[redacted]

LINEAR AND BILINEAR INFINITE-DIMENSIONAL REPRESENTATIONS OF FINITE-DIMENSIONAL NONLINEAR SYSTEMS

by

S. P. BANKS and A. ASHTIANI

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

RESEARCH REPORT NO 259

1984

Abstract

The linear representation of nonlinear systems is considered and applied to obtain a bilinear representation of a nonlinear control system. Finally we consider the bilinear-quadratic regulator problem using a certain class of suboptimal controls.

1. Introduction

Nonlinear systems have been studied extensively for many years by a large number of authors and although our understanding of many specific types of nonlinearities is now considerable we are still a long way from a unified theory of such systems. Because of the existence of a unified theory of linear systems many attempts have been made to find good linearisations or linear representations of nonlinear systems. The best known engineering technique is the describing function which has been verified rigorously in certain cases (see Mees and Bergen, 1975 and Bergen and Franks, 1971). Linearisations from a differential geometric approach have also recently been proposed (see Hunt et al, 1983, Sandberg 1981, Krener, 1973, Brockett, 1978). An elementary method using linearly independent or orthogonal functions has been considered for finite dimensional systems by Takata, 1979 and for infinite dimensional systems by Banks, 1984. The main drawback with the latter methods is that for systems of dimension greater than 1 obtaining a linear infinite dimensional system with state vector consisting of the linearly independent functions in a linear order is not easy and leads to a system matrix which has little recognisable structure. This leads to difficulties in the inductive generation of the system matrix. In this paper we propose a tensor operator representation which leads to a system operator in a completely transparent form for polynomial systems.

The second main objective of this paper is to apply the above ideas to show that any nonlinear system which has an analytic vector field can be replaced by an infinite dimensional bilinear system. We shall then apply some recent results on suboptimal controls for such systems obtained by Banks and Yew (1984).

2. Notation and Terminology

Much of the notation used in this paper is standard or will be introduced as we proceed. However, note that we use $C_n^{\omega}(R)$ to denote the space of real analytic functions of n variables which have a convergent Taylor series in the ball of radius R. Later in the paper we shall use some tensor theory on a Hilbert soace H. Recall that if E and F are vector spaces and G is any vector space, then the tensor product (denoted $E\otimes F$) of E and F is the pair $(E\otimes F,\otimes)$, where $\otimes$ is a bilinear mapping, with the following universal property: if ϕ is a bilinear mapping from ExF into G 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. $H\otimes\dots\otimes H \triangleq \otimes_i H$. $\otimes_i H$ is a Hilbert space with inner product defined by

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

and extended by linearity.

Finally, if $\{e_i\}$ is a basis of H then

$$\{e_{k_1}\otimes\dots\otimes e_{k_i}\} \qquad (1\le k_j<\infty,\ 1\le j\le i)$$

is a basis of $\otimes_i H$ and any tensor Ξ in $\otimes_i H$ may be written

$$\Xi = \sum_{k_1=1}^{\infty} \dots \sum_{k_i=1}^{\infty} \xi_{k_1\dots k_i}\, e_{k_1}\otimes\dots\otimes e_{k_i} \quad ;$$

see also Greub, 1978.

3. General Nonlinear Systems - Without Control

Consider the general autonomous nonlinear (finite-dimensional) differential equation

$$\dot{x} = f(x) \tag{3.1}$$

where f is (real) analytic. Then f has a Taylor series which converges to f in some ball $\{x\varepsilon\mathbb{R}^n : ||x|| < R\}$. Hence, for such a function we may write

$$f(x) = f(o) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(o)x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{1}{2!}\frac{\partial^2 f}{\partial x_i \partial x_j}(o)\, x_i x_j$$

$$+ \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{1}{3!}\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(o)\, x_i x_j x_k + \ldots \tag{3.2}$$

We define the functions

$$\phi_{i_1 \ldots i_n}(x) = x_1^{i_1} \ldots . x_n^{i_n} \tag{3.3}$$

for $i_j \geq 0$, $1 \leq j \leq n$. If $i_1 + \ldots + i_n = m$, then $\phi_{i_1 \ldots i_n}(x)$ appears in the (m+1)th term of the series in (3.2). Hence we may write (3.2) in the form

$$f(x) = f(o) + \sum_{m=1}^{\infty} \sum_{i_1+\ldots+i_n=m} \frac{1}{m!}\left\{\frac{\partial^{i_1+\ldots+i_n} f}{\partial x_1^{i_1} \ldots . \partial x_n^{i_n}}\right\}_s (o)\,\phi_{i_1 \ldots i_n}(x) \tag{3.4}$$

where

$$\left\{\frac{\partial^{i_1+\ldots+i_n} f}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}}\right\}_s = N(i_1, \ldots, i_n)\, \frac{\partial^{i_1+\ldots+i_n}}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}}(f)$$

and $N(i_1, \ldots, i_n)$ is the number of distinct sequences of length $m = i_1 + \ldots + i_n$ containing i_j x_j's $(1 \leq j \leq n)$.

Note that (3.4) merely states that f(x) can be written in terms of the basis $\{\phi_{i_1 \ldots i_n}(x)\}$ of $C_n^{\omega}(R-\varepsilon)$, if $R<\infty$, for any $\varepsilon>0$.

We can 'count' the basis functions in many different ways; however, the two most useful methods may be described as follows. Firstly, we can count the ϕ's antilexicographically within each fixed total order m; i.e. terms of the form $x_{i_1}x_{i_2}\ldots x_{i_m}$, $i_1\geq i_2\geq\ldots\geq i_m$ are counted by allowing i_1 to vary first, then i_2, etc. Secondly, we can count the ϕ's antilexicographically without regard to order; i.e. in (3.3) we let i_1 vary from 0 to ∞, then i_2, etc. We denote the sequence of ϕ's, when totally ordered in either of these ways by $\{\psi_i\}$.

For example, when n=2 the first method corresponds to counting the ϕ's as in the matrix

$$\begin{pmatrix} \psi_o=\phi_{oo} & \psi_2=\phi_{o1} & \psi_5=\phi_{o2} & \psi_9=\phi_{o3} & \cdot \\ \psi_1=\phi_{10} & \psi_4=\phi_{11} & \psi_8=\phi_{12} & \cdot & \\ \psi_3=\phi_{20} & \psi_7=\phi_{21} & \cdot & & \\ \psi_6=\phi_{30} & \cdot & & & \end{pmatrix} \tag{3.5}$$

Similarly the second method corresponds to counting down the columns, i.e.

$$\psi = (\psi_o,\psi_1,\psi_2,\ldots)^T = (\phi_o^T,\phi_1^T,\phi_2^T,\ldots)^T, \tag{3.6}$$

where

$$\phi_j^T = (\phi_{oj},\phi_{1j},\phi_{2j},\ldots)^T, \quad 0\leq j<\infty.$$

Now, using (3.1) it follows that for $i\geq 1$,

$$\frac{d\psi_i}{dt} = \text{grad } \psi_i \cdot \frac{dx}{dt} = \text{grad } \psi_i . f \quad ,$$

(cf. Takata, 1979). Since f is analytic in $\{x:\|x\|<R\}$, so is grad $\psi_i\cdot f$ and so we may write

$$\text{grad}\psi_i \,.f = \sum_{j=o}^{\infty} a_{ij}\psi_j \,.$$

Hence we obtain the equation

$$\dot{\psi} = A\psi \tag{3.7}$$

where

$$A = (a_{ij})_{o\leq i<\infty, o\leq j<\infty}.$$

Hence we can replace the nonlinear equation (3.1) by an infinite dimensional linear equation. It is natural to consider the conditions under which these equations are equivalent in the sense of generating the same solutions. Before making some observations on the semigroup properties of the equation (3.7) we consider the following simple example which illustrates some of the problems which can arise.

Example 3.1 Consider the scalar equation

$$\dot{x} = x^2, \quad x(o) = x_o$$

The solution of this equation is, of course,

$$x(t) = \frac{x_o}{1-tx_o} \quad , \qquad \text{for} \quad o\leq t<1/x_o \tag{3.8}$$

In the scalar case we have the functions

$$\psi_o(x) = 1, \quad \psi_1(x)=x,\ldots,\psi_n(x) = x^n,\ldots,$$

Hence,

$$\dot{\psi}_n(x) = \psi_n'(x)\dot{x} = n\psi_{n+1}(x) \;,$$

and equation (3.7) takes the form

$$\dot{\psi} = \begin{pmatrix} 0 & 0 & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 1 & 0 & \cdots & \cdots \\ 0 & 0 & 0 & 2 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 3 & 0\cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \psi \tag{3.9}$$

Consider this equation to be defined on ℓ^2, and suppose that $o<x_o<1$.

The initial condition for ψ is then

$$\psi(o) = (1, x_o, x_o^2, \ldots) \varepsilon\, \ell^2.$$

From (3.8) the solution remains less than unity until $t = (1-x_o)/x_o$, and so the solution of (3.9) must belong to ℓ^2 on the interval $[o, (1-x_o)/x_o)$. Moreover, although the 'A' matrix of (3.9) is an unbounded operator on ℓ^2 we can calculate the solution of the equation by evaluating

$$e^{At} \triangleq I + At + \frac{A^2 t^2}{2!} + \ldots \tag{3.10}$$

$$= \begin{pmatrix} 1 & \underline{0} \\ \underline{0} & \begin{pmatrix} 1 & t & t^2 & t^3 & \ldots & \ldots & \ldots \\ 0 & 1 & 2t & 2.(3/2!)t^2 & \frac{2.3.4}{3!}t^3 & \ldots & \ldots \\ 0 & 0 & 1 & 3t & \frac{3.4}{2!}t^2 & \frac{3.4.5}{3!}t^3 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix} \end{pmatrix}$$

Note, however, that the series in (3.10) does not converge uniformly for any positive t. In fact it only converges strongly on sets of the form

$$\ell^2(x_o) \triangleq \{\phi \,\varepsilon\, \ell^2 : \phi = (1, x, x^2, x^3, \ldots) \text{ for some } x \text{ with } |x| \leq |x_o|\}$$

for $|x_o| < 1$ and $t < (1-|x_o|)/|x_o|$. (A sequence of linear operators $\{S_i\}$ is strongly convergent on a set X if $S_i x$ converges for any $x \varepsilon X$.) $\ell^2(x_o)$ is not a linear space but we may consider the linear operator $T(t) \triangleq e^{At}$ defined on the linear span of $\ell^2(x_o)$, which we denote by sp $\{\ell^2(x_o)\}$. T(t) is not a semigroup (of bounded operators) on $sp\{\ell^2(x_o)\}$ since it is defined only for $t \varepsilon [o, (1-|x_o|)/|x_o|) \triangleq I_{x_o}$. However, T(t) can be extended to a semigroup of unbounded operators on $[o, \infty)$ in the sense of Hughes (1977), with T(t) bounded on the interval I_{x_o}.

We can associate a semigroup of bounded operators with the equation (3.1) when the origin is asymptotically stable in a region Ω. In fact we have

Lemma 3.2 The matrix A in (3.7) defined as above generates a semigroup of bounded operators on $sp\{\ell^2(\Omega)\}$ where Ω is an invariant set for (3.1), and

$$\ell^2(\Omega) = \{\phi\varepsilon\ell^2 : \phi = (1,x,x^2,\ldots) \quad , \quad x\varepsilon\Omega\},$$

provided $\Omega \subseteq \{x : \|x\| < 1\}$.

Proof. If Ω is an invariant set for the system (3.1) the solution of (3.7) exists for all $t \geq o$ with initial condition $\psi(o)\varepsilon\ell^2(\Omega)$. Writing the solution as $T(t)\psi(o)$, it is clear that $T(t)$ has the semigroup property and the strong continuity of $T(t)$ at $t=0$ follows from the continuity of solutions of (3.1). It is easy to see that $T(t)$ can be extended to a semigroup of bounded operator on $sp\{\ell^2(\Omega)\}$. □

4. Polynomial Systems

The main problem with the theory of section 3 is that the A matrix in (3.7) will have little or no structural pattern in general and will therefore be difficult even to write down. In this section we shall develop a general method for polynomial systems which will make the A operator have a simple form. We use the term 'operator' here rather than 'matrix' since we shall be using tensor representations of the A operator. In order to clarify the development we shall consider first one dimensional systems, then two-dimensional systems and finally the general case.

Let $p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_o$ be a polynomial in the single variable x and consider the equation

$$\dot{x} = p(x).$$

Define the functions $\phi_i(x) = x^i$, $i \geq o$. Then,

$$\frac{d\phi_i(x)}{dt} = \frac{d\phi_i}{dx} \cdot \frac{dx}{dt} = ix^{i-1} \sum_{j=o}^{n} a_j x^j$$

$$= i \sum_{j=o}^{n} a_j x^{j+i-1}$$

$$= i \sum_{j=o}^{n} a_j \phi_{j+i-1}(x).$$

Hence we obtain the linear system

$$\dot{\phi} = A\phi$$

where $\phi=(\phi_o,\phi_1,\phi_2,\ldots)^T$ and

$$A = \begin{pmatrix} 0 & 0 & 0 \ldots\ldots\ldots\ldots & & & \\ a_o & a_1 & a_2 \ldots\ldots\ldots a_n & 0 \ldots 0 \ldots & & \\ 0 & 2a_o & 2a_1 \ldots\ldots\ldots & 2a_n & 0 \ldots & \\ 0 & 0 & 3a_o & 3a_1 \ldots\ldots\ldots & 3a_n & 0 \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix}$$

Now let $\underline{a} = (a_o,a_1,\ldots a_n)$ and let $\underline{\phi}_i = (\phi_i,\phi_{i+1},\ldots,\phi_{i+n})$ be regarded as Cartesian tensors of rank 1. Then with the matrix A we can associate a linear operator L such that

$$(L\phi)_\alpha = \alpha\{C(\underline{a}\otimes\underline{\phi}_{\alpha-1}\} \quad , \quad \alpha\geq 0 \tag{4.1}$$

where C is the contraction operator on n+1 dimensional tensors of rank 1 defined by

$$C(\underline{x}\otimes\underline{y}) = \sum_{i=1}^{n+1} x_i y_i \ .$$

Consider next the two-dimensional system

$$\begin{aligned} \dot{x}_1 &= p_1(x_1,x_2) = \sum_{i=o}^{m_1}\sum_{j=o}^{n_1} a^1_{ij}x_1^i x_2^j \\ \dot{x}_2 &= p_2(x_1,x_2) = \sum_{i=o}^{m_2}\sum_{j=o}^{n_2} a^2_{ij}x_1^i x_2^j \end{aligned} \tag{4.2}$$

where p_1 and p_2 are polynomials. Then, instead of ordering the functions

$$\phi_{i_1 i_2}(x) = x_1^{i_1} x_2^{i_2}$$

as a linear array as in section 3, we consider the ϕ's as a two dimensional array, i.e. as a tensor of rank 2. Then ,

$$\frac{d\phi}{dt}_{\alpha\beta}(x) = \frac{\partial\phi_{\alpha\beta}}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial\phi_{\alpha\beta}}{\partial x_2}\cdot\frac{dx_2}{dt}$$

$$= \alpha x_1^{\alpha-1} x_2^{\beta} \sum_{i=o}^{m_1}\sum_{j=o}^{n_1} a^1_{ij} x_1^i x_2^j + \beta x_1^{\alpha} x_2^{\beta-1} \sum_{i=o}^{m_2}\sum_{j=o}^{n_2} a^2_{ij} x_1^i x_2^j$$

$$= \alpha \sum_{i=o}^{m_1}\sum_{j=o}^{n_1} a^1_{ij} x_1^{i+\alpha-1} x_2^{j+\beta} + \beta \sum_{i=o}^{m_2}\sum_{j=o}^{n_2} a^2_{ij} x_1^{i+\alpha} x_2^{j+\beta-1}$$

$$= \alpha \sum_{i=o}^{m_1}\sum_{j=o}^{n_1} a^1_{ij}\phi_{i+\alpha-1,j+\beta} + \beta \sum_{i=o}^{m_2}\sum_{j=o}^{n_2} a^2_{ij}\phi_{i+\alpha,\, j+\beta-1}$$

$$= \alpha\{C(\underline{a}^1 \otimes \underline{\phi}^1_{\alpha-1,\beta})\} + \beta\{C(\underline{a}^2 \otimes \underline{\phi}^2_{\alpha,\beta-1})\},$$

where $\underline{a}^1 = (a^1_{ij})$, $\underline{a}^2 = (a^2_{ij})$ are tensors of rank 2, and

$$\underline{\phi}^k_{\ell,m} = (\phi_{i+\ell,\, j+m})_{o\le i\le m_k,\ o\le j\le n_k}, \quad k=1,2$$

is also a tensor of rank 2 formed from a submatrix of (ϕ_{ij}).

Moreover, C is the complete contraction operation on cartesian tensors of rank 2, i.e.

$$C(\underline{a} \otimes \underline{b}) = \sum_{i=o}^{m_1}\sum_{j=o}^{n_1} a_{ij}b_{ij},$$

where $\underline{a} = (a_{ij})$, $\underline{b} = (b_{ij})$.

Hence we can write the equations (4.2) in the form

$$\dot{\phi} = L\phi$$

where $\phi = (\phi_{ij})$ and L is the linear operator defined by

$$(L\phi)_{\alpha\beta} = \alpha\{C(\underline{a}^1 \otimes \underline{\phi}^1_{\alpha-1,\beta})\} + \beta\{C(\underline{a}^2 \otimes \underline{\phi}^2_{\alpha,\beta-1})\}$$

Example 4.1 Consider the Van der Pol oscillator

$$\dot{x}_1 = x_2 - x_1^3 + x_1$$
$$\dot{x}_2 = -x_1$$

In this case it is easy to see that

$$\underline{a}^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \\ -1 & 0 \end{pmatrix}, \quad \underline{a}^2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

(since $m_1 = 3$, $n_1 = 1$, $m_2 = 1$, $n_2 = 0$). The solution of the Van der Pol oscillator is given by the elements (1,0) and (0,1) of the tensor $e^{Lt}\,\phi(o)$, where

$$(\phi(o))_{ij} = x_1^i(o)x_2^j(o) \quad , i , j \geq o.$$

If Ω denotes the union of the limit cycle of the system and its interior then e^{Lt} is a semigroup of bounded operators on $sp\{\ell^2(\Omega)\}$ by lemma 3.2 (if Ω is scaled so that $\Omega\subseteq\{x:||x||<1\}$.)

Finally we consider the general case of the polynomial system

$$\dot{x}_j = p_j(x_1,\dots, x_n) = \sum_{i_1=o}^{m_1^j} \sum_{i_2=0}^{m_2^j} \dots \sum_{i_n=0}^{m_n^j} a^j_{i_1 i_2\dots i_n} x_1^{i_1} x_2^{i_2}\dots x_n^{i_n} \quad (4.3)$$

$(1\leq j\leq n)$. Then we let $\phi_{i_1\dots i_n}(x)$ be the functions defined by (3.3), and so

$$\frac{d}{dt}\phi_{\alpha_1\alpha_2\dots\alpha_n}(x) = \sum_{k=1}^{n} \frac{\partial\phi_{\alpha_1\alpha_2\dots\alpha_n}}{\partial x_k}\cdot\frac{dx_k}{dt}$$

$$= \sum_{k=1}^{n} \alpha_k x_1^{\alpha_1}\dots x_{k-1}^{\alpha_{k-1}} x_{k+1}^{\alpha_{k+1}}\dots x_n^{\alpha_n} p_k(x_1,\dots,x_n)$$

$$= \sum_{k=1}^{n} \alpha_k \sum_{i_1=o}^{m_1^k} \sum_{i_2=o}^{m_2^k}\dots \sum_{i_n=o}^{m_n^k} a^k_{i_1 i_2\dots i_n} \phi_{i_1+\alpha_1,\dots,i_k+\alpha_k-1,\dots,i_n+\alpha_n}.$$

As before, let $\underline{a}^k = (a^k_{i_1 i_2\dots i_n})$ be a tensor of rank n and let

$$\underline{\phi}^k_{\ell_1\ell_2\dots\ell_n} = (\phi_{i_1+\ell_1,\dots,i_n+\ell_n}) \quad 0\leq i_1\leq m_1^k, \dots, 0\leq i_n\leq m_n^k \qquad (1\leq k\leq n).$$

Then we have

$$\dot{\phi} = L\phi$$

where $\phi = (\phi_{i_1\dots i_n})$ and L is the linear operator defined by

$$(L\phi)_{\alpha_1\dots\alpha_n} = \sum_{k=1}^{n} \alpha_k\{C(\underline{a}^k \otimes \underline{\phi}^k_{\alpha_1,\dots,\alpha_{k-1},\alpha_k-1,\alpha_{k+1},\dots,\alpha_n})\} \qquad (4.4)$$

Example 4.2 An interesting example in three dimensions is the so-called Lorenz or strange attractor, defined by the equations

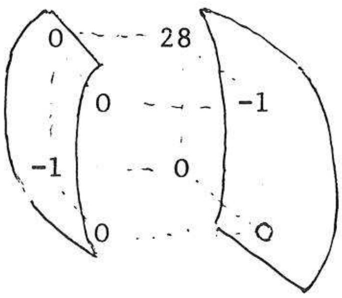
$$\dot{x}_1 = -10x_1 + 10x_2$$

$$\dot{x}_2 = -x_1x_3 + 28x_1 - x_2$$

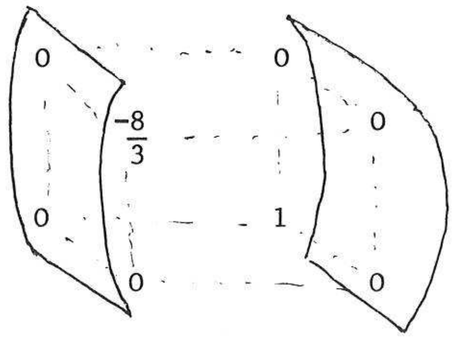
$$\dot{x}_3 = x_1x_2 - \frac{8}{3}x_3$$

(See Marsden and McCracken, 1976). In this case we have

$$\underline{a}^1 = \begin{pmatrix} 0 & -10 \\ 10 & 0 \end{pmatrix} \quad , \quad \underline{a}^2 =$$



$$\underline{a}^3 =$$



where we have indicated pictorially that $\underline{a}^2$ and $\underline{a}^3$ are tensors of rank 3 (of orders 2x2x2).

In order to proceed with the theory we can write (4.4) in a different form. Let $\{e_i\}$ denote the standard basis of ℓ^2; i.e.

$$e_i = (0,0, \ldots, 0, \overset{\substack{i^{th}\text{ place}\\ \downarrow}}{1}, 0, \ldots)$$

Then the space of tensors of rank n defined on ℓ^2 is spanned by the vectors

$$e_{i_1} \otimes e_{i_2} \otimes \ldots \otimes e_{i_n} \quad , \quad \le i_j < \infty \, .$$

Let $P^k_{\alpha_1, \ldots, \alpha_n}$ denote the projection on the subspace spanned by the vectors

$$e_{i_1} \otimes e_{i_2} \otimes \ldots \otimes e_{i_n} \quad , \quad \alpha_\ell + 1 < i_\ell \le \alpha_\ell + 1 + m^k_\ell \quad \text{for } 1 \le \ell \le n.$$

Then we can write (4.4) in the form

$$(L\phi)_{\alpha_1 \dots \alpha_n} = \sum_{k=1}^{n} \alpha_k \{ C(\underline{a}^k \otimes P^k_{\alpha_1, \dots, \alpha_{k-1}, \alpha_k -1, \alpha_{k+1}, \dots, \alpha_n}) \} \phi$$

(again interpreting $P^k_{\alpha_1, \dots, \alpha_n} = 0$ if any $\alpha_i < 0$). Writing

$$P^k_{\underline{\alpha}} = \alpha_k P^k_{\alpha_1, \dots, \alpha_{k-1}, \alpha_k -1, \alpha_{k+1}, \dots, \alpha_n}$$

we can express L simply as the operator

$$L = \sum_{k=1}^{n} \{ C(\underline{a}^k \otimes P^k) \}$$

where $\underline{a}^k \otimes P^k$ is the tensor operator given by

$$(\underline{a}^k \otimes P^k)\alpha_1, \dots, \alpha_n = \underline{a}^k \otimes P^k_{\underline{\alpha}}$$

$$= \alpha_k \underline{a}^k \otimes P^k_{\alpha_1, \dots, \alpha_{k-1}, \alpha_k -1, \alpha_{k+1}, \dots, \alpha_n} .$$

Now let $\otimes_n(\ell^2)$ denote the tensor product of n copies of ℓ^2. This is, of course, a vector space. We can make it into a Hilbert space by defining the inner product

$$\langle \underline{a}, \underline{b} \rangle = \sum_{i_1=1}^{\infty} \dots \sum_{i_n=1}^{\infty} a_{i_1 \dots i_n} b_{i_1 \dots i_n} = C(\underline{a} \otimes \underline{b})$$

where

$$\underline{a} = a_{i_1 \dots i_n} e_{i_1} \otimes \dots \otimes e_{i_n}$$

$$\underline{b} = b_{i_1 \dots i_n} e_{i_1} \otimes \dots \otimes e_{i_n}$$

(using the summation convention).

For any tensor ϕ of the form

$$\phi = \phi_{i_1 \dots i_n}(x) \, e_{i_1} \otimes \dots \otimes e_{i_n} = x_1^{i_1 - 1} \dots x_n^{i_n - 1} \, e_{i_1} \otimes \dots \otimes e_{i_n}$$

we have $\phi \varepsilon \otimes_n(\ell^2)$ if $|x_i| < 1$, $1 \le i \le n$. For,

$$\langle \phi , \phi \rangle = \sum_{i_1 = o}^{\infty} \dots \sum_{i_n = o}^{\infty} x_1^{2i_1} \dots x_n^{2i_n}$$

$$= \prod_{i=1}^{n} \frac{1}{1 - x_i^2} .$$

As we saw earlier, therefore, we can see that if $\Omega \subseteq \{\underline{x}; |x_i| < 1, 1 \leq i \leq n\}$ is an invariant set for the flow of the equation (4.3), then the operator L generates a semigroup on $sp\{\ell^2(\Omega)\}$, given by

$$T(t) = e^{Lt}$$

where the series for e^{Lt} converges strongly to $T(t)$ on $sp\{\ell^2(\Omega)\}$.

Hence,

$$T(t) = \sum_{i=o}^{\infty} \frac{L^i t^i}{i!} \tag{4.5}$$

where

$$L^i = \left(\sum_{k=1}^{n} \{C(\underline{a}^k \otimes P^k)\} \right)^i$$

$$= \sum_{\substack{k_1=o \\ k_1+k_2+\ldots+k_n = i}}^{i} \sum_{k_2=o}^{i} \cdots \sum_{k_n=o}^{i} \binom{i}{k_1 k_2 \ldots k_n} \prod_{j=1}^{n} \{C(\underline{a}^j \otimes P^j)\}^{k_j}.$$

Recall that

$$\binom{i}{k_1\ k_2 \ldots k_n} = \frac{i!}{k_1! k_2! \ldots k_n!}$$

is the multinomial coefficient and we have used the multinomial theorem.

It is interesting to note the amount of numerical computation and storage requirements necessary to evaluate (4.5) to a given power of t.

If $(x_{10}, x_{20}, \ldots, x_{n0})^T$ is the initial condition of the equation (3.1), then that of the equation $\dot{\phi} = L\phi$ is $\phi(o) = (x_{10}^{i_1} \ldots x_{n0}^{i_n})_{o \leq i_1, i_2, \ldots, i_n < \infty}$. Now the solution of (3.1) is given by

$$x_i(t) = \phi_{o,o,\ldots,\underset{\substack{\uparrow \\ i^{th}\text{place}}}{1},o,\ldots,o}(t) \quad , \quad 1 \leq i \leq n \ ,$$

For ℓ given, let

$$\ell_1 = \max\{(\ell-1)m_1^1, \ell m_1^2 + 1, \ell m_1^3 + 1, \ldots, \ell m_1^n\}$$

$$\ell_2 = \max\{\ell m_2^1 + 1, (\ell-1)m_2^2, \ell m_2^3 + 1, \ldots, \ell m_2^n\}$$

$$\cdots\cdots$$

$$\ell_n = \max\{\ell m_n^1 + 1, \ell m_n^2 + 1, \ldots, (\ell-1)m_n^n\}$$

Then it is easy to see that, in order to calculate $\underline{x}(t)$ to order ℓ in t we must evaluate the subtensor of $\phi(o)$ with $\phi_{ooo\ldots o}(0)$ in the 'top left hand corner' of size $\ell_1 x \ell_2 x \ldots x \ell_n$.

5. Nonlinear Control Systems

Consider now the general nonlinear control system

$$\dot{x}(t) = f(x(t), u(t)) \tag{5.1}$$

where $x(t)\epsilon\mathbb{R}^n$, $u(t)\epsilon\mathbb{R}^m$ for each t. Suppose that we restrict attention to differentiable controllers u(t) so that we may write

$$\dot{u}(t) = v(t) \tag{5.2}$$

for some function v. Then we can consider the augmented system

$$\begin{pmatrix} \dot{x}(t) \\ \dot{u}(t) \end{pmatrix} = \begin{pmatrix} f(x(t),u(t)) \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ v(t) \end{pmatrix} \tag{5.3}$$

where the control v now appears linearly. Introduce the sequence of Taylor polynomials $x_1^{i_1}\ldots x_n^{i_n} u_1^{j_1}\ldots u_m^{j_m}$ arranged in some order, say $\phi_o(x,u),\phi_1(x,u)\ldots$, as before. Then

$$\begin{aligned}
\frac{d\phi_k}{dt} &= \frac{\partial\phi_k}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial\phi_k}{\partial u}\cdot\frac{du}{dt} \\
&= \frac{\partial\phi_k}{\partial x}\cdot f(x,u) + \frac{\partial\phi_k}{\partial u}\cdot v \\
&= \sum_{\ell=o}^{\infty} A_{k\ell}\phi_\ell(x,u) + \sum_{j=1}^{m}\left(\sum_{\ell=o}^{\infty} B^j_{k\ell}\,\phi_\ell(x,u)\right) v_j
\end{aligned}$$

for some infinite matrices $A = (A_{k\ell})$, $B^j = (B^j_{k\ell})$, $1\leq j\leq m$. Hence we obtain the infinite dimensional bilinear system

$$\dot{\phi} = A\phi + \sum_{j=1}^{m} (B^j\phi)\cdot v_j$$

For simplicity of exposition assume that m=1 (i.e. a scalar control). Then we have the system

$$\dot{\phi} = A\phi + B\phi v. \tag{5.4}$$

We can now apply various results on the control of infinite-dimensional bilinear systems which have already been developed. Here we shall consider the application of optimal control of bilinear systems to general nonlinear systems.

Suppose we consider the optimal control problem of minimising the quadratic cost functional

$$J = \phi^T(t_f)F\phi(t_f) + \int_{t_o}^{t_f} (\phi^T(t)\,Q\,\phi(t) + rv^2)\ dt \tag{5.5}$$

subject to the dynamics (5.4). Note that this is a reasonable functional to minimise with respect to system (5.3) since the ϕ vector contains all the monomials $x_1^{i_1} \ldots x_n^{i_h} u^j$, $i_1, \ldots, i_n, j \geq 0$ and so we are weighting all the states, the control and its derivative v simultaneously in (5.5). The 'bilinear-quadratic' regulator problem for infinite dimensional systems has recently been considered (Banks and Yew, 1984) and we may state the result as follows. Suppose we regard the system (5.4) and the cost (5.5) as being defined on a Hilbert space H. Then it can be shown that the optimal control v is given formally by

$$v = -r^{-1} \sum_{i=1}^{\infty} \langle \otimes_i \phi, (P_i B) \otimes_i \phi \rangle_{\otimes_i H} \tag{5.6}$$

where $\otimes_i H = H \otimes \ldots \otimes H$ is the tensor product of i copies of H, $\otimes_i \phi = \phi \otimes \ldots \otimes \phi \in \otimes_i H$, $P_i \in \mathcal{L}(\otimes_i H)$ and $P_i B$ is defined by

$$P_i B = \sum_{j=1}^{i} \left(\sum_{k_j=1}^{\infty} P_{i,k_1 \ldots k_j \ldots k_i}^{k_1 \ldots k_i} B_{k_j \ell_1} \right)_{(1 \leq \ell_1 < \infty, 1 \leq k_1 < \infty, \ldots, 1 \leq k_i < \infty, 1 \leq \alpha < \infty)}$$

where $P_{i,k_1,\dots,k_i}^{\kappa_1\dots\kappa_i}$, $B_{k_j}^{\ell_1}$ are the components of the tensors P_i, B with respect to some (fixed) basis of H. Moreover, the tensor operators P_i are given by

$$P_1(t) = e^{\mathcal{A}_1(t_f-t)} G e^{\mathcal{A}_1^T(t_f-t)} + \int_0^{t_f-t} e^{\mathcal{A}_1(t_f-t-s)} M e^{\mathcal{A}_1^T(t_f-t-s)} ds$$

and

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

$m\geq 2$ where $\mathcal{A}_i P_i = P_i A$. Hence the control u of the nonlinear system (5.1) is given by

$$u(t) = u_o + \int_o^t v(t)dt = u_o - r^{-1} \sum_{i=1}^{\infty} \int_o^t \langle \otimes_i \phi, (P_i B) \otimes_i \phi \rangle_{\otimes_i H}$$

where the ϕ vector is formed from elements of the form $x_1^{i_1} \dots\dots x_n^{i_n} u^j$. Equation (5.6) may therefore be written in the form

$$\dot{u} = g(u,x,t) \tag{5.7}$$

for some function g. (The explicit t dependence comes from P_i)

The control system (5.1),(5.7) can be implemented as in fig. 1.

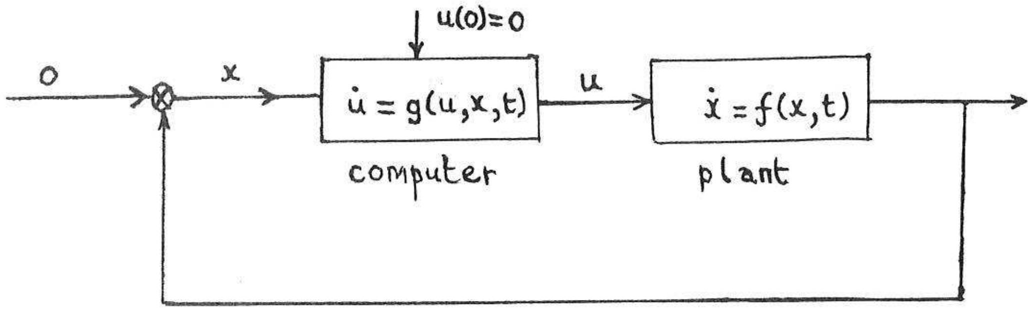


fig. 1.

We have taken u(o) = o since we are trying to minimise u. However any other value for u could be chosen. Note that in a real implementation we must consider only a finite number of the ϕ_i functions in which case (5.4) will be only a finite dimensional approximation to (5.1). In addition we can truncate the power series in (5.6) for v.

6. Example

Consider the simple scalar bilinear system

$$\dot{x} = ax + ubx \tag{6.1}$$

Let the ϕ_i's be ordered as in (3.5), i.e.

$$\phi = 1,\ \phi_2 = x,\ \phi_3 = u,\ \ \phi_4 = x^2,\ \phi_5 = xu,\ \ \phi_6 = u^2, \ldots$$

To evaluate a general expression for ϕ_i note that there exists ℓ such that $\frac{1}{2}\ell(\ell+1) < i \leq \frac{1}{2}(\ell+1)(\ell+2)$. Put $q = i - \frac{1}{2}\ell\ (\ell+1)$.

Then $\phi_i = x^{\ell-q}\ u^q$, $i \geq 2$ and $\phi_1 = 1$. To illustrate the application of the control (5.7) we shall truncate the ϕ vector to the first six terms; i.e. $\phi = (1,x,u,x^2,xu,u^2)^T$. It is easy to see that (5.4) becomes

$$\dot{\phi} = A\phi + vB\phi$$

where

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a & 0 & 0 & b & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2a & 0 & 0 \\ 0 & 0 & 0 & 0 & a & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \end{pmatrix}$$

For simplicity let G = I, M = I. Then

$$e^{A(t_f-t)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & e^{a(t_f-t)} & 0 & 0 & be^{a(t_f-t)} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & e^{2a(t_f-t)} & 0 & 0 \\ 0 & 0 & 0 & 0 & e^{a(t_f-t)} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Hence

$$P_1(t) = \begin{pmatrix} 1+t_f-t & 0 & 0 & 0 & 0 & 0 \\ 0 & e_2 & 0 & 0 & be_2 & 0 \\ 0 & 0 & 1+t_f-t & 0 & 0 & 0 \\ 0 & 0 & 0 & e_4 & 0 & 0 \\ 0 & be_2 & 0 & 0 & e_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1+t_f-t \end{pmatrix}$$

where

$$e_k = e^{ka(t_f-t)} + \frac{1}{ka}\cdot(e^{ka(t_f-t)}-1) \quad , \; k = 2 \text{ or } 4.$$

$P_2(t)$ can now be found in a similar way by using $P_1(t)$.

Note that, instead of counting the ϕ's linearly as above we could have defined $\phi_{ij} = x^i u^j$ as in section 4. In this case the Hilbert space H is itself a space of tensors (for the example above, they are just matrices). H has a basis consisting of tensors e_{ij} with $k\ell^{th}$ element $\delta_{ij}^{k\ell}$ and any tensor in H can be represented in terms of this basis. If $\zeta \in H$ is given by $\zeta = \Sigma\zeta_{ij}e_{ij}$, then the norm on ζ is just $(\Sigma_i\Sigma_j\zeta_{ij}^2)^{\frac{1}{2}}$.

7. Conclusions

In this paper we have considered the representation of nonlinear systems by linear infinite dimensional systems and we have obtained a particularly simple tensor operator representation for polynomial systems. It appears that for finite dimensional systems with finite escape times the infinite dimensional representation defines a bounded semigroup of operators on an appropriate Hilbert space until the escape time, when the simigroup becomes unbounded.

We have applied the representation theory to show that any nonlinear control system can be replaced by an equivalent infinite dimensional bilinear system and applied some recent results on suboptimal control to the resulting system. Note that it has also been shown recently (Banks,1984) that we may even apply these ideas to infinite dimensional nonlinear systems.

8. References

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 Yew, M.K. (1984) On a Class of Suboptimal Controls for Infinite-Dimensional Bilinear Systems, Research Report No. 258 , Sheffield.

Bergen, A.R. and Franks, R.L. (1971). Justification of the Describing Function Method, SIAM J. Contr., 9, pp 568-589.

Brockett, R. (1978) Feedback Invariants for Nonlinear Systems, IFAC Congr. Helsinki, Finland.

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

Hughes, R.J. (1977) Semigroups of Unbounded Linear Operators in Banach Space, Trans. Am. Math. Soc., 230, pp 113-145.

Hunt, L.R., Su. R. and Meyer, G. (1983). Global Transformations of Nonlinear Systems, IEEE Trans. Aut. Control, AC-28, pp 24-31.

Krener, A.J. (1973) On the Equivalence of Control Systems and the Linearization of Nonlinear Systems, SIAM J. Contr., 11, pp 670-676.

Mees, A.I. and Bergen, A.R. (1975). Describing Functions Revisited, IEEE Trans. Aut. Contr., AC-20, pp 473-478.

Sandberg. I.W. (1981) Global Implicit Function Theorems, IEEE Trans. Circuits Syst., CAS-28, pp 145-149.