High-Order Small-Time Local Controllability

Matthias Kawski *
Department of Mathematics
Arizona State University
Tempe, Arizona 85287

May 18, 1987

Abstract

In this paper we give a survey of the recent progress in the search for both necessary and sufficient conditions for small-time local controllability of affine control systems, and discuss in detail various high-order phenomena – which are some of the obstacles a more complete theory has to overcome – and exhibit the techniques which allow one to show (non-)controllability of individual such systems, thereby proving new conditions, both necessary and sufficient ones.

1 Introduction

The object of our studies are affine control systems of the form

$$\begin{aligned} \dot{x}(t) &= f^0(x(t)) + \sum_{i=1}^{\kappa} u_i(t) f^i(x(t)) \\ x(0) &= 0 \\ u(\cdot) &\in \mathcal{U} \subset\subset \mathbf{R}^\kappa \end{aligned} \tag{1}$$

where $x \in \mathbf{R}^n$, and $f^0, \ldots, f^\kappa$ are smooth (typically real analytic) vectorfields on $\mathbf{R}^n$, $f^0(0) = 0$, and the control u is a measurable function taking values in a compact subset $\mathcal{U}$ of $\mathbf{R}^\kappa$ containing zero in its interior. We mostly consider single-input systems, i.e. $\kappa = 1$, and then usually have $\mathcal{U}$ an symmetric interval about zero.

Such systems are a natural generalization of the linear system

$$\dot{x} = Ax + Bu \tag{2}$$

which are so often employed in engineering applications. One ultimate goal of the study of systems of form (1) is to find stabilizing feedback laws. A solution to this problem is of utmost importance for a successful implementation in automatic control into nonlinear environments. Standing in the

*This work was partially supported through NSF grants DMS-8500941 and DMS-8603156

way to such a solution are many difficult problems, which are of high mathematical interest in their own right. One of these problems is (small-time) local controllability (or reachability), which essentially asks for the set of points which may be reached from a given initial point by solutions of system (1). On a purely formal level any solution to this problem may be thought of as a nonlinear open mapping theorem (for mappings from a function space into a finite dimensional space), thus explaining the deepness and intricacy of this problem.

Many different notions of controllability have been developped, for a survey see section 2.3 of [23]. Here we only will use accessibility and small-time local controllability (STLC) (about an equilibrium point). The attainable set $\mathcal{A}_{\mathcal{U}}(t)$ at time $t > 0$ is the set of all points which can be reached in time t by solutions of (1). We say (1) is *accessible* if $\mathcal{A}_{\mathcal{U}}(t)$ has n-dimensional interior for all $t > 0$, and we say (1) is *small-time locally controllable* if $\mathcal{A}_{\mathcal{U}}(t)$ contains the equilibrium (i.e. the reference solution) $x \equiv 0$ in its interior for all $t > 0$. (Here a regrettable confusion of terminology has occurred, as intuitively one would like to call this property small-time local reachability from zero, while small-time local controllability should refer to that the equilibrium can be reached from a neighbourhood in small time. However, in our setting these two notions are symmetric under time reversal, and we shall stick to the now established definition as given above.)

We recall the *Kalman-Rank-Condition* which is both a necessary and sufficient condition for both local controllability and accessibility of the linear system (2):

Theorem 1.1 *If and only if rank*$\{B, AB, \ldots, A^{n-1}B\} = n$ *then* $0 \in int\mathcal{A}_{\mathcal{U}}(t)$ *for all* $t > 0$.

In the nonlinear case, however, a system may be accessible without being locally controllable; for example consider the system

Example 1.1

$$\begin{array}{rclrcl} \dot{x}_1(t) & = & u(t) & x(0) & = & 0 \\ \dot{x}_2(t) & = & x_1^2 & |u(\cdot)| & \leq & 1. \end{array} \tag{3}$$

The for engineering applications important and for mathematicians more interesting property is, of course, local controllability. Indeed, it is not too hard to find necessary and sufficient conditions for accessibility (at least in the analytic case) once the *matrix language* involving the matrices A and B has been replaced by the corresponding *Lie algebra language* involving the vectorfields $f^0, \ldots, f^\kappa$. (For a statement of the theorem see section 2.)

Thus while accessibility is in itself of no major importance for nonlinear systems, the easily checkable conditions for it remain important as a very first test for local controllability (since a locally controllable system clearly must be accessible.)

There are many possible approaches to finding conditions for local controllability, leading to different results and requiring different hypotheses. We here shall follow the road taken by Hermes, Sussmann and others in the early 70s. The underlying philosophy is that for (analytic) systems of form (1) the entire information about local properties of the system (e.g. local controllability) is contained in the values of the iterated Lie brackets of the vectorfields $f^0, \ldots, f^\kappa$ evaluated at the initial point $x = 0$. Moreover, these values are in principle easily computable (possibly involving some symbolic computations in the case of higher order Lie brackets). Therefore it is a very natural approach to look for conditions for small-time local controllability in terms of the elements of the Lie algebra generated by the vector fields $f^0, \ldots, f^\kappa$ evaluated at $x = 0$.

In other cases, e.g. nonsmooth vectorfields, different approaches are necessary as the Lie series techniques will often only yield little information. To only name a few examples, powerful techniques have been developed for systems posessing sufficient symmetries (e.g. Crouch and Byrnes [6]), so that actions of a symmetry group group of the system allow one to deduce controllability. As another example, results allowing nonsmooth systems (e.g. only Lipschitz) have been obtained by Frankowska [7, 8]. Alternatively, a more algebraic approach has been taken by Knobloch [16] and Wagner [29].

We already mentioned that from a purely formal point of view any result for local controllability may be viewed as a nonlinear open mapping theorem. But also local controllability is very closely related to optimal control: On one side one looks for conditions for a reference trajectory to lie in the interior, on the other side on looks for conditions for a reference trajectory to lie on the boundary of the attainable sets (or the attainable funnel in (t, x)-space), and naturally sufficient conditions for one problem readily translate into necessary conditions for the other one, and conversely.

Finally, even in the study of second order differential operators (and stochastic processes) our problem of local controllability plays an important role, as for example what we call attainable set is precisely the domain on which a strong maximum principle for hypoelliptic operators holds [28].

2 Terminology and technical preliminaries

For the most part we will consider single-input systems, i.e. $\kappa = 1$. In this case we will write u and g instead of u_1 and f^1. The solution to system (1) at time $t > 0$ corresponding to the control u is denoted by $x(t, u) = x(t, u)(0)$. In the few cases that the initial condition $x(0) = 0$ is replaced by $x(0) = x_0$ we write $x(t, u)(x_0)$ for the solution.) The attainable set at time $t > 0$ (corresponding to the control set $\mathcal{U}$) is $\mathcal{A}_{\mathcal{U}}(t) = \{x(t, u) : u(s) \in \mathcal{U} \text{ for } s \in [0, t]\}$. In the single-input case we will typically consider $\mathcal{U} = [-\varepsilon_0, \varepsilon_0]$, i.e. the control-set a symmetric interval about the origin. In this case we write $\mathcal{A}_{\varepsilon_0}(t)$ instead of $\mathcal{A}_{[-\varepsilon_0,\varepsilon_0]}(t)$.

The system (1) is accessible (from zero) if $\text{int}\mathcal{A}_{\mathcal{U}}(t) \neq \emptyset$ for all $t > 0$. The system (1) is *small-time locally controllable* (STLC) if $0 \in \text{int}\mathcal{A}_{\mathcal{U}}(t)$ for all $t > 0$, and the system (1) is *small-time locally controllable with small controls* $STLC_\varepsilon$ (about zero) if $0 \in \text{int}\mathcal{A}_{\varepsilon_0}(t)$ for all $\varepsilon_0, t > 0$. Thus, STLC_ε is stronger than STLC, which in turn implies accessibility.

The Lie bracket of two smooth vectorfields v and w is $[v, w] = (Dw)v - (Dv)w$. (Note, that in some places the negative of this expression is used.) We also use $(\text{ad}\, v, w) = [v, w]$ and $(\text{ad}^{i+1} v, w) = [v, (\text{ad}^i v, w)]$ for $i = 1, 2, \ldots$. The Lie algebra generated by the set $\mathcal{F}$ of vectorfields $\mathcal{F} = \{f^0, f^1, \ldots, f^\kappa\}$ is $L(\mathcal{F}) = L(f^1, \ldots, f^\kappa, f^0)$ (or $L(g, f)$ if $\kappa = 1$). Its homogeneous components $L^{(k,l)}(\mathcal{F})$ are spanned by all iterated brackets containing k_j factors f^j, $j = 1, \ldots, \kappa$, and l factors f_0. Here $k \in Z_0^+$ if $\kappa = 1$ and k is the multi-index $(k_1, \ldots, k_\kappa) \in (Z_0^+)^\kappa$ otherwise.

In the single-input case the following notation is also standard: $\mathcal{S}^k(g, f) = \sum_{j \leq k} \sum_{l=1}^{\infty} L^{(j,l)}(g, f)$, the subspace of $L(g,f)$ spanned by all brackets containing at most k factors of the *controlled field* g.

Finally for $p \in \mathbf{R}^n$ and $\mathcal{C}$ any set of vectorfields let $\mathcal{C}(p) = \{v(p) : v \in \mathcal{C}\}$. At several places we will have to carefully distinguish between formal Lie brackets and the vectorfields obtained by bracketing the fields f and g. We will use the convention that $X^0, X^1, \ldots, X^\kappa$ (or X, Y in the single input case) denote indeterminates, which generate the free Lie algebra $L(\mathcal{X})$ (or $L(Y,X)$) with homogeneous components $L^{(k,l)}(\mathcal{X})$ (or $L^{(k,l)}(Y,X)$) defined in the obvious way. The envelopping algebra is $A(\mathcal{X})$ with homogeneous components $A^{(k,l)}(\mathcal{X})$. Finally, we denote particular iterated Lie brackets generally by using greek superscipts, e.g. X^π, X^σ, X^{π_j}, etc. The natural *evaluation homomorphism* $\Phi = \Phi^{\mathcal{F}} : L(\mathcal{X}) \longrightarrow L(\mathcal{F})$ which is defined on the generators by $\Phi(X^j) = f^j$, $j = 0, 1, \ldots \kappa$ shall in general not be mentioned explicitly, unless necessary to avoid confusion. Rather, we will understand that if X^π denotes a particular element of $L(Y,X)$, then we use f^π for $\Phi(X^\pi)$.

An important role in the theory is played by families of dilations (on a vector space). (For an exhaustive treatment see [9].) Specifically, a one-parameter family of dilations on $\mathbf{R}^n$ (w.r.t. fixed coordinates $(x_1, \ldots, x_n)$) parametrized by $t \in \mathbf{R}^+$ is a family of maps $\Delta_t(x) = (t^{r_1}x_1, \ldots, t^{r_n}x_n)$ where $r_1, \ldots, r_n$ are nonnegative integers. A polynomial p is homogeneous of degree m, $p \in H_m$, (w.r.t. the family Δ_t of dilations) if $p(\Delta_t(x)) = t^m p(x)$. We let $H_m = 0$ if $m < 0$. For example, if $\Delta_t(x) = (tx_1, t^2x_2, t^6x_3)$, then $p(x) = x_1^8 + x_1^4x_2^2 + x_3x_2 \in H_8$. The subspaces H_m form a gradation of the algebra $\mathcal{P}$ of polynomials in $x = (x_1, \ldots, x_n)$, i.e. $H_m \cdot H_{m'} \subseteq H_{m+m'}$, $H_m \cap H_{m'} = \emptyset$ if $m \neq m'$, and $H = \bigoplus_{m=0}^\infty H_m$. (One may also consider the associated filtration $\mathcal{P} = \bigcup_{m=0}^\infty \mathcal{P}_m$ with $\mathcal{P}_m = \bigoplus_{j=0}^m H_j$).

Corresponding to this notion of homogeneous polynomials there is a notion of homogeneous vectorfields with polynomial coefficients: A vectorfield $v(x) = \sum_{j=0}^n v_j(x)\frac{\partial}{\partial x_j}$ (with polynomial coefficients $v_j(x)$ is homogeneous of degree $-k$ w.r.t. Δ_t if $vH_m \subseteq H_{m-k}$ for all $m \in Z$. We then write $v \in \underline{\mathfrak{n}}_{-k}$. Note that $v \in \underline{\mathfrak{n}}_{-k}$ if and only if $v_j \in H_{r_j-k}$ for $j = 1, \ldots, n$. For example with Δ_t as above $v(x) = \frac{\partial}{\partial x_1} + x_1\frac{\partial}{\partial x_2} + x_1^3x_2\frac{\partial}{\partial x_3} \in \underline{\mathfrak{n}}_{-1}$.

Observe that $\underline{\mathfrak{n}}_{-k}(\underline{\mathfrak{n}}_{-l}H_m) \subseteq H_{m-k-l}$ and thus $[\underline{\mathfrak{n}}_{-k}, \underline{\mathfrak{n}}_{-l}] \subseteq \underline{\mathfrak{n}}_{-k-l}$. Also $\underline{\mathfrak{n}}_{-k} \cap \underline{\mathfrak{n}}_{-l} = \emptyset$ if $k \neq l$ and $\underline{\mathfrak{n}} = \sum_{-\infty}^\infty \underline{\mathfrak{n}}_{-k}$, and thus $\{\underline{\mathfrak{n}}_{-k}\}_{k\in\mathbf{Z}}$ form a gradation of the Lie algebra $\underline{\mathfrak{n}}$ of vectorfields with polynomial coefficients. Finally, observe that $\underline{\mathfrak{n}}_{-\rho} = \{0\}$ if $\rho > \max_{j=1}^n r_j$, and thus $L(\underline{\mathfrak{n}}_{-1})$, the Lie algebra generated by $\underline{\mathfrak{n}}_{-1}$ is nilpotent.

At some times it is very convenient to also consider compositions of dilations. In particular we will use two-parameter families of dilations $\{\Delta_{\varepsilon,\delta}\}_{\varepsilon,\delta>0}$ defined by $\Delta_{\varepsilon,\delta}(x) = \Delta_\varepsilon^{(1)}(\Delta_\delta^{(2)}(x) = (\varepsilon^{r'_1}\delta^{r''_1}x_1, \ldots, \varepsilon^{r'_n}\delta^{r''_n}x_n)$ with nonnegative integers $r'_j, r''_j, j = 1, \ldots, n$. The sets $H_{m,m'}$ and $\underline{\mathfrak{n}}_{-k,-l}$ are defined in the obvious way.

Fundamental to this approach of small-time local controllability is the Lie-series formalism, in particular the Chen-Fliess-series. (Again, for a comprehensive coverage of this topic we refer the reader to [23].) Specifically, one considers the differential equation (or control-system)

$$\begin{array}{rcl} \dot{S} & = & S(X^0 + \sum_{j=1}^\kappa u_j(t)X^j) \\ S(0) & = & \mathbf{1} \end{array} \qquad (4)$$

(with $u(\cdot)$ integrable) on $A(X^1, \ldots, X^\kappa, X^0)$, and one finds the unique solution

$$\mathrm{Ser}(t,u) = \sum_I a_I(t,u) X^I \tag{5}$$

where I ranges over all multi-indices $(i_1, \ldots, i_r)$ with $r \in Z^+$ and $i_j \in \{0, \ldots, \kappa\}$, $X^I = X^{i_1} X^{i_2} \cdots X^{i_r}$ and the coeffcent $a_I(t,u)$ is the iterated integral

$$a_I(t,u) = \int_0^t u(t_r) \int_0^{t_r} u(t_{r-1}) \int_0^{t_{r-1}} \ldots \int_0^{t_1} u(t_2) u(t_1)\, dt_1 dt_2 \ldots dt_r. \tag{6}$$

The importance of this series for our purposes lies in its role as an *asymptotic series for the propagation of an analytic function ψ along trajectories of (1).* Specifically, if $\psi : \mathbf{R}^n \longrightarrow \mathbf{R}$ is analytic, then

$$\psi(x(t,u)) = ((\Phi(\mathrm{Ser}(t,u))\psi)(0) \tag{7}$$

or equivalently

$$\psi(x(t,u)) = \sum_I a_I(t,u)(f^I \psi)(0). \tag{8}$$

For small times $t \geq 0$ and analytic vectorfields f^i this series converges locally uniformly (for details see [23]).

Note that the iterated integrals $a_I(t,u)$ are homogeneous in the following sense. (Here we only consider the single-input case $\kappa = 1$, the multi-input case being similar in proper multi-index notation.) For a fixed control $u = u_{1,1} : [0,T] \longrightarrow [-\varepsilon_0, \varepsilon_0]$ and $\varepsilon, \delta > 0$ define $u_{\varepsilon,\delta} : [0, \delta T] \longrightarrow [-\varepsilon\varepsilon_0, \varepsilon\varepsilon_0]$ by $u_{\varepsilon,\delta}(\delta t) = \varepsilon u_{1,1}(t)$. Letting I_j the number of i_j in I which are equal to $j = 0, 1$ one obtains

$$a_I(\delta t, u_{\varepsilon,\delta}) = \varepsilon^{I_1} \delta^{I_0 + I_1} a_I(t, u_{1,1}). \tag{9}$$

Closely related is the following *scaling lemma* which applies to systems which are homogeneous w.r.t. a two-parameter family of dilations. Here we only give the single-input version, the extension to several controlled fields being simple, but a notational nuisance.

Lemma 2.1 *If $\Delta_{\varepsilon,\delta}$ is a two parameter family of dilations on $\mathbf{R}^n$ such that $f^0 \in \underline{\mathrm{n}}_{0,-1}$ and $f^1 = \frac{\partial}{\partial x_1} \in \underline{\mathrm{n}}_{-1,-1}$, then $x(\delta t, u_{\varepsilon,\delta}) = \Delta_{\varepsilon,\delta}(x(t, u_{1,1})$ for every admissible control $u = u_{1,1}$.*

Corollary 2.2 *Under the same hypotheses as in the lemma the attainable sets are homogeneous in the following sense $\Delta_{\varepsilon,\delta} \mathcal{A}_{\varepsilon_0}(t) = \mathcal{A}_{\varepsilon\varepsilon_0}(\delta t)$.*

Proof (of the lemma): We show that $t \longrightarrow x(\delta t, u_{\varepsilon,\delta})$ and $t \longrightarrow \Delta_{\varepsilon,\delta}(x(t, u_{1,1}))$ are both solution to the same initial value problem

$$\begin{aligned} \dot{y}(t) &= \delta f(y(t)) + \delta u_{\varepsilon,\delta}(\delta t) f^1(y(t)) \\ &= \delta f(y(t)) + \delta u_{1,1}(t) f^1(y(t)); \\ y(0) &= 0. \end{aligned} \tag{10}$$

Observe that from the hypothesis necessarily $< dx_1; f >\equiv 0$ must hold, and we may write $f(x) = \sum_{j=2}^n a_j(x) \frac{\partial}{\partial x_j}$ with polynomials $a_j(x) \in H_{r_j, s_j - 1}$, where r_j, s_j are nonegative integers with

$\Delta_{\varepsilon,\delta}(x) = (\varepsilon^{r_1}\delta^{s_1}x_1, \ldots, \varepsilon^{r_n}\delta^{s_n}x_n)$. Verify $x(\delta 0, u_{\varepsilon,\delta}) = 0 = \Delta_{\varepsilon,\delta}(x(0, u_{1,1}))$. Also, in the first component $\frac{d}{dt}x_1(\delta t, u_{\varepsilon,\delta}) = \delta u_{\varepsilon,\delta}(\delta t) = \varepsilon\delta u_{1,1}(t)$ and $\frac{d}{dt}(\Delta_{\varepsilon,\delta}(x(t, u_{1,1})))_1 = \varepsilon\delta\frac{d}{dt}x_1(t, u_{1,1}) = \varepsilon\delta u_{1,1}(t)$. For $j > 1$ we have $\frac{d}{dt}x_j(\delta t, u_{\varepsilon,\delta}) = \delta a_j(x(\delta t, u_{\varepsilon,\delta}))$ and

$$\begin{aligned}
\tfrac{d}{dt}(\Delta_{\varepsilon,\delta}(x(t, u_{1,1})))_j &= \varepsilon^{r_j}\delta^{s_j}\tfrac{d}{dt}x_j(t, u_{1,1}) \\
&= \delta\, \varepsilon^{r_j}\delta^{s_j-1}a_j(x(t, u_{1,1})) \\
&= \delta\, a_j(\Delta(x(t, u_{1,1}))).
\end{aligned} \tag{11}$$

A basic tool to prove general theorems about local controllability as well as optimality, but also to obtain information about the geometry of individual systems where the known theorems fail, are local approximation cones to the attainable set, and the associated families of admissible control variations.

Definition. *A vector $\xi \in \mathbf{R}^n$ is an m-th order tangent vector to the attainable set(s) at zero, if there is a one-parameter family of points $q(s) \in \mathcal{A}_{\mathcal{U}}(s)$, $s \geq 0$ such that $q(s) = \xi s^m + o(s^m)$. Here $o(s)$ is such that $lim_{s\to 0}\frac{O(s^m)}{s^m} = 0$.*

The set of all m-th order tangent vectors is denoted $\mathbf{K}^m$ and $\mathbf{K} = \bigcup_{m=1}^{\infty}\mathbf{K}^m$. Sometimes it is more convenient to work with the cones generated by these tangent vectors, and we use $\overline{\mathbf{K}^m} = \bigcup_{\lambda>0}\lambda\mathbf{K}^m$ and $\overline{\mathbf{K}} = \bigcup_{\lambda>0}\lambda\mathbf{K}$.

As a consequence of $f^0(0) = 0$ the sets $\overline{\mathbf{K}^m}$, $m > 0$ form an increasing sequence of convex cones. We may remind the reader that in general the attainable set may well have inward corners, and even inward cusps, and and that because of this the local approximating cones as employed say in the various versions of the Maximal Principle or when considering controllability about a nonstationary reference trajectory have to be defined in a much more restrictive way than in above definition. Specifically one finds in this special case:

Lemma 2.3 *Let $\mathbf{K}^m$ be defined as above.*
(i) If $\lambda \in [0, 1]$, then $\lambda^m\mathbf{K}^m \subseteq \mathbf{K}^m$.
(ii) If $m \leq l$, then $\mathbf{K}^m \subseteq \mathbf{K}^l$.
(iii) If $v^{(1)}, v^{(2)} \in \mathbf{K}^m$ and $\lambda \in [0, 1]$, then $\lambda^m v^{(1)} + (1-\lambda)^m v^{(2)} \in \mathbf{K}^m$.

Thus the sets $\mathbf{K}^m$ form an increasing sequence of truncated cones which are almost convex.

Proof: (i) Let $\lambda \in [0, 1]$ and $v \in \mathbf{K}^m$. Fix a one-parameter family $\{u_s\}_{s\geq 0}$ of control variations $u_s : [0, s] \longrightarrow \mathcal{U}$ such that $x(s, u_s) = vs^m + o(s^m)$, $s \geq 0$. Define the controls $\tilde{u}_s : [0, s] \longrightarrow \mathcal{U}$ by

$$\tilde{u}_s(t) = \begin{cases} 0 & \text{if} \quad 0 \leq t < (1-\lambda)s \\ u_s(t - (1-\lambda)s) & \text{if} \quad (1-\lambda)s \leq t \leq s, \end{cases} \tag{12}$$

and clearly $x(s, \tilde{u}_s) = x(\lambda s, u_{\lambda s}) = (\lambda^m v)s^m + \lambda^m o(s^m)$.

(ii) Let λ, v and $\{u_s\}_{s\geq 0}$ be as in (i). Also let $l > m$. We only have to consider $0 \leq s \leq 1$ and thus may define:

$$\bar{u}_s(t) = \begin{cases} 0 & \text{if} \quad 0 \leq t < s - s^{l/m} \\ u_{s^{m/l}}(t - s^{l/m}) & \text{if} \quad s - s^{l/m} \leq t \leq s. \end{cases} \tag{13}$$

and clearly $x(s, \bar{u}_s) = x(s^{m/l}, u_{s^{m/l}}) = v(s^{m/l})^l + o((s^{m/l})^l) = vs^m + o(s^m)$.

(iii) Let $\{u_s^{(1)}\}_{s\geq 0}$ and $\{u_s^{(2)}\}_{s\geq 0}$ be families of control variations generating the m-th order tangent vectors $v^{(1)}$ and $v^{(2)}$, respectively. For $\lambda \in [0,1]$ define the family of control variations $\{u_s^\lambda\}_{s\geq 0}$ by

$$u_s^\lambda(t) = \begin{cases} u_{\lambda s}^{(1)}(t) & \text{if} \quad o \leq t < \lambda s \\ u_{(1-\lambda)s}^{(2)}(t - \lambda s) & \text{if} \quad \lambda s \leq t \leq s. \end{cases} \tag{14}$$

The conclusion $x(s, u_s^\lambda) = (\lambda^m v^{(1)} + (1-\lambda)^m v^{(2)})s^m + os^m)$ is a direct consequence of Gronwall's lemma. (See also the appendix.)

The main value of the cones $\overline{\mathbf{K}^m}$ is that they are approximating cones in the following sense:

Theorem 2.4 *If $\overline{\mathbf{K}'}$ is a closed convex cone (with vertex $0 \in \mathbf{R}^n$) such that $\overline{\mathbf{K}'} \setminus \{0\} \subseteq \mathrm{int}\overline{\mathbf{K}^m}$ for some $m < \infty$ then there are constants $C, T > 0$ such that $\overline{\mathbf{K}'} \cap B(0, Ct^m) \subseteq \mathcal{A}_\mathcal{U}(t)$ for all $0 \leq t \leq T$.*

Corollary 2.5 *If $\overline{\mathbf{K}^m} = \mathbf{R}^n$ then there are constants $C, T > 0$ such that $B(0, Ct^m) \subseteq \mathcal{A}_\mathcal{U}(t)$ for all $0 \leq t \leq T$.*

The theorem can be seen as the statement that the minimum-time map is locally Hölder continuous with exponent $1/m$ in the direction of $\overline{\mathbf{K}'}$.

The proof of this theorem is constructive and can be used to synthesize some stabilizing feedback laws (for the time-reversed system). An outline of the proof is given in the appendix. A similar theorem for a more general setting with an indirect proof is given in Frankowska [7, 8].

To explicitly generate and compute tangent vectors to the attainable sets it is sufficient and usual to consider one-parameter families of control variations $u_s : [0, s] \longrightarrow \mathcal{U}$ (of the zero reference control) and the terminal points $q(s) = x(s, u_s)$ of the associated trajectories.

Note that we do not require any continuity of the selection $s \longrightarrow u_s$ whatsoever in order to obtain above convexity results. This makes the cones we consider here very different from those usually employed in the theory of optimality, e.g. in the many versions of the Maximal Principle.

3 Survey of conditions for STLC

Before we come to conditions for small-time local controllability we need the following criterion for accessibility:

Theorem 3.1 (Hermann-Nagano) *If $\dim L(f^0, \ldots, f^\kappa)(0) = n$ then $\mathrm{int}\mathcal{A}_\mathcal{U}(t) \neq \emptyset$ for all $t > 0$. Moreover if the vectorfields $f^i, i = 0, 1, \ldots, \kappa$ are analytic then $\dim L(f^0, \ldots, f^\kappa)(0) = n$ is also necessary for $\mathrm{int}\mathcal{A}_\mathcal{U}(t) \neq \emptyset$ for all $t > 0$.*

Blanket Hypothesis: Unless otherwise stated we shall from now on always assume that $\dim L(f^0, \ldots, f^\kappa)(0) = n$ since otherwise (in the case of analytic vectorfields) we we may restrict our attention to the smaller integral manifold of $L(f^0, \ldots, f^\kappa)$ through zero.

The first sufficient condition for STLC is the *Linear Test*:

Theorem 3.2 *If* $span\{(ad^{\nu} f^0, f^j)(0) : \nu = 0, 1, \ldots;\ j = 1, \ldots, \kappa\} = \mathbf{R}^n$ *then the system (1) is STLC.*

The close relation of this theorem to the Kalman Rank condition (Theorem 1.1) is immediate, and indeed, this theorem may be thought of as the statement: If the linearization of the system (1) is STLC then so is the system (1) itself. (The converse is, of course, not true.)

The most simple system for which the linear test fails to hold is (3) which clearly is not controllable. Here $g = \frac{\partial}{\partial x_1}$, $f = x_1^2 \frac{\partial}{\partial x_2}$, $[g, f] = 2x_1 \frac{\partial}{\partial x_2}$, $[g, [g, f]] = 2\frac{\partial}{\partial x_2}$, all other brackets (modulo anticommutativity and the Jacobi identity) are identically zero. The only brackets which do not vanish at $x = 0$ are $g(0) = \frac{\partial}{\partial x_1}$ and $[g, [g, f]](0) = 2\frac{\partial}{\partial x_2}$; the second bracket being nonzero at $x = 0$ essentially determines the non-controllability of the system. This example motivates the following necessary condition for STLC which in optimal control theory is known as the Legendre-Clebsch condition.

Theorem 3.3 (Hermes [10], Sussmann [24]) *If the system (1) is STLC then* $[g, [g, f]](0) \in \mathcal{S}^1(g, f)(0)$.

Example 3.1

$$\begin{array}{rclcrcl} \dot{x}_1 & = & u & \qquad & x(0) & = & 0 \\ \dot{x}_2 & = & x_1^m & \qquad & |u(\cdot)| & \leq & 1 \end{array} \tag{15}$$

For this system one easily calculates that $g(0) = \frac{\partial}{\partial x_1}$, and $(ad^m g, f)(0) = m!\frac{\partial}{\partial x_2}$, are the only brackets not vanishing at zero. Also it is not too difficult to see that the system (15) is STLC if and only if m is an odd integer. This example motivates the *Hermes Condition*:

Theorem 3.4 (Hermes [10], Sussmann [24])
If $\mathcal{S}^{2k}(g, f)(0) \subseteq \mathcal{S}^{2k-1}(g, f)(0)$ *for all* $k \in \mathbf{Z}+$ *then system (1) is STLC.*

Note that this condition does not require brackets containing an even number of factors of the controlled field g to vanish at the origin, but rather only requires them to be (at zero) expressible as linear combinations of brackets containing fewer factors g.

The basic idea why brackets containing an odd number of factors of the controlled field g are never obstructions to controllability is as follows: Those iterated integrals $a_I(t, u)$ in the Chen-Fliess-series with I_1 odd are homogeneous of an odd degree in the control u, and thus $a_I(t, -u) = -a_I(t, u)$. Consequently, if all brackets containing an even number of factors of the controlled field g vanish at the origin, and the family of control variations $\tilde{u}_s$ generates the tangent vector $f^{\pi}(0)$, then the family $-\tilde{u}_s$ generates the tangent vector $-f^{\pi}(0)$. Here $u \longrightarrow -u$ is an input symmetry which leads to a symmetry of the attainable set if all brackets with an even number of factors g vanish at zero.

If not all these brackets vanish at zero, but they are linear combinations of brackets with fewer factors g, then one considers families of control variations essentially parametrized by their amplitude, e.g. $u_{\varepsilon} = \varepsilon u_1$, $\varepsilon > 0$. Now one uses that the iterated integrals $a_I(t, u)$ are homogeneous of order I_1 in the control-amplitude, i.e. $a_I(t, u_{\varepsilon}) = \varepsilon^{I_1} a_I(t, u_1)$, and clearly as ε goes to zero those coefficients $a_I(t,u)$ with I_1 small will dominate those with I_1 large. The hard part in the proof is to show that several such neutralizations can be done simultaneously.

The necessary condition corresponding to Theorem 3.4 is:

Theorem 3.5 (Stefani [21]) *If (1) is STLC then $(ad^{2m}g, f)(0) \in \mathcal{S}^{2m-1}(0)$ for all $m \in \mathbf{Z}^+$.*

The proof of this theorem uses that there is precisely one term corresponding to the brackets $(\mathrm{ad}^{2m}Y, X)$ in the Chen-Fliess series, this iterated integral being always nonnegative, and moreover it can be shown that for sufficiently small times $t > 0$ it always dominates the remaining part of the series corresponding to terms with at least $2m$ factors Y. The proof of this last assertion is based on a repeated use of the Hölder inequality to estimate the size of all $a_I(t,u)$ with $I_1 \geq 2m$ and a rough estimate of how many such coefficients for each fixed length $|I|$ have to be considered.

Geometrically, if $(\mathrm{ad}^{2m}g, f)(0) \notin \mathcal{S}^{2m-1}(g,f)(0)$, then in first approximation the attainable set lies (for small times) on one side of the hyperplane through zero perpendicular to $(\mathrm{ad}^{2m}g, f)(0)$. In other words every tangent vector in $\mathbf{K}$ has a nonnegative projection on $(\mathrm{ad}^{2m}g, f)(0)$.

However not all brackets containing an even number of factors Y need to be neutralized to give STLC as was first shown in the following example:

Example 3.2 (Stefani [19])

$$\begin{array}{rclrcl} \dot{x}_1 & = & u & x(0) & = & 0 \\ \dot{x}_2 & = & x_1 & |u(\cdot)| & \leq & 1 \\ \dot{x}_3 & = & x_1^3 x_2 & & & \end{array} \tag{16}$$

One easily computes that the only brackets (modulo the Jacobi identity) which do not vanish at zero are: $g(0) = \frac{\partial}{\partial x_1}$, $[g,f](0) = \frac{\partial}{\partial x_2}$, and $[[g,f],(\mathrm{ad}^g, f)(0) = 6\frac{\partial}{\partial x_3},$. The last bracket contains 4 factors of the controlled field, and thus is a possible obstruction to STLC. However, an explicit construction of controls shows that $0 \in \mathrm{int}\mathcal{A}_1(t)$ for all $t > 0$.

A close examination of this and similar systems revealed the following reasoning why these are controllable: If $\tilde{u} : [0,T] \longrightarrow [-1,1]$ is a control such that $x_1(T,\tilde{u}) = x_2(T,\tilde{u}) = 0$, and $x_3(T,\tilde{u}) \neq 0$, then for the time-reversed control $\tilde{u}^{-1}$ (which is defined by $\tilde{u}^{-1}(t) = \tilde{u}(T-t)$) one easily verifies that $x_1(T,\tilde{u}^{-1}) = x_2(T,\tilde{u}^{-1}) = 0$, but also $x_3(T,\tilde{u})x_3(T,\tilde{u}^{-1}) < 0$. The change of sign in the last component occurs *because* the bracket $X^\pi = [[Y,X],(\mathrm{ad}^3Y,X)]$ of type $(4,2)$ contains an even total number of factors X and Y. This corresponds to an even number of integrations in the corresponding iterated integral, which leads to a change of sign under the *symmetry operation* time reversal. ($x_i(T,\tilde{u}) = 0$, $i = 1,2$ is necessary in order for the boundary terms to vanish, when manipulating the integral.)

The time reversal as illustrated above is one main ingredient to Sussmann's general theorem (s.b.), another one is illustrated in the following example.

Example 3.3 (Sussmann [24])

$$\begin{array}{rclrcl} \dot{x}_1 & = & u & x(0) & = & 0 \\ \dot{x}_2 & = & x_1 & |u(\cdot)| & \leq & 1 \\ \dot{x}_3 & = & x_2^2 + x_1^3 & & & \end{array} \tag{17}$$

In (17) the important brackets are $(\mathrm{ad}^2[g,f],f)(0) = 2\frac{\partial}{\partial x_3}$ and $(\mathrm{ad}^3 g,f)(0) = 6\frac{\partial}{\partial x_3}$, which are of types $(2,3)$ and $(3,1)$ respectively, so that the Hermes condition (Theorem 3.4) fails However, one can show that this system is STLC. The idea is as follows: Even though the possible obstruction to STLC $(\mathrm{ad}^2[Y,X],X)$ contains fewer factors Y than the bracket$(\mathrm{ad}^3 Y,X)$, it is of greater total length, and thus the associated integrals in the Chen-Fliess series are homogeneous of order 5 in t as opposed to order 4 for the second bracket.

Thus instead of considering families of control variations parametrized by the amplitude, here one considers controls parametrized by the length of time that they are different from the zero reference control. Specifically, for any fixed control $u_{1,1} : [0,T] \longrightarrow \mathcal{U}$ consider the one-parameter family of controls $u_{1,\delta} : [0,\delta T] \longrightarrow \mathcal{U}$ defined by $u_{1,\delta}(\delta t) = u_{1,1}(t)$, and one finds $a_I(\delta T, u_{1,\delta}) = \delta^{|I|} a_I(t, u_{1,1})$. Intuitively, for $\delta \longrightarrow 0$ the integrals corresponding to brackets of shorter length dominate those corresponding to brackets of greater total length.

Combining these families of control variations with those introduced earlier (those parametrized by the amplitude), and also the input symmetries $u \longrightarrow -u$ and $u \longrightarrow u^{-1}$ one arrives after meticulous work at the following special version of Sussmann's general theorem on sufficient conditions for STLC:

Theorem 3.6 (Sussmann [26]) *If there is a weight $\Theta \in [0,1]$ such that for all brackets X^π of type (k,l) with k even and l odd there are brackets X^{π_j} of type (k_j,l_j) with $k_j + \Theta l_j < k + \Theta l$ such that $f^\pi(0)$ is a linear combination of $f^{\pi_j}(0)$ then the system (1) is STLC.*

The general theorem of Sussmann applies to the multi-input case (allowing different weights for the various controlled fields), case, and also allows the brackets of type (even,odd) in above statement to be replaced by more general "fixed elements of a group of input symmetries". Since presently we do not know of any other practical input symmetries, we here shall only use above version and refer the interested reader to the original paper for the statement of the general version of the theorem.

Pictorially, the statement of Theorem 3.6 is as follows: Graph the homogeneous subspaces $L^{(k,l)}(Y,X)$ as the (k,l)-th entries of an infinite matrix. The only "bad" brackets (i.e.possible obstructions to STLC) are of type (even,odd). But the possible obstructions in $(2k, 2l+1)$ may be neutralized by all brackets in types which lie above the line through $(2k, 2l+1)$ with slope Θ, see figure 3.

Note, that several bad brackets in a given system may be neutralized simultaneously in above manner, but the weight Θ has to be the same in all these neutralizations.

In terms of families of control variations the weight Θ determines how the two different families of control variations from above are to be combined. In particular, starting with the two-parameter family of control variations $\{u_{\varepsilon,\delta}\}_{\varepsilon,\delta>0}\}$ defined by $u_{\varepsilon,\delta}(\delta t) = \varepsilon u_{1,1}(t)$ one sets $\varepsilon = s^{1-\Theta}$ and $\delta = s^\Theta$ to obtain a one-parameter family of control variations $\{u_{s^{1-\Theta},s^\Theta}\}_{s\geq 0}$, which finally leads to $a_I(s^\Theta t, u_{s^{1-\Theta},s^\Theta}) = s^{I_1 + \Theta I_0} a_I(t, u_{1,1})$.

Above theorem has been slightly extended in such a way that brackets of the form $(\mathrm{ad}^\nu X^0, X^\pi)$ of type $(k+\nu, l)$ may be treated in certain cases like brackets of type (k,l), i.e. *outside* factors of the uncontrolled (or drift-) vector field f^0 do not count:

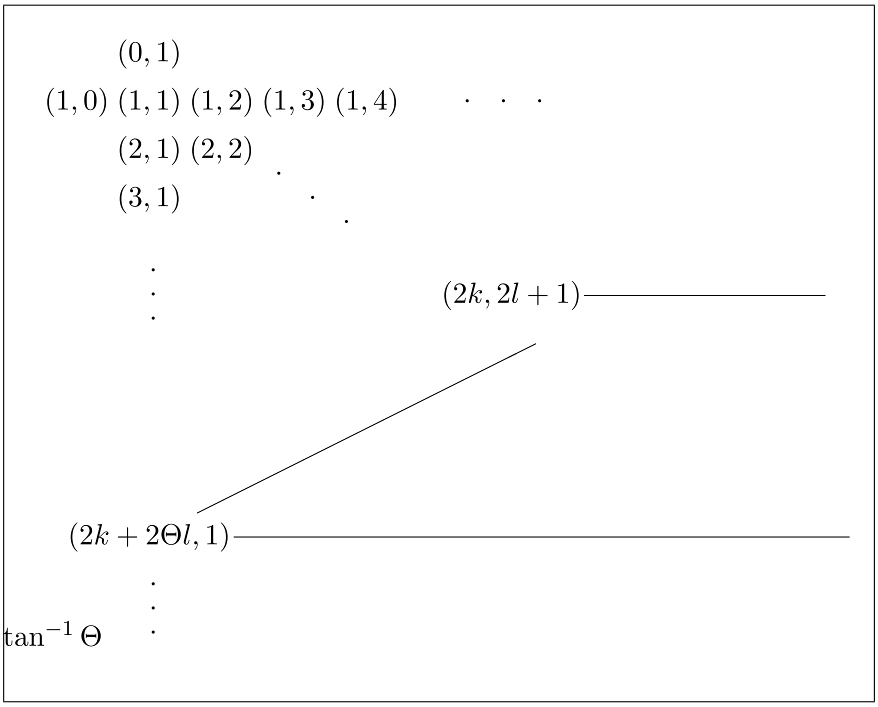

Figure 1: Neutralizing bad brackets

Theorem 3.7 (Bianchini-Stefani [1], Hermes-Kawski [11]) *If there is a weight $\Theta \in [0,1]$ such that for all brackets X^π of type (k,l) with k even and l odd there are brackets X^{π_j} of type (k_j,l_j) with $k_j+\Theta l_j < k+\Theta l$ and $\nu_j \geq 0$ such that $f^\pi(0)$ is a linear combination of $(ad^{\nu_j} f^0, f^{\pi_j})(0)$ then the system (1) is STLC.*

For brackets of type $(2,3)$ the complementary necessary condition has been proven (s.b.). However, in general brackets of type (even,odd) need not be *bad* brackets as will be shown in the next section.

Theorem 3.8 (Kawski [13, 14]) *If (1) is STLC$_\varepsilon$ and $X^\pi \in L^{(2,3)}(Y,X)$ then $f^\pi(0) \in (\mathcal{S}^1(Y,X)(0) + \{(ad^\nu f, (ad^3 g, f))(0) \,:\, \nu \geq 1\,\})$.*

Note, that the two last theorems specify a certain set of brackets of type $(3,\cdot)$ which can neutralize brackets of type $(2,3)$, while the other brackets of the same typee $(3,\cdot)$ cannot neutralize brackets of type $(2,3)$. This necessary distinction of brackets of the same causes the proof to become even more technical, and also makes one expect much increased difficulties when trying to prove similar necessasry conditions for STLC pertaining to brackets of type (even,odd) of higher order (greater length).

Comparing theorems 3.5 and 3.6 on one side and 3.4 and 3.8 on the other side, one observes a large gap betwen necessary and sufficient conditions for STLC (even for the single-input system). In the following sections we shall venture into this gap from various sides, and explore the high order phenomena which make it so difficult to find simultaneously sharp necessary and sufficient conditions for STLC.

At this point we may remark that some of the conditions for STLC about an equilibrium point given above carry over to small-time local controllability about a (not necessarily constant) reference trajectory. One of the main differences is that in that case even brackets of type (even,even) may be the first obstructions to small-time local controllability In the case of $f(0) = 0$, if a bracket f^π is neutralized by the brackets f^{π_j}, $j = 1, 2, \ldots$, then $(\text{ad}^\nu f, f^\pi)$ is neutralized by $(\text{ad}^\nu f, f^{\pi_j})$, $j = 1, 2, \ldots$. This is not necessarily true in the case of a nonconstant reference trajectory $\tilde{x}$, and in Bianchini-Stefani [2] an example of a system is given in which all brackets of type (even,odd) are neutralized as in theorem 3.6, but the system is nonetheless not *locally controllable about a reference trajectory* due to the bracket $[X, [Y, [Y, X]]]$ of type (2,2) which is not neutralized.

4 Good brackets of type (even,odd)

In this section we show that there are more brackets of type (even,odd) than those of form $(\text{ad}^\nu X^0, X^\pi)$ (with $\nu \geq 1$) that need not be neutralized in order to give small-time local controllability.

We first consider the following system on $\mathbf{R}^4$:

Example 4.1.

$$\begin{array}{rclcrcl} \dot{x}_1 & = & u & & x(0) & = & 0 \\ \dot{x}_2 & = & x_1 & & |u(\cdot)| & \leq & 1 \\ \dot{x}_3 & = & \frac{1}{6}x_1^3 & & & & \\ \dot{x}_4 & = & x_2 x_3 & & & & \end{array} \qquad (18)$$

One easily computes the at zero nonvanishing brackets: $g(0) = \frac{\partial}{\partial x_1}$, $[g, f](0) = \frac{\partial}{\partial x_2}$, $(\text{ad}^3 g, f)(0) = \frac{\partial}{\partial x_3}$, and $f^\pi(0) = [(\text{ad}^3 g, f), (\text{ad}^2 f, g)](0) = \frac{\partial}{\partial x_4}$. Note that the bracket X^π is of type $(4, 3)$, not of the form $(\text{ad}^\nu X, X^\sigma)$ for any $\nu \geq 1$, and $f^\pi(0)$ is not a linear combination of any other brackets at zero, and thus X^π has to be considered a potential obstruction to STLC by the theorems in the previous section. However, the system (18) is STLC. To show this, observe that the system (18) is homogeneous w.r.t. the two-parameter family of dilations $\Delta_{\varepsilon,\delta}(x) = (\varepsilon\delta x_1, \varepsilon\delta^2 x_2, \varepsilon^3\delta^4 x_3, \varepsilon^4\delta^7 x_4)$ in the sense that w.r.t. this dilation $f \in \underline{\text{n}}_{0,-1}$ and $g \in \underline{\text{n}}_{-1,-1}$. Therefore $\mathcal{A}_\varepsilon(\delta t) = \Delta_{\varepsilon,\delta}\mathcal{A}_1(t)$ for all $\varepsilon, \delta \in (0, 1]$.

By the Hermes Condition (Theorem 3.4) the (x_1, x_2, x_3)-subsystem is STLC, and one easily generates suitable positive multiples of $\pm\frac{\partial}{\partial x_j}$, $j = 1, 2, 3$, as $1^{st}, 2^{nd}$ and 4^{th} order tangent vectors, respectively. To generate a positive multiple of $+\frac{\partial}{\partial x_4}$ as a 7^{th} order tangent vector consider the family of control variations $u_\delta : [0, 4\delta] \longrightarrow [-1, 1]$, $\delta \in (0, 1]$ defined by $u_1(t) = 1$ if $t \in [0, 1] \cup [3, 4]$ and $u_1(t) = -1$ if $t \in (1, 3]$, and $u_\delta(\delta t) = u_1(t)$. One easily verifies that $x(4\delta, u_\delta) = (0, 0, 0, \delta^7 y_4)$ for some suitable constant $y_4 > 0$.

To generate $-\frac{\partial}{\partial x_4}$ as a 7^{th} order tangent vector consider the following control variation $u^{\xi,\eta}$:

$[0, 2T + \eta] \longrightarrow [-1, 1]$ (with ξ, η as chosen below) defined by (here $T_\xi = 4 + 2\xi$)

$$u^{\xi,\eta}(t) = \begin{cases} 1 & \text{if } t \in [0,1) \cup [2+\xi, 3+2\xi) \\ -1 & \text{if } t \in [1, 2+\xi) \cup [3+2\xi, T_\xi) \\ 0 & \text{if } t \in [T_\xi, t_\xi + \eta] \\ -u^{\xi,\eta}(t - T_\xi - \eta) & \text{if } t \in [T_\xi + \eta, 2T_\xi + \eta]. \end{cases} \tag{19}$$

Then some elementary calculations give that for all $s \in [0, \eta]$:

$$\begin{array}{rcccl} x_1(T_\xi, u^{\xi,\eta}) & = & x_1(T_\xi + s, u^{\xi,\eta}) & = & 0, \\ x_2(T_\xi, u^{\xi,\eta}) & = & x_2(T_\xi + s, u^{\xi,\eta}) & = & 2 - \xi^2, \text{ and} \\ x_3(T_\xi, u^{\xi,\eta}) & = & x_3(T_\xi + s, u^{\xi,\eta}) & = & 2 - \xi^4 \end{array} \tag{20}$$

Also $x_j(2T_\xi + \eta, u^{\xi,\eta}) = 0$ for $j = 1, 2, 3$ and all choices of $\xi, \eta \geq 0$. We let $C(\xi) = x_4(T_\xi, u^{\xi,\eta}) = \int_0^{T_\xi} x_2(s, u^{\xi,\eta}) x_3(s, u^{\xi,\eta})\, ds$. One easily verifies that $C(0) > 0$ and $x_4(2T_\xi + \eta) = 2C(\xi) + \eta(2-\xi^2)(2-\xi^4)$. Choose $\sqrt[4]{2} < \xi < \sqrt[2]{2}$ and $\eta > 2C(\xi)(2-\xi^2)^{-1}(2-\xi^4)^{-1}$ and one obtains $x_4(2T_\xi + \eta, u^{\xi,\eta}) < 0$. Now use the homogenuity of the system to obtain a suitable positive multiple of $-\frac{\partial}{\partial x_4}$ as a 7^{th} order tangent vector via the usual scaling of the control, i.e. $u_\delta^{\xi,\eta}(\delta t) = u^{\xi,\eta}(t)$ for $\delta \in (0, 1]$.

In our construction we made use of that $f_j(E_1) = 0$ for $j = 1, 2, 3$ where E_1 is the hyperplane $\{ x : x_1 = 0 \}$, and that on the intersection of E_1 with any neighbourhood of $x = 0$ $f_4(x) = x_2 x_3$ takes both negative and positive values.

Note that Example 4.1 generalizes in an obvious way, namely by replacing $\dot{x}_3 = \frac{1}{6} x_1^3$ by $\dot{x}_3 = \frac{1}{(2m-1)!} x_1^{2m-1}$ for any integer $m \geq 2$. The computations remain essentially unchanged, and if one chooses $\sqrt[2m]{2} < \xi < \sqrt[2]{2}$ then one obtains $-f^{\pi_m} = -[(\mathrm{ad}^{2m-1} g, f), (\mathrm{ad}^2 f, g)] = \frac{\partial}{\partial x_4}$ as a $(2m+3)^{th}$ order tangent vector. Summarizing, we have:

Proposition 4.1 *For every integer $m \geq 2$ the systems*

$$\begin{array}{rclcrcl} \dot{x}_1 & = & u & & x(0) & = & 0 \\ \dot{x}_2 & = & x_1 & & |u(\cdot)| & \leq & 1 \\ \dot{x}_3 & = & x_1^{2m-1} & & & & \\ \dot{x}_4 & = & x_2 x_3 & & & & \end{array} \tag{21}$$

are STLC and thus the associated brackets $[(\mathrm{ad}^{2m-1} Y, X), (\mathrm{ad}^2 X, Y)]$ of type $(2m, 3)$ need not be obstructions that have to be neutralized in order to give STLC.

To obtain new general sufficient or necessary conditions for STLC one has to consider all brackets in one homogeneous component at one time. We illustrate this by considering $L^{(4,3)}(Y, X)$ which is free of rank five. We consider the following system on $\mathbf{R}^{12}$:

$$\begin{cases} \dot{x}_1 = u & \dot{x}_7 = \frac{1}{6} x_1^3 x_2 & x(0) = 0 \\ \dot{x}_2 = x_1 & \dot{y}_1 = x_6 & |u(\cdot)| \leq 1 \\ \dot{x}_3 = \frac{1}{2} x_1^2 & \dot{y}_2 = \frac{1}{2} x_3^2 & \\ \dot{x}_4 = \frac{1}{6} x_1^3 & \dot{y}_3 = \frac{1}{4} x_1^2 x_2^2 & \\ \dot{x}_5 = \frac{1}{24} x_1^4 & \dot{y}_4 = x_7 & \\ \dot{x}_6 = x_5 & \dot{y}_5 = x_2 x_4 & \end{cases} \tag{22}$$

This system is clearly not STLC (e.g. because of $x_3(t,u) \geq 0$ for all $t \geq 0$, u admissible). In this case we are interested in the projection of the attainable sets on the y-subspace. We compute the values of the following brackets at zero:

$$\begin{array}{rcl} f^{\pi_1}(0) & = & (\mathrm{ad}^2 f, (\mathrm{ad}^4 g, f))(0) = \frac{\partial}{\partial y_1} \\ f^{\pi_2}(0) & = & (\mathrm{ad}^2 (\mathrm{ad}^2 g, f), f)(0) = \frac{\partial}{\partial y_2} \\ f^{\pi_3}(0) & = & (\mathrm{ad}^2 [g,f], (\mathrm{ad}^2 g, f))(0) = \frac{\partial}{\partial y_3} \\ f^{\pi_4}(0) & = & [f, [[f,g], (\mathrm{ad}^3 g, f)]](0) = \frac{\partial}{\partial y_4} \\ f^{\pi_5}(0) & = & [(\mathrm{ad}^2 f, g), (\mathrm{ad}^3 g, f)](0) = \frac{\partial}{\partial y_5} \end{array} \tag{23}$$

Note, that all these brackets are of type $(4,3)$, and they are linearly independent vectors at $x = 0$. Also, $L^{(4,3)}(Y,X)$ is free of rank five, and therefore $X^{\pi_1}, \ldots, X^{\pi_5}$ form a basis of this homogeneous subspace of the free Lie algebra. Observe, that there are at least three independent supporting hyperplanes of the attainable set in the y-subspace. Also, one may show that the $(x_1, x_2, x_4, x_7, y_5, y_6)$-subsystem is STLC, and therefore three is the maximal number of independent supporting hyperplanes of the attainable set in the y-subspace.

Thus, there are three *bad* and two *good* brackets of type $(4,3)$. Above basis for $L^{(4,3)}(Y,X)$ is *nice* because it splits into bases $\{X^{\pi_1}, X^{\pi_2}, X^{\pi_3}\}$ for the bad and $\{X^{\pi_4}, X^{\pi_5}\}$ for the good subspace. For the higher order homogeneous subspaces of the free Lie algebra on two generators it is not known how to find similar nice basis. (Already in the next interesting homogeneous subspace $L^{(4,5)}(Y,X)$, which is free of rank 14, the Philipp-Hall- and the Chen-Fox-Lyndon bases no longer split in such a way.)

5 Neutralizing and balancing bad brackets

In this section we give a new necessary condition for neutralizing bad brackets and also consider the case when a bad bracket is not neutralized by lower order brackets, but a pair of bad brackets is balancing each other to possibly give STLC. We shall give examples for various cases, in particular pairs of brackets of different type and pairs of brackets of the same type.

Also, the controllability of the system may depend on the relative lengths of the vectors obtained by evaluating the brackets at zero. In this case one may consider a one-parameter family of such control systems, and in particular look at the critical value when the system changes from STLC to not STLC. In the three-dimensional example we give here, we will find nontrivial periodic extremals (solutions to the Hamilton-Jacobi-Belmann equations) which foliate some portion of the boundary of the attainable set.

First consider the following two control systems on $\mathbf{R}^4$:

Example 5.1

$$\begin{array}{rclcrcl} \dot{x}_1 & = & u & \quad & x(0) & = & 0 \\ \dot{x}_2 & = & x_1 & & |u(\cdot)| & \leq & \varepsilon_0 \\ \dot{x}_3 & = & x_2 & & \lambda & > & 0 \\ \dot{x}_4 & = & x_2^2 - \lambda x_1^4 & & & & \end{array} \tag{24}$$

Example 5.2

$$\begin{array}{rcl} \dot{x}_1 & = & u \\ \dot{x}_2 & = & x_1 \\ \dot{x}_3 & = & x_2 \\ \dot{x}_4 & = & x_3^2 - \lambda x_1^4 \end{array} \qquad \begin{array}{rcl} x(0) & = & 0 \\ |u(\cdot)| & \leq & \varepsilon_0 \\ \lambda & > & 0 \end{array} \tag{25}$$

Certain aspects of the first system have been discussed in [20]. (Note, that in the first system the x_3-component is insignificant for questions about controllability and has only been introduced to facilitate the comparison of these two systems.) The brackets which do not vanish at zero are $g(0) = \frac{\partial}{\partial x_1}$, $[g, f](0) = \frac{\partial}{\partial x_2}$, $(\text{ad}^2 f, g)(0) = \frac{\partial}{\partial x_3}$, and $(\text{ad}^4 g, f)(0) = -24\lambda\frac{\partial}{\partial x_4}$ in both systems; and $(\text{ad}^2[g, f], f)(0) = 2\frac{\partial}{\partial x_4}$ in the first system, and $(\text{ad}^2(\text{ad}^2 f, g), f)(0) = 2\frac{\partial}{\partial x_4}$ in the second system.

From the linear theory it is clear that suitable positive multiples of $\pm\frac{\partial}{\partial x_j}$, $j = 1, 2, 3$ are j-th order tangent vectors to the attainable set. We now show explicitly when and how one can obtain (multiples of) $\pm\frac{\partial}{\partial x_4}$ as tangent vectors. It is more convenient to do these calculations for the system

$$\begin{array}{rcl} \dot{y}_1 & = & u \\ \dot{y}_2 & = & y_1 \\ \dot{y}_3 & = & y_2 \\ \dot{y}_4 & = & y_2^2 \\ \dot{y}_5 & = & y_3^2 \\ \dot{y}_6 & = & y_1^4 \end{array} \qquad \begin{array}{rcl} y(0) & = & 0 \\ |u(\cdot)| & \leq & \varepsilon_0 \end{array} \tag{26}$$

which has the advantage of being homogeneous w.r.t. the two-parameter family of dilations $\Delta_{\varepsilon,\delta}(y) = (\varepsilon\delta y_1, \varepsilon\delta^2 y_2, \varepsilon\delta^3 y_3, \varepsilon^2\delta^5 y_4, \varepsilon^2\delta^7 y_5, \varepsilon^4\delta^5 y_6)$ in the sense that w.r.t. this dilation $f \in \underline{\mathrm{n}}_{0,-1}$ and $g \in \underline{\mathrm{n}}_{-1,-1}$, and therefore the attainable set is homogeneous, also.

First choose a control $u = u_{1,1} : [0, T] \longrightarrow [-1, 1]$ for some $T > 0$ such that $u \not\equiv 0$ and $y_j(t, u_{1,1}) = 0$ for $j = 1, 2, 3$. For example the control $u_{1,1}$ defined on $[0, 4 + 2\sqrt{2}]$ by $u_{1,1}(t) = 1$ if $t \in [0, 1] \cup [2 + \sqrt{2}, 3 + 2\sqrt{2}]$ and $u_{1,1}(t) = -1$ else satisfies these requirements.

Let $C_j = y_j(T, u_{1,1})$ for $j = 4, 5, 6$. Then

$$\begin{array}{rcl} C_4 & = & \int_0^T y_2^2(s, u_{1,1})\, ds > 0 \\ C_5 & = & \int_0^T y_3^2(s, u_{1,1})\, ds > 0 \\ C_6 & = & \int_0^T y_1^4(s, u_{1,1})\, ds > 0 \end{array} \tag{27}$$

Define $u_{\varepsilon,\delta} : [0, \delta T] \longrightarrow [-\varepsilon, \varepsilon]$ as usual by $u_{\varepsilon,\delta}(\delta t) = \varepsilon u_{1,1}(t)$, $(u_{\varepsilon,0} \equiv 0)$. Using the scaling-lemma (Lemma 2.1), i.e. $y(\delta t, u_{\varepsilon,\delta}) = \Delta_{\varepsilon,\delta}(y(t, u_[1, 1]))$, for system (24) we obtain for system (25)

$$x(\delta T, u_{\varepsilon,\delta}) = (0, 0, 0, \varepsilon^2\delta^5(C_4 - \varepsilon^2\lambda C_6)), \tag{28}$$

and for system (25)

$$x(\delta T, u_{\varepsilon,\delta}) = (0, 0, 0, \varepsilon^2\delta^7(C_5 - (\frac{\varepsilon}{\delta})^2\lambda C_6)) \tag{29}$$

Let $\eta_1 = \left(\frac{C_4}{\lambda C_6}\right)^{\frac{1}{2}}$ and $\eta_2 = \left(\frac{C_5}{\lambda C_6}\right)^{\frac{1}{2}}$. Then in the second system the family $\{u_{\eta_2 s,2s}\}_{s>0}$ generates $+k_1\frac{\partial}{\partial x_4}$, and the family $\{u_{2\eta_2 s,s}\}_{s>0}$ generates $-k_2\frac{\partial}{\partial x_4}$ as a tangent vector (both of order nine)(with suitable positive constants k_1, k_2.

In the first sytem the situation is more delicate. Here one can fix $\varepsilon > 0$ so small that $(C_4 - \varepsilon^2\lambda C_6) > 0$ and then consider the one-parameter family of control variations $\{u_{\varepsilon,s}\}_{s\geq 0}$ to generate $+\frac{\partial}{\partial x_4}$ as a fifth order tangent vector. But in general it will not be possible to generate $-\frac{\partial}{\partial x_4}$ as tangent vector. Here the only way to reach points on the negative x_4-axis is to modify the basic control $u = u_{1,1}$ in order to change the relative sizes of C_4 and C_6.

However, consider the C^1-function $\phi : \mathbf{R}^4 \longrightarrow \mathbf{R}$ defined by $\phi(x) = x_4 + \frac{1}{\varepsilon_0}|x_2|x_2x_1 + \frac{2}{3\varepsilon_0^2}x_2x_1^3$, and one easily verifies

$$\varepsilon_0^2(f+ug)(\phi(x(t,u)) = (\varepsilon_0|x_2| + 2x_1^2)(\varepsilon_0|x_2| + ux_2) + (\frac{2}{3} - \varepsilon_0^2\lambda)x_1^4. \tag{30}$$

Thus ϕ is nondecreasing along trajectories if $\lambda\varepsilon_0^2 \leq \frac{2}{3}$, and in this case the system is uncontrollable. (Then clearly no point in the x_1-plane with negative x_4-coordinate may be reached.) On the other hand one can explicitly show that the system is STLC if $\lambda\varepsilon_0^2 > \frac{2}{3}$. The most interesting case is however when equality holds. In this case we have that ϕ is constant along trajectories if the control satisfies the feedback-law $u = u(x) = -\text{sgn}x_2$. One easily sees that the corresponding trajectories are periodic, and that they are solutions to the Hamilton-Jacobi equations.

To view this critical case, e.g. $\varepsilon_0 = 1$ and $\lambda = \lambda_0 = \sqrt{\frac{2}{3}}$ from a different perspective one may first consider the following C^1 change of coordinates $(z_1, z_2, z_3, z_4) = (x_1, x_2, x_3, \phi(x))$, and then form the Hamiltonian $H(p, z, u) = p_1u + p_2z_1 + p_3z_2 + p_4\phi(x(z))$. If one interprets the derivatives in the system of equations

$$\dot{z}_j = \frac{\partial H(p,z,u)}{\partial p_j}, \;\; \dot{p}_j = -\frac{\partial H(p,z,u)}{\partial z_j}, j = 1,2,3,4. \tag{31}$$

in a generalized sense, then one finds that above periodic trajectories foliating part of the boundary of the attainable set correspond to solutions $u = -\text{sgn}z_2$, $p \equiv (0,0,0,-1)$, and $z(t)$ satisfies $|z_2| - 2z_1^2 \equiv 0$ and $z_4 = \mathit{const.}$. Remarkable is, that (only) after this nonsmooth change of coordinates these distinguished extremals correspond to singular solutions of the Hamilton Jacobi equations with constant co-state vector p.

The system (24) is closely related to the necessary condition Theorem 3.8, as essentially the construction of the (along trajectories nondecreasing) function $\phi(x)$ together with the estimates used to prove Theorem 3.5 show that the bad bracket $(\text{ad}^2[g,f]), f)$ cannot be neutralized by brackets containing four or more factors g. If one increases the number of integrators before *taking* *a square* then the situation becomes more intricate:

$$
\begin{array}{rclrcl}
\dot{x}_1 & = & u & x(0) & = & 0 \\
\dot{x}_2 & = & x_1 & |u(\cdot)| & \leq & \varepsilon_0 \\
\dot{x}_3 & = & x_2 & \lambda & > & 0 \\
 & \vdots & & & & \\
\dot{x}_{n-1} & = & x_{n-2} & & & \\
\dot{x}_n & = & x_{n-1}^2 - \lambda x_1^m & & &
\end{array}
\tag{32}
$$

Here the important brackets are $(\mathrm{ad}^2(\mathrm{ad}^{n-1}X, Y)X)$ and $(\mathrm{ad}^m Y, X)$ which are of types $(2, 2n-1)$ and $(m, 1)$, respectively. By Sussmann's theorem (Theorem 3.6) this system is STLC if $m < 2n$. On the other hand if one finds that this system is not $\mathrm{STLC}_\varepsilon$ for some $m = m(n) < \infty$, then by a similar reasoning as in Theorem 3.8 one may conclude that the bad bracket cannot be neutralized by brackets containing more than $m-1$ factors Y.

Due to the more intricate geometry of the boundary of the attainable set of system (32) the construction of a similar function as above ϕ becomes much more difficult, and we only arrive at the following condition which probably can be improved.

Proposition 5.1 *The system (32) is not $STLC_\varepsilon$ if $m \geq 2^n$.*

This proposition follows directly from the following lemma.

Lemma 5.2 *If $f \in C^{m-1}(\mathbf{R}, \mathbf{R})$ has an absolutely continuous $(m-1)$th derivative, $|f^{(m)}(\cdot)| \leq M$ a.e. and $|f^{(j)}(0) = f^{(j)}(t) = 0$ for $j = 0, 1, \ldots,$ then*

$$\int_0^t f^2(s)\, ds \geq K t^{2^m - 2m} \int_0^t (f^{(m-1)}(s))^{2^m}\, ds \tag{33}$$

where K is a constant depending on M and m only. (An estimate for K can be found in the proof, s.b.)

Proof. First consider a differentiable function $\xi \in C^1$ with absolutely continuous derivative $\xi' = \eta$, $|\eta'| \leq L$. Then for $m \geq 1$ and $k \geq 0$ compute

$$\frac{d}{dt}(|\xi|\xi^{2m-1}\frac{\eta^{2k+1}}{(2k+1)L}) = |\xi|\xi^{2m-1}\eta^{2k}(\frac{\eta'}{L}) + |\xi|^{2m-1}\eta^{2k+2}(\frac{2m}{(2k+1)L}), \tag{34}$$

and

$$
\begin{aligned}
\frac{d}{dt}\left(\xi^{2m-1}\frac{(2m)\eta^{2k+3}}{(2k+1)(2k+3)L}\right) &= \xi^{2m-1}\eta^{2k+2}\frac{2m}{(2k+1)L}(\frac{\eta'}{L}) \\
&\quad + \xi^{2m-2}\eta^{2k+4}(\frac{2m(2m-1)}{(2k+1)(2k+3)L^2}),
\end{aligned}
$$

Combining the appropriate terms and using that $(\frac{\eta'}{L}) \leq 1$ we may conclude that

$$
\begin{aligned}
\frac{d}{dt}\left\{\frac{(2m)(2m-1)}{(2k+1)(2k+3)L^2}\int_0^t \xi^{2m-2}(s)\eta^{2k+4}(s)\, ds\right\} \leq \frac{d}{dt}\left\{\int_0^t \xi^{2m}(s)\eta^{2k}(s)\, ds\right. \\
\left. + \frac{1}{(2k+1)L}|\xi|\xi^{2m-1}(t)\eta^{2k+1}(t) + \frac{2m}{(2k+1)(2k+3)L}\xi^{2m-1}(t)\eta^{2k+3}(t)\right\}
\end{aligned}
$$

Thus if $\eta(0) = \eta(t) = 0$ then upon integrating one obtains

$$\int_0^t \xi^{2m}(s)\eta^{2k}(s)\,ds \geq \frac{(2m)(2m-1)}{(2k+1)(2k+3)L^2}\int_0^t \xi^{2m-2}(s)\eta^{2k+4}(s)\,ds \tag{35}$$

and inductively

$$\int_0^t \xi^{2m}(s)\,ds \geq \frac{(2m)!}{(4m-1)!!}L^{2m}\int_0^t \eta^{2k+4}(s)\,ds \tag{36}$$

The lemma follows by successively applying inequality (36) to the derivatives of the function f, observing that if $|f^{(j)}(\cdot)| \leq L_j$, and $f^{(j-1)}(0) = f^{(j-1)}(t) = 0$, then $|f^{(j-1)}(s)| \leq \frac{1}{2}L_j t$ (this last expression may be taken as L_{j-1} in the next step).

In both the examples 5.1 and 5.2 the bad brackets which are to balance each other are of different types, and thus the systems are sensitive w.r.t. the usual scaling operations on the control $u_{1,1} \longrightarrow u_{\varepsilon,\delta}$. In the following we consider a system in which two different bad brackets of the same type have to balance each other in order to obtain STLC. The lowest order such brackets are of length nine, as one can check readily by writing out all possible lower order systems.

Example 5.3.

$$\begin{array}{rclcrcl} \dot{x}_1 &=& u & \qquad & x(0) &=& 0 \\ \dot{x}_2 &=& x_1 & & |u(\cdot)| &\leq& \varepsilon_0 \\ \dot{x}_3 &=& x_1^3 & & \lambda &>& 0 \\ \dot{x}_4 &=& x_3^2 - \lambda x_2^2 x_1^4 & & & & \end{array} \tag{37}$$

In this case the brackets which do not vanish at zero are $g(0) = \frac{\partial}{\partial x_1}$, $[g,f](0) = \frac{\partial}{\partial x_2}$, $(\mathrm{ad}^3 g, f)(0) = 6\frac{\partial}{\partial x_3}$, which are of types $(1,0), (1,1)$ and $(3,1)$, respectively, and $(\mathrm{ad}^2[g,f], (\mathrm{ad}^4 g, f))(0) = -48\lambda\frac{\partial}{\partial x_4}$, and $(\mathrm{ad}^2(\mathrm{ad}^3 g, f), f)(0) = 72\frac{\partial}{\partial x_4}$, which are both of type $(6,3)$. Thus the system is homogeneous w.r.t. the two-parameter of dilations $\Delta_{\varepsilon,\delta}(x) = (\varepsilon\delta x_1, \varepsilon\delta^2 x_2, \varepsilon^3\delta^4 x_3, \varepsilon^6\delta^9 x_4)$ and thus the system is STLC or equivalently STLC$_\varepsilon$ if and only if zero lies in the interior of the attainable set at any positive time for any positive control bound.

As before one easily obtains $\frac{\partial}{\partial x_j}$, $j = 1,2,3$ as tangent vectors of order $1, 2$, and 4, respectively.

Note, that because of the homogenuity of the attainable set (or the control system) it suffices to show that for some time $T > 0$ and some control bound $\varepsilon_0 > 0$ the attainable set $\mathcal{A}_{\varepsilon_0}(t)$ intersects both the negative and the positive x_4-axis in order to conclude that the system is STLC (even STLC$_\varepsilon$) (noting that we already obtained the tangent vectors $\pm\frac{\partial}{\partial x_j}$ $j = 1,2,3$).

First consider the following one-parameter family of controls $u^\eta : [0, 4+\eta] \longrightarrow [-1,1]$, $\eta \geq 0$, defined by

$$u^\eta(t) = \begin{cases} 1 & \text{if} \quad t \in [0,1) \cup [3+\eta, 4+\eta] \\ 0 & \text{if} \quad t \in [2, 2+\eta) \\ -1 & \text{if} \quad t \in [1,2) \cup [2+\eta, 3+\eta) \end{cases} \tag{38}$$

Then $x_j(4+\eta, u^\eta) = 0$ for $j = 1,2,3$ and every $\eta \geq 0$, and $x_4(4+\eta, u^\eta) = x_4(4, u^0) + \eta x_3^2(1, u^0)$. Note, that $x_3(1, u^0) > 0$ and $x_1(t, u^\eta) = 0$ for $t \in [2, 2+\eta]$) and that therefore $x_4(4+\eta, u^\eta)$ eventually becomes positive for η sufficiently large. This means that we found a control steering to a point on the positive x_4-axis.

For $\lambda \gg 1$ one also has $x_4(4, u^0) < 0$ and one is done.

The interesting case however is $0 < \lambda \ll 1$. (The intermediate case can be handled by a combination of the controls given above with those introduced below.) We first define the following basic control element $u^1 : [0,T] \longrightarrow [-1,1]$ where $T = 4 + 2\sqrt[4]{2}$ defined by

$$u^1(t) = \begin{cases} 1 & \text{if} \quad t \in [0,1) \cup [2+\sqrt[4]{2}, 3+\sqrt[4]{2}) \\ -1 & \text{if} \quad t \in [1, 2+\sqrt[4]{2}) \cup [3+\sqrt[4]{2}), T] \end{cases} \tag{39}$$

For two controls $v_j : [0,T_j] \longrightarrow \mathcal{U}$ we define their concatenation $v_1 * v_2 : [0, T_1+T_2] \longrightarrow \mathcal{U}$ by

$$v_1 * v_2(t) \;=\; \begin{cases} v_1(t) & \text{if} \quad t \in [0,T_1) \\ v_2(t) & \text{if} \quad t \in [T_1, T_1+T_2] \end{cases} \tag{40}$$

The time-reversed control u^{-1} is defined by $u^{-1}(t) = u^1(T-t)$, and inductively let $u^{k+1} = u^1 * u^k$, $u^{-k-1} = u^{-1} * u^{-k}$, and finally $u_k = u^k * u^{-k}$.

From $x_j(T,u^1) = x_j(T,u^{-1}) = 0$ for $j = 1,3$ it follows that $x_j(2kT, u_k) = 0$ for $j = 1,3$. Similarily, from $x_2(T,u^1) = -x_2(T,u^{-1}) = 2-\sqrt{2}$ we obtain $x_2(jT,u_k) = x_2((2k-j)T,u_k) = j(2-\sqrt{2})$, $j = 0,1,\ldots,k$. We let $C_{41} = \int_0^T x_3^2(s,u^1)\,ds$ and compute

$$\begin{aligned} \int_0^{2kT} x_1^4(s,u_k)x_2^2(s,u_k)\,ds \;&=\; 2\textstyle\sum_{j=1}^{k}\int_0^T x_1^4(s,u^1)(j(2-\sqrt{2}) + x_2^2(s,u^1)\,ds \\ &=\; 2\textstyle\sum_{j=0}^{k-1}(A_2 j^2 + A_1 j + A_0) \\ &\qquad \text{for some constants } A_0, A_1, A_2 \text{ with } A_2 > 0 \\ &=\; C_{42}k^3 + o(k^3) \text{ with } C_42 > 0. \end{aligned} \tag{41}$$

(Here $o(k^m)$ stands for terms such that $\lim_{k\to\infty} \frac{o(k^m)}{k^m} = 0$.) Thus $x_4(2kT,u_k) = 2kC_{41} - \lambda(k^3C_{42} - o(k^3))$, which is negative for sufficiently large k. Detailed analysis shows $C_{41} \approx C_{42} \approx 1$, and thus $k \approx \lambda^{-1/2}$ gives an approximation for the number of switchings employed to reach a point on the negative x_4-axis. Obviously this growing number of switchings reflects that the system becomes *increasingly uncontrollable* as the parameter λ approaches 0 from above.

(Here we are only concerned about small-time controllability of the system with no claims about optimality. However, it seems reasonable to expect that any time-optimal control steering to a point on the negative axis will also have to change the sign more and more often as the parameter λ becomes small. Since we do not have a priori regularity estimates for such optimal control, a reasonable choice to measure the regularity of the control is the number of changes of sign of the absolutely continuous function $x(\cdot,\tilde{u})$ where $\tilde{u}$ is such an optimal control. For details see following section.)

6 Fast switching control variations

In the previous section in Example 5 we used control variations whose number of switchings depended on the parameter λ in the way that if λ decreases, i.e. the system becomes increasingly uncontrollable, then the number of switchings needed to control the system increases. However, for each system (i.e. each fixed λ) the number of switchings was the same for all times $t > 0$.

In this section we shall consider systems which only can be controlled by means of increasingly faster switching control variations; the number of switchings increasing to infinity as the disposable time $t > 0$ decreases to 0.

The importance of such controls for the purpose of controllability lies in that they allow one to control systems which cannot be controlled by means of the usual controls, and for which in particular Sussmann's theorem (Theorem 3.6) fails; but also these controls promise to have some impact on the theory of (regularity of) time-optimal controls, as in particular for such systems there cannot be any bound on the number of switchings of optimal controls. An interesting question is whether also the *switching surfaces* of such systems are irregular, in the sense that they may be no more locally finite unions of embedded manifolds.

Before explicitly discussing the systems requiring such controls, recall the usual construction of families of control variations, as for example in any of the standard proofs of the different versions of the maximum principle:

Usually one starts with a reference control u_0 defined on some time interval $[0, t]$ (here we may for ease assume that $u_0 \equiv 0$) and perturbation data $\Gamma = \{t_1, \ldots, t_r, \delta_1, \ldots, \delta_r, \varepsilon_1, \ldots, \varepsilon_r\}$, where the t_j are times in the interval $[0, T]$ at which the perturbations are to be taken, and these are specified by defining $u^\Gamma(t) = \varepsilon_j$ if $t \in [t_j - \delta_j, t_j]$ and $u^\Gamma(t) = u_0(t)$ else. Then one considers the curve generated by the terminal points $x(T, u^\Gamma)$ when $\delta_j, \varepsilon_j \longrightarrow 0$, $j = 1, \ldots, r$. Finally one considers the tangent vectors to this curve at $x(T, u_0)$ which under suitable hypotheses form an infinitesimal convex cone, which via an associated open mapping theorem gives the desired sufficient conditions for local controllability or necessary conditions for optimality.

Recently many refinements of the outlined technique have been introduced, for example one may also consider control variations obtained by deleting and inserting pieces of the control. (Here it is important to make sure that the time intervals $[t_j - \delta_j, t_j]$ do not overlap, and that insertions and deletions are done the proper order.) Moreover, one may obtain refined results by letting the various parameters go to zero at different rates, etc. All of the control variations obtained in this way have in common that one always works with a finite number of times t_j, and in particular essentially always works with control variations with a fixed number of switchings, and these families of control variations are parametrized by the continuous parameters δ_j, ϵ_j. Here we show that in order to obtain sharper results for both controllability and optimality it is essential to also consider control variations obtained by *increasing* the number of switchings; the technique is outlined in the following example. The difference between these new and the standard control variations is sketched in the figures (6) and (6), both illustrating the case of $r = 1$.

Consider the following system on $\mathbf{R}^4$, which for reasons explained in [13] is the *lowest order* system for which such phenomena occur.

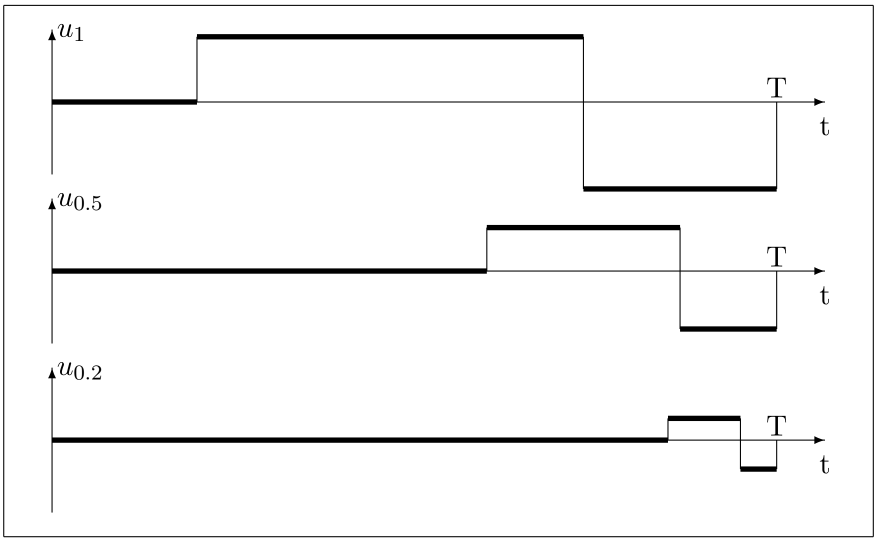

Figure 2: Standard Control Variation

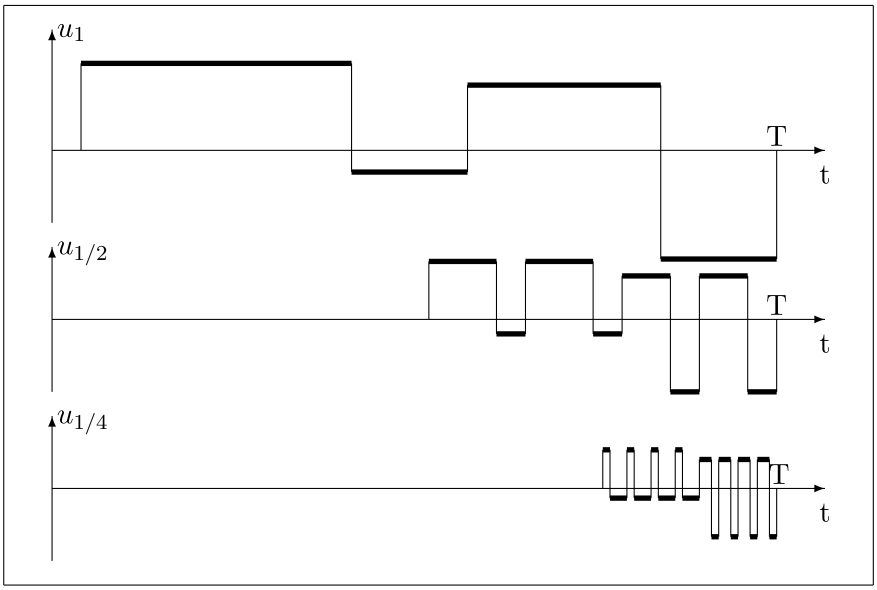

Figure 3: New Control Variation

Example 6.1

$$\begin{array}{rclcrcl} \dot{x_1} & = & u & \qquad & x(0) & = & 0 \\ \dot{x_2} & = & x_1 & & |u(\cdot)| & \leq & \varepsilon_0 \\ \dot{x_3} & = & x_1^3 & & & & \\ \dot{x_4} & = & x_3^2 - x_2^7 & & & & \end{array} \tag{42}$$

In this case the behaviour of the system is determined by the brackets $g(0) = \frac{\partial}{\partial x_1}$, $[g,f](0) = \frac{\partial}{\partial x_2}$, $(\mathrm{ad}^3 g, f)(0) = 6\frac{\partial}{\partial x_3}$, $f^\beta(0) = (\mathrm{ad}^2(\mathrm{ad}^3 g, f), f)(0) == 72\frac{\partial}{\partial x_4}$, and $f^\gamma(0) = (\mathrm{ad}^7[g,f], f)(0) = -7!\frac{\partial}{\partial x_4}$. The first three brackets are of types $(1,0)$, $(1,1)$, and $(3,1)$, respectively, and in particular contain an odd number of factors of the controlled field g. The fourth bracket f^β is of type $(6,3)$, and clearly an possible obstruction to STLC. It is only linear dependent on the fifth bracket f^γ, which is of type $(7,8)$, i.e. contains both more factors g and more factors f, and thus by the standard theorems can not neutralize the bracket f^β.

For $N \in \mathbf{Z}^+$, T, $\varepsilon_0 > 0$ we denote by $\mathcal{U}_N(T,\varepsilon_0)$ the set of measurable controls $u : [0,T] \longrightarrow [-\varepsilon,\varepsilon]$ which are such that $t \longrightarrow x_1(t,u) = \int_0^t u(s)\,ds$ changes the sign at most $(N-1)$ times. With this notation we claim the following:

Claim 6.1 *The system (42) is STLC$_\varepsilon$.*

Claim 6.2 *If $u \in \mathcal{U}_N(T,\varepsilon_0)$, $x_1(T,u) = 0$, and $N^7 \leq \varepsilon^{\frac{3}{4}} T^{\frac{7}{2}}$, then $x_4(T,u) \geq 0$.*

Claim 6.3 *If x_2^7in the last equation of (42) is replaced by x_2^8, then the system is not STLC.*

Note, that in particular if $u : [0,T] \longrightarrow [-\varepsilon_0, \varepsilon_0]$ is piecewise constant with at most N jumps, or is piecewise smooth and changes the sign at most N times, then $u \in \mathcal{U}_N(T,\varepsilon_0)$. However, in general the control u is only required to be integrable, and thus the number of changes of sign of $x_1(\cdot,u) = \int_0^t u(s)\,ds$ is the meaningful parameter.

To prove Claim 6.1, first note that the (x_1,x_2,x_3)-subsystem is STLC by Theorem 3.4, and thus one easily obtains $\pm\frac{\partial}{\partial x_j}$, $j = 1,2,3$ as tangent vectors of respective orders $1,2,4$. Also, because the bracket X^β is *shorter* than the bracket X^γ, one easily obtains $+\frac{\partial}{\partial x_4}$ as 9th order tangent vector. (For an explicit construction of the associted families of control variations see Kawski [15].) To generate $-\frac{\partial}{\partial x_4}$ as a tangent vector to the attainable set, construct a family of control variations with an increasing number of switchings as follows: Fix a control $v : [0,T] \longrightarrow [-1,1]$ (for some $T > 0$) such that $x_1(T,v) = x_3(T,v) = 0$ and $x_2(T,v) > 0$. (For example one may take $T = 4 + 2\sqrt{2}$, $v(t) = 1$ if $t \in [0,1) \cup [2+\sqrt{2}, 3+2\sqrt{2})$ and $v(t) = -1$ otherwise.)

Denote by v^{-1} the *time reversed* control (i.e. $v^{-1}(t) = v(T-t)$) and inductively define via concatenation $u^{(1)} = v * v^{-1}$ (i.e. $u^{(1)}(t) = v(t)$ if $0 \leq t \leq T$ and $u^{(1)}(t) = v^{-1}(t-T)$ if $T \leq t \leq 2T$), and $u^{(k+1)} = v * u^{(k)} * v^{-1} : [0, 2kT] \longrightarrow [-1,1]$. For given $t_0, \varepsilon_0 > 0$ let $\delta = \delta(t_0,k) = t_0/(2kT)$ and define $u^{(k)}_{\varepsilon_0,\delta} : [0,t_0] \longrightarrow [\varepsilon_0, \varepsilon_0]$ by $u^{(k)}_{\varepsilon_0,\delta}(\delta t) = \varepsilon_0 u^{(k)}(t)$. One easily verifies that $x_j(t_0, u^{(k)}_{\varepsilon_0,\delta}) = 0$ for $j = 1,2,3$ and computes

$$x_4(t_0, u^{(k)}_{\varepsilon_0,\delta}) = \int_0^{t_0} x_3^2(s, u^{(k)}_{\varepsilon_0,\delta})\,ds - \int_0^{t_0} x_2^7(s, u^{(k)}_{\varepsilon_0,\delta})\,ds$$

$$
\begin{aligned}
&= 2\,{\varepsilon_0}^6\delta^9 k\int_0^T x_2(s,v)\,ds - 2\,{\varepsilon_0}^7\delta^1 5\sum_{j=0}^{k-1}\int_0^T (jx_2(T,v)+x_2(s,v))^7\,ds \\
&= {\varepsilon_0}^6 t_0 {k_{-1}}^9 kC_1 - {\varepsilon_0}^7 t_0 {k_{-1}}^1 5k^8(C_2+O(\frac{1}{k})) \\
&= {\varepsilon_0}^6 t_0^9 k^{-8} C_1 - {\varepsilon_0}^7 t_0^1 5k^{-7}(C_2+O(\frac{1}{k})) \qquad (43)
\end{aligned}
$$

Here C_1 and C_2 are some positive constant depending on the choice of the initial control v and $O(\frac{1}{k})$ stands for terms such that $O(\frac{1}{k}) \longrightarrow 0$ when $k \longrightarrow \infty$. Now choosing $k = k(\varepsilon_0, t_0) = C{\varepsilon_0}^{-1}t_0^{-6}$ (with C a sufficiently large constant) one obtains $x(t_0, u^{(k)}_{\varepsilon_0,\delta}) = (0,0,0,-Kt_0^{57} + o(t_0^{57}))$ for some constant $K > 0$, and thus generates $-K\frac{\partial}{\partial x_4}$ as a 57-th order tangent vector. Controllability follows from Theorem 2.4.

To prove claim 6.2, let $N \in \mathbf{Z}^+$ and $\varepsilon_0, T > 0$ satisfy $N^7T^{7/2}\varepsilon_0^{3/4} < 1$ and let $u \in \mathcal{U}_N(T,\varepsilon_0)$ be such that $x_1(T,u) = 0$. Choose $0 = t_0 \le t_1 \le \ldots \le t_r = T$, $(r \le N)$ such that $x_1|_[t_j, t_{j+1}]$ is of constant sign for $j = 0,1,\ldots,r-1$. Let $A = \int_0^T x_2^7(s,u)\,ds$; we may assume $0 < A < 1$. Clearly there is $\bar{t} \in [0,T]$ such that $x_2(\bar{t},u) > (A/T)^{1/7} = B$ and thus $x_2(\cdot,u)$ increases by at least $C = B/N$ on at least one subinterval $I_{j_0} = [t_{j_0}, t_{j_0+1}]$.

Since $x_1(\cdot,u)$ is nonnegative on this subinterval we may use the Hölder inequality to conclude that

$$
x_3(t_{j_0+1},u) - x_3(t_{j_0},u) = \int_{t_{j_0}}^{t_{j_0}+1} x_1^3(s,u)\,ds \ge (C^3/T^2) = D. \qquad (44)
$$

We may assume that $x_3(t_{j_0},u) \le (-1/2)D < 0$ (otherwise $x_3(t_{j_0+1},u) \ge (+1/2)D < 0$ leads via similar reasoning to the same conclusion). From $|u(\cdot)| \le \varepsilon_0$ and $x_1(t_{j_0},u) = 0$ we we know that $x_3(t_{j_0}+s,u) \le (-1/2)D + \frac{1}{4}\varepsilon_0^3 s^4$ for $0 \le s \le s_0 = (2D\varepsilon_0^{-3})^{1/4}$ and finally

$$
\int_0^T x_3^2(s,u)\,ds \ge \int_0^{s_0}(-\frac{1}{2}D + \frac{1}{4}\varepsilon_0^3 s^4)^2\,ds \ge (\frac{A}{T}^{\frac{27}{28}} N^{-\frac{27}{4}}{\varepsilon_0}^{-\frac{3}{4}}T^{-\frac{9}{2}} \ge A. \qquad (45)
$$

To prove Claim 6.3 one proceds similar as in Section 5 by constructing a C^1-function $\phi : \mathbf{R}^4 \longrightarrow \mathbf{R}$, here $\phi(x) = x_4 + \frac{1}{\varepsilon_0}x_1x_3|x_3| + \frac{2}{\varepsilon_0^2}x_1^5x_3$. Note that $\phi(0) = 0$ and $\nabla\phi(0) \ne 0$ and thus $\phi^{-1}(0)$ locally defines a C^1 hypersurface $\mathcal{H}$ through the origin. We show that for $t, \varepsilon_0 > 0$ sufficiently small the intersection of the attainable set (of 42 with x_2^7 replaced by x_2^8) with the subspace $E_1 = \{x \in \mathbf{R}^4 : x_1 = 0\}$ lies on one side of $\mathcal{H}$, and thus the system is not STLC.

The function ϕ does not necessarily increase along trajectories, but

$$
\frac{d}{dt}\phi(x) = (|x_3| + \frac{u}{\varepsilon_0}x_3)(\frac{2}{\varepsilon_0^2}x_1^4 + |x_3|) + (\frac{2}{7\varepsilon_0^2}x_1^8 - x_2^8). \qquad (46)
$$

The first summand on the right hand side of 46 is clearly nonnegative. Also, a simple application of the Hölder inequality gives

$$
\int_0^t x_2^8(s,u)\,ds = \int_0^T \cdot (\int_0^s x_1(\sigma,u)\,d\sigma)^8\,ds \le t^8\int_0^8 x_1^8(s,u)\,ds. \qquad (47)
$$

Thus for any admissible control $u : [0,t] \longrightarrow [\varepsilon_0, \varepsilon_0]$ we may conclude that $\phi(x(t,u)) \ge (\frac{2}{7\varepsilon_0^2} - t^8)\int_0^t x_1^8(s,u)\,ds \ge 0$ if $\varepsilon_0, t > 0$ are sufficiently small.

For a first generalization of this example see Kawski [15]. (The detailed analysis of what precisely are the admissible partial orderings on the homogeneous elements in the free Lie algebra, induced by these new families of control variations, is subject of ongoing research.)

7 Appendix

The proof of theorem 2.4 uses Gronwall's inequality in several places in the following form:
Gronwall's lemma. *If the vector field system (1) has Lipschitz constant L on $\mathcal{A}_\mathcal{U}(2T)$, i.e.*

$$|f^0(x_0) - f^0(y_0) + \sum_{i=1}^{\kappa} u_i(f^i(x_0) - f^i(y_0)| \leq L|x_0 - y_0| \tag{48}$$

for all $x_0, y_0 \in \mathcal{A}_\mathcal{U}(2T)$ and all $u \in \mathcal{U}$ then $|x(t,u)(p) - x(t,u)(0) - p| \leq |p|te^{Lt}$ for all $0 \leq t \leq T$, $u(\cdot) \in \mathcal{U}$ and $p \in \mathcal{A}_\mathcal{U}(T)$. In particular, for any $\delta > 0$ one may choose $T = T(\delta) > 0$ such that $Te^{LT} \leq \delta$ and thus $|x(t,u)(p) - x(t,u)(0) - p| \leq \delta|p|$ for all $0 \leq t \leq T$ and $p \in \mathcal{A}_\mathcal{U}(T)$.
Proof of theorem 2.4 It suffices to consider for $\overline{\mathbf{K}'}$ *elementary simplicial cones* of the form $\overline{C}(v^1,\ldots,v^n) = \{\sum_{j=1}^n \lambda_j v^j : \lambda_j \geq 0\}$ by means of the following
Fact. *If $\overline{\mathbf{K}'}$ and $\overline{\mathbf{K}}$ are closed convex cones (both with vertex 0) such that $\overline{\mathbf{K}'} \setminus \{0\} \subseteq int\overline{\mathbf{K}}$, then there is a finite number of simplicial cones $\overline{C^i}(v^{i,1},\ldots,v^{i,n})$, $i = 1,\ldots,\rho$, such that $\overline{\mathbf{K}'} \subseteq \cup_{i=1}^{\rho}\overline{C^i}$ and $\overline{C^i} \setminus \{0\} \subseteq int\overline{\mathbf{K}}$ for each $i = 1,\ldots,\rho$.* (To prove this fact note that for example the intersection of the cone $\overline{\mathbf{K}'}$ with the unit $(n-1)$-sphere S^{n-1} is compact and consider the open cover of this set by the interiors of all elementary simplicial cones lying inside $\overline{\mathbf{K}}$.) Note that the number ρ of simplicial cones may be very large as is illustrated in $\overline{\mathbf{K}} = \{v \in \mathbf{R}^3 : v_3 \geq (v_1^2 + v_2^2)^{1/2}\}$ and $\overline{\mathbf{K}'} = \{v \in \mathbf{R}^3 : v_3 \geq (1+\varepsilon)(v_1^2 + v_2^2)^{1/2}\}$ for $\varepsilon > 0$ small.

Next we assume that the elementary simplicial cone $\overline{\mathbf{K}'} = \overline{C}(v^1,\ldots,v^n)$ has nonempty interior, and thus there is a linear change of coordinates such that $e^i = (0,\ldots,0,1,0,\ldots,0) \in \overline{\mathbf{K}}$ and $v^i = (1-n\varepsilon)e^i + \sum_{j=1}^n \varepsilon e^j = (\varepsilon,\ldots,\varepsilon,1-(n-1)\varepsilon,\varepsilon,\ldots,\varepsilon)$, $i = 1,\ldots n$ for some $\varepsilon > 0$ which we may take so small that $4\varepsilon n^2 \leq 1$.

It is convenient to work with the associated one-norm, i.e. $\|\sum_{j=1}^n \alpha_j e^j\| := \sum_{j=1}^n |\alpha_j|$. This, and the change of basis of $\mathbf{R}^n$ clearly affect the size of the constant C in the theorem by a multiplicative factor. Note, that in these local coordinates a vector $w = (w_1,\ldots,w_n)$ is in $\overline{\mathbf{K}'}$ if and only if $w_i \geq \varepsilon\|w\| = \varepsilon\sum_{j=1}^n |w_j|$ for each $i = 1,\ldots,n$. First fix $T_1 > 0$ such that for all $0 \leq t \leq T_1$, $p \in \mathcal{A}_\mathcal{U}(T_1)$ and $u(\cdot)$ admissible

$$\|x(t,u)(p) - x(t,u)(0) - p\| \leq \frac{\varepsilon}{4n}\|p\|. \tag{49}$$

Fix families $\{u_s^{(i)}\}_{s\geq 0}$ of control variations generating the m-th order tangent vectors e^i, $i = 1,\ldots,n$. Fix $0 < T \leq T_1$ such that

$$\|x(t,u_t^{(i)})(0) - t^m e^i\| \leq \varepsilon^2 t^m \text{ for all } 0 \leq t \leq T \text{ and } i = 1,\ldots,n. \tag{50}$$

Let $C = (2nm)^{-m}$. For every $t \leq T$ and for every point $z^\infty \in \overline{\mathbf{K}'} \cap B(0,Ct^m)$ we will construct a control $\overline{v^\infty}$ (as a L^1-limit) such that $x(t,\overline{v^\infty})(0) = z^\infty$. Fix $z^\infty \in \overline{\mathbf{K}'} \cap B(0,Ct^m)$ and let

$\tau_\infty = 2nm\|z^\infty\|^{1/m}$.

Notation: For any control u defined on $[0,\tau']$ with $\tau' \le \tau_\infty$ we define the control $\overline{u}$ on $[0,\tau^\infty]$ by

$$\overline{u}(t) = \begin{cases} 0 & \text{if} \quad 0 \le t < \tau_\infty - \tau' \\ u(t-(\tau_\infty - \tau')) & \text{if} \quad \tau_\infty - \tau' \le t \le \tau_\infty. \end{cases} \tag{51}$$

Let $z^0 = 0$, $\tau_0 = 0$, and $\overline{v^0} \equiv 0$. Inductively, suppose z^i, τ^i and v^i (defined on $[0,\tau_i]$ have been defined. Choose $\mu(i+1) \in \{1,\ldots,n\}$ such that

$$(z^\infty - z^i)_{\mu(i+1)} \ge (z^\infty - z^i)_l \text{ for } l = 1,\ldots,n. \tag{52}$$

Let $t_{i+1} = (1-n\varepsilon)(z^\infty - z^i)^{\frac{1}{m}}_{\mu(i+1)}$ and $\tau_{i+1} = \tau_i + t_{i+1}$ $(= \sum_{j=0}^{i+1} t_j)$

$$v^{i+1}(s) = \begin{cases} u^{\mu(i+1)}_{t_{i+1}}(s) & \text{if} \quad 0 \le s < t_{i+1} \\ v^I(s - t_{i+1}) & \text{if} \quad t_{i+1} \le s \le \tau_{i+1}. \end{cases} \tag{53}$$

Let $y^{i+1} = x(t_{i+1}, u^{\mu(i+1)}_{t_{i+1}})(0)$ and $z^{i+1} = x(\tau_{i+1}, v^{i+1})(0) = x(\tau_i, v^i)(y^{i+1})$. To verify that $\|z^\infty - z^i\| \longrightarrow 0$ as $i \longrightarrow \infty$ first observe that

$$\|y^{i+1} - t^m_{i+1} e^{\mu(i+1)}\| \le \varepsilon^2 t^m_{i+1}. \tag{54}$$

and thus

$$\|y^{i+1}\| \le (1+\varepsilon^2) t^m_{i+1} \le \|z^\infty - z^i\|. \tag{55}$$

Using Gronwall's inequality and the last inequality as an estimate for $\|y^{i+1} - 0\|$ we have

$$\|z^{i+1} - (y^{i+1} + z^i)\| \le \frac{\varepsilon}{4n}\|y^{i+1}\|. \tag{56}$$

By choice of $\mu(i+1)$

$$\begin{aligned} \|(1-n\varepsilon)(z^\infty - z^i) - t^m_{i+1} e^{\mu(i+1)}\| &= (1-n\varepsilon)\textstyle\sum_{l \ne \mu(i+1)} (z^\infty - z^i)_l \\ &\le (1-n\varepsilon)\tfrac{n-1}{n}\|z^\infty - z^i\|. \end{aligned} \tag{57}$$

Combining the inequalities (54),(55) and (57) results in

$$\begin{aligned} \|z^\infty - z^{i+1}\| &\le n\varepsilon\|z^\infty - z^i\| + \|(1-n\varepsilon)(z^\infty - z^i) - t^m_{i+1} e^{\mu(i+1)}\| \\ &\quad + \|t^m_{i+1} e^{\mu(i+1)} - y^{i+1}\| + \|z^{i+1} - (z^i + y^{i+1})\| \\ &\le (n\varepsilon + (1-n\varepsilon)\tfrac{n-1}{n} + \varepsilon^2 + \tfrac{\varepsilon}{4n})\|z^\infty - z^i\| \\ &\le \tfrac{2n-1}{2n} \text{ by choice of } \varepsilon < \tfrac{1}{4n^2}. \end{aligned} \tag{58}$$

Therefore inductively $\|z^\infty - z^{i+1}\| \le (\frac{2n-1}{2n})^{i+1}\|z^\infty\| \longrightarrow)$ as $i \longrightarrow \infty$. Next observe that $t_{i+1} \le \|z^\infty - z^i\|^{1/m} \le ((\frac{2n-1}{2n})^i\|z^\infty\|)^{1/m}$ and thus

$$\begin{aligned} \tau_i \quad + \quad & \textstyle\sum_{j=o}^{i} \le \|z^\infty\|^{1/m} \sum_{j=0}^{i} (\frac{2n-1}{2n})^{j/m} \\ < \quad & \|z^\infty\|^{1/m} \frac{1}{1-(1-\frac{1}{2n})^{1/m}} \le 2nm\|z^\infty\|^{1/m} \le \tau_\infty \\ & \text{using } (1-\tfrac{1}{2n})^{1/m} \le 1 - \tfrac{1}{2nm}. \end{aligned} \tag{59}$$

At this point it becomes apparent why we tangent vectors of some finite order m to compose infinitely many corrections in a finite time in order to reach the point z^∞. Finally we have to verify that $(z^\infty - z^{i+1}) \in \overline{\mathbf{K'}}$ for all $i = 1, \dots$. (Since one cannot go backwards we need $t_{i+1} \geq 0$.)

By hypothesis $(z^\infty - z^0) = z^\infty \in \overline{\mathbf{K'}}$. Suppose $(z^\infty - z^i) \in \overline{\mathbf{K'}}$, i.e. $(z^\infty - z^i)_l \geq \varepsilon \| z^\infty - z^i \|$, $l = 1, \dots, n$. Also $(z^\infty - z^i)_{(\mu(i+1)} \geq \frac{1}{n} \| (z^\infty - z^i) \|$ by choice of $\mu(i+1)$. Using these together with the inequalities (54),(55) and (57) componentwise, one finds:

$$(z^\infty - z^{i+1})_{\mu(i+1)} - n\varepsilon (z^\infty - z^i)_{\mu(i+1)} \leq (\varepsilon^2 + \frac{\varepsilon}{4n}) \| z^\infty - z^i \|, andthus \tag{60}$$

$$\begin{array}{rcl} \frac{(z^\infty - z^{i+1})_{\mu(i+1)}}{\| z^\infty - z^{i+1} \|} & \geq & \frac{n\varepsilon (z^\infty - z^i)_{\mu(i+1)} - (\varepsilon^2 + \frac{\varepsilon}{4n}) \| z^\infty - z^i \|}{\frac{2n-1}{2n} \| z^\infty - z^i \|} \\ & \geq & \frac{2n}{2n-1} (\frac{n\varepsilon}{n} - \varepsilon^2 - \frac{\varepsilon}{4n}) \\ & \geq & \varepsilon \frac{4n-1-\varepsilon}{2(2n-1)} \geq \varepsilon \text{ by choice of } \varepsilon. \end{array} \tag{61}$$

Similarily, for $l \neq \mu(i+1)$:

$$(z^\infty - z^{i+1})_l = (z^\infty - z^i)_l + (z^i + y^{i+1} - z^{i+1})_l - y_l^{i+1}, \tag{62}$$

and thus

$$\begin{array}{rcl} \frac{(z^\infty - z^{i+1})_l}{\| z^\infty - z^{i+1} \|} & \geq & \frac{(z^\infty - z^i)_l - (\varepsilon^2 + \frac{\varepsilon}{4n}) \| z^\infty - z^i \|}{\frac{2n-1}{2n} \| z^\infty - z^i \|} \\ & \geq & \frac{2n}{2n-1} (\varepsilon - \varepsilon^2 - \frac{\varepsilon}{4n}) \geq \varepsilon. \end{array} \tag{63}$$

Here it becomes apparent, that to make the inductive step work one needs that $\overline{\mathbf{K'}} \setminus \{0\}$ is properly contained in the interior of $\overline{\mathbf{K}^m}$. Also, one has to aim at a point short of $(z^\infty - z^i)$ in order to avoid *overshooting* (because of the several errors incurred in each step), so that one can still do further corrections.

Finally, note that the order in which the controls $u_{t_{i+1}}^{\mu(i+1)}$ and v^i are composed is the natural one when trying to synthesize stabilizing feedback for the *time reversed* system, i.e. first one takes a rough shot at the equilibrium, and later when the system is already close to the equilibrium one uses successively finer corrections. The advantage of synthesizing stabilizing feedback as in this proof lies in that this construction only requires the knowledge of $2n$ one-parameter families of controls, rather than one $2n$-parameter (or one $n-$parameter) family of controls.

References

[1] R. M. Bianchini and G. Stefani, *Sufficient conditions of local controllability*, Proc. $25^t h$ IEEE-Conf. Decision & Control, (1986) Athens.

[2] R. M. Bianchini and G. Stefani, *Local controllability about a reference trajectory* Analysis and Optimization of Systems, A. Bensoussand & J. L. Lions,eds., Lect. Notes Control, **83** Berlin (1986), Springer.

[3] A. Bressan, *The generic local time-optimal stabilizing controls in dimension* 3, SIAM J. Cntrl. & Opt., **24** (1986) 177-190.

[4] N. Bourbaki, Lie Groups and Lie Algebras, Part I., Paris (1975) Hermann.

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

[6] P. E. Crouch and C. I. Byrnes, *Local accessibility, local reachability and representations of local groups*, Math. Systems Theory **19** (1986) 43-65.

[7] H. Frankowska, *An open mapping principle for set-valued maps*, Centre des Récherches Mathématiques, Univ. Montréal, rapport CRM-1344 (1986).

[8] H. Frankowska, *Local controllability of control systems with feedback*, Centre des Récherches Mathématiques, Univ. Montréal, rapport CRM-1364 (1986).

[9] R. Goodman, *Nilpotent Lie Groups: Structure and Applications to Analysis*, Lect. Notes Math. **562** (1976) Springer.

[10] H. G. Hermes, *Controlled Stability*, Annali di Matematica pura ed applicata IV, **CXIV** (1977) 103-119.

[11] H. G. Hermes and M. Kawski, *Local controllability of a single-input, affine system*, to appear in Proc. 7^{th} Int. Conf. Nonlinear Analysis, Dallas (1986).

[12] V. Jurdjevic and I. Kupka, *Polynomial control systems*, Math. Ann. **272** (1985) 361-368.

[13] M. Kawski, *Nilpotent Lie algebras of vectorfields and local controllability of nonlinear systems*, Ph. D. Dissertation, Univ. of Colorado, Boulder (1986).

[14] M. Kawski, *A new necessary condition for local controllability*, to appear in AMS Contemporary Math. Series, Proc. Conference on Diff. Geometry, San Antonio (1986).

[15] M. Kawski, *Control variations and local controllability*, Proc. MTNS 87, to appear.

[16] H. W. Knobloch, *Higher Order Necessary Conditions in Optimal Control Theory*, Lect. Notes Ctrl. & Inform. Sciences **34** (1981) Springer.

[17] C. Marchall, *Chattering arcs and chattering controls*, J. Opt. Theory & Appl., **11** (1973) 441-467.

[18] H. Schättler and H. J. Sussmann, *On the regularity of optimal controls*, J. Appl. Math. & Physics (ZAMP) **38** (1987) 292-301.

[19] G. Stefani, *Local controllability of nonlinear systems: an example*, Systems Control Letters **6** (1985) no. 2, 123-125.

[20] G. Stefani, *Polynomial approximations to control systems and local controllability*, Proc. 24^{th} IEEE Conf. Decision & Control, (1985) Ft. Lauderdale.

[21] *On the local controllability of a scalar-input control system*, Theory and Applications of Nonlinear Control Systems, C. I. Byrnes and A. Lindquist, eds., Elsevier Science Publ. (1986)

[22] H. J. Sussmann, *An extension of a theorem of Nagano on transitive Lie algebras*, Proc. Amer. Math. Soc. **45** (1974) 349-356.

[23] H. J. Sussmann, *Lie brackets, real analyticity and geometric control*, Differential Geometric Control Theory, R. W. Brockett,R. S. Millmann and H.J. Sussmann, eds., Boston (1982) Birkhäuser.

[24] H. J. Sussmann, *Lie brackets and local controllability: a sufficient condition for scalar-input systems*, SIAM J. Cntrl. &Opt. **21** (1983) 686-713.

[25] H. J. Sussmann, *A product expansion for the Chen-series*, Theory and Applications of Nonlinear Control Systems, C. I. Byrnes and A. Lindquist, eds., Elsevier Science Publ. (1986)

[26] H. J. Sussmann, *A general theorem on local controllability*, SIAM J. Cntrl. &Opt. **25** (1987) 158-194.

[27] H. J. Sussmann and V. Jurdjevic, *Controllability of nonlinear systems*, J. Diff. Equns. **12** (1972) 95-116.

[28] D. Stroock and S. Varadhan, *On degenerate elliptic-parabolic operators of second order and their associated diffusions*, Comm. Pure & Appl. Math. **25** (1972) 651-713.

[29] K. Wagner, *Über den Steuerbarkeitsbegriff bei nichtlinearen Kontrollsystemen*, Arch. Math. **47** (1986) 29-40.