Cooperative Control of Multiple Agents and Search Strategy

Vitaliy A. Yatsenko[1,*], Michael J. Hirsch[2], and Panos M. Pardalos[3]

[1] Space Research Institute of NASU and NSAU, Kiev, 03022, Ukraine
yatsenko@ufl.edu
[2] Raytheon, Inc., Net-Centric Systems, P.O. Box 12248, St. Petersburg, FL, 33733 U.S.A.
michael_j_ hirsch@raytheon.com
[3] Center for Applied Optimization, Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611 U.S.A.
pardalos@ufl.edu

Abstract. This chapter discusses problems dealing with cooperative control of multiple agents moving in a region. An appropriate search strategy for the whole system can be embodied: hierarchical, coordinated, or cooperative. Geometrical and computational aspects of many-target search problems are considered. Nonlinear and bilinear processes of search for moving objects are proposed. Search problems of ecological danger objects and detection of biological and chemical agents using multi-spectral information are also considered.

Multiagent coordination problems are studied in detail. This problem is addressed for a class of targets for which control Lyapunov functions can be derived. We describe such a multiagent system 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 agents 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 bundle representation. In the case of two agents, these optimality conditions allow us to construct optimal maneuvers geometrically.

1 Introduction

Modern game theory basically deals with dynamical systems on smooth manifolds. However, many practical systems like multiple agents do not have such structures. Axiomatic control theories should adequately be reflected, in terms of their internal language of notions and control problems [1]. In terms of these theories, the control structures can make up various hierarchies. According to [2], the most general structure is represented by a controllability-reachability structure

* Corresponding author.

M.J. Hirsch et al. (Eds.): Adv. in Cooper. Ctrl. & Optimization, LNCIS 369, pp. 231–263, 2007.
springerlink.com

over which the optimal control structure is built. This approach regarding the structure of optimal control and Yang–Mills Fields was discussed in [3] and [4].

In this chapter, the multiagent coordination problem is studied. This problem is addressed for a class of targets in which control Lyapunov functions can be found. The main result is a suite of prepositions about formation maintenance, task completion time, and formation velocity. It is also shown how to moderate the requirement that, for each individual target, there exists a control Lyapunov function.

We discuss mathematical aspects of Unified Game Theory (UGT) and the Theory of 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 Figure 1. Each agent of the system can be described by a stochastic or deterministic differential equation with a control. In this chapter, 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 a single model. For example, the interaction between six robots, as seen in Figure 2, can be described by a hierarchical structure. This integrated model allows for solving of controllability, observability, and cooperative control problems.

In Section 2, we demonstrate the power of the satisficing solution methodology for cooperative control problems regarding many-target search. An appropriate search strategy for the whole system can be embodied: hierarchical, coordinated, or cooperative. Geometrical and computational aspects of many-target search problems are considered. In Section 3, we analyze in detail the relationship between gauge fields, identification problems, and control systems. 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 4 focuses on nonlinear control systems and Yang–Mills fields. This section 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 analysis of symmetries. Conclusions are drawn in Section 5.

2 Coordination for Different-Type Objects and Search Strategies

This section is dedicated to the development of methods for solving the problems of interception of multiple mobile targets by a group of unmanned vehicles.

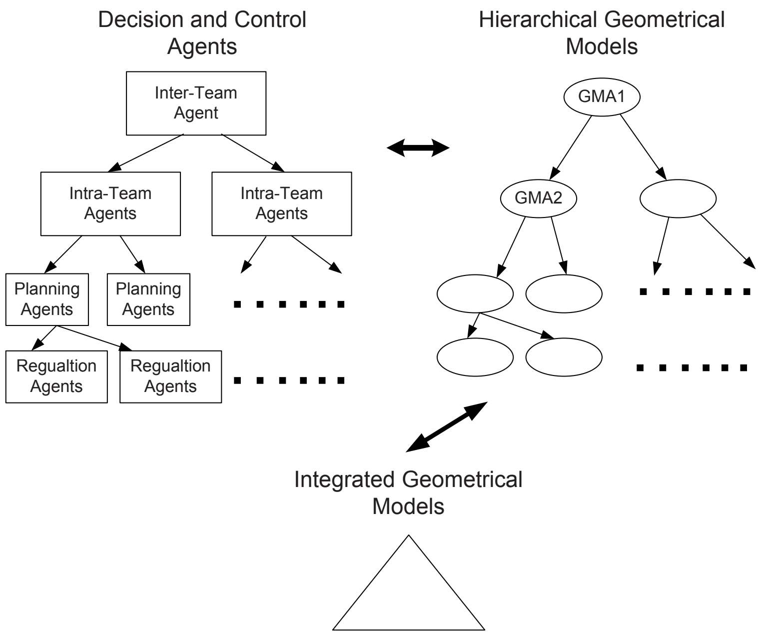


Fig. 1. Hierarchical structure of multiple agents

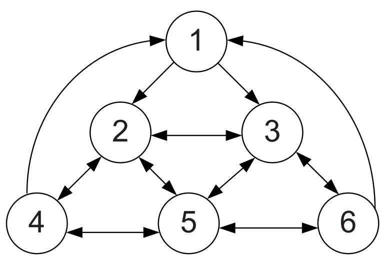


Fig. 2. Hierarchical structure of multiple robot

Emphasis will be paid on the following aspects: Interaction of controlled object groups; Active coordination for different-type objects; Implementation of new pursuit strategies; Investigation of group pursuit problems; and Search strategies.

2.1 Interaction of Controlled Object Groups

Methods and strategies will be devised for interception of multiple mobile targets (evaders), on the basis of the by-interval decomposition principle. This principle assumes that at the initial instant of time the interceptors (pursuers) and the targets are divided into subgroups, each consisting of either multiple pursuers and single target or single pursuer and multiple targets. Such targets' distribution can be performed either on the basis of certain experience or by using discrete optimization methods. As a result, the complicated process of a groups interaction is decomposed into a number of independent subproblems of group or successive pursuit, whereby the term 'group pursuit' is meant the pursuit of a single evader by multiple pursuers, and by the term 'successive pursuit' we mean the pursuit of multiple evaders by a single pursuer.

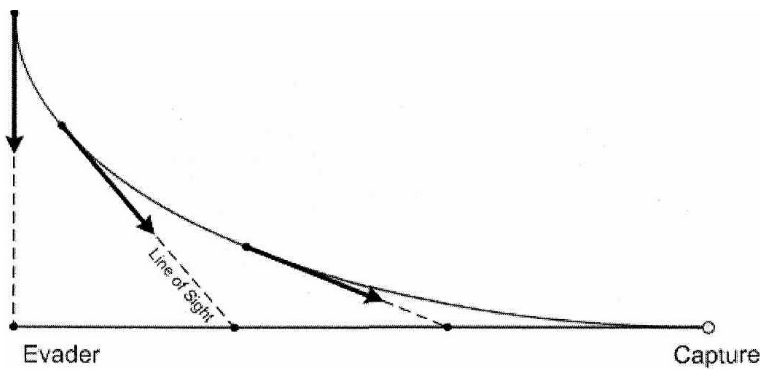


Fig. 3. Pursuit along the 'pursuit curve'

Let us fix the first instant of time when one of the mobile targets is intercepted and therefore can be excluded from further analysis. As a result, the newly freed pursuers can be included into other subgroups. At the instant t_k, let us perform a new decomposition of the pursuers and the remaining targets into subgroups, analogous to the first step. Analyzing the obtained problems of group and successive pursuit, we find the next instant t_{k+1} of interception of next target(s). At the instant t_{k+1}, a new target distribution is performed, and the process repeats.

It this manner, the process of optimization of controlled object interaction is reduced to the iterative procedure, which assumes solving the following typical problems:

1. Target distribution problem.
2. Group pursuit problem.
3. Successive pursuit problem.

The suggested procedure is particularly advantageous for sufficiently large numbers of mobile targets and pursuers, because in this case it reduces a complex original problem into a number of considerably less complicated processes, evolving in parallel, and makes it feasible for computer simulation on parallel computers. The problems of group and successive pursuit can be solved by using the Method of Resolving Functions (MRF) [5,6].

This method proved to be efficient in solving the group pursuit problem. It makes it feasible to study all known (classical) methods of pursuit from the unified standpoint. In particular, the MRF fully substantiates Parallel Pursuit Guidance Law, well known to designers of rocket and aerospace techniques [6].

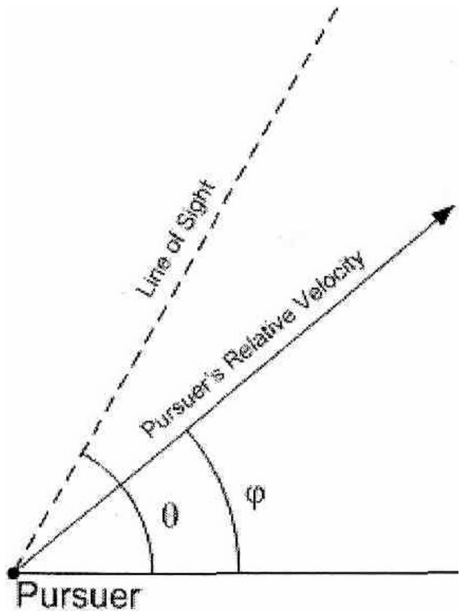


Fig. 4. Proportional navigation guidance law

Line of Sight (LOS) Guidance Law. This strategy has been long known from Euler's time [7]. It implies that at each instant of time the pursuer's velocity is directed along the Line of Sight (LOS) (Figure 3). The problem of finding the form of a trajectory of the pursuer, moving in the plane under the LOS Guidance Law was first formulated by Leonardo da Vinci [7]. It was solved by Pierre Bouguer in 1732 [8]. Despite simplicity in realization, this strategy fails to account for possible mistakes of the evader and frequently yields capture times significantly longer then optimal. The LOS strategy appears as a specific case of the Extremal Targeting Rule (ETR) in [9].

It is possible to formulate ETR in terms of support functions that essentially facilitate constructing the pursuer control and make it feasible to present the latter in explicit form. A modified ETR version for solving the problems of group and successive pursuit is discussed in [5].

Proportional Navigation (PN) Guidance Law. This method is well known to engineers involved in design of aerospace techniques [6]. The basic idea is that the angular velocity of the bearing, φ', varies proportionally with the angular velocity of the LOS, θ', $\varphi' = k\theta'$, where k is a navigation constant. The geometry behind this method is shown on Figure 4.

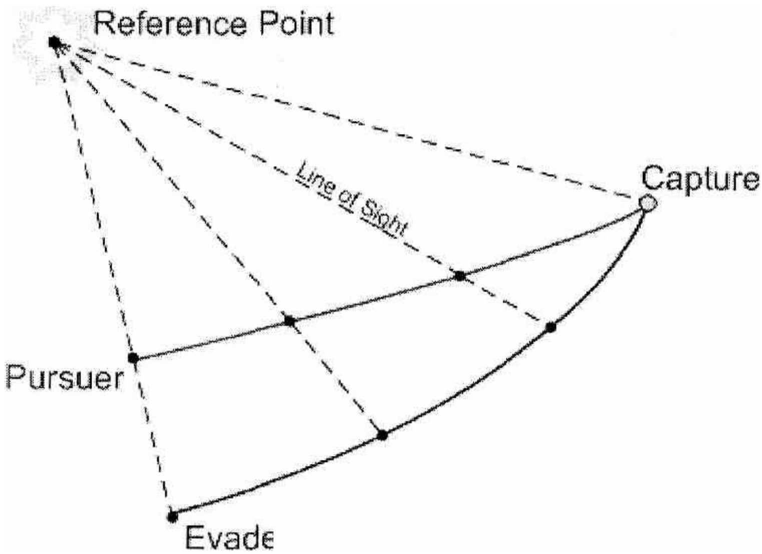


Fig. 5. Parallel pursuit

Note that when $k = 1$, the PN and LOS guidance laws coincide. Furthermore, as k approaches infinity the PN guidance law turns into the strategy of Parallel Pursuit (PP) [10]. The PP strategy implies that the lines of sight are parallel in the course of pursuit.

It is known that, in the case of simple motions, the strategy of parallel pursuit yields the optimal capture time (Figure 5) [5,6]. The MRF approach allows one to extend the ideology of parallel pursuit to wide classes of pursuit problems, where parallel pursuit is meant in a generalized sense. On the MRF basis, the authors have obtained important results concerning both the group and the successive pursuit [6]. In addition, necessary and sufficient conditions for solvability of the group pursuit problem were derived, together with explicit formulas for controlling functions [6].

2.2 Actions Coordination for Different-Type Objects

In the case of unmanned aerial vehicles (UAV) and unmanned ground vehicles (UGV), this problem becomes more complicated in view of state constraints, as one group of objects is moving in the air, while the other in the plane. This difficulty was successfully overcome in solving the problem of soft landing (e.g., airplanes landing on an aircraft carrier) [11].

2.3 Implementation of New Pursuit Strategies

In practice, when pursuing a moving target it is sometimes important for the pursuer not only to rapidly intercept the target, but also to conceal its approach.

This requirement is ensured by the pursuit strategy called *Motion Camouflage* (Figure 6). *Motion camouflage* is observed in insects, especially dragonflies [12].

In this strategy the pursuer camouflages itself against a fixed background object (reference point) so that the evader observes no relative motion between the pursuer and the fixed object (e.g., the sun). The pursuer simply remains on the line between the evader and the reference point, so it seems to be stationary from the evaders perspective [13,14]. If the reference point is at infinity, we obtain the parallel pursuit strategy described above. The motion camouflage strategy can be immediately applied to autonomous system control. For example, low observability behaviors have obvious applications in UAVs and guided missiles.

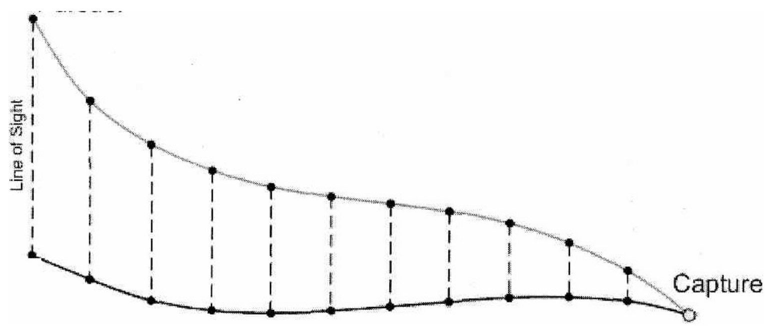


Fig. 6. Motion camouflage

2.4 Investigation of Group Pursuit Problems

The problems of search and observation for mobile targets constitute an important branch of the theory of conflict-controlled processes. Fundamental studies are provided in and [15,16,17,18]. The main feature of such problems is that only information on probability density of the current target state is available to the pursuing object(s), rather than the exact position. Using the probability density evolving according to the Fokker-Planck-Kolmogorov equation [17,18], we developed an approach (cell model of search), which is based on discretization of the search process both temporally and spatially. This process is bilinear and may appear as a Markovian or semi-Markovian chain. The detection probability or the average detection time are utilized as the performance criteria. The Pontryagin's discrete maximum principle and the Bellman dynamic programming method, respectively, were used to optimize the performance criteria. Game problems for the processes, described by the Ito equation, are studied in [18].

The problem of determining sea clutter dynamics and the application of reconstruction methodology in detection and classification of small targets has been considered in [19]. We explore the use of dynamical system techniques, optimization methods and statistical methods to estimate the dynamical characteristics of sea clutters. We assume that radar information is in a form of nonlinear time series. Hence we employ a dynamical approach for characterizing a radar signal, based on nonlinear estimation of dynamical characteristics, by forming a vector of these characteristics, and modeling the evolution of dynamical processes over time.

For the Navy domain, [6] created a decision support system tailored for the search for submarines in various tactical episodes. It is based on the above mentioned scheme of searching for mobile targets. Cases of discrete, continuous, and cyclic search, in their number conducted by a tactical group, were treated, as well as search performed in a hidden way, with the use of contemporary tools (e.g., UAV, UGV). For searches, performed by a tactical group, cases of information exchange within a group and individual search were analyzed. The problem of search for multiple targets was also studied. In this framework, we are planning to apply the achieved theoretical results and gained experience for the creation of methods and algorithms of search for multiple mobile targets by multiple-agent unmanned vehicles (both aerial and ground).

2.5 Cellular Search Model

Let us consider a search region which can be divided into a finite number of cells (states) $i = 1, \ldots, n$. A pursuer in state i at time t is able to move with probability $p_i(t)$, thus

$$p_i(t) \geq 0 \ \forall \ i = 1, \ldots, n \tag{1}$$

$$\sum_{i=1}^{n} p_i(t) = 1, \ \ t = 0, 1, \ldots \ . \tag{2}$$

Denote $p(t) = \Big(p_1(t), \ldots, p_n(t)\Big)$. The dynamics of the pursuer can be described by the discrete differential equation

$$p(t+1) \ = \ U^*(t)p(t), \ \ t = 0, 1, \ldots \ , \tag{3}$$

where $U(t)$ is a stochastic square matrix of order n, and $U^*(t)$ is the conjugate matrix, which play the role of control parameters and satisfy the constraints

$$u_{i,i_1}(t) \geq 0 \ \forall \ i, i_1 = 1, \ldots, n \tag{4}$$

$$\sum_{i_1=1}^{n} u_{i,i_1}(t) = 1 \ \forall \ i = 1, \ldots, n. \tag{5}$$

Suppose that an evasion object can be found in any state $j = 1, \ldots, m$ at time t with probability $q_j(t)$, i.e.,

$$q_j(t) \geq 0 \ \forall \ j = 1, \ldots, m \tag{6}$$

$$\sum_{j=1}^{m} q_j(t) = 1, \ \ t = 0, 1, \ldots \ . \tag{7}$$

Denote $q(t) = \Big(q_1(t), \ldots, q_m(t)\Big)$. The dynamics of the evasion object can be described by the discrete differential equation

$$q(t+1) = V^*(t)q(t),\ t = 0, 1, \ldots\ , \tag{8}$$

where $V(t)$ is a stochastic square matrix of order m, and $V^*(t)$ is the conjugate matrix elements, which play the role of control parameters and satisfy the constraints

$$v_{jj_1}(t) \geq 0,\ \forall\ j, j_1 = 1, \ldots, m \tag{9}$$

$$\sum_{j_1=1}^{m} v_{j,j_1}(t) = 1, \quad j = 1, \ldots, m. \tag{10}$$

The problem of optimal probability detection can be reduced to a conflict control problem of finite state

$$W_0(T) = (c, x(T)),\ c = (0, \ldots, 0, 1) \tag{11}$$

of the bilinear discrete process

$$x(t+1) = A(U(t), V(t))x(t), \quad t = 0, 1, \ldots, \tag{12}$$

where $W_0(T)$ is the probability of detection for time T.

Let r_{ij} be the probability of detection of the evasion object for the ith pursuer state and jth evasion state. Then the joint probability of evasion transition from j to j_1 at the moment t under undetected condition of the evasion object until time t is determined by the equation

$$f(i, i_1, j, j_1) = u_{ii_1}(t)v_{jj_1}(t)(1 - r_{ij}). \tag{13}$$

Denote by $F(u(t), v(t))$ the matrix function of dimension $m \cdot n$ with elements $f(i, i_1, j, j_1, t)$, where $u(t)$ is vector function with n^2 components $\{u_{ii_1}(t)\}$, $v(t)$ is vector function with m^2 components $\{v_{jj_1}(t)\}$.

This problem can be described by the optimization model

$$\begin{aligned} \omega_+ &= \min_{V} \max_{U} W_0(T) \\ &= \min_{V(0)} \max_{U(0)} \ldots \min_{V(T-1)} \max_{U(T-1)} W_0(T) \end{aligned} \tag{14}$$

$$\begin{aligned} \omega_- &= \max_{U} \min_{V} W_0(T) \\ &= \max_{U(0)} \min_{V(0)} \ldots \max_{U(T-1)} \min_{V(T-1)} W_0(T) \end{aligned} \tag{15}$$

where $U = U(0), \ldots, U(T-1)$, $V = V(0), \ldots, V(T-1)$, and $W_0(T)$ is the detection probability in time T.

The mean value of the target detection time is determined by the equation

$$\tau(u, v) = (W(0), N\xi), \tag{16}$$

where $N = \sum_{t=0}^{\infty} F^t(u,v)$, and ξ is the column vector with all components equal to one. It is evident that matrix $N = (1-F(u,v))^{-1}$ exists and can be considered the problem of optimization of the mean target detection time.

The deviation of target detection for fixed control is given by the equation

$$D(u,v) = (W(0), (2N - E)N\xi - (N\xi)_{sq}), \tag{17}$$

where E is the single matrix and $(N\xi)_{sq}$ is the vector with components which equal the square of components of vector $N\xi$.

3 Geometrical Aspects of Multiagent Coordination

Investigations of controlled multiagent objects 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 is the differential geometry methods of system theory [20,21]. 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 [20,21] which is based on the lamination concept, we propose a geometric method for "identifying" its hidden structure.

Investigation of geometrical aspects of multiagent coordination is one of the most essential stages in the creation of new strategies. The goal of the experimental and theoretical research is the implementation of optimal strategy using complex structure non-equilibrium 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 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 (e.g., 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

[22], 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 [23]. Continuous symmetry group dynamic systems were considered by [20]. While substantive results on the decomposability of systems with symmetries have been obtained [24], this question remains open for DITAs.

In this section, we investigate the problem of how to coordinate a collection of targets in such a way that they maintain a given formation relative to each other. The main assumption about the dynamics of the individual robots that we initially make in this paper is that they have control Lyapunov functions (CLFs). Based on this assumption, an abstract and theoretically sound coordination strategy can be developed.

3.1 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, connectedness, and distributions are given in [25]. We introduce the definition of a nonlinear DITA.

Definition 1. *We refer a triple, $F(B, M, \psi)$, where B is a smooth fiber over M with the projection $\pi : B \to M$; π_M is the natural projection of TM on M; and ψ is a smooth mapping such that the diagram presented in Figure 7 is commutative, by a ''geometrical model of the agent''.*

We interpret the M manifold as the 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 agent, F, corresponds locally to the nonlinear transformation $\psi : (x, u) \to (x, \psi(x, u))$ and the dynamic system

$$\dot{x}(t) = \psi(x(t), u(t)), \quad u(t) \in U, \tag{18}$$

where x is the 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 agents 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.

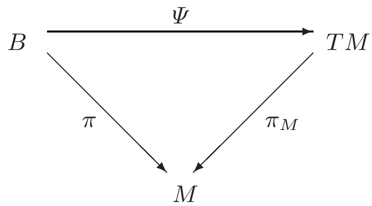


Fig. 7. Diagram of a nonlinear controlled DITA

Our primary object of study is a collection of targets, whose dynamics can be described by the following set of controlled differential equations:

$$\dot{x_i}(t) = \psi(x(t), u(t)) = f_i(x_i) + g_i(x_i)u_i, \quad i = 1, \ldots, n, \tag{19}$$

where $f_i, g_i \in C^\infty$, $x_i \in \mathbb{R}^n$, and $u_i = \mathbb{R}^{p_i}$. Now, a desired formation in $\mathbb{R}^{nm}$ is simply a set $(x_{10}, \ldots, x_{m0}) \in \mathbb{R}^{nm}$, and we define this set implicitly through the null set of a so-called formation function.

3.2 Coordinated Control

By using the Lyapunov formation functions derived from the individual target, we can now shift our attention to actually controlling the evolution of the formation. The one parameter that we can control is the time evolution of the desired positions. We do this by specifying the trajectory that we want the so-called virtual leader, $x_0(s(t))$, to follow.

This nonphysical leader is a reference point in the state space with respect to which we can define the rest of the formation. We denote the trajectory executed by the virtual leader as $x_0(s(t)) = p(s(t))$. Intuitively, one might want to set $s(t) = t$. But, due to robustness considerations, we incorporate error feedback into the time evolution of s and let $\dot{s}$ be given by

$$\dot{s} = \min\Big[\frac{v_0}{\delta + \|\frac{\partial p(s)}{\partial s}\|}, \frac{-(\frac{\partial F}{\partial x})^T}{\delta + |\frac{\partial F}{\partial s}|}\big[\frac{\sigma(F_U)}{\sigma(F(s,x))}\big]\Big]. \tag{20}$$

Here, $\delta > 0$ is a small positive constant that prevents $\dot{s}$ from becoming singular, and F_U is the bound or something smaller chosen by the user. It can be shown to be an upper bound on the Lyapunov formation function $F(s, x)$. The idea is that the formation is being respected as long as $F(s, x) \leq F_U$. Furthermore, v_0 is the nominal velocity that we want the formation to move with, and it holds that $\|\dot{x_0}(s(t))\| \approx m_0$ when small.

3.3 Symmetry of Multiagent Coordination

Definition 2. *Let M be a smooth manifold. We say that the smooth mapping $Q : G \times M \to M$ such that: i) $Q(e, x) = x$ for all $x \in M$, and ii) $Q(g, Q(h, x)) = Q(gh, x)$ for any $g, h \in G$, 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 $Q_x : G \to M$ denote the function $g \mapsto Q(g, x)$. We note that since $(Q_g)^{-1} = Q_g^{-1}$, Q_g is a diffeomorphism. We introduce the definition of group action on a manifold.

Definition 3. *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 $g \mapsto Q_g(x)$ 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 4. *Let $\hat{F}(B, M, \psi)$ be a nonlinear controlled DITA, and θ and Q be actions of G on B and M, respectively. Then, F has symmetry (G, θ, Q) if the diagram presented in Figure 8 is commutative for all $g \in G$.*

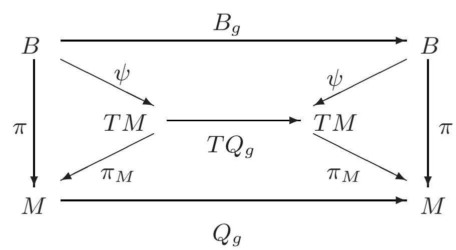


Fig. 8. A commutative diagram of an DITA with symmetries

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

Definition 5. *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, θ, 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 a DITA B_x, 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 the definition of local symmetry.

Definition 6. *We assume that $Q : G \times M \to M$ is an action and that $\varepsilon \in T_e G$. Then, $Q^\xi(R \times M \to M) : (t, x) \mapsto Q(\exp t\xi, x)$, where $\exp : T_e G \to G$ is the usual exponential mapping, is the $\mathbb{R}$-action on M, and Q^ξ is the complete flow on M. We say that the corresponding vector field on M, which is defined by the expression*

$$\xi_m(x) = \frac{d}{dt} Q(\exp t\xi, x)\Big|_{t=0}, \tag{21}$$

is the infinitesimal action generator, which corresponds to ξ.

Let X_t denote the flow of the vector field X, i.e., $X_t = F_t(X_0)$. It is obvious from the definition of the infinitesimal generator that if (G, θ, Q) is a symmetry of the $\hat{F}(B, M, \psi)$ system, then the diagram presented in Figure 9 is commutative for all $t \in \mathbb{R}$ and $\xi \in T_eG$.

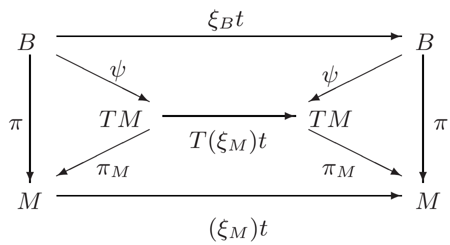


Fig. 9. Diagram of a symmetric DITA

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

Definition 7. *Let $\hat{F}(B, M, \psi)$ be a nonlinear DITA. Then, (G, θ, 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*

$$(\xi_M)_t * \psi(\xi) = \psi((\xi_b)_t(b)), \tag{22}$$

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.

One can define an infinitely small state space symmetry for nontrivial laminations of the input actions manifold when one can introduce integrable connectivity. For this we introduce Definition 8.

Definition 8. *Let $H(\cdot)$ be an integrable connectivity on B and (G, θ, Q) be a symmetry of F. Then, (G, θ, 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 θ are horizontal.*

We introduce a definition of feedback equivalence of two DITAs in analogy with [20].

Definition 9. *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 Figure 10 is commutative.*

Isomorphism means that, for $x \in M$, γ_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".

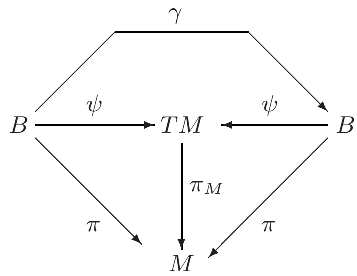


Fig. 10. Diagram of feedback-equivalent DITAs

3.4 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.

In addition, let $C = \{h \in G \mid hg = gh \ \forall \ g \in G\}$ 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 $1 \le i \le l$ and $1 \le j \le k$.

Using the results in [20,26] that deal with the properties of systems with symmetries as applied to DITAs, one can formulate the following theorems.

Theorem 1. *Let us assume that $\hat{F}(B, M, \xi)$ is a controlled DITA with an infinitesimal state space symmetry, (G, θ, 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 with infinitesimal state space symmetries, one can propose the structure of the decomposed system. It suffices to demonstrate that the decomposed system with infinitesimal symmetry is locally feedback-equivalent to the original system with infinitesimal state space symmetry.

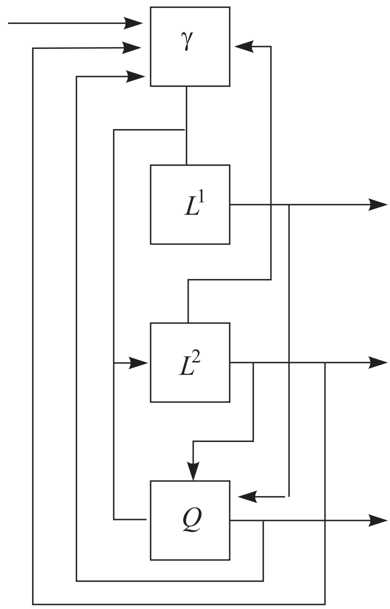


Fig. 11. Local structure of DITA with infinitesimal symmetries

Definition 10. *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 2. *Let $\hat{F}(B, M, \psi)$ have an infinitesimal symmetry (G, θ, Q) and Q be nonsingular at the point m. There exists a neighborhood of m and a system F with infinitesimal symmetry (G, θ, Q) such that $\hat{F}|\hat{O}$ is feedback equivalent to the system $\hat{F}$.*

Let $\hat{F}(B, M, \psi)$ be a controlled DITA with symmetry (G, θ, 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 Figure 11, where γ is the feedback function, the L^i are nonlinear subsystems of dimensions $n-k$ and $k-l$, respectively, and Q is an l-dimensional "quadrature" system

$$\dot{x}_i = f_i(x_1, \ldots, x_{n-k}, u), \quad i = 1, \ldots, n-k \tag{23}$$

$$\dot{x}_j = f_j(x_1, \ldots, x_{n-1}, u), \quad i = n-k+1, \ldots, k. \tag{24}$$

3.5 The Global Structure of DITA

The decomposability of a 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 11. *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 $p \cdot \sigma$ 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 3. *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 σ, then $\hat{F}$ is isomorphic to the system*

$$\dot{y} = \Psi(y, u) \tag{25}$$

$$\dot{g} = (T_e L_g)(T_e Q_{\sigma(y)})^{-1} \left[\Psi(\sigma(y), u) - (T_y \sigma)\Psi(y, u)\right], \tag{26}$$

defined on $M/G \times G$.

Assertion 1. *Let the DITA $F(M \times U, M, \psi)$ have a symmetry (G, θ, 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 σ.*

Combining Theorem 3 and Assertion 1, we obtain the following corollary:

Corollary 1. *Let DITA $\hat{F}(M \times U, M, \psi)$ have a symmetry (G, θ, 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.*

3.6 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

$$\dot{x}(t) = \psi(x(t), u(t)). \tag{27}$$

One can define for Equation (27) 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

$$\dot{x}(t) = f(x) + \sum_{i=1}^{k} u_i g_i(x). \tag{28}$$

One can formulate 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, θ, Q), local commutativity of the diagram means that $\psi * \varepsilon_B = \varepsilon_m$ and $\pi * \varepsilon_B = \varepsilon_n$. Let $\Delta_B = \text{span}\{\varepsilon \mid \varepsilon_B \in T_e G\}$ and the same hold for Δ_m. Then, $\psi * \Delta B \subset \Delta_m$, $\pi * \Delta_B = \Delta$, and Δ_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 groups, 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 a 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.

4 Fiber Bundles and Observability

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

$$\dot{x} = g(x, u), \quad y = h(x), \tag{29}$$

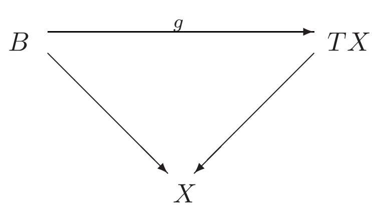
where x is the state of the system, u is the input, and y is the output. Most of the attention has been directed to the formulation in this context of fundamental system theoretic concepts like controllability, observability, minimality, and realization theory.

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

1. Normally, the equations

$$\dot{x} = g(x, u) \tag{30}$$

are interpreted as a family of vector fields on a manifold parameterized 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 (30) as a coordinization of the following diagram.



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).

2. The "usual" definition of *observability* has some drawbacks. In fact, observability is defined as *distinguishable*; i.e., for every $x_1, x_2 \in X$, 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.

3. In the class of nonlinear systems (29), *memoryless* systems

$$y = h(u) \tag{31}$$

are not included. Of course, one could extend the system (29) to the form

$$\dot{x} = g(x, u), \quad y = h(x, u), \tag{32}$$

but this gives, if one wants to regard observability as distinguishability, the following rather complicated notion of observability. As can be seen, distinguishability of (32) with $y \in \mathbb{R}^p$, $u \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$ is equivalent to distinguishability of

$$\dot{x} = g(x, u), \quad \overline{y} = \overline{h}(x), \tag{33}$$

where $\overline{h} : \mathbb{R}^n \to (\mathbb{R}^p)^{\mathbb{R}^m}$ is defined by $\overline{h}(x)(u) = h(x, u)$. Also, checking the Lie algebra conditions for distinguishability for the system (33) is not very easy.

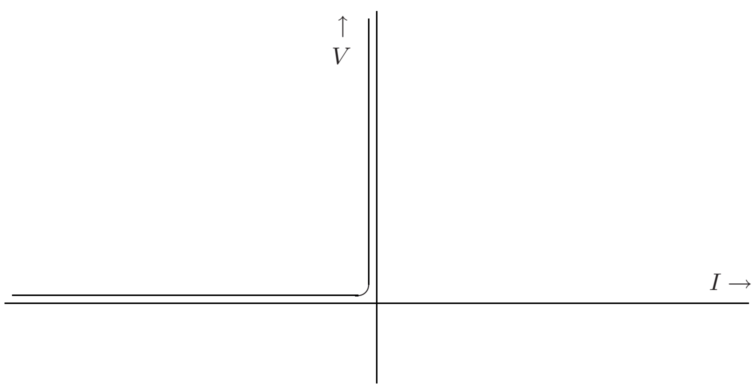


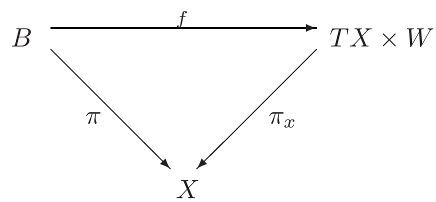
Fig. 12. Ideal diode for the $I - V$ characteristic

4. 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 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 Figure 12. 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. 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.

4.1 Nonlinear Model of Agents

The (say C^∞) agents can be represented in the commutative diagram



(34)

where (all spaces are smooth manifolds) B is a fiber bundle above X with projection π, TX is the tangent bundle of X, π_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 X represents the space of inputs (to be seen initially as *dummy* variables), which is state dependent e.g., 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

$$f(\pi^{-1}(x)) \subset T_x X \times W. \tag{35}$$

We denote the system (34) 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 as $f = (g, h)$, the system is given by

$$\dot{x} = g(x, v), \quad w = h(x, v). \tag{36}$$

Of course one should ask 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 (34) we shall distinguish among three important situations, of which the last is equivalent to the "usual" interpretation of system (29).

1. 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 B, be an immersion into W. Let $(\overline{x}, \overline{v})$ 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* $(x, v) \in B$, *coordinates* $(w_1, w_2) \in W$ *and a map* $\overline{h}$ *such that* h *has the form*

$$(x, v) \gg h > (w_1, w_2) = (\overline{h}(x, v), v). \tag{37}$$

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 (36) has the form, only locally,

$$\dot{x} = y(x, u), \quad y = \overline{h}(x, u). \tag{38}$$

2. Now we not only assume that $\partial h/\partial v$ is injective, which results in a *local* input-output parametrization (38), 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, so that $p : W \to Y$, 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* $x \in X$ *around* $\overline{x}$ *and coordinates* $y \in Y$ *around* $\overline{y}$. *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*

$$(x, v) \gg h > (y, u) = (\overline{h}(x), v). \tag{39}$$

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

$$(x, \overline{v}) \gg h > (y, u) = (\overline{h}(x), h'(x, \overline{v})), \tag{40}$$

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

$$v = (h')^{-1}(x, u) \quad \text{with } x \text{ fixed}. \tag{41}$$

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

$$(x, v) \gg h > (y, u) = (\overline{h}(x), u). \tag{42}$$

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

$$\dot{x} = g(x, u), \quad y = \overline{h}(x). \tag{43}$$

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.

3. Take the same assumptions as in 2 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

$$\dot{x} = g(x, u), \quad y = \overline{h}(x), \tag{44}$$

where for each fixed $\overline{u}, g(\cdot, \overline{u})$ is a globally defined vector field on X. This is the "usual" interpretation of (29).

Some remarks are in order:

Remark 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.*

Remark 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 is beyond the scope of this chapter.*

Remark 3. *From Lemma 2 it is clear that the coordinization of the fibers of the bundle W uniquely determines, via h, the coordinization of the fibers of B. It should be remarked that a coordinization of the fibers of W is locally equivalent to the existence of an (integrable)* connection *on the bundle W, and that one coordinization is linked to another by what is essentially an output feedback transformation, i.e., a bundle isomorphism from W into itself.*

Remark 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 ρ, then B is just ρ^*T^*Q (the pullback bundle via ρ). 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*

$$\ddot{q} = F(q, \dot{q}) + \sum_j u_j Z_j(q, \dot{q}), \quad y = q, \tag{45}$$

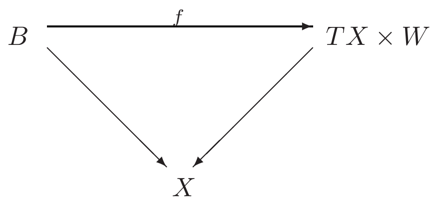
with the vector fields $F(q, \dot{q})$ and $Z_j(q, \dot{q})$ satisfying certain conditions. Moreover the vector fields Z_j commute, i.e., $[Z_i, Z_j] = 0$ for all i, j, a fact which has a very interesting interpretation.

Remark 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.*

4.2 Minimality

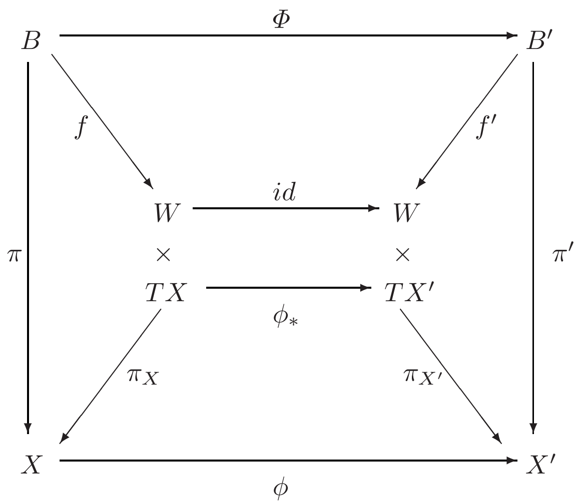
We want to give a definition of minimality for a general nonlinear agent.

Definition 12. *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 following diagram commutes.*



(46)

Σ is called *equivalent to* Σ' (denoted $\Sigma \sim \Sigma'$) if ϕ and Φ are diffeomorphisms. We call Σ *minimal* if $\Sigma' \leqq \Sigma \Rightarrow \Sigma' \sim \Sigma$.



Remark 6. *This definition formalizes the idea that we call Σ'* less complicated *than Σ ($\Sigma' \leqq \Sigma$) if Σ' consists of a set of trajectories in the state space, smaller than the set of trajectories of Σ, but which generates the same* external behavior. *(The external behavior Σ_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 $\Sigma := \{(x, w) : \mathbb{R} \to X \times W|_x$ that are absolutely continuous and $(\dot{x}(t), w(t)) \inf(\pi^{-1}(x(t)))$ a.e.$\}$, then Σ_e is just the projection of Σ on $W^{\mathbb{R}}$).*

Remark 7. *Notice that we only formalize the* regular *case by asking that Φ and ϕ be surjective as well as submersive. In fact we could, for instance, allow that at isolated points ϕ or Φ are not submersive. However, we do not discuss this problem here, and treat only the regular case as described in Definition 12.*

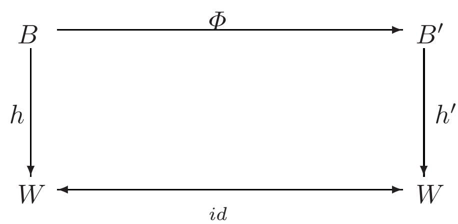
Remark 8. *Notice that $\Sigma_1 \leqq \Sigma_2$ and $\Sigma_2 \leqq \Sigma_1$ need not imply $\Sigma_1 \sim \Sigma_2$. This fact leads to very interesting problems, which again are out of scope for this chapter.*

Of course, Definition 12 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 5 that locally these conditions of commutativity do have a very direct interpretation. But first we have to state some preparatory lemmas and theorems.

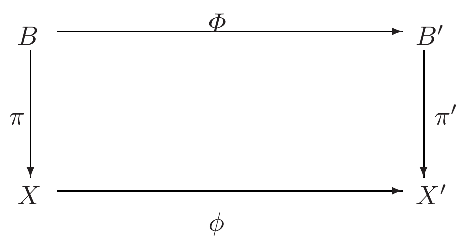
Let us look at Diagram (46). Because Φ is a submersion it induces an involutive distribution D on B given by

$$D := \{Z \in TB | \Phi_* \dot{Z} = 0\} \tag{47}$$

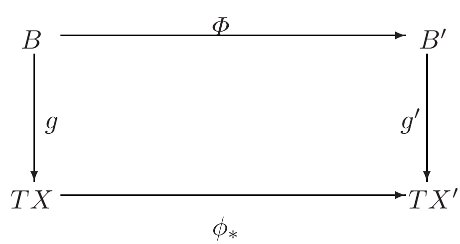
(the foliation generated by D is of the form $\Phi^{-1}(c)$ with c constant). In the same way ϕ induces an involutive distribution E on X. Now the information in the diagram (46) is contained in three subdiagrams (we assume $f = (g, h)$ and $f' = (g', h')$):



I



II



III

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

$$I' : \quad D \subset \ker\, dh \tag{48}$$

$$II' : \quad \pi_* D = E \tag{49}$$

$$III' : \quad g_* D \subset TE = T\pi_*(D). \tag{50}$$

Proof. I' and II' are trivial. For III' observe that, when ϕ induces a distribution E on X, then ϕ_* 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

$$\dot{x} = f(x) + \sum_{i=1}^{m} u_i g_i(x) \quad \text{on a manifold } X. \tag{51}$$

We can interpret this as an affine distribution on manifold.

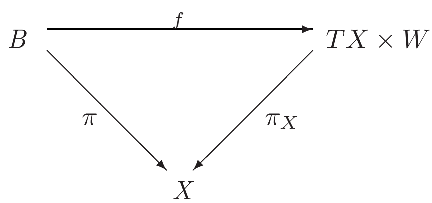
Theorem 4. *Let $D \in A(\varDelta_0)$. Then the condition*

$$[\varDelta, D] \subseteq D + \varDelta_0 \tag{52}$$

(we call such a $D \in A(\varDelta_0)\varDelta(\mathrm{mod}\, \varDelta_0)$ invariant) is equivalent to the two conditions: a) there exists a vector field $F \in \varDelta$ such that $[F, D] \subseteq D$ and b) there exist vector fields $B_i \in \varDelta_0$ such that span $\{B_i\} = \varDelta_0$ 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 13. (Extended system). *Let*



Then we define the extended system of $\Sigma(X, W, B, f)$ as follows: We define $\varDelta_0$ as the vertical tangent space of B, i.e.,

$$\varDelta_0 := \{Z \in TB | \pi_* Z = 0\}. \tag{53}$$

Note that $\varDelta_0$ 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

$$\varDelta(\overline{x}, \overline{v}) := \{Z \in T_{(\overline{x},\overline{v})} | \pi_* Z = g(\overline{x}, \overline{v})\}. \tag{54}$$

So $\varDelta(\overline{x}, \overline{v})$ consists of the possible lifts of $g(\overline{x}, \overline{v})$ in $(\overline{x}, \overline{v})$. Then it is easy to see that $\varDelta$ is an affine distribution on B, and that $\varDelta - \varDelta = \varDelta_0$. We call the affine system $(\varDelta, \varDelta_0)$ 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. *i) 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. ii) Let D be an involutive distribution on B and let $D \cap \Delta_0$ have constant dimension. Then the following two conditions are equivalent: a)$\pi_* D$ is a well-defined and involutive distribution on X, and $g_* D \subset T\pi_* D$ and b) $[\Delta, D] \subset D + \Delta_0$.*

Proof. i) Let $D + \Delta_0$ be involutive. Because D and Δ_0 are involutive this is equivalent to $[D, \Delta_0] \subset D + \Delta_0$. Applying Theorem 4 to this case gives a basis $\{Z_1, \ldots, Z_k\}$ of D such that $[Z_i, \Delta_0] \subseteq \Delta_0$. 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.

ii) 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

$$\left(\frac{\partial g}{\partial x_i}\right)_{j^e \text{comp}} = 0 \tag{55}$$

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 14. (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 12 that minimality of $\Sigma(X, W, B, f)$ *locally* implies local minimality (locally every involutive distribution can be factored out). Combining Lemma 3, Definition 13 and Lemma 4 we can state:

Theorem 5. *$\Sigma(X, W, B, f = (g, h))$ 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*

$$i) \quad [\Delta, D] \subset D + \Delta_0, \tag{56}$$

$$ii) \quad D \subset \ker dh. \tag{57}$$

Remark 9. *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 10. *Actually there is a conceptual algorithm to check local minimality. Define*

$$\Delta^{-1}(\Delta_0 + D) := \{\textit{vector fields } Z \textit{ on } B \mid [\Delta, Z] \subseteq \Delta_0 + D\}. \tag{58}$$

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

$$D^0 = \ker dh, \tag{59}$$

$$D^\mu = D^{\mu-1} \cap \Delta^{-1}(\Delta_0 + D^{\mu-1}), \quad \mu = 1, 2, \ldots. \tag{60}$$

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

$$i) \qquad [\Delta, D^k] \subset D^k + \Delta_0, \tag{61}$$

$$ii) \quad D^k \subset \ker dh. \tag{62}$$

From Theorem 5 it follows that $\Sigma(X, W, B, f)$ is locally minimal if and only if $D^k = O$.

4.3 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 (34) 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

$$\dot{x} = Ax, \quad y = Cx \tag{63}$$

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 (34) 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

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

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

Proposition 1. *Choose a local input-output parametrization as in (64). 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 $g(\cdot, u)$) and 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 $E = 0$.

Corollary 2. *Suppose there exists an input-output coordinization*

$$\dot{x} = g(x, u), \quad y = \overline{h}(x). \tag{65}$$

Then local minimality implies local weak observability.

Proof. As can be seen from Proposition 1, local minimality in this more restricted case implies that the only involutive distribution E on X which satisfies *i)* $[g(\cdot, u), E] \subset E$ for all u and *ii)*$E \subset \ker d\overline{h}$, 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

$$L_{g(\cdot,u^1)} L_{g(\cdot,u^2)} \cdots L_{g(\cdot,u^k)} d\overline{h}, \quad \text{with } \ u^j \text{ arbitrary.} \tag{66}$$

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 (65) be locally weakly observable. Then all feedback transformations $u \mapsto v = \alpha(x, u)$ which leave the form (65) 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 1 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 (64)).

Definition 15. *Let $\Sigma(X, W, B, f) = (g, h)$ be a smooth nonlinear system. It is called strongly observable if for every possible input-output coordinization (64) the autonomous system*

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

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

Remark 11. *Let $\Sigma(X, W, B, f = (g, h))$ be strongly observable. Take one input-output coordinization (y, u). The system has the form (in these coordinates)*

$$\dot{x} = g(x, u), \quad y = \overline{h}(x, u). \tag{68}$$

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

Proposition 2. *Consider the Pfaffian system constructed as follows:*

$$P = dh^{\overline{u}} + L_{g^{\overline{u}}} dh^{\overline{u}} + L_{g^{\overline{u}}}(L_{g^{\overline{u}}} dh^{\overline{u}}) + \cdots + L_{g^{\overline{u}}}^{n-1} dh^{\overline{u}}, \tag{69}$$

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

$$\dot{x} = g^{\overline{u}}(x), \quad y = h^{\overline{u}}(x) \tag{70}$$

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 12. *As is known, local weak observability of the system*

$$\dot{x} = g^{\overline{u}}(x), \quad y = h^{\overline{u}}(x) \tag{71}$$

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

4.4 Controllability

The aim of this section is to define a kind of controllability which is "dual" to the definition of local distinguishability (Proposition 1). The notion of controllability we shall use is the so-called "strong accessibility".

Definition 16. *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,*

$$\begin{aligned} R(T, x_0) := \{x_1 \in X \mid \exists \text{ state trajectory } x(t) \text{ generated by } g \\ \ni \ x(0) = x_0 \wedge x(T) = x_1\}. \end{aligned} \tag{72}$$

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)

$$\dot{x} = f(x) + \sum_{i=1}^{m} u_i g_i(x) \tag{73}$$

(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 (73) 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 (73) this kind of controllability is an elegant "dual" of local weak observability.

It is well known that the extended system (see Definition 13) 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 3. *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 Σ^e and Σ are given by

$$I \quad \dot{x} = g(x, u) \quad (\Sigma), \tag{74}$$

$$II \quad \dot{x} = g(x, v) \quad (\Sigma^e), \tag{75}$$

$$\dot{v} = u. \tag{76}$$

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

The converse is more difficult to prove:

Proposition 4. *Let $\Sigma(X, W, B, f = (g, h))$ be strongly accessible. In addition, let the fibers of B be connected. Then $\Sigma^e(X, W, B, f = (g, h))$ is strongly accessible.*

Proof. Consider the same representation of Σ and Σ^e as in the proof of Proposition 3. Let $x_0 \in X$ and x_1 be in the (nonempty) interior of $R_\Sigma(x_0, T)$ (the reachable set of system Σ). 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

$$\dot{x} = g(x, v), \quad \dot{v} = u \tag{77}$$

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

5 Conclusions

In this chapter, cooperative control of multiple agents was studied. Methods and algorithms were explored for solving the problem of vehicle group interaction, when one group of vehicles is moving in a plane (UGV) and another in a halfspace (UAV-s). We have already analyzed an analogous situation, when one object (a pursuer) is moving in a halfspace while the other (an evader) - in a plane, in solving the problem of ‘‘soft meeting’’. Nonlinear and bilinear Markovian models are proposed for solution of the game theoretic problem of searching for a moving object in discrete time over a finite set of states.

The multiagent coordination problem has been studied. This problem is addressed for a class of targets for which control Lyapunov functions can be found. The main result is a suite of propositions about formation maintenance, task completion time, and formation velocity. It is also shown how to moderate the requirement that, for each individual target, there exists a control Lyapunov function.

The connection between cooperative control and Yang–Mills fields has been established. A geometric model of a controlled agent as dynamic information-transforming system was examined. A description of the information-transforming system within the framework of the geometric formalism was 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. Summing up, we can say that the combined problems of ‘‘search and tracking’’ and ‘‘pursuit and evasion’’ for multiple different-type pursuing objects and multiple evaders will be solved in the next step.

References

1. Cressman, R.: Evolutionary Dynamics and Extensive Form Games. MIT Press (2003)
2. Kalman, R., Falb, S., Arbib, M.: Topics in Mathematical System Theory. McGraw Hall (1969)
3. Yatsenko, V.: Control systems and fiber bundles. Avtomatika **5** (1985) 25–28
4. Butkovskiy, A., Samoilenko, Y.: Control of Quantum-Mechanical Processes and Systems. Kluwer Academic Publishers (1990)

5. Chikrii, A.: Differential games with multiple pursuers. In: of the Banach International Mathematical Centre. Volume 14. (1985) 81–107
6. Chikrii, A.: Conflict-Controlled Processes. Kluwer Academic Publishers (1997)
7. Gall, S.: Search Games. Academic Press (1980)
8. Larrie, F.: Ships and Science: The Birth of Naval Architecture in the Scientific Revolution, 1600-1800. MIT Press (2007)
9. Krasovskii, A., Subbotin, A.: Game-theoretical control problems. Springer-Verlag (1988)
10. Imado, F.: The features of optimal avoidance in two dimensional pursuit-evasion dynamic games. Technical report, Information, Technology, and Management (ITEM) (2002)
11. Albus, J., Meystel, A., Chikrii, A., Belousov, A.: Analytic method for solution of the game problem of soft landing for moving objects. Cybernetics and Systems Analysis **37**(1) (2001) 75–91
12. Mizutani, A., Chahl, J., Srinivasan, M.: Insect behavior: Motion camouflage in dragonflies. Nature **423**(6940) (2003) 604
13. Matichin, A., Chikriy, A.: Motion camouflage in differential games of pursuit. Journal of Automation and Information Science **37**(3) (2003) 1–5
14. Matichin, I., Chikriy, A.: Motion camouflage in differential games of pursuit (in Russian). Journal of Automation and Information Science **37**(3) (2005) 1–5
15. Koopman, B.: The theory of search. Operations Research **4** (1956) 324–346
16. Stone, L.: Theory of optimal search. Academic Press (1975)
17. Hellman, O.: Optimization of search for an object drifting in outer space. Journal of Spacecraft and Rockets **7**(7) (1970) 886–889
18. Hellman, O.: On the optimal search for a randomly moving target. SIAM Journal on Applied Mathematics **22**(4) (1972) 545–552
19. Pardalos, P., Yatsenko, V., Grundel, D.: Nonlinear dynamics of sea clutter and detection of small targets. In Pardalos, P., Murphey, R., Butenko, S., eds.: Recent Developments in Cooperative Control and Optimization. Kluwer Academic Publishers (2004) 407–426
20. Van der Shaft, A.: Controllability and observability for affine nonlinear hamiltonian systems. IEEE Transactions on Automatic Control **27** (1982) 490–494
21. Van der Shaft, A.: Equations of motion for Hamiltonian systems with constraints. Journal of Physics A **11** (1987) 3271–3277
22. Arnold, V.: Mathematical Methods of Classical Methods. Springer–Verlag (1983)
23. Samoilenko, Y.: Reduction to the elementary cell of the linear system with discrete symmetry. Cybernetics and Computer Techniques **3** (1970) 48–53
24. Krener, A.: On the equivalence of control systems and the linearization of nonlinear systems. SIAM Journal of Control **11** (1973) 670–676
25. Griffiths, P., Coates, J., Helgason, S.: Exterior Differential Systems and the Calculus of Variations. Birkhauser Verlag (1983)
26. Marcus, L.: General theory of global dynamics. In Mayne, D., Brockett, R.W., eds.: Geometric Methods in System Theory. (1973) 150–158