Monograph:

Banks, S.A. and Yew, M.K. (1984) On a Class of Suboptimal Controls for Infinite-Dimensional Bilinear Systems. Research Report. Acse Report 258 . Dept of Automatic Control and System Engineering. University of Sheffield

[redacted]

ON A CLASS OF SUBOPTIMAL CONTROLS FOR INFINITE-DIMENSIONAL BILINEAR SYSTEMS.

by

S.P.Banks and M.K.Yew

RESEARCH REPORT NO. 258

University of Sheffield
Department of Control Engineering
Mappin Street,
SHEFFIELD S1 3JD

1984

ABSTRACT

The suboptimal control of a bilinear system is considered with respect to a quadratic cost criterion. The feedback control is in the space of formal power series on a Hilbert space.

1. Introduction

Bilinear systems have been considered very extensively by many authors; see, for example, Brockett (1976), Gutman (1981), Banks (1984). The main reason for this is that they form the most simple generalization of linear systems. However it can also be shown that any system which is analytic has a bilinear approximation and so bilinear systems do not comprise too restrictive a class of nonlinearities.

Although a great deal of attention has been given to controllability and observability and stabilization of bilinear systems (see for example, Murthy, 1979, Ball and Slemrod, 1979 and Grasselli and Isidori, 1981) there does not seem to have been much published work on the optimal control of bilinear systems subject to a quadratic cost functional. In this paper we shall consider the bilinear system

$$\dot{x} = Ax + uBx$$

for a scalar control u, where A and B are bounded operators on a separable Hilbert space H, together with the quadratic cost

$$J = \langle x, Gx \rangle + \int_0^{t_f} \{\langle x, Mx \rangle + ru^2\}\, dt$$

We shall determine the optimal control in a certain class of controls by extending the linear-quadratic dynamic programming argument. This will require the notion of tensors and tensor operators on H and so in the next section we shall give a brief introduction to these ideas.

The control will turn out to be given by a power series whose tensor coefficients can be determined recursively. When the series is truncated, we obtain a control which is suboptimal in the class of admissible controls.

2. Tensor Theory in Hilbert Space

In this section we shall give a brief introduction to the theory of tensors on a Hilbert space H. For more details see Greub, 1978. First recall that if E and F are vector spaces and G is any vector space then the tensor product of E and F is defined as the pair $(E\otimes F,\otimes)$ ($\otimes$ a bilinear mapping) with the following universal property: if ϕ is a bilinear mapping then there exists a unique linear mapping $f:E\otimes F\to G$ such that the diagram

$$\begin{tikzcd} E\times F \arrow[r, "\phi"] \arrow[d, "\otimes"'] & G \\ E\otimes F \arrow[ur, "f"'] & \end{tikzcd}$$

commutes. By induction we can define the tensor product of i copies of H; i.e. $\underbrace{H\otimes\ldots\otimes H}_{i}$ which we shall denote by $\otimes_i H$. We can make $\otimes_i H$ into a Hilbert space by defining

$$\langle x_1 \ldots\ldots x_i, y_1 \ldots\ldots y_i\rangle_{\otimes_i H} = \prod_{j=1}^{i} \langle x_j, y_j\rangle_H ,$$

and extending by linearity.

It is convenient to consider the space $H = \bigoplus_{i=1}^{\infty} (\otimes_i H)$ as a graded Hilbert space.

Let $\{e_k\}_{k\geq 1}$ be an orthonormal basis of H which will be fixed throughout the discussion. Then $\{e_{k_1}\otimes\ldots\otimes e_{k_i}\}$ $(1\leq k_j<\infty,\ 1\leq j\leq i)$ is an orthonormal basis of $\otimes_i H$ and so any tensor $\Xi \in \otimes_i H$ can be written in the form

$$\Xi = \sum_{k_1=1}^{\infty} \ldots \sum_{k_i=1}^{\infty} \xi_{k_1\ldots k_i} (e_{k_1}\otimes\ldots\otimes e_{k_i})$$

Since $\otimes_i H$ is a Hilbert space we can consider linear operators defined on $\otimes_i H$. The space of all bounded linear operators on a space X will be denoted by $\mathcal{L}(X)$. Let $P\in\mathcal{L}(\otimes_i H)$. Then the matrix representation of P with respect to the above basis of $\otimes_i H$ will be written $p_{k_1 \ldots k_i}^{\ell_1 \ldots \ell_i}$; i.e.

$$P(e_{k_1} \otimes \cdots \otimes e_{k_i}) = \sum_{\substack{\ell_j = 1 \\ (j=1,i)}}^{\infty} p_{k_1 \ldots k_i}^{\ell_1 \ldots \ell_i} (e_{\ell_1} \otimes \cdots \otimes e_{\ell_i}).$$

(Since $\otimes_i H$ is a 'flat' space writing indices contra- or co-variantly makes no difference.)

The dual or adjoint P^* of $P \varepsilon \mathcal{L}(\otimes_i H)$ is defined in the usual way:

$$\langle P^*(x_1 \otimes \ldots \otimes x_i), (y_1 \otimes \ldots \otimes y_i) \rangle = \langle (x_1 \otimes \ldots \otimes x_i), P(y_1 \otimes \ldots \otimes y_i) \rangle$$

for all $x_j, y_j \varepsilon H$. Clearly, P is self-adjoint if

$$P_{k_1 \ldots k_i}^{\ell_1 \ldots \ell_i} = P_{\ell_1 \ldots \ell_i}^{k_1 \ldots k_i} ,$$

and such an operator P will be said to be symmetric (This should not be confused with the usual definition of symmetric tensor.)

3. Optimal Control of Bilinear Systems

We shall consider the bilinear system

$$\dot{x} = Ax + uBx \tag{3.1}$$

where $x \in H$ (a separable Hilbert space) and u is a scalar control (the latter assumption being purely for notational convenience - the general case presents no further difficulties). However, we shall assume here, for simplicity, that A and B are bounded operators. The generalisation to unbounded operators will be considered in a future paper. We shall determine the control u which minimises the quadratic cost functional

$$J = \langle x, Gx \rangle + \int_o^{t_f} \{\langle x, Mx \rangle + ru^2\} \, dt \tag{3.2}$$

for controls which belong to a certain class, to be introduced shortly. In (3.2), G and M are nonnegative definite bounded linear operators on H and $r > o$.

If $V(t,x)$ denotes the usual value function, then the dynamic programming equation for V is

$$\langle x,Mx\rangle + V_t + (\mathcal{F}_x V)Ax + \min_u (ru^2+(\mathcal{F}_x V)Bxu) = 0 \qquad (3.3)$$

where $\mathcal{F}_x V$ is the Fréchet derivative of V (which we assume for the moment exists).

Now, as in the linear-quadratic regular problem, if $c = (\mathcal{F}_x V)Bx$, then

$$ru^2 + cu = (u+\tfrac{1}{2}r^{-1}c)^2 r - \tfrac{1}{4}c^2 r^{-1}$$

and so the minimum is attained when $u = -\frac{1}{2}r^{-1}c$. Then (3.3) becomes

$$V_t+\langle x,Mx\rangle + (\mathcal{F}_x V)Ax - \frac{1}{4}\langle(\mathcal{F}_x V)Bx,\ r^{-1}(\mathcal{F}_x V)Bx\rangle = 0 \qquad (3.4)$$

Now let $\Phi = \mathbb{R}^e[[x]]$ denote the ring of formal power series in the indeterminate x (εH) which have only even order powers; i.e. we may write, for any $\phi\varepsilon\Phi$,

$$\phi = \sum_{i=1}^{\infty} \langle \otimes_i x, \phi_i \otimes_i x\rangle_{\otimes_i H} \qquad (3.5)$$

where $\otimes_i H$ is the tensor product of i copies of H, $\otimes_i x = \underbrace{x\otimes x \ldots \otimes x}_{i}$

and $\phi_i \ \varepsilon\mathcal{L}(\otimes_i H)$. (Recall that the inner product on $\otimes_i H$ is given by

$$\langle x_1\otimes\ldots\otimes x_i, y_1\otimes\ldots\otimes y_i\rangle_{\otimes_i H} = \prod_{j=1}^{i} \langle x_j,y_j\rangle_H \ , \ x_j,y_j\varepsilon H.)$$

We shall need the following lemma, whose proof is trivial:

<u>Lemma 3.1</u>. Let $P\varepsilon\mathcal{L}(\otimes_i H)$, $Q\varepsilon\mathcal{L}(\otimes_j H)$. Then,

$$\left[\mathcal{F}_x \langle\otimes_i x,\ P\otimes_i x\rangle\right] x = 2i\langle\otimes_i x, P\otimes_i x\rangle \ , \qquad (3.6)$$

(if P is symmetric) and

$$\langle\otimes_i x,\ P\otimes_i x\rangle\langle\otimes_j x, Q\otimes_j x\rangle = \langle\otimes_{i+j} x, (P\otimes Q)\otimes_{i+j} x\rangle \qquad (3.7)$$

Moreover, we have

$$\| P\otimes Q \|_{\mathcal{L}(\otimes_{i+j} H)} \le \| P\|_{\mathcal{L}(\otimes_i H)} \| Q\|_{\mathcal{L}(\otimes_j H)} \ . \ \square \qquad (3.8)$$

It follows from this lemma that, for a bounded operator $C\in\mathcal{L}(H)$, we have for any symmetric P,

$$\mathcal{D}_x \langle \otimes_i x, P \otimes_i x \rangle Cx = 2 \langle \otimes_i x, (PC) \otimes_i x \rangle \tag{3.9}$$

where

$$PC = \sum_{j=1}^{i} \left(\sum_{k_j=1}^{\infty} P^{\kappa_1 \ldots \kappa_i}_{k_1 \ldots k_j \ldots k_i} \cdot C_{k_j \ell_1} \right)_{(1\le \ell_1<\infty, 1\le k_1<\infty, \ldots, 1\le k_i<\infty, 1\le \kappa_\alpha<\infty)}$$

and $P^{\kappa_1 \ldots \kappa_i}_{k_1 \ldots k_i}$, $C_{k_j \ell_1}$ are the components of the tensors P,C with respect to some (fixed) orthonormal basis of H. A similar definition can be given for CP. Also, it is clear that

$$\| (PC) \otimes_i x \|_{\otimes_i H} = \| P(Cx \otimes x \ldots \otimes x) + P(x \otimes Cx \ldots \otimes x) + \ldots + P(x \otimes \ldots \otimes Cx) \|_{\otimes_i H}$$

$$\le i \| P \|_{\mathcal{L}(\otimes_i H)} \| Cx \|_H \| x \|_H^{i-1}$$

and so

$$\| PC \|_{\mathcal{L}(\otimes_i H)} \le i \| P \|_{\mathcal{L}(\otimes_i H)} \| C \|_{\mathcal{L}(H)} . \tag{3.10}$$

We can now substitute (3.5) into (3.4) (with $\phi_i = P_i$), to obtain

$$\sum_{i=1}^{\infty} \left\langle \otimes_i x, \dot{P}_i(t) \otimes_i x \right\rangle_{\otimes_i H} + \sum_{i=1}^{\infty} 2 \left\langle \otimes_i x, (P_i A) \otimes_i x \right\rangle_{\otimes_i H} + \langle x, Mx \rangle_H$$
$$- \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} r^{-1} \left\langle \otimes_{i+j} x, (P_i B \otimes P_j B) \otimes_{i+j} x \right\rangle_{\otimes_{i+j} H} = 0 \tag{3.11}$$

with the final conditions

$$P_1(t_f) = G$$
$$P_i(t_f) = 0 \quad , \quad i>1$$

(Using (3.7),(3.9)). Equating like 'powers' in x in (3.11) we obtain the equations

$$\dot{P}_1(t)+P_1(t)A+A^TP_1(t)+M = 0 \quad , \qquad P_1(t_f) = G$$

$$\dot{P}_m(t)+P_m(t)A+A^TP_m(t)-r^{-1} \sum_{\substack{i+j=m \\ i,j\geq 1}} \left\{ P_iB\otimes P_jB \right\} = 0, \qquad P_m(t_f) = 0 \tag{3.12}$$

for m>1.

Note that the latter equation can also be written in the form

$$\dot{P}_m(t) + P_m(t)\mathcal{A} + A^TP_m(t) - \frac{1}{2r} \left\{ \sum_{\substack{i+j=m \\ i,j\geq 1}} (B^TP_i\otimes B^TP_j+P_iB\otimes P_jB) \right\} = 0$$

since, clearly $(P_iB\otimes P_jB)^T = B^TP_i\otimes B^TP_j$, so that P_i is indeed symmetric.

Consider the operators $\mathcal{A}_i$ defined on the Banach spaces $\mathcal{L}(\underset{i}{\otimes}H)$ by

$$\mathcal{A}_iP_i = P_iA \quad , \quad P_i\varepsilon\mathcal{L}(\underset{i}{\otimes}H) \quad , \quad i\geq 1 \ .$$

where P_iA is defined as in (3.9). Then $\mathcal{A}_i$ is clearly a bounded operator and

$$||\mathcal{A}_iP_i|| \leq i||P_i||\cdot||A||$$

by (3.10), whence

$$||\mathcal{A}_i||_{\mathcal{L}(\mathcal{L}(\underset{i}{\otimes}H))} \leq i||A||_{\mathcal{L}(H)}$$

Hence we can define the operator

$$e^{\mathcal{A}_it} \ \varepsilon\mathcal{L}(\mathcal{L}(\underset{i}{\otimes}H))$$

and the solution of (3.12(1)) is then

$$P_1(t) = e^{\mathcal{A}_1(t_f-t)}Ge^{\mathcal{A}_1^T(t_f-t)} + \int_o^{t_f-t} e^{\mathcal{A}_1(t_f-t-s)}Me^{\mathcal{A}_1^T(t_f-t-s)}ds \tag{3.13}$$

Similarly, from (3.12(m)), we have

$$P_m(t) = -r^{-1} \sum_{\substack{i+j=m \\ i,j\geq 1}} \int_o^{t_f-t} e^{\mathcal{A}_m(t_f-t-s)}P_i(t_f-s)B\otimes P_j(t_f-s)Be^{\mathcal{A}_m^T(t_f-t-s)}ds \tag{3.14}$$

The optimal control is then formally

$$u(t) = -\tfrac{1}{2}r^{-1}(\mathcal{J}_x V)Bx = -r^{-1} \sum_{i=1}^{\infty} < \otimes_i x, (P_i B)\otimes_i x > \tag{3.15}$$

However this series may not converge and so we propose the following sub-optimal controls:

$$u_m(t) = -r^{-1} \sum_{i=1}^{m} < \otimes_i x, (P_i B)\otimes_i x> \tag{3.16}$$

These controls have been shown to be effective for finite dimensional systems (see Banks and Yew, 1984), for example in stabilising unstable bilinear control systems.

4. Example

As a simple example of the above theory we shall consider the system

$$\dot{x} = Ax + ux \quad , \quad x\varepsilon\ell^2 \tag{4.1}$$

where B = I and A is the left shift operator. This is not too restrictive on the operator A since any bounded operator has the left shift as a model on some Hilbert space. (see Rota (1960), Banks and Abbasi-Ghelmansarai, 1983). Recall that the left shift operator has the matrix representation

$$\begin{pmatrix} 0 & 1 & 0 & 0 \cdots\cdots \\ 0 & 0 & 1 & 0 & 0 \cdots \\ 0 & 0 & 0 & 1 & 0 \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

on ℓ^2.

Before considering this particular system let us examine the operator $e^{\mathcal{A}_m t}$ in more detail. Recall that $\mathcal{A}_m$ is defined on $\mathcal{L}(\otimes_m H)$ by

$$\mathcal{A}_m P = PA \quad , \quad P\varepsilon\mathcal{L}(\otimes_m H),$$

where

$$PA = \sum_{j=1}^{m} \left\{ \sum_{k_j=1}^{\infty} \left(P_{k_1 \ldots k_j \ldots k_m}^{\kappa_1 \ldots\ldots \kappa_m} \cdot A_{k_j \ell_1} \right) \right\}_{(1\le\ell_1<\infty, 1\le k_1<\infty, \ldots, 1\le k_i<\infty, 1\le\kappa_\alpha<\infty)}$$

Write

$$\mathcal{A}_m^j P = \sum_{k_j=1}^{\infty} P_{k_1 \ldots k_j \ldots k_m}^{\varkappa_1 \ldots\ldots \varkappa_m} \cdot A_{k_j \ell_1} .$$

Then

$$\mathcal{A}_m P = (\Sigma \mathcal{A}_m^j) P .$$

Note that $\mathcal{A}_m^j, \mathcal{A}_m^k$ commute for all j,k and so

$$e^{\mathcal{A}_m t} = e^{\mathcal{A}_m^1 t} e^{\mathcal{A}_m^2 t} \ldots\ldots e^{\mathcal{A}_m^m t} .$$

Also,

$$e^{\mathcal{A}_j t} = \sum_{k_j=1}^{\infty} P_{k_1 \ldots k_j \ldots k_m}^{\varkappa_1 \ldots \varkappa_m} (e^{A t})_{k_j \ell_1}$$

and it is easy to see that

$$e^{\mathcal{A}_m t} P = P(\underbrace{e^{At} \otimes \ldots \otimes e^{At}}_{m}) \tag{4.3}$$

From (3.14) with B = I we have

$$e^{\mathcal{A}_m t} P_m(t) e^{\mathcal{A}_m^T t} = -r^{-1} \sum_{\substack{i+j=m \\ i,j \geq 1}} \int_0^{t_f - t} e^{\mathcal{A}_m (t_f - s)} P_i(t_f - s) \otimes P_j(t_f - s) e^{\mathcal{A}_m^T (t_f - s)} ds$$

$$= -r^{-1} \sum_{\substack{i+j=m \\ i,j \geq 1}} \int_0^{t_f - t} \{ e^{\mathcal{A}_i (t_f - s)} P_i (t_f - s) e^{\mathcal{A}_i^T (t_f - s)} \} \otimes \{ e^{\mathcal{A}_j (t_f - s)} P_j (t_f - s) e^{\mathcal{A}_j^T (t_f - s)} \} ds$$

where the latter equality follows from (4.3), and the fact that

$$(A \otimes B) \cdot (P \otimes Q) = (AP \otimes BQ)$$

for any operators A,B. Hence, writing

$$Q_m = e^{\mathcal{A}_m t} P_m(t) e^{\mathcal{A}_m^T t}$$

we have

$$Q_m(t) = -r^{-1} \sum_{\substack{i+j=m \\ i,j \geq 1}} \int_0^{t_f - t} Q_i(t_f - s) \otimes Q_j(t_f - s)\,ds$$
$$Q_1(t) = e^{\mathcal{A}_1 t_f} G e^{\mathcal{A}_1^T t_f} + \int_0^{t_f - t} e^{\mathcal{A}_1 (t_f - s)} M e^{\mathcal{A}_1^T (t_f - s)}\,ds \tag{4.4}$$

Now, if A is the left shift operator, then

$$e^{At} = \begin{pmatrix} 1 & t & t^2/2! & t^3/3! & \ldots\ldots & \\ 0 & 1 & t & t^2/2! & t^3/3! & \ldots\ldots \\ 0 & 0 & 1 & t & t^2/2! & \ldots\ldots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}.$$

Equations (4.4) can be solved recursively for Q_m. Of course we must terminate at some finite value of m and thus obtain a suboptimal control. If $G = M = I$, we clearly have, for example,

$$Q^j_{1,i}(t) = \sum_{n=0}^{\infty} \frac{t_f^{2n+|i-j|}}{(n+|i-j|!)n!} + \sum_{n=0}^{\infty} \frac{(t_f - t)^{2n+|i-j|+1}}{(2n+|i-j|+1)(n+|i-j|!)n!}$$

for the $(i,j)^{th}$ component of Q_1 when A is the left shift operator.

5. Conclusions

In this paper we have derived a class of suboptimal controls for a belinear system subject to a quadratic cost criterion. The control is a nonlinear feedback which is a power series in the state and can be calculated recursively. We have considered the case of bounded operators here - the unbounded case will be considered in a future paper.

6. References

Ball J.M. and Slemrod, M. (1979) Feedback Stabilization of Distributed Semilinear Control Systems, App. Math.Optim., 5,pp.169-179.

Banks, S.P. (1984). On the Generation of Infinite Dimensional Bilinear Systems and Volterra Series, to be published in Int. J.Sys.Sci.

Banks S.P. and Abbasi-Ghelmansarai, F. (1983). Realization Theory and the Infinite-Dimensional Root Locus, Int. J. Control, 38, pp 589-606.

Banks, S.P. and Yew, M.K. (1984) Suboptimal Control of Bilinear Systems - Some Examples, in preparation.

Brockett, R.W. (1976) Volterra Series and Geometric Control Theory, Automatica, 12, pp. 167-176.

Grasselli, O.M. and Isidori, A. (1981) An Existence Theorem for Observers of Bilinear Systems, IEEE Trans. Aut. Control, AC-26,pp.1299-1300

Greub, W. (1978) Multilinear Algebra, 2nd Ed., Springer-Verlag.

Gutman, P.O. (1981) Stabilizing Controllers for Bilinear Systems, IEEE Trans. Aut. Contr., AC-26, pp 917-922.

Murty, D.N.P., (1979) Controllability of a Discrete-Time Bilinear System, IEEE Trans. Aut. Control, AC-24, pp. 974-975

Rota, G-C, (1960) Universal Models for Linear Operators, Comm. Pure Appl. Math. 13, pp. 469-472.