Scuola Internazionale Superiore di Studi Avanzati - Trieste

Area of Mathematics
Ph.D. in Applied Mathematics

Thesis

Topics in sub-Riemannian Geometry

Candidate
Francesco Boarotto

Supervisor
prof. Andrei Agrachev

Thesis submitted in partial fulfillment of the requirements for the degree of Philosophiæ Doctor

academic year 2015-16

SISSA - Via Bonomea 265 - 34136 TRIESTE - ITALY

The present Thesis, defended on September 30th, 2016, constitutes the work presented by Francesco Boarotto, under the supervision of Prof. Andrei A. Agrachev at SISSA.

Trieste, academic year 2015-2016.

A mamma e papà, perché mi fanno
sentire il loro affetto ogni giorno.
A Francesca, perché questo lavoro
è frutto anche della serenità che mi dai.
A tutti gli amici che hanno reso questi
quattro anni bellissimi.

Abstract

The central object of this thesis is sub-Riemannian geometry. By its very own nature, this relatively new subject can be investigated under very different perspectives: still all of the question that we may pose have a common root in the deceivingly simple differential system

$$\dot{x}(t) = \sum_{i=1}^{k} u_i(t) X_i(x(t)),$$

where $u \in L^p([0,T], \mathbb{R}^k)$, and the set $\{X_i\}_{i=1}^k$, with $k < \dim(M)$, indicates a family of smooth vector fields on a manifold M, which satisfies the Hörmander bracket generating condition

$$\operatorname{span}\left\{[X_{i_1}, \dots, [X_{i_{j-1}}, X_{i_j}]](x) \mid X_{i_l} \in \mathcal{F}, j \in \mathbb{N}\right\} = T_x M, \quad \forall x \in M.$$

Then, in essence, a sub-Riemannian structure is the datum $(M, \Delta, \mathbf{g})$, where the distribution $\Delta \subset TM$ is equipped with a scalar product $\mathbf{g}$ that smoothly varies with respect to base point. The freedom we have on Δ and its low dimensionality with respect to $\dim(M)$, pose a series of new questions which range from differential geometry to nonlinear analysis, from topology to dynamical systems. What we try to do here is to give original contributions in some of these areas, and in fact every chapter of this work addresses a different topic.

The first work presented in Chapter 2 deals with differential geometry and concerns the (local) conformal classification of the simplest sub-Riemannian structures we can consider, namely contact three dimensional sub-Riemannian structures on Lie Groups. In Chapter 3 we turn instead to topology, and we investigate the structure of the fibers of the Endpoint map from the homotopy point of view; indeed, as opposed to the Riemannian case, the non surjectivity of the differential of the Endpoint map allows these subsets to be possibly terrible objects from the differentiable point of view, but still if we weaken a bit our perspective and work with less sophisticated tools, again many nice conclusions can be drawn. Pursuing a bit further this philosophy we finally come to Chapter 4, where we show that abnormal curves do not affect the topology of the manifold M and are "invisible" to the classical Morse theory, at least for the generic sub-Riemannian structure. On the other hand our techniques provide an effective counterpart to the standard nonlinear analysis tools we use in the Riemannian setting, and can be seen as a real starting point to develop variational calculus within this framework; in particular we give two very practical applications to exemplify the effectiveness of the machinery developed so far.

All the works collected in this thesis are either submitted or already accepted in different mathematical journals: here we list the bibliographical references

Chapter 2: F. Boarotto, *Conformal Equivalence of 3D Contact Structures on Lie Groups*, Journal of Dynamical and Control Systems, vol. 22, no. 2, 2016.

Chapter 3: F. Boarotto and A. Lerario, *Homotopy Properties of Horizontal Path Spaces and a Theorem of Serre in sub-Riemannian Geometry*, to appear on Communications in Analysis and Geometry.

Chapter 4: A. A. Agrachev, F. Boarotto and A. Lerario, *Homotopically invisible singular curves*, submitted to Calculus of Variations, preprint on arXiv at `https://arxiv.org/pdf/1603.08937.pdf`.

Contents

CHAPTER 1

Introduction

1. Optimal Control Problems

Optimal control problems have attracted an increasing attention in the last decades and have been proved to be flexible enough to cover both classical and new fields of mathematics. In particular, we are much concerned with the possibility they offer to describe geometry, both in the classical Riemannian and in the more recent sub-Riemannian viewpoint.

In fact, the idea that the *distance* between two points x and y should be quantified by the *length* of the shortest path that connects them, leads immediately to think about a minimization problem with constraints. In the classical Riemannian context, all we need to measure the length of a curve is the knowledge of the length of its velocities. This is well formalized with the concept of a Riemannian manifold, where each tangent space is equipped with an Euclidean structure, which smoothly varies with respect on the base point. An observer standing on a fixed point on the manifold, should be able to send and receive informations using all the directions available in the tangent space; if all of them are at disposal, then he should be able to reconstruct the whole geometry of the ambient space.

Assume that for some reasons there are restrictions on the admissible directions we have at disposal. Then the "admissible" tangent space should in principle have the freedom of change smoothly from point to point, but this operation may sound a bit awkward to be justified within the frame of pure geometry.

A control-theoretic approach has the enormous advantage of changing this point of view from the beginning. Instead of speaking about admissible tangent spaces we bring into the picture the global space of *admissible vector fields*; if we pick generators $X_1, \dots, X_k$, then admissible paths are described as solutions to time-varying differential equations of the form

$$\dot{x}(t) = \sum_{i=1}^{k} u_i(t) X_i(x(t)), \tag{1.1}$$

where the map u belong in general to some $L^p([0,T],\mathbb{R}^k)$ space of functions, and is called the *control*.

It is somewhat natural at this point to ask whether any pair of points x and y can be joined by at least an admissible path, for otherwise the definition of distance could be vacuous. It turns out that this request is indeed satisfied by a large class of geometric structures, that is those satisfying the so called *bracket generating condition*, which is defined in terms of the vector fields $X_1, \dots, X_k$ as follows:

Definition 1. *Let M be a smooth connected manifold and consider on it the family $\mathcal{F} = \{X_1, \dots, X_k\}$ of smooth vector fields. We call the* Lie algebra generated by $\mathcal{F}$ *the smallest subalgebra of* $\mathrm{Vec}(M)$ *containing $\mathcal{F}$. In particular, we have the equality*

$$Lie\mathcal{F} = \mathrm{span}\left\{[X_{i_1}, \dots, [X_{i_{j-1}}, X_{i_j}]] \;\middle|\; X_{i_l} \in \mathcal{F},\, j \in \mathbb{N}\right\}.$$

We say that $\mathcal{F}$ *is* bracket generating *on* M *if*

$$Lie_x\mathcal{F} = \{X(x) \,|\, X \in Lie\mathcal{F}\} = T_xM$$

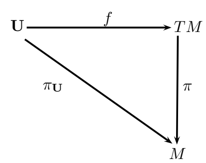
holds for every point $x \in M$.

It is clear, at least heuristically, why we should expect such a condition. We already explained that some directions are forbidden; yet by analogy with the Riemannian setting we would like that an observer standing on a point of a sub-Riemannian manifold should be able to reconstruct his space by sending and receiving signals from his position. If this is not possible using the admissible vector fields, we should require that, upon paying some "penalties" (i.e. using the brackets), we can in fact reach any point in the space.

We are finally in the position to give a precise definition of a sub-Riemannian geometry:

Definition 2. *Let* M *be a connected smooth manifold. A* sub-Riemannian structure *on* M *is a pair* $(\mathbf{U}, f)$ *where:*

a) $\mathbf{U}$ *is an Euclidean bundle over* M *whose fibers* U_q *are vector spaces endowed with a scalar product* $\mathbf{g}_q$ *smoothly varying with respect to the base point* q.
b) $f : \mathbf{U} \to TM$ *is a smooth morphism of vector bundles which makes the following diagram*



commutative. Moreover, we require that f *is* linear *on fibers.*
c) *The set of* horizontal vector fields $\Delta = \{f(\sigma) \mid \sigma : M \to \mathbf{U}$ *is a smooth section*$\}$ *is a bracket generating family of vector fields.*

In this language, to be able to write down something like (1.1), is equivalent to have a local trivialization chosen for the bundle $\mathbf{U}$; moreover if $\mathbf{U}$ is *globally trivializable*, we speak of a *free* sub-Riemannian structure. It turns out that every sub-Riemannian structure is in some sense equivalent to a free one, and therefore we will always present it as in (1.1).

Finally, notice that control theory deals with a larger class of bundle maps f, which are very far from having any linearity property. For the scope of sub-Riemannian geometry this generality is not needed; we just mention that a nice theory can be developed even for *affine* bundle maps, that is for systems whose admissible trajectories are presented as solution to the affine control system

$$\dot{x}(t) = X_0(x(t)) + \sum_{i=1}^{k} u_i(t)X_i(x(t)).$$

In this thesis however, we will touch only marginally these kind of structures.

2. Geodesics

Arguably, the most important concept in geometry is that of a *geodesic*. A geodesic γ can be described as a curve whose sufficiently small pieces are length minimizers, that is the length between sufficiently close points on γ equals their distance.

To be a bit more precise on these concepts some technical details are in order: we have already spoken of admissible paths as integral curves of the differential system (3.1), and moreover we know that every fiber U_q of the Euclidean bundle $\mathbf{U}$ is endowed with a scalar product $\mathbf{g}_q$. Then we can define the *length* of an admissible path γ as follows:

$$l(\gamma) = \int_0^T \sqrt{\mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))}dt,$$

so that the sub-Riemannian d_{SR} distance between a given pair of points x and y on M is given by

$$d_{SR}(x, y) = \inf \{l(\gamma) \mid \gamma \text{ is admissible an admissible path joining } x \text{ and } y\}.$$

In the Riemannian setting, any geodesic γ is uniquely described by the knowledge of the initial point x and the initial velocity $v \in T_xM$. This is not quite the case in sub-Riemannian geometry, since there are more geodesics than initial velocities: on the one hand, indeed, we know that between any pair of points there exists a length minimizer, which means that geodesics do fill a full neighborhood of the initial point; on the other, the space of admissible velocities is a proper subspace of T_xM.

One may then wonder what is the correct way of parametrizing the set of geodesics emanating from a point. It turns out that it is convenient to adopt a "dual" point of view. In Riemannian geometry, if we parameterize geodesics $\gamma : t \mapsto \gamma(t)$ by their length, we see that the set of their endpoints is nothing but a sphere of fixed radius, called the *wave front.* Moreover, we have that $\dot{\gamma}(t)$ is transversal to the wave front, and we can use the covector $p(t)$, orthogonal to the sphere, to have the desired dual description of geodesics. Of course this will be no longer true in sub-Riemannian geometry, where in general we cannot expect the wave front to be a sphere, and the transversality of $\dot{\gamma}(t)$ holds just for the *generic* geodesic. Nevertheless, we can define the sub-Riemannian Hamiltonian $H : T^*M \to \mathbb{R}$ by the formula $H(p, x) = \frac{1}{2}\langle p, v\rangle^2$, where v is a vector of length one which maximizes the inner product among all admissible velocities in T_xM. Any smooth function on the cotangent bundle defines a Hamiltonian vector field and, in turn, this defines a Hamiltonian flow. We call *sub-Riemannian geodesic flow* the Hamiltonian flow associated with H, which is a generalization of the Riemannian one.

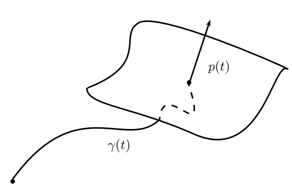


FIGURE 1. The "front wave" parametrized with the normal covector

There is a big gap between Riemannian and sub-Riemannian geometry at this point. Namely, in the first case the geodesics coincides exactly with the projections of integral curves of H, and many of their properties (e.g. their regularity) can be deduced by a quantitative study of the dynamic on T^*M. This is not possible in sub-Riemannian geometry, where there are the so called *abnormal geodesics*, which are mysteriously unrelated to the sub-Riemannian geodesic flow. Recall in fact that whenever ω is a two form defined on the cotangent space T^*M, a curve

$t \mapsto \gamma(t)$, $t \in [0,T]$, is a *characteristic* curve for ω if

$$\dot\gamma(t) \in \ker \omega_{\gamma(t)}, \ \forall\, t \in [0,T] \quad \text{(i.e. } \omega_{\gamma(t)}(\dot\gamma(t),\cdot) = 0);$$

moreover, we can find within T^*M an intrinsically defined subspace $\Delta^\perp$, called the *annihilator of the distribution* Δ, so that $\lambda \in \Delta^\perp$ if, and only if, $\langle \lambda, v\rangle = 0$ for every $v \in \Delta$. Consider now the standard symplectic form σ defined on the cotangent space: then abnormal geodesics are exactly characteristic curves of $\sigma\big|_{\Delta^\perp}$.

Abnormal geodesics comes into existence as soon as we give the sub-Riemannian structure, and have nothing to do with the metric we define on it. This lack of structure makes their nature very difficult to grasp, but at the same time very intriguing and fascinating, since new mathematical tools have to be developed to deal with them.

3. The Endpoint Map

It is probably not too far-fetched to say that, essentially, sub-Riemannian geometry is the study of the Endpoint map F. The idea behind F is incredibly simple: fix an origin x, and consider all the admissible curves which emanate from x and are defined on the interval $[0,T]$; the Endpoint map then parametrize the final points of these curves. If we let Ω to be the set of all admissible curves of a given sub-Riemannian structure, then F can be seen formally as the application $F : \Omega \to M$ such that $F(\gamma) = \gamma(T)$. Abusing of the notations, in many occasion we see the Endpoint map defined on the space of controls rather than the space of curves: this identification is made possible, up to some technicalities we don't want to discuss here, by equation (1.1), where essentially controls $u \in L^p([0,T],\mathbb{R}^k)$ play the role of coordinates, and we can identify an admissible curve with its associated control.

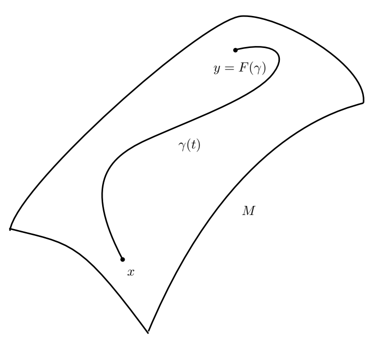


FIGURE 2. The Endpoint Map $F : \Omega \to M$

In practice, this second point of view is much more widely used, since it lends itself well to computations: it is quite often the case to meet the expression $F(u)$ instead of $F(\gamma)$, and its meaning is now clear. The Endpoint map is surjective as soon as the sub-Riemannian structure is bracket-generating, as we have already mentioned; this is indeed a reformulation of the standard Chow-Rashevskii Theorem [**AS04, Ras38**]. Most notably, the Endpoint map, together with its differential dF is weakly continuous in the L^p topology: given any sequence of controls u_n, weakly convergent in some L^p space to u, then there hold the limits

$$\lim_{n\to\infty} F(u_n) = F(u), \quad \text{and} \quad \lim_{n\to\infty} d_{u_n}F = d_uF.$$

Unfortunately, there is also some bad news: if we let $P^u_{s,t}$ denote the diffeomorphism induced on M by flowing from time t to time s along admissible paths driven by the control u, then we have the explicit formula [**ABB, AS04**]:

$$d_uF[v] = \int_0^T \sum_{i=1}^k v_i(t)(P^u_{t,T})_* X_i(x_u(t))dt, \quad v \in L^2([0,T], \mathbb{R}^k).$$

In particular Im $d_0F = \text{span}\{f_1, \dots, f_k\}$, and d_0F cannot be surjective as soon as the modulus Δ has not full dimension, which is what differs most from Riemannian geometry. In some sense, the biggest challenge of sub-Riemannian geometry is precisely to understand under which conditions Im d_uF becomes as big as the whole tangent space.

4. Structure of The Thesis

As we have seen, sub-Riemannian geometry opens the door to an enormous range of problems, going from differential geometry to analysis, from topology to dynamical systems. In this thesis we try to give original contributions in some of these aspects.

4.1. Chapter 2: Differential Geometry. As we have seen a sub-Riemannian structure is essentially given by a bundle $\mathbf{U}$ and a bundle map f, linear on the fibers of $\mathbf{U}$. A natural question would be then to classify, in some sense, what classes of structures arise in this way. This is of course an impossible task if we don't impose some restrictions on the structures we consider; at least in the simplest cases, one would nonetheless be able to understand the whole picture.

In this work we thus focus on the *local conformal classification* of sub-Riemannian structures on a three dimensional manifold M. This situation is particularly favorable since the dimension of the horizontal distribution has to be equal to two, for otherwise the bracket generating condition would be violated. In particular there is just one non admissible direction, which then has to be recovered using only one Lie Bracket. Moreover, the module Δ of the horizontal vector fields can be seen on some open set U as the kernel of a one form $\omega \in \bigwedge^1(M)$ meeting the condition

$$\omega \wedge d\omega \neq 0 \quad \text{on } U,$$

that is $\Delta\big|_U = \text{Span}\{f_1, f_2\}$.

The contact form ω is defined only up to multiplication by a non zero real valued smooth function f. A natural normalization is to impose that the restriction to the contact planes $\omega|_{\Delta_x}$ equals the area form (notice that $d\omega$, when restricted to the contact planes Δ_x, becomes a symplectic form). The remaining direction is determined intrinsically (that is, just using ω) as follows: we declare the *Reeb* vector field f_0 to satisfy the conditions $\omega(f_0) = 1$ and $d\omega(f_0, \cdot) = 0$.

Using just these informations, we can describe all sub-Riemannian structures on M by writing down the so called *structural equations*

$$\begin{cases} [f_2, f_1] = f_0 + c^1_{12} f_1 + c^2_{12} f_2, \\ [f_1, f_0] = c^1_{10} f_1 + c^2_{10} f_2, \\ [f_2, f_0] = c^1_{20} f_1 + c^2_{20} f_2. \end{cases} \tag{1.2}$$

All the information is contained in the *structural coefficients* c^k_{ij}, which are smooth functions on M; therefore in principle any classification task can be carried out starting from (1.2). But this is not quite the whole story: indeed the possible non constancy of the structural coefficients makes the analysis extremely difficult, therefore we would like to assume that they remain constant when passing between different fibers. There is indeed a very common situation when this is actually true, that is the case of a three dimensional *Lie Group* M. Using the left-multiplication map $L_y : M \to M$, $L_y(x) = yx$, we can simply define the contact plane Δ_e at the identity e, equip it with its scalar product $\mathbf{g}_e$, and then bring these objects around the whole of M by using the differential $(L_y)_*$, that is we declare:

$$\Delta_y = (L_y)_*\Delta_e \quad \text{and} \quad \mathbf{g}_y(v, w) = \mathbf{g}_e((L_y)_*^{-1}v, (L_y)_*^{-1}w).$$

Now that we have the right setting to work with, we can ask ourselves how to carry out a conformal classification. In classical geometry we look for *invariants*: by a simple dimension counting argument we can expect that in our case there will be just one of them. Indeed since Δ is two dimensional, we need six scalar equations to describe such a sub-Riemannian structure; three of them can be normalized by a change of basis, one more by a rotation of the frame $\{f_1, f_2\}$ of an angle θ and a last one by a rescaling of the metric $\mathbf{g}$. Finally, the last question concerns the flat model: it is possible to define a flat structure just in terms of being isometric to its tangent space. In Riemannian Geometry this definition is equivalent to the vanishing of the Riemann tensor, and so the problem is to define correctly what is the tangent space, to find in other words who should play the role of the Euclidean space $\mathbb{R}^3$. We will see that the right construction was introduced by Mitchell in [**Mit85**], and that a sub-Riemannian tangent space is a *Carnot Group*. In principle, for a fixed dimension of the manifold M and of the structure Δ there can be more than one tangent space, but this is not the case if M is three dimensional, where we have just the Heisenberg group. Thus, the question of flatness reduces to find which structures are locally conformally equivalent to the three dimensional Heisenberg group $\mathbf{H}_3$, and more generally, we would like to decide when two sub-Riemannian structures on a three dimensional Lie Group M are locally conformally equivalent.

This question was already addressed and answered in [**FGV95**], but the approach we propose here is more concrete and introduces a new construction, the *ambient metric* of [**FG12**], which we hope could be useful also for further questions in the future. In essence, this technique adds artificially one more dimension, and extend the given sub-Riemannian metric $\mathbf{g}$ to a Lorentz pseudoRiemannian one which scales correctly with a rescaling $e^{2\varphi}\mathbf{g}$ of the original metric. The drawback of this construction is that we live now on a circle bundle $\pi : Z \to M$, which is four dimensional. Nonetheless, we can apply the classical theory (Levi-Civita connection, Riemann Tensor, Weyl Tensor), to deduce the results in this bigger space; once we translate them back correctly on M, then our classification problem is completely solved.

A similar line of research was already carried out by Agrachev and Barilari in [**AB12**] where they performed a *metric classification* of sub-Riemannian structures on three dimensional Lie Groups. What is surprising is that, apart from a minor difference in what concerns flat structures, the conformal classification essentially coincides with the metric one, that is the situation is still very rigid.

4.2. Chapter 3: Topology. In this second work we switch to topological issues. We have already explained that the bundle map $f : \mathbf{U} \to TM$ is linear on the fibers of $\mathbf{U}$. In a local trivialization, the equations describing admissible horizontal curves are comprised in the following system of differential equations:

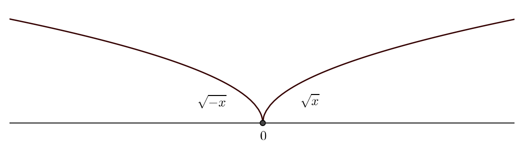
$$\dot{x}(t) = \sum_{i=1}^{k} u_i(t) X_i(x(t)); \tag{1.3}$$

moreover, we commonly assume that the control $u : [0, T] \to \mathbb{R}^k$ belongs to some L^p space of functions. The first natural question would be then to decide whether all these choices are equivalent: this is not so clear and obvious, since it is a well-known result in the field [**BH93**] that if we work within the class of essentially bounded measurable controls, then there are admissible curves which are *rigid*, that is isolated up to reparametrizations (in the corresponding $W^{1,\infty}$-topology of curves). This is of course in contrast with the classical Riemannian setting, where all these distinctions essentially irrelevant since the Endpoint map is a submersion.

The major outcome of working with a "convenient" L^p space concerns the possibility of carrying out analysis on the sub-Riemannian structure. Indeed, as soon as p can be chosen greater than one, then one can define the *p-Energy functional* $J_p(u) = \|u\|_{L^p}^p$, and try to use all the tools coming from the standard nonlinear analysis.

The primary focus of this part is the comprehension of some topological properties on the space of horizontal curves Ω: we hope to deduce some of them using the information on the topology of M. The situation is indeed nice in the sub-Riemannian case: it turns out, in fact, that the Endpoint map $F : \Omega \to M$ is a *Hurewicz fibration* for every $1 \leq p < \infty$, which means that we can read the topological properties of M in a more treatable L^p space of controls; it is, on the other hand, more delicate in the case of *affine* control systems, with a free drift term $X_0(x(t))$ appearing in front of the right-hand side of (1.3). Still there is some good news: it is indeed again possible to choose some exponent $p > 1$ so that the Endpoint map remains a Hurewicz Fibration whenever the controls are elements of this L^p space. The exponent p depends on the distribution Δ, more precisely on its step, that is how many Lie brackets of vector fields we need to consider to be sure to span the tangent space TM at every point. In particular, this constraint prevents the possibility of choosing $p = 2$, and makes it impossible to work with Hilbert spaces, so that the resulting analysis becomes slightly more complicated. But again, we will not focus on the affine case in this thesis.

Fix some point $y \in M$; then we also define $\Omega(y) \subset \Omega$ as the subset of all horizontal curves ending at y. The biggest issue of sub-Riemannian geometry concerns the non regularity of the Endpoint map: the presence of abnormal controls possibly joining the origin x to y makes the preimage $\Omega(y) = F^{-1}(y)$ a terrible set from the differentiable viewpoint, with wild singularities. Our results recognize that this is indeed not the case if we look at it from the homotopy point of view. Indeed the space $\Omega(y)$ has the homotopy type of a CW-complex, independently on whether y is regular or critical; in particular any curve $\gamma \in \Omega(y)$ possesses a contractible neighborhood. The simplest (and closest) example which describes this situation is depicted by joining the two branches of the square root on the real line: in zero we have a "bad" singularity from the



differentiable point of view, nevertheless the homotopy type of this graph is the same as that of $\mathbb{R}$, as can be seen by ideally stretching the ends on both sides.

As a final application of the machinery developed so far, we provide an existence result of sub-Riemannian geodesics on a compact manifold M. This is indeed the counterpart of the theorem of Serre [**Ser51**] on the existence of infinitely many Riemannian geodesics between two given points. We stress once more that we cannot drop the assumption of $\Omega(y)$ to be a regular fiber. The existence of at least one regular value for the endpoint map F seems not very demanding; still we have no idea whether this hypothesis is indeed satisfied by every sub-Riemannian structure. In fact, such a result can be seen as a possible formulation of an analogue of the Sard lemma for the Endpoint map F (for which the standard, i.e. finite dimensional, theory is not applicable).

4.3. Chapter 4: Analysis. We have shown how, from a homotopical point of view, any fiber $\Omega(y)$ of the Endpoint map is nice, regardless of y being either a regular or a critical value. However, if we bring an *Energy* functional J into the picture, then it is no longer true in general that along the deformation process we can preserve the Lebesgue sets $\{J \leq E\}$. In this third and last work we want to show that there is in fact a large class of sub-Riemannian structures in which this is actually possible, and we explain how to carry out the deformation by developing a sort of a sub-Riemannian Morse Theory, which, inexorably, has to take care of abnormal curves. Moreover, since we have proved that all the L^p topologies are essentially equivalent in the sub-Riemannian case, we will always work with controls living in the Hilbert space $L^2([0,T],\mathbb{R}^k)$.

To understand what assumptions we have to use, a brief digression on critical points is due. Let us consider the so-called *extended Endpoint Map*, $\Phi = (F, J) : \Omega \to M \times \mathbb{R}$, which associates to a given horizontal curve γ its final point and its Energy. To simplify the exposition we assume to work in a coordinate chart; this means that we can locally identify the horizontal curves with the control u appearing in (1.3), and work with the Hilbert space $L^2([0,T],\mathbb{R}^k)$. Then the optimality condition translates into the Lagrange multipliers' rule for Φ, and implies the existence of a covector $\overline{\lambda} = (\lambda_F, \lambda_J)$ annihilating the differential $d_u\Phi$. If $\lambda_J \neq 0$, then the differential d_uF is surjective, while if $\lambda_J = 0$, then d_uF is not a submersion and the fiber $\Omega(y)$ may be wildly singular.

Nonetheless, there is still some hope: first of all, we call the corank of the control u the dimension of vector space $\{\xi \in T^*_yM \mid \xi d_uF = 0\}$. In particular the corank of u equals the codimension of d_uF, and counts how many directions are "missed" at the first order by the differential; the best we can expect for a critical control u is that its corank is equal to one. On the other hand there exists a *second order* necessary optimality condition for the Endpoint map F, called the *Goh condition*, and whose violation implies that the Hessian of F is positive definite on an infinite dimensional subspace H^+ and is negative definite on another infinite dimensional subspace H^-. Both the corank one and the absence of curves satisfying the Goh condition are in some sense "generic" when the dimension of Δ is greater than or equal to three [**CJT06, CJT08**]; any small perturbation of a given sub-Riemannian structure satisfies them both.

Around critical points (which from now one are assumed to be of corank one and not Goh) a deformation using the *pseudo gradient* technique is not possible, since the fiber is not a smooth manifold and it is not even clear what a gradient flow should be! But we can do something weaker, yet as effective: we can construct a *cross section*, that is a map $\alpha : \mathbb{R}^{k+1} \to L^2([0,T],\mathbb{R}^k)$ extending continuously the initial data u, and try to solve the system of equations defining the deformation using the implicit function theorem. In practice we add some "extra" dimensions on which we perform the deformation, embedding a copy of $M \times \mathbb{R} \simeq \mathbb{R}^{k+1}$ into the space of controls L^2. Of course this embedding cannot be smooth, but the good news is that it is Lipschitzian: the only direction missing can be parametrized by means of two different directions joined at 0, whose existence is granted by the non Goh property, which permits to find a direction on which the Hessian is positive definite and one on which it is instead negative.

Outside the set of critical points we use the gradient flow of the Energy J. There is some weak topology going on behind all this: if we work inside some big L^2-ball $B = \{J \leq E\}$, which is weakly compact, the set of all critical controls inside B is weakly compact, and the implicit function just described works on a full weak neighborhood A of the abnormal set. This means that if we consider the complement R of A, then R contains just regular controls; moreover it is *strongly* separated from the abnormal set, and we can effectively construct a gradient flow which

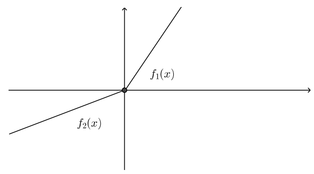


FIGURE 3. The function obtained composing $f_1(x)$ and $f_2(x)$ is invertible at the origin even though it does not satisfy the classical implicit function theorem

remains separated from critical points. This concept of weak compactness comes in handy again to globalize estimates on the whole ball B; in some sense it can be seen as a weak counterpart of the Palais-Smale condition on regular controls.

If we glue together these two constructions, we arrive at a sub-Riemannian deformation lemma: assume there are no *normal geodesics* (that is for which the Lagrange multiplier λ_J is different from zero) between two Energy sublevel sets $\{J \leq E_1\}$ and $\{J \leq E_2\}$. Then for every $\varepsilon > 0$, every compact manifold X and any continuous map

$$h : X \to \Omega(y) \cap \{J \leq E_2\}$$

there exists an homotopy $h_t : X \to \Omega(y) \cap \{J \leq E_2\}$ so that $h_0 = h$ and $h_1(X) \subset \Omega(y) \cap \{J \leq E_1 + \varepsilon\}$. A few comments on this result are in order: first of all the condition "for every $\varepsilon > 0$" is unavoidable and relates to the impossibility of coming too close to critical energy levels, since the more we near to critical levels, the more topology is influenced by the presence of singularities. Moreover, we want to stress that we don't make any strong deformation retract of sublevel sets: we are just able to deform a much weaker class of objects, namely continuous maps (but more generally singular chains representing homology classes). Surprisingly, this is still enough to predict, even in the Riemannian case, the existence of critical points for the Energy.

We finally apply all this machinery to give a sub-Riemannian Min-Max principle and, lastly, a true counterpart of the Serre theorem [**Ser51**], even in the presence of singular curves (at least for the generic sub-Riemannian structure). The Min-Max however just grasps normal geodesics: the nature of abnormals remains slippery. Maybe we can detect those of them which are either Goh abnormals or not of corank one, because they should be perceived as a change in the topology along the flow of the energy J. But how to catch what we have called *soft* abnormals remains up to now very unclear, since it seems that they do not affect topology and, in some sense, they remain "invisible".

Acknowledgements: I would like to thank my supervisor Andrei Agrachev for having suggested these problems and for having introduced me to the wonderful topic of Control Theory; his dedication and enthusiasm for mathematics have fascinated me since the beginning. More importantly, he taught me that there is no easy way to overcome a problem, and that sometimes, if it seems too difficult, you just have to look at it with a different perspective; this a lesson that I will use forever in my life. I'm also grateful to Antonio Lerario, who shared with me many parts of this work: he taught me the importance of hard working, and that technique is fundamental to polish raw ideas that otherwise nobody would read.

CHAPTER 2

Local Conformal Equivalence of sub-Riemannian Three Dimensional Structures on Lie Groups

1. Introduction

A three dimensional sub-Riemannian manifold is a triplet $(M, \Delta, \mathbf{g})$ where

a) M is a smooth connected three dimensional manifold,
b) Δ is a smooth rank two vector sub-bundle of TM
c) $\mathbf{g}_q$ is an Euclidean metric on Δ_q, which varies smoothly with respect to the base point $q \in M$.

If M is a Lie group and both the metric $\mathbf{g}$ and Δ are preserved by the left translations defined on M, then we say that $(M, \Delta, \mathbf{g})$ is *left invariant*.

In what follows, we assume that Δ is *bracket generating*, that is for every $q \in M$ the Lie algebra of horizontal vector fields evaluated at q equals T_qM. Under this assumption, M is endowed with a natural structure of metric space, where the distance is the so called *Carnot-Carathéodory* distance

$$d(p,q) \doteq \inf \left\{ \int_0^T \sqrt{\mathbf{g}_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))}dt,\ \gamma : [0,T] \to M \text{ is a Lipschitz curve,} \right.$$
$$\left. \gamma(0) = p,\ \gamma(T) = q,\ \dot\gamma(t) \in \Delta_{\gamma(t)},\ \text{a.e. } t \in [0,T] \right\}.$$

As a consequence of the bracket generating condition, d is always finite and continuous, and induces on M its original topology [**Ras38, AS04**].

A sub-Riemannian manifold is said to be *contact* if Δ can be locally described as the kernel of a contact differential one form ω, i.e. satisfying $d\omega \wedge \omega \neq 0$.

Three dimensional sub-Riemannian contact manifolds are the simplest examples of sub-Riemannian geometries; they possess two basic functional invariants χ and κ which appear in the expansion of the Exponential map [**Ag96**]. It is natural to expect, at least heuristically, why there must be two such invariants: locally the distribution Δ is defined by a pair of orthonormal vector fields in $\mathbb{R}^3$, that is by six scalar equations. One of them can be normalized by a smooth rotation of the frame within its linear hull, while three more are normalized through a smooth change of variables. What remains are indeed two scalar functions.

By a well known classification result [**Jac62, Kir08**], the analysis can be restricted to the Lie algebras of the following Lie groups

i) $\mathbf{H}_3$, the Heisenberg Group,
ii) $A(\mathbb{R}) \oplus \mathbb{R}$, where $A(\mathbb{R})$ is the group of orientation-preserving affine maps on $\mathbb{R}$,
iii) $SOLV^+$ and $SOLV^-$ are Lie groups whose Lie algebra is solvable and has a two dimensional square,
iv) $SE(2)$ and $SH(2)$ are the groups of motion of the Euclidean and the Hyperbolic plane respectively,

v) The simple Lie groups $SL(2)$ and $SU(2)$.

Moreover, it can be shown that in each of these cases but one all left invariant bracket generating distributions are equivalent by automorphisms of the Lie algebra. The only case where there exist two nonequivalent distributions is the Lie algebra $\mathfrak{sl}(2)$. More precisely a two dimensional subspace of $\mathfrak{sl}(2)$ is called *elliptic* (resp. *hyperbolic*) if the restriction of the Killing form on this subspace is sign-definite (resp. sign-indefinite). Accordingly, the notation $SL_e(2)$ and $SL_h(2)$ are used to specify on which subspace the sub-Riemannian structure on $SL(2)$ is defined.

In [**AB12**] it is proved that left invariant sub-Riemannian structures on three dimensional Lie groups are classified by local isometries, i.e. smooth maps preserving the sub-Riemannian metric on $(M, \Delta, \mathbf{g})$. The classification is as in figure 1, where a structure is identified by a point (χ, κ) and two distinct points represent non locally isometric structures. In particular we have:

THEOREM 1. *Let $(M, \Delta, \mathbf{g})$ be a three dimensional left invariant contact sub-Riemannian manifold.*

- *If $\chi = \kappa = 0$ then the manifold is locally isometric to the Heisenberg Group,*
- *If $\chi^2 + \kappa^2 > 0$ then there exist no more than three non isometric normalized structures having these invariants,*
- *If $\chi \neq 0$ or $\chi = 0$ and $\kappa \geq 0$ then the structures are locally isometric if, and only if their Lie algebras are isomorphic.*

As a byproduct of this classification, it turns out that there exist non isomorphic Lie groups with locally isometric sub-Riemannian structures, as it is the case of $A(\mathbb{R}) \oplus \mathbb{R}$ and $SL(2)$ with elliptic type killing metric, with the sub-Riemannian structure defined by $\chi = 0$, $\kappa < 0$.

The aim of this paper is to carry out a similar classification task if we assume that we can act on a three dimensional contact manifold $(M, \Delta, \mathbf{g})$ using the sub-Riemannian conformal group, that is we allow all those smooth maps which preserve angles on the distribution Δ, and not necessarily distances (i.e. we allow multiplication of the metric by smooth functions $e^{2\varphi}$, $\varphi \in C^\infty(M)$); notice that this additional degree of freedom permits to normalize one more equation among those that define $(M, \Delta, \mathbf{g})$, hence it is natural to expect the existence of just one functional conformal invariant.

Some of the results presented in this work were already investigated in [**FGV95**]: in particular they also present the conformal invariant and pay special attention to flat structures, although the exposition follows arguments which are different from those presented here.

1.1. Structure of the Paper. We begin with a brief review of the basic terminology of the sub-Riemannian geometry, and immediately in the section 3 we introduce the main technical tool used in the sequel: the so called *Fefferman ambient metric.*

Using the same approach of [**CM08**] we construct on $(M, \Delta, \mathbf{g})$ its associated Fefferman metric g_F, which is a Lorentz pseudometric living on a circle bundle over M, and which provides the right conformal setting to work with since the beginning. In particular, the unique conformal invariant associated to $(M, \Delta, \mathbf{g})$ is presented explicitly in section 4 as the ratio of the only two non zero entries of the Weyl tensor associated to g_F.

Section 5 is devoted to the investigation of the locally conformally flat left invariant sub-Riemannian structures, where the definition of flatness we give is as follows:

DEFINITION 3. *We say that a contact three dimensional sub-Riemannian manifold $(M, \Delta, \mathbf{g})$ is locally conformally flat if, fixed any point $q \in M$ there exists a neighborhood U of q and a function $\varphi : U \to \mathbb{R}$ so that the rescaled sub-Riemannian structure $\left(M, \Delta, e^{2\varphi}\mathbf{g}\right)$ becomes locally isometric in U to the Heisenberg Group.*

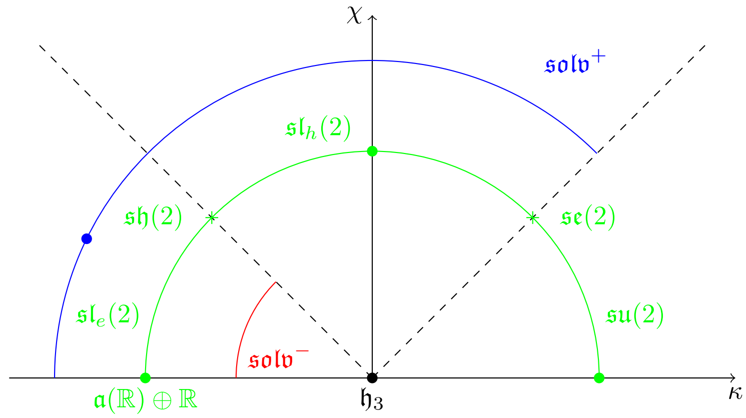


FIGURE 1. Conformal classification of 3D left-invariant structures. Points on different circles denotes different classes of equivalence of structures under the normalization condition $\chi^2 + \kappa^2 = 1$. Circled structures are locally conformally flat. Unimodular structures are those in the middle circle.

If φ can be defined globally on M we say that the manifold is conformally flat.

By a direct computation based upon the realization of explicit models for such structures, we will show that the local conformal flatness of the Fefferman metric associated to $(M, \Delta, \mathbf{g})$ (i.e. the vanishing of its Weyl tensor) is equivalent to the local conformal flatness of the contact sub-Riemannian manifold itself.

In section 6 we study the *chains' equation* associated to g_F [**FA86, Lee86, FG12**], that is the zero level set for the Hamiltonian flow generated by this pseudometric. Chains, considered as a set of unparametrized curves, are invariant under conformal rescalings of $(M, \Delta, \mathbf{g})$. We will explicitly integrate their flow and, consequently, compute the tangent space to the chains' set. It will turn to be dependent on the metric invariants χ and κ and, as any conformal map must preserve it, we will be able to deduce that whenever g_F is not locally conformally flat, then the conformal classification of $(M, \Delta, \mathbf{g})$ coincides with the metric one.

THEOREM 2. *Let $(M, \Delta, \mathbf{g})$ be a left invariant three dimensional sub-Riemannian contact manifold. If the Fefferman metric associated to $(M, \Delta, \mathbf{g})$ is not locally conformally flat, then its conformal classification is uniquely determined by the pair χ and κ; in particular it coincides with the metric one.*

Finally, section 7 is devoted to the study of the conformal group $Conf(\mathbf{H}_3)$, of the three dimensional Heisenberg Group. We start with its explicit computation using the Hamiltonian viewpoint of [**AS04**]

DEFINITION 4. *Given $X \in \mathrm{Vec}(M)$, we say that X is a sub-Riemannian conformal symmetry if the flow generated by X is a conformal map, i.e. preserves the conformal class of the sub-Riemannian metric $\mathbf{g}$.*

Our calculations generalize previous results like those in [**FMP99**] where the group of isometries of $\mathbf{H}_3$ was described.

We also show that the Lie Algebra of conformal symmetries, $\mathfrak{conf}(\mathbf{H}_3)$, is in fact isomorphic to $\mathfrak{su}(2,1)$. This proposition combines our calculations with the purely algebraic Tanaka's prolongation procedure [**Tan79, Zel09**]. Since locally conformally equivalent structures have isomorphic conformal groups, we deduce that the conformal group of a locally conformally flat manifold $(M, \Delta, \mathbf{g})$ is isomorphic to $SU(2,1)$. This in turn completes the task which motivated the paper.

The main result of the last section is the following:

THEOREM 3. *Let $(M, \Delta, \mathbf{g})$ be a left invariant three dimensional contact manifold. Then $(M, \Delta, \mathbf{g})$ is locally conformally flat if, and only if its associated Fefferman metric is locally conformally flat; if this is the case, then*

$$\mathfrak{conf}\,(M, \Delta, \mathbf{g}) \cong \mathfrak{su}(2,1).$$

2. Three Dimensional sub-Riemannian Structures

2.1. Notations and Basic Definitions.

DEFINITION 5. *A three dimensional* sub-Riemannian *structure is a triplet $(M, \Delta, \mathbf{g})$, where*

a) *M is a three dimensional smooth connected manifold,*
b) *Δ is a smooth distribution in TM of constant rank $k = 2$, that is a smooth map which associates to any $q \in M$ a plane $\Delta_q \subset T_qM$,*
c) *$\mathbf{g}_q$ is an Euclidean metric on Δ_q, which varies smoothly with respect to the base point.*

We distinguish among the set of all smooth vector fields of M, the subspace $\mathcal{H}$ of the *horizontal* vector fields defined by:

$$\mathcal{H} = \{f \in \mathrm{Vec}(M) \mid f(q) \in \Delta_q, \ \forall q \in M\}.$$

A sub-Riemannian structure is then said to be *bracket generating* if $\mathrm{Lie}\,(\mathcal{H})_q = T_qM$ for any $q \in M$.

An absolutely continuous curve $\gamma : [0,T] \to M$ is said to be *admissible* if its derivative is a.e. horizontal, that is $\dot\gamma(t) \in \Delta_{\gamma(t)}$ a.e. $t \in [0,T]$. For an admissible curve γ, its length is defined by

$$l(\gamma) = \int_0^T \sqrt{\mathbf{g}_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))}dt.$$

If we let Ω_{xy} to be the set of all the admissible curves joining two given points x and y, their *Carnot-Carathéodory* distance is then computed as the infimum of the length of all the curves comprised in Ω_{xy}, that is $d(p,q) = \inf_{\gamma \in \Omega_{xy}} l(\gamma)$. Under the bracket generating assumption, the classical Chow-Rashevski theorem [**Ras38, AS04**] ensures that for any pair of points x and y there holds $d(x,y) < +\infty$; moreover it is well known that d defines a metric which induces on M its standard manifold topology.

DEFINITION 6 (Contact Structure). *A sub-Riemannian structure $(M, \Delta, \mathbf{g})$ is said to be a* contact *structure if there exists a one form ω such that i) Δ is locally a contact structure, that is $\Delta = \mathrm{Ker}(\omega)$, and ii) ω satisfies the* contact condition $d\omega \wedge \omega \neq 0$.

REMARK 1. A contact structure is forced to be bracket generating.

If, in addition, M is also a Lie Group, there is a natural way to construct a contact structure: indeed it is sufficient to fix a plane on the Lie Algebra $\mathfrak{m}$ and an inner product on it. The construction is carried onto all the other points using the left multiplication map $L_s : M \to M$, $L_s(g) = sg$, that is declaring that $\Delta_{sh} = L_{s^*}\Delta_h$ and $\mathbf{g}_h(v,w) = \mathbf{g}_{sh}(L_{s^*}v, L_{s^*}w)$ for any s and h in M. Such contact structures will be called *left invariant*.

The contact condition implies that $d\omega$ is a symplectic form when restricted on the contact planes Δ_q, and it is immediate to observe that any contact form is defined up to multiplication by a nonzero real-valued function f; a natural choice of f is given by the requirement that $d\omega|_{\Delta_q}$

coincides with the area form on Δ_q (the ambiguity on the sign can be avoided assuming that an orientation is assigned on the contact planes).

Fix a point $\overline{q} \in M$. It is always possible to find a neighborhood U of $\overline{q}$ in M where Δ can be described in terms of an orthonormal frame $\{f_1, f_2\}$, that is $\Delta_q = \text{span}\{f_1(q), f_2(q)\}$ for any $q \in U$. In the left invariant case the frame can be defined globally and the above equality holds everywhere on the manifold.

Naturally associated to a contact structure there is the so called *Reeb vector field* f_0, uniquely characterized via the requirements $\iota_{f_0}\omega = 1$ and $\iota_{f_0}d\omega = 0$. The contact condition implies that $\{f_0, f_1, f_2\}$ is a local basis for the tangent space TM (global in the left invariant case).

The sub-Riemannian Hamiltonian is a smooth function $h \in C^\infty(T^*M)$ defined by:

$$h(\lambda) = \max_{u \in \Delta_q} \left\{ \langle \lambda, u \rangle - \frac{1}{2}|u|^2 \right\}, \quad \lambda \in T^*M, q = \pi(\lambda)^1.$$

If we consider the linear on fibers functions $h_i \in C^\infty(T^*M)$ defined, for $i = 0, 1, 2$, as

$$h_i(\lambda) = \langle \lambda, f_i(q) \rangle, \quad \lambda \in T^*M, q = \pi(\lambda),$$

the sub-Riemannian Hamiltonian can be rewritten as

$$h(\lambda) = \frac{1}{2}\left(h_1^2(\lambda) + h_2^2(\lambda)\right). \tag{2.1}$$

All the information about the sub-Riemannian structure is encoded in this Hamiltonian: it is indeed independent on the chosen orthonormal frame, and from (2.1) we may reconstruct both the Riemannian metric $\mathbf{g}_q$ on Δ_q and the annihilator $\Delta_q^\perp \subset T_q^*M$ which is nothing but the kernel of h restricted to the fiber T_q^*M.

2.2. Metric Invariants of 3D sub-Riemannian Structures. Let $\{\nu^0, \nu^1, \nu^2\}$ be the coframe dual to $\{f_0, f_1, f_2\}$; since this is a basis of one forms on T^*M, we have

$$\begin{cases} d\nu^0 = \nu^1 \wedge \nu^2, \\ d\nu^1 = c_{10}^1 \nu^0 \wedge \nu^1 + c_{20}^1 \nu^0 \wedge \nu^2 + c_{12}^1 \nu^1 \wedge \nu^2, \\ d\nu^2 = c_{10}^2 \nu^0 \wedge \nu^1 + c_{20}^2 \nu^0 \wedge \nu^2 + c_{12}^2 \nu^1 \wedge \nu^2, \end{cases} \tag{2.2}$$

where the c_{ij}^k are smooth functions on M. Cartan's Formula $d\omega(X,Y) = X\omega(Y) - Y\omega(X) - \omega([X,Y])$ yields the dual version of (2.2), the so called *structural equations* for $(M, \Delta, \mathbf{g})$, which read

$$\begin{cases} [f_2, f_1] = f_0 + c_{12}^1 f_1 + c_{12}^2 f_2, \\ [f_1, f_0] = c_{10}^1 f_1 + c_{10}^2 f_2, \\ [f_2, f_0] = c_{20}^1 f_1 + c_{20}^2 f_2. \end{cases} \tag{2.3}$$

Contact sub-Riemannian three dimensional structures possess two metric scalar invariants χ and κ, which are independent on the orthonormal frame chosen for Δ. Consider first the Poisson bracket $\{h, h_0\}_q$: it can be shown that this is a traceless, quadratic on fibers map on the cotangent space T_q^*M. The first invariant

$$\chi(q) = \sqrt{-\det\{h, h_0\}_q} = \sqrt{-c_{10}^1 c_{20}^2 + \left(\frac{c_{10}^2 + c_{20}^1}{2}\right)^2}$$

[1]Here $\langle \cdot, \cdot \rangle$ stands for the usual duality between the tangent and the cotangent space.

is zero if and only if the the one-parametric flow generated by f_0 is a sub-Riemannian isometry; the second invariant

$$\kappa(q) = f_2(c_{12}^1) - f_1(c_{12}^2) - (c_{12}^1)^2 - (c_{12}^2)^2 + \frac{1}{2}(c_{10}^2 - c_{20}^1) \tag{2.4}$$

is the analogue of the curvature for a two-dimensional Riemannian surface, and appears as a term in the asymptotic expansion of the cut locus of our structure; they are found and studied in greater detail in [**Ag96**].

Left invariant three dimensional contact structures can be classified through these invariants, as it is done in [**AB12**]. For future purposes we recall how it is possible to choose a *canonical frame*; to do so, we have to distinguish between two cases

PROPOSITION 4. *Let* $(M, \Delta, \mathbf{g})$ *be a three dimensional contact sub-Riemannian structure and let* $q \in M$*. If* $\chi(q) \neq 0$*, then there exists a local frame for which there holds the equality*

$$\{h, h_0\} = 2\chi h_1 h_2.$$

In particular, in the Lie Group case endowed with a left invariant structure, there exists a unique (up to sign) canonical frame $\{f_0, f_1, f_2\}$ *so that the structural equations* (2.3) *become*

$$\begin{cases} [f_2, f_1] = f_0 + c_{12}^1 f_1 + c_{12}^2 f_2, \\ [f_1, f_0] = c_{10}^2 f_2, \\ [f_2, f_0] = c_{20}^1 f_1. \end{cases}$$

PROPOSITION 5. *Let* $(M, \Delta, \mathbf{g})$ *be a three dimensional contact sub-Riemannian structure, and let* $q \in M$ *be such that* $\chi = 0$ *in a neighborhood* U *of* q*. Then there exists a rotation of the frame associated to* $(M, \Delta, \mathbf{g})$ *such that the structural equations* (2.3) *on* U *take the form (cfr. the defintion of* κ *in* (2.4)*)*

$$\begin{cases} [f_2, f_1] = f_0, \\ [f_1, f_0] = \kappa f_2, \\ [f_2, f_0] = -\kappa f_1. \end{cases}$$

REMARK 2. A three dimensional contact sub-Riemannian structure can be seen as an $SO(2)$-principal bundle over the base manifold M. We naturally define the $SO(2)$-action on the fiber over the point q by rotating the orthonormal frame which spans Δ_q. Under this point of view, a canonical frame permits to select a preferred section within this bundle.

Moreover, if we write the structural equations as indicated by the canonical frame, it turns out that

$$\chi = \frac{|c_{10}^2 + c_{20}^1|}{2};$$

by possibly changing the orientation on the contact planes, we can always assume that $c_{10}^2 + c_{20}^1 \geq 0$, so that the absolute value can be dropped in the previous formula. On the other hand, it is easy to see that κ remains unaffected by this operation.

REMARK 3. We close this introductory exposition on contact structures by recalling the possibility to define an intrinsic volume form. This is the so called *Popp's volume* and in terms of a (local) coframe it may be expressed as

$$\text{vol} = \nu^0 \wedge \nu^1 \wedge \nu^2.$$

For further details we refer the reader to [**BR13**].

3. The Fefferman Metric

3.1. CR structures. Let M be an n-dimensional manifold. A *Cauchy-Riemann* (CR in short) structure modeled on M is the datum of a subbundle L of the complexified tangent bundle $\mathbb{C}TM = \mathbb{C} \otimes_{\mathbb{R}} TM$ which is integrable, i.e. $[L, L] \subseteq L$, and which is nowhere real, that is meets the condition $\bar{L} \cap L = \{0\}$. The codimension of a CR structure is defined as $k = n - (2 \dim L)$: in particular, if $k = 1$, we say that our structure is of *hypersurface type*, and it is immediate to observe that if the underlying manifold M is three dimensional, necessarily any CR structure on it must be of hypersurface type.

A *strictly pseudoconvex* CR structure on a three dimensional manifold M consists of a contact structure Δ defined on M, together with an almost complex map J defined on the contact planes $\{\Delta_q\}_{q \in M}$. We will adopt in the sequel an alternative local characterization of such structures, which better suits our purposes, namely

Definition 7. *A strictly pseudoconvex CR structure on a three dimensional manifold M consists of an oriented contact structure Δ defined on M together with a conformal class of metrics defined on the contact planes $\{\Delta_q\}_{q \in M}$.*

To see the why these two definitions are the same we proceed as follows: let us choose a contact form θ satisfying locally the equality $\Delta = \ker \theta$. Given the almost complex map J, we define the *Levi form* as:

$$L_\theta(v, w) = d\theta(v, Jw), \quad \forall v, w \in \Delta. \tag{2.5}$$

The condition of non degeneracy of $d\theta$ restricted to Δ implies that the Levi form is either positive or negative definite. We insist that, up to work with $-\theta$ instead of θ, the choice of the contact form and the almost complex map J can be made so that the Levi form is positive definite, which is the same as to require that the orientation induced on the contact planes Δ_q by θ and J coincide. We notice that if we multiply the contact form by a factor $e^{2\varphi}$, $\varphi : M \to \mathbb{R}$, then the Levi form rescales as

$$L_{e^{2\varphi}\theta} = e^{2\varphi} L_\theta,$$

that is, the definition of the Levi form depends only on the conformal class of the metric and not on the (oriented) contact form θ.

To go the other way round, we build the almost complex map J by declaring how it behaves on a frame, that is for a representative metric $\mathbf{g}$ in the conformal class and two $\mathbf{g}$-orthonormal vector fields f_1 and f_2, we declare that

$$J(f_1) = f_2, \ J(f_2) = -f_1. \tag{2.6}$$

The definition of a CR structure depends therefore either on the choice of an oriented contact form θ or on the definition of an almost complex map J on the contact planes; these choices however are unique up to a conformal factor, and, as we have seen, are linked one to the other via the Levi form as in (2.5).

Finally, one more way to see a strictly pseudoconvex CR structure on a three dimensional manifold M is the datum of a complex line field, that is a rank one subbundle of $\mathbb{C}TM$, which is nowhere real.

3.2. Holomorphic and anti-Holomorphic forms. Extend the almost complex map J to the whole of $\mathbb{C}TM$ by complex linearity. We define as *holomorphic* directions those which correspond to eigenvectors relative to the imaginary eigenvalue i; analogously, the anti-holomorphic directions constitute the $-i$-eigenspace. In three dimensions there is exactly one holomorphic and one anti-holomorphic direction, and their generators are easily seen to be by $f_1 - if_2$ and $f_1 + if_2$ respectively.

Holomorphic and anti-holomorphic forms are declared dually on $\mathbb{C}T^*M$: a one form λ is holomorphic if and only if it annihilates the anti-holomorphic direction, i.e. if and only if $\lambda(f_1 + if_2) = 0$. Anti-holomorphic forms are defined in the obvious way. The almost complex map J induces therefore a splitting in the space of complex differential forms; we may speak of a (p, q) form, as to indicate a complex form which has degree p holomorphic part and degree q anti-holomorphic part. In the present situation we may choose an orthonormal frame for Δ, say $\{f_1, f_2\}$ and build J as indicated in (2.6). Then the complex holomorphic one forms are spanned over $\mathbb{C}$ by ν^0 and $\nu^1 + i\nu^2$.

We extend the Levi form to the whole complexified tangent space $\mathbb{C}TM$ by insisting that

$$L_{\nu^0}(f_0, v) = 0,\ \forall v \in TM;$$

by an abuse of notation we will continue to write L_{ν^0} for this extended form.

3.3. The Ambient Metric. The ambient metric has been introduced by Fefferman and Graham in [**FG12**]. The approach we will follow in this paper is however due to Farris [**FA86**] and Lee [**Lee86**] who developed an intrinsic construction which works also for abstract (i.e. not embedded) CR manifolds; we also refer the reader to [**CM08**], whose exposition suits surprisingly well our necessities.

Let $\pi : Z \to M$ be a circle bundle over M and let σ be a one form on Z such that it is non zero on vertical vectors, i.e. $\sigma\big|_{\ker(\pi)}$ is non zero. We try to build a Lorentz pseudometric $g = g_{\nu^0}$ on Z in such a way that

$$g_{\nu^0} = \pi^* L_{\nu^0} + 4\left(\pi^*\nu^0\right) \odot \sigma \tag{2.7}$$

and so that a conformal rescaling of the contact form $\nu^0 \mapsto e^{2\varphi}\nu^0$, with $\varphi \in C^\infty(M)$, induces a conformal rescaling of (2.7) as:

$$g_{e^{2\varphi}\nu^0} = e^{2\varphi} g_{\nu^0}.$$

Remark 4. The convention here is that $\odot$ denotes the symmetric product of one forms; namely, for any pair of one forms ν and η, $\nu \odot \eta = \frac{1}{2}(\nu \otimes \eta + \eta \otimes \nu)$.

Holomorphic $(2, 0)$-forms on M, which are spanned over $\mathbb{C}$ by $\nu^0 \wedge (\nu^1 + i\nu^2)$, if considered pointwise form a complex line bundle which we will denote by K; if we forget about the zero section, this is commonly referred to as the *canonical bundle* K^*. We begin by defining Z as the ray projectivization of K^*, that is

$$Z = K^*/\mathbb{R}^+.$$

We now want to build the one form σ on Z in a canonical way, i.e. depending just on the choice of the contact form ν^0.

To this end, let us look for holomorphic $(2, 0)$ solutions to the so called *volume normalization equation*

$$i\nu^0 \wedge \iota_{f_0}\psi \wedge \iota_{f_0}\overline{\psi} = \nu^0 \wedge d\nu^0 = \nu^0 \wedge \nu^1 \wedge \nu^2. \tag{2.8}$$

We recognize Popp's volume in the right hand side of the equation, while in the left hand side the two form ψ is the unknown to be determined. Notice that (2.8) depends only on the conformal class of metrics, and that it is quadratic in ψ, which means that a conformal multiplication

$\psi \mapsto f\psi$, $f > 0$, implies that the volume normalization equation rescales by a factor $|f|^2$. It is then evident that the solution to (2.8) is defined just up to a complex unit multiple

$$\psi \mapsto e^{i\gamma}\psi.$$

To put it in other words, we are defining a section

$$s_{\nu^0} : Z \to K^*,$$

since once we have decided the complex phase of ψ, the volume normalization equation uniquely determines the real scale factor.

Fix any solution to the volume normalization equation, i.e. a smoothly varying family of solutions $\psi_0 : M \to K$. This determines a trivialization of Z, since once ψ_0 is chosen, we may identify any point $z \in Z$ as a pair (q, γ), with $q = \pi(z)$, and

$$s_{\nu^0}(z) = e^{i\gamma}\psi_0(\pi(z)).$$

On $Z \cong M \times S^1$ we may define the following two form:

$$\zeta(q, \gamma) = e^{i\gamma}\psi_0(q); \tag{2.9}$$

what we have to check is that it depends just on the choice of the contact form ν^0 and that is therefore, up to this choice, intrinsic to the bundle Z. As any total space constructed as a bundle of differential two-forms on M, K^* possesses a tautological two form Σ, which may be described as follows: choose any point in K^*, say $k = (q, \eta)$, where $q \in M$ and $\eta \in \bigwedge^{(2,0)} T_qM$ is a holomorphic $(2, 0)$ form; then

$$\Sigma(q, \eta) = \pi_q^*(\eta); \tag{2.10}$$

we use s_{ν^0} to pull the tautological form Σ back to a $(2, 0)$ holomorphic form

$$\zeta = s_{\nu^0}^*\Sigma$$

on Z. The reproducing property as described in (2.10) shows that, under the local trivialization induced by ψ_0, ζ is indeed given by (2.9).

The local trivialization on Z is induced by ψ_0, and depends only on the choice of the contact form ν^0; consequently we can express ζ as

$$\zeta = \frac{1}{\sqrt{2}} e^{i\gamma} \left(\nu^0 \wedge (\nu^1 + i\nu^2)\right),$$

where the factor $\frac{1}{\sqrt{2}}$ guarantees the validity of (2.8). The following proposition is the fundamental step in order to build the one form σ needed in the definition of the Fefferman metric (2.7). Since its proof is quite technical we refer the reader to [**Lee86**].

PROPOSITION 6. *Let ν^0 be a fixed contact form for $(M, \Delta, \mathbf{g})$, let f_0 be the associated Reeb vector field and let ζ be the two form on Z constructed as in* (2.9). *Then*

a) *There exists a unique complex valued one form η on Z determined by the conditions*

$$\zeta = \nu^0 \wedge \eta, \quad \text{and} \quad \iota_v\eta = 0 \;\forall\, v \in TZ \text{ such that } \pi_*(v) = f_0.$$

b) *With η chosen as in a) there exists a unique real valued one form σ on Z determined by the equations*

$$\begin{cases} d\zeta = 3i\sigma \wedge \zeta, \\ \sigma \wedge d\eta \wedge \bar{\eta} = \mathrm{Tr}(d\sigma) i\sigma \wedge \nu^0 \wedge \eta \wedge \bar{\eta}. \end{cases} \tag{2.11}$$

---

[1] We identify ν^0 with $\pi^*(\nu^0)$ using the projection $\pi : Z \to M$.

The meaning of Tr *is to be intended as follows: any solution σ to* (2.11) *has the property that $d\sigma$ is the pullback of a two form on M, which by an abuse of notation we still denote by $d\sigma$. There exist a one form β on M and a real valued function f on M for which the following decomposition holds true*

$$d\sigma = f d\nu^0 + \beta \wedge \nu^0.$$

We therefore set $\mathrm{Tr}(d\sigma) = f$.

c) *The form σ satisfying both points a) and b) meets the conditions for g_{ν^0}, as defined in* (2.7)*, to be a Fefferman metric.*

Specializing the construction to the case of a three dimensional contact sub-Riemannian structure $(M, \Delta, \mathbf{g})$, we obtain the following:

$$\begin{aligned} \sigma =& \frac{d\gamma}{3} - \frac{c_{12}^1}{3}\nu^1 - \frac{c_{12}^2}{3}\nu^2 + f\nu^0, \\ \mathrm{Tr}(d\sigma) =& \frac{f_2(c_{12}^1)}{3} - \frac{f_1(c_{12}^2)}{3} - \frac{(c_{12}^1)^2}{3} - \frac{(c_{12}^2)^2}{3} + f, \\ f =& \frac{3}{4}\left(\frac{c_{10}^2 - c_{20}^1}{6} - \frac{f_2(c_{12}^1) - f_1(c_{12}^2) - (c_{12}^1)^2 - (c_{12}^2)^2}{9}\right); \end{aligned} \tag{2.12}$$

if we substitute the expression for f in (2.12), it turns out that

$$\mathrm{Tr}(d\sigma) = \frac{1}{4}\left(f_2(c_{12}^1) - f_1(c_{12}^2) - (c_{12}^1)^2 - (c_{12}^2)^2 + \frac{c_{10}^2 - c_{20}^1}{2}\right) = \frac{1}{4}\kappa.$$

We are finally in the position to give the explicit form of the Fefferman metric: with the choice of the frame $\{f_0, f_1, f_2, f_\infty\}$ dual to the orthonormal basis of T^*Z $\{\nu^0, \nu^1, \nu^2, \nu^\infty = d\gamma\}$ its matrix is given by:

$$g_{ij} = \begin{pmatrix} \frac{c_{10}^2 - c_{20}^1}{2} + \frac{(c_{12}^1)^2 + (c_{12}^2)^2 + f_1(c_{12}^2) - f_2(c_{12}^1)}{3} & -\frac{2c_{12}^1}{3} & -\frac{2c_{12}^2}{3} & \frac{2}{3} \\ -\frac{2c_{12}^1}{3} & 1 & 0 & 0 \\ -\frac{2c_{12}^2}{3} & 0 & 1 & 0 \\ \frac{2}{3} & 0 & 0 & 0 \end{pmatrix}$$

For future purposes, we write explicitly also the matrix expression for the inverse metric

$$g^{ij} = \begin{pmatrix} 0 & 0 & 0 & \frac{3}{2} \\ 0 & 1 & 0 & c_{12}^1 \\ 0 & 0 & 1 & c_{12}^2 \\ \frac{3}{2} & c_{12}^1 & c_{12}^2 & -\frac{9}{4}\left(-\frac{(c_{12}^1)^2 + (c_{12}^2)^2}{9} + \frac{c_{10}^2 - c_{20}^1}{2} + \frac{f_1(c_{12}^2) - f_2(c_{12}^1)}{3}\right) \end{pmatrix} \tag{2.13}$$

Remark 5. We emphasize the fact that the Fefferman metric constructed in this section is independent on the choice of the orthonormal frame.

Moreover, we will always implicitly assume that all the coefficients which appear in the definition of the Fefferman metric are to be intended as lifts of the original structural coefficients on M; their extension to the product manifold $Z \cong M \times S^1$ is constant on S^1.

4. Conformal Invariants

4.1. The Levi Civita Connection. For any pseudoRiemannian metric, its conformal information is contained in the Weyl Tensor. As a first step it is thus necessary to determine the Levi Civita connection associated to the ambient metric. We refer in this sense to the well-known *Koszul formula*, valid for any three smooth vector fields on M X, Y and Z

$$2g(\nabla_X Y, Z) = Xg(Y,Z) + Yg(X,Z) - Zg(X,Y) + g([X,Y],Z) - g([X,Z],Y) - g([Y,Z],X).$$

We find:

$$\begin{aligned}
&\nabla_{f_0} f_0 = \frac{1}{12}\big(-8c^1_{12}c^1_{10} - 8c^2_{12}c^2_{10} - 8f_0(c^1_{12}) - 2f_1(f_1(c^2_{12})) + 2f_1(f_2(c^1_{12}))\\
&-4c^1_{12}f_1(c^1_{12}) - 4c^2_{12}f_1(c^2_{12}) - 3f_1(c^2_{10}) + 3f_1(c^1_{20})\big)\, f_1\\
&+\frac{1}{12}\big(-8c^1_{12}c^1_{20} - 8c^2_{12}c^2_{20} - 8f_0(c^2_{12}) - 2f_2(f_1(c^2_{12})) + 2f_2(f_2(c^1_{12}))\\
&-4c^1_{12}f_2(c^1_{12}) - 4c^2_{12}f_2(c^2_{12}) - 3f_2(c^2_{10}) + 3f_2(c^1_{20})\big)\, f_2\\
&+\frac{1}{24}\big(-16(c^1_{12})^2c^1_{10} - 16c^1_{12}c^2_{12}c^2_{10} - 16c^1_{12}c^2_{12}c^1_{20} - 16(c^2_{12})^2c^2_{20}\\
&+6f_0(f_1(c^2_{12})) - 6f_0(f_2(c^1_{12})) - 4c^1_{12}f_0(c^1_{12}) - 4c^2_{12}f_0(c^2_{12}) + 9f_0(c^2_{10})\\
&-9f_0(c^1_{20}) - 4c^1_{12}f_1(f_1(c^2_{12})) + 4c^1_{12}f_1(f_2(c^1_{12})) - 8(c^1_{12})^2f_1(c^1_{12})\\
&-8c^1_{12}c^2_{12}f_1(c^2_{12}) - 8c^1_{12}c^2_{12}f_2(c^1_{12})\\
&-6c^1_{12}f_1(c^2_{10}) + 6c^1_{12}f_1(c^1_{20}) - 4c^2_{12}f_2(f_1(c^2_{12})) + 4c^2_{12}f_2(f_2(c^1_{12}))\\
&-8(c^2_{12})^2f_2(c^2_{12}) - 6c^2_{12}f_2(c^2_{10}) + 6c^2_{12}f_2(c^1_{20})\big)\, f_\infty,\\
&\nabla_{f_0} f_1 = \frac{1}{12}\big(-2(c^1_{12})^2 - 2(c^2_{12})^2 - 3c^2_{10} + 3c^1_{20} - 2f_1(c^2_{12}) + 2f_2(c^1_{12})\big)\, f_2\\
&+\frac{1}{24}\big(-4(c^1_{12})^2c^2_{12} - 4(c^2_{12})^3 + 24c^1_{12}c^1_{10} + 18c^2_{12}c^2_{10} + 6c^2_{12}c^1_{20} + 6f_1(f_1(c^2_{12}))\\
&-6f_1(f_2(c^1_{12})) + 12c^1_{12}f_1(c^1_{12}) + 8c^2_{12}f_1(c^2_{12}) + 9f_1(c^2_{10}) - 9f_1(c^1_{20})\\
&+4c^2_{12}f_2(c^1_{12})\big)\, f_\infty,\\
&\nabla_{f_0} f_2 = \frac{1}{12}\big(2(c^1_{12})^2 + 2(c^2_{12})^2 + 3c^2_{10} - 3c^1_{20} + 2f_1(c^2_{12}) - 2f_2(c^1_{12})\big)\, f_1\\
&+\frac{1}{24}\big(4(c^1_{12})^3 + 4c^1_{12}(c^2_{12})^2 + 6c^1_{12}c^2_{10} + 18c^1_{12}c^1_{20} + 24c^2_{12}c^2_{20} + 4c^1_{12}f_1(c^2_{12})\\
&+6f_2(f_1(c^2_{12})) - 6f_2(f_2(c^1_{12})) + 8c^1_{12}f_2(c^1_{12}) + 12c^2_{12}f_2(c^2_{12}) + 9f_2(c^2_{10})\\
&-9f_2(c^1_{20})\big)\, f_\infty,\\
&\nabla_{f_0} f_\infty = 0,\\
&\nabla_{f_1} f_1 = \frac{c^1_{12}}{3}f_2 + \frac{1}{6}\big(2c^1_{12}c^2_{12} - 9c^1_{10} - 6f_1(c^1_{12})\big)\, f_\infty,\\
&\nabla_{f_1} f_2 = -\frac{1}{2}f_0 - \frac{2c^1_{12}}{3}f_1 - \frac{c^2_{12}}{3}f_2\\
&+\frac{1}{12}\big(-2(c^1_{12})^2 + 2(c^2_{12})^2 - 9c^2_{10} - 9c^1_{20} - 6f_1(c^2_{12}) - 6f_2(c^1_{12})\big)\, f_\infty,
\end{aligned}$$

$$\nabla_{f_1} f_\infty = \frac{1}{3} f_2 + \frac{c_{12}^2}{3} f_\infty,$$
$$\nabla_{f_2} f_2 = -\frac{c_{12}^2}{3} f_1 + \frac{1}{6}\left(-2c_{12}^1 c_{12}^2 - 9c_{20}^2 - 6f_2(c_{12}^2)\right) f_\infty,$$
$$\nabla_{f_2} f_\infty = -\frac{1}{3} f_1 - \frac{c_{12}^1}{3} f_\infty.$$

The remaining covariant derivatives are filled by means of the commutator rules imposed by (2.3) in M, extended to Z by declaring that f_∞ commutes with all the other vector fields: $[f_\infty, f_i] = 0$, for $i = 0, 1, 2$.

Remark 6. Differentiating the second and the third relation in (2.3) one sees that

$$c_{12}^2 c_{10}^1 - c_{12}^1 c_{10}^2 + f_0(c_{12}^2) - f_1(c_{20}^2) + f_2(c_{10}^2) = 0$$
$$c_{12}^2 c_{20}^1 - c_{12}^1 c_{20}^2 - f_0(c_{12}^1) + f_1(c_{20}^1) - f_2(c_{10}^1) = 0.$$

We will extensively use these relations in the following.

Starting from the Riemann tensor

$$R(X,Y)W = [\nabla_X, \nabla_Y]W - \nabla_{[X,Y]}W,$$
$$(R^i)_{jkl} = \langle R(f_k, f_l) f_j, \nu^i \rangle,$$

the Weyl tensor associated to the Fefferman metric can be computed by:

$$W_{ijkl} = R_{ijkl} - \frac{1}{2}\left(g_{ik}R_{jl} - g_{il}R_{jk} - g_{jk}R_{il} + g_{jl}R_{ik}\right) + \frac{R}{6}\left(g_{ik}g_{jl} - g_{il}g_{jk}\right),$$

where R_{ij} is the Ricci tensor and R is the scalar curvature, that is

$$R_{ij} = R^k_{ikj}, \quad R = R^i_i.$$

Remark 7. Incidentally, observe that both the metric invariants κ and χ can be recovered considering suitable complete contractions of the Riemann tensor. In particular

$$R = \frac{3}{2}\left(f_2(c_{12}^1) - f_1(c_{12}^2) - (c_{12}^1)^2 - (c_{12}^2)^2 + \frac{1}{2}(c_{10}^2 - c_{20}^1)\right) = \frac{3}{2}\kappa$$

and

$$\chi^2 = \frac{9}{16}\left\|\nabla_{f_\infty} R_{ijkl}\right\|^2.$$

4.2. Conformal Invariants. Heuristically speaking, it is natural to expect the existence of a single conformal invariant. On the one hand indeed, any three dimensional contact sub-Riemannian structure $(M, \Delta, \mathbf{g})$ is indeed specified giving two vector fields on M, that is by six scalar equations. On the other hand five of these equations can be normalized: three of them by a suitable change of coordinates on the manifold, one by the choice of a rotation angle θ within the group $SO(2)$, and finally one more equation is defined by a suitable rescaling factor φ.

It turns out that there are just two nonzero linearly independent entries in the Weyl tensor, and each of them rescales by a factor $e^{-4\varphi}$; their ratio is then the conformal invariant associated to $(M, \Delta, \mathbf{g})$. We have:

$$
\begin{aligned}
\alpha = &-\frac{(c^1_{12})^2c^2_{10}}{12}+\frac{(c^2_{12})^2c^2_{10}}{12}-\frac{3(c^2_{10})^2}{8}+\frac{(c^1_{12})^2c^1_{20}}{12}-\frac{(c^2_{12})^2c^1_{20}}{12}+\frac{3(c^1_{20})^2}{8}\\
&-\frac{f_0(f_1(c^1_{12}))}{3}+\frac{f_0(f_2(c^2_{12}))}{3}+\frac{c^2_{12}f_0(c^1_{12})}{6}+\frac{c^1_{12}f_0(c^2_{12})}{6}-\frac{f_0(c^1_{10})}{2}+\frac{f_0(c^2_{20})}{2}\\
&-\frac{f_1(f_1(f_1(c^2_{12})))}{12}+\frac{f_1(f_1(f_2(c^1_{12})))}{12}-\frac{c^1_{12}f_1(f_1(c^1_{12}))}{6}-\frac{c^2_{12}f_1(f_1(c^2_{12}))}{12}\\
&+\frac{f_1(f_1(c^1_{20}))}{8}-\frac{c^2_{12}f_1(f_2(c^1_{12}))}{12}+\frac{c^1_{12}c^2_{12}f_1(c^1_{12})}{6}-\frac{2c^1_{10}f_1(c^1_{12})}{3}-\frac{f_1(c^1_{12})^2}{6}\\
&+\frac{(c^2_{12})^2f_1(c^2_{12})}{6}-\frac{7c^2_{10}f_1(c^2_{12})}{12}+\frac{c^1_{20}f_1(c^2_{12})}{12}-\frac{f_1(c^2_{12})^2}{6}-\frac{c^1_{12}f_1(c^1_{10})}{3}\\
&+\frac{c^2_{12}f_1(c^1_{20})}{24}+\frac{c^1_{12}f_1(c^2_{20})}{6}+\frac{c^1_{12}f_2(f_1(c^2_{12}))}{12}+\frac{f_2(f_2(f_1(c^2_{12})))}{12}\\
&+\frac{c^1_{12}f_2(f_2(c^1_{12}))}{12}+\frac{c^2_{12}f_2(f_2(c^2_{12}))}{6}+\frac{f_2(f_2(c^2_{10}))}{8}-\frac{f_2(f_2(c^1_{20}))}{8}\\
&-\frac{c^2_{10}f_2(c^1_{12})}{12}+\frac{7c^1_{20}f_2(c^1_{12})}{12}+\frac{f_2(c^1_{12})^2}{6}+\frac{c^1_{12}c^2_{12}f_2(c^2_{12})}{6}+\frac{2c^2_{20}f_2(c^2_{12})}{3}\\
&-\frac{f_1(f_1(c^2_{10}))}{8}-\frac{5c^2_{12}f_1(c^2_{10})}{24}-\frac{f_2(f_2(f_2(c^1_{12})))}{12}+\frac{(c^1_{12})^2f_2(c^1_{12})}{6}\\
&-\frac{c^2_{12}f_2(c^1_{10})}{6}-\frac{c^1_{12}f_2(c^2_{10})}{24}+\frac{5c^1_{12}f_2(c^1_{20})}{24}+\frac{c^2_{12}f_2(c^2_{20})}{3}+\frac{f_2(c^2_{12})^2}{6},
\end{aligned}
$$

$$
\begin{aligned}
\beta = &-\frac{c^1_{12}c^2_{12}c^2_{10}}{6}-\frac{c^1_{10}c^2_{10}}{8}+\frac{c^1_{12}c^2_{12}c^1_{20}}{6}-\frac{7c^1_{10}c^1_{20}}{8}-\frac{7c^2_{10}c^2_{20}}{8}-\frac{c^1_{20}c^2_{20}}{8}\\
&-\frac{f_0(f_2(c^1_{12}))}{3}-\frac{c^1_{12}f_0(c^1_{12})}{6}+\frac{c^2_{12}f_0(c^2_{12})}{6}-\frac{f_0(c^2_{10})}{2}-\frac{f_1(c^1_{12})f_2(c^1_{12})}{3}\\
&-\frac{f_1(f_2(f_1(c^2_{12})))}{12}+\frac{f_1(f_2(f_2(c^1_{12})))}{12}-\frac{c^1_{12}f_1(f_2(c^1_{12}))}{12}-\frac{c^2_{12}f_1(f_2(c^2_{12}))}{6}\\
&+\frac{f_1(f_2(c^1_{20}))}{8}-\frac{(c^1_{12})^2f_1(c^1_{12})}{6}-\frac{2c^1_{20}f_1(c^1_{12})}{3}-\frac{c^1_{12}c^2_{12}f_1(c^2_{12})}{6}-\frac{c^1_{10}f_1(c^2_{12})}{12}\\
&-\frac{7c^2_{20}f_1(c^2_{12})}{12}-\frac{c^1_{12}f_1(c^2_{10})}{8}-\frac{3c^1_{12}f_1(c^1_{20})}{8}-\frac{c^2_{12}f_1(c^2_{20})}{6}-\frac{f_2(f_1(f_1(c^2_{12})))}{12}\\
&+\frac{f_2(f_1(f_2(c^1_{12})))}{12}-\frac{c^1_{12}f_2(f_1(c^1_{12}))}{6}-\frac{c^2_{12}f_2(f_1(c^2_{12}))}{12}-\frac{f_2(f_1(c^2_{10}))}{8}\\
&-\frac{c^2_{12}f_2(f_2(c^1_{12}))}{12}+\frac{c^1_{12}c^2_{12}f_2(c^1_{12})}{6}-\frac{7c^1_{10}f_2(c^1_{12})}{12}-\frac{c^2_{20}f_2(c^1_{12})}{12}-\frac{f_0(c^1_{20})}{2}\\
&+\frac{(c^2_{12})^2f_2(c^2_{12})}{6}-\frac{2c^2_{10}f_2(c^2_{12})}{3}-\frac{f_1(c^2_{12})f_2(c^2_{12})}{3}-\frac{c^1_{12}f_2(c^1_{10})}{6}-\frac{3c^2_{12}f_2(c^2_{10})}{8}\\
&-\frac{c^2_{12}f_2(c^1_{20})}{8}-\frac{f_0(f_1(c^2_{12}))}{3}-\frac{c^1_{12}f_1(f_1(c^2_{12}))}{12}-\frac{f_1(f_2(c^2_{10}))}{8}+\frac{f_2(f_1(c^1_{20}))}{8}.
\end{aligned}
$$

To carry out the calculations needed to verify that both α and β rescale by a factor $e^{-4\varphi}$ as claimed we need the following lemmas:

LEMMA 7. *Let $\overline{q} \in M$ and let U be a neighborhood of $\overline{q}$. Let $\varphi : U \to \mathbb{R}$ be any smooth rescaling function for the sub-Riemannian metric i.e. assume $\Delta_q = \operatorname{span}\{e^{-\varphi}f_1(q), e^{-\varphi}f_2(q)\}$,* *for every $q \in U$. Then the local frame of T_qM, $q \in U$ transforms as*

$$\begin{cases} f_{0,\varphi} = e^{-2\varphi}\left(f_0 + 2f_2(\varphi)f_1 - 2f_1(\varphi)f_2\right), \\ f_{1,\varphi} = e^{-\varphi}f_1, \\ f_{2,\varphi} = e^{-\varphi}f_2. \end{cases} \tag{2.14}$$

Proof. Let $\nu^0_\varphi = h\nu^0$, $C^\infty(U) \ni h : U \to \mathbb{R}$ denote the rescaled contact form. Since $d\nu^0_\varphi\big|_\Delta$ is normalized to be the area form on Δ, and it is uniquely fixed by this requirement, h can be determined from the following two equations, that is:

$$d\nu^0_\varphi(e^{-\varphi}f_1, e^{-\varphi}f_2) = 1$$

and

$$\begin{aligned} d\nu^0_\varphi\big|_\Delta &= d(h\nu^0)\big|_\Delta \\ &= \left(d(h) \wedge \nu^0 + hd\nu^0\right)\big|_\Delta \\ &= hd\nu^0 \\ &= h\nu^1 \wedge \nu^2, \end{aligned}$$

where the equality between the second and the third line follows since $\nu^0\big|_\Delta \equiv 0$. Then $h = e^{2\varphi}$. Moreover, since

$$\begin{aligned} d\nu^0_\varphi &= e^{2\varphi}\left(2d(\varphi) \wedge \nu^0 + d\nu^0\right) \\ &= e^{2\varphi}\left(-2f_1(\varphi)\nu^0 \wedge \nu^1 - 2f_2(\varphi)\nu^0 \wedge \nu^2 + \nu^1 \wedge \nu^2\right), \end{aligned}$$

using the definition of the Reeb vector field $\iota_{f_{0,\varphi}}\nu^0_\varphi = 1$ and $\iota_{f_{0,\varphi}}d\nu^0_\varphi = 0$, we find the expression for $f_{0,\varphi}$ as claimed in (2.14). □

Remark 8. The rescaled dual frame is completed by

$$\begin{cases} \nu^1_\varphi = e^\varphi\nu^1 - 2e^\varphi f_2(\varphi)\nu^0, \\ \nu^2_\varphi = e^\varphi\nu^2 + 2e^\varphi f_1(\varphi)\nu^0. \end{cases}$$

As expected,

$$\nu^1_\varphi \wedge \nu^2_\varphi\big|_\Delta = e^{2\varphi}\nu^1 \wedge \nu^2 = d\nu^0_\varphi\big|_\Delta.$$

Lemma 8. *Let $f_{0,\varphi}, f_{1,\varphi}, f_{2,\varphi}$ be the rescaled frame as in Lemma 7. Then the structural coefficients transform according to:*

$$\begin{cases} c^1_{12,\varphi} = e^{-\varphi}\left(c^1_{12} - 3f_2(\varphi)\right), \\ c^2_{12,\varphi} = e^{-\varphi}\left(c^2_{12} + 3f_1(\varphi)\right), \\ c^1_{10,\varphi} = e^{-2\varphi}\left(-4f_1(\varphi)f_2(\varphi) + c^1_{10} + 2f_1(f_2(\varphi)) + 2c^1_{12}f_1(\varphi) + f_0(\varphi)\right), \\ c^2_{10,\varphi} = e^{-2\varphi}\left(4f_1(\varphi)^2 + c^2_{10} - 2f_1(f_1(\varphi)) + 2c^2_{12}f_1(\varphi)\right), \\ c^1_{20,\varphi} = e^{-2\varphi}\left(-4f_2(\varphi)^2 + c^1_{20} + 2f_2(f_2(\varphi)) + 2c^1_{12}f_2(\varphi)\right), \\ c^2_{20,\varphi} = e^{-2\varphi}\left(4f_1(\varphi)f_2(\varphi) + c^2_{20} + 2c^2_{12}f_2(\varphi) - 2f_2(f_1(\varphi)) + f_0(\varphi)\right). \end{cases} \tag{2.15}$$

Proof. It is a straightforward computation using lemma 7 and the structural equations (2.3). □

In the case of a left invariant contact three dimensional sub-Riemannian structure $(M, \Delta, \mathbf{g})$, under the choice of the canonical frame of either Proposition 4 or 5, equalities in (2.15) become

$$c_{12}^1 c_{10}^2 = 0 \text{ and } c_{12}^2 c_{20}^1 = 0. \tag{2.16}$$

On the other hand, in this case α and β read

$$\alpha = \frac{(c_{12}^2)^2 c_{10}^2}{12} - \frac{3(c_{10}^2)^2}{8} + \frac{(c_{12}^1)^2 c_{20}^1}{12} + \frac{3(c_{20}^1)^2}{8} \text{ and } \beta = 0. \tag{2.17}$$

Therefore, when well-defined, the ratio $\frac{\beta}{\alpha}$ is equal to 0.

5. Local Conformal Flatness of Left Invariant Structures

5.1. Local Conformal Flatness. A pseudoRiemannian metric g is said to be locally conformally flat around a point $q \in M$ if there exist a neighborhood U of q and a local rescaling function $\varphi : M \supseteq U \to \mathbb{R}$ so that the Riemann tensor of $\tilde{g} = e^{2\varphi} g$ is zero: this is equivalent to say that a pseudoRiemannian manifold M is (locally) conformally flat at q if, and only if, after a suitable rescaling, it becomes locally isometric to its tangent space $T_q M$; this second formulation can indeed be used to define flatness in the sub-Riemannian setting.

Remark 9. It is known that the right notion of tangent space in the sense of Gromov for the sub-Riemannian case is that of the *nilpotent approximation* introduced by Mitchell [**Mit85**]. For a three dimensional contact structure $(M, \Delta, \mathbf{g})$ the nilpotent approximation is unique and it is precisely the Heisenberg Group $\mathbf{H}_3$.

Definition 8. *A contact three dimensional sub-Riemannian structure* $(M, \Delta, \mathbf{g})$ *is locally conformally flat at* $q \in M$ *if there exist a neighborhood* U *of* q *and a function* $\varphi : U \to \mathbb{R}$ *so that the rescaled structure* $\left(M, \Delta, e^{2\varphi}\mathbf{g}\right)$ *becomes locally isometric to the Heisenberg Group. If* φ *can be defined globally on* M *we say that* $(M, \Delta, \mathbf{g})$ *is conformally flat.*

We know from the previous section that the scalar conformal invariant associated to a three dimensional left invariant contact sub-Riemannian structure is zero, nonetheless any locally conformally flat structure necessarily satisfies $\alpha = 0$; we are now going to analyze in greater details all the situations in which this actually occurs.

Remark 10. A contact three dimensional sub-Riemannian structure is locally isometric to the Heisenberg group if and only if both its metric invariants χ and κ are zero. If we denote with κ_φ and χ_φ the metric invariants of the rescaled structure, around any point $q \in M$ there have to be a neighborhood U of q and a suitable rescaling φ such that the equalities $\kappa_\varphi = \chi_\varphi = 0$ hold on U.

5.2. Non Unimodular Structures. We recall that a Lie Group M is *unimodular* if the Haar measure on it is both left and right invariant. Then a unimodular (resp. non unimodular) contact structure $(M, \Delta, \mathbf{g})$ is a left invariant contact structure defined on a unimodular (resp. non unimodular) lie group M. Equations (2.16) reveal that non unimodular structures are divided into two distinct subfamilies, namely

a) $\mathfrak{solv}^+$ structures satisfying $c_{12}^1 = c_{20}^1 = 0$,
b) $\mathfrak{solv}^-$ structures satisfying $c_{12}^2 = c_{10}^2 = 0$.

Notice that the choice of the canonical frame forces $\chi \neq 0$; therefore $c_{10}^2 \neq 0$ in case *a*), and $c_{20}^1 \neq 0$ in *b*). Observe moreover that from (2.17) we have $\alpha = -\frac{\chi}{6}(\kappa + 8\chi)$ for $\mathfrak{solv}^+$ structures, and $\alpha = -\frac{\chi}{6}(\kappa - 8\chi)$ for $\mathfrak{solv}^-$ structures.

The main result for non unimodular structures is the following:

Theorem 9. *Let $(M, \Delta, \mathbf{g})$ be a three dimensional left invariant non unimodular sub-Riemannian contact structure. Then $(M, \Delta, \mathbf{g})$ is locally conformally flat if and only if its canonical frame satisfies one of the following set of structural equations*

$$\mathfrak{solv}^+ \begin{cases} [f_2, f_1] = f_0 + c_{12}^2 f_2, \\ [f_1, f_0] = \frac{2}{9}\left(c_{12}^2\right)^2 f_2, \\ [f_2, f_0] = 0 \end{cases} \qquad \mathfrak{solv}^- \begin{cases} [f_2, f_1] = f_0 + c_{12}^1 f_1, \\ [f_2, f_0] = -\frac{2}{9}\left(c_{12}^1\right)^2 f_2, \\ [f_1, f_0] = 0. \end{cases} \tag{2.18}$$

In particular a rescaling function φ can be chosen requiring that

$$\mathfrak{solv}^+ \begin{cases} f_1(\varphi) = -\frac{c_{12}^2}{3} \\ f_2(\varphi) = 0 \end{cases} \qquad \mathfrak{solv}^- \begin{cases} f_2(\varphi) = \frac{c_{12}^1}{3} \\ f_1(\varphi) = 0. \end{cases} \tag{2.19}$$

Remark 11. The canonical frame in the $\mathfrak{solv}^-$ case does not satisfy $c_{10}^2 + c_{20}^1 \geq 0$; however if we change the orientation on the contact planes, that is $\tilde{f}_1 = f_2$ and $\tilde{f}_2 = f_1$, we fall back in the $\mathfrak{solv}^+$ case.

Proof. It is not restrictive to assume $c_{12}^2 \neq 0$ for the $\mathfrak{solv}^+$ case or $c_{12}^1 \neq 0$ for the $\mathfrak{solv}^-$ case, for otherwise everything reduces to the Heisenberg case, which is flat by definition.

Observe that the structural equations (2.18) are necessary for a non unimodular structure $(M, \Delta, \mathbf{g})$ to have $\alpha = 0$. On the other hand assume that the structural equations of $(M, \Delta, \mathbf{g})$ satisfies (2.18). If we were to find a smooth rescaling function as in (2.19), then Lemma 8 would force all the rescaled structural coefficients of $\left(M, \Delta, e^{2\varphi}\mathbf{g}\right)$ to be sent to zero (indeed the commutator equality $[f_2, f_1] = f_0$ combined with the constancy of the structural coefficients of $(M, \Delta, \mathbf{g})$ imply in both cases the further relation $f_0(\varphi) = 0$).

Thus we are left with proving the existence of φ satisfying (2.19); however, a trivial computation shows that the one forms

$$\mathfrak{solv}^+ \ \gamma = -\frac{c_{12}^2}{3}\nu^1 \ \text{and} \ \ \mathfrak{solv}^- \ \gamma = \frac{c_{12}^1}{3}\nu^2$$

are closed; the result then follows by Poincaré's Lemma, and the theorem is proved. □

Remark 12. It is possible to give an explicit coordinate description of φ. We go through all the details just for $\mathfrak{solv}^+$ case, the other being entirely analogous.

Since $[f_2, f_0] = 0$, there exists a diffeomorphism (cfr. [**AS04**])

$$\Phi : M \ni O_q \to O_0 \in \mathbb{R}^3, \quad q \in M$$

such that

$$\begin{aligned} \Phi_*(f_0) &= \frac{\partial}{\partial z}, \\ \Phi_*(f_1) &= \frac{\partial}{\partial x} + a_1(x,y,z)\frac{\partial}{\partial y} + a_2(x,y,z)\frac{\partial}{\partial z}, \\ \Phi_*(f_2) &= \frac{\partial}{\partial y}. \end{aligned}$$

From the structural equations (2.18) it follows that:

$$\Phi_*(f_1) = \frac{\partial}{\partial x} + \left(c_{12}^2 y - \frac{2}{9}\left(c_{12}^2\right)^2 z + b_1(x)\right)\frac{\partial}{\partial y} + (y + b_2(x))\frac{\partial}{\partial z}.$$

If we now write in coordinates $f_1(\varphi) = -\frac{c_{12}^2}{3}$ subject to the conditions $f_2(\varphi) = f_0(\varphi) = 0$ it is easy toconclude that

$$\varphi(x) = A_1 - \frac{c_{12}^2}{3}x, \quad A_1 \in \mathbb{R}.$$

Analogously, for the $\mathfrak{solv}^-$ case we find that φ must be of the form

$$\varphi(y) = B_1 + \frac{c_{12}^1}{3}y, \quad B_1 \in \mathbb{R}.$$

5.3. Unimodular Structures. A connected Lie Group M is unimodular if, and only if

$$\mathrm{trace}(\mathrm{ad}(m)) = 0, \ \forall m \in \mathfrak{m},$$

where $\mathfrak{m}$ denotes the Lie Algebra of M. In terms of the canonical frame, this translates into the conditions $c_{12}^1 = c_{12}^2 = 0$. Observe that then we have:

$$\alpha = 0 \Leftrightarrow \frac{3}{8}(c_{10}^2 + c_{20}^1)(c_{10}^2 - c_{20}^1) = \frac{3}{2}\kappa\chi = 0,$$

that is, $\alpha = 0$ if, and only if, either $\kappa = 0$ or $\chi = 0$. To verify that a unimodular structure $(M, \Delta, \mathbf{g})$ is locally conformally flat, it is sufficient to exhibit a rescaling function φ which solves locally the following system of equations (cfr. Lemma 8):

$$\begin{cases} \kappa_\varphi = -4f_1(f_1(\varphi)) - 4f_2(f_2(\varphi)) - 4f_1(\varphi)^2 - 4f_2(\varphi)^2 + \kappa = 0, \\ \chi_\varphi = 2f_1(\varphi)^2 - f_1(f_1(\varphi)) - 2f_2(\varphi)^2 + f_2(f_2(\varphi)) + \chi = 0, \end{cases} \tag{2.20}$$

Our first proposition concerns the realization of explicit models for unimodular structures; indeed computations becomes easier to handle in coordinates. We work, as customary, under the assumption $\kappa^2 + \chi^2 = 1$; in particular we will analyze three different cases:

a) $\kappa = -1$, $\chi = 0$,
b) $\kappa = 1$, $\chi = 0$,
c) $\kappa = 0$, $\chi = 1$.

Proposition 10. *Let $(M, \Delta, \mathbf{g})$ be a left invariant three dimensional unimodular sub-Riemannian contact satisfying one of the following set of structural equations*

$$a) \begin{cases} [f_2, f_1] = f_0, \\ [f_1, f_0] = -f_2, \\ [f_2, f_0] = f_1. \end{cases} \quad b) \begin{cases} [g_2, g_1] = g_0, \\ [g_1, g_0] = g_2, \\ [g_2, g_0] = -g_1, \end{cases} \quad c) \begin{cases} [l_2, l_1] = l_0, \\ [l_1, l_0] = l_2, \\ [l_2, l_0] = l_1. \end{cases} \tag{2.21}$$

An explicit model for each of these structures is given by:

$$
(2.22)\qquad a)\quad \begin{cases} f_0 = \dfrac{\partial}{\partial z} \\ f_1 = e^y \cos(z)\dfrac{\partial}{\partial x} - \sin(z)\dfrac{\partial}{\partial y} + \cos(z)\dfrac{\partial}{\partial z}, \\ f_2 = -e^y \sin(z)\dfrac{\partial}{\partial x} - \cos(z)\dfrac{\partial}{\partial y} - \sin(z)\dfrac{\partial}{\partial z}; \end{cases}
$$

$$
b)\quad \begin{cases} g_0 = \dfrac{\partial}{\partial z} \\ g_1 = ie^y \cos(z)\dfrac{\partial}{\partial x} - i\sin(z)\dfrac{\partial}{\partial y} + i\cos(z)\dfrac{\partial}{\partial z}, \\ g_2 = ie^y \sin(z)\dfrac{\partial}{\partial x} + i\cos(z)\dfrac{\partial}{\partial y} + i\sin(z)\dfrac{\partial}{\partial z}; \end{cases}
$$

$$
c)\quad \begin{cases} l_0 = -i\dfrac{\partial}{\partial z} \\ l_1 = -ie^y \sin(z)\dfrac{\partial}{\partial x} - i\cos(z)\dfrac{\partial}{\partial y} - i\sin(z)\dfrac{\partial}{\partial z}, \\ l_2 = e^y \cos(z)\dfrac{\partial}{\partial x} - \sin(z)\dfrac{\partial}{\partial y} + \cos(z)\dfrac{\partial}{\partial z}. \end{cases}
$$

PROOF. Notice that the structures in b) and c) are isomorphic to the structure in a), and the correspondence is given by

$$
b)\quad \begin{array}{l} g_0 \mapsto f_0, \\ g_1 \mapsto if_1, \\ g_2 \mapsto -if_2, \end{array} \qquad c)\quad \begin{array}{l} l_0 \mapsto -if_0, \\ l_1 \mapsto if_2, \\ l_2 \mapsto f_1. \end{array}
$$

It is then sufficient to observe that the coordinates given for the model in a) satisfy the structural equations (2.21) to conclude. □

With explicit models at hand, it is immediate to prove the following:

THEOREM 11. *Let* $(M, \Delta, \mathbf{g})$ *be a left invariant three dimensional unimodular structure whose associated ambient metric is locally conformally flat, that is let* $(M, \Delta, \mathbf{g})$ *be of type a), b) or c) as in Proposition 10. Then* $(M, \Delta, \mathbf{g})$ *is locally conformally flat, and if we choose the model representation of suggested in* (2.22)*, an admissible rescaling function* φ *is given by:*

$$
\begin{aligned} a, b)\quad & \varphi = \frac{1}{2}y + C, \quad C \in \mathbb{R}, \\ c)\quad & \varphi = -y + C, \quad C \in \mathbb{R}. \end{aligned}
$$

PROOF. It is just a matter of computations. Notice that in case a) our choice of φ solves (2.20) since the following relations hold true

$$
\begin{aligned} f_1(f_1(\varphi)) &= -\frac{1}{2}\cos^2(z), \quad f_1(\varphi)^2 = \frac{1}{4}\sin^2(z), \\ f_2(f_2(\varphi)) &= -\frac{1}{2}\sin^2(z), \quad f_2(\varphi)^2 = \frac{1}{4}\cos^2(z). \end{aligned}
$$

For cases b) and c) we proceed similarly, and this concludes the proof. □

6. Classification of Non Flat Left Invariant Structures

6.1. Generalities. Geodesics of the ambient metric can be characterized as projections of integral lines of the Hamiltonian vector field $\overrightarrow{H}$, obtained as the symplectic lift of the function

$$H(q,h) = \frac{1}{2}\sum_{i,j} g^{ij} h_i h_j, \quad h_i(\lambda) = \langle \lambda, f_i(q)\rangle,\ q = \pi(\lambda),$$

which is quadratic on the fibers of the cotangent bundle, and g^{ij} is the inverse ambient metric defined coordinate-wise in (2.13). In particular, we are interested just in light-like geodesics, the so called *chains*, that is the set of solutions corresponding to the zero level set $H(q,h) = 0$.

Chains are conformally invariant: any rescaling of the metric on the contact planes $\{\Delta_q\}_{q\in M}$ induces a rescaling for the Fefferman metric $\widetilde{g^{ij}} = e^{-2\varphi} g^{ij}$; accordingly, the new Hamiltonian geodesic field is related to the old one on the set $\{H = 0\}$ by

$$\overrightarrow{\widetilde{H}} = e^{-2\varphi}\overrightarrow{H};$$

this proportionality relation says that light-like geodesics for the two metrics are the same if considered as sets of unparametrized curves. Here we recall that the symplectic lift of an Hamiltonian function is defined by the equality

$$\iota_{\vec{H}}\sigma = -dH,$$

where $\sigma = ds$ denotes the symplectic form of the cotangent bundle T^*Z and

$$s = h_0\nu^0 + h_1\nu^1 + h_2\nu^2 + h_\infty\nu^\infty$$

is its tautological one form. Notice also that the left invariance of our structures imply that H is actually independent on the base point, that is $H(q,h) = H(h)$.

Remark 13. It is worth recalling that the momentum scaling property

$$H(q,\lambda h) = \lambda^2 H(q,h), \quad h \in T_q^*Z,$$

corresponds to the fact that the geodesic $\tilde{\gamma}(t)$ with initial conditions $(q, \lambda h)$ is the same as the geodesic $\gamma(t)$ as represented by the initial conditions (q,h), but parametrized with a different speed, that is $\tilde{\gamma}(t) = \gamma(\lambda t)$.

Remark 14. It is well-known [**AS04, Ar89**] that, since geodesics are solutions of the ODE

$$\lambda(t) = e^{t\overrightarrow{H}}\lambda(0),$$

differentiating with respect to time both sides of this equation, it follows:

$$\dot{\lambda}(t) = \overrightarrow{H}e^{t\overrightarrow{H}}\lambda(0) = \overrightarrow{H}\lambda(t) = \{H, \lambda(t)\},$$

from which it is immediate to read in fiber coordinates $\{h_0, h_1, h_2, h_\infty\}$ that $\dot{h}_\infty \equiv 0$. Indeed by construction this coordinate commutes with all the others, that is it belongs to the center of the Lie Algebra generated by the h_i's. In particular, since H is also independent on the base point, h_∞ remains constant along chains.

By the scaling momentum property recalled above, we may distinguish between the cases $h_\infty \equiv 0$ or $h_\infty \equiv 1$; only the second instance will matter to us.

6.2. Conformal Invariance.

LEMMA 12. *Let $(M, \Delta, \mathbf{g})$ be a left invariant three dimensional contact sub-Riemannian structure. Then any conformal rescaling on $(M, \Delta, \mathbf{g})$ preserves the algebra of integral of motions for the chains' flow on the level set $\{H = 0\}$, where H is the Hamiltonian relative to the Fefferman ambient metric built on $(M, \Delta, \mathbf{g})$.*

PROOF. If P is such that

$$\{P, H\}\big|_{H=0} = 0,$$

since a conformal rescaling φ on $(M, \Delta, \mathbf{g})$ maps H into $e^{-2\varphi}H$, then

$$\{P, e^{-2\varphi}H\}\big|_{H=0} = \{P, e^{-2\varphi}\}H\big|_{H=0} + e^{-2\varphi}\{P, H\}\big|_{H=0} = 0^2,$$

where as usual we are identifying the function $e^{-2\varphi}$ on the manifold M with its pullback $\pi^*(e^{-2\varphi})$ as a constant on fiber function on T^*Z. □

REMARK 15. A similar statement holds true even for the algebra of the elements commuting with h_∞. The argument in this case follows directly because h_∞ is not affected by conformal rescalings on $(M, \Delta, \mathbf{g})$.

COROLLARY 13. *Let $(M, \Delta, \mathbf{g})$ be a left invariant three dimensional sub-Riemannian contact structure. Assume that P is a smooth function on $T^*Z \cong T^*(M \times S^1)$ commuting both with H and h_∞ on the set $\{H = 0\}$. Then any conformal rescaling on $(M, \Delta, \mathbf{g})$ fiberwise preserves the foliation induced on the $\{h_1, h_2\}$-plane generated by the intersection of $H = 0$, $h_\infty = 1$ and the level sets of P.*

PROOF. Let φ be any conformal rescaling of $(M, \Delta, \mathbf{g})$ and let us fix a fiber $T^*_{(q,\gamma)}Z$. H rescales as $e^{-2\varphi(q)}H$, therefore its zero level set remains unchanged; moreover h_∞ is not affected by φ, hence via momentum scaling can always be renormalized to be equal to one. Using Lemma 12 and the subsequent remark, we conclude that the foliation generated on the $\{h_1, h_2\}$-plane by the level sets of P, $H = 0$ and $h_\infty = 1$ must be preserved by φ. □

REMARK 16. In general it is not true that each leaf will be preserved by φ; in fact it may very well happen that leaves, which were originally indexed along the level sets $P = \text{const}$, are now mixed according to some smooth automorphism $a : \mathbb{C} \to \mathbb{C}$ induced by the rescaling φ and are accordingly labeled as level sets $a(P) = \text{const}$.

In any case, since any conformal rescaling of $(M, \Delta, \mathbf{g})$ must preserve the foliation induced by P on the set $\{H = 0\} \cap \{h_\infty = 1\}$, we deduce that it must also preserve the (normalized) distribution $\frac{\nabla P}{|\nabla P|}$ attached to any of these points. Since the distribution is not affected by the mixing of the leaves, we have found an effective tool in order to carry out the classification of left invariant three dimensional contact structures.

Possibly changing the orientation on the contact planes, we will assume $c_{10}^2 + c_{20}^1 \geq 0$. Although the analysis can be carried out along the same lines both for unimodular and non unimodular structures, we prefer to split the presentation into two parts since the calculations involved are quite different. Also, for non unimodular structures we will furnish all the details just for those in $\mathfrak{solv}^+$, since the handling of the $\mathfrak{solv}^-$ case is almost identical.

[2] The way φ acts on the level sets of P may be nontrivial, i.e. $\{P, e^{-2\varphi}\} \neq 0$; this is irrelevant if we restrict to the zero level set of H.

6.3. Non Unimodular Structures. By the choice of a canonical frame for a $\mathfrak{solv}^+$ structure, that is

$$\begin{cases} [f_2, f_1] = f_0 + c_{12}^2 f_2, \\ [f_1, f_0] = c_{10}^2 f_2, \\ [f_2, f_0] = 0, \end{cases} \tag{2.23}$$

the condition $2\chi = c_{10}^2 \neq 0$ holds. Further, by possibly change the signs both for f_1 and f_2, we may also assume $\sqrt{\chi - \kappa} = c_{12}^2 > 0$. Moreover, notice that $\mathfrak{solv}^+$ structures may be interpreted as the semidirect product of the two dimensional Lie algebra spanned by f_0 and f_2 by the one dimensional Lie algebra spanned by f_1, where the action of f_1 is given through the matrix

$$\mathrm{ad}_{f_1} = \begin{pmatrix} 0 & -1 \\ c_{10}^2 & -c_{12}^2 \end{pmatrix}$$

PROPOSITION 14. *Let $(M, \Delta, \mathbf{g})$ be a $\mathfrak{solv}^+$ left invariant three dimensional contact structure. Define*

$$\delta = (c_{12}^2)^2 - 4c_{10}^2 = -\kappa - 7\chi.$$

Then there exists a frame $\tilde{f}_0, \tilde{f}_1, \tilde{f}_2$ for $(M, \Delta, \mathbf{g})$ satisfying the following relations, depending on whether a) $\delta = 0$, b) $\delta > 0$ or c) $\delta < 0$ respectively:

$$a) \begin{cases} [\tilde{f}_2, \tilde{f}_1] = \tilde{f}_0 + \tilde{f}_2, \\ [\tilde{f}_0, \tilde{f}_1] = \tilde{f}_0, \\ [\tilde{f}_2, \tilde{f}_0] = 0; \end{cases} \quad b) \begin{cases} [\tilde{f}_2, \tilde{f}_1] = \tilde{f}_2, \\ [\tilde{f}_0, \tilde{f}_1] = a\tilde{f}_0, \\ [\tilde{f}_2, \tilde{f}_0] = 0; \end{cases} \quad c) \begin{cases} [\tilde{f}_2, \tilde{f}_1] = \tilde{f}_0 + b\tilde{f}_2, \\ [\tilde{f}_0, \tilde{f}_1] = b\tilde{f}_0 - \tilde{f}_2. \\ [\tilde{f}_2, \tilde{f}_0] = 0; \end{cases}$$

where

$$0 < a = \frac{c_{12}^2 - \sqrt{\delta}}{c_{12}^2 + \sqrt{\delta}} < 1, \quad \textit{and} \quad 0 < b = \frac{c_{12}^2}{\sqrt{-\delta}} < +\infty.$$

PROOF. It suffices to bring the matrix $A = -\mathrm{ad}_{f_1}$ into its real canonical Jordan form, and subsequently to rescale f_1 to normalize the structural equations. The explicit change of basis is given by

$$a) \begin{cases} \tilde{f}_0 = f_0 + \dfrac{c_{12}^2}{2} f_2, \\ \tilde{f}_1 = \dfrac{2}{c_{12}^2} f_1, \\ \tilde{f}_2 = f_0; \end{cases} \quad b) \begin{cases} \tilde{f}_0 = f_0 + \dfrac{c_{12}^2 - \sqrt{\delta}}{2} f_2, \\ \tilde{f}_1 = \dfrac{2}{c_{12}^2 + \sqrt{\delta}} f_1, \\ \tilde{f}_2 = f_0 + \dfrac{c_{12}^2 + \sqrt{\delta}}{2} f_2; \end{cases} \quad c) \begin{cases} \tilde{f}_0 = 2f_0 + c_{12}^2 f_2, \\ \tilde{f}_1 = \dfrac{2}{\sqrt{-\delta}} f_1, \\ \tilde{f}_2 = \sqrt{-\delta} f_2. \end{cases}$$

□

REMARK 17. The Lie algebras in the previous Proposition correspond respectively to the elements $A_{3,2}$, $A_{3,5}^a$ and $A_{3,7}^b$ in [**PSWZ76**]; we recover from there the invariants for the Lie algebra generated by $\{h_0, h_1, h_2, h_\infty\}$, endowed with the Poisson brackets. In terms of the metric invariants χ and κ they read

$$\begin{aligned}
a)\quad & J = \left(h_0 + \frac{\sqrt{\chi-\kappa}}{2}h_2\right)\exp\left(\frac{2h_0}{2h_0+\sqrt{\chi-\kappa}h_2}\right)\\
b)\quad & K = \left(h_0 + \frac{\sqrt{\chi-\kappa}-\sqrt{-\kappa-7\chi}}{2}h_2\right)\times\\
& \times\left(h_0 + \frac{\sqrt{\chi-\kappa}+\sqrt{-\kappa-7\chi}}{2}h_2\right)^{-\frac{\sqrt{\chi-\kappa}-\sqrt{-\kappa-7\chi}}{\sqrt{\chi-\kappa}+\sqrt{-\kappa-7\chi}}}\\
c)\quad & L = \left((2h_0+\sqrt{\chi-\kappa}h_2)^2 + (\kappa+7\chi)h_2^2\right)\times\\
& \times\left(\frac{2h_0+\sqrt{\chi-\kappa}h_2 + i\sqrt{\kappa+7\chi}h_2}{2h_0+\sqrt{\chi-\kappa}h_2 - i\sqrt{\kappa+7\chi}h_2}\right)^{i\frac{\sqrt{\chi-\kappa}}{\sqrt{\kappa+7\chi}}}
\end{aligned}$$

For $\mathfrak{solv}^+$ structures the chain equation is

$$H(h) = 3h_0h_\infty + h_1^2 + h_2^2 + 2\sqrt{\chi-\kappa}h_2h_\infty - \frac{1}{4}\left(\kappa+8\chi\right)h_\infty^2 = 0.$$

PROPOSITION 15. *If $(M,\Delta,\mathbf{g})$ is a $\mathfrak{solv}^+$ structure, then the distribution of Remark 16 associated with the foliation generated by $H=0$, $h_\infty=1$ and either a) J, b) K or c) L on the $\{h_1,h_2\}$-plane and attached to the point $(1,1)$, is functionally dependent on*

$$\begin{aligned}
a)\quad & \sqrt{\chi-\kappa} \quad \textit{and} \quad \kappa+8\chi\\
b)\quad & \sqrt{\chi-\kappa},\quad \sqrt{-\kappa-7\chi} \quad \textit{and} \quad \kappa+8\chi\\
c)\quad & \sqrt{\chi-\kappa},\quad \sqrt{\kappa+7\chi} \quad \textit{and} \quad \kappa+8\chi.
\end{aligned}$$

PROOF. Let $(M,\Delta,\mathbf{g})$ be any $\mathfrak{solv}^+$ structure, then the foliation is given considering the level sets of either $J=\text{const}$, $K=\text{const}$ or $L=\text{const}$ (according to the sign of δ), subject to the constraints

$$h_\infty = 1 \quad \text{and} \quad h_0 = -\frac{1}{3}(h_1^2+h_2^2+2\sqrt{\chi-\kappa}h_2) + \frac{1}{12}(\kappa+8\chi).$$

Moreover, an explicit (although quite messy) computation of

$$\nabla X_{h_1,h_2}\Big|_{(1,1)}, \quad X = \{J,K,L\},$$

shows that its components functionally depends on the metric constants announced in the statement of the proposition. □

6.4. Unimodular Structures. The canonical frame for a unimodular structure can be expressed as

$$\begin{cases}
[f_2,f_1]=f_0,\\
[f_1,f_0]=(\chi+\kappa)f_2,\\
[f_2,f_0]=(\chi-\kappa)f_1.
\end{cases}$$

Moreover, the chain equation reads in this case

$$H(h) = 3h_0h_\infty + h_1^2 + h_2^2 - \frac{9}{4}\kappa h_\infty^2 = 0,$$

and from the structural equations it is easy to deduce that

$$I = h_0^2 - (\chi - \kappa)h_1^2 + (\chi + \kappa)h_2^2$$

is a *Casimir* for the Lie Algebra, and as such it commutes with all the coordinates h_0, h_1, h_2 and h_∞.

PROPOSITION 16. *Let $(M, \Delta, \mathbf{g})$ be a unimodular left invariant three dimensional contact structure. Then the distribution of Remark 16 associated with the foliation generated by $H = 0$, $h_\infty = 1$ and I on the $\{h_1, h_2\}$-plane and attached to the point $(1, 1)$, functionally depends on*

$$\chi + \frac{1}{2}\kappa \quad \textit{and} \quad \chi - \frac{1}{2}\kappa.$$

PROOF. The foliation on $(M, \Delta, \mathbf{g})$ induced by I, $H = 0$ and $h_\infty = 1$ is given by

$$I = \frac{9}{16}\kappa^2 - \left(\chi - \frac{1}{2}\kappa\right)h_1^2 + \left(\chi + \frac{1}{2}\kappa\right)h_2^2 + \frac{1}{9}(h_1^2 + h_2^2)^2.$$

We then evaluate

$$\begin{pmatrix} \partial_{h_1} I \\ \\ \partial_{h_2} I \end{pmatrix}_{(1,1)} = \begin{pmatrix} \frac{8}{9} - 2\left(\chi - \frac{\kappa}{2}\right) \\ \\ \frac{8}{9} + 2\left(\chi + \frac{\kappa}{2}\right); \end{pmatrix}$$

and this proves the proposition. □

6.5. Classification.

THEOREM 17. *Let $(M, \Delta, \mathbf{g})$ be a left invariant three dimensional contact structure satisfying $\alpha \neq 0$, i.e. let the Fefferman metric associated to $(M, \Delta, \mathbf{g})$ be non locally conformally flat. The conformal classification of $(M, \Delta, \mathbf{g})$ is then determined by the metric invariants χ and κ; in particular, once a normalization is fixed, it coincides with the metric one.*

PROOF. Let us recall that for a $\mathfrak{solv}^+$ structure

$$\alpha = -\frac{\chi}{6}(\kappa + 8\chi).$$

Since $\chi \neq 0$ by the choice of a canonical frame (2.23) for $(M, \Delta, \mathbf{g})$, the condition $\alpha \neq 0$ implies that the coefficient $\kappa + 8\chi$ is nonzero. Normalize $(M, \Delta, \mathbf{g})$ so that $c_{12}^2 = \sqrt{\chi - \kappa} = 1$; there are then two cases to be distinguished.

If $\delta = 0$, by Proposition 15 we know that the distribution tangent to the foliation J and attached to the point $(1, 1)$ functionally depends on $\kappa + 8\chi$; moreover it is preserved by any conformal map Φ on $(M, \Delta, \mathbf{g})$, which means that Φ does not affect some function $f = f(\kappa + 8\chi)$. Since the condition $\delta = 0$ says that κ and χ are linearly dependent, this implies that the conformal classification of $(M, \Delta, \mathbf{g})$ is determined by κ, and coincides with the metric one.

If $\delta \neq 0$, Proposition 15 says that the tangent space to the foliation induced by either K or L, and attached to the point $(1, 1)$ in the $\{h_1, h_2\}$-plane, functionally depends on $\kappa + 7\chi$ and $\kappa + 8\chi$. Any conformal map Φ on $(M, \Delta, \mathbf{g})$ preserves then some independent functions $g = g(\kappa + 7\chi)$ and $f = f(\kappa + 8\chi)$. If $\alpha \neq 0$, as above we thus conclude that the conformal classification of $(M, \Delta, \mathbf{g})$ is determined by its metric invariants.

Switch now to a unimodular structure $(M, \Delta, \mathbf{g})$, and normalize it so that $\kappa^2 + \chi^2 = 1$. In this case

$$\alpha = -\frac{3}{2}\kappa\chi,$$

so that if $\alpha \neq 0$, neither κ nor χ can be equal to zero. The foliation is indexed on the $\{h_1, h_2\}$-plane by the level sets of the smooth function I, and the tangent distribution attached at the point $(1,1)$ to the foliation functionally depends on $\kappa \pm \frac{1}{2}\chi$ by Proposition 16. Since any conformal map Φ on $(M, \Delta, \mathbf{g})$ must preserve some independent functions $f = f(\chi + \frac{1}{2}\kappa)$ and $g = g(\chi - \frac{1}{2}\kappa)$, again we deduce that the conformal classification coincides with the metric one. □

7. Flat Group of Conformal Symmetries

7.1. Preliminary Observations. In this section we will work out the full computation of the conformal group $Conf(\mathbf{H}_3)$ and, in turn, of all the locally conformally flat structures found in the previous sections. We recall that the group multiplication law for $\mathbf{H}_3$ reads

$$(x_1, y_1, z_1) \star (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2 + \frac{1}{2}(x_1 y_2 - x_2 y_1))$$

and that a local frame for the Lie algebra $\mathfrak{h}_3$ is given by

$$f_1 = \frac{\partial}{\partial y} + \frac{x}{2}\frac{\partial}{\partial z}, \quad f_2 = \frac{\partial}{\partial x} - \frac{y}{2}\frac{\partial}{\partial z}, \quad f_0 = \frac{\partial}{\partial z}.$$

Return for the moment to the general setting. Assume that $X \in \mathrm{Vec}(M)$; we can associate to X a linear on fibers function $X^* \in C^\infty(T^*M)$ by:

$$X^*(\lambda) = \langle \lambda, X(q) \rangle, \quad \lambda \in T_q^*M, \quad q = \pi(\lambda).$$

Consider the Hamiltonian lift $\overrightarrow{X^*}$ of X^*, which is defined by

$$\iota_{\overrightarrow{X^*}} \sigma = -dX^*, \tag{2.24}$$

where $\sigma = ds$ denotes the symplectic form on the cotangent bundle T^*M and $s = \sum_{i=0}^2 h_i \nu^i$ denotes its tautological one form.

It turns out in this case that the lift to the cotangent bundle of the flow generated by X is nicely related to the flow generated on T^*M by $\overrightarrow{X^*}$:

Proposition 18. *Let $X \in \mathrm{Vec}(M)$ be an autonomous vector field. Then*

$$\left(e^{tX}\right)^* = e^{-t\overrightarrow{X^*}}. \tag{2.25}$$

For a proof and further details we invite the reader to consult [**AS04**], chapter 11.

Definition 9. *Given $X \in \mathrm{Vec}(M)$, we say that X is an* infinitesimal conformal sub-Riemannian symmetry *if the flow generated by X preserves the conformal class of the sub-Riemannian Hamiltonian h, i.e. it is a conformal sub-Riemannian map. In particular*

$$\Phi_t = e^{tX}$$

satisfies

$$(\Phi_t)^* h = g^t h, \quad g^t : M \to \mathbb{R}, \quad g^t \in C^\infty(M) \ \ \forall t \in \mathbb{R}.$$

Thanks to (2.25) *we have also the infinitesimal counterpart of this definition, that is*

$$\{X^*, h\} = 2\eta h, \quad \eta : M \to \mathbb{R}, \quad \eta = -\frac{1}{2}\frac{d}{dt}\bigg|_{t=0} g^t \in C^\infty(M).$$

The subgroup of conformal sub-Riemannian maps for which there holds

$$(\Phi_t)^* h \equiv h, \quad \forall t$$

are called sub-Riemannian isometries*, and their generators are called* infinitesimal sub-Riemannian isometries.

PROPOSITION 19. *Let*

$$X = \sum_{i=0}^{2} a^i f_i \in \mathrm{Vec}(M)$$

be an infinitesimal conformal sub-Riemannian symmetry. Its coefficients $a^i : M \to \mathbb{R}$ *then satisfy the following system of differential equations:*

$$\begin{cases} f_1(a^0) - a^2 = 0, \\ f_2(a^0) + a^1 = 0, \\ f_1(a^1) - c_{12}^1 a^2 = -\eta, \\ f_2(a^2) + c_{12}^2 a^1 = -\eta, \\ f_1(a^2) + f_2(a^1) + c_{12}^1 a^1 - c_{12}^2 a^2 + 2\chi a^0 = 0. \end{cases} \tag{2.26}$$

PROOF. By definition 9, X is an infinitesimal conformal sub-Riemannian symmetry if, and only if, $\{X^*, h\} = 2\eta h$.

By equation (2.24) we have

$$X^* = \sum_{i=0}^{2} a^i h_i.$$

Recall that we may interpret the smooth functions $a^i : M \to \mathbb{R}$ as constant on fibers functions on the cotangent bundle via the pullback provided by the natural projection $\pi : T^*M \to M$; by an abuse of notation we will continue to denote these functions with $a^i = \pi^*\left(a^i\right)$.

By the Leibnitz rule for the Poisson brackets and the fact that $\{h_i, h_j\} = [f_i, f_j]^*$, the equation $\{X^*, h\} = 2\eta h$ then becomes

$$\begin{aligned} \{X^*, h\} &= \left(-f_1(a^0) + a^2\right) h_0 h_1 + \left(-a^1 - f_2(a^0)\right) h_0 h_2 \\ &+ \left(-f_1(a^1) + c_{12}^1 a^2\right) h_1^2 + \left(-f_2(a^2) - c_{12}^2 a^1\right) h_2^2 \\ &+ \left(-f_1(a^2) - f_2(a^1) - c_{12}^1 a^1 + c_{12}^2 a^2 - 2\chi a^0\right) h_1 h_2 \\ &= \eta h_1^2 + \eta h_2^2. \end{aligned}$$

The conclusion is now evident. □

7.2. Computations. The first step is to characterize all the possible functions η which appear in definition 9. Since for the Lie Algebra $\mathfrak{h}_3$ all the structural coefficients are 0, it is possible to rewrite (2.26) in the equivalent way

$$\begin{cases} f_1(a^0) - a^2 = 0, \\ f_2(a^0) + a^1 = 0, \\ f_1(a^1) = -\eta, \\ f_2(a^2) = -\eta, \\ f_2(a^1) = \omega, \\ f_1(a^2) = -\omega, \end{cases} \tag{2.27}$$

where ω is some smooth scalar function defined on M.

In particular we have the following

THEOREM 20. *Assume that $\{f_0,\ f_1,\ f_2\}$ span the Lie algebra $\mathfrak{h}_3$. The system (2.27) has at least a local solution if, and only if, the smooth function $\eta : M \to \mathbb{R}$ satisfies the following set of partial differential equations:*

$$\begin{cases} 0 = f_0(f_1(\eta)), \\ 0 = f_0(f_2(\eta)), \\ 0 = f_1(f_1(\eta)), \\ 0 = f_2(f_2(\eta)), \\ 0 = f_1(f_2(\eta)) + f_2(f_1(\eta)) = 0. \end{cases} \tag{2.28}$$

PROOF. Let us assume that (2.27) has at least one solution. Then $f_1(a^0) = a^2$ and $f_2(a^0) = -a^1$. Combine with $f_1(a^1) = f_2(a^2) = -\eta$ to find

$$f_0(a^0) = [f_2, f_1](a^0) = -2\eta.$$

Imposing the exactness of $\alpha = -2\eta\nu^0 + a^2\nu^1 - a^1\nu^2$, we recover two non trivial integrability conditions

$$2f_1(\eta) + f_0(a^2) = 0,$$
$$2f_2(\eta) - f_0(a^1) = 0.$$

To guarantee the existence of the functions $a^1,\ a^2 : M \to \mathbb{R}$ in (2.27), these two differential one forms also have to be exact

$$\beta = 2f_2(\eta)\nu^0 - \eta\nu^1 + \omega\nu^2,$$
$$\gamma = -2f_1(\eta)\nu^0 - \omega\nu^1 - \eta\nu^2;$$

this leads to the following system of integrability conditions

$$\begin{cases} 0 = 3f_2(\eta) + f_1(\omega), \\ 0 = -3f_1(\eta) + f_2(\omega), \\ 0 = -2f_1(f_2(\eta)) - f_0(\eta), \\ 0 = 2f_2(f_1(\eta)) - f_0(\eta), \\ 0 = -2f_2(f_2(\eta)) + f_0(\omega), \\ 0 = 2f_1(f_1(\eta)) - f_0(\omega). \end{cases} \tag{2.29}$$

From the first two and the last two equations it follows that both the equalities

$$[f_2, f_1](\omega) = -3f_2(f_2(\eta)) - 3f_1(f_1(\eta)),$$
$$f_0(\omega) = f_2(f_2(\eta)) + f_1(f_1(\eta))$$

are true, hence $f_0(\omega) = 0$, which implies

$$f_2(f_2(\eta)) = 0, \quad f_1(f_1(\eta)) = 0. \tag{2.30}$$

Using the fourth equation in (2.29) and the commutator relation $[f_2, f_1](\eta) = f_0(\eta)$ we further obtain

$$f_1(f_2(\eta)) + f_2(f_1(\eta)) = 0. \tag{2.31}$$

From the existence assumption for $\omega : M \to \mathbb{R}$ we deduce the exactness of the differential one form

$$\delta = -3f_2(\eta)\nu^1 + 3f_1(\eta)\nu^2,$$

which yields the non trivial conditions

$$f_0(f_1(\eta)) = 0, \quad f_0(f_2(\eta)) = 0. \tag{2.32}$$

Collected together, (2.30),(2.31) and (2.32) are exactly the conditions given in (2.28).

Conversely, it is easy to repeat backwards the previous argument assuming the validity of (2.28). This completes the proof. □

COROLLARY 21. *Under the local parametrization of* $\mathfrak{h}_3$ *given by*

$$f_0 = \frac{\partial}{\partial z}, \quad f_1 = \frac{\partial}{\partial y} + \frac{x}{2}\frac{\partial}{\partial z}, \quad f_2 = \frac{\partial}{\partial x} - \frac{y}{2}\frac{\partial}{\partial z},$$

the system (2.27) *has at least a local solution if, and only if*

$$\eta = \xi_1 x + \xi_2 y + \xi_3 z + \xi_4, \quad \xi_i \in \mathbb{R}, \text{ for } i = 1, 2, 3, 4. \tag{2.33}$$

PROOF. Since $[f_1, f_0](\eta) = [f_2, f_0](\eta) = 0$, we have $f_0(f_0(\eta)) = [f_2, f_1]f_0(\eta) = 0$, that is

$$\frac{\partial^2}{\partial z^2}\eta = 0.$$

Rewrite in coordinates the conditions of Theorem 20 to deduce that all possible second derivatives of η must vanish; this is possible if, and only if, η has the form of (2.33). □

Using the linearity of the Poisson brackets, it is sufficient to analyze (2.27) just for

$$\eta = \xi_1 x, \quad \eta = \xi_2 y, \quad \eta = \xi_3 z, \quad \eta = \xi_4, \quad \xi_i \in \mathbb{R}.$$

It is well known [**FMP99**] that the algebra of sub-Riemannian isometries of the group $\mathbf{H}_3$ is four dimensional and generated by

$$F_1 = \frac{\partial}{\partial z}, \quad F_2 = \frac{\partial}{\partial x} + \frac{y}{2}\frac{\partial}{\partial z}, \quad F_3 = \frac{\partial}{\partial y} - \frac{x}{2}\frac{\partial}{\partial z}, \quad F_4 = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x};$$

the computations below extend those calculations to the whole conformal group.

COROLLARY 22. *The Lie Algebra* $\mathfrak{conf}(\mathbf{H}_3)$ *of conformal sub-Riemannian symmetries is eight dimensional. A complete set of generators is given by*

$$\begin{aligned}
\eta = 0) \quad & F_1 = \frac{\partial}{\partial z},\ F_2 = \frac{\partial}{\partial x} + \frac{y}{2}\frac{\partial}{\partial z},\ F_3 = \frac{\partial}{\partial y} - \frac{x}{2}\frac{\partial}{\partial z},\ F_4 = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x},\\
\eta = \xi_4) \quad & F_5 = -2z\frac{\partial}{\partial z} - x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y},\\
\eta = \xi_1 x) \quad & F_6 = \frac{1}{2}(3y^2 - x^2)\frac{\partial}{\partial x} + 2(z - xy)\frac{\partial}{\partial y} + \left(-xz - \frac{x^2 y}{4} - \frac{y^3}{4}\right)\frac{\partial}{\partial z},\\
\eta = \xi_2 y) \quad & F_7 = -2(z + xy)\frac{\partial}{\partial x} + \frac{1}{2}(3x^2 - y^2)\frac{\partial}{\partial y} + \left(-yz + \frac{x^3}{4} + \frac{xy^2}{4}\right)\frac{\partial}{\partial z},\\
\eta = \xi_3 z) \quad & F_8 = \left(-xz - \frac{x^2 y}{4} - \frac{y^3}{4}\right)\frac{\partial}{\partial x} + \left(-yz + \frac{xy^2}{4} + \frac{x^3}{4}\right)\frac{\partial}{\partial y} +\\
& \quad + \left(-z^2 + \frac{1}{16}(x^2 + y^2)^2\right)\frac{\partial}{\partial z}.
\end{aligned}$$

PROOF. We will go through all the details just for $\eta = \xi_3 z$. Use the equations in (2.29) to deduce

$$\frac{\partial}{\partial z}\omega = 0, \quad \frac{\partial}{\partial y}\omega = \frac{3}{2}\xi_3 y, \quad \frac{\partial}{\partial x}\omega = \frac{3}{2}\xi_3 x,$$

from which it follows

$$\omega = \frac{3}{4}\xi_3\left(x^2+y^2\right) + c^\omega, \quad c^\omega \in \mathbb{R}.$$

From here it is immediate to solve the systems for a^1, a^2 and a^0; we find

$$\begin{aligned} a^1 &= -\xi_3 yz + \frac{1}{4}\xi_3 xy^2 + \frac{1}{4}\xi_3 x^3 + c^\omega x + c^1, \quad c^1 \in \mathbb{R} \\ a^2 &= -\xi_3 xz - \frac{1}{4}\xi_3 x^2 y - \frac{1}{4}\xi_3 y^3 - c^\omega y + c^2, \quad c^2 \in \mathbb{R} \end{aligned}$$

and

$$a^0 = -\xi_3 z^2 - \frac{1}{8}\xi_3 x^2y^2 - \frac{1}{16}\xi_3\left(x^4+y^4\right) - \frac{c^\omega}{2}\left(x^2+y^2\right) + c^2 y - c^1 x + c^0, \quad c^0 \in \mathbb{R}.$$

Any $X \in \mathrm{Vec}(M)$ has the form $X = \sum_{i=0}^2 a^i f_i$; collecting similar terms and simplifying, we find that

$$\begin{aligned} X = {} & c^0\frac{\partial}{\partial z} + c^1\left(\frac{\partial}{\partial y} - \frac{x}{2}\frac{\partial}{\partial z}\right) + c^2\left(\frac{\partial}{\partial x} + \frac{y}{2}\frac{\partial}{\partial z}\right) + c^\omega\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \\ & + \xi_3\left(\left(-xz - \frac{x^2y+y^3}{4}\right)\frac{\partial}{\partial x} + \left(-yz + \frac{xy^2+x^3}{4}\right)\frac{\partial}{\partial y} + \right. \\ & \left. + \left(-z^2 + \frac{1}{16}(x^2+y^2)^2\right)\frac{\partial}{\partial z}\right). \end{aligned}$$

We already know that the first four terms are the generators for the sub-Riemannian isometries; the non trivial solution to the system (2.27) is then just the last one, which we call F_8. The explicit computation shows that it has the correct expression indicated in the statement of the theorem. $\square$

REMARK 18. It is known [**ABB12**] that the triplet (x, y, z), endowed with the weight w so that

$$w(x) = w(y) = 1, \;\; w(z) = 2, \;\; w\left(\frac{\partial}{\partial x}\right) = w\left(\frac{\partial}{\partial y}\right) = -1, \;\; w\left(\frac{\partial}{\partial z}\right) = -2,$$

together with the local parametrization

$$f_1 = \frac{\partial}{\partial y} + \frac{x}{2}\frac{\partial}{\partial z}, \quad f_2 = \frac{\partial}{\partial x} - \frac{y}{2}\frac{\partial}{\partial z}, \quad f_0 = \frac{\partial}{\partial z},$$

gives a system of *privileged coordinates* for the Heisenberg Group. The Lie algebra of the conformal group, $\mathfrak{conf}(\mathbf{H}_3)$, becomes then a *graded* Lie algebra, that is

$$\mathfrak{conf}(\mathbf{H}_3) = \mathfrak{h}^{-2} \oplus \mathfrak{h}^{-1} \oplus \mathfrak{h}^0 \oplus \mathfrak{h}^1 \oplus \mathfrak{h}^2,$$

where the superscript of each $\mathfrak{h}^i$ corresponds to the weight of its generators.

7.3. Tanaka Prolongation. $\mathfrak{conf}(\mathbf{H}_3)$ can be described purely algebraically applying Tanaka's prolongation. For a detailed exposition of this topic one can refer to [**Zel09**]; we recall here just how to proceed with this construction.

Let $\mathfrak{h}^0$ be the subalgebra of the derivations of $\mathfrak{h}_3$ generated by rotations and dilations; $\mathfrak{h}_3 \oplus \mathfrak{h}^0$ becomes endowed with the structure of a graded Lie algebra where

$$[f, v] = f(v), \quad \forall f \in \mathfrak{h}^0, \; v \in \mathfrak{h}_3.$$

Denote, for any positive integer $l > 0$,

$$\mathfrak{h}^l = \Big\{ f \in \bigoplus_{i<0} \mathrm{Hom}(\mathfrak{h}^i, \mathfrak{h}^{i+l}) : \\ f([v_1, v_2]) = [f(v_1), v_2] + [v_1, f(v_2)], \forall v_1, v_2 \in \mathfrak{h}_3 \Big\}, \tag{2.34}$$

and define inductively the brackets via the position

$$[f_1, f_2]v = [f_1(v), f_2] + [f_1, f_2(v)] \quad \forall v \in \mathfrak{h}_3,\ f_1 \in \mathfrak{h}^l,\ f_2 \in \mathfrak{h}^k. \tag{2.35}$$

Using (2.34) and (2.35), one can show that the structure of $\mathfrak{conf}(\mathbf{H}_3)$ is as follows

$$\mathfrak{conf}(\mathbf{H}_3) = \mathfrak{h}^{-2} \oplus \mathfrak{h}^{-1} \oplus \mathfrak{h}^0 \oplus \mathfrak{h}^1 \oplus \mathfrak{h}^2,$$

where

$$\mathfrak{h}^{-2} = \mathrm{span}\{f_0\}, \qquad \mathfrak{h}^{-1} = \mathrm{span}\{f_2, f_1\}, \quad \mathfrak{h}^0 = \mathrm{span}\{\Lambda_1^0, \Lambda_2^0\}$$

$$\mathfrak{h}^1 = \mathrm{span}\{\Lambda_1^1, \Lambda_2^1\}, \quad \mathfrak{h}^2 = \mathrm{span}\{\Lambda\};$$

and the nontrivial brackets are

$$\begin{array}{lll}
[f_2, f_1] = f_0, & [\Lambda_2^0, f_0] = 0, & [\Lambda, f_1] = -\Lambda_1^1, \\
[f_2, f_0] = 0, & [\Lambda_1^1, f_2] = \Lambda_1^0, & [\Lambda, f_0] = 2\Lambda_1^0, \\
[f_1, f_0] = 0, & [\Lambda_1^1, f_1] = 3\Lambda_2^0, & [\Lambda_1^1, \Lambda_1^0] = \Lambda_1^1, \\
[\Lambda_1^0, f_2] = f_2, & [\Lambda_1^1, f_0] = -2f_1, & [\Lambda_1^1, \Lambda_2^0] = -\Lambda_2^1, \\
[\Lambda_1^0, f_1] = f_1, & [\Lambda_2^1, f_2] = -3\Lambda_2^0, & [\Lambda_2^1, \Lambda_1^0] = \Lambda_2^1, \\
[\Lambda_1^0, f_0] = 2f_0, & [\Lambda_2^1, f_1] = \Lambda_1^0, & [\Lambda_2^1, \Lambda_2^0] = \Lambda_1^1, \\
[\Lambda_2^0, f_2] = f_1, & [\Lambda_2^1, f_0] = 2f_2, & [\Lambda_1^1, \Lambda_2^1] = 2\Lambda, \\
[\Lambda_2^0, f_1] = -f_2, & [\Lambda, f_2] = \Lambda_2^1, & [\Lambda, \Lambda_1^0] = 2\Lambda.
\end{array}$$

Remark 19. The isomorphism between the Lie algebra found in corollary 22 and the one above is given by

$$\begin{array}{lll}
F_1 \mapsto f_0, & F_4 \mapsto \Lambda_2^0, & F_7 \mapsto \Lambda_2^1, \\
F_2 \mapsto f_2, & F_5 \mapsto \Lambda_1^0, & 2F_8 \mapsto -\Lambda, \\
F_3 \mapsto -f_1, & F_6 \mapsto -\Lambda_1^1. &
\end{array}$$

Theorem 23.

$$\mathfrak{conf}(\mathbf{H}_3) \cong \mathfrak{su}(2,1).$$

Proof. Denoting with E_{ij} the 3×3 real-valued matrix whose only nonzero entry is $e_{ij} = 1$, we have that the claimed isomorphism is

$$\begin{array}{lll}
f_2 \mapsto E_{12} + E_{23}, & \Lambda_1^0 \mapsto E_{11} - E_{33}, & \Lambda_2^1 \mapsto -iE_{21} + iE_{32}, \\
f_1 \mapsto -iE_{12} + iE_{23}, & 3\Lambda_2^0 \mapsto -iE_{11} + 2iE_{22} + -iE_{33}, & \Lambda \mapsto iE_{31}, \\
f_0 \mapsto 2iE_{13}, & \Lambda_1^1 \mapsto -E_{21} - E_{32}. &
\end{array}$$

□

Remark 20. By virtue of the previous discussion, we deduce that the local conformal flatness of the Fefferman metric associated to a sub-Riemannian three dimensional left invariant contact structure identifies those cases where the conformal group has the maximal possible dimension. In particular there is a dimension gap between the flat and the non-flat case; in the first case the conformal group has dimension eight, while in the latter it has dimension three, with no intermediate possibilities in between.

CHAPTER 3

Homotopy Properties of Horizontal Path Spaces and a Theorem of Serre in sub-Riemannian Geometry

1. Introduction

We study homotopy properties of the set of those curves on a manifold M whose velocities are constrained in a nonholonomic way (these curves are called *horizontal*). The nonholonomic constraint is made explicit by requiring that the curves should be tangent to a totally non-integrable distribution (for example a contact distribution, whose horizontal curves are called *legendrian*). More generally we will allow *affine* constraints, by considering a set of vector fields $\mathcal{F} = \{X_0, X_1, \dots, X_d\}$ and defining a horizontal curve $\gamma : I = [0, 1] \to M$ to be an *absolutely continuous* curve (hence differentiable almost everywhere) solving the equation:

$$\dot{\gamma} = X_0(\gamma) + \sum_{i=1}^{d} u_i X_i(\gamma), \quad \gamma(0) = x \tag{3.1}$$

for functions $u_1, \dots, u_d$ called *controls* ($x \in M$ is a point that we fix from the very beginning).

The vector field X_0 is special (it plays the role of a "drift") and in many interesting cases, like the sub-Riemannian, it is assumed to be zero; the remaining vector fields satisfy the totally nonintegrable *Hörmander* condition: a finite number of their iterated brackets should span the whole tangent space TM (this is also called the *bracket generating* condition).

The regularity we impose on the controls determines the topology on the space Ω of all horizontal curves (called also *trajectories*). In this paper we will assume $u = (u_1, \dots, u_d) \in L^p(I, \mathbb{R}^d)$ for some $1 < p < \infty$ (thus we consider the $W^{1,p}$ topology on the space of trajectories). The correspondence between a curve and its controls defines local coordinates on Ω, which in turn becomes a Banach manifold modeled on $L^p = L^p(I, \mathbb{R}^d)$ (in fact this manifold is just the open subset of L^p consisting of all controls whose corresponding trajectory is defined on the whole interval I, see the Appendix of this paper or [**Mon02**] for more details); as a byproduct of this identification we will often replace a curve with the d-tuple of controls describing it in local coordinates.

The *Endpoint map* is the map that associates to each trajectory its final point:

$$F : \Omega \to M \quad \gamma \mapsto \gamma(1).$$

This map is differentiable (smooth in the $W^{1,2}$ case [**ABB**]), and the set:

$$\Omega(y) = F^{-1}(y)$$

with the induced topology coincides with the set of horizontal curves joining x to y. In the Riemannian case, these spaces are well understood and their topological properties are related to those of the manifold M via the *path fibration* (see [**BT82, Hat02**]), which in our setting we discuss below.

The uniform convergence topology on Ω has been studied in [**Sar91**] and the $W^{1,1}$ in [**DR12**]. For the scopes of calculus of variations the case $W^{1,p}$ with $p > 1$ is especially interesting as the analysis becomes more pleasant: for example the p-th power of the L^p norm becomes a C^1 function and one can apply classical techniques from critical point theory to many problems of interest. Also, it is worth recalling that already in the sub-Riemannian case not all topologies on Ω are equivalent a priori: for example in the $W^{1,\infty}$ case the so-called *rigidity* phenomenon appear: some curves might be isolated (up to reparametrization) in the $W^{1,\infty}$ topology [**BH93**].

The key property for studying the topology of horizontal path spaces is the homotopy lifting property for the Endpoint map. Our first result generalizes the main results from [**DR12, Sar91**], proving that there exists $p_c > 1$ (depending on $\mathcal{F}$) such that Endpoint map is a *Hurewicz fibration* for the $W^{1,p}$ topology for all $1 \leq p < p_c$ (i.e. F has the homotopy lifting property with respect to any space for these topologies).

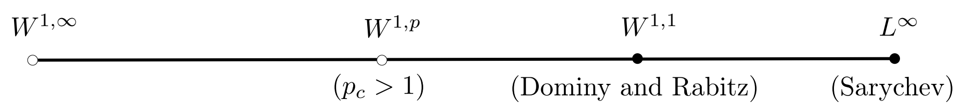


FIGURE 1. A picture of the continuous inclusions (from left to right) of the various $W^{1,p}([0,1])$ spaces.

THEOREM (The Endpoint map is a Hurewicz fibration). *There exists an interval* $[1, p_c) \subseteq [1, \infty)$ *(depending on* $\mathcal{F}$*), such that if* $p \in [1, p_c)$ *the Endpoint map* $F : \Omega \to M$ *is a Hurewicz fibration for the* $W^{1,p}$ *topology on* Ω*. Moreover if* $X_0 = 0$ *then* $p_c = \infty$*.*

It is remarkable that the sub-Riemannian case ($X_0 = 0$) has the Hurewicz fibration property for all $1 \leq p < \infty$, as in general if $X_0 \neq 0$ the Endpoint map can fail to have the homotopy lifting property for some finite $p < \infty$, as shown in the next example [**AL10**].

EXAMPLE 1. Consider $M = \mathbb{R}^2$ with coordinates (x_1, x_2) and:

$$X_0 = x_1^2 \partial_{x_2}, \quad X_1 = \partial_{x_1}, \quad X_2 = x_1^k \partial_{x_2}, \quad k \geq 3$$

We consider on Ω the function $u \mapsto J(u) = \|u\|_2^2$, which is continuous for every $p \geq 2$ for the $W^{1,p}$ topology. Let us also consider the function $c_1 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$:

$$c_1(x, y) = \inf\{J(\gamma) \in \Omega \,|\, \gamma(0) = x,\, \gamma(1) = y\}.$$

In [**AL10**, Proposition 2.1] it is proved that there exists $K > 0$ such that for all $w \in \mathbb{R}$ and all $z < 0$:

$$c_1((0, w), (0, w)) = 0 \quad \text{and} \quad c_1((0, w), (0, z)) \geq K.$$

Consider now the path $g_s = (0, -s)$ and let $u_0 \in \Omega$ be a lift for g_0 (i.e. $F(u_0) = g_0$). Now this path (a homotopy of inclusions of a single point) cannot be lifted: an existence of such a lift would be a continuous path u_s on Ω with $u_s \in \Omega(g_s)$, and in particular:

$$\lim_{s \to 0} J(u_s) = 0,$$

which contradicts the fact that $J|_{\Omega(g_s)} \geq K > 0$ for all $s > 0$.

Our proof of the previous theorem is much inspired from [**Sar91, DR12**] and in fact consists in a simple (but important) modification of the proof from [**Sar91**]. This theorem has important consequences for the topology of fibers of the Endpoint map.

THEOREM (The homotopy type of the fiber). *Any two fibers of the Endpoint map, endowed with the $W^{1,p}$ topology ($p < p_c$), are homotopy equivalent. Moreover each fiber $\Omega(y)$ has the homotopy type of a CW-complex and its inclusion in the ordinary space of curves (i.e. curves without the nonholonomic constraint) is a homotopy equivalence.*

We should stress at this point that the space $\Omega(y)$ might be highly singular, because of the possible existence of *abnormal curves* (curves γ such that the differential $d_\gamma F$ is not a submersion). The existence of these curves *is not* excluded in our setting. It is remarkable that even if abnormal curves might influence the differential structure of $\Omega(y)$, still its homotopy remains very controlled: *any* two fibers of F are homotopy equivalent, regardless them being singular or regular, and what is known for the homotopy of the standard loop space can be deduced also for our horizontal one.

COROLLARY (Some topological implications). *For every $k \in \mathbb{N}$, every $1 \leq p < p_c$ and every $y \in M$ the following isomorphism between homotopy groups holds for the $W^{1,p}$ topology:*

$$\pi_k(\Omega(y)) \simeq \pi_{k+1}(M).$$

Moreover if M is compact and simply connected, then the Lusternik-Schnirelmann category of the space $\Omega(y)$ is infinite.

Once there is some information available for the topology of $\Omega(y)$, it can be used to study critical points of functionals, the classical example being the study of geodesics between two points. A celebrated theorem of Serre [**Ser51**] states that if a Riemannian manifold M is compact, then every two points are joined by infinitely many geodesics; the proof of this theorem essentially uses the topology of $\Omega(y)$ to force the existence of critical points of the Energy functional, which in the Riemannian case are exactly geodesics.

More generally one can study critical points of the p-Energy $J_p : u \mapsto \|u\|_p^p$ on $\Omega(y)$ for affine control systems on *regular fibers* $\Omega(y)$: as long as $1 < p < p_c$ this function is C^1 (Lemma 32) and when restricted to $\Omega(y)$ it satisfies the Palais-Smale condition (Proposition 33). These two properties allow to use classical results to force the existence of critical points.

THEOREM (On the critical points for the p-Energy). *Let y be a regular value for the Endpoint map of the control system* (3.1)*, $1 < p < p_c$ and consider $f = J_p|_{\Omega(y)}$. If the base manifold M is compact then f has infinitely many critical points.*

As a corollary, we thus obtain a sub-Riemannian version of the Serre's theorem: given x and any regular point y for the Endpoint map on a compact sub-Riemannian manifold there are infinitely many geodesics connecting them. In some cases (e.g. contact or fat distributions) the assumption of y being a regular value may be dropped: in these situations there are no abnormal curves other than the trivial ones, and our arguments are essentially not affected (here the fact that $\Omega(y)$ has the homotopy type of a CW-complex plays a crucial role, see the end of the proof of Theorem 39).

1.1. Related Work. The problem of understanding the topology of the space of maps with some restrictions on their differential goes back to the works on immersions of S. Smale [**Sma58**], for the case of curves on a manifold the author considers spherical-type constraints on the velocities (i.e. immersions and regular homotopies). Hurewicz properties for Endpoint maps of affine control systems were studied first by A. V. Sarychev [**Sar91**] for the uniform convergence topology and by J. Dominy and H. Rabitz [**DR12**] for the $W^{1,1}$ topology. The quantitative study of the interaction between the topology of the horizontal loop space and the set of geodesics was initiated by the second author together with A. Agrachev and A. Gentile in [**AGL15**]. In the contact case a "local" version of Serre's theorem was investigated by the second author and L. Rizzi in [**LR**] (the authors perform an asymptotic count of the number of geodesics between two point on a contractible contact manifold, using the relation between a sub-Riemannian manifold and its nilpotent approximation).

1.2. Structure of the Paper. Section 2 is devoted to the proof of the Hurewicz fibration property (Theorem 27): the crucial ingredient is the construction of a continuous cross-section for the Endpoint map (Proposition 25). The topological implications are discussed in Section 2.1. In section 3 we study critical points of geometric costs: the Palais-Smale property is proved in Proposition 33 and applications via Lusternik-Schnirelmann method are discussed in Section 3.2. The sub-Riemannian case is discussed in Section 4. The Appendix contains some additional technical results, mostly known to experts.

2. Homotopy Properties of the Endpoint Map

LEMMA 24. *Let $0 < \beta < \frac{p}{p-1}$ and for every $j = 1, \ldots, N$ define the map $\rho_j : \mathbb{R}^N \to L^p([0,\infty))$ by: $\rho_j(r) = 0$ if $r_j = 0$ and $\rho_j(r) = \chi_j r_j |r_j|^{-\beta}$ otherwise (χ_j is the characteristic function of the interval $[|r_{j-1}|^\beta, |r_{j-1}|^\beta + |r_j|^\beta]$ and $r_0 = 0$). Then the map ρ_j is continuous.*

PROOF. The only needed verification is continuity at zero:

$$\lim_{r_j \to 0} \|\chi_j r_j |r_j|^{-\beta}\|_p = \lim_{r_j \to 0} \left(\int_{|r_{j-1}|^\beta}^{|r_{j-1}|^\beta + |r_j|^\beta} \left| r_j |r_j|^{-\beta} \right|^p dt \right)^{1/p}$$
$$= \lim_{r_j \to 0} |r_j|^{\frac{\beta + p - \beta p}{p}} = 0$$

since $\beta + p - \beta p > 0$. □

PROPOSITION 25 (The cross-section). *Given the manifold M and the family of vector fields $\mathcal{F}$, there exists an interval $[1, p_c) \subset [1, \infty)$ such that for every $1 \le p < p_c$ every point in M has a neighborhood W and a continuous map:*

$$\hat{\sigma} : W \times W \to L^p([0,\infty), \mathbb{R}^d) \times \mathbb{R}$$
$$(x, y) \mapsto (\sigma(x,y), T(x,y))$$

such that $F_x^{T(x,y)}(\sigma(x,y)) = y$ and $\hat{\sigma}(x,x) = (0,0)$ for every $x, y \in W$. Moreover, if $X_0 = 0$ then $p_c = \infty$.

PROOF. We first work out the case $X_0 = 0$ and $p > 1$ (the case $p = 1$ and $X_0 = 0$ is a special case of [**DR12**, Lemma 1], whose notation we follow closely). Given the vector fields $\{Y_1, \ldots, Y_k\}$, define inductively $Q^1(Y_1) = e^{Y_1}$ and:

$$Q^\nu(Y_1, \ldots, Y_\nu) = e^{Y_\nu} \circ Q^{\nu-1}(Y_1, \ldots, Y_{\nu-1}) \circ e^{-Y_\nu} \circ (Q^{\nu-1}(Y_1, \ldots, Y_{\nu-1}))^{-1}, \quad \nu \ge 1.$$

Given a real number r we define also:

$$P^\nu(Y_1, \ldots, Y_\nu, r) = Q^\nu(rY_1, \ldots, rY_\nu).$$

It follows from the Baker-Campbell-Hausdorff formula that, for r sufficiently small,

$$P^\nu(Y_1, \ldots, Y_\nu, r^{1/\nu}) = e^{r\, \mathrm{ad}Y_\nu \ldots \mathrm{ad}Y_2 Y_1 + \text{higher order terms in } r}.$$

The bracket generating condition on $\mathcal{F}$ implies that (see [**Jea14**, Section 2.1] or the proof of [**DR12**, Lemma 1]) every point in M has a neighborhood W and a continuous[1] map $\phi : W \times W \to$

[1]The k-th component of $\phi = (\phi_1, \ldots, \phi_n)$ is the ν_k-th root of a C^1 function.

$\mathbb{R}^n$ such that $\phi(x,x)=0$ for all $x \in W$ and:

$$\left(\prod_{k=1}^{n} P^{\nu_k}(X_{k_1},\dots,X_{k_{\nu_k}},\phi_k(x,y))\right)(x)=y \quad \forall x,y\in W. \tag{3.2}$$

Now we notice that the product in (3.2) can be written as:

$$\left(\prod_{k=1}^{n} P^{\nu_k}(X_{k_1},\dots,X_{k_{\nu_k}},\phi_k(x,y))\right)=\prod_{j=1}^{N} e^{\phi_{a_j}(x,y)X_{b_j}}$$

where N is a given number and $a_j, b_j \in \{1,\dots,d\}$ for $j=1,\dots,N$ (these numbers are fixed and depend on the neighborhood W only).

Given $p>1$ choose β satisfying the hypothesis of Lemma 24. Using the notation of Lemma 24 we can now interpret $y=(\prod_{j=1}^{N} e^{\phi_{k_j}(x,y)X_{k_j}})(x)$ as the solution at time:

$$T(x,y)=\sum_{j=1}^{N}|\phi_{k_j}(x,y)|^{\beta}$$

of the control problem with initial datum $y(0)=x$ and control:

$$\sigma(x,y)=\left(\sum_{\{j\,|\,k_j=1\}}\rho_j(\phi_{k_j}(x,y)),\ \sum_{\{j\,|\,k_j=2\}}\rho_j(\phi_{k_j}(x,y)),\dots,\ \sum_{\{j\,|\,k_j=d\}}\rho_j(\phi_{k_j}(x,y))\right).$$

By Lemma 24 it follows that the map $\hat{\sigma}=(\sigma,T)$ defined in this way is continuous: each component is the sum of compositions of continuous functions ($T(x,y)$ is continuous since $\beta>0$) and $\hat{\sigma}(x,x)=(0,0)$.

For the case $X_0\neq 0$ we notice that the proof of [**DR12**, Lemma 1] produces indeed the continuity of the cross section for some $1<p<p_c$ (as we will see, a lower bound for p_c in this case is given by $\sigma/(\sigma-1)$, where σ is the step of the distribution $\mathcal{F}$). We simply check the needed details. The sequence of exponentials (3.2) now has to be replaced with [**DR12**, Equation 6.a] (using the same notation as the mentioned paper):

$$\left(\prod_{k=1}^{n} R^{\nu_k}(X_0,X_{k_1},\dots,X_{k_{\nu_k}},\pm\phi_{k_j}(x,y),\phi_{k_j}(x,y),\dots,\phi_{k_j}(x,y))\right). \tag{3.3}$$

The construction in [**DR12**] works in such a way that given $\alpha>\nu_k/2$, using BCH formula, R^{ν_k} can be written as the exponential of a series of terms from $\{\phi_{k_1}^{2\alpha}X_0,\dots,\phi_{k_\nu}^{2\alpha}X_0,\phi_{k_1}X_{k_1},\dots,\phi_{k_\nu}X_{k_\nu}\}$ and their Lie brackets. We choose thus $\alpha>\sigma/2$ which guarantees $\alpha>\nu_k/2$ for all $k=1,\dots,n$. The product in (3.3) can thus be regarded as the solution at time $T=\sum_j \nu_k\phi_{k_j}^{2\alpha}$ of a control problem with initial datum $y(0)=x$ and locally constant controls $\sigma=(\sigma_1,\dots,\sigma_d)$ taking values on an interval of length $\phi_{k_j}^{2\alpha}$. The continuity of the final time T follows from the fact that $\alpha>0$; for the continuity of the corresponding σ we argue as in [**DR12**, Appendix C]. Each component of σ is the concatenation of some fixed number of locally constant controls (some of them can possibly be zero) each one defined on an interval of length $\phi_{k_j}^{2\alpha}$ and taking a value proportional to $\phi_{k_j}^{1-2\alpha}$. Then it is enough to check the continuity of this control at zero for the L^p-topology. If we choose $p<\frac{2\alpha}{2\alpha-1}$ then:

$$\lim_{\phi_{k_j}\to 0}\int_c^{c+\phi_{k_j}^{2\alpha}}\left|\phi_{k_j}^{1-2\alpha}\right|^p dt=\lim_{\phi_{k_j}\to 0}\phi_{k_j}^{2\alpha+p-2p\alpha}=0.$$

(Notice in particular that, because of the way we chose α, a lower bound for p_c is given by $\sigma/(\sigma-1)$.) □

PROPOSITION 26 (Rescaled concatenation). *Let $p \in [1,\infty)$, then the map $\mathcal{C} : L^p(I) \times L^p([0,+\infty)) \times \mathbb{R} \to L^p(I)$ defined below is continuous:*

$$\mathcal{C}(u,v,T)(t) = \begin{cases} (T+1)u(t(T+1)) & 0 \le t \le \dfrac{1}{T+1} \\ (T+1)v((T+1)t-1), & \dfrac{1}{T+1} < t \le 1. \end{cases}$$

Moreover (extending the definition componentwise to controls with value in $\mathbb{R}^d$) we also have $F_x^{1+T}(u*v) = F_x^1(\mathcal{C}(u,v,T))$ for every $x \in M$ (here $u*v$ denotes the usual concatenation).

PROOF. As $L^p(I) \times L^p([0,+\infty)) \times \mathbb{R}$ is a metric space, it is sufficient to prove that if $(u_k, v_k, T_k) \to (u,v,T)$, then $\|\mathcal{C}(u_k,v_k,T_k) - \mathcal{C}(u,v,T)\|_p \to 0$.

Assume for simplicity that $T_k \ge T$ (we can split the sequence $\{T_k\}_{k\in\mathbb{N}}$ into two monotone subsequences and work the case $T_k \le T$ separately, it is completely analogous). Start with:

$$\begin{aligned} &\|\mathcal{C}(u_k,v_k,T_k) - \mathcal{C}(u,v,T)\|_p^p \\ &= \int_0^{1/(T_k+1)} |(T_k+1)u_k(t(T_k+1)) - (T+1)u(t(T+1))|^p dt \\ &+ \int_{1/(T_k+1)}^{1/(T+1)} |(T_k+1)v_k(t(T_k+1)-1) - (T+1)u(t(T+1))|^p dt \\ &+ \int_{1/(1+T)}^{1} |(T_k+1)v_k(t(T_k+1)-1) - (T+1)v(t(T+1)-1)|^p dt. \end{aligned} \tag{3.4}$$

Fix $\varepsilon > 0$ and let g be a smooth function compactly supported on $[0,3/2)$ such that $\|g-u\|_p \le \varepsilon$. Observe that for k sufficiently large we have $\|u_k - g\|_p \le \|u-u_k\|_p + \varepsilon \le 2\varepsilon$. We can bound the first integral in (3.4) as:

$$\begin{aligned} &\int_0^{1/(T_k+1)} |(T_k+1)u_k(t(T_k+1)) - (T+1)u(t(T+1))|^p dt \\ &\le 2^{2(p-1)}\Big(\int_0^{1/(T_k+1)} |(T_k+1)u_k(t(T_k+1)) - (T_k+1)g(t(T_k+1))|^p dt + \\ &\qquad \int_0^{1/(T_k+1)} |(T_k+1)g(t(T_k+1)) - (T+1)g(t(T+1))|^p dt + \\ &\qquad \int_0^{1/(T_k+1)} |(T+1)g(t(T+1)) - (T+1)u(t(T+1))|^p dt \Big) \\ &\le 2^{2(p-1)}\Big(|T_k+1|^{p-1}\|u_k-g\|_p^p + |T+1|^{p-1}\|u-g\|_p^p + \\ &\qquad \int_0^{1/(T_k+1)} |(T_k+1)g(t(T_k+1)) - (T+1)g(t(T+1))|^p dt \Big). \end{aligned} \tag{3.5}$$

Since g is uniformly continuous in $[0,1]$, the last integral in (3.5) can also be made as small as we wish as $k \to \infty$ as it is evident from:

$$\int_0^{1/(T_k+1)} |(T_k+1)g(t(T_k+1)) - (T+1)g(t(T+1))|^p dt$$
$$\leq 2^{p-1}\Big(\int_0^{1/(T_k+1)} |(T_k+1)g(t(T_k+1)) - (T_k+1)g(t(T+1))|^p dt +$$
$$\int_0^{1/(T_k+1)} |(T_k+1)g(t(T+1)) - (T+1)g(t(T+1))|^p dt\Big).$$

The third integral in (3.4) is formally the same as the one just handled; a similar reasoning proves that it goes to zero as $k \to \infty$. We are left to deal with the middle one. In this case as $k \to \infty$ by the dominated convergence theorem we have both

$$\int_{1/(T_k+1)}^{1/(T+1)} |(T_k+1)v_k(t(T_k+1)-1)|^p dt = |T_k+1|^{p-1} \int_0^{(T_k+1)/(T+1)-1} |v_k(z)|^p dz \to 0$$

and

$$\int_{1/(T_k+1)}^{1/(T+1)} |(T+1)u(t(T+1))|^p dt = |T+1|^{p-1} \int_{(T+1)/(T_k+1)}^1 |u(z)|^p dz \to 0.$$

Finally this yields:

$$\int_{1/(T_k+1)}^{1/(T+1)} |(T_k+1)v_k(t(T_k+1)-1) - (T+1)u(t(T+1))|^p dt$$
$$\leq 2^{p-1}\left(|T_k+1|^{p-1} \int_0^{(T_k+1)/(T+1)-1} |v_k(z)|^p dz + |T+1|^{p-1} \int_{(T+1)/(T_k+1)}^1 |u(z)|^p dz \right),$$

and with this we can eventually conclude that:

$$\lim_{k\to\infty} \|\mathcal{C}(u_k, v_k, T_k) - \mathcal{C}(u, v, T)\|_p^p = 0.$$

□

2.1. The Hurewicz Fibration Property and its Consequences.

THEOREM 27. *There exists an interval*[2] $[1, p_c) \subseteq [1, \infty)$, *such that if* $p \in [1, p_c)$ *the Endpoint map* $F : \Omega \to M$ *is a Hurewicz fibration for the* $W^{1,p}$ *topology on* Ω. *Moreover if* $X_0 = 0$ *then* $p_c = \infty$.

REMARK 21. In general the family of vector fields $\{X_1, \ldots, X_d\}$ generating the distribution cannot be chosen such that $d = \text{rank}(\Delta)$ (some topological obstructions might occur), unless we restrict to a small contractible neighborhood in M.

The correspondence $A : L^2(I, \mathbb{R}^d) \to \Omega$ associating to a control its trajectory might not be injective, but still it is a Hurewicz fibration: the fibers of this map are convex sets and the map $\mu : \Omega \to L^2(I, \mathbb{R}^d)$ giving the *minimal control* [**ABB**] is a continuous section of this fibration (the reader is referred to [**LM**] for a detailed discussion of this map). In particular, the Hurewicz fibration property for $F \circ A$ implies the Hurewicz fibration property for F and we can reduce to study the case $F : L^p(I, \mathbb{R}^d) \to M$ (this is the definition we considered, using the control system in (3.1)).

[2] Depending on $(M, X_0, X_1, \ldots, X_d)$.

Proof. Recall that *Hurewicz fibration* means that F has the homotopy lifting property with respect to every space Z. By Hurewicz uniformization theorem [**Hur55**], it is enough to show that the homotopy lift property holds locally, i.e. every point $x \in M$ has a neighborhood W such that $F|_{F^{-1}(W)}$ has the homotopy lifting property with respect to any space.

The case $p = 1$ is proved in [**DR12**], thus let $1 < p < p_c$, W and $\hat{\sigma}$ be given as in Proposition 25. Consider a continuous map $g : Z \times I \to W$ and a lift $\tilde{g}_0 : Z \to \Omega$ such that $F(\tilde{g}_0(z)) = g(z, 0)$ for all $z \in Z$. We define the lifting homotopy $\tilde{g} : Z \times I \to \Omega$ by:

$$\tilde{g}(z, s) = \mathcal{C}(\tilde{g}_0(z), \underbrace{\sigma(g(z, 0), g(z, s)), T(g(z, 0), g(z, s))}_{\hat{\sigma}(g(z,0),g(z,s))})$$

(here $\mathcal{C}$ is defined as in Proposition 26 componentwise).

The defined function $\tilde{g}$ is the composition of continuous functions (by Propositions 25 and 26). Moreover by the second assertions in Propositions 25 and 26:

$$F(\tilde{g}(z, s)) = g(z, s) \quad \forall (z, s) \in Z \times I,$$

which proves the claim. □

2.2. The Homotopy Type of the Fibers. As a consequence of Theorem 27 all fibers of F (even the singular fibers) have the same homotopy type [**Spa94**]. Moreover, by the long exact homotopy sequence of Hurewicz fibrations [**Spa94**] one also obtains the following isomorphisms between homotopy groups:

$$\pi_k(\Omega(y)) \simeq \pi_{k+1}(M) \quad \forall k \geq 0. \tag{3.6}$$

In the case the domain of the Hurewicz fibration is contractible we can be more precise about the homotopy type of the fiber.

Theorem 28. *For every $p < p_c$ and $y \in M$ the space $\Omega(y)$ with the $W^{1,p}$ topology has the homotopy type of a CW-complex. In particular the inclusion $\Omega(y) \hookrightarrow \Omega(y)_{std}$ in the standard loop space with the $W^{1,p}$ topology is a homotopy equivalence.*

Proof. First we recall that given the Hurewicz fibration $F : \mathcal{U} \to M$ (in fact any Hurewicz fibration with Ω contractible), then any two fibers are homotopy equivalent to:

$$\Omega M = \{\text{loop spaces in } M \text{ based at } x \text{ with the compact-open topology}\}.$$

The Hurewicz fibration condition is indeed equivalent [**Arn72**] to the existence of a map:

$$\lambda : \{(u, \omega) \in \mathcal{U} \times M^I \mid F(u) = \omega(0)\} \to \mathcal{U}^I$$

where[3] the map λ satisfies:

$$\lambda(u, \omega)(0) = u \quad \text{and} \quad F(\lambda(u, \omega)(t)) = \omega(t).$$

The map $\eta : \Omega M \to F^{-1}(y)$ defined by $\eta(\omega) = \lambda(x, \omega)(1)$ is proved to be a homotopy equivalence in [**DR12**, Lemma 2].

The inclusion $i : \Omega(x)_{\text{std}} \hookrightarrow \Omega M$ is a weak homotopy equivalence: the corresponding Hurewicz fibrations of Endpoint maps for $\Omega(x)_{\text{std}}$ with the $W^{1,p}$ and ΩM with the compact open topology give rise to two long exact sequence of homotopy groups; the map i induces an isomorphism between these long exact sequences.

The space ΩM has the homotopy type of a CW-complex [**Mil59**, Corollary 2] and $\Omega(x)^{1,p}_{\text{std}}$ also have the homotopy type of a CW-complex, since it is a Banach manifold modeled on a

[3] for a topological space X we denoted by X^I the space of paths in $\omega : I \to X$ endowed with the compact-open topology.

metrizable space. In particular [**Mil59**, Lemma 1] the weak homotopy equivalence $\Omega(x)_{\mathrm{std}} \hookrightarrow \Omega M$ is indeed a homotopy equivalence.

Finally, $\Omega(y)$ has the homotopy type of a CW-complex, since:

$$\Omega(y) \sim \Omega(x) \sim \Omega M \sim \Omega(x)_{\mathrm{std}}$$

(the first homotopy equivalence follows from the fact that all fibers of a Hurewicz fibration have the same homotopy type) and consequently:

$$\Omega(y) \hookrightarrow \Omega(y)_{\mathrm{std}} \quad \text{is a homotopy equivalence.}$$

□

Corollary 29. *If the base manifold M is compact and simply connected, then for every $p < p_c$ (where p_c is given by Theorem 27) and every $y \in M$ the Lusternik-Schnirelmann category of the space $\Omega(y)$ with respect to the $W^{1,p}$ topology is infinite.*

Proof. Let $1 < p < p_c$ be given by Theorem 27. Then $\Omega(y)$ and $\Omega(y)_{\mathrm{std}}$ are homotopy equivalent by the previous theorem (no matter the $W^{1,p}$ topology, as long as $p < p_c$). Since the cup length of the $W^{1,2}$-*ordinary* loop space of a compact simply connected manifold is infinite (see [**Sch64**, Corollary 20] or the classical work of Serre [**Ser51**]), so it is for $\Omega(y)$ with the $W^{1,p}$-topology. The cup-length is a lower bound for the Lusternik-Schnirelmann category, hence the result follows. □

A sufficiently small neighborhood of a nonsingular point in $\Omega(y)$ looks like a Hilbert space (hence it is contractible), but the structure near an abnormal curve can be fairly more complicated. This local structure is sharpened by the following result.

Corollary 30. *Every point $\gamma \in \Omega(y)$ (in particular an abnormal curve) has a neighborhood U such that the inclusion $U \hookrightarrow \Omega(y)$ is homotopic to a constant.*

Proof. Since $\Omega(y)$ has the homotopy type of a CW-complex by Theorem 28 above, then the result follows from [**FP90**, Proposition 5.1.2]. □

3. Critical Points of Geometric Costs

3.1. The Regularity of the Energy. For $p > 1$ we define the p-Energy $J_p : L^p(I, \mathbb{R}^d) \to \mathbb{R}$ by (for simplicity we omit to make explicit the dependence of J_p on p, when it will be clear from the context):

$$J_p(u) = \sum_{i=1}^{d} \|u_i\|_p^p, \quad u = (u_1, \ldots, u_d).$$

To simplify notations below we will simply denote $L^p = L^p(I, \mathbb{R}^d)$, also we will omit the subscript notation for $u = (u_1, \ldots, u_d)$ when not needed (the corresponding equations should thus be interpreted componentwise).

We will need the following result on Nemytskii operators.

Theorem 31 (Theorem 2.2 [**AP93**]). *Let $g : I \times \mathbb{R} \to \mathbb{R}$ be a function such that (i) the function $v \mapsto g(t, v)$ is continuous for almost every $t \in I$; (ii) the function $t \mapsto g(t, v)$ is measurable for all $v \in \mathbb{R}$. Assume also there exists $a, b > 0$ such that:*

$$|g(t, v)| \le a + b|v|^\alpha, \quad \alpha = \frac{p}{q}.$$

Then the map $u(\cdot) \mapsto g(\cdot, u(\cdot))$ (a Nemytskii operator*) is continuous from $L^p(I)$ to $L^q(I)$.*

As a corollary we derive the following elementary lemma.

LEMMA 32. *The map $u \mapsto u|u|^{p-2}$ is a continuous map from $L^p(I)$ to $L^{\frac{p}{p-1}}(I)$. In particular, if y is a regular value of the Endpoint map, then $f = J\big|_{\Omega(y)}$ is a C^1 function.*

PROOF. The continuity of $u \mapsto u|u|^{p-2}$ is immediate from the previous Theorem. Now, if y is a regular value of the Endpoint, the differential $d_u f$ coincides with $d_u J|_{T_u\Omega(y)}$ thus to prove that it is differentiable with continuous derivative it is enough to prove it for J. The differential $d_u J$ as a linear functional on $L^p(I, \mathbb{R}^d)$ is easily computed to be (componentwise):

$$\langle d_u J, h\rangle = \int_0^1 pu(t)|u(t)|^{p-2}h(t)dt, \quad \text{for all } h \in L^p,$$

i.e. $d_u J = pu|u|^{p-2} \in L^q = (L^p)^*$, then the result is clear from the previous claim. □

PROPOSITION 33 (Palais-Smale condition). *Let y be a regular value of the Endpoint map and $p > 1$. Then the function $f = J|_{\Omega(y)}$ satisfies the Palais-Smale condition, i.e. any sequence $\{\gamma_k\}_{k\in\mathbb{N}} \subset \Omega(y)$ on which f is bounded and such that $d_{\gamma_k} f \to 0$ has a convergent subsequence.*

PROOF. Consider the differential $d_u F$ of the Endpoint map at a point u. Using the notations of Theorem 45 we can write it, for any $v \in L^p$ as:

$$(d_u F)v = \int_0^1 M_u(1)M_u(s)^{-1}B_u(s)v(s)ds.$$

Denote by $w_1(t; u), \dots, w_n(t; u)$ the rows of the matrix $M_u(1)M_u(t)^{-1}B_u(t)$; notice that for $j = 1, \dots, d$ we have $w_j(\cdot; u) \in L^q$. If $u \in \Omega(y)$, then we can write:

$$T_u\Omega(y) = \ker d_u F = \operatorname{span}\{w_1(\cdot; u), \dots, w_n(\cdot; u)\}^\perp;$$

as the latter is a linear subspace, we also deduce:

$$T_u\Omega(y)^\perp = \operatorname{span}\{w_1(\cdot; u), \dots, w_n(\cdot; u)\}.$$

In particular, for any $u \in \Omega(y)$, $T_u\Omega(y)$ is a closed subspace of codimension n in L^p and therefore it is complemented, i.e. there exists a closed and finite dimensional subspace W_u such that

$$L^p = T_u\Omega(y) \oplus W_u; \tag{3.7}$$

finally, observe that there exists a continuous linear projection $\pi_u : L^p \to W_u$ subordinated to this splitting, that is $\ker(\pi_u) = T_u\Omega(y)$, see [**Car05**, Chapter 2].

Let now $\{u_k\}_{k\in\mathbb{N}} \subset \Omega(y)$ be a bounded sequence such that $d_{u_k} f \to 0$. Since $d_u f = (d_u J)|_{T_u\Omega(y)}$ then by definition of the projections π_{u_k} we have:

$$\langle d_{u_k} J, (\mathrm{Id} - \pi_{u_k})v\rangle \to 0, \quad \forall v \in L^p.$$

The space L^p is uniformly convex, hence reflexive by the Milman-Pettis theorem; the sequence $\{u_k\}$ is bounded by assumption and invoking Banach-Alaoglu we deduce the existence of a subsequence $\{u_{k_l}\}_{l\in\mathbb{N}}$ and $\overline{u} \in L^p$ such that $u_{k_l} \rightharpoonup \overline{u}$. Furthermore, observe that if $q = p^* = \frac{p}{p-1}$ is the conjugate exponent of p, then:

$$d_u J = pu|u|^{p-2} \Rightarrow \|d_u J\|_q^q = \|u\|_p^p.$$

By the above discussion, up to subsequences, we may thus assume that $\|u_k\|_p < C$ and $u_k \rightharpoonup \overline{u}$ in L^p. There exists then $K \in \mathbb{N}$ sufficiently large so that for any norm-one $v \in L^p$ and $k > K$ the following holds:

$$|\langle d_{u_k} J, \pi_{u_k}(v)\rangle| \leq |\langle d_{u_k} J, v\rangle| + |\langle d_{u_k} J, v - \pi_{u_k}(v)\rangle| < C + 1. \tag{3.8}$$

It is well-known [**Car05**, Section 3] that the splitting in (3.7) induces a dual splitting on L^q, namely for any $u \in \Omega(y)$ we have

$$L^q = (T_u\Omega(y))^* \oplus W_u^*;$$

moreover the adjoint operator $\pi_{u_k}^*$ is still a projection with kernel $W_{u_k}^\perp$ and range $(T_{u_k}\Omega(y))^\perp \cong W_{u_k}^* \cong L^q/W_{u_k}^\perp = \operatorname{span}\{w_1(\cdot;u_k),\dots,w_n(\cdot;u_k)\}$. In particular, (3.8) shows that

$$\|\pi_{u_k}^*(d_{u_k}J)\|_q < C+1, \quad \forall k > K.$$

Write:

$$\pi_{u_k}^*(d_{u_k}J) = \sum_{j=1}^n a_{j,k} w_j(\cdot;u_k);$$

since the projections have finite ranges, and all norms are equivalent on finite-dimensional spaces, by the above we deduce that there exists $C' > 0$ so that

$$\sum_{j,l} a_{j,k}a_{l,k}\langle w_j(\cdot;u_k), w_l(\cdot;u_k)\rangle = \|\pi_{u_k}^*(d_{u_k}J)\|_2^2 < C'. \tag{3.9}$$

Because of Lemma 47 and Theorem 42 and the fact that $u_k \to \overline{u}$ weakly in L^p, then for every $j = 1,\dots,n$ the function $w_j(\cdot;u_k) : [0,1] \to \mathbb{R}^d$ converges strongly (and hence in any L^p norm) to a function $\overline{w}_j : [0,1] \to \mathbb{R}^d$. Also, $F(\overline{u}) = y$ and since y is a regular value, then $\{\overline{w}_1,\dots,\overline{w}_n\}$ is a linearly independent set.

By (3.9) we have $\sum_{j,l} a_{j,k}a_{l,k}\langle \overline{w}_j, \overline{w}_l\rangle < C'$, which tells the sequence:

$$\left\{ z_k = \sum_j a_{j,k}\overline{w}_j \right\}_{k\in\mathbb{N}} \subset \operatorname{span}\{\overline{w}_1,\dots,\overline{w}_n\} \quad \text{is bounded.}$$

Since $\operatorname{span}\{\overline{w}_1,\dots,\overline{w}_n\}$ is finite dimensional we can then assume $z_k \to \overline{z}$; since $\{\overline{w}_1,\dots,\overline{w}_n\}$ is a linearly independent set then the sequences $\{a_{j,k}\}_{k\in\mathbb{N}}$ for $j = 1,\dots,n$ are bounded and we can assume they converge. Consequently also $\pi_{u_k}^*(d_{u_k}J) \to \overline{z}$ (all this up to subsequences).

Finally we have:

$$\begin{aligned}\lim_{k\to\infty} \|d_{u_k}J - \overline{z}\|_q \le& \lim_{k\to\infty}\left(\|d_{u_k}J - \pi_{u_k}^*(d_{u_k}J)\|_q\right)\\ &+ \lim_{k\to\infty}\left(\|\pi_{u_k}^*(d_{u_k}J) - \overline{z}\|_q\right) = 0.\end{aligned}$$

This proves that $u_k|u_k|^{p-2} = d_{u_k}J \xrightarrow{L^q} \overline{z}$ (up to subsequences), and the result follows now from the next Lemma 34. $\square$

LEMMA 34. *Let* $\{u_n\}_{n\in\mathbb{N}} \subset L^p$ *such that:*

$$u_n|u_n|^{p-2} \xrightarrow{L^q} z.$$

Then $u_n \xrightarrow{L^p} z|z|^{(2-p)/(p-1)}$.

PROOF. Consider the Nemytskii operator $N : L^q \to L^p$ defined by $v \mapsto v|v|^{(2-p)/(p-1)}$. Since:

$$\left| v|v|^{\frac{2-p}{p-1}} \right| \le |v|^{\frac{1}{p-1}} = |v|^{\frac{p}{p-1}\cdot\frac{1}{p}}$$

then $N \in C^0(L^q, L^p)$ by Theorem 31. In particular $u_n = N(u_n|u_n|^{p-2}) \xrightarrow{L^p} N(z)$, and the claim follows. $\square$

3.2. Critical Points.

THEOREM 35. *Let y be a regular value for the Endpoint map of the control system* (3.1), $1 < p < p_c$ *(where p_c is given by Theorem 27) and consider $f = J_p|_{\Omega(y)}$. Then f has infinitely many critical points.*

PROOF. The first part of the proof follows the lines of the classical argument. Assume first that the fundamental group of M is infinite. Then by (3.6) $\Omega(y)$ has infinitely many components. Lemma 32 tells that f is C^1 and Proposition 33 that it satisfies the Palais-Smale condition. Assume that one component of $\Omega(y)$ does not contain any critical point of f. Then we can apply the deformation lemma [**Cha93**, Lemma 3.2] and conclude that f needs to be unbounded from below, which is in contradiction with the definition $f = J_p|_{\Omega(y)} \geq 0$.

Assume now the fundamental group of M is finite. Let us call $r : \overline{M} \to M$ the universal covering map. Then $\overline{M}$ is also compact, and the structure $\mathcal{F}$ can be lifted to a structure $\overline{\mathcal{F}} = \{\overline{X}_0, \dots, \overline{X}_d\}$ by setting:

$$d_{\overline{x}} r \overline{X}_i(\overline{x}) = X_i(r(\overline{x})).$$

Let $\overline{x}$ be a lift of x and $\{\overline{y}_1, \dots, \overline{y}_k\}$ be the lifts of y (here $k = \#\pi_1(M)$, the number of sheets of the covering map). Denote by $\overline{\Omega}$ the set of horizontal curves on $\overline{M}$ leaving from $\overline{x}$, by $\overline{F}$ the corresponding Endpoint map and by $\overline{\Omega}(\overline{y})$ the set of horizontal curves on $\overline{M}$ between $\overline{x}$ and $\overline{y} \in \overline{M}$. We denote by $\overline{r} : \overline{\Omega} \to \Omega$ the smooth map that associates to a horizontal trajectory $\overline{\gamma}$ on $\overline{M}$ the trajectory $r \circ \overline{\gamma}$ on M. Notice that in coordinates this map is the identity maps on controls (hence it is a local diffeomorphism), and in particular:

$$J(\overline{\gamma}) = J(\overline{r}(\overline{\gamma})).$$

Moreover, by construction the following diagram is commutative:

$$\begin{array}{ccc} \overline{\Omega}(\overline{y}) & \xrightarrow{\ \overline{r}\ } & \Omega(y) \\ \Big\downarrow{\scriptstyle \overline{F}} & & \Big\downarrow{\scriptstyle F} \\ \overline{M} & \xrightarrow{\ r\ } & M \end{array}$$

and since r and $\overline{r}$ are local diffeomorphism, then $\overline{y}$ is a regular value of $\overline{F}$.

If we prove the statement for $\overline{M}$, then we are done: in fact given a critical point $\overline{u}$ for the geometric cost $\overline{f} = J|_{\overline{\Omega}(\overline{y})}$ then $\overline{r}(\overline{u})$ is a critical point for f (hence we would obtain an infinite numbers of distinct critical points for f). To see this fact let us use the Lagrange multiplier formulation: $\overline{u}$ is a critical point of $\overline{f}$ if and only if there exists $\overline{\lambda} \in T^*_{\overline{y}}\overline{M}$ such that:

$$\overline{\lambda} \circ d_{\overline{u}} \overline{F} = d_{\overline{u}} J.$$

Using the commutativity of the above diagram, and the fact that r is a local diffeomorphism we see that this implies the existence of a $\lambda \in T^*_y M$ such that

$$\lambda \circ d_{\overline{r}(\overline{u})} F \circ d_{\overline{u}} \overline{r} = d_{\overline{r}(\overline{u})} J \circ d_{\overline{u}} \overline{r} : \tag{3.10}$$

in fact

$$\begin{aligned} d_{\overline{r}(\overline{u})}J \circ d_{\overline{u}}\overline{r} &= d_{\overline{u}}J \\ &= \overline{\lambda} \circ d_{\overline{u}}\overline{F} \\ &= \overline{\lambda} \circ d_{r(\overline{F}(\overline{u}))}r^{-1} \circ d_{\overline{F}(\overline{u})}r \circ d_{\overline{u}}\overline{F} \\ &= \lambda \circ d_{\overline{r}(\overline{u})}F \circ d_{\overline{u}}\overline{r}. \end{aligned}$$

On the other hand, being $\overline{r}$ a local diffeomorphism, $d_{\overline{u}}\overline{r}$ is also an isomorphism of vector spaces; consequently simplifying it from (3.10) we can write:

$$\lambda \circ d_{\overline{r}(\overline{u})}F = d_{\overline{r}(\overline{u})}J$$

which tells exactly that $\overline{r}(\overline{u})$ is a critical point for f.

We are left with the case M compact and *simply connected*. Let y be a regular value of the Endpoint map and consider the horizontal path space $\Omega(y)$ endowed with the $W^{1,p}$ topology (recall that we are assuming $1 < p < p_c$ with p_c given by Theorem 27). Since y is a regular value of the Endpoint map, $\Omega(y)$ is a smooth Banach manifold modeled on $L^p = L^p([0,1], \mathbb{R}^d)$ (here d is the rank of the distribution). The function f is C^1 (by Lemma 32) and it satisfies the Palais-Smale condition (by Proposition 33 above), hence the results follows from Corollary 29 and the following Proposition.

PROPOSITION 36 (Corollary 3.4 from [**Cha93**]). *Let $\Omega(y)$ be Banach manifold and $f \in C^1(\Omega(y), \mathbb{R})$ bounded from below and satisfying the Palais-Smale condition. Then f has at least as many critical points as the Lusternik-Schnirelmann category of $\Omega(y)$.*

□

4. The sub-Riemannian Case

In this section we discuss applications of the previous results to the sub-Riemannian case, in particular we will always make the assumption $X_0 = 0$.

4.1. Geodesics. Given two points x, y in a sub-Riemannian manifold M, a *sub-Riemannian geodesic* is a curve $\gamma : I \to M$ satisfying the following properties: (i) it is absolutely continuous; (ii) its derivative (which exists almost everywhere) belongs to the sub-Riemannian distribution; (iii) it is parametrized by constant speed; (iv) $\gamma(0) = x$ and $y(1) = y$; (v) it is locally length minimizer, i.e. for every $t \in [0,1]$ there exists $\delta(t) > 0$ such that $\gamma|_{[t-\delta(t), t+\delta(t)]}$ has minimal length among all horizontal curves joining $\gamma(t - \delta(t))$ with $\gamma(t + \delta(t))$.

PROPOSITION 37. *Let y be a regular value of the Endpoint map centered at x. For every $p > 1$ all critical points of $f = J_p\big|_{\Omega(y)}$ are sub-Riemannian geodesics joining x to y.*

PROOF. First let us notice that curves that are *locally* J_p-minimizers are parametrized by constant speed and are locally length minimizer (the proof of this fact is the same as the classical proof for $p = 2$ as in [**Mil73**, Section 12] and essentially uses the fact that $\left(\int |u|\right)^p \leq \int |u|^p$ with equality if and only if $|u| \equiv c$). Also, being locally length minimizer and parametrized by constant speed implies that *globally* the parametrization is with constant speed.

Let us consider the equation for $u \in L^p$ to be a critical point of $f = J_p|_{F^{-1}(y)}$ (using Lagrange multipliers rule):

$$\exists \lambda \in T_y^*M \quad \text{such that} \quad \lambda \circ d_u F = pu|u|^{p-2}. \tag{3.11}$$

In particular since a critical point u of f is a *local* length minimizer (this can be seen by considering variations of only a small portion of the corresponding curve), we must have $|u| \equiv c > 0$ and we can rewrite (3.11) as:

$$\exists \eta = \frac{\lambda}{pc} \in T_y^* M \quad \text{such that} \quad \eta \circ d_u F = u,$$

which is the equation for the critical points of J_2 on $\Omega(y)$.

Thus if y is a regular value of the Endpoint map, the critical points of J_2 and J_p on $\Omega(y)$ are the same; since critical points of $J_2\big|_{\Omega(y)}$ are sub-Riemannian geodesics joining x to y (see [**ABB**, Section 4]), the result follows. □

As a corollary of Proposition 37 and Theorem 35, we obtain the sub-Riemannian version of Serre's theorem.

Theorem 38 (SubRiemannian Serre's Theorem). *If y is a regular value of the Endpoint map centered at a point x in a compact* sub-Riemannian *manifold, the set of sub-Riemannian geodesics joining x and y is infinite.*

4.2. The Contact Case. In the contact case we can remove from the sub-Riemannian Serre's theorem the *regularity* assumption on the two points. In fact the same proof works in the slightly more general case of *fat* distributions (see [**Mon02**] for more details on these distributions), as the only property that we are going to use is that there are no nontrivial abnormal curves.

Theorem 39. *For every two points on a compact, contact sub-Riemannian manifold the set of sub-Riemannian geodesics joining them is infinite.*

Proof. We prove that J_p (with $p > 1$) has infinitely many critical points when restricted to each $\Omega(y)$. Because of Theorem 35 the only case that we have to cover is the case the final point y is the same point as the initial point x (in which case it is not a regular value for F).

Recall that on a contact manifold there are no *nontrivial* abnormal extremals (i.e. critical points of the Endpoint map), see [**ABB, Mon02**], the trivial one being the one with zero control.

The case when the base manifold is not simply connected can be treated as in the proof of Theorem 35: if the fundamental group is infinite, then only one of the infinitely many components of $\Omega(x)$ contains the zero control; if the fundamental group is finite, we pass to the universal cover (which is still compact) and notice that the projection of a geodesic is still a geodesic (no matter if it is a singular point of the Endpoint map, as in the sub-Riemannian case geodesics are locally length minimizers and length is preserved by projection).

Thus we assume our manifold M is compact and simply connected. Consider $\tilde{F}$, the restriction to $L^p \backslash \{0\}$ of the Endpoint map centered at x. Then $\tilde{F}^{-1}(x)$ is a smooth Banach manifold and:

$$\Omega(x) = \tilde{F}^{-1}(x) \cup \{0\}$$

($\Omega(x)$ has its only singularity at zero).

We prove that the Lusternik-Schnirelmann category of $\tilde{F}^{-1}(x)$ is infinite. Combining this with the fact that the p-Energy $f : \tilde{F}^{-1}(x) \to \mathbb{R}$ is C^1 and satisfies Palais-Smale for every level $c > 0$ (by [**LM**, Theorem 19]), implies that f has infinitely many critical points.

Assume that the Lusternik-Schnirelmann category of $\tilde{F}^{-1}(x)$ is finite and let $U_1, \ldots, U_k$ be contractible open sets covering $\tilde{F}^{-1}(x)$. By Corollary 30 there exists an open neighborhood U_0 of the zero control (the singular point of $\Omega(x)$) such that the inclusion $U_0 \hookrightarrow \Omega(x)$ is homotopic to a constant map. As a consequence $\{U_0, \ldots, U_k\}$ would be an open cover of $\Omega(x)$ made of sets contractible in $\Omega(x)$, hence Lusternik-Schnirelmann category of $\Omega(x)$ would be finite, contradicting Corollary 29. $\square$

5. Appendix

In this section we collect a list of technical results that we use in the proofs. Most of these results are well known to experts, but it is often not easy to find an appropriate reference. Some proofs are adaptations from [**Tre00**] to the general case $p \in (1, \infty)$.

LEMMA 40 (Gronwall inequality). *Assume $\varphi : [0, T] \to \mathbb{R}$ to be a bounded nonnegative measurable function, $\alpha : [0, T] \to \mathbb{R}$ to be a nonnegative integrable function and $B : [0, T] \to \mathbb{R}$ to be non decreasing such that*

$$\varphi(t) \leq B(t) + \int_0^t \alpha(\tau)\varphi(\tau)d\tau, \quad \forall t \in [0, T];$$

then

$$\varphi(t) \leq B(t)e^{\int_0^t \alpha(\tau)d\tau}, \quad \forall t \in [0, T].$$

PROPOSITION 41. *Let $T > 0$ be fixed. Then the domain of the Endpoint map is open in $L^p([0, T], \mathbb{R}^d)$.*

PROOF. The strategy of the proof consists in showing that if v belongs to a sufficiently small neighborhood of u in $L^p([0, T], \mathbb{R}^d)$, then the corresponding trajectories γ_u and γ_v remain uniformly close. It is not restrictive to prove the theorem for small $T > 0$, which in turn allows us to work inside a coordinate chart. Also, we assume that the vector fields X_i, $i = 0, 1, \ldots, d$ have compact support in $\mathbb{R}^n$; Lemma 3.2 in the aforementioned paper yields that they are therefore globally Lipschitzian. For any $t \in [0, T]$ we have the following:

$$\begin{aligned}
\|\gamma_u(t) - \gamma_v(t)\| \leq \Bigg\| & \int_0^t (X_0(\gamma_u(\tau)) - X_0(\gamma_v(\tau)))d\tau \\
& + \int_0^t \sum_{i=1}^d v_i(\tau)(X_i(\gamma_u(\tau)) - X_i(\gamma_v(\tau)))d\tau \\
& - \int_0^t \sum_{i=1}^d (v_i(\tau) - u_i(\tau))X_i(\gamma_u(\tau))d\tau \Bigg\| \\
\leq C & \int_0^t (1 + \sum_{i=1}^d |u_i(\tau)|)\|\gamma_u(\tau) - \gamma_v(\tau)\|d\tau + h_v(t),
\end{aligned}$$

with

$$h_v(t) = \left\| \int_0^t \sum_{i=1}^d (v_i(\tau) - u_i(\tau))X_i(\gamma_u(\tau))d\tau \right\|.$$

By Hölder inequality we obtain

$$h_v(t) \leq C' T^{1/q}\|u - v\|_p, \quad \forall t \in [0, T];$$

moreover we deduce that for any $\varepsilon > 0$ there exists a neighborhood U of u in $L^p([0, T], \mathbb{R}^d)$ such that $h_v(t) \leq \varepsilon$, for any $v \in U$ and $t \in [0, T]$. We conclude using Gronwall inequality that

$$\|\gamma_u(t) - \gamma_u(v)\| \leq \varepsilon e^{C(T + T^{1/q}K)}, \quad \forall t \in [0, T].$$

$\square$

THEOREM 42. *Let* $u = (u_1, \dots, u_d) \in L^p([0,T], \mathbb{R}^d)$ *be a control in the domain of the Endpoint map* F*, and let* γ_u *be the corresponding solution to* (3.7)*. Let* $(u_n)_{n\in\mathbb{N}}$ *be a sequence in* $L^p([0,T], \mathbb{R}^d)$*. If* $u_n \overset{L^p}{\rightharpoonup} u$*, then for* n *large enough* γ_{u_n} *is well-defined on* $[0,T]$ *and moreover* γ_{u_n} *converges to* γ_u*, uniformly on* $[0,T]$.

PROOF. It suffices to prove the proposition when T is close to zero; this in turn permits to work in a coordinate chart, that is we may suppose the vector fields X_i to have compact support in $\mathbb{R}^n$. Moreover, let K be a compact neighborhood of x such that there exists $C > 0$ for which

$$\|X_i(z_1) - X_i(z_2)\| \le C\|z_1 - z_2\|$$

holds for any $z_1, z_2 \in K$ and any $i = 0, 1, \dots, d$. For all $t \in [0,T]$ we have:

$$\begin{aligned}\|\gamma_u(t) - \gamma_{u_n}(t)\| \le & \int_0^t \|(X_0(\gamma_u(\tau)) - X_0(\gamma_{u_n}(\tau))\| d\tau \\ & + \int_0^t \sum_{i=1}^d |u_{n,i}(\tau)| \|X_i(\gamma_u(\tau)) - X_i(\gamma_{u_n}(\tau))\| d\tau \\ & + \int_0^t \sum_{i=1}^d |u_{n,i}(\tau) - u_i(\tau)| \|X_i(\gamma_u(\tau))\| d\tau \\ & \le C \int_0^1 (1 + \sum_{i=1}^d |u_{n,i}(\tau)|) \|\gamma_u(\tau) - \gamma_{u_n}(\tau)\| d\tau + h_n(t),\end{aligned}$$

where

$$h_n(t) = \int_0^t \sum_{i=1}^d |u_{n,i}(\tau) - u_i(\tau)| \|X_i(\gamma_u(\tau))\| d\tau.$$

The uniform boundedness principle of Banach and Steinhaus ensures that $\sup_{n\in\mathbb{N}} \|u_n\|_p \le M$; if we can prove that h_n tends *uniformly* on $[0,T]$ to the zero function, then we would finish the argument using the Gronwall inequality.

Observe that h_n tends pointwise to the zero function; it is also uniformly $1/q$-Hölderian, where $q = \frac{p}{p-1}$, indeed if $L = \sup_i \sup_{p\in\mathbb{R}^n} \|X_i(p)\|$ we have

$$\begin{aligned}\|h_n(t_1) - h_n(t_2)\| &\le L \int_{t_1}^{t_2} \sum_{i=1}^d (|u_{n,i}(\tau)| + |u_i(\tau)|) d\tau \\ &\le L(M + \|u\|_p)|t_1 - t_2|^{1/q}.\end{aligned}$$

The proof is then concluded by the next lemma. □

LEMMA 43 (Uniform convergence of Hölderian maps). *Let* $\{f_k\}_{k\in\mathbb{N}} : [a,b] \to \mathbb{R}^n$ *be a uniformly* α*-Hölderian sequence of functions which converges pointwise to a limit function* f*. Then* f *is* α*-Hölderian and* $f_k \to f$ *uniformly on* $[a,b]$.

PROOF. The relation $\|f_k(x) - f_k(y)\| \le M|x-y|^\alpha$ immediately yields that the limit function f is also α-Hölderian.

Next, let $\varepsilon > 0$ be arbitrary and let accordingly $\rho = \left(\frac{\varepsilon}{3M}\right)^{1/\alpha}$. As $[a,b]$ is compact, it can be covered by a finite collection $\{B_i\}_{i=1}^l$ of balls of radius ρ, whose centers will be denoted by x_i; this means that for any $x \in [a,b]$ there exists $i \in \{1, \dots, l\}$ such that $|x - x_i| \le \rho$. Let $K \in \mathbb{N}$

be such that $\|f_k(x_i) - f(x_i)\| \le \varepsilon/3$ for all $i = 1, \ldots, l$ if $k > K$. The following holds true for $k \in \mathbb{N}$ sufficiently large:

$$\begin{aligned}\|f_k(x) - f(x)\| &\le \|f_k(x) - f_k(x_i)\| + \|f_k(x_i) - f(x_i)\| + \|f(x_i) - f(x)\| \\ &\le 2M|x - x_i|^\alpha + \frac{\varepsilon}{3} \le \varepsilon,\end{aligned}$$

and this finishes the proof. □

We turn now to the issue of the differentiability of the Endpoint map F, i.e. we want to determine its Fréchet differential and prove some of its continuity properties.

PROPOSITION 44. *Let u be in the domain of the Endpoint map $F : L^p([0,T], \mathbb{R}^d)$ and let γ_u be the associated trajectory. Then for any bounded neighborhood U of u in $L^p([0,T], \mathbb{R}^d)$, there exists a constant $C = C(U)$ such that whenever $v, w \in U$ and $t \in [0,T]$ we have*

$$\|\gamma_v(t) - \gamma_w(t)\| \le C\|v - w\|_p.$$

PROOF. Using (3.7) we derive the following estimate

$$\begin{aligned}\|\gamma_v(t) - \gamma_w(t)\| \le &\sum_{i=1}^{d} \int_0^t |v_i - w_i|\|X_i(\gamma_v(s))\|ds + \int_0^t \|X_0(\gamma_v(s)) - X_0(\gamma_w(s))\|ds \\ &+ \sum_{i=1}^{d} \int_0^t |w_i|\|X_i(\gamma_v(s)) - X_i(\gamma_w(s))\|ds.\end{aligned} \tag{3.12}$$

Theorem 42 ensures that γ_v and γ_w take values in a compact K which depends just on U[4]; as $X_0, X_1, \ldots, X_d$ are smooth, we have the existence of a constant M such that for all $v, w \in U$ and for all $i \in 1, \ldots, d$ there holds

$$\begin{aligned}\|X_i(\gamma_v)\| &\le M, \\ \|X_i(\gamma_v) - X_i(\gamma_w)\| &\le M\|\gamma_v - \gamma_w\|, \quad \forall t \in [0,T];\end{aligned}$$

lastly we may assume that U is contained in a ball of radius R, that is $\|w\|_p \le R$ for all $w \in U$. We proceed with the estimate in (3.12) as

$$\|\gamma_v(t) - \gamma_w(t)\| \le B\|v - w\|_p + M \int_0^t (1 + \sum_{i=1}^{d} |w_i|)\|\gamma_v(s) - \gamma_w(s)\|ds, \quad \forall t \in [0,T],$$

where $B = MT^{1/q}$; finally, Gronwall inequality yields

$$\|\gamma_v(t) - \gamma_w(t)\| \le Be^{M(T + RT^{1/q})}\|v - w\|_p, \quad \forall t \in [0,T].$$

□

We fix now some notations used in the next theorem: let $A_u(t) = dX_0(\gamma_u) + \sum_{i=1}^d u_i dX_i(\gamma_u)$, $B_u(t) = (X_1(\gamma_u), \ldots, X_d(\gamma_u))$, and let M_u be the $n \times n$ matrix solution of $M_u' = A_u M_u$ satisfying $M_u(0) = I$; we have

[4]By Banach-Alaoglu U is sequentially weakly compact, hence weakly compact by the Eberlein-Smulian theorem. On the other hand theorem 42 implies that for any $\varepsilon > 0$, whenever u, v belong to a sufficiently small open set, $\|\gamma_u(t) - \gamma_v(t)\| \le \varepsilon$ on $[0,T]$. The statement follows since whenever we cover U with a collection of open sets of arbitrary small size, we may always extract a finite subcover and then proceed via the triangular inequality.

THEOREM 45 (Differentiability of the Endpoint map). *The Endpoint map F is L^p-Fréchet differentiable; its differential at u is the linear map $dF(u) : L^p \to \mathbb{R}^n$ defined by*

$$(d_u F)v = \int_0^T M_u(T) M_u(s)^{-1} B_u(s) v(s) ds.$$

PROOF. Let $u \in L^p([0,T], \mathbb{R}^d)$ be fixed in the domain of F. Let us consider a neighborhood U of u in L^p; without loss of generality we may assume that there exists $R > 0$ such that $\|v\|_p \leq R$ for any $v \in U$. Let γ_u and γ_{u+v} be the solutions to (3.7) with respect to the controls u and $u+v$ respectively. We have

$$\dot{\gamma}_{u+v} - \dot{\gamma}_u = X_0(\gamma_{u+v}) - X_0(\gamma_u) + \sum_{i=1}^d v_i X_i(\gamma_{u+v}) + \sum_{i=1}^d u_i (X_i(\gamma_{u+v}) - X_i(\gamma_u)). \tag{3.13}$$

For all $i = 0, 1, \ldots, d$ there hold the expansions

$$\begin{aligned} X_i(\gamma_{u+v}) - X_i(\gamma_u) &= dX_i(\gamma_u)(\gamma_{u+v} - \gamma_u) \\ &\quad + \int_0^1 (1-t) d^2 X_i(t\gamma_u + (1-t)\gamma_{u+v})(\gamma_{u+v} - \gamma_u, \gamma_{u+v} - \gamma_u) dt, \\ X_i(\gamma_{u+v}) &= X_i(\gamma_u) + \int_0^1 (1-t) dX_i(t\gamma_u + (1-t)\gamma_{u+v})(\gamma_{u+v} - \gamma_u) dt; \end{aligned}$$

plug the above into (3.13) to rewrite that equation as

$$\dot{\omega} = A_u \omega + B_u v + \xi, \tag{3.14}$$

where $\omega(t) = \gamma_{u+v}(t) - \gamma_u(t)$ and

$$\begin{aligned} \xi(t) = &\sum_{i=1}^d v_i(t) \int_0^1 (1-s) dX_i(s\gamma_u + (1-s)\gamma_{u+v})(\gamma_{u+v} - \gamma_u) ds \\ &+ \int_0^1 (1-s) d^2 X_0(s\gamma_u + (1-s)\gamma_{u+v})(\gamma_{u+v} - \gamma_u, \gamma_{u+v} - \gamma_u) ds \\ &+ \sum_{i=1}^d u_i(t) \int_0^1 (1-s) d^2 X_i(s\gamma_u + (1-s)\gamma_{u+v})(\gamma_{u+v} - \gamma_u, \gamma_{u+v} - \gamma_u) ds. \end{aligned}$$

We have $\|v\|_p \leq R$ for all $v \in U$; the previous proposition and the estimate

$$\|s\gamma_u(s) + (1-s)\gamma_{u+v}(s)\| \leq \|\gamma_u(s)\| + (1-s)\|\gamma_{u+v}(s) - \gamma_u(s)\| \leq \|\gamma_u(s)\| + CR$$

imply that there exists a compact $K \subset \mathbb{R}^n$ such that $s\gamma_u(s) + (1-s)\gamma_{u+v}(s) \in K$ for any $s \in [0,1]$ and any $v \in U$. Since the X_i are smooth, again by the proposition above we have we the estimate

$$\|\xi(t)\| \leq c_1 \|v\|_p \sum_{i=1}^d |v_i(t)| + c_2 \|v\|_p^2 (1 + \sum_{i=1}^d |u_i(t)|).$$

We solve (3.14) to obtain

$$\omega(t) = \int_0^t M_u(t) M_u(s)^{-1} B_u(s) v(s) ds + \int_0^t M_u(t) M_u(s)^{-1} \xi(s) ds;$$

in particular for $t = T$

$$\left\|\gamma_{u+v}(T) - \gamma_u(T) - \int_0^T M_u(T)M_u(s)^{-1}(s)B_u(s)v(s)ds\right\| \tag{3.15}$$

$$\leq C'\left(c_1\|v\|_p \int_0^T \sum_{i=1}^d |v_i(s)|ds + c_2\|v\|_p^2 \int_0^T (1 + \sum_{i=1}^d |u_i(s)|)ds\right)$$

$$\leq C'\left(c_1 T^{1/q} + c_2(T + \|u\|_p T^{1/q})\right)\|v\|_p^2.$$

The map

$$\mathcal{F}_u : L^p \ni v \mapsto \int_0^T M_u(T)M_u(s)^{-1}B_u(s)v(s)ds \in \mathbb{R}^n$$

is evidently linear and by (3.15) also continuous. It then follows that the Endpoint map F is differentiable at u and $d_u F u = \mathcal{F}_u$. □

THEOREM 46. *Let $u = (u_1, \dots, u_d) \in L^p([0,T], \mathbb{R}^d)$ be a control in the domain of the Endpoint map F. Let $(u_n)_{n\in\mathbb{N}}$ be a sequence in $L^p([0,T], \mathbb{R}^d)$ such that $u_n \overset{L^p}{\rightharpoonup} u$ for some $u \in L^p([0,T], \mathbb{R}^d)$. Then $d_{u_n}F \to d_u F$.*

The proof of this theorem needs a series of preliminary lemmas; for $s \in [0,T]$, set $N_u(s) = M_u(T)M_u(s)^{-1}$. Since $N_u(s)M_u(s) = M_u(T)$, upon differentiation and using the definition of M_u, we obtain $N_u'(s)M_u(s) + N_u(s)A_u(s)M_u(s) = 0$, that is

$$N_u'(s) = -N_u(s)A_u(s), \quad N_u(T) = I.$$

LEMMA 47. *Let $\{u_n\}_{n\in\mathbb{N}}$ and u be as in the statement of theorem 46. Then $N_{u_n} \to N_u$ uniformly on $[0,T]$.*

PROOF.

$$\begin{aligned} N_u(t) - N_{u_n}(t) = \int_0^t \Bigg(& N_{u_n}(s)(dX_0(\gamma_{u_n}(s)) + \sum_{i=1}^d u_{n,i} dX_i(\gamma_{u_n}(s))) \\ & - N_u(s)(dX_0(\gamma_u(s)) + \sum_{i=1}^d u_i(s) dX_i(\gamma_u(s)))\Bigg)ds \end{aligned} \tag{3.16}$$

$$\begin{aligned} = \int_0^t \Bigg(& (N_{u_n}(s) - N_u(s))dX_0(\gamma_u(s)) + N_u(s)(dX_0(\gamma_{u_n}(s)) - dX_0(\gamma_u(s))) \\ & + (N_{u_n}(s) - N_u(s))\sum_{i=1}^d u_{n,i}(s) dX_i(\gamma_{u_n}(s)) \\ & + N_u(s)\sum_{i=1}^d u_{n,i}(s)(dX_i(\gamma_{u_n}(s)) - dX_i(\gamma_u(s))) \\ & + N_u(s)\sum_{i=1}^d (u_{n,i}(s) - u_i(s)) dX_i(\gamma_u(s))\Bigg)ds. \end{aligned}$$

By virtue of theorem 42, $\gamma_{u_n} \to \gamma_u$ uniformly on $[0,T]$; moreover if

$$h_n(t) = \int_0^1 N_u(s)\sum_{i=1}^d (u_{n,i}(s) - u_i(s)) dX_i(\gamma_u(s))ds,$$

then $\|h_n\| \to 0$ uniformly on $[0,T]$ by lemma 43: indeed the sequence $\{h_n\}_{n\in\mathbb{N}}$ is $1/q$- Hölderian and converges pointwise to 0, moreover the factor $N_u(s)$ does not depend on n. Then (3.16) can be estimated for n sufficiently large as

$$\|N_u(t) - N_{u_n}(t)\| \leq C \int_0^t \|N_u(s) - N_{u_n}(s)\| ds + \varepsilon,$$

and the theorem follows using the Gronwall inequality, as desired. □

Proof of theorem 46. Theorem 45 yields that the differential of the Endpoint map at the point w has the form

$$(d_w F)v = \int_0^T N_w(s) B_w(s) v(s) ds.$$

We know from theorem 42 that $\gamma_{u_n} \to \gamma_u$ uniformly on $[0,T]$; then $B_{u_n} \to B_u$ uniformly on $[0,T]$. As lemma 47 shows that also $N_{u_n} \to N_u$ uniformly on $[0,T]$, we deduce that

$$d_{u_n} F v \to d_u F v$$

uniformly on $[0,T]$, for any $v \in L^p([0,T], \mathbb{R}^d)$, and this finishes the proof. □

CHAPTER 4

Homotopically Invisibile Singular Curves

1. Introduction

1.1. Horizontal path spaces and singular curves. Let M be a smooth manifold of dimension m and $\Delta \subset TM$ be a smooth, totally nonholonomic distribution of rank d. Given a point $x \in M$ (which we will assume fixed once and for all) the horizontal path space Ω of *admissible* curves is defined[1] by:

$$\Omega = \{\gamma : [0,1] \to M \mid \gamma(0) = x,\ \gamma \text{ is absolutely continuous, } \dot{\gamma} \in \Delta \text{ a.e. and is } L^2\text{-integrable}\}.$$

The $W^{1,2}$ topology endows Ω with a Hilbert manifold structure, locally modeled on $L^2(I, \mathbb{R}^d)$.

The *endpoint map* is the smooth map assigning to each curve its final point:

$$F : \Omega \to M, \quad F(\gamma) = \gamma(1).$$

A *singular* curve is a critical point of F. Singular curves are at the core of nonholonomic geometry, but some natural questions about these curves remain open. In fact even the existence of a point $y \in M$ not joined by singular curves sorting from x is still an open problem (the "sub-Riemannian Sard's conjecture").

Given $y \in M$ we will denote by $\Omega(y)$ the set of horizontal curves joining x and y:

$$\Omega(y) = F^{-1}(y), \quad y \in M.$$

If y is a regular value of F, then there are no singular curves between x and y and the space $\Omega(y)$ is a smooth Hilbert manifold. As we just said, the absence of singular curves can not be granted in general and the space $\Omega(y)$ might be very singular. Despite this fact, from a homotopy theory point of view $\Omega(y)$ still turns out to be a "nice" space.

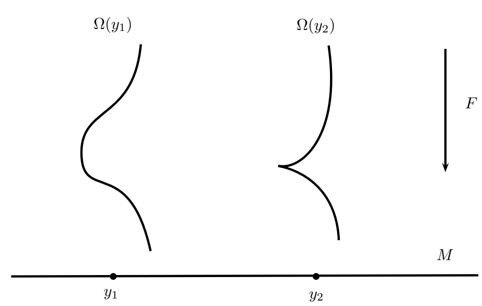


FIGURE 1. A simplified picture of two fibers of the enpdoint map $F : \Omega \to M$. The path space $\Omega(y_2)$ is singular, but it is homotopy equivalent to $\Omega(y_1)$.

---

[1] This definition requires the choice of an inner product in each fiber (a sub-Riemannian structure on Δ) in order to integrate the square of the norm of $\dot{\gamma}$, but the fact of being *integrable* is independent of the chosen structure (we refer the reader to [**ABB, Mon02**] for more details).

Theorem (Theorem 5, [**BL**]). *For every $y \in M$ the space $\Omega(y)$ has the homotopy type of a CW-complex and its inclusion in the standard path space (i.e. the space of $W^{1,2}$ curves on M with no nonholonomic constraint on their velocity) is a homotopy equivalence. In particular for a given manifold M and totally nonholonomic distribution $\Delta \subset TM$, all the spaces $\Omega(y)$ (regardless $y \in M$ being a regular value of F or not) have the same homotopy type.*

Thus, globally the homotopy type of $\Omega(y)$, $y \in M$, is not influenced by the fact of being singular. In fact all fibers of the endpoint map can be continuously deformed one into another, but if an additional function $J : \Omega \to \mathbb{R}$ (an energy functional) is considered, during the deformation we cannot preserve the Lebesgue sets of J.

1.2. Soft Singular Curves and Homotopically Invisible Curves. In this framework, the main interest of calculus of variations is to determine the existence of critical points[2] of a functional:

$$J : \Omega(y) \to \mathbb{R}.$$

Using the homotopy of the space $\Omega(y)$ (which in this case is known by the previous theorem) one is often able to force the existence of such critical points (typically via minimax procedures).

In our case the role of J will be played by a *sub-Riemannian Energy.* In other words, we fix an inner product on Δ smoothly depending on the base-point and define for $\gamma \in \Omega$:

$$J(\gamma) = \frac{1}{2}\int_I |\dot{\gamma}(t)|^2 dt.$$

If y is a regular value of F, then $\Omega(y)$ is smooth and a critical point of J is simply a curve γ for which there exists a nonzero $\lambda \in T^*M$ such that:

$$\lambda d_\gamma F = d_\gamma J. \tag{4.1}$$

In the language of sub-Riemannian geometry curves satisfying (4.1) are called *normal geodesics*; their short segments are length minimizers for the corresponding Carnot-Caratheodory distance on M (not all length minimizers are normal geodesics though).

In the spirit of Morse theory, when $\Omega(y)$ is smooth, normal geodesics with energy in $[E_1, E_2]$ are precisely the obstruction to deform the Lebesgue set $\{J \leq E_2\}$ to $\{J \leq E_1\}$ following the gradient flow of $-J$ (if y is a regular value of F, the function J satisfies the Palais-Smale condition [**BL**, Proposition 10] and the classical theory can be used; we refer the reader to [**BL**]).

If $\Omega(y)$ is singular, even if there are no normal geodesics with energy in $[E_1, E_2]$, the same deformation is in general not possible (in fact it is not even clear what a gradient flow should be on this singular space). Nevertheless, as we will see, for a generic (in the sense of [**CJT06**]) sub-Riemannian structure of rank $d \geq 3$ the absence of normal geodesics with energy in $[E_1, E_2]$ is enough to guarantee a *weak* deformation.

Denoting by $\Omega(y)^E$ the set $\{\gamma \in \Omega(y) \,|\, J(\gamma) \leq E\}$, the main result of this paper implies Theorem 48 below. The technical condition that we need in order to guarantee the conclusion of the results of this paper is that all singular curves with $J \leq E_2$ satisfy the following three properties: (a) they have corank one, (b) they are not Goh and (c) they are strictly abnormal.

When $d \geq 3$ sub-Riemannian structures whose all singular curves satisfy these conditions form an open dense set in the C^∞-Whitney topology by a result of Y. Chitour, F. Jean, and E. Trélat [**CJT06**] (see Theorem 51 below for a more precise statement).

[2]Note that, when $\Omega(y)$ is singular it is not clear yet what a "critical point" for J should be; this will be clarified below.

We call a singular curve satisfying conditions (a), (b) and (c) a *soft* singular curve; thus generic structures only have soft singular curves. The message of this paper is to show that soft singular curves are homotopically invisible.

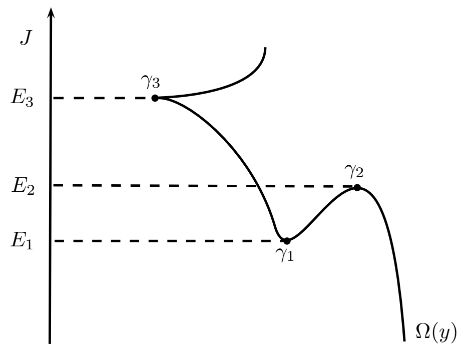


FIGURE 2. A simplified picture of a fiber $\Omega(y)$ and the Energy function (in the vertical direction). The curves γ_1 and γ_2 represent normal geodesics (they are obstructions to deform the Lebesgue sets). The curve γ_3 represents a soft singular curve: any cycle $X \subset \Omega(y)^{E_3}$ can be deformed a bit below the level E_3.

THEOREM 48 (sub-Riemannian Deformation Lemma). *Assume that all singular curves with energy $J \leq E_2$ are soft and that there are no normal geodesics in $\Omega(y)$ with energy in $[E_1, E_2]$. Then for every compact manifold X, every continuous map $f : X \to \Omega(y)^{E_2}$ and every $\epsilon > 0$ there exists a homotopy $f_t : X \to \Omega(y)^{E_2}$ such that:*

$$f_0 = f \quad \text{and} \quad f_1(X) \subset \Omega(y)^{E_1+\epsilon}.$$

In particular $\Omega(y)^{E_2}$ and $\Omega(y)^{E_1+\epsilon}$ are weakly homotopy equivalent.

The previous theorem says that (at least in the generic situation) singular curves with energy $E_1 \leq J \leq E_2$ aren't obstacles for the deformation of continuous maps (see Figure 2). They are "homotopically invisible".

1.3. The Calculus of Variations on the Horizontal Path Space. Going back to the calculus of variation, what we are interested in is the existence of *sub-Riemannian geodesics*: horizontal curves whose short segments are length minimizers for the Carnot-Caratheodory distance (this is in fact what we mean by a "critical point" for J).

As we already noted, curves satisfying (4.1) are geodesics – and these are all geodesics if $\Omega(y)$ is smooth, but in principle singular curves can also be geodesics.

On the other hand, condition (b) above (the non-Goh hypothesis) is a necessary optimality condition for singular curves: if a sub-Riemannian structure satisfies this condition, nonconstant singular curves cannot be geodesics (in other words: soft singular curves cannot be geodesics). Theorem 48 above can thus be considered as the starting point for the variational analysis on the space of horizontal curves.

One comment on the fact that deformation in Theorem 48 is a *weak* deformation and on the appearance of the "for every $\epsilon > 0$" in the statement. As a matter of fact a *strong* deformation retraction of Lebesgue sets is not needed even in the classical theory: all one needs is to be able to deform continuous maps (and more generally singular chains representing homology classes: we will prove that this is possible in Section 7). Neither one needs to deform up to the level E_1

included: the possibility of getting arbitrarily close (with the Energy) is still enough to use the results and predict the existence of critical points.

If two functions $f, g : X \to Y$ between topological spaces are homotopic, we will write $f \sim g$. The following statement is a sub-Riemannian version of the Minimax principle.

THEOREM 49 (sub-Riemannian Minimax principle). *Let X be a compact manifold and $f : X \to \Omega(y)$ be a continuous map which is not homotopic to a constant map. Consider:*

$$c = \inf_{g\sim f} \sup_{\theta\in X} J(g(\theta)).$$

Assume that there exists $\delta > 0$ such that all singular curves with energy $J \leq c+\delta$ are soft. Then for every $\epsilon > 0$ there exists a normal geodesic $\gamma_\epsilon \in \Omega(y)$ such that:

$$c - \epsilon \leq J(\gamma_\epsilon) \leq c + \epsilon.$$

It should be clear at this point that in principle, using Theorem 48, the powerful machinery of classical critical point theory can be adapted to the generic sub-Riemannian framework – to mention a specific example, we will show how to prove an analogue of Serre's theorem on the existence of infinitely many normal geodesics joining two points on a compact manifold (Corollary 72 below).

REMARK 22. The study of the space of maps with some restrictions on their differential goes back to the work on immersions of S. Smale [**Sma58**]; the case of trajectories of affine control systems were studied first by A. V. Sarychev [**Sar91**] for the uniform convergence topology, by J. Dominy and H. Rabitz [**DR12**] for the $W^{1,1}$ topology and by the last two authors of this paper for the $W^{1,p}, p > 1$ topology [**BL**]. The case of *closed* $W^{1,2}$ curves on nonholonomic distribution has been addressed by the last author and A. Mondino on [**LM**]. A sub-Riemannian version of Serre's theorem was proved by the last two authors of the current paper under the assumption that $y \in M$ is a regular value of F [**BL**]. The paper [**LM**] also contains various related results on variational problems on the closed horizontal loop space in the contact case.

1.4. Structure of the Paper. The paper is organized as follows. In Section 2 we recall the main definitions and properties of the objects we use. We note that Section 2.6 contains an interesting tool (the "global chart") which allows to switch from the space of curves to the space of controls (where the objects can be handled easier; we will switch back to the language of curves in the last section). Section 3 contains the main technical idea, which is the construction of a *cross-section* (a map "parametrizing" all possible values) for the pair (F, J) near a singular soft curve. This is used in Section 4 to prove a quantitative *non-smooth* implicit function theorem (Theorem 59) near a soft singular curve. In Section 5 we introduce the technical ingredients for handling the deformation on a singular space but away from singular curves. The sub-Riemannian Deformation Lemma (Theorem 68) is proved in Section 6.1. Indeed the invisibility of soft singular curves is a corollary of the Serre fibration property (Theorem 67). In Section 7 we discuss some applications and prove the sub-Riemannian Minimax principle (Theorem 69) and Serre's theorem on the existence of infinitely many normal geodesics on a compact sub-Riemannian manifold (Corollary 72).

2. Preliminaries

2.1. The Endpoint Map. Recall that we are considering a smooth totally nonholonomic distribution $\Delta \subset TM$ and that we have denoted by Ω the space of *horizontal* curves:

$$\Omega = \{\gamma : [0, 1] \to M \mid \gamma(0) = x,\ \gamma \text{ is absolutely continuous, } \dot\gamma \in \Delta \text{ a.e. and is } L^2\text{-integrable}\}.$$

This space endowed with the $W^{1,2}$ topology has a Hilbert manifold structure, locally modeled on $L^2(I,\mathbb{R}^d)$; we will call this topology the *strong* topology – the *weak* topology can also be considered on Ω (see [**BL, LM, Mon02**] for more details on these topologies).

The *endpoint map* is the map that gives the final point of each horiziontal curve:

$$F:\Omega\to M,\quad F(\gamma)=\gamma(1).$$

If $y\in M$, we indicate by:

$$\Omega(y)=F^{-1}(y)=\{\text{horizontal curves joining } x \text{ and } y\}.$$

We recall some properties of the Endpoint map, see [**BL**, Theorem 19 and Theorem 23].

PROPOSITION 50. *The endpoint map $F:\Omega\to M$ is smooth (with respect to the Hilbert manifold structure on Ω). Moreover if $\gamma_n\rightharpoonup\gamma$ weakly, then $F(\gamma_n)\to F(\gamma)$ uniformly (in particular F is continuous for the weak topology) and $d_{\gamma_n}F\to d_\gamma F$ in the operator norm.*

A horizontal (admissible) curve γ is said to be *singular* if it is a critical point of the endpoint map. The codimension of the singularity is called the *corank* of the singular curve.

2.2. Abnormal Extremals. Let us consider the cotangent bundle T^*M, and let us fix on it the standard symplectic structure ω. It is possible to find a distinguished subspace $\Delta^\perp$ within T^*M, called the annihilator of the distribution Δ, and accordingly we can restrict ω to a two form $\overline{\omega}$ on $\Delta^\perp$, which may now have characteristics.

An absolutely continuous curve $\lambda:[0,1]\to\Delta^\perp$ is an *abnormal* extremal of Δ if $\dot\lambda(t)\in\ker\overline{\omega}(\lambda(t))$ for any $t\in[0,1]$.

There exists a remarkable connection between singular curves (which are defined on the manifold M), and abnormal extremals (on T^*M): a horizontal curve γ is singular if and only if it is the projection of an abnormal extremal λ on the cotangent space. If this is the case, we say that λ is an *abnormal extremal lift* of γ.

2.3. The Energy and the Extended Endpoint Map. A sub-Riemannian structure on (M,Δ) is a Riemannian metric on Δ, i.e. a scalar product on Δ which smoothly depends on the base point. If Δ is endowed with a sub-Riemannian structure, we can define the *Energy* of horizontal paths:

$$J:\Omega\to\mathbb{R},\quad J(\gamma)=\int_0^1|\dot\gamma(t)|^2dt.$$

The Energy is a smooth (hence continuous) map on Ω, but it is only lower semicontinous with respect to the weak topology. Throughout this paper we will use the shorthand notation:

$$\Omega^E=\{\gamma\in\Omega\,|\,J(\gamma)\le E\}.$$

The *extended* endpoint map $\varphi:\Omega\to M\times\mathbb{R}$ is defined by (notice that it depends on the choice of the sub-Riemannian structure on Δ):

$$\gamma\mapsto(F(\gamma),J(\gamma)).$$

2.4. Normal Extremals. Once we have fixed a Riemannian metric g on Δ, we can define the sub-Riemannian Hamiltonian $H:T^*M\to\mathbb{R}$ as follows. We require that for every $x\in M$, the restriction of H to the fiber T_x^*M coincides with the nonnegative quadratic form

$$\lambda\mapsto\frac{1}{2}\max\left\{\frac{\langle\lambda,v\rangle^2}{g_q(v,v)}\,\middle|\,v\in\Delta_x\setminus\{0\}\right\}.$$

We call *normal extremal* any integral curve of the vector field[3] $\overrightarrow{H}$, that is any curve $\lambda : [0,1] \to T^*M$ such that $\dot{\lambda}(t) = \overrightarrow{H}(\lambda(t))$. The Pontryagin maximum principle [**AS04**] states that a necessary condition for a horizontal curve to be locally minimizing is to be either the projection of a normal or an abnormal extremal (the two possibilities are not mutually exclusive); accordingly, a singular curve γ which is not the projection of a normal extremal will be said *strictly abnormal*.

2.5. The Goh Condition. Let us consider a singular curve $\gamma : I \to M$, and let $\lambda : I \to T^*M$ an abnormal lift for γ. We say that γ is a *Goh* singular curve if $\lambda(t) \in \left(\Delta^2_{\gamma(t)}\right)^\perp$. A detailed discussion of the Goh condition can be found in [**ABB**, Chapter 12]; here we just briefly recall the salient facts. Since γ is a critical point of the endpoint map, there is a well-defined map (the *Hessian*):

$$\mathrm{Hess}_\gamma F : \ker d_\gamma F \to \operatorname{coker} d_\gamma F = T_{F(\gamma)}M/\mathrm{im}\, d_\gamma F.$$

If we precompose the Hessian with $\lambda = \lambda(1) \in T^*_{F(\gamma)}M$ what we get is a real-valued quadratic form defined on $\ker d_\gamma F$. A necessary condition for this map to have finite negative inertia index, i.e. $\mathrm{ind}^- \lambda \mathrm{Hess}_\gamma F < +\infty$, is that γ is a Goh singular curve. In this sense we may think of the Goh condition as a necessary second-order optimality condition for singular curves.

2.6. The Global "Chart" and the Minimal Control. We discuss in this section a useful construction introduced in [**LM**] in order to switch from curves to controls.

Assume M is a compact manifold. Given a sub-Riemannian structure on $\Delta \subset TM$, there exists a family of vector fields $X_1, \ldots, X_l$ with $l \geq d$ such that:

$$\Delta_x = \mathrm{span}\{X_1(x), \ldots, X_l(x)\}, \qquad \forall x \in M.$$

Moreover the previous family of vector fields can be chosen such that for all $x \in M$ and $u \in \Delta_x$ we have [**ABB**, Corollary 3.26]:

$$|u|^2 = \inf\left\{ u_1^2 + \cdots + u_l^2 \,\middle|\, u = \sum_{i=1}^{l} u_i X_i(x) \right\}, \tag{4.2}$$

where $|\cdot|$ denotes the modulus w.r.t. the fixed sub-Riemannian structure.

Denoting by $\mathcal{U} = L^2(I, \mathbb{R}^l)$, we define the map $A : \mathcal{U} \to \Omega$ by:

$$A(u) = \text{the curve solving the Cauchy problem } \dot{\gamma} = \sum_{i=1}^{l} u_i(t) X_i(\gamma(t)) \text{ and } \gamma(0) = x.$$

(We can use the compactness of M to guarantee that the solution to the Cauchy problem is defined for all $t \in [0,1]$, otherwise we need to define $\mathcal{U}$ as the open set of controls for which the solution $A(u)$ is defined up to time $t = 1$).

We will consider this construction fixed once and for all, and call it the "global chart". Abusing of notation, the endpoint map for this global chart will still be denoted by $F : \mathcal{U} \to M$:

$$F : \mathcal{U} \to M, \quad F(u) = F(A(u)).$$

The map A is continuous (both for the strong and the weak topologies on $L^2(I, \mathbb{R}^l)$) and has a right inverse $\mu : \Omega \to \mathcal{U}$ defined by:

$$\mu(\gamma) = u^*(\gamma)$$

[3] We recall that $\overrightarrow{H}$ is defined by the equation $\iota_{\overrightarrow{H}}\omega = -dH$

where $u^*(\gamma)$ is the control realizing the minimum of $\|\cdot\|^2$ on $A^{-1}(\gamma)$ (notice that this in particular implies $J(A(u)) \leq J(u)$). This control is called the *minimal control* [**ABB**, Remark 3.9]. The minimal control exists and is unique by [**LM**, Lemma 2]; it depends continuously on the curve γ (for the strong topologies) by [**LM**, Proposition 4]. Moreover [**LM**, Lemma 3] guarantees that:

$$J(\gamma) = \frac{1}{2}\|u^*(\gamma)\|^2.$$

In the sequel, given a point $y \in M$ we will fix a set of coordinates on a neighborhood $U \simeq \mathbb{R}^m$ of the point y, denote by $\mathcal{U} = F^{-1}(U)$ and simply write:

$$\varphi : \mathcal{U} \to U \times \mathbb{R} \subset \mathbb{R}^{m+1}$$

for the extended endpoint map *already in coordinates.*

A singular control u is a critical point of $F : \mathcal{U} \to M$; its corank is the corank of d_uF (notice that the corank of u can also be defined as the corank of $F|_{\{J=J(u)\}}$).

2.7. How to Build a sub-Riemannian Manifold. Following the notation of [**ABB**], we will assume that the distribution $\Delta \subset TM$ is defined as the image of a bundle map with constant rank:

$$f : M \times \mathbb{R}^l \to TM, \quad \Delta_x = f(U_x)$$

In this way Δ can be endowed with a sub-Riemannian metric by simply taking (4.2) as a definition. By virtue of [**ABB**, Corollary 3.26], this construction is indeed equivalent to the standard one. Denoting by $\{e_1, \ldots, e_l\}$ the standard basis of $\mathbb{R}^l$, this approach also has the advantage that the vector fields generating the distribution are naturally defined as:

$$X_i(x) = f(x, e_i), \quad x \in M.$$

2.8. Generic Properties and Soft Curves. In addition to the totally nonholonomic condition on Δ (also called the Hörmander condition [**ABB**]), in this paper we will consider sub-Riemannian structures whose singular curves satisfy the following properties.

DEFINITION 10 (Soft singular curves). *We will say that a singular curve is* soft *if:* (a) *it has corank one,* (b) *it is not Goh and* (c) *it is strictly abnormal. We will use the same terminology for singular controls.*

For *generic* sub-Riemannian structures all singular curves are soft, as it is clarified by the following result from [**CJT06**]. We denote by $\mathcal{D}_d$ the set of rank d distributions on M endowed with the Whitney C^∞ topology and by $\mathcal{G}_d$ the set of couples (Δ, g) where Δ is a distribution on M and g is a Riemannian metric on Δ, endowed with the Whitney C^∞ topology. We will say that a distribution $\Delta \subset TM$ satisfy a property from (a), (b), (c) if all its singular curves satisfy this property.

THEOREM 51 (Chitour, Jean, Trélat). *If $d \geq 2$ there exists an open dense set $O_{a,d} \subset \mathcal{D}_m$ where condition* (a) *is satisfied* [**CJT06**, Theorem 2.4]*; if $d \geq 3$ there exists an open dense set $O_{b,d} \subset \mathcal{D}_d$ where also condition* (b) *is satisfied* [**CJT06**, Corollary 2.5]*. Moreover, if $d \geq 2$ there exists an open dense set $O_{c,d} \subset \mathcal{G}_m$ where condition* (c) *is satisfied* [**CJT06**, Proposition 2.7]*. In particular for a generic sub-Riemannian structure of rank $d \geq 3$ all singular curves are soft.*

Using the approach of viewing the sub-Riemannian structure as the image of a bundle map, the above conditions are still generically satisfied. Specifically, when working in the global chart, we consider the control system $\{X_1, \ldots, X_l\}$ whose *trajectories* are solutions to:

$$\dot{x} = \sum_{i=1}^{l} u_i X_i, \quad x(0) = x.$$

The fact that generically in the Whitney topology for l-tuples of vector fields on M all singular controls are soft is granted by [**CJT08**, Theorem 2.6, Corollary 2.7]. Moreover it is not difficult to verify that if a control is not soft then the same is true for the associated trajectory. In fact let $G : \mathcal{U} \to M$ be defined as $G = F \circ A$, and assume that u is a critical point for G. Then the following chain of implications holds

$$0 = \lambda d_u G \Leftrightarrow \lambda d_{A(u)} F \circ d_u A = 0 \Leftrightarrow \lambda d_{A(u)} F = 0,$$

where we have used in the last implication that A is a submersion. Then any abnormal lift of u is an abnormal lift for the (singular) curve γ, hence it must be unique, and it must annihilate $\Delta^2_{\gamma(t)}$ for every $t \in [0,1]$ whence (a) and (b) follows also for controls. To prove (c) observe that if the singular control u admits a normal extremal lift, then we must have $\lambda d_u G = d_u J = u$, and in particular the curve $A(u)$ has to be a local minimizer of the length, parametrized with constant velocity. The Pontryagin maximum principle says that in this case $A(u)$ is either the projection of a normal or an abnormal extremal, but the first possibility is excluded by point (c) for curves. Hence u cannot admit a normal extremal lift (it would also be a normal extremal lift of $A(u)$), and it is therefore strictly abnormal. (As a corollary, if all singular curves are soft, then the same is true for all singular controls.)

The set Ω can be decomposed as follows:

$$\Omega = \mathcal{R} \cup \mathcal{C} \cup \mathcal{A},$$

where $\mathcal{R}$ is the set of regular points for φ, $\mathcal{C}$ consists of *strictly normal* curves, i.e. regular points γ of F for which there exists (λ_0, λ) with $\lambda_0 \neq 0$ such that $\lambda d_\gamma F = \lambda_0 d_\gamma J$, and $\mathcal{A}$ are the *abnormal* curves, for which there exists $(0, \lambda)$ such that $\lambda d_\gamma F = 0$. If properties (a) and (c) are verified, these three sets are indeed disjoint.

All the "technical" results in the sequel will be proved for the global chart, but we will go back to the general setting of horizontal curves for the main theorems. Abusing of notations, we will still denote by $\mathcal{C}$ the set of strictly normal *controls* and by $\mathcal{A}$ the set of abnormal *controls*.

3. Soft Abnormal Controls: The Cross Section

Let u_0 be a *soft* abnormal control with Energy $J(u_0) \leq E$. Then the corank of φ at a u_0 is one and there exist $e_1(u_0), \ldots, e_m(u_0)$ such that:

$$\mathrm{im} d_{u_0}\varphi = \mathrm{span}\{d_{u_0}\varphi\, e_1(u_0), \ldots, d_{u_0}\varphi\, e_m(u_0)\}.$$

Consider now the finite-codimensional set:

$$\mathcal{P} = \left\{ v \in L^2(I, \mathbb{R}^l) \,\middle|\, \|v\|_{L^2} \leq 1 \text{ and } \int_0^1 v(t)dt = 0 \right\}.$$

For $v \in \mathcal{P}$, $\bar{t} \in [0,1]$ and $s \in \mathbb{R}$ small enough, we set:

$$v_s(t) = \begin{cases} \frac{1}{|s|^{1/4}} v\left(\frac{t-\bar{t}}{|s|^{3/4}}\right) & \bar{t} \leq t \leq \bar{t} + |s|^{3/4} \\ 0 & \text{otherwise} \end{cases} \tag{4.3}$$

An easy computation shows that for any $v \in \mathcal{P}$ we have:

$$\|v_s\|_{L^2} = |s|^{1/8} \|v\|_{L^2} \xrightarrow{s \to 0} 0$$

For any $u \in L^2(I, \mathbb{R}^l)$ we define the non autonomous horizontal vector field $f_u = \sum_{i=1}^l u^i X_i$; if $0 \le t_1 \le t_2 \le 1$, its flow is given by the diffeomorphism (we highlight from here on the explicit dependence on u):

$$P_{t_1}^{t_2,u} = \overrightarrow{\exp} \int_{t_1}^{t_2} f_{u(t)} dt.$$

Since u_0 is not a Goh singular control, there exist[4] $\bar{t} \in [0, 1/2]$ and $a, b \in \mathbb{R}^m$ such that

$$\langle (P_{\bar{t}}^{1,u_0})^* \lambda, [f_a, f_b] \rangle \neq 0. \tag{4.4}$$

This in turn yields [**ABB**, Lemma 11.21] that the map $Q^{u_0} : \mathcal{P} \to \mathbb{R}$, defined by

$$Q^{u_0}(v(\cdot)) = \int_0^1 \langle (P_{\bar{t}}^{1,u_0})^* \lambda, [f_{w(\theta)}, f_{v(\theta)}] \rangle d\theta, \tag{4.5}$$

with $w(\theta) = \int_0^\theta v(\zeta) d\zeta$, has infinite positive and negative index. It is then possible to choose v^+ and v^- in $\mathcal{P}$ so that $\operatorname{sign}(Q^{u_0}(v^\pm)) = \pm 1$.

With this notation, we define:

$$\alpha_{u_0}(x, y) = v_{|x|}^{\operatorname{sgn}(x)} + y_1 e_1(u_0) + \cdots + y_m e_m(u_0),$$

where in the definition of $v_{|x|}^{\operatorname{sgn}(x)}$ (compare with (4.3)) we choose $\bar{t}$ as a time for which (4.4) holds.

The goal of this section is to prove the following key proposition.

Proposition 52. *For every soft $u_0 \in \mathcal{A}^E$ consider the function:*

$$G_{u_0}(u, x, y) = \varphi(u + \alpha_{u_0}(x, y)) - \varphi(u)$$

where α_{u_0} is defined as above. There exist weak neighborhoods $\mathcal{V} \subset \mathcal{W}$ of u_0, positive constants $r_1, r_2, r_3 > 0$ and a function:

$$g : B(0, r_1) \times \mathcal{W} \to B(0, r_3)$$

which is continuous for the strong topology and such that for every $(w, u) \in B(0, r_1) \times \mathcal{W}$

$$G_{u_0}(u, g(w, u)) = w.$$

Moreover for every $u \in \mathcal{V} \cap \{J \le E\}$ we have $B(u, r_2) \subset \mathcal{W}$.

We postpone the proof to Section 3.3, because it will require some preliminary results.

3.1. A Change of Coordinates. Let us start by writing:

$$G_{u_0}(u, x, y) = (\varphi_0(u, x, y), \varphi_1(u, x, y), \ldots, \varphi_m(u, x, y)) - \varphi(u)$$

and for $i = 1, \ldots, m$ let us denote by $\psi_i^u : \mathbb{R}^{m+1} \to \mathbb{R}$ the function:

$$\psi_i^u(x, y) = \varphi_i(u, x, y) - \varphi_i(u, 0, 0).$$

Notice that the map G_{u_0} is weak continuous: the first components are continuous because the Endpoint map itself is weak continuous; for the last component we have:

$$\begin{aligned}
\psi_m^u(x, y) &= \frac{1}{2} \left(\|u + \alpha_{u_0}(x, y)\|^2 - \|u\|^2 \right) \\
&= \frac{1}{2} \left(2\langle u, \alpha_{u_0}(x, y) \rangle + \|\alpha_{u_0}(x, y)\|^2 \right)
\end{aligned}$$

[4]Even small pieces of singular curves cannot be Goh curves.

which is weak continuous as a function of (u,x,y). We will consider local coordinates on a neighborhood $W \simeq \mathbb{R}^{m+1}$ of zero induced by the splitting:

$$\mathbb{R}^{m+1} \simeq \mathrm{coker} d_{u_0}\varphi \oplus \mathrm{im} d_{u_0}\varphi. \tag{4.6}$$

LEMMA 53. *For every soft* $u_0 \in \mathcal{A}^E$ *there exist a weak neighborhood* $\mathcal{W}_1$ *of* u_0*, a ball* $B_1 \subset \mathbb{R}^m$ *centered at zero and an interval* $I_1 \subset \mathbb{R}$ *centered at zero such that for every* $u \in \mathcal{W}_1$ *the map:*

$$\phi^u = (x, \psi_1^u, \dots, \psi_m^u) : I_1 \times B_1 \to \mathbb{R}^{m+1}$$
$$(x,y) \mapsto (x, \psi_1^u(x,y), \dots, \psi_m^u(x,y))$$

is a coordinate chart (i.e. a homeomorphism onto its image).

Moreover there exist positive numbers $r_1', r_1'' > 0$ *and a weak neighborhood* $\mathcal{V}_1 \subset \mathcal{W}_1$ *such that* $\phi^u(I_1 \times B_1) \supset I_1 \times B(0, r_1')$ *and* $B(u, r_1'') \subset \mathcal{W}_1$ *for every* $u \in \mathcal{V}_1 \cap \{J \leq E\}$.

REMARK 23. Essentially working in the coordinate chart induced by ϕ^u is equivalent to changing the differentiable structure on the target space $\mathbb{R}^{m+1}$. The last condition (the existence of $r_1' > 0$) tells that the size of the neighborhood where we use these new coordinates is uniform on a weak open set in the u-parameter.

PROOF. Denoting by $(G_0, G_1, \dots, G_m)$ the components of G_{u_0}, notice that the partial derivatives $\partial_{y_j} G_i|_{(u,x,y)}$ exist for $i, j = 1, \dots, m$ and are continuous functions for the *weak* topology on the u-variable. In fact there exists a matrix A (because we have performed the change of coordinates (4.6) in the target space) such that:

$$\left(\frac{\partial G_i}{\partial y_j}\bigg|_{(u,x,y)} \right)_{i,j=1,\dots,m} = A \cdot \begin{pmatrix} \frac{1}{2}\langle u + \alpha_{u_0}(x,y), e_1(u_0)\rangle & \cdots & \frac{1}{2}\langle u + \alpha_{u_0}(x,y), e_m(u_0)\rangle \\ d_{u+\alpha_{u_0}(x,y)}F\, e_1(u_0) & \cdots & d_{u+\alpha_{u_0}(x,y)}F\, e_m(u_0) \\ \vdots & \vdots & \vdots \\ d_{u+\alpha_{u_0}(x,y)}F\, e_1(u_0) & \cdots & d_{u+\alpha_{u_0}(x,y)}F\, e_m(u_0) \end{pmatrix}. \tag{4.7}$$

The elements of the first row in the above matrix are fixed linear functional evaluated on $u + \alpha_{u_0}(x,y)$ and are continuous in (u,x,y) even when considering the weak topology for the u-variable; for all the other elements, the weak-continuity follows from the fact that the map $u \mapsto d_u F$ is weak-strong continuous. We denote by:

$$D(u,x,y) = \left(\frac{\partial G_i}{\partial y_j}\bigg|_{(u,x,y)} \right)_{i,j=1,\dots,m}.$$

By assumption the matrix $D(u_0,0,0)$ is of rank m and there exists a ball W centered at zero in $\mathbb{R}^m$ such that $\psi^{u_0}(0,\cdot) : W \to \mathbb{R}^m$ is a diffeomorphism onto its image.

Consider the function $H : L^2 \times \mathbb{R} \times \overline{W} \times S^{m-1} \to \mathbb{R}$ defined by:

$$H(u,x,y,v) = \max_{i=1,\dots,m} \left\{ |\langle \nabla_y \varphi_i|_{(u,x,y)}, v\rangle| \right\}$$

where $\nabla_y \varphi_i|_{(u,x,y)}$ denotes the i-th row of $D(u,x,y)$. Note that this function is continuous with respect to the weak topology in the u-variable.

We claim that there exists $a > 0$ such that $H(u_0,0,0,v) > a$ for every $v \in S^{m-1}$. Assume that this is false. Then, for every $n \in \mathbb{N}$ there exists $v_n \in S^{m-1}$ such that $|\langle \nabla_y \varphi_i|_{(u_0,0,0)}, v_n\rangle| \leq 1/n$ for every $i = 1, \dots, m$. By compactness of S^{m-1} this gives the existence of a nonzero vector $\lim_k v_{n_k} = \overline{v} \in S^{m-1}$ such that $|\langle \nabla \varphi_i|_{(u_0,0,0)}, \overline{v}\rangle| = 0$ for every $i = 1, \dots, m$, which means that $\overline{v}$ is orthogonal to all the rows of $D(u_0,0,0)$ which is impossible since this matrix is invertible.

We claim that there exists a weak-open neighborhood $U^{\text{weak}} = \mathcal{U}_1^{\text{weak}} \times U_1 \times W_1$ of $(u_0, 0, 0)$ (here we take W_1 to be a ball for simplicity) such that for every $v \in S^{m-1}$ there exists $i = i(v) \in \{1, \ldots, m\}$ such that $|\langle \nabla_y \varphi_i|_{(u,x,y)}, v\rangle| > a$ for all $(u, x, y) \in U^{\text{weak}}$. In fact, by weak-continuity of H, for every $v \in S^{m-1}$ there exists a weak-open neighborhood $W(v)^{\text{weak}} = \mathcal{U}_1(v)^{\text{weak}} \times U_1(v) \times W_1(v) \times U_{S^{m-1}}(v)$ with the property that there exists $i \in \{1, \ldots, m\}$ such that for every $(u, x, y, w) \in W(v)^{\text{weak}}$ we have $|\langle \nabla_y \varphi_i|_{(u,x,y)}, w\rangle| > a$. By compactness of S^{m-1} there exist $v_1, \ldots, v_N$ such that $\{U_{S^{m-1}}(v_j)\}_{j=1,\ldots,N}$ is an open cover of S^{m-1}. The open set U^{weak} is defined as:

$$U^{\text{weak}} = \bigcap_{j=1}^{N} \left(\mathcal{U}_1(v_j)^{\text{weak}} \times U_1(v_j) \times W_1(v_j)\right).$$

We use this to prove that for every $(u, x) \in \mathcal{U}_1^{\text{weak}} \times U_1$ the following map is injective:

$$\psi^u(x, \cdot) = (\psi_1(u, x, \cdot), \ldots, \psi_m(u, x, \cdot)) : W_1 \to \mathbb{R}^m.$$

To this end consider:

$$\begin{aligned} \|\psi^u(x, y_1) - \psi^u(x, y_2)\|_1 &= \sum_{i=1}^{m} |\varphi_i(u, x, y_1) - \varphi_i(u, x, y_2)| \\ &= \sum_{i=1}^{m} \left| \int_0^1 \partial_t f_i(u, x, t) dt \right| = (*) \end{aligned}$$

where $f_i(u, x, t) = \varphi_i(u, x, y_2 + t(y_1 - y_2))$. Consequently:

$$\partial_t f_i(u, x, t) = \langle \nabla_y \varphi_i|_{(u,x,y_2+t(y_1-y_2))}, y_1 - y_2\rangle = \|y_1 - y_2\| \langle \nabla_y \varphi_i|_{(u,x,y(t))}, v\rangle$$

where we have set $y(t) = y_2 + t(y_1 - y_2) \in W_1$ and $v = \frac{y_1 - y_2}{\|y_1 - y_2\|} \in S^{m-1}$. Thus we can write:

$$\begin{aligned} (*) &= \sum_{i=1}^{m} \left| \int_0^1 \|y_1 - y_2\| \langle \nabla_y \varphi_i|_{(u,x,y(t))}, v\rangle dt \right| \\ &\geq \left| \int_0^1 \|y_1 - y_2\| \langle \nabla_y \varphi_{i(v)}|_{(u,x,y(t))}, v\rangle dt \right| \\ &\geq \|y_1 - y_2\| a \end{aligned}$$

which proves the injectivity of $\psi^u(x, \cdot)$.

As a consequence for every $u \in \mathcal{U}_1^{\text{weak}}$ the map:

$$\phi^u = (x, \psi^u) : U_1 \times W_1 \to \mathbb{R}^{m+1}$$

is continuous and injective; by the Invariance of Domain Theorem this map is a homeomorphism onto its image.

We prove now that (up to restricting the open set U^{weak}) the image of (x, ψ^u) contains a ball of uniform radius.

First, notice that the above chain of inequalities implies that for every $u \in \mathcal{U}_1^{\text{weak}}$ and $x \in U_1$ we have:

$$\|\psi^u(x, y_1) - \psi^u(x, y_2)\| \geq ca\|y_1 - y_2\|$$

where a constant $c > 0$ appears, but it only depends on m (because all norms on a finite dimensional space are equivalent).

We are now in the position of using [**?**, Lemma 5], which guarantees that if $rB \subset W_1$ (here rB is a shorthand notation for $B(0, r)$), then:

$$\psi^u(x, \cdot)(rB) \supset \psi^u(x, 0) + rcaB.$$

The map G_{u_0} is weak continuous and $\psi^u(0,0)=0$. Hence for every $r>0$ there exists a weak open set $\mathcal{O}_1^{\text{weak}}(r)\times O_1(r)$ such that for every $(u,x)\in\mathcal{O}_1^{\text{weak}}(r)\times O_1(r)$ we have $\|\psi^u(x,0)\|<\frac{1}{3}acr$. In particular for every $u\in\mathcal{U}_1^{\text{weak}}\cap\mathcal{O}_1^{\text{weak}}(r)$ and for every $x\in U_1\cap O_1(r)$, if $rB\subset W_1$ then:

$$\psi^u(x,\cdot)(rB)\supset\psi^u(x,0)+racB\supset\frac{2}{3}racB.$$

Let now $\tilde r>0$ be such that[5] $2\tilde rB\subset W_1$ and denote by $\mathcal{U}_3^{\text{weak}}=\mathcal{U}_1^{\text{weak}}\cap\mathcal{O}_1^{\text{weak}}(\tilde r)$ and by $U_3=U_1\cap O_1(\tilde r)$. Then we have just proved that, for every $u\in\mathcal{U}_3^{\text{weak}}$, the map:

$$\phi^u:U_3\times\tilde rB\to\mathbb{R}^{m+1}$$

is a homeomorphism onto its image, and this image contains $U_3\times\frac{2}{3}\tilde racB$.

Consider now the two functions:

$$\alpha_1(u)=\min_{(x,y,v)\in\overline{U}_3\times\tilde r\overline{B}\times S^{m-1}}H(u,x,y,v)\quad\text{and}\quad\alpha_2(u)=\max_{x\in\overline{U}_3}\|\psi^u(x,0)\|.$$

These functions are well defined (the max and the min are taken over compacts) and are continuous for the weak topology[6]. Moreover:

$$\mathcal{U}_1^{\text{weak}}\supset\{\alpha_1>a\}\quad\text{and}\quad\mathcal{O}_1^{\text{weak}}(\tilde r)\supset\{\alpha_2<ac\tilde r/3\}.$$

Define finally, for $\epsilon>0$ small enough, the open sets:

$$\mathcal{W}_1=\{\alpha_1>a,\alpha_2<ac\tilde r/3\},\quad\mathcal{V}_1=\{\alpha_1>a+\epsilon,\alpha_2<ac\tilde r/3-\epsilon\},\quad I_1=U_3\quad B_1=\tilde rB.$$

The two open sets $\mathcal{W}_1,\mathcal{V}_1$ are weakly open because α_1,α_2 are weak continuous; $\epsilon>0$ is taken small enough in order to guarantee that $u_0\in\mathcal{W}_1,\mathcal{V}_1$ (such an ϵ exists because $\alpha_1(u_0)>a$ and $\alpha_2(u_0)<ac\tilde r/3$).

Then $r_1'=\frac{2}{3}\tilde rac$ satisfies the requirements from the statement. For the existence of $r_1''>0$ we argue as follows. We consider the weak closed set $C=(\mathcal{W}_1)^c$ and the weak compact:

$$K=\{\alpha_1\geq a+\epsilon,\alpha_2\leq ac\tilde r/3-\epsilon\}\cap\{J\leq E\}.$$

These two sets are disjoint and by Lemma 54 below there exists $r_1''>0$ such that each ball of radius r_1'' centered on some $u\in K$ is entirely disjoint from C, which means it is contained in $\mathcal{W}_1$. This concludes the proof. $\square$

Lemma 54. *Let X be a normed space, $C\subset X$ be weakly closed and $K\subset X$ weakly compact, and assume that $C\cap K=\emptyset$. Then there exists $\nu>0$ such that*

$$dist(C,K)=\inf\{\|u-v\|_X\mid u\in C,\,v\in K\}>\nu.$$

Proof. Let us suppose, on the contrary, that $\mathrm{dist}(C,K)=0$. Then we can find sequences $\{u_n\}_{n\in\mathbb{N}}\subset C$ and $\{v_n\}_{n\in\mathbb{N}}\subset K$ such that, for every $n\in\mathbb{N}$, we have

$$\|u_n-v_n\|_X<\frac{1}{n}.$$

Since K is weakly compact, by the Eberlein-Smulian theorem it is also sequentially weakly compact, and there exists $v\in K$ such that $v_n\rightharpoonup v$. We claim that actually v is a weak limit also for the sequence $\{u_n\}_{n\in\mathbb{N}}$, and we have the absurd since then v is forced to be an element

[5]Here we chose $\tilde r$ such that $2\tilde r\subset W_1$ in order to guarantee that:

$$\mathrm{clos}\left((\phi^u)^{-1}(U_3\times 2\tilde rac/3B)\right)\subset\mathrm{clos}\,(U_3\times\tilde rB)\,.\tag{4.8}$$

We will need this property in the proof of Corollary 55.

[6]This follows from this elementary fact. Let $F:P\times K\to\mathbb{R}$ be a continuous function, where P and K are (just) *topological spaces* and K is compact. Define $f(p)=\max_{k\in K}F(p,k)$. Then f is continuous (and the analogue statement with max replaced with min is also true). The proof is easy and left to the reader.

of C. But this easily follows from the fact that, if Λ is any norm-one linear functional on X, the following line holds

$$\Lambda(u_n - v) = \Lambda(u_n - v_n) + \Lambda(v_n - v) \le \|\Lambda\|_{X^*}\|u_n - v_n\|_X + \Lambda(v_n - v) \le \frac{1}{n} + \Lambda(v_n - v) \to 0.$$

□

Corollary 55. *Keeping the notation of Lemma 53, for every soft* $u_0 \in \mathcal{A}^E$ *the function:*

$$g_1 : I_1 \times B(0, r_1') \times \mathcal{W}_1 \to \overline{I}_1 \times \overline{B}_1$$

giving for every $(x, y, u) \in I_1 \times B(0, r_1') \times \mathcal{W}_1$ *the unique solution to:*

$$\phi^u(g_1(x, y, u)) = (x, y), \tag{4.9}$$

is well defined and continuous (for the strong topologies).

Proof. For every $u \in \mathcal{W}_1$ the inverse of ϕ^u is defined on $\overline{I}_1 \times \overline{B}_1$ and continuous, hence $g_1(x, y, u)$ is well defined. To prove that it is continuous for the strong topology, consider a sequence $\{(x_n, y_n, u_n)\}_n \subset I_1 \times B(0, r_1') \times \mathcal{W}_1$ converging to $(\overline{x}, \overline{y}, \overline{u}) \in I_1 \times B(0, r_1') \times \mathcal{W}_1$. Then:

$$\phi^{u_n}(g_1(x_n, y_n, u_n)) = (x_n, y_n).$$

Denote by $g^n = g_1(x_n, y_n, u_n)$; since $\overline{I}_1 \times \overline{B}_1$ is compact, let $g^{n_k} \to \overline{g} \in \overline{I}_1 \times \overline{B}_1$. Then, by continuity of ϕ we have:

$$(\overline{x}, \overline{y}) = \lim_{k\to\infty} \phi^{u_{n_k}}(g^{n_k}) = \lim_{k\to\infty} \phi^{\overline{u}}(\overline{g})$$

which proves $\overline{g}$ is a solution to (4.9); this solution is unique (because of (4.8)), hence:

$$\overline{g} = g_1(\overline{x}, \overline{y}, \overline{u}).$$

This proves that the bounded sequence $\{g^n\}_n$ has only one accumulation point $\overline{g}$, hence the all sequence converges itself to $\overline{g}$. □

3.2. Lipschitz Inverses. The previous section provided a convenient change of coordinates to express the last m components of $G_{u_0}(u, x, y)$; indeed these were linearized after we changed the differentiable structure (Lemma 53), and the equation $G_{u_0}(u, x, y) = w$ can now be reinterpreted as

$$G_{u_0}(u, x, \psi^u) = (\varphi_0(u, x, \psi^u) - \varphi_0(u, 0, 0), \psi_1^u, \ldots, \psi_m^u) = w. \tag{4.10}$$

Using Corollary 55 it is evident how to choose the ψ coordinates in order to solve (4.10): this means that the problem is reduced at this point in finding a continuous solution to the single real equation

$$x \mapsto \varphi_0(u, x, \psi^u) - \varphi_0(u, 0, 0)$$

where, we stress again this point, the ψ^u-variables are treated as *parameters*.

Observe at first that if we want to study the x-derivative of $G_{u_0}(u, x, \psi)$, we may reduce ourselves to the case $\psi^u = 0$. Indeed, if we switch back for a moment to the (x, y)-coordinates, we see that

$$\begin{aligned}\varphi_0(u, x, y) &= \overline{\lambda}\varphi(u + \alpha_{u_0}(x, y)) = \lambda F(u + \alpha_{u_0}(x, y)) \\ &= \lambda F(u + v_{|x|}^{\operatorname{sgn}(x)} + y_1 e_1(u_0) + \ldots + y_m e_m(u_0)),\end{aligned}$$

where the equality in the second line follows from the fact that $\overline{\lambda} = (\lambda, 0)$, by the corank one assumption on singular curves. Then, the identity

$$\left.\frac{\partial \varphi_0}{\partial x}\right|_{(u,x,y)} = \left.\frac{\partial \varphi_0}{\partial x}\right|_{(\widetilde{u},x,0)}, \quad \text{with } \widetilde{u} = u + y_1 e_1(u_0) + \ldots + y_m e_m(u_0),$$

shows that the x-derivative of G_{u_0} at the point (u,x,y) coincides with the x-derivative of G_{u_0} evaluated at the point $(\widetilde{u}, x, 0)$. Finally, notice that equation (4.7) implicitly defines $\psi^u(x,0) = 0$.

Let us fix $v \in \mathcal{P}$ and $s_0 \in \mathbb{R}$; our goal is to prove the following Proposition.

Proposition 56. *For every $v \in \mathcal{P}$, the following estimates hold:*

$$\lim_{s\to 0} \frac{F(u+v_{s+s_0}) - F(u+v_{s_0})}{s} = \frac{sgn(s_0)}{2}\int_0^1 (P_0^{1,u})_*[g^{\bar{t},u}_{w(\theta)}, g^{\bar{t},u}_{v(\theta)}]d\theta + R(u,s_0), \tag{4.11}$$

$$\limsup_{s\to 0} \frac{F(u+v_s) - F(u)}{s} = \int_0^1 (P_0^{1,u})_*[g^{\bar{t},u}_{w(\theta)}, g^{\bar{t},u}_{v(\theta)}]d\theta,$$

$$\liminf_{s\to 0} \frac{F(u+v_s) - F(u)}{s} = -\int_0^1 (P_0^{1,u})_*[g^{\bar{t},u}_{w(\theta)}, g^{\bar{t},u}_{v(\theta)}]d\theta.$$

Moreover, the map $(u,s_0) \mapsto R(u,s_0)$ is weakly continuous with respect to u, satisfies the equality $R(u,0) = 0$ for every u and, for $|s_0|$ sufficiently small, there holds the estimate

$$|R(u,s_0)| \leq a(u)(|s_0|^{1/4}).$$

Proof. By the results in [**AS04**, Section 2.6], the limit (4.11) exists if and only if

$$\lim_{s\to 0} \frac{G(u+v_{s+s_0}) - G(u+v_{s_0})}{s}$$

exists, with (observe that we give the inline definition of $g^{t,u}_v$)

$$G(v_q) = \overrightarrow{\exp}\int_{\bar{t}}^{\bar{t}+|q|^{3/4}} (P_t^{0,u})_* f_{v_q(t)}dt = \overrightarrow{\exp}\int_{\bar{t}}^{\bar{t}+|q|^{3/4}} g^{t,u}_{v_q(t)}dt.$$

Moreover, the following identity holds:

$$\lim_{s\to 0} \frac{F(u+v_{s+s_0}) - F(u+v_{s_0})}{s} = \left(P_0^{1,u}\right)_* \left(\lim_{s\to 0} \frac{G(v_{s+s_0}) - G(v_{s_0})}{s}\right).$$

The Volterra series [**AS04**, Section 2.4] provides the expansion:

$$G(v_q) = Id + \int_{\bar{t}}^{\bar{t}+|q|^{3/4}} g^{t,u}_{v_q(t)}dt + \iint\limits_{\bar{t}\leq\tau\leq t\leq\bar{t}+|q|^{3/4}} g^{\tau,u}_{v_q(\tau)} \circ g^{t,u}_{v_q(t)}d\tau dt + |q|^{3/2}O(1),$$

where the last term $|q|^{3/2}O(1)$ is a consequence of the fact that $\|v\|_{L^2} \leq 1$ and the change of variables:

$$\theta = \frac{t-\bar{t}}{|q|^{3/4}}.$$

This term weakly depends on u, since so does the flow $u \mapsto P_{t_1}^{t_2,u}$. The behavior of (4.11) will be thus determined by a careful analysis, as $s \to 0$, of the expression:

(4.12)

$$\begin{aligned}
\frac{G(v_{s+s_0}) - G(v_{s_0})}{s} &= \frac{1}{s}\Big(\underbrace{\int_{\bar t}^{\bar t+|s+s_0|^{3/4}} g^{t,u}_{v_{s+s_0}(t)}dt - \int_{\bar t}^{\bar t+|s_0|^{3/4}} g^{t,u}_{v_{s_0}(t)}dt}_{I'(s)} \\
&+ \underbrace{\iint_{\bar t\le\tau\le t\le\bar t+|s+s_0|^{3/4}} g^{\tau,u}_{v_{s+s_0}(\tau)}\circ g^{t,u}_{v_{s+s_0}(t)}d\tau dt - \iint_{\bar t\le\tau\le t\le\bar t+|s_0|^{3/4}} g^{\tau,u}_{v_{s_0}(\tau)}\circ g^{t,u}_{v_{s_0}(t)}d\tau dt}_{I''(s)} \\
&+ (|s+s_0|^{3/2} - |s_0|^{3/2})O(1)\Big) \\
&= \frac{1}{s}\Big(I'(s) + I''(s) + (|s+s_0|^{3/2} - |s_0|^{3/2})O(1)\Big)
\end{aligned}$$

3.2.1. *Estimate on the Remainders.* In subsequent calculations, we will frequently use the fact that, for any fixed $u, v \in L^2(I, \mathbb{R}^l)$, the map $t \mapsto g_v^{t,u}$ is Lipschitzian. The following lemma gives a quantitative version of this statement.

LEMMA 57. *Let $u, v \in L^2(I, \mathbb{R}^l)$, and let $g_1(s, s_0)$ and $g_2(s, s_0)$ be any two real numbers depending on s and s_0. Then we have the estimate*

$$\int_0^1 \left| g_{v(\theta)}^{\bar t+g_1(s,s_0)\theta} - g_{v(\theta)}^{\bar t+g_2(s,s_0)\theta}\right| d\theta \le \sqrt{l}a(u)|g_1(s,s_0) - g_2(s,s_0)|\|v\|_{L^2},$$

where a is some weakly continuous function of u satisfying $a(0) = 0$.

PROOF. We have

$$\begin{aligned}
&\int_0^1 \left| g_{v(\theta)}^{\bar t+g_1(s,s_0)\theta} - g_{v(\theta)}^{\bar t+g_2(s,s_0)\theta}\right| d\theta \\
&= \int_0^1 \left|(P^{0,u}_{\bar t+g_1(s,s_0)\theta} - P^{0,u}_{\bar t+g_2(s,s_0)\theta})_* f_{v(\theta)}\right| d\theta \\
&\le \int_0^1 \sum_{i=1}^l |v_i(\theta)| \left|(P^{0,u}_{\bar t+g_1(s,s_0)\theta} - P^{0,u}_{\bar t+g_2(s,s_0)\theta})_* X_i\right| d\theta \\
&\le \max_{i=1,\dots,l}|X_i| \int_0^1 \sum_{i=1}^l |v_i(\theta)| \left\|(P^{0,u}_{\bar t+g_1(s,s_0)\theta} - P^{0,u}_{\bar t+g_2(s,s_0)\theta})_*\right\|_\infty d\theta \\
&\le a(u)|g_1(s,s_0) - g_2(s,s_0)| \left(\int_0^1 \left(\sum_{i=1}^l |v_i(\theta)|\right)^2 d\theta\right)^{1/2} \\
&\le \sqrt{l}a(u)|g_1(s,s_0) - g_2(s,s_0)|\|v\|_{L^2},
\end{aligned}$$

where the second-last line follows by the Lipschitz continuity of the flow and the Cauchy-Schwartz inequality. □

3.2.2. *Expansion of the First-Order Term.* We consider here the asymptotic expansion of the first-order term in (4.12). We have:

$$\begin{aligned} I'(s) =& |s+s_0|^{1/2}\int_0^1 g_{v(\theta)}^{\bar t+|s+s_0|^{3/4}\theta,u}d\theta - |s_0|^{1/2}\int_0^1 g_{v(\theta)}^{\bar t+|s_0|^{3/4}\theta,u}d\theta \\ =& \underbrace{\left(|s+s_0|^{1/2}-|s_0|^{1/2}\right)\int_0^1 g_{v(\theta)}^{\bar t+|s+s_0|^{3/4}\theta,u}d\theta}_{A} - \underbrace{|s_0|^{1/2}\int_0^1 g_{v(\theta)}^{\bar t+|s_0|^{3/4}\theta,u} - g_{v(\theta)}^{\bar t+|s+s_0|^{3/4}\theta,u}d\theta}_{B}. \end{aligned}$$

If we now apply the conclusions of Lemma 57, to both the terms denoted by A and B (observe that for the first one we are implicitly using the fact that, as $v \in \mathcal{P}$, we have $\int_0^1 g_{v(\theta)}^{\bar t}dt = 0$), we deduce that:

$$\begin{aligned} \lim_{s\to 0}\frac{|A|}{s} &\le a(u)\lim_{s\to 0}\frac{\left||s+s_0|^{1/2}-|s_0|^{1/2}\right|}{s}|s+s_0|^{3/4} = \frac{1}{2}a(u)|s_0|^{1/4}, \\ \lim_{s\to 0}\frac{|B|}{s} &\le a(u)\lim_{s\to 0}\frac{\left||s+s_0|^{3/4}-|s_0|^{3/4}\right|}{s}|s_0|^{1/2} = \frac{3}{4}a(u)|s_0|^{1/4}. \end{aligned}$$

3.2.3. *Expansion of the Second-Order Term.* We turn now to the more complicated second order term in (4.12). Let us first recall the equality $w(t) = \int_0^t v(\tau)d\tau$; we have by construction that w is a Lipschitz function from I into $\mathbb{R}^l$. Then we recall two useful formulas [**ABB**, Theorem 11.13]:

$$\int_{t_1}^{t_2} g_{v(t)}^{t,u}dt = g_{w(t_2)}^{t_2,u} - g_{w(t_1)}^{t_1,u} - \int_{t_1}^{t_2}\dot g_{w(t)}^{t,u}dt, \tag{4.13}$$

$$\begin{aligned} \iint\limits_{t_1\le\tau\le t\le t_2}[g_{v(\tau)}^{\tau,u}, g_{v(t)}^{t,u}]d\tau dt &= \int_{t_1}^{t_2}[\int_{t_1}^{t} g_{v(\tau)}^{\tau,u}d\tau, g_{v(t)}^{t,u}]dt \\ &= \int_{t_1}^{t_2}[g_{v(\tau)}^{\tau,u}, \int_{\tau}^{t_2} g_{v(t)}^{t,u}dt]d\tau. \end{aligned}$$

We start with the observation that:

$$\begin{aligned} \iint\limits_{\bar t\le\tau\le t\le\bar t+|q|^{3/4}} g_{v_q(\tau)}^{\tau,u}\circ g_{v_q(t)}^{t,u}d\tau dt =& \frac{1}{2}\int_{\bar t}^{\bar t+|q|^{3/4}} g_{v_q(\tau)}^{\tau,u}d\tau \circ \int_{\bar t}^{\bar t+|q|^{3/4}} g_{v_q(t)}^{t,u}dt \\ &+ \frac{1}{2}\iint\limits_{\bar t\le\tau\le t\le\bar t+|q|^{3/4}}[g_{v_q(\tau)}^{\tau,u}, g_{v_q(t)}^{t,u}]d\tau dt. \end{aligned}$$

Then (let us call for a moment $s' = s+s_0$):

$$\begin{aligned} I''(s) =& \frac{1}{2}\left(\int_{\bar t}^{\bar t+|s'|^{3/4}} g_{v_{s'}(\tau)}^{\tau,u}d\tau\circ\int_{\bar t}^{\bar t+|s'|^{3/4}} g_{v_{s'}(t)}^{t,u}dt + \iint\limits_{\bar t\le\tau\le t\le\bar t+|s'|^{3/4}}[g_{v_{s'}(\tau)}^{\tau,u}, g_{v_{s'}(t)}^{t,u}]d\tau dt\right) \\ &- \frac{1}{2}\left(\int_{\bar t}^{\bar t+|s_0|^{3/4}} g_{v_{s_0}(\tau)}^{\tau,u}d\tau\circ\int_{\bar t}^{\bar t+|s_0|^{3/4}} g_{v_{s_0}(t)}^{t,u}dt + \iint\limits_{\bar t\le\tau\le t\le\bar t+|s_0|^{3/4}}[g_{v_{s_0}(\tau)}^{\tau,u}, g_{v_{s_0}(t)}^{t,u}]d\tau dt\right) \end{aligned}$$

We take into considerations the terms without the commutator: if we add and subtract the term

$$\int_{\bar t}^{\bar t+|s'|^{3/4}} g^{\tau,u}_{v_{s'}(\tau)}d\tau \circ \int_{\bar t}^{\bar t+|s_0|^{3/4}} g^{t,u}_{v_{s_0}(t)}dt$$

we obtain the expression:

$$C = \int_{\bar t}^{\bar t+|s'|^{3/4}} g^{\tau,u}_{v_{s'}(\tau)}d\tau \left(\int_{\bar t}^{\bar t+|s'|^{3/4}} g^{t,u}_{v_{s'}(t)}dt - \int_{\bar t}^{\bar t+|s_0|^{3/4}} g^{t,u}_{v_{s_0}(t)}dt \right) + \int_{\bar t}^{\bar t+|s_0|^{3/4}} g^{t,u}_{v_{s_0}(t)}dt \left(\int_{\bar t}^{\bar t+|s'|^{3/4}} g^{\tau,u}_{v_{s'}(\tau)}d\tau - \int_{\bar t}^{\bar t+|s_0|^{3/4}} g^{\tau,u}_{v_{s_0}(\tau)}d\tau \right).$$

Using again Lemma 57 and the same arguments as in the expansion of the first-order terms, we easily deduce that:

$$\lim_{s\to 0} \frac{|C|}{s} \leq \frac{3}{4} a(u)|s_0|^{1/2}.$$

We move now to the commutator term: again we begin with a general computation, namely using (4.13) and the fact that $w_q(\bar t) = w_q(\bar t + |q|^{3/4}) = 0$, we can write:

$$\begin{aligned}\int_{\bar t}^{\bar t+|q|^{3/4}} [\int_{\bar t}^{t} g^{\tau,u}_{v_q(\tau)}d\tau, g^{t,u}_{v_q(t)}]dt &= \int_{\bar t}^{\bar t+|q|^{3/4}} [g^{t,u}_{w_q(t)}, g^{t,u}_{v_q(t)}]dt - \int_{\bar t}^{\bar t+|q|^{3/4}} [\int_{\bar t}^{t} \dot g^{\tau,u}_{w_q(\tau)}d\tau, g^{t,u}_{v_q(t)}]dt \\ &= \int_{\bar t}^{\bar t+|q|^{3/4}} [g^{t,u}_{w_q(t)}, g^{t,u}_{v_q(t)}]dt - \int_{\bar t}^{\bar t+|q|^{3/4}} [\dot g^{\tau,u}_{w_q(\tau)}, \int_{\tau}^{\bar t+|q|^{3/4}} g^{t,u}_{v_q(t)}dt]d\tau \\ &= \int_{\bar t}^{\bar t+|q|^{3/4}} [g^{t,u}_{w_q(t)}, g^{t,u}_{v_q(t)}]dt + \int_{\bar t}^{\bar t+|q|^{3/4}} [\dot g^{\tau,u}_{w_q(\tau)}, g^{\tau,u}_{w_q(\tau)}]d\tau \\ &+ \int_{\bar t}^{\bar t+|q|^{3/4}} [\dot g^{\tau,u}_{w_q(\tau)}, \int_{\tau}^{\bar t+|q|^{3/4}} \dot g^{t,u}_{w_q(t)}dt]d\tau.\end{aligned}$$

We immediately realize that just the first summand matters for our purposes. Indeed in both the other summands there is at least a double change of variables $\theta = \frac{t-\bar t}{|q|^{3/4}}$ plus a differentiation, which yields a power of $|q|$ not smaller than $7/4$. By virtue of the equalities

$$\begin{aligned} v_q(\bar t + |q|^{3/4}\theta) &= \frac{1}{|q|^{1/4}} v(\theta), \\ w_q(\bar t + |q|^{3/4}\theta) &= \int_{\bar t}^{\bar t+|q|^{3/4}\theta} \frac{1}{|q|^{1/4}} v\left(\frac{\tau - \bar t}{|q|^{3/4}}\right) d\tau \\ &= |q|^{1/2} \int_0^{\theta} v(\zeta)d\zeta = |q|^{1/2} w(\theta),\end{aligned}$$

we have

$$\int_{\bar t}^{\bar t+|q|^{3/4}} [g^{t,u}_{w_q(t)}, g^{t,u}_{v_q(t)}]dt = = |q|^{3/4} \int_0^1 [g^{\bar t+|q|^{3/4}\theta,u}_{w_q(\bar t+|q|^{3/4}\theta)}, g^{\bar t+|q|^{3/4}\theta,u}_{v_q(\bar t+|q|^{3/4}\theta)}]d\theta = |q| \int_0^1 [g^{\bar t+|q|^{3/4}\theta,u}_{w(\theta)}, g^{\bar t+|q|^{3/4}\theta,u}_{v(\theta)}].$$

Call

$$D=|s'|\int_0^1[g_{w(\theta)}^{\bar t+|s'|^{3/4}\theta,u},g_{v(\theta)}^{\bar t+|s'|^{3/4}\theta,u}]d\theta-|s_0|\int_0^1[g_{w(\theta)}^{\bar t+|s_0|^{3/4}\theta,u},g_{v(\theta)}^{\bar t+|s_0|^{3/4}\theta,u}]d\theta.$$

Adding and subtracing the common term

$$|s'|\int_0^1[g_{w(\theta)}^{\bar t+|s_0|^{3/4}\theta,u},g_{v(\theta)}^{\bar t+|s_0|^{3/4}\theta,u}]d\theta,$$

we end up with

$$D=\underbrace{|s'|\int_0^1[g_{w(\theta)}^{\bar t+|s'|^{3/4}\theta,u},g_{v(\theta)}^{\bar t+|s'|^{3/4}\theta,u}]-[g_{w(\theta)}^{\bar t+|s_0|^{3/4}\theta,u},g_{v(\theta)}^{\bar t+|s_0|^{3/4}\theta,u}]d\theta}_{L}$$
$$+\underbrace{(|s'|-|s_0|)\int_0^1[g_{w(\theta)}^{\bar t+|s_0|^{3/4}\theta,u},g_{v(\theta)}^{\bar t+|s_0|^{3/4}\theta,u}]d\theta}_{H}.$$

On the one hand, by Lemma 57, there holds

$$\begin{aligned}L&=|s'|\int_0^1[g_{w(\theta)}^{\bar t+|s'|^{3/4}\theta,u},g_{v(\theta)}^{\bar t+|s'|^{3/4}\theta,u}-g_{v(\theta)}^{\bar t+|s_0|^{3/4}\theta,u}]+[g_{w(\theta)}^{\bar t+|s'|^{3/4}\theta,u}-g_{w(\theta)}^{\bar t+|s_0|^{3/4}\theta,u},g_{v(\theta)}^{\bar t+|s_0|^{3/4}\theta,u}]d\theta\\&\le a(u)|s'|(|s'|^{3/4}-|s_0|^{3/4}),\end{aligned}$$

which implies that

$$\lim_{s\to0}\frac{|L|}{s}\le\frac{3}{4}a(u)|s_0|^{3/4};$$

on the other hand, whenever $s_0\neq0$, we have by similar reasonings (notice that $\|v\|_{L^2}\le1$ implies that also $\|w\|_{L^2}\le1$):

$$\lim_{s\to0}\frac{H}{s}=\operatorname{sgn}(s_0)\int_0^1[g_{w(\theta)}^{\bar t,u},g_{v(\theta)}^{\bar t,u}]d\theta+R(u,s_0),\quad\text{with }|R(u,s_0)|\le a(u)|s_0|^{3/4},$$

while if $s_0=0$ the following is true:

$$\limsup_{s\to0}\frac{H}{s}=\int_0^1[g_{w(\theta)}^{\bar t,u},g_{v(\theta)}^{\bar t,u}]d\theta,\quad\liminf_{s\to0}\frac{H}{s}=-\int_0^1[g_{w(\theta)}^{\bar t,u},g_{v(\theta)}^{\bar t,u}]d\theta.$$

This conlcudes the proof of Proposition 56. □

LEMMA 58. *For every soft $u_0\in\mathcal{A}^E$ there exists a weak neighborhood $\mathcal{W}_2$ of u_0, a neighborhood $I_2\subset\mathbb{R}$ of zero, a neighborhood $W_2\subset\mathbb{R}^m$ of zero and a positive number $r_2'>0$ such that for every $(u,\psi)\in\mathcal{W}_2\times W_2$ the function:*

$$x\mapsto\varphi_0(u,x,\psi)-\varphi_0(u,0,0),\quad x\in I_2$$

is invertible with inverse defined on $(-r_2',r_2')$ and which depends continuously on u and ψ. More precisely, there exists:

$$g_2:\mathcal{W}_2\times(-r_2',r_2')\times W_2\to I_2$$

which is continuous for the strong topology and such that:

$$g_2(u,s,\psi)\text{ is the unique solution to }\varphi_0(u,g_2(u,s,\psi),\psi)-\varphi_0(u,0,0)=s$$

Moreover there exist $r_2''>0$ and a weak neighborhood $\mathcal{V}_2\subset\mathcal{W}_2$ such that $B(u,r_2'')\subset\mathcal{W}_2$ for every $u\in\mathcal{V}_2\cap\{J\le E\}$.

PROOF. Consider the map $G_0(u,x,\psi) = \varphi_0(u,x,\psi) - \varphi_0(u,0,0) = \lambda F(u,x,\psi) - \lambda F(u,0,0)$. Proposition 56 yields that $G_0(u,x,0)$ is both weakly continuous in the u-variable and Lipschitz continuous in the x-variable. Let us define:

$$H_0(u,x,\psi) = \min \left| \frac{\partial G_0}{\partial x}\bigg|_{(u,x,\psi)} \right|$$

where the min is taken over all the elements in the Clarke x-subderivative of G_0. Our choice of $v^\pm$ in (4.5) ensures that:

$$\liminf_{x\to 0^\pm} \frac{\partial G_0(u,x,0)}{\partial x} = \pm \int_0^1 \langle (P_{\bar{t}}^{1,u_0})^* \lambda, [f_{w^\pm(\theta)}, f_{v^\pm(\theta)}] \rangle d\theta > 0.$$

In particular the subdifferential $H(u,x,0)$ is not zero. By Clarke's Implicit Function Theorem [**?**] and the weak continuity of $u \mapsto H_0(u,x,0)$, we deduce that there exist a weak neighborhood $\mathcal{W}_2'$ of u_0 and $r>0$ such that $G_0(u,x,0) : (-r,r) \to \mathbb{R}$ is an homeomorphism onto its image, for every $u \in \mathcal{W}_2'$.

Consider the linear map

$$\Phi(y_1,\ldots,y_m) = y_1 e_1(u_0) + \cdots + y_m e_m(u_0);$$

being (weakly) continuous, there exist a neighborhood $W_2' \subset \mathbb{R}^m$ of zero and a weak neighborhood $\mathcal{W}_2''$ of u_0 such that $u + \Phi(y)$ belongs to $\mathcal{W}_2'$ for every $(u,y) \in \mathcal{W}_2'' \times W_2'$.

Finally, as the map $(\psi_1,\ldots,\psi_m) \mapsto (y_1,\ldots,y_m)$ is also continuous, we conclude that there exists a further neighborhood $W_2'' \subset \mathbb{R}^m$ of zero such that $H_0(u,x,\psi)$ is of maximal rank whenever $(u,\psi) \in \mathcal{W}_2'' \times W_2''$.

We take advantage of this fact to show that, for every $(u,\psi) \in \mathcal{W}_2'' \times W_2''$, the map $G_0(u,\cdot,\psi) : (-r,r) \to \mathbb{R}$ is injective. Indeed, let a be any positive number such that $H(u,x,\psi) > a$ on $\mathcal{W}_2'' \times (-r,r) \times W_2''$; then, by the Lipschitz version of the mean value theorem [**Cla83**, Theorem 2.6.5], we infer that:

$$|G_0(u,x_1,\psi) - G_0(u,x_2,\psi)| > a|x_1 - x_2|.$$

Again we can use [**Cla76**, Lemma 5] to conclude that, whenever $r' < r$, then:

$$G_0(u,\cdot,\psi)(-r',r') \supset G_0(u,0,\psi) + (-r'a, r'a).$$

Moreover, the weak continuity of $u \mapsto G_0(u,x,\psi)$, and the fact that $G_0(u,0,0) = 0$, imply that for every $r>0$, there exists a weak open set $\mathcal{W}_2^{\text{weak}}(r) \times W_2(r)$ on which $|G_0(u,0,\psi)| < ar'/3$. Then, if $r' < r$, for every $(u,\psi) \in (\mathcal{W}_2'' \cap \mathcal{W}_2^{\text{weak}}(r')) \times (W_2'' \cap W_2(r'))$ we will have that:

$$G_0(u,\cdot,\psi) \supset G_0(u,0,\psi) + (-r'a, r'a) \supset (-2/3r'a, 2/3r'a).$$

Choose $0 < \widetilde{r} < r/2$, and let $\mathcal{W}_3^{\text{weak}} = \mathcal{W}_2'' \cap \mathcal{W}_2^{\text{weak}}(\widetilde{r})$ and $W_3 = W_2'' \cap W_2(\widetilde{r})$; our previous arguments then show that, for every $(u,\psi) \in \mathcal{W}_3^{\text{weak}} \times W_3$, the map

$$G_0(u,\cdot,\psi) : (-\widetilde{r},\widetilde{r}) \to \mathbb{R}$$

is an homeomorphism onto its image, and its image contains $(-2/3\widetilde{r}a, 2/3\widetilde{r}a)$.

Similarly as in Lemma 53, we define the weakly continuous functions:

$$\alpha_1(u) = \min_{(x,\psi)\in[-\widetilde{r},\widetilde{r}]\times\overline{W}_3} H_0(u,x,\psi), \quad \text{and} \quad \alpha_2(u) = \max_{\psi\in\overline{W}_3} |G_0(u,0,\psi)|$$

and, for $\epsilon > 0$ small enough, the weakly open sets:

$$\mathcal{W}_2 = \{\alpha_1 > a, \alpha_2 < a\tilde{r}/3\}, \quad \mathcal{V}_2 = \{\alpha_1 > a+\epsilon, \alpha_2 < a\tilde{r}/3 - \epsilon\}, \quad W_2 = W_3 \quad I_2 = (-\widetilde{r},\widetilde{r}).$$

Then we conclude as in Lemma 53, by choosing e.g. $r_2' = 2/3\tilde{r}a$, and where the existence of r_2'' is guaranteed by applying Lemma 54 to the sets $(\mathcal{W}_2)^c$ and $\overline{\mathcal{V}}_2 \cap \{J \le E\}$. Finally, the assertion on the existence and the continuity of the function g_2 follows literally as in Corollary 55. □

3.3. Proof of Proposition 52. Let $u_0 \in \mathcal{A}^E$ be soft and define $\mathcal{W} = \mathcal{W}_1 \cap \mathcal{W}_2$ and $\mathcal{V} = \mathcal{V}_1 \cap \mathcal{V}_2$. Let $r_2 = \min\{r_1'', r_2''\}$ and $r_1, r_3 > 0$ be such that:

$$B(0, r_1) \subset (I_1 \cap (-r_2', r_2')) \times (B(0, r_1') \cap W_2) \quad \text{and} \quad B(0, r_3) \supset I_2 \times B_1.$$

(All these objects have been constructed in Lemma 53 and Lemma 58.) Then the function g defined by:

$$g(x, y, u) = (g_2(x, u), g_1(g_2(x, u), y, u))$$

verifies the required properties.

4. An Implicit Function Theorem

THEOREM 59. *Assume all abnormal controls with $J \le E$ are soft. There exists a neighborhood $\mathcal{W}(\mathcal{A}^E)$ and positive numbers $r_1, r_2, r_3 > 0$ such that for every $u_0 \in \mathcal{W}(\mathcal{A}^E)$ there exists a function:*

$$\sigma_{u_0} : B(0, r_1) \times B(u_0, r_2) \to L^2(I, \mathbb{R}^l)$$

which is continuous for the strong topology and such that:

$$\sigma_{u_0}(0, u) = 0 \quad \text{and} \quad \varphi(u + \sigma_{u_0}(w, u)) = \varphi(u) + w \quad \forall (w, u) \in B(0, r_1) \times B(u_0, r_2).$$

Moreover the family $\{\sigma_{u_0}\}_{u_0 \in \mathcal{W}(\mathcal{A}^E)}$ is equicontinuous.

PROOF. Every $u \in \mathcal{A}^E$ is soft and we can consider the weak open sets $\mathcal{V}(u) \subset \mathcal{W}(u)$, the positive numbers $r_1(u), r_2(u) > 0$ and the functions $g = g_u$ and α_u constructed in Section 3. Notice that $\mathcal{A}^E$ is weakly compact (it is a weakly closed set in the weakly closed ball $\{J \le E\}$). Then $\{\mathcal{V}(u)\}_{u \in \mathcal{A}^{\mathcal{E}}}$ is a weak cover of $\mathcal{A}^E$ and consequently there exist $u_1, \ldots, u_k \in \mathcal{A}^E$ such that:

$$\mathcal{W}(\mathcal{A}^E) = \mathcal{V}_{u_1} \cup \cdots \cup \mathcal{V}_{u_k}$$

is an open neighborhood of $\mathcal{A}^E$. Set $r_1 = \min_i\{r_1(u_i)\}$ and $r_2 = \min_i\{r_2(u_i)\}$.

Pick $u \in \mathcal{W}(\mathcal{A}^E)$. Then $u \in \mathcal{W}_{u_i}$ for some $i \in \{1, \ldots, k\}$ and we define the function σ_u by:

$$\sigma_u(w, v) = \alpha_{u_i}(g_{u_i}(w, v)) \quad \text{for } (w, v) \in B(0, r_1) \times B(u, r_2).$$

Notice that g_{u_i} was defined on $B(0, r_1(u_i)) \times \mathcal{W}_{u_i}$; on the other hand since $u \in \mathcal{V}_{u_i}$ then $B(u, r_2) \subset \mathcal{W}_{u_i}$ and the domain of σ_u is contained in the domain of definition of g_{u_i}.

The family $\{\sigma_u\}_{u \in \mathcal{W}(\mathcal{A}^E)}$ is equicontinuous simply because it is finite. □

5. Regular Controls: Gradient Flow

Throughout this section we will assume that all abnormal controls with $J \le E$ are soft.

REMARK 24. As the endpoint map is weakly continuous, the set $F^{-1}(y)$ is weakly closed. Then $\mathcal{U}(y)^E = \{u \in \mathcal{U} \mid J(u) \le E\} \cap F^{-1}(y)$ is weakly compact in $\mathcal{U}^E$, this latter set also being weakly compact.

Let $\mathcal{W}(\mathcal{A}^E)$ be the weak neighborhood of the abnormal controls with energy less than E constructed in the previous section (abusing of the notation, we tacitly adopt the convention that $\mathcal{W}(\mathcal{A}^E)$ and $\mathcal{A}^E$ are to be intended in the relative topology of $\mathcal{U}(y)^E$; they are, respectively, weakly open and weakly compact), and let $\mathcal{B}^E = \mathcal{U}(y)^E \setminus \mathcal{W}(\mathcal{A}^E)$. Now, $\mathcal{B}^E$ and $\mathcal{A}^E$ are two disjoint weakly compact subsets in $\mathcal{U}(y)^E$, therefore they can be separated by means of weak neighborhoods $\mathcal{V}(\mathcal{A}^E)$ and $\mathcal{V}(\mathcal{B}^E)$. In particular we have the following sequence of inclusions:

$$\mathcal{B}^E \subset \mathcal{V}(\mathcal{B}^E) \subset \mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E),$$

and, by Lemma 54, the weakly compact set $\mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E)$ is strongly separated from $\mathcal{A}^E$. Exploiting this distance from the abnormal set, the idea is to mimic the classical deformation theory via the gradient flow. If we call

$$f = J\big|_{\mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E)}$$

the restriction of the Energy, we can apply verbatim the arguments of [**BL**, Proposition 10] in this case: the only salient fact to be observed is that the set $\mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E)$ is weakly closed in $\mathcal{U}(y)^E$, and, as such, it contains all the weak limits of its sequences (in particular they are all regular points of the endpoint map). Having this in mind, we obtain the following.

PROPOSITION 60 (Palais-Smale condition). *The function f satisfies the Palais-Smale condition, i.e any sequence $\{u_n\}_{n\in\mathbb{N}} \subset \mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E)$ such that*

$$\lim_{n\to\infty} d_{u_n} f \to 0$$

admits a convergent subsequence.

As an immediate consequence we also recover the following adapted version of the standard Palais-Smale Lemma.

COROLLARY 61 (Palais-Smale lemma). *Let $0 < s < E$, and let $\mathcal{N}$ be any open set contained in $\mathcal{U}(y)^E$ (possibly empty). Assume that*

$$(\mathcal{U}(y)^E \setminus (\mathcal{V}(\mathcal{A}^E) \cup \mathcal{N})) \cap \{J = s\} \cap \mathcal{C} = \emptyset;$$

then there exists a positive constant $0 < \eta(s) < 1$ such that

$$\|d_u f\| > \eta, \quad \forall u \in (\mathcal{U}(y)^E \setminus (\mathcal{V}(\mathcal{A}^E) \cup \mathcal{N})) \cap \{s - \eta < J < s + \eta\}.$$

PROOF. We argue by contradiction and assume $\{u_n\}_{n\in\mathbb{N}}$ to be a sequence in $\mathcal{U}(y)^E \setminus (\mathcal{V}(\mathcal{A}^E) \cup \mathcal{N})$ such that

1. $u_n \in \{s - 1/n < J < s + 1/n\}$
2. $\|d_{u_n} f\| < 1/n$.

Then the assumptions of Proposition 60 are satisfied and therefore, passing possibly to a subsequence, we may assume that $\overline{u} = \lim_n u_n$ exists in $\mathcal{U}(y)^E \setminus (\mathcal{V}(\mathcal{A}^E) \cup \mathcal{N})$, as this set is closed. By point 1., $J(\overline{u}) = \lim_n J(u_n) = s$; moreover, by point 2., $\overline{u}$ has to be a critical point of f, which leads to an absurd. □

A further application of Lemma 54 gives $\beta > 0$ such that

$$\text{dist}\left(\mathcal{B}^E, \mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{B}^E)\right) > \beta; \tag{4.14}$$

in this situation we define the (strong) open set

$$\mathcal{W}(\mathcal{B}^E) = \bigcup_{u\in\mathcal{B}^E} B\left(u, \frac{\beta}{2}\right) \tag{4.15}$$

as the union of open balls of radius $\beta/2$, centered on elements of $\mathcal{B}^E$.

The next proposition is a refinement of a classical result, which can for instance be found in [**Cha93**, Chapter 1, Theorem 3.4], adapted to our setting. Notice that this statement is stronger than the analogous statement as in [**Cha93**, Chapter 1, Theorem 3.3], which is in fact a consequence of [**Cha93**, Chapter 1, Theorem 3.4] (the constants depend on the Palais-Smale condition).

PROPOSITION 62. *Let $0 < s < E$, $0 < \delta < \min\{\beta, 1\}$ and $\mathcal{N}' \subset \mathcal{N}$ be two open sets in $\mathcal{U}(y)^E$ such that dist$(\overline{\mathcal{N}'}, \mathcal{U}(y)^E \setminus \mathcal{N}) > \delta$. Assume that there exists $0 < \eta < 1$ such that*

$$\|d_u f\| > \eta, \quad \forall u \in (\mathcal{U}(y)^E \setminus (\mathcal{V}(\mathcal{A}^E) \cup \mathcal{N}')) \cap \{s - \eta < J < s + \eta\},$$

and let

$$t = \eta \frac{\beta}{2}, \quad 0 < \varepsilon' < \eta \frac{\delta}{4}, \quad \text{and } \varepsilon' < \varepsilon'' < \eta.$$

Then there exist a continuous map

$$\Theta : [0, t] \times \mathcal{W}(\mathcal{B}^E) \to \mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E),$$

such that

a) $\Theta(0, \cdot) = Id$,
b) $\Theta(\tau, \cdot) = Id$ *on the set* $\{J \leq s - \varepsilon''\} \cup \{J \geq s + \varepsilon''\}$,
c) *For every* $u \in (\mathcal{W}(\mathcal{B}^E) \setminus \mathcal{N}) \cap \{J \leq s + \varepsilon'\}$, $\Theta(t, u) \in \mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E) \cap \{J \leq s - \varepsilon'\}$,
d) $J(\Theta(\tau, u))$ *is nonincreasing, for any* $(\tau, u) \in [0, t] \times \mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E)$.

PROOF. Define

$$\chi(r) = \begin{cases} 0 & \text{if } r \notin (s - \varepsilon'', s + \varepsilon'') \\ 1 & \text{if } r \in [s - \varepsilon', s + \varepsilon'] \end{cases}$$

to be a smooth function satisfying $0 \leq \chi(r) \leq 1$. Consider two closed subsets $C_1 = \mathcal{U}(y)^E \setminus (\mathcal{V}(\mathcal{A}^E) \cup \mathcal{N}'_{\delta/2})$, where $\mathcal{N}'_p = \{u \in \mathcal{U}(y)^E \,|\, \mathrm{dist}(u, \overline{\mathcal{N}'}) < p\}$, and $C_2 = \overline{\mathcal{N}'} \cap (\mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E))$. Then we can construct the function

$$g(u) = \frac{\mathrm{dist}(u, C_2)}{\mathrm{dist}(u, C_1) + \mathrm{dist}(u, C_2)},$$

so that $0 \leq g(u) \leq 1$, $g \equiv 1$ in C_1 and $g \equiv 0$ in C_2. On the set $\mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E)$, we define the vector field

$$Y(u) = -g(u)\chi(J(u)) \frac{d_u f}{\|d_u f\|^2}.$$

By the standard theory of differential equations on the real line, since $\|Y\| < \frac{1}{\eta}$, it is well-defined the time-x flow of Y starting from the point u_0 for any time $x > 0$, which we will indicate by $\psi^Y_x(u_0)$. Pick $u \in \mathcal{W}(\mathcal{B}^E) \setminus \mathcal{N}$, and assume that it belongs to a ball centered at u_0. If we let $\psi^Y_\cdot(u)$ flow for a time $T_1 \leq t$, by virtue of (4.14) and the inequality

$$\|\psi^Y_{T_1}(u) - u\| \leq \int_0^{T_1} \|Y(\psi^Y_\tau(u))\| d\tau < \frac{\beta}{2},$$

we see that:

$$\|\psi^Y_{T_1}(u) - u_0\| \leq \|u - u_0\| + \|\psi^Y_{T_1}(u) - u\| < \frac{\beta}{2} + \frac{\beta}{2} = \beta,$$

that is $\psi^Y_{T_1}(u)$ belongs to $\mathcal{V}(\mathcal{B}^E) \subset \mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E)$. Then we define the deformation map Θ by

$$\Theta(\tau, u) = \psi^Y_\tau(u), \quad \forall\, (\tau, u) \in [0, t] \times \mathcal{W}(\mathcal{B}^E).$$

Points a), b) and d) are almost immediate: the only non trivial point to verify is c). Let $T_2 = \frac{\delta}{2}\eta < \frac{\beta}{2}\eta = t$; we claim that flowing for time T_2 suffices for our purposes, that is we want to show that:

$$J(\psi^Y_{T_2}(u)) \leq s - \varepsilon', \quad \forall\, u \in (\mathcal{W}(\mathcal{B}^E) \setminus \mathcal{N}) \cap \{s - \varepsilon' < J \leq s + \varepsilon'\}.$$

We observe at first that

$$\frac{d}{dt}J(\psi_\tau^Y(u)) = -g(\psi_\tau^Y(u))\chi(J(\psi_\tau^Y(u))), \tag{4.16}$$

To prove our claim we argue by contradiction, and we assume

$$s - \varepsilon' < J(\psi_\tau^Y(u)) \le s + \varepsilon', \quad \forall\, \tau \in \left[0, \frac{\delta}{2}\eta\right], \tag{4.17}$$

so that $\chi(J(\psi_\tau^Y(u))) \equiv 1$. Moreover, since $\|\psi_\tau^Y(u) - u\| < \frac{\tau}{\eta}$, we also have

$$\begin{aligned}\operatorname{dist}(\psi_\tau^Y(u), \mathcal{N}'_{\delta/2}) &> \operatorname{dist}(u, \mathcal{N}'_{\delta/2}) - \frac{\tau}{\eta}\\ &> \delta - \frac{1}{2}\delta - \frac{1}{\eta}\frac{\eta\delta}{2} = 0,\end{aligned}$$

so that we even have the equality $g(\psi_\tau^Y(u)) \equiv 1$. Finally, it follows from (4.16) that

$$\frac{d}{dt}J(\psi_\tau^Y(u)) = -1,$$

which implies, combined with our choice of $\varepsilon' < \frac{\delta}{4}\eta$, that the following line is true

$$J(\psi_{T_2}^Y(u)) = J(u) - \frac{\delta}{2}\eta \le s + \varepsilon' - \frac{\delta}{2}\eta < s - \varepsilon'.$$

Since this contradicts (4.17), the proof is complete. □

Corollary 63. *Let $\mathcal{N}$ be a neighborhood of*

$$\mathcal{C}_s = (\mathcal{U}^E \setminus \mathcal{V}(\mathcal{A}^E)) \cap \{J = s\} \cap \mathcal{C}.$$

Then there exist $0 < \eta < 1$ and a deformation map Θ satisfying the conclusions of Proposition 62.

Proof. The Palais Smale condition implies that the set $\mathcal{C}_s$ is (sequentially) compact; therefore for sufficiently small values of the parameter $\nu > 0$ the closure of the set

$$\mathcal{C}_s(\nu) = \{u \in \mathcal{U}^E \,|\, \operatorname{dist}(u, \mathcal{C}_s) < \nu\}$$

is contained in $\mathcal{N}$. Let $\overline{\nu}$ be such that the inclusion holds, and define $\mathcal{N}' = \mathcal{C}_s(\overline{\nu})$. Then by construction

$$\mathcal{U}(y)^E \setminus (\mathcal{V}(\mathcal{A}^E) \cup \mathcal{N}') \cap \{J = s\} \cap \mathcal{C} = \emptyset;$$

therefore, by Corollary 61, there exists $0 < \eta < 1$ satisfying the assumptions of Proposition 62, and then we conclude. □

Using somewhat the same ideas as in Proposition 62 we also have the following proposition.

Proposition 64. *Let $0 < E_1 < E_2 < E$ and $\epsilon > 0$. Assume that there exists $0 < \eta < \min\{\epsilon/2, 1\}$ such that*

$$\|d_u f\| > \eta, \quad \forall u \in (\mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E)) \cap \{E_1 + \epsilon - \eta < J < E_2 + \eta\},$$

and let

$$t = \eta\frac{\beta}{2}, \quad \textit{and } 0 < \varepsilon' < \varepsilon'' < \eta.$$

Then there exist a continuous map

$$\Theta : [0, t] \times \mathcal{W}(\mathcal{B}^E) \to \mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E),$$

such that

a) $\Theta(0,\cdot) = Id$,
b) $\Theta(\tau,\cdot) = Id$ *on the set* $\{J \le E_1 + \epsilon - \varepsilon''\} \cup \{J \ge E_2 + \varepsilon''\}$,
c) *For any* $u \in \mathcal{W}(\mathcal{B}^E) \cap \{E_1 + \epsilon \le J \le E_2\}$ *and* $0 \le \tau < \varepsilon'$, $\Theta(\tau, u) \in \mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E)$ *and* $J(\Theta(\tau, u)) = J(u) - \tau$.

Sketch of proof. In this case (the notations are mutuated from Proposition 62), we have

$$\chi(r) = \begin{cases} 0 & \text{if } r \notin (E_1 + \epsilon - \varepsilon'', E_2 + \varepsilon'') \\ 1 & \text{if } r \in [E_1 + \epsilon - \varepsilon', E_2 + \varepsilon'] \end{cases}$$

and

$$Y(u) = -\chi(J(u)) \frac{d_u f}{\|d_u f\|^2}.$$

The flow up to time t is well-defined, since for any point u in $\mathcal{W}(\mathcal{B}^E)$, $\psi^Y_\tau(u)$ stays within $\mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E)$, and we define the deformation map Θ by

$$\Theta(\tau, u) = \psi^Y_\tau(u), \quad \forall\, (\tau, u) \in [0, t] \times \mathcal{W}(\mathcal{B}^E).$$

Only point c) needs a verification: but, since for any $u \in \mathcal{W}(\mathcal{B}^E) \cap \{E_1 + \epsilon \le J \le E_2\}$ and $0 \le \tau < \varepsilon'$

$$J(\psi^Y_\tau(u)) \ge J(u) - \varepsilon' \ge E_1 + \epsilon - \varepsilon',$$

then $\chi(J(\psi^Y_\tau(u))) \equiv 1$, and the claim follows. □

Corollary 65. *Let* $0 < E_1 < E_2 < E$ *be such that*

$$(\mathcal{U}(y)^E \setminus \mathcal{V}(\mathcal{A}^E)) \cap \{E_1 < J \le E_2\} \cap \mathcal{C} = \emptyset.$$

Then, for any $\epsilon > 0$*, the conclusions of Proposition 64 hold on the strip* $\{E_1 + \epsilon \le J \le E_2\}$*.*

Proof. Let $\epsilon > 0$ be fixed. Then, for any $E_1 + \epsilon \le s \le E_2$, Corollary 61 applies, and permits to find the corresponding parameter $0 < \eta(s) < 1$. Since

$$[E_1 + \epsilon, E_2] \subset \bigcup_{s \in [E_1+\epsilon, E_2]} \underbrace{(s - \eta(s), s + \eta(s))}_{I(s)},$$

by compactness we may extract a finite subcover consisting of the open intervals $I(s_1), \dots, I(s_p)$. Now it is sufficient to choose $\eta = \min\{\eta_1, \dots, \eta_p, \epsilon/2\}$ (notice in particular that $0 < \eta < \min\{1, \epsilon/2\}$) to see that the assumption of Proposition 64 are satisfied. □

6. The Serre Fibration Property

We will need the following preliminary lemma, asserting that the property of being a Serre fibration can be verified locally. Notice that if the open cover in the statement of Lemma 66 did not depend on n, then this is classical (see [**Hur55**] for a proof in the more general case of a Hurewicz fibration); instead here we have to work with an open cover that might depend on n, but in the case of a *Serre* fibration this is not an obstacle and the proof remains essentially unchanged.

Lemma 66. *Let* $p : U \to Y$ *be a continuous function between topological spaces such that for every* $n \in \mathbb{N}$ *there exists an open cover* $\mathfrak{U}_n = \{Y_\alpha\}_{\alpha \in A}$ *of* Y *with the property that for every* $\alpha \in A$ *the map*

$$p|_{p^{-1}(Y_\alpha)} : p^{-1}(Y_\alpha) \to Y_\alpha$$

has the homotopy lifting property with respect to all n-dimensional CW-complexes. Then p is a Serre fibration.

PROOF. Recall that in order for p to be a Serre fibration, it is enough to check that it has the homotopy lifting property with respect to all cubes. Let $\tilde{H} : I^n \times I \to Y$ be a homotopy, $\tilde{h}_t = \tilde{H}(\cdot, t)$ and $h_0 : I^n \to U$ be a lift; consider the open cover $\mathfrak{U}_n$ and a subdivision of I^n into small cubes $\{C_\beta\}_{\beta \in B}$ and of I into small intervals $\{I_j\}_{j=1,\dots,N}$ with the property that for every $(\beta, j) \in B \times \{1, \dots, N\}$ there exists $\alpha \in A$ such that:

$$\tilde{H}(C_\beta \times I_j) \subset Y_\alpha. \tag{4.18}$$

We can assume by induction that h_t has been constructed over ∂C_β for every C_β. To extend h_t over the all cube C_β we use (4.18) and notice that it means that we are given a homotopy $\tilde{h}_t|_{C_\beta} : C_\beta \to Y_\alpha$ with a lift:

$$h_t|_{\partial C_\beta} : \partial C_\beta \to p^{-1}(Y_\alpha). \tag{4.19}$$

By assumption $p|_{p^{-1}(Y_\alpha)}$ has the homotopy lifting property with respect to all n-dimensional CW-complexes. This is equivalent to the fact that $p|_{p^{-1}(Y_\alpha)}$ has the homotopy lifting property with respect to all n-dimensional CW-pairs (see [**Hat02**, Section 4.2]); that means exactly that we can lift the homotopy $\tilde{h}_t|_{C_\beta} : C_\beta \to Y_\alpha$ with the constraint (4.19). □

THEOREM 67 (Serre fibration property). *Assume all singular curves with $J \leq E_2$ are soft. If there are no normal geodesics in $\mathcal{U}(y)$ with Energy $E_1 < J \leq E_2$, then for every $\epsilon > 0$ sufficiently small the restriction of the Energy to $\mathcal{U}(y)_{E_1+\epsilon}^{E_2}$ is a Serre fibration.*

PROOF. We will prove that for every $n \in \mathbb{N}$ there exists $\delta = \delta(n)$ such that for every point $s \in [E_1 + \epsilon, E_2]$, denoting by:

$$I(s) = [E_1 + \epsilon, E_2] \cap (s - \delta, s + \delta) \quad \text{and} \quad \mathcal{U}(y)_{I(s)} = \mathcal{U}(y) \cap J^{-1}(I(s)),$$

then $J|_{J^{-1}(I(s)) \cap \mathcal{U}(y)}$ has the homotopy lifting property with respect to all n-dimensional CW-complexes (or, equivalently with respect to all n-dimensional disks). The result will then follow from Lemma 66.

We denote by $\psi_t(\cdot)$ the deformation $\Theta(t, \cdot)$ coming from Proposition 64; it has the property that for every $u \in \mathcal{W}(\mathcal{B}^E) \cap \{E_1 + \epsilon, J \leq E_2\}$ and $t < \epsilon'$:

$$\psi_t(u) \in \mathcal{U}(y)^E \backslash \mathcal{V}(A^\epsilon) \quad \text{and} \quad J(\psi_t(u)) = J(u) - t.$$

Recalling the definition of $\mathcal{W}(\mathcal{B}^E)$ given in (4.15), we define the open set:

$$\mathcal{K}(\mathcal{B}^E) = \bigcup_{u \in \mathcal{B}^E} B(u, \beta/4).$$

Let r_2 be given by Theorem 59 and define the number:

$$\mu = \min\left\{\frac{\beta}{8}, \frac{r_2}{3}\right\}.$$

By Proposition 64, if $u \in \mathcal{K}(\mathcal{B}^E)$, and $t < \min\{\epsilon', \frac{\mu}{n}\frac{1}{\eta}\}$, then:

$$\|u - \psi_t(u)\| \leq \frac{\mu}{n}. \tag{4.20}$$

Moreover by the equicontinuity (at zero) of the family of functions $\{\sigma_{u_0}\}$ from Theorem 59, there exists c such that if $\|w - \varphi(u_0)\| + \|u - u_0\| \leq c$ then:

$$\|u + \sigma_{u_0}(w, u) - u_0\| \leq \frac{\mu}{n}. \tag{4.21}$$

We define accordingly:

$$\delta = \delta(n) = \min\left\{\frac{\epsilon'}{2}, \frac{\mu}{2n}\frac{1}{\eta}, \frac{c}{2}\right\}.$$

Consider then a map $h_0 : D^n \to \mathcal{U}(y)_{I(s)}$ lifting the homotopy $\tilde{h}_t : D^n \to I(s)$ at time $t = 0$. Endow D^n with a CW-complex structure such that each cell is either entirely contained in $\mathcal{K}(\mathcal{B}^E)$ or entirely contained in $h_0^{-1}(B_{\frac{\mu}{n}}(h(x)))$ for some $x \in D^n$.

We lift the homotopy inductively on the skeleta of D^n, starting from the zero skeleton. If $x \in D^n$ is a point in the zero skeleton such that $h(x) \in \mathcal{K}(\mathcal{B}^E)$, then we define the homotopy $h_t|_{\{x\}}$ by:

$$h_t(x) = \psi_{\tilde{h}_0(x)-\tilde{h}_t(x)}(h_0(x)),$$

in such a way that $J(h_t(x)) = J(h_0(x)) - \tilde{h}_0(x) + \tilde{h}_t(x) = \tilde{h}_t(x)$.

If otherwise $x \notin \mathcal{K}(\mathcal{B}^E)$, then $x \in B_{\frac{\mu}{n}}(u_0(x))$ for some $u_0 = u_0(x)$. We consider the corresponding function $\sigma = \sigma_{u_0(x)} : B(0, r_1) \times B(u_0(x), r_2) \to L^2$ given by Theorem 59. Then $h_t|_{\{x\}}$ is defined as:

$$h_t(x) = h_0(x) + \sigma_{u_0(x)}(\underbrace{(\tilde{h}_t(x), y), h_0(x)}_{\in \mathbb{R}\times\mathbb{R}^m\times L^2})$$

which lifts $\tilde{h}_t(x)$ by Theorem 59.

Notice that, because of (4.20) and (4.21), during the homotopy the point $h_t(x)$ has been moved from its original location at most at a distance:

$$\|h_0(x) - h_t(x)\| \leq \frac{\mu}{n}. \tag{4.22}$$

Assume now that the homotopy has been lifted to the $(k-1)$-skeleton of D^n (with the CW-complex structure defined before). Composing with the characteristic map $\phi : D^k \to D^n$ of a cell, reduces to the case when we have to extend to the whole disk D^k a homotopy which has been defined already on ∂D^k.

If $h_0(D^k)$ is entirely contained in $\mathcal{K}(\mathcal{B}^E)$, we simply define the homotopy using the flow as above:

$$h_t(x) = \psi_{\tilde{h}_0(x)-\tilde{h}_t(x)}(h_0(x)), \quad x \in D^k.$$

This homotopy glues on the boundary ∂D^k, which by (4.22) is also entirely contained in $\mathcal{K}(\mathcal{B}^E)$ and for which the homotopy was defined by the flow.

Otherwise there exists u_0 such that $h_0(D^k) \subset B_{\frac{\mu}{n}}(u_0)$. We need to extend to the all disk a homotopy that has been defined already on ∂D^k; notice that since at each previous inductive step the homotopy moved the points from their original location at a distance at most $\frac{\mu}{n}$, then:

$$h_t(\partial D^k) \subset B_{\frac{k\mu}{n}}(u_0) \subset B_{r_2}(u_0) \quad \forall t \in I.$$

In other words, denoting by $H : D^n \times I \to \mathcal{U}(y)$ the partially defined homotopy, we have:

$$H(D^k \times I, D^k \times \{0\} \cup \partial D^k \times I) \subset B_{r_2}(u_0).$$

The pairs $(D^k \times I, D^k \times \{0\} \cup \partial D^k \times I)$ and $(D^k \times I, D^k \times \{0\})$ are homeomorphic and the extension problem is equivalent to just the homotopy lifting property for a map from the k-dimensional disk to $I(s)$ with a time-zero lift all in $B_{r_2}(u_0)$. For such a map $h_0 : D^k \to B_{r_2}(u_0)$ we define again the lifting homotopy as:

$$h_t(x) = h_0 + \sigma_{u_0(x)}((\tilde{h}_t(x), y), h_0(x)).$$

This concludes the proof. □

6.1. The Deformation Lemma.

THEOREM 68 (sub-Riemannian Deformation Lemma). *Assume that all singular curves with $J \leq E_2$ are soft and that there are no normal geodesics in $\Omega(y)$ with Energy between E_1 and E_2. Then for every $\epsilon > 0$, every compact manifold X and any continuous map $h : X \to \Omega(y)^{E_2}$ there exists a homotopy $h_t : X \to \Omega(y)^{E_2}$ such that $h_0 = h$ and $h_1(X) \subset \Omega(y)^{E_1+\epsilon}$.*

PROOF. Let us first show how to reduce the proof for horizontal curves to the proof of the exact same statement for the global chart. Notice first that, as we have shown in Section 2.8, if a singular curve γ is soft, then the same is true for any control whose associated trajectory is γ; moreover if u is a singular control with $J(u) \leq E_2$ which is not soft, then the corresponding trajectory γ_u is also singular, it is not soft and its energy satisfies $J(\gamma_u) \leq E_2$. Also if u is a normal control, then $A(u)$ is a normal geodesic. This can be seen as follows: being locally length minimizing, $A(u)$ can be either the projection of a normal or an abnormal extremal. If it were the projection of an abnormal extremal, it would be a singular curve, hence of corank one, and strictly abnormal, contradicting the existence of a normal extremal lift for u. In particular the hypothesis at the level of horizontal curves imply the same hypothesis for the set of controls.

Given h, we can consider the function:

$$\overline{h} : X \to \mathcal{U}, \quad \overline{h}(\theta) = \mu(h(\theta))$$

(recall that $\mu : \Omega \to \mathcal{U}$ denotes the minimal control). Then, since $J(\mu(\gamma)) = J(\gamma)$, we have $\overline{h}(X) \subset \mathcal{U}(y)^{E_2}$.

If we can find a homotopy $\overline{h}_t : X \to \mathcal{U}(y)^{E_2}$ with the property that $\overline{h}_1(X) \subset \mathcal{U}(y)^{E_1+\epsilon}$, then the function:

$$h_t = A \circ \overline{h}_t : X \to \Omega(y)$$

defines the desired homotopy. In fact $h_0(\theta) = A(\overline{h}_0(\theta)) = A(\mu(h(\theta))) = h(\theta)$ for every $\theta \in X$; moreover since $J(A(\overline{h}_t(\theta))) \leq J(\overline{h}_t(\theta))$ then we also have $h_t(X) \subset \Omega(y)^{E_2}$ and $h_1(X) \subset \Omega(y)^{E_1+\epsilon}$.

This reduces to prove that the same statement holds true if $\Omega(y)$ is replaced with $\mathcal{U}(y)$. Consider the two open sets of $V_1 = h^{-1}(\{J < E_1 + \epsilon\})$ and $V_2 = h^{-1}(\{J > E_1 + \epsilon/2\})$. Let $\{\rho_1, \rho_2\}$ be a smooth partition of unity subordinated to the open cover $\{V_1, V_2\}$ of X. Then $\rho_2|_{h^{-1}(\{J \geq E_1+\epsilon\})} \equiv 1$ and $\rho_2|_{h^{-1}(\{J \leq E_1+\epsilon/2\})} \equiv 0$. Let $c \in (E_1 + \epsilon/2, E_1 + \epsilon)$ be a regular value of ρ_2 and consider the smooth submanifold:

$$M = \{\rho_2 \geq c\} \quad \text{with} \quad \partial M = \{\rho_2 = c\}.$$

The pair $(M, \partial M)$ is a CW-complex pair and $h(M) \subset \{J \geq c\}$. Then by Theorem 67 there exists a homotopy

$$\tilde{H} : M \times I \to \Omega(y) \cap \{c \leq J \leq E_2\}$$

such that $\tilde{H}(\cdot, 0) = h|_M$, $\tilde{h}_1(M) \subset \{J \leq \epsilon\}$ and $\tilde{H}(\cdot, t)|_{\partial M} \equiv h|_{\partial M}$. The desired homotopy $H : X \times I \to \mathcal{U}(y)^{E_2}$ is defined by:

$$H(x,t) = \begin{cases} h(x) & \overline{x} \in X \backslash M \\ \tilde{H}(x,t) & x \in M \end{cases}$$

□

7. Applications

7.1. A sub-Riemannian Minimax Principle. In this section we prove a sub-Riemannian version of the classical Minimax principle for variational problems. If X is a compact manifold and two continuous maps $f, g : X \to \Omega(y)$ are homotopic we will write $f \sim g$.

THEOREM 69 (sub-Riemannian Minimax Principle). *Let X be a compact manifold and $f : X \to \Omega(y)$ be a continuous map which is not homotopic to a constant map. Consider:*

$$c = \inf_{g\sim f} \sup_{\theta\in X} J(g(\theta)).$$

Assume that there exists $\delta > 0$ such that all singular curves with energy $J \leq c + \delta$ are soft (a generic condition if $d \geq 3$). Then for every $\delta > \epsilon > 0$ there exists a normal geodesic $\gamma_\epsilon \in \Omega(y)$ such that:

$$c - \epsilon \leq J(\gamma_\epsilon) \leq c + \epsilon.$$

PROOF. First note that $c > 0$. In fact, if $x \neq y$, then $c \geq d(x, y) > 0$; if $x = y$ this follows from Proposition 70 below.

Assume that the claim of the theorem is false. Then there exists $\delta > \epsilon > 0$ such that any curve in $\Omega(y)^{c+\epsilon}_{c-\epsilon}$ is either regular or a soft abnormal. Then let $g \sim f$ such that $\sup_{\theta\in X} J(g(\theta)) \leq c + \epsilon$. By Theorem 68 the map g is homotopic to a map $g' : X \to \Omega(y)^{c-\epsilon/2}$, which contradicts the definition of c. □

Note that in the case $x = y$ the following proposition requires no assumption on the type of singular curves..

PROPOSITION 70. *Assume $x = y$. Then there exists $c > 0$ such that any map $f : X \to \Omega(x)$ satisfying* $\sup_{\theta\in X} J(f(\theta)) \leq c$ *is homotopic to a constant map.*

PROOF. By [**BL**, Corollary 7] the space $\Omega(x)$ has the homotopy type of a CW-complex, and in particular any point in it has a contractible neighborhood. Consider the constant curve $\gamma(y) \equiv x$, and a neighborhood $U_\gamma \subset \Omega(x)$ which is contractible in $\Omega(x)$. Since the family $\{J < t\}_{t\in\mathbb{R}}$ is a local basis for $\Omega(x)$ at the constant curve γ, then there exists ϵ such that $\{J \leq \epsilon\} \subset U_\gamma$. As a consequence, if $\mathrm{im}(f) \subset \{J \leq c\}$, then f is homotopic to a constant map. □

7.2. Serre's Theorem and Another Deformation Lemma. We finish this section with a proof of a sub-Riemannian version of Serre's Theorem, providing the existence of infinitely many geodesics between any two points x and y on a compact sub-Riemannian manifold. The case when y is a regular value for the endpoint map centered at x and the case of a contact manifolds are proved in [**BL**]. We will need the following variation of the deformation lemma.

LEMMA 71. *Assume all singular curves with $J \leq E$ are soft. Let $0 < s < E$ and $\mathcal{M}$ be a neighborhood of $\mathcal{C} \cap \{J = s\}$. Then for every $n \in \mathbb{N}$ there exists $\epsilon = \epsilon(n)$ such that for every n-dimensional simplicial complex X and any continuous map:*

$$f : X \to \Omega(y)^{s+\epsilon}\backslash\mathcal{M}$$

there exists a homotopy of maps $f_t : X \to \Omega(y)^{s+\epsilon}\backslash\mathcal{M}$, $t \in [0, 1]$, such that $f_0 = f$ and:

$$f_1 : X \to \Omega^{s-\epsilon}.$$

PROOF. First, arguing as in the proof of Theorem 68, we reduce to prove the statement for $\mathcal{U}$ instead of Ω, thus working with $\mathcal{N} = A^{-1}(\mathcal{M})$ instead of $\mathcal{M}$.

We claim that the statement follows from the following fact (whose proof we postpone): there exists $\epsilon > 0$ such that for every n-dimensional simplicial pair (Y, Z), $Z \subset Y$ and any continuous map $f : Y \to (\mathcal{U}(y)\backslash\mathcal{M}) \cap \{s-2\epsilon \le J \le s+\epsilon\}$ such that $f(Z) \subset \{J \le s-\epsilon\}$ we can find a homotopy $f_t : Y \to (\mathcal{U}(y)\backslash\mathcal{M}) \cap \{s-2\epsilon \le J \le s+\epsilon\}$ with $f_1(Y) \subset \{J \le s-\epsilon\}$ and $f_t|_Z \equiv f$. To see that this implies the statement consider $c \in (s-2\epsilon, s-\epsilon)$ and the two open sets $V_1 = f^{-1}(\{J < s-\epsilon\})$ and $V_2 = f^{-1}(\{J > c\})$. Since X is a simplicial complex, there exist subcomplexes $Y_1, Y_2 \subset X$ such that $X = Y_1 \cup Y_2$ and $Y_1 \subset V_1$, $Y_2 \subset V_2$. Applying the claim to the map $f|_{Y_2}$ and the pair $(Y_2, Y_2 \cap Y_1)$ provides a homotopy $\tilde{f}_t : Y_2 \to (\mathcal{U}(y)\backslash\mathcal{M}) \cap \{s-2\epsilon \le J \le s+\epsilon\}$ which is stationary on $Y_1 \cap Y_2$ and which consequently glues to a homotopy $f_t : X \to (\mathcal{U}(y)\backslash\mathcal{M}) \cap \{s-2\epsilon \le J \le s+\epsilon\}$ which is just f on Y_2 and which satisfies the needed requirements.

It remains to prove the claim. Arguing as in Lemma 66, we see that instead of proving it for any simplicial pair, it is enough to do it for a map from the disk D^n:

$$f : D^n \to (\mathcal{U}(y)\backslash\mathcal{M}) \cap \{s-2\epsilon \le J \le s+\epsilon\}$$

(ϵ will be chosen below).

First we use Corollary 63, which provides us with $0 < \eta' < \eta$ such that for every $\epsilon' < \epsilon'' < \eta' < \eta$ we have a deforming map:

$$\psi_t : \mathcal{W}(\mathcal{B}^E) \to \mathcal{U}(y)\backslash\mathcal{V}(\mathcal{A}^E)$$

with the properties: $\psi_0 \equiv \text{id}$, $\psi_t|_{\{J \le s-\epsilon''\}\cup\{J \ge s+\epsilon''\}} \equiv \text{id}$ and:

$$\psi_{\bar{t}}\left((\mathcal{W}(\mathcal{B}^E)\backslash\mathcal{N}) \cap \{J \le s+\epsilon'\}\right) \subset \mathcal{U}^E\backslash\mathcal{V}(\mathcal{A}^E) \cap \{J \le s-\epsilon'\}$$

(here $\bar{t} = \eta\beta/2$). We define $\tilde{\psi}_t = \psi_{t/\bar{t}}$, $t \in [0,1]$ so that we have completely deformed at time $t = 1$; moreover (using the notation of Proposition 62) if $t < \frac{\mu\bar{t}\eta}{n}$ we also have:

$$\|\psi_t(u) - u\| \le \int_0^{\frac{t}{\bar{t}}} \|Y(\psi_\tau^Y(u))\| d\tau \le \frac{t}{\bar{t}\eta} \le \frac{\mu}{n}.$$

As before, by the equicontinuity (at zero) of the family of functions $\{\sigma_{u_0}\}$ from Theorem 59, there exists c such that if $\|w - \varphi(u_0)\| + \|u - u_0\| \le c$ then:

$$\|u + \sigma_{u_0}(w,u) - u_0\| \le \frac{\mu}{n}.$$

We define accordingly:

$$\epsilon = \epsilon(n) = \min\left\{\frac{\epsilon'}{2}, \frac{\mu}{2n}\frac{1}{\eta}, \frac{c}{2}, \frac{r_1}{2}\right\}.$$

The proof now proceeds exactly as the proof of Theorem 67, with the choice $h_0 = f$ and $\tilde{h}_t$ defined by:

$$\tilde{h}_t(z) = J(f(z))(1-t) + t(s-\epsilon), \quad z \in D^n.$$

□

COROLLARY 72 (Serre's Theorem). *Let x, y be any two point on a* compact *sub-Riemannian manifold whose singular curves are all soft. Then there are infinitely many normal geodesics joining x and y.*

PROOF. The scheme of the proof is identical to the classical proof (see for example [**Cha93**, Section 3.2]), except that here we have to work with the (possibly singular) space $\Omega(y)$ of horizontal curves joining x and y and with the functional $J : \Omega(y) \to \mathbb{R}$. The main difference is to replace the classical deformation lemma with the one above. We only give a sketch, leaving the details to the reader.

In the case the fundamental group of M is infinite, it is enough to apply a Minimax procedure on each homotopy class (as it is done in [**LM**]); otherwise one passes to the universal cover, which is locally isometric to the manifold itself, and proves the statement for the case M is compact and simply connected.

The homotopy of $\Omega(y)$ is the same as the homotopy of the ordinary path-space (in fact the inclusion is a homotopy equivalence, see [**BL**]), and under the assumption that M is compact and simply connected the cup-length of $\Omega(y)$ is infinite [**Ser51**]. The theorem follows then by applying the same argument as in the proof of [**Cha93**, Lemma 3.1], with the following modification. Using the notation of [**Cha93**, Lemma 3.1, pag. 106], the simplicial set $|\tau_1|$ is deformed below the level $c_2 - \epsilon$ using the previous Lemma. □

Bibliography

[AB12] A. A. Agrachev and D. Barilari. Sub-Riemannian Structures on 3D Lie Groups. *J. Dynam Control Systems*, 18:21–44, 2012.

[ABB] Andrei Agrachev, Davide Barilari, and Ugo Boascain. *Introduction to Riemannian and subRiemannian Geometry.* notes available at https://webusers.imj-prg.fr/ davide.barilari/SRnotes-v17-04-2016.pdf.

[ABB12] A. A. Agrachev, D. Barilari, and U. Boscain. On the Hausdorff Volume in sub-Riemannian Geometry. *Calc. Var. and PDE*, 43:355–388, 2012.

[AGL15] Andrei Agrachev, Antonio Lerario, and Alessandro Gentile. Geodesics and Horizontal-Paths Spaces in Carnot Groups. *Geom. Topol.*, 19(3):1569–1630, 2015.

[AL10] Andrei Agrachev and Paul Lee. Continuity of Optimal Control Cost and Its Application to Weak KAM Theory. *Calc. Var. Partial Differential Equations*, 39(1-2):213–232, 2010.

[AP93] Antonio Ambrosetti and Giovanni Prodi. *A Primer in Nonlinear Analysis.* Cambridge University Press, 1993.

[AS04] A. A. Agrachev and Y. Sachkov. *Control Theory from the Geometric Viewpoint*, volume 87. Springer - Verlag, Berlin, 2004.

[Ag96] A. A. Agrachev. Exponential Mappings for Contact sub-Riemannian Structures. *J. Dynam Control Systems*, 238:321–358, 1996.

[Ar89] V. I. Arnol'd. *Mathematical Methods of Classical Mechanics*, volume 60. Springer - Verlag, New York, 1989.

[Arn72] James Arnold. Local to Global Theorems in the Theory of Hurewicz Fibrations. *Trans. Amer. Math. Soc.*, 164:179–188, 1972.

[BH93] Robert Bryant and Lucas Hsu. Rigidity of Integral Curves of Rank 2 distributions. *Invent. Math.*, 114(2):435–461, 1993.

[BL] Antonio Lerario and Francesco Boarotto. Homotopy Properties of Horizontal Path Spaces and a Theorem of Serre in subRiemannian Geometry. *To appear on Communication in Analysis and Geometry.*

[BR13] D. Barilari and L. Rizzi. A Formula for Popp's Volume in sub-Riemannian Geometry. *Analysis and Geometry in Metric Spaces*, 123:42–57, 2013.

[BT82] Raoul Bott and Loring Tu. *Differential Forms in Algebraic Topology.* Springer - Verlag, 1982.

[CJT06] Yacine Chitour, Emmanuel Trélat, and Frédéric Jean. Genericity Results for Singular Curves. *J. Differential Geom.*, 73(1):45–73, 2006.

[CJT08] Yacine Chitour, Emmanuel Trélat, and Frédéric Jean. Singular Trajectories of Control-Affine Systems. *SIAM J. Control Optim.*, 47(2):1078–1095, 2008.

[CM08] A. L. Castro and R. Montgomery. The Chains of left-invariant Cauchy-Riemann structures on $SU(2)$. *Pacific J. Math*, 238:41–71, 2008.

[Car05] Neal Carothers. *A Short Course on Banach Space Theory.* Cambridge University Press, 2005.

[Cha93] Kung Chan. *Infinite Dimensional Morse Theory and Multiple Solution Problems.* Birkäuser, 1993.

[Cla76] Francis Clarke. On the Inverse Function Theorem. *Pacific J. Math.*, 64(1):97–102, 1976.

[Cla83] Francis Clarke. *Optimization and Nonsmooth Analysis.* John Wiley & Sons, New York, 1983.

[DR12] Jason Dominy and Hershel Rabitz. Dynamic Homotopy and Landscape Dynamical Set Topology in Quantum Control. *J. Math. Physics*, 53(8):17 pp., 2012.

[FA86] F. A. Farris. An Intrinsic Construction of Fefferman's CR Metric. *Pacific J. Math*, 123:33–45, 1986.

[FG12] C. Fefferman and C. R. Graham. *The Ambient Metric*, volume 178.

[FGV95] E. Falbel, C. Gorodski, and J. M. Veloso. Conformal sub-Riemannian Geometry in Dimension 3. *Mat. Contemp.*, 9:61–73, 1995.

[FMP99] C. B. Figueroa, F. Mercuri, and R. H. Pedrosa. Invariant Surfaces of the Heisenberg Groups. *Ann.*, Ann. Mat. Pura Appl.:173–194, 1999.

[FP90] Rudolph Fritsch and Renzo Piccinini. *Cellular Structures in Topology.* Number 19. 1990.

[Hat02] Allen Hatcher. *Algebraic Topology.* Cambridge University Press, 2002.

[Hur55] Witold Hurewicz. On the concept of fiber space. *Proc. Nat. Acad. Sci. U. S. A.*, 41:956–961, 1955.

[Jac62] N. Jacobson. *Lie Algebras*. Number 10. Interscience Publisher, 1962.

[Jea14] Frédéric Jean. *Control of nonholonomic systems: from sub-Riemannian geometry to motion planning*. Springer, Cham, 2014.

[Kir08] A. A. Kirillov. *An Introduction to Lie Groups and Lie Algebras*. Cambridge University Press, 2008.

[LM] Antonio Lerario and Andrea Mondino. Homotopy Properties of Horizontal Loop Spaces. *https://arxiv.org/pdf/1509.07000v1.pdf*.

[LR] Antonio Lerario and Luca Rizzi. How Many Geodesics join Two Points on a subRiemannian Manifold? *To appear on Journal of Symplectic Geometry*.

[Lee86] J. M. Lee. The Fefferman Metric and Pseudo-Hermitian Invariants. *Trans. Amer. Math. Soc*, 296:411–429, 1986.

[Mil59] John Milnor. On spaces having the homotopy type of a CW-complex. *Trans. Amer. Math. Soc.*, 90:272–280, 1959.

[Mil73] John Milnor. *Morse Theory*. Princeton University Press, 1973.

[Mit85] J. Mitchell. On Carnot-Carathéodory Metrics. *J. Differential Geom.*, 21:35–45, 1985.

[Mon02] Richard Montgomery. *A Tour of SubRiemannian Geometries, Their Geodesics and Applications*, volume 91 of *AMS Mathematical Surveys and Monographs*. 2002.

[PSWZ76] J. Patera, R. T. Sharp, P. Winternitz, and H. Zassenhaus. Invariants of Real Low Dimension Lie Algebras. *Journal of Math. Physics*, 178:986–994, 1976.

[Ras38] P. K. Rashevski. Any Two Points of a Totally Nonholonomic Space May Be Connected By an Admissible Line . *Uch. Zap. Ped Inst. im. Liebknechta*, 21:83–94, 1938.

[Sar91] Andrey Sarychev. Homotopy Properties of The Space of Trajectories of a Completely Nonholonomic Differential System. *Soviet Math. Dokl.*, 42(2):674–678, 1991.

[Sch64] Jacob Schwartz. Generalizing the Lusternik-Schnirelmann Theory of Critical Points. *Comm. Pure Appl. Math.*, 17:307–315, 1964.

[Ser51] Jean-Pierre Serre. Homologie Singulière des Espaces Fibrés. *Ann. of Math.*, 54(2):425–505, 1951.

[Sma58] Stephen Smale. Regular Curves on Riemannian Manifolds. *Trans. Amer. Math. Soc.*, 87:492–512, 1958.

[Spa94] Edwin Spanier. *Algebraic Topology*. Springer Science and Business Media, 1994.

[Tan79] N. Tanaka. On the Equivalence Problems Associated with Simple Lie Algebras. *Hokkaido Math. J.*, 8:23–84, 1979.

[Tre00] Emmanuel Trélat. Some Properties of the Value Function and its Level Sets for Affine Control Systems with Quadratic Cost. *J. Dynam. Control Systems*, 6(4):511–541, 2000.

[Zel09] I. Zelenko. On Tanaka's Prolongation Procedure for Filtered Structures of Constant Type. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 5(94):21 pages, 2009.