united nations educational, scientific and cultural organization

international atomic energy agency

the
abdus salam
international centre for theoretical physics

SMR1327/3

Summer School on Mathematical Control Theory

(3 - 28 September 2001)

Introduction to Geometric Nonlinear Control I

Bronislaw Jakubczyk
Institute of Mathematics
Polish Academy of Sciences
Sniadeckich 8
00-950 Warsaw
Poland

Introduction to Geometric Nonlinear Control I *

Bronislaw Jakubczyk

Institute of Mathematics
Polish Academy of Sciences
00-950 Warsaw, Śniadeckich 8, Poland

*Lecture notes of a minicourse to be delivered jointly with W. Respondek at the Summer School on Mathematical Control Theory, International Centre for Theoretical Physics, Trieste-Miramare, September 2001.

1 Controllability and Lie bracket

Controllability properties of a control system are properties related to the following questions. (Q1) Can the system be steered from a given initial state x_0 to a given final state x_1? (Q2) Can this be done for any pair of initial and final states? (Q3) How large is the the set of points to which the system can be steered from a given initial state x_0? (Q4) Which trajectories of the system are realizable and how do we find controls realizing them?

Such questions can be motivated by practical problems and they are basic for any qualitative study of control systems. Our aim in these lectures will be to develope tools which will anable us to answer such questions and to understand qualitative properties of nonlinear control systems. We will see that for a large class of problems a control system can be represented by a family of vector fields (dynamical systems). The qualitative properties of the control system depend on the properties of the vector fields (dynamical systems) and interactions between them. The basic tool which will anable us to understand the interactions between different vector fields will be the Lie bracket.

1.1 Control systems and controllability problems

By a control system we shall mean a system of the form

$$\Sigma : \quad \dot{x} = f(x, u),$$

where x, called *state* of Σ, takes values in an open subset X of $\mathbb{R}^n$ (or in a differentiable manifold X of dimension n) and u, called *control*, takes values in a set U. We call X the *state space* of the system and U the *control set.* When the control u is fixed the system equation $\dot{x} = f(x, u)$ defines a single dynamical system. Thus, the control system Σ can be viewed as a collection of dynamical systems parametrized by the control as parameter. We will see later that this interpretation is fruitful.

Example 1.1 *Boat on a lake.* Consider a motor boat on a lake. We can choose some coordinate system in which the lake is identified with a subset X of $\mathbb{R}^2$ and the state of the boat with a point $x = (x_1, x_2) \in X$. The simplest mathematical model of the motion of the boat is the following control system

$$\dot{x} = u$$

where the control $u = (u_1, u_2)$ is the velocity vector which belongs to the set $U = \{u \in \mathbb{R}^2 : ||u|| \leq m\}$, where $||u|| = \sqrt{u_1^2 + u_2^2}$ is the norm of u and m is the maximal possible velocity of the boat.

A different version of the problem is obtained if we consider a motor boat (or a rowing boat) on a river. Then the set of velocities of the boat $F(x)$ depends on the current of the river at this point. This means that in our model we have to change the equation $\dot{x} = u$ for

$$\dot{x} = f(x) + u,$$

where the control u is in the set $U = \{u \in \mathbb{R}^2 : ||u|| \leq m\}$ and $f(x)$ denotes the velocity vector of the current of the river at the point x. We could also keep the equation $\dot{x} = u$ and choose the control set $\tilde{U}(x) = f(x) + U$ depending on x (we will usually try to avoid the latter possibility as more complicated). Clearly, if the set of available velocities $F(x) = f(x) + U$ contains 0 in its interior then the boat can be steered from any intitial position to any final position if we use enough time.

Example 1.2 *Sailing boat.* A more interesting system is obtained when the boat is a sailing boat. Assuming that the wind is stable (of constant direction and force) we can model the motion of the boat on a lake by the equation

$$\dot{x} = v(\theta),$$

where θ is the angle of the axis of the boat with respect to the wind. The angle θ is treated as control and takes values in the set $U = (\alpha, 2\pi - \alpha)$, where α is the minimal angle with which the boat can sail against the wind. The velocity v, as a function of θ, depends on the characteristics of the boat related to the wind and it usually looks like in Figure 1 (a). An interesting problem for a sailor appears when the target is placed in the "dead cone" of the boat, when we look at it from the starting point. In that case sailing consists of a series of tacks chosen in such a way that the target is reached even if it is placed in the dead cone. In fact, sailing against the wind can be restricted to using only two values of the control $\theta = \pm\theta_{opt}$, where θ_{opt} maximizes the parallel to the wind component of $v(\theta)$ (directed against the wind). In this case the system reduces to two dynamical systems with two available velocities $v^+ = v(\theta_{opt})$ and $v^- = v(-\theta_{opt})$. By changing the tacks (Fig. 1 (b)) with the time spent for each (left and right) tack proportional, respectively, to constants λ_+ and λ_- (where $\lambda_+ + \lambda_- = 1$) the sailing boat changes its position as it was sailing with the average velocity $v = \lambda_+ v(\theta_{opt}) + \lambda_- v(-\theta_{opt})$.

The observation of the above example can be generalized to the following informal (but intuitively plausible)

Conclusion (principle of convexification). In analyzing controllability properties of systems Σ we can replace the set of available velocities $F(x) = \{f(x,u) : u \in U\}$ by its convex hull, the trajectories of the convexified system can be approximated (in C^0 topology) by the trajectories of the original system. In particular, if

$$0 \in int\, co\, F(x)$$

for all $x \in X$, then the system is completely controllable (any state can be reached from any other state).

Example 1.3 *Car parking I.* Suppose we would like to unpark our car blocked by two other cars parked on the side of the street (Fig. 2 (a)). The simplest but not always applicable strategy is to use a series of moves that gradually turn the car until it points to the free part of the street (Fig. 2 (b)).

We use the following mathematical model of our problem. We let x_1 and x_2 denote the Euclidean coordinates of the geometric center of the back axle of the car and ϕ will denote the angle between the axis of the car and the x_1-axis. We assume that the street is parallel to the x_1-axis. It is enough to consider movements with two extreme positions of the steering wheel. If we assume that the car moves with a constant angular velocity $\pm b$ then the velocity of the center of the rear axle moves along a circle (at each position of the steering wheel). The kinematic movements of the car in coordinates $x = (x_1, x_1, \phi)$ can be described by the following two vector fields on $\mathbb{R}^2 \times (-\pi, \pi) \subset \mathbb{R}^3$

$$f = (r\cos\phi, r\sin\phi, b)^T, \quad g = (r\cos\phi, r\sin\phi, -b)^T,$$

where r is a constant. Our strategy is to use a series of short moves (with equal length) where we interchange moving forwards with the leftmost position of the steering wheel (the vector field f) and moving backwards with the rigthmost position of the steering wheel (the vector field $-g$). Intuitively, the overal movement should be approximately described by the vector which is a linear combination of the vectors

f and $-g$. We have $(1/2)f - (1/2)g = (0, 0, b)$ which suggests that our series of movements can be approximated by a pure turn.

We shall later show that our approximation is justified by a suitable mathematical result (Proposition 1.8). The above strategy can not be used if the cars are approximately rectangular and the blocking cars are parked very close to our car (then their geometry will not allow for the turn of our car). In this case we have to use a more sophisticated strategy (Example 1.10) based on the notion of Lie bracket of vector fields. This strategy allows, approximately, to drive our car almost parallely in the direction perpendicular to the street (Fig. 2 (c)).

In fact, we shall be able to show later the following much stronger controllability property of the car. "Given $\epsilon > 0$ and any compact curve in the state space $X = \{(x_1, x_2, \phi) \in \mathbb{R}^2 \times S^1\}$, there exist admissible moves of the car which approximately follow the curve. More precisely, they bring it from the initial position of the curve to the final position of the curve and the car is never at a distance (in the state space) larger then ϵ from the curve."

1.2 Vector fields and flows

Let X denote an open subset of $\mathbb{R}^n$, possibly equal to $\mathbb{R}^n$ (the reader familiar with the theory of differentiable manifolds may assume from the beginning that X is a manifold). We denote by T_pX the space of tangent vectors to X at the point p. In the case where X is an open subset of $\mathbb{R}^n$ one can identify T_pX with $\mathbb{R}^n$ (this identification depends on the coordinate system).

A *vector field* on X is a mapping

$$X \ni p \longrightarrow f(p) \in T_pX$$

which assigns a tangent vector at p to any point p in X (Fig. 3). An analogous mapping defined on an open subset of X, only, will be called *partial vector field.* In a given system of coordinates f can be expressed as a column vector

$$f = (f_1, \ldots, f_n)^T,$$

where "T" stands for transposition. We say that f is of class C^k if its components are of class C^k.

We shall usually assume that the vector fields considered here are of class C^∞. The space of such vector fields forms a linear space (with natural, pointwise operations of summation and multiplication by numbers) denoted by $V(X)$.

For any vector field (or partial vector field) f we can write the differential equation

$$\dot{x} = f(x).$$

From theorems on existence of solutions of ordinary differential equations it follows that, if f is of class C^k and $k \geq 1$, then for any initial point p in the domain of f there is an open interval I containing zero and a differentiable curve $t \to x(t) = \gamma_t(p), t \in I$, which satisfies the above equation and $x(0) = \gamma_0(p) = p$. If f is of class C^∞, then from elementary properties of differential equations it follows that the map

$$(t, p) \longrightarrow \gamma_t(p)$$

is also of class C^∞ and is well defined on a maximal open subset of $\mathbb{R} \times X$. The resulted family γ_t of local maps of X (Fig. 4), called the *local flow* or simply the *flow* of the vector field f, has the following group type properties ("$\circ$" denotes composition of maps)

$$\gamma_{t_1} \circ \gamma_{t_2} = \gamma_{t_1+t_2}, \quad \gamma_{-t} = (\gamma_t)^{-1}, \quad \gamma_0 = id. \tag{1}$$

If the solution $\gamma_t(p)$ is well defined for all $t \in \mathbb{R}$ and $p \in X$, then the vector field f is called *complete* and its flow forms a one parameter group of (global) diffeomorphisms of X. Any one parameter family of maps which satisfies conditions (1) defines a unique vector field through the formula

$$f(p) = \left.\frac{\partial}{\partial t}\right|_{t=0} \gamma_t(p),$$

and the flow of this vector field coincides with γ_t.

We shall denote the local flow of a vecor field f by γ_t^f or by $\exp(tf)$. A reason for the latter notation will become clear later.

Example 1.4 The linear vector field $f(x) = Ax$ is complete and the corresponding flow is the one-parameter group of linear transformations

$$p \longrightarrow e^{At}p, \quad \text{i.e.} \quad \gamma_t = e^{At}.$$

1.3 Lie bracket and its properties

A nonlinear control system can be considered as a collection of dynamical systems (vector fields) parametrized by a parameter called control. It is natural to expect that basic properties of such a system depend on interconnections between the different dynamical systems corresponding to different controls. We represent our dynamical systems by vector fields as this allows us to perform algebraic operations on them such as taking linear combinations and a taking a product called Lie bracket. It is the Lie product which allows studying interconnections between different dynamical systems in a coordinate independent way.

The Lie bracket of two vector fields is another vector field which, roughly speaking, measures noncommutativenes of the flows of both vector fields. Noncommutativeness here means dependence of the result of applying the flows on the order of aplying these flows. These remark, as well as the definition of Lie bracket is made precise below.

There are three equivalent definitions of Lie bracket and each of them will be useful to us later. We start with the easiest (but coordinate dependent) definition in $\mathbb{R}^n$. Let $X \subset \mathbb{R}^n$, and let f and g be vector fields on X. The *Lie bracket* of f and g is another vector field on X defined as follows

$$[f,g](x) = \frac{\partial g}{\partial x}(x)f(p) - \frac{\partial f}{\partial x}(x)g(x), \tag{2}$$

where $\partial f/\partial x$ and $\partial g/\partial x$ denote the Jacobi matrices of f and g. We will call this *the Jacobian definition of Lie bracket.*

Example 1.5 For the vector fields $f = (1,0)^T$ and $g = (0, x_1)^T$ on $\mathbb{R}^n$ one easily finds that $[f,g] = (0,1)^T$. Note that the Lie bracket of f and g adds a new direction to the space spanned by f and g at the origin.

Let $f = b$ be a constant vector field and $g = Ax$ be a linear vector field. Then $[f,g] = [b, Ax] = Ab - 0 = Ab$. Similar trivial calculations show that the following holds.

Proposition 1.6 *The Lie bracket of two constant vector fields is zero. The Lie bracket of a constant vector field with a linear vector field is a constant vector field. Finally, the Lie bracket of two linear vector fields is a linear vector field.*

The basic geometric properties of Lie bracket are stated in the following propositions. The first one says that vanishing of Lie bracket $[f,g]$ is equivalent to the fact that starting from a point p and going along trajectory of f for time t and then along trajectory of g for time s gives always the same result as with the order of taking f and g reversed (Fig. 5).

Proposition 1.7 *The Lie bracket of vector fields f and g is equal identically to zero if and only if their flows commute, i.e.*

$$[f,g] \equiv 0 \quad \Longleftrightarrow \quad \gamma_t^f \circ \gamma_s^g(p) = \gamma_s^g \circ \gamma_t^f(p) \quad \forall s,t \in \mathbb{R}, \forall p \in X,$$

where the equality on the right should be satisfied for those s,t and p for which both sides are well defined.

Proof. To prove the implication "$\Longleftarrow$" it is enough to note that by computing the partial derivatives $(\partial/\partial t)(\partial/\partial s)$ at $t = s = 0$ of the left side of the equality $\gamma_t^f \circ \gamma_s^g(p) = \gamma_s^g \circ \gamma_t^f(p)$ and the same partial derivatives (but in reverse order) of the right side gives the equality $(\partial f/\partial x)g = (\partial g/\partial x)f$. The converse implication will be shown after Proposition 1.13 ■

Two vector fields having the property of Proposition 1.7 will be called *commuting.*

Proposition 1.8 *Let us fix a $p \in X$ and consider the curve (Fig. 6)*

$$\alpha(t) = \gamma_{-t}^g \circ \gamma{-}t^f \circ \gamma_t^g \circ \gamma_t^f(p).$$

Then we have that its first derivative at zero vanishes, $\alpha'(0) = 0$ and the second derivative is given by the Lie bracket:

$$\alpha''(0) = 2[f,g](p).$$

The above means that, after a reparametrization, the tangent vector at zero to the curve $t \to \alpha(t)$ is equal to $2[f,g](p)$ (see Fig. 6). This implies that the points attainable from p by means of the vector fields f and g lie not only in the "directions" $f(p)$ and $g(p)$, but also in the "direction" of the Lie bracket $[f,g](p)$. This fact will be of basic importance for studing controllability properties of nonlinear control systems.

The proof of the above proposition requires a lengthy calculation and is omited here (see e.g. Spivak [Sp], page 224). Note that the formula in Proposition 1.8 can be used for defining the Lie bracket $[f, g]$.

Proposition 1.9 *Suppose we are given two vector fields f and g on X and a point $p \in X$ and let λ_1, λ_2 be real constants. Define the following (local) diffeomorphisms of X*

$$\phi_t = \gamma^f_{\lambda_1 t} \circ \gamma^g_{\lambda_2 t}, \quad \psi_t = \gamma^g_{-t} \circ \gamma^f_{-t} \circ \gamma^g_t \circ \gamma^f_t .$$

Then the families of curves (Fig. 7)

$$\begin{aligned} \alpha_k(t) &= \phi_{t/k} \circ \cdots \circ \phi_{t/k}(p), \quad \text{k} - \text{times} \\ \beta_k(t) &= \psi_{t/k} \circ \cdots \circ \psi_{t/k}(p), \quad \text{k}^2 - \text{times} \end{aligned}$$

converge to the trajectories of the vector fields $\lambda_1 f + \lambda_2 g$ and $[f, g]$, respectively. More precisely, we have the convergence

$$\alpha_k(t) \longrightarrow \gamma^{\lambda_1 f + \lambda_2 g}_t(p), \quad \text{and} \quad \beta_k(t) \longrightarrow \gamma^{[f,g]}_{t^2}(p) \quad \text{as } k \longrightarrow \infty.$$

We will not prove this proposition here. However, the reader should find the first property about the convergence of α_k intuitively clear (compare the principle of convexification from Section 1.1). Namely, the movement which jumps sufficiently often between trajectories of two vector fields (and the time spent for these vector fields is proportional to some weights) follows, approximately, a trajectory of the linear combination of these vector fields (with the same weights). This property is used, for example, by sailors passing through narrow rivers or canals. A sailing boat can go against the wind only with certain minimal positive or negative angle (Example 1.2). But, even if the direction of the canal is in the "dead" cone and the boat can not go straight in this direction, the sailor tacks sufficiently often spending suitable amount of time for the left and the right tacks to reach the desired direction.

The property of convergence of β_k can be illustrated by the following example.

Example 1.10 *Car parking II.* Suppose the strategy of turning the car in Example 1.3 is inadmissible because the blocking cars are too close. There is a better strategy for unparking which works in any situation. Namely, we use repeatidly the following series of 4 moves: LF, RF, LB, RB, where "L" and "R" stand for the leftmost and rightmost positions of the steering wheel while "F" and "B" stand for forward and bacward motions. This means that our strategy is precisely the zig-zaging strategy described by $\beta_k(t)$ in Proposition 1.9. Therefore, the resulting movement follows approximately the Lie bracket of the vector fields

$$f = (r\cos\phi, r\sin\phi, b)^T, \quad g = (r\cos\phi, r\sin\phi, -b)^T.$$

We compute

$$\frac{\partial f}{\partial x} = \begin{pmatrix} 0 & 0 & -r\sin\phi \\ 0 & 0 & r\cos\phi \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \frac{\partial g}{\partial x} = \begin{pmatrix} 0 & 0 & -r\sin\phi \\ 0 & 0 & r\cos\phi \\ 0 & 0 & 0 \end{pmatrix}$$

and the Lie bracket of f and g equals to

$$[f, g] = br^2(-\sin\phi - \sin\phi, \cos\phi + \cos\phi, 0)^T.$$

In particular, at $\phi = 0$ we have that

$$[f, g] = (0, 2br^2, 0)^T.$$

The zig-zaging strategy produces movement approximating the trajectory of the Lie bracket $[f, g]$, that is the movement keeping the axis of the car approximately constant ($\phi = 0$) and changing its x_2-coordinate only (Fig. 3 (c)). This means that we should be able to unpark the car no matter how close the other cars are.

1.4 Coordinate changes and Lie bracket

To study what happens with vector fields and flows under coordinate changes let us consider a global diffeomorphism $\Phi : X \longrightarrow X$ (or a partial diffeomorphism i.e. a diffeomorphism between two open subsets of X). As tangent vectors are transformed through the Jacobian map of a diffeomorphism, our diffeomorphism defines the following transformation of a vector field f (see Fig. 8)

$$\mathrm{Ad}_\Phi(f)(p) = D\Phi(q)\, f(q), \quad q = \Phi^{-1}(p),$$

where $D\Phi$ denotes the tangent map of Φ (Jacobian mapping of Φ represented, in coordinates, by the Jacobi matrix $\partial\Phi/\partial x$). Another commonly used notation for the linear operator on $V(X)$ corresponding to the change of coordinates Φ is

$$\Phi_* f = \mathrm{Ad}_\Phi.$$

Note that the coordinate change $p = \Phi(q)$ transforms the differential equation $\dot{p} = f(p)$ into the equation $\dot{q} = \tilde{f}(q)$ where $\tilde{f} = \mathrm{Ad}_\Phi f$.

If Φ is a global diffeomorphism of X, then the operation Ad_Φ is a linear operator on the space of vector fields on X, i.e. $\mathrm{Ad}_\Phi(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 \mathrm{Ad}_\Phi(f_1) + \lambda_2 \mathrm{Ad}_\Phi(f_2)$. Additionally, if Ψ is another global diffeomorphism of X, then

$$\mathrm{Ad}_{\Phi\circ\Psi}(f) = \mathrm{Ad}_\Phi \mathrm{Ad}_\Psi(f),$$

where "$\circ$" denotes composition of maps.

For further reference we state the following fact.

Proposition 1.11 *Consider the vector field* $\mathrm{Ad}_\Phi(f)$. *The local flow of this vector field is given by*

$$\sigma_t = \Phi \circ \gamma_t \circ \Phi^{-1}.$$

Proof. It is easy to see that σ_t satisfies the group conditions (1) and we have

$$\left.\frac{\partial}{\partial t}\right|_{t=0} \Phi \circ \gamma_t \circ \Phi^{-1}(p) = D\Phi(\Phi^{-1}(p))\, f(\Phi^{-1}(p)) = (\mathrm{Ad}_\phi f)(p).$$

■

It is not immediately clear from the definition of Lie bracket in Section 1.3 that so defined $[f,g]$ is a vector field, that is, it is transformed with coordinate changes like a vector field. There are also other disadvantages of this definition which are not shared by the following *geometric definition of Lie bracket.* We define the Lie bracket of f and g as the derivative with respect to t, at $t = 0$, of the vector field g transformed by the flow of the field f. More precisely, we define (Fig. 9)

$$[f,g](p) = \frac{\partial}{\partial t} D\gamma^f_{-t}(\gamma^f_t(p))\, g(\gamma^f_t(p)) = \frac{\partial}{\partial t}(\mathrm{Ad}_{\gamma^f_{-t}} g)(p). \tag{3}$$

Let us check that this definition coincides with the Jacobian definition from Section 1.3. By taking the partial derivative $\partial/\partial t$ at $t = 0$ and taking into account that $\gamma_0^f = id$ and $\gamma_0^f(p) = p$ we find that the obove definition, where where t appears three times, gives

$$[f,g](p) = (D\frac{\partial}{\partial t}\Big|_{t=0}\gamma_{-t})(p)g(p) + \frac{\partial}{\partial t}\Big|_{t=0} D(id)(\gamma_t(p))g(p) + id\frac{\partial}{\partial t}\Big|_{t=0} g(\gamma_t(p)),$$

where we interchanged the order of taking the tangent map "D" (which is a matrix of partial derivatives with respect to the coordinates) and the partial derivative $\partial/\partial t$ in the first expression. The first term gives $-Df(p)g(p)$, the second is equal to zero, and the third equals to $Dg(p)f(p)$, which means that this definition coincides with the previous one.

It follows from the second definition of Lie bracket that $[f,g]$ transforms with coordinate changes like a vector field, that is via the Jacobi matrix of the coordinate change. Namely, we have the following basic property of equivariance of Lie bracket with coordinate changes.

Proposition 1.12 *If Φ is a (partial or global) diffeomorphism of X then*

$$[\mathrm{Ad}_\Phi f, \mathrm{Ad}_\Phi g] = \mathrm{Ad}_\Phi [f,g].$$

Proof. As we have established earlier, the flow of the vector field $\mathrm{Ad}_\Phi f$ is equal to $\sigma_t = \Phi \circ \gamma_t^f \circ \Phi^{-1}$. Thus, applying the geometric definition of Lie bracket gives

$$[\mathrm{Ad}_\Phi f, \mathrm{Ad}_\Phi g](p) = \tfrac{\partial}{\partial t}\Big|_{t=0}(\mathrm{Ad}_{\Phi\circ\gamma^f_{-t}\circ\Phi^{-1}}\mathrm{Ad}_\Phi g)(p) =$$

$$\frac{\partial}{\partial t}\Big|_{t=0}(\mathrm{Ad}_\Phi \mathrm{Ad}_{\gamma^f_{-t}}\mathrm{Ad}_{\Phi^{-1}}\mathrm{Ad}_\Phi g)(p) = \tfrac{\partial}{\partial t}\Big|_{t=0}(\mathrm{Ad}_\Phi \mathrm{Ad}_{\gamma^f_{-t}} g)(p) = \mathrm{Ad}_\Phi[f,g].$$

From the geometric definition of Lie bracket we deduce the following relation.

Proposition 1.13 *We have*

$$\frac{\partial}{\partial t}\mathrm{Ad}_{\gamma_t^f} g = -[f, \mathrm{Ad}_{\gamma_t^f} g] = -\mathrm{Ad}_{\gamma_t^f}([f,g]).$$

Proof. To show the first equality it is enough to note that

$$\frac{\partial}{\partial t}\mathrm{Ad}_{\gamma_t^f} g = \frac{\partial}{\partial h}\Big|_{h=0} \mathrm{Ad}_{\gamma_h^f}\mathrm{Ad}_{\gamma_t^f} g$$

and apply the geometric definition of Lie bracket to the vector fields $-f$ and $\mathrm{Ad}_{\gamma_t^f} g$. The second equality follows analogously from $\frac{\partial}{\partial t}\big|_{t=0} \mathrm{Ad}_{\gamma_t^f} g = \frac{\partial}{\partial h}\big|_{h=0} \mathrm{Ad}_{\gamma_t^f} \mathrm{Ad}_{\gamma_h^f} g$. ■

Proof of Proposition 1.7. To show the converse implication note that from $[f,g] \equiv 0$ and the equalities in Proposition 1.13 it follows that $\mathrm{Ad}_{\gamma_t^f} g$ is independent of t, i.e. $\mathrm{Ad}_{\gamma_t^f} g = \mathrm{Ad}_{\gamma_0^f} g = g$. Therefore, the flow of g is equal to the flow of the vector field $\mathrm{Ad}_{\gamma_t^f} g$, i.e. $\gamma_t^f \circ \gamma_s^g \circ \gamma_{-t}^f = \gamma_s^g$, by Proposition 1.11. This implies that $\gamma_t^f \circ \gamma_s^g = \gamma_s^g \circ \gamma_t^f$ and the proposition is proved. ■

Below and in the following sections we shall use the following notation. We denote $\mathrm{ad}_f g = [f,g]$. Thus, ad_f is a linear operator in the space of vector fields $V(X)$. We also consider its iterations

$$\mathrm{ad}_f^0 g = g \quad \text{and} \quad \mathrm{ad}_f^i g = \mathrm{ad}_f \cdots \mathrm{ad}_f g \quad \mathrm{i-times}.$$

The following dependence between the operations Ad and ad follows from the formula in Proposition (1.13)

$$\frac{\partial}{\partial t}(\mathrm{Ad}_{\gamma_t^f} g)(p) = -(\mathrm{ad}_f(\mathrm{Ad}_{\gamma_t^f} g))(p). \tag{4}$$

In the analytic case we also have an expansion formula which follows from this relation.

Proposition 1.14 *If the vector fields f and g are real analytic, then we have the following expansion formula for the vector field g transformed by the flow of the vector field f:*

$$(\mathrm{Ad}_{\gamma_t^f} g)(p) = \sum_{i=0}^{\infty} \frac{(-t)^i}{i!} (\mathrm{ad}_f^i g)(p),$$

where the series converges absolutely for t in a neighborhood of zero (more precisely, each of n components of this series converges absolutely).

Proof. Applying iteratively the formula (4) and taking into account that $\gamma_0^f = id$ we find that

$$(\frac{\partial}{\partial t})^i (\mathrm{Ad}_{\gamma_t^f} g)(p)|_{t=0} = (-1)^i \mathrm{ad}_f^i g(p).$$

Therefore, our equality is simply the Maclaurin series of the left hand side. ■

1.5 Vector fields as differential operators

A smooth vector field f on X defines a linear operator L_f on the space of smooth functions $C^\infty(X)$ in the following way

$$(L_f\phi)(p) = \frac{\partial}{\partial t}\bigg|_{t=0} \phi(\gamma_t(p)) = \sum_{i=1}^{n} f_i(p)\frac{\partial}{\partial x_i}\phi(p). \tag{5}$$

This operator is called *directional derivative along* f or *Lie derivative along* f and it is a differential operator of order one.

Conversely, any differential operator of order one with no zero order term can be written as

$$L = \sum_{i=1}^{n} a_i(x)\frac{\partial}{\partial x_i}$$

and it defines a unique vector field given in coordinates as $f = (a_1, \ldots, a_n)^T$. (We can easily check that the coordinate vector $(a_1, \ldots, a_n)$ of the operator L transforms with a coordinate change Φ by the Jacobi matrix $\partial\Phi/\partial x$. Thus so defined f is a vector field on X, cf. Remark ??.) This means that there is a unique correspondence

$$f \to L_f$$

between vector fields and differential operators of order one (with no zero order term)

Because of the above correspondence mathematicians often identify vector fields f with the corresponding differential operators L_f and write

$$L_f = f = \sum_{i=1}^{n} f_i\frac{\partial}{\partial x_i}.$$

We will rather try to distinguish between these two objects.

We shall close this subsection with a third definition of Lie bracket and some useful corollaries to it. Let f, g be vector fields and L_f, L_g the corresponding differential operators. Consider the commutator of these operators defined by

$$[L_f, L_g] := L_f L_g - L_g L_f.$$

Proposition 1.15 *The commutator* $[L_f, L_g]$ *is a differential operator of order one which corresponds to the Lie bracket* $[f, g]$, *i.e.,*

$$[L_f, L_g] = L_{[f,g]}.$$

Proof. Given any smooth function ϕ, we compute the composed differential operator on ϕ

$$L_f L_g \phi = \sum_i f_i \frac{\partial}{\partial x_i}(\sum_j g_j \frac{\partial}{\partial x_j}\phi) = \sum_{ij} f_i g_j \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\phi + \sum_{ij} f_i \frac{\partial g_j}{\partial x_i}\frac{\partial \phi}{\partial x_j}.$$

The analogous expression for $L_g L_f \phi$ has the same first summand, due to commutativeness of partial derivatives with respect to x_i and x_j, thus we have

$$[L_f, L_g]\phi = L_f L_g \phi - L_g L_f \phi = \sum_{ij} f_i \frac{\partial g_j}{\partial x_i}\frac{\partial \phi}{\partial x_j} - \sum_{ij} g_i \frac{\partial f_j}{\partial x_i}\frac{\partial \phi}{\partial x_j}.$$

We see that $[L_f, L_g]$ is a differential operator of order one. Using the Jacobian definition of Lie bracket from Section 1.3 we see that $L_{[f,g]}\phi$ gives the same expression

$$L_{[f,g]}\phi = \sum_j (\sum_i f_i \frac{\partial g_j}{\partial x_i} - g_i \frac{\partial f_j}{\partial x_i})\frac{\partial \phi}{\partial x_j},$$

which means that $[L_f, L_g] = L_{[f,g]}$. ■

If we identify vector fields f with the corresponding differential operators L_f, i.e. write $f = L_f = \sum_i f_i \partial / x_i$, then Proposition 1.15 suggests that we can equivalently define the *Lie bracket* as the *commutator*

$$[f,g] = fg - gf = \sum_j (\sum_i \frac{\partial g_j}{\partial x_i} f_i - \frac{\partial f_j}{\partial x_i} g_i)\frac{\partial}{\partial x_j},$$

where $g = \sum_j g_j \partial / \partial x_j$. We shall call this the *algebraic definition of Lie bracket.* Clearly, this definition coincides in a given coordinate system with the Jacobian definition, if we use the identifications $f = L_f$, $g = L_g$.

Commutator of linear operators is antisymmetric and satisfies the Jacobi identity $[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0$ (verify this using the definition $[A,B] = AB - BA$ of commutator). Therefore, we have the following properties of Lie bracket

$$[f,g] = -[g,f], \qquad \text{(antisymmetry)}$$

$$[f,[g,h]] + [g,[h,f]] + [h,[f,g]] = 0. \quad \text{(Jacobi identity)}$$

for any vector fields f, g, in $V(X)$. The former property also follows easily from the first definition of Lie bracket. Because of the above properties the linear space

$V(X)$ of smooth vector fields on X, with the Lie bracket as product, is called the *Lie algebra of vector fields on* X.

Appendix 1: Lie Algebras

A *Lie algebra* is a linear space L with a product $[\cdot, c] : L \times L \longrightarrow L$ which satisfies the following properties

$$[f,g] = -[g,f] \qquad \text{(antisymmetry)},$$

$$[f,[g,h]] + [g,[h,f]] + [h,[f,g]] = 0 \qquad \text{(Jacobi condition)}.$$

The Jacobi condition can be equivalently written as the following Leibnitz type condition

$$[f,[g,h]] = [[f,g],h] + [g,[f,h]]$$

or equivalently

$$\mathrm{ad}_f[g,h] = [\mathrm{ad}_f g, h] + [g, \mathrm{ad}_f h],$$

where ad_f denotes the linear operator in L defined by the formula $\mathrm{ad}_f g = [f,g]$. There is still another form of writing the Jacobi condition which will be useful to us

$$\mathrm{ad}_{[g,h]} f = \mathrm{ad}_g \mathrm{ad}_h f - \mathrm{ad}_h \mathrm{ad}_g f = [\mathrm{ad}_g, \mathrm{ad}_h] f, ()$$

where the square bracket on the right denotes the commutator of linear operators in L: $[\mathrm{ad}_g, \mathrm{ad}_h] = \mathrm{ad}_g \mathrm{ad}_h - \mathrm{ad}_h \mathrm{ad}_g$.

A linear subspace K of L which is closed under the product $[;] : L \times L \longrightarrow L$ is called a ıLie subalgebra of L. A *Lie subalgebra generated by a subset* or simply *Lie algebra generated by a subset* $S \subset L$ is the smallest Lie subalgebra of L which contains S. A *Lie ideal of* L is a linear subspace $I \subset L$ such that $[f,g] \in I$, whenever $f \in L$ and $g \in I$.

Example 1.16 The space $gl(n)$ of all square $n \times n$ matrices with the commutator

$$[A,B] = AB - BA$$

forms a Lie algebra. There are various Lie subalgebras of this algebra which are interesting and important for mathematics and physics. For example, skew symmetric matrices form a Lie subalgebra of this Lie algebra.

Example 1.17 The space $V(X)$ of smooth vector fields on a smooth manifold X (or simply on $X = \mathbb{R}^n$) forms a Lie algebra with Lie bracket as product. When the vector fields are treated as differential operators of order one, then the Lie bracket becomes the commutator of operators, as in the above case of square matrices (treated as linear operators). There is no surprise about this, namely, there is a Lie subalgebra of the algebra of vector fields which is formed by the space of linear vector fields: $f = Ax$, or in the operator form

$$f = \sum_{i,j} a_{ij} x_j \frac{\partial}{\partial x_i}.$$

Here, the Lie bracket corresponds to taking commutators of the corresponding matrices $[Ax, Bx] = (BA - AB)x = [B, A]x$.

Example 1.18 In the Lie algebra of linear vector fields as defined above there is an ideal which consists of all constant vector fields.

An iterative application of the Jacobi identity in the form (1.5) and of anti-symmetry of Lie bracket leads to the following general property. Let $f_1,\ldots,f_k$ be elements of a Lie algebra L. We shall call an iterated Lie bracket of these elements any element of L obtained from these elements by applying iteratively the operation of Lie bracket in any possible order, e.g. $[[f_1, f_4], [f_3, f_1]]$. Left iterated Lie brackets will be brackets of the form $[f_{i_1}, \ldots, [f_{i_{k-1}}, f_{i_k}] \ldots]$.

Proposition 1.19 *Any iterated Lie bracket of $f_1, \ldots, f_k$ is a linear combination of left iterated Lie brackets of $f_1, \ldots, f_k$.*

For example

$$[[f_1, f_4], [f_3, f_1]] = [\mathrm{ad}_{f_1}, \mathrm{ad}_{f_4}]\ [f_3, f_1] = [f_1, [f_4, [f_3, f_1]]] - [f_4, [f_1, [f_3, f_1]]].$$

Exercise. Prove the above proposition (you may use induction with respect to the order of Lie bracket).

Appendix 2: Equivalence of families of vector fields

To close this chapter we shall show that the Lie brackets taken at a point of an analytic family of vector fields form a complete set of its invariants. As a control system can be represented by a family of vector fields, this will have direct applications to control systems. In another version of this result we will define a family of functions which forms a set of complete invariants for state equivalence.

Consider two general families of analytic vector fields on X and $\tilde{X}$, respectively, parametrized by the same parameter $u \in U$

$$F = \{f_u\}_{u \in U}, \qquad \tilde{F} = \{\tilde{f}_u\}_{u \in U}.$$

We shall call these families *locally equivalent* at the points p and $\tilde{p}$, respectively, if there is a local analytic diffeomorphism $\Phi : X \longrightarrow \tilde{X}$, $\Phi(p) = \tilde{p}$ which transforms the vector fields f_u into $\tilde{f}_u$ locally, i.e.

$$\mathrm{Ad}_\Phi f_u = \tilde{f}_u, \quad \text{for} \;\; u \in U$$

locally around $\tilde{p}$.

Denote by $\mathcal{L}$ and $\tilde{\mathcal{L}}$ the Lie algebras of vector fields generated by the families F and $\tilde{F}$. Recall that a family of vector fields is called transitive at a point if its Lie algebra is of full rank at this point, i.e. the vector fields in this Lie algebra span the whole tangent space at this point.

We shall use the following notation for left iterated Lie brackets

$$f_{[u_1 u_2 \cdots u_k]} = [f_{u_1}, [f_{u_2}, \ldots, [f_{u_{k-1}}, f_{u_k}] \cdots]]$$

and analogous for the tilded family. In particular, $f_{[u_1]} = f_{u_1}$.

Theorem 1.20 *If the families F and $\tilde{F}$ are transitive at the points p and $\tilde{p}$, respectively, then they are locally equivalent at these points if and only if there exixts a linear map between the tangent spaces $L : T_p \longrightarrow T_{\tilde{p}}$ such that*

$$L \; f_{[u_1 u_2 \cdots u_k]}(p) = \tilde{f}_{[u_1 u_2 \cdots u_k]}(\tilde{p}) \tag{6}$$

for any $k \geq 1$ and any $u_1, \ldots, u_k \in U$.

Proof. Necessity. If $\tilde{f}_u = \mathrm{Ad}_\Phi f_u$, then $\tilde{f}_u(\tilde{p}) = L \; f_u(p)$ where $L = d\Phi(p)$. To prove condition (6) in general it is enough to use iteratively the property of Lie bracket

$$[\mathrm{Ad}_\Phi f, \mathrm{Ad}_\Phi g] = \mathrm{Ad}_\Phi [f, g]$$

from which we get $\tilde{f}_{[u_1 \cdots u_k]} = \mathrm{Ad}_\Phi f_{[u_1 \ldots u_k]}$ and so the condition (6). ∎

The proof of sufficiency is more involved an will be presented in the next section together with other versions of the above result.

2 Orbits, distributions, and foliations

2.1 Distributions and local Frobenius theorem

In this chapter we introduce notions and results which play a basic role in analysis and understanding the structure of nonlinear control systems. Their are directly related to controllability properties of such systems. We denote by X an open subset of $\mathbb{R}^n$ or a diferentiable manifold of dimension n.

Definition 2.1 A *distribution on* X is, by definition, a map Δ which assigns to each point in X a subspace of the tangent space at this point, i.e.

$$X \ni p \longrightarrow \Delta(p) \subset T_pX.$$

The distribution Δ is called of class C^∞ if, locally around each point in X, there is a family of vector fields $\{f_\alpha\}$ (called local generators of Δ) which spans Δ, i.e. $\Delta(p) = \text{span}_\alpha f_\alpha(p)$. Δ is called *locally finitely generated* if the above family of vector fields is finite. Finally, the distribution Δ is called *of dimension* k if $\dim\Delta(p) = k$ for all points p in X, and of *constant dimension* if it is of dimension k, for some k.

We will tacitly assume that our distributions are of class C^∞.

Definition 2.2 We say that a vector field f belongs to a distribution Δ and write $f \in \Delta$ if $f(p) \in \Delta(p)$ for all p in X. A distribution Δ is called *involutive* if for any vector fields $f, g \in \Delta$ the Lie bracket is also in Δ; $[f, g] \in \Delta$. If the distribution has, locally, a finite number of generators $f_1, ..., f_m$ then involutivity of Δ means that

$$[f_i, f_j](p) = \sum_{k=1}^{m} \phi_{ij}^k(p) f_k(p), \qquad i, j = 1, ..., m,$$

where ϕ_{ij}^k are C^∞ functions.

Involutivity plays a fundamental role in the following Frobenius theorem.

Theorem 2.3 *If Δ is an involutive distribution of class C^∞ and of dimension k on X then, locally around any point in X, there exists a smooth change of coordinates which transforms the distribution Δ to the following constant distribution*

$$\text{span}\{e_1, \ldots, e_k\},$$

where $e_1, \dots, e_k$ *are the constant versors* $e_i = (0, \dots, 1, \dots, 0)^T$, *with* 1 *at i-th place.*

Proof. The proof will consist of two steps.

Step 1. We shall first show that the distribution Δ is locally generated by k pairwise commuting vector fields. Let us fix a point p in X and let $f_1, ..., f_k$ be any vector fields which generate the distribution Δ in a neighborhood of p. Treating f_i as column vectors, we form the $n \times k$ matrix $F = (f_1, ..., f_k)$. Note that multiplying F from the right by an invertible $k \times k$ matrix of smooth functions does not change the distribution spanned by the columns of F (it changes its generators, only). By a possible permutation of variables we achieve that the upper $k \times k$ submatrix of the matrix F is nonsingular. Multiplying F from the right by a suitable invertible matrix we obtain that this submatrix is equal to the identity, i.e. the new matrix F takes the form

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ * & * & \dots & * \\ \vdots & \vdots & & \vdots \\ * & * & \dots & * \end{pmatrix},$$

where "*" denote unknown coefficients. The new vector fields formed by the columns of this matrix commute. In fact, since their first k coefficients are constant, the first k coefficients of any Lie bracket $[f_i, f_j]$ vanish. On the other hand, from involutivity it follows that this Lie bracket is a linear combination of the columns of F. Both these facts can only hold when the coefficients of this linear combination are equal to zero. This shows that the new vector fields commute.

Step 2. Assume that the vector fields $f_1, \dots, f_k$ generate the distribution Δ, locally around p, and they commute. We can choose other $n - k$ vector fields $f_{k+1}, \dots, f_n$ so that $f_1, \dots, f_n$ are linearly independent at p. Define a map Φ by

$$(t_1, \dots, t_n) \longrightarrow \exp(t_1 f_1) \exp(t_2 f_2) \cdots \exp(t_n f_n) p.$$

As the flows of the vector fields $f_1, \dots, f_k$ commute, we see that the order of taking these flows in the above definition can be changed. Therefore, an integral curve of a vector field $e_i = (0, \ldots, 1, \ldots, 0)^T$, $1 \le i \le k$ is transformed to an integral curve of the vector field f_i (as we may place the flow of f_i to the most left place). It follows that the map Φ sends the vector fields $e_1, \ldots, e_k$ to the vector fields $f_1, \ldots, f_k$ and conversely does the inverse map Φ^{-1}. This inverse map is the desired map which transforms the distribution Δ spanned by $f_1, \ldots, f_k$ to the constant distribution spanned by $e_1, \ldots, e_k$. ∎

In order to state a global version of this theorem as well as other theorems related to transitivity of families of vector fields and integrability of distributions we need more definitions.

2.2 Submanifolds and foliations

Definition 2.4 A subset $S \subset X$ is called a *regular submanifold* of X of dimension k if for any $x \in S$ there exists a neighborhood U of x and a diffeomorphism $\Phi : U \longrightarrow V \subset \mathbb{R}^n$ onto an open subset V such that

$$\Phi(U \cap S) = \{x = (x_1, \ldots, x_n) \in V \mid x_{k+1} = 0, \ldots, x_n = 0\}$$

(see Fig. 10). The regularity class of this submanifold is by definition the regularity class of the diffeomorphism Φ (we shall assume that this regularity is C^∞ or C^ω).

In other words, a regular submanifolds of dimension k is a subset which locally looks like a piece of of subspace of dimension k, up to a change of coordinates. A slightly weaker notion of a submanifold is introduced in the following definition.

Definition 2.5 We call a subset $S \subset X$ an *immersed submanifold* of X of dimension k if

$$S = \cup_{i=1}^{\infty} S_i, \qquad \text{where} \qquad S_1 \subset S_2 \subset S_3 \subset \cdots \subset S$$

and S_i are regular submanifolds of X of dimension k.

In the case when S itself is a regular submanifold we can take $S_i = S$ and so S is also an immersed submanifold.

Example 2.6 In Figure 11 (a) and (b) are regular submanifolds of $\mathbb{R}^2$ while (b) and (d) are only immersed submanifolds.

We shall later need two geometric properties of Lie bracket.

Property 1. If two vector fields f, g are tangent to an (immersed) submanifold S then also their Lie bracket $[f, g]$ is tangent to this submanifold.

This follows from the geometric definition of Lie bracket. In fact, if f is tangent to S, then its flow transforms points of S into points of S when the time is sufficiently small. Therefore, the tangent map to the flow $D\gamma_t^f$ transforms the tangent subspaces of S into tangent subspaces of S, in particular, it transforms the tangent vectors $g(p)$ into vectors tangent to S. Moreover, the vectors $v(t) = (\mathrm{Ad}^f_{\gamma_{-t}} g)(p)$ are all in the tangent space T_pS. Taking derivative with respect to t of this expression, which appears in the geometric definition of $[f, g]$, gives a tangent vector to S.

Definition 2.7 A *foliation* $\{S_\alpha\}_{\alpha\in A}$ of X of dimension k is a partition

$$X = \cup_{\alpha\in A} S_\alpha$$

of X into disjoint connected (immersed) submanifolds S_α, called *leaves*, which has the following property. For any $x \in X$ there exists a neighborhood U of x and a diffeomorphism $\Phi : U \longrightarrow V \subset \mathbb{R}^n$ onto an open subset V such that

$$\Phi((U \cap S_\alpha)_{cc}) = \{x = (x_1, \ldots, x_n) \in V \mid x_{k+1} = c_\alpha^{k+1}, \ldots, x_n = c_\alpha^n\},$$

where P_{cc} denotes a connected component of the set P and the above property should hold for any such connected componentm, with the constants c_α^i depending on the leaf and the choice of the connected component (Fig. 12). Similarly as for submanifolds, the regularity of the foliation is defined by the regularity of the diffeomorphism Φ.

Examples of foliations on subsets of $\mathbb{R}^2$ are presented in Figure 13. A general example of a foliation of dimension $k = n - r$ is given by the following equations for leaves

$$S_\alpha = \{x \in X \mid h_1(x) = c_\alpha^1, \ldots, h_r(x) = c_\alpha^r\},$$

where c_α^i are arbitrary constants and $h = (h_1, \ldots, h_r)$ is a smooth map of constant rank r (i.e. its Jacobi map is of rank r).

Property 2. Assume that a vector field g is tangent to a foliation $\{S_\alpha\}_{\alpha\in A}$, that is, it is tangent to its leaves. Then, if the flow of another vector field f locally preserves this foliation, the Lie bracket $[f, g]$ is tangent to this foliation.

Here by saying that the flow of f locally preserves the foliation $\{S_\alpha\}_{\alpha\in A}$ we mean that for any point $p \in S_\alpha$ there is a neighborhood U of p such that the image of a piece of a leaf $\gamma_t^f(S_\alpha \cap U)$ is contained in a leaf of the foliation (which depends on t), for any t sufficiently small.

To prove this property let us choose coordinates as in the definition of the foliation and assume that γ_t^f locally preserves $\{S_\alpha\}_{\alpha\in A}$. It follows that the tangent map to γ_t^f maps tangent spaces to leaves into tangent spaces to leaves. Therefore the vector $D\gamma_t^f(p)g(p)$ is tangent to leaves and, in particular, its last $n-k$ components are zero (here we use our special coordinates). Differentiating with respect to t at $t=0$ gives a vector with the last $n-k$ components equal to zero (and so tangent to a leaf), which by the geometric definition of Lie bracket is equal to $[f,g](p)$.

2.3 Orbits of families of vector fields

Consider a family of (global or partial) vector fields $\mathcal{F} = \{f_u\}_{u\in U}$ on X.

Definition 2.8 We define the *orbit of a point* $p \in X$ of this family as the set of points of X reachable from p piecewice by trajectories of vector fields in the family, i.e.

$$\mathrm{Orb}(p) = \{\gamma_{t_k}^{u_k} \circ \cdots \circ \gamma_{t_1}^{u_1} \mid k \geq 1,\ u_1,\ldots,u_k \in U,\ t_1,\ldots,t_k \in \mathbb{R}\},$$

where γ_t^u denotes the flow of the vector field f_u. Of course, if some of our vector fields are not complete then we consider only such $t_1,\ldots,t_k$ for which the above expression has sense.

The relation: "q belongs to the orbit of p" is an equivalence relation on the space X. In fact, a point q belongs to the orbit $\mathrm{Orb}(p)$ if and only if it is reachable from p piecewise by trajectories of the vector fields in the family $\mathcal{F}$. It is evident that q is reachable from p if and only if p is reachable from q (symmetry). Also, if q is reachable from p and r is reachable from q, then r is reachable from p (transitivity).

It follows then that the space X is a disjoint union of orbits (equivalence classes).

Definition 2.9 Let Γ be the smallest distribution on X which contains the vector fields in the family $\mathcal{F}$ (i.e. $f_u(p) \in \Gamma(p)$ for all $u \in U$) and is invariant under any flow γ_t^u, $u \in U$, that is

$$D\gamma_t^u(p)\Gamma(p) \subset \Gamma(\gamma_t^u(p)).$$

for all $p \in X$, $u \in U$ and t for which the above expression is well defined.

Equivalently, we can write the invariance property (using partial vector fields) in the form:

$$g \in \Gamma \quad \Longrightarrow \quad \mathrm{Ad}_{\gamma_t^u} g \in \Gamma, \qquad \text{for any } u \in U \text{ and } t \in \mathbb{R}.$$

The following theorem was proved independently by H.J. Sussmann and P. Stefan. We state it here without proof.

Theorem 2.10 *(Orbit Theorem) Each orbit $S = \mathrm{Orb}(p)$ of a family of vector fields $\mathcal{F} = \{f_u\}_{u\in U}$ is an immersed submanifold (of class C^k if the vector fields f_u are of class C^k). Moreover, the tangent space to this submanifold is given by the distribution Γ,*

$$T_pS = \Gamma(p), \quad \textit{for all} \;\; p \in X.$$

Corollary 2.11 *If the vector fields f_u are analytic, then the tangent space to the orbit is given by*

$$T_pS = L(p) := \{g(p) \mid g \in \mathrm{Lie}\{f_u\}_{u\in U}\ \},$$

where $\mathrm{Lie}\{f_u\}_{u\in U}$ *denotes smallest family of (partial) vector fields which contains the family $\mathcal{F}$ and is closed under taking linear combinations and Lie bracket (this is the Lie algebra of vector fields generated by the family $\mathcal{F} = \{f_u\}_{u\in U}$ in the case when f_u are global vector fields). In the smooth case the following inclusion holds*

$$L(p) \subset \Gamma(p).$$

Proof. We shall first prove the inclusion. Using the second form of the invariance property of Γ and the geometric definition of Lie bracket we obtain the following implication

$$g \in \Gamma \Longrightarrow [f_u, g] \in \Gamma.$$

Applying this implicaction iteratively, we deduce that the left iterated Lie brackets

$$[f_{u_k}, \cdots, [f_{u_2}, f_{u_1}] \cdots]$$

are in Γ. As all iterated Lie brackets are linear combinations of left iterated Lie brackets (1.19), it follows that $L(p) \subset \Gamma(p)$ for $p \in X$.

To prove the equality in the analytic case it is enough to use the formula

$$D\gamma_t^u(q) f_v(q) = \sum_{i \geq 0} \frac{(-t)^i}{i!} \mathrm{ad}_{f_u}^i f_v(p), \quad p = \gamma_t^u(q),$$

(1.14) which shows that transformations of vectors under the tangent maps to flows of f_u can be expressed by taking (infinite) linear combinations of Lie brackets. This implies that $\Gamma(p) \subset L(p)$. ■

Example 2.12 The following system in the plane

$$\begin{aligned} \dot{x}_1 &= u_1 x_1, \qquad |u_1| \leq 1, \\ \dot{x}_2 &= u_2 x_2, \qquad |u_2| \leq 1, \end{aligned}$$

represented by the family of vector fields

$$f_u = (u_1 x_1, u_2 x_2)^T$$

has four 2-dimensional orbits (the open octants), four 1-dimensional orbits (open half-axes) and one zero dimensional orbit which is the origin.

Example 2.13 The family of three vector fields which represent rotations around the three axes

$$f_1 = (0, x_3, -x_2)^T, \quad f_2 = (x_3, 0, -x_1)^T, \quad f_3 = (x_2, -x_1, 0)^T$$

has a continuum of 2-dimensional orbits which are spheres with the center at the origin and one zero dimensional orbit which is the origin itself. Note that the orbits form a 2-dimensional foliation on the set $X = \mathbb{R}^3 \setminus \{0\}$.

The following example shows that in the nonanalytic case the equality $\Gamma(p) = L(p)$ may not hold.

Example 2.14 Consider the family of the following two C^∞ vector fields in the plane

$$f_1 = (1, 0)^T, \quad f_2 = (0, \phi(x_1))^T,$$

where $\phi(y)$ is a smooth function on $\mathbb{R}$ positive for $y < 0$ (for example $\phi(y) = \exp(1/y)$ and equal to zero for $y \geq 0$. Then the orbit of any point is equal to the whole $\mathbb{R}^2$ and from the orbit theorem it follows that $\dim\Gamma(p) = 2$ for any p. On the other hand, we have that $L(p)$ is spanned by the first vector field only, when $x_1 \geq 0$, so $\dim L(p) = 1$.

Corollary 2.15 *(Chow and Rashevski) If* $\dim L(p) = n$ *for any* $p \in X$*, then any point of* X *is reachable from any other point piecewise by trajectories of* $\mathcal{F} = \{f_u\}_{u\in U}$ *(allowing positive and negative times), i.e.* $\mathrm{Orb}(p) = X$ *for any* p.

Proof. It follows from our assumption and the above corollary that $\Gamma(p)$ is equal to the whole tangent space T_pX for any p. From the orbit theorem it follows then that the orbit of any point is of full dimension, so it is an open subset of X. We conclude that X is a union of disjoint open subsets and, as X is connected, only one of them can be nonempty. Therefore, X consists of a single orbit and any point is reachable from any other point piecewice by trajectories of our family of vector fields. ■

2.4 Integrability of distributions and foliations

The above results, especially the orbit theorem, allow us to give criteria for integrability of distributions and prove some classicical theorems.

Definition 2.16 We say that a distribution of constant dimension $p \longrightarrow \Delta(p)$ on X is *integrable* if there exists a foliation $\{S_\alpha\}_{\alpha\in A}$ on X such that for any $p \in X$

$$T_pS = \Delta(p),$$

where S is the leaf passing through p.

Finding the foliation which satisfies the condition of the above definition is usually called integrating this distribution, while the foliation and its leaves are called integral foliation and integral (sub)manifolds of the distribution.

Theorem 2.17 *(Global Frobenius theorem) A smooth distribution of constant dimension* Δ *is integrable if and only if it is involutive. The integral foliation of* Δ *is the partition of* X *into orbits of the family of (partial) vector fields* $\{g \mid g \in \Delta\}$.

Proof. Assume that our distribution is integrable and choose two vector fields $f, g \in \Delta$ and any point $p \in X$. Then f and g are tangent to the leaf S passing through p, therefore their Lie bracket $[f, g]$ is also tangent to this leaf by Property 1. As this happens for any p, it follows that $[f, g](p) \in T_pS = \Delta(p)$ for all p and so $[f, g] \in \Delta$.

Assume now that our distribution is involutive. Consider the family of partial vector fields $\mathcal{F} = \{f \mid f \in \Delta\}$. We shall prove that the partition of X into orbits of this family gives the desired foliation.

Let $f_1, \ldots, f_k, \in \Delta$ span this distribution in a neighborhood of p. We shall show that Δ is invariant under the flows of the vector fields $f \in \Delta$, that is the distribution Γ in the orbit theorem coincides with Δ. We have to prove that

$$D\gamma_t^f(p)\Delta(p) = \Delta(q), \quad q = \gamma_t^f(p),$$

for $f \in \Delta$. The left hand side subspace is spanned by the vector fields

$$g_t^i = \mathrm{Ad}_{\gamma_t^f} f_i, \quad i = 1, \ldots, k.$$

From the involutiveness assumption we have that $[f, f_i] = \sum_j \phi_{ij} f_j$. Denote the functions $a_t^{ij} = -\phi_{ij} \circ \gamma_{-t}^f$. From Proposition 1.13 it follows that the spanning vector fields satisfy pointwise the following system of linear differential equations

$$\frac{d}{dt} g_t^i = -\mathrm{Ad}_{\gamma_t^f}[f, f_i] = \sum_j g_t^j a_t^{ij}.$$

As the solution of a linear differential equation depends linearly on its initial conditions, it follows that

$$g_t^i = \sum_j \psi_t^{ij} g_0^j = \sum_j \psi_t^{ij} f_j,$$

where ψ_t^{ij} are functions. Therefore, the subspace $D\gamma_t^f(p)\Delta(p)$ is spanned by the vectors $f_1(p), \ldots, f_k(p)$ and so it is equal to $\Delta(p)$.

It follows from the orbit theorem that Δ gives the tangent space to the orbits and completes the proof.

To complete the prove it is enough to show that the orbits indeed form a foliation of X. This follows immediately from the local version of the Frobenius theorem. In fact, our distribution is constant in appropriate coordinates and so the connected components of intersections of leaves look like in the definition of a foliation. ∎

In order to define integrability of distributions which are not of constant dimension we have to weaken the notion of foliation. We will do this in such a way that partitions of the space X into orbits of a family of vector fields form foliations in this weaker sense.

Definition 2.18 A *foliation with singularities* is a partition

$$X = \cup_{\alpha \in A} S_\alpha$$

of X into immersed submanifolds such that, locally, there is a family of vector fields $\{g_b\}_{\beta \in B}$ such that $T_p S_\alpha = \text{span}\{g_\beta(p) \mid \beta \in B\}$ for all p and α.

A distribution on X is called *integrable* if there exists a foliation with singularities $\{S_\alpha\}_{\alpha \in A}$ which satisfies $T_p S = \Delta(p)$ for any p and S denoting the leaf which passes through p.

Theorem 2.19 *(Nagano) Any analytic distribution Δ is involutive.*

Proof. We take the partition of X into orbits of the famify of vector fields $\{f \mid f \in \Delta\}$ as a candidate for the integral foliation. From the orbit theorem and the corollary to it it follows that the tangent space to the leaf passing through p is equal to $\Gamma(p) = \Delta(p)$. This means that the partition into orbits is the integral foliation of Δ indeed. ■

Appendix: Global equivalence of families of vector fields

We close this chapter with a proof of sufficiency of the theorem about equivalence of families of vector fields and a global version of this result. The theorem of Nagano will be helpful in this proof.

Proof of Theorem 1.20. Sufficiency. In the proof we shall use the method of graph of Cartan and the theorem of Nagano. The method of graph consists of considering the product space $Z = X \times \tilde{X}$ and constructing the graph of the desired diffeomorphism $\Phi : X \longrightarrow \tilde{X}$ as an integral manifold of a distribution of vector fields on Z.

Define the product vector fields on Z by $h_u = f_u \times \tilde{f}_u$, $u \in U$, where, in $\mathbb{R}^n$,

$$f_u \times \tilde{f}_u = (f_u^1, \dots, f_u^n, \tilde{f}_u^1, \dots, \tilde{f}_u^n)^T.$$

Consider the distribution spanned by the Lie algebra Lie$\{H\}$ generated by the family $H = \{h_u\}_{u \in U}$ of these vector fields. The Nagano's theorem says that the distribution $Z \ni z \longrightarrow \text{Lie}\{H\}(z)$ is integrable, i.e. for each point $z \in Z$ there is an integral manifold of Lie$\{H\}$ passing through this point.

Take the point $z_0 = (p, \tilde{p}) \in Z$. We claim that the integral submanifold S passing through z_0 is of dimension n and it is the graph of a local diffeomorphism between X and $\tilde{X}$. Since S is the integral manifold, its dimension is equal to the dimension of the distribution Lie$\{H\}$ at z_0. But the vectors defining Lie$\{H\}(z_0)$ are of the form

$$h_{[u_1 \cdots u_k]} = (f_{[u_1 \cdots u_k]}, \tilde{f}_{[u_1 \cdots u_k]}) = (f_{[u_1 \cdots u_k]}, Lf_{[u_1 \cdots u_k]}),$$

the latter equality following from the assumption. From transitivity of Lie$\{\mathcal{F}\}$ at p and the above form of the vector fields $h_{[u_1 \cdots u_k]}$ it follows that the dimension of Lie$\{H\}(z_0)$ is at least n. On the other hand, since the second component of these vector fields depends on the first through the same linear map L, it follows that this dimension is precisely equal to n.

It follows that the integral submanifold S is of dimension n. To show that it defines a graph of a local diffeomorphism between X and $\tilde{X}$ we should check that the projections of the tangent space to S onto the tangent spaces of X and $\tilde{X}$ are "onto". From continuity, it is enough to show this at the point z_0. But $T_{z_0}S = \text{Lie}\{H\}(z_0)$ and the "ontoness" follows immediately from the above form of vectors $h_{[u_1 \cdots u_k]}$ and the transitivity of $\mathcal{F}$ and $\tilde{\mathcal{F}}$.

Let Φ be the local diffeomorphism from X to $\tilde{X}$ defined in a neighborhood of p via the submanifold S, $\Phi(p) = \tilde{p}$. Since among the vectors tangent to S there are vectors $h_u = (f_u, \tilde{f}_u)$, and S is the graph of Φ, it follows that there is the following relation between the domain component f_u of h_u and its codomain component $\tilde{f}_u$:

$$\tilde{f}_u(\Phi(x)) = D\Phi(x)f_u(x), \qquad or \qquad \tilde{f}_u(\tilde{x}) = D\Phi(x)f_u(x), \quad x = \Phi^{-1}(\tilde{x}).$$

The latter equality means that $\tilde{f}_u = \text{Ad}_\Phi f_u$, $u \in U$. The proof of sufficency is complete. ∎

Theorem 2.20 *(Sussmann) Assume that $\mathcal{F}$ and $\tilde{\mathcal{F}}$ are analytic transitive families of vector fields on compact, simply connected, analytic manifolds X and $\tilde{X}$ and the relation between the Lie brackets as in the local theorem holds. Then there exists a global diffeomorphism $\Phi : X \longrightarrow \tilde{X}$ such that $\text{Ad}_\Phi f_u = \tilde{f}_u$, $u \in U$.*

Proof. The proof of this theorem is an extension of the obove proof. Namely, it is enough to prove that under our assumptions the map Φ defined above is a global diffeomorphism.

We shall first show that the projection maps from S to X and $\tilde{X}$ are onto (and so they are coverings of X and of $\tilde{X}$, respectively). It is enough to show this for X. Take any point q on X. From the theorem of Chow and Rashevskii it follows that this point is reachable from p piecewice by the trajectories of the vector fields in F. Consider the point z_1 on S which corresponds to q and is reachable from z_0 piecewice by the lifted trajectories of the corresponding vector fields in H. It is easy to see that the projection of z_1 onto X is equal to q. Therefore, S is a covering of X.

As X is simply connected, it follows that this covering is a single covering, i.e. a diffeomorphism of S and X. In a similar way we show that the projection of S onto $\tilde{X}$ is adiffeomorphism. We conclude that the families F and $\tilde{F}$ are diffeomorphic. ∎

3 Controllability and accessibility

3.1 Basic definitions

We shall be dealing with two classes of control systems, the general nonlinear systems

$$\Sigma: \quad \dot{x} = f(x, u),$$

where $x(t) \in X$ and $u(t) \in U$, and the control-affine systems

$$\Sigma_{aff}: \quad \dot{x} = f(x) + \sum_{i=1}^{m} u_i g_i(x),$$

where $x(t) \in X$ and $u(t) = (u_1(t), \ldots, u_m(t)) \in U$. The state space X is assumed to be an open subset of $\mathbb{R}^n$ or a smooth differential manifold of dimension n. The control set U is an arbitrary set (with at least two elements), in the case of system Σ, and a subset of $\mathbb{R}^m$, in the case of Σ_{aff}. The vector fields $f_u = f(\cdot, u)$, defined by Σ, are assumed to be smooth (of class C^∞). Similarly, we assume that the vector fields f, $g_1, \ldots, g_m$ defined by Σ_{aff} are smooth. We will not need regularity of $f(x, u)$ in Σ with respect to u when we will use piecewise constant controls. Otherwise, we will assume that $f(x, u)$ together with the first partial derivatives with respect to u are smooth as functions of x and continuous with respect to (x, u).

We begin with the formal definition of reachable sets.

Definition 3.1 We shall call the set of points reachable from $x_0 \in X$ for system Σ its *reachable set from* x_0 and denote it by $\mathcal{R}(x_0)$. For the class of piecewise constant controls this is the set of points

$$\gamma_{t_k}^{u_k} \circ \cdots \circ \gamma_{t_1}^{u_1} x_0, \quad k \geq 1, \; u_1, \ldots, u_k \in U, \quad t_1, \ldots, t_k \geq 0.$$

Similarly, the set of above points with $t_1 + \cdots + t_k = t$ will be called the *reachable set at time t from* x_0 and denoted by $\mathcal{R}_t(x_0)$, and the set of such points with $t_1 + \cdots + t_k \leq t$ will be refered to as the *reachable set up to time t from* x_0 and denoted by $\mathcal{R}_{\leq t}(x_0)$.

It is unreasonable to expect that the reachable set of a nonlinear control system will have a simple structure, in general. Almost never it will be a linear subspace, even if $X = \mathbb{R}^n$ and $U = \mathbb{R}^m$. For example, for the system in the plane

$$\dot{x}_1 = u_1^2, \quad \dot{x}_2 = u_2^2$$

with $U = \mathbb{R}^2$ the reachable set from the origin is the positive ortant.

Therefore, our aim will be to establish qualitative properties of the reachable sets. One of such basic properties is the following.

Definition 3.2 We shall say that the system Σ is *accessible from* x_0 if its reachable set $\mathcal{R}(x_0)$ has a nonempty interior. Similarly, we will call this system *strongly accessible from* x_0 if the reachable set $\mathcal{R}_t(x_0)$ has a nonempty interion for any $t > 0$.

3.2 Taylor linearization

We begin with a presentation of a rough sufficient condition for strong accessibility.

Let (x_0, u_0) be an equilibrium point of our system Σ, i.e. $f(x_0, u_0) = 0$. Denote

$$A(x,u) = \frac{\partial f}{\partial x}(x,u), \quad B(x,u) = \frac{\partial f}{\partial u}(x,u),$$

and let $A_0 = A(x_0, u_0)$, $B_0 = B(x_0, u_0)$.

Theorem 3.3 *If $u_0 \in intU$ and the pair (A_0, B_0) satisfies the controllability rank condition, then the system is strongly accessible from x_0.*

A corresponding result result outside an equilibrium can be stated as follows. Let $u^*(\cdot)$ be an admissible control and let $x^*(\cdot)$ be the corresponding trajectory of system Σ. Denote

$$A(t) = \frac{\partial f}{\partial x}(x^*(t), u^*(t)), \quad B(t) = \frac{\partial f}{\partial u}(x^*(t), u^*(t)).$$

Theorem 3.4 *If $u^*(t) \in intU$ and the linear system $\dot{x} = A(t)x + B(t)x$, $x(0) = 0$ without constraints is controllable on the interval $[0, T]$, then the reachable set $\mathcal{R}_T(x_0)$ of system Σ has a nonempty interior. In particular, if the Grammian rank condition* $\mathrm{rank}G(0,t) = n$ *is satisfied for our linear system for any $t > 0$, then system Σ is strongly accessible.*

For the proof we shall need the following lemma of the theory of ordinary differential equations, which will be stated without proof.

Let $\bar{u}$ be a measurable, essentially bounded control and consider an admissible control in the form of the following variation

$$u_\epsilon = u^* + \epsilon \bar{u}.$$

Denote by x_ϵ the trajectory of system Σ, $x_\epsilon(0) = x_0$, corresponding to the control u_ϵ. Introduce the variation of the trajectory by

$$\bar{x}(t) = \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} x_\epsilon(t).$$

Lemma 3.5 *If $f = f(x,u)$ is of class C^1, then the variation of the trajectory satisfies the following equation, called variational equation*

$$\dot{\bar{x}} = A(x^*(t), u^*(t))\bar{x} + B(x^*(t), u^*(t))\bar{u}, \quad \bar{x}(0) = 0.$$

Both above theorems follow from the criteria on controllability of linear systems without constraints (presented in the section on controllability of linear systems) and from the following lemma.

Lemma 3.6 *If the variational system (treated as a linear system without constraints on the control) is controllable, then the original system is strongly accessible.*

Proof. Denote the matrices $A(t)$ and $B(t)$ as above. As the variational system is controllable, there exist (bounded) controls v^i which steer this system from 0 to $e_i = (0, \ldots, 1, \ldots, 0)^T$ (with 1 at i-th place) at time T, $i = 1, \ldots, n$. Take the control

$$u = u(\lambda_1, \ldots, \lambda_n) = \lambda_1 v^1 + \cdots + \lambda_n v^n.$$

When applied to the original system with the initial condition $x(0) = x_0$ it gives a final state $x(T)$ dependent on the parameters $\lambda(\lambda_1, \ldots, \lambda_n)$ in a differentiable way. In particular, the variation

$$\frac{\partial x(T)}{\partial \lambda_i}\bigg|_{\lambda=0} = \bar{x}^i$$

satisfies the variational equation with the control $\bar{u} = v^i$. As $\bar{x}^i(T) = e_i$, it follows that the Jacobi map of the nonlinear mapping

$$(\lambda_1, \ldots, \lambda_n) \longrightarrow x(T)$$

is of full rank. Therefore, it follows from the inverse function theorem that this mapping maps a neighborhood of the origin to a neighborhood of the origin. As $u(\lambda_1, \ldots, \lambda_n)$ form admissible controls, for λ_i small, it follows that the reachable set $\mathcal{R}_T(x_0)$ contains a neighborhood of the point $x^*(T)$. ■

3.3 Lie algebras of control system

We shall be using the following families of vector fields associated to the system Σ. Denote

$$f_u = f(\cdot, u),$$

and define the following families of vector fields

$$\mathcal{F} = \{f_u\}_{u \in U}, \quad \text{and} \quad \mathcal{G} = \{f_u - f_v \mid u, v \in U\}.$$

We define the *Lie algebra of system* Σ as the smallest linear space $\mathcal{L}$ of vector fields on X which contains the family $\mathcal{F}$ and is closed under Lie bracket:

$$f_1,\ f_2 \in \mathcal{L} \Rightarrow [f_1, f_2] \in \mathcal{L},$$

or equivalently

$$f_1 \in F,\ f_2 \in \mathcal{L} \Rightarrow [f_1, f_2] \in \mathcal{L}.$$

Exercise. Prove equivalence of both conditions using the property of Lie bracket expressed in the Proposition in the Appendix on Lie algebras, Section 1.1.

We also define the *Lie ideal of system* Σ as the smallest linear space $\mathcal{L}_0$ of vector fields on X which contains the family $\mathcal{G}$ and is closed under taking Lie brackets with the elements of $\mathcal{F}$:

$$f_1 \in F,\ f_2 \in \mathcal{L}_0 \Rightarrow [f_1, f_2] \in \mathcal{L}_0.$$

$\mathcal{L}_0$ is closed under Lie bracket and so is a Lie algebra in the usual sense (cf. Appendix: Lie Algebras in Section 1).

From both definitions it follows immediately that $\mathcal{L}$ and $\mathcal{L}_0$ can be equivalently defined through the iterative Lie brackets as follows

$$\mathcal{L} = \text{span}\{[f_{u_1}, \cdots, [f_{u_{k-1}}, f_{u_k}] \cdots] \mid k \geq 1,\ u_1, \cdots, u_k \in U\ \},$$

$$\mathcal{L}_0 = \text{span}\{[f_{u_1}, \cdots, [f_{u_{k-1}}, f_{u_k} - f_{u_{k+1}}] \cdots] \mid k \geq 2,\ u_1, \cdots, u_{k+1} \in U\ \}.$$

It follows then that

$$\mathcal{L} = \text{span}\{f_{u^*},\ \mathcal{L}_0\ \},$$

where u^* is any fixed element of U. In fact, directly from the definitions we obtain that $\mathcal{L}_0 \subset \mathcal{L}$, and also $f_{u^*} \in \mathcal{L}$. The converse inclusion $\mathcal{L} \subset \text{span}\{f_{u^*}, \mathcal{L}_0\}$ follows from the equalities

$$f_{u_1} = f_{u^*} + f_{u_1} - f_{u_2}, \quad u_2 = u^*,$$

$$[f_{u_1}, \cdots, [f_{u_{k-1}}, f_{u_k}] \cdots] = [f_{u_1}, \cdots, [f_{u_{k-1}}, f_{u_k} - f_{u_{k+1}}] \cdots],$$

where $u_{k+1} = u_{k-1}$.

For the control-affine system Σ_{aff} the corresponding Lie algebras can be expressed as

$$\mathcal{L} = \text{Lie}\{f, g_1, \ldots, g_m\} = \text{span}\{[g_{i_1}, \cdots, [g_{i_{k-1}}, g_{i_k}] \cdots] \mid k \geq 1,\ 0 \leq i_1, \ldots, i_k \leq m\},$$

$$\mathcal{L}_0 = \text{span}\{[g_{i_1}, \cdots, [g_{i_{k-1}}, g_{i_k}] \cdots] \mid k \geq 1,\ 0 \leq i_1, \ldots, i_k \leq m,\ i_k \neq 0\},$$

where $g_0 = f$.

Example 3.7 For illustration and also for further use we shall compute the Lie algebra and the Lie ideal of the linear system

$$\dot{x} = Ax + Bu = Ax + \sum_{i=1}^{m} u_i b_i,$$

where b_i are constant vector fields being columns of the matrix B. Taking into account that $g_1 = b_1, \ldots, g_m = b_m$, $f = g_0 = Ax$, and that Lie bracket of constant vector fields is zero, we find that in the above formula for $\mathcal{L}_0$ the only nonzero iterated Lie brackets are

$$[Ax, b_i] = -Ab_i, \quad [Ax, [Ax, b_i]] = [Ax, -Ab_i] = A^2 b_i, \ldots,$$

$$\text{ad}_{Ax} \cdots \text{ad}_{Ax} b_i = \text{ad}^j_{Ax} b_i = (-1)^j A^j b_i.$$

Therefore, the Lie ideal $\mathcal{L}_0$ consists of constant vector fields only,

$$\mathcal{L}_0 = \text{span}\{A^j b_i \mid j \geq 0,\ 1 \leq i \leq m\} = \text{span}\{A^j b_i \mid 0 \leq j \leq n-1,\ 1 \leq i \leq m\},$$

and $\mathcal{L} = \text{span}\{Ax, \mathcal{L}_0\}$.

3.4 Accessibility criteria

Given a family of vector fields $\mathcal{H}$, we shall use the notation

$$\mathcal{H}(x) = \text{span}\{\ h(x) \mid h \in \mathcal{H}\ \}.$$

In particular, $\mathcal{L}(x)$ and $\mathcal{L}_0(x)$ will denote the space of tangent vectors at x defined by the Lie algebra and the Lie ideal of system Σ. The following result was first proved by Jurdjevic and Sussmann [?].

Theorem 3.8 *(a) If for a state smooth system Σ the Lie algebra is of full rank at x_0, $\dim\mathcal{L}(x_0) = n$, then the attainable set up to time t from x_0 has the nonempty interior and so the system is accessible from x_0.*

(b) If the system is state analytic and $\dim\mathcal{L}(x_0) < n$, then the system is not accessible from x_0.

We present a proof of the first statement (due to A. Krener) which is very simple and gives insight to the problem of accessibility.

Proof of (a). It follows from the assumption $\dim\mathcal{L}(x_0) = n$ that $\dim\mathcal{L}(x) = n$ for x in a neighborhood of x_0 (the full rank is realized by n vector fields which are linearly independent in a neighborhood of x_0). It also follows from the same assumption that there is a $u_1 \in U$ such that $f_{u_1}(x_0) \neq 0$. Otherwise, it would follow from the Jacobian definition of Lie bracket that all the vector fields in $\mathcal{L}$ vanished at x_0 and so $\dim\mathcal{L}(x_0) = 0$. The trajectory $\gamma_{t_1}^{u_1} x_0$, $t \in V_1 = (0, \epsilon_1)$, $\epsilon_1 > 0$, forms a one dimensional submanifold of X which we denote by S_1.

We now claim that there is a $u_2 \in U$ such that the vector fields f_{u_1} and f_{u_2} are linearly independent at a point $x_1 \in S_1$. Otherwise, all the vector fields in $\mathcal{F}$ would be tangent to the submanifold S_1. As taking linear combinations and Lie bracket of vector fields tangent to a submanifold gives vector fields tangent to this submanifold, we would have that all the vector fields in $\mathcal{L}$ were tangent to S_1 which would contradict $\dim\mathcal{L}(x_0) = n$ (if $n > 1$).

Let f_{u_1} and f_{u_2} be linearly independent at $x_1 = \gamma_{t_1}^{u_1} x_0 \in S_1$, $0 < t_1 < \epsilon_1$. Define the map

$$V_2 \ni (t_1, t_2) \longrightarrow x = \gamma_{t_2}^{u_2} \circ \gamma_{t_1}^{u_1} x_0,$$

where V_2 is an open subset of $\mathbb{R}^2$: $V_2 = (0, \epsilon_1) \times (0, \epsilon_2)$, $\epsilon_2 > 0$. For ϵ_2 sufficiently small the image of this map contains a submanifold of X of dimension 2 (this follows from linear independence of f_{u_1} and f_{u_2}) which we denote by S_2.

By an argument analogous to the above there exists a $u_3 \in U$ and a point $x_2 \in S_2$ such that the vector field f_{u_3} is not tangent to S_2 at x_2. Thus the image of the map

$$V_3 \ni (t_1, t_2, t_3) \longrightarrow x = \gamma_{t_3}^{u_3} \circ \gamma_{t_2}^{u_2} \gamma_{t_1}^{u_1} x_0$$

(where $V_3 = (0, \epsilon_1) \times (0, \epsilon_2) \times (0, \epsilon_3)$) contains a submanifold S_3 of X of dimension 3. Of course, S_i, $i = 1, 2, 3$ are subsets of the reachable set.

After n steps of such a construction we obtain a submanifold S_n of X of dimension n, i.e. an open subset of X, which is contained in the reachable set $\mathcal{R}(x_0)$ and, more precisely, in the reachable set $\mathcal{R}_{\leq t}(x_0)$, where $t = \epsilon_1 + \cdots + \epsilon_n$. Since $\epsilon_1, \ldots, \epsilon_n$ could have been taken arbitrarily small, it follows that any attainable set $\mathcal{R}_t$, $t > 0$ has the nonempty interior.

Proof of (b). From the corollary to the orbit theorem it follows that the tangent space to the orbit from x_0 is equal to $L(x_0)$. When $\dim L(x_0) < n$, it follows that this orbit is a submanifold of dimension smaller then n. Thus, its interior is empty. As the reachable set is a subset of the orbit, its interior is empty also. ■

The analyticity assumption in statement (b) can not be dropped. This can be seen in the example presented after the orbit theorem in Section 2 (showing that in the smooth case we can have $\Gamma(x) \neq L(x)$) by taking an initial point with positive second coordinate.

If the dimension of the Lie algebra of the system is not full at some point, still we have the following positive result.

Corollary 3.9 *If the system* Π *is state analytic, then the interior in the orbit* $\mathrm{Orb}(x_0)$ *of the reachable set* $\mathcal{R}(x_0)$ *is nonempty.*

Proof. If $\dim\mathcal{L}(x_0) = n$, then this is simply statement (a) of our theorem. When this dimension is smaller we can restrict our system to the orbit passing through the initial point. The corollary to the orbit theorem says that $\dim\mathcal{L}(x_0)$ is equal to the dimension of the orbit. Thus, our system reduced to the orbit satisfies the assumptions of statement (a) of our theorem and our result follows. ■

Example 3.10 Consider the system with the scalar control $u \in U = \mathbb{R}$

$$\dot{x}_1 = u, \quad \dot{x}_2 = x_1^k, \qquad k \geq 2.$$

It is easy to check that the Taylor linearization of this system, at the equilibrium $x_0 = 0$ and $u_0 = 0$, is not controllable. Our system is control-affine with $f = (0, x_1^k)^T$ and $g = (1,0)^T$. Then

$$[g, f] = (0, kx_1^{k-1})^T, \quad [g, [g, f]] = (0, k(k-1)x_1^{k-2})^T, \quad \mathrm{ad}_g^k f = (0, k!)^T,$$

and so $\dim\mathcal{L}_0(x) = \dim\mathcal{L}(x) = 2$ for all x, in particular the system is strongly accessible from the origin.

There is an analogous relation between the Lie ideal $\mathcal{L}_0$ and the attainable set at time t which is established by the following theorem.

Theorem 3.11 *(a) If the system is state smooth and* $\dim\mathcal{L}_0(x_0) = n$, *then the attainable set* $\mathcal{R}_t(x_0)$ *has a nonempty interior for any* $t > 0$.

(b) If $\dim\mathcal{L}_0(x_0) < n$, *then* $\text{int } \mathcal{R}_t(x_0) = \emptyset$ *for any* $t > 0$.

Example 3.12 Consider the system on $\mathbb{R}^2$

$$\dot{x}_1 = 1, \quad \dot{x}_2 = u\, x_1^2,$$

and take $x_0 = (0,0)$, and $U = \mathbb{R}$. We have

$$\mathcal{F} = \{(1, u\, x_1^2)^T \mid u \in \mathbb{R}\}, \quad \mathcal{G} = \text{span}\{(0, x_1^2)^T\}.$$

The Lie algebra $\mathcal{L}$ contains the vector fields

$$f_1 = (1,0)^T, \quad f_2 = (1, x_1^2)^T, \quad f_3 = [f_1, f_2] = (0, 2x_1)^T, \quad [f_1, f_3] = (0,2)^T.$$

Therefore, $\dim\mathcal{L}(x_0) = 2$ and so the system is accessible from x_0. (Note that one gets the same result if the set U is restricted to two values $U = \{0, 1\}$). On the other hand $\mathcal{L}_0(x_0) = \text{span}\{(0,1)^T\}$ and so the interior of the attainable set at time t, $t > 0$, is empty. In fact, it can be proved that the attainable set $\mathcal{R}(x_0)$ is equal to the open right half plain including the origin and the set $\mathcal{R}_t(x_0)$ is equal to the set $x_1 = t$, $x_2 \in \mathbb{R}$.

Example 3.13 *Accessibility of linear systems without constraints.* As we computed earlier, for autonomous linear system Λ with unconstrained control we have $\mathcal{L}(x) = Im[B, AB, \ldots, A^{n-1}B]$ and $\mathcal{L}(x) = \text{span}\{Ax, \mathcal{L}(x)\}$. Thus, such a system is strongly accessible from x if and only if the controllability matrix

$$[B, AB, \ldots, A^{n-1}B]$$

is of rank n (such linear systems are called controllable).

Noncontrollable linear system may be accessible from x. This happens when $\dim\mathcal{L}(x) = n$ but $Ax \notin Im[B, AB, \ldots, A^{n-1}B]$. Then the system is accessible from those x at which Ax is not in the image of the controllability matrix. The system is not accessible from the linear subspace of points at which Ax is in this image (this subspace is the counterimage under A of the image of the controllability matrix).

Exercise. Analyse the orbits of the linear system without constrains on the control. Show that this system may have one orbit, three orbits, or a continuum of orbits. (The Kalman decomposition theorem from the first section is helpful here).

Example 3.14 *Accessibility of linear systems with constraints.* Consider now a linear autonomous system with the constraints $u(t) \in U \subset \mathbb{R}^m$. If the interior of U is nonempty, then the controllability rank condition implies that the system is strongly accessible, as we have already established in the section about linear systems. When U has the empty interior then the situation is more complicated. To analyse this case it is more convenient to consider our system in the form

$$\dot{x} = Ax + v, \quad v \in V,$$

where V is the image of U under the linear map $B : \mathbb{R}^m \longrightarrow \mathbb{R}^n$. Let us introduce the set

$$W = \{v' - v'' | v', v'' \in V\}.$$

The one can easily compute that the Lie algebra of our system contains the vector fields $Ax + v' - (Ax + v'') = v' - v'' \in W$, i.e. all constant vector fields $f = w$, $w \in W$. Thus, it contains also the Lie brackets $[w, Ax + v] = Aw$, $w \in W$, and by induction it contains all the constant vector fields $A^i w$, $i \geq 0$, $w \in W$. It follows then from the Cayley-Hamilton theorem that the linear system with constraints is strongly accessible from x_0 if and only if

$$\text{dimspan}\{A^i w | 0 \leq i \leq n-1, w \in W\} = n.$$

It is accessible from x_0 if and only if the same collection of vectors together with any fixed vector $Ax_0 + v$, $v \in V$, span the whole space.

Example 3.15 *Space-craft with two jets* Consider a spacecraft with two pairs of jets placed so that they angular momenta are parallel to principal axes of the spacecraft. Then, the equations of motion for the angular velocities take the form

$$\begin{aligned} \dot{\omega}_1 &= a_1\omega_2\omega_3 + u_1, \\ \dot{\omega}_2 &= a_2\omega_3\omega_1 + u_2, \\ \dot{\omega}_3 &= a_3\omega_1\omega_2. \end{aligned}$$

Our system is control-affine with

$$f = (a_1\omega_2\omega_3, a_2\omega_3\omega_1, a_3\omega_1\omega_2)^T, \quad g_1 = (1,0,0)^T, \quad g_2 = (0,1,0)^T.$$

We compute

$$[f, g_1] = -(0, a_2\omega_3, a_3\omega_2)^T, \quad [f, g_2] = -(a_1\omega_3, 0, a_3\omega_1)^T,$$

$$[g_1, [g_2, f]] = (0, 0, a_3)^T.$$

It follows easily that

$$\dim\mathcal{L}_0(x) = 2 \Longleftrightarrow \dim\mathcal{L}(x) = 2 \Longleftrightarrow a_3 \neq 0,$$

for any $x = (\omega_1, \omega_2, \omega_3)$. Here the coefficient a_3 is equal to $(I_1 - I_2)/I_3$, where I_1, I_2, I_3 are the momenta of inertia along the principal axes. It follows then that the above system is accessible (equivalently, strongly accessible) if and only if the momenta of inertia of the space-craft along the axes with two pairs of jets are different.

Example 3.16 *Space-craft with one jet.* The analysis of the space-craft with one jet, with the equations

$$\begin{aligned} \dot{\omega}_1 &= a_1\omega_2\omega_3 + u, \\ \dot{\omega}_2 &= a_2\omega_3\omega_1, \\ \dot{\omega}_3 &= a_3\omega_1\omega_2, \end{aligned}$$

gives a different result. We have that

$$f = (a_1\omega_2\omega_3, a_2\omega_3\omega_1, a_3\omega_1\omega_2)^T, \quad g = (1,0,0)^T,$$

$$[f, g] = -(0, a_2\omega_3, a_3\omega_2)^T = -(\omega_1)^{-1} f + (\omega_1)^{-1} a_1\omega_2\omega_3 g.$$

Computing the higher order Lie brackets does not give anything new:

$$[g, [f, g]] = 0, \quad [f, [f, g]] = (*, 0, 0)^T = \phi g,$$

where ϕ is a function. It follows that

$$\mathcal{L}(x) = \text{span}\{g(x), [f, g](x)\}$$

and these two vector fields span an involutive distribution. From the form of g and $[f, g]$ it follows that the orbits of the system consist of the Cartesian product of lines along the first coordinate and the trajectories of the vector field $(a_2\omega_3, a_3\omega_2)^T$ along the last two coordinates. In particular, if $a_2 \neq 0 \neq a_3$, then there is one 1-dimensional orbit of the system (the first coordinate axis) corresponding to the equilibrium of the vector field $(a_2\omega_3, a_3\omega_2)^T$, and continuum of 2-dimensional orbits. Our system is not accessible from any $x = (\omega_1, \omega_2, \omega_3)$. We conclude that if there is only one pair of jets which gives the angular momentum parallel to one of the principal axes of inertia of the space-craft then, contrary to the case of two pairs of jets, the system is never accessible.

3.5 Time-symmetric systems

In general, the reachable set is a proper subset of the orbit. It is reasonable to ask for which systems the reachable set coincides with the orbit. One class of such systems is called time-symmetric systems.

Definition 3.17 A system Σ is called *time-symmetric* if for any value $u \in U$ of the control there is another value $v \in U$ such that

$$f(x, u) = -f(x, v), \quad \text{for any} \quad x \in X.$$

Proposition 3.18 *For any time-symmetric system and piecewise constant controls we have that* $\mathcal{R}(x_0) = \mathrm{Orb}(x_0)$.

Proof. The definition of the reachable set (with constant controls) and the definition od the orbit differ in the fact that it is not allowed to go forward in time along trajectories of the vector fields $f_u = f(\cdot, u)$, in the definition of the reachable set. For a time-symmetric system going forward with the control v in the above definition is the same as going backward with the control u. Therefore, for a time-symmetric system the points which are forward-backward reachable are also forward reachable and the proposition follows. ■

Proposition 3.19 *For any state smooth time-symmetric system we have that*

$$\dim\mathcal{L}(x_0) = n \Longrightarrow x_0 \in \text{int}\mathcal{R}(x_0).$$

Proof. From our theorem on accessibility of systems which satisfy the above Lie algebra rank condition it follows that the reachable set has a nonempty interior. Let x_1 be a point in this interior, where

$$x_1 = \gamma_{t_k}^{u_k} \circ \cdots \circ \gamma_{t_1}^{f_1}(x_0).$$

Let $v_1, \ldots, v_k$ be the controls corresponding to $u_1, \ldots, u_k$ in our definition of time-symmetric systems. Then the point

$$x_2 = \gamma_{-t_k}^{v_1} \circ \cdots \circ \gamma_{-t_1}^{v_k} \circ \gamma_{t_k}^{u_k} \circ \cdots \circ \gamma_{t_1}^{u_1}(x_0)$$

coincides with x_0. This point is also in the interior of the reachable set as flows $\gamma_{-t_i}^{v_i}$ are local diffeomorphisms and map neigborhood of points into neighborhood of points. It follows that x_0 lies in the interior of the reachable set from x_0. ■

As a corollary we obtain another proof of the Chow's theorem (this proof is independent of the orbit theorem).

Corollary 3.20 *If our system is time-symmetric and* $\dim\mathcal{L}(x) = n$ *for all* $x \in X$*, then any point of* X *is forward reachable from any other by piecewise controls, i.e.* $\mathcal{R}(x) = X$ *for any* $x \in X$.

Proof. From the above propositions it follows that the reachable set coincides with the orbit and is open (as after reaching any point we can also reach a neighborhood of it be the second proposition). As X is a disjoint union of orbits, it is a disjoint union of open orbits. From connectedness of X it follows that X consists of one orbit, which is the orbit of any x_0 at the same time. As the reachable set of x_0 coincides with the orbit, it equals to X. ■

Example 3.21 Our example of the motion of a car (Examples 1.3 and 1.10) gives a time-symmetric system if, together with the forward motions given by the vector fields f and g we introduce also backward motions $-f$ and $-g$. It follows from the above result that the reachable set is the whole X, which means that we can reach any position of the car. In fact, a much stronger result can be proved. Namely, the movement of the car can "approximately follow" any trajectory in its state space.

In fact, a more general fact takes place. Namely, it can be proved that for any time-symmetric system which satisfies the Lie algebra rank condition any curve in the state space X can be approximately followed by trajectories of the system.

References

[Bo] William M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press, New York 1975.

[Sp] Michael Spivak, *Differential Geometry*, Vol.1 Publish or Perish, Inc. Berkeley 1979.

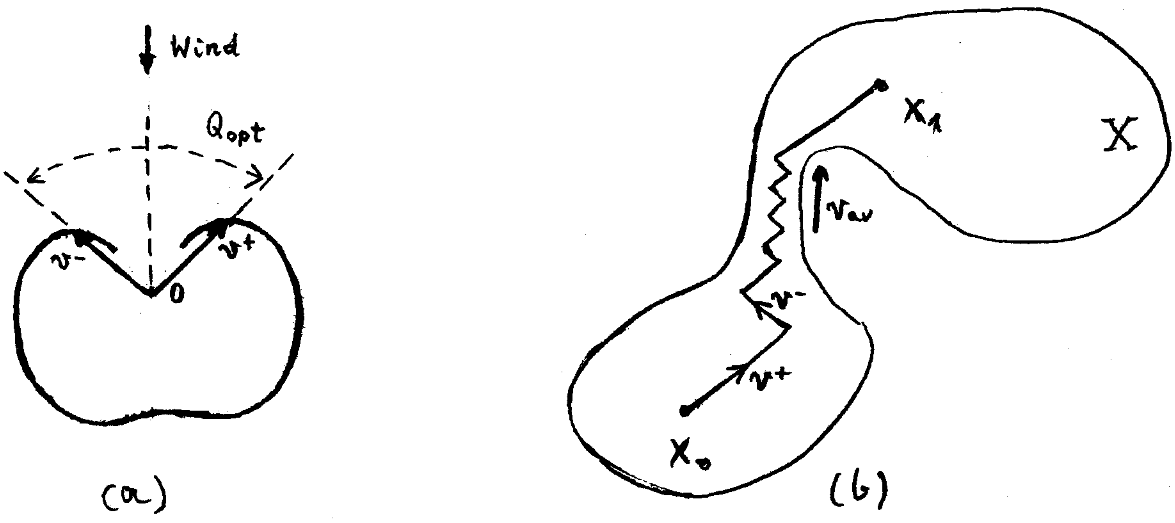

fig.1

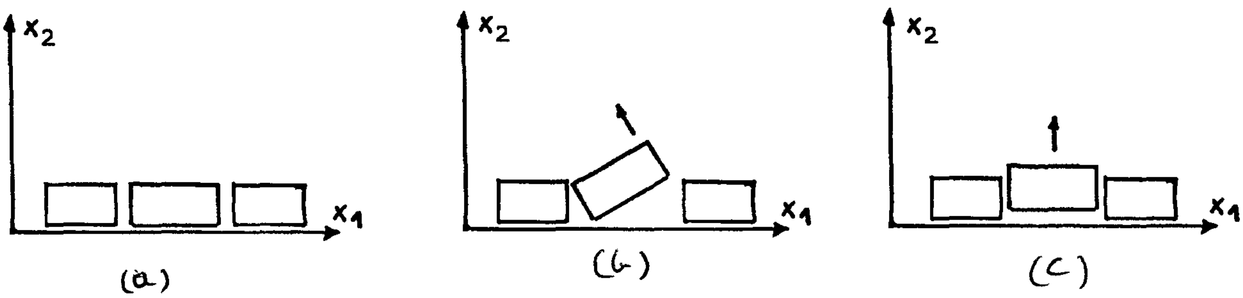

fig.2

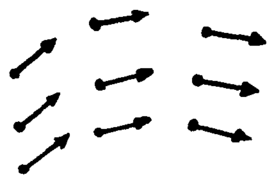

fig.3

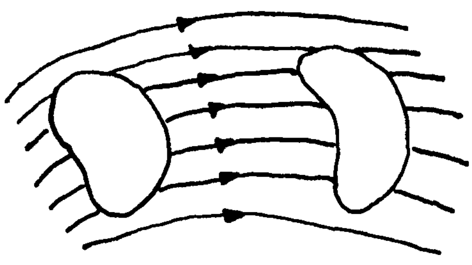

fig.4

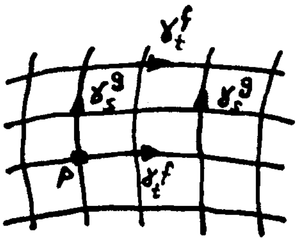

fig.5

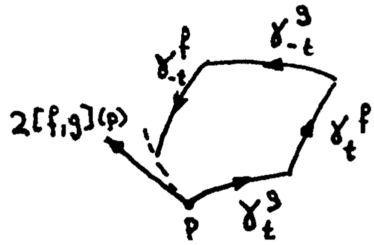

fig.6

fig.7

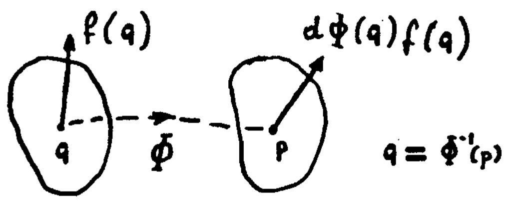

fig.8

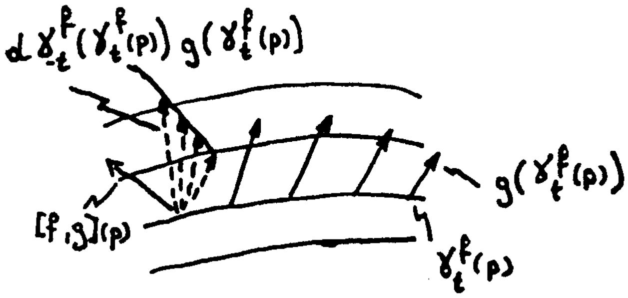

fig.9

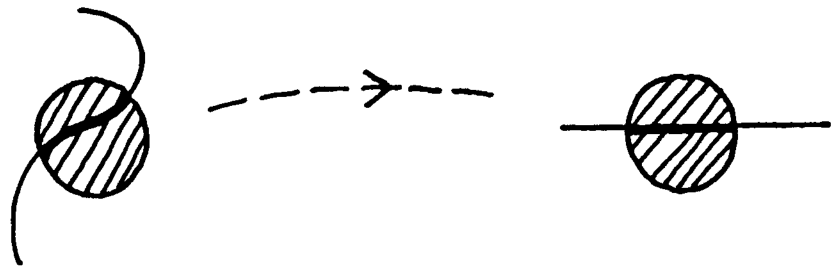

fig.10

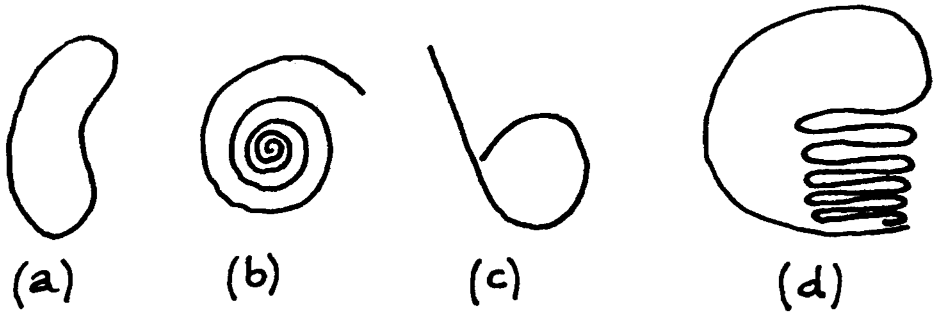

fig.11

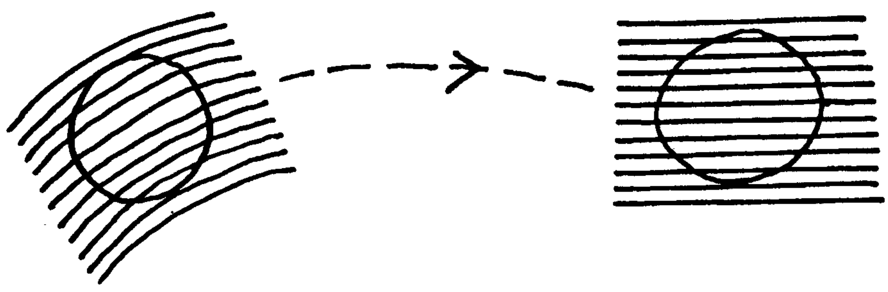

fig.12

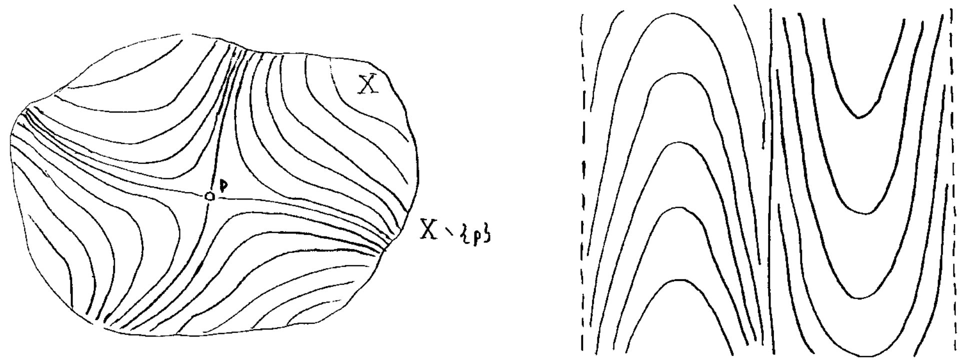

fig.13