Topological **M**ethods in **N**onlinear **A**nalysis
Journal of the Juliusz Schauder Center
Volume 21, 2003, 273–291

MORSE THEORY FOR NORMAL GEODESICS IN SUB-RIEMANNIAN MANIFOLDS WITH CODIMENSION ONE DISTRIBUTIONS

R. Giambò — F. Giannoni — P. Piccione — D. V. Tausk

Abstract. We consider a Riemannian manifold $(\mathcal{M}, g)$ and a codimension one distribution $\Delta \subset T\mathcal{M}$ on $\mathcal{M}$ which is the orthogonal of a unit vector field Y on $\mathcal{M}$. We do not make any nonintegrability assumption on Δ. The aim of the paper is to develop a Morse Theory for the sub-Riemannian action functional E on the space of horizontal curves, i.e. everywhere tangent to the distribution Δ. We consider the case of horizontal curves joining a smooth submanifold $\mathcal{P}$ of $\mathcal{M}$ and a fixed point $q \in \mathcal{M}$. Under the assumption that $\mathcal{P}$ is transversal to Δ, it is known (see [19]) that the set of such curves has the structure of an infinite dimensional Hilbert manifold and that the critical points of E are the so called *normal extremals* (see [10]). We compute the second variation of E at its critical points, we define the notions of $\mathcal{P}$-Jacobi field, of $\mathcal{P}$-focal point and of exponential map and we prove a Morse Index Theorem. Finally, we prove the Morse relations for the critical points of E under the assumption of completeness for $(\mathcal{M}, g)$.

1. Introduction

Sub-Riemannian geometry is the geometry of of manifolds endowed with a partially defined positive definite metric tensor. More precisely, a sub-Riemannian manifold consists of a triple $(\mathcal{M}, \Delta, g)$, where $\mathcal{M}$ is a smooth manifold,

2000 *Mathematics Subject Classification.* 53C17, 53C22, 58E05, 58E10, 58E25.

Key words and phrases. Morse theory, sub-Riemannian geometry, nomal sub-Remannian minimizers.

During the writing of the paper, the third author was partially sponsored by CNPq, (Processo 301410/95 and 200615/01-7). The fourth author was sponsored by FAPESP, São Paulo, Brazil (Processo n. 98/12530-2).

$\Delta \subset TM$ is a (non necessarily integrable) smooth distribution in $\mathcal{M}$, and g is a smoothly varying positive definite inner product in Δ. Such partially defined metric tensors are also called *Carnot–Carathéodory metrics* in the literature (see [7], [8]). The interest in the study of sub-Riemannian geometry comes essentially from Control Theory, from the study of mechanics of systems subject to (non holonomic) linear constraint, and also from other applications of Critical Point Theory when one considers solutions of constrained variational problems (see [5], [6] for an example in general relativity).

Many aspects of both the local and the global geometry of sub-Riemannian manifolds are drastically different from the case of Riemannian geomery; for instance, the exponential mapping is never a local diffeomorphism on a neighbourhood of the point at which it is based, the space of paths tangent to the distribution and joining two fixed points may contain singularities which may happen to be minimizing geodesics, the Hausdorff dimension of the induced metric space structure is always strictly bigger than the manifold dimension, etc.

There is nowadays a quite extensive literature on sub-Riemannian geometry (see for instance [1], [10], [12], [14] and the references therein); in this paper we continue the development of a variational theory for curves that are local minimizers of the sub-Riemannian length functional started in [3]. More precisely, the aim of this paper is to develop the basis of a Morse theory for sub-Riemannian *normal* geodesics; recall that the normal geodesics are the curves that are regular points of the set of horizontal curves and that are stationary for the sub-Riemannian energy functional. Under strong non integrability assumptions for the distribution Δ, for instance when Δ is *strongly bracket generating*, it is well known that every sub-Riemannian geodesic is normal. In this context, a first version of the Morse index theorem was proven by Kishimoto in [8].

The aim of this paper is to develop an infinite dimensional Morse theory for sub-Riemannian geodesics. More specifically, we will consider geodesics in a sub-Riemannian manifold $(\mathcal{M}, \Delta, g)$, where Δ is a codimension one transversally oriented distribution in $\mathcal{M}$, starting at a given submanifold $\mathcal{P}$ of $\mathcal{M}$ which is transversal to Δ. By this, we mean that the sub-Riemannian geodesics in consideration are obtained as stationary points of the sub-Riemannian action functional in the space of curves whose initial edpoint is left free to move on $\mathcal{P}$ and whose final endpoint is fixed. Such geodesics are called *$\mathcal{P}$-normal geodesics* in this paper. In this context, we will do the following:

- study the second variation of the sub-Riemannian action functional at a given geodesic,
- define a suitable notion of exponential map, Jacobi fields and focal points along a geodesic relativey to the initial submanifold $\mathcal{P}$,
- prove a Morse index theorem,

- prove the Morse relations for sub-Riemannian geodesics.

Observe that no assumption is made on the non integrability of Δ. Also, it will be quite clear that the codimension one assumption makes our computation much easier, although it is certainly non crucial in the theory presented; virtually all the results of the paper can be proven in the case of distribution of arbitrary codimension.

The crucial assumption in our setup is that we allow variations of geodesics whose initial endpoint is *not* fixed, but left free to move on the submanifold $\mathcal{P}$. When $\mathcal{P}$ is transversal to Δ, this assumption implies that the space of trial paths for our variational problem does not contain singularity (see [19]), that the exponential map defined by the $\mathcal{P}$-normal geodesics is indeed a local diffeomorphism (see Section 3), and that many of the classical results in Riemannian geometry can be extended to this situation. It is interesting to observe that, despite the analogies with the Riemannian case, in our context sub-Riemannian geodesics may admit a continuum of focal points; the set of focal points consists of a finite union of isolated points and of segments of a $\mathcal{P}$-normal geodesic (Proposition 4.4). However, we prove that sufficiently small portions of a $\mathcal{P}$-normal geodesic does not contain focal points, and as an application of the Morse index theorem we prove the minimality of sufficiently small initial portions of $\mathcal{P}$-normal geodesics (Proposition 5.2).

The discreteness of the set of focal points holds under *generic* circumstances, for instance when the data are real-analytic (see Proposition 4.4). Our formulation of the Morse Index Theorem (Section 4) holds only when one assumes finiteness of the set of focal points. A more general version of the theorem involving the so-called *Maslov index* (rather than the number of focal points) of the corresponding solution of the Hamiltonian geodesic (see [18], [17] for some index theorems involving the Maslov index); the statement of this generalization is quite involved (see [20], [21]) and it will not be discussed in this paper.

Finally, in Section 6 we use techniques of Critical Point Theory to prove the Morse relations for $\mathcal{P}$-normal geodesics under a suitable completeness assumption. These relations give an estimate on the number of $\mathcal{P}$-normal geodesics from $\mathcal{P}$ and a fixed point $q \in \mathcal{M}$ in terms of some topological invariants of the space of horizontal curves from $\mathcal{P}$ to q.

2. The second variation of the sub-Riemannian action functional

We will consider throughout the article the following setup: $(\mathcal{M}, g)$ is a Riemannian manifold with $\dim(\mathcal{M}) = n$, $\Delta \subset T\mathcal{M}$ is a smooth distribution of rank $n-1$ on $\mathcal{M}$. We assume that Δ is *transversally oriented*, i.e. Δ is the orthogonal distribution to a nonvanishing smooth vector field Y on $\mathcal{M}$; we can clearly assume $g(Y, Y) \equiv 1$. Let $\mathcal{P}$ be a smooth submanifold of $\mathcal{M}$ which is everywhere *transversal* to Δ, i.e. $T_p\mathcal{P} + \Delta_p = T_p\mathcal{M}$ for all $p \in \mathcal{P}$; we set $d = \dim(\mathcal{P})$. Let $q \in \mathcal{M}$ be a fixed point; we denote by $\Omega_{\mathcal{P},q}$ the set of curves $x\colon [0,1] \to \mathcal{M}$ of Sobolev class H^1 with $x(0) \in \mathcal{P}$ and $x(1) = q$. Recall that a curve is of class H^1 if it is absolutely continuous and if its derivative $\dot{x}$ is square integrable; we refer to [2] for the basics on Sobolev spaces. It is well known that $\Omega_{\mathcal{P},q}$ has the structure of an infinite dimensional Hilbert manifold and that, for $x \in \Omega_{\mathcal{P},q}$, the tangent space $T_x\Omega_{\mathcal{P},q}$ can be identified with the space of vector fields V along x of class H^1 with $V(0) \in T_{x(0)}\mathcal{P}$ and $V(1) = 0$.

We denote by ∇ the covariant derivative of the Levi–Civita connection of g and by R its curvature tensor, chosen with the following sign convention: $R(X,Z) = \nabla_X\nabla_Z - \nabla_Z\nabla_X - \nabla_{[X,Z]}$. When there is no danger of confusion, we will denote with a dot the derivative of curves in $\mathcal{M}$ and with a prime the covariant derivative of vector fields along curves. A possible exception to this convention will be made in the case of the covariant derivative of the tangent field along a curve z: rather than the awkward notation $\dot{z}'$ we will use the double prime notation z''.

Given a smooth vector field W on $\mathcal{M}$, we denote by $(\nabla W)^*$ the g-transpose of the covariant derivative of W, which is the $(1,1)$ tensor field on $\mathcal{M}$ whose value at a point $p \in \mathcal{M}$ is the linear map on $T_p\mathcal{M}$ defined by:

$$g((\nabla W)^* v_1, v_2) = g(\nabla_{v_2} W, v_1) \quad \text{for all } v_1, v_2 \in T_p\mathcal{M}. \tag{2.1}$$

For all $p \in \mathcal{P}$ and all $n \in T_p\mathcal{P}^\perp$, let $\mathcal{S}^{\mathcal{P}}_n$ be the second fundamental form of $\mathcal{P}$ in the orthogonal direction n, which is the symmetric bilinear form on $T_p\mathcal{P}$ given by $\mathcal{S}^{\mathcal{P}}_n(v_1,v_2) = g(n, \nabla_{v_1}V_2)$, where V_2 is any extension of v_2 to a vector field tangent to $\mathcal{P}$. We will also look at the second fundamental form as the linear map $\mathcal{S}^{\mathcal{P}}_n\colon T_p\mathcal{P} \mapsto T_p\mathcal{P}$ such that $g(\mathcal{S}^{\mathcal{P}}_n(v_1), v_2) = \mathcal{S}^{\mathcal{P}}_n(v_1,v_2)$ and we observe that, if N is a normal field along $\mathcal{P}$ and $v \in T_p\mathcal{P}$, then $\nabla_v N + \mathcal{S}^{\mathcal{P}}_{N(p)}(v)$ is in $T_p\mathcal{P}^\perp$ (this will be used in the proof of Lemma 2.5).

By $\Omega_{\mathcal{P},q}(\Delta)$ we will mean the subset of $\Omega_{\mathcal{P},q}$ consisting of *horizontal* curves, i.e. those curves x for which $\dot{x} \in \Delta$ almost everywhere. It is proven in [19] that, since $\mathcal{P}$ is transversal to Δ, $\Omega_{\mathcal{P},q}(\Delta)$ is a smooth submanifold of $\Omega_{\mathcal{P},q}$. More precisely, it is shown in [19] that 0 is a regular value for the map $F\colon \Omega_{\mathcal{P},q} \to L^2([0,1],\mathbb{R})$ given by $F(x) = g(\dot{x}, Y)$; the tangent space $T_x\Omega_{\mathcal{P},q}(\Delta)$ is given by the kernel of $dF(x)$:

$$T_x\Omega_{\mathcal{P},q}(\Delta) = \{V \in T_x\Omega_{\mathcal{P},q} : g(V', Y) + g(\dot{x}, \nabla_V Y) = 0\}. \tag{2.2}$$

We will consider the following Hilbert space inner product on $T_x\Omega_{\mathcal{P},q}(\Delta)$:

$$\langle V, W \rangle = \int_0^1 g(V', W')\, dt.$$

We will consider the action functional on $\Omega_{\mathcal{P},q}$ and on $\Omega_{\mathcal{P},q}(\Delta)$, denoted by E and defined by $E(x) = (1/2)\int_0^1 g(\dot{x},\dot{x})\,dt$. It is a smooth map and its critical points in $\Omega_{\mathcal{P},q}(\Delta)$ are the so called *normal* sub-Riemannian geodesics x between $\mathcal{P}$ and q whose Hamiltonian lifts $X\colon[0,1]\to T\mathcal{M}^*$ annihilate $T_{x(0)}\mathcal{P}$ (see [19]). The critical points x of E in $\Omega_{\mathcal{P},q}(\Delta)$ can be thought as critical points in $\Omega_{\mathcal{P},q}$ subject to the constraint $F(x)=0$.

In this Section we want to study the second variation of E at its critical points in $\Omega_{\mathcal{P},q}(\Delta)$; we will use the method of Lagrange multipliers. To this aim, let $x\in\Omega_{\mathcal{P},q}(\Delta)$ be a fixed critical point of E and let $\lambda\in L^2([0,1],\mathbb{R})$ be the corresponding Lagrange multiplier, i.e. λ is the unique map such that x is a critical point in $\Omega_{\mathcal{P},q}$ of the functional $E_\lambda = E-\lambda\circ F$:

$$E_\lambda(x)=\int_0^1\left(\frac{1}{2}\,g(\dot{x},\dot{x})-\lambda\,g(\dot{x},Y)\right)dt.$$

The following result is basically contained in [3] and [19]:

Proposition 2.1. *Let $x\in\Omega_{\mathcal{P},q}(\Delta)$ be fixed. Then, x is a critical point of E in $\Omega_{\mathcal{P},q}(\Delta)$ if and only if x is of class C^2 and there exists $\lambda\in C^1([0,1],\mathbb{R})$ such that the following are satisfied:*

$$x''=(\lambda\,Y)'-\lambda(\nabla Y)^*\dot{x}, \tag{2.5}$$

$$\lambda(0)\,Y(x(0))-\dot{x}(0)\in T_{x(0)}\mathcal{P}^\perp. \tag{2.6}$$

Proof. The regularity of x and λ is obtained by standard bootstrap arguments (see [3], [19]). Equations (2.5) and (2.6) are obtained as the Euler–Lagrange equations of E_λ, using integration by parts of the terms containing V' in the following formula for the first variation of E_λ:

$$dE_\lambda(x)[V]=\int_0^1\big(g(V',\dot{x})-\lambda\,g(V',Y)-\lambda\,g(\dot{x},\nabla_V Y)\big)\,dt, \tag{2.7}$$

for all $V\in T_x\Omega_{\mathcal{P},q}$. $\square$

Remark. We observe that the map λ in the statement of Proposition 2.1 is precisely the Lagrangian multiplier that appears in (2.4). Multiplying (2.5) by Y we obtain the following differential equation for λ:

$$\lambda'-g(\nabla_Y Y,\dot{x})\,\lambda+g(Y',\dot{x})=0. \tag{2.8}$$

Moreover, since $Y(x(0))\notin T_{x(0)}\mathcal{P}^\perp$, observe that (2.6) allows at most one value for $\lambda(0)$.

Definition 2.3. A horizontal curve x of class C^2 (defined in an interval containing 0) with $x(0)\in\mathcal{P}$ and that satisfies (2.5) and (2.6) for some map λ will be called a *$\mathcal{P}$-normal geodesic*.

Multiplying (2.5) by $\dot x$ we obtain $(d/dt)g(\dot x,\dot x)=0$, hence if x is a $\mathcal{P}$-normal geodesic, then $g(\dot x,\dot x)$ is constant.

By the abstract theory, the second variation of E at x in $\Omega_{\mathcal{P},q}(\Delta)$ is given by the restriction to $T_x\Omega_{\mathcal{P},q}(\Delta)$ of the second variation of E_λ at x in $\Omega_{\mathcal{P},q}$.

We recall that the *Hessian* $\mathrm{Hess}Z$ of a smooth vector field Z on $\mathcal{M}$ is the (2,1)-tensor field on $\mathcal{M}$ given by $\nabla\nabla Z$; more explicitly, it is computed as:

$$\mathrm{Hess}\,Z(v_1,v_2)=\nabla_{v_1}\nabla_{V_2}Z-\nabla_{\nabla_{v_1}V_2}Z,$$

where V_2 is any extension of v_2. The symmetric and the anti-symmetric part of $\mathrm{Hess}\,Z$ are easily computed:

$$\begin{aligned}\mathrm{Hess}_{\mathrm{a}}Z(v_1,v_2)&=\frac12\,R(v_1,v_2)Z,\\ \mathrm{Hess}_{\mathrm{s}}Z(v_1,v_2)&=\mathrm{Hess}\,Z(v_1,v_2)-\frac12\,R(v_1,v_2)\,Z.\end{aligned}\tag{2.9}$$

Given a tangent vector $v_1\in T_p\mathcal{M}$, we will consider $\mathrm{Hess}Y(v_1)$ and $\mathrm{Hess}_{\mathrm{s}}Y(v_1)$ as linear maps on the tangent space $T_p\mathcal{M}$; for the computation of the kernel of the second variation on E_λ we will need the adjoint of $\mathrm{Hess}_{\mathrm{s}}Y(v_1)$. This is easily computed from (2.9) as:

$$(\mathrm{Hess}_{\mathrm{s}}Y(v_1))^*(v_2)=(\mathrm{Hess}Y(v_1))^*(v_2)-\frac12 R(Y,v_2)v_1.\tag{2.10}$$

Proposition 2.4. *Let x be a critical point of E in $\Omega_{\mathcal{P},q}(\Delta)$, λ the associated Lagrange multiplier and let $n_x\in T_{x(0)}\mathcal{P}^\perp$ denote the vector given in* (2.6)*. Then, the second variation d^2E_x of E at x is given by the following symmetric bilinear form on $T_x\Omega_{\mathcal{P},q}(\Delta)$:*

$$\begin{aligned}d^2E_x(V,W)&\\ =&\int_0^1(g(V',W')+g(R(\dot x,V)\dot x,W))\,dt\\ &-\int_0^1(\lambda g(V',\nabla_W Y)+\lambda g(W',\nabla_V Y)+\lambda g(\mathrm{Hess}_{\mathrm{s}}Y(V,W),\dot x))\,dt\\ &-\int_0^1\Big(\frac12\lambda g(R(V,\dot x)\,W,Y)+\frac12\lambda g(R(W,\dot x)V,Y)\Big)\,dt\\ &+\mathcal{S}^{\mathcal{P}}_{n_x}(V(0),W(0)).\end{aligned}\tag{2.11}$$

Proof. The second variation $d^2E_x(V,V)$ of E in $T_x\Omega_{\mathcal{P},q}(\Delta)$ is computed as the second derivative

$$\frac{d^2}{ds^2}E_\lambda(x_s)\Big|_{s=0},$$

where $\{x_s\}_{s\in\,]-\varepsilon,\varepsilon[}$ is a variation of x in $\Omega_{\mathcal{P},q}$ with variational vector field $V\in T_x\Omega_{\mathcal{P},q}(\Delta)$, i.e. $(d/ds)x_s|_{s=0}=V$. Such computation is straightforward, it uses the equation (2.5) satisfied by x and the definitions of R and of $\mathcal{S}^{\mathcal{P}}$. Formula (2.11) is then obtained by polarization. $\square$

LEMMA 2.5. *Let $p \in \mathcal{P}$, $v, z \in T_p\mathcal{P}$ and $n_0 \in T_p\mathcal{P}^\perp$ be fixed. Then, the following two statements are equivalent:*

(a) *there exists a smooth curve $\mu\colon]-\varepsilon,\varepsilon[\to \mathcal{P}$ and a smooth vector field $n\colon]-\varepsilon,\varepsilon[\to T\mathcal{P}^\perp$ along μ with*

$$\mu(0) = p, \quad n(0) = n_0, \quad \dot\mu(0) = v, \quad n'(0) = z,$$

(b) *$v \in T_p\mathcal{P}$ and $z + \mathcal{S}^{\mathcal{P}}_{n_0}(v) \in T_p\mathcal{P}^\perp$.*

PROOF. If (a) holds, then clearly $v \in T_p\mathcal{P}$. Moreover, let X be an arbitrary vector field tangent to $\mathcal{P}$. Differentiating the expression $g(n(s), X) \equiv 0$ at $s = 0$ we get $g(n'(0), X) + g(n_0, \nabla_{\dot\mu(0)}X) = g(z + \mathcal{S}^{\mathcal{P}}_{n_0}(v), X) = 0$. Hence, (a) implies (b).

Conversely, suppose that (b) holds; let $\mu\colon]-\varepsilon,\varepsilon[\to \mathcal{P}$ be any smooth curve with $\mu(0) = p$ and $\dot\mu(0) = v$. Let $X_1, \dots, X_{n-d}$ be a local referential of $T\mathcal{P}^\perp$ around p and choose numbers $\alpha_1, \dots, \alpha_{n-d} \in \mathbb{R}$ so that $n_0 = \sum_i \alpha_i X_i(p)$. As we have observed, $\nabla_v(\sum_i \alpha_i X_i) = \mathcal{S}^{\mathcal{P}}_{n_0}(v) + v_0$, with $v_0 \in T_p\mathcal{P}^\perp$, hence:

$$z - \sum_i \alpha_i \nabla_v X_i = z - \nabla_v\Big(\sum_i \alpha_i X_i\Big) \in T_p\mathcal{P}^\perp.$$

Now, choose numbers $\beta_1, \dots, \beta_{n-d}$ in such a way that

$$z - \sum_i \alpha_i \nabla_v X_i = \sum_i \beta_i X_i(p)$$

and let $c_1, \dots, c_{n-d}$ be smooth maps on $]-\varepsilon,\varepsilon[$ such that $c_i(0) = \alpha_i$ and $c_i'(0) = \beta_i$ for all i. Finally, define $n(s) = \sum_i c_i(s) X_i(\mu(s))$, which gives the required vector field along μ. $\square$

PROPOSITION 2.6. *Let x be a critical point of E in $\Omega_{\mathcal{P},q}(\Delta)$ with Lagrange multiplier λ, let $n_x \in T_{x(0)}\mathcal{P}^\perp$ be the vector given in (2.6) and let $V \in T_x\Omega_{\mathcal{P},q}(\Delta)$. Then, V is in the kernel of d^2E_x if and only if V is of class C^2 and there exists a map $\delta \in C^1([0,1], \mathbb{R})$ such that the following linear differential equation with initial condition is satisfied:*

$$\begin{aligned} V'' = &-\lambda(\nabla Y)^*V' + (\lambda\nabla_V Y)' + R(\dot x, V)\dot x - \delta(\nabla Y)^*\dot x + (\delta\, Y)' \\ &- \lambda\,\mathrm{Hess}\,Y(V)^*\dot x - \lambda\, R(\dot x, V)Y, \end{aligned} \tag{2.12}$$

$$-V'(0) + \lambda(0)\nabla_{V(0)}Y + \mathcal{S}^{\mathcal{P}}_{n_x}(V(0)) + \delta(0)Y(x(0)) \in T_{x(0)}\mathcal{P}^\perp. \tag{2.13}$$

PROOF. It is easy to see that the elements in the kernel of d^2E_x in $T_x\Omega_{\mathcal{P},q}(\Delta)$ are the critical points of the quadratic form $d^2E_x(V,V)/2$ subject to the constraint $dF(x)[V] = 0$. Since $dF(x)$ is surjective, we can use again the method of

Lagrange multipliers. Given V in the kernel d^2E_x in $T_x\Omega_{\mathcal{P},q}(\Delta)$, we denote by δ the corresponding Lagrange multiplier in L^2. Then, V is a critical point of the map:

$$G(V) = \frac{1}{2}d^2E_x(V,V) - \int_0^1 (\delta g(V',Y) + \delta g(\dot{x}, \nabla_V Y))\,dt, \tag{2.14}$$

in $T_x\Omega_{\mathcal{P},q}$, and d^2E_x is given in (2.11). The regularity of V and δ is obtained by a bootstrap argument, and the equations (2.12), (2.13) are the Euler–Lagrange equations of the functional G. They are computed using integration by parts of the terms containing W' in the following formula for the first variation of G:

$$dG_V(W) = d^2E_x(V,W) - \int_0^1 (\delta\, g(W',Y) + \delta g(\dot{x}, \nabla_W Y))\,dt \tag{2.15}$$

To obtain the result in the final form (2.12) one needs to use the first Bianchi identity for R:

$$R(\dot{x},V)Y + R(Y,\dot{x})V + R(V,Y)\dot{x} = 0. \qquad \square$$

Remark 2.7. Multiplying (2.12) by Y, using (2.5) and the fact that $V \in T_x\Omega_{\mathcal{P},q}(\Delta)$ (see (2.2)) we obtain the following differential equation for δ:

$$\begin{aligned} \delta' = {} & g(\nabla_Y Y, \dot{x})\delta + g(\nabla_V \nabla_Y Y, \dot{x})\lambda + g(\nabla_Y Y, V')\lambda \\ & - g((\nabla_V Y)', \dot{x}) - g(R(\dot{x},V)\,\dot{x}, Y)) - g(Y', V'). \end{aligned} \tag{2.16}$$

Moreover, since $Y(x(0)) \notin T_{x(0)}\mathcal{P}^\perp$, observe that (2.13) allows at most one value for $\delta(0)$.

3. Exponential map, Jacobi fields and focal points

A $\mathcal{P}$-normal geodesic x is uniquely determined by the choice of some $p \in \mathcal{P}$ and $n \in T_p\mathcal{P}^\perp$ in the following way. Consider the unique solution (x,λ) of the system of differential equations (2.5) and (2.8) with initial conditions $\lambda(0) = g(n, Y(p))$, $x(0) = p$, $\dot{x}(0) = \lambda(0)\,Y(p) - n$. Clearly, with such a choice (2.6) is also satisfied and that $\dot{x}(0) \in \Delta_p$. By differentiating $g(\dot{x}, Y)$ and using (2.5) and (2.8), it is easy to see that x is horizontal and it is therefore a $\mathcal{P}$-normal geodesic. Conversely, any $\mathcal{P}$-normal geodesic can be obtained in this way by setting n equal to the vector in (2.6).

Remark 3.1. A straightforward calculation shows that, if x is a $\mathcal{P}$-normal geodesic with Lagrange multiplier λ, then given $c \in \mathbb{R}$, the linear reparameterization $t \mapsto x(ct)$ is again a $\mathcal{P}$-normal geodesic with Lagrange multiplier $t \mapsto c\lambda(ct)$.

We can therefore give the following:

DEFINITION 3.2. We define the map exp in a subset of $T\mathcal{P}^{\perp}$ taking values in $\mathcal{M}$ by $\exp(n) = x(1)$, where x is the unique $\mathcal{P}$-normal geodesic determined by n.

By standard results in ordinary differential equations, exp is a smooth map and its domain is an open subset of $T\mathcal{P}^{\perp}$ that contains the zero section. Moreover, it follows from Remark 3.1 that for all $n \in T\mathcal{P}^{\perp}$, the curve $t \mapsto \exp(tn)$ is a $\mathcal{P}$-normal geodesic.

DEFINITION 3.3. Given a $\mathcal{P}$-normal geodesic $x(t) = \exp(tn_0)$, a vector field V along x is a *$\mathcal{P}$-Jacobi field* along x if there exists a smooth curve $s \mapsto n(s) \in T\mathcal{P}^{\perp}$ with $n(0) = n_0$ and

$$V(t) = \frac{d}{ds}\exp(tn(s))\Big|_{s=0}. \tag{3.1}$$

PROPOSITION 3.4. *V is $\mathcal{P}$-Jacobi along the $\mathcal{P}$-normal geodesic x if and only if, for some δ of class C^1, (V, δ) satisfies the system of differential equations* (2.12)*,* (2.16) *with initial conditions* (2.13)*.*

PROOF. It is easy to see that (2.12) and (2.16) can be obtained respectively as the linearizations of (2.5) and (2.8).

If V is $\mathcal{P}$-Jacobi, then (2.12) and (2.16) are satisfied. Now, Lemma 2.5 is used to prove that (2.13) is the linearized of (2.6), and therefore (2.13) is satisfied by (V, δ).

Conversely, if (V, δ) satisfies (2.12) and (2.16), Lemma 2.5 is used to construct a curve $n(s)$ in $T\mathcal{P}^{\perp}$ so that we can define

$$W(t) = \frac{d}{ds}\exp(tn(s))\Big|_{s=0}.$$

Then, W is a $\mathcal{P}$-Jacobi field. By the first part of the proof V and W satisfy the same differential equation with initial conditions, hence $V = W$ and we are done. $\square$

Observe that, by Remark 2.7, every $\mathcal{P}$-Jacobi field V determines uniquely the map δ appearing in (2.12).

COROLLARY 3.5. *Let $x \in \Omega_{\mathcal{P},q}(\Delta)$ be a $\mathcal{P}$-normal geodesic. Then, the kernel of d^2E_x in $T_x\Omega_{\mathcal{P},q}(\Delta)$ is the set of $\mathcal{P}$-Jacobi fields V along x with $V(1) = 0$.*

PROOF. By linearizing the condition $g(\dot{x}, Y) = 0$, it is easily seen that a $\mathcal{P}$-Jacobi field V with $V(1) = 0$ is in $T_x\Omega_{\mathcal{P},q}(\Delta)$. The conclusion follows from Propositions 2.6 and 3.4. $\square$

Definition 3.6. Let x be a $\mathcal{P}$-normal geodesic. We say that $x(t)$ is a *$\mathcal{P}$-focal point along* x if there exists a non zero $\mathcal{P}$-Jacobi field V along x with $V(t) = 0$. The dimension $\mathrm{mul}(t)$ of the vector space of $\mathcal{P}$-Jacobi fields vanishing at t is called the *multiplicity* of the $\mathcal{P}$-focal point $x(t)$.

In Section 4 we will prove that a sufficiently small initial portion of a $\mathcal{P}$-normal geodesic does not contain $\mathcal{P}$-focal points. As an application of this fact, we will show in Section 5 that a sufficiently small initial portion of a $\mathcal{P}$-normal geodesic is a *global* arc length minimizer.

The set of $\mathcal{P}$-focal points along a $\mathcal{P}$-normal geodesic is characterized in Proposition 4.4.

4. The Morse Index Theorem

In this section we prove that the second variation d^2E_x of E at a $\mathcal{P}$-normal geodesic x has finite index, and that such index equals the number of $\mathcal{P}$-focal points along x counted with multiplicity, provided that the latter number is finite.

We recall that the *index* $n_-(B)$ of a symmetric bilinear form defined on the vector space $\mathfrak{V}$ is the (possibly infinite) supremum of the dimensions of the subspaces $\mathfrak{W}$ of $\mathfrak{V}$ on which B is negative definite. If $\mathfrak{V}$ is a Hilbert space with inner product $\langle\,\cdot\,,\,\cdot\,\rangle$ and B is bounded, then, by Riesz's Theorem, B is *represented* by a linear operator T_B, i.e. $\langle T_B x, y\rangle = B(x,y)$ for all $x, y \in \mathfrak{V}$. In this case, the symmetry of B implies that T_B is self-adjoint. If T_B is diagonalizable in a Hilbert basis, then $n_-(B)$ equals the number of negative eigenvalues of T_B counted with multiplicity.

Throughout this section we will assume that $x\colon [0,1] \to \mathcal{M}$ is a $\mathcal{P}$-normal geodesic, with associated Lagrange multiplier λ. We choose the following Hilbert space inner product on $T_x\Omega_{\mathcal{P},q}(\Delta)$:

$$\langle V_1, V_2\rangle = \int_0^1 g(V_1', V_2')\,dt. \tag{4.1}$$

Proposition 4.1. *The symmetric bilinear form d^2E_x is represented on the Hilbert space $T_x\Omega_{\mathcal{P},q}(\Delta)$ by a self-adjoint operator of the form $I - K$, where I is the identity and K is compact. In particular, d^2E_x has finite index on $T_x\Omega_{\mathcal{P},q}(\Delta)$.*

Proof. The first term in the first integral of (2.11) is the inner product (4.1) of $T_x\Omega_{\mathcal{P},q}(\Delta)$, and therefore it is represented by the identity operator. All the remaining terms in formula (2.11) are continuous with respect to the C^0-topology in at least one of the variables V_1, V_2. By the compact embedding of H^1 in C^0, the bilinear form corresponding to such terms is represented by a compact operator (see [4]). □

Let K be a compact self-adjoint operator on a Hilbert space H. By the spectral theorem, we can write the set of positive eigenvalues of K as a (possibly finite) nonincreasing sequence $\sigma_1 \geq \sigma_2 \geq \ldots$, where each eigenvalue is repeated according to its multiplicity. If such sequence is finite, we extend it to an infinite sequence by setting $\sigma_k = 0$ whenever σ_k is undefined (observe that 0 may not be an eigenvalue of K). Then we have the following *minimax* formula:

$$\sigma_k(K) = \sup_{\substack{\mathfrak{V} \text{ subspace of } H \\ \dim(\mathfrak{V})=k}} \min_{\substack{\xi\in\mathfrak{V} \\ \langle\xi,\xi\rangle=1}} \langle K\xi, \xi\rangle. \tag{4.2}$$

Observe that the supremum in (4.2) is actually attained if $\sigma_k > 0$.

Lemma 4.2. *Given $a, b, c \in \mathbb{R}$, for $t > 0$ sufficiently small and for all $V\colon [0,t] \to \mathbb{R}^n$ of class H^1 with $V(t) = 0$ and $V \neq 0$, we have:*

$$\int_0^t (\|V'\|^2 + a\|V'\|\|V\| + b\|V\|^2)\,dr + c\|V(0)\|^2 > 0, \tag{4.3}$$

where $\|\cdot\|$ denotes the Euclidean norm.

Proof. Let $t \in \,]0,1]$ be fixed and V be as in the statement. Using Young's inequality, for all $\varepsilon > 0$ we have

$$\|V'\| \cdot \|V\| \leq \frac{\varepsilon}{2}\|V'\|^2 + \frac{1}{2\varepsilon}\|V\|^2. \tag{4.4}$$

Since $V(t) = 0$, for $s \in [0,t]$ we have $V(s) = -\int_s^t V'\,dr$, hence, using Schwarz's inequality, we obtain:

$$\|V(s)\| \leq \int_s^t \|V'\|\,dr \leq \left(\int_0^t \|V'\|^2\,dr\right)^{1/2} \cdot \sqrt{t}, \tag{4.5}$$

hence

$$\int_0^t \|V'\|^2\,dr \geq \frac{1}{t^2}\int_0^t \|V\|^2\,dr. \tag{4.6}$$

Inequality (4.3) follows at once from (4.4) and (4.6). □

Theorem 4.3 (Morse Index Theorem). *Let $x\colon [0,1] \to \mathcal{M}$ be a $\mathcal{P}$-normal geodesic. If the number of $\mathcal{P}$-focal points along x is finite then the following identity holds:*

$$n_-(d^2E_x) = \sum_{t\in]0,1[} \operatorname{mul}(t). \tag{4.7}$$

Proof. For $t \in \,]0,1]$, let x_t denote the restriction of x to $[0,t]$, which is again a $\mathcal{P}$-normal geodesic, and by λ_t the corresponding Lagrangian multiplier.

Observe that λ_t is the restriction of λ to $[0,t]$. Denote by H_t the Hilbert space $T_{x_t}\Omega_{\mathcal{P},x(t)}(\Delta)$. By Proposition 4.1, $d^2E_{x_t}$ is represented on H_t by a self-adjoint operator of the form $I-K_t$, where K_t is compact. By considering extensions to zero in the interval $[t,1]$, we can regard each H_t as an isometrically embedded Hilbert subspace of $H_1 = T_x\Omega_{\mathcal{P},q}(\Delta)$, and, with this identification, we have the following identity:

$$\langle K_t\xi_1,\xi_2\rangle = \langle K_s\xi_1,\xi_2\rangle, \quad 0<t\le s\le 1, \text{ for all } \xi_1,\xi_2\in H_t.$$

For $k\ge 1$, set $\sigma_k(t)=\sigma_k(K_t)$ as in formula (4.2); since the map σ_k of (4.2) is continuous in the norm topology, for each k the continuity of the map $t\mapsto\sigma_k(t)$ can be deduced with arguments similar to [4, Proposition 3.3 and Corollary 4.5].

The left hand side of equality (4.7) is precisely the total number of indices k for which $\sigma_k(1)$ is greater than 1. By Corollary 3.5, the dimension of the kernel of $d^2E_{x_t}$ is precisely $\mathrm{mul}(t)$. It follows that $x(t)$ is a $\mathcal{P}$-focal point if and only if for some k it is $\sigma_k(t)=1$; in this case, the multiplicity $\mathrm{mul}(t)$ is equal to the total number of indices k for which $\sigma_k(t)=1$.

Using (2.11) and Lemma 4.2, it is easy to see that $d^2E_{x_t}$ is positive definite for $t>0$ sufficiently small, hence, for such t, $\sigma_k(t)<1$ for all k.

The function $\sigma_k(t)$ is nondecreasing on $]0,1]$; this is easily seen from formulas (4.2) and (4.8). The conclusion now follows from our assumption of finiteness for the number of $\mathcal{P}$-focal points, which implies that each σ_k assumes the value 1 exactly once whenever $\sigma_k(1)>1$. □

Using the ideas appearing in the proof of the index theorem we also get the following characterization of the set of $\mathcal{P}$-focal points. We recall that Δ is said to be a *contact distribution* if the following skew-symmetric bilinear form is nondegenerate on Δ:

$$\Delta\times\Delta\ni(v,w)\mapsto g(\nabla_vY,w)-g(\nabla_wY,w). \tag{4.9}$$

Proposition 4.4. *Let $x\colon[0,1]\to\mathcal{M}$ be a $\mathcal{P}$-normal geodesic. Then*

- *there are no $\mathcal{P}$-focal points along a sufficiently small initial portion of x,*
- *the set of $\mathcal{P}$-focal points consists of a finite union of isolated points and closed segments of x,*
- *if $(\mathcal{M},g)$ and Y are real-analytic, or if Δ is a contact distribution, then the set of $\mathcal{P}$-focal points along x is finite.*

Proof. The first statement is an immediate corollary of Lemma 4.2 and Corollary 3.5.

In the notations of the proof of Theorem 4.3, the second statement follows from the fact that the $\mathcal{P}$-focal points correspond to the zeroes of the continuous nondecreasing functions $\sigma_k(t)-1$. Observe also that there are only a finite number of indexes k for which $\sigma_k(1) \geq 1$; only for such indexes k it is possible for σ_k to assume the value 1 on $]0,1]$.

To prove the last statement, we start with the observation that if there is an infinite number of $\mathcal{P}$-focal points along x, then there exists a non zero $\mathcal{P}$-Jacobi field V along x that vanishes on a non trivial subinterval $[t,s] \subset]0,1]$. To prove this claim, observe that if there are infinitely many $\mathcal{P}$-focal points along x then there exists $k \in I\!N$ such that $\sigma_k = 1$ on an interval $[t,s]$.

Let $\mathfrak{W} \subset H_t$ be a k-dimensional subspace such that:

$$\min_{\substack{\xi\in\mathfrak{W}\\ \langle\xi,\xi\rangle=1}} \langle K_t\xi,\xi\rangle = 1.$$

Write H_s as an orthogonal sum $H_s = H^0 \oplus H^+ \oplus H^-$, where H^0, H^+, H^- are respectively the null, positive and negative eigenspace of $I - K_s$. Let $\pi\colon H_s \to H^-$ be the orthogonal projection; observe that π cannot be injective on $\mathfrak{W}$ because this would imply $\dim(\pi(\mathfrak{W})) = k$, and therefore

$$\sigma_k(s) \geq \min_{\substack{\xi\in\pi(\mathfrak{W})\\ \langle\xi,\xi\rangle=1}} \langle K_t\xi,\xi\rangle > 1 = \sigma_k(t),$$

contradicting $\sigma_k(s) = 1$. Let $\xi \in \mathfrak{W}$ be a non zero vector in $\mathrm{Ker}(\pi) = H^0 \oplus H^+$; write $\xi = \xi^0 + \xi^+ \in \mathfrak{W}$, with $\xi^0 \in H^0$ and $\xi^+ \in H^+$. If $\xi^+ \neq 0$, then $\langle K_s\xi^+,\xi^+\rangle < \langle\xi^+,\xi^+\rangle$. Hence,

$$\langle K_s\xi,\xi\rangle = \langle K_s\xi^+,\xi^+\rangle + \langle K_s\xi^0,\xi^0\rangle < \langle\xi^+,\xi^+\rangle + \langle\xi^0,\xi^0\rangle = \langle\xi,\xi\rangle,$$

which contradicts the fact that $\xi \in \mathfrak{W}$. Now, it follows that $\xi \in H^0$, i.e. ξ is a non zero $\mathcal{P}$-Jacobi field that vanishes on $[t,s]$.

Under the assumption that $(\mathcal{M}, g)$ and Y are real-analytic, then clearly ξ is also real-analytic, and we obtain a contradiction.

Let us consider now the case that Δ is a contact distribution. We have observed that non isolated $\mathcal{P}$-focal points determine non trivial $\mathcal{P}$-Jacobi fields V that vanish in some interval $[t,s]$. Hence, to conclude the proof we must show that every $\mathcal{P}$-Jacobi field that vanishes on a non trivial interval $[t,s]$ must vanish on $[0,1]$; to this aim, by the uniqueness of the solution of the system of ODE's given by equations (2.12) and (2.16), it suffices to show that the Lagrangian multiplier δ corresponding to V vanishes somewhere in $[t,s]$. To see this, observe that, since $V'' = 0$ on $[t,s]$, from (2.12) we get:

$$-\delta\,(\nabla Y)^*\dot{x} + (\delta\,Y)' = 0 \tag{4.10}$$

on $[t,s]$. If we multiply (4.10) by some vector $v \in \Delta = Y^\perp$ we get:

$$\delta(-g(\nabla_v Y, \dot{x}) + g(\nabla_{\dot{x}} Y, v)) = 0$$

and since (4.9) is nondegenerate on Δ it follows that $\delta = 0$ on $[t, s]$. This concludes the proof. $\square$

Observe that in the above proof, the finiteness of the number of $\mathcal{P}$-focal points along x could be obtained under the weaker assumption that x is never tangent to the kernel of the bilinear form (4.9) on Δ.

We conclude the section with the observation that, using Theorem 4.3 and a result of [16], it is easy to determine a formula for the index of the second variation of the sub-Riemannian action functional in the space of horizontal paths with *both* endpoints variable in two submanifolds of $\mathcal{M}$.

5. Minimality of $\mathcal{P}$-normal geodesics

In this section we prove the following minimality property as an application of the Morse Index Theorem. Given a $\mathcal{P}$-normal geodesic $x\colon [0, T] \to \mathcal{M}$, then, if $T > 0$ is sufficiently small, x is a *global* arc length minimizer among all horizontal curves in $\mathcal{M}$ connecting $\mathcal{P}$ and $x(T)$. The result is indeed not new, it is already proven in a very different way in [19, Proposition B.1]; the proof in [19] uses the Hamiltonian formalism and is an adaptation of the proof of minimality of normal geodesics between fixed endpoints presented in [10, Appendix C].

Let L denote the length functional on $\mathcal{M}$. In first place, we observe that a horizontal curve x is an *action minimizer* if and only if it is a length minimizer and $g(\dot{x}, \dot{x})$ is constant (see for instance [19, Appendix A]).

From the Morse Index Theorem we deduce that $x\colon [0, T] \to \mathcal{M}$, a $\mathcal{P}$-normal geodesic, that does not contain any $\mathcal{P}$-focal point is a *local* action minimizer with respect to the H^1-topology of $\Omega_{\mathcal{P},x(T)}(\Delta)$. Our first Lemma tells us that x is also a local action minimizer with respect to the C^0-topology of $\Omega_{\mathcal{P},x(T)}(\Delta)$.

Let dist denote the distance function induced by the Riemannian metric g on $\mathcal{M}$.

LEMMA 5.1. *Let $x\colon [0, T] \to \mathcal{M}$ be a $\mathcal{P}$-normal geodesic. For $T > 0$ small enough there exists $\varepsilon > 0$ such that if $y\colon [0, T] \to \mathcal{M}$ is an horizontal curve of class H^1 with $y(0) \in \mathcal{P}$, $y(T) = x(T)$ and $\operatorname{dist}(y(s), x(s)) < \varepsilon$ for all s, then $E(x) \leq E(y)$.*

PROOF. Let $T > 0$ be small enough so that there are no $\mathcal{P}$-focal points along x (Proposition 4.4) and such that the image of x is contained in the domain of a local chart of $\mathcal{M}$. For simplicity, the coordinate representation of the objects appearing in this proof will be denoted by the same letter as the original objects.

We can assume $x(T) = 0$. We argue by contradiction: suppose there exists a sequence $\{x_n\}$ in $\Omega_{\mathcal{P},x(T)}(\Delta)$ which is uniformly convergent to x and with

$$E(x_n) < E(x) \quad \text{for all } n. \tag{5.1}$$

By a reparameterization argument, we can assume that $g(\dot{x}_n, \dot{x}_n) = c_n$ is constant. Indeed it is possible to reparameterize the curves using arc length (cf. [19, Appendix A]) and then to see that, for any given curve, the parameterization by arc length minimizes the functional E among all reparameterizations. Since $E(x_n)$ is bounded, then c_n is a bounded sequence.

We consider the Hilbert space $\mathfrak{H}$ of $\mathbb{R}^n$-valued maps of class H^1 on $[0,T]$ vanishing at T, endowed with the following inner product:

$$\langle u_1, u_2\rangle = \frac{1}{2}\int_0^T g_{x(t)}(\dot{u}_1, \dot{u}_2)\,dt.$$

Observe that $E(x) = \langle x, x\rangle$. Moreover,

$$E(x_n) - \langle x_n, x_n\rangle = \frac{1}{2}\int_0^T (g_{x_n(t)} - g_{x(t)})(\dot{x}_n, \dot{x}_n)\,dt \to 0, \tag{5.2}$$

because $g_{x_n(t)} - g_{x(t)}$ is uniformly convergent to 0 and $\dot{x}_n$ is bounded in L^∞.

Since $E(x_n)$ is bounded, by (5.2) x_n is bounded in $\mathfrak{H}$, and, passing to a subsequence, we can assume that x_n is weakly convergent to x in $\mathfrak{H}$. By the lower semi-continuity of the norm with respect to the weak topology, we get:

$$E(x) = \langle x, x\rangle \le \liminf_{n\to\infty}\langle x_n, x_n\rangle = \liminf_{n\to\infty} E(x_n) \le \limsup_{n\to\infty} E(x_n) \le E(x).$$

By the above formula and (5.2) it follows that $\langle x_n, x_n\rangle$ converges to $\langle x, x\rangle$, hence x_n converges to x in $\mathfrak{H}$. This means that x_n tends to x in the H^1-topology of $\Omega_{\mathcal{P},x(T)}(\Delta)$, and therefore $E(x) \le E(x_n)$ for n sufficiently large, which contradicts (5.1) and proves the lemma. □

Proposition 5.2. *Let $x\colon [0,T] \to \mathcal{M}$ be a $\mathcal{P}$-normal geodesic. If $T > 0$ is small enough, then x is a global minimum point of E (and L) in $\Omega_{\mathcal{P},x(T)}(\Delta)$.*

Proof. Let T be small enough so that we can take $\varepsilon > 0$ as in Lemma 5.1. Now choose a possibly smaller T so that $L(x) \le \frac{\varepsilon}{2}$ and let $y \in \Omega_{\mathcal{P},x(T)}(\Delta)$. If $\operatorname{dist}(y(s), x(s)) < \varepsilon$ for all s, then by construction $E(y) \ge E(x)$.

If for some $s \in [0,T]$ we have $\operatorname{dist}(y(s), x(s)) \ge \varepsilon$, then by the triangle inequality we have:

$$L(y) \ge \operatorname{dist}(y(s), x(T)) \ge \operatorname{dist}(y(s), x(s)) - L(x) \ge \frac{\varepsilon}{2} \ge L(x).$$

Since $g(\dot{x}, \dot{x})$ is constant, a simple computation using the Cauchy–Schwarz inequality shows that also $E(y) \ge E(x)$. □

6. The Palais–Smale condition and the Morse relations for sub-Riemannian geodesics

In this section we prove a multiplicity result for sub-Riemannian geodesics joining the submanifold $\mathcal{P}$ and the point q. We use the Morse theory for the sub-Riemannian action functional E defined on the Hilbert manifold $\Omega_{\mathcal{P},q}(\Delta)$, endowed with the Riemannian structure:

$$\langle V,W\rangle_x = \int_0^1 g(V',W')\,dt, \quad x\in\Omega_{\mathcal{P},q}(\Delta),\ V,W\in T_x\Omega_{\mathcal{P},q}(\Delta). \tag{6.1}$$

The main technical assumption that allows the development of an infinite dimensional Morse Theory is the *Palais–Smale* condition.

We recall that, given a C^1 functional $F\colon X\to\mathbb{R}$ on a Hilbert manifold $(X,\mathfrak{h})$, then F satisfies the *Palais–Smale condition at level* $c\in\mathbb{R}$ if every sequence (x_k) in X such that

$$\lim_{k\to\infty} F(x_k)=c, \quad \lim_{k\to\infty}\mathfrak{h}(\nabla F(x_k),\nabla F(x_k))=0, \tag{6.2}$$

has a subsequence converging in X. By $\nabla F(x)$ we denote the *gradient* of F at $x\in X$, defined by $\mathfrak{h}(\nabla F(x),\,\cdot\,)=dF(x)$. A sequence (x_k) satisfying (6.2) is called a *Palais–Smale sequence* for F in X at the level c.

In order to prove the Palais–Smale condition for the sub-Riemannian action functional E we will assume that $(\mathcal{M},g)$ is a complete Riemannian manifold; in order to simplify our proof, we will also make the assumption that $Y(p)\in T_p\mathcal{P}$ for all $p\in\mathcal{P}$, without a significant loss of generality. Note that this assumption implies $\lambda(0)=0$ in the initial condition (2.6) for critical points of E in $\Omega_{\mathcal{P},q}(\Delta)$.

Proposition 6.1. *The functional E satisfies the Palais–Smale condition in $\Omega_{\mathcal{P},q}(\Delta)$ with respect to the Riemannian metric* (6.1) *at every level $c\in\mathbb{R}$.*

Proof. Let (x_k) a Palais–Smale sequence for E in $\Omega_{\mathcal{P},q}(\Delta)$ at the level $c\in\mathbb{R}$. By the completeness of $(\mathcal{M},g)$, the Theorem of Ascoli–Arzelà implies that a subsequence of (x_k) (still denoted by (x_k)) is uniformly convergent to some continuous curve $x\colon[0,1]\to\mathcal{M}$. Using local charts around the points of the image of x, by the boundedness of (x_k) in H^1 we get that x is also in $\Omega_{\mathcal{P},q}$.

Moreover, it follows that, up to subsequences, $\dot x_k$ converges to $\dot x$ weakly in L^2, in the following sense. Given a sequence (V_k), with V_k a vector field along x_k, we say that V_k tends to a vector field V along x *weakly in L^2* if for any subinterval $[t_0,t_1]\subset[0,1]$ and any chart whose domain contains $x([t_0,t_1])$, the coordinates of V_k tend to the coordinates of V weakly in $L^2([t_0,t_1],\mathbb{R}^n)$. Other types of convergence of vector fields along x_k to a vector field along x are defined similarly; in particular, we will consider norm and weak convergence in the spaces L^p and in the Sobolev spaces $W^{k,p}$.

For each k, let $u_k\in T_{x_k}\Omega_{\mathcal{P},q}(\Delta)$ be the gradient of E at x_k and let $a_k=u_k'$ be the covariant derivative of u_k along x_k. Then, a_k tends to 0 in L^2 and

$$\int_0^1 g(\dot x_k,\xi')\,dt=\int_0^1 g(a_k,\xi')\,dt \quad \text{for all } \xi\in T_{x_k}\Omega_{\mathcal{P},q}(\Delta). \tag{6.3}$$

It is not difficult to see that, for every $y \in \Omega_{\mathcal{P},q}(\Delta)$, there exists a projection operator

$$V \in T_y\Omega_{\mathcal{P},q} \mapsto (V - Y\psi_{y,V}) \in T_y\Omega_{\mathcal{P},q}(\Delta),$$

where $\psi_{y,V}$ is given by (see [3]):

$$\psi_{y,V}(t) = -\int_t^1 (g(V', Y) + g((\nabla Y)^* \dot{y}, V)) e^{\int_t^s g(\nabla_Y Y, \dot{y})\, dr}\, ds.$$

Setting $\xi = V - Y\psi_{y,V}$ in (6.3) and using integration by parts, we obtain the existence of a sequence (b_k) converging to 0 in L^2 and such that $\dot{x}_k - b_k = d_k$, where[1]

$$d_k(t) = \lambda_k Y + \int_0^t \lambda_k (\nabla Y)^* \dot{x}_k\, ds + z_k, \tag{6.4}$$

$$\lambda_k(t) = -\int_0^t g(\dot{x}_k, Y') e^{-\int_t^s g(\nabla_Y Y, \dot{x}_k)\, dr}\, ds, \tag{6.5}$$

and (z_k) is a sequence such that $z_k' = 0$.

We have the following facts:

- $g(\nabla_Y Y, \dot{x}_k)$ converges weakly in L^2,
- $\int_t^s g(\nabla_Y Y, \dot{x}_k)\, dr$ converges uniformly in s for each t,
- $g(\dot{x}_k, Y') e^{-\int_t^s g(\nabla_Y Y, \dot{x}_k)\, dr}$ is bounded in L^1,
- from (6.5) we get that (λ_k) is bounded in $W^{1,1}$,
- $\int_0^t \lambda_k (\nabla Y)^* \dot{x}_k\, ds$ is bounded in H^1,
- from (6.4), z_k is bounded in L^2,
- since $z_k' = 0$, z_k is bounded in $W^{1,1}$,
- from (6.4), d_k is bounded in $W^{1,1}$.

By the compact inclusion of $W^{1,1}$ in L^2 (see [2]), d_k has a converging subsequence in L^2. This fact implies that (x_k) has a converging subsequence in L^2, and the proposition is proved. □

Remark 6.2. Let us observe that, by classical results of Critical Point Theory (see e.g. [6]), Proposition 6.1 implies the existence of minimizers for E in $\Omega_{\mathcal{P},q}(\Delta)$.

The Global Morse Relations can be written in the following way. First recall that, given a topological space X, an algebraic field $\mathbb{K}$ and a natural number i, the *i-th Betti number $\beta_i(X;\mathbb{K})$ of X relative to* $\mathbb{K}$ is the $\mathbb{K}$-dimension of the i-th singular vector space $H_i(X;\mathbb{K})$ of X with coefficients in $\mathbb{K}$. The Poincaré polynomial $P_\lambda(X;\mathbb{K})$ of X with coefficients in $\mathbb{K}$ is the formal power series in $\lambda \in \mathbb{K}$ given by:

$$P_\lambda(X;\mathbb{K}) = \sum_i \beta_i(X;\mathbb{K})\lambda^i.$$

[1] the integral in (6.4) is meant as a "covariant primitive" with initial value 0, i.e. $\int_0^t \lambda_k (\nabla Y)^* \cdot \dot{x}_k\, ds$ is the unique vector field W along x_k such that $W(0) = 0$ and $W' = \lambda_k (\nabla Y)^* \dot{x}_k$.

Thanks to the completeness of $\mathcal{M}$ (and therefore of $\Omega_{\mathcal{P},q}(\Delta)$), the Palais–Smale condition (cf. Theorem 6.1) and the classical Morse relations (see e.g. [11]), we have the following theorem:

Theorem 6.3. *Suppose that $\mathcal{P}$ and q are not conjugate by sub-Riemannian geodesics. Then for every field $\mathbb{K}$ there exists a formal power series $Q_{\mathbb{K}}(\lambda)$ in the variable λ, with coefficients in $\mathbb{N}\cup\{\infty\}$, such that the following identity between formal series is satisfied:*

$$\sum_{z\in\mathcal{G}_{\mathcal{P},q}} \lambda^{m(z)} = P_\lambda(\Omega_{\mathcal{P},q}(\Delta);\mathbb{K}) + (1+\lambda)Q_{\mathbb{K}}(\lambda),$$

where $\mathcal{G}_{\mathcal{P},q}$ is the set of sub-Riemannian geodesics joining $\mathcal{P}$ with q and $m(z) = n_-(d^2E_z)$ is the Morse index of z as critical point of E on $\Omega_{\mathcal{P},q}(\Delta)$.

Remark 6.4. If Y does not have closed integral lines , using the flow of Y it is easy to prove that the inclusion of $\Omega_{\mathcal{P},q}(\Delta)$ in $\Omega_{\mathcal{P},q}$ is a homotopy equivalence. Moreover it is well known that $\Omega_{\mathcal{P},q}$ is homotopically equivalent to $\Omega^0_{\mathcal{P},q}$, the space of continuous path joining $\mathcal{P}$ with q. Then, by the Morse Index Theorem 4.3 we see that, if the number of $\mathcal{P}$-focal points along any sub-Riemannian geodesic is finite, the above relations can be written as

$$\sum_{z\in\mathcal{G}_{\mathcal{P},q}} \lambda^{\mu(z)} = P_\lambda(\Omega^0_{\mathcal{P},q};\mathbb{K}) + (1+\lambda)Q_{\mathbb{K}}(\lambda),$$

where $\mu(z)$ is the geometric index of z, which is the number of its $\mathcal{P}$-focal points counted with multiplicity.

References

[1] A. A. Agrachev and A. V. Sarychev, *Abnormal sub-Riemannian geodesics: Morse index and rigidity*, Ann. Inst. H. Poincaré **13** (1996), 635–690.

[2] H. Brezis, *Analyse Fonctionelle*, Masson, Paris, 1983.

[3] R. Giambó, F. Giannoni and P. Piccione, *Existence, multiplicity and regularity for sub-Riemannian geodesics by variational methods*, SIAM J. Control Optim. **40** (2002), 1840–1857.

[4] F. Giannoni, A. Masiello, P. Piccione and D. Tausk, *A Generalized Index Theorem for Morse-Sturm systems and applications to semi-Riemannian geometry*, Asian J. Math. **5** (2001), 441–472.

[5] F. Giannoni and P. Piccione, *An existence theory for relativistic brachistochrones in stationary spacetimes*, J. Math. Phys. **39** (1998), 6137–6152.

[6] F. Giannoni, P. Piccione and J. A. Verderesi, *An approach to the relativistic brachistochrone problem by sub–Riemannian geometry*, J. Math. Phys. **38** (1997), 6367–6381.

[7] U. Hamenstädt, *Some regularity theorems for Carnot–Carathéodory metrics*, J. Differential Geom. **32** (1990), 819–850.

[8] I. Kishimoto, *The Morse Index Theorem for Carnot–Carathéodory spaces*, J. Math. Kyoto Univ. **38** (1998), 287–293.

[9] S. Lang, *Differential Manifolds*, Springer–Verlag, Berlin, 1985.

[10] W. Liu and H. J. Sussmann, *Shortest paths for sub-Riemannian metrics on rank–2 distribution*, Mem. Amer. Math. Soc. **564** (1995).

[11] J. Mawhin and M. Willem, *Critical Point Theory and Hamiltonian Systems*, Springer–Verlag, Berlin, 1989.

[12] R. Montgomery, *Singular extremals on Lie groups*, Math. Control Signals **7** (1994), 217–234.

[13] ______, *Abnormal minimizers*, SIAM J. Control Optim. **32** (1994), 1605–1620.

[14] ______, *A survey of singular curves in sub-Riemannian geometry*, J. Dynam. Control Systems **1** (1995), 49–90.

[15] J. Milnor, *Morse Theory*, Princeton Univ. Press, Princeton, 1969.

[16] P. Piccione and D. V. Tausk, *A note on the Morse Index Theorem for geodesics between submanifolds in semi-Riemannian geometry*, J. Math. Phys. **40** (1999), 6682–6688.

[17] ______, *An Index Theorem for non periodic solutions of Hamiltonian systems*, Proc. London Math. Soc. **83** (2001), 351–389.

[18] ______, *The Maslov index and a generalized Morse Index Theorem for non positive definite metrics*, Comptes Rendus de l'Académie de Sciences de Paris **331** (2000), 385–389.

[19] ______, *Variational aspects of the geodesic problem is sub-Riemannian geometry*, J. Geom. Phys. **39** (2001), 183–206.

[20] ______, *Constrained Lagragians and degenerate Hamiltonians on manifolds: an Index Theorem*, Differential Equations and Dynamical Systems (Lisbon, 2000), pp. 309–324; Fields Inst. Commun. **31** (2002).

[21] ______, *On the Maslov and the Morse index for constrained variational problems*, J. Math. Pures Appl. (9) **81** (2002), 403–437.

Manuscript received October 10, 2002

Roberto Giambò and Fabio Giannoni
Dipartimento di Matematica e Informatica
Universitá di Camerino
Camerino, ITALY

E-mail address: roberto.giambo@unicam.it, fabio.giannoni@unicam.it

Paolo Piccione and Daniel V. Tausk
Departamento de Matemática
Universidade de São Paulo
Rua do Matao 1010, Cidade Universitaria São Paulo, SP, BRAZIL

E-mail address: piccione@ime.usp.br, tausk@ime.usp.br