# Differential Games Of Multiple Agents and Geometrical Structures

P. M. Pardalos[1], V. A. Yatsenko[2], A. Chinchuluun[3], and A. G. Nahapetyan[4]

[1] Center for Applied Optimization, Department of Industrial and Systems Engineering, University of Florida,Gainesville, FL 32611, USA pardalos@ufl.edu
[2] Center for Applied Optimization, Department of Industrial and Systems Engineering, University of Florida,Gainesville, FL 32611, USA yatsenko@ufl.edu
[3] Center for Applied Optimization, Department of Industrial and Systems Engineering, University of Florida,Gainesville, FL 32611, USA altannar@ufl.edu
[4] Center for Applied Optimization, Department of Industrial and Systems Engineering, University of Florida,Gainesville, FL 32611, USA artyom@ufl.edu

**Summary.** This chapter deals with problems of differential games of multiple agents moving in a region. We describe such a game by a hierarchical structure, which can be simplified using a fiber bundle. Then, using geometrical techniques, we study controllability, observability, and optimality problems. In addition we also consider a cooperative problem when the agent's motions must satisfy a separation constraint throughout the encounter to be conflict-free. A classification of maneuvers based on different commutative diagrams is introduced using their fiber bundles representation. In the case of two agents, these optimality conditions allow us to construct the optimal maneuvers geometrically.

## 1 Introduction

The modern game theory basically deals with dynamical systems on smooth manifolds. However, many practical systems like multiple agents do not have such structures. The axiomatic control theories should adequately reflect in terms of their internal language of notions and control problems (Cressman, 2003). In terms of these theories, the control structures can make up various hierarchies. According to Kalman, for example, the most general structure is represented by a controllability-reachability structure over which the optimal control structure is built. This approach regarding the structure of optimal control and Yang–Mills Fields was discussed in (Yatsenko, 1989; Butkovskiy, 1990) .

In this chapter, the geometrical description problem of multiple agents is studied. We discuss mathematical aspects of the "Unified game (UGT)" and "Theory of the control structures (TCS)". We consider a game as a hierarchical structure. It is assumed that each agent can be described by a fiber bundle. A joint maneuver has to be chosen to guide each agent from its starting position to its target position while avoiding conflicts. Among all the conflict-free joint maneuvers, we aim to determine the one with the least overall cost. The cost of an agents maneuver is its energy, and the overall cost is a weighted sum of the maneuver energies of all individual agents, where the weights represent priorities of the agents.

As an example, we consider the hierarchical structure of such multi-agent system on Fig. 1. Each agent of the system can be described by stochastic or deterministic differential equation with a control. In the paper we first reduce the model to a hierarchical geometric representation using fiber bundles. Then we consider an integrated geometrical model where the separated model of agents are integrated into single model. For example, the interaction between six robots on Figure 2 can be described by a hierarchical structure. The integrated model allows solving controllability, observability, and cooperative control problems.

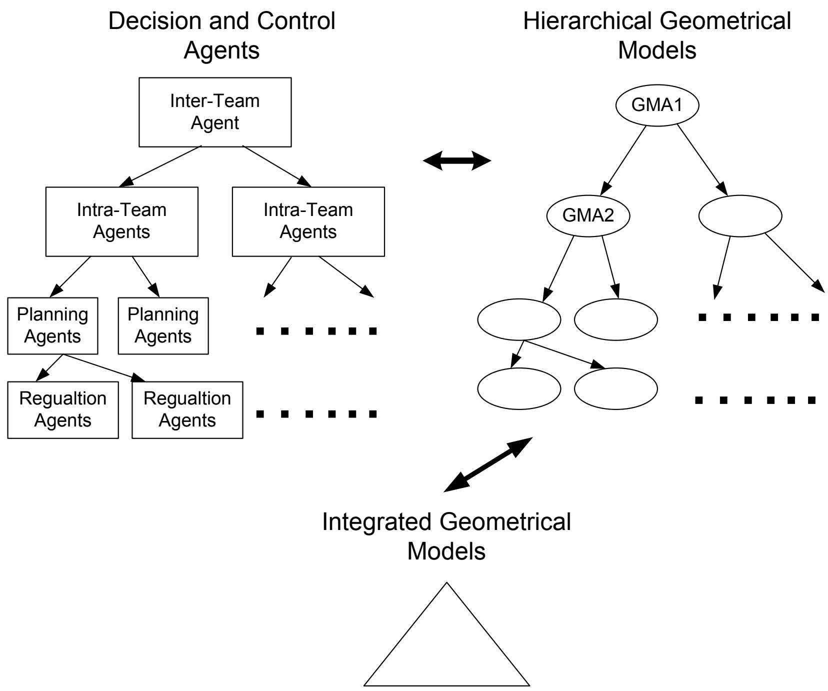



**Fig. 1.** Hierarchical structure of multiple agents.

In Section 2, we consider geometric aspects of the nonlinear control systems. The section constructs a formal model, where the optimal control structure appears independently from the controllability-reachability structure and that of the space of local system states. The efficiency of this axiomatic approach is illustrated using structural analysis of a general problem of the optimal control. In Section 3, we analyze in details the relationship between gauge fields, identification problems and control systems. The result of the analysis is an estimation algebra of a nonlinear estimation problem. The estimation algebra turns out to be a useful concept to explore finite-dimensional nonlinear filters. In Section 4, we consider a Lie group related to Yang–Mills gauge groups. We show that the estimation algebra of the identification problem is a subalgebra of the current algebra. Section 5 focuses on nonlinear control systems and Yang–Mills fields. Section 6 is devoted to geometric models of multiagent systems as controlled dynamical-information objects. It is shown that these systems can be described by commutative diagrams which allow to analyze a symmetry.

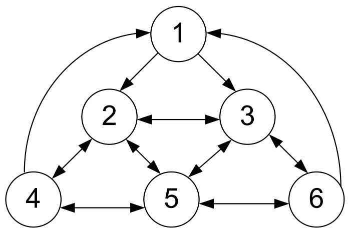



**Fig. 2.** Hierarchical structure of multiple robot.

## 2 Geometric Structures

We briefly describe the role of topological, metric and orderness structures. Note that each standard ordinary differential control system or inclusion $x' \in I(x)$, $x \in X$, generates two independent topological structures on $X$. One of them is generated by a family of inclusions of $x \in X$, i.e., the family of reachability sets $O(x_0, \varepsilon)$ from $x_0$ for time $\varepsilon \geq 0$, and another one by a family of a controllability area $O(x_0, \varepsilon)$ to $x_0$ for time $\varepsilon \geq 0$ (observability topology).

Let $(X, \tau)$ be a topological space, where $X$ is an abstract non-empty set and $\tau$ is a topology on $X$.

**Definition 1** *Control (or admissible control) $\gamma(a, b)$ in $(X, \tau)$ is an image of the continuous (in sense of topology $\tau$) map $\varphi$: $[0, 1] \to X$,*

$$x = \varphi(t), \quad 0 \leq t \leq T, \quad x \in X, \tag{1}$$

$$a = \varphi(0), \quad b = \varphi(T). \tag{2}$$

*$a \in X$ is an initial point and $b \in X$ is a final point of the control $\gamma(a,b)$.*

Thus, the control $\gamma(a,b)$ is pathwise connected and linearly ordered subset (sequence) of $X$ where $a \in \gamma(a,b)$ and $b \in \gamma(a,b)$ are the smallest and the largest of its elements, respectively. As results, maps (1) and (2) are admissible parameterizations of the control $\gamma(a,b)$.

The verification of Definition 1 consists of a validation of the controllability and finding optimal control of systems without using any differential or difference structure. Furthermore, we shall consider that there is a metric in topological spaces, which allows to analyze control problems at various levels of generality. We shall be looking for the "minimal" but not trivial structures, which can be responsible for controls.

### 2.1 Metric Spaces

The concepts of a metric and a metric space are introduced by the following definitions.

**Definition 2** *Metric space $(X,\rho)$ is a pair $(X,\rho)$ where $X$ is an arbitrary non-empty set and p is a metric structure of $X$, i.e., $\rho$ is a real valued function $\rho = \rho(x,y)$, $(x,y) \in X^2 = X \times X$, or map*

$$\rho : X^2 \to \mathbb{R} \tag{3}$$

*with the metric axioms:*

$$\rho(a,b) \geq 0 \quad \textit{for} \;\; \forall\, (a,b) \in X^2, \tag{4}$$

$$\rho(a,a) = 0 \quad \textit{for} \;\; \forall\, a \in X, \tag{5}$$

$$\rho(a,b) < \rho(a,c) + \rho(c,b) \tag{6}$$

*for any $a \in X$, $b \in X$, $c \in X$, and (6) is called "triangle inequality". Sometimes $\rho$ is also called a global metric on $X$ or distance in $X$.*

The metric introduced by Definition 2 differs from the usual concept of metric: there is neither the symmetry axiom ($\rho(a,b) = \rho(b,a)$ for $\forall\, a \in X$, $\forall\, b \in X$) nor the requirement: $\rho(a,b) > 0$ if $a \neq b$. So, the given concept of metric is more adequate to the situation in typical control problems. As is known the metric space $(X,\rho)$ can also be considered as a topological space $(X,\tau)$, where topology $T$ is born by metric $\rho$.

But the metric can measure control $\gamma(a,b)$ introduced by Definition 1. This can be done by the following definition.

**Definition 3** *The length $l[\gamma(a,b)]$ of the control $\gamma(a,b)$ is a real valued function*

$$l\left[\gamma(a,b)\right] = \lim_{N\to\infty} l_N\left[\gamma(a,b)\right], \tag{7}$$

*where*

$$l_N\left[\gamma(a,b)\right] = \sum_{i=0}^{N} \rho(x_i, x_{i+1}), \tag{8}$$

*where $a = x_0 < x_1 < \dots < x_N < x_{N+1} = b$ is the N-th partition $T_N$ of $\gamma(a,b)$, and the partition $T_N$ becomes a finer with $N \to \infty$. Of course, it is necessary to prove or admit the existence and uniqueness of (7). If so, $\gamma(a,b)$ is called measurable (in metric $\rho$). The set $\gamma(a,b)$ of all measurable $\gamma(a,b)$ is denoted by $\Gamma(a,b)$:*

$$\Gamma(a,b) = \{\gamma(a,b)\}. \tag{9}$$

So, in the metric space $(X,\rho)$ the *admissible control* $\gamma(a,b)$ is just a measurable (in sense of metric $\rho$) sequence and vice versa.

If we have several sequences in $X$:

$$\gamma_1(x_0,x_1), \gamma_2(x_1,x_2), \dots, \gamma_n(x_n,x_{n+1}) \tag{10}$$

then we can define their sum

$$\gamma(x_0,x_{n+1}) = \sum_{i=1}^{n} \gamma_i(x_{i-1},x_i) \tag{11}$$

which is also a sequence.

Inversely, if $\gamma(a,b)$ is a sequence and $x_i \in \gamma(a,b)$, $i = 1,\dots,n$, $x_1 < \dots < x_n$, then $\gamma(a,b)$ can be represented as the sum of sequences:

$$\gamma(a,b) = \gamma_1(a,x_1) + \gamma_2(x_1,x_2) + \dots + \gamma_n(x_n,b). \tag{12}$$

**Definition 4** *The sequence $\gamma_i(x_{i-1},x_i)$ in (11) is called a piece of the sequence $\gamma(a,b)$. We accept that functional (7) is additive one:*

$$\rho(a,b) \geq 0 \quad \textit{for} \ \ \forall\ (a,b) \in X^2, \tag{13}$$

$$\rho(a,a) = 0 \quad \textit{for} \ \ \forall\ a \in X. \tag{14}$$

### 2.2 Optimal Control

Consider the following problem of optimal control in $(X,\rho)$.

1. Determine
$$\bar{l}(a,b) = \{\inf l[\gamma(a,b)] : \ \gamma(a,b) \in \Gamma(a,b)\}. \tag{15}$$
2. Determine $\bar\gamma = \bar\gamma(a,b)$, if exists, such that
$$l[\bar\gamma(a,b)] = \bar{l}(a,b). \tag{16}$$
This admissible $\bar\gamma(a,b)$ will be called the minimal of the optimal control problem.
3. Describe all set $\{\bar\gamma(a,b)\}$ for fixed $(a,b) \in X^2$ and for all $(a,b) \in X^2$.

A simple but an important property of the minimal $\bar{\gamma}(a,b)$ is given by the following theorem.

**Theorem 1** *If the admissible $\gamma(c,d)$ is the minimal of the optimal control problem, then the sequence $\bar{\gamma}(a,b)$ is also minimal.*

This is a consequence of the additivity property of (12). If any admissible sequence $\gamma(a,b)$ is minimal, it does not mean that $\gamma(a,b)$ is also minimal.

It is easy to prove the inequality

$$\rho(a,b) \leq \bar{l}(a,b). \tag{17}$$

**Definition 5** *The metric space $(X,\rho)$ is an obstacleness metric space iff there exists at least one point $(a,b) \in X^2$ such that*

$$\rho(a,b) \leq \bar{l}(a,b).$$

*The metric space $(X,\rho)$ is a generalized metric space iff*

$$\rho(a,b) = \bar{l}(a,b) \quad \textit{for} \ \ \forall\ (a,b) \in X^2.$$

An example of a generalized space is an Euclidean space $\mathbb{R}^n$.

The following theorem is valid:

**Theorem 2** *$\bar{l} = \bar{l}(a,b)$ is also metric on $X$, and $(X,\bar{l})$ is also a metric space.*

**Definition 6** *Metric $\bar{l} = \bar{l}(a,b)$ is called a secondary metric.*

Generally, $\bar{l}(a,b)$ is distinguished from initial metric $\rho(a,b)$ on $X$.

**Definition 7** *If the secondary metric $\bar{l}$ coincides with the metric $\rho$ then $\rho$ is called a self-secondary metric.*

The following theorems are valid:

**Theorem 3** *The secondary metric is a self-secondary metric.*

This is similar to the property of projection operator $P : P^2 = P$.

**Theorem 4** *The metric space $(X,\rho)$ is a generalized space if the metric $\rho$ is the self-secondary metric.*

We illustrate the application of the above introduced concepts by the following:

**Theorem 5** *(Sufficient condition for minimal). The sequence $\gamma(a,b)$ in $(X,\rho)$ is minimal if for any of its admissible $\gamma(c,d)$ the next relation is true:*

$$\bar{l}(c,d) = \bar{l}(c,x) + \bar{l}(x,d) \quad \textit{for} \ \ \forall\ x \in \gamma(c,d), \tag{18}$$

*where $\bar{l}$ is a secondary metric of $\rho$.*

It might seem that for $\gamma(a,b)$ to be minimal just one identity is sufficient:

$$\bar{l}(a,b) = \bar{l}(a,x) + \bar{l}(x,b) \quad \text{for} \ \forall\, x \in \gamma(a,b). \tag{19}$$

But it is not true, there exists a contrary example.

From topology standpoint the secondary metric $\bar{l}$ generally is weaker (rougher) than "initial" or "first" metric $\rho$. In other words, topology $(X,\rho)$ is stronger (thinner) than secondary topology $(X,\bar{l})$.

## 3 Identification of Agents and Yang–Mills Fields

In this section we consider the models where each agents of the hierarchical system is described by a stochastic differential equation.

### 3.1 Stochastic Agents

Consider the stochastic differential system:

$$d\theta = 0, \tag{20}$$

$$dx_t = A(\theta)x_t dt + b(\theta)dw_t, \tag{21}$$

$$dy_t = \langle(\theta), x_t\rangle dt + dv_t. \tag{22}$$

Here $\{w_t\}$ and $\{v_t\}$ are independent, scalar and standard Wiener processes, and $\{x_t\}$ is an $\mathbb{R}^n$-valued process. Assume that $\theta$ takes values in a smooth manifold $\Theta \to \mathbb{R}^N$, and the map $\theta \to \Sigma(\theta) := (A(\theta), b(\theta), c(\theta))$ in a smooth map taking values in minimal triples. By the identification problem we shall mean the nonlinear filtering problem associated with equation (1); i.e. the problem of recursively computing conditional expectations of the form $\pi_t(\phi) \,\underline{\Delta}\, E[\phi(x_t,\theta)|Y_t]$, where $Y_t$ is the $\sigma$-algebra generated by the observations $\{y_s : 0 \le s \le t\}$ and $\phi$ belongs to a suitable class of functions on $\mathbb{R}^n \times \Theta$.

For given $y_t$, the joint unnormalized conditional density $\rho \,\underline{\Delta}\, \rho(t,x,\theta)$ of $x_t$ and $\theta$ satisfy the stochastic partial differential Stratonovitch equation

$$d\rho = A_0 \rho dt + B_0 \rho dy_t, \tag{23}$$

where the operators $A_0$ and $B_0$ are given by

$$A_0 := \frac{1}{2}\left\langle b(\theta), \frac{\partial}{\partial x}^{2}\right\rangle - \left\langle \frac{\partial}{\partial x}, A(\theta)x\right\rangle - \langle c(\theta), x\rangle^2/2, \tag{24}$$

$$B_0 := \langle c(\theta), x\rangle. \tag{25}$$

From the Bayes formula it follows that

$$\pi_t(\phi) = \sigma_t(\phi)/\sigma_t(l), \tag{26}$$

where

$$\sigma_t(\phi) = \int_{\Theta} \int_{\mathbb{R}^n} \phi(x, \theta)\rho(t, x, \theta)|dx||d\theta|, \tag{27}$$

where $|dx|$ and $|d\theta|$ are fixed volume elements on $\mathbb{R}^n$ and $\Theta$, respectively. Further if $Q(t, \theta)$ denotes the unnormalized posterior density of $\theta$ given $t$, then it satisfies the equation:

$$dQ = E[\langle c(\theta), x_t|\theta\rangle, Y_t]Q(t, \theta)dy_t. \tag{28}$$

The paper on nonlinear filtering theory (Hazewinkel, 1981) shows that it is natural to look at equation (23) formally as a deterministic partial differential equation,

$$\frac{\partial \rho}{\partial t} = A_0\rho + \dot{y}B_0\rho. \tag{29}$$

By the Lie algebra of the identification problem, we shall mean the operator Lie algebra $\bar{\mathrm{G}}$ generated by $A_0$ and $B_0$. For more general nonlinear filtering problems, estimation algebras analogous to $\bar{\mathrm{G}}$ have been emphasized by Brockett (Mitter, 1990) and others as being objects of central interest. In the papers (Krishnaprasad, 1981) the Lie algebra $\bar{\mathrm{G}}$ is used to classify identification problems and to understand the role of certain sufficient statistics.

### 3.2 The Estimation Algebra of Nonlinear Filtering Systems

To understand the structure of the estimation algebra it is well-worth considering an example.

**Example 1** Let $dx_t = \theta dw_t$; $d\theta = 0$; $dy_t = x_t dt + dv_t$. Then $A_0 = \frac{\theta^2}{2}\frac{\partial^2}{\partial t^2} - \frac{x^2}{2}$ and $B_0 = x$, and $\bar{\mathrm{G}} = \{A_0, B_0\}_{I..A}$ is spanned by the set of operators $\left(\frac{\theta^2}{2} - \frac{x^2}{2}\right)$, $\left(\theta^{2n}x\right)_{n=0}^{\infty}$, $\left(\theta^{2n}\frac{\partial}{\partial x}\right)_{n=1}^{\infty}$ and $\{\theta^{2n}1\}_{n=1}^{\infty}$. We then notice that,

$$\tilde{G} \subseteq \mathbb{R}\left[\theta^2\right] \otimes \left\{\frac{\partial^2}{\partial x^2}, x\frac{\partial}{\partial x}, \frac{\partial}{\partial x}, x^2, x, 1\right\} L.A.$$

is a subalgebra of the Lie algebra obtained by tensoring the polynomial ring $\mathbb{R}\left[\theta^2\right]$ with a 6 dimensional Lie algebra. Here, L.A. stands for the Lie algebra generated by the elements in the brackets.

The general situation is very much as in this example. Consider the vector space (over the reals) of operators spanned by the set,

$$S := \left\{\frac{\partial^2}{\partial x_i \partial x_j}, x_i\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_i}, x_i x_i, x_j, 1\right\}, \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, n. \tag{30}$$

This space of operators has the structure of a Lie algebra henceforth denoted as $\tilde{\mathrm{G}}_0$ (of dimension $3n^2 + 2n + l$) under operator commutation (the commutation rules being $\frac{\partial^2}{\partial x_i \partial x_j}$, $x_k = \delta jk \frac{\partial}{\partial x_i} + \delta ik \frac{\partial}{\partial x_j}$ etc., where $\delta jk$ denotes the Kronecker symbol). For each choice $\Theta$, $A_0$ and $B_0$ take values in $\tilde{\mathrm{G}}_0$. It follows that in general $A_0$ and $B_0$ are smooth maps from $\Theta$ into $\tilde{\mathrm{G}}_0$. Thus, let us consider the space of smooth maps $C^\infty(\Theta; \tilde{\mathrm{G}}_0)$. This space can be given by the structure of a Lie algebra (over the reals) in the following way:

$$\text{given } \varphi,\ \phi \varepsilon C^\infty(\Theta; \tilde{\mathrm{G}}_0),$$

define the Lie bracket $[\cdot, \cdot]_C$ on $C^\infty(\Theta; \tilde{\mathrm{G}}_0)$ by

$$[\phi, \psi]_C(P) = [\phi(P), \psi(P)] \quad \text{for every} \quad P \in \Theta. \tag{31}$$

Here the bracket on the right hand side of equation (31) is in $\tilde{\mathrm{G}}_0$. We denote as $\tilde{\mathrm{G}}_0$ the Lie algebra $(C^\infty(\Theta; \tilde{\mathrm{G}}_0); [.,.]_C)$. Whenever the dimension of $\Theta$ is greater than zero, $\tilde{\mathrm{G}}_0)$ is infinite dimensional and is an example of a *current algebra*. Current algebras play a fundamental role in the physics of Yang–Mills fields where they occur as Lie algebras of gauge transformations. Elsewhere in mathematics they are studied under the guise of local Lie algebras. The following is immediate.

**Proposition 1** *The Lie algebra* $\tilde{\mathrm{G}})$ *of operators generated by*

$$A_0 := \frac{1}{2}\left\langle b(\theta), \frac{\partial}{\partial x}\right\rangle^2 - \left\langle \frac{\partial}{\partial x}, A(\theta)x\right\rangle - \langle c(\theta), x\rangle^2/2 \tag{32}$$

*and* $B_0 := \langle c(\theta), x\rangle$*, is a subalgebra of the current algebra* $C^\infty(\Theta; \tilde{\mathrm{G}}_0)$.

### 3.3 Estimation Algebra and Identification Problem

It is known (Markus, 1976) that $\tilde{\mathrm{G}}$ admits a faithful representation as a Lie algebra of vector fields on a finite dimensional manifold. Specifically, consider the system of equations,

$$\begin{aligned} d\theta &= 0, \\ dz &= \left[A(\theta) - Pc(\theta)c^T(\theta)\right] z dt + Pc(\theta) dy_t, \\ \frac{dP}{dt} &= A(\theta)P + PA^T(\theta) + b(\theta)b^T(\theta) - Pc(\theta)c^T(\theta)P, \\ ds &= \frac{1}{2}\langle c(\theta), z\rangle^2 dt - \langle c(\theta), z\rangle dy_t. \end{aligned} \tag{33}$$

The system of equations (33) evolves on the product manifold $\Theta \times \mathbb{R}^{n(n+3)/2+1}$. Associated with equations (33) there are the pair of vector fields (first order differential operators),

$$a_0^* = \langle (A(\theta) - Pc(\theta)c^T(\theta))z, \partial/\partial z\rangle +$$
$$+ \operatorname{tr}((A(\theta)P + PA^T(\theta) + b(\theta)b^T(\theta) - Pc(\theta)c^T(\theta)P), \partial/\partial P) +$$
$$+ 1/2\langle c(\theta), z\rangle^2 \partial/\partial s$$

and

$$b_0^* = \langle P(\theta), \partial/\partial z\rangle - \langle c(\theta), z\rangle \partial/\partial z.$$

Here $\partial/\partial P = [\partial/\partial P_{ij}] = (\partial/\partial P)^T = n \times n$ symmetric matrix of differential operators. Consider the Lie algebra of vector fields generated by $a_0^*$ and $b_0^*$. Since $a_0^*$ and $b_0^*$ are vertical vector fields with respect to the fibering $\Theta \times \mathbb{R}^{n(n+3)/2+1} \to \Theta$, then every vector field is in this Lie algebra. One of the main results is the following (Markus, 1980):

**Theorem 6** *The map*

$$\Phi_k : \tilde{\mathrm{G}}_0 \to \bigcup \Theta \times \mathbb{R}^{n(n+3)/2+1}$$

*defined by*

$$b_0^* = \langle P(\theta), \partial/\partial z\rangle 1/2\langle c(\theta), z\rangle \partial/\partial s$$

*is a faithful representation of the Lie algebra of the identification problem as a Lie algebra of (vertical) vector fields on a finite dimensional manifold fibered over $\Theta$.*

**Example 2** To illustrate Theorem 5, consider the Lie algebra of Example 1. The embedding equations (33) take the form

$$\begin{aligned} d\theta &= 0,\\ dp &= \left(\theta^2 - p^2\right) dt,\\ dz &= -pzdt + pdy_t,\\ ds &= z^2/2dt - zdy_t. \end{aligned}$$

Then

$$\Phi_k(B_0) = \Phi_k(x) = b_0^* = p\frac{\partial}{\partial z} + (-z)\frac{\partial}{\partial s}.$$

The induced maps on Lie brackets are given by

$$\begin{aligned} \Phi_k(\theta^{2k}\partial/\partial z) &= \theta^{2k}\partial/\partial z, \quad k = 0, 1, 2, \dots,\\ \Phi_k(\theta^{2k}x) &= \theta^{2k}(p\partial/\partial z - z\partial/\partial s), \quad k = 1, 2, \dots,\\ \Phi_k(\theta^{2k}l) &= \theta^{2k}\partial/\partial s, \quad k = 1, 2, \dots. \end{aligned}$$

The embedding equations have the following statistical interpretation. Assume that the initial condition for (12) is of the form

$$\rho_0(x,\theta) = \Big(2\pi \det \sum(\theta)\Big)^{-n/2} \times$$
$$\times \exp\left(-\left\langle x-\mu(\theta), \sum^{-1}(\theta)(x-\mu(\theta))\right\rangle\right)\cdot Q_\theta,$$

where $\theta \to (\mu(\theta), \Sigma(\theta), Q_0(\theta))$ is a smooth map, $\sum(\theta) > 0$, $\theta \in \Theta$ and $Q_0 > 0$ for $\theta \in \Theta$. Suppose equation (11) is initialized at,

$$(\theta_0, z_0, P_0, s_0) = \Big(\theta_0, \mu(\theta_0), \sum(\theta_0), -\log(Q_0(\theta))\Big) \tag{34}$$

Append to the system (11) an output equation,

$$\bar{Q}_t = e^{-s_t}. \tag{35}$$

Now if (11) is solved with initial condition (14) one can show by differentiating $\bar{Q}_t$ that $\bar{Q}_t$ satisfies the equation (7). In other words, the system (11)–(15) with initial condition (14) is a finite dimensional recursive estimation for the posterior density $Q(t,\theta_0)$. We have thus verified the homomorphism principle of Brockett (Brockett, 1979): that finite dimensional recursive estimators must involve Lie algebras of vector fields that are homomorphic images of the Lie algebra of operators associated with the unnormalized conditional density equation.

## 4 Sobolev Lie Group and Yang–Mills Fields

It has been remarked elsewhere that the Cauchy problem associated with (8) may be viewed as a problem of integrating a Lie algebra representation. In this connection one should be interested whether there is an appropriate topological group associated with $\tilde{G}$. We have the following general procedure.

Let $M$ be a compact Riemannian manifold of dimension $d$. Let $L$ be a Lie algebra of dimension $n < \infty$. We can always view $L$ as a subalgebra of the general linear Lie algebra $g\ell(m;\mathbb{R}), m > n$ (Ado's theorem).

**Assumption 1** *Let $G = \{\exp(L)\}_G \subset g\ell(m;\mathbb{R})$ be the smallest Lie group containing the exponentials of elements of $L$. We assume that $G$ is a closed subset of $g\ell(m;\mathbb{R})$.*

Define,

$$\mathcal{R} = \mathrm{C}^\infty(\mathrm{M}; g\ell(\mathrm{m};\mathbb{R})),$$
$$\mathcal{L} = \mathrm{C}^\infty(\mathrm{M}; \mathrm{L}),$$
$$\mathcal{D} = \mathrm{C}^\infty(\mathrm{M}; \mathrm{G}).$$

Clearly $\mathcal{R}$ is an algebra under pointwise multiplication and

$$\mathcal{L} \subset \mathcal{R}, \quad \mathcal{D} \subset \mathcal{R}$$

Let $(U\alpha, \varphi_\alpha)$ be a $C^\infty$ atlas for M. Then for a $f_1, f_2 \in \mathcal{R}$, define

$$\| f_1 - f_2 \| = \left[ \int_{\varphi_\alpha(U_\alpha)} d\text{vol} \sum_{\ell=0}^{k} \left| D^\ell (f_1 - f_2) \varphi_\alpha^{-1} \right|^2 \right]^{1/2}, \tag{36}$$

where

$$|f|^2 = \text{tr}\,(f' f). \tag{37}$$

(Here $k = d/2 + s$, $s > 0$). Let $\mathcal{R}_k$ be the completion of $\mathcal{R}$ and $\mathcal{D}_k$, the completion of $\mathcal{D}$ in the norm $\| \cdots \|_k$ ($\mathcal{D}_k$ is closed in $\mathcal{R}_k$). By the Sobolev theorem, $\mathcal{R}_k$ is a Banach algebra and the group operation

$$\begin{aligned} &\mathcal{D}_k \times \mathcal{D}_k \to \mathcal{D}_k, \\ &(f_1, f_2) \to f_1 f_2 \end{aligned} \tag{38}$$

when $(f_1 f_2)(m) = f_1(m) f_2(m)$ is continuous. Thus $\mathcal{D}_k$, is a topological group.

By proceeding as before, one can give a Sobolev completion of $\mathcal{L}$ to obtain $\mathcal{L}_k$, an infinite dimensional Lie algebra, where once again by the Sobolev theorem the bracket operation

$$\begin{aligned} &[.,.]\mathcal{L}_\mathrm{k} \times C_\mathrm{k} \to \mathcal{L}_\mathrm{k}, \\ &(f_1, f_2) \to [f_1, f_2] \end{aligned}$$

with $[f_1, f_2](m) = [f_1(m), f_2(m)]$ is continuous. Now, for a small enough neighborhood $V(0)$ of $0 \in \mathcal{L}$, one can define

$$\begin{aligned} \exp : V(0) &\to \mathcal{D}_\mathrm{k}, \\ \xi &\to \exp(\xi) \end{aligned}$$

by pointwise exponentiation. This permits us to provide a Lie group structure on $\mathcal{D}_\mathrm{k}$ with $\mathcal{L}_\mathrm{k}$ canonically identified as the Lie algebra of $\mathcal{D}_\mathrm{k}$.

The procedure outlined above appears to play a significant role in several contexts (the index theorem Yang–Mills fields (Milter, 1980, 1981)).

For our purposes $\mathcal{L}$ will be identified with a faithful matrix representation of $\tilde{G}_0$. Thus we associate with the identification problem a Sobolev Lie group, which is a subgroup of $\mathcal{D}_\mathrm{k}$ corresponding to $\tilde{G}_0$

**Remark 1** One of the important differences between the problem of filtering and the problems of Yang–Mills theories is that in the latter case there are natural norms for Sobolev completion. This follows from the fact that in Yang–Mills theories the algebra $\mathcal{L}$ is compact (semi-simple) and one has the Killing form to work with. In filtering problems, $\tilde{G}_0$ is never compact.

We use a representation of the form

$$\rho(t,x,\theta)=\exp(g_1(t,\theta)\mathrm{A}^1)\ldots\exp(g_n(t,\theta)A^n)\rho_0 \tag{39}$$

for the solution to the equation (8). In the case of example (1), this takes the form

$$\rho(t,x\theta)=\exp\left(g_1(t,\theta)\left(\frac{\theta^2}{2}\frac{\theta^2}{\theta_x}-\frac{x^2}{2}\right)\right)\exp\left(g_2(t,\theta)\theta^2\frac{\partial}{\partial x}\right)\times$$
$$\times\exp\left(g_3(t,\theta)x\right)\exp\left(g_4(t,\theta)l\right)\rho_0.$$

Differentiating and substituting in (29), we can obtain

$$\begin{aligned}\frac{\partial g}{\partial t}(t,\theta)&=1,\\ \frac{\partial g_2}{\partial t}(t,\theta)&=\cosh(g_1,\theta)\dot{y},\\ \frac{\partial g_3}{\partial t}&=-\frac{1}{\theta}\sinh(g_1,\theta)\dot{y},\\ \frac{\partial g_4}{\partial t}&=\frac{\partial q_3}{\partial t}(t,\theta)g_2(t,\theta)\end{aligned} \tag{40}$$

and $g_i(0,\theta)=0$ for $i=1,2,3,4$, $\theta\in\mathit{\Theta}$. The above first-order partial differential equations may be easily solved by quadrature and one has the representation

$$\begin{aligned}\rho(t,x,\theta)=&\int_{-\infty}^{\infty}\sqrt{\frac{1}{2\pi\sinh(|\theta|t)}}\exp\left(-\frac{1}{2}\coth^2\left(\frac{|x|}{|\theta|}^2+z\right)t|\theta|\right)\times\\ &\times\exp\left(\frac{xz}{\sqrt{|\theta|\sinh(|\theta|t)}}\right)\exp\left(g_4\left(t,\theta\right)\theta^2\right)\times\\ &\times\exp\left(g_2(t,\theta)\sqrt{|\theta|z}\right)\rho_0\left(g_3(t,\theta)\theta^2\sqrt{|\theta|z},\theta\right)dz,\end{aligned} \tag{41}$$

where $\rho_0(,\theta)\in L_2(\mathbb{R})$ for every $\theta\in\mathit{\Theta}$ and is smooth in $\theta$. Further $\mathit{\Theta}\mathbb{R}$ is a bounded set and 0 closure $\mathit{\Theta}$.

In equation (39), $g_1$ should be viewed as canonical coordinates of the second kind on the corresponding Sobolev Lie group. Now expand $g_2$, and $g_3$ to obtain

$$\begin{aligned}g_2(t,\theta)&=\sum_{k=0}^{\infty}\theta^2\int_0^t\frac{\sigma^{2k}}{(2k)!}\dot{y}_\sigma d\sigma,\quad k=1,2,\ldots,\\ g_3(t,\theta)&=-\sum_{k=0}^{\infty}\theta^{2k}\int_0^t\frac{\sigma^{2k+1}}{(2k+1)!}\dot{y}_\sigma d\sigma,\quad k=1,2,\ldots.\end{aligned} \tag{42}$$

It follows that all the "information" contained by the observations $\{y_\sigma:0\leq\sigma\leq t\}$ about the joint unnormalized conditional density is contained in the sequence

$$\mathrm{T}\Delta\left\{\int \frac{\sigma^k}{k!}\dot{y}_\sigma d\sigma;\quad k=0,1,2,\dots\right\}. \tag{43}$$

Thus T is nothing but a joint sufficient statistic for the identification problem.

## 5 Control Agents and Yang–Mills Fields

Consider an object, the motion equation for which can be represented as

$$\dot{x} = r(x,u), \tag{44}$$

where $x=(x_1,x_2,x_3)\in Q\subset\mathbb{R}^3$; a function $r(x,u)$ is derived when an equation for dynamics of a particle in a field is reduced to Cauchy form, and the field is characterized by a variable $u$. The equations similar to (44) are widely used in physics and its applications. The equations of the concrete particle dynamics are considered in (Daniel,1988) and in many other papers. At present, control dynamics equation construction problems deserve a great attention. For instance, these problems include controllable models of dynamics of particles in scalar, vector and spinor fields.

This section builds up a controllable model for dynamics of a particle in electromagnetic and charged fields. The model is based on the gauge field concept (Daniel, 1980), which allows us to formulate different principles for an automatic control of the dynamics of the particles.

Constructing a controllable model means creating a transformation from a field $u$ to *Yang–Mills field*. The essence of this transition is as follows (Mitter, 1979). Instead of $u$, consider an $n$-component vector field $\widehat{f}(\widehat{x})$, $\widehat{x}\in T^1$ in a 4D space-time $T^1$. Let $M(\widehat{x})$ be local gauge transformations such that

$$\widehat{f}(\widehat{x}) = M(\widehat{x})\widehat{f}'(\widehat{x}) \tag{45}$$

and, under a fixed $x$, $M(x)$ form a group $G_1\in GL(n)$. Introduce an operator $\nabla_\alpha$, i.e.

$$\nabla_\alpha\widehat{f} = \left[\partial_\alpha + K_\alpha(\widehat{x})\widehat{f}(\widehat{x})\right], \tag{46}$$

which satisfies the conditions

$$M(\widehat{x})\nabla'_\alpha\widehat{f}'(\widehat{x}) = \nabla_\alpha\widehat{f}(\widehat{x}),\quad \nabla'_\alpha = \partial_\alpha + K'_\alpha, \tag{47}$$

where $K_\alpha = -Q_bC^b_\alpha$; $\{Q_b\}$ is a basis of Lie algebra $\widehat{g}$ for a group $G_1$, $[G_a,Q_b] = g^c_{ab}Q_c$; $G^c_{ab}$ are structural constants of Lie algebra $\widehat{g}$. The equations for the values $C^b_\alpha$ are derived from the Lagrangian $Y^a_{\alpha\beta}Y^{\alpha\beta}_a$, where

$$Y^a_{\alpha\beta} = \frac{\partial C^a_\beta}{\partial\widehat{x}^a} - \frac{\partial C^a_\alpha}{\partial\widehat{x}^\beta} - \frac{1}{2}g^a_{bc}\left(C^b_\alpha C^c_\beta - C^b_\beta C^c_\alpha\right), \tag{48}$$

and the Lagrangian has the following form:

$$\partial_\beta Y^{\alpha\beta} = Y_b^{\alpha\beta} g_{ac}^b C_\beta^c.$$

Relation (47) yields the law of transformation for a field of matrices $K_\alpha$:

$$K'_\alpha(\widehat{x}) = M^{-1}(\widehat{x}) K_\alpha(\widehat{x}) M(\widehat{x}) + M(x)^{-1} \frac{\partial M(\widehat{x})}{\partial \widehat{x}^\alpha}.$$

Such transformation satisfies the group law $g$. A set of these transformations forms a gauge group, formally denoted as

$$\widetilde{g} = \prod_x g.$$

It is shown in (Fadeev, 1982), that the values $C_\alpha^b$ are Yang–Mills fields. The Yang–Mills field describes a parallel transfer in a charge field and states its curvature. Such field can be brought in correspondence with the notion of connectedness in some main fiber bundle $(P, T^1, \widetilde{g})$, $\pi\colon P \to T^1$, where $T^1$ is a base and $\widetilde{g}$ is a structure group.

A control in $(P, T^1, \widetilde{g})$, $\pi\colon P \to T^1$ is understood as a connectedness $C_\alpha^b$. Notice that one can consider a projection $\pi$ as a control. Thus, it is possible to deal with a "controllable" fiber bundle $(P, T^1, \widetilde{g})$, $\pi\colon P \to T^1$ and a vector field $r(\widehat{x}, u(C_\alpha^b))$ on $P$ instead of the initial object described by equation (44). To make this interpretation more concrete, some additional investigations are required, and their results are expected to be published separately.

To solve control problems, it is necessary to construct equivalent and aggregated models. We construct an equivalent model of a controllable object $(P, T^1, \widetilde{g})$, $\pi\colon P \to T^1$ as follows. Let $\pi\colon P \to T^1$ be a main $\widetilde{g}$-fiber-bundle and let $l\colon Z \to T^1$ be some $m$-dimensional $\widetilde{g}$-vector fiber bundle with a trivial action, exerted by $\widetilde{g}$ onto $Z$. Assume also that a structure of a $k > 1$-dimensional cellular set can be introduced on $T^1$. An equivariant embedding of $\pi$ into 1 is understood as an embedding $h\colon P \to Z$, commutating with projections. If $K > m$, i.e., an action, exerted by $\widetilde{h}$ onto $Z$ is free outside a zero section for 1, then the main $\widetilde{g}$-fiber-bundle $\pi\colon P \to T^1$ can be equivariantly embedded into $l\colon Z \to T^1$. An equivalent model of a controlled process is understood as a ternary $(Z, T^1, \widetilde{g})$. In its turn, an equivalent model admits an exact aggregation, performed by means of a factorization of an induced *vector fiber bundle*. In this case, it is possible to assume, that a vector fiber bundle is specified by an interrelation system $\omega$ on some set $X^1$. Introduce an equivalence relation $S$ on $X^1$. This relation generates an object of the same nature, as the initial object $X^1$, and a factor-object ($F$-object) is obtained, which possesses a factorizing equivalence relation $S$. If $(X_1, \omega)$ generates an object of the same nature, possessed by an initial object, and this generation is carried out on a subset $X_1$ of $X_1$, then a subobject $X_1(\widetilde{\omega})$ ($P$-object) is derived. By using the mathematical structure theory language, it is possible to create a general theory of aggregation of invariant models for nonlinear systems.

Consider the main automatic particle dynamics control principles, when electromechanical systems with distributed parameters are taken as an example. It is shown in (Samoilenko, 1982), that a closed distributed automatic control system can be represented by two subsystems $S_1$ and $S_2$, interrelated by electromagnetic field

$$S = S_1 \cup S_2.$$

Represent $B$ of fields of a whole control system state by a field of internal states of each subsystem $B_1$ and $B_2$, of an interaction field $B_0$ and a by-side field $B_3$. In addition, represents $B_0$ by two components $X$ and $U$, i.e., $B_0 = X + U$, where $X$ is an information carrier and $U$ is a control field. Consider $U$ as the result of an influence exerted by $X$ onto the control medium and simulated by an operator dependence $U = \widehat{B}(X, E)$, where $\widehat{B}$ is a control operator and $E$ is the control medium power supply field. A by-side field $B$ also falls into two components, viz. into $V$ and $N$, where $V$ is a field of control which is carried out according to a fixed space-time program, and $N$ is a field of disturbing effects. The control object is described by a fiber bundle $(P, T^1, \widetilde{g})$ with a control $C^b_\alpha(X, Y, U, V, N)$, where $Y$ is a field of an internal field state. It is clear that a section is only one in $(P, T^1, \widetilde{g})$, if $U$, $V$ and $N$ are physically implementable and uniquely specified. The general problem, concerning calculation of an electromagnetic field of control system, consists of finding such physically implementable operator $\widehat{B}$ and programmed controlling influence $V$, under which the particle dynamics would meet certain previously formulated requirements.

## 6 Multiagent Systems and Fiber Bundes

Investigations of controlled multiagent objects as information-transforming systems have been under active development for last few years. Despite the achievements that have been made in this area, effective mathematical methods for investigating such systems have not yet been developed. One possible approach to close is the differential geometry methods of the system theory (Van der Shaft, 1982, 1987). This section is devoted to one of the problems of this area of research, that of developing a method for analyzing a class of mathematical models of symmetric controlled processes. Assuming that the process is described by a commutative diagram (Van der Shaft, 1982, 1987) which is based on the lamination concept, we propose a geometric method for "identifying" its hidden structure.

Investigation of the information-transformation laws in various systems is one of the most essential stages in the creation of new agents. The goal of the experimental and theoretical research is the implementation of optimal strategy using complex structure nonequilibrium processes in such systems. To investigate these processes it is required to develop the corresponding mathematical methods. In this context we propose an approach, which is based on the assumption that one can use models from the mathematical system theory to adequately describe informational processes. The essence of this approach is in the following.

Some dynamic system, $S$, which implements a transformation, $F$, or an input informational action, $U$, into an output one, $X$, is considered. It is assumed that one can affect the information-transforming process by a reconfiguring action that changes the dynamic behavior, structure, symmetry, etc. of the process. We refer to the objects described in the preceding $S$ as dynamic information-transforming agents (DITA).

The connection between the input and output actions is necessary for obtaining answers to questions about the method of programming the entire system, optimizing the flow of informational signals, and the interconnections among the global system properties (stability, controllability, etc.) and the corresponding local properties of the various subsystems. One has to answer those questions also when solving pattern–recognition problems, constructing an associative memory. A generalized description of an DITA that contains a large number of subsystems (for example, a neural network) is postulated in this section: the controlled process in the DITA is described adequately by a commutative diagram which generalizes the concept of a nonlinear controlled dynamic system on a manifold. Taking into account the symmetry concept which is characteristic of classical mechanics (Arnold, 1974), one has to transfer it to the DITA, "identify" the hidden structure of the informational process, and demonstrate that the proposed model admits local and/or global decompositions into smaller-dimensionality feedback subsystems.

We note that the decomposition idea was first applied to discretely symmetric automatic control systems by Yu. Samoylenko (the elementary cell method) (Samoilenko, 1972). Continuous symmetry group dynamic systems were considered by Van der Shaft (Van der Shaft, 1987). Substantive results on the decomposability of systems with symmetries have been obtained by A.Y. Krener (Krener, 1973) and others. However, this question remains open for DITAs.

**Necessary concepts and definitions.** Some definitions and concepts that are necessary for describing the DITA structure and the conditions for its decomposability are presented in this section. The necessary notions about manifolds, connectivities, and distributions are given in (Griffiths, 1983). We introduce the definition of a nonlinear DITA.

**Definition 8** *We refer a triple, $F(B, M, \psi)$, where $B$ is a smooth fiber over $M$ with the projection $\pi : B \to M$; $\pi_M$ is the natural projection of $TM$ on $M$; and $\psi$ is a smooth mapping such that the diagram presented in Fig. 3 is commutative, by a "geometrical model of the agent".*

We interpret the $M$ manifold as the DITA state space and the $\pi^{-1}(x) \in B$ layer as the space of input action values which depends in the general case on the current system state. If one chooses the coordinates $(x, u)$, which correspond to the $B_x$ layer, then this definition of the DITA, $F$, corresponds

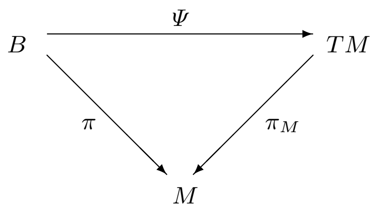



**Fig. 3.** Diagram of a nonlinear controlled DITA.

locally to the nonlinear transformation $\psi : (x, u) \to (x, \psi(x, u))$ and the dynamic system

$$x(t) = \psi(x(t), u(t)), \quad u(t) \in U. \tag{49}$$

where $x$ is the DITA state vector, $u = (u^1, u^2)$ are the control actions, $u^1(\cdot, \cdot)$ is the vector of the coded input informational action which depends in general on time and on the current state, and $u^2(\cdot, \cdot)$ is the action used to reconfigure the dynamic properties of the DITA and to train it.

The control algorithm, $u^2$, inputs to the system the capability of transforming the set of input actions into a set of output signals that allows one to identify the input images uniquely. In essence, it realizes the decoding process, which identifies the input images. In the simplest case, it can be realized on the basis of the successive input action segmentation method. Such a method facilitates a unique separation of the input images by the use of the simplest binary decoding rule.

**Definition 9** *Let $M$ be a smooth manifold. We say that the smooth mapping $Q : G \times M$ such that:*
*1. $Q(e, x) = x$ for all $x \in M$, and*
*2. $Q(g, Q(h, x)) = Q(gh, x)$ for any $g$ and $h \in G$, and all $x \in M$, is the left action (or $G$-action) of the $G$ Lie group on $M$.*

We fix one of the variables for various time instants and examine the $Q$ action as a function of the remaining variables. Let $Q_g : M \to M$ denote the function $x \ \mapsto Q(g, x)$ and $Q_x : G \to M$ the function $g| \to Q(g, x)$. We note that since $(Q_g)^{-1} = Q_g^{-1}$, $Q_g$ is a diffemorphism.

We introduce the definition of group action on a manifold.

**Definition 10** *Let $Q$ be the action of $G$ on $M$. We say that the set $G \cdot x = \{Q_g(x) | g \in G\}$ is the orbit ($Q$-orbit) of the point $x \in M$. The action is free at $x$ if $g| \to Q_g(x)$ is one-to-one. It is free on $M$ if and only if it is free at all $x \in M$.*

We now introduce the concept of global symmetry of a controlled DITA.

**Definition 11** *Let $\hat{F}(B, M, \psi)$ be a nonlinear controlled DITA, and $\theta$ and $Q$ be actions of $G$ on $B$ and $M$, respectively. Then, $F$ has symmetry $(G, \theta, Q)$ if the diagram presented in Fig. 4 is cummutative for all $g \in G$.*

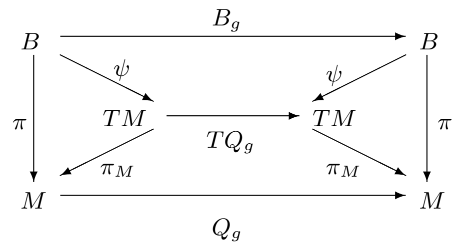



**Fig. 4.** A commutative diagram of an DITA with symmetries.

We consider, within the framework of the presented definition, the special case in which the symmetry lies "entirely within the state space".

**Definition 12** *Let $B = M \times U$, where $U$ is some manifold. Then, $(G, Q)$ is a symmetry of the state space of system $\hat{F}(B, M, \psi)$ if $(G, \theta, Q)$ is a symmetry of $\hat{F}$ for $\theta_g = (Q_g, Id_U) : (x, u) \to (Q_g(x), u)$.*

Global state space symmetry can be defined only for an DITA $B_x$ of which is a trivial lamination since otherwise the input spaces would depend on the state and the problem is made substantially more complicated.

We introduce now the definition of local symmetry.

**Definition 13** *We assume that $Q : G \times M \to M$ is an action and that $\varepsilon \in T_eG$. Then, $Q^\xi(R \times M \to M) : (t, x)| \to Q(\exp t\xi, x)$, where $\exp : T_eG \to G$ is the usual exponential mapping, is the $\mathbb{R}$-action on $M$, and $Q^\xi$ is the complete flow on $M$. We say that the corresponding vector field on $M$, which is defined by the expression*

$$\xi_m(x) = \frac{d}{dt} Q(\exp t\xi, x)\Big|_{t=0}, \tag{50}$$

*is the infinitesimal action generator, which corresponds to $\xi$.*

Let $X_t$ denote the flow of the vector field $X$, that is, $X_t = F_t(X_0)$. It is obvious from the definition of the infinitesimal generator that if $(G, \theta, Q)$ is a symmetry of the $\hat{F}(B, M, \psi)$ system, then the diagram presented in Fig. 5 is commutative for all $t \in \mathbb{R}$ and $\xi \in T_eG$.

On the basis of the local commutativity property we present the following definition of infinitesimal DITA symmetry.

**Definition 14** *Let $\hat{F}(B, M, \psi)$ be a nonlinear DITA. Then, $(G, \theta, Q)$ is an infinitesimal symmetry of $F$ if, for each $x_0 \in M$, there exist an open neighborhood $\hat{O}$ of the point $x_O$ and $\xi > 0$ such that*

$$(\xi_M)_t * \psi(\xi) = \psi((\xi_b)_t(b)), \tag{51}$$

*for all $b \in \pi^{-1}(\hat{O})$, $|t| < \xi$, and $\| \xi \| < 1$, $\xi \in T_eG$, where $\| \cdot \|$ is an arbitrary fixed norm on $T_eG$.*

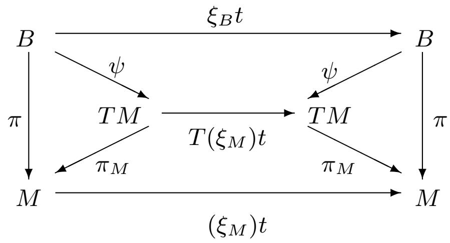



**Fig. 5.** Diagram of a symmetric DITA.

One can define an infinitely small state space symmetry for nontrivial laminations of the input actions manifold when one can introduce integratable connectivity. For this we introduce Definition 15.

**Definition 15** *Let $H(\cdot)$ be an integratable connectivity on $B$ and $(G, \theta, Q)$ be a symmetry of $F$. Then, $(G, \theta, Q)$ is an infinitesimal state space symmetry if $\xi_B(b) \in H(b)$ for all $\xi \in T_eG$, that is, the infinitesimal generators $\theta$ are horizontal.*

We introduce a definition of feedback equivalence of two DITAs in analogy with (Van der Shaft, 1982).

**Definition 16** *A system, $F(B, M, \psi)$, is feedback equivalent to a system, $F'(B, M, \widetilde{\psi})$, if there exists an isomorphism, $\gamma : B \to B$, such that the diagram presented in Fig. 6 is commutative.*

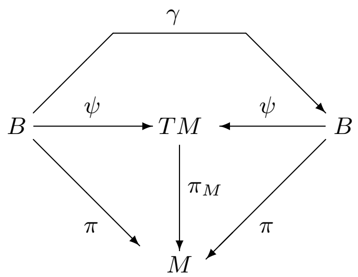



**Fig. 6.** Diagram of feedback-equivalent DITAs.

Isomorphism means that, for $x \in M$, $\gamma_x$ is a mapping from the layer over $x'$ into the layer over $x'$, and it is a diffeomorphism. Consequently, this corresponds to a "control feedback".

**The local structure of DITAs with symmetries.** Since we are interested in the local structure of an DITA, we have to assume that the system has an infinitesimal symmetry, which satisfies some nonsingularity condition. For this, we set the dimensionality of $M$ to $n$ and that of $G$ to $k$, where

$k < n$. We note that the action $Q : G \times M \to M$ is free at the point $m \in M$ if $Q_m : G \to M$ is one-to-one. This is equivalent to saying that the tangent mapping $Q$ is of full rank, that is, rank $Q = \dim G$. Hence, $Q$ is free on $M$ if and only if it is free in some neighborhood of $m$. We say that an action which satisfies this condition is nonsingular at the point $m$.

The basic result of this section is that the existence of an infinitesimal symmetry in a neighborhood of a singular point in an DITA makes it possible to decompose the system into a cascade union of simpler subsystems. The structure of these subsystems depends, in general, on the symmetry group $G$. If, for example, $G$ has a nontrivial center, then one of the subsystems is in fact a quadrature subsystem.

Let, in addition, $C = h \in G|jg = gh$ for all $g \in G$ be the center of the $G$ group to which the kernel, $C_+$, of the Lie semialgebra $T_eG$, which has the same dimensionality as $C$, corresponds. Hence, if $G$ has an $l$-dimensional center, there exist linearly independent vectors $\xi^1, \ldots, \xi^k \in T_eG$ such that $[\xi^i, \xi^j] = 0$ for all $1 \leq i \leq l$ and $1 \leq j \leq k$.

Using the results of Van der Shaft, Markus, and Grizzle's investigations (Van der Shaft, 1982, 1987; Markus, 1973, 1976; Grizzle, 1972) that deal with the properties of systems with symmetries as applied to DITAs, one can formulate the following theorems.

**Theorem 7** *Let us assume that $\hat{F}(B, M, \xi)$ is a controlled DITA with an infinitesimal state space symmetry, $(G, \theta, Q)$, that $G$ has an $l$-dimensional center, and that $Q$ is nonsingular at the point $m \in M$. Then, the $B$ coordinates $(x_1, \ldots, x_n, u)$ in a neighborhood of $m$ exist such that $\hat{F}$ is given in these coordinates by the expression.*

Using the obtained results for systems for infinitesimal state space symmetries, one can propose the structure of the decomposed system. It suffices to demonstrate for this that the decomposed system with infinitesimal symmetry is locally feedback-equivalent to the original system with infinitesimal state space symmetry.

**Definition 17** *Let $\hat{F}(B, M, \psi)$ be a controlled DITA and $\hat{O}$ be an open subset of $M$. Then, we say that a system of the form $\hat{F}(\pi^{-1}(\hat{O}), \hat{O}, \psi)|\pi^{-1}(O)$ is $\hat{F}|\hat{O}$ ($\hat{F}$ bounded on $\hat{O}$).*

**Theorem 8** *Let $\hat{F}(B, M, \psi)$ have an infinitesimal symmetry $(G, \theta, Q)$ and $Q$ be nonsingular at the point $m$. There exist a neighborhood of $m$ and a system $F$ with infinitesimal symmetry $(G, \theta, Q)$ such that $\hat{F}|\hat{O}$ is feedback equivalent to the $\hat{F}$ system.*

Let $\hat{F}(B, M, \psi)$ be a controlled DITA with symmetry $(G, \theta, Q)$ and $Q$ be nonsingular at the point $m$. Then, in a neighborhood of $m$, $\hat{F}$ is feedback-equivalent to $\hat{F}$ with infinitesimal symmetry and has the structure shown in Fig. 7, where $\gamma$ is the feedback function, the $L^i$ are nonlinear subsystems of

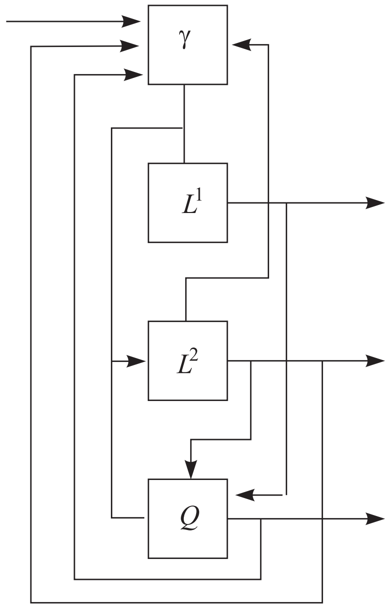



**Fig. 7.** Local structure of DITA with infinitesimal symmetries.

dimensions $n-k$ and $k-l$, respectively, and $Q$ is an $l$-dimensional "quadrature" system

$$\begin{aligned} \dot{x}_i &= f_i(x_1, \ldots, x_{n-k}, u), \quad i = 1, \ldots, n-k, \\ \dot{x}_j &= f_j(x_1, \ldots, x_{n-1}, u), \quad i = n-k+1, \ldots, k. \end{aligned} \tag{52}$$

**The global structure of DITA.** The decomposability of an DITA with global symmetries is the result of factoring the DITA state space, which follows from the properties of a symmetry.

We introduce the definition of proper action.

**Definition 18** *Let $Q$ be a $G$-action on $M$. We say that $Q$ acts properly if $(g, m) \to m$ is a proper mapping, that is, if the pre-images of compact sets are compact.*

This definition is equivalent to the following assertion: whenever $x_n$ converges on $M$ and $Q_{g_n}(x_n)$ converges on $M$, $g_n$ includes a subsequence, which converges in $G$. Hence, if $G$ is compact, this condition is satisfied automatically. Membership in the same $Q$-orbit is an equivalence relation on $M$. Let $M/G$ be the set of equivalence classes and $p : M \to M/G$ be specified by the relation $p(m) = Gm$. We introduce on $M/G$ a relations topology, that is, $V \subset M/G$ is open if and only if $p^{-1}(V)$ is open on $M$. In general, $M/G$ can be a rather poor space.

If $G$ acts freely and properly on $M$, then $M/G$ is a smooth manifold and $p : M \to M/G$ is the principal lamination with Lie group $G$.

We introduce the following constraints on the principal lamination:

1) $p$ is a smooth full-rank function;
2) $p : M \to M/G$ has a cross section (that is, a smooth mapping $\sigma : M/G \to M$ such that $p \cdot \sigma$ a is the identity mapping on $M/G$ if and only if M is equivalent to $M/G \times G$;
3) the topological conditions which guarantee the existence of a section, that is, if $M/G$ or $G$ is a contraction mapping, a cross section must exist, are specified.

We formulate a theorem, which is necessary for obtaining a global factorization of the DITA state space.

Let $Q_m : G \to G \cdot m$ be specified by $g \to Q(G, m)$. The following result about the global structure of a DITA with symmetries holds.

**Theorem 9** *We assume that $\hat{F}(M \times U, M, \psi)$ is a controlled DITA with a state space symmetry $(C, Q)$. Then, if $Q$ is free and proper, and $p : M \to M/G$ has a cross section $\sigma$, then $\hat{F}$ is isomorphic to the system*

$$\begin{aligned} \dot{y} &= \Psi(y, u), \\ \dot{g} &= (T_e L_g)(T_e Q_{\sigma(y)})^{-1} \left[\Psi(\sigma(y), u) - (T_y \sigma)\Psi(y, u)\right], \end{aligned} \tag{53}$$

*defined on $M/G \times G$.*

We formulate an assertion on feedback equivalence of DITAs with symmetries.

**Assertion 1** *Let the DITA $F(M \times U, M, \psi)$ have a symmetry $(G, \theta, Q)$ such that $Q$ is free and proper. Then, there exists a system $F$ with symmetry $(G, Q)$ to which $F$ is feedback equivalent under the condition that $p : M \to M/G$ has a cross section $\sigma$.*

Combining Theorem 9 and Assertion 1, we obtain the following corollary

**Corollary 1** *Let DITA $\hat{F}(M \times U, M, \psi)$ have a symmetry $(G, \theta, Q)$, $Q$ be free and proper, and $p : M \to M/G$ have a cross section. Then, there exists a model of DITA $F$ with state space symmetry $(G, Q)$ to which $\hat{F}$ is feedback-equivalent. Consequently, $F$ has a global structure.*

**The feasibility of applying the results to the investigation of agents.** It is of interest to investigate the decomposability of DITAs composed of neural-like agents that are described by the system of equations

$$\dot{x}(t) = \psi(x(t), u(t)). \tag{54}$$

One can define for (54) a decomposed system $L$ as a nontrivial cascade of subsystem $L^1$ and $L^2$. If the Lie algebra $\hat{L}(L)$ is the semidirect sum of finite-dimensional subalgebra $L^1$ and the ideal of $L^2$, it has a nontrivial cascade decomposition into subsystems $L^1$ and $L^2$ such that $\hat{L}(L^1) = L^1$, and $\hat{L}(L^2) = L^2$. Using this fact and Levy's theorem one can demonstrate that if $\hat{L}(L)$ is finite-dimensional, the DITA admits a nontrivial decomposition into a parallel cascade of $L^i$ systems with simple Lie algebras followed by a cascade of one-dimensional systems, $L^j$. As a result, the basic informational transformation is done in subsystems with simple Lie algebras. The state space, $M$, of the original system, $L$, is adopted here as the state space of these systems. Therefore, despite the fact that the system has been partitioned into simpler parts, the overall dimensionality of these parts is, in general, larger than that of the original system. (One can reduce at the local level this dimensionality by replacing the $L^i$ system by matrix equivalents defined on the exponential functions of the Lie algebras that correspond to them.) These results can be compared with the conditions for decomposability obtained by analyzing the DITA symmetries described in this section for which the subsystem dimensionality equals that of the original system. No assumptions about the finite dimensionality of the Lie algebra are required here. We consider a class of neural nets described by the linear-analytic equations

$$\dot{x}(t) = f(x) + \sum_{i=1}^{k} u_i g_i(x). \tag{55}$$

One can formulate for it the necessary and sufficient conditions for parallel-cascade decomposability by Lie algebras. In doing so, one can pose the condition that each component of the input action be applied to only one of the subsystems, that is, the decomposition procedure partition the inputs into disjoint subsets. However, such an approach cannot be applied to the decomposition of an DITA with scalar input.

If DITA $\hat{F}(B, M, \psi)$ has an infinitesimal symmetry $(G, \theta, Q)$, local commutativity of the diagram means that $\psi * \varepsilon_B = \varepsilon_m$ and $\pi * \varepsilon_B = \varepsilon_n$. Let $\Delta_B = \mathrm{span}\{\varepsilon \| \varepsilon_B \in T_e G\}$ and the same hold for $\Delta_m$. Then, $\psi * \Delta B \subset \Delta_m$ and $\pi * \Delta_B = \Delta$, and $\Delta_m$, is a controlled invariant distribution. Models of neural networks, including affine ones, have invariant distributions that induce decompositions of the system into simpler subsystems. However, since the symmetry conditions are constraints, the decompositions are obtained as more detailed and structured.

A class of dynamic information-transforming systems that are described by a commutative diagram is examined in this section. Constraints on systems with symmetry under which one can expose, explicitly the hidden structure of the controlled process are formulated. We show that the effect of the DITA on the information-transforming process depends substantially on the type of system symmetry. The informational process is subject here to the action of cascade group, transformations or the action of a dynamic-transformation operator with feedback. The obtained results can be expanded to adaptive learning systems by introducing the corresponding optimization models. When doing so, one can expect that an DITA of which the quality functional is invariant in symmetry-conserving transformations will be described adequately by a nonlinear system with optimal feedback and will have a differential-geometric structure, which is of interest from the point of view of applications. We plan to use the results of the investigations presented here in the study of a synergetic model of a neural network on the basis of potential-dependent ion channels in biomembranes.

## 7 Fiber Bundles and Observability

In the last decade, an important work has been done on a differential geometric approach to nonlinear input state-output systems, which in local coordinates have the form

$$\dot{x} = g(x, u), \quad y = h(x), \tag{56}$$

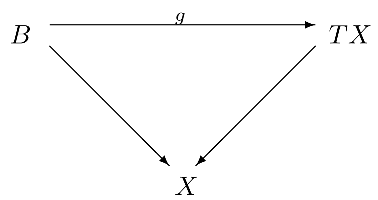
where $x$ is the *state* of the system, $u$ is the *input* and $y$ is the *output*. Most of the attention has been directed to the formulation in this context of fundamental system theoretic concepts like controllability, observability, minimality and realization theory.

In spite of some very natural formulations and elegant results, which have been achieved, there are certain disadvantages in the whole approach, from which we summarize the following points,

a) Normally the equations

$$\dot{x} = g(x, u) \tag{57}$$

are interpreted as a family of vector fields on a manifold parametrized by $u$; i.e., for every fixed $\overline{u}$, $g(\cdot, \overline{u})$ is a globally defined vector field. We propose another framework by looking at (57) as a coordinatization of




where $B$ is a *fiber bundle* above the state space manifold $X$ and the fibers of $B$ are the *state dependent* input spaces, while $TX$ is as usual the tangent bundle of $X$ (the possible velocities at every point of $X$).

b) The "usual" definition of *observability* has some drawbacks. In fact, observability is defined as *distinguishability*; i.e., for every $x_1$ and $x_2$ (elements of $X$) there exists a *certain* input function (in principle dependent on $x_1$ and $x_2$) such that the output function of the system starting from $x_1$ under the influence of this input function is different from the output function of the system starting from $x_2$ under the influence of the same input function. Of course, from a practical point of view this notion of observability is not very useful, and also is not in accord with the usual definition of observability or reconstructibility for general systems.

Hence, despite the work of Sussmann (Sussman, 1983) on *universal* inputs, i.e., input functions, which distinguish between every two states $x_1$ and $x_2$, this approach remains unsatisfactory.

c) In the class of nonlinear systems (56), *memoryless* systems

$$y = h(u) \tag{58}$$

are not included. Of course, one could extend the system (56) to the form

$$\dot{x} = g(x, u), \quad y = h(x, u), \tag{59}$$

but this gives, if one wants to regard observability as distinguishability, the following rather complicated notion of observability. As can be seen, distinguishability of (59) with $y \in \mathbb{R}^p$, $u \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$ is equivalent to distinguishability of

$$\dot{x} = g(x, u), \quad \overline{y} = \overline{h}(x), \tag{60}$$

where $\overline{h} : \mathbb{R}^n \to (\mathbb{R}^p)^{\mathbb{R}^m}$ is defined by $\overline{h}(x)(u) = h(x, u)$.

Checking the Lie algebra conditions for distinguishability for the system (60) is not very easy.

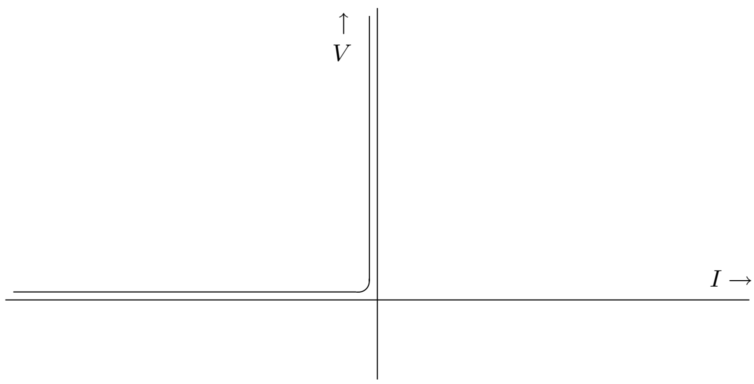
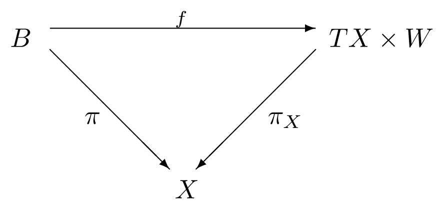



**Fig. 8.**

d) It is often not clear how to distinguish a priori between inputs and outputs. Especially in the case of a nonlinear system, it could be possible that a separation of what we shall call *external variables* in input variables and output variables should be interpreted only *locally*. An example is the (nearly) ideal diode given by the $I - V$ characteristic in Fig. 8. For $I < 0$ it is natural to regard $I$ as the input and $V$ as the output,while for $V > 0$ it is natural to see $V$ as the input and $I$ as the output. Around) an input-output description should be given in the scattering variables $(I-V, I+V)$. Moreover, in the case of nonlinear systems it can happen that a global separation of the external variables in inputs and outputs is simply not possible! This results in a definition of a system, which is a generalization of the usual input-output framework. It appears that various notions like the definitions of autonomous (i.e., without inputs), memoryless, time-reversible, Hamiltonian and gradient systems are very natural in this framework.

### 7.1 Nonlinear Model of Agents

The (say $C^\infty$) agents can be represented in the commutative diagram




(61)

where (all spaces are smooth manifolds) $B$ is a fiber bundle above $X$ with projection $\pi$, $TX$ is the tangent bundle of $X$, $\pi_x$ the natural projection of $TX$ on $X$ and $f$ is a smooth map. $W$ is the space of external variables (think of the inputs *and* the outputs). $X$ is the state space and the fiber $\pi_{-1}(x)$ in $B$ above $\in X$ represents the space of inputs (to be seen initially as *dummy* variables), which is state dependent (think of forces acting at different points of a curved surface).

This definition formalizes the idea that at every point $x \in X$ we have a set of possible velocities (elements of $TX$) and possible values of the external variables (elements of $W$), namely the space

$$f(\pi^{-1}(x)) \subset T_x X \times W.$$

We denote the system (61) by $\Sigma(X, W, B, f)$. It is easily seen that in local coordinates $x$ for $X$, $v$ for the fibers of $B$, $w$ for $W$, and with $f$ factored in $f = (g, h)$, the system is given by

$$\dot{x} = g(x, v), \quad w = h(x, v). \tag{62}$$

Of course one should ask oneself how this kind of system formulation is connected with the usual input-output setting. In fact, by adding more and more assumptions successively to the very general formulation (61) we shall distinguish among three important situations, of which the last is equivalent to the "usual" interpretation of system (56).

(i) Suppose the map $h$ restricted to the fibers of $B$ is an *immersive* map into $W$ (this is equivalent to assuming that the matrix $\partial h/\partial v$ is injective). Then:

**Lemma 1** *Let $h$ restricted to the fibers of 4 be an immersion into $W$. Let $(\overline{x}, \overline{u})$ and $\overline{w}$ be points in $B$ and $W$ respectively such that $h(\overline{x}, \overline{v}) = \overline{w}$. Then locally around $(\overline{x}, \overline{v})$ and $\overline{w}$ there are coordinates $(x, v)$ for $B$ (such that $v$ are coordinates for the fibers of), coordinates $(w_1, w_2)$ for $W$ and a map $\overline{h}$ such that $h$ has the form*

$$(x, v) \gg h > (w_1, w_2) = (\overline{h}(x, v), v). \tag{63}$$

*Proof.* The lemma follows from the implicit function theorem.

Hence *locally* we can interpret a part of the external variables, i.e., $w_1$, as the outputs, and a complementary part, i.e., $w_2$, as the inputs! If we denote $w_1$ by $y$ and $w_2$ by $u$, then system (62) has the form (of course only locally)

$$\dot{x} = y(x, u), \quad y = \overline{h}(x, u). \tag{64}$$

(ii) Now we not only assume that $\partial h / \partial v$ is injective, which results in a *local* input-output parametrization (64), but we also assume that the output set denoted by $Y$ is *globally* defined. Moreover, we assume that $W$ is a fiber bundle above $Y$, which we call $p : W \to Y$, and that $h$ is a bundle morphism (i.e., maps fibers of $B$ into fibers of $W$). Then:

**Lemma 2** *Let $h : B \to W$ be a bundle morphism, which is a diffeomorphism restricted to the fibers. Let $\overline{x} \in X$ and $\overline{y} \in Y$ be such that $h(\pi^{-1}(\overline{x})) = p^{-1}(\overline{y})$. Take coordinates $x$ around $\overline{x}$ for $X$ and coordinates $y$ around $\overline{y}$ for $Y$. Let $(\overline{x}, \overline{v})$ be a point in the fiber above $\overline{x}$ and let $(\overline{y}, \overline{u})$ be a point in the fiber above $\overline{y}$ such that $h(\overline{x}, \overline{v}) = (\overline{y}, \overline{u})$. Then there are local coordinates $v$ around $\overline{v}$ for the fibers of $B$, coordinates $u$ around $\overline{u}$ for the fibers of $W$ and a map $\overline{h} : X \to Y$ such that $h$ has the form*

$$(x, v) \gg h > (y, u) = (\overline{h}(x), v). \tag{65}$$

*Proof.* Choose a locally trivializing chart $(0, \phi)$ of $W$ around $\overline{y}$. Then $\phi : p^{-1}(0) \to 0 \times U$, with $U$ the standard fiber of $W$. Take local coordinates $u$ around $\overline{u} \in U$. Then $(y, u)$ forms a coordinate system for $W$ around $(\overline{y}, \overline{u})$. Because $h$ is a bundle morphism, it has the form

$$(x, \overline{v}) \gg h > (y, u) = (\overline{h}(x), h'(x, \overline{v})).$$

where $(x, \overline{v})$ is a coordinate system for $B$ around $(\overline{x}, \overline{v})$. Now adapt this last coordinate system by defining

$$v = (h')^{-1}(x, u) \quad \text{with } x \text{ fixed.}$$

Because $h$ restricted to the fibers is a diffeomorphism, $v$ is well defined and $(x, v)$ forms a coordinate system for $B$ in which $h$ has the form

$$(x, v) \gg h > (y, u) = (\overline{h}(x), u).$$

Hence under the conditions of Lemma 2 our system is locally (around $\overline{x} \in X$ and $\overline{y} \in Y$) described by

$$\dot{x} = g(x, u), \quad y = \overline{h}(x). \tag{66}$$

This input-output formulation is essentially the same as the one proposed by Brockett and Takens, who take the input spaces as the fibers of a bundle above a globally defined output space $Y$. In fact, this situation should be regarded as the normal setting for nonlinear control systems.

(iii) Take the same assumptions as in (ii) and assume moreover that $W$ is a *trivial* bundle, i.e., $W = Y \times U$, and that $B$ is a trivial bundle, i.e., $B = X \times V$. Because $h$ is a diffeomorphism on the fibers, we can identify $U$ and $V$. In this case the output set $Y$ and the input set $U$ are *globally* defined, and the system is described by

$$\dot{x} = g(x, u), \quad y = \overline{h}(x), \tag{67}$$

where for each fixed $\overline{u}, g(\cdot, \overline{u})$ is a globally defined vector field on $X$. This is the "usual" interpretation of (56).

**Remark 2** 1. When $h$ restricted to the fibers of $B$ is *not* an immersion we have a situation where we could speak of "hidden inputs". In fact, in this case there are variables in the fibers of $B$ which can affect the internal state behavior via the equation $\dot{x} = g(x, v)$ but which cannot be directly identified with some of the external variables.

2. The splitting of the external variables into inputs and outputs as described in Lemma 1 is of course by no means unique! This fact has interesting implications, even in the linear case, which we shall not pursue further here.

3. From Lemma 2 it is clear that the coordinatization of the fibers of the bundle $W$ uniquely determines, via $h$, the coordinatization of the fibers of $B$. It should be remarked that a coordinatization of the fibers of $W$ is locally equivalent to the existence of an (integrable) *connection* on the bundle $W$, and that one coordinatization is linked with another by what is essentially an output feedback transformation, i.e., a bundle isomorphism from $W$ into itself. Again we do not comment further on this point.

4. A beautiful example of this kind of system is the Lagrangian system. Here the output space is equal to the configuration space $Q$ of a mechanical system. The state space $X$ is the configuration space with the velocity space, so $X = TQ$. The space $W$ is equal to $T^*Q$ (the cotangent bundle of $Q$), with the fibers of $T^*Q$ representing the external forces. When we denote the natural projection of $TQ$ on $Q$ by $\rho$, then $B$ is just $\rho^* T^*Q$ (the pullback bundle via $\rho$). Now given a function $L : TQ \to \mathbb{R}$ (called the Lagrangian) we can construct a symplectic form $d(\partial L/\partial \dot{q}) \wedge dq$ (with $(q, \dot{q})$ coordinates for $TQ$) on $TQ$, which uniquely determines a map $g : B \to TTQ$. Finally, in coordinates the system is given by

$$\ddot{q} = F(q, \dot{q}) + \sum_j u_j Z_j(q, \dot{q}), \quad y = q, \tag{68}$$

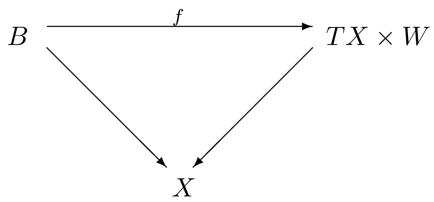
with the vector fields $F(q, \dot{q})$ and $Z_j(q, \dot{q})$ satisfying certain conditions. Moreover the vector fields $Z_j$ commute, i.e., $[Z_i, Z_j] = 0$ for all $i, j$, a fact which has a very interesting interpretation.

5. Most cases where $B$ can be taken as trivial are generated by a space $X$ such that $TX$ is a trivial bundle. For instance, when $X$ is a Lie group $TX$ is automatically trivial.

### 7.2 Minimality and Observability

**Minimality.** We want to give a definition of minimality for a general nonlinear agent

**Definition 19** *Let $\Sigma(X, W, B, f)$ and $\Sigma'(X', W, B', f')$ be two smooth systems. Then we say $\Sigma' \leqq \Sigma$ if there exist surjective submersions $\phi : X \to X', \Phi : B \to B'$ such that the diagram*

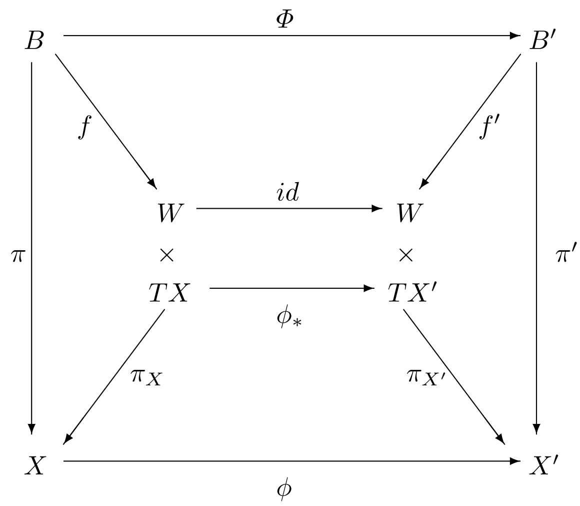



(69)

*commutes.*

$\Sigma$ is called *equivalent to* $\Sigma'$ (denoted $\Sigma \sim \Sigma'$) if $\phi$ and $\Phi$ are diffeomorphisms.

We call $\Sigma$ *minimal* if $\Sigma' \leqq \Sigma \Rightarrow \Sigma' \sim \Sigma$.




**Remark 3** This definition formalizes the idea that we call $\Sigma'$ *less complicated* than $\Sigma(\Sigma' \leqq \Sigma)$ if $\Sigma'$ consists of a set of trajectories in the state space, smaller than the set of trajectories of $\Sigma$, but which generates the same *external behavior*. (The external behavior $\Sigma_e$ of $\Sigma(X, W, B, f)$ consists

of the possible functions $w : \mathbb{R} \to W$ generated by $\Sigma(X, W, B, f)$. Hence, when we define $\Sigma := \{(x, w) : \mathbb{R} \to X \times W|_x$ absolutely continuous and $(\dot{x}(t), w(t)) \inf(\pi^{-1}(x(t)))a.e.\}$, then $\Sigma_e$, is just the projection of $\Sigma$ on $W^{\mathbb{R}}$).

**Remark 4** Notice that we only formalize the *regular* case by asking that $\Phi$ and $\phi$ be surjective as well as submersive. In fact we could, for instance, allow that at isolated points $\phi$ or $\Phi$ are not submersive. However, we do not discuss this problem here, and treat only the regular case as described in Definition 19.

**Remark 5** Notice that $\Sigma_1 \leqq \Sigma_2$ and $\Sigma_2 \leqq \Sigma_2$ need not imply $\Sigma_1 \sim \Sigma_1$. This fact leads to very interesting problems, which we do not pursue further at this time.

Of course, Definition 19 is an elegant but rather abstract definition of minimality. From a differential geometric point of view it is very natural to see what these conditions of commutativity mean *locally.* In fact, we will see in Theorem 11 that locally these conditions of commutativity do have a very direct interpretation. But first we have to state some preparatory lemmas and theorems.

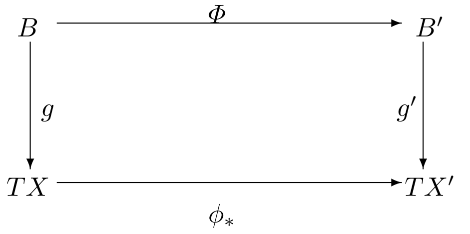
Let us look at (69). Because $\Phi$ is a submersion it induces an involutive distribution $D$ on $B$ given by

$$D := \{Z \in TB | \Phi_* \dot{Z} = 0\}$$

(the foliation generated by $D$ is of the form $\Phi^{-1}(c)$ with $c$ constant). In the same way $\phi$ induces an involutive distribution $E$ on $X$. Now the information in the diagram (69) is contained in three subdiagrams (we assume $f = (g, h)$ and $f' = (g', h')$):

$$\begin{array}{ccc} B & \xrightarrow{\Phi} & B' \\ \downarrow h & & \downarrow h' \\ W & \xleftrightarrow[id]{} & W \end{array} \qquad \text{I}$$

$$\begin{array}{ccc} B & \xrightarrow{\Phi} & B' \\ \downarrow \pi & & \downarrow \pi' \\ X & \xrightarrow[\phi]{} & X' \end{array} \qquad \text{II}$$




III

**Lemma 3** *Locally the diagrams I, II, III are equivalent, respectively, to*

$$
\begin{aligned}
I' : &\quad D \subset \ker dh, \\
II' : &\quad \pi_* D = E, \\
III' : &\quad g_* D \subset TE = T\pi_*(D).
\end{aligned} \tag{70}
$$

*Proof.* $I'$ and $II'$ are trivial. For $III'$ observe that, when $\phi$ induces a distribution $E$ on $X$, then $\phi_*$ induces the distribution $TE$ on $TX$.

Now we want to relate conditions $I'$, $II'$, $III'$ with the theory of nonlinear disturbance decoupling. Consider in local coordinates the system

$$\dot{x} = f(x) + \sum_{i=1}^{m} u_i g_i(x) \quad \text{on a manifold } X.$$

We can interpret this as an affine distribution on manifold.

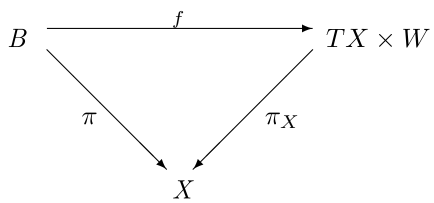
**Theorem 10** *Let $D \in A(\Delta_0)$. Then the condition*

$$[\Delta, D] \subseteq D + \Delta_0 \tag{71}$$

*(we call such a $D \in A(\Delta_0)\Delta(\operatorname{mod} \Delta_0)$ invariant) is equivalent to the two conditions a) there exists a vector field $F \in \Delta$ such that $[F, D] \subseteq D$; b) there exist vector fields $B_i \in \Delta_0$ such that span $\{B_i\} = \Delta_0$ and $[B_i, D] \subset D$.*

With the aid of this theorem the disturbance decoupling problem is readily solved. The key to connecting our situation with this theory is given by the concept of the *extended system*, which is of interest in itself.

**Definition 20** (Extended system). *Let*



*Then we define the extended system of $\Sigma(X,W,B,f)$ as follows: We define $\Delta_0$ as the vertical tangent space of $B$, i.e.,*

$$\Delta_0 := \{Z \in TB | \pi_* Z = 0\}.$$

Note that $\Delta_0$ is *automatically involutive.*

Now take a point $(\overline{x},\overline{v}) \in B$. Then $g(\overline{x},\overline{v})$ is an element of $T_{\overline{x}}X$. Now define

$$\Delta(\overline{x},\overline{v}) := \{Z \in T_{(\overline{x},\overline{v})} | \pi_* Z = g(\overline{x},\overline{v})\}.$$

So $\Delta(\overline{x},\overline{v})$ consists of the possible lifts of $g(\overline{x},\overline{v})$ in $(\overline{x},\overline{v})$. Then it is easy to see that $\Delta$ is an affine distribution on $B$, and that $\Delta - \Delta = \Delta_0$. We call the affine system $(\Delta, \Delta_0)$ on $B$ constructed in this way, together with the output function $h : B \to W$, the extended system $\Sigma^e(X,W,B,f)$.

We have the following:

**Lemma 4** *a) Let $D$ be an involutive distribution on $B$ such that $D \cap \Delta_0$ has constant dimension. Then $\pi_* D$ is a well-defined and involutive distribution on $X$ if and only if $D + \Delta_0$ is an involutive distribution.*

*b) Let $D$ be an involutive distribution on $B$ and let $D \cap \Delta_0$ have constant dimension. Then the following two conditions are equivalent: i) $\pi_* D$ is a well-defined and involutive distribution on $X$, and $g_* D \subset T\pi_* D$. ii) $[\Delta, D] \subset D + \Delta_0$.*

*Proof.* a) Let $D + \Delta_0$ be involutive. Because $D$ and $\Delta_0$ are involutive this is equivalent to $[D, \Delta_0] \subset D + \Delta_0$. Applying Theorem 10 to this case gives a basis $\{Z_1, \ldots, Z_k\}$ of $D$ such that $[Z_i, \Delta_0] \subset= \Delta_0$. In coordinates $(x,u)$ for $B$, the last expression is equivalent to $Z_i(x,u) = (Z_{ix}, Z_{iu}(x,u))$, where $Z_{ix}$ and $Z_{iu}$ are the components of $Z_i$ in the $x$- and $u$-directions, respectively. Hence $\pi_* D = \text{span}\{Z_{1x}, \ldots, Z_{kx}\}$ and is easily seen to be involutive. The converse statement is trivial.

b) Assume i); then there exist coordinates $(x,u)$ for $B$ such that $D = \{\partial/\partial x_1, \ldots, \partial/\partial x_x\}$ (the integral manifolds of $D$ are contained in the sections $u = \text{const}$). Then $g_* D \subset T\pi_* D$ is equivalent to

$$\left(\frac{\partial g}{\partial x_i}\right)_{j^e\text{comp}} = 0$$

with $i = 1, \ldots, k$ and $j = k + l, \ldots, n$ ($n$ is the dimension of $X$). From these expressions $[\Delta, D] \subset D + \Delta_0$ readily follows. The converse statement is based on the same argument.

Now we are prepared to state the main theorem of this section. First we have to give another definition.

**Definition 21** (Local minimality). *Let $\Sigma(X,W,B,f)$ be a smooth system. Let $\overline{x} \in X$. Then $\Sigma(X,W,B,f)$ is called locally minimal (around $\overline{x}$) if when $D$ and $E$ are distributions (around $\overline{x}$) which satisfy conditions $I'$, $II'$, $III'$ of Lemma 3, then $D$ and $E$ must be the zero distributions.*

It is readily seen from Definition 19 that minimality of $\Sigma(X,W,B,f)$ *locally* implies local minimality (locally every involutive distribution can be factored out).

Combining Lemma 3, Definition 20 and Lemma 4 we can state:

**Theorem 11** *$\Sigma(X,W,B,f=(g,h))$ is locally minimal if and only if the extended system $\Sigma^e(X,W,B,f=(g,h))$ satisfies the condition that there exist no nonzero involutive distribution $D$ on $B$ such that*

$$\begin{aligned} (i) \quad & [\Delta, D] \subset D + \Delta_0, \\ (ii) \quad & D \subset \ker dh. \end{aligned} \tag{72}$$

**Remark 6** It is very surprising that the condition of minimality locally comes down to a condition on the extended system, which is in some sense an infinitesimal version of the original system.

**Remark 7** Actually there is a conceptual algorithm to check local minimality. Define

$$\Delta^{-1}(\Delta_0 + D) := \{\text{vector fields } Z \text{ on } B \mid [\Delta, Z] \subseteq \Delta_0 + D\}.$$

Then we can define the sequence $\{D^{mu}\}, \mu = 0,1,2,\ldots$ as follows:

$$D^0 = \ker dh,$$
$$D^\mu = D^{\mu-1} \cap \Delta^{-1}(\Delta_0 + D^{\mu-1}), \quad \mu = 1,2,\ldots.$$

Then $\{D^\mu\}$, $\mu = 0,1,2,\ldots$, is a decreasing sequence of involutive distributions, and for some $k \geqq \dim(\ker dh) D^k = D^\mu$ for all $\mu \geqq k$. Then $D^k$ is the *maximal* involutive distribution which satisfies

$$\begin{aligned} (i) \quad & [\Delta, D^k] \subset D^k + \Delta_0, \\ (ii) \quad & D^k \subset \ker dh. \end{aligned}$$

From Theorem 11 it follows that $\Sigma(X,W,B,f)$ is locally minimal if and only if $D^k = O$.

**Observability.** It is natural to suppose that our definition of minimality has something to do with controllability and observability. However, because the definition of a nonlinear system (61) also includes autonomous systems, (i.e., no inputs), minimality cannot be expected to imply, in general, some kind of controllability. In fact an autonomous linear system

$$\dot{x} = Ax, \quad y = Cx$$

is easily seen to be minimal if and only if $(A,C)$ is observable. Moreover, it seems natural to define a notion of *observability* only in the case that the system (61) has at least a local input-output representation; i.e., we make the standing assumption that $(\partial h/\partial v)$ is injective (see Lemma 1). Therefore, *locally* we have as our system

$$\dot{x} = g(x, u), \quad y = \overline{h}(x, u) \tag{73}$$

for every possible input-output coordinatization $(y, u)$ of $W$. For such an input-output system local minimality implies the following notion of observability, which we call *local distinguishability.*

**Proposition 2** *Choose a local input-output parametrization as in (73). Then local minimality implies that the only involutive distribution $E$ on $X$ which satisfies*

*i) $[g(\cdot, u), E] \subset E$ for all $u$ ($E$ is invariant under $g(\cdot, u)$),*

*ii) $E \subset \ker\, d_x h(\cdot, u)$ for all $u$ ($d_x \overline{h}$ means differentiation with respect to $x$) is the zero distribution.*

*Proof.* Let $E$ be a distribution on $X$ which satisfies i) and ii). Then we can lift $E$ in a trivial way to a distribution $D$ on $B$ by requiring that the integral manifolds of $D$ be contained in the sections $u = \text{const}$. Then one can see that $D$ satisfies $[\Delta, D] \subset D + \Delta_0$ and $D \subset \ker\, dh$. Hence $D = 0$ and $E = 0$.

**Remark 8** It is easily seen that, under the condition $(\partial h/\partial v)$ injective local minimality. We can state the following Corollary 2.

**Corollary 2** *Suppose there exists an input-output coordinatization*

$$\dot{x} = g(x, u), \quad y = \overline{h}(x). \tag{74}$$

*Then heal minimality implies local weak observability.*

*Proof.* As can be seen from Proposition 2, local minimality in this more restricted case implies that the only involutive distribution $E$ on $X$ which satisfies

$$i) \quad [g(\cdot, u), E] \subset E \quad \text{for all} \quad u,$$
$$ii) \quad E \subset \ker\, d\overline{h}$$

is the zero distribution. It can be seen that the biggest distribution which satisfies i) and ii) is given by the null space of the codistribution $P$ generated by elements of the form

$$L_{g(\cdot, u^1)} L_{g(\cdot, u^2)} \cdots L_{g(\cdot, u^k)} d\overline{h}, \quad \text{with} \quad u^j \text{ arbitrary.}$$

Because this distribution has to be zero, the codistribution $P$ equals $T_x^* X$, in every $\in X$. This is, apart from singularities (which we don't want to consider), equivalent to local weak observability.

Moreover, let (74) be locally weakly observable. Then all feedback transformations $u \mapsto v = \alpha(x, u)$ which leave the form (74) invariant (i.e., $y$ is only the function $x$) are exactly the output feedback transformations $u \mapsto v = \alpha(y, u)$. It can be easily seen in local coordinates that after such output feedback is applied the modified system is still locally weakly observable.

In Proposition 2 and its corollary we have shown that local minimality implies a notion of observability, which generalizes the usual notion of local weak observability. Now we will define a much stronger notion. Let us denote the (defined only locally) vector field $\dot{x} = g(x, \overline{u})$ for fixed $\overline{u}$ by $g^{\overline{u}}$ and the function $\overline{h}(x, \overline{u})$ by $h^{\overline{u}}$ (with $g$ and $\overline{h}$ as in (73)).

**Definition 22** *Let $\Sigma(X, W, B, f) = (g, h)$ be a smooth nonlinear system. It is called strongly observable if for every possible input-output coordinatization (73) the autonomous system*

$$\dot{x} = g^{\overline{u}}(x), \quad y = h^{\overline{u}}(x) \tag{75}$$

*with $\overline{u}$ constant is locally weakly observable, for all $\overline{u}$.*

**Remark 9** Let $\Sigma(X, W, B, f = (g, h))$ be strongly observable. Take one input-output coordinatization $(y, u)$. The system has the form (in these coordinates)

$$\dot{x} = g(x, u), \quad y = \overline{h}(x, u).$$

Because the system is strongly observable, every *constant* input-function (constant in *this* coordinatization) distinguishes between two nearby states. However, in every other input-output coordinatization every constant (i.e., in *this* coordinatization) input function also distinguishes. This implies that in the first coordinatization every $C^\infty$ input function distinguishes. Because the $C^\infty$ input functions are dense in a reasonable set of input functions, every input function in this coordinatization distinguishes.

**Proposition 3** *Consider the Pfaffian system constructed as follows:*

$$P = dh^{\overline{u}} + L_{g^{\overline{u}}} dh^{\overline{u}} + L_{g^{\overline{u}}}(L_{g^{\overline{u}}} dh^{\overline{u}}) + \cdots + L_{g^{\overline{u}}}^{n-1} dh^{\overline{u}},$$

*with $n$ the dimension of $X$ and $L_{g^{\overline{u}}}$ the Lie derivative with respect to $g^{\overline{u}}$. As is well known, the condition that the Pfaffian system $P$ as defined above satisfies the condition $P_x = T_x^* X$ for all $x \in X$ (the so called observability rank condition) implies that the system*

$$\dot{x} = g^{\overline{u}}(x), \quad y = h^{\overline{u}}(x)$$

*is locally weakly observable. Hence, when the observability rank condition is satisfied for all $u$, the system is strongly observable.*

We will call the Pfaffian system $P$ the *observability codistribution.*

**Remark 10** As is known, local weak observability of the system

$$\dot{x} = g^{\overline{u}}(x), \quad y = h^{\overline{u}}(x)$$

implies that the observability rank condition (i.e., $\dim P_x = T_x^* X$) is satisfied almost everywhere (in fact, in the analytic case everywhere). Because we don't want to go into singularity problems, for us local weak observability and the observability rank condition are the same.

**Remark 11** It is easily seen that when for one input-output coordinatization the observability rank condition for all $u$ is satisfied, then for *every* input-output coordinatization the observability rank condition for all $u$ is satisfied. This follows from the fact that the observability rank condition is an open condition.

**Controllability.** The aim of this section is to define a kind of controllability which is "dual" to the definition of local distinguishability (Proposition 2) and which we shall use in the following section. The notion of controllability we shall use is the so-called "strong accessibility".

**Definition 23** *Let $\dot{x} = g(x, u)$ be a nonlinear system in local coordinates. Define $R(T, x_0)$ as the set of points reachable from $x_0$ in exactly time $T$; in other words,*

$$R(T, x_0) := \{x_1 \in X \mid \exists \text{ state trajectory } x(t) \text{ generated by } g \\ \text{such that } \ x(0) = x_0 \ \text{ and } \ x(T) = x_1\}.$$

We call the system *strongly accessible* if for all $x_0 \in X$, and for all $T > 0$ the set $R(T, x_0)$ has a nonempty interior.

For systems of the form (in local coordinates)

$$\dot{x} = f(x) + \sum_{i=1}^{m} u_i g_i(x) \tag{76}$$

(i.e., affine systems) we can define $A$ as the smallest Lie algebra which contains $\{g_1, \ldots, g_m\}$ and which is invariant under $f$ (i.e., $[f, A] \subset A$). It is known that $A_x = T_x X$ for every $x \in X$ implies that the system (76) is strongly accessible. In fact, when the system is analytic, strong accessibility and the rank condition $A_x = T_x X$ for every $x \in X$, are equivalent. We call $A$ the *controllability distribution* and the rank condition the controllability rank condition. Now it is clear that for affine systems (76) this kind of controllability is an elegant "dual" of local weak observability.

It is well known that the extended system (see Definition 20) is an affine system. Hence for this system we can apply the rank condition described above. This makes sense because the strong accessibility of $\Sigma(X, W, B, f)$ is very much related to the strong accessibility of $\Sigma^e(X, W, B, f)$, which can be seen from the following two propositions.

**Proposition 4** *If $\Sigma^e(X, W, B, f = (g, h))$ is strongly accessible, then $\Sigma(X, W, B, f = (g, h))$ is strongly accessible as well.*

*Proof.* In local coordinates the dynamics of $\Sigma^e$ and $\Sigma$ are given by

$$\begin{aligned} I \quad \dot{x} &= g(x, u) \quad (\Sigma), \\ II \quad \dot{x} &= g(x, v) \quad (\Sigma^e), \\ \dot{v} &= u. \end{aligned}$$

It is easy to show that if for $\Sigma^e$ one can steer to a point $x_1$ then the same is possible for $\Sigma$ (even with an input that is smoother).

The converse is harder:

**Proposition 5** *Let $\Sigma(X, W, B, f = (g, h))$ be strongly accessible. In addition if the fibers of $B$ are connected. Then $\Sigma^e(X, W, B,\ f = (g, h))$ is strongly accessible.*

*Proof.* Consider the same representation of $\Sigma$ and $\Sigma^e$ as in the proof of Proposition 4. Let $x_0 \in X$ and $x_1$ be in the (nonempty) interior of $R_\Sigma(x_0, T)$ (the reachable set of system $\Sigma$). Then it is possible to reach $x_1$ from $x_0$ by an input function $v(t)$ which cannot be generated by the differential equation $\dot{v} = u$ . However, we know that the set of the $v$ generated in this way is dense in $L^2$. (For this we certainly need that the fibers of $B$ are connected.) Because we only have to prove that the interior of a set is nonempty, this makes no difference. Now it is obvious from the equations

$$\dot{x} = g(x, v), \quad \dot{v} = u$$

that if we can reach an open set in the $x$-part of the (extended) state, then it is surely possible in the hole $(x, v)$-state.

## 8 Conclusions

In this chapter, the problem of geometrical description of multiple agents is studied. The connection of the optimal game and Yang–Mills fields has been established. A geometric model of a controlled agent as dynamic information-transforming system is examined. A description of the information-transforming system within the framework of the geometric formalism is also proposed. After a classification of the fiber bundle types of conflict and conflict-free maneuvers, a weighted energy can be proposed as the cost function to select the optimal one. Various local and global controllability and observability conditions are derived. For the general multi-agent case, a convex optimization algorithm is proposed to find the optimal multi-legged maneuvers. To completely characterize the optimal conflict-free maneuvers, many issues remain to be addressed. Possible directions of future research include the analysis of the proposed mathematical models in terms of its performance and its robustness with respect to uncertainty of the agents positions and velocities, and a more realistic study for the agent dynamics.

## References


1. Butkovskiy A, Samoilenko Yu (1990) Control of Quantum-Mechanical Processes and Systems, Kluwer Academic Publishers.
2. Brockett R (1981) Nonlinear systems and nonlinear estimation theory, In Hazewinkel M, Willems JC (eds), Stochastic Systems: The Mathematics of Filtering and Identification and Applications, Reidel, Dordrecht.
3. Griffiths P (1983) Exterior Differential Systems and the Calculus of Variatlons, Birkhuser, Boston.
4. Daniel M, Viallet C (1980) The geometrical setting of guage theories of the Yang–Mills type, Reviews of Modern Physics, 52:175–197.
5. Dubrovin B, Novikov S, Fomenko A (1984) Modern Geometry – Methods and Applications, Part. 1, Graduate text in mathematics, Vol. 93. Springer–Verlag, New York.
6. Cressman R (2003) Evolutionary Dynamics and Extensive Form Games, The MIT Press, Massachusetts, London.
7. Isidory A, Krener A (1984) On the svnthesis of linear input-output responses for nonlinear systems, Systems and Control Letters, 4(1):17–22..
8. Isidory A (1995) Nonlinear Control Systems, Springer–Verlag, Berlin.
9. Marcus L (1973) General theory of global dynamics, In Mayne DQ, Brockett RW (eds), Geometric Methods in System Theory, pages 150–158, D. Reidel Publisching Company, Dodrecht – Boston.
10. Mitter P (1980) Geometry of the space of gauge orbits and the Yang–Mills dynamical system. In Hoof GT, et. al. (eds), Recent Developments in Gauge Theories, (Cargese School, Corsica, August 26 – September 8, 1979), Plenum Press, New York.
11. Van der Shaft A (1982) Controllability and observability for affine nonlinear Hamiltonian systems, IEEE Trans. Aut. Cont., AC–27:490–494.
12. Yatsenko V (1984) Dynamic equivalent systems in the solution of some optimal control problems, Avtomatika, No.4, 59–65.