LOCAL CONTACT SUB-FINSLERIAN GEOMETRY FOR MAXIMUM NORMS IN DIMENSION 3

ENTISAR A.-L. ALI
Univ. Grenoble Alpes, CNRS, Institut Fourier
F-38000 Grenoble, France
and
Diyala University, Baquba, Diyala Province, Iraq

G. CHARLOT
Univ. Grenoble Alpes, CNRS, Institut Fourier
F-38000 Grenoble, France

(Communicated by Ugo Vittorio Boscain)

Abstract. The local geometry of sub-Finslerian structures in dimension 3 associated with a maximum norm is studied in the contact case. A normal form is given. The short extremals, the local switching, conjugate and cut loci, and the small spheres are described in the generic case.

1. Introduction. From a geometric point of view, the sub-Finslerian (SF) structure we are interested in here is a triplet $(\mathcal{M}, \Delta, |.|_\infty)$ where $\mathcal{M}$ is a connected manifold, $\Delta$ is a sub-bundle of the tangent bundle, and $|.|_\infty$ is a maximum norm on $\Delta$. With such a structure we can define

**Definition 1.** Let $\gamma : [0, T] \to \mathcal{M}$ be an absolutely continuous curve in $\mathcal{M}$. It is said admissible if $\dot{\gamma}(t) \in \Delta_{\gamma(t)}$ a.e. The length of an admissible $\gamma$ is defined as

$$\ell(\gamma) := \int_0^T |\dot{\gamma}(t)|_\infty dt.$$

If $\text{Lie}_q(\Delta) = T_q\mathcal{M}$ for any $q$ then, for any couple of points $(q_1, q_2)$, it exists an admissible curve joining $q_1$ and $q_2$. In this last case, we define the distance between two points $p$ and $q$ in $\mathcal{M}$ as the infimum of the lengths of the curves joining $p$ to $q$

$$d(p, q) = \inf\{\ell(\gamma) \mid \gamma(t) \in \Delta_{\gamma(t)} \text{ a.e., } \gamma(0) = p, \gamma(T) = q\}.$$

When we are only interested in local issues, we can define the structure by the data of $k$ linearly independent vector fields $F_1, \cdots, F_k$ and by the standard maximum norm defined on $\text{span}(F_1, \cdots, F_k)$ by

$$|G| = \max_i \{|u_i| \mid G = \sum_i u_i F_i\}.$$
For the control point of view, we are considering the dynamics
\[ \frac{dq}{dt}(t) = \sum_{i=1}^{k} u_i(t) F_i(q(t)), \quad \forall t \in [0, T] \quad (1) \]
with the constraints
\[ |u_i(t)| \leq 1, \forall i \leq k, \forall t \in [0, T] \quad (2) \]
and we are interested in the optimal synthesis for the problem of minimizing time.

In this situation \( \Delta = \text{span}\{F_1, \ldots, F_k\} \).

In this article, we are interested only in the local version of this problem, that is to understand the local synthesis for small times (or small distances). Moreover we fix our attention on the case of constant rank of smallest dimension namely \( \text{dim}(M) = 3, k = 2 \). Hence, in the following we work in the neighborhood of \( 0 \in \mathbb{R}^3 \).

We say that a property is generic for this class of sub-Finslerian metrics if it is true on a residual set of couples of vector fields (defining the metric) for the \( C^{\infty} \)-Whitney topology on the set of couple of vector fields. Genericity is usually proven using the Thom transversality theorem (see [26], page 82). It is a classical fact that generically the set of points \( q \) where a distribution \( \Delta \) of dimension 2, on a manifold of dimension 3, satisfies \( [\Delta, \Delta]_q = \Delta_q \) is a sub-manifold of dimension 2 called the Martinet surface. Outside this surface, the distribution is contact: \( [\Delta, \Delta]_q = T_q M \).

We are interested in describing the local objects, such as the optimal trajectories, the cut locus, the conjugate locus, the switching locus and the small spheres at contact points.

Few publications exist about sub-Finslerian geometry since it is a new subject. Let mention the works [22, 23] for dimension 3, considering norms which are smooth outside the zero section. In [19], the sphere of a left invariant sub-Finsler structure associated to a maximum norm in the Heisenberg group is described. In the preprint [7], the authors describe the extremals (and discuss in particular their number of switches before the loss of optimality) for the Heisenberg, Grushin and Martinet distributions. In [6], the authors describe, in the 2d generic case, the small spheres and the local cut locus. In this last preprint, the distribution is not supposed of constant rank and it can be related to the almost-riemannian case, see [2, 18, 17, 14, 16].

The work we propose here is a continuation of what has been done in sub-Riemannian geometry at the end of the nineties for the codimension one distributions in the contact, quasi-contact and Martinet cases (see [1, 15, 13, 24, 4, 20]). These works, in addition to the interest of understanding the local geometry, were in particular motivated by results on the heat kernel asymptotics in the sub-Riemannian context (see [12, 27, 28, 11]). They allowed recently to give new results on the asymptotics of the heat kernel associated to the metric (see [9, 8]).

In section 2, we construct a normal form for couples \( (F_1, F_2) \) defining the contact distribution \( \Delta \). In section 3, we establish some properties of the minimizing trajectories and construct the exponential map. In section 4 we present the optimal synthesis of the nilpotent case. In section 5, we present the jets of the extremals, the switching times and the conjugate times, the switching locus and the conjugate locus for extremals whose controls are both switching in short time. In section 6, we calculate the cut locus generated by these extremals, similar to the cut locus in the 3D contact sub-Riemannian case. In section 7 we discribe the part of the
optimal synthesis linked to extremals with only one control switching several times, which is very different from the sub-Riemannian case.

2. Normal form. In this section, the goal is to construct a normal form for the couple \((G_1, G_2)\) defined by \(G_1 = F_1 + F_2\) and \(G_2 = F_1 - F_2\). As we will see later, \(\pm G_1\) and \(\pm G_2\) are the velocities of the non singular extremals of the optimal control system defined by (1) and (2).

Since we consider only points \(q\) where the distribution is contact then \(G_1, G_2\) and \([G_1, G_2] = -2[F_1, F_2]\) form a basis of \(T_q\mathbb{R}^3\). Hence, we can build a coordinate system centered at \(q\), by the following way. Denoting \(e^tX\) the flow at time \(t\) of a vector field \(X\), we can define

\[
\Xi : (x, y, z) \mapsto e^{xG_1}e^{yG_2}e^{z[G_1, G_2]}q,
\]

which to \((x, y, z)\) associates the point reached by starting at \(q\) and following \([G_1, G_2]\) during time \(z\), then \(G_2\) during time \(y\) and finally \(G_1\) during time \(x\). The map \(\Xi\) is smooth and satisfies

\[
\frac{\partial \Xi}{\partial x}(x, y, z) = G_1, \quad \frac{\partial \Xi}{\partial y}(0, y, z) = G_2, \quad \text{and} \quad \frac{\partial \Xi}{\partial z}(0, 0, z) = [G_1, G_2].
\]

As a consequence \(\Xi\) is not degenerate at \((0, 0, 0)\) and defines a coordinate system in a neighborhood of \(q\). Such coordinates are called normal coordinates and \(G_1\) and \(G_2\) satisfy

\[
G_1(x, y, z) = \partial_x, \quad G_2(x, y, z) = x\epsilon_x(x, y, z)\partial_x + (1 + x\epsilon_y(x, y, z))\partial_y + x(1 + \epsilon_z(x, y, z))\partial_z
\]

where \(\epsilon_x, \epsilon_y, \epsilon_z\) are smooth functions satisfying \(\epsilon_x(0, 0, z) = \epsilon_y(0, 0, z) = \epsilon_z(0, 0, z) = 0\). Hence we can give the following expressions of \(G_2\)

\[
G_2(x, y, z) = (a_{200}x^2 + a_{110}xy + x\theta_x(x, y, z))\partial_x + (1 + b_{200}x^2 + b_{110}xy + x\theta_y(x, y, z))\partial_y
\]

\[
+ (x + c_{200}x^2 + c_{110}xy + c_{100}x^3 + 210x^2y + c_{120}xy^2 + x\theta_z(x, y, z))\partial_z
\]

where \(\theta_x, \theta_y\) and \(\theta_z\) are smooth functions such that \(\theta_x(0, 0, z) = \theta_y(0, 0, z) = \theta_z(0, 0, z) = 0\) and whose Taylor series of respective order 1, 1, 2 are null with \(x\) and \(y\) of order 1 and \(z\) of order 2.

3. General facts about the computation of the optimal synthesis. In the following of the paper we are going to study the local geometry for a generic class of 3D sub-Finslerian metric defined by a maximum norm, that is for a residual set for the Whitney \(C^\infty\) topology on the set of such metrics. And, for this residual set of metrics, we are going to consider the local geometry only at points in the complement of a finite union of codimension 1 submanifolds.

For example, we consider only contact points (points \(q\) where \([\Delta, \Delta]_q = T_qM\)) when generically the Martinet surface (points \(q\) where \([\Delta, \Delta]_q = \Delta_q\) has codimension 1. We may also ask that an invariant appearing in the normal form is not zero, which happens outside a codimension 1 submanifold. All along the paper we will assume only a finite number of such assumptions.
3.1. Controllability and existence of minimizers. The contact hypothesis is 
span\(F_1, F_2, [F_1, F_2]\) = \(\mathbb{R}^3\).

Hence, as a consequence of Chow-Rashevski theorem (see [5, 29, 21]), such a control 
system is controllable that is, for any two points, it always exists an admissible curve 
joining the two points.

Moreover, since at each point the set of admissible velocities is convex and com-
 pact (in the control version), thanks to Filippov theorem (see [5, 25]), for any two 
points, it always exists at least a minimizer.

3.2. Pontryagin maximum principle (PMP) and switching functions. The Pontryagin 
Maximum Principle (PMP) gives necessary conditions for a curve to be 
a minimizer of the SF distance. For our problem it takes the following form.

**Theorem 2 (PMP).** Let define the Hamiltonian:

\[ H(q, \lambda, u, \lambda_0) = u_1\lambda F_1(q) + u_2\lambda F_2(q) + \lambda_0 \]

where \(q \in \mathbb{R}^3\), \(\lambda \in T^*\mathbb{R}^3\), \(u \in \mathbb{R}^2\) and \(\lambda_0 \in \mathbb{R}^+\). For any minimizer \((q(t), u(t))\) there
exist a never vanishing Lipschitz continuous covector \(\lambda : t \mapsto \lambda(t) \in T^*\mathbb{R}^3\) and a constant \(\lambda_0 \leq 0\) such that for a.e. \(t \in [0, T]\) we have

1. \(\frac{d\lambda}{dt}(t) = \frac{\partial H}{\partial q}(q(t), \lambda(t), u(t), \lambda_0)\),
2. \(\frac{d\lambda}{dt}(t) = -\frac{\partial H}{\partial \lambda}(q(t), \lambda(t), u(t), \lambda_0)\),
3. \(H(q(t), \lambda(t), u(t), \lambda_0) = \max\{H(q(t), \lambda(t), v, \lambda_0) \mid \max_{i=1,2}|v_i| \leq 1\}\),
4. \(H(q(t), \lambda(t), u(t), \lambda_0) = 0\).

If \(\lambda_0 = 0\) then \(q\) is said abnormal, if not \(q\) is said normal. A solution of the PMP is 
called an extremal.

**Remark 3.** It is well known that for a contact distribution there is no non-trivial 
abnormal extremal. In the following we fix \(\lambda_0 = -1\).

In the following, we will have to consider the vector fields \(F_3 = [F_1, F_2]\), \(F_4 = 
[F_1, [F_1, F_2]]\) and \(F_5 = [F_2, [F_1, F_2]]\). We can now define

**Definition 4.** For an extremal triplet \((q(\cdot), \lambda(\cdot), u(\cdot))\), we define the functions 
\(\phi_i(t) = \langle \lambda(t), F_i(q(t)) \rangle, i = 1 \cdots 5\).

The functions \(\phi_1\) and \(\phi_2\) are called the switching functions. All the \(\phi_i\) are absolutely continuous functions.

**Proposition 5.** For \(i = 1, 2\)

1. If \(\phi_i(t) > 0\) (resp \(\phi_i(t) < 0\)) then \(u_i(t) = 1\) (resp \(u_i(t) = -1\)).
2. If \(\phi_i(t) = 0\) and \(\frac{d\phi_i}{dt}(t) > 0\) (resp \(\frac{d\phi_i}{dt}(t) < 0\)) then \(\phi_i\) changes sign at time \(t\) and the control \(u_i\) switches from \(-1\) to \(+1\) (resp from \(+1\) to \(-1\)).

**Proof.** Point 1 is a direct consequence of the maximality condition of the PMP.
Point 2 is a direct consequence of point 1. \(\Box\)

**Remark 6.** One can compute easily that along a bang arc

\[ \frac{d\phi_1}{dt} = -u_2\phi_3\quad \text{and} \quad \frac{d\phi_2}{dt} = u_1\phi_3.\]

and moreover, since \((F_1, F_2, F_3)\) is a frame of the tangent space, we can define the function \(f_{ij}\) for \(i = 4, 5\) and \(j = 1, 2, 3\) by setting

\[ F_4 = [F_1, [F_1, F_2]] = f_{41}F_1 + f_{42}F_2 + f_{43}[F_1, F_2], \]
Definition 7. We call a null switching function $\phi$ vanish (with constant controls with value 1 or −1) and bang-bang an extremal which is a finite concatenation of bangs. We call $u_i$-singular an extremal corresponding to a null switching function $\phi_i(.)$. A time $t$ is said to be a switching time if $u$ is not bang in any neighborhood of $t$.

Remark 8. Notice that the switching functions $\phi_i(.)$ are at least Lipschitz continuous. Moreover thanks to condition 4 of PMP and $\lambda_0 = -1$ we have that $u_1(t)\phi_1(t) + u_2(t)\phi_2(t) = 1$, for all $t$ which implies

$|$\phi_1(t)\vert + |\phi_2(t)\vert = 1.$

Remark 9. Along a $u_1$-singular, $\phi_1 \equiv 0$, $\phi_3 \equiv 0$ and $|\phi_2| \equiv 1$. If $\phi_2 \equiv \pm 1$ then $u_2 \equiv \pm 1$ and, thanks to equation (4), we get that

$u_1f_{42} \pm f_{52} \equiv 0,$

which determines entirely the control $u_1$ only provided that $f_{42} \neq 0$ and $|f_{52}| \leq |f_{42}|$.

3.3. Change of coordinates. We first concentrate our attention on extremals with initial $|\lambda_z|$ very large corresponding to short cut times (as we will see later).

Following the techniques used in the 3d-contact case in sub-Riemannian geometry (see Agrachev et al [3]), one can make the following change of coordinates and time

$r = \frac{1}{\lambda_z}, \quad s = \int_0^t \frac{1}{r(\tau)} d\tau, \quad p_x = r\lambda_x, \quad p_y = r\lambda_y.$

Denoting $p = (p_x, p_y, 1)$ and $q = (x, y, z)$ one gets the equations for the extremals

$\dot{q} = \frac{dq}{ds} = r(u_1F_1(q) + u_2F_2(q)),$

$\dot{p} = \frac{dp}{ds} = p(-p(u_1dF_1(q) + u_2dF_2(q)) + (p(u_1\frac{\partial F_1(q)}{\partial z} + u_2\frac{\partial F_2(q)}{\partial z}))p),$ 

$\dot{r} = \frac{dr}{ds} = r^2(p(u_1\frac{\partial F_1(q)}{\partial z} + u_2\frac{\partial F_2(q)}{\partial z})).$

In the following, we denote $\dot{g}$ the derivate with respect to the new time $s$ of a function $g$.

3.4. Exponential map and conjugate locus. The set of initial condition is determined by

$H = u_1\lambda(0)F_1(0) + u_2\lambda(0)F_2(0) - 1 = 0$

which implies max{|$\lambda_x(0)$|, |$\lambda_y(0)$|} = 1. This implies that max{|$p_x(0)$|, |$p_y(0)$|} = r(0).

If an extremal is not singular, then it starts by a first bang and hence by the speed $\pm G_1$ or $\pm G_2$. Assume $r_0 > 0$. If the first bang follows $\pm G_1$ then $p_x(0) = \pm r_0$ and we define $\alpha_2$ by setting $p_y(0) = \mp r_0 \alpha_2$ with $\alpha_2 \in [-1, 1]$. If the first bang follows $\pm G_2$ then $p_y(0) = \pm r_0$ and we define $\alpha_1$ by setting $p_x(0) = \pm r_0 \alpha_1$ with $\alpha_1 \in [-1, 1]$. With this convention, among the extremals starting with $r_0$ fixed and
following $\pm G_1$ (resp $\pm G_2$), the last one to switch is the one with initial condition $\alpha_2 = 1$ (resp. $\alpha_1 = 1$).

We can hence define 4 exponential maps corresponding to the 4 initial speed $\pm G_1$ and $\mp G_2$ and describing the bang-bang extremals. For these maps, depending on $r_0$, $\alpha_i$, and $s$, when $\alpha_i \neq 1$ and when $s$ is not a switching time of the extremal with initial condition $(r_0, \alpha_i)$, we can compute the Jacobian with respect to the parameters $(r_0, \alpha_i, s)$.

Recall that we denote by $t$ the time and $s$ the new time after reparameterization.

**Definition 10.** The first conjugate time along an extremal is the infimum of the times $t$ such that there exist $t_1$ and $t_2$ with $0 < t_1 < t_2 < t$ such that $\text{Jac}(t_1)\text{Jac}(t_2) < 0$. The first conjugate point along an extremal is the point reached at first conjugate time and the first conjugate locus is the set of the first conjugate points.

The cut locus is the set of points where an extremal curve loses optimality.

The Maxwell set is the set of points where two optimal curves meet.

The sphere at time $t$ is the collection of all end points at time $t$ of the optimal extremals.

**Remark 11.** With this definition, it will happen that the Maxwell set is not always included in the cut locus (which is very different from the Riemannian case) : as we will see in particular in the nilpotent case, different optimal curves may join and stay the same during a non trivial interval of time, staying optimal. It is a classical phenomena with such metrics built with a non strictly convex norm.

4. Nilpotent case. This part is not entirely new since this case has been studied in [7, 19].

As in sub-Riemannian geometry (see [10, 3]), the nilpotent approximation plays an important role as “good estimation” of the real situation. The nilpotent approximation at $(0, 0, 0)$ of $G_1$, $G_2$ given in the normal form is

$$\hat{G}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \hat{G}_2 = \begin{pmatrix} 0 \\ 1 \\ x \end{pmatrix}.$$ 

It is a left invariant sub-Finslerian metric defined on the Heisenberg group with the representation

$$(x, y, z) \ast (x', y', z') = (x + x', y + y', z + z' + xy').$$

It is the tangent space in the sense of Gromov. See [10].

The Hamiltonian for the nilpotent case is

$$H = \frac{u_1 + u_2}{2} \lambda_x + \frac{u_1 - u_2}{2} (\lambda_y + x \lambda_z) - 1.$$ 

Thus the Hamiltonian equations are given by

$$\frac{dx}{dt} = \frac{u_1 + u_2}{2}, \quad \frac{d\lambda_x}{dt} = -\frac{u_1 - u_2}{2} \lambda_z,$$

$$\frac{dy}{dt} = \frac{u_1 - u_2}{2}, \quad \frac{d\lambda_y}{dt} = 0,$$

$$\frac{dz}{dt} = \frac{u_1 - u_2}{2} x, \quad \frac{d\lambda_z}{dt} = 0,$$

which implies that $\lambda_y$ and $\lambda_z$ are constants.

Before entering the computations, one can think that, thanks to the PMP, most of the optimal trajectories will be concatenations of bang arcs of $\pm G_1$ and $\pm G_2$. 

Moreover, one shows relatively easily that the extremals are solutions of an isoperimetric problem, the $z$ coordinate of an optimal curve from 0 being a certain area defined from the projection on the $(x, y)$-plane of the trajectory, as it is in the Heisenberg case in subriemannian geometry. Hence it seems clear that many optimal curves project on squares. As we will see, a large class of optimal curves satisfy this property but many others, the singular ones, do not satisfy it which is very different from the subriemannian case.

4.1. **Extremals with** $\lambda_z \neq 0$. Following the change of coordinates of section 3.3, and still denoting $\dot{g}$ the derivative of a function $g$ with respect to $s$, we get

\[
\begin{align*}
\dot{x} &= r \frac{u_1 + u_2}{u_2}, \\
\dot{y} &= r \frac{u_1 - u_2}{u_2}, \\
\dot{z} &= r \frac{2 - u_2}{u_2} x, \\
\dot{r} &= 0.
\end{align*}
\]

Let present, for example, the computation of extremals with $\lambda_z \equiv \lambda_z(0) > 0$, $\lambda_y \equiv \lambda_y(0) = 1$, $\lambda_x(0) \in [-1, 1]$. In $x, y, z, p_x, p_y, r, s$ coordinates, one gets $p_y = r$, $p_x = r \alpha$ with $\alpha \in [-1, 1]$ and $\phi_1(s) = p_x(s) + p_y(s) + 2(s)$ and $\phi_2(s) = p_y(s) - z(s)$. We denote $s_1, s_2$, etc. the sequence of switching times along an extremal. During the first bang, since $\phi_1(0) = \frac{\alpha r - p_x}{2r} > 0$ hence $u_1 = 1$, and since $\phi_2(0) = \frac{\alpha r - p_y}{2r} < 0$ and $\phi_2(0) = -\frac{u_2^2}{2} \lambda_z < 0$, the controls satisfy $u_1 = 1$ and $u_2 = -1$. Moreover

\[
\begin{align*}
x(s) &= 0, \\
p_x(s) &= r \alpha_1 - rs, \\
\phi_1(s) &= \frac{\alpha_1 - s + 1}{2}, \\
p_y(s) &= p_y(0) = r, \\
\phi_2(s) &= \frac{\alpha_1 - s - 1}{2}, \\
z(s) &= 0.
\end{align*}
\]

The first switching time $s_1$ corresponds to $\phi_1(s_1) = 0$ hence $s_1 = 1 + \alpha_1$.

During the second bang, the controls satisfy $u_1 = -1$ and $u_2 = -1$ and

\[
\begin{align*}
x(s) &= -sr + r + \alpha_1 r, \\
p_x(s) &= -r, \\
\phi_1(s) &= \frac{-s + 1 + \alpha_1}{2}, \\
p_y(s) &= p_y(0) = r, \\
\phi_2(s) &= \frac{s - 3 - \alpha_1}{2}, \\
z(s) &= 0.
\end{align*}
\]

The second switching time $s_2$ corresponds to $\phi_2(s_2) = 0$ hence $s_2 = 3 + \alpha_1$.

Along the third bang, the controls satisfy $u_1 = -1$ and $u_2 = 1$ and

\[
\begin{align*}
x(s) &= -2r, \\
p_x(s) &= -\alpha_1 r - 4r + sr, \\
\phi_1(s) &= \frac{-\alpha_1 - 5 + s}{2}, \\
p_y(s) &= p_y(0) = r, \\
\phi_2(s) &= \frac{-\alpha_1 - 3 + s}{2}, \\
z(s) &= 2r(3 + 3 + \alpha_1)).
\end{align*}
\]

The third switching time $s_3$ corresponds to $\phi_1(s_3) = 0$ hence $s_3 = 5 + \alpha_1$.

During the fourth bang, the controls satisfy $u_1 = 1$ and $u_2 = 1$ and

\[
\begin{align*}
x(s) &= -7r - \alpha_1 r + sr, \\
p_x(s) &= r, \\
\phi_1(s) &= \frac{-5 - \alpha_1 + s}{2}, \\
p_y(s) &= p_y(0) = r, \\
\phi_2(s) &= \frac{7 + \alpha_1 - s}{2}, \\
z(s) &= 4r^2.
\end{align*}
\]

The fourth switching time $s_4$ corresponds to $\phi_2(s_4) = 0$ hence $s_4 = 7 + \alpha_1$.

Along the fifth bang, the controls satisfy $u_1 = 1$ and $u_2 = -1$ and

\[
\begin{align*}
x(s) &= 0, \\
p_x(s) &= 8r + \alpha_1 r - sr, \\
\phi_1(s) &= \frac{9 + \alpha_1 - s}{2}, \\
p_y(s) &= p_y(0) = r, \\
\phi_2(s) &= \frac{7 + \alpha_1 + s}{2}, \\
z(s) &= 4r^2.
\end{align*}
\]

The fifth switching time $s_5$ corresponds to $\phi_1(s_5) = 0$ hence $s_5 = 9 + \alpha_1$.

The other extremals with $\lambda_z \neq 0$ can be computed the same way and are very similar. Finally, extremals with $\lambda_z > 0$ have projections in the $(x, y)$-plane which
are squares and the z-coordinate after one turn of the square is equal to the area of the square. This implies that they are all optimal until the end of this turn that is until \( s = 8 \) or \( t = \frac{8}{p_z} \). After they lose optimality, crossing one each other transversally. As a consequence the cut time is \( s = 8 \) or \( t = 8r \) and the cut locus is the vertical axis (as in the Heisenberg case in sub-riemannian geometry).

4.2. Extremal with \( \lambda_z = 0 \). What about the extremals with \( \lambda_z = 0 \)? For such an extremal, \( \lambda \) is constant and \( \phi_1 = \frac{\lambda_x + \lambda_y}{2} \) and \( \phi_2 = \frac{\lambda_x - \lambda_y}{2} \) are also constant. If both are not zero hence \( u_1 \) and \( u_2 \) are constants along the extremal, the corresponding curve is optimal and is an extremal. If \( \phi_1 \equiv 0 \) and \( \phi_2 \equiv 1 \) then the extremal is \( u_1 \)-singular and the control \( u_1 \) is not determined by the max condition of the PMP. In fact in this case, one proves easily that for any choice of \( u_1(\cdot) \) such that \( |u_1(t)| \leq 1 \), one gets for any \( T > 0 \), a minimizer from \((0,0,0)\) to \((\frac{\int_0^T u_1(t) dt + T}{2}, \frac{\int_0^T u_1(t) dt - T}{2}, z)\) where
$z = \int_0^T \left( u_1(t) - 1 \right) \left( \int_0^t u_1(\tau) d\tau + t \right) \frac{dt}{4}.$

The proof comes from the fact that the projection of this point on the $(x, y)$-plane is on the segment between the two points $(T, 0)$ and $(0, -T)$. The same kind of computation can be done for $\phi_1 \equiv 0$ and $\phi_2 \equiv -1$ or $\phi_1 \equiv 1$ and $\phi_2 \equiv 0$.

4.3. **Exponential map.** Let us concentrate again on the extremals with $\lambda_z \neq 0$. One can consider the exponential map which to $(r, \alpha, s)$ where $\alpha \in [-1, 1[, r > 0$, $s \geq 0$ associates the end point of the extremal with initial condition $\lambda_x = \alpha$, $\lambda_y = 1$ and $\lambda_z = \frac{1}{r}$ for the time $t = rs$. This map is smooth at points with $-1 < \alpha < 1$, $s_i(p_x, r) < s < s_{i+1}(p_x, r)$ for a certain $i$ where $s_j(p_x, r)$ is the $j^{th}$ switching time of the extremal with initial condition $p_x$, $p_y = 1$ and $r$. The same can be done for $\lambda_y = -1$ or $\lambda_z = \pm 1$ and $\lambda_y \in [-1, 1]$. Since it is smooth for $-r < p_x < r$ and $s \neq s_i$, we can compute its Jacobian. It happens that it is null during the two first bangs, and that it has opposite sign to the one of $r$ during the third and fourth bangs. It is again null during the fifth bang. As we will see later for $r$ small in the generic cases, the Jacobian will not be null during the third and fourth bangs also. In the nilpotent case, the first conjugate time is $t_5 = rs_5$ and for $t \in ]rs_4, rs_5[$, $Jac(t) = 0$.

4.4. **Geometric objects.** Since the conjugate time is $t_5$, the first conjugate locus is the set of points where an extremal switches for the fifth time. The first conjugate locus is

$\{(2\delta r, 0, \pm 4r^2)|r \in \mathbb{R}, \delta \in ]-1, 1[\} \cup \{(0, 2\delta r, \pm 4r^2)|r \in \mathbb{R}, \delta \in ]-1, 1[\}.$

The Maxwell set is exactly the same set.

Figure 2 shows the conjugate locus and three points of view of the part of the sphere that is reached by non singular extremals.

5. **Extremals with both controls switching.** In this section, we present the computation of jets of extremals with large covector $|\lambda|$ and of geometric objects attached to them: switching locus and conjugate locus. As in the nilpotent case, we can define a Hamiltonian flow which, to an initial condition $(\lambda_x, \lambda_y, \lambda_z)$ (with $\max(|\lambda_x|, |\lambda_y|) = 1$) associates the end point at time $t$ of the solution of the dynamics

$\dot{x} = \frac{u_1 + u_2}{2} + \frac{u_1 - u_2}{2} (a_{200}x^2 + a_{110}xy + \theta_x),$

$\dot{y} = \frac{u_1 - u_2}{2} (1 + b_{200}x^2 + b_{110}xy + \theta_y),$

$\dot{z} = \frac{u_1 - u_2}{2} (x + c_{200}x^2 + c_{110}xy + c_{300}x^3 + c_{210}x^2y + c_{120}y^2x + \theta_z),$

$\dot{\lambda}_x = -\frac{u_1 - u_2}{2} (\lambda_x(2a_{200}x + a_{110}y) + \lambda_y(2b_{200}x + b_{110}y) + \lambda_z(1 + 2c_{200}x + 3c_{300}x^2 + c_{110}y + 2c_{210}xy + c_{120}y^2)),$

$\dot{\lambda}_y = -\frac{u_1 - u_2}{2} (a_{110}x\lambda_x + b_{110}x\lambda_y + \lambda_z(c_{110}x + c_{210}x^2 + 2c_{120}xy)),$

$\dot{\lambda}_z = \frac{u_1 - u_2}{2} \lambda_z(c_{201}x + c_{111}y),$

$u_1(t) = sign(\phi_1(t)), \quad u_2(t) = sign(\phi_2(t)), \quad \phi_1(t) = \lambda(t)F_1(q(t)), \quad \phi_2(t) = \lambda(t)F_2(q(t)).$
From now $\dot{x}$ denotes $\frac{dx}{dt}$. Using the change of coordinates for $(x, y, z, p, r, s)$, we can define a new Hamiltonian flow by the dynamics

\[ \dot{x} = \frac{u_1 + u_2}{2} r + \frac{u_1 + u_2}{2} r (a_{200} x^2 + a_{110} xy + \theta_x), \]
\[ \dot{y} = \frac{u_1 - u_2}{2} r (1 + b_{200} x^2 + b_{110} xy + \theta_y), \]
\[ \dot{z} = \frac{1}{2} r (\theta_z (u_1 + u_2) + (u_1 - u_2) (x + c_{200} x^2 + c_{300} x^3 + c_{110} xy + c_{210} x^2 y + c_{120} xy^2)), \]
\[ \dot{p}_x = -\frac{u_1 - u_2}{2} r (1 + 2 c_{200} x + p_x (2 a_{200} x + a_{110} y) + p_y (2 b_{200} x + b_{110} y) + 3 c_{300} x^2 + c_{110} y + 2 c_{210} xy + c_{120} xy^2), \]
\[ \dot{p}_y = -\frac{u_1 - u_2}{2} r (c_{110} x + a_{110} p_x x + b_{110} p_y x + c_{210} x^2 + 2 c_{120} xy), \]
\[ \dot{r} = \frac{u_1 - u_2}{2} r^2 x (c_{201} x + c_{111} y). \]

Since the set of initial condition is a square for $(p_x, p_y)$, we define in fact four Hamiltonian flows for each initial speed $(G_1, -G_1, G_2, -G_2)$. For example, for the extremals with initial speed equal to $G_2$ we have $p_y(0) = r$ and $p_x = \alpha r$ with...
\(\alpha \in [-1, 1]\). The new Hamiltonian flow as for variables \((r_0, \alpha, s)\) where \(r_0 = r(0)\), \(p_x(0) = \alpha r\) and s = \(\frac{1}{\alpha^2}\). In order to compute jets of the Hamiltonian flow we write

\[
\begin{align*}
x(r_0, \alpha, s) &= x_1(\alpha, s)r_0 + x_2(\alpha, s)r_0^2 + x_3(\alpha, s)r_0^3 + X_4(r_0, \alpha, s)r_0^4, \\
y(r_0, \alpha, s) &= y_1(\alpha, s)r_0 + y_2(\alpha, s)r_0^2 + y_3(\alpha, s)r_0^3 + Y_4(r_0, \alpha, s)r_0^4, \\
z(r_0, \alpha, s) &= z_2(\alpha, s)r_0^2 + z_3(\alpha, s)r_0^3 + z_4(\alpha, s)r_0^4 + Z_5(r_0, \alpha, s)r_0^5, \\
p_x(r_0, \alpha, s) &= p_{x_1}(\alpha, s)r_0 + p_{x_2}(\alpha, s)r_0^2 + p_{x_3}(\alpha, s)r_0^3 + P_{x_4}(r_0, \alpha, s)r_0^4, \\
p_y(r_0, \alpha, s) &= p_{y_1}(\alpha, s)r_0 + p_{y_2}(\alpha, s)r_0^2 + p_{y_3}(\alpha, s)r_0^3 + P_{y_4}(r_0, \alpha, s)r_0^4, \\
r(r_0, \alpha, s) &= r_0 + r_2(\alpha, s)r_0^2 + r_3(\alpha, s)r_0^3 + R_4(r_0, \alpha, s)r_0^4.
\end{align*}
\]

where all the new functions are smooth functions of their variables. Using this dynamics we find the following. For the first order

\[
\begin{align*}
\dot{x}_1 &= \frac{u_1 + u_2}{2}, & \dot{p}_{x_1} &= -\frac{u_1 + u_2}{2}, \\
\dot{y}_1 &= \frac{u_1 - u_2}{2}, & \dot{p}_{y_1} &= 0, \\
\dot{z}_2 &= \frac{u_1 - u_2}{2}x_1.
\end{align*}
\]

For the second order

\[
\begin{align*}
\dot{x}_2 &= 0, & \dot{p}_{x_2} &= -\frac{u_1 - u_2}{2}(2c_{200}x_1 + c_{110}y_1), \\
\dot{y}_2 &= 0, & \dot{p}_{y_2} &= -\frac{u_1 - u_2}{2}c_{110}x_1, \\
\dot{z}_3 &= \frac{u_1 - u_2}{2}(x_2 + x_1(c_{200}x_1 + c_{110}y_1)), & \dot{r}_2 &= 0.
\end{align*}
\]

For the third order

\[
\begin{align*}
\dot{x}_3 &= \frac{u_1 - u_2}{2}(a_{200}x_1 + a_{110}x_1y_1), \\
\dot{y}_3 &= \frac{u_1 - u_2}{2}(b_{200}x_1 + b_{110}x_1y_1), \\
\dot{z}_4 &= \frac{u_1 - u_2}{2}(c_{300}x_1^3 + 2c_{200}x_1x_2 + x_3 + c_{110}x_2y_1 + c_{110}x_1y_2 + c_{210}x_1^2y_1 + c_{120}x_1y_1^2), \\
\dot{p}_{x_3} &= -\frac{u_1 - u_2}{2}(2b_{200}p_{x_1}x_1 + 2b_{110}p_{x_1}y_1 + 2c_{200}x_2 + 3c_{300}x_1^2 \\
&\quad+ a_{110}p_{x_1}y_1 + b_{110}p_{y_1}y_1 + c_{110}y_2 + 2c_{210}x_1y_1 + c_{120}y_1^2), \\
\dot{p}_{y_3} &= \frac{u_1 - u_2}{2}(-c_{110}x_2 - x_1(a_{110}p_{x_1} + b_{110}p_{y_1} + c_{210}x_1 + 2c_{120}y_1)), \\
\dot{r}_3 &= 0.
\end{align*}
\]

Recall that the extremals we are interested in have initial condition

\[
\begin{align*}
x(r_0, \alpha, 0) &= 0, & p_x(r_0, \alpha, 0) &= r_0p_{x_1}(\alpha, 0), \\
y(r_0, \alpha, 0) &= 0, & p_y(r_0, \alpha, 0) &= r_0p_{y_1}(\alpha, 0), \\
z(r_0, \alpha, 0) &= 0, & r(r_0, \alpha, 0) &= r_0.
\end{align*}
\]

These equations are integrable hence we can compute the jets of the switching functions and hence the jets of the switching times. Finally, we are able to compute the jets of the different bangs of the extremals. We do not give the expressions of these calculus since they would take too much place. If we restrict the computation to \(x, y, z\) as functions of \((r_0, \alpha, s)\) for the four Hamiltonian flows, we get four exponential maps that we denote \(\text{Exp}_\beta\) where \(\beta = -1, 1, -2\) or 2 depending on if the initial velocity is \(-G_1, G_1, -G_2, G_2\).

It happens that all the extremals computed that way have \((p_x, p_y)\) rotating around \((0, 0)\) like in the 3D contact sub-riemannian geometry. As a consequence, if \(r_0 > 0\) and if the extremal starts with \(G_1\) then after it switches to \(G_2\), then \(-G_1, -G_2, G_1\) and so on.
In [30], M. Sigalotti proves, studying second order optimality conditions, that this family of extremals cannot be optimal after the fifth switch.

For these exponential maps, one can compute their Jacobian for each bang arc. One finds

- \( \text{Jac}(\text{Exp}_{\pm 2}) = 0 \) for \( 0 < s < s_2, s \neq s_1 \),
- \( \text{Jac}(\text{Exp}_{\pm 2}) = -8r_0^3 + o(r_0^3) \) for \( s_2 < s < s_3 \),
- \( \text{Jac}(\text{Exp}_{\pm 2}) = -8r_0^3 + o(r_0^3) \) for \( s_3 < s < s_4 \),
- \( \text{Jac}(\text{Exp}_{\pm 2}) = 32(2c_{1120} - c_{1110}^2)r_0^3 + o(r_0^3) \) for \( s_4 < s < s_5 \),
- \( \text{Jac}(\text{Exp}_{\pm 2}) = 8r_0^3 + o(r_0^3) \) for \( s_5 < s < s_6 \),

and

- \( \text{Jac}(\text{Exp}_{\pm 1}) = 0 \) if \( 0 < s < s_1 \) or \( s_1 < s < s_2 \),
- \( \text{Jac}(\text{Exp}_{\pm 1}) = 4r_0^3 + o(r_0^3) \) if \( s_2 < s < s_3 \),
- \( \text{Jac}(\text{Exp}_{\pm 1}) = -8r_0^3 + o(r_0^3) \) if \( s_3 < s < s_4 \),
- \( \text{Jac}(\text{Exp}_{\pm 1}) = 64(3c_{300} - 2b_{200} - 2c_{200}^2)r_0^3 + o(r_0^3) \) if \( s_4 < s < s_5 \),
- \( \text{Jac}(\text{Exp}_{\pm 1}) = 8r_0^3 + o(r_0^3) \) if \( s_5 < s < s_6 \).

We can now state the following proposition which shows that the sign of the Jacobian is an important invariant which determines the conjugate time.

**Proposition 12.** Let \( G_1 \) and \( G_2 \) as in the normal form given in section 2.

- If \( C_1 = 3c_{300} - 2b_{200} - 2c_{200}^2 > 0 \) then the fourth switching time \( t_4 \) is the first conjugate time for extremal with initial velocity \( \pm G_1 \). If \( C_1 < 0 \) then it is the fifth \( t_5 \).
- If \( C_2 = 2c_{1120} - c_{1110}^2 > 0 \) then the fourth switching time \( t_4 \) is the first conjugate time for extremals with initial velocity \( \pm G_2 \). If \( C_2 < 0 \) then it is the fifth \( t_5 \).

Still using the expressions given in Appendix, we can give the expressions of the upper part of the first conjugate locus for the four exponential maps.

For \( \text{Exp}_{\pm 1} \), if \( C_1 > 0 \)

\[
\begin{align*}
x_{\text{conj}} &= \pm (\alpha_2 - 1)r_0 + (4c_{1110} - c_{200}(\alpha_2 - 1)^2)r_0^3 + o(r_0^3), \\
y_{\text{conj}} &= -8c_{200}r_0^3 \pm 4(b_{110} + 6c_{110}c_{200} - 2c_{210}) \\
&\quad + (4b_{200} + 12c_{200}^2 - 6c_{300})\alpha_2 r_0^3 + o(r_0^3), \\
z_{\text{conj}} &= 4r_0^2 \mp 8(c_{1110} + 2c_{200}\alpha_2)r_0^3 + o(r_0^3),
\end{align*}
\]

and if \( C_1 < 0 \)

\[
\begin{align*}
x_{\text{conj}} &= \pm (1 + \alpha_2)r_0 + (4c_{1110} - c_{200}(1 + \alpha_2)^2)r_0^3 + o(r_0^3), \\
y_{\text{conj}} &= -8c_{200}r_0^3 \pm 4(b_{110} + 6c_{110}c_{200} - 2c_{210}) \\
&\quad + (4b_{200} + 12c_{200}^2 - 6c_{300})\alpha_2 r_0^3 + o(r_0^3), \\
z_{\text{conj}} &= 4r_0^2 \mp 8(c_{1110} + 2c_{200}\alpha_2)r_0^3 + o(r_0^3),
\end{align*}
\]

and for \( \text{Exp}_{\pm 2} \), if \( C_2 > 0 \)

\[
\begin{align*}
x_{\text{conj}} &= 4c_{1110}r_0^2 \pm 4(b_{110} + 6c_{110}c_{200} - 2c_{210}) \\
&\quad + \alpha_1(2c_{1120} - 3c_{1110}r_0^3 + o(r_0^3)), \\
y_{\text{conj}} &= \pm (-1 + \alpha_1)r_0 - \frac{1}{2}(16c_{200} + c_{110}(\alpha_1 - 1)^2)r_0^3 + o(r_0^3), \\
z_{\text{conj}} &= 4r_0^2 \pm 4(4c_{200} - c_{110}(1 + \alpha_1))r_0^3 + o(r_0^3),
\end{align*}
\]
and if $C_2 < 0$
\[
\begin{align*}
x_{\text{conj}} &= 4c_{110}r_0^2 ± 4(b_{110} + 6c_{110}c_{200} - 2c_{210} + \alpha_1(2c_{120} - 3c_{110}^2))r_0^3 + o(r_0^3), \\
y_{\text{conj}} &= ±(1 + \alpha_1)r_0 - \frac{1}{2}(16c_{200} + c_{110}(1 + \alpha_1)^2)r_0^2 + o(r_0^2), \\
z_{\text{conj}} &= 4r_0^2 ± 4(4c_{200} + c_{110}(1 - \alpha_1))r_0^3 + o(r_0^3).
\end{align*}
\]

6. **Local cut locus of extremals with $\lambda_z(0) >> 1$**. In the nilpotent case, the extremals with $|\alpha| < 1$ reach the Maxwell set at the fourth switch. When $C_1 \neq 0$ and $C_2 \neq 0$ we will see that the cut locus is reached during the fourth or fifth bang.

From section 4, since the optimal synthesis of the nilpotent case is a good estimation of the optimal synthesis of the general case, we can deduce that the cut time will be close to a Maxwell time for the nilpotent approximation that is a time at which, for the same initial condition, the geodesic of the nilpotent case is in the Maxwell set. Hence, since in the nilpotent approximation, the Maxwell times are between $s = 7 + \alpha_1$ (or $s = 7 + \alpha_2$) and $s = 8$.

We can conclude that the loss of optimality may come during the fourth bang (close to the fourth switching time) or during the fifth bang. Moreover, in [30] the author proves that the extremals we are considering cannot be optimal after the fifth switch. Hence we can conclude that the cut locus comes from the intersection of two fourth bangs of different $exp_i$, the intersection of two fifth bangs of different $exp_i$, the intersection of a fourth bang and a fifth bang of two different $exp_i$.

In the following we compute, for the jets of order 3, 4 and 4 of $x$, $y$ and $z$ in $r_0$, the possible intersections listed previously, and finally describe the possible pictures of the cut locus depending on the values of invariants of the structure appearing in the normal form. Finally we discuss the stability of the pictures.

6.1. **Intersections of fourth bangs**.

6.1.1. **Intersection of an extremal starting with $±G_1$ with another starting with $±G_2$**.

As seen in the nilpotent case, an extremal starting with $±G_1$ and $|\alpha_2| < 1$ meets the Maxwell set at $s = s_4$ and intersect at this time the extremal starting with $±G_2$ and $\alpha_1 = 1$. Hence, we compute the jets of $Exp_{±1}$ close to the fourth switch time that is at $s = 7 + \alpha_2 + T_2r_0 + T_3r_0^2$ and the jets of $Exp_{±2}$ for $r_0' = r_0 + R_2r_0^2 + R_3r_0^3$, $\alpha_1 = 1 - \alpha_{11}r_0 - \alpha_{12}r_0^2$ and at time $s' = \frac{r_0'^4}{r_0^4}$. Asking that the corresponding points are the same, one gets
\[
\begin{align*}
R_2 &= ±2c_{200}(1 + \alpha_2) \\
T_2 &= ±8c_{110} - c_{200}(1 + 14\alpha_2 + \alpha_2^2) \\
\alpha_{11} &= 0
\end{align*}
\]
and
\[
R_3 = \frac{(1 + \alpha_2)}{2} (3b_{110} + 6c_{110}c_{200} + 4c_{200}(1 + 3\alpha_2)) + 4b_{200}(1 - \alpha_2) + 6c_{300}(1 - \alpha_2) - 6c_{210})
\]
We see here that in order the intersection exists, \(\alpha_1 = 1 - \alpha_{11} r_0 - \alpha_{12} r_0^2\) should be less or equal to 1 hence, since \(\alpha_{11} = 0\), one should have \(\alpha_{12} > 0\) which implies \(C_1 > 0\).

When \(C_1 > 0\), once computed the corresponding points (depending on \(r_0\) and \(\alpha_2\)) one can compute the suspension of this part of the cut locus by looking at its intersection with \(z = 4\rho^2\) for \(\rho\) small. One gets

\[
\begin{align*}
T_3 &= \frac{16}{3} a_{110} + 20 c_{110}^2 - a_{200} + 8 b_{200} + 38 c_{200}^2 - \frac{40}{3} c_{120} - 27 c_{300} \\
&\quad + (8 b_{110} + 2 a_{200} + 9 b_{200} + 48 c_{110} c_{200} + 14 c_{200}^2 - 15 c_{300} - 16 c_{210}) \alpha_2 \\
&\quad + ((-a_{200} + 14 b_{200} + 4 a_{200}^2 - 21 c_{300}) \alpha_2^2 \\
&\quad + (b_{200} + 2 a_{200}^2 - c_{300}) \alpha_3^2) \\
\alpha_{12} &= 4(1 + \alpha_2)(3 c_{300} - 2 b_{200} - 2 c_{200}^2) = 4(1 + \alpha_2) C_1
\end{align*}
\]

6.1.2. Intersection of an extremal starting with \(\pm G_2\) with another starting with \(\mp G_1\).

The same computations can be done for extremals starting by \(\pm G_2\) and intersecting \(\mp G_1\) and one gets that \(C_2\) should be positive. Hence

\[
\begin{align*}
x_{\text{cut}} &= \pm(-1 + \alpha_2) \rho + (3 c_{110} - c_{200} + c_{110} \alpha_2 + c_{200} \alpha_2) \rho^2 \\
&\quad \pm \frac{1}{2} \left(a_{110} - 7 c_{110}^2 - 2 a_{200} + 2 b_{200} - 8 c_{110} c_{200} + 4 c_{200}^2 + 12 c_{120} - 4 c_{300} \right) \\
&\quad + ((4 a_{200} - a_{110} - 5 b_{110} - c_{110} - 6 c_{110} c_{200} - 4 c_{200}^2 + 4 c_{120} + 10 c_{210}) \alpha_2 \\
&\quad + (6 c_{210} - 3 b_{110} - 2 a_{200} + 2 c_{110} c_{200}) \alpha_2^2 + (4 c_{300} - 2 b_{200}) \alpha_2^3) \rho^3 \\
y_{\text{cut}} &= -8 c_{200} \rho^2 \pm (4 b_{110} + 8 c_{110} c_{200} + (8 b_{200} - 12 c_{300} + 8 c_{200}^2)(\alpha_2 - 1)) \rho^3 \\
z_{\text{cut}} &= 4 \rho^2
\end{align*}
\]

6.1.3. Intersection of the front starting with \(G_1\) with the one starting with \(\mp G_1\).

Such a self-intersection of the front can take place only at \(s = 8 + O(r_0)\) as in the nilpotent case. In order to compute such intersection close to \(s = 8\), we proceed as follows. We compute the intersection of these parts of the front with \(z = 4\rho^2\). In order to do this, we fix \(t = 8 \rho + T_2 \rho^2 + T_3 \rho^3\), for each type of extremal fix \(\alpha_2 = 1 - \alpha_{21} \rho - \alpha_{22} \rho^2\) and find the \(r_0\) such that the corresponding point \(\text{Exp}_{\pm 1}(r_0, \alpha, t/r_0)\)
satisfies \( z = 4\rho^2 \). For the extremals starting by \( \pm G_1 \) one finds

\[
\begin{align*}
x_{sus} &= (4c_{110} + \alpha_{11} 2 + 2\rho^2 \pm (4b_{110} + 4c_{110}^2 + 4c_{200}^2 3)
+ 2c_{110}(4c_{200} 131) + \alpha_{22} - 8c_{120} - 8c_{210})\rho^3
\end{align*}
\]

\[
\begin{align*}
y_{sus} &= (-8c_{200} 12 + \alpha_{21} 2 110 12 + 8b_{200} 12 21c_{110}
+ 8c_{120} + 4\alpha_{21}c_{200} + 16c_{110}c_{200} + 16c_{200} c_{200} - 16c_{300} - \alpha_{21}T_2 4c_{110}T_2 - T_3)\rho^3
\end{align*}
\]

\[
\begin{align*}
z_{sus} &= 4\rho^2
\end{align*}
\]

It is then easy to show that, in order to get a contact between these two fronts, \( T_2 \) should be equal to 0 and \( \alpha_{21} = -\alpha_{21} \). But, since both should be positive hence \( \alpha_{21} = \alpha_{21} = 0 \) and this implies that \( T_3 \) should be equal to

\[
T_{3b} = \frac{4}{3}(a_{110} 3b_{110} 6b_{200} + 3c_{110}^2 12 - 4c_{120} + 18c_{110}c_{200} + 12c_{200}^2 6c_{200} - 6c_{210} - 12c_{300}).
\]

At this time, with \( \alpha_{21} = \alpha_{21} = 0 \), the two fronts are segments belonging to the same line.

6.1.4. Intersection of the front starting with \( G_2 \) with the one starting with \( -G_2 \). We proceed the same way. For the extremals starting by \( \pm G_2 \) one finds

\[
\begin{align*}
x_{sus} &= (4c_{110} 16 + \alpha_{11} 2 + 2\rho^2 \pm (4b_{110} - 4c_{110}^2 + 4b_{200}^2 4)
- 2c_{110}(8c_{200} 131) + \alpha_{12} 16 c_{120} - 8c_{300} + T_3)\rho^3
\end{align*}
\]

\[
\begin{align*}
y_{sus} &= -(8c_{200} 12 + \alpha_{11} 2 \pm (4b_{110} - 16b_{200} + 8c_{110}c_{200} - 16c_{200}^2
+ 2c_{110}c_{11} 4c_{200}c_{111} - \alpha_{12} 24c_{300} - 8c_{210})\rho^3
\end{align*}
\]

\[
\begin{align*}
z_{sus} &= 4\rho^2
\end{align*}
\]

It is then easy to show that, in order to get a contact between these two fronts, \( T_2 \) should be equal to 0 and \( \alpha_{11} = -\alpha_{11} \). But, since both should be positive hence \( \alpha_{11} = \alpha_{11} = 0 \) and this implies that \( T_3 \) should be equal to

\[
T_{3a} = \frac{4}{3}(a_{110} 3b_{110} 6b_{200} + 3c_{110}^2 12 - 4c_{120} + 6c_{110}c_{200} + 12c_{200}^2 6c_{210} - 12c_{300}).
\]

6.2. Cut locus when \( C_1 > 0 \) and \( C_2 > 0 \). With the considerations given before, if \( C_1 > 0, C_2 > 0 \) and \( T_{3a} 13b \), the intersection of the cut locus with \( \{ z = 4\rho^2 \} \) is constituted of 5 branches as in the Figure 3.

The four external branches comes from the intersection of the fourth bangs of \( \exp_{\pm 1} \) with \( \exp_{\pm 2} \) and of the fourth bangs of \( \exp_{\pm 1} \) with \( \exp_{\pm 2} \), see Figure 3. The central branch is the intersection of the fourth bangs of \( \exp_{\pm 1} \) with \( \exp_{\pm 1} \) if \( T_{3b} < T_{3a} \) or of the fourth bangs of \( \exp_{\pm 2} \) with \( \exp_{\pm 2} \) if \( T_{3a} < T_{3b} \), see Figure 4.

After \( \min\{T_{3a}, T_{3b}\} \) all the extremals participating to the construction of this part of the cut locus have lost optimality.

Finally the picture of the cut depends on the sign of

\[
T_{3a} - T_{3b} = -8(b_{110} + 2c_{110}c_{200} - 2c_{210}).
\]
Figure 3. $C_1 > 0$ and $C_2 > 0$: closure of the cut locus at $z$ fixed.

Figure 4. $C_1 > 0$ and $C_2 > 0$: closure of the cut locus at $z$ fixed.

If $T_{3a} > T_{3b}$ then the two points of the cut locus that connect three branches are with

\[
\begin{align*}
    x &= 4c_{110}\rho^2 \pm C\rho^3 + o(\rho^3) \\
    y &= -8c_{200}\rho^2 \pm C\rho^3 + o(\rho^3) \\
    z &= 4\rho^2
\end{align*}
\]

with $C = 4(b_{110} + 2c_{110}c_{200} - 2c_{210})$, when if $T_{3a} < T_{3b}$ then the two points of the cut locus that connect three branches satisfy

\[
\begin{align*}
    x &= 4c_{110}\rho^2 \pm C\rho^3 + o(\rho^3) \\
    y &= -8c_{200}\rho^2 \pm C\rho^3 + o(\rho^3) \\
    z &= 4\rho^2
\end{align*}
\]
Finally we can present the upper part of the cut locus when $C_1 > 0$ and $C_2 > 0$ in Figure 5.

6.3. Suspension of fifth bang front. At $6 < s < 8$, the part of the front corresponding to the fifth bang is close to $(\pm (s - 8)\rho, 0, 4\rho^2)$ for the front starting with $\pm G_1$ and close to $(0, \pm (s - 8)\rho, 4\rho^2)$ for the front starting with $\pm G_2$. Hence the intersections come at $s$ close to 8.

In order to compute these intersections we fix a small $\rho$, consider a time $t = 8\rho + T_2\rho^2 + T_3\rho^3$, and for each type of extremal find the $r_0$ such that the corresponding point $\text{Exp}\, \pm 1(r_0, \alpha, t/r_0)$ satisfies $z = 4\rho^2$. For the extremals starting by $\pm G_1$ one finds $\text{exp}_{\pm 1}$

$$x_{\pm 1\text{us}} = (4c_{110} \pm T_2)\rho^2 \pm \left(\frac{16}{3}c_{120} - \frac{4}{3}a_{110} - 4c_{110}^2 - 16c_{110}c_{200} - 8c_{200}^2\right) + 4c_{300} + T_3 - 4C_1\alpha_2)\rho^3$$

$$y_{\pm 1\text{us}} = -8c_{200}\rho^2 \pm (4b_{110} + 8c_{110}c_{200} - 8c_{210} - 8C_1\alpha_2)\rho^3$$

$$z_{\pm 1\text{us}} = 4\rho^2$$

For the extremals starting by $\pm G_2$ one finds $\text{exp}_{\pm 2}$

$$x_{\pm 2\text{us}} = 4c_{110}\rho^2 \pm (4b_{110} + 8c_{110}c_{200} - 8c_{210} + 4C_2\alpha_1)\rho^3$$

$$y_{\pm 2\text{us}} = (-8c_{200} \pm T_2)\rho^2 \pm \left(\frac{4}{3}c_{120} - \frac{4}{3}a_{110} - 8b_{200} - 16c_{200}^2 - 2c_{110}^2\right) + 16c_{300} + 4c_{110}(-4c_{200} \pm T_2) + T_3 - 2C_2\alpha_1^2)\rho^3$$

$$z_{\pm 2\text{us}} = 4\rho^2$$

As one can see, the intersection of the fifth bang front at $t$ with the plane $z = 4\rho^2$ is the union of arc of parabolas. If we consider all these curves for $\alpha_i \in [-1, 1]$ we can observe that the tangents at $\alpha_i = \pm 1$ are lines with equations of the type $x + y = c$ or $x - y = c$. Moreover, this tangent at $\alpha_2 = -1$ of the fifth bang front of $\text{exp}_{\pm 1}$ is tangent to the fourth bang at the corresponding $\alpha_1$ of $\text{exp}_{\pm 2}$, and the tangent at $\alpha_1 = -1$ of the fifth bang front of $\text{exp}_{\pm 2}$ is tangent to the fourth bang at the corresponding $\alpha_2$ of $\text{exp}_{\pm 1}$.

Moreover remark that, at $T_2 = 0$, the intersection of the front with $z = 4\rho^2$ still has a central symmetry at this order of jets, centred at $(x, y) = (4c_{110}\rho^2, -8c_{200}\rho^2)$.
6.4. Cut locus when $C_1 > 0$ and $C_2 < 0$. If $C_1 > 0$ and $C_2 < 0$ then the picture of the front at $t < 8\rho$ is as in the Figure 6. The fifth bang of $\exp_{\pm 1}$ do not participate to the optimal synthesis and the fourth bang front of $\exp_{\pm 1}$ intersect the fourth bang front of $\exp_{\pm 2}$. The fifth bang front of $\exp_{\pm 2}$ is optimal.

![Figure 6. The front before $t = 8\rho$ when $C_1 > 0$ and $C_2 < 0$](image)

Let consider the closure of the cut, that it when $t = 8\rho + T_2\rho^2 + T_3\rho^3$. We can identify the following subcases

- When $4b_{110} + 8c_{110}c_{200} - 8c_{210} + 4C_2 < 0$ then all the fifth bang of $\exp_2$ satisfies $x < 4c_{110}\rho^2$ when all the fifth bang of $\exp_{-2}$ satisfies $x > 4c_{110}\rho^2$. This implies that the sequel of the self intersections of the front is the following: first the fourth bang front of $\exp_{\pm 1}$ intersect the fourth bang front of $\exp_{\pm 2}$; then at time $T_2 = 0$, $T_3 = T_{3c} = T_{3b} + \frac{4}{3}C_2 - \frac{8}{3}c_{110} < T_{3b}$ the fourth bang of $\exp_{\pm 1}$ intersects the fifth bang of $\exp_{\pm 2}$; finally the fourth bang of $\exp_1$ intersects the fourth bang of $\exp_{-1}$ at $T_2 = 0$ and $T_3 = T_{3b}$. See Figure 7.

![Figure 7. $C_1 > 0$ and $C_2 < 0$: picture of the front at times with $T_2 = 0$ and $T_3 < T_{3c}$, $T_3 = T_{3c}$ and $T_3 = T_{3b}$ when $4b_{110} + 8c_{110}c_{200} - 8c_{210} + 4C_2 < 0$](image)

- When $4b_{110} + 8c_{110}c_{200} - 8c_{210} < 0$ and $4b_{110} + 8c_{110}c_{200} - 8c_{210} + 4C_2 > 0$ then the relative position of the two parabola of the fifth bang of $\exp_2$ and $\exp_{-2}$ implies that the sequel of the self intersections of the front is the following: first the fourth bang front of $\exp_{\pm 1}$ intersect the fourth bang front of $\exp_{\pm 2}$; then at time $T_2 = 0$, $T_3 = T_{3c} = T_{3b} + \frac{4}{3}C_2 - \frac{8}{3}c_{110} < T_{3b}$ the fourth bang of $\exp_{\pm 1}$ intersects the fifth bang of $\exp_{\pm 2}$; finally the fifth bang of $\exp_2$ intersects the fifth bang of $\exp_{-2}$ at a time with $T_2 = 0$ and $T_3 = T_{3g}$ between $T_{3c}$ and $T_{3b}$. See Figure 8.

- When $4b_{110} + 8c_{110}c_{200} - 8c_{210} > 0$ and $4b_{110} + 8c_{110}c_{200} - 8c_{210} + 4C_2 < 0$ then the relative position of the two parabola of the fifth bang of $\exp_2$ and $\exp_{-2}$ implies that the sequel of the self intersections of the front is the following: first the fourth bang front of $\exp_{\pm 1}$ intersect the fourth bang front of $\exp_{\pm 2}$;
Figure 8. $C_1 > 0$ and $C_2 < 0$: picture of the front at times with $T_2 = 0$ and $T_3 < T_{3c}$, $T_3 = T_{3c}$ and $T_3 = T_{3g}$ when $4b_{110} + 8c_{110}c_{200} - 8c_{210} < 0$ and $4b_{110} + 8c_{110}c_{200} - 8c_{210} + 4C_2 > 0$

then at time $T_2 = 0$ and $T_3 = T_{3d} = T_{3a} + 2C_2 - 2c_{110}^2 < T_{3a}$ the fourth bang of $\exp_{\pm 2}$ intersects the fifth bang of $\exp_{\mp 2}$; finally the fifth bang of $\exp_2$ intersects the fifth bang of $\exp_{-2}$. The picture is similar to the one of Figure 7.

- When $4b_{110} + 8c_{110}c_{200} - 8c_{210} + 4C_2 > 0$ then all the fifth bang of $\exp_2$ satisfies $x > 4c_{110}\rho^2$ when all the fifth bang of $\exp_{-2}$ satisfies $x < 4c_{110}\rho^2$. This implies that the sequel of the self intersections of the front is the following: first the fourth bang front of $\exp_{\pm 1}$ intersect the fourth bang front of $\exp_{\pm 2}$; then at time $T_2 = 0$ and $T_3 = T_{3d} = T_{3a} + 2C_2 - 2c_{110}^2 < T_{3a}$ the fourth bang of $\exp_{\pm 2}$ intersects the fifth bang of $\exp_{\pm 2}$; finally the fourth bang of $\exp_2$ intersects the fourth bang of $\exp_{-2}$ at $T_2 = 0$ and $T_3 = T_{3a}$. The picture is similar to the one of Figure 8.

In the four cases, the cut locus has only one branch, which is continuous and piecewise smooth. And the proportions are those given in the Figure 9.

Figure 9. Picture of the cut locus when $C_1 > 0$ and $C_2 < 0$

6.5. **Cut locus when $C_1 < 0$ and $C_2 > 0$.** The same kind of computations can be done in this case as in the previous case. For the picture of the cut locus we refer to the same figure 9 where the $x$-axis should be replaced by the $y$-axis.

6.6. **Intersections of fifth bangs.** In the case $C_1 < 0$ and $C_2 < 0$, the fifth bang fronts intersect before losing optimality. As before this happens for $t \sim 8\rho$ and we write $t = 8\rho + T_2\rho^2 + T_3\rho^3$.

As seen before, each fifth bang front is a part of parabola. For $T_2 < 0$, or $T_2 = 0$ and $T_3$ small enough, the four parabolas are not intersecting, are positioned as in the figure 10 and they are linked by the part of the front constituted of fourth bangs, and the front do not self intersect.

One way to build the optimal part of the front is to consider the expressions of the fifth bangs and of the four bangs, to consider them for all the values of $\alpha_i \in [-1, 1]$
and to keep only the part which constitutes the boundary of the “central” domain (see Figure 10). The dynamics with respect to $T_3$ of each of these expressions consist only on translations of $\pm T_3$ along $x$ or $y$. Hence to identify the optimal part of these expressions, we just have to understand what are the consecutive intersections when $T_3$ varies.

Figure 10. The front before $t = 8\rho$ when $C_1 < 0$ and $C_2 < 0$

- The first intersection is of the fifth bang front of $\exp_{\pm 1}$ with the one of $\exp_{\pm 2}$ at $T_2 = 0$ and $T_3 = T_{3e}$ or with the one of $\exp_{\mp 2}$ at $T_2 = 0$ and $T_3 = T_{3f}$.
  
  When writing the intersection of the fifth fronts, that is for example that $x_{1sus}|_{\alpha_2 = -1} = x_{2sus}|_{\alpha_1 = 1}$ and $y_{1sus}|_{\alpha_2 = -1} = y_{2sus}|_{\alpha_1 = 1}$, one finds
  
  $$T_{3e} = \frac{4}{3}(a_{110} + 3b_{110} - 6b_{200} + 18c_{110}c_{200} + 2c_{120} + 6c_{300} - 6c_{210})$$
  
  and
  
  $$T_{3f} = \frac{4}{3}(a_{110} - 3b_{110} - 6b_{200} + 6c_{110}c_{200} + 2c_{120} + 6c_{300} + 6c_{210}).$$
  
  After that time, the fifth bang fronts that connected self intersect, until a next event.

Case 1 The next event can be that all the front corresponding to the fifth bang of $\exp_{\pm 1}$ (resp. $\exp_{\pm 2}$) is no more optimal. This comes from the fact that
  
  - the entire arc of parabolas of the fifth bang front of $\exp_{\pm 1}$ crossed the parabolas of $\exp_{\pm 2}$ which occurs if $2|C_1| < |C_2|$, 
  - the entire arc of parabolas of the fifth bang front of $\exp_{\pm 2}$ crossed the parabolas of $\exp_{\pm 1}$ which occurs if $2|C_1| > |C_2|$, 
  
  see figure 11. The corresponding time can be computed in the following way. Assume that $T_{3e} < T_{3f}$ and hence that the first event was the contact of the fifth bang front of $\exp_{\pm 1}$ with the one of $\exp_{\pm 2}$ at one of their extremity. Then, the second event will happen at $T_3$ such that one of the other extremities, let call it $q(T_3)$ crosses the other parabola at $p(T_3)$, see Figure 11. Thanks to the dynamics with respect to $T_3$, $p(T_3)$ and $q(T_3)$ belongs for all $T_3$ at the line $x + y = c + T_3$ where $c \in \mathbb{R}$. Together with the expressions of the parabolas one find that the corresponding time is $T_3 = T_{3e} + \tau_3$ with
  
  $$\tau_3 = 8\sqrt{2C_1C_2}.$$ 
  
  If $T_{3f} < T_{3e}$ then it happens at $T_3 = T_{3f} + \tau_3$. 

Case 2 An other event, that can occurs after the first intersection, is the other contact between fifth bang fronts occurs. If \( T_{3e} < T_{3f} \) then this event is at \( T_3 = T_{3f} \) and if \( T_{3f} < T_{3e} \) it is at \( T_3 = T_{3e} \). See Figure 12

Case 1.1 In the case 1, the next event can be the closure of the synthesis by the contact of the four bangs. If \( T_{3e} < T_{3f} \) the fourth bang fronts of \( \text{exp}_1 \) and \( \text{exp}_{-1} \) can intersect at time \( T_{3b} \). If \( T_{3f} < T_{3e} \) the fourth bang fronts of \( \text{exp}_2 \) and \( \text{exp}_{-2} \) can intersect at time \( T_{3e} \). This case occurs only if the arc of parabolas of \( \text{exp}_{\pm 1} \) from one part, and the arc of parabolas of \( \text{exp}_{\pm 2} \) from the other part, do not intersect at any time \( T_3 \).

Case 1.2 In the case 1, another possibility is that the four bang front loses entirely its optimality. If \( 2|C_1| < |C_2| \) it correspond to the time at which an extremity of the fifth bang front of \( \text{exp}_2 \) touches the fifth bang front of \( \text{exp}_{-2} \). If \( 2|C_1| > |C_2| \) it correspond to the time at which an extremity of the fifth bang front of \( \text{exp}_1 \) touches the fifth bang front of \( \text{exp}_{-1} \). These times can be computed by translating in the calculus these intersection and we get in the different cases

- If \( T_{3e} < T_{3f} \) and \( |C_2| < 2|C_1| \) then \( T_3 = T_{3b} = -K_1 + 2C_1(1 + \alpha_1^2) \) with \( \alpha_i = -1 + \frac{1}{C_1} \left(b_{110} + 2c_{110c200} - 2c_{210}\right) \) and

  \[
  K_1 = \frac{16}{3}c_{120} - \frac{4}{3}a_{110} - 4c_{110}^2 - 16c_{110c200} - 8c_{200}^2 + 4c_{300}.
  \]

- If \( T_{3e} < T_{3f} \) and \( |C_2| > 2|C_1| \) then \( T_3 = T_{3b} = -K_2 + 2C_2(1 + \alpha_2^2) \) with \( \alpha_i = 1 - \frac{2}{C_2} \left(b_{110} + 2c_{110c200} - 2c_{210}\right) \) and

  \[
  K_2 = \frac{4}{3}c_{120} - \frac{4}{3}a_{110} - 2c_{110}^2 - 16c_{110c200} - 16c_{200}^2 + 16c_{300} - 8b_{200}.
  \]

- If \( T_{3e} > T_{3f} \) and \( |C_2| < 2|C_1| \) then \( T_3 = T_{3i} = -K_1 + 2C_1(1 + \alpha_1^2) \) with \( \alpha_i = 1 + \frac{1}{C_1} \left(b_{110} + 2c_{110c200} - 2c_{210}\right) \).

- If \( T_{3e} > T_{3f} \) and \( |C_2| > 2|C_1| \) then \( T_3 = T_{3i} = -K_2 + 2C_2(1 + \alpha_2^2) \) with \( \alpha_i = -1 - \frac{2}{C_2} \left(b_{110} + 2c_{110c200} - 2c_{210}\right) \).

After the fourth bang front lost optimality the optimal synthesis finishes by the last self intersection of the fifth bang front.

Case 2.2 In case two, after \( \max\{T_{3e}, T_{3f}\} \), the optimal synthesis closes as follows.

If \( |C_2| < 2|C_1| \), then the next event is the loss of optimality of the entire fifth bang front of \( \text{exp}_{\pm 2} \), and the optimal synthesis finishes by the intersection of the parabolas of \( \text{exp}_{\pm 1} \). If \( |C_2| > 2|C_1| \), then the next event is the loss of optimality of the entire fifth bang front of \( \text{exp}_{\pm 1} \), and the optimal synthesis finishes by the intersection of the parabolas of \( \text{exp}_{\pm 2} \).

6.7. **Cut locus when** \( C_1 < 0 \) and \( C_2 < 0 \). Thanks to the description of the different steps that can occur along the dynamics of the front, we can conclude by claming

- If \( |T_{3e} - T_{3f}| < \tau_3 \) then the cut locus has 5 smooth branches as in Figure 12.
- If not it has only one branch which is continuous and smooth by arcs, see Figure 11.

Finally we can give the picture of the cut locus in this two cases in Figure 13.

6.8. **Singularities and stability, open question.** For extremals that do not satisfy the initial condition \( |\lambda_x| = |\lambda_y| = 1 \), the computations of the cut locus and the conjugate locus are stable in the following sense: the cut locus corresponds to a transversal intersection of two fronts computed on jets which stay transversal when
we consider the true front; the conjugate locus corresponds to the computation of a certain switching locus which is transversal to the front before and after the switch when computed on the jets and which stays transversal when we consider the true front. Hence we computed a true estimation of the conjugate and cut times and hence of the conjugate and cut loci. The proof of these facts (the permanence of transversality) is simply that the consideration of the true front do not change the calculus at the order where appears the transversality hence, up to taking smaller $r$, the transversality is still true.

For the initial conditions $|\lambda_x| = |\lambda_y| = 1$, things are more complicated. A further study should be done in order to find a good notion of stability, which is itself not clear, and to study it in this case. In the case $C_1 > 0$ and $C_2 > 0$, the corresponding singularity of the front in the sub-Riemannian contact case, corresponding to the
extremity of the cut locus, is a cusp \( A_3 \) (in the classification of Arnol’d) and it is stable as smooth or lagrangian singularity. We may propose the conjecture that a good theory of stability should find in our context that the singularity is stable. If this conjecture is valid then the pictures of the cut locus and the conjugate locus are stable and valid not only for the jets of the dynamics we have computed but also for the true dynamics.

7. **Extremals with only one control switching several times.** For \(|\lambda_z|\) large enough the dynamics is described in the previous sections. We can now choose a constant \( \Lambda_z > 0 \) large enough and considering only the extremal satisfying \(|\lambda_z| < \Lambda_z\). As seen before, along an extremal

\[ \dot{\phi}_3 = u_1(f_{41}\phi_1 + f_{42}\phi_2 + f_{43}\phi_3) + u_2(f_{51}\phi_1 + f_{52}\phi_2 + f_{53}\phi_3). \]

One computes easily that

\[
\begin{align*}
    f_{41}(0) &= -\frac{a_{110} + 2a_{200} + b_{110} + 2b_{200}}{4}, \\
    f_{42}(0) &= -\frac{a_{110} + 2a_{200} - b_{110} - 2b_{200}}{4}, \\
    f_{43}(0) &= \frac{1}{2} c_{110} + c_{200}, \\
    f_{51}(0) &= \frac{a_{110} - 2a_{200} + b_{110} - 2b_{200}}{4}, \\
    f_{52}(0) &= \frac{a_{110} - 2a_{200} - b_{110} + 2b_{200}}{4}, \\
    f_{53}(0) &= -\frac{1}{2} c_{110} + c_{200}.
\end{align*}
\]

With \(|\phi_1| \leq 1\) and \(|\phi_2| \leq 1\), we get

\[ |\dot{\phi}_3| \leq |f_{41}| + |f_{42}| + |f_{51}| + |f_{52}| + (|f_{53}| + |f_{43}|)|\phi_3| \leq 4M' + 2M'\Lambda_z \]

where \(M'\) is a local bound of the \(f_{ij}\). This implies that, for the extremals we are considering, the possibility of switching in short time implies that the corresponding switching function starts close to 0. Which implies that in short time only one control switches. And if in short time a control switches twice hence \(\phi_3\) should change sign and hence starts close to 0 that is \(\lambda_z\) should starts close to 0.

In the following, we will be interested only in finding extremals with only one control switching in short time since the ones with both controls switching in short time are yet obtained with initial conditions with large \(|\lambda_z|\).

We will consider only extremals with \(u_1 \equiv 1\), the study of the other ones being equivalent. Along such an extremal

\[ \ddot{\phi}_2 = u_1 \dot{\phi}_3 = \dot{\phi}_3 \]

and since \(u_1 \equiv 1\) one gets

\[ \ddot{\phi}_2 = (f_{41} + u_2 f_{51})\phi_1 + (f_{42} + u_2 f_{52})\phi_2 + (f_{43} + u_2 f_{53})\phi_3. \]

Since \(\phi_3(t) = O(t)\), \(\phi_2 = O(t)\) and \(\phi_1(t) = 1 + O(t)\) we get that

\[ \ddot{\phi}_2(t) = (f_{41} + u_2 f_{51}) + O(t). \]

In the following we assume that we are considering a point where \(f_{41} + f_{51} \neq 0\) and \(f_{41} - f_{51} \neq 0\). We consider then the four following cases

1. If \(|f_{51}| < f_{41}\) then \(f_{41} + u_2 f_{51} > 0\) for all \(u_2 \in [-1,1]\) and \(\ddot{\phi}_2(t) > 0\) for all \(t\). As a consequence the only possible behaviours of the control \(u_2\) are (see Figure 14)
   (a) \(u_2 \equiv 1\),
   (b) \(u_2 = -1\) during a first interval of time and switches to 1,
   (c) \(u_2 = 1\) during a first interval of time, then \(-1\) during a second one, and finally switches to 1.
Figure 14. Extremals when $|f_{51}| < f_{41}$

2. If $|f_{51}| < -f_{41}$ then $f_{41} + u_2 f_{51} < 0$ for all $u_2 \in [-1,1]$ and $\ddot{\phi}_2(t) < 0$ for all $t$. As a consequence the only possible behaviours of the control $u_2$ are (see Figure 15)

(a) $u_2 \equiv -1$,
(b) $u_2 = 1$ during a first interval of time and switches to $-1$,
(c) $u_2 = -1$ during a first interval of time, then $1$ during a second one, and finally switches to $-1$.

Figure 15. Extremals when $|f_{51}| < -f_{41}$

3. If $|f_{41}| < f_{51}$ then $f_{41} + f_{51} > 0$ hence $\ddot{\phi}_2(t) > 0$ when $\phi_2(t) > 0$ (since then $u_2(t) = 1$) and $f_{41} - f_{51} < 0$ hence $\ddot{\phi}_2(t) < 0$ when $\phi_2(t) < 0$ (since then $u_2(t) = -1$). In that case the possible behaviours of the control $u_2$ are (see Figure 16)

(a) $u_2$ is constant and equal to $\pm 1$,
(b) $u_2$ is equal to $1$ or $-1$ during a first interval of time and switches to $\mp 1$,
(c) $u_2$ is equal to $1$ or $-1$ during a first interval of time, then $\phi_2 = 0$ during a second interval where $u_2(t) = \mp f_{41}(q(t)) f_{51}(q(t)) + O(t)$, and finally $u_2$ switches to $1$ or $-1$, independently of the value during the first interval of time.

4. If $|f_{41}| < -f_{51}$ then $f_{41} + f_{51} < 0$ hence $\ddot{\phi}_2(t) < 0$ when $\phi_2(t) > 0$ (since then $u_2(t) = 1$) and $f_{41} - f_{51} > 0$ hence $\ddot{\phi}_2(t) > 0$ when $\phi_2(t) < 0$ (since then $u_2(t) = -1$). In that case the list of possible behaviours may be very large. In the following we analyse more deeply to prove that the possible behaviours are

(a) $u_2$ is constant and equal to $\pm 1$,
(b) $u_2$ is constant and equal to $\pm 1$ during a first interval of time and switches to $\mp 1$, 

(c) $u_2$ is constant and equal to $\pm 1$ during a first interval of time and switches to $\mp 1$, and finally switches again to $\pm 1$.

A more precise description of the optimal ones is given in the following analysis. In particular, in this case, appears a cut locus.

7.1. **Extremals when** $|f_{41}| < -f_{51}$. In the following we prove that, in the case $|f_{41}| < -f_{51}$, an extremal with $u_1 \equiv 1$ with four bangs is not optimal.

An easy computation shows that

\[
\begin{align*}
    f_{41}(0) &= -\frac{1}{2}(a_{200} + b_{200} + \frac{a_{110} + b_{110}}{2}) \\
    f_{51}(0) &= -\frac{1}{2}(a_{200} + b_{200} - \frac{a_{110} + b_{110}}{2})
\end{align*}
\]

The hypothesis $|f_{41}| < -f_{51}$ is equivalent to $a_{200} + b_{200} > 0$ and $\frac{a_{110} + b_{110}}{2} < 0$.

Consider the three following extremals from $(0,0,0)$ to $(x,y,z)$. The first one, denoted $\epsilon$, has $u_2 = 1$ during time $\epsilon_1$ then $u_2 = -1$ during time $\epsilon_2$ and finally $u_2 = 1$ during time $\epsilon_3$. The second one, denoted $\theta(t)$, has $u_2 = -1$ during time $\theta_1$ then $u_2 = 1$ during time $\theta_2$ and finally $u_2 = -1$ during time $\theta_3$. The last one, denoted $\gamma(t)$, has $u_2 = -1$ during time $\gamma_1$ then $u_2 = 1$ during time $\gamma_2$ then $u_2 = -1$ during time $\gamma_3$ and finally $u_2 = 1$ during time $\gamma_4$. One prove easily that, denoting $s_{\epsilon} = \epsilon_1 + \epsilon_2 + \epsilon_3$, $s_{\theta} = \theta_1 + \theta_2 + \theta_3$ and $s_{\gamma} = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$,

\[
\begin{align*}
    x(\epsilon) &= \epsilon_1 + \epsilon_3 + a_{200}\epsilon_1^2\epsilon_2 + a_{110}\epsilon_1\frac{\epsilon_2^2}{2} + o(s_{\epsilon}^3), \\
    y(\epsilon) &= \epsilon_2 + b_{200}\epsilon_1^2\epsilon_2 + b_{110}\epsilon_1\frac{\epsilon_2^2}{2} + o(s_{\epsilon}^3), \\
    z(\epsilon) &= \epsilon_1\epsilon_2 + o(s_{\epsilon}^3),
\end{align*}
\]
hence it exists easily that, since along the curve $\gamma$

$$a_{110}\theta_2\theta_3(\theta_1 + \theta_3) + o(s_\theta^3),$$

$$b_{110}\theta_2\theta_3(\theta_1 + \theta_3) + o(s_\theta^3),$$

$$\gamma_2 + \gamma_4 + a_{200}\gamma_2^2\gamma_3 + a_{110}\gamma_2\gamma_3(\gamma_1 + \frac{\gamma_3}{2}) + o(s_\theta^3),$$

$$\gamma_1 + \gamma_3 + b_{200}\gamma_2^2\gamma_3 + b_{110}\gamma_2\gamma_3(\gamma_1 + \frac{\gamma_3}{2}) + o(s_\theta^3),$$

$$\gamma_2\gamma_3 + o(s_\theta^3),$$

but since at $s_\theta$ and $s_r$ the extremals are supposed to be at $(x, y, z)$ then one gets

$$x + y = s_\theta + (a_{200} + b_{200})\theta_2\theta_3 + \frac{a_{110} + b_{110}}{2}\theta_2\theta_3(\theta_1 + \frac{\theta_3}{2}) + o(s_\theta^3)$$

$$= s_\theta + (a_{200} + b_{200})\theta_2\theta_3 + \frac{a_{110} + b_{110}}{2}(2\theta_1 + \theta_3)$$

$$x + y = s_\gamma + (a_{200} + b_{200})\gamma_2\gamma_3 + \frac{a_{110} + b_{110}}{2}\gamma_2\gamma_3(\gamma_1 + \frac{\gamma_3}{2}) + o(s_\gamma^3)$$

$$= s_\gamma + (a_{200} + b_{200})\gamma_2\gamma_3 + \frac{a_{110} + b_{110}}{2}(2\gamma_1 + \gamma_3).$$

Hence we deduce

$$\frac{s_\gamma - s_\theta}{z} = (a_{200} + b_{200})(\theta_2 - \gamma_2) + \frac{a_{110} + b_{110}}{2}(2\theta_1 + \theta_3 - (2\gamma_1 + \gamma_3)) + o(x + y)$$

$$= (a_{200} + b_{200})(\theta_2 - \gamma_2) + \frac{a_{110} + b_{110}}{2}(\theta_1 - \gamma_1) + o(x + y)$$

since $\theta_1 + \theta_3 = y + o(x + y)$ and $\gamma_1 + \gamma_3 = y + o(x + y)$.

Now, we should analyse the relation between $\gamma_2$ and $\gamma_3$. One can prove that along the curve $\gamma$, during the second bang, $\phi_2 = f_{41} - f_{51} + o(t) = -\frac{a_{110} + b_{110}}{2} + o(t)$ and during the second bang $\phi_2 = f_{41} + f_{51} + o(t) = -(a_{200} + b_{200}) + o(t)$. One proves easily that, since $\phi_2 = 0$ at the extremity of each of these intervals, this implies that

$$\gamma_3 - \gamma_2 = \frac{a_{200} + b_{200}}{a_{110} + b_{110}} + o(x + y),$$

hence it exists $\lambda > 0$ such that $\gamma_3 = \lambda(a_{200} + b_{200}) + o((x + y)^2)$ and $\gamma_2 = -\lambda\frac{a_{110} + b_{110}}{2} + o((x + y)^2)$. As a consequence

$$\lambda\frac{s_\gamma - s_\theta}{z} = \gamma_3(\theta_2 - \gamma_2) - \gamma_2(\theta_1 - \gamma_1) + o((x + y)^2)$$

$$= \gamma_3(x - \gamma_2) - \gamma_2(\gamma_3 - \theta_3) + o((x + y)^2)$$

$$= \gamma_3x + \gamma_2\frac{z}{x} - 2\gamma_2\gamma_3 + o((x + y)^2)$$

$$= \gamma_3x + \frac{z}{\gamma_3x} - 2z + o((x + y)^2)$$

hence

$$\lambda\frac{s_\gamma - s_\theta}{z^2} = \frac{\gamma_3x}{z} + \frac{z}{x\gamma_3} - 2 + o(1)$$

hence $s_\gamma - s_\theta$ is strictly positive except maybe when $\gamma_3 \sim \theta_3$ and $\gamma_2 \sim x$. 

But comparing with the curve \( \epsilon \) we get that \( s_\epsilon - s_\theta > 0 \) except maybe when \( \gamma_2 \sim \epsilon_1 \) and \( \gamma_3 \sim y \). Finally we can conclude that such an extremal \( \gamma \) is not optimal. The same proof can be done for the extremals with four bangs following first \( G_1 \), then \( G_2 \), then \( G_1 \) and finally \( G_2 \). And no extremal with three switches on the same control can be optimal.

Comparing the curves \( \epsilon \) and \( \theta \) one gets

\[
\frac{s_\epsilon - s_\theta}{z(1 - \frac{x}{xy})} = (a_{200} + b_{200})x + \frac{a_{110} + b_{110}}{2}y + o(x + y).
\]

Hence, since \( a_{200} + b_{200} > 0 \) and \( \frac{a_{110} + b_{110}}{2} < 0 \) the curve \( \epsilon \) is optimal for \( y > -2\frac{a_{200} + b_{200}}{a_{110} + b_{110}}x + o(x) \) and we find that there is a cut locus which is tangent at 0 to the plane

\[
(a_{200} + b_{200})x + \frac{a_{110} + b_{110}}{2}y = 0.
\]

### 7.2. Other extremals generating cut locus.

One show easily that, for extremals with \( u_1 \equiv 1 \), there is also cut locus only if \( |f_{41}| < -f_{51} \), that is if \( a_{200} + b_{200} > 0 \) and \( a_{110} + b_{11} < 0 \), and the tangent plane is the same.

In the cases \( u_2 \equiv 1 \) and \( u_2 \equiv -1 \) then there is cut locus only if \( |f_{52}| < f_{42} \), that is if \( b_{110} - a_{110} > 0 \) and \( b_{200} - a_{200} > 0 \). In this last case the tangent plane at 0 is

\[
(a_{200} - b_{200})x + \frac{a_{110} - b_{110}}{2}y = 0.
\]

### 7.3. Cut locus generated by extremals with \( \lambda_z(0) \sim 0 \).

As a consequence of the previous computations, we can describe the part of the local cut locus generated by the extremals with \( \lambda_0(0) \sim 0 \).

- if \( a_{200} + b_{200} < 0 \) or \( a_{110} + b_{11} > 0 \) and \( b_{110} - a_{110} < 0 \) or \( b_{200} - a_{200} < 0 \) then this part of the local cut locus is empty.
- if \( a_{200} + b_{200} > 0 \) and \( a_{110} + b_{11} < 0 \) and \( b_{110} - a_{110} < 0 \) or \( b_{200} - a_{200} < 0 \) then this part of the cut locus writes

\[
\{(x, -2\frac{a_{200} + b_{200}}{a_{110} + b_{110}}x + o(x), z) \mid 0 \leq z \leq -2\frac{a_{200} + b_{200}}{a_{110} + b_{110}}x^2 + o(x^2)\}
\]

- if \( a_{200} + b_{200} < 0 \) or \( a_{110} + b_{11} > 0 \) and \( b_{110} - a_{110} > 0 \) and \( b_{200} - a_{200} > 0 \) then this part of the cut locus writes

\[
\{(x, -2\frac{a_{200} - b_{200}}{a_{110} - b_{110}}x + o(x), z) \mid 0 \geq z \geq -2\frac{a_{200} - b_{200}}{a_{110} - b_{110}}x^2 + o(x^2)\}
\]

- if \( a_{200} + b_{200} > 0 \) and \( a_{110} + b_{11} < 0 \) and \( b_{110} - a_{110} > 0 \) and \( b_{200} - a_{200} > 0 \) then this part of the local cut locus is the union of the two previous sets.

Finally we can propose the picture of this part of the cut locus in Figure 18

### 7.4. Stability of the pictures.

The same questions as in subsection 6.8 are open here: is the picture we obtain with this level of developments of series stable? Are the singularities we find stable for a good class of functions and a good notion of stability? Again, our conviction is that the picture we give of the front stay the same generically but it is still an open question.
Figure 18. Part of the cut locus generated by the extremal with $\lambda_2(0) \sim 0$ when $|f_{41}| < -f_{51}$ and $|f_{52}| < f_{42}$

REFERENCES

[1] A. Agrachev, B. Bonnard, M. Chyba and I. Kupka, Sub-Riemannian sphere in Martinet flat case, ESAIM Control Optim. Calc. Var., 2 (1997), 377–448.
[2] A. A. Agrachev, U. Boscain, G. Charlot, R. Ghezzi and M. Sigalotti, Two-dimensional almost-Riemannian structures with tangency points, Ann. Inst. H. Poincaré Anal. Non Linéaire, 27 (2010), 793–807.
[3] A. A. Agrachev, El-H. Chakir El-A. and J. P. Gauthier, Sub-Riemannian metrics on $\mathbb{R}^3$, In Geometric Control and Non-Holonomic Mechanics (Mexico City, 1996), CMS Conf. Proc., Vol. 25, Amer. Math. Soc., Providence, RI, 1998, 29–78.
[4] A. A. Agrachev and J.-P. Gauthier, On the subanalyticity of Carnot-Caratheodory distances, Ann. Inst. H. Poincaré Anal. Non Linéaire, 18 (2001), 359–382.
[5] A. A. Agrachev and Y. L. Sachkov, Control Theory from the Geometric Viewpoint, Encyclopedia of Mathematical Sciences, Vol. 87, Springer-Verlag, Berlin, 2004.
[6] E. A.-L. Ali and G. Charlot, Local (sub)-Finslerian geometry for the maximum norms in dimension 2, J. Dyn. Control Syst., 25 (2017), 457–490.
[7] D. Barilari, U. Boscain, E. Le Donne and M. Sigalotti, Sub-Finsler structures from the time-optimal control viewpoint for some nilpotent distributions, J. Dyn. Control Syst., 23 (2017), 547–575.
[8] D. Barilari, U. Boscain, G. Charlot and R. W. Neel, On the heat diffusion for generic Riemannian and sub-Riemannian structures, Int. Math. Res. Not. IMRN, (2017), 4639–4672.
[9] D. Barilari, U. Boscain and R. W. Neel, Small-time heat kernel asymptotics at the sub-Riemannian cut locus, J. Differential Geom., 92 (2012), 373–416.
[10] A. Bellaïche, The tangent space in sub-Riemannian geometry, In Sub-Riemannian Geometry, Progr. Math., Vol. 144, Birkhäuser, Basel, 1996, 1–78.
[11] G. Ben Arous, Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus, Ann. Sci. École Norm. Sup. (4), 21 (1988), 307–331.
[12] G. Ben Arous and R. Léandre, Décroissance exponentielle du noyau de la chaleur sur la diagonale. II, Probab. Theory Related Fields, 90 (1991), 377–402.
[13] B. Bonnard, M. Chyba and E. Trelat, Sub-Riemannian geometry, one-parameter deformation of the martinet flat case, J. Dynam. Control Systems, 4 (1998), 59–76.
[14] B. Bonnard, G. Charlot, R. Ghezzi and G. Janin, The sphere and the cut locus at a tangency point in two-dimensional almost-Riemannian geometry, J. Dyn. Control Syst., 17 (2011), 141–161.
[15] B. Bonnard and M. Chyba, Méthodes géométriques et analytiques pour étudier l’application exponentielle, la sphère et le front d’onde en géométrie sous-Riemannienne dans le cas Martinet, ESAIM Control Optim. Calc. Var., 4 (1999), 245–334.
[16] U. Boscain, G. Charlot and R. Ghezzi, Normal forms and invariants for 2-dimensional almost-Riemannian structures, Differential Geom. Appl., 31 (2013), 41–62.
[17] U. Boscain, G. Charlot, R. Ghezzi and M. Sigalotti, Lipschitz classification of almost-Riemannian distances on compact oriented surfaces, *Journal of Geometric Analysis*, 23 (2013), 438–455.

[18] U. Boscain, T. Chambrion and G. Charlot, Nonisotropic 3-level quantum systems: Complete solutions for minimum time and minimum energy, *Discrete Contin. Dyn. Syst. Ser. B*, 5 (2005), 957–990.

[19] E. Breuillard and E. Le Donne, On the rate of convergence to the asymptotic cone for nilpotent groups and subFinsler geometry, *Proc. Natl. Acad. Sci. USA*, 110 (2013), 19220–19226.

[20] G. Charlot, Quasi-contact s-r metrics: Normal form in \( \mathbb{R}^{2n} \), wave front and caustic in \( \mathbb{R}^4 \), *Acta App. Math.*, 74 (2002), 217–263.

[21] W.-L. Chow, Über systeme von linearen partiellen differentialgleichungen erster ordnung, *Math. Ann.*, 117 (1939), 98–105.

[22] J. N. Clelland and C. G. Moseley, Sub-finsler geometry in dimension three, *Differ. Geom. Appl.*, 24 (2006), 628–651.

[23] J. N. Clelland, C. G. Moseley and G. R. Wilkens, Geometry of sub-Finsler Engel manifolds, *Asian J. Math.*, 11 (2007), 699–726.

[24] E. El-Alaoui, J.-P. Gauthier and I. Kupka, Small sub-Riemannian balls on \( \mathbb{R}^3 \), *J. Dynam. Control Systems*, 2 (1996), 359–421.

[25] A. F. Filippov, On some questions in the theory of optimal regulation: Existence of a solution of the problem of optimal regulation in the class of bounded measurable functions, *Vestnik Moskov. Univ. Ser. Mat. Meh. Astr. Fiz. Him.*, 1959 (1959), 25–32.

[26] M. W. Hirsch, *Differential Topology*, Graduate Texts in Mathematics, Vol. 33 Springer-Verlag, New York, 1994.

[27] R. Léandre, Majoration en temps petit de la densité d’une diffusion dégénérée, *Probab. Theory Related Fields*, 74 (1987), 289–294.

[28] R. Léandre, Minoration en temps petit de la densité d’une diffusion dégénérée, *J. Funct. Anal.*, 74 (1987), 399–414.

[29] P. K. Rashevsky, About connecting two points of complete nonholonomic space by admissible curve, *Uch. Zap. Ped. Inst. Libknehta*, 2 (1938), 83–94.

[30] M. Sigalotti, Bounds on time-optimal concatenations of arcs for two-input driftless 3D systems, preprint, 2019, arXiv:1911.10811.

Received April 2019; revised August 2020.

E-mail address: gregoire.charlot@univ-grenoble-alpes.fr
E-mail address: entisar.al@univ-grenoble-alpes.fr