Thirty years of Optimal Control: was the path unique?[†]

H. J. SUSSMANN*, Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903.

1 INTRODUCTION

In this note I intend to take seriously the title of this conference —"Thirty years of modern optimal control"— and regard it as an opportunity to speculate on one aspect of the past and present of our discipline. Specifically, I want to look at the evolution of our field, regarded as a path in the space of conceptual structures, and at the point in this space where we find ourselves now, and I want to ask to what extent the path that has been followed was the one that we were *bound* to follow, and to what extent it is a special path, perhaps a product of peculiar initial conditions and historical circumstances. In asking this question —and arguing for a particular answer— I will proceed in a spirit similar to D. Ruelle's recent Gibbs lecture (cf. [17]). Ruelle asked: "*How natural is our human 20th-century mathematics?* ... How much does the existing construct depend on human nature and condition?" This is more or less equivalent to asking whether, if we were to establish contact with "little green men from outer space, ... [they] would have a similar central core for their mathematics; ... in their own language they would have a theorem saying that the image of a compact set by a continuous map is compact."

A "yes" answer to this question essentially means that mathematical concepts force themselves upon us by virtue of their internal logical structure and the logical relations among them, that the evolution of mathematical knowledge follows a "natural" path which is essentially determined by this structure. A "no" answer means that, just because mathematics has developed in a particular way, it does not follow that this is

† This paper has appeared in the book *Modern Optimal Control* (Emilio O. Roxin Ed.), Marcel Dekker Inc., New York-Basel (1989), pp. 359-375. The book contains the proceedings of the conference "Thirty years of Optimal Control," held in Kingston, R.I., in June 1988.

* Partially supported by NSF grant DMS83-01678-01, and by the CAIP Center, Rutgers University, with funds provided by the New Jersey Commission on Science and Technology and by CAIP's industrial members.

how it inexorably had to develop. It also means that the development *may have taken a wrong turn*, in that important concepts could have been missed and unimportant ones may have become the focus of everybody's attention.

For the specific case of Optimal Control Theory, I want to argue that there is a set of mathematical concepts, such as *coordinate-free properties, Lie brackets, real analyticity, integral manifolds, Lie algebras of vector fields*, that most civilizations of "little green men from outer space" would probably have come across during the very earliest stages of their development of the field, and immediately incorporated into the central core of their Control Theory, which would, for that reason, look quite different from ours. In most cases, the little green men would have been able to borrow these concepts from other areas of their mathematics, where they would have already been discovered (invented?) for different purposes. But, even in those planets where other mathematicians failed to hit upon the Lie bracket, the control theorists would have discovered it, exactly as on Earth we control theorists would have discovered the variations-of-constants formula if it hadn't been known before.

This is not to deny that the importance of these ideas has also been recognized on Earth, as shown by the tremendous amount of activity in Optimal Control in recent years that makes systematic use of them. However, in spite of this activity, these concepts are still often regarded by other control theorists as belonging to a somewhat esoteric fringe of our field called "differential geometric methods in control," or "differential geometric control theory." (I believe this name to be inappropriate, because *Differential Geometry* usually starts when one studies differentiable manifolds *with some extra structure*, such as a Riemann or Lorenz metric, or at least a connection, and this is seldom what one encounters in Optimal Control.) The implication of such terminology is the belief that this esoteric subfield somehow asks its own esoteric questions and then introduces esoteric concepts in order to solve them. Lie brackets presumably occur in Control Theory because one asks questions about Lie brackets, subanalytic sets occur because one wants to know whether certain reachable sets are subanalytic, etc. I will argue that this belief is false. My main thesis is that *the concepts that I listed above are forced upon almost anyone who asks the most basic, most natural questions in Optimal Control.* In particular, I will mostly focus on the concept of the *Lie bracket*, and argue that it is as entitled to belonging to the mainstream of our field as, say, concepts that we all accept as basic, such as the exponential of a matrix, Lagrange multipliers, convexity, or weak* convergence. (Those readers who think this is so obvious and well known that it does not need proof should look at the large number of mathematical books on Optimal Control where *the Lie bracket is not mentioned.* Cf., e.g., [1], [11], [12], [13], [16], [37].)

2 AN ALTERNATIVE SCENARIO: EXTRATERRESTRIAL OPTIMAL CONTROL

Our little green men's thoughts might have evolved as follows. The mathematical object that describes the evolution of a physical system is an *ordinary differential equation*, i.e. a *vector field* or *dynamical system.* When such a system evolves, its future path is uniquely determined by its present state. Suppose now that we want to study systems whose evolution can be affected by a controlling agent, i.e. we want to incorporate the element of *choice*. We should then look at systems of two or more vector fields. If, as is natural, we want to start by studying the simplest possible case, then it clear that we have to begin our development of control theory by studying the systems arising from *pairs of vector fields*. (We shall refer to these as *two-vector-field systems*.)

What questions would the little green men ask about such pairs? If X and Y are our two vector fields, then a *control* is clearly specified by giving a time interval $[a, b]$, and specifying the subset S_X of $[a, b]$ during which we follow X (or, equivalently, the subset $S_Y = [a, b] - S_X$ during which we follow Y). An obvious question is to what extent the point q that we get at the end —starting from a point p— depends on the set S_X. For instance, it might be that all we need to know in order to know q is *how much time was spent controlling in each of the two directions*, i.e. *the Lebesgue measure of the sets* S_X, S_Y. Or perhaps we need more information, i.e. it is not enough to know how much X-control and how much Y-control we have applied, and we also need to know the *order* in which these two controls were applied. So let us consider the most extreme situation, and compare two controls for which the sets S_X and S_Y have exactly the same size, but such that, in one case, all the X-controlling is done first, whereas in the other case the Y-controlling is done first. For these two controls, let us ask whether the point q will be the same.

To formulate this mathematically, let us introduce some notations. For a smooth (i.e. C^∞) vector field X on $\mathbb{R}^n$ (or on an open subset Ω of $\mathbb{R}^n$), let us use exponential notation for the flow, so that $t \to e^{tX}p$ denotes the integral curve of X that goes through p at time 0. If X and Y are two smooth vector fields, then the question we have asked above is, simply, whether the identity

$$e^{tX}e^{\tau Y}p = e^{\tau Y}e^{tX}p \tag{2.1}$$

holds for all choices of t, τ and p. If this happens, we say that the flows of X and Y *commute*. To be precise, we have to take into account the possibility that the integral curves of X or Y may fail to be defined for all times, i.e. that the flows of X or Y may exhibit explosions. The precise definition of commuting flows is as follows. We say that X and Y *have commuting flows* if for every $p \in \Omega$ there is an $\varepsilon > 0$ such that 2.1 holds whenever $t \in \mathbb{R}$, $\tau \in \mathbb{R}$, $|t| < \varepsilon$, $\tau < \varepsilon$.

One can then show that,

A. *If X and Y have commuting flows, then*

- **a.** *ε can be chosen to be independent of p as long as p stays in a compact set, and*
- **b.** *2.1 actually holds for all t and τ, with no restrictions, if both X and Y are complete.*

(Recall that a vector field X is *complete* if the integral trajectories of X are defined for all times.)

Our question now is how we can tell whether X and Y have commuting flows. But, before we answer this, it is worth showing that commutation has many interesting consequences. To begin with, we can broaden our analysis and study the two-vector-field system

$$\dot{x} \in \mathrm{co}(X(x), Y(x)) , \tag{2.2}$$

i.e.

$$\dot{x} = vX(x) + (1-v)Y(x) , \ 0 \leq v \leq 1 \tag{2.3}$$

or, equivalently (if we write $f = \frac{1}{2}(X+Y)$, $g = \frac{1}{2}(Y-X)$ and $u = 1-2v$),

$$\dot{x} = f(x) + ug(x) , \ -1 \leq u \leq 1 , \tag{2.4}$$

where the controls are required to be measurable, but we allow the values of u to be not just 1 and -1 but an arbitrary number in $[-1,1]$. (Our little green men would presumably have noticed, as we have on Earth, that it is always nice to have a closed reachable set, so that chattering controls should be permitted, and in particular the control set $\{-1,1\}$ might as well be replaced by the interval $[-1,1]$.)

Assume now —for the moment— that X and Y are complete. Then the assertion that the flows of X and Y commute is easily seen to be exactly equivalent to the statement that, if $\gamma : [a,b] \to \Omega$ is a bang-bang trajectory corresponding to a control $\bar{u}(\cdot) : [a,b] \to [-1,1]$, and α, β are, respectively, the total amount of time during which the control u is equal to 1 and -1 (i.e $\alpha = |\{t : \bar{u}(t) = 1\}|$ and $\beta = b - a - \alpha$, where $|\ldots|$ denotes Lebesgue measure) then

$$\gamma(b) = \mathrm{e}^{\alpha Y} \mathrm{e}^{\beta X} \gamma(a) . \tag{2.5}$$

Since $\alpha = \frac{1}{2}(b - a + \int_a^b \bar{u}(t)\, dt)$, we can rewrite 2.5 as

$$\gamma(b) = \mathrm{e}^{\frac{1}{2}(b-a+\int_a^b \bar{u}(t)\, dt)Y} \mathrm{e}^{\frac{1}{2}(b-a-\int_a^b \bar{u}(t)\, dt)X} \gamma(a) . \tag{2.6}$$

It then follows easily from the continuity of the integral as a functional of the control with respect to the weak topology, together with the fact that, in this topology, the bang-bang controls are dense in the class of all controls, that Formula 2.6 actually holds for an arbitrary control. What this says is that, when the flows commute, then all that matters is how much control you apply in each direction, and the order in which these controls are applied makes no difference. In particular, *the point $\gamma(b)$ to which a control $\bar{u}(\cdot)$ steers the initial point $p = \gamma(a)$ depends on $\bar{u}(\cdot)$ through two parameters only, namely, the total time and the integral of the control.* Formula 2.5 implies that $\mathcal{R}(p) \subseteq S$, where $S = \{\mathrm{e}^{\tau Y}\mathrm{e}^{tX}p : t \in \mathbb{R}, \tau \in \mathbb{R}, t \geq 0, \tau \geq 0\}$, so that S is the image of a subset of the plane under a smooth map. Therefore

B. *If X and Y have commuting flows, then the reachable set $\mathcal{R}(p)$ from a point p is at most two-dimensional.*

The precise meaning of "at most two dimensional" is that $\mathcal{R}(p) \subseteq S$, where S is the image of a two-dimensional manifold under a smooth map. With this interpretation we have actually established **B** when X and Y are complete. However, with a little bit more work one can generalize this to the noncomplete case, and **B** is still valid, and in fact the result is even stronger, because we can interpret "$\mathcal{R}(p)$ is at most two dimensional" to mean "$\mathcal{R}(p)$ is a subset of a two-dimensional connected submanifold of Ω."

Also, if we assume that X and Y are linearly independent at a point p, then the map $(\tau, t) \to \mathrm{e}^{\tau Y}\mathrm{e}^{t\mathrm{X}}p$ is one-to-one for small enough τ, t, by the Implicit Function Theorem. This implies, in particular, that α and β can be recovered from a knowledge of $q = \mathrm{e}^{\alpha Y}\mathrm{e}^{\beta X}p = \gamma(b)$. Since $b - a = \alpha + \beta$, it follows that the time $b - a$ can be recovered from the knowledge of q, i.e. *all trajectories going from p to q take exactly the same time*, at least as long as this time is small. In particular, this means that

(#) *there is an $\varepsilon > 0$ such that all trajectories that start at p and take time $\leq \varepsilon$ are time-optimal.*

If $X(p)$ and $Y(p)$ are linearly independent, then Condition (#) is equivalent to

(##) *there exists a neighborhood N of p such that every trajectory through p which is entirely contained in N is time-optimal.*

The implication (##) $\Rightarrow$ (#) is trivial. To show the converse, pick a smooth function φ such that the derivatives of φ in the X- and Y- direction are > 0 at p. (This is possible because of the linear independence.) Then we can pick a neighborhood N' such that the X- and Y-derivatives of φ are $> c$ throughout N', where c is some positive constant. Let $N = N' \cap \{q : \varphi(q) < \varphi(p) + c\varepsilon\}$. Then every trajectory through p that is contained in N must take time $\leq \varepsilon$.

So we have established:

C. *if X and Y have commuting flows and $X(p)$, $Y(p)$ are linearly independent, then the minimum time problem associated with 2.4 is degenerate at p,*

where we define a minimum time control problem to be *degenerate* at a point p if (##) holds, and to be *degenerate* if it is degenerate at every p. In order words, except for the minor detail that degeneracy, as defined, only refers to "small time trajectories," a degenerate system is one for which all trajectories are time-optimal. A minimum-time problem for such a system is such that all trajectories that satisfy the constraints of interest are solutions. Clearly, the question whether a particular system is degenerate is the most basic one we can pose about the minimum time problem, in the same way as, for the problem of minimizing a function, the first basic question one should ask is whether the function is a constant. Any theorem that enables us to characterize degenerate problems will be about as basic as the Calculus theorem that says that a function on an open connected set is a constant if and only if it is everywhere differentiable and the gradient vanishes. Such a characterization will be described later and, not surprisingly, Lie brackets will be involved.

The hypothesis that $X(p)$ and $Y(p)$ are linearly independent is crucial, as can be seen by considering the system $\dot{x} = u$, $x \in \mathbb{R}$, $|u| \leq 1$. The restriction to small-time trajectories is also unavoidable, even if X and Y are complete. (This can easily be seen from the following example. Let $\Omega = \mathbb{R}^2 - \{(0,0)\}$, and let X, Y have components $(y, -x)$, (x, y), respectively. Then the trajectories of X are periodic with period 2π. Also, it is clear that X and Y have commuting flows. But a trajectory of X that goes from a point p back to p in time 2π is not time-optimal.)

Finally, one can strengthen **C** to

D. *If X and Y have commuting flows, $K \subseteq \Omega$ is compact, and $X(p)$ and $Y(p)$ are linearly independent for every $p \in K$, then there is an $\varepsilon > 0$ such that all trajectories that start at a point $p \in K$ and take time $\leq \varepsilon$ are time-optimal.*

3 THE LIE BRACKET

Now that our little green men have established that the commutation of the flows is of paramount control-theoretic significance, they should naturally turn to the question of how to give an *infinitesimal characterization* of this property. If we write X and Y as *column-vector-valued functions*, we can define the *Lie bracket* of X and Y to be the vector field $[X, Y]$ given by

$$[X, Y] = DY \cdot X - DX \cdot Y \ , \tag{3.1}$$

where DX is the square-matrix-valued dunction whose rows are the gradients of the components of X.

It then turns out that

E. *X and Y have commuting flows if and only if $[X, Y] \equiv 0$.*

More generally, using a Taylor expansion up to order two, one can prove the following *geometric characterization of the Lie bracket*:

$$[X, Y](p) = \lim_{t \to 0+\,,\, \tau \to 0+} \frac{1}{t\tau} \left[\mathrm{e}^{\tau Y} \mathrm{e}^{tX} \mathrm{e}^{-\tau Y} \mathrm{e}^{-tX} p - p \right] . \tag{3.2}$$

From 3.2 one can easily derive the "if" part of **E**. To derive the "only if" part, one shows the following result, which is important in its own right:

F. *if X and Y satisfy $[X, Y] \equiv 0$ on some neighborhood of p, and the vectors $X(p)$, $Y(p)$ are linearly independent, then there exists a smooth change of coordinates near p such that, in the new coordinates, X and Y are constant vector fields.*

Since constant vector fields obviously have commuting flows, the "if" part of **E** follows from **F**, at least near points p where X and Y are independent. The general case can be reduced to this one by the following procedure. Work in $\Omega \times \mathbb{R} \times \mathbb{R}$, i.e. add two new variables t and τ. Let $\tilde{X}(x, t, \tau)$ have components $(X(x), 1, 0)$. Similarly, let $\tilde{Y}$ have components $(Y(x), 0, 1)$. Then one can easily verify that $[\tilde{X}, \tilde{Y}] \equiv 0$ if $[X, Y] \equiv 0$. Moreover, if $\pi : \Omega \times \mathbb{R} \times \mathbb{R} \to \Omega$ denotes the projection, then $\pi(\mathrm{e}^{t\tilde{X}}(p, s, \sigma)) = \mathrm{e}^{tX}(p)$, and similarly for Y. Therefore, if $\tilde{X}$ and $\tilde{Y}$ have commuting flows, the same is true for X and Y. This establishes **E** in general.

The Lie bracket of two vector fields therefore has the advantage of being an object that can be *explicitly computed from the components of the vector fields by means of algebraic operations* (without having to solve differential equations), and yields information about a property of the solutions. So it is a natural object to associate to a pair of vector fields.

4 LOCAL CLASSIFICATION AND COORDINATE FREE OBJECTS

While some of our little green men are busy asking some specific questions of interest about two-vector-field systems, other may take a broader point of view and ask for a *classification theory* of such pairs. Following the model of single vector fields, one may

(a) look first for a *local classification.* i.e. a classification of pairs of vector fields on a neighborhood of a given point p,

(b) try to start by looking at the least degenerate cases, and progress towards the more degenerate ones,

(c) wonder what *group* of transformations should be used in the classification theory,

(d) look for *canonical forms.*

The point of *(c)* is that, as is the case elsewhere in Mathematics, to each problem there is attached a group, namely, the group of transformations that preserve the properties of interest for the problem. It is then often convenient, for a particular problem, to look for a transformation in the group that will put the problem in a simpler form, from which it might be easy to see the solution. In our case, if we ask questions such as the ones discussed above (whether the minimum time control problem is degenerate, or whether the point q just depends on α and β as above), or many other related questions (such as whether the reachable sets have nonvoid interior, or whether all time-optimal trajectories are bang-bang, or whether the reachable set from a point p contains a neighborhood of p, or whether there is a smooth feedback control defined near p that makes p an asymptotically stable equilibrium, or whether the boundary of the reachable sets is a finite union of submanifolds), *the appropriate group is that of smooth tranformations of the state space*, variously referred to as *diffeomorphisms* or *coordinate changes*, depending on the point of view. So the classification theory one wants to develop is the local classification of pairs of vector fields under *smooth coordinate changes.*

For the case of a *single* vector field X, the "least degenerate" case is that of a point p that is not an equilibrium, i.e. a p such that $X(p) \neq 0$. The classification problem for this case is solved by the most basic result in the qualitative theory of ordinary differential equations, namely, the existence of a "flow box":

G. *if X is a smooth vector field near p, and $X(p) \neq 0$, then there exists a smooth change of coordinates on a neighborhood of p such that in the new coordinates X becomes a constant vector field.*

(Notice that, by making a further *linear* change of coordinates we could equally well have stated the conclusion of **G** to be that, in the new coordinates, $X = (1, 0, \ldots, 0)$.)

In the qualitative theory of differential equations, one then proceeds to the "degenerate case" when $X(p) = 0$, and within that one studies first the "nondegenerate case" of a *hyperbolic equilibrium* (obtaining the theory of *stable and unstable manifolds*), and then the more degenerate case when the linearization has some pure imaginary eigenvalues (so that one has to study *center manifolds*).

For our two-vector-field systems, Statement **F** is none other than the extension to the case of two vector fields of the flow-box theorem **G**. However, in this case the simultaneous "flow-boxing" of two vector fields is not possible unless their Lie bracket vanishes. Hence *the Lie bracket is the obstruction that prevents simultaneous flow-boxing.* When the Lie bracket does not vanish, *there is no classification theorem analogous to* **G***, even in the "nonequilibrium" case.* (Naturally, the two-vector-field analogue of "nonequilibrium" is the case when the vector fields are linearly independent at a point p.) So, while the problem of the local classification of ordinary differential equation is completely solved by the flow-box theorem, and therefore does not require any further analysis, *the analogous problem for pairs of vector fields is very difficult and far from solved.*

5 RESULTS INVOLVING LIE BRACKETS

So far, our little green men have stumbled upon the Lie bracket, but most of our discussion has not made use of all the information it contains. We have only focused on the implications of the *vanishing* of the Lie bracket. However, once we have discovered this important object, it is possible to derive immediately a plethora of results that solve basic control-theoretic problems in terms of properties of Lie brackets. We list just a few of them, without proof.

First, let us ask the question when the minimum time problem for a two-vector-field system is degenerate. As explained above, the condition that the Lie bracket vanishes is *sufficient* for this. It's not hard to see that this is too strong a condition. The appropriate necessary and sufficient condition turns out to involve Lie brackets as well. Moreover, this condition turns out to be exactly the same for general control systems, i.e. systems involving more than two vector fields. To state the condition, let us first use $\mathcal{V}(\Omega)$ to denote the class of all C^∞ vector fields on Ω. If we define addition and multiplication by scalars in the obvious way, and then let $[\ldots,\ldots]$ be the Lie bracket as defined above, it turns out that $\mathcal{V}(\Omega)$ is a *Lie algebra.* (In general, a Lie algebra is a linear space V endowed with a bilinear map $V \times V \ni (x, y) \to [x, y] \in V$ such that $[x, y] = -[y, x]$ and $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for all $x, y, z \in V$.) If $\mathcal{F} \subseteq \mathcal{V}(\Omega)$ is a set of smooth vector fields, let us use $\Lambda(\mathcal{F})$ to denote the smallest Lie subalgebra of $\mathcal{V}(\Omega)$ that contains $\mathcal{F}$. Also, let us use $\Lambda_0(\mathcal{F})$ to denote the smallest ideal of $\Lambda(\mathcal{F})$ that contains all the differences $X - Y$, $X \in \mathcal{F}$, $Y \in \mathcal{F}$. (A *Lie subalgebra* of a Lie algebra L is a linear subspace of L that is closed under the map $[\ldots,\ldots]$, and an *ideal* is a linear subspace I of L such that $x \in L$, $y \in I$ imply $[x, y] \in I$.) Also, we let $\mathcal{F}(p) = \{X(p) : X \in \mathcal{F}\}$. Call $\mathcal{F}$ *infinitesimally degenerate* at a point p if $\Lambda_0(p) \neq \Lambda(p)$. (One can easily see that

$\Lambda_0(p)$ is a subspace of $\Lambda(p)$ of codimension 0 or 1, so infinitesimal degeneracy is the condition that the codimension is 1.)

To a control system $\Sigma_f : \dot{x} = f(x,u)$, $u \in U$, in which f is assumed to be of class C^∞ in x for each fixed $u \in U$, we can associate the family $\mathcal{F}(f) = \{f(\cdot,u) : u \in U\}$, i.e. the set of all vector fields that correspond to all the possible constant values of u.

Let us call a control problem Σ_f *analytic* if all the vector fields $f(\cdot,u)$ are real analytic. We then have the following (cf. [24]):

Theorem 5.1. *An analytic control problem Σ_f is degenerate at a point p if and only if the corresponding family $\mathcal{F}(f)$ is infinitesimally degenerate at p.*

So, at least for the analytic case, we have a complete characterization of degeneracy in terms of Lie brackets. To understand what this means, and how this generalizes **C**, let us return to the two-vector-field case. In this case, we can write all the *iterated* brackets, i.e. f, g, $[f,g]$, $[f,[f,g]]$, $[g,[f,g]]$, $[f,[f,[f,g]]]$, $[g,[f,[f,g]]]$, $[[f,g],[f,[f,g]]]$, ... etc. Degeneracy will occur, in this case, if and only if, for each point p, the vector $f(p)$ does not belong to the linear hull of the values at p of all the other brackets. In particular, this includes **C** because $[X,Y] = 2[f,g]$, so that, if the hypothesis of **C** holds, then the only brackets that might be $\neq 0$ are f and g. Therefore, if $f(p)$ and $g(p)$ are independent, then $f(p)$ cannot belong to linear hull of the "other" brackets, and Theorem 5.1 applies.

Now let us turn to the problem of understanding what the reachable sets look like. Let $\mathcal{R}_f(p)$ denote the reachable set from p for the system Σ_f, and let $\mathcal{R}_f^T$, $\mathcal{R}_f^{\leq T}$ denote, respectively, the time T and the time $\leq T$ reachable sets. We could ask whether these sets are smooth submanifolds, but trivial examples show that this need not be so. The next best thing is if the sets are "close" to being submanifolds in some sense. For our purposes, let us call a subset S of a smooth manifold M *thick in M* if $S \subseteq \operatorname{Clos}\operatorname{Int} S$, where both the closure and the interior are taken relative to M. (This implies, among other things, that the set has *integer dimension* in any reasonable dimension theory, and cannot consist of a piece of some dimension together with some smaller dimensional "tail" sticking out. A convex set is thick in its affine hull.) Let us call a smooth submanifold M of Ω an *integral manifold* of a set $\mathcal{F}$ of smooth vector fields if M is connected and the tangent space T_xM of M at x equals $\mathcal{F}(x)$ for every $x \in M$. A *maximal* integral manifold is one that is maximal with respect to inclusion. We then have (cf. [6], [15]):

Theorem 5.2. *If $\mathcal{F}$ is a Lie algebra of real-analytic vector fields on Ω, then for every $p \in \Omega$ there exists a unique maximal integral manifold M of $\mathcal{F}$ such that $p \in M$.*

If we use $I(\mathcal{F},p)$ to denote the maximal integral manifold of $\mathcal{F}$ through p, then the following holds (cf. [24]):

Theorem 5.3. *For an analytic system Σ_f, the sets $\mathcal{R}(p)$, $\mathcal{R}_f^{\leq T}(p)$ (for $T > 0$) are thick subsets of the integral manifold $I(\Lambda(\mathcal{F}(f)),p)$. If the vector fields in $\mathcal{F}(f)$ are complete (or, more generally, if $\mathcal{R}_f^T(p)$ is connected), then $\mathcal{R}_f^T(p)$ is a thick subset of the submanifold $I(\Lambda_0(\mathcal{F}(f)),q)$, if q is any point of $\mathcal{R}_f^T(p)$.*

In particular, Theorem 5.3 can be used to give a Lie bracket characterization of the *accessibility property* (AP) from a point p, i.e. the property that the interior of $\mathcal{R}_f(p)$ is nonvoid (cf. [24]):

Theorem 5.4. *For an analytic system Σ_f, on Ω, with Ω open in $\mathbb{R}^n$, the AP holds from a point p iff $\Lambda(p) = \mathbb{R}^n$.*

One may then ask whether, if the AP from p holds, one might not be able to conclude something stronger, namely, that not only is the interior of $\mathcal{R}_f(p)$ nonvoid, but the point p itself is in the interior. (This property is called *local controllability from p.*) This seems to be a very difficult problem, and no complete answer is known, but a number of results have been proved. Let us limit ourselves to systems of the form

$$\dot{x} = f(x) + \sum_{i=1}^{m} u_i g_i(x) \ , \ |u_i| \leq 1 \ , \ i = 1, \ldots, m \ . \tag{5.1}$$

One can then see that the obstruction to local controllability comes from brackets such as $[g_i, g_j]$, $[g_i, [f, g_j]]$, $[[g_i, [f, g_j]], [g_r, g_s]]$, etc., i.e. iterated brackets with an even number of g's. It is then easy to prove:

Theorem 5.5. *Consider a system 5.1 such that the vector fields f, $g_1, \ldots, g_m$ are real analytic. Assume that the AP from p holds. Assume that all the iterated brackets with an even number of g's vanish at p. Then, for every $T > 0$, p is an interior point of the time T reachable set from p.*

(Notice that f itself is a bracket with 0 g's, and 0 is even, so the hypothesis of Theorem 5.5 implies in particular that $f(p) = 0$, i.e. p is an equilibrium point. In particular, the time T and time $\leq T$ reachable sets from p coincide.)

It turns out that the vanishing of the brackets with an even number of g's is a very interesting property. Actually, one can show that these systems can be put in a very useful canonical form by changing coordinates.

Theorem 5.6. *If a system 5.1 satisfies the hypotheses of Theorem 5.5, then one can make a real analytic change of coordinates on a neighborhood of p, that sends p to 0, such that in the new coordinates the system has the form*

$$\dot{x} = Ax + \sum_{i=1}^{m} u_i b_i \ , \tag{5.2}$$

where A is a square matrix and the b_i are vectors.

Such systems are called *linear systems.* For our little green men, a linear system is a "canonical form," under nonlinear changes of coordinates, of a general system for which the brackets with an even number of g's happen to vanish.

Theorem 5.5 is far from being the best one can prove. A lot of effort has recently been devoted to proving more general sufficient conditions (cf. [2], [5], [7], [8], [9], [10], [21], [22],[23], [25], [27], [29], [33], [36]). The common feature of all these results is that they involve some kind of *bracket neutralization condition.* That is, rather than require that certain brackets be zero, all that one asks is that these brackets be "neutralized," by being expressible, at p, as linear combinations of brackets that are "smaller" in some sense. For instance, one can prove (cf. [27]):

Theorem 5.7. *Consider a system 5.1 such that the vector fields f, $g_1, \ldots, g_m$ are real analytic. Assume that the AP from p holds. Assume that every iterated bracket with an even number of g's is equal, at p, to a linear combination of brackets with fewer g's. Then, for every $T > 0$, p is an interior point of the time T reachable set from p.*

Another interesting control problem, that the little green men would have rightly regarded as one of the most natural questions to be asked at the very beginning of the development of Control Theory, is the question of the *structure of optimal trajectories and optimal controls*, i.e. *what do time-optimal controls look like?* Assuming that our little green men have discovered the Pontryagin Maximum Principle, let us explore how this result can be used to obtain information about optimal controls. For simplicity, let us return now to two-vector-field systems, written in the form 2.4. To study optimal trajectories, we apply the Maximum Principle and end up concluding that one has to study the *switching function* $t \to \varphi(t)$ corresponding to an extremal γ, where

$$\varphi(t) = \Big\langle \lambda(t), g\big(\gamma(t)\big) \Big\rangle \,, \tag{5.3}$$

and λ is the adjoint variable. It then follows that the optimal control will be equal to 1 whenever $\varphi(t) < 0$, and to -1 when $\varphi(t) > 0$. This appears to suggest that optimal controls will tend to be "bang-bang," in the sense that their values will be 1 or -1. The switchings will occur at the zeros of the switching function. "Singular trajectories" can occur also, provided we manage to make φ vanish identically on an interval. And we should not forget about the possibility of even worse pathology, e.g. a switching function whose set of zeros is a Cantor-like set of positive measure.

In any case, the simplest possibility is for the zeros of the switching function to be isolated. Whenever this happens, we will have a "bang-bang theorem." The easiest way to make sure that the zeros of φ are isolated is to differentiate φ. If we can show that $\varphi(t)$ and $\dot{\varphi}(t)$ never vanish simultaneously, then the zeros of φ will be isolated, as desired. If we compute the derivative of the switching function we find ... a Lie bracket once again! Precisely, we get

$$\dot{\varphi}(t) = \Big\langle \lambda(t), [f,g]\big(\gamma(t)\big) \Big\rangle \,. \tag{5.4}$$

This already makes it possible to get *something*. If we are *in the plane*, and the vectors $g(x)$ and $[f,g](x)$ are linearly independent at every point x, then the conditions $\varphi(t) = 0$ and $\dot{\varphi}(t) = 0$ can never hold simultaneously, because $\lambda(t) \neq 0$. So we have:

Theorem 5.8. *For a system 2.4 in the plane, if $g(x)$ and $[f,g](x)$ are linearly independent for every x, then every time-optimal trajectory is bang-bang.*

If we want to prove a similar result in higher dimensions, then we have to keep differentiating the switching function. Unfortunately, we obtain

$$\frac{d^2\varphi}{dt^2}(t) = \Big\langle \lambda(t), \big([f,[f,g]] + u(t)[g,[f,g]]\big)\big(\gamma(t)\big) \Big\rangle \,, \tag{5.5}$$

which has the drawback that the control appears explicitly. The easiest way to get around this problem is to assume that the coefficient of the control —which happens to be a bracket with an even number of g's— vanishes identically. If we keep on differentiating, we see that the simplest case to study is that when all brackets with an even number of g's vanish. In that case, the successive derivatives of the switching function are of the form

$$\frac{d^j \varphi}{dt^j} = \left\langle \lambda(t), \mathrm{ad}_f^j(g)\big(\gamma(t)\big) \right\rangle , \tag{5.6}$$

where ad_f is the operator $h \to [f, h]$. Eventually, one of the $\mathrm{ad}_f^j(g)$ will be a linear combination of the preceding ones, and therefore φ will satisfy a linear homogeneous differential equation (possibly with nonconstant coefficients). This implies, in particular, that its zeros are isolated, so we get bang-bang trajectories once again. The conclusion is none other than the *bang-bang theorem for linear systems.* However, for our little green men this is just the simplest, crudest result that you can prove by just assuming that every bracket that bothers you actually vanishes. This, of course, is too much. As in the case of local controllability, one can see that it is not necessary for all the offending brackets to vanish. Using this, one can prove genuinely *nonlinear* bang-bang theorems (cf. [26], [18], [3]).

Clearly, bang-bang theorems are only the beginning of the study of the general problem of the structure of optimal controls. If, as one can easily see from trivial examples, time-optimal controls are not always bang-bang, then what can be said about them? Are they, for instance, finite concatenations of "bang-bang" and "singular" arcs? Can they have a more complicated structure involving, e.g., an uncountable set of points of discontinuity?

It turns out that, for this question, the Lie brackets play a fundamental role as well. Moreover, *the difference between C^∞ and real-analytic systems is crucial.* To see this, let us observe first that, in order to study these questions, the obvious thing to do is to apply the Pontryagin Maximum Principle, or whatever other necessary conditions for optimality might be available, in order to prove something about our optimal controls. As we have seen in our discussion of bang-bang theorems, the Lie brackets show up as soon as you write down what the Maximum Principle gives, and try to use it. To understand more complicated possibilities, let us go back to the case of systems in the plane. As explained above, for such systems the time-optimal trajectories are bang-bang, *as long as we stay away from the set* $S = \{x : \Delta(x) = 0\}$, where Δ is the determinant of the matrix whose columns are g and $[f, g]$. To understand "singular" extremals one needs to understand the set S better. For general C^∞ systems one cannot say much about this set, other than the not very useful fact that it is closed. For real-analytic systems, the picture is much more interesting. The set S is the zero set of a *real analytic function* and therefore it has a very special structure. There are two possibilities (assuming that Ω is connected): either $S = \Omega$, i.e. $\Delta \equiv 0$, or S is a locally finite union of real-analytic arcs and branch points. The first case turns out to be exactly that of a *degenerate problem* in the sense defined earlier, so that in this case all trajectories are time-optimal, at least locally. In the second case, the "singular trajectories" will have to live in the arcs that make up the set S, and this imposes severe restrictions on what may happen. One can also analyze the junctions between singular and bang-bang arcs, and carry the analysis to completion (cf. [30], [31], [32]), proving that

Theorem 5.9. *For a nondegenerate real-analytic system 2.4 in an open connected subset Ω of the plane, every time-optimal control is a finite concatenation of real-analytic controls that are either "bang" (i.e. $\equiv 1$ or $\equiv -1$) or singular.*

One can attempt to pursue a similar analysis in higher dimensions. It turns out that the problem becomes much more difficult, but a very rich structure is beginning to emerge, in which the Lie brackets naturally play a central role. Significant results on this problem have recently been obtained by Bressan (cf. [4]) and Schättler (cf. [18], [19], [20]).

We conclude our list of examples of Lie bracket results by stating a theorem that explains, at least in principle, the surprising ubiquitousness of the Lie bracket. Suppose we are given a collection $\mathcal{F}$ of smooth vector fields. A *Lie bracket relation* (LBR) at p among the vector fields in $\mathcal{F}$ is a formal linear combination R of iterated formal brackets of symbols —i.e. indeterminates— that stand for the vector fields in $\mathcal{F}$ (technically, an element of the free Lie algebra generated by indeterminates $\{X_f : f \in \mathcal{F}\}$) that vanishes when the actual vector fields of $\mathcal{F}$ are substituted for the symbols representing them. (For instance, if $f \in \mathcal{F}$ and $g \in \mathcal{F}$, and F, G are the corresponding indeterminates, then $F+[F,G]-3[G,[F,G]]$ is a LBR at p iff $f(p)+[f,g](p)-3[g,[f,g]](p)=0$.) Assume now that we are given two families $\mathcal{F}^i = \{f_a^i : a \in A\}$, $i=1,2$ of vector fields, *indexed by the same set* A, and points $p^i \in \Omega^i$. Use the same indeterminates F_a for both systems. Let us say that System $\mathcal{F}^1$ near p^1 is *equivalent* to System $\mathcal{F}^2$ near p^2 if there exists a diffeomorphism Φ from a neighborhood N^1 of p^1 onto a neighborhood N^2 of p^2 that transforms f_a^1 to f_a^2 for each a. In other words, the two systems are equivalent if they really are the same up to a change of coordinates. Then one has the following (cf. [28]):

Theorem 5.10. *Assume that the systems $\mathcal{F}^i$ as above are analytic and satisfy the accessibility property from p^i. Then $\mathcal{F}^1$ near p^1 is equivalent to $\mathcal{F}^2$ near p^2 if and only if both systems have exactly the same LBR's at the correponding points p^i.*

This theorem says that, at least for the case of analytic systems, all local properties at a point are determined by the LBR's. This explains why in so many cases the results are naturally stated in terms of Lie brackets.

6 THE ROLE OF REAL ANALYTICITY

Our main emphasis so far has been on the role of the Lie bracket. However, we have also encountered at various points the concept of real analyticity. It should come as no surprise that this matters. Take, for instance, Theorem 5.3. This says, in particular, that the reachable set $\mathcal{R}_f(p)$ is entirely contained in a submanifold whose dimension is that of $\Lambda(\mathcal{F}(f))(p)$. But $\Lambda(\mathcal{F}(f))(p)$ is an object that can be computed by just looking at the vector fields $f(\cdot,u)$ and their derivatives of all orders *at the point* p. So we are saying, among other things, that if we are in, say, $\mathbb{R}^4$, but the dimension of $\Lambda(\mathcal{F}(f))(p)$

turns out to be 3, then it follows that there is no way that one can fill up a 4-dimensional set by means of trajectories starting at p, *even if one is allowed to go far away.* So the values of the vector fields and their derivatives at a point p impose severe restrictions on what these vector fields can do even far away from p. Clearly, this sort of thing cannot possibly happen for general C^∞ systems, and it is not surprising that real analyticity is needed to derive it.

Another striking illustration of the difference between the C^∞ and the real-analytic case is provided by the following two facts (cf. [34] and [35]).

H. *Let $a < b$, and let $\eta : [a,b] \to [-1,1]$ be an arbitrary measurable function. Then there exist a single-input C^∞ system of the form 2.4 in $\mathbb{R}^3$, and points p, q in $\mathbb{R}^3$, such that η steers p to q and no other control does.*

K. *Consider a real analytic system of the form 2.4. Then, given any pair of points p, q such that there is a control η that steers p to q in time T, it follows that there is a (possibly different) $\tilde{\eta}$ that also steers p to q in time T and is analytic on an open dense subset J of its interval of definition.*

The contrast between **H** and **K** is quite remarkable. For C^∞ systems, **H** says that there is no general *a priori* regularity theorem for optimal controls. If we know that a control η is the unique time-optimal control steering p to q for some C^∞ problem, then there is nothing we can infer about η from this information, since every η can arise in this way for some problem. (Notice that the η of **H** *is* time-optimal, since it is the only control that steers p to q.) On the other hand, for *real-analytic problems* we can infer *something.* It is not much, of course, but it is enough to show that there is a fundamental difference between the two cases. At the moment, it is still a mystery how big this difference is. For instance, it does not seem to be known whether we can strengthen the conclusion of **K** to make the complement of J a null set, or even a countable one. (That it cannot be taken to be finite follows from "Fuller's problem," cf. [14].)

7 CONCLUSION

In the planet of the little green men, the development of control theory began by studying the simplest control systems where there really is some control (i.e. two-vector-field system) and asking the most natural questions, such as

1. when is the minimum-time problem degenerate?
2. when does the effect of a control that involves alternation between two different actions actually depend on the order in which these actions are carried out rather than just how much time has been alotted to each one?
3. when does a bang-bang theorem hold?
4. what do reachable sets look like, and how can one calculate their dimension?
5. when does the reachable set have a nonvoid interior?
6. when does local controllability hold?

7. what are optimal controls like?

Having asked these questions, they immediately found that the effort to answer them leads right away to the Lie bracket, and to drawing a sharp line of demarcation between C^∞ and real analytic systems, the latter having many interesting properties that do not hold in the C^∞ case. This is why all their textbooks on Control Theory start with the study of vector fields, Lie brackets, Lie algebras of vector fields, integral manifolds, and real analytic systems. This path seems to me to be much more "natural" than the one that has been followed on Earth. So the path we have followed is not the only one we could have followed, and perhaps not the one we *should* have followed. However, we are free to make choices now that will affect our future path. We can (and probably should) make an effort to incorporate the missing elements (geometry and topology) into the mainstream of our field, so as to make the next thirty years even more exciting than the first ones.

REFERENCES

1 L. D. Berkovitz, *Optimal Control Theory*, Springer-Verlag, New York, 1974.

2 R. M. Bianchini and G. Stefani, Sufficient conditions of local controllability, *Proc. 25th. I.E.E.E. Conf. Decision and Control*, Athens, Greece (December 1986).

3 A. Bressan, A high-order test for optimality of bang-bang controls, *S.I.A.M. J. Control and Optimization*, 23: 38-48 (1985).

4 A. Bressan, The generic local time-optimal stabilizing controls in dimension 3, *S.I.A.M. J. Control and Optimization*, 24: 177-190 (1986).

5 P. Brunovsky, Local controllability of odd systems, *Banach Center Publications* 1: 39-45, Warsaw, Poland (1974).

6 R. Hermann, On the accessibility problem in control theory, in *International Symposium on Nonlinear Differential Equations and Nonlinear Mechanics*, J.P. LaSalle and S. Lefschetz Eds., Academic Press, New York (1963).

7 H. Hermes, On local and global controllability, *S.I.A.M. J. Control* 12: 252-261 (1974).

8 H. Hermes, Lie algebras of vector fields and local approximation of attainable sets, *S.I.A.M. J. Control and Optimization* 16: 715-727 (1974).

9 M. Kawski, A new necessary condition for local controllability, in *Proceedings of the Conference on Differential Geometry*, San Antonio, 1986.

10 M. Kawski, High-order small-time local controllability, to appear in *Nonlinear Controllability and Optimal Control*, H.J. Sussmann ed., to be published by M. Dekker, Inc.

11 E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*, J. Wiley, New York, 1967.

12 G. Leitmann, *An Introduction to Optimal Control*, McGraw-Hill, New York, 1966.

13 J. Macki and A. Strauss, *Introduction to Optimal Control Theory*, Springer-Verlag, New York, 1982.

14 C. Marchal, Chattering arcs and chattering controls, *J. Optim. Theory and Applic.*, 11: 441-468 (1973).

15 T. Nagano, Linear differential systems with singularities and an application to transitive Lie algebras, *J. Math. Soc. Japan* 18: 398-404 (1966).

16 L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko, *The Mathematical Theory of Optimal Processes* (translated by K. N. Trigoroff, L. W. Neustadt, editor), J. Wiley, 1962.

17 D. Ruelle, Is our mathematics natural? The case of equilibrium statistical mechanics, *Bull. Amer. Math. Soc.*, 19: 259-268 (July 1988).

18 H. Schättler, On the local structure of time-optimal bang-bang trajectories in $\mathbb{R}^3$, *S.I.A.M. J. Control and Optim.*, 26, No. 1: 186-204 (1988).

19 H. Schättler, The local structure of time-optimal trajectories in dimension 3 under generic conditions, *S.I.A.M. J. Control and Optim.*, 26, No. 1: 186-204 (1988).

20 H. Schättler and H. J. Sussmann, On the regularity of optimal controls, *Journal of Applied Mathematics and Physics (ZAMP)* 38: 292-301, (1987).

21 G. Stefani, Local controllability of nonlinear systems: an example, *Systems and Control Letters*, 6, No. 2: 123-125 (1985).

22 G. Stefani, On the local controllability of a scalar-input control system, in *Theory and Applications of Nonlinear control Systems*, C.I. Byrnes and A. Lindquist eds., Elsevier Science Publ. (1986).

23 G. Stefani, Polynomial approximations to control systems and local controllability, *Proc. 24th. I.E.E.E. Conference on Decision and Control*, Ft. Lauderdale, Fla. (1986).

24 H. J. Sussmann and V. Jurdjevic, Controllability of nonlinear systems, *J. Diff. Equations* 12: 95-116 (1972).

25 H. J. Sussmann, A sufficient condition for local controllability, *S.I.A.M. J. Control* 16, No.5: 790-802 (1978).

26 H. J. Sussmann, A bang-bang theorem with bounds on the number of switchings, *S.I.A.M. J. Control and Optimization* 17, No. 5: 629-651 (1979).

27 H. J. Sussmann, Lie brackets and local controllability: a sufficient condition for scalar input systems, *S.I.A.M. J. Control and Optimization* 21, No. 5: 686-713 (1983).

28 H. J. Sussmann, Lie brackets, real analyticity and geometric control theory, in *Differential Geometric Control Theory*: 1-115, R.W. Brockett, R.S. Millman and H.J. Sussmann Eds., Birkhäuser Boston Inc. (1983),

29 H. J. Sussmann, A general theorem on local controllability, *S.I.A.M. J. Control and Optimization* 25, No. 1: 158-194 (1987).

30 H. J. Sussmann, The structure of time-optimal trajectories for single-input systems in the plane : the C^∞ nonsingular case, *S.I.A.M. J. Control and Optimization* 25: 433-465 (1987).

31 H. J. Sussmann, The structure of time-optimal trajectories for single-input systems in the plane: the general real-analytic case, *S.I.A.M. J. Control and Optimization* 25, No. 4: 868-904, (1987).

32 H. J. Sussmann, Regular synthesis for time-optimal control of single-input real-analytic systems in the plane, *S.I.A.M. J. Control and Optimization* 25, No. 5: 1145-1162 (1987).

33 H. J. Sussmann, A general theorem on symmetries and local controllability, in *Proceedings of the 24th I.E.E.E. Conference on Decision and Control*: 27-32, Fort Lauderdale, Florida, (1985).

34 H. J. Sussmann, A weak regularity theorem for real analytic optimal control problems, *Revista Matemática Iberoamericana* 2, No. 3: 307-317 (1986).

35 H. J. Sussmann, Recent developments in the regularity theory of optimal trajectories, in *Linear and Nonlinear Mathematical Control Theory*, Rendiconti del Seminario Matematico, Università e Politecnico di Torino, Fascicolo Speciale 1987, 149-182.

36 K. Wagner, Uber den Steuerbarkeitsbegriff bei nichtlinearen Kontrollsystemen, *Arch. Math.*, 47: 29-40 (1986).

37 J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York (1972).