INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONES: 224281/2/3/4/5/6
CABLE: CENTRATOM - TELEX 460392 - I

SMR.201/18

SECOND WORKSHOP ON MATHEMATICS IN INDUSTRY

(2 - 27 February 1987)

NONLINEAR CONTROL SYSTEMS: AN INTRODUCTION.

A. ISIDORI
Dipartimento di Informatica e Sistematica
Università degli Studi di Roma "La Sapienza"
Via Eudossiano, 18
00184 Roma
Italy

Lecture Notes in Control and Information Sciences

Edited by M. Thoma

Vol. 22: Optimization Techniques
Proceedings of the 9th IFIP Conference on Optimization Techniques,
Warsaw, September 4–8, 1979
Part I
Edited by K. Iracki, K. Malanowski, S. Walukiewicz
XVI, 569 pages. 1980

Vol. 23: Optimization Techniques
Proceedings of the 9th IFIP Conference on Optimization Techniques,
Warsaw, September 4–8, 1979
Part 2
Edited by K. Iracki, K. Malanowski, S. Walukiewicz
XV, 621 pages. 1980

Vol. 24: Methods and Applications in Adaptive Control
Proceedings of an International Symposium
Bochum, 1980
Edited by H. Unbehauen
VI, 309 pages. 1980

Vol. 25: Stochastic Differential Systems – Filtering and Control
Proceedings of the IFIP-WG7/1 Working Conference
Vilnius, Lithuania, USSR, Aug. 28 – Sept. 2, 1978
Edited by B. Grigelionis
X, 362 pages. 1980

Vol. 26: D. L. Iglehart, G. S. Shedler
Regenerative Simulation of Response Times in Networks of Queues
XII, 204 pages. 1980

Vol. 27: D. H. Jacobson, D. H. Martin, M. Pachter, T. Geveci
Extensions of Linear-Quadratic Control Theory
XI, 288 pages. 1980

Vol. 28: Analysis and Optimization of Systems
Proceedings of the Fourth International Conference on Analysis and Optimization of Systems
Versailles, December 16–19, 1980
Edited by A. Bensoussan and J. L. Lions
XIV, 999 pages. 1980

Vol. 29: M. Vidyasagar,
Input-Output Analysis of Large-Scale Interconnected Systems –
Decomposition, Well-Posedness and Stability
VI, 221 pages. 1981

Vol. 30: Optimization and Optimal Control
Proceedings of a Conference Held at Oberwolfach, March 16–22, 1980
Edited by A. Auslender, W. Oettli, and J. Stoer
VIII, 254 pages. 1981

Vol. 31: Berc Rustem
Projection Methods in Constrained Optimisation and Applications to Optimal Policy Decisions
XV, 315 pages. 1981

Vol. 32: Tsuyoshi Matsuo,
Realization Theory of Continuous-Time Dynamical Systems
VI, 329 pages, 1981

Vol. 33: Peter Dransfield
Hydraulic Control Systems – Design and Analysis of Their Dynamics
VII, 227 pages, 1981

Vol. 34: H.W. Knobloch
Higher Order Necessary Conditions in Optimal Control Theory
V, 173 pages, 1981

Vol. 35: Global Modelling
Proceedings of the IFIP-WG 7/1 Working Conference Dubrovnik, Yugoslavia,
Sept. 1–5, 1980
Edited by S. Krčevinac
VIII, 232 pages, 1981

Vol. 36: Stochastic Differential Systems
Proceedings of the 3rd IFIP-WG 7/1 Working Conference
Visegrád, Hungary, Sept. 15–20, 1980
Edited by M. Arató, D. Vermes, A.V. Balakrishnan
VI, 238 pages, 1981

Vol. 37: Rüdiger Schmidt
Advances in Nonlinear Parameter Optimization
VI, 159 pages, 1982

Vol. 38: System Modeling and Optimization
Proceedings of the 10th IFIP Conference
New York City, USA, Aug. 31 – Sept. 4, 1981
Edited by R.F. Drenick and F. Kozin
XI, 894 pages. 1982

Vol. 39: Feedback Control of Linear and Nonlinear Systems
Proceedings of the Joint Workshop on Feedback and Synthesis of Linear and Nonlinear Systems
Bielefeld/Rom
XIII, 284 pages. 1982

Vol. 40: Y.S. Hung, A.G.J. MacFarlane
Multivariable Feedback: A Quasi-Classical Approach
X, 182 pages. 1982

Vol. 41: M. Gössel
Nonlinear Time-Discrete Systems – A General Approach by Nonlinear Superposition
VIII, 112 pages. 1982

Vol. 42: Advances in Filtering and Optimal Stochastic Control
Proceedings of the IFIP-WG 7/1 Working Conference
Cocoyoc, Mexico, February 1–6, 1982
VIII, 391 pages. 1982

For information about Vols. 1–21 please contact your bookseller or Springer-Verlag.

Lecture Notes in Control and Information Sciences

Edited by M.Thoma

72

A. Isidori

Nonlinear Control Systems: An Introduction

Springer-Verlag
Berlin Heidelberg New York Tokyo

CHAPTER I

LOCAL DECOMPOSITIONS OF CONTROL SYSTEMS

*. Introduction

In this section we review some basic results from the theory of linear systems, with the purpose of describing some fundamental properties which find close analogues in the theory of nonlinear systems.

Usually, a linear control system is described by equations of the form

$$\dot{x} = Ax + Bu$$
$$y = Cx$$

in which the state x belongs to X, an n-dimensional vector space and the input u and the output y belong respectively to an m-dimensional vector space U and ℓ-dimensional vector space Y. The mappings $A : X \to X$, $B : U \to X$, $C : X \to Y$ are linear mappings.

Suppose that there exists a d-dimensional subspace V of X with the following property:

(i) V is invariant under the mapping A, i.e. is such that $Ax \in V$ for all $x \in V$;

then, it is known from linear algebra that there exists a basis for X (namely, any basis $(v_1,\ldots,v_n)$ with the property that $(v_1,\ldots,v_d)$ is also a basis for V) in which A is represented by means of a block-triangular matrix

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

whose elements on the lower (n-d) rows and left d columns are vanishing.

Moreover, if this subspace V is such that:

(ii V contains the image of the mapping B, i.e. is such that $Bu \in V$ for all $u \in U$;

then, choosing again the same basis as before for X, regardless of the choice of basis in U, the mapping B is represented by a matrix

$$\begin{pmatrix} B_1 \\ 0 \end{pmatrix}$$

whose last n-d rows are vanishing.

Thus, if there exists a subspace V which satisfies (i) and (ii), then there exists a choice of coordinates for X in which the control system is described by a set of differential equations of the form

$$\dot{x}_1 = A_{11}x_1 + A_{12}x_2 + B_1u$$

$$\dot{x}_2 = A_{22}x_2$$

By x_1 and x_2 we denote the d-vector and, respectively, the n-d vector formed by taking the first d and, respectively, the last n-d coordinates of a point x of X in the selected basis.

The representation thus obtained is particularly interesting when studying the behavior of the system under the action of the control u. At time T, the coordinates of x(T) are

$$x_1(T) = \exp(A_{11}T)x_1(0) + \int_0^T \exp(A_{11}(T-\tau))A_{12}\exp(A_{22}\tau)d\tau x_2(0) + \int_0^T \exp(A_{11}(T-\tau))B_1u(\tau)d\tau$$

$$x_2(T) = \exp(A_{22}T)x_2(0)$$

From this we see that the set of coordinates denoted with x_2 does not depend on the input u but only on the time T. The set of points that can be reached at time T, starting from x(0), under the action of the input lies inside the set of points of X whose x_2 coordinate is equal to $\exp(A_{22}T)x_2(0)$. In other words, if we let $x^o(T)$ denote the point of X reached at time T when u(t) = 0 for all $t \in [0,T]$, we observe that the state x(T) may be expressed as

$$x(T) = x^o(T) + v$$

where v is a vector in V. Therefore, the set of points that can be reached at time T, starting from x(0), lies inside the set

$$S_T = x^o(T) + V$$

Let us now make the additional assumption that the subspace V, which is the starting point of our considerations, is such that:

(iii) V is the smallest subspace which satisfies (i) and (ii) (i.e. is contained in any other subspace of X which satisfies both (i) and (ii)).

It is known from the linear theory that this happens if and only if

$$V = \sum_{i=0}^{n-1} \mathrm{Im}(A^iB)$$

and, moreover, that in this case the pair (A_{11},B_1) is a reachable pair, i.e. satisfies the condition

$$\mathrm{rank}(B_1 \; A_{11}B_1 \; \ldots \; A_{11}^{d-1}B_1) = d$$

or, in other words, for each $x_1 \in \mathbb{R}^d$ there exists an input u, defined on [0,T], such that

$$x_1 = \int_0^T \exp(A_{11}(T-\tau))B_1u(\tau)d\tau$$

Then, if V is such that the condition (iii) is also satisfied, starting from x(0) we can reach at time T any state of the form $x^o(T) + v$ with $v \in V$ or, in other words, any state belonging to the set S_T. This set is therefore exactly the set of the states reachable at time T starting from x(0).

This result suggests the following considerations. Given a linear control system, let V be the smallest subspace of X satisfying (i) and (ii). Associated with V there is a *partition* of X into subsets of the form

$$x + V$$

with the property that each one of these subsets coincides with the set of points reachable at some time T starting from a suitable point of X. Moreover, these subsets have the structure of a d-dimensional *flat submanifold* of X.

An analysis similar to the one developed so far can be carried out by examining the interaction between state and output. In this case we consider a d-dimensional subspace W of X such that

(i) W is invariant under the mapping A
(ii) W is contained into the kernel of the mapping C (i.e. is such that Cx = 0 for all $x \in W$)
(iii) W is the largest subspace which satisfies (i) and (ii) (i.e. contains any other subspace of X which satisfies both (i) and (ii)).

Then, there is a choice of coordinates for X in which the control system is described by equations of the form

$$\dot{x}_1 = A_{11}x_1 + A_{12}x_2 + B_1u$$

$$\dot{x}_2 = A_{22}x_2 + B_2u$$

$$y = C_2x_2$$

From this we see that the set of coordinates denoted with x_1 has no influence on the output y. Thus any two initial states whose last n-d coordinates coincide produce two identical outputs under any input, i.e. are indistinguishable. Actually, any two states whose last n-d coordinates coincide are such that their difference is an element of W and, then, we may conclude that any two states belonging to a set of the form x+W are indistinguishable.

Moreover, we know that the condition (iii) is satisfied if and only if

$$W = \bigcap_{i=0}^{n-1} \ker(CA^i)$$

and, if this is the case, the pair (C_2, A_{22}) is observable, i.e. satisfies the condition

$$\text{rank}(C_2' \quad A_{22}'C_2' \ \ldots (A_{22}')^{d-1}C_2') = d$$

or, in other words,

$$C_{22}\exp(A_{22}t)x_2 \equiv 0 \rightarrow x_2 = 0$$

Then, if two initial states are such that their difference does not belong to W, they may be distinguished from each other by the output produced under zero input.

Again we may synthesize the above discussion with the following considerations. Given a linear control system, let W be the largest subspace of X satisfying (i) and (ii). Associated with W there is a *partition* of X into subsets of the form

$$x + W$$

with the property that each one of these subsets coincides with the set of points that are indistinguishable from a fixed point of X. Moreover, these subsets have the structure of a d-dimensional *flat submanifold* of X.

In the following sections of this chapter and in the following chapter we shall deduce similar decompositions for nonlinear control

systems.

2. Distributions on a Manifold

The easiest way to introduce the notion of distribution Δ on a manifold N is to consider a mapping assigning to each point p of N a *subspace* Δ(p) of the tangent space T_pN to N at p. This is not a rigorous definition, in the sense that we have only defined the domain N of Δ without giving a precise characterization of its codomain. Deferring for a moment the need for a more rigorous definition, we proceed by adding some conditions of regularity. This is imposed by assuming that for each point p of N there exist a neighborhood U of p and a set of smooth vector fields defined on U, denoted $\{\tau_i : i \in I\}$, with the property that,

$$\Delta(q) = \text{span}\{\tau_i(q) : i \in I\}$$

for all $q \in U$. Such an object will be called a smooth *distribution* on N. Unless otherwise noted, in the following sections we will use the term "distribution" to mean a smooth distribution.

Pointwise, a distribution is a linear object. Based on this property, it is possible to extend a number of elementary concepts related to the notion of subspace. Thus, if $\{\tau_i : i \in I\}$ is a set of vector fields defined on N, their *span*, written $\text{sp}\{\tau_i : i \in I\}$, is the distribution defined by the rule(*)

$$\text{sp}\{\tau_i : i \in I\}: p \mapsto \text{span}\{\tau_i(p) : i \in I\}$$

If Δ_1 and Δ_2 are two distributions, their *sum* $\Delta_1 + \Delta_2$ is defined by taking

$$\Delta_1 + \Delta_2 : p \mapsto \Delta_1(p) + \Delta_2(p)$$

and their *intersection* $\Delta_1 \cap \Delta_2$ by taking

$$\Delta_1 \cap \Delta_2 : p \mapsto \Delta_1(p) \cap \Delta_2(p)$$

(*) In order to avoid confusions, we use the symbol span{·} to denote any ℝ-linear combination of elements of some ℝ-vector space (in particular, tangent vectors *at a point*). The symbol sp{·} is used to denote a *distribution* (or a *codistribution*, see later).

A distribution Δ_1 *is contained* in the distribution Δ_2 and is written $\Delta_1 \subset \Delta_2$ if $\Delta_1(p) \subset \Delta_2(p)$ for all $p \in N$. A vector field τ *belongs* to a distribution Δ and is written $\tau \in \Delta$ if $\tau(p) \in \Delta(p)$ for all $p \in N$.

The *dimension* of a distribution Δ at $p \in N$ is the dimension of the subspace $\Delta(p)$ of T_pN.

Note that the span of a given set of smooth vector fields is a smooth distribution. Likewise, the sum of two smooth distributions is smooth. However, the intersection of two such distributions may fail to be smooth. This may be seen in the following example.

(2.1) *Example.* Let $M = \mathbb{R}^2$, and

$$\Delta_1 = \mathrm{sp}\{\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\}$$

$$\Delta_2 = \mathrm{sp}\{(1+x_1)\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\}$$

Then we have

$$(\Delta_1 \cap \Delta_2)(x) = \{0\} \quad \text{if} \quad x_1 \neq 0$$

$$(\Delta_1 \cap \Delta_2)(x) = \Delta_1(x) = \Delta_2(x) \quad \text{if} \quad x_1 = 0$$

This distribution is not smooth because it is not possible to find a smooth vector field on $\mathbb{R}^2$ which is zero everywhere but on the line $x_1 = 0$. □

Since sometimes it is useful to take the intersection of smooth distributions Δ_1 and Δ_2 , one may overcome the problem that $\Delta_1 \cap \Delta_2$ is possibly non-smooth with the aid of the following concepts. Suppose Δ is a mapping which assigns to each point $p \in N$ a subspace $\Delta(p)$ of T_pN and let $M(\Delta)$ be the set of all smooth vector fields defined on N which at p take values in $\Delta(p)$, i.e.

$$M(\Delta) = \{\tau \in V(N) : \tau(p) \in \Delta(p) \text{ for all } p \in N\}$$

Then, it is not difficult to see that the span of $M(\Delta)$, in the sense defined before, is a smooth distribution contained in Δ.

(2.2) *Remark.* Recall that the set V(N) of all smooth vector fields defined on N may be given the structure of a vector space over $\mathbb{R}$ and, also, the structure of a module over $C^\infty(N)$, the ring of all smooth real-valued functions defined on N. The set $M(\Delta)$ defined before (which is non-empty because the zero element of V(N) belongs to $M(\Delta)$ for any Δ) is a subspace of the vector space V(N) and a submodule of the module V(N). From this is it easily seen that the span of $M(\Delta)$ is contained in Δ. □

Note that if Δ' is any smooth distribution contained in Δ, then Δ' is contained in the span of $M(\Delta)$, so the span of $M(\Delta)$ is actually the *largest* smooth distribution contained in Δ. To identify this distribution we shall henceforth use the notation

$$\mathrm{smt}(\Delta) \triangleq \mathrm{sp}\ M(\Delta)$$

i.e. we look at the span of $M(\Delta)$ as the "smoothing" of Δ. Note also that if Δ is smooth, then $\mathrm{smt}(\Delta) = \Delta$.

Thus, if $\Delta_1 \cap \Delta_2$ is non-smooth, we shall rather consider the distribution $\mathrm{smt}(\Delta_1 \cap \Delta_2)$.

(2.3) *Remark.* Note that $M(\Delta)$ may not be the unique subspace of V(N), or submodule of V(N), whose span coincides with $\mathrm{smt}(\Delta)$. But if M' is any other subspace of V(N), or submodule of V(N), with the property that $\mathrm{sp}\ M' = \mathrm{smt}(\Delta)$, then $M' \subset M(\Delta)$.

(2.4) *Example.* Let $N = \mathbb{R}$, and

$$\Delta = \mathrm{sp}\{x\frac{\partial}{\partial x}\}$$

Then $M(\Delta)$ is the set of all vector fields of the form $c(x)\frac{\partial}{\partial x}$ where $c(x)$ is a smooth function defined on $\mathbb{R}$ which vanishes at $x = 0$. Clearly Δ is smooth and coincides with $\mathrm{smt}(\Delta)$. There are many submodules of $V(\mathbb{R})$ which span Δ, for instance

$$M_1 = \{\tau \in V(\mathbb{R}) : \tau(x) = c(x)x\frac{\partial}{\partial x} \text{ and } c \in C^\infty(\mathbb{R})\}$$

$$M_2 = \{\tau \in V(\mathbb{R}) : \tau(x) = c(x)x^2\frac{\partial}{\partial x} \text{ and } c \in C^\infty(\mathbb{R})\}$$

Both are submodules of $M(\Delta)$, M_2 is a submodule of M_1 but M_1 is not a submodule of M_2 because is not possible to express every function $c(x)x$ as $\tilde{c}(x)x^2$ with $\tilde{c} \in C^\infty(\mathbb{R})$. □

(2.5) *Remark.* The previous considerations enable us to give a rigorous definition of a smooth distribution in the following way. A smooth distribution is a submodule M of V(N) with the following property: if θ is a smooth vector field such that for all $p \in N$

$$\theta(p) \in \mathrm{span}\{\tau(p) : \tau \in M\}$$

then θ belongs to M. □

Other important concepts associated with the notion of distribution are the ones related to the "behavior" of a given Δ as a "function" of p. We have already seen how it is possible to characterize the quality of being smooth, but there are other properties to be considered.

A distribution Δ is *nonsingular* if there exists an integer d such that

$$\dim \Delta(p) = d \tag{2.6}$$

for all $p \in N$. A singular distribution, i.e. a distribution for which the above condition is not satisfied, is sometimes called a distribution of variable dimension. If a distribution Δ is such that the condition (2.6) is satisfied for all p belonging to an open subset U of N, then we say that Δ is nonsingular on U. A point p is a *regular* point of a distribution Δ if there exists a neighborhood U of p with the property that Δ is nonsingular on U.

There are some interesting properties related to these notions, whose proof is left to the reader.

(2.7) *Lemma*. Let Δ be a smooth distribution and p a regular point of Δ. Suppose $\dim \Delta(p) = d$. Then there exist an open neighborhood U of p and a set $\{\tau_1,\ldots,\tau_d\}$ of smooth vector fields defined on U with the property that every smooth vector field τ belonging to Δ admits on U a representation of the form

$$\tau = \sum_{i=1}^{d} c_i \tau_i \tag{2.8}$$

where each c_i is a real-valued smooth function defined on U. □

A set of d vector fields which makes (2.8) satisfied will be called *a set of local generators* for Δ at p.

(2.9) *Lemma*. The set of all regular points of a distribution Δ is an open and dense submanifold of N.

(2.10) *Lemma*. Let Δ_1 and Δ_2 be two smooth distributions with the property that Δ_2 is nonsingular and $\Delta_1(p) \subset \Delta_2(p)$ at each point p of a dense submanifold of N. Then $\Delta_1 \subset \Delta_2$.

(2.11) *Lemma*. Let Δ_1 and Δ_2 be two smooth distributions with the property that Δ_1 is nonsingular, $\Delta_1 \subset \Delta_2$ and $\Delta_1(p) = \Delta_2(p)$ at each point p of a dense submanifold of N. Then $\Delta_1 = \Delta_2$. □

We have seen before that the intersection of two smooth distributions may fail to be smooth. However, around a regular point this cannot happen, as we see from the following result.

(2.12) *Lemma*. Let p be a regular point of Δ_1, Δ_2 and $\Delta_1 \cap \Delta_2$. Then there exists a neighborhood U of p with the property that $\Delta_1 \cap \Delta_2$ restricted to U is smooth. □

A distribution is *involutive* if the Lie bracket $[\tau_1,\tau_2]$ of any pair of vector fields τ_1 and τ_2 belonging to Δ is a vector field which belongs to Δ, i.e. if

$$\tau_1 \in \Delta,\ \tau_2 \in \Delta \Rightarrow [\tau_1,\tau_2] \in \Delta$$

(2.13) *Remark*. It is easy to see that a nonsingular distribution of dimension d is involutive if and only if, at each point p, any set of local generators $\tau_1,\ldots,\tau_d$ defined on a neighborhood U of p is such that

$$[\tau_i,\tau_j] = \sum_{\ell=1}^{d} c_{ij}^{\ell} \tau_\ell$$

where each c_{ij}^{ℓ} is a real-valued smooth function defined on U. □

If f is a vector field and Δ a distribution on N we denote by $[f,\Delta]$ the distribution

$$[f,\Delta] = \mathrm{sp}\{[f,\tau] \in V(N) : \tau \in \Delta\} \tag{2.14}$$

Note that $[f,\Delta]$ is a smooth distribution, even if Δ is not. Using this notation, one can say that a distribution is involutive if and only if $[f,\Delta] \subset \Delta$ for all $f \in \Delta$.

Sometimes, it is useful to work with objects that are dual to the ones defined above. In the same spirit of the definition given at the beginning of this section, we say that a *codistribution* Ω on N is a mapping assigning to each point p of N a subspace $\Omega(p)$ of the cotangent space $T_p^*(N)$. A smooth codistribution is a codistribution Ω on N with the property that for each point p of N there exist a neighborhood U of p and a set of smooth covector fields (smooth one-forms) defined on U, denoted $\{\omega_i : i \in I\}$, such that

$$\Omega(q) = \mathrm{span}\{\omega_i(q) : i \in I\}$$

for all $q \in U$.

In the same manner as we did for distributions we may define the dimension of a codistribution at p, and construct codistributions by taking the span of a given set of covector fields, or else by adding or intersecting two given codistributions, etc. always looking at a pointwise characterization of the objects we are dealing with.

Sometimes, one can construct codistributions starting from given distributions and conversely. The natural way to do this is the following: given a distribution Δ on N, the *annihilator* of Δ, denoted $\Delta^{\perp}$, is the codistribution on N defined by the rule

$$\Delta^{\perp} : p \mapsto \{ v^* \in T_p^*N : \langle v^*, v \rangle = 0 \text{ for all } v \in \Delta(p) \}$$

Conversely, the *annihilator* of Ω, denoted $\Omega^{\perp}$, is the distribution defined by the rule

$$\Omega^{\perp} : p \mapsto \{ v \in T_pN : \langle v^*, v \rangle = 0 \text{ for all } v^* \in \Omega(p) \}$$

Distributions and codistributions thus related possess a number of interesting properties. In particular, the sum of the dimensions of Δ and of $\Delta^{\perp}$ is equal to the dimension of N. The inclusion $\Delta_1 \subset \Delta_2$ is satisfied if and only if the inclusion $\Delta_1^{\perp} \supset \Delta_2^{\perp}$ is satisfied. The annihilator $(\Delta_1 \cap \Delta_2)^{\perp}$ of an intersection of distributions is equal to the sum $\Delta_1^{\perp} + \Delta_2^{\perp}$.

Like in the case of the distributions, some care is required when dealing with the quality of being smooth for codistributions constructed in some of the ways we described before. Thus it is easily seen that the span of a given set of smooth covector fields, as well as sum of two smooth codistributions is again smooth. But the intersection of two such codistributions may not need to be smooth.

Moreover, the annihilator of a smooth distribution may fail to be smooth, as it is shown in the following example.

(2.15) *Example.* Let $N = \mathbb{R}$

$$\Delta = \text{sp}\{x \frac{\partial}{\partial x}\}$$

Then

$$\Delta^{\perp}(x) = \{0\} \quad \text{if } x \neq 0$$

$$\Delta^{\perp}(x) = T_x^*N \quad \text{if } x = 0$$

and we see that $\Delta^{\perp}$ is not smooth because it is not possible to find a smooth covector field on $\mathbb{R}$ which is zero everywhere but on the point

$x = 0$. □

Or, else, the annihilator of a smooth codistribution may not be smooth, as in the following example.

(2.16) *Example.* Consider again the two distributions Δ_1 and Δ_2 described in the Example (2.1). One may easily check that

$$\Delta_1^{\perp} = \text{sp}\{dx_1 - dx_2\}$$

$$\Delta_2^{\perp} = \text{sp}\{dx_1 - (1 + x_1)dx_2\}$$

The intersection $\Delta_1 \cap \Delta_2$ is not smooth but its annihilator $\Delta_1^{\perp} + \Delta_2^{\perp}$ is smooth, because both $\Delta_1^{\perp}$ and $\Delta_2^{\perp}$ are smooth. □

One may easily extend Lemmas (2.7) to (2.12). In particular, if p is a regular point of a codistribution Ω and $\dim \Omega(p) = d$, then it is possible to find an open neighborhood U of p and a set $\{\omega_1, \ldots, \omega_d\}$ of smooth covector fields defined on U, such that every smooth covector field ω belonging to Ω can be expressed on U as

$$\omega = \sum_{i=1}^{d} c_i \omega_i$$

where each c_i is a real-valued smooth function defined on U. The set $\{\omega_1, \ldots, \omega_d\}$ is called a set of local generators for Ω at p.

We have seen before that the annihilator of a smooth distribution Δ may fail to be smooth. However, around a regular point of Δ this cannot happen, as we see from the following result.

(2.17) *Lemma.* Let p be a regular point of Δ. Then p is a regular point of $\Delta^{\perp}$ and there exists a neighborhood U of p with the property that $\Delta^{\perp}$ restricted to U is smooth. □

We conclude this section with some notations that are frequently used. If f is a vector field and Ω a codistribution on N we denote by $L_f\Omega$ the smooth codistribution

(2.18) $$L_f\Omega = \text{sp}\{L_f\omega \in V^*(N) : \omega \in \Omega\}$$

If h is a real-valued smooth function defined on N, one may associate with h a distribution, written $\ker(h_*)$, defined by

$$\ker(h_*) : p \mapsto \{v \in T_pN : h_*v = 0\}$$

One may also associate with h a codistribution, taking the span of the covector field dh. It is easy to verify that the two objects thus defined are one the annihilator of the other, i.e. that

$$(sp(dh))^{\perp} = \ker(h_*).$$

3. Frobenius Theorem

In this section we shall establish a correspondence between the notion of distribution on a manifold N and the existence of partitions of N into lower dimensional submanifolds. As we have seen at the beginning of this chapter, partitions of the state space into lower dimensional submanifolds are often encountered when dealing with reachability and/or observability of control systems.

We begin our analysis with the following definition. A *nonsingular* d-dimensional distribution Δ on N is *completely integrable* if at each $p \in N$ there exists a cubic coordinate chart (V,ξ) with coordinate functions $\xi_1,\dots,\xi_n$, such that

(3.1) $$\Delta(q) = \text{span}\{(\frac{\partial}{\partial\xi_1})_q,\dots,(\frac{\partial}{\partial\xi_d})_q\}$$

for all $q \in V$.

There are two important consequences related to the notion of completely integrable distribution. First of all, observe that if there exists a cubic coordinate chart (V,ξ), with coordinate functions $\xi_1,\dots,\xi_n$, such that (3.1) is satisfied, then any *slice* of V passing through any point p of V and defined by

(3.2) $$S_p = \{q \in V: \xi_i(q) = \xi_i(p);\ i = d+1,\dots,n\}$$

(which is a d-dimensional imbedded submanifold of N), has a tangent space which, at any point q, coincides with the subspace $\Delta(q)$ of T_qN.

Since the set of all such slices is a *partition* of V, we may see that a completely integrable distribution Δ induces, locally around each point $p \in N$, a partition into lower dimensional submanifolds, and each of these submanifolds is such that its tangent space, at each point, agrees with the distribution Δ at that point.

The second consequence is that a completely integrable distribution is *involutive*. In order to see this we use the definition of involutivity and compute the Lie bracket of any pair of vector fields belonging to Δ. For, recall that in the ξ coordinates, any vector field τ defined on N is represented by a vector of the form

$$\tau(\xi) = (\tau_1(\xi)\dots\tau_n(\xi))'$$

The components of this vector are related to the value of the vector field τ at a point q by the expression

$$\tau(q) = \tau_1(\xi(q))(\frac{\partial}{\partial\xi_1})_q+\dots+\tau_n(\xi(q))(\frac{\partial}{\partial\xi_n})_q$$

If τ is a vector field of Δ and (3.1) is satisfied, the last n-d components $\tau_{d+1}(\xi),\dots,\tau_n(\xi)$ must vanish. Moreover, if θ is any other vector field of Δ, also the last n-d components of its local representation

$$\theta(\xi) = (\theta_1(\xi)\dots\theta_n(\xi))'$$

must vanish. From this one deduces immediately that also the last n-d components of the vector

$$\frac{\partial\theta}{\partial\xi}\tau(\xi) - \frac{\partial\tau}{\partial\xi}\theta(\xi)$$

are vanishing. Since this vector represents locally the vector field $[\tau,\theta]$ one may conclude that $[\tau,\theta]$ belongs to Δ, i.e. that Δ is involutive.

We have seen that involutivity is a *necessary* condition for the complete integrability of a distribution. However, it can be proved that this condition is also sufficient, as it is stated below

(3.3) *Theorem* (Frobenius). A nonsingular distribution is completely integrable if and only if it is involutive

Proof. Let d denote the dimension of Δ. Since Δ is nonsingular, given any point $p \in N$ it is possible to find d vector fields $\tau_1,\dots,\tau_d \in \Delta$ with the property that $\tau_1(q),\dots,\tau_d(q)$ are linearly independent for all q in a suitable neighborhood U of p. In other words, these vector fields are such that

$$\Delta(q) = \text{span}\{\tau_1(q),\dots,\tau_d(q)\}$$

for all $q \in U$.

Moreover, let $\tau_{d+1},\dots,\tau_n$ be any other set of vector fields with the property that $\text{span}\{\tau_i(p): i = 1,\dots,n\} = T_pN$. With each vector field τ_i, $i = 1,\dots,n$, we associate its flow $\Phi_t^{\tau_i}$ and we consider the mapping

$$F : C_\varepsilon(0) \longrightarrow N$$

$$: (\xi_1, \ldots, \xi_n) \longmapsto \Phi_{\xi_1}^{\tau_1} \circ \Phi_{\xi_2}^{\tau_2} \circ \ldots \circ \Phi_{\xi_n}^{\tau_n}(p)$$

where $C_\varepsilon(0) = \{\xi \in \mathbb{R}^n : |\xi_i| < \varepsilon,\ 1 \leq i \leq n\}$.
If ε is sufficiently small, this mapping:

(i) is defined for all $\xi \in C_\varepsilon(0)$ and is a diffeomorphism onto its image

(ii) is such that for all $\xi \in C_\varepsilon(0)$

$$F_*\left(\frac{\partial}{\partial \xi_i}\right)_\xi \in \Delta(F(\xi)) \qquad i = 1, \ldots, d \quad (*)$$

We show now that (i) and (ii) are true and, later, that both imply the thesis.

Proof of (i). We know that for each $p \in N$ and sufficiently small $|t|$ the flow $\Phi_t^\tau(p)$ of a vector field τ is defined and this makes the function F defined for all $(\xi_1, \ldots, \xi_n)$ with sufficiently small $|\xi_i|$. Moreover, since a flow is smooth, so is F. We prove that F is a local diffeomorphism by showing that the rank of F at 0 is equal to n.

To this purpose, we first compute the image under F_* of the tangent vector $\left(\frac{\partial}{\partial \xi_i}\right)_\xi$ at a point $\xi \in C_\varepsilon(0)$. Suppose F is expressed in local coordinates. Then, it is known that the coordinates of $F_*\left(\frac{\partial}{\partial \xi_i}\right)_\xi$ in the basis $\{\left(\frac{\partial}{\partial \varphi_1}\right)_q, \ldots, \left(\frac{\partial}{\partial \varphi_n}\right)_q\}$ of the tangent space to N at the point $q = F(\xi)$ coincide with the elements of the i-th column of the jacobian matrix

$$\frac{\partial F}{\partial \xi}$$

By taking the partial derivative of F with respect to ξ_i we obtain

$$\frac{\partial F}{\partial \xi_i} = (\Phi_{\xi_1}^{\tau_1})_* \ldots (\Phi_{\xi_{i-1}}^{\tau_{i-1}})_* \frac{\partial}{\partial \xi_i}(\Phi_{\xi_i}^{\tau_i} \circ \ldots \circ \Phi_{\xi_n}^{\tau_n}(p)) =$$

$$= (\Phi_{\xi_1}^{\tau_1})_* \ldots (\Phi_{\xi_{i-1}}^{\tau_{i-1}})_* \tau_i \circ \Phi_{\xi_i}^{\tau_i} \circ \ldots \circ \Phi_{\xi_n}^{\tau_n}(p) =$$

$$= (\Phi_{\xi_1}^{\tau_1})_* \ldots (\Phi_{\xi_{i-1}}^{\tau_{i-1}})_* \tau_i \circ \Phi_{-\xi_{i-1}}^{\tau_{i-1}} \circ \ldots \circ \Phi_{-\xi_1}^{\tau_1}(F(\xi))$$

(*) Note that $\left(\frac{\partial}{\partial \xi_i}\right)_\xi$ is a tangent vector at the point ξ of $C_\varepsilon(0)$.

In particular, at $\xi = 0$, since $F(0) = p$,

$$F_*\left(\frac{\partial}{\partial \xi_i}\right)_0 = \tau_i(p)$$

The tangent vectors $\tau_1(p), \ldots, \tau_n(p)$ are by assumption linearly independent and this proves that F_* has rank n at p.

Proof of (ii). From the previous computations, we deduce also that, at any $\xi \in C_\varepsilon(0)$,

$$F_*\left(\frac{\partial}{\partial \xi_i}\right)_\xi = (\Phi_{\xi_1}^{\tau_1})_* \ldots (\Phi_{\xi_{i-1}}^{\tau_{i-1}})_* \tau_i \circ \Phi_{-\xi_{i-1}}^{\tau_{i-1}} \circ \ldots \circ \Phi_{-\xi_1}^{\tau_1}(q)$$

where $q = F(\xi)$.
If we are able to prove that for all q in a neighborhood of p, for $|t|$ small, and for any two vector fields τ and θ belonging to Δ,

$$(\Phi_t^\theta)_* \tau \circ \Phi_{-t}^\theta(q) \in \Delta(q)$$

i.e. that $(\Phi_t^\theta)_* \tau \circ \Phi_{-t}^\theta$ is a (locally defined) vector field of Δ, then we easily see that (ii) is true.

To prove the above, one proceeds as follows. Let θ be a vector field of Δ and set

$$V_i(t) = (\Phi_{-t}^\theta)_* \tau_i \circ \Phi_t^\theta(q)$$

for $i = 1, \ldots, d$.
Then, from a well known property of the Lie bracket we have

$$\frac{dV_i}{dt} = (\Phi_{-t}^\theta)_* [\theta, \tau_i] \circ \Phi_t^\theta(q)$$

Since both τ_i and θ belong to Δ and Δ is involutive, there exist functions λ_{ij} defined locally around p such that

$$[\theta, \tau_i] = \sum_{j=1}^{d} \lambda_{ij} \tau_j$$

and, therefore,

$$\frac{dV_i}{dt} = (\Phi_{-t}^\theta)_* [\sum_{j=1}^{d} \lambda_{ij}(\Phi_t^\theta(q))] \tau_j \circ \Phi_t^\theta(q) = \sum_{j=1}^{d} \lambda_{ij}(\Phi_t^\theta(q)) V_j(t)$$

The functions $V_i(t)$ are seen as solutions of a linear differential equation and, therefore, it is possible to set

$$|V_1(t)\dots V_d(t)| = |V_1(0)\dots V_d(0)|X(t)$$

where X(t) is a d×d fundamental matrix of solutions. By multiplying on the left both sides of this equality by $(\phi_t^\theta)_*$ we get

$$|\tau_1\circ\phi_t^\theta(q)\dots\tau_d\circ\phi_t^\theta(q)| = |(\phi_t^\theta)_*\tau_1(q)\dots(\phi_t^\theta)_*\tau_d(q)|X(t)$$

and also, by replacing q with $\phi_{-t}^\theta(q)$

$$|\tau_1(q)\dots\tau_d(q)| = |(\phi_t^\theta)_*\tau_1\circ\phi_{-t}^\theta(q)\dots(\phi_t^\theta)_*\tau_d\circ\phi_{-t}^\theta(q)|X(t)$$

Since X(t) is nonsingular for all t we have that, for $i = 1,\dots,d$,

$$(\phi_t^\theta)_*\tau_i\circ\phi_{-t}^\theta(q) \in \text{span}\{\tau_1(q),\dots,\tau_p(q)\}$$

i.e.

$$(\phi_t^\theta)_*\tau_i\circ\phi_{-t}^\theta(q) \in \Delta(q)$$

This result, bearing in mind the possibility of expressing any vector τ of Δ in the form

$$\tau = \sum_{i=1}^{d} c_i\tau_i$$

completes the proof of (ii).

From (i) and (ii) the thesis follows easily. Actually, (i) makes it possible to consider on the neighborhood $V = F(C_\varepsilon(0))$ of p the coordinate chart (V,F^{-1}). By definition, the tangent vector $(\frac{\partial}{\partial\xi_i})_q$ at a point $q \in V$ coincides with the image under F_* of the tangent vector $(\frac{\partial}{\partial\xi_i})_\xi$ at the point $\xi = F^{-1}(q) \in C_\varepsilon(0)$. From (ii) we see that the tangent vectors

$$(\frac{\partial}{\partial\xi_1})_q, \dots, (\frac{\partial}{\partial\xi_d})_q$$

are elements of $\Delta(q)$. Since these vectors are linearly independent, they span $\Delta(q)$ and (3.1) is satisfied. □

There are several interesting system-theoretic consequences of Frobenius' Theorem. The most important one is found in the correspondence, established by this Theorem, between involutive distributions and local partitions of a manifold into lower dimensional submanifolds. As we have seen, given a nonsingular and completely integrable, i.e. involutive, d-dimensional distribution Δ on a manifold N, around each $p \in N$ it is possible to find a coordinate neighborhood V on which Δ induces a partition into submanifolds of dimension d, which are slices (and, therefore, imbedded submanifolds) of V. Conversely, given any coordinate neighborhood V, a partition of V into d-dimensional slices defines on V a nonsingular completely integral distribution of dimension d.

We examine some examples in order to further clarify these concepts

(3.4) *Example*. Let $N = \mathbb{R}^n$ and let $x = (x_1,\dots,x_n)$ be a point on $\mathbb{R}^n$. Suppose V is a subspace of $\mathbb{R}^n$, of dimension d, spanned by the vectors

$$v_i = (v_{i1},\dots,v_{in}) \qquad 1 \le i \le d$$

We may associate with V a distribution, denoted Δ_V, in the following way. At each $x \in \mathbb{R}^n$, $\Delta_V(x)$ is the subspace of $T_x\mathbb{R}^n$ spanned by the tangent vectors

$$\sum_{j=1}^{n} v_{ij}(\frac{\partial}{\partial x_j})_x \qquad 1 \le i \le d$$

It is easily seen that this distribution is nonsingular and involutive, thus completely integrable.

Now, suppose we perform a (linear) change of coordinates in $\mathbb{R}^n$, $\xi = \xi(x)$ such that

$$\xi_i(v_j) = \delta_{ij}$$

In the ξ coordinates, the subspace V will be spanned by vectors of the form (1,0,...,0), (0,1,...,0), etc., while the subspace $\Delta_V(x)$ by the tangent vectors $(\frac{\partial}{\partial\xi_1})_x,\dots,(\frac{\partial}{\partial\xi_d})_x$. Thus, we see that the condition (3.1) is satisfied globally on $\mathbb{R}^n$ in the ξ coordinates.

The slices

$$S = \{x \in \mathbb{R}^n: \xi_i(x) = c_i, \quad i = d+1,\dots,n\}$$

characterize a global partition of $\mathbb{R}^n$ and each of these is such that its tangent space, at each point x, is exactly $\Delta_V(x)$. It is worth noting that each of these slices corresponds to a set of the form

$$x + V$$

Thus, the partitions of the state space X discussed in section 1 may be thought of as global partitions induced by a distribution associated with a given subspace of X.

(3.5) *Example.* Let $N = \mathbb{R}^2$ and let $x = (x_1, x_2)$ be a point on $\mathbb{R}^2$. Consider the one-dimensional nonsingular distribution

$$\Delta = \mathrm{sp}\{(\exp x_2)\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\}$$

If we want to find a change of coordinates that makes (3.1) satisfied, we may proceed as follows. Recall that, given a coordinate chart with coordinate functions ξ_1, ξ_2, a tangent vector v at x may be represented as

$$v = v_1(\frac{\partial}{\partial \xi_1})_x + v_2(\frac{\partial}{\partial \xi_2})_x$$

where the coefficients v_1 and v_2 are such that $v_1 = L_v\xi_1$ and $v_2 = L_v\xi_2$.

Since the tangent vector

$$\tau(x) = (\exp x_2)(\frac{\partial}{\partial x_1})_x + (\frac{\partial}{\partial x_2})_x$$

spans $\Delta(x)$ at each $x \in \mathbb{R}^2$, if we want that (3.1) is satisfied we have to have

$$\tau(x) = (L_\tau\xi_1)(\frac{\partial}{\partial \xi_1})_x + (L_\tau\xi_2)(\frac{\partial}{\partial \xi_2})_x = (\frac{\partial}{\partial \xi_1})_x$$

for all $x \in U$, or

$$1 = (L_\tau\xi_1) = (\exp x_2)\frac{\partial \xi_1}{\partial x_1} + \frac{\partial \xi_1}{\partial x_2}$$

$$0 = (L_\tau\xi_2) = (\exp x_2)\frac{\partial \xi_2}{\partial x_1} + \frac{\partial \xi_2}{\partial x_2}$$

A solution of this set of partial differential equations is given by

$$\xi_1 = \xi_1(x) = x_2$$

$$\xi_2 = \xi_2(x) = x_1 - \exp(x_2)$$

The mapping $\xi = \xi(x)$ is a diffeomorphism $\xi : \mathbb{R}^2 \to \mathbb{R}^2$ and solves the problem of finding the change of coordinates that makes (3.1) satisfied. Note that $\mathbb{R}^2$ is globally partitioned into one-dimensional slices, each one being the locus where the function $\xi_2(x)$ is constant, i.e. the locus of points (x_1, x_2) such that

$$x_1 = \exp(x_2) + \text{constant} \quad \square$$

The procedure described in the Example (3.5) may easily be extended. For, let Δ be a nonsingular involutive distribution of dimension d. Let (U, φ) be a coordinate chart with coordinate functions $\varphi_1, \ldots, \varphi_n$. Given any point $p \in U$ it is possible to find d vector fields $\tau_1, \ldots, \tau_d \in \Delta$ with the property that $\tau_1(q), \ldots, \tau_d(q)$ are linearly independent for all q in a suitable neighborhood $U' \subset U$ of p. In other words, these vectors are such that

$$\Delta(q) = \mathrm{span}\{\tau_1(q), \ldots, \tau_d(q)\}$$

for all $q \in U'$.

In the coordinates $\varphi_1, \ldots, \varphi_n$, each of these vector fields is locally expressed in the form

$$\tau_i = \sum_{j=1}^{n} (L_{\tau_i}\varphi_j)(\frac{\partial}{\partial \varphi_j})$$

If (V, ξ) is another coordinate chart around p with coordinate functions $\xi_1, \ldots, \xi_n$ the corresponding expressions for τ_i has the form

$$\tau_i = \sum_{j=1}^{n} (L_{\tau_i}\xi_j)(\frac{\partial}{\partial \xi_j})$$

For (3.1) to be satisfied, i.e. for

$$\mathrm{sp}\{\tau_1, \ldots, \tau_d\} = \mathrm{sp}\{\frac{\partial}{\partial \xi_1}, \ldots, \frac{\partial}{\partial \xi_d}\}$$

on V, we must have

$$L_{\tau_i}\xi_j = 0$$

on V, for $i = 1, \ldots, d$ and $j = d+1, \ldots, n$ and, moreover

$$\mathrm{rank}\begin{pmatrix} L_{\tau_1}\xi_1 & \ldots & L_{\tau_1}\xi_d \\ & \ldots & \\ L_{\tau_d}\xi_1 & \ldots & L_{\tau_d}\xi_d \end{pmatrix} = d$$

on V. These conditions characterize a set of partial differential equations on V, which has to be satisfied by the new coordinate functions $\xi_1,\ldots,\xi_n$.

Setting, as usual

$$\tau_{ij}(x) = (L_{\tau_i}\varphi_j)\circ\varphi^{-1}(x)$$

where $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$, it is possible to express the functions $L_{\tau_i}\xi_j$ involved in the previous conditions as follows

$$L_{\tau_i}\xi_j(x) = \sum_{k=1}^{n}\tau_{ik}(x)\frac{\partial(\xi_j\circ\varphi^{-1})}{\partial x_k}.$$

Therefore, using just $\xi_j(x)$ to denote the composite function $\xi_j\circ\varphi^{-1}(x)$, one has

$$L_{\tau_i}\xi_j(x) = \sum_{k=1}^{n}\tau_{ik}(x)\frac{\partial\xi_j}{\partial x_k}$$

Setting

$$T(x) = \begin{pmatrix} \tau_{11}(x)\ldots\tau_{d1}(x) \\ \cdot \quad \ldots \quad \cdot \\ \tau_{1n}(x)\ldots\tau_{dn}(x) \end{pmatrix}$$

the previous equations for $L_{\tau_i}\xi_j$ become

(3.6) $$\frac{\partial\xi}{\partial x}T(x) = \begin{pmatrix} K(x) \\ \\ 0_{(n-d)\times d} \end{pmatrix}$$

in which $K(x)$ is some $d\times d$ matrix of real valued functions, nonsingular for all $x \in \varphi(V)$.

Thus, we may conclude that finding a coordinate transformation $\xi = \xi(x)$ that makes (3.1) satisfied corresponds to solving a partial differential equation of the form (3.6).

Note that the matrix $T(x)$ is a matrix of rank d at $x = \varphi(p)$ because the tangent vectors $\tau_1(p),\ldots,\tau_d(p)$ are linearly independent. Therefore the matrix $\frac{\partial\xi}{\partial x}$ can be nonsingular at $x = \varphi(p)$ and this, according to the rank Theorem, guarantees that $\xi = \xi(x)$ is a local diffeomorphism.

(3.7) *Remark.* There are alternative ways to describe the equation (3.6).

For instance, one may easily check that solving these equations corresponds to find n-d functions $\lambda_1,\ldots,\lambda_{n-d}$ defined on a neighborhood V of p with values in $\mathbb{R}$ with the following properties

(i) the tangent covectors $d\lambda_1(p),\ldots,d\lambda_{n-d}(p)$ are linearly independent

(ii) $\langle d\lambda_i(q),\tau_j(q)\rangle = 0$ for all $q \in V$, $i = 1,\ldots,n-d$ and $j = 1,\ldots,d$.

In fact, if (V,ξ) is a coordinate chart that makes (3.6) satisfied, then the functions

$$\lambda_i = \xi_{i+d}$$

will satisfy (i) and (ii). Conversely, if $\lambda_1,\ldots,\lambda_{n-d}$ is a set of functions that satisfies (i) and (ii), then it is always possible to find d functions $\xi_1,\ldots,\xi_d$ defined on V and with values in $\mathbb{R}$ which, together with the functions $\xi_{d+1} = \lambda_1,\ldots,\xi_n = \lambda_{n-d}$, define a coordinate chart (V,ξ) with ξ solving the equations (3.6).

From (ii) we deduce also that there is a set of covector fields $\{d\lambda_1,\ldots,d\lambda_{n-d}\}$ with the property that at each $q \in V$, $\langle d\lambda_i(q),v\rangle = 0$ for all $v \in \Delta(q)$. Thus

$$d\lambda_i(q) \in \Delta^\perp(q) \qquad i = 1,\ldots,n-d$$

Moreover, the tangent covectors $d\lambda_1(q),\ldots,d\lambda_{n-d}(q)$ are linearly independent for all q in a neighborhood of p and $\Delta^\perp(q)$ has exactly dimension n-d. Therefore, we may conclude that the set of covector fields $\{d\lambda_1,\ldots,d\lambda_{n-d}\}$ spans $\Delta^\perp$ locally around p.

In short, we may state this result by saying that a nonsingular distribution of dimension d is integrable if and only if its annihilator is locally spanned by n-d *exact* one-forms.

(3.8) *Remark.* We note that the involutivity of Δ corresponds to the property that any two columns $\tau_i(x)$ and $\tau_j(x)$ of the matrix $T(x)$ are such that

$$\left(\frac{\partial\tau_i}{\partial x}\tau_j(x) - \frac{\partial\tau_j}{\partial x}\tau_i(x)\right) \in \mathrm{Im}(T(x))$$

for all $x \in \varphi(V)$.

(3.9) *Remark.* We know that, given a set of functions $\{\lambda_i : i \in I\}$, defined on N and with values in $\mathbb{R}$, we can define a codistribution $\Omega = \mathrm{sp}\{d\lambda_i : i \in I\}$. It is easily seen that if Ω is nonsingular then $\Omega^\perp$ is completely integrable. For, let d denote the dimension of Ω, take a point $p \in N$ and a set of functions $\lambda_1,\ldots,\lambda_d$ with the property that

$$\Omega(p) = \mathrm{span}\{d\lambda_1(p),\dots,d\lambda_d(p)\}$$

If U is a neighborhood of p with the property that $d\lambda_1(q),\dots,d\lambda_d(q)$ are linearly independent at all $q \in U$, it is seen that Ω is spanned on U by the exact forms $d\lambda_1,\dots,d\lambda_d$. As a consequence of our earlier discussions, $\Omega^\perp$ is completely integrable. □

The notion of complete integrability can be extended to a given collection of distributions. There are two cases of special importance in the applications.

Let $\Delta_1,\Delta_2,\dots,\Delta_r$ be a collection of *nested* distributions, i.e. a set of distributions with the property

$$\Delta_1 \subset \Delta_2 \subset \dots \subset \Delta_r .$$

A collection of nested nonsingular distributions on N is completely integrable if at each point $p \in N$ there exists a coordinate chart (V,ξ) with coordinate functions $\xi_1,\dots,\xi_n$ such that

$$\Delta_i(q) = \mathrm{span}\{(\frac{\partial}{\partial\xi_1})_q,\dots,(\frac{\partial}{\partial\xi_{d_i}})_q\}$$

for all $q \in V$, where d_i denotes the dimension of Δ_i.

The following results extends Frobenius Theorem

(3.10) *Theorem.* A collection $\Delta_1 \subset \Delta_2 \subset \dots \subset \Delta_r$ of nested nonsingular distributions is completely integrable if and only if each distribution of the collection is involutive.

Proof. The same construction described in the proof of Theorem (3.3) can be used. □

A collection $\Delta_1,\dots,\Delta_r$ of distributions on N is said to be *independent* if

(i) Δ_i is nonsingular, for all $i = 1,\dots,r$

(ii) $\Delta_i \cap (\sum_{j\neq i} \Delta_j) = 0$, for all $i = 1,\dots,r$

A collection of distributions $\Delta_1,\dots,\Delta_r$ is said to *span the tangent space* if for all $q \in N$

$$\Delta_1(q) + \Delta_2(q) + \dots + \Delta_r(q) = T_qN.$$

An independent collection of distributions $\Delta_1,\dots,\Delta_r$ which spans the tangent space is said to be *simultaneously integrable* if at each point $p \in N$ there exists a coordinate chart (V,ξ), with coordinate

functions $\xi_1,\dots,\xi_n$ such that

(3.11) $$\Delta_i(q) = \mathrm{span}\{(\frac{\partial}{\partial\xi_{s_i+1}})_q,\dots,(\frac{\partial}{\partial\xi_{s_{i+1}}})_q\}$$

for all $q \in V$, where $s_1 = 0$ and

$$s_i = \dim(\Delta_1 + \dots + \Delta_{i-1})$$

for $i = 2,\dots,r+1$.

The following result is an additional extension of Frobenius Theorem

(3.12) *Theorem.* An independent collection of distributions $\Delta_1,\dots,\Delta_r$ which spans the tangent space is simultaneously integrable if and only if, for all $1 \le i \le r$, the distribution

(3.13) $$D_i = \sum_{\substack{j=1\\ j\neq i}}^{r} \Delta_j$$

is involutive.

Proof. Sufficiency. Let $n_i = \dim(\Delta_i)$. Using Theorem(3.3), at each point p one may find a neighborhood V of p and, for each $1 \le i \le r$, a set of coordinate functions ξ^i_j, $1 \le j \le n$, defined on V with the property that

$$D_i = \mathrm{sp}\{\frac{\partial}{\partial\xi^i_j} : 1 \le j \le n-n_i\}$$

An easy computation shows that the covector fields

$$d\xi^1_{n-n_1+1},\dots,d\xi^1_n,\dots,d\xi^r_{n-n_r+1},\dots,d\xi^r_n$$

are linearly independent at p. Thus, the set of functions $\{\xi^i_j : n-n_i+1 \le j \le n;\ 1 \le i \le r\}$ defines on V a set of coordinate functions.

Since D_i is tangent to the slice of V where all the coordinate functions $\xi^i_{n-n_i+1},\dots,\xi^i_n$ are held constant, one deduces that Δ_i is tangent to the slice of V where all the coordinate functions $\xi^k_{n-n_k+1},\dots,\xi^k_n$, for all $k \neq i$, are held constant. This yields (3.11).

The necessity is a straightforward consequence of the definition.

4. Invariant Distributions

The notion of distribution invariant under a vector field plays, in the theory of nonlinear control systems, a role similar to the one played in the theory of linear systems by the notion of subspace invariant under a linear mapping.

A distribution Δ on N is *invariant* under a vector field f if the Lie bracket $[f,\tau]$ of f with every vector field $\tau \in \Delta$ is a vector field which belongs to Δ, i.e. if

$$(4.1) \qquad [f,\Delta] \subset \Delta$$

(4.2) *Remark.* There is a natural way to see that the previous definition generalizes the notion of invariant subspace. Let $N = \mathbb{R}^n$, A a linear mapping $A : \mathbb{R}^n \to \mathbb{R}^n$ and V a subspace of $\mathbb{R}^n$ invariant under A, i.e. such that $AV \subset V$. Suppose V is spanned by the vectors

$$v_i = (v_{i1},\ldots,v_{in}) \qquad 1 \le i \le d$$

and consider, as in the Example (3.4), the flat distribution Δ_V spanned by the vector fields

$$\tau_i = \sum_{j=1}^{n} v_{ij}\frac{\partial}{\partial x_j} \qquad 1 \le i \le d$$

With the mapping A we associate a vector field f_A represented, in the canonical basis $(\frac{\partial}{\partial x_1})_x,\ldots,(\frac{\partial}{\partial x_n})_x$ of $T_x\mathbb{R}^n$ by the vector

$$f_A(x) = Ax$$

(note that the right-hand-side of this expression represent of coordinates of an element of the tangent space at x to $\mathbb{R}^n$ and not a vector of coordinates of a point in $\mathbb{R}^n$).

It is easily seen that the distribution Δ_V is invariant under the vector field f_A in the sense of our previous definition. For, observe that any vector field τ in Δ_V can be represented in the form (2.8) where $c_1,\ldots,c_d$ is any set of real-valued functions defined locally around x.

Computing the Lie bracket of f_A and τ we have

$$[f_A,\tau] = \sum_{i=1}^{d}[f_A,c_i\tau_i] = \sum_{i=1}^{d}c_i[f_A,\tau_i] + \sum_{i=1}^{d}(L_{f_A}c_i)\tau_i$$

Moreover,

$$[f_A,\tau_i](x) = \frac{\partial\tau_i}{\partial x}f_A(x) - \frac{\partial f_A(x)}{\partial x}\tau_i(x) = -A\tau_i(x)$$

Note that $\tau_i(x)$, regarded as a point of $\mathbb{R}^n$, is an element of V, so also $A\tau_i(x) \in V$. Then, for each x, $[f_A,\tau_i](x) \in \Delta_V(x)$ and

$$[f_A,\tau](x) \in \Delta_V(x)$$

that proves the assertion. □

The notion of invariance under a vector field is particularly useful when referred to completely integrable distributions, because it provides a way of simplifying the local representation of the given vector field.

(4.3) *Lemma.* Let Δ be a nonsingular involutive distribution of dimension d and assume that Δ is invariant under the vector field f. Then, at each point $p \in N$ there exists a coordinate chart (U,ξ) with coordinate functions $\xi_1,\ldots,\xi_n$, in which the vector field f is represented by a vector of the form

$$(4.4) \qquad f(\xi) = \begin{pmatrix} f_1(\xi_1,\ldots,\xi_d,\xi_{d+1},\ldots,\xi_n) \\ \cdot\ \cdot\ \cdot \\ f_d(\xi_1,\ldots,\xi_d,\xi_{d+1},\ldots,\xi_n) \\ f_{d+1}(\xi_{d+1},\ldots,\xi_n) \\ \cdot\ \cdot\ \cdot \\ f_n(\xi_{d+1},\ldots,\xi_n) \end{pmatrix}$$

Proof. The distribution Δ, being nonsingular and involutive, is integrable and, therefore, at each point $p \in N$ there exists a coordinate chart (U,ξ) that makes (3.1) satisfied for all $q \in U$. Now, let $f_1(\xi),\ldots,f_n(\xi)$ denote the coordinates of $f(q)$ in the canonical basis of T_qN associated with (U,ξ), and recall that

$$f(q) = \sum_{i=1}^{n} f_i(\xi(q))(\frac{\partial}{\partial\xi_i})_q$$

The invariance condition (4.1) implies, in particular, that

$$[f,\frac{\partial}{\partial\xi_j}](q) \in \text{span}\{(\frac{\partial}{\partial\xi_1})_q,\ldots,(\frac{\partial}{\partial\xi_d})_q\}$$

for all $q \in U$ and $j = 1,\ldots,d$. Therefore we must have that

$$[f,\frac{\partial}{\partial\xi_j}] = \sum_{i=1}^{n}[f_i\frac{\partial}{\partial\xi_i},\frac{\partial}{\partial\xi_j}] = -\sum_{i=1}^{n}(\frac{\partial f_i}{\partial\xi_j})\frac{\partial}{\partial\xi_i} \in \text{sp}\{\frac{\partial}{\partial\xi_1},\ldots,\frac{\partial}{\partial\xi_d}\}$$

From this we see that the coefficients $\frac{\partial f_i}{\partial \xi_j}$ are such that

$$\frac{\partial f_i}{\partial \xi_j} = 0$$

for all $i = d+1,\dots,n$ and $j = 1,\dots,d$ and all $\xi \in \xi(U)$. The components $f_{d+1},\dots,f_n$ are thus independent of the coordinates $\xi_1,\dots,\xi_d$, and the (4.4) are proved. □

The following properties of invariant distributions will be also used later on.

(4.5) *Lemma*. Let Δ be a distribution invariant under the vector fields f_1 and f_2. Then Δ is also invariant under the vector field $[f_1,f_2]$.

Proof. Suppose τ is a vector field in Δ. Then, from the Jacobi identity we get

$$[[f_1,f_2],\tau] = [f_1,[f_2,\tau]]-[f_2,[f_1,\tau]]$$

By assumption $[f_2,\tau] \in \Delta$ and so is $[f_1,[f_2,\tau]]$. For the very same reasons $[f_2,[f_1,\tau]] \in \Delta$ and thus from the above equality we conclude that $[[f_1,f_2],\tau] \in \Delta$. □

(4.6) *Remark*. Note that the notion of invariance under a given vector field f is still meaningful in the case of a distribution Δ which is not smooth. In this case, it is simply required that the Lie bracket $[f,\tau]$ of f with every smooth vector field in Δ be a vector field in Δ. Since $[f,\tau]$ is a smooth vector field, it follows that if Δ is a (possibly) non-smooth distribution invariant under the vector field f, then also $\text{smt}(\Delta)$ is invariant under f. □

When dealing with codistributions, one can as well introduce the notion of invariance under a vector field in the following way.

A codistribution Ω on M is *invariant* under a vector field f if the Lie derivative along f of any covector field $\omega \in \Omega$ is a covector field which belongs to Ω, i.e. if

(4.7) $$L_f\Omega \subset \Omega$$

It is easily seen that this is the dual version of the notion of invariance of a distribution.

(4.8) *Lemma*. If a smooth distribution Δ is invariant under the vector field f, then the codistribution $\Omega = \Delta^{\perp}$ is invariant under f. If a smooth codistribution Ω is invariant under the vector field f, then the distribution $\Delta = \Omega^{\perp}$ is invariant under f.

Proof. We shall make use of the identity

$$\langle L_f\omega,\tau\rangle = L_f\langle\omega,\tau\rangle - \langle\omega,[f,\tau]\rangle$$

Suppose Δ is invariant under f and let τ be any vector field of Δ. Then $[f,\tau] \in \Delta$. Let ω be any covector field in Ω. Then, by definition

$$\langle\omega,\tau\rangle(p) = 0$$

for all $p \in N$, and also

$$\langle\omega,[f,\tau]\rangle(p) = 0$$

This yields

$$\langle L_f\omega,\tau\rangle(p) = 0$$

Since Δ is a smooth distribution, given any vector v in $\Delta(p)$ we may find a vector field τ in Δ with the property that $\tau(p) = v$ and, then, the previous result shows that

$$\langle L_f\omega(p),v\rangle = 0$$

for all $v \in \Delta(p)$, i.e. that $L_f\omega(p) \in \Omega(p)$. From this it is concluded that $L_f\omega$ is a covector field in Ω.

The second part of the statement is proved in the same way. □

(4.9) *Remark*. Note that in the previous Lemma, first part, we don't need to assume that the annihilator $\Delta^{\perp}$ of Δ is smooth, nor, in the second part, that the annihilator $\Omega^{\perp}$ of Ω is smooth. However, if both Δ and $\Delta^{\perp}$ are smooth, we conclude from the Lemma that the invariance of Δ under f implies and is implied by the invariance of $\Delta^{\perp}$ under the same vector field. In view of Lemma (2.17) this is true, in particular, whenever Δ is nonsingular. □

By making use of these notions one may give a dual formulation of Lemma (4.3). Instead of a nonsingular and involutive distribution Δ, we have to consider (see Remark (3.7)) a nonsingular codistribution Ω of dimension $n-d$ with the property that for each $p \in N$ there exist a neighborhood U of p and $n-d$ functions $\xi_{d+1},\dots,\xi_n$ defined on U with values in $\mathbb{R}$ such that

$$\Omega(q) = \text{span}\{d\xi_{d+1}(q),\ldots,d\xi_n(q)\}$$

for all $q \in U$.

If Ω satisfies these assumptions and if Ω is also invariant under f, then it is possible to find d more real-valued functions $\xi_1,\ldots,\xi_d$ defined on U with the property that, choosing as local coordinates on U the functions ξ_i , $1 \le i \le n$, for each $q \in U$ the vector field f is represented by a vector $f(\xi)$ of the form (4.4).

5. Local Decompositions of Control Systems

Throughout these notes we deal with nonlinear control systems described by equations of the form

(5.1a) $$\dot{x} = f(x) + \sum_{i=1}^{m} g_i(x)u_i$$

(5.1b) $$y_i = h_i(x) \qquad (i = 1,\ldots,\ell)$$

The state x of this system belongs to an open subset N of $\mathbb{R}^n$, while the m components $u_1,\ldots,u_m$ of the input and, respectively, the ℓ components $y_1,\ldots,y_\ell$ of the output are real-valued functions of time. We shall make later on some further assumptions on the class of admissible input functions to be considered. The vector fields $f,g_1,\ldots,g_m$ are smooth vector fields defined on N and assumed to be complete. The output maps $h_1,\ldots,h_\ell$ are real-valued smooth functions defined on N.

(5.2) *Remark.* One may define systems with the same structure as (5.1), with the state evolving on some abstract manifold N (not necessarily diffeomorphic to an open subset of $\mathbb{R}^n$). In this case, instead of (5.1), which is an ordinary differential equation defined on an open subset of $\mathbb{R}^n$, one should consider a description based upon an ordinary differential equation defined on the abstract manifold N. The vector fields $f,g_1,\ldots,g_m$ will be defined on N and so the output functions $h_1,\ldots,h_\ell$. If we let p denote a point in N then, instead of (5.1), we may use a description of the form

(5.3a) $$\dot{p} = f(p) + \sum_{i=1}^{m} g_i(p)u_i$$

(5.3b) $$y_i = h_i(p) \qquad (i = 1,\ldots,\ell)$$

with the understanding that $\dot{p}$ stands for the tangent vector at the point p to the smooth curve which characterizes the solution of (5.3a) for some fixed initial condition.

If this is the case, then (5.1) may be regarded as a local representation of (5.3) in some coordinate chart (U,φ) with the understanding that $x = \varphi(p)$. □

The theory developed so far enables us to obtain for this class of systems decompositions similar to those described at the beginning of the Chapter. The relevant results may be formalized in the following way.

(5.4) *Proposition.* Let Δ be a nonsingular involutive distribution of dimension d and assume that Δ is invariant under the vector fields $f,g_1,\ldots,g_m$. Moreover, suppose that the distribution $\text{sp}\{g_1,\ldots,g_m\}$ is contained in Δ. Then, for each point $\bar{x} \in N$ it is possible to find an open subset U of $\bar{x}$ and a local coordinates transformation $\xi = \xi(x)$ defined on U, such that, in the new coordinates, the control system (5.1a) is represented by equations of the form

(5.5a) $$\dot{\xi}_1 = f_1(\xi_1,\xi_2) + \sum_{i=1}^{m} g_{i1}(\xi_1,\xi_2)u_i$$

(5.5b) $$\dot{\xi}_2 = f_2(\xi_2)$$

where (ξ_1,ξ_2) is a partition of ξ and $\dim(\xi_1) = d$.

Proof. From Lemma (4.3) it is known that there exists, around each $\bar{x} \in N$, a coordinate chart (U,ξ) with coordinate functions $\xi_1,\ldots,\xi_n$ with the property that the vector fields $f,g_1,\ldots,g_m$ are represented in form (4.4). Moreover, since by assumption $g_i \in \Delta$ for all $i=1,\ldots,m$, then the vector fields $g_1,\ldots,g_m$ in the same coordinate chart are represented by vectors whose last $(n-d)$-components are vanishing. This coordinate chart (U,ξ) may obviously be considered as a local change of coordinates around $\bar{x}$ and therefore the Proposition is proved. □

(5.6) *Proposition.* Let Δ be a nonsingular involutive distribution of dimension d and assume that Δ is invariant under the vector fields $f,g_1,\ldots,g_m$. Moreover, assume that the codistribution $\text{sp}\{dh_1,\ldots,dh_\ell\}$ is contained in the codistribution $\Delta^\perp$. Then, for each $\bar{x} \in N$ it is possible to find an open subset U of $\bar{x}$ and a local coordinates transformation $\xi = \xi(x)$ defined on U, such that, in the new coordinates, the control system (5.1) is represented by equations of the form

(5.7a) $$\dot{\xi}_1 = f_1(\xi_1,\xi_2) + \sum_{i=1}^{m} g_{i1}(\xi_1,\xi_2)u_i$$

(5.7b) $$\dot{\xi}_2 = f_2(\xi_2) + \sum_{i=1}^{m} g_{i2}(\xi_2)u_i$$

(5.7c) $$y_i = h_i(\xi_2)$$

where (ξ_1,ξ_2) is a partition of ξ and $\dim(\xi_1) = d$.

Proof. As before, we know that there exists, around each $\bar{x} \in N$, a coordinate chart (U,ξ), with coordinate functions $\xi_1,\ldots,\xi_n$, with the property that the vector fields $f,g_1,\ldots,g_m$ are represented in the form (4.4). Moreover, we have assumed that

$$\Delta \subset [\mathrm{sp}\{dh_1,\ldots,dh_\ell\}]^\perp$$

For each point x of the selected coordinate chart we have in particular, for $j = 1,\ldots,d$,

$$\left(\frac{\partial}{\partial\xi_j}\right)_x \in \Delta(x) \subset [\sum_{i=1}^{\ell} \mathrm{span}\{dh_i(x)\}]^\perp = \bigcap_{i=1}^{\ell} [\mathrm{span}\{dh_i(x)\}]^\perp$$

As a consequence, for $j = 1,\ldots,d$ and $i = 1,\ldots,\ell$ and for all $x \in U$

$$\langle dh_i(x), \left(\frac{\partial}{\partial\xi_j}\right)_x \rangle = 0$$

or, in other words, we see that the local representation of h_i in the selected coordinate chart is such that

$$\frac{\partial h_i}{\partial \xi_j} = 0$$

for all $j = 1,\ldots,d$ and $i = 1,\ldots,\ell$ and for all $\xi \in \xi(U)$. We conclude that h_i depends only on the local coordinates $\xi_{d+1},\ldots,\xi_n$ on U and this completes the proof. □

The two local decompositions thus obtained are very useful in understanding the input-state and state-output behavior of the control system (5.1).

Suppose that the inputs u_i are piecewise constant functions of time, i.e. that there exist real numbers $T_o = 0 < T_1 < T_2 < \ldots$ such that

$$u_i(t) = \bar{u}_i^k \text{ for } T_k \le t < T_{k+1}$$

Then, on the time interval $[T_k,T_{k+1})$, the state of the system evolves

along the integral curve of the vector field

$$f + g_1\bar{u}_1^k + \ldots + g_m\bar{u}_m^k$$

passing through the point $x(T_k)$. In particular, if the initial state x^o at time $t = 0$ is contained in some neighborhood U of N, then for small t the state x(t) evolves in U.

Suppose now that the assumptions of the Proposition (5.4) are satisfied and that x^o belongs to the domain U of the coordinate transformation $\xi(x)$. If the input u is such that the x(t) evolves in U, we may use the equations (5.5) to describe the behavior of the system. From these we see that the local coordinates $(\xi_1(t),\xi_2(t))$ of x(t) are such that $\xi_2(t)$ is not affected by the input. In particular, let $x^o(T)$ denote the point of U reached at time T when $u(t) = 0$ for all $t\in[0,T]$, i.e. the point

$$x^o(T) = \Phi_T^f(x^o)$$

Φ_T^f being the flow of the vector field f, and let $(\xi_1^o(T),\xi_2^o(T))$ denote the local coordinates of $x^o(T)$. We see that the set of points that can be reached at time T, starting from x^o, lies inside the set of points whose local coordinates ξ_2 are equal to $\xi_2^o(T)$. This set is actually a *slice* of U passing through the point $x^o(T)$.

Thus, we see that locally the system displays a behavior strictly analogous to the one described in section 1. Locally, the state space may be partitioned into submanifolds (the slices of U), all of dimension d, and the points reachable at time T, along trajectories that stay in U for all $t \in [0,T]$, lie inside the slice passing through the point $x^o(T)$ reached under zero input.

The Proposition (5.6) is useful in studying state-output interactions. Suppose we take two initial states x^a and x^b belonging to U with local coordinates (ξ_1^a,ξ_2^a) and (ξ_1^b,ξ_2^b) such that

$$\xi_2^a = \xi_2^b$$

i.e. two initial states belonging to the same *slice* of U. Let $x_u^a(t)$ and $x_u^b(t)$ denote the values of the states reached at time t, starting from x^a and x^b, under the action of the same input u. From the equation (5.7b) we see immediately that, if the input u is such that $x_u^a(t)$ and $x_u^b(t)$ both evolve in U, the ξ_2 coordinates of $x_u^a(t)$ and of $x_u^b(t)$ are the same, no matter which input u we consider. Actually these coordinates $\xi_2^a(t)$ and $\xi_2^b(t)$ are solutions of the same differential equation (the equation (5.7b)) with the same initial condition. If we take into account also the (5.7c) we have the equality

$$h_i(\xi_2^a(t)) = h_i(\xi_2^b(t))$$

which holds for every input u. We may conclude that x^a and x^b are indistinguishable.

Again, we find that locally the state space may be partitioned into submanifolds (the slices of U), all of dimension d, and pair of points of each slice both produce the same output (i.e. are indistinguishable) under any input u which keeps the state trajectory evolving on U.

In the next sections we shall reach stronger conclusions, showing that if we add to the hypotheses contained in the Propositions (5.4) and (5.6) the further assumption that the distribution Δ is "minimal" (in the case of Proposition (5.4)) or "maximal" (in the case of Proposition (5.6)), then from the decompositions (5.5) and (5.7) one may obtain more informations about the set of states reachable from x^o and, respectively, indistinguishable from x^o.

We conclude this section with a remark about a dual version of Proposition (5.6).

(5.8) *Remark.* Suppose that Ω is a nonsingular codistribution of dimension n-d with the property that for each $x \in M$ there exist a neighborhood U of x and n-d real-valued functions $\xi_{d+1},\dots,\xi_n$ defined on U such that

$$\Omega(x) = \text{span}\{d\xi_{d+1}(x),\dots,d\xi_n(x)\}$$

for all $x \in U$. Let $\xi_1,\dots,\xi_d$ be other functions defining, together with $\xi_{d+1},\dots,\xi_n$, a coordinate transformation on U. In these coordinates, the one-form dh_i will be represented by a row vector

$$dh_i(\xi) = (\gamma_{i1}(\xi)\dots\gamma_{in}(\xi))$$

whose components are related to the value of dh_i at x by the expression

$$dh_i(x) = \gamma_{i1}(\xi(x))(d\xi_1)_x + \dots + \gamma_{in}(\xi(x))(d\xi_n)_x$$

If we assume that the covector fields $dh_1,\dots,dh_\ell$ belong to Ω, then, since Ω is spanned by $d\xi_{d+1},\dots,d\xi_n$ on U, we must have

$$\gamma_{ij}(\xi) = 0$$

for all $1 \le i \le \ell$, $1 \le j \le d$ and all ξ in ξ(U). But since

$$\gamma_{ij}(\xi) = \frac{\partial h_i}{\partial \xi_j}$$

one concludes that $h_1,\dots,h_\ell$ are independent of $\xi_1,\dots,\xi_d$ on U, like in (5.7c). □

6. Local Reachability

In the previous section we have seen that if there is a nonsingular distribution Δ of dimension d with the properties that:

(i) Δ is involutive
(ii) Δ contains the distribution $sp\{g_1,\dots,g_m\}$
(iii) Δ is invariant under the vector fields $f,g_1,\dots,g_m$

then at each point $\bar{x} \in N$ it is possible to find a coordinate transformation defined on a neighborhood U of $\bar{x}$ and a partition of U into slices of dimension d, such that the points reachable at some time T, starting from some initial state $x^o \in U$, along trajectories that stay in U for all $t \in [0,T]$, lie inside a slice of U. Now we want to investigate the actual "thickness" of the subset of points of a slice reached at time T.

The obvious suggestion that comes from the decomposition (5.5) is to look at the "minimal" distribution, if any, that satisfies (ii), (iii) and, then, to examine what can be said about the properties of points which belong to the same slice in the corresponding local decomposition of N. It turns out that this program can be carried out in a rather satisfactory way.

We need first some additional results on invariant distributions. If $\mathcal{D}$ is a family of distributions on N, we define the *smallest* or *minimal* element as the member of $\mathcal{D}$ (when it exists) which is contained in every other element of $\mathcal{D}$.

(6.1) *Lemma.* Let Δ be a given smooth distribution and $\tau_1,\dots,\tau_q$ a given set of vector fields. The family of all distributions which are invariant under $\tau_1,\dots,\tau_q$ and contain Δ has a minimal element, which is a smooth distribution.

Proof. The family in question is nonempty because the distribution $sp\{V(N)\}$ clearly belongs to it. Let Δ_1 and Δ_2 be two elements of this family, then it is easily seen that their intersection $\Delta_1 \cap \Delta_2$ contains Δ and, being invariant under $\tau_1,\dots,\tau_q$, is an element of the same family. This argument shows that the intersection $\hat{\Delta}$ of all elements in the family contains Δ, is invariant under $\tau_1,\dots,\tau_q$ and is contained in any other element of the family. Thus is its minimal element. $\hat{\Delta}$ must be smooth because otherwise $\mathrm{smt}(\hat{\Delta})$ would be a smooth distribution containing Δ (because Δ is smooth by assumption), invariant under $\tau_1,\dots,\tau_q$ (see Remark (4.6)) and possibly contained in $\hat{\Delta}$. □

In what follows, the smallest distribution which contains Δ and is invariant under the vector fields $\tau_1,\dots,\tau_q$ will be denoted by the symbol

$$\langle \tau_1,\dots,\tau_q | \Delta \rangle$$

While the existence of a minimal element in the family of distributions which satisfy (ii) and (iii) is always guaranteed, the nonsingularity and the involutivity require some additional assumptions. We deal with the problem in the following way. Given a distribution Δ and a set $\tau_1,\dots,\tau_q$ of vector fields we define the nondecreasing sequence of distributions

(6.2a) $$\Delta_0 = \Delta$$

(6.2b) $$\Delta_k = \Delta_{k-1} + \sum_{i=1}^{q} [\tau_i, \Delta_{k-1}]$$

There is a simple consequence of this definition

(6.3) *Lemma.* The distributions $\Delta_0,\Delta_1,\dots$ generated with the algorithm (6.2) are such that

$$\Delta_k \subset \langle \tau_1,\dots,\tau_q | \Delta \rangle$$

for all k. If there exists an integer k^* such that $\Delta_{k*} = \Delta_{k*+1}$, then

$$\Delta_{k*} = \langle \tau_1,\dots,\tau_q | \Delta \rangle$$

Proof. If Δ' is any distribution which contains Δ and is invariant under τ_i, then it is easy to see that $\Delta' \supset \Delta_k$ implies $\Delta' \supset \Delta_{k+1}$. For, we have

$$\Delta_{k+1} = \Delta_k + \sum_{i=1}^{q} [\tau_i, \Delta_k] = \Delta_k + \sum_{i=1}^{q} \mathrm{sp}\{[\tau_i,\tau] : \tau \in \Delta_k\}$$

$$\subset \Delta_k + \sum_{i=1}^{q} \mathrm{sp}\{[\tau_i,\tau] : \tau \in \Delta'\} \subset \Delta'$$

Since $\Delta' \supset \Delta_0$, by induction we see that $\Delta' \supset \Delta_k$ for all k.

If $\Delta_{k*} = \Delta_{k*+1}$ for some k^* we easily see that $\Delta_{k*} \supset \Delta$ (by definition) and Δ_{k*} is invariant under $\tau_1,\dots,\tau_q$ (because $[\tau_i,\Delta_{k*}] \subset \Delta_{k*+1} = \Delta_{k*}$ for all $1 \le i \le q$). Thus Δ_{k*} must coincide with $\langle \tau_1,\dots,\tau_q | \Delta \rangle$. □

The property $\Delta_{k*} = \Delta_{k*+1}$ expresses a sort of finiteness quality of the sequence $\Delta_0,\Delta_1,\dots$, and such a property is clearly useful from a computational point of view. The simplest practical situation in which the chain of distributions (6.2) satisfies the assumption of Lemma (6.3) arises when all the distributions of the chain are nonsingular. In this case, in fact, since by construction

$$\dim \Delta_k \le \dim \Delta_{k+1} \le n$$

it is easily seen that there exists an integer $k^* < n$ such that $\Delta_{k*} = \Delta_{k*+1}$.

If the distributions $\Delta_0,\Delta_1,\dots$ are singular, one has the following weaker result.

(6.4) *Lemma.* There exist an open and dense subset N^* of N with the property that at each point $p \in N^*$

$$\langle \tau_1,\dots,\tau_q | \Delta \rangle (p) = \Delta_{n-1}(p)$$

Proof. Suppose U is an open set with the property that, for some k^*, $\Delta_{k*}(p) = \Delta_{k*+1}(p)$ for all $p \in U$. Then, it is possible to show that $\langle \tau_1,\dots,\tau_q | \Delta \rangle (p) = \Delta_{k*}(p)$ for all $p \in U$. For, we already know from Lemma (6.3) that $\langle \tau_1,\dots,\tau_q | \Delta \rangle \supset \Delta_{k*}$. Suppose the inclusion is proper at some $\bar{p} \in U$ and define a new distribution $\bar{\Delta}$ by setting

$$\bar{\Delta}(p) = \Delta_{k*}(p) \qquad \text{if } p \in U$$

$$\bar{\Delta}(p) = \langle \tau_1,\dots,\tau_q | \Delta \rangle (p) \qquad \text{if } p \notin U$$

This distribution contains Δ and is invariant under $\tau_1,\dots,\tau_q$.

For, if τ is a vector field in $\bar{\Delta}$, then $[\tau_i,\tau] \in \langle \tau_1,\dots,\tau_q | \Delta \rangle$ (because $\bar{\Delta} \subset \langle \tau_1,\dots,\tau_q | \Delta \rangle$) and, moreover, $[\tau_i,\tau](p) \in \Delta_{k*}(p)$ for all $p \in U$ (because, in a neighborhood of p, $\tau \in \Delta_{k*}$ and $[\tau_i,\Delta^*_k] \subset \Delta^*_k$). Since $\bar{\Delta}$ is properly contained in $\langle \tau_1,\dots,\tau_q | \Delta \rangle$, this would contradict the minimality of $\langle \tau_1,\dots,\tau_q | \Delta \rangle$.

Now, let N_k be the set of regular points of Δ_k. This set is an open and dense submanifold of N (see Lemma (2.9) and so is the set $N^* = N_0 \cap N_1 \cap \dots \cap N_{n-1}$. In a neighborhood of every point $p \in N^*$ the distributions $\Delta_0,\ldots,\Delta_{n-1}$ are nonsingular. This, together with the previous discussion and a dimensionality argument, shows that $\Delta_{n-1} = \langle \tau_1,\ldots,\tau_q|\Delta \rangle$ on N^* and completes the proof. □

(6.5) *Remark.* If the distribution Δ is spanned by some of the vector fields of the set $\{\tau_1,\ldots,\tau_q\}$, then, it is possible to show that there exists an open and dense submanifold N^* of N with the following property. For each $p \in N^*$ there exist a neighborhood U of p and d vector fields (with $d = \dim\langle \tau_1,\ldots,\tau_q|\Delta \rangle(p)$) $\theta_1,\ldots,\theta_d$ of the form

$$\theta_i = [v_r,[v_{r-1},\ldots,[v_1,v_0]]]$$

where $r \leq n-1$ is an integer which may depend on i and $v_0,\ldots,v_r$ are vector fields in the set $\{\tau_1,\ldots,\tau_q\}$, such that

$$\langle \tau_1,\ldots,\tau_q|\Delta \rangle(q) = \text{span}\{\theta_1(q),\ldots,\theta_d(q)\}$$

for all $q \in U$.

This fact may be proved by induction using as N^* the subset of N defined in the proof of Lemma (6.4). Let d_0 denote the dimension of Δ_0 (which may depend on p but is constant locally around p). Since, by assumption, Δ_0 is the span of some vector fields in the set $\{\tau_1,\ldots,\tau_q\}$, there exist exactly d_0 vector fields in this set that span Δ_0 locally around p. Let d_k denote the dimension of Δ_k (constant around p) and suppose Δ_k is spanned locally around p by d_k vector fields $\theta_1,\ldots,\theta_{d_k}$ of the form

$$\theta_i = [v_r,[v_{r-1},\ldots,[v_1,v_0]]]$$

where $v_0,\ldots,v_r$ (with $r \leq k$ and possibly depending on i) are vector fields in the set $\{\tau_1,\ldots,\tau_q\}$. Then, a similar result holds for Δ_{k+1}. For, let τ be any vector field in Δ_k. From Lemma (2.7) it is known that there exists real-valued smooth functions $c_1,\ldots,c_{d_k}$ defined locally around p such that τ may be expressed, locally around p, as $\tau = c_1\theta_1 + \ldots + c_{d_k}\theta_{d_k}$. If τ_j is any vector in the set $\{\tau_1,\ldots,\tau_q\}$ we have

$$[\tau_j, c_1\theta_1+\ldots+c_{d_k}\theta_{d_k}] = c_1[\tau_j,\theta]+\ldots+c_{d_k}[\tau_j,\theta_{d_k}]+(L_{\tau_j}c_1)\theta_1+\ldots+(L_{\tau_j}c_{d_k})\theta_{d_k}$$

As a consequence

$$\Delta_{k+1} = \Delta_k + [\tau_1,\Delta_k] + \ldots + [\tau_q,\Delta_k] =$$

$$\text{sp}\{\theta_i,[\tau_1,\theta_i],\ldots,[\tau_q,\theta_i]: i = 1,\ldots,d_k\}$$

Since Δ_{k+1} is nonsingular around p, then it is possible to find exactly d_{k+1} vector fields of the form

$$\theta_i = [v_r,[v_{r-1},\ldots,[v_1,v_0]]]$$

where $v_0,\ldots,v_r$ (with $r \leq k+1$ and possibly depending on i) are vector fields in the set $\{\tau_1,\ldots,\tau_q\}$, which span Δ_{k+1} locally around p. □

The previous remark is useful in getting involutivity for the distribution $\langle \tau_1,\ldots,\tau_q|\Delta \rangle$.

(6.6) *Lemma.* Suppose Δ is spanned by some of the vector fields $\tau_1,\ldots,\tau_q$ and that $\langle \tau_1,\ldots,\tau_q|\Delta \rangle$ is nonsingular. Then $\langle \tau_1,\ldots,\tau_q|\Delta \rangle$ is involutive.

Proof. We use first the conclusion of Remark (6.5) to prove that if τ_1 and τ_2 are two vector fields in Δ_{n-1}, then their Lie bracket $[\tau_1,\tau_2]$ is such that $[\tau_1,\tau_2](p) \in \Delta_{n-1}(p)$ for all $p \in N^*$. Using again Lemma (2.7) and the previous result we deduce, in fact, that in a neighborhood U of p

$$[\tau_1,\tau_2] = [\sum_{i=1}^{d} c_i^1\theta_i, \sum_{j=1}^{d} c_j^1\theta_j] \in \text{sp}\{\theta_i,\theta_j,[\theta_i,\theta_j]: i,j = 1,\ldots,d\}$$

where θ_i,θ_j are vector fields of the form described before.

In order to prove the claim, we have only to show that $[\theta_i,\theta_j](p)$ is a tangent vector in $\Delta_{n-1}(p)$. For this purpose, we recall that on N^* the distribution Δ_{n-1} is invariant under the vector fields $\tau_1,\ldots,\tau_q$ (see Lemma (6.4)) and that any distribution invariant under vector fields τ_1 and τ_2 is also invariant under their Lie bracket $[\tau_1,\tau_2]$ (see Lemma (4.5)). Since each θ_i is a repeated Lie bracket of the vector fields $\tau_1,\ldots,\tau_q$, $[\theta_i,\Delta_{n-1}](p) \subset \Delta_{n-1}(p)$ for all $1 \leq i \leq d$ and, thus, in particular $[\theta_i,\theta_j](p)$ is a tangent vector which belongs to $\Delta_{n-1}(p)$.

Thus the Lie bracket of two vector fields τ_1,τ_2 in Δ_{n-1} is such that $[\tau_1,\tau_2](p) \in \Delta_{n-1}(p)$. Moreover, it has already been observed that $\langle \tau_1,\ldots,\tau_q|\Delta \rangle = \Delta_{n-1}$ in a neighborhood of p and, therefore, we conclude that at any point p of N^* the Lie bracket of any two vector fields τ_1,τ_2 in $\langle \tau_1,\ldots,\tau_q|\Delta \rangle$ is such that $[\tau_1,\tau_2](p) \in \langle \tau_1,\ldots,\tau_q|\Delta \rangle(p)$.

Consider now the distribution

$$\bar{\Delta} = \langle \tau_1,\dots,\tau_q|\Delta\rangle + sp\{[\theta_i,\theta_j] : \theta_i,\theta_j \in \langle \tau_1,\dots,\tau_q|\Delta\rangle\}$$

which, by construction, is such that

$$\bar{\Delta} \supset \langle \tau_1,\dots,\tau_q|\Delta\rangle$$

From the previous result it is seen that $\bar{\Delta}(p) = \langle \tau_1,\dots,\tau_q|\Delta\rangle(p)$ at each point p of N^*, which is a dense set in N. By assumption, $\langle \tau_1,\dots,\tau_q|\Delta\rangle$ is nonsingular. So, by Lemma (2.11) we deduce that $\bar{\Delta} = \langle \tau_1,\dots,\tau_q|\Delta\rangle$, and, therefore, that $[\theta_i,\theta_j] \in \langle \tau_1,\dots,\tau_q|\Delta\rangle$ for all pair $\theta_i,\theta_j \in \langle \tau_1,\dots,\tau_q|\Delta\rangle$. This concludes the proof. □

(6.7) *Remark.* From Lemmas (6.4),(6.6) and (2.11) it may also be deduced that if Δ is spanned by some of the vector fields $\tau_1,\dots,\tau_q$ and Δ_{n-1} is nonsingular, then

$$\langle \tau_1,\dots,\tau_q|\Delta\rangle = \Delta_{n-1}$$

and $\langle \tau_1,\dots,\tau_q|\Delta\rangle$ is involutive. □

We now come back to the original problem of the study the smallest distribution which contains $sp\{g_1,\dots,g_m\}$ and is invariant under the vector fields $f,g_1,\dots,g_m$. From the previous Lemma it is seen that if $\langle f,g_1,\dots,g_m|\ sp\{g_1,\dots,g_m\}\rangle$ is nonsingular, then it is also involutive and, therefore, the decomposition (5.5) may be performed. We will see later that the minimality of $\langle f,g_1,\dots,g_m|sp\{g_1,\dots,g_m\}\rangle$ makes it possible to deduce an interesting topological property of the set of points reached at some fixed time T starting from a given point x^o. However, before doing this, it is convenient to analyze some other characteristics of the decomposition (5.5).

Consider the distribution $\langle f,g_1,\dots,g_m|sp\{f,g_1,\dots,g_m\}\rangle$, i.e. the smallest distribution invariant under $f,g_1,\dots,g_m$ and which contains $sp\{f,g_1,\dots,g_m\}$ (note that now not only the vector fields $g_1,\dots,g_m$ but also the vector field f is assumed to belong to this distribution).

If this distribution is nonsingular, and therefore involutive by Lemma (6.6), it may indeed be used in defining a local decomposition of the control system (5.1) similar to the decomposition (5.5). We are going to see in which way this new decomposition is related to the decomposition (5.5) and why it may be of interest.

In order to simplify the notation, we set

(6.8a) $$P = \langle f,g_1,\dots,g_m|sp\{g_1,\dots,g_m\}\rangle$$

(6.8b) $$R = \langle f,g_1,\dots,g_m|sp\{f,g_1,\dots,g_m\}\rangle$$

The relation between P and R is described in the following statement

(6.9) *Lemma.* The distributions P and R are such that

(a) $P + sp\{f\} \subset R$

(b) if x is a regular point of $P + sp\{f\}$, then

$$(P + sp\{f\})(x) = R(x)$$

Proof. By definition, $P \subset R$ and $f \in R$, so (a) is true.

It is known from the proof of Lemma (6.6) that, around each point x of an open dense submanifold N^* of N, R is spanned by vector fields of the form

$$\theta_i = [v_r,\dots,[v_1,v_0]]$$

where $r \le n-1$ is an integer which may depend on i, and $v_r,\dots,v_1,v_0$ are vector fields in the set $\{f,g_1,\dots,g_m\}$.

It is easy to see that all such vector fields belong to P+sp{f}. For, if θ_i is just one of the vector fields in the set $\{f,g_1,\dots,g_m\}$ it either belongs to P (which contains $g_1,\dots,g_m$) or to sp{f}. If θ_i has the general form shown above we may, without loss of generality, assume that v_0 is in the set $\{g_1,\dots,g_m\}$. For, if $v_0 = f$ and $v_1 = f$, then $\theta_i = 0$. Otherwise, if $v_0 = f$ and $v_1 = g_j$, then $-\theta_i = [v_r,\dots,[f,g_j]]$ has the desired form. Any vector of the form

$$\theta_i = [v_r,\dots,[v_1,g_j]]$$

with $v_r,\dots,v_1$ in the set $\{f,g_1,\dots,g_m\}$ is in P because P contains g_j and is invariant under $f,g_1,\dots,g_m$ and so the claim is proved.

From this fact we deduce that on an open and dense submanifold N^* of N,

$$R \subset P + sp\{f\}$$

and therefore, since $R \supset P + sp\{f\}$ on N, that on N^*

$$R = P + sp\{f\}$$

Suppose that P + span f has constant dimension on some neighborhood U. Then, from Lemma (2.11) we conclude that the two distributions R and P + sp{f} coincide on U. □

(6.10) *Corollary.* If P and P + sp{f} are nonsingular, then

$$\dim(R) - \dim(P) \le 1. \; \square$$

If P and P + sp{f} are both nonsingular, so is R and, by Lemma (6.6), both P and R are involutive. Suppose that P is properly contained in R. Then, using Theorem (3.10),one can find, locally around each $\bar{x} \in N$, a neighborhood U of $\bar{x}$ and a coordinate transformation $\xi = \xi(x)$ defined on U such that

$$P(x) = sp\{(\frac{\partial}{\partial\xi_1})_x, \ldots, (\frac{\partial}{\partial\xi_{r-1}})_x\} \tag{6.11a}$$

$$R(x) = sp\{(\frac{\partial}{\partial\xi_1})_x, \ldots, (\frac{\partial}{\partial\xi_{r-1}})_x, (\frac{\partial}{\partial\xi_r})_x\} \tag{6.11b}$$

for all $x \in U$, where $r = \dim(R)$.

In the ξ coordinates the control system (5.1a) is represented by equations of the form

$$\dot{\xi}_1 = f_1(\xi_1,\ldots,\xi_n) + \sum_{i=1}^{m} g_{i1}(\xi_1,\ldots,\xi_n)u_i$$

$$\cdots$$

$$\dot{\xi}_{r-1} = f_{r-1}(\xi_1,\ldots,\xi_n) + \sum_{i=1}^{m} g_{i,r-1}(\xi_1,\ldots,\xi_n)u_i$$

$$\dot{\xi}_r = f_r(\xi_r,\ldots,\xi_n) \tag{6.12}$$

$$\dot{\xi}_{r+1} = 0$$

$$\cdots$$

$$\dot{\xi}_n = 0$$

The last components of the vector field f are vanishing because, by construction, $f \in R$. In the particular case where R = P also the r-th component of f vanishes and the corresponding equation for ξ_r is

$$\dot{\xi}_r = 0$$

From the equation (6.12) we see that any trajectory $x(t)$ evolving on the neighborhood U actually belongs to an r-dimensional slice of U

passing through the initial point. This slice is in turn partitioned into (r-1)-dimensional slices, each one including the set of points reached at a prescribed time T.

(6.13) *Remark.* A further change of local coordinates makes it possible to better understand the role of the time in the behavior of the control system (6.12). We may assume, without loss of generality, that the initial point x^o is such that $\xi(x^o) = 0$. Therefore we have $\xi_i(t) = 0$ for all $i = r+1,\ldots,n$ and

$$\dot{\xi}_r = f_r(\xi_r,0,\ldots,0)$$

Moreover, if we make the assumption that $f \notin P$, then the function f_r is nonzero everywhere on the neighborhood U. Now, let $\xi_r(t)$ denote the solution of this differential equation which passes through 0 at t = 0. Clearly, the mapping

$$\mu : t \longmapsto \xi_r(t)$$

is a diffeomorphism from an open interval $(-\varepsilon,\varepsilon)$ of the time axis onto the open interval of the ξ_r axis $(\xi_r(-\varepsilon),\xi_r(\varepsilon))$. If its inverse μ^{-1} is used as a local coordinate transformation on the ξ_r axis one easily sees that the new coordinate

$$\tilde{\xi}_r = \mu^{-1}(\xi_r) = t$$

satisfies the differential equation

$$\dot{\tilde{\xi}}_r = 1$$

In these new coordinates, points on the r-dimensional slice of U passing through the initial state are parametrized by $(\xi_1,\ldots,\xi_{r-1},t)$. In particular, the points reached at time T belong to the (r-1)-dimensional slice

$$S = \{x \in U: \; \xi_r(x) = T, \; \xi_{r+1}(x) = 0,\ldots,\xi_n(x) = 0\}. \; \square$$

(6.14) *Remark.* If f is a vector field of P then the local representation (6.12) is such that f_r vanishes on U. Therefore, starting from a point x^o such that $\xi(x^o) = 0$ we shall have $\xi_i(t) = 0$ for all $i=r,\ldots,n$ and the state x(t) shall evolve on a (r-1)-dimensional slice of U passing through x^o. □

By definition the distribution R is the smallest distribution which contains $f, g_1, \ldots, g_m$ and is invariant under $f, g_1, \ldots, g_m$. Thus, we may say that in the associated decomposition (6.12) the dimension r is "minimal", in the sense that it is not possible to find another set of local coordinates $\tilde{\xi}_1, \ldots, \tilde{\xi}_{\tilde{r}}, \ldots, \tilde{\xi}_n$, with $\tilde{r}$ strictly less than r, with the property that the last $n-\tilde{r}$ coordinates remain constant with the time. We shall now show that, from the point of view of the interaction between input and state, the decomposition (6.12) has even stronger properties. Actually, we are going to prove that the states reachable from the initial state x^o fill up at least an open subset of the r-dimensional slice of in which they are contained.

(6.15) *Theorem.* Suppose the distribution R (i.e. the smallest distribution invariant under $f, g_1, \ldots, g_m$ which contains $f, g_1, \ldots, g_m$) is nonsingular. Let r denote the dimension of R. Then, for each $x^o \in N$ it is possible to find a neighborhood U of x^o and a coordinate transformation $\xi = \xi(x)$ defined on U with the following properties

(a) the set $R(x^o)$ of states reachable starting from x^o along trajectories entirely contained in U and under the action of piecewise constant input functions is a subset of the slice

$$S_{x^o} = \{x \in U : \xi_{r+1}(x) = \xi_{r+1}(x^o), \ldots, \xi_n(x) = \xi_n(x^o)\}$$

(b) the set $R(x^o)$ contains an open subset of S_{x^o}.

Proof. The proof of the statement (a) follows from the previous discussion. We proceed directly to the proof of (b), assuming throughout the proof to operate on the neighborhood U on which the coordinate transformation $\xi(x)$ is defined. For convenience, we break up the proof in several steps.

(i) Let $\theta_1, \ldots, \theta_k$ be a set of vector fields, with $k < r$, and let $\phi_t^1, \ldots, \phi_t^k$ denote the corresponding flows. Consider the mapping

$$F : (-\varepsilon, \varepsilon)^k \to N$$

$$(t_1, \ldots, t_k) \longmapsto \phi^k_{t_k} \circ \ldots \circ \phi^1_{t_1}(x^o)$$

where x^o is a point of N and suppose that its differential has rank k at some $s_1, \ldots, s_k$, with $0 \le s_i < \varepsilon$ for $1 \le i \le k$. For ε sufficiently small the mapping

(6.16) $$\tilde{F} : (s_1, \varepsilon) \times \ldots \times (s_k, \varepsilon) \to N$$

$$(t_1, \ldots, t_k) \longmapsto F(t_1, \ldots, t_k)$$

is an embedding.

Let M denote the image of the mapping (6.16) (which depends on the point x^o). Consider the slice of U

$$S_{x^o} = \{x \in U : \xi_i(x) = \xi_i(x^o),\ r+1 \le i \le n\}$$

If the vector fields $\theta_1, \ldots, \theta_k$ have the form

$$\theta_j = f + \sum_{i=1}^{m} g_i u_i^j$$

with $u_i^j \in \mathbb{R}$ for $1 \le i \le m$ and $1 \le j \le k$, then for ε small M is an embedded submanifold of S_{x^o}. This implies, in particular, that for each $x \in M$

(6.17) $$T_x M \subset R(x)$$

where R, as before, is the smallest distribution invariant under $f, g_1, \ldots, g_m$ which contains $f, g_1, \ldots, g_m$ (recall that R(x) is the tangent space to S_{x^o} at x).

(ii) Suppose that the vector fields $f, g_1, \ldots, g_m$ are such that

(6.18a) $$f(x) \in T_x M$$

(6.18b) $$g_i(x) \in T_x M \qquad 1 \le i \le m$$

for all $x \in M$. We shall show that this contradicts the assumption $k < r$. For, consider the distribution $\bar{\Delta}$ defined by setting

$$\bar{\Delta}(x) = T_x M \qquad \text{for all } x \in M$$

$$\bar{\Delta}(x) = R(x) \qquad \text{for all } x \in (N \setminus M)$$

This distribution is contained in R (because of (6.17)) and contains the vector fields $f, g_1, \ldots, g_m$ (because these vector fields are in R and, moreover, it is assumed that (6.18) are true).

Let τ be any vector field of $\bar{\Delta}$. Then $\tau \in R$ and since R is invariant under $f, g_1, \ldots, g_m$, then for all $x \in (N \setminus M)$

(6.19a) $$[f,\tau](x) \in \bar{\Delta}(x)$$

(6.19b) $$[g_i,\tau](x) \in \bar{\Delta}(x) \qquad 1 \leq i \leq m$$

Moreover since $\tau, f, g_1, \ldots, g_m$ are vector fields which are tangent to M at each $x \in M$, we have also that (6.19) hold for all $x \in M$, and therefore for all $x \in N$.

Having shown $\bar{\Delta}$ is invariant under $f, g_1, \ldots, g_m$ and contains $f, g_1, \ldots, g_m$, we deduce that $\bar{\Delta}$ must coincide with R. But this is a contradiction since for all $x \in M$

$$\dim \bar{\Delta}(x) = k$$

$$\dim R(x) = r > k$$

(iii) If (6.18) are not true, then it is possible to find m real numbers $u_1^{k+1}, \ldots, u_m^{k+1}$ and a point $\bar{x} \in M$ such that the vector field

$$\theta_{k+1} = f + \sum_{i=1}^{m} g_i u_i^{k+1}$$

satisfies the condition $\theta_{k+1}(\bar{x}) \notin T_{\bar{x}} M$.

Let $\bar{x} = \bar{F}(s_1', \ldots, s_k')$ be this point and ϕ_t^{k+1} denote the flow of θ_{k+1}. Then the mapping

$$F' : (-\varepsilon, \varepsilon)^{k+1} \to N$$

$$(t_1, \ldots, t_k, t_{k+1}) \longmapsto \phi^{k+1}_{t_{k+1}} \circ F(t_1, \ldots, t_k)$$

at the point $(s_1', \ldots, s_k', 0)$ has rank k+1.

For, note that

$$(F')_*(\frac{\partial}{\partial t_i})(s_1', \ldots, s_k', 0) = (F)_*(\frac{\partial}{\partial t_i})(s_1', \ldots, s_k')$$

for $i = 1, \ldots, k$ and that

$$(F')_*(\frac{\partial}{\partial t_{k+1}})(s_1', \ldots, s_k', 0) = \theta_{k+1}(\bar{x})$$

The first k tangent vectors at $\bar{x}$ are linearly independent, because F has rank k at all points of $(s_1, \varepsilon) \times \ldots \times (s_k, \varepsilon)$. The (k+1)-th one is independent from the first k by construction and therefore F' has rank k+1 at $(s_1, \ldots, s_k, 0)$.

Since $s_i' > s_i$, we may conclude that the mapping F' has rank k+1 at a point $(s_1', \ldots, s_{k+1}')$, with $0 \leq s_i' < \varepsilon$ for $1 \leq i \leq k+1$.

Note that given any real number $T > 0$ it is always possible to choose the point $\bar{x}$ in such a way that

$$(s_1' - s_1) + \ldots + (s_k' - s_k) < T$$

For, otherwise, we had that any vector field of the form

$$\theta = f + \sum_{i=1}^{m} g_i u_i$$

would be tangent to the image under $\bar{F}$ of the open set

$$\{(t_1, \ldots, t_k) \in (s_1, \varepsilon) \times \ldots \times (s_k, \varepsilon) : (t_1 - s_1) + \ldots + (t_k - s_k) < T\}$$

and this, as in (ii), would be a contradiction.

(iv) We can now construct a sequence of mappings of the form (6.16).

Let $\theta_1 = f + \sum_{i=1}^{m} g_i u_i^1$ be a vector field which is not zero at x^o (such a vector field can always be found because, otherwise, we would have $R(x^o) = \{0\}$) and let M_1 denote the image of the mapping

$$\bar{F}_1 : (0, \varepsilon) \to N$$

$$t_1 \longmapsto \phi^1_{t_1}(x^o)$$

Let $\bar{x} = \bar{F}_1(s_1^1)$ be a point of M_1 in which a vector field of the form $\theta_2 = f + \sum_{i=1}^{m} g_i u_i^2$ is such that $\theta_2(\bar{x}) \notin T_{\bar{x}} M_1$. Then we may define the mapping

$$\bar{F}_2 : (s_1^1, \varepsilon) \times (0, \varepsilon) \to N$$

$$(t_1, t_2) \longmapsto \phi^2_{t_2} \circ \phi^1_{t_1}(x^o)$$

Iterating this procedure, at stage k we start with a mapping

$$\bar{F}_k : (s_1^{k-1}, \varepsilon) \times \ldots \times (s_{k-1}^{k-1}, \varepsilon) \times (0, \varepsilon) \to N$$

$$(t_1, \ldots, t_{k-1}, t_k) \longmapsto \phi^k_{t_k} \circ \ldots \circ \phi^1_{t_1}(x^o)$$

and we find a point $\bar{x} = \bar{F}_k(s_1^k,\dots,s_k^k)$ of its image M_k and a vector field $\theta_{k+1} = f + \sum_{i=1}^{m} g_i u_i^{k+1}$ such that $\theta_{k+1}(\bar{x}) \notin T_{\bar{x}}M_k$. This makes it possible to define the next mapping $\bar{F}_{k+1}$. Note that $s_i^k > s_i^{k-1}$ for $i = 1,\dots,k-1$ and $s_k^k > 0$.

The procedure clearly stops at the stage r, when a mapping $\bar{F}_r$ is defined

$$\bar{F}_r : (s_1^{r-1},\varepsilon)\times\dots\times(s_{r-1}^{r-1},\varepsilon)\times(0,\varepsilon) \to N$$

$$(t_1,\dots,t_{r-1},t_r) \longmapsto \Phi_{t_r}^r \circ\dots\circ\Phi_{t_1}^1(x^o)$$

(v) Observe that a point $x = \bar{F}_r(t_1,\dots,t_r)$ in the image M_r of the embedding $\bar{F}_r$ can be reached, starting from the state x^o at time t=0, under the action of the piecewise constant control defined by

$$u_i(t) = u_i^k \text{ for } t \in [t_1+\dots+t_{k-1}, t_1+t_2+\dots+t_k)$$

Thus, we know from our previous discussions that M_r must be contained in the slice of U

$$S_{x^o} = \{x \in U: \xi_i(x) = \xi_i(x^o),\ r+1 \le i \le n\}$$

The images under $\bar{F}_r$ of the open sets of

$$U_r = (s_1^{r-1},\varepsilon)\times\dots\times(s_{r-1}^{r-1},\varepsilon)\times(0,\varepsilon)$$

are open in the topology of M_r as a subset of U (because $\bar{F}_r$ is an embedding) and therefore they are also open in the topology of M_r as a subset of S_{x^o} (because S_{x^o} is an embedded submanifold of U). Therefore we have that M_r is an emebedded submanifold of S_{x^o} and a dimensionality argument tell us that M_r is actually an open submanifold of S_{x^o}. □

(6.20) *Theorem*. Suppose the distributions P (i.e. the smallest distribution invariant under $f,g_1,\dots,g_m$ which contains $g_1,\dots,g_m$) and P+sp{f} are nonsingular. Let p denote the dimension of P. Then, for each $x^o \in M$ it is possible to find a neighborhood U of x^o and a coordinate transformation $\xi = \xi(x)$ defined on U with the following properties:

(a) the set $R(x^o,T)$ of states reachable at time t = T starting from x^o at t = 0, along trajectories entirely contained in U and under the action of piecewise constant input functions, is a subset of the slice

$$S_{x^o,T} = \{x \in U: \xi_{p+1} = \xi_{p+1}(\Phi_T^f(x^o)); \xi_{p+2}(x) = \xi_{p+2}(x^o),\dots,\xi_n(x) = \xi_n(x^o)\}$$

(b) the set $R(x^o,T)$ contains an open subset of $S_{x^o,T}$.

Proof. We know from Lemma (6.9) that R is nonsingular. Therefore one can repeat the construction used to prove the part (b) of Theorem (6.15). Moreover, from Corollary (6.10) it follows that r, the dimension of R, is equal either to p+1 or to p.

Suppose the first situation happens. Given any real number $T \in (0,\varepsilon)$, consider the set

$$U_r^T = \{(t_1,\dots,t_r) \in U_r : t_1+\dots+t_r = T\}$$

where U_r is as defined at the step (v) in the proof of Theorem (6.15). From the last remark at the step (iii) we know that there exists always a suitable choice of $s_1^{r-1},\dots,s_{r-1}^{r-1}$ after which this set is not empty.

Clearly the image $\bar{F}_r(U_r^T)$ consists of points reachable at time T and therefore is contained in $R(x^o,T)$. Moreover, using the same arguments as in (v) , we deduce that the set $\bar{F}_r(U_r^T)$ is an open subset of $S_{x^o,T}$.

If p = r, i.e. if P = R, the proof can be carried out by simply adding an extra state variable satisfying the equation

$$\dot{\xi}_{n+1} = 1$$

and showing that this reduces the problem to the previous one. The details are left to the reader. □

7. Local Observability

We have seen in section 5 that if there is a nonsingular distribution Δ of dimension d with the properties that

(i) Δ is involutive

(ii) Δ is contained in the distribution $sp\{dh_1,\dots,dh_\ell\}^\perp$

(iii) Δ is invariant under the vector fields $f,g_1,\dots,g_m$

then, at each point $\bar{x} \in N$ it is possible to find a coordinate transformation defined in a neighborhood U of $\bar{x}$ and a partition of U into slices of dimension d, such that points on each slice produce the same output under any input u which keeps the state trajectory evolving on U. We want now to find conditions under which points belonging to different slices of U produce different outputs, i.e. are distinguishable.

In this case we see from the decomposition (5.7) that the right object to look for is now the "largest" distribution which satisfies (ii),(iii). Since the existence of a nonsingular distribution Δ which satisfies (i),(ii),(iii) implies and is implied by the existence of a codistribution Ω (namely $\Delta^{\perp}$) with the properties that

(i') Ω is spanned, locally around each point $p \in N$, by n-d exact covector fields

(ii') Ω contains the codistribution $sp\{dh_1,\dots,dh_\ell\}$

(iii') Ω is invariant under the vector fields $f, g_1,\dots,g_m$

we may as well look for the "smallest" codistribution which satisfies (ii'),(iii').

Like in the previous section, we need some background material. However, most of the results stated below require proofs which is are similar to those of the corresponding results stated before and, for this reason, will be omitted.

(7.1) *Lemma.* Let Ω be a given smooth codistribution and $\tau_1,\dots,\tau_q$ a given set of vector fields. The family of all codistributions which are invariant under $\tau_1,\dots,\tau_q$ and contain Ω has a minimal element, which is a smooth codistribution. □

We shall use the symbol $\langle \tau_1,\dots,\tau_q|\Omega\rangle$ to denote the smallest codistribution which contains Ω and is invariant under $\tau_1,\dots,\tau_q$.

Given a codistribution Ω and a set of vector fields $\tau_1,\dots,\tau_q$ one can consider the following dual version of the algorithm (6.2)

(7.2a) $$\Omega_0 = \Omega$$

(7.2b) $$\Omega_k = \Omega_{k-1} + \sum_{i=1}^{q} L_{\tau_i}\Omega_{k-1}$$

and have the following result.

(7.3) *Lemma.* The codistributions $\Omega_0, \Omega_1, \dots$ generated with the algorithm (7.2) are such that

$$\Omega_k \subset \langle \tau_1,\dots,\tau_q|\Omega\rangle$$

for all k. If there exists an integer k^* such that $\Omega_{k^*} = \Omega_{k^*+1}$, then

$$\Omega_k^* = \langle \tau_1,\dots,\tau_q|\Omega\rangle \qquad \square$$

The dual version of Lemma (6.4) is the following one

(7.4) *Lemma.* There exists an open and dense subset N^* of N with the property that at each point $p \in N^*$

$$\langle \tau_1,\dots,\tau_q|\Omega\rangle = \Omega_{n-1}(p)$$

(7.5) *Remark.* If the codistribution Ω is spanned by a set $d\lambda_1,\dots,d\lambda_s$ of exact covector fields, then there exists an open and dense submanifold N^* of N with the following property. For each $p \in N^*$ there exists a neighborhood U of p and d exact covector fields (with $d = \dim\langle \tau_1,\dots,\tau_q|\Omega\rangle(p)$) $\omega_1,\dots,\omega_d$ which have the form

$$\omega_i = d(L_{v_r}\dots L_{v_1}\lambda_j)$$

where $r \le n-1$ is an integer which may depend on i, $v_1,\dots,v_r$ are vector fields in the set $\{\tau_1,\dots,\tau_q\}$ and λ_j is a function in the set $\{\lambda_1,\dots,\lambda_s\}$, such that

$$\langle \tau_1,\dots,\tau_q|\Omega\rangle(q) = sp\{\omega_1(q),\dots,\omega_d(q)\}$$

for all $q \in U$.

This may easily be proved by induction as for the corresponding statement in Remark (6.5). □

(7.6) *Lemma.* Suppose Ω is spanned by a set $d\lambda_1,\dots,d\lambda_s$ of exact covector fields and that $\langle \tau_1,\dots,\tau_q|\Omega\rangle$ is nonsingular. Then $\langle \tau_1,\dots,\tau_q|\Omega\rangle^{\perp}$ is involutive.

Proof. From the previous Remark, it is seen that in a neighborhood of each point p in an open and dense submanifold N^*, the codistribution $\langle \tau_1,\dots,\tau_q|\Delta\rangle$ is spanned by exact covector fields.

Therefore, the Lie bracket of any two vector fields τ_1, τ_2 in $\langle \tau_1,\dots,\tau_q|\Omega\rangle^{\perp}$ is such that $[\tau_1,\tau_2](p) \in \langle \tau_1,\dots,\tau_q|\Omega\rangle^{\perp}(p)$ (see Remark (3.9)).

From this result, using again Lemma (2.11) as in the proof of Lemma (6.6), one deduces that $\langle \tau_1,\dots,\tau_q|\Omega\rangle^{\perp}$ is involutive. □

(7.7) *Remark.* From Lemmas (7.4),(7.6) and (2.11) one may also deduce that if Ω is spanned by a set $d\lambda_1,\dots,d\lambda_s$ of exact covector fields and Ω_{n-1} is nonsingular, then

$$\langle \tau_1,\dots,\tau_q|\Omega\rangle = \Omega_{n-1}$$

and $\langle \tau_1,\dots,\tau_q|\Omega\rangle^{\perp}$ is involutive. □

In the study of the state-output interactions in a control system of the form (5.1), we consider the distribution

$$Q = \langle f,g_1,\dots,g_m|\text{sp}\{dh_1,\dots,dh_\ell\}\rangle^{\perp}$$

From Lemma (4.8) we deduce that this distribution is invariant under $f,g_1,\dots,g_m$ and we also see that, by definition, it is contained in $\text{sp}\{dh_1,\dots,dh_\ell\}^{\perp}$. If nonsingular, then, according to Lemma (7.6) is also involutive.

Invoking Proposition (5.6), this distribution may be used in order to find locally around each $\bar{x} \in N$ an open neighborhood U of $\bar{x}$ and a coordinate transformation yielding a decomposition of the form (5.7). Let s denote the dimension of Q. Since $Q^{\perp}$ is the smallest codistribution invariant under $f,g_1,\dots,g_m$ which contains $dh_1,\dots,dh_\ell$, then in this case the decomposition we find is maximal, in the sense that it is not possible to find another set of local coordinates $\tilde{\xi}_1,\dots,\tilde{\xi}_{\tilde{s}},\tilde{\xi}_{\tilde{s}+1},\dots,\tilde{\xi}_n$ with $\tilde{s}$ strictly larger than s, with the property that only the last $n-\tilde{s}$ coordinates influence the output. We show now that this corresponds to the fact that points belonging to different slices of the neighborhood U are distinguishable.

(7.8) *Theorem*. Suppose the distribution Q (i.e. the annihilator of the smallest codistribution invariant under $f,g_1,\dots,g_m$ and which contains $dh_1,\dots,dh_\ell$) is nonsingular. Let s denote the dimension of Q. Then, for each $\bar{x} \in N$ it is possible to find a neighborhood U of $\bar{x}$ and a coordinate transformation $\xi = \xi(x)$ defined on U with the following properties

(a) Any two initial states x^a and x^b of U such that

$$\xi_i(x^a) = \xi_i(x^b) \ , \ i = s+1,\dots,n$$

produce identical output functions under any input which keeps the state trajectories evolving on U

(b) Any initial state x of U which cannot be distinguished from $\bar{x}$ under piecewise constant input functions belongs to the slice

$$S_{\bar{x}} = \{x \in U: \xi_i(x) = \xi_i(\bar{x}),\ s+1 \le i \le n\}.$$

Proof. We need only to prove (b). For simplicity, we break up the proof in various steps.

(i) consider a piecewise-constant input function

$$u_i(t) = u_i^k \quad \text{for} \quad t \in [t_1+\dots+t_{k-1},t_1+\dots+t_k)$$

Define the vector field

$$\theta_k = f + \sum_{i=1}^{m} g_i u_i^k$$

and let Φ_t^k denote the corresponding flow. Then, the state reached at time t_k starting from x^o at time $t = 0$ under this input may be expressed as

$$x(t_k) = \Phi_{t_k}^k \circ \dots \circ \Phi_{t_1}^1 (x^o)$$

and the corresponding output y as

$$y_i(t_k) = h_i(x(t_k))$$

Note that this output may be regarded as the value of a mapping

$$F_i^{x^o}: (-\varepsilon,\varepsilon)^k \longrightarrow \mathbb{R}$$

$$(t_1,\dots,t_k) \longmapsto h_i \circ \Phi_{t_k}^k \circ \dots \circ \Phi_{t_1}^1 (x^o)$$

If two initial states x^a and x^b are such that they produce two identical outputs for any possible piecewise constant input, we must have

$$F_i^{x^a}(t_1,\dots,t_k) = F_i^{x^b}(t_1,\dots,t_k)$$

for all possible $(t_1,\dots,t_k)$, with $0 \le t_i < \varepsilon$ for $1 \le i \le k$. From this we deduce that

$$\left(\frac{\partial F_i^{x^a}}{\partial t_1 \dots \partial t_k}\right)_{t_1=\dots=t_k=0} = \left(\frac{\partial F_i^{x^b}}{\partial t_1 \dots \partial t_k}\right)_{t_1=\dots=t_k=0}$$

An easy calculation shows that

$$\left(\frac{\partial F_i^{x^o}}{\partial t_1 \ldots \partial t_k}\right)_{t_1=\ldots=t_k=0} = (L_{\theta_1} \ldots L_{\theta_k} h_i(x))_{x^o}$$

and, therefore, we must have

$$(L_{\theta_1} \ldots L_{\theta_k} h_i(x))_{x^a} = (L_{\theta_1} \ldots L_{\theta_k} h_i(x))_{x^b}$$

(ii) Now, remember that θ_j , $j = 1,\ldots,k$, depends on $(u_1^j,\ldots,u_m^j)$ and that the above equality must hold for all possible choices of $(u_1^j,\ldots,u_m^j) \in \mathbb{R}^m$. By appropriately selecting these $(u_1^j,\ldots,u_m^j)$ one easily arrives at an equality of the form

(7.9) $$(L_{v_1} \ldots L_{v_k} h_i)_{x^a} = (L_{v_1} \ldots L_{v_k} h_i)_{x^b}$$

where $v_1,\ldots,v_k$ are vector fields belonging to the set $\{f,g_1,\ldots,g_m\}$.

For, set $\gamma_2 = L_{\theta_2} \ldots L_{\theta_k} h$. From the equality $(L_{\theta_1}\gamma_2)_{x^a} = (L_{\theta_1}\gamma_2)_{x^b}$ we obtain

$$(L_f\gamma_2)_{x^a} + \sum_{i=1}^{m} (L_{g_i}\gamma_2)_{x^a} u_i^1 = (L_f\gamma_2)_{x^b} + \sum_{i=1}^{m} (L_{g_i}\gamma_2)_{x^b} u_i^1$$

This, due to the arbitrariness of the $u_1^1,\ldots,u_m^1$, implies that

$$(L_v\gamma_2)_{x_a} = (L_v\gamma_2)_{x_b}$$

where v is any vector in the set $\{f,g_1,\ldots,g_m\}$. This procedure can be iterated, by setting $\gamma_3 = L_{\theta_3} \ldots L_{\theta_k} h$. From the above equality one gets

$$(L_v L_f \gamma_3)_{x^a} + \sum_{i=1}^{m} (L_v L_{g_i} \gamma_3)_{x^a} u_i^2 = (L_v L_f \gamma_3)_{x^b} + \sum_{i=1}^{m} (L_v L_{g_i} \gamma_3)_{x^b} u_i^2$$

and, therefore,

$$(L_{v_1} L_{v_2} \gamma_3)_{x^a} = (L_{v_1} L_{v_2} \gamma_3)_{x^b}$$

for all v_1,v_2 belonging to the set $\{f,g_1,\ldots,g_m\}$. Finally, one arrives at (7.9).

(iii) Let U be a neighborhood of $\bar{x}$ on which a coordinate transformation $\xi(x)$ is defined which makes the condition

(7.10) $$Q(x) = \text{span}\left\{\left(\frac{\partial}{\partial \xi_1}\right)_x, \ldots, \left(\frac{\partial}{\partial \xi_s}\right)_x\right\}$$

satisfied for all $x \in U$. From Remark (7.5), we know that there exists an open subset U^* of U, dense in U , with the property that, around each $x' \in U^*$ it is possible to find a set of $n-s$ real-valued functions $\lambda_1,\ldots,\lambda_{n-s}$ which have the form

(7.11) $$\lambda_i = L_{v_r} \ldots L_{v_1} h_j$$

with $v_1,\ldots,v_r$ vector fields in $\{f,g_1,\ldots,g_m\}$ and $1 \leq j \leq \ell$, such that

$$Q^{\perp}(x') = \text{span}\{d\lambda_1(x'),\ldots,d\lambda_{n-s}(x')\}$$

Since $Q^{\perp}(x')$ has dimension $n-s$, it follows that the tangent covectors $d\lambda_1(x'),\ldots,d\lambda_{n-s}(x')$ are linearly independent.

In the local coordinates which satisfy (7.10), $\lambda_1,\ldots,\lambda_{n-s}$ are functions only of $\xi_{s+1},\ldots,\xi_n$ (see (5.7)). Therefore, we may deduce that the mapping

$$\Lambda : (\xi_{s+1},\ldots,\xi_n) \longmapsto (\lambda_1(\xi_{s+1},\ldots,\xi_n),\ldots,\lambda_{n-s}(\xi_{s+1},\ldots,\xi_n))$$

has a jacobian matrix which is square and nonsingular at $(\xi_{s+1}(x'),\ldots,\xi_n(x'))$.

The mapping Λ is thus locally injective. We may use this property to deduce that, for some suitable neighborhood U' of x', any other point x'' of U' such that

$$\lambda_i(x') = \lambda_i(x'')$$

for $1 \leq i \leq n-s$, must be such that

$$\xi_{s+i}(x'') = \xi_{s+i}(x')$$

for $1 \leq i \leq n-s$, i.e. must belong to the slice of U passing through x'. This, in view of the results proved in (ii) completes the proof in the case where $\bar{x} \in U^*$.

(iv) Suppose $\bar{x} \notin U^*$. Let $x(\bar{x},T,u)$ denote the state reached at time $t=T$ under the action of the piecewise constant input function u. If T is sufficiently small, $x(\bar{x},T,u)$ is still in U. Suppose $x(\bar{x},T,u) \in U^*$. Then, using the conclusions of (iii), we deduce that in some neighborhood U' of $x' = x(\bar{x},T,u)$, the states indistinguishable from x' lie on the slice of U passing through x'.

Now, recall that the mapping

$$\Phi : x^o \longmapsto x(x^o,T,u)$$

is a local diffeomorphism. Thus, there exists a neighborhood $\tilde{U}$ of $\bar{x}$ whose (diffeomorphic) image under Φ is a neighborhood $U^* \subset U'$ of x'.

Let $\bar{\bar{x}}$ denote a point of $\tilde{U}$ indistinguishable from $\bar{x}$ under piecewise constant inputs. Then, clearly, also $x" = x(\bar{\bar{x}},T,u)$ is indistinguishable from $x(\bar{x},T,u) = x'$. From the previous discussion we know that $x"$ and x' belong to the same slice of U. But this implies also that $\bar{x}$ and $\bar{\bar{x}}$ belong to the same slice of U. Thus the proof is completed, provided that

(7.12) $$x(\bar{x},T,u) \in U^*$$

(v) All we have to show now is that (7.12) can be satisfied. For, suppose $R(\bar{x})$, the set of states reachable from $\bar{x}$ under piecewise constant control along trajectories entirely contained in U, is such that

(7.13) $$R(\bar{x}) \cap U^* = \emptyset$$

If this is true, we know from Theorem (6.15) that it is possible to find an r-dimensional embedded submanifold V of U entirely contained in $R(\bar{x})$ and therefore such that $V \cap U^* = \emptyset$. For any choice of functions $\lambda_1,\ldots,\lambda_{n-s}$ of the form (7.11), at any point $x \in V$ the covectors $d\lambda_1(x),\ldots,d\lambda_{n-s}(x)$ are linearly dependent. Thus, without loss of generality, we may assume that there exist $d < n-s$ functions $\gamma_1,\ldots,\gamma_d$ still of the form (7.11) such that, for some open subset V' of V,

- $\text{span}\{dh_1(x),\ldots,dh_\ell(x)\} \subset \text{span}\{d\gamma_1(x),\ldots,d\gamma_d(x)\}$ for all $x \in V'$
- $d\gamma_1(x),\ldots,d\gamma_d(x)$ are linearly independent covectors at all $x \in V'$,
- $dL_v\gamma_j(x) \in \text{span}\{d\gamma_1(x),\ldots,d\gamma_d(x)\}$ for all $x \in V'$ and $v \in \{f,g_1,\ldots,g_m\}$.

Now, we define a codistribution on N as follows

$$\Omega(x) = Q^\perp(x) \qquad \text{for } x \notin V'$$

$$\Omega(x) = \text{span}\{d\gamma_1(x),\ldots,d\gamma_d(x)\} \text{ for } x \in V'$$

Using the fact that $f,g_1,\ldots,g_m$ are tangent to V', it is not difficult to verify that this codistribution is invariant under $f,g_1,\ldots,g_m$, contains $\text{sp}\{dh_1,\ldots,dh_\ell\}$ and is smaller than $\langle f,g_1,\ldots,g_m|\text{sp}\{dh_1,\ldots,dh_\ell\}\rangle$. This is a contradiction and therefore (7.13) must be false. □

CHAPTER IV

DISTURBANCE DECOUPLING AND NON INTERACTING CONTROL

1. Nonlinear Feedback and Controlled Invariant Distributions

In this and in the following chapters, we assume that in the control system

(1.1) $$\dot{x} = f(x) + \sum_{i=1}^{m} g_i(x)u_i$$

it is possible to assign the values of the inputs $u_1,\ldots,u_m$ at each time t as functions of the value at t of the state x and, possibly, of some other real-valued functions $v_1,\ldots,v_m$. This control mode is called a static state-feedback control. In order to preserve the structure of (1.1), we let u_i depend on x and $v_1,\ldots,v_m$ in the following form

(1.2) $$u_i = \alpha_i(x) + \sum_{j=1}^{m} \beta_{ij}(x)v_j$$

where $\alpha_i(x)$ and $\beta_{ij}(x)$, $1 \le i,j \le m$, are real-valued smooth functions defined on the same open subset N of $\mathbb{R}^n$ on which (1.1) is defined.

In doing this we modify the original dynamics (1.1) and obtain the control system

(1.3) $$\dot{x} = \tilde{f}(x) + \sum_{i=1}^{m} \tilde{g}_i(x)v_i$$

in which

(1.4a) $$\tilde{f}(x) = f(x) + \sum_{i=1}^{m} g_i(x)\alpha_i(x)$$

(1.4b) $$\tilde{g}_i(x) = \sum_{j=1}^{m} g_j(x)\beta_{ji}(x)$$

For reasons of notational simplicity, most of the times we consider $\alpha_i(x)$ as the i-th entry of an m-dimensional vector $\alpha(x)$, $\beta_{ij}(x)$ as the (i,j)-th entry of an m×m-dimensional matrix $\beta(x)$ and we consider the vector fields $g_j(x)$ and $\tilde{g}_j(x)$ as j-th columns of n×m-dimensional matrices $g(x)$ and $\tilde{g}(x)$.

In this way we may replace (1.4) with the shorter expressions

(1.5a) $$\tilde{f}(x) = f(x) + g(x)\alpha(x)$$

(1.5b) $$\tilde{g}(x) = g(x)\beta(x)$$

We also systematically assume that the m×m matrix $\beta(x)$ is invertible for all x. This makes it possible to invert the transformation (1.5), and to obtain

(1.6a) $$f(x) = \tilde{f}(x) - \tilde{g}(x)\beta^{-1}(x)\alpha(x)$$

(1.6b) $$g(x) = \tilde{g}(x)\beta^{-1}(x)$$

(1.8) *Remark.* Strictly speaking, only (1.5a) may be regarded as a "feedback", while (1.5b) should be regarded as a change of coordinates in the space of input values, depending on x. □

The purpose for which feedback is introduced is to obtain a dynamics with some nice properties that the original dynamics does not have. As we shall see later on, a typical situation is the one in which a modification is required in order to obtain the invariance of a given distribution Δ under the vector fields which characterize the new dynamics. This kind of problem is usually dealt with in the following way.

A distribution Δ is said to be *controlled invariant* on N if there exists a feedback pair (α,β) defined on N with the property that Δ is invariant under the vector fields $\tilde{f},\tilde{g}_1,\ldots,\tilde{g}_m$ (see (1.4)), i.e. if

(1.9a) $$[\tilde{f},\Delta](x) \subset \Delta(x)$$

(1.9b) $$[\tilde{g}_i,\Delta](x) \subset \Delta(x) \quad \text{for} \quad 1 \le i \le m$$

for all x ∈ N.

A distribution Δ is said to be *locally controlled invariant* if for each x ∈ N there exists a neighborhood U of x with the property that Δ is controlled invariant on U. In view of the previous definition, this requires the existence of a feedback pair (α,β) defined on U such that (1.9) is true for all x ∈ U.

The notion of local controlled invariance lends itself to a simple geometric test. If we set

$$G = \mathrm{sp}\{g_1,\ldots,g_m\}$$

we may express the test in question in the following terms.

(1.10) *Lemma*. Let Δ be an involutive distribution. Suppose Δ, G and $\Delta+G$ are nonsingular on N. Then Δ is locally controlled invariant if and only if

$$(1.11a) \qquad [f,\Delta] \subset \Delta + G$$

$$(1.11b) \qquad [g_i,\Delta] \subset \Delta + G \quad \text{for} \quad 1 \le i \le m$$

Proof. Necessity. Suppose Δ is locally controlled invariant. Let $x \in N$, U a neighborhood of x and (α,β) a feedback pair defined on U which makes (1.9) satisfied on U. Let τ be any vector field of Δ. Then we have

$$[\tilde{f},\tau] = [f+g\alpha,\tau] = [f,\tau] + \sum_{j=1}^{m}[g_j,\tau]\alpha_j + \sum_{j=1}^{m}(L_\tau\alpha_j)g_j$$

$$[\tilde{g}_i,\tau] = \sum_{j=1}^{m}[g_j\beta_{ji},\tau] = \sum_{j=1}^{m}[g_j,\tau]\beta_{ji} + \sum_{j=1}^{m}(L_\tau\beta_{ji})g_j$$

for $1 \le i \le m$.

Since β is invertible, one may solve the last m equalities for $[g_j,\tau]$, obtaining

$$[g_j,\tau] \in \sum_{i=1}^{m}[\tilde{g}_i,\Delta] + G$$

for $1 \le j \le m$. Therefore, from (1.9b) we deduce (1.11b). Moreover, since

$$[f,\tau] \in [\tilde{f},\Delta] + \sum_{i=1}^{m}[g_i,\Delta] + G$$

again from (1.9) and (1.11b) we deduce (1.11a). □

In order to prove the sufficiency, we first need the following interesting result, which is a consequence of Frobenius Theorem.

(1.12) *Theorem*. Let U and V be open sets in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. Let $x_1,\ldots,x_m$ denote coordinates of a point x in $\mathbb{R}^m$ and $y_1,\ldots,y_n$ coordinates of a point y in $\mathbb{R}^n$. Let $\Gamma^1,\ldots,\Gamma^m$ be smooth functions

$$\Gamma^i : U \to \mathbb{R}^{n\times n}$$

Consider the set of partial differential equations

$$(1.13) \qquad \frac{\partial y(x)}{\partial x_i} = \Gamma^i(x)y(x) \qquad 1 \le i \le m$$

where y denotes a function

$$y : U \to V$$

Given a point $(x^o,y^o) \in U\times V$ there exist a neighborhood U_o of x in U and a unique smooth function

$$y : U_o \to V$$

which satisfies the equations (1.13) and is such that $y(x^o) = y^o$ if and only if the functions $\Gamma^1,\ldots,\Gamma^n$ satisfy the conditions

$$(1.14) \qquad \frac{\partial\Gamma^i}{\partial x_k} - \frac{\partial\Gamma^k}{\partial x_i} + \Gamma^i\Gamma^k - \Gamma^k\Gamma^i = 0 \qquad 1 \le i,k \le m$$

for all $x \in U$.

Proof. Necessity. Suppose that for all (x^o,y^o) there is a function y which satisfies (1.13). Then from the property

$$\frac{\partial^2 y}{\partial x_i\partial x_k} = \frac{\partial^2 y}{\partial x_k\partial x_i}$$

one has

$$\frac{\partial}{\partial x_i}(\Gamma^k(x)y(x)) = \frac{\partial}{\partial x_k}(\Gamma^i(x)y(x))$$

Expanding the derivatives on both sides and evaluating them at $x = x^o$ one obtains

$$[(\frac{\partial\Gamma^k}{\partial x_i})_{x^o} + \Gamma^k(x^o)\Gamma^i(x^o)]y^o = [(\frac{\partial\Gamma^i}{\partial x_k})_{x^o} + \Gamma^i(x^o)\Gamma^k(x^o)]y^o$$

which, due to arbitrariness of x^o,y^o, yields the condition (1.14).

Sufficiency. The proof of this part consists of the following steps.

(i) It is shown that the fulfillment of (1.14) enables us to define on $U\times V$ a certain involutive distribution Δ, of dimension m.

(ii) Using Frobenius Theorem, one can find a neighborhood $U'\times V'$ of

(x^o,y^o) and a local coordinates tranformation

$$F : (x,y) \mapsto \xi$$

defined on $U'\times V'$, with the property that

$$\Delta(x,y) = \mathrm{span}\{(\frac{\partial}{\partial\xi_1})_{(x,y)},\ldots,(\frac{\partial}{\partial\xi_m})_{(x,y)}\}$$

for all $(x,y) \in U'\times V'$

(iii) From the transformation F one constructs a solution of (1.13).

As for the step (i), the distribution Δ is defined, at each $(x,y) \in U\times V$, by

$$\Delta(x,y) = \mathrm{span}\{(\frac{\partial}{\partial x_i}) + \sum_{h=0}^{n}\sum_{k=0}^{n}\Gamma^i_{hk}(x)y_k(\frac{\partial}{\partial y_h}) : 1 \le i \le m\}$$

In other words, $\Delta(x,y)$ is spanned by m tangent vectors whose coordinates with respect to the canonical basis $\{(\frac{\partial}{\partial x_1}),\ldots,(\frac{\partial}{\partial x_m})$, $(\frac{\partial}{\partial y_1}),\ldots,(\frac{\partial}{\partial y_n})\}$ of the tangent space to $U\times V$ at (x,y) have the form

$$\begin{pmatrix}1\\0\\\vdots\\0\\\Gamma^1(x)y\end{pmatrix},\begin{pmatrix}0\\1\\\vdots\\0\\\Gamma^2(x)y\end{pmatrix},\ldots,\begin{pmatrix}0\\0\\\vdots\\1\\\Gamma^m(x)y\end{pmatrix}$$

These m vectors are linearly independent at all (x,y) and so the distribution Δ is nonsingular and of dimension m. Moreover, it is an easy computation to check that if the "integrability" condition (1.14) is satisfied, then Δ is involutive.

The possibility of constructing the coordinate transformation described in (ii) is a straightforward consequence of Frobenius theorem. The function F thus defined is such that if v is a vector in Δ, the last n components of F_*v are vanishing. Since, moreover, the tangent vectors $(\frac{\partial}{\partial y_1}),\ldots,(\frac{\partial}{\partial y_n})$ span a subspace which is complementary to $\Delta(x,y)$ at all (x,y) and F is nonsingular, one may easily conclude that the function

$$\xi = F(x,y)$$

is such that the jacobian matrix

$$(1.15) \qquad \begin{pmatrix}\frac{\partial\xi_{m+1}}{\partial y_1} & \cdots & \frac{\partial\xi_{m+1}}{\partial y_n}\\ & \cdots & \\ \frac{\partial\xi_{m+n}}{\partial y_1} & \cdots & \frac{\partial\xi_{m+n}}{\partial y_n}\end{pmatrix}$$

is nonsingular at all $(x,y) \in U'\times V'$.

Without loss of generality we may assume that

$$\xi_i(x^o,y^o) = 0$$

for all $m+1 \le i \le m+n$. As a consequence, the integral submanifold of Δ passing through (x^o,y^o) is defined by the set of equations

$$\xi_{m+i}(x,y) = 0 \qquad 1 \le i \le n$$

Since the matrix (1.15) is nonsingular, thanks to the implicit function theorem the above equations may be solved for y, yielding a set of functions

$$(1.16) \qquad y_i = \eta_i(x) \qquad 1 \le i \le n$$

defined in a neighborhood $U_o \subset U'$ of x^o. Moreover

$$\eta_i(x^o) = y_i^o \qquad 1 \le i \le n$$

The functions (1.16) satisfy the differential equations (1.13) and therefore, are the required solutions. As a matter of fact, the functions

$$\varphi_i(x,y) = y_i - \eta_i(x) \qquad 1 \le i \le n$$

are constant on the integral submanifold of Δ passing through (x^o,y^o) and, therefore, if v is a vector in Δ,

$$d\varphi_i v = 0 \qquad 1 \le i \le n$$

at all pairs $(x,\eta(x))$. These equations, taking for v each one of the n vectors used to define Δ, yield exactly

$$\frac{\partial \eta_i}{\partial x_j} = (\Gamma^j(x)\eta(x))_i \qquad 1 \le i \le n, \quad 1 \le j \le m. \ \square$$

Proof. (of Lemma 1.10). Sufficiency. Recall that, by assumption, Δ, G and $\Delta+G$ are nonsingular; let d denote the dimension of Δ and let

$$p = \dim G - \dim \Delta \cap G$$

Given any $x^o \in N$ it is possible to find a neighborhood U of x^o and an $m \times m$ nonsingular matrix B, whose (i,j)-th element b_{ij} is a smooth real-valued function defined on U, such that, for

$$\hat{g}_i = \sum_{j=1}^{m} g_j b_{ji} \qquad 1 \le i \le m$$

the following is true

$$\mathrm{sp}\{\hat{g}_{p+1}, \ldots, \hat{g}_m\} \subset \Delta$$

$$(1.17) \qquad (\Delta+G) = \Delta \oplus \mathrm{sp}\{\hat{g}_1, \ldots, \hat{g}_p\}$$

The tangent vectors $\hat{g}_1(x), \ldots, \hat{g}_p(x)$ are clearly linearly independent at all $x \in U$.

Now, observe that if the assumption (1.11b) is satisfied, then also

$$(1.18) \qquad [\hat{g}_i, \Delta] \subset \Delta + G$$

and let $\tau_1, \ldots, \tau_d$ be a set of vector fields which locally span Δ around x^o. From (1.17) and (1.18) we deduce the existence of a unique set of smooth real-valued functions c_{ji}^k, defined locally around x^o, and a vector field $\delta_i^k \in \Delta$ defined locally around x^o such that

$$(1.11b') \qquad [\hat{g}_i, \tau_k] = \sum_{j=1}^{p} c_{ji}^k \hat{g}_j + \delta_i^k$$

for all $1 \le i \le m$ and $1 \le k \le d$. Using the same arguments and setting

$$\hat{g}_0 = f$$

from (1.11a) and (1.18) we deduce the existence of a unique set of real-valued smooth functions c_{j0}^k and a vector field $\delta_0^k \in \Delta$, defined locally around x^o, such that

$$(1.11a') \qquad [\hat{g}_0, \tau_k] = \sum_{j=1}^{p} c_{j0}^k \hat{g}_j + \delta_0^k$$

Now, suppose there exists a nonsingular $m \times m$ matrix $\hat{B}$, whose (i,j)-th element $\hat{b}_{ij}$ is a smooth real-valued function defined locally around x, such that

$$(1.19) \qquad -L_{\tau_k} \hat{b}_{hi} + \sum_{j=1}^{m} c_{hj}^k \hat{b}_{ji} = 0$$

for $1 \le k \le d$, $1 \le h \le p$, $1 \le i \le m$. Then, it is easy to see that

$$(1.20) \qquad [\sum_{h=1}^{m} \hat{g}_h \hat{b}_{hi}, \tau_k] \in \Delta$$

for $1 \le i \le m$, $1 \le k \le d$. For,

$$[\sum_{h=1}^{m} \hat{g}_h \hat{b}_{hi}, \tau_k] = -\sum_{h=1}^{m} (L_{\tau_k} \hat{b}_{hi}) \hat{g}_h + \sum_{j=1}^{m} \hat{b}_{ji} [\hat{g}_j, \tau_k]$$

$$= -\sum_{h=1}^{p} (L_{\tau_k} \hat{b}_{hi}) \hat{g}_h + \sum_{j=1}^{m} \hat{b}_{ji} \sum_{h=1}^{p} c_{hj}^k \hat{g}_h + \bar{\delta}_i^k = \bar{\delta}_i^k$$

where $\bar{\delta}_i^k$ is a vector field in Δ. Since $\tau_1, \ldots, \tau_k$ locally span Δ, (1.20) implies that

$$[\sum_{h=1}^{m} \hat{g}_h \hat{b}_{hi}, \Delta] \subset \Delta$$

Therefore, the matrix

$$\beta = B\hat{B}$$

is such that (1.9b) is satisfied.

Using similar arguments, one can see that if there exists an $m \times 1$ vector $\hat{a}$, whose i-th element $\hat{a}_i$ is a smooth real-valued function defined locally around x^o, such that

$$(1.21) \qquad -L_{\tau_k} \hat{a}_h + \sum_{j=1}^{m} c_{hj}^k \hat{a}_j + c_{h0}^k = 0$$

for $1 \le k \le d$, $1 \le h \le p$, then

$$(1.22) \qquad [\hat{g}_0 + \sum_{h=1}^{m} \hat{g}_h \hat{a}_h, \tau_k] \in \Delta$$

for $1 \leq k \leq d$. For,

$$[\hat{g}_0 + \sum_{h=1}^{m} \hat{g}_h \bar{a}_h, \tau_k] = -\sum_{h=1}^{p} (L_{\tau_k} \bar{a}_h)\bar{g}_h + \sum_{j=1}^{m} \hat{a}_j \sum_{h=1}^{p} c_{hj}^k \bar{g}_h + \sum_{h=1}^{p} c_{h0}^k \hat{g}_h + \bar{\delta}^k = \bar{\delta}^k$$

where $\bar{\delta}^k$ is a vector field in Δ. From this one deduces that the vector

$$\alpha = B\hat{a}$$

is such that (1.9a) is satisfied.

Thus, we have seen that the possibility of finding $\hat{B}$ and $\hat{a}$ which satisfy (1.19) and (1.21) enables us to construct a pair of feedback functions that makes (1.9) satisfied. In order to complete the proof, we have to show that (1.19) and (1.21) can be solved for $\hat{B}$ and $\hat{a}$.

Since Δ is nonsingular and involutive, we may assume, without loss of generality, that our choice of local coordinates is such that

$$\tau_k = \frac{\partial}{\partial x_k} \qquad 1 \leq k \leq d.$$

The equations (1.19) and (1.21) may be rewritten as a set of partial differential equations of the form (1.13) by simply setting

$$\Gamma^k = \begin{pmatrix} c_{11}^k & \cdots & c_{1m}^k & c_{10}^k \\ \cdot & \cdots & \cdot & \cdot \\ c_{p1}^k & \cdots & c_{pm}^k & c_{p0}^k \\ 0 & \cdots & 0 & 0 \\ \cdot & \cdots & \cdot & \cdot \\ 0 & \cdots & 0 & 0 \end{pmatrix} \qquad 1 \leq k \leq d$$

As a matter of fact, for each fixed i, the equations (1.19) correspond to an equation for the i-th column of $\hat{B}$, of the form

$$\text{(1.23)} \qquad \frac{\partial}{\partial x_k}\begin{pmatrix} \hat{b}_i \\ 0 \end{pmatrix} = \Gamma^k \begin{pmatrix} \hat{b}_i \\ 0 \end{pmatrix} \qquad 1 \leq k \leq d$$

(where $\hat{b}_i$ stands for the i-th column of $\hat{B}$) and the equations (1.21) correspond to

$$\text{(1.24)} \qquad \frac{\partial}{\partial x_k}\begin{pmatrix} \hat{a} \\ 1 \end{pmatrix} = \Gamma^k \begin{pmatrix} \hat{a} \\ 1 \end{pmatrix} \qquad 1 \leq k \leq d$$

Both these equations have exactly the form

$$\text{(1.25)} \qquad \frac{\partial y}{\partial x_k} = \Gamma^k y \qquad 1 \leq k \leq d$$

the unknown vector y being $m+1$ dimensional. Since now the functions Γ^k depend also on the coordinates $x_{d+1}, \ldots, x_n$ (with respect to which no derivative of y is considered), in order to achieve uniqueness, the value of y must be specified, for a given $x_1^o, \ldots, x_d^o$, at each $x_{d+1}, \ldots, x_n$. For consistency, the last component of the initial value of the solution sought for the equations (1.23) must be set equal to zero, whereas the last component of the initial value of the solution sought for the equation (1.24) must be set equal to 1. In addition, the first m components of the initial values of the solutions sought for each of the equations (1.23) must be columns of a nonsingular $m \times m$ matrix, in order to let $\hat{B}$ be nonsingular.

The solvability of an equation of the form (1.25) depends, as we have seen, on the fulfillment of the integrability conditions (1.14). This, in turn, is implied by (1.11). Consider the Jacobi identity

$$-[[\hat{g}_i, \tau_k], \tau_h] + [[\hat{g}_i, \tau_h], \tau_k] = [\hat{g}_i, [\tau_h, \tau_k]]$$

for any $0 \leq i \leq m$. Using for $[\hat{g}_i, \tau_k]$ and $[\hat{g}_i, \tau_h]$ the expressions given by (1.11a') or (1.11b') and taking $\tau_k = \frac{\partial}{\partial x_k}$, $\tau_h = \frac{\partial}{\partial x_h}$ one easily obtains

$$[\sum_{j=1}^{p} c_{ji}^k \hat{g}_j + \delta_i^k, \frac{\partial}{\partial x_h}] - [\sum_{j=1}^{p} c_{ji}^h \hat{g}_j + \delta_i^h, \frac{\partial}{\partial x_k}] = 0$$

This yields

$$-\sum_{j=1}^{p} \frac{\partial c_{ji}^k}{\partial x_h} \hat{g}_j + \sum_{j=1}^{p} c_{ji}^k (\sum_{\ell=1}^{p} c_{\ell j}^h \hat{g}_\ell + \delta_j^h) + [\delta_i^k, \frac{\partial}{\partial x_h}]$$

$$+ \sum_{j=1}^{p} \frac{\partial c_{ji}^h}{\partial x_k} \hat{g}_j - \sum_{j=1}^{p} c_{ji}^h (\sum_{\ell=1}^{p} c_{\ell j}^k \hat{g}_\ell + \delta_j^k) - [\delta_i^h, \frac{\partial}{\partial x_k}] = 0$$

Now, recall that $\frac{\partial}{\partial x_h}$ and $\frac{\partial}{\partial x_k}$ are both vector fields of Δ, which is involutive. Therefore, also $[\delta_i^k, \frac{\partial}{\partial x_h}]$ and $[\delta_i^h, \frac{\partial}{\partial x_k}]$ are in Δ. Since Δ and $\text{sp}\{\hat{g}_1, \ldots, \hat{g}_p\}$ are direct summands and $\hat{g}_1, \ldots, \hat{g}_p$ are linearly independent, the previous equality implies

$$-\frac{\partial c_{ji}^k}{\partial x_h} + \frac{\partial c_{ji}^h}{\partial x_k} + \sum_{\ell=1}^{p} c_{j\ell}^h c_{\ell i}^k - \sum_{\ell=1}^{p} c_{j\ell}^k c_{\ell i}^h = 0$$

for $1 \leq j \leq p$, $0 \leq i \leq m$, $1 \leq h,k \leq d$, which is easily seen to be identical to the condition (1.14). □

We see from this Lemma that, under reasonable assumptions (namely, the nonsingularity of Δ, G and $\Delta+G$) an involutive distribution is locally controlled invariant if and only if the conditions (1.11) are satisfied. These conditions are of special interest because they don't invoke the existence of feedback functions α and β, as the definition does, but are expressed only in terms of the vector fields $f, g_1, \ldots, g_m$ which characterize the given control system and of the distribution itself. The fulfillment of conditions (1.11) implies the existence of a pair of feedback functions which make Δ invariant under the new dynamics but the actual construction of such a feedback pair generally involves the solution of a set of partial differential equations, as we have seen in the proof of Lemma (1.10). There are cases, however, in which the solution of partial differential equations may be avoided and these, luckyly enough, include some situations of great importance in control theory. These will be examined later on in this chapter.

2. The Disturbance Decoupling Problem

The notion of locally controlled invariance will now be used in order to solve the following control problem. Consider a control system

(2.1a) $$\dot{x} = f(x) + \sum_{i=1}^{m} g_i(x)u_i + p(x)w$$

(2.1b) $$y = h(x)$$

where the additional input w represents an undesired perturbation, which influences the behavior of the system through the vector field p. The system is to be modified, via static state-feedback control on the inputs $u_1, \ldots, u_m$, in such a way that the disturbance w has no influence on the output y.

In view of some earlier results (Theorem III.(3.12) and Remark III.(3.13)) this problem consists in finding a feedback pair (α,β) and a distribution Δ which is invariant under $\tilde{f} = f+g\alpha$ and $\tilde{g}_i = (g\beta)_i$, $1 \leq i \leq m$, contains the vector field p and is contained in $(\mathrm{sp}\{dh_j\})^{\perp}$ for all $1 \leq j \leq \ell$.

According to the terminology introduced in the previous section, a distribution Δ which is invariant under $\tilde{f} = f+g\alpha$ and $\tilde{g}_i = (g\beta)_i$, $1 \leq i \leq m$, for some feedback (α,β) is *controlled invariant*. If we set

$$H = \bigcap_{j=1}^{\ell} (\mathrm{sp}\{dh_j\})^{\perp} = (\mathrm{sp}\{dh_1, \ldots, dh_\ell\})^{\perp}$$

we may express the problem in question in the following terms.

Disturbance decoupling problem. Find a distribution Δ which

(i) is controlled invariant

(ii) is such that $p \in \Delta \subset H$. □

As we have seen in the previous section, the notion of *local* controlled invariance is sometimes easier to deal with than (global) controlled invariance. This motivates the consideration of the following problem.

Local disturbance decoupling problem. Find a distribution Δ which

(i) is locally controlled invariant

(ii) is such that $p \in \Delta \subset H$.

(2.2) *Remark.* Note that the distribution Δ is not required to be nonsingular, neither involutive. However, nonsingularity and involutivity may be needed in order to construct the pair of feedback functions (α,β) which make it possible to implement the disturbance-decoupling control mode. This typically happens when one has found a distribution Δ which satisfies (ii) and, instead of (i), satisfies the condition

(i') $$[f,\Delta] \subset \Delta + G$$

$$[g_i,\Delta] \subset \Delta + G \qquad 1 \leq i \leq m$$

In this case, we know from Lemma (1.10) that nonsingularity of Δ, G and $\Delta+G$ helps in finding at least locally a pair of feedback functions (α,β) with the desired properties.

If Δ is nonsingular and involutive, invariant under $\tilde{f}$ and $\tilde{g}_i$, $1 \leq i \leq m$, and satisfies (ii), then it is known from the analysis developed in chapter I that there exist local coordinate transformations which put the closed-loop system into the form

(2.3) $$\begin{aligned} \dot{x}_1 &= \tilde{f}_1(x_1,x_2) + \sum_{i=1}^{m} \tilde{g}_{i1}(x_1,x_2)u_i + p_1(x_1,x_2)w \\ \dot{x}_2 &= \tilde{f}_2(x_2) + \sum_{i=1}^{m} \tilde{g}_{i2}(x_2)u_i \\ y &= h(x_2) \end{aligned}$$

Here, once again, one sees that the disturbance w has no influence on the output y. □

A systematic way to deal with the Disturbance Decoupling Problem is to examine first whether or not the family of all controlled invariant distributions contained in H has a "maximal" element (an element which contains all other members of the family). For, if this is true, then the problem is solved if and only if this maximal element contains the vector field p.

If, rather than controlled invariant distributions, we look at *locally* controlled invariant distributions, then the existence of such a maximal element may be shown under rather mild assumptions. To this end, we introduce a notation and an algorithm. Let $\mathbb{I}(f,g;K)$ denote the collection of all smooth distributions which are contained in a given distribution K and satisfy the conditions (1.11). In view of Lemma (1.10), the maximal element of $\mathbb{I}(f,g;K)$ is the natural candidate for the maximal locally controlled invariant distribution in K. As a matter of fact, the maximal element of $\mathbb{I}(f,g;K)$ may be found by means of the following algorithm.

(2.4) *Lemma* (Controlled Invariant Distribution Algorithm). Let

$$\Omega_0 = K^{\perp}$$

(2.5)

$$\Omega_k = \Omega_{k-1} + L_f(G^{\perp} \cap \Omega_{k-1}) + \sum_{i=1}^{m} L_{g_i}(G^{\perp} \cap \Omega_{k-1})$$

Suppose there exists an integer k^* such that $\Omega_{k^*} = \Omega_{k^*+1}$. Then $\Omega_k = \Omega_{k^*}$ for all $k > k^*$.

If $\Omega_{k^*} \cap G^{\perp}$ and $\Omega_{k^*}^{\perp}$ are smooth, then $\Omega_{k^*}^{\perp}$ is the maximal element of $\mathbb{I}(f,g;K)$.

Proof. The first part of the statement is a trivial consequence of the definitions. As for the other, note first that from the equality $\Omega_{k^*+1} = \Omega_{k^*}$ we deduce

$$L_{g_i}(G^{\perp} \cap \Omega_{k^*}) \subset \Omega_{k^*}$$

for $1 \le i \le m$ and also for $i = 0$ if we set $f = g_0$, as sometimes we did before. Let ω be a one-form in $G^{\perp} \cap \Omega_{k^*}$, and τ a vector field in $\Omega_{k^*}^{\perp}$. In the expression

$$\langle L_{g_i}\omega, \tau \rangle = L_{g_i}\langle \omega, \tau \rangle - \langle \omega, [g_i, \tau] \rangle$$

we have

$$\langle L_{g_i}\omega, \tau \rangle = 0$$

because $L_{g_i}\omega \in \Omega_{k^*}$ and

$$\langle \omega, \tau \rangle = 0$$

because $\tau \in \Omega_{k^*}^{\perp} + G$. Thus

$$\langle \omega, [g_i, \tau] \rangle = 0$$

Since $G^{\perp} \cap \Omega_{k^*}$ is smooth by assumption, $[g_i, \tau]$ annihilates every covector in $G^{\perp} \cap \Omega_{k^*}$, i.e.

$$[g_i, \tau] \in \Omega_{k^*}^{\perp} + G$$

for $0 \le i \le m$. Thus, $\Omega_{k^*}^{\perp}$ is a member of $\mathbb{I}(f,g;K)$. Let $\bar{\Delta}$ be any other element of this collection. We will prove that $\bar{\Delta} \subset \Omega_{k^*}^{\perp}$. First of all, note that if ω is a one-form in $\bar{\Delta}^{\perp} \cap G^{\perp}$ and τ a vector field in $\bar{\Delta}$ we have

$$\langle L_{g_i}\omega, \tau \rangle = 0$$

so that (recall that $\bar{\Delta}$ is a smooth distribution)

$$L_{g_i}(\bar{\Delta}^{\perp} \cap G^{\perp}) \subset \bar{\Delta}^{\perp}$$

Suppose

$$\bar{\Delta}^{\perp} \supset \Omega_k$$

for some $k \ge 0$. Then

$$\Omega_{k+1} \subset \Omega_k + L_f(\bar{\Delta}^{\perp} \cap G^{\perp}) + \sum_{i=1}^{m} L_{g_i}(\bar{\Delta}^{\perp} \cap G^{\perp}) \subset \bar{\Delta}^{\perp}$$

Thus, since $\Omega_0 = K^{\perp} \subset \bar{\Delta}^{\perp}$, we deduce that

$$\bar{\Delta} \subset \Omega_{k^*}^{\perp}$$

and $\Omega_{k^*}^{\perp}$ is the maximal element of $\mathbb{I}(f,g;K)$. □

For convenience, we introduce a terminology which is useful to remind both the convergence of the sequence (2.5) in a finite number of stages and the dependence of its final element on the distribution K. We set

$$J(K) = (\Omega_0+\Omega_1+\ldots+\Omega_k+\ldots)^{\perp} \tag{2.6}$$

and we say that $J(K)$ is *finitely computable* if there exists an integer k^* such that, in the sequence (2.5), $\Omega_{k^*} = \Omega_{k^*+1}$. If this is the case, then obviously $J(K) = \Omega^{\perp}_{k^*}$.

In the Lemma (2.4) we have seen that if $J(K)$ is finitely computable and if $J(K)^{\perp} \cap G^{\perp}$ and $J(K)$ are smooth, then $J(K)$ is the maximal element of $\mathbb{I}(f,g;K)$. In order to let this distribution be locally controlled invariant all we need are the assumptions of Lemma (1.10), as stated below.

(2.7) *Lemma.* Suppose $J(K)$ is finitely computable. Suppose K is an involutive distribution and G, $J(K)$, $J(K)+G$ are nonsingular. Then $J(K)$ is involutive and is the largest locally controlled invariant distribution contained in K.

Proof. First, observe that the assumption of nonsingularity on G, $J(K)$, $J(K)+G$ indeed implies the smoothness of $J(K)^{\perp} \cap G^{\perp}$ and $J(K)$. So, in view of Lemma (1.10) we need only to show that $J(K)$ is involutive.

For, let d denote the dimension of $J(K)$. At any point x^o one may find a neighborhood U of x^o and vector fields $\tau_1,\ldots,\tau_d$ such that

$$J(K) = \mathrm{sp}\{\tau_1,\ldots,\tau_d\}$$

on U. Consider the distribution

$$D = \mathrm{sp}\{\tau_i\colon 1 \le i \le d\} + \mathrm{sp}\{[\tau_i,\tau_j]\colon 1 \le i,j \le d\}$$

and suppose, for the moment, that D is nonsingular on U. Then, every vector field τ in D can be expressed as the sum of a vector field τ' in $J(K)$ and a vector field τ'' of the form

$$\tau'' = \sum_{i=1}^{d}\sum_{j=1}^{d} c_{ij}[\tau_i,\tau_j]$$

where c_{ij}, $1 \le i,j \le d$, are smooth real-valued functions defined on U.

We want to show that

$$[g_k,D] \subset D + G$$

for all $0 \le k \le m$. In view of the above decomposition of any vector field τ in D, this amounts to show that

$$[g_k,[\tau_i,\tau_j]] \subset D + G$$

The expression of the vector field on the left-hand-side via Jacobi identity yields

$$[g_k,[\tau_i,\tau_j]] = [\tau_i,[g_k,\tau_j]] - [\tau_j,[g_k,\tau_i]]$$

The vector field $[g_k,\tau_j]$ is in $J(K) + G$ and therefore, because of the nonsingularity of $J(K)$ and $J(K) + G$, it can be written as the sum of a vector field τ in $J(K)$ and a vector field g in G. Since, $[\tau_i,g] \in J(K) + G$ for any $g \in G$, we have

$$[\tau_i,[g_k,\tau_j]] = [\tau_i,\tau+g] \in D + J(K) + G = D + G$$

and we conclude that D is such that

$$[g_k,D] \subset D + G$$

for all $0 \le k \le m$.

Now, recall that K is involutive by assumption, and therefore that

$$D \subset K$$

From this and from the previous inclusions we deduce that D is an element of $\mathbb{I}(f,g;K)$. Since $D \supset J(K)$ by construction and $J(K)$ is the maximal element of $\mathbb{I}(f,g;K)$, we see that

$$D = J(K)$$

Thus, any Lie bracket of vector fields of $J(K)$, which is in D by construction, is still in $J(K)$ and the latter is an involutive distribution.

If we drop the assumption that D has constant dimension on U, we can still conclude that D coincides with $J(K)$ on the subset $\tilde{U} \subset U$ consisting of all regular points of D. Then, using Lemma I.(2.11), we can as well prove that $D = J(K)$ on the whole of U. □

In the Local Disturbance Decoupling Problem one is interested in the largest locally controlled invariant distribution contained in H. Since this latter is involutive (see chapter I), in order to be able to use the previous Lemma, we need to assume that the distribution $J(H)$ is finitely computable and that G, $J(H)$, $J(H) + G$ are nonsingular. If this is the case, then, as we said before, the Local Disturbance Decoupling Problem is solvable if and only if

$$p \in J(H)$$

We conclude the section with a remark about the invariance of the algorithm (2.5) under feedback transformation.

(2.8) *Lemma.* Let $\tilde{f},\tilde{g}_1,\dots,\tilde{g}_m$ be any set of vector fields deduced from $f,g_1,\dots,g_m$ by setting $\tilde{f} = f + g\alpha$, $\tilde{g}_i = (g\beta)_i$, $1 \le i \le m$; then each codistribution Ω_k of the sequence (2.5) is such that

$$\Omega_k = \Omega_{k-1} + L_{\tilde{f}}(G^\perp \cap \Omega_{k-1}) + \sum_{i=1}^{m} L_{\tilde{g}_i}(G^\perp \cap \Omega_{k-1})$$

Proof. Recall that, given a covector field ω, a vector field τ and a scalar function γ,

$$L_{(\tau\gamma)}\omega = (L_\tau\omega)\gamma + \langle\omega,\tau\rangle d\gamma$$

If ω is a covector field in $G^\perp \cap \Omega_{k-1}$, then

$$L_{\tilde{f}}\omega = L_f\omega + \sum_{i=1}^{m}(L_{g_i}\omega)\alpha_i + \sum_{i=1}^{m}\langle\omega,g_i\rangle d\alpha_i$$

$$L_{\tilde{g}_i}\omega = \sum_{j=1}^{m}(L_{g_j}\omega)\beta_{ji} + \sum_{j=1}^{m}\langle\omega,g_j\rangle d\beta_{ji}$$

But $\langle\omega,g_j\rangle = 0$ because $\omega \in G^\perp$ and therefore

$$L_{\tilde{f}}(G^\perp \cap \Omega_{k-1}) + \sum_{i=1}^{m} L_{\tilde{g}_i}(G^\perp \cap \Omega_{k-1}) \subset L_f(G^\perp \cap \Omega_{k-1}) + \sum_{i=1}^{m} L_{g_i}(G^\perp \cap \Omega_{k-1})$$

Since β is invertible, one may also write $f = \tilde{f} - \tilde{g}\beta^{-1}\alpha$ and $g_i = (\tilde{g}\beta^{-1})_i$ and, using the same arguments, prove the reverse inclusion. The two sides of inclusion are thus equal and the Lemma is proved. □

3. Some Useful Algorithms

In this section we describe a practical implementation of the algorithm yielding the largest locally controlled invariant distribution contained in H. Moreover, we show that in some particular cases the construction of this distribution may be obtained with simpler methods.

We begin with the easiest situation, first. For each output function $h_i(x)$ we define an integer ρ_i, called the *characteristic number* of y_i, as the integer identified by the conditions

(3.1a) $$L_{g_j}L_f^k h_i(x) = 0$$

for all $k < \rho_i$, all $1 \le j \le m$, all $x \in N$ and

(3.1b) $$L_{g_j}L_f^{\rho_i} h_i(x) \neq 0$$

for some j and x.

Note that if for some output y_i the characteristic number is not defined (i.e. (3.1a) holds for all k, all j and all x), then the output y_j is in no way affected by any of the inputs $u_1,\dots,u_m$. The expansions described in chapter III show that if this is the case

$$y_i(t) = \sum_{k=0}^{\infty} L_f^k h_i(x^o)\frac{t^k}{k!} = h_i(\Phi_t^f(x^o))$$

Thus, it seems reasonable to assume that our control system is such that the characteristic numbers are defined for each output.

Once the characteristic numbers are known, we may define an $\ell\times m$ matrix $A(x)$ whose element $a_{ij}(x)$ on the i-th row and j-th column is

(3.2) $$a_{ij}(x) = L_{g_j}L_f^{\rho_i}h_i(x) = \langle dL_f^{\rho_i}h_i(x), g_j(x)\rangle$$

and an ℓ-vector $b(x)$ whose element $b_i(x)$ on the i-th row is

(3.3) $$b_i(x) = L_f^{\rho_i+1}h_i(x) = \langle dL_f^{\rho_i}h_i(x), f(x)\rangle$$

We point out first of all an interesting property of the objects defined so far

(3.4) *Lemma.* Let (α,β) be any pair of feedback functions and let $\tilde{f} = f + g\alpha$, $\tilde{g}_i = (g\beta)_i$. Then

$$L_{\tilde f}^k h_i(x) = L_f^k h_i(x)$$

for all $k \le \rho_i$ and all $x \in N$. Moreover, let $\tilde A(x)$ be the $\ell\times m$ matrix whose (i,j)-th element a_{ij} is

$$\tilde a_{ij}(x) = L_{\tilde g_j} L_{\tilde f}^{\rho_i} h_i(x)$$

and $\tilde b(x)$ the ℓ-vector whose i-th element b_i is

$$\tilde b_i(x) = L_{\tilde f}^{\rho_i+1} h_i(x)$$

Then

$$\tilde A(x) = A(x)\beta(x)$$

$$\tilde b(x) = A(x)\alpha(x)+b(x)$$

Proof. The first equality is easily proved by induction. It is true for $k = 0$ and, if true for some $0 < k < \rho_i$, yields

$$L_{\tilde f}^{k+1} h_i(x) = L_{\tilde f} L_{\tilde f}^k h_i(x) = L_f^{k+1} h_i(x) + \sum_{j=1}^m L_{g_j} L_f^k h_i(x)\alpha_j(x) = L_f^{k+1} h_i(x)$$

The other equalities are straightforward consequences of the first one.

(3.5) *Remark*. Note that the invertibility of β implies the invariance of the integers $\rho_1,\dots,\rho_\ell$ as well as that of rank of $A(x)$ under feedback transformations. □

From this one can deduce the following interesting result.

(3.6) *Lemma*. Every locally controlled invariant distribution contained in H is also contained in the distribution Δ_{sup} defined by

(3.7)
$$\Delta_{sup} = \bigcap_{i=1}^{\ell} \bigcap_{k=0}^{\rho_i} (sp\{dL_f^k h_i\})^\perp$$

Suppose Δ_{sup} is a smooth distribution. A pair of feedback functions (α,β) is such that

(3.8a)
$$[f + g\alpha, \Delta_{sup}] \subset \Delta_{sup}$$

(3.8b)
$$[(g\beta)_i, \Delta_{sup}] \subset \Delta_{sup} \qquad 1 \le i \le m$$

if and only if the differentials of each entry of the column vector $A(x)\alpha(x) + b(x)$ and those of each entry of the matrix $A(x)\beta(x)$ belong to the codistribution $\Delta_{sup}^\perp$.

Proof. Let Δ be a locally controlled invariant distribution contained in H. Then, by definition, $\Delta \subset (sp\{dh_i\})^\perp$ for all $1 \le i \le \ell$. Moreover, for some local feedback α, $[\tilde f,\Delta] \subset \Delta$. Suppose $\Delta \subset (sp\{dL_f^k h_i\})^\perp$ for some $k < \rho_i$; then using Lemma (3.4) we have for any vector field $\tau \in \Delta$

$$0 = \langle dL_f^k h_i, [\tilde f,\tau]\rangle = L_{\tilde f}\langle dL_f^k h_i,\tau\rangle - \langle dL_{\tilde f} L_f^k h_i,\tau\rangle = \langle dL_f^{k+1} h_i,\tau\rangle$$

i.e. $\Delta \subset (sp\{dL_f^{k+1} h_i\})^\perp$. This proves that

$$\Delta \subset \bigcap_{i=1}^{\ell} \bigcap_{k=0}^{\rho_i} (sp\{dL_f^k h_i\})^\perp$$

and therefore the distribution (3.7) contains every locally controlled invariant distribution.

Now, suppose there exists a pair of feedback functions that makes (3.8) satisfied. Let τ be a vector field in Δ_{sup}. Then

(3.9a)
$$\langle dL_f^k h_i, \tau\rangle = 0$$

(3.9b)
$$\langle dL_f^k h_i, [\tilde f,\tau]\rangle = 0$$

(3.9c)
$$\langle dL_f^k h_i, [\tilde g_j,\tau]\rangle = 0$$

for all $1 \le i \le \ell$, $0 \le k \le \rho_i$, $1 \le j \le m$. From (3.9b) written for $k = \rho_i$, we deduce, using Lemma (3.4),

$$0 = L_{\tilde f}\langle dL_f^{\rho_i} h_i,\tau\rangle - \langle dL_{\tilde f} L_f^{\rho_i} h_i,\tau\rangle = \langle dL_{\tilde f}^{\rho_i+1} h_i,\tau\rangle = \langle d\tilde b_i,\tau\rangle$$

Similarly, for (3.9c) written for $k = \rho_i$ we deduce that

$$0 = \langle d\tilde a_{ij},\tau\rangle$$

Therefore, the differentials of $\tilde b_i$ and $\tilde a_{ij}$ belong to the codistribution $\Delta_{sup}^\perp$. Conversely, if the differentials of $\tilde b_i$ and $\tilde a_{ij}$ belong to the codistribution $\Delta_{sup}^\perp$, we have that (3.9b) and (3.9c) hold for $k = \rho_i$. For values $k < \rho_i$ (3.9b) and (3.9c) hold for any feedback (α,β) because of Lemma (3.4) and, therefore, we deduce that Δ_{sup} is invariant under $\tilde f$ and $\tilde g_i$. □

From this result we see that there are cases in which the computation of the largest controlled invariant distribution contained in H is not terribly difficult. An interesting special case is the one in which the matrix A(x) has a rank equal to the number of its rows (i.e. the number of the output channels); this is explained in the following results.

(3.10) *Lemma.* Suppose that the matrix A(x) has rank ℓ at x^o. Then the covectors

$$dh_1(x^o),\ldots,dL_f^{\rho_1}h_1(x^o),\ldots,dh_\ell(x^o),\ldots,dL_f^{\rho_\ell}h_\ell(x^o)$$

are linearly independent. As a consequence, the distribution Δ_{sup} is nonsingular in a neighborhood U of x^o and

(3.11) $$\dim \Delta_{sup}^{\perp}(x) = \rho_1+\ldots+\rho_\ell+\ell \le n$$

Proof. Suppose that the differentials are linearly dependent at x^o. Then there exist real numbers c_{ik}, $1 \le i \le \ell$, $0 \le k \le \rho_i$ such that

(3.12) $$\sum_{i=1}^{\ell}\sum_{k=0}^{\rho_i} c_{ik}dL_f^k h_i(x^o) = 0$$

Now consider the function

$$\lambda(x) = \sum_{i=0}^{\ell}\sum_{k=0}^{\rho_i} c_{ik}L_f^k h_i(x)$$

According to the definition of $\rho_1,\ldots,\rho_m$, this function is such that

$$\langle d\lambda,g_j\rangle(x) = \sum_{i=1}^{\ell} c_{i\rho_i}\langle dL_f^{\rho_i}h_i,g_j\rangle(x) = \sum_{i=1}^{\ell} c_{i\rho_i}a_{ij}(x)$$

But, on the other hand, (3.12) shows that $d\lambda(x^o) = 0$ and therefore the above equality implies the linear dependence of the rows of the matrix $A(x^o)$, i.e. a contradiction. Therefore we conclude that if (3.12) holds, we must have $c_{1\rho_1} = \ldots = c_{\ell\rho_\ell} = 0$.

Now consider the function

$$\gamma(x) = \sum_{i=0}^{\ell}\sum_{k=0}^{\rho_i-1} c_{ik}L_f^k h_i(x)$$

(with the understanding that the above sum is exendend over all non-negative k's) and observe that, if $0 \le k \le \rho_i-1$, then(*)

$$-\langle dL_f^k h_i,[f,g_j]\rangle = \langle dL_f^{k+1}h_i,g_j\rangle$$

Now, by the definition of $\rho_1,\ldots,\rho_m$ and from this formula, we have

$$\langle d\gamma,[f,g_j]\rangle(x) = -\sum_{i=0}^{\ell}\sum_{k=0}^{\rho_i-1} c_{ik}\langle dL_f^{k+1},g_j\rangle = -\sum_{i=1}^{\ell} c_{i,\rho_i-1}a_{ij}(x)$$

But since in the (3.12) the coefficients $c_{1\rho_1},\ldots,c_{\ell\rho_\ell}$ have already been proved being equal to 0, the function $\gamma(x)$ is such that $d\gamma(x^o)=0$ and the above equality implies again the linear dependence of the rows of the matrix $A(x^o)$, i.e. a contradiction. Therefore $c_{1,\rho_1-1} = \ldots = c_{\ell,\rho_\ell-1} = 0$ (for all c_{i,ρ_i-1} defined, i.e. such that $\rho_i \ge 1$).

By repeating the procedure one completes the proof.

(3.13) *Remark.* As a consequence of this Lemma, if the matrix $A(x^o)$ has rank ℓ, the functions $L_f^k h_i(x)$, $1 \le i \le \ell$, $0 \le k \le \rho_i$ are part of a coordinate system in a neighborhood U of x^o. This fact will be extensively used in the sequel. □

The assumption on the rank of A(x) identifies a special case in which the computation of the largest controlled invariant distribution contained in H is particularly simple.

(3.14) *Corollary.* Suppose the matrix A(x) has rank ℓ at x^o. Then in a neighborhood U of x^o the distribution Δ_{sup} coincides with the largest locally controlled invariant distribution contained in H.

Proof. If A(x) has rank ℓ at x^o, in a neighborhood U' of x^o the distribution Δ_{sup} is nonsingular and therefore smooth. Moreover, in a neighborhood $U \subset U'$ of x^o the equations

(3.15a) $$A(x)\alpha(x) + b(x) = \gamma(x)$$

(3.15b) $$A(x)\beta(x) = \delta(x)$$

where $\gamma(x)$ and $\delta(x)$ are an arbitrary ℓ-vector and respectively an

---

(*)

$$-\langle dL_f^k h_i,[f,g_j]\rangle = \langle dL_f^{k+1}h_i,g_j\rangle - L_f\langle dL_f^k h_i,g_j\rangle$$

and the last term is zero because $k \le \rho_i-1$.

arbitrary $\ell \times m$ matrix, have smooth solutions. If the entries of γ and δ are such that their differentials belong to $\Delta^{\perp}_{sup}$, then the feedback (α,β) is such that (3.8) are satisfied on U. In particular this is true if the entries of α and β are constants. Note that the matrix δ must have rank ℓ in order to let β be nonsingular.

(3.16) *Remark.* Recall that any pair of feedback functions α and β which makes Δ_{sup} invariant is a solution of (3.15), provided that γ and δ have entries with differentials in $\Delta^{\perp}_{sup}$ (see Lemma (3.6)). □

The procedures outlined so far are not always usable, because $A(x)$ may fail to have rank ℓ or, more in general, Δ_{sup} may not be a locally controlled invariant distribution. In this case one may still use the general algorithm (2.4). A practical implementation of this algorithm can be obtained in the following way.

(3.17) *Algorithm* (Construction of the largest locally controlled invariant distribution contained in H).

Suppose that in a neighborhood of the point x^{o} the codistribution $sp\{dh_1,\ldots,dh_\ell\}$ has constant dimension, say s_0. Let $\lambda_0(x)$ be an s_0-vector whose entries $\lambda_{01},\ldots,\lambda_{0s_0}$ are entries of h, with the property that $d\lambda_{01},\ldots,d\lambda_{0s_0}$ are linearly independent at all x in a neighborhood of x^{o}.

The algorithm consists of a finite number of iterations, each one defined as follows.

Iteration (k). Consider the $s_k \times m$ matrix $A_k(x)$ whose (i,j)-th entry is $\langle d\lambda_{ki}(x), g_j(x)\rangle$. Suppose that in a neighborhood U_k of the point x^{o} the rank of $A_k(x)$ is constant and equal to r_k. Then it is possible to find r_k rows of $A_k(x)$ which, for all x in a neighborhood $U'_k \subset U_k$ of x^{o}, are linearly independent. Let

$$P_k = \begin{pmatrix} P_{k1} \\ P_{k2} \end{pmatrix}$$

be a $s_k \times s_k$ permutation matrix, chosen in such a way that the r_k rows of $P_{k1}A_k(x)$ are linearly independent at all $x \in U'_k$. Let $B_k(x)$ be an s_k-vector whose i-th element is $\langle d\lambda_{ki}, f\rangle(x)$. As a consequence of previous positions, the equations

$$\text{(3.18a)}\qquad P_{k1}A_k(x)\alpha(x) = -P_{k1}B_k(x)$$

$$\text{(3.18b)}\qquad P_{k1}A_k(x)\beta(x) = K$$

(where K is a matrix of real numbers, of rank r_k) may be solved for α and β, an m-vector and an $m\times m$ invertible matrix whose entries are real-valued smooth functions defined in a neighborhood U''_k of x^{o}.

Set $\tilde{g}_0 = f + g\alpha$ and $\tilde{g}_i = (g\beta)_i$, $1 \le i \le m$.

Consider the set of functions

$$\Lambda_k = \{\lambda = L_{\tilde{g}_i}\lambda_{kj} : 1 \le j \le s_k,\ 0 \le i \le m\}$$

and the codistributions

$$\Omega_{k1} = \sum_{j=1}^{s_k} sp\{d\lambda_{kj}\}$$

$$\Omega_{k2} = sp\{d\lambda : \lambda \in \Lambda_k\}$$

Suppose the codistribution $\Omega_{k1} + \Omega_{k2}$ has constant dimension, say s_{k+1}, in a neighborhood $U'''_k \subset U''_k$ of x^{o}. This integer s_{k+1} is necessarily larger than or equal to r_k because the r_k entries of $P_{k1}\lambda_k$ have linearly independent differentials at all $x \in U'_k$, otherwise $A_k(x)$ would not have rank r_k. Let $\lambda_{k+1,1},\ldots,\lambda_{k+1,s_{k+1}}$ be entries of λ_k and/or elements of Λ_k with the property that the differentials $d\lambda_{k+1,1},\ldots,d\lambda_{k+1,s_{k+1}}$ are linearly independent at all x in neighborhood $U''''_k \subset U'''_k$ of x^{o}. Thus

$$\Omega_{k1} + \Omega_{k2} = \sum_{j=1}^{s_{k+1}} sp\{d\lambda_{k+1,j}\}$$

Define the s_{k+1}-vector λ_{k+1} whose i-th entry is the function $\lambda_{k+1,i}$.

This concludes the description of the algorithm. □

As a matter of fact, it is possible to show that the operations thus described are exactly the ones required in order to compute the codistribution Ω_k from codistribution Ω_{k-1} and therefore that, under suitable assumptions, the algorithm ends at a certain stage, yielding the required distribution. Since the possibility of completing the operations defined at the k-th stage depends on assumptions on the rank of A_k and on the dimension of $\Omega_{k1} + \Omega_{k2}$, we set for convenience all these assumptions in a suitable definition. We say that x^{o} is a *regular point* for the algorithm (3.17) if, for all $k \ge 0$, the matrix

A_k has constant rank in a neighborhood of x^o and the codistribution $\Omega_{k1} + \Omega_{k2}$ has constant dimension in a neighborhood of x.

In this case r_k, the rank of A_k, and s_{k+1}, the dimension of $\Omega_{k1} + \Omega_{k2}$ are well-defined quantities in a neighborhood of x^o. Note, however, that around a regular point x^1 other than x^o, r_k and s_{k-1} might be different.

The following statement shows that the algorithm in question provides the largest locally controlled invariant distribution contained in H.

(3.19) *Proposition.* Suppose x^o is a regular point for the algorithm (3.17). Then, there exists an integer k^* with the property that $s_{k^*+1} = s_{k^*}$ and, therefore, the algorithm terminates at the (k^*)-th iteration. Suppose also G is nonsingular. Then on a suitable neighborhood U of x^o distribution

$$\Delta^* = \bigcap_{i=0}^{s_{k^*}} (\mathrm{sp}\{d\lambda_{k^*,i}\})^\perp$$

coincides with the largest locally controlled invariant distribution contained in H. The pair of feedback functions that solve (3.18) for $k = k^*$ is such that

$$[f + g\alpha, \Delta^*] \subset \Delta^*$$

$$[(g\beta)_i, \Delta^*] \subset \Delta^* \qquad 1 \leq i \leq m$$

Proof. We shall prove by induction that the assumptions of Lemma (2.7) are satisfied and that

$$\Omega_k = \sum_{j=1}^{s_k} \mathrm{sp}\{d\lambda_{kj}\}$$

This is true for $k = 0$, by definition.

Suppose it is true for some k. To compute Ω_{k+1} we need to compute first $\Omega_k \cap G^\perp$. Note that Ω_k is nonsingular around x^o because the differentials $d\lambda_{kj}$, $1 \leq j \leq s_k$, are linearly independent at all $x \in U_k^{**}$. The intersection $\Omega_k \cap G^\perp$ at x is defined as the set of all linear combinations of the form

$$\sum_{i=1}^{s_k} c_i d\lambda_{kj}(x)$$

which annihilates $g_1(x),\dots,g_m(x)$. Therefore, it is easily seen that the coefficients $c_1,\dots,c_{s_k}$ of this combination must be solutions of the equation

$$(c_1 \dots c_{s_k}) A_k(x) = 0$$

Since $A_k(x)$ has constant rank r_k in a neighborhood of x^o, $\Omega_k \cap G^\perp$ is nonsingular around x^o, has dimension $s_k - r_k$ and is spanned by covector fields which may be expressed as

$$\omega = (\gamma_1(x)P_{k2} + \gamma_2(x)P_{k1})d\lambda_k \tag{3.20}$$

$\gamma_1(x)$ being an arbitrary (s_k-r_k)-row vector of smooth functions. With $d\lambda_k$ we denote an s_k-column whose i-th entry is the covector field $d\lambda_{ki}$.

In computing Ω_{k+1}, we make also use of the fact that, if (α,β) is any feedback pair, then (see Lemma (2.8))

$$\Omega_k + \sum_{i=0}^m L_{g_i}(\Omega_k \cap G^\perp) = \Omega_k + \sum_{i=0}^m L_{\tilde g_i}(\Omega_k \cap G^\perp)$$

Now, take the Lie derivative of (3.20) along $\tilde g_i$, with α and β solutions of (3.18). As a result one obtains

$$L_{\tilde g_i}\omega = ((L_{\tilde g_i}\gamma_1)P_{k2} + (L_{\tilde g_i}\gamma_2)P_{k1})d\lambda_k + \gamma_1 P_{k2} dL_{\tilde g_i}\lambda_k + \gamma_2 P_{k1} dL_{\tilde g_i}\lambda_k$$

But the way the $\tilde g_i$ are defined is such that

$$P_{k1}L_{\tilde g_0}\lambda_k = P_{k1}\langle d\lambda_k, \tilde g_0\rangle = 0$$

$$P_{k1}L_{\tilde g_i}\lambda_k = P_{k1}\langle d\lambda_k, \tilde g_i\rangle = \text{constant}$$

for all $1 \leq i \leq m$. Thus, in the above expression we may replace γ_2 with any arbitrary r_k-row vector $\tilde\gamma_2$ of smooth functions. This makes it possible to express $L_{\tilde g_i}\omega$ in the form

$$L_{\tilde g_i}\omega = \gamma_3 d\lambda_k + \gamma_4 dL_{\tilde g_i}\lambda_k \tag{3.21}$$

where γ_3 is some s_k-row vector and γ_4 is an arbitrary s_k-row vector of smooth functions. The first term of this sum is already an element of Ω_k, by assumption, while the second, due to the arbitrariness of γ_4, spans the codistribution Ω_{k2} (see above). Thus, we may conclude that

$$\Omega_{k+1} = \Omega_k + \sum_{i=0}^{m} L_{\tilde{g}_i}(\Omega_k \cap G^{\perp}) = \Omega_k + \Omega_{k2} = \Omega_{k1} + \Omega_{k2} = \sum_{j=1}^{s_{k+1}} sp\{d\lambda_{k+1,j}\}$$

By assumption, the codistributions Ω_k, $k \geq 0$, are nonsingular around x^o (their dimension is s_k). Thus, there exists an integer k^* such that

$$\Omega_{k^*} = \Omega_{k^*+1}$$

This clearly implies the termination of the algorithm at the k^*-th step. We have also assumed that $\Omega_k \cap G^{\perp}$ are nonsingular around x^o (their dimension is r_k). So in particular $\Omega_{k^*}^{\perp} + G$ is nonsingular. If G is also nonsingular all the assumptions of Lemma (2.7) are satisfied and $\Omega_{k^*}^{\perp}$ is the required distribution.

In order to complete the proof, we have to show that the feedback pair which solves (3.18) for $k = k^*$ is such as to make $\Omega_{k^*}^{\perp}$ invariant under the new dynamics. To this end, consider again the expression (3.21) of the Lie derivative along $\tilde{g}_i$ of a covector field ω of $\Omega_k \cap G^{\perp}$. If the algorithm terminates at k^*, then

$$L_{\tilde{g}_i}(\Omega_{k^*} \cap G^{\perp}) \subset \Omega_{k^*}$$

and, therefore, we see from (3.21) that every entry of $dL_{\tilde{g}_i}\lambda_{k^*}$ (due to the arbitraryness of γ_4) is a covector field of Ω_{k^*}. But, since the entries of $d\lambda_{k^*}$ span Ω_{k^*}, this implies

$$L_{\tilde{g}_i}\Omega_{k^*} \subset \Omega_{k^*} \qquad 0 \leq i \leq m$$

and thus Ω_{k^*} is invariant under $\tilde{g}_0, \tilde{g}_1, \ldots, \tilde{g}_m$. Ω_{k^*} being nonsingular and therefore smooth, we may conclude that $\Omega_{k^*}^{\perp}$ is invariant under the new dynamics. □

This result is very important because it shows that, under suitable regularity assumptions, it is possible to find the largest locally controlled invariant distribution contained in H, and also a (locally defined) feedback pair α and β which makes it invariant under the new dynamics. The latter is particularly useful because we see that, as far as one is concerned with the maximal locally controlled invariant distribution contained in H, the computation of a such a feedback pair does not require solving partial differential equations (like in the general case, as seen from Lemma (1.10)) but may be carried out essentially solving x-dependent linear algebraic equations.

We conclude the section with some additional considerations about the properties of the algorithm (3.17). It is observed that, if the algorithm may be carried out until its final stage (i.e. if x^o is a regular point for the algorithm), as a by-product one obtains, for all $k \geq 0$, not only the dimension s_k of each codistribution Ω_k of the sequence (2.5) but also the dimension $s_k - r_k$ of the codistribution $\Omega_k \cap G^{\perp}$.

Thus the rank r_k of A_k may be interpreted as

$$r_k = \dim \frac{\Omega_k}{\Omega_k \cap G^{\perp}} \tag{3.22}$$

The integers $r_0, r_1, \ldots, r_{k^*}$ are rather important also for reasons not directly related to the construction of the distribution Δ^*. We will see in the next chapter, for instance, that the sequence of integers defined by setting

$$\begin{aligned} \delta_1 &= r_0 \\ \delta_2 &= r_1 - r_0 \\ &\ldots \\ \delta_{k^*+1} &= r_{k^*} - r_{k^*-1} \end{aligned} \tag{3.23}$$

may be, in a special case, directly evaluated starting from the coefficients of the functional expansion of the input-output behavior and plays an essential role in the problem of matching linear models.

It is also possible to relate the integers r_i, $0 \leq i \leq k^*$, to the characteristic numbers ρ_i, $1 \leq i \leq \ell$, as stated below.

(3.24) *Proposition*. Suppose that the outputs have been renumbered in such a way that the sequence of the characteristic numbers $\rho_1, \ldots, \rho_\ell$ is increasing. Let x^o be a regular point for the Algorithm (3.17). If rank $A(x^o) = \ell$, then the integers $\delta_1, \ldots, \delta_{k^*+1}$ defined by (3.23) are such that δ_i is equal to the number of outputs whose characteristic number is $(i-1)$ and, moreover, $\delta_1+\ldots+\delta_{k^*+1} = \ell$. □

The proof of this proposition is left as an exercise to the reader.

4. Noninteracting Control

Consider again a control system of the form

$$\dot{x} = f(x) + \sum_{i=1}^{m} g_i(x)u_i$$

$$y_i = h_i(x) \qquad 1 \leq i \leq \ell$$

and suppose $\ell \leq m$.

It is required to modify the system, via static state-feedback control, in order to obtain a closed loop system

$$\dot{x} = \tilde{f}(x) + \sum_{i=1}^{m} \tilde{g}_i(x)v_i$$

$$y_i = h_i(x) \qquad 1 \leq i \leq \ell$$

in which, for some suitable partition of the inputs $v_1,\ldots,v_m$ into ℓ disjoint sets, the i-th output is influenced only by the i-th set of inputs.

This control problem may easily be dealt with on the basis of the results discussed in chapter III (Theorem III.(3.12)) and its solution has interesting connections with the analysis developed so far in this chapter. In the present case, in order to ensure the independence of y_i from a set of inputs $v_{j_1},\ldots,v_{j_k}$ we have to find a distribution Δ_i which is invariant under $\tilde{f}$ and $\tilde{g}_j$, $1 \leq j \leq m$, is contained in $(sp\{dh_i\})^{\perp}$, and contains the vector fields $\tilde{g}_{j_1},\ldots,\tilde{g}_{j_k}$. Since this is required to hold for each individual output, one has to find ℓ distributions $\Delta_1,\ldots,\Delta_\ell$ all invariant with respect to the same set of vector fields $\tilde{f},\tilde{g}_1,\ldots,\tilde{g}_m$.

A set of distributions $\Delta_1,\ldots,\Delta_\ell$ with the property that

(4.1a) $$[\tilde{f},\Delta_i] \subset \Delta_i$$

(4.1b) $$[\tilde{g}_j,\Delta_i] \subset \Delta_i \qquad 1 \leq j \leq m$$

for all $1 \leq i \leq \ell$ is called a set of *compatible* controlled invariant distributions. The feedback pair which makes (4.1) satisfied is called a compatible feedback. Obviously, in the very same way one can introduce the notion of compatible local controlled invariance.

Thus, the problem we face is the following one.

(Local) single-outputs noninteracting control problem. Find a set of distributions $\Delta_1,\ldots,\Delta_\ell$ which:

(i) are compatibly (locally) controlled invariant

(iia) satisfy the conditions $\Delta_i \subset (sp\{dh_i\})^{\perp}$

(iib) for some partition $I_1 \cup I_2 \cup \ldots \cup I_\ell$ of the index set $\{1,\ldots,m\}$ and for some compatible feedback, satisfy the conditions

$$(g\beta)_j \in \Delta_i$$

for all $j \notin I_i$. □

The existence of a solution to this problem is characterized as follows

(4.2) *Theorem.* The Local Single-Outputs Noninteracting Control Problem is solvable if and only if the matrix $A(x)$ has rank ℓ for all x.

Proof. (Necessity). Suppose there exists a pair of feedback functions which solves the Single-Outputs Noninteracting Control Problem. Then, we know from the analysis of chapter III, section 3, that, in particular, for all k and all $1 \leq i \leq \ell$

$$L_{\tilde{g}_j} L_{\tilde{f}}^k h_i(x) = 0$$

whenever $j \notin I_i$. Without loss of generality we may assume the inputs $v_1,\ldots,v_m$ being renumbered in such a way that

$$I_i = \{m_{i-1} + 1,\ldots,m_i\} \qquad 1 \leq i \leq \ell$$

with $m_0 = 1$ and $m_\ell = m$. The above condition, written for $k = \rho_i$ shows that the matrix $\tilde{A}(x)$ has a block-diagonal structure: on the i-th row only the elements whose indexes belong to the set I_i are nonzero. But we have also that

$$\tilde{A}(x) = A(x)\beta(x)$$

Thus, since the matrix β is nonsingular and each row of $A(x)$ is nonzero by construction (we assumed that all ρ_i's are defined), each row of $\tilde{A}(x)$ is nonzero. $\tilde{A}(x)$ being block-diagonal, this implies that the ℓ rows of $\tilde{A}(x)$ are linearly independent and so are the ℓ rows of $A(x)$.

(Sufficiency). It is known from the analysis given in the previous section that if the i-th row of $A(x)$ is nonzero for all x, the largest locally controlled invariant distribution contained in $(\mathrm{sp}\{dh_i\})^{\perp}$ is nonsingular and given by

$$\Delta_i^* = \bigcap_{k=0}^{\rho_i} (\mathrm{sp}\{dL_f^k h_i\})^{\perp}$$

A pair of feedback functions (α,β) such that

$$[f + g\alpha, \Delta_i^*] \subset \Delta_i^* \tag{4.3a}$$

$$[(g\beta)_j, \Delta_i^*] \subset \Delta_i^* \qquad 1 \le j \le m \tag{4.3b}$$

is a solution of equations of the form

$$A_i(x)\alpha(x) + b_i(x) = \gamma_i(x) \tag{4.4a}$$

$$A_i(x)\beta(x) = \delta_i(x) \tag{4.4b}$$

where $A_i(x)$ and $b_i(x)$ denote the i-th rows of $A(x)$ and $b(x)$. The scalar $\gamma_i(x)$ and the $1\times m$ row vector $\delta_i(x)$ are functions whose differentials belong to $(\Delta_i^*)^{\perp}$: in particular, real numbers.

Considering the equations (4.4) all together, for all $1 \le i \le \ell$, one sees immediately that, thanks to the assumption on the rank of $A(x)$, there exists a pair of feedback functions (α,β) that makes (4.3) satisfied simultaneously for all Δ_i^*, i.e. that $\Delta_1^*,\ldots,\Delta_\ell^*$ are compatible locally controlled invariant distributions. In particular if the right-hand-side of (4.4b) is chosen to be the i-th row of a block diagonal matrix, one has that in the i-th row of $A(x)\beta(x)$, i.e. in the i-th row of $\tilde{A}(x)$, the only elements whose indexes belong to the set I_i are nonzero. This proves that a compatible β exists with the property that

$$L_{\tilde{g}_j} L_{\tilde{f}}^{\rho_i} h_i = 0$$

for all $j \notin I_i$. But this, in view of Lemma (3.4) is equivalent to

$$L_{\tilde{g}_j} L_f^{\rho_i} h_i = 0$$

i.e., because by definition $L_{\tilde{g}_j} L_f^k h_i = 0$ for $0 \le k < \rho_i$,

$$\tilde{g}_j \in \Delta_i^*$$

for all $j \notin I_i$.

This proves that the Local Single-Outputs Noninteracting Control Problem is solved. □

It may be interesting to look at the internal structure of the decoupled system obtained in the proof of this theorem. Suppose again that $A(x)$ has rank ℓ on some neighborhood U and let α and β be solutions of the equations (4.4) on U. One knows from Lemma (3.10) (see also Remark (3.13)) that the functions $L_f^k h_i(x)$, $1 \le i \le \ell$, $0 \le k \le \rho_i$, are part of a local coordinate system. Without loss of generality we may assume that they are coordinate functions exactly on the neighborhood U. We want to examine the special structure of the control system in the new coordinates, after the introduction of the decoupling feedback.

To this end, we set the new coordinates in the following way. Let

$$\xi_i(x) = \begin{pmatrix} z_{i0} \\ z_{i1} \\ \cdot \\ \cdot \\ \cdot \\ z_{i\rho_i} \end{pmatrix} = \begin{pmatrix} h_i(x) \\ L_f h_i(x) \\ \cdot \\ \cdot \\ \cdot \\ L_f^{\rho_i} h_i(x) \end{pmatrix}$$

for $1 \le i \le \ell$. If $\rho_1 + \ldots + \rho_\ell + \ell$ is strictly less than n, an extra set of coordinates, say $\xi_{\ell+1}$, is needed.

The computation of the form taken by the differential equations describing the system in the new coordinates is rather easy. For $1 \le i \le \ell$ and $k < \rho_i$

$$\begin{aligned} \dot{z}_{ik} &= \frac{\partial z_{ik}}{\partial x}(\tilde{f} + \sum_{j=1}^{m} \tilde{g}_j v_j) = L_{\tilde{f}} z_{ik} + \sum_{j=1}^{m} L_{\tilde{g}_j} z_{ik} v_j \\ &= L_{\tilde{f}} L_f^k h_i + \sum_{j=1}^{m} L_{\tilde{g}_j} L_f^k h_i v_j = L_f^{k+1} h_i = z_{i,k+1} \end{aligned} \tag{4.5}$$

Whereas, for $k = \rho_i$ (Lemma (3.4))

$$\dot z_{i\rho_i} = \gamma_i(x) + \sum_{j=1}^{m} \delta_{ij}(x) v_j$$

where $\gamma_i(x)$ is the right-hand-side of (4.4a) and $\delta_{ij}(x)$ is the j-th element of the right-hand-side of (4.4b). If this latter is chosen to be as the i-th row of a block-diagonal matrix, as in the proof of Theorem (4.2), then the above equation reduces to

(4.6) $$\dot z_{i\rho_i} = \gamma_i + \sum_{j \in I_i} \delta_{ij} v_j$$

Again from the proof of Theorem (4.2), it is seen that γ_i and δ_{ij} depend only on $z_{i0}, \dots, z_{i\rho_i}$ (*). As a matter of fact, γ_i and δ_{ij} may be simply real numbers.

Finally, by definition, for all $1 \le i \le \ell$

(4.7) $$y_i = z_{i0}$$

As a result, we see that in the new coordinates the closed loop system may be described in the form

(4.8)
$$\dot\xi_i = \bar f_i(\xi_i) + \sum_{j \in I_i} \bar g_{ij}(\xi_i) v_j \qquad 1 \le i \le \ell$$
$$\dot\xi_{\ell+1} = \bar f_{\ell+1}(\xi_1, \dots, \xi_{\ell+1}) + \sum_{j=1}^{m} \bar g_{\ell+1,j}(\xi_1, \dots, \xi_{\ell+1}) v_j$$
$$y_i = \bar h_i(\xi_i)$$

with

$$\bar f_i(\xi_i) = \begin{pmatrix} z_{i1} \\ \cdot \\ \cdot \\ \cdot \\ z_{i\rho_i} \\ \gamma_i(\xi_i) \end{pmatrix} \qquad \bar g_{ij}(\xi_i) = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ \delta_{ij}(\xi_i) \end{pmatrix}$$

$$\bar h_i(\xi_i) = z_{i0}$$

(*) Let $\gamma_i(z) = \gamma_i \circ x(z)$. Then

$$\frac{\partial \gamma_i}{\partial z_{jk}} = d\gamma_i \frac{\partial x}{\partial z_{jk}} = \sum_{s=0}^{\rho_i} c_{is} \frac{\partial z_{is}}{\partial x} \frac{\partial x}{\partial z_{jk}} = 0$$

because $i \neq j$.

These equations clearly stress the decoupled structure of the closed loop system.

(4.9) *Remark*. The choice of $\gamma_i(\xi_i)$ linear in ξ_i, i.e. the choice

$$\gamma_i(x) = a_{i0} h_i(x) + a_{i1} L_f h_i(x) + \dots + a_{i\rho_i} L_f^{\rho_i} h_i(x)$$

with $a_{i0}, \dots, a_{i\rho_i}$ real numbers, is admissible, because $d\gamma_i(x)$ in this case belongs to $sp\{dh_i, \dots, dL_f^{\rho_i} h_i\}$. It is also possible to choose δ_{ij} constant, provided that, for some $j \in I_i$, δ_{ij} is nonzero because this is required for the solution β of the (4.4b) be nonsingular. The two facts show that a suitable choice of decoupling feedback makes linear the first ℓ subsystems of (4.8).

(4.10) *Remark*. Note that Δ_i^*, the largest locally controlled invariant distribution contained in $(sp\{dh_i\})^\perp$, in the coordinates is expressed as

$$\Delta_i^* = sp\{\frac{\partial}{\partial z_{jk}} : j \neq i,\ 0 \le k \le \rho_j\} + sp\{\frac{\partial}{\partial z_{\ell+1,k}} : 1 \le k \le d\}$$

where d denotes the dimension of $\xi_{\ell+1}$ (see chapter I, section 3). □

At the beginning of this section, we have formulated the Noninteracting Control Problem looking at the existence of a set of compatible controlled invariant distributions, each one contained in $(sp\{dh_i\})^\perp$ and containing the vector fields $\tilde g_j = (g\beta)_j$ for all $j \notin I_i$. One can also consider a complementary formulation in the following terms.

Local single-outputs noninteracting control problem. Find a set of distributions $\Delta_1, \dots, \Delta_\ell$ which

(i) are compatibly locally controlled invariant

(iia) satisfy the conditions $\Delta_i \subset (sp\{dh_j\})^\perp$ for all $j \neq i$

(iib) for some partition $I_1 \cup I_2 \cup \dots \cup I_\ell$ of the index set $\{1, \dots, m\}$ and for some compatible feedback, satisfy the conditions

$$(g\beta)_j \in \Delta_i$$

for all $j \in I_i$. □

Also in this case, in fact, the output y_i of the closed-loop system will be affected only by the inputs whose index belongs to the set I_i.

Clearly the condition that the rank of $A(x)$ is equal to ℓ remains necessary and sufficient for the existence of a solution to the problem. If desired, one could directly prove the sufficiency in terms of the complementary formulation discussed above. As in Theorem (4.2), it is easy to prove that the assumption on $A(x)$ makes it possible to express the largest locally controlled invariant distribution contained in $\bigcap_{j\neq i}(\mathrm{sp}\{dh_j\})^{\perp}$ as

$$K_i^* = \bigcap_{j\neq i}\ \bigcap_{k=0}^{\rho_j}(\mathrm{sp}\{dL_f^k h_j\})^{\perp}$$

The distributions $K_1^*,\ldots,K_\ell^*$ are compatible and a compatible feedback is exactly the one that makes $\Delta_1^*,\ldots,\Delta_\ell^*$ compatible.

(4.11) *Remark.* Note that in the new coordinate system

$$K_i^* = \mathrm{sp}\{\frac{\partial}{\partial z_{ik}} : 0 \le k \le \rho_i\} + \mathrm{sp}\{\frac{\partial}{\partial z_{\ell+1,k}} : 1 \le k \le d\}$$

(4.12) *Remark.* Summarizing some of the above results, one may observe that if $A(x)$ has rank ℓ, there is a set of distributions $D_1,\ldots,D_{\ell+1}$, namely

$$D_i = \mathrm{sp}\{\frac{\partial}{\partial z_{ik}} : 0 \le k \le \rho_i\} \qquad 1 \le i \le \ell$$

$$D_{\ell+1} = \mathrm{sp}\{\frac{\partial}{\partial z_{\ell+1,k}} : 1 \le k \le d\}$$

which are independent, i.e. such that

$$D_i \cap (\sum_{j\neq i} D_j) = 0$$

and span the tangent space, i.e. are such that

$$D_1 + D_2 + \ldots + D_{\ell+1} = TM$$

Moreover,

$$\Delta_i^* = \bigoplus_{j\neq i} D_j$$

$$K_i^* = D_i + D_{\ell+1}$$

$$(5.22)\qquad y(t) = Q(t,(\xi^o,x^o)) + \int_0^t W_M(t-\tau)v(\tau)d\tau$$

Let y^* be any smooth ℓ-vector-valued function, defined on $\mathbb{R}$, such that

$$(5.23)\qquad \left(\frac{d^k y^*(t)}{dt^k}\right)_0 = \left(\frac{d^k Q(t,(\xi^o,x^o))}{dt^k}\right)_0$$

for $0 \le k \le \delta-1$.

From the Remark (5.21) we easily deduce that there exists an input v under which the right-hand-side of (5.22) becomes exactly y^*. Thus, the composition of (1.1) with the dynamic state-feedback compensator which solves the problem of matching the transfer function (5.20) is a system that, in the initial state (ξ^o,x^o), *can reproduce any output function* which satisfies the conditions (5.23).

Moreover, we note that in a linear system with transfer function (5.20) each output component is influenced only by the corresponding component of the input. Thus, we also see that if the condition $r_{k^*} = \ell$ holds, we can *achieve non-interaction via dynamic state-feedback*.

6. State-space linearization

In the first section of this chapter, we examined the problem of achieving, via feedback, a linear *input-output* response. The subsequent analysis developed in the second section showed that, from the point of view of a state-space description, in suitable local coordinates, the system thus linearized assumes (at least in the special case where $r_{q-1} = \ell$) the form

$$\dot z = Fz + Gv$$
$$\dot\eta = f(z,\eta) + g(z,\eta)v$$
$$y = Hz$$

In other words, the input-output-wise linear system one obtains by means of the techniques in question may be interpreted, at a state-space level, as the interconnection of a (possibly) nonlinear unobservable subsystem with a system that, in suitable local coordinates, is state-space-wise linear. Moreover, the latter subsystem was also shown being both reachable and observable (Remark (2.21)).

In other words again, we may say that the techniques developed at the beginning of this chapter modify the behavior of the original system in a way such as to make a part of it (i.e. the observable one) locally diffeomorphic to a reachable linear system.

Motivated by these considerations, we want to examine now the problem of modifying, via feedback, a given nonlinear system in a way such that not simply a part, but the whole of it, is locally diffeomorphic to a reachable linear system. In formal terms, the problem thus introduced may be characterized as follows.

State-Space Linearization Problem. Given a collection of vector fields $f,g_1,\ldots,g_m$ and an initial state x^o, find (if possible) a neighborhood U of x^o, a pair of feedback functions α and β (with invertible β) defined on U, a coordinates transformation $z = F(x)$ defined on U, a matrix $A \in \mathbb{R}^{n\times n}$ and a set of vectors $b_1 \in \mathbb{R}^n,\ldots,b_m \in \mathbb{R}^n$ such that

$$(6.1)\qquad F_*(f+g\alpha)\circ F^{-1}(z) = Az$$

$$(6.2)\qquad F_*(g\beta)_i\circ F^{-1}(z) = b_i \qquad 1 \le i \le m$$

for all $z \in F(U)$, and

$$(6.3)\qquad \sum_{k=0}^{n-1}\sum_{i=1}^{m} \mathrm{Im}(A^k b_i) = \mathbb{R}^n$$

(6.4) *Remark.* Let $x(t)$ denote a state trajectory of the system

$$\dot x = (f+g\alpha)(x) + \sum_{i=1}^{m}(g\beta_i)(x)u_i$$

and suppose $x(t) \in U$ for all $t \in [0,T]$ for some $T > 0$. If (6.1) and (6.2) hold, then for all $t \in [0,T]$

$$z(t) = F(x(t))$$

is a state trajectory of the linear system

$$\dot z = Az + \sum_{i=1}^{m} b_i u_i$$

Moreover, if (6.3) also holds, the latter is a reachable linear system. □

We shall describe first the solution of this problem in the special case of a system with a single input, which is rather easy. Then, we make some remarks about the usefulness of this linearization technique in problems of asymptotic stabilization. Finally, we conclude the section with the analysis of the (general) multi-input systems.

For the sake of simplicity, we state some intermediate results which may have their own independent interest.

(6.5) *Lemma.* Suppose $m = 1$ and let $g = g_1$. The State-Space Linearization Problem is solvable if and only if there exists a neighborhood V of x^o and a function $\varphi : V \to \mathbb{R}$ such that

$$L_g\varphi(x) = L_gL_f\varphi(x) = \dots = L_gL_f^{n-2}\varphi(x) = 0 \tag{6.6}$$

for all $x \in V$, and

$$L_gL_f^{n-1}\varphi(x^o) \neq 0 \tag{6.7}$$

Proof. Necessity. Let (A,b) a reachable pair. Then, it is well known from the theory of linear system that there exist a nonsingular $n\times n$ matrix T and a $1\times n$ row vector k such that

$$T(A+bk)T^{-1} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \qquad Tb = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ 0 \\ 1 \end{pmatrix} \tag{6.8}$$

Suppose (6.1) and (6.2) hold, and set

$$\hat{z} = \hat{F}(x) = TF(x)$$

$$\hat{\alpha}(x) = \alpha(x) + \beta(x)kF(x)$$

$$\hat{\beta}(x) = \beta(x)$$

Then, it is easily seen that

$$\hat{F}_*(g\hat{\beta})\circ\hat{F}^{-1}(\hat{z}) = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ 0 \\ 1 \end{pmatrix}$$

$$\hat{F}_*(f+g\hat{\alpha})\circ\hat{F}^{-1}(\hat{z}) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \hat{z}$$

From this, we deduce that there is no loss of generality in assuming that the pair A,b which makes (6.1) and (6.2) satisfied has directly the form specified in the right-hand-sides of (6.8).

Now, set

$$z = F(x) = \mathrm{col}(z_1(x),\dots,z_n(x))$$

If (6.1) holds (with A and b in the form of the right-hand-sides of (6.8)), we have for all $x \in U$,

$$F_*(f(x)+g(x)\alpha(x)) = AF(x)$$

that is

$$\frac{\partial z_1}{\partial x}(f(x)+g(x)\alpha(x)) = z_2(x)$$

$$\dots$$

$$\frac{\partial z_{n-1}}{\partial x}(f(x)+g(x)\alpha(x)) = z_n(x)$$

$$\frac{\partial z_n}{\partial x}(f(x)+g(x)\alpha(x)) = 0$$

If also (6.2) holds we have

$$F_*g(x)\beta(x) = b$$

that is

$$\frac{\partial z_1}{\partial x}\, g(x)\beta(x) = 0$$

$$\dots$$

$$\frac{\partial z_{n-1}}{\partial x}\, g(x)\beta(x) = 0$$

$$\frac{\partial z_n}{\partial x}\, g(x)\beta(x) = 1$$

Since $\beta(x)$ is nonzero for all $x \in U$, the second set of conditions imply

$$\frac{\partial z_i}{\partial x} g(x) = L_g z_i(x) = 0 \qquad 1 \le i \le n-1 \tag{6.9}$$

$$\frac{\partial z_n}{\partial x} g(x) = L_g z_n(x) = \frac{1}{\beta(x)} \tag{6.10}$$

for all $x \in U$. These, in turn, together with the first set of conditions imply

$$L_f z_i(x) = z_{i+1}(x) \qquad 1 \le i \le n-1 \tag{6.11}$$

$$L_f z_n(x) = -\frac{\alpha(x)}{\beta(x)} \tag{6.12}$$

for all $x \in U$.

If one sets

$$\varphi(x) = z_1(x) \tag{6.13a}$$

the conditions (6.11) yield

$$z_{i+1}(x) = L_f^i \varphi(x) \qquad 0 \le i \le n-1 \tag{6.13b}$$

Thus, from (6.9) one obtains

$$L_g\varphi(x) = L_g L_f \varphi(x) = \ldots = L_g L_f^{n-2}\varphi(x) = 0$$

for all $x \in U$ and, from (6.10),

$$L_g L_f^{n-1}\varphi(x) \neq 0$$

for all $x \in U$. This completes the proof of the necessity.

Sufficiency. Suppose (6.6) and (6.7) are true and let $U \subset V$ be a neighborhood of x^o such that $L_g L_f^{n-1}\varphi(x) \neq 0$ for all $x \in U$. Use (6.13) in order to define a set of functions $z_1,\ldots,z_n$ on U. The functions thus defined are clearly such that (6.9) and (6.11) hold. Moreover, since $L_g z_n(x)$ is nonzero on U, one can define a nonzero function $\beta(x)$ and a function $\alpha(x)$ by means of (6.10) and (6.12). This pair of functions α and β and the mapping

$$F : x \longmapsto \mathrm{col}(z_1(x),\ldots,z_n(x))$$

are clearly such that

$$F_*(f(x) + g(x)\alpha(x)) = AF(z)$$

$$F_*(f(x)\beta(x)) = b$$

with A and b in the form of the right-hand-sides of (6.8). Thus, in order to complete the proof, we only have to show that F qualifies as a local coordinates transformation around x^o, i.e. that its differential F_* is nonsingular at x^o.

For, observe that the vector fields $\tilde f = f + g\alpha$ and $\tilde g = g\beta$ are such that

$$F_*\tilde f(x) = AF(x)$$

$$F_*\tilde g(x) = b$$

or, in other words, that $\tilde f$ is F-related to the vector field f' defined by

$$f'(z) = Az$$

and that $\tilde g$ is F-related to the vector field g' defined by

$$g'(z) = b$$

As a consequence, we have that the Lie bracket $[\tilde f, \tilde g]$ is F-related to the Lie bracket $[f', g']$. Using this fact repeatedly, one may check that

$$F_*(ad^i_{\tilde f}\tilde g)(x) = (ad^i_{f'} g') \circ F(x)$$

for all $0 \le i \le n-1$. The special form of f' and g' is such that

$$(ad^i_{f'} g') = (-1)^i A^i b$$

All together, these yield

$$F_*(g \quad ad_{\tilde f}\tilde g \ \ldots \ ad^{n-1}_{\tilde f}\tilde g) =$$

$$= (b \quad -Ab \ \ldots \ (-1)^{n-1}A^{n-1}\, b)$$

The matrix on the right-hand-side is nonsingular, because (A,b) is a reachable pair, and so is Γ_*. This completes the proof of the sufficiency. □

(6.14) *Lemma*. Let φ be a real-valued function defined on an open set V. Then the conditions (6.6) and (6.7) hold if and only if

(6.6') $$L_g\varphi(x) = L_{[f,g]}\varphi(x) = \dots = L_{(ad_f^{n-2}g)}\varphi(x) = 0$$

for all $x \in V$, and

(6.7') $$L_{(ad_f^{n-1}g)}\varphi(x^o) \neq 0$$

Proof. We show, by induction, that the set of conditions

(6.15a) $$L_gL_f^0\varphi = \dots = L_gL_f^k\varphi = 0$$

is equivalent to the set of conditions

(6.15b) $$L_{(ad_f^0g)}\varphi = \dots = L_{(ad_f^kg)}\varphi = 0$$

and both imply

(6.16) $$L_{(ad_f^ig)}L_f^j\varphi = (-1)^iL_gL_f^{i+j}\varphi$$

for all i,j such that $i+j = k+1$.

This is clearly true for $k = 0$. In this case (6.15a) and (6.15b) reduces to $L_g\varphi = 0$ and

$$L_{[f,g]}\varphi = L_fL_g\varphi - L_gL_f\varphi = -L_gL_f\varphi$$

Suppose (6.15a) and (6.15b) true for some k and (6.16) true for all i,j such that $i+j = k+1$. The latter yields, in particular,

$$L_{(ad_f^{k+1}g)}\varphi = (-1)^{k+1}L_gL_f^{k+1}\varphi$$

so that $L_gL_f^{k+1}\varphi = 0$ if and only if $L_{(ad_f^{k+1}g)}\varphi = 0$. Assume either one of these conditions holds. Then

$$L_{(ad_f^{k+2}g)}\varphi = L_fL_{(ad_f^{k+1}g)}\varphi - L_{(ad_f^{k+1}g)}L_f\varphi =$$

$$= (-1)L_fL_{(ad_f^kg)}L_f\varphi + (-1)^2L_{(ad_f^kg)}L_f^2\varphi$$

$$= (-1)^{k+1}L_fL_gL_f^{k+1}\varphi + (-1)^2L_{(ad_f^kg)}L_f^2\varphi$$

$$= (-1)^2L_{(ad_f^kg)}L_f^2\varphi$$

$$= (-1)^2L_fL_{(ad_f^{k-1}g)}L_f^2\varphi + (-1)^3L_{(ad_f^{k-1}g)}L_f^3\varphi$$

$$= (-1)^{k+1}L_fL_gL_f^{k+1}\varphi + (-1)^3L_{(ad_f^{k+1}g)}L_f^3\varphi$$

$$= (-1)^3L_{(ad_f^{k-1}g)}L_f^3\varphi = \dots$$

We see in this way that for all $0 \le j \le k+2$

$$L_{(ad_f^{k+2}g)}\varphi = (-1)^jL_{(ad_f^{k+2-j}g)}L_f^j\varphi = (-1)^{k+2}L_gL_f^{k+2}\varphi$$

and therefore that (6.16) is true for all i,j such that $i+j = k+2$.

From (6.15) and (6.16) the statement follows immediately. □

(6.17) *Remark*. We have proved, by the way, that either one of the two equivalent sets of conditions (6.15) imply

$$L_{(ad_f^ig)}L_f^j\varphi = (-1)^jL_{(ad_f^{i+j}g)}\varphi$$

for all i,j such that $i+j \le k+1$. This fact will be used in the sequel. □

(6.18) *Theorem*. Suppose $m = 1$ and let $g = g_1$. The State-Space Linearization Problem is solvable if and only if:

(i) $\dim(\mathrm{span}\{g(x^o), ad_fg(x^o), \dots, ad_f^{n-1}g(x^o)\}) = n$

(ii) the distribution

(6.19) $$\Delta = \mathrm{sp}\{g, ad_fg, \dots, ad_f^{n-2}g\}$$

is involutive in a neighborhood U of x^o.

(6.20) *Remark.* Note that the condition (i) implies that the tangent vectors $g(x), ad_f g(x), \dots, ad_f^{n-1} g(x)$ are linearly independent for all x in a suitable neighborhood of x^o. Therefore the distribution (6.19) is nonsingular around x^o and has dimension $(n-1)$.

Proof. We know from the previous Lemmas that the problem is solvable if and only if there exists a real-valued function φ defined in a neighborhood V of x^o such that the conditions (6.6') and (6.7') hold. These may be rewritten as

$$\langle d\varphi, ad_f^i g\rangle(x) = 0 \qquad (6.6")$$

for all $0 \le i \le n-2$ and all $x \in V$, and

$$\langle d\varphi, ad_f^{n-1} g\rangle(x^o) \ne 0 \qquad (6.7")$$

If both these conditions hold, then necessarily the tangent vectors $g(x^o), ad_f g(x^o), \dots, ad_f^{n-1} g(x^o)$ are linearly independent. For, we see from Remark (6.17) that (6.6") implies

$$\langle dL_f^j\varphi, ad_f^i g\rangle = L_{(ad_f^i g)} L_f^j \varphi =$$

$$= (-1)^j L_{(ad_f^{i+j} g)} \varphi = (-1)^j \langle d\varphi, ad_f^{i+j} g\rangle$$

for all $i+j \le n-1$. Therefore, using again (6.6") and (6.7") we have

$$\langle dL_f^j\varphi, ad_f^i g\rangle(x) = 0$$

for all i,j such that $i+j \le n-2$ and all $x \in V$ and

$$\langle dL_f^j\varphi, ad_f^i g\rangle(x^o) \ne 0$$

for all i,j such that $i+j = n-1$.

The above conditions, all together, show that the matrix

$$\begin{pmatrix} d\varphi(x^o) \\ dL_f\varphi(x^o) \\ \vdots \\ dL_f^{n-1}\varphi(x^o) \end{pmatrix} (g(x^o) \;\; ad_f g(x^o) \dots ad_f^{n-1} g(x^o)) = \qquad (6.21)$$

$$= \begin{pmatrix} \langle d\varphi, g\rangle(x^o) & \langle d\varphi, ad_f g\rangle(x^o) & \dots & \langle d\varphi, ad_f^{n-1} g\rangle(x^o) \\ \langle dL_f\varphi, g\rangle(x^o) & \langle dL_f\varphi, ad_f g\rangle(x^o) & \dots & \langle dL_f\varphi, ad_f^{n-1} g\rangle(x^o) \\ \cdot & \cdot & \dots & \\ \langle dL_f^{n-1}\varphi, g\rangle(x^o) & \langle dL_f^{n-1}\varphi, ad_f g\rangle(x^o) & \dots & \langle dL_f^{n-1}\varphi, ad_f^{n-1} g\rangle(x^o) \end{pmatrix}$$

has rank n and, therefore, that the vectors $g(x^o), ad_f g(x^o), \dots, ad_f^{n-1} g(x^o)$ are linearly independent.

This proves the necessity of (i). If (i) holds then the distribution (6.19) has dimension $n-1$ around x^o and (6.6") tell us that the exact covector field $d\varphi$ spans $\Delta^\perp$ around x^o. So, because of Frobenius theorem (see Remark I.(3.7)) we conclude that Δ is completely integrable and thus involutive, i.e. the necessity of (ii).

Conversely, suppose (i) holds. Then the distribution (6.19) is nonsingular around x^o. If also (ii) holds, Δ is completely integrable around x^o and there exists a real-valued function φ, defined in a neighborhood V of x^o, such that $d\varphi$ spans $\Delta^\perp$ on V, i.e. such that (6.6") are satisfied. Moreover, the covector field $d\varphi$ is such that (6.7") also is satisfied, because otherwise $d\varphi$ would be annihilated by a set of n linearly independent vectors. This, in view of the previous Lemmas, completes the proof of the sufficiency. □

For the sake of convenience, we summarize now the *procedure* leading to the construction of the feedback α and β which solves the State-Space Linearization Problem in the case of a single input channel.

Suppose (i) and (ii) hold. Then, using Frobenius Theorem one constructs a function φ, defined in a neighborhood V of x^o, such that (6.6") and (6.7") hold. Then, one sets

$$\beta(x) = \frac{1}{L_g L_f^{n-1}\varphi(x)} \qquad (6.22a)$$

and

$$\alpha(x) = \frac{-L_f^n\varphi(x)}{L_g L_f^{n-1}\varphi(x)} \qquad (6.22b)$$

for all $x \in V$. This pair of feedback functions, together with the local coordinates transformation defined by

$$z_i(x) = L_f^{i-1}\varphi(x)$$

for $1 \le i \le n$, is such as to satisfy (6.1) and (6.2) with A and $b_1 = b$ in the form of the right-hand-sides of (6.8).

(6.23) *Remark*. There is a surprising affinity between some results described in this section and the ones described in the sections IV.3 and IV.4. For instance one may rephrase Lemma (6.5) by saying that the State-Space Linearization Problem is solvable if and only if one may define, for the system

$$\dot{x} = f(x) + g(x)u$$

a (dummy) output function

$$y = \varphi(x)$$

whose characteristic number is *exactly* n-1. Of course, this will be possible if and only if the conditions (i) and (ii) are satisfied.

Once such a dummy output function has been found, then the solution of a State-Space Linearization Problem proceeds like a solution of a (degenerate, because both ℓ and m are equal to 1) noninteracting control problem. As a matter of fact, we have from Lemma IV.(3.10) that the differentials $d\varphi, dL_f\varphi, \ldots, dL_f^{n-1}\varphi$ are linearly independent at x^o and thus that the mapping

$$F : x \longmapsto \text{col}(\varphi(x), L_f\varphi(x), \ldots, L_f^{n-1}\varphi(x))$$

qualifies as a local coordinates transformation. Then, from Corollary IV.(3.14) we learn that

$$\Delta^* = \bigcap_{i=0}^{n-1} (\text{sp}\{dL_f^i\varphi\})^\perp = 0$$

is the largest locally controlled invariant distribution contained in $(\text{sp}\{d\varphi\})^\perp$.

The feedback (6.22) coincides with a solution of IV.(3.15) (with $\gamma(x) = 0$ and $\delta(x) = 1$). Under this feedback the system becomes linear in the new coordinates, as it is seen from the constructions given in the section IV.4 (see Remark IV.(4.9)). □

We note that the formal statement of the State-Space Linearization Problem, given at the beginning of the section, does not incorporate any requirement about the image F(U) of the coordinates transformation that makes it possible (6.1) and (6.2) to hold. However, one may wish to impose the additional requirement that the image F(U) contains the origin of $\mathbb{R}^n$. In this case, the condition of Theorem (6.18) must be strenghtened a little bit.

Suppose the coordinates transformation $z = F(x)$ solving the State-Space Linearization Problem is such that

(6.24) $$z^o = F(x^o) = 0$$

Then, from (6.1) we deduce that necessarily

(6.25) $$f(x^o) + g(x^o)\alpha(x^o) = 0$$

If $f(x^o) = 0$, then the construction already proposed for the solution of the problem may be adapted to make (6.24) and (6.25) satisfied. As a matter of fact, one may always choose a function φ satisfying (6.6") and (6.7") in such a way that $\varphi(x^o) = 0$ (see, e.g., the construction proposed along the proof of Theorem I.(3.3)). If this is the case, then

$$z_1(x^o) = \varphi(x^o) = 0$$

and also, for $2 \le i \le n$,

$$z_i(x^o) = L_f^{i-1}\varphi(x^o) = \langle dL_f^{i-2}\varphi(x^o), f(x^o)\rangle = 0$$

because we have assumed $f(x^o) = 0$. Thus the proposed coordinates transformation satisfies (6.24). Moreover,

$$\alpha(x^o) = -\frac{L_f^n\varphi(x^o)}{L_gL_f^{n-1}\varphi(x^o)} = -\frac{\langle dL_f^{n-1}\varphi(x^o), f(x^o)\rangle}{L_gL_f^{n-1}\varphi(x^o)} = 0$$

and also (6.25) holds.

One may thus assert that if $f(x^o) = 0$, i.e. if the initial state x^o is an equilibrium state for the autonomous system

$$\dot{x} = f(x)$$

and if the State-Space Linearization Problem is solvable, one may always find a solution such that $\alpha(x^o) = 0$ and $F(x^o) = 0$.

If $f(x^o) \neq 0$, the condition (6.25) may be rewritten as

$$f(x^o) = cg(x^o) \tag{6.26}$$

where c is a nonzero real number. Again, if the State-Space Linearization Problem is solvable one may find a function φ such that $z_1(x^o) = \varphi(x^o) = 0$. But also, for $2 \leq i \leq n$, (6.26) ensures that

$$z_i(x^o) = \langle dL_f^{i-2}\varphi, f(x^o)\rangle = cL_gL_f^{i-2}\varphi(x^o) = 0$$

and thus (6.24) still holds. Moreover, the proposed α is such that

$$\alpha(x^o) = -\frac{\langle dL_f^{n-1}\varphi(x^o), f(x^o)\rangle}{L_gL_f^{n-1}\varphi(x^o)} = -c$$

as expected.

In this case, the initial state x^o is not an equilibrium state for the original system, but an α may be found such that x^o is an equilibrium state for the system

$$\dot{x} = f(x) + g(x)\alpha(x)$$

In summary, we have the following result.

(6.27) *Corollary*. Suppose m = 1 and let $g = g_1$. Suppose the State-Space Linearization Problem is solvable. Then, a solution with $F(x^o) = 0$ exists if and only if

$$f(x^o) \in \mathrm{sp}\{g(x^o)\} \qquad \square$$

(6.28) *Remark*. When $F(x^o) = 0$, one may use the solution of the State-Space Linearization Problem for local stabilization purposes. Indeed, since (A,b) is a reachable pair, one may arbitrarily assign the eigenvalues to the matrix (A + bk), via suitable choice of the 1×n row vector k. If this is the case, the feedback control law

$$u = \alpha(x) + \beta(x)kF(x) + \beta(x)v$$

makes the system locally diffeomorphic, on U, to the asymptotically stable system

$$\dot{z} = (A + bk)z + bv \qquad \square$$

We now describe the extension of the previous discussion to the case of many inputs. This requires the introduction of some further notations, but the substance of the procedure is essentially the same as the one examined so far.

Given a set of vector fields $f, g_1, \ldots, g_m$ we define a sequence of distributions as follows

$$G_0 = \mathrm{sp}\{g_1, \ldots, g_m\}$$

$$G_i = G_{i-1} + [f, G_{i-1}]$$

The following Lemma describes the possibility of computing all G_i's in a simple way.

(6.29) *Lemma*. Suppose all G_i's are nonsingular. Then

$$G_i = \mathrm{sp}\{ad_f^k g_j : 0 \leq k \leq i,\ 1 \leq j \leq m\} \tag{6.30}$$

Proof. Suppose $G_i = \mathrm{sp}\{ad_f^k g_j : 0 \leq k \leq i,\ 1 \leq j \leq m\}$. Suppose that at x^o some vectors $ad_f^{k_1} g_{j_1}, \ldots, ad_f^{k_r} g_{j_r}$ are linearly independent and span $G_i(x^o)$. Then on a neighborhood U of x^o any vector field τ in G_i may be written as $\tau = \sum_{\alpha=1}^{r} c_\alpha ad_f^{k_\alpha} g_{j_\alpha}$, with $c_\alpha \in C^\infty(U)$. Then $[f,\tau] = \sum_{\alpha=1}^{r} (c_\alpha ad_f^{k_\alpha+1} g_{j_\alpha} + (L_f c_\alpha) ad_f^{k_\alpha} g_{j_\alpha})$. Therefore, on U, $G_{i+1} = \mathrm{sp}\{ad_f^{k_\alpha+1} g_{j_\alpha}, ad_f^{k_\alpha} g_{j_\alpha} : 1 \leq \alpha \leq r\}$. Since, by construction, all $ad_f^k g_j$, $0 \leq k \leq i+1$ and $1 \leq j \leq m$ are in G_{i+1}, this proves that $G_{i+1} = \mathrm{sp}\{ad_f^k g_j : 0 \leq k \leq i+1,\ 1 \leq j \leq m\}$. $\square$

Since $G_i \subset G_{i+1}$ by definition, if the G_i's are nonsingular we have that

$$\dim \frac{G_{i+1}(x)}{G_i(x)} = \text{independent of } x$$

Thus we may define a sequence of integers $\nu_0, \nu_1, \ldots$ by setting

$$\nu_0 = \dim G_0 \tag{6.31a}$$

$$\nu_i = \dim \frac{G_i}{G_{i-1}} \qquad i \geq 1 \tag{6.31b}$$

The integers thus defined have the following property

(6.32) *Lemma.* The following condition holds

$$\nu_i \geq \nu_{i+1}$$

for all $i \geq 0$. Let ν_{i^*} denote the last nonzero element in the sequence $\{\nu_i : i \geq 0\}$. If

$$\dim G_{i^*} = n$$

then

$$\nu_0 + \nu_1 + \ldots + \nu_{i^*} = n$$

Proof. Consider G_i and G_{i-1}. By definition

$$\dim G_i(x) = \dim G_{i-1}(x) + \nu_i$$

From (6.30), we deduce that, given a point x^o, there will be ν_i vectors $ad_f^i g_{j_1}(x^o), \ldots, ad_f^i g_{j_{\nu_i}}(x^o)$ linearly independent and with the property that all vector fields in G_i may be written as linear combinations, with smooth coefficients, of vectors of G_{i-1} and of $sp\{ad_f^i g_{j_s} : 1 \leq s \leq \nu_i\}$. Thus

$$G_{i+1} = G_i + sp\{ad_f^{i+1} g_{j_s} : 1 \leq s \leq \nu_i\}$$

and

$$\nu_{i+1} \leq \nu_i \, . \quad \square$$

From the sequence $\{\nu_i : 0 \leq i \leq i^*\}$ we define another sequence of integers $m_0, m_1, \ldots, m_{i^*}$, setting

$$m_0 = \nu_{i^*}$$

$$m_0 + m_1 = \nu_{i^*-1}$$

(6.33) $$m_0 + m_1 + m_2 = \nu_{i^*-2}$$

$$\ldots$$

$$m_0 + m_1 + \ldots + m_{i^*} = \nu_0$$

(6.34) *Lemma.* The following conditions hold

$$m_0 > 0$$

$$m_i \geq 0 \qquad 1 \leq i \leq i^*$$

Moreover, if $\dim G_{i^*} = n$, then

(6.35) $$\dim G^{\perp}_{i^*-i} = im_0 + \ldots + 2m_{i-2} + m_{i-1}$$

for $1 \leq i \leq i^*$. $\square$

There is a need for a third sequence of integers $\{\kappa_i : 1 \leq i \leq \nu_0\}$ related to the previous ones by the following relations

$$\kappa_i = i^*+1 \quad \text{if} \quad 1 \leq i \leq \nu_{i^*}$$

$$\kappa_i = i^* \quad \text{if } m_1 > 0 \text{ and } \nu_{i^*}+1 \leq i \leq \nu_{i^*-1}$$

(6.36) $$\kappa_i = i^*-1 \quad \text{if } m_2 > 0 \text{ and } \nu_{i^*-1}+1 \leq i \leq \nu_{i^*-2}$$

$$\ldots \qquad \ldots$$

$$\kappa_i = 1 \quad \text{if } m_{i^*} > 0 \text{ and } \nu_1+1 \leq i \leq \nu_0$$

With the help of these notations it is rather simple to state the necessary and sufficient conditions for the existence of a solution to the State-Space Linearization Problem in the general case where $m \geq 1$.

(6.37) *Theorem.* The State-Space Linearization Problem is solvable if and only if

(i) x^o is a regular point of the distribution G_i, for all $i \geq 0$

(ii) $\dim G_{i^*}(x^o) = n$

(iii) the distribution G_i is involutive, for all i such that $m_{i^*-i-1} \neq 0$.

Proof. We restrict ourselves to the proof of the sufficiency, which is constructive. Without loss of generality, we may assume that

$$\nu_0 = m$$

For, if this is not the case, since G_0 by assumption is nonsingular around x^o, we may always find a nonsingular $m \times m$ matrix $\bar{\beta}(x)$, defined in a neighborhood U of x^o, such that

$$G_0(x) = \text{span}\{\bar{g}_1(x),\ldots,\bar{g}_{\nu_0}(x)\}$$

and

$$\bar{g}_{\nu_0+1}(x) = \ldots = \bar{g}_m(x) = 0$$

for all $x \in U$, where

$$\bar{g}_i(x) = (g(x)\bar{\beta}(x))_i$$

for $1 \le i \le m$. If a feedback (α,β) solves the State-Space Linearization Problem for the set $f,\bar{g}_1,\ldots,\bar{g}_{\nu_0}$, then it is easily seen that a feedback of the form

$$\alpha' = \bar{\beta}\begin{pmatrix}\alpha\\0\end{pmatrix} \qquad \beta' = \bar{\beta}\begin{pmatrix}\beta & 0\\0 & I\end{pmatrix}$$

solves the problem for the original set $f,g_1,\ldots,g_m$.

For the sake of simplicity, we break up the construction in two stages.

(i) Recursive construction of a coordinates transformation around the point x^o.

Step (1): By assumption

$$\dim G_{i^*} = n$$

and $\dim G^{\perp}_{i^*-1} = m_0 > 0$. Moreover, G_{i^*-1} is assumed to be involutive. Then, by Frobenius theorem, we know that there exist a neighborhood U_1 of x^o and m_0 functions $h_{01},\ldots,h_{0m_0}$ defined on U_1, whose differentials span $G^{\perp}_{i^*-1}(x)$ at all $x \in U_1$. In particular,

$$\langle dh_{0i}, ad_f^{\alpha}g_j\rangle(x) = 0 \tag{6.38}$$

for all $1 \le j \le m$, $1 \le i \le m_0$, $0 \le \alpha \le i^*-1$ and all $x \in U_1$. Moreover, the differentials $dh_{01}(x),\ldots,dh_{0m_0}(x)$ are linearly independent at all $x \in U_1$.

We claim that the $m_0 \times m$ matrix

$$M_0 = \{m^{(0)}_{ij}(x)\} = \{\langle dh_{0i}, ad_f^{i^*}g_j\rangle(x)\}$$

has rank m_0 at all $x \in U_1$. For, suppose it is false at some $\bar{x} \in U_1$.

Then, there exist real numbers $c_1,\ldots,c_{m_0}$ such that

$$\langle \sum_{i=1}^{m_0} c_i dh_{0i}, ad_f^{i^*}g_j\rangle(\bar{x}) = 0$$

for all $1 \le j \le m$. This, together with (6.38), implies

$$\langle \sum_{i=1}^{m_0} c_i dh_{0i}, ad_f^{\alpha}g_j\rangle(\bar{x}) = 0 \tag{6.39}$$

for all $1 \le j \le m$, $0 \le \alpha \le i^*$, and this in turn implies

$$\langle \sum_{i=1}^{m_0} c_i dh_{0i}(\bar{x}), v\rangle = 0$$

for all $v \in G_{i^*}(\bar{x})$. Since $\dim G_{i^*} = n$, then $\sum_{i=1}^{m_0} c_i dh_{0i}(\bar{x})$ must be a zero covector, but since $dh_{01}(\bar{x}),\ldots,dh_{0m_0}(\bar{x})$ are independent, then $c_1 = \ldots = c_{m_0} = 0$.

Step (2): Consider the distribution G_{i^*-2}, which is such that

$$\dim G^{\perp}_{i^*-2} = 2m_0 + m_1$$

We claim that $dL_fh_{01},\ldots,dL_fh_{0m_0}$ are such that

$$\langle dL_fh_{0i}, ad_f^{\alpha}g_j\rangle(x) = 0$$

for all $1 \le j \le m$, $1 \le i \le m_0$, $0 \le \alpha \le i^*-2$ and all $x \in U_1$. This comes from the property

$$-\langle dL_fh_{0i}, ad_f^{\alpha} g_j\rangle = \langle dh_{0i}, ad_f^{\alpha+1}g_j\rangle - L_f\langle dh_{0i}, ad_f^{\alpha}g_j\rangle$$

in which both the terms are zero on U_1 because $\alpha \le i^*-2$.

We claim also that the $2m_0$ differentials

$$\{dh_{01}(x),\ldots,dh_{0m_0}(x), dL_fh_{01}(x),\ldots,dL_fh_{0m_0}(x)\} \tag{6.40}$$

are linearly independent all $x \in U_1$. For, suppose this is false; then, for suitable reals c_{1i}, c_{2i}, we had

$$\sum_{i=1}^{m_0} c_{1i}dh_{0i}(\bar{x}) + \sum_{i=1}^{m_0} c_{2i}dL_fh_{0i}(\bar{x}) = 0 \tag{6.41}$$

at some $\bar{x} \in U$. This would imply

$$\langle \sum_{i=1}^{m_0} (c_{1i} dh_{0i} + c_{2i} dL_f h_{0i}), ad_f^{i^*-1} g_j \rangle(\bar{x}) = 0$$

for all $1 \leq j \leq m$. This in turn implies (because of (6.38))

$$\langle \sum_{i=1}^{m_0} c_{2i} dL_f h_{0i}, ad_f^{i^*-1} g_j \rangle(\bar{x}) = -\langle \sum_{i=1}^{m_0} c_{2i} dh_{0i}, ad_f^{i^*} g_j \rangle(\bar{x}) = 0$$

i.e. a contradiction, like in (6.39). Therefore $c_{21} = \dots = c_{2m_0} = 0$ in (6.41), and also $c_{11} = \dots = c_{1m_0} = 0$ because $dh_{01}(\bar{x}), \dots, dh_{0m_0}(\bar{x})$ are linearly independent.

If $m_1 = 0$, the $2m_0$ covectors (6.40) span $G^{\perp}_{i^*-2}$. If $m_1 > 0$, using again Frobenius theorem (because G_{i^*-2} is involutive), we may find m_1 more functions $h_{11}(x), \dots, h_{1m_1}(x)$, defined in a neighborhood $U_2 \subset U_1$ of x^o, such that the $2m_0+m_1$ differentials

(6.43) $$\{dh_{01}(x), \dots, dh_{0m_0}(x), dL_f h_{01}(x), \dots, dL_f h_{0m_0}(x), dh_{11}(x), \dots, dh_{1m_1}(x)\}$$

are linearly independent and

$$\langle dh_{1i}, ad_f^{\alpha} g_j \rangle(x) = 0$$

for all $1 \leq j \leq m$, $1 \leq i \leq m_1$, $0 \leq \alpha \leq i^*-2$ and all $x \in U_2$.

We claim that the $(m_0+m_1) \times m$ matrix

$$\begin{pmatrix} M_0 \\ M_1 \end{pmatrix}$$

where M_0 is as before and M_1 defined as

$$M_1 = \{m_{ij}^{(1)}(x)\} = \{\langle dh_{1i}, ad_f^{i^*-1} g_j \rangle(x)\}$$

has rank m_0+m_1 at all $x \in U_2$.

For suppose for some reals $c_{01}, \dots, c_{0m_0}, c_{11}, \dots, c_{1m_1}$ we had

$$\langle \sum_{i=1}^{m_0} c_{0i} dh_{0i}(\bar{x}), ad_f^{i^*} g_j(\bar{x}) \rangle + \langle \sum_{i=1}^{m_1} c_{1i} dh_{1i}(\bar{x}), ad_f^{i^*-1} g_j(\bar{x}) \rangle = 0$$

at some $\bar{x} \in U_2$. Then (recall Remark (6.17))

(6.44) $$\langle \sum_{i=1}^{m_0} c_{0i} dL_f h_{0i}(\bar{x}) + \sum_{i=1}^{m_1} c_{1i} dh_{1i}(\bar{x}), ad_f^{i^*-1} g_j(\bar{x}) \rangle = 0$$

The covector $\sum_{i=1}^{m_0} c_{0i} dL_f h_{0i}(\bar{x}) + \sum_{i=1}^{m_1} c_{1i} dh_{1i}(\bar{x})$ annihilates, as we have seen before, all $ad_f^{\alpha} g_j(\bar{x})$, $\alpha \leq i^*-2$, $1 \leq j \leq m$, but (6.44) tells us that it also annihilates $ad_f^{i^*-1} g_j(\bar{x})$. Thus, this covector annihilates all vectors in G_{i^*-1}.

From the previous discussion, we conclude that this covector must belong to span$\{dh_{01}(\bar{x}), \dots, dh_{0m_0}(\bar{x})\}$, but this is a contradiction, because the covectors (6.43) are linearly independent. Therefore, the c_{0i}'s and c_{1i}'s of (6.44) must be zero.

Eventually, with this procedure we end up with a set of functions

(6.45) $$\begin{array}{l} h_{01}, \dots, h_{0m_0}, L_f h_{01}, \dots, L_f h_{0m_0}, \dots, L_f^{i^*-1} h_{01}, \dots, L_f^{i^*-1} h_{0m_0} \\ h_{11}, \dots, h_{1m_1}, \dots, L_f^{i^*-2} h_{11}, \dots, L_f^{i^*-2} h_{1m_1} \\ \dots \\ h_{i^*-1,1}, \dots, h_{i^*-1,m_{i^*-1}} \end{array}$$

(of course, some of these lines may be missing if some m_i is zero) with the following properties:

- the total number of functions is

$$i^* m_0 + (i^*-1) m_1 + \dots + 2m_{i^*-2} + m_{i^*-1} = n-m$$

- the n-m differentials of these functions are independent at all $x \in U$, a neighborhood of x^o,
- the $\nu_1 \times m$ matrix

$$\begin{pmatrix} M_0 \\ \vdots \\ M_{i^*-1} \end{pmatrix}$$

where M_i is $m_i \times m$ and

$$m_{\ell j}^{(i)}(x) = \langle dh_{i\ell}, ad_f^{i^*-i} g_j \rangle(x)$$

has rank ν_1 at all $x \in U$.

If $m_{i^*} > 0$, one may still find m_{i^*} more functions $h_{i^*1},\dots,h_{i^*m_{i^*}}$, that, together with the functions (6.45) and with the additional functions $L_f^{i^*}h_{01},\dots,L_f^{i^*}h_{0m_0},\dots,L_f h_{i^*-1,1},\dots,L_f h_{i^*-1,m_{i^*-1}}$, give rise to a set of n linearly independent differentials at x^o.

For convenience, let us relabel the functions h_{ij} and set

$$\varphi_i = h_{0i} \qquad \text{if} \qquad 1 \le i \le \nu_{i^*}$$

$$\varphi_i = h_{1,i-\nu_{i^*}} \qquad \text{if} \quad m_1 > 0 \quad \text{and} \quad \nu_{i^*}+1 \le i \le \nu_{i^*-1}$$

$$\dots \qquad \dots$$

$$\varphi_i = h_{i^*,i-\nu_1} \qquad \text{if} \quad m_{i^*} > 0 \quad \text{and} \quad \nu_1+1 \le i \le \nu_0 = m$$

The previous constructions tell us that the mapping $F : x \mapsto \mathrm{col}(\xi_1(x),\dots,\xi_m(x))$, where

$$\xi_i(x) = \begin{pmatrix} \varphi_i(x) \\ L_f\varphi_i(x) \\ \vdots \\ L_f^{\kappa_i-1}\varphi_i(x) \end{pmatrix}$$

qualifies as a local diffeomorphism around x^o.

Moreover, by construction,

(6.46) $$\langle dL_f^{\alpha}\varphi_i, g_j\rangle(x) = 0$$

for all $0 \le \alpha \le \kappa_i-2$, $1 \le i,j \le m$, at all x around x^o, and the m×m matrix

(6.47) $$A(x) = \{a_{ij}(x)\} = \{\langle dL_f^{\kappa_i-1}\varphi_i, g_j\rangle(x)\}$$

is nonsingular at $x = x^o$.

(ii) Construction of the linearizing feedback. From the conditions (6.46) and (6.47), we see that the control system

(6.48a) $$\dot{x} = f(x) + \sum_{i=1}^{m} g_i(x)u_i$$

with (dummy) outputs

(6.48b) $$y_i = \varphi_i(x) \qquad 1 \le i \le m$$

is such that:

- the characteristic number ρ_i associated with the i-th output channel is exactly equal to κ_i-1,
- the single-outputs noninteracting control problem is solvable around x^o.

Choose a feedback α and β as a solution of the equations

$$A(x)\alpha(x) = -\begin{pmatrix} L_f^{\kappa_1}\varphi_1(x) \\ \vdots \\ L_f^{\kappa_m}\varphi_m(x) \end{pmatrix}$$

$$A(x)\beta(x) = I$$

(they correspond to the equations IV.(4.4a) with $\gamma_i = 0$ and IV.(4.4b) with δ_i the i-th row of an m×m identity matrix). Under this feedback, the system (6.48) splits into m noninteracting single-input single-output channels. In particular, in the new coordinates defined at the previous stage, each subsystem is described by equations of the form (see IV.(4.8))

$$\dot{\xi}_i = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}\xi_i + \begin{pmatrix} 0 \\ 0 \\ \cdot \\ 0 \\ 1 \end{pmatrix} v_i$$

$$y_i = (1 \quad 0 \quad 0 \quad \dots \quad 0 \quad 0)\xi_i$$

This completes the proof. □

APPENDIX

BACKGROUND MATERIAL IN DIFFERENTIAL GEOMETRY

1. Some facts from advanced calculus

Let A be an open subset of $\mathbb{R}^n$ and f: A $\to$ $\mathbb{R}$ a function. The value of f at $x = (x_1,\dots,x_n)$ is denoted $f(x) = f(x_1,\dots,x_n)$. The function f is said to be a function of class C^∞ (or, simply, C^∞ or also, a *smooth* function) if its partial derivatives of any order with respect to $x_1,\dots,x_n$ exist and are continuous. A function f is said to be analytic (sometimes noted as C^ω) if it is C^∞ and for each point $x^o \in A$ there exists a neighborhood U of x^o, such that the Taylor series expansion of f at x^o converges to f(x) for all x $\in$ U.

Example. A typical example of a function which is C^∞ but not analytic is the function f: $\mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = 0 \quad \text{if} \quad x \le 0$$

$$f(x) = \exp(-\frac{1}{x}) \quad \text{if} \quad x > 0 \qquad \square$$

A mapping F: A $\to \mathbb{R}^m$ is a collection $(f_1,\dots,f_m)$ of functions f_i: A $\to \mathbb{R}$. The mapping F is C^∞ if all f_i's are C^∞.

Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^n$ be open sets. A mapping F: U $\to$ V is a diffeomorphism if is bijective (i.e. one-to-one and onto) and both F and F^{-1} are of class C^∞. The *jacobian matrix* of F at a point x is the matrix

$$\frac{\partial F}{\partial x} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \cdot & \cdots & \cdot \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

The value of $\frac{\partial F}{\partial x}$ at a point $x = x^o$ is sometimes denoted $(\frac{\partial F}{\partial x})_{x^o}$.

Theorem. (Inverse function theorem). Let A be an open set of $\mathbb{R}^n$ and F: A $\to \mathbb{R}^n$ a C^∞ mapping. If $(\frac{\partial F}{\partial x})_{x^o}$ is nonsingular at some $x^o \in A$, then there exists an open neighborhood U of x^o in A such that V = F(U) is open in $\mathbb{R}^n$ and the restriction of F to U is a diffeomorphism onto V.

Theorem. (Rank theorem). Let $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^m$ be open sets, F:A $\to$ B a C^∞ mapping. Suppose $(\frac{\partial F}{\partial x})_x$ has rank k for all x $\in$ A. For each point $x^o \in A$ there exist a neighborhood A_o of x^o in A and an open neighborhood B_o of $F(x^o)$ in B, two open sets $U \in \mathbb{R}^n$ and $V \subset \mathbb{R}^m$, and two diffeomorphisms G: U $\to A_o$ and H: $B_o \to$ V such that $H\circ F\circ G(U) \subset V$ and such that for all $(x_1,\dots,x_n) \in U$

$$(H\circ F\circ G)(x_1,\dots,x_n) = (x_1,\dots,x_k,0,\dots,0)$$

Remark. Let P_k denote the mapping P_k: $\mathbb{R}^n \to \mathbb{R}^m$ defined by

$$P_k(x_1,\dots,x_n) = (x_1,\dots,x_k,0,\dots,0)$$

Then, since H and G are invertible, one may restate the previous expression as

$$F = H^{-1}\circ P_k\circ G^{-1}$$

which holds at all points of A_o.

Theorem. (Implicit function theorem). Let $A \subset \mathbb{R}^m$ and $B \subset \mathbb{R}^n$ be open sets. Let F: A$\times$B $\to \mathbb{R}^n$ be a C^∞ mapping. Let $(x,y)=(x_1,\dots,x_m,y_1,\dots,y_n)$ denote a point of A$\times$B. Suppose that for some $(x^o,y^o) \in A\times B$

$$F(x^o,y^o) = 0$$

and that the matrix

$$\frac{\partial F}{\partial y} = \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \cdot & \cdots & \cdot \\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_n} \end{pmatrix}$$

is nonsingular at (x^o,y^o). Then, there exists open neighborhoods A_o of x^o in A and B_o of y^o in B and a unique C^∞ mapping G: $A_o \to B_o$ such that

$$F(x,G(x)) = 0$$

for all $x \in A_o$.

Remark. As an application of the implicit function theorem, consider the following corollary. Let A be an open set in $\mathbb{R}^m$, let M be a k×n matrix whose entries are real-valued C^∞ functions defined on A and b a k-vector whose entries are also real-valued C^∞ functions defined on A. Suppose that for some $x^o \in A$

$$\text{rank } M(x^o) = k$$

Then, there exist an open neighborhood U of x^o and a C^∞ mapping $G : U \to \mathbb{R}^n$ such that

$$M(x)G(x) = b(x)$$

for all $x \in U$.

In other words, the equation

$$M(x)y = b(x)$$

has at least a solution which is a C^∞ function of x in a neighborhood of x^o. If k = n this solution is unique.

2. Some elementary notions of topology

This section is a review of the most elementary topological concepts that will be encountered later on.

Let S be a set. A *topological structure*, or a *topology*, on S is a collection of subsets of S, called *open* sets, satisfying the axioms

(i) the union of any number of open sets is open
(ii) the intersection of any finite number of open sets is open
(iii) the set S and the empty set $\emptyset$ are open

A set S with a topology is called a *topological space*.

A *basis* for a topology is a collection of open sets, called *basic open sets*, with the following properties

(i) S is the union of basic open sets
(ii) a nonempty intersection of two basic open sets is a union of basic open sets.

A *neighborhood* of a point p of a topological space is any open set which contains p.

Let S_1 and S_2 be topological spaces and F a mapping $F: S_1 \to S_2$. The mapping F is *continuous* if the inverse image of every open set of S_2 is an open set of S_1. The mapping F is *open* if the image of an open

set of S_1 is an open set of S_2. The mapping F is an *homeomorphism* if is a bijection and both continuous and open.

If F is an homeomorphism, the inverse mapping F^{-1} is also an homeomorphism.

Two topological spaces S_1, S_2 such that there is an homeomorphism $F: S_1 \to S_2$ are said to be *homeomorphic*.

A subset U of a topological space is said to be *closed* if its complement $\tilde{U}$ in S is open. It is easy to see that the intersection of any number of closed sets is closed, the union of any finite number of closed sets is closed, and both S and $\emptyset$ are closed.

If S_o is a subset of a topological space S, there is a unique open set, noted $\text{int}(S_o)$ and called the *interior* of S_o, which is contained in S_o and contains any other open set contained in S_o. As a matter of fact, $\text{int}(S_o)$ is the union of all open sets contained in S_o. Likewise, there is a unique closed set, noted $\text{cl}(S_o)$ and called the *closure* of S_o, which contains S_o and is contained in any other closed set which contains S_o. Actually, $\text{cl}(S_o)$ is the intersection of all closed sets which contain S_o.

A subset of S is said to be *dense* in S if its closure coincides with S.

If S_1 and S_2 are topological spaces, then the cartesian product $S_1 \times S_2$ can be given a topology taking as a basis the collection of all subsets of the form $U_1 \times U_2$, with U_1 a basic open set of S_1 and U_2 a basic open set of S_2. This topology on $S_1 \times S_2$ is sometimes called the *product topology*.

If S is a topological space and S_1 a subset of S, then S_1 can be given a topology taking as open sets the subsets of the form $S_1 \cap U$, with U any open set in S. This topology on S_1 is sometimes called the *subset topology*.

Let $F: S_1 \to S_2$ be a continuous mapping of topological spaces, and let $F(S_1)$ denote the image of F. Clearly, $F(S_1)$ with the subset topology is a topological space . Since F is continuous, the inverse image of any open set of $F(S_1)$ is an open set of S_1. However, not all open sets of S_1 are taken onto open sets of $F(S_1)$. In other words, the mapping $F': S_1 \to F(S_1)$ defined by $F'(p) = F(p)$ is continuous but not necessarily open. The set $F(S_1)$ can be given another topology, taking as open sets in $F(S_1)$ the images of open sets in S_1. It is easily seen that this new topology, sometimes called the *induced topology*, contains the subset topology (i.e. any set which is open in the subset topology is open also in the induced topology), and that the mapping F' is now open. If F is an injection, then S_1 and $F(S_1)$ endowed with the induced topology are homeomorphic.

A topological space S is said to satisfy the *Hausdorff separation axiom* (or, briefly, to be an Hausdorff space) if any two different points p_1 and p_2 have disjoint neighborhoods.

3. Smooth manifolds

Definition. A *locally Euclidean space* X of dimension n is a topological space such that, for each $p \in X$, there exists a homeomorphism φ mapping some open neighborhood of p onto an open set in $\mathbb{R}^n$. □

Definition. A *Manifold* N of dimension n is a topological space which is locally Euclidean of dimension n, is Hausdorff and has a countable basis. □

It is not possible that an open subset U of $\mathbb{R}^n$ be homeomorphic to an open subset V of $\mathbb{R}^m$, if $n \neq m$ (Brouwer's theorem on invariance of domain). Therefore, the *dimension* of a locally Euclidean space is a well-defined object.

A *coordinate chart* on a manifold N is a pair (U,φ), where U is an open set of N and φ a homeomorphism of U onto an open set of $\mathbb{R}^n$. Sometimes φ is represented as a set $(\varphi_1,\dots,\varphi_n)$, and $\varphi_i: U \to \mathbb{R}$ is called the i-th *coordinate function*. If $p \in U$, the n-tuple of real numbers $(\varphi_1(p),\dots,\varphi_n(p))$ is called the set of *local coordinates* of p in the coordinate chart (U,φ). A coordinate chart (U,φ) is called a *cubic* coordinate chart if $\varphi(U)$ is an open cube about the origin in $\mathbb{R}^n$. If $p \in U$ and $\varphi(p) = 0$, then the coordinate chart is said to be *centered* at p.

Let (U,φ) and (V,ψ) be two coordinate charts on a manifold N, with $U \cap V \neq \emptyset$. Let $(\psi_1,\dots,\psi_n)$ be the set of coordinate functions associated with the mapping ψ. The homeomorphism

$$\psi\circ\varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$$

taking, for each $p \in U \cap V$, the set of local coordinates $(\varphi_1(p),\dots,\varphi_n(p))$ into the set of local coordinates $(\psi_1(p),\dots,\psi_n(p))$, is called a *coordinates transformation* on $U \cap V$. Clearly, $\varphi\circ\psi^{-1}$ gives the inverse mapping, which expresses $(\varphi_1(p),\dots,\varphi_n(p))$ in terms of $(\psi_1(p),\dots,\psi_n(p))$.

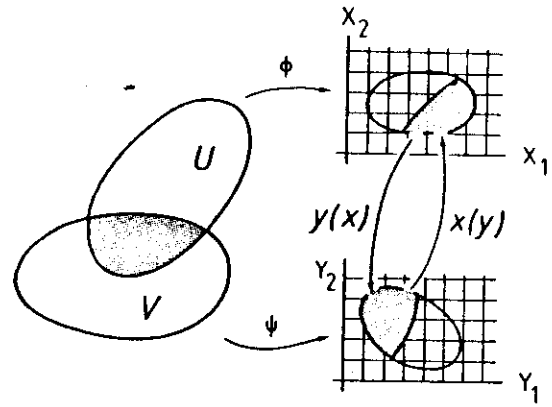
Frequently, the set $(\varphi_1(p),\dots,\varphi_n(p))$ is represented as an n-vector $x = \mathrm{col}(x_1,\dots,x_n)$, and the set $(\psi_1(p),\dots,\psi_n(p))$ as an n-vector $y = \mathrm{col}(y_1,\dots,y_n)$. Consistently, the coordinate transforma-

tion $\psi\circ\varphi^{-1}$ can be represented in the form

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} y_1(x_1,\dots,x_n) \\ \vdots \\ y_n(x_1,\dots,x_n) \end{pmatrix} = y(x)$$

and the inverse transformation $\varphi\circ\psi^{-1}$ in the form

$$x = x(y)$$



Two coordinate charts (U,φ) and (V,ψ) are C^∞-*Compatible* if, whenever $U \cap V \neq \emptyset$, the coordinate transformation $\psi\circ\varphi^{-1}$ is a diffeomorphism, i.e. if y(x) and x(y) are both C^∞ maps.

A C^∞ *atlas* on a manifold N is a collection $A = \{(U_i,\varphi_i)\}_{i\in I}$ of pairwise C^∞-compatible coordinate charts, with the property that $\bigcup_{i\in I} U_i = N$. An atlas is *complete* if not properly contained in any other atlas.

Definition. A smooth or C^∞ *manifold* is a manifold equipped with a complete C^∞ atlas. □

Remark. If A is any C^∞ atlas on a manifold N, there exists a *unique* complete C^∞ atlas A^* containing A. The latter is defined as the set of all coordinate charts (U,φ) which are compatible with every coordinate chart (U_i,φ_i) of A. This set contains A, is a C^∞ atlas, and is complete by construction. □

Some elementary examples of smooth manifolds are the ones described below.

Example. Any open set U of $\mathbb{R}^n$ is a smooth manifold, of dimension n. For, consider the atlas A consisting of the (single) coordinate chart (U, identity map on U) and let A^* denote the unique complete atlas containing A. In particular, $\mathbb{R}^n$ is a smooth manifold.

Remark. One may define different complete C^∞ atlases on the same manifold, as the following example shows. Let $N = \mathbb{R}$, and consider the coordinate charts $(\mathbb{R},\varphi)$ and $(\mathbb{R},\psi)$, with

$$\varphi(x) = x$$

$$\psi(x) = x^3$$

Since $\varphi^{-1}(x) = x$ and $\psi^{-1}(x) = x^{1/3}$,

$$\varphi\circ\psi^{-1}(x) = x^{1/3}$$

and the two charts are not compatible. Therefore the unique complete atlas A^*_φ which includes $(\mathbb{R},\varphi)$ and the unique complete atlas A^*_ψ which includes $(\mathbb{R},\psi)$ are different. This means that the same manifold N may be considered as a substrate of two different objects (two *smooth* manifolds), one arising with the atlas A^*_φ and the other with the atlas A^*_ψ . □

Example. Let U be an open set of $\mathbb{R}^m$ and let $\lambda_1,\ldots,\lambda_{m-n}$ be real-valued C^∞ functions defined on U. Let N denote the (closed) subset of U on which all functions $\lambda_1,\ldots,\lambda_{m-n}$ vanish, i.e. let

$$N = \{x \in U : \lambda_i(x) = 0,\ 1 \le i \le m-n\}$$

Suppose the rank of the jacobian matrix

$$\begin{pmatrix} \frac{\partial\lambda_1}{\partial x_1} & \cdots & \frac{\partial\lambda_1}{\partial x_m} \\ \cdot & \cdots & \cdot \\ \frac{\partial\lambda_{m-n}}{\partial x_1} & \cdots & \frac{\partial\lambda_{m-n}}{\partial x_m} \end{pmatrix}$$

is m-n at all $x \in N$. Then N is a smooth manifold of dimension n.

The proof of this essentially depends on the Implicit Function Theorem, and uses the following arguments. Let $x^o=(x^o_1,\ldots,x^o_n,x^o_{n+1},\ldots,x^o_m)$ be a point of N and assume, without loss of generality, that the matrix

$$\begin{pmatrix} \frac{\partial\lambda_1}{\partial x_{n+1}} & \cdots & \frac{\partial\lambda_1}{\partial x_m} \\ \cdot & \cdots & \cdot \\ \frac{\partial\lambda_{m-n}}{\partial x_{n+1}} & \cdots & \frac{\partial\lambda_{m-n}}{\partial x_m} \end{pmatrix}$$

is nonsingular at x^o. Then, there exist neighborhoods A_o of $(x^o_1,\ldots,x^o_n)$ in $\mathbb{R}^n$ and B_o of $(x^o_{n+1},\ldots,x^o_m)$ in $\mathbb{R}^{m-n}$ and a C^∞ mapping $G: A_o \to B_o$ such that

$$\lambda_i(x_1,\ldots,x_n,g_1(x_1,\ldots,x_n),\ldots,g_{m-n}(x_1,\ldots,x_n)) = 0$$

for all $1 \le i \le m-n$. This makes it possible to describe points of N around x^o as n-tuples $(x_1,\ldots,x_m)$ such that $x_{n+i} = g_i(x_1,\ldots,x_n)$ for $1 \le i \le m-n$. In this way one can construct a coordinate chart around each point x^o of N and the coordinate charts thus defined form a C^∞ atlas.

A manifold of this type is sometimes called a smooth *hypersurface* in $\mathbb{R}^m$. An important example of hypersurface is the *sphere* S^{m-1}, defined by taking n = m-1 and

$$\lambda_1 = x_1^2 + x_2^2 + \ldots + x_m^2 - 1$$

The set of points of $\mathbb{R}^m$ on which $f_1(x) = 0$ consists of all the points on a sphere of radius 1 centered at the origin. Since

$$\left(\frac{\partial\lambda_1}{\partial x_1} \cdots \frac{\partial\lambda_1}{\partial x_m}\right)$$

never vanishes on this set, the required conditions are satisfied and the set is a smooth manifold, of dimension m-1.

Example. An open subset N' of a smooth manifold N is itself a smooth manifold. The topology of N' is the subset topology. If (U,φ) is a coordinate chart of a complete C^∞ atlas of N, such that $U \cap N' \neq \emptyset$, then the pair (U',φ') defined as

$$U' = U \cap N'$$

$$\varphi' = \text{restriction of } \varphi \text{ to } U'$$

is a coordinate chart of N'. In this way, one may define a complete C^∞ atlas of N'. The dimension of N' is the same as that of N.

Example. Let M and N be smooth manifolds, of dimension m and n. Then the cartesian product M×N is a smooth manifold. The topology of M×N is the product topology. If (U,φ) and (V,ψ) are coordinate charts of M and N, the pair $(U\times V,(\varphi,\psi))$ is a coordinate chart of M×N. The dimension of M×N is clearly m+n.

An important example of this type of manifold is the *torus* $T^2 = S^1\times S^1$, the cartesian product of two circles. □

Let λ be a real-valued function defined on a manifold N. If (U,φ) is a coordinate chart on N, the composed function

$$\hat{\lambda} = \lambda\circ\varphi^{-1} : \varphi(U) \to \mathbb{R}$$

taking, for each $p \in U$, the set of local coordinates $(x_1,\ldots,x_n)$ of p into the real number $\lambda(p)$, is called an *expression of λ in local coordinates*.

In practice, whenever no confusion arises, one often uses the same symbol λ to denote $\lambda\circ\varphi^{-1}$, and write $\lambda(x_1,\ldots,x_n)$ to denote the value of λ at a point p of local coordinates $(x_1,\ldots,x_n)$.

If N and M are manifolds, of dimension n and m, $F : N \to M$ is a mapping, (U,φ) a coordinate chart on N and (V,ψ) a coordinate chart on M, the composed mapping

$$\hat{F} = \psi\circ F\circ\varphi^{-1}$$

is called an expression of F in local coordinates. Note that this definition make sense only if $F(U) \cap V \neq \emptyset$. If this is the case, then $\hat{F}$ is well defined for all n-tuples $(x_1,\ldots,x_n)$ whose image under $F\circ\varphi^{-1}$ is a point in V.

Here again, one often uses F to denote $\psi\circ F\circ\varphi^{-1}$, writes $y_i = f_i(x_1,\ldots,x_n)$ to denote the value of the i-th coordinate of F(p), p being a point of local coordinates $(x_1,\ldots,x_n)$, and also

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} f_1(x_1,\ldots,x_n) \\ \vdots \\ f_m(x_1,\ldots,x_n) \end{pmatrix} = F(x)$$

Definition. Let N and M be smooth manifolds. A mapping $F: N \to M$ is a *smooth* mapping if for each $p \in N$ there exists coordinate charts (U,φ)

of N and (V,ψ) of M, with $p \in U$ and $F(p) \in V$, such that the expression of F in local coordinates is C^∞.

Remark. Note that the property of being smooth is independent of the choice of the coordinate charts on N and M. Different coordinate charts (U',φ') and (V',ψ') are by definition C^∞ compatible with the former and

$$\hat{F}' = \psi'\circ F\circ\varphi'^{-1} =$$

$$= \psi'\circ\psi^{-1}\circ\psi\circ F\circ\varphi^{-1}\circ\varphi\circ\varphi'^{-1} =$$

$$= (\psi'\circ\psi^{-1})\circ\hat{F}\circ(\varphi'\circ\varphi^{-1})^{-1}$$

being a composition of C^∞ functions is still C^∞. □

Definition. Let N and M be smooth manifolds, both of dimension n. A mapping $F : N \to M$ is a *diffeomorphsim* if F is bijective and both F and F^{-1} are smooth mappings. Two manifolds N and M are *diffeomorphic* if there exists a diffeomorphism $F : N \to M$. □

The *rank* of a mapping $F : N \to M$ at a point $p \in N$ is the rank of the jacobian matrix

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \cdot & \cdots & \cdot \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

at $x = \varphi(p)$. It must be stressed that, although apparently dependent on the choice of local coordinates, the notion of rank thus defined is actually coordinate-independent. The reader may easily verify that the ranks of the jacobian matrices of two different expressions of F in local coordinates are equal.

Theorem. Let N and M be smooth manifolds both of dimension n. A mapping $F : N \to M$ is a diffeomorphism if and only if F is bijective, F is smooth and rank(F) = n at all points of N.

Remark. In some cases, the assumption that functions, mappings, etc. are C^∞, may be replaced by the stronger assumption that functions, mappings, etc. are analytic. In this way one may define the notion of analytic manifold, analytic mappings of manifolds, and so on. We shall make this assumption explicitly whenever needed.

4. Submanifolds

Definitions. Let $F : N \to M$ be a smooth mapping of manifolds.

(i) F is an *immersion* if $\mathrm{rank}(F) = \dim(N)$ for all $p \in N$.

(ii) F is a *univalent immersion* if F is an immersion and is injective.

(iii) F is an *embedding* if F is a univalent immersion and the topology induced on F(N) by the one of N coincides with the topology of F(N) as a subset of M. □

Remark. The mapping F, being smooth, is in particular a continuous mapping of topological spaces. Therefore (see section 2) the topology induced on F(N) by the one of N may properly contain the topology of F(N) as a subset of M. This motivates the definition (iii). □

The difference between (i),(ii) and (iii) is clarified in the following examples.

Examples. Let $N = \mathbb{R}$ and $M = \mathbb{R}^2$. Let t denote a point in N and (x_1, x_2) a point in M. The mapping F is defined by

$$x_1(t) = at - \sin t$$
$$x_2(t) = \cos t$$

and, then,

$$\mathrm{rank}(F) = \mathrm{rank}\begin{pmatrix} a - \cos t \\ \\ -\sin t \end{pmatrix}$$

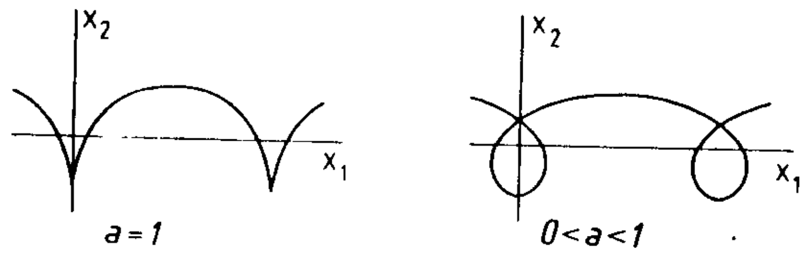
If a = 1 this mapping is *not* an immersion because rank(F) = 0 at $t = 2k\pi$ (for any integer k).

If $0 < a < 1$ the mapping is an immersion, because rank(F) = 1 for all t, but *not* a univalent immersion, because $F(t_1) = F(t_2)$ for all t_1, t_2 such that $t_1 = 2k\pi - \tau$, $t_2 = 2k\pi + \tau$ and $\sin \tau = a\tau$.

As a second example we consider the so-called "figure-eight". Let N be the open interval $(0, 2\pi)$ of the real line and $M = \mathbb{R}^2$. Let t denote a point in N and (x_1, x_2) a point in M. The mapping F is defined by

$$x_1(t) = \sin 2t$$

$$x_2(t) = \sin t$$



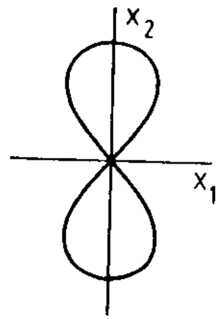
This mapping is an immersion because

$$\mathrm{rank}(F) = \mathrm{rank}\begin{pmatrix} \frac{dx_1}{dt} \\ \\ \frac{dx_2}{dt} \end{pmatrix} = \begin{pmatrix} 2\cos 2t \\ \\ \cos t \end{pmatrix} = 1$$

for all $0 < t < 2\pi$. It is also univalent because

$$F(t_1) = F(t_2) \Rightarrow t_1 = t_2$$

However, the mapping is *not* an embedding. For, consider the image of F.



The mapping F takes the open set $(\pi-\epsilon, \pi+\epsilon)$ of N onto a subset U' of F(N) which is open *by definition* in the topology induced by the one of N, but is not an open set in the topology of F(N) as a subset of M. This is because U' cannot be seen as the intersection of F(N) with an open set of $\mathbb{R}^2$.

As a third example one may consider the mapping $F : \mathbb{R} \to \mathbb{R}^3$ given by

$$x_1(t) = \cos 2\pi t$$

$$x_2(t) = \sin 2\pi t$$

$$x_3(t) = t$$

whose image is an "helix" winding on an infinite cylinder whose axis is the x_3 axis. The reader may easily check that is an embedding. □

The following theorem shows that the every immersion *locally* is an embedding.

Theorem. Let $F : N \to M$ be an immersion. For each $p \in N$ there exists a neighborhood U of p with the property that the restriction of F to U is an embedding.

Example. Consider again the "figure eight" discussed above. If U is any interval of the type $(\delta, 2\pi-\delta)$, then the critical situation we had before disappears and the image U' of $(\pi-\epsilon, \pi+\epsilon)$ is now open also in the topology of F(N) as a subset of $\mathbb{R}^2$. □

The notions of univalent immersion and of embedding are used in the following way.

Definition. The image F(N) of a univalent immersion is called an *immersed submanifold* of M. The image F(N) of an embedding is called an *embedded submanifold* of M.

Remark. Conversely, one may say that a subset M' of M is an immersed (respectively, embedded) submanifold of M if there is another manifold N and a univalent immersion (respectively, embedding) $F: N \to M$ such that $F(N) = M'$. □

The use of the word "submanifold" in the above definition clearly indicates the possibility of giving F(N) the structure of a smooth manifold, and this may actually be done in the following way. Let $M' = F(N)$ and $F': N \to M'$ denote the mapping defined by

$$F'(p) = F(p)$$

for all $p \in N$. Clearly, F' is a bijection. If the topology of M' is the one induced by that of N (i.e. open sets of M' are the images under F' of open sets of N), F' is a homeomorphism. Consequently, any coordinate chart (U, φ) of N induces a coordinate chart (V, ψ) of M', defined as

$$V = F'(U) \quad , \quad \psi = \varphi \circ (F')^{-1}$$

C^∞-compatible charts of N induce C^∞-compatible charts of M' and so complete C^∞-atlases induce complete C^∞-atlases. This gives M' the structure of a smooth manifold.

The smooth manifold M' thus defined is *diffeomorphic* to the smooth manifold N. A diffeomorphism between M' and N is indeed F' itself, which is bijective, smooth, and has rank equal to the dimension of N at each $p \in N$.

Embedded submanifolds can also be characterized in a different way, based on the following considerations.

Let M be a smooth manifold of dimension m and (U, φ) a cubic coordinate chart. Let n be an integer, $0 \le n < m$, and p a point of U. The subset of U

$$S_p = \{q \in U: x_i(q) = x_i(p), \; i = n+1, \ldots, m\}$$

is called an n-dimensional *slice* of U passing through p. In other words a slice of U is the locus of all points of U for which some coordinates (e.g. the last m-n) are constant.

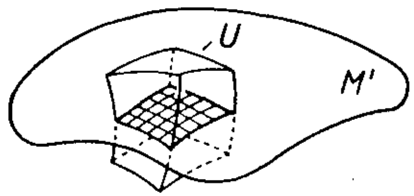
Theorem. Let M be a smooth manifold of dimension m. A subset M' of M is an embedded submanifold of dimension $n < m$ if and only if for each $p \in M'$ there exists a cubic coordinate chart (U, φ) of M, with $p \in U$, such that $U \cap M'$ coincides with an n-dimensional slice of U passing through p. □

This theorem provides a more "intrinsic" characterization of the notion of an embedded submanifold (of a manifold M), directly related to the existence of special coordinate charts (of M). Note that, if (U, φ) is a coordinate chart of M such that $U \cap M'$ is an n-dimensional slice of U, the pair (U', φ') defined as

$$U' = U \cap M'$$

$$\varphi'(p) = (x_1(p), \ldots, x_n(p))$$

is a coordinate chart of M'. This is illustrated in the following figure (where $M = \mathbb{R}^3$ and $n = 2$).



Remark. Note that an open subset M' of M is indeed an embedded submanifold of M, of the same dimension m. Thus, a submanifold M' of M may be a *proper* subset of M, although being a manifold of the same dimension.

Remark. It can be proven that any smooth hypersurface in $\mathbb{R}^m$ is an embedded submanifold of $\mathbb{R}^m$. Moreover it has also been shown that if N is an n-dimensional smooth manifold, there exist an integer $m \ge n$ and a mapping $F: N \to \mathbb{R}^m$ which is an embedding (Whitney's embedding theorem). In other words, any manifold is diffeomorphic to an embedded submanifold of $\mathbb{R}^m$, for a suitably large m.

Remark. Let V be an n-dimensional subspace of $\mathbb{R}^m$. Any subset of $\mathbb{R}^m$ of the form

$$x^o+V = \{x \in \mathbb{R}^m : x = x'+x^o;\ x' \in V\},$$

where x^o is some fixed point of $\mathbb{R}^m$, is indeed a smooth hypersurface and so an embedded submanifold of $\mathbb{R}^m$, of dimension n. This is sometimes called a *flat* submanifold of $\mathbb{R}^n$.

5. Tangent vectors

Let N be a smooth manifold of dimension n. A real-valued function λ is said to be *smooth in a neighborhood* of p, if the domain of λ includes an open set U of N containing p and the restriction of λ to U is a smooth function. The set of all smooth functions in a neighborhood of p is denoted $C^\infty(p)$. Note that $C^\infty(p)$ forms a vector space over the field $\mathbb{R}$. For, if λ,γ are functions in $C^\infty(p)$ and a,b are real numbers, the function $a\lambda+b\gamma$ defined as

$$(a\lambda+b\gamma)(q) = a\lambda(q)+b\gamma(q)$$

for all q in a neighborhood of p, is again a function in $C^\infty(p)$. Note also that two functions $\lambda,\gamma \in C^\infty(p)$ may be multiplied to give another element of $C^\infty(p)$, written $\lambda\gamma$ and defined as

$$(\lambda\gamma)(q) = \lambda(q)\cdot\gamma(q)$$

for all q in a neighborhood of p.

Definition. A *tangent vector* v at p is a map $v: C^\infty(p) \to \mathbb{R}$ with the following properties:

(i) (linearity) : $v(a\lambda+b\gamma) = av(\lambda)+bv(\gamma)$ for all $\lambda,\gamma \in C^\infty(p)$ and $a,b \in \mathbb{R}$

(ii) (Leibnitz rule): $v(\lambda\gamma) = \gamma(p)v(\lambda)+\lambda(p)v(\gamma)$ for all $\lambda,\gamma \in C^\infty(p)$.

Definition. Let N be a smooth manifold. The *tangent space* to N at p, written T_pN, is the set of all tangent vectors at p.

Remark. A map which satisfies the properties (i) and (ii) is also called a *derivation*.

Remark. The set T_pN forms a vector space over the field $\mathbb{R}$ under the rules of scalar multiplication and addition defined in the following way. If v_1,v_2 are tangent vectors and c_1,c_2 real numbers, $c_1v_1+c_2v_2$ is a new tangent vector which takes the function $\lambda \in C^\infty(p)$ into the real number

$$(c_1v_1+c_2v_2)(\lambda) = c_1v_1(\lambda) + c_2v_2(\lambda)$$

Remark. We shall see later on that, if the manifold N is a smooth hypersurface in $\mathbb{R}^m$, the object previously defined may be naturally identified with the intuitive notion of "tangent hyperplane" at a point. □

Let (U,φ) be a (fixed) coordinate chart around p. With this coordinate chart one may associate n tangent vectors at p, denoted

$$\left(\frac{\partial}{\partial\varphi_1}\right)_p, \ldots, \left(\frac{\partial}{\partial\varphi_n}\right)_p$$

defined in the following way

$$\left(\frac{\partial}{\partial\varphi_i}\right)_p(\lambda) = \left(\frac{\partial(\lambda\circ\varphi^{-1})}{\partial x_i}\right)_{x=\varphi(p)}$$

for $1 \le i \le n$. The right-hand-side is the value taken at $x = (x_1,\ldots,x_n) = \varphi(p)$ of the partial derivative of the function $\lambda\circ\varphi^{-1}(x_1,\ldots,x_n)$ with respect to x_i (recall that the function $\lambda\circ\varphi^{-1}$ is an expression of λ in local coordinates).

Theorem. Let N be a smooth manifold of dimension n. Let p be any point of N. The tangent space T_pN to N at p is an n-dimensional vector space over the field $\mathbb{R}$. If (U,φ) is a coordinate chart around p, then the tangent vectors $\left(\frac{\partial}{\partial\varphi_1}\right)_p,\ldots,\left(\frac{\partial}{\partial\varphi_n}\right)_p$ form a basis of T_pN. □

The basis $\{\left(\frac{\partial}{\partial\varphi_1}\right)_p,\ldots,\left(\frac{\partial}{\partial\varphi_n}\right)_p\}$ of T_pN is sometimes called the *natural basis* induced by the coordinate chart (U,φ).

Let v be a tangent vector at p. From the above theorem it is seen that

$$v = \sum_{i=1}^{n} v_i\left(\frac{\partial}{\partial\varphi_i}\right)_p$$

where $v_1,\ldots,v_n$ are real numbers. One may compute the v_i's explicitly in the following way. Let φ_i be the i-th coordinate function. Clearly

$\varphi_i \in C^\infty(p)$, and then

$$v(\varphi_i) = \sum_{j=1}^{n} v_j \left(\frac{\partial}{\partial\varphi_j}\right)_p (\varphi_i) = \sum_{j=1}^{n} v_j \left(\frac{\partial(\varphi_i \circ \varphi^{-1})}{\partial x_j}\right)_{x=\varphi(p)} = v_i$$

because $\varphi_i \circ \varphi^{-1}(x_1,\dots,x_n) = x_i$. Thus the real number v_i coincides with the value of v at φ_i, the i-th coordinate function.

A change of coordinates around p clearly induces a *change of basis* in T_pN. The computations involved are the following ones. Let (U,φ) and (V,ψ) be coordinate charts around p. Let $\{(\frac{\partial}{\partial\psi_1})_p,\dots,(\frac{\partial}{\partial\psi_n})_p\}$ denote the natural basis of T_pN induced by the coordinate chart (V,ψ). Then

$$\left(\frac{\partial}{\partial\psi_i}\right)_p(\lambda) = \left(\frac{\partial(\lambda\circ\psi^{-1})}{\partial y_i}\right)_{y=\psi(p)} = \left(\frac{\partial(\lambda\circ\varphi^{-1}\circ\varphi\circ\psi^{-1})}{\partial y_i}\right)_{y=\psi(p)} =$$

$$= \sum_{j=1}^{n} \left(\frac{\partial(\lambda\circ\varphi^{-1})}{\partial x_j}\right)_{x=\varphi(p)} \cdot \left(\frac{\partial(\varphi_j\circ\psi^{-1})}{\partial y_i}\right)_{y=\psi(p)} =$$

$$= \sum_{j=1}^{n} \left(\left(\frac{\partial}{\partial\varphi_j}\right)_p(\lambda)\right) \left(\frac{\partial(\varphi_j\circ\psi^{-1})}{\partial y_i}\right)_{y=\psi(p)}$$

In other words

$$\left(\frac{\partial}{\partial\psi_i}\right)_p = \sum_{j=1}^{n} \left(\frac{\partial(\varphi_j\circ\psi^{-1})}{\partial y_i}\right)_{y=\psi(p)} \left(\frac{\partial}{\partial\varphi_j}\right)_p$$

Note that the quantity

$$\frac{\partial(\varphi_j\circ\psi^{-1})}{\partial y_i}$$

is the element on the j-th row and i-th column of the *jacobian matrix* of the coordinate transformation

$$x = x(y)$$

So the elements of the columns of the jacobian matrix of $x = x(y)$ are the coefficients which express the vectors of the "new" basis as linear combinations of the vectors of the "old" basis.

If v is a tangent vector, and $(v_1,\dots,v_n)$, $(w_1,\dots,w_n)$ the n-tuples of real numbers which express v in the form

$$v = \sum_{i=1}^{n} v_i \left(\frac{\partial}{\partial\varphi_i}\right)_p = \sum_{i=1}^{n} w_i \left(\frac{\partial}{\partial\psi_i}\right)_p$$

then

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \cdot & \cdots & \cdot \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

Definition. Let N and M be smooth manifolds. Let $F : N \to M$ be a smooth mapping. The *differential* of F at $p \in N$ is the map

$$F_* : T_pN \to T_{F(p)}M$$

defined as follows. For $v \in T_pN$ and $\lambda \in C^\infty(F(p))$,

$$(F_*(v))(\lambda) = v(\lambda\circ F)$$

Remark. F_* is a map of the tangent space of N at a point p into the tangent space of M at the point F(p). If $v \in T_pN$, the value $F_*(v)$ of F_* at v is a tangent vector in $T_{F(p)}M$. So one has to express the way in which $F_*(v)$ maps the set $C^\infty(F(p))$, of all functions which are smooth a neighborhood of F(p), into $\mathbb{R}$. This is actually what the definition specifies. Note that there is one of such maps for *each* point p of N.

Theorem. The differential F_* is a linear map. □

Since F_* is a linear map, given a basis for T_pN and a basis for $T_{F(p)}M$ one may wish to find its matrix representation. Let (U,φ) be a coordinate chart around p, (V,ψ) a coordinate chart around $q = F(p)$, $\{(\frac{\partial}{\partial\varphi_1})_p,\dots,(\frac{\partial}{\partial\varphi_n})_p\}$ the natural basis of T_pN and $\{(\frac{\partial}{\partial\psi_1})_q,\dots,(\frac{\partial}{\partial\psi_m})_q\}$ the natural basis of T_qM. In order to find a matrix representation of F_*, one has simply to see how F_* maps $(\frac{\partial}{\partial\varphi_i})_p$ for each $1 \le i \le n$.

$$\left(F_*\left(\frac{\partial}{\partial\varphi_i}\right)_p\right)(\lambda) = \left(\frac{\partial}{\partial\varphi_i}\right)_p(\lambda\circ F) = \left(\frac{\partial(\lambda\circ F\circ\varphi^{-1})}{\partial x_i}\right)_{x=\varphi(p)} =$$

$$= \left(\frac{\partial(\lambda\circ\psi^{-1}\circ\psi\circ F\circ\varphi^{-1})}{\partial x_i}\right)_{x=\varphi(p)} = \sum_{j=1}^{m}\left(\frac{\partial(\lambda\circ\psi^{-1})}{\partial y_j}\right)_{y=\psi(q)}\left(\frac{\partial(\psi_j\circ F\circ\varphi^{-1})}{\partial x_i}\right)_{x=\varphi(p)}$$

$$= \sum_{j=1}^{m}\left(\left(\frac{\partial}{\partial\psi_j}\right)_q(\lambda)\right)\left(\frac{\partial(\psi_j\circ F\circ\varphi^{-1})}{\partial x_i}\right)_{x=\varphi(p)}$$

In other words

$$F_*\left(\frac{\partial}{\partial\varphi_i}\right)_p = \sum_{j=1}^{m}\left(\frac{\partial(\psi_j\circ F\circ\varphi^{-1})}{\partial x_i}\right)_{x=\varphi(p)}\left(\frac{\partial}{\partial\psi_j}\right)_q$$

Now, recall that $\psi\circ F\circ\varphi^{-1}$ is an expression of F in local coordinates. Then, the quantity

$$\frac{\partial(\psi_j\circ F\circ\varphi^{-1})}{\partial x_i}$$

is the element on the j-th row and i-th column of the *jacobian matrix* of the mapping expressing F in local coordinates. Using again

$$F(x) = F(x_1,\ldots,x_n) = \begin{pmatrix} F_1(x_1,\ldots,x_n) \\ \vdots \\ F_m(x_1,\ldots,x_n)\end{pmatrix}$$

to denote $\psi\circ F\circ\varphi^{-1}$, one has simply

$$F_*\left(\frac{\partial}{\partial\varphi_i}\right)_p = \sum_{j=1}^{m}\left(\frac{\partial F_j}{\partial x_i}\right)\left(\frac{\partial}{\partial\psi_j}\right)_q$$

If $v \in T_pN$ and $w = F_*(v) \in T_{F(p)}M$ are expressed as

$$v = \sum_{i=1}^{n} v_i\left(\frac{\partial}{\partial\varphi_i}\right)_p \qquad w = \sum_{i=1}^{m} w_i\left(\frac{\partial}{\partial\psi_i}\right)_q$$

then

$$\begin{pmatrix} w_1 \\ \vdots \\ w_m\end{pmatrix} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \cdot & \cdots & \cdot \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n}\end{pmatrix}\begin{pmatrix} v_1 \\ \vdots \\ v_n\end{pmatrix}$$

Remark. The matrix representation of F_* is exactly the jacobian of its expression in local coordinates. From this, it is seen that the rank of a mapping coincides with the rank of the corresponding differential.

Remark (Chain rule). It is easily seen that, if F and G are smooth mappings, then

$$(G\circ F)_* = G_*F_*$$

The following examples may clarify the notion of tangent space and the one of differential.

Example. The tangent vectors on $\mathbb{R}^n$. Let $\mathbb{R}^n$ be equipped with the "natural" complete atlas already considered in previous examples (i.e. the one including the chart ($\mathbb{R}^n$, identity map on $\mathbb{R}^n$)). Then, if v is a tangent vector at a point x and λ a smooth function

$$v(\lambda) = \sum_{i=1}^{n} v_i\left(\frac{\partial}{\partial x_i}\right)_x(\lambda) = \sum_{i=1}^{n}\left(\frac{\partial\lambda}{\partial x_i}\right)_x v_i$$

So, $v(\lambda)$ is just the value of the *derivative* of λ *along the direction* of the vector

$$\mathrm{col}(v_1,\ldots,v_n)$$

at the point x.

Remark. Let $F : N \to M$ be a univalent immersion. Let $n = \dim(N)$ and $m = \dim(M)$. By definition, F_* has rank n at each point. Therefore the image $F_*(T_pN)$ of F_*, at each point p, is a subspace of $T_{F(p)}M$ isomorphic to T_pN. The subspace $F_*(T_pN)$ can actually be *identified* with the tangent space at F(p) to the submanifold $M' = F(N)$. In order to understand this point, let F' denote the function $F' : N \to M'$ defined as

$$F'(p) = F(p)$$

for all $p \in N$. F' is a diffeomorphism and so F'_* is an isomorphism. Therefore the image $F'_*(T_pN)$ is exactly the tangent space at F'(p) to M'. Any tangent vector in $T_{F(p)}M'$ is the image $F'_*(v)$ of a (unique) vector $v \in T_pN$ and can be identified with the (unique) vector $F_*(v)$ of $F_*(T_pN)$.

In other words, the tangent space at p to a *submanifold* M' of M can be identified with a *subspace* of the tangent space at p to M.

The same considerations can be repeated in local coordinates. It is known that an immersion is locally an embedding. Therefore, around every point $p \in M'$ it is possible to find a coordinate chart (U,φ) of M, with the property that the pair (U',φ') defined by

$$U' = \{q \in U : \varphi_i(q) = \varphi_i(p),\ i = n+1,\dots,m\}$$

$$\varphi' = (\varphi_1,\dots,\varphi_n)$$

is a coordinate chart of M'. According to this choice, the tangent space to M' at p is identified with the n-dimensional subspace of T_pM spanned by the tangent vectors $\{(\frac{\partial}{\partial\varphi_1})_p,\dots,(\frac{\partial}{\partial\varphi_n})_p\}$. □

Example. The tangent vector to a smooth curve in $\mathbb{R}^n$. We define first the notion of a smooth curve in $\mathbb{R}^n$. Let $N = (t_1,t_2)$ be an open interval on the real line. A *smooth curve* in $\mathbb{R}^n$ is the image of a univalent immersion $\sigma : N \to \mathbb{R}^n$. Thus, a smooth curve is an immersed submanifold of $\mathbb{R}^n$. In N and $\mathbb{R}^n$ one may choose natural local coordinates as usual and, letting t denote an element of N, express σ by means of an n-tuple of scalar-valued functions $\sigma_1,\dots,\sigma_n$ of t.

A smooth curve is a 1-dimensional immersed submanifold of $\mathbb{R}^n$. At a point $\sigma(t_o)$, the tangent space to the curve is a 1-dimensional vector space which, as we have seen, may be identified with a subspace of the tangent space to $\mathbb{R}^n$ at this point. A basis of the tangent space to the curve at $\sigma(t_o)$ is given by the image under σ_* of $(\frac{d}{dt})_{t_o}$, a tangent vector at t_o to N. This image is computed as follows

$$\sigma_*(\frac{d}{dt})_{t_o} = \sum_{i=1}^{n} (\frac{d\sigma_i}{dt})_{t_o} (\frac{\partial}{\partial x_i})_{\sigma(t_o)}$$

Thinking of $t \in N$ as time and $\sigma(t)$ as a point moving in $\mathbb{R}^n$, we may interpret the vector

$$\mathrm{col}((\frac{d\sigma_1}{dt})_{t_o},\dots,(\frac{d\sigma_n}{dt})_{t_o})$$

as the velocity along the curve, evaluated at the point $\sigma(t^o)$. So, we have that the velocity vector at a point of the curve spans the tangent space to the curve at this point. From this point of view, we see that the notion of tangent space to a 1-dimensional manifold may be identified with the geometric notion of tangent line to a curve in a Euclidean space.

Example. Let h be a smooth function $h : \mathbb{R}^2 \to \mathbb{R}$ and $F : \mathbb{R}^2 \to \mathbb{R}^3$ a mapping defined by

$$F(x_1,x_2) = (x_1,x_2,h(x_1,x_2))$$

This mapping is an embedding and therefore $F(\mathbb{R}^2)$, a surface in $\mathbb{R}^2$, is an embedded submanifold of $\mathbb{R}^3$. At each point $F(x)$ of this surface, the tangent space, identified as a subspace of the tangent space to $\mathbb{R}^3$ at this point, may be computed as

$$\mathrm{span}\{F_*(\frac{\partial}{\partial x_1})_x\ ,\ F_*(\frac{\partial}{\partial x_2})_x\}$$

Now,

$$F_*(\frac{\partial}{\partial x_1})_x = \sum_{i=1}^{3} (\frac{\partial F_i}{\partial x_i})(\frac{\partial}{\partial x_i})_{F(x)} = (\frac{\partial}{\partial x_1})_{F(x)} + (\frac{\partial h}{\partial x_1})(\frac{\partial}{\partial x_3})_{F(x)}$$

$$F_*(\frac{\partial}{\partial x_2})_x = (\frac{\partial}{\partial x_2})_{F(x)} + (\frac{\partial h}{\partial x_2})(\frac{\partial}{\partial x_3})_{F(x)}$$

This tangent space to $F(\mathbb{R}^2)$ at some point $(x_1^o,x_2^o,h(x_1^o,x_2^o))$ is the set of tangent vectors whose expressions in local coordinates are of the form

$$v = \begin{pmatrix} \alpha \\ \beta \\ (\frac{\partial h}{\partial x_1})\alpha + (\frac{\partial h}{\partial x_2})\beta \end{pmatrix}$$

α,β being real numbers and $\frac{\partial h}{\partial x_1}, \frac{\partial h}{\partial x_2}$ being evaluated at $x_1 = x_1^o$ and $x_2 = x_2^o$. From this point of view, we see that the notion of tangent space to a 2-dimensional manifold may be identified with the geometric notion of tangent plane to a surface in a Euclidean space. □

One may define objects dual to the ones considered so far.

Definition. Let N be a smooth manifold. The *cotangent space* to N at p, written T_p^*N, is the dual space of T_pN. Elements of the cotangent space are called *tangent covectors*.

Remark. Recall that a dual space V^* of a vector space V is the space of all linear functions from V to $\mathbb{R}$. If $v^* \in V^*$, then $v^* : V \to \mathbb{R}$ and the value of v^* at $v \in V$ is written as $\langle v^*,v\rangle$. V^* forms a vector space over the field $\mathbb{R}$, with rules of scalar multiplication and addition which define $c_1v_1^*+c_2v_2^*$ in the following terms

$$\langle c_1v_1^*+c_2v_2^*,v\rangle = c_1\langle v_1^*,v\rangle + c_2\langle v_2^*,v\rangle$$

If $e_1,\dots,e_n$ is a basis of V , the unique basis $e_1^*,\dots,e_n^*$ of V^* which satisfies

$$\langle e_i^*,e_j\rangle = \delta_{ij}$$

is called a *dual basis*.

If V and W are vector spaces, $F : V \to W$ a linear mapping and $w \in W$, $v \in V$, the mapping $F^* : W^* \to V^*$ defined by

$$\langle F^*(w^*),v\rangle = \langle w^*,F(v)\rangle$$

is called the *dual* mapping (of F). □

Let λ be a smooth function $\lambda : N \to \mathbb{R}$. There is a natural way of identifying the differential λ_* of λ at p with an element of T_p^*N. For, observe that λ_* is a linear mapping

$$\lambda_* : T_pN \to T_{\lambda(p)}\mathbb{R}$$

and that $T_{\lambda(p)}\mathbb{R}$ is isomorphic to $\mathbb{R}$. The natural isomorphism between $\mathbb{R}$ and $T_{\lambda(p)}\mathbb{R}$ is the one in which the element c of $\mathbb{R}$ corresponds to the tangent vector $c(\frac{d}{dt})_t$. If $c(\frac{d}{dt})_t$ is the value at v of the differential λ_* at p, then c must depend linearly on v, i.e. there must exist a covector, denoted $(d\lambda)_p$, such that

$$\lambda_*(v) = \langle (d\lambda)_p,v\rangle(\frac{d}{dt})_t$$

Given a basis of T_pN, the covector $(d\lambda)_p$ (like any other covector), may be represented in matrix form. Let $\{(\frac{\partial}{\partial\varphi_1})_p,\dots,(\frac{\partial}{\partial\varphi_n})_p\}$ be the natural basis of T_pN induced by the coordinate chart (U,φ). The image under λ_* of a vector

$$v = \sum_{i=1}^n v_i(\frac{\partial}{\partial\varphi_i})_p$$

is the vector

$$\lambda_*(v) = (\sum_{i=1}^n \frac{\partial\lambda}{\partial x_i} v_i)(\frac{d}{dt})_t$$

and this shows that

$$\langle (d\lambda)_p,v\rangle = \begin{pmatrix}\frac{\partial\lambda}{\partial x_1} & \dots & \frac{\partial\lambda}{\partial x_n}\end{pmatrix}\begin{pmatrix}v_1\\ \vdots\\ v_n\end{pmatrix}$$

Remark. Note also that the value at λ of a tangent vector v is equal to the value at v of the tangent covector $(d\lambda)_p$, i.e.

$$v(\lambda) = \langle (d\lambda)_p,v\rangle \quad □$$

The dual basis of $\{(\frac{\partial}{\partial\varphi_1})_p,\dots,(\frac{\partial}{\partial\varphi_n})_p\}$ is computed as follows. From the equality $v(\lambda) = \langle (d\lambda)_p,v\rangle$ we deduce that

$$\langle (d\varphi_i)_p,(\frac{\partial}{\partial\varphi_j})_p\rangle = (\frac{\partial}{\partial\varphi_j})_p(\varphi_i) = \frac{\partial(\varphi_i\circ\varphi^{-1})}{\partial x_j} = \frac{\partial x_i}{\partial x_j} = \delta_{ij}$$

so that the desired dual basis is exactly provided by the set of tangent covectors $\{(d\varphi_1)_p,\dots,(d\varphi_n)_p\}$.

If v^* is any tangent covector, expressed as

$$v^* = \sum_{i=1}^n v_i^*(d\varphi_i)_p \ ,$$

the real numbers $v_1^*,\dots,v_n^*$ are such that

$$v_i^* = \langle v^*,(\frac{\partial}{\partial\varphi_i})_p\rangle$$

Note also that, if v is any tangent vector expressed as

$$v = \sum_{i=1}^n v_i(\frac{\partial}{\partial\varphi_i})_p$$

the real numbers $v_1,\dots,v_n$ are such that

$$v_i = \langle (d\varphi_i)_p,v\rangle .$$

6. Vector fields

Definition. Let N be a smooth manifold, of dimension n. A *vector field* f on N is a mapping assigning to each point $p \in N$ a tangent vector $f(p)$ in T_pN. A vector field f is *smooth* if for each $p \in N$ there exists a coordinate chart (U,φ) about p and n real-valued smooth functions $f_1,\dots,f_n$ defined on U, such that, for all $q \in U$,

$$f(q) = \sum_{i=1}^{n} f_i(q)\left(\frac{\partial}{\partial \varphi_i}\right)_q$$

Remark. Because of C^{∞}-compatibility of coordinate charts, given any coordinate chart (V,ψ) about p other than (U,φ), one may find a neighborhood $V' \subset V$ of p and n real-valued smooth functions $f'_1,\dots,f'_n$ defined on V', such that, for all $q \in V'$,

$$f(q) = \sum_{i=1}^{n} f'_i(q)\left(\frac{\partial}{\partial \psi_i}\right)_q$$

Thus, the notion of smooth vector field is independent of the coordinates used.

Remark. If (U,φ) is a coordinate chart of N, on the *submanifold* U of N one may define a special set of smooth vector fields, denoted $\frac{\partial}{\partial \varphi_1},\dots,\frac{\partial}{\partial \varphi_n}$, in the following way

$$\left(\frac{\partial}{\partial \varphi_i}\right): p \longmapsto \left(\frac{\partial}{\partial \varphi_i}\right)_p$$

It must be stressed, however, that such a set of vector fields is an object defined only on U. □

For any fixed coordinate chart (U,φ), the set of tangent vectors $\{(\frac{\partial}{\partial \varphi_1})_q,\dots,(\frac{\partial}{\partial \varphi_n})_q\}$ is a basis of T_qN at each $q \in U$, and therefore there is a *unique* set of smooth functions $\{f_1,\dots,f_n\}$ that makes it possible to express the value of a vector field f at q in the form

$$f(q) = \sum_{i=1}^{n} f_i(q)\left(\frac{\partial}{\partial \varphi_i}\right)_q$$

Expressing each f_i in local coordinates, as

$$\hat{f}_i = f_i \circ \varphi^{-1}$$

provides an expression in local coordinates of the vector field f itself. So, if p is a point of coordinates $(x_1,\dots,x_n)$ in the chart (U,φ), f(p) is a tangent vector of coefficients $(\hat{f}_1(x_1,\dots,x_n),\dots,\hat{f}_n(x_1,\dots,x_1))$ in the natural basis $\{(\frac{\partial}{\partial \varphi_1})_p,\dots,(\frac{\partial}{\partial \varphi_n})_p\}$ of T_pN

induced by (U,φ). Most of the times, whenever possible, the symbol f_i replaces $f_i \circ \varphi^{-1}$ and the expression of f in local coordinates is given a form of an n-vector $f = \mathrm{col}(f_1,\dots,f_n)$.

Remark. Let f be a smooth vector field, (U,φ) and (V,ψ) two coordinate charts about p and $f(x) = f(x_1,\dots,x_n)$, $f'(y) = f'(y_1,\dots,y_n)$ the corresponding expressions of f in local coordinates. Then

$$f'(y) = \left(\frac{\partial y}{\partial x} f(x)\right)_{x=x(y)} \qquad \square$$

The notion of vector field makes it possible to introduce the concept of a *differential equation on a manifold* N. For, let f be a smooth vector field. A smooth curve $\sigma:(t_1,t_2) \to N$ is an *integral curve* of f if

$$\sigma_*\left(\frac{d}{dt}\right)_t = f(\sigma(t))$$

for all $t \in (t_1,t_2)$. The left-hand-side is a tangent vector to the submanifold $\sigma((t_1,t_2))$ at the point $\sigma(t)$; the right-hand-side is a tangent vector to N at $\sigma(t)$. As usual, we identify the tangent space to a submanifold of N at a point with a subspace of the tangent space to N at this point.

In local coordinates, $\sigma(t)$ is expressed as an n-tuple $(\sigma_1(t),\dots,\sigma_n(t))$, and $f(\sigma(t))$ as

$$f(\sigma(t)) = \sum_{i=1}^{n} f_i(\sigma_1(t),\dots,\sigma_n(t))\left(\frac{\partial}{\partial \varphi_i}\right)_{\sigma(t)}$$

Moreover

$$\sigma_*\left(\frac{d}{dt}\right)_t = \sum_{i=1}^{n} \frac{d\sigma_i}{dt}\left(\frac{\partial}{\partial \varphi_i}\right)_{\sigma(t)}$$

Therefore, the expression of σ in local coordinates is such that

$$\frac{d\sigma_i}{dt} = f_i(\sigma_1(t),\dots,\sigma_n(t))$$

for all $1 \le i \le n$. This shows that the notion of integral curve of a vector field corresponds to the notion of solution of a set of n ordinary differential equations of the first order.

For this reason, one often uses the notation

$$\dot{\sigma}(t) = \sigma_*(\frac{d}{dt})_t$$

to indicate the image of $(\frac{d}{dt})_t$ under the differential σ_* at t.

The following theorem contains all relevant informations about the properties of integral curves of vector fields.

Theorem. Let f be a smooth vector field on a manifold N. For each $p \in N$, there exist an open interval - depending on p and written I_p - of $\mathbb{R}$ such that $0 \in I_p$ and a smooth mapping

$$\Phi : W \to N$$

defined on the subset W of $\mathbb{R}\times N$

$$W = \{(t,p) \in \mathbb{R}\times N:\ t \in I_p\}$$

with the following properties:

(i) $\Phi(0,p) = p$,

(ii) for each p the mapping $\sigma_p : I_p \to N$ defined by

$$\sigma_p(t) = \Phi(t,p)$$

is an integral curve of f,

(iii) if $\mu : (t_1,t_2) \to N$ is another integral curve of f satisfying the condition $\mu(0) = p$, then $(t_1,t_2) \subset I_p$ and the restriction of σ_p to (t_1,t_2) coincides with μ,

(iv) $\Phi(s,\Phi(t,p)) = \Phi(s+t,p)$ whenever both sides are defined,

(v) whenever $\Phi(t,p)$ is defined, there exists an open neighborhood U of p such that the mapping $\Phi_t : U \to N$ defined by

$$\Phi_t(q) = \Phi(t,q)$$

is a diffeomorphism onto its image, and

$$\Phi_t^{-1} = \Phi_{-t}$$

Remark. Properties (i) and (ii) say that σ_p is an integral curve of f passing through p at t = 0. Property (iii) says that this curve is unique and that the domain I_p on which σ_p defined is maximal. Property (iv) and (v) say that the family of mappings $\{\Phi_t\}$ is a one-parameter (namely, the parameter t) group of local diffeomorphisms, under the operation of composition. □

Example. Let $N = \mathbb{R}$ and use x to denote a point in N. Consider the vector field

$$f(x) = (x^2+1)(\frac{\partial}{\partial x})_x$$

An integral curve σ of f must be such that

$$\dot{\sigma}(t) = (\frac{d\sigma}{dt})(\frac{\partial}{\partial x})_x = (\sigma^2(t)+1)(\frac{\partial}{\partial x})_x$$

so

$$\frac{d\sigma}{dt} = \sigma^2 + 1$$

A solution of this equation has the form

$$\sigma(t) = tg(t+tg^{-1}(x^o))$$

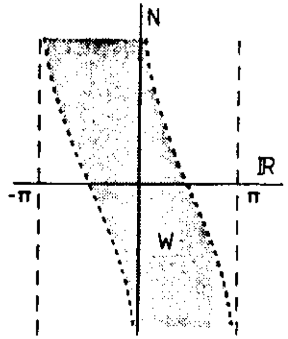
with x^o being indeed the value of σ at t = 0. Clearly, for each x^o the solution is defined for

$$-\frac{\pi}{2} < t+tg^{-1}(x^o) < \frac{\pi}{2}$$

Thus W is the set

$$W = \{(t,x^o) : t \in (-\frac{\pi}{2} - tg^{-1}(x^o),\ \frac{\pi}{2} - tg^{-1}(x^o))\}$$

which has the form indicated below. □



The mapping Φ is called the *flow* of f. Often, for practical purposes, the notation Φ_t replaces Φ, with the understanding that t is a variable. To stress the dependence on f, sometimes Φ_t is written as Φ_t^f.

Definition. A vector field f is *complete* if, for all $p \in N$, the interval I_p coincides with $\mathbb{R}$, i.e. - in other words - if the flow Φ of f is defined on the whole cartesian product $\mathbb{R}\times N$. □

The integral curves of a complete vector field are thus defined, whatever the initial point p is, for all $t \in \mathbb{R}$.

Definition. Let f be a smooth vector field on N and λ a smooth real-valued function on N. The *Lie derivative of* λ *along* f is a function $N \to \mathbb{R}$, written $L_f\lambda$ and defined by

$$(L_f\lambda)(p) = (f(p))(\lambda)$$

(i.e. $(L_f\lambda)(p)$ is the value at λ of the tangent *vector* f(p) at p). □

The function $L_f\lambda$ is a smooth function. In local coordinates, $L_f\lambda$ is represented by

$$(L_f\lambda)(x_1,\dots,x_n) = \left(\frac{\partial\lambda}{\partial x_1} \ \cdots \ \frac{\partial\lambda}{\partial x_n}\right)\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

If f_1, f_2 are vector fields and λ a real-valued function, we denote

$$L_{f_1}L_{f_2}\lambda = L_{f_1}(L_{f_2}\lambda)$$

The set of all smooth vector fields on a manifold N is denoted by the symbol V(N). This set is a *vector space* over $\mathbb{R}$ since if f, g are vector fields and a,b are real numbers, their linear combination af+bg is a vector field defined by

$$(af+bg)(p) = af(p)+bg(p)$$

If a,b are smooth real-valued functions on N, one may still define a linear combination af+bg by

$$(af+bg)(p) = a(p)f(p)+b(p)f(p)$$

and this gives V(N) the structure of a *module* over the ring, denoted $C^\infty(N)$, of all smooth real-valued functions defined on N. The set V(N) can be given, however, a more interesting algebraic structure in this way.

Definition. A vector space V over $\mathbb{R}$ is a *Lie algebra* if in addition to its vector space structure it is possible to define a binary operation $V\times V \to V$, called a product and written $[\cdot,\cdot]$, which has the following properties

(i) it is skew commutative, i.e.

$$[v,w] = -[w,v]$$

(ii) it is bilinear over $\mathbb{R}$, i.e.

$$[\alpha_1 v_1+\alpha_2 v_2, w] = \alpha_1[v_1,w]+\alpha_2[v_2,w]$$

(where α_1, α_2 are real numbers)

(iii) it satisfies the so called *Jacobi identity*, i.e.

$$[v,[w,z]]+[w,[z,v]]+[z,[v,w]] = 0. \quad □$$

The set V(N) forms a Lie algebra with the vector space structure already discussed and a product $[\cdot,\cdot]$ defined in the following way. If f and g are vector fields, [f,g] is a new vector field whose value at p, a tangent vector in T_pN, maps $C^\infty(p)$ into $\mathbb{R}$ according to the rule

$$([f,g](p))(\lambda) = (L_fL_g\lambda)(p)-(L_gL_f\lambda)(p)$$

In other words, $[f,g](p)$ takes λ into the real number $(L_fL_g\lambda)(p)-$ $+(L_gL_f\lambda)(p)$. Note that one may write more simply

$$L_{[f,g]}\lambda = L_fL_g\lambda-L_gL_f\lambda$$

Theorem. V(M) with the product [f,g] thus defined is a Lie algebra. □

The product [f,g] is called the *Lie bracket* of the two vector fields f and g.

Remark. If f,g are smooth vector fields and λ,γ smooth real-valued functions, then

$$[\lambda f,\gamma g] = \lambda\cdot\gamma\cdot[f,g] + \lambda\cdot L_f\gamma\cdot g - \gamma\cdot L_g\lambda\cdot f$$

Note that $\lambda,\gamma,L_f\gamma,L_g\lambda$ are elements of $C^\infty(N)$, and g,f,[f,g] elements of V(N). □

The reader may easily find that the expression of [f,g] in local coordinates is given by the n-vector

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \cdot & \cdots & \cdot \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{pmatrix}\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} - \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \cdot & \cdots & \cdot \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}\begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix} = \frac{\partial g}{\partial x} f - \frac{\partial f}{\partial x} g$$

If, in particular, $N = \mathbb{R}^n$ and

$$f(x) = Ax \ , \ g(x) = Bx$$

then

$$[f,g](x) = (BA-AB)x$$

The matrix $[A,B] = (BA-AB)$ is called the *commutator* of A,B.

The importance of the notion of Lie bracket of vector fields is very much related to its applications in the study of nonlinear control systems. For the moment, we give hereafter two interesting properties.

Theorem. Let N' be an embedded submanifold of N. Let U' be an open set of N' and f,g two smooth vector fields of N such that for all $p \in U'$

$$f(p) \in T_pN' \quad \text{and} \quad g(p) \in T_pN'.$$

Then also

$$[f,g](p) \in T_pN'$$

for all $p \in U'$. □

In other words, the Lie bracket of two vector fields "tangent" to a fixed submanifold is still tangent to that submanifold.

Theorem. Let f,g be two smooth vector fields on N. Let Φ_t^f denote the flow of f. For each $p \in N$.

$$\lim_{t\to 0} \frac{1}{t}[(\Phi_{-t}^f)_* g(\Phi_t^f(p)) - g(p)] = [f,g](p)$$

Remark. The first term of the expression under bracket is a tangent vector at p, obtained in the following way. With p, the mapping Φ_t^f (always defined for sufficiently small t) associates a point $q=\Phi_t^f(p)$. The vector field g is evaluated at q, and the value $g(q) \in T_qN$ is taken back to T_pN via the differential $(\Phi_{-t}^f)_*$ (which maps the tangent space at q onto the tangent space at $p=\Phi_{-t}^f(q)$). Thus, the mapping $p \mapsto (\Phi_{-t}^f)_* g(\Phi_t^f(p))$ defines a vector field, on the domain of Φ_t^f.

Remark. Let f be a smooth vector field on N, g a smooth vector field on M and $F : N \to M$ a smooth function. The vector fields f,g are said to

be *F-related* if

$$F_* f = g \circ F$$

Note that the vector field $(\Phi_{-t}^f)_* g(\Phi_t^f(p))$ considered in the above Remark is Φ_{-t}^f-related to g.

Remark. If $\bar{f}$ is F-related to f and $\bar{g}$ is F-related to g, then $[\bar{f},\bar{g}]$ is F-related to $[f,g]$.

Definition. Let f,g be two smooth vector fields on N. The *Lie derivative of* g *along* f is a vector field on N, written $L_f g$ and defined by

$$(L_f g)(p) = \lim_{t\to 0} \frac{1}{t}[(\Phi_{-t}^f)_* g(\Phi_t^f(p)) - g(p)] . \ \square$$

Thus, by definition, the Lie derivative $L_f g$ of g along f coincides with the Lie bracket $[f,g]$. There is also a third notation often used, which expreses the Lie derivative of g along f as

$$L_f g = ad_f g$$

Both notations may be used recurrently, taking

$$L_f^0 g = g \quad \text{and} \quad L_f^k g = L_f(L_f^{k-1} g)$$

or

$$ad_f^0 g = g \quad \text{and} \quad ad_f^k g = ad_f(ad_f^{k-1} g)$$

Remark. The Lie derivative of g along f may be interpreted as the value at t = 0 of the derivative with respect to t of a function defined as

$$W(t) = (\Phi_{-t}^f)_* g(\Phi_t^f(p))$$

Moreover, it is easily seen that for any $k \geq 0$

$$\left(\frac{d^k W(t)}{dt^k}\right)_{t=0} = L_f^k g(p) = ad_f^k g(p)$$

If W(t) is analytic in a neighborhood of t = 0, then W(t) can be expanded in the form

$$W(t) = \sum_{k=0}^{\infty} ad_f^k g(p) \frac{t^k}{k!}$$

known as *Campbell-Baker-Hausdorff formula*. □

One may define an object which dualizes the notion of a vector field.

Definition. Let N be a smooth manifold of dimension n. A *covector field* (also called *one-form*) ω on N is a mapping assigning to each point $p \in N$ a tangent covector $\omega(p)$ in T_p^*N. A covector field f is *smooth* if for each $p \in N$ there exists a coordinate chart (U,φ) about p and n real-valued smooth functions $\omega_1,\dots,\omega_n$ defined on U, such that, for all $q \in U$

$$\omega(q) = \sum_{i=1}^{n} \omega_i(q)(d\varphi_i)_q \qquad \square$$

The notion of smooth covector field is clearly independent of the coordinate used. The expression of a covector field in local coordinates is often given the form of a row vector $\omega = \text{row}(\omega_1,\dots,\omega_n)$ in which the ω_i's are real-valued functions of $x_1,\dots,x_n$.

If ω is a covector field and f is a vector field, $\langle \omega, f \rangle$ denotes the smooth real-valued function defined by

$$\langle \omega, f \rangle (p) = \langle \omega(p), f(p) \rangle$$

With any smooth function $\lambda: N \to \mathbb{R}$ one may associate a covector field by taking at each p the cotangent vector $(d\lambda)_p$. The covector field thus defined is usually still represented by the symbol $d\lambda$. However, the converse is *not always true*.

Definition. A covector field ω is *exact* if there exists a smooth real valued function $\lambda: N \to \mathbb{R}$ such that

$$\omega = d\lambda \qquad \square$$

The set of all smooth covector fields on a manifold N is denoted by the symbol $V^*(N)$.

One may also define the notion of Lie derivative of a covector field ω along a vector field f. In order to do this, one has to introduce first the notion of a covector field Φ_t^f-*related* to a given covector field ω. Let p be a point of the domain of Φ_t^f. Recall that $(\Phi_t^f)_*: T_pN \to T_{\Phi_t^f(p)}N$ is a linear mapping and let $(\Phi_t^f)^*: T^*_{\Phi_t^f(p)}N \to T_p^*N$ denote the dual mapping. With ω and Φ_t^f we associate a new covector field, whose value at a point p in the domain of Φ_t^f is defined by

$$(\Phi_t^f)^* \omega(\Phi_t^f(p))$$

The covector field thus defined is said to be Φ_t^f-related to ω.

Theorem. Let f be a smooth vector field and ω a smooth covector field on N. For each $p \in N$ the limit

$$\lim_{t \to 0} \frac{1}{t}[(\Phi_t^f)^* \omega(\Phi_t^f(p)) - \omega(p)]$$

exists.

Definition. The *Lie derivative of* ω *along* f is a covector field on N, written $L_f\omega$, whose value at p is set equal to the value of the limit

$$\lim_{t \to 0} \frac{1}{t}[(\Phi_t^f)^* \omega(\Phi_t^f(p)) - \omega(p)] \qquad \square$$

The expression of $L_f\omega$ in local coordinates is given by the (row) n-vector

$$(f_1 \dots f_n)\begin{pmatrix} \frac{\partial \omega_1}{\partial x_1} & \dots & \frac{\partial \omega_n}{\partial x_1} \\ \cdot & \dots & \cdot \\ \frac{\partial \omega_1}{\partial x_n} & \dots & \frac{\partial \omega_n}{\partial x_n} \end{pmatrix} + (\omega_1 \dots \omega_n)\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \cdot & \dots & \cdot \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} = \left(\frac{\partial \omega^T}{\partial x} f\right)^T + \omega \frac{\partial f}{\partial x}$$

where the superscript "T" denotes "transpose".

Remark. The three types of Lie derivatives $L_f\lambda, L_fg, L_f\omega$ defined above are related by the following Leibnitz-type relation

$$L_f\langle \omega, g \rangle = \langle L_f\omega, g \rangle + \langle \omega, L_fg \rangle$$

Remark. If ω is an exact covector field, i.e. if $\omega = d\lambda$ for some λ,

$$\langle d\lambda, f \rangle = L_f\lambda$$

and

$$L_f d\lambda = d(L_f\lambda)$$

Remark. If ω is a covector field, f a vector field, λ and γ real-valued functions, then

$$L_{\lambda f}\gamma\omega = \lambda\cdot\gamma\cdot L_f\omega + \lambda\cdot L_f\gamma\cdot\omega + \gamma\,\langle\omega, f\rangle\, d\lambda$$

Note that $\lambda, \gamma, L_f\gamma, \langle\omega, f\rangle$ are elements of $C^\infty(N)$ and $\omega, L_f\omega, d\backslash$ elements of $V^*(N)$.