Controllability and Observability of Nonlinear Systems

M. R. James[1]

October, 1986
Revised October, 1987

[1]Mathematics Department and Systems Research Center, University of Maryland, College Park, MD 20742, USA.

Abstract

These *tutorial notes* discuss the basic ideas in the theory of controllability and observability for nonlinear control systems. The theory treated is primarily due to Hermann and Krener. The first section gives a short overview of the issues, followed by section 2 which reviews distributions, codistributions, and the Frobenius Theorem. Section 3 deals with controllability. Chow's Theorem is presented, before beginning the Hermann–Krener theory. Finally, section 4 discusses the Hermann–Krener formulation of observability. A number of examples and illustrations are provided.

Key words: Controllability, observability, nonlinear systems.

Controllability and Observability of Nonlinear Systems

1. Introduction

Let M be a smooth manifold of dimension n. Denote by U an open neighborhood of $x^0 \in M$. We consider a control system Σ, described in local coordinates by

$$\Sigma : \begin{cases} \dot{x}(t) = f(x(t), u(t)) = f(x(t)) + \Sigma_{i=1}^{m} g_i(x(t))u_i(t), \\ y(t) = h(x(t)), \qquad\qquad x(0) = x^0, \end{cases}$$

where $t \mapsto u_i(t)$ is a *control* function with values in a convex set $\Omega \subset \mathbb{R}$, $t \mapsto x(t)$ is the *state* trajectory with $x(t) \in M$ and $t \mapsto y(t)$ is the *output* curve with $y(t) \in \mathbb{R}^p$.

Given system Σ initialized at x^0, the map

$$S_{x^0} : \{t \mapsto u(t),\ t \in [0,T]\} \longrightarrow \{t \mapsto y(t), t \in [0,T]\}$$

is called the *input-output map.*

Given a control $t \mapsto u(t)$, let γ_u denote the corresponding flow:

$$x(t) = \gamma_u(t)x^0.$$

We consider the following two concepts, with an emphasis on local ideas.

a. Controllability

Suppose we are given a system Σ and an initial state x^0. Let x^1 be another state. Is it possible to choose a control $t \mapsto u(t)$ to steer Σ from x^0 to x^1?

(This is often referred to as *reachability*, here x^1 is reachable from x^0.)

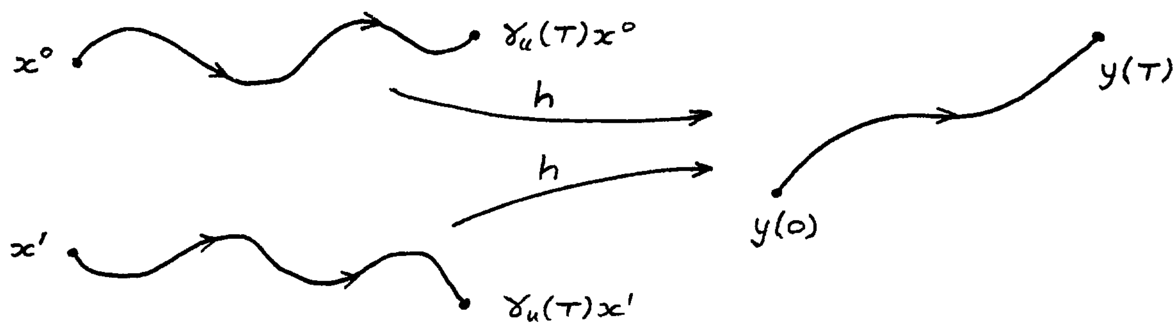
If so, x^1 is *accessible* from x^0. What are the accessible states? Is x^0 accessible from x^1? Is x^1 accessible from x^0 locally?

Σ is *controllable* if every state is accessible from every other state. What criteria (for example, algebraic) tell us when Σ is controllable (or has some weaker property)?

b. Observability

This time we are given an output "record" $t \mapsto y(t)$, $t \in [0,T]$. We ask what information about the states can be obtained from such a record.

Two initial states x^0, x^1 are *indistinguishable* if no matter what control we use, the corresponding trajectories always produce the same output record.



What are the indistinguishable/distinguishable states? Can states be locally distinguished?

Σ is *observable* if any state is distinguishable from any other state. What criteria is available here, perhaps for weaker concepts?

In a sense, observability is a "dual" notion to controllability.

What follows is based on Hermann and Krener [1], and Isidori [2].

2. Distributions and Codistributions

Notation:

$\mathcal{F}(M) =$ ring of smooth functions $M \to \mathbb{R}$

$\mathcal{X}(M) =$ smooth vector fields on M.

(a Lie algebra and a module over $\mathcal{F}(M)$)

$\mathcal{X}^*(M) =$ smooth 1-forms on M

(a module over $\mathcal{F}(M)$)

Definitions

1. A *distribution* $\mathcal{D}$ is a submodule of $\mathcal{X}(M)$.
2. $D(x) \equiv \{X(x) : X \in \mathcal{D}\} \leq T_xM$ (subspace of)
3. $D = \cup_{x \in M} D(x)$ (a *(singular) subbundle* of TM)
4. If $\dim D(x)$ is constant, we say D is *nonsingular.*
5. A set of linearly independent vector fields $\{X_1, \ldots, X_k\}$ in a neighborhood U of x is called a *local basis* (frame) if

$$D(y) = \operatorname{span}\{X_1(y), \ldots, X_k(y)\} \text{ for all } y \in U.$$

6. A point x is called a *regular point* of $\mathcal{D}$ if $\dim \mathcal{D}(y) = \dim \mathcal{D}(x)$ for all $y \in U, U$ a neighborhood of x. Otherwise, x is a *singular point.*
7. $\Gamma(D) = \{X \in \mathcal{X}(M) : X(x) \in D(x)\ \forall\, x \in M\}$ (a distribution)
8. $\mathcal{D}$ is *complete* if $\mathcal{D} = \Gamma(D)$. We shall assume that all distributions are complete.

Lemma *A point x is a regular point of $\mathcal{D}$ if and only if there exists a local basis in a neighborhood of x.*

Definitions

1. A *codistribution* $\mathcal{E}$ is a submodule of $\mathcal{X}^*(M)$. $\mathcal{E}$ is sometimes called a *Pfaffian system.*
2. $E(x) = \{\omega(x) : \omega \in \mathcal{E}\} \leq T_x^*M$
3. $E = \cup_{x \in M} E(x)$ (a *(singular) subbundle* of T^*M)
4. Analogous definitions for nonsingularity, completeness, local frame, etc.

Duality

$$\mathcal{D}^{\perp} = \{\omega \in \mathcal{X}^*(M) : \langle \omega, X \rangle = 0 \ \forall \ X \in \mathcal{D}\}$$
$$\mathcal{E}^{\perp} = \{X \in \mathcal{X}(M) : \langle \omega, X \rangle = 0 \ \forall \ \omega \in \mathcal{E}\}$$

$\mathcal{D}^{\perp}$ (resp. $\mathcal{E}^{\perp}$) is called the *annihilator* of $\mathcal{D}$ (resp. $\mathcal{E}$) and is a codistribution (distribution).

Invariance Let $X, Y \in \mathcal{X}(M)$, $\omega \in \mathcal{X}^*(M)$.

$$ad_X Y = L_X Y = [X, Y] \qquad \text{(Lie derivative)}$$

$\mathcal{D}$ is *ad_X invariant* (or, *invariant under X*) if $Y \in \mathcal{D}$ implies $ad_X Y \in \mathcal{D}$.

$$ad_X \omega = L_X \omega \qquad \text{(Lie derivative)}$$

$\mathcal{E}$ is *ad_X invariant* (*invariant under X*) if $\omega \in \mathcal{E}$ implies $ad_X \omega \in \mathcal{E}$.

Integrability

Definitions

1. A distribution $\mathcal{D}$ is *involutive* if for all $X \in \mathcal{D}$, $\mathcal{D}$ is ad_X invariant.
2. An *integral submanifold* N of $\mathcal{D}$ is a connected immersed submanifold $N \subset M$ such that for all $x \in N$, $T_x N \leq D(x)$. A *maximal integral submanifold* is an integral submanifold not properly contained in any other integral submanifold.
3. A distribution $\mathcal{D}$ is *integrable* if its maximal integral submanifolds define a partition of M. This is called a *foliation*, the maximal integral submanifolds being called *leaves*.

Theorem *(Frobenius) Let $\mathcal{D}$ be a nonsinguar distribution. Then $\mathcal{D}$ is integrable if and only if $\mathcal{D}$ is involutive.*

Following Boothby [6], we state and prove (sketch) a local version of the Frobenius theorem.

Theorem *(A local version of Frobenius Theorem) Let $p \in M$ be a regular point of $\mathcal{D}$. Then, on a neighborhood U of p, $\mathcal{D}$ is (completely) integrable if and only if $\mathcal{D}$ is involutive.*

Since p is regular point, if $\mathcal{D}$ is k-dimensional, there exists a local basis $\{X_1, \ldots, X_k\}$ in a neighborhood U of p. To say that $\mathcal{D}$ is involutive means that

$$[X_i, X_j] = \Sigma_{l=1}^{k} c_{ij}^{l} X_l, \ \ 1 \leq i, j \leq k$$

for some $c_{ij}^l \in \mathcal{F}(U)$.

$\mathcal{D}$ is (completely) integrable if there exists local coordinates $x_1, \ldots, x_n$ in U such that $\{\frac{\partial}{\partial x_i},\ \ i = 1, \ldots, k\}$ form a local basis for $\mathcal{D}$. In this case the k-dimensional submanifold

$$\{x_1, \ldots, x_n : x^{k+1} = c^1, \ldots, x^n = c^{n-k}\}$$

is an integral submanifold of $\mathcal{D}$ in U. A submanifold expressed in this form is called a *slice.*

Proof If $\mathcal{D}$ is (completely) integrable, then it is involutive, since

$$[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}] = 0, \qquad 1 \leq i,\ j \leq k.$$

Next, suppose that $\mathcal{D}$ is involutive. Proceed by induction. (The ideas are sketched).

$k = 1$. In this case $\mathcal{D}$ is determined by a vector field on U, call it X. Let integral curves of X define the coordinate y_1, that is, choose coordinates $y_1, \ldots, y_n$ such that $X = \frac{\partial}{\partial y_1}$. Hence $\mathcal{D}$ is integrable. The integral submanifolds are the integral curves. So in this case the theorem is just existence of local flows.

Suppose that the theorem is true for distributions of dimensions $1, \ldots, k-1$.

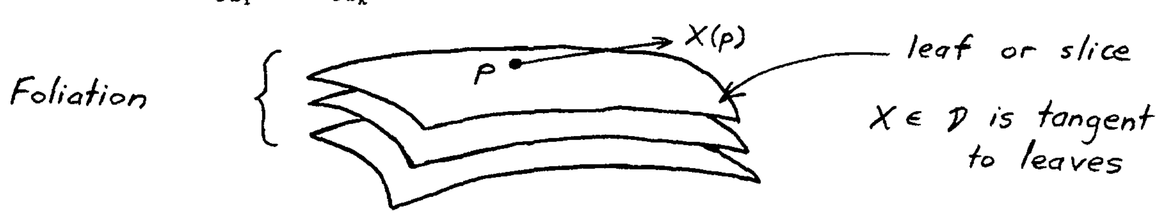
Let $X_1, \ldots, X_k$ be a local basis for $\mathcal{D}$ and $y_1, \ldots, y_n$ coordinates with $X_1 = \frac{\partial}{\partial y_1}$. Change basis

$$\begin{aligned} Y_1 &= X_1 \\ Y_i &= X_i - (X_i\, y_1)X_1\ , \qquad 2 \leq i \leq k. \end{aligned}$$

Then $Y_2, \ldots, Y_k$ are involutive and set $N_0 = \{y_1 = 0\}$. Change coordinates on N_0 : $y_2, \ldots, y_n \mapsto x_2, \ldots, x_n$ so that

$$\text{span } \{Y_2, \ldots, Y_k\} = \text{ span } \{\frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_k}\},$$

(by induction hypothesis). Extend this to a change of coordinates on U by setting $X_1 = y_1$. Then check that $\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_k}\}$ forms a local basis for $\mathcal{D}$ on U. $\square$



Definitions

1. We say that a codistribution $\mathcal{E}$ is *integrable* if its annihilator $\mathcal{E}^\perp$ is integrable.
2. Let $h : M \to \mathbb{R}^p$ be smooth.

$$\mathcal{R}(dh) = \text{ codistribution spanned by } dh_1, \ldots, dh_p.$$

Lemma *$\mathcal{R}(dh)$ is integrable.*

Proof First, we check if $\mathcal{E}$ is ad_X invariant, then $\mathcal{E}^\perp$ is ad_X invariant. Let $\omega \in \mathcal{E}$, $Y \in \mathcal{E}^\perp$. Then $\langle \omega, Y \rangle = 0$ and $ad_X \omega \in \mathcal{E}$. Now

$$\begin{aligned} 0 = L_X \langle \omega, Y \rangle &= \langle \omega, Y \rangle + \langle \omega, L_X, Y \rangle \\ &= 0 + \langle \omega, L_X Y \rangle. \end{aligned}$$

Hence $L_X Y \in \mathcal{E}^\perp$. So we must show that

$$ad_X \mathcal{R}(dh) \subset \mathcal{R}(dh) \qquad \text{for all } X \in \mathcal{R}(dh)^\perp.$$

But this follows from

$$ad_X dh_i = \langle dh_i, X \rangle = 0. \qquad \square$$

Remarks

1. A converse is also true. A nonsingular distribution is integrable if its annihilator is locally spanned by exact 1-forms. (See Isidori, p.21.)
2. If $\mathcal{E}$ is integrable, then it defines a foliation via $\mathcal{E}^\perp$. If $\omega \in \mathcal{E}$ its restriction to a leaf is the zero 1-form.
3. If $X \in \mathcal{R}(dh)^\perp$, then X is tangent to leaves of the foliation. In particular, $\langle dh_i, X \rangle = 0$, so that h_i is constant on leaves.

Local Representation (Isidori, pp.25)

Lemma *Let $\mathcal{D}$ be a nonsingular involutive distribution of dimension k and assume $\mathcal{D}$ is invariant under a vector field X. Then at each $p \in M$, there exist coordinates (U, ξ) in* *which the vector field X can be represented as*

$$X(\xi) = \begin{bmatrix} X_1(\xi_1, \ldots, \xi_k, \xi_{k+1}, \ldots, \xi_n) \\ \vdots \\ X_k(\xi_1, \ldots, \xi_k, \xi_{k+1}, \ldots, \xi_n) \\ X_{k+1}(\xi_{k+1}, \ldots, \xi_n) \\ \vdots \\ X_n(\xi_{k+1}, \ldots, \xi_n) \end{bmatrix}$$

Proof By Frobenius' theorem, there exist coordinates (U, ξ) about $p \in M$ such that

$$\mathcal{D}(q) = \text{span}\{\frac{\partial}{\partial \xi_1}(q), \ldots, \frac{\partial}{\partial \xi_k}(q)\}, \quad q \in U.$$

Now $ad_X \mathcal{D} \subset \mathcal{D}$ implies

$$ad_X \frac{\partial}{\partial \xi_j} \in \text{ span } \{\frac{\partial}{\partial \xi_1}, \ldots, \frac{\partial}{\partial \xi_k}\}.$$

But

$$ad_X \frac{\partial}{\partial \xi_j} = -\Sigma_{i=1}^n (\frac{\partial X_i}{\partial \xi_j}) \frac{\partial}{\partial \xi_i}.$$

Hence must have

$$\frac{\partial X_i}{\partial \xi_j} = 0 \text{ for } i = k+1, \ldots, n; \ j = 1, \ldots, k. \qquad \square$$

Remark If $\mathcal{E}$ is a nonsingular involutive codistribution of dimension $(n-k)$, there exist local coordinates such that

$$\mathcal{E} = \text{ span } \{d\xi_{k+1}, \ldots, d\xi_n\}.$$

If also $\mathcal{E}$ is invariant under X, then we can choose $\xi_1, \ldots, \xi_k$ such that X has a representation of the above form.

These representations are useful in visualising controllability and observability.

Notation

$\langle ad_X | \mathcal{D} \rangle$ = smallest ad_X invariant distribution containing $\mathcal{D}$.

$\langle ad_X | \mathcal{E} \rangle$ = smallest ad_X invariant codistribution containing $\mathcal{E}$.

Refering to our control system Σ, we say that $\mathcal{D}$ (or $\mathcal{E}$) is *ad_f invariant* if $\mathcal{D}$ (or $\mathcal{E}$) is $ad_{f(\cdot,u)}$ invariant for all $u \in \Omega^m$.

$$\mathcal{R}(f) = \text{ distribution spanned by } \{f(\cdot,u) : u \in \Omega^m\}$$

$\langle ad_f | \mathcal{R}(f) \rangle$ is called the *controllability distribution.*

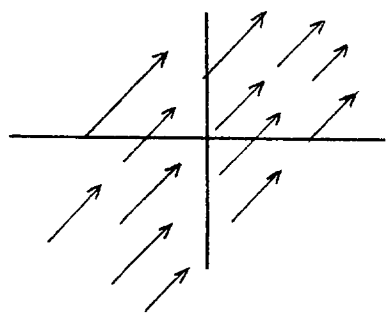
$\langle ad_f | \mathcal{R}(dh) \rangle$ is called the *observability codistribution.*

Examples

1. $M = \mathbb{R}^2$ $\qquad \Delta_1 = \text{ sp } \{\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\}$

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

These are smooth distributions.



However,

$$\Delta_1 \cap \Delta_2(x) = \begin{cases} \{0\} & \text{if } x_1 \neq 0 \\ \Delta_1(x) = \Delta_2(x) & \text{if } x_1 = 0 \end{cases}$$

is not a smooth distribution.

2. $M = \mathbb{R}$, $\qquad \Delta = \text{ sp } \{x\frac{\partial}{\partial x}\}$.

$$\dim \Delta(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x \neq 0 \end{cases}$$

So $x = 0$ is a singular point of Δ.

3. $M = \mathbb{R}$, $\quad \Delta = \text{sp } \{x\frac{\partial}{\partial x}\}$. Then

$$\Delta^{\perp}(x) = \begin{cases} \{0\} & \text{if } x \neq 0 \\ T_x^* M & \text{if } x = 0 \end{cases}$$

Then $\Delta^{\perp}$ is not a smooth codistribution.

4. $M = I\!R^2$, $\Delta = \{\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\}$. Then

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

is a smooth codistribution.

3. Controllability

3.1. Chow's Theorem

Consider a system Σ, defined in a neighborhood U, without drift that is $f(x) \equiv 0$. Define

$$\begin{aligned} \mathcal{A}(x^0, U) &= \{\gamma_u(s)x^0 : \ s \in I\!R, \ r \mapsto u(r) \\ &\quad \text{piecewise constant}, \ u(r) \in \Omega^m, \ \gamma_u(r)x^0 \in U \\ &\quad \text{for all } \ 0 \leq |r| \leq s.\} \end{aligned}$$

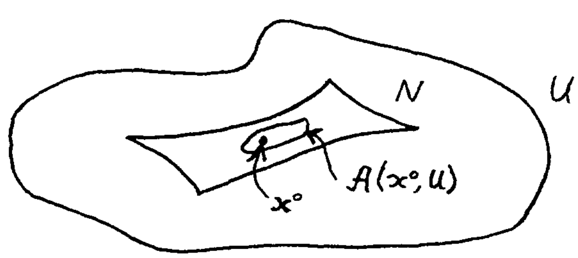
This is the set of states in U *accessible from* $x^0 \in U$. Define $\mathcal{F}$ to be the Lie algebra spanned by the vector fields $g_1, \ldots, g_m$ on U. Note that $\mathcal{F}$ is a distribution.

Assume that $\mathcal{F}$ is nonsingular. Let N be the corresponding (maximal in U) integral submanifold of $\mathcal{F}$ passing through x^0.

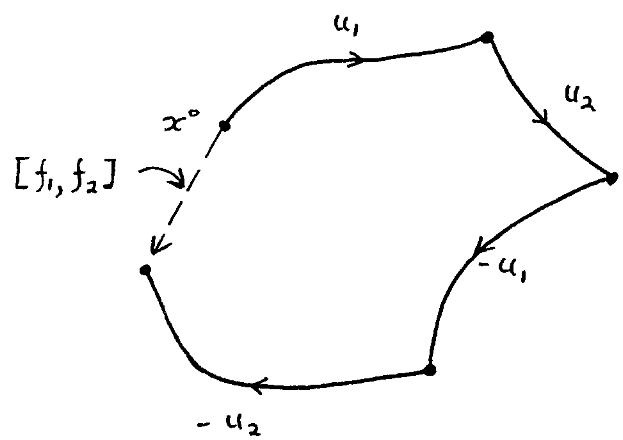
Theorem *Suppose* $\dim \mathcal{F}(x) = k \leq n$ *on* U. *Then* $\mathcal{A}(x^0, U) \subset N$ *contains a relatively open subset of* N.

If $\dim \mathcal{F}(x) = n$ on U, then $N = U$ and we have:

Corollary *(Classical Chow's theorem) If* $\dim \mathcal{F}(x) = n$ *on* U, *then* $\mathcal{A}(x^0, U)$ *contains an open subset of* U.



The Lie algebra $\mathcal{F}$ gives the possible directions in which the system can evolve. Let u^1 and u^2 be two controls, and $f^j = \Sigma_{i=1}^m g_i u_i^j$ be the corresponding vector fields. Now Σ can evolve in the directions f^1, f^2. The theorem says it can evolve in the direction $[f^1, f^2] \in \mathcal{F}$ also (Brockett [5]).



All such trajectories lie in the integral submanifold N.

Proof (Based on Krener [3], Isidori [2], p. 43) Choose $u^1 \in \Omega^m$ and set $f^1(x) = g(x)u^1$. We can assume $f^1(x_0) \neq 0$, otherwise choose another u^1. So $f^1(x) \neq 0$ for x near x^0, and there exists $\delta_1 > 0$ such that $\phi_1 : V_1 = (-\delta_1, \delta_1) \longrightarrow U$, where $\phi_1(s_1) = \gamma_1(s_1)x^0$, is an injective immersion. Thus $N_1 = \phi_1(V_1)$ is a 1-dimensional integral submanifold of $\mathcal{F}$ in U. Suppose we have constructed $N_{j-1} = \phi_{j-1}(V_{j-1})$ and $j \leq k$. Note $\dim N_{j-1} = j - 1$.
Claim: Given $x \in N_{j-1}$, we can choose $u^j \in \Omega^m$ such that

$$f^j(x) = g(x)u^j \notin T_x N_{j-1}.$$

Suppose not. Then $g(x)u \in T_x N_{j-1} \ \forall \ u \in \Omega^m$. This implies $\mathcal{F}(x) \subset T_x N_{j-1}$ for all $x \in N_{j-1}$. Define

$$\overline{\mathcal{F}}(x) = \begin{cases} T_x N_{j-1} & \text{if} \quad x \in N_{j-1} \\ \mathcal{F}(x), & \text{if} \quad x \in U \backslash N_{j-1} \end{cases}$$

Then by construction $g_1(x), \ldots, g_m(x) \in \overline{\mathcal{F}}(x)$ $, x \in U$, and $\overline{\mathcal{F}} \subset \mathcal{F}$. Let X_1, X_2 be vector fields on U with $X_i(x) \in \overline{\mathcal{F}}(x)$. Then X_1, X_2 are vector fields on N_{j-1}, so

$$[X_1, X_2](x) \in T_x N_{j-1} = \overline{\mathcal{F}}(x), \quad x \in N_{j-1}.$$

Also,

$$[X_1, X_2](x) \in \mathcal{F}(x) \text{ for } x \in U \backslash N_{j-1}.$$

Hence $\overline{\mathcal{F}}$ is involutive, and so $\overline{\mathcal{F}} = \mathcal{F}$.

But $k = \dim \mathcal{F}(x) \leq \dim \overline{\mathcal{F}}(x) = \dim T_x N_{j-1} = j - 1, \ \ x \in N_{j-1}$.

This is a contradiction since $j \leq k$, proving the claim.

By continuity, we can shrink V^{j-1}, N^{j-1} if necessary so that $f^j(x) \notin T_x N_{j-1}$ for all $x \in N_{j-1}$. Also, shrinking further if necessary, there exits $\delta_j > 0$ such that $\gamma_j(s_j)\ x \in U$ for $x \in N_{j-1},\ \ s_j \in (-\delta_j, \delta_j)$. Set $V_j = V_{j-1} \times (-\delta_j, \delta_j)$ and define

$$\phi_j(s_1, \ldots, s_j) = \gamma_j(s_j)\phi_{j-1}(s_1, \ldots, s_{j-1}).$$

By assumption, ϕ_{j-1} has rank $j-1$. It remains to verify that ϕ_j has rank j, from which it follows that N_j is a j-dimensional integral submanifold of $\mathcal{F}$ in U.

Now $\phi_{j*} = \gamma_{j*}\phi_{j-1*}$. So for $1 \leq i \leq j-1$,

$$\phi_{j*}(\frac{\partial}{\partial s_i})(s_1, \ldots, s_{j-1}, 0) = id\phi_{j-1*}(\frac{\partial}{\partial s_i})(s_1, \ldots, s_{j-1})$$

$$\phi_{j*}(\frac{\partial}{\partial s_j})(s_1, \ldots, s_{j-1}, 0) = \gamma_{j*}(0)id(\frac{\partial}{\partial s_j}) = f^j(\bar{x}),$$

where $\bar{x} = \phi_{j-1}(s_1, \ldots, s_{j-1}) \in N_{j-1}$. Thus if δ_j is sufficiently small,

$$\{\phi_{j*}(\frac{\partial}{\partial s_i})(s_1, \ldots, s_j)\}_{i=1}^{j}$$

are j linearly independent vectors at $\phi_j(s_1, \ldots, s_j) \in N_j$.

This process terminates at $j = k$, and therefore N_k is the desired relatively open subset of N. □

The situation is more complicated when Σ has drift, that is, $f(x) \not\equiv 0$. In particular, distinction must be made between forward and reverse time. Without drift, reversing time amounts to replacing u by $-u$.

Define

$$\begin{aligned} \mathcal{A}(x^0, U) \;&=\; \{\gamma_u(s)x^0 : \; s \geq 0, \;\; r \mapsto u(r) \;\text{ piecewise} \\ &\quad\;\; \text{constant}, \;\; u(r) \in \Omega^m, \; \gamma_u(r)x^0 \in U \\ &\quad\;\; \text{for all} \;\; 0 \leq r \leq s\} \end{aligned}$$

This is the set of states in U accessible from $x^0 \in U$ by going *forward* in time only. Let $\mathcal{F}$ be the Lie algebra generated by $f, g_1, \ldots, g_m$ on U.

Again, assume that $\mathcal{F}$ is nonsingular on U and let N be the corresponding (maximal in U) integral submanifold of $\mathcal{F}$, passing through x^0.

The following generalization of Chow's theorem is due to Krener, [3].

Theorem (*Krener*) *Suppose* $\dim \mathcal{F}(x) = k$ *on* U. *Then* $\mathcal{A}(x^0, u) \subset N$ ***contains a relatively open subset of*** N.

Proof Refer to [3]. □

This time asymmetry is reflected in the following. Define

$$\begin{aligned} \mathcal{C}(x^0, U) \;&=\; \{\gamma_u(s)x^0 : \; s \leq 0, \; r \mapsto u(r) \;\text{ piecewise constant} \\ &\quad\;\; u(r) \in \Omega^m, \; \gamma_u(r)x^0 \in U \text{ for all } \; s \leq r \leq 0\}. \end{aligned}$$

If $x^1 \in \mathcal{C}(x^0, U)$, then it is possible to steer the system from x^1 to x^0 by going forward in time, that is going backwards from x^0 to x^1. This is sometimes stated "x^1 is *controllable* to x^0", distinguishing between accessibility (reachability) and (this notion of) controllability. In general $\mathcal{A}(x^0, U) \neq \mathcal{C}(x^0, U)$, however they coincide when Σ has no drift, or when $f \in$ span $\{g_1, \ldots, g_m\}$.

3.2. Hermann-Krener Formulation

We summarize the ideas and results of Hermann and Krener [1], modified a little by the more recent ideas in Krener [4].

Controllability and Local Controllability

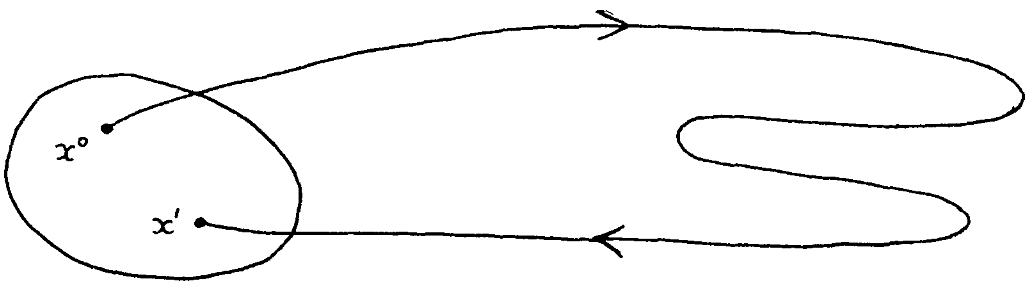
A state x^1 is *U-accessible* from x^0 if there exits a bounded measurable control $t \mapsto u(t)$, defined on some interval $[0, T]$, such that the corresponding trajectory $t \mapsto x(t), \;\; x(t) \in$

U, for all $t \in [0,T]$, $x(0) = x^0$, $x(T) = x^1$. We define the *accessible* sets by

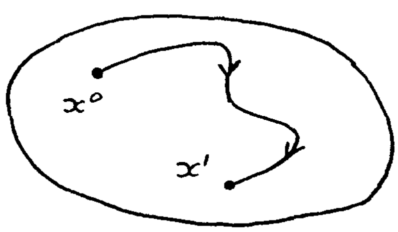
$$\begin{aligned} \mathcal{A}(x^0, U) &= \{x^1 \in U : x^1 \text{ is } U\text{-accessible from } x^0\}, \\ \mathcal{A}(x^0) &= \mathcal{A}(x^0, M). \end{aligned}$$

If $x^1 \in \mathcal{A}(x^0, U)$, in general it is *not* true that $x^0 \in \mathcal{A}(x^1, U)$. So accessibility is a reflexive, transitive but not symmetric relation.

We say Σ is *controllable at* x^0 if $\mathcal{A}(x^0) = M$, and *controllable* if $\mathcal{A}(x^0) = M$ for all $x^0 \in M$. However, it may be necessary to go a long way or for a long time to reach points near x^0:



A stronger notion of controllability would require that the trajectory stay near x^0:



Thus, we say that Σ is *locally controllable at* x^0 if for all neighborhoods U of x^0, $\mathcal{A}(x^0, U)$ is also a neighborhood of x^0. Σ is *locally controllable* if Σ is locally controllable at every $x^0 \in M$.

The above definitions consider the ability of Σ to steer from one state to another.

Accessibility Property

We noted above that U-accessibility is not an equivalence relation. According to Hermann-Krener, it is possible to define an equivalence relation on U containing all U-accessible *pairs*. They call this *weak U-accessibility*.

Write

$$\mathcal{WA}(x^0, U) = \{x^1 \in U: \; x^1, x^0 \text{ weakly } U\text{-accessible}\},$$

$$\mathcal{WA}(x^0) = W\mathcal{A}(x^0, M).$$

Here, $x^1 \in \mathcal{WA}(x^0, U)$ if and only if $x^0 \in \mathcal{WA}(x^1, U)$.

Analogously, one can define concepts of *(local) weak controllability.*

Another aspect of controllability is the ability of controls to influence all modes. Thus: We say that Σ has the *accessibility property* if $\mathcal{A}(x^0)$ has nonempty interior for all $x^0 \in M$. Σ has the *local accessibility property* if for every $x^0 \in M$, and every neighborhood U of x^0, $\mathcal{A}(x^0, U)$ has nonempty interior.

Theorem *If Σ is locally weakly controllable, then Σ has the local accessibility property.*

Proof Suppose Σ is locally weakly controllable. The argument is similar to that used in the proof of Chow's theorem, only the Claim is true for a different reason.

If $f(x,u) \in T_x N_{j-1}$ for all $u \in \Omega^m$, then $\gamma_u(t)x \in N_{j-1}$ for all t, for all $u \in \Omega^m$. This contradictors local weak controllability. □

Controllability Rank Condition

We say that Σ satisfies the *controllability rank condition at x^0* if in a neighborhood of x^0, $\dim\langle ad_f | \mathcal{R}(f)\rangle(x) = n$. If this holds for all $x^0 \in M$, we say that Σ satisfies the *controllability rank condition.*

Theorem *If Σ satisfies the controllability rank condition at $x^0 \in M$, then Σ has the local accessibility property at x^0.*

Proof By assumption, x^0 is a regular point for $\langle ad_f | \mathcal{R}(f)\rangle$. By Chow's theorem, $\mathcal{A}(x^0, U)$ contains an open subset of U. □

There is almost a converse:

Theorem *If Σ has the local accessibility property, then the controllability rank condition is satisfied generically.*

Proof Suppose there exits $U \subset M$ such that $\dim\langle ad_f | \mathcal{R}(f)\rangle(x) = k < n$ for $x \in U$. Let $x^0 \in U$ and U' be the corresponding maximal integral submanifold passing through x^0. Then $\mathcal{A}(x^0, U) \subset U'$, contradicting the local accessibility property. □

Remarks

1. The rank condition is an *algebraic test* for a form of controllability.
2. We have the following implications:

$$\begin{array}{ccc} \text{local controllability} & \Longrightarrow & \text{controllability} \\ \Downarrow & & \Downarrow \\ \text{local accessibility property} & \Longrightarrow & \text{accessibility property.} \end{array}$$

3.3. Local Representations (Isidori, p. 29, 40)

Write

$$\begin{aligned} P = & \ \langle ad_f | sp\{g_1, \dots, g_m\} \rangle, \\ R = & \ \langle ad_f | \mathcal{R}(f) \rangle. \end{aligned}$$

If x is a regular point of $P + sp\{f\}$, then

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

Let $r = \dim R$.

Theorem *Let P, $P + sp\{f\}$, R be nonsingular, and suppose $P \subset R, P \neq R$. Then, at each $p \in M$, there exist coordinates (U, ξ) in which Σ is represented by:*

$$\left\{ \begin{array}{rcl} \dot{\xi}_1 & = & f_1(\xi_1, \dots, \xi_n) + \Sigma_{i=1}^m g_{i1}(\xi_1, \dots, x_n)\, u_i \\ & \vdots & \\ & \vdots & \\ \dot{\xi}_{r-1} & = & f_{r-1}(\xi_1, \dots, \xi_n) + \Sigma_{i=1}^m g_{i,r-1}(\xi_1, \dots, \xi_n) u_i \\ \dot{\xi}_r & = & f_r(\xi_r, \dots, \xi_n) \\ \dot{\xi}_{r+1} & = & 0 \\ & \vdots & \\ \dot{\xi}_n & = & 0 \end{array} \right.$$

Proof According to the representation theorem in §2, there exits local coordinates such that

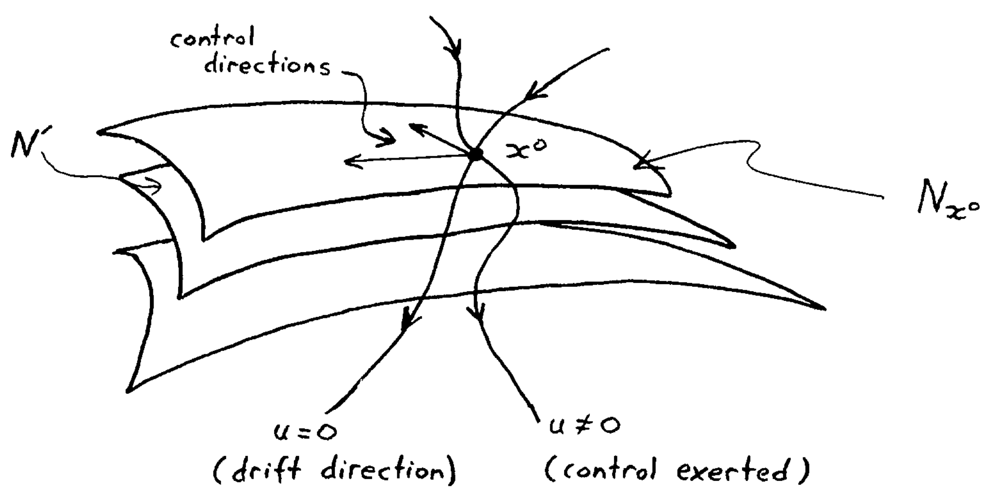
$$P = \{\frac{\partial}{\partial \xi_1}, \dots, \frac{\partial}{\partial \xi_{r-1}}\}, \ \ R = \{\frac{\partial}{\partial \xi_1}, \dots, \frac{\partial}{\partial \xi_{r-1}}, \frac{\partial}{\partial \xi_r}\}.$$

The first $(r-1)$ coordinates represent the (maximal in U) integral submanifolds of P, while the first r coordinates represent those of R. Since $f, g_1, \ldots, g_m \in R$, the components for $r+1, \ldots, n$ are zero. □

This gives us a geometrical picture of the behavior of Σ. All trajectories of Σ are contained in slices of the form

$$N_{x^0} = \{\xi_{r+1} = c_1, \ldots, \quad \xi_n = c_{n-r}\}, \qquad (r - \text{dimensional})$$

depending on the initial condition. The controls can affect the $(r-1)$ directions $\frac{\partial}{\partial \xi_1}, \ldots, \frac{\partial}{\partial \xi_{r-1}}$ only. The drift f causes Σ to move from one $(r-1)$ dimensional slice $N' = \{\xi_r = c_0, \ldots, \ \xi_n = c_{n-r}\}$ to another. In this sense ξ_r is analogous to time.



If $f_r = 0$, all trajectories are contained in an $(r-1)$ dimensional slice N', and there is no drift effect. This corresponds to $f(x) = 0$ or $f \in R$.

3.4. Controllable Subsystems

As the above geometrical description suggests, it may be possible to restrict Σ to a submanifold on which it is controllable. We mention one result in this direction.

Theorem *Suppose* $\dim\langle ad_f|\mathcal{R}(f)\rangle(x) = k \leq n$ *for all* $x \in M$. *Fix* $x^0 \in M$. *Then there exits a system* Σ' *defined on the maximal integral submanifold* N *of* $\langle ad_f|\mathcal{R}(f)\rangle$ *passing* *through x^0 which has the local accessibility property. Further, Σ' satisfies the controllability rank condition.*

3.5. Examples

1. Linear systems. $\dot{x} = Ax + Bu, \qquad M = I\!R^n.$

$$f(x) = Ax \qquad g^i(x) = B_i \qquad (B_i = \text{ith column of } B)$$

$$[Ax, B_i] = -AB_i, \qquad [Ax,[Ax,B_i]] = -A^2B_i, \ \ etc.$$

$$\mathcal{R}(f)(x) = \text{ span } \{Ax, B_1, \ldots, B_m\}.$$

$$\langle ad_f|\mathcal{R}(f)\rangle = \text{ span } \{B_1,\ldots,B_m,\ AB_1,\ldots,AB_m,\ldots,A^{n-1}B_1,\ldots,A^{n-1}B_m\}$$

$$\dim\langle ad_f|\mathcal{R}(f)\rangle = \text{ rank } \{B, AB, \ldots, A^{n-1}B\}$$

Thus the controllability rank condition is equivalent to the requirement that rank $\{B, AB, \ldots, A^{n-1}B\} = n$. In this case, this is equivalent to global controllability.

2. Bilinear systems $\dot{x} = Ax + \Sigma_{i=1}^m (N_i x) u_i$

$$f(x) = Ax, \qquad g^i(x) = N_i x$$

$$[f, g^i] = -[A, N_i] \qquad \text{(matrix commutator).}$$

$$\mathcal{R}(f)(x) = \text{ span } \{Ax, N_i x, \ldots, N_m x\}$$

$$\langle ad_f|\mathcal{R}(f)\rangle(x) = \text{ span } \{\mathcal{R}(f)(x),\ [A,N^i](x),\ [A,[A,N^i]](x), [N^i,N^s](x),\ \text{etc}\ldots\}$$

Note that $\mathcal{R}(f)(0) = \langle ad_f|\mathcal{R}(f)\rangle(0) = 0$, so bilinear systems are not controllable at 0. This example is discussed in detail in Brockett [8], where N^i are skew symmetric matrices. Then the trajectories evolve on the sphere $M : |x(t)| = |x(0)|$. Controllability is studied in terms of the matrix equations

$$\dot{X}(t) = (A + \Sigma_{i=1}^m\ u_i(t)N_i)X(t) \ \ , \quad X(0) = I.$$

It turns out that $X(t)$ is contained in a subgroup of $SO(n)$. If this subgroup acts transitively on S^{n-1}, then the bilinear systems is controllable on M.

Now the Lie algebra of $SO(n)$ is $O(n)$, the real skew symmetric matrices. If

$$\{A, N_1, \ldots, N_m\}_{LA} = O(n),$$

then the above mentioned subgroup is $SO(n)$, which acts transitively on S^{n-1}, and so the bilinear system is controllable on M.

3. $$M = I\!R^2 \qquad \dot{x}_1 = u \qquad u \in I\!R$$

$$\dot{x}_2 = x_1^2$$

$$x^0 = (x_1(0) = 0\ ,\ \ x_2(0) = 0).$$

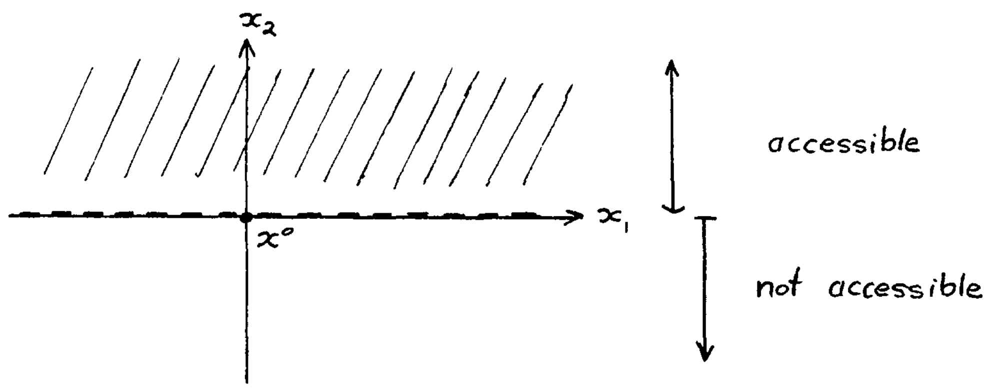
$$f(x) = (0, x_1^2)^T\ , \qquad g(x) = (1,0)^T.$$

$$\mathcal{R}(f)(x) = \begin{cases} I\!R\ , & x_1 = 0 \\ I\!R^2\ , & x_1 \neq 0 \end{cases}$$

$$\langle ad_f | \mathcal{R}(f) \rangle (x) = I\!R^2 \ \text{ for all } \ x.$$

Thus the controllability rank condition is satisfied everywhere. However,

$$\mathcal{A}(x^0) = \{(x_1, x_2) : \ x_2 > 0\} \cup \{x^0\}$$



Clearly $\mathcal{A}(x^0)$ has non-empty interior.

This example is due to Crouch and Byrnes [9]. They remark that this system is invariant under the Z_2 action on $I\!R^2$ defined by

$$(x_1, x_2) \longmapsto (-x_1, x_2).$$

4. Observability

4.1. Hermann – Krener Formulation

Once again we review the ideas and results in [1], influenced by the more recent work in [4].

Observability and Local Observability

Two states x^0, x^1 are *U-distinguishable* if there exits a bounded measurable control $t \mapsto u(t)$, defined on some interval $[0,T]$, such that the corresponding trajectories $t \mapsto x^i(t)$ satisfy $\dot{x}^i(0) = x^i$, $x^i(t) \in U$ for all $t \in [0,T]$, and $h(x^1(t)) \neq h(x^2(t))$ for some $t \in [0,T]$. Define indistinguishability sets

$$\begin{aligned} I(x^0,U) &= \{x^1 \in U : x^1 \text{ is not } U \text{ distinguishable from } x^0\}, \\ I(x^0) &= I(x^0,M). \end{aligned}$$

If $x^1 \in I(x^0,U)$, $x^2 \in I(x^1,U)$, then in general $x^2 \notin I(x^0,U)$. Thus indistinguishability defines a reflexive, symmetric but not transitive relation. However, when $U = M$, we get an equivalence relation.

We say that Σ is *observable at* x^0 if $I(x^0) = \{x^0\}$, and *observable* if $I(x^0) = \{x^0\}$ for all $x^0 \in M$.

Thus, for an observable system Σ, all the input-output maps S_{x^0}, $x^0 \in M$, are distinct.

A stronger concept is the following. Σ is *locally observable at* x^0 if for all neighborhood U of x^0, $I(x^0,U) = \{x^0\}$; and Σ is *locally observable* if this is true for all $x^0 \in M$.

Notice that this requires that states be distinguishable by local experiments.

Distinguishability Property

It may suffice to distinguish locally between points, either by local or global experiments. We shall discuss an appropriate equivalence relation in section 4.4.

We say that Σ has the *distinguishability property* if every $x^0 \in M$ has an open neighborhood U such that $I(x^0) \cap U = \{x^0\}$.

Σ has the *local distinguishability property* if every $x^0 \in M$ has an open neighborhood V such that for all open neighborhoods U of x^0, $U \subset V$, one has $I(x^0,U) \cap V = \{x^0\}$.

Note These concepts were called *weakly observable* and *locally weakly observable* in [1].

Observability Rank Condition

We say that Σ satisfies the *observability rank condition at* x^0 if in a neighborhood of x^0, $\dim\langle ad_f, \mathcal{R}(dh)\rangle(x) = n$. If this holds for all $x^0 \in M$, we say that Σ satisfies the observability rank condition.

Theorem *If Σ satisfies the observability rank condition at $x^0 \in M$, then Σ has the local distinguishability property at x^0.*

Proof First, let U be any neighborhood of x^0. Suppose that $\mathcal{I}(x^0, U) \neq \emptyset$ and let $x^1 \in \mathcal{I}(x^0, U)$. Then we claim that $\phi(x^0) = \phi(x^1)$ for every $\phi \in \mathcal{G}$, where $\mathcal{G} = \langle ad_f | \{h_1, \ldots, h_p\}\rangle$. (Note $d\mathcal{G} = \langle ad_f | \mathcal{R}(dh)\rangle$.)

To see this: let $u^1, \ldots, u^k \in \Omega^m$ and $s_1, \ldots, s_k \geq 0$ be sufficiently small. Since $x^1 \in \mathcal{I}(x^0, U)$,

$$h_i(\gamma_{u_k}(s_k) \circ \ldots \circ \gamma_{u_1}(s_1)x^0) = h_i(\gamma_{u_k}(s_k) \circ \ldots \circ \gamma_{u_1}(s_1)x^1).$$

Differentiate with respect to $s_k, \ldots, s_1$ and evaluate at 0 gives

$$ad_{f_1} \circ \ldots \circ ad_{f_k}(h_i)x^0 = ad_{f_1} \circ \ldots \circ ad_{f_k}(h_i)x^1.$$

But $\mathcal{G}$ is spanned by such functions. Hence the claim.

Since $\dim d\mathcal{G} = n$ around x^0, there exits $\phi_1, \ldots, \phi_n \in \mathcal{G}$ such that $d\phi_1, \ldots, d\phi_n$ are linearly independent. Define

$$\Phi : x \mapsto (\phi_1(x), \ldots, \phi_n(x))^T.$$

Now $D\Phi(x^0)$ is nonsingular, so by the inverse function theorem, Φ is locally injective, say in a neighborhood V. Then if $U \subset V$ is a neighborhood of x^0, the claim implies $\mathcal{I}(x^0, U) = \{x^0\}$. $\square$

Once again, we have a partial converse:

Theorem *If Σ has the local distinguishability property, then the observability rank condition is satisfied generically.*

Proof (Sketch). Suppose there exits $U \subset M$ such that $\dim\langle ad_f | \mathcal{R}(dh)\rangle(x) = k < n$ for $x \in U$. Let $x^0 \in U$, and consider Σ restricted to U, that is, $\Sigma\lceil_U$. Now $\langle ad_f | \mathcal{R}(dh)\rangle^\perp$ is a $n - k > 0$ dimensional integrable distribution. If x^1 is in the same leaf as x^0, then

$S_{x_0}\lceil_U = S_{x_1}\lceil_U$, so $\Sigma\lceil_U$ does not have the local distinguishability property. Hence neither does Σ. □

Remarks

1. The rank condition is an *algebraic* test for a (weak) form of observability.
2. We have:

$$\begin{array}{ccc} \text{local observability} & \Longrightarrow & \text{observability} \\ \Downarrow & & \Downarrow \\ \text{local distinguishability property} & \Longrightarrow & \text{distinguishability property} \end{array}$$

4.2. Local Representations (Isidori, p. 29, 50)

Write $\qquad Q = \langle ad_f | \mathcal{R}(dh) \rangle, \qquad s = \dim Q, \;\; d = n - s.$

Theorem *Let Q be nonsingular. Then, at each $p \in M$, there exist coordinates (U, ξ) in which Σ is represented by*

$$\begin{cases} \dot{\xi}_1 = f_1(\xi_1, \ldots, \xi_n) + \sum_{i=1}^m g_{i,1}(\xi_1, \ldots, x_n) u_i \\ \vdots \\ \dot{\xi}_s = f_s(\xi_1, \ldots, \xi_n) + \sum_{i=1}^m g_{i,s}(\xi_1, \ldots, \xi_n) u_i \\ \dot{\xi}_{s+1} = f_{s+1}(\xi_{s+1}, \ldots, \xi_n) + \sum_{i=1}^m g_{i,s+1}(\xi_{s+1}, \ldots, \xi_n) u_i \\ \vdots \\ \dot{\xi}_n = f_n(\xi_{s+1}, \ldots, \xi_n) + \sum_{i=1}^m g_{i,n}(\xi_{s+1}, \ldots, \xi_n) u_i \\ y_i = h_i(\xi_{s+1}, \ldots, \xi_n), \qquad i = 1, \ldots, p. \end{cases}$$

Proof This follows from the representation theorem in §2. □

Notice that the outputs depend only on $\xi_{s+1}, \ldots, \xi_n$. Leaves of $Q^\perp$ are $(n-s)$ dimensional, given by slices of the form

$$N = \{\xi_{s+1} = c_1, \ldots, \xi_n = c_{n-s}\}.$$

If $x^0, x^1 \in N$, then the last $(n-s)$ coordinates of the trajectories $t \mapsto x^0(t)$, $t \mapsto x^1(t)$, agree at time t, for all t, that is $\xi_i^0(t) = \xi_i^1(t)$, $i = s+1, \ldots, n$. Hence they produce the same output, and are indistinguishable. Σ moves slice to slice.

4.3. Duality

The above discussion parallels somewhat the discussion on controllability, and some duality is evident. A duality is well known for linear systems, and a corresponding notion for nonlinear systems is expressed in terms of the duality between vector fields and 1-forms. This idea was developed by Krener and Hermann [7].

In linear system theory, a pair (A, B) is controllable if

$$\mathcal{C} = \text{ span } \{B, AB, \ldots, A^{n-1}B\} = I\!R^n.$$

This corresponds to the controllability rank condition, and $\mathcal{C}$ can be identified with the controllability distribution. A pair (C, A) is observable if

$$\mathcal{O} = \text{ span } \{C, CA, \ldots, CA^{n-1}\} = I\!R^{n*}.$$

This is the observability rank condition, and $\mathcal{O}$ is the observability codistribution.
This is equivalent to requiring

$$\mathcal{O}^{\perp} = \bigcap_{i=0}^{n-1} \ker(CA^i) = \{0\},$$

which says that the annihilator of the observability codistribution is zero.

4.4. Observable Quotient Systems

Even if a system Σ on M is not observable, it may be possible to define a *"quotient system"* Σ' on M/I which is observable, for an appropriate equivalence relation I. The equivalence classes ought to be the leaves of the above mentioned foliation.
In this section we shall simply state two results.

$x^1 I x^0$ if and only if $x^1 \in \mathcal{I}(x^0)$, or equivalently, $x^0 \in \mathcal{I}(x^1)$. Then I is an equivalence relation on M. In fact, I is closed (continuity of ODEs on initial conditions). Then M/I is Hausdorff. In general, I need not be regular. (Recall that I is regular if M/I admits a C^∞ structure for which $\pi : M \to M/I$ is a submersion.).

<u>Theorem</u> *(Sussman) Let Σ be symmetric (that is, for all $u \in \Omega^m$ there exits $v \in \Omega^m$ such that $f(x, u) = -f(x, v)$ for all $x \in M$). If (Σ, x^0) has the local accessibility condition,*

then I is closed and regular. Also, there exits a system Σ' defined on $M' = M/I$ such that $(\Sigma', I(x^0))$ is observable, has the local accessibility property, and realizes the same input-output map.

We say that x^0, x^1 are *strongly indistinguishable*, written $x^0 SI x^1$, if there exits a continuous $\alpha : [0,1] \to M$ such that $\alpha(0) = x^0$, $\alpha(1) = x^1$, and $x^0 I \alpha(s)$, for all $s \in [0,1]$. Then SI is an equivalence relation and $x^0 SI x^1$ implies $x^0 I x^1$. If Σ has the local distinguishability property at x^0, then $SI(x^0) = \{x^1 : x^1 SI x^0\} = \{x^0\}$.

Theorem *Suppose* $\dim\langle ad_f | \mathcal{R}(dh)\rangle(x) = k \leq n$, *for all* $x \in M$. *Then:*

(i) *SI is a regular equivalence relation;*

(ii) *there exits a system Σ' on $M' = M/SI$ which has the local distinguishability property;*

(iii) (Σ, x^0) and $(\Sigma', SI(x^0))$ *release the same input-output map, for all* $x^0 \in M$;

(iv) *if Σ is (locally) controllable, then so is Σ';*

(v) *if Σ has the (local) accessibility property, then so does Σ';*

(vi) *if Σ satisfies the controllability rank condition, then so does Σ', and moreover, M' is Hausdorff.*

4.5. Examples

1.
$$\dot{x} = f(x), \quad x \in \mathbb{R}^n$$
$$y = h(x), \quad y \in \mathbb{R}.$$

$$f(x) = \textstyle\sum_{i=1}^{n} f^i(x)\frac{\partial}{\partial x_i} \qquad dh(x) = \sum_{i=1}^{n} \frac{\partial h}{\partial x_i} dx_i$$
$$L_f h(x) = \textstyle\sum_{i=1}^{n} f^i(x)\frac{\partial h}{\partial x_i}(x) \quad L_f dh = dL_f h$$

$$\mathcal{G} = \text{span}\{h, L_f h, L_f^2 h, \text{ etc}\} = \mathcal{R}(h)$$

$$\langle ad_f | \mathcal{R}(dh)\rangle = d\mathcal{G} = \text{span}\,\{dh, L_f dh, L_f^2 dh, \text{ etc}\}$$

2.Linear Systems
$$\dot{x} = Ax$$
$$y = Cx$$

$$f(x) = Ax \qquad f^i(x) = \sum_{j=1}^{n} a_{ij}\, x_j \qquad h(x) = Cx = \sum_{j=1}^{n} c_j\, x_j$$

$$
\begin{aligned}
L_f h(x) &= \sum_{i=1}^{n}(\sum_{j=1}^{n} a_{ij}\, x_j)C_i = \sum_{i=1}^{m}(A^T C^T)_i x^i \\
L_f^k h(x) &= \sum_{i=1}^{n}((A^T)^k C^T)_i x_i \\
L_f^k dh &= dL_f h = \sum_{i=1}^{n}((A^T)^k C_i^T dx_i \\
\langle ad_f | \mathcal{R}(dh)\rangle &= \text{ span } \{C, CA, \ldots, CA^{n-1}\}
\end{aligned}
$$

Thus the observability rank condition is equivalent to the requirement that rank $\{C, CA, \ldots, CA^{n-1}\} = n$. Here, this is equivalent to global observability.

The Lie differentiation is just differentiating the output $(n-1)$ times:

$$
x(t) = e^{At}x_0, \qquad\qquad y(t) = Ce^{At}x_0.
$$

$$
\begin{cases}
y(0) &= Cx_0 \\
\dot{y}(0) &= CAx_0 \\
 & \ddots \\
y^{(n-1)}(0) &= CA^{n-1}x_0
\end{cases}
$$

This system of n equations can be solved uniquely for x_0 in terms of $(y(0), \ldots, y^{(n-1)}(0))$ if rank $\{C, CA, \ldots, CA^{n-1}\} = n$.

3. $M = I\!R$, $\qquad \dot{x} = u$

$$
y = \sin x
$$

$$
f(x) = 0, \qquad g(x) = 1, \qquad h(x) = \sin x
$$

$$
\mathcal{R}(f) = \text{ sp } \{\frac{\partial}{\partial x}\}, \ \langle ad_f | \mathcal{R}(f)\rangle = \text{ span } \{\frac{\partial}{\partial x}\} \simeq I\!R
$$

$$
dh = (\cos x)dx \qquad\qquad L_f dh = (\sin x)dx
$$

$$
\langle ad_f | \mathcal{R}(dh)\rangle(x) = \text{ span}\{(\cos x)dx, (\sin x)dx\} \simeq I\!R
$$

Therefore both the controllability and observability rank conditions are satisfied. The system is in fact controllable, but not observable.

Let $x_0 = 0$. Then $\mathcal{I}(x^0) = \{2k\pi, k \in Z\} \simeq Z$.

However, on $M' = S^1 = I\!R/Z$, the system is observable. (It is again controllable, and we have "minimal realization".)

References

[1] R. Hermann and A.J. Krener, *"Nonlinear Controllability and Observability"*, IEEE Trans. Automatic Control, Vol AC-2, No. 5, Oct. 1977, pp. 728-740.

[2] A. Isidori, *"Nonlinear Control Systems: An Introduction"*, Springer Verlag, 1985.

[3] A.J. Krener, *"A Generalization of Chow's Theorem and the Bang-Bang Theorem to Nonlinear Control Problems"*, SIAM J. Control, Vol. 12, No. 1, Feb. 1974, pp. 43-52.

[4] A.J. Krener, "$(Ad_{f,g}), (ad_{f,g})$ *and Locally* $(ad_{f,g})$ *Invariant and Controllability Distributions*", SIAM J. Control and Optimization, Vol. 23, No. 4, July 1985, pp. 523-549.

[5] R. Brockett, *"Nonlinear Systems and Differential Geometry"*, Proc. IEEE, Vol. 64, No. 1, Jan 1976, pp. 61 - 72.

[6] W.M. Boothby, *"In Introduction to Differentiable Manifolds and Riemannian Geometry"*, Academic Press, 1975.

[7] A.J. Krener and R. Hermann *"The Differential Geometric Duality Between Controllability and Observability for Nonlinear Systems"*.

[8] R.W. Brockett, *"Lie Theory and Control System Defined on Spheres"*, SIAM J. Appl. Math, Vol. 25, No. 2, Sept. 1973.

[9] P.E. Crouch and C.I. Byrnes, *"Local Assessibility, Local Reachability, and Representations of Compact Groups,"* Math. Sys. Theory, 19 (1986), 43-65.