A CONTINUATION METHOD FOR NONHOLONOMIC PATH-FINDING PROBLEMS

H.J. SUSSMANN*

March 16, 1993

1 Introduction

In this note we study the mathematical theory of one particular algorithm for nonholonomic point-to-point path finding. The algorithm considered here is a "continuation" or "deformation" method, in which one starts with an admissible trajectory that goes from the desired initial point $\bar{p}$ to some point $\bar{q}_0$, and then tries to construct a one-parameter family of trajectories, whose terminal points $\bar{q}_s$ describe, as s varies, a path that joins $\bar{q}_0$ to the desired terminal point $\bar{q}$.

The use of a continuation algorithm has been proposed by S. Seereeram and J. Wen in [4], who carried out extensive simulations and reported very good results. Independently, a similar proposal was made in [7], where a proof

*Mathematics Department, Rutgers University, New Brunswick, New Jersey 08903. E-mail: sussmann@hamilton.rutgers. edu. (Address from September 1992 to June 1993: Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, Minnesota 55455.)

was outlined, for a very simple case, of the fact that the differential equation describing the path deformation (the "path-lifting equation") is well defined. The theoretical problem of the global existence of solutions of this equation for general nonholonomic systems without drift is still open and appears to be very hard. The purpose of this note is to prove global existence for the class —defined below— of "strongly bracket generating systems." The proof shows how the properties of the path-lifting equation are related to the so-called "abnormal extremals," that have recently attracted considerable interest for different reasons (cf. [1], [2], [3]). (Work is now in progress on more general, not strongly bracket-generating, systems.)

2 Outline of the results

The purpose of the algorithm is to compute, for a given pair $(\bar{p}, \bar{q})$ of points in the state space M of the given nonholonomic system Σ, a trajectory ξ of Σ that goes from $\bar{p}$ to $\bar{q}$. If we let $\mathcal{U}$ denote the space of admissible open-loop controls defined on $[a, b]$, and use $\mathcal{E}_{\bar{p}}$ to denote the *endpoint map* corresponding to the starting point $\bar{p}$ (i.e. the map that assigns, to each control $\eta \in \mathcal{U}$, the terminal point $\xi_{\eta,\bar{p}}(b)$ of the trajectory $\xi_{\eta,\bar{p}} : [a, b] \to M$ of Σ that corresponds to η and is such that $\xi_{\eta,\bar{p}}(a) = \bar{p}$), then our problem is to find η that solves the equation $\mathcal{E}_{\bar{p}}(\eta) = \bar{q}$. This is one particular instance of the general problem of solving a nonlinear equation $F(X) = Y$, where F is a map from a (possibly infinite-dimensional) manifold N_1 to another manifold N_2, Y is given, and X is the unknown. A general strategy for solving equations of this kind is the *continuation method*: (i) start with a pair $(\bar{X}, \bar{Y})$ such that $F(\bar{X}) = \bar{Y}$; (ii) find a path $\pi : [0, 1] \to N_2$ that goes from $\bar{Y}$ to Y, and (iii) try to lift π to a path $\Pi : [0, 1] \to N_1$ such that $\Pi(0) = \bar{X}$ and $F(\Pi(s)) = \pi(s)$ for all s. Then we can take $X = \Pi(1)$.

In our situation, the resulting algorithm proceeds as follows: one starts with a trajectory $t \to \xi_0(t)$ of Σ, defined on an interval $[a, b]$, that goes from $\bar{p}$ to some point $\bar{q}_0 = \xi_0(b)$, and then tries to deform ξ_0 by constructing a one-parameter family $\{\xi_s : 0 \leq s \leq 1\}$ of trajectories $\xi_s : [a, b] \to M$ such that $\xi_s(0) = \bar{p}$, in such a way that the terminal points $\xi_s(b)$ follow a given path $s \to \pi(s)$ in M. If $\pi : [0,1] \to M$ is chosen so that $\pi(0) = \bar{q}_0$ and $\pi(1) = \bar{q}$, then the trajectory ξ_1 will solve our problem.

If the starting trajectory ξ_0 corresponds to an open-loop control η_0, then $\bar{q}_0$ is equal to $\mathcal{E}_{\bar{p}}(\eta_0)$. Since the path $\pi : [0,1] \to M$ is such that $\pi(0) = \bar{q}_0$, finding trajectories $\xi_s : [a,b] \to M$ such that $\xi_s(0) = \bar{p}$ and $\xi_s(b) = \pi(s)$ for $s \in [0,1]$ is tantamount to finding controls $\eta_s \in \mathcal{U}$ such that $\mathcal{E}_{\bar{p}}(\eta_s) = \pi(s)$. If we write $\Pi(s) = \eta_s$, then our problem is precisely that of "lifting" —with respect to $\mathcal{E}_{\bar{p}}$— the path $\pi : [0,1] \to M$ to a path $\Pi : [0,1] \to \mathcal{U}$ i.e. of constructing a path $\Pi : [0,1] \to \mathcal{U}$ such that $\Pi(0) = \eta_0$ and $\mathcal{E}_{\bar{p}}(\Pi(s)) = \pi(s)$. The lifting is found by solving an ordinary differential equation —the "Path-lifting Equation," abbreviated PLE— in $\mathcal{U}$. The PLE expresses the lifting property of the path on an infinitesimal level. Intuitively, if we have found a lifting Π on the interval $[0,s]$, and want to extend this lifting a little further, say to $[0,s+\varepsilon]$, then we should find a direction $v \in \mathcal{U}$ such that $\mathcal{E}_{\bar{p}}(\Pi(s)+hv) \sim \pi(s+h)$ (i.e. $\mathcal{E}_{\bar{p}}(\Pi(s)+hv) \sim \pi(s)+h\dot{\pi}(s))$ as $h \to 0$, and then make sure that $\Pi(s+h) \sim \Pi(s)+hv$. (This will at least yield a first-order approximation to the desired lifting.) Since $\mathcal{E}_{\bar{p}}(\Pi(s)+hv) \sim \pi(s)+hd\mathcal{E}_{\bar{p}}(\Pi(s))\cdot v$ (where $d\mathcal{E}_{\bar{p}}(\Pi(s))$ denotes the differential of $\mathcal{E}_{\bar{p}}$ at $\Pi(s)$), the condition that must be satisfied by the vector $v = \dot{\Pi}(s)$ is:

$$d\mathcal{E}_{\bar{p}}(\Pi(s)) \cdot v = \dot{\pi}(s) \ . \tag{1}$$

One is then led to consider the question whether, once we have constructed the lifting up to a time s, a vector v can be found that satisfies (1). An obvious sufficient condition for this to be possible is that $d\mathcal{E}_{\bar{p}}(\Pi(s))$ be surjective. In that case, we can find one particular solution v of (1) by letting $v = P(\Pi(s))$, where $P(\eta)$ is the *Moore-Penrose pseudoinverse* of $d\mathcal{E}_{\bar{p}}(\eta)$, i.e.

$$P(\eta) = d\mathcal{E}_{\bar{p}}(\eta)^*\Big(d\mathcal{E}_{\bar{p}}(\eta)d\mathcal{E}_{\bar{p}}(\eta)^*\Big)^{-1} \ . \tag{2}$$

With this choice of v, the fact that $v = \dot{\Pi}(s)$ implies that

$$\dot{\Pi}(s) = P(\Pi(s)) \cdot \dot{\pi}(s) \ . \tag{3}$$

Equation (3) is the PLE. Clearly, the PLE is well defined as long as $d\mathcal{E}_{\bar{p}}$ is onto, i.e. that $\mathcal{E}_{\bar{p}}$ is a submersion. It turns out that the controls η where $\mathcal{E}_{\bar{p}}(\eta)$ fails to be a submersion correspond precisely to certain exceptional trajectories (referred to in the literature by various names, such as "abnormal extremals" or "singular trajectories"), and the PLE yields local lifts of paths wherever these abnormal extremals can be avoided. But, even after this particular issue is resolved, the possibility remains that the PLE might have explosions, and therefore fail to have a global solution $\Pi : [0,1] \to \mathcal{U}$. Our goal is to investigate in detail the properties of the PLE, and in particular to show that *the same conditions that imply the nonexistence of abnormal trajectories also suffice to guarantee that the PLE has no explosions.* The natural condition that guarantees nonexistence of abnormal extremals is the so-called *Strong Bracket-Generating Condition* (SBG), and the main result of this paper is Theorem 5, which says precisely that under the SBG condition the PLE has global existence of solutions. The proof of Theorem 5 is based on a somewhat delicate estimate, proved in Theorem 4.

3 Basic definitions

We consider driftless control systems of the form

$$\Sigma : \quad \dot{x} = \sum_{i=1}^{m} u_i f_i(x) \tag{4}$$

where the state variable x takes values in a smooth (i.e. C^∞), connected n-dimensional manifold M, the control $u = (u_1, \ldots, u_m)$ takes values in $\mathbb{R}^m$, and $f_1, \ldots, f_m$ are smooth vector fields on M.

Regarding manifolds, we follow the usual conventions of Differential Geometry, and agree that manifolds are, by definition, Hausdorff and paracompact. This implies that every manifold M admits a Riemannian metric, a fact that will be used below, as will the seemingly stronger —though in fact equivalent— property that M admits a *complete* Riemannian metric. For $x \in M$, we use T_xM, T_x^*M, TM, T^*M to denote, respectively, the tangent and cotangent spaces of M at x, and the tangent and cotangent bundles of M. We use $T^\# M$ to denote T^*M with the zero section removed, i.e.

$$T^\# M = \{(x,z) \in T^*M \;\ni\; z \neq 0\}.$$

For $y \in T_xM$ and $z \in T_x^*M$, we write the duality product $z(y)$ as $\langle z, y\rangle$. Relative to a coordinate chart κ of M, tangent vectors y at points x in the domain of κ will be viewed as column vectors. and cotangent vectors z as row vectors, so we can also write zy as another expression equivalent to $z(y)$ or $\langle z, y\rangle$. (More precisely, if $\kappa = (\kappa_1, \ldots, \kappa_n)$, then to a tangent vector $y \in T_xM$ and a cotangent vector $z \in T_x^*M$ we assign the column vector y^κ whose components are the numbers $y_i^\kappa = d\kappa_i(y)$ and the row vector z^κ whose components are the $z_i^\kappa = z(\frac{\partial}{\partial \kappa_i})$. Then $z(y) = z^\kappa y^\kappa$.) If M is equipped with a Riemannian metric $\langle\cdot,\cdot\rangle$, then we write $\langle y_1, y_2\rangle$ for the inner product of two tangent vectors at a point $x \in M$ and, using the identification between T_xM and T_x^*M induced by the metric, $\langle z_1, z_2\rangle$ is also well defined when both z_1 and z_2 are covectors at x. We also write $\|y\|^2 = \langle y, y\rangle$ and $\|z\|^2 = \langle z, z\rangle$.

We fix once and for all a time interval $[a, b]$, and take $\mathcal{U}$ —the space of open-loop controls— to be $L^2([a, b], \mathbb{R}^m)$. Then $\mathcal{U}$ is a real Hilbert space, with the inner product of two elements $\eta = (\eta_1, ..., \eta_m)$, $\tilde{\eta} = (\tilde{\eta}_1, ..., \tilde{\eta}_m)$ of $\mathcal{U}$ given by

$$\langle \eta, \tilde{\eta}\rangle = \sum_{i=1}^{m} \int_a^b \eta_i(t)\tilde{\eta}_i(t)\, dt \ .$$

If $\eta = (\eta_1, ..., \eta_m) \in \mathcal{U}$, a *trajectory of* Σ *corresponding to* η is an absolutely continuous curve $\xi : [a, b] \to M$ such that

$$\dot{\xi}(t) = \sum_{i=1}^{m} \eta_i(t) f_i(\xi(t)) \tag{5}$$

for a.e. $t \in [a, b]$. A *controlled trajectory* of Σ is a pair (ξ, η), where $\eta \in \mathcal{U}$ and ξ is a trajectory of Σ corresponding to η. (In many cases —e.g. if the vectors $f_1(x), \ldots, f_m(x)$ are linearly independent for every x— a trajectory ξ cannot arise from more than one control η, so it is not necessary to distinguish between "trajectories" and "controlled trajectories.")

Carathéodory's Existence and Uniqueness Theorem guarantees that, for any $\eta \in \mathcal{U}$ and any $p \in M$: (i) there exists, on some subinterval I of $[a, b]$ such that $a \in I$, an absolutely continuous solution $\xi : I \to M$ of the initial value problem consisting of Equation (5) plus the initial condition $\xi(a) = p$; (ii) if $\xi_i : I_i \to M$, $i = 1, 2$ are two such solutions, then $\xi_1 = \xi_2$ on $I_1 \cap I_2$.

We will assume that the solution actually exists on the whole interval $[a, b]$, i.e. that the following *No Explosions Condition* holds:

(NE) For every $\eta \in \mathcal{U}$, $p \in M$, there exists a trajectory $\xi : [a, b] \to M$ of Σ that corresponds to η and is such that $\xi(a) = p$.

The trajectory ξ is obviously unique, and will be denoted by $\xi_{\eta,p}$.

We also assume that the vector fields f_i satisfy the following *Lie Algebra Rank Condition*:

(LARC) Let L_Σ denote the Lie algebra of vector fields on M generated by $f_1, \ldots, f_m$. For $x \in M$, define $L_\Sigma(x) = \{X(x) : X \in L_\Sigma\}$. Then $L_\Sigma(x) = T_x M$ for every $x \in M$.

It is well known that **(LARC)** implies that Σ is *completely controllable* (CC), i.e. that given any two points p, q of M there exists a trajectory $\xi : [a, b] \to M$ of Σ such that $\xi(a) = p$ and $\xi(b) = q$. (The complete controllability condition is often referred to as a *nonholonomy condition*, and a CC system is called a *nonholonomic system*.)

Remark 1 It is not hard to see that, if **(LARC)** holds for the vector fields f_i, $i = 1, \ldots, m$, then it still holds for the new system $\tilde{\Sigma}_\varphi$ in which each f_i is replaced by $\tilde{f}_i = \varphi f_i$, $i = 1, \ldots, m$, where φ is an arbitrary C^∞ real-valued strictly positive function. On the other hand, the trajectories of the two systems are the same, up to a reparametrization of time, so solving the path-finding problem for $\tilde{\Sigma}$ is equivalent to solving it for Σ. It is also easy to see that, if M is an *arbitrary* connected manifold, and Σ is a system of the form (4) that satisfies **(LARC)** but not **(NE)**, then we can find φ such that Σ_φ satisfies **(NE)**. (Indeed, since M is paracompact, it admits a complete Riemannian metric. One can then take φ to be any function such that the resulting $\tilde{f}_i$ are bounded, e.g. $\varphi = (1 + \sum_{i=1}^m ||f_i||^2)^{-1/2}$.) For example, if we start with a system (4) in $\mathbb{R}^n$, and then are interested in path-finding with obstacle avoidance, the "obstacle" being a closed subset C of $\mathbb{R}^n$, then we can take M to be $\mathbb{R}^n \backslash C$, if this set is connected. We can then choose φ to be a smooth function on M such that $\varphi(x)$ goes to zero sufficiently fast as x approaches C, and also as $||x|| \to \infty$. ∎

4 The continuation algorithm

Fix a point $\bar{p} \in M$. As explained above, we define the *endpoint map* $\mathcal{E}_{\bar{p}}$ by letting

$$\mathcal{E}_{\bar{p}}(\eta) = \xi_{\eta,\bar{p}}(b) \ . \tag{6}$$

Since **(NE)** holds, $\mathcal{E}_{\bar{p}}(\eta)$ is defined for every $\eta \in \mathcal{U}$. The complete controllability condition implies that $\mathcal{E}_{\bar{p}}$ is surjective. Our problem now becomes that of computing, for given $\bar{p}$, $\bar{q}$ in M, an $\eta \in \mathcal{U}$ that satisfies the equation

$$\mathcal{E}_{\bar{p}}(\eta) = \bar{q} \ . \tag{7}$$

The continuation procedure outlined in the introduction requires that we start with an η_0 of $\mathcal{U}$ and a path $\pi : [0,1] \to M$ such that $\pi(0) = \bar{q}_0$, $\pi(1) = \bar{q}$, where $\bar{q}_0 = \mathcal{E}_{\bar{p}}(\eta_0)$. We then try to find a path $\Pi : [0,1] \to \mathcal{U}$ such that $\Pi(0) = \eta_0$ and $\mathcal{E}_{\bar{p}}(\Pi(s)) = \pi(s)$. To find this lifting, we use the PLE, which is an ordinary differential equation on $\mathcal{U}$. We now describe the PLE in detail. For this purpose, we assume that *the path π is of class C^1.*

In general, if N_1 and N_2 are two smooth manifolds (with N_1 possibly infinite-dimensional, but N_2 finite-dimensional), and $F : N_1 \to N_2$ is a surjective map of class C^1, then one can lift C^1 paths $\pi : [0,1] \to N_2$ to paths $\Pi : [0,1] \to N_1$ as follows. Suppose we had, for each point w of N_1, a linear map $P(w)$ from $T_{F(w)}N_2$ to $T_w N_1$ which is a right inverse of $dF(w)$, i.e. is such that $dF(w) \circ P(w)$ is the identity map. (Here $dF(w)$ denotes the differential of F at w.) Consider first the case when $N_2 = \mathbb{R}^n$. For a given path $\pi : [0,1] \to N_2$, consider the differential equation

$$\dot{w}(s) = P(w(s)) \cdot \dot{\pi}(s) \ , \tag{8}$$

which has local existence and uniqueness of solutions if the map P is locally Lipschitz with respect to some Riemannian metric on N_1. It is easy to see that, if $\omega : [0,1] \to N_1$ is a solution of (8), and $F(\omega(0)) = \pi(0)$, then $F(\omega(s)) = \pi(s)$ for all s. (To see this, just observe that the derivative of $F(\omega(s))$ with respect to s is $dF(\omega(s)) \cdot P(\omega(s)) \cdot \dot{\pi}(s)$, i.e. $\dot{\pi}(s)$.) Therefore the path Π given by $\Pi(s) = \omega(s)$ is a lifting of π. Now assume that N_2 is a general finite-dimensional manifold, not necessarily equal to $\mathbb{R}^n$. An ordinary differential equation on N_1 is an equation of the form $\dot{w} = X(w,s)$, where $X(w,s)$ is defined for all $w \in N_1$ and all s in some interval. Equation (8) is not of that form, because the quantity $P(w) \cdot \dot{\pi}(s)$ (which is the obvious candidate for the role of $X(w,s)$) only makes sense if $F(w) = \pi(s)$, since the domain of $P(w)$ is $T_{F(w)}N_2$ and $\dot{\pi}(s) \in T_{\pi(s)}N_2$. We can get around this problem by finding an s-dependent vector field $(x,s) \to h(x,s)$ on N_2 (i.e. a map $(x,s) \to h(x,s)$ such that, for each fixed s, $h(\cdot,s)$ is a vector field on N_2) such that $h(\pi(s),s) = \dot{\pi}(s)$. We can then consider the equation

$$\dot{w}(s) = P(w(s)) \cdot h(F(w(s)),s) \ , \tag{9}$$

corresponding to the choice $X(w,s) = P(w) \cdot h(F(w),s)$, which is now perfectly meaningful, since $h(F(w),s)$ belongs to $T_{F(w)}N_2$, which is precisely the domain of $P(w)$. If P is locally Lipschitz as before, and in addition $h(x,s)$ is continuous with respect to (x,s), and locally Lipschitz with respect to x, locally uniformly in s, then (9) also has local existence and uniqueness of solutions. Once again, one verifies that a solution $\omega : [0,1] \to N_1$ of (9) for which $F(\omega(0)) = \pi(0)$ necessarily satisfies $F(\omega(s)) = \pi(s)$ for all s. (Indeed, the curve $s \to \xi(s) = F(\omega(s))$ is a solution of $\dot{x} = h(x,s)$, and satisfies $\xi(0) = \pi(0)$. Our hypotheses on h imply uniqueness of solutions for $\dot{x} = h(x,s)$, so $\xi(s) \equiv \pi(s)$.)

Equation (9) is the *Path-lifting Equation* (PLE) for F. The PLE can be written provided that a map P with the above properties exists, and in that case it depends on the choice of P. It might seem that, for the PLE to make sense, one also needs to require the existence of an s-dependent vector field h with the above properties. And, in that case, one may think that whether or not a particular function $\omega(\cdot)$ is a solution might depend on the choice of h. An easy construction, using partitions of unity, shows that an h with the required properties always exists, and that the set of solutions $\omega(\cdot)$ of (9) that satisfy $F(\omega(0)) = \pi(0)$ is in fact independent of the choice of h.

We now turn to the problem of how to choose P. An obvious necessary condition for the existence of P is:

(SUB) The map F is a submersion.

(This means that $dF(w)$ is surjective for every $w \in N_1$.) If **(SUB)** holds, and N_1 is equipped with a Riemannian metric, in such a way that every tangent space $T_w N_1$ is a Hilbert space, then we can define $P(w) : T_{F(w)} N_2 \to T_w N_1$ by letting $P(w) \cdot y$ be, for each $y \in T_{F(w)} N_2$, the vector $v \in T_w N_1$ of smallest length such that $dF(w) \cdot v = y$.

The linear map $P(w) : T_{F(w)} N_2 \to T_w N_1$ defined above is the *Moore-Penrose pseudo-inverse* (MPPI) (or generalized pseudo-inverse) of $dF(w)$. We now recall some general properties of the MPPI.

Let H_1 and H_2 be two Hilbert spaces. Assume that $\dim H_2 < \infty$. Let $L : H_1 \to H_2$ be a continuous linear mapping and let $L^* : H_2 \to H_1$ be its adjoint, so that $\langle Lx, y \rangle = \langle x, L^* y \rangle$ for all $x \in H_1$, $y \in H_2$. Let $W = LL^*$, so W is a linear map from $H_2 \to H_2$, and $W \geq 0$. It is clear that L is onto if and only if $W > 0$. In that case, if we define $L^\# : H_2 \to H_1$ by $L^\# = L^* W^{-1}$, we have the identity $LL^\# x = x$ for all $x \in H_1$, so $L^\#$ is a right inverse of L. By definition, the operator $L^\#$ is the MPPI of L. If $x \in H_2$, $v = L^\# x$, and $\tilde{v} \in H_1$ is such that $L\tilde{v} = x$, then

$$\langle \tilde{v} - v, v \rangle = \langle \tilde{v} - v, L^* W x \rangle = \langle L(\tilde{v} - v), W x \rangle = 0 \ ,$$

so $||\tilde{v}|| \geq ||v||$, with equality holding if and only if $\tilde{v} = v$. So $L^\# x$ is the vector v of smallest length such that $Lv = x$. (This shows, in particular, that the operator $L^\#$ does not depend on the inner product of H_2.) Clearly,

$$||L^\# x||^2 = \langle L^* W^{-1} x, L^* W^{-1} x \rangle = \langle W^{-1} x, LL^* W^{-1} x \rangle = \langle W^{-1} x, x \rangle$$

for every $x \in H_2$, so $||L^\#||^2 = ||W^{-1}||$. On the other hand, the norm of the positive-definite operator W^{-1} is equal to its largest eigenvalue, which is the inverse of $\sigma_{\min}(W)$, the smallest eigenvalue of W. Since $\sigma_{\min}(W)$ is the infimum of the numbers $\langle z, Wz\rangle$, ranging over all $z \in H_2$ such that $||z|| = 1$, we conclude that

$$\|L^{\#}\|^2 = \left(\inf_{\substack{z\in H_2\\ \|z\|=1}} \langle z, LL^* z\rangle\right)^{-1} . \tag{10}$$

Returning now to the abstract path-lifting situation considered above, of two manifolds N_1 and N_2, and a surjective submersion $F : N_1 \to N_2$, let us specialize even further by assuming that N_1 is actually an open subset of a Hilbert space $\mathcal{H}$ and $F \in C^2$. In that case, we can define $P(w)$, for each $w \in N_1$, to be the MPPI of the map $dF(w) : \mathcal{H} \to T_{F(w)}N_2$, and it is clear that $P \in C^1$. The PLE equation will therefore have local solutions for any C^1 path $\pi : [0,1] \to N_2$ and any initial condition $w(0)$ such that $F(w(0)) = \pi(0)$. If, moreover, P satisfies a *linear growth estimate* (LGE):

(LGE) For every compact subset K of N_2 there exists a $C > 0$ such that

$$||P(w)|| \leq C(1 + ||w||) \tag{11}$$

for all $w \in N_1$ for which $F(w) \in K$.

then it follows easily from Gronwall's inequality that the PLE admits global solutions for any C^1 path $\pi : [0,1] \to N_2$, provided that F satisfies the following property

(CL) For every compact subset K of N_2, the set $F^{-1}(K)$ is closed in $\mathcal{H}$.

(Indeed, let $\omega : I \to N_1$ be a solution of the PLE for π on some subinterval I of $[0,1]$ that contains 0, and assume that $F(\omega(0)) = \pi(0)$, and that ω is maximal, in the sense that ω cannot be extended to a larger subinterval of $[0,1]$. Condition **(LGE)**, together with Gronwall's inequality, imply that ω is bounded, and then **(LGE)** implies that ω is Lipschitz. (Recall that, as long as ω exists, we have $F(\omega(s)) = \pi(s)$, so that $\dot{\omega}(s) = P(\omega(s)) \cdot h(\omega(s), s)$, and then $||\dot{\omega}(s)||$ is bounded by a constant times $1+||\omega(s)||$, since h is continuous, and hence bounded, on the compact set $\pi([0,1]) \times [0,1]$.) If $I \neq [0,1]$, then I must be open to the right, because the PLE has local solutions. So $I = [0, \alpha)$

for some $\alpha \in (0,1]$. Since ω is Lipschitz, the limit $\omega(\alpha) = \lim_{s\to\alpha-}\omega(s)$ exists. By **(CL)**, $\omega(\alpha) \in N_1$, since $F(\omega(s))$ belongs to the compact set $\pi([0,1])$. So ω has an extension to the larger interval $[0,\alpha]$, contradicting the maximality of I.)

Remark 2 The LGE involves the operator norm of the linear map $P(w) : T_{F(w)}N_2 \to \mathcal{H}$. For this norm to be well defined, we need in principle to specify norms in the domain and the range of $P(w)$. The norm on $\mathcal{H}$ poses no problem, since $\mathcal{H}$ is a Hilbert space. To have a norm on $T_{F(w)}N_2$ we can equip N_2 with a Riemannian metric. It is easy to see that, if K is a compact subset of N_2, then any two Riemannian metrics $\langle\cdot,\cdot\rangle_1$, $\langle\cdot,\cdot\rangle_2$ on N_2 are necessarily equivalent on K, in the sense that there exists constants $c > 0$, $C > 0$ such that $c\langle v,v\rangle_1 \le \langle v,v\rangle_2 \le C\langle v,v\rangle_1$ for all $v \in T_xN_2$, $x \in K$. This implies, in particular, that *although the norm $||P(w)||$ depends on the choice of a metric on N_2, the validity of Condition* **(LGE)** *is in fact independent of the metric.* A similar remark applies to the equivalent conditions **(LGE')** and **(LGE")** introduced below. ▮

It will be convenient to use Formula (10) to rewrite **(LGE)**. Since $P(w) = dF(w)^{\#}$, the bound $||P(w)|| \le C(1+||w||)$ is equivalent to the requirement that

$$\inf_{\substack{z\in H_2\\ \|z\|=1}} \langle z, dF(w)dF(w)^* z\rangle \ge \frac{c}{(1+||w||)^2} \tag{12}$$

(where $c = C^{-2}$), so **(LGE)** is equivalent to

(LGE') For every compact subset K of N_2 there exists a $c > 0$ such that

$$||dF(w)(z)||^2 \ge \frac{c||z||^2}{1+||w||^2} \tag{13}$$

for all $w \in N_1$ for which $F(w) \in K$ and all $z \in T_{F(w)}N_2$.

Summarizing, we have shown:

Theorem 1 *Let N_1 be an open subset of a real Hilbert space $\mathcal{H}$. Let N_2 be a finite-dimensional manifold, and let $F : N_2 \to N_1$ be a surjective submersion of class C^2 such that Condition* **(CL)** *holds. For $w \in N_1$, let $P(w)$ be the Moore-Penrose pseudoinverse of $dF(w)$. Assume that F satisfies an estimate* **(LGE')**. *Then, for each C^1 path $\pi : [0,1] \to N_2$ and each initial condition $\omega(0)$ such that $F(\omega(0)) = \pi(0)$, the solution of the corresponding path-lifting equation exists on the whole interval $[0,1]$, and yields a lifting of π.* ∎

Let us now return to our original path finding problem for a nonholonomic system Σ. We start by taking $N_1^0 = \mathcal{U}$, $N_2^0 = M$, and try to apply the preceding considerations to the map $F = \mathcal{E}_{\bar{p}}$. The complete controllability condition implies that $\mathcal{E}_{\bar{p}}$ is surjective. Moreover, $\mathcal{E}_{\bar{p}}$ is of class C^∞. To decide whether $\mathcal{E}_{\bar{p}}$ is a submersion we have to compute its differential.

It turns out that, for $\eta \in \mathcal{U}$, the differential $d\mathcal{E}_{\bar{p}}(\eta)$ is given in terms of the solutions of the *variational equation* along the trajectory $\xi_{\eta,\bar{p}}$. To make this precise, let $v = (v_1, \ldots, v_m) \in \mathcal{U}$. Let $[a,b] \ni t \to y_{\eta,\bar{p},v}(t) \in T_{\xi_{\eta,\bar{p}}(t)}M$ be the solution of the *variational equation*

$$\dot{y}(t) = (\eta(t) \cdot Df)(\xi_{\eta,\bar{p}}(t)) \cdot y(t) + (v(t) \cdot f)(\xi_u(t)) \tag{14}$$

with initial condition $y(0) = 0$. (As usual, a *solution* of (14) is an absolutely continuous function y such that (14) holds for almost every t. The expressions $\eta(t) \cdot Df$, $v(t) \cdot f$ are abbreviations for $\sum_{i=1}^m \eta_i(t) Df_i$, $\sum_{i=1}^m v_i(t) f_i$.) Here, if X is any smooth vector field on M, DX denotes —relative to a given coordinate chart κ— the Jacobian matrix of the column-vector-valued map X that represents X. Equation (14) makes sense in coordinates, but it is easy to see that, if $y(t)$ is thought of as a *tangent vector at* $\xi_{\eta,\bar{p}}(t)$, then the property that $y(\cdot)$ is a solution of (14) is independent of the choice of κ.

Having defined $y_{\eta,\bar{p},v}$, the map $d\mathcal{E}_{\bar{p}}(\eta) : \mathcal{U} \to T_qM$ (where $q = \mathcal{E}_{\bar{p}}(\eta)$) is simply given by

$$d\mathcal{E}_{\bar{p}}(\eta)(v) = y_{\eta,\bar{p},v}(1)\ . \tag{15}$$

Notice that (14) is the *linearization* of our system at the reference trajectory $\xi_{\eta,\bar{p}}$. Hence

The differential $d\mathcal{E}_{\bar{p}}(\eta)$ is surjective if and only if the linearized (time-varying) system (14) is controllable.

Naturally, $d\mathcal{E}_{\bar{p}}(\eta)$ is surjective if and only if its adjoint $d\mathcal{E}_{\bar{p}}(\eta)^*$ is one-to-one. It will be convenient to determine $d\mathcal{E}_{\bar{p}}(\eta)^*$ explicitly. As expected, this requires that we consider the *adjoint variational equation along* $\xi_{\eta,\bar{p}}$.

We use $V(M)$ to denote the set of smooth vector fields on M. For $X \in V(M)$, we define the *variational covector field of* X to be the element X^* of $V(T^*M)$ whose expression in coordinates is given by

$$X^*(x,z) = (X(x), -z \cdot DX(x)) \ , \tag{16}$$

where, as before, DX is the Jacobian matrix of the column-vector-valued map X. The vector field X^* is also known as the *Hamiltonian lift* of X. Notice that the integral curves of X^* are the solutions $t \to (\xi(t), \zeta(t))$ of the system of differential equations

$$\dot{x} = \frac{\partial h}{\partial z} \quad , \quad \dot{z} = -\frac{\partial h}{\partial x} \ , \tag{17}$$

that corresponds to the Hamiltonian function $h : T^*M \to \mathbb{R}$ given by $h(x,z) = \langle z, X(x) \rangle$ for $z \in T^*_x M$. It is well known that, if h is any smooth function on T^*M, then the vector field whose trajectories are given in coordinates by (17) is in fact intrinsically defined on T^*M.

Similarly, Σ^* will denote the *Hamiltonian lift of* Σ, i.e. the system

$$\Sigma^* : \quad \dot{x}^* = \sum_{i=1}^{m} u_i f_i^*(x^*) \ , \tag{18}$$

where the variable $x^* = (x,z)$ evolves in T^*M. A trajectory $t \to \xi^*(t)$ of Σ^* corresponding to $\eta = (\eta_1, \cdots, \eta_m) \in \mathcal{U}$ is a pair $\xi^* = (\xi, \zeta)$ where ξ is a trajectory of Σ corresponding to η and the map $[a,b] \ni t \to \zeta(t) \in T^*_{\xi(t)}M$ satisfies, in coordinates, the equation

$$\dot{\zeta} = -\zeta \left(\sum_{i=1}^{m} \eta_i(t)(Df_i)(\xi(t)) \right) \qquad \text{for a.e. } t \in [a,b] \ . \tag{19}$$

(Equation (19) is the well known *adjoint equation* along ξ. A map ζ that picks, for each t, a covector $\zeta(t) \in T^*_{\xi(t)}M$, is called a *covector along* ξ. If, in addition, ζ satisfies (19), then ζ is called an *adjoint vector* along ξ.) Since (19) is linear in ζ, the solution of (19) is well-defined on $[a,b]$ for any terminal condition $\zeta(b) = \bar{z} \in T^*_{\xi(b)}M$. Then $\xi^* = (\xi, \zeta)$ is a trajectory of $T^*\Sigma$ corresponding to $\eta = (\eta_1, \cdots, \eta_m)$ if

$$\dot{\xi}^* = \sum_{i=1}^{m} u_i f_i^*(\xi) \quad \text{for a.e. } t \in [a,b]. \tag{20}$$

Equation (20) is the well known system

$$\dot{\xi}(t) = \frac{\partial H}{\partial z}(\xi(t), \zeta(t), \eta(t)) \ , \ \dot{\zeta}(t) = -\frac{\partial H}{\partial x}(\xi(t), \zeta(t), \eta(t)) \, , \tag{21}$$

where H is the *Hamiltonian* of Σ, given by

$$H(x, z, u) = \sum_{i=1}^{m} u_i \langle z, f_i(x) \rangle \ .$$

With these definitions, we can finally write a formula for the map $d\mathcal{E}_{\bar{p}}(\eta)^*$. As before, let $q = \mathcal{E}_{\bar{p}}(\eta)$, and pick a covector $z \in T^*_q M$. Let $\zeta_{\eta,\bar{p},z}$ be the solution of (19) that satisfies the *terminal condition* $\zeta(b) = z$. Define the *switching functions* $\varphi_{\eta,\bar{p},z,i}$, for $i = 1, \ldots, m$, by

$$\varphi_{\eta,\bar{p},z,i}(t) = \langle \zeta(t), f_i(\xi_{\eta,\bar{p}}(t)) \rangle \ . \tag{22}$$

We can then associate, to each $\eta, \bar{p}, z$, the m-tuple of functions

$$\Phi_{\eta,\bar{p},z} = (\varphi_{\eta,\bar{p},z,1}, \ldots, \varphi_{\eta,\bar{p},z,m}) \ . \tag{23}$$

We then have

The adjoint $d\mathcal{E}_{\bar{p}}(\eta)^* : T^*_q M \to \mathcal{U}$ of $d\mathcal{E}_{\bar{p}}(\eta)$ is given by

$$d\mathcal{E}_{\bar{p}}(\eta)^*(z) = \Phi_{\eta,\bar{p},z} \quad \text{for } \ z \in T^*_q M \ , \ q = \mathcal{E}_{\bar{p}}(\eta) \ . \tag{24}$$

In particular, $d\mathcal{E}_{\bar{p}}(\eta)^*$ is one-to-one if and only if there does not exist any nonzero covector $z \in T^*_q M$ for which the switching functions $\varphi_{\eta,\bar{p},z,i}$ vanish identically for $i = 1, \ldots, m$. Equivalently, $d\mathcal{E}_{\bar{p}}(\eta)^*$ fails to be one-to-one if and only if there exists a nontrivial solution of the corresponding adjoint equation, for which the switching functions $\varphi_{\eta,\bar{p},z,i}$ vanish identically for $i = 1, \ldots, m$.

A controlled trajectory (ξ, η) for which there exists a nontrivial solution of the adjoint equation whose associated switching functions vanish identically is called an *abnormal extremal of* Σ.

We have therefore shown:

Theorem 2 *The map $d\mathcal{E}_{\bar{p}}(\eta)^*$ is one-to-one if and only if $(\xi_{\eta,\bar{p}}, \eta)$ is not an abnormal extremal.* ∎

Now consider an open subset Ω of M with the property that

(NAE) there exist no abnormal extremals going from $\bar{p}$ to a point in Ω.

If we now take $N_1 = \mathcal{E}_{\bar{p}}^{-1}(\Omega)$, $N_2 = \Omega$, then the restriction to N_1 of the map $\mathcal{E}_{\bar{p}}$ is a C^2 surjective submersion onto N_2. The path-lifting equation is therefore well defined for paths in Ω. Global solutions will exist if the linear growth estimate **(LGE')** is satisfied. In our case, we see that **(LGE')** will hold if we can find, for every compact subset K of Ω, a constant $c > 0$ such that

$$||d\mathcal{E}_{\bar{p}}(\eta)^*(z)||^2 \geq \frac{c||z||^2}{1 + ||\eta||^2} \tag{25}$$

for all $\eta \in \mathcal{U}$ such that $\mathcal{E}_{\bar{p}}(\eta) \in \Omega$.

The map

$$G_{\eta,\bar{p}} \stackrel{\text{def}}{=} d\mathcal{E}_{\bar{p}}(\eta) d\mathcal{E}_{\bar{p}}(\eta)^* \tag{26}$$

is the *controllability gramian* of the linearized system (14). A simple calculation yields:

$$\langle z, G_{\eta,\bar{p}}\, z\rangle = \sum_{i=1}^{m} \int_a^b \varphi_{\eta,\bar{p},z,i}(t)^2\, dt\ . \tag{27}$$

Moreover, it is clear that $||d\mathcal{E}_{\bar{p}}(\eta)^*(z)||^2 = \langle z, G_{\eta,\bar{p}}\, z\rangle$. Hence the linear growth estimate is equivalent to the following:

(LGE") For every compact subset K of Ω there exists a $c > 0$ such that

$$\sum_{i=1}^{m} \int_a^b \varphi_{\eta,\bar{p},z,i}(t)^2 \, dt \geq \frac{c||z||^2}{1 + ||\eta||^2} \tag{28}$$

for all $\eta \in \mathcal{U}$ for which $\mathcal{E}_{\bar{p}}(\eta) \in K$, and all $z \in T_{\mathcal{E}_{\bar{p}}(\eta)}M$.

5 SBG systems

For a system Σ of the form (4) we define the functions $\varphi_{\Sigma,i} : T^*M \to \mathbb{R}$, $\psi_{\Sigma,ij} : T^*M \to \mathbb{R}$ by letting

$$\varphi_{\Sigma,i}(x,z) = \langle z, f_i(x) \rangle \tag{29}$$

$$\psi_{\Sigma,ij}(x,z) = \langle z, [f_i, f_j](x) \rangle \ . \tag{30}$$

We then define the vector-valued function $\Phi_\Sigma : T^*M \to \mathbb{R}^m$ and the matrix-valued function $\Psi_\Sigma : T^*M \to \mathbb{R}^{m \times m}$ by

$$\Phi_\Sigma = (\varphi_{\Sigma,1}, \ldots, \varphi_{\Sigma,m}) \ , \tag{31}$$

$$\Psi_\Sigma = (\psi_{\Sigma,ij})_{1 \leq i,j \leq m} \ . \tag{32}$$

It is clear that, for every (x,z), the matrix $\Psi_\Sigma(x,z)$ is skew-symmetric.

A system Σ of the form (4) is said to be *Strong Bracket Generating* (SBG) if the following condition is true:

(SBG) For every $(x,z) \in T^\# M$, if $\Phi_\Sigma(x,z) = 0$, then the matrix $\Psi_\Sigma(x,z)$ is nonsingular.

The class of SBG systems is extremely restrictive. For example, a system for which m is odd can never be SBG unless $f_1(x), \ldots, f_m(x)$ span T_xM for all x, since an $m \times m$ skew-symmetric matrix cannot be nonsingular if m is odd. One interesting example of SBG systems is the following:

EXAMPLE. *Let f_1, f_2 be two smooth vector fields on a 3-dimensional manifold M, such that $f_1(x)$, $f_2(x)$ and $[f_1, f_2](x)$ span T_xM for each $x \in M$. Then the system $\dot{x} = u_1 f_1(x) + u_2 f_2(x)$ is SBG.*

To see this, write $f_3 = [f_1, f_2]$, and observe that in this case

$$\Psi_\Sigma(x,z) = \begin{pmatrix} 0 & -\varphi_3(x,z) \\ \varphi_3(x,z) & 0 \end{pmatrix} , \tag{33}$$

where $\varphi_i(x,z) = \langle z, f_i(x)\rangle$ for $i = 1,2,3$. If $\varphi_1(x,z) = \varphi_2(x,z) = 0$, and $(x,z) \in T^\# M$, then $z \neq 0$, so $\varphi_3(x,z) \neq 0$, since $f_1(x)$, $f_2(x)$ and $f_3(x)$ form a basis of T_xM. Then the matrix $\Psi_\Sigma(x,z)$ is obviously nonsingular. ∎

The SBG condition was introduced by R. Strichartz in [5] and [6] as the most natural assumption that excludes the possibility of abnormal extremals. Indeed, it is easy to show that

Theorem 3 *If a system of the form (4) is SBG, then the only abnormal extremals are the trajectories that correspond to a control $\eta(t) \equiv 0$.*

PROOF. If (ξ, η) is an abnormal extremal, then there is a nonzero solution $t \to \zeta(t)$ of the adjoint equation such that $\varphi_{\Sigma,i}(\xi(t), \zeta(t)) \equiv 0$. A simple calculation shows that the derivative of $\varphi_{\Sigma,i}(\xi(t), \zeta(t))$ is $\sum_j \psi_{\Sigma,ij}(\xi(t), \zeta(t))\eta_j(t)$. Therefore $\Psi_\Sigma(\xi(t), \zeta(t))\eta(t) = 0$. Since $\Psi_\Sigma(\xi(t), \zeta(t))$ is nonsingular, we conclude that $\eta(t) \equiv 0$. ∎

It follows from Theorem 3 that, for an SBG system, if we fix a starting point $\bar{p}$, then the set $\Omega = M \backslash \{\bar{p}\}$ satisfies Condition **(NAE)**. Therefore the PLE is well defined for every path π that does not go through $\bar{p}$. In particular, to construct a trajectory of Σ from $\bar{p}$ to another point $\bar{q} \neq \bar{p}$, it suffices to find one control η_0 that steers $\bar{p}$ to some point $\bar{q}_0$ other than $\bar{p}$, and then find a path π in M that goes from $\bar{q}_0$ to $\bar{q}$ *and does not go through $\bar{p}$*. (Such a path always exists if $\dim M > 1$.) The PLE then enables us to lift π, at least locally. The only remaining obstruction to a global lift is the possibility that solutions of the PLE might explode. It turns out that this does not happen, as we now show.

6 Global liftings

Theorem 4 *Let Σ be a system of the form (4), and assume that Σ is SBG. Let $\bar{p} \in M$, and write $\Omega = M \backslash \{\bar{p}\}$. Let $a < b$, and let $\mathcal{U} = L^2([a,b], \mathbb{R}^m)$. Then for every compact subset K of Ω there exists a constant $c > 0$ such that, if $\eta \in \mathcal{U}$ is such that $\mathcal{E}_{\bar{p}}(\eta) \in K$, and $\hat{z}$ is any covector at $\mathcal{E}_{\bar{p}}(\eta)$ such that $||\hat{z}|| = 1$, then the inequality*

$$\Big(\int_a^b ||\eta(t)||^2 \, dt\Big)\Big(\int_a^b (\varphi_{\eta,\bar{p},\hat{z},1}(t)^2 + \ldots + \varphi_{\eta,\bar{p},\hat{z},m}(t)^2) \, dt\Big) \geq c \,, \tag{34}$$

holds, where $\varphi_{\eta,\bar{p},\hat{z},i}(t) \stackrel{\mathrm{def}}{=} \langle \zeta(t) \,,\, f_i(\xi_{\eta,\bar{p}}(t)) \rangle$, and ζ is the solution of the adjoint equation along $(\xi_{\eta,\bar{p}}, \eta)$ such that $\zeta(b) = \hat{z}$.

PROOF. Let K be a compact subset of Ω. As explained in Remark 2, we can assume that M is equipped with a Riemannian metric. From now on, $d(x, \tilde{x})$ denotes, for $x, \tilde{x} \in M$, the distance with respect to this metric. Also, if v is a tangent vector or a covector, then $||v||$ is its norm relative to the metric.

For $\alpha \geq 0$, let

$$K_\alpha = \{x \in M \,:\, d(x, K) \leq \alpha\} \tag{35}$$

Then, if $\alpha > 0$ is sufficiently small, the set K_α is compact and contained in Ω. From now on, we fix an $\alpha > 0$ with these properties. We let

$$\begin{aligned} K^* &= \{(x,z) \in T^*M \,:\, x \in K \,,\, ||z|| = 1\} \,, \\ K^*_\alpha &= \{(x,z) \in T^*M \,:\, x \in K_\alpha \,,\, \frac{1}{2} \leq ||z|| \leq 2\} \,. \end{aligned}$$

Notice that the closure of $T^*M \backslash K^*_\alpha$ is disjoint from K^*. So we can choose a smooth function $\theta : T^*M \to \mathbb{R}$ such that $\theta \equiv 0$ on K^* and $\theta > 1$ on $T^*M \backslash K^*_\alpha$. We let $\nu_i(x,z)$ denote the directional derivative $f_i^*\theta(x,z)$ of θ at (x,z) in the direction of the vector field f_i^* (so that $\nu_i : T^*M \to \mathbb{R}$ is the function that in the control literature is sometimes known as the "Lie derivative" of θ in the direction of f_i^*, and denoted by $L_{f_i^*}\theta$). We let $C_1 > 0$ be such that $\nu_1^2 + \ldots + \nu_m^2 \leq C_1^2$ throughout K^*_α. We write φ_i, ψ_{ij}, Φ, Ψ for $\varphi_{\Sigma,i}$, $\psi_{\Sigma,ij}$,

Φ_Σ, Ψ_Σ. We let $\Delta : T^*M \to \mathbb{R}$ be the square of the determinant of Ψ. The SBG hypothesis then implies that the function $||\Phi||^2 + \Delta$ never vanishes on $T^\# M$. Let 2β be the infimum of this function on the compact subset K_α^* of $T^\# M$. Then $\beta > 0$ and, if $(x, z) \in K_\alpha^*$ is such that $||\Phi(x, z)|| \leq \sqrt{\beta}$, then we can conclude that $\Delta(x, z) \geq \beta$.

Now let $\eta \in \mathcal{U}$ be such that $q \stackrel{\text{def}}{=} \mathcal{E}_{\bar{p}}(\eta) \in K$. Use ξ to denote the trajectory $\xi_{\eta,\bar{p}}$. Fix a $\hat{z} \in T_q^*M$ such that $||\hat{z}|| = 1$. Let ζ be the solution of the adjoint equation along (ξ, η) such that $\zeta(b) = \hat{z}$. Let $\xi^*(t) = (\xi(t), \zeta(t))$, so that $\xi^* : [a, b] \to T^\# M$ is a trajectory of the Hamiltonian lift Σ^* of the system Σ. Then $\xi^*(b) \in K^*$, and $\xi^*(a) \notin K_\alpha^*$, since $\xi(a) = \bar{p} \notin K$. So $\theta(\xi^*(b)) = 0$, and $\theta(\xi^*(a)) > 1$. Therefore there exists a τ such that $a < \tau < b$, with the property that $\theta(\xi^*(\tau)) = 1$ and $\theta(\xi^*(t)) < 1$ for $\tau < t \leq b$. It follows in particular that $\xi^*(t) \in K_\alpha^*$ for $\tau \leq t \leq b$.

From now on, we use the convention that, whenever h is a smooth scalar- or vector-valued function on T^*M, then $\underline{h}$ denotes h evaluated along the trajectory ξ^*, i.e. $\underline{h}(t) = h(\xi^*(t))$. We have

$$\dot{\xi}^*(t) = \sum_{i=1}^{m} \eta_i(t) f_i^*(\xi^*(t)) \quad \text{for a.e. } t \in [a, b] \, , \tag{36}$$

and therefore

$$\dot{\underline{\theta}}(t) = \sum_{i=1}^{m} \eta_i(t) \underline{\nu}_i(t) \quad \text{for a.e. } t \in [a, b] \, , \tag{37}$$

where $\underline{\nu}_i(t)$ stands for $\nu_i(\xi^*(t))$, as explained before. Since $\underline{\theta}$ is absolutely continuous, $\underline{\theta}(b) = 0$, and $\underline{\theta}(\tau) = 1$, we easily conclude from (37) that

$$\Big| \int_\tau^b \langle \eta(t), \underline{\nu}(t) \rangle \, dt \Big| \geq 1 \ , \tag{38}$$

where $\langle \eta(t), \underline{\nu}(t) \rangle \stackrel{\text{def}}{=} \sum_{i=1}^m \eta_i(t) \underline{\nu}_i(t)$. If we had an estimate of the form

$$||\underline{\nu}|| \leq \text{constant.} ||\underline{\Phi}|| \ , \tag{39}$$

then (38), together with the Cauchy-Schwarz inequality, would yield the estimate $\left(\int_\tau^b ||\eta(t)||^2 \, dt \right) \left(\int_\tau^b ||\underline{\Phi}(t)||^2 \, dt \right) \geq$ constant, which would prove our conclusion. On the other hand, it is clear that an estimate (39) holds as long as $||\underline{\Phi}||$ is "large." So the real difficulty arises when $||\underline{\Phi}||$ is "small."

To take care of this difficulty, we introduce a "cutoff function" $\sigma_0 : \mathbb{R} \to \mathbb{R}$ by letting $\sigma_0(s) = 0$ for $s \leq 0$, $\sigma_0(s) = \frac{s}{\rho}$ for $0 \leq s \leq \rho$, and $\sigma_0(s) = 1$ for $s \geq \rho$. Here $\rho > 0$ is a constant to be chosen later, and required to satisfy $\rho^2 \leq \beta$. Define

$$\sigma(x,z) = \sigma_0(|\varphi_1(x,z)| + \ldots + |\varphi_m(x,z)|). \tag{40}$$

Notice that the function σ is Lipschitz, since (i) σ_0 itself is Lipschitz, and (ii) the φ_i are smooth, so that the functions $|\varphi_i|$ are Lipschitz. We now write

$$\langle \eta(t), \underline{\nu}(t) \rangle = \underline{\sigma}(t) \langle \eta(t), \underline{\nu}(t) \rangle + (1 - \underline{\sigma}(t)) \langle \eta(t), \underline{\nu}(t) \rangle \ . \tag{41}$$

and decompose the integral of (38) as a sum

$$\int_\tau^b \langle \eta(t), \underline{\nu}(t) \rangle \, dt = I_1 + I_2 \ , \tag{42}$$

where

$$I_1 \ = \ \int_\tau^b \underline{\sigma}(t) \langle \eta(t), \underline{\nu}(t) \rangle \, dt \ \ , \tag{43}$$

$$I_2 \ = \ \int_\tau^b \Big(1 - \underline{\sigma}(t)\Big) \langle \eta(t), \underline{\nu}(t) \rangle \, dt \ \ . \tag{44}$$

It is clear that $\sigma_0(s) \leq \frac{s}{\rho}$ for all $s \in [0, \infty)$. Therefore σ is pointwise bounded by $\frac{\sqrt{m}}{\rho} ||\Phi||$. So I_1 satisfies

$$|I_1| \leq \frac{C_1 \sqrt{m}}{\rho} \Big(\int_\tau^b ||\eta(t)||^2 \, dt \Big) \Big(\int_\tau^b ||\underline{\Phi}(t)||^2 \, dt \Big) \ . \tag{45}$$

To estimate I_2, we first recall the fact (already used in the proof of Theorem 3), that the derivative of $\underline{\varphi}_i(t)$ is $\sum_j \underline{\psi}_{ij}(t) \eta_j(t)$. In vector notation, this says that $\dot{\underline{\Phi}} = \eta \underline{\Psi}$, where we are writing both Φ and η as row vectors. If we use $\Psi^\dagger$ to denote the matrix of complementary minors of Ψ, then we have $\underline{\Delta} \eta = \eta \underline{\Psi} \, \underline{\Psi}^\dagger$. We also notice that the integrand in the formula defining I_2 vanishes unless $|\underline{\varphi}_1| + \ldots + |\underline{\varphi}_m| \leq \rho$. Since $\xi^*(t) \in K^*_\alpha$ for $\tau \leq t \leq b$, and

$\rho^2 \leq \beta$, the bound $|\underline{\varphi}_1| + \ldots + |\underline{\varphi}_m| \leq \rho$ implies $||\underline{\Phi}||^2 \leq \rho^2$, and therefore $\underline{\Delta} \geq \beta$. If, in the expression for I_2, we replace η by $\frac{1}{\underline{\Delta}}\eta\underline{\Psi}\,\underline{\Psi}^\dagger$, and use the fact that $\dot{\underline{\Phi}} = \eta\underline{\Psi}$, we end up with the equality:

$$I_2 = \int_\tau^b \left(\frac{1-\underline{\sigma}(t)}{\underline{\Delta}(t)}\right) \langle \dot{\underline{\Phi}}(t)\underline{\Psi}^\dagger(t), \underline{\nu}(t)\rangle \, dt \ , \tag{46}$$

and the bound $\underline{\Delta} \geq \beta$ holds whenever the integrand is not zero.

We are now ready to integrate by parts. Define a function $A(\mathbf{v}, t)$, where $\mathbf{v}$ is a m-dimensional row-vector variable, by

$$A(\mathbf{v}, t) = \left(\frac{1-\underline{\sigma}(t)}{\underline{\Delta}(t)}\right) \langle \mathbf{v}\underline{\Psi}^\dagger(t), \underline{\nu}(t)\rangle \ . \tag{47}$$

Then (46) says:

$$I_2 = \int_\tau^b A(\dot{\underline{\Phi}}(t), t)\, dt \ . \tag{48}$$

Since A is linear with respect to $\mathbf{v}$, we have:

$$\frac{d}{dt}\Big[A(\underline{\Phi}(t), t)\Big] = A(\dot{\underline{\Phi}}(t), t) + D_2 A(\underline{\Phi}(t), t) \ , \tag{49}$$

where $D_2 A$ denotes the derivative of A with respect to the second component, i.e.

$$\begin{aligned} D_2 A(\mathbf{v}, t) \ &= \ -\left(\frac{\dot{\underline{\sigma}}(t)}{\underline{\Delta}(t)}\right)\langle \mathbf{v}\underline{\Psi}^\dagger(t), \underline{\nu}(t)\rangle \\ &\quad -\left(\frac{(1-\underline{\sigma}(t))\dot{\underline{\Delta}}(t)}{\underline{\Delta}(t)^2}\right)\langle \mathbf{v}\underline{\Psi}^\dagger(t), \underline{\nu}(t)\rangle \\ &\quad +\left(\frac{1-\underline{\sigma}(t)}{\underline{\Delta}(t)}\right)\langle \mathbf{v}\dot{\underline{\Psi}}^\dagger(t), \underline{\nu}(t)\rangle \\ &\quad +\left(\frac{1-\underline{\sigma}(t)}{\underline{\Delta}(t)}\right)\langle \mathbf{v}\Psi^\dagger(t), \dot{\underline{\nu}}(t)\rangle \ . \end{aligned}$$

Therefore $I_2 = I_{21} - I_{22}$, where

$$I_{21} \ = \ A(\underline{\Phi}(b), b) - A(\underline{\Phi}(\tau), \tau) \ , \tag{50}$$

$$I_{22} \ = \ \int_\tau^b D_2 A(\underline{\Phi}(t), t)\, dt \ . \tag{51}$$

It is easy to see by direct inspection that $||D_2 A(\mathbf{v},t)||$ is bounded by a fixed constant C_2 times $||\mathbf{v}||$ times $||\eta(t)||$ times $\frac{1}{\rho}$. (To see this, first let Q be the set of those $(x,z) \in K_\alpha^*$ such that $\Delta(x,z) \leq \beta$. Observe that $1 - \underline{\sigma}(t)$ and $\dot{\underline{\sigma}}(t)$ vanish unless $\xi^*(t) \in Q$. The functions ν, $\Psi^\dagger$, σ and $\frac{1}{\Delta}$ are bounded on Q. If h is any of these four functions (so h can be scalar- or vector- or matrix-valued), then $\dot{\underline{h}}(t)$ is equal to $\sum_i \eta_i(t)(f_i^* h)(\xi^*(t))$, and the functions $f_i^* h$ are bounded on Q. (In the case when $h = \sigma$, then h is the composite of the Lipschitz function $g = |\varphi_1| + \ldots + |\varphi_m|$ with the function σ_0. The derivatives $f_i^* g$ are of course bounded on K_α^*, and the derivative of σ_0 is bounded by $\frac{1}{\rho}$, which is why $\frac{1}{\rho}$ appears in our bound.) We therefore get the bound

$$|I_2| \leq \frac{C_2}{\rho}\Big(\int_\tau^b ||\eta(t)||^2\,dt\Big)\Big(\int_\tau^b ||\underline{\Phi}(t)||^2\,dt\Big)\ . \tag{52}$$

On the other hand, $A(\mathbf{v},t)$ is bounded by a fixed constant C_3 times $||\mathbf{v}||$, since $\underline{\Delta}(t) \geq \beta$ whenever $\underline{\sigma}(t) \neq 1$. But $\underline{\sigma}(t) \neq 1$ only if $|\underline{\varphi}_1| + \ldots + |\underline{\varphi}_m| \leq \rho$, and in that case $||\underline{\Phi}|| \leq \rho$. So $|A(\underline{\Phi}(t),t)| \leq C_3\rho$. Therefore

$$|I_{21}| \leq 2C_3\rho\ . \tag{53}$$

If we now combine (38), (42, (45), (52), and (53), and let $C_4 = C_1\sqrt{m} + C_2$, we get the bound

$$1 \leq 2C_3\rho + \frac{C_4}{\rho}\Big(\int_\tau^b ||\eta(t)||^2\,dt\Big)\Big(\int_\tau^b ||\underline{\Phi}(t)||^2\,dt\Big)\ . \tag{54}$$

Next, choose ρ such that $2C_3\rho \leq \frac{1}{2}$ (in addition to the requirement imposed before, namely, $\rho^2 \leq \beta$). We then get

$$\frac{1}{2} \leq \frac{C_4}{\rho}\Big(\int_\tau^b ||\eta(t)||^2\,dt\Big)\Big(\int_\tau^b ||\underline{\Phi}(t)||^2\,dt\Big)\ , \tag{55}$$

i.e.

$$\Big(\int_\tau^b ||\eta(t)||^2\,dt\Big)\Big(\int_\tau^b ||\underline{\Phi}(t)||^2\,dt\Big) \geq c\ , \tag{56}$$

where

$$c = \frac{\rho}{2C_4}\ . \tag{57}$$

Clearly, (56) implies the bound

$$\Big(\int_a^b ||\eta(t)||^2\,dt\Big)\Big(\int_a^b ||\underline{\Phi}(t)||^2\,dt\Big) \geq c\ . \tag{58}$$

Our proof is therefore complete. ∎

Theorem 4 implies that Condition **(LGE")** holds. We have therefore proved:

Theorem 5 *Let Σ be a strongly bracket-generating system that satisfies the LARC and the No Explosions Condition. Then the path-lifting equation corresponding to an initial point $\bar{p}$ yields global liftings for any path π that does not go through $\bar{p}$.*

References

[1] Kupka, I., "Abnormal extremals," preprint, Univ. of Toronto, 1992.

[2] Liu, W., and H.J. Sussmann, "Abnormal sub-Riemannian Minimizers," IMA preprint series, #1059, 1992.

[3] Montgomery, R., "Geodesics which do not Satisfy the Geodesic Equations," preprint, Univ. of California at Sta. Cruz, 1991.

[4] Seereeram, S., and J. Wen, "A Global Approach to Path Planning for Redundant Manipulators," submitted to IEEE Robotics and Automation Conference, Atlanta, GA, 1993.

[5] Strichartz, R., "Sub-Riemannian Geometry," *J. Diff. Geom.* 24, 221-263, (1983).

[6] Strichartz, R., "Corrections to 'Sub-Riemannian Geometry'," *J. Diff. Geom.* 30, no. 2, 595-596, (1989).

[7] Sussmann, H.J., "New differential geometric methods in nonholonomic path finding," in *Systems, Models, and Feedback,* A. Isidori and T.J. Tarn Eds., Birkäuser Boston, Inc. 1992.