5-1976

Interpretation and Application of Elements of Differential Geometry and Lie Theory

James R. Brannan
Utah State University

INTERPRETATION AND APPLICATION OF ELEMENTS OF DIFFERENTIAL GEOMETRY AND LIE THEORY

by

James R. Brannan

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Mathematics

Plan A

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1976

ACKNOWLEDGMENTS

I thank the members of my graduate committee for their criticism and advice on the writing of this thesis. In particular, I would like to thank my major professor, Dr. Clyde Martin, for his guidance and encouragement during the preparation of the thesis.

And to my wife, Cheryl, I express my gratitude for the preliminary typings of the manuscript and for enduring my absence with understanding.

James R. Brannan

TABLE OF CONTENTS

ABSTRACT

Interpretation and Application of Elements of

Differential Geometry and Lie Theory

by

James R. Brannan, Master of Science

Utah State University, 1976

Major Professor: Dr. Clyde Martin
Department: Mathematics

Basic concepts of differential geometry and Lie theory are introduced. Lie transformation groups are applied to linear systems of differential equations and the problem of describing rigid body orientation. Linear Hamiltonian systems are then treated as a Lie system of differential equations. This theory is applied to a particular Hamiltonian system arising from a problem in control theory, the linear state regulator problem.

(40 pages)

CHAPTER I

INTRODUCTION

The objective of this thesis is to extract viable concepts from differential geometry and Lie theory which will be of use in the treatment of real problems. The main topics considered are the differentiable manifold, Lie transformation groups, Hamiltonian systems, and the linear state-regulator problem. It is the manifold construct which links these topics.

The second chapter develops the idea of differentiable manifolds. Then, by attaching vector spaces to each point in the manifold, new manifolds, called vector bundles, are constructed. This allows one to consider cross-sections, which are maps from the original manifold to the vector bundle. Specific examples of cross-sections that will be introduced are vectorfields, covectorfields, and more generally, tensorfields. Emphasis is given to local coordinate representations of tensorfields in order to develop some familiarity in working with these functions which are widely used in physics and continuum mechanics.

In chapter three a group structure is assigned to manifold point sets. When the group elements are associated with transformations which act on a space in a continuous way, the group-manifold structure becomes a Lie transformation group. While matrix Lie groups are the primary consideration, their connection with linear first order systems of differential equations is also mentioned. A specific example of a matrix Lie group is presented and applied to the problem of rigid body orientation.

Linear Hamiltonian systems of differential equations are considered in chapter four. This is a Lie system of differential equations which evolves in a manifold where their form never changes, the symplectic manifold. These equations define a bilinear form which determines the Lie transformation group of admissible curvilinear coordinate transformations which connect the local regions of the symplectic manifold.

Chapter five treats a particular linear Hamiltonian system which arises from a problem in control theory, the linear state-regulator problem. The system is treated as a Lie system of differential equations. This leads to expression of the linear transformation between the state vector and costate vector in terms of a generalized linear fractional transformation.

CHAPTER II

GENERAL STRUCTURES AND OBJECTS OF DIFFERENTIAL GEOMETRY

The principal object of investigation in differential geometry is the n-dimensional differentiable manifold. This is not necessarily a Euclidean n-space but for an observer in the manifold there is a small region about his position that appears to be a part of R^n. Differentiable manifolds are said to be locally Euclidean.

Some preliminary definitions are required in order to define a differentiable manifold. Let S be an open subset of R^n. A function $f:S \to R^n$ is said to be of <u>class C^k</u> iff each component function of f has continuous partial derivatives of all orders $r \leq k$. The function $f:S \to R^n$ is said to be of <u>class C^∞</u> iff it is of class C^k for every positive integer k. The function $f:S \to R^n$ is of <u>class C^ω</u> iff each of its component functions is analytic. A map $f:S \to T$ where S and T are open subsets of R^n is a <u>C^r diffeomorphism</u> iff f is of class C^r, is a bijection and f^{-1} is also of class C^r.

A <u>C^k n-dimensional differentiable manifold</u> consists of a topological space M together with a countable collection of open sets $U_1, U_2, \ldots$ such that each point of M lies in at least one of these U_i. Associated with each U_i is a homeomorphism f of U_i onto an open subset of R^n such that if $U_i \cap U_j \neq \emptyset$, then

$$f_i \circ f_j^{-1} : f_j(U_i \cap U_j) \rightarrow f_i(U_i \cap U_j)$$

is a C^k diffeomorphism. The ordered pairs (U_i, f_i) are called charts. If (U,f) is a chart containing the point m, then the local coordinates of m are given by $f(m) = (x_1(m), \ldots, x_n(m))$. Suppose that U with coordinate system $x_1, \ldots, x_n$ and V with coordinate system $y_1, \ldots, y_n$ overlap. Then the map which relates the coordinate systems

$$x_i = x_i(y_1, \ldots, y_n) \qquad i = 1, \ldots, n$$

is a C^k diffeomorphism called a curvilinear coordinate transformation. From now on, the capital letter M will be used to denote a C^∞– differentiable n-manifold. Although locally M appears to be a part of R^n, it is not a vector space in general because there may not be closure under addition or scalar multiplication of ordered sets of numbers, even if such operations can be defined.

The definition of a C^k map $g:M \rightarrow N$ where M and N are manifolds will be required later. The map g is of class C^k iff for each $m \in M$ and admissible chart (V, h) of N with $g(m) \in V$, there is a chart (U, f) of M with $m \in U$ and $g(U) \subset V$ and the local representative of g, $h \circ g \circ f^{-1}$, is of class C^k.

Historically, differential geometry began as the study of properties of curves and surfaces imbedded in 3-dimensional Euclidean space. An example of a surface is the unit 2-sphere defined by

$$x^2 + y^2 + z^2 - 1 = 0.$$

In such a case the geometric properties of the manifold can be studied extrinsically, with the points of the manifold being located by coordinates of the imbedding space. More generally, an n-manifold in R^{n+k} may be represented implicitly as the inverse image set, $F^{-1}(0)$, of $F:R^{n+k} \rightarrow R^k$:

$$F_i(x_1, \ldots, x_{n+k}) = 0 \qquad i=1, \ldots, k.$$

The Jacobian matrix of F is required to have rank k. Alternatively, an n-manifold in R^m may be represented by the imbedding $f:R^n \rightarrow R^m, n \leq m$:

$$f_i = f_i(U_1, \ldots, U_n) \qquad i=1, \ldots, m.$$

The rank of the Jacobian matrix of f is required to have rank n. The definition of a differentiable manifold given in this thesis is independent of any imbedding. The geometry which concerns itself with the study of properties determined entirely within the manifold is called intrinsic geometry. A beautiful example of application of the theory of intrinsic differential geometry is to general relativity. Both the intrinsic and extrinsic viewpoints are important and whenever possible it is usually advantageous to picture the manifold as being imbedded in a higher dimensional Euclidean space. A theorem from dimension theory states that every n-manifold may be imbedded as a closed subset of R^m for some $m \leq 2n+1$ [9]. For more information on intrinsic geometry see [10].

A linear vector space will now be attached to $m \in M$. Let

$$c(t) = (x_1(t), \ldots, x_n(t))$$

be a curve in local coordinates such that $c(0) = x(m)$. The tangent vector at m is the ordered set of first derivatives $(\dot{x}_1(0), \ldots, \dot{x}_n(0))$ evaluated at $t=0$. Every differentiable curve through m defines a vector and conversely every ordered n-tuple is the tangent vector to some curve. The totality of these vectors form a vector space over the field of real numbers. However, a vector space isomorphic to this one is needed. To each tangent vector $(a_1, \ldots, a_n)$ at $m \in M$, associate the partial derivative operator

$$a_1 \frac{\partial}{\partial x_1} + \ldots + a_n \frac{\partial}{\partial x_n} .$$

The operators $\left\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right\}$ will be considered as the basis of this space called the <u>tangent space</u> of M at m, denoted by T(M, m). If (U, x) is the chart containing m, let F(U) denote the set of real-valued C^{∞} functions on U. If $L \in T(M,m)$ and $f \in F(U)$ then L(f)(m) is often called the derivative of f in the direction L.

Consider the union $TM = \bigcup_{m \in M} T(M,m)$ of the tangent spaces to M at all points $m \in M$. The set TM has the natural structure of a C^{∞} 2n-manifold and is called the <u>tangent bundle</u> of M. Suppose U is a neighborhood of m with local coordinates $x_1, \ldots, x_n$, and a is a tangent vector at m with components $(a_1, \ldots, a_n)$. Then the tangent vector a has local coordinates in a neighborhood TU given by $(x_1, \ldots, x_n, a_1, \ldots, a_n)$. Let $y_i = y_i(x_1, \ldots, x_n)$ be a curvilinear coordinate transformation between neighborhoods and suppose

$$\sum_{i=1}^{n} a_i \frac{\partial}{\partial x_i} \quad \text{and} \quad \sum_{j=1}^{n} b_j \frac{\partial}{\partial y_j}$$

represent the same tangent vector at m. Since

$$b_j = \frac{d}{dt} y_j(c(t))$$

for some curve c(t) passing through m, we have

$$b_j = \sum_{i=1}^{n} \frac{\partial y_j}{\partial x_i} \bigg|_m \frac{d}{dt} x_i(c(t))$$

which implies that

$$b_j = \sum_{i=1}^{n} \frac{\partial y_j}{\partial x_i}\Big|_m a_i,$$

the transformation law for contravariant vectors [5].

Consider the map $\pi: TM \to M$ such that $\pi(m, a)$ is the point m at which a is tangent to M. The preimages of the points $m \in M$ under π are called fibres of the bundle TM. M is called the base space of the bundle TM. Each fibre has the structure of a vector space. By a vector field X on M is meant a C^∞ mapping $X: M \to TM$ such that the mapping $\pi \circ X: M \to M$ is the identity mapping. A vector field X is merely an assignment of a tangent vector to each point $m \in M$. The general form of X in local coordinates is

$$a_1(x_1, \ldots, x_n) \frac{\partial}{\partial x_1} + \ldots + a_n(x_1, \ldots, x_n) \frac{\partial}{\partial x_n} .$$

The space of all vector fields on M will be denoted by X(M).

Let $f \in F(U)$. The differential of f at m will be defined as a linear mapping of the tangent space T(M, m) into R.

For $L \in T(M, m)$, $df(L) = L(f)$. If $f_1, f_2 \in F(U)$, then

$$d(c_1 f_1 + c_2 f_2)(L) = c_1 df_1(L) + c_2\, df_2(L)$$

implies that the differentials df at m of $f \in F(U)$ form a linear subspace of all linear functions on $T(M,m)$. Let $x_1, \ldots, x_n$ be a local coordinate system at m. Each x_i is a map from U into R and the set $\langle dx_1, \ldots, dx_n \rangle$ forms a basis for the space of linear functionals on $T(M,m)$. This space will be denoted by $T^*(M,m)$ and is called the cotangent space of M at m. It is clearly dual to $T(M,m)$. Let $y_i = y_i(x_1, \ldots, x_n)$ be a curvilinear coordinate transformation between neighborhoods of m. If

$$\sum_{i=1}^{n} a_i dx_i \text{ and } \sum_{j=1}^{n} b_j dy_j$$

represent the same cotangent vector, then

$$\sum_{j=1}^{n} b_j dy_j = \sum_{j=1}^{n} \sum_{i=1}^{n} b_j \left.\frac{\partial y_j}{\partial x_j}\right|_m dx_i$$

implies that the components are related by the linear transformation

$$a_i = \sum_{j=1}^{n} \left.\frac{\partial y_j}{\partial x_i}\right|_m b_j,$$

the transformation law for covariant vectors [5].

The union $T^*M = \bigcup_{m \in M} T^*(M,m)$ is called the cotangent bundle of M and is a C^∞ 2n-manifold in the same way that the tangent bundle of M is a C^∞ 2n-manifold. By a covectorfield X^* on M is meant a C^∞ mapping $X^*: M \to T^*M$ such that the mapping $\pi \circ X^*: M \to M$ is the identity. A covectorfield merely assigns a cotangent vector to each point $m \in M$. The general form of X^* in local coordinates is

$$X^* = b_1(x_1, \ldots, x_n)dx_1 + \ldots + b_n(x_1, \ldots, x_n)dx_n.$$

The space of all covectorfields is denoted by $X^*(M)$.

TM and T^*M are special cases of a more general structure called a vector bundle [1]. Intuitively, a vector bundle may be thought of as a manifold with a vector space attached to each point. More precisely, a vector bundle over M is a C^∞ map $\pi: E \to M$ of an (n+k)-manifold E onto M such that for each $m \in M$ the fibre above m, $\pi^{-1}(m) \subset E$ is a k-dimensional real vector space. A C^∞ cross-section is a C^∞ map $\Psi: M \to E$ such that $\pi \circ \Psi(m) = m$ for each $m \in M$. The set of cross sections is denoted by $\Gamma(E)$. Two cross sections Ψ_1 and Ψ_2 can be added at each $m \in M$ since $\Psi_1(m)$ and $\Psi_2(m)$ lie in the same vector space. Also, $\Psi \in \Gamma(E)$ can be multiplied by $f \in F(M)$:

$$f\Psi(m) = f(m)\Psi(m).$$

It should be pointed out that cross-sections are globally defined. Local coordinates only provide a local representation of the cross-section and the real-valued functions are elements of the set F(U) where U is a coordinate neighborhood. To patch these neighborhoods together requires knowledge of the interconnecting curvilinear coordinate transformations.

The concept of vector bundles and cross sections allows us to assign more complex objects than just tangent vectors and cotangent vectors to points $m \in M$. Denote by $T^r_s(M)$ the space of multi-linear maps of the fibres of $T^*M \times \ldots \times T^*M \times TM \times \ldots \times TM$ (r copies of T^*M and s copies of TM) into R. $T^r_s(M)$ is called the vector bundle of tensors of contravariant order r and covariant order s, or simply of type $\binom{r}{s}$. Clearly $T^0_1(M) = T^*M$ and $T^1_0(M)$ may be identified with TM. A tensorfield of type $\binom{r}{s}$ on M is a C^∞ cross-section of $T^r_s(M)$. The set of all C^∞ cross-sections of $T^r_s(M)$ will be denoted by $\Gamma^r_s(M)$. Then

$$X(M) = \Gamma^1_0(M) \text{ and } X^*(M) = \Gamma^0_1(M).$$

It is beneficial to observe the form of these cross-sections in local coordinates. Take, for example, a local tensorfield of type $\binom{0}{2}$, that is, a second rank covariant tensorfield. Such an object can be created by forming the tensor product, or direct product, of two local covariant vector fields [6]. Let $x = (x_1, \ldots, x_n)$. If $\omega = \sum_{i=1}^{n} a_i(x)\, dx_i$ and $\Theta = \sum_{j=1}^{n} b_j(x)\, dx_j$, then their tensor product is

$$\alpha = \omega \otimes \Theta = \sum_{j=1}^{n} \sum_{i=1}^{n} a_i b_j\, dx_i \otimes dx_j = \sum_{j=1}^{n} \sum_{i=1}^{n} C_{ij}(x)\, dx_i \otimes dx_j.$$

The symbol $\otimes$ is the tensor product symbol and $dx_i \otimes dx_j$ is a basis element on the fibers of TU X TU. α is a bilinear map from TU X TU into R. Let

$$X = \sum_{k=1}^{n} e_k \frac{\partial}{\partial x_k} \qquad \text{and} \qquad Y = \sum_{m=1}^{n} f_m \frac{\partial}{\partial x_m}$$

be vectorfields. Then

$$\alpha(X, Y) = \omega(X) \cdot \Theta(Y) = \left(\sum_{i=1}^{n} \sum_{k=1}^{n} a_i e_k \delta_{ik}\right) \cdot \left(\sum_{j=1}^{n} \sum_{m=1}^{n} b_j f_m \delta_{jm}\right)$$

$$= \left(\sum_{i=1}^{n} a_i e_i\right)\left(\sum_{j=1}^{n} b_j f_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} e_i a_i b_j f_j = \sum_{i=1}^{n} \sum_{j=1}^{n} e_i C_{ij} f_j .$$

All this can be represented as a matrix operation:

$$\alpha(X, Y) = [e_1, \ldots, e_n] \begin{bmatrix} C_{11} & \cdot & \cdot & \cdot & C_{1n} \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ C_{n1} & \cdot & \cdot & \cdot & C_{nn} \end{bmatrix} \begin{bmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_n \end{bmatrix}$$

A particular class of covariant tensors has been found to be very important. These are the covariant tensors which are antisymmetric under exchange of any pair of indices. The formalism developed for these tensorfields is called the theory of differential forms [5]. Continuing with the example, suppose

$$\alpha = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij}(x)\, dx_i \otimes dx_j$$

is antisymmetric. Under the formalism α is called a 2-form and represented as

$$\alpha = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij}(x)\, dx_i \wedge dx_j$$

where the symbol $\wedge$ is called the wedge or exterior product. The fact that $c_{ij} = -c_{ij}$ motivates the following rules:

$$dx_i \wedge dx_i = 0$$

$$dx_i \wedge dx_j = - dx_i \wedge dx_j. \qquad (2\text{-}1)$$

If β is a 1-form, $\beta = \sum_{k=1}^{n} b_k(x)\, dx$, then the wedge product of α and β is the 3-form

$$\alpha \wedge \beta = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} C_{ij}(x)\, b_k(x)\, (dx_i \wedge dx_j) \wedge dx_k$$

which can be simplified by using rules (2-1) and the associativity rule:

$$(dx_i \wedge dx_j) \wedge dx_k = dx_i \wedge (dx_j \wedge dx_k) \quad dx_i \wedge dx_j \wedge dx_k .$$

A 3-form called the <u>exterior derivative</u> of α, denoted $d\alpha$, can be constructed as follows. Each coefficient $c_{ij}(x)$ of α is an element of F(U), that is, a 0-form. Then the differential of c_{ij}, dc_{ij} is a 1-form. The exterior derivative of α is defined as

$$d\alpha = \sum_{i=1}^{n} \sum_{j=1}^{n} dc_{ij} \wedge (dx_i \wedge dx_j).$$

For more details and generalization of the algebra of tensors and differential forms see [1] or [5].

CHAPTER III

LIE GROUPS OF TRANSFORMATIONS

The theory of Lie groups and Lie algebras is an area where modern algebra, classical analysis, differential geometry, and topology interact to give the user a powerful mathematical structure with which to work. The Lie theory has been applied to such areas as differential equations, special functions, perturbation theory, continuum mechanics, and control theory [3]. Gilmore [6] implies that the Lie theory may serve as a tool for studying the overall structure of dynamical systems. In that capacity, the theory is in an embryonic stage. A fairly complete bibliography on theory and application of Lie groups and Lie algebras may be found in [3] and [6]. This chapter is concerned with defining some basic elements of Lie theory, implicating the relationship with differential equations, and considering a specific example of a Lie group.

A Lie group consists of an analytic manifold G which has a group structure

$$(x, y) \rightarrow xy = z$$

with the group operation being analytic. Each element of the group is specified by its local coordinates. Let the coordinates in a neighborhood of the identity be chosen so that the coordinates of the identity are zero. Then, if

$$x = (x_1, \ldots, x_n) \text{ and } y = (y_1, \ldots, y_n), z_i(x_1, \ldots, x_n, y_1, \ldots y_n)$$

can be expanded in a convergent Taylor series about the origin:

$$z_i(x, y) = z_i(0, 0) + \sum_{j=1}^{n} \left(\frac{\partial z_i(0, 0)}{\partial x_j} x_j + \frac{\partial z_i(0, 0)}{\partial y_j} y_j \right)$$

$$+ \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \left(\frac{\partial^2 z_i(0, 0)}{\partial x_j \, \partial x_k} x_j x_k + 2 \frac{\partial^2 z_i(0, 0)}{\partial x_j \partial y_k} x_j y_k + \frac{\partial^2 z_i(0, 0)}{\partial y_j \, \partial y_k} y_j y_k + \ldots \right.$$

Because G is a group, all the manifold charts can be generated from the charts at the identity element. If U is a neighborhood of $g \in G$ then

$$g^{-1} U = \{ g^{-1} h \mid h \in U \}$$

contains a coordinate neighborhood of the identity. For this reason, it suffices to study Lie groups in a neighborhood of the identity.

Given a group G and a space M, the <u>action</u> of G on M is a function assigning to each element g of G, a continuous map

$$f_g : M \to M$$

so that

(1) if e is the identity element of G, f_e is the identity map of M,

(2) if g = hk, then $f_g = f_h \circ f_k$.

If G is a Lie group which acts on a space M according to this definition then G is called a <u>Lie group of transformations</u>.

For simplicity, elements of G will now be identified with their image under the action function. Given an action of G on M, a <u>flow</u> on the space M (relative to G) is a curve $t \to g(t)$ in G such that $g(0) = e$. An <u>orbit</u> or <u>path</u> of the flow is a curve x(t) in M of the form

$$x(t) = g(t)x_0.$$

The discussion will now be restricted to matrix Lie groups. Let M(n;R) denote the set of real-valued nxn matrices. A <u>Lie algebra</u> L in M(n;R) is a subspace of M(n;R) with a multiplication operation defined for $B, C \in L$ by

$$[B, C] = BC - CB.$$

This is called the <u>Lie bracket</u> of B and C. The Lie bracket is skew-symmetric

$$[B, C] = -[C, B]$$

and satisfies the <u>Jacobi identity</u>

$$[B, [C,D]] + [C, [D,B]] + [D,[B,C]] = 0.$$

If S is a subset of M(n;R), the Lie algebra generated by S, denoted $\langle S \rangle_A$, is the smallest Lie algebra containing S. It is generated with the Lie bracket operation. A matrix group is a subset of M(n;R) that is a group under multiplication. Let exp:M(n;R)→M(n;R) denote the matrix exponential map

$$\exp(B) = \sum_{n=0}^{\infty} \frac{B^n}{n!} .$$

A matrix group G is a matrix Lie group if for some Lie algebra L

$$G = \langle \exp(L) \rangle_G,$$

that is, G is the group generated by exp (L) under matrix multiplication [11]. To see that matrix groups are analytic manifolds, consider GL(n;R), the group of nonsingular nxn real matrices. Each element of GL(n;R) can be considered a point in a Euclidean space of dimension n^2. Each point in this space lies in an open set contained in the space since the determinant function is continuous in the coordinates of the space, i.e., all points in a neighborhood of a point representing a nonsingular matrix also represent nonsingular matrices. Euclidean coordinates serve as curvilinear coordinates for the group. Since the group operation is matrix multiplication, the coordinates of the product of two matrices are polynomials in the coordinates of the two factors, hence the group operation is analytic. All the classical matrix Lie groups are subgroups of the complex general linear group GL(n;C) and can be represented as hypersurfaces in Euclidean space. Whereas the exponential map is an algebraic relationship between the group and its algebra, in terms of differential geometry it is the tangent space to the group manifold at the identity, T(G, e), that corresponds to the Lie algebra.

A flow $t \to A(t)$ in GL(n;R) is called a linear flow.

The matrix

$$B(t) = \frac{dA}{dt}(t)\, A^{-1}(t)$$

defines a curve in M(n;R) called the infinitesimal generator of the flow A(t). Now consider the orbit of a flow in a topological space M:

$$x(t) = A(t)x(0).$$

Differentiating obtains

$$\frac{dx}{dt}(t) = \frac{dA}{dt}(t)\, x(0) = B(t)\, A(t)\, x(0) = B(t)\, x(t)$$

and shows the relationship of a system of linear differential equations to the infinitesimal generator of a flow. If a flow $t \to A(t)$ in GL(n;R) satisfies

$$A(t_1 + t_2) = A(t_1)\, A(t_2)$$

and

$$A(0) = e$$

then it is a one-parameter subgroup of GL(n;R). In such a case the infinitesimal generator is constant. For a linear autonomous system of first order differential equations,

$$\mathring{x}(t) = Bx(t),$$

the solution is given by an orbit of the flow $t \rightarrow \exp(Bt) = A(t)$,

$$x(t) = A(t)x_0.$$

A specific example of a matrix Lie group will now be considered, the special orthogonal group SO(3;R). The matrices $A \in SO(3;R)$ are characterized by

$$\det A = 1$$

$$A^*A = I$$

The action of this group on R^3 leaves distances fixed, hence SO(3;R) is often called the 3-dimensional rotation group. A representation of SO(3;R) is given by the product ABC where

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} \quad B = \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \quad C = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Each of the matrices A, B, and C correspond to a one-parameter subgroup of SO(3;R). Differentiating each curve with respect to its parameter and evaluating at 0 gives the following basis for the tangent space of SO(3;R):

$$e_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \qquad e_2 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad e_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Any element of the Lie algebra is a linear combination of e_1, e_2, and e_3. The effect of an element of the Lie algebra of SO(3;R), denoted by so(3, R), is to assign a vector at each point of R^3 which points in the direction the point is being rotated under the action of the associated group element. Thus, a vectorfield is defined on R^3. The Lie bracket operation gives

$$[e_2, e_1] = e_3 \qquad [e_3, e_2] = e_1 \qquad [e_1, e_3] = e_2.$$

Notice that the motion of points in R^3 is symmetrical about the origin under the action of $A \in$ SO(3;R). For this reason, SO(3;R) is particularly well suited for description of motions on the 2-manifold S^2, i.e., the unit two-sphere defined by $x^2 + y^2 + z^2 = 1$. SO(3;R) is said to act <u>transitively</u> on S^2 since the orbit of a point in S^2 is the entire space. In such a case, S^2 is called <u>homogeneous</u> with respect to SO(3;R) and can be identified with the underlying manifold of G, since the three parameters α, β, γ can be used to unambiguously specify any point of

S^2. An example of application of this group is to the differential equations describing the orientation of a rigid body relative to a fixed set of axes. The system may be thought of as evolving on SO(3;R). The differential equation for such a system is given by

$$\dot{A}(t) = \left(\sum_{i=1}^{3} \omega_i(t)\, e_i\right) A(t) \qquad A(0) = I$$

where $A(t) \in SO(3;R)$ and the ω_i are angular velocities. It has been shown that there exists a time interval $[0, T]$ and real functions $h_1(t)$, $h_2(t)$, $h_3(t)$ such that

$$A(t) = \exp[h_1(t)\, e_1] \exp[h_2(t)\, e_2] \exp[h_3(t)\, e_3]$$

for each $t \in [0, T]$. See [11]. For information concerning controllability and observability of this system see [4] and [7].

CHAPTER IV

SYMPLECTIC MANIFOLDS AND HAMILTONIAN SYSTEMS

In classical mechanics the Lagrangian of a conservative mechanical system is a function of the generalized position coordinates $x_1, \ldots, x_n$ and their time derivatives $\dot{x}_1, \ldots, \dot{x}_n$. It is defined as

$$L = T - V$$

where T is the kinetic energy of the system and V is the potential energy of the system. The Hamiltonian function, H, is defined in terms of the Lagrangian as

$$H = \sum_{i=1}^{n} p_i \dot{x}_i - L \qquad (4\text{-}1)$$

and must be expressed in terms of the generalized coordinates $x_1, \ldots, x_n$ and the generalized momenta $p_1, \ldots, p_n$ defined by

$$p_i = \frac{\partial T}{\partial \dot{x}_i} .$$

In order to simplify the discussion, only autonomous systems will be considered.

In equation (4-1), for each t the vector $(\dot{x}_1(t), \ldots, \dot{x}_n(t))$ is a tangent vector to a curve in a configuration space, M, which is assumed to be a differentiable manifold. The vector $(p_1(t), \ldots, p_n(t))$ may be thought of as a vector dual to $(\dot{x}_1(t), \ldots, \dot{x}_n(t))$ since it maps it into the real numbers. Then (x, p) can be considered a local coordinate system of the cotangent bundle T^*M with H(x, p) an element of F(U). The equations

$$\dot{x}_i(t) = \frac{\partial H}{\partial p_i} \qquad i = 1, \ldots, n$$

$$\dot{p}_i(t) = \frac{-\partial H}{\partial x_i} \qquad i = 1, \ldots, n \tag{4-2}$$

are called Hamilton's equations, a local system of first order ordinary differential equations which evolve in T^*M. Given an initial value (x(0), p(0), equations (4-2) define a curve (x(t), p(t)) in $U \subset T^*M$. By a solution of (4-2) is meant the projection of this curve down to a region of the base space M. A method of handling this projection will be considered in the next chapter.

Hamilton's equations may be written more suggestively as

$$\begin{bmatrix} \dot{x}(t) \\ \dot{p}(t) \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \nabla H \tag{4-3}$$

where

$$\nabla H = \left[\frac{\partial H}{\partial x_1}, \ldots, \frac{\partial H}{\partial x_n}, \frac{\partial H}{\partial p_1}, \ldots \frac{\partial H}{\partial p_n}\right]^* .$$

The matrix

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

defines an antisymmetric bilinear form on the fibers of $T(T^*M) X T(T^*M)$ called a symplectic form. ∇H is a local representation of a covariant vectorfield and $(\dot{x}(t), \dot{p}(t))$, as tangent vectors to a family of curves in T^*M, is a local contravariant vectorfield. Thus, for each $(x, p) \in U \subset T^*M$ Hamilton's equations express a canonical relationship between a vector in $T(U, (x, p))$ and a covector in $T^*(U, (x, p))$.

It is natural to ask what curvilinear coordinate transformations leave the form of equations (4-2) invariant. Suppose $f: U \subset R^{2n} \rightarrow V \subset R^{2n}$ given by

$$y_i = y_i(x_1, \ldots, x_n, p_1, \ldots, p_n) \quad i=1, \ldots, n$$

$$s_i = s_i(x_1, \ldots, x_n, p_1, \ldots, p_n) \quad i=1, \ldots, n$$

is such a transformation. It will be beneficial to pause and consider again the transformation laws concerning contravariant and covariant vectorfields. The map f induces a map on the local tangent bundle and local cotangent bundle of the 2n-manifold T^*M.

The induced map is the Jacobian of f, Df.

$$T(T^*M, (y, s)) \xleftarrow{\;Df^*\big|_{(x,p)}\;} T(T^*M, (x, p))$$

$$T^*(T^*M, (y, s)) \xrightarrow{\;Df\big|_{(x,p)}\;} T^*(T^*M, (x, p)).$$

Since $T^*(T^*M, (x, p))$ is dual to $T(T^*M, (x, p))$, $Df^*\big|_{(x, p)}$ maps opposite to $Df\big|_{(x, p)}$. Hence, if $H_1(y, s)$ is the Hamiltonian in the new coordinate system, and ∇H_1 transforms according to

$$\nabla H(x, p) = Df\big|_{(x, p)} \nabla H_1(y, s) \qquad (4\text{-}4)$$

the left hand side of (4-3) must transform according to

$$\begin{bmatrix} \dot{x} \\ \\ \dot{p} \end{bmatrix} = \begin{bmatrix} Df^* & \\ & (x, p) \end{bmatrix}^{-1} \begin{bmatrix} \mathring{y} \\ \\ \dot{s} \end{bmatrix} \qquad (4\text{-}5)$$

Substitution of (4-4) and (4-5) into (4-3) gives

$$\begin{bmatrix} \mathring{y} \\ \\ \mathring{u} \end{bmatrix} = Df^*\big|_{(x, p)} J \;\; Df\big|_{(x, p)} \nabla H_1(y, u). \qquad (4\text{-}6)$$

The form of Hamilton's equations will remain invariant iff given a curvilinear coordinate transformation f,

$$Df^* J \; Df = J. \tag{4-7}$$

The set of all curvilinear coordinate maps satisfying (4-7) form a Lie group called the <u>symplectic group</u> [1]. Such transformations are called <u>homogeneous canonical</u> or <u>contact transformations.</u>

The condition (4-7) is identical to the Lagrange bracket conditions

$$[x_j, p_k] = \delta_{jk} \tag{4-8}$$

$$[x_j, x_k] = 0 \tag{4-9}$$

$$[p_j, p_k] = 0 \tag{4-10}$$

where

$$[x_j, p_k] = \sum_{i=1}^{n} \left(\frac{\partial y_i}{\partial x_j} \frac{\partial u_i}{\partial p_k} - \frac{\partial y_i}{\partial p_j} \frac{\partial u_i}{\partial x_k} \right)$$

with similar definitions for (4-9) and (4-10). The manifold T^*M with this sympletic form is called a <u>symplectic manifold.</u> The symplectic group provides an elegant method of discussing all admissible curvilinear coordinate transformations for this manifold. It is this group which ties the manifold together.

If f happens to be linear, then Df = f. This space of linear symplectic maps is a subgroup of the symplectic group and is classically denoted by Sp(2n;R). It is called the linear symplectic group.

In the case that Hamilton's equations are linear, they can be written in the form

$$\begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix} . \qquad (4\text{-}11)$$

This is possible iff H(x, p) is quadratic in x_i and p_j. It can be shown that the matrix

$$B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$$

must satisfy

$$B^* J + J B = 0 \qquad (4\text{-}12)$$

which implies

$$B_1 = -B_4^*$$

$$B_2 = B_2^*$$

$$B_3 = B_3^*$$

The set of all linear maps $B:R^{2n} \to R^{2n}$ which satisfy (4-12) are called Hamiltonian matrices. It can also be shown that

$$(\exp B)^{*} \; J \; (\exp B) = J,$$

that is, the exponential map associates B with some element of Sp(2n;R). When the set of matrices satisfying (4-12) is equipped with the product

$$[B, C] = BC - CB$$

it becomes a Lie algebra, denoted by sp(2n;R), called the symplectic algebra.

For linear Hamiltonian systems, the Hamiltonian matrix is the infinitesimal generator of a flow $t \to A(t)$ whose orbits are the level sets of the Hamiltonian H in phase space. However, to solve the linear Hamiltonian system, one needs to know the specific relationship between the state vector x(t) and its dual state vector p(t). This will be taken up in the next section.

CHAPTER V

THE LINEAR REGULATOR AND THE SYMPLECTIC GROUP

This chapter treats a linear autonomous Hamiltonian system arising from a problem in optimal control theory. In this case the symplectic manifold is R^{2n}. A general method of finding the map which relates the orbit of the system in the tangent bundle to the orbit in the base space is derived in terms of a generalized linear fractional transformation. An alternative method of obtaining this map results in a matrix Ricatti system of differential equations.

It is desired to find the control function u(t) which minimizes the functional

$$J(u) = \frac{1}{2}\int_0^T (x^*(t)\, Qx(t) + u^*(t)Ru(t))\, dt$$

subject to the linear autonomous system constraint

$$\dot{x}(t) = Fx(t) + Gu(t)$$

with the arbitrary initial condition $x(0) = x_0$. Q is a positive definite nxn matrix and R is a positive definite mxm matrix. F is an nxn matrix and G is an nxm matrix. In physical terms this may be interpreted as finding the control which keeps the state x(t) near zero with minimum energy expenditure.

It is a result of optimal control theory [2] that the problem may be reformulated as a Hamiltonian system. The Hamiltonian function H(x, p, u) is given by

$$H(x, p, u) = \tfrac{1}{2}\,(x^{*}Qx + u^{*}Ru) + p^{*}Fx + p^{*}Gu)$$

where p(t) is the costate n-vector associated with x(t). The extremal path in state space is the solution to Hamilton's equations:

$$\begin{bmatrix} \dot{x}(t) \\ \dot{p}(t) \end{bmatrix} = \begin{bmatrix} F & -GR^{-1}G^{*} \\ Q & -F^{*} \end{bmatrix} = \begin{bmatrix} x \\ p \end{bmatrix} \qquad (5\text{-}1)$$

This is a linear autonomous system of 2n differential equations. The initial state x_0 furnishes n boundary conditions and the remaining n boundary conditions are given by p(T) = 0. It is also a fact that p and x are related by an equation of the form

$$p(t) = K(t)\,x(t) \qquad (5\text{-}2)$$

for all $t \in [0, T]$. See [2].

Let W_1 be the subspace of R^{2n} spanned by the standard basis vectors $\langle e_1, \ldots, e_n \rangle$ and let W_2 be the subspace of R^{2n} spanned by the standard basis vectors $\langle e_{n+1}, \ldots, e_{2n} \rangle$. Denote by $L(W_1, W_2)$ the space of linear maps of W_1 into W_2. Then $K(t) \in L(W_1, W_2)$ for every $t \in [0, T]$ and at time $t = 0$ the position of the system in phase space is a point $(x(0), p(0))$ in the subspace

$$S_0 = \langle (x, p) \mid p = K(0)s, \ x \in W_1 \rangle$$

Let

$$B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$$

be the Hamiltonian state matrix of (5-1). Let $t \rightarrow A(t) = \exp(Bt)$ be a flow in $Sp(2n;R)$ and partition $A(t)$ into four nxn submatrices:

$$A = \begin{bmatrix} A_1(t) & A_2(t) \\ A_3(t) & A_4(t) \end{bmatrix} .$$

Consider the action of $A(t)$ on S_0. $A(t)$ must map S_0 to another subspace S_t,

$$S_t = \left\{ (x, p) \mid p = K(t)x, \ x \in W_1 \right\}.$$

At this time the position of the system in phase space is the point $(x(t), p(t))$ in S_t. Since $S_t = K(t)S_0$, it follows that

$$\begin{bmatrix} A_1(t) & A_2(t) \\ A_3(t) & A_4(t) \end{bmatrix} \begin{bmatrix} x(0) \\ K(0)\ x(0) \end{bmatrix} = \begin{bmatrix} x(t) \\ K(t)\ x(t) \end{bmatrix} .$$

Then

$$\begin{aligned} A_1(t)\, x(0) + A_2(t)\, K(0)\, x(0) &= x(t) \\ A_3(t)\, x(0) + A_4(t)\, K(0)\, x(0) &= K(t)\, x(t) \end{aligned} \qquad (5\text{-}3)$$

and

$$K(t) = [\, A_3(t) + A_4(t)\, K(0)]\ [\, A_1(t) + A_2(t)\, K(0)]^{-1} \qquad (5\text{-}4)$$

when this inverse exists. Equation (5-4) is a special example of a generalized linear fractional transformation [8]. The symplectic automorphism $A(t)$ thus induces an action on $L(W_1, W_2)$ as well as on R^{2n}. Equation (5-4) defines a flow $t \rightarrow K(t)$ acting on the costate space W_1.

Since $K(0)$ is unknown, an alternative method of solving for $K(t)$ must be found.

Consider equations (5-3). Differentiating these equations gives

$$\dot{x}(t) = (\dot{A}_1(t) + \dot{A}_2(t)\, K(0))\, x(0) \qquad (5\text{-}5)$$

$$\dot{K}(t)\, x(t) + K(t)\, \dot{x}(t) = (\dot{A}_3(t) + \dot{A}_4(t)\, K(0))\, x(0). \qquad (5\text{-}6)$$

Substitution of (5-5) into (5-6) obtains

$$\dot{K}(t)\, x(t) + K(t)\, (\dot{A}_1(t) + \dot{A}_2(t)\, K(0))x(0) = (\dot{A}_3(t) + \dot{A}_4(t)\, K(0))\, x(0). \quad (5\text{-}7)$$

But $BA = \dot{A}$ implies that

$$\begin{aligned} \dot{A}_1 &= B_1A_1 + B_2A_3 \\ \dot{A}_2 &= B_1A_2 + B_2A_4 \\ \dot{A}_3 &= B_3A_1 + B_4A_3 \\ \dot{A}_4 &= B_3A_2 + B_4A_4 \end{aligned} \qquad (5\text{-}8)$$

Substituting equations (5-8) into (5-7) and simplifying gives a differential equation that $K(t)$ must satisfy:

$$\mathring{K}(t) = B_3 + B_4 K(t) - K(t)(B_1 + B_2 K(t)). \qquad (5\text{-}9)$$

This is a <u>matrix Ricatti equation.</u> The right hand side generates the flow $t \rightarrow K(t)$ in $L(W_1, W_2)$. Since $K(T)$ is known, the solution to (5-9) exists and is unique [2].

Hamilton's equations (5-1) may now be written

$$\begin{bmatrix} \mathring{x}(t) \\ \mathring{p}(t) \end{bmatrix} = \begin{bmatrix} F & -GR^{-1}G^* \\ -Q & -F^* \end{bmatrix} \begin{bmatrix} x(t) \\ K(t)\ x(t) \end{bmatrix}$$

Comparing

$$\mathring{x}(t) = Fx(t) - GR^{-1}G^* K(t)\ x(t)$$

with the system constraint

$$\mathring{x}(t) = Fx(t) + Gu(t)$$

implies that

$$u(t) = -R^{-1}G^* K(t)\ x(t).$$

It is a fact of control theory that this is the unique optimal control [2].

BIBLIOGRAPHY

1. Abraham, Ralph, and Jerrold E. Marsden. 1967. Foundations of mechanics. W. A. Benjamin, Inc., New York, N.Y.

2. Athans, Michael, and Peter L. Falb. 1966. Optimal control. McGraw-Hill, Inc., New York, N.Y.

3. Belinfante, Johan G. F., and Bernard Kolman. 1972. A survey of Lie groups and Lie algebras with applications and computational methods. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

4. Brockett, R. W. System theory on group manifolds and coset spaces. Siam Journal on Control, Vol. 10, No. 2, May 1972, pp. 265-284.

5. Flanders, Harvey. 1963. Differential forms. Academic Press, New York and London.

6. Gilmore, Robert. 1974. Lie groups, Lie algebras, and some of their applications. John Wiley and Sons, Inc., New York, London, Sidney, and Toronto.

7. Jurdjevic, Velimir, and Hector J. Sussman. Control systems on Lie groups. Journal of Differential Equations. 12, pp. 313-329 (1972).

8. Hermann, Robert. 1973. Algebraic topics in systems theory. Rutger's University, New Brunswick, New Jersey.

9. Munkres, James R. 1975. Topology, a first course. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

10. Stoker, J. J. 1969. Differential geometry. Wiley-Interscience, New York, London, Sydney, and Toronto.

11. Willsky, Alan S. Some results on the estimation of the angular velocity and orientation of a rigid body. (Submitted for publication to Automatica. November 27, 1973).