LIE ALGEBRAS AND LIE GROUPS IN CONTROL THEORY

Roger W. Brockett

Harvard University

Cambridge, Massachusetts

I. THE ALGEBRAIC THEORY OF LINEAR DIFFERENTIAL EQUATIONS
1.1 Lie Algebras and Linear Differential Equations
1.2 The $x^{[p]}$ and $x^{(p)}$ Equations
1.3 Matrix Lie Algebras and The Matrix Exponential
1.4 Cones and Semigroups

II. CONTROLLABILITY, OBSERVABILITY, AND SYSTEMS
2.1 Controllability
2.2 Observability
2.3 Isomorphic Systems

III. OPTIMAL CONTROL
3.1 Bang-Bang Theorems
3.2 Least Squares Theory

IV. STOCHASTIC DIFFERENTIAL EQUATIONS
4.1 Bilinear Stochastic Equations
4.2 The Moment Equations
4.3 Fokker-Planck Equations
4.4 Calculation of Diffusion Times

V. STABILITY THEORY
5.1 Stability of the $x^{[p]}$ Equations
5.2 Periodic Self-Contragradient Systems
5.3 Symplectic Systems

VI. REFERENCES

---

*This work was supported in part by the U.S. Office of Naval Research under the Joint Services Electronics Program by Contract N00014-67-A-0298-0006 and by the National Aeronautics and Space Administration under Grant NGR 22-007-172.

PREFACE

The theory of differential equations and control have been linked very closely because most of the early applications of control theory were to engineering problems of the type which are most naturally described by ordinary differential equations. The questions of importance in control have helped to revitalize certain problem areas in differential equations and methods and tools from control have been useful in obtaining new results in differential equation theory. On the other hand, going back to the era of Lie himself, there has been close ties between Lie theory and differential equations. Thus it is not surprising that one finds that Lie theory and control are also closely connected. This "triangle" is the subject of this set of notes.

In control theory, Lie algebras make their appearance as Lie algebras of vector fields. Topological properties associated with Lie groups show up in the study of controllability and stability. Partial differential operators arise in the Fokker-Planck equations modeling the uncertainty of the environment and our uncertainty about the measurements we make of it. The problems which are of interest in control frequently require a generalization of the usual treatment of topics such as existence of geodesics, expressions for the spectrum of the Laplacian etc. The modification is, roughly speaking, to include the possibility of a metric which is "infinite" in certain directions, subject only to the condition that the directions along which it is finite can be combined in such a way as to make the distance between any two points finite. These notes contain a brief account of some of these topics, together with references where complete proofs can be found.

I have included a few exercises for the reader, both to indicate some results which do not exactly fit the format chosen here and to indicate some partial results and suggestions on additional problems of interest. Most of the examples are to be found in the exercises as well.

It is a pleasure to thank Prof. David Mayne for organizing such a stimulating forum for the exchange of ideas on system theory.

I. THE ALGEBRAIC THEORY OF LINEAR DIFFERENTIAL EQUATIONS

1.1 Lie Algebras and Linear Differential Equations

Clearly any linear differential equation of the form

$$\dot{x}(t) = A(t)x(t); \quad x(t) \in \mathbb{R}^n$$

can be expressed as

$$\dot{x}(t) = \left(\sum_{i=1}^{m} u_i(t)A_i\right)x(t)$$

with the A_i constant matrices and the $u_i(t)$ scalar functions of time. In view of the fact that the solution of the equation with a single A_i, i.e.

$$\dot{x}(t) = u(t)Ax(t)$$

is

$$x(t) = e^{A\int_0^t u(\sigma)d\sigma} x(0)$$

the question arises as to when the solution of the general problem can be written as the composition of a number of such solutions

$$x(t) = e^{A_1 g_1(t)} e^{A_2 g_2(t)} \dots e^{A_m g_m(t)} x(0)$$

for a suitable choice of the $g_i(\cdot)$. Otherwise stated, we would like to know if the solutions of the matrix differential equation

$$\dot{X}(t) = (\sum_{i=1}^{m} u_i(t)A_i)X(t); \quad X(0) = I \quad \text{(identity)}$$

can be written as

$$X(t) = e^{A_1 g_1(t)} e^{A_2 g_2(t)} \dots e^{A_m g_m(t)}$$

for a reasonably wide class of $u_i(t)$ and over some interval of time, say $|t| < \varepsilon$.

The above question is basically answered by a classical theorem of Frobenius [1]. However the theorem of Frobenius which applied here is a theorem in differential geometry. To use the insight of his result we need to look at the problem posed from a geometrical point of view. Consider the identity matrix as a point in the set of all nonsingular n by n matrices. Suppose that the one parameter curves $e^{A_i t}$ leave the identity as indicated in figure 1.

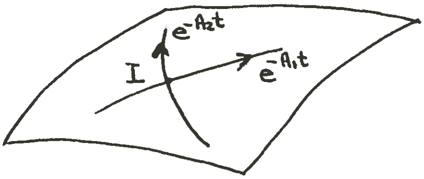

Figure 1: Neighborhood of I in the set of all n by n matrices

We regard the set of all points of the form

$$S = \{X : X = \prod_{i=1}^{m} e^{A_i \alpha_i}; \quad \alpha_i \in \mathbb{R}\}$$

as a subset of the set of all nonsingular n by n matrices. Our question is, when do the integral curves of the given matrix differential equation corresponding to a wide class of $u_i(\cdot)$ lie in S? In order for this to be true for all piecewise continuous u's we require, for example, that

$$e^{A_1 t} e^{A_2 t} e^{-A_1 t} e^{-A_2 t}$$

be expressible as an element of S. To see why this is so we point out that the choice

$$u_1(\sigma) = \begin{cases} -1 & t \leqslant \sigma < 2t \\ 0 & 0 \leqslant \sigma < t; \quad 2t \leqslant \sigma < 3t \\ 1 & 3t \leqslant \sigma < 4t \end{cases}$$

$$u_2(\sigma) = \begin{cases} -1 & 0 \leqslant \sigma < t \\ 0 & t \leqslant \sigma < 2t; \quad 3t \leqslant \sigma < 4t \\ 1 & 2t \leqslant \sigma < 3t \end{cases}$$

$$u_i(\sigma) = 0 \quad i > 2$$

yields

$$X(4t) = e^{A_1 t} e^{A_2 t} e^{-A_1 t} e^{-A_2 t}$$

Geometrically, what we are asking is that in following the 4-sided path shown in figure 2 we should not be lead out of the set S.

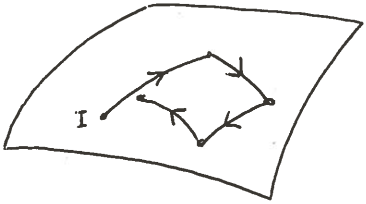

Figure 2: Illustrating the path leading to $e^{A_1 t} e^{A_2 t} e^{-A_1 t} e^{-A_2 t}$

More generally if f_1 and f_2 are smooth maps of $\mathbb{R}^n$ into $\mathbb{R}^n$ and if we apply the above choice of $u(\cdot)$ to the system

$$\dot{x}(t) = u_1(t)f[x(t)]+u_2(t)g[x(t)]; \quad x(0) = x_0$$

then a slightly messy calculation shows that to second order in t we have

$$x(4t) \simeq x_0 + \{(\frac{\partial f}{\partial x})_{x=x_0} g(x_0) - (\frac{\partial g}{\partial x})_{x=x_0} f(x_0)\}t^2$$

The quantity $\frac{\partial f}{\partial x} g(x) - \frac{\partial g}{\partial x} f(x)$ is usually written as $[f,g]$ and is called the Lie bracket of f and g. One calls a set of vectors $f_i : \mathbb{R}^n \to \mathbb{R}^n$ <u>involutive</u> if the Lie bracket of any two is a linear combination of the $\{f_i\}$. Frobenius showed that the set of points near x_0 which can be reached from x_0 along integral curves of

$$\dot{x}(t) = \sum_{i=1}^{m} u_i(t) f_i(x)$$

with $\{f_i\}$ involutive can be expressed as

$$\phi_m(t_m, .., \phi_3(t_3, \phi_2(t_2, \phi_1, x)))...)$$

where $\phi_i(t,x)$ are the solutions of

$$\dot{x}(t) = f_i[x(t)]$$

The reason the set $\{f_i\}$ must be involutive is that otherwise the special choice of $u(\cdot)$ outlined above will, for small t, surely lead out of set of points expressible as $\phi_m(t_m,\phi(t_{m-1}, \ldots\phi(t,x_o))\ldots))$.

Applying this type of thinking to the linear case, we see first of all that the Lie bracket of A_1x and A_2x is $[A_1x, A_2x]=(A_1A_2-A_2A_1)x$ That is, the Lie bracket of the vector fields is expressible as the commutator of the matrices. We write $[A_i,A_j]$ for $A_iA_j-A_jA_i$. Thus if the set of matrices $\{A_i\}$ have the property that

$$[A_i,A_j] = \sum_{k=1}^{m} \gamma_{ijk}A_k$$

then the theorem of Frobenius would imply that for small $|t|$ we can write

$$X(t)x_o = \prod_{i=1}^{m} e^{A_ig_i(t)} x_o$$

A linear space of square matrices which is closed under $[\cdot,\cdot]$ is a matrix Lie algebra. Of course if the original set $\{A_i\}$ does not form a basis for a Lie algebra we simply supplement it with additional A's until it does. If x is of dimension n then there are only n^2 linearly independent matrices so this process always results in a finite set.

Wei and Norman [2] have given a direct verification of the above representation based on the implicit function theorem and have developed a set of nonlinear differential equations for the $g_i(\cdot)$. The basis for their derivation is the Baker-Campbell-Hausdorff formula

$$e^ABe^{-A} = B+[A,B]+ \frac{1}{2} [A,[A,B]+ \frac{1}{3!} [A,[A,[A,B]]\ldots]$$

Thus if one assumes a solution of the form

$$X(t)X_o = e^{A_1g_1(t)} e^{A_2g_2(t)} \ldots e^{A_mg_m(t)}$$

and then differentiates, the result is

$$\dot{X}(t) = A_1\dot{g}_1(t)e^{A_1g_1(t)} e^{A_2g_2(t)} \ldots e^{A_mg_m(t)}$$

$$+ e^{A_1g_1(t)} A_2\dot{g}_2(t)e^{A_2g_2(t)} \ldots e^{A_mg_m(t)}$$

$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot$$

$$+ e^{A_1g_1(t)} e^{A_2g_2(t)} \ldots A_m\dot{g}_m(t)e^{A_mg_m(t)}$$

Now we must collect all the A's together at the left in order to compare this expression for $\dot{X}$ with that given by the differential equation. The Baker-Campbell-Hausdorff formula provides the means to do this. To see how this happens, observe that by inserting $e^{-A_ig_i(t)} e^{A_ig_i(t)}$ freely we can arrive at

$$\dot{g}_1A_1+\dot{g}_2e^{A_1g_1(t)}A_2e^{-A_1g_1(t)}+\ldots\dot{g}_me^{A_1g_1(t)}e^{A_2g_2(t)}\ldots A_m\ldots e^{-A_1g_1(t)}$$

$$= A_1u_1(t)+A_2u_2(t) + \ldots + A_mu_m(t)$$

We apply the Baker-Campbell-Hausdorff expansion to each term on the left. If the set $\{A_i\}$ is a basis for a Lie algebra then we can express the result as a linear combination of the A_i. Since the A_i are linearly independent we can equate coefficients on each side and thereby get a set of differential equations for the g_i. It is important to note that the differential equations for the g_i only depend on the A_i through the commutation rules

$$[A_i,A_j] = \sum_{k=1}^{m} \gamma_{ijk}A_k$$

Thus when a differential equation is solved by this method a whole class of differential equations are solved at the same time -- one for each set of A's which satisfy the given commutation relation.

Exercises

1. Show that if the A_i in

$$\dot{X}(t) = \sum_{i=1}^{m} u_i(t)A_iX(t)$$

are all upper triangular then it is possible to express the solution of the differential equations for the $g_i(\cdot)$ explicitly in terms of integrals.

2. Show that the smallest Lie algebra of matrices which contains A_1 and A_2

$$A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} ; \quad A_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

is 4 dimensional.

3. Study the definition of Euler angles from the point of view of the Wei-Norman equations. In particular explain why it is generally not possible to obtain a Wei-Norman representation the entire half-line $[0,\infty)$ in terms of the degeneracy of the Euler angles.

4. Show that for any square matrix P the set of all solutions of $PA+A'P = 0$ from a Lie algebra.

1.2 The $x^{[p]}$ and $x^{(p)}$ Equations

Associated with each linear map of $\mathbb{R}^n$ into $\mathbb{R}^n$ are two families of linear maps which may be described as follows. Choose a basis in $\mathbb{R}^n$ and let the original map be represented by the matrix A. Then we easily see that

$$y_i = \sum_{j=1}^{n} a_{ij}x_j$$

implies that the n(n+1)/2 linearly independent terms of the form $y_i y_j$ depend linearly on the n(n+1)/2 linearly independent terms of the form $x_i x_j$. More generally the set of all linearly independent p-degree terms $y_i y_j \ldots y_k$ depend linearly on the set of all linearly independent p-degree terms $x_i x_j \ldots x_k$. How many linearly independent terms of degree p are there in n variables? If we denote this integer by N_n^p then it is easy to see that

$$N_{n+1}^{p+1} = N_{n+1}^{p} + N_{n}^{p+1}$$

from which an induction gives $N_n^p = \binom{n+p-1}{p}$. Thus associated with each map of $\mathbb{R}^n$ into $\mathbb{R}^n$ is a sequence of maps, the pth one mapping $\mathbb{R}^{N_n^p}$ into $\mathbb{R}^{N_n^p}$.

In order to give this family of maps a matrix description we need to choose a basis in $\mathbb{R}^{N_n^p}$ which is in some way convenient. The principle which guides our choice of basis is this: let $\langle x,y\rangle$ be the ordinary inner product

$$\langle x,y\rangle = \sum_{i=1}^{n} x_i y_i$$

If the map of $\mathbb{R}^n$ into $\mathbb{R}^n$ defined by A preserves length, we would like the maps of $\mathbb{R}^{N_n^p}$ into $\mathbb{R}^{N_n^p}$ to preserve length as well. To achieve this we introduce the basis elements

$$\sqrt{\binom{p}{p_1}\binom{p-p_1}{p_2}\cdots\binom{p-p_1-\cdots p_{p-1}}{p_p}}\; x_1^{p_1} x_2^{p_2} \ldots x_n^{p_n};\quad \sum_{i=1}^{n} p_i = p;\; p_i \geq 0$$

For example if n=p=3 we have basis elements

$$x_1^3,\ \sqrt{3}x_1^2x_2,\ \sqrt{3}x_1^2x_3,\ \sqrt{3}x_1x_2^2,\ \sqrt{6}x_1x_2x_3,\ \sqrt{3}x_1x_3^2,\ x_2^3,\ \sqrt{3}x_2^2x_3,\sqrt{3}x_2x_3^2,x_3^3$$

If we denote this vector, ordered lexigraphically, by $x^{[p]}$ then the choice of basis is such that $(||x|| = (\langle x,x\rangle)^{1/2})$

$$||x^{[p]}|| = ||x||^p$$

More generally, we have

$$\langle x,y\rangle^p = \langle x^{[p]}, y^{[p]}\rangle$$

We denote by $A^{[p]}$ the map, or matrix, which verifies

$$y = Ax \Rightarrow y^{[p]} = A^{[p]}x^{[p]}$$

The principle properties of $A^{[p]}$ are covered by the following theorem.

Theorem 1: Suppose we are given A and B. $A : \mathbb{R}^n \to \mathbb{R}^n$ and $B : \mathbb{R}^n \to \mathbb{R}^n$. Then $A^{[p]}$ and $B^{[p]}$ satisfy

i) $I_n^{[p]} = I_{N_n^p}$

ii) $(AB)^{[p]} = A^{[p]}B^{[p]}$

iii) $(A^q)^{[p]} = (A^{[p]})^q$; q integer; A^q defined

iv) $(A')^{[p]} = (A^{[p]})'$

Proof: i) Clear from definition. ii) Let z=Ay=ABx. Then $z^{[p]}=A^{[p]}y^{[p]}=A^{[p]}B^{[p]}x^{[p]}=[AB]^p x^{[p]}$. iii) This follows from ii) on letting B=A (or $B=A^{-1}$ if A is invertible) and using induction. iv) This follows from the identity $\langle x,y\rangle^p=\langle x^{[p]},y^{[p]}\rangle$ and $\langle x,Ay\rangle = \langle A'x,y\rangle$.

A second series of maps associated with A are the so called compounds of A which we write as $A^{(p)}$ and define in terms of matrices as

$$A^{(p)} = \begin{pmatrix} \text{matrix of all p by p minors} \\ \text{of A ordered lexographically} \end{pmatrix}$$

Since there are $\binom{n}{p}$ ways to select the rows and $\binom{n}{p}$ ways to select the columns in a p by p minor of an n by n matrix we see that $A^{(p)}$ is an $\binom{n}{p}$ by $\binom{n}{p}$ matrix. The following properties of $A^{(p)}$ are well known. See for example [2] or [3].

Theorem 2: Let A and B be given; $A: \mathbb{R}^n \to \mathbb{R}^n$ and $B: \mathbb{R}^n \to \mathbb{R}^n$. Then $A^{(p)}$ and $B^{(p)}$ for $0 \leq p \leq n$ maps $\mathbb{R}^{\binom{n}{p}}$ into $\mathbb{R}^{\binom{n}{p}}$ and

i) $I_n^{(p)} = I_{\binom{n}{p}}$

ii) $(AB)^{(p)} = A^{(p)}B^{(p)}$

iii) $(A^q)^{(p)} = (A^{(p)})^q$ q integer; A^q defined

iv) $(A')^{(p)} = (A^{(p)})'$

We have used two different points of view in defining $A^{[p]}$ and $A^{(p)}$. The construction of $A^{[p]}$ from A was described in terms of linear maps whereas in the definition of $A^{(p)}$ we used matrices exclusively. Alternative approaches are available which give $A^{(p)}$ a geometric meaning in terms of skew symmetric forms of degree p in n variables.

These two constructions are specializations of the tensor product in the following way. If $A: \mathbb{R}^n \to \mathbb{R}^n$ and $B: \mathbb{R}^n \to \mathbb{R}^n$ then we may identify the tensor product of $A\eta$ and $B\gamma$ with $A\eta(B\gamma)'$; i.e.

$$A\eta \otimes B\gamma = A\eta(B\gamma)' = A(\eta\gamma')B'$$

If we consider the linear map of the space of n by n matrices into itself defined by L(Q)=AQB' then $L_A(Q)=AQA'$ when restricted to act on symmetric matrices has $A^{[2]}$ as a matrix representation and when restricted to the complementary space of skew symmetric matrices, it has $A^{(2)}$ as its matrix representation. Thus if we let $\simeq$ indicate "similar to" then we have

$$A \otimes A \simeq A(\cdot)A' \simeq \begin{bmatrix} A^{[2]} & 0 \\ 0 & A^{(2)} \end{bmatrix}$$

One can also see that $A \otimes A \otimes A$ "contains" $A^{[3]}$ and $A^{(3)}$ but there are more than 2 symmetry types for a 3 index tensor so that $A^{[3]} \oplus A^{(3)}$ is only part of $A \otimes A \otimes A$. (Check the dimensionality; $n(n+1)(n+2)/6$ and $n(n-1)(n-2)/6$ does not add up to n^3.)

Now consider a linear differential equation in $\mathbb{R}^n$

$$\dot{x}(t) = A(t)x(t)$$

Observe that

$$x^{[p]}(t+h)=(I+hA(t))^{[p]}x^{[p]}(t)+0(h^2)$$

so that

$$x^{[p]}(t+h)-x^{[p]}(t)=[(I+hA)(t))^{[p]}-I]x^{[p]}(t)+0(h^2)$$

Thus

$$\frac{d}{dt}x^{[p]}(t) = (\lim_{h\to 0}\frac{1}{h}[(I+hA(t))^{[p]}-I])x^{[p]}(t)$$

(Note that the dimensions of the identity matrices in these equations are n and N_n^p respectively.) We define $A_{[p]}$ to be the coefficient matrix in this differential equation.

$$\frac{d}{dt}x^{[p]}(t) = A_{[p]}(t)x^{[p]}(t); \quad p=1,2,3,...$$

Thus the set of all p-degrees forms in $\{x_1,x_2,...,x_n\}$ satisfies a <u>linear</u> differential equation with a coefficient matrix which is easily derived from A.

Starting with a matrix equation

$$\dot{X}(t) = A(t)X(t)$$

we can make an analogous construction using compound matrices (round brackets). The estimate

$$X^{(p)}(t+h) = (I+hA(t))^{(p)}X^{(p)}(t) + 0(h^2)$$

leads to

$$\frac{d}{dt}X^{(p)}(t) = (\lim_{h\to 0}\frac{1}{h}[(I+hA(t))^{(p)}-I])X^{(p)}(t)$$

which we write as

$$\frac{d}{dt}X^{(p)}(t) = A_{(p)}(t)X^{(p)}(t); \quad p=1,2,...,n$$

The special case in which p=n is the basis for well known Able-Jacobi-Liouville formula obtained by integrating the scalar equation

$$\frac{d}{dt}(\det X) = (\operatorname{tr} A(t))\det X(t)$$

Thus we see that $A_{[p]}$ and $A_{(p)}$ are infinitesimal versions of

$A^{[p]}$ and $A^{(p)}$ respectively. As such, they depend linearly on the elements of A. This has some significant implications.

We also have the infinitesimal version of the tensor product reduction given above. It takes the form

$$A(\cdot)+(\cdot)A' \simeq I \otimes A + A \otimes I \simeq \begin{bmatrix} A_{[2]} & 0 \\ 0 & A_{(2)} \end{bmatrix}$$

There are important relationships between A, $A_{[p]}$ and $A_{(p)}$ which are more or less clear from derivation. First of all, if A has all distinct eigenvalues $\{\lambda_i\}$ then the solutions of $\dot{x}(t)=Ax(t)$ consists of a sum of terms of the form $\alpha_i e^{\lambda_i t}$. Thus $x^{[p]}$ consists of products, p at a time, of such terms

$$x^{[p]} = \Sigma\beta_{ij\ldots k} e^{(\lambda_i+\lambda_j+\ldots\lambda_k)t}$$

Thus the eigenvalues of the $\binom{n+p-1}{p}$ by $\binom{n+p-1}{p}$ matrix $A_{[p]}$ are the $\binom{n+p-1}{p}$ sums over distinct (unordered) index sets

$$\lambda_i+\lambda_j+\ldots\lambda_k; \quad p \text{ terms}$$

The same is true for the case where A has eigenvalues of higher multiplicity. Similarly, the eigenvalues of $A_{(p)}$ consist of sums p at a time of the eigenvalues of A but in this case the indices i,j,..,k must all be distinct.

A second fact involves the transition matrix $\Phi_A(t)$ which satisfies

$$\dot{\Phi}(t) = A(t)\Phi(t); \quad \Phi(0) = I$$

By the above construction we see that

$$\Phi_{A_{[p]}}(t) = \Phi_A^{[p]}(t)$$

and

$$\Phi_{A_{(p)}}(t) = \Phi_A^{(p)}(t)$$

(Again, the last of these is the Able-Jacobi-Liouville formula if p=n.)

Finally, if $\{A_i\}$ is a basis for a Lie algebra and if

$$[A_i, A_j] = \sum_{k=1}^{m} \gamma_{ijk} A_k$$

then

$$[A_{i_{[p]}}, A_{j_{[p]}}] = \sum_{i=1}^{m} \gamma_{ijk} A_{k_{[p]}}$$

That is, the $\{A_{i_{[p]}}\}$ form a Lie algebra with the same structural constants. To see this we need to show that

$$[A,B]_{[p]} = [A_{[p]},B_{[p]}]$$

but this can be seen from the approximations

$$\begin{aligned} e^{[A,B]_{[p]}t^2} &= (e^{[A,B]t^2})^{[p]} \\ &\approx e^{A_{[p]}t}e^{B_{[p]}t}e^{-A_{[p]}t}e^{-B_{[p]}t} \\ &\approx e^{[A_{[p]},B_{[p]}]t^2} \end{aligned}$$

where in all cases the approximations are valid up to and including terms of second order in t. Identical formulas hold with [p] replaced by (p).

This circle of ideas is of great importance in the theory of representations of Lie algebras; see [4] or [5]. However in control theory and differential equations there exist many problems where one can use these ideas, and other ideas from representation theory, to simplify calculations and to provide insight. A particular example is the study of the moment equations for stochastic differential equations. See, for example, reference [6].

Exercises

1. Show that

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -k(t) & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

and

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 0 \\ -\sqrt{2}k(t) & -1 & \sqrt{2} \\ 0 & -\sqrt{2}k(t) & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

are an A, $A_{[2]}$ pair.

2. Show that $A^{[p]}$ is orthogonal if A is orthogonal. What about $A^{(p)}$?

3. Describe in full the decomposition of $A \otimes A \otimes A$.

4. Give a definition of $A^{[p]}$ for which $z = Ax$ implies $z^{[p]}=A^{[p]}x^{[p]}$ but which does not require A to be square.

1.3 Matrix Lie Algebras and the Matrix Exponential

In section 1 we saw that the solution of the differential equation

$$\dot{x}(t) = (\sum_{i=1}^{m} u_i(t)A_i)x(t); \quad x(0) = x_o$$

could be expressed for small $|t|$ as

$$x(t) = e^{A_1g_1(t)}e^{A_2g_2(t)}\ldots e^{A_mg_m(t)}x_o$$

provided the A_i form a basis for a Lie algebra. On the strength of the theorem of Frobenius, similar statements can be made for

$$\dot{x}(t) = \sum_{i=1}^{m} u_i(t)f_i[x(t)]; \quad x(0) = x_o$$

provided the set of vectors $\{f_i(\cdot)\}$ are involutive. There is a sort of converse question. If the set $\{A_i\}$ does not form the basis for a Lie algebra to what extent is it necessary to add elements to these sets in order to cover all possibilities? We know already that by adding enough elements to $\{A_i\}$ so as to obtain a basis for a Lie algebra we can be assured of a representation of the above form. However, it might happen that for

$$\dot{x}(t) = u_1(t)A_1x(t)+u_2(t)A_2x(t); \quad x(t) \in \mathbb{R}^n$$

the smallest Lie algebra which contains A_1 and A_2 is of dimension n^2. Are all of the n^2-2 elements which we add in order to get a Lie algebra really necessary?

In 1939 Chow [7] published a generalization of an earlier theorem of Caratheodory proving that if some regularity conditions hold, then along solution curves of

$$\dot{x}(t) = \sum_{i=1}^{m} u_i(t)f_i[x(t)]; \quad x_o = x(0)$$

one can reach the same points as one can along the solution curves of

$$\dot{x}(t) = \sum_{i=1}^{m} u_i(t)f_i[x(t)] + \sum_{i=1}^{\nu} v_i(t)g_i[x(t)]$$

where $g_i(x)$ are obtained as Lie brackets of the f_i, Lie brackets of these Lie brackets, etc. Thus on the basis of this "reachability" theorem of Chow we see that no matter how many elements we must add to get a basis for a Lie algebra, nothing short of the full set will suffice.

We formalize this discussion as follows. Let B denote any subspace of $g\ell(n)$. Let $\{B\}_A$ denote the smallest Lie algebra which contains B. Let C be any subset of $G\ell(n)$ and let $\{C\}_G$ denote the smallest group which contains C.

Theorem 1: With the above definitions

$$\{\exp B\}_G = \{\exp \{B\}_A\}_G$$

Perhaps the most elementary proof of this result appears in [8].

After sufficient insight is built up it is frequently possible to evaluate $\{\exp \{B\}_A\}_G$ by inspection. The insight comes from a handful of special cases and general formulas such as $\exp A_{[p]}$ $(\exp A)^{[p]}$. The notation for the principle special casesis this:

We take the field to be $\mathbb{R}$ and let $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$.

$$g\ell(n) = \{X : X = n \text{ by } n \text{ matrices}\}$$
$$s\ell(n) = \{X : X \in g\ell(n);\ \text{tr } X = 0\}$$
$$so(n) = \{X : X \in g\ell(n);\ X'+X = 0\}$$
$$sp(n) = \{X : X \in g\ell(n);\ X'J+JX = 0\}$$

Matrices satisfying the last condition are often called Hamiltonian because they take the form familiar in Hamiltonian mechanics

$$\begin{bmatrix} A & Q \\ R & -A' \end{bmatrix} \quad ; \quad Q = Q'; \quad R = R'$$

It is very important to keep in mind that $J^2 = -I$ so that $J^{-1} = -J$.

Associated with each of these algebras is a multiplicative group of matrices which are defined in a corresponding way

$$G\ell(n) = \{X : X \text{ is } n \text{ by } n \text{ matrix; } \det X \neq 0\}$$
$$S\ell(n) = \{X : X \in G\ell(n);\ \det X = 1\}$$
$$So(n) = \{X : X \in G\ell(n);\ X'X = I\}$$
$$Sp(n) = \{X : X \in G\ell(n);\ X'JX = J\}$$

These groups are called the general linear group, the special linear group, the special orthogonal group and the symplectic group, respectively.

It is easy to verify that in any of these cases exp X belongs to a particular group if X belongs to the corresponding algebra. This corresponds to the following well known facts

i) exp M is nonsingular for all M
ii) $\det(\exp M) = \exp(\text{tr } M) = 1$ if $\text{tr } M = 0$
iii) exp A is orthogonal if A is skew symmetric since $(e^A)' = e^{A'} = e^{-A} = (e^A)^{-1}$ if $A = -A'$.
iv) exp A is symplectic if A is Hamiltonian since $e^{A'}Je^A = Je^{J'A'J}e^A = J$ if $A'J+JA = 0$.

Notice that the set of n by n symmetric matrices do not form a Lie algebra; alternatively, the nonsingular symmetric matrices do not form a group.

The implication for the study of differential equations is as follows. If X is an n by n matrix which satisfies the equation

$$\dot{X}(t) = A(t)X(t)$$

Then of course the fundamental solution $\Phi_n(t)$ is going to belong to the general linear group. But if A at all points in time belongs to one of the above subalgebras of $g\ell(n)$ then $\Phi_A(t)$ will belong to the corresponding subgroup of $G\ell(n)$. This group-algebra relationship provides qualitative information about the solution without actually solving the equations of motion.

To what extent are the above maps of the algebra into the group actually onto the group? It is well known that a real nonsingular matrix need not have a real logarithm. Thus as far as the real field is concerned, exp does not map $g\ell(n)$ onto $G\ell(n)$. However if the field is either the reals or the complexes, then every matrix sufficiently close to the identity does have a logarithm in the appropriate field and it is easy to see that exp maps a neighborhood of zero in the algebra onto a neighborhood of the identity in the group in a one to one way.

Exercises

1. Consider the set of n by n matrices whose column sums are zero. Show that they form a Lie algebra. If we denote this algebra by L then characterize $\{\exp L\}_G$.

2. Let so(p,q) denote the set of matrices satisfying

$$A'\Sigma(p,q) + \Sigma(p,q)A = 0$$

where $\Sigma(p,q)$ is defined by

$$\Sigma(p,q) = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$$

Show that this set of matrices forms a Lie algebra and show that for all matrices M in $\exp\{so(p,q)\}$ we have

$$\Sigma(p,q) = M'\Sigma(p,q)M$$

These are often called the pseudo orthogonal groups since they preserve the pseudo length $x'\Sigma(p,q)x$.

1.4 Cones and Semigroups

A semigroup of real n by n matrices is simply a subset of the n by n matrices which is closed under matrix multiplication. A cone in a real vector space is a subset closed under addition and multiplication by positive real numbers. Consider a real Lie algebra L in the set of n by n matrices. Let K be a conical subset of L. In general K will not be closed under Lie bracketing but it could be. Let $\{\exp K\}_{SG}$ indicate the smallest semigroup which contains exp K. As we will see, a number of problems in control lead to the question of characterizing $\{\exp K\}_{SG}$ in terms of K. The connection between a Lie algebra and its corresponding Lie group suggests analogous relationships between cones in the algebra and semigroups in the corresponding group. This kind of relationship is illustrated in the following example.

Example: Let K be the cone in $g\ell(n)$ consisting of all n by n matrices A such that A'+A is nonnegative definite. Then $\{\exp K\}_{SG}$ includes all orthogonal matrices since all skew symmetric matrices belong to K. Moreover, all symmetric matrices with eigenvalues greater than or equal to one belong to $\{\exp K\}_{SG}$ by well known properties of the exponential map. Thus by appealing to the fact that any matrix can be written in polar form $M = \theta R$ with θ orthogonal and R positive definite we see that if for all vectors x of unit length $||Mx||^2 = ||\theta Rx||^2 = ||Rx||^2 \geq 1$ then M belongs to $\{\exp K\}_{SG}$. It is easy to see that if $||Mx|| < 1$ for some x of unit length then we can not express M in the required way thus this condition is necessary and sufficient. We conclude that the semigroup of "expansive" matrices is the exponential of the non-negative definite ones. Likewise, the semigroup of (nonsingular) "contractive" matrices is the exponential of the cone of non-positive definite matrices.

This example can be generalized somewhat to give a theorem with broader scope.

Theorem 1: Let K be as above and let L_P be the Lie algebra of matrices satisfying $A'P+PA = 0$ with $P'P = I$. Then $\{\exp K \cap L_P\}_{SG} = \{\exp K\}_{SG} \cap \{\exp L_P\}_G$ i.e. the expansive matrices in $\{\exp L_P\}_G$.

Proof: Given any orthogonal matrix P, the group of matrices satisfying $M'PM = P$ has the property that the polar representations of each element has both its factors in the group. That is, if $M = e^{\Omega}e^R$ with e^{Ω} orthogonal and e^R positive definite and symmetric, then $e^{\Omega'}Pe^{\Omega} = P$, $e^RPe^R = P$. To prove this we note that if $e^Re^{\Omega'}Pe^{\Omega}e^{R\Omega} = P$ then $e^Re^{\Omega'} = Pe^{-R}P'Pe^{-\Omega}P'$. However the term of the right is a polar decomposition since $Pe^{-R}P'$ is symmetric and positive definite and $Pe^{-\Omega}P'$ is orthogonal. Thus by uniqueness of the polar decomposition we see that $e^R = Pe^{-R}P'$ and $e^{\Omega} = Pe^{\Omega}P'$ which shows that each factor belongs to the given group.

Now if M has the polar form $M = e^{\Omega}e^R$ and if M belongs to $\{\exp K\}_{SG} \cap \{\exp L_P\}_G$ then $R \geqslant 0$ and Ω and R belong to L_P. Thus Ω belongs to $L_P \cap K$ and so does R.

Typically the relationship between a cone in the Lie algebra and the semigroup which the exponential maps it into is very difficult to describe. One problem of this type which has been investigated extensively arises in probability theory. Let $x_o \in \mathbb{R}^n$ have nonnegative components which sum to one. Suppose that x(t) evolves in time according to

$$\dot{x}(t) = A(t)x(t); \quad x(0) = x_o$$

If $A(\cdot)$ has the two properties:

(i) the off-diagonal elements of A(t) are nonnegative for all t
(ii) the sums of the columns of A(t) are zero for all t,

then x(t) will have nonnegative components which sum to one for all $t \geqslant 0$. This is equivalent to saying that subject to the above restrictions on $A(\cdot)$ the solution of the matrix equation

$$\dot{X}(t) = A(t)X(t); \quad X(0) = I \qquad (*)$$

is a stochastic matrix; i.e. a matrix with nonnegative entries whose columns sum to 1. The imbedding problem [9] is that of determining which stochastic matrices ϕ can be reached from the identity along solutions of (*) given only that A(t) must satisfy (i) and (ii). Of course the set of matrices which satisfy (i) and (ii) form a cone and the set of reachable matrices form a semigroup. It is not true however that for n > 2 this semigroup consists of all stochastic matrices.

In control applications there is particular interest in the case of cones of the form

$$K = \{X : X = \alpha A + \Sigma\beta_i B_i; \quad \alpha \geq 0; \ \beta_i \text{ unrestricted}\}$$

i.e. cones which are half spaces.The first point to make is that by virtue of theorem 3.1 we may as well assume that the B_i form a basis for a Lie algebra since by adding elements to $\{B_i\}$ to make the basis of the Lie algebra generated by $\{B_i\}$ we do not enlarge the reachable set. Moreover, it is also clear from theorem 3.1 that

$$\{\exp\{A,B_i\}_A\}_G \supseteq \{\exp K\}_{SG} \supseteq \{\exp\{B_i\}_A\}_G$$

It is more or less clear that if e^{At} is periodic then

$$\{\exp\{A,B_i\}_A\}_G = \{\exp K\}_{SG}$$

and Jurdjevic and Sussmann [10] have shown that this is also true if e^{At} is almost periodic.

It is also true that $Ad_A^k B_i$ belongs to the Lie algebra generated by the B_i's then

$$\exp K = e^{\alpha A}\{\exp\{B_i\}_A\}_G$$

For a proof and some generalizations see the thesis of Hirschorn [11].

<u>Exercises</u>

1. Calculate $\{\exp N\}_{SG}$ where N is the cone

$$N = \{X : X = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}; \quad X+X' \leq 0\}$$

2. It is well known that the elements of $\Phi_A(t)$ are nonnegative for all $t \geq 0$ if A(t) itself as elements which are nonnegative off the diagonal -- the diagonals may have any sign. Give an example which shows that $\{\exp K\}_{SG}$ is not the entire semigroup of square matrices with nonnegative entries if K is the cone of A's described above. (Find a matrix with positive entires and negative determinant.)

3. Explore the relationship between #2 and the imbedding problem.

II. INPUT-OUTPUT SYSTEMS

In this chapter we consider input/output systems which can be represented by a pair of equations of the form

$$\dot{X}(t) = (A + \sum_{i=1}^{m} u_i(t)B_i)X(t); \quad y(t) = C(X(t)) \qquad (*)$$

Here X is an n by n matrix as are A and $B_1, B_2, \ldots, B_m$; the map C is subject to certain restrictions to be described later. The differential equation is said to be of the "right invariant type" because a multiplication on the right by a fixed element of $G\ell(n)$ gives an equation

$$\dot{X}(t)M = (A+\sum_{i=1}^{m} u_i(t)B_i)X(t)M$$

which is again of the same form and with the same coefficient matrices. This is to be contrasted with an equation such as

$$\dot{X}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)X(t)+X(t)(D+\sum_{i=1}^{m} u_i(t)E_i)$$

which does not have this invariance property. The basic idea is to understand as well as possible the properties of input-output maps which can be represented by equation (*). We will study controllability, observability and state space isomorphism theorems.

2.1 Controllability

If u_i is an m-dimensional piecewise continuous function of time and if t_i is a nonnegative number, then we given the pairs (u_i,t_i) a semigroup structure by defining

$$(u_1,t_1) \circ (u_2,t_2) = (u_1|u_2,t_1+t_2)$$

whereby $u_1|u_2 = u_3$ we mean

$$u_3(t) = \begin{cases} u_1(t); & 0 \leq t < t_1 \\ u_2(t-t_1); & t_1 \leq t < t_2 \end{cases}$$

This is the concatenation semigroup with due regard for the domain of definition of the functions being concatenated. We denote it by U^m.

Consider the time invariant control system

$$\dot{x}(t) = f[x(t),u(t)] \quad ; \quad x(t) \in \mathbb{R}^n \qquad (**)$$

with f well enough behaved so as to guarantee the existence of a unique solution for each starting point $x_o \in \mathbb{R}^n$ and each $(u,t) \in U^m$. Let T^n be the semigroup of one to one continuous maps of $\mathbb{R}^n$ into $\mathbb{R}^n$ with composition as the semigroup operation. Then the control system (**) defines a homomorphism of U^m into T^n. We denote this homomorphism by ϕ and, by analogy with automata theory, call the image of U^m under ϕ the Myhill semigroup of the system.

The main thing which is special about bilinear systems is that the Myhill semigroup is easily identified with a matrix semigroup. That is, if we have a system in $\mathbb{R}^n$

$$\dot{x}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)x(t)$$

then the matrix equation

$$\dot{X}(t) = (A + \sum_{i=1}^{m} u_i(t)B_i)X(t); \quad X(0) = I$$

describes the relationship between U^m and T^n -- each matrix being associated with an element of T^n in the standard way

$$M \mapsto f(x) = Mx$$

If A is absent in the above equation then it is clear that the Myhill semigroup is actually a group since if $u(\cdot) \in U^m$ steers the system from I to M at time t_1 then $v(\cdot) \in U^m$ and defined by

$$v(t) = -u(t_1 - t)$$

steers the system to M^{-1} at $t = t_1$.

Given an initial state x_o, the set of states reachable from x_o can be identified with the set of points which x_o is mapped into by the various elements of the Myhill semigroup. That is, the Myhill semigroup acts on the state space

$$S : \Sigma \to \Sigma$$

The reachable set from x_o is the "orbit" through x_o defined by this action.

We now give various examples of reachability theorems.

Theorem 1: There exists a control which steers the system

$$\dot{X}(t) = (\sum_{i=1}^{m} u_i(t)B_i)X(t)$$

from X_o to X_1 in time $t_1 > 0$ if and only if $X_1 X_o^{-1}$ belongs to $\{\exp\{B_i^o\}_A\}_G$.

Proof: This is an immediate consequence of Theorem 1.3.1.

It is also easy to see that if A belongs to $\{B_i\}_A$ then the reachable set for

$$\dot{X}(t) = (A + \sum_{i=1}^{m} u_i(t)B_i)X(t)$$

is just the same as it would be if A were absent.

Notice that the reachable set does not depend on t_1 as long as t_1 is positive. If A is absent and if one restricts the controls to be bounded, say $|u_i(t)| \leq 1$ then all points of the above form are reachable after a suitably long time but the time required will depend on the point to be reached.

A second result which we want to use in a moment is this.

Theorem 2: The reachable set at time t for

$$\dot{X}(t) = (\tilde{A} + \sum_{i=1}^{m} u_i(t)\tilde{B}_i)X(t); \quad X(0) = I$$

and

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \qquad \tilde{B}_i = \begin{bmatrix} 0 & B_i \\ 0 & 0 \end{bmatrix}$$

with A square is

$$R(t) = e^{At}\{\exp\{Ad_{\tilde{A}}, \tilde{B}_i\}_A\}_G$$

Here $\{Ad_{\tilde{A}}, \tilde{B}_i\}_A$ indicates the smallest Lie algebra which contains $\{\tilde{B}_i\}_A$ and is closed under the action of $Ad_{\tilde{A}}$.

Proof: See reference [8], Theorem 7.

We can combine theorems 1 and 2 in an obvious way to get the following more general result.

Theorem 3: The reachable set at time t for

$$\dot{X}(t) = \tilde{A}x(t) + \sum_{i=1}^{m} u_i(t)\tilde{B}_i X(t) + \sum_{i=1}^{q} v_i(t)\tilde{C}_i X(t)$$

where

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}; \quad \tilde{B}_i = \begin{bmatrix} 0 & 0 \\ 0 & B_i \end{bmatrix}; \quad \tilde{C}_i = \begin{bmatrix} 0 & C_i \\ 0 & 0 \end{bmatrix}$$

with A and B_i square is

$$R(t) = \exp At\{\exp\{Ad_{\tilde{A}}, \tilde{B}_i, \tilde{C}_i\}_A\}_G$$

Finally, one can get additional results by using a nice lemma of Jurdjevic and Sussmann [10].

Theorem 4: The reachable set for the $\mathbb{R}^n$ system at time t starting from x=0 at t=0 and governed by

$$\dot{x}(t) = (A + \sum_{i=1}^{m} u_i(t)B_i)x(t) + \sum_{i=1}^{p} v_i(t)g_i; \quad x(t) \in \mathbb{R}^n$$

is the vector space generated by $\{L_i^k g_i\}$ where k indicates powers and L_i is a basis for the associative algebra generated by $\{A, B_i\}$.

Proof: To begin we observe that if x_1 is reached at $t=t_1$ starting from x=0 at t=0 using the control (u,v) then the control $(u,\alpha v)$ steers the system to αx_1 at $t=t_1$. Also, we know that if we write the system as

$$\frac{d}{dt}\begin{bmatrix} x(t) \\ 1 \end{bmatrix} = \left(\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} + u_i(t)\begin{bmatrix} B_i & 0 \\ 0 & 0 \end{bmatrix} + v_i\begin{bmatrix} 0 & g_i \\ 0 & 0 \end{bmatrix}\right)\begin{bmatrix} x(t) \\ 1 \end{bmatrix}$$

then the reachable set has a nonempty interior in

$$R = \{\exp\{\tilde{A}, \tilde{B}, \tilde{G}\}_{LA}\}_G \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

where

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}; \quad \tilde{B}_i = \begin{bmatrix} B_i & 0 \\ 0 & 0 \end{bmatrix}; \quad \tilde{G}_i = \begin{bmatrix} 0 & g_i \\ 0 & 0 \end{bmatrix}$$

There exists a nonzero control of the form (0,v) which steers the system back to zero at time $t=t_1$ from 0 at t=0 -- use u=0 and invoke standard linear theory. According to lemma 6.1 of [10] we obtain on taking perturbations about this control an open set in R containing 0. Using the cone property mentioned in the first sentence we see that the reachable set is a vector space. Lie algebras tell us which one.

A particular problem in controllability theory which has received a good deal of attention is

$$\dot{x}(t) = Ax(t)+u(t)b\langle c,x(t)\rangle \quad ; \quad x(t) \in \mathbb{R}^n$$

where $u(\cdot)$ is a scalar, and b is a column vector. Of course the linear system

$$\dot{x}(t) = Ax(t)+bv(t)$$

is controllable in $\mathbb{R}^n$ if and only if $(b,Ab,..,A^{n-1}b)$ is of full rank. If the linear system is controllable it might be supposed that the bilinear one is also controllable since if v is a control which drives the state of the linear system from x_o to x_1 then the control

$$u(t) = v(t)/\langle c,x(t)\rangle$$

drives the bilinear system from x_o to x_1. This argument has the obvious fallacy that $\langle c,x(t)\rangle$ might vanish along the trajectory leaving u(t) undefined. In particular, if $x(0) = 0$ then of course x vanishes identically for all future time. Thus the most one could hope for is that any nonzero state could be steered to any nonzero state. It turns out that this is too much to hope for also. A simple pair of examples which illustrate that no amount of work can salvage this argument and which at the same time suggest the nature of the problem are these.

Consider the system

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} + u(t)\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}$$

which has the form

$$\dot{x}(t) = Ax(t) + u(t)b\langle c,x(t)\rangle$$

with [A,b,c] a minimal realization of $s/(s^2-1)$. However for any given x_o there exists x_1 such that x_1 is not reachable from x_o, because regardless of k, the off-diagonal elements of $(A+k(t)bc')$ are always positive so that $\phi(t,t_o)$, the transition matrix, has all entries nonnegative for $t > t_o$. Thus if x(0) has nonnegative entries for all $t > 0$. This argument shows that the system is not controllable.

Consider the system

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} + k(t)\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}$$

which has the form $\dot{x}(t) = Ax(t)+k(t)bc'x(t)$ with [A,b,c] a minimal realization of $s/(s^2+1)$. In this case we see that the system is controllable on $R^2-\{0\}$. (See reference [12] for details.)

<u>Exercises</u>

1. Show that the Myhill semigroup for the linear system

$$\dot{x}(t) = Ax(t)+bu(t); \quad x(t) \in \mathbb{R}^n$$

can be identified with the multiplicative matrix semigroup

$$S = \{X : X = \begin{bmatrix} e^{At} & x \\ 0 & 1 \end{bmatrix} ; \quad t \geq 0; \ x \in \mathrm{span}(b, Ab, \ldots A^{n-1}b)\}$$

2. Consider a bilinear system

$$\dot{x}(t) = Ax(t) + u(t)Bx(t)$$

on $\mathbb{R}^n - \{0\}$. Is it true that if there exists any state x_o such that all points in $\mathbb{R}^n - \{0\}$ are reachable from x_o then all states have this property?

3. Consider the linear system

$$\dot{X}(t) = A_\ell X(t) + X(t)A_r + \sum_{i=1}^{m} u_i(t)B_i$$

Here $X(t)$ is an n by q matrix and A_ℓ and A_r are n by n and q by q respectively; the B_i are n by q. Show that the Myhill semigroup equation can be identified with

$$\frac{d}{dt}\begin{bmatrix} S_1(t) & S_3(t) \\ 0 & S_2(t) \end{bmatrix} = \left(\begin{bmatrix} A_\ell & 0 \\ 0 & A_r \end{bmatrix} + \sum_{i=1}^{m} u_i(t)\begin{bmatrix} 0 & B_i \\ 0 & 0 \end{bmatrix}\right)\begin{bmatrix} S_1(t) & S_3(t) \\ 0 & S_2(t) \end{bmatrix}$$

Show that the reachable set at time t for the Myhill equation is

$$\exp At \cdot \exp\{Ad_A, \tilde{B}_i\}$$

2.2 Observability

We now consider systems with an output

$$\dot{X}(t) = (A(t) + \sum_{i=1}^{m} u_i(t)B_i(t))X(t); \quad y(t) = C(X(t)); \quad X(t) \in G\ell(n)$$

The exact nature of the output map is not essential. We give the output space no structure -- it is just a set. The critical assumption is that there should exist subgroups H_ℓ and H_r of $G\ell(n)$ such that $C(X_1) = C(X_2)$ if and only if

$$H_1 X_1 H_2 = X_2$$

for some H_1 in H_ℓ and some H_2 in H_r. Under this assumption $C(X)$ identifies X to within a multiplication on the left by an element of H_ℓ and a multiplication on the right with an element of H_r. We call systems of this form homogeneous.

In such a set up, the observation of y, even over a period of time, can at most determine X to within a right multiplication by an element of H_r. Thus we might as well regard the system as evolving on the coset space $G\ell(n)/H_r$. Whether or not the observation of y and the knowledge of u over the interval $[0,\infty)$ serves to identify uniquely an element of X/H_r as a starting state is then subject to investigation.

<u>Theorem 1</u>: Consider the above system with H_r and H_ℓ given. Let R denote the set of X's reachable from I. Suppose that R is a group.

Then two points X_1H_r and X_2H_r in $G\ell(n)/H_r$ give rise to the same input/output map if and only if for each R_1 in R there exists $H_1(R)$ in H_ℓ such that

$$R^{-1}H_1(R)RX_1H_r = X_2H_r$$

If we denote by P the subgroup

$$P = \{X : R^{-1}XR \in H_\ell; \quad \forall R \in S\}$$

then any two elements of the form X_1H_r and $P_1X_1H_r$ with P_1 in P are not distinguishable.

Proof: If X_1H_r and X_2H_r are to be indistinguishable as starting states we must have

$$H_\ell R_i X_1 H_r = H_\ell R_i X_2 H_r$$

for all R_i in R. Since H_ℓ and H_r are groups and since R is a subgroup of $G\ell(n)$, the above condition is equivalent to asking that for each R_i in R there exist $H_1(R)$ in H_ℓ such that

$$R_i^{-1}H_1(R_i)R_iX_1H_r = X_2H_r$$

The remainder of the conclusions are clear.

Exercises

1. Assuming that the evolution equations are of the form

$$\dot{x}(t) = Ax(t) + \sum_{i=1}^{m} u_i(t)B_ix(t); \quad y(t) = H_\ell x(t)H_r$$

with

$$H_\ell = \{\exp\{C_i\}_A\}_G; \quad H_r = \{\exp\{D_i\}_A\}_G$$

give an observability condition in terms of Lie algebras. (See ref. [8] for some results along this line.)

2. Apply the results of problem 1 to the bilinear problem

$$\dot{x}(t) = Ax(t) + \sum_{i=1}^{m} u_i(t)B_ix(t); \quad y(t) = c[x(t)]$$

by identifying $\mathbb{R}^n$ with the n dimensional affine group modulo $G\ell(n)$.

2.3 Isomorphic Systems

The two scalar realizations

$$\dot{x}(t) = x(t)+u(t)x(t); \quad y(t) = x^3(t); \quad x(0) = 1$$

and

$$\dot{z}(t) = 3z(t)+3u(t)z(t); \quad y(t) = z(t); \quad z(0) = 1$$

realize the same input-output map. They are each controllable on $(0,\infty)$ and any two reachable states are distinguishable. They are related by the automorphism of the multiplicative group $(0,\infty)$ defined by

$$z = x^3$$

Thus despite the apparent differences between these two realizations they are closely related. The following theorem describes a general result of this type.

Theorem 1: Consider the two homogeneous realizations of the same input-output map

$$\dot{X}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)X(t); \quad y(t) = c[X(t)]$$

$$\dot{Z}(t) = (F+\sum_{i=1}^{m} u_i(t)G_i)Z(t); \quad y(t) = h[Z(t)]$$

which evolve in $G\ell(n_1)$ and $G\ell(n_2)$ respectively and which have reachable sets from the identity, **R and** $\hat{R}$, which are groups. Suppose H_ℓ, H_r and $\hat{H}_\ell$, $\hat{H}_r$ are given subgroups of $G\ell(n_1)$ and $G\ell(n_2)$ respectively such that c and h are one to one on $H_\ell R H_r$ and $\hat{H}_\ell \hat{R} \hat{H}_r$ and such that the systems are observable on RH_r and $\hat{R}\hat{H}_r$. Finally, suppose that there is no normal subgroup of R which has a non-trivial intersection with $R \cap H_r$ and the same for $\hat{R}$ and $\hat{H}_r$. Then there exists an isomorphism $\phi : R \to \hat{R}$ such that

$$\phi(e^{At}) = e^{Ft}; \quad \phi(e^{B_i t}) = e^{G_i t}$$

Proof: Suppose that there exists a control (u,T) in U^m which takes the first system from I to $D_1 \neq I$ and takes the second system from I to I. Let D denote the set of all such points. By virtue of the observability hypothesis we see that D is a subset of H_r and, in fact, a subgroup of H_r. Moreover it is easily seen to be a normal subgroup of R and hence of $R \cap H_r$. By hypothesis D is trivial. This implies that there is a one to one correspondence between points in $R \cap H_r$ and $\hat{R} \cap \hat{H}_r$ which is, in fact, a homomorphism.

We see that R and $\hat{R}$ are both homomorphic images of U^m. If a pre-image of R in U^m is in U_R then what is the image under the action of the second system of U_R? It is clearly $\hat{R}$ or else a subgroup of $\hat{R}$. If it is a subgroup then the subgroup must contain $R \cap \hat{H}_r$ but there is a one to one and onto correspondence between $\hat{R}/R \cap H_r$ and $\hat{R}/\hat{R} \cap \hat{H}_r$ and an isomorphism between $R \cap H_r$ and $\hat{R} \cap \hat{H}_r$ Using the properties of the system maps we see that the above map must be onto R and thus it establishes an isomorphism. The remaining claims then follow.

Exercises

1. Develop the Lie algebra analog of Theorem 1.

2. Apply the above results to bilinear systems of the form

$$\dot{x}(t) = Ax(t)+\sum_{i=1}^{m} u_i(t)B_i x(t); \quad y(t) = cx(t); \; x(0) = x_o$$

See P. d'Alessandro, A. Isidori and A. Ruberti [13] and Brockett [14].

III. OPTIMAL CONTROL

This chapter is quite brief due to the absence in the literature of results relating specifically to the Lie group case. We discuss only two problem areas -- the question of existence of optimal controls in the bang bang case and questions centering around minimum "energy" transfer.

3.1 Bang-Bang Theorems

It is well known that under very weak assumptions on the matrices $A(\cdot)$ and $B(\cdot)$ the linear system

$$\dot{x}(t) = A(t)x(t)+B(t)u(t);\; x(0) = \text{given}$$

with controls constrained by

$$|u_i(t)| = 1$$

has a set of reachable points at any time $t_1 > 0$ which is the same as the set of points reachable with the constraint relaxed to

$$|u_i(t)| \leq 1$$

This is called a "bang-bang theorem" because the controls u_i need only take on their extreme values and not intermediate ones. Some generalizations of this have been investigated by Krenner [15] and Sussmann [16]. We examine only an easy case here.

<u>Theorem 1</u>: Let X satisfy the differential equation in $G\ell(n)$

$$\dot{X}(t) = AX(t) + \Big(\sum_{i=1}^{m} B_i u_i(t)\Big)X(t)$$

Then if $[\mathrm{Ad}_A^k(B_i),B_j]$ is zero for all i and j and $k=0,1,\ldots n^2-1$ then the set of states reachable at time t for $|u_i(t)| = 1$ is the same as the set reachable for $|u_i(t)| \leq 1$.

<u>Proof</u>: In view of the commutativity condition we can express the solution of the given equation as

$$X(t) = e^{At} e^{\int_0^t \sum_{i=1}^{m} e^{-A\sigma}B_i e^{A\sigma} u_i(\sigma)\,d\sigma} X(0)$$

See [8] Theorem 7 for details. Now since the bang-bang theorem is valid for the linear system

$$\dot{F}(t) = \sum_{i=1}^{m} e^{-At}B_i e^{At} u_i(t)$$

and since $X(t) = e^{At}e^{F(t)}$ we see that it holds for the systems defined here as well.

<u>Exercises</u>

1. The solution of the scalar differential equation

$$\dot{x}(t) = u(t)x(t)+v(t)$$

is

$$x(t) = e^{\int_0^t u(\sigma)\,d\sigma} x(0) + \int_0^t e^{\int_\sigma^t u(\rho)\,d\rho} v(\sigma)\,d\sigma$$

Is the bang-bang theorem valid if we regard u and v as controls?

2. Is the bang-bang theorem valid for the pair of scalar equations

$$\dot{z}(t) = u(t)z(t)$$

$$\dot{x}(t) = (u(t)+v(t))x(t)$$

3. Show that the bang-bang theorem is valid for

$$\dot{x}(t) = u(t)x(t)$$

$$\dot{y}(t) = -y(t)+u(t)$$

Generalize this result.

3.2 Least Squares Theory

Under the assumption used in the previous section we can develop a satisfactory theory for minimizing

$$\eta = \int_0^t \sum_{i=1}^m u_i^2(t)\,dt$$

subject to the constraint that the system

$$\dot{X}(t) = (A + \sum_{i=1}^m u_i(t)B_i)X(t) \qquad (*)$$

should be transferred from the state X_o at t=0 to the state X_1 at $t=t_1$.

Theorem 1: Let X(t) satisfy the $G\ell(n)$ equation (*). Suppose that $[Ad_A^k B_i, B_j] = 0$ for all i and j and $k=0,1,2,\ldots n^2-1$. Suppose that

$$X_1X_0^{-1} \in e^{At_1}\{\exp\{Ad_A, B_i\}_A\}_G$$

Then there exists a control $u(\cdot)$ which steers the system from X_o at t=0 to X_1 at $t=t_1$ and minimizes η. This control is the same as the control which steers the linear system

$$\dot{F}(t) = \sum_{i=1}^m e^{-At}B_ie^{At}u_i(t)$$

from 0 at t=0 to $\ell n(e^{-At_1}X_1X_o^{-1})$ at $t=t_1$ and minimizes η where ℓn denotes the real solution of

$$e^M = e^{-At_1}X_1X_o^{-1}$$

which results in the smallest value of η. The optimal control is of the form

$$u_i(t) = tr(M_ie^{-At}B_ie^{At})$$

for some constant matrices M_i.

Proof: As in the proof of the bang-bang theorem we see that

$$X(t) = e^{At}e^{F(t)}$$

where F(t) satisfies

$$\dot{F}(t) = \sum_{i=1}^{m} e^{-At}B_i e^{At}u_i(t)$$

From this point on everything follows from standard linear theory. See [17], section 22.

Exercises

1. Consider the system

$$\dot{x}(t) = x(t)+u(t)$$
$$\dot{y}(t) = u(t)y(t)$$

Suppose we want to steer this system from (α,β) to (γ,δ) in t_1 units of time and to minimize

$$\eta = \int_0^{t_1} u^2(t)dt$$

If δ/β is positive this transfer is possible and the $u(\cdot)$ which achieves the optimal is of the form ae^t+b. Generalize Theorem 1 in such a way as to capture this example.

2. If B_1 and B_2 commute, describe the solutions of

$$\prod_{i=1}^{\nu} (e^{B_1 u_i} e^{B_2 v_i}) = N$$

IV. STOCHASTIC DIFFERENTIAL EQUATIONS

Stochastic processes on spheres has been of interest in physics for some time. Debye [18] in his book on statistical mechanics gives one application of S^2 stochastic processes. Nuclear magnetic resonance phenomena account for some more recent interest in diffusions on S^2. See Chapter 15 of the recent text [19]. The French mathematical physicist Perin wrote a classical paper [20] on diffusion on SO(3). Recent interest in physics regarding models of the type under study here is discussed in Fox [21]. Transmission of electromagnetic waves through random media leads to stochastic processes on the symplectic group -- distance playing the role usually assumed by time. Tutubalin [22] can be consulted for recent results and references. Carrier [23] has examined an equation of this general type in connection with a gravity wave propagation problem. One can think of this study as a stochastic process on the two dimensional symplectic group. An engineering problem for which the theory is potentially interesting is the randomly switched electrical circuit.

4.1 Bilinear Stochastic Equations

In this paper all stochastic differential equations are to be interpreted in the Ito sense. All Wiener processes are of unity variance and Wiener processes with distinct indices are assumed to be uncorrelated. The reader is encouraged to study Clark [24] for more details on stochastic calculus.

Under what circumstances does the Ito equation

$$dx(t) = Ax(t)dt + \sum_{i=1}^{m} dw_i(t)B_i x(t) \qquad (*)$$

evolve on the manifold defined by $x'Qx$ = constant? If we expand to second order keeping in mind that $dw_i dw_j = \frac{1}{2}\delta_{ij}dt$ we get

$$dx'Qx = x'(A'Q+QA)xdt + \sum_{i=1}^{m} x'(B_i'Q + QB_i)xdw + \frac{1}{2}\sum_{i=1}^{m} x'B_i'QB_i x\ dt$$

Thus in order for the derivative of $x'Qx$ to vanish we require

$$A'Q + QA + \sum_{i=1}^{m} B_i'QB_i = 0$$

and also we require

$$B_i'Q + QB_i = 0$$

We see that the drift term A needs to be "corrected" by a term coming from the white noise. For example, if we want equation (*) to evolve on a sphere then A is not skew symmetric as it would be in the deterministic case but rather it has a correction term whose size depends on the B_i. On the other hand, the B_i must be skew symmetric.

In order to evolve on the symplectic group it is a skew symmetric form which must be preserved. Repeating the above with Hamiltonian matrices gives rise to the conditions that B_i and $A + \frac{1}{2}\sum_i B_i^2$ should be Hamiltonian.

<u>Exercises</u>

1. Show that the Ito equation

$$\begin{bmatrix} dx_1 & dx_2 \\ dx_3 & dx_4 \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} dt + \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}\begin{bmatrix} dw_1 & dw_2 \\ dw_3 & -dw_1 \end{bmatrix}$$

evolves on the special linear group $S\ell(2)$ if suitable restrictions are placed on α, β, γ, δ.

2. Generalize the previous problem to $S\ell(n)$.

4.2 The Moment Equations

Associated with the stochastic equation

$$dx(t) = Ax(t)dt + \sum_{i=1}^{m} B_i x(t)dw_i(t) \qquad (*)$$

is a family of higher order equations analogous to those given in section 1.2. These are the equations for $x^{[p]}$. In order to display their form it is necessary to work out section 1.2 using the Ito calculus. As an alternative, suggested to me by Martin Clark, one can convert (*) into an analogous Stratonovich equation, use the ordinary calculus to get the $x^{[p]}$ equation, and then convert back to the Ito form. This idea is particularly attractive in the present setup since we have the deterministic results already.

The Stratonovich analog of (*) is simply

$$\dot{d}x(t) = (A - \frac{1}{2}\sum_{i=1}^{m} B_i^2)x(t)dt + \sum_{i=1}^{m} B_i x(t)\dot{d}w_i(t)$$

where $\dot{d}$ indicates Stratonovich differentials. Applying ordinary calculus we get

$$\dot{d}x^{[p]}(t) = (A - \frac{1}{2}\sum_{i=1}^{m} B_i^2)_{[p]}x^{[p]}dt + \sum_{i=1}^{m} B_{i[p]}x^{[p]}(t)\dot{d}w_i(t)$$

Now if we want to convert this back to an Ito form we must correct the drift term to get

$$dx^{[p]}(t) = [(A - \frac{1}{2}\sum_{i=1}^{m} B_i^2)_{[p]} + \sum_{i=1}^{m}(B_{i[p]})^2]x^{[p]}(t)dt + \sum_{i=1}^{m} B_{i[p]}x^{[p]}(t)dw_i(t)$$

We can easily take expectations to get the moment equation

$$\frac{d}{dt}(\mathscr{E}x^{[p]}(t)) = [(A - \frac{1}{2}\sum_{i=1}^{m} B_i^2)_{[p]} + \sum_{i=1}^{m}(B_{i[p]})^2]\,\mathscr{E}x^{[p]}(t)$$

Notice that the apparently more general equation

$$dx(t) = Ax(t)dt + \sum_{i=1}^{m} B_i x(t)dw_i(t) + \sum_{i=1}^{m} e_i dw_2(t) \qquad (**)$$

is covered by these equations as well. To see this we let

$$\tilde{x} = \begin{bmatrix} x \\ 1 \end{bmatrix}$$

then x satisfies an equation of the form

$$d\tilde{x}(t) = \tilde{A}\tilde{x}(t)dt + \sum_{i=1}^{m}(\tilde{B}_i + \tilde{C}_i)\tilde{x}(t)dw_i(t)$$

There are many papers in the literature which analyze the stability of of these equations under various assumptions -- particular emphasis being placed on the case p=2. See, e.g. [25]. In reference [6] it is shown that under a suitable hypothesis all the moment equations are stable.

<u>Exercises</u>

1. Show that in the scalar case the moment equations for

$$dx(t) = \alpha(t)x(t)dt + \beta(t)x(t)dw(t)$$

are
$$\frac{d}{dt}\mathscr{E}x^p(t) = [p(\alpha(t) - \frac{1}{2}\beta^2(t)) + \frac{1}{2}p^2\beta^2(t)]\,\mathscr{E}x^p(t)$$

Notice that if α and $\beta \neq 0$ are constant then it can never happen that all moment equations are stable.

2. A problem of interest in geophysics leads to the stochastic equation

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} = \begin{bmatrix} 0 & dt \\ -dt+\varepsilon dw(t) & 0 \end{bmatrix} \begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} ; \quad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Show that the autocorrelation is, for small ϵ, approximated by

$$\mathscr{E} x_1(t)x_1(\tau) \simeq e^{(\varepsilon^2/4)(t+\tau)} e^{-(\varepsilon^2/4)|t-\tau|} \cos(t-\tau)$$

(See Carrier [23]).

4.3 Fokker-Planck Equations

Associated with the Ito equation

$$dx(t) = Ax(t)dt + \sum_{i=1}^{m} dw_i B_i x(t)$$

is the formal Fokker-Plank equation

$$\frac{\partial \rho}{\partial t} - \frac{1}{2} tr(\sum_{i=1}^{m} B_i xx' B_i', \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j})\rho - \nabla_x \rho Ax = 0$$

However, if x evolves on a manifold then this equation will not be especially useful unless the redundant variables are eliminated. In order to carry out this reduction it is necessary to coordinatize the manifold in some natural way. This coordinatization necessarily proceeds in a case by case way. To illustrate we work out four cases on the two-sphere S^2.

Consider the stochastic equations (Compare with McKean [26] who considers case b, case a being classical.)

$$\begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} = \begin{bmatrix} -dt & -dw_3 & dw_2 \\ dw_3 & -dt & -dw_1 \\ -dw_2 & dw_1 & -dt \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (a)$$

$$\begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} dt & 0 & dw_2 \\ 0 & -\frac{1}{2} dt & -dw_1 \\ -dw_2 & dw_1 & -dt \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (b)$$

$$\begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} dt & -dt & dw_2 \\ +dt & -\frac{1}{2} dt & -dw_1 \\ -dw_2 & dw_1 & -dt \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (c)$$

$$\begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} = \begin{bmatrix} 0 & -dt & 0 \\ dt & -\frac{1}{2} dt & -dw_1 \\ 0 & dw_1 & -\frac{1}{2} dt \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (d)$$

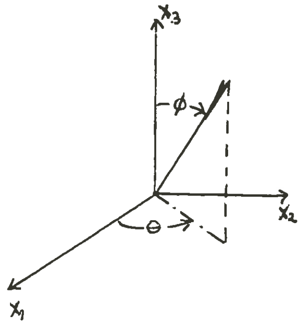

Figure 3:
Spherical Coordinates

We introduce polar coordinates according to figure 3. The Fokker-Planck equations corresponding to the above cases are then

$$\left[\frac{\partial}{\partial t} - \frac{1}{2}\left(\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}\sin\phi\frac{\partial}{\partial\phi} + \frac{1}{\sin^2\phi}\frac{\partial^2}{\partial\theta^2}\right)\right]\rho(t,\phi,\theta) = 0 \qquad \text{(a)}$$

$$\left[\frac{\partial}{\partial t} - \frac{1}{2}\left(\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}\sin\phi\frac{\partial}{\partial\phi} + \frac{1}{\tan^2\phi}\frac{\partial^2}{\partial\theta^2}\right)\right]\rho(t,\phi,\theta) = 0 \qquad \text{(b)}$$

$$\left[\frac{\partial}{\partial t} - \frac{1}{2}\left(\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}\sin\phi\frac{\partial}{\partial\phi} + \frac{1}{\tan^2\phi}\frac{\partial^2}{\partial\theta^2}\right) + \frac{\partial}{\partial\theta}\right]\rho(t,\phi,\theta) = 0 \qquad \text{(c)}$$

$$\left[\frac{\partial}{\partial t} - \frac{1}{2}\left(\sin\theta\frac{\partial}{\partial\phi} + \cot\phi\cos\theta\frac{\partial}{\partial\theta}\right)^2 + \frac{\partial}{\partial\theta}\right]\rho(t,\phi,\theta) = 0 \qquad \text{(d)}$$

The idea behind the derivation of these equations is that each of the three generators

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

can be associated with a first order partial differential operator which describes the effect of a drift around the corresponding axis of rotation and also with a second order partial differential operator which describes the effect of a diffusion around the corresponding axis of rotation. The derivation of these operators is an exercise in differential geometry, however the following insight is useful.

On a manifold with a Riemannian metric $(g_{ij}(x))$, the Laplace-Beltrami operator [27]

$$\nabla^2 = \frac{1}{\sqrt{\det(g_{ij}(x))}}\frac{\partial}{\partial x_i}(g_{ij}(x))^{-1}\sqrt{\det(g_{ij}(x))}\frac{\partial}{\partial x_j}$$

serves as the Laplacian, in that the basic heat equation, assuming constant conductivity, is

$$\left(\frac{\partial}{\partial t} - \frac{1}{2}\nabla^2\right)\phi(t,x) = 0$$

On S^2, in terms of the given coordinates, the usual metric is

$$(ds)^2 = [d\phi, \ d\theta]\begin{bmatrix} 1 & 0 \\ 0 & \sin^2\phi \end{bmatrix}\begin{bmatrix} d\phi \\ d\theta \end{bmatrix}$$

one sees easily that case a above corresponds to the heat equation.

As for case b, it is obtained from case a by removing one of the generators -- the one which corresponds to a diffusion about the x_3-axis. This is equivalent to subtracting $\frac{1}{2}(\partial^2/\partial\theta^2)$ from the operator appearing in case a.

Case c is obtained in an analogous way. We must add a drift term to the operator appearing in b corresponding to a rotation about the x_3-axis. Thus we add a $(\partial/\partial\theta)$ term to the operator appearing in b.

Case d is the most degenerate of all in that there is now only diffusion about one axis. There is a $(\partial/\partial\theta)$ drift term as in case c together with the operator which corresponds to diffusion about the x_1-axis.

It is of some interest to note that all these operators are studied in quantum theory. See Rose [28], appendix A.

Exercises

1. Consider the stochastic equation

$$\begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} \quad \begin{bmatrix} -\frac{1}{2}\,dt & -dw & 0 \\ dw & -\frac{1}{2}\,dt & dt \\ 0 & dt & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \; ; \; x_1^2(0)+x_2^2(0)-x_3^2(0)=1$$

Show that it evolves on the manifold defined by $x_1^2+x_2^2-x_3^2=1$. Introduce coordinates in this manifold and work out the Fokker-Planck equation. Is there a limiting distribution?

2. Show that the moment equations associated with each of the four cases analyzed here are stable. (see [26])

4.4 Calculation of Diffusion Times

We continue with the analysis of the four cases of diffusions on spheres, now with a view toward determining, if possible, a complete solution to the Fokker-Plank equation. In cases where that proves too difficult we look for some measure of the relaxation time of the process.

To begin with, the standard S^2 diffusion (case a above) leads to the Fokker-Plank equation

$$\frac{\partial\rho(t,x)}{\partial t} - \frac{1}{2}\,\nabla^2\rho(t,x) = 0$$

Where ∇^2 is the usual Laplacian on the sphere. It is, of course, well known that the eigenvalues of the Laplacian on the sphere are $n(n+1)$, $n=0,1,2,\ldots$ with the nth being of multiplicity $2n+1$. Thus the general solution of the above equation starting from the singular distribution concentrated at $\theta = \phi = 0$ is

$$\rho(t,\theta,\phi) = \sum_{n=1}^{\infty}\sum_{k=-n}^{n} P_{nk}(\cos\phi)e^{ik\theta}e^{-n(n+1)t}$$

where P_{nk} are the spherical harmonics. We also see that the eigenvalues are a measure of the speed with which the density approaches steady state.

On the basis of this Green's function one can, of course, express the general solution of the Fokker-Plank equation in terms of its initial value. Thus we have, in terms of the spherical harmonics, a complete solution to the Fokker-Planck equation. This is classical.

On the other hand, it is possible to be almost as explicit in the other cases as well. This comes about because the 2n+1 equations for the coefficients of the spherical harmonics of the form $P_{nk}(\cos\ )e^{ik\theta}$ $k=0,\pm 1,\ldots\pm n$ are decoupled from those corresponding to $P_{n'k}(\cos\ )e^{ik\theta}$ for $n\neq n'$. Thus the solution of the Fokker-Planck equation reduces to a sequence of linear differential equations; the nth entry in the sequence being a coupled set of 2n+1 equations. It happens, however, that there is a simple connection between the moment equations of section 4.2 and the equations for the coefficients of the spherical harmonics. We describe this for the S^2 situation but similar results hold on spheres of any dimension.

For an S^2 equation x is a 3-vector and $x^{[p]}$ is of dimension $(p+1)(p+2)/2$. The equation for $x^{[p]}$ includes all linearly independent p-forms in x; thus it includes $(p-1)(p)/2$ terms of the form

$$(x_1^2+x_2^2+x_3^2)\,x^{[p-2]}$$

Hence we can partition $x^{[p]}$ into two parts of dimension $(p-1)p/2$ and $(p+1)(p+2)/2 - (p-1)p/2 = 2p+1$, respectively according to whether the components have a factor of $x_1^2+x_2^2+x_3^2$ or not. Now of course $x_1^2+x_2^2+x_3^2 = 1$ so that the components which do contain this factor can be thought of as moment equations of a lower order and hence they evolve independently of the second part of the equation. On the other hand, the 2p+1 components which do not contain $x_1^2+x_2^2+x_3^2$ as a factor evolve independently as well. Collecting these facts we see that the moment equations have the structure

$$\frac{d}{dt}\mathscr{E}x^{[p]}(t)=\begin{bmatrix} A_\delta & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & A_{\delta+2} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \tilde{A}_{p-2} & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \tilde{A}_p \end{bmatrix}\mathscr{E}x^{[p]}(t)$$

where δ is zero or one depending on whether p is even or odd. The dimension of A_p is (2p+1) by (2p+1) and the coefficients of the spherical harmonics of type P_{nk}, n fixed, $k=0,\pm 1,\pm 2,\ldots\pm n$ are governed by the differential equation

$$\dot{y}(t) = \tilde{A}_p y(t)$$

Thus the spectrum of the operators

$$(A - \frac{1}{2}\sum_{i=1}^{m} B_i^2)_{[p]} + \sum_{i=1}^{m}(B_{i[p]})^2$$

which were derived in section 4.2, governs the relaxation time of the process. In case a above we have already commented that the spectrum is $\frac{1}{2}(n(n+1))$ with the nth term being of multiplicity 2n+1. In case b there is less diffusion and one would expect the relaxation to be slower. This is the case; a calculation shows that the first few entries of the spectrum compares with case as follows.

$$\frac{1}{2}\begin{cases} 0, & 2, \; 2, \; 2 \quad 6, \; 6, \; 6, \; 6 \quad 6 \;\dots\; \text{case a} \\ \underbrace{0,}_{I} & \underbrace{1, \; 1, \; 2}_{II} \quad \underbrace{2, \; 2, \; 5, \; 5,}_{III} \; 6 \;\dots\; \text{case b} \end{cases}$$

Finally, we remark that examples b, c, and d are specific cases of the hypoelliptic operators of Hormander [29].

Exercises

1. Consider the linear stochastic equation

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} = \begin{bmatrix} -1/2 & 1 \\ -1 & -1/2 \end{bmatrix} dt + \begin{bmatrix} dw_1(t) \\ dw_2(t) \end{bmatrix}; \quad x(0) = 0$$

as an approximation to the first two components of the S^2 equation

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \\ dx_3(t) \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}dt & dt & dw_1 \\ -dt & -\frac{1}{2}dt & dw_2 \\ -dw_1 & -dw_2 & -dt \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}; \quad x(0) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Compute the second moment in each case and compare.

2. Consider the stochastic equation on S^2 defined by

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \\ dx_3(t) \end{bmatrix} = \begin{bmatrix} -dt/2 & dw_1 & 0 \\ -dw_1 & -(1+\rho)dt/2 & \rho dw_2 \\ 0 & -\rho dw_2 & -\frac{\rho}{2} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$$

Find the first few eigenvalues of corresponding Fokker-Planck operator as a function of ρ.

V. STABILITY THEORY

In the study of ordinary differential equations on Lie groups both linear and nonlinear problems are of interest, however in these notes we discuss linear problems only. Of course the most common stability problems encountered in control concern the general linear group. However in the study of specific applications other groups may occur. For example, in the case of problems arising in classical mechanics the symplectic group plays a major role. Moreover since tensoring will typically transform a system evolving in $G\ell(n)$ into one which evolves on some subgroup of $G\ell(q)$ is desirable to take a general point of view.

5.1 Stability of the $x^{[p]}$ Equations

The following theorem is an obvious consequence of the calculations in section 1.2.

Theorem 1: The null solution of the system

$$\dot{x}(t) = A(t)x(t)$$

is stable (asymptotically stable) if and only if the null solution of the equation

$$\dot{y}(t) = A_{[p]}(t)y(t)$$

is stable (asymptotically stable). Moreover if all solutions of the first equation are bounded by $|x(t)| \leq Me^{-\lambda t}$ then all solutions of the second are bounded by $|y(t)| \leq M_1e^{-p\lambda t}$.

When combined with standard estimates this theorem can give very precise information about high order systems which are either in the form of $\dot{y}(t) = A_{[p]}(t)y(t)$ or else in the form

$$\dot{y}(t) = A_{[p]}(t)y(t) + D(t)y(t)$$

with $D(t)$ small in some sense.

Example: We know from Liapunov [see e.g. [30]] that all solutions of the $Sp(2)$ equation

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -p(t) & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

are bounded if $p(\cdot)$ is pointwise nonnegative, periodic of period T with positive average value and with

$$\int_0^T p(t)dt \leq 4/T$$

Thus we see that all solutions of the $x^{[2]}$ equation

$$\begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \\ \dot{y}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -2p(t) & 0 & 2 \\ 0 & -p(t) & 0 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix}$$

are also bounded under the same hypothesis. (Here we have taken $y_2 = 2x_1x_2$ instead of $\sqrt{2}\, x_1x_2$). A change of basis puts this equation in a more symmetric form

$$\begin{bmatrix} \dot z_1(t) \\ \dot z_2(t) \\ \dot z_3(t) \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2}(1-p(t)) & 0 \\ -\frac{1}{2}(1-p(t)) & 0 & \frac{1}{2}(1+p(t)) \\ 0 & \frac{1}{2}(1+p(t)) & 0 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \end{bmatrix}$$

This equation evolves on the pseudo-orthogonal group SO(2,1).

One particular fact which should be mentioned here is that systems with a single time varying parameter, say

$$\dot x(t) = Ax(t)+k(t)Bx(t) \qquad (*)$$

go into systems with a single time varying parameter e.g.

$$\dot x^{[p]}(t) = (A_{[p]}+k(t)B_{[p]})x^{[p]}(t)$$

Thus the many useful results about (*) (circle criterion, [17], etc.) can be extended in a nontrivial way.

Exercises

1. It is known that all solutions of the differential equation

$$\ddot x + \dot x + k(t)x(t) = 0$$

remain bounded if $0 \leq k(t) \leq \sim 3.9$ (see [17]). On the other hand, if one picks a positive definite quadratic form in x and $\dot x$ say $v(x,\dot x)$ and computes its derivative along solutions of the given differential equation then there exists one quadratic form which implies stability via Liapunov theory, for $0 \leq k(t) \leq 1$ but the constant 1 cannot be improved on using a quadratic Liapunov function. However, if we look at the $x^{[p]}$ version of the differential equation then a quadratic Liapunov function for $x^{[p]}$ is a 2p-degree Liapunov function for the original equation and a more suitable Liapunov function can be found. Work out the details.

2. Consider a differential equation in $\mathbb{R}^n$

$$\dot x(t) = Ax(t)+k(t)Bx(t)$$

Suppose that A and B generate a four dimensional Lie algebra which is isomorphic with $g\ell(2)$. Use the theory of the representations of $g\ell(2)$ (see, e.g. Samelson [4] page 114) and the circle criterion (see, e.g. [17]) to derive stability criteria for the given system.

5.2 Periodic Self-Contragradient Systems

A matrix Lie algebra is said to be self-contragradient if there exists a matrix P such that

$$PLP^{-1} = -L'$$

for all L in the Lie algebra. For example, so(n) is self-contragradient with P=I and sp(n) is self-contragradient with P=J. As far as the stability of periodic systems is concerned, the important consequence of this assumption is that if A(t) satisfies $PA(t)P^{-1} = -A'(t)$ then the transition matrix for

$$\dot{x}(t) = A(t)x(t)$$

satisfies

$$\Phi_A'(t)P\Phi_A(t) = P$$

since

$$\frac{d}{dt}(\Phi_A'(t)P\Phi_A(t)) = \Phi_A'(t)(A'(t)P+PA(t))\Phi_A(t) = 0$$

Thus $\Phi_A(t)$ similar to $(\Phi_A^{-1})'$. As an immediate consequence of this fact we see that the eigenvalues of $\Phi_A(t)$ occur in reciprocal pairs -- if λ is an eigenvalue then so is $1/\lambda$. If we assume we are dealing with real systems then of course the eigenvalues occur in complex conjugate pairs as well.

If $A(t) = A(t+T)$ then the well known Floquet theory insures that the transition matrix for

$$\dot{x}(t) = A(t)x(t)$$

can be expressed as

$$\Phi_A(t) = Q(t)e^{Rt}; \quad Q(0) = I$$

with $Q(t+T) = Q(t)$ and R constant, though not necessarily real. The value of $\Phi_A(T)$ is decisive as far as the stability of a periodic system is concerned since $\Phi_A(nT) = [\Phi_A(T)]^n$.

If $A(t)$ is given by

$$A(t) = \sum_{i=1}^{m} a_i(t)A_i$$

with the A_i being a basis for a self-contragradient representation of a Lie algebra, then of course

$$\Phi_A'(t)P\Phi_A(t) = P$$

for all t. If $(\Phi_A(T))^n$ is bounded for n=1,2,... then we call $\Phi_A(T)$ stable. We call it P-strongly stable if it happens that for all sufficiently small R such that $R'P+PR = 0$, the matrix $e^R\Phi_A(T)$ is also stable. (Compare with [31].) In view of the fact that the eigenvalues of a matrix depend continuously on the elements of the matrix and in view of the fact that the eigenvalues of Φ_A must occur in reciprocal pairs, we see that if the eigenvalues of $\Phi_A(T)$ are distinct and if $\Phi_A(T)$ is stable, then it is P-strongly stable. However it can happen that $\Phi_A(T)$ is P-strongly stable even if the eigenvalues of $\Phi_A(T)$ are not distinct.

Theorem 1: If $\{A_i\}$ is the basis for a self-contragradient matrix Lie algebra, $A_i'P+PA_i = 0$, and if

$$\dot{x}(t) = (\sum_{i=1}^{m} a_i(t)A_i)x(t)$$

is periodic and if $\Phi_A(T)$ is P-strongly stable, then there exists $\varepsilon > 0$ such that for $|b_i(t)-a_i(t)| < \varepsilon$ and $b_i(t)$ periodic of period T the system

$$\dot{x}(t) = (\sum_{i=1}^{m} b_i(t)A_i)x(t)$$

is stable.

Exercises

1. Determine if for P = J the matrix

$$M = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & \cos & 0 & \sin\theta \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & -\sin & 0 & \cos\theta \end{bmatrix} ; \; 0 < \theta < \pi$$

is P-strongly stable or not. See [30], theorem 8.

2. Show that if p(t) is periodic of period T with average value zero and if

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -p(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

then $\Phi_A(T)$ is symplectic although $\Phi_A(t)$ for $t \neq T$ need not be. The corresponding $x^{[2]}$ equation is expressible as

$$\frac{d}{dt} \begin{bmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 0 \\ -1 & -p(t) & -1 \\ 0 & -2 & -2p(t) \end{bmatrix} \begin{bmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix}$$

Use the idea of strong stability to investigate the stability of these systems.

3. If D is diagonal then D+H is similar to a diagonal matrix if H is any symmetric matrix. However if D is diagonal there may exist an n(n-1)/2 dimensional set, the upper triangular matrices, such that D+T is not diagonalizable; consider the identity. Relate this to strong stability.

5.3 The Symplectic Case

In the special case of the symplectic group Krein [30] has given an elegant theorem on how large the perturbation in Theorem 1 of the previous section can be. We give an application of this theorem and some facts about realizations of feedback systems as well.

Notice that the second order system with Q(t) symmetric

$$\ddot{x}(t) + Q(t)x(t) = 0; \quad x(t) \in \mathbb{R}^n$$

is equivalent to the symplectic system

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & I \\ -Q(t) & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

Krein has investigated this set of equations and more general ones. One of his results reads as follows.

Theorem 1: Let P(t) = P(t+T) = P'(t), then all solutions of the equation

$$\ddot{x}(t)+P(t)x(t) = 0$$

are bounded if

i) $P(t) \geq 0$ all t

ii) $\int_0^T P(t)dt > 0$ (positive definite)

iii) $(4/T)I-\int_0^T P(t)dt > 0$ (positive definite)

Proof: See Krein [30], page 165.

As an example of an application of this result to problems of the type which arise frequently in system theory we prove the following theorem. (Compare with [32])

Theorem 2: Suppose that $q(s)$ and $p(s)$ are polynomials without common factors. Suppose further that $q(s)/p(s)$ is an even function of s with all its poles and zeros on the imaginary axis and assume that the poles and zeros of $sq(s)/p(s)$ interlace. Let $D = d/dt$ and let $k(\ )$ be periodic of period T. Then all solutions of the nth order differential equation

$$p(D)x(t) + k(t)q(D)x(t) = 0$$

are bounded provided

$$0 < \int |\lambda(t)|^2 dt < 4/T$$

where $\lambda(t)$ denotes the zero of $p(s)+k(t)q(s) = 0$ which has the largest magnitude.

Proof: Write $q(s)/p(s)$ as $r(s^2)/m(s^2)$ with r and m polynomials. This is possible because $q(s)/p(s)$ is even. Write $r(s)/m(s)$ as $b'(Is-A)^{-1}b$ with $A = A'$. This is possible because the poles and zeros of $r(s)/m(s)$ interlace, (See [25]). Thus

$$q(s)/p(s) = b'(Is^2-A)^{-1}b$$

and the differential equation in the theorem statement is equivalent to the system

$$\ddot{x} + (A+k(t)bb')x(t) = 0$$

Krein's result implies stability if

$$I(T/4) - \int_0^T (A+k(t)bb')dt > 0$$

But since the largest eigenvalue of the sum of two positive definite symmetric matrices is less than or equal to the sum of the largest eigenvalues of the respective matrices there is a corresponding inequality for integrals and we see that

$$\lambda_{max}\int_0^T (A+k(t)bb')dt \leq \int_0^1 |\lambda(t)|^2 dt$$

The result then follows.

It is interesting to compare this result with the analogous facts about completely symmetric systems investigated in [25]. Also notice that this theorem captures Liapunov's original theorem as a special case, as does the basic theorem of Krein.

Exercises

1. Use these results to investigate the stability of the scalar equation

$$x^{(4)}+4x^{(2)}+3x+k(t)(x^{(2)}+x) = 0$$

with k(t) periodic.

2. Derive a matrix version of Theorem 2.

VI. REFERENCES

1. H. Flanders, Differential Forms, Academic Press, 1963, New York.
2. J. Wei and E. Norman, "On the Global Representation of the Solutions of Linear Differential Equations as a Product of Exponentials," Proc. Am. Math. Soc., April 1964.
3. R. F. Gantmacher, Theory of Matrices, Chelsea, New York, 1959.
4. H. Samelson, Notes on Lie Algebras, Van Nostrand, New York, 1969.
5. M. Hausner and J. T. Schwartz, Lie Groups ; Lie Algebras, Gordon and Breach, London, 1968.
6. R. W. Brockett, "Lie Theory and Control Systems Defined on Spheres," SIAM J. on Applied Math., Vol. 24, No. 5, Sept. 1973.
7. W. L. Chow, "Über Systeme Von linearen partiellen Differentialgleichungen erster Ordinung," Math. Ann. Vol. 117, (1939), pp. 98-105.
8. R. W. Brockett, "System Theory on Group Manifolds and Coset Spaces," SIAM Journal on Control, Vol. 10, No. 2, May 1972, pp. 265-284.
9. S. Jonhansen (This Volume).
10. V. Jurdjevic and H. J. Sussmann, "Control Systems on Lie Groups," J. of Differential Equations, Vol. 12, No. 2, (1972) pp. 313-329.
11. R. Hirschorn, Ph.D. Thesis, Applied Mathematics, Harvard University, 1973.
12. R. E. Rink and R. R. Mohler, "Completely Controllable Bilinear Systems," SIAM J. on Control, Vol. 6, No. 3, 1968.
13. P. d'Alessandro, A. Isidori and A. Ruberti, "Realization and Structure Theory of Bilinear Dynamical Systems," (to appear, SIAM J. on Control).
14. R. W. Brockett, "On the Algebraic Structure of Bilinear Systems," in Variable Structure Control Systems, (R. Mohler and A. Ruberti, eds.) Academic Press, 1972, pp. 153-168.

15. A. J. Krener, A Generalization of the Pontryagin Maximal Principle and the Bang-Bang Principle, Ph.D. Thesis, Dept. of Mathematics, University of California, Berkeley, 1971.
16. H. J. Sussmann, "The Bang-Bang Problem for Linear Control Systems in $G\ell(n)$," SIAM J. on Control, Vol. 10, No. 3, Aug. 1972
17. R. W. Brockett, Finite Dimensional Linear Systems, John Wiley and Sons, New York, 1970.
18. R. Debye, "Polar Molecules," The Chemical Catalogue Co., New York, 1929.
19. C. P. Poole and H. A. Farach, Relaxation in Magnetic Resonance, Academic Press, 1971.
20. F. Perrin, "Etude Mathematique du Mouvement Brownien de Rotation," Ann. Ecole Norm. Sup., Vol. 45, (1928), pp. 1-51.
21. R. F. Fox, Physical Applications of Multiplicative Stochastic Processes, J. of Mathematical Physics, Vol. 14, No. 1, Jan. 1973, pp. 20-25.
22. V. A. Tutubalin, "Multimode Waveguides and Probability Distributions on a Symplectic Group," Theory of Probability and Its Applications, Vol. 16, No. 4, pp. 631-642, 1971.
23. G. F. Carrier, "Stochastically Driven Dynamical Systems," J. of Fluid Mechanics, Vol. 44, Part 2 (1970) pp. 249-264.
24. Martin Clark, (This Volume).
25. R. W. Brockett and J. C. Willems, "Average Value Criteria for Stochastic Stability," Stability of Stochastic Dynamical Systems, (ed. Ruth Curtain), Springer-Verlag, 1972.
26. H. P. McKean, Stochastic Integrals, Academic Press, New York, 1969.
27. S. Helgason, Differential Geometry and Symmetric Spaces, Academic Press, New York, 1960.
28. M. E. Rose, Elementary Theory of Angular Momentum, J. Wiley, New York, 1957.
29. L. Hormander, "Hypoelliptic Second Order Differential Equations," Acta. Math., 119: 147-171, 1967.
30. M. G. Krein, "Introduction to the Geometry of Indefinite J-Spaces and to the Theory of Operators in Those Spaces," Am. Math. Soc. Translations, (Series 2) Vol. 93, 1970, pp. 103-176.
31. I. M. Gel'fand and V. B. Lidskii, "On the Structure of the Regions of Stability of Linear Canonical Systems of Differential Equations with Periodic Coefficients," Am. Math. Soc. Transl. Series 2, Vol. 8, (1958) pp. 143-182.
32. R. W. Brockett and A. Rahimi, "Lie Algebras and Linear Differential Equations," in Ordinary Differential Equations, (L. Weiss ed.) Academic Press, New York, 1972, pp. 379.