Systems and Control Theory of Nonlinear Systems

H. Nijmeijer, A.J. van der Schaft

March 18, 2003

Addendum on MANIFOLDS, TANGENT SPACES, VECTOR FIELDS, LIE BRACKETS, DISTRIBUTIONS, FROBENIUS' THEOREM

As already alluded to in Chapter 1, the state space of a nonlinear system is usually not $\mathbb{R}^n$, nor an open part of $\mathbb{R}^n$. In fact, in Example 1.1 the state space variables are the angles θ_1, θ_2 and correspondingly the angular velocities $\dot{\theta}_1, \dot{\theta}_2$. Naturally the angles belong to the interval $[0, 2\pi]$ where the point 2π is identified with 0, i.e., the circle S^1. Thus the state space in this case is $S^1 \times S^1 \times \mathbb{R} \times \mathbb{R}$. These spaces are all examples of what is usually called a (smooth) *manifold*.
We will give the following definition of a manifold.

Definition 0.1. Let $f_1, \cdots, f_m, m \leq n$, be smooth functions on an open part V of $\mathbb{R}^n$. Define the set

$$M = \{x \in V | f_1(x) = \cdots = f_m(x) = 0\} \tag{1}$$

Suppose that the rank of the Jacobian matrix of $f = (f_1, \cdots, f_m)^T$

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix} =: \frac{\partial f}{\partial x}(x) \tag{2}$$

is m at each $x \in M$. Then M is a *manifold* of *dimension* $n - m$ (if M is non-empty).

Example 0.2. Every open subset V of $\mathbb{R}^n$ is a manifold of dimension n (Take $m = 0$).

Example 0.3. The circle S^1 is a manifold of dimension 1, since $S^1 = \{(x_1, x_2) \in \mathbb{R}^2 | x_1^2 + x_2^2 - 1 = 0\}$.

Example 0.4. Consider the group $O(n)$ of orthogonal (n, n)-matrices (i.e. $A \in O(n)$ satisfies $A^T A = I_n$). This is a manifold as can be seen as follows. Consider the set $g\ell(n)$ of *all* (n, n) matrices. Obviously $g\ell(n)$ can be identified with $\mathbb{R}^{n^2}$. Now define the map f from $g\ell(n)$ to the space of *symmetric* (n, n) matrices (which in the obvious way can be identified with $\mathbb{R}^{\frac{1}{2}n(n+1)}$) as follows

$$f(A) = A^T A \tag{3}$$

Then $O(n) = \{A \in g\ell(n) | f(A) = I_n\}$. It can be checked that the rank of the Jacobian matrix of f (seen as a map from $\mathbb{R}^{n^2}$ to $\mathbb{R}^{\frac{1}{2}n(n+1)}$) equals $\frac{1}{2}n(n+1)$ at every point $A \in O(n)$. (One may easily do this for $n = 2$ or $n = 3$.) Therefore $O(n)$ is a smooth manifold of dimension

$$n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1) \tag{4}$$

The basic feature of a manifold M of dimension $n - m$ is that it is "locally $\mathbb{R}^{n-m}$" in the following sense. Let M be given as in Definition 0.1 and let $x^o \in M$. By permuting the coordinates $x_1, \cdots, x_n$ for $\mathbb{R}^n$ if necessary we may assume that the (m, m) matrix

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & & \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_m} \end{bmatrix} \tag{5}$$

is non-singular at x^o. By the implicit function theorem there now exists a neighborhood $W_1 \subset \mathbb{R}^n$ of x^o, a neighborhood $W_2 \subset \mathbb{R}^{n-m}$ of $(x^o_{m+1}, \cdots, x^o_n)$, and a smooth map $g : W_2 \to \mathbb{R}^m$ such that $M \cap W_1$ equals

$$\{[g_1(x_{m+1}, \cdots, x_n), \cdots, g_m(x_{m+1}, \cdots, x_n), x_{m+1}, \cdots, x_n] \, | (x_{m+1}, \cdots, x_n) \in W_2\}$$

Then on $U := M \cap W_1$ we define *coordinate functions* $\varphi_i, \quad i = 1, \cdots, n - m$, by

$$\varphi_i \, [g_1(x_{m+1}, \cdots, x_n), \cdots, g_m(x_{m+1}, \cdots, x_n), x_{m+1}, \cdots, x_n] = x_{m+i} \tag{6}$$

U is called a *coordinate neighborhood* of x^o. In this way the neighborhood U of x^o becomes identified with an open part of $\mathbb{R}^{n-m}$. Since we can do this for every x^o we conclude that "locally $M = \mathbb{R}^{n-m}$".

Example 0.5. Consider the circle $S^1 = \{(x_1, x_2) | x_1^2 + x_2^2 - 1 = 0\}$. Take any point $x^o = (x^o_1, x^o_2) \in S^1$. If $x^o_1 \neq 0$ we have that $\frac{\partial}{\partial x_1}(x_1^2 + x_2^2 - 1)|_{(x^o_1, x^o_2)} \neq 0$, and thus we can solve for x_1, i.e. $x_1 = \pm\sqrt{1 - x_2^2}$ (with sign depending on the sign of x^o_1). The x_2-coordinate will serve as coordinate function (6) in both cases. Alternatively, if $x^o_2 \neq 0$ we solve for x_2, i.e. $x_2 = \pm\sqrt{1 - x_1^2}$, leading to neighborhoods $\tilde{U}_1$ and $\tilde{U}_2$ which are respectively in the upper- and the lower half-plane (see Figure 3.1). □

The idea is now to look at a manifold M as a space *on itself*, i.e. without reference to the embedding space $\mathbb{R}^n$. In particular we want to define *smooth* (i.e. C^∞) *functions* on M, and independence of them on M.

Definition 0.6. Let M be given as in Definition 0.1, and let $h : M \to \mathbb{R}$ be a function on M. Let U be a coordinate neighborhood of $x^o \in M$ as in (6). Then h is *smooth on* U if the function

$$h \, [g_1(x_{m+1}, \cdots, x_n), \cdots, g_m(x_{m+1}, \cdots, x_n), x_{m+1}, \cdots, x_n] \tag{7}$$

depends smoothly on its arguments $x_{m+1}, \cdots, x_n$. (It can be easily seen that this does not depend on the particular choice of g.) The function h is *smooth on* M if it

Figure 1: Coordinate neighborhoods for S^1

is smooth on a covering set of coordinate neighborhoods of M.
Let $h_1, \cdots, h_k$ be smooth functions on M. Then $h_1, \cdots, h_k$ are called *independent* on U if the functions

$$h_i\left[g_1(x_{m+1}, \cdots, x_n), \cdots, g_m(x_{m+1}, \cdots, x_n), x_{m+1}, \cdots, x_n\right], \quad i = 1, \cdots, k \tag{8}$$

are independent as functions of $x_{m+1}, \cdots, x_n$. The functions $h_1, \cdots, h_k$ are independent on M if they are so on a covering set of coordinate neighborhoods of M.

With the aid of the above definition the notion of a *coordinate neighborhood* and of *coordinate functions* defined on it can be immediately generalized. Indeed, any open subset V of M with $n - m$ ($= \dim M$) independent smooth functions $(h_1, \cdots, h_{n-m})$ defined on it defines a coordinate neighborhood and coordinate functions for M, or, briefly, a *coordinate system*

$$(V, (h_1, \cdots, h_{n-m})) \tag{9}$$

Example 0.7. Let $M = \mathbb{R}^2 \setminus \{(x_1, x_2) | x_1 \geq 0, x_2 = 0\}$, and let $V = M$. Define the coordinate functions

$$\begin{array}{rcl} h_1(x_1, x_2) & = & \sqrt{x_1^2 + x_2^2} \\ h_2(x_1, x_2) & = & \arctan \frac{x_1}{x_2} \end{array} \tag{10}$$

i.e., polar coordinates.

Remark In Chapter 2 (as in most treatises) manifolds are defined in a more abstract and intrinsic way. Indeed one considers an abstract topological (Hausdorff) space M, and by identifying M locally with a linear space $\mathbb{R}^{n-m}$ one provides M with a "differentiable structure", i.e., a notion of what are the smooth functions on M. However it can be shown that to every manifold M defined in this abstract way there corresponds some $\mathbb{R}^n$ and functions $f_1, \cdots, f_m$ defined on $\mathbb{R}^n$ such that M is given as in Definition 0.1. In this sense the definition given in this Addendum does not restrict generality.

Also the following notion will be needed in the sequel.

Definition 0.8. Let M be a manifold of dimension n. A subset $P \subset M$ is called a *submanifold* of dimension $k < n$ if for each $p \in P$ there exists a coordinate system $(V, \varphi_1, \cdots, \varphi_n)$ for M about p such that

$$P \cap V = \{q \in V | \varphi_i(q) = \varphi_i(p), \quad i = k+1, \cdots, n\} \tag{11}$$

Example 0.9. Consider Example 0.7. All the circles with positive radius are submanifolds of M (see Figure 3.2).

Figure 2: Submanifold

Notice that a submanifold P of a manifold M is a manifold in its own right, with coordinate system $(P \cap V, (\varphi_1, \cdots, \varphi_k))$. Now we proceed to the definition of *tangent space*.

Definition 0.10. Let M be an $(n-m)$-dimensional manifold as given in Definition 0.1. Let $x^o \in M$, then the tangent space $T_{x_0}M$ at x_0 to the manifold M is given as the linear space

$$T_{x_0}M = \left\{ z \in \mathbb{R}^n | \frac{\partial f}{\partial x}(x_0) z = 0 \right\} = \ker \frac{\partial f}{\partial x}(x_0) \tag{12}$$

(Notice that because the rank of $\frac{\partial f}{\partial x}(x_0)$ equals m the dimension of $T_{x_0}M$ equals $n - m$, i.e. the dimension of the manifold M.) Furthermore the *tangent bundle* TM is defined as the manifold

$$TM = \{(x, z) \in V \times \mathbb{R}^n | f_1(x) = \cdots = f_m(x) = 0, \frac{\partial f}{\partial x}(x) z = 0\} \tag{13}$$

and equals $\bigcup_{x \in M} T_x M$.

Example 0.11. Consider Example 0.4. We want to compute $T_{I_n}O(n)$, i.e. the tangent space to the manifold $O(n)$ in the point $I_n \in O(n)$. In principle this can be done using (12) differentiating the mapping f defined in (3), regarded as a map from $\mathbb{R}^{n^2}$ to $\mathbb{R}^{\frac{1}{2}n(n+1)}$. A more slick way is to compute (3) for $\epsilon \in \mathbb{R}, X \in g\ell(n)$

$$\frac{f(I_n + \epsilon X) - f(I_n)}{\epsilon} = (X^T + X) + \epsilon X^T X \tag{14}$$

and to conclude that $T_{I_n}O(n)$ can be identified with the space of all (n, n) matrices X satisfying

$$X^T + X = 0 \tag{15}$$

i.e. the space of $n \times n$ skew-symmetric matrices. Notice that the dimension of this linear space is $\frac{1}{2}n(n-1)$, in accordance with the dimension of the manifold $O(n)$, cf. (4).

Consider now a vector $X_p \in T_pM, p \in M$. Clearly, we can define a smooth mapping

$$c : (-\epsilon, \epsilon) \longrightarrow M, \quad \epsilon > 0, \quad c(0) = p \tag{16}$$

such that $c'(0) = \frac{dc}{dt}(0) = X_p$. For any smooth function $h : M \to \mathbb{R}$ we now define the *derivative* of h *in the direction* X_p at the point $p \in M$ as

$$X_p(h) := \frac{d}{dt}h(c(t))|_{t=o} \tag{17}$$

(By the chain-rule it is easily seen that the definition of $X_p(h)$ is independent of the particular choice of c satisfying $x(0) = p$ and $c'(0) = X_p$.) Hence any tangent vector $X_p \in T_pM$ can be regarded as an *operator* acting on smooth functions on M, i.e.

$$X_p : C^\infty(p) \to \mathbb{R} \tag{18}$$

where $C^\infty(p)$ denotes the set of smooth functions defined on a neighborhood of p in M, and where $X_p(h)$ for any $h \in C^\infty(p)$ is defined as in (17).
One readily verifies the properties

$$\begin{aligned} X_p(\alpha_1 h_1 + \alpha_2 h_2) &= \alpha_1 X_p(h_1) + \alpha_2 X_p(h_2) \quad h_1, h_2 \in C^\infty(p), \\ X_p(h_1 \cdot h_2) &= X_p(h_1)h_2(p) + h_1(p)X_p(h_2) \quad , \alpha_1, \alpha_2 \in \mathbb{R} \end{aligned} \tag{19}$$

In general, we can define the linear space of *all* operators $D : C^\infty(p) \to \mathbb{R}$ satisfying property (19), i.e.

$$\begin{aligned} D(\alpha_1 h_1 + \alpha_2 h_2) &= \alpha_1 D(h_1) + \alpha_2 D(h_2) \\ D(h_1 \cdot h_2) &= D(h_1)h_2(p) + h_1(p)D(h_2) \end{aligned} \tag{20}$$

This space is called the space of *derivations* at p, and we conclude that every tangent vector $X_p \in T_pM$ can be identified with an element in this space. Conversely, we

will now show that every derivation D can be identified with some tangent vector X_p.
Indeed, let $(V, (\varphi_1, \cdots, \varphi_n))$ be a coordinate system for M around the point p. Without loss of generality we may assume that $\varphi_i(p) = 0, i = 1, \cdots, n$. For simplicity of notation let us denote the local coordinates defined by $\varphi_1, \cdots, \varphi_n$ as $z_1, \cdots, z_n$.

Lemma 0.12. Let f be a smooth function defined on a neighborhood of $0 \in \mathbb{R}^n$. Then on this neighborhood

$$f(x_1, \cdots, x_n) = \sum_{i=1}^{n} x_i g_i(x_1, \cdots, x_n) + f(0) \tag{21}$$

for certain smooth functions $g_1, \cdots, g_n$ satisfying $g_i(0) = \frac{\partial f}{\partial x_i}(0)$.

Proof Define $h(t, x_1, \cdots, x_n) = f(tx_1, \cdots, tx_n), 0 \leq t \leq 1$. Since $h(1, x_1 \cdots, x_n) = f(x_1, \cdots, x_n)$ and $h(0, x_1, \cdots, x_n) = f(0, \cdots, 0)$ we have

$$f(x_1, \cdots, x_n) - f(0, \cdots, 0) = \int_0^1 \frac{\partial h}{\partial t}(t, x_1, \cdots, x_n)dt =$$

$$= \int_0^1 \sum_{i=1}^{n} \frac{\partial f}{\partial r_i}(tx_1, \cdots, tx_n) \cdot x_i dt = \sum_{i=1}^{n} x_i \int_0^1 \frac{\partial f}{\partial r_i}(tx_1, \cdots, tx_n)dt$$

$$=: \sum_{i=1}^{n} x_i g_i(x)$$

□

Now consider an arbitrary derivation D at p, i.e. D satisfies (20) for all $h_1, h_2 \in C^\infty(p)$ and $\alpha_1, \alpha_2 \in \mathbb{R}$. Note that by the second line of (20) we have

$$D(1) = D(1.1) = D(1).1 + 1.D(1) = 2D(1) \tag{22}$$

so that $D(1) = 0$, and thus $D(c) = 0$ for every constant c. Now take an arbitrary element $h \in C^\infty(p)$. By Lemma 0.12 we may write in the coordinates $z_1, \cdots, z_n$

$$h(z_1, \cdots, z_n) = \sum_{i=1}^{n} z_i g_i(z_1, \cdots, z_n) + h(0)$$

where $g_i(0) = \frac{\partial h}{\partial z_i}(0)$. Then by application of (20)

$$\begin{aligned} D(h) &= \sum_{i=1}^{n} [D(z_i)g_i(0) + 0 \cdot D(g_i)] \\ &= \sum_{i=1}^{n} D(z_i)\frac{\partial h}{\partial z_i}(0) \end{aligned} \tag{23}$$

From this it follows that the space of derivations at p is n-dimensional, and thus can be identified with T_pM. Indeed, we define the n derivations $D_1, \cdots, D_n$ as

$$D_i(h) = \frac{\partial h}{\partial z_i}(0) \tag{24}$$

(It is easily seen that D_i satisfies (20).) It follows that $D_1, \cdots, D_n$ are independent since

$$D_i(z_j) = \delta_{ij}, \quad i, j \in \underline{n} \quad (z_1, \cdots, z_n \text{ coordinate functions })$$

and by (23) we have

$$D(h) = \sum_{i=1}^{n} \alpha_i D_i(h), \quad \alpha_i = D(z_i) \in \mathbb{R}, \tag{25}$$

so that actually $D_1, \cdots, D_n$ is a *basis* for the linear space of derivations at p. Usually this basis $D_1, \cdots, D_n$ for the linear space of derivations at p corresponding to the above coordinate system $(V, (\varphi_1, \cdots, \varphi_n)) = (V, z_1, \cdots, z_n)$ is denoted as

$$\left.\frac{\partial}{\partial \varphi_1}\right|_p, \cdots, \left.\frac{\partial}{\partial \varphi_n}\right|_p, \text{ or, } \left.\frac{\partial}{\partial z_1}\right|_p, \cdots, \left.\frac{\partial}{\partial z_n}\right|_p \tag{26}$$

By the above identification of T_pM with the space of derivations at p this basis also serves as a basis for T_pM. Note that this basis crucially depends on the choice of the coordinate system. Indeed, let $(U, (\psi_1, \cdots, \psi_m))$ be another coordinate system for M around p. Then one obtains another basis $\left(\left.\frac{\partial}{\partial \psi_1}\right|_p, \cdots, \left.\frac{\partial}{\partial \psi_n}\right|_p\right)$ for T_pM. A straightforward calculation yields that if $X_p \in T_pM$ is represented in the first basis as $\sum_{i=1}^{n} \alpha_i \frac{\partial}{\partial \varphi_i} |_p$ and in the second basis as $\sum_{i=1}^{n} \beta_i \frac{\partial}{\partial \psi_i} |_p$, then the column vectors $\alpha = (\alpha_1, \cdots, \alpha_n)^T$ and $\beta = (\beta_1, \cdots, \beta_n)^T$ are related as

$$\beta = \frac{\partial S}{\partial x}(\varphi(p))\alpha \tag{27}$$

where $S(x) := \psi \circ \varphi^{-1}(x)$ is the *coordinate transformation* $(\psi = (\psi_1, \cdots, \psi_n)^T, \varphi = (\varphi_1, \cdots, \varphi_n)^T)$. (We will rederive this later in the context of coordinate transformations for differential equations.) We conclude that the choice of a local coordinate system $(V, (\varphi_1, \cdots, \varphi_n))$ uniquely specifies a basis, and hence a system of *linear* coordinates, for any tangent space T_pM with $p \in V$.

In the same manner also the tangent bundle TM becomes endowed with a natural set of coordinates. Indeed, an arbitrary point in TM is given as X_p where $p \in M$ and $X_p \in T_pM$, cf. (13). Given the local coordinate system $(V, (\varphi_1, \cdots, \varphi_n))$ we may write for $p \in V, X_p = \sum_{i=1}^{n} \alpha_i \frac{\partial}{\partial \varphi_i} |_p$, and thus we naturally assign to the point X_p in TM the coordinate values

$$(\varphi_1(p), \cdots, \varphi_n(p), \alpha_1, \cdots, \alpha_n) \tag{28}$$

Recall that in Definition 0.6 we have defined what we mean by a smooth *function* on a manifold M. Quite similarly we now define what we mean by a *smooth* mapping

$$F : M_1 \to M_2 \tag{29}$$

with M_1 and M_2 manifolds. Indeed, let M_1 and M_2 be manifolds of dimension n_1 and n_2, respectively. Then for any $p \in M_1$ there exist local coordinate systems $(U, (\varphi_1, \cdots, \varphi_{n_1}))$ for p and $(V, (\psi_1, \cdots, \psi_{n_2}))$ for $F(p) \in M_2$. We now require that the maps

$$\hat{F} := \psi \circ F \circ \varphi^{-1} : \varphi(U) \subset \mathbb{R}^{n_1} \to \psi(V) \subset \mathbb{R}^{n_2} \tag{30}$$

where $\varphi = (\varphi_1, \cdots, \varphi_{n_1})^T, \psi = (\psi_1, \cdots, \psi_{n_2})^T$, are smooth maps. (In the ordinary sense of smooth mappings from (an open part of) $\mathbb{R}^{n_1}$ to $\mathbb{R}^{n_2}$.) Note that $\hat{F}$ is nothing else than the *local coordinate expression* of the map $F : M \to N$. Also note that in a similar way we may rephrase the definition of a smooth function $h : M \to \mathbb{R}$ (cf. (7)) by requiring that the functions

$$\hat{h} := h \circ \varphi^{-1} : \varphi(U) \subset \mathbb{R}^n \to \mathbb{R} \tag{31}$$

are smooth, where $(U, (\varphi_1, \cdots, \varphi_n))$ is a local coordinate system for M.
Let now $F : M \to N$ be a smooth map. Then we define a linear map (called the *tangent map* of F at $p \in M$)

$$F_{*p} : T_pM \to T_{F(p)}N \tag{32}$$

as follows. Let $X_p \in T_pM$. For any $f \in C^\infty(F(p))$ set

$$F_{*p}X_p(f) = X_p(f \circ F) \tag{33}$$

where $X_p \in T_pM$ is identified with the corresponding derivation at $p \in M$. It is easily verified that $F_{*p}X_p$ defines a derivation at $F(p) \in N$, and thus can be regarded as an element of $T_{F(p)}N$. Indeed, the only non-trivial verification is (let $f, g \in C^\infty(F(p))$)

$$\begin{gathered} F_{*p}X_p(fg) = X_p\left[(f \circ F)(g \circ F)\right] = \\ \\ X_p(f \circ F) \cdot g[F(p)] + f[F(p)] \cdot X_p(g \circ F) = \\ \\ F_{*p}X_p(f) \cdot g[F(p)] + f[F(p)] \cdot F_{*p}X_p(g) \end{gathered} \tag{34}$$

The following properties are immediate.

Proposition 0.13. (a) If $H = G \circ F$ is a composition of smooth maps, then $H_{*p} = G_{*F(p)} \circ F_{*p}$

(b) Let $id : M \to M$ be the identity mapping, then $id_{*p} : T_pM \to T_pM$ is the identity matrix I for any $p \in M$.

(c) For a *diffeomorphism* $F : M \to N$ (i.e., F is smooth, $F^{-1} : N \to M$ exists and is smooth) we have

$$I = F^{-1}_{*F(p)} \circ F_{*p} = F_{*p} \circ F^{-1}_{*F(p)}$$

Hence $\dim T_pM = \dim T_{F(p)}N = \mathrm{rank}(F_{*p}) = \mathrm{rank}\left(F^{-1}_{*F(p)}\right)$.

Finally, we define the *tangent mapping*

$$F_* : TM_1 \to TM_2 \tag{35}$$

as the union of all tangent mappings $F_{*p} : T_pM_1 \to T_{F(p)}M_2$ for $p \in M_1$. Let $x = (x_1, \cdots, x_{n_1})$ be local coordinates for M, and $z = (z_1, \cdots, z_{n_2})$ for M_2, then in the above defined natural coordinates $(x, v) = (x_1, \cdots, x_{n_1}, v_1, \cdots, v_{n_1})$ for M_1, respectively $(z, w) = (z_1, \cdots, z_{n_2}, w_1, \cdots, w_{n_2})$ for M_2, the tangent mapping F_* is simply given as

$$F_*(x, v) = (\hat{F}(x), \frac{\partial \hat{F}}{\partial x}(x)v) \tag{36}$$

where $\hat{F}$ is the local coordinate expression of F. Usually we will omit the caret $\wedge$, and simply write $F(x)$ instead of $\hat{F}(x)$.
Now we come to the definition of (smooth) *vector fields* on a manifold M.

Definition 0.14. A (smooth) *vector field* X on a manifold M is defined as a smooth mapping

$$X : M \longrightarrow TM \tag{37}$$

satisfying $\pi(X(p)) = p, \forall p \in M$, where $\pi : TM \to M$ is the canonical projection mapping $(p, X_p) \in TM$ to $p \in M$.

Thus a vector field X on M assigns to every point $p \in M$ an element of T_pM :

$$X(p) \in T_pM \tag{38}$$

Let now $(U, \varphi_1, \cdots, \varphi_n) = (U, x_1, \cdots, x_n)$ be a coordinate system for M. For every $p \in U$ this yields a basis $\left\{ \frac{\partial}{\partial x_1}\Big|_p, \cdots, \frac{\partial}{\partial x_n}\Big|_p \right\}$ for T_pM. It thus follows that locally on U the vector field X can be expressed by a column-vector

$$X(x) = [X_1(x_1, \cdots, x_n), \cdots, X_n(x_1, \cdots, x_n)]^T \tag{39}$$

where $[X_1(p), \cdots, X_n(p)]^T$ is the vector of coefficients of the tangent vector $X(p)$ in T_pM.
(In fact, there is some abuse of notation, since in the corresponding natural coordinates $(x_1, \cdots, x_n, v_1, \cdots, v_n)$ for TM the mapping $X : M \to TM$ is given as

$$(x_1, \cdots, x_n)^T \longmapsto (x_1, \cdots, x_n, X_1(x_1, \cdots, x_n), \cdots, X_n(x_1, \cdots, x_n))^T \tag{40}$$

and thus the coordinates of the base point in the right-hand side of (40) are “forgotten”.) It follows that in local coordinates $x_1, \cdots, x_n$ a vector field X corresponds to the n-dimensional set of first-order *differential equations*

$$\begin{aligned} \dot{x}_1 &= X_1(x_1, \cdots, x_n) \\ &\vdots \\ \dot{x}_n &= X_n(x_1, \cdots, x_n) \end{aligned} \tag{41}$$

Using this machinery we are now able to give a *coordinate-free definition* of a nonlinear state space system

$$\begin{array}{rcll} \dot{x} & = & f(x)+g(x)u & ,u \in \mathbb{R}^m \quad ,x \in \mathcal{X}, \\ y & = & h(x) & ,y \in \mathbb{R}^p, \end{array} \tag{42}$$

living on a state space $\mathcal{X}$ that is a *manifold.* Indeed, $f(x)$ is the local coordinate expression of a *vector field* on $\mathcal{X}$ (called the *drift* vector field), and also the columns of $g(x)$ are local coordinate expressions of vector fields on $\mathcal{X}$ (the input-vector fields), while h is a smooth mapping from $\mathcal{X}$ to $\mathbb{R}^p$.

In the case of the more general form

$$\begin{array}{rcl} \dot{x} & = & f(x,u) \\ y & = & h(x,u) \end{array} \tag{43}$$

one considers f to be a smooth bundle mapping $f : M \times \mathbb{R}^m \to TM$, and $h : M \times \mathbb{R}^m \to \mathbb{R}^p$ a smooth mapping.

The next important notion is that of a *Lie bracket* of vector fields. For X and Y any two vectorfields on M we define the *Lie bracket* of X and Y, denoted $[X,Y]$, by setting

$$[X,Y]_p(f) = X_p(Y(f)) - Y_p(X(f)) \tag{44}$$

for every function $f : M \to \mathbb{R}$. It can be checked that $[X,Y]_p$ belongs to the space of derivations at p. Indeed, the first part of (20) is trivial, while the second part follows from

$$\begin{array}{rcl} [X,Y]_p(fg) & = & X_p(Y(fg)) - Y_p(X(fg)) = \\ \\ & = & X_p\{Y(f)\cdot g + f\cdot Y(g)\} - Y_p\{X(f)\cdot g + f \cdot X(g)\} = \\ \\ & = & X_p[Y(f)]g(p) + Y_p(f)X_p(g) + X_p(f)Y_p(g) + f(p)X_p(Y(g)) \\ \\ & - & Y_p(X(f))g(p) - X_p(f)Y_p(g) - Y_p(f)X_p(g) - f(p)Y_p(X(g)) \\ \\ & = & [X,Y]_p(f)\cdot g(p) + f(p)\cdot [X,Y]_p(g) \end{array} \tag{45}$$

Thus $[X,Y]_p$ can be uniquely identified with an element in the tangent space T_pM, and $[X,Y]$ defines a new vectorfield on M.
In local coordinates the Lie bracket takes the following form:

Proposition 0.15. Let X and Y be vectorfields on M, given in local coordinates $(x_1,\cdots,x_n)$ as $X(x) = (X_1(x),\cdots,X_n(x))^T$ and $Y(x) = (Y_1(x),\cdots,Y_n(x))^T$. Then the local coordinate expression of $[X,Y]$ is given as

$$[X,Y](x) = \frac{\partial Y}{\partial x}(x)X(x) - \frac{\partial X}{\partial x}(x)Y(x) \tag{46}$$

with $\frac{\partial Y}{\partial x}, \frac{\partial X}{\partial x}$ denoting the Jacobian matrices.

Proof Compute for any $j = 1, \cdots, n$

$$\begin{aligned} [X,Y]_p(x_j) &= X_p(Y(x_j)) - Y_p(X(x_j)) = \\ &= X_p(Y_j) - Y_p(X_j) = \sum_{i=1}^{n} \left[\frac{\partial Y_j}{\partial x_i} X_i - \frac{\partial X_j}{\partial x_i} Y_i \right] (x(p)) \end{aligned} \tag{47}$$

Since $[X,Y]_p(x_j)$ is the j-th component of $[X,Y]_p$ in these coordinates the result follows. □

It readily follows from the coordinate expression (46) (or from the definition (44)) that the Lie bracket satisfies the following properties

$$\begin{array}{llll} (a) & [fX, gY] & = & fg[X,Y] + f \cdot L_X g \cdot Y - g \cdot L_Y f \cdot X \quad f, g \in C^\infty(M) \\ (b) & [X,Y] & = & -[Y,X] \\ (c) & [X, Y_1 + Y_2] & = & [X, Y_1] + [X, Y_2] \end{array} \tag{48}$$

Furthermore, the following property can be checked

$$[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0 \quad X, Y, Z \text{ vectorfields.} \tag{49}$$

Let us now come to the question of generalizing the linear notion of an A-invariant subspace to the nonlinear case. The proper generalization of the concept of subspace will be that of a *"distribution"*.

Definition 0.16. Let M be an n-dimensional manifold. A *distribution* D on M is a map which assigns to each $p \in M$ a linear subspace $D(p)$ of the tangent space T_pM. D is called a smooth distribution if around any point these subspaces are spanned by a set of vectorfields; i.e., for each $p \in M$ there exists a neighborhood U of p and a set of (smooth) vectorfields $X_i, i \in I$, with I some index set, such that

$$D(q) = \text{span}\{X_i(q) | i \in I\}, \quad q \in U. \tag{50}$$

In the sequel "distribution" will always mean *"smooth* distribution".

The definition of *invariance* of a distribution is defined as follows. We say that a vector field Z belongs to a distribution D, denoted as $Z \in D$, if $Z(p) \in D(p)$ for all $p \in M$.

Definition 0.17. Let D be a vector field on M. Then D is *invariant* under f if

$$[f, X] \in D, \quad \forall \text{ vectorfields } X \in D \tag{51}$$

It follows from the properties of the Lie bracket (cf. (48(a)) that if D is locally given as span$\{X_i | i \in I\}$ then (51) holds iff $[f, X_i] \in D, i \in I$.
We will use throughout the following shorthand notation for (51)

$$[f, D] \subset D \tag{52}$$

Example 0.18. Let $M = \mathbb{R}^n$ with natural linear coordinates $x_1, \cdots, x_n$. Define the distribution $D = \text{span}\left\{\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_k}\right\}, k \leq n$, and let $f(x)$ be the linear vector field Ax. Then

$$\left[Ax, \frac{\partial}{\partial x_i}\right] = 0 - Ae_i = i - th \text{ column of A} \quad (e_i = i - th \text{ basis vector }) \tag{53}$$

where the $i-th$ column of A is interpreted as a *constant* vector field on $M = \mathbb{R}^n$. It immediately follows that (53) is equivalent to $AV \subset V$, where $V = \text{span}\{e_1, \cdots, e_k\}$.

Alternatively, let f be a general nonlinear vector field, given in the local coordinates as above by $f(x) = [f_1(x), \cdots, f_n(x)]^T$. Then invariance with regard to $D = \text{span}\{\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_k}\}$ implies

$$[f, \frac{\partial}{\partial x_i}](X) = 0 - \frac{\partial f}{\partial x}(x)e_i = \left(\frac{\partial f_1}{\partial x_i}, \cdots, \frac{\partial f_n}{\partial x_i}\right)^T (x), \quad i = 1, \cdots, k \tag{54}$$

and thus (52) amounts to

$$\frac{\partial f_j}{\partial x_i}(x) = 0, \quad j = k+1, \cdots, n, i = 1, \cdots, k, \tag{55}$$

i.e., the set of differential equations corresponding to f has the block-triangular structure

$$\begin{array}{rcl} \dot{x}_1 & = & f_1(x_1, \cdots, x_k, x_{k+1}, \cdots, x_n) \\ \dot{x}_k & = & f_k(x_1, \cdots, x_k, x_{k+1}, \cdots, x_n) \\ \dot{x}_{k+1} & = & f_{k+1}(x_{k+1}, \cdots, x_n) \\ \dot{x}_n & = & f_n(x_{k+1}, \cdots, x_n) \end{array} \tag{56}$$

or in shorthand notation, denoting $x^1 = (x_1, \cdots, x_k)^T, x^2 = (x_{k+1}, \cdots, x_n)^T$, $f^1 = (f_1, \cdots, f_k)^T, f^2 = (f_{k+1}, \cdots, f_n)^T$

$$\begin{array}{rcl} \dot{x}^1 & = & f^1(x^1, x^2) \\ \dot{x}^2 & = & f^2(x^2) \end{array} \tag{57}$$

This local triangular form is what we are aiming for, and so the logical question is under what conditions a distribution D can be brought into the form $\text{span}\{\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_k}\}$ by a suitable choice of local coordinates $(x_1, \cdots, x_n)$. This question is largely answered by the following theorem which we state without proof. First, we call a distribution D *involutive* if

$$[X, Y] \in D, \quad \text{for all } X, Y \text{ vectorfields in } D \tag{58}$$

(i.e., D is invariant with respect to all vectorfields contained in D)

Theorem 0.19. *(Frobenius) Let D be a distribution on a manifold M. Suppose that*
(i) D is involutive
(ii) D is constant dimensional, i.e. $\dim D(p), p \in M$, *is constant.*
Then for any $x^o \in M$ there exists a coordinate system $(U, x_1, \cdots, x_n)$ such that

$$D(q) = \text{span}\left\{ \left.\frac{\partial}{\partial x_1}\right|_q, \cdots, \left.\frac{\partial}{\partial x_k}\right|_q \right\}, \quad q \in U \tag{59}$$

(with $k = \dim D$). For brevity we will simply write

$$D = \text{span}\left\{ \frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_k} \right\} \quad (\text{ on } U) \tag{60}$$

Remark It is easily seen that (i) is a *necessary* condition in order to bring D into the form (59). Indeed,

$$\left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right] = 0, \quad i, j = 1, \cdots, n \tag{61}$$

and in view of (48) (a), (c) this implies that D given as in (59) is involutive. Also (ii) is clearly a necessary condition (if we assume that M is connected.)

Example 0.20. Consider on $M = \{(x_1, x_2, x_3)^T \in \mathbb{R}^3 | x_i > 0, i = 1, 2, 3,\}$ the distribution $D(x) = \text{span}\{X_1(x), X_2(x)\}$ where

$$X_1(x) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad X_2(x) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Then

$$[X_1, X_2](x) = 0 - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = -X_2(x)$$

and it follows that D is involutive, as well as constant-dimensional. How do we construct new coordinates z_1, z_2, z_3 such that $D = \text{span}\{\frac{\partial}{\partial z_1}, \frac{\partial}{\partial z_2}\}$? (By Frobenius' theorem this is possible!)
This is solved as follows. If $D = \text{span}\{\frac{\partial}{\partial z_1}, \frac{\partial}{\partial z_2}\}$ then clearly $\frac{\partial}{\partial z_1} z_3 = \frac{\partial}{\partial z_2} z_3 = 0$, and thus we have to find $\varphi(x)$ such that

$$L_{X_1}\varphi(x) = 0, \quad L_{X_2}\varphi(x) = 0 \tag{62}$$

and then set $z_3 := \varphi(x)$. $L_{X_2}\varphi = 0$ simply means that φ should only depend on x_1 and x_2, and a possible solution of $L_{X_1}\varphi(x) = 0$ is

$$\varphi(x_1, x_2) = \ell n \frac{x_1}{x_2} \tag{63}$$

Denote $z_3 := \varphi(x), z_1 := x_1, z_2 := x_3$ (for instance), then it is checked that z_1, z_2, z_3 are independent, and it follows that $D = \text{span}\{\frac{\partial}{\partial z_1}, \frac{\partial}{\partial z_2}\}$. (N.B. Check that φ as given in (63) is not the only solution of (62); for example also $\varphi(x) = \arctan \frac{x_2}{x_1}$ does the job.)