Subriemannian minimizers

H. J. Sussmann

Department of Mathematics — Rutgers University
Piscataway, NJ 08854, USA
sussmann@math.rutgers.edu

Conference on Nonlinear Partial Differential Equations and Applications

In honor of Jean-Michel Coron's 60th birthday

Paris, June 20-24, 2016

HAPPY BIRTHDAY

JEAN – MICHEL !!

VECTOR DISTRIBUTIONS AND SECTIONS

A vector distribution on a smooth manifold M is a vector subbundle E of the tangent bundle TM of M. (That is, E assigns to each $x \in M$ a linear subspace $E(x)$ of the tangent space $T_x M$, in such a way that the dimension of $E(x)$ is the same for all $x \in M$.

The fiber dimension of E is the dimension of the spaces $E(x)$.

A section of a vector distribution E on M over an open subset Ω of M is a vector field V on Ω such that $V(x) \in E(x)$ for each $x \in \Omega$.

If $\kappa = \infty$ or $\kappa = \omega$, and M is of class C^κ, we use $\Gamma^\kappa(E, \Omega)$ to denote the set of all sections of E over Ω that are of class C^κ.

In this lecture,

a. "manifold" means "finite-dimensional paracompact manifold without boundary",

b. "smooth" means "of class C^∞",

c. If M is a smooth manifold, then TM, T^*M are, respectively, the tangent and cotangent bundles of M. If $x \in M$, then $T_x M$, $T_x^* M$ are, respectively, the tangent and cotangent spaces of M at x.

SMOOTH AND REAL ANALYTIC DISTRIBUTIONS

The vector distribution E is smooth if E is a smooth submanifold of TM. (Equivalently, E is smooth iff for every $x \in M$ and every $v \in E(x)$ there exists a smooth section V on some neighborhood Ω of x such that $V(x) = v$.)

If M is real analytic, then the vector distribution E is real analytic if E is a real analytic submanifold of TM. (Equivalently, E is real analytic iff for every $x \in M$ and every $v \in E(x)$ there exists a real analytic section V on some neighborhood Ω of x such that $V(x) = v$.)

A TRIVIAL WELL-KNOWN FACT: If E is a smooth (or real analytic) distribution on M with fiber dimension d, then for every $x \in M$ there exists a local basis of smooth (or real analytic) sections of E near x, that is, a d-tuple $(X_1, \ldots, X_d)$ of smooth (or real analytic) sections of E defined on an open neighborhood Ω of x such that $(X_1(y), \ldots, X_d(y))$ is a basis of $E(y)$ for every $y \in \Omega$.

GLOBAL SECTIONS

A global section of E is a section of E over M.

A WELL-KNOWN FACT: If E is a smooth distribution on M and $\dim M = n$, then E has $2n+1$ smooth global sections $X_1, \ldots, X_{2n+1}$ such that the vectors $X_1(x), \ldots, X_{2n+1}(x)$ linearly span $E(x)$ for each $x \in M$. (Proof: Use Whitney's mbedding theorem.)

ANOTHER WELL-KNOWN FACT: If E is a real analytic distribution on the real analytic manifold M and $\dim M = n$, then E has $4n+2$ real analytic global sections $X_1, \ldots, X_{4n+2}$ such that the vectors $X_1(x), \ldots, X_{4n+2}(x)$ linearly span $E(x)$ for each $x \in M$. (Proof: Use the Morrey-Grauert embedding theorem.)

RIEMANNIAN METRICS

A smooth Riemannian metric on a smooth distribution E on a smooth manifold M is a map $M \ni x \mapsto G_x$ such that

(i) each G_x is a strictly positive definite symmetric bilinear form on $E(x)$,

(ii) whenever V, W are smooth sections on E over an open subset Ω of M, it follows that the function $\Omega \ni x \mapsto G_x(V(x), W(x))$ is smooth.

If M and E are real analytic then the metric G is real analytic if

(ii') whenever V, W are real analytic sections on E over an open subset Ω of M, it follows that the function $\Omega \ni x \mapsto G_x(V(x), W(x))$ is real analytic.

A WELL-KNOWN FACT: If E is a smooth (or real analytic) distribution on the smooth (or real analytic) manifold M and G is a smooth (or real analytic) Riemannian metric on E, then G is the restriction to E of a smooth (or real analytic) Riemannian metric on M (that is, a Riemannian metric on the full tangent bundle TM).

A TRIVIAL WELL-KNOWN FACT: If E is a smooth (or real analytic) distribution on M with fiber dimension d, and G is a smooth (or real analytic) Riemannian metric on E, then for every $x \in M$ there exists an orthonormal local basis of smooth (or real analytic) sections of E near x, that is, a d-tuple $(X_1, \ldots, X_d)$ of smooth (or real analytic) sections of E defined on an open neighborhood Ω of x such that, for every $y \in \Omega$, $(X_1(y), \ldots, X_d(y))$ is a basis of $E(y)$ for which

$$G_y(X_i(y), X_j(y)) = \delta_{ij} \, .$$

BRACKET-GENERATING DISTRIBUTIONS

If L is any linear space of smooth vector field on a manifold M, and $x \in M$, we define $\boxed{L(x) \stackrel{\text{def}}{=} \{V(x) : V \in L\}\,.}$

We say that L has full rank at x if $L(x) = T_xM$.

If S is any set of smooth vector fields on M, then $L[S]$ will denote the Lie algebra of vector fields generated by S, that is, the smallest Lie algebra (over $\mathbb{R}$) of vector fields that contains S.

A smooth distribution E of fiber dimension d on a manifold M is bracket-generating if one of the following equivalent conditions hold:

(1) $L[\Gamma^\infty(E, M)]$ has full rank at every $x \in M$.

(2) For every $x \in M$, if $(X_1, \ldots, X_d)$ is any basis of sections of E defined on a neighborhood Ω of x, then $L[X_1, \ldots, X_d]$ has full rank at x.

SUBRIEMANNIAN MANIFOLDS

A smooth subriemannian manifold is a triple (M, E, G) such that

(1) M is a smooth manifold,

(2) E is a smooth vector distribution on M,

(3) G is a smooth Riemannian metric on E.

Naturally, we call (M, E, G) real analytic if M, E and G are real analytic.

ADMISSIBLE ARCS

If (M, E, G) is a subriemannian manifold, an admissible arc is an absolutely continuous arc $\xi : [a, b] \mapsto M$, defined on some compact interval $[a, b]$, such that

$$\dot{\xi}(t) \in E(\xi(t)) \qquad \text{for almost every } t \in [a, b] .$$

If we define the G-length $\|v\|_G$ of a tangent vector $v \in T_x M$ by

$$\begin{aligned} \|v\|_G &= \sqrt{G_x(v, v)} \qquad &\text{if} \quad v \in E(x) , \\ \|v\|_G &= +\infty &\text{if} \quad v \notin E(x) , \end{aligned}$$

and let the G-length of an arbitrary absolutely continuous arc $\xi : [a, b] \mapsto M$ be the (finite or infinite) number

$$\|\xi\|_G \stackrel{\text{def}}{=} \int_a^b \|\dot{\xi}(t)\|_G dt ,$$

then it is easy to see that

an absolutely continuous arc $\xi : [a, b] \mapsto M$ is admissible if and only if $\|\xi\|_G < \infty$.

THE SUBRIEMANNIAN DISTANCE

If $\mathcal{M} = (M, E, G)$ is a subriemannian manifold, and $\xi : [a, b] \mapsto M$ is an absolutely continuous arc in M, we use $\partial\xi$ to denote the ordered pair $(\xi(a), \xi(b))$, and refer to $\partial\xi$ as the endpoint value, or boundary value, of ξ.

We then define the distance $d_{\mathcal{M}}(x, y)$ between two points x, y of M to be the infimum of $\|\xi\|_G$, taken over all absolutely continuous arcs ξ in M such that $\partial\xi = (x, y)$.

A WELL-KNOWN FACT: If M is connected, then $(M, d_{\mathcal{M}})$ is a metric space, whose topology is the same as the manifold topology of M. (So, in particular, $d_{\mathcal{M}}$ is continuous on $M \times M$.) (Reason: The bracket-generating condition and Chow's Theorem imply that $d_{\mathcal{M}}(x, y) < \infty$ for all x, y.)

MINIMIZERS AND PAL MINIMIZERS

An absolutely continuous arc $\xi : [a,b] \mapsto M$ is a length minimizer if $\|\xi\|_G = d_{\mathcal{M}}(\partial\xi)$.

An admissible arc $\xi : [a,b] \mapsto M$ is parametrized by arc-length, or a PAL arc, if $\|\dot{\xi}(t)\|_G = 1$ for almost all $t \in [a,b]$.

TRIVIAL FACT: Every admissible arc can be reparametrized to become a PAL arc. (Reason: Just use arc-length as the new time parameter.)

REMARK: In the real analytic case, for the purpose of the results considered here, we may always assume, without loss of generality, that the bracket-generating condition holds even if it originally does not hold.

The reason is: through every point of M there passes a unique maximal integral submanifold of the Lie algebra of vector fields generated by the real analytic global sections of E.

And, in addition: all the results of this talk are about regularity of distance-minimizing arcs, and every such arc is entirely contained in an integral manifold, so in order to study such arcs we can restrict ourselves to an integral manifold.

AN IMPORTANT OPEN QUESTION

Are all PAL minimizers smooth?

This is well known to be true in the Riemannian case (i.e. when $E = TM$), because in that case the PAL minimizers satisfy (in coordinates) the geodesic equation

$$\ddot{\xi}^i(t) + \sum_{j,k} \Gamma^i{}_{jk}(\xi(t))\dot{\xi}^j(t)\dot{\xi}^k(t) = 0\,,$$

where the $\Gamma^i{}_{jk}$ are the Christoffel symbols, which are smooth functions on M.

THE SMOOTHNEESS PROBLEM

To study the smoothness problem, it suffices to work with sufficiently short arcs, so we are allowed to assume

(*) M is an open subset of $\mathbb{R}^n$, and E has an orthonormal basis $(f_1, f_2, \ldots, f_d)$ of smooth (or real analytic) sections.

It follows that the admissible arcs are exactly the arcs ξ that are trajectories of the control system

$$\dot{x} = \sum_{i=1}^{d} u_i f_i(x)\,. \tag{1}$$

Precisely, an arc $\xi : [a,b] \mapsto M$ is admissible if and only if it is absolutely continuous and satisfies

$$\dot{\xi}(t) = \sum_{i=1}^{d} \eta_i(t) f_i(\xi(t)) \text{ for a.e. } t\,, \tag{2}$$

for some integrable function $\eta = (\eta_1, \ldots, \eta_d) : [a,b] \mapsto \mathbb{R}^d$.

Furthermore, ξ is PAL if and only if $\sum_{i=1}^{d} \eta_i(t)^2 = 1$ for a.e. t.

And, for a PAL arc $\xi : [a,b] \mapsto M$, the length of ξ is exactly $b - a$. So the PAL minimizers are the minimum time arcs for the control system (1) with contol constraint $\sum_{i=1}^{d} u_i(t)^2 = 1$.

For convenience, we use instead the control constraint

$$\sum_{i=1}^{d} u_i(t)^2 \leq 1\,. \tag{3}$$

(This does not change the minimizers.)

So, from now on we study the minimum-time trajectories for the system

$$\dot{x} = \sum_{i=1}^{d} u_i f_i(x) \,, \tag{4}$$

with control constraint $\boxed{\sum_{i=1}^{d} u_i^2 \leq 1}$.

Here,
(i) the state x belongs to M, an open subset of $\mathbb{R}^n$,
(ii) $f_1, \ldots, f_d$ are smooth (or real-analytic) vector fields on M,
(iii) if L is the Lie algebra of vector fields on M generated by the f_i, then $L(f_1, \ldots, f_d)(x) = T_x M$ for every point $x \in M$.

To prove regularity theorems for optimal trajectories, we use necessary conditions for optimality, of which the most famous one is the Pontryagin Maximum Principle (PMP). According to the PMP, if $\xi : [a,b] \mapsto M$ is a time-optimal trajectory for the system

$$\dot{x} = \sum_{i=1}^{d} u_i f_i(x)\,, \quad x \in M\,, \quad \sum_{i=1}^{d} u_i^2 \leq 1\,, \tag{5}$$

and $\eta = (\eta_1, \ldots, \eta_d)$ is the corresponding control, then

ξ is the projection of a Hamiltonian-maximizing trajectory $\Xi : [a,b] \mapsto T^{\#}M$ of the Hamiltonian lift of (5), i.e. the system

$$\dot{X} = \sum_{i=1}^{d} u_i \vec{F}_i(X)\,, \quad X \in T^{\#}M\,, \quad \sum_{i=1}^{d} u_i^2 \leq 1\,. \tag{6}$$

Here

(i) $T^{\#}M$ is the cotangent bundle of M with the zero section removed, that is, the set of all pairs (x, p) such that $x \in M$, $p \in T^*_x M$, and $p \neq 0$.

(ii) For $j = 1, \ldots, d$, $\vec{F}_j$ is the Hamiltonian lift of f_j, that is, the Hamilton vector field on $T^{\#}M$ corresponding to the function

$$T^{\#}M \ni (x, p) \mapsto \langle p, f_j(x) \rangle \stackrel{\text{def}}{=} F_j(x, p) \,.$$

Recall that, if v is any vector field on M, then the "momentum function" μ_v is the function on T^*M given by

$$\mu_v(x,p) = \langle p, v(x) \rangle \quad \text{for} \quad x \in M\,,\ p \in T_x^*M\,.$$

Also, φ is a smooth function on an open subset U of T^*M, then the Hamilton vector field arising from φ is the vector field $\vec{\varphi}$ on U given in local coordinates by

$$\vec{\varphi} = \sum_{i=1}^{n} \frac{\partial \varphi}{\partial p_i} \cdot \frac{\partial}{\partial x_i} - \sum_{i=1}^{n} \frac{\partial \varphi}{\partial x_i} \cdot \frac{\partial}{\partial p_i}\,,$$

so that the integral curves $t \mapsto (x(t), p(t))$ of $\vec{\varphi}$ on U satisfy the "Hamilton equations"

$$\dot{x}_i = \frac{\partial \varphi}{\partial p_i}\,, \qquad \dot{p}_i = -\frac{\partial \varphi}{\partial x_i}\,.$$

So the $\vec{F}_j$ are the Hamilton vector fields arising from the momentum functions $F_j = \mu_{f_j}$.

The Hamiltonian of our control system $\dot{x} = \sum_{i=1}^{d} u_i f_i(x)$ is the momentum function of the u-dependent vector field $M \ni x \mapsto \sum_{i=1}^{d} u_i f_i(x) \in T_x M$.

So the Hamiltonian H is a u-dependent function on $T^{\#}M$, given by

$$H_u(x,p) = \sum_{j=1}^{d} u_j \langle p, f_j(x) \rangle \,,$$

for $u = (u_1, \ldots, u_d)$, $x \in M$, $p \in T_x^* M$, $p \neq 0$.

A trajectory of the lifted system for a control $\eta = (\eta_1, \ldots, \eta_d)$, consists of a pair (ξ, π) such that

(i) ξ is a trajectory of the original system for the control η,

(ii) π is a field of nonzero covectors along ξ (that is, $\pi(t)$ is a nonzero convector at $\xi(t)$ for each t),

(iii) π satisfies, in coordinates, the "adjoint equatoon"

$$\dot{\pi}(t) = -\sum_{j=1}^{d} \eta_j(t)\pi(t) \cdot \frac{\partial f_j}{\partial x}\,.$$

(In that case, π is called an "adjoint vector" for (ξ, η).)

HAMILTONIAN MAXIMIZATION

A trajectory Ξ of the lifted system is Hamiltonian-maximizing if there exists a real constant π_0 such that

$$\pi_0 = H_{\eta(t)}(\Xi(t)) = \max\{H_u(\Xi(t)) : u \in \mathbb{B}^d\}$$

for almost every t.

(Here $\mathbb{B}^d = \{u = (u_1, \ldots, u_d) \in \mathbb{R}^d : \sum_{j=1}^d u_j^2 \leq 1\}$.)

Therefore, if (ξ, η) is an optimal trajectory-control pair, there exist an adjoint vector π and a constant π_0 such that

$$\pi_0 = H_{\eta(t)}(\Xi(t)) = \max\{H_u(\Xi(t)) : u \in \mathbb{B}^d\} \quad \text{for a.e. } t\,.$$

Since $u \mapsto H_u(X)$ is a linear function $\sum_{j=1}^d u_j a_j(X)$ for each X, the maximum of this function on the unit ball $\mathbb{B}^d$ is attained for

$$u_j = \frac{a_j(X)}{\sqrt{\sum_{k=1}^d a_k(X)^2}} \quad \text{if} \quad \sum_{k=1}^d a_k(X)^2 \neq 0\,,$$

and u an arbitrary member of $\mathbb{B}^d$ if $\sum_{k=1}^d a_k(X)^2 = 0$.

Furthermore, the maximum value of $H_u(X)$ is $\sqrt{\sum_{k=1}^d a_k(X)^2}$.

Since $a_j(X) = \langle p, f_j(x)\rangle$ if $X = (x,p)$, we obtain

If (ξ, η) is an optimal trajectory-control pair, then there exists a solution π of the adjoint equation along (ξ, η) such that

(i) the number $\pi_0 = \sqrt{\sum_{j=1}^{d} \langle \pi(t), f_j(\xi(t)) \rangle^2}$ is constant (i.e. does not depend on t),

(ii) If $\pi_0 > 0$, then the control η is given by

$$\eta_j(t) = \frac{\langle \pi(t), f_j(\xi(t)) \rangle}{\sqrt{\sum_{k=1}^{d} \langle \pi(t), f_j(\xi(t)) \rangle^2}} = \frac{1}{\pi_0} \langle \pi(t), f_j(\xi(t)) \rangle \, .$$

A Hamiltonian-maximizing adjoint vector π along a trajectory-control pair (ξ, η) is normal if $\pi_0 > 0$, and abnormal if $\pi_0 = 0$.

A minimizer (ξ, η) is normal if it has a normal Hamiltonian-maximizing adjoint vector, and abnormal if it has an abnormal Hamiltonian-maximizing adjoint vector. (Notice that a minimizer can be both normal and abnormal.)

The minimizer (ξ, η) is strictly abnormal if it is abnormal and not normal (that is, if all the Hamiltonian-maximizing adjoint vectors along (ξ, η) are abnormal).

Then the following follows from the PMP:

Every normal minimizer is smooth.

This is because if (ξ,η) is a normal minimizer, and π is a normal Hamiltonian-maximizing adjoint vector along (ξ,η), then the pair (ξ,π) is (in coordinates) a solution of the system of ordinary differential equations

$$\begin{aligned}\dot{\xi} &= \frac{1}{\pi_0}\sum_{j=1}^{d}\langle\pi, f_j(\xi)\rangle f_j(\xi)\,,\\ \dot{\pi} &= -\frac{1}{\pi_0}\sum_{j=1}^{d}\langle\pi, f_j(\xi)\rangle\frac{\partial f_j}{\partial x}(\xi)\,,\end{aligned}$$

from it follows easily that ξ and π are smooth functions.

And then the control η is also smooth, because

$$\eta_j(t) = \frac{1}{\pi_0}\langle\pi(t), f_j(\xi(t))\rangle\,.$$

REMARK: The previous result contains, in particular, the theorem on the smoothness of Riemannian geodesics.

That's because in the Riemannian case every minimizer is normal.

Indeed, if $d = n = \dim M$, then the sum $\sum_{j=1}^{d} \langle \pi(t), f_j(\xi(t)) \rangle^2$ cannot vanish, because if it did then all the numbers $\langle \pi(t), f_j(\xi(t)) \rangle$ would vanish, so $\pi(t)$ would be the zero covector, because the vectors $f_j(\xi(t))$ span the tangent space $T_{\xi(t)}M$. But $\pi(t) \neq 0$, because $(\pi(t), \xi(t)) \in T^{\#}M$.

Clearly, now, the problem of the smoothness of optimal trajectories reduces to that of the smoothness of strictly abnormal extremals.

For this problem, there is another useful necessary condition called the **Goh condition**:

If (ξ, η) is a strictly abnormal minimizer, then there exists a Hamiltonian maximizing adjoint vector π along (ξ, η) such that

$$\langle \pi(t), [f_j, f_k](\xi(t)) \rangle = 0$$

for all t and all j, k.

Call a smooth distribution E 2-generating if, for every $x \in M$, the vectors $f(x)$, for f a smooth section of E near x, together with the vectors $[f, g](x)$, for f, g smooth sections of E near x, span the full tangent space $T_x M$.

It the follows from the Goh condition that

If E is 2-generating, then every PAL minimizer is smooth (and real analytic in the real analytic case).

Is it possible that the Goh condition can be generalized to the following statement?

> If (ξ, η) is a strictly abnormal minimizer, then there exists a Hamiltonian-maximizing adjoint vector π along (ξ, η) such that
>
> $$\langle \pi(t), [[f_{j_1}, [f_{j_2}, \ldots, [f_{j_{k-1}}, f_{j_k}] \ldots]](\xi(t)) \rangle = 0$$
>
> for all t, all k, and all $(j_1, \ldots, j_k)$.

The answer is NO, because if such a condition was true, it would follow from the bracket-generating condition that strictly abnormal minimizers do not exist.

But:

Strictly abnormal miminizers exist. (Examples due to R. Montgomery, Sussmann-Liu, Kupka.)

A RECENT THEOREM

If $\mathcal{M} = (M, E, G)$ is a real analytic sub-riemannian manifold, then every PAL minimizer is real analytic on an open dense subset of its interval of definition.

The proof is by induction on the fiber dimension d.

The case $d = 1$ is trivial.

So we assume that the theorem is true for fiber dimensions $\leq d - 1$, and prove it's true for fiber dimension d.

Assume that $\mathcal{M} = (M, E, G)$ is a real analytic subriemannian manifold with fiber dimension d.

Also, as explained before, we assume that M is an open subset of a Euclidean space $\mathbb{R}^n$, and E has an orthonormal basis $(f_1, \ldots, f_d)$ of real analytic sections.

The key to the proof will the construction of a stratification of $T^{\#}M$ with certain special properties.

So we have to start by explaining what a "stratification" is.

A stratification of a smooth manifold Q is a locally finite partition $\mathcal{P}$ of Q into smooth, connected, relatively compact embedded submanifolds (called the "strata" of $\mathcal{P}$) such that the following conditions hold:

(1) the closure $\operatorname{Clos} S$ of each stratum is a union of strata,

(2) if S is a stratum, then the "frontier strata" of S—that is, the strata that are contained in $(\operatorname{Clos} S)\backslash S$—are of dimension strictly smaller than $\dim S$.

A stratification $\mathcal{P}$ of Q is compatible with a family $\mathcal{A}$ of subsets of Q if every $A \in \mathcal{A}$ is a union of strata of $\mathcal{P}$.

Now, in the special case when $Q = T^{\#}M$, we consider stratifications whose strata are "nice", in the sense described below.

For a submanifold S of $T^{\#}M$, we let $U_S(X)$ be, for $X \in S$, the set of all $u = (u_1, \ldots, u_d) \in \mathbb{R}^d$ such that the vector $u.\vec{F}(X)$ (defined by $u.\vec{F}(X) = \sum_{i=1}^{d} u_i \vec{F}_i(X)$) is tangent to S. It is clear that each $U_S(X)$ is a linear subspace of $\mathbb{R}^d$.

We call a submanifold S of $T^{\#}M$ "nice" if S is a real analytic submanifold, the subspace $U_S(X)$ is of constant dimension (as X varies over S), and $U_S(X)$ depends real analytically on X, in the sense that there exist real analytic maps $e_j : S \mapsto \mathbb{R}^d$, for $j = 1, \ldots, \delta_S$ (where $\delta_S = \dim U_S(X)$ for $X \in S$) such that $(e_1(X), \ldots, e_{\delta_S}(X))$ is a basis of $U_S(X)$ for each $X \in S$.

THE KEY TECHNICAL LEMMA

LEMMA: If $\mathcal{M} = (M, E, G)$ is a real analytic subriemannian manifold of fiber dimension d, then there exists a stratification $\mathcal{P}$ of $T^{\#}M$ such that

(1) $\mathcal{P}$ is compatible with $\{A\}$, where A is the "abnormal set", given by

$$A = \{(x, p) \in T^{\#}M : \sum_{j=1}^{d} \langle p, f_j(x) \rangle^2 = 0\}.$$

(2) Every stratum of $\mathcal{P}$ is nice.

HOW THE LEMMA IMPLIES OUR THEOREM

Let $\mathcal{P}$ be a stratification of $T^{\#}M$ having the properties of the lemma.

Suppose (ξ, η) is a PAL trajectory-control pair which is a minimizer, and is defined on an interval $[a, b]$.

It suffices to prove that there exists τ such that $a < \tau < b$ and ξ is real analytic on a neighborhood $]\tau - \varepsilon, \tau + \varepsilon[$ of τ.

We may assume that (ξ, η) is strictly abnormal.

Let π be a Hamiltonian-maximizing adjoint vector, and let $\Xi = (\xi, \pi)$ be the corresponding Hamiltonian lift.

Then Ξ is entirely contained in the abnormal set A. So each point $\Xi(t)$ of Ξ belongs to a unique stratum $S(t)$ of $\mathcal{P}$, and $S(t) \subseteq A$.

Let $\delta(t) = \dim S(t)$.

Pick $\bar{t}$ such that $\delta(\bar{t}) = \max\{\delta(t) : a < t < b\}$.

Let $S = S(\bar{t})$.

We claim that $\Xi(t) \in S$ for t in some neighborhood of $\bar{t}$.

To see this suppose otherwise. Then there exists a sequence $\{t_k\}_{k=1}^{\infty}$, converging to $\bar{t}$, such that $S(t_k) \neq S$ for every k.

Since $\mathcal{P}$ is locally finite, the set of all the $S(t_k)$ is finite, so we may assume, after passing to a subsequence, that all the $S(t_k)$ are one and the same stratum $\check{S}$. Naturally, $\check{S} \neq S$. Since the point $\Xi(\bar{t})$ is in S, so $\Xi(\bar{t}) \notin \tilde{S}$, but $\Xi(\bar{t})$ is a limit of points of $\check{S}$, it follows that $\Xi(\bar{t}) \in (\text{Clos}\,\tilde{S})\backslash\tilde{S}$. So $S \subseteq (\text{Clos}\,\check{S})\backslash\check{S}$, and this implies that $\dim S < \dim \check{S}$. (Notice the crucial role played by the condition on the dimension of the frontier strata!) Hence $\delta(\bar{t}) < \delta(t_k)$ contradicting our choice of $\bar{t}$ as a time when δ has its maximum value.

We now know that, for some positive ε,

$$\Xi(t) \in S \text{ for } t \in I \text{, where } I = [\bar{t}-\varepsilon, \bar{t}+\varepsilon].$$

Since S is a nice stratum, the spaces $U_S(X)$, for all $X \in S$, have the same dimension $\bar{d}$, and there exist real analytic maps $e_j : S \mapsto \mathbb{R}^d$, for $j = 1, \ldots, \bar{d}$, such that

$$\mathbf{e}(X) \stackrel{\text{def}}{=} (e_1(X), \ldots, e_{\bar{d}}(X))$$

is a basis of $U_S(X)$ for each X. Furthemore, we can apply the Gram-Schmidt orthogonalization procedure and assune that $\mathbf{e}(X)$ is an orthonormal basis of $U_S(X)$.

Let Ξ_I be the restriction of Ξ to I. Then Ξ_I is a trajectory of the system

$$\dot{X} = \sum_{j=1}^{d} u_j \vec{F}_j(X)\,,$$

but the vectors $\dot{\Xi}_I(t)$, for $t \in I$, are tangent to S.

Therefore the control η is such that $\eta(t) \in U_S(\Xi_I(t))$ for each $t\,\mathrm{ln}\,I$, and then $\eta(t)$ is a linear combination

$$\eta(t) = \sum_{k=1}^{\bar{d}} c_k(t) e_k(t)\,,$$

with coefficients $c_k(t)$ such that $\sum_{k=1}^{\bar{d}} c_k(t)^2 = 1$.

So Ξ_I is a PAL trajectory of the system

$$\dot{X} = \sum_{k=1}^{\bar{d}} v_k G_k(X)\,, \tag{7}$$

with state space S and control constraint $\sum_{k=1}^{\bar{d}} v_k^2 \leq 1$.

(Here $G_k = e_k \cdot \vec{F}$.)

Furthermore, Ξ_I is a time-optimal trajectory of (7). (This is easy.)

We now prove the

KEY OBSERVATION: $\boxed{\bar{d} < d}$

To prove this, assume $\bar{d} = d$. ($\bar{d}$ cannot be larger than d, because $\bar{d} = \dim U_S(X)$, and $U_S(X)$ is a subspace of $\mathbb{R}^d$.)

This would mean that $U_S(X) = \mathbb{R}^d$ for every $X \in S$.

Then every vector field $\vec{F}_j$ would be tangent to S.

Then every iterated Lie bracket $[\vec{F}_{i_1}, [\vec{F}_{i_2}, \ldots, [\vec{F}_{i_{\ell-1}}, \vec{F}_{i_\ell}] \ldots]]$ would be tangent to S as well. Since the momentum functions F_i vanish on S (because $S \subseteq A$), it follows that the directional derivatives $[\vec{F}_{i_1}, [\vec{F}_{i_2}, \ldots, [\vec{F}_{i_{\ell-1}}, \vec{F}_{i_\ell}] \ldots]]F_j$ vanish on S.

That is, all the iterated Poisson brackets

$$\Psi = \{\{F_{i_1}, \{F_{i_2}, \ldots, \{F_{i_{\ell-1}}, F_{i_\ell}\} \ldots\}\}, F_j\}$$

vanish.

But $\Psi(x,p) = \langle p, [[f_{i_1}, [f_{i_2}, \ldots, [f_{i_{\ell-1}}, f_{i_\ell}] \ldots]], f_j](x)\rangle$.

Hence p vanishes against all the iterated brackets of the f_k. By the bracket-generating assumption, it follows that $p = 0$, contradicting the fact that $S \subseteq T^{\#}M$.

So Ξ_I is a PAL solution of a minimum-time optimal control problem exactly like our original problem, except that the fiber dimension of the new problem is smaller.

Hence, by the inductive hypothesis, there is a point τ of I such that Ξ_I is real analytic near $\bar{\tau}$.

And then of course, ξ is real analytic near $\bar{\tau}$.

Q.E.D.

This was the easy part of the proof.

Now we need the technical part:

How do we prove the existence of a nice stratification compatible with $\{A\}$?

This is done using the machinery of subanalytic sets.

The construction is long and tedious, but conceptually quite simple.

SUBANALYTIC SETS

A semianalytic subset of a real analytic manifold Q is a subset S of Q such that every point x of Q has an open neighborhood U on which there exist a finite collection $\{f_1, \ldots, f_m\}$ of real analytic functions such that $S \cap U$ belongs to the Boolean algebra generated by the sets $\{y \in U : f_j(y) > 0\}$ and $\{y \in U : f_j(y) = 0\}$.

A subanalytic subset of Q is a subset S of QQ such that there exist a real analytic manifold R, a semianalytic subset T of R, and a real analytic map $f : R \mapsto Q$, such that f is proper on $\text{Clos}_R T$ and $f(T) = S$.

THE BASIC STRATIFICATION THEOREM

Given a locally finite collection $\mathcal{A}$ of subanalytic subsets of a real analytic manifold Q, there exists a stratification $\mathcal{P}$ of Q by real analytic submanifolds that are subanalytic subsets of Q, such that $\mathcal{P}$ is compatible with $\mathcal{A}$.

Using this fact, we can start the construction of the stratification we want by observing that the abnormal set A is obviously subanalytic (because it is in fact a real analytic set) and then constructing a stratification $\mathcal{P}_0$ compatible with $\{A\}$ and consisting of real analytic submanifolds that are subanalytic subsets of $T^{\#}M$.

Next we use various techniques to construct refinements of $\mathcal{P}_0$ until we get one that has the desired properties.

The key fact to be proved first is:

If S is a connected embedded relatively compact real analytic submanifold of $T^{\#}M$ which is a subanalytic subset of $T^{\#}M$, then the graph $G(U_S)$ of U_S is a subanalytic subset of $T^{\#}M\times\mathbb{R}^d$.

Here

$$G(U_S) \stackrel{\text{def}}{=} \{(x,u) : x \in S \text{ and } u \cdot \vec{F}(x) \in T_x S\}\,,$$

so $G(U_S) \subseteq T^{\#}M \times \mathbb{R}^d$.

Then you take each stratum S of $\mathcal{P}_0$, and prove that for each dimension δ between 0 and d, the set

$$\Sigma_\delta(S) = \{x \in S : \dim U_S(x) = \delta\}$$

is subanalytic.

Clearly, the $\Sigma_\delta(S)$, for $\delta = 0, 1, \ldots, d$, form a finite partition of S into subanalytic sets.

Also, the $\Sigma_\delta(S)$, for all $S \in \mathcal{P}_0$ and all $\delta \in \{0, 1, \ldots, d\}$, form a locally finite family of subanalytic subsets of $T^\# M$.

So we can construct a stratification $\mathcal{P}_1$ by real analytic submanifolds that are subanalytic subsets, such that $\mathcal{P}_1$ is compatible with the $\Sigma_\delta(S)$.

So now we have a stratification such that, if $T \in \mathcal{P}_1$, $T \subseteq S$, and $S \in \mathcal{P}_0$, then the function $T \ni x \mapsto \dim U_S(x)$ is constant.

Then, with a lot of extra work (including the inductiove construction of several refinements of $\mathcal{P}_1$), we end up with a stratification $\tilde{\mathcal{P}}$ such that the function $S \ni x \mapsto \dim U_S(x)$ is constant for each $S \in \tilde{\mathcal{P}}$.

Now we know that for each stratum S of $\tilde{\mathcal{P}}$, the dimension d_S of $U_S(x)$ is the same for all $x \in S$.

Let $\mathcal{E}_S(x)$ be the set of all orthonormal bases $(e_1, \ldots, e_{d_s})$ of $U_S(x)$.

Then we prove that the graph

$$G(\mathcal{E}_S) \stackrel{\text{def}}{=} \{(x, \mathbf{e}) : x \in S \text{ and } \mathbf{e} \in \mathcal{E}_S(x)\}$$

is a subanalytic subset of $T^{\#}M \times (\mathbb{R}^d)^{d_s}$. (This requires some work.)

We regard $\mathcal{E}_S$ as a set-valued function from S to $(\mathbb{R}^d)^{d_s}$, and make a single-valued subanalytic selection E_S, so E_S is an ordinary single-valued function from S to $(\mathbb{R}^d)^{d_S}$.

Then you need a few more refinements to end up with a stratification in which the functions E_S are not just subanalytic (i.e. with a subanalytic graph) but actually real analytic.

And this gives the desired nice stratification.

Can one do something similar in the smooth (i.e., C^∞) case?.

Yes, but only (so far) for $d = 2$.

ANOTHER RECENT THEOREM

If $\mathcal{M} = (M, E, G)$ is a smooth subriemannian manifold, and $d = 2$, then every PAL minimizer is smooth on an open dense subset of its interval of definition.

REMARK: This argument does not seem to work for $d > 2$.

PROOF: Write our system (locally) as

$$\dot{x} = u_1 f_1(x) + u_2 f_2(x)\,, \quad x \in U\,, \quad u_1^2 + u_2^2 \leq 1\,,$$

where U is an open subset of $\mathbb{R}^n$.

Let $\mathcal{B}$ be the set of all iterated Lie brackets B of the form

$$B = [f_{i_m}, [f_{i_{m-1}}, [\cdots, [f_2, f_1] \cdots]]]$$

where each i_j is 1 or 2.

(The number m is the degree of B.)

It is easy to see, using the Jacobi identity, that every iterated of f_1 and f_2 is a linear combination of the brackets in $\mathcal{B}$. (For example,

$$[[f_1, f_2], [f_1, [f_1, f_2]]] = [f_1, [f_2, [f_1, [f_1, f_2]]]] - [f_2, [f_1, [f_1, f_1, f_2]]]\,,$$

and so on.)

It then follows from the bracket-generating condition that

> For every $x \in U$, the linear span of the vectors $B(x)$, $B \in \mathcal{B}$, is the whole space.

To each vector field X on $\mathbb{R}^n$, associate the momentum function

$$\mu_X : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}\,,$$

given by

$$\mu_X(p, x) = p^\dagger \cdot X(x)\rangle\,.$$

(Here p is a covector, and $\mathbb{R}^n \times \mathbb{R}^n$ is the cotangent bundle of the state space.)

Then there is a simple formula for differentiating a momentum function μ_X along Hamiltonian lifts of trajectories.

Roughly:

$$\dot{\mu}_X = \eta_1 \mu_{[f_1, X]} + \eta_2 \mu_{[f_2, X]}\,.$$

More precisely, if $\xi : [a,b] \mapsto \mathbb{R}^n$ is a trajectory, corresponding to a control

$$[a,b] \ni t \mapsto \eta(t) = (\eta_1(t), \eta_2(t)) \in \mathbb{R}^2 ,$$

and

$$\Xi : [a,b] \mapsto \mathbb{R}^n \times \mathbb{R}^n ,$$

is a Hamiltonian lift of ξ, given by

$$\Xi(t) = (\pi(t), \xi(t)) ,$$

where

$$\dot{\pi}(t) = -\eta_1(t)\pi(t)^\dagger \cdot \frac{\partial f_1}{\partial x}(\xi(t)) - \eta_2(t)\pi(t)^\dagger \cdot \frac{\partial f_2}{\partial x}(\xi(t))$$

we have

$$\frac{d}{dt}\Big(\mu_X(\Xi(t))\Big) = \eta_1(t)\mu_{[f_1,X]}(\Xi(t)) + \eta_2(t)\mu_{[f_2,X]}(\Xi(t))$$

for every smooth vector field X

If (ξ, η) is a PAL minimizer, then it is either a normal extremal, in which case it is smooth, or a strictly abnormal one, in which case there exists a Hamiltonian lift $\Xi = (\pi, \xi)$ such that

$$\mu_{f_1}(\Xi(t)) \equiv \mu_{f_2}(\Xi(t)) \equiv \mu_{[f_1,f_2]}(\Xi(t)) \equiv 0\,.$$

For every $t \in [a, b]$, there exists a $B \in \mathcal{B}$ such that

$$\mu_B(\Xi(t)) = \pi(t)^\dagger \cdot B(\xi(t)) \neq 0\,.$$

(Reason: if $\pi(t)^\dagger \cdot B(\xi(t)) = 0$ for every $B \in \mathcal{B}$, then $\pi(t)^\dagger v = 0$ for every $v \in \mathbb{R}^n$, so $\pi(t) = 0$.)

Given t, let $m_*(t)$ be the smallest m such that $\mu_B(\Xi(t)) \neq 0$ for some $B \in \mathcal{B}$ of degree m.

Pick a $t_0 \in [a,b]$ such that

$$m_*(t_0) \leq m_*(t) \quad \text{for every } t \in [a,b]\,.$$

Let $\widehat{m} = m_*(t_0)$. Let B be a bracket in $\mathcal{B}$ of degree $\widehat{m}$ such that $\mu_B(t_0) \neq 0$.

Write $B = [f_{i_{\widehat{m}}}, C]$, where $C \in \mathcal{B}$ is of degree $\widehat{m} - 1$.

Then $\mu_C(\Xi(t)) = 0$ for all $t \in [a,b]$.

Assume $i_{\widehat{m}} = 1$. (The case when $i_{\widehat{m}} = 2$ is identical.)

Then

$$\frac{d}{dt}\Big(\mu_C(\xi(t))\Big) \equiv 0\,.$$

Therefore

$$\eta_1(t)\mu_{[f_1,C]}(\Xi(t)) + \eta_2(t)\mu_{[f_2,C]}(\Xi(t)) \equiv 0\,,$$

that is

$$\eta_1(t)\mu_B(\Xi(t)) + \eta_2(t)\mu_{[f_2,C]}(\Xi(t)) \equiv 0\,,$$

Let ε be a positive number such that

$$\mu_B(\Xi(t)) \neq 0 \quad \text{for } t \in (t_0\varepsilon, t_0+\varepsilon)\,.$$

Let

$$I = (t_0\varepsilon, t_0+\varepsilon)\,.$$

Since $\eta_1(t)^2 + \eta_2(t)^2 = 1$, the vector $\eta(t)$ is determined uniquely, up to a sign, for $t \in I$.

More precisely: let Ω be the set of points $(p,x) \in \mathbb{R}^n \times \mathbb{R}^n$ such that $\mu_B(p,x) \neq 0$. Then $\Xi(t) \in \Omega$ for all $t \in I$.

On Ω, define "feedback controls" $\varphi_j : \Omega :\mapsto \mathbb{R}$ by letting

$$\begin{aligned}
\varphi_1(p,x) &= -\frac{\mu_{[f_2,C]}(p,x)}{\sqrt{\mu_B(p,x)^2 + \mu_{[f_2,C]}(p,x)^2}}\,, \\
\varphi_2(p,x) &= \frac{\mu_B(p,x)}{\sqrt{\mu_B(p,x)^2 + \mu_{[f_2,C]}(p,x)^2}}\,.
\end{aligned}$$

(These are smooth functions on Ω because μ_B and $\mu_{[f_2,C]}$ are smooth functions and μ_B is $\neq 0$ on Ω.)

Then define a "feedback Hamiltonian vector field" V on Ω by letting

$$V(p,x) = \varphi_1(p,x)F_1(p,x) + \varphi_2(p,x)F_2(p,x)\,,$$

where F_1, F_2 are the Hamiltonian lifts of f_1 and f_2.

Then our Hamiltonian lift Ξ satisfies

$$\dot{\Xi}(t) = \pm(V(\Xi(t))) \text{ for a.e. } t \in I\,.$$

So Ξ is "almost" a trajectory of a smooth vector field. So ξ is smooth, if we can make sure that the sign in the previous equation does not change.

Furthermore, the sign cannot change, because if it did our trajectory would not be optimal. (Easy proof.)

Hence Ξ is smooth on I, and so is ξ.

So we have proved that ξ is smooth on a nonempty open subinterval of $[a, b]$.

This result can be applied to the restriction of ξ to any nontrivial subinterval of $[a, b]$, and our conclusion follows.

One can also prove that for every optimal trajectory the control is obtained by "shuffling smooth feedback controls", in the following sense:

There exists a sequence $(S_j)_{j=0}^{\infty}$ of measurable sets and a sequence $(V_j)_{j=1}^{\infty}$ of smooth feedback vector fields defined on open subsets Ω_j of $\mathbb{R}^n \times \mathbb{R}^n$, such that S_0 has measure zero, and

$$
\begin{aligned}
[a,b] &= \bigcup_{j=0}^{\infty} S_j \,, \\
\Xi(t) &\in \Omega_j \qquad \text{whenever } j \geq 1 \text{ and } t \in S_j \,, \\
\dot{\Xi}(t) &= \pm V_j(\Xi(t)) \qquad \text{whenever } j \geq 1 \text{ and } t \in S_j \,.
\end{aligned}
$$

The proof is essentially as before:

1. For each $B \in \mathcal{B}$, let A_B be the set of all points $t \in [a, b]$ such that $\mu_B(\Xi(t)) \neq 0$ but $\mu_C(\Xi(t)) = 0$ for all $C \in \mathcal{B}$ of smaller degree.

2. For each $B \in \mathcal{B}$, construct a smooth feedback vector field V_B as before.

3. Fix B. Let $B = [f_i, C]$, where $i = 1$ or $i = 2$, and $C \in \mathcal{B}$.

4. Let S_B be the set of all points of density of A_B that are also Lebesgue points of the control η. (This is a huge subset of A_B. In particular, $A_B \backslash S_B$ is a null set.)

5. Assume $i = 1$.

6. For a point $t \in S_B$, the numbers $\mu_C(\Xi(s_j))$ vanish for the points s_j in a sequence $(s_j)_{j=1}^{\infty}$ of points that converge to t and are distinct from t.

7. So we can compute the derivative $\dot{\mu}_C(\Xi(t))$ as the limit as $j \to \infty$ of the quotient

$$\frac{\mu_C(\Xi(s_j)) - \mu_C(\Xi(t))}{s_j - t}.$$

8. This derivative vanishes, so

$$\eta_1(t)\mu_B(\Xi(t)) + \eta_2(t)\mu_{[f_2,C]}(\Xi(t)) = 0.$$

9. Hence $\dot{\Xi}(t) = \pm V_B(\Xi(t))$.

10. Clearly, the set $S_0 = [a,b] \backslash \bigcup_{B \in \mathcal{B}} S_B$ has measure zero. So we are done.

What is to be done next?

If you want to prove that all minimizers are smooth, then you must continue the analysis that was started in work by Leonardi and Monti, and more recent work by Hakavuori and Le Donne, who proved that PAL subriemannian minimizers cannot have corners.

If, on the other hand, you want to prove that it is not true that all minimizers are smooth, then you have to find an example of a nonsmooth minimizer. This will be, of course, a strictly abnormal extremal.

Finding nonsmooth abnormal extremals is easy.

And finding lots of strictly abnormal minimizers is also easy. (See HJS "A cornucopia of abnormal minimizers".)

The hard part is to prove that a given extremal is actually a minimizer.

For normal extremals there are methods, especially the construction of fields of extremals. But these do not work for abnormal extremals.

The methods of the "cornucopia" paper do work for lots of families of abnormal extremals, but I do not know how to construct nonsmooth abnormal extremals to which those methods apply.