SIAM J. APPL. MATH.
Vol. 25, No. 2, September 1973

LIE THEORY AND CONTROL SYSTEMS DEFINED ON SPHERES*

R. W. BROCKETT†

Abstract. We show in this paper that in constructing a theory for the most elementary class of control problems defined on spheres, some results from Lie theory play a natural role. In particular to understand controllability, optimal control, and certain properties of stochastic equations, Lie theoretic ideas are needed. The framework considered here is probably the most natural departure from the usual linear system/vector space problems which have dominated the control systems literature. For this reason our results are compared with those previously available for the finite-dimensional vector space case.

1. Introduction. Specific results about control systems whose state spaces are spheres have been useful in understanding problems in energy conversion, controlled rigid body dynamics, etc. Some examples are mentioned in our earlier paper [1]. Here we work out in more detail, and in greater generality, the theory for a class of problems of this type and compare our results with the case where the state space is a vector space. To carry out this program requires some results from Lie theory, Lie groups acting on spheres, etc. There has been no attempt here to discuss the most general setting in which techniques which we use are applicable. Instead we have taken the sphere problems as a model and have studied a range of control-theoretic questions in that setting. A number of possible generalizations will be apparent.

To begin with, we mention some well-known facts about linear system theory. We do this to make the paper a little more accessible to those not familiar with control problems and to sensitize the reader to certain issues important in control. For a more complete account and references to the literature one can consult [2] for the deterministic results and [3] for the stochastic results.

Linear system theory deals with the pair of equations

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad y(t) = Cx(t), \tag{1.1}$$

where $\dot{x}$ denotes a time derivative. It is assumed that $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$. For simplicity we take A, B, C to be constant matrices. One calls u the *control*, x the *state* and y the *output*. The theory of linear system is extensive but for our present purposes we point out only the following five results.

(i) The pair of equations (1.1) is said to be *controllable* if for every x_0 and x_1 in $\mathbb{R}^n$ and every $t_1 > 0$ there exists a piecewise continuous control $u(\cdot)$ such that if $x(0) = x_0$ then $x(t_1) = x_1$. A necessary and sufficient condition for controllability is that Rank $(B, AB, \cdots, A^{n-1}B) = n$ where , indicates a column partition.

* Received by the editors November 27, 1972. Presented by invitation at the Drexel University–Society for Industrial and Applied Mathematics Conference on Lie Algebras: Applications and Computational Methods, sponsored in part by a grant from the United States Air Force Office of Scientific Research, held at Drexel University, Philadelphia, Pennsylvania, June 15 and 16, 1972.

† Division of Engineering and Applied Physics, Harvard University, Cambridge, Mass. 02138. 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. It was partially written while the author held a Science Research Council (U.K.) Senior Visiting Fellowship at Imperial College, London.

(ii) The pair of equations (1.1) is said to be *observable* if for every $x_1 \neq x_2$ and every $t_1 > 0$ the outputs corresponding to x_1 and x_2 differ on the interval $[0, t_1]$. A necessary and sufficient condition for observability is that rank $(C; CA; \cdots; CA^{n-1}) = n$ where ; indicates a row partition.

(iii) If the pair (1.1) is controllable, then for every given x_0 and x_1 in $\mathbb{R}^n$ and every $t_1 > 0$ there exists a piecewise continuous control u defined on $[0, t_1]$ which transfers the state from x_0 at $t = 0$ to x_1 at $t = t_1$ and minimizes

$$\eta(t) = \int_0^{t_1} u'(t)u(t)\, dt \tag{1.2}$$

relative to all other piecewise continuous controls which accomplish the same transfer.

(iv) If there exists a linear feedback control law $u = Fx$ such that $\dot{x} = (A + BF)x$ has a null solution which is asymptotically stable, then there exists a control law $u = Kx$ such that $\lim_{t\to\infty} x(t) = 0$ and the functional

$$\eta = \int_0^{\infty} u'(t)u(t) + y'(t)y(t)\, dt$$

is minimized by setting $u(t) = Kx(t)$.

(v) If (1.1) is controllable and if the differential equation $\dot{x} = Ax$ is asymptotically stable, then the associated stochastic equation (for notation, see [3])

$$dx(t) = Ax(t)\, dt + B\, dw(t) \tag{1.3}$$

has a unique invariant Gaussian measure which has zero mean and variance Q satisfying

$$QA + A'Q = -BB'. \tag{1.4}$$

In this paper we establish analogues for each of these results for systems of the type

$$\dot{x}(t) = \left(A + \sum_{i=1}^{m} u_i(t)B_i\right)x(t), \qquad y(t) = Cx(t), \tag{1.5}$$

where $A, B_1, B_2, \cdots, B_m$ are skew symmetric matrices and the system can be thought of as evolving on the sphere $\|x(t)\| = \|x(0)\|$.

One significant point in the linear theory is that the matrix B is generally not invertible and cases for which it is invertible are so infrequent as to be virtually without interest. If B is invertible, then by an appropriate choice of basis equation (1.1) becomes

$$\dot{x}(t) = Ax(t) + u(t) \tag{1.6}$$

and controllability is automatic. Moreover, in this case problems (iii) and (iv) are easily reduced to variational problems of the classical type

$$\eta = \int_0^{t_1} L(x, \dot{x})\, dt \tag{1.7}$$

with L quadratic in x and $\dot{x}$ and $L_{\dot{x}\dot{x}}$ positive definite. Control theory works with the more general "degenerate" case where $L_{\dot{x}\dot{x}}$ is only nonnegative definite but certain constraints are in effect. If the above integral is thought of as the action integral in a mechanics problem, then the case treated in control theory allows for the possibility of certain zero masses provided there are appropriate linear constraints between position and velocity. It can also be thought of as a limiting case of an unconstrained dynamical problem where certain masses and associated energies go to infinity. This second interpretation is generally more useful. Remarks of the same type apply to equation (1.3) where existence of a smooth transition density is well known if B is invertible whereas the same is true, but for rather more subtle reasons, if we assume controllability instead of invertibility of B.

2. Controllability. One of the main areas of applicability of Lie theory in control has been that of determining the set of points reachable along solution curves of $\dot{x}(t) = f(x(t), u(t), t)$ for the set of all piecewise continuous controls $u(\cdot)$. For studies of this kind see [4]–[10]. If the control equations are of the form

$$\dot{x}(t) = \left(A + \sum_{i=1}^{m} u_i(t)B_i\right)x(t), \qquad x(t) \in \mathbb{R}^n, \tag{2.1}$$

then the system typically evolves on a manifold in $\mathbb{R}^n$. The determination of the set of points reachable from a given point x_0 can be accomplished by the determination of the set of matrices reachable from the identity for the matrix equation

$$\dot{X}(t) = \left(A + \sum_{i=1}^{m} u_i(t)B_i\right)X(t), \qquad X(0) = I, \tag{2.2}$$

and then letting this set act on x_0 via ordinary matrix-vector multiplication. Equation (2.2) can be thought of as defining a control problem on a matrix Lie group. The question of determining what matrices are reachable from the identity along solutions of (2.2) has been the subject of a number of papers [1], [7]–[10]. Following Jurdjevic and Sussmann, we term systems of the form of (2.2) *right invariant.* This is appropriate because the vector fields defined on the $Gl(n)$ by the right side of (2.2) are invariant under the translation defined by right multiplication with an element of $Gl(n)$. We shall say that (2.2) is *controllable* on a group $\mathscr{G}$ if any two points in $\mathscr{G}$ can be joined by a solution curve generated by some piecewise continuous control $u(\cdot)$.

Suppose that A and $B_1, B_2, \cdots, B_m$ are all skew symmetric. Then regardless of the choice of u the solutions of (2.1) remain on the sphere defined by $\|x(t)\| = \|x(0)\|$. We shall say that the system (2.1) is *controllable* on the sphere if any two points on the sphere be joined by a solution curve generated by some piecewise continuous curve $u(\cdot)$. Phrased another way, the system is controllable if the set of matrices reachable from the identity along solutions of (2.2) act transitively on S^{n-1}. From earlier results [10] we know that since the motion is confined to a subgroup of $SO(n)$ the set of matrices reachable from I is the matrix Lie group consisting of all the matrices which can be expressed as products of the form $\exp H_1, \exp H_2, \cdots, \exp H_n$, where $H_1, H_2, \cdots, H_n$ belong to the Lie algebra generated by $A, B_1, B_2, \cdots, B_n$.

Now of course the orthogonal group $SO(n)$ acts transitively on S^{n-1} so that if the algebra generated by $A, B_1, B_2, \cdots, B_m$ is the full set of skew symmetric matrices, then the system (2.1) is controllable on S^{n-1}. However, there are certain subgroups of $SO(n)$ which act transitively on S^{n-1} as well. The real compact forms of the classical Lie groups are all candidates. The results are well known [11] but we repeat them here. For example, it is clear that both the full unitary group and the special unitary group of dimension n act transitively on the set of complex n-vectors whose Hermitian length is one. But this set is just a set of vectors with components $(x_i + \sqrt{-1}y_i)$ such that

$$\sum_{i=1}^{n} (x_i^2 + y_i^2) = 1, \tag{2.3}$$

which is a $(2n - 1)$-dimensional sphere. Thus by defining the realification [12] of the unitary algebras by the Lie algebra homomorphism

$$B \to \begin{bmatrix} \operatorname{Re} B & \operatorname{Im} B \\ -\operatorname{Im} B & \operatorname{Re} B \end{bmatrix} \tag{2.4}$$

we obtain a set of real matrices whose associated group acts transitively on S^{2n-1}. The real compact form of C_n is the intersection of a special unitary group and the symplectic groups. Naturally this representation is in terms of matrices of even dimension so that they can act on even dimensional complex vectors only. Thus, by analogy with the unitary case, the real compact form of C_n acts on the sphere of dimension S^{4n-1}. This action is known to be transitive and of course we can add to the algebra real multiples of $\sqrt{-1}I$ to get the "full quaterion-unitary group" which acts transitively as well. These four cases, each valid for all integer n, together with three particular ones account for all possibilities. The particular cases may be explained as follows. The exceptional algebra G_2 admits a 7-dimensional skew-symmetric representation whose exponential acts transitively on S^6. The spin representation of $SO(7)$ is 8-dimensional and it acts transitively on S^7. The spin representation of $SO(9)$ is 16-dimensional and it acts transitively on S^{15}. With this explanation we can state the following result.

THEOREM 1. *Let $A, B_1, \cdots, B_m$ be a collection of $n \times n$ skew symmetric matrices. The control system*

$$\dot{x}(t) = \left(A + \sum_{i=1}^{m} u_i(t)B_i\right)x(t) \tag{2.5}$$

is controllable on S^{n-1} if the algebra generated by $A, B_1, B_2, \cdots, B_m$ is:

(i) *$SO(n)$ for $n = 0 \bmod 2$;*

(ii) *$SO(n)$ or the realification of $SU(n/2)$ or $U(n)$ for $n = 1 \bmod (2)$;*

(iii) *the realification of $Sp(n/2)$ for $n = 1 \bmod (4)$;*

(iv) *G_2 if $n = 6$, Spin (8) if $n = 7$ or Spin (16) if $n = 15$.*

Moreover, if the Lie algebra is not one of these cases the system (2.8) is not controllable.

If the system is not controllable on S^{n-1} it is sometimes of interest to compute exactly what points can be reached from a given initial state. The determination of what points belong to this set is facilitated by a knowledge of the structure of the representation defined by the matrices in the algebra generated by $A, B_1, B_2, \cdots, B_m$. If this representation is not irreducible, then its reduction is clearly the first step in the determination of the reachable set. The properties of the irreducible pieces may reveal the form of the reachable set in a straightforward way. For example, if the evolution equation can be decomposed as

$$\dot{x} = \left[I \otimes A^1 + A^2 \otimes I + \sum_{i=1}^{m} u_i(I \otimes B_i^1 + B_i^2 \otimes I)\right]x(t), \tag{2.6}$$

then the Kronecker product of the reachable group for

$$\dot{X}(t) = \left(A^1 + \sum_{i=1}^{m} u_i(t)B_i^1\right)X(t) \tag{2.7}$$

and the reachable group for

$$\dot{X}(t) = \left(A^2 + \sum_{i=1}^{m} u_i(t)B^2\right)X(t) \tag{2.8}$$

contains the reachable group for (2.2). The reachable group will not, in general, simply be the Kronecker product of the reachable groups unless the effects of the u's are decoupled.

For the linear evolution equation (1.1) it happens that if it is possible to transfer any state to any other state then this transfer can be done in arbitrarily small time. This is not the case for systems defined by (2.1). Jurdjevic and Sussmann [10] give an example of a system defined on S^2 which is controllable but certain transfers cannot be made in less than 1 unit of time. Thus if (1.1) is controllable on S^n the strongest statement we can make on the basis of the present analysis is that for t_1 sufficiently large every state can be transferred to every other state in t_1 units of time. Estimates on this time have not yet been worked out.

In the vector space case controllability is closely related to the concept of observability as mentioned in the Introduction. In the present setting this is not the case at all. We say that the system

$$\dot{x}(t) = \left(A + \sum_{i=1}^{m} u_i(t)B_i\right)x(t), \qquad y(t) = Cx(t) \tag{2.9}$$

is *observable* on S^{n-1} if no two distinct initial states on S^{n-1} give rise to the same response y for all controls $u(\cdot)$. The following theorem gives a necessary and sufficient condition for observability.

THEOREM 2. *Let $A, B_1, B_2, \cdots, B_m$ be a collection of skew symmetric matrices and let c be a unit vector. The control system*

$$\dot{x}(t) = \left(A + \sum_{i=1}^{m} u_i(t)B_i\right)x(t), \qquad y(t) = cx(t)$$

is observable on S^{n-1} *if and only if the set of matrices* $\{A, B_1, B_2, \cdots, B_m, cc'\}$ *is irreducible.*

For a proof of this theorem and more general results of this type, see [13].

3. Optimal control. Consider again the evolution equation (2.2) defined on matrix group $\mathscr{G}$. Let there be given a time $t_1 > 0$ and boundary conditions of the form $X(0) = X_0$; $X(t_1) = X_1$. Suppose that in addition there is given a functional which is of the *action* type

$$\eta_1 = \frac{1}{2}\int_{t_0}^{t_1} \sum_{i=1}^{m} u_i^2(t)\,dt \tag{3.1}$$

as opposed to the *geodesic* type

$$\eta_2 = \int_{t_0}^{t_1} \left(\sum_{i=1}^{m} u_i^2(t)\right)^{1/2} dt. \tag{3.2}$$

Our problem is to determine if there exists a control $u(\cdot)$ such that the boundary conditions are met and the given functional is minimized and, if such a control exists, to characterize it. Just as with controllability, there is an obvious connection between problems defined on a group and problems defined on a manifold on which that group acts. This would no longer be the case if η depended on x in a general way.

We shall use the formalism of the maximum principle of Pontryagin [14] rather than the calculus of variations to attack this problem because it handles the degeneracy which is built into the problem in a natural way. Applied to the present problem, Pontryagin's maximum principle asserts that if $u(\cdot)$ is an optimizing control then there exists a matrix P such that

$$\dot{P}(t) = -A'P(t) - \sum_{i=1}^{m} u_i(t)B_i'P(t) \tag{3.3}$$

and H, defined by

$$H(P, X, u) = \langle P, AX\rangle + \sum_{i=1}^{m} u_i\langle P, B_iX\rangle + \sum_{i=1}^{m} \frac{1}{2}u_i^2 \tag{3.4}$$

is minimized with respect to u by the optimal control. Thus we have the optimal control given by

$$u_i(t) = \langle -P(t), B_iX(t)\rangle. \tag{3.5}$$

This choice of u gives a pair of differential equations with split boundary conditions

$$\frac{d}{dt}\begin{bmatrix} X(t) \\ P(t) \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & -A' \end{bmatrix}\begin{bmatrix} X(t) \\ P(t) \end{bmatrix} - \sum_{i=1}^{m} \langle P, B_iX\rangle \begin{bmatrix} B_i & 0 \\ 0 & -B_i' \end{bmatrix}\begin{bmatrix} X(t) \\ P(t) \end{bmatrix}. \tag{3.6}$$

The problem can be reduced to a single quadratic equation with split boundary conditions by introducing $K = XP'$. An easy calculation shows that

$$\dot{K}(t) = AK(t) - K(t)A' - \sum_{i=1}^{m} \langle B_i', K(t) \rangle (B_i K(t) - K(t)B_i'). \tag{3.7}$$

So far everything is valid for an arbitrary subgroup of $Gl(n)$. If $A, B_1, B_2, \cdots, B_m$ are self-contragredient, then a simplification occurs. In that case any solution of the differential equation for P can be expressed in terms of a solution of the differential equation for X with nonsingular boundary conditions; that is, $P(t) = NX(t)M$ for some constant matrices M and N. Specializing to the skew symmetric case gives the following result.

THEOREM 4. *Suppose that $A, B_1, B_2, \cdots, B_m$ are skew symmetric $n \times n$ matrices and suppose that there exists a piecewise continuous control $u(\cdot)$ which transfers the state of the matrix system*

$$\dot{X}(t) = \left(A + \sum_{i=1}^{m} u_i(t) B_i \right) X(t) \tag{3.8}$$

from X_0 at $t = 0$ to X_1 at $t = t_1 > 0$. Then there exist constant matrices M and N such that the solution of

$$\dot{X}(t) = \left(A + \sum_{i=1}^{m} \langle B_i, X(t)MX'(t)N \rangle B_i \right) X(t), \qquad X(0) = X_0 \tag{3.9}$$

passes through X_1 at $t = t_1$. Moreover, there exists one such pair M, N which minimizes η_1 relative to any other continuous $u(\cdot)$ which steers the system to X_1 from X_0 in the same period of time.

Proof. That there exists an optimal control follows from Theorem 6 of Cesari [15]. The rest follows from the maximum principle as discussed above.

There is an alternative point of view available for these problems which makes a little closer contact with both physics and Lie theory but which is not so useful here. Consider the right-invariant control equation in $SO(n)$ with control Ω:

$$\dot{X}(t) = \Omega(t)X(t), \qquad X(0) = X_0. \tag{3.10}$$

Let the problem be to pick Ω in the space of skew symmetric matrices such that $X(t_1) = X_1$ and the trace form

$$\eta = \int_0^{t_1} - \operatorname{tr} (I^{-1}\Omega)^2 \, dt \tag{3.11}$$

is minimized. Elementary variational arguments with due regard for the admissibility of variations lead to the Euler equation

$$\dot{\Omega} = \Omega I \Omega I^{-1} - I^{-1} \Omega I \Omega. \tag{3.12}$$

In $SO(3)$ this matrix equation is equivalent to the familiar Euler equations for a rigid body

$$\begin{aligned} I_1 \dot{\omega}_1 &= (I_2 - I_3)\omega_2\omega_3, \\ I_2 \dot{\omega}_2 &= (I_3 - I_1)\omega_1\omega_3, \\ I_3 \dot{\omega}_3 &= (I_1 - I_2)\omega_1\omega_2, \end{aligned} \tag{3.13}$$

which, after all, come from minimizing the action integral on $SO(3)$. (Note that the kinetic energy of a rigid body can be expressed by the trace form $(\det I)\operatorname{tr}(I^{-1}\Omega)^2$, where I is the usual inertia tensor (see [2, p. 64]). Incidentally, this also serves to define the degree of difficulty of actually solving the control problem mentioned above. Since it is well known that the solution of the Euler equations generally involves elliptic functions, the solution of the optimal control problems cannot be expressed in terms of elementary functions except in special cases.

By far the simplest special case on $SO(n)$ occurs when η_1 is the negative of the integral of the Killing form. That is, given $X(0)$ and $X(1)$ and given the evolution equation

$$\dot{X}(t) = \sum_{i=1}^{n(n-1)/2} u_i(t)B_iX(t), \qquad X \in SO(n), \tag{3.14}$$

where $B_i = -B_i'$ and for all i and j

$$\langle B_i, B_j\rangle = \operatorname{tr} B_iB_i' = \delta_{ij}, \tag{3.15}$$

one finds that the optimal trajectory is

$$X(t) = e^{\Omega t}X(0), \tag{3.16}$$

where Ω is the solution of $e^{\Omega} = X(1)X^{-1}(0)$ which has the smallest Frobenius norm.

We turn now to applying the above results to the problem of optimizing trajectories on spheres. Note that trajectories on spheres can be optimized for fixed endpoints by solving an associated right invariant group problem and then picking the minimizing element in the group for transferring x_0 to x_1. The following theorem expresses this.

THEOREM 5. *Let $A, B_1, B_2, \cdots, B_m$ be skew symmetric matrices. Suppose that the system*

$$\dot{x}(t) = \left(A + \sum_{i=1}^{n} u_i(t)B_i\right)x(t) \tag{3.17}$$

is controllable on S^n. Then given a sufficiently large time $t_1 > 0$ and given points x_0 and x_1 in S^{n-1}, there exists a control which transfers the system from x_0 at $t = 0$ to x_1 at $t = t_1$ and minimizes

$$\eta = \int_0^{t_1} u'(t)u(t)\,dt. \tag{3.18}$$

Moreover, there exists a matrix K_0 such that the optimal control is given by $u_i(t) = \langle K(t), B_i\rangle$, where K is defined by the matrix differential equation

$$\dot{K}(t) = [A, K(t)] + \sum_{i=1}^{m} \langle K(t), B_i\rangle[K(t), B_i], \qquad K(0) = K_0. \tag{3.19}$$

We complete this section on optimal control with a result of the type which plays a major role in linear system theory in connection with the regulator problem.

THEOREM 6. *Let A and B be $n \times n$ skew symmetric matrices and consider the system*

$$\dot{x}(t) = Ax(t) + u(t)Bx(t). \tag{3.20}$$

Let a be a unit vector in the null space of A such that A and $Baa'B'$ are a pair of matrices which act irreducibly on the orthogonal complement of the one-dimensional subspace defined by a. Then the control law $u(t) = a'Bx(t)$ steers the system from any initial state $x_0 \neq -a$ to a and minimizes the integral

$$\eta = \int_0^\infty u^2(t) + [a'Bx(t)]^2 \, dt$$

relative to any other continuous control $u(\cdot)$.

Proof. We can write η as

$$\eta = \int_0^\infty u^2(t) - 2a'\dot{x}(t) + [a'Bx(t)]^2 \, dt + 2a'x(t)\Big|_0^\infty ;$$

since $Aa = 0$ we have

$$\eta = \int_0^\infty (u(t) - a'Bx(t))^2 \, dt + 2a'x(t)\Big|_0^\infty .$$

Thus if the control law $u(t) = a'Bx(t)$ actually drives the state x to a then it is optimal. However, observing that $a'x(t)$ has a derivative along the given solution which is equal to $-[a'Bx(t)]^2$, we see by LaSalle's theorem (see, for example, [2]) that the solution $x = a$ can fail to be stable if and only if $a'B\, e^{At}x$ vanishes identically for some $x \neq \pm a$. By looking at the derivatives at $t = 0$ we see that this can happen if and only if $(Ba, ABa, \cdots, A^{n-1}Ba)$ fails to span the orthogonal complement of the one-dimensional subspace defined by a.

4. Stochastic differential equations. We consider now a third aspect of control theory on spheres. This has to do with the analogue of property (v) mentioned in the Introduction. What we show is that controllability implies the existence of a unique invariant measure for a stochastic equation on S^{n-1}. We use Itô notation for stochastic differential equations. Wong [3] can be consulted for an explanation of both the mathematics and the notation.

Let $w_1, w_2, \cdots, w_m$ denote independent Wiener (Brownian motion) processes of unitary variance. In giving a precise meaning to differential equations in which something like "white noise" appears, K. Itô [16] invented what has proven to be a very successful calculus in which the standard differentiation rule is significantly modified insofar as differentials of Wiener processes are concerned. In this calculus $dw_i\, dw_j = \delta_{ij}\, dt$, a first order term; $dw_i\, dt$, and $(dt)^2$ are both higher than first order. We discuss the implication of this in one important special case. If x and y are vectors satisfying the Itô differential equations

$$dx(t) = Ax(t)\, dt + Bx(t)\, dw(t), \tag{4.1}$$

$$dy(t) = Fy(t)\, dt + Gy(t)\, dw(t), \tag{4.2}$$

then $z(t) = x(t)y'(t)$ satisfies the Itô equation

$$dz(t) = (Az(t) + z(t)F' + Bz(t)G')\,dt + (Bz(t) + z(t)G')\,dw. \tag{4.3}$$

The only other fact we need about Itô equations concerns the associated mean equation. If x and y satisfy equations (4.1) and (4.2), then $\bar{x}(t) = \mathscr{E}x(t)$ and $\bar{y}(t) = \mathscr{E}y(t)$ satisfy the ordinary differential equation

$$\frac{d}{dt}\bar{x}(t) = A\bar{x}(t), \tag{4.4}$$

$$\frac{d}{dt}\bar{y}(t) = F\bar{y}(t). \tag{4.5}$$

We shall see that these two results permit the derivation of equations for all moments and imply that the moment equations are decoupled from each other.

Recall that the number of linearly independent degree p forms in n variables is given by

$$N(n, p) = \binom{n + p - 1}{p}. \tag{4.6}$$

We can therefore associate with each n tuple $(x_1, x_2, \cdots, x_n)$ a $N(n, p)$-tuple $x^{[p]} = (x_1^p, \sqrt{p}x_1^{p-1}x_2, \cdots, x_n^p)$ where the coefficients are chosen in such a way as to validate the equality

$$\|x^{[p]}\|^2 = \|x\|^{2p}. \tag{4.7}$$

It is clear that if x satisfies an ordinary differential equation which is linear, say,

$$\frac{d}{dt}x(t) = Ax(t), \tag{4.8}$$

then $x^{[p]}$ also satisfies a linear differential equation

$$\frac{d}{dt}x^{[p]}(t) = A^{[p]}x(t). \tag{4.9}$$

We regard this as a definition of $A^{[p]}$. It is related to the classical idea of an induced representation. Of course if there are controls present a similar set of equations follows; that is, (2.1) implies

$$\frac{d}{dt}x^{[p]}(t) = A^{[p]}x^{[p]}(t) + \sum_{i=1}^{m} u_i(t)B_i^{[p]}x^{[p]}(t). \tag{4.10}$$

Similar remarks hold for stochastic equations of the type under consideration here, provided suitable allowance is made for the Itô calculus. Associated with the Itô equation

$$dx(t) = Ax(t)\,dt + \sum_{i=1}^{m} B_i x(t)\,dw_i \tag{4.11}$$

is the family of equations

$$(4.12)\quad dx^{[p]}(t) = \left(\left(A - \sum_{i=1}^{m} \tfrac{1}{2}B_i^2\right)^{[p]} + \sum_{i=1}^{m} \tfrac{1}{2}(B_i^{[p]})^2\right)x^{[p]}(t)\,dt + \sum_{i=1}^{m} B_i^{[p]}x^{[p]}(t)\,dw_i.$$

The derivation of this is a straightforward exercise using the properties of dw_i outlined above. Finally, we have the moment equations associated with (4.11):

$$(4.13)\quad \frac{d}{dt}\bar{x}^{[p]}(t) = \left(\left(A - \sum_{i=1}^{m} \tfrac{1}{2}B_i^2\right)^{[p]} + \sum_{i=1}^{m} \tfrac{1}{2}(B_i^{[p]})^2\right)\bar{x}^{[p]},$$

where $\bar{x}^{[p]}(t) = \mathscr{E}x^{[p]}(t)$. (Compare with [17].)

In terms of the Itô calculus when can the matrix stochastic equation

$$(4.14)\quad dX(t) = AX(t)\,dt + \cdot\sum_{i=1}^{m} dw_i(t)B_iX(t)$$

be thought of as evolving the orthogonal group? This will be the case when the associated vector equation (4.11) evolves on the sphere defined by $\|x(t)\| = \|x(0)\|$ for all $x(0)$. Using the facts outlined above we see that $d(x'x) = 0$ if and only if for all i

$$(4.15)\quad B_i = -B_i', \qquad A - \sum_{i=1}^{m} \tfrac{1}{2}B_i^2 = -\left(A - \sum_{i=1}^{m} \tfrac{1}{2}B_i^2\right)'.$$

Thus these are the conditions under which (4.14) evolves in the orthogonal group and the conditions under which (4.11) evolves on the sphere.

It is apparent that the measure associated with the uniform density on the sphere is an invariant measure for the process defined by equation (4.11). Since the area of the $(n-1)$-sphere is $2\pi^{n/2}/\Gamma(n/2)$ the uniform density is

$$(4.16)\quad \rho_0(x) = \Gamma(n/2)/2\pi^{n/2}.$$

The corresponding values of the odd moments are zero by symmetry but the even moments are not. The following theorem claims that all the moments approach the moments associated with a uniform distribution if we have controllability. Incidentally, equation (4.13) provides a means for actually computing the moments for all time in terms of their values at $t = 0$.

THEOREM 7. *Suppose that $A, B_1, B_2, \cdots, B_m$ are all skew symmetric and suppose that*

$$(4.17)\quad \dot{x}(t) = \left(A + \sum_{i=1}^{m} u_i(t)B_i\right)x(t)$$

is controllable on S^{n-1}. Then the solution of the Itô differential equation defined on the sphere by

$$(4.18)\quad dx(t) = \left(A + \sum_{i=1}^{m} \tfrac{1}{2}B_i^2\right)x(t)\,dt + \sum_{i=1}^{m} B_ix(t)\,dw_i$$

is such that all moments approach the moments associated with a uniform distribution on the $n-1$ sphere as t approaches infinity.

Proof. First of all, note the shift in notation from (4.11) to (4.18). In (4.11) $A + \frac{1}{2}\sum B_i^2$ is playing the role played by A alone here. It is not difficult to show that because $A, B_1, B_2, \cdots, B_m$ are skew symmetric it follows that $A^{[p]}, B_1^{[p]}, B_2^{[p]}, \cdots, B_m^{[p]}$ are also skew symmetric. A second observation concerns stability. If $A = -A'$ and $B_i = -B_i'$, then all solutions of the ordinary differential equation

$$\dot{x}(t) = \left(A + \sum_{i=1}^{m} \tfrac{1}{2} B_i^2\right) x(t) \tag{4.19}$$

are bounded. Moreover, each solution approaches zero as t approaches infinity provided $B_i e^{At}x$ does not vanish identically for any $x \neq 0$ and there will exist nonzero vectors such that $B_i e^{At}x$ vanishes identically if and only if A and B_i can be put in the form

$$\theta' A \theta = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \qquad \theta' B_i \theta = \begin{bmatrix} B_i & 0 \\ 0 & 0 \end{bmatrix}. \tag{4.20}$$

To prove the first of these facts we notice that since $A = -A'$,

$$\frac{d}{dt}\|x(t)\|^2 = -\sum_{i=1}^{m} \|B_i x(t)\|^2. \tag{4.21}$$

Thus by LaSalle's theorem (see, for example, [2]) the solution either goes to zero or else there is a solution along which $\|B_i x(t)\|$ vanishes identically for all i. That solution would have to be of the form $e^{At}x_0$. As for the conditions on A and B_i, they follow from considering the subspace of vectors such that $B_i e^{At}x$ vanishes, together with its orthogonal complement, making use of the skew symmetry of $A, B_1, B_2, \cdots, B_m$.

Clearly controllability implies that all solutions of the mean equation approach zero as t approaches infinity because controllable systems cannot be decomposed as indicated. As for the higher moments, we must distinguish between the even and odd cases. For the odd cases if there is a decomposition, then controllability of the equation (4.17) is clearly impossible. For the even moments, we have in view of the identity $\|x^{[p]}\|^2 = \|x\|^{2p}$, a decomposition of the type given by (4.20) but with the zero block in B_i being one-dimensional. The one-dimensional subspace defines the steady state value of the even moments. On the orthogonal complement the equation (4.18) is asymptotically stable. These remarks are related to some well-known properties of orthogonal representations of Lie algebras.

As is well known, the moments $x^{[p]}$ are related to the spherical harmonics in a direct way. Thus by working with equation (4.13) it is possible to obtain a full solution to the Fokker–Plank equation associated with the Ito equation (4.18). The interpretation of the moments in terms of spherical harmonics also allows one to establish some qualitative features of the probability density. In particular its smoothness and convergence to the steady state can be easily studied.

REFERENCES

[1] R. W. BROCKETT, *System theory on group manifolds and coset spaces*, SIAM J. Control, 10 (1972), pp. 265–284.

[2] ———, *Finite Dimensional Linear Systems*, John Wiley, New York, 1970.

[3] E. WONG, *Stochastic Processes in Information and Dynamical Systems*, McGraw-Hill, New York, 1971.

[4] R. HERMANN, *On the accessibility problem in control theory*, International Symposium on Nonlinear Differential Equations and Nonlinear Mechanics, Academic Press, New York, 1963, pp. 325–332.

[5] C. LOBRY, *Contrôlabilité des systèmes non linéares*, SIAM J. Control, 8 (1970), pp. 573–605.

[6] G. W. HAYNES AND H. HERMES, *Nonlinear controllability via Lie theory*, Ibid., 8 (1970), pp. 450–460.

[7] J. KUCERA, *Solution in large of control problem*: $\dot{x} = (A(1 - u) + Bu)x$, Czech. Math. J., 16 (1966), no. 91, pp. 600–623.

[8] ———, *Solution in large of control problem*: $\dot{x} = (Au + Bu)x$, Ibid., 17 (1967), no. 92, pp. 91–96.

[9] ———, *On accessibility of bilinear systems*, Ibid., 20 (1970), no. 95, pp. 160–168.

[10] V. JURDJEVIC AND H. J. SUSSMANN, *Control systems on Lie groups*, J. Differential Equations, 12 (1972), no. 2, pp. 313–329.

[11] H. SAMUELSON, *Topology of Lie groups*, Bull. Amer. Math. Soc., 58 (1952), pp. 2–37.

[12] ———, *Notes on Lie Algebras*, Van Nostrand Reinhold, New York, 1969.

[13] R. W. BROCKETT, *On the algebraic structure of bilinear systems*, Theory and Applications of Variable Structure Systems, R. Mohler and A. Ruberti, eds. Academic Press, New York, 1972.

[14] L. S. PONTRYAGIN, V. BOLTYANSKII, R. GAMKRELIDZE AND E. MISHCHENKO, *The Mathematical Theory of Optimal Processes*, Interscience, New York, 1962.

[15] L. CESARI, *Existence theorems for optimal solutions in Lagrange and Pontryagin problems*, SIAM J. Control, 3 (1965), pp. 475–498.

[16] K. ITÔ, *Stochastic differential equations on a differentiable manifold*, Nagoya Math. J., 1 (1950), pp. 35–47.

[17] R. W. BROCKETT AND J. C. WILLEMS, *Average value criteria for stochastic stability*, Stability of Stochastic Dynamical Systems, Springer Verlag Lecture Notes on Mathematics, vol. 294, 1972.