MATHEMATICAL CONTROL THEORY
BANACH CENTER PUBLICATIONS, VOLUME 14
PWN – POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

LIE BRACKETS AND REAL ANALYTICITY IN CONTROL THEORY

HÉCTOR J. SUSSMANN

Department of Mathematics, Rutgers University, New Brunswick, N. J., U.S.A.

§ 1. Introduction

The purpose of these lectures is to show how two mathematical tools not traditionally used by control theorists have become important in recent years. The tools are

(1) Real analyticity, and

(2) The Lie bracket of vector fields.

It is not our goal here to provide a comprehensive survey but only to show, by means of several examples, why these tools are useful. We discuss real analyticity and the Lie bracket together because many of the applications actually involve both.

In § 2 we describe the class of control systems to be considered. In § 3 we take one problem—that of generalizing to the nonlinear case the familiar controllability criteria for linear systems—and we use it as motivation for introducing the Lie bracket and for considering real analyticity. Then we state (without proof) the important theorem of Nagano on Lie algebras of real analytic vector fields, as well as some facts on orbits, and we give Krener's proof of the "positive form of Chow's theorem". In § 4 we show how, in a very precise sense, the set of all Lie bracket relations between the vector fields of an analytic control system at a point p determines the system in a neighborhood of p, so that one can think of this set as a kind of Taylor series for the system. In § 5 we apply the results of § 4 to give a simple proof of a necessary and sufficient condition for a system to be locally equivalent to a linear system. In § 6 we state a bang-bang theorem where the system is required to satisfy a certain Lie-theoretic condition. In § 7 we develop a formalism based on exponential Lie series and we sketch a proof based on this formalism of a local controllability theorem conjectured by Hermes. Finally, in § 8 we mention some other applications.

§ 2. General definitions

A *general control system* is an ordinary differential equation of the form

$$\dot{x} = f(x, u), \quad x \in M, \ u \in U. \tag{1}$$

Here M, the "state space", could be taken to be a Euclidean space $\mathbf{R}^n$ or, more generally, an open subset of some $\mathbf{R}^n$. However, it turns out that we get a nicer theory if we allow M to be a more general object. Specifically, we will assume

(I) *M is a C^∞ manifold.*

Remark 1. We include in the definition of "manifold" the requirement that M be finite-dimensional, Hausdorff and a countable union of compact sets. ■

When M is an open subset of $\mathbf{R}^n$, then f should be required to be a mapping which to each $x \in M$, $u \in U$, assigns a vector $f(x, u) \in \mathbf{R}^n$. In the more general case where M is a manifold we have to assume that $f(x, u)$ is a tangent vector to M at x. Moreover, we will assume that the dependence on x is smooth, i.e.,

(II) *For each $u \in U$, the map $x \to f(x, u)$ is a C^∞ vector field on M.*

In order to talk about trajectories of (1) for general controls, one has to assume something about the control space U, and about the dependence of f jointly on x and u. Then one has to make assumptions on the class of admissible controls. This gives rise to technical problems that are easy to settle but uninteresting. So we shall limit ourselves to a particular type of situation, namely, the case when the control enters linearly in (1). That is, we will assume

(III) *U is a convex subset of $\mathbf{R}^m$ and $f(x, u)$ has the form*

$$f(x, u) = f_0(x) + \sum_{i=1}^{m} u_i f_i(x). \tag{2}$$

But we emphasize that:

Remark 2. Much of what will be done here under assumption III generalizes to the case when linearity in u is not assumed. ■

The following assumption can be made with no loss of generality:

(IV) *U has a nonempty interior in $\mathbf{R}^m$.*

(Indeed, if (IV) were not satisfied, we can always redefine our system so that (IV) holds.)

An *admissible control* is a bounded, measurable U-valued function defined on some interval $I \subseteq [0, \infty)$ such that $0 \in I$. A *trajectory* for an admissible control $u(\cdot)\colon I \to U$ is an absolutely continuous curve $x(\cdot)\colon I \to M$ such that

$$\dot{x}(t) = f\big(x(t), u(t)\big) \tag{3}$$

for almost all $t \in I$.

If $x_0 \in M$, $x_1 \in M$, $T \geqslant 0$ are such that there is a trajectory $x(\cdot)\colon [0, T] \to M$ corresponding to some admissible $u(\cdot)\colon [0, T] \to U$ for which $x(0) = x_0$, $x(T) = x_1$, we say that x_1 is *reachable from x_0 in T units of time*. The set of all points x_1 that are reachable from x_0 in T units of time is the *time T reachable set from x_0* and we denote it $\mathrm{Reach}_T(x_0)$. Also we let

$$\mathrm{Reach}(x_0) = \bigcup_{T \geqslant 0} \mathrm{Reach}_T(x_0). \tag{4}$$

§ 3. Integral manifolds and Nagano's theorem

In this section we state a very natural question (referred to as (Q1)) about nonlinear systems and show how, in order to answer it, one is naturally lead to the study of Lie brackets, and to paying special attention to the real analytic case. We state an important theorem due to Nagano (cf. [5]), as well as the so-called "positive form of Chow's theorem". These two results together enable us to answer the original question. Moreover, we derive a nonlinear analogue of Kalman's controllability criterion and we show how the property that the ring of germs of analytic functions is Noetherian plays an important role. Thanks to it, the "nonlinear controllability criterion" becomes a test which gives a definitive yes or no answer in a finite number of steps. We emphasize that question (Q1), as stated, does not involve any mention of Lie brackets or real analyticity. In our view the true criterion to decide whether or not a mathematical theory A has significant applications to an area of research B is whether, using A, one can answer questions about *B that make sense before A is brought in*, but cannot be answered without A. The discussion that follows shows that, in our particular case, the criterion is met.

(Q1) *What is the analogue, for the systems considered here, of the basic results of controllability theory for linear systems?*

First, let us clarify what Q1 means. For a linear system

$$\dot{x} = Ax + \sum_{i=1}^{m} u_i b_i \quad (M = \boldsymbol{R}^n,\ U = \boldsymbol{R}^m) \tag{5}$$

there is a well-defined "controllability" concept, which can be defined in many equivalent ways. For instance, we may call (5) controllable if either

(Ca) $\operatorname{Reach}(x_0) = \boldsymbol{R}^n$ for all x_0,

or

(Cb) $\operatorname{Reach}(0) = \boldsymbol{R}^n$,

or

(Cc) $\operatorname{Reach}(x_0)$ has a nonempty interior in $\boldsymbol{R}^n$ for all x_0,

or

(Cd) There exists no linear subspace $S \subseteq \boldsymbol{R}^n$ such that $S \neq \boldsymbol{R}^n$ but $\operatorname{Reach}(x_0) \subseteq S$ for all $x_0 \subseteq S$,

or

(Ce) There exists no linear subspace $S \subseteq \boldsymbol{R}^n$ such that $S \neq \boldsymbol{R}^n$ and that, whenever $x_0 \in M$, $x_1 \in M$, $x_1 \in \operatorname{Reach}(x_0)$, then $x_0 \in S$ iff $x_1 \in S$.

It is a trivial matter to prove that (Ca), (Cb), (Cc), (Cd) and (Ce) are equivalent. Moreover, any system (5), even if it does not satisfy these conditions, can always be "reduced" to one that does. (Just let S be the linear span of all the vectors $A^k b_i$ for all k, i. Then (5) can be restricted to S, and the restricted system satisfies (Ca), (Cb), (Cc), (Cd), (Ce).)

For our more general systems one sees right away that, in general, it is not true that if $\operatorname{Reach}(x_0)$ has a nonempty interior in M for all $x_0 \in M$, then $\operatorname{Reach}(x_0) = M$ for all x_0. (For example, let $M = \boldsymbol{R}$ and take the system $\dot{x} = 1$.) So the nonlinear analogues of (Ca), (Cb), (Cc), (Cd), (Ce) are no longer equivalent.

Let us call a subset S of M *forward invariant* for the system (1) if $\operatorname{Reach}(x_0) \subseteq S$ for all $x_0 \in S$. Let us call S *bi-invariant* if, whenever $x_0 \in M$, $x_0 \in M$, $x_1 \in \operatorname{Reach}(x_0)$, then $x_0 \in S$ iff $x_1 \in S$. With this terminology the definition of controllability for linear systems can be restated by saying that (5) is controllable iff there is no proper bi-invariant subspace of M or, equivalently, if there is no proper forward invariant subspace. Moreover, since $\operatorname{Reach}(0)$ is always a linear subspace, an equivalent condition for controllability is simply that there is no proper *forward invariant set* or that there is no proper *bi-invariant set*. When a linear system is not controllable, then there is a unique *minimal forward invariant set* (MFIS) through 0. This set turns out to be a linear subspace and a minimal bi-invariant set (MBIS). Since it is a linear subspace, the restriction of (5) to this set turns out to be another system in the class we started with, i.e., a linear system.

For nonlinear systems it is therefore natural to try to answer (Q1) by first asking:

(Q1a) *Given a system* (1) *and an* $x_0 \in M$, *is there always an* MFIS *containing* x_0?

(Q1b) *Is there always an* MBIS *through* x_0?

(Q1c) *If the answer to* (Q1a) *or* (Q1b) *is "yes", is the* MFIS (*or the* MBIS) *a subset S of M such that the restriction of* (1) *to S is well defined and is another system in the class we started with?*

The answers to both (Q1a) and (Q1b) are clearly "yes". (For (Q1a) take $S = \mathrm{Reach}(x_0)$. For (Q1b) let S be the intersection of all bi-invariant sets through x_0.) In order to answer (Q1c) we have to find out whether the MFIS or the MBIS through x_0 is a manifold. The answer is obviously "no" for the MFIS (example: $\dot{x} = 1$ once again).

It turns out that *for real analytic systems* the MBIS through any $x_0 \in M$ is always an analytic submanifold of M, and that the restriction of (1) to this submanifold is another analytic system. Moreover, the MBIS through x_0 can be characterized as an integral manifold of the family of vector fields consisting of the f_i and their Lie brackets of all orders. *This is the first reason why real analyticity and Lie brackets are important.*

Remark 3. Incidentally, this also shows one reason why we have to allow our state spaces to be general manifolds. Had we started with open subsets of Euclidean spaces, we would have run into trouble with systems such as $\dot{x} = uy$, $\dot{y} = -ux$, $(x, y) \in \boldsymbol{R}^2$. Here the MBIS through $(1, 0)$ is a circle which is not a state space of the kind we were willing to allow. ∎

We now make the preceding considerations precise. An *analytic system* is a system (1) which satisfies our hypotheses (I), (II), (III), (IV) and for which, in addition:

(AS1) M is a real analytic manifold, and

(AS2) The vector fields $f_0, \ldots, f_m$ are real analytic.

If M is a C^∞ manifold, let $V(M)$ denote the set of all C^∞ vector fields on M. If $f, g \in V(M)$, the *Lie bracket* of f and g is another vector field, denoted by $[f, g]$, which can be defined in at least two ways. Since both definitions are of interest to us, we will give them both. First, let us use the notation $\Phi^f(t)$ for the flow of f (that is, if $x_0 \in M$, then $t \to \Phi^f(t)(x_0)$ is the integral curve of f which goes through x_0 when $t = 0$). Recall that $V(M)$ can be thought of as the set $D_1(M)$ of all first order differential operators F: $C^\infty(M) \to C^\infty(M)$, where $C^\infty(M) = \{\varphi: M \to \boldsymbol{R}, \varphi \in C^\infty\}$. (That is, $F \in V(M)$ iff F is a map $C^\infty(\boldsymbol{R}) \to C^\infty(\boldsymbol{R})$ such that $F(\alpha\varphi + \beta\psi) = \alpha F\varphi + \beta F\psi$, $F(\varphi\cdot\psi) = \varphi\cdot F\psi + \psi\cdot F\varphi$ for all $\varphi, \psi \in C^\infty(M)$, $\alpha, \beta \in \boldsymbol{R}$). The identification $V(M) \sim D_1(M)$ is the map which to an $f \in V(M)$ assigns the $F \in D_1(M)$ given by

$$(6) \qquad (F\varphi)(x) = \lim_{t\to 0} \frac{1}{t}\left[\varphi\left(\Phi^f(t)(x)\right) - \varphi(x)\right].$$

The *Lie bracket* is defined as follows:

DEFINITION 1. If $F \in D_1(M)$, $G \in D_1(M)$, we let $[F, G]: C^\infty(M) \to C^\infty(M)$ be the map $FG - GF$.

DEFINITION 2. If $f \in V(M)$, $g \in V(M)$, $x \in M$, we let $[f, g](x)$ be the tangent vector at $t = 0$ to the curve

$$t \to \Phi^g(-\sqrt{t})\Phi^f(-\sqrt{t})\Phi^g(\sqrt{t})\Phi^f(\sqrt{t})(x). \tag{7}$$

Remark 4. The equivalence of these two definitions is a standard fact which, in any case, will be proved below (cf. § 7). ■

A *Lie algebra of vector fields on* M is a subset A of $V(M)$ which is a linear space (over $\mathbf{R}$) and satisfies $f \in A$, $g \in A \Rightarrow [f, g] \in A$. If $A \in V(M)$, there is a smallest Lie algebra L of vector fields that contains A. It is called the *Lie algebra generated by* A and we use $\mathrm{Lie}(A)$ to denote it.

Now let us return to the study of the MBIS S through a point $x_0 \in M$ for a system (1). If we wish to prove that S is a submanifold of M, it is reasonable to try to determine its tangent space $T_x(S)$ at every $x \in S$. Since S is bi-invariant, the curve $t \to \Phi^g(t)(x)$ must be contained in S if g is any vector field of the form

$$g = f_0 + \sum_{i=1}^{m} a_i f_i \tag{8}$$

for which $(a_1, \ldots, a_m) \in U$. Therefore every such g must be tangent to S. In view of assumption (IV), it follows that every linear combination of $f_0, \ldots, f_m$ is tangent to M. From this one can conclude that every $g \in \mathrm{Lie}(\{f_0, \ldots, f_m\})$ must be tangent to S. (Reason: if f, g are tangent to S then the curve (7) lies in S if $x \in S$. So $[f, g](x) \in T_x S$.) So $T_x S$ must contain the space $\mathrm{Lie}(\{f_0, \ldots, f_m\})(x)$, where if $A \in V(M)$, $x \in M$, we let

$$A(x) = \{g(x) \colon g \in A\}. \tag{9}$$

An *integral manifold* (IM) of a set $A \subseteq V(M)$ is a connected submanifold S of M such that

$$T_x S = \text{linear span}\big(A(x)\big) \quad \text{for all } x \in S. \tag{10}$$

A *maximal integral manifold* (MIM) of A is an IM S of A such that whenever S' is an IM of A such that $S \cap S' \neq \emptyset$, it follows that $S' \subseteq S$.

NAGANO'S THEOREM. *Let M be an analytic manifold and let $L \subseteq V(M)$ be a Lie algebra of analytic vector fields. Then for every $x \in M$ there exists a MIM of L through x.* ■

Remark 5. The MIM of the preceding statement is obviously unique. ■

Returning to our system (1), it is reasonable to expect, in view of Nagano's theorem plus our earlier remarks, that the MIM of $\mathrm{Lie}(\{f_0, \ldots, f_m\})$ through x_0 will turn out to be the same as the MBIS through x_0 if (1) is analytic. This can actually be proved as follows: the set Σ of MIM's of $\mathrm{Lie}(\{f_0, \ldots, f_m\})$ is a partition of M. Every f_i is tangent to every

$S \in \Sigma$. Therefore, if $x(\cdot)\colon I \to M$ is a trajectory of (1) and if $t_0 \in I$, $x(t_0) \in S \in \Sigma$, it follows that $x(t) \in S$ for $t \in (t_0 - \varepsilon, t_0 + \varepsilon) \cap I$ for some $\varepsilon > 0$. So $\{t\colon x(t) \in S\}$ is relatively open in I for each $S \in \Sigma$. Therefore every trajectory of (1) is entirely contained in one $S \in \Sigma$. So all the members of Σ are bi-invariant for (1). To prove that they are MBIS's, we will use another important result. First we need two definitions. We say that a system (1) has the *accessibility property* (AP) from x_0 if $\mathrm{Reach}(x_0)$ has a nonempty interior in M, and that (1) satisfies the *rank condition* (RC) at x_0 if

$$\dim \mathrm{Lie}(\{f_0, \ldots, f_m\})(x_0) = \dim M. \tag{11}$$

POSITIVE FORM OF CHOW'S THEOREM (PFCT). *Let* (1) *be analytic and let $x_0 \in M$. Then* (1) *has the* AP *from x_0 iff* (1) *satisfies the* RC *at x_0.*

Proof (cf. Krener [4]). Suppose that (11) holds. Let $W \subseteq M$ be open such that $x_0 \in W$ and that the RC holds at every $x \in W$. Let Q be a submanifold of M of the largest possible dimension such that $Q \subseteq \mathrm{Reach}(x_0) \cap W$. If g is any vector field of the form (8) with $(a_1, \ldots, a_m) \in U$, and if $g(x) \notin T_x Q$ for some $x \in Q$, then $(t, y) \to \Phi^g(t)(y)$ has rank $(\dim Q) + 1$ at $(0, x)$, and it maps $(0, \varepsilon) \times Z$ diffeomorphically onto a submanifold Q' such that $Q' \subseteq \mathrm{Reach}(x_0) \cap W$, if $\varepsilon > 0$ is small enough and Z is a sufficiently small neighborhood of x. Then $\dim Q' > \dim Q$, contradicting the maximality of $\dim Q$. So g is tangent to Q. Since this is true for every g of the form (8), with $(a_1, \ldots, a_m) \in U$, we conclude as before (cf. the argument between formulas (8) and (9)) that

$$\mathrm{Lie}(\{f_0, \ldots, f_m\})(x) \subseteq T_x Q$$

for all $x \in Q$. Since the RC holds at every $x \in Q$, we conclude that $\dim Q = \dim M$. So Q is open and therefore (1) has the AP from x_0.

For the converse, if the AP from x_0 holds, we already know that there is a MIM of $\mathrm{Lie}(\{f_0, \ldots, f_m\})$ through x_0 and that $\mathrm{Reach}(x_0)$ is contained in this MIM. The AP from x_0 then implies that the MIM has dimension $\dim M$ and so the RC holds at x_0. This concludes the proof of the PFCT. ∎

We now return to the proof that for an analytic system (1) the MIM's of $\mathrm{Lie}(\{f_0, \ldots, f_m\})$ are MBIS's. We already know that they are bi-invariant sets, so all we need is to prove minimality. Let S be a MIM for $\mathrm{Lie}(\{f_0, \ldots, f_m\})$. The system (1) can be restricted to S, and the restricted system has the RC at every point of S, and therefore it has the AP from every $x \in S$. Moreover, S is connected. So our conclusion follows from:

LEMMA 1. *If a system* (1) *has the* AP *from every $x \in M$ and if M is connected, then M is the only bi-invariant set for* (1).

Proof. Let $S \subseteq M$ be a MBIS for (1). If $x \in S$, let $W \subseteq \operatorname{Reach}(x)$ be open, $W \neq \emptyset$. Then $W \subseteq S$. Take some $u(\cdot)$: $[0, T] \to U$ that steers x to some $y \in W$. For $z \in W$, let γ_z be the trajectory for $u(\cdot)$ such that $\gamma_z(T) = z$. Then γ_z is defined on $[0, T]$ for z in some open W' such that $y \in W' \subseteq W$. The set $\{\gamma_z(0): z \in W'\}$ is a neighborhood of x and by the bi-invariance of S it is a subset of S. So $x \in \operatorname{Int} S$. So S is open. So the MBIS's of (1) form a partition of M whose members are open. Since M is connected, M is itself a MBIS. The proof of Lemma 1 is complete. ■

As indicated earlier, we have now proved:

THEOREM A. *For an analytic system* (1), *if* $x_0 \in M$, *then there exists an analytic submanifold* S *such that* $x_0 \in S$, *and that* S *is a minimal bi-invariant set for* (1). ■

This result is our nonlinear analogue of the linear controllability result. Call a system (1) *controllable* if M is itself a MBIS for (1). Then Theorem A asserts that (if (1) is analytic) for every $x_0 \in M$ there is a submanifold S (the "controllable piece through x_0") such that (1) can be restricted to S, and that this restriction is controllable. Moreover, controllability can be characterized quite easily, as follows:

THEOREM B. *Consider an analytic system* (1). *Then the following are equivalent*:

(a) M *is a* MBIS.

(b) M *is connected and the* RC *holds at every* $x \in M$.

Proof. That (b) ⇒ (a) follows from the PFCT and Lemma 1. Assume (a) holds. Since every connected component is bi-invariant, M must be connected. Since M is a MBIS, then M is an integral manifold of $\operatorname{Lie}(\{f_0, \ldots, f_m\})$ by Theorem A. Therefore the RC holds at every $x \in M$. ■

Remark 6. Theorem B is the nonlinear analogue of the Kalman controllability criterion. It is easy to verify that for linear systems the RC yields exactly the Kalman criterion. ■

Remark 7. Simple examples show that both Nagano's theorem and the PFCT, as stated here, fail if the vector fields under consideration are only C^∞ but not analytic. Theorem B is also false for general C^∞ systems. (Example: consider the system $\dot{x} = 1$, $\dot{y} = u\varphi(x)$, where φ: $\boldsymbol{R} \to \boldsymbol{R}$ is a C^∞ function such that $\varphi(x) = 0$ for $x \leqslant 0$, $\varphi(x) > 0$ for $x > 0$. Then $\boldsymbol{R}^2$ is the only MBIS but the RC only holds on $\{(x, y): x > 0\}$.) The analogue of Theorem A is true (cf. Sussmann [6]). ■

Remark 8. An important feature of Kalman's controllability criterion is that it involves a finite number of steps. The system (5) is controllable iff the vectors $A^k b_i$, $i = 1, \ldots, m$, $k = 0, \ldots, n-1$, span $\boldsymbol{R}^n$. (Thanks to the Cayley–Hamilton theorem.)

For the nonlinear case we can ask whether, if we compute successive brackets of the f_i at a particular x_0, a point is ever reached when we know for sure whether the RC holds at x_0. The answer is "yes". To see this, let $\mathscr{A}_{x_0}$ denote the ring of germs of analytic functions at x_0, and let $\mathscr{V}_{x_0}$ denote the set of germs at x_0 of analytic vector fields. Then $\mathscr{V}_{x_0}$ is a finitely generated $\mathscr{A}_{x_0}$-module. Since $\mathscr{A}_{x_0}$ is a noetherian ring, every submodule of $\mathscr{V}_{x_0}$ is finitely generated. In particular, let $B_k(f_0, \ldots, f_m)$ denote the set of all brackets of the f_i of degree $\leqslant k$ (i.e., $B_1(f_0, \ldots, f_m) = \{f_0, \ldots \ldots, f_m\}$ and, for $k > 1$, $g \in B_k(f_0, \ldots, f_m)$ iff $g \in B_{k-1}(f_0, \ldots, f_m)$ or $g = [h_1, h_2]$ with $h_i \in B_{k_i}(f_0, \ldots, f_m)$, $k_1 + k_2 = k$). Then let $M^k_{x_0}(f_0, \ldots, f_m)$ be the submodule of $\mathscr{V}_{x_0}$ generated by the germs at x_0 of the $g \in B_k(f_0, \ldots \ldots, f_m)$. Then the $M^k_{x_0}(f_0, \ldots, f_m)$ are an increasing sequence of submodules and so there exists N such that

$$M^N_{x_0}(f_0, \ldots, f_m) = M^{N+1}_{x_0}(f_0, \ldots, f_m). \tag{12}$$

Then it is easy to verify that the RC holds at x_0 if $B_N(f_0, \ldots, f_m)(x_0)$ spans $T_{x_0}M$. That is, if we want to find out whether or not the RC holds at x_0, we compute successive brackets of the f_i, and we see whether they span $T_{x_0}M$. As soon as we have computed all brackets of degree not greater than N, where N is such that (12) holds, we do not need to compute more. If the brackets already computed do not span $T_{x_0}M$, we know that the RC at x_0 does not hold. ∎

§ 4. Local equivalence of nonlinear systems

An important feature of analyticity is that analytic objects can be expanded in a "Taylor series" about a point, and that the object's properties in a neighborhood of the point are completely determined by the Taylor coefficients at the point. We now ask

(Q2) *What is the analogue, for a nonlinear system such as* (1), *of the Taylor series?*

More precisely, we would like to obtain information about the trajectories emanating from x_0 (e.g.: Is every time-optimal trajectory bang-bang? Are the reachable sets $\mathrm{Reach}_T(x_0)$ finite unions of submanifolds? Is it possible to reach a full neighborhood of x_0 from x_0?) using values of the vector fields f_i, and of their derivatives, at x_0. It turns out that what we need is an object called the *set of Lie relations between the* f_i *at* x_0 which we denote by $\mathrm{Rel}_{x_0}(f_0, \ldots, f_m)$. To define this, we introduce $m+1$ symbols ("indeterminates") $X_0, \ldots, X_m$, and define the *free associative algebra* $\mathrm{Assoc}(X_0, \ldots, X_m)$ to be the set of all formal polynomials in the

X_i, with real coefficients, and with the X_i *not* commuting. (Precisely: a *monomial* is any finite sequence

$$X_I = X_{i_1} X_{i_2} \dots X_{i_k}, \quad 0 \leqslant i_j \leqslant m, \tag{13}$$

the elements of $\mathrm{Assoc}(X_0, \dots, X_m)$ are the sums $\sum_I a_I X_I$ such that $a_I \in \boldsymbol{R}$, and that $a_I = 0$ for all but finitely many multiindices I. Elements of $\mathrm{Assoc}(X_0, \dots, X_m)$ are multiplied using the rule $X_I X_J = X_{IJ}$, where IJ is the concatenation of the multiindices I, J. We let $X_\emptyset = 1$.) Then we define $\mathrm{Lie}(X_0, \dots, X_m)$, the *free Lie algebra* in $X_0, \dots, X_m$, to be the smallest linear subspace of $\mathrm{Assoc}(X_0, \dots, X_m)$ that contains $X_0, \dots, X_m$, and is closed under the Lie bracket. (The Lie bracket $[Y, Z]$ of two elements of $\mathrm{Assoc}(X_0, \dots, X_m)$ is simply $YZ - ZY$.) Given $f_0, \dots, f_m$, we can define a map

$$\mu(f_0, \dots, f_m)\colon \mathrm{Assoc}(X_0, \dots, X_m) \to D(M) \tag{14}$$

(where $D(M)$ is the set of all linear differential operators $C^\infty(M) \to C^\infty(M)$) by

$$\mu(f_0, \dots, f_m)(X_I) = F_I, \tag{15}$$

where, for $I = (i_1, \dots, i_k)$,

$$F_I = F_{i_1} \dots F_{i_k}. \tag{16}$$

and $F_i \in D_1(M)$ is the operator that corresponds to the vector field f_i as before.

It is easy to see that $\mu(f_0, \dots, f_m)$ maps $\mathrm{Lie}(X_0, \dots, X_m)$ into $D_1(M)$ and that $\mu(f_0, \dots, f_m)$ is an *algebra homomorphism* (i.e., it is linear and it satisfies $B_3 = B_1 B_2$ whenever $A_3 = A_1 A_2$ and $B_i = \mu(f_0, \dots, f_m)(A_i)$ for $i = 1, 2, 3$).

Let $\nu(f_0, \dots, f_m)$ denote the restriction of $\mu(f_0, \dots, f_m)$ to $\mathrm{Lie}(X_0, \dots, X_m)$. Then $\nu(f_0, \dots, f_m)$ is a *Lie algebra homomorphism* (i.e., it is $\boldsymbol{R}$-linear, and it satisfies $B_3 = [B_1, B_2]$ whenever $A_3 = [A_1, A_2]$ and $B_i = \nu(f_0, \dots, f_m)(A_i)$ for $i = 1, 2, 3$). For $x \in M$, let $\mathrm{Ev}_x\colon D_1(M) \to T_x M$ be the evaluation map (recall that $D_1(M) \sim V(M)$). Then we define

$$\mathrm{Rel}_x(f_0, \dots, f_m) = \mathrm{Ker}\big(\mathrm{Ev}_x \circ \nu(f_0, \dots, f_m)\big). \tag{17}$$

That is, the elements of $\mathrm{Rel}_x(f_0, \dots, f_m)$ are those finite linear combinations of the X_i and their brackets which vanish when the F_i are plugged in for the X_i and the resulting vector field is evaluated at x.

A *Lie subalgebra* of $\mathrm{Lie}(X_0, \dots, X_m)$ (or of any Lie algebra) is a subset S which is a linear subspace and is closed under the Lie bracket operation. Then we have

LEMMA 2. *For every system* (1) *and every state* x_0, $\mathrm{Rel}_{x_0}(f_0, \dots, f_m)$ *is a Lie subalgebra of* $\mathrm{Lie}(X_0, \dots, X_m)$ *of finite codimension.*

Proof. The first part follows from the fact that if g, h are vector fields such that $g(p) = h(p) = 0$, then $[g, h](p) = 0$. The second part is trivial. ■

THEOREM C. *Consider two systems:*

$$x^i = f_0^i(x^i) + \sum_{j=1}^{m} u_j f_j^i(x^i), \qquad x^i \in M^i \tag{18.i}$$

and initial points $x_0^i \in M^i$, $i = 1, 2$. (The control $u = (u_1, \dots, u_m)$ is required to belong to U and both systems are supposed to satisfy (I)–(IV).) *Assume both systems are analytic. Let S^i be the* MBIS *through x_0^i. Then the following two properties are equivalent:*

(a) *There exist neighborhoods W^i of x_0^i in S^i and a diffeomorphism Δ: $W^1 \to W^2$ that maps trajectories of* (18.1) *to trajectories of* (18.2)*(corresponding to the same control).*

(b) $\operatorname{Rel}_{x_0^1}(f_0^1, \dots, f_m^1) = \operatorname{Rel}_{x_0^2}(f_0^2, \dots, f_m^2)$.

Proof. Suppose (a) holds. Let $\tilde{f}_j^i$ denote the restriction of f_j^i to W^i. Since Δ maps trajectories to trajectories, it follows that

$$\Delta_*\big(\tilde{f}_j^1(x)\big) = \tilde{f}_j^2\big(\Delta(x)\big) \tag{19}$$

for all $x \in W^1$. (Here Δ_*: $T_xW^1 \to T_xW^2$ is the differential of Δ.) If we let $\tilde{F}_j^i \in D_1(W^i)$ be the corresponding first order differential operators, we can then conclude that

$$\tilde{F}_j^1 \circ \Delta^\# = \Delta^\# \circ \tilde{F}_j^2, \tag{20}$$

where $\Delta^\#$: $C^\infty(W^2) \to C^\infty(W^1)$ is the map $\varphi \to \varphi \circ \Delta$. Therefore

$$(\tilde{F}_{i_1}^1 \dots \tilde{F}_{i_k}^1) \circ \Delta^\# = \Delta^\# \circ (\tilde{F}_{i_1}^2 \dots \tilde{F}_{i_k}^2) \tag{21}$$

for all $i_1, \dots, i_k$. If

$$A = \sum a_I X_I$$

is an element of $\operatorname{Assoc}(X_0, \dots, X_m)$, then (21) implies that

$$\mu(\tilde{f}_0^1, \dots, \tilde{f}_m^1)(A) \circ \Delta^\# = \Delta^\# \circ \mu(\tilde{f}_0^2, \dots, \tilde{f}_m^2)(A). \tag{22}$$

Equation (22) holds in particular if $A \in \operatorname{Lie}(X_0, \dots, X_m)$. Therefore, for any such A, we have

$$\nu(\tilde{f}_0^1, \dots, \tilde{f}_m^1)(A)(\varphi \circ \Delta) = [\nu(\tilde{f}_0^2, \dots, \tilde{f}_m^2)(A)\varphi] \circ \Delta \tag{23}$$

if $\varphi \in C^\infty(W^2)$. Now we have:

A belongs to $\operatorname{Rel}_{x_0^1}(f_0, \dots, f_m)$ iff $\nu(f_0^1, \dots, f_m^1)(A)(x_0^1) = 0$,

i.e., iff $\nu(\tilde{f}_0^1, \ldots, \tilde{f}_m^1)(A)(x_0^1) = 0,$
i.e., iff $[\nu(\tilde{f}_0^1, \ldots, \tilde{f}_m^1)(A)\psi](x_0^1) = 0$ for all $\psi \in C^\infty(W^1)$,
i.e., iff $\nu(\tilde{f}_0^1, \ldots, \tilde{f}_m^1)(A)(\varphi \circ \Delta)(x_0^1) = 0$ for all $\varphi \in C^\infty(W^2)$,
i.e., iff $[\nu(\tilde{f}_0^2, \ldots, \tilde{f}_m^2)(A)\varphi](x_0^2) = 0$ for all $\varphi \in C^\infty(W^2)$,
i.e., iff $A \in \mathrm{Rel}_{x_0^2}(f_0^2, \ldots, f_m^2)$.

So (a) implies (b).

Let us show that (b) implies (a). Assuming (b) holds, let $M = M^1 \times M^2$. For $f \in V(M^1)$, $g \in V(M^2)$, let us define a vector field $f \times g \in V(M)$ by

$$(24) \qquad (f \times g)(x^1, x^2) = \big(f(x^1), g(x^2)\big)$$

using the canonical identification

$$(25) \qquad T_{(x^1,x^2)}M \sim T_{x^1}M^1 \times T_{x^2}M^2$$

Then it is easy to verify that

$$(26) \qquad [f \times g, f' \times g'] = [f, f'] \times [g, g']$$

for all $f, f' \in V(M^1)$, $g, g' \in V(M^2)$. Therefore the identity

$$(27) \qquad \nu(f_0^1 \times f_0^2, \ldots, f_m^1 \times f_m^2)(A) = \nu(f_0^1, \ldots, f_m^1)(A) \times \nu(f_0^2, \ldots, f_m^2)(A)$$

holds for every $A \in \mathrm{Lie}(X_0, \ldots, X_m)$.

Let L be the set of all vector fields on $M^1 \times M^2$ that are of the form $\nu(f_0^1 \times f_0^2, \ldots, f_m^1 \times f_m^2)(A)$ for some $A \in \mathrm{Lie}(X_0, \ldots, X_m)$. Then the elements of L are analytic vector fields. Moreover, the fact that $\nu(f_0^1 \times f_0^2, \ldots, f_m^1 \times f_m^2)$ is a Lie algebra homomorphism implies that L is a Lie algebra. Therefore we can conclude from Nagano's theorem that there exists a MIM of L through the point (x_0^1, x_0^2). Let this MIM be denoted by S. We can define two maps p^1: $S \to M^1$, p^2: $S \to M^2$, by $p^i(x^1, x^2) = x^i$. If v is a tangent vector to S at (x_0^1, x_0^2), then $v = (v^1, v^2)$, where

$$v^i = \nu(f_0^i, \ldots, f_m^i)(A)(x_0^i)$$

for some A. Then v^i vanishes iff $A \in \mathrm{Rel}_{x_0^i}(f_0^i, \ldots, f_m^i)$. In view of Hypothesis (b), we conclude that $v^1 = 0$ iff $v^2 = 0$. Clearly, $p_*^i(v) = v^i$. Therefore if $p_*^1(v) = 0$, it follows that $v = 0$. So the differential of p^1 is injective at (x_0^1, x_0^2). A similar conclusion holds for p^2. Therefore, it follows from the implicit function theorem that there is a connected open neighborhood W of (x_0^1, x_0^2) which is mapped diffeomorphically by p^1, p^2 onto submanifolds W^1, W^2 of M^1, M^2. The map Δ: $W^1 \to W^2$ given by $\Delta = p^2 \circ (p^1)^{-1}$ is therefore a diffeomorphism.

If $x^1 \in W^1$, let $x^2 = \Delta(x^1)$. A vector $v^1 \in T_{x^1}M^1$ is tangent to W^1 iff $v^1 = p_*^1(v)$ for some v in $T_{(x^1,x^2)}W$. Since W is open in S, and S is an integral manifold for L, v belongs to $T_{(x^1,x^2)}W$ iff v is of the form $\nu(f_0^1\times f_0^2,\dots,f_m^1\times f_m^2)(A)(x_0^1)$ for some $A\in\mathrm{Lie}(X_0,\dots,X_m)$. So $v^1\in T_{x^1}W^1$ iff $v^1=g(x_0^1)$ for some g in $\mathrm{Lie}(f_0^1,\dots,f_m^1)$. This shows that W^1 is an integral manifold of $\mathrm{Lie}(f_0^1,\dots,f_m^1)$. So $W^1\subseteq S^1$ and W^1 is open in S^1. A similar argument shows that $W^2\subseteq S^2$ and W^2 is open in S^2.

Now let $x^1(\cdot)\colon I\to W^1$ be a trajectory of (18.1) corresponding to a control $u(\cdot)\colon I\to U$. Let

$$u(\cdot)=\big(u_1(\cdot),\dots,u_m(\cdot)\big).$$

Let $x^2(\cdot)\colon I\to W^2$ be given by $x^2(t)=\Delta\big(x^1(t)\big)$. Then $x^2(\cdot)$ is absolutely continuous because Δ is a diffeomorphism. For almost every t, the tangent vector $\dot x^2(t)$ is $\Delta_*\big(\dot x^1(t)\big)$, i.e.,

$$\Delta_*\Big(f_0^1\big(x^1(t)\big)+\sum_{j=1}^m u_j(t)f_j^1\big(x^1(t)\big)\Big).$$

But our construction of Δ clearly implies that $\Delta_*\big(f_j^1(x)\big)=f_j^2\big(\Delta(x)\big)$ for $j=0,\dots,m$, $x\in W^1$. Therefore

$$\dot x^2(t)=f_0^2\big(x^2(t)\big)+\sum_{j=1}^m u_j(t)f_j^2\big(x^2(t)\big).$$

So $x^2(\cdot)$ is a trajectory for $u(\cdot)$. This means that Δ maps trajectories to trajectories and our proof is complete. ∎

We now consider the special case when both systems satisty the AP from x_0^i. In this case the MBIS through x_0^i is open in M^i. (Actually, the MBIS is the connected component of M^i through x_0^i.) So Theorem C gives:

THEOREM D. *In addition to the hypotheses of Theorem* C, *assume that both systems* (18.1), (18.2) *have the accessibility property from* x_0^1, x_0^2, *respectively. Then the following are equivalent*:

(a) *There are neighborhoods* W^i *of* x_0^i *in* M^i, *and a diffeomorphism* $\Delta\colon W^1\to W^2$ *that maps trajectories of* (18.1) *to trajectories of* (18.2) *corresponding to the same control.*

(b) $\mathrm{Rel}_{x_0^1}(f_0^1,\dots,f_m^1)=\mathrm{Rel}_{x_0^2}(f_0^2,\dots,f_m^2)$.

Theorem D shows that for analytic systems (1) that have the AP from a point p the object $\mathrm{Rel}_p(f_0,\dots,f_m)$ plays a role similar to that of the family of Taylor coefficients for an analytic function.

§ 5. Local equivalence to linear systems

Let us say that two systems (18.1), (18.2) are *locally equivalent* about points x_0^1, x_0^2 if property (a) of the statement of Theorem D holds. We now ask:

(Q3) *When is a system* (1), *with an initial state* x_0 *such that the* AP *from* x_0 *holds, locally equivalent to a linear system about* 0?

To answer this question we first observe that if two analytic systems (18.1) (18.2) are locally equivalent about points x_0^1, x_0^2, and if (18.1) has the AP from x_0^1, then (18.2) has the AP from x_0^2. So in order to answer question (Q3) we may limit ourselves to considering linear systems which have the AP from 0, i.e., to systems of the form (5) that are completely controllable. Let $\mathrm{Lin}(m)$ be the class of all these systems. For a system $S \in \mathrm{Lin}(m)$, let $\mathrm{Rel}(S)$ be the set $\mathrm{Rel}_0(\tilde{f}_0, \ldots, \tilde{f}_m)$, where $\tilde{f}_0(x) = Ax$, $\tilde{f}_i(x) = b_i$ for $i > 0$.

Theorem D tells us that a necessary and sufficient condition for a system (1), which has the AP from x_0, to be locally equivalent to an $S \in \mathrm{Lin}(m)$ about 0 is that

$$\mathrm{Rel}_{x_0}(f_0, \ldots, f_m) = \mathrm{Rel}(S) \tag{28}$$

for some $S \in \mathrm{Lin}(m)$. So our first task is to determine which Lie subalgebras of $\mathrm{Lie}(X_0, \ldots, X_m)$ can be equal to $\mathrm{Rel}(S)$ for some $S \in \mathrm{Lin}(m)$. We begin by listing some properties that $\mathrm{Rel}(S)$ must have for every $S \in \mathrm{Lin}(m)$.

We claim that if $S \in \mathrm{Lin}(m)$ and $\Lambda = \mathrm{Rel}(S)$, then

(I) $X_0 \in \Lambda$,

(II) $(\mathrm{ad}\, X_{i_1}) \ldots (\mathrm{ad}\, X_{i_r})(X_{i_{r+1}}) \in \Lambda$

for every $r \in \mathbf{Z}$, $r \geqslant 0$, and every sequence $i_1, \ldots, i_{r+1}$ of integers between 0 and m such that at least two of the i_j are > 0 (here if $Z \in \mathrm{Lie}(X_0, \ldots, X_m)$, $\mathrm{ad}\, Z$ denotes the linear map $X \to [Z, X]$).

That (I) holds is simply a consequence of the fact that the vector field $\tilde{f}_0$ vanishes at 0. To prove (II), just observe that any time a linear vector field (i.e., a vector field of the form $g(x) = Px$, P a square matrix) is bracketed with a constant vector field, the result is again constant. Also, any bracket of two constant vector fields vanishes. From this it follows that any time we bracket several of the $\tilde{f}_i$ in such a way that at least two of them are taken from $\{\tilde{f}_1, \ldots, \tilde{f}_m\}$, the result will have to vanish. So (II) is proved.

There is a third condition which Λ must satisfy, namely,

(III) Λ *is a subalgebra of* $\mathrm{Lie}(X_0, \ldots, X_m)$ *of finite codimension.*

LEMMA 3. *Assume that* $\Lambda \subseteq \mathrm{Lie}(X_0, \ldots, X_m)$. *Then* $\Lambda = \mathrm{Rel}(S)$ *for some* $S \in \mathrm{Lin}(m)$ *if and only if* (I), (II), (III) *above hold.*

Proof. Only the "if" part requires proof. Let

$$V = \mathrm{Lie}(X_0, \ldots, X_m)/\Lambda.$$

Then V is a finite-dimensional linear space. Let

$$\pi\colon \mathrm{Lie}(X_0, \ldots, X_m) \to V$$

be the canonical projection. Define vectors $b_1, \ldots, b_m$ in V by

$$b_i = \pi(X_i).$$

Also, define a linear map $A\colon V \to V$ by

$$(29) \qquad A\pi(Z) = -\pi([X_0, Z]).$$

This map is well defined because if $\pi(Z) = \pi(Z')$, then $Z - Z' \in \Lambda$ and so $[X_0, Z - Z'] \in \Lambda$ since $X_0 \in \Lambda$ and Λ is a Lie subalgebra.

Then we have defined a linear system S. (The reader who so wishes may choose a basis for V, and think of A as a matrix and the b_i as row vectors.) We now have to show that $\mathrm{Rel}(S) = \Lambda$.

As before, let $\tilde{f}_0(x) = Ax$, $\tilde{f}_i(x) = b_i$ for $i > 0$. Let μ denote the map $\mu(\tilde{f}_0, \ldots, \tilde{f}_m)$, so that $\mu(X_i) = \tilde{f}_i$, and μ is a homomorphism from $\mathrm{Lie}(X_0, \ldots$ $\ldots, X_m)$ into the Lie algebra of vector fields on V.

Let $Z \in \mathrm{Lie}(X_0, \ldots, X_m)$. Then Z has a unique expression of the form

$$Z = aX_0 + \sum_i \sum_{j=1}^{m} a_{ij} (\mathrm{ad}\, X_0)^i (X_j) + Z',$$

where Z' is a linear combination of brackets which involve at least two X_j, $j > 0$. It is a simple exercise to show that if $g(x) = Px$, $h(x) = c$ (P a linear map $V \to V$, c a constant vector), then $[g, h](x) = -Pc$. Therefore

$$\mu\big((\mathrm{ad}\, X_0)^i (X_j)\big) = (\mathrm{ad} \tilde{f}_0)^i (f_j)$$

is the constant vector field whose value is $(-A)^i b_j$. But formula (29) implies that

$$(-A)^i b_j = \pi\big((\mathrm{ad}\, X_0)^i (X_j)\big).$$

Therefore

$$\mu\big((\mathrm{ad}\, X_0)^i (X_j)\big)(x) = \pi\big((\mathrm{ad}\, X_0)^i (X_j)\big)$$

for all $x \in V$. In particular, $\mathrm{Ev}_0\big(\mu\big((\mathrm{ad}\, X)^i (X_j)\big)\big)$ equals $\pi\big((\mathrm{ad}\, X_0)^i (X_j)\big)$. On the other hand, $\mathrm{Ev}_0\big(\mu(X_0)\big) = \tilde{f}_0(0) = 0$, and $\pi(X_0) = 0$ because $X_0 \in \Lambda$. Finally, $\pi(Z') = 0$ because $Z' \in \Lambda$, and $\mu(Z') = 0$ because every bracket $\big[\tilde{f}_{i_1}[\tilde{f}_{i_2} \ldots [\tilde{f}_{i_{r-1}}, \tilde{f}_{i_r}] \ldots]\big]$, where at least two i_j are greater than 0, must vanish. So we can conclude that

$$\pi(Z) = \mathrm{Ev}_0\big(\mu(Z)\big).$$

But then $\mathrm{Ker}\, \pi = \mathrm{Ker}\,(\mathrm{Ev}_0 \mu)$, i.e., $\Lambda = \mathrm{Rel}(S)$. ∎

Now we can answer question (Q3):

THEOREM E. *Consider an analytic system* (1) *and a state* x_0 *such that* $f_0(x_0) = 0$ *and* (1) *has the* AP *from* x_0. *Then* (1) *is locally equivalent to a linear system by an equivalence map which takes* x_0 *to* 0 *if and only if every bracket* $g = \big[f_{i_1}\big[f_{i_2}\big[\ldots[f_{i_{r-1}}, f_{i_r}]\big]\big]\big]$ *for which at least two* i_j *are* $\neq 0$ *satisfies* $g(x_0) = 0$.

In a similar fashion, it is easy to derive conditions for local equivalence to a linear system by means of a map which does not necessarily send x_0 to 0 (and without assuming that $f(x_0) = 0$). It turns out that it is better to study equivalence to systems of the form

$$\dot{x} = Ax + c + \sum_{i=1}^{m} u_i b_i, \tag{30}$$

and that every such system initialized at a point $\bar{x}$ is equivalent to a similar system initialized at 0.

THEOREM F. *Consider an analytic system* (1), *and let* x_0 *be a state. Then* (1) *is locally equivalent to some system* (30), *with an equivalence that takes* x_0 *to some point* $\bar{x}_0$, *if and only if the following two conditions hold*:

(a) *Every bracket* $g = (\mathrm{ad} f_{i_1}) \ldots (\mathrm{ad} f_{i_{r-1}})(f_{i_r})$ *for which at least two* i_j*'s are greater than* 0 *satisfies* $g(x_0) = 0$.

(b) *Whenever a linear combination*

$$h = \sum_{i} \sum_{j>0} a_{ij} (\mathrm{ad} f_0)^i (f_j) \tag{31}$$

satisfies $h(x_0) = 0$, *then necessarily* $[f_0, h](x_0) = 0$.

Proof. If (1) is indeed a system of the form (30), then every vector field h of the type (31) is constant. So, if h vanishes at one point, it follows that $h \equiv 0$, and so $[f_0, h] \equiv 0$. So (b) holds. Also, every vector field g of the type considered in (a) is a bracket of two or more h's of the form (31). So the g's must vanish and (a) holds as well. So (a) and (b) hold for systems of the form (30) and therefore they hold for every system which is equivalent to one of the form (30).

To prove the converse, suppose (a) and (b) hold. Let $L = \mathrm{Lie}(X_0, \ldots, X_m)$, $\Lambda = \mathrm{Rel}_{x_0}(f_0, \ldots, f_m)$. Let L_0 be the Lie subalgebra of L generated by the X_i, $i > 0$, and by all the brackets $(\mathrm{ad} f_{i_1}) \ldots (\mathrm{ad} f_{i_r})(f_{i_{r+1}})$ with $0 \leqslant i_j \leqslant m$, $r \geqslant 1$. Then L_0 has codimension one in L. Let $\Lambda_0 = L_0 \cap \Lambda$.

Let $V_0 = L_0 / L_0 \cap \Lambda$. Then V_0 is a finite-dimensional linear space. Define vectors $b_i^0 \in V_0$ by

$$b_i^0 = \pi_0(X_i), \quad i = 1, \ldots, m,$$

where $\pi_0\colon L_0 \to V_0$ is the canonical projection. Then define a linear map $A_0 = V_0 \to V_0$ by

$$A_0\big(\pi_0(Z)\big) = -\pi_0([X_0, Z]), \quad Z \in L_0.$$

This map is well defined because hypotheses (a), (b) imply that Λ_0 is invariant under $\operatorname{ad} X_0$. Exactly as in the proof of Theorem E, one shows that if $\tilde{f}_0^0(x) = A_0 x, \tilde{f}_i^0(x) = b_i^0$ for $i > 0$, then $\mu(\tilde{f}_0^0, \ldots, \tilde{f}_m^0)(Z)(0) = \pi_0(Z)$ for all $z \in L_0$, so that

$$\Lambda_0 = \operatorname{Rel}_0(\tilde{f}_0^0, \ldots, \tilde{f}_m^0).$$

Now let us distinguish two cases.

Case 1. $f_0(x_0)$ belongs to the linear span Σ of the vectors $f_i(x_0)$ $(i > 0)$, $(\operatorname{ad} f_{i_1}) \ldots (\operatorname{ad} f_{i_r})(f_{i_{r+1}})(0)$ $(r > 0, 0 \leqslant i_j \leqslant m)$.

Case 2. $f_0^0(x_0) \notin \Sigma$.

In Case 1 we take $V = V_0$, $A = A_0$, $b_i = b_i^0$. We choose some relation of the form

$$\bar{Z} = Z_0 = \sum_i \sum_{j>0} \bar{a}_{ij} (\operatorname{ad} X_0)^i (X_j)$$

such that $\bar{Z} \in \Lambda$, and we let $c = \sum_i a_{ij} A^i b_j$. Then we have defined a system of the form (30). Let $\tilde{f}_0(x) = Ax + c$, $\tilde{f}_i(x) = b_i$ for $i > 0$. Then if $Z \in L$, we can write Z as a sum

$$Z = aX_0 + \sum_i \sum_{j>0} a_{ij} (\operatorname{ad} X_0)^i (X_j) + Z', \tag{32}$$

where Z' involves brackets with two or more f_j, $j > 0$. Then it is clear that $Z' \in \Lambda_0$, and that $\mu(\tilde{f}_0, \ldots, \tilde{f}_m)(Z') = 0$. Also,

$$\mu(\tilde{f}_0, \ldots, \tilde{f}_m)\big((\operatorname{ad} X_0)^i X_j\big)(0) = A^i b_j = \pi_0\big((\operatorname{ad} X_0)^i X_j\big).$$

Finally,

$$\mu(\tilde{f}_0, \ldots, \tilde{f}_m)(X_0)(0) = c = \pi_0(X_0 - \bar{Z}).$$

So we have

$$\mu(\tilde{f}_0, \ldots, \tilde{f}_m)(Z)(0) = \pi_0(Z - a\bar{Z}).$$

Since $\bar{Z} \in \Lambda$, we see that $z \in \Lambda$ iff $\mu(\tilde{f}_0, \ldots, \tilde{f}_m)(Z)(0) = 0$. Therefore

$$\Lambda = \operatorname{Rel}_0(\tilde{f}_0, \ldots, \tilde{f}_m). \tag{33}$$

In Case 2 we take $V = \boldsymbol{R} \oplus V_0$, $A = (0, A_0)$, $b_i = (0, b_i^0)$, $c = (1, 0)$. Then $\Lambda = \Lambda_0$. If Z is as in (32), we see that $\mu(\tilde{f}_0, \ldots, \tilde{f}_m)(Z)(0)$ equals

$$\big(a, \pi_0(Z - aX_0)\big).$$

Therefore $\mu(\tilde{f}_0, \ldots, \tilde{f}_m)(Z)(0) = 0$ iff $a = 0$ and $Z - aX_0 \in \Lambda_0$, i.e., iff $Z \in \Lambda$. So again (33) holds.

We have proved that both in Case 1 and Case 2 $\operatorname{Rel}_{x_0}(f_0, \ldots, f_m)$ is equal to $\operatorname{Rel}_0(\tilde{f}_0, \ldots, \tilde{f}_m)$ for some system (30). The conclusion follows. ■

§ 6. Bang-bang theorems

It is well known that linear systems satisfy the bang-bang property. Therefore any system which is equivalent to a linear system also satisfies a bang-bang property. Now, the obstruction for a system to be equivalent to a linear system is the existence of brackets $[f_j, (\mathrm{ad} f_0)^i(f_k)]$, $j > 0$, $k > 0$, that do not vanish. So it is reasonable to expect that some condition on these brackets, weaker than the requirement that they vanish, might suffice to prove a bang-bang theorem. This can be substantiated, at least for $m = 1$.

Precisely, consider a system

$$\dot{x} = f_0(x) + uf_1(x), \qquad |u| \leqslant 1. \tag{34}$$

THEOREM G. *Suppose the system* (34) *is analytic. Assume that for every state x_0 and every i there is a neighborhood U of x_0 such that*

$$[f_1, (\mathrm{ad} f_0)^i(f_1)] = \sum_{k=0}^{i+1} \varphi_{ik}(\mathrm{ad} f_0)^k(f_1) \tag{35}$$

on U, where the φ_{ik} are analytic functions on U such that $|\varphi_{i,i+1}(x)| < 1$ for all $x \in U$. Then (34) *satisfies the following property:*

(BBBNS) (Bang-bang with bounds on the number of switchings) *For every compact set $K \subseteq M$, and every time $T > 0$, there exists a positive integer N such that, whenever γ is a time-optimal trajectory of* (1) *which is entirely contained in K, and goes from a point $p \in K$ to a point $q \in K$ in time not greater than T, then there exists a time-optimal trajectory γ' from p to q which is bang-bang with at most N switchings.*

The proof of this theorem is quite long and it is given in Sussmann [10]. An explanation of why this result is useful for nonlinear synthesis theory can be found in Sussmann [12]. An important open problem is to find a good generalization of Theorem G to the multiinput case.

§ 7. Lie series and local controllability

Consider a system

$$\dot{x} = f_0(x) + \sum_{i=1}^{m} u_i f_i(x), \qquad |u_i| \leqslant 1. \tag{36}$$

We say that (39) is *small-time locally controllable* (STLC) from a point x_0 if, for every $T > 0$, the set

$$\mathrm{Reach}_{\leqslant T}(x_0) = \bigcup_{0 \leqslant t \leqslant T} \mathrm{Reach}_t(x_0) \tag{37}$$

contains x_0 in its interior. Small-time local controllability is a local property, and it is invariant under changes of coordinates, so it should be possible to characterize it in terms of Lie bracket relations. A good necessary and sufficient condition for STLC is not known at present but substantial progress has been made. Here we present a rough sketch of one method that has been used. It is based on the idea of associating with every admissible control an *exponential Lie series.*

Let $\widehat{\mathrm{Assoc}}(X_0, \ldots, X_m)$ be the set of all *formal power series* in the indeterminates $X_0, \ldots, X_m$. The elements of $\widehat{\mathrm{Assoc}}(X_0, \ldots, X_m)$ are all formal sums $\sum a_I X_I$, where the sum runs over all multiindices $I = (i_1, \ldots, i_r)$ of arbitrary length r such that $0 \leqslant i_j \leqslant m$ for all j. It is not required that $a_I = 0$ for all but finitely many I's.

For each r, let $\mathrm{Assoc}_r(X_0, \ldots, X_m)$ denote the set of all *homogeneous* elements of degree r of $\mathrm{Assoc}(X_0, \ldots, X_m)$, i.e., the set of all formal sums $\sum a_I X_I$, where I runs over all multiindices of length r. Then $\widehat{\mathrm{Assoc}}(X_0, \ldots, X_m)$ can be identified with the infinite product $\prod_{r=0}^{\infty} \mathrm{Assoc}_r(X_0, \ldots, X_m)$, i.e., with the set of all formal infinite sums

$$\sum_{r=0}^{\infty} Z_r, \qquad Z_r \in \mathrm{Assoc}_r(X_0, \ldots, X_m).$$

Let

$$\mathrm{Lie}_r(X_0, \ldots, X_m) = \mathrm{Assoc}_r(X_0, \ldots, X_m) \cap \mathrm{Lie}(X_0, \ldots, X_m)$$

and let $\widehat{\mathrm{Lie}}(X_0, \ldots, X_m)$ be the set of all sums $\sum_{r=1}^{\infty} Z_r$ with $Z_r \in \mathrm{Lie}_r(X_0, \ldots, X_m)$.

The elements of $\widehat{\mathrm{Lie}}(X_0, \ldots, X_m)$ are called *Lie series* in $X_0, \ldots, X_m$, and the elements of $\widehat{\mathrm{Assoc}}(X_0, \ldots, X_m)$ are the *noncommutative formal power series* in $X_0, \ldots, X_m$.

We let $\widehat{\mathrm{Assoc}}^0(X_0, \ldots, X_m)$ denote the set of formal power series with no constant term and then $1 + \widehat{\mathrm{Assoc}}^0(X_0, \ldots, X_m)$ is the set of all formal power series whose constant term is equal to 1. If $S \in \widehat{\mathrm{Assoc}}^0(X_0, \ldots, X_m)$, then the series

$$\exp(S) = \sum_{i=0}^{\infty} \frac{1}{i!} S^i \tag{38}$$

is well defined because, for each r, if

$$p_r \colon \widehat{\mathrm{Assoc}}(X_0, \ldots, X_m) \to \widehat{\mathrm{Assoc}}_r(X_0, \ldots, X_m)$$

is the obvious projection, then $p_r(S^i) = 0$ for all but finitely many i.

Similarly, the series

$$\log(1+S) = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i} S^i \tag{39}$$

is also well defined. It is easy to see that exp maps $\widehat{\mathrm{Assoc}}^0(X_0, \ldots, X_m)$ onto $1+\widehat{\mathrm{Assoc}}^0(X_0, \ldots, X_m)$, log maps $1+\widehat{\mathrm{Assoc}}^0(X_0, \ldots, X_m)$ onto $\widehat{\mathrm{Assoc}}^0(X_0, \ldots, X_m)$, and exp and log are inverse maps.

A series $Z \in \widehat{\mathrm{Assoc}}(X_0, \ldots, X_m)$ which is of the form $\exp(S)$, $S \in \mathrm{Lie}(X_0, \ldots, X_m)$, is called an *exponential Lie series* (ELS) in $X_0, \ldots, X_m$.

We use $\mathrm{ELS}(X_0, \ldots, X_m)$ to denote the set of all exponential Lie series in $X_0, \ldots, X_m$. One of the most useful results of the theory of Lie series is the *Campbell–Hausdorff formula* which says that $\mathrm{ELS}(X_0, \ldots, X_m)$ is a group under multiplication. Precisely, if $\exp(Z)$ and $\exp(Z')$ are ELS's, then

$$\exp(Z)\exp(Z') = \exp(Z''), \tag{40}$$

where Z'' is a Lie series, whose first few terms are

$$Z'' = Z+Z'+\tfrac{1}{2}[Z, Z']+\tfrac{1}{12}[Z, [Z, Z']]+\tfrac{1}{12}[Z', [Z', Z]]+\ldots \tag{41}$$

Now let $u\colon [0, T]\to \boldsymbol{R}^m$ be bounded and measurable. We define a series, which we denote by $\mathrm{Ser}(u)$, and a series-valued curve

$$S_u\colon [0, T]\to \widehat{\mathrm{Assoc}}(X_0, \ldots, X_m)$$

as follows: S_u is the solution of the differential equation

$$\dot{S}(t) = S(t)\Big(X_0+\sum u_i(t)X_i\Big) \tag{42}$$

for which $S(0) = 1$. Then

$$\mathrm{Ser}(u) = S_u(T). \tag{43}$$

Then an explicit expression for S_u can be given in terms of *iterated integrals*. For any $u\colon [0, T]\to \boldsymbol{R}^m$, let u_0 always denote the function constantly equal to 1. Then define, for any multiindex $I = (i_1, \ldots, i_r)$, and $t \in [0, T]$:

$$\int_0^t u_I = \int_0^t u_{i_1}(s_1) \int_0^{s_1} u_{i_2}(s_2) \ldots \int_0^{s_{r-1}} u_{i_r}(s_r)\,ds_r \ldots ds_1.$$

Also, let I^* denote the multiindex I in reverse order. Then S_u turns out to be given by the formula

$$S_u(t) = \sum_I \Big(\int_0^t u_I\Big) X_{I^*}. \tag{44}$$

Also, one can prove (e.g., by the Campbell–Hausdorff formula) that $S_u(t)$ is always an ELS. Therefore we have defined a map

$$\mathrm{Ser}\colon \mathscr{U}_m \to \mathrm{ELS}(X_0, \ldots, X_m),$$

where $\mathscr{U}_m$ is the class of all bounded, measurable, $\boldsymbol{R}^m$-valued functions on intervals $[0, T]$, $T \geqslant 0$. Moreover, if $\mathscr{U}_m$ is equipped with the operation of concatenation (which turns $\mathscr{U}_m$ into a semigroup) and $\mathrm{ELS}(X_0, \ldots, X_m)$ with multiplication, then Ser is a semigroup homomorphism. In addition, the map Ser is one-to-one.

Now let $f_0, \ldots, f_m$ be smooth vector fields on a manifold M. If $u \in \mathscr{U}_m$, $u\colon [0, T] \to \boldsymbol{R}^m$, then we define $S_u(f_0, \ldots, f_m)(t)$ and $\mathrm{Ser}(u)(f_0, \ldots, f_m)$ to be the results of "plugging in" the f_i for the X_i in $S_u(t)$, $\mathrm{Ser}(u)$, respectively. Then $S_u(f_0, \ldots, f_m)(t)$ and $\mathrm{Ser}(u)(f_0, \ldots, f_m)$ are *formal series of partial differential operators*. Precisely, if we let

$$f_I = f_{i_1} f_{i_2} \cdots f_{i_r}$$

for $I = (i_1, \ldots, i_r)$, then

$$S_u(f_0, \ldots, f_m)(t) = \sum_I \Big(\int_0^t u_I\Big) f_I, \tag{45}$$

and

$$\mathrm{Ser}(u)(f_0, \ldots, f_m) = \sum_I \Big(\int_0^T u_I\Big) f_I. \tag{46}$$

Each $u \in \mathscr{U}_m$ is defined on an interval $[0, T]$. Let us use T_u to denote the T that corresponds to a given u. For $u \in \mathscr{U}_m$, $x \in M$, let $t \to \pi(u, x, t)$ denote the trajectory corresponding to the control u which goes through x when $T = 0$. Then $\pi(u, x, t)$ is defined for all t in some interval $J(x, u) \subseteq [0, T]$ such that $0 \in J(x, u)$ and that if $\sup J(x, u) = t < T$, then $t \notin J(x, u)$. If $K \subseteq M$ is compact and $a > 0$ is a constant, then there is an $\varepsilon > 0$ such that $J(x, u) \supseteq [0, \varepsilon]$ for all $x \in K$ and all $u \in \mathscr{U}_m$ such that $\varepsilon \leqslant T_u$ and

$$\sup\{\|u(t)\|\colon 0 \leqslant t \leqslant \varepsilon\} \leqslant a.$$

Let us refer to any ε with these properties as a *good* ε for the given K, a.

If $\varphi\colon M \to \boldsymbol{R}$ is a smooth function, and if $u \in \mathscr{U}_m$, then we can define a function $P_u(\varphi)$ by

$$P_u(\varphi)(x) = \varphi\big(\pi(u, x, T_u)\big).$$

The domain of $P_u(\varphi)$ is the open set

$$Q_u = \{x\colon J(x, u) = [0, T_u]\}.$$

Now consider a compact set K and an $a > 0$. If $\varepsilon > 0$ is good for K, a, then $K \subseteq Q_u$ for all $u \in \mathscr{U}_m(\varepsilon, a)$, where

$$\mathscr{U}_m(\varepsilon, a) = \{u\colon T_u \leqslant \varepsilon, \sup \|u, (t)\| \leqslant a\}.$$

Hence, for any φ: $M\to\mathbf{R}$, the function $P_u(\varphi)$ is well defined on K for all $u\in\mathscr{U}_m(\varepsilon, a)$. If φ is smooth, then we can consider the series

$$\mathrm{Ser}(u)(f_0,\dots,f_m)(\varphi)=\sum_I\Big(\int_0^{T_u}u_I\Big)f_{I\bullet}\varphi, \tag{47}$$

where $f_{I\bullet}\varphi$ is the result of applying to φ the partial differential operator $f_{I\bullet}$.

It turns out that $\mathrm{Ser}(u)(f_0,\dots,f_m)(\varphi)$ *is an asymptotic series for* $P_u(\varphi)$ *as* $T_u\to 0$ while $u\in\mathscr{U}_m(\varepsilon, a)$. Precisely, let us define the "truncated" series $\mathrm{Ser}_N(u)$ to be the sum $\sum_{|I|\leq N}\Big(\int_0^{T_u}u_I\Big)X_{I\bullet}$, where $|I|$ is the length of I. Then define $\mathrm{Ser}_N(u)(f_0,\dots,f_m)$ and $\mathrm{Ser}_N(u)(f_0,\dots,f_m)(\varphi)$ in an obvious way.

LEMMA 4. *Consider a system* (1), *where the f_i are C^∞ vector fields. Let $K\subseteq M$ be compact, $a>0$, ε good for K, a. Then for every smooth φ: $M\to\mathbf{R}$ and every nonnegative integer N, there exists a constant $C>0$ such that*

$$\big|\varphi\big(\pi(x,u,T_u)\big)-\mathrm{Ser}_N(u)(f_0,\dots,f_m)(\varphi)(x)\big|\leq CT_u^{N+1}$$

for all $x\in K$ and all $u\in\mathscr{U}_m(\varepsilon, a)$.

The proof of this lemma is straightforward and we omit it. We now explain how the lemma can be used to generate "control variations." Suppose $\{u_\sigma\}$ is a family of controls depending on a real parameter σ and belonging to $\mathscr{U}_m(\varepsilon, a)$ for some fixed ε, a. Assume that $T_{u_\sigma}\to 0$ as $\varphi\to 0_+$ and that the coefficients of $\mathrm{Ser}(u_\sigma)$ are finite linear combinations of (not necessarily integral) powers of σ. Since $\mathrm{Ser}(u_\sigma)$ is an exponential Lie series, we have

$$\mathrm{Ser}(u_\sigma)=\exp\big(Z(u_\sigma)\big),$$

where

$$Z(u_\sigma)=\log\big(\mathrm{Ser}(u_\sigma)\big),$$

so that $Z(u_\sigma)$ is also a series in powers of σ. Then we can write

$$Z(u_\sigma)=\sigma^\varrho Z'+o(\sigma^\varrho),$$

where Z' is a Lie polynomial in $X_0,\dots,X_m$. From this it follows that

$$\mathrm{Ser}(u_\sigma)=1+\sigma^\varrho Z'+o(\sigma^\varrho).$$

Since $\mathrm{Ser}(u_\sigma)$ is a series in powers of σ, the coefficient of X_0 in $\mathrm{Ser}(u_\sigma)$ is also a series in powers of σ, and so

$$T_{u_\sigma}\sim\beta\sigma^\lambda\quad\text{as}\quad\sigma\to 0_+,$$

for some $\beta>0$, $\lambda>0$. Then the lemma implies that for any smooth function φ

$$\varphi\big(\pi(x,u_\sigma,T_{u_\sigma})\big)=\varphi(x)+\sigma^\varrho Z'(f_0,\dots,f_m)(\varphi)(x)+o(\sigma^\varrho) \tag{48}$$

as $\sigma \to 0$, uniformly as long as x stays in a fixed compact set. (Indeed, take N so large that $N >$ degree (Z'), and that $\lambda(N+1) > \varrho$. Then

$$\mathrm{Ser}_N(u_\sigma) = 1 + \sigma^\varrho Z' + o(\sigma^\varrho)$$

and so (48) holds if $\varphi(\pi(x, u_\sigma, T_{u_\sigma}))$ is replaced by $\mathrm{Ser}_N(u_\sigma)(f_0, \ldots, f_m)(\varphi)(x)$. On the other hand, the lemma implies that $\mathrm{Ser}_N(u_\sigma)(f_0, \ldots, f_m)(\varphi)(x)$ and $\varphi(\pi(x, u_\sigma, T_{u_\sigma}))$ differ by an $O(T_{u_\sigma}^{N+1})$, i.e. by an $o(\sigma_\sigma)$).

The vector field $Z'(f_0, \ldots, f_m)$ then has the property that for every x the tangent vector at $t = 0$ to the curve

$$t \to \pi(x, u_{t^{1/\varrho}}, T_{u_{t^{1/\varrho}}})$$

is precisely $Z'(f_0, \ldots, f_m)(x)$.

As an illustration, let us take $m = 2$. For each $\sigma > 0$, let u_σ be defined on $[0, 4\sigma]$, by letting $(u_\sigma)_1$ equal 1, 0, -1, 0, respectively, on $[0, \sigma]$, $(\sigma, 2\sigma]$, $(2\sigma, 3\sigma]$, $(3\sigma, 4\sigma]$, and $(u_\sigma)_2$ equal 0, 1, 0, -1, on the same intervals. Let $f_0 = 0$, $f_1 = f$, $f_2 = g$, where f and g are smooth vector fields. Then (using the notation of § 3),

$$\pi(x, u_\sigma, 4\sigma) = \Phi^g(-\sigma)\Phi^f(-\sigma)\Phi^g(\sigma)\Phi^f(\sigma)(x)$$

and

$$\mathrm{Ser}(u_\sigma) = \exp(\sigma X_0 - \sigma X_1)\exp(\sigma X_0 - \sigma X_2)\exp(\sigma X_0 + \sigma X_1)\cdot \cdot\exp(\sigma X_0 + \sigma X_2). \tag{49}$$

Since we are going to "plug in" 0, f, g for X_0, X_1, X_2, we can forget about X_0 in (49).

From the Campbell–Hausdorff formula, we get

$$\mathrm{Ser}(u_\sigma) = \exp(\sigma^2[X_1, X_2]) + P,$$

where P involves monomials that contain X_0. So the tangent vector to the curve

$$t \to \pi(x, u_{\sqrt{t}}, 4\sqrt{t})$$

at $t = 0$ is $[f, g](x)$. This proves the equivalence of Definition 1 and Definition 2 of § 3.

Now let us consider a system

$$\dot{x} = f_0(x) + uf_1(x), \qquad |u| \leqslant 1. \tag{50}$$

Hermes introduced a condition which, he conjectured, implied small-time local controllability. Recently, we have been able to prove Hermes' conjecture, using the exponential Lie series formalism. Let us first state Hermes' condition.

For each $k \geqslant 0$, let $\mathcal{S}^k(f, g)$ denote the linear span of all brackets of f's and g's that involve no more than k g's. Then

$$\mathcal{S}^0(f, g) \subseteq \mathcal{S}^1(f, g) \subseteq \mathcal{S}^2(f, g) \ldots$$

At any point x_0 let $\mathscr{S}^k(f, g)(x_0)$ denote the set of all $X(x_0)$ with $X \in \mathscr{S}^k(f, g)$. The following is Hermes' conjecture:

THEOREM H. *Suppose that the system* (50) *satisfies the* RC *at* x_0, *that* $f(x_0) = 0$, *and that*

$$\mathscr{S}^k(f, g)(x_0) = \mathscr{S}^{k+1}(f, g)(x_0)$$

for every odd k. *Then* (50) *is* STLC *from* x_0.

The proof of this result is too long to be given here. We shall limit ourselves to a brief sketch. Pick a very large N. Let $\mathrm{ELS}_N(X_0, X_1)$ denote the set of all exponential Lie series in X_0, X_1, truncated at N. Then $\mathrm{ELS}_N(X_0, X_1)$ is a nilpotent Lie group, whose Lie algebra is $\mathrm{Lie}_N(X_0, X_1)$, the set of all Lie polynomials in X_0, X_1, of degree not greater than N. The system

$$\dot{S} = S(X_0 + uX_1) \tag{51}$$

can be regarded as evolving in $\mathrm{ELS}_N(X_0, X_1)$. The trajectory of (51) for a given control u, starting at $S = 1$ when $t = 0$, is the curve

$$t \to [S_u(t)]_N,$$

where $[\cdot]_N$ denotes truncation at N.

So the accessible set from 1 is $\{\mathrm{Ser}_N(u)\colon u \in \mathscr{U}_1\}$. On the other hand, it is easy to see that (51) has the AP from 0 and so $\{\mathrm{Ser}_N(u)\colon u \in \mathscr{U}_1\}$ has a nonempty interior.

Pick a control $\bar{u}\colon [0, T] \to \boldsymbol{R}$ such that $P_1 = \mathrm{Ser}_N(\bar{u})$ is in the interior (relative to $\mathrm{ELS}_N(X_0, X_1)$) of the reachable set from 1. Now P_1 is the exponential of an element Z_1 of $\mathrm{Lie}_N(X_0, X_1)$. If we let Z_1' denote the result of replacing X_1 by $-X_1$ in Z_1, and if

$$P_1' = \exp(Z_1'),$$

then P_1' is also reachable from 1 and so $P_2 = P_1 P_1'$ is in the interior of the reachable set from 1.

Now we can write

$$Z_1 = (Z_1)_{\text{even}} + (Z_1)_{\text{odd}},$$

where $(Z_1)_{\text{even}}$, $(Z_1)_{\text{odd}}$ are the parts of Z_1 that involve, respectively, brackets with an even number of X_1's and brackets with an odd number. Then

$$Z_1' = (Z_1)_{\text{even}} - (Z_1)_{\text{odd}}.$$

Then

$$P_2 = \exp(Z_1)\exp(Z_1') = \exp(Z_2),$$

where

$$\begin{aligned} Z_2 &= Z_1 + Z_1' + \tfrac{1}{2}[Z_1, Z_1'] + \dots \\ &= 2(Z_1)_{\text{even}} + [(Z_1)_{\text{odd}}, (Z_1)_{\text{even}}] + \dots \end{aligned}$$

so that Z_2 is the sum of a part that is even in X_1, and a part which has total degree not less than 2. Repeating this procedure, one can produce $P_3, P_4, \ldots$ such that $P_i = \exp(Z_i)$, where Z_i is the sum of a part which is even in X_1 and a part which has total degree not less than i. Since $\mathrm{Lie}_N(X_0, X_1)$ is nilpotent, we conclude that there is a Q which is of the form $\mathrm{Ser}_N(u)$ for some $u \in \mathcal{U}_1$ and satisfies

$$Q = \exp(Y)$$

with Y even in X_1, and which belongs to the interior of the reachable set from 1.

Now suppose

$$B = (\mathrm{ad}\, X_{i_1}) \ldots (\mathrm{ad}\, X_{i_{r-1}})\, (X_{i_r}),$$

where each i_j is either 0 or 1. Assume that B is even in X_1. By the hypothesis, $B(f_0, f_1)\,(x_0)$ is equal to a linear combination

$$\Big(\sum_C a_{BC} C\Big)(f_0, f_1)\,(x_0)$$

where C runs over brackets with fewer X_1's than those that appear in B.

We can write

$$Y = \sum_B y_B B,$$

where the sum suns over brackets that are even in X_1. Let D be a Lie monomial in X_0, X_1, not necessarily even in X_1. Since $\exp Y$ is in the interior of the reachable set from 1, there is an $\varepsilon > 0$ such that the point

$$Q(\lambda, \eta) = \exp\big(Y(\lambda, \eta)\big),$$

where

$$Y(\lambda, \eta) = \lambda D + \sum_B y_B\Big(B - \sum_C \eta_{BC} a_{BC} C\Big),$$

is reachable from 1, whenever $\lambda \in \boldsymbol{R}$, $\eta = \{\eta_{BC}\}$, satisfy

$$|\lambda| < \varepsilon, \qquad \|\eta\| = \max\{|\eta_{BC}|:\ B, C\} < \varepsilon.$$

Now fix, for each λ, η for which $|\lambda| < \varepsilon$, $\|\eta\| < \varepsilon$, a control $u_{\lambda,\eta}\colon [0, T_{\lambda,\eta}] \to \boldsymbol{R}$ such that

$$\mathrm{Ser}_N(u_{\lambda,\eta}) = Q(\lambda, \eta).$$

Let r be a fixed positive integer. For $0 < t \leqslant 1$, $|\lambda| < \varepsilon$, $\|\eta\| < \varepsilon$, define

$$u^t_{\lambda,\eta}\colon\ [0, tT_{\lambda,\eta}] \to \boldsymbol{R}$$

by $u^t_{\lambda,\eta}(s) = t^r u_{\lambda,\eta}(s/t)$. Then

$$\mathrm{Ser}_N(u^t_{\lambda,\eta}) = \exp\Big(\lambda t^{\alpha(D)} D + \sum_B y_B\Big(t^{\alpha(B)} B - \sum_C t^{\alpha(C)} \eta_{BC} a_{BC} D\Big)\Big), \tag{52}$$

where, for each bracket C,

$$\delta(C) = r\mu(C) - \nu(C).$$

Here $\nu(C)$ is the total degree of C, and $\mu(C)$ is the degree in X_1.

Each term $t^{\delta(C)}\eta_{BC}a_{BC}C$ that appears in (52) corresponds to B, C such that

$$\mu(C) < \mu(B).$$

Since only a finite number of B's and C's are involved, it is possible to choose r such that $\delta(C) < \delta(B)$ for all the B, C that occur in (52), and that $\delta(D) \geqslant 0$.

Then we can choose

$$\eta_{BC}(t) = t^{\delta(B)-\delta(C)}, \quad \lambda(t) = t, \quad \eta(t) = \{\eta_{BC}(t)\}.$$

Then, if t is small enough, $|\lambda(t)| < \varepsilon$ and $\|\eta(t)\| < \varepsilon$. So $u^t_{\lambda(t),\eta(t)}$ is well defined, and

$$\mathrm{Ser}_N(u^t_{\lambda(t),\eta(t)}) = \exp\Big[t^{\delta(D)+1}D + \sum_B y_B t^{\delta(B)} R_B\Big], \tag{53}$$

where

$$R_B = B - \sum_C a_{BC}C.$$

If we expand the exponential in (53), we find that

$$\mathrm{Ser}_N(u^t_{\lambda(t),\eta(t)}) = 1 + t^{\delta(D)+1}D + H + o(t^{\delta(D)+1}),$$

where H is a linear combination of products $R_{B_1}\dots R_{B_s}$, $s \geqslant 1$. If we plug in f_0, f_1 for X_0, X_1 and evaluate at x_0, then H will vanish because all the $R_B(f_0, f_1)(x_0)$ vanish. So if φ: $M \to \boldsymbol{R}$ is smooth, we find

$$\varphi\big(\gamma(t)\big) = \varphi(x_0) + t^{\delta(D)+1}[D(f_0, f_1)\varphi](x_0) + o\big(t^{\delta(D)+1}\big) + o(T_t^{N+1}),$$

where

$$T_t = T_{u^t_{\lambda(t),\eta(t)}}$$

and

$$\gamma(t) = \pi(u^t_{\lambda(t),\eta(t)}, x_0, T_t).$$

Now if N is sufficiently large, the $o(T_t^{N+1})$ will be an $o(t^{\delta(D)+1})$ as well. Then the tangent vector to the curve $\sigma \to \gamma(\sigma^{1/1+\delta(D)})$ at $\sigma = 0$ is $D(f_0, f_1)(x_0)$.

But $D(f_0, f_1)$ was an arbitrary bracket of f_0, f_1, and these brackets span the space of tangent directions to x_0 because of the rank condition. So we can build control variations in all directions, and therefore it is possible to reach, from x_0, a full neighborhood of x_0.

This completes our sketch of the proof of Theorem H. The details will be published elsewhere (Sussmann [11]).

§ 8. Concluding remarks

Lack of space prevents us from surveying other applications of Lie brackets and real analyticity. In particular, we have not discussed observability and realization theory. A theory of observability of nonlinear systems has been developed, as well as a theory of minimal realizations (cf. Sussmann [7], [8]). Also, Fliess has developed a local theory where, to each system with an observation $y = h(x)$, one associates a formal power series in noncommutative indeterminates. In [3] Fliess has found a necessary and sufficient condition for a series to arise from an analytic system.

A deep development in the theory of real analytic functions is the theory of analytic stratifications and of analytic, semianalytic, and subanalytic sets. One application of this theory to a system-theoretic problem was given in [9], where we proved that for an analytic system there exists a ***universal input***, that is, an input u with the property that, whenever two states produce different outputs for some input v, then they produce different outputs for u.

The hardest (and possibly the deepest) use of real analyticity in control theory is the work (begun by Brunovsky in [1]) on the existence of regular synthesis. This work makes use of the theory of subanalytic sets. Due to the length and the technical complexity of the proofs, a detailed account of the results is not yet available. However, a long paper by Brunovsky and this author, now in preparation, will, we hope, fill this gap.

References

[1] P. Brunovsky, *Every normal linear system has a regular time-optimal synthesis*, Math. Slovaca 28 (1978), 81–100.

[2] —, *On the structure of optimal feedback systems*, in: Proc. Int. Congress of Mathematicians, Helsinki 1978, 841–846.

[3] M. Fliess, *Realizations of nonlinear systems and abstract transitive Lie algebras*, Bull. Amer. Math. Soc. 2, 3 (1980), 444–446.

[4] A. Krener, *A generalization of Chow's theorem and the bang-bang theorem to nonlinear control problems*, SIAM J. Control Optimization 12 (1974), 398–404.

[5] T. Nagano, *Linear differential systems with singularities and an application to transitive Lie algebras*, J. Math. Soc. Japan 18 (1966), 398–404.

[6] H. J. Sussmann, *Orbits of families of vector fields and integrability of distributions*, Trans. Amer. Math. Soc. 180 (1973), 171–188.

[7] —, *Existence and uniqueness of minimal realizations of nonlinear systems*, Math. Systems Theory 10 (1977), 263–284.

[8] —, *A generalization of the closed subgroup theorem to quotients of arbitrary manifolds*, J. Differential Geometry 10 (1975), 151–166.

[9] H. J. Sussmann, *Single-input observability of continuous-time systems*, Math. Systems Theory 12 (1979), 371–393.

[10] —, *A bang-bang theorem with bounds on the number of switchings*, SIAM J. Control Optimization 17, 5 (1979), 629–651.

[11] —, *Formal exponential Lie series and local controllability*, to appear.

[12] —, *Analytic stratifications and control theory*, in: Proc. Int. Congress of Mathematicians, Helsinki 1978, 865–871.