JOURNAL OF DIFFERENTIAL EQUATIONS 55, 385–400 (1984)

Nilpotent Bases for Distributions and Control Systems

HENRY HERMES,* ALBERT LUNDELL, AND DENNIS SULLIVAN

Department of Mathematics, University of Colorado, Boulder, Colorado

Received February 11, 1983; revised July 18, 1983

Let $V(M)$ be the Lie algebra (infinite dimensional) of real analytic vector fields on the n-dimensional manifold M. Necessary conditions that a real analytic k-dimensional distibution on M have a local basis which generates a nilpotent subalgebra of $V(M)$ are derived. Two methods for sufficient conditions are given, the first depending on the existence of a solution to a system of partial differential equations, the second using Darboux's theorem to give a computable test for an $(n-1)$-dimensional distribution. A nonlinear control system in which the control variables appear linearly can be transformed into an orbit equivalent system whose describing vector fields generate a nilpotent algebra if the distribution generated by the original describing vector fields admits a nilpotent basis. When this is the case, local analysis of the control system is greatly simplified.

0. INTRODUCTION

Let M be a real analytic, n-dimensional manifold, $V(M)$ denote the real vector space of real analytic vector fields on M considered as a Lie algebra (infinite dimensional) under the Lie product $[X, Y]$, $X, Y \in V(M)$. If $Y^1,..., Y^k \in V(M)$ and are linearly independent at $p \in M$ then for $x \in M$ near p the map $x \to \operatorname{span}\{Y^1(x),..., Y^k(x)\}$ defines a k-dimensional distribution denoted D^k having $Y^1,..., Y^k$ as a basis. Our major concern is with the question of when D^k admits a basis which generates a nilpotent subalgebra.

In Section 1 we first derive properties of a distribution which are basis invariant. If a distribution admits a nilpotent basis the local action of the associated nilpotent Lie group forces homogeneity for a nbd of p, i.e., the invariants of $D^k(x)$ must be consistent with those of $D^k(p)$ for x near p. By violating this homogeneity (Example 1.1) for any integers $2 \leqslant k < n$ we construct a distribution D^k on M^n which locally admits no nilpotent basis.

Section 2 obtains several positive results. The methods utilized are, briefly, as follows. First, given a distribution $D^k(x) = \operatorname{span}\{Y^1(x),..., Y^k(x)\}$ locally on M we construct a distribution $\tilde{D}^k(s) = \operatorname{span}\{X^1(s),..., X^k(s)\}$ on $\mathbb{R}^n$ with

* This research was supported by NSF Grant MCS79-26316.

0022-0396/84 $3.00

the Lie algebra generated by $X^1,\ldots,X^k$, denoted $L(X^1,\ldots,X^k)$, nilpotent and such that the invariants of D^k and $\tilde{D}^k$ are compatible. We next attempt to construct a diffeomorphism ϕ from $\mathbb{R}^n$ to M such that the induced tangent space isomorphism ϕ_* carries $\tilde{D}^k(s)$ onto $D^k(\phi(s))$. This leads to a system of first-order partial differential equations which, if solvable, yield ϕ and the nilpotent basis $\phi_* X^1,\ldots,\phi_* X^k$ for D^k. This is analogous to the usual method of proof of the Frobenius theorem, i.e., if D^k is an involutive distribution one can lift the standard abelian basis for $\mathbb{R}^k$ to a basis for D^k via ϕ_*.

The second method is via the use of Darboux's theorem and yields a computable condition (Theorem 3) that an n-dimensional distribution $D^n(x) = \text{span}\{Y^1(x),\ldots, Y^n(x)\}$ on M^{n+1} admits a nilpotent basis. This is obtained by requiring that a nonzero one form ω "perpendicular" to the distribution have constant rank in an nbd of p, a condition which can be described in terms of a sufficient number (depending on the rank) of products $[Y^i, Y^j](p)$ being independent of $Y^1(p),\ldots, Y^n(p)$. We end this section with Example 2.3 of a two distribution on $\mathbb{R}^3$ to which the Darboux method cannot be applied but the differential equations of the first method can be solved to show the existence of a nilpotent basis.

The motivation for this work originated in feedback control problems for systems of the form

$$\dot{x}(t) = Y^0(x(t)) + \sum_{i=1}^{k} u_i Y^i(x(t)), \qquad x(0) = p \ (\dot{x} = dx/dt) \tag{0.1}$$

with $Y^0,\ldots, Y^k \in V(M)$ and the control components u_i having values in $\mathbb{R}^1$. If $L(Y^0,\ldots, Y^k)$ is nilpotent, system (0.1) lends itself nicely to analysis. Indeed, general results for vector field systems are often obtained by first deriving these for systems which generate nilpotent Lie algebras; the general result then following from a "nilpotent approximation." Examples occur in the study of a parametrix for hypoelliptic operators of the form $\sum_{i=0}^{k} (Y^i)^2$ and in local controllability for scalar input, control linear, systems of the form (0.1), see [1; 2].

Control systems which generate the same trajectories may have a variety of discriptions. One of the purposes of feedback is to obtain a "simple" description or representation. Specifically, let $G(x)$ denote the $n \times k$ matrix with columns $Y^1(x),\ldots, Y^k(x)$ and u the column vector $(u_1,\ldots, u_k)$. Rewrite system (0.1) as $\dot{x} = Y^0(x) + G(x)u$. If one admits feedback control, specifically $u(x) = h(x) + H(x)v$, where the column vector function $h(x) = (h_{10}(x),\ldots, h_{k0}(x))$ is arbitrary, $H(x) = (h_{ij}(x))_{i,j=1,\ldots,k}$ is a $k \times k$ nonsingular matrix valued function and v a new control, the system (0.1) is transformed into $\dot{x} = (Y^0(x) + G(x)\,h(x)) + G(x)\,H(x)v$, i.e., a system again of the form

$$\dot{x} = \tilde{Y}^0(x) + \sum_{i=1}^{k} v_i \tilde{Y}^i(x). \tag{0.2}$$

Letting $Y(x)$ be the $n \times (k+1)$ matrix with columns $Y^0(x),..., Y^k(x)$ and $\tilde{Y}(x)$ the $n \times (k+1)$ matrix with columns $\tilde{Y}^0(x),..., \tilde{Y}^k(x)$, we find the relationship

$$Y(x)\,A(x) = \tilde{Y}(x), \tag{0.3}$$

where $A(x)$ is the $(k+1) \times (k+1)$ nonsingular matrix

$$A(x) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ h_{10}(x) & h_{11}(x) & \cdots & h_{1k}(x) \\ \vdots & \vdots & & \vdots \\ h_{k0}(x) & h_{k1}(x) & \cdots & h_{kk}(x) \end{pmatrix}. \tag{0.4}$$

Matrices of this form give a representation of the affine (or in systems theory, feedback) group. Two systems such as (0.1), (0.2) whose describing vector fields are related by (0.3) are called feedback equivalent. The problem of when the system (0.1) can be transformed into a linear system via state feedback has been studied in [3] while linearization via a diffeomorphism of $M \times \mathbb{R}^k$ ($\mathbb{R}^k$ the control space) was studied in [4; 5]. The notion of a linear system is not coordinate free, hence one must first answer the question of when, with proper choice of local coordinates, the system (0.1) is linear. This was accomplished in [6]. In [3], local coordinate changes are included in the definition of the feedback group. We are interested in when system (0.1) is "equivalent" to a nilpotent system, i.e., a system of the form (0.1) described by vector fields which generate a nilpotent Lie algebra. This is a coordinate-free concept.

Assume that $Y^0(p),..., Y^k(p)$ are linearly independent so $x \to D^{k+1}(x) = \text{span}\{Y^0(x),..., Y^k(x)\}$ locally defines a $(k+1)$-dimensional distribution. Letting $Y(x)$, $\tilde{Y}(x)$ be as above one easily sees that any other basis $\tilde{Y}^0,..., \tilde{Y}^k$ for D^{k+1} has the form $\tilde{Y}(x) = Y(x)\,M(x)$ for $M(x) \in Gl(k+1, \mathbb{R})$. From the special form of the matrix $A(x)$ in (0.4) it follows that: *a necessary condition for system* (0.1) *to be feedback equivalent to a nilpotent system is that both of the distributions* $x \to D^{k+1}(x) = \text{span}\{Y^0(x),..., Y^k(x)\}$ *and* $x \to D^k(x) = \text{span}\{Y^1(x),..., Y^k(x)\}$ *admit nilpotent bases.* This is, however, not a sufficient condition.

One may extend the feedback group by also allowing reparametrization of trajectories which is equivalent to introducing a scalar function $x \to \gamma(x) > 0$ on the right side of (0.2). Since the h_{ij} are arbitrary, rename $\gamma(x)\,h_{ij}(x)$ as $h_{ij}(x)$ and one obtains that, via feedback and reparametrization, the systems (0.1) and (0.2) are related by (0.3) with the 1 which appears as the upper left entry of $A(x)$ in (0.4) replaced by $\gamma(x)$. Such matrices again form a subgroup of $Gl(k+1, \mathbb{R})$ which we shall call the F/R group. F/R equivalent systems have the same orbit structure, i.e., are *orbit equivalent.* It is, however, still not the case that if $x \to D^{k+1}(x)$ has a nilpotent basis $X^0,..., X^k$ then system

(0.1) is F/R equivalent to a system with $X^0,\dots,X^k$ as its describing vector fields, a property desirable for applications.

State feedback was replaced by a local diffeomorphism of $M \times \mathbb{R}^k$ in [4; 5], which led to a notion the authors called $\mathscr{F}$ equivalence of systems. These papers give sufficient conditions that system (0.1) be $\mathscr{F}$ equivalent to a linear system. Our approach, here, will be related but more geometrical.

For the sake of discussion, assume in (0.1) that the control values lie in the unit cube and let

$$R(x) = \left\{ Y^0(x) + \sum_{i=1}^{k} u_i Y^i(x),\ |u_i| \leqslant 1,\ i = 1,\dots,k \right\}.$$

The set valued function $x \to R(x)$ is often called the local direction cone (or vectogram) and carries the basic information of the system. If two systems, such as (0.1), (0.2), have the same direction cones they have the same trajectories, i.e., are *trajectory equivalent.* Now suppose $x \to D^{k+1}(x) = \operatorname{span}\{Y^0(x),\dots, Y^k(x)\}$ has a nilpotent basis $X^0,\dots,X^k$. Then (locally near p) one can write

$$R(x) = \left\{ \sum_{j=0}^{k} v_j X^j(x) \colon v = (v_0,\dots,v_k) \in S(x) \right\} \tag{0.5}$$

with the set $S(x)$ readily determined. Indeed, if we write $Y^i(x) = \sum_{j=0}^{k} m_{ij}(x) X^j(x)$ then $v_j = m_{0j}(x) + \sum_{i=1}^{k} m_{ij}(x)\, u_j$. Replacing X^0 by $-X^0$ if necessary and choosing $C_\varepsilon = \{u \in \mathbb{R}^k \colon |u_i| \leqslant \varepsilon,\ i = 1,\dots,k\}$, where $\varepsilon > 0$ may depend on x, we may assume $v_0 > 0$. By a local reparametrization, which preserves orbits, we can achieve $v_0 = 1$, hence $S(x)$ is the (convex) affine image of a small cube for each x. This yields

PROPOSITION 0.1. *Assume, in system* (0.1), *that* $Y^0(p),\dots, Y^k(p)$ *are linearly independent so that the map* $x \to \operatorname{span}\{Y^0(x),\dots, Y^k(x)\}$ *defines a* $(k+1)$*-dimensional distribution* D^{k+1}. *Then* D^{k+1} *admits a nilpotent basis* $\{X^0,\dots,X^k\}$ *if and only if system* (0.1) *is orbit equivalent to a nilpotent system of the form* $\dot{x} = X^0(x) + \sum_{i=1}^{k} v_i X^i(x)$, $x(0) = p$.

In view of Proposition 0.1, we turn our attention to the problem of determining when a distribution $x \to D^k(x)$ admits a nilpotent basis locally. Notationally, M_p will denote the tangent space to M at p, $L(Y^1,\dots, Y^k)$ the Lie subalgebra of $V(M)$ generated by $Y^1,\dots, Y^k \in V(M)$, and $L(Y^1,\dots, Y^k)(p)$ the subspace of M_p obtained by evaluating the elements of $L(Y^1,\dots, Y^k)$ at p. For $X, Y \in V(M)$ and $f\colon M \to \mathbb{R}$ smooth we choose the convention $[X, Y]f = Y(Xf) - X(Yf)$ which, in local coordinates $x_1,\dots, x_n$ on M, becomes $[X, Y](x) = X_x(x)\, Y(x) - Y_x(x) X(x)$ the $X_x(x)$, $Y_x(x)$ being Jacobian matrices of partial derivatives. This is the negative of the Lie product often used by differential geometers but is the definition relative to which the

Campbell–Baker–Hausdorff formula is usually given. We also let $[X, Y] = (\operatorname{ad} X, Y)$ and inductively $(\operatorname{ad}^{m+1} X, Y) = [X, (\operatorname{ad}^m X, Y)]$, and for $W \in V(M)$, $(\exp tW)(p)$ will denote the solution, at time t, of $dx/dt = W(x)$, $x(0) = p$.

Our concern is with the case when D^k is not involutive. If $D^k(x) = \operatorname{span}\{Y^1(x),..., Y^k(x)\}$ and $\dim L(Y^1,..., Y^k)(p) = m < n$, (we are in the analytic category) the Hermann–Nagano theorem [7; 8], yields the existence of an m-dimensional integral manifold for $L(Y^1,..., Y^k)$ through p. For control systems, all solutions initiating from p would then remain on this integral manifold, i.e., one could as well replace the original manifold M by this integral manifold. For this reason we assume throughout that $\dim L(Y^1,..., Y^k)(p) = n$.

1. Necessary Conditions That a Distribution Admit a Nilpotent Basis

Let $Y^1,..., Y^k \in V(M)$ be linearly independent at p and such that $\dim L(Y^1,..., Y^k)(p) = n$. Then $x \to D^k(x) = \operatorname{span}\{Y^1(x),..., Y^k(x)\}$ locally (near p) defines a k-dimensional distribution. If $X^1,..., X^k \in V(M)$ is any other basis for D^k there exists a smooth $k \times k$ matrix valued function $A = (a_{ij})$, $(A(x) \in Gl(k, \mathbb{R}))$ such that $Y^i(x) = \sum_{j=1}^k a_{ij}(x) X^j(x)$. Let $\mathscr{Y}^1(x)$ be the set $\{Y^1(x),..., Y^k(x)\}$ and inductively $\mathscr{Y}^m(x)$ is the set of products of l-tuples of $Y^1,..., Y^k$ evaluated at x with $l \leqslant m$. Similarly, $\mathscr{X}^m(x)$ is the set of products of l-tuples of $X^1,..., X^k$ evaluated at x with $l \leqslant m$.

Proposition 1.1. *The integer valued functions $x \to \dim \operatorname{span} \mathscr{Y}^m(x)$ are independent of the basis $\mathscr{Y}^1$ for D^k.*

To see this let $\mathscr{Y}^1 = \{Y^1,..., Y^k\}$ and $\mathscr{X}^1 = \{X^1,..., X^k\}$ be bases for D^k, as above. Clearly $\dim \operatorname{span} \mathscr{Y}^1(x) = \dim \operatorname{span} \mathscr{X}^1(x) = k$. Next, if $Y^i = \sum a_{ij} X^j$ and $Z \in \mathscr{Y}^{m-1} \subset \operatorname{span} \mathscr{X}^{m-1}$ the formula $[Y^i, Z] = \sum ((Za_{ij}) X^j + a_{ij}[X^j, Z])$ provides an inductive proof that $\mathscr{Y}^m(x) \subset \mathscr{X}^m(x)$. Similarly $\mathscr{X}^m(x) \subset \mathscr{Y}^m(x)$.

The next structure theorem for nilpotent Lie algebras of vector fields is necessary for our development.

Theorem 1. *Let $X^1,..., X^k \in V(M)$ with $L = L(X^1,..., X^k)$ nilpotent and $\dim L(p) = n$. Define $\mathscr{K}_0 = \{V \in L: V(p) = 0\}$, $\mathscr{K}_1 = \{V \in L: [V, \mathscr{K}_0] \subset \mathscr{K}_0\}$ and inductively $\mathscr{K}_i = \{V \in L: [V, \mathscr{K}_{i-1}] \subset \mathscr{K}_{i-1}\}$, $i = 1, 2,...$. Then each $\mathscr{K}_i$ is a subalgebra; $\mathscr{K}_{i-1}$ is an ideal in $\mathscr{K}_i$ and if $r_i = \dim \mathscr{K}_i(p)$, $r_0 < r_1 < \cdots < r_m = n$ for some m.*

Proof. Clearly $\mathscr{K}_0$ is a subalgebra. (Indeed one may easily show that if L

has dimension l as a real Lie algebra then $\dim L(p)=n$ implies $\dim \mathscr{H}_0 = l-n$.) Clearly $r_0=\dim \mathscr{H}_0(p)=0$. To see that $\mathscr{H}_1$ is a subalgebra, suppose $V^1, V^2 \in \mathscr{H}_1$, $H \in \mathscr{H}_0$ is arbitrary and denote $[V^1, H]=H^1 \in \mathscr{H}_0$, $[V^2, H]=H^2 \in \mathscr{H}_0$. Then by the Jacobi identity $[[V^1, V^2], H] = [V^1, [V^2, H]] + [V^2, [V^1, H]] = [V^1, H^2] + [V^2, H^1] \in \mathscr{H}_0$. This shows $\mathscr{H}_1$ is a subalgebra; $\mathscr{H}_0$ is an ideal in $\mathscr{H}_1$ by definition. Next, and here we use the fact that L is nilpotent, we show that $r_1 > r_0 = 0$. L nilpotent implies there exists an integer s such that any product of $(s+1)$ elements of L is zero, and hence in $\mathscr{H}_0$. Define $\Lambda_0 = \{$integers l: any product of $(l+1)$ elements of L is in $\mathscr{H}_0\}$. Then $\Lambda_0 \notin \phi$ since $s \in \Lambda_0$. Also $\dim L(p)=n>0$ implies $0 \notin \Lambda_0$. Then Λ_0 has a least element $l^* > 0$; i.e., there exist $V^1,..., V^{l^*} \in L$ such that $V = [\cdots[V^1, V^2],\ V^3]\cdots],\ V^{l^*}] \notin \mathscr{H}_0$ but $[V, W] \in \mathscr{H}_0$ for any $W \in L$, in particular $[V, W] \in \mathscr{H}_0$ for any $W \in \mathscr{H}_0$. Thus $V \in \mathscr{H}_1$, $V \notin \mathscr{H}_0$ so $V(p) \neq 0$ and $r_1 = \dim \mathscr{H}_1(p) > r_0 = 0$.

The argument that $\mathscr{H}_i$ is a subalgebra proceeds via the Jacobi identity as for $\mathscr{H}_1$, while $\mathscr{H}_{i-1}$ is an ideal in $\mathscr{H}_i$ by definition. Define $\Lambda_{i-1} = \{$integers $l \geqslant 0$: any product of $(l+1)$ elements of L is in $\mathscr{H}_{i-1}\}$. Since $\mathscr{H}_0 \subset \mathscr{H}_{i-1}$ we always have $s \in \Lambda_{i-1}$ and if $\dim \mathscr{H}_{i-1}(p) < n$, $0 \notin \Lambda_{i-1}$ so Λ_{i-1} has a least element $l^* > 0$. (If $\dim \mathscr{H}_{i-1}(p) = n$ the proof is complete.) Thus, as in the case $i=1$, we obtain an element $V \in \mathscr{H}_i$ with $V \notin \mathscr{H}_{i-1}$. Finally, if for any such V we were to have $V(p) = \sum a_j W^j(p)$ with $W^j \in \mathscr{H}_{i-1}$ (i.e., $\dim \mathscr{H}_i(p) = \dim \mathscr{H}_{i-1}(p)$) then $(V - \sum a_j W^j) = H \in \mathscr{H}_0 \subset \mathscr{H}_{i-1}$ hence $V = H + \sum a_j W^j \in \mathscr{H}_{i-1}$ since $\mathscr{H}_{i-1}$ is a subalgebra. This contradiction shows $\dim \mathscr{H}_i(p) = r_i > r_{i-1}$ if $r_{i-1} < n$, completing the proof. ∎

Let $X^1,..., X^k \in V(M)$ with $L(X^1,..., X^k)$ nilpotent. Then the $\mathscr{H}_i$, as defined in Theorem 1, are subalgebras of vector fields, hence, by the Hermann–Nagano theorem, each $\mathscr{H}_i$ has an integral manifold of dimension r_i thru p.

The next proposition shows the homogeneity forced by nilpotency. This will be used to construct examples of distributions which do not admit nilpotent bases.

PROPOSITION 1.2. *For each $i = 1, 2,..., \dim \mathscr{H}_{i-1}(x)$ is constant on the integral manifold of $\mathscr{H}_i$ through p.*

Proof. It is well known that for any subalgebra $\mathscr{H} \subset V(M)$, $\dim \mathscr{H}(x)$ is constant on the integral manifold of $\mathscr{H}$ through p. The interesting feature, here, is that $\dim \mathscr{H}_{i-1}(x)$ is constant on the integral manifold of $\mathscr{H}_i$ (which is larger than that of $\mathscr{H}_{i-1}$) through p.

Let M^{r_i} denote the integral manifold, having dimension r_i, of $\mathscr{H}_i$ through p. For any $x \in M^{r_i}$ near p, there is a $W \in \mathscr{H}_i$ and real s such that $x = (\exp sW)(p)$. Let $V \in \mathscr{H}_{i-1}$ so $(\operatorname{ad}^\nu W, V) \in \mathscr{H}_{i-1}$ for all $\nu = 0, 1,...,$ and $(\exp -sW)_*: M_x \to M_p$ denote the tangent space isomorphism induced by the diffeomorphism $x \to (\exp -sW)(x)$. Then $(\exp -sW)_*\ V((\exp sW)(p)) =$

$\sum_{\nu=0}^{\infty}((-s)^{\nu}/\nu!)(\mathrm{ad}^{\nu} W, V)(p) \in \mathscr{H}_{i-1}(p)$. One concludes $(\exp - sW)_* \mathscr{H}_{i-1}(x) = \mathscr{H}_{i-1}(p)$ showing $\dim \mathscr{H}_{i-1}(p)$ is constant, locally, on M^{r_i}. ∎

The typical use of Proposition 1.2 to construct a distribution which does not admit a nilpotent basis proceeds as follows. Let $x \to D^2(x) = \mathrm{span}\{Y^1(x), Y^2(x)\}$ be a two-distribution such that $Y^1(p)$, $Y^2(p)$ are independent, $\dim L(Y^1, Y^2)(p) = n$, $[Y^1, Y^2](p) = 0$ but $Y^1(x), Y^2(x), [Y^1, Y^2](x)$ are independent if $x \neq p$. Then from Proposition 1.1, for any basis $\mathscr{X}^1 = \{X^1, X^2\}$ for D^2, $\dim \mathrm{span}\, \mathscr{X}^1(x) = \dim \mathrm{span}\, \mathscr{Y}^1(x) = 2$ while $\dim \mathrm{span}\, \mathscr{X}^2(x) = \dim \mathrm{span}\, \mathscr{Y}^2(x)$ is 2 if $x = p$ and 3 if $x \neq p$. Suppose $\mathscr{X}^1 = \{X^1, X^2\}$ is a nilpotent basis, i.e., $L(X^1, X^2)$ is nilpotent. Then $[X^1, X^2](p) = \alpha_1 X^1(p) + \alpha_2 X^2(p)$ or $V = [X^1, X^2] - \alpha_1 X^1 - \alpha_2 X^2 \in \mathscr{H}_0$, where the $\mathscr{H}_i$ are defined as in Theorem 1 relative to the nilpotent algebra $L(X^1, X^2)$. But then $\dim \mathscr{H}_1(p) = r_1 \geqslant 1$ and V must vanish on the integral manifold M^{r_1} of $\mathscr{H}_1$ through p, i.e., $\dim \mathrm{span}\, \mathscr{X}^2(x) = 2$ for $x \in M^{r_1}$. This contradiction implies D^2 could not admit a nilpotent basis.

EXAMPLE 1.1. For any integer $2 \leqslant k < n$ there exists a k-dimensional distribution D^k on M^n which does not admit a nilpotent basis locally.

Let $M = \mathbb{R}^n$ and define $D^k(x) = \mathrm{span}\{X^1(x), \ldots, X^k(x)\}$, where

$$X^i = \partial/\partial x_i, \qquad 1 \leqslant i < k-1,$$
$$X^k = \partial/\partial x_k - x_1(x_1^2/6 + x_2^2/2 + \cdots + x_{k+1}^2/2)\,\partial/\partial x_{k+1} + \sum_{l=4}^{n-k+2} ((-1)^l x_1^l/l! - x_1 x_{k+l-2})\,\partial/\partial x_{k+l-2}.$$

One calculates the relevant brackets for this example:

$$[X^1, X^k] = (\tfrac{1}{2})(x_1^2 + \cdots + x_{k+1}^2)\,\partial/\partial_{k+1} + \sum_{l=4}^{n-k+2} ((-1)^{l-1} x_1^{l-1}/(l-1)! + x_{k+l-2})\,\partial/\partial x_{k+l-2},$$

$$[X^1, [X^1, X^k]] = -x_1\,\partial/\partial x_{k+1} + \sum_{l=4}^{n-k+2} ((-1)^{l-2} x^{l-2}/(l-2)!)\,\partial/\partial x_{k+l-2},$$

and for $3 \leqslant m \leqslant n-k+2$,

$$(\mathrm{ad}^m X^1, X^k) = \partial/\partial x_{k+m-2} + \sum_{l=1}^{n-k+2-m} ((-1)^l x_1^l/l!)\,\partial/\partial x_{k+m+l-2}.$$

Clearly $\{X^1, \ldots, X^k, (\mathrm{ad}^3 X^1, X^k), \ldots, (\mathrm{ad}^{n-k+2} X^1, X^k)\}$ span $\mathbb{R}_p^n$ at all points p. The calculation of $[X^1, X^k]$ shows that $[X^1, X^k](x_1, \ldots, x_n) = 0$ if and only if $x_1 = x_2 = \cdots = x_n = 0$.

2. The Construction of Nilpotent Bases for Distributions

The first of two methods which we consider for the construction of nilpotent bases proceeds as follows. Given a distribution $y \to D^k(y) = \text{span}\{Y^1(y),..., Y^k(y)\}$ on M we construct a nilpotent basis $\mathscr{X}^1 = \{X^1,..., X^k\}$ for a distribution on $\mathbb{R}^n$ such that $\dim \text{span}\, \mathscr{X}^i(x) = \dim \text{span}\, \mathscr{Y}^i(y)$, $i = 1, 2,...$. We then attempt to construct a diffeomorphism $\phi: \mathbb{R}^n \to M$ such that $\phi_* X^1,..., \phi_* X^k$ is a basis (hence a nilpotent basis) for D^k. This construction leads to a system of first-order partial differential equations which, if solvable, produce ϕ. The second method is the use of Darboux's theorem to obtain a preferred local coordinate system.

The approach of attempting to realize a nilpotent basis for D^k on M as the image of a nilpotent basis for a distribution on $\mathbb{R}^n$ by the induced map of a diffeomorphism is general. Indeed, Sussmann [9] shows that if $L(Y^1,..., Y^k)$ and $L(X^1,..., X^k)$ are isomorphic Lie algebras of real analytic vector fields on, respectively, M and $\mathbb{R}^n$, with $\dim L(Y^1,..., Y^k)(p) = n = \dim L(X^1,..., X^k)(0)$ then the Lie algebra isomorphism can be realized as the induced map of a (local) diffeomorphism $\phi: \mathbb{R}^n \to M$ with $\phi(0) = p$.

Given, locally, vector fields $V^1,..., V^n \in V(M)$ and $W^1,..., W^n \in V(\mathbb{R}^n)$ which are linearly independent, respectively, at $p \in M$ and $0 \in \mathbb{R}^n$ (these will later be related to the $Y^1,..., Y^k$ and $X^1,..., X^k$ above) the first goal is to construct an arbitrary diffeomorphism $\phi: \mathbb{R}^n \to M$ with $\phi(0) = p$ and such that $\phi_* W^i$ is expressed in terms of the V^i. Introduce local coordinates of the first kind on $\mathbb{R}^n$ via the map

$$s = (s_1,..., s_n) \to g(s) = \exp(s_1 W^1 + \cdots + s_n W^n)(0). \tag{2.1}$$

Our notation will be $x = g(s)$; note that $g^{-1}(x)$ exists locally. Let $f = (f_1,...,f_n): \mathbb{R}^n \to \mathbb{R}^n$ with $f(0) = 0$ and Jacobian $(\partial f_i/\partial s_j(0))$, denoted $f_s(0)$, nonsingular. Then if $x = g(s)$, any diffeomorphism $\phi: \mathbb{R}^n \to M$ with $\phi(0) = p$ can be expressed as

$$\phi(x) = (\exp f_1(s)\, V^1) \circ \cdots \circ (\exp f_n(s)\, V^n)(p) \tag{2.2}$$

for some f as above. Consider f free to be later chosen. In order to compute ϕ_* we first note that for any $1 \leqslant i \leqslant n$ there will exist a smooth function $r^i(t, s) = (r_1^i(t, s),..., r_n^i(t, s))$ such that $r_j^i(0, s) = s_j$ and

$$(\exp tW^i) \circ \exp(s_1 W^1 + \cdots + s_n W^n)(0)$$
$$= \exp(r_1^i(t, s)\, W^1 + \cdots + r_n^i(t, s)\, W^n)(0).$$

The functions $r^i(t, s)$ are completely determined by the structure of $L(W^1,..., W^n)$ and are quite computable if this algebra is nilpotent. In the calculation to follow, terms such as the inner product $\dot{r}^i(0,s)\cdot \partial f_j(s)/\partial s$ often appear. Our notation will be to associate an operator

$$\mathscr{R}^i = \dot{r}^i_1(0,s)\,\partial/\partial s_1 + \cdots + \dot{r}^i_n(0,s)\,\partial/\partial s_n \tag{2.3}$$

with $\dot{r}^i(0,s)$ and write the previous inner product as $(\mathscr{R}^i f_j)(s)$. Thus

$$\begin{aligned}
\phi_*(x)\, W^i(x) &= (d/dt)\,\phi((\exp tW^i)\circ(\exp(s_1 W^1 + \cdots + s_n W^n))(0))\,|_{t=0}\\
&= d/dt\{(\exp f_1(r^i(t,s))\,V^1)\circ\cdots\circ(\exp f_n(r^i(t,s))\,V^n)(p)\}\,|_{t=0}\\
&= (\mathscr{R}^i f_1)(s)\,V^1(\phi(x)) + (\mathscr{R}^i f_2)(s)(\exp f_1(s)\,V^1)_*\\
&\quad\times V^2((\exp -f_1(s)\,V^1)\circ\phi(x))\\
&\quad + (\mathscr{R}^i f_3)(s)(\exp f_1(s)\,V^1)_*\,(\exp f_2(s)\,V^2)_*\\
&\quad\times V^3((\exp -f_2(s)\,V^2)\circ(\exp -f_1(s)\,V^1)\circ\phi(x)) + \cdots\\
&= (\mathscr{R}^i f_1)(s)\,V^1(\phi(x)) + (\mathscr{R}^i f_2)(s)\sum_{\nu_1=0}^{\infty}\frac{f_1^{\nu_1}(s)}{\nu_1!}(\mathrm{ad}^{\nu_1} V^1, V^2)(\phi(x))\\
&\quad + (\mathscr{R}^i f_3)(s)\sum_{\nu_1,\nu_2=0}^{\infty}\frac{f_1^{\nu_1}(s)\,f_2^{\nu_2}(s)}{\nu_1!\,\nu_2!}\\
&\quad\times(\mathrm{ad}^{\nu_1} V^1, (\mathrm{ad}^{\nu_2} V^2, V^3))(\phi(x)) + \cdots.
\end{aligned}\tag{2.4}$$

The basic idea in the use of this formula is as follows. Suppose $\{Y^1,\dots,Y^k\}$ is a local basis for the distribution D^k on M and $\dim L(Y^1,\dots,Y^k)(p) = n$. Choose $V^1 = Y^1,\dots, V^k = Y^k$ and $V^{k+1},\dots,V^n \in V(M)$ so that $V^1(p),\dots,V^n(p)$ are independent. Next select an appropriate nilpotent model Lie algebra on $\mathbb{R}^n$, say generated by $X^1,\dots,X^k$. Let $W^1 = X^1,\dots,W^k = X^k$ and $W^{k+1},\dots,W^n \in V(\mathbb{R}^n)$ be such that $W^1(0),\dots,W^n(0)$ are independent. In applications it is often useful to choose $V^{k+1},\dots,V^n \in L(Y^1,\dots,Y^k)$ and $W^{k+1},\dots,W^n \in L(X^1,\dots,X^k)$. The conditions $\phi_*(x)\,W^i(x) \in \operatorname{span}\{V^1(\phi(x)),\dots, V^k(\phi(x))\}$, i.e., that the coefficients of $V^{k+1}(\phi(x)),\dots,V^n(\phi(x))$ in Eq. (2.4) vanish, yield partial differential equations for the f_i. In order to exhibit these coefficients we express the right side of Eq. (2.4) as a linear combination of $V^1(\phi(x)),\dots,V^n(\phi(x))$. To this purpose, let

$$\begin{aligned}
(\mathrm{ad}^{\nu_1} V^1, V^2)(\phi(x)) &= \sum_{l=1}^{n}\beta_{\nu_1,l}(\phi(g(s)))\,V^l(\phi(x))\\
(\mathrm{ad}^{\nu_1} V^1, (\mathrm{ad}^{\nu_2} V^2, V^3))(\phi(x)) &= \sum_{l=1}^{n}\beta_{\nu_1,\nu_2,l}(\phi(g(s)))\,V^l(\phi(x))\\
(\mathrm{ad}^{\nu_1} V^1, (\cdots(\mathrm{ad}^{\nu_{n-1}} V^{n-1}, V^n)\cdots)(\phi(x)) &= \sum_{l=1}^{n}\beta_{\nu_1,\dots,\nu_{n-1},l}(\phi(g(s)))\,V^l(\phi(x)).
\end{aligned}\tag{2.5}$$

Note that

$$\beta_{0,l}=1 \quad \text{if} \quad l=2, \qquad =0 \quad \text{if} \quad l\neq 2;$$

$$\beta_{0,0,l}=1 \quad \text{if} \quad l=3, \qquad =0 \quad \text{if} \quad l\neq 3; \quad \cdots; \quad \beta_{0,0,\ldots,0,l}=1 \quad \text{if} \quad l=n, \qquad =0 \quad \text{if} \quad l\neq n.$$

Substituting in the right side of (2.4) for $1\leqslant l\leqslant n$, the coefficient of $V^l(\phi(x))$ is

$$(\mathscr{R}^l f_2)(s)\sum_{\nu_1=1}^{\infty}\frac{f_1^{\nu_1}}{\nu_1!}\beta_{\nu_1,l}+(\mathscr{R}^l f_3)(s)\sum_{\nu_1+\nu_2=1}^{\infty}\frac{f_1^{\nu_1}f_2^{\nu_2}}{\nu_1!\,\nu_2!}\beta_{\nu_1,\nu_2,l}+\cdots$$
$$+(\mathscr{R}^l f_l)(s)\left[1+\sum_{\nu_1+\cdots+\nu_{l-1}=1}^{\infty}\frac{f_1^{\nu_1}\cdots f_{l-1}^{\nu_{l-1}}}{\nu_1!\cdots\nu_{l-1}!}\beta_{\nu_1,\ldots,\nu_{l-1},l}\right]+\cdots$$
$$+(\mathscr{R}_i f_n)(s)\sum_{\nu_1+\cdots+\nu_{n-1}=1}^{\infty}\frac{f_1^{\nu_1}\cdots f_{n-1}^{\nu_{n-1}}}{\nu_1!\cdots\nu_{n-1}!}\beta_{\nu_1,\ldots,\nu_{n-1},n}. \tag{2.6}$$

The goal is to choose f, subject to the condition $f(0)=0$ and the Jacobian $f_s(0)$ nonsingular, to make the coefficients of $V^{k+1},\ldots, V^n$ zero for each $i=1,\ldots,k$. The process is best illustrated by its use in

THEOREM 2. *Let $x\to D^2(x)=\text{span}\{Y^1(x), Y^2(x)\}$ be a two-dimensional distribution on $\mathbb{R}^3$ and suppose $Y^1(0)$, $Y^2(0)$, $[Y^1,Y^2](0)$ are independent. Then D^2 admits a nilpotent basis which generates a three-dimensional nilpotent algebra.*

Proof. Choose the model nilpotent algebra on $\mathbb{R}^3$ generated by vector fields X^1, X^2 with structure: $X^1(0)$, $X^2(0)$, $[X^1,X^2](0)$ are independent and all other products vanish. Let $W^1=X^1$, $W^2=X^2$, $W^3=[X^1,X^2]$. Then

$$(\exp tW^1)\circ\exp(s_1W^1+s_2W^2+s_3W^3)$$
$$=\exp\left((s_1+t)\,W^1+s^2\,W^2+\left(s_3+\frac{ts_2}{2}\right)W^3\right)$$

so $r^1(t,s)=(s_1+t,\ s_2,\ s_3+ts_2/2)$, $\dot r^1(0,s)=(1,0,s_2/2)$, and $\mathscr{R}^1=\partial/\partial s_1+(s_2/2)\,\partial/\partial s_3$. A similar computation gives $\mathscr{R}^2=\partial/\partial s_2-(s_1/2)\,\partial/\partial s_3$. Next, choose $V^1=Y^1$, $V^2=Y^2$, and $V^3=[Y^1,Y^2]$. Our goal is to have the coefficient of V^3 in (2.4) to be zero for $i=1,2$. Explicitly, from (2.6), this leads to the equations

$$\left[1+\sum_{\nu_1+\nu_2=1}^{\infty}\frac{f_1^{\nu_1}(s)f_2^{\nu_2}(s)}{\nu_1!\,\nu_2!}\beta_{\nu_1,\nu_2,3}\right](\mathscr{R}^1 f_3)(s) = -(\mathscr{R}^1 f_2)(s)\sum_{\nu_1=1}^{\infty}\frac{f_1^{\nu_1}(s)}{\nu_1!}\beta_{\nu_1,3},$$
$$\left[1+\sum_{\nu_1+\nu_2=1}^{\infty}\frac{f_1^{\nu_1}(s)f_2^{\nu_2}(s)}{\nu_1!\,\nu_2!}\beta_{\nu_1,\nu_2,3}\right](\mathscr{R}^2 f_3)(s) = -(\mathscr{R}^2 f_2)(s)\sum_{\nu_1=1}^{\infty}\frac{f_1^{\nu_1}(s)}{\nu_1!}\beta_{\nu_1,3}. \tag{2.7}$$

Choose $f_2(s)=s_2$ which gives $\mathscr{R}^1 f_2=0$, $\mathscr{R}^2 f_2=1$ and the first of the above equations becomes $\mathscr{R}^3 f_3=0$, which has a solution $f_3(s)=s_3-s_1s_2/2$. Then the first equation is satisfied. In the second equation, $\mathscr{R}^2 f_3=-s_1$. Since $V^3=[Y^1,Y^2]$ and $\beta_{\nu_1,3}$ is the component of $(\mathrm{ad}^{\nu_1} Y^1, Y^2)$ on V^3, we have $\beta_{1,3}=1$. The second becomes

$$f_1(s)+\sum_{\nu_1=2}^{\infty}\frac{f_1^{\nu_1}(s)}{\nu_1!}\beta_{\nu_1,3}+s_1\left[1+\sum_{\nu_1+\nu_2=1}^{\infty}\frac{f_1^{\nu_1}(s)f_2^{\nu_2}(s)}{\nu_1!\,\nu_2!}\beta_{\nu_1,\nu_2,3}\right]=0. \tag{2.8}$$

Letting $\gamma=f_1$, this has the form $\Phi(s,\gamma)=0$, where $\Phi(0,0)=0$ and $\partial\Phi/\partial\gamma(0,0)=1$. By the implicit function theorem, a local solution $\gamma=f_1(s)$ exists such that $f_1(0)=0$ and $\Phi(s,f_1(s))\equiv 0$. Furthermore we may now differentiate (2.8) with respect s_1 and find $\partial f_1(0)/\partial s_1=-1$. With $f_1(s)$ as above, $f_2(s)=s_2$ and $f_3(s)=s_3-s_1s_2/2$, $\det f_s(0)=-1$ and $\phi\colon \mathbb{R}^3\to\mathbb{R}^3$ has been determined as a diffeomorphism such that ϕ_*X^1, ϕ_*X^2 is a nilpotent basis for D^2. ∎

It is interesting to use this method to show the well-known fact that: If $x\to D^k(x)=\mathrm{span}\{Y^1(x),\ldots,Y^k(x)\}$ is an involutive distribution on M^n (say $k<n$ for interest) then D^k admits an abelian basis.

To verify this, choose $V^1=Y^1,\ldots,V^k=Y^k$ and let $V^{k+1},\ldots,V^n\in V(M)$ be so that $V^1(p),\ldots,V^n(p)$ are linearly independent. Choose local coordinates $(z_1,\ldots,z_n)$ on $\mathbb{R}^n$ and let $W^1=X^1=\partial/\partial z_1,\ldots,$ $W^k=X^k=\partial/\partial z_k$ while $W^{k+1},\ldots,W^n$ are arbitrary but so that $W^1(0),\ldots,W^n(0)$ are independent. Then $X^1,\ldots,X^k$ is an abelian basis for a k-dimensional distribution on $\mathbb{R}^n$; the associated operators are $\mathscr{R}^i=\partial/\partial s_i$, $i=1,\ldots,k$. The distribution D^k involutive implies (see Eq. (2.5)) that $\beta_{\nu_1,l}=0$, $\beta_{\nu_1,\nu_2,l}=0,\ldots,\beta_{\nu_1,\ldots,\nu_{n-1},l}=0$ whenever $l>k$. Thus the coefficients of $V^{k+1},\ldots,V^n$ (see Eq. (2.6)) are zero for $i=1,\ldots,k$. We may choose $f_1(s)=s_1,\ldots,f_n(s)=s_n$ and this determines a diffeomorphism $\phi\colon\mathbb{R}^n\to M$ with $\phi_*(x)X^i(x)\in D^k(\phi(x))$, $i=1,\ldots,k$, giving an abelian basis for D^k.

The Use of Darboux's Theorem

Let $M^n = M$ be an n-dimensional manifold with M_x, M_x^*, respectively, the tangent and cotangent spaces of M at x. Our notation is $\Lambda(M_x)$, $\Lambda(M_x^*)$ for, respectively, the exterior or Grassmann algebras over M_x, M_x^*. We write $\Lambda(M_x) = \Lambda^0(M_x) \oplus \cdots \oplus \Lambda^n(M_x)$, where $\Lambda^0(M_x) = \mathbb{R}$, $\Lambda^1(M_x) = M_x \cong \mathbb{R}^n$, etc. The bilinear pairing of $\Lambda^p(M_x)$ with $\Lambda^p(M_x^*)$ is denoted $\langle \alpha(x), \xi(x) \rangle$, $\xi(x) \in \Lambda^p(M_x)$, $\alpha(x) \in \Lambda^p(M_x^*)$. Recall that the dimension of $\Lambda^p(M_x)$ is the binomial coefficient $\binom{n}{p}$; $\Lambda^p(M_x) = 0$ if $p > n$. We write $d\omega \wedge d\omega$ as $(d\omega)^2$, etc., for a 1-form ω, i.e., $\omega(x) \in M_x^*$.

DEFINITION. The rank of the 1-form ω at p is r if $(\omega \wedge (d\omega)^r)(p) \neq 0$ but $(\omega \wedge (d\omega)^{r+1})(p) = 0$. (Note $2r + 1 \leqslant n$.) Thus if $\omega \wedge (d\omega)^{r+1} \equiv 0$ in an nbd of p while $(\omega \wedge (d\omega)^r)(p) \neq 0$, the 1-form ω will have rank r in an nbd of p. The following form of Darboux's theorem is as given in [10].

THEOREM (Darboux). *If the 1-form ω has constant rank r in an* nbd *of p, there exist local coordinates $x_1, \ldots, x_n$ such that*

$$\omega = dx_1 + x_2\, dx_3 + \cdots + x_{2r}\, dx_{2r+1}. \tag{2.9}$$

The general use of Darboux's theorem to obtain a nilpotent basis for a distribution is illustrated in

EXAMPLE 2.1. Let $Y^1, \ldots, Y^{n-1}$ be vector fields on M^n which are linearly independent at p, so $x \to D^{n-1}(x) = \operatorname{span}\{Y^1(x), \ldots, Y^{n-1}(x)\}$ is an $(n-1)$-distribution. Suppose $\omega \neq 0$ is a 1-form such that $\langle \omega(x), Y^i(x) \rangle \equiv 0$ for x in an nbd of p. If ω has constant rank r in an nbd of p, (2.9) holds and (for $r \geqslant 1$) the vector fields $X^1(x) = -x_2\, \partial/\partial x_1 + \partial/\partial x_3, \ldots, X^r = -x_{2r}\, \partial/\partial x_1 + \partial/\partial x_{2r+1}$, $X^{r+1} = \partial/\partial x_2$, $X^{r+2} = \partial/\partial x_4, \ldots, X^{2r} = \partial/\partial x_{2r}$, $X^{2r+1} = \partial/\partial x_{2r+2}$, $X^{2r+2} = \partial/\partial x_{2r+3}, \ldots, X^{n-1} = \partial/\partial x_n$ also satisfy $\langle \omega, X^i \rangle \equiv 0$, $i = 1, \ldots, n-1$, hence $X^1, \ldots, X^{n-1}$ is again a basis for D^{n-1} and, in fact, a nilpotent basis with $\dim L(X^1, \ldots, X^{n-1}) = n$. (If $r = 0$, the distribution D^{n-1} is involutive.) Notice that if ω has rank r, the above nilpotent basis shows D^{n-1} contains an involutive subdistribution of dimension $(n - r - 1)$.

Remark. Note that the vector fields $X^1, \ldots, X^{n-1}$ obtained by the method of Example 2.1 are affine in the local coordinates. This leads to a linear–bilinear representation of a control system.

EXAMPLE 2.2. The purpose, here, is to stress the importance of the rank of ω being constant in an nbd of p. Let $M = \mathbb{R}^3$; Y^1, Y^2 be two vector fields (as in Example 1.1) which are linearly independent at p while $[Y^1, Y^2](x)$ is linearly independent of $Y^1(x)$, $Y^2(x)$ if $x \neq p$, $[Y^1, Y^2](p) = 0$. Let $\omega \neq 0$ be a 1-form such that $\langle \omega, Y^i \rangle \equiv 0$, $i = 1, 2$. Then from dimensional considerations $\omega \wedge (d\omega)^2 \equiv 0$. Also $\langle \omega \wedge d\omega, [Y^1, Y^2] \wedge Y^1 \wedge Y^2 \rangle = 2\langle \omega, [Y^1, Y^2] \rangle \langle d\omega, Y^1 \wedge Y^2 \rangle$. But from the standard "local Stokes' formula"

$$\langle d\omega, Y^1 \wedge Y^2 \rangle = Y^1(\langle \omega, Y^2 \rangle) - Y^2(\langle \omega, Y^1 \rangle) + \langle \omega, [Y^1, Y^2] \rangle. \tag{2.10}$$

Thus

$$\langle (\omega \wedge d\omega)(x), [Y^1, Y^2](x) \wedge Y^1(x) \wedge Y^2(x) \rangle = 2\langle \omega(x), [Y^1, Y^2](x) \rangle^2. \tag{2.11}$$

This shows

$$\begin{aligned} \operatorname{rank} \omega(x) &= 0 \qquad \text{if} \quad x = p, \\ &= 1 \qquad \text{if} \quad x \neq p. \end{aligned}$$

Darboux's theorem does not apply; indeed the distribution $x \to D^2(x) = \operatorname{span}\{Y^1(x), Y^2(x)\}$ does not admit a nilpotent basis.

LEMMA. *Let V be vector space with basis $\{V_1, \ldots, V_n\}$ and let $\theta \in \Lambda^2(V^*)$ be a skew-symmetric 2-form on V. Then $\theta^r \neq 0$ and $\theta^{r+1} = 0$ if and only if the skew-symmetric matrix $\Theta = (\theta(V_i, V_j))$ has rank $2r$.*

Proof. Let $\{\bar{V}^1, \ldots, \bar{V}^n\}$ be a basis for V^* dual to $\{V_1, \ldots, V_n\}$. According to Sternberg [11, p. 24] there is a basis $\{\bar{W}^1, \ldots, \bar{W}^n\}$ for V^* so that $\theta = \bar{W}^1 \wedge \bar{W}^2 + \cdots + \bar{W}^{2r-1} \wedge \bar{W}^{2r}$ and r depends only on θ. Indeed, r is characterized by $\theta^r = r!\, \bar{W}^1 \wedge \bar{W}^2 \wedge \cdots \wedge \bar{W}^{2r} \neq 0$ and $\theta^{r+1} = 0$. Let $A = (a_{ij})$ be the nonsingular matrix such that $\bar{V}^i = \sum a_{ij} \bar{W}^j$. Then

$$\begin{aligned} \theta &= \sum_{i<j} \theta(V_i, V_j)\, \bar{V}^i \wedge \bar{V}^j = \tfrac{1}{2} \sum_{i,j} \theta(V_i, V_j)\, \bar{V}^i \wedge \bar{V}^j \\ &= \tfrac{1}{2} \sum_{i,j} \sum_{h,k} a_{ih} \theta(V_i, V_j)\, a_{jk} \bar{W}^h \wedge \bar{W}^k \\ &= \sum_{h<k} \left(\sum_{i,j} a_{ih} \theta(V_i, V_j)\, a_{jk} \right) \bar{W}^h \wedge \bar{W}^k \\ &= \bar{W}^1 \wedge \bar{W}^2 + \cdots + \bar{W}^{2r-1} \wedge \bar{W}^{2r}. \end{aligned}$$

Thus there is a nonsingular matrix A such that $A \Theta A^T$ has r blocks $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ down the diagonal and zeros elsewhere, showing Θ is of rank $2r$.

If Θ is skew-symmetric of rank $2r$, there is an orthogonal matrix Q and a nonsingular diagonal matrix D such that $DQ\, \Theta\, Q^{-1} D = (DQ)\, \Theta\, (DQ)^T$ has 2×2 blocks $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ down the diagonal and zeros elsewhere. If we let $A = DQ$ and change basis in V^* by A, the equations above show $\theta = \bar{W}^1 \wedge \bar{W}^2 + \cdots + \bar{W}^{2r-1} \wedge \bar{W}^{2r}$ so $\theta^r \neq 0$ and $\theta^{r+1} = 0$. ∎

THEOREM 3. *Let M be an $(n+1)$-manifold, let $\mathscr{Y} = \{Y^1,\dots, Y^n\}$ be a local basis for an n-distribution D^n in a neighborhood of $p \in M$, and let ω be a 1-form such that $\langle \omega, Y^i\rangle = 0$ near p for $i = 1, 2,\dots, n$. If the rank of the $n \times n$ skew-symmetric matrix $S = (\langle \omega(p), [Y^i, Y^j](p)\rangle)$ is $2r > 0$, then the rank of ω is r at p, hence ω is of rank $\geqslant r$ in a neighborhood of p.*

If in addition r is such that $\omega \wedge (d\omega)^{r+1}(x) = 0$, or equivalently rank $S(x) = 2r$ in a neighborhood of p, then ω is of rank r in a neighborhood of p and D^n has a nilpotent basis near p which generates an $(n+1)$-dimensional algebra.

Proof. By formula (2.10) $\langle d\omega, Y^i \wedge Y^j\rangle = \langle \omega, [Y^iY^j]\rangle$ and since the rank of S is positive at p we can choose i, j so $\langle \omega(p), [Y^i, Y^j](p)\rangle \neq 0$. By the lemma, $(d\omega)^r(p) \neq 0$ and $(d\omega)^{r+1}(p) = 0$. Thus we can choose $Y^{i_1},\dots, Y^{i_{2r}}$ so that $\langle (d\omega)^r(p), Y^{i_1} \wedge \cdots \wedge Y^{i_{2r}}(p)\rangle \neq 0$. But then

$$\begin{aligned}&\langle \omega \wedge (d\omega)^r(p), [Y^i, Y^j] \wedge Y^i, \wedge \cdots \wedge Y^{i_{2r}}(p)\rangle\\ &\quad = \langle \omega(p), [Y^i, Y^j](p)\rangle\langle (d\omega)^r(p), Y^{i_1} \wedge \cdots \wedge Y^{i_{2r}}(p)\rangle \neq 0.\end{aligned}$$

This shows that $\omega \wedge (d\omega)^r(p) \neq 0$ and clearly $\omega \wedge (d\omega)^{r+1}(p) = 0$ since $(d\omega)^{r+1}(p) = 0$, which means ω is of rank r at p and of rank $\geqslant r$ near p.

The assumption $\omega \wedge (d\omega)^{r+1} = 0$ in a neighborhood means ω is of rank r near p and Darboux's theorem applies as in Example 2.1 to give the result. ∎

Remark. In the case of maximal rank one can reduce computations to a point. Specifically, with D^n and ω as above, we have

COROLLARY. *If the rank of the skew-symmetric matrix $S = (\langle \omega(p), [Y^i, Y^j](p)\rangle)$ is n for n even or $n-1$ for n odd then D^n has a nilpotent basis.*

Proof. For n even $\omega \wedge (d\omega)^{n/2+1}$ is an $(n+2)$-form and for n odd $\omega \wedge (d\omega)^{n/2+1/2}$ is an $(n+2)$-form. Since M is an $(n+1)$-manifold both are zero. This is tantamount to rank $S(p)$ being maximal. ∎

EXAMPLE 2.2. Let Y^1, Y^2, Y^3 be a basis for a 3-dimensional distribution D^3 on M^4. Then $\omega \wedge (d\omega)^2 \equiv 0$ while if D^3 is not involutive, $\omega \wedge d\omega \not\equiv 0$ and $r = 1$. If some $[Y^i, Y^j](p)$ is linearly independent of $Y^1(p), Y^2(p), Y^3(p)$, it follows that the 3×3 skew-symmetric matrix S has rank 2 and D^3 admits a nilpotent basis. (A similar argument gives an alternative proof of Theorem 2.)

EXAMPLE 2.3. This is an example of a 2-dimensional distribution $x \to D^2(x) = \operatorname{span}\{Y^1(x), Y^2(x)\}$ on $\mathbb{R}^3$ for which $[Y^1, Y^2](x) \in D^2(x)$ for x on a 2-manifold through p. The Darboux approach cannot be used but the first method of this section gives the existence of a nilpotent basis for D^2.

Let $Y^1(x) = \partial/\partial x_1 + x_2\,\partial/\partial x_2 + x_1^2 x_2\,\partial/\partial x_3$, $Y^2(x) = \partial/\partial x_2 + x_1^2\,\partial/\partial x_3$, and $p = 0$. Then $(adY^1, Y^2)(x) = \partial/\partial x_2 + (x_1^2 - 2x_1)\,\partial/\partial x_3$, $(ad^\nu Y^1, Y^2)(x) = \partial/\partial x_2 + (x_1^2 - 2x_1 + 2)\,\partial/\partial x_3$ if $\nu \geqslant 2$, so $L(Y^1, Y^2)$ is not nilpotent; $\dim L(Y^1, Y^2)(0) = 3$ and $[Y^1, Y^2](x) = \partial/\partial x_2$ if $x_1 = 0$. We choose $V^1 = Y^1$, $V^2 = Y^2$, and $V^3 = \partial/\partial x_3$.

For the model nilpotent basis, choose $X^1 = \partial/\partial x_1$, $X^2 = \partial/\partial x_2 + x_1^2\,\partial/\partial x_3$ so $L(X^1, X^2)$ is nilpotent of dimension 4, $[X^1, X^2](x) = -2x_1\,\partial/\partial x_3$, hence also vanishes for $x_1 = 0$; $(\text{ad}^2 X^1, X^2) = 2\,\partial/\partial x_3$. Choose $W^1 = X^1$, $W^2 = X^2$, $W^3 = 2\,\partial/\partial x_3$. Explicitly,

$$(\exp(s_1 W^1 + s_2 W^2 + s_3 W^3))(0) = \left(s_1, s_2, \frac{s_1^2 s_2}{3} + 2s_3\right) = g(s). \quad (2.15)$$

It is now easiest to compute $(\exp tW^i)\,g(s)$, for $i = 1, 2$ and write this as $(\exp(r_1^i(t, s)\,W^1 + r_2^i(t, s)\,W^2 + r_3^i(t, s)\,W^3))(0)$. Explicitly, $r^1(t, s) = (t + s_1, s_2, s_3 - (ts_1 s_2)/3 + t^2 s_3/6)$, $r^2(t, s) = (s_1, s_2 + t, (s_1^2 t/3) + s_3)$, hence

$$\mathscr{R}^1 = \partial/\partial s_1 - (s_1 s_2/3)\,\partial/\partial s_3, \qquad \mathscr{R}^2 = \partial/\partial s_2 + (s_1^2/3)\,\partial/\partial s_3.$$

The basic formulae to insure $\phi_*(x)\,W^i(x) \in \text{span}\{Y^1(\phi(x)), Y^2(\phi(x))\}$ for $i = 1, 2$, i.e., for a 2-distribution on $\mathbb{R}^3$, are Eqs. (2.7). To compute the coefficients β we use Eq. (2.5); here

$$(\text{ad } V^1, V^2)(x) = V^2(x) - 2x_1 V^3(x) \qquad \text{or} \quad \beta_{1,3}(x) = -2x_1,$$

$$(\text{ad}^\nu V^1, V^2)(x) = V^2(x) - (2x_1 - 2)\,V^3(x) \quad \text{or} \quad \beta_{\nu,3}(x) = -2x_1 + 2 \text{ if } \nu \geqslant 2.$$

Also

$$(\text{ad } V^2, V^3) = 0, (\text{ad } V^1, V^3) = 0 \qquad \text{or} \qquad \beta_{\nu_1,\nu_2,3} = 0, \nu_1 + \nu_2 \geqslant 1.$$

A major difficulty is that $\beta_{\nu,3}$ should be evaluated at $\phi(g(s))$ in Eqs. (2.7) but ϕ is to be determined. If we make the a priori choice $f_1(s) = s_1$, since the first component of g (see (2.15)) is s_1, then we will have the first component of ϕ, call it ϕ_1, such that $\phi_1(x) = x_1$. Then $\beta_{1,3}(\phi(g(s))) = -2s_1$, $\beta_{\nu,3}(\phi(g(s))) = -2s_1 + 2$ if $\nu \geqslant 2$ and Eqs. (2.7) become

$$\mathscr{R}^1 f_3(s) = -\mathscr{R}^1 f_2(s)[(-2s_1 + 2)(e^{s_1} - 1) - 2s_1]$$

$$\mathscr{R}^2 f_3(s) = -\mathscr{R}^2 f_2(s)[(-2s_1 + 2)(e^{s_1} - 1) - 2s_1].$$

A solution, with Jacobian at zero nonsingular, is

$$f_1(s) = s_1$$

$$f_2(s) = s_2 e^{-s_1}$$

$$f_3(s) = 2s_3 - (2s_1^2 s_2/3) + 2s_2 e^{-s_1} - 2s_2 + 2s_1 s_2.$$

This actually gives ϕ as the identity map relative to coordinates of the first kind generated by W^1, W^2, W^3 on the domain and coördinates of the second kind generated by V^1, V^2, V^3 on the range.

REFERENCES

1. H. J. SUSSMANN, Lie brackets and local controllability: A sufficient condition for scalar-input systems, *SIAM J. Control Optim.* **21** (1983), 686–713.
2. H. HERMES, Control systems which generate decomposable Lie algebras, *J. Differential Equations* **44** (1982), 166–187.
3. B. JAKUBCZYK AND W. RESPONDEK, On linearization of control systems, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astro. Phys.* **28** (1980), 517–522.
4. L. R. HUNT, R. SU, AND G. MEYER, Global transformations of nonlinear systems, *IEEE Trans. Automat. Control* **AC-28** No. 1, (1983), 24–31.
5. L. R. HUNT, R. SU, AND G. MEYER, Design for multi-input nonlinear systems, *in* "Differential Geometric Control Theory" (Brockett, Millman, Sussmann, Eds.) Birkhäuser, Boston, 1983; *Progr. in Math.* **27** (1983), 268–298.
6. A. J. KRENER, On the equivalence of control systems and the linearization of nonlinear systems, *SIAM J. Control* **11** (1973), 670–676.
7. R. HERMANN, Cartan connections and the equivalence problems for geometric structures, *Contrib. Differential Equations* **3** (1964), 199–248.
8. T. NAGANO, Linear differential systems with singularities and an application of transitive Lie algebras, *J. Math. Soc. Japan* **18** (1966), 398–404.
9. H. J. SUSSMANN, An extension of a theorem of Nagano on transitive Lie algebras, *Proc. Amer. Math. Soc.* **45** (1974), 349–356.
10. R. BRYANT, S. S. CHERN, AND P. A. GRIFFITHS, Exterior differential systems, NSF Regional Conference Lecture Notes, Albuquerque, N.M., 1981.
11. S. STERNBERG, "Lectures on Differential Geometry," Prentice-Hall, Englewood Cliffs, N.J., 1964.