

# GEODESICS AND ADMISSIBLE-PATH SPACES IN CARNOT GROUPS

A. A. AGRACHEV, A. GENTILE, AND A. LERARIO

ABSTRACT. We study the topology of the space  $\Omega_p$  of *admissible* paths between two points  $e$  (the origin) and  $p$  on a step-two Carnot group  $G$ :

$$\Omega_p = \{\gamma : I \rightarrow G \mid \gamma \text{ admissible, } \gamma(0) = e, \gamma(1) = p\}.$$

As it turns out,  $\Omega_p$  is homotopy equivalent to an infinite dimensional sphere and in particular it is contractible. The *energy* function:

$$J : \Omega_p \rightarrow \mathbb{R}$$

is defined by  $J(\gamma) = \frac{1}{2} \int_I \|\dot{\gamma}\|^2$ ; critical points of this function are sub-Riemannian geodesics between  $e$  and  $p$ . We study the asymptotic of the number of geodesics and the topology of the sublevel sets:

$$\Omega_p^s = \{\gamma \in \Omega_p \mid J(\gamma) \leq s\} \quad \text{as } s \rightarrow \infty.$$

If  $p$  is *not* a vertical point in  $G$ , the number of geodesics joining  $e$  and  $p$  is bounded and the homology of  $\Omega_p^s$  stabilizes to zero for  $s$  large enough.

A completely different behavior is experienced for the generic *vertical*  $p$ . In this case we show that  $J$  is a Morse-Bott function: geodesics appear in isolated *families* (critical manifolds), indexed by their energy. Denoting by  $l$  the corank of the horizontal distribution on  $G$ , we prove that:

$$\text{Card}\{\text{Critical manifolds with energy less than } s\} \leq O(s)^l.$$

Despite this evidence, Morse-Bott inequalities  $b(\Omega_p^s) \leq O(s)^l$  are far from being sharp and we show that the following stronger estimate holds:

$$b(\Omega_p^s) \leq O(s)^{l-1}.$$

Thus each single Betti number  $b_i(\Omega_p^s)$  ( $i > 0$ ) becomes eventually zero as  $s \rightarrow \infty$ , but the sum of all of them can possibly increase as fast as  $O(s)^{l-1}$ . In the case  $l = 2$  we show that indeed

$$b(\Omega_p^s) = \tau(p)s + o(s) \quad (l = 2).$$

The leading order coefficient  $\tau(p)$  can be *analytically* computed using the structure constants of the Lie algebra of  $G$ .

Using a dilation procedure, reminiscent to the rescaling for Gromov-Hausdorff limits, we interpret these results as giving some local information on the geometry of  $G$  (e.g. we derive for  $l = 2$  the rate of growth of the number of geodesics with bounded energy as  $p$  approaches  $e$  along a vertical direction).

## 1. INTRODUCTION

How many geodesics are there between two fixed points on a Riemannian manifold? The classical way to give an answer to this question, due to M. Morse, is to consider the space  $\Omega$  of *all* curves (defined on the same interval) joining the two points, together with the function  $J : \Omega \rightarrow \mathbb{R}$  that associates to each curve  $\gamma$  the number  $\frac{1}{2} \int_I \|\dot{\gamma}\|^2$  (the *Energy*). Geodesics joining the two points are thus regarded as *critical points* of the Energy, and the celebrated *Morse inequalities* state that, in the generic case, the number of such geodesics with Energy bounded by  $s$  is *at least* the number of “holes” of the topological space  $\{J \leq s\}$ .

In this Riemannian framework, if the two points  $p_0$  and  $p$  are close enough and we only allow curves with energy bounded by a small constant  $c > 0$ , there is only one geodesic joining them and the space of curves connecting  $p_0$  with  $p$  with energy bounded by  $c$  is contractible (bounding the energy ensures that the curves we consider are contained in a small neighborhood of  $p_0$ ).

In fact Morse inequalities can be used both ways: geometers study the topology of  $\{J \leq s\}$  to get information on geodesics, topologists use properties of geodesics in order to understand

the path space  $\Omega$ . This subject has been widely investigated, see for instance the classical works [15, 7, 8]

On a *sub-Riemannian* manifold  $M$  the situation gets more complicated. Being the velocities of curves constrained to the subbundle  $\Delta \subset TM$ , the *admissible*-path space is no longer the same as above and new phenomena can occur.

Typically, the space of admissible curves with energy less than  $s$  joining  $p_0$  and  $p$  has finite dimensional homology but its homological dimension and total Betti number may grow to infinity as  $p \rightarrow p_0$ . We are interested in the asymptotic behavior of these quantities and related properties of the space of sub-Riemannian geodesics joining  $p_0$  and  $p$ .

In this paper, we focus on the case of a step-two Carnot group where the situation is well controlled. A step-one Carnot group is just a Euclidean space. The general step- $k$  Carnot group  $G$  is a simply connected step- $k$  nilpotent Lie group equipped with a “dilation”. Such a dilation is a one-parametric group of automorphisms that generalizes homotheties of the Euclidean space. A typical example is the group of  $n \times n$  lower triangular matrices with units on the main diagonal, zeros over the diagonal and dilation  $\delta_t : \{a_{ij}\}_{i,j=1}^n \mapsto \{t^{i-j} a_{ij}\}_{i,j=1}^n, t > 0$  (Example 1 below is the case  $n = 3$ ).

First order elements with respect to the dilation form a vector subspace of the Lie algebra of the group and thus define a left-invariant vector distribution  $\Delta \subset TG$ . It is assumed that the first order elements generate the whole Lie algebra (see [18] and [2] for details and for the explanation of the role of Carnot groups in sub-Riemannian geometry).

There are no one or two dimensional nonabelian Carnot groups. The only three dimensional nonabelian Carnot group is the Heisenberg group: we discuss its geometry as it will be an enlightening, leading example for the rest of the paper.

*Example 1* (The Heisenberg group). The *Heisenberg group*  $G$  is the smooth manifold  $\mathbb{R}^3$  with the usual coordinates  $(x, y, z)$  and the distribution:

$$(1) \quad \Delta = \text{span} \left\{ \frac{\partial}{\partial x} - \frac{y}{2} \frac{\partial}{\partial z}, \frac{\partial}{\partial y} + \frac{x}{2} \frac{\partial}{\partial z} \right\}.$$

This means that one is allowed to move only along curves  $\gamma : I = [0, 2\pi] \rightarrow G$  whose velocity pointwise belongs to  $\Delta$  (these curves are called admissible and they are all defined on the same interval; we pick this interval to be  $[0, 2\pi]$  because it will simplify notations later). In fact, for a technical reason, one considers curves whose derivative is defined a.e. and is square integrable, ending up with absolutely continuous curves. The sub-Riemannian structure is given by declaring the above vector fields to be an orthonormal basis.

We assume one of the two points is the origin  $e = (0, 0, 0)$  (the structure is invariant by translation) and call the other  $p$ ; thus we consider:

$$\Omega_p = \{\text{admissible curves starting at } e \text{ and ending at } p\}.$$

As we will show in a while, there is a natural inclusion:

$$\Omega_p \hookrightarrow L^2(I, \mathbb{R}^2),$$

and we endow it with the induced topology. Recall that  $\Delta_p = \text{span}\{X(p), Y(p)\}$ , where  $X, Y$  are the two vector fields given in (1); in particular for every admissible curve  $\gamma$  we can write  $\dot{\gamma} = u_x X + u_y Y$ , where  $u_x, u_y \in L^2(I)$ . This correspondence is one-to-one: because of Cauchy’s theorem given  $u \in L^2(I, \mathbb{R}^2)$ , there is only one solution  $\gamma_u$  to the system  $\dot{\gamma} = u, \gamma(0) = e$ . Thus we can identify the space of *all* admissible curves starting at  $e$  with  $L^2(I, \mathbb{R}^2)$  and  $\Omega_p$  coincides with the set of curves such that  $\gamma_u$  is defined for  $t = 2\pi$  and  $\gamma_u(2\pi) = p$  (thus  $u$  represents the “coordinates” of  $\gamma_u$  and it is usually called its *control*).

The energy of an admissible curve is by definition (one half) the square of the  $L^2$  norm of its control:

$$J(\gamma_u) = \frac{1}{2} \int_0^{2\pi} \|\dot{\gamma}_u(t)\|^2 dt = \frac{\|u\|^2}{2}.$$

Critical points of this function restricted to the various  $\Omega_p$  (i.e. geodesics between  $e$  and  $p$ ) are curves whose projection on the  $(x, y)$ -plane is an arc of a circle (possibly with infinite radius, i.e. an interval on a straight line); the signed area swept out on the circle by this projection equals the  $z$ -coordinate of the final point.

We see that if  $p$  belongs to the  $(x, y)$ -plane there is only one geodesic joining it with the origin (this is precisely the segment trough  $e$  and  $p$ ), as if we were in a riemannian manifold; if  $p$  has both nonzero components in the  $(x, y)$  plane and the  $z$  axis, the number of geodesics is finite; finally, if  $p$  belongs to the  $z$ -axis there are infinitely many such geodesics (moreover given one we obtain infinitely many others by composing it with a rotation in  $\mathbb{R}^3$  along the axis).

The behavior of geodesics ending at a point  $(0, 0, z)$  (slightly abusing of notation we call this point  $z$ ) is really new if compared to the classical Riemannian situation and is worth discussing it in more details.

Given a natural number  $n \in \mathbb{N}_0$  let us pick a circle in the  $(x, y)$ -plane of area  $|z|/n$  passing through the origin; starting at  $e$  and moving clockwise around the circle  $n$  times, we obtain the projection of a geodesic ending at  $z$  and whose energy is  $2\pi n|z|$ ; rotating this curve along the  $z$ -axis of an angle  $\theta \in SO(2) \simeq S^1$  gives another such geodesic (thus we have an  $S^1$  of geodesics with the same energy and endpoints). Hence geodesics ending at  $z$  (i.e. critical points of  $J$  on  $\Omega_z$ ) arrange into *infinitely many families*, indexed by the values of their energy  $\{2\pi n|z|\}_{n \in \mathbb{N}}$ , each family being homeomorphic to an  $S^1$  in  $\Omega_z$ .

It is not possible, therefore, to study the topology of  $\Omega_z$  using  $f = J|_{\Omega_z}$  as a Morse function: simply because its critical points are not isolated. Despite this, as we have noted these critical points appear in nice families and in fact these families are *nondegenerate* in a sense that will be specified later, allowing to use the Morse theory machinery (this generalization is usually called *Morse-Bott* theory).

Before proceeding, let us show how to determine the topology of  $\Omega_z$ , by giving its equations in  $L^2(I, \mathbb{R}^2)$ . The map that associates to each  $u$  the final point  $\gamma_u(2\pi)$  is called the *End-point map*; it is a smooth map and will be denoted by  $F : L^2(I, \mathbb{R}^2) \rightarrow \mathbb{R}^3$ . A careful analysis of the definition of  $\gamma_u$  provides the explicit expression for  $F = (F_1, F_2, F_3)$ , see (11). In particular we can write the equations for  $\Omega_z$  as:

$$\Omega_z = \{F_1(u) = 0, F_2(u) = 0, F_3(u) = z\}.$$

The first two equations define a linear subspace  $H = \{F_1 = F_2 = 0\}$  and restricted to this space  $F_3$  is a quadratic form  $q = F_3|_H$  with *infinitely many* positive and *infinitely many* negative eigenvalues. In particular  $\Omega_z$  is homotopy equivalent to an infinite dimensional sphere:

$$\Omega_z \sim S^\infty$$

and is contractible. What kind of information can therefore a Morse theoretical study of  $\Omega_z$  give? As a starting point, the investigation of critical points of  $f = J|_{\Omega_z}$  gives information on the structure of geodesics; but we will see there is also more.

Let us find all the critical manifolds of  $f$  with energy bounded by  $s$ . Given a circle trough the origin of area  $|z|/n$ , it is the projection of a geodesic between  $e$  and  $z$  with energy bounded by  $s$  if and only if  $n \leq \frac{s}{2\pi|z|}$ . Rotating each such geodesic we get a whole critical manifold homeomorphic to  $S^1$ . In particular:

$$(2) \quad \{\text{geodesics between } e \text{ and } z \text{ with } J \leq s\} = \underbrace{S^1 \cup \dots \cup S^1}_{\lfloor \frac{s}{2\pi|z|} \rfloor \text{ copies}}$$

In this context the “amount” of geodesics can still be used as a measure of the topology of  $\Omega_z \cap \{J \leq s\}$ , in a form that generalizes Morse inequalities. Specifically, whenever the critical points of  $f$  arrange into smooth manifolds (with some nondegeneracy conditions), then the sum of the Betti numbers (i.e. the number of “holes”) of  $\{f \leq s\}$  is bounded by the sum of the Betti numbers of all the critical manifolds with  $f \leq s$ . The inequalities we get are called *Morse-Bott*

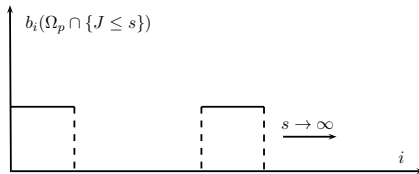


FIGURE 1. The homology of  $\Omega_p \cap \{J \leq s\}$  in the Heisenberg group: the top dimensional homology moves to infinity as  $s$  grows and eventually “disappears”.

inequalities, and in our case they give<sup>1</sup>

$$(3) \quad b(\Omega_z \cap \{J \leq s\}) \leq b(S^1 \cup \dots \cup S^1) \sim \frac{s}{2\pi|z|}.$$

Thus, as  $s \rightarrow \infty$  there are more and more critical manifolds but, surprisingly enough, the above inequality is far from being sharp. One can in fact show that  $\Omega_z \cap \{J \leq s\}$  has the homotopy type of a finite-dimensional sphere  $S^{d_s}$  in  $\Omega_z$ :

$$\Omega_z \cap \{J \leq s\} \sim S^{d_s} \quad \text{and} \quad b(\Omega_z \cap \{J \leq s\}) = 2.$$

In fact the dimension  $d_s$  of this sphere is increasing at a constant rate linear in  $s$  but the sum of its Betti numbers is constant: this shows that in (3) using Morse-Bott inequalities we were overcounting.

Hence the picture is the same as of a family of finite-dimensional spheres “approaching” an infinite dimensional one; the dimension of the spheres increases with  $s$  and the topology eventually vanishes (i.e.  $\lim_{s \rightarrow \infty} b_i(\Omega_z \cap \{J \leq s\}) = 0$  for every  $i \in \mathbb{N}_0$ ). Nonetheless there is something that “persists” at infinity: it is the *sum* of all Betti numbers of  $\Omega_z \cap \{J \leq s\}$ . What is the meaning of this invariant number?

A possible way to answer this question is to consider a local problem near  $e$ : assuming  $d(e, p_\epsilon) = \epsilon$  how many geodesics are there between  $e$  and  $p_\epsilon$  with *bounded* energy (say by a constant  $c$ )? As we have seen, if  $p_\epsilon$  is not a verical point, the answer is finitely many (depending on the upper bound on the energy). On the other hand, what happens if  $p_\epsilon$  approaches  $e$  along the non-riemannian  $z$ -axis? To answer this question let us take the point:

$$p_\epsilon = \left(0, 0, \frac{\epsilon^2}{4\pi}\right) \quad \text{such that} \quad d(e, p_\epsilon) = \epsilon.$$

Arguing as in (2), we see that the critical manifolds of  $J|_{\Omega_{p_\epsilon}}$  with energy bounded by  $c$  are  $\lfloor \frac{2c}{\epsilon^2} \rfloor$  copies of  $S^1$ : as we let  $p_\epsilon$  closer and closer to  $e$  the amount of geodesics keeps on increasing. The equations for  $\Omega_{p_\epsilon} \cap \{J \leq c\}$  are:

$$\Omega_{p_\epsilon} \cap \{J \leq c\} = \left\{ F_1(u) = F_2(u) = 0, q(u) = \frac{\epsilon^2}{2\pi}, \|u\|^2 \leq c \right\},$$

and the transformation  $u \mapsto u/(\epsilon\sqrt{2})$  gives a homeomorphism between  $\Omega_{p_\epsilon} \cap \{J \leq c\}$  and  $\Omega_{(0,0,\frac{1}{4\pi})} \cap \{J \leq \frac{2c}{\epsilon^2}\}$ . Hence the above discussion applies with  $z = (0, 0, \frac{1}{4\pi})$  and  $s = \frac{2c}{\epsilon^2}$ , giving:

$$\lim_{\epsilon \rightarrow 0} b(\Omega_{p_\epsilon} \cap \{J \leq c\}) = 2.$$

Thus, this persisting number can be interpreted as a local invariant of the Heisenberg group “in the direction” of  $z$ , as we will make more clear later.

This example shows that the sub-Riemannian case can be much richer than the riemannian, even at the level of the topology of the set of curves in  $\Omega_p$  when  $p$  is close to the initial point and the energy is bounded. Let us continue along these lines and ask if we can find, in the general

<sup>1</sup>For the rest of the paper,  $b(X)$  will indicate the *sum* of the  $\mathbb{Z}_2$ -Betti numbers of  $X$ , i.e.  $b(X) = \sum_i \text{rank} H_i(X; \mathbb{Z}_2)$ ; this sum might a priori be infinite, depending on the topological space  $X$ , but in the case of our interest it will always be finite.

case, some asymptotic behavior when the distance of the final point become smaller and the energy stays bounded. To this end we need a procedure for taking limit.

The “time-one” picture is the following: we have two points  $e$  and  $p$  at distance  $s$  on a sub-Riemannian manifold  $M$ , and the space  $\Omega_p \cap \{J \leq c\}$  of curves reaching a neighborhood of  $p$ . Now, on a neighborhood  $U$  of  $e$  there is a well defined family of non-isotropic “dilations” centered at  $e$  (we assume  $p$  is close enough to be in  $U$ ):

$$\delta_\lambda : U \rightarrow U, \quad \lambda > 0$$

with the property that  $d(e, \delta_\lambda(p)) = \lambda d(e, p)$  (for example, in the Heisenberg group this family operates as:  $\delta_\lambda(x, y, z) = (\lambda x, \lambda y, \lambda^2 z)$ ). We are thus led to consider the family of spaces:

$$(4) \quad \Omega_{\delta_\epsilon(p)} \cap \{J \leq c\}, \quad \epsilon \rightarrow 0$$

(this is exactly what we did in the Heisenberg case). This limiting procedure is reminiscent of the one used to construct the so called *sub-Riemannian* tangent space at  $e$ : as a vector space it is simply  $T_e M$ , but it comes endowed with much more structure that makes of it a Carnot group.

*Step-two* Carnot groups are “tangent spaces” to a generic sub-Riemannian manifold (the step-two condition needs some assumptions on the corank of  $\Delta$  to be generic).

Thus, on one hand our results should be interpreted as an “infinitesimal” version of a more general local theory (that the authors plan to discuss in a subsequent paper); on the other hand, Carnot groups are sub-Riemannian manifolds by themselves (plenty of literature has been devoted to them) and we believe this study is of autonomous interest.

The questions we will be interested in are:

- (i) What is the structure of geodesics connecting two points  $e$  and  $p$  (i.e. critical points of the energy on  $\Omega_p$ ) in a step-two Carnot group?
- (ii) What is the growth rate of the “number” of critical points at infinity? How does the topology of  $\Omega_p \cap \{J \leq s\}$  behave as  $s \rightarrow \infty$ ?
- (iii) Is there some invariant that can be captured for  $\Omega_{\delta_\epsilon(p)} \cap \{J \leq c\}$  when we perform the limit  $\epsilon \rightarrow 0$ ?

To start with we need to make our definition of a step-two Carnot group more precise. As a differentiable manifold it will simply be  $\mathbb{R}^{d+l}$ , together with the distribution  $\Delta \subset T\mathbb{R}^{d+l}$  generated by the span of the vector fields defined in coordinates  $(x, y)$ ,  $x \in \mathbb{R}^d$  and  $y \in \mathbb{R}^l$ , by:

$$E_i(x, y) = \frac{\partial}{\partial x_i}(x, y) - \frac{1}{2} \sum_{k=1}^l \sum_{j=1}^d a_{ij}^k x_j \frac{\partial}{\partial y_k}(x, y), \quad i = 1, \dots, d.$$

The numbers  $\{a_{ij}^k\}_{i,j,k}$  come from skew symmetric matrices  $A_k = (a_{ij}^k) \in \mathfrak{so}(d)$ ,  $k = 1, \dots, l$ ; these matrices span a  $l$ -dimensional vector space:

$$W = \text{span}\{A_1, \dots, A_l\} \subset \mathfrak{so}(d).$$

The sub-Riemannian structure is given by declaring the  $E_i$  to be an orthonormal basis of  $\Delta$ . Two different choices of bases for  $W$  will produce *isomorphic* Carnot groups; for this reason we will call  $W \subset \mathfrak{so}(d)$  the *Carnot group structure*.

Admissible curves will be absolutely continuous curve, whose velocities (a.e. defined) are in  $\Delta$ . As we did for the Heisenberg group,  $\Omega_p$  will denote the set of admissible curves joining the origin  $e$  with  $p$  and there will be a natural inclusion:

$$\Omega_p \hookrightarrow L^2(I, \mathbb{R}^d).$$

In this way the *Energy* of a curve  $\gamma_u$  with velocity  $u$  will simply be  $J(\gamma_u) = \frac{\|u\|^2}{2}$ .

With this notation we have the homeomorphism:

$$(5) \quad \Omega_{\delta_\epsilon(p)} \cap \{J \leq c\} = \delta_\epsilon(\Omega_p \cap \{J \leq c/\epsilon^2\}).$$

If  $p$  is not a vertical point (and  $d > l$ ), then the number of geodesics joining it with the origin is finite; in other words, the energy of geodesics connecting the origin with  $p$  is uniformly bounded

and the homology of the space  $\Omega_{\delta_\epsilon(p)} \cap \{J \leq c\}$  stabilizes as  $\epsilon \rightarrow 0$  (this is not true for a vertical  $p$ ).

**Proposition.** *If  $p$  is not a vertical point and  $d > l$ , the cohomology of  $\Omega_{\delta_\epsilon(p)} \cap \{J \leq c\}$  stabilizes for  $\epsilon$  small enough.*

Because of this, we are going to study asymptotics of the homological dimension and total Betti number of  $\Omega_{\delta_\epsilon(p)} \cap \{J \leq c\}$  as  $\epsilon \rightarrow 0$ , under the assumption:

$$(6) \quad p \text{ is a vertical point.}$$

Using the  $(x, y)$  coordinates as above the “vertical” direction is identified with the  $y$ -space and we denote such space by  $\Delta^2 \simeq \mathbb{R}^l$ ; in this way the condition (6) can be rewritten as  $p \in \Delta^2$ .

All the above questions can be addressed by studying the Morse-Bott theory of  $J|_{\Omega_p}$ , by taking appropriate rescalings. In fact, assuming (6), we have  $\delta_\epsilon(p) = \epsilon^2 p$  and the homogeneity of the end-point map on the vertical directions gives the homeomorphism (5). In particular the limit (4) can be studied by simply fixing  $p$  and letting the energy grow (at a rate  $\epsilon^{-2}$ ).

Our first result concerns the topology of the space  $\Omega_p$ : first we show that its homotopy type does not depend on  $p$ : in particular  $\Omega_p \sim \Omega_e$  and since the latter is defined by homogeneous equations, then these spaces are contractible. The homotopy equivalence is realized by a simple modification of the construction for ordinary loop spaces. Notice that  $\Omega_e$  is a singular space (it is homeomorphic to a cone based at the constant curve  $\gamma \equiv e$ ); despite this, for the generic choice of the Carnot group structure and the final point in the vertical direction,  $\Omega_p$  is a smooth (infinite-dimensional) manifold.

**Theorem** (The topology of  $\Omega_p$ ). *The topological space  $\Omega_p$  is homotopy equivalent to  $\Omega_e$  and is therefore contractible; moreover for the generic  $p \in \Delta^2$  it is a Hilbert manifold.*

In particular, analysis can be performed over  $\Omega_p$  and critical points of  $J|_{\Omega_p}$  (which is a smooth function) can be found using the Lagrange multipliers rule. As for Heisenberg, we can write down explicit equations: here the crucial point is that the End-point map sends a curve to  $\Delta^2$  if and only if its control has zero mean (i.e.  $0 = \int_I u \in \mathbb{R}^d$ ) and the restriction of this map to the space of such controls is *quadratic*.

A careful investigation of the critical points of  $f$  shows that they appear in families, i.e. they arrange into critical manifolds. These manifolds are tori  $S^1 \times \cdots \times S^1$  and are indexed by their Lagrange multipliers. The whole structure can be recovered by  $W \subset \mathfrak{so}(d)$  only, as follows.

Consider the infinite union of algebraic sets:

$$\Lambda = \bigcup_{n \in \mathbb{N}} \Lambda_n, \quad \Lambda_n = \{A \in W \mid \det(A - in\mathbb{1}) = 0\}.$$

Each set  $\Lambda_n$  is an hypersurface in  $W$ : generically it is smooth, but there can be points where  $iA$  has multiple integer eigenvalues and the corresponding hypersurface is singular. In fact all these hypersurfaces (except  $\Lambda_0$ ) can be obtained by dilations of  $\Lambda_1$  (i.e.  $\Lambda_n = n\Lambda_1$ ). Hence  $\Lambda$  looks like an “infinite net” with two kinds of singularities (they might appear at the same time): one kind comes from the singularities of each  $\Lambda_n$  and the other from the intersections  $\Lambda_{n_1} \cap \cdots \cap \Lambda_{n_\nu}$ , for  $\nu \leq l$ .

As we will see  $\Lambda$  represents the set of all possible Lagrange multipliers, for all possible final points, hence to each point of  $\Lambda$  there corresponds a family of geodesics. Notice that, once the basis  $\{A_1, \dots, A_l\}$  of  $W$  and coordinates are fixed, the correspondence:

$$\omega = (\omega_1, \dots, \omega_l) \mapsto \omega A = \omega_1 A_1 + \cdots + \omega_l A_l$$

defines a linear isomorphism  $(\Delta^2)^* \simeq W$ .

The next theorem gives a detailed answer to question (i) above on the structure of geodesics.

**Theorem** (The structure of geodesics). *Let  $u$  be the control associated to a geodesic from  $e$  to a point in  $\Delta^2$  with Lagrange multiplier  $\omega$ . Then:*

1.  $u(t) = e^{-t\omega A} u_0$  with  $u_0 = e^{-2\pi\omega A} u_0$ .

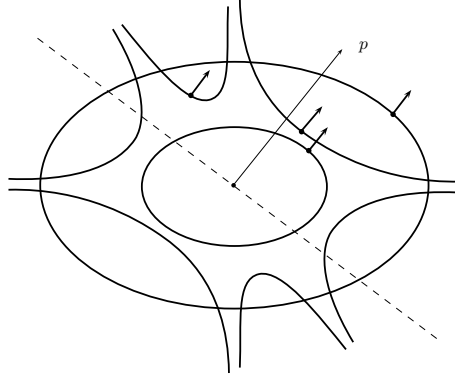


FIGURE 2. The set  $\bigcup_{n \in \mathbb{N}_0} \{\omega \in \Lambda_n \mid \omega(p) \geq 0, p \in (T_\omega \Lambda_1)^\perp\}$

2. The final point  $q(u) = (q_1(u), \dots, q_l(u)) \in \Delta^2$  of such geodesic is given by:

$$q_i(u) = \langle u_0, L_i(\omega)u_0 \rangle, \quad L_i(\omega) = \frac{1}{2} \int_0^{2\pi} \left( \int_0^t e^{\tau\omega A} d\tau \right) A_i e^{-t\omega A} dt.$$

3. The energy of such geodesic is given by  $J(u) = \omega(F(u))$ .

In addition, for the generic choice of  $W \subset \mathfrak{so}(d)$  and of  $p \in \Delta^2$ :

4. The set  $\Lambda(p) \subset \Lambda$  of Lagrange multipliers for geodesics whose final point is  $p$  is a discrete set.
5. If  $\omega \in \Lambda(p)$  then the integer eigenvalues of the matrix  $i\omega A$  are simple and there exist  $n_1, \dots, n_\nu \in \mathbb{N}_0$  such that  $\omega$  belongs to  $\Lambda_{n_1} \cap \dots \cap \Lambda_{n_\nu}$  (the number  $\nu \leq l$  will be called the number of resonances of  $\omega$ ).
6. If  $\omega$  has  $\nu$  resonances, the union of all the geodesics from  $e$  to  $p$  with Lagrange multiplier  $\omega$  is a smooth manifold  $C_\omega$  homeomorphic to:

$$C_\omega \simeq \underbrace{S^1 \times \dots \times S^1}_{\nu \text{ times}}.$$

Before proceeding any further, it is interesting to discuss one more example.

*Example 2* (Corank-two distributions). Let us consider the case of  $\mathbb{R}^{d+2}$  with Carnot group structure  $W = \text{span}\{A_1, A_2\} \subset \mathfrak{so}(d)$  having the above genericity property.

Let us consider the curve  $\Lambda_1 \subset W$  first: since the set of matrices with double eigenvalues has codimension three in  $\mathfrak{so}(d)$ ,  $\Lambda_1$  is a smooth curve (with possibly many components). In this case the set  $\{\omega \in \Lambda_1 \mid \omega(p) \geq 0, p \in (T_\omega \Lambda_1)^\perp\}$  consists of Lagrange multipliers whose critical manifold  $C_\omega$  is homeomorphic to  $S^1$ ; all the other Lagrange multipliers whose associated critical manifold is a circle are obtained multiplying these covectors by a positive natural number:

$$\{\omega \in \Lambda(p) \mid C_\omega \simeq S^1\} = \bigcup_{n \in \mathbb{N}_0} \{\omega \in \Lambda_n \mid \omega(p) \geq 0, p \in (T_\omega \Lambda_1)^\perp\}.$$

Let us now consider two natural numbers  $n_1, n_2$  such that  $\Lambda_{n_1} \cap \Lambda_{n_2} \neq \emptyset$  and let  $\omega$  be a point in this intersection such that  $\omega(p) \geq 0$ . Let  $E(\omega) = \{v \in \mathbb{R}^d \mid e^{2\pi\omega A} v = v\}$  and consider the quadratic forms  $q_i|_\omega : E(\omega) \rightarrow \mathbb{R}$  defined by:

$$q_i|_\omega : v \mapsto \langle v, L_i(\omega)v \rangle, \quad i = 1, 2.$$

Let  $q|_\omega : E(\omega) \rightarrow \mathbb{R}^2$  be the quadratic map whose components are the above  $q_i|_\omega$ . The set of Lagrange multipliers whose associated critical manifold is  $S^1 \times S^1$  coincides with:

$$\{\omega \in \Lambda(p) \mid C_\omega = S^1 \times S^1\} = \bigcup_{n_1, n_2 \in \mathbb{N}_0} \{\omega \in \Lambda_{n_1} \cap \Lambda_{n_2} \mid p \in \text{im}(q|_\omega)\}.$$

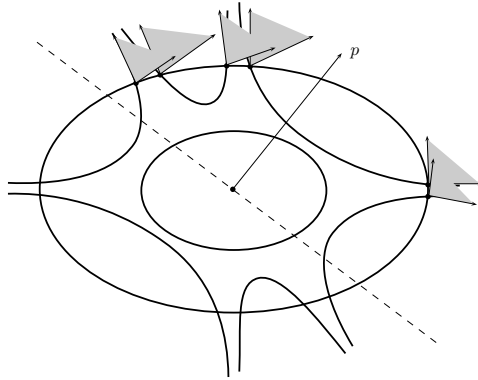


FIGURE 3. The set  $\bigcup_{n_1, n_2 \in \mathbb{N}_0} \{\omega \in \Lambda_{n_1} \cap \Lambda_{n_2} \mid p \in \text{im}(q|_\omega)\}$ .

Let us move now to question number (ii) above, on the growth rate of the number of critical manifolds of  $J|_{\Omega_p}$  as the energy goes to infinity. As we have seen for the Heisenbger group in (2), once  $p$  is fixed the number of critical manifolds with energy bounded by  $s$  increases linearly in  $s$ :

$$\text{Card}\{\text{critical manifolds of } J|_{\Omega_p} \text{ such that } J \leq s\} \sim O(s) \quad (\text{Heisenberg}).$$

The previous example shows that in the corank two case the critical manifolds of  $J|_{\Omega_p}$  can be indexed by points on an infinite two-dimensional net and, naively, it is reasonable to guess that they increase quadratically in  $s$ . Indeed it is a general fact that this number increases at most as a polynomial in  $s$  of degree  $l = \text{corank}(\Delta)$  (see Figure 4).

**Theorem** (The growth rate of the number of critical manifolds). *For the generic choice of  $W \subset \mathfrak{so}(d)$  and of  $p \in \Delta^2$  we have:*

$$(7) \quad \text{Card}\{\text{critical manifolds of } J|_{\Omega_p} \text{ such that } J \leq s\} \leq O(s)^l, \quad l = \dim(\Delta^2).$$

Using the inequality in (7) and the fact that the topology of the critical manifold is uniformly bounded by a constant depending on  $d$  and  $l$  only, one is tempted to answer the second part of question (ii), concerning the topology of  $\Omega_p \cap \{J \leq s\}$ , using Morse-Bott inequalities. What we would get with this strategy is, for a corank  $l$  distribution:

$$(8) \quad b(\Omega_p \cap \{J \leq s\}) \leq O(s)^l.$$

Surprisingly enough, this estimate can be improved up to an  $O(s)^{l-1}$  (as we already noticed, in the corank-one case  $b(\Omega_p \cap \{J \leq s\})$  is indeed constant). Before discussing the general case, let us continue the corank-two example.

*Example 3* (Corank-two: a topological coarea formula). In this case not only we can show that  $b(\Omega_p \cap \{J \leq s\})$  grows as  $O(s)$ , but also we can analytically compute the leading coefficient.

To explain the result, consider a unit circle  $S^1 \subset W \simeq (\Delta^2)^*$ . As the parameter  $\theta$  varies on  $S^1$  let us consider the positive eigenvalues  $\alpha_1(\theta), \dots, \alpha_j(\theta)$  of the matrix  $i \cos(\theta)A_1 + i \sin(\theta)A_2$ . The genericity assumption ensures that these numbers can be taken as the value of semialgebraic functions  $\alpha_j : S^1 \rightarrow \mathbb{R}$ . Given  $p \in \Delta$  we consider the *rational* functions  $\lambda_j : S^1 \rightarrow \mathbb{R} \cup \{\infty\}$  given by:

$$\lambda_j : \theta \mapsto \left| \frac{\alpha_j(\theta)}{p_1 \cos(\theta) + p_2 \sin(\theta)} \right| \quad \text{for } j = 1, \dots, d.$$

Notice that when  $\omega$  approaches  $p^\perp$  these functions might explode, that is why they are rational in  $\theta$ ; on the other hand they are semialgebraic and differentiable almost everywhere and it makes sense to consider the integral:

$$\tau(p) \doteq \frac{1}{2} \int_{S^1} \sum_{j=1}^d \left| \frac{\partial \lambda_j}{\partial \theta}(\theta) \right| - \left| \sum_{j=1}^d \frac{\partial \lambda_j}{\partial \theta}(\theta) \right| d\theta.$$



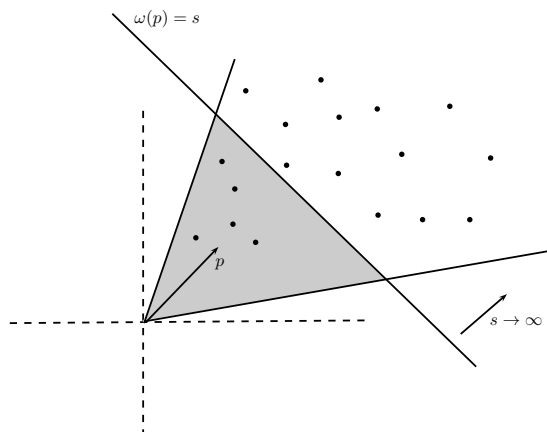


FIGURE 4. (Corank two) The number of critical manifolds with energy less than  $s$  can grow quadratically in  $s$  (this number is proportional to the volume of the shaded region; dots are Lagrange multipliers).

The convergence of the integral follows from the fact that where the derivatives explode, they all have the same sign and the integrand vanishes.

*Theorem* (Topological coarea formula). *If the corank  $l = 2$ , for a generic choice of  $W \subset \mathfrak{so}(d)$  and  $p \in \Delta^2$  we have:*

$$(9) \quad b(\Omega_p \cap \{J \leq s\}) = \tau(p)s + o(s).$$

The name “topological coarea” reminds of integral geometry: in fact the coefficient  $\tau(p)$  is computed by considering the sum of the Betti numbers  $b = \sum_i b_i$  as an integral over the index set, regarding  $i$  as the “variable” of integration; rearranging this sum in an appropriate way, in the limit gives exactly the coarea formula for functions of one variable. In fact, looking at the inequalities defining  $\Omega_p \cap \{J \leq s\}$ , one immediately sees that they are quadratic. There is a general theory (see Appendix C or [3] for more details) for studying the topology of sets defined by quadratic inequalities: the idea is to consider quadratic forms obtained by taking linear combinations of the equations defining the set, i.e. considering quadratic forms depending on some (homogeneous) parameters (there are as many parameters as the number of inequalities minus one). The main ingredient is the function on this parameter space that counts the number of positive eigenvalues (the *positive inertia index*) of the corresponding quadratic form. Roughly, the theory says that for each “change” in the monotony of this function, there corresponds a “hole” in the set. Thus one can count, for each level of the positive inertia index, how many homology classes are at that level, and this gives another rearrangement of the above sum.

In the general case the upper bound we get is the following (see Figure 5).

**Theorem** (Bound on the growth rate of the topology). *For the generic choice of  $W \subset \mathfrak{so}(d)$  and of  $p \in \Delta^2$  we have:*

$$b(\Omega_p \cap \{J \leq s\}) \leq O(s)^{l-1}, \quad l = \dim(\Delta^2).$$

Here the idea is to use the function  $-J|_{\Omega_p}$ , which is again Morse-Bott, whose critical manifolds are the same of  $J|_{\Omega_p}$  but with *infinite* index. Thus passing one of these critical manifolds amounts of attaching an infinite dimensional cell (i.e the homotopy of the sublevel set doesn’t change). Thus one can start with  $\Omega_p \cap \{J \leq s\}$  and “push” its topology to its boundary  $\Omega_p \cap \{J = s\}$  via  $\nabla J|_{\Omega_p}$ . This observation will have important consequences for our study, including the above result.

If we look back again at the Heisenberg example, we see that  $b_i(\Omega_p \cap \{J \leq s\})$  is zero for  $i > O(s)$ , simply because  $\Omega_p \cap \{J \leq s\}$  has the homotopy type of a sphere whose dimension grows linearly in  $s$ . It is in fact a general phenomenon that the maximum nonzero Betti number increases at most linearly in  $s$  (with no restrictions on the corank).

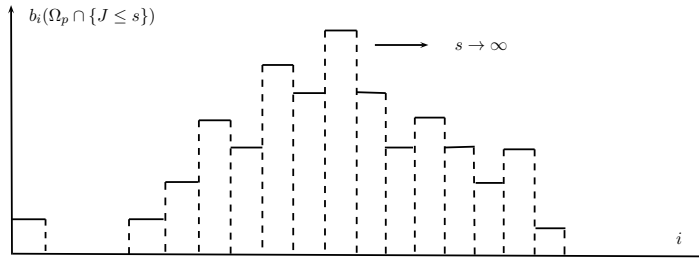


FIGURE 5. The Betti numbers of  $\Omega_p \cap \{J \leq s\}$ . As  $s$  goes to infinity the “wave” moves to the right and the sum of all the Betti numbers (the area below the wave) can increase as fast as  $O(s^{l-1})$ , but eventually everything “disappears”.

**Theorem** (The vanishing rate of Betti numbers). *For the generic choice of  $W \subset \mathfrak{so}(d)$  and of  $p \in \Delta^2$  we have:*

$$\max\{i \mid b_i(\Omega_p \cap \{J \leq s\}) \neq 0\} \leq O(s).$$

At this point the study of question (ii) above is complete and one can address question (iii). What we know so far is that the whole  $\Omega_p$  is contractible and we have rather precise asymptotics on the behavior of the number of geodesics in  $\Omega_p \cap \{J \leq s\}$  and its total Betti number as  $s \rightarrow \infty$ . What we have to do is simply to translate the results using the homeomorphism (5):

$$\Omega_{\delta_\epsilon(p)} \cap \{J \leq c\} \simeq \Omega_p \cap \{J \leq s\} \quad \text{as } s = c/\epsilon^2 \rightarrow \infty.$$

For example one might ask whether the estimate for the number of critical points and the topology of  $\Omega_{\delta_\epsilon(p)} \cap \{J \leq c\}$  can be actually attained. In the case  $l = 2$  we have seen that the leading coefficient  $\tau(p)$  of  $b(\Omega_{\delta_\epsilon(p)})$  can be analytically computed and in fact for the generic Carnot group structure and  $p \in \Delta^2$  it is not difficult to show that  $\tau(p) \neq 0$ . In particular as we let  $\epsilon \rightarrow 0$ , the topology of  $\Omega_{\delta_\epsilon(p)}$  explodes; as a corollary the number of geodesics grows unbounded as well (this is an example of a backward use of Morse-Bott inequalities).

For the general corank  $l$  we do not know whether the limit:

$$\tau(p) = \limsup_{\epsilon \rightarrow 0} b(\Omega_{\delta_\epsilon(p)} \cap \{J \leq 1\})\epsilon^{2l-2} = \limsup_{\epsilon \rightarrow 0} b(\Omega_p \cap \{J \leq 1/\epsilon\})\epsilon^{l-1}$$

is different from zero, but it is natural to guess so (we only know it is finite).

As a concluding remark, we believe these asymptotic Morse inequalities can be useful for generalizations and, in particular, for the next paper that we plan to devote to the general step-two sub-Riemannian structure. Of course, in general, we cannot guarantee that the functional is Morse-Bott. The quantity to estimate could be a “virtual number of geodesics”: the minimal number of critical points for a Morsification of the functional.

**1.1. Structure of the paper.** The paper is organized as follows. In section 2 we discuss some preliminary material and give the main definitions. In Section 3 we study the structure of geodesics: the theorem on the smoothness and the topology of  $\Omega_p$  is Theorem 9 and the theorem on the structure of geodesics is a combination of Lemmas 10, 11 and 12 and Theorem 13. The growth of the number of critical manifolds is computed in Theorem 21; the bound on the growth rate of the topology of  $\Omega_p \cap \{J \leq s\}$  is proved in Theorem 26 and the exact asymptotic for the case  $l = 2$  is the content of Theorem 27; the theorem on the vanishing rate of the Betti numbers is Corollary 25. The Appendix contains, for the reader’s convenience, some useful results we will use in the main body.

## 2. PRELIMINARIES

**2.1. Step two Carnot Groups and their geometry.** Here we briefly recall the main definitions related to Carnot groups; the reader is referred to [2, 19, 16] for more details.

As a differentiable manifold a *step two Carnot group* is a connected, simply connected Lie group  $G$  whose Lie algebra  $\mathfrak{g} = T_e G$  decomposes as:

$$\mathfrak{g} = \Delta \oplus \Delta^2, \quad \text{with} \quad [\Delta, \Delta] = \Delta^2, \quad [\mathfrak{g}, \Delta] = 0;$$

as vector spaces here we have  $\Delta \cong \mathbb{R}^d$  and  $\Delta^2 \cong \mathbb{R}^l$ .

Whenever a Lie algebra  $\mathfrak{g}$  as above is given, the existence (and uniqueness) of such a group  $G$  is guaranteed by Lie's theorem. We recall that in fact under the above assumption on the structure of  $\mathfrak{g}$ , the exponential map  $\exp : \mathfrak{g} \rightarrow G$  is an analytic diffeomorphism, hence in particular  $G \simeq \mathbb{R}^{d+l}$ .

The geometric structure on  $G$  is given by fixing a scalar product  $h$  on  $\Delta$  and considering the distribution  $\Delta_g = dL_g \Delta$  together with the extension of  $h$  by left translation; in this way the triple  $(G, \Delta, h)$  defines a *sub-Riemannian manifold*. Notice that the distribution  $\Delta$  is by assumption bracket generating, hence it satisfies Hörmander's condition, in particular meaning that  $G$  is *admissible-path-connected* (see below).

Two such Carnot groups  $(G_1, \Delta_1, h_1)$  and  $(G_2, \Delta_2, h_2)$  are considered to be isomorphic if there exists a Lie algebra isomorphism  $L : \mathfrak{g}_1 \rightarrow \mathfrak{g}_2$  such that  $L\Delta_1 = \Delta_2$  and  $L|_{\Delta_1}^* h_2 = h_1$  : in fact by the simple connectedness assumption the linear map  $L$  integrates to a Lie group isomorphism  $\phi : G_1 \rightarrow G_2$  and the global geometric structures are related by  $(G_1, \Delta_1, h_1) = (G_2, \phi^{-1}\Delta_2, \phi^*h_2)$ .

If we fix now an *orthonormal* basis  $\{e_1, \dots, e_d\}$  of  $\Delta$  and a basis  $\{f_1, \dots, f_l\}$  of  $\Delta^2$ , then the bracket structure can be written as:

$$(10) \quad [e_i, e_j] = \sum_{k=1}^l a_{ij}^k f_k, \quad \text{for all } i, j \in \{1, \dots, d\}$$

where each matrix  $A_k = (a_{ij}^k)$  belongs to  $\mathfrak{so}(d)$ . In particular we can consider the vector space:

$$W = \text{span}\{A_1, \dots, A_l\} \subset \mathfrak{so}(d).$$

Each vector space  $W$  of dimension  $l$  in  $\mathfrak{so}(d)$  defines a sub-Riemannian structure on  $G$  by fixing a basis  $\{A_1, \dots, A_l\}$  for  $W$  and declaring that the corresponding matrices define the bracket structure in an orthonormal basis  $\{e_1, \dots, e_d\}$  for  $\Delta$ . The isomorphism class of the Carnot group does not depend on the choice of the basis of  $W$ : indeed, let  $\{A'_1, \dots, A'_l\}$  be another basis of  $W$  and  $B = (b_{hk})$  the basis-change matrix, such that  $A'_h = \sum_{k=1}^l b_{hk} A_k$ . Now we can build another Carnot group by defining its Lie algebra  $\mathfrak{g}'$  with basis  $\{e'_1, \dots, e'_d, f'_1, \dots, f'_l\}$  such that  $\{e'_1, \dots, e'_d\}$  is orthonormal; the structure constants are given by the entries of the matrices  $A'_h$  as in equations (10). The map  $\phi$ , defined on the basis elements by  $e'_i \mapsto e_i$  and  $f'_h \mapsto \sum_{k=1}^l b_{hk} f_k$ , gives an isomorphism between  $\mathfrak{g}'$  and  $\mathfrak{g}$ .

*Remark 1* (The moduli space of Carnot Groups). Once a sub-Riemannian structure is given, changing  $\{e_1, \dots, e_d\}$  to another orthonormal basis  $\{Me_1, \dots, Me_d\}$  (where  $M$  is an orthogonal matrix in  $O(d)$ ) changes  $W$  into  $W' = MWM^T$ . Thus denoting by  $G(l, \mathfrak{so}(d))$  the Grassmannian of  $l$ -planes in  $\mathfrak{so}(d)$ , the (naive) moduli space of step two Carnot Groups is represented by the quotient  $\mathcal{M}_{l,d} = G(l, \mathfrak{so}(d))/O(d)$ . Since  $\mathcal{M}_{l,d}$  is the quotient of a manifold by a Lie group action, the quotient map is open and perturbing  $W$  defines a "genuine" perturbation of the isomorphism class of the corresponding Carnot group; in particular this means that a generic choice of  $W$  results in a generic choice of an isomorphism class of Carnot groups.

Motivated by the above remark, once  $W \in G(l, \mathfrak{so}(d))$  is fixed we consider the Carnot group given by exponentiating  $\mathfrak{g} = \mathbb{R}^d \oplus \mathbb{R}^l$ , whose Lie algebra is given as follows:  $\{e_1, \dots, e_d\}$  is the standard orthonormal basis for  $\mathbb{R}^d$ ,  $\{f_1, \dots, f_l\}$  is the standard basis for  $\mathbb{R}^l$  and fixing a basis  $\{A_1, \dots, A_l\}$  for  $W$ , the Lie brackets are given by equation (13); we will call  $W$  the *Carnot group structure*.

The following theorem gives a geometric realization of Carnot groups.

**Theorem 1.** *Let  $\{A_1, \dots, A_l\}$  be a basis for  $W \subset \mathfrak{so}(d)$  and for  $i = 1, \dots, d$  consider the vector fields  $E_i$  on  $\mathbb{R}^{d+l}$  defined in coordinates  $(x, y)$  by:*

$$E_i(x, y) = \frac{\partial}{\partial x_i}(x, y) - \frac{1}{2} \sum_{k=1}^l \sum_{j=1}^d a_{ij}^k x_j \frac{\partial}{\partial y_k}(x, y).$$

*Then the sub-Riemannian manifold  $(\mathbb{R}^{d+l}, \Delta = \text{span}\{E_1, \dots, E_l\}, g)$ , where  $g$  is the standard Euclidean metric, is isomorphic to the Carnot group defined by  $W$ .*

**2.2. The End-point map.** *Admissible paths* on  $G$  are defined to be curves  $\gamma : I = [0, 2\pi] \rightarrow G$  whose derivative exists almost everywhere, is square integrable and belongs to the distribution  $\Delta$ . We denote the set of such paths by  $\Omega$ . Our choice of the interval  $I = [0, 2\pi]$  is motivated by simplicity of notation: we will later need to expand the components of an admissible path into their Fourier series; a different choice of the interval will produce a completely equivalent theory.

As earlier noticed, the bracket generating condition implies that any two points in  $G$  can be joined by an admissible path. The set  $\Omega$  can be given a *Hilbert manifold structure* as follows. Let  $u = (u_1, \dots, u_d) \in L^2(I, \mathbb{R}^d)$  and consider the Cauchy problem:

$$\dot{\gamma}(t) = \sum_{i=1}^d u_i(t) E_i(\gamma(t)), \quad \gamma(0) = e.$$

The reader can assume this ODE problem is set on  $\mathbb{R}^{d+l}$ , using the above geometric realization Theorem; in this case the identity element  $e \in G$  corresponds to the zero of  $\mathbb{R}^{d+l}$ . By Caratheodory's Theorem the above Cauchy problem has a local solution  $\gamma_u$  and we consider the set:

$$\mathcal{U} = \{u \in L^2(I, \mathbb{R}^d) \mid \gamma_u \text{ is defined for } t = 2\pi\}.$$

For general subriemannian manifolds the set  $\mathcal{U}$  is an open subset of  $L^2(I, \mathbb{R}^d)$  (by ODE's continuous dependence theorem) and is called the set of *controls*; in our case the estimates for the final time can be made uniform and we actually have  $\mathcal{U} = L^2(I, \mathbb{R}^d)$ . Associating to each  $u$  the corresponding path  $\gamma_u$  gives thus a local coordinate chart and by slightly abusing of notation in the sequel we will often identify  $\Omega$  with  $\mathcal{U}$ .

Once we are given the Carnot group structure  $W = \text{span}\{A_1, \dots, A_l\}$ , we can use Theorem 1 to write down the above ODE in a more explicit form:

$$\begin{cases} \dot{x} = u \\ \dot{y}_i = \frac{1}{2} x^T A_i u \end{cases} \quad \text{and} \quad \gamma(0) = 0$$

In this framework the *End-point map* is the *smooth* map:

$$F : \Omega \longrightarrow G,$$

which associates to each curve  $\gamma$  its final point  $\gamma(2\pi)$ .

If we define  $A = (A_1, \dots, A_l)$ , we can use again Theorem 1 and write the End-point map as:

$$(11) \quad F(u) = \left( \int_0^{2\pi} u(t) dt, \frac{1}{2} \int_0^{2\pi} \left\langle \int_0^t u(\tau) d\tau, A u(t) \right\rangle dt \right);$$

(here the brackets denote the sub-Riemannian scalar product on the Lie algebra  $\mathfrak{g}$ ).

In the sequel we will mainly be interested in admissible paths whose endpoints lie on  $\Delta^2$  (see Proposition 7 below). Being  $G$  of step two, we know that  $\Delta^2$  is an abelian subalgebra of  $\mathfrak{g}$ ; therefore we can identify  $\Delta^2$  with the submanifold  $\exp(\Delta^2) \subset G$  and using Theorem 1 we can write this identification as:

$$\xi_1 f_1 + \dots + \xi_l f_l \mapsto (\underbrace{0, \dots, 0}_x, \underbrace{\xi_1, \dots, \xi_l}_y),$$

(here as above  $\{f_1, \dots, f_l\}$  is a basis of  $\Delta^2$ ).

We study now the structure of the set of admissible paths whose endpoints are on  $\Delta^2$ . It turns out that in the local coordinates given by the controls  $\mathcal{U}$  it coincides with the kernel of the differential of  $F$  at  $0 \in \mathcal{U}$

$$H = \ker D_0F,$$

as described by the following proposition.

**Proposition 2.** *The following properties hold:*

- (a)  $H = \{u \in L^2(I, \mathbb{R}^d) \mid \int_I u \, dt = 0\}$ ;
- (b)  $u \in H \Leftrightarrow F(u) \in \exp(\Delta^2)$ ;
- (c)  $F|_H = \text{He}_0F$ .

*Proof.* For point (a) we compute the differential  $D_0F$ : by taking a variation  $\varepsilon v$  of the constant curve  $\gamma \equiv 0$  we easily see that

$$D_0Fv = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \left( \varepsilon \int_0^{2\pi} v(t) \, dt, \frac{1}{2} \varepsilon^2 \int_0^{2\pi} \left\langle \int_0^t v(\tau) \, d\tau, A v(t) \right\rangle dt \right) = \left( \int_0^{2\pi} v(t) \, dt, 0 \right) \in \mathfrak{g}.$$

which proves property (a).

Point (b) is a direct consequence of equation (11).

For point (c) we notice that the Hessian  $\text{He}_0F$  is defined on  $H = \ker D_0F$  with values in  $\text{coker } D_0F = \Delta^2$ , thus has the same range as  $F|_H$ . As in the proof of (a) if we consider the second derivative of a variation and we easily obtain  $\text{He}_0F$  has the same expression of  $F$  when restricted to  $H$ .  $\square$

We denote by  $q$  the Hessian of  $F$  at zero, i.e. the quadratic map:

$$q \doteq F|_H : H \rightarrow \mathbb{R}^l.$$

Every component  $q_i$  of  $q$  is a quadratic form on  $H$  and its explicit expression is given by:

$$(12) \quad q_i(u) = \frac{1}{2} \int_0^{2\pi} \left\langle \int_0^t u(\tau) \, d\tau, A_i u(t) \right\rangle dt.$$

By polarization we obtain the expression for the associated bilinear form:

$$\begin{aligned} q_i(u, v) &= \frac{1}{4} \left( \int_0^{2\pi} \left\langle \int_0^t u(\tau) + v(\tau) \, d\tau, A_i(u(t) + v(t)) \right\rangle dt - \int_0^{2\pi} \left\langle \int_0^t u(\tau) \, d\tau, A_i u(t) \right\rangle dt + \right. \\ &\quad \left. - \int_0^{2\pi} \left\langle \int_0^t v(\tau) \, d\tau, A_i v(t) \right\rangle dt \right) = \\ &= \frac{1}{4} \left( \int_0^{2\pi} \left\langle \int_0^t u(\tau) \, d\tau, A_i v(t) \right\rangle dt + \int_0^{2\pi} \left\langle \int_0^t v(\tau) \, d\tau, A_i u(t) \right\rangle dt \right) = \\ &= \frac{1}{4} \left( \int_0^{2\pi} \left\langle \int_0^t u(\tau) \, d\tau, A_i v(t) \right\rangle dt - \int_0^{2\pi} \left\langle v(t), A_i \int_0^t u(\tau) \, d\tau \right\rangle dt \right) = \\ &= \frac{1}{2} \int_0^{2\pi} \left\langle \int_0^t u(\tau) \, d\tau, A_i v(t) \right\rangle dt, \end{aligned}$$

where the fourth row follows from integration by parts.

Moreover to every  $q_i$  it corresponds a symmetric operator  $Q_i : H \rightarrow H$  defined by:

$$q_i(u) = \langle u, Q_i u \rangle_H \quad \text{for all } u \in H.$$

(recall that  $\langle u, v \rangle_H = \int_I \langle u, v \rangle dt$ ). We will use the notation  $Q$  for the map  $(Q_1, \dots, Q_l) : H \rightarrow H \otimes \mathbb{R}^l$ . Also, given a covector  $\omega \in \mathbb{R}^{l*}$  we will denote by  $\omega q, \omega Q, \omega A$  the compositions of  $\omega : \mathbb{R}^l \rightarrow \mathbb{R}$  respectively with  $q, Q$  and  $A$ .

In the sequel we will need to expand a control  $u \in H$  into its Fourier series: we will write  $u = \sum_{k \in \mathbb{N}_0} U_k \frac{1}{\sqrt{\pi}} \cos kt + V_k \frac{1}{\sqrt{\pi}} \sin kt$  where  $U_k, V_k \in \Delta$ ; the constant term is zero because of part (a) of Proposition 2 (mean zero condition).

**Proposition 3.** Let  $T_k$  be the subspace of  $H$  with “wave number”  $k$ , namely

$$T_k = \Delta \otimes \text{span}\{\cos kt, \sin kt\}.$$

Then we have the following:

- (a)  $H = \bigoplus_{k \geq 1} T_k$ , and the sum is orthogonal with respect to the scalar product;
- (b) For every  $\omega \in \mathbb{R}^{l^*}$  we have  $\omega Q T_k \subset T_k$  (i.e. each subspace  $T_k$  is invariant by  $\omega Q$ );
- (c) Consider the orthonormal basis  $\{e_i \otimes \frac{1}{\sqrt{\pi}} \cos kt, e_i \otimes \frac{1}{\sqrt{\pi}} \sin kt\}_{i=1}^d$  for  $T_k$ ; in this basis the matrix associated to  $\omega Q|_{T_k}$  is:

$$\frac{1}{k}(\omega P) \doteq \frac{1}{k} \begin{pmatrix} 0 & \frac{1}{2}\omega A \\ -\frac{1}{2}\omega A & 0 \end{pmatrix}.$$

*Proof.* Point (a) is just Fourier decomposition theorem;  $k \geq 1$  expresses the mean zero condition. For the other two points, let us consider  $u \in T_n$  and  $v \in H$ , with Fourier series respectively  $u = U \frac{1}{\sqrt{\pi}} \cos kt + V \frac{1}{\sqrt{\pi}} \sin kt$  and  $v = \sum_{n \geq 1} U_n \frac{1}{\sqrt{\pi}} \cos nt + V_n \frac{1}{\sqrt{\pi}} \sin nt$ . By a direct computation we have

$$\begin{aligned} \langle u, \omega Q v \rangle_H &= \int_0^{2\pi} \left\langle \int_0^t U \frac{1}{\sqrt{\pi}} \cos k\tau + V \frac{1}{\sqrt{\pi}} \sin k\tau \, d\tau, \frac{1}{2}\omega A v \right\rangle dt = \\ &= \sum_{n \geq 1} \int_0^{2\pi} \frac{1}{k} \left\langle U \frac{1}{\sqrt{\pi}} \sin kt - V \frac{1}{\sqrt{\pi}} \cos kt, \frac{1}{2}\omega A \left( U_n \frac{1}{\sqrt{\pi}} \cos nt + V_n \frac{1}{\sqrt{\pi}} \sin nt \right) \right\rangle dt = \\ &= \frac{1}{k} \int_0^{2\pi} -\frac{1}{\pi} (\cos kt)^2 \left\langle V, \frac{1}{2}\omega A U_k \right\rangle + \frac{1}{\pi} (\sin kt)^2 \left\langle U, \frac{1}{2}\omega A V_k \right\rangle dt = \\ &= \frac{1}{k} \left( -\left\langle V, \frac{1}{2}\omega A U_k \right\rangle + \left\langle U, \frac{1}{2}\omega A V_k \right\rangle \right), \end{aligned}$$

where the equality between second and third row holds because the only non-zero integrals of products of sines/cosines are  $\int_0^{2\pi} \frac{1}{\pi} (\cos kt)^2 dt = \int_0^{2\pi} \frac{1}{\pi} (\sin kt)^2 dt = 1$ .  $\square$

*Remark 2.* We notice that for every  $\omega \in (\Delta^2)^*$  the operator  $\omega Q$  is compact. Indeed, it is the limit of a converging series of operators with finite-dimensional image:

$$S_n = \sum_{i=1}^n \omega Q|_{T_i}.$$

Let us prove that the operator norm of  $\omega Q - \omega S_n$  goes to zero. Given a norm one  $v = \sum_{k \geq 1} v_k$ , we have:

$$\begin{aligned} \|(\omega Q - \omega S_n)v\|^2 &= \sum_{k \geq n+1} \|\omega Q|_{T_k} v\|^2 \leq \sum_{k \geq n+1} \frac{4\|\omega P\|_{\text{op}}^2 \|v_k\|^2}{k^2} \\ &\leq \frac{4\|\omega P\|_{\text{op}}^2}{(n+1)^2} \sum_{k \geq n+1} \|v_k\|^2 \leq \frac{4\|\omega P\|_{\text{op}}^2}{(n+1)^2}. \end{aligned}$$

In particular, taking square roots,  $\|(\omega Q - \omega S_n)v\| \leq \frac{4\|\omega P\|_{\text{op}}}{(n+1)}$ , i.e.  $\|\omega Q - \omega S_n\|_{\text{op}} \rightarrow 0$

We conclude this chapter by describing the spectrum of the operator  $\omega Q$ . Given  $\omega Q$  we consider as above the skew-symmetric matrix  $\omega A$ , which can be put in canonical form as a block matrix of the form  $\text{Diag}(\alpha_1(\omega)J_2, \dots, \alpha_m(\omega)J_2, 0_n)$ , where  $J_2 \in \mathfrak{so}(2)$  is the standard symplectic matrix and  $0_n$  is the  $n \times n$  zero matrix. More precisely on  $\Delta$  we can find an orthonormal basis  $\{X_i, Y_i, Z_j, i = 1, \dots, m, j = 1, \dots, n\}$  for suitable  $m, n \in \mathbb{N}$  satisfying  $2m + n = d$  such that:

$$\omega A X_i = -\alpha_i(\omega) Y_i, \quad \omega A Y_i = \alpha_i(\omega) X_i, \quad \omega A Z_j = 0.$$

Let us consider now the operator  $\omega Q$  restricted to  $T_k$ . Using the basis for  $\Delta$  defined above we get the orthogonal basis

$$\left\{ \begin{pmatrix} X_i \\ Y_i \end{pmatrix}, \begin{pmatrix} -Y_i \\ X_i \end{pmatrix}, \begin{pmatrix} X_i \\ -Y_i \end{pmatrix}, \begin{pmatrix} Y_i \\ X_i \end{pmatrix}, \begin{pmatrix} Z_j \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ Z_j \end{pmatrix} \right\}, \quad i = 1, \dots, m, \quad j = 1, \dots, n,$$

with eigenvalues  $\frac{\alpha_i(\omega)}{k}, \frac{\alpha_i(\omega)}{k}, -\frac{\alpha_i(\omega)}{k}, -\frac{\alpha_i(\omega)}{k}, 0, 0$  respectively. Thus we have proved:

**Proposition 4.** *The non-zero eigenvalues of the operator  $\omega Q$  are  $\pm \frac{\alpha_i(\omega)}{k}$  with multiplicity two, where  $\alpha_i(\omega)$  are the coefficients of the canonical form of  $\omega A$  and  $k \in \mathbb{N}_0$ .*

### 3. GEODESICS

**3.1. The general structure of geodesics.** In this section we introduce the *admissible path space*; given a point  $p \in G$  it is defined as:

$$\Omega_p = \{\text{admissible curves starting at the origin and ending at } p\} = F^{-1}(p).$$

*Remark 3.* It will be useful for us to set  $\Sigma_1 = \overline{\Sigma'_1}$ . Since the euclidean closure is smaller than the Zariski closure and dimension of a semialgebraic set is preserved after taking its Zariski closure, then  $\Sigma_1$  is *closed* and contained in a semialgebraic set of codimension one. Points in the complement of  $\Sigma_1$  are an open dense set of regular values of  $q = F|_H$ .

Since in the space  $\Omega$  of admissible paths we are allowed to compute velocities and their lengths, we define the *Energy* functional:

$$J : \Omega \longrightarrow \mathbb{R},$$

by associating (as in Riemannian geometry) to each curve  $\gamma$  the integral  $\int_0^{2\pi} \|\dot{\gamma}(t)\|^2 dt$ .

In the case  $p$  is a regular value of the End-point map  $F$ , then a critical point of  $J|_{F^{-1}(p)}$  is called a *normal geodesic*.

Theorem 9 allows then to study the structure of (normal) geodesics whose endpoint is the generic  $p$ ; in fact they are by definition critical points of the restriction of  $J$  to  $\Omega_p$  and in the case the latter is a Hilbert manifold are defined by the Lagrange multipliers rule. Using the above coordinates  $G \simeq \mathbb{R}^d \oplus \mathbb{R}^l$  we can decompose a vector  $\lambda \in T^*G$  as  $\lambda = \eta + \omega$ , where  $\eta$  is the ‘‘horizontal’’ part and  $\omega \in (\Delta^2)^*$  is the ‘‘vertical’’ one. We have the following proposition.

**Proposition 5.** *Let  $u$  be the control associated to a geodesic with Lagrange multiplier  $\lambda \in T^*G$  (i.e.  $\lambda d_u F = d_u J$ ). Then:*

$$u(t) = e^{-(\omega A)t} u_0 \quad \text{and} \quad 2\eta = (e^{-2\pi\omega A} + \mathbb{1})u_0.$$

*Proof.* A simple computation using (11) in  $\lambda d_u F = d_u J$  gives:

$$u(t) = \eta - (\omega A) \int_0^t u(s) ds + \frac{1}{2} \int_0^{2\pi} (\omega A) u(s) ds.$$

Differentiating the above equation provides  $\dot{u} = -(\omega A)u$ , which in turn implies  $u(t) = e^{-(\omega A)t} u_0$ . Substituting the explicit expression  $u(t) = e^{-(\omega A)t} u_0$  into the same equation and evaluating at zero gives  $u_0 = \eta - \frac{1}{2}u(2\pi) + \frac{1}{2}u_0$ .  $\square$

If  $p$  is not a vertical point and the rank is sufficiently big ( $d > l$ ), then the number of geodesics joining the origin to  $p$  is bounded. In order to prove this statement we need a preliminary lemma.

**Lemma 6.** *For the generic point  $p \in G$ , the set of geodesics starting at the origin and ending at  $p$  is discrete.*

*Proof.* We can parametrize geodesics with their initial covector  $\omega$  and the initial velocity  $u_0$ ; in this way we obtain a smooth map:

$$f : \mathbb{R}^l \times \mathbb{R}^d \rightarrow G$$

defined by:

$$(\omega, u_0) \mapsto \left( \int_I e^{-t\omega A} u_0 dt, \int_I \left\langle \int_0^t e^{-s\omega A} u_0 ds, A e^{-t\omega A} u_0 \right\rangle dt \right).$$

If  $p$  is a regular value of  $f$  (and the set of such  $p$  is a residual set) then  $f^{-1}(p)$  is a submanifold of  $\mathbb{R}^{d+l}$  of dimension zero (possibly noncompact).  $\square$

**Proposition 7.** *Assume that  $d > l$ . Then for the generic  $p \notin \Delta^2$  there is a finite number of geodesics between  $e$  and  $p$ .*

*Proof.* We already know by the previous lemma that the set of geodesics (i.e. pairs  $(\omega, u_0)$  such that  $f(\omega, u_0) = p$ ) is discrete for the generic  $p$ . We will exclude that the set of possible Lagrange multipliers is unbounded: this will imply (see below) that the set of initial velocities is bounded as well, hence the set of geodesics to  $p$  is a discrete set in compact, i.e. it is finite.

If  $\omega$  is a Lagrange Multiplier, we can always choose a basis for  $\Delta$  such that the matrix  $\omega A$  appears in canonical form: so we get  $k$  subspaces of  $\Delta$  of dimension two on which the matrix  $\omega A$  is of the form  $\alpha_i J_2$  (where  $J_2$  is the standard  $2 \times 2$  symplectic matrix on the  $i$ -th eigenspace and  $\alpha_i > 0$ ), for  $i = 1, \dots, k$ ; for every subspace we take the component  $u_0^i$  of the initial velocity  $u_0$ . Since the eigenspaces are orthogonal, the computations for the end-point of the geodesic can be performed separately: thus we split the horizontal part of the end-point  $p_\Delta$  into the components of the eigenspaces of  $\omega A$ , say  $p_\Delta^i$ . These components are:

$$p_\Delta^i = \int_0^{2\pi} e^{-t\omega A} u_0^i dt = \left[ \frac{1}{\alpha_j} J_2 e^{-t\alpha_j J_2} u_0^i \right]_0^{2\pi} = \frac{1}{\alpha_j} J_2 (e^{-2\pi\alpha_j J_2} u_0^i - u_0^i),$$

on the eigenspaces of  $\omega A$ . The norm squared of the component  $p_\Delta^i$  is given by:

$$\begin{aligned} \|p_\Delta^i\|^2 &= \frac{1}{\alpha_i^2} \langle J_2 (e^{-2\pi\alpha_j J_2} u_0^i - u_0^i), J_2 (e^{-2\pi\alpha_j J_2} u_0^i - u_0^i) \rangle = \\ &= \frac{2}{\alpha_i^2} (\|u_0^i\|^2 - \langle u_0^i, e^{-2\pi\alpha_i J_2} u_0^i \rangle) = \frac{2}{\alpha_i^2} \|u_0^i\|^2 (1 - \cos(2\pi\alpha_i)) \end{aligned}$$

The vertical part is more complicated, but we only need to compute its value in the direction of the Lagrange multiplier  $\omega$ :

$$\begin{aligned} \omega(p_{\Delta^2}) &= \frac{1}{2} \int_0^{2\pi} \left\langle \int_0^t e^{-\omega A s} u_0 ds, \omega A e^{-\omega A t} u_0 \right\rangle dt = \\ &= \frac{1}{2} \int_0^{2\pi} \left\langle \int_0^t -\omega A e^{-\omega A s} u_0 ds, e^{-\omega A t} u_0 \right\rangle dt = \\ &= \frac{1}{2} \int_0^{2\pi} \langle e^{-\omega A t} u_0 - u_0, e^{-\omega A t} u_0 \rangle dt = \pi \|u_0\|^2 - \frac{1}{2} \langle u_0, p_\Delta \rangle. \end{aligned}$$

The equation

$$(13) \quad \omega(p_{\Delta^2}) = \pi \|u_0\|^2 - \frac{1}{2} \langle u_0, p_\Delta \rangle$$

is important at first because it tells that it is enough to prove that the set of Lagrange multipliers is bounded. Indeed, we have

$$\|\omega\| \|p_{\Delta^2}\| \geq \pi \|u_0\|^2 - \frac{1}{2} \langle u_0, p_\Delta \rangle;$$

if the set of Lagrange multipliers is bounded, the set of initial velocities cannot be unbounded otherwise the second term of the inequality would diverge while being limited by a constant.

In order to prove that the set of Lagrange multipliers of the generic end-point is bounded, we suppose on the contrary that we have a sequence  $\omega_n$  of Lagrange multipliers; this sequence is forced to diverge since the set of the initial data  $(\omega, u_0)$  for the generic end-point  $p$  is discrete. Up to subsequences, we may assume that the normalized Lagrange multipliers  $\hat{\omega}_n = \omega_n / \|\omega_n\|$  converge to a covector  $\lambda$ ; moreover we can assume the rescaled eigenvalues  $\hat{\alpha}_{i,n} = \alpha_{i,n} / \|\omega_n\|$  and the corresponding eigenspaces converge. The first one is true since every eigenvalue is bounded by the norm of the matrix which is 1 by definition; the second one follows from the fact that the set of changes of basis is the orthogonal group which is compact. For every sequence of eigenvalues  $\hat{\alpha}_{i,n}$  that doesn't converge to 0 we find a positive real number  $c$  such that  $c\omega_n \leq \alpha_{i,n}$ :



if we split the second term of the equation (13) into its components given by the eigenspaces of  $\omega A$ :

$$\omega(p_{\Delta^2}) = \sum_{i=1}^k \pi \|u_0^i\|^2 - \frac{1}{2} \langle u_0^i, p_{\Delta}^i \rangle.$$

Therefore we have:

$$\|p_{\Delta^2}\| \geq \hat{\omega}_n(p_{\Delta^2}) \geq \pi \frac{\|u_{0,n}^i\|^2}{\|\omega_n\|} - \frac{\langle u_{0,n}^i, p_{\Delta} \rangle}{\|\omega_n\|}$$

and if  $\|u_{0,n}^i\|$  diverges, the last term of the previous inequality is asymptotic to its first addendum, which turns out to be bounded. Then for the corresponding component of the horizontal part of the end-point we have that:

$$\|p_{\Delta,n}^i\|^2 = \frac{2}{\alpha_{i,n}^2} \|u_{0,n}^i\|^2 (1 - \cos(2\pi\alpha_{i,n})) \leq 2 \frac{\|u_{0,n}^i\|^2}{\|\omega_n\|^2}.$$

If  $\|u_{0,n}^i\|$  is bounded the last term of the inequality converges to 0; if it diverges we have that

$$\|p_{\Delta,n}^i\|^2 \leq 2 \frac{\|u_{0,n}^i\|^2}{\|\omega_n\|^2} \leq \text{const.} \frac{1}{\|\omega_n\|}$$

which again converges to 0. So, in the limit, the components of the horizontal part of the end-point are orthogonal to the eigenspaces of the limit matrix  $\lambda A$  with non-zero eigenvalues.

Now, if the matrix  $\lambda A$  is not singular the horizontal part has to be zero contradicting the hypothesis; thus in order to end the argument, we will prove that the horizontal part of the generic point does not belong to the kernel of any degenerate matrix in the span  $W$ .

Let us consider the semialgebraic set of the couples of degenerate matrices together with the vectors in their kernels, namely

$$X \doteq \{(\eta, x) \in S^{l-1}(W) \times \Delta \mid \eta Ax = 0, \dim \ker \eta A \geq 2\}$$

stratified in subsets with constant dimension of the kernel. Assuming that the Carnot algebra structure defined by  $W$  is transversal to every subset of the stratification described in the Appendix A (condition satisfied by the generic  $W$ ), the stratum  $W_k$  of the matrices of  $k$ -dimensional kernel has codimension  $k \frac{k-1}{2}$ . Therefore the semialgebraic set  $X$  has dimension not greater than  $\max_{k \geq 2} l - 1 - k \frac{k-1}{2} + k \leq l$ . If we project  $X$  on  $\Delta$  we get the set  $\tilde{X}$  of points belonging to the kernel of a degenerate matrix of  $W$ : this is still a semialgebraic set and the projection does not increase the dimension. In particular the set

$$Y = \{(p_{\Delta}, p_{\Delta^2}) \in G \mid p_{\Delta} \in \ker \eta A \text{ for some } \eta \neq 0\}$$

has dimension less than  $2l$ , which is smaller than  $d + l = \dim(G)$  under the assumption  $d > l$ . This concludes the proof.  $\square$

**Corollary 8.** *If  $p \notin \Delta^2$  the homology of  $\Omega_p \cap \{J \leq s\}$  stabilizes as  $s \rightarrow \infty$ .*

**3.2. Geodesics ending at  $p \in \Delta^2$ .** In this section we study in more detail the case  $p \in \Delta^2$ . To start with, we prove that for the generic choice of  $p \in \Delta^2$  the set  $\Omega_p$  is a nice object, more precisely we have the following theorem.

**Theorem 9.** *For the generic choice of  $W \subset \mathfrak{so}(d)$  and a generic  $p \in \Delta^2$  the topological space  $\Omega_p$  is a Hilbert manifold.*

*Proof.* We will prove that the set of critical values of  $q = F|_{F^{-1}(\Delta^2)}$  is contained in a semialgebraic subset of codimension one, from this the conclusion follows.

We first notice that the condition for  $p \in \Delta^2$  to be a critical value of  $q$  is that there exist  $u \in H$  and  $\omega \in (\Delta^2)^*$  such that  $q(u) = p$  and  $dq_u = 2\omega Qu = 0$ . In particular if  $u = \sum_{k=1}^{\infty} u_k$ , where  $u_k \in T_k \simeq \mathbb{R}^{2d}$ , then  $q(u) = \sum_{k=1}^{\infty} \langle u_k, \frac{1}{k}(\omega P)u_k \rangle$  and the condition that  $\omega Qu = 0$ , by invariance of the spaces  $T_k$ , reads  $\omega P u_k = 0$  for all  $k \geq 1$ .

Let us consider the stratification  $\mathfrak{so}(d) = \coprod S_r$ , where  $S_r$  is the set of matrices with constant rank  $r$ . Each  $S_r$  is smooth and over it we have the smooth bundle  $K_r = \{(A, v) \in \mathfrak{so}(d) \times \mathbb{R}^n \mid Av = 0\}$ . Since this stratification is homogeneous, then the generic  $W$  is transversal to all strata and we have an induced stratification:

$$W = \coprod_{r=0}^d W_r, \quad W_r = S_r \cap W.$$

Consider now the vector bundle  $K|_{W_r}$  over  $W_r$  (the restriction of  $K_r$ ); notice that for every  $\omega A \in W$  we have  $\ker(\omega P) = \{(x, y) \in \mathbb{R}^{2d} \mid \omega Ax = \omega Ay = 0\}$ . In particular a smooth section of  $K|_{W_r}$  produces also a smooth section of  $\{(\omega, z) \in W_r \times \mathbb{R}^{2d} \mid \omega Pz = 0\}$  over  $W_r$ .

We notice now the following interesting property: if  $p \in \Delta^2$  is a critical point for  $q$  with Lagrange multiplier  $\omega \in W_r$ , then  $p \in (T_\omega W_r)^\perp$ . In fact for every  $k \geq 1$  let us consider a smooth curve  $\omega(t)$  in  $W_r$  with  $\omega(0) = \omega$ ,  $\dot{\omega}(0) = \eta \in T_\omega W_r$  and a smooth  $z(t) \in \ker(\omega(t)\omega P)$  with  $z(0) = u_k$  such that

$$\omega(t)\omega Pz(t) = 0,$$

(the existence of such a smooth  $z(t)$  follows from the above discussion). Then deriving the above equation we get  $\dot{\omega}(0)\omega Pz(0) + \omega(0)\omega P\dot{z}(0) = 0$  and considering the scalar product with  $z$  gives:

$$\langle z, \dot{\omega}(0)\omega Pz \rangle = \dot{\omega}(0)(q(u_k)) = 0$$

which tells  $\eta(p) = \sum_{k=1}^{\infty} \eta(q(u_k))$  vanishes for every  $\eta$  in  $T_\omega W_r$ . We consider now the semialgebraic set:

$$\Sigma'_1 = \bigcup_{r=0}^d \left( \bigcup_{\omega \in W_r} (T_\omega W_r)^\perp \right).$$

Because of the above argument all critical points of  $q$  are contained in  $\Sigma'_1$  and we want to show this set is of dimension strictly less than  $l$ .

We first check that  $\Sigma'_1$  is indeed semialgebraic (and as a result we compute its dimension); being a finite union, it is enough to prove that each  $\bigcup_{\omega \in W_r} (T_\omega W_r)^\perp$  is semialgebraic. To this end consider  $T_r = \{(\omega, \eta) \in W_r \times W^* \mid (\eta \in T_\omega W_r)^\perp\}$ , which is clearly semialgebraic, and the semialgebraic projection  $\pi_2 : T_r \rightarrow W^*$  to the second factor. The image of  $\pi_2$  is semialgebraic and coincides with  $\bigcup_{\omega \in W_r} (T_\omega W_r)^\perp$ .

Now, each  $\bigcup_{\omega \in W_r} (T_\omega W_r)^\perp$  has dimension less than  $\dim(W_r) + l - \dim(W_r) - 1 \leq l - 1$ , where the  $-1$  comes from the fact that by homogeneity  $(T_{\frac{\omega}{|\omega|}} W_r)^\perp = (T_\omega W_r)^\perp$  (notice in particular that the stratum of maximal dimension is open and produces only the zero, since the orthogonal complement of its tangent space is the zero only).

In particular the dimension of  $\Sigma'_1$  is strictly less than  $l$  and the generic  $p \in \Delta^2$  is a regular value of  $q$ .  $\square$

Following up the discussion after Proposition 5, we see that in the case the final point of  $\gamma_u$  is in  $\Delta^2$ , which we know it is equivalent to  $\int_I u = 0$ , we can apply the Lagrange multiplier rule to the map  $q$ . More precisely  $u$  is the control associated to a curve which is a geodesic with endpoint  $p \in \Delta^2$  if:

$$q(u) = p \quad \text{and there exists } \omega \text{ such that } \omega Qu = u.$$

The covector  $\omega \in (\Delta^2)^*$  is called the *Lagrange multiplier* associated to  $u$ . Using this remark we see that  $u$  is a geodesics with Lagrange multiplier  $\omega$  iff:

$$\langle u, v \rangle_H = \langle \omega Qu, v \rangle_H \quad \text{for all } v \in H$$

(here the final point is not specified, i.e. we are considering all possible geodesics with Lagrange multiplier  $\omega$ ; the final point is recovered by simply applying the expression given in (11) to  $u$ ). Thus in particular we have that for all  $v$  in  $H = \{\int v = 0\}$ :

$$\int_0^{2\pi} \langle u(t), v(t) \rangle dt = - \int_0^{2\pi} \langle \omega AU(t), v(t) \rangle dt = 0,$$

where  $U(t) = \int_0^t u(s)ds$ . The previous condition tells that  $u + \omega AU$  is a constant function, or equivalently that  $\dot{u} = -\omega Au$ . This implies that  $u$  must be of the form:

$$u(t) = e^{-t(\omega A)}u_0,$$

and since  $u \in H$ , then  $u_0$  must be in the *integer* eigenspace of  $i\omega A$ , i.e.:

$$u_0 = e^{-2\pi\omega A}u_0.$$

In particular notice that in this case the complete Lagrange multiplier (i.e. the one arising by using the map  $F$ , as in Proposition 5), is  $\lambda = (u_0, \omega)$ .

We collect the result for a geodesic ending at  $p \in \Delta^2$  in a lemma.

**Lemma 10.** *Let  $u$  be the control associated to a geodesic whose final point is in  $\Delta^2$  with Lagrange multiplier  $\omega$ . Then:*

$$u(t) = e^{-t\omega A}u_0 \quad \text{with} \quad u_0 = e^{-2\pi\omega A}u_0.$$

Motivated by the previous lemma, for every  $\omega \in W$  we define:

$$E(\omega) = \{v \in \mathbb{R}^d \mid e^{-2\pi\omega A}v = v, v \notin \ker(\omega A)\}.$$

Thus  $E(\omega)$  is the set of possible initial data for *non constant* geodesics with Lagrange multiplier  $\omega$ . In particular we see that in order to have a nonzero initial datum the matrix  $i\omega A$  must have nonzero integer eigenvalues, thus the set of all possible Lagrange multipliers coincides with the set:

$$\Lambda = \{\omega \in (\Delta^2)^* \mid \det(\omega A - in\mathbb{1}) = 0 \quad \text{for some } n \in \mathbb{N}_0\}.$$

Notice that  $\Lambda$  is *not* an algebraic (or a semialgebraic set): it is indeed given by the infinite union of algebraic sets  $\Lambda_n = \{\det(\omega A - in\mathbb{1}) = 0\}$ . However  $\Lambda$  is locally algebraic: if we intersect it with a ball, then only a finite number of  $\Lambda_n$  show up.

We discuss now in more detail the structure of the set:

$$E = \{(\omega, v) \in W \times \mathbb{R}^d \mid v \in E(\omega)\}.$$

As for  $\Lambda$ , this set is *not* semialgebraic, although if we take the “restriction”  $E|_B$  to a compact semialgebraic set  $B$ , i.e. we only allow  $\omega$  to vary on a compact semialgebraic set  $B \subset W$ , then  $E|_B$  becomes semialgebraic.

First for every  $\omega$  let us consider the canonical skew-symmetric form of  $\omega A$ :

$$M(\omega)^T(\omega A)M(\omega) = \text{Diag}(\alpha_1(\omega)J_2, \dots, \alpha_s(\omega)J_2, 0, \dots, 0)$$

where  $M(\omega)$  is an orthogonal matrix,  $\alpha_1(\omega), \dots, \alpha_s(\omega)$  are the *positive nonzero* eigenvalues of  $i\omega A$  and  $J_2 \in \mathfrak{so}(2)$  is the canonical symplectic matrix. Thus  $\mathbb{R}^d$  decomposes as the orthogonal sum  $\mathbb{R}^d = V_1 \oplus \dots \oplus V_s \oplus K$ , where the  $V_i$ s are the coordinate two planes and  $K$  is the vector space of the last  $d - 2s$  coordinates. Using this notation we set  $V_i(\omega) = M(\omega)V_i$ : it is the invariant subspace of  $\omega A$  associated to the eigenvalue  $\alpha_i(\omega)$ . In particular we see that:

$$E(\omega) = \bigoplus_{\alpha_i(\omega) \in \mathbb{N}_0} V_i(\omega).$$

Associating to each  $v \in E(\omega)$  the control  $e^{t\omega A}v$  defines a linear injection of  $E(\omega)$  into  $H = \bigoplus_{k \geq 1} T_k$ ; in particular the previous curve admits the Fourier series decomposition

$$(14) \quad e^{-t\omega A}v = \sum_{k \geq 1} X_k(v) \frac{1}{\sqrt{\pi}} \cos kt - Y_k(v) \frac{1}{\sqrt{\pi}} \sin kt$$

and we denote by  $\phi_k$  the linear map  $v \mapsto (X_k(v), Y_k(v))$  (the  $k$ -th component of the Fourier series of  $e^{t\omega A}v$  written in coordinates  $T_k \simeq \mathbb{R}^{2d}$ ).

Let now  $v \in V_i(\omega)$  with  $\alpha_i(\omega) = k \in \mathbb{N}$ ; in order to get the expression for  $\phi_k(v)$  we compute the Taylor series of  $e^{\omega A t} v$ . We have:

$$\begin{aligned} e^{-t\omega A} v &= \left( \sum_{n \geq 0} \frac{(\omega A)^n t^n}{n!} \right) v = \left( \sum_{m \geq 0} \frac{(\omega A)^{2l} t^{2l}}{(2l)!} \right) v - \left( \sum_{m \geq 0} \frac{(\omega A)^{2l+1} t^{2l+1}}{(2l+1)!} \right) v = \\ &= \left( \sum_{m \geq 0} \frac{(-1)^l k^{2l} t^{2l}}{(2l)!} \right) v - \left( \sum_{m \geq 0} \frac{(-1)^l k^{2l+1} t^{2l+1}}{(2l+1)!} \right) \frac{\omega A}{k} v = \\ &= v \cos kt - \frac{\omega A}{k} v \sin kt \end{aligned}$$

where in the second line we have used the fact that  $(\omega A)^2 v = -k^2 v$  (being  $V_i(\omega)$  the space associated to the eigenvalue  $\alpha_i(\omega) = k$ ). This computation implies that:

$$\phi_k(v) = \sqrt{\pi} \begin{pmatrix} v \\ -\frac{\omega A}{k} v \end{pmatrix}.$$

Notice that the same construction can be performed using the linear immersion  $v \mapsto e^{t\omega A} v$ , which gives the above control with *backward* time; the Lagrange multiplier for the corresponding geodesic is  $-\omega$  and the final point is  $-q(e^{-t\omega A} v)$ .

Slightly abusing of notation, we will still denote by  $q$  the map obtained by composing the endpoint map with the linear immersion  $E(\omega) \hookrightarrow H$ .

We recall from the Appendix (see Proposition A.1) that the Lie algebra  $\mathfrak{so}(d)$  of skew-symmetric matrices of size  $d$  is stratified by the sets  $\Gamma_{k|m_1, \dots, m_r}$  consisting of those matrices  $A$  satisfying:  $\dim \ker(A) = k$ ; the numbers  $m_1, \dots, m_r$  are natural non-increasing (they are the multiplicities in the positive spectrum of  $iA$ ). Each one of these strata is smooth and has codimension  $\sum_{i=1}^r (m_i^2 - 1) + \frac{k(k-1)}{2}$ . Since (by construction) this stratification is homogeneous, then a generic choice of the Carnot structure  $W \subset \mathfrak{so}(d)$  will be transversal to all of the strata and will inherit the stratification (in particular respecting the codimensions and smoothness).

To deal with integer eigenvalues we need to refine this stratification, unfortunately ending up with an infinite number of strata, but still with nice properties. More specifically for every  $r \geq 0$  we consider  $\vec{n} = (n_1, \dots, n_r) \in \mathbb{N}^r$  with distinct nonzero components and define the *semialgebraic* set  $\Gamma_{k|m_1, \dots, m_r|\vec{n}}$  as follows: we look at the nonzero components of  $\vec{n}$ , say  $n_{j_1}, \dots, n_{j_\nu}$ , and we take those matrices in  $\Gamma_{k|m_1, \dots, m_r}$  such that the eigenvalue with multiplicity  $m_{j_1}$  equals  $n_{j_1}$ , the one with multiplicity  $m_{j_2}$  equals  $n_{j_2}$ , and so on.

For example  $(0, 2, 2, 0) \in \mathbb{N}^4$  is not an admissible  $\vec{n}$  (since there are two equal nonzero entries); on the other hand if  $\vec{n} = (0, 1, 2, 0)$  then  $\Gamma_{k|m_1, m_2, m_3, m_4|\vec{n}}$  equals the set of all matrices in  $\mathfrak{so}(d)$  with multiplicities of the spectrum  $\{m_1, m_2, m_3, m_4\}$  and with one eigenvalue equal to  $i$  (the imaginary unit) with multiplicity  $m_2$  and another equal to  $2i$  and with multiplicity  $m_3$ .

This operation of fixing some eigenvalues to some integer numbers increases the codimension by  $\nu$  (the number of nonzero components of  $\vec{n}$ ).

Using this notation we see that one can stratify  $\Lambda$  as:

$$\Lambda = \coprod_r \coprod_{\{k, m_1, \dots, m_r\}} \coprod_{\{\vec{n} \in \mathbb{N}^r \text{ admissible}\}} \Gamma_{k|m_1, \dots, m_r|\vec{n}} \cap W.$$

As we already noticed, this stratification is not finite even though each stratum is semialgebraic. Nevertheless if we intersect  $\Lambda$  with a compact ball  $B \subset W$  only a finite number of the above strata appear and we are locally semialgebraic.

The next lemma tells that for the generic choice of  $W \subset \mathfrak{so}(d)$  and a generic  $p \in \Delta^2$ , the Lagrange multipliers have simple integer spectrum.

**Lemma 11.** *For a generic Carnot group structure  $W \subset \mathfrak{so}(d)$ , the generic  $p \in \Delta^2$  is not the final point of a geodesic with lagrange multiplier  $\omega$  such that  $\omega A$  has multiple eigenvalues in  $i\mathbb{Z}$ .*

*Proof.* First we pick the structure  $W$  to be transversal to all strata of the first one of the above stratifications, the one using only the multiplicities in the spectrum (and we know such a property is generic). We stratify now the set  $\Lambda$  by intersecting it with the different strata  $\Gamma_{k|m_1, \dots, m_r}$ ; we are interested only in those strata for which there is at least a multiple integer eigenvalues and we refine the stratification to the above infinite one, by indexing with the admissible  $\vec{n} \in \mathbb{N}^r$ .

Thus we let  $\Lambda_{\{m_j \geq 2\}}$  be one stratum  $\Gamma_{k|m_1, \dots, m_r|\vec{n}}$  such that  $m_j \geq 2$  for at least one index  $j$  with  $n_j \neq 0$ . Each  $\Lambda_{\{m_j \geq 2\}}$  obtained in this way has codimension:

$$\text{codim}_W \Lambda_{\{m_j \geq 2\}} = \sum_{i=1}^r (m_i^2 - 1) + \frac{k(k-1)}{2} + \nu.$$

We consider now as above the set  $E = \{(\omega, v) \in \Lambda \times \mathbb{R}^d \mid e^{2\pi\omega A}v = v, v \notin \ker(\omega A)\}$ . Over each stratum  $\Lambda_{\{m_j \geq 2\}}$  the set  $E|_{\Lambda_{\{m_j \geq 2\}}}$  is a smooth vector bundle (it is the restriction to  $\Lambda_{\{m_j \geq 2\}}$ , which is smooth, of a smooth vector bundle); moreover  $E|_{\Lambda_{\{m_j \geq 2\}}}$  is semialgebraic as well (here the vector  $\vec{n}$  is fixed).

Consider the smooth map:

$$f : E|_{\Lambda_{\{m_j \geq 2\}}} \rightarrow \Delta^2$$

defined by  $(\omega, v) \mapsto q(v)$ , where  $q(v)$  is the final point of the geodesic associated to the control  $v(t) = e^{t\omega A}v$ . We compute the rank of the differential of  $f$  and show that the assumption  $m_j \geq 2$  implies this rank is less than  $l-1$ ; in particular the image of  $f$  has measure zero. Since the set of final points of geodesics with Lagrange multipliers with multiple eigenvalues in  $i\mathbb{Z}$  is the countable union of the images of the different  $f$  obtained as  $\vec{n}$  varies over  $\mathbb{N}^r$ , the result follows.

The differential of  $f$  restricted to the base  $\Lambda_{\{m_j \geq 2\}}$  has rank smaller than the dimension of  $\Lambda_{\{m_j \geq 2\}}$ , which is  $l - \sum_{i=1}^r (m_i^2 - 1) - \frac{k(k-1)}{2} - \nu$ . For the rank of  $f$  restricted to the fibers we argue as follows. For every  $\omega \in \Lambda_{\{m_j \geq 2\}}$  we consider the invariant subspaces of  $\omega A$ ; for each natural nonzero eigenvalue  $\lambda_j(\omega)$  of  $i\omega A$  we find an invariant space  $V_j(\omega)$  (the real part of the  $\lambda_j(\omega)$ -eigenspace of  $i\omega A$ ) of dimension  $2\mu_j$ , twice the multiplicity of  $\lambda_j(\omega)$ ; let's call  $I \subset \{1, \dots, r\}$  the index set for such spaces  $V_j(\omega)$  (notice that  $I = \{j_1, \dots, j_\nu\}$ ).

The restriction of  $f$  to each such  $V_j(\omega)$  maps  $\mu_j$  unit circles (lying on distinct orthogonal planes) to a point, in particular the dimension of the kernel of the differential of  $f$  on each  $V_j(\omega)$  is at least  $\mu_j$ . Since the dimension of  $E(\omega)$  is  $2 \sum_{j \in I} \mu_j$ , we see that the rank of the differential of  $f$  on the fibers is at most  $\sum_{j \in I} \mu_j$ .

In particular we can bound the rank of the differential of  $f$  as:

$$\begin{aligned} \text{rk}(df) &\leq l - \nu - \sum_{j=1}^r (m_j^2 - 1) - \frac{k(k-1)}{2} + \sum_{j \in I} \mu_j \\ &\leq l - \nu - \sum_{j \in I} (m_j^2 - 1 - m_j) - \frac{k(k-1)}{2} \\ &\leq l - \nu - \sum_{j \in I, m_j \geq 2} (m_j^2 - 1 - m_j) - \sum_{j \in I, m_j = 1} (m_j^2 - 1 - m_j) \\ &\leq l - \nu - \sum_{j \in I, m_j \geq 2} 1 - \sum_{j \in I, m_j = 1} (-1) \\ &\leq l - \nu - 1 + (\nu - 1) < l - 1. \end{aligned}$$

□

We define now the set  $\Sigma_2 \subset \Delta^2$  to be the union of the various  $f(E|_{\Lambda_{k|m_1, \dots, m_r, \vec{n}}})$  where  $\vec{n} = (n_1, \dots, n_r) \in \mathbb{N}^r$  is admissible and  $m_j \geq 2$  for at least one index  $j$  with  $n_j \neq 0$ . The above lemma says that  $\Sigma_2$  is the countable union of semialgebraic sets of codimension at least 2 (in particular, for example,  $\Sigma_2$  has measure zero).

We study now what happens for a  $p \in \Delta^2 \setminus (\Sigma_1 \cup \Sigma_2)$  (because of the above argument such  $p$  is generic). Such a  $p$  has Lagrange multipliers  $\omega$  with simple spectrum, i.e.  $\omega A$  belongs to a stratum  $\Lambda_{k|m_1, \dots, m_r|\vec{n}}$  with all multiplicities equal to 1; for simplicity of notation we omit the string of multiplicities and denote such stratum simply by  $\Lambda_{\vec{n}}$ . In other words  $\Lambda_{\vec{n}}$  is one of the above strata where all eigenvalues are distinct and we have fixed  $\nu$  of them to be equal to  $in_{j_1}, \dots, in_{j_\nu}$  (the nonzero entries of  $\vec{n}$ ).

**Lemma 12.** *Let  $\omega \in \Lambda_{\vec{n}}$  and  $n_1, \dots, n_\nu$  be the nonzero eigenvalues of  $i\omega A$  in  $\mathbb{N}$ ; for  $j = 1, \dots, \nu$  let also  $V_j(\omega)$  be the (two dimensional) invariant subspace of  $\omega A$  associated to  $n_j$ . Then  $E(\omega)$  splits as the direct orthogonal sum:*

$$E(\omega) = \bigoplus_{j=1}^{\nu} V_j(\omega).$$

Moreover the image of  $q|_{V_j(\omega)}$  is a half line  $l_j^+(\omega)$  and:

$$\text{im}(q|_{E(\omega)}) = \text{cone}\{l_1^+(\omega), \dots, l_\nu^+(\omega)\}.$$

*Proof.* Recall that the space  $E(\omega)$  is defined to be  $\{v \in \mathbb{R}^d \mid e^{-2\pi\omega A}v = v, v \notin \ker(\omega A)\}$ ; the map that associates to a vector  $v \in E(\omega)$  the curve  $e^{-t\omega A}v$  defines an embedding of  $E(\omega)$  into  $H$  and if  $v \in V_j(\omega)$  then the resulting control must be a linear combination of  $\sin(n_j t)$  and  $\cos(n_j t)$ ; in particular  $V_{n_j}(\omega) \subset T_{n_j}$ . Since the  $T_k$  are pairwise orthogonal for each operator  $Q_1, \dots, Q_l$ , then decomposing  $v \in E(\omega)$  into its pieces  $v = v_1 + \dots + v_\nu$  with  $v_j \in V_j(\omega)$ , we get:

$$q(v) = q(v_1) + \dots + q(v_\nu),$$

which proves the image of  $q|_{E(\omega)}$  is the cone spanned by  $\{q(V_1(\omega)), \dots, q(V_\nu(\omega))\}$ ; the orthogonality of the  $V_j$  follows from the one of the  $T_{n_j}$ .

It remains to prove that the image of  $q|_{V_j(\omega)}$  is a half line. By assumption  $\omega$  belongs to a smooth stratum of codimension  $\nu$  in  $W$  and recalling the definition of  $\Lambda_{n_j} = \{\det(\omega A - in_j \mathbb{1}) = 0\}$ , we have that:

$$\Lambda_{\vec{n}} = \bigcap_{j=1}^{\nu} \Lambda_{n_j} \cap \Gamma_{k|1, \dots, 1}.$$

The bundle  $\bigsqcup_{\omega \in \Lambda_{n_j}} V_j(\omega)$  is smooth (being the restriction of a smooth bundle). In particular for every  $\eta \in T_\omega \Lambda_{n_j}$  there are curves  $\omega(t) \in \Lambda_{n_j}$  and  $v(t) \in V_j(\omega(t))$  such that  $\omega(0) = \omega$ ,  $\dot{\omega}(0) = \eta$  and  $v(0) = v$ . Deriving the equation  $\omega(t)\omega P v(t) = v(t)$  and taking inner product with  $v$  we get:

$$0 = \left\langle v, \frac{\dot{\omega}(0)\omega P}{k} v \right\rangle = \eta(q(v))$$

which tells the final point of the geodesic  $u$  associated to  $e^{\omega A t}v$  is orthogonal to  $T_\omega \Lambda_{n_j}$ , hence is it contained in a line. On the other hand since  $\omega$  is the Lagrange multiplier for the geodesic  $u$ , we get  $\omega(q(u)) = J(u) > 0$ , which concludes the proof.  $\square$

Everything now is ready for the proof of the Theorem that describes the structure of geodesics.

**Theorem 13.** *For a generic Carnot group structure  $W \subset \mathfrak{so}(d)$  and a generic  $p \in \Delta^2$  the set  $\Lambda(p)$  of Lagrange multipliers of geodesic whose final point is  $p$  is discrete. Moreover every  $\eta \in \Lambda(p)$  belongs to some  $\Lambda_{\vec{n}}$  and the set of all geodesics whose endpoint is  $p$  with Lagrange multiplier  $\eta$  is a compact manifold of dimension  $\nu \leq l$  ( $\nu$  is the number of nonzero entries of  $\vec{n}$ ) diffeomorphic to the torus  $\underbrace{S^1 \times \dots \times S^1}_{\nu \text{ times}}$ .*

*Proof.* By Lemma 11 we know that for the generic choice of  $W \subset \mathfrak{so}(d)$  the generic  $p \in \Delta^2$  is a final point only of geodesics with Lagrange multipliers in  $\Lambda_{\vec{n}}$  for some  $\vec{n} \in \mathbb{N}^{\lfloor d/2 \rfloor}$ .

Moreover for every  $\eta$  in  $\Lambda(p)$  the set of all geodesics with Lagrange multipliers  $\eta$  and final point  $p$  is the preimage of  $p$  under the map  $q : E|_{\Lambda(p)} \rightarrow \Delta^2$ .

For every admissible  $\vec{n} \in \mathbb{N}^{\lfloor d/2 \rfloor}$  we consider the semialgebraic set  $F_{\vec{n}}$  defined by:

$$F_{\vec{n}} = \{(\omega, p) \in \Lambda_{\vec{n}} \times \Delta^2 \mid p \in \text{im}(q|_{E(\omega)})\}$$

together with the semialgebraic map  $g : F_{\vec{n}} \rightarrow \Delta^2$  defined by  $(\omega, p) \mapsto p$ . Since each point  $(\omega, p)$  in  $F_{\vec{n}}$  has  $\omega$  in  $\Lambda_{\vec{n}}$ , then by Lemma 12 the dimension of  $F_{\vec{n}}$  is at most  $l$ . In fact  $\omega$  varies on a set of dimension  $l - \nu$  and the image of  $q|_{E(\omega)}$  is a cone of dimension at most  $\nu$ . Since  $F_{\vec{n}}$  is semialgebraic we stratify it as  $F_{\vec{n}} = \coprod_{j=1}^s F_{\vec{n},j}$ , where each stratum is smooth semialgebraic of dimension at most  $l$  (in fact here the index  $s$  depends on  $\vec{n}$  as well, but we omit this dependence to simplify notations). Notice that if  $(\omega, p)$  belongs to a stratum of maximal dimension  $l$ , then the cone  $q(E(\omega))$  must have maximal dimension  $\nu$  and  $p$  must be in its interior.

The restriction  $g_{\vec{n},j} = g|_{F_{\vec{n},j}}$  is smooth semialgebraic, thus by the semialgebraic Sard's lemma (see [5]) the set  $C_{\vec{n},j}$  of its critical values is a semialgebraic set of dimension at most  $l - 1$ . If  $p$  is not one of these critical values then  $g_{\vec{n},j}^{-1}(p)$  consists of isolated points if  $\dim(F_{\vec{n},j}) = l$ , and is empty otherwise.

We set  $\Sigma_3$  to be the union of the critical values of  $g_{\vec{n},j}$  ( $\vec{n}$  varies over  $\mathbb{N}^{\lfloor d/2 \rfloor}$  and  $j$  is the stratifying index for  $F_{\vec{n}}$  as above); such a union, being a countable union of semialgebraic set of dimension at most  $l - 1$  has measure zero, hence points belonging to its complement are generic.

On the other hand  $\Lambda(p)$  equals the union of the projections on  $\Lambda$  of the various  $g_{\vec{n},j}^{-1}(p)$ . If we intersect  $\Lambda$  with a compact ball  $B$ , we hit only a finite number of strata  $\Lambda_{\vec{n}}$  and  $\Lambda(p) \cap B$  is discrete; thus for a generic  $p$  the set  $\Lambda(p)$  is discrete set (possibly infinite).

From Lemma 12 we recall that  $q(E(\omega))$  is the cone spanned by the half-lines  $l_j^+(\omega) = q(E_j(\omega))$ : moreover  $p$  is the sum of nonzero vectors belonging to these half-lines,  $p = p_1 + \dots + p_\nu$  with  $p_j \in l_j^+(\omega)$ . Since the spaces  $E_j(\omega)$  are orthogonal with respect to the operators  $J, Q_1, \dots, Q_l$ , the condition  $q(u) = p$  with  $u \in E(\omega)$  can be split up as  $q(u_j) = p_j$  with  $u_j \in E_j(\omega)$ . The condition  $q(u_j) = p_j$  is equivalent to  $\omega q(u_j) = \omega(p_j)$  since for every covector  $\eta$  orthogonal to  $p_j$  the condition  $\eta q(u_j) = \eta(p_j)$  is automatically satisfied; on the other hand by the Lagrange Multiplier condition we have that  $\omega q(u_j) = J(u_j)$ , so that the condition is a positive definite one,  $J(u_j) = \omega(p_j)$ . This implies that every component  $u_j \in E_j(\omega)$  of the geodesic  $u$  going to  $p$  is constrained on a circle  $S^1 \in E_j(\omega)$ , from which follows that the critical manifold  $C_\omega$  is a  $\nu$ -dimensional torus.  $\square$

Up to now the genericity assumptions for  $p$  come from Theorem 9, Lemma 11 and Theorem 13: specifically we require  $p \notin \Sigma_1 \cup \Sigma_2 \cup \Sigma_3$  (where  $\Sigma_3$  is defined in the proof of Theorem 13).

**3.3. Morse-Bott theory.** Since critical points of the Energy functional  $J$  are far from being isolated (they arrange themselves into compact manifolds) we cannot apply Morse Theory in its standard version. What we need is a generalization of it called *Morse-Bott Theory*: it allows to prove the same results as for the ordinary theory if in the definitions nondegenerate critical points are replaced by *nondegenerate critical manifolds* (see [7, 12]; the basic definitions and results are recalled in the Appendix B).

**Theorem 14.** *For a generic Carnot Group structure  $W \subset \mathfrak{so}(d)$  and a generic point  $p \in \Delta^2$ , the Energy functional  $J$  restricted to  $\Omega_p$  is a Morse-Bott function.*

*Proof.* We assume all the genericity conditions of the previous theorems to be satisfied. We need to prove that we can possibly restrict the set of “good” final points  $p$  to a smaller (but still dense) set for which  $J|_{\Omega_p}$  is Morse-Bott.

Part (a) of the definition of a Morse-Bott function given in the Appendix B, immediately follows from Theorem 13.

We proceed to prove part (b). Let us take a Lagrange multiplier  $\eta$ , its corresponding critical manifold  $C_\eta$  and a point  $u \in C_\eta$ . Since we already know that the critical manifold is compact, it remains to show that the Hessian of the energy  $J$  is non-degenerate outside the tangent space to the critical manifold  $C_\eta$ . Since both  $q$  and  $J$  are quadratic, they coincide with their second

derivatives. By the Lagrange multiplier rule we get the expression for the Hessian:

$$\text{He}_u(q) = (d^2J - \eta D^2q_u)|_{T_u\Omega_p} = (J - \eta q)|_{T_u\Omega_p},$$

where we have  $T_u\Omega_p = \ker D_uq$ . Notice that to the quadratic form defined by the Hessian it corresponds the self-adjoint operator  $\mathbb{1} - \eta Q$ .

The Hessian is degenerate in the direction of  $v \in T_u\Omega_p$  if and only if:

$$\langle v - \eta Qv, x \rangle = 0 \quad \forall x \in T_u\Omega_p,$$

meaning that  $v - \eta Qv$  is orthogonal to  $T_u\Omega_p$ . Since the tangent space  $T_u\Omega_p$  is the orthogonal space to  $\text{span}\{Q_1u, \dots, Q_lu\}$ , then  $v - \eta Qv$  is a linear combination of the vectors  $Q_iu$ , namely  $\lambda_1 Q_1u + \dots + \lambda_l Q_lu$ .

We can eventually restate the degeneracy condition by the following equations:

$$(15) \quad \begin{cases} v - \eta Qv = \lambda Qu \\ \langle v, Qu \rangle = 0 \end{cases}$$

where  $\lambda = (\lambda_1, \dots, \lambda_l) \in (\Delta^2)^*$  as above.

If we take  $\lambda = 0$  we see that the degeneracy condition is satisfied by the vectors in:

$$E(\eta) \cap T_u\Omega_p = T_uC_\eta,$$

and we have to prove that for the generic choice of  $p$  the degeneracy equation (15) does not admit other solutions.

Let us consider the smooth manifold  $\Lambda_{\bar{n}}$  of Lagrange multipliers containing  $\eta$  (the definition of  $\Lambda_{\bar{n}}$  is given before Lemma 12). Let us call as before  $E_{\bar{n}}$  the fiber bundle with base space  $\Lambda_{\bar{n}}$  and fiber  $E(\omega)$  with  $\omega \in \Lambda_{\bar{n}}$  (see the above discussion).

The tangent space to  $E_{\bar{n}}$  at  $(u, \eta)$  is determined as follows: take a curve  $(u(t), \eta(t))$  in  $E_{\bar{n}}$  based on  $(u, \eta)$  and compute its tangent vector in  $t = 0$ . Differentiating the condition  $\eta(t)Qu(t) = u(t)$ , we get:

$$\dot{u} - \eta Q\dot{u} = \dot{\eta}Qu,$$

which is the same condition as for the degeneracy of the Hessian (the first equation in (15)).

We consider now the smooth semialgebraic map:

$$f : E_{\bar{n}} \rightarrow \Delta^2 \quad \text{given by} \quad (\omega, v) \mapsto q(v).$$

The set of regular values of  $f$  is a dense subset of  $\Delta^2$  (it is the complement of a semialgebraic set of codimension at least one): this subset is the good one we want to restrict to. In other words we consider  $\Sigma_4$  to be the union of the set of critical values of the various  $f : E_{\bar{n}} \rightarrow \Delta^2$  as  $\bar{n}$ ; the complement of  $\Sigma_4$  contains generic points<sup>2</sup>. On the preimage of a ‘‘good’’  $p$  we know that the differential of  $f$  is surjective with rank  $l$ ; moreover:

$$\dim E_{\bar{n}} = \dim \Lambda_{\bar{n}} + \dim E(\eta) = l - \nu + 2\nu = l + \nu.$$

Looking at the dimensions of the domain and the range of  $d_u f$  we get that the kernel has dimension  $\nu$ . On one hand, the kernel of  $d_u f$  is the vector space satisfying both the equations (15) for the degeneracy of the Hessian; on the other hand we have the inclusion:

$$E(\eta) \cap T_u\Omega_p \subset \ker d_u f.$$

Since both these spaces have dimension  $\nu$ , they must be equal. It follows that with all the above generic restrictions on  $p$ , the only directions of degeneracy for the Hessian are in  $T_uC_\eta$ .

*Remark 4.* We know that the operator  $\eta Q$  is compact, and that the eigenvalues are of the form  $\pm \frac{\alpha_i}{k}$  with  $k \in \mathbb{N}$  non-zero and  $i = 1, \dots, s < \infty$ . Then the eigenvalues of the Hessian of the energy on a critical point are  $1 \pm \frac{\alpha_i}{k}$ ; it follows that the number of negative eigenvalues is always finite, so that the index of every critical manifold is finite.

<sup>2</sup>Thus at this stage  $p \in \Delta^2 \setminus (\Sigma_1 \cup \Sigma_2 \cup \Sigma_3 \cup \Sigma_4)$ .



It remains to prove property (c) of the definition (the Palais-Smale property). Let us consider a sequence  $\{u_k\}$  in  $\Omega_p$  with energy  $\|u_k\|^2$  bounded by  $E$  and such that  $\nabla\psi_{u_k} \rightarrow 0$ , where  $\psi \doteq J|_{\Omega_p}$ .

The gradient  $\nabla\psi_u$  is the orthogonal projection of  $\nabla J_u = u$  on the space  $\text{span}\{Q_1u, \dots, Q_lu\}^\perp$ , then if we define  $\pi_u$  to be the orthogonal projection on the space  $\text{span}\{Q_1u, \dots, Q_lu\}$  we have

$$\nabla\psi_u = u - \pi_u u.$$

From  $\{u_k\}$  we can extract a subsequence (we keep calling it  $\{u_k\}$ ) such that  $\|u_k\|^2 \rightarrow L$ . Now we can compute

$$\|\nabla\psi_u\|^2 = \langle u, u \rangle - 2\langle u, \pi_u u \rangle + \langle \pi_u u, \pi_u u \rangle = \langle u, u \rangle - \langle \pi_u u, \pi_u u \rangle,$$

where the second equality follows from  $\langle u, \pi_u u \rangle = \langle \pi_u u, \pi_u u \rangle$  being  $\pi_u$  an orthogonal projection. Since  $\langle u_k, u_k \rangle \rightarrow L$ , and  $\|\nabla\psi_{u_k}\|^2 = \langle u_k, u_k \rangle - \langle \pi_{u_k} u_k, \pi_{u_k} u_k \rangle \rightarrow 0$ , it follows that

$$\langle \pi_{u_k} u_k, \pi_{u_k} u_k \rangle \rightarrow L.$$

Now, every  $\pi_{u_k} u_k$  is a linear combination of the vectors  $Q_i u_k$ , namely

$$\pi_{u_k} u_k = \sum_{i=1}^l \eta_k^i Q_i u_k,$$

and by computing its norm we get

$$(16) \quad \sum_{i,j=1}^l \eta_k^i \eta_k^j \langle Q_i u_k, Q_j u_k \rangle \rightarrow L.$$

Since the operators  $Q_i$  are compact, they map the bounded sequence  $\{u_k\}$  to a sequence  $\{Q_i u_k\}$  with limit points, so we can iteratively extract converging subsequences (again we keep calling them  $\{u_k\}$ ) and we have  $Q_i u_k \rightarrow v_i$ . In this way the equation (16) becomes

$$\sum_{i,j=1}^l \eta_k^i \eta_k^j \langle v_i, v_j \rangle \rightarrow L,$$

where the coefficients  $\langle v_i, v_j \rangle$  give the scalar product of the whole Hilbert space  $H$  restricted to the finite dimensional subspace  $V = \text{span}\{v_1, \dots, v_l\}$ . Now we have a bounded sequence of vectors  $\eta_n = \sum_{i=1}^l \eta_n^i v_i$  in  $\mathbb{R}^l$  from which we can extract a converging sequence with limit  $\eta$ .

So the sequence  $\pi_{u_k} u_k = \sum_{i=1}^l \eta_k^i Q_i u_k$  tends to  $v = \sum_{i=1}^l \eta^i v_i$ , and since

$$0 = \lim_{k \rightarrow \infty} \|\nabla\psi_{u_k}\|^2 = \lim_{k \rightarrow \infty} \|u_k - \pi_{u_k} u_k\|^2 = \lim_{k \rightarrow \infty} \|u_k - v\|^2,$$

also the sequence  $\{u_k\}$  tends to  $v$ . Moreover, since  $v_i = \lim_{k \rightarrow \infty} Q_i u_k = Q_i v$  we have that  $v = \eta Q v$  which is the condition for  $v$  to be a critical point of the Energy.  $\square$

#### 4. ADMISSIBLE-PATH SPACE AND ITS TOPOLOGY

**4.1. Paths with bounded energy.** Despite Theorem 9 shows that (a priori) only some of the  $\Omega_p$  are Hilbert manifolds, they are in fact all homotopy equivalent to each other; the argument is a simple modification of the standard one for loop spaces and appeared first in [10]; we recall it here for convenience of the reader.

**Theorem 15.** *For every  $p_1, p_2 \in G$  the spaces  $\Omega_{p_1}$  and  $\Omega_{p_2}$  are homotopy equivalent.*

*Proof.* It is sufficient to prove that for every  $p \in G$  the space  $\Omega_p$  is homotopy equivalent to  $\Omega_e$ . To this end let  $\gamma_0 \in \Omega_p$  be a fixed *admissible* path and define the map:

$$A : \Omega_e \rightarrow \Omega_p$$

by concatenation of loops in  $\Omega_e$  with  $\gamma$ :  $A(\gamma) = \gamma_0 \gamma$  (velocities have to be rescaled). Let also  $\hat{\gamma}_0$  be  $\gamma_0$  with backward time (it connects  $p$  to  $e$ ); then define

$$B : \Omega_p \rightarrow \Omega_e$$

by conatenation with  $\hat{\gamma}_0 : B(\gamma) = \hat{\gamma}_0\gamma$ . Let now  $\gamma_\epsilon, \epsilon \in [0, 1]$  be the paths:

$$\gamma_\epsilon(t) = \gamma_0(\epsilon(1-t))$$

and  $L_\epsilon : \gamma \rightarrow \gamma_\epsilon \hat{\gamma}_\epsilon \gamma$ . the maps  $L_\epsilon$  give a homotopy between is the identity  $L_0 = \text{id} : \Omega_e \rightarrow \Omega_e$  and  $L_1 = AB$ . In a similar way  $BA$  is homotopic to the identity on  $\Omega_p$  and the two spaces are homotopy equivalent.  $\square$

As a corollary we see that  $\Omega_p$  is contractible (in particular all its nonzero Betti numbers vanish).

**Corollary 16.** *For every  $p \in G$  the topological space  $\Omega_p$  is contractible.*

*Proof.* By the above Theorem it is enough to show that  $\Omega_e$  is contractible, and this is obvious since it is given by *homogeneous* equations.  $\square$

We start now the study of Morse-Bott theory of the function  $J|_{\Omega_p} = f$  on  $\Omega_p$ . In the sequel we will use the notation:

$$\Omega_p^s \doteq \Omega_p \cap \{J \leq s\}.$$

Proposition 7 tells us that in the case  $p$  is not a vertical point the number of geodesics joining  $e$  to  $p$  is finite; in particular if  $s > 0$  is large enough  $f$  does not have critical points on  $\{f \geq s\}$  and the topology of  $\Omega_p^s$  stabilizes.

**Proposition 17.** *If  $p$  is not a vertical point, then for every  $s$  large enough and  $t > 0$  the inclusion:*

$$\Omega_p^s \hookrightarrow \Omega_p^{s+t} \quad \text{is a homotopy equivalence.}$$

Thus we focus on the case  $p \in \Delta^2$ . As we already mentioned, we *will not* use Morse-Bott theory to give a lower bound on  $b(\Omega_p^s)$  by counting critical manifolds; we will instead use it to reduce the problem to the study of intersection of real quadrics. In fact the following proposition shows that  $\Omega_p^s$  is homotopy equivalent to its boundary  $\partial\Omega_p^s$ ; since the map  $q$  is quadratic, the space  $\partial\Omega_p^s$  is an intersection of infinite-dimensional quadrics (it is given by the quadratic equations  $\|u\|^2 = 2s$  and  $q(u) = p$ ).

**Proposition 18.** *For a generic choice of the Carnot group structure  $W \subset \mathfrak{so}(d)$  and a generic point  $p \in \Delta^2$ , for almost every  $s$  the following isomorphism holds:*

$$H_*(\Omega_p^s) \simeq H_*(\partial\Omega_p^s).$$

*Proof.* We first notice that the generic  $s$  is not a critical value for the energy. Let us consider now the Morse-Bott function  $g = -J$  and let us denote by  $X^a$  the set  $\{g \leq a\}$ . The critical manifolds of  $g$  are the same as for  $J$ , except that the *index* of each one of them for  $g$  is infinite (since these manifolds have *finite* index for  $J$ , then they must have *infinite* index for  $g$ ).

After passing a critical value  $c$  with corresponding critical manifold  $C$ , the relative homology of the Lebesgue set is given by (i.e. “the homology changes by”):

$$H_*(X^{c+\delta}, X^{c-\delta}) \simeq H_*(D_C^-, \partial D_C^-).$$

We recall that  $D_C^-$  is the unit disk bundle in the fiber bundle over  $C$  on which the Hessian of the Morse-Bott function is negative definite; see Theorem B.3 and the subsequent discussion from Appendix B. Notice that here the choice of the coefficients field  $\mathbb{Z}_2$  prevents us from the problem of orientability of this bundle.

This relative homology is zero: since the index of  $C$  is infinite, then both  $D_C^-$  and  $\partial D_C^-$  retract on  $C$  (this follows from the fact that the infinite dimensional sphere is contractible).

We can conclude our proof by observing that even though we pass critical values for  $-J$ , the homology remains the same of  $\partial\Omega_p^s$  until we get the whole  $\Omega_p^s$ .  $\square$

Thus we see that, being  $\Omega_p$  contractible, each of the Betti numbers  $b_i(\Omega_p^s)$  ( $i > 0$ ) eventually vanishes as  $s \rightarrow \infty$ . Despite this their sum can still grow: the smallest  $i > 0$  for which  $b_i(\Omega_p^s) \neq 0$  will get bigger and bigger and the amount of topology can increase as well: we are interested in understanding quantitatively this phenomem.

**4.2. Asymptotic Morse-Bott inequalities.** Before giving an explicit bound to  $b(\Omega_p^s)$ , we will see what *would* this bound be if we were to use Morse-Bott inequalities only. The Morse-Bott inequalities bound will follow from the count of the number of critical manifolds with energy less than  $s$ . It turns out that this bound is much worse than the actual one: in fact one has:

$$\text{Card}\{\text{critical manifolds with energy less than } s\} \leq O(s)^l$$

against the actual bound  $b(\Omega_p^s) \leq O(s)^{l-1}$  (this will be proved in the next section).

The following proposition will be fundamental for the sequel: essentially it allows to turn the direct limits arguments into a quantitative form. Roughly it says that the wave numbers of the controls associated to the geodesics grow at most with the order of their energy.

To deal with this ideas, we introduce the following useful notation: for every  $L \in \mathbb{N}$  we define:

$$T^L \doteq \bigoplus_{k \leq L} T_k.$$

**Proposition 19.** *For the generic choice of the Carnot group structure  $W \subset \mathfrak{so}(d)$  and the generic point  $p \in \Delta^2$  there exists a constant  $c_p > 0$  such that for every geodesic  $\gamma \in \Omega_p \cap \{J \leq s\}$ , its associated control belongs to  $T^{\lfloor sc_p \rfloor}$ .*

In order to prove the previous proposition we first need the following lemma:

**Lemma 20.** *For the generic choice of the Carnot group structure  $W \subset \mathfrak{so}(d)$  and the generic point  $p \in \Delta^2$  there exists a constant  $c_p > 0$  such that for every Lagrange multiplier  $\omega$  associated to  $p$ , the following inequality holds:*

$$\frac{\langle \omega, p \rangle}{\|\omega\|} \geq \frac{1}{c_p}.$$

*Remark 5.* Being the quantity  $\langle \omega, p \rangle / \|\omega\|$  the cosinus between  $\omega$  and  $p$  times the norm of  $p$ , the lemma says that the Lagrange multipliers for  $p$  are contained in a convex acute cone in  $W$ .

The norm on the space of covectors  $(\Delta^2)^* \cong W$  is the one induced by the inclusion  $W \hookrightarrow \mathfrak{so}(d)$  where  $\langle X, Y \rangle = \text{Trace}(X^T Y)$ .

Before giving the proof of the lemma we show how it implies Proposition 19.

*Proof of proposition 19.* Lemma 20 is equivalent to

$$\|\omega\| \leq c_p \langle \omega, p \rangle.$$

The norm of  $\omega A \in W$  can be written in terms of its eigenvalues  $\alpha_1, \dots, \alpha_s$ :  $\|\omega\| = \|\omega A\| = \sqrt{2\alpha_1^2 + \dots + 2\alpha_s^2}$ ; it follows that every eigenvalue of  $\omega A$  is smaller than the norm of  $\omega A$ . Since  $\omega$  is a Lagrange multiplier, if  $u = u_{k_1} + \dots + u_{k_l}$  then the  $k_j$  are integer eigenvalues of  $i\omega A$ . Since the energy of a geodesic  $u$  associated to  $\omega$  is  $J(u) = \langle \omega, p \rangle$ , we have

$$(17) \quad k_j \leq \|\omega\| \leq c_p \langle \omega, p \rangle \leq c_p s.$$

□

Now we go back to the proof of Lemma 20.

*Proof of Lemma 20.* Suppose on the contrary that the constant bounding  $\langle \omega, p \rangle / \|\omega\|$  from below doesn't exist, so that we can find a sequence of Lagrange multipliers  $\omega_n$  such that, setting  $\hat{\omega}_n \doteq \omega_n / \|\omega_n\|$ , the sequence  $\hat{\omega}_n(p) \rightarrow 0$ . Since the sequence  $\hat{\omega}_n$  is contained in  $S^{l-1}$  which is compact, we can assume (up to subsequences) that it converges, with limit  $\lambda$  such that  $\langle \lambda, p \rangle = 0$  by hypothesis. Up to subsequences we can also assume that every Lagrange multiplier  $\omega_n$  has the same number of integer eigenvalues (all distinct by Lemma 11), say  $\nu$ . For every Lagrange multiplier  $\omega_n$  we have the cone of the endpoints of the geodesics associated to  $\omega_n$ ; the Lagrange multiplier  $\omega_n$  is contained in the intersection of  $\nu$  hypersurfaces of matrices with constant eigenvalue equal the imaginary integers  $ik_1(n), \dots, ik_r(n)$ . The direct sum  $E(\omega)$  of the associated eigenvalues contains all the geodesics with Lagrange multiplier  $\omega$  and  $q(E(\omega))$  is the cone spanned by the normal vectors to each of these  $\nu$  surfaces (Proposition 12), where

the normal vector has to be chosen with positive scalar product with  $\omega_n$  (since the Energy is positive).

Let us call  $l_j^+$  the normal vectors to the surfaces with eigenvalue equal to  $k_j(n)$  respectively; up to subsequences again we can assume that every direction  $l_j^+(n)$  converges to some  $l_j^+$ . The point  $p$  is contained in the interior of every cone (Theorem 13) and it can be written as  $p = \sum_{j=1}^{\nu} c_j(n) l_j^+(n)$  with  $c_j(n) > 0$  for every  $j$  and every  $n$ . Since  $\langle \lambda, p \rangle = 0$ , we have

$$0 = \langle \lambda, p \rangle = \sum_{j=1}^{\nu} \langle \lambda, l_j^+ \rangle = \lim_{n \rightarrow \infty} \sum_{j=1}^{\nu} c_j(n) \langle \hat{\omega}_n, l_j^+(n) \rangle;$$

the terms of the sum above must converge to 0 one by one because they are non-negative. Not every  $c_j(n)$  can converge to 0, otherwise  $p$  would be 0 as well. Therefore at least one of the terms  $\langle \hat{\omega}_n, l_j^+(n) \rangle$  converges to 0; in the limit  $p$  is a linear combination of directions  $l_j^+$  with  $j \in I$  for a set of indexes  $I$  such that  $\langle \lambda, l_j^+ \rangle = 0$  for every  $j \in I$ .

Now we are going to see what happens to Lagrange multipliers associated to directions  $l_j^+$  such that  $\lim_{n \rightarrow \infty} \langle \hat{\omega}_n, l_j^+(n) \rangle = 0$ . Let us take the smooth hypersurface  $S_\alpha \subset W$  of the matrices with an eigenvalue equal to a given  $i\alpha \in i\mathbb{R}$  and all the other eigenvalues different from  $i\alpha$ . Then we take a point  $\omega A \in S_\alpha$  and we compute the angle between  $\omega A$  and the normal vector to the surface  $S_\alpha$  at  $\omega A$ . The surface  $S_\alpha$  can be given as a zero locus of the real valued function  $s_\alpha(\eta) \doteq \det(\eta A - i\alpha \mathbb{1})$  in the even-dimensional case and  $i \det(\eta A - i\alpha \mathbb{1})$  in the odd-dimensional case (here for simplicity we discuss the case  $d$  is even, but the proof for the odd-dimensional case is analogous).

We may assume that the matrices  $A^1, \dots, A^l$  form an *orthonormal* basis for  $W$ ; moreover we can choose an orthonormal basis for  $\Delta$  such that the matrix  $\omega A$  is written in canonical form, i.e.  $\omega A = \text{Diag}(\alpha J_2, \alpha_2 J_2, \dots, \alpha_{d/2} J_2)$ . Now we compute the differential of  $s_\alpha$  at  $\omega A$ :

$$(ds_\alpha)_{\omega A} = \sum_{i=1, j=1}^d \sum_{k=1}^l \frac{\partial s_\alpha}{\partial m_{ij}} \frac{\partial m_{ij}}{\partial \eta_k} d\eta_k = \sum_{i=1, j=1}^d \sum_{k=1}^l \text{adj}(\omega A - i\alpha \mathbb{1})^{ij} a_{ij}^k d\eta_k$$

where  $m_{ij}$  are the variables for the entries of the matrices,  $\eta_k$  are the coordinates on  $W$  given by the components of the covectors in  $(\Delta^2)^*$ ,  $\text{adj}(\omega A - i\alpha \mathbb{1})^{ij}$  is the  $ij$  entry of the adjugate matrix of  $\omega A - i\alpha \mathbb{1}$  and  $a_{ij}^k$  are the entries of the matrix  $A^k$ . The matrix  $\omega A - i\alpha \mathbb{1}$  takes the form

$$\omega A - i\alpha \mathbb{1} = \text{Diag} \left( \begin{pmatrix} -i\alpha & \alpha \\ -\alpha & -i\alpha \end{pmatrix}, \begin{pmatrix} -i\alpha & \alpha_2 \\ -\alpha_2 & -i\alpha \end{pmatrix}, \dots, \begin{pmatrix} -i\alpha & \alpha_{d/2} \\ -\alpha_{d/2} & -i\alpha \end{pmatrix} \right);$$

setting  $\beta \doteq \prod_{i=2}^{d/2} (a_i^2 - \alpha^2)$ , the adjugate matrix is

$$\text{adj}(\omega A - i\alpha \mathbb{1}) = \text{Diag} \left( \begin{pmatrix} -i\alpha\beta & \alpha\beta \\ -\alpha\beta & -i\alpha\beta \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right),$$

and so we get

$$(ds_\alpha)_{\omega A} = 2\alpha\beta a_{12}^k d\eta_k.$$

Therefore we have

$$\langle (ds_\alpha)_{\omega A}, \omega A \rangle = 2\alpha\beta \omega_k a_{12}^k = 2\alpha^2\beta.$$

Since the basis  $(A^1, \dots, A^l)$  is orthonormal for  $W$  and so is  $(d\lambda_1, \dots, d\lambda_l)$  for  $W^*$ , the norm  $\|(ds_\alpha)_{\omega A}\|$  is easily computed:

$$\|(ds_\alpha)_{\omega A}\| = 2|\alpha\beta| \sqrt{(a_{12}^1)^2 + \dots + (a_{12}^l)^2} = 2|\alpha\beta| \|a_{12}\|$$

where  $a_{12} \doteq (a_{12}^1, \dots, a_{12}^l)$ . Now we can compute the cosinus of the angle  $\theta$  between  $\omega$  and the normal to  $S_\alpha$  at  $\omega A$ :

$$\cos \theta = \frac{\langle (ds_\alpha)_{\omega A}, \omega A \rangle}{\|(ds_\alpha)_{\omega A}\| \|\omega A\|} = \frac{2\alpha^2\beta}{2|\alpha\beta| \|a_{12}\| \|\omega\|} = \pm \frac{\alpha}{\|a_{12}\| \|\omega\|}.$$

Let us go back to the directions  $l_j^+$  such that  $\langle \lambda, l_j^+ \rangle = \lim_{n \rightarrow \infty} \langle \hat{\omega}_n, l_j^+(n) \rangle = 0$ ; recall that the imaginary integers  $ik_j(n)$  are the eigenvalue corresponding to the eigenspaces mapped on  $\mathbb{R}l_j^+(n)$ . Notice that the constant  $\|a_{12}\|$  in the previous computation depends on the space  $W$  and on the basis we choose in order to put  $\omega A$  in canonical form, but now the basis is a priori different for every  $\omega_n A$ ; indeed we have that  $\|a_{12}\| \leq \|A\|$  where the norm  $\|A\|$  is the one induced by the norm on  $\Delta$  on its exterior algebra. It follows that

$$|\langle \hat{\omega}_n, l_j^+(n) \rangle| \geq \left| \frac{\alpha_i(n)}{\|A\| \|\omega_n\|} \right|,$$

and since the first term converges to 0, so does the second one.

Notice that the vectors  $l_j^+(c)$ , orthogonal to the surfaces with eigenvalue equal to  $k_i(n)$ , are orthogonal to the surfaces with eigenvalue equal to  $\frac{k_i(n)}{\|\omega_n\|}$  as well, since the surfaces are the same up to homothety. The point  $p$  is a linear combination of the directions  $l_j^+$  with  $j \in I$  with the property that  $\langle \lambda, l_j^+ \rangle = 0$ : these directions are the images of the limit eigenspaces  $E_j(\lambda)$  of the sequences  $E_j(\hat{\omega}_n)$ . Since their eigenvalues are  $\frac{k_j(n)}{\|\omega_n\|}$  which converge to 0, the eigenvalues of the eigenspaces  $E_j(\lambda)$  for  $j \in I$  are 0. It follows that  $p$  belongs to the image by  $q$  of the kernel of  $\omega Q$ , which means that  $p$  is a critical value for  $q$ , and this is absurd by hypothesis.  $\square$

We prove now the theorem that gives an upper bound for the number of critical manifolds with energy bounded by  $s$ .

**Theorem 21.** *For the generic choice of the Carnot group structure  $W \subset \mathfrak{so}(d)$  and of  $p \in \Delta^2$ :*

$$\text{Card}\{\text{critical manifolds with energy less than } s\} \leq O(s)^l.$$

*Proof.* Let us start by considering the vector space  $P_{l,d}$  of real polynomials of degree  $d$  in  $l$  variables. For every  $\nu = 1, \dots, l$  we set  $f = (f_1, \dots, f_\nu) \in P_{l,d}^\nu$  and :

$$Y_\nu = \{(f, V, \omega) \in P_{l,d}^\nu \times G(l - \nu, l) \times \mathbb{R}^l \mid \omega \in Z(f) \setminus \text{Sing}(Z(f)), V = (T_\omega Z(f))^\perp, p \in V\}.$$

Since  $Y_\nu$  is semialgebraic, we can consider the semialgebraic projection to the first factor  $\pi : Y_\nu \rightarrow P_{l,d}^\nu$  and stratify  $P_{l,d}^\nu = \coprod_{j=1}^s P_j$  such that  $\pi$  is semialgebraically trivial over each stratum (see [5]). In particular since there are finitely many strata, there exists a number  $\beta_\nu$  such that for every  $f \in P_{l,d}^\nu$ :

$$b_0(\pi^{-1}(f)) \leq \beta_\nu$$

(only a finite number of fibers, up to semialgebraic homeomorphism, appear). In particular, in the case when there are finitely many  $\omega$  with normals  $V = (T_\omega Z(f))^\perp$  containing  $p$ , this construction implies their number is bounded by  $\beta_\nu$ .

Consider now for every  $n \in \mathbb{N}$  the polynomial  $f_n \in P_{l,d}$  defined by:

$$f_n(\omega) = i^d \det(\omega A - in\mathbb{1})$$

(the  $i^d$  factor has the only scope of turning  $f_n$  into a *real* polynomial in the case  $d$  is odd). We know from Theorem 11 that for the generic choice of  $W \subset \mathfrak{so}(d)$  and  $p \in \Delta^2$  each Lagrange multiplier belongs to some  $Z(f_{k_1}, \dots, f_{k_\nu})$ , where  $\nu$  is the number of positive integer eigenvalues of the matrix  $i\omega A$ . Now if  $u = u_1 + \dots + u_\nu$  is the control associated to a geodesic with final point  $p$  and energy less than  $s$ , Proposition 19 implies:

$$k_j \leq sc_p \quad \text{for } j = 1, \dots, \nu$$

for a constant  $c_p$  depending only on  $W$  and  $p$ .

Thus the way to get all *possible* Lagrange multipliers associated to geodesics ending at  $p$  with energy less than  $s$  is by intersecting  $\nu$  of the hypersurfaces  $Z(f_k)$  for  $k = 1, \dots, \lfloor sc_p \rfloor$  and  $\nu = 1, \dots, l$  and considering the points in this intersection where the normal space contains  $p$ . There are

$$\binom{\lfloor sc_p \rfloor}{\nu} = O(s)^\nu$$

possible ways of choosing the hypersurfaces and the above argument implies each such choice can contribute by at most  $\beta_\nu$  Lagrange multipliers.

In particular the set of Lagrange multipliers for  $p$  with energy less than  $s$  is bounded by:

$$\text{Card}\{\text{Lagrange multipliers } \omega \text{ such that } \omega(p) \leq s\} \leq \sum_{\nu=1}^l O(s)^\nu = O(s)^l.$$

To each critical manifold with energy less than  $s$  there corresponds one and only one Lagrange multiplier whose scalar product with  $p$  is less than  $s$ , hence the conclusion follows.  $\square$

As a corollary we derive now Morse-Bott inequalities; as already stated the following bound will be improved in the next section.

**Corollary 22** (Morse-Bott inequalities). *For the generic choice of  $W \subset \mathfrak{so}(d)$  and of the point  $p \in \Delta^2$  we have:*

$$b(\Omega_p^s) \leq \sum_{J(C_\omega) \leq s} b(C_\omega) \leq O(s)^l.$$

*Proof.* By Lemma 12 and Theorem 13 we know that each critical manifold is an intersection of  $l$  quadrics in  $\mathbb{R}^{2d}$  (it is the preimage of  $p$  under  $q|_{E(\omega)}$ ): in particular the possible homeomorphism types of such manifold are finite and there is a constant  $\beta$  such that  $b(C_\omega) \leq \beta$  for every critical manifold  $C_\omega$ .

Since the sum  $\sum_{J(C) \leq s} b(C)$  contains at most  $O(s)^l$  terms (by Theorem 21), the conclusion follows from Morse-Bott inequalities (see Appendix B).  $\square$

**4.3. Asymptotic total Betti number.** The aim of this section is to refine the bound for  $b(\Omega_p^s)$  given in Corollary 22.

We start by proving some technical results.

**Proposition 23.** *For the generic choice of  $W \in \mathfrak{so}(d)$  and the generic point  $p \in \Delta^2$  we have: (a) the critical manifolds of  $J$  on  $\Omega_p$  with Energy less than  $s$  coincide with the critical manifolds of  $J$  restricted to  $\Omega_p \cap T^{\lfloor sc_p \rfloor}$ ; (b) their index is the same either if they are considered critical manifolds for  $J$  or if they are considered critical manifolds for  $J$  restricted to  $\Omega_p \cap T^{\lfloor sc_p \rfloor}$ .*

*Proof.* We have already proved that there exists a constant  $c_p$  such that all the critical points with energy less than  $s$  are contained in  $\Omega_p \cap T^{\lfloor sc_p \rfloor}$  (Proposition 19); since the spaces  $T_k$  are orthogonal with respect to both the quadratic maps  $J$  and  $q$ , the critical points of  $J|_{\Omega_p \cap T^{\lfloor sc_p \rfloor}}$  are given by the same equations as for critical points of  $J|_{\Omega_p}$  using the Lagrange multipliers rule.

Let  $\omega$  be a Lagrange multiplier for  $p$  with energy less than  $s$ . The Hessian of the energy  $J$  is  $\langle \text{Id} - \omega Q \cdot, \cdot \rangle$ , so its eigenvalues are  $1 - \frac{\alpha_i(\omega)}{k}$  with  $k \in \mathbb{N}_0$  and  $\alpha_i(\omega)$  eigenvalues of  $\omega A$  as usual. Therefore we have negative eigenvalues for every integer  $k$  such that  $k < \alpha_i(\omega)$  for at least one  $i$ ; since  $\alpha_i(\omega) \leq \|\omega\|$ , it follows from Lemma 20 (as in the proof of Proposition 19) that

$$k \leq \|\omega\| \leq c_p \langle \omega, p \rangle \leq sc_p$$

with the same constant  $c_p$  as in Proposition 19.  $\square$

The next proposition tells that if we want to compute  $b(\Omega_p^s)$  we can restrict ourselves to the intersection with a finite dimensional subspace of the form  $T^L$  (and indeed we have a quantitative control on the dimension).

**Proposition 24.** *For a generic choice of  $W \subset \mathfrak{so}(d)$  and of  $p \in \Delta^2$  there exists a constant  $r_p > 0$  such that for every  $m \in \mathbb{N}$ :*

$$\Omega_p^s \text{ deformation retracts to } \Omega_p^s \cap T^{\lfloor sr_p \rfloor + m}.$$

*In particular:  $H_*(\Omega_p^s) \simeq H_*(\Omega_p^s \cap T^{\lfloor sr_p \rfloor + m})$ .*

*Proof.* Given  $L \in \mathbb{N}$  we can define the function  $f_L$ , “distance from  $T^L$ ” in the following way: every  $u \in \Omega_p^s$  can be uniquely written as  $u = \bar{u} + v$  where  $\bar{u} \in T^L$  and  $v \in (T^L)^\perp$ ; then we define  $f_L(u) \doteq \|v\|^2$ . Assume that for a suitable  $L$  there are no critical points for  $f_L$  outside  $\Omega_p^s \cap T^L$  and the function  $f_L$  satisfies the Palais-Smale condition: then we can retract the manifold  $f_L^{-1}([0, s]) = \Omega_p^s$  on the sublevel set  $f_L^{-1}(0) = \Omega_p^s \cap T^L$  by Theorem B.2 (the function  $f_L$  is bounded on  $\widehat{\Omega}_{ep}$  since  $f_L \leq J \leq s$ ). The manifolds we are considering have boundary, but we can apply the argument above to  $\Omega_p \cap \{J < s + \delta\}$ : it deformation retracts to  $\Omega_p^s$  again by Theorem B.2 if we choose  $\delta$  small enough not to have new critical values for the Energy; we can indeed choose such  $\delta$  small enough not to have new critical values for  $f_L$  as well.

We are now going to prove that  $f_L$  satisfies the Palais-Smale condition and to find for which  $L$  we are sure not to have critical points for  $f_L$  in  $\Omega_p^s \setminus (\Omega_p^s \cap T^L)$ .

The gradient of the function  $f_L$  at  $u = \bar{u} + v$  is  $2v$  restricted to  $T_u\Omega_p$ : this means that there exists  $\eta \in (\Delta^2)^*$  such that  $\nabla_u f_L = 2v - 2\eta Qu$ . For every critical point  $u$  for  $f_L$  there exists  $\eta$  such that  $v - \eta Qu = v - \eta Qv - \eta Q\bar{u} = 0$ : since the space  $T^L$  and its orthogonal are invariant with respect to  $\eta Q$  we have the equivalent couple of conditions

$$v = \eta Qv, \quad \eta Q\bar{u} = 0;$$

the first condition tells that  $v$  is a geodesic going somewhere, the second one tells that  $q(\bar{u})$  is a critical value for  $q$ .

We need  $f_L$  to satisfy the Palais-Smale condition: having the explicit expression of its gradient  $\nabla f_L$ , we omit this verification whose proof is analogous to the one for  $J$  in Theorem 14.

We are going to prove that there exists a constant  $r_p$  such that if we take  $L = \lfloor sr_p \rfloor$  there are no critical points for  $f_L$  in  $\Omega_p^s \setminus (\Omega_p^s \cap T^L)$ .

Assume that such a constant does not exist: then for every term  $\rho_n$  of a diverging sequence of positive real numbers, we find  $s_n > 0$  and a critical point  $u(n)$  for the function  $f_{\lfloor s_n \rho_n \rfloor}$  outside  $T^{\lfloor \rho_n s_n \rfloor}$  with Energy less than  $s_n$ . By hypothesis  $v(n) \in (T^{\lfloor \rho_n s_n \rfloor})^\perp$  so that its Fourier expansion is  $v(n) = \sum_{k > \rho_n s_n} v_k(n)$ . Recall that  $P \doteq Q|_{T_1}$ : then we have:

$$\begin{aligned} \|q(v(n))\| &= \left\| \sum_{k > \rho_n s_n} \frac{1}{k} \langle P v_k(n), v_k(n) \rangle \right\| \leq \sum_{k > \rho_n s_n} \frac{1}{k} \|P\| \|u_k(n)\|^2 \leq \\ &\leq \frac{1}{\rho_n s_n} \|P\| \|v(n)\|^2 \leq \frac{1}{\rho_n s_n} \|P\| s_n = \frac{\|P\|}{\rho_n}. \end{aligned}$$

The last term of the chain of inequalities converges to zero: it follows that  $q(v(n))$  goes to zero as well. Since  $p = q(\bar{u}(n)) + q(v(n))$ , we get

$$\lim_{n \rightarrow \infty} q(\bar{u}(n)) = p,$$

and as we noticed before  $\bar{u}(n)$  is a critical point for  $q$ . This means that we have a sequence of critical values for  $q$  converging to  $p$ , which is impossible since we picked our  $p$  into an open subset of regular values (i.e.  $p \in \Delta^2 \setminus \Sigma_1$ ). We end the proof by noticing that the condition for having critical points for  $f_{\lfloor r_p s_n \rfloor}$  has to be verified on every  $T_k$  separately (they are orthogonal and invariant with respect to  $Q$ ): since there are no critical points of  $f_{\lfloor r_p s_n \rfloor}$  outside  $T^{\lfloor r_p s_n \rfloor}$ , there are no critical points of the same function restricted to  $T^{\lfloor r_p s_n \rfloor + m}$  (similarly to what happens to the critical points of  $J$  in Proposition 23) and it follows eventually that  $\Omega_p^s \cap T^{\lfloor r_p s_n \rfloor + m}$  deformation retracts onto  $\Omega_p^s \cap T^{\lfloor r_p s_n \rfloor}$  for every  $m \in \mathbb{N}$ .  $\square$

As a corollary we get the following interesting result, that controls the growth rate of the index of the highest nonzero Betti number of  $\Omega_p^s$ .

**Corollary 25.** *For a generic  $W \subset \mathfrak{so}(d)$  and  $p \in \Delta^2$ , there exists a constant  $r_p > 0$  such that:*

$$\max_i \{i \mid b_i(\Omega_p^s) \neq 0\} \leq 2d \lfloor r_p s \rfloor.$$

*Proof.* By Proposition 24 there exists  $c_p > 0$  such that  $H_*(\Omega_p^s) \simeq H_*(\Omega_p^s \cap T^{\lfloor c_p s \rfloor})$ . In particular  $\Omega_p^s$  has the homology of a semialgebraic subset of  $\mathbb{R}^{2d\lfloor c_p s \rfloor}$  (namely  $\Omega_p^s \cap T^{\lfloor c_p s \rfloor}$ ) and its  $j$ -th Betti number must be zero for  $j > 2d\lfloor c_p s \rfloor$ .  $\square$

Everything is now ready for the proof of the main theorem of this section.

**Theorem 26** (Strong Morse-Bott inequalities). *For the generic choice of  $W \subset \mathfrak{so}(d)$  and of  $p \in \Delta^2$  we have:*

$$b(\Omega_p^s) \leq O(s)^{l-1}.$$

*Proof.* First we know from Proposition 24 that there exists  $r_p > 0$  such that:

$$b(\Omega_p^s) = b(\Omega_p^s \cap T^{\lfloor r_p s \rfloor}) \quad \text{for all } s > 0 \text{ and } m \in \mathbb{N}.$$

It means that  $\Omega_p^s$  has the same Betti numbers as

$$\{x \in T^{\lfloor r_p s \rfloor} \mid q_1(x) = p_1, \dots, q_l(x) = p_l\} \cap \{\|x\|^2 \leq 2s\}.$$

Notice that the dimension of  $T^{\lfloor r_p s \rfloor}$  is a  $O(s)$ .

Let us consider the semialgebraic set (a level set of a quadratic map with  $l$  components):

$$X = \{x \in T^{\lfloor r_p s \rfloor} \mid q_1(x) = p_1, \dots, q_l(x) = p_l\}$$

and  $\epsilon > 0$  small enough such that  $X \cap \{2s - \epsilon \leq \|x\|^2 \leq 2s + \epsilon\}$  deformation retracts onto  $X$  (the existence of such  $\epsilon$  is guaranteed by semialgebraic triviality, see [5]). Let also  $A = X \cap \{\|x\|^2 \leq 2s + \epsilon\}$  (which deformation retracts onto  $X \cap \{\|x\|^2 \leq 2s\}$ ) and  $B = X \cap \{2s - \epsilon \leq \|x\|^2\}$ . The Mayer-Vietoris exact sequence of the pair  $(A, B)$  gives  $b(A) + b(B) \leq b(A \cap B) + b(A \cup B)$ , which implies:

$$b(\Omega_p^s) = b(A) \leq b(A \cap B) + b(A \cup B).$$

Since  $A \cap B$  deformation retracts onto  $X \cap \{\|x\|^2 = 2s\}$ , then it is defined by  $l$  quadratic equations on a sphere of dimension  $\dim(T^{\lfloor r_p s \rfloor}) - 1 = O(s)$ ; on the other hand  $X = A \cup B$  is given by  $l$  quadratic equations in a vector space of dimension  $\dim(T^{\lfloor r_p s \rfloor}) = O(s)$ ; hence by Proposition C.1 the total Betti numbers of these spaces are bounded by:

$$b(A \cap B) \leq O(s)^{l-1} \quad \text{and} \quad b(A \cup B) \leq O(s)^{l-1}$$

and the conclusion follows.  $\square$

**4.4. A topological coarea formula.** In this section we compute the first order asymptotic of  $b(\Omega_p^s)$  in  $s$ , for the case  $l = 2$ . It turns out that for a generic choice of the Carnot group structure  $W \subset \mathfrak{so}(d)$  and the point  $p \in \Delta^2$ , the leading term is a *real* number and can be analytically computed using only the data  $W = \text{span}\{A_1, A_2\}$ .

Consider a unit circle  $S^1 \subset W$ . For a generic  $W$  the eigenvalues of  $\omega A$  are distinct and differentiable almost everywhere (the set of matrices in  $\mathfrak{so}(d)$  with multiple eigenvalues is a cone with codimension 3, and the eigenvalues are semialgebraic functions of the parameter  $\omega \in S^1$ ). Thus there exist semialgebraic functions  $\alpha_j : S^1 \rightarrow \mathbb{R}$  such that the  $\alpha_j(\omega)$ , for  $j = 1, \dots, d$ , are the coefficients of the canonical form of  $\omega A$ . Given  $p \in \Delta$  we consider the *rational* functions  $\lambda_j : S^1 \rightarrow \mathbb{R} \cup \{\infty\}$  given by:

$$\lambda_j : \omega \mapsto \left| \frac{\alpha_j(\omega)}{\langle \omega, p \rangle} \right| \quad \text{for } j = 1, \dots, d.$$

Notice that when  $\omega$  approaches  $p^\perp$  these functions might explode, that is why they are rational in  $\omega$ ; on the other hand they are semialgebraic and differentiable almost everywhere and it makes sense to consider the integral:

$$(18) \quad \tau(p) \doteq \frac{1}{2} \int_{S^1} \sum_{j=1}^d \left| \dot{\lambda}_j(\omega) \right| - \left| \sum_{j=1}^d \dot{\lambda}_j(\omega) \right| d\omega.$$



The convergence of the integral follows from the fact that where the derivatives of the  $\lambda_j$ s explode, they all have the same sign and the integrand vanishes.

The next theorem proves that for a fixed  $p$ , as a function of  $s$ :

$$b(\Omega_p^s) = \tau(p)s + o(s) \quad \text{as } s \rightarrow \infty.$$

**Theorem 27.** *If the corank  $l = 2$ , for a generic choice of  $W \subset \mathfrak{so}(d)$  and  $p \in \Delta^2$  we have:*

$$\lim_{s \rightarrow \infty} \frac{b(\Omega_p^s)}{s} = \tau(p).$$

*Proof.* In order to compute the asymptotic for  $b(\Omega_p^s)$  for  $s \rightarrow \infty$  we use (28). In fact we have seen that  $\Omega_p^s$  is homotopy equivalent to  $\{v \in H \mid q(v) = p/s, \|v\|^2 = 1\}$  and the latter can be rewritten as:

$$\left\{ v \in H \mid q(v) - \frac{\|v\|^2}{s} p = 0 \right\} \cap S$$

where  $S$  is the infinite dimensional sphere  $S = \{\|v\|^2 = 1\}$ . In particular we can present our set as the intersection of two quadrics in  $H$  on the unit sphere  $S$ ; thus we let  $i_s^-$  for the index function of the quadratic map  $q - p\| \cdot \|^2/s$  (we are using the notation of C.2). In this setting the set  $P \subset S^1$  coincides with  $\{\omega \in S^1 \mid \langle \omega, p \rangle < 0\}$ . In fact if we let  $p = (p_1, p_2)$ , here the two quadrics we are considering are  $q_1 - p_1\| \cdot \|^2/s$  and  $q_2 - p_2\| \cdot \|^2/s$  and for every  $\omega$  the selfadjoint operator on  $H$  corresponding to the quadratic form  $\omega q$  is  $\omega Q - \omega(p)\mathbb{1}/s$ . In particular the spectrum of  $\omega Q - \omega(p)\mathbb{1}/s$  is obtained by translating the spectrum of  $\omega Q$  by  $\langle \omega, p \rangle/s$  and since  $\omega Q$  is compact and its spectrum is symmetric with respect to the origin, we see that in order to have finitely many negative eigenvalues we need  $\langle \omega, p \rangle < 0$ .

On the other hand the subspaces  $T_k$  are invariant by both  $\omega Q$  and  $\langle \omega, p \rangle\mathbb{1}/s$ , thus the index function can be computed as:

$$\begin{aligned} i_s^-(\omega) &= \sum_{k \geq 1} i^- \left( \omega Q - \frac{\langle \omega, p \rangle}{s} \mathbb{1}|_{T_k} \right) = \sum_{k \geq 1} i^- \left( \frac{\omega Q_0}{k} - \frac{\langle \omega, p \rangle}{s} \mathbb{1} \right) = \\ &= \sum_{j=1}^d \left[ \frac{s\alpha_j(\omega)}{\langle \omega, p \rangle} \right] = \sum_{j=1}^d \lfloor s\lambda_j(\omega) \rfloor \end{aligned}$$

where in the second line we have used the fact that the spectrum of  $\frac{\omega Q_0}{k} - \frac{\langle \omega, p \rangle}{s} \mathbb{1}$  is of the form  $\frac{\alpha_j(\omega)}{k} - \frac{\langle \omega, p \rangle}{s}$ . In the sequel we also identify  $P \subset S^1$  with a subset of  $[0, 2\pi]$  in the standard way.

Denoting now by  $\mu(s)$  the number of local maxima of  $i_s^-$  on  $P$ , we see that formula (28) implies:

$$(19) \quad b(\Omega_p^s) = 2\mu(s) + 1 - b_0(P_0) = 2\mu(s) + O(1)$$

In fact, using the long exact sequence of the pair  $(P_{j+1}, P_j)$ , we can rewrite  $b_0(P_{j+1}, P_j) = b_0(P_{j+1}) - b_0(P_j) + b_1(P_{j+1}, P_j)$ ; substituting these identities into  $b(\hat{\Omega}_{\epsilon p}) = 1 + \sum_{j \geq 1} b_0(P_{j+1}, P_j) + b_1(P_{j+2}, P_{j+1})$  we get  $b(\hat{\Omega}_{\epsilon p}) = 1 - b_0(P_0) + 2 \sum_{j \geq 1} b_1(P_{j+1}, P_j)$ . Since each local maximum of  $i_s^-$  contributes by 1 to one of the  $b_1(P_{j+1}, P_j)$  and  $b(P_0) \leq 1$  (since  $P_0$  is convex), then (19) follows.

In order to compute the asymptotic of the number of maxima of  $i_s^-$  we introduce the following auxiliary data. First we let  $\lambda = \sum_{j=1}^d \lambda_j(\omega)$  and notice that this is a semialgebraic function. In particular we can divide  $P$  into a finite number of intervals (arcs):

$$P = (\omega_0, \omega_1] \cup [\omega_1, \omega_2] \cup \dots \cup [\omega_m, \omega_{m+1}] \cup [\omega_{m+1}, \omega_{m+2})$$

such that for every  $j, k$  the functions  $\alpha_j$  as well as  $\alpha$  are monotone on  $(\omega_k, \omega_{k+1})$ . Labeling  $I_k = [\omega_k, \omega_{k+1}]$  we see that also each  $\lfloor s\lambda_j \rfloor$  is monotone on  $I_k$ . On the other hand monotonicity of  $i_s^-$  is granted only where the signs of the derivatives of the  $\lambda_j$  all agree. Since for the generic choice of  $p$  the functions  $\alpha_j$  do not vanish on  $\{\omega_0, \omega_{m+2}\}$  (the orthogonal complement of  $p$  on  $S^1$ ),  $\lambda_j$  approaches infinity when approaching  $\omega_0$  or  $\omega_{m+2}$ ; in particular  $i_s^-$  is monotone on  $I_0$  and  $I_{m+1}$  and has no local maxima on them.

For every  $j \in \{1, \dots, m\}$  let us denote respectively by  $\mu_j(s)$  and  $\sigma_j(s)$  the number of local maxima of  $i_s^-$  on  $I_j$  and the number of subintervals of  $I_j$  where  $i_s^-$  is constant (thus  $\sigma_j(s)$  equals the number of ‘‘jumps’’ of the integer valued function  $i_s^-$  on  $I_j$ ).

For every interval  $I_j = [\omega_j, \omega_{j+1}]$  we see that:

$$|i_s^-(\omega_{j+1}) - i_s^-(\omega_j)| = \sigma_j(s) - 2\mu_j(s)$$

In particular summing all these equations and using the fact that  $i_s^-$  is monotone on  $I_0$  and  $I_{m+1}$ , combining with (19) we get:

$$\begin{aligned} \frac{b(\Omega_p^s)}{s} &= 2\frac{\mu(\epsilon)}{s} + O(1/s) = 2\sum_{j=1}^m \frac{\mu_j(s)}{s} + O(1/s) = \\ &= \sum_{j=1}^m \frac{\sigma_j(s)}{s} - \sum_{j=1}^m \left| \frac{i_s^-(\omega_{j+1}) - i_s^-(\omega_j)}{s} \right| + O(1/s). \end{aligned}$$

Now we notice that as  $s \rightarrow \infty$ , the function  $i_s^-/s$  converges uniformly to  $\lambda$ , thus:

$$(20) \quad \lim_{s \rightarrow \infty} \sum_{j=1}^m \left| \frac{i_s^-(\omega_{j+1}) - i_s^-(\omega_j)}{s} \right| + O(1/s) = \sum_{j=1}^m |\lambda(\omega_{j+1}) - \lambda(\omega_j)| = \int_{\omega_1}^{\omega_{m+1}} |\dot{\lambda}(\omega)| d\omega.$$

It remains to evaluate  $\lim_s \sum_j \frac{\sigma_j(s)}{s}$ . To this end we let  $\sigma_j^i(s)$  be the number of jumps of  $\lfloor s\lambda_i \rfloor$  on the interval  $I_j$ . We notice that  $\sigma_j^i(s) = \sum_{i=1}^d \sigma_j^i(s) + O(1)$ : in fact  $i_s^-$  jumps exactly when one of the  $\lfloor s\lambda_j \rfloor$  jumps and these function all jump at different points (except for the points where two eigenvalues are in resonance, but these are in finite number bounded independently of  $s$ ); we also notice that each function  $\lfloor s\lambda_i \rfloor / s$  converges uniformly to  $\lambda_i$ . Thus we get:

$$\begin{aligned} \lim_{s \rightarrow \infty} \sum_{j=1}^m \frac{\sigma_j(s)}{s} &= \lim_{s \rightarrow \infty} \sum_{j=1}^m \left( \sum_{i=1}^d \frac{\sigma_j^i(s)}{s} \right) = \lim_{s \rightarrow \infty} \sum_{j=1}^m \left( \sum_{i=1}^d \left| \frac{\lfloor s\lambda_i(\omega_{j+1}) \rfloor}{s} - \frac{\lfloor s\lambda_i(\omega_j) \rfloor}{s} \right| \right) = \\ &= \sum_{j=1}^m \left( \sum_{i=1}^d |\lambda_i(\omega_{j+1}) - \lambda_i(\omega_j)| \right) = \sum_{j=1}^m \int_{\omega_j}^{\omega_{j+1}} \left( \sum_{i=1}^d |\dot{\lambda}_i(\omega)| \right) d\omega = \\ (21) \quad &= \int_{\omega_1}^{\omega_{m+1}} \left( \sum_{i=1}^d |\dot{\lambda}_i(\omega)| \right) d\omega. \end{aligned}$$

Combining (20) and (21) we finally get:

$$\lim_{s \rightarrow \infty} \frac{b(\Omega_p^s)}{s} = \int_{\omega_1}^{\omega_{m+1}} \sum_{i=1}^d |\dot{\lambda}_i(\omega)| - \left| \sum_{i=1}^d \dot{\lambda}_i(\omega) \right| d\omega = \int_P \sum_{i=1}^d |\dot{\lambda}_i(\omega)| - \left| \sum_{i=1}^d \dot{\lambda}_i(\omega) \right| d\omega$$

where the last identity follows from the fact that on  $I_0$  and  $I_{m+1}$  the two functions  $\sum_{i=1}^d |\dot{\lambda}_i(\omega)|$  and  $\left| \sum_{i=1}^d \dot{\lambda}_i(\omega) \right|$  are equal. The limit of the statement simply follows by noticing that  $\alpha_i(\omega) = \alpha_i(-\omega)$  (i.e. the positive eigenvalues of  $i\omega A$  are  $\pi$ -periodic).  $\square$

As a corollary we get the following result: it says that the topology of the set of paths reaching the point  $\epsilon p$  with energy  $J \leq 1$  explodes; in particular the number of geodesics gets unbounded as well.

**Corollary 28.** *For a generic choice of  $W \subset \mathfrak{so}(d)$  and the point  $p \in \Delta^2$ :*

$$\lim_{\epsilon \rightarrow 0} b(\Omega_{\epsilon p} \cap \{J \leq 1\}) = \infty.$$

*Proof.* First notice that  $\Omega_{\epsilon p} \cap \{J \leq 1\}$  is homeomorphic to  $\Omega_p \cap \{J \leq 1/\epsilon\}$ ; thus we can apply the above theorem.

In order to prove the limit it is enough to show that for the generic choice of  $W$  and  $p$  the integral  $\tau(p)$  is not zero. Since the integrand function is always nonnegative, it's enough to prove

it doesn't vanish identically. Pick two distinct eigenvalues (functions), say  $i\alpha_1$  and  $i\alpha_2$ , for the family  $\{\omega A\}_{\omega \in S^1}$ . Since these functions are continuous semialgebraic,  $i\alpha_1$  has a maximum point  $\bar{\omega}$  and we can assume this is not a critical point for  $\alpha_2$  also (this is a generic condition). Then in a neighborhood of  $\bar{\omega}$  the derivatives of the corresponding  $\lambda_1$  and  $\lambda_2$  have different signs and the integrand is nonzero.  $\square$

We conclude the section with an example where the topological coarea formula can be computed directly.

*Example 4* (Commuting matrices, corank  $l = 2$ ). Let us fix the corank  $l = 2$ . If the matrices  $A_1$  and  $A_2$  commute, they can be written simultaneously in their canonical form

$$A_i = \text{diag}(v_1^i J_2, \dots, v_k^i, 0_h),$$

where as usual  $J_2$  is the  $2 \times 2$  symplectic matrix and  $0_h$  is the  $h \times h$  zero matrix (possibly with  $h = 0$ ). Setting  $v_j \doteq (v_j^1, v_j^2)$  and given  $\omega \in \mathbb{R}^{2*}$ , the eigenvalues of the matrix  $\omega A$  are  $\pm \langle \omega, v_j \rangle$ . Now we pick a generic  $p \in \Delta^2$ : having parametrized by  $t$  the unit circle in  $\mathbb{R}^{2*}$ , the functions we need in order to compute  $\tau(p)$  are  $\lambda_j(t) \doteq \left| \frac{\langle \omega(t), v_j \rangle}{\langle \omega(t), p \rangle} \right|$ , their derivatives being

$$\dot{\lambda}_j(t) = \frac{\text{sgn}(\langle \omega(t), v_j \rangle)}{\text{sgn}(\langle \omega(t), p \rangle)} \cdot \frac{\langle \dot{\omega}(t), v_j \rangle \langle \omega(t), p \rangle - \langle \omega(t), v_j \rangle \langle \dot{\omega}(t), p \rangle}{\langle \omega(t), p \rangle^2}.$$

Since the curve  $\omega(t)$  is the arc-length parametrization of the unit circle, for every  $t$  the covectors  $\omega(t), \dot{\omega}(t)$  form an orthonormal basis for  $(\Delta^2)^*$ ; it follows that the term

$$m_j \doteq \langle \dot{\omega}(t), v_j \rangle \langle \omega(t), p \rangle - \langle \omega(t), v_j \rangle \langle \dot{\omega}(t), p \rangle$$

is the determinant of the matrix  $\begin{pmatrix} p^1 & a_j^1 \\ p^2 & a_j^2 \end{pmatrix}$  which does not depend on  $t$ . Since the functions  $\lambda_j$  are periodic by  $\pi$ , we can modify the formula (18) by integrating on the semicircle  $\{\langle \omega(t), p \rangle > 0\}$  and eliminating the coefficient of  $1/2$  before the integral sign.

These two remarks allow us to simplify the expression of  $\dot{\lambda}(t)$ :

$$\dot{\lambda}_j(t) = \text{sgn}(\langle \omega(t), v_j \rangle) \cdot \frac{m_j}{\langle \omega(t), p \rangle^2}.$$

Now, in order to have the term  $\tau(p) > 0$ , the integrand of

$$\begin{aligned} \tau(p) &= \int_{\langle \omega(t), p \rangle > 0} \sum_{j=1}^d |\dot{\lambda}_j(\omega)| - \left| \sum_{j=1}^d \dot{\lambda}_j(\omega) \right| d\omega = \\ &= \int_{\langle \omega(t), p \rangle > 0} \frac{1}{\langle \omega(t), p \rangle^2} \cdot \left( \sum_{i=1}^d |m_i| - \left| \sum_{i=1}^d \text{sgn}(\langle \omega(t), v_i \rangle) m_i \right| \right) d\omega \end{aligned}$$

must be strictly positive somewhere. This happens if and only if there exist  $i$  and  $j$  such that for some  $t_0$  the terms  $\text{sgn}(\langle \omega(t_0), v_i \rangle) m_i$  and  $\text{sgn}(\langle \omega(t_0), v_j \rangle) m_j$  have opposite sign. If  $m_i \cdot m_j > 0$  we must have  $t_0$  such that  $\text{sgn}(\langle \omega(t), v_i \rangle) \cdot \text{sgn}(\langle \omega(t), v_j \rangle) < 0$ : if such a  $t_0$  does not exist, we have  $\text{sgn}(\langle \omega(t), v_i \rangle) \cdot \text{sgn}(\langle \omega(t), v_j \rangle) > 0$  for all  $t$  such that  $\langle \omega(t), p \rangle > 0$ , meaning that  $v_i$  is proportional to  $v_j$ ; a similar argument holds if  $m_i \cdot m_j < 0$ .

In order to have  $\tau(p) = 0$  the only possibility is that every  $v_i$  is proportional to every other  $v_j$ , which is equivalent to say that  $A_1$  and  $A_2$  are proportional: this implies that  $\dim \Delta^2 - \dim \Delta = 1$  which contradicts the hypothesis.

## APPENDIX A. STRATIFICATIONS OF $\mathfrak{so}(d)$

Here we construct a useful stratification of  $\mathfrak{so}(d)$ , generalizing the results from the appendix of [6]; for general results on stratifications and semialgebraic sets the reader is referred to [5].

We are interested in studying the dimensions of the semi-algebraic sets with generalized eigenvalues of given multiplicities and with given dimension of the kernel. Every skew-symmetric

matrix  $A$  can be written in its canonical form as a block matrix, with blocks on the diagonal of the form

$$\alpha J_2 = \begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix},$$

( $J_2$  being the canonical symplectic matrix in  $\mathfrak{so}(2)$ ) and a 0-block of the dimension of the kernel. By *generalized eigenvalues* we mean the entries like  $\alpha$ .

We introduce the set:

$$\Gamma_{k|m_1, \dots, m_r} \subset \mathfrak{so}(d)$$

defined to be the set of skew-symmetric matrices in  $\mathfrak{so}(d)$  with: (a) dimension of the kernel equal to  $k$  and (b) multiplicities of the generalized eigenvalues  $m_1, \dots, m_r$  (with  $m_1 \geq m_2 \geq \dots \geq m_r$ ).

By acting with  $SO(d)$  on a matrix  $A \in \Gamma_{k|m_1, \dots, m_r}$  with eigenvalues  $\alpha_1, \dots, \alpha_r$  (corresponding to the ordered multiplicities), we can put it in the form:

$$\text{Diag}(\alpha_1 J_{2m_1}, \dots, \alpha_r J_{2m_r}, 0_k),$$

Let us look at the stabilizer  $SO(d)_A$  of  $A$ : first of all it has to fix every eigenspace of  $A$ , since eigenspaces with different eigenvalues are orthogonal. On the kernel  $K$  the stabilizer is the restriction of  $SO(d)$  on  $K$ , i.e. a copy of  $SO(k)$ ; on the eigenspace with eigenvalue  $\alpha_i$  the restriction of  $SO(d)$  is a copy of  $SO(2m_i)$ , but the stabilizer has to fix the symplectic matrix; it follows that the stabilizer act as a copy of:

$$SO(2m_i) \cap Sp(2m_i) = U(m_i).$$

where  $J_{2n}$  is the symplectic matrix in  $\mathbb{R}^{2n}$  and  $0_k$  is the null matrix on  $\mathbb{R}^k$ .

Now it is possible to compute the codimension of the orbit  $\text{Ad}(SO(d))A$  of  $A$  by the adjoint action  $\text{Ad}$  of  $SO(d)$  on  $\mathfrak{so}(d)$  with known stabilizer:

$$\begin{aligned} \text{codim } \text{Ad}(SO(d))A &= \dim \mathfrak{so}(d) - \dim \text{Ad}(SO(d))A = \\ &= \dim \mathfrak{so}(d) - \dim SO(d) + \dim SO(d)_A = \dim SO(d)_A. \end{aligned}$$

We know that the stabilizer  $SO(d)_A$  is:

$$SO(d)_A = SO(k) \times U(m_1) \times \dots \times U(m_r),$$

and its dimension is:

$$\dim SO(d)_A = \frac{k(k-1)}{2} + \sum_{i=1}^r m_i^2.$$

Let us now consider the eigenvalues  $\alpha_i$ : as long as they are distinct (so they preserve their multiplicities) they are smooth functions of the matrices [11]. On the set  $\Gamma_{k|m_1, \dots, m_r}$  this condition holds true (by definition), hence we have a smooth map:

$$\psi : \Gamma_{k|m_1, \dots, m_r} \rightarrow \mathbb{R}^r$$

given by  $A \mapsto (\alpha_i)$ . This map is indeed a submersion on the open subset  $\mathcal{O}$  of vectors in  $\mathbb{R}^r$  with distinct entries. The fibers of the map  $\psi$  are the orbits of the adjoint action and they are diffeomorphic to a fixed manifold  $SO(d)/SO(d)_A$ ; in particular  $\Gamma_{k|m_1, \dots, m_r}$  is a fiber bundle over  $\mathcal{O}$  with fibers diffeomorphic to  $SO(d)/SO(d)_A$ . Now we can compute the codimension of  $\Gamma_{k|m_1, \dots, m_r}$  in  $\mathfrak{so}(d)$ :

$$(22) \quad \text{codim}_{\mathfrak{so}(d)} \Gamma_{k|m_1, \dots, m_r} = \text{codim}_{\mathfrak{so}(d)} \text{Ad}(SO(d))A - r = \frac{k(k-1)}{2} - r + \sum_{i=1}^r m_i^2$$

$$(23) \quad = \frac{k(k-1)}{2} + \sum_{i=1}^r (m_i^2 - 1).$$

Since we are interested in the matrices with integer eigenvalues, we will need to stratify  $\Gamma_{k|m_1, \dots, m_r}$  in infinite semialgebraic sets with given integer eigenvalues. Take  $\vec{n} = (n_1, \dots, n_r) \in \mathbb{N}^r$  with non-negative entries such that all the non-zero entries are distinct. Then by  $\Gamma_{k|m_1, \dots, m_r|\vec{n}}$  we will mean the stratum in  $\Gamma_{k|m_1, \dots, m_r}$  with the eigenvalue of multiplicity  $m_i$  equal to  $in_i$  if and

only if  $n_i > 0$ . The eigenvalues corresponding to zero entries of  $\vec{n}$  vary in  $\mathbb{R}$ . Since by fixing an eigenvalue we drop the dimension of the stratum by 1, we have the following:

**Proposition A.1.** *Given  $\vec{n} \in \mathbb{N}^r$  with the properties described above and with  $\nu$  non-zero entries, the submanifold  $\Gamma_{k|m_1, \dots, m_r|\vec{n}}$  has codimension  $\nu$  in  $\Gamma_{k|m_1, \dots, m_r}$ , thus its codimension in the set of all matrices  $\mathfrak{so}(d)$  is*

$$(24) \quad \text{codim}_{\mathfrak{so}(d)} \Gamma_{k|m_1, \dots, m_r|\vec{n}} = \frac{k(k-1)}{2} + \sum_{i=1}^r (\mu_i^2 - 1) + \nu.$$

## APPENDIX B. MORSE-BOTT FUNCTIONS

In this section we give a short review of Morse-Bott theory; the interested reader is referred to the original paper by Bott [7] and to the books [12] and [9] for more details (especially for the infinite dimensional case).

We recall that a function  $f : X \rightarrow \mathbb{R}$  on the Hilbert manifold  $X$  is called a *Morse-Bott function* if: (a) the critical set is the disjoint union of compact smooth manifolds; (b) if  $x$  is a critical point belonging to the critical manifold  $C$  then  $\ker \text{He}_x f = T_x C$ ; (c) for every sequence  $\{x_k\}$  on  $X$  such that  $f(x_k)$  is bounded and  $\|\nabla f_{x_k}\| \rightarrow 0$ , then the sequence  $\{x_k\}$  has limit points and every limit point is critical for  $f$ .

Condition (c) is usually referred as *Palais-Smale condition* and is automatically satisfied in the finite dimensional case. The smooth manifolds of critical points are called *nondegenerate critical manifolds*; notice that admitting only zero-dimensional critical manifolds we get classical Morse functions.

The second condition is equivalent to the non-degeneracy of the Hessian on the normal space  $N_x C$  for  $x \in C$ . The *index* of the critical manifold  $C$  is defined as the maximum of the dimensions of subspaces  $V \subset N_x C$  where the Hessian is negative-definite; since the Hessian is nondegenerate in the all normal bundle, this number does not depend on the point  $x \in C$  and it is denoted by  $\text{ind}(C)$ .

Under this assumptions it is still possible to describe what happens to the topology of the sublevels  $X^c \doteq f^{-1}(-\infty, c)$  of the Morse-Bott function  $f$  by increasing  $c$ .

Let us consider a Morse-Bott function  $f : X \rightarrow \mathbb{R}$  together with a Riemannian metric  $g$  on  $X$  (the choice of  $g$  it's only a technical convenience and in fact the following results do not depend on it). The first fundamental theorem of Morse theory describe how the sublevels (don't) change when  $c$  increases without passing critical values:

**Theorem B.1.** *If  $f : X \rightarrow \mathbb{R}$  is a Morse-Bott function on  $X$  satisfying the Palais-Smale condition and the interval  $[a, b] \subset \mathbb{R}$  does not contain critical values,  $X^b$  is diffeomorphic to  $X^a$ .*

Roughly speaking  $X^b$  is deformed to  $X^a$  along the integral curves of the gradient flow  $\nabla f$ . Moreover if we let  $a$  be a critical value for the Morse-Bott function  $f$  the sublevels aren't diffeomorphic one to another anymore, but still there exists a deformation.

**Theorem B.2.** *If  $f : X \rightarrow \mathbb{R}$  is a  $C^1$  function satisfying the Palais-Smale condition and  $(a, b] \subset \mathbb{R}$  does not contain critical values,  $X^a$  is a strong deformation retract of  $X^b$ .*

For the proof see Lemma 3.2 in Chapter 1 of [9].

It only remains to recall what happens when we pass a critical value for a Morse-Bott function. Given a critical manifold  $C$  we restrict the tangent bundle  $TX$  to  $C$  and consider the sub-bundle:

$$E_C^- = \{\text{directions where the Hessian of the function } f \text{ is negative-definite}\}.$$

together with the unit disk bundle  $D_C^- \subset E_C^-$ . With this notation the following theorem generalizes the classical one (the statement we present here is actually the one in [12]).

**Theorem B.3** (Bott). *Let  $f : X \rightarrow \mathbb{R}$  be a Morse-Bott function, and let  $c$  be a critical value. For  $\delta > 0$  sufficiently small the sublevel  $X^{c+\delta} = X \cap \{f \leq c + \epsilon\}$  is homotopic to the sublevel  $X^{c-\epsilon}$  with the unit disk bundle  $D_C^-$  glued along the boundary.*

Since in our setting we consider homology with  $\mathbb{Z}_2$  coefficients, it follows from the Thom isomorphism that:

$$H_*(X^{c+\delta}, X^{c-\delta}) \simeq H_*(D_C^-, \partial D_C^-),$$

where in the last equation we allow the critical manifold  $C$  to be nonconnected, in which case we actually have a disjoint union of different bundles (with possibly different rank, corresponding to the possibly different indexes of the components of  $C$ ).

Moreover we can state *Morse-Bott inequalities* in terms of the Poincaré polynomial of  $X^s$  and Morse (Bott) polynomial of  $f$ , which is defined by

$$M_f^s(t) = \sum_{\{C \mid f(C) \leq s\}} P_C(t) t^{\text{ind}(C)},$$

where the sum is taken amongst the critical manifolds  $C$  contained in  $\{f \leq s\}$ ; evaluations of this sum at  $t = 1$  gives the following.

**Proposition B.4** (Morse-Bott inequalities).

$$b(X^s) \leq \sum_{\{C \mid f(C) \leq s\}} b(C).$$

### APPENDIX C. THE COHOMOLOGY OF THE INTERSECTION OF REAL QUADRICS

In this section we present useful results from [3, 13, 14] for the study of the Betti numbers of the intersection of real quadrics on the sphere  $S^n$ . The motivating example is the case of the zero locus  $Y$  of one single nondegenerate quadratic form  $q$  on the sphere  $S^n$ : if  $i^-(q)$  denotes the negative inertia index of  $q$ , then:

$$Y \simeq S^{i^-(q)-1} \times S^{n-i^-(q)}.$$

In particular we see that the knowledge of the *index* function on the whole line spanned by  $q$  in the space of all quadratic forms determines the topology (since by nondegeneracy  $n - i^-(q) = i^-(-q) - 1$ ).

More generally if we have  $l$  quadratic forms  $q_1, \dots, q_l$  in  $n + 1$  variables, then we consider the function  $\eta \mapsto i^-(\eta q)$ ; in the generic case this function is the restriction of the negative inertia index function to the span of  $q_1, \dots, q_l$  in the space of all quadratic forms. Although the general theory is more detailed, for our purposes we need explicit computations only in the case  $l = 2$  and in the general case it will suffice to have quantitative bounds on the topology of:

$$Y = \{x \in S^n \mid q_1(x) = \dots = q_l(x) = 0\}.$$

In the case  $l = 2$  we consider a unit circle  $S^1$  in  $\mathbb{R}^2$  and the restriction  $i^-|_{S^1}$ ; also for  $j \geq 0$  we let:

$$P_j = \{\eta \in S^1 \mid i^-(\eta q) \leq j\}.$$

The following formula (25) is proved in [13] and relates the Betti numbers of  $Y$  to the topology of the sets  $P_j$ ; we denote by  $\tilde{b}_j(Y)$  the rank of  $\tilde{H}^j(Y; \mathbb{Z}_2)$ . The general bound (26) for the topology of  $Y$  in the case  $l \geq 2$  is proved in [14]. The reader is referred to [3, 13, 14] for more details.

**Proposition C.1.** *If  $Y$  is the intersection of two quadrics on the sphere  $S^n$  and  $0 \leq j \leq n - 3$ , then:*

$$(25) \quad \tilde{b}_j(Y) = \tilde{b}_{n-j-1}(S^n \setminus Y) = b_0(P_{j+1}, P_j) + b_1(P_{j+2}, P_{j+1}).$$

*Moreover if  $Y$  is defined by  $l \geq 2$  quadratic equations on  $S^n$ , on  $\mathbb{R}P^n$  or in  $\mathbb{R}^n$ , then:*

$$(26) \quad b(Y) \leq O(n)^{l-1}.$$

It is possible to apply the above technique also in the case  $Y$  is the intersection of quadrics on the unit sphere in some (infinite dimensional) Hilbert space  $H$ . The main differences for this infinite dimensional case are the following:  $Y$  must be nonsingular;  $\tilde{b}_i$  denotes the rank of the  $i$ -th Čech cohomology group; the negative inertia index might be infinite for some  $\eta \in S^1$ , but these  $\eta$  are already excluded by the condition  $i^-(\eta) \leq j < \infty$ . With these modification we have

the following result from [1]; formula (28) is the analogue of (25), but the condition that  $H$  is infinite dimensional allows to remove the restriction on the range for  $j$ .

**Theorem C.2.** *Let  $q_1, \dots, q_l$  be continuous quadratic forms on the Hilbert space  $H$  and  $Y$  be their (nondegenerate) common zero locus on the unit sphere. Then:*

$$(27) \quad H_*(Y) = \varinjlim_{V \in \mathcal{F}} \{H_*(Y \cap V)\}.$$

where  $\mathcal{F}$  denotes the family of all finite dimensional subspace of  $H$ . Moreover in the case  $l = 2$ :

$$(28) \quad \tilde{b}_j(Y) = \check{b}_0(P_{j+1}, P_j) + \check{b}_1(P_{j+2}, P_{j+1})$$

*Sketch.* Since  $Y$  is assumed to be nonsingular, then it has a tubular neighborhood  $U$  in  $H$  and  $H_*(Y) \simeq H_*(U)$ . In particular every singular chain in  $U$  is homotopic to one whose image is contained in a finite dimensional subspace and (27) follows. To prove (28) we fix a  $j \geq 0$ ; then using (25) we have:

$$\tilde{b}_j(Y \cap V) = b_0(P_{j+1}(V), P_j(V)) + b_1(P_{j+2}(V), P_{j+1}(V))$$

where  $V \subset H$  is a sufficiently big finite dimensional subspace (the condition required on the dimension is  $\dim(V) - 3 \geq j$ ) and  $P_j(V) = \{\eta \in S^1 \mid i^-(\eta q|_V) \leq j\}$ . Now the sets  $\{P_j(V)\}_{V \in \mathcal{F}}$  are also partially ordered by inclusion: if  $V_1 \subset V_2$ , then  $P_j(V_2) \subset P_j(V_1)$ . It is not difficult to show that under the isomorphism  $\tilde{H}_j(Y \cap V) \simeq H^0(P_{j+1}(V), P_j(V)) \oplus H^1(P_{j+2}(V), P_{j+1}(V))$  the inclusion morphism on the homology  $H_j(Y \cap V_1) \rightarrow H_j(Y \cap V_2)$  is induced by the restriction morphism (see [1]):

$$\bigoplus_{i=0,1} H^i(P_{j+i+1}(V_1), P_{j+i}(V_1)) \rightarrow \bigoplus_{i=0,1} H^i(P_{j+i+1}(V_2), P_{j+i}(V_2))$$

Since the sets  $\{P_j(V)\}_{V \in \mathcal{F}}$  are Euclidean Neighborhood Retracts (being semialgebraic sets), then by the continuity property of Čech cohomology:

$$\tilde{H}_*(Y) = \varinjlim_{V \in \mathcal{F}} \{H^*(P_{j+1}(V), P_j(V))\} = \check{H}^*\left(\bigcap_{V \in \mathcal{F}} P_{j+1}(V), \bigcap_{V \in \mathcal{F}} P_j(V)\right).$$

Finally  $P_j$  equals by construction  $\bigcap_{V \in \mathcal{F}} P_j(V)$  and the conclusion follows. □

**Acknowledgements.** The first author is supported by Grant of the Russian Federation for the State Support of Researches (Agreement No 14.B25.31.0029)

## REFERENCES

- [1] A. A. Agrachev: *Topology of quadratic maps and Hessians of smooth maps*, Itogi nauki. VINITI. Algebra. Topologiya. Geometriya, 1988, v.26, 85–124 (Translated in J. Soviet Math. 49, 1990, no. 3, 990–1013).
- [2] A. A. Agrachev, U. Boscain, D. Barilari: *Introduction to Riemannian and sub-Riemannian geometry*, lectures notes, available at: <http://www.cmapx.polytechnique.fr/~barilari/Notes.php>
- [3] A. A. Agrachev, A. Lerario: *Systems of quadratic inequalities*, Proceedings of the London Mathematical Society, (2012) 105 (3).
- [4] D. Barilari, U. Boscain, J. P. Gauthier: *On 2-step, corank 2 sub-Riemannian metrics*, SIAM Journal of Control and Optimization, 50 (2012), no. 1, 559–582
- [5] J. Bochnak, M. Coste, M-F. Roy: *Real Algebraic Geometry*, Springer-Verlag, 1998.
- [6] U. Boscain, J. P. Gauthier: *On the Spherical Hausdorff Measure in Step 2 Corank 2 sub-Riemannian Geometry*, arXiv:1210.2615
- [7] R. Bott: *Nondegenerate critical manifolds*, Annals of Mathematics, Second Series, Vol. 60, No. 2 (Sep., 1954), pp. 248–261
- [8] R. Bott: *Lectures on Morse Theory, old and new* Bulletin of the American Mathematical Society, Volume 7, Number 2, September 1982
- [9] K. Chang: *Infinite dimensional Morse Theory and Multiple Solution Problems*, Birkhäuser, 1993
- [10] Z. Ge: *Horizontal path spaces and Carnot-Carathéodory metrics*, Pacific Journal of Mathematics, volume 161, no.2, 1993
- [11] T. Kato: *Perturbation theory for linear operators*, Springer, 1995
- [12] W. Klingenberg: *Lectures on closed geodesics*, Springer, 1978

- [13] A. Lerario: *Convex Pencils of real quadratic forms*, Discrete and computational Geometry, Volume 48, Number 4 (2012), 1025-1047.
- [14] A. Lerario: *Complexity of intersection of real quadrics and topology of symmetric determinantal varieties*, arXiv:1211.1444
- [15] J. W. Milnor: *Morse theory*, Princeton University Press
- [16] A. L. Onishchick, E. B. Vinberg: *Lie groups and algebraic groups*, Springer, 1980
- [17] R. Palais: *Morse theory on Hilbert manifolds*
- [18] P. Pansu *Metriques de Carnot-Caratheodory et quasi isometries des espaces symetriques de rang un*, Annals of Math. 129 , 1-60, 1989
- [19] S. Semmes: *An Introduction to Heisenberg Groups in Analysis and Geometry*, Notices of the AMS, volume 50, number 6.