CONTROL THEORY FOR DISTRIBUTED PARAMETER SYSTEMS

JENNIFER C. POLLOCK

Submitted for the degree of Doctor of Philosophy at the University of Warwick.

This research was conducted initially in the Control Theory Centre, University of Warwick, and then in the Faculty of Commerce, University College, Dublin.

June 1980.

CONTENTS

Introduction	1
1. Standard Concepts	3
2. System Formulation	10
3. Perturbation Results	17
4. The Linear Quadratic Cost Regulator Problem	25
5. The Tracking Problem	44
6. The Infinite Time Quadratic Cost Control Problem	59
7. A Stabilizability Result	73
8. Alternative Approaches	74
9. Applications	98
9.1 Optimum Inventory Control	98
9.2 Optimum Advertising Policy	115
9.3 Pollution Control	125
9.4 Traffic Flow Control	127
10. Comments on Hyperbolic Systems	133
11. Observer Theory	137
11.1 Observers for Systems Described by Semigroups	139
12. The Effect of Observers in the Linear Quadratic Cost Control Problem	160
13. Observers for Systems with Unbounded Control Action	166
Conclusions	175
References	180

ACKNOWLEDGEMENTS

I wish to express my grateful thanks to the partners of Coopers and Lybrand for the provision of financing that enabled this research to be completed, and also to the members of the Faculty of Commerce, University College, Dublin, namely Professor M. MacCormac, Professor D. Hally and Professor H. Harrison, for arranging this financial support.

I am also indebted to Professor M. Hayes and Professor D. Judge, Department of Mathematical Physics, University College, Dublin, for the provision of office and secretarial facilities, and to Miss M. Mitchell for the assistance she rendered.

Finally I wish to express my sincere thanks to my supervisor, Dr. A.J. Pritchard, and to acknowledge my indebtedness to him, for his patient help, encouragement and direction.

The material contained in sections 2 - 7 of this thesis will shortly appear, in abridged form, in two papers {30} and {31} accepted for publication in the Journal of the Institute of Mathematics and its Applications. These papers were prepared for publication with the assistance of my supervisor, Dr. A.J. Pritchard.

SUMMARY.

In this thesis problems in control theory for distributed parameter systems are studied, using a semigroup approach.

Firstly the control problem is formulated for systems $\dot{z} = Az + Bu$ on a Banach space Z, when the control operator B is unbounded on Z. The semigroup T_t generated by the system operator A is required to be smoothing so that the resultant operator T_tB is bounded by an L^P function.

The finite and infinite time regulator, and the tracking problems for such systems are then solved. By constructing an iterative sequence of sub-optimal controls it is shown that for the regulator problems the optimal control is feedback, and for the tracking problem is feedback plus open-loop. It is further shown that the feedback operator, time independent in the infinite time case, is the unique solution to an integral Riccati equation which is differentiable. The differential equation has unique solution also, where we make additional assumptions on the system operators in the infinite time case. The open-loop control of the tracking problem is also shown to be associated with the unique solutions to integral and differential equations.

Arising out of the solution to the infinite time regulator problem, the stabilizability result, exact null controllability implies stabilizability, is also proved for these systems.

The results obtained are then compared with those of other authors and applications given.

Observer theory for distributed parameter systems described by semigroups is then considered. Conditions are found, in terms of the system operators, for an observer to be an asymptotic state estimator under feedback and general control action. The increase in cost due to using an observer as feedback in the regulator is studied and found to be dependent on the initial state of the system, in general unknown.

INTRODUCTION

In this thesis we examine problems in control theory for distributed parameter systems.

After detailing the standard concepts used in the thesis we will formulate the control problem for systems with unbounded control action. The systems considered are those given by the mild solutions to abstract equations of the form $z_t = Az + Bu$ on the Banach space Z,, where U the control space is also a Banach space (though in consideration of the linear quadratic cost control problem we require further that Z and U be Hilbert spaces). An assumption is made, essentially that the system semigroup be smoothing, so that the resultant operator T_tB is bounded.

We will then consider the linear quadratic cost control problem and, using an iterative sequence of sub-optimal controls, seek the optimal control which minimizes the quadratic performance index. Three aspects of the problem will be considered, the finite time regulator, the tracking problem and the infinite time regulator.

We will then examine how the formulation and results obtained compare with those of others considering control problems with unbounded control action. The application of the results to specific examples will also be considered.

Finally we consider observer theory for systems described by semigroups. For systems with bounded control action we will examine conditions under which an observer is also an asymptotic estimator and under which an observer will stabilize the original system. The effect of using the state of the observer in the feedback control of the regulator problem will also be considered. Lastly we extend the results on observer theory to systems with unbounded control action.

1. STANDARD CONCEPTS - CONTROLLABILITY, OBSERVABILITY, STABILIZABILITY AND THE LINEAR QUADRATIC COST PROBLEM.

In this thesis we consider control systems of the form

$$\begin{aligned} \dot{z} &= Az + Bu \\ z(o) &= z_o \end{aligned} \tag{1.1}$$

where A generates a strongly continuous semigroup T_t on a Banach space Z and for the purposes of this section, B is a bounded operator from a control space U to Z. U is also a Banach space.

If z is a strict solution of (1.1) then $z(t) \in D(A)$ for all $t \in [o,t_1]$ and so, in the general case where A is unbounded and hence $D(A) \neq Z$, the system cannot be steered to all of Z.

For $z_o \in Z$ and $u \in L^q[o,t_1;U]$ consider the mild solution of (1.1) defined by

$$z(t) = T_t z_o + \int_o^t T_{t-s} Bu(s)ds \tag{1.2}$$

then $z(t)$ is well defined by (1.2) though we are not able, in general, to differentiate it to obtain (1.1) unless $z_o \in D(A)$ and u is suitably smooth, e.g. $u \in C^1[o,t_1;U]$. We thus chose to take as our system description the more general mild solution (1.2) rather than the abstract equation (1.1) and we now introduce some of the standard concepts and definitions refered to in this thesis.

Remark

For a general exposition on semigroup theory in relation to control problems see Curtain and Pritchard {11}. Also in this book may

1. STANDARD CONCEPTS - CONTROLLABILITY, OBSERVABILITY, STABILIZABILITY AND THE LINEAR QUADRATIC COST PROBLEM.

In this thesis we consider control systems of the form

$$(1.1)\qquad \dot{z} = Az + Bu$$
$$z(o) = z_o$$

where A generates a strongly continuous semigroup T_t on a Banach space Z and for the purposes of this section, B is a bounded operator from a control space U to Z . U is also a Banach space.

If z is a strict solution of (1.1) then $z(t) \in D(A)$ for all $t \in [o,t_1]$ and so, in the general case where A is unbounded and hence $D(A) \neq Z$, the system cannot be steered to all of Z .

For $z_o \in Z$ and $u \in L^q[o,t_1;U]$ consider the mild solution of (1.1) defined by

$$(1.2)\qquad z(t) = T_t z_o + \int_o^t T_{t-s} Bu(s)ds$$

then $z(t)$ is well defined by (1.2) though we are not able, in general, to differentiate it to obtain (1.1) unless $z_o \in D(A)$ and u is suitably smooth, e.g. $u \in C^1[o,t_1;U]$. We thus chose to take as our system description the more general mild solution (1.2) rather than the abstract equation (1.1) and we now introduce some of the standard concepts and definitions refered to in this thesis.

<u>Remark</u>

For a general exposition on semigroup theory in relation to control problems see Curtain and Pritchard {11} . Also in this book may be found details and examples relating to the definitions introduced below.

Firstly we consider the concept of controllability, i.e. whether given two points $z_o, z_1 \in Z$ it is possible to find a control which steers the system from z_o to z_1, and we make the following definition,

<u>Definition 1.1</u> - <u>Exact Controllability on $[o,t_1]$</u>.

Given any two points $z_o, z_1 \in Z$ we say (1.2) is exactly controllable on $[o,t_1]$ if there exists a control $u \in L^q[o,t_1;U]$ such that $z(t_1) = z_1$.

In those cases where it is not possible to exactly control the system to all points in Z (Curtain and Pritchard {11} give an example of such a system) it may be possible to control the system to points which form a dense set in Z. Thus

<u>Definition 1.2</u> - <u>Approximate Controllability on $[o,t_1]$</u>.

We say (1.2) is approximately controllable on $[o,t_1]$ if

$$\text{Range } \overline{\{\int_o^{t_1} T_{t_1-s} Bu(s)ds\}} = Z$$

i.e. (1.2) is approximately controllable if for any $z_1 \in Z$ and any $\varepsilon > 0$ there exists a control $u \in L^q[o,t_1;U]$ such that $||z(t_1) - z_1||_Z \leq \varepsilon$.

Since in many problems of practical interest the null state plays an important role special definitions are introduced for systems which can achieve or approximately achieve the null state, viz:-

<u>Definition 1.3</u> - <u>Exact Null Controllability on $[o,t_1]$</u>.

We say that (1.2) is exactly null controllable on $[o,t_1]$ if

$$\text{Range }\{\int_0^{t_1} T_{t_1-s}Bu(s)ds\} \supset \text{Range }\{T_{t_1}\}$$

i.e. for any $z_o \in Z$ there exists a control $u \in L^q[0,t_1;U]$ such that $\int_0^{t_1} T_{t_1-s}Bu(s)ds = -T_{t_1}z_o$ as then

$$z(t_1) = T_{t_1}z_o + \int_0^{t_1} T_{t_1-s}Bu(s)ds = 0$$

<u>Definition 1.4</u> - <u>Approximate Null Controllability on $[0,t_1]$</u>.

We say that (1.2) is approximately null controllable on $[0,t_1]$ if

$$\overline{\text{Range }\{\int_0^{t_1} T_{t_1-s}Bu(s)ds\}} \supset \text{Range }\{T_{t_1}\}$$

i.e. for any $z_o \in Z$, $\varepsilon > 0$, there exists a control $u \in L^q[0,t_1;U]$ such that $||z(t_1) - 0||_Z \leq \varepsilon$.

Suppose now that in addition to the system equation

(1.3) $\dot{x} = \mathcal{A}x$, $x(o) = x_o$,

where $\mathcal{A}$ is the infinitesimal generator of a strongly continuous semigroup S_t on a reflexive Banach space X, we have the observation equation

(1.4) $y = Cx$.

Here $C \in \mathcal{L}(X,Y)$ where Y, the observation space, is also a reflexive Banach space. Then the mild solution of (1.3) - (1.4) is

(1.5) $y = CS_tx_o$

and we may define a map $\mathcal{C}$ by

(1.6) $\mathcal{C} : X \to L^p[0,t_1;Y] : \mathcal{C}x_o = CS_tx_o$.

The observation problem is concerned with whether or not it is possible to reconstruct the initial state at time t_1 of a system in a unique fashion from the observations. As with controllability we make four definitions.

Knowledge of the initial state enables us to obtain the whole state at all times $t \geq 0$ and hence initial observation is important.

<u>Definition 1.5</u> - <u>Initially Observable on $[0,t_1]$</u> .

We say (1.5) is initially observable on $[0,t_1]$ if $\ker\{\mathcal{C}\} = \{0\}$ so then the initial state is distinguishable, i.e. $y(t) = 0$ on $[0,t_1]$ implies $x_o = 0$.

<u>Definition 1.6</u> - <u>Continuously Initially Observable on $[0,t_1]$</u> .

We say (1.6) is continuously initially observable on $[0,t_1]$ if there exists a $\gamma > 0$ such that

$$\gamma \|\mathcal{C} x\|_{L^p[0,t_1;Y]} \geq \|x\|_X \qquad \forall\, x \in X ,$$

which implies the existence of a continuous reconstruction operator R_o , R_o : Range $\{\mathcal{C}\} \to X$, such that $R_o \mathcal{C} = I$.

Since in practice it is not generally possible to construct the initial state until a finite time t_1 has elapsed, and, for times greater than t_1 the whole state can be determined by the state at time t_1 , being able to distinguish the state at time t_1 is a useful concept.

<u>Definition 1.7</u> - <u>Finally Observable on $[0,t_1]$</u> .

We say that (1.6) is finally observable on $[0,t_1]$ if $\ker\{\mathcal{C}\} \subset \ker\{S_{t_1}\}$ so that then the state at time t_1 can be distinguished, i.e. $y(t) = 0$ on $[0,t_1]$ implies $x(t_1) = 0$.

<u>Definition 1.8</u> - <u>Continuously Finally Observable on $[0,t_1]$</u> .

We say that (1.6) is continuously finally observable on $[0,t_1]$ if there exists a $\gamma > 0$ such that

$$\gamma \|\mathcal{C}\, x\|_{L^p[0,t_1;Y]} \geq \|S_{t_1}\|_X \qquad \forall\, x \in X$$

which implies the existence of a continuous reconstruction operator R_{t_1} , R_{t_1} : Range $\{\mathcal{C}\} \to X$, such that $R_{t_1}\mathcal{C} = S_{t_1}$

By making dual identifications between the spaces and operators of the systems (1.2) and (1.6) we can show that controllability and observability are dual concepts. The following theorem is proved in {11}.

<u>Theorem 1.9</u>

If we make the identifications

$U = Y^*$, $B = C^*$, $T_t = S_t^*$, $Z = X^*$, $A = \mathcal{A}^*$, $\frac{1}{p} + \frac{1}{q} = 1$

then

(a) (1.6) is initially observable on $[0,t_1]$ iff (1.2) is approximately controllable on $[0,t_1]$.

(b) (1.6) is continuously initially observable on $[0,t_1]$ iff (1.2) is exactly controllable on $[0,t_1]$.

(c) (1.6) is finally observable on $[0,t_1]$ iff (1.2) is approximately null controllable on $[0,t_1]$.

(d) (1.6) is continuously finally observable on $[0,t_1]$ iff (1.2) is exactly null controllable on $[0,t_1]$.

Another important question is whether it is possible to design a feedback controller so that the controlled system is asymptotically stable, and thus

Definition 1.10 - **Exponential Stabilizability**.

Let A be the infinitesimal generator of a strongly continuous semigroup T_t on a Banach space Z and $B \varepsilon \mathcal{L}(U,Z)$. If there exists a $D \varepsilon \mathcal{L}(Z,U)$ such that $A + BD$ generates a strongly continuous semigroup T_t^D with

$$||T_t^D|| \leq Ke^{-\omega t} \quad , \quad \omega > 0,$$

then the pair $\{A,B\}$ is said to be exponentially stabilizable.

For finite dimensional spaces the nice result that controllability implies stabilizability holds . The situation however is much more complicated for infinite dimensional Banach spaces, and in {11} Curtain and Pritchard give an example of a system which is approximately controllable but not stabilizable. In {11} they also prove (by exploiting the linear quadratic cost problem) that exact null controllability implies stabilizability. In section 7 we show that this result holds for the class of systems with unbounded operators B considered there.

Finally we outline the nature of the linear quadratic cost problem for control systems of the form (1.2) but now defined on a Hilbert space H .

Definition 1.11 - **The Linear Quadratic Cost Control Problem** .

The linear quadratic cost problem is to find the control u^* which minimizes a performance index of the form

$$(1.7) \qquad J(u;t_o,z_o) = \langle z(t_1) - r(t_1) \ , G[z(t_1) - r(t_1)]\rangle_H$$

$$+ \int_{t_o}^{t_1} \{\langle z(t) - r(t) \ , M[z(t) - r(t)]\rangle_H + \langle u(t) \ , Ru(t)\rangle_U \}dt$$

on a Hilbert space H , with controls u in the Hilbert space U . Here $z \in H$ is the response through (1.2) to $u \in L^2[0,t_1;U]$ and $r \in C[0,t_1;H]$ is a given H-valued function. We take $G, M \in \mathcal{L}(H)$ and $R, R^{-1} \in \mathcal{L}(U)$ such that $G, M \geq 0$ and $R > 0$. Also G, M and R are self-adjoint.

The minimization problem with cost functional given by (1.7) is known as the tracking problem. For the finite time regulator problem we set $r(t) \equiv 0$, i.e. we track the zero function. In the infinite time regulator problem the cost functional is given by (1.7) with $r(t) \equiv 0$ and $t_1 = \infty$.

In this thesis these three aspects of the linear quadratic cost control problem are considered for a class of control problems with unbounded control action.

Before we are able to examine the quadratic cost control problem we must detail and formulate the control problem for the class of systems under consideration and prove perturbation results. This is done in the following two sections

2. SYSTEM FORMULATION.

In many examples of practical importance, as in the case of most systems described by partial differential equations, the control is not distributed over the whole space but is restricted to subsets or to the boundary. It is then not always possible to formulate a control problem of the form (1.1) with B a bounded operator, $B \in \mathcal{L}(U,H)$. An example of such a process is

Example 2.1

The controlled diffusion equation

(2.1) $$z_t = Az = z_{xx}$$

with

(2.2) $$z(o,t) = u(t) , \quad z(1,t) = 0 , \quad z(x,o) = z_o(x) .$$

Alternatively we might consider (2.1) with the boundary conditions

(2.3) $$[z_x]_o = u , \quad z_x(1,t) = 0 , \quad z(x,o) = z_o(x) .$$

where $[\ \ ]_o$ denotes the change from 0^- to 0^+ .

Equations (2.1) - (2.2) are not in the normal form (1.1) for which we can write down the mild solution (1.2) and it is not immediately obvious how the operator B should be chosen.. We will see that the course to follow is to define a weak solution to (2.1) - (2.2) , establish the operator B via a Green's formula, and then show that the mild solution (1.2) is a weak solution of (2.1) - (2.2).

In order to establish the formal framework in which to consider the problem we assume the existence of a Banach space $\tilde{W}$, with H

dense in $\tilde{W}$, such that

(2.4)
- (a) $\tilde{W} \supset R(B)$
- (b) $B \in \mathcal{L}(U,\tilde{W})$
- (c) $T_t \in \mathcal{L}(\tilde{W},H)$, $t > 0$
- (d) $||T_t w||_H \leq g(t)||w||_{\tilde{W}}$ for all $w \in \tilde{W}$ with $g \in L^p[0,t_1]$, $p \geq 1$.

Consider the control system

$$z(t) = T_t z_o + \int_o^t T_{t-s}Bu(s)ds \tag{2.5}$$

for $z_o \in H$ and $u \in L^q[0,t_1;U]$ where $\frac{1}{p} + \frac{1}{q} = 1$ then we have the following propositions, proofs of which may be found in {11}.

Proposition 2.2

$z(t)$ is well defined by (2.5) and furthermore $z \in C[0,t_1;H]$.

Proposition 2.3

If $f \in C[0,t_1;H]$ and $x(t) = -\int_t^{t_1} T^*_{s-t}f(s)ds$ then $z(t)$ satisfies (2.5) iff it satisfies

$$\int_o^{t_1} \langle f(t), z(t)\rangle_H dt + \int_o^{t_1} \langle Cx(t), u(t)\rangle_U dt + \langle x(o), z_o\rangle_H = 0 \tag{2.6}$$

where $B = C^*$.

Remark

If f is smooth, namely $f \in C^1[0,t_1;H]$, then

$$\dot{x} + A^*x = f$$
$$x(t_1) = 0 .$$

We now show that (2.5) is a weak solution to a controlled abstract differential equation. To do this let A the infinitesimal generator of

T_t and A^* the generator of T_t^* be defined on appropriate function spaces on an open bounded set Ω with boundary conditions on Γ, the boundary of Ω. If U is a Hilbert space of functions on a subset Ω_1 of Ω we denote by $\tilde{A}$ the same formal operator as A, but now defined on the restriction of the function space to $\Omega\backslash\Omega_1$, with the same boundary conditions on $\Gamma\backslash\Omega_1$.

If we also assume the existence of a Green's formula

$$(2.7) \qquad \langle\phi, \tilde{A}\psi\rangle_H = \langle A^*\phi, \psi\rangle_H + \langle C\phi, D\psi\rangle_U + \langle G\phi, E\psi\rangle_U$$

for $\phi \in D(A^*)$, $\psi \in D(\tilde{A})$, then by means of proposition 2.3 the following result can be shown to hold.

<u>Theorem 2.4</u>

Under the assumptions (2.4) and (2.7) with $C = B^*$, (2.5) i.e.

$$z(t) = T_t z_o + \int_o^t T_{t-s}Bu(s)ds$$

is a weak solution of

$$(2.8) \qquad \dot{z} = \tilde{A}z$$

$$Dz = u\,, \quad Ez = 0\,, \quad z(o) = z_o\,,$$

where we make the following definition.

<u>Definition 2.5</u>

A <u>weak solution</u> of (2.8) is a function $z \in C[0,t_1;H]$ such that

$$(2.9) \qquad \int_o^{t_1}\langle f(t), z(t)\rangle_H dt + \int_o^{t_1}\langle Cx(t), u(t)\rangle_U dt + \langle x(o), z_o\rangle_H = 0$$

where

$$\dot{x} + A^* x = f$$

$$x(t_1) = 0 \;, \qquad f \in C^1[0,t_1;H] \;.$$

Then, since from (2.8)

$$0 = \int_o^{t_1} \langle x(t),\, \dot{z}(t) - \tilde{A}z(t)\rangle_H \, dt$$

applying the Green's formula (2.7) with $\phi = x$, $\psi = z$

$$0 = \langle x(t_1),\, z(t_1)\rangle - \langle x(o),\, z_o\rangle - \int_o^{t_1} \langle \dot{x}(t) + A^* x(t),\, z(t)\rangle_H \, dt$$

$$- \int_o^{t_1} \langle Cx(t),\, Dz(t)\rangle_U \, dt - \int_o^{t_1} \langle Gx(t),\, Ez(t)\rangle_U \, dt \;.$$

As $Dz = u$, $Ez = 0$ taking $f = \dot{x} + A^* x$ with $x(t_1) = 0$ yields (2.9). Thus, from proposition 2.3, (2.5) is a weak solution of (2.8).

<u>Remark</u>

Assumption (2.7) on the existence of a Green's formula is realistic since, from Aubin { 1 }, at least in the following case, an operator can always be found such that a Green's formula holds.

<u>Proposition 2.6</u> - (Aubin { 1 })

If K is a Hilbert space such that

$$D(\tilde{A}) \subset K \subset H$$

and there exists an operator α which maps K onto the Hilbert space U such that $K_o = \ker \alpha$ is dense in H, then there exist operators α_o, α_o^*, α_1, α_1^*, associated with α, such that the following Green's formula holds

$$\langle A^* u,\, v\rangle - \langle u,\, \tilde{A}v\rangle = \langle \alpha_o^* u,\, \alpha_o v\rangle - \langle \alpha_1^* u,\, \alpha_1 v\rangle \quad \text{for } u \in D(A^*) \;,\; v \in D(\tilde{A})$$

where A^* is the adjoint to $\tilde{A}$.

Thus if we take $K = \tilde{W}^*$ and $C = B^* \varepsilon \mathcal{L}(\tilde{W}^*, U^*)$, where $\tilde{W}^*$ denotes the dual of $\tilde{W}$, and we have proposition 2.6 holding, we know that there exists a Green's formula for our system of the form (2.7).

The following examples illustrate how the abstract theory works in practice.

<u>Example 2.7</u>

Consider again the control system (2.1) - (2.2) i.e.

$$(2.10) \qquad \begin{aligned} &z_t = z_{xx} \\ &z(o,t) = u \ , \quad z(1,t) = 0 \ , \quad z(x,o) = z_o(x) \ . \end{aligned}$$

We wish to determine the operator B such that the mild solution of

$$(2.11) \qquad \begin{aligned} &z_t = z_{xx} + Bu \\ &z(o,t) = 0 = z(1,t) \ , \quad z(x,o) = z_o(x) \ , \end{aligned}$$

is a weak solution of (2.10). To employ the abstract theory we set $\Omega = (0,1)$, $H = L^2(\Omega)$, $U = R$, $A\phi = A^*\phi = \phi_{xx}$ on $D(A) = H^2(0,1) \cap H_o^1(0,1)$ and $\tilde{A}\psi = \psi_{xx}$ on $D(\tilde{A}) = \{\psi \ \varepsilon \ L^2(\Omega);\ \psi_x, \psi_{xx} \ \varepsilon \ L^2(\Omega),\ \psi(1) = 0\}$.

The Green's formula for the system is

$$\int_o^1 (\phi\psi_{xx} - \phi_{xx}\psi)dx = \phi_x(o)\psi(o)$$

which comparing with the general Green's formula (2.7) gives $C\phi = \phi_x(o)$, and thus letting $B^* = C$ we have, formally,

$$\begin{aligned} \langle\phi, Bu\rangle_H &= \langle C\phi, u\rangle_U \\ &= u\phi_x(o) \\ &= u\int_o^1 \delta(x)\phi_x(x)dx \end{aligned}$$

$$= -u\int_0^1 \delta'(x)\phi(x)dx$$

$$= \langle\phi, -\delta'(x)u\rangle_H$$

Hence, $(Bu)(x) = -u\delta'(x)$. i.e. $B = -\delta'$, the negative derivative of the Dirac delta function.

It is easy to show that the semigroup T_t is given by

(2.12) $$(T_t z_o)(x) = \sum_{n=1}^{\infty} 2e^{-n^2\pi^2 t} \sin n\pi x \langle z_o(\cdot), \sin n\pi\cdot\rangle$$

and then

$$||T_t Bu||^2 = \langle T_t Bu, T_t Bu\rangle$$

$$\leq \sum_{n=1}^{\infty} 2n^2\pi^2 e^{-2n^2\pi^2 t} \left[\int_0^1 \cos n\pi y\ u(y,t)dy\right]^2$$

$$\leq \max_n\ n\pi^2 e^{-2n^2\pi^2 t}\ ||u||^2_{L^2}$$

so

$$||T_t Bu|| \leq \frac{M}{t^{1/2}} ||u||_{L^2} \quad .$$

From the work of Lions and Magenes {44} we know that $B \in \mathcal{L}(U,\tilde{W})$ where $\tilde{W} = (H^{3/2+\epsilon}(0,1))^*$ and so

$$||T_t w||_H \leq \frac{M}{t^{3/4+\epsilon/2}} ||w||_{\tilde{W}} \qquad \text{for } w \in W \quad .$$

i.e. condition (2.4)(d) is satisfied with $g(t) = \dfrac{M}{t^{3/4+\epsilon/2}}$ and by a p such that $4/3 > p \geq 1$, for this $\tilde{W}$.

Example 2.8

Consider the system (2.1) now with boundary conditions (2.3) i.e.

$$z_t = z_{xx}$$

$$\left[z_x\right]_o = u \ , \quad z_x(1,t) = 0 \ , \quad z(x,o) = z_o(x) \ .$$

If we set $H = L^2(0,1)$ and $U = \mathbb{R}$ we can this time show that $(Bu)(x) = -\delta(x)u$ and T_t is given by

$$(T_t z_o)(x) = \langle z_o, 1\rangle + \sum_{n=1}^{\infty} 2e^{-n^2\pi^2 t} \cos n\pi x \langle z_o(\cdot), \cos n\pi\cdot\rangle \ ,$$

so condition (2.4) is satisfied this time with $4 > p \geq 1$ if we take $\tilde{W} = (H^{1/2+\epsilon}(0,1))^*$ and $g(t) = \dfrac{M}{t^{1/4+\epsilon/2}}$.

Remarks

1. This formulation remains valid for the more general case of control systems defined on a Banach space Z with controls in a Banach space U if, for example, in the Green's formula (2.7) instead of the inner product on H we consider the duality pairing between Z and Z^* .

2. The standard definitions in section 1 can be carried over and their concepts considered for these systems with unbounded control operators B .

3. Curtain and Pritchard {11} formulate the problem only for systems satisfying (2.4) with $p \geq 2$ and are thus unable to consider systems such as that considered here in examples 2.1 and 2.7.

3. PERTURBATION RESULTS.

Consider the control system, as formulated in the previous section, over the interval $[0,t_1]$,

$$(3.1)\qquad z(t) = T_t z_o + \int_o^t T_{t-s} Bu(s)ds$$

on a Hilbert space H, with controls taking values in a Hilbert space U, and where condition (2.4) is assumed to hold.

Since for the linear quadratic cost problem we wish to consider controls of the form $u(t) = F(t)z(t)$ we are led to considering perturbations of T_{t-s} defined by

$$(3.2)\qquad V(t,s)h = T_{t-s}h + \int_s^t T_{t-r} BF(r)V(r,s)hdr\ , \quad h \in H$$

where $F(t)$ satisfies the following condition

$$(3.3)\qquad \begin{array}{ll} \text{(a)} & F(t) \in \mathcal{L}(H,U) \quad \text{for all} \quad t \in [0,t_1] \\ \text{(b)} & ||F(t)||_{\mathcal{L}(H,U)} \leq g(t) \in L^q[0,t_1] \end{array}$$

Notation

For the remainder of this thesis, where no confusion can arise, we will drop the subscripts from $||\cdot||$ if it is the norm in H, U, $\mathcal{L}(H)$ or $\mathcal{L}(U)$ and also from $\langle\cdot,\cdot\rangle_H$ and $\langle\cdot,\cdot\rangle_U$.

Curtain and Pritchard {11} establish the following perturbation result for $F(t) = F$ independent of time which we now extend to the time dependent case.

<u>Theorem 3.1</u>

If $F(t)$ satisfies (3.3) and T_t , B satisfy (2.4) with $\frac{1}{p} + \frac{1}{q} = 1$, then the controlled system

$$z(t) = T_t z_o + \int_o^t T_{t-s} F(s) z(s) ds$$

has a unique solution $z(t) = V(t,o)z_o$ where $V(t,s)$ is a mild evolution operator (i.e. a two parameter semigroup) which is the unique solution of

$$V(t,s)z(s) = T_{t-s}z(s) + \int_s^t T_{t-r} BF(r)V(r,s)z(s)dr \quad . \tag{3.4}$$

<u>Proof</u>

In order to prove theorem 3.1 we require the following lemma and corollary, proofs of which may be found in {11}.

<u>Lemma 3.2</u>

Consider the integral equation

$$f(t) = h(t) + \int_o^t g(t-s)f(s)ds \tag{3.5}$$

where $h \in L^p[0,t_1]$, $g \in L^1[0,t_1]$ and are positive, then (3.5) has unique solution which is given by

$$f(t) = h(t) + \sum_{n=1}^{\infty} (G^n h)(t) \tag{3.6}$$

with

$$(3.7) \qquad (G^n h)(t) = \int_o^t g_n(t-s)h(s)ds$$

where

$$(3.8) \qquad g_n(t) = \int_o^t g(t-s)g_{n-1}(s)ds \quad , \quad n > 1$$

$$g_1(t) = g(t) \quad .$$

<u>Corollary 3.3</u> - <u>Generalised Gronwall's Inequality</u>.

Suppose

$$(3.9) \qquad f(t) \leq h(t) + \int_o^t g(t-s)f(s)ds$$

with $h \in L^p[0,t_1]$, $g \in L^1[0,t_1]$ both positive, then

$$(3.10) \qquad f(t) \leq h(t) + \sum_{n=1}^{\infty}(G^n h)(t) \ ,$$

and in particular,

if $h = 0$, then $f = 0$.

Returning to the proof of theorem 3.1 we construct the evolution operator by means of the iterative scheme

$$(3.11) \qquad \begin{aligned} V_o(t,s) &= T_{t-s} \\ V_n(t,s)z(s) &= \int_s^t T_{t-\rho}BF(\rho)V_{n-1}(\rho,s)z(s)d\rho \ . \end{aligned}$$

Firstly note that

$$(3.12) \qquad \begin{aligned} ||T_{t-s}BF(s)h|| &\leq g(t-s)||B||_{\mathcal{L}(U,W)} {}^{\sim}\underline{g}(s)||h|| \\ &= G(t-s)||h|| \end{aligned}$$

and then $G \in L^1$ since $g \in L^p$, $\underline{g} \in L^q$ with $\frac{1}{p} + \frac{1}{q} = 1$.

By induction we prove that

(3.13) $$||V_n(t,s)z|| \leq \bar{M}\int_s^t G_n(t-\alpha)d\alpha||z|| \quad \text{for all} \quad n \geq 1$$

where $G_1(t) = G(t)$

$$G_n(t) = \int_s^t G(t-r)G_{n-1}(r)dr$$

and $\bar{M}$ is a constant such that $||T_{t-s}|| \leq \bar{M}$.

For $n = 1$, we have

$$||V_1(t,s)z|| \leq \int_s^t ||T_{t-r}BF(r)V_o(r,s)z||dr$$

$$\leq \int_s^t G(t-r)||V_o(r,s)z||dr$$

$$\leq \bar{M}\int_s^t G(t-r)dr||z|| \quad \text{since} \quad V_o(t,s) = T_{t-s}$$

Now assuming (3.13) holds for $n = k-1$, we have

$$||V_k(t,s)z|| \leq \int_s^t G(t-r)||V_{k-1}(r,s)z||dr$$

$$\leq \bar{M}\int_s^t G(t-r)\int_s^r G_{k-1}(r-\alpha)d\alpha dr||z||$$

$$\leq \bar{M}\int_s^t\int_\alpha^t G(t-r)G_{k-1}(r-\alpha)drd\alpha||z||$$

$$\leq \bar{M}\int_s^t G_k(t-\alpha)d\alpha||z|| \ .$$

Thus (3.13) is established, and so using lemma 3.2 and its corollary 3.3 we have

(3.14) $$\sum_{n=0}^{\infty} ||V_n(t,s)|| \leq \bar{M} + \bar{M}\sum_{n=1}^{\infty} \int_s^t G_n(t-\alpha)d\alpha < \infty \ .$$

By induction we prove that

$$(3.13) \qquad ||V_n(t,s)z|| \leq \bar{M}\int_s^t G_n(t-\alpha)d\alpha||z|| \qquad \text{for all} \quad n \geq 1$$

where $\quad G_1(t) = G(t)$

$$G_n(t) = \int_s^t G(t-r)G_{n-1}(r)dr$$

and $\bar{M}$ is a constant such that $||T_{t-s}|| \leq \bar{M}$.

For $n = 1$, we have

$$||V_1(t,s)z|| \leq \int_s^t ||T_{t-r}BF(r)V_o(r,s)z||dr$$

$$\leq \int_s^t G(t-r)||V_o(r,s)z||dr$$

$$\leq \bar{M}\int_s^t G(t-r)dr||z|| \qquad \text{since} \quad V_o(t,s) = T_{t-s}$$

Now assuming (3.13) holds for $n = k-1$, we have

$$||V_k(t,s)z|| \leq \int_s^t G(t-r)||V_{k-1}(r,s)z||dr$$

$$\leq \bar{M}\int_s^t G(t-r)\int_s^r G_{k-1}(r-\alpha)d\alpha dr||z||$$

$$\leq \bar{M}\int_s^t\int_\alpha^t G(t-r)G_{k-1}(r-\alpha)drd\alpha||z||$$

$$\leq \bar{M}\int_s^t G_k(t-\alpha)d\alpha||z|| \; .$$

Thus (3.13) is established, and so using lemma 3.2 and its corollary 3.3 we have

$$(3.14) \qquad \sum_{n=0}^{\infty} ||V_n(t,s)|| \leq \bar{M} + \bar{M}\sum_{n=1}^{\infty} \int_s^t G_n(t-\alpha)d\alpha < \infty \; .$$

Hence $V(t,s) = \sum_{n=0}^{\infty} V_n(t,s)$ converges absolutely in the uniform topology and the convergence is uniform in s and t. $V(t,s)$ clearly satisfies (3.11).

For uniqueness:- Suppose $V^2(t,s)$ is another solution and let $\overline{V}(t,s) = V(t,s) - V^2(t,s)$ then

$$\overline{V}(t,s)z(s) = \int_s^t T_{t-r} BF(r)\overline{V}(r,s)z(s)dr$$

and

$$||\overline{V}(t,s)z(s)|| \le \int_s^t G(t-r)||\overline{V}(r,s)z(s)||dr \ .$$

So $\overline{V}(t,s) = 0$ by corollary 3.3.

The strong continuity of $V(t,s)$ in t and s may be proved either via the construction or directly as follows:- Let $h > 0$, then

$$V(t+h,s)z - V(t,s)z$$
$$= (T_{t+h-s} - T_{t-s})z + (T_h - I)\int_s^t T_{t-r}BF(r)V(r,s)zdr + \int_t^{t+h} T_{t+h-s}BF(r)V(r,s)zdr$$

from which the continuity on the right follows using the strong continuity of T_t and the fact that $T_{t-r}BF(r)V(r,s)z \in L^1[0,t_1;H]$ since $||T_{t-r}BF(r)V(r,s)z|| \le g(t-r)||B||_{\mathcal{L}(U,W)} \underline{g}(r)||V(r,s)|| \, ||z||$, with $g\underline{g} \in L^1$ and from (3.14) there exists an $M^F < \infty$ such that $\operatorname{ess\,sup}_{0\le s\le t\le T} ||V(t,s)|| \le M^F$.
Similar arguments hold for continuity on the left in t and continuity in s.

For the semigroup property we have that

$$V(t,r)V(r,s)z$$
$$= T_{t-r}V(r,s)z + \int_s^t T_{t-\rho}BF(\rho)V(\rho,r)V(r,s)zd\rho$$

$$= T_{t-r}T_{r-s}z + T_{t-r}\int_s^r T_{r-\rho}BF(\rho)V(\rho,s)zd\rho + \int_r^t T_{t-\rho}BF(\rho)V(\rho,r)V(r,s)zd\rho$$

Hence

$V(t,s)z - V(t,r)V(r,s)z$

$$= (T_{t-s}z - T_{t-r}T_{r-s})z + \int_s^t T_{t-r}BF(r)V(r,s)zdr - T_{t-r}\int_s^r T_{r-\rho}BF(\rho)V(\rho,s)zd\rho$$
$$- \int_r^t T_{t-\rho}BF(\rho)V(\rho,r)V(r,s)zd\rho$$

$$= \int_s^t T_{t-r}BF(r)V(r,s)zdr - \int_s^r T_{r-\rho}BF(\rho)V(\rho,s)zd\rho - \int_r^t T_{r-\rho}BF(\rho)V(\rho,r)V(r,s)zd\rho$$

using the semigroup property of T_t

$$= \int_r^t T_{t-\rho}BF(\rho)V(\rho,s)zd\rho - \int_r^t T_{t-\rho}BF(\rho)V(\rho,r)V(r,s)zd\rho$$

$$= \int_r^t T_{t-\rho}BF(\rho)[V(\rho,s) - V(\rho,r)V(r,s)]zd\rho$$

and thus by corollary 3.3

$$V(t,s) - V(t,r)V(r,s) = 0$$

i.e. $\quad V(t,s) = V(t,r)V(r,s)$.

Theorem 3.1 shows that perturbing the original semigroup by an operator of the form $BF(t)$ where $F(t)$ satisfies (3.3) leads to a mild evolution operator $V(t,s)$. We have further the following proposition which is again an extension of the time dependent result to be found in {11}.

<u>Proposition 3.4</u>

$V(t,s)$ is a quasi-evolution operator in the sense that

(3.15) $\quad \int_o^t V(t,\rho)[A + BF(\rho)]z_o d\rho = V(t,o)z_o - z_o \quad$ for $z_o \in D(A)$

which implies

(3.16) $\quad \frac{\partial}{\partial s} V(t,s)z_o = - V(t,s)[A + BF(s)]z_o \quad$ a.e. for $z_o \in D(A)$, $t > s$,

and if we assume the existence of a Green's formula (2.7) with $B^* = C$ then $V(t,s)$ satisfies

(3.17) $$\int_o^t V(t,\rho)\tilde{A}z_o d\rho = V(t,o)z_o - z_o$$

for $z_o \in \mathcal{D} = D(\tilde{A}) \cap \ker\{D - F(\rho)\} \cap \ker\{E\}$, for almost all $\rho \in [0,t_1]$ which implies

(3.18) $\quad \frac{\partial}{\partial s} V(t,s)z_o = - V(t,s)\tilde{A}z_o \quad$ a.e. for $z_o \in \mathcal{D}$, $t \geq s$.

<u>Proof</u>

For $z_o \in D(A)$,

$$\int_o^t V(t,\rho)[A + BF(\rho)]z_o d\rho$$

$$= \int_o^t T_{t-\rho}[A + BF(\rho)]z_o d\rho + \int_o^t\int_\rho^t T_{t-s}BF(s)V(s,\rho)[A + BF(\rho)]z_o ds d\rho$$

$$= \int_o^t T_{t-\rho}Az_o d\rho + \int_o^t T_{t-\rho}BF(\rho)z_o d\rho + \int_o^t T_{t-s}BF(s)\int_o^s V(s,\rho)[A + BF(\rho)]z_o d\rho ds$$

$$= T_t z_o - z_o + \int_o^t T_{t-s}BF(s)\{\int_o^s V(s,\rho)[A + BF(\rho)]z_o d\rho + z_o\}ds$$

since A the infinitesimal generator of T_t .

Hence $\int_o^t V(t,\rho)[A + BF(\rho)]z_o d\rho + z_o$ satisfies (3.4) which we know has unique solution $V(t,o)z_o$: and thus

$$\int_o^t V(t,\rho)[A + BF(\rho)]z_o d\rho = V(t,o)z_o - z_o \quad \text{for} \quad z_o \in D(A) .$$

(3.15) is thus established.

Now let $z_o \in \mathcal{D}$, $x \in H$ then

$$<x, \int_o^t V(t,\rho)\tilde{A}z_o d\rho> = <x, \int_o^t T_{t-\rho}\tilde{A}z_o d\rho> + <x, \int_o^t\int_\rho^t T_{t-s}BF(s)V(s,\rho)\tilde{A}dsd\rho>$$

$$= <\int_o^t T^*_{t-\rho}xd\rho, \tilde{A}z_o> + <x, \int_o^t\int_o^s T_{t-s}BF(s)V(s,\rho)\tilde{A}z_o d\rho ds>$$

which gives, if we assume a Green's formula (2.7), since $\int_o^t T^*_{t-\rho}xd\rho \in D(A^*)$ for all $x \in H$,

$$<x, \int_o^t V(t,\rho)\tilde{A}z_o d\rho>$$

$$= <A^*\int_o^t T^*_{t-\rho}xd\rho, z_o> + <C\int_o^t T^*_{t-\rho}xd\rho, Dz_o> + <x, \int_o^t T_{t-s}BF(s)\int_o^s V(s,\rho)\tilde{A}z_o d\rho ds>$$

$$= <T^*_t x - x, z_o> + <x, \int_o^t T_{t-\rho}BF(\rho)z_o d\rho> + <x, \int_o^t T_{t-s}BF(s)\int_o^s V(s,\rho)\tilde{A}z_o d\rho ds>$$

since $A^*\int_o^t T^*_{t-\rho}xd\rho = T^*_t x - x$ and $D = F(\rho)$, $E = 0$ in the Green's formula

$$= <x, T_t z_o - z_o> + <x, \int_o^t T_{t-\rho}BF(\rho)z_o d\rho> + <x, \int_o^t T_{t-s}BF(s)\int_o^s V(s,\rho)\tilde{A}z_o d\rho ds> .$$

Hence

$$<x, \int_o^t V(t,\rho)\tilde{A}z_o d\rho + z_o> = <x, T_t z_o + \int_o^t T_{t-s}BF(s)\{\int_o^s V(s,\rho)\tilde{A}z_o d\rho + z_o\}ds>$$

for all $x \in H$, and so for $z_o \in \mathcal{D}$, $\int_o^t V(t,\rho)\tilde{A}z_o d\rho + z_o$ satisfies (3.4) and thus

$$\int_o^t V(t,\rho)\tilde{A}z_o d\rho = V(t,o)z_o - z_o$$

establishing (3.17).

4. THE LINEAR QUADRATIC COST REGULATOR PROBLEM.

In {11} Curtain and Pritchard consider the linear quadratic cost problem and develop the Riccati equation for the control problem on a Hilbert space H

$$(4.1) \qquad z(t) = T_{t-t_o}z_o + \int_{t_o}^{t} T_{t-s}Bu(s)ds \quad , \quad z(t_o) = z_o \in H$$

where T_t and B are assumed to satisfy (2.4) with $p \geq 2$ and $u \in L^2[t_o,t_1;U]$, where U is also a Hilbert space.

The performance index is taken to be

$$(4.2) \qquad J(u;t_o,z_o) = \langle z(t_1), Gz(t_1)\rangle + \int_{t_o}^{t_1}\{\langle z(s), Mz(s)\rangle + \langle u(s), Ru(s)\rangle\}ds$$

where M , $G \in \mathcal{L}(H)$ are self-adjoint and non-negative and R^{-1}, $R \in \mathcal{L}(U)$ is self-adjoint and strictly positive.

If we now assume instead T_t and B satisfy (2.4) with $2 > p \geq 1$ and still consider controls $u \in L^2[t_o,t_1;U]$, then solutions $z(t)$ to (4.1) are no longer continuous but we do have $z \in L^2[t_o,t_1;H]$, since, from (4.1)

$$||z(t)|| \leq ||T_{t-t_o}z_o|| + \int_{t_o}^{t} g(t-s)||B||_{\mathcal{L}(U,\tilde{W})}||u(s)||ds$$

Thus

$$\int_{t_o}^{t_1}||z(t)||^2dt \leq 2\int_{t_o}^{t_1}||T_{t-t_o}z_o||^2dt + 2\int_{t_o}^{t_1}\left|\left|\int_{t_o}^{t} g(t-s)||B||_{\mathcal{L}(U,\tilde{W})}||u(s)||ds\right|\right|^2dt < \infty$$

with the second term on the right hand side being finite since the convolution of an L^2 and an L^1 function is L^2 .

Since for a general $u \in L^2[t_o,t_1;U]$ the solutions z are only $L^2[t_o,t_1;H]$, not continuous, a terminal cost in the performance index would be meaningless so we consider a performance index of pure integral form, namely

$$J(u;t_o,z_o) = \int_{t_o}^{t_1}\{\langle z(s), Mz(s)\rangle + \langle u(s), Ru(s)\rangle\}ds$$

where M and R satisfy the same conditions as before.

We will see later that for a pure integral cost functional the perturbation operators $BF(t)$ will be such that $F(t)$ satisfies (3.3) with $q = \infty$ and thus the perturbation result, theorem 3.1, holds for $p \geq 1$.

Consider now the sequence of feedback controls for $k = 0,1,2,\ldots$

(4.3) $\quad u_k(t) = - F_k(t)z(t)$

where

(4.4) $\quad F_k(t) = R^{-1}B^*Q_{k-1}(t) \quad , \quad F_o = 0$

(4.5) $\quad M_k(t) = M + F_k^*(t)RF_k(t)$

(4.6) $\quad Q_k(t)x = \int_t^{t_1}\mathcal{U}_k^*(s,t)M_k(s)\mathcal{U}_k(s,t)xds \quad , \quad x \in H$

and where $\mathcal{U}_k(t,s)$ is the perturbation of T_t by $- BF_k(t)$, that is

(4.7) $\quad \mathcal{U}_k(t,s)x = T_{t-s}x - \int_s^t T_{t-\rho}BF_k(\rho)\mathcal{U}_k(\rho,s)xd\rho \quad , \quad x \in H$.

The iterative scheme is well defined since, for $k = 1$, we have $F_1(t) = R^{-1}B^*Q_o(t)$ with $Q_o(t)x = \int_t^{t_1}T_{s-t}^*MT_{s-t}xds$ and as

$$||T_t||_{\bar{\mathcal{L}}(W,H)} \leq g(t) \quad , \quad g \in L^1[t_o,t_1]$$

$$||T_t||_{\bar{\mathcal{L}}(H)} \leq \overline{M} \qquad \text{for} \quad t \in [t_o,t_1]$$

we find

$$||Q_o||_{\bar{\mathcal{L}}(W,H)} \leq \int_t^{t_1}\overline{M}||M||g(s-t)ds \leq \text{constant} < \infty \; .$$

Thus, by theorem 3.1, $\mathcal{U}_1(t,s)$ is well defined by (4.7) with

$||\mathcal{U}_1(t,s)|| \leq \bar{M}_1$, $t_1 \geq t \geq s \geq t_o$, and

$$||\mathcal{U}_1(t,s)||_{\mathcal{L}(\tilde{W},H)} \leq g(t-s) + \int_s^t g(t-\rho)||B||_{\mathcal{L}(U,\tilde{W})}||F_1(\rho)||_{\mathcal{L}(H,U)}||\mathcal{U}_1(\rho,s)||_{\mathcal{L}(\tilde{W},H)}d\rho$$

so using the Generalised Gronwall's Inequality, corollary 3.3,

$$||\mathcal{U}_1(t,s)||_{\mathcal{L}(\tilde{W},H)} \leq g_1(t-s) \quad \text{where} \quad g_1 \in L^1[t_o,t_1] \quad .$$

From (4.5)

$$||M_1(t)|| \leq ||M|| + ||R||\ ||F_1(t)||^2 \leq \text{constant} < \infty ,$$

thus (4.6) is well defined for $k = 1$ and

$$||Q_1(t)||_{\mathcal{L}(\tilde{W},H)} \leq \int_t^{t_1} \bar{M}_1||M_1(t)||g_1(s-t)ds \leq \text{constant} < \infty .$$

Hence, also, $||F_2(t)||_{\mathcal{L}(H,U)} \leq \text{constant} < \infty$.

So the iterative scheme is well defined for $k = 1$ and the estimates on $\mathcal{U}_1(t,s)$ and $F_2(t)$ are similar to those assumed for T_{t-s} and $F_1(t)$.

Using a simple inductive argument we have that the iterative scheme is well defined for all k .

With the control given by (4.3), (4.1) becomes

(4.8) $\quad z(t) = \mathcal{U}_k(t,t_o)z_o \quad .$

If we also consider the controlled version of (4.8)

(4.9) $\quad z(t) = \mathcal{U}_k(t,t_o)z_o + \int_{t_o}^t \mathcal{U}_k(t,s)B\bar{u}(s)ds$

for some $\bar{u} \in L^2[t_o,t_1;U]$, then the following lemma is easy to prove by direct substitution for $z(t)$ given by (4.9) and $Q_k(t)$ given by (4.6).

<u>Lemma 4.1</u>

$$\langle z(t),\ Q_k(t)z(t)\rangle = \int_t^{t_1}\{\langle z(s),\ M_k(s)z(s)\rangle - 2\langle z(s),\ Q_k(s)B\bar{u}(s)\rangle\}ds$$

where $z(t)$ is given by (4.9).

Putting $t = t_o$ and $\bar{u} = 0$ in lemma 4.1 gives

$$\langle z_o,\ Q_k(t_o)z_o\rangle = \int_{t_o}^{t_1}\langle z(s),\ \{M + F_k^*(s)RF_k(s)\}z(s)\rangle ds = J(u_k;t_o,z_o)$$

and lemma 4.1 with $t = t_o$, $\bar{u}(t) = \{F_k(t) - F_{k+1}(t)\}z(t)$ yields

$$\langle z_o,\ Q_k(t_o)z_o\rangle$$

$$= \int_{t_o}^{t_1}\{\langle z(s),\ Mz(s)\rangle + \langle F_k(s)z(s),\ RF_k(s)\rangle + 2\langle F_{k+1}(s)z(s),\ R\{F_{k+1}(s) - F_k(s)\}z(s)\rangle\}ds$$

$$= \int_{t_o}^{t_1}\{\langle z(s),\ Mz(s)\rangle + \langle F_{k+1}(s)z(s),\ RF_{k+1}(s)z(s)\rangle$$
$$+ \langle\{F_{k+1}(s) - F_k(s)\}z(s),\ R\{F_{k+1}(s) - F_k(s)\}z(s)\rangle\}ds$$

$$= J(u_{k+1};t_o,z_o) + \int_{t_o}^{t_1}\langle\{F_{k+1}(s) - F_k(s)\}z(s),\ R\{F_{k+1}(s) - F_k(s)\}z(s)\rangle ds$$

$\geq J(u_{k+1};t_o,z_o)$ since R strictly positive

$= \langle z_o,\ Q_{k+1}(t_o)z_o\rangle$ from above.

Thus we have shown

<u>Lemma 4.2</u>

The cost for (4.1) with feedback control (4.3) is

$$J(u_k;t_o,z_o) = \langle z_o,\ Q_k(t_o)z_o\rangle\ .$$

<u>Remark</u> Curtain and Pritchard {11} define $u_k = F_k z_k$ and $F_k = -R^{-1}B^*Q_{k+1}$ and hence the proof of their similar result, for bounded control problems, differs slightly.

For fixed $z_o \in H$ and fixed $t_o < t_1$, $\langle z_o, Q_k(t_o)z_o \rangle$ is monotonically decreasing in k, with $\langle z_o, Q_k(t_o)z_o \rangle \leq \langle z_o, Q_o(t_o)z_o \rangle$ for all $z_o \in H$, and hence $Q_k(t)$ converges strongly as $k \to \infty$ to a self-adjoint operator $Q(t)$ for each $t \in [t_o, t_1]$.

Before we can derive an integral expression for $Q(t)$ it is necessary to show that the estimates obtained are uniform. In this regard we have,

__Lemma 4.3__

$F_k(t)$ and $\mathcal{U}_k(t,s)$ are uniform in the sense that there exists a function $g_\infty \in L^1[t_o, t_1]$ and constants $\bar{M}_\infty$, f_∞ such that

$$||\mathcal{U}_k(t,s)||_{\mathcal{L}(H)} \leq \bar{M}_\infty \qquad t_1 \geq t \geq s \geq t_o, \ \forall k = 0,1,$$

$$||F_k(t)||_{\mathcal{L}(H,U)} \leq f_\infty \qquad \forall k = 0,1,...$$

$$||\mathcal{U}_k(t,s)||_{\mathcal{L}(\tilde{W},H)} \leq g_\infty(t-s) \qquad \forall k = 0,1,...$$

__Proof__

From lemma 4.2 there exists a constant c such that

$$\sup_{t \in [t_o, t_1]} ||Q_k(t)|| \leq c \quad , \quad k = 0,1,...$$

Also from lemma 4.2, $\langle Q_k(t)x, x \rangle \leq \langle Q_{k-1}(t)x, x \rangle$, so

(4.10) $\quad |\langle Q_k(t)x, x\rangle| \leq |\langle Q_{k-1}(t)x, x\rangle| \quad , \quad x \in H$

But, since $Q_k(t)$ is self-adjoint,

$$||Q_k(t)|| = \sup_{||x||=1} |\langle Q_k(t)x, x\rangle| \leq \sup_{||x||=1} |\langle Q_{k-1}(t)x, x\rangle|$$

by (4.10).

Hence, $||Q_k(t)|| \leq ||Q_{k-1}(t)||$.

We can extend $Q_k(t)$ to an operator from $\tilde{W}$ to H with the same norm (see for example, Kantovorich and Akilov {17}), thus

$$||Q_k(t)||_{\mathcal{L}(\tilde{W},H)} \leq ||Q_{k-1}(t)||_{\mathcal{L}(\tilde{W},H)}$$

and we have already shown that there exists a constant σ such that $||Q_1(t)||_{\mathcal{L}(\tilde{W},H)} \leq \sigma$, independent of t .

Thus, $||Q_k(t)||_{\mathcal{L}(\tilde{W},H)}$ is uniformly bounded in k .

Now since $F_k(t) = R^{-1}B^*Q_{k-1}(t)$

$$||F_k(t)||_{\mathcal{L}(H,U)} = ||F_k^*(t)\ ||_{\mathcal{L}(U,H)} \leq ||R^{-1}||\ ||B||_{\mathcal{L}(U,\tilde{W})} ||Q_{k-1}(t)||_{\mathcal{L}(\tilde{W},H)}$$

and hence there exists a f_∞ such that

$$||F_k(t)||_{\mathcal{L}(H,U)} \leq f_\infty \qquad k = 0,1,\ldots \quad .$$

From (4.7)

$$||U_k(t,s)|| \leq \bar{M} + \int_s^t g(t-\rho)f_\infty ||B||_{\mathcal{L}(U,\tilde{W})} ||U_k(\rho,s)||d\rho$$

and so by the Generalised Gronwall's inequality,

$$||U_k(t,s)|| \leq \bar{M} + \bar{M}\sum_{n=1}^{\infty} G^n(t-s) \leq \bar{M}_\infty \quad \text{independent of } k \ .$$

Similarly,

$$||U_k(t,s)||_{\mathcal{L}(\tilde{W},H)} \leq g(t-s) + \int_s^t g(t-\rho)||B||_{\mathcal{L}(U,\tilde{W})} f_\infty ||U_k(\rho,s)||_{\mathcal{L}(\tilde{W},H)} d\rho$$

so again by the Generalised Gronwall's Inequality there exists a function $g_\infty \in L^2[t_o,t_1]$ such that

$$||U_k(t,s)||_{\mathcal{L}(\tilde{W},H)} \leq g_\infty(t-s) \quad , \qquad k = 0,1,\ldots \ .$$

Thus lemma 4.3 is proved and we are in a position to prove the following result on the existence of a unique optimizing control.

__Theorem 4.4__

The optimal control which minimizes $J(u;t_o,z_o)$ is the feedback control

$$(4.11) \quad u^*(t) = -R^{-1}B^*Q(t)z(t)$$

where $Q(t)$ is the unique solution of

$$(4.12) \quad Q(t)x = \int_t^{t_1} \mathcal{U}^*(s,t)[M + Q(s)BR^{-1}B^*Q(s)]\mathcal{U}(s,t)x\,ds$$

and the evolution operator $\mathcal{U}(t,s)$ is given by

$$(4.13) \quad \mathcal{U}(t,s)x = T_{t-s}x - \int_s^t T_{t-\rho}BR^{-1}B^*Q(\rho)\mathcal{U}(\rho,s)x\,d\rho \ .$$

Furthermore, the cost of this optimal control is

$$J(u^*;t_o,z_o) = \langle z_o, Q(t_o)z_o\rangle \ .$$

__Proof__

Since $||Q_k(t)||_{\mathcal{L}(\tilde{W},H)}$ is uniformly bounded and $Q_k(t) \to Q(t)$ strongly on H as $k \to \infty$, then $||Q_k(t)w - Q(t)w|| \to 0$ as $k \to \infty$ for all $w \in \tilde{W}$.

It follows that $F_k(t)$ and $M_k(t)$ converge strongly to $F_\infty(t) = R^{-1}B^*Q(t)$ and $M_\infty(t) = M + Q(t)BR^{-1}B^*Q(t)$ respectively, and we have $\mathcal{U}_k(t,s) \to \mathcal{U}(t,s)$ strongly by the Generalised Gronwall's Inequality, using the uniform bounds established in lemma 4.3 and the strong convergence of $F_k(t)$ to $F_\infty(t)$.

Employing the Lebesgue Dominated Convergence Theorem we have that $Q(t)$ satisfies (4.12).

Consider now an arbitrary admissible control $u \in L^2[t_o,t_1;U]$ so that the controlled system is

$$z(t) = T_{t-t_o}z_o + \int_{t_o}^{t} T_{t-s}Bu(s)ds \ .$$

Since $\mathcal{U}(t,s)$ is the perturbation of T_t by $-BR^{-1}B^*Q(t)$ we also have

$$z(t) = \mathcal{U}(t,t_o)z_o + \int_{t_o}^{t} \mathcal{U}(t,s)B\bar{u}(s)ds$$

where $\bar{u}(t) = u(t) + R^{-1}B^*Q(t)z(t)$.

Applying lemma 4.1 with $t = t_o$ gives

$<z_o, Q(t_o)z_o>$

$$= \int_{t_o}^{t_1}\{<z(s), Mz(s)> + <R^{-1}B^*Q(s)z(s), B^*Q(s)z(s)> - 2<z(s), Q(s)B\bar{u}(s)>\}ds$$

$$= \int_{t_o}^{t_1}\{<z(s), Mz(s)> + <u(s), Ru(s)>\}ds - \int_{t_o}^{t_1}<\bar{u}(s), R\bar{u}(s)>ds \ .$$

From (4.12)

$$<z_o, Q(t_o)z_o> = \int_{t_o}^{t_1}\{<z(s), Mz(s)> + <u^*(s), Ru^*(s)>\}ds$$

and so

$$J(u^*;t_o,z_o) = J(u;t_o,z_o) - \int_{t_o}^{t_1}<\bar{u}(s), R\bar{u}(s)>ds$$

and hence u^* is optimal with optimal cost

$$J(u^*;t_o,z_o) = <z_o, Q(t_o)z_o> \ .$$

For uniqueness:- Suppose $P(t)$ is another solution of (4.12), that is,

(4.14) $$P(t)x = \int_{t}^{t_1}\mathcal{U}_p^*(s,t)[M + P(s)BR^{-1}B^*P(s)]\mathcal{U}_p(s,t)xds$$

where $\mathcal{U}_p(t,s)$ is the perturbation of T_t by $-BR^{-1}B^*P(t)$, i.e.

(4.15) $$\mathcal{U}_p(t,s)x = T_{t-s}x - \int_{s}^{t}T_{t-\rho}BR^{-1}B^*P(\rho)\mathcal{U}_p(\rho,s)xd\rho \ .$$

Consider the system

$$(4.16)\qquad z(t) = T_{t-t_o} z_o + \int_{t_o}^{t} T_{t-s} Bu(s)ds$$

for some $u \in L^2[t_o, t_1; U]$. Then, using (4.15) and substituting u_o for u in (4.16) we obtain

$$(4.17)\qquad z(t) = \mathcal{U}_p(t,t_o)z_o + \int_{t_o}^{t} \mathcal{U}_p(t,s)B\bar{u}_o(s)ds$$

where $\bar{u}_o(t) = u_o(t) + R^{-1}B^*P(t)z(t)$. Lemma 4.1 for $P(t)$ given by (4.14) and $z(t)$ given by (4.17) gives

$$(4.18)\qquad \langle z(t), P(t)z(t)\rangle = \int_{t}^{t_1} \{\langle z(s), Mz(s)\rangle + \langle B^*P(s)z(s), R^{-1}B^*P(s)z(s)\rangle - 2\langle z(s), P(s)Bu_o(s)\rangle\}ds$$

The cost of the control $u_o(t)$ is given by

$$J(u_o;t_o,z_o)$$

$$= \int_{t_o}^{t_1} \{\langle z(s), Mz(s)\rangle + \langle u_o(s), Ru_o(s)\rangle\}ds$$

$$= \int_{t_o}^{t_1} \{\langle z(s), Mz(s)\rangle + \langle B^*P(s)z(s), R^{-1}B^*P(s)z(s)\rangle + \langle \bar{u}_o(s), R\bar{u}_o(s)\rangle - 2\langle \bar{u}_o(s), B^*P(s)z(s)\rangle\}ds \ .$$

Hence, from (4.18)

$$(4.19)\qquad J(u_o;t_o,z_o) = \langle z_o, P(t_o)z_o\rangle + \int_{t_o}^{t_1} \langle \bar{u}_o(s), R\bar{u}_o(s)\rangle ds \ .$$

Since $\mathcal{U}(t,s)$ is the perturbation of T_t by $-BR^{-1}B^*Q(t)$, (4.16) is equivalent to

$$(4.20)\qquad z(t) = \mathcal{U}(t,t_o)z_o + \int_{t_o}^{t} \mathcal{U}(t,s)B\bar{u}(s)ds$$

where $\bar{u}(t) = u(t) + R^{-1}B^*Q(t)z(t)$. So, arguing as above, for $Q(t)$ given by (4.12) and $z(t)$ given by (4.20) we can show that

(4.21) $$J(u;t_o,z_o) = \langle z_o, Q(t_o)z_o\rangle + \int_{t_o}^{t_1} \langle \bar{u}(s), R\bar{u}(s)\rangle ds$$

now, if $u(t) = -R^{-1}B^*P(t)z(t)$ in (4.16) we have $\bar{u}(t) = R^{-1}B^*[Q(t) - P(t)]z(t)$ in (4.20) or $z(t) = \mathcal{U}_p(t,t_o)z_o$. For this particular choice of u, by lemma 4.2 and (4.21) we obtain

(4.22) $$\begin{aligned} J(u;t_o,z_o) &= \langle z_o, Q(t_o)z_o\rangle \\ &\quad + \int_{t_o}^{t_1} \langle R^{-1}B^*[Q(t) - P(t)]\mathcal{U}_p(t,t_o)z_o, B^*[Q(t) - P(t)]\mathcal{U}_p(t,t_o)z_o\rangle dt \\ &= \langle z_o, P(t_o)z_o\rangle . \end{aligned}$$

Similarly, choosing $u_o(t) = -R^{-1}B^*Q(t)z(t)$ in (4.16), and using lemma 4.2 and (4.19) we obtain

(4.23) $$\begin{aligned} J(u_o;t_o,z_o) &= \langle z_o, P(t_o)z_o\rangle \\ &\quad + \int_{t_o}^{t_1} \langle R^{-1}B^*[P(t) - Q(t)]\mathcal{U}(t,t_o)z_o, B^*[P(t) - Q(t)]\mathcal{U}(t,t_o)z_o\rangle dt \\ &= \langle z_o, Q(t_o)z_o\rangle . \end{aligned}$$

Adding (4.22) and (4.23) gives

$$\int_{t_o}^{t_1} ||R^{-1/2}B^*[P(t) - Q(t)]\mathcal{U}(t,t_o)z_o||^2 dt = 0$$

and

$$\int_{t_o}^{t_1} ||R^{-1/2}B^*[Q(t) - P(t)]\mathcal{U}_p(t,t_o)z_o||^2 dt = 0 .$$

Since R is strictly positive we have

$$B^*Q(t)\mathcal{U}_p(t,t_o)z_o = B^*P(t)\mathcal{U}_p(t,t_o)z_o$$

and

$$B^*Q(t)\mathcal{U}(t,t_o)z_o = B^*P(t)\mathcal{U}(t,t_o)z_o$$

on $[t_o,t_1]$, which implies $\mathcal{U}(t,t_o) = \mathcal{U}_p(t,t_o)$ by theorem 3.3, and so

$$B^*Q(t)\mathcal{U}(t,t_o)z_o = B^*P(t)\mathcal{U}(t,t_o)z_o = B^*P(t)\mathcal{U}_p(t,t_o)z_o \quad .$$

Thus from (4.12) and (4.14)

$$[P(t) - Q(t)]x$$

$$= \int_t^{t_1}\mathcal{U}_p^*(s,t)P(s)BR^{-1}B^*P(s)\mathcal{U}_p(s,t)xds - \int_t^{t_1}\mathcal{U}^*(s,t)Q(s)BR^{-1}B^*Q(s)\mathcal{U}(s,t)xds$$

$$= 0 \quad .$$

This completes the proof.

Remarks

The results of theorem 4.4 are the same as those for the case $p \geq 2$ to be found in {11}.

As in {11} we can also show (by substitution for $\mathcal{U}(t,s)$ from (4.13) into (4.12)) that the Riccati equation has the following alternative form

$$(4.24) \qquad Q(t)x = \int_t^{t_1}\mathcal{U}^*(s,t)MT_{t-s}xds \quad .$$

Also since, from proposition 3.4, $\mathcal{U}(t,s)$ is a quasi-evolution operator, we can differentiate the integral Riccati equation (4.12), two different expressions being obtained depending on whether we use (3.16) or (3.18) (where we take $F(t) = - R^{-1}B^*Q(t)$).

Theorem 4.5

$Q(t)$ is the unique solution to the differential equation

$$(4.25) \qquad \frac{d}{dt}\langle Q(t)x,\ y\rangle + \langle Q(t)x,\ Ay\rangle + \langle Ax,\ Q(t)y\rangle = \langle Q(t)BR^{-1}B^*Q(t)x,\ y\rangle - \langle Mx,\ y\rangle$$

for $x, y \in D(A)$, with $Q(t_1) = 0$.

If we assume the existence of a Green's formula (2.7) with $B^* = C$, then, for $x \in \mathcal{D} = D(\tilde{A}) \cap \ker\{D - R^{-1}B^*Q(\rho)\}$, $Q(t)$ satisfies the differential equation

$$(4.26) \qquad \frac{d}{dt} \langle Q(t)x, y\rangle + \langle Q(t)x, \tilde{A}y\rangle + \langle \tilde{A}x, Q(t)y\rangle = - \langle Q(t)BR^{-1}B^*Q(t)x, y\rangle - \langle Mx, y\rangle$$

$$Q(t_1) = 0 \quad .$$

Further, if $\overline{\mathcal{D}} = H$, then $Q(t)$ is the unique solution to (4.26).

Proof

In order to justify the formal differentiation we need the following lemmas, proofs of which may be found in {9} ,

Lemma 4.6

Let $f : (0,t_1)\times(0,t_1) \to R$ be integrable and suppose

(i) for almost all s , $f(s,\cdot)$ is absolutely continuous

(ii) $f(s,s) \in L^1[0,t_1]$

(iii) $\int_0^{t_1} ds \int_s^{t_1} \left|\frac{\partial f}{\partial s}(s,\tau)\right| d\tau < \infty$

then $g(t) = \int_0^t f(s,t)ds$ is absolutely continuous with

$$g'(t) = f(t,t) + \int_0^t \frac{\partial f}{\partial t}(s,t)ds \quad \text{a.e.}$$

Lemma 4.7

Let H be a real Hilbert space and $W \in \mathcal{L}(H)$. Suppose $g_1(\cdot)$ and $g_2(\cdot)$ are weakly absolutely continuous H-valued functions on $[0,t_1]$ such that

$$\langle g_i(t),\ x\rangle = \langle g_i(o),\ x\rangle + \int_o^t \frac{\partial}{\partial s}\langle g_i(s),\ x\rangle ds \qquad \forall\ x \in H\ ,\ i = 1,2,\ldots\ .$$

Then $f(t) = \langle Wg_1(t),\ g_2(t)\rangle$ is an absolutely continuous function with

$$\langle Wg_1(t),\ g_2(t)\rangle = \langle Wg_1(o),\ g_2(o)\rangle + \int_o^t \frac{\partial}{\partial s}\langle Wg_1(s),\ g_2(s)\rangle ds\ .$$

<u>Lemma 4.8</u>

Let H be a real Hilbert space and suppose $P(\cdot)$ is a weakly absolutely continuous $\mathcal{L}(H)$-valued function and $g(\cdot)$ is strongly differentiable with the representation

$$g(t) = g(o) + \int_o^t g'(s)ds$$

then $P(\cdot)g(\cdot)$ is weakly absolutely continuous with

$$\frac{d}{dt}\langle P(t)g(t),\ x\rangle = \langle P(t)g'(t),\ x\rangle + \frac{\partial}{\partial t}\langle P(t)g(s),\ x\rangle\Big|_{s=t} \quad \text{a.e. on } [0,t_1].$$

Thus differentiating (4.12) we have

$$\frac{d}{dt}\langle Q(t)x,\ y\rangle = \frac{d}{dt}\left\langle\int_t^{t_1} \mathcal{U}^*(s,t)[M + Q(s)BR^{-1}B^*Q(s)]\mathcal{U}(s,t)x ds,\ y\right\rangle$$

$$= \frac{d}{dt}\int_t^{t_1}\langle M\mathcal{U}(s,t)x,\ \mathcal{U}(s,t)y\rangle ds + \frac{d}{dt}\int_t^{t_1}\langle Q(s)BR^{-1}B^*Q(s)\mathcal{U}(s,t)x,\ \mathcal{U}(s,t)y\rangle ds$$

which by the above lemmas

$$= -\langle Mx,\ y\rangle + \int_t^{t_1}\frac{d}{dt}\langle M\mathcal{U}(s,t)x,\ \mathcal{U}(s,t)y\rangle ds - \langle Q(t)BR^{-1}B^*Q(t)x,\ y\rangle$$

$$+ \int_t^{t_1}\frac{d}{dt}\langle Q(s)BR^{-1}B^*Q(s)\mathcal{U}(s,t)x,\ \mathcal{U}(s,t)y\rangle ds\ .$$

So applying (3.16) we obtain

$$\frac{d}{dt}\langle Q(t)x,\ y\rangle = -\langle Mx,\ y\rangle - \langle Q(t)BR^{-1}B^*Q(t)x,\ y\rangle$$

$$- \int_t^{t_1}\langle M\mathcal{U}(s,t)x,\ \mathcal{U}(s,t)[A - BR^{-1}B^*Q(t)]y\rangle ds$$

$$- \int_t^{t_1} \langle M\mathcal{U}(s,t)[A - BR^{-1}B^*Q(t)]x,\ \mathcal{U}(s,t)y\rangle ds$$

$$- \int_t^{t_1} \langle Q(s)BR^{-1}B^*Q(s)\mathcal{U}(s,t)[A - BR^{-1}B^*Q(t)]x,\ \mathcal{U}(s,t)y\rangle ds$$

$$- \int_t^{t_1} \langle Q(s)BR^{-1}B^*Q(s)\mathcal{U}(s,t)x,\ \mathcal{U}(s,t)[A - BR^{-1}B^*Q(t)]y\rangle ds$$

$$= - \langle Mx, y\rangle - \langle Q(t)BR^{-1}B^*Q(t)x, y\rangle - \langle Q(t)x, [A - BR^{-1}B^*Q(t)]y\rangle$$
$$- \langle [A - BR^{-1}B^*Q(t)]x, Q(t)y\rangle$$

as required, and obviously $Q(t_1) = 0$.

For uniqueness: Let $Q_1(t)$ and $Q_2(t)$ be two solutions of (4.25) i.e.

$$\frac{d}{dt}\langle Q_i(t)x, y\rangle = \langle Q_i(t)BR^{-1}B^*Q_i(t)x, y\rangle - \langle Mx, y\rangle - \langle Q_i(t)x, Ay\rangle - \langle Ax, Q_i(t)y\rangle$$
for $i = 1,2$.

So writing $P(t) = Q_1(t) - Q_2(t)$ we have

$$(4.26)\quad \frac{d}{dt}\langle P(t)x, y\rangle = \langle P(t)BR^{-1}B^*Q_1(t)x, y\rangle - \langle Ax, P(t)y\rangle - \langle P(t)x, Ay\rangle$$
$$+ \langle x, P(t)BR^{-1}B^*Q_1(t)y\rangle - \langle P(t)BR^{-1}B^*P(t)x, y\rangle$$

and

$$(4.27)\quad \frac{d}{dt}\langle P(t)x, y\rangle = \langle P(t)BR^{-1}B^*Q_2(t)x, y\rangle - \langle Ax, P(t)y\rangle - \langle P(t)x, Ay\rangle$$
$$+ \langle x, P(t)BR^{-1}B^*Q_2(t)y\rangle - \langle P(t)BR^{-1}B^*P(t)x, y\rangle\ .$$

Define $F(t)x = \int_t^{t_1} \mathcal{U}_1^*(s,t)P(s)BR^{-1}B^*P(s)\mathcal{U}_1(s,t)x\,ds$ where $\mathcal{U}_1(t,s)$ is the perturbation of T_t by $- BR^{-1}B^*Q_1(t)$ then, for $x, y \in D(A)$, we may differentiate $\langle F(t)x, y\rangle$ to obtain

$$(4.28)\quad \frac{d}{dt}\langle F(t)x, y\rangle = \langle F(t)BR^{-1}B^*Q_1(t)x, y\rangle - \langle Ax, F(t)y\rangle - \langle F(t)x, Ay\rangle$$
$$+ \langle x, F(t)BR^{-1}B^*Q_1(t)y\rangle - \langle P(t)BR^{-1}B^*P(t)x, y\rangle$$

with $F(t_1) = 0$.

If we assume (4.28) has unique solution on $D(A)$ we have $F(t) = P(t)$ and

$$\langle P(t)x,\ x\rangle = \int_t^{t_1} \langle \mathcal{U}_1^*(s,t)P(s)BR^{-1}B^*P(s)\mathcal{U}_1(s,t)x,\ x\rangle ds$$

$$\geq 0 \quad \text{for all } x \in D(A) \text{ , dense in } H \text{ .}$$

Similarly if we use (4.27) and $Q_2(t)$ perturbations we find $\langle P(t)x,\ x\rangle \leq 0$ for all $x \in D(A)$, dense in H , and hence $P(t) = 0$. All that remains is to show the uniqueness of (4.28).

Let $S(t) = T^*_{t-s}F(t)T_{t-s}$, then, for $x,\ y \in D(A)$, $\langle x,\ S(t)y\rangle$ is differentiable in t and

$$\frac{d}{dt}\langle x,\ S(t)y\rangle = \langle F(t)T_{t-s}x,\ BR^{-1}B^*Q_1(t)T_{t-s}y\rangle + \langle BR^{-1}B^*Q_1(t)T_{t-s}x,\ F(t)T_{t-s}y\rangle$$

so $\langle T^*_{t-s}F(t)T_{t-s}x,\ y\rangle = \langle x,\ S(t)y\rangle$

$$= -\int_t^{t_1} \langle [Q_1(\rho)BR^{-1}B^*F(\rho) + F(\rho)BR^{-1}B^*Q_1(\rho)]T_{\rho-s}x,\ T_{\rho-s}y\rangle d\rho$$

Hence, since $\overline{D(A)} = H$

$$||B^*F(t)x|| \leq 2\int_t^{t_1} ||Q_1(\rho)||_{\bar{\mathcal{L}}(W,H)} ||B||^2_{\bar{\mathcal{L}}(U,W)} ||R^{-1}||\ ||T_{\rho-t}||_{\bar{\mathcal{L}}(W,H)} ||B^*F(\rho)x||d\rho$$

By the Generalised Gronwall's Inequality we have $||B^*F(t)x|| = 0$ and thus $||F(t)|| = 0$.

Proof that $Q(t)$ satisfies (4.26) under the assumption that there exists a Green's formula (2.7) with $C = B^*$ follows similarly, with uniqueness if $\overline{\mathcal{D}} = H$.

<u>Example 4.9</u>

Suppose we consider again the controlled system of example 2.7, i.e.

(4.29) $z_t = z_{xx}$

$z(o,t) = u \quad , \quad z(1,t) = 0 \quad , \quad z(x,o) = z_o(x) \quad ,$

and take as cost functional

(4.30) $J(u) = \int_o^{t_1} [\int_o^1 z^2(x,t)dx + cu^2(t)]dt$

then we know from example 2.7 that by setting $\Omega = (0,1)$, $H = L^2(\Omega)$, $U = R$ and $A = \frac{\partial^2}{\partial x^2}$ on $D(A) = H^2(0,1) \cap H_o^1(0,1)$, A generates the strongly continuous semigroup T_t given by

$$(T_t z_o)(x) = \sum_{n=1}^{\infty} 2e^{-n^2\pi^2 t} \sin n\pi x <z_o(\cdot), \sin n\pi\cdot> .$$

Also $B = -\delta'$ and T_t, B satisfy (2.4) with $4/3 > p \geq 1$ so

(4.31) $z(t) = T_t z_o + \int_o^t T_{t-s} Bu(s)ds$

is a weak solution to (4.29). Furthermore (4.30) is in the form (4.1) with $G = 0$, $M = I$ and $R = c$ thus we know there exists a unique optimal control $u^* = -R^{-1}B^*Qz$ where Q satisfies the differential equation (4.25). If we assume Q has the form

(4.32) $(Qz)(\zeta) = \int_o^1 K(\zeta,\eta,t)z(\eta)d\eta$

then

$$<Qh, Ak> = \int_o^1 (Qh)(\zeta)(Ak)(\zeta)d\zeta$$

$$= \int_o^1 (\int_o^1 K(\zeta,\eta,t)h(\eta)d\eta)(\frac{\partial^2}{\partial\zeta^2} k(\zeta))d\zeta$$

$$= [\int_o^1 K(\zeta,\eta,t)h(\eta)\frac{\partial k}{\partial\zeta} d\eta]_o^1 - \int_o^1\int_o^1 K_\zeta(\zeta,\eta,t)h(\eta)\frac{\partial k}{\partial\zeta} d\eta d\zeta$$

$$= [\int_0^1 K(\zeta,\eta,t)h(\eta)\frac{\partial k}{\partial \zeta}\,d\eta]_0^1 - [\int_0^1 K_\zeta(\zeta,\eta,t)h(\eta)k(\zeta)d\eta]_0^1$$

$$+ \int_0^1\int_0^1 K_{\zeta\zeta}(\zeta,\eta,t)h(\eta)k(\zeta)d\eta d\zeta$$

$$= \int_0^1\int_0^1 K_{\zeta\zeta}(\zeta,\eta,t)h(\eta)k(\zeta)d\eta d\zeta \quad \text{when} \quad K(1,\eta,t) = 0 = K(o,\eta,t) \ .$$

Similarly,

$$\langle Ah,\ Qk\rangle = \int_0^1\int_0^1 K_{\eta\eta}(\zeta,\eta,t)h(\eta)k(\zeta)d\eta d\zeta \quad \text{when} \quad K(\zeta,1,t) = 0 = K(\zeta,o,t)$$

and

$$\frac{d}{dt}\langle Qh,\ k\rangle = \int_0^1\int_0^1 K_t(\zeta,\eta,t)h(\eta)k(\zeta)d\eta d\zeta \ .$$

Also

$$\langle Bu,\ Qk\rangle = \int_0^1 (Bu)(\zeta)(Qk)(\zeta)d\zeta$$

$$= - u\int_0^1\int_0^1 \delta'(\zeta)K(\zeta,\eta,t)k(\eta)d\eta d\zeta$$

$$= - [\int_0^1 \delta(\zeta)K(\zeta,\eta,t)k(\eta)d\eta]_0^1 + u\int_0^1\int_0^1 \delta(\zeta)K_\zeta(\zeta,\eta,t)k(\eta)d\eta d\zeta$$

$$= u\int_0^1 K_\zeta(o,\eta,t)k(\eta)d\eta$$

so $B^*Qk = \int_0^1 K_\zeta(o,\eta,t)k(\eta)d\eta$ and thus

$$\frac{1}{c}\langle B^*Qh,\ B^*Qk\rangle_U = \frac{1}{c}\langle\int_0^1 K_\zeta(o,\eta,t)h(\eta)d\eta,\ \int_0^1 K_\eta(\zeta,o,t)k(\zeta)d\zeta\rangle$$

$$= \frac{1}{c}\int_0^1\int_0^1 K_\zeta(o,\eta,t)K_\eta(\zeta,o,t)h(\eta)k(\zeta)d\eta d\zeta \ .$$

Finally

$$\langle h,\ Mk\rangle = \langle h,\ k\rangle = \int_0^1 h(\zeta)k(\zeta)d\zeta$$

$$= \int_0^1\int_0^1 \delta(\zeta-\eta)h(\eta)k(\zeta)d\eta d\zeta \ .$$

Thus substituting for this Q given by (4.32) with $K(\zeta,1,t) = K(1,\eta,t) = K(\zeta,o,t) = K(o,\eta,t) = 0$ into the differential equation (4.25) leads to the following equation for $K(\zeta,\eta,t)$

(4.33) $$\int_o^1\int_o^1 \{K_t(\zeta,\eta,t) + K_{\zeta\zeta}(\zeta,\eta,t) + K_{\eta\eta}(\zeta,\eta,t) - \frac{1}{c} K_\zeta(o,\eta,t)K_\eta(\zeta,o,t) + \delta(\zeta-\eta)\}k(\zeta)h(\eta)d\eta d\zeta = 0$$

$$K(\zeta,\eta,t_1) = 0$$

which is satisfied if we choose $K(\zeta,\eta,t)$ to satisfy

(4.34) $$K_t + K_{\zeta\zeta} + K_{\eta\eta} - \frac{1}{c} K_\zeta(o,\eta,t)K_\eta(\zeta,o,t) + \delta(\zeta-\eta) = 0\ ,\ K(\zeta,\eta,t_1) = 0$$

[with $K(o,\eta,t) = K(1,\eta,t) = K(\zeta,o,t) = K(\zeta,1,t) = 0$] .

Thus we have shown that the optimal control for (4.29), (4.30) is $u^*(t) = - R^{-1}B^*Q(t)z(t) = - \frac{1}{c}\int_o^1 K_\zeta(o,\eta,t)z(\eta)d\eta$ where K is given by (4.34).

It is natural to try and write K in terms of the basis for the semigroup T_t , i.e. if we let

$$K(\zeta,\eta,t) = 2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn}\sin m\pi\zeta \sin n\pi\eta$$

and substitute into (4.33) we have

$$2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \dot{a}_{mn}\sin m\pi\zeta \sin n\pi\eta - 2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn}[(m\pi)^2 + (n\pi)^2]\sin m\pi\zeta \sin n\pi\eta$$

$$- \frac{4}{c}\sum_{i=1}^{\infty} a_{in}(i\pi)\sin n\pi\eta \sum_{j=1}^{\infty} a_{mj}(j\pi)\sin m\pi\zeta + 2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \delta_m^n \sin m\pi\zeta \sin n\pi\eta = 0$$

where $\delta_m^n = \begin{cases} 0 \text{ if } m \neq n \\ 1 \text{ if } m = n \end{cases}$.

The coefficients a_{mn} must therefore satisfy

(4.35) $$\dot{a}_{mn} - [(m\pi)^2 + (n\pi)^2]a_{mn} - \frac{2}{c}\sum_{i=1}^{\infty} i\pi a_{in}\sum_{j=1}^{\infty} j\pi a_{mj} + \delta_m^n = 0\ .$$

Remark

It is not shown whether coefficients can be found to satisfy (4.35), i.e. whether the assumptions that Q has the form (4.32) and that we can write K in terms of the basis for the semigroup T_t are justified.

Solutions to the Riccati equation, of this form, have however been shown to exist for bounded control problems. (see Curtain and Pritchard {11}).

Note - This remark applies equally to all the following examples.

5. THE TRACKING PROBLEM.

Here we do not wish to bring the system to the origin but to some preassigned final state, or we wish to follow some preassigned trajectory as closely as possible and thus the performance index considered is

$$(5.1) \qquad J(u;t_o,z_o) = \langle z(t_1) - r(t_1), G[z(t_1) - r(t_1)]\rangle + \int_{t_o}^{t_1} \{\langle z(s) - r(s), M[z(s) - r(s)]\rangle + \langle u(s), Ru(s)\rangle\} ds$$

where $r(t)$ is a given continuous H-valued function on $[t_o,t_1]$, G , M , R as before.

We seek to minimize the performance index (5.1) over all controls $u \in L^2[t_o,t_1;U]$ where $z(t)$ is given by

$$(5.2) \qquad z(t) = T_{t-t_o} z_o + \int_{t_o}^{t} T_{t-s} Bu(s)ds \quad , \quad z(o) = z_o \in H$$

and T_t and B are assumed to satisfy condition (2.4) with either $p \geq 2$ or more generally, $p \geq 1$. Again when considering $p \geq 1$ we must take $G = 0$ in (5.1) and consider a pure integral cost functional.

We will consider the two cases separately.

The case $p \geq 2$.

As in section 4 we construct a sequence of controls u_k in such a way that $u_k \to u^*$ $(k \to \infty)$ and u^* is the unique optimal control.

Set

$$(5.3) \qquad u_k(t) = F_k(t)z(t) - R^{-1}B^*S_{k-1}(t)$$

where $F_k(t)$, $S_{k-1}(t)$ are defined recursively by

$$(5.4) \qquad F_k(t) = - R^{-1}B^*Q_{k-1}(t) \quad , \quad F_o = 0$$

(5.5) $$Q_k(t)x = \mathcal{U}_k^*(t_1,t)G\mathcal{U}_k(t_1,t)x + \int_t^{t_1} \mathcal{U}_k^*(s,t)[M + F_k^*(s)RF_k(s)]\mathcal{U}_k(s,t)x\,ds$$

(5.6) $$S_k(t) = -\mathcal{U}_k^*(t_1,t)Gr(t_1) - \int_t^{t_1} \mathcal{U}_k^*(s,t)[Mr(s) + \{Q_k(s) - Q_{k-1}(s)\} BR^{-1}B^*S_{k-1}(s)]ds$$

$$S_o(t) = 0$$

and where

(5.7) $$\mathcal{U}_k(t,s)x = T_{t-s}x + \int_s^t T_{t-\rho}BF_k(\rho)\mathcal{U}_k(\rho,s)x\,d\rho \quad , \quad x \in H \ .$$

Firstly we must show that the sequence is well defined. From existing results on the regulator problem, given by Curtain and Pritchard {11} for the case $p \geq 2$, we know that:-

(i) $Q_k(t)$ is well defined by (5.5)

(ii) $F_k(t)$ satisfies an estimate of the form

$$||F_k(t)||_{\mathcal{L}(H,U)} \leq f_k(t) \quad \text{where} \quad f_k \in L^2[t_o,t_1]$$

and

(iii) $\mathcal{U}_k(t,s)$ is well defined by (5.7), satisfying estimates of the form

$$||\mathcal{U}_k(t,s)||_{\mathcal{L}(H)} \leq \bar{M}_k$$

$$||\mathcal{U}_k(t,s)||_{\mathcal{L}(\bar{W},H)} \leq g_k(t-s) \ , \quad g_k \in L^2[t_o,t_1] \ .$$

Similarly we can show

<u>Lemma 5.1</u>

$S_k(t)$ satisfies an estimate of the form

$$||S_k(t)||_{\bar{W}^*} \leq l_k(t) \ , \quad l_k \in L^2[t_o,t_1] \ .$$

<u>Proof</u>

$$S_1(t) = - T^*_{t_1-t}Gr(t_1) - \int_t^{t_1} T^*_{s-t}Mr(s)ds \quad \text{therefore}$$

$$||S_1(t)||_{\bar{W}^*} \leq g(t_1-t)\beta + \gamma\int_t^{t_1} g(s-t)ds \quad \text{where} \quad \beta = ||Gr(t_1)|| \quad \text{and}$$

$$\gamma = \operatorname*{ess\,sup}_{s\epsilon[t_o,t_1]} ||Mr(s)|| \quad \text{so} \quad ||S_1(t)||_{\tilde{W}^*} \le l_1(t),\ l_1 \in L^2[t_o,t_1] \ .$$

Now suppose $||S_{k-1}(t)||_{\tilde{W}^*} \le l_{k-1}(t)$ with $l_{k-1} \in L^2[t_o,t_1]$ then

$$||S_k(t)||_{\tilde{W}^*} \le ||\mathcal{U}_k^*(t_1,t)Gr(t_1)||_{\tilde{W}^*} + \int_t^{t_1} ||\mathcal{U}_k^*(s,t)Mr(s)||_{\tilde{W}^*}ds$$

$$+ \int_t^{t_1} ||\mathcal{U}_k^*(s,t)[Q_k(s) - Q_{k-1}(s)]BR^{-1}B^*S_{k-1}(s)||ds$$

$$\le \beta g_k(t_1-t) + \int_t^{t_1} g_k(s-t)ds + \int_t^{t_1} g_k(s-t)f_k(s)||B^*||_{\mathcal{L}(\tilde{W}^*,U)}\, l_{k-1}(s)ds$$

$$+ \int_t^{t_1} g_k(s-t)f_{k-1}(s)||B^*||_{\mathcal{L}(\tilde{W}^*,U)}\, l_{k-1}(s)ds \ .$$

Using the fact that $f_k l_{k-1} \in L^1$ since it is the product of two L^2-functions and so $\int_t^{t_1} g_k(s-t)f_k(s)l_{k-1}(s)ds \in L^2$ as it is the convolution of an L^1 and an L^2 function, (similarly for the other term), we have

$$||S_k(t)||_{\tilde{W}^*} \le l_k(t) \ , \quad l_k \in L^2[t_o,t_1] \ .$$

Also,

__Lemma 5.2__

S_k is continuous in $t \in [t_o,t_1]$ with values in H .

__Proof__

$$||S_k(t)|| \le \bar{M}_k\beta + \bar{M}_k\int_t^{t_1}[\gamma + \{f_{k-1}(s) + f_k(s)\}||B^*||_{\mathcal{L}(\tilde{W}^*,U)}\, l_{k-1}(s)]ds$$

Hence $S_k(t)$ is bounded on H . The continuity follows from the strong continuity of $\mathcal{U}_k^*(\cdot,\cdot)$ since, for $h > 0$

$$||S_k(t+h) - S_k(t)||$$

$$\le ||[\mathcal{U}_k^*(t_1,t+h) - \mathcal{U}_k^*(t_1,t)]Gr(t_1)||$$

$$- \int_{t+h}^{t_1} ||[\mathcal{U}_k^*(s,t+h) - \mathcal{U}_k^*(s,t)][Mr(s) + \{Q_k(s) - Q_{k-1}(s)\}BR^{-1}B^*S_{k-1}(s)]||ds$$

$$+ \int_t^{t+h} ||\mathcal{U}_k^*(s,t)[Mr(s) + \{Q_k(s) - Q_{k-1}(s)\}BR^{-1}B^*S_{k-1}(s)]||ds \ .$$

Hence by the strong continuity of $\mathcal{U}_k^*(\cdot,\cdot)$ and the bounds on G , M , R , $r(t)$, $Q_k(t)$, $S_k(t)$ we have that $||S_k(t+h) - S_k(t)|| \to 0$ as $h \to 0$, i.e. S_k is continuous on the right.

A similar argument holds for continuity on the left.

If we consider now the sequence of control problems

$$(5.8) \qquad z(t) = \mathcal{U}_k(t,t_o)z_o + \int_{t_o}^{t} \mathcal{U}_k(t,s)B[\bar{u}(s) - R^{-1}B^*S_{k-1}(s)]ds \ ,$$

the following lemmas can easily be proved by direct substitution for $z(t)$, $Q_k(t)$ and $S_k(t)$, given by (5.8), (5.5) and (5.6) respectively.

<u>Lemma 5.3</u>

$$(5.9) \qquad \langle z(t), Q_k(t)z(t)\rangle = \langle z(t_1), Gz(t_1)\rangle - 2\int_t^{t_1}\langle z(s), Q_k(s)B[\bar{u}(s) - R^{-1}B^*S_{k-1}(s)]\rangle ds + \int_t^{t_1}\langle z(s), [M + Q_{k-1}(s)BR^{-1}B^*Q_{k-1}(s)]z(s)\rangle ds$$

and

$$(5.10) \qquad \langle z(t), S_k(t)\rangle = - \langle z(t_1), Gr(t_1)\rangle - \int_t^{t_1}\langle [\bar{u}(s) - R^{-1}B^*S_{k-1}(s)], B^*S_k(s)\rangle ds - \int_t^{t_1}\langle z(s), Mr(s) + \{Q_k(s) - Q_{k-1}(s)\}BR^{-1}B^*S_{k-1}(s)\rangle ds$$

which give, adding (5.9) + 2×(5.10) and putting $t = t_o$, $\bar{u} = 0$,

$J(u_k;t_o,z_o)$

$$= \langle z_o, Q_k(t_o)z_o\rangle + \langle r(t_1), Gr(t_1)\rangle + \int_{t_o}^{t_1}\langle r(s), Mr(s)\rangle ds + 2\langle z_o, S_k(t_o)\rangle$$

$$- 2\int_{t_o}^{t_1}\langle B^*S_{k-1}(s), R^{-1}B^*S_k(s)\rangle ds + \int_{t_o}^{t_1}\langle B^*S_{k-1}(s), R^{-1}B^*S_{k-1}(s)\rangle ds$$

and hence, similarly to lemma 4.2, we can prove

<u>Lemma 5.4</u>

Letting $\bar{u}(t) = R^{-1}B^*[\{Q_k(t) - Q_{k-1}(t)\}z(t) + \{S_k(t) - S_{k-1}(t)\}]$, so that $u_{k-1} + \bar{u} = u_k$, we find that

$$J(u_k;t_o,z_o) = J(u_{k-1};t_o,z_o) - \int_{t_o}^{t_1}\langle \bar{u}(s), R\bar{u}(s)\rangle ds$$

and thus

$$J(u_k;t_o,z_o) \leq J(u_{k-1};t_o,z_o) \quad \text{for each } t_o \text{ and } z_o \in H .$$
$$\leq \text{constant}\cdot||z_o||^2$$

Again from existing regulator results in {11} we have the following two lemmas and theorem

<u>Lemma 5.5</u>

$Q_k(t)$ converges strongly as $k \to \infty$ to a self-adjoint non-negative definite bounded linear operator $Q(t)$ on H such that $\sup_{t\in[t_o,t_1]} ||Q(t)|| \leq c$.

<u>Lemma 5.6</u>

$F_k(t)$ and $\mathcal{U}_k(t,s)$ are uniformly bounded in the sense that there exist functions f_∞ and $g_\infty \in L^2[t_o,t_1]$ and a constant $\bar{M}_\infty$ such that

$$||\mathcal{U}_k(t,s)||_{\mathcal{L}(H)} \leq \bar{M}_\infty \quad t_1 \geq t \geq s \geq t_o \quad \forall\, k = 0,1,...$$

$$||F_k(t)||_{\mathcal{L}(H,U)} \leq f_\infty(t) \quad \forall\, k = 0,1,...$$

$$||\mathcal{U}_k(t,s)||_{\mathcal{L}(\tilde{W},H)} \leq g_\infty(t-s) \quad \forall\, k = 0,1,...$$

yielding

Theorem 5.7

$\mathcal{U}_k(t,s)$ converges in H to a mild evolution operator $\mathcal{U}(t,s)$ given by

$$(5.11) \qquad \mathcal{U}(t,s)x = T_{t-s}x - \int_s^t T_{t-\rho}BR^{-1}B^*Q(\rho)\mathcal{U}(\rho,s)x\,d\rho$$

and Q is the unique solution to the integral equation

$$(5.12) \qquad Q(t)x = \mathcal{U}^*(t_1,t)G\mathcal{U}(t_1,t)x + \int_t^{t_1}\mathcal{U}^*(s,t)[M + Q(s)BR^{-1}B^*Q(s)]\mathcal{U}(s,t)x\,ds$$

In order to prove that the limit as $k \to \infty$ of $S_k(t)$ also satisfies an integral equation we need the following lemma.

Lemma 5.8

$S_k(t)$ is uniformly bounded by estimates of the form

$$||S_k(t)||_{\tilde{W}^*} \leq l_\infty \qquad l_\infty \in L^2[t_o,t_1] \qquad \forall\ k = 0,1,\ldots$$

$$||S_k(t)||_H \leq m_\infty \qquad \text{a constant} \qquad \forall\ k = 0,1,\ldots \quad .$$

Proof

From lemma 5.4, $J(u_k;t_o,z_o) \leq J(u_{k-1};t_o,z_o) \leq c||z_o||^2$, therefore

$$\langle z(t_1) - r(t_1),\ G[z(t_1) - r(t_1)]\rangle + \int_{t_o}^{t_1}\langle u_k(s),\ Ru_k(s)\rangle ds + \int_{t_o}^{t_1}\langle z(s) - r(s),\ M[z(s) - r(s)]\rangle ds \leq c||z_o||^2.$$

But, since G , M are non-negative definite and R is strictly positive we have

$$(5.13) \qquad \int_{t_o}^{t_1}||u_k(s)||^2 ds \leq c||z_o||^2 \ .$$

In consideration of the regulator problem Curtain and Pritchard {11} further show that

(5.14) $\quad \int_t^{t_1} ||F_k(s)\mathcal{U}_k(s,t)x||^2 ds \le \text{constant}\cdot||x||^2 \;; \quad x \in H$

and

(5.15) $\quad \int_t^{t_1} ||F_k(s)\mathcal{U}_k(s,t)w||^2 ds \le \text{constant}\cdot d(t)||w||^2_{\tilde{W}} \quad , \quad d \in L^1[t_o,t_1] \;.$

From (5.8) with $\bar{u} = 0$, $z(t) = \mathcal{U}_k(t,t_o)z_o - \int_{t_o}^t \mathcal{U}_k(t,s)BR^{-1}B^*S_{k-1}(s)ds$

thus

$$R^{-1}B^*S_{k-1}(t) = -u_k(t) + F_k(t)z_k(t)$$

$$= -u_k(t) + F_k(t)\mathcal{U}_k(t,t_o)z_o - \int_{t_o}^t F_k(t)\mathcal{U}_k(t,s)BR^{-1}B^*S_{k-1}(s)ds \;.$$

So letting $||R^{-1}B^*S_{k-1}(t)|| = P(t)$,

$$P(t) \le \{||u_k(t)|| + ||F_k(t)\mathcal{U}_k(t,t_o)z_o||\} + \int_{t_o}^t ||F_k(t)\mathcal{U}_k(t,s)BR^{-1}B^*S_{k-1}(s)||ds \,.$$

Here the first term can be bounded by an L^2 function from (5.13) and (5.14). In the second term we have $f_\infty g_\infty \in L^1[t_o,t_1]$ as it is the product of two L^2 functions, so by the Generalised Gronwall's Inequality $P \in L^2[t_o,t_1]$, independent of k . This gives

$$||S_k(t)||_{\tilde{W}^*} \le g_\infty(t_1-t) + \int_t^{t_1} g_\infty(s-t)ds + \int_t^{t_1} 2g_\infty(s-t)f_\infty(s)||R||P(s)ds$$

$$\le l_\infty(t) \qquad l_\infty \in L^2[t_o,t_1] \;,$$

since $f_\infty P \in L^1[t_o,t_1]$, as it is the product of two L^2 functions and then $\int_t^{t_1} g_\infty(s-t)f_\infty(s)P(s)ds$ is the convolution of an L^1 function and an L^2 function, so is L^2. Also

$$||S_k(t)||_H \le ||\mathcal{U}_k^*(t_1,t)Gr(t_1)|| + \int_t^{t_1} ||\mathcal{U}_k^*(s,t)Mr(s)||ds$$

$$+ \int_t^{t_1} ||\mathcal{U}_k^*(s,t)[Q_k(s) - Q_{k-1}(s)]BR^{-1}B^*S_{k-1}(s)||ds$$

$$\leq \overline{M}_\infty \beta + \int_t^{t_1} \overline{M}_\infty \gamma ds + 2\int_t^{t_1} \overline{M}_\infty f_\infty(s)||R||P(s)ds$$

$\leq$ constant independent of k .

The above results yield the following theorem.

Theorem 5.9

If we define

$$(5.16) \quad S(t) = -\,U^*(t_1,t)Gr(t_1) - \int_t^{t_1} U^*(s,t)Mr(s)ds$$

then $S_k(t)$ converges strongly to $S(t)$ as $k \to \infty$.

Proof

$$||S_k(t) - S(t)|| \leq ||[U^*(t_1,t) - U_k^*(t_1,t)]Gr(t_1)||$$
$$+ \int_t^{t_1} ||[U^*(s,t) - U_k^*(s,t)]Mr(s)||ds$$
$$+ \int_t^{t_1} ||U_k^*(s,t)[Q_k(s) - Q_{k-1}(s)]BR^{-1}B^*S_{k-1}(s)||ds \quad .$$

Since all the terms are uniformly bounded in k , using the strong convergence of $U_k(\cdot,\cdot)$ to $U(\cdot,\cdot)$ and $Q_k(\cdot)$ to $Q(\cdot)$ we have by the Lebesgue Dominated Convergence Theorem that $||S_k(t) - S(t)|| \to 0$ as $k \to \infty$.

Thus we have,

Theorem 5.10

The optimal control which minimizes $J(u;t_o,z_o)$ is the control

$$u^* = -\,R^{-1}B^*[Q(t)z(t) + S(t)]$$

which is a combination of feedback and open-loop control.

Proof

Consider any arbitrary admissible control $u \in L^2[t_o,t_1;U]$ then letting $\overline{u}(t) = u^*(t) - u(t)$ in lemma 5.3 gives

$$J(u^*;t_o,z_o) = J(u;t_o,z_o) - \int_{t_o}^{t_1} \langle \bar{u}(s), R\bar{u}(s)\rangle ds$$

i.e. u^* is optimal.

Letting $k \to \infty$ in lemma 5.3 we find that the cost of this optimal control is given by

$$J(u^*;t_o,z_o) = \langle z_o, Q(t_o)z_o\rangle + \langle r(t_1), Gr(t_1)\rangle + \int_{t_o}^{t_1} \langle r(s), Mr(s)\rangle ds$$

$$+ 2\langle z_o, S(t_o)\rangle - \int_{t_o}^{t_1} \langle B^*S(s), R^{-1}B^*S(s)\rangle ds \quad .$$

In {11} Curtain and Pritchard show that it is possible to differentiate the integral Riccati equation (5.12) and that $Q(t)$ is the unique solution to (4.25) with $Q(t_1) = G$. Further, assuming the existence of a Green's formula (2.7), $Q(t)$ satisfies (4.26) with $Q(t_1) = G$, uniquely if $\bar{\mathcal{D}} = H$. Similarly we may obtain two differential equations for $S(t)$.

<u>Theorem 5.11</u>

For $x \in D(A)$, $S(t)$ is the unique solution to

(5.17) $$\frac{d}{dt}\langle x, S(t)\rangle = -\langle Ax, S(t)\rangle + \langle x, Mr(t)\rangle + \langle R^{-1}B^*Q(t)x, B^*S(t)\rangle$$

a.e. on $[t_o,t_1]$

$$S(t_1) = Gr(t_1)$$

or for $x \in \mathcal{D}$, $S(t)$ is a solution to the differential equation

(5.18) $$\frac{d}{dt}\langle x, S(t)\rangle = \langle x, Mr(t)\rangle - \langle \tilde{A}x, S(t)\rangle$$

$$S(t_1) = Gr(t_1) \;,$$

uniquely if $\bar{\mathcal{D}} = H$.

<u>Proof</u>

We will prove uniqueness only, since the proof that $S(t)$ satisfies (5.17) and (5.18) is similar to that for the equations in $Q(t)$, (4.25) and (4.26), given for the case $p \geq 1$ in theorem 4.5.

Let $S_1(t)$ and $S_2(t)$ be two solutions of (5.17) then $P(t) = S_1(t) - S_2(t)$ satisfies

$$\frac{d}{dt} \langle P(t), x\rangle = - \langle P(t), Ax\rangle + \langle B^*P(t), R^{-1}B^*Q(t)\rangle$$

$$P(t_1) = 0$$

Now let $\Psi(t) = \langle T_{t-s}x, P(t)\rangle$ and differentiate giving

$$\dot{\Psi}(t) = \langle AT_{t-s}x, P(t)\rangle - \langle AT_{t-s}x, P(t)\rangle + \langle R^{-1}B^*Q(t)T_{t-s}x, B^*P(t)\rangle$$

$$= \langle R^{-1}B^*Q(t)T_{t-s}x, B^*P(t)\rangle \quad .$$

Then integrating

$$\Psi(t) = \langle T_{t-s}x, P(t)\rangle = -\int_t^{t_1} \langle R^{-1}B^*Q(\rho)T_{\rho-s}x, B^*P(\rho)\rangle d\rho$$

and so putting $x = P(t)$ and letting $s \to t$

$$||P(t)|| \leq \int_t^{t_1} f_\infty(t)\bar{M}_\infty ||B||_{\tilde{\mathcal{L}}(U,W)} ||P(\rho)|| d\rho \ .$$

Hence, by the Generalised Gronwall's Inequality, $P(t) = 0$, i.e. $S_1(t) = S_2(t)$.

Similarly we may prove uniqueness of solutions to (5.18) when $\bar{\mathcal{D}} = H$.

<u>The case $p \geq 1$</u> .

As previously stated, in this case we must take $G = 0$ in (5.1) and consider a pure integral cost functional. We can then show that the optimal control is $u^*(t) = - R^{-1}B^*[Q(t)z(t) + S(t)]$ where $Q(t)$ and $S(t)$ are the unique solutions to the integral equations

$$(5.19) \qquad Q(t)x = \int_t^{t_1} \mathcal{U}^*(s,t)[M + Q(s)BR^{-1}B^*Q(s)]\mathcal{U}(s,t)x ds$$

and

(5.20) $S(t) = -\int_t^{t_1} \mathcal{U}^*(s,t)Mr(s)ds$

respectively, with $\mathcal{U}(s,t)$ satisfying (5.11). Since z is not in general continuous but belongs to $L^2[t_o,t_1;H]$ it is natural to take r also in this class.

The proof mirrors that for the case $p \geq 2$ with $G = 0$, using the results for the quadratic cost regulator problem given in section 4, the only difference being in the estimates obtained.

In section 4 we saw that $F_k(t)$ satisfies an estimate of the form

$$||F_k(t)||_{\mathcal{L}(H,U)} \leq f_k \text{ , a constant}$$

and that $\mathcal{U}_k(t,s)$ satisfies estimates of the form

$$||\mathcal{U}_k(t,s)||_{\mathcal{L}(H)} \leq \bar{M}_k \qquad \text{a constant}$$

$$||\mathcal{U}_k(t,s)||_{\mathcal{L}(\tilde{W},H)} \leq g_k(t-s) \qquad g_k \in L^1[t_o,t_1] \text{ .}$$

Hence, analagous to lemmas 5.1 and 5.2 we have ,

<u>Lemma 5.12</u>

$S_k(t)$ satisfies an estimate of the form

$$||S_k(t)||_{\tilde{W}^*} \leq l_k \qquad l_k \in L^2[t_o,t_1]$$

$$||S_k(t)||_H \leq m_k \qquad \text{a constant .}$$

<u>Proof</u>

$S_1(t) = -\int_t^{t_1} T^*_{s-t}Mr(s)ds$ therefore

$$||S_1(t)||_{\tilde{W}^*} \leq ||M||\int_t^{t_1} g(s-t)||r(s)||ds \leq l_1(t)$$

and $l_1 \in L^2[t_o,t_1]$ since the convolution of an L^1 function and an L^2 function is L^2 . Now suppose $||S_{k-1}(t)||_{\tilde{W}^*} \leq l_{k-1}(t)$ with $l_{k-1} \in L^2[t_o,t_1]$, then

$$||S_k(t)||_{\bar{W}^*}$$

$$\leq \int_t^{t_1} ||\mathcal{U}_k^*(s,t)Mr(s)||_{W^*}ds + \int_t^{t_1} ||\mathcal{U}_k^*(s,t)[Q_k(s) - Q_{k-1}(s)]BR^{-1}B^*S_{k-1}(s)||_{\bar{W}^*}ds$$

$$\leq ||M|| \int_t^{t_1} g_k(s-t)||r(s)||ds + \int_t^{t_1} g_k(s-t)[f_k + f_{k-1}]||B^*||_{\mathcal{L}(\bar{W}^*,U)} l_{k-1}(s)ds$$

$$\leq l_k(t) \qquad l_k \in L^2[t_o,t_1]$$

using again the same convolution result. Also,

$$||S_k(t)||_H$$

$$\leq \int_t^{t_1} ||\mathcal{U}_k^*(s,t)Mr(s)||ds + \int_t^{t_1} ||\mathcal{U}_k^*(s,t)[Q_k(s) - Q_k(s)]BR^{-1}B^*S_{k-1}||ds$$

$$\leq ||M||\bar{M}_k \int_t^{t_1} ||r(s)||ds + ||B^*||_{\mathcal{L}(\bar{W}^*,U)} \bar{M}_k[f_k + f_{k-1}] \int_t^{t_1} l_{k-1}(s)ds$$

$$\leq m_k \qquad \text{a constant.}$$

We also have from section 4 that the estimates on F_k and $\mathcal{U}_k$ are uniformly bounded in that there exists a function $g_\infty \in L^1[t_o,t_1]$ and constants $\bar{M}_\infty$, f_∞ such that

$$||\mathcal{U}_k(t,s)||_{\mathcal{L}(H)} \leq \bar{M}_\infty \qquad t_1 \geq t \geq s \geq t_o \qquad \forall\ k = 0,1,\ldots$$

$$||F_k(t)||_{\mathcal{L}(H,U)} \leq f_\infty \qquad \forall\ k = 0,1,\ldots$$

$$||\mathcal{U}_k(t,s)||_{\mathcal{L}(\bar{W},H)} \leq g_\infty(t-s) \qquad \forall\ k = 0,1,\ldots$$

Similarly for S_k we have,

<u>Lemma 5.13</u>

$S_k(t)$ is uniformly bounded in k by estimates of the form

$$||S_k(t)||_{\bar{W}^*} \leq l_\infty(t) \qquad l_\infty \in L^2[t_o,t_1] \qquad \forall\ k = 0,1,\ldots$$

$$||S_k(t)||_H \leq m_\infty \qquad \text{a constant} \qquad \forall\ k = 0,1,\ldots$$

Proof

Consider $F_k(t)z(t) = F_k(t)\mathcal{U}_k(t,t_o)z_o - \int_{t_o}^{t} F_k(t)\mathcal{U}_k(t,s)BR^{-1}B^*S_{k-1}(s)ds$,

then since

$$R^{-1}B^*S_{k-1}(t) = -u_k(t) + F_k(t)z(t)$$

$$= -u_k(t) + F_k(t)\mathcal{U}_k(t,t_o)z_o - \int_{t_o}^{t} F_k(t)\mathcal{U}_k(t,s)BR^{-1}B^*S_{k-1}(s)ds ,$$

letting $||R^{-1}B^*S_{k-1}(t)|| = P(t)$,

$$||P(t)|| \leq \{||u_k(t)|| + ||F_k(t)\mathcal{U}_k(t,t_o)z_o||\} + \int_{t_o}^{t} ||F_k(t)\mathcal{U}_k(t,s)BR^{-1}B^*S_{k-1}(s)||ds$$

$$\leq v(t) + \int_{t_o}^{t} f_\infty g_\infty(t-s)||B||_{\tilde{\mathcal{L}}(U,W)} P(s)ds$$

where $v \in L^2[t_o,t_1]$ as in lemma 5.8. Thus by the Generalised Gronwall's Inequality $P \in L^2[t_o,t_1]$, independent of k . Therefore

$$||S_k(t)||_{\tilde{W}^*}$$

$$\leq \int_t^{t_1} ||\mathcal{U}_k^*(s,t)Mr(s)||_{\tilde{W}^*} + \int_t^{t_1} ||\mathcal{U}_k^*(s,t)\{Q_k(s) - Q_{k-1}(s)\}BR^{-1}B^*S_{k-1}(s)||_{\tilde{W}^*}ds$$

$$\leq ||M||\int_t^{t_1} g_\infty(s-t)||r(s)||ds + 2\int_t^{t_1} g_\infty(s-t)f_\infty||R||P(s)ds$$

$$\leq l_\infty(t) \qquad l_\infty \in L^2[t_o,t_1]$$

and

$$||S_k(t)|| \leq ||M||\int_t^{t_1} \bar{M}_\infty||r(s)||ds + 2\int_t^{t_1} \bar{M}_\infty f_\infty||R||P(s)ds$$

$$\leq m_\infty \quad \text{a constant independent of } k .$$

With these new estimates and $G = 0$ in the cost functional all the remaining results for the case $p \geq 2$ remain valid.

<u>**Example 5.14**</u>

Consider again the case of the controlled heat equation

(5.21) $$z_t = z_{xx}$$
$$z(o,t) = u\ ,\quad z(1,t) = 0\ ,\quad z(x,o) = z_o$$

and suppose we wish to minimize a cost functional of the form

(5.22) $$J(u) = \int_o^{t_1}\int_o^1 \{z(x,t) - r(x,t)\}^2 dxdt + c\int_o^{t_1} u^2(t)dt\ .$$

It has been shown in the previous examples, example 2.7 and example 4.9 that by taking $H = L^2(0,1)$, $U = R$ and $A = \frac{\partial^2}{\partial x^2}$ on $D(A) = H^2(0,1) \cap H_o^1(0,1)$ the conditions of section 2 are satisfied with $4/3 > p \geq 1$, and where $B = -\delta'$. Thus since (5.22) is of the form (5.1) with $G = 0$, $M = I$ and $R = c$ we know there exists a unique optimal control $u^* = -R^{-1}[B^*Qz + B^*S]$ with Q satisfying (4.25) and S satisfying (5.17) with $S(t_1) = 0$.

From example 4.9 we know that if we write $(Qz)(\zeta) = \int_o^1 K(\zeta,\eta,t)z(\eta)d\eta$ then K satisfies (4.34). Similarly by writing $S(t)(\zeta) = S(\zeta,t)$ and substituting for this S into (5.17), since,

$$\frac{d}{dt}\langle h,\ S(t)\rangle = \int_o^1 S_t(\zeta,t)h(\zeta)d\zeta$$

$$\langle Ah,\ S(t)\rangle = \int_o^1 S_{\zeta\zeta}(\zeta,t)h(\zeta)d\zeta \qquad \text{when } S(1,t) = 0 = S(o,t)$$

$$\langle h,\ Mr\rangle = \langle h,\ r\rangle = \int_o^1 h(\zeta)r(\zeta)d$$

and

$$\langle u(t),\ B^*S(t)\rangle = \langle Bu(t),\ S(t)\rangle = \int_o^1 (Bu)(\zeta)S(\zeta,t)d\zeta = -u\int_o^1 \delta'(\zeta)S(\zeta,t)d\zeta$$
$$= uS_\zeta(o,t)$$

so that

$$\langle R^{-1}B^*Q(t)h,\ B^*S(t)\rangle = \frac{1}{c}\int_0^1 K_\eta(\zeta,o,t)h(\zeta)S_\zeta(o,t)d\zeta$$

we find that

$$\int_0^1 \{S_t(\zeta,t) + S_{\zeta\zeta}(\zeta,t) - \frac{1}{c}K_\eta(\zeta,o,t)S_\zeta(o,t) - r(\zeta)\}h(\zeta)d\zeta = 0 ,$$

which is satisfied if S satisfies

$$(5.23) \quad S_t(\zeta,t) + S_{\zeta\zeta}(\zeta,t) - \frac{1}{c}K_\eta(\zeta,o,t)S_\zeta(o,t) - r(\zeta) = 0 .$$

Thus the optimal control is

$$u^*(t) = - R^{-1}[B^*Q(t)z(t) + B^*S(t)] = -\frac{1}{c}\{\int_0^1 K_\zeta(o,\eta,t)z(\eta)d\eta + S_\zeta(o,t)\}$$

where K satisfies (4.34) and S satisfies (5.23).

In example 4.9 we saw that by writing K in terms of the basis $\{\sqrt{2}\sin n\pi x\}$ for the semigroup T_t, i.e.

$$K(\zeta,\eta,t) = 2\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} a_{mn}\sin m\pi\zeta\sin n\pi\eta$$

that the coefficients a_{mn} satisfy the differential equation (4.35). Similarly we may write

$$S(\zeta,t) = \sqrt{2}\sum_{m=0}^{\infty} b_m(t)\sin m\pi\zeta$$

and then from (5.23) the coefficients b_m satisfy

$$\dot{b}_m - (m^2\pi^2)b_m - \frac{2}{c}\sum_{i=0}^{\infty} i\pi a_{mi}\sum_{j=0}^{\infty} j\pi b_j - r_m = 0$$

where r_m are the coefficients such that $r = \sqrt{2}\sum_{m=0}^{\infty} r_m\sin m\pi\zeta$.

6. THE INFINITE TIME QUADRATIC COST CONTROL PROBLEM

We now consider the control system

(6.1) $$z(t) = T_{t-t_o} z_o + \int_{t_o}^{t} T_{t-s} Bu(s)ds \ , \quad z(t_o) = z_o \ ,$$

over the infinite interval, with the cost functional

(6.2) $$J(u;t_o,z_o) = \int_{t_o}^{\infty} \{\langle z(s), Mz(s)\rangle + \langle u(s), Ru(s)\rangle\}ds$$

where $z_o \in H$, M , R are as before, $u \in L^2[t_o,\infty;U]$ and T_t and B are assumed to satisfy condition (2.4) with $p \geq 1$.

Firstly consider the problem over the finite interval $[t_o,t_n]$ $t_o \geq 0$, denoting the performance index by

(6.3) $$J^n(u;t_o,z_o) = \int_{t_o}^{t_n} \{\langle z(s), Mz(s)\rangle + \langle u(s), Ru(s)\rangle\}ds$$

with $z(t)$ given by (6.1).

We know from section 4 that the optimal control is feedback and given by

(6.4) $$u^n(t) = -R^{-1}B^*Q^n(t)z(t)$$

where $Q^n(t)$ is the unique solution of

(6.5) $$Q^n(t)x = \int_t^{t_n} \mathcal{U}^{n*}(s,t)[M + Q^n(s)BR^{-1}B^*Q^n(s)]\mathcal{U}^n(s,t)xds$$

and $\mathcal{U}^n(s,t)$ is the mild evolution operator given by

(6.6) $$\mathcal{U}^n(t,s)x = T_{t-s}x - \int_s^t T_{t-\rho}BR^{-1}B^*Q^n(\rho)\mathcal{U}^n(\rho,s)xd\rho \ .$$

__Lemma 6.1__

$Q^n(t)$ is also given by

(6.7) $$Q^n(t)x = \int_t^{t_n} T^*_{s-t}[M - Q^n(s)BR^{-1}B^*Q^n(s)]T_{s-t}xds$$

<u>Proof</u>

Firstly we can show that $Q^n(t)$ is also given by

$$(6.8)\qquad Q^n(t)x = \int_t^{t_n} \mathcal{U}^{n^*}(s,t)MT_{s-t}x\,ds \quad ,$$

by substituting for $\mathcal{U}^n(t,s)$ from (6.6) into the first term on the right hand side of (6.5) and employing the Generalised Gronwall's Inequality. Substitution into the right hand side of (6.8) for $\mathcal{U}^{n^*}(t,s)$ given by

$$(6.9)\qquad \mathcal{U}^{n^*}(s,t)x = T^*_{s-t}x - \int_t^s T^*_{s-\rho}Q^n(\rho)BR^{-1}B^*\mathcal{U}^n(\rho,t)x\,d\rho$$

yields the result.

We also know that $J^n(u^n;t_o,z_o)$, i.e. the cost of the optimal control $u^n(t)$ over the interval $[t_o,t_n]$, is given by

$$(6.10)\qquad J^n(u^n;t_o,z_o) = \langle z_o,\, Q^n(t_o)z_o\rangle \ .$$

$J^n(u^n;t_o,z_o)$ is increasing in n for any $t_o \geq 0$ since the cost of the optimal control over the interval $[t_o,t_n]$ must be greater than, or equal to, the cost of the optimal control over the interval $[t_o,t_k]$ if $t_n > t_k$. Thus $Q^n(t)$ is increasing in n , for all t .

In order to show that the infinite time problem is well defined we need to make the following assumption.

<u>Optimizability Assumption</u>

There exists a control $u \in L^2[t_o,\infty;U]$ such that the performance index

$$J(u;t_o,z_o) = \int_o^\infty \{\langle z(s),\, Mz(s)\rangle + \langle u(s),\, Ru(s)\rangle\}ds$$

is finite.

Suppose $\bar{u}(t)$ is one such control, then

$$J(\bar{u};t_o,z_o) = \int_{t_o}^{\infty} \{<z(s),\ Mz(s)> + <\bar{u}(s),\ R\bar{u}(s)>\}ds$$

$$= \int_{t_o}^{t_n}\{<z(s),\ Mz(s)> + <\bar{u}(s),\ R\bar{u}(s)>\}ds + \int_{t_n}^{\infty} \{<z(s),\ Mz(s)> + <\bar{u}(s),\ R\bar{u}(s)>\}ds$$

$$= J^n(\bar{u};t_o,z_o) + \int_{t_n}^{\infty} \{<z(s),\ Mz(s)> + <\bar{u}(s),\ R\bar{u}(s)>\}ds \ .$$

But $J^n(u^n;t_o,z_o)$ is the minimum cost over $[t_o,t_n]$, hence

$$J(\bar{u};t_o,z_o) \geq J^n(u^n;t_o,z_o) + \int_{t_n}^{\infty} \{<z(s),\ Mz(s)> + <\bar{u}(s),\ R\bar{u}(s)>\}ds \ .$$

Since $\bar{u}(t)$ satisfies the optimizability assumption we have $J(\bar{u};t_o,z_o) < \infty$. Thus

(6.12) $\quad Q^{\infty}(t) = \lim_{n\to\infty} Q^n(t)$ exists

i.e. $\lim_{n\to\infty} \int_t^{t_n} T^*_{s-t}[M - Q^n(s)BR^{-1}B^*Q^n(s)]T_{s-t}x\,ds$ is well defined.

<u>**LEMMA 6.2**</u>

$Q^{\infty}(t) = Q^{\infty}$ independent of t .

<u>**Proof**</u>

$$Q^n(t+\alpha)x = \int_{t+\alpha}^{t_n} T^*_{s-t-\alpha}[M - Q^n(s)BR^{-1}B^*Q^n(s)]T_{s-t-\alpha}x\,ds$$

$$= \int_t^{t_n-\alpha} T^*_{\rho-t}[M - Q^n(\rho+\alpha)BR^{-1}B^*Q^n(\rho+\alpha)]T_{\rho-t}x\,d\rho$$

i.e. $Q^n(t+\alpha)$ is the solution of

(6.13) $\quad P(t)x = \int_t^{t_n-\alpha} T^*_{\rho-t}[M - P(\rho)BR^{-1}B^*P(\rho)]T_{\rho-t}x\,d\rho$.

But we know (6.13) has unique solution $Q^{n-\alpha}(t)$, and hence

$Q^n(t+\alpha) = Q^{n-\alpha}(t)$. Letting $n \to \infty$ we find $Q^\infty(t+\alpha)x = Q^\infty(t)x$ and hence Q^∞ is independent of t .

Theorem 6.3

$$Q^\infty x = \lim_{n\to\infty} Q^n(t)x = \lim_{n\to\infty} \int_t^{t_n} \mathcal{U}^{n*}(s,t)[M + Q^n(s)BR^{-1}B^*Q^n(s)]\mathcal{U}^n(s,t)xds$$

and so

$$(6.14) \qquad Q^\infty x = \int_t^\infty T^{\infty *}_{s-t}[M + Q^\infty BR^{-1}B^*Q^\infty]T^\infty_{s-t}xds$$

Proof

Define

$$(6.15) \qquad \bar{Q}x = \int_{t_o}^\infty T^{\infty *}_{s-t_o}[M + Q^\infty BR^{-1}B^*Q^\infty]T^\infty_{s-t_o}xds$$

where T^∞_t is the perturbation of T_t by $-BR^{-1}B^*Q^\infty$

i.e. $\quad T^\infty_{t-s}x = T_{t-s}x - \int_s^t T_{t-\rho}BR^{-1}B^*Q^\infty T^\infty_{\rho-s}xd\rho$.

We have that $\mathcal{U}^n(t,s) \to T^\infty_{t-s}$ as $n \to \infty$ where $\mathcal{U}^n(t,s)$ is the perturbation of T_{t-s} by $-BR^{-1}B^*Q^n(t)$ since

$||\mathcal{U}^n(t,s)x - T^\infty_{t-s}x||$

$$= ||\int_s^t T_{t-\rho}BR^{-1}B^*Q^n(\rho)\mathcal{U}^n(\rho,s)xd\rho - \int_s^t T_{t-\rho}BR^{-1}B^*Q^\infty T^\infty_{\rho-s}xd\rho||$$

$$\leq \int_s^t ||T_{t-\rho}BR^{-1}B^*Q^n(\rho)[\mathcal{U}^n(\rho,s)x - T^\infty_{\rho-s}x]||d\rho + \int_s^t ||T_{t-\rho}BR^{-1}B^*[Q^n(\rho) - Q^\infty]T^\infty_{\rho-s}x||d\rho .$$

The second term tends to 0 as $n \to \infty$ by Lebesgue Dominated Convergence and so we may use the Generalised Gronwall's Inequality to show that

$||\mathcal{U}^n(t,s)x - T^\infty_{t-s}x|| \to 0$ as $n \to \infty$.

In order to complete the proof of theorem 6.3 we further require the following lemma.

Lemma 6.4

$$\lim_{n\to\infty} \langle x, \mathcal{U}^{n^*}(s,t)[M + Q^n(s)BR^{-1}B^*Q^n(s)]\mathcal{U}^n(s,t)x\rangle = \langle x, T^{\infty^*}_{s-t}[M + Q^\infty BR^{-1}B^*Q^\infty]T^\infty_{s-t}x\rangle$$

Proof

$$\langle x, \mathcal{U}^{n^*}(s,t)[M + Q^n(s)BR^{-1}B^*Q^n(s)]\mathcal{U}^n(s,t)x\rangle - \langle x, T^{\infty^*}_{s-t}[M + Q^\infty BR^{-1}B^*Q^\infty]T^\infty_{s-t}x\rangle$$

$$= \langle\{\mathcal{U}^n(s,t) - T^\infty_{s-t}\}x, \ [M + Q^n(s)BR^{-1}B^*Q^n(s)]\mathcal{U}^n(s,t)x\rangle$$

$$+ \langle T^\infty_{s-t}x, \ [M + \{Q^n(s) - Q^\infty\}BR^{-1}B^*Q^n(s)]\mathcal{U}^n(s,t)x\rangle$$

$$+ \langle T^\infty_{s-t}x, \ [M + Q^\infty BR^{-1}B^*\{Q^n(s) - Q^\infty\}]\mathcal{U}^n(s,t)x\rangle$$

$$+ \langle T^\infty_{s-t}x, \ [M + Q^\infty BR^{-1}B^*Q^\infty]\{\mathcal{U}^n(s,t) - T^\infty_{s-t}\}x\rangle \ .$$

Since $Q^n(t)$ is an increasing sequence bounded above (from the optimizability assumption) it is uniformly bounded in n and so using the Generalised Gronwall's Inequality we have a uniform bound on $\mathcal{U}^n(s,t)$. We may thus employ the Lebesgue Dominated Convergence Theorem, using the strong convergence of $\mathcal{U}^n(t,s) \to T^\infty_{t-s}$ and of $Q^n(t)$ to Q^∞, to show the right hand side tends to 0 as $n \to \infty$ and hence so too does the left hand side. Thus the lemma is proved.

Using lemma 6.4 we have that, for $t_1 < t_n$

(6.16)
$$\int_{t_o}^{t_1} \langle T^\infty_{s-t_o}z_o, \ [M + Q^\infty BR^{-1}B^*Q^\infty]T^\infty_{s-t_o}z_o\rangle ds$$

$$= \lim_{n\to\infty} \int_{t_o}^{t_1} \langle \mathcal{U}^n(s,t_o)z_o, \ [M + Q^n(s)BR^{-1}B^*Q^n(s)]\mathcal{U}^n(s,t_o)z_o\rangle ds$$

$$\leq \lim_{n\to\infty} \int_{t_o}^{t_n} \langle \mathcal{U}^n(s,t_o)z_o, \ [M + Q^n(s)BR^{-1}B^*Q^n(s)]\mathcal{U}^n(s,t_o)z_o\rangle ds$$

$$= \lim_{n\to\infty} J^n(u^n;t_o,z_o)$$

$$= \langle z_o,\ Q^\infty z_o\rangle \quad \text{by (6.12)} .$$

Hence (6.16) exists for all t_1 . Now taking the limit as $t_1 \to \infty$,

$$\langle z_o,\ \bar{Q}z_o\rangle = \int_{t_o}^{\infty} \langle T^\infty_{s-t_o} z_o,\ [M + Q^\infty BR^{-1}B^*Q^\infty]T^\infty_{s-t_o} z_o\rangle ds$$

$$= \lim_{n\to\infty} \int_{t_o}^{t_n} \langle T^\infty_{s-t_o} z_o,\ [M + Q^\infty BR^{-1}B^*Q^\infty]T^\infty_{s-t_o} z_o\rangle ds$$

so, from the above

(6.17) $\quad \langle z_o,\ \bar{Q}z_o\rangle \le \langle z_o,\ Q^\infty z_o\rangle$.

But

$$\langle z_o,\ Q^\infty z_o\rangle = \lim_{n\to\infty} \langle z_o,\ Q^n(t)z_o\rangle$$

$$= \lim_{n\to\infty} J^n(u^n;t_o,z_o)$$

$$\le \lim_{n\to\infty} J^n(u^\infty;t_o,z_o) \qquad \text{since } u^n \text{ is optimal on } [t_o,t_n]$$

$$= \lim_{n\to\infty} \int_{t_o}^{t_n} \langle T^\infty_{s-t_o} z_o,\ [M + Q^\infty BR^{-1}B^*Q^\infty]T^\infty_{s-t_o} z_o\rangle ds$$

$$= \int_{t_o}^{\infty} \langle T_{s-t_o} z_o,\ [M + Q^\infty BR^{-1}B^*Q^\infty]T^\infty_{s-t_o} z_o\rangle ds$$

$$= \langle z_o,\ \bar{Q}z_o\rangle$$

i.e.

(6.18) $\quad \langle z_o,\ Q^\infty z_o\rangle \le \langle z_o,\ \bar{Q}z_o\rangle$.

Combining (6.17) and (6.18) we have that $Q^\infty = \bar{Q}$ and hence the result (6.14) holds.

We have therefore shown that the infinite time problem (6.1) - (6.2) is well defined and that the optimal control is the feedback control

$u^*(t) = -R^{-1}B^*Q^\infty z(t)$ where Q^∞ is a solution to the integral equation

$$(6.19) \qquad Q^\infty x = \int_{t_o}^{\infty} T^\infty_{s-t_o}[M + Q^\infty BR^{-1}B^*Q^\infty]T^\infty_{s-t_o} x ds$$

with

$$(6.20) \qquad T^\infty_{t-s}x = T_{t-s}x - \int_s^t T_{t-\rho}BR^{-1}B^*Q^\infty T^\infty_{\rho-s} x\, d\rho \quad .$$

Furthermore we can show, as for the finite time regulator that Q^∞ is the unique solution to (6.19) in the class of time-independent self-afjoint bounded linear operators on H .

As for the finite time problem we can differentiate the integral Riccati equation (6.19) yielding

Theorem 6.5

Q^∞ is a solution to the algebraic Riccati equation

$$(6.21) \qquad \langle Q^\infty x, Ay\rangle + \langle Ax, Q^\infty y\rangle + \langle x, My\rangle = \langle Q^\infty BR^{-1}B^*Q^\infty x, y\rangle \quad \text{for } x, y \in D(A) .$$

Alternatively, assuming the existence of a Green's formula (2.7) with $B^* = C$, Q^∞ is a solution of

$$(6.22) \qquad \langle Q^\infty x, \bar{A}y\rangle + \langle \bar{A}x, Q^\infty y\rangle + \langle x, My\rangle + \langle Q^\infty BR^{-1}B^*Q^\infty x, y\rangle = 0$$
$$\text{for } x, y \in \mathcal{D} = D(\bar{A}) \cap \ker\{D - R^{-1}B^*Q^\infty\} .$$

We know that $\bar{A}$ is the quasi-generator of T^∞_t on $\mathcal{D}$, i.e.

$$\frac{d}{dt} T^\infty_t y = T^\infty_t \bar{A}y \qquad \text{for } y \in \mathcal{D}, \quad t > 0 \quad ,$$

but we do not have in general that $\bar{A}$ is the infinitesimal generator of T^∞_t . Suppose now that T^∞_t has infinitesimal generator A^Q , then for $x \in D(A^Q)$

$$(6.23) \qquad \frac{d}{dt} T^\infty_t x = T^\infty_t A^Q x = A^Q T^\infty_t x , \qquad t > 0 \quad ,$$

so we have that A^Q is a closed extension of $\bar{A}$, and on $\mathcal{D}$, $A^Q = \bar{A}$.

Hence, since Q^∞ satisfies (6.22) on $\mathcal{D}$, Q^∞ also satisfies

(6.24) $\langle Q^\infty x, A^Q y\rangle + \langle A^Q x, Q^\infty y\rangle + \langle x, My\rangle + \langle Q^\infty BR^{-1}B^* Q^\infty x, y\rangle = 0$

for $x, y \in \mathcal{D}$ which can be extended to hold for all x , $y \in D(A^Q)$ if $\overline{\mathcal{D}} = H$.

With the aid of this equivalent form of the algebraic Riccati equation (6.24) we can examine uniqueness of solutions to (6.22) when $\overline{\mathcal{D}} = H$, and show that if either a) the pair $\{A^*, M^{\frac{1}{2}}\}$ is stabilizable or b) $M > 0$ then Q^∞ is the unique solution to (6.22).

Definition 6.6

The pair $\{A^*, M^{\frac{1}{2}}\}$ is stabilizable, where A^* is the infinitesimal generator of a strongly continuous semigroup T_t^* and $M^{\frac{1}{2}} \in \mathcal{L}(H)$, if there exists an operator $S \in \mathcal{L}(H)$ such that the semigroup given by

(6.25) $T^S_{t-s}x = T^*_{t-s}x - \int_s^t T^*_{t-\rho}M^{\frac{1}{2}}ST^S_{\rho-s}x d\rho$

is exponentially stable, or equivalently, the semigroup T^1_t given by

(6.26) $T^1_{t-s}x = T_{t-s}x - \int_s^t T_{t-\rho}S^*M^{\frac{1}{2}}T^1_{\rho-s}x d\rho$

is exponentially stable.
i.e. there exists K , $\omega > 0$ such that $||T^1_t|| \leq Ke^{-\omega t}$.

Proposition 6.7

Suppose the pair $\{A^*, M^{\frac{1}{2}}\}$ is stabilisable and that $Q \geq 0$ is a solution to the algebraic Riccati equation (6.22), then the semigroup given by

(6.27) $T^2_{t-s}x = T_{t-s}x - \int_s^t T_{t-\rho}BR^{-1}B^*QT^2_{\rho-s}x d\rho$

is exponentially stable.

<u>**Proof**</u>

Since $\{A^*, M^{\frac{1}{2}}\}$ is stabilizable there exists an operator S such that the semigroup T^1_t given by (6.26) is exponentially stable, so from (6.26) and (6.27) we can easily see that

(6.28) $$T^2_{t-s}x = T^1_{t-s}x + \int_s^t T^1_{t-\rho}\{S^*M^{\frac{1}{2}} - BR^{-1}B^*Q\}T^2_{\rho-s}x\,d\rho \quad .$$

Now, Q satisfies the algebraic Riccati equation (6.22), i.e. for $x, y \in \mathcal{D}$

(6.29) $$\langle Qx, \tilde{A}y\rangle + \langle \tilde{A}x, Qy\rangle + \langle x, My\rangle + \langle x, QBR^{-1}B^*Qy\rangle = 0 \ ,$$

or using the equivalent form (6.24) where A^2 is the infinitesimal generator of T^2_t, so that $A^2 = \tilde{A}$ on $\mathcal{D}$,

(6.30) $$\langle Qx, A^2y\rangle + \langle A^2x, Qy\rangle + \langle x, My\rangle + \langle x, QBR^{-1}B^*Qy\rangle = 0 \quad .$$

If we let $x = y = T^2_s k$ in (6.30) we find

(6.31) $$\langle QT^2_sk, A^2T^2_sk\rangle + \langle A^2T^2_sk, QT^2_sk\rangle + \langle T^2_sk, MT^2_sk\rangle + \langle T^2_sk, QBR^{-1}B^*QT^2_sk\rangle = 0,$$

then since A^2 is the infinitesimal generator of T^2_t,

$$\frac{d}{dt}\langle QT^2_sk, T^2_sk\rangle = \langle QA^2T^2_sk, T^2_sk\rangle + \langle QT^2_sk, A^2T^2_sk\rangle$$

and thus (6.31) implies

(6.32) $$\frac{d}{ds}\langle QT^2_sk, T^2_sk\rangle + ||M^{\frac{1}{2}}T^2_sk||^2 + ||R^{-\frac{1}{2}}B^*QT^2_sk||^2 = 0 \quad \text{for all} \quad s \geq 0 \ ,$$

and as $Q \geq 0$ and R invertible we can deduce that

(6.33)

a) $$\int_{t_0}^{\infty} ||M^{\frac{1}{2}}T^2_sk||^2 ds < \infty$$

b) $$\int_{t_0}^{\infty} ||R^{-\frac{1}{2}}B^*QT^2_sk||^2 ds < \infty$$

since (6.32) implies

$$0 \leq \int_{t_o}^{\infty} ||M^{\frac{1}{2}}T_s^2k||^2 ds + \int_{t_o}^{\infty} ||R^{-\frac{1}{2}}B^*QT_s^2k||^2 ds = -\int_{t_o}^{\infty} \frac{d}{ds} \langle QT_s^2k, T_s^2k\rangle ds .$$

From (6.28) we have

$$||T_{t-t_o}^2 x|| \leq ||T_{t-t_o}^1 x|| + \int_{t_o}^{t} ||T_{t-\rho}^1\{S^*M^{\frac{1}{2}} - BR^{-1}B^*Q\}T_{\rho-t_o}^2 x||d\rho$$

$$\leq ||T_{t-t_o}^1 x|| + \int_{t_o}^{t} ||T_{t-\rho}^1||_{\mathcal{L}(\tilde{W},H)} \Big\{||S^*||_{\mathcal{L}(H,\tilde{W})} ||M^{\frac{1}{2}}T_{\rho-t_o}^2 x|| + ||B||_{\mathcal{L}(U,\tilde{W})} ||R^{-1}B^*QT_{\rho-t_o}^2 x||\Big\} d\rho ,$$

so Young's inequality implies

$$\Big(\int_{t_o}^{\infty} ||T_{t-t_o}^2 x||^2 dt\Big)^{\frac{1}{2}}$$

$$\leq \Big(\int_{t_o}^{\infty} ||T_{t-t_o}^1 x||^2 dt\Big)^{\frac{1}{2}} + \int_{t_o}^{\infty} ||T_{\rho-t_o}^1||_{\mathcal{L}(\tilde{W},H)} d\rho \Big(\int_{t_o}^{\infty} \Big(||S^*||_{\mathcal{L}(H,\tilde{W})} ||M^{\frac{1}{2}}T_{\rho-t_o}^2 x|| + ||B||_{\mathcal{L}(U,\tilde{W})} ||R^{-1}B^*QT_{\rho-t_o}^2 x||\Big)^2 d\rho\Big)^{\frac{1}{2}}$$

and hence T_t^2 is exponentially stable provided $\int_{t_o}^{\infty} ||T_{\rho-t_o}^1||_{\mathcal{L}(\tilde{W},H)} d\rho < \infty$, using the inequalities (6.33), the fact that T_t^1 is exponentially stable and the following lemma - proof of which may be found in {11}.

<u>Lemma 6.8</u>

A strongly continuous semigroup T_t, defined on a Hilbert space H is exponentially stable if and only if for each $x \in H$ the integral $\int_0^{\infty} ||T_t x||^2 dt$ is convergent.

Since T_t^1 is the semigroup given by the perturbation of T_t by $-S^*M^{\frac{1}{2}}$ we have from the perturbation theory that there exists a function $\bar{g} \in L^p[0,t_1]$, for all $t_1 < \infty$, such that $||T_{t-s}^1||_{\mathcal{L}(\tilde{W},H)} \leq \bar{g}(t-s)$ and since T_t^1 is exponentially stable there exist constants K , $\omega > 0$ such that $||T_t^1|| \leq Ke^{-\omega t}$. Also by the semigroup property of T_t^1

$$||T_{t+s}^1||_{\mathcal{L}(\tilde{W},H)} \leq ||T_s^1||_{\mathcal{L}(\tilde{W},H)} ||T_t^1||_{\mathcal{L}(H)}$$

and so

$$(6.34) \qquad ||T_{t+s}^1||_{\mathcal{L}(\tilde{W},H)} \leq Ke^{-\omega t}\,\bar{g}(s) \quad .$$

Consider now

$$\int_s^{t_1} ||T_\rho^1||_{\mathcal{L}(\tilde{W},H)} d\rho = \int_0^{t_1-s} ||T_{s+\rho'}^1||_{\mathcal{L}(\tilde{W},H)} d\rho' \quad \text{putting } \rho = s+\rho'$$

$$\leq Ke^{-\omega s}\int_0^{t_1-s} \bar{g}(\rho')d\rho' \qquad \text{from (6.34)}$$

$$= Ke^{-\omega(t_1-\beta)}\int_0^{\beta} \bar{g}(\rho')d\rho' \qquad \text{letting } t_1-s = \beta$$

then, as $s = t_1-\beta$, we have shown that

$$\int_s^{t_1} ||T_\rho^1||_{\mathcal{L}(\tilde{W},H)} d\rho = \int_{t_1-\beta}^{t_1} ||T_\rho^1||_{\mathcal{L}(\tilde{W},H)} d\rho$$

$$\leq Ke^{-\omega(t_1-\beta)}\int_0^{\beta} \bar{g}(\rho')d\rho' \quad .$$

Now, if we take $t_1 = n\beta$

$$\int_{(n-1)\beta}^{n\beta} ||T_\rho^1||_{\mathcal{L}(\tilde{W},H)} d\rho \leq Ke^{-\omega(n-1)\beta}\int_0^{\beta} \bar{g}(\rho')d\rho'$$

so

$$\int_0^\infty ||T_\rho^1||_{\mathcal{L}(\tilde{W},H)} = \sum_{n=1}^{\infty} \int_{(n-1)\beta}^{n\beta} ||T_\rho^1||_{\mathcal{L}(\tilde{W},H)} d\rho \leq \sum_{n=1}^{\infty} Ke^{-\omega(n-1)\beta}\int_0^{\beta} \bar{g}(\rho')d\rho' < \infty$$

thus completing the proof of proposition 6.7.

Proposition 6.9

If the pair $\{A^*, M^{\frac{1}{2}}\}$ is stabilizable, $\overline{\mathcal{D}} = H$, then the solution to the algebraic Riccati equation (6.22) is unique.

Proof

Let $Q \geq 0$ be another solution to the Riccati equation (6.22) different from Q^∞ the unique solution to the integral Riccati equation (6.19). Q therefore satisfies

$$(6.35) \quad \langle Qx, \tilde{A}y\rangle + \langle \tilde{A}x, Qy\rangle + \langle x, My\rangle + \langle QBR^{-1}B^*Qx, y\rangle = 0$$

for x , $y \in \mathcal{D} = D(\tilde{A}) \cap \ker\{D - R^{-1}B^*Q\}$, or using the equivalent form (6.24) Q satisfies

$$(6.36) \quad \langle Qx, A^2y\rangle + \langle A^2x, Qy\rangle + \langle x, My\rangle + \langle QBR^{-1}B^*Qx, y\rangle = 0$$

for x , $y \in D(A^2)$ where $A^2 = \tilde{A}$ on $\mathcal{D}$ and A^2 is the infinitesimal generator of T_t^2 given by (6.27). Thus

$$\frac{d}{ds}\langle T_s^2k, QT_s^2k\rangle = \langle A^2T_s^2k, QT_s^2k\rangle + \langle QT_s^2k, A^2T_s^2k\rangle$$

so putting $x = y = T_s^2k$ in (6.36) gives

$$\langle QBR^{-1}B^*QT_s^2k, T_s^2k\rangle + \langle T_s^2k, MT_s^2k\rangle = -\frac{d}{ds}\langle T_s^2k, QT_s^2k\rangle$$

and hence

$$\int_0^\infty \langle T_s^2k, \{M + QBR^{-1}B^*Q\}T_s^2k\rangle ds = -\int_0^\infty \frac{d}{ds}\langle T_s^2k, T_s^2k\rangle ds$$

$$= \langle k, Qk\rangle$$

since, by proposition 6.7, T_s^2 is exponentially stable. Thus

$$Qy = \int_0^\infty T_s^{2^*}\{M + QBR^{-1}B^*Q\}T_s^2y\,ds$$

i.e. Q also satisfies the integral Riccati equation (6.19) which we know has unique solution Q^∞. Hence the solution to the algebraic Riccati equation is unique.

Proposition 6.10

If $M^{-1}, M > 0$ and $\overline{\mathcal{D}} = H$, then the solution to the algebraic Riccati equation (6.22) is unique.

Proof

Firstly we show that T_t^∞ is exponentially stable, where T_t^∞ is the perturbed semigroup corresponding to the perturbation of T_t by $-BR^{-1}B^*Q^\infty$ and where Q^∞ is the unique solution to the integral Riccati equation (6.19) and so,

$$<x, Q^\infty x> = <x, \int_{t_o}^{\infty} T_{s-t_o}^{\infty *}\{M + Q^\infty BR^{-1}B^*Q^\infty\}T_{s-t_o}^\infty x ds> < \infty \;, \quad x \in H .$$

This implies $<x, \int_{t_o}^{\infty} T_{s-t_o}^{\infty *} M T_{s-t_o}^\infty x ds> < \infty$, but since $M > 0$ this gives $\int_{t_o}^{\infty} ||T_{s-t_o}^\infty x||^2 ds < \infty$ thus, by lemma 6.8, T_t^∞ is exponentially stable.

Since $M > 0$, there exists a unique positive self-adjoint, bounded operator C, such that $C = M^{\frac{1}{2}}$ i.e. $C^*C = M$, with bounded inverse C^{-1}. Let $S^* = BR^{-1}B^*Q^\infty C^{-1}$ then $S^*C = BR^{-1}B^*Q^\infty$ and hence the perturbed semigroup corresponding to the perturbation of T_t by $-S^*C$ is the same as T_t^∞, which is exponentially stable. So the pair $\{A^*, M^{\frac{1}{2}}\}$ is stabilisable and we have already shown that this implies that the algebraic Riccati equation has unique solution.

Example 6.11

Consider again the quadratic cost control problem of example 4.9, but now over the infinite interval, i.e. we wish to minimize

(6.37) $J(u) = \int_0^\infty\int_0^1 z^2(x,t)dxdt + c\int_0^\infty u^2(t)dt$

where z is given by the solution to the controlled heat equation

(6.38) $z_t = z_{xx}$

$z(o,t) = u(t)$, $z(1,t) = 0$, $z(x,o) = z_o(x)$.

From example 2.7 we know that for this system our abstract formulation is valid and furthermore the system operator for (6.38) is self-adjoint and generates a stable semigroup ({11}) so the infinite time problem is well posed. Thus we know that there exists a unique optimal control $u^* = -R^{-1}B^*Qz$ where Q satisfies the algebraic Riccati equation (6.21). (6.21) is however just the time independent version of (4.25) hence again, writing $Q(z)(\zeta) = \int_0^1 K(\zeta,\eta)z(\eta)d\eta$, from example 4.9 with K independent of t , the optimal control is $u^* = -\frac{1}{c}\int_0^1 K_\zeta(o,\eta)z(\eta)d\eta$ where K satisfies

$$K_{\zeta\zeta} + K_{\eta\eta} - \frac{1}{c}K_\zeta(o,\eta)K_\eta(\zeta,o) + \delta(\zeta-\eta) = 0$$

with $K(o,\eta) = K(1,\eta) = K(\zeta,o) = K(\zeta,1) = 0$.

If we expand K in terms of the basis for T_t , i.e. $\{\sqrt{2}\sin m\pi x\}$ so then

$$K(\zeta,\eta) = 2\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} a_{mn} \sin m\pi\zeta \sin n\pi\eta$$

and again using example 4.9 we have that the coefficients a_{mn} satisfy

$$(m^2\pi^2 + n^2\pi^2)a_{mn} + \frac{2}{c}\sum_{i=0}^{\infty} i\pi a_{in} \sum_{j=0}^{\infty} j\pi a_{mj} - \delta_m^n = 0 \quad .$$

7. A STABILIZABILITY RESULT

Using the results of the infinite time problem from section 6 we can prove the following proposition.

Proposition 7.1

If the system

$$(7.1) \qquad z(t) = T_t z_o + \int_o^t T_{t-s} Bu(s)ds \ , \qquad z(o) = z_o \ ,$$

where T_t and B satisfy (2.4) with $p \geq 1$, is exactly null controllable on $[0,t_1]$ with controls $u \in L^2[0,t_1;U]$ then the system is exponentially stabilizable.

Proof

Since the system is exactly null controllable on $[0,t_1]$ there exists a control $u \in L^2[0,t_1;U]$ such that $z(t_1) = 0$. If we now play this control on $[0,t_1]$ and the zero control for $t > t_1$ so that $z(t) = 0$ for $t > t_1$, then

$$J(u) = \int_o^\infty \{||z(s)||_H^2 + ||u||_U^2\}ds < \infty \ .$$

We saw in section 6 that this condition is sufficient for the existence of a unique optimizing control $u^* \in L^2[0,\infty;U]$ which minimizes $J(u)$ and, furthermore, u^* is bounded feedback. Implementing this optimal feedback control ensures the existence of a $D \in \mathcal{L}(H,U)$ such that

$$(7.2) \qquad \int_o^\infty \{||T_t^D z_o||_H^2 + ||DT_t^D z_o||_U^2\}dt \leq k||z_o||^2$$

where T_t^D is the perturbed semigroup corresponding to the perturbation of T_t by BD, i.e.

$$T_t^D x = T_t x + \int_o^t T_{t-s} BDT_s^D x ds$$

and (7.2) implies

$$\int_0^\infty ||T_t^D z_o||^2 dt \leq k||z_o||^2 \quad .$$

Lemma 6.8 completes the proof.

Remarks

Even though in general for $p \geq 1$ solutions to (7.1) are only L^2, and not continuous, we note that for a certain class of controls continuous solutions to (7.1) are obtained. The requirement that the system be exactly null controllable is therefore tenable.

The result of proposition 7.1 is identical to that proved for bounded operators B in {11} .

8. ALTERNATIVE APPROACHES

A semigroup approach to boundary control systems is also adopted by Balakrishnan { 2 } and { 3 }, Barbu { 4 } and Washburn {39} following basically the model of Fattorini {13}, and, with a slightly different formulation by Zabczyk {43} and {32}. Lions {24} has developed a general theory for boundary control action where the operators are assumed to satisfy a coercivity condition. In this section we investigate how the approach of these authors compares with that taken in this thesis.

(a) Semigroup Approach Based on Fattorini's Model

In order to compare with our approach that of Balakrishnan, Barbu, Fattorini and Washburn we firstly detail the model of Fattorini {13} on which the other authors base their work.

Fattorini's Model

Let Z be a Banach space, σ a closed linear densely defined operator in Z, and let τ be a linear operator with domain in Z and range some Banach space Y. Also let U be a Banach space, the control space of the system. Consider now the system,

$$(8.1) \qquad \left| \begin{array}{l} y'(t) = \sigma y(t) \\ \tau y(t) = F_1 u(t) \end{array} \right. \qquad \text{over } [o,t_1]$$

with initial condition

$$(8.2) \qquad y(o) = y^o$$

where $F_1 : U \to Y$ is a continuous linear operator and $[o,t_1]$ a fixed interval and where $u(\cdot)$ is a summable function on $[o,t_1]$ with values in U. Fattorini makes the following assumptions,

Assumption 8.1

$D(\sigma) \subset D(\tau)$ and the restriction of τ to $D(\sigma)$ is continuous relative to the graph norm of $D(\sigma)$.

Let $A : Z \to Z$ be the linear operator defined by

$$(8.3) \qquad D(A) = \{y \in D(\sigma);\ \tau y = 0\}\ , \quad Ay = \sigma y \quad \text{for} \quad y \in D(A)\ ,$$

then

Assumption 8.2

The operator A is the infinitesimal generator of a strongly continuous semigroup $\{T_t;\ t \geq 0\}$ on Z.

and

Assumption 8.3

There exists a continuous linear operator $F : U \to Z$ such that

(8.4) $\quad \sigma F \in \mathcal{L}(U,Z) \quad , \quad \tau(Fu) = F_1 u \qquad$ for all $u \in U$,

(8.5) $\quad ||Fu||_Z \leq k||F_1 u||_Y \qquad$ for all $u \in U$, k a positive constant .

In terms of A and F the system (8.1) can be written as

$$(8.6) \qquad \begin{cases} y' = Az + \sigma Fu \\ y = z + Fu \end{cases} \qquad o \leq t \leq t_1 \ ,$$

If $u(\cdot)$ is continuously differentiable on $[o,t_1]$ then z can be defined as a mild solution to the Cauchy problem

$$\begin{cases} z' = Az + \sigma Fu - Fu' \\ z(o) = y^o - Fu(o) \end{cases}$$

and thus we may define the solution y to the system (8.1) - (8.2) by the variation of constants formula

$$(8.7) \qquad y(t) = T_t\{y^o - Fu(o)\} + Fu(t) + \int_o^t T_{t-s}\{\sigma Fu(s) - Fu'(s)\}ds \ .$$

Since differentiability of the controller u represents an unrealistic and severe requirement the other authors are led to extend the concept of solutions to (8.1) - (8.2) for general $u \in L^1[o,t_1;U]$ (or $u \in L^2[o,t_1;U]$). Barbu { 4 } extends the notion of a solution to hold for all $u \in L^1[o,t_1;U]$ in the following way:-

Integrating (formally) by parts in (8.7) we obtain

$$(8.8) \qquad y(t) = T_t y^o - \int_o^t AT_{t-s}Fu(s)ds + \int_o^t T_{t-s}\sigma Fu(s)ds \ ,$$

though in general, unless we impose further assumptions on T_t and F , the right hand side of (8.8) is not well defined, hence

<u>Assumption 8.4</u>

For each $t \in [o,t_1]$ and $u \in U$, $T_t Fu \in D(A)$, and there exists a positive function $\gamma \in L^1[o,t_1]$ such that

$$(8.9) \qquad ||AT_tF||_{\mathcal{L}(U,Z)} \le \gamma(t) \quad \text{a.e. } t \in [o,t_1] .$$

Since $T_tFu \in D(A)$ for all $u \in U$, by the closed graph theorem, the operator AT_tF is continuous from U to Z so that (8.9) makes sense. Assumption 8.4 then implies that for every $u \in L^1[o,t_1;U]$ the function $t \to \int_o^t AT_{t-s}Fu(s)ds$ is well defined as an element of $L^1[o,t_1;Z]$. Then, by definition, for each $y^o \in Z$, $u \in L^1[o,t_1;U]$, the function $y \in L^1[o,t_1;Z]$ defined by (8.8) is the solution of the boundary control system (8.1) - (8.2).

Also, since the function $t \to \int_o^t T_{t-s}Fu(s)ds$ belongs to $L^1[o,t_1;D(A)]$, $y(\cdot)$ may be expressed in the following equivalent form

$$(8.10) \qquad y(t) = T_ty^o - A\int_o^t T_{t-s}Fu(s)ds + \int_o^t T_{t-s}\sigma Fu(s)ds \quad \text{a.e. } t \in [o,t_1] .$$

Washburn {39} considers those operators F for which $\sigma F = 0$ in Z and takes $y^o = 0$ so then the generalised solution (8.10) has the form

$$(8.11) \qquad y(t) = -\int_o^t AT_{t-s}Fu(s)ds .$$

For arbitrary $u \in L^2[o,t_1;U]$, y given by (8.11) is not defined pointwise, i.e. $y(t) \in Z$ only for almost all t. In order to find solutions with $y(t) \in Z$ for all t Washburn considers the following:-

Define L by $Lu = y$ where

$$y(t) = \int_o^t AT_{t-s}(Fu)(s)ds$$

and define

$$L_{t_1}u = Lu(t_1) = \int_o^{t_1} AT_{t_1-s}(Fu)(s)ds$$

when the above makes sense. Washburn then gives the following theorem.

Theorem 8.5

If

(8.12) $|AT_tF| = O(t^{-\theta})$

then $L_{t_1} : L^p[o,t_1;U] \to Z$ for all p such that $\frac{1}{1-\theta} < p \leq \infty$.

We have already noted that $L : L^2[o,t_1;U] \to L^2[o,t_1,Z]$ continuously so we now further note that $L : L^p[o,t_1;U] \to L^\infty[o,t_1;Z]$ continuously when $\frac{1}{1-\theta} < p \leq \infty$.

We now consider how the system of example 2.1 would be formulated in this approach.

Example 8.6

Consider again the system (2.1) - (2.2), i.e.

(8.13) $z_t = z_{xx}$

$z(o,t) = u(t)$, $z(1,t) = 0$, $z(x,o) = z_o(x)$ on $Z = L^2(0,1)$.

If we take $\sigma z = z_{xx}$ with $D(\sigma) = \{z \in Z : z_{xx} \in Z , z(1) = 0\}$ and define τ by $\tau z = z(o,t)$, then (8.13) can be written as

$z_t = \sigma z$

$\tau z(t) = u(t)$, $z(x,o) = z_o(x)$,

i.e. in the form (8.1) with $F_1 = I$.

Let A be the operator defined from σ by (8.3), i.e. $D(A) = \{z \in D(\sigma) , \tau z = 0\} = \{z : z, z_{xx} \in Z , z(o) = z(1) = 0\}$ and $Az = \sigma z = z_{xx}$ on $D(A)$. Finally let F be given by $Fu = (1-x)u$. Then assumptions 8.1 - 8.3 are satisfied and furthermore $\sigma F = 0$.

We now construct a solution to (8.13) following Fattorini's method. From (8.13)

$$\frac{\partial}{\partial t}(z - Fu) = \frac{\partial^2}{\partial x^2}(z - Fu) - \frac{\partial}{\partial t} Fu$$

as $\frac{\partial^2}{\partial x^2} Fu = 0$. Since $z - Fu \in D(A)$, if we put $w = z - Fu$, then we have that

$$w_t = Aw - \frac{d}{dt} Fu$$

which gives integrating

$$w(t) = T_t w_o - \int_o^t T_{t-\rho} \frac{d}{d\rho}(Fu)(\rho)d\rho$$

where T_t is the semigroup generated by A , given by (2.12), and

$$w_o = (z - Fu)\big|_{t=0} = z_o(x) - Fu\big|_{t=0} = z_o(x) - Fu_o \ .$$

So

$$\begin{aligned} z(t) &= w(t) + Fu(t) \\ &= T_t w_o - \int_o^t T_{t-\rho} \frac{d}{d\rho}(Fu)(\rho)d\rho + (Fu)(t) \end{aligned}$$

which gives, integrating by parts ,

$$z(t) = T_t z_o - \int_o^t AT_{t-\rho}(Fu)(\rho)d\rho$$

i.e.

$$z(t) = T_t z_o - \int_o^t AT_{t-\rho}(1 - x)u(\rho)d\rho \ . \tag{8.14}$$

Note: $F \not\in D(A)$ so we cannot commute A and T_t .

This alternative approach places restrictions on the class of operators different to those required by our approach. Barbu { 4 } and Washburn {39} apply an analyticity type condition, i.e. $T_t Fu \in D(A)$ whilst Balakrishnan { 3 } considers only second order strongly elliptic operators with controls $u \in L^2[o,t_1;U]$ and, in { 2}, restricts consideration still further to the case of the Laplacian, (using then the fact that the semigroup generated is analytic the assumptions on A are automatically satisfied.)

Note that we require, from (2.4)(c) that $T_t^* \in \mathcal{L}(H,\bar{W}^*)$ for $t > 0$ which is similar to requiring that T_t^* be an analytic semigroup for which $T_t^* : H \to D(A^*)$, $t > 0$. We are however not able to use $\bar{W}^* = D(A^*)$ with the graph norm since then $||A^*T_t^*|| \leq \frac{M}{t}$ for some $M > 0$ and so (2.4)(d) will not hold.

We now show that when the alternative Fattorini approach is valid the two formulations (i.e. the Fattorini formulation and our formulation) are equivalent.

<u>Theorem 8.7</u>

Suppose $z(t)$ is given by the Fattorini formulation with $\sigma F = 0$, i.e.

(8.15) $$z(t) = T_t z_o + \int_o^t AT_{t-s}Fu(s)ds$$

then $z(t)$ is a weak solution of

$$\dot{z} = \bar{A}z$$

$$Dz = u \quad , \quad Ez = 0 \quad , \; z(o) = z_o \quad .$$

The proof of theorem 8.7 is similar to the proof of theorem 2.4 using the following proposition (cf. Proposition 2.3).

<u>Proposition 8.8</u>

Consider $f \in C[o,t_1;D(A^*)]$, $x(t) = -\int_t^{t_1} T_{s-t}^* f(s)ds$, and $z(t)$ satisfying (8.15), then $z(t)$ also satisfies (2.6), i.e.

(8.16) $$\int_o^{t_1} \langle f(t), z(t)\rangle_{Z^*,Z} dt + \int_o^{t_1} \langle Cx(t), u(t)\rangle_{U^*,U} dt + \langle x(o), z_o\rangle_{Z^*,Z} = 0$$

assuming there exists a Green's formula (2.7) with $Dz = u$, $Ez = 0$.

<u>Proof</u>

Substituting for $z(t)$ from (8.15) into the left hand side of (8.16) gives

$$\int_0^{t_1}\langle f(t),\ T_t z_o + \int_0^t AT_{t-s}Fu(s)ds\rangle dt + \int_0^{t_1}\langle -\ C\int_t^{t_1}T^*_{s-t}f(s)ds,\ u(t)\rangle dt$$

$$- \langle\int_0^{t_1}T^*_s f(s)ds,\ z_o\rangle$$

$$= \int_0^{t_1}\langle f(t),\ \{T_t - T_t\}z_o\rangle dt - \int_0^{t_1}\int_0^s\langle CT^*_{s-t}f(s),\ u(t)\rangle dtds$$

$$+ \int_0^{t_1}\int_0^t\langle f(t),\ AT_{t-s}Fu(s)\rangle dsdt$$

since:-

(a) $$\langle\int_0^{t_1}T^*_s f(s)ds,\ z_o\rangle = \int_0^{t_1}\langle T^*_s f(s),\ z_o\rangle ds$$

as $$\int_0^{t_1}|\langle f(s),\ T_s z_o\rangle|ds \le \int_0^{t_1}||f(s)||\ ||T_s z_o||ds < \infty\ .$$

(b) $$\int_0^t\int_0^{t_1}\langle CT^*_{s-t}f(s),\ u(t)\rangle dtds = \int_0^{t_1}\int_0^s\langle CT^*_{s-t}f(s),\ u(t)\rangle dsdt$$

changing the order of integration

$$= \int_0^t\langle C\int_t^{t_1}T^*_{s-t}f(s)ds,\ u(t)\rangle dt$$

as $$\int_0^t\int_0^{t_1}|\langle CT^*_{s-t}f(s),\ u(s)\rangle|dsdt \le \int_0^t\int_0^{t_1}||CT^*_{s-t}||\ ||f(s)||\ ||u(t)||dsdt < \infty$$

because $||CT^*_{s-t}||$ is bounded by an L^2-function from assumption (2.4) and we have $||u|| \in L^2$ and f is continuous.

(c) $$\int_0^{t_1}\langle f(t),\ \int_0^t AT_{t-s}Fu(s)ds\rangle dt = \int_0^{t_1}\int_0^t\langle f(t),\ AT_{t-s}Fu(s)\rangle dsdt$$

as $$\int_0^{t_1}\int_0^t|\langle f(t),\ AT_{t-s}Fu(s)\rangle|dsdt \le \int_0^{t_1}\int_0^t||f(t)||\ ||AT_{t-s}F||\ ||u(s)||dsdt < \infty$$

with $||AT_{t-s}F|| \in L^q$ by assumption.

Now $$\int_o^{t_1}\int_o^t \langle f(t),\ AT_{t-s}Fu(s)\rangle ds dt$$

$$= \int_o^{t_1}\int_o^t \langle T^*_{t-s}A^* f(t),\ Fu(s)\rangle ds dt$$

$$= \int_o^{t_1}\int_o^t \langle A^* T^*_{t-s} f(t),\ Fu(s)\rangle ds dt$$

$$= \int_o^{t_1}\int_o^t \langle T^*_{t-s} f(t),\ \tilde{A}Fu(s)\rangle ds dt - \int_o^{t_1}\int_o^t \langle CT^*_{t-s} f(t),\ DFu(s)\rangle ds dt$$

$$- \int_o^{t_1}\int_o^t \langle GT^*_{t-s} f(t),\ EFu(s)\rangle ds dt$$

from the Green's formula (2.7)

$$= \int_o^{t_1}\int_o^t \langle T^*_{t-s} f(s),\ \tilde{A}Fu(s)\rangle ds dt - \int_o^{t_1}\int_o^t \langle CT^*_{t-s} f(t),\ u(s)\rangle ds dt$$

as $Dz = u$, $Ez = 0$

$$= \int_o^{t_1}\int_o^t \langle CT^*_{t-s} f(t),\ u(s)\rangle ds dt$$

since $\tilde{A} = \sigma$ implies $\tilde{A}Fu(s) = 0$.

Thus the left hand side of (8.16) = 0, and so proposition 8.8 is established.

$\int_o^{t_1} \langle Cx(t),\ u(t)\rangle dt$ is well defined for all $f \in C[o,t_1;Z^*]$ as

$$\int_o^{t_1} \langle Cx(t),\ u(t)\rangle dt = - \int_o^{t_1} \langle C\int_t^{t_1} T^*_{s-t} f(s) ds,\ u(t)\rangle dt$$

and

$$\int_o^{t_1}\int_t^{t_1} |\langle CT^*_{s-t} f(s),\ u(t)\rangle| ds\ dt \leq \int_o^{t_1}\int_t^{t_1} ||CT^*_{s-t}||\ ||f(s)||\ ||u(t)|| ds dt < \infty$$

with $||\ CT^*_{s-t}||$ bounded by an L^2-function, $||u|| \in L^2$, and f is continuous.
Thus, since $D(A^*)$ is dense in Z^* we can extend (8.16) to hold for all $f \in C[o,t_1;Z^*]$ and so we have that if $z(t)$ satisfies (8.15) it satisfies

(8.16) for $f \in C[o,t_1;Z^*]$. Hence, as in the proof of theorem 2.4, we have that $z(t)$ given by (8.15) is a weak solution of

$$\dot{z} = \tilde{A}z$$

$$Dz = u \quad , \quad Ez = 0 \quad , z(o) = z_o \ .$$

Using the green's formula (2.7) it is easy to see directly, formally, the equivalence of the two forms. For $x \in D(A^*)$ consider

$$<x, AT_{t-s}Fu> = <(AT_{t-s})^* x, Fu>$$

$$= <A^* T^*_{t-s} x, Fu>$$

$$= <T^*_{t-s} x, \tilde{A}Fu> \quad <CT^*_{t-s} x, DFu>$$

from (2.7) with $Ez = 0$

$$= - <CT^*_{t-s} x, u>$$

since $\tilde{A}Fu = \sigma Fu = 0$ and $Dz = u$

$$= - <T^*_{t-s} x, Bu>$$

$$= - <x, T_{t-s}Bu>$$

giving, since $\overline{D(A^*)} = Z^*$,

$$AT_{t-s}Fu = - T_{t-s}Bu \ .$$

This equivalence is demonstrated by considering again the system of examples 2.1 and 8.6.

<u>Example 8.9</u>

Consider the system

(8.17) $z_t = z_{xx}$

$$z(o,t) = u \quad , \quad z(1,t) = 0 \quad , \quad z(x,o) = z_o(x) \quad , \text{ on } \ Z = L^2(0,1) \ ,$$

which we have shown in example 2.1 to have solution

$$(8.18) \quad z(t) = T_t z_o + \int_o^t T_{t-s} Bu(s)ds$$

$$= T_t z_o - \int_o^t T_{t-s} \delta' u(s)ds$$

and in example 8.6 to have solution

$$(8.19) \quad z(t) = T_t z_o - \int_o^t AT_{t-s} Fu(s)ds$$

$$= T_t z_o - \int_o^t AT_{t-s}(1 - x)u(s)ds$$

where T_t is the semigroup generated by $A = \frac{d^2}{dx^2}$ with $D(A) = \{z \in Z\ ,\ z_{xx} \in Z\ ,\ z(1) = 0 = z(o)\}$, and so $A^* = A$ with $D(A^*) = D(A)$.

Let $z \in D(A^*)$ then

$$<A^*z,\ Fu> = \int_o^1 \frac{d^2}{dx^2} z(1 - x)udx$$

$$= \left[\frac{d}{dx} z\ (1 - x)u\right]_o^1 - \int_o^1 \frac{d}{dx} z \cdot \frac{d}{dx} (1 - x)udx$$

$$= -\ z_x(o)u - \int_o^1 \frac{d}{dx} z \cdot \frac{d}{dx} (1 - x)udx$$

$$= -\ z_x(o)u + \int_o^1 z_x udx$$

$$= -\ z_x(o)u \quad .$$

From example 2.7

$$<z,\ Bu> = <Cz,\ u> = z_x(o)u$$

Hence

$$<A^*z,\ Fu> = -<z,\ Bu>$$

Let $z = T^*_{t-s}z'$, $z' \in D(A^*)$, then $z \in D(A^*)$ and

$$<A^*T^*_{t-s}z',\ Fu> = -<T^*_{t-s}z',\ Bu>$$

i.e. $\langle (AT_{t-s})^* z', Fu\rangle = - \langle z', T_{t-s}Bu\rangle$

so $\langle z', AT_{t-s}Fu\rangle = - \langle z', T_{t-s}Bu\rangle$, $z' \in D(A^*)$,

and then since $\overline{D(A^*)} = Z$, $AT_{t-s}Fu = - T_{t-s}Bu$ and so (8.18) and (8.19) are equivalent.

(b) Zabczyk's Approach

Zabczyk's approach ({43} and {32}) is slightly different to that of Balakrishnan, Barbu, Fattorini and Washburn considered above. He considers U , Z again two Banach spaces with A the infinitesimal generator of a strongly continuous semigroup T_t , on Z , and he takes F and G as bounded operators from U into Z . He then calls a function $u(\cdot)$ from $[0,\infty)$ into U a boundary control if it is locally integrable, for $t \geq 0$ $\int_0^t T_{t-s}Fu(s)ds \in D(A)$ and

(8.20) $$y(t) = T_t y_0 - A\{\int_0^t T_{t-s}Fu(s)ds\} + \int_0^t T_{t-s}Gu(s)ds$$

is continuous. Then if we assume $u(\cdot)$ is twice continuously differentiable and $y_0 - Fu_0 \in D(A)$, (8.20) is the unique continuous solution of

(8.21) $$\dot{y}(t) = A\{y(t) - Fu(t)\} + Gu(t) \quad , \quad t \geq 0 ,$$

and $y(t) - Fu(t) \in D(A)$ for $t \geq 0$. It follows that if $u(\cdot)$ is sufficiently smooth and $y_0 - Fu_0 \in D(A)$ then the elements $y(t)$ and $Fu(t)$ satisfy the same boundary conditions as formulated in the definition of the domain $D(A)$. Usually, in applications, the generator A is a restriction of the operator σ defined on a domain $D(\sigma) \supset D(A)$. If now, for every $u \in U$, $Fu \in D(\sigma)$, $\sigma(Fu) = Gu$ then (8.21) becomes

$$\dot{y}(t) = \sigma y(t) \quad , \quad t \geq 0$$
$$y(t) - Fu(t) \in D(A)$$

As with the Fattorini approach we can consider the system of example 2.1 and see how this would be formulated in the Zabczyk approach.

__Example 8.10__

Consider again the system of example 2.1, i.e.

(8.22) $z_t = z_{xx}$

$z(o,t) = u \quad , \quad z(1,t) = 0 \quad , \quad z(x,o) = z_o(x) \quad , \quad \text{on} \quad L^2(0,1) = Z \; .$

Let $\sigma = \dfrac{d^2}{dx^2}$ on $D(\sigma) = \{z \in Z \; ; \; z_{xx} \in Z \; , \; z(1) = 0\}$ and let A be the restriction of σ to $D(A) = \{z \in Z \; ; \; z_{xx} \in Z \; , \; z(o) = z(1) = 0\}$ so then A is the generator of a strongly continuous semigroup T_t (given by (2.12)). Let F be given by $Fu = (1 - x)u$ and G by $Gu = 0$ then (8.22) can be written as

(8.23) $\dot{z} = \sigma z$

$z - Fu \in D(A) \; ,$

since $(z - Fu)_{xx} = (z - (1 - x)u)_{xx} = z_{xx} \in Z$ for $z \in D(\sigma)$ and

$z(o) = u$ implies $(z - Fu)(o) = z(o) - u = 0$

$z(1) = 0$ implies $(z - Fu)(1) = z(1) = 0 \; .$

But, since $\sigma(Fu) = Gu = 0$, (8.23) is equivalent to

(8.24) $\dot{z} = A(z - Fu)$

$z - Fu \in D(A) \; .$

Thus, $\dfrac{d}{ds} T_{t-s} z(s) = - AT_{t-s} z(s) + T_{t-s}\{A(z - Fu)(s)\}$

$= - T_{t-s} AFu$

$= - AT_{t-s} Fu$

and so integrating from 0 to t ,

$$z(t) = T_t z_o - A\int_o^t T_{t-s} Fu(s)ds$$

$$= T_t z_o - A\int_o^t T_{t-s}(1 - x)u(s)ds \ .$$

This class of problems, namely those such that $\int_o^t T_{t-s}Fu(s)ds \in D(A)$, for which the Zabczyk approach is valid is wider than those which the Fattorini approach is able to handle - recall assumption (8.4) requires $T_t Fu \in D(A)$.

In order that the solution $y(t)$ to (8.20) be continuous Zabczyk requires that $A\{\int_o^t T_{t-s}Fu(s)ds\}$ be continuous and he investigates when this holds, proving the following {43}.

<u>Proposition 8.11</u>

Let A be self-adjoint and the generator of a strongly continuous semigroup T_t on a Hilbert space H . If for some $b \in H$ and $p > 1$

$$(8.25) \qquad \int_{-\infty}^{\infty} |\lambda|^{2/p} \langle \underline{P}(d\lambda)b, b\rangle < \infty \ ,$$

where $\underline{P}(\cdot)$ is the spectral measure of A , then for every $u(\cdot) \in L^p(0,\infty)$ the function

$$y(t) = A\{\int_o^t T_{t-s}bu(s)ds\}$$

is continuous.

In order to show that the solution $z(t)$ given by the Zabczyk formulation is a weak solution of

$$\dot{z} = \bar{A}\, z$$

$$Dz = u \ , \quad Ez = 0 \ , \quad z(o) = z_o \ ,$$

we need to impose a further condition on the operator A , namely the condition (8.9) required in the Fattorini approach, i.e.

$$(8.26) \qquad ||AT_{t-s}F|| \le g(t-s) \ , \qquad g \in L^2[0,t_1] \ ,$$

in consideration of controls $u \in L^2[o,t_1;U]$. We can compare this condition with that given by Zabczyk as a sufficient condition for the solutions $z(t)$ to (8.20) to be continuous when A is self-adjoint, namely proposition 8.11, and prove the following.

Proposition 8.12

$$\int_{-\infty}^{\infty} |\lambda|^{2/p} <\underline{P}(d\lambda)b,\ b> \ < \infty \qquad \text{when} \quad p = 2$$

implies

$$\int_o^{t_1} ||AT_t b||^2 dt < \infty$$

i.e. $\quad ||AT_{t-s}b|| \leq g(t-s) \qquad$ with $g \in L^2[o,t_1]$

where $\underline{P}(\cdot)$ is the spectral measure of A , a self-adjoint operator.

Proof

Since A is self-adjoint the spectrum of A is contained in $[-\infty,\lambda_o]$ for some $\lambda_o < \infty$ and thus, ,

$$Az = \int_{-\infty}^{\lambda_o} \lambda \underline{P}(d\lambda)z \qquad \text{with } D(A) = \{z \in H\ ;\ \int_{-\infty}^{\lambda_o} |\lambda|^2 <\underline{P}(d\lambda)z,\ z> \ < \infty\} \ .$$

Also

$$T_t b = \int_{-\infty}^{\lambda_o} e^{\lambda t} \underline{P}(d\lambda)b$$

and hence

$$AT_t b = \int_{-\infty}^{\lambda_o} \lambda \underline{P}(d\lambda) \int_{-\infty}^{\lambda_o} e^{\lambda t} \underline{P}(d\lambda)b \ .$$

From {15}

$$\{\int f dE\}\{\int g dE\} = \{\int fg dE\} \qquad \text{where } E \text{ is a spectral measure, so}$$

$$AT_t b = \int_{-\infty}^{\lambda_o} \lambda e^{\lambda t} \underline{P}(d\lambda)b$$

and

$$||AT_t b||^2 = \int_{-\infty}^{\lambda_o} |\lambda|^2 \{e^{\lambda t}\}^2 <\underline{P}(d\lambda)b,\ b>$$

since from {18},

$$Hu = \int_{-\infty}^{\infty} \lambda dE(\lambda)u \quad \text{implies} \quad ||Hu||^2 = \int_{-\infty}^{\infty} \lambda^2 d\langle E(\lambda)u, u\rangle \quad .$$

Thus

$$||AT_t b||^2 = \int_{-\infty}^{\lambda_o} |\lambda|^2 e^{2\lambda t}\langle \underline{P}(d\lambda)b, b\rangle$$

and hence

$$\int_o^{t_1} ||AT_t b||^2 dt = \int_o^{t_1}\int_{-\infty}^{\lambda_o} e^{2\lambda t}|\lambda| \ |\lambda|\langle \underline{P}(d\lambda)b, b\rangle dt$$

$$= \int_{-\infty}^{\lambda_o}\{\int_o^{t_1}|\lambda| e^{2\lambda t}dt\}|\lambda|\langle \underline{P}(d\lambda)b, b\rangle$$

$$= \int_{-\infty}^{\lambda_o} \frac{1}{2\lambda}\, |\lambda|\{e^{2\lambda t_1} - 1\}|\lambda|\langle \underline{P}(d\lambda)b, b\rangle$$

$$\leq \max_{\lambda\varepsilon[-\infty,\lambda_o]} \frac{1}{2}\{e^{2\lambda t_1} - 1\}\int_{-\infty}^{\lambda_o}|\lambda|\langle \underline{P}(d\lambda)b, b\rangle$$

$$\leq K\int_{-\infty}^{\lambda_o}|\lambda|\langle \underline{P}(d\lambda)b, b\rangle \quad \text{some} \quad K < \infty$$

$$< \infty \quad \text{since} \quad \int_{-\infty}^{\lambda_o}|\lambda|\langle \underline{P}(d\lambda)b, b\rangle < \infty \quad .$$

Thus we have shown that condition (8.25), when $p = 2$ is sufficient for (8.26) to hold.

Similarly to proposition 8.8 we now prove.

<u>Proposition 8.13</u>

Consider $f \in C[o,t_1;D(A^*)]$, $x(t) = -\int_t^{t_1} T^*_{s-t} f(s)ds$ and $z(t)$ satisfying

$$z(t) = T_t z_o + A\int_o^t T_{t-s} Fu(s)ds \quad ,$$

then $z(t)$ also satisfies (2.6), i.e.

$$(8.27) \quad \int_o^{t_1}\langle f(t), z(t)\rangle_{Z^*,Z} + \int_o^{t_1}\langle Cx(t), u(t)\rangle_{U^*,U} + \langle x(o), z_o\rangle_{Z^*,Z} = 0$$

assuming that there exists a Green's formula (2.7) with $Dz = u$, $Ez = 0$, and that (8.26) holds.

<u>Proof</u>

Substituting for $z(t)$ from (8.8) into the left hand side of (8.27) gives

$$\int_0^{t_1} \langle f(t),\ T_t z_0 + A\int_0^t T_{t-s}Fu(s)ds\rangle dt + \int_0^{t_1}\langle -\ C\int_t^{t_1} T^*_{s-t}f(s)ds,\ u(t)\rangle dt$$

$$-\ \langle\int_0^{t_1} T^*_s f(s)ds,\ z_0\rangle$$

which equals, as in the proof of proposition 8.8,

$$\int_0^{t_1}\langle f(t),\ (T_t - T_t)z_0\rangle - \int_0^{t_1}\int_0^s \langle CT^*_{s-t}f(s),\ u(t)\rangle dtds + \int_0^{t_1}\langle f(t)\ ,\ A\int_0^t T_{t-s}Fu(s)ds\rangle dt\ .$$

Now

$$\int_0^{t_1}\langle f(t),\ A\int_0^t T_{t-s}Fu(s)ds\rangle dt = \int_0^{t_1}\langle A^*f(t),\ \int_0^t T_{t-s}Fu(s)ds\rangle dt$$

$$= \int_0^{t_1}\int_0^t \langle A^*f(t),\ T_{t-s}Fu(s)\rangle dsdt$$

since

$$\int_0^{t_1}\int_0^t |\langle A^*f(t),\ T_{t-s}Fu(s)\rangle| dsdt = \int_0^{t_1}\int_0^t |\langle f(t),\ AT_{t-s}Fu(s)\rangle| dsdt$$

$$\leq \int_0^{t_1}\int_0^t ||f(t)||\ ||AT_{t-s}F||\ ||u|| dsdt < \infty$$

since $||AT_{t-s}F|| \in L^2$ by assumption (8.26). So

$$\int_0^{t_1}\langle f(t),\ A\int_0^t T_{t-s}Fu(s)ds\rangle dt = \int_0^{t_1}\int_0^t \langle T^*_{t-s}A^*f(t),\ Fu(s)\rangle dsdt$$

$$= \int_0^{t_1}\int_0^t \langle A^*T^*_{t-s}f(t),\ Fu(s)\rangle dsdt$$

$$= \int_0^{t_1}\int_0^t \langle T^*_{t-s}f(t),\ \bar{A}Fu(s)\rangle dsdt - \int_0^{t_1}\int_0^t \langle CT^*_{t-s}f(t),\ u(s)\rangle dsdt$$

from the Green's formula with $Dz = u$, $Ez = 0$.

But, $\tilde{A}Fu(s) = \sigma Fu(s) = 0$, hence

$$\int_o^{t_1} <f(t), \tilde{A}\int_o^t T_{t-s}Fu(s)ds>dt = \int_o^{t_1}\int_o^t <CT^*_{s-t}f(t), u(s)>dsdt$$

and thus (8.27) is established.

From proposition 8.13 we can prove, in the same way as theorem 2.4, the following theorem.

Theorem 8.14

Suppose $z(t)$ is given by the Zabczyk formulation, with $Gu = 0$,

(8.28) $$z(t) = T_t z_o + \int_o^t T_{t-s}Fu(s)ds \quad ,$$

with (8.26) holding, then $z(t)$ is a weak solution of

$$\dot{z} = \tilde{A}z$$

$$Dz = u \quad , Ez = 0 \quad , z(o) = z_o \quad .$$

The linear Quadratic Cost Control Problem

Only Balakrishnan { 2 } and Barbu { 4 } consider the linear quadratic cost or a similar problem. Fattorini, Washburn and Zabczyk confine themselves to system formulation and controllability (with Washburn taking a brief look at the time optimal problem).

Balakrishnan considers the optimization problem of minimizing a quadratic functional of the form

(8.29) $$J(u) = \int_o^{t_1} <Kx(t), Kx(t)>dt + \int_o^{t_1} <u(t), Ru(t)>dt$$

where K is a linear bounded map from $H = L^2(\Omega)$ into another Hilbert space and $u(\cdot) \in L^2[o,t_1;U]$ with $U = L^2(\Gamma)$. Here Ω is a bounded domain in R^n

with boundary Γ . $x(t)$ is assumed to be given by

$$(8.30) \quad x(t) = T_t x_o - \int_o^t AT_{t-s}Fu(s)ds \quad \text{a.e.} \quad t > 0 \quad ,$$

i.e. $x(t)$ satisfies (8.8) where $\sigma Fu = 0$.

Balakrishnan also makes the following "smoothness" assumption on K , which implies A^*K^* is bounded.

Assumption 8.15

K^* maps into the domain of A^* .

Hence in addition to the semigroup being smoothing (he assumes A is analytic) he also requires K to be smoothing, though he leaves open to question whether this assumption is in fact necessary.

By exploiting the dual filtering problem he shows that under assumption 8.15 there exists a unique optimal control given by $u^*(t) = - F^*A^*P(t)x(t)$ where $P(t)$ satisfies a differential equation of the form

$$(8.31) \quad <\frac{d}{dt} P(t)x, y> + <P(t)Ay, x> + <P(t)x, Ay> - <F^*A^*P(t)x, F^*A^*P(t)y> + <Kx, Ky> = 0$$

$$P(t_1) = 0 \quad .$$

The optimal cost is

$$J(u^*) = <P(o)x(o), x(o)> \quad .$$

Note that (8.13) is the same equation as (4.25) with $K = M^{\frac{1}{2}}$, $G = 0$ and $R = I$ and since, recall, we can consider $F^*A^* = B^*$ in some sense.

The problem Barbu { 4 } considers is that of minimizing a functional with a convex rather than a quadratic integrand, i.e. he minimizes,

$$(8.32) \quad \int_o^{t_1} L(t,y,u)dt + \ell(y(o),\ y(t_1))$$

where y is given by (8.8) and where L and ℓ are lower semicontinuous functionals. Utilizing Gateaux differentiation he obtains necessary and sufficient conditions for a given control to be optimal in terms of the adjoint system. Balakrishnan {2} and we here, in section 4 of this thesis, give an explicit form for the optimal control in terms of the solution to a Riccati equation. Owing to the greater generality of his problem Barbu's result is less specific. He shows that a given control u^* is optimal, in the sense that it minimizes (8.32), if and only if there exist functionals p and q such that the following conditions are satisfied,

(8.33)

(a) $y^*(t)$ is the response to $u^*(t)$ via the system equation (8.8)

(b) $p(t)$ satisfies the adjoint equation

$$p(t) = T^*_{t_1-t}p(t_1) - \int_t^{t_1} T^*_{s-t}q(s)ds$$

with

(c) $(q(t),\ F^*A^*p(t)) \in \partial L(t,y^*(t),u^*(t))$

and

(d) $(p(o),\ -\ p(t_1)) \in \partial\ell(y^*(o),\ y^*(t_1))$.

In the above ∂L and $\partial\ell$ are the subdifferentials of $L(t)$ and ℓ respectively, a subdifferential being defined by

<u>Definition 8.16</u>

Let $\Phi : X \to \overline{R} =]-\infty, \infty]$ be a lower semicontinuous convex functional, then the subdifferential $\partial\Phi : X \to X^*$ is defined by

$$\partial\Phi(x) = \{x^* \in X^* ;\ \phi(x) - \phi(h) \leq \langle x^*,\ x - h\rangle \ \forall\ h \in X\}.$$

(c) Lions Approach

Lions {24} considers evolution systems in which the operator $A(t)$ is linked to a bilinear form $a(t;\phi,\psi)$ on a Hilbert space V.

If we let V and H be Hilbert spaces, V dense in H, such that H is identified with its dual, then, $V \subset H \subset V^*$. Suppose that the family of bilinear forms on V are such that

(a) $a(t;\phi,\psi)$ is measurable on $[o,t_1]$ for all ϕ, $\psi \in V$

(8.34) (b) $|a(t;\phi,\psi)| \leq C||\phi||_V\,||\psi||_V$

(c) $a(t;\phi,\phi) + \lambda||\phi||_H^2 \geq \alpha||\phi||_V$ for all $\phi \in V$, $t \in [o,t_1]$,

then for each t it is possible to write

$$a(t;\phi,\psi) = -\langle A(t)\phi,\ \psi\rangle_{V^*,V}$$

where V^*,V denotes the duality pairing between V^* and V.

Lions shows that for such an operator $A(t)$ there is a unique solution in $W(o,t_1)$ of

$$\frac{dz}{dt} = A(t)z \tag{8.35}$$

$$z(o) = z_o \in H,$$

where (8.35) is to be interpreted in the sense of distributions, and $W(o,t_1)$ is the Hilbert space

$$W(o,t_1) = \{x\ ;\ x \in L^2[o,t_1;V]\ ,\ \frac{dx}{dt} \in L^2[o,t_1;V^*]\}$$

with norm

$$||x||^2_{W(o,t_1)} = ||x||^2_{L^2[o,t_1;V]} + \left|\left|\frac{dx}{dt}\right|\right|^2_{L^2[o,t_1;V^*]} .$$

Moreover, the solution depends continuously on the initial data in the sense that the map $x_o \to x(\cdot)$ from $H \to W(o,t_1)$ is continuous.

In the case under consideration in this thesis, $A(t) = A$, independent of t, (8.35) becomes

(8.36) $\frac{dz}{dt} = Az$, $z(o) = z_o \in H$,

with unique solution $z(t)$ in $W(o,t_1)$ and we may then define a continuous linear map of $H \to H$, $T(t)$ by $z(t) = T(t)z_o$. The family of operators $T(t)$ then constitute a strongly continuous semigroup on H, with infinitesimal generator A with domain $D(A) = \{h \in V ; Ah \in H\}$.

The optimal control problem Lions considers is that of minimizing the quadratic form on the Hilbert spaces K and U

(8.37) $J(u) = ||Cy - z_d||^2_K + \langle Nu, u\rangle_U$

for systems of the form

(8.38) $\begin{cases} \dot{y}(t) = Ay(t) + f + Bu(t) \\ y(o) = y_o \end{cases}$

where $y \in L^2[o,t_1;V]$, $f \in L^2[o,t_1;V^*]$ and $y_o \in H$ are given, and $B \in \mathcal{L}(U, L^2[o,t_1;V^*])$. The observation is given by

$z(t) = Cy(t)$, $C \in \mathcal{L}(W(o,t_1),K)$

and $N \in \mathcal{L}(U,U)$ with $\langle Nu, u\rangle_U \geq k||u||^2_U$, $k > 0$.

By using variational inequalities and results on the minimizing of coercive forms Lions shows that there exists a unique minimizing control $u \in U$ of the form

(8,39) $u = - N^{-1}\Lambda_U^{-1}B^*p$

where p is given by the solution of the adjoint system

(8.40) $$-\frac{dp}{dt} = A^*p + C^*\Lambda(Cy - z_d)$$
$$p(t_1) = 0$$

and where Λ is the canonical isomorphism of K onto K^*, and Λ_U the canonical isomorphism of U onto U^*.

If we have the further condition holding

(8.41) $U = L^2[o,t_1;E]$, $K = L^2[o,t_1;F]$ with E and F separable Hilbert spaces, then Lions shows that the equations in y and p , (8.38) and (8.40) respectively, can be decoupled and that

(8.42) $$p(t) = P(t)y(t) + r(t)$$

where $P(t)$ satisfies a differential Riccati equation and $r(t)$ is determined by the solution of an (abstract) parabolic equation.

Also, when $f = 0 = z_d$ so that the cost functional becomes

(8.43) $$J(u) = \int_o^{t_1} ||Cy(t)||_F^2 dt + \int_o^{t_1} \langle Nu, u\rangle_E \, dt$$

then the optimal control is

$$J(u^*) = \langle p(o), y(o)\rangle = \langle P(o)y(o), y(o)\rangle$$

where P satisfies the Riccati equation

(8.44) $$\langle\frac{dP}{dt} x, y\rangle + \langle PAx, y\rangle + \langle Px, Ay\rangle + \langle Cx, Cy\rangle - \langle B^*Px, N^{-1}B^*Py\rangle = 0$$
$$x , y \in D(A) .$$

In the case where U , K are given by (8.41) Lions also investigates the infinite time problem and shows that now the optimal control is given by $u^* = - N^{-1}B^*P^\infty y$ with P^∞ independent of time.

<u>**Remarks**</u>

1. Lions shows that if A is given by $a(\phi,\psi)$ with

$$(8.37) \quad a(\phi,\phi) + \lambda||\phi||^2_H \geq \alpha||\phi||^2_V \quad \forall\, \phi \in V ,$$

i.e. the time independent version of (8.34)(c), then A is the infinitesimal generator of a strongly continuous semigroup, but, the converse does not hold. Thus Lions' approach is valid only for operators A given by such an $a(\cdot,\cdot)$ and not for all infinitesimal generators of semigroups.

2. Since $\int_o^{t_1} ||T_s z_o||^2_V ds < \infty$, $z(t)$ being the solution to (8.35) in $W(o,t_1)$, assumption (2.4) is satisfied with $p = 2$ if we take $V^* = \bar{W}$.

3. Note that from the above remark Lions systems satisfy our formulation also, and that by putting $G = 0$, $M^{\frac{1}{2}} = C$ and $R = N$ in (4.2) this cost functional is the same as the special case of Lions (8.44). For this cost functional, and its infinite time equivalent, the result Lions obtains is the same as ours in that the optimal control is of the form $u^* = - N^{-1}B^*P(t)y$ with $P(t)$ the solution to a Riccati equation, time independent for the infinite time problem.

4. An advantage of the approach taken here in this thesis to that of Lions is that not only is it possible to give the equations the optimal control must satisfy but also an iterative sequence which converges to the optimal control.

9. APPLICATIONS

The common areas of application of control theory are in the engineering field. Bensoussan et al., however, in their book { 5 } detail applications of finite dimensional control theory relevant to modern management. Here we concentrate mainly on some of the areas suggested in { 5 } to highlight the role distributed parameter control theory can also play in modern management.

The areas of application considered are inventory control, advertising, pollution control, traffic flow control and population control.

9.1. Optimum Inventory Control - Deteriorating Product

Here we consider the problem of manufacturing and storing a product which is in continuous production and whose state deteriorates over time. We assume there exists a continuous demand for the product $f(x,t)$, where x denotes the quality or state of deterioration and we wish to find the replenishment or production policy which maintains a given (quality dependent) stock level (the demand having been satisfied). Both deterministic and stochastic deterioration rates are considered.

Bensoussan, Nissen and Tapiero { 6 } solve these problems using a Lions type approach to give the equations the optimal policy must satisfy. The equations they give differ slightly from those found here by the semigroup approach indicating that perhaps an error has occurred in their work (not detailed in the paper).

The system model we consider is that as given by Bensoussan, Nissen and Tapiero { 6 } viz:-

Let $y(x,t)$ denote the amount of an item in stock at time t having a deterioration equal to x , $x \in [0,1]$. If μ is the deterioration rate, $f(x,t)$ the demand rate, then y satisfies the following partial differential equation

$$(9.1) \qquad \frac{\partial y(x,t)}{\partial t} + \mu \frac{\partial y(x,t)}{\partial x} + f(x,t) = 0 \qquad x \in (0,1) \quad .$$

The amount of a product at time t in state o , corresponding to replenished items, is a control variable $u(t)$, i.e.

$$(9.2) \qquad y(o,t) = u(t)$$

and the initial inventory at time $t = 0$ is

$$(9.3) \qquad y(x,o) = y_o(x) \quad .$$

If we let the replenishment rate incur a quadratic cost an optimum replenishment policy may be found by minimizing the following cost function over the planning time $[o,t_1]$,

$$(9.4) \qquad J(u;o,y_o) = \int_o^{t_1} \{ \int_o^1 \langle y(x,t) - y_d(x,t),\ y(x,t) - y_d(x,t) \rangle dx + \langle u(t),\ cu(t) \rangle \} dt$$

where y_d is the desired inventory level.

Suppose now that the deterioration rate is stochastic, i.e. let $\mu\Delta t + \sigma\Delta w(t)$ be the deterioration in a time interval $(t,t+\Delta t)$ where $w(t)$ is a standard Brownian motion. The quantity of a product with deterioration state x , at time t , is now a random variable denoted by $\tilde{y}(x,t)$. Assuming $\tilde{y}(x,t)$ is independent of $\Delta w(t)$ for any x , t we denote the expected value of $\tilde{y}(x,t)$ by $y(x,t)$ i.e.

$$(9.5) \qquad y(x,t) = E\{\tilde{y}(x,t)\}$$

and then y satisfies

$$(9.6) \qquad \frac{\partial y}{\partial t} + \mu\frac{\partial y}{\partial x} - \frac{1}{2}\sigma^2\frac{\partial^2 y}{\partial x^2} + f(x,t) = 0 \qquad x \in (0,1)$$

by using a Taylor expansion and deleting the terms in Δt of order greater than 1 . As in the deterministic case

$$(9.7) \qquad y(o,t) = u(t) \quad , \quad y(x,o) = y_o(x) \quad ,$$

but we need a further boundary condition such as

$$(9.8) \qquad y(1,t) = 0 \quad ,$$

i.e. items reaching the worst deterioration state $x = 1$ are rejected or

$$(9.9) \qquad \frac{\partial y}{\partial x}(1,t) = 0 \quad ,$$

i.e. items in quality state $x = 1$ can no longer deteriorate.

For obtaining an optimum replenishment policy we again consider the cost functional (9.4).

Solution of the Deterministic Problem

Here the problem is to minimize (9.4) where y satisfies (9.1) - (9.3) i.e.

$$(9.10) \qquad \frac{\partial y}{\partial t} + \mu\frac{\partial y}{\partial x} + f(x,t) = 0 \qquad x \in (0,1)$$

$$y(o,t) = u(t) \quad , \quad y(x,o) = y_o(x) \ .$$

Following our general theory of system formulation, as in section 2, we wish to determine a B such that a mild solution of (9.11) is a weak solution of (9.10) where

$$(9.11) \qquad y_t = -\mu y_x + Bu - f$$

$$y(o,t) = 0 \quad , \ y(x,o) = y_o(x) \quad .$$

Abstracting (9.11) we obtain

$$(9.12) \qquad \dot{y} = Ay - f + Bu \quad , \quad y(o) = y_o$$

and if we take $A = -\mu\frac{\partial}{\partial x}$ with $D(A) = \{y \in H;\ y_x \in H\ ,\ y(o) = 0\}$ where $H = L^2(0,1)$, then (see for example {16} p.530) A generates a strongly continuous semigroup on H given by

$$(T_t y)(x) = \begin{cases} y(x - \mu t) & 0 \leq x - \mu t \\ 0 & \text{else} \end{cases} .$$

Furthermore, $A^* = -A = \mu\frac{\partial}{\partial x}$ with $D(A^*) = \{y \in H\ ;\ y_x \in H\ ,\ y(1) = 0\}$, since then , for $y_1 \in D(A)$, $y_2 \in D(A^*)$

$$\langle Ay_1, y_2\rangle = \int_o^1 -\mu\frac{\partial y_1}{\partial y} y_2\, dx$$

$$= -\Big[\mu y_1 y_2\Big]_o^1 + \int_o^1 \mu y_1 \frac{\partial y_2}{\partial x}\, dx$$

$$= \langle y_1, \mu\frac{\partial y_2}{\partial x}\rangle = \langle y_1, A^* y_2\rangle .$$

Now, let $\tilde{A}$ be the same formal operator as A but defined on $D(\tilde{A}) = \{y \in H\ ;\ y_x \in H\}$ then from the above with $y_1 \in D(\tilde{A})$, $y_2 \in D(A^*)$

$$\langle \tilde{A}y_1, y_2\rangle = \mu y_1(o)y_2(o) + \langle y_1, A^* y_2\rangle \quad .$$

Thus comparing with the general Green's formula (2.7) i.e.

$$\langle \tilde{A}y_1, y_2\rangle = \langle y_1, A^* y_2\rangle + \langle Dy_1, Cy_2\rangle + \langle Ey_1, Gy_2\rangle$$

we have $Cy = \mu y(o)$, and so if $B^* = C$,

$$\langle Bu, y\rangle_H = \langle u, Cy\rangle_U$$

$$= \mu\langle u, y(o)\rangle$$

$$= \mu\int_o^1 u\delta(x)y(x)dx$$

$$= \mu\langle \delta(\cdot)u, y(\cdot)\rangle_H$$

i.e. $(Bu)(x) = \mu\delta(x)u$.

Now $A^* = \frac{\partial}{\partial x}$ on $D(A^*) = \{y \in H\ ;\ y_x \in H\ ,\ y(1) = 0\}$ generates the semigroup T_t^* where

$$(T_t^* y)(x) = \begin{cases} y(x + \mu t) & x + \mu t \leq 1 \\ 0 & \text{else} \end{cases} ,$$

which implies that

$$B^* T_t^* y = \begin{cases} B^* y(x + \mu t) & x + \mu t \leq 1 \\ 0 & \text{else} \end{cases} .$$

But $B^* y(x,t) = Cy(x,t) = \mu y(o,t)$ and so

$$B^* T_t^* y = \begin{cases} \mu y(\mu t) & \mu t \leq 1 \\ 0 & \text{else} \end{cases} .$$

Hence

$$\begin{aligned} ||B^* T_t^* y||_{L^2[o,t_1;U^*]} &= \int_o^{t_1} ||B^* T_t^* y||^2 dt \\ &= \int_o^{1/\mu} \mu^2 y^2(\mu t) dt \\ &= \int_o^1 \frac{\mu^2}{\mu} y^2(s) ds \qquad \text{by putting } \mu t = s \\ &= \mu ||y||_{L^2[0,1]} \end{aligned}$$

i.e.

$$||B^* T_t^* y||_{L^2[o,t_1;U^*]} \leq \mu ||y||_{L^2[0,1]}$$

thus

$$||T_t Bu||_{L^2[0,1]} \leq \mu ||u||_{L^2[o,t_1;U]}$$

and hence condition (2.4) holds with $p \geq 2$. Thus we can employ the general theory to the problem of minimizing (9.4) where $y(t)$ is given by (9.12), which has mild solution

$$(9.13) \qquad y(x,t) = T_t y_o + \int_o^t T_{t-s} Bu(s) ds - \int_o^t T_{t-s} f(s) ds \quad .$$

If we take

$$r(t) = y_d(t) + \int_o^t T_{t-s} f(s)ds \quad , \quad \bar{y}(t) = y(t) + \int_o^t T_{t-s} f(s)ds \quad .$$

then the problem is equivalent to minimizing

$$(9.14) \qquad J(u) = \int_o^{t_1}\int_o^1 \langle \bar{y}(t) - r(t), \bar{y}(t) - r(t)\rangle dxdt + c\int_o^{t_1} \langle u(t), u(t)\rangle dt$$

subject to

$$(9.15) \qquad \bar{y}(t) = T_t y_o + \int_o^t T_{t-s} Bu(s)ds \quad .$$

Thus the problem is in the general form for the tracking problem as given in section 5, where we take $G = 0$, $M = 0$ and $R = c$ in the cost functional 5.1. Hence we know there exists a unique optimal control $u^*(t) = - R^{-1}[B^*Q(t)\bar{y}(t) + B^*S(t)]$ where $Q(t)$ satisfies the differential equation

$$(9.16) \qquad \frac{d}{dt}\langle Q(t)h, k\rangle + \langle Q(t)h, Ak\rangle + \langle Ah, Q(t)k\rangle - \frac{1}{c}\langle B^*Q(t)h, B^*Q(t)k\rangle + \langle h, k\rangle = 0 \qquad \text{for } h , k \in D(A)$$

$$Q(t_1) = 0$$

and $S(t)$ satisfies

$$(9.17) \qquad \frac{d}{dt}\langle x, S(t)\rangle = - \langle Ax, S(t)\rangle + \frac{1}{c}\langle B^*Q(t)x, B^*S(t)\rangle + \langle x, r(t)\rangle \qquad x \in D(A)$$

$$S(t_1) = 0 \ .$$

If we assume that $Q(t)$ has the form

$$(9.18) \qquad (Q\bar{y})(\zeta) = \int_o^1 K(\zeta,\eta,t)\bar{y}(\eta)d\eta$$

then

$$\langle Qh, Ak\rangle = \int_o^1 (Qh)(\zeta)Ak(\zeta)d\zeta$$

$$= \int_o^1 \{\int_o^1 K(\zeta,\eta,t)h(\eta)d\eta\}\{- \mu\frac{\partial k(\zeta)}{\partial\zeta}\}d\zeta$$

$$= \left(- \mu\int_o^1 K(\zeta,\eta,t)h(\eta)k(\zeta)d\eta\right)_o^1 + \mu\int_o^1\int_o^1 K_\zeta(\zeta,\eta,t)h(\eta)k(\zeta)d\eta d\zeta$$

$$= \int_o^1\int_o^1 K_\zeta(\zeta,\eta,t)h(\eta)k(\zeta)d\eta d\zeta$$

when $K(1,\eta,t) = 0$, since $h \in D(A)$ implies $h(o) = 0$. Similarly,

$$\langle Ah, Qk\rangle = \left[-\int_o^1 K(\zeta,\eta,t)k(\zeta)h(\eta)d\eta\right]_0^1 + \mu\int_o^1\int_o^1 K_\eta(\zeta,\eta,t)h(\eta)k(\zeta)d\eta d\zeta$$

$$= \mu\int_o^1\int_o^1 K_\eta(\zeta,\eta,t)h(\eta)k(\zeta)d\eta d\zeta$$

when $K(\zeta,1,t) = 0$. Also

$$\frac{d}{dt}\langle Q(t)h, k\rangle = \int_o^1\int_o^1 K_t(\zeta,\eta,t)h(\eta)k(\zeta)d\eta d\zeta$$

and for the term $-\frac{1}{c}\langle B^*Q(t)h, B^*Q(t)k\rangle$,

$$\langle Bu, Qk\rangle = \int_o^1 (Bu)(\zeta)(Qk)(\zeta)d\zeta$$

$$= \int_o^1 \mu\delta(\zeta)(Qk)(\zeta)d\zeta\cdot u$$

$$= \mu\cdot u\cdot[(Qk)(\zeta)]_{\zeta=0}$$

So

$$\langle u, B^*Qk\rangle_U = \langle Bu, Qk\rangle_H = \mu\cdot u\int_o^1 K(o,\eta,t)k(\eta)d\eta$$

i.e.

$$B^*Qk = \mu\int_o^1 K(o,\eta,t)k(\eta)d\eta$$

and thus

$$-\frac{1}{c}\langle B^*Qh, B^*Qk\rangle = -\frac{1}{c}\langle\mu\int_o^1 K(o,\eta,t)h(\eta)d\eta,\ \mu\int_o^1 K(\zeta,o,t)k(\zeta)d\zeta\rangle$$

$$= -\frac{\mu^2}{c}\int_o^1\int_o^1 K(o,\eta,t)K(\zeta,o,t)h(\eta)k(\zeta)d\eta d\zeta \quad .$$

Finally

$$\langle h, k\rangle = \int_o^1 h(\zeta)k(\zeta)d\zeta = \int_o^1\int_o^1 \delta(\zeta-\eta)h(\eta)k(\zeta)d\eta d\zeta \quad .$$

Thus substituting for $Q(t)$ given by (9.18), with $K(1,\eta,t) = 0 = K(\zeta,1,t)$, in the differential equation (9.16) gives

$$\int_o^1\int_o^1 \{K_t(\zeta,\eta,t) + K_\zeta(\zeta,\eta,t) + K_\eta(\zeta,\eta,t) - \frac{\mu^2}{c} K(o,\eta,t)K(\zeta,o,t) + \delta(\zeta-\eta)\}h(\eta)k(\zeta)d\eta d\zeta = 0$$

which is satisfied if we chose K to satisfy

$$K_t(\zeta,\eta,t) + K_\zeta(\zeta,\eta,t) + K_\eta(\zeta,\eta,t) - \frac{\mu^2}{c} K(o,\eta,t)K(\zeta,o,t) + \delta(\zeta-\eta) = 0 .$$

Similarly, if we assume

(9.19) $S(t) = S(\zeta,t)$

then

$$\frac{d}{dt} \langle x, S(t)\rangle = \int_o^1 S_t(\zeta,t)x(\zeta)d\zeta ,$$

$$\langle Ax, S(t)\rangle = \int_o^1 - \mu\frac{\partial x(\zeta)}{\partial\zeta} S(\zeta,t)d\zeta$$

$$= \Big[- \mu x(\zeta)S(\zeta,t)\Big]_o^1 + \int_o^1 \mu x(\zeta)S_\zeta(\zeta,t)d\zeta$$

$$= \int_o^1 \mu x(\zeta)S_\zeta(\zeta,t)d\zeta$$

when $S(1,t) = 0$, as $x \in D(A)$ implies $x(o) = 0$. Also

$$\langle B^*Qx, B^*S\rangle = \frac{1}{c} \langle\mu\int_o^1 K(\zeta,o,t)x(\zeta)d\zeta, B^*S\rangle$$

and B^*S is given by

$$\langle Bu, S\rangle = \int_o^1 (Bu)(\zeta)S(\zeta,t)d\zeta$$

$$= \int_o^1 \mu\delta(\zeta)u\cdot S(\zeta,t)d\zeta$$

$$= \mu u S(o,t) ,$$

i.e. $\langle u, B^*S\rangle = \langle u, \mu S(o,t)\rangle$.

Hence,

$$\frac{1}{c}\langle B^*Qx, B^*S\rangle = \frac{\mu^2}{c}\int_o^1 K(\zeta,o,t)x(\zeta)d\zeta\cdot S(o,t) \ .$$

Finally

$$\langle x, r\rangle = \int_o^1 x(\zeta)r(\zeta)d\zeta \ .$$

Thus substituting into the differential equation (9.17) for $S(t)$ given by (9.19) with $S(1,t) = 0$, we find

$$\int_0^1 \{S_t(\zeta,t) - \frac{\mu^2}{c}K(\zeta,o,t)S(o,t) + S_\zeta(\zeta,t) - r(\zeta)\}x(\zeta)d\zeta = 0$$

which is satisfied if we take S to satisfy

$$S_t(\zeta,t) + S_\zeta(\zeta,t) - \frac{\mu^2}{c}K(\zeta,o,t)S(o,t) - r(\zeta) = 0 \ ,$$

i.e.

$$S_t(\zeta,t) + S_\zeta(\zeta,t) - \frac{\mu^2}{c}K(\zeta,o,t)S(o,t) - y_d(\zeta,t) - \int_o^t T_{t-s}f(\zeta,s)ds = 0$$

since $r(t) = y_d(t) + \int_o^t T_{t-s}f(s)ds$.

We have therefore shown that the optimal control u^* can be written as

$$u^*(t) = -\frac{1}{c}\mu\left(\int_o^1 K(o,\eta,t)\bar{y}(\eta)d\eta + S(o,t)\right)$$

$$= -\frac{1}{c}\mu\left(\int_o^1 K(o,\eta,t)y(\eta)d\eta + \int_o^1 K(o,\eta,t)\int_o^t T_{t-s}f(\eta,s)dsd\eta + S(o,t)\right)$$

$$= -\frac{1}{c}\mu\left(\int_o^1 K(o,\eta,t)y(\eta)d\eta + \int_o^1 K(o,\eta,t)\int_o^t f(\eta-\mu(t-s),s)dsd\eta + S(o,t)\right) ,$$

since $T_{t-s}y(\alpha) = y(\alpha-\mu(t-s))$ so that $\int_o^t T_{t-s}f(\alpha,s)ds = \int_o^t f(\alpha-\mu(t-s),s)ds$,

where K and S satisfy

$$K_t(\zeta,\eta,t) + \mu K_\zeta(\zeta,\eta,t) + \mu K_\eta(\zeta,\eta,t) - \frac{\mu^2}{c} K(o,\eta,t)K(\zeta,o,t) + \delta(\zeta-\eta) = 0$$

$$K(1,\eta,t) = K(\zeta,1,t) = 0$$

and

$$S_t(\zeta,t) - \mu S_\zeta(\zeta,t) - \frac{\mu^2}{c} K(\zeta,o,t)S(o,t) - y_d(\zeta,t) - \int_o^t f(\zeta-\mu(t-s),s)ds = 0$$

$$S(1,t) = 0 ,$$

respectively.

<u>Remark</u>

Bensoussan, Nissen and Tapiero {5}, using a Lions type approach, obtain the solution for the optimal control as being

$$u^* = \frac{1}{c}\int_o^1 K(o,\eta,t)z(\eta)d\eta + \frac{1}{c} S(o,t)$$

where K satisfies

$$K_t(\zeta,\eta,t) + \mu K_\zeta(\zeta,\eta,t) + \mu K_\eta(\zeta,\eta,t) + \frac{\mu^2}{c} K(\zeta,o,t)K(o,\eta,t) = \delta(\zeta-\eta)$$

$$K(\zeta,1,t) = 0$$

and S satisfies

$$S_t(\zeta,t) + \mu S_\zeta(\zeta,t) + y_d(\zeta,t) + \frac{\mu^2}{c} K(\zeta,o,t)S(o,t) - \int_o^1 K(\zeta,\eta,t)f(\zeta,t)d\zeta = 0 .$$

<u>Solution of the Stochastic Problem</u>

Recall that in this case the problem is to minimize (9.4) subject to $y(t)$ satisfying

(9.20) $$\frac{\partial y}{\partial t} + \mu\frac{\partial y}{\partial x} - \frac{1}{2}\sigma^2 \frac{\partial^2 y}{\partial x^2} + f(x,t) = 0$$

$$y(o,t) = u(t) \quad , \quad y(1,t) = 0 \quad , \quad y(x,o) = y_o(x) .$$

In order to facilitate obtaining the estimates required to show that (2.4) holds we rewrite the system in terms of a self-adjoint operator. To do this let

$$(9.21)\qquad z(x,t) = e^{-(\mu/\sigma^2)x}y(x,t) \quad \text{with} \quad z_o(x) = e^{-(\mu/\sigma^2)x}y_o(x)$$

then (9.20) is equivalent to

$$(9.22)\qquad z_t = \frac{1}{2}\sigma^2 z_{xx} - \frac{\mu^2}{2\sigma^2} z - fe^{-(\mu/\sigma^2)x}$$

$$z(x,o) = z_o(x) \quad , \quad z(o,t) = v(t) \quad , \quad z(1,t) = 0 \quad .$$

To employ the general theory we set $H = L^2(0,1)$, $U = \mathcal{R}$ and $A = \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2} - \frac{\mu^2}{2\sigma^2}$ with $D(A) = \{y \in H \text{ , } y_{xx} \in H \text{ , } y(o) = y(1) = 0\}$

and so A generates the semigroup T_t given by

$$T_t = 2\sum_{n=1}^{\infty} e^{-c_n t}\langle\cdot, \sin n\pi(\cdot)\rangle \sin n\pi x$$

with $c_n = \frac{\mu^2}{2\sigma^2} + \frac{n^2\pi^2\sigma^2}{2}$. Furthermore A is self-adjoint ,i.e. $A^* = A$ with $D(A^*) = D(A)$. Now taking $\tilde{A}$ as the same formal operator as A but with $D(\tilde{A}) = \{y \in H \text{ , } y_{xx} \in H \text{ , } y(1) = 0\}$ then for $y_1 \in D(\tilde{A})$, $y_2 \in D(A^*)$ we have

$$\langle Ay_1, y_2\rangle = \langle y_1, A^*y_2\rangle + \frac{1}{2}\sigma^2 y_1(o)\frac{\partial y_2}{\partial x}(o)$$

which comparing with the general Green's formula (2.7) implies $B^*y = Cy = \frac{1}{2}\sigma^2\frac{\partial y}{\partial x}(o)$ and so, as in example 2.7, $B = -\frac{1}{2}\sigma^2\delta'(x)$. Hence

$$T_tB\,v = 2\sum_{n=1}^{\infty} e^{-c_n t}\langle (Bv)(x), \sin n\pi x\rangle \sin n\pi x$$

$$= 2\sum_{n=1}^{\infty} e^{-c_n t} v\,\frac{1}{2}\sigma^2 n\pi \sin n\pi x$$

and so

$$||T_tBv||_H^2 = \langle T_tBv, T_tBv\rangle$$

$$= \langle 2\sum_{n=1}^{\infty} e^{-c_n t}\frac{1}{2}\sigma^2 n\pi \sin n\pi x\cdot v, 2\sum_{n=1}^{\infty} e^{-c_n t}\frac{1}{2}\sigma^2 m\pi \sin m\pi x\cdot v\rangle$$

$$\leq \frac{1}{2} \sum_{n=1}^{\infty} e^{-n^2\pi^2\sigma^2 t} n^2\pi^2\sigma^4 ||v||^2_U$$

$$\leq \frac{K}{t} ||v||^2_U$$

thus $||T_t Bv||_H \leq \frac{\bar{K}}{t^{1/2}} ||v||_U$.

Condition (2.4) is therefore satisfied, this time with $2 > p \geq 1$, where we take $\tilde{W} = (H^{3/2 + \epsilon}(0,1))^*$ and $g = \frac{M}{t^{1/2 + \epsilon/2}}$. Furthermore, the mild solution to

(9.23) $$z_t = Az + Bv - fe^{-(\mu/\sigma^2)x}$$
$$z(x,o) = z_o(x) \quad , \quad z(o,t) = 0 = z(1,t)$$

i.e.

(9.24) $$z(t) = T_t z_o + \int_o^t T_{t-s} Bv(s)ds - \int_o^t T_{t-s} f(s)e^{-(\mu/\sigma^2)x} ds$$

is a weak solution to (9.22), and thus $e^{(\mu/\sigma^2)x} z(t)$ is a weak solution to (9.20).

Similarly abstracting (9.20) and using the Green's formula (2.7) it is easy to see that we can write (9.20) in the form

(9.25) $$y_t = A'y + B'u - f$$
$$y(o,t) = 0 = y(1,t) \quad , \quad y(x,o) = y_o(x)$$

where $A' = \frac{1}{2}\frac{\partial^2}{\partial x^2} - \mu\frac{\partial}{\partial x}$, and where $B' = -\frac{1}{2}\sigma^2\delta'(x)$. Furthermore the mild solution to (9.25) is a weak solution to (9.20). From (9.25), using the fact that $z = e^{-(\mu/\sigma^2)x}y$ and $Az = e^{-(\mu/\sigma^2)x}A'y$ we have

$$z_t = e^{-(\mu/\sigma^2)x} y_t$$

$$= e^{-(\mu/\sigma^2)x} A'y + e^{-(\mu/\sigma^2)x} B'u - e^{-(\mu/\sigma^2)x} f$$

$$= Az + e^{-(\mu/\sigma^2)x} B'u - e^{-(\mu/\sigma^2)x} f$$

which comparing with (9.23) gives $v = e^{-(\mu/\sigma^2)x}u$ since $B' = B$.

The original problem is to minimize the cost functional (9.4) with y given by (9.20). With $z = e^{-(\mu/\sigma^2)x}y$,

$$y - y_d = e^{(\mu/\sigma^2)x}z - y_d = e^{(\mu/\sigma^2)x}(z - z_d)$$

if we let $z_d = e^{-(\mu/\sigma^2)x}y_d$, and recall $v = e^{-(\mu/\sigma^2)x}u$, so the problem is equivalent to minimizing

$$(9.26) \qquad J(v) = \int_0^{t_1}\int_0^1 \langle z(x,t) - z_d(x,t),\ e^{2(\mu/\sigma^2)x}[z(x,t) - z_d(x,t)]\rangle dx dt + \int_0^{t_1} \langle v(t),\ ce^{2(\mu/\sigma^2)x}v(t)\rangle dt$$

with $z(t)$ given by (9.24).

If we now let

$$(9.27) \qquad r(t) = z_d + \int_0^t T_{t-s}e^{-(\mu/\sigma^2)x}f(s)ds$$

then the problem is equivalent to minimizing

$$(9.28) \qquad J(v) = \int_0^{t_1}\int_0^1 \langle \bar{z}(t) - r(t),\ e^{2(\mu/\sigma^2)x}[\bar{z}(t) - r(t)]\rangle dx dt + \int_0^{t_1} \langle v(t),\ ce^{2(\mu/\sigma^2)x}v(t)\rangle dt$$

where

$$(9.29) \qquad \bar{z}(t) = T_t z_o + \int_0^t T_{t-s}Bv(s)ds$$

so then

$$\bar{z}(t) = z(t) + \int_0^t T_{t-s}e^{-(\mu/\sigma^2)x}f(s)ds \quad .$$

Since we have shown that T_t , B satisfy (2.4) with $2 > p \geq 1$ we can use the general theory of the tracking problem, as given in section 5, taking $G = 0$, $R = ce^{2(\mu/\sigma^2)x}$ and $M = e^{2(\mu/\sigma^2)x}$ in the cost functional (5.1). Hence we know there exists a unique optimizing control v^* which minimizes (9.28) and furthermore $v^* = -R^{-1}[B^*Q\bar{z} + B^*S]$ where Q satisfies (4.25) with $Q(t_1) = 0$, and S satisfies (5.17) with M and R as above.

If we now assume we can write Q in the form

$$(9.30) \qquad (Q\bar{z})(\zeta) = \int_o^1 K(\zeta,\eta,t)\bar{z}(\eta)d\eta$$

then, as for the deterministic case, we can substitute for this Q into (4.25) and then K satisfies, for h, $k \in D(A)$

$$(9.31) \qquad \int_o^1\int_o^1\Big(K_t(\zeta,\eta,t) + \frac{\sigma^2}{2}\{K_{\zeta\zeta}(\zeta,\eta,t) + K_{\eta\eta}(\zeta,\eta,t)\} - \frac{\mu^2}{\sigma^2}K(\zeta,\eta,t)$$
$$- \frac{\sigma^4}{4c}\{K_\zeta(o,\eta,t) - \frac{\mu}{\sigma^2}K(o,\eta,t)\}\{K_\zeta(\zeta,o,t) - \frac{\mu}{\sigma^2}K(\zeta,o,t)\}$$
$$+ \delta(\zeta-\eta)e^{(\mu/\sigma^2)(\zeta+\eta)}\Big)h(\eta)k(\zeta)d\eta d\zeta = 0,$$

when $K(o,\eta,t) = K(1,\eta,t) = K(\zeta,o,t) = K(\zeta,1,t) = 0$,
which is satisfied by $K(\zeta,\eta,t)$ if

$$(9.32) \qquad K_t(\zeta,\eta,t) + \frac{\sigma^2}{2}\{K_{\zeta\zeta}(\zeta,\eta,t) + K_{\eta\eta}(\zeta,\eta,t)\} - \frac{\mu^2}{\sigma^2}K(\zeta,\eta,t) + \delta(\zeta-\eta)e^{(\mu/\sigma^2)(\zeta+\eta)}$$
$$- \frac{\sigma^4}{4c}\{K_\zeta(o,\eta,t) - \frac{\mu}{\sigma^2}K(o,\eta,t)\}\{K_\eta(\zeta,o,t) - \frac{\mu}{\sigma^2}K(\eta,o,t)\} = 0.$$

Similarly if we set $S(t) = S(\zeta,t)$ and substitute into equation (5.17) then we find that S satisfies

$$(9.33) \qquad \int_o^1\Big(S_t(\zeta,t) + \frac{1}{2}\sigma^2 S_{\zeta\zeta}(\zeta,t) - \frac{\mu^2}{2\sigma^2}S(\zeta,t) + e^{2(\mu/\sigma^2)\zeta}r(\zeta)$$
$$- \frac{\sigma^4}{4c}\{K_\eta(\zeta,o,t) - \frac{\mu}{\sigma^2}K(\zeta,o,t)\}\{S_\zeta(o,t) - \frac{\mu}{\sigma^2}S(o,t)\}\Big)h(\zeta)d\zeta = 0,$$

when $S(o,t) = S(1,t) = 0$,
which is satisfied if S satisfies

$$(9.34)\quad S_t(\zeta,t) + \frac{\sigma^2}{2} S_{\zeta\zeta}(\zeta,t) - \frac{\mu^2}{2\sigma^2} S(\zeta,t) + e^{2(\mu/\sigma^2)\zeta} r(\zeta)$$

$$- \frac{\sigma^4}{4c} \{K_\eta(\zeta,o,t) - \frac{\mu}{\sigma^2} K(\zeta,o,t)\}\{S_\zeta(o,t) - \frac{\mu}{\sigma^2} S(o,t)\} = 0 .$$

Thus we have shown that the optimal control for (9.28) - (9.29) is

$$v^*(t) = - R^{-1}[B^*(Q(t)\bar{z}(t))(\zeta) + B^*S(t)(\zeta)]$$

$$= - \frac{1}{c} e^{-2(\mu/\sigma^2)\zeta}\{\int_o^1 K_\zeta(o,\eta,t)\bar{z}(t)d\eta + S_\zeta(o,t)\}$$

where $K(\zeta,\eta,t)$ and $S(\zeta,t)$ satisfy (9.32) and (9.34) respectively, and so the optimal control for (9.26) - (9.24) is

$$(9.35)\quad v^*(t) = - \frac{1}{c} e^{-(\mu/\sigma^2)\zeta}\{\int_o^1 K_\zeta(o,\eta,t)\{z(\eta) + \int_o^t T_{t-s}f(\eta,s)ds\}d\eta + S_\zeta(o,t)\} .$$

Furthermore, since the original problem (9.4) - (9.20) is equivalent to (9.26) - (9.24) with $u^* = e^{(\mu/\sigma^2)x}v^*$ and $y = e^{(\mu/\sigma^2)x}z$, the optimal control for (9.4) - (9.20) is

$$(9.36)\quad u^*(t) = - \frac{1}{c} e^{-(\mu/\sigma^2)x}B^*[Q(t)e^{-(\mu/\sigma^2)x}\{y(t) + \int_o^t T_{t-s}f(s)ds\}(x) + S(t)(x)] .$$

Recall that Q has the form (9.30) i.e.

$$(Qz)(\zeta) = \int_o^1 K(\zeta,\eta,t)z(\eta)d\eta \quad .$$

If we now let

$$L(\zeta,\eta,t) = K(\zeta,\eta,t)e^{-(\mu/\sigma^2)\zeta}e^{-(\mu/\sigma^2)\eta}$$

and set

$$(9.37)\quad (Py)(\zeta) = \int_o^1 L(\zeta,\eta,t)y(\eta)d\eta$$

and similarly set

$$(9.38)\quad N(\zeta,t) = S(\zeta,t)e^{-(\mu/\sigma^2)\zeta}$$

then $u^* = -\frac{1}{c} B^*[Py + N]$

with P , N given by (9.37) and (9.38) respectively. Since K satisfies (9.32), substituting for $K = L\, e^{(\mu/\sigma^2)\zeta} e^{(\mu/\sigma^2)\eta}$, L satisfies

$$(9.39)\quad \int_o^1\int_o^1 \Big[L_t(\zeta,\eta,t) + \frac{\sigma^2}{2}\{L_{\zeta\zeta}(\zeta,\eta,t) + L_{\eta\eta}(\zeta,\eta,t)\} + \mu\{L_\zeta(\zeta,\eta,t) + L_\eta(\zeta,\eta,t)\} - \frac{\sigma^4}{4c} L_\zeta(o,\eta,t)L_\eta(\zeta,o,t) + \delta(\zeta-\eta)\Big] h(\eta)e^{(\mu/\sigma^2)\eta}k(\zeta)e^{(\mu/\sigma^2)\zeta}d\eta d\zeta$$

which is satisfied by an $L(\zeta,\eta,t)$ such that

$$(9.40)\quad L_t(\zeta,\eta,t) + \frac{\sigma^2}{2}\{L_{\zeta\zeta}(\zeta,\eta,t) + L_{\eta\eta}(\zeta,\eta,t)\} + \mu\{L_\zeta(\zeta,\eta,t) + L_\eta(\zeta,\eta,t)\} - \frac{\sigma^4}{4c} L_\zeta(o,\eta,t)L_\eta(\zeta,o,t) + \delta(\zeta-\eta) = 0$$

and from the conditions on K we require

$$(9.41)\quad L(o,\eta,t) = L(1,\eta,t) = L(\zeta,o,t) = L(\zeta,1,t) = 0 .$$

Similarly substitution for $S(\zeta,t) = N(\zeta,t)e^{(\mu/\sigma^2)\zeta}$ in (9.34) leads to

$$(9.42)\quad \int_o^1\int_o^1 \Big[N_t(\zeta,t) + \frac{\sigma^2}{2} N_{\zeta\zeta}(\zeta,t) + \mu N_\zeta(\zeta,t) - \frac{\sigma^4}{4c} L_\eta(\zeta,o,t)N_\zeta(o,t) + e^{(\mu/\sigma^2)\zeta}r(\zeta)\Big] e^{(\mu/\sigma^2)\zeta}h(\zeta)d\zeta = 0$$

with $N(o,t) = N(1,t) = 0$,

which is satisfied by an N such that

$$(9.43)\quad N_t(\zeta,t) + \frac{\sigma^2}{2} N_{\zeta\zeta}(\zeta,t) - \frac{\sigma^4}{4c} L_\eta(\zeta,o,t)N_\zeta(o,t) + \mu N_\zeta(\zeta,t) + e^{(\mu/\sigma^2)\zeta}r(\zeta) = 0 ,$$

$$N(o,t) = N(1,t) = 0 .$$

But, recall,

$$e^{(\mu/\sigma^2)\zeta}r(\zeta) = e^{(\mu/\sigma^2)\zeta}\{z_d + \int_o^{t_1} T_{t-s}e^{-(\mu/\sigma^2)\zeta}f(s)ds\}$$

$$= y_d + \int_0^t {}^1T_{t-s} f(s)ds$$

$$= \bar{y}_d \quad .$$

Hence the optimal policy for minimizing (9.4) is $u^* = -\frac{1}{c} B^*[Py + N]$ with P of the form (9.37), where L satisfies (9.40) - (9.41), and where N satisfies (9.43). Thus, since $B^*y = \frac{1}{2}\sigma^2 \frac{\partial y}{\partial \zeta}(o)$, u^* has the form

$$(9.44) \qquad u^* = -\frac{1}{c}\{\int_0^1 L_\zeta(o,\eta,t)[y(\eta) + \int_0^t T_{t-s} f(\eta,s)ds]d\eta + N_\zeta(o,t)\} \ .$$

We can again compare these results with those given by Bensoussan, Nissen and Tapiero { 6 } who gave the optimal control as

$$(9.45) \qquad u(t) = \frac{1}{c}\int_0^1 P(o,\zeta,t)y(\zeta,t)d\zeta + \frac{1}{c} r(o,t)$$

with P satisfying

$$(9.46) \qquad P_t(x,\zeta,t) + \mu\{P_x(x,\zeta,t) + P_\zeta(x,\zeta,t)\} + \frac{\sigma^2}{2}\{P_{xx}(x,\zeta,t) + P_{\zeta\zeta}(x,\zeta,t)\}$$
$$- \frac{\sigma^2}{2c}\{P_x(o,\zeta,t) + P_\zeta(x,o,t)\} = \delta(x-\zeta)$$

and r satisfying

$$(9.47) \qquad r_t(x,t) + \mu r_x(x,t) + \frac{\mu^2}{c} P(x,o,t)r(o,t) - \int_0^1 P(x,\zeta,t)f(\zeta,t)d\zeta = 0$$

as before. This solution again differs from the one obtained here.

If we express K in terms of the basis for the semigroup T_t , i.e. let

$$(9.48) \qquad K(\zeta,\eta,t) = 2 \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{mn}(t) \sin m\pi\zeta \sin n\pi\eta$$

so that

$$(9.49) \qquad L(\zeta,\eta,t) = 2 \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{mn}(t)\, e^{-(\mu/\sigma^2)\zeta} e^{-(\mu/\sigma^2)\eta} \sin m\pi\zeta \sin n\pi\eta$$

then substituting into the differential equation for K , (9.32), or for L , (9.40), leads to the following equation for the coefficients a_{mn}

(9.50) $$\dot{a}_{mn} - \{\frac{1}{2}\sigma^2(m^2\pi^2 + n^2\pi^2) + \frac{\mu^2}{\sigma^2}\} a_{mn} - \frac{\sigma^4}{4c}\sum_{i=0}^{\infty} i\pi a_{in} \sum_{j=0}^{\infty} j\pi a_{mj} + \delta_m^n = 0$$

and similarly, if we write

$$N(\zeta,t) = \sqrt{2}\sum_{n=0}^{\infty} b_n(t)\sin n\pi\zeta\, e^{-(\mu/\sigma^2)\zeta}$$

i.e.

(9.51) $$S(\zeta,t) = \sqrt{2}\sum_{n=0}^{\infty} b_n(t)\sin n\pi\zeta$$

then substituting for N in (9.43), or S in (9.33), gives

(9.52) $$\dot{b}_n - \{n^2\pi^2 - \frac{\mu^2}{2\sigma^2}\}b_n - \frac{\sigma^4}{4c}\sum_{i=0}^{\infty} i\pi a_{ni} \sum_{j=0}^{\infty} j\pi b_j - \bar{y}_d^n = 0$$

as the equation which the b_j's must satisfy, where

$$\bar{y}_d = y_d + \int_o^{t_1} T_{t-s} f(s)ds$$

and $\bar{y}_d^n$ is given by

$$\bar{y}_d = \sqrt{2}\sum_{n=0}^{\infty} \bar{y}_d^n \sin n\pi\zeta \quad .$$

9.2. Optimum Advertising Policy - With Either Deterministic or Stochastic Rate of Loss of Goodwill.

Here we examine the optimium advertising policy, to achieve a given sales profile, assuming that sales are directly related to goodwill. Advertising is directed solely at those who have no initial knowledge of the product, and thus minimum goodwill towards it.

If we denote by $y(x,t)$ the number of people at time t with goodwill x towards a product, $x \in (0,1)$ where 0 and 1 are the minimum and maximum goodwill states, and assume the loss of goodwill by forgetting is a at a rate $\bar{\mu}$, then

(9.53) $\quad y(x,t+\Delta t) = y(x+\bar{\mu}\Delta t,t)$.

For Δt small enough and $1 \geq x > 0$, a Taylor series expansion of $y(x+\bar{\mu}\Delta t,t)$ yields

(9.54) $\quad y(x+\bar{\mu}\Delta t,t) = y(x,t) + \bar{\mu}\Delta t \frac{\partial y}{\partial x}$.

Substituting (9.54) into (9.53) and taking the limit as $\Delta t \to 0$ gives

$$\lim_{\Delta t \to 0} \left(\frac{y(x,t+\Delta t) - y(x,t)}{t} \right) = \bar{\mu} \frac{\partial y}{\partial x}$$

i.e.

(9.55) $\quad \frac{\partial y}{\partial t} = \bar{\mu} \frac{\partial y}{\partial x}$.

Let the initial goodwill profile, i.e. the number of people with goodwill x at time $t = 0$, be given by

(9.56) $\quad y(x,o) = y_o(x)$.

We suppose now that advertising is directed towards those who have no previous knowledge of the product, and thus have minimum (zero) goodwill towards it. If we assume that 1 unit of expenditure on advertising results in $B(x)$ people with goodwill x , $x \in (0,1)$, then we can take the total expenditure $u(t)$ as the control variable and so the evolution of y is given by

(9.57) $\quad \frac{\partial y}{\partial t} = \bar{\mu} \frac{\partial y}{\partial x} + B(x)u(t)$.

The problem is thus one of distributed control and so falls within the scope of the book by Curtain and Pritchard {11}.

If however we assume that 1 unit of expenditure on advertising results in giving F people (with no previous goodwill towards the product) a goodwill of x_f , for some fixed x_f, then we have an unbounded control problem with

(9.58) $\quad [y(x,t)]_{x_f} = Fu(t)$.

We also have that

(9.59) $\quad y(1,t) = 0 \qquad t \in (o,t_1]$

since people forget and no people with maximum goodwill are being created.

In the above we assumed that the rate of loss of goodwill, due to forgetting, was deterministic. If we now let $\bar{\mu}\Delta t + \sigma\Delta\omega(t)$ be the rate of loss of goodwill in the time interval $(t,t+\Delta t)$, where $\omega(t)$ is a standard Brownian motion, then the number of people with goodwill x at time t is now a random variable which we denote by $\tilde{y}(x,t)$. If we also assume that $\tilde{y}(x,t)$ is independent of $\Delta\omega(t)$ for any x , t and set its expected value as $E\{\tilde{y}(x,t)\} = y(x,t)$ then we have, analagous to (9.53)

$$\tilde{y}(x,t+\Delta t) = \tilde{y}(x+\bar{\mu}\Delta t+\sigma\Delta\omega(t),t) \quad .$$

Taking expected values we obtain

$$y(x,t+\Delta t) = E\{\tilde{y}(x+\bar{\mu}\Delta t+\sigma\Delta\omega(t),t) \quad .$$

Using the assumption that $\tilde{y}$ is independent of $\Delta\omega(t)$ we can replace $\tilde{y}$ by y , and so replacing $y(x+\bar{\mu}\Delta t+\sigma\Delta\omega(t),t)$ by a Taylor series expansion and deleting the terms in Δt of order higher that 1 , we find that

(9.60) $$\frac{\partial y}{\partial t} = \bar{\mu}\,\frac{\partial y}{\partial t} + \frac{1}{2}\,\sigma^2\,\frac{\partial^2 y}{\partial x^2}$$

and again we take the initial goodwill as

(9.61) $\quad y(x,o) = y_o(x)$.

As in the deterministic case we take

(9.62) $\quad [y(x,t)]_{x_f} = Fu(t) \quad , \quad y(1,t) = 0$

and we need a further boundary condition, such as

$$(9.63) \qquad \frac{\partial y}{\partial x}(o,t) = 0$$

reflecting the fact that people with no goodwill cannot forget.

If we assume that sales are directly related to goodwill then we may wish to achieve some desired final goodwill profile $y_d(x,t_1)$, at the end of the planing time t_1, so then future sales can be calculated from the system equation with no control (i.e. no advertising expenditure). We may also wish to follow as closely as possible some desired trajectory $y_d(x,t)$ during the planing time $[o,t_1]$ so that sales over the planing time also follow some desired trajectory. To achieve these two aims with minimum cost we take as cost functional

$$(9.64) \qquad g\int_o^1 \{y(x,t_1) - y_d(x,t_1)\}^2 dx + d\int_o^{t_1}\int_o^1 \{y(x,t) - y_d(x,t)\}^2 dxdt + c\int_o^{t_1} u^2(t)dt \quad ,$$

where g, d and c are appropriate weighting constants.

Remark

Obviously the optimum advertising problem can be considered as an optimum inventory control problem, as given in section 9.1 (by taking $\bar{\mu}$; the forgetting rate in the advertising problem, as $= -\mu$; the deterioration rate in the inventory control problem), and then the control (9.58), $[y(x,t)]_{x_f} = Fu(t)$ would represent 1 unit of expenditure resulting in the production of F goods in deterioration state x_f. The solution to the advertising problem is thus similar to that of the inventory problem of section 9.1 with modifications to allow for the slightly different control function and boundary conditions.

<u>**Solution to the Deterministic Problem**</u>

We wish to minimize (9.64) where y is given by the solution to (9.55), (9.56), (9.58) and (9.59), i.e.

$$(9.65)\qquad \frac{\partial y}{\partial t} = \bar{\mu}\frac{\partial y}{\partial x}$$

$$y(x,o) = y_o(x)\quad , \quad y(1,t) = 0\quad , \quad [y(x,t)]_{x_f} = Fu(t) = \bar{u}(t)\quad .$$

To employ the general theory we take $Ay = \bar{\mu}\frac{\partial y}{\partial x}$ with $D(A) = \{y \in H,\ y_x \in H,\ y(1) = 0\}$ where $H \in L^2(\Omega)$ and $\Omega = (0,1)$, so then (as in the inventory control problem) $A^* = A = -\bar{\mu}\frac{\partial}{\partial x}$ with $D(A^*) = \{y \in H,\ y_x \in H,\ y(o) = 0\}$. If we let $\tilde{A}$ be the same formal operator as A but now defined on $D(\tilde{A}) = \{y \in L^2(\Omega/\{x_f\}),\ y_x \in L^2(\Omega/\{x_f\}),\ y(1) = 0\}$ then we have the Green's formula

$$\langle \tilde{A}y_1, y_2\rangle = \bar{\mu}y_2(x_f)[y_1]_{x_f} + \langle y_1, A^*y_2\rangle$$

for $y_1 \in D(\tilde{A})$, $y_2 \in D(A^*)$, which comparing with the general Green's formula (2.7) implies that $B^* = C = \bar{\mu}y(x_f)$ so that $B = \bar{\mu}\delta(x-x_f)$. We also have that A^* on $D(A^*)$ generates the semigroup T_t^* where T_t^* is given by

$$(T_t^*y)(x) = \begin{cases} y(x-\bar{\mu}t) & 0 \le x - \bar{\mu}t \\ 0 & \text{else} \end{cases}$$

and thus

$$B^*T_t^* = \begin{cases} B^*y(x-\bar{\mu}t) & 0 \le x - \bar{\mu}t \\ 0 & \text{else} \end{cases}$$

$$= \begin{cases} \bar{\mu}y(x_f-\bar{\mu}t) & 0 \le x_f - \bar{\mu}t \\ 0 & \text{else} \end{cases}$$

so

$$||B^*T_t^*y||_{L^2[o,t_1;U^*]} = \int_o^{t_1}||B^*T_t^*y||^2dt$$

$$= \int_o^{(x_f/\bar{\mu})} \bar{\mu}^2y^2(x_f-\bar{\mu}t)dt$$

$$= \int_o^{x_f} \bar{\mu}^2 y^2(s)ds \qquad \text{putting} \quad x_f - \bar{\mu}t = s$$

$$\leq \int_o^1 \bar{\mu}y^2(s)ds \qquad \text{since} \quad x_f \in [0,1]$$

and hence

$$||B^*T_t^*y||_{L^2[o,t_1;U^*]} \leq \bar{\mu}||y||_{L^2[0,1]} \quad .$$

Therefore

$$||T_tBu||_{L^2[0,1]} \leq \bar{\mu}||u||_{L^2[o,t_1;U]}$$

and condition (2.4) is satisfied with $p \geq 2$. Hence we have that y given by

$$(9.66) \qquad y(x,t) = T_ty_o + \int_o^t T_{t-s}B\bar{u}(s)ds$$

is a weak solution to (9.65), and so we seek to minimize (9.64) with y given by (9.66). We know from the general theory that there exists a unique optimal control $u^* = - R^{-1}[B^*Q(t)y(t) + B^*S(t)]$ where $Q(t)$ satisfies the differential equation

$$(9.67) \qquad \frac{d}{dt} \langle Q(t)h, k\rangle + \langle Q(t)h, Ak\rangle + \langle Ah, Q(t)k\rangle - \langle B^*Q(t)h, R^{-1}B^*Q(t)k\rangle + \langle h, Mk\rangle = 0 \quad , \quad h , k \in D(A)$$

$$Q(t_1) = G$$

and $S(t)$ satisfies

$$(9.68) \qquad \frac{d}{dt} \langle x, S(t)\rangle = - \langle Ax, S(t)\rangle + \langle B^*Q(t)x, R^{-1}B^*S(t)\rangle + \langle x, My_d\rangle$$

$$S(t_1) = - Gy_d(t_1) \qquad\qquad x \in D(A)$$

with $G = g$, $M = d$ and $R = \frac{c}{F^{1/2}}$.

If as for the inventory control problem we assume that $Q(t)$ has the form

$$(9.69) \qquad (Qy)(\zeta) = \int_o^1 K(\zeta,\eta,t)y(\eta)d\eta$$

and substitute for this Q in (9.67) we find that K must satisfy

$$\int_o^1\int_o^1\Big[K_t(\zeta,\eta,t) - \bar{\mu}\{K_\zeta(\zeta,\eta,t) + K_\eta(\zeta,\eta,t)\} - \bar{\mu}^2 \frac{F^{1/2}}{c} K(x_f,\eta,t)K(\zeta,x_f,t) + d\delta(\zeta-\eta)\Big]h(\eta)k(\zeta)d\eta d\zeta$$

with $K(o,\eta,t) = 0 = K(\zeta,o,t)$,

which is satisfied if we chose K to satisfy

(9.70) $$K_t(\zeta,\eta,t) - \bar{\mu}\{K_\zeta(\zeta,\eta,t) + K_\eta(\zeta,\eta,t)\} - \bar{\mu}^2 \frac{F^{1/2}}{c} K(x_f,\eta,t)K(\zeta,x_f,t) + d\delta(\zeta-\eta) = 0$$

$$K(o,\eta,t) = 0 = K(\zeta,o,t) \quad .$$

Similarly writing $S(t)(\zeta)$ as $S(\zeta,t)$ and substituting into the differential equation for S , (9.68), gives

$$\int_o^1\Big[S_t(\zeta,t) - \bar{\mu}^2 \frac{F^{1/2}}{c} K(\zeta,x_f,t)S(x_f,t) - \bar{\mu}S_\zeta(\zeta,t) - dy_d(\zeta)\Big]x(\zeta)d\zeta = 0$$

when $S(o,t) = 0$, which is satisfied if S is chosen to satisfy

(9.71) $$S_t(\zeta,t) - \bar{\mu}S_\zeta(\zeta,t) - \bar{\mu}^2 \frac{F^{1/2}}{c} K(\zeta,x_f,t)S(x_f,t) - dy_d(\zeta) = 0$$

$$S(o,t) = 0 \quad .$$

Thus the optimal control u^* can be written as

$$u^* = - R^{-1}[B^*Qy + B^*S]$$

$$= \frac{F^{1/2}}{c} \bar{\mu}[\int_o^1 K(x_f,\eta,t)y(\eta)d\eta + S(x_f,t)]$$

where K satisfies (9.70) and S satisfies (9.71).

<u>Solution to the Stochastic Problem</u>

As in the stochastic inventory control problem this leads to an estimate with $2 > p \geq 1$ and thus we can only consider cost functionals (9.64) with

$g = 0$. Hence in this case we seak to minimize (9.64), with $g = 0$, where y satisfies

$$(9.72)\quad \frac{\partial y}{\partial t} = \bar{\mu}\frac{\partial y}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 y}{\partial x^2}$$

$$y(x,o) = y_o(x) \quad , \quad y(1,t) = 0 \quad , \quad \frac{\partial y}{\partial x}(o,t) = 0$$

$$[y(x,t)]_{x_f} = Fu(t) = \bar{u}(t) \quad .$$

As in the solution to the stochastic inventory problem we write

$$z(x,t) = e^{(\bar{\mu}/\sigma^2)x}y(x,t) \quad , \quad v(t) = e^{(\bar{\mu}/\sigma^2)x}u(t) \quad , \quad z_o(x) = e^{(\bar{\mu}/\sigma^2)x}y_o(x)$$

and then (9.72) is equivalent to

$$z_t = \frac{\sigma^2}{2} z_{xx} - \frac{\bar{\mu}^2}{2\sigma^2} z$$

$$z(x,o) = z_o(x) \quad , \quad z(1,t) = 0 \quad , \quad z_x(o,t) = 0$$

$$[z(x,t)]_{x_f} = \bar{v}(t) \quad .$$

If we now take $\Omega = [0,1]$ and $H = L^2(\Omega)$, and consider the operator $A = \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2} - \frac{\bar{\mu}}{2\sigma^2}$ on $D(A) = \{y \, , \, y_{xx} \in H \, , \, y(1) = 0 \, , \, y_x(o) = 0\}$ then A generates the semigroup T_t given by

$$T_t y_o = 2\sum_{n=1}^{\infty} e^{-c_n t}\langle y_o(\cdot), \cos n\pi\cdot\rangle \cos n\pi x \quad , \quad c_n = \frac{\bar{\mu}^2}{2\sigma^2} + \frac{\sigma^2}{8}(2n\pi+\pi)^2 \, .$$

We also have that $A^* = A$ with $D(A^*) = D(A)$. If $\tilde{A}$ is the same formal operator as A but with $D(\tilde{A}) = \{y \, , \, y_{xx} \in L^2(\Omega/\{x_f\}) \, , \, y(1) = 0 \, , \, y_x(o) = 0\}$ then we have the following Green's formula for $y_1 \in D(\tilde{A})$, $y_2 \in D(A^*)$

$$\langle \tilde{A}y_1, y_2\rangle = -\frac{\sigma^2}{2}[y_1]_{x_f}\frac{\partial y_2}{\partial x}(x_f) + \frac{\sigma^2}{2}\left[\frac{\partial y_1}{\partial x}\right]_{x_f} y_2(x_f)$$

which comparing with the general Green's formula (2.7) gives $B = C^* = -\frac{\sigma^2}{2}\delta'(x - x_f)$.

We can now, as for the case of the stochastic inventory control problem show that T_t , B satisfy condition (2.4) with a p such that $2 > p \geq 1$ and hence, by taking $g = 0$ in (9.64), we may employ the general theory of the tracking problem. In this case we seak to minimize

$$(9.73) \qquad J(u) = d\int_0^{t_1}\int_0^1 \langle z - z_d,\ e^{-2(\bar{\mu}/\sigma^2)x}(z - z_d)\rangle dxdt + \frac{ce^{-2(\bar{\mu}/\sigma^2)x}}{F^{1/2}}\int_0^{t_1}\langle u, u\rangle dt$$

where z is given by

$$(9.74) \qquad z(t) = T_t z_o + \int_0^t T_{t-s}Bu(s)ds$$

and $z_d = e^{(\bar{\mu}/\sigma^2)x}y_d$. The unique optimal control is then given by $u^* = - R^{-1}[B^*Qz + B^*S]$ where Q satisfies (9.67) and S satisfies (9.68) with $G = 0$, $M = de^{-2(\bar{\mu}/\sigma^2)x}$ and $R = \dfrac{ce^{-2(\bar{\mu}/\sigma^2)x}}{F^{1/2}}$.

If we again assume that we can write Q in the form (9.69) then substituting for this Q into (9.67) gives

$$(9.75) \qquad \int_0^1\int_0^1 \Big(K_t(\zeta,\eta,t) + \frac{\sigma^2}{2}\{K_{\zeta\zeta}(\zeta,\eta,t) + K_{\eta\eta}(\zeta,\eta,t)\} - \frac{\bar{\mu}^2}{\sigma^2}K(\zeta,\eta,t)$$
$$- \frac{\sigma^4}{4}\frac{F^{1/2}}{c}\{K_\zeta(x_f,\eta,t) - \frac{\bar{\mu}}{\sigma^2}K(x_f,\eta,t)\}\{K_\eta(\zeta,x_f,t) - \frac{\bar{\mu}}{\sigma^2}K(\zeta,x_f,t)\}$$
$$+ de^{-(\bar{\mu}/\sigma^2)(\zeta+\eta)}\delta(\zeta-\eta)\Big)h(\eta)k(\zeta)d\eta d\zeta = 0$$

when $K(1,\eta,t) = 0 = K_\zeta(o,\eta,t) = K(\zeta,1,t) = K_\eta(\zeta,o,t)$

which is satisfied by a K such that

$$(9.76) \qquad K_t(\zeta,\eta,t) + \frac{\sigma^2}{2}\{K_{\zeta\zeta}(\zeta,\eta,t) + K_{\eta\eta}(\zeta,\eta,t)\} - \frac{\bar{\mu}^2}{\sigma^2}K(\zeta,\eta,t)$$
$$+ de^{-(\bar{\mu}/\sigma^2)(\zeta+\eta)}\delta(\zeta-\eta)$$
$$- \frac{\sigma^4}{4}\frac{F^{1/2}}{c}\{K_\zeta(x_f,\eta,t) - \frac{\bar{\mu}}{\sigma^2}K(x_f,\eta,t)\}\{K_\eta(\zeta,x_f,t) - \frac{\bar{\mu}}{\sigma^2}K(\zeta,x_f,t)\} = 0$$

Similarly, writing $(S(t))(\zeta) = S(\zeta,t)$ and substituting into (9.68) leads to

$$\int_o^1\Big[S_t(\zeta,t) + \frac{\sigma^2}{2} S_{\zeta\zeta}(\zeta,t) - \frac{\bar{\mu}^2}{2\sigma^2} S(\zeta,t) + de^{-2(\bar{\mu}/\sigma^2)\zeta} z_d$$

$$- \frac{\sigma^4}{4} \frac{F^{1/2}}{c} \{K_\eta(\zeta,x_f,t) - \frac{\bar{\mu}}{\sigma^2} K(\zeta,x_f,t)\}\{S_\zeta(x_f,t) - \frac{\bar{\mu}}{\sigma^2} S(x_f,t)\}\Big] h(\zeta)d\zeta = 0$$

when $S_\zeta(o,t) = S(1,t) = 0$

which is satisfied if S satisfies

$$\text{(9.77)} \quad S_t(\zeta,t) + \frac{\sigma^2}{2} S_{\zeta\zeta}(\zeta,t) - \frac{\bar{\mu}^2}{2\sigma^2} S(\zeta,t) + de^{-2(\bar{\mu}/\sigma^2)\zeta} z_d$$

$$- \frac{\sigma^4}{4} \frac{F^{1/2}}{c} \{K_\eta(\zeta,x_f,t) - \frac{\bar{\mu}}{\sigma^2} K(\zeta,x_f,t)\}\{S(x_f,t) - \frac{\bar{\mu}}{\sigma^2} S(x_f,t)\} = 0.$$

Hence the optimal control for (9.73) - (9.74) is

$$u^* = - R^{-1}[B^*Qz + B^*S]$$

$$= \frac{F^{1/2}}{-c} e^{2(\bar{\mu}/\sigma^2)} \left\{ \int_o^1 K_\zeta(x_f,\eta,t)z(\eta)d\eta + S_\zeta(x_f,t) \right\}$$

where K , S satisfy (9.76) and (9.77) respectively.

As in the case of the stochastic inventory problem we can easily show that the solution to the original problem of minimizing (9.64) with $g = 0$, and y given by (9.72) is the optimal control $v^* = e^{(\bar{\mu}/\sigma^2)x} u^* = - R^{-1}B^*[Py + N]$ where if we write

$$(Py)(\zeta) = \int_o^1 L(\zeta,\eta,t)y(\eta)d\eta$$

then

$$L(\zeta,\eta,t) = K(\zeta,\eta,t)e^{(\bar{\mu}/\sigma^2)\zeta}e^{(\bar{\mu}/\sigma^2)\eta} ,$$

and

$$N(\zeta,t) = S(\zeta,t)e^{(\bar{\mu}/\sigma^2)\zeta} .$$

Furthermore L and N satisfy

$$\text{(9.78)} \quad L_t(\zeta,\eta,t) + \frac{\sigma^2}{2}\{L_{\zeta\zeta}(\zeta,\eta,t) + L_{\eta\eta}(\zeta,\eta,t)\} - \bar{\mu}\{L_\zeta(\zeta,\eta,t) + L_\eta(\zeta,\eta,t)\}$$

$$- \frac{\sigma^4}{4} \frac{F^{1/2}}{c} L_\zeta(x_f,\eta,t)L_\eta(\zeta,x_f,t) + d\delta(\zeta-\eta) = 0$$

with $L(1,\eta,t) = L_\zeta(o,\eta,t) = L(\zeta,1,t) = L_\eta(\zeta,o,t) = 0$

and

$$(9.79)\qquad N_t(\zeta,t) + \frac{\sigma^2}{2} N_{\zeta\zeta}(\zeta,t) - \bar{\mu}\, N_\zeta(\zeta,t) - \frac{\sigma^4}{4}\,\frac{F^{1/2}}{c}\, L_\eta(\zeta,x_f,t)N_\zeta(x_f,t) + dy_d = 0 ,$$

with $N(o,t) = N(1,t) = 0$

and then v^* is given by

$$v^* = \frac{F^{1/2}}{-c}\left(\int_o^1 L_\zeta(x_f,\eta,t)y(\eta)d\eta + N_\zeta(x_f,t)\right) .$$

Remarks

The problem as posed could be modified to take into account the increase in goodwill due to word of mouth recommendation. For instance, suppose the rate of increase in goodwill due to word of mouth is also stochastic, given by $\rho\Delta t + s\Delta\bar{\omega}(t)$ with $\bar{\omega}(t)$ again Brownian motion, then (9.56) would become

$$\frac{\partial y}{\partial t} = (\bar{\mu} - \rho)\frac{\partial y}{\partial x} + \frac{1}{2}(\sigma^2 - s^2)\frac{\partial^2 y}{\partial x^2}$$

For a survey of results relating to the use of finite dimensional control theory in optimal advertising see the paper of Suresh Sethi {36}.

9.3. Pollution Control

Bensoussan, Hurst and Näslund { 5 } give four reasons why pollution control is very important for many business firms.

1. The laws, taxes, grants and subsidies associated with pollution and pollution control

2. Pollution may enter as a benefit (or cost) in production.

3. Relations with other firms:- e.g. stockholders who own stock in two companies, one polluting the waters used by the other, will be interested in pollution

control.

4. Public Pressure:- Here pollution can be considered as the opposite of advertising, too much pollution creates badwill.

The aspect of pollution considered here is water pollution, in particular river and lake pollution. In the case of river pollution the amount of pollutant at the point (x,t) , $y(x,t)$, is given by

$$(9.80)\qquad \frac{\partial y}{\partial t}(x,t) = D\,\frac{\partial^2 y}{\partial x^2}(x,t) - V\,\frac{\partial y}{\partial x}(x,t) + f(x,t) \qquad 0 \leq x \leq 1$$

where D is the dispersion coefficient, V is the water velocity and $f(x,t)$ is the rate of increase of the concentration at (x,t) due to deposits of chemical wastes. This is the model Kawakernaak {19} and Curtain { 8 } consider for $f(x,t)$ a stochastic process.

Suppose we have an initial pollution level

$$(9.81)\qquad y(x,o) = y_o(x)$$

and that the rate of increase in pollution $f(x,t)$ is a given, known, function (possibly zero). We might wish to remove pollutants from the river following some planned pattern, i.e. keeping the pollutant level following as closely as possible some desired trajectory $y_d(x,t)$ with minimum cost. One possible means is by using a filter to remove pollutants at the point x_1 . If we let

$$(9.82)\qquad [y(x,t)]_{x_1} = u(t)$$

then $u(t)$ becomes the control variable and we might take as cost functional

$$(9.83)\qquad J(u) = \int_o^{t_1}\int_o^1 \{y_d(x,t) - y(x,t)\}^2 dxdt + c\int_o^{t_1} u^2(t)dt \quad .$$

By taking f into the cost functional, as was done in the case of the inventory control problem, the problem is the same as that of optimal advertising considered in section 9.2, and thus the results of that section apply here (by taking $D = \sigma^2/2$ and $V = -\bar{\mu}$).

In the case of a lake, which we might consider as a stagnant river, the equation for the pollution level then becomes

$$(9.84) \qquad \frac{\partial y}{\partial t}(x,t) = D\,\frac{\partial^2 y}{\partial x^2}(x,t) + f(x,t)$$

since V , the water velocity, is zero.

Suppose $f = 0$, the dispersion coefficient $D = 1$ and we control the pollution level by controlling the level at the point $x = 0$ (the pollution level at $x = 1$ being kept at zero throughout), then, if the cost functional is again given by (9.83) we are in the situation of example 5.14, and so the results for that example apply here.

9.4. Traffic Flow Control

The model for traffic flow control considered in this section is based on the work of Tabac {37}.

For roads carrying a large population of automobiles we can regard the flow of automobiles as a continuous flow. If we consider a segment of road dx centred on x during the time interval of length dt , and define the flow, q , as the number of vehicles per unit time crossing the interval dx at x , the concentration, c , as the number of vehicles per unit of road, and the space-mean speed v as the average of vehicle speeds weighted according to the time they remain on the road interval dx , then

(9.85) $v = \frac{q}{c}$, i.e. $q = vc$.

If the flow entering the segment is q then the flow at the end is $q + \frac{\partial q}{\partial x}dx$; if the concentration at the begining of the time interval is c , at the end it is $c - \frac{\partial c}{\partial t}dt$; and if the number of vehicles at the begining of the time interval is $c\,dx$ at the end it is $(c - \frac{\partial c}{\partial t}dt)dx$. The number of vehicles entering the segment within dt is $q\,dt$ and the number exiting is $(q + \frac{\partial q}{\partial x}dx)dt$. Thus the balance of vehicles may be expressed as

$$c\,dx - (c - \frac{\partial c}{\partial t})dx = q\,dt - (q + \frac{\partial q}{\partial x}dx)dt$$

i.e. $\frac{\partial c}{\partial t} + \frac{\partial q}{\partial x} = 0$,

which using (9.85) becomes

(9.86) $$\frac{\partial c}{\partial t} + v\frac{\partial c}{\partial x} = 0 .$$

Hence the concentration satisfies the one dimensional wave equation.

During periods of heavy traffic we might wish to control the traffic flow so that the concentration is as near as possible to some desired level c_d which is the optimum for safe, efficient travel, and so we take as performance index

(9.87) $$J(u) = \int_o^{t_1}\{\int_{x_o}^{x_1}(c(x,t) - c_d(x,t))^2dx + Ru^2(t)\}dt$$

As a means of control we might use the flow of traffic entering the road so that

$$q(x_o,t) = u(t)$$

i.e.

(9.88) $$c(x_o,t) = \frac{1}{v}u(t) = \bar{u}(t) .$$

The problem as formulated is as for the inventory control problem with deterministic rate of deterioration, considered in section 9.1, and thus the results of that section , for that problem, apply here.

9.5. Population Control

The evolution of a country (or indeed any animal colony with non-seasonal birth pattern) can be descrided by the following partial differential equation

(9.89) $$\frac{\partial p}{\partial t}(t,r) + \frac{\partial p}{\partial r}(t,r) = -\mu(t,r)p(t,r)$$

$$p(o,r) = p_o(r) \qquad 0 \le r \le 1$$

$$p(t,o) = u(t) \qquad 0 \le t \le t_1$$

where $p(t,r)$ represents the population density of individuals of age r at time t , $\mu(t,r)$ is the mortality function, $p_o(r)$ is the given initial age distribution and $u(t)$ is the birth rate which is assumed to be the control variable.

One problem might be to chose the birth rate u so as to achieve a desired age profile $q(r)$ at the final time t_1 and thus we take as cost functional

(9.90) $$J(u) = \int_o^1 (p(t_1,r) - q(r))^2 dr + \int_o^{t_1} \lambda u^2(s)ds$$

where the second term measures the social cost of controlling the birth rate.

Alternatively we might wish to control the population growth and to choose u such that the population follows some predetermined profile $q(t,r)$ and so we would take as cost functional

(9.91) $$J(u) = \int_o^{t_1}\int_o^1 (p(t,r) - q(t,r))^2 drdt + \int_o^{t_1} \lambda u^2(s)ds \quad .$$

Solution of the Population Control Problem

If we take A as the operator $A = -\frac{\partial}{\partial x} - \mu$ on $D(A) = \{z \in H^1(0,1)$, $z(o) = 0\}$ then (9.89) may be written as

$$\frac{\partial p}{\partial t} = Ap$$

$$p(o,r) = p_o(r) \quad , \quad p(t,o) = u(t)$$

and then if $H = L^2(0,1)$, $U = R$, we have also (see Zabczyk {43}) that A generates a strongly continuous semigroup T_t on H given by

$$T_t p(x) = \begin{cases} p(x-t) \exp\left(-\int_{x-t}^{x} \mu(s)ds\right) & t \leq x \\ 0 & t \geq x \end{cases} .$$

It is easy to see that $A^* = \frac{\partial}{\partial x} - \mu$ with $D(A^*) = \{z \in H^1(0,1)$, $z(1) = 0\}$. If we take $\tilde{A}$ as the same formal operator as A but with $D(\tilde{A}) = \{z \in H^1(0,1)\}$ then we have the following Green's formula for $p_1 \in D(\tilde{A})$, $p_2 \in D(A^*)$

$$<\tilde{A}p_1, p_2> = p_1(o)p_2(o) + <p_1, A^*p_2>$$

which comparing with the general Green's formula (2.7) implies that $Cp = p(o)$ and hence taking $B^* = C$, $B = \delta$, the dirac delta function. Therefore

$$T_t Bu = \begin{cases} Bu(x-t) \exp\left(-\int_{x-t}^{x} \mu(s)ds\right) & t \leq x \\ 0 & t \geq x \end{cases}$$

$$= \begin{cases} \delta(x-t)\ u \exp\left(-\int_{x-t}^{x} \mu(s)ds\right) & t \leq x \\ 0 & t \geq x \end{cases} ,$$

and as for the inventory control problem with deterministic deterioration rate, in order to obtain an estimate for $T_t Bu$ we consider the adjoint system. Since $A^* = \frac{\partial}{\partial x} - \mu$ on $D(A^*) = \{z \in H^1(0,1)$, $z(1) = 0\}$, A^* generates the strongly continuous semigroup T_t^*

$$T_t^*p(x) = \begin{cases} p(x+t)\exp\left(-\int_x^{x+t} \mu(s)ds\right) & x+t \leq 1 \\ 0 & \text{else} \end{cases}$$

Thus

$$B^*T_t^*p = \begin{cases} p(t)\exp\left(-\int_x^{t} \mu(s)ds\right) & t \leq 1 \\ 0 & \text{else} \end{cases}$$

since $B^* = C$ so that $B^*(T_t^*p) = (T_t^*p)(o)$, and hence

$$||B^*T_t^*p||_{L^2[o,t_1;U^*]} = \int_o^{t_1} ||B^*T_t^*p||^2 dt$$

$$= \int_o^1 p^2(t)\exp\left(-2\int_o^t \mu(s)ds\right)dt$$

$$\leq \int_o^1 p^2(t)dt$$

$$= ||p||_{L^2[0,1]} \quad .$$

Hence

$$||T_tBu||_{L^2[0,1]} \leq ||p||_{L^2[0,t_1;U]}$$

and so condition (2.4) holds with $p \geq 2$ and thus the control problem (9.89) with either cost functional (9.90) or (9.91) is well defined. We thus know they lead to an optimal control $u^* = -R^{-1}[B^*Q(t)p(t) + B^*S(t)]$ where Q satisfies

$$\frac{d}{dt}\langle Q(t)h, k\rangle + \langle Q(t)h, Ak\rangle + \langle Ah, Q(t)k\rangle - \langle B^*Q(t)h, R^{-1}B^*Q(t)k\rangle + \langle h, Mk\rangle = 0 ,$$

$$Q(t_1) = G$$

and S satisfies

$$\frac{d}{dt}\langle x, S(t)\rangle = -\langle Ax, S(t)\rangle + \langle B^*Q(t)x, R^{-1}B^*S(t)\rangle + \langle x, Mq\rangle$$

$$S(t_1) = -Gq(t_1),$$

where in the case of the cost functional (9.90), $G = I$, $R = \lambda$ and $M = 0$ and for (9.91), $G = 0$, $R = \lambda$ and $M = I$.

As previously if we let

$$Q(t)h(\zeta) = \int_0^1 K(\zeta,\eta,t)h(\eta)d\eta \quad \text{and} \quad S(t)(\zeta) = S(\zeta,t)$$

then for the problem (9.89) - (9.90) we have that the equations satisfied by K and S are the following

$$K_t(\zeta,\eta,t) + K_\zeta(\zeta,\eta,t) + K_\eta(\zeta,\eta,t) - 2\mu K(\zeta,\eta,t) - \frac{1}{\lambda} K(o,\lambda,t)K(\zeta,o,t) = 0$$

$$\text{if} \quad K(1,\eta,t) = 0 = K(\zeta,1,t)$$

and the final time condition

$$\int_0^1 K(\zeta,\eta,t_1)h(\eta)d\eta = h(\zeta) .$$

Also

$$S_t(\zeta,t) + S_\zeta(\zeta,t) - S(\zeta,t) - \frac{1}{\lambda} K(\zeta,o,t)S(o,t) = 0$$

$$S(1,t) = 0$$

with the final time condition

$$S(\zeta,t_1) = q(\zeta) .$$

The optimal control $u^* = - R^{-1}[B^*Q(t)p + B^*S(t)]$ is then given by

$$u^*(\zeta) = - \frac{1}{\lambda} \left(\int_0^1 K(o,\eta,t)p(\eta)d\eta + S(\zeta,t)\right) .$$

For the problem with cost functional (9.19) the two equations are

$$K_t(\zeta,\eta,t) + K_\zeta(\zeta,\eta,t) + K_\eta(\zeta,\eta,t) + K(\zeta,\eta,t) - \frac{1}{\lambda} K(o,\eta,t)K(\zeta,o,t) = 0$$

$$K(1,\eta,t) = 0 = K(\zeta,1,t)$$

with

$$\int_0^1 K(\zeta,\eta,t_1)h(\eta)d\eta = 0$$

and

$$S_t(\zeta,t) + S_\zeta(\zeta,t) - S(\zeta,t) - \frac{1}{\lambda} K(\zeta,o,t)S(o,t) - q(\zeta) = 0$$

$$S(1,t) = 0$$

with $S(\zeta,t_1) = 0$.

This time the optimal control is given by

$$u^*(\zeta,t) = -\frac{1}{\lambda}\left[\int_o^1 K(o,\eta,t)p(\eta)d\eta + S(\zeta,o)\right] .$$

10. COMMENTS ON HYPERBOLIC SYSTEMS

We would, of course, like to be able to apply the results of the previous sections on the linear quadratic cost control problem, to general hyperbolic systems. This does not however seem to be possible owing to the failure of the semigroups generated by such systems to be adequately smoothing.

In consideration of distributed control problems, Curtain and Pritchard {11}, examine systems of the form

(10.1) $z_{tt} + \alpha z + Az = Bu$

$z(x,o) = z_o$, $z_t(x,o) = z_1$

where A is a positive, self-adjoint operator on the Hilbert space H , with domain $D(A)$, and B is a bounded linear operator from the control space u , also a Hilbert space, to H . α is a constant. Letting

(10.2) $z_t = y$

so that

(10.3) $y_t = -\alpha y - Az + Bu$

we may write (10.1) as the following first order system

(10.4) $w_t = \mathcal{A}w + \mathcal{B}u$

$$w(x,o) = w_o = \begin{bmatrix} z_o \\ z_1 \end{bmatrix}$$

where

$$w = \begin{bmatrix} z \\ y \end{bmatrix} \quad , \quad \mathcal{A}w = \mathcal{A}\begin{bmatrix} z \\ y \end{bmatrix} = \begin{bmatrix} 0 & I \\ -A & -\alpha \end{bmatrix}\begin{bmatrix} z \\ y \end{bmatrix} \quad \text{and} \quad \mathcal{B} = \begin{bmatrix} 0 \\ B \end{bmatrix} .$$

Taking X to be the Hilbert space $X = D(A^{\frac{1}{2}})\times H$, with inner product

(10.5) $\langle w_1, w_2\rangle = \langle A^{\frac{1}{2}}z_1, A^{\frac{1}{2}}z_2\rangle + \langle y_1, y_2\rangle$ for $w_i = \begin{bmatrix} z_i \\ y_i \end{bmatrix}$,

Curtain and Pritchard show that $\mathcal{A}$ generates a strongly continuous semigroup T_t on X . They are thus able to solve the linear quadratic cost control problem for such systems.

Consider now systems such as (10.1) but with boundary rather than distributed control action, e.g. systems of the form

(10.6) $z_{tt} + \alpha z + Az = 0$

$z(x,o) = z_1 \quad , \quad z_t(x,o) = z_1$,

with either

(10.7) $z(o,t) = u \quad , \quad z(1,t) = 0$

or

(10.8) $z_x(o,t) = u \quad , \quad z_x(1,t) = 0$.

As previously demonstrated, for first order systems, such control action gives rise to an unbounded control operator. In order to apply the results of this thesis we require that T_t , $\mathcal{B}$ satisfy condition (2.4), i.e. that the semigroup T_t be smoothing. Unfortunately this does not seem to be the case and so we are unable to apply our theory for the solution of the linear quadratic cost control problem to such systems.

One method for tackling linear quadratic cost problems with boundary control action, for such hyperbolic systems, is to reformulate them, in a restricted sense, as problems with bounded, distributed, control action, and then by the results of {11} we know that there exist exists a unique solution to the linear quadratic cost control problem. There are three possibilities, we may either, restrict the control space, restrict G and W, or restrict R (where G, W and R are the weighting operators in the quadratic cost functional (1.7)).

Curtain and Pritchard {10} consider the first possibility, i.e. restricting the control space, for a particular example of a controlled wave equation. They show that by taking $H = L^2(\Omega)$ and choosing the smooth control space $U = H^{\frac{1}{2}}(\partial\Omega)$ they can reformulate the problem as one with bounded control action since then $B \in \mathcal{L}(U, L^2(\Omega))$. (The problem for this same wave equation with bounded, distributed control action is solved by Curtain and Pritchard { 9 }, and also Vinter and Johnson {38} using a Lions type approach.)

Zabczyk {32} also restricts the control space in order to include hyperbolic systems in his formulation for boundary control problems. He gives as an example how the same wave equation can be formulated as a boundary control problem (satisfying his formulation as outlined in section 8) by again restricting the control space to $H^{\frac{1}{2}}(\partial\Omega)$.

Curtain and Pritchard {10} note that alternatively we could restrict G and W so that the feedback controls always lie in $H^{\frac{1}{2}}(\partial\Omega)$ and that this is in fact the case if G and $W \in \mathcal{L}(L^2(\Omega), H_0^1(\Omega))$, and provided $R^{-1} \in \mathcal{L}(H^{\frac{1}{2}}(\partial\Omega))$. They also note that this is essentially what Vinter and Johnson {38} have done by assuming $W = 0$, R the identity on $L^2(\partial\Omega)$

and $G \in \mathcal{L}(L^2(\Omega), H_o^1(\Omega))$. Vinter and Johnson's approach, however is quite different from the semigroup approach adopted by Curtain and Pritchard and in this thesis.

The third alternative is to restrict the choice of weighting operators R to those strictly positive, bounded, linear operators on $U = L^2(\partial\Omega)$ which are also strictly positive, bounded, linear operators when considered as operators from $H^{\frac{1}{2}}(\partial\Omega)$ to $(H^{\frac{1}{2}}(\partial\Omega))^*$ so then R^{-1} exists and $R^{-1} \in \mathcal{L}((H^{\frac{1}{2}}(\partial\Omega))^*, H^{\frac{1}{2}}(\partial\Omega))$. This also has the effect of assuring that the feedback controls always remain in $H^{\frac{1}{2}}(\partial\Omega)$.

Lions {24} also considers the linear quadratic cost control problem for a hyperbolic, second order, system, using his methods as outlined in section 8. He shows that the optimal control is given by the simultaneous solution of the system and adjoint equations but he does not give any results on decoupling the two equations (in this case) to provide the solution in the form of a Riccati equation.

Linear symmetric hyperbolic systems in two independent variables, are considered by Russell {33} using a Lions type approach. He shows that the optimal control is feedback and given by the solution of a Riccati type equation for both the finite time and infinite time problems.

Kim and Erzberg {20} consider a system, similar to that of Russell, of an N-dimensional wave equation. Using dynamic programing techniques they, formally, derive the optimal control and the Riccati equations for the system. They do not however cite any existence or uniqueness results.

11. OBSERVER THEORY

Observers of the type considered here are known as Luenberger observers after D.C. Luenberger (see {25}) and their theory for finite dimensional linear systems is well developed (see for example Wonham or Wolovich {40}).

The use of observers in the finite dimensional regulator problem has been considered by Sarma and Jayaraj {35}, Newmann {27}, Borigiorno Jr. and Youla { 7 } and Yuksel and Borigiorno Jr. {42} amongst others. In all these papers an expression for the increase in cost, ΔJ , is obtained which depends on the initial state of the system.

For infinite dimensional systems one approach is that taken by, for instance, Orner and Foster {28}. They define an optimal approximating finite dimensional model for the infinite dimensional system and then consider observers for this finite dimensional model.

Working directly with the infinite dimensional system Gressang and Lamont {14} consider observers for systems characterised by semigroups, and give conditions for an observer to asymptotically estimate the state of such a system. They confine themselves to controls $u(t)$ which are strongly continuous and are thus able to use the differential form (1.1) of the system equations. Here we examine the conditions for an observer to asymptotically estimate a system given by a mild solution of the form (1.2). We also consider the use of observers in the linear quadratic regulator problem and the increase in cost this leads to when an observer is used rather than the system state feedback.

El Jai {12} and Prado {29} consider finite dimensional observers for a particular class of systems, namely those for which the operator A can be partitioned in the form $A = \begin{bmatrix} A_N & 0 \\ 0 & A_r \end{bmatrix}$ where the subscript N denotes the projection onto a subspace of dimension N and the subscript r the residual part. Here we examine the increase in regulator cost for these observers and for the particular case of systems with an expansion in terms of distinct eigenfunctions.

Also using the semigroup approach, Salmon {34} considers the problem of constructing an observer for systems with general time delays in state and output. In this report, using the spectral decomposition of the original system, he decomposes the observer into finite and infinite dimensional parts, and considers when the observer can be reduced to its finite dimensional component only.

Liu and Lapidus {26} also examine observer theory for distributed parameter systems in which the operators are the infinitesimal generators of strongly continuous semigroups. Using the direct method of Lyapunov they analyse the stability characteristics of the observer error for an infinite dimensional observer.

Conditions for observers of infinite dimensional systems in differential form, are examined by Kitamura, Sakairi and Nishimura {21}, with special reference to diffusion systems.

In his two papers {22} and {23}, Köhne gives examples of finite dimensional approximating observers for infinite dimensional systems. In {23} he considers their use in estimating the state of a heat conductor and in {22} he applies them to elastic systems giving the particular example of a transverse vibrating beam.

11.1 OBSERVERS FOR SYSTEMS DESCRIBED BY SEMIGROUPS AND WITH DISTRIBUTED CONTROL ACTION.

Gressang and Lamont {14} consider the observed system

(11.1) $$z_t = Az + Bu$$
$$z(o) = z_o \in D(A)$$

(11.2) $$y(t) = Cz(t)$$

where A is the infinitesimal generator of a strongly continuous semigroup T_t on a Banach space Z, $B \in \mathcal{L}(U,Z)$, U the control space also a Banach space, and the controls $u(t)$ are assumed to be strongly continuously differentiable. $C \in \mathcal{L}(Z,Y)$ where Y is the Banach space of observations.

They consider also the observer defined by

(11.3) $$x_t = Fx + Gy + Hu$$
$$x(o) = x_o \in D(F)$$

where F is the infinitesimal generator of a strongly continuous semigroup S_t on a Banach space X and $G \in \mathcal{L}(Y,X)$, $H \in \mathcal{L}(U,X)$.

For $P \in \mathcal{L}(Z,X)$ they call the observer (11.3) an asymptotic state estimator of $Pz(t)$ if

(11.4)
(i) $\lim_{t\to\infty} [x(t) - Pz(t)] = 0$
(ii) P maps $D(A)$ into $D(F)$,

where $z(t)$ is the solution of (11.1).

Now (11.1) and (11.3) have mild solutions

(11.5) $$z(t) = T_t z_o + \int_o^t T_{t-s} Bu(s)ds$$

and

(11.6) $$x(t) = S_t x_o + \int_o^t S_{t-s}[Gy(s) + Hu(s)]ds$$

respectively. Furthermore (11.5) and (11.6) are well defined for all $z_o \in Z$ and all $x_o \in X$ and for all controls $u \in L^2[o,t_1;U]$, and so we shall again work with these mild solutions as our system definitions. We can then remove the condition (ii) of (11.4) and make the following definition, which we use as our definition from now on.

<u>Definition 11.1</u>

For $P \in \mathcal{L}(Z,X)$, (11.6) defines an asymptotic state estimator of $Pz(t)$, where $z(t)$ given by (11.5), if

$$\lim_{t\to\infty} [x(t) - Pz(t)] = 0 .$$

We can now examine conditions for (11.6) to be an asymptotic state estimator of $Pz(t)$ and prove the following.

<u>Proposition 11.2</u>

For $P \in \mathcal{L}(Z,X)$

$$x(t) - Pz(t) = S_t(x_o - Pz_o)$$

if

(11.7)

(i) $\langle PAz, \bar{x}\rangle_{X,X^*} - \langle Pz, F^*\bar{x}\rangle_{X,X^*} = \langle GCz, \bar{x}\rangle_{X,X^*}$, $z \in D(A)$, $\bar{x} \in D(F^*)$

(ii) $Hu - PBu = 0 \quad \forall\, u \in U$,

so then, if F is stable, $\lim_{t\to\infty} [x(t) - Pz(t)] = 0$, i.e. $x(t)$ is an asymptotic state estimator of $Pz(t)$.

<u>Proof</u>

$$x(t) - Pz(t) = S_t x_o + \int_o^t S_{t-s}[Gy(s) + Hu(s)]ds - P[T_t z_o + \int_o^t T_{t-s}Bu(s)ds]$$

Rearranging the terms and using the fact that $y = Cz$, we have, for any $\bar{x} \varepsilon D(F^*)$,

$$(11.8) \quad \langle x(t) - Pz(t), \bar{x}\rangle = \langle S_t[x_o - Pz_o], \bar{x}\rangle + \langle [S_t P - PT_t]z_o, \bar{x}\rangle + \int_o^t \langle S_{t-s}GCz(s), \bar{x}\rangle ds + \int_o^t \langle [S_{t-s}H - PT_{t-s}B]u(s), \bar{x}\rangle ds \quad .$$

If we now restrict consideration to initial states $z_o \varepsilon D(A)$ and controls $u \varepsilon C^1[o,t_1;U]$ (from {11}) we know that $T_t z_o$ and $\int_o^t T_{t-s}Bu(s)ds \varepsilon D(A)$ for $t \geq 0$, and hence then $z(t)$, given by (11.5), belongs to $D(A)$ also.

Now

$$\frac{d}{d\rho} \langle S_{t-\rho}PT_\rho z_o, \bar{x}\rangle = \langle S_{t-\rho}PAT_\rho z_o, \bar{x}\rangle - \langle S_{t-\rho}PT_\rho z_o, F^*\bar{x}\rangle$$

and thus, integrating both sides with respect to ρ

$$(11.9) \quad \langle PT_t z_o, \bar{x}\rangle - \langle S_t Pz_o, \bar{x}\rangle = \int_o^t \langle S_{t-\rho}PAT_\rho z_o, \bar{x}\rangle d\rho - \int_o^t \langle S_{t-\rho}PT_\rho z_o, F^*\bar{x}\rangle d\rho .$$

Also,

$$\frac{d}{ds} \langle S_{t-s}P\int_o^s T_{s-r}Bu(r)dr, \bar{x}\rangle = \langle S_{t-s}PA\int_o^s T_{s-r}Bu(r)dr, \bar{x}\rangle + \langle S_{t-s}PBu(s), \bar{x}\rangle - \langle S_{t-s}P\int_o^s T_{s-r}Bu(r)dr, F^*\bar{x}\rangle$$

and hence

$$\int_0^t \langle S_{t-s}PA\int_0^s T_{s-r}Bu(r)dr, \bar{x}\rangle ds - \int_0^t \langle S_{t-s}P\int_0^s T_{s-r}Bu(r)dr, F^*\bar{x}\rangle ds$$

$$= \int_0^t \frac{d}{ds}\langle S_{t-s}P\int_0^s T_{s-r}Bu(r)dr, \bar{x}\rangle ds - \int_0^t \langle S_{t-s}PBu(s), \bar{x}\rangle ds$$

$$= \int_0^t\int_r^t \frac{d}{ds}\langle S_{t-s}PT_{s-r}Bu(r), \bar{x}\rangle dsdr - \int_0^t \langle S_{t-s}PBu(s), \bar{x}\rangle ds \quad ,$$

i.e.

$$(11.10) \quad \int_0^t \langle S_{t-s}PA\int_0^s T_{s-r}Bu(r)dr, \bar{x}\rangle ds - \int_0^t \langle S_{t-s}P\int_0^s T_{s-r}Bu(r)dr, F^*\bar{x}\rangle ds$$

$$= \int_0^t \langle PT_{t-r}Bu(r), \bar{x}\rangle dr - \int_0^t \langle S_{t-r}PBu(r), \bar{x}\rangle dr \ .$$

Adding (11.9) and (11.10) and combining the terms gives

$$(11.11) \quad \int_0^t \langle S_{t-s}PAz(s), \bar{x}\rangle ds - \int_0^t \langle S_{t-s}Pz(s), F^*\bar{x}\rangle ds$$

$$= \langle PT_tz_o, \bar{x}\rangle - \langle S_tPz_o, \bar{x}\rangle + \int_0^t \langle PT_{t-r}Bu(r), \bar{x}\rangle dr - \int_0^t \langle S_{t-r}PBu(r), \bar{x}\rangle dr.$$

Using (11.11), (11.8) becomes

$$(11.12) \quad \langle x(t) - Pz(t), \bar{x}\rangle$$

$$= \langle S_t[x_o - Pz_o], \bar{x}\rangle$$

$$+ \int_0^t \{\langle S_{t-s}GCz(s), \bar{x}\rangle - \langle S_{t-s}PAz(s), \bar{x}\rangle + \langle S_{t-s}Pz(s), F^*\bar{x}\rangle\}ds$$

$$+ \int_0^t \{\langle S_{t-s}Hu(s), \bar{x}\rangle - \langle S_{t-s}PBu(s), \bar{x}\rangle\}ds$$

It is easy to see that conditions (11.7) then imply that

$$(11.13)\quad \langle x(t) - Pz(t), \bar{x}\rangle = \langle S_t[x_o - Pz_o], \bar{x}\rangle \qquad \forall\, \bar{x} \in D(F^*) \quad .$$

(11.13) is well defined for all $z_o \in Z$ and all $u \in L^2[o,t_1;U]$. Thus since $D(A)$ is dense in Z and $C^1[o,t_1;U]$ is dense in $L^2[o,t_1;U]$ we may extend (11.13) to hold for all $z(t)$ given by (11.5) with $z_o \in Z$ and $u \in L^2[o,t_1;U]$. Then since $\overline{D(F^*)} = X^*$ we have that

$$x(t) - Pz(t) = S_t[x_o - Pz_o]$$

and the proposition is proved.

__Remark__

The conditions analagous to (11.7) obtained by Gressang and Lamont are

$$PA - FP = GC$$

$$H = PB$$

since they require P to map $D(A)$ into $D(F)$.

If we now consider feedback controls of the form

$$(11.14)\quad u(t) = My(t) + Nx(t) \qquad M \in \mathcal{L}(Y,U) \quad , \quad N \in \mathcal{L}(X,U)$$

i.e. we feedback a combination of the state of the observer and the output of the original system, the "error", $x(t) - Pz(t)$ then becomes

$$x(t) - Pz(t) = S_t[x_o - Pz_o] + [S_tP - PT_t]z_o + \int_o^t S_{t-s}GCz(s)ds +$$
$$+ \int_o^t [S_{t-s}H - PT_{t-s}B][My(s) + Nx(s)]ds$$

and since $y(t) = Cz(t)$ we have, for $\bar{x} \in D(F^*)$, (11.15) $\langle x(t) - Pz(t), \bar{x}\rangle$

$$= \langle S_t[x_o - Pz_o], \bar{x}\rangle + \langle [S_tP - PT_t]z_o, \bar{x}\rangle + \int_o^t \langle S_{t-s}GCz(s), \bar{x}\rangle ds$$

$$+ \int_o^t \langle S_{t-s}HMCz(s), \bar{x}\rangle ds - \int_o^t \langle PT_{t-s}BMCz(s), \bar{x}\rangle ds$$

$$+ \int_o^t \langle S_{t-s}HNx(s), \bar{x}\rangle ds - \int_o^t \langle PT_{t-s}BNx(s), \bar{x}\rangle ds ,$$

and we can now prove the following.

<u>Proposition 11.3</u>

The conditions sufficient for

$$x(t) - Pz(t) = S_t[x_o - Pz_o]$$

become, under feedback control of the form (11.14)

(11.16)

(i) $\langle PAz, \bar{x}\rangle - \langle Pz, F^*\bar{x}\rangle - \langle GCz, \bar{x}\rangle + \langle PBMCz, \bar{x}\rangle - \langle HMCz, \bar{x}\rangle = 0$ for $z \in D(A)$, $\bar{x} \in D(F^*)$

(ii) $HNx = PBNx \quad \forall x \in X$.

<u>Proof</u>

If we initially restrict consideration to those controls $u(t)$ given by (11.14) which are continuously differentiable, then substituting for such a u into (11.12) gives, for $z_o \in D(A)$, $\bar{x} \in D(F^*)$

$$\langle x(t) - Pz(t), \bar{x}\rangle$$

$$= \langle S_t[x_o - Pz_o], \bar{x}\rangle$$

$$+ \int_o^t \{\langle S_{t-s}Pz(s), F^*\bar{x}\rangle - \langle S_{t-s}PAz(s), \bar{x}\rangle - \langle S_{t-s}PBMCz(s), \bar{x}\rangle + \langle S_{t-s}GCz(s), \bar{x}\rangle$$

$$+ \langle S_{t-s}HMCz(s), \bar{x}\rangle\} ds$$

$$+ \int_o^t \{\langle S_{t-s}HNx(s), \bar{x}\rangle - \langle S_{t-s}PBNx(s), \bar{x}\rangle\} ds$$

so if (11.16) holds

$$<x(t) - Pz(t), \bar{x}> = <S_t[x_o - Pz_o], \bar{x}> \quad .$$

Again, since the continuously differentiable controls are dense in L^2 and $D(A)$ is dense in Z , this result can be extended to hold for all controls of the form (11.14), and then as $\overline{D(F^*)} = X^*$,

$$x(t) - Pz(t) = S_t[x_o - Pz_o]$$

as required.

In propositions 11.2 and 11.3 we have found conditions for the observer $x(t)$ given by (11.6) to asymptotically estimate a function of the state of the origional system. Suppose now, however, that we want to estimate the origional state of the system, $z(t)$. If we take as our estimator $\hat{z}$, a combination of the state of the observer $x(t)$ and the output from the origional system $y(t)$, so that

$$\hat{z}(t) = \hat{M}y(t) + \hat{N}x(t) \quad \text{with } \hat{M} \in \mathcal{L}(Y,Z) \quad , \quad \hat{N} \in \mathcal{L}(X,Z)$$

then

$$\hat{z}(t) - z(t) = \hat{M}y(t) + \hat{N}x(t) - z(t) \quad .$$

We know from proposition 11.2 that, under conditions (11.7),

$$x(t) - Pz(t) = S_t[x_o - Pz_o]$$

so that

$$\hat{z}(t) - z(t) = \hat{M}Cz(t) + \hat{N}Pz(t) - z(t) + \hat{N}S_t[x_o - Pz_o] \quad .$$

Hence, if we require in addition to (11.7) that

(11.17) $\hat{M}C + \hat{N}P = I$

then

$$||\hat{z}(t) - z(t)|| \leq ||\hat{N}|| \; ||S_t[x_o - Pz_o]||$$

and thus if S_t is asymptotically stable

$$\lim_{t\to\infty} ||\hat{z}(t) - z(t)|| = 0 \ .$$

We have therefore proved the following proposition.

<u>Proposition 11.4</u>

$$\hat{z}(t) = \hat{M}y(t) + \hat{N}x(t) \qquad \hat{M} \in \mathcal{L}(Y,Z) \ , \ \hat{N} \in \mathcal{L}(X,Z)$$

where y is given by (11.2) and x is given by (11.6), is an asymptotic estimator of $z(t)$ given by (11.5) if, (11,7) and (11.17) hold, and S_t is an asymptotically stable semigroup.

Propositions 11.2 and 11.3 give conditions under which $x(t) - Pz(t) = S_t[x_o - Pz_o]$ so that if S_t is asymptotically stable then $\lim_{t\to\infty} \{x(t) - Pz(t)\} = 0$ with decay rate given by ω where $||S_t|| \leq Me^{-\omega t}$. Suppose we wish the error to decay at a faster rate than that given by S_t . That means we require to find an operator Q , infinitesimal generator of a strongly continuous semigroup R_t , such that

$$x(t) - Pz(t) = R_t[x_o - Pz_o] \quad \text{where} \quad ||R_t|| \leq \bar{M}e^{-\rho t} \leq Me^{-\omega t} \quad \forall\, t \ .$$

We may then ask the question - Can we choose Q , and thus R_t , arbitrarily and so have the error decaying arbitrarily fast ? In this regard we prove the following.

<u>Proposition 11.5</u>

For general control action $u(t)$ in (11.5), in order that

$$\text{(11.18)} \qquad x(t) - Pz(t) = R_t[x_o - Pz_o]$$

where $z(t)$ is given by (11.5), $x(t)$ is given by (11.6), X a reflexive Banach space, and R_t a strongly continuous semigroup with generator Q , we require that $Q = F$, when we have (11.7) holding. i.e. we can do no better that the decay rate given by F .

If we restrict consideration to feedback controls of the form $u(t) = MCz(t) + Nx(t)$ then (11.18) holds if we have (11.19) holding.

$$(11.19) \quad (i) \quad \langle Pz, Q^*\bar{q}\rangle - \langle PAz, \bar{q}\rangle + \langle GCz, \bar{q}\rangle + \langle HMCz, \bar{q}\rangle - \langle PBMCz, \bar{q}\rangle = 0$$

$$(ii) \quad \langle Fx, \bar{q}\rangle - \langle x, Q^*\bar{q}\rangle + \langle HNx, \bar{q}\rangle - \langle PBNx, \bar{q}\rangle$$

for $z \in D(A)$, $x \in D(F)$, $\bar{q} \in D(Q^*)$.

Thus only if we can satisfy (11.19)(i) with an H such that $HNx \neq PBNx$ will we be able to take Q different from F in this case.

Proof

In this case we have that

$$(11.20) \quad \langle x(t) - Pz(t), \bar{q}\rangle = \langle R_t[x_o - Pz_o], \bar{q}\rangle + \langle [S_t - R_t]x_o, \bar{q}\rangle$$
$$+ \langle [R_tP - PT_t]z_o, \bar{q}\rangle - \int_o^t \langle PT_{t-s}Bu(s), \bar{q}\rangle ds$$
$$+ \int_o^t \langle S_{t-s}[GCz(s) + Hu(s)], \bar{q}\rangle ds \qquad \bar{q} \in D(Q^*).$$

Using the same arguments as were employed in the proof of proposition 11.2 we have, for $\bar{q} \in D(Q^*)$, $z_o \in D(A)$, $x_o \in D(F)$, and $u \in C^1[o,t_1;U]$,

$$\frac{d}{d\rho} \langle R_{t-\rho}PT_\rho z_o, \bar{q}\rangle = \langle R_{t-\rho}PAT_\rho z_o, \bar{q}\rangle - \langle R_{t-\rho}PT_\rho z_o, Q^*\bar{q}\rangle$$

so

$$(11.21) \quad \int_o^t \langle R_{t-\rho}PAT_\rho z_o, \bar{q}\rangle d\rho - \int_o^t \langle R_{t-\rho}PT_\rho z_o, Q^*\bar{q}\rangle d\rho$$

$$= \langle PT_tz_o, \bar{q}\rangle - \langle R_tPz_o, \bar{q}\rangle \quad .$$

Also since

$$\frac{d}{ds} \langle R_{t-s}P\int_o^s T_{s-r}Bu(r)dr, \bar{q}\rangle$$

$$= - \langle R_{t-s}P\int_o^s T_{s-r}Bu(r)dr, Q^*\bar{q}\rangle + \langle R_{t-s}PA\int_o^s T_{s-r}Bu(r)dr, \bar{q}\rangle + \langle R_{t-s}PBu(s), \bar{q}\rangle$$

we have

(11.22) $$\int_o^t \langle R_{t-s}PA\int_o^s T_{s-r}Bu(r)dr, \bar{q}\rangle ds - \int_o^t \langle R_{t-s}P\int_o^s T_{s-r}Bu(r)dr, Q^*\bar{q}\rangle ds$$

$$= \int_o^t \langle PT_{t-s}Bu(s), \bar{q}\rangle ds - \int_o^t \langle R_{t-s}PBu(s), \bar{q}\rangle ds \quad .$$

Similarly,

$$\frac{d}{d\rho} \langle R_{t-\rho}S_\rho x_o, \bar{q}\rangle = - \langle R_{t-\rho}S_\rho x_o, Q^*\bar{q}\rangle + \langle R_{t-\rho}FS_\rho x_o, \bar{q}\rangle$$

leads to

(11.23) $$\int_o^t R_{t-\rho}FS_\rho x_o, \bar{q}\rangle d\rho - \int_o^t \langle R_{t-\rho}S_\rho x_o, Q^*\bar{q}\rangle d\rho$$

$$= \langle S_t x_o, \bar{q}\rangle - \langle R_t x_o, \bar{q}\rangle$$

and as

$$\frac{d}{ds} \langle R_{t-s}\int_o^s S_{s-r}[GCz(r) + Hu(r)]dr, \bar{q}\rangle$$

$$= \langle R_{t-s}F\int_o^s S_{s-r}[GCz(r) + Hu(r)]dr, \bar{q}\rangle + \langle R_{t-s}[GCz(s) + Hu(s)], \bar{q}\rangle$$

$$- \langle R_{t-s} \int_o^s S_{s-r}[GCz(r) + Hu(r)]dr, Q^*\bar{q}\rangle$$

we have

(11.24) $$\int_o^t R_{t-s}F\int_o^s S_{s-r}[GCz(r) + Hu(r)]dr, \bar{q}\rangle ds$$

$$- \int_o^t \langle R_{t-s}\int_o^s S_{s-r}[GCz(r) + Hu(r)]dr, Q^*\bar{q}\rangle ds$$

$$= \int_o^t \langle S_{t-s}[GCz(s) + Hu(s)], \bar{q}\rangle ds$$

$$- \int_o^t \langle R_{t-s}[GCz(s) + Hu(s)], \bar{q}\rangle ds.$$

If we now consider

$$(11.25) \quad <[S_t - R_t]x_o, \bar{q}> + <[R_tP - PT_t]z_o, \bar{q}> - \int_o^t <PT_{t-s}Bu(s), \bar{q}>ds$$
$$+ \int_o^t <S_{t-s}[GCz(s) + Hu(s)], \bar{q}>ds$$

then by adding (11.21) and (11.22) we find that

$$<x(t) - Pz(t), \bar{q}> = <R_t[x_o - Pz_o], \bar{q}>$$

if (11.25) = 0 .

But, combining (11.21), (11.22), (11.23) and (11.24) gives

$$(11.25) = \int_o^t \{<R_{t-s}Pz(s), Q^*\bar{q}> - <R_{t-s}PAz(s), \bar{q}> + <R_{t-s}GCz(s), \bar{q}>\}ds$$
$$+ \int_o^t \{<R_{t-s}Fx(s), \bar{q}> - <R_{t-s}x(s), Q^*\bar{q}>\}ds$$
$$+ \int_o^t \{<R_{t-s}Hu(s), \bar{q}> - <R_{t-s}PBu(s), \bar{q}>\}ds \quad .$$

Thus if we have (11.7) holding, in order that (11.25) = 0 , we further require that

$$\int_o^t \{<R_{t-s}Fx(s), \bar{q}> - <R_{t-s}x(s), Q^*\bar{q}>\}ds = 0$$

which implies $F = Q$ since $R_{t-s}Fx(s) = FR_{t-s}x(s)$ for $x(s) \in D(F)$.
Under these condition we have therefore shown that

$$<x(t) - Pz(t), \bar{q}> = <R_t[x_o - Pz_o], \bar{q}> \quad , \bar{q} \in D(F^*) \ .$$

As in the case of proposition 11.2, the above result was proved for $z_o \in D(A)$, $x_o \in D(F)$ and $u \in C^1[o,t_1;U]$. But since $\overline{D(A)} = Z$, $\overline{D(F)} = X$ and $C^1[o,t_1;U]$ dense in $L^2[o,t_1;U]$ we can by taking sequences of such functions extend the result to hold for all $z_o \in Z$, $x_o \in X$ and $u \in L^2[o,t_1;U]$. Then since $\overline{D(F^*)} = X^*$

$$x(t) - Pz(t) = R_t[x_o - Pz_o] \qquad \text{as required.}$$

If we now consider feedback controls only, of the form $u(t) = MCz(t) + Nx(t)$ (11.25) becomes

$$\text{(11.26)} \quad \int_o^t \{\langle R_{t-s}Pz(s), Q^*\bar{q}\rangle - \langle R_{t-s}PAz(s), \bar{q}\rangle + \langle R_{t-s}GCz(s), \bar{q}\rangle + \langle R_{t-s}HMCz(s), Q^*\bar{q}\rangle - \langle R_{t-s}PBMCz(s), \bar{q}\rangle\}ds$$
$$+ \int_o^t \{\langle R_{t-s}Fx(s), \bar{q}\rangle - \langle R_{t-s}x(s), Q^*\bar{q}\rangle + \langle R_{t-s}HNx(s), \bar{q}\rangle - \langle R_{t-s}PBNx(s), \bar{q}\rangle\}ds$$

and so (11.25) = 0 if (11.19) holds. As previously we can then extend the result so that (11.25) = 0 implies

$$x(t) - Pz(t) = R_t[x_o - Pz_o] \quad .$$

Remark

There is no reason why, in order to satisfy the conditions of propositions 11.2 to 11.5 , X should not be a finite dimensional space. Then $P \in \mathcal{L}(Z,X)$ would have to map the infinite dimensional space Z into the finite dimensional space X, e.g. be some finite dimensional approximation or truncation.

We now consider whether it is possible to design an observer of the form (11.6) to stabilize the original system and prove the following.

Theorem 11.6

If the pair $\{A,B\}$ is stabilizable then there exists an observer of the form (11.6) which stabilizes the system (11.5) when applied as a feedback operator of the form (11.14), provided

(a) S_t is exponentially stable , i.e. there exist constants K , $\omega > 0$ such that $||S_t|| \leq Ke^{-\omega t}$,

(11.27) (b) $MC + NP = D$ where $D \in \mathcal{L}(Z,U)$ and the semigroup generated by $A + BD$ is stable,

(c) (11.16) holds.

<u>Proof</u>

Let V_t be the semigroup generated by $A + BD$ so then

$$V_t z_o = T_t z_o + \int_o^t T_{t-s} BDV_s z_o ds$$

or

$$V_t z_o = T_t z_o + \int_o^t V_{t-s} BDT_s z_o ds$$

using the equivalent, alternative, formulation for the perturbed semigroup. The method for proving the two formulations is identical, and their equivalence is proved for the more general case of evolution operators by Curtain and Pritchard in {9}.

Now, since V_t is stable, there exist constants $\overline{K}$, $\lambda > 0$ such that $||V_t|| \leq \overline{K}e^{-\lambda t}$. $z(t)$ given by (11.5) can then be written in terms of the semigroup V_t as

$$z(t) = V_t z_o + \int_o^t V_{t-s} Bu(s)ds - \int_o^t V_{t-s} BDz(s)ds$$

$$= V_t z_o + \int_o^t V_{t-s} BN[x(s) - Pz(s)]ds$$

since $D = MC + NP$ and $u(t) = MCz(t) + Nx(t)$

$$= V_t z_o + \int_o^t V_{t-s} BNS_s [x_o - Pz_o] ds$$

since (11.16) holds, which implies $x(s) - Pz(s) = S_s[x_o - Pz_o]$. Thus

$$||z(t)|| \leq ||V_t z_o|| + ||\int_o^t V_{t-s} BNS_s [x_o - Pz_o] ds|| .$$

The first term on the right hand side is bounded by $\overline{K}e^{-\lambda t}||z_o||$ and for the second term we have

$$||\int_o^t V_{t-s} BNS_s [x_o - Pz_o] ds|| \leq \int_o^t ||V_{t-s}|| \; ||BN|| \; ||S_s|| \; ||x_o - Pz_o|| ds$$

$$\leq ||BN|| K\overline{K} \int_o^t e^{-\omega(t-s)} e^{-\lambda s} ds \; ||x_o - Pz_o||$$

$$\leq \text{constant } \{e^{-\omega t} - e^{-\lambda t}\} \; ||x_o - Pz_o|| .$$

Hence the result follows.

<u>Remark</u>

Provided we can find an F which satisfies the conditions of theorem 11.6 we can stabilize the system (11.5) by a finite dimensional observer.

As remarked earlier the results of the propositions and theorem of this section remain valid if we consider a finite dimensional observer, but (i.e. X a finite dimensional space), but that then P must map the infinite dimensional space Z into the finite dimensional space X . Pz could be a finite dimensional approximation, or a truncation, of z .

An example of such an observer is given by El Jai in his thesis {12}. He constructs, using the results of Gressang and Lamont {14}, an observer which estimates the first N components of an infinite dimensional system of the form (11.1), with observations (11.2), where the control space $U = \mathbb{R}^p$ and the observation space $Y = \mathbb{R}^q$. He takes his observer to be defined by an equation of the form (11.3) with $G \in \mathcal{L}(L^2[t_o,t_1;\mathbb{R}^q],l^2)$ and $H \in \mathcal{L}(L^2[t_o,t_1;\mathbb{R}^p],l^2)$.

Since he takes the differential form of the equations as his system description he requires, for Pz to be asymptotically estimated by x, that $P \in \mathcal{L}(D(A),D(F))$ and that the controls $u(t)$ be sufficiently smooth so that $z(t)$ and $x(t)$ are well defined by (11.1) and (11.3) respectively. We have shown that by considering the mild solutions (11.5) and (11.6) we can consider controls $u(t) \in L^2[o,t_1;U]$ and that then if condition (11.7) holds, $x(t) - Pz(t) = S_t[x_o - Pz_o]$.

El Jaï assumes that A, B and C in (11.1) can be written as

$$A = \begin{bmatrix} A_N & 0 \\ 0 & A_r \end{bmatrix} \qquad B = \begin{bmatrix} B_N \\ B_r \end{bmatrix} \qquad C = [C_N \;,\; C_r]$$

where the subscript N denotes the projection onto a subspace of dimension N and the subscript r denotes the residual part. Then writing

$z(t) = \begin{bmatrix} z_N(t) \\ z_r(t) \end{bmatrix}$ the mild solution to (11.1) can be represented in the form

$$\text{(11.28)} \quad \begin{aligned} &\text{(i)} \quad z_N(t) = T_{A_N}(t-t_o)z_N(t_o) + X_N^t(u) \\ &\text{(ii)} \quad z_r(t) = T_{A_r}(t-t_o)z_r(t_o) + X_r^t(u) \end{aligned}$$

where $T_{A_N}(t)$ is the semigroup generated by A_N and similarly T_{A_r} is the semigroup generated by A_r, and where

$$(11.29)\quad \begin{aligned} &(i) \quad X_N^t(u) = \int_{t_o}^{t} T_{A_N}(t-\tau)B_N u(\tau)d\tau \\ &(ii) \quad X_r^t(u) = \int_{t_o}^{t} T_{A_r}(t-\tau)B_r u(\tau)d\tau \quad . \end{aligned}$$

If we similarly choose

$$F = \begin{bmatrix} F_N & 0 \\ 0 & F_r \end{bmatrix} \qquad G = \begin{bmatrix} G_N \\ G_r \end{bmatrix} \qquad H = \begin{bmatrix} H_N \\ H_r \end{bmatrix}$$

then the mild solution to (11.6), writing $x(t) = \begin{bmatrix} x_N(t) \\ x_r(t) \end{bmatrix}$ has the form

$$(11.30)\quad \begin{aligned} &(i) \quad x_N(t) = S_{F_N}(t-t_o)x_N(t_o) + \int_{t_o}^{t} S_{F_N}(t-\tau)[G_N y(\tau) + H_N u(\tau)]d\tau \\ &(ii) \quad x_r(t) = S_{F_r}(t-t_o)x_r(t_o) + \int_{t_o}^{t} S_{F_r}(t-\tau)[G_r y(\tau) + H_r u(\tau)]d\tau \end{aligned}$$

where S_{F_N} and S_{F_r} are the semigroups generated by F_N and F_r respectively. Also let

$$P = \begin{bmatrix} P_N & P_r \\ P_s & P_t \end{bmatrix} \quad .$$

As an immediate consequence of propositon 11.2 we have the following.

<u>Theorem 11.7</u>

$x_N(t)$ given by (11.30)(i) (i.e. the projection of $x(t)$ onto the subspace of dimension N) is an asymptotic estimator for $P_N z_N(t) + P_r z_r(t)$ provided S_{F_N} is asymptotically stable and

$$\text{(i)} \quad \langle P_N z_N, \overset{*}{F}_N \bar{x}_N \rangle - \langle P_N A_N z_N, \bar{x}_N \rangle + \langle G_N C_N z_N, \bar{x}_N \rangle = 0$$

$$\text{(11.31)} \quad \text{(ii)} \quad \langle P_r z_r, \overset{*}{F}_N \bar{x}_N \rangle - \langle P_r A_r z_r, \bar{x}_N \rangle + \langle G_N C_r z_r, \bar{x}_N \rangle = 0$$

$$\text{(iii)} \quad H_N = P_N B_N + P_r B_r$$

for all $z_N \in D(A_N)$, $z_r \in D(A_r)$, $\bar{x}_N \in D(F_N^*)$.

<u>Remark</u>

Theorem 11.7 extends the result given by El Jai (in which equations (11.31) are written in operator form) to systems described by the more general mild solutions rather than those which satisfy the differential form.

Prado {29} considers observers of the same form, but takes P_N = identity, and so for his type of systems we have that x_N estimates $z_N(t) + P_r z_r(t)$ asymptotically if, S_{F_N} is stable and

$$\text{(i)} \quad \langle z_N, \overset{*}{F}_N \bar{x}_N \rangle - \langle A_N z_N, \bar{x}_N \rangle + \langle G_N C_N z_N, \bar{x}_N \rangle = 0$$

$$\text{(11.32)} \quad \text{(ii)} \quad \langle P_r z_r, \overset{*}{F}_N \bar{x}_N \rangle - \langle P_r A_r z_r, \bar{x}_N \rangle + \langle G_N C_r z_r, \bar{x}_N \rangle = 0$$

$$\text{(iii)} \quad H_N = B_N + P_r B_r$$

for $z_N \in D(A_N)$, $z_r \in D(A_r)$, $\bar{x}_N \in D(F_N^*)$.

Then, in the case considered by Prado, where $P: D(A) \rightarrow D(F)$, $A = (A_{ij})$ and $F = (f_{ij})$ are diagonal, (11.32)(ii) implies that

$$f_{ii}P_{ij} - P_{ij}A_{jj} = - \sum_{l=1}^{\infty} G_{il}C_{lj} \qquad \begin{array}{l} i = 1,\ldots N \\ j = N+1,\ldots \infty \end{array}$$

for $G = (G_{ij})$ and $C = (C_{ij})$

and so

$$P_{ij} = - (f_{ii} - A_{jj})^{-1} \sum_{l=1}^{\infty} G_{il}C_{lj}$$

and thus he shows that P has been found explicitly, with conditions for rendering the term $P_r z_r$ negligable .

El Jai also considers under what conditions x_N asymptotically estimates $P_N z_N$ and proves a result equivalent to the following, for the systems he considers.

<u>Theorem 11.8</u>

If S_{F_N} is asymptotically stable and

(11.33)

(i) $<P_N z_N, \overleftarrow{F_N^* \bar{x}_N}> - <P_N A_N z_N, \bar{x}_N> + <G_N C_N z_N, \bar{x}_N>$

(ii) $H_N = P_N B_N + G_N C_r X_r$

for all $z_N \in D(A_N)$, $\bar{x}_N \in D(F_N^*)$, with X_r given by (11.29), and if

$$\lim_{t\to\infty} V_N(t) = 0$$

where

(11.34) $$V_N(t) = \int_{t_o}^{t} S_{F_N}(t-\tau)G_N C_r T_{A_r}(\tau-t_o)x_r(t_o)d\tau$$

then $x_N(t)$ given by (11.30)(i) estimates $P_N z_N(t)$ asymptotically.

<u>Proof</u>

$x_N(t) - P_N z_N(t)$

$$= S_{F_N}(t-t_o)x_N(t_o) + \int_{t_o}^{t} S_{F_N}(t-\tau)[G_N\{C_N z_N(\tau) + C_r z_r(\tau)\} + H_N u(\tau)]d\tau$$

$$- P_N T_{A_N}(t-t_o) z_N(t_o) - \int_{t_o}^{t} T_{A_N}(t-\tau) B_N u(\tau) d\tau$$

$$= S_{F_N}(t-t_o) x_N(t_o) + \int_{t_o}^{t} S_{F_N}(t-\tau)[G_N\{C_N z_N(\tau) + C_r X_r^{\tau}(u)\} + H_N u(\tau)] d\tau$$

$$+ \int_{t_o}^{t} S_{F_N}(t-\tau) G_N C_r T_{A_r}(\tau-t_o) z_r(t_o) d\tau - P_N T_{A_N}(t-t_o) z_N(t_o)$$

$$- \int_{t_o}^{t} T_{A_N}(t-\tau) B_N u(\tau) d\tau \quad .$$

Using the results of proposition 11.2, if (11.33) holds then,

$$x_N(t) - P_N z_N(t) = S_{F_N}[x_o - P_N z_N(t_o)] + \int_{t_o}^{t} S_{F_N}(t-\tau) G_N G_r T_{A_r}(\tau-t_o) z_r(t_o)$$

$$= S_{F_N}[x_o - P_N z_N(t_o)] + V_N(t) \quad .$$

Thus if S_{F_N} is asymptotically stable, and $V_N(t) \to 0 \ (t \to \infty)$, then x_N estimates $P_N z_N$ asymptotically.

El Jaï further proves the following lemma for the parabolic systems he considers in his thesis.

<u>Lemma 11.9</u>

If F is stable and A is parabolic then $\lim_{t\to\infty} V_N(t) = 0$.

Remark

El Jaï notes the following two points.

1. The asymptotic convergence is independent of the initial state of the observer and of the system.
2. The observer answers a real problem - How to reconstruct N components of the state of a system from q continuous observations at the same time not ignoring completely the dynamics of the neglected modes.

Example 11.10 - Estimator of the first N components in an eigenfunction expansion.

Let A, the infinitesimal generator of a strongly continuous semigroup, be self-adjoint on the Hilbert space Z, with $R(\lambda_o, A)$ compact for some $\lambda_o \in \rho(A)$. Then we know (see for example {17}):

(a) there exists an infinite sequence λ_j of distinct eigenvalues of A such that $\lambda_j \to \infty$ as $j \to \infty$, each with finite multiplicity r_j. Furthermore, since A generates a strongly continuous semigroup, the λ_j are bounded above and can be ordered, i.e. $\lambda_n < \ldots < \lambda_2 < \lambda_1 \leq$ constant ;

(b) there exists a complete orthonormal set of eigenvectors $\{\phi_{jk}\}$ of A ;.

(c) from the unique expansion

$$z = \sum_{j=1}^{\infty} \sum_{k=1}^{r_j} \langle z, \phi_{jk} \rangle \phi_{jk}$$

we have, for $z \in D(A)$

$$Az = \sum_{j=1}^{\infty} \lambda_j \sum_{k=1}^{r_j} \langle z, \phi_{jk} \rangle \phi_{jk} \qquad ;$$

(d) the semigroup generated by A is given by

$$T_t z = \sum_{j=1}^{\infty} e^{\lambda_j t} \sum_{k=1}^{r_j} \langle z, \phi_{jk} \rangle \phi_{jk} \qquad .$$

Suppose now that the eigenvalues all have multiplicity 1, then the control problem

$$z(t) = T_t z_o + \int_o^t T_{t-s} Bu(s) ds \tag{11.35}$$

can be written as

$$\sum_{j=1}^{\infty} z_j \phi_j = \sum_{j=1}^{\infty} e^{\lambda_j t} z_{oj} \phi_j + \int_o^t \sum_{j=1}^{\infty} e^{\lambda_j (t-s)} (Bu(s))_j \phi_j ds \tag{11.36}$$

where

$$z_j = \langle z, \phi_j \rangle \quad , \quad (Bu)_j = \langle Bu, \phi_j \rangle \quad \text{etc.} \quad .$$

Let the observation equation be

(11.37) $y(t) = Cz(t) = C \sum_{j=1}^{\infty} z_j \phi_j$.

If we consider the first N eigenvectors $\{\phi_n , n = 1,\ldots,N\}$, they span an N-dimensional space X . Let F , the infinitesimal generator of the semigroup S_t , be the operator on X with eigenvalue expansion

(11.38) $F\phi_n = \sigma_n \phi_n$

where σ_n , $n = 1,\ldots,N$ are distinct. Any $x \in X$ has the expansion

(11.39) $x = \sum_{j=1}^{N} \langle x, \phi_j \rangle \phi_j = \sum_{j=1}^{N} x_j \phi_j$,

and the semigroup S_t is given by

$$S_t x = \sum_{j=1}^{N} e^{\sigma_j t} \langle x, \phi_j \rangle \phi_j \quad .$$

The observer

(11.40) $x(t) = S_t x_o + \int_o^t S_{t-s}[Gy(s) + Hu(s)]ds$

can therefore be written as

(11.41) $\sum_{j=1}^{N} x_j(t)\phi_j = \sum_{j=1}^{N} e^{\sigma_j t} x_{oj}\phi_j + \int_o^t \sum_{j=1}^{N} e^{\sigma_j (t-s)} [(GC)_j z_j(s) + (Hu(s))_j]ds$

where $(Hu)_j = \langle Hu, \phi_j \rangle$ etc. as before, and where we assume we can write $GC \sum_{j=1}^{\infty} z_j \phi_j = \sum_{j=1}^{\infty} (GC)_j z_j \phi_j$.

If we now take

$$Pz(t) = \sum_{j=1}^{N} z_j \phi_j$$

then from propositon 11.2

$$x(t) - Pz(t) = S_t[x_o - Pz_o]$$

if

$$(11.42)\quad \begin{cases} \text{(i)} \quad \langle PAz, \bar{x}\rangle - \langle Pz, F^* \bar{x}\rangle = \langle GCz, \bar{x}\rangle \ , \ \bar{x} \in D(F^*) \\ \text{(ii)} \quad Hu - PBu = 0 \quad \forall\, u \in U \ . \end{cases}$$

Re-writing this result in terms of the eigenvalues and eigenvectors for the systems we have

$$\sum_{j=1}^{N} x_j \phi_j - \sum_{j=1}^{N} z_j \phi_j = \sum_{j=1}^{N} e^{\sigma_j t}[x_{oj} - z_{oj}]\phi_j$$

i.e.

$$x_j - z_j = e^{\sigma_j t}[x_{oj} - z_{oj}] \qquad j = 1,\ldots,N \ ,$$

if

$$\text{(i)} \quad \sum_{j=1}^{N} \lambda_j z_j \phi_j - \sum_{j=1}^{N} \sigma_j z_j \phi_j = \sum_{j=1}^{N} (GC)_j z_j \phi_j$$

i.e.

$$\lambda_j - \sigma_j = (GC)_j \qquad j = 1,\ldots,N$$

and

$$\text{(ii)} \quad (Hu)_j = (Bu)_j \qquad j = 1,\ldots,N \ .$$

12. THE EFFECT OF OBSERVERS IN THE LINEAR QUADRATIC COST CONTROL PROBLEM

We now turn our attention to the use of observers in the linear quadratic cost control problem, and investigate the increase in cost which arises when an observer is used to drive the original system, rather than feeding back the state of the original system. This is an area which, to our knowledge, is as yet uninvestigated for distributed parameter systems, but which has been studied for finite dimensional systems by, amongst others, Sarma and Jayaraj {35}, Newmann {27}, Borigiorno Jr. and Youla { 7 } and Yüksel and Borigiorno Jr. {42} .

We consider the control problem of section 1, (12.1) $z(t) = T_t z_o + \int_o^t T_{t-s} Bu(s)ds$

with the performance index given by

(12.2) $J(u) = \langle z(t_1), Gz(t_1)\rangle + \int_o^{t_1} \{\langle z(s), Wz(s)\rangle + \langle u(s)\ Ru(s)\rangle\} ds$

where, as in section 1, $B \in \mathcal{L}(U,H)$, and G , W and R satisfy the usual conditions. Then we know, {11}, there exists a unique optimal control $u^*(t) = - R^{-1}B^*Q(t)z(t)$ where $Q(t)$ is a bounded linear self-adjoint operator, the unique solution to the Riccati equation

(12.3) $Q(t)h = \mathcal{U}^*(t_1,t)G\mathcal{U}(t_1,t)h + \int_t^{t_1} \mathcal{U}^*(s,t)\{W + Q(s)BR^{-1}B^*Q(s)\}\mathcal{U}(s,t)hds$

with $\mathcal{U}(s,t)$ the evolution operator defined by

(12.4) $\mathcal{U}(t,s)h = T_{t-s}h - \int_s^t T_{t-\rho}BR^{-1}B^*Q(\rho)\mathcal{U}(\rho,s)hd\rho$.

Furthermore the cost of the control u^* is given by

(12.5) $J(u^*) = \langle z_o, Q(o)z_o\rangle$

and if u is any other control the difference in costs is

(12.6) $J(u) - J(u^*) = \int_o^{t_1} \langle \bar{u}, R\bar{u}\ ds\rangle$

where $u = u^* + \bar{u}$.

Thus if we take as u the control $u(t) = My(t) + Nx(t)$ where y is given by (11.2) and x given by (11.6), the increase in cost due to not using the optimal control u^* is

$J(u) - J(u^*)$

$= \int_o^{t_1} \langle u(t) - u^*(t), R\{u(t) - u^*(t)\}\rangle dt$

$$= \int_0^{t_1} \langle MCz(t) + Nx(t) + R^{-1}B^*Q(t)z(t),\ R\{MCz(t) + Nx(t) + R^{-1}B^*Q(t)z(t)\}\rangle dt$$

$$= \int_0^{t_1} \langle MCz(t) + N\{e(t) + Pz(t)\} + R^{-1}B^*Q(t)z(t),$$
$$R\{MCz(t) + N\{e(t) + Pz(t)\} + R^{-1}B^*Q(t)z(t)\}\rangle dt$$

where $e(t) = x(t) - Pz(t)$.

Hence

(12.7) $$J(u) - J(u^*) = \int_0^{t_1} \langle \{MC + NP\}z(t) + R^{-1}B^*Q(t)z(t) + Ne(t),$$
$$R\{\{MC + NP\}z(t) + R^{-1}B^*Q(t)z(t) + Ne(t)\}\rangle dt\ .$$

If we choose the observer to stabilize the system we require from (11.27)(b) that $MC + NP = D$ where $A + BD$ stable, so then

(12.8) $$J(u) - J(u^*) = \int_0^{t_1} \langle \{D + R^{-1}B^*Q(t)\}z(t) + Ne(t),$$
$$R\{\{D + R^{-1}B^*Q(t)\}z(t) + Ne(t)\}\rangle dt\ .$$

Over the infinite time interval, where $G = 0$ in (12.2), $Q(t) = Q$, independent of t and we know that $D = - R^{-1}B^*Q$ stabilizes the system. This D is therefore an obvious choice in this case and then

$$J(u) - J(u^*) = \int_0^{t_1} \langle Ne(t),\ RNe(t)\rangle dt$$

$$= \int_0^{t_1} \langle N[x(t) - Pz(t)],\ RN[x(t) - Pz(t)]\rangle dt$$

$$= \int_0^{t_1} \langle NS_t[x_o - Pz_o],\ RNS_t[x_o - Pz_o]\rangle dt$$

Clearly the increase in cost depends on the initial state of the system which is, in general, unknown.

<u>Remark</u>

As with the case of the results of section 11, the above remains valid when x is a finite dimensional observer and P maps the infinite dimensional state space Z into the finite dimensional space X.

<u>Example 12.1</u> - Observers of Prado type.

Consider $x_N(t)$ given by (11.31), then from theorem 11.7,

(12.9) $$x_N(t) = z_N(t) + P_r z_r(t) + S_{F_N}(t-t_o)[x_N(t_o) - z_N(t_o) + P_r z_r(t_o)]$$

where $z_N(t)$ and $z_r(t)$ are given by (11.28), provided the conditions (11.32) are satisfied.

If u^* is the unique optimal control which minimizes (11.2) with

$z(t) = \begin{bmatrix} z_N(t) \\ z_r(t) \end{bmatrix}$ given by (11.28) then $u^* = - Kz = - R^{-1}B^*Q(t)z$

and under the partitioning assumptions of Prado on the operators A, B and C it is easy to see that K must have the form $[K_N \ . \ K_r]$ so that $u^* = - \{K_N z_N + K_r z_r\}$. Suppose that instead of using z to feedback we use x_N and so take as our control $u = - K_N x_N$ where we assume that (11.32) holds so the observer is one of Prado type satisfying (12.9).

Then from (12.6)

$$\Delta J = J(u) - J(u^*) = \int_o^{t_1} ||R^{\frac{1}{2}}(u^* - u)||^2 dt$$

$$= \int_o^{t_1} ||R^{\frac{1}{2}}\{K_N x_N(t) - K_N z_N(t) - K_r z_r(t)\}||^2 dt$$

so

$$\Delta J = \int_o^{t_1} ||R^{\frac{1}{2}}\{K_N\{P_r z_r(t) + S_{F_N}(t-t_o)[x_N(t_o) - z_N(t_o) - P_r z_r(t_o)]\} - K_r z_r(t)\}||^2 dt$$

from (12.9)

$$\leq \int_o^{t_1} ||R^{\frac{1}{2}}\{K_N P_r z_r(t) - K_r z_r(t)\}||^2 dt$$

$$+ \int_o^{t_1} ||R^{\frac{1}{2}} K_N S_{F_N}(t-t_o)[x_N(t_o) - z_N(t_o) - P_r z_r(t_o)]||^2 dt \quad .$$

In the above the first term represents the increase in cost due to the neglected modes. The second term is the increase due to using an observer as feedback, rather than the state of the system, in the first N modes. Therefore, to minimize ΔJ, the best we can do is to minimize

$$\int_0^{t_1} ||R^{\frac{1}{2}} K_N S_{F_N}(t-t_o)[x_N(t_o) - z_N(t_o) - P_r z_r(t_o)]||^2 dt$$

but this requires knowledge of the initial state of the system $\begin{bmatrix} z_N(t_o) \\ z_r(t_o) \end{bmatrix}$ which is, in general, unknown.

Example 12.2 - Observers of El Jai type.

El Jai considers observers that are estimators of $P_N z_N(t)$ and from the proof of theorem 11.8 we have that provided (11.33) holds,

$$x_N(t) - P_N z_N(t) = S_{F_N}(t-t_o)[x_N(t_o) - P_N z_N(t_o)] + V_N(t)$$

where

$$V_N(t) = \int_{t_o}^{t} S_{F_N}(t-\tau) G_N C_r T_{A_r}(\tau - t_o) d\tau \; z_r(t_o)$$

with z_N given by (11.28) and x_N given by (11.30), and where the operators are as in the previous section. As in example 12.1 the optimal control for the regulator problem has the form $u^* = -\{K_N z_N + K_r z_r\}$. Suppose however we take as our control $u = x_N$ and we let $P_N = K_N$ so that x_N is an estimator of $K_N z_N$, then the increase in cost due to using the control u, rather than the optimal control, in the regulator problem is

$$\Delta J = \int_0^{t_1} ||R^{\frac{1}{2}}[x_N(t) - P_N z_N(t) - K_r z_r(t)]||^2 dt$$

$$\leq \int_0^{t_1} ||R^{\frac{1}{2}}\{V_N(t) - K_r x_r(t)\}||^2 dt + \int_0^{t_1} ||R^{\frac{1}{2}} S_{F_N}(t-t_o)[x_N(t_o) - K_N z_N(t_o)]||^2 dt$$

Again the first term represents the increase in cost due to the neglected modes and the second that due to using the observer rather than state feedback on the first N modes, dependent on the initial state of the system.

Example 12.3

We now consider again the system of example 11.10. The optimal control for the regulator problem is of the form $u^* = -Kz$ where z is given by (11.36). Suppose now that instead of feeding back z we feedback the observer x given by (11.41), and so $u = -Kx$. Then

$$\Delta J = \int_0^{t_1} ||R^{\frac{1}{2}}K\{\sum_{j=1}^{\infty} z_j\phi_j - \sum_{j=1}^{N} x_j\phi_j\}||^2 ds$$

$$\leq \int_0^{t_1} ||R^{\frac{1}{2}}K\sum_{j=1}^{N} e^{\sigma_j s}\{x_{oj} - z_{oj}\}\phi_j||^2 ds + \int_0^{t_1} ||R^{\frac{1}{2}}K\sum_{j=N+1}^{\infty} z_j\phi_j||^2 ds \quad .$$

Here the first term represents the error due to using an observer and the second that due to the neglected states z_j , $j = N+1,\ldots$.

As has been illustrated by the above three examples, in order to minimize the increase in cost due to using an observer rather than the optimal control in the regulator problem we need to minimize a term of the form

$$(12.10) \qquad \int_0^{t_1} ||R^{\frac{1}{2}}KS_t[x_o - Pz_o]||^2 dt \quad .$$

If we know the initial state of the system z_o then by choosing $x_o = Pz_o$ we automatically have this term zero. In general however the initial state z_o is unknown and thus we want to choose x_o to "minimize", in some sense, (12.10) , over all initial states z_o .

If the statistics of the probability distribution for the initial values, z_o, is known then we might choose x_o so as to minimize the expected value of (12.10). Alternatively we might consider minimizing

$$\frac{(12.10)}{\int_o^{t_1} ||R^{\frac{1}{2}}KS_tPz_o||^2 dt} .$$

13. OBSERVERS FOR SYSTEMS WITH UNBOUNDED CONTROL ACTION

Consider the system

$$(13.1) \qquad z(t) = T_t z_o + \int_o^t T_{t-s} Bu(s)ds$$

where T_t is a strongly continuous semigroup, with generator A, on the Banach space Z. Suppose also that there exists a Banach space $\tilde{W}$, with Z dense in $\tilde{W}$, such that (2.4) holds. We then know from section 2 that $z(t)$ is well defined for $u \in L^q[o,t_1;U]$ and if $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ (where recall $g \in L^p[o,t_1;Z]$ in (2.4)(d)), then $z \in L^r[o,t_1;Z]$; with $z \in C[o,t_1;Z]$ when $\frac{1}{p} + \frac{1}{q} = 1$. Furthermore $z(t)$ given by (13.1) is the weak solution to a boundary control problem if we assume the existence of a Green's formula of the form (2.7).

Suppose now we have the observation equation

$$(13.2) \qquad y(t) = Cz(t)$$

where $C \in \mathcal{L}(Z,Y)$, Y a Banach space, and we define an observer by

$$(13.3) \qquad x(t) = S_t x_o + \int_o^t S_{t-s}\{Gy(s) + Hu(s)\}ds$$

where S_t is a strongly continuous semigroup with generator F on the Banach space X and $G \in \mathcal{L}(Y,X)$.

There are two alternatives, we may consider (13.3) as a distributed control system with $H \in \mathcal{L}(U,X)$ or again as the weak solution to a boundary control problem. In the latter case we assume the existence of a Banach space $\tilde{W}$ with

$$(13.4)\quad \begin{array}{ll} \text{(a)} & \tilde{W} \supset R(H) \\ \text{(b)} & H \in \mathcal{L}(U,\tilde{W}) \\ \text{(c)} & S_t \in \mathcal{L}(\tilde{W},X) \quad , \quad t > 0 \\ \text{(d)} & ||S_t\tilde{w}||_X \leq h(t)||\tilde{w}||_{\tilde{W}} \quad \text{for all } \tilde{w} \in \tilde{W} \text{ with } h \in L^p[o,t_1] \end{array}$$

Again, from section 2, we have that $x(t)$ is well defined by (13.3) when condition (13.4) holds, for $u \in L^q[o,t_1;U]$ with $x \in C[o,t_1;X]$ when $\frac{1}{p} + \frac{1}{q} = 1$.

We now consider what are conditions which are sufficient for $x(t)$ given by (13.3), with either distributed or unbounded (boundary) control action, to be an asymptotic estimator of $Pz(t)$, where $z(t)$ is the solution to (13.1) with T_t , B satisfying (2.4).

<u>Observers with Distributed Control Action</u>

Let $x(t)$ be given by (13.3) where $H \in \mathcal{L}(U,X)$. In order that $x(t)$ be an asymptotic estimator of $Pz(t)$ take

$$(13.5)\quad P \in \mathcal{L}(\tilde{W},X)$$

and then we can easily show that the results of propositions 11.2 and 11.3 remain valid. The formal proofs of the two propositions are the same and then condition (13.5) ensures that the formal manipulations are justified and that all the terms are well defined.

Observers with Unbounded Control Action

Let $x(t)$ be given by (13.3) where S_t, H satisfy (13.4). In this case in order that $x(t)$ be an asymptotic estimator of $Pz(t)$ we take

$$\text{(13.6)} \qquad P \in \mathcal{L}(\bar{W}, \underset{\sim}{W})$$

and then again we have that the results of propositions 11.2 and 11.3 remain valid, the formal proofs once more being the same with the formal manipulations being justified by lemmas 4.6, 4.7, and 4.8 and condition (13.6) . (13.6) also ensures that all the terms are well defined.

Remark

We require condition (13.5) in the distributed control case and (13.6) when considering observers with unbounded control action in order that terms of the form

$$\int_o^t \{\langle S_{t-s}Hu(s), \bar{x}\rangle - \langle S_{t-s}PBu(s), \bar{x}\rangle\}ds$$

are well defined.

Example 13.1

Now we consider the construction of an observer for the inventory control problem with stochastic deterioration rate, as presented in section 9.1. Recall that in this problem the quantity, in stock, of a product with deterioration state ζ at time t is $z(\zeta,t)$ satisfying

$$\text{(13.7)} \qquad \frac{\partial z}{\partial t} + \mu\frac{\partial z}{\partial \zeta} - \frac{\sigma^2}{2}\frac{\partial^2 z}{\partial \zeta^2} + f(\zeta,t) = 0 \qquad \zeta \in (0,1)$$

$$z(o,t) = u(t) \quad , \quad z(\zeta,o) = z_o(\zeta) \quad , \quad z(1,t) = 0$$

where $f(\zeta,t)$ is the demand function, the deterioration in the time interval $(t,t+\Delta t)$ being given by $\mu\Delta t + \sigma\Delta\omega(t)$, $\omega(t)$ standard Brownian motion.

From section 9 we know that (13.7) has a weak solution given by

$$(13.8) \qquad z(t) = T_t z_o + \int_o^t T_{t-s} Bu(s) ds$$

which is the mild solution to

$$(13.9) \qquad z_t = Az + Bu$$

$$z(o,t) = 0 \quad , \quad z(1,t) = 0 \quad , \quad z(\zeta,o) = z_o(\zeta) \quad ,$$

where T_t is the semigroup generated by $A = \frac{\sigma^2}{2} \frac{\partial^2}{\partial\zeta^2} - \mu\frac{\partial}{\partial\zeta}$, and $B = -\frac{1}{2}\sigma^2\delta'(\zeta)$.

It is easy to show that A has eigenvalues $\lambda_n = -\{\frac{\mu^2}{\sigma^2} + \frac{n^2\pi^2\sigma^2}{2}\}$ with corresponding eigenvectors $\phi_n = \sqrt{2}\, e^{(\mu/\sigma^2)\zeta}\sin n\pi\zeta$. If we also consider the adjoint operator A^*, then A^* has eigenvectors $\psi_n = \sqrt{2}\, e^{-(\mu/\sigma^2)\zeta}\sin n\pi\zeta$ (corresponding to λ_n) and $\{\phi_n,\psi_n\}$ form a biorthogonal sequence, i.e. $<\phi_n, \psi_m> = \delta_m^n$. The system (13.9) can thus be written as

$$(13.10) \qquad \sum_{j=1}^{\infty} \dot{z}_j\phi_j = \sum_{j=1}^{\infty} \lambda_j z_j \phi_j + \sum_{j=1}^{\infty} (Bu)_j\phi_j$$

where $z_j = <z, \psi_j>$, $(Bu)_j = <Bu, \psi_j>$, also the mild solution then takes the form

$$(13.11) \qquad z(t) = \sum_{j=1}^{\infty} z_j\phi_j = \sum_{j=1}^{\infty} e^{\lambda_j t} z_{oj}\phi_j + \int_o^t \sum_{j=1}^{\infty} e^{\lambda_j(t-s)} (Bu(s))_j\phi_j \, ds \quad .$$

Suppose now that we wish to construct an observer to asymptotically estimate the first N states of $z(t)$, given by (13.11), i.e. we want $x(t)$ to be such that $x(t) - Pz(t)$ tends to zero as $t \to \infty$, where

$$(13.12) \qquad Pz(t) = \sum_{j=1}^{N} z_j \phi_j \quad .$$

We also assume that we have an observation equation

$$(13.13) \qquad y(t) = Cz(t) \quad .$$

If we consider the first N eigenfunctions of A, $\{\phi_n, n = 1,\ldots,N\}$, they span an N-dimensional space X . Let F be the operator on X with eigenfunction expansion

$$(13.14) \qquad F\phi_n = \sigma_n \phi_n \quad .$$

Any $x \in X$ has the expansion

$$x = \sum_{j=1}^{N} \langle x, \psi_j \rangle \phi_j = \sum_{j=1}^{N} x_j \phi_j \quad .$$

Consider an observer for $Pz(t)$ of the form (13.13) where S_t is the semigroup generated by F , i.e.

$$(13.15) \qquad x(t) = S_t x_o + \int_o^t S_{t-s}\{Gy(s) + Hu(s)\}ds$$

which can be written in the form

$$(13.16) \qquad x(t) = \sum_{j=1}^{N} x_j \phi_j = \sum_{j=1}^{N} e^{\sigma_j t} x_{oj} \phi_j + \int_o^t \sum_{j=1}^{N} e^{\sigma_j (t-s)} \{(GCz)_j + (Hu)_j\}ds$$

where again $x_{oj} = \langle x_o, \psi_j \rangle$, $(GCz)_j = \langle GCz(s), \psi_j \rangle$, and $(Hu)_j = \langle Hu(s), \psi_j \rangle$.

If we take $P\tilde{w} = \sum_{j=1}^{N} \tilde{w}_j \phi_j$, where $\tilde{w}_j = \langle \tilde{w}, \psi_j \rangle$, for all $\tilde{w} \in \tilde{W}$ then $P \in \mathcal{L}(\tilde{W},X)$. Condition (13.5) thus holds and so we know, using proposition 11.2 that

$$(13.17) \qquad x(t) - Pz(t) = S_t[x_o - Pz_o]$$

if

$$(13.18)\quad \langle PAz, \bar{x}\rangle_{X,X^*} - \langle Pz, F^*\bar{x}\rangle_{X,X^*} = \langle GCz, \bar{x}\rangle_{X,X^*}$$

and

$$(13.19)\quad Hu - PBu = 0 \quad \forall\, u \in U \,.$$

Then, in this case, (13.17) is equivalent to

$$\sum_{j=1}^{N} x_j\phi_j - \sum_{j=1}^{N} z_j\phi_j = \sum_{j=1}^{N} \{x_{oj} - z_{oj}\}\phi_j e^{\sigma_j t}$$

i.e.

$$(13.20)\quad x_j - z_j = e^{\sigma_j t}\{x_{oj} - z_{oj}\} \;,\quad j = 1,\ldots,N \;.$$

Similarly (13.18) becomes

$$\sum_{j=1}^{N} \lambda_j z_j\phi_j - \sum_{j=1}^{N} \sigma_j z_j\phi_j = \sum_{j=1}^{N} (GCz)_j\phi_j$$

which implies

$$(13.21)\quad \lambda_j z_j - \sigma_j z_j = (GCz)_j \;,\quad j = 1,\ldots,N \;.$$

Also (13.19) reduces to

$$(13.22)\quad (Hu)_j = (Bu)_j \;,\quad j = 1,\ldots,N \;.$$

If we further assume that GC can be chosen such that $(GCz)_j = (GC)_j z_j$ then (13.21) reduces still further to

$$(13.23)\quad \lambda_j - \sigma_j = (GC)_j \;.$$

Hence we have shown that $x(t)$ given by (13.15) is an asymptotic estimator of $Pz(t)$ given by (13.12), (13.11), if we choose the operators F, G and H such that (13.14), (13.21) (or (13.23)) and (13.22) hold with F the generator of a stable semigroup.

One possible form for the observation equation (13.2) is

$$(13.24)\quad y(\zeta,t) = \int_0^1 b(\zeta,\zeta')z(\zeta',t)d\zeta'$$

i.e.

$$Cz(\zeta) = \int_0^1 b(\zeta,\zeta')z(\zeta')d\zeta'$$

and we might take G as

(13.25) $G = g(\zeta)$ with $g \in L^2(0,1)$.

From (13.24), since $z = \sum_{j=1}^{\infty} z_j\phi_j$

$$Cz(\zeta) = \int_0^1 b(\zeta,\zeta') \sum_{j=1}^{\infty} z_j\phi_j(\zeta')d\zeta'$$

$$= \sum_{j=1}^{\infty} z_j\int_0^1 b(\zeta,\zeta')\phi_j(\zeta')d\zeta'$$

and since $g \in L^2(0,1)$, we may expand G as

$$G = g(\zeta) = \sum_{j=1}^{\infty} g_j\phi_j(\zeta) \quad \text{where} \quad g_j = \langle g, \psi_j\rangle$$

then

$$GCz = \sum_{j=1}^{\infty} g_j\{ \sum_{k=1}^{\infty} z_k\int_0^1 b(\zeta,\zeta')\phi_k(\zeta')d\zeta'\}\phi_j$$

and thus (13.21) becomes

$$\lambda_j z_j - \sigma_j z_j = g_j \sum_{k=1}^{\infty} z_k\int_0^1 b(\zeta,\zeta')\phi_k(\zeta')d\zeta' \quad . \tag{13.26}$$

In the case of point observations - though we must note that it has not been proved here that the theory can be extended to cover unbounded observation -

$$Cz = z(\zeta_1) \quad , \quad \zeta_1 \in (0,1)$$

we can write (13.24), formally, as

$$y(\zeta,t) = Cx(\zeta,t) = \int_0^1 \delta(\zeta_1-\zeta')z(\zeta',t)d\zeta'$$

i.e.

$$b(\zeta,\zeta') = \delta(\zeta_1-\zeta')$$

and hence (13.26) becomes

$$\lambda_j z_j - \sigma_j z_j = g_j \sum_{k=1}^{\infty} z_k\phi_k(\zeta_1)$$

that is

$$(13.27) \qquad \lambda_j z_j - \sigma_j z_j = g_j \sum_{k=1}^{\infty} \sqrt{2}\, e^{(\mu/\sigma^2)\zeta_1} \sin k\pi\zeta_1 \quad .$$

From (13.22),

$$(Hu)_j = (Bu)_j = \langle Bu, \psi_j \rangle$$

$$= \langle u, B^* \psi_j \rangle$$

$$= \langle u, \frac{\sigma^2}{2} \frac{\partial}{\partial\zeta} \psi_j(o) \rangle$$

$$= \langle u, \frac{\sigma^2}{2} j\pi \rangle$$

so $(Hu)_j = \frac{\sigma^2}{2} j\pi u$.

In section 9.1 we saw that the tracking problem for the inventory control problem with stochastic deterioration rate and cost functional

$$J(u) = \int_o^{t_1} \{ \int_o^1 \langle z(\zeta,t) - z_d(\zeta,t), z(\zeta,t) - z_d(\zeta,t) \rangle d\zeta + \langle u(t), cu(t) \rangle \} dt$$

on the Hilbert space H, is well defined and that the optimal control is $u^* = -\{Kz + E\} = -\frac{1}{c} B^* \{Q(t)z(t) + N(t)\}$ with Q of the form $(Qz)(\zeta) = \int_o^1 L(\zeta,\eta,t)z(\eta)d\eta$, L satisfying (9.40), (9.41), and with $N(t)(\zeta) = N(\zeta,t)$ satisfying (9.43).

If instead of feeding back z we feedback x so that the control becomes $u = -\{Kx + E\}$ then from the previous section, section 12, we know that the cost suffers an increase given by

$\Delta J = \int_o^{t_1} \langle \bar{u}, c\bar{u} \rangle dt$ where $\bar{u} = u - u^*$. Thus

$$\Delta J = \int_o^{t_1} || c^{\frac{1}{2}} K\{z - x\} ||^2 dt$$

$$\leq \int_o^{t_1} ||c^{\frac{1}{2}} K\{\sum_{j=1}^{N} z_j\phi_j - \sum_{j=1}^{N} x_j\phi_j\}||^2 dt + \int_o^{t_1} ||c^{\frac{1}{2}} K \sum_{j=N+1}^{\infty} z_j\phi_j||^2 dt$$

$$= \int_o^{t_1} ||c^{\frac{1}{2}} K\{Pz - x\}||^2 dt + \int_o^{t_1} ||c^{\frac{1}{2}} K \sum_{j=N+1}^{\infty} z_j\phi_j||^2 dt$$

$$= \int_o^{t_1} ||c^{\frac{1}{2}} K \sum_{j=1}^{N} e^{\sigma_j}\{x_{oj} - z_{oj}\}\phi_j||^2 dt + \int_o^{t_1} ||c^{\frac{1}{2}} K \sum_{j=N+1}^{\infty} z_j\phi_j||^2 dt$$

and in this case $K = - c^{-\frac{1}{2}} B^* Q(t)$ with Q as above and $B^* y = \frac{\sigma^2}{2} \frac{\partial y}{\partial \zeta}(o)$. σ_j is given by (13.21), or in the case of observations of the form (13.24) with $G = g(\zeta)$, $g \in L^2(0,1)$, by (13.26), and in the case of point observations $Cz = z(\zeta_1)$, by (13.27).

CONCLUSIONS

Firstly in this thesis we examined the linear quadratic cost control problem for a class of distributed parameter systems with unbounded control action. The class of systems considered were those which satisfy a technical condition, namely condition (2.4). This condition in effect requires the semigroup, T_t , of the system to be smoothing to the extent that it nullifies the unboundedness of the control operator B , that is so that the resultant operator T_tB is bounded.

For this class of control systems three aspects of the linear quadratic cost control problem were examined, the finite time regulator problem, the tracking problem and the infinite time regulator problem.

Curtain and Pritchard {11} have shown that for the finite time regulator, when condition (2.4) is satisfied with $p \geq 2$, there exists an optimal control, which is feedback in nature, of the form $u^*(t) = - R^{-1}B^*Q(t)z(t)$. Further $Q(t)$ is the unique solution to both integral and differential Riccati equations.

We considered a wider class of systems, those for which (2.4) is satisfied with just $p \geq 1$, and in section 4 we have shown that for the finite time regulator with no final time penelty Curtain and Pritchard's results remain valid.

For the tracking problem, with no final time penelty when $p \geq 1$, and allowing final time penalty if $p \geq 2$, in condition (2.4), the optimal control was shown, in section 5, to be a combination of feedback and openloop control :- $u^*(t) = - R^{-1}B^*\{Q(t)z(t) + S(t)\}$. Here $Q(t)$ is the unique solution of the integral and differential Riccati equations associated with the finite time regulator and $S(t)$ also satisfies integral and differential equations.

For the infinite time regulator, which has pure integral cost functional, all the results obtained are for systems satisfying (2.4) with $p \geq 1$. It was shown that under an optimizability assumption, the unique optimal feedback control is again feedback, $u^* = -R^{-1}B^*Qz(t)$. In this case the feedback operator Q is time independent, and is the unique solution to an integral Riccati equation. The integral Riccati equation can be differentiated as for the finite time regulator and tracking problems but uniqueness of the resultant algebraic equation does not automatically follow. We however proved that when the adjoint system was stabilizable (i.e. $\{A^*, M^{\frac{1}{2}}\}$ was stabilizable) the algebraic Riccati equation has unique solution.

In establishing this uniqueness result a problem was mentioned which perhaps needs further highlighting and is a possible subject for further research. The problem is in establishing exactly what is the infinitesimal generator of the perturbed semigroup, T_t^Q, corresponding to the perturbation of the semigroup, T_t, by $-BR^{-1}B^*Q$.

We know that $\tilde{A}$ is the quasi-generator of T_t^Q but we do not have in general that $\tilde{A}$ is the infinitesimal generator of T_t^Q (where $\tilde{A}$ is the operator as defined in section 6). In section 6 all we have shown is that if A^Q is the infinitesimal generator of T_t^Q then A^Q is a closed extension of $\tilde{A}$ and on $\tilde{\mathcal{D}} = D(A) \cap \ker\{D - R^{-1}B^*Q\}$, $A^Q = \tilde{A}$. In order to obtain this result we have assumed the existence of a Green's formula (2.7) with $B^* = C$.

Arising out of the solution of the infinite time regulator problem we were able, in section 7, to prove the stabilizability result, exact null controllability implies stabilisability, for the systems under consideration.

The following section was devoted to a comparison of the approach of this thesis with those of others examining control problems with unbounded control action. Three alternative approaches were examined, that based on Fattorini's model, Zabczyk's approach (a variation of the Fattorini approach), and Lions' approach.

For the Fattorini approach it was shown that when his formulation is valid the two formulations (i.e. the one based on his model and the formulation of this thesis) are equivalent. Zabczyk's approach, like Fattorini's and that adopted in this thesis, is a semigroup approach. His formulation is valid for a wider class of systems than is ours but we show that when restricted so that our formulation is also valid, the two are equivalent.

It was noted that the majority of the work published using these two alternative approaches is in establishing the formulation and controllability results, and that very little consideration has been given to the linear quadratic cost control problem. Lions however does consider the linear quadratic cost control problem and obtains results essentially the same as ours. He considers systems where A satisfies a coercivity assumption. This assumption implies that A generates a strongly continuous semigroup but the reverse is not in general true.

We then, in section 9, saw how the results obtained in previous sections could be applied to specific examples.

Attention was then drawn to the fact that we could not apply our formulation in the case of hyperbolic systems. The problem arises in that in general the system semigroup for such systems is not smoothing, so that the condition, condition (2.4), crucial for our formulation to be valid, is not satisfied. As a consequence we noted how for a particular system, by restricting the problem, either by restricting the control space or by restricting the weighting operatos in the performance index, the problem could be reformulated as one of bounded, distributed, control action, which we know has unique solution.

This area, the area of hyperbolic systems with unbounded (boundary type) control action is one which would lend itself well to further research. It seems that the formulation and methods of this thesis are not suitable for such systems described by semigroups. The question of what is the right approach remains unanswered.

For the remainder of the thesis observer theory for distributed parameter systems described by semigroups was considered, particular attention being given to the use of observers in the quadratic cost control problem. Firstly we found conditions, in terms of the system operators, for an observer given by the mild solution to a system of the form $x_t = Fx + Gy + Hu$, to be an asymptotic state estimator of Pz. Here z is the mild solution to a system of the form $z_t = Az + Bu$ with observation equation $y = Cz$. Both general control action and feedback control $u = My + Nx$ were considered.

The conditions on the system operators we found to be such that if Pz was a finite dimensional approximation to z (or indeed any mapping of z on to a finite dimensional space) then we could construct a finite dimensional observer which asymptotically estimated Pz .

We also proved the stabilizability result that if the system $\{A,B\}$ is stabilizable then there exists an observer such that the feedback control $u = My + Nx$ stabilizes the system provided the observer operator F generates an exponentially stable semigroup; $MC + NP = D$ with $A + BD$ stable; and the observer is an asymptotic estimator of Pz .

The effect of using the feedback control obtained from an observer, $u = My + Nx$, rather than the optimal control, in the regulator problem was then considered. The observer was chosen to stabilize the system and it was shown that the resultant increase in cost depends on the initial state of the system which is, in general, unknown.

Finally we extended the theory to cover observers for systems with unbounded control action as formulated in section 2. It was shown that under different conditions on the operator P for the two cases, an observer with either distributed or unbounded control action could be found that asymptotically estimates Pz where z is the solution to a system with unbounded control action.

REFERENCES

1. AUBIN, J.P. Abstract Boundary Value Operators and Their Adjoints. Rendiconti del Seminario Matematico Padova. Vol.43, 1970.

2. BALAKRISHNAN, A.V. Filtering and Control Problems for Partial Differential Equations. Differential Games and Control Theory, Proceedings of the 2nd. Kingston Conference, University of Rhode Island. June 1976.

3. BALAKRISHNAN, A.V. Applied Functional Analysis. Springer-Verlag, New York. 1976.

4. BARBU, V. Boundary Control Problems with Convex Cost Criterion. Reprint, Institute of Mathematics, Bucharest. 1975.

5. BENSOUSSAN, A., HURST, E.G. and NASSLUND, B. Managerial Applications of Modern Control Theory. Studies in Mathematical and Managerial Economics. Vol.18, 1974. North Holland.

6. BENSOUSSAN, A., NISSEN, G. and TAPIERO, C. Optimum Inventory and Product Quality Control with Deterministic and Stochastic Deterioration - An Application of Distributed Parameter Control Systems. IEEE Transactions, AC-20, 1975, pp.407-412.

7. BORIGIORNO JR., J.J. and YOULA, D.C. On Observers in Multi-variable Control Systems. International Journal of Control, Vol.8, pp.221-243.

8. CURTAIN, R.F. Infinite Dimensional Estimation Theory Applied to a Water Pollution Problem. Optimization Techniques, Proceedings of the 7th. IFIP Conference, Nice, 1975, Part 2. Lecture Notes in Computer Science, Vol.41, 1975. Springer-Verlag.

9. CURTAIN, R.F. and PRITCHARD, A.J. The Infinite Dimensional Riccati Equation for Systems Described by Evolution Operators. SIAM J. Control, Vol.14, 1975, pp.951-983.

10. CURTAIN, R.F. and PRITCHARD, A.J. An Abstract Theory for Unbounded Control Action for Distributed Parameter Systems. SIAM J. Control Vol.15, 1977, pp.566-611.

11. CURTAIN, R.F. and PRITCHARD, A.J. Infinite Dimensional Linear Systems Theory. Lecture Notes in Control and Information Sciences. No.8, 1978, Springer-Verlag.

12. EL JAI, A. Etude D'Algorithmes pour la Commande Optimal De Systems a Parametres Repartis de Type Parabolique. Ph.D. Thesis, Paul Sabatier University, Toulouse,1978.

13. FATTORINI, H.O. Boundary Control Systems. SIAM J. Control, Vol.6, 1968, pp. 349-385.

14. GRESSANG, R.V. and LAMONT, G.B. Observers for Systems Characterised by Semigroups. IEEE Transactions, AC-20, 1975, pp.523-528.

15. HALMOS, P.R. Introduction to Hilbert Space and the Theory of Spectral Measure.

16. HILLE, E. and PHILLIPS, R.S. Functional Analysis and Semigroups AMS Colloquium Publications. Vol.31, 1957.

17. KANTOROVICH, L.V. and AKILOV, G.P. Functional Analysis in Normed Spaces. Pergamon Press, 1964.

18. KATO, T. Perturbation Theory for Linear Operators. Springer-Verlag, 1966.

19. KAWAKERNAAK, H. Filtering for Systems Excited by Poisson White Noise. Int. Symposium on Control Theory, Numerical Methods and Computer Systems Modelling, 1974. Lecture Notes in Economics and Mathematical systems, Vol.107. Springer-Verlag, 1974.

20. KIM, M. and ERZBERGER, H. On the design of Optimum Distributed Parameter Systems with Boundary Control Function. IEEE Transactions, AC-12, 1967, pp.22-28.

21. KITAMURA, S., SAKAIRI, S. and NISHIMURA, M. Observer for Distributed Parameter Diffusion System. Electrical Engineering in Japan, Vol.92, 1972, pp.142-149.

22. KOHNE, M. The Control of Vibrating Elastic Systems, in, Distributed Parameter Systems, Ray and Lainiotis (Eds.). Marcel Dekker, 1977.

23. KOHNE, M. Implementation of Distributed Parameter State Observers. Distributed Parameter System Model Identification, Proceedings of IFIP Conference, Rome, 1976. Lecture Notes in Control and Information Science, Vol.1, 1978, pp.310-324.

24. LIONS, J.L. Optimal Control of Systems Governed by Partial Differential Equations. Springer-Verlag, 1971.

25. LUENBERGER, D.C. Observers for Multivariable Systems. IEEE Transactions, AC-11, 1966, pp.190-197.

26. LUI, Y.A. and LAPIDUS, L. Observer Throry for Distributed Parameter Systems. Int. Journal of Systems Science, Vol.7, 1976, pp.731-742.

27. NEWMANN, M.M. Optimal and Sub-optimal Control Using an Observer when some of the State Variables are not Mearsurable. Int. J. Control, Vol.9, 1969, pp.281-290.

28. ONER, P.A. and FOSTER, A.M. A Design Procedure for a Class of Distributed Parameter Control Systems. J. of Dynamic Systems Measurement and Control. ASME, 1971, pp.86-93.

29. PRADO, G. Observability, Estimation and Control of Distributed Parameter Systems. Ph.D. Thesis, MIT, 1971.

30. POLLOCK, J. and PRITCHARD, A.J. The Infinite Time Quadratic Cost Control Problem for Distributed Parameter Systems with Unbounded Control Action. J. Inst. Maths. Applics. To be published.

31. POLLOCK, J. and PRITCHARD, A.J. The Tracking Problem for Distributed Systems with Unbounded Control Action. J. Inst. Maths. Applics. To be published.

32. PRITCHARD, A.J. and ZABCZYK, J. Stability and Stabilizability of Infinite Dimensional Systems. Control Theory Centre Report, No.70, University of Warwick, 1977.

33. RUSSELL, D.L. Quadratic Performance Criteria in Boundary Control of Linear Symmetric Hyperbolic Systems. SIAM J. Control, Vol.11, 1973, pp.475-509.

34. SALAMON, D. Observers and Duality between Observation and State Feedback for Time Delay Systems. Department of Dynamical Systems Report, No.5, University of Bremen, 1979.

35. SAMARA, I.G. and JAYARAJ, C. On the Use of Observers in Finite-Time Optimal Regulator Problems. Int. J. Control, Vol.11, 1970, pp.489-497.

36. SETHI, S.P. Dynamic Optimal Control Models in Advertising: A Survey. SIAM Review, Vol.19, 1977.

37. TABAC, D. Application of Modern Control and Optimization Techniques to Transportation Systems. Control and Dynamical Systems: Advances in Theory and Applications. Vol.10, 1973, pp.345-432.

38. VINTER, R.B.,and JOHNSON, T.L. Optimal Control of Non-symmetric Hyperbolic Systems in n Variables on the Half-space. SIAM J. Control, Vol.15, 1977, pp.129-133.

39. WASHBURN, D.C. A Semigroup Theoretic Approach to Modelling of Boundary Input Problems. Proceedings of the IFIP Working Conference, Rome. Lecture Notes in Control and Information Science, Springer-Verlag, 1977.

40. WOLOVICH, W.A. Linear Multivariable Systems. Applied Mathematical Sciences, Vol.11, Springer-Verlag, 1974.

41. WONHAM, W.M. Linear Multivariable Control. A Geometric Approach. Lecture Notes in Economics and Mathematical Systems, Vol.101, Springer-Verlag, 1974.

42. YUKSEL, Y.O. and BORIGIORNO JR., J.J. Observers for Linear Multi-Variable Systems with Applications. IEEE Transactions, AC-16, 1971, pp.603-613.

43. ZABCZYK, J. Infinite Dimensional Systems in Control Theory. Control Theory Centre Report, No.66, University of Warwick, 1977.

44. LIONS, J.L. and MAGENES, E. Non-Homogeneous Boundary Value Problems and Applications, Vols. I - III, Springer-Verlag, 1972.

ADDITIONAL REFERENCES

ARTHUR, W.B. and McNICOLL, G. Optimal Time Paths with Age-Dependence: A Theory of Population Policy. Review of Economic Studies,Vol.44, 1977, pp.111-123.

AUBIN, J. Approximation of Elliptic Boundary-Value Problems. Wiley-Interscience, 1972.

CHICHESTER, F.D. and PERLIS, N.J. Optimal Dynamic Control of a Tidal River Quality System Subject to Economic Constraints. Proc. 1976 Joint Automatic Control Conference, American Society of Mechanical Engineers.

CURTAIN, R.F. and PRITCHARD, A.J. The Infinite Dimensional Riccati Equation. J. Math. Anal. and Appl., Vol.49, 1974, pp.43-57.

DATKO, R. A Linear Control Problem in an Abstract Hilbert Space. J. of Diff. Eq. Vol.9, 1971, pp.346-359.

DATKO, R. Uniform Asymptotic Stability of Evolutionary Processes in a Banach Space. SIAM J. Math. Anal. Vol.3, 1972, pp.428-445.

DATKO, R. Unconstrained Control Problems With Quadratic Cost. SIAM J. Control, Vol.11, 1973, pp.32-52.

DELFOUR, M.C. The Linear Quadratic Optimal Control Problem Over an Infinite Horizon for a Class of Distributed Parameter Systems. Control of Distributed Parameter Systems, Proc. of 2nd IFAC Symp., Coventry, 1977. Eds. Banks and Pritchard. Pergamon Press.

DELFOUR, M.C., McCALLA, C. and MITTER, K. Stability and The Infinite-Time Quadratic Cost Problem For Linear Hereditary Differential Systems. SIAM J. Control, Vol.13, 1975, pp.48-85.

DELFOUR, M.C. and MITTER, S.K. Controllability, Observability and Optimal Feedback Control of Affine Hereditary Differential Systems. SIAM J. Control, Vol.10, 1972, pp.298-328.

DI BLASIO, G. and LAMBERTI, L. An Initial-Boundary Value Problem for Age-Dependent Population Diffusion. SIAM J. Appl.Math., Vol.35, 1978, pp.593-615.

DOLECKI, S. and RUSSELL, D.L. A General Theory of Observation and Control. SIAM J. Control. Vol.15, 1977, pp.185-219.

DUNFORD, N. and SCHWARTZ, J.T. Linear Operators, Interscience, 1959, 1963.

ERZBERGER, H. and KIM, M. Optimum Boundary Control of Distributed Parameter Systems. Information and Control, Vol.9, 1966, pp.265-278.

FATTORINI, H.O. Boundary Control of Temperature Distributions in a Parallelepipedon. SIAM J. Control, Vol.13, 1975, pp.1-13.

FATTORINI, H.O. The Time-Optimal Problem for Boundary Control of the Heat Equation. Calculus of Variations and Control Theory, Proc. of Symp. Maths. Research Centre, University of Wisconsin, 1975. Ed. D.L. Russell. pp.305-320.

GIBSON, J.S. The Riccati Equation for Optimal Control Problems on Hilbert Spaces. SIAM J. Control, Vol.17, 1979, pp.537-565.

GIURGIU, M. A Feedback Solution of a Linear Quadratic Problem for Boundary Control of Heat Equation. Rev. Roum. Math. Pures et Appl. Vol.20, 1975, pp. 928-954.

GOULD, J.P. Diffusion Processes and Optimal Advertising Policy. Microeconomic Foundations of Employment and Inflation Theory, Eds. Phelps et al, Macmillan, 1970.

GRAHAM, J.W. Optimal Boundary Control of a Class of Distributed Parameter Systems. Int. J. Control, Vol.14, 1971, pp.937-949.

HULLETT, W. Optimal Estuary Aeration: An Application of Distributed Parameter Control Theory. Appl. Math. and Opt. Vol.1, 1974, pp.20-63.

KHATSKEVICH, V.P. On the Problem of the Analytic Design of Regulators for Distributed-Parameter Systems Under Boundary Control. PMM., Vol.35, 1971, pp.598-608.

LUKES, D.L. and RUSSELL, D.L. The Quadratic Criterion for Distributed Parameter Systems. SIAM J. Control, Vol.7, 1969, pp.101-121.

MacCAMY, R.C., MIZEL, V.J. and SEIDMAN, T.I. Approximate Boundary Control of the Heat Equation, I and II, J. Math. Anal. and Appl., Vol.23, 1968, pp.699-703, and Vol.28, 1969, pp.482-492.

PRITCHARD, A.J. and CROUCH, P.E. Modelling and Control for Distributed Parameter Systems. Optimization Techniques, Proc. 7th IFIP Conf., Nice, 1975. Lecture Notes in Computer Science, Vol.41. Springer-verlag, 1975.

PRITCHARD, A.J. and WIRTH, A. Unbounded Control and Observation Systems and Their Duality. SIAM J. Control, Vol.16, 1978, pp.535-545.

RUSSELL, D.L. On Boundary-Value Controllability of Linear Symmetric Hyperbolic Systems. Mathematical Theory of Control, Eds. A.V. Balakrishnan and L.W. Neustadt, Academic Press, 1967, pp.312-321.

RUSSELL, D.L. Boundary Value Control of the Higher-Dimensional Wave Equation, I and II. SIAM J. Control, Vol.9, 1971, pp.29-42, and pp.401-419.

RUSSELL, D.L. A Unified Boundary Controllability Theory for Hyperbolic and Parabolic Partial Differential Equations. Studies in Applied Maths., Vol.52, 1973, pp.189-211.

SEIDMAN, T.I. Boundary Observation and Control for the Heat Equation. Calculus of Variations and Control Theory, Proc. of Symp. Maths. Research Centre, University of Wisconsin, 1975. Ed. D.L. Russell. pp. 321-351.

SEIDMAN, T.I. Observation and Prediction for the Heat Equation,III. J. of Diff. Eq., Vol.20, 1976, pp.18-27.

SLEMROD, M. Stabilization of Boundary Control Systems. J. of Diff. Eq., Vol.22, 1976, 402-415.

TRIGGIANI, R. A Cosine Operator Approach to Modelling Boundary Input Hyperbolic Systems. Proc. 8th IFIP Conf. on Optimization, University of Wurzburg, West Germany, 1977.

TRIGGIANI, R. On the Relationship Between First and Second Order Controllable Systems in Banach Space. Distributed Parameter Model Identification, Proc. of IFIP Conf. Rome, 1976. Lecture Notes in Control Inf. Sci., Vol.1, 1978, pp.370-393.

VINTER, R. Filter Stability for Stochastic Evolution Equations. SIAM J. Control, Vol.15, 1977, 465-485.

YOSIDA, K. Functional Analysis. Springer-Verlag, 1966.

ZABCZYK, J. A Semigroup Approach to Boundary Value Control. Control of Distributed Parameter Systems, Proc. 2nd IFAC Symp., Coventry, 1977. Eds. Banks and Pritchard, Pergamon Press.