SIAM J. CONTROL AND OPTIMIZATION
Vol. 25, No. 1, January 1987

011

A GENERAL THEOREM ON LOCAL CONTROLLABILITY*

H. J. SUSSMANN†

Abstract. We prove a general sufficient condition for local controllability of a nonlinear system at an equilibrium point. Earlier results of Brunovsky, Hermes, Jurdjevic, Crouch and Byrnes, Sussmann and Grossmann, are shown to be particular cases of this result. Also, a number of new sufficient conditions are obtained. All these results follow from one simple general principle, namely, that local controllability follows whenever brackets with certain symmetries can be "neutralized," in a suitable way, by writing them as linear combinations of brackets of a lower degree. Both the class of symmetries and the definition of "degree" can be chosen to suit the problem.

Key words. nonlinear control, local controllability, nilpotent approximation, symmetries

AMS(MOS) subject classifications. 49B10, 93C10

Introduction. In recent years, several papers have been published giving sufficient conditions for a nonlinear control system to be locally controllable from a point. (Cf. Brunovsky [3], Crouch and Byrnes [5], Grossmann [8], Hermes [10]-[13], Jurdjevic [14], Stefani [21], [22], Sussmann [24], [25].) The purpose of this article is to prove a general theorem which contains all these results as particular cases and, in addition, gives stronger results. Our result (Theorem 2.4) shows that many known sufficient conditions can be derived in a unified way from a single general principle, namely, the combination of a nilpotent approximation with the use of input symmetries. Section 2 is devoted to the statement of the main theorem, preceded by an outline of the basic facts and definitions needed for its formulation. Section 3 reviews the basic formalism needed to set up the nilpotent approximation, and proves a number of technical lemmas needed to turn this approximation into a tool for establishing local controllability results. Sections 4 and 5 introduce the basic ingredients of our main result, namely, dilations and invariant elements. The proof of the main theorem is then given in § 6. In § 7, we review in detail the various controllability results referred to above, and explain how they all follow from our result. We also prove stronger versions of several of those theorems, and some new sufficient conditions. Finally, in § 8 we discuss some of the limitations of our method. In addition to the observations of § 8, we remark that there are recent results by R. M. Bianchini and G. Stefani, as well as work by H. Knobloch and K. Wagner, which provide new sufficient conditions that are not contained in the ones given here.

1. Preliminaries. The local controllability problem has a long history, beginning with the classical controllability theory for linear systems, and the first nonlinear local controllability result, namely, the theorem which states that if the linearization of a system at an equilibrium point p is controllable, then the system itself is locally controllable from p in small time (i.e. for every $T>0$, the time T reachable set from p contains p in its interior; cf., for example, Lee and Markus [17]). This "small-time local controllability" property, henceforth abbreviated as STLC, is of interest to control theorists for a number of reasons, such as: (a) that a sufficient condition for STLC is obviously equivalent to a necessary condition for the constant trajectory $x(t)\equiv p$ to

* Received by the editors April 3, 1985. This research was partially supported by the National Science Foundation under grant MCS 78-02442.

† Mathematics Department, Rutgers University, New Brunswick, New Jersey 08903.

lie on the boundary of the attainable set from p; since the simplest form of the Pontryagin Maximum Principle is precisely one such necessary condition, the STLC problem can be viewed as a particular case of the general problem of "high order optimality conditions;" (b) that STLC is equivalent to an important property of the optimal time function V, namely, continuity at p. (Here V is defined by letting $V(q)$ be the infimum of all the times T such that q can be reached from p in time T. If we wish to study the more common V function, defined in terms of the time it takes to steer q to p, then the continuity of V at p is equivalent to the STLC property for the system obtained by running the original system backwards.)

More recently, the problem has attracted the interest of "differential geometric control theorists," i.e. of those who take the point of view that a control system is primarily a family of vector fields on a manifold, and a lot of the control-theoretically interesting information about the system should be contained in the Lie brackets of these vector fields (cf. [9], [16], [19], [23], [26]). At the early stages of the development of this "Lie theoretic approach," attention was concentrated on proving those results that followed most naturally from the method. In particular, it was recognized right away that the Lie algebraic method yielded a complete characterization of local controllability for real analytic systems with the somewhat unnatural property of being "symmetric," i.e. such that every trajectory run backwards is also a trajectory. (Hermann [9], Lobry [19]; the result is known as "Chow's Theorem.") On the other hand, for "reasonable" (i.e. not necessarily symmetric, but real-analytic) systems, the method yielded a complete characterization of a property which is related to, but not quite the same as, STLC. D. Elliott introduced the name "accessibility property" to refer to the property that the reachable set from p has an interior point. The so-called "positive form of Chow's Theorem" (Krener [16]; cf. also [23]) characterizes this property in terms of Lie brackets. In 1974, P. Brunovsky [3] started from the observation that the "Lie theoretic" theorem about symmetric systems does not even give the most classical of all local controllability theorems, namely, the one for linear systems. He then proceeded to single out a class of systems (called "odd systems") which could be proved to be STLC by Lie theoretic methods, and contained the class of linear systems. Since then, other local controllability results have been proved, as indicated above. The common feature of all these results is the exploitation of certain "structural symmetries" of a problem.

The traditional approach towards proving local controllability theorems has been to construct "control variations." Heuristically, if one can construct control variations in all possible directions, then the reachable set ought to be a full neighborhood of the starting point. The argument can usually be made rigorous by some topological consideration. Ideally, the construction of variations in various directions should involve Lie bracket calculations. In practice, however, these calculations become rather cumbersome, and a different method is desirable which would construct, once and for all, a large collection of variations. One such method was used by us in [25], to prove a conjecture of H. Hermes. (Our earlier paper [24], which proved a different sufficient condition for STLC, was based on constructing variations, and it has only recently become clear to us that the result of [24] also follows using the method of [25].) The goal of the present paper is to prove the most general result that can be obtained by means of the method of [25].

A brief outline of the approach is as follows. Since a control system is primarily a family $\mathcal{V} = \{V_i : i \in I\}$ of vector fields, one can associate with it the Lie algebra $L(\mathcal{V})$ of vector fields generated by $\mathcal{V}$. Forgetting about rigor, one can think about "the Lie group" $G(\mathcal{V})$ with Lie algebra $L(\mathcal{V})$, and obtain an "action" of $G(\mathcal{V})$ on the state space of the system. A $g \in G(\mathcal{V})$ is a product of exponentials $\exp(t_j V_{i_j})$, and therefore the result gp of acting on p by g is a point obtained by starting from p and following integral trajectories of the V_i, with switchings of vector fields allowed, and with motion "backwards in time" permitted as well. Those g's for which all the t_j are positive constitute a subsemigroup S of $G(\mathcal{V})$, which gives rise to the true trajectories of the control system. The reachable set from p is $S \cdot p$. For $\mathcal{V}$ to be locally controllable from p, $S \cdot p$ has to have a nonempty interior, and so $G(\mathcal{V})p$ must be open, which means that, at least locally, $G(\mathcal{V})$ has to act transitively. If H is the isotropy group at p of this action, then a sufficient condition for p to be an interior point of $S \cdot p$ is that the interior of S in $G(\mathcal{V})$ should contain an element of H.

To make this rigorous, an algebraic formalism is needed to surmount the obstacles arising from the fact that $L(\mathcal{V})$ is, in general, infinite dimensional, and therefore $G(\mathcal{V})$ is not a well defined "Lie group." Rather than work with $L(\mathcal{V})$ one works formally, with a free Lie algebra $L(\mathbf{X})$ in indeterminates X_i, and with its completion, the Lie algebra $\hat{L}(\mathbf{X})$ of formal Lie series in the X_i. Then there is a well defined group $\hat{G}(\mathbf{X})$, the group of exponentials of Lie series (cf., for example, Serre [20]). The controls can be embedded in $\hat{G}(\mathbf{X})$ as a subsemigroup S, by means of a map which assigns to each control a noncommutative formal power series, obtained by solving the differential equation of the system formally, using the indeterminates rather than the vector fields. (This map, introduced by Chen in [4], has been extensively used in control theory by M. Fliess, under the name of "Chen series," cf. [6], [7].) Although obvious convergence and integrability difficulties arise if one tries to make $\hat{G}(\mathbf{X})$ act on the state space, the subsemigroup S does act in an obvious way, since S is identified with the set of admissible controls. And the series of a control $u(\cdot)$ contains a lot of information about the action of $u(\cdot)$. (More precisely, it is an asymptotic series, and it converges in the analytic case if $u(\cdot)$ is sufficiently small, cf. [1], [2], [7], [18], [25].) Since $\hat{G}(\mathbf{X})$ is not yet a true Lie group, one then makes a nilpotent approximation $G^N(\mathbf{X})$ of $\hat{G}(\mathbf{X})$ by killing all brackets of degree $>N$. If I is finite, $G^N(\mathbf{X})$ is now a Lie group in the usual sense. Then there is a corresponding approximating semigroup S^N. Although it is not possible in general to have $G^N(\mathbf{X})$ act on the state space, one can still define an "approximate action" and an "approximate isotropy group." To get local controllability one must be able to prove (modulo technicalities) that the interior of S^N intersects the isotropy group. This we do by proving a general lemma that says that the interior of S^N always contains an element of a "very special form." It then follows that, if one hypothesizes that all these "very special" elements are in the isotropy group, one gets controllability. As will be made clear in § 7, all known local controllability theorems amount to various forms of this hypothesis.

The special elements are obtained as the fixed points of the action of a finite group Λ on "input symmetries." An input symmetry is, roughly, a linear map from $L(\mathbf{X})$ to $L(\mathbf{X})$ whose exponential maps S to S. Examples of such symmetries are: (a) multiplying a control by -1, if its range of values permits it; (b) interchanging two controls; (c) time reversal. If a system has many symmetries, then there will be few Λ-fixed elements, and the resulting local controllability theorem will be very strong. As an example, we remark that the introduction of time-reversal, which was not used in [25], enables us here to prove a result which is considerably stronger than the Hermes conjecture proved in [25].

It turns out that the condition that certain "special elements" of the semigroups S^N be in the "isotropy group" can be rephrased, by passing to the logarithms, as the requirement that certain Lie brackets should vanish at p. It then becomes apparent that one can do slightly better. The brackets need not vanish. It suffices for them to be "neutralized," i.e. expressible as linear combinations of brackets of lower degree. And there is a certain amount of freedom as to the concept of "degree" to be used. One can use any one-parameter group of dilations to define "degree," provided that certain technical conditions hold.

In order to avoid unnecessary complications, we will only work with systems that can be studied using a free Lie algebra generated by a *finite* set of indeterminates. That is, we will only study systems where the collection $\mathcal{V}$ of associated vector fields is either finite, or a set of linear combinations of a finite set of vector fields. That is, we will only work with systems of the form

$$\dot{x} = \sum_{i=1}^{k} v_i g_i(x) \tag{1.1}$$

where the control $v = (v_1, \cdots, v_k)$ is required to satisfy a constraint $v \in J$, where J is some subset of $\mathbb{R}^k$. It is then clear that we can assume that J linearly spans $\mathbb{R}^k$. If J does not affinely span $\mathbb{R}^k$, let A be the affine hull of J. By making a linear change of coordinates, we may assume that A is the set $\{1\} \times \mathbb{R}^{k-1}$. Then the system (1.1) becomes

$$\dot{x} = f_0(x) + \sum_{i=1}^{m} u_i f_i(x) \tag{1.2}$$

with control constraint $u = (u_1, \cdots, u_m) \in K$. (Here K is such that $J = \{1\} \times K$, and $m = k-1$.)

If J affinely spans $\mathbb{R}^k$, then we let $f_0 \equiv 0$, $m = k$, $g_i = f_i$ for $i = 1, \cdots, m$, $K = J$, $u_i = v_i$ for $i = 1, \cdots, m$. Then our system is also of the form (1.2), with a control constraint $u \in K$, where K affinely spans $\mathbb{R}^m$. It is in this form that, from now on, all our systems will be expressed.

2. Statement of the main theorem.

In this section we will state our main local controllability theorem. In order to get to the statement as quickly as possible, we will omit a number of definitions. Detailed definitions of all the concepts occurring in the statement are given in subsequent sections.

We consider control systems of the form

$$\dot{x} = f_0(x) + \sum_{i=1}^{m} u_i f_i(x), \qquad x \in M \tag{2.1}$$

with a control constraint

$$u = (u_1, \cdots, u_m) \in K \tag{2.2}$$

where

(CS1) *M is a smooth (i.e. C^∞) manifold,*

(CS2) *$\mathbf{f} = (f_0, \cdots, f_m)$ is an $(m+1)$-tuple of C^∞ vector fields on M,*

and

(CS3) *K is a subset of $\mathbb{R}^m$ such that*

$$\mathrm{Aff}\,(K) = \mathbb{R}^m. \tag{2.3}$$

Here Aff (K) denotes the affine hull of K, i.e. the set of all finite linear combinations $\sum \alpha_i u^i$ with the $u^i \in K$, $\alpha_i \in \mathbb{R}$, and $\sum_i \alpha_i = 1$.

To specify a system we must give M, $\mathbf{f}$ and K. So we will simply refer to the triple $\Sigma=(M, \mathbf{f}, K)$ as the control system, it being understood that M, $\mathbf{f}$ and K are supposed to satisfy (CS1), (CS2) and (CS3).

An *admissible control for* Σ is a Lebesgue integrable, K-valued function defined on some interval $[0, T]$. If $u(\cdot):[0, T]\to K$ is an admissible control, a *trajectory* for $u(\cdot)$ is an absolutely continuous curve $x(\cdot):[0, T]\to M$ such that

$$\dot{x}(t)=f_0(x(t))+\sum_{i=1}^{m} u_i(t)f_i(x(t)) \tag{2.4}$$

for almost all $t\in[0, T]$.

If $q\in M$ is of the form $x(T)$ for some trajectory such that $x(0)=p$, then q will be said to be *reachable from p in time T*. The set of all q that are reachable from p in time T for the system $\Sigma=(M, \mathbf{f}, K)$ is the *time T reachable set from p, and will be denoted by* Reach (Σ, T, p). Also we write

$$\text{Reach}\,(\Sigma, \leqq T, p)=\bigcup_{0\leqq t\leqq T} \text{Reach}\,(\Sigma, t, p) \tag{2.5}$$

for $T\geqq 0$.

The system Σ is *small-time locally controllable* (STLC) from p if p is an interior point of Reach $(\Sigma, \leqq T, p)$ for all $T>0$. An equivalent characterization of this condition involves the optimal time function $V_{\Sigma,p}$. We define $V_{\Sigma,p}(q)$ to be the infimum of those T such that q is reachable from p in time T. (If no such T exists, then $V_{\Sigma,p}(q)=+\infty$.) Then Σ is STLC from p if and only if $V_{\Sigma,p}$ is continuous at p.

One can also consider the reachable sets obtained by restricting the class of admissible controls. For instance, we let $\text{Reach}_{pc}(\Sigma, T, p)$, $\text{Reach}_{pc}(\Sigma, \leqq T, p)$ be the reachable sets obtained by using *piecewise constant* controls, and we say that Σ is STLC_{pc} from p if p is an interior point of $\text{Reach}_{pc}(\Sigma, \leqq T, p)$ for all $T>0$. The sufficient condition stated below in our main theorem is for STLC. However, under the hypotheses of the theorem, STLC and STLC_{pc} are equivalent, as will be observed below (cf. Proposition 2.3), so that the distinction between these two types of controllability need not worry us here.

If K is compact and convex, then one can also consider the sets reachable by bang-bang controls. (A bang-bang control is a piecewise constant control with values in the set of extreme points of K.) The corresponding small-time local controllability property is denoted by STLC_{bb}. Again, Proposition 2.3 will show that STLC and STLC_{bb} are equivalent under the hypotheses of our main theorem.

If $\mathcal{F}$ is a family of C^∞ vector fields on a manifold M, then $L(\mathcal{F})$ denotes the Lie algebra of vector fields generated by the elements of $\mathcal{F}$. If $\mathcal{V}$ is any set of vector fields on M, and $p\in M$, then we write

$$\mathcal{V}(p)=\{V(p)\colon V\in\mathcal{V}\}. \tag{2.6}$$

The family $\mathcal{F}$ is said to satisfy the *Lie algebra rank condition* (LARC) at p if $L(\mathcal{F})(p)$ is the whole tangent space of M at p. An *$\mathcal{F}$-trajectory* is a curve $x(\cdot)$ which is a finite concatenation of integral arcs of members of $\mathcal{F}$. (Note: if an integral arc γ of a member f of $\mathcal{F}$ is reparametrized by reversing the sense of time, then the resulting curve is an integral arc of $-f$, and need not be an $\mathcal{F}$-trajectory, since $-f$ need not belong to $\mathcal{F}$.) The family $\mathcal{F}$ has the *accessibility property* (AP) from p if, for every $T>0$, the set of points that can be reached from p by $\mathcal{F}$-trajectories in time $\leqq T$ has a nonempty interior. The following is a standard result from accessibility theory (the "positive form of Chow's Theorem," cf. Krener [16], Sussmann and Jurdjevic [23]).

PROPOSITION 2.1. *Let $\mathcal{F}$ be a family of C^∞ vector fields on a C^∞ manifold M. Then the* LARC *at p implies the* AP *from p. Conversely, the* AP *from p implies the* LARC *at p if M is a real-analytic manifold and the members of $\mathcal{F}$ are real-analytic.*

To a system Σ of the form (2.1), with a control constraint (2.2), we associate the family $\mathcal{F}_\Sigma$ whose members are all the vector fields $f_0+\sum_{i=1}^m u_i f_i$, for $(u_1, \cdots, u_m) \in K$. The hypothesis that $\mathrm{Aff}(K) = \mathbb{R}^m$ implies that the linear span of the members of $\mathcal{F}_\Sigma$ is precisely the same as the linear span of $f_0, \cdots, f_m$. Therefore $L(\mathcal{F}_\Sigma) = L(\mathbf{f})$, so that $\mathcal{F}_\Sigma$ satisfies the LARC at p if and only if $\mathbf{f}$ does. On the other hand, it is easy to see that an $\mathcal{F}_\Sigma$ trajectory is precisely the same as a trajectory of Σ for a piecewise constant control. Hence Σ cannot be STLC_{pc} from p unless $\mathcal{F}_\Sigma$ satisfies the AP from p. On the other hand, if $f_0, \cdots, f_m$ are real analytic vector fields, $\mathcal{F}_\Sigma$ satisfies the AP from p if and only if $\mathbf{f}$ satisfies the LARC from p. Therefore, in the analytic case, it is no restriction to assume that $\mathbf{f}$ satisfies the LARC from p, if we seek to characterize the STLC_{pc} property. Actually, it is easy to prove:

PROPOSITION 2.2. *A system Σ of the form* (2.1), *with a control constraint* (2.2), *and $f_0, \cdots, f_m$ real analytic, cannot be* STLC *from a point p unless* $\mathbf{f}$ *satisfies the* LARC *from p.*

Moreover, when the LARC from p holds, the distinction between STLC, STLC_{pc} and STLC_{bb} disappears, as shown by the following result, whose proof is given in the Appendix.

PROPOSITION 2.3. *Let Σ be a system of the form* (2.1), *with a control constraint* (2.2) *that satisfies* (2.3). *Assume that* $\mathbf{f}$ *satisfies the* LARC *at p.*

Let $\tilde{K}$ be the closure of the convex hull of K, and let $\tilde{\Sigma}$ be the system $(M, \mathbf{f}, \tilde{K})$. Then $\tilde{\Sigma}$ is SLTC *from p if and only if Σ is* SLTC_{pc} *from p.*

In particular, Proposition 2.3 implies that, for an arbitrary Σ, the STLC and STLC_{pc} properties from p are equivalent, if $\mathbf{f}$ satisfies the LARC at p. Also, if K is compact and convex, STLC and STLC_{bb} are equivalent.

The sufficient condition for STLC to be proved here involves two main ingredients, namely, a finite group of symmetries and a one-parameter group of dilations. The symmetries considered will be mappings of a Lie algebra which is naturally associated to our problem. Precisely, we consider $L(\mathbf{X})$, the free Lie algebra in the indeterminates $\mathbf{X} = (X_0, \cdots, X_m)$. We will be interested in linear maps $\lambda : L(\mathbf{X}) \to L(\mathbf{X})$ which are not necessarily Lie algebra automorphisms, but have a weaker property which we now define.

Let L be a Lie algebra over $\mathbb{R}$. We define $[L]^k$ for $k = 1, 2, \cdots$ by $[L]^1 = L$, $[L]^{k+1} = [L, L^k]$. Clearly, any Lie algebra automorphism of L maps each $[L]^k$ into itself. A linear mapping $\lambda : L \to L$ which is a linear isomorphism and satisfies $\lambda([L]^k) \subseteq [L]^k$ for each k will be called a *pseudoautomorphism* of L.

In the particular case when L is $L(\mathbf{X})$, the $[L]^k$ are the ideals $L_k(\mathbf{X})$, where $L_k(\mathbf{X})$ is the sum of all the homogeneous components $L^{j,\mathrm{hom}}(\mathbf{X})$ of degree j, for $j \geqq k$. If $\lambda : L(\mathbf{X}) \to L(\mathbf{X})$ is a pseudoautomorphism, then λ gives rise to a linear map $\hat{\lambda}$ from $\hat{L}(\mathbf{X})$ to $\hat{L}(\mathbf{X})$, where $\hat{L}(\mathbf{X})$ is the Lie algebra of formal Lie series in $X_0, \cdots, X_m$. (If $S \in \hat{L}(\mathbf{X})$, and $S = \sum_{j=1}^\infty S_j$, where S_j is homogeneous of degree j, then $\hat{\lambda}$ is defined by $\hat{\lambda}(S) = \sum_j \lambda(S_j)$. The sum is well defined because, for each k, the only terms that may contribute to the homogeneous component of degree k are the $\lambda(S_j)$ for $j \leqq k$.) It is clear that $\widehat{\lambda\mu} = \hat{\lambda}\hat{\mu}$ if λ, μ are pseudoautomorphisms.

The class of controls is embedded as a subsemigroup $\hat{S}(\mathbf{X}, K)$ of the group $\hat{G}(\mathbf{X}) = \{\exp(Z) \colon Z \in \hat{L}(\mathbf{X})\}$. A pseudoautomorphism λ of $L(\mathbf{X})$ gives rise to a mapping

$\lambda^{\#}$ from $\hat{G}(\mathbf{X})$ to $\hat{G}(\mathbf{X})$ by letting

$$\lambda^{\#}(\exp(Z)) = \exp(\hat{\lambda}(Z)) \quad \text{for } Z \in \hat{L}(X). \tag{2.7}$$

An *input symmetry* for Σ is a pseudoautomorphism λ of $L(\mathbf{X})$ such that the corresponding map $\lambda^{\#}$ maps $\hat{S}(\mathbf{X}, K)$ to $\hat{S}(\mathbf{X}, K)$. (Actually, the definition of input symmetry only depends on m and K, and not on the particular choice of $M, f_0, \cdots, f_m$.)

The second important ingredient is a one parameter group of dilations $\{\Delta(\rho): 0<\rho<\infty\}$ of the linear space $V = L^{1,\text{hom}}(\mathbf{X})$. Then Δ gives rise to groups of dilations Δ^A, Δ^L of the free associative algebra $A(\mathbf{X})$ in the indeterminates $X_0, \cdots, X_m$, and of $L(\mathbf{X})$, respectively. Also, one obtains a one-parameter group $\hat{\Delta}^A$ of automorphisms of the algebra $\hat{A}(\mathbf{X})$ of formal power series in $X_0, \cdots, X_m$. We call Δ *compatible* with $\hat{S}(\mathbf{X}, K)$ if the $\hat{\Delta}^A(\rho)$ map $\hat{S}(\mathbf{X}, K)$ into itself for $0<\rho\leqq 1$. (Equivalently, Δ is compatible with $\hat{S}(\mathbf{X}, K)$ if and only if, for every $u=(u_1, \cdots, u_m) \in K$ and every ρ such that $0<\rho\leqq 1$, $\Delta(\rho)(X_0+\sum_{i=1}^m u_i X_i)$ is of the form $T(X_0+\sum_{i=1}^m v_i X_i)$ for some $T>0$, $v=(v_1, \cdots, v_m) \in K$.)

A group of dilations Δ as above can be used to define the Δ*-degree* of an element Z of $A(\mathbf{X})$. We call Z Δ*-homogeneous* of degree r if $\Delta^A(\rho)(Z) = \rho^r Z$ for every ρ. If Z is arbitrary, then Z is a finite sum of homogeneous elements, and the Δ*-degree* of Z (denoted by $\deg_\Delta(Z)$) is the largest of the degrees of the homogeneous components of Z.

If $\mathbf{f} = (f_0, \cdots, f_m)$ is an $(m+1)$-tuple of C^∞ vector fields on a C^∞ manifold M, then we can consider the map $\text{Ev}(\mathbf{f})$ which assigns to every element P of $L(\mathbf{X})$ the vector field obtained by plugging in each f_i for the corresponding indeterminate X_i. If $p \in M$, then we also define the map $\text{Ev}_p(\mathbf{f})$, from $L(\mathbf{X})$ to the tangent space T_pM, given by $\text{Ev}_p(\mathbf{f})(P) = \text{Ev}(\mathbf{f})(P)(p)$.

We now define what it means for a $Z \in L(\mathbf{X})$ to be Δ*-neutralized for* $\mathbf{f}$ *at* p. If Z is Δ-homogeneous, we say that Z is Δ-neutralized for $\mathbf{f}$ at p if $\text{Ev}_p(\mathbf{f})(Z)$ can be expressed as a sum of vectors $\text{Ev}_p(\mathbf{f})(Q_j)$, where the Q_j are elements of $L(\mathbf{X})$ such that $\deg_\Delta(Q_j) < \deg_\Delta(Z)$. (Clearly, the Q_j can always be chosen to be Δ-homogeneous.) If Z is not necessarily homogeneous, then we write Z as a sum of homogeneous components, and we say that Z is Δ-neutralized for $\mathbf{f}$ at p if each homogeneous component is.

With these definitions, our main result is the following.

THEOREM 2.4. *Let $\Sigma = (M, \mathbf{f}, K)$ be a control system, and let $p \in M$. Assume that*:

(i) *Σ satisfies the Lie algebra rank condition at p,*

(ii) *there exist* (a) *a finite group Λ of input symmetries and* (b) *a one-parameter group of dilations $\Delta = \{\Delta(\rho): \rho>0\}$ of $L^{1,\text{hom}}(\mathbf{X})$ which is compatible with $\hat{S}(\mathbf{X}, K)$, such that every Λ-fixed element of $L(\mathbf{X})$ is Δ-neutralized for $\mathbf{f}$ at p.*

Then Σ is small-time locally controllable at p.

3. Exponential Lie series and the nilpotent approximation. We review the basic facts about the formalism of noncommutative power series and nilpotent approximation (cf. [4], [6], [25]). The idea is to solve (2.1) formally, by using indeterminates $X_0, \cdots, X_m$ rather than the vector fields $f_0, \cdots, f_m$, and then regard a given control system as an action of a "Lie group" $\hat{G}(\mathbf{X})$ of exponential Lie series, together with the specification of a subsemigroup $\hat{S}(\mathbf{X}, K)$ which is identified with the class of controls. We now make this precise.

Let $\mathbf{X} = (X_0, \cdots, X_m)$ be a finite sequence of indeterminates. We let $A(\mathbf{X})$ denote the free associative algebra over $\mathbb{R}$ generated by the X_j. For any multiindex $I = (i_1, \cdots, i_k)$, with $i_j \in \{0, \cdots, m\}$ for $j = 1, \cdots, k$, we let $X_I = X_{i_1} \cdots X_{i_k}$. Then $A(\mathbf{X})$ is the set of all sums $\sum_I a_I X_I$, where the coefficients a_I are real numbers, the summation runs over all possible multiindices I, and all but finitely many a_I vanish. (It is understood that $X_\phi = 1$.)

We also let $\hat{A}(\mathbf{X})$ denote the set of all *formal power series* in the noncommuting indeterminates X_j, i.e. the set of all sums $\sum_I a_I X_I$ as above, except that the a_I are no longer required to vanish for all but finitely many I. In both $A(\mathbf{X})$ and $\hat{A}(\mathbf{X})$, addition is done componentwise, and multiplication is carried out using the formula $X_I X_J = X_{I*J}$, where $I * J$ is the concatenation of I and J (i.e. the multiindex obtained by writing, in order, first the components of I and then those of J).

For any nonnegative integer N, we use $A^{N,\mathrm{hom}}(\mathbf{X})$ to denote the homogeneous component of degree N of $A(\mathbf{X})$, and $A^N(\mathbf{X})$ to denote the sum of the $A^{j,\mathrm{hom}}(\mathbf{X})$ for $j = 0, \cdots, N$. The space $A^N(\mathbf{X})$ is embedded as a linear subspace of $A(\mathbf{X})$ but, naturally, it is not a subalgebra. On the other hand, $A^N(\mathbf{X})$ is an algebra if one defines multiplication as in $A(\mathbf{X})$, with the extra proviso that monomials of degree greater than N are set equal to zero. Thus regarded, $A^N(\mathbf{X})$ is the *free nilpotent associative algebra of step* $N+1$ *in the indeterminates* $X_0, \cdots, X_m$. Then $A^N(X)$ can be identified with the quotient of $A(\mathbf{X})$ by the ideal of all linear combinations of monomials of degree strictly larger than N. The canonical projection from $A(\mathbf{X})$ onto $A^N(\mathbf{X})$ is the *truncation map* $\tau_{\mathbf{X}}^N$. We will write τ^N rather than $\tau_{\mathbf{X}}^N$ whenever the context makes it clear which $\mathbf{X}$ is being referred to. Clearly, one can also think of $A^N(\mathbf{X})$ as a quotient of $\hat{A}(\mathbf{X})$. The corresponding truncation map from $\hat{A}(\mathbf{X})$ onto $A^N(\mathbf{X})$ will be denoted by $\hat{\tau}_{\mathbf{X}}^N$ or $\hat{\tau}^N$. The kernels of τ^N, $\hat{\tau}^N$ are denoted by $A_N(\mathbf{X})$, $\hat{A}_N(\mathbf{X})$, respectively. In particular, $\hat{A}_0(\mathbf{X})$ is the set of formal power series $\sum_I a_I X_I$ for which $a_\phi = 0$. The *exponential map* is a well defined bijection

$$\exp: \hat{A}_0(\mathbf{X}) \to 1 + \hat{A}_0(\mathbf{X}) \tag{3.1}$$

whose inverse is a map from $1 + \hat{A}_0(\mathbf{X})$ to $\hat{A}_0(\mathbf{X})$ denoted by "log." If $S \in \hat{A}_0(\mathbf{X})$, then $\exp(S)$ and $\log(1+S)$ are given by the usual power series.

One can also define $A_k^N(\mathbf{X})$ to be the set of all elements of $A^N(\mathbf{X})$ that are linear combinations of monomials of degree $> k$. Then

$$A_k^N(\mathbf{X}) = \tau^N(A_k(\mathbf{X})) = \hat{\tau}^N(\hat{A}_k(\mathbf{X})). \tag{3.2}$$

The exponential map

$$\exp_N: A_0^N(\mathbf{X}) \to 1 + A_0^N(\mathbf{X}) \tag{3.3}$$

and its inverse $\log_N$ are given, in this case, by power series that are actually finite sums, due to the nilpotency of $A^N(\mathbf{X})$.

The algebras $A(\mathbf{X})$, $\hat{A}(\mathbf{X})$, $A^N(\mathbf{X})$ are Lie algebras in the usual way. We let $L(\mathbf{X})$ denote the Lie subalgebra of $A(\mathbf{X})$ generated by $X_0, \cdots, X_m$. An element S of $A(\mathbf{X})$ will be said to be a *Lie element* iff $S \in L(\mathbf{X})$. It is clear that S is a Lie element iff all the homogeneous components of S are Lie elements. Therefore, if we let

$$L^{N,\mathrm{hom}}(\mathbf{X}) = L(\mathbf{X}) \cap A^{N,\mathrm{hom}}(\mathbf{X}), \tag{3.4}$$

we see that $L(\mathbf{X})$ is the direct sum of the $L^{N,\mathrm{hom}}(\mathbf{X})$, $N = 1, 2, 3, \cdots$.

The Lie algebra $L(\mathbf{X})$ is spanned by the *formal brackets* of $X_0, \cdots, X_m$. Precisely, we define $\mathrm{Br}(\mathbf{X})$ to be the smallest subset of $L(\mathbf{X})$ that contains $X_0, X_1, \cdots, X_m$ and is closed under bracketing. The elements of $\mathrm{Br}(\mathbf{X})$ will be referred to as *brackets* of $\mathbf{X}$. It is clear that every $B \in \mathrm{Br}(\mathbf{X})$ is homogeneous. (Notice that we have chosen not to define a "bracket" as a formal expression but as an element of $L(\mathbf{X})$ so that, for example, $[[X_0, X_1], [X_0, X_2]]$ and $[[X_1, X_0], [X_2, X_0]]$ are the same element of $\mathrm{Br}(\mathbf{X})$. Naturally, the elements of $\mathrm{Br}(\mathbf{X})$ are not linearly independent. Several systematic procedures for singling out subsets of Br (X) that form bases of $L(\mathbf{X})$ can be found in the literature, cf., for example, [20], [27], but we shall not need those results here.)

We can also define $\hat{L}(\mathbf{X})$ to be the set of all formal sums $\sum_{N=1}^{\infty} S_N$ such that each S_N is in $L^{N,\mathrm{hom}}(\mathbf{X})$, i.e. the set of those elements of $\hat{A}(\mathbf{X})$ all of whose homogeneous components are Lie. The members of $\hat{L}(\mathbf{X})$ will be referred to as *Lie elements* of $\hat{A}(\mathbf{X})$, and they clearly form a Lie subalgebra of $\hat{A}(\mathbf{X})$. The Lie algebras $L(\mathbf{X})$, $\hat{L}(\mathbf{X})$ are known, respectively, as the *free Lie algebra in the indeterminates* $X_0, \cdots, X_m$ and the *algebra of Lie series in* $X_0, \cdots, X_m$.

Since $\hat{L}(\mathbf{X}) \subseteq \hat{A}_0(\mathbf{X})$, the exponential map is well defined on $\hat{L}(\mathbf{X})$. The elements of $\hat{A}(\mathbf{X})$ that are of the form $\exp(S)$ for some $S \in \hat{L}(\mathbf{X})$ are the *exponential Lie series in* $X_0, \cdots, X_m$. The set of all such series is denoted by $\hat{G}(\mathbf{X})$. It follows from the Campbell-Hausdorff formula that $\hat{G}(\mathbf{X})$ is a group under multiplication. The exponential map, restricted to $\hat{L}(\mathbf{X})$, is a bijection from $\hat{L}(\mathbf{X})$ onto $\hat{G}(\mathbf{X})$, which will also be denoted by "exp," while we will use "log" to denote the inverse map.

The group $\hat{G}(\mathbf{X})$ is almost "a Lie group whose Lie algebra is $\hat{L}(\mathbf{X})$," but it fails to be a true Lie group, since it is infinite-dimensional. However, its truncated versions

$$G^N(\mathbf{X}) = \hat{\tau}^N(\hat{G}(\mathbf{X})) \tag{3.5}$$

are true Lie groups. (As for $\hat{G}(\mathbf{X})$ itself, it is a projective limit of the $G^N(\mathbf{X})$, but we will not make use of this fact.) Each $G^N(\mathbf{X})$ is a connected, simply connected, nilpotent Lie group, with Lie algebra $L^N(\mathbf{X})$, where

$$L^N(\mathbf{X}) = \tau^N(L(\mathbf{X})) = \hat{\tau}^N(\hat{L}(\mathbf{X})). \tag{3.6}$$

The exponential map from $L^N(\mathbf{X})$ to $G^N(\mathbf{X})$ is none other than the restriction of $\exp_N$ to $L^N(\mathbf{X})$ (which is a subset of $A_0^N(\mathbf{X})$). We will therefore also use $\exp_N$ to denote this map. Then $\exp_N$ is a bijection from $L^N(\mathbf{X})$ onto $G^N(\mathbf{X})$, whose inverse map will, as expected, be denoted by $\log_N$. Then $L^N(\mathbf{X})$ is the *free nilpotent Lie algebra of step* $N+1$ *in* $X_0, \cdots, X_m$, and we shall refer to the group $G^N(\mathbf{X})$ as the *free nilpotent Lie group of step* $N+1$ *infinitesimally generated by* $X_0, \cdots, X_m$.

Now suppose that we are given a C^∞ manifold M and an $(m+1)$-tuple $\mathbf{f} = (f_0, \cdots, f_m)$ of C^∞ vector fields on M. Each f_j is therefore a member of $D(M)$, the algebra of all partial differential operators $P: C^\infty(M) \to C^\infty(M)$. (Here $C^\infty(M)$ denotes the space of C^∞ real-valued functions on M.) There is therefore a well defined *evaluation map*

$$\mathrm{Ev}(\mathbf{f}): A(\mathbf{X}) \to D(M) \tag{3.7}$$

obtained by "plugging in the f_j for the X_j," so that

$$\mathrm{Ev}(\mathbf{f})\left(\sum_I a_I X_I\right) = \sum_I a_I f_I, \tag{3.8}$$

where, if $I = (i_1, \cdots, i_k)$, we write

$$f_I = f_{i_1} f_{i_2} \cdots f_{i_k}. \tag{3.9}$$

The image $\mathrm{Ev}(\mathbf{f})(A(\mathbf{X}))$ will be denoted by $A(\mathbf{f})$. Then $A(\mathbf{f})$ is the subalgebra of $D(M)$ generated by $f_0, \cdots, f_m$. The evaluation map $\mathrm{Ev}(\mathbf{f})$ can be restricted to $L(\mathbf{X})$. The corresponding map, which we will also denote by $\mathrm{Ev}(\mathbf{f})$, is a surjective homomorphism from $L(\mathbf{X})$ onto $L(\mathbf{f})$, where $L(\mathbf{f})$ is the Lie algebra of vector fields generated by $f_0, \cdots, f_m$.

The kernel of $\mathrm{Ev}(\mathbf{f}): A(\mathbf{X}) \to A(\mathbf{f})$ is the set of all *algebraic identities satisfied by* $f_0, \cdots, f_m$, and we will denote it by AI $(\mathbf{f})$. Similarly, the kernel of $\mathrm{Ev}(\mathbf{f}): L(\mathbf{X}) \to L(\mathbf{f})$ is the set of *Lie algebraic identities satisfied by* $f_0, \cdots, f_m$, and we denote it by LI $(\mathbf{f})$.

If p is a point in M, then we use $D_p(M)$ to denote the set of all partial differential operators at p, i.e. the quotient of $D(M)$ modulo the set of $P \in D(M)$ such that $(P\phi)(p)=0$ for every $\phi \in C^\infty(M)$. Also, we let $T_p(M)$ denote the tangent space of M at p. We then have the *evaluation at p map* $\mathrm{Ev}_p(\mathbf{f}): A(\mathbf{X}) \to D_p(M)$ given by $\mathrm{Ev}_p(\mathbf{f})(S) = (\mathrm{Ev}(\mathbf{f})(S))(p)$. The kernel of this map is the set of *algebraic relations among the* f_j *at* p, and will be denoted by $\mathrm{AR}(\mathbf{f}, p)$. Similarly, $\mathrm{Ev}_p(\mathbf{f})$ maps $L(\mathbf{X})$ to $T_p(M)$. The kernel of this map, denoted by $\mathrm{LR}(f, p)$, is the set of *Lie algebraic relations* (or, simply, *Lie relations*) among the f_j at p. (For instance, $[X_0, X_1]+X_2$ is a Lie identity satisfied by f_0, f_1, f_2 iff the vector field $[f_0, f_1]+f_2$ vanishes identically. Similarly, $[X_0, X_1]+X_2$ is a Lie relation among the f_j at p iff $[f_0, f_1]+f_2$ vanishes at p.)

The image $\mathrm{Ev}_p(\mathbf{f})(L(\mathbf{X}))$ is precisely the subspace $L(\mathbf{f})(p)$ of $T_p(M)$, where

$$L(\mathbf{f})(p)=\{V(p): V \in L(\mathbf{f})\}. \tag{3.10}$$

The system $\mathbf{f}$ satisfies the *Lie algebra rank condition* (LARC) at p if $L(\mathbf{f})(p)=T_p(M)$, i.e. if $\mathrm{Ev}_p(\mathbf{f})$ maps $L(\mathbf{X})$ onto $T_p(M)$.

The evaluation maps $\mathrm{Ev}(\mathbf{f})$, $\mathrm{Ev}_p(\mathbf{f})$ can formally be applied to series S in $\hat{A}(\mathbf{X})$, giving rise to formal infinite sums of partial differential operators (which, if $S \in \hat{L}(\mathbf{X})$, are vector fields). However, if one wishes to make sense of $\mathrm{Ev}(\mathbf{f})(S)$ as a mathematical object in a rigorous way, technical difficulties arise. (For instance, suppose that f_0, f_1 are C^∞ vector fields that satisfy $[f_0, f_1]=f_1$, and S is the Lie series $\sum_{k=0}^\infty (-1)^k(\mathrm{ad}\, X_0)^k(X_1)$. Should the general definition of $\mathrm{Ev}(\mathbf{f})(S)$ be such that, in this particular case, $\mathrm{Ev}(\mathbf{f})(S)$ is the zero series?) Rather than attempt to overcome these difficulties, we shall avoid them, by agreeing to refer to the series $\mathrm{Ev}(\mathbf{f})(S)$ (or $\mathrm{Ev}_p(\mathbf{f})(S)$) *only* as part of purely heuristic discussions which are not expected to be rigorous anyhow, *or* as part of statements that are given a precise mathematical translation. (For instance, the phrase "$\mathrm{Ev}_p(f)(S)$, applied to a function ϕ, is asymptotic to ..." will be translated into a collection of inequalities involving only the truncations $\mathrm{Ev}_p(\mathbf{f})(\hat{\tau}^N(S))$, in which only finite sums occur.)

We can also define truncated evaluation maps $\mathrm{Ev}^N(\mathbf{f})$, $\mathrm{Ev}_p^N(\mathbf{f})$ by restricting $\mathrm{Ev}(\mathbf{f})$ and $\mathrm{Ev}_p(\mathbf{f})$ to $A^N(\mathbf{X})$ or to $L^N(\mathbf{X})$. However, the algebra structure of $A^N(\mathbf{X})$ and the Lie algebra structure of $L^N(\mathbf{X})$ do not turn $A^N(\mathbf{X})$, $L^N(\mathbf{X})$ into subalgebras of $A(\mathbf{X})$, $L(\mathbf{X})$. This implies that $\mathrm{Ev}^N(\mathbf{f})$ need not be an algebra homomorphism from $A^N(\mathbf{X})$ to $A(\mathbf{f})$ or from $L^N(\mathbf{X})$ to $L(f)$. Also, the point evaluation maps $\mathrm{Ev}_p^N(\mathbf{f})$ are defined in an obvious way as maps from $A^N(\mathbf{X})$ to $D_p(M)$ and from $L^N(\mathbf{X})$ to $T_p(M)$.

If $\mathscr{E}_p: D(M) \to D_p(M)$ is the map $Q \to Q(p)$, then $\mathrm{Ev}_p^N(\mathbf{f}) = \mathscr{E}_p \circ \mathrm{Ev}^N(\mathbf{f})$. We use $\mathrm{AI}^N(\mathbf{f})$, $\mathrm{LI}^N(\mathbf{f})$, $\mathrm{AR}^N(\mathbf{f}, p)$, $\mathrm{LR}^N(\mathbf{f}, p)$ to denote, respectively, the kernels of the maps $\mathrm{Ev}^N(\mathbf{f}): A^N(\mathbf{X}) \to D(M)$, $\mathrm{Ev}^N(\mathbf{f}): \mathrm{L}^N(\mathbf{X}) \to L(\mathbf{f})$, $\mathrm{Ev}_p^N(\mathbf{f}): A^N(\mathbf{X}) \to D_p(M)$ and $\mathrm{Ev}_p^N(\mathbf{f}): L^N(\mathbf{X}) \to T_p(M)$. Then $\mathrm{AI}^N(\mathbf{f})$ is the set of *algebraic identities of degree* $\leqq N$ *among the* f_i, and similar self-explanatory names will be used for the other sets $\mathrm{LI}^N(\mathbf{f})$, $\mathrm{AR}^N(\mathbf{f}, p)$, $\mathrm{LR}^N(\mathbf{f}, p)$. Since, as indicated earlier, Ev^N need not be a homomorphism, the sets $\mathrm{AI}^N(\mathbf{f})$ may fail to be ideals of $A^N(\mathbf{X})$, and the $\mathrm{LI}^N(\mathbf{f})$ need not be ideals of $L^N(\mathbf{X})$. Also, $\mathrm{AR}^N(\mathbf{f}, p)$ can fail to be a subalgebra of $A^N(\mathbf{X})$, and $\mathrm{LR}^N(\mathbf{f}, p)$ may fail to be a Lie subalgebra of $L^N(\mathbf{X})$. (For instance, let $\mathbf{f}=(f_0, f_1, f_2)$, and suppose that $[f_0, f_1](p)=f_1(p)$, and $f_2(p)=0$. Then $[X_0, X_1]-X_1 \in \mathrm{LR}^2(\mathbf{f}, p)$ and $X_2 \in \mathrm{LR}^2(\mathbf{f}, p)$. If $\mathrm{LR}^2(\mathbf{f}, p)$ were a Lie subalgebra of $L^2(X_0, X_1, X_2)$, it would follow that $[[X_0, X_1], X_2]-[X_1, X_2]$ is in $\mathrm{LR}^2(\mathbf{f}, p)$, i.e. that $[X_1, X_2]$ is in $\mathrm{LR}^2(\mathbf{f}, p)$, since $[[X_0, X_1], X_2]=0$ in $L^2(X_0, X_1, X_2)$. So $[f_1, f_2](p)=0$. However, it is easy to construct f_0, f_1, f_2 that satisfy the conditions stated above as well as $[f_1, f_2](p) \neq 0$.)

As in [25], $\mathscr{U}_m$ will denote the set of all functions $u(\cdot)$ whose domain $\mathrm{Dom}(u(\cdot))$ is a compact interval of the form $[0, T]$, such that $u(\cdot)$ takes values in $\mathbb{R}^m$ and is

Lebesgue integrable on $[0, T]$. The time T is the *terminal time* of $u(\cdot)$ and is denoted by $T(u(\cdot))$. If $0 \leqq t \leqq T(u(\cdot))$, then the restriction of $u(\cdot)$ to $[0, t]$ is denoted by $u^t(\cdot)$. The components of $u(\cdot)$ are $u_1(\cdot), \cdots, u_m(\cdot)$, and we write $u_0(t) \equiv 1$.

If we consider the differential equation

$$\dot{S} = S\left(X_0 + \sum_{i=1}^{\infty} u_i X_i\right) \tag{3.11}$$

for an $\hat{A}(\mathbf{X})$-valued function $t \to S(t), 0 \leqq t \leqq T(u(\cdot))$, with the initial condition $S(0) = 1$, then the solution is

$$S(t) = \sum_I \left(\int_0^t u_I\right) X_I, \tag{3.12}$$

where $\int_0^t u_I$ is the iterated integral

$$\int_0^t u_I = \int_0^t \int_0^{\tau_k} \int_0^{\tau_{k-1}} \cdots \int_0^{\tau_2} u_{i_k}(\tau_k) u_{i_{k-1}}(\tau_{k-1}) \cdots u_{i_1}(\tau_1)\, d\tau_1 \cdots d\tau_k \tag{3.13}$$

if $\phi \neq I = (i_1, \cdots, i_k)$. (We let $\int_0^t u_\phi = 1$.)

The series $S(T(u(\cdot)))$, with $t \to S(t)$ given as above, is the *formal power series associated with the control* $u(\cdot)$, and will be denoted by $\operatorname{Ser}(u(\cdot))$. The mapping $\operatorname{Ser}: \mathcal{U}_m \to \hat{A}(\mathbf{X})$ is injective and, if $\mathcal{U}_m$ is regarded as a semigroup under the operation of concatenation, and $\hat{A}(\mathbf{X})$ is equipped with multiplication, then Ser is a semigroup homomorphism (cf. [25, Lemma 3.1]). Moreover, $\operatorname{Ser}(u(\cdot))$ is always an exponential Lie series (cf. [25, Prop. 3.1]), so that Ser actually takes values in $\hat{G}(\mathbf{X})$. The subsemigroup $\operatorname{Ser}(\mathcal{U}_m)$ of $\hat{G}(\mathbf{X})$ will be denoted by $\hat{S}(\mathbf{X})$. Since Ser is injective, one should think of $\hat{S}(\mathbf{X})$ as being just another way of realizing the control semigroup $\mathcal{U}_m$, which has the particular advantage of exhibiting $\mathcal{U}_m$ as embedded in a group.

If K is an arbitrary subset of $\mathbb{R}^m$, then we can consider $\mathcal{U}_m(K)$, the subsemigroup of $\mathcal{U}_m$ whose elements are the K-valued controls. The image of $\mathcal{U}_m(K)$ under Ser will be denoted by $\hat{S}(\mathbf{X}, K)$.

One can also consider the truncated versions of the map Ser and the semigroups $\hat{S}(\mathbf{X})$, $\hat{S}(\mathbf{X}, K)$. The truncation map $\hat{\tau}^N$ maps solutions of (3.11) to solutions of the same equation, regarded now as evolving in $A^N(X)$. Hence, if we let

$$\operatorname{Ser}_N(u(\cdot)) = \hat{\tau}^N(\operatorname{Ser}(u(\cdot))) \tag{3.14}$$

we find that

$$\operatorname{Ser}_N(u(\cdot)) = \sum_{|I| \leqq N} \left(\int_0^{T(u(\cdot))} u_I\right) X_I. \tag{3.15}$$

Moreover, $\operatorname{Ser}_N(u(\cdot)) \in G^N(\mathbf{X})$. The sets $\operatorname{Ser}_N(\mathcal{U}_m)$, $\operatorname{Ser}_N(\mathcal{U}_m(K))$ will be denoted by $S^N(\mathbf{X})$, $S^N(\mathbf{X}, K)$, respectively. Clearly, these subsets are subsemigroups of $G^N(\mathbf{X})$. Moreover, $S^N(\mathbf{X})$ is the set of points that can be reached from the identity element of $G^N(\mathbf{X})$ by trajectories of the system

$$\dot{S} = \tilde{F}_0^N(S) + \sum_{i=1}^{m} u_i \tilde{F}_i^N(S), \tag{3.16}$$

where $\tilde{F}_i^N$ is the restriction to $G^N(\mathbf{X})$ of the linear vector field F_i^N on $A^N(\mathbf{X})$, given by $F_i^N(S) = SX_i$. (It is clear that $\tilde{F}_i^N$ is tangent to $G^N(\mathbf{X})$, and therefore $\tilde{F}_i^N$ is well defined.) The Lie algebra of vector fields generated by $F_0^N, \cdots, F_m^N$ is isomorphic to $L^N(\mathbf{X})$ in an obvious way, and therefore acts transitively on $G^N(\mathbf{X})$. From this it follows, using general results from accessibility theory, that $S^N(\mathbf{X})$ has a nonempty interior relative to $G^N(\mathbf{X})$ and, moreover, this interior is dense in $S^N(\mathbf{X})$. More generally, $S^N(\mathbf{X}, K)$ is the reachable set from the identity corresponding to the system (3.16) with the additional control constraint $(u_1, \cdots, u_m) \in K$. The Lie algebra associated with this system is the Lie algebra $\Lambda^N(\mathbf{X}, K)$ generated by the vector fields $u \cdot F^N$, for $u \in K$, where we use the abbreviation $u \cdot F^N$ for $F_0^N + \sum_{i=1}^m u_i F_i^N$. (Recall that $u_0 = 1$.) Then $\Lambda^N(\mathbf{X}, K)$ acts transitively iff

$$\text{Aff}\,(K) = \mathbb{R}^m. \tag{3.17}$$

Since we are assuming that (3.17) holds, we can conclude that $\Lambda^N(\mathbf{X}, K)$ is indeed transitive. We then have:

LEMMA 3.1. *For every N,*

$$\emptyset \neq \mathring{S}^N(\mathbf{X}, K) \subseteq S^N(\mathbf{X}, K) \subseteq \text{Clos}\,\mathring{S}^N(\mathbf{X}, K).$$

(*Here* "$^\circ$," *and* "Clos" *mean interior and closure relative to* $G^N(\mathbf{X})$.)

The semigroup $S^N(\mathbf{X}, K)$ is the image of $\mathcal{U}_m(K)$ under the map Ser_N. We need nice inverses of this map, i.e. ways of selecting, for $S \in S^N(\mathbf{X}, K)$, a control $u_S(\cdot) \in \mathcal{U}_m(\mathbf{X}, K)$ which "depends smoothly on S" and is such that $\text{Ser}_N(u_S(\cdot)) = S$. The construction of such inverses was already done in [25]. However, we shall need a slightly stronger result, which we now state.

As in [25], we let Γ be any finite sequence $(\gamma^1, \cdots, \gamma^r)$ of points of $\mathbb{R}^m$, such that $\text{Aff}\,(\gamma^1, \cdots, \gamma^r) = \mathbb{R}^m$. We let $\mathbb{R}_+^k$ denote the set of k-tuples of nonnegative numbers. If $\mathbf{t} = (t_1, \cdots, t_k)$ is in $\mathbb{R}_+^k$, then we define $\{\Gamma, \mathbf{t}\}$ to be the piecewise constant control which is equal to γ^1 during the first t_1 units of time, then to γ^2 during time t_2, and so on. (This control is well defined even if $k > r$, because we extend the definition of γ^j to all positive integers j, by making $j \to \gamma^j$ periodic with period r, i.e., we let $\gamma^{r+1} = \gamma^1$, $\gamma^{r+2} = \gamma^2$, and so on.) Any control of the form $\{\Gamma, \mathbf{t}\}$ for some k and some $\mathbf{t} \in \mathbb{R}_+^k$ will be called a Γ-*control.* If $K \subseteq \mathbb{R}^m$ and Γ consists of elements of K, then Γ will be said to be a K-*sequence.*

The map $\nu_{k,\Gamma}^N$, defined by $\nu_{k,\Gamma}^N(\mathbf{t}) = \text{Ser}_N(\{\Gamma, \mathbf{t}\})$, takes $\mathbb{R}_+^k$ to $S^N(\mathbf{X})$. Moreover, if Γ is a K-sequence, then $\nu_{k,\Gamma}^N$ maps $\mathbb{R}_+^k$ to $S^N(\mathbf{X}, K)$. If $\mathbf{t}^0 \in \mathbb{R}_+^k$ is such that the differential $d\nu_{k,\Gamma}^N(\mathbf{t}_0)$ has rank equal to the dimension of $G^N(\mathbf{X})$, then the Γ-control $\{\Gamma, \mathbf{t}_0\}$ is said to be N-*normal.* Clearly, if Γ is a K-sequence and $\{\Gamma, \mathbf{t}_0\}$ is N-normal, then $\nu_{k,\Gamma}^N(\mathbf{t}_0) \in \mathring{S}^N(\mathbf{X}, K)$. Conversely, suppose that $S \in \mathring{S}^N(\mathbf{X}, K)$. We claim that $S = \nu_{k,\Gamma}^N(\mathbf{t}_0)$ for some K-sequence Γ and some N-normal Γ-control $\{\Gamma, \mathbf{t}_0\}$. To see this, observe first that the system (3.16), with the restriction $u \in K$, necessarily has the accessibility property from S, and the same is therefore true for the "backward system" whose trajectories are those of (3.16) run in reverse. It then follows from standard accessibility theory that, if U is any open subset of $G^N(\mathbf{X})$ containing S, then U contains a nonempty open set V such that, for the reverse system, every $S' \in V$ can be reached from S by means of a piecewise constant control. If we apply this with $U = \mathring{S}^N(\mathbf{X}, K)$, we get an open subset V of $\mathring{S}^N(\mathbf{X}, K)$ such that every $S' \in V$ can be steered to S by means of a trajectory of (3.16) that corresponds to a piecewise constant K-valued control. On the other hand, if $S' \in V$ then S' can be reached from the identity element $\mathbb{1}$ of $G^N(\mathbf{X})$ by means of some K-valued control. This control can be approximated by piecewise constant ones. Since V is open, we conclude that some $S' \in V$ is reachable from $\mathbb{1}$ by means of some piecewise constant control. This control is then necessarily of the form $\{\Gamma, \mathbf{t}^0\}$ for some sequence $\Gamma = (\gamma^1, \cdots, \gamma^k)$ and some $\mathbf{t}^0 = (t_1^0, \cdots, t_k^0) \in \mathbb{R}_+^k$ such that $t_j^0 > 0$ for all j. Since $\text{Aff}\,(K) = \mathbb{R}^m$, the sequence Γ can be assumed to be such that

$\mathrm{Aff}(\gamma^1, \cdots, \gamma^k) = \mathbb{R}^m$. (This may require that some new γ's be added at the end of Γ, and then the control $\{\Gamma, \mathbf{t}^0\}$ has to be continued by assigning positive times t_j to the new γ^j's. However, the t^j can be taken to be arbitrarily small, and then the new S' will still be in V, since V is open.) We then get a Γ-control $\{\Gamma, \mathbf{t}^0\}$ that steers $\mathbb{1}$ to an $S' \in V$, and is such that Γ is a K-sequence and the affine hull of the elements of Γ is $\mathbb{R}^m$. The proof of [25, Prop. 3.3] then implies that V contains a point S'' which is of the form $\nu^N_{l,\Gamma(\mathbf{t})}$ for some l and some N-normal Γ-control $\{\Gamma, \mathbf{t}\}$. (The proof of [25, Prop. 3.3] shows that, if $\Gamma = (\gamma^1, \cdots, \gamma^r)$ is such that $\mathrm{Aff}(\gamma^1, \cdots, \gamma^r) = \mathbb{R}^m$, then an N-normal Γ-control exists. This was shown by choosing an l and a $\mathbf{t}$ such that $d\nu^N_{l,\Gamma}(\mathbf{t})$ had the largest possible rank $\bar{\rho}$, and then constructing a submanifold M of $G^N(\mathbf{X})$ such that $\dim M = \bar{\rho}$, with the property that all the vector fields in the Lie algebra generated by the $\tilde{F}^N_j$ are tangent to M, from which it follows that $\bar{\rho} = \dim G^N(\mathbf{X})$. The same proof applies if we now choose $l, \mathbf{t}$ to be such that $d\nu^N_{l,\Gamma}(\mathbf{t})$ has the largest possible rank $\bar{\rho}$ among all $l, \mathbf{t}$ such that $\nu^N_{l,\Gamma}(\mathbf{t}) \in V$. Such an $l, \mathbf{t}$ exists because there is some $l, \mathbf{t}$ such that $\nu^N_{l,\Gamma}(\mathbf{t}) \in V$, namely, $l = k$ and $\mathbf{t} = \mathbf{t}^0$. The conclusion that $\bar{\rho} = \dim G^N(X)$ follows exactly as in [25].) If we now concatenate this N-normal control $\{\Gamma, \mathbf{t}\}$ with a piecewise constant control that steers S'' to S, it follows easily that the resulting control is a $\tilde{\Gamma}$-control for some $\tilde{\Gamma}$, and is N-normal. So, we have shown:

LEMMA 3.2. *Let $K \subseteq \mathbb{R}^m$, and let $S \in G^N(\mathbf{X})$. Then $S \in \mathring{S}^N(\mathbf{X}, K)$ if and only if there exist*

(a) *a K-sequence $\Gamma = (\gamma^1, \cdots, \gamma^r)$ such that* $\mathrm{Aff}(\gamma^1, \cdots, \gamma^r) = \mathbb{R}^m$,

(b) *a k and a $\mathbf{t} \in \mathbb{R}^k_+$ such that $\{\Gamma, \mathbf{t}\}$ is N-normal and $\nu^N_{k,\Gamma}(\mathbf{t}) = S$.*

The existence of "nice local inverses" to the map Ser_N follows easily.

COROLLARY 3.3. *Let $K \subseteq \mathbb{R}^m$, and let $S \in \mathring{S}^N(\mathbf{X}, K)$. Then there exist:*

(*a*) *a K-sequence $\Gamma = (\gamma^1, \cdots, \gamma^r)$ such that* $\mathrm{Aff}(\gamma^1, \cdots, \gamma^r) = \mathbb{R}^m$,

(b) *a positive integer k,*

(c) *an open subset W of $G^N(\mathbf{X})$ such that $S \in W$,*

(d) *a real analytic map $\psi: W \to \mathbb{R}^k_+$, such that*

$$\mathrm{Ser}_N(\{\Gamma, \psi(S')\}) = S' \quad \text{for all } S' \in W. \tag{3.18}$$

The proof is just a straightforward application of the Implicit Function theorem.

The group $\hat{G}(\mathbf{X})$ is the "Lie group" described at the beginning of this section. Formally, an element S of $\hat{G}(\mathbf{X})$ is an exponential of a Lie series in the indeterminates $X_0, \cdots, X_m$, and therefore $\mathrm{Ev}(\mathbf{f})(S)$ is the exponential of a vector field on M, i.e. a map from M to M. If $S \in \hat{S}(\mathbf{X}, K)$, then S can be thought of as a control, and $\mathrm{Ev}_p(\mathbf{f})(S)$ is the point of M to which p is steered by this control. Then $\hat{L}(\mathbf{X})$ is the "Lie algebra" of the Lie group $\hat{G}(\mathbf{X})$. Those elements $Z \in \hat{L}(\mathbf{X})$ such that $\mathrm{Ev}_p(\mathbf{f})(Z) = 0$ constitute the "isotropy subalgebra," and their exponentials are the "isotropy subgroup." The reachable set from p is $\mathrm{Ev}_p(\mathbf{f})(\hat{S}(\mathbf{X}, K))$. The Lie algebra rank condition says that $\hat{G}(\mathbf{X})$ "acts transitively on M near p." Hence p will be an interior point of the reachable set if the interior of $\hat{S}(\mathbf{X}, K)$ intersects the isotropy subgroup.

The preceding formal considerations are not rigorous, because $\hat{G}(\mathbf{X})$ is not a true Lie group and, as explained above, $\mathrm{Ev}(\mathbf{f})$ is not well defined on $\hat{G}(\mathbf{X})$. In order to obtain a rigorous local controllability theorem one has to consider the nilpotent approximations $G^N(\mathbf{X})$ to $\hat{G}(\mathbf{X})$. The $G^N(\mathbf{X})$ are true Lie groups, with Lie algebra $L^N(\mathbf{X})$, and the subsemigroups $S^N(\mathbf{X}, K)$ represent the nilpotent approximations to $\hat{S}(\mathbf{X}, K)$. Pursuing the analogy with our earlier discussion, we may think of $\mathrm{LR}^N(\mathbf{f}, p)$ as the "isotropy subalgebra" corresponding to the "action" of $G^N(\mathbf{X})$, and of $H^N(\mathbf{f}, p) = \exp_N(\mathrm{LR}^N(\mathbf{f}, p))$ as the "isotropy group." If N is large enough (so that $\mathrm{Ev}^N_p(\mathbf{f})(L^N(\mathbf{X}))$ is the whole tangent space T_pM), then the "action" of $G^N(\mathbf{X})$ on M

is transitive. So we might expect to be able to prove that, if the interior of $S^N(\mathbf{X}, K)$ intersects H^N, then p is in the interior of the reachable set from p. Also, it should follow that, if $\mathring{S}^N(\mathbf{X}, K) \cap H^N$ contains points reachable from the identity in arbitrarily small time, then $(M, \mathbf{f}, K)$ is STLC from p. However, this reasoning is not valid, since $\mathrm{Ev}^N(\mathbf{f})$ need not be a true Lie algebra homomorphism, $\mathrm{LR}^N(\mathbf{f}, p)$ need not be a Lie subalgebra of $L^N(\mathbf{X})$, and $G^N(\mathbf{X})$ does not really act on M. If $S \in \mathring{S}^N(\mathbf{X}, K) \cap H^N(\mathbf{f}, p)$ and we write $S = \exp_N(Z)$, then $\mathrm{Ev}_p^N(\mathbf{f})(Z) = 0$, and so $\mathrm{Ev}_p(\mathbf{f})(Z) = 0$. Therefore $\exp(Z)$ is equal to the identity map plus a series of differential operators that vanish at p. However, there is no reason for $\exp(Z)$ to be the series of a control $u(\cdot)$. What can be said is that $\exp_N(Z) = \mathrm{Ser}_N(u(\cdot))$ for some $u(\cdot)$. But then $\mathrm{Ser}(u(\cdot))$ will not necessarily be equal to $\exp(Z)$, although it will be equal to $\exp(Z)$ up to terms of degree N. So $u(\cdot)$ will not necessarily steer p to p. However, it will steer p to a point q which is close to p. If U is a neighborhood of Z in $L^N(\mathbf{X})$, and $\exp(U)$ is small enough so that $\exp(U) \subseteq S^N(\mathbf{X}, K)$, then one can choose a $u'(\cdot)$ such that $\mathrm{Ser}_N(u'(\cdot)) = \exp_N(Z')$ for each $Z' \in U$. Then the controls $u'(\cdot)$ will steer p to a neighborhood V of q. If U is large enough, then we may expect V to be such that $p \in V$. To make all this rigorous, we have to be able to choose $u'(\cdot)$ in a continuous fashion as a function of Z'. This requires that we confine ourselves to neighborhoods U such that, if $W = \exp(U)$, then there is a map ψ that satisfies the conditions of Corollary 3.3. So we define a *normal neighborhood* of a point $S \in \mathring{S}^N(\mathbf{X}, K)$ to be an open subset W of $G^N(\mathbf{X})$ such that there exist Γ, k, ψ for which the conditions of Corollary 3.3 hold. Then Corollary 3.3 simply says that every point of $\mathring{S}^N(\mathbf{X}, K)$ has a normal neighborhood. The sufficient condition for STLC from p will then say that, if $\mathring{S}^N(\mathbf{X}, K) \cap H^N(\mathbf{f}, p)$ contains points S_t reachable from the identity in arbitrarily small time t, then $(M, \mathbf{f}, K)$ is STLC from p, provided that N is sufficiently large, and that these points have normal neighborhoods whose size does not decrease too fast as $t \to 0$. It will be clear from the proof that it is not necessary to have a lower bound for the size of the neighborhood in all directions, but only in directions transversal to $H^N(\mathbf{f}, p)$. To make this precise, let $\mathscr{E} = (E_1, \cdots, E_k)$ be a finite sequence of elements of $L^N(\mathbf{X})$, and let $Z \in L^N(\mathbf{X})$. We define, for $r > 0$

$$B_{\mathscr{E}}(Z, r) = \left\{ Z + \sum_{i=1}^{k} x_i E_i : \sum_{i=1}^{k} x_i^2 \leqq r^2 \right\}. \tag{3.19}$$

($B_{\mathscr{E}}(Z, r)$ is the *$\mathscr{E}$-ball of radius r* and center Z. We will only use this definition for sequences $\mathscr{E}$ such that $E_1, \cdots, E_k$ are linearly independent.)

Also, we define a function $T^N : A^N(\mathbf{X}) \to \mathbb{R}$ by letting $T^N(S)$ be the coefficient of X_0 in S. (In particular, if $S = \mathrm{Ser}_N(u(\cdot))$ for some control $u(\cdot)$, then $T^N(S)$ is the terminal time of $u(\cdot)$.) We then have:

THEOREM 3.4. *Let $(M, \mathbf{f}, K)$ be a control system, and let $p \in M$. Assume that K is a bounded set. Let N be a positive integer, and let $\mathscr{E} = (E_1, \cdots, E_n)$ be a sequence of elements of $L^N(\mathbf{X})$ such that $(\mathrm{Ev}_p(\mathbf{f})(E_1), \cdots, \mathrm{Ev}_p(\mathbf{f})(E_n))$ is a basis of the tangent space T_pM. Assume that there is a sequence of points S_j, $j = 1, 2, \cdots$, such that:*

(i) *$S_j \in \mathring{S}^N(\mathbf{X}, K) \cap H^N(\mathbf{f}, p)$ for all j,*

(ii) *$T^N(S_j) \to 0$ as $j \to \infty$,*

(iii) *If $S_j = \exp_N(Z_j)$, then there are normal neighborhoods W_j of S_j, and a constant $\alpha > 0$, such that*

$$\log_N W_j \supseteq B_{\mathscr{E}}(Z_j, \alpha[T^N(S_j)]^N) \tag{3.20}$$

for all j. Then $(M, \mathbf{f}, K)$ is STLC from p.

Proof. Let $\rho_j = T^N(S_j)$. For each j, choose a K-sequence Γ_j, a positive integer k_j, and a real-analytic map $\psi_j: W_j \to \mathbb{R}_+^{k_j}$ such that

$$\mathrm{Ser}_N(\{\Gamma_j, \psi_j(S)\}) = S \tag{3.21}$$

whenever $S \in W_j$.

Choose coordinates on a neighborhood $\mathcal{N}$ of p such that p becomes $(0, \cdots, 0)$ and the vectors $\mathrm{Ev}_p(\mathbf{f})(E_i)$ are the members e_i of the canonical basis of $\mathbb{R}^n$. For each control $u(\cdot)$, let $\pi_p(u(\cdot))$ be the point to which $u(\cdot)$ steers p (i.e. $\pi_p(u(\cdot)) = x(T)$, if $u(\cdot)$ is defined on $[0, T]$, and $x(\cdot)$ is the trajectory for $u(\cdot)$ such that $x(0) = p$). Proposition 4.1 of [25] implies that the series $\mathrm{Ev}_p(\mathbf{f})(\mathrm{Ser}\, u(\cdot))$ gives an asymptotic expansion for $\pi_p(u(\cdot))$ in the following sense: if ϕ is an arbitrary C^∞ function on $\mathcal{N}$, then there are constants β_ν and times τ_ν such that

$$\|\phi(\pi_p(u(\cdot)) - \mathrm{Ev}_p(\mathbf{f})(\mathrm{Ser}_\nu(u(\cdot)))\phi\| < \beta_\nu T(u(\cdot))^{\nu+1} \tag{3.22}$$

for all ν and all controls $u(\cdot)$ such that $T(u(\cdot)) \leqq \tau_\nu$. (Here $\mathrm{Ev}_p(\mathbf{f})(\mathrm{Ser}_\nu(u(\cdot)))$ is a finite sum of partial differential operators evaluated at p, and so $\mathrm{Ev}_p(\mathbf{f})(\mathrm{Ser}_\nu(u(\cdot)))\phi$ is a finite sum of numbers, namely, the results of applying those partial differential operators to ϕ. The result from [25] gives constants β_ν that also depend on a bound A for the controls, but here we are assuming that K is bounded, so that β_ν only depends on ν.)

Inequality (3.22) clearly holds for vector functions as well, so we can apply it to the identity map $\phi: \mathcal{N} \to \mathcal{N}$. From now on, ϕ denotes this map. Therefore $\phi(\pi_p(u(\cdot)) = \pi_p(u(\cdot)))$, and so (3.22) becomes

$$\|\pi_p(u(\cdot)) - \mathrm{Ev}_p(\mathbf{f})(\mathrm{Ser}_\nu(u(\cdot)))\phi\| \leqq \beta_\nu T(u(\cdot))^{\nu+1}. \tag{3.23}$$

Now define maps μ_j from the closed unit ball $\mathbb{B}$ of $\mathbb{R}^n$ into M, by

$$\mu_j(x_1, \cdots, x_n) = \pi_p\left(\left\{\Gamma_j, \psi_j\left(\exp_N\left(Z_j + \alpha\rho_j^N \sum_{i=1}^n x_i E_i\right)\right)\right\}\right). \tag{3.24}$$

The definition is possible because $Z_j + \alpha\rho_j^N \sum_{i=1}^n x_i E_i$ is in $B_{\mathscr{E}}(Z_j, \alpha\rho_j^N)$, and so its exponential in $A^n(\mathbf{X})$ is in W_j. By construction, $\mu_j(x_1, \cdots, x_n)$ is reachable from p, by means of the control

$$u_{j,x_1,\cdots,x_n}(\cdot) = \left\{\Gamma_j, \psi_j\left(\exp_N\left(Z_j + \alpha\rho_j^N \sum_{i=1}^n x_i E_i\right)\right)\right\}. \tag{3.25}$$

The truncated series $\mathrm{Ser}_N(u_{j,x_1,\cdots,x_n}(\cdot))$ is then

$$\exp_N\left(Z_j + \alpha\rho_j^N \sum_{i=1}^n x_i E_i\right),$$

and so the terminal time $T(u_{j,x_1,\cdots,x_n}(\cdot))$ is equal to $T^N(\mathrm{Ser}_N(u_{j,x_1,\cdots,x_n}(\cdot)))$, i.e. to

$$\rho_j + \alpha\rho_j^N \sum_{i=1}^n x_i\theta_i,$$

where θ_i is the coefficient of X_0 in E_i. In particular, all the points $\mu_j(x_1, \cdots, x_n)$, for $(x_1, \cdots, x_n) \in \mathbb{B}$, and fixed j, are reachable from p in time not greater than $\hat{\alpha}\rho_j$, where $\hat{\alpha}$ is some fixed constant which does not depend on j. Since $\rho_j \to 0$ as $j \to \infty$, our theorem will be proved if we show that $\mu_j(\mathbb{B})$ contains a neighborhood of p for sufficiently large j.

In view of (3.23), we have

$$\left\|\mu_j(x_1,\cdots,x_n)-\mathrm{Ev}_p(\mathbf{f})\left(\exp_N\left(Z_j+\alpha\rho_j^N\sum_{i=1}^n x_iE_i\right)\right)\phi\right\|\leqq\beta_N\hat{\alpha}^{N+1}\rho_j^{N+1}; \tag{3.26}$$

provided that j is large enough, so that $\hat{\alpha}\rho_j\leqq\tau_N$.

For $Q\in L^N(\mathbf{X})$, define

$$\widetilde{\exp}_N(Q)=\sum_{k=0}^N\frac{1}{k!}Q^k. \tag{3.27}$$

(Here we identify $L^N(\mathbf{X})$ with the subspace $\sum_{k=1}^N L^{k,\mathrm{hom}}(\mathbf{X})$, but the powers Q^k are computed in $L(\mathbf{X})$, so that $\widetilde{\exp}_N(Q)$ is allowed to contain terms of degree greater than N.) We claim that all the coefficients of the finite series

$$\widetilde{\exp}_N\left(Z_j+\alpha\pi_j^N\sum_{i=1}^n x_iE_i\right)-\exp_N\left(Z_j+\alpha\rho_j^N\sum_{i=1}^n x_iE_i\right)$$

are bounded by a fixed constant times ρ_j^{N+1}. To see this, observe first that, if we write

$$Z_j=\sum_I z_I^jX_I, \tag{3.28}$$

then $z_I^j=0$ for $|I|>N$, and there is a constant c such that $|z_I^j|\leqq C\rho_j^{|I|}$ for all j, I. (Here $|I|$ is the length of the multiindex I, i.e. the degree of the monomial X_I. The first assertion follows because $Z_j\in L^N(\mathbf{X})$. The second one holds because $\exp_N(Z_j)=\mathrm{Ser}_N(u_j(\cdot))$ for some K-valued control $u_j(\cdot)$ with terminal time ρ_j. Since the coefficients σ_I of $\mathrm{Ser}(u_j(\cdot))$ are iterated integrals, as shown in (3.12) and (3.13), they satisfy bounds $|\sigma_I^j|\leqq\text{constant}\times\rho_j^{|I|}$. Similar bounds then hold for the coefficients of the series $\log(\mathrm{Ser}(u_j(\cdot)))$, and for those of its truncation $Z_j=\hat{\tau}^N(\log(\mathrm{Ser}(u_j(\cdot))))$.) The coefficients of $\alpha\rho_j^N\sum_{i=1}x_iE_i$ also satisfy a similar bound, since they all contain a factor ρ_j^N, and those of degree $>N$ vanish. So the coefficients of

$$\exp_N\left(Z_j+\alpha\rho_j^N\sum_{i=1}^n x_iE_i\right)\quad\text{and}\quad\widetilde{\exp}_N\left(Z_j+\alpha\rho_j^N\sum_{i=1}^n x_iE_i\right)$$

also satisfy these bounds, and then the same is true for those of the difference of these two series. However, the coefficients of this difference vanish whenever $|I|\leqq N$. Hence they are bounded by a constant times ρ_j^{N+1}.

It then follows that (3.26) remains valid (possibly with a different constant in the right side) if "$\exp_N$" is replaced by "$\widetilde{\exp}_N$." Now $\widetilde{\exp}_N(Z_j+\alpha\rho_j^N\sum_{i=1}^n x_iE_i)$ can be written out by applying (3.27) and then expanding the powers of $Z_j+\alpha\rho_j^N\sum_{i=1}^n x_iE_i$. This leads to a finite sum of terms of the following five kinds: (i) the term $\alpha\rho_j^N\sum_{i=1}^n x_iE_i$, (ii) powers of Z_j, (iii) products of at least one Z_j factor and at least one $\alpha\rho_j^N\sum_{i=1}^n x_iE_i$, (iv) powers of $\alpha\rho_j^N\sum_{i=1}^n x_iE_i$ other than the first power, (v) the identity.

When evaluated at p, all the terms $\mathrm{Ev}(\mathbf{f})(Z_j^k)$ vanish, because $Z_j\in\mathrm{LR}^N(\mathbf{f},p)$. The terms of type (iii) are $O(\rho_j^{N+1})$, because Z_j is $O(\rho_j)$. The terms of type (iv) are also $O(\rho_j^{N+1})$. Hence, modulo $O(\rho_j^{N+1})$, only the terms of types (i) and (v) count. So we get the bound

$$\left\|\mu_j(x_1,\cdots,x_n)-\mathrm{Ev}_p(\mathbf{f})\left(\mathbb{1}+\alpha\rho_j^N\sum_{i=1}^n x_iE_i\right)\phi\right\|\leqq\gamma\rho_j^{N+1} \tag{3.29}$$

for some constant γ. Then (3.29) implies (using the facts that $\mathrm{Ev}_p(\mathbf{f})(\mathbb{1})\phi = \phi(p) = p = 0$, and $\mathrm{Ev}_p(\mathbf{f})(E_i)\phi = e_i\phi = e_i$)

$$\left\| \mu_j(x_1, \cdots, x_n) - \alpha\rho_j^N \sum_{i=1}^n x_i e_i \right\| \leqq \gamma \rho_j^{N+1}. \tag{3.30}$$

Let

$$\nu_j(x_1, \cdots, x_n) = \frac{1}{\alpha\rho_j^N}\mu_j(x_1, \cdots, x_n). \tag{3.31}$$

Then $\mu_j(\mathbb{B})$ contains a neighborhood of 0 if $\nu_j(\mathbb{B})$ does. It follows from (3.30) that

$$\|\nu_j(x_1, \cdots, x_n) - (x_1, \cdots, x_n)\| \leqq \text{constant} \times \rho_j. \tag{3.32}$$

Therefore the ν_j are continuous maps from $\mathbb{B}$ to $\mathbb{R}^n$ that converge uniformly to the identity map of $\mathbb{B}$ as $j \to \infty$. This implies that $\nu_j(\mathbb{B})$ contains a neighborhood of 0 for large enough j. The proof is then complete.

Theorem 3.4 gives a sufficient condition for local controllability, but not one that is easy to check in practice. The next two sections will be devoted to providing more easily checkable conditions. Here we will just give a simple example that follows directly from Theorem 3.4.

THEOREM 3.5. *Let $(M, \mathbf{f}, K)$ be a control system, and let $p \in M$. Let N be a positive integer such that* $\mathrm{Ev}_p(\mathbf{f})(L^N(\mathbf{X})) = T_pM$. *Assume that $\mathring{S}^N(\mathbf{X}, K)$ contains an element $S = \exp_N(Z)$ such that all the homogeneous components of Z are in* $\mathrm{LR}(\mathbf{f}, p)$. *Then $(M, \mathbf{f}, K)$ is* STLC *from p.*

Proof. For each $p > 0$, let $\Delta(\rho)$ be the automorphism of $A(\mathbf{X})$ which sends X_i to ρX_i for $i = 0, \cdots, M$. Then $\Delta(\rho)$ gives rise to an automorphism $\hat{\Delta}(\rho)$ of $\hat{A}(\mathbf{X})$, defined by sending a series $S = \sum_{j=0}^\infty S_j$, with $S_j \in A^{j,\mathrm{hom}}(\mathbf{X})$ to the series

$$\hat{\Delta}(\rho)(S) = \sum_{j=0}^\infty \rho^j S_j. \tag{3.33}$$

Clearly, $\hat{\Delta}(\rho)$ induces an automorphism of $L(\mathbf{X})$, $\hat{L}(\mathbf{X})$, and $\hat{G}(\mathbf{X})$. Moreover, since $\hat{\Delta}(\rho)$ maps $\hat{A}_N(\mathbf{X})$ to $\hat{A}_N(\mathbf{X})$, it induces an automorphism $\Delta^N(\rho)$ of $A^N(\mathbf{X})$ and, in particular, an automorphism of $G^N(\mathbf{X})$ and one of $L^N(\mathbf{X})$.

Moreover, if $0 < \rho \leqq 1$, then $\hat{\Delta}(\rho)$ maps $S(\mathbf{X}, K)$ into $S(\mathbf{X}, K)$, and $\Delta^N(\rho)$ maps $S^N(\mathbf{X}, K)$ into $S^N(\mathbf{X}, K)$. (Indeed, if $t \to S(t)$, $0 \leqq t \leqq T$, is a solution of (3.11) corresponding to a K-valued control $t \to u(t) = (u_1(t), \cdots, u_m(t))$, then $\tau \to \hat{\Delta}(\rho)(S(\tau/\rho))$ is a solution of (3.11) on the interval $[0, \rho T]$, corresponding to the control $\tau \to u(\tau/\rho)$.)

Now suppose that S is an element of $\mathring{S}^N(\mathbf{X}, K)$ such that $S = \exp_N(Z)$, where

$$Z = \sum_{j=1}^N Z^j, \qquad Z^j \in L^{j,\mathrm{hom}}(\mathbf{X}), \tag{3.34}$$

and $Z \in \mathrm{LR}(\mathbf{f}, p)$. Pick $\mathscr{E} = (E_1, \cdots, E_n)$ such that

(a) the E_i are members of $L^N(\mathbf{X})$,

(b) each E_i is homogeneous of degree θ_i (with $\theta_i \leqq N$),

(c) $\mathrm{Ev}_p(\mathbf{f})(E_1), \cdots, E_p(\mathbf{f})(E_n)$ form a basis of T_pM.

Since $S \in \mathring{S}^N(\mathbf{X}, K)$, there exists a normal neighborhood W of S. If we let $S_\rho = \Delta^N(\rho)(S)$, $W_\rho = \Delta^N(\rho)(W)$, then W_ρ is a normal neighborhood of S_ρ. Let $\tilde{\alpha} > 0$ be such that $\exp_N(Z + \sum_{i=1}^n y_i E_i) \in W_\rho$ whenever $|y_i| \leqq \tilde{\alpha}$. Let $Z_\rho = \Delta(\rho)Z$. Then $\exp_N(Z_\rho + \sum_{i=1}^n \rho^{\theta_i} y_i E_i) \in W_\rho$ whenever $|y_i| \leqq \tilde{\alpha}$. Since $\theta_i \leqq N$, it follows that

$$B_{\mathscr{E}}(Z_\rho, \tilde{\alpha}\rho^N) \subseteq \log_N(W_\rho) \tag{3.35}$$

whenever $0 < \rho \leqq 1$.

Since $Z^j \in L^{j,\mathrm{hom}}(\mathbf{X}) \cap \mathrm{LR}(\mathbf{f}, p)$, it follows that $Z_\rho \in \mathrm{LR}(\mathbf{f}, p)$ for all ρ. Finally, it is clear that $T^N(S_\rho) = \rho c$, where $c = T^N(S)$. Therefore

(3.36) $$B_{\mathscr{g}}(Z_\rho, \alpha[T^N(S)]^N) \subseteq \log_N(W_\rho)$$

for $0 \leqq \rho \leqq 1$, if $\alpha = \tilde{\alpha} c^{-N}$. So the conditions of Theorems 3.4 hold, and our desired conclusion follows.

The preceding result is too weak for applications. In the following section we will strengthen it in two ways. First, the requirement that each homogeneous component Z^j of Z be a Lie relation at p will be replaced by the weaker condition that Z^j be equal to a Lie relation plus an element of lower degree. Second, the "degree" will be allowed to be a more general one, arising from a one-parameter group of dilations which is not necessarily the family $\{\Delta(\rho): \rho > 0\}$ considered in the proof of Theorem 3.5.

4. Dilations. We now define the concept of a "group of dilations," and prove a generalization of Theorem 3.5.

If V is a linear space over the reals, a *group of dilations* of V is a mapping $\rho \to \Delta(\rho)$ that assigns to every real $\rho > 0$ a linear endomorphism $\Delta(\rho): V \to V$, in such a way that

(DIL1) $\Delta(1) =$ *identity*,

(DIL2) $\Delta(\rho_1)\Delta(\rho_2) = \Delta(\rho_1\rho_2)$ *for all* ρ_1, ρ_1,

(DIL3) *V has a direct sum decomposition*

(4.1) $$V = \bigoplus_j V_j$$

such that the subspaces V_j are invariant under the $\Delta(\rho)$, and the action of $\Delta(\rho)$ on each V_j is given by multiplication by ρ_j^α for some $\alpha_j \geqq 0$.

The decomposition (4.1) is clearly unique if, in addition, we require that $\alpha_j \neq \alpha_k$ whenever $j \neq k$. In this case, the V_j are referred to as *the homogeneous components of V with respect to* Δ. If $v \in V$ is such that $v \in V_j$ for some j, then v is said to be Δ*-homogeneous.* If $v \neq 0$, then V_j is uniquely determined by v, and the corresponding α_j is the Δ*-degree of* v. More generally, any $v \in V$ can be expressed in a unique way as a sum $\sum_j v_j$, $v_j \in V_j$. The Δ*-degree of* v is the largest α_j such that $v_j \neq 0$, and is denoted by $\deg_\Delta(v)$.

If Δ is a group of dilations of V, then Δ gives rise to groups of dilations Δ^A of $A(V)$, the free associative $\mathbb{R}$-algebra generated by V (i.e. the tensor algebra over V) and Δ^L of $L(V)$, the free Lie algebra generated by V. In both cases, the new group of dilations consists of automorphism of the algebraic structure, which in addition leave invariant the usual homogeneous components of $A(V)$, $L(V)$ (i.e. the homogeneous components with respect to the groups of dilations induced by $\{\Delta_0(\rho): \rho > 0\}$, where $\Delta_0(\rho): V \to V$ is multiplication by ρ).

A group of dilations Δ of V will be called *strict* if it has no component of degree zero. If Δ is strict, then Δ^L is also strict, and Δ^A is strict on $A_0(V)$, the set of elements of $A(V)$ with no constant term. (But $\Delta^A(\rho)(1) = 1$ so Δ^A is not strict on $A(V)$.)

We will use $\nu(\Delta)$ to denote the infimum of the degrees of the homogeneous components of Δ. Then $\nu(\Delta) = \nu(\Delta^L)$ for every Δ. If V is finite-dimensional, then the infimum considered above is actually a minimum, and Δ is strict if and only if $\nu(\Delta) > 0$.

In the particular case when $V = L^{1,\mathrm{hom}}(\mathbf{X})$ (i.e. the linear span of $X_0, \cdots, X_m$), we let $\Delta_{1,m}(\rho): V \to V$ be multiplication by ρ as above. Let Δ be any group of dilations of V. Then Δ gives rise to groups of dilations Δ^A, Δ^L of $A(\mathbf{X})$ and $L(\mathbf{X})$. Any group of dilations of $A(\mathbf{X})$ or of $L(\mathbf{X})$ which arises in this fashion from a strict group of dilations of V will be called an *admissible group of dilations.* (Clearly, a group of dilations $\Delta^{\#}$ of $L(\mathbf{X})$, is admissible iff $\Delta^{\#}$ is strict and the $\Delta(\rho)$ are automorphisms which leave the usual homogeneous components invariant. If $\Delta^{\#}$ is a group of dilations of $A(\mathbf{X})$, then $\Delta^{\#}$ is admissible if and only if $\Delta^{\#}$ consists of automorphisms which leave the usual homogeneous components invariant, and the only elements of $\Delta^{\#}$-degree zero are the constants.)

If Δ is a strict group of dilations of $L^{1,\text{hom}}(\mathbf{X})$—so that Δ gives rise to admissible groups of dilations Δ^A, Δ^L— we will also refer to Δ itself as an admissible group of dilations.

In particular, the groups that arise from $\Delta_{1,m}$, denoted by $\Delta^A_{1,m}, \Delta^L_{1,m}$, are clearly admissible.

If Δ^A is any admissible group of dilations of $A(\mathbf{X})$ as above, arising from a group of dilations Δ of $L^{1,\text{hom}}(\mathbf{X})$, then every $\Delta^A(\rho)$ gives rise in an obvious way to an automorphism $\hat{\Delta}^A(\rho)$ of $\hat{A}(\mathbf{X})$. The $\hat{\Delta}^A(\rho)$ map $\hat{L}(\mathbf{X})$ to $\hat{L}(\mathbf{X})$ and therefore $\hat{G}(\mathbf{X})$ to $\hat{G}(\mathbf{X})$. Since $\Delta^A(\rho)$ maps $A_N(\mathbf{X})$ to $A_N(\mathbf{X})$ for each N, there are induced automorphisms $\Delta^{A,N}(\rho)$ of the algebra $A^N(\mathbf{X})$, which gives rise to automorphisms of the Lie algebra $L^N(\mathbf{X})$ and of the Lie group $G^N(\mathbf{X})$.

We will say that Δ is *compatible with the semigroup* $\hat{S}(\mathbf{X}, K)$ if

$$\hat{\Delta}^A(\rho)\hat{S}(\mathbf{X}, K) \subseteq \hat{S}(\mathbf{X}, K) \quad \text{for every } \rho \leqq 1. \tag{4.2}$$

Compatibility can be described more directly as follows. The map $\Delta(\rho)$ takes $L^{1,\text{hom}}(\mathbf{X})$ into itself. For $u \in \mathbb{R}^m$, $u=(u_1, \cdots, u_m)$, let $\mathbf{X}(u)=X_0+\sum_{i=1}^m u_iX_i$. If $u \in K$, then $\exp(\mathbf{X}(u)) \in \hat{S}(\mathbf{X}, K)$. Therefore $\hat{\Delta}^A(\rho)(\exp(\mathbf{X}(u)))$ must belong to $\hat{S}(\mathbf{X}, K)$, if Δ is compatible with $\hat{S}(\mathbf{X}, K)$ and $0<\rho\leqq 1$. That is, $\exp(\hat{\Delta}^A(\rho)(\mathbf{X}(u))) \in \hat{S}(\mathbf{X}, K)$ and so $\exp(\hat{\Delta}^A(\rho)(\mathbf{X}(u))) = \operatorname{Ser}(v(\cdot))$ for some $v(\cdot) \in \mathcal{U}_m(K)$. Let T be the terminal time of $v(\cdot)$. Then T cannot equal zero for, if $T=0$, then we would have $\operatorname{Ser}(v(\cdot))=\mathbb{1}$, and so $\hat{\Delta}^A(\rho)(\mathbf{X}(u))=0$, contradicting the fact that $\mathbf{X}(u)\neq 0$ and $\hat{\Delta}^A(\rho)$ is an automorphism. It follows from the construction of $\operatorname{Ser}(v(\cdot))$ that the coefficient of X_0 in $\operatorname{Ser}(v(\cdot))$ is precisely T. Moreover, the coefficient of X_0 in $\exp Z$ is the same as the coefficient of X_0 in Z. So

$$\hat{\Delta}^A(\rho)(\mathbf{X}(u)) = TX_0+\sum_{i=1}^m \alpha_iX_i, \tag{4.3}$$

for some choice of $(\alpha_1, \cdots, \alpha_m)\in\mathbb{R}^m$. If we let $\beta=\alpha/T$, we see that

$$\hat{\Delta}^A(\rho)(\mathbf{X}(u)) = T\mathbf{X}(\beta). \tag{4.4}$$

Let $v':[0, T]\to\mathbb{R}^m$ be such that $v'(t)=\beta$ for $0\leqq t\leqq T$. Then $\operatorname{Ser}(v'(\cdot))=\exp(T\mathbf{X}(\beta))$. Since Ser is injective as a map from $\mathcal{U}_m$ into $\hat{G}(\mathbf{X})$, we conclude that $v(\cdot)=v'(\cdot)$. Since $v(\cdot)$ is K-valued, we conclude that $\beta\in K$. Hence $\hat{\Delta}^A(\rho)(\mathbf{X}(u))$ is of the form $T\mathbf{X}(\beta)$ for some $T>0$ and some $\beta\in K$. Conversely, if $\{\Delta(\rho)\}$ has the property that $\Delta(\rho)(\mathbf{X}(u))$ is of the form $T\mathbf{X}(\beta)$ for some $T>0$, $\beta\in K$, whenever $0<\rho\leqq 1$ and $u\in K$, then it is easy to see that $\hat{\Delta}^A(\rho)\hat{S}(\mathbf{X}, K)\subseteq\hat{S}(\mathbf{X}, K)$ whenever $0<\rho\leqq 1$.

To see this, write

$$\Delta(\rho)(\mathbf{X}(u)) = \sum_{i=0}^m \theta_i(u)X_i \tag{4.5}$$

for $u\in\mathbb{R}^m$. Then

$$\Delta(\rho)(X_0)=\sum_{i=0}^m \theta_i(0)X_i \quad \text{and} \quad \Delta(\rho)\left(\sum_{j=1}^m u_jX_j\right)=\sum_{i=0}^m [\theta_i(u)-\theta_i(0)]X_i,$$

so that each of the functions $u \to \theta_i(u) - \theta_i(0)$ is linear. Moreover, $\theta_0(u) > 0$ for every $u \in K$ and, if we define $\beta_i(u) = \theta_i(u)/\theta_0(u)$ for $i = 1, \cdots, m$, $u \in \mathbb{R}^m$, $\theta_0(u) \neq 0$, then we have $\beta(u) \in K$ whenever $u \in K$, if $\beta(u) = (\beta_1(u), \cdots, \beta_m(u))$. Let $S \in \hat{S}(\mathbf{X}, K)$. Then $S = S(T)$, for some $\hat{A}(\mathbf{X})$-valued function $t \to S(t)$, $0 \leqq t \leqq T$, that satisfies

$$\dot{S}(t) = S(t)\left(X_0 + \sum_{i=1}^{m} u_i(t)X_i\right), \tag{4.6}$$

where the u_i are Lebesgue integrable functions such that $(u_1(t), \cdots, u_m(t)) \in K$ for all t. Let $S^{\#}(t) = \hat{\Delta}^A(\rho)S(t)$. Then

$$\begin{aligned} \dot{S}^{\#}(t) &= S^{\#}(t)\left(\theta_0(u(t))X_0 + \sum_{i=1}^{m} \theta_i(u(t))X_i\right) \\ &= S^{\#}(t)\left(\theta_0(u(t))\left[X_0 + \sum_{i=1}^{m} \beta_i(u(t))X_i\right]\right). \end{aligned} \tag{4.7}$$

Let $\tau(t) = \int_0^t \theta_0(u(s))\,ds$. (The integral exists because $u(\cdot)$ is Lebesgue integrable and θ_0 is affine linear.) Since $\theta_0(u(t)) > 0$ for $0 \leqq t \leqq T$, τ is a strictly increasing function of t. Let $S^*(\tau) = S^{\#}(t)$, if $\tau = \tau(t)$. Then

$$\dot{S}^*(\tau) = S^*(\tau)\left(X_0 + \sum_{i=1}^{m} v_i(\tau)X_i\right), \tag{4.8}$$

where the dot now denotes differentiation with respect to τ, and $v_i(\tau) = \beta_i(u(t))$ whenever $\tau = \tau(t)$. The vector-valued function $v(\cdot)$ is Lebesgue integrable. (Notice that $\beta(u(\cdot))$ might fail to be integrable, since we do not know that $\theta_0(\cdot)$ is bounded away from zero. However, $v(\cdot)$ is necessarily integrable because, whenever ϕ is a strictly positive integrable function on $[0, T]$, and $\tau(t) = \int_0^t \phi(s)\,ds$ for $0 \leqq t \leqq T$, then the function $g: [0, \tau(T)] \to R$ defined by $g(\tau(t)) = f(t)/\phi(t)$ is integrable whenever f is integrable.) Since $(v_1(\tau), \cdots, v_m(\tau)) \in K$ for every τ, we see that $S^*(\tau) \in \hat{S}(\mathbf{X}, K)$. In particular, since $S^*(\tau(T)) = S^{\#}(T) = \hat{\Delta}^A(\rho)S$, we see that $\hat{\Delta}^A(\rho)S \in \hat{S}(\mathbf{X}, K)$. Therefore $\hat{\Delta}^A(\rho)$ maps $\hat{S}(\mathbf{X}, K)$ to $\hat{K}(\mathbf{X}, K)$. So we have shown:

LEMMA 4.1. *Let $\Delta = \{\Delta(\rho): 0 < \rho < \infty\}$ be a one-parameter group of dilations of $V = L^{1,\mathrm{hom}}(\mathbf{X})$, and let Δ^A, $\hat{\Delta}^A$ be the corresponding groups of automorphisms of $A(\mathbf{X})$, $\hat{A}(\mathbf{X})$. Then Δ is compatible with the semigroup $\hat{S}(\mathbf{X}, K)$ if and only if $\Delta(\rho)$ $(X_0 + \sum_{i=1}^m u_iX_i)$ is of the form $T(X_0 + \sum_{i=1}^m v_iX_i)$ for some $T > 0$, $(v_1, \cdots, v_m) \in K$, whenever $0 < \rho \leqq 1$ and $(u_1, \cdots, u_m) \in K$.*

If $\hat{\Delta}^A$ is compatible with $\hat{S}(\mathbf{X}, K)$ then $\hat{\Delta}^A(\rho)$ gives rise to a map $\Delta_K^{\hat{S}}(\rho): \hat{S}(\mathbf{X}, K) \to \hat{S}(\mathbf{X}, K)$ whenever $0 < \rho \leqq 1$, and hence to a map $\Delta_K^{\mathcal{U}}(\rho): \mathcal{U}_m(K) \to \mathcal{U}_m(K)$, since $\mathcal{U}_m(K)$ is identified with $\hat{S}(\mathbf{X}, K)$ by means of the bijection Ser. An explicit description of this map follows from the reasoning preceding the statement of Lemma 4.1. If we write

$$\Delta(\rho)\left(X_0 + \sum u_iX_i\right) = \sum_{i=0}^{m} \theta_i^\rho(u)X_i, \tag{4.9}$$

for $u \in \mathbb{R}^m$, then the control $v(\cdot) = \Delta_K^{\mathcal{U}}(\rho)(u(\cdot))$ that corresponds to a given $u(\cdot) \in \mathcal{U}_m(K)$, defined on an interval $[0, T]$, is obtained from the K-valued map $t \to \beta^\rho(u(t))$, $0 \leqq t \leqq T$, by reparametrizing time, using $\tau = \tau(t) = \int_0^t \theta_0^\rho(u(s))\,ds$ as the new time parameter. (Here $\beta^\rho(u) = (\beta_1^\rho(u), \cdots, \beta_m^\rho(u))$, $\beta_i^\rho(u) = \theta_i^\rho(u)/\theta_0^\rho(u)$.) This explicit description implies, in particular, that $\Delta_K^{\mathcal{U}}(\rho)$ is continuous with respect to some natural topologies on $\mathcal{U}(T)$ (for example, L^1, pointwise convergence), and that $\Delta_K^{\mathcal{U}}(\rho)$ maps piecewise constant controls to piecewise constant controls. More precisely, if $u(\cdot)$ is a piecewise constant control whose values are $u^1, \cdots, u^k$, on intervals of length

$t^1, \cdots, t^k$, then $\Delta_K^{\mathcal{U}}(\rho)(u(\cdot))$ is piecewise constant with values $v^1, \cdots, v^k$ on intervals of length $\tau^1, \cdots, \tau^k$, where $v^i = \beta^\rho(u^i)$, and $\tau^i = \theta_0^\rho(u^i)t^i$.

This implies, in particular:

LEMMA 4.2. *If Δ is an admissible group of dilations of $L^{1,\text{hom}}(X)$, which is compatible with K, and W is a normal neighborhood of an $S \in \mathring{S}^N(\mathbf{X}, K)$, then $\Delta^{A,N}(\rho)(W)$ is a normal neighborhood $\Delta^{A,N}(\rho)(S)$ for every $\rho \in (0, 1]$.*

Now suppose that an $(m+1)$-tuple $\mathbf{f} = (f_0, \cdots, f_m)$ of smooth vector fields on a manifold M is given, as well as a point $p \in M$. We can then define $N_0(\mathbf{f}, p)$ to be the smallest integer N such that

$$\text{Ev}_p(\mathbf{f})(L^N(\mathbf{X})) = T_pM. \tag{4.10}$$

If, in addition, an admissible group of dilations Δ on $L^{1,\text{hom}}(\mathbf{X})$ is given, we can also define $\nu_0(\mathbf{f}, p, \Delta)$ to be the largest of the Δ-degrees of all the elements of $L^{N_0(\mathbf{f},p)}(\mathbf{X})$.

An element Z of $L(\mathbf{X})$ is said to be Δ-*neutralized* for $\mathbf{f}$ at p if each Δ-homogeneous component Z_j of Z is the sum of an $R_j \in L(\mathbf{X})$ which belongs to $\text{LR}(\mathbf{f}, p)$ and a $Q_j \in L(\mathbf{X})$ such that

$$\deg_\Delta(Q_j) < \deg_\Delta(Z_j). \tag{4.11}$$

Our generalization of Theorem 3.5 is then the following

THEOREM 4.3. *Let $(M, \mathbf{f}, K)$ be a control system, and let $p \in M$. Let Δ be an admissible group of dilations of $L^{1,\text{hom}}(\mathbf{X})$ which is compatible with $\hat{S}(\mathbf{X}, K)$. Let N be a positive integer that satisfies*

$$N \geqq N_0(\mathbf{f}, p), \tag{4.12}$$

and

$$N\nu(\Delta) \geqq \nu_0(\mathbf{f}, p, \Delta). \tag{4.13}$$

Assume that there exists an element Z of $L^N(\mathbf{X})$ which is Δ-neutralized for $\mathbf{f}$ at p and satisfies $\exp_N(Z) \in \mathring{S}^N(X, K)$. Then (M, f, K) is STLC *from p.*

Remark. Theorem 3.5 is a particular case of this result. Indeed, to get Theorem 3.5 it suffices to let Δ be the group of dilations defined by $\Delta(\rho)(P) = \rho P$ for $P \in L^{1,\text{hom}}(\mathbf{X})$. A Z that satisfies the condition of Theorem 3.5 is clearly Δ-neutralized for $\mathbf{f}$ at p.

Proof. Let $S \in \mathring{S}^N(\mathbf{X}, K)$ be such that $S = \exp_N(Z)$, $Z \in L^N(\mathbf{X})$, and Z is Δ-neutralized for $\mathbf{f}$ at p. Let $\mathscr{E} = (E_1, \cdots, E_n)$ consist of elements of $L^{N_0(\mathbf{f},p)}(\mathbf{X})$ which are Δ-homogeneous of degrees $\sigma_1, \cdots, \sigma_n$ and are such that the vectors $\text{Ev}_p(\mathbf{f})(E_i)$, $i = 1, \cdots, n$ span T_pM. (Then, in particular, $\sigma_j \leqq \nu_0(\mathbf{f}, p, \Delta)$ for $j = 1, \cdots, n$.) Let W be a normal neighborhood of S. Then we can pick a neighborhood W_0 of S and a $\beta > 0$ such that $\exp_N(Z' + \sum_{j=1}^n y_jE_j) \in W$ whenever $\exp_N(Z') \in W_0$ and $|y_j| \leqq \beta$ for $j = 1, \cdots, n$.

Since Z is Δ-neutralized for $\mathbf{f}$ at p, we can write

$$Z = \sum_i Z_i, \tag{4.14}$$

where the Z_i are elements of $L(\mathbf{X})$, are Δ-homogeneous of degree θ_i (with $\theta_i \neq \theta_j$ if $i \neq j$), and satisfy

$$Z_i = R_i + \sum_k Q_{ik}, \tag{4.15}$$

where the Q_{ik} are Δ-homogeneous of degree η_{ik}, the R_i belong to $\text{LR}(\mathbf{f}, p)$, and the η_{ik} satisfy $\eta_{ik} < \theta_i$.

It then follows that the Z_i belong to $L^N(\mathbf{X})$, for we can write $Z=\sum_{j=1}^N Z^j$, with $Z^j \in L^{j,\mathrm{hom}}(\mathbf{X})$, and then each Z^j is a sum of Δ-homogeneous components Z_k^j, which must necessarily belong to $L^{j,\mathrm{hom}}(\mathbf{X})$. The Z_i are then obtained by grouping together all the Z_k^j that have the same Δ-degree, and therefore belong to $L^N(\mathbf{X})$, as stated.

The Q_{ik} can also be assumed to belong to $L^N(\mathbf{X})$. Indeed, suppose that one Q_{ik} was not in $L^N(\mathbf{X})$. Since Q_{ik} is Δ-homogeneous, we must have

$$\deg_\Delta (Q_{ik}) \geqq (N+1)\times \nu(\Delta). \tag{4.16}$$

On the other hand, we can write

$$\mathrm{Ev}_p (\mathbf{f})(Q_{ik}) = \sum_l q_{ikl} \,\mathrm{Ev}_p (\mathbf{f})(E_l) \tag{4.17}$$

for appropriate coefficients q_{ikl}. Therefore

$$Q_{ik} = R_{ik} + \sum_l q_{ikl} E_l \tag{4.18}$$

where $R_{ik} \in \mathrm{LR}(\mathbf{f}, p)$. The $q_{ikl}E_l$ are Δ-homogeneous of degree σ_l. Since

$$\sigma_l \leqq \nu_0(f, p, \Delta) \leqq (N+1)\nu(\Delta) \leqq \eta_{ik} < \theta_i, \tag{4.19}$$

we can replace each Q_{ik} that occurs in (4.15) but does not belong to $L^N(\mathbf{X})$ by the sum of the $q_{ikl}E_l$, and add R_{ik} to R_i. This leads to an expression for Z_i for the form (4.15), with all the Q_{ik} in $L^N(\mathbf{X})$.

It then follows that the R_i are in $L^N(\mathbf{X})$ as well. Define

$$\tilde{Z}_\rho = Z - \sum_{ik} \rho^{\theta_i - \eta_{ik}} Q_{ik}. \tag{4.20}$$

Then $\exp_N (\tilde{Z}_\rho) \in W_0$ if ρ is small enough. Therefore, if ρ is small, W is a normal neighborhood of $\exp_N (\tilde{Z}_\rho)$ such that

$$\exp_N \left(\tilde{Z}_\rho + \sum_{i=1}^n y_i E_i \right) \in W \tag{4.21}$$

whenever $|y_i| \leqq \beta$ for $i=1, \cdots, n$. Let $Z_\rho = \Delta(\rho)(\tilde{Z}_\rho)$. Let $W_\rho = \Delta(\rho) W$. Then, if ρ is sufficiently small, W_ρ is a normal neighborhood of $\exp_N (Z_\rho)$ such that

$$\exp_N \left(Z_\rho + \sum_{i=1}^n y_i \rho^{\sigma_i} E_i \right) \in W_\rho \tag{4.22}$$

whenever $|y_i| \leqq \beta$ for $i=1, \cdots, n$. Let

$$S_\rho = \exp_N (Z_\rho). \tag{4.23}$$

Then

$$\log_N (W_\rho) \supseteq B_{\mathscr{E}}(Z_\rho, \beta \rho^{\nu_0(\mathbf{f}, p, \Delta)}). \tag{4.24}$$

On the other hand, S_ρ satisfies

$$T^N(S_\rho) \leqq c\rho^{\nu(\Delta)} \quad \text{for } 0 < \rho \leqq 1, \tag{4.25}$$

for some $c > 0$. (This is because $\tilde{Z}_\rho$ is a sum of Δ-homogeneous components, each of which has Δ-degree at least equal to $\nu(\Delta)$, and coefficients that are bounded as $\rho \to 0$. Therefore all the coefficients of $\Delta(\rho)\tilde{Z}_\rho$ are bounded by a constant times $\rho^{\nu(\Delta)}$. In particular, this is true for the coefficient of X_0, and so (4.25) follows.)

From (4.25) we conclude that

$$\rho^{\nu_0(\mathbf{f}, p, \Delta)} \geqq \rho^{N\nu(\Delta)} \geqq c^{-N}[T^N(S_\rho)]^N \tag{4.26}$$

and so

$$\log_N (W_\rho) \supseteq B_{\mathscr{E}}(Z_\rho, \alpha[T^N(S_\rho)]^N) \tag{4.27}$$

if ρ is sufficiently small, and $\alpha = \beta c^{-N}$.

Finally, we have

$$\begin{aligned} Z_\rho = \Delta(\rho)\tilde{Z}_\rho = \Delta(\rho)\Big(\sum_i \Big(Z_i - \sum_k \rho^{\theta_i - \eta_{ik}} Q_{ik}\Big)\Big) \\ = \sum_i \Big[\rho^{\theta_i} Z_i - \sum_k \rho^{\theta_i - \eta_{ik}} \rho^{\eta_{ik}} Q_{ik}\Big] \\ = \sum_i \rho^{\theta_i}\Big[Z_i - \sum_k Q_{ik}\Big] \\ = \sum_i \rho^{\theta_i} R_i. \end{aligned}$$

Therefore $Z_\rho \in \mathrm{LR}(\mathbf{f}, p)$. Hence all the hypotheses of Theorem 3.4 are satisfied, and the desired conclusion follows.

5. Invariant elements. Theorem 4.3 says that $(M, \mathbf{f}, K)$ is STLC from p if, for some sufficiently large N, $\mathring{S}^N(\mathbf{X}, K)$ contains an element S such that $\log_N (S)$ is Δ-neutralized for $\mathbf{f}$ at p. In order to be able to use this result, we need to know that $\mathring{S}^N(\mathbf{X}, K)$ necessarily will contain elements of some very special kind, for then STLC will follow if we hypothesize that these special elements are exponentials of Δ-neutralized members of $L^N(\mathbf{X})$.

To get these "special elements" we exploit a general result about existence of points that are invariant under certain finite groups of pseudoautomorphisms (cf. § 2 for the definition of "pseudoautomorphism").

Let L be a finite-dimensional, nilpotent Lie algebra over $\mathbb{R}$, and let G_L be its corresponding connected, simply connected Lie group. Then the exponential map $\exp: L \to G_L$ is a diffeomorphism onto. Therefore, if $\lambda: L \to L$ is an arbitray map, then λ gives rise to a map $\tilde{\lambda}: G_L \to G_L$, defined by letting

$$\tilde{\lambda}(\exp(z)) = \exp(\lambda(z)). \tag{5.1}$$

PROPOSITION 5.1. *Let L be a finite-dimensional, nilpotent Lie algebra over $\mathbb{R}$, and let G_L be the corresponding connected, simply connected Lie group. Let Λ be a finite group of pseudoautomorphisms of L, and let $\tilde{\Lambda} = \{\tilde{\lambda}: \lambda \in \Lambda\}$ be the group of bijections of G_L induced by Λ. Let S be a nonempty subset of G_L which is closed under multiplication. Suppose that every $\tilde{\lambda} \in \tilde{\Lambda}$ maps S into S. Then S contains an element s such that $\tilde{\lambda}(s) = s$ for all $\lambda \in \Lambda$.*

Proof. Start with an element $s_1 \in S$, and write $s_1 = \exp(b_1)$, where $b_1 \in L$. Let $\Lambda = \{\lambda_1, \cdots, \lambda_n\}$, with $\lambda_i \neq \lambda_j$ whenever $i \neq j$. Define s_2 by

$$s_2 = \tilde{\lambda}_1(s_1)\tilde{\lambda}_2(s_1) \cdot \cdots \cdot \tilde{\lambda}_n(s_1). \tag{5.2}$$

Then $s_2 \in S$, because $\tilde{\lambda}_j(s_1) \in S$ for each j, and S is closed under multiplication. On the other hand, we have

$$\tilde{\lambda}_j(s_1) = \exp(\lambda_j(b_1)) \quad \text{for } j = 1, \cdots, n. \tag{5.3}$$

Therefore the Campbell-Hausdorff formula gives

$$s_2 = \exp(z_2 + b_2) \tag{5.4}$$

where

$$z_2=\lambda_1(b_1)+\cdots+\lambda_n(b_1) \tag{5.5}$$

and $b_2\in[L]^2$. In view of (5.5), z_2 satisfies $\lambda(z_2)=z_2$ for every $\lambda\in\Lambda$. Assume we have proved, for some k, that there exists an $s_k\in S$ which is of the form $\exp(z_k+b_k)$, with $\lambda(z_k)=z_k$ for all $\lambda\in\Lambda$, and $b_k\in[L]^k$. Then we can define s_{k+1} by

$$s_{k+1}=\lambda_1(s_k)\lambda_2(s_k)\cdots\lambda_n(s_k) \tag{5.6}$$

and conclude from the Campbell-Hausdorff formula that

$$s_{k+1}=\exp(z_{k+1}+b_{k+1}), \tag{5.7}$$

where

$$z_{k+1}=\sum_{i=1}^{n}\lambda_i(z_k+b_k) \tag{5.8}$$

and b_{k+1} is a linear combination of terms, each of which is a Lie bracket of two or more elements of L of the form $\lambda_j(z_k+b_k)$ for some j. But $\lambda_j(z_k)=z_k$ and, if we let $b_{jk}=\lambda_j(b_k)$, we have $b_{jk}\in[L]^k$, because $b_k\in[L]^k$ and λ is a pseudoautomorphism. So the brackets that appear in b_{k+1} are brackets of two or more terms of the form z_k+b_{jk}. Now $[z_k+b_{ik}, z_k-b_{jk}]=[z_k,b_{jk}]+[b_{ik},z_k]+[b_{ik},b_{jk}]$, and so $[z_k+b_{ik},z_k+b_{jk}]\in[L]^{k+1}$. So $b_{k+1}\in[L]^{k+1}$. This proves, by induction, that an $s_k\in S$ of the desired form exists for every k. Since L is nilpotent, we can take k such that $[L]^k=\{0\}$. Then $b_k=0$, and so, if we let $s=s_k$, the condition that $\lambda(s)=s$ holds for all $\lambda\in\Lambda$.

6. End of the proof of Theorem 2.4. Assume that the conditions of Theorem 2.4 hold. Pick N so large that (4.12) and (4.13) hold. The group Λ obviously induces a group Λ_N of pseudoautomorphisms of the Lie algebra $L^N(\mathbf{X})$. The maps $\tilde{\lambda}$, for $\lambda\in\Lambda_N$, clearly map $\mathring{S}^N(\mathbf{X},K)$ into itself. The set $\mathring{S}^N(\mathbf{X},K)$ is nonempty and closed under multiplication. Proposition 5.1 then implies that $\mathring{S}^N(\mathbf{X},K)$ contains an element $S=\exp_N(Z)$, where $z\in L^N(\mathbf{X})$ is Λ_N-fixed. Then Z is Λ-fixed, and therefore Z is Δ-neutralized for $\mathbf{f}$ at p. Theorem 4.3 then says that $(M,\mathbf{f},K)$ is STLC from p.

7. Applications. In all the applications discussed here, Λ will be a group obtained from a group of automorphisms Λ_0 of $L(\mathbf{X})$, by adding to it the "time reversal" map. Precisely, let $\mathbb{T}^A: A(\mathbf{X})\to A(\mathbf{X})$ be the linear map which sends each monomial $X_{i_1}X_{i_2}\cdots X_{i_k}$ to the "reversed" monomial $X_{i_k}\cdots X_{i_2}X_{i_1}$. Then $\mathbb{T}^A$ is an antiautomorphism of $A(\mathbf{X})$ (i.e. $\mathbb{T}^A(PQ)=\mathbb{T}^A(Q)\mathbb{T}^A(P)$ for all P, Q in $A(\mathbf{X})$). It then follows easily that $\mathbb{T}^A([P,Q])=[\mathbb{T}^A(Q),\mathbb{T}^A(P)]$, i.e.

$$\mathbb{T}^A([P,Q])=-[\mathbb{T}^A(P),\mathbb{T}^A(Q)], \tag{7.1}$$

for P, Q in $A(\mathbf{X})$. Then $\mathbb{T}^A$ maps $L(\mathbf{X})$ to $L(X)$, and

$$\mathbb{T}(P)=(-1)^{1+k}P \quad \text{for } P\in L^{k,\mathrm{hom}}(\mathbf{X}), \tag{7.2}$$

where $\mathbb{T}$ denotes the restriction of $\mathbb{T}^A$ to $L(\mathbf{X})$.

It is clear that $\mathbb{T}$ is a pseudoautomorphism of $L(\mathbf{X})$. On the other hand $\mathbb{T}^A$ gives rise in an obvious way to a map $\hat{\mathbb{T}}^A:\hat{A}(\mathbf{X})\to\hat{A}(\mathbf{X})$. Clearly, $\hat{\mathbb{T}}^A(P^k)=P^k$ if $P\in A^{1,\mathrm{hom}}(\mathbf{X})$. Therefore

$$\hat{\mathbb{T}}^A(\exp P)=\exp P \tag{7.3}$$

if $P\in A^{1,\mathrm{hom}}(\mathbf{X})$. So, if $P_1,\cdots,P_k$ are elements of $L^{1,\mathrm{hom}}(\mathbf{X})$, we have

$$\hat{\mathbb{T}}^A(\exp(P_1)\cdots\exp(P_k))=\exp(P_k)\cdots\exp(P_1). \tag{7.4}$$

This implies that, if $u(\cdot)$ is a piecewise constant K-valued control, defined on $[0, T]$, then

$$\hat{\mathbb{T}}^A(\mathrm{Ser}\,(u(\cdot)) = \mathrm{Ser}\,(u^{\mathrm{rev}}(\cdot)) \tag{7.5}$$

where $u^{\mathrm{rev}}(t) = u(T-t)$ for $0 \leqq t \leqq T$. By an elementary continuity argument, (7.5) holds for all controls $u(\cdot)$. Therefore

$$\hat{\mathbb{T}}^A(\hat{S}(\mathbf{X}, K)) = \hat{S}(\mathbf{X}, K). \tag{7.6}$$

On the other hand, $\mathbb{T}$ gives rise to a map $\hat{\mathbb{T}} : \hat{L}(\mathbf{X}) \to \hat{L}(\mathbf{X})$, which is obviously equal to the restriction of $\hat{\mathbb{T}}^A$ to $\hat{L}(\mathbf{X})$. If P is any element of $\hat{A}_0(\mathbf{X})$, then $\hat{\mathbb{T}}^A(P^k) = [\hat{\mathbb{T}}^A(P)]^k$ for every k, and therefore

$$\hat{\mathbb{T}}^A(\exp\,(P)) = \exp\,(\hat{\mathbb{T}}^A(P)). \tag{7.7}$$

In particular, if $P \in \hat{L}(\mathbf{X})$, we get the equality

$$\hat{\mathbb{T}}^A(\exp\,(P)) = \exp\,(\hat{\mathbb{T}}(P)), \tag{7.8}$$

which implies

$$\hat{\mathbb{T}}^A(\hat{G}(\mathbf{X})) = \hat{G}(\mathbf{X}). \tag{7.9}$$

In the terminology of § 2 (cf. especially (2.7)), (7.8) shows that the restriction of $\hat{\mathbb{T}}^A$ to $\hat{G}(\mathbf{X})$ is precisely the map $\mathbb{T}^{\#} : \hat{G}(\mathbf{X}) \to \hat{G}(\mathbf{X})$. Hence (7.6) says that $\mathbb{T}^{\#}$ maps $\hat{S}(\mathbf{X}, K)$ to $\hat{S}(\mathbf{X}, K)$. So we have proved:

LEMMA 7.1. *$\mathbb{T}$ is an input symmetry.*

Now suppose that Λ_0 is a finite group of *graded* linear maps from $L(\mathbf{X})$ to $L(\mathbf{X})$. (A linear map $\lambda : L(\mathbf{X}) \to L(\mathbf{X})$ is *graded* if λ maps $L^{j,\mathrm{hom}}(\mathbf{X})$ into $L^{j,\mathrm{hom}}(\mathbf{X})$ for each j.) Then every $\lambda \in \Lambda_0$ commutes with $\mathbb{T}$. Since $\mathbb{T}^2$ is the identity map, the set

$$\Lambda = \Lambda_0 \cup \{\lambda\mathbb{T} : \lambda \in \Lambda_0\}, \tag{7.10}$$

is a finite group of pseudoautomorphisms. If Λ_0 is a group of input symmetries, then Λ is a group of input symmetries as well. We shall refer to the input symmetry $\mathbb{T}$ as "time reversal," and to the group Λ defined by (7.10) as "the augmentation of Λ_0 by time reversal."

Let us call an element of $L(\mathbf{X})$ *totally odd* if all its homogeneous components have odd degree. Then it is clear that the totally odd elements of $L(\mathbf{X})$ are precisely those $P \in L(\mathbf{X})$ that satisfy $\mathbb{T}(P) = P$. If Λ_0 is a finite group of graded linear maps of $L(\mathbf{X})$, and Λ is its augmentation by time reversal, then the Λ-fixed elements of $L(\mathbf{X})$ are precisely those $P \in L(\mathbf{X})$ that are Λ_0-fixed and totally odd. So we can conclude from Theorem 2.4 the following:

COROLLARY 7.2. *Let $(M, \mathbf{f}, K)$ be a control system, and let $p \in M$. Assume that* $\mathbf{f}$ *satisfies the* LARC *at p, and that there exist* (a) *an admissible group of dilations Δ of* $L^{1,\mathrm{hom}}(\mathbf{X})$ *which is compatible with* $\hat{S}(\mathbf{X}, K)$, (b) *a finite group Λ_0 of graded linear maps from $L(\mathbf{X})$ to $L(\mathbf{X})$ that are input symmetries, such that every totally odd Λ_0-fixed element of $L(\mathbf{X})$ is Δ-neutralized for* $\mathbf{f}$ *at p. Then* $(M, \mathbf{f}, K)$ *is* STLC *from p.*

7.1. Symmetric systems. A symmetric system is a family $\mathcal{V} = \{V_i : i \in I\}$ of vector fields on a manifold M, such that for every $i \in I$ there is a $j \in I$ such that $V_j = -V_i$. It is well known that, if a symmetric system satisfies the LARC at p, then the system is STLC from p. For completeness, we show that our theorem implies this result. First, it is clear that we can pick vector fields $f_1, \cdots, f_m$ in this family such that the m-tuple $(f_1, \cdots, f_m)$ satisfies the LARC at p. Then we can let $f_0 = 0$. Also, we take K to be the set of all points of $\mathbb{R}^m$ of the form $(0, 0, \cdots, 0, \pm 1, 0, \cdots, 0)$. We let Δ be the group of dilations such that $\Delta(\rho)(P)=\rho P$ for $P\in L^{1,\mathrm{hom}}(X)$, so that the Δ-degree is just the ordinary degree. We let Λ_0 be the group of automorphisms of $L(\mathbf{X})$ generated by $\lambda_1,\cdots,\lambda_m$, where λ_i is the automorphism that takes X_j to X_j for $j\neq i$, and X_i to $-X_i$. Since the λ_i commute, Λ_0 is finite. The Λ_0-fixed elements of $L(\mathbf{X})$ are those that are linear combinations of brackets where each X_i, $i=1,\cdots,m$, occurs an even number of times. Such a bracket cannot be totally odd unless it contains X_0. But then, when the bracket is evaluated by plugging in the f_j for the X_j, the result must be zero, because $f_0=0$. Hence every totally odd Λ_0-fixed element of $L(\mathbf{X})$ is actually in LR$(\mathbf{f},p)$, and therefore is Δ-neutralized for $\mathbf{f}$ at p. So we can apply Corollary 7.2 and conclude that $\Sigma=(M,\mathbf{f},K)$ is STLC from p. Since every trajectory of Σ is a trajectory of $\mathcal{V}$, the small-time local controllability of $\mathcal{V}$ follows.

7.2. The results of Brunovsky, Crouch and Byrnes. In [3], Brunovsky defined an *odd family* $\mathcal{V}=\{V_i: i\in I\}$ of vector fields on a symmetric neighborhood M of 0 to be a family such that for every $i\in I$ there is a $j\in I$ such that $V_j(-x)=-V_i(x)$ for $x\in M$. He then proved that, if $\mathcal{V}$ is odd and satisfies the LARC, then $\mathcal{V}$ is STLC from 0. Crouch and Byrnes [5] provided a coordinate-free generalization of this result. Suppose that $\mathcal{V}=\{V_i: i\in I\}$ is a collection of vector fields on a manifold M, and $p\in M$. Suppose that $\mathcal{V}$ satisfies the LARC at p. Assume that there is a finite group Λ_0 of diffeomorphisms of M such that

(i) each $\lambda\in\Lambda_0$ maps p to p,
(ii) each $\lambda\in\Lambda_0$ maps each V_i to some V_j in the family,
(iii) the differentials at p of the maps $\lambda\in\Lambda_0$ have no common invariant half space.

The result of [5] then says that $\mathcal{V}$ is small-time locally controllable from p. Brunovsky's theorem is a particular case of this, obtained by letting Λ_0 consist of the identity and the map $\lambda: x\to -x$. (If $V_j(-x)=-V_i(x)$, and λ_* denotes the differential of λ, so that $\lambda_*(v)=-v$, then λ_* maps V_i to V_j.)

We show that the result of [5] is a particular case of our Corollary 7.2. Let $f_1,\cdots,f_m$ be members of the family $\mathcal{V}$, chosen so that: (i) $(f_1,\cdots,f_m)$ satisfies the LARC, (ii) $f_i\neq f_j$ whenever $i\neq j$, (iii) the set $\{f_1,\cdots,f_m\}$ is mapped to itself by the maps λ_*, $\lambda\in\Lambda_0$. Let $f_0\equiv 0$, $\mathbf{f}=(f_0,\cdots,f_m)$. Then consider the system $(M,\mathbf{f},K)$, where $K=\{0\}\cup\tilde{K}$, and $\tilde{K}$ is the set of all vectors of $\mathbb{R}^m$ of the form $(0,0,\cdots,0,1,0,\cdots,0)$. If $\Sigma=(M,\mathbf{f},K)$ is STLC from p, then $\mathcal{V}$ is. (Indeed, let q be reachable from p by a trajectory of Σ that corresponds to a piecewise constant $u(\cdot):[0,T]\to K$. Then q can also be reached by a trajectory that corresponds to a $\tilde{K}$-valued control, in time $T'\leqq T$, by simply eliminating from $u(\cdot)$ all the pieces for which $u(\cdot)$ has the value 0.) The group Λ_0 acts on $L(\mathbf{X})$ for, if $\lambda\in\Lambda_0$, then λ_* permutes the elements of $\{f_1,\cdots,f_m\}$, and so we can define an automorphism $g(\lambda)$ of $L(\mathbf{X})$ by

$$g(\lambda)(X_0)=X_0 \tag{7.11}$$

and

$$g(\lambda)(X_i)=X_j \quad \text{if } \lambda_*(f_i)=f_j. \tag{7.12}$$

If $\mathcal{V}$ is any element of $L(\mathbf{X})$, it follows from (7.11) and (7.12) that

$$\mathrm{Ev}(\mathbf{f})(g(\lambda)(V))=\lambda_*(\mathrm{Ev}(\mathbf{f})(V)). \tag{7.13}$$

In particular, this implies that, if V is Λ_0-fixed, then the vector $\mathrm{Ev}_p(\mathbf{f})(V)$ is invariant under the differentials at p of all the maps $\lambda\in\Lambda_0$. Therefore $\mathrm{Ev}_p(\mathbf{f})(V)=0$, and so $V\in\mathrm{LR}(\mathbf{f},p)$.

So, if Δ is any group of dilations whatsoever, all the Λ_0-fixed elements of $L(\mathbf{X})$ are Δ-neutralized for $\mathbf{f}$ at p, and so Σ is STLC from p.

7.3. The Hermes condition and some generalizations. Consider a system

$$\dot{x} = f_0(x) + \sum_{i=1}^{m} u_i f_i(x), \qquad |u_i| \leqq 1, \tag{7.14}$$

and assume that p is an equilibrium point of f_0, i.e. that $f_0(p) = 0$. We let Λ_0 be the group of automorphisms of $L(\mathbf{X})$ generated by $\sigma_1, \cdots, \sigma_m$ and all the $\tilde{\pi}$, $\pi \in S_m$, where: (a) S_m is the group of permutations of $\{1, \cdots, m\}$, (b) for $\pi \in S_m$, $\tilde{\pi}$ is the automorphism of $L(\mathbf{X})$ which maps X_0 to X_0 and X_i to $X_{\pi(i)}$ for $i = 1, \cdots, m$, (c) σ_i is the automorphism that sends X_j to X_j for $j \neq i$, and X_i to $-X_i$. It is clear that Λ_0 is finite. The Λ_0-fixed elements are those that are linear combinations of elements of the form $\alpha(B)$, where B is a bracket of $X_0, \cdots, X_m$, and α is the Λ_0-symmetrization operator, i.e.

$$\alpha(V) = \sum_{\lambda \in \Lambda_0} \lambda(V). \tag{7.15}$$

It is clear that $\alpha(V) = 0$, if $\sigma_i(V) = -V$ for some i. Therefore, $\alpha(B) = 0$ if B is a bracket in which one of the X_i, $i > 0$, appears an odd number of times. Hence, in order to find the Λ_0-fixed elements, we may limit ourselves to considering the symmetrizations of brackets B where, for $i = 1, \cdots, m$, X_i appears an even number of times. For such a B, one may use the symmetrization operator β given by

$$\beta(V) = \sum_{\pi \in S_m} \tilde{\pi}(V). \tag{7.16}$$

Next, let $\theta_1, \cdots, \theta_m$ be arbitrary real numbers such that $\theta_i \geqq 1$ for $i = 1, \cdots, m$. Define $\Delta(\rho)$ by

$$\Delta(\rho): (X_0, \cdots, X_m) \to (\rho X_0, \rho^{\theta_1} X_1, \rho^{\theta_2} X_2, \cdots, \rho^{\theta_m} X_m). \tag{7.17}$$

Then Δ is compatible with $\tilde{S}(\mathbf{X}, K)$, where $K = \{(u_1, \cdots, u_m): |u_i| \leqq 1 \text{ for } i = 1, \cdots, m\}$. Then Corollary 7.2 implies that the system is STLC if, whenever B is a bracket with an odd number of X_0's, and an even number of X_i's for each $i \in \{1, \cdots, m\}$, it follows that every Δ-homogeneous component of $\beta(B)$ is equal, when evaluated at p, to a linear combination of brackets of lower Δ-degree. When the θ_i are different, this condition requires too much, for $\beta(B)$ will in general fail to be homogeneous. So the most interesting case obtains when all the θ_i are equal. Let $1 \leqq \theta < \infty$. Define the *θ-degree* δ_θ of a bracket $B \in \mathrm{Br}(\mathbf{X})$ to be the sum

$$\delta_\theta(B) = \delta^0(B) + \theta \sum_{i=1}^{m} \delta^i(B) \tag{7.18}$$

where $\delta^i(B)$ is the number of times that X_i occurs in B. Then δ_θ is the degree that arises, in an obvious way, from a group of dilations Δ_θ. We can then apply Corollary 7.2 to get a local controllability theorem involving the group Δ_θ. However, Δ_θ only enters the theorem via the concept of Δ-neutralization, and this concept is unchanged if we multiply all the degrees by a fixed number $\nu > 0$. Hence we can use, instead of δ_θ, the degree $\hat{\delta}_\theta$ defined by $\hat{\delta}_\theta(B) = (1/\theta)\delta_\theta(B)$, i.e.

$$\hat{\delta}_\theta(B) = \frac{1}{\theta}\delta^0(B) + \sum_{i=1}^{m} \delta^i(B). \tag{7.19}$$

The new definition now has the advantage that $\hat{\delta}_\theta$ also makes sense for $\theta=\infty$, in which case $\hat{\delta}_\infty(B)$ is, simply, the total number of occurrences in B of the X_i for $i=1,\cdots,m$. (But $\hat{\delta}_\infty$ does not arise from an admissible group of dilations.) We can then state the following

THEOREM 7.3. *Consider a system*

$$\dot{x}=f_0(x)+\sum_{i=1}^{m} u_i f_i(x), \qquad x\in M, \quad |u_i|\leqq 1, \tag{7.20}$$

and a point $p\in M$ such that $f_0(p)=0$. Assume that $(f_0,\cdots,f_m)$ satisfies the LARC *at p. Assume that there is a $\theta\in[1,\infty]$ such that, whenever $B\in \mathrm{Br}(\mathbf{X})$ is a bracket for which $\delta^0(B)$ is odd and $\delta^1(B),\cdots,\delta^m(B)$ are even, then there are brackets $C_1,\cdots,C_k$ in* $\mathrm{Br}(\mathbf{X})$ *such that*

$$\mathrm{Ev}_p(\mathbf{f})(\beta(B))=\sum_{i=1}^{k}\xi_i\,\mathrm{Ev}_p(\mathbf{f})(C_i) \tag{7.21}$$

for some $\xi_1,\cdots,\xi_k\in\mathbb{R}$, and

$$\hat{\delta}_\theta(C_i)<\hat{\delta}_\theta(B) \quad \textit{for } i=1,\cdots,m. \tag{7.22}$$

Then the system (7.20) *is* STLC *from p.*

For $\theta<\infty$, this theorem is just the result of applying Corollary 7.2 to our situation, using the group of dilations Δ_θ. To prove the theorem for $\theta=\infty$ we just "take the limit as $\theta\to\infty$." Rigorously, this means that, if the hypotheses are satisfied for $\theta=\infty$, then they are also satisfied for some large finite θ. To see this, let us use $S_{k,l}$ to denote the linear span of all the brackets B such that $\delta^0(B)\leqq k$ and $\delta^1(B)+\cdots+\delta^m(B)\leqq l$. If $\mathbf{l}$ is an m-tuple $(l_1,\cdots,l_m)$ of nonnegative integers, we define $\tilde{S}_{k,\mathbf{l}}$ to be the linear span of those B's for which $\delta^0(B)\leqq k$, $\delta^1(B)=l_1,\cdots,\delta^m(B)=l_m$. Also, we define

$$S_{\infty,l}=\bigcup_{k=0}^{\infty} S_{k,l}, \tag{7.23}$$

$$\tilde{S}_{\infty,\mathbf{l}}=\bigcup_{k=0}^{\infty}\tilde{S}_{k,\mathbf{l}}, \tag{7.24}$$

and we define spaces $S^{\text{odd}}_{\infty,l}$, $\tilde{S}^{\text{odd}}_{\infty,\mathbf{l}}$ in exactly the same way, except that the unions are only taken over odd values of k. We call the m-tuple $\mathbf{l}$ *even* if all its components $l_1,\cdots,l_m$ are even. Then the hypothesis of Theorem 7.3 for $\theta=\infty$ says that

$$\mathrm{Ev}_p(\mathbf{f})\beta(\tilde{S}^{\text{odd}}_{\infty,\mathbf{l}})\subseteq \mathrm{Ev}_p(\mathbf{f})(S_{\infty,|\mathbf{l}|-1}) \tag{7.25}$$

for all even $\mathbf{l}$. Pick an $\bar{l}$ such that

$$\mathrm{Ev}_p(\mathbf{f})(S_{\infty,\bar{l}})=T_pM. \tag{7.26}$$

For $l=0,1,\cdots,\bar{l}$, pick $k(l)$ such that

$$\mathrm{Ev}_p(\mathbf{f})(S_{\infty,l})=\mathrm{Ev}_p(\mathbf{f})(S_{k(l),l}). \tag{7.27}$$

Then pick $\theta\in[1,\infty)$ such that $k(l)<\theta$ for $l=0,\cdots,\bar{l}$. We claim that, with this choice of θ, the hypothesis of Theorem 7.3 holds. To see this, let $B\in\tilde{S}^{\text{odd}}_{\infty,\mathbf{l}}$. Assume first that $|\mathbf{l}|>\bar{l}$. Then $\mathrm{Ev}_p(\mathbf{f})(B)\in\mathrm{Ev}_p(\mathbf{f})(S_{\infty,\bar{l}})$, so that $\mathrm{Ev}_p(\mathbf{f})(B)\in\mathrm{Ev}_p(\mathbf{f})(S_{k(\bar{l}),\bar{l}})$. Therefore $\mathrm{Ev}_p(\mathbf{f})(B)$ is a linear combination of vectors $\mathrm{Ev}_p(\mathbf{f})(C_i)$, where the C_i are in $S_{k(\bar{l}),\bar{l}}$.

But then

$$\hat{\delta}_\theta(C_i) \leqq \frac{k(\bar{l})}{\theta} + \bar{l} < \bar{l} + 1 \leqq |\mathbf{l}| \leqq \hat{\delta}_\theta(B). \tag{7.28}$$

Next assume that $|\mathbf{l}| \leqq \bar{l}$ and $\mathbf{l}$ is even. By (7.25) and (7.27), $\mathrm{Ev}_p(\mathbf{f})(\beta(B))$ is a linear combination of vectors $\mathrm{Ev}_p(\mathbf{f})(C_i)$ with $C_i \in S_{k(\lambda),\lambda}$, where $\lambda = |\mathbf{l}| - 1$. But then

$$\hat{\delta}_\theta(C_i) \leqq \frac{k(\lambda)}{\theta} + \lambda < \lambda + 1 = |\mathbf{l}| < \hat{\delta}_\theta(B). \tag{7.29}$$

The proof of Theorem 7.3 is now complete.

Theorem 7.3 contains as a particular case a result for single-input systems was conjectured by H. Hermes and proved by us in [25]. Precisely, the Hermes condition (HC) for a system

$$\dot{x} = f(x) + ug(x), \qquad |u| \leqq 1, \tag{7.30}$$

at a point p, is the condition that, if B is an arbitrary bracket of f's and g's with an even number of g's, then $B(p)$ is a linear combination of values at p of brackets with fewer g's. (In particular, by taking $B = f$, we see that the HC implies that $f(p) = 0$.) The result proved in [25] says that, if the system (7.14) satisfies the LARC and the HC at p, then it is STLC from p. If we apply Theorem 7.3 with $m = 1$ (in which case, of course, the symmetrization operator β is just the identity), we obtain a strengthened version of the theorem of [25]. The HC corresponds to $\theta = \infty$ whereas Theorem 7.3 allows other values of θ. Moreover, even if we apply Theorem 7.3 with $\theta = \infty$, the condition that has to be satisfied to get controllability is weaker than the HC, and therefore the resulting controllability theorem is stronger. (The HC demands that *every* bracket with an even number of g's be neutralized, whereas our result only requires this for brackets with an even number of g's *and an odd number of* f's. As will be shown in examples below, these refinements make it possible to handle cases where the HC is unsufficient.)

For general m, R. Grossmann [8] states a sufficient condition for controllability, namely, that every bracket where each of the f_i for $i = 1, \cdots, m$ occurs an even number of times be expressible, at p, as a linear combination of brackets of lower total degree. This condition amounts to a weaker form of the case $\theta = 1$ of our theorem. (Theorem 7.3 only requires that the symmetrized brackets, which in addition have an odd number of f_0's, be expressible as linear combination of lower-degree elements.)

7.4. Low order sufficient conditions for systems with a cubic control set. We now illustrate the use of Theorem 7.3 by deriving some sufficient conditions for systems of the form (7.14), in terms of brackets of low degree. Assume that p is an equilibrium point of (7.14), i.e. that $f_0(p) = 0$. Also, assume that (7.14) satisfies the LARC at p. The simplest sufficient condition for STLC is the one obtained from the Pontryagin Maximum Principle, which says that (7.14) is STLC from p if, for every $\varepsilon > 0$, the adjoint equation along the trajectory $t \to x(t) = p$, $0 \leqq t \leqq \varepsilon$, has no nontrivial solution $t \to \lambda(t)$ such that $\langle \lambda(t), f_i(p)\rangle = 0$ for $0 \leqq t \leqq \varepsilon$. The adjoint equation in this case, written in coordinates, is simply the equation

$$\dot{\lambda} = -\lambda A, \tag{7.31}$$

where A is the Jacobian matrix of f_0 at 0. If $\lambda(\cdot)$ is a solution of (7.31) such that $\langle \lambda(t), f_i(p)\rangle \equiv 0$, then $\langle \lambda(0), A^k f_i(p)\rangle = 0$ for all k. Hence, if the vectors $A^k f_i(p)$ $i = 1, \cdots, m$, $k = 0, 1, \cdots$ span T_pM, there will not exist a $\lambda(\cdot)$ with the desired properties, and so

(7.14) will be STLC from p. Clearly, $A^k f_i(p) = (\text{ad} f_0)^k (f_i)(p)$. Therefore the sufficient condition obtained from the Maximum Principle simply says that (7.14) will be STLC from p if the vectors $(\text{ad} f_0)^k (f_i)(p)$, $i = 1, \cdots, m$, $k = 0, 1, \cdots$ span T_pM. Theorem 7.3 implies a stronger result, namely

PROPOSITION 7.4. *Assume that* $f_0(p) = 0$, *and the vectors* $(\text{ad} f_0)^k (f_i)(p)$, $i = 1, \cdots, m$, $k = 0, 1, \cdots$, *together with the vectors* $[f_i, f_j](p)$, $i, j \in \{1, \cdots, m\}$, *span* T_pM. *Then* (7.14) *is* STLC *from* p.

Proof. Our hypotheses imply in particular that (7.14) satisfies the LARC from p. Let $\mu > 0$ be such that the span of the vectors $(\text{ad} f_0)^k (f_i)(p)$, $i \in \{1, \cdots, m\}$, $k = 0, 1, \cdots$, is actually spanned by vectors of this same form with $k \leqq \mu$. Pick θ such that $\mu \leqq \theta < \infty$. Then T_pM is spanned by vectors $\text{Ev}_p(\mathbf{f})(B)$, where the B's are brackets such that $\hat{\delta}_\theta(B) \leqq 2$. On the other hand, if C is any bracket with an odd number of X_0's and an even number of X_i's for each $i \in \{1, \cdots, m\}$, then either $B = X_0$, in which case $\text{Ev}_p(\mathbf{f})(B) = 0$, or $\hat{\delta}_\theta(B) > 2$, in which case $\text{Ev}_p(\mathbf{f})(C)$ is certainly a linear combination of vectors $\text{Ev}_p(\mathbf{f})(B)$ with $\hat{\delta}_\theta(B) < \hat{\delta}(C)$. Hence the conditions of Theorem 7.3 are satisfied, and (7.14) is STLC from p.

If the sufficient condition of Proposition 7.4 is not satisfied, then it will be necessary to "neutralize" some brackets in order to be able to apply Theorem 7.3. The lowest total degree d where there may exist brackets to be neutralized is $d = 3$. (The case $d = 1$ is disposed of by the assumption that $f_0(p) = 0$.) The only brackets of total degree 3 where X_0 occurs an odd number of times, and each of the other X_i's an even number of times, are the expressions $[X_i, [X_i, X_0]]$. Symmetrization yields the element

$$H = \sum_{i=1}^{m} [X_i, [X_i, X_0]]. \tag{7.32}$$

We write $h = \text{Ev}(\mathbf{f})(H)$.

If h is "neutralized," in the sense that $h(p)$ is a linear combination of vectors $g_j(p)$, where the g_j are brackets of "lower degree," then that "releases" a whole collection of new brackets. If these brackets now span T_pM, then we get controllability again. Exactly which brackets are released by the neutralization of h will depend on how h is neutralized. Suppose that

$$h(p) = \sum_{i=1}^{m} \sum_{k=0}^{\nu} \alpha_{ik} (\text{ad} f_0)^k (f_i)(p) + \sum_{i=1}^{m} \sum_{j=1}^{m} \beta_{ij} [f_i, f_j](p) \tag{7.33}$$

for some choice of coefficients α_{ik}, β_{ij}.

Then, if we choose any θ such that $\theta \geqq 1$, $\theta > \nu - 1$, we see that $\text{Ev}_p(\mathbf{f})(H)$ is a linear combination of vectors $\text{Ev}_p(\mathbf{f})(B)$ with $\hat{\delta}_\theta(B) < \hat{\delta}_\theta(H)$. The next value of the total degree d for which there may be brackets B to be neutralized is $d = 5$. And the lowest possible value $\hat{\delta}_\theta(B)$ for such brackets is $2 + (3/\theta)$. If the brackets for which $\hat{\delta}_\theta < 2 + (3/\theta)$ span T_pM, the system will be STLC from p. So we get

PROPOSITION 7.5. *Assume that* (i) $f_0(p) = 0$, (ii) (7.33) *holds for some* ν *and some choice of coefficients* α_{ik}, β_{ij}. *Assume that there is a number* $\theta \in [1, \infty]$ *such that* $\theta > \nu - 1$, *with the property that the brackets B with* k_1 f_0*'s and* k_2 f_i*'s with* $i > 0$ *for all* k_1, k_2 *such that* $k_1 + \theta k_2 < 2\theta + 3$, *span* T_pM. *Then* (7.14) *is* STLC *from* p.

As a simple example, suppose that $h(p) = 0$ or, more generally, that (7.33) holds with $\nu = 1$. Then θ can be chosen to be an arbitrary number in $[1, \infty]$. In particular, we can conclude that the system (7.14) is STLC from p if either (i) T_pM is spanned by all the brackets of total degree $\leqq 4$ or (ii) T_pM is spanned by all the brackets with $\delta^+ = 1$, $\delta^0 \leqq 4$, together with those with $\delta^+ = 2$, $\delta^0 \leqq 2$, those with $\delta^+ = 3$ and $\delta^0 \leqq 1$, and those with $\delta^+=4$ and $\delta^0=0$, or (iii) T_pM is spanned by the brackets with $\delta^+=1$, $\delta^0\leqq 5$, together with those with $\delta^+=2$, $\delta^0\leqq 2$, and those with $\delta^+=3$, $\delta^0=0$, or (iv) T_pM is spanned by the brackets with $\delta^+=1$, δ^0 arbitrary, together with those with $\delta^+=2$, $\delta^0\leqq 2$. (Here, for a bracket B, $\delta^i(B)$ is the number of occurrences of f^i in B, and $\delta^+(B)=\sum_{i=1}^m \delta^i(B)$. The four results stated above are obtained by taking, respectively, $\theta=1$, $\theta=1.1$, $\theta=2.2$, and θ very large.) If (7.33) holds with $\nu=2$, then we have to choose $\theta>1$, and so we can conclude that (7.14) is STLC from p if (ii), (iii) or (iv) above hold. If (7.33) holds with $\nu=3$, then we must choose $\theta>2$, and we get that (7.14) is STLC from p if (iii) or (iv) hold. Finally, if (7.33) holds with some $\nu\geqq 4$, we get small-time local controllability if (iv) holds.

Notice, in particular, that if H is neutralized, in the sense that (7.33) holds for some ν, then this has the effect of unconditionally releasing a number of brackets, namely, all the brackets with $\delta^+=2$, $\delta^0\leqq 2$.

Finally, we illustrate the result of Theorem 7.3 in the case $m=2$, by giving some simple sufficient conditions in terms of brackets up to degree 6. The algebra to be considered here is $L(X_0, X_1, X_2)$. The homogeneous components $L^{j,\text{hom}}(X_0, X_1, X_2)$ have dimensions 3, 3, 8, 18, 48, 116, for $j=1$, 2, 3, 4, 5, 6, respectively. So there is a total of 196 potentially linearly independent brackets of degree $\leqq$6. After we eliminate those brackets that are totally even, or odd in either X_1 or X_2, and symmetrize, we are left with exactly eight linearly independent elements to be neutralized, namely, (a) X_0, (b) H, and (c) six elements B_1, B_2, B_3, B_4, B_5, B_6 of degree five, given by

$$B_1=[[X_1, X_2],[[X_1, X_2], X_0]], \tag{7.34}$$

$$B_2=\sum_{i=1}^{2}[X_i, (\text{ad } X_0)^3(X_i)], \tag{7.35}$$

$$B_3=\sum_{i=1}^{2}[[X_0, X_i],[X_0,[X_0, X_i]]], \tag{7.36}$$

$$B_4=\sum_{i=1}^{2}(\text{ad } X_i)^4(X_0), \tag{7.37}$$

$$B_5=[X_1,[X_1,[X_2,[X_0, X_2]]]]+[X_2,[X_2,[X_1,[X_0, X_1]]]], \tag{7.38}$$

$$B_6=[[X_0, X_1],[X_2,[X_1, X_2]]]+[[X_0, X_2],[X_1,[X_2, X_1]]]. \tag{7.39}$$

If we apply Theorem 7.3 with $\theta=1$, we can conclude that our system is STLC from p if the brackets of total degree $\leqq 6$ span T_pM, provided that (i) $f_0(p)=0$, (ii) $h(p)$ is a linear combination of values at p of brackets of degree <3, (iii) each vector $\text{Ev}_p(\mathbf{f})(B_i)$, $i=1,\cdots,6$, is a linear combination of values at p of brackets of degree <5.

7.5. Two single-input examples. We now analyze from the point of view of Theorem 7.3 two examples where the Hermes condition fails to hold but the system is STLC from p.

In [22], G. Stefani discusses an example of a system $\dot{x}=f(x)+ug(x)$ which is STLC from 0 even though (a) the Hermes condition is not satisfied, (b) the first bracket B needed to span the whole tangent space is one where g occurs four times. (This shows that not all brackets that are even in g are obstructions to local controllability.) We will show that Stefani's example fits the framework of our Theorem 7.3, since B is also even in f, and therefore there is no need for it to be neutralized. Stefani's example is the system $\dot{x}=u$, $\dot{y}=x$, $\dot{z}=x^3y$, in $\mathbb{R}^3$, with control constraint $|u|\leqq 1$. Then

$g=(1,0,0)$, $f=(0,x,x^3y)$. The relevant Lie brackets are as follows:

$$[g,f]=(0,1,3x^2y),\ [f,[g,f]]=(0,0,2x^3),$$
$$[g,[g,f]]=(0,0,6xy),\ [f,[f,[g,f]]]=0,$$
$$[g,[f,[g,f]]]=[f,[g,[g,f]]]=(0,0,6x^2),$$
$$[g,[g,[g,f]]]=(0,0,6y),$$
$$[g,[g,[f,[g,f]]]]=(0,0,12x),$$
$$[g,[g,[g,[f,[g,f]]]]]=(0,0,12).$$

(We omit brackets that vanish or that are trivially expressed in terms of the ones listed here.) In particular, the vectors $g(0)$ and $[g,f](0)$ span a two-dimensional space S. If B is any bracket of f's and g's of degree $\leqq 5$, then $B(0)\in S$. In particular, if we take $\theta=1$ in Theorem 7.3, we see that every bracket of degree 3 or 5 is equal, when evaluated at 0, to a linear combination of brackets of lower degree. If we now add the bracket $[g,[g,[g,[f,[g,f]]]]]$, which has degree 6, we span the whole space. Notice that, since this bracket is of even total degree, Theorem 7.3 does not require that it be neutralized. Hence Theorem 7.3 implies the fact—proved by Stefani—that this system is STLC from 0.

The preceding example shows that it is possible for controllability to be achieved thanks to the effect of some brackets that are even in g, so that not all such brackets are "obstructions." We now briefly review another example, already discussed in [25], which shows that a bracket which is even in g may be "neutralized" by a bracket with more g's but lower total degree. Consider the system $\dot{x}=u$, $\dot{y}=x$, $\dot{z}=x^3+y^2$, $|u|<1$. Here $f=(0,x,x^3+y^2)$ and $g=(1,0,0)$. The vectors $g(0)$ and $[f,g](0)$ span a two-dimensional space S, and all the $(\operatorname{ad} f)^k g(0)$, $k\geqq 2$, are in S. The vector $[g,[f,g]](0)$ belongs to S. However, $[g,[f,[f,[g,f]]]](0)$ is not in S, so that the Hermes condition fails to hold. On the other hand, $[g,[g,[g,f]]](0)$ is not in S either, and therefore we can apply Theorem 7.3 with $\theta=1$ and conclude that our system is STLC from 0.

7.6. Polynomial control systems. In [14], V. Jurdjevic studied control problems of the form

$$\dot{x}=P(x)+\sum_{i=1}^{m} u_i b_i, \qquad u=(u_1,\cdots,u_m)\in K, \tag{7.40}$$

where the state variable x takes values in $\mathbb{R}^n$, $P:\mathbb{R}^n\to\mathbb{R}^n$ is a polynomial map each of whose components is homogeneous of degree d, and K is either a cube centered at 0 (the "restricted controls" case) or the whole space $\mathbb{R}^m$ (the "unrestricted case").

Jurdjevic proved that, if d is odd, then (7.40) is STLC from 0 if and only if $S=\mathbb{R}^n$, where S is the smallest linear subspace of $\mathbb{R}^n$ which is invariant under the map P and contains $b_1,\cdots,b_m$. Moreover, in the unrestricted case it follows that every $x\in\mathbb{R}^n$ can be reached from 0 in time T, for every $T>0$. We show that this result follows from our general theorem. Actually, we show that it follows from Brunovsky's theorem on odd systems. First we observe that, if d is odd, then (7.40) is an "odd system" in Brunovsky's sense. Therefore, the characterization of small-time local controllability from 0 will follow if we show that the Lie algebra L generated by the vector fields $x\to P(x)$, $x\to b_i$, $x\to -b_i$ satisfies $L(0)=S$.

It follows from the definition of S that, if $x\in S$, then all the vectors b_i belong to S, and so does $P(x)$. Therefore all the members of L are tangent to S, and so $L(0)\subseteq S$. On the other hand, if $Q:\mathbb{R}^n\to\mathbb{R}^n$ is a polynomial map and v is a vector in $\mathbb{R}^n$, then the Lie bracket $[v, Q]$, of the constant vector field $x \to v$ and the vector field $x \to Q(x)$, is the vector field $x \to D_vQ(x)$, where $D_vQ(x)$ is the directional derivative of Q at x in the direction of v. In particular, this implies that $(\text{ad } v)^d(P)$ is the constant vector field $x \to P(v)$. Hence, if we let Σ be set of all vectors v such that the constant vector field $x \to v$ belongs to L, we see that Σ is invariant under the map P. Therefore $S \subseteq \Sigma$. This implies that $S \subseteq L(0)$, and so $L(0) = S$.

To complete the proof, we must show that, in the unrestricted case, the condition $S = \mathbb{R}^n$ implies that the time T reachable sets are equal to $\mathbb{R}^n$ for all $T > 0$. Assume that $S = \mathbb{R}^n$. Let $T > 0$. Then we already know that there is a neighborhood U of 0 that can be reached in time T. Let $p \in \mathbb{R}^n$. Pick r such that $0 < r < 1$ and $rp \in U$. Let $t \to x(t)$ be a trajectory such that $x(0) = 0$, $x(r^{1-d}T) = rp$. (Since rp is reachable in time T, $r^{1-d} \geqq 1$, and 0 is an equilibrium, it follows that rp is also reachable in time $r^{1-d}T$.) Let $y(t) = r^{-1}x(r^{1-d}t)$ for $0 \leqq t \leqq T$. Then $y(0) = 0$, $y(T) = p$. Let $u_1, \cdots, u_m$ be functions on $[0, r^{d-1}T]$ such that

$$\dot{x}(s) = P(x(s)) + \sum_{i=1}^{m} u_i(s) b_i \quad \text{for } 0 \leqq s \leqq r^{d-1}T.$$

Then

$$y(t) = r^d P(x(r^{1-d}t)) + \sum_{i=1}^{m} r^{-d} u_i(t) b_i,$$

i.e.

$$\dot{y}(t) = P(y(t)) + \sum_{i=1}^{m} r^{-d} u(t) b_i.$$

Since $t \to (r^{-d}u_1(t), \cdots, r^{-d}u_m(t))$ is an admissible control, it follows that p is reachable from 0 in time T.

7.7. Low order conditions with a general polyhedral control set. Consider a finite sequence $\mathcal{V} = (V_1, \cdots, V_m)$ of vector fields on a manifold M. We want conditions for $\mathcal{V}$ to be STLC from a point p. Equivalently, we want to know when the system

$$\dot{x} = \sum_{i=1}^{m} w_i V_i(x), \qquad w = (w_1, \cdots, w_m) \in J \tag{7.41}$$

is STLC from p, where J is the set of vectors $(0, 0, \cdots, 0, 1, 0, \cdots, 0) \in \mathbb{R}^m$.

Let $S_0(\mathcal{V}, p)$ denote the convex hull of the vectors $V_1(p), \cdots, V_m(p)$. Let $I_0(\mathcal{V}, p)$ denote the largest subset I of the index set $\{1, \cdots, m\}$, such that 0 is a convex combination of the vectors $V_i(p)$, $i \in I$, with strictly positive coefficients. In [24], we proved that, if (7.41) is STLC from p, then $I_0(\mathcal{V}, p)$ has to be nonempty and, moreover, there have to exist indices $i \in I_0(\mathcal{V}, p)$ such that $V_i(p) \neq 0$. The main result of [24] was a sufficient condition for (7.41) to be STLC from p. Let $S_1(\mathcal{V}, p)$ denote the convex hull of the vectors $V_i(p)$, $i \in \{1, \cdots, m\}$, $[V_i, V_j](p)$, $i, j \in I_0(\mathcal{V}, p)$. The result of [24] says that, if $S_1(\mathcal{V}, p)$ contains a neighborhood of the origin in the full tangent space T_pM, then (7.14) is STLC from p. We now show that this result, as well as some stronger conditions, can be derived from Corollary 7.2.

First, we observe that, instead of (7.41), we can consider the system

$$\dot{x} = \sum_{i=1}^{m} w_i V_i(x), \qquad w = (w_1, \cdots, w_m) \in K_m, \tag{7.42}$$

where K_m is the convex hull of J and the vector $(0, 0, \cdots, 0)$. (Indeed, let $\tilde{J}$ be the union of J and $\{(0, \cdots, 0)\}$. It is clear that, if small-time local controllability holds with J replaced by $\tilde{J}$, then it holds for the system (7.41). On the other hand, Proposition 2.3 says that $\tilde{J}$ can be replaced by its convex hull as well.)

Let us assume that the origin is an interior point of $S_1(\mathcal{V}, p)$. Also, let us relabel the indices so that $I_0(\mathcal{V}, p) = \{1, \cdots, \mu\}$, where $2 \leqq \mu \leqq m$.

Then we can express 0 as a convex combination

$$0 = \sum_{i=1}^{\mu} \lambda_i V_i(p), \quad \lambda_i > 0, \quad \sum_{i=1}^{\mu} \lambda_i = 1. \tag{7.43}$$

On the other hand, the hypothesis that 0 is an interior point of $S_1(\mathcal{V}, p)$ implies that: (i) *the vectors* $V_1(p), \cdots, V_m(p)$, *together with the* $[V_i, V_j](p)$, *for* $i = 1, \cdots, \mu$; $j = 1, \cdots, \mu$ *span the tangent space* T_pM, (ii) *it is possible to express* 0 *as a convex combination*

$$0 = \sum_{i=1}^{m} \alpha_i V_i(p) + \sum_{1 \leqq i < j \leqq \mu} \beta_{ij} [V_i, V_j](p) \tag{7.44}$$

where the α_i, β_{ij} *are strictly positive, and* $\Sigma \alpha_i + \Sigma \beta_{ij} = 1$.

(To see that (ii) follows, pick $\delta > 0$ so small that $-\delta Z(p) \in S_1(\mathcal{V}, p)$ whenever $Z = V_i$ for some i or $Z = [V_i, V_j]$ for some $i, j \in \{1, \cdots, \mu\}$. Then each $-\delta V_i(p)$, $i \in \{1, \cdots, m\}$, can be written as a convex combination of the $V_j(p)$, $j \in \{1, \cdots, m\}$, and the $[V_j, V_k](p)$, $j, k \in \{1, \cdots, \mu\}$. So 0 can be written as a linear combination of these same vectors in which all the coefficients are nonnegative and the coefficient of $V_i(p)$ is strictly positive. The same is true for $[V_i, V_j](p)$, if $i, j \in \{1, \cdots, \mu\}$. If we then add all these expressions and divide by the sum of the coefficients, we obtain an expression of the desired form.)

Now define $f_0 \equiv 0$, $f_i = \lambda_i V_i$ for $i = 1, \cdots, \mu$, $f_i = \alpha_i V_i$ for $i = \mu + 1, \cdots, m$. Consider the system

$$\dot{x} = f_0(x) + \sum_{i=1}^{m} u_i f_i(x), \qquad u = (u_1, \cdots, u_m) \in K_m. \tag{7.45}$$

It is clear that every trajectory of (7.45) is a trajectory of (7.42). Hence it suffices to prove that (7.45) is STLC from p.

To prove that (7.45) is STLC from p, we work with the free Lie algebra $L(X_0, \cdots, X_m)$. We let Λ_0 be the group of all automorphisms g_π of $L(X_0, \cdots, X_m)$ that are induced by a permutation π of the indices $\{0, \cdots, m\}$ that satisfies $\pi(0) = 0$, $\pi(\{1, \cdots, \mu\}) = \{1, \cdots, \mu\}$. It is clear that all the g_π are input symmetries. We define dilations $\Delta(\rho)$ by assigning Δ-degree one to $X_0, \cdots, X_\mu$, and Δ-degree δ to $X_{\mu+1}, \cdots, X_m$, where δ is some number such that $2 < \delta < 3$.

We now show that all the totally odd Λ_0-fixed elements of $L(X_0, \cdots, X_m)$ are Δ-neutralized for $\mathbf{f}$ at p. Since $f_0(p), \cdots, f_m(p)$, and the $[f_i, f_j](p)$ with $i, j \in \{1, \cdots, \mu\}$ span T_pM, it is clear that T_pM is spanned by the evaluations at $\mathbf{f}, p$ of elements of $L(X_0, \cdots, X_m)$ of Δ-degree not greater than δ. Among these, the Λ_0-fixed elements of Δ-degree one are spanned by X_0 and $X_1 + \cdots + X_\mu$. But

$$f_0(p) = (f_1 + \cdots + f_\mu)(p) = 0, \tag{7.46}$$

and so these elements are neutralized. The elements of Δ-degree 2 are totally odd and therefore need not be considered. The Λ_0-fixed elements of Δ-degree δ are spanned by $X_{\mu+1} + \cdots + X_m$, and (7.44) shows that $f_{\mu+1}(p) + \cdots + f_m(p)$ is equal to the value at p of an element of Δ-degree $< \delta$. Hence Corollary 7.2 can be applied. This completes the proof that the result of [24] is a particular case of Corollary 7.2.

It should be clear from the preceding proof that one can get more sophisticated results by just applying the same method. A detailed analysis of what can be so obtained will be the subject of a future paper. At the moment, we limit ourselves to two examples. In these examples, if $\lambda_1, \cdots, \lambda_\mu$ are such that (7.43) holds, we let g_1, g_2 denote the vector fields

$$g_1 = \lambda_1 V_1 + \cdots + \lambda_\mu V_\mu, \tag{7.47}$$

$$g_2 = \sum_{i=1}^{\mu} \lambda_i^2 (\operatorname{ad} V_i)^2 (g_1). \tag{7.48}$$

(That is, $g_1 = f_1 + \cdots + f_\mu$, $g_2 = \sum_{i=1}^{\mu} [f_i, [f_i, g_1]]$.)

Then, if (7.43) holds, the system (7.41) is STLC from p if one of the following conditions holds:

(I) (a) *$g_2(p)$ is a linear combination of the vectors $V_i(p)$, $i \in \{1, \cdots, \mu\}$ and the $[V_i, V_j](p)$, $i, j \in \{1, \cdots, \mu\}$,*
(b) *0 is a convex combination, with strictly positive coefficients, of the $V_i(p)$, $i \in \{1, \cdots, m\}$, the $[V_i, V_j](p)$, $i, j \in \{1, \cdots, \mu\}$, and the $[V_i, [V_j, V_k]](p)$, $i, j, k \in \{1, \cdots, \mu\}$,*
(c) *T_pM is spanned by the $V_i(p)$, $i \in \{1, \cdots, m\}$, the $[V_i, V_j](p)$, $i \in \{1, \cdots, m\}$, $j \in \{1, \cdots, \mu\}$, the $[V_i, [V_j, V_k]](p)$, $i, j, k \in \{1, \cdots, \mu\}$, and the $[V_i, [V_j, [V_k, V_l]]](p)$, $i, j, k, l \in \{1, \cdots, \mu\}$;*

(II) (a) *0 is a convex combination with strictly positive coefficients of the $V_i(p)$, $i \in \{1, \cdots, m\}$ and the $[V_i, V_j](p)$, $i, j \in \{1, \cdots, \mu\}$,*
(b) *$g_2(p)$ is a linear combination of the $V_i(p)$, $i \in \{1, \cdots, m\}$ and the $[V_i, V_j](p)$, $i, j \in \{1, \cdots, \mu\}$,*
(c) *T_pM is spanned by the $V_i(p)$, $i \in \{1, \cdots, m\}$, the $[V_i, V_j](p)$, $i, j \in \{1, \cdots, m\}$, the $[V_i, [V_j, V_k]](p))$, where i, j, k are in $\{1, \cdots, \mu\}$, and the $[V_i, [V_j, [V_k, V_l]]](p)$, $i, j, k, l \in \{1, \cdots, \mu\}$.*

To see that (I) implies small-time local controllability from p, we reason as before, but with $3 < \delta < 4$. Condition (c) of (I) says that T_pM is spanned by the brackets of Δ-degree not greater than $1 + \delta$. The Λ_0-fixed elements of Δ-degree one are spanned by X_0 and $X_1 + \cdots + X_\mu$, which are obviously neutralized, since $f_0 \equiv 0$ and (7.43) holds. The Λ_0-fixed elements of Δ-degree 2 do not matter, because they are totally even. In Δ-degree 3 there is only one Λ_0-fixed element, namely G_2, where we let $G_1 = X_1 + \cdots + X_\mu$, $G_2 = \sum_{i=1}^{\mu} [X_i, [X_i, G_1]]$. Condition (a) then says that this element is neutralized. In Δ-degree δ there is one Λ_0-fixed element, namely, $X_{\mu+1} + \cdots + X_m$. Condition (b) then says that this element is neutralized. Finally, all the elements of Δ-degree 4 or $1 + \delta$ are totally even, and therefore need not be considered.

To see the sufficiency of (II) we again use the same argument, with δ such that $2 < \delta < 2.5$. Then condition (c) says that T_pM is spanned by evaluations of brackets of Δ-degree not greater than 2δ. The possible Δ-degrees of such brackets are 1, 2, δ, 3, $1 + \delta$, 4, $2 + \delta$ and 2δ. The only Δ-degrees where totally odd brackets occur are 1, δ, 3 and $2 + \delta$. In Δ-degree 1, the Λ_0-fixed elements are X_0 and $X_1 + \cdots + X_\mu$, which are neutralized. In degree δ, the Λ_0-fixed element is $X_{\mu+1} + \cdots + X_m$, which is neutralized by condition (a). In Δ-degree 3, the Λ_0-fixed element is G_2, which is neutralized by condition (b). So Corollary 7.2 applies, and small-time local controllability follows.

8. Conclusion. The main implication of the results proven here is that, so far, one method appears to suffice to prove most known small-time local controllability results. It seems to us that this method is still very special, and it should be possible to obtain better results by making a more detailed analysis of the semigroups $S^N(\mathbf{X}, K)$.

One important application of the theory developed in this paper is to the problem of High Order Optimality Conditions. Small-time local controllability is a particular instance of this general problem, in which we are concerned with finding sufficient conditions for a particular trajectory (given by $x(t) \equiv p = \text{constant}$) to lie in the interior of the attainable set from p, which is the same as finding necessary conditions for the trajectory to lie on the boundary of the reachable set. The methods of this paper can be used to prove results on the construction of control variations for more general trajectories, and to obtain necessary conditions for optimality. The results will be reported in subsequent papers.

Appendix.

Proof of Proposition 2.3. Clearly, all that needs to be shown is that, if $\tilde{\Sigma}$ is STLC from p, it follows that Σ is STLC_{pc} from p. Suppose that $\tilde{\Sigma}$ is STLC from p. Let $T > 0$. Pick T' such that $0 < T' < T$, and let U be an open set such that $p \in U$ and $U \subseteq \text{Reach}(\tilde{\Sigma}, \leqq T', p)$. Shrink U, if necessary, so that $L(\mathbf{f})(q)$ is the full tangent space at q for every $q \in U$. Let $q \in U$. Let $\mathcal{F}_\Sigma$ be the family of vector fields associated with Σ, and let $-\mathcal{F}_\Sigma = \{-V: V \in \mathcal{F}_\Sigma\}$. Then $L(\mathbf{f}) = L(\mathcal{F}_\Sigma) = L(-\mathcal{F}_\Sigma)$. Let $\mathcal{H}$ be the family of restrictions to U of the members of $-\mathcal{F}_\Sigma$. Then $\mathcal{H}$ has the AP from q. So there is a nonempty open subset W of U such that every $r \in W$ is reachable from q by an $\mathcal{H}$-trajectory in time not greater than $T - T'$, so that q is reachable from every $r \in W$ by an $\mathcal{F}_\Sigma$-trajectory in time not greater than $T - T'$.

Now pick an $r \in W$. Since $W \subseteq U$, r is reachable from p in time τ, for some $\tau \in [0, T']$ by means of a trajectory of $\tilde{\Sigma}$ that corresponds to a control $u(\cdot): [0, \tau] \to \tilde{K}$. Then $u(\cdot)$ can be approximated in $L^1([0, \tau], \mathbb{R}^m)$ by a sequence $\{u_n(\cdot)\}$ of piecewise constant K-valued controls. If $x_n(\cdot)$ is the trajectory for $u_n(\cdot)$ such that $x_n(0) = p$, and we let $r_n = x_n(\tau)$, then $r_n \in W$ for sufficiently large n. Therefore W contains a point r' which is reachable from p in time τ by means of a piecewise constant $\tilde{K}$-valued control. Let $\text{co}(K)$ denote the convex hull of K. Since $\tilde{K}$ is the closure of $\text{co}(K)$, every piecewise constant $\tilde{K}$-valued control can be approximated in $L^1([0, \tau], \mathbb{R}^m)$ by piecewise constant $\text{co}(K)$-valued controls. Therefore W must contain a point r'' which is reachable from p in time τ by means of a piecewise constant $\text{co}(K)$-valued control. Finally, a piecewise constant $\text{co}(K)$-valued control can be approximated weakly by piecewise constant K-valued controls. Hence W contains a point r''' which is reachable from p in time τ by a piecewise constant K-valued control. So $r''' \in \text{Reach}_{pc}(\Sigma, \leqq T', p)$. Since $q \in \text{Reach}_{pc}(\Sigma, \leqq (T - T'), r''')$, we conclude that $q \in \text{Reach}_{pc}(\Sigma, \leqq T, p)$. Since q was an arbitrary point of U, we see that $U \subseteq \text{Reach}_{pc}(\Sigma, \leqq T, p)$. Since U is open, $p \in U$, and T was an arbitrary positive number, it follows that Σ is STLC_{pc} from p.

Acknowledgments. We thank an anonymous referee for very helpful suggestions and remarks. Also, we are especially grateful to Klaus Wagner who read the paper with great care and pinpointed a large number of misprints as well as some mathematical inaccuracies.

REFERENCES

[1] A. AGRACHEV AND R. GAMKRELIDZE, *The exponential representation of flows and the chronological calculus*, Math. USSR Sbornik, 35 (1979), pp. 727–785.

[2] R. BROCKETT, *Volterra series and geometric control theory*, Automatica, 12 (1976), pp. 167–176.

[3] P. BRUNOVSKY, *Local controllability of odd systems*, Banach Center Publications, Warsaw, Poland, 1 (1974), pp. 39–45.

[4] K. J. CHEN, *Integration of paths, geometric invariants and a generalized Baker–Hausdorff formula*, Ann. Math., 65 (1957), pp. 163–178.

[5] P. CROUCH AND C. BYRNES, *Symmetries and local controllability*, to appear.
[6] M. FLIESS, *Fonctionnelles causales non linéaires et indetéterminées non commutatives*, Bull. Soc. Math. France, 109 (1981), pp. 3-40.
[7] ———, *Développements fonctionnels en indéterminées non commutatives des solutions d'équations différentielles non linéaires forcées*, C. R. Acad. Sci. Paris, Ser. A, 287 (1978), pp. 1133-1135.
[8] R. GROSSMANN, Unpublished doctoral dissertation, Princeton Univ., Princeton, NJ, 1985.
[9] R. HERMANN, *On the accessibility problem in control theory*, in International Symposium on Nonlinear Differential Equations and Nonlinear Mechanics, Academic Press, New York, 1963, pp. 325-332.
[10] H. HERMES, *Lie algebras of vector fields and local approximation of attainable sets*, this Journal, 16 (1978), pp. 715-727.
[11] ———, *On local controllability*, this Journal, 20 (1982), pp. 211-220.
[12] ———, *Local controllability and sufficient conditions in singular problems*, J. Differential Equations, 20 (1976), pp. 213-232.
[13] ———, *Control systems which generate decomposable Lie algebras*, J. Differential Equations, 44 (1982), pp. 166-187.
[14] V. JURDJEVIC, *Polynomial control systems*, Proc. 22nd IEEE CDC, Vol. 2, 1983, pp. 904-906.
[15] A. KRENER, *Local approximation of control systems*, J. Differential Equations, 19 (1975), pp. 125-133.
[16] ———, *A generalization of Chow's theorem and the bang-bang theorem to nonlinear control systems*, this Journal, 12 (1974), pp. 43-52.
[17] E. B. LEE AND L. MARKUS, *Foundations of Optimal Control Theory*, John Wiley, New York, 1976.
[18] C. LESIAK AND A. KRENER, *The existence and uniqueness of Volterra series for nonlinear systems*, IEEE Trans. Automat. Control, 23 (1978), pp. 1090-1095.
[19] C. LOBRY, *Controlabilité des systèmes non linéaires*, this Journal, 8 (1970), pp. 573-605.
[20] J. P. SERRE, *Lie Algebras and Lie Groups*, W. A. Benjamin, New York, 1965.
[21] G. STEFANI, *On local controllability and related topics*, preprint, Facoltà di Ingegneria, Università di Firenze.
[22] ———, *Local properties of nonlinear control systems*, preprint, Facoltà di Ingegneria, Università di Firenze.
[23] H. J. SUSSMANN AND V. JURDJEVIC, *Controllability of nonlinear systems*, J. Differential Equations, 12 (1972), pp. 95-116.
[24] H. J. SUSSMANN, *A sufficient condition for local controllability*, this Journal, 16 (1978), pp. 790-802.
[25] ———, *Lie brackets and local controllability: a sufficient condition for scalar-input systems*, this Journal, 21 (1983), pp. 686-713.
[26] ———, *Lie brackets, real analyticity and geometric control*, in Differential Geometric Control Theory, R. W. Brockett, R. S. Millman and H. J. Sussmann, eds., Birkhauser, Boston, 1983, pp. 1-116.
[27] G. VIENNOT, *Algèbres de Lie Libres et Monoides Libres*, Lecture Notes in Mathematics 691, Springer-Verlag, New York, 1978.