HORIZONTAL MEAN CURVATURE FLOW AND NONLINEAR PDEs IN SUB-RIEMANNIAN SPACES

Raffaele Grande

Submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy.

Supervisors: Dr.Federica Dragoni and Prof.Nicolas Dirr

March 2021

Contents

Introduction

The *horizontal mean curvature flow* is an evolution of a hypersurface, which is interesting not only in a theoretical framework but also in applied science, for example in neurogeometry and computer science (e.g. [13, 14, 15]). The associated equation, roughly speaking, describes the motion of a hypersurface embedded in a *sub-Riemannian geometry* (e.g. the Heisenberg group or a Carnot group, see [4, 43]) in relation to its horizontal mean curvature.

A sub-Riemannian geometry is defined as a triple $(M, \mathcal{H}, g)$, where M is a smooth manifold (usually $\mathbb{R}^N$), $\mathcal{H}$ is a distribution [1], i.e. $\mathcal{H} \subset TM$ and g is a sub-Riemannian metric defined on the fibres of $\mathcal{H}$. In this kind of geometries not all the curves are admissible. The admissible curves are the curves $\gamma : [0, T] \to M$ absolutely continuous and such that $\dot{\gamma}(t) \in \mathcal{H}_{\gamma(t)}$, for almost every $t \in [0, T]$.

The presence of non admissible curves leads to some technical difficulties at some points. In particular considering a hypersurface embedded in a sub-Riemannian geometry, there are some points called *characteristic* in which the *horizontal normal* (i.e. the renormalized projection of the Euclidean normal onto the horizontal geometry) is not defined, even if the Euclidean normal is well defined. In order to deal with these points, we introduce a *Riemannian*

[1] In mathematics the word *distribution* has different meanings, depending on if we are talking about, for example, a geometrical distribution, an analytical distribution or the distribution of a (probabilistic) random variable

approximation of the sub-Riemannian geometry $(M, \mathcal{H}, g)$, i.e. we complete the distribution $\mathcal{H}$ with a suitable number of vector fields depending on a parameter $\varepsilon > 0$ to generate the whole tangent space at any point, in order to obtain a Riemannian geometry.

Another problem that is present in the evolution of these manifolds is the change of topological properties of the starting manifolds. To be more precise, the surfaces which evolve by the mean curvature flow, even in the Euclidean and Riemannian case, may develop singularities, as in the case of the dumbbell (for further details [31]).

In order to deal with these singularities, we will consider the *generalized mean curvature flow* and we will use the notion of *viscosity solution* (for further details [17, 18, 19, 20]) in order to solve it.

There are two established approaches to solve mean curvature flow equation (which is called level set equation for the evolution by mean curvature flow): the approach of Evans and Spruck ([31]) is, roughly speaking, to complete the equation using an approximation which depends on an $\varepsilon > 0$ in order to avoid singularities and, using the property of stability of viscosity solutions, find a solution; the approach of Chen, Giga and Goto ([11]) is related to solve in the viscosity sense for upper and lower semicontinuous envelopes of the PDEs.

Furthermore, Buckdahn, Cardaliaguet and Quincampoix ([27]) and, independently, Soner and Touzi ([49, 52]) in 2001 found a stochastic representation for the solution of the horizontal mean curvature flow, a result which was extended to the sub-Riemannian case by Dirr, Dragoni and Von Renesse ([22]) in 2006.

Using the Riemannian approximation it is possible to define the *approximated Riemannian mean curvature flow*. This is a different approach from the one used in [22], to study the same equation and it was introduced by Citti and collaborators (see e.g. [13]).

The *aim of this thesis* is to find the relation between the solution of approximated mean curvature flow and the solution of the horizontal one as $\varepsilon \to 0^+$ by using for the both cases the associated stochastic formulas. More precisely, we define the function $V : [0,T] \times \mathbb{R}^N \to \mathbb{R}$ as

$$V(t,x) = \inf_{\nu \in \mathcal{A}} ess \sup_{\omega \in \Omega} g(\xi^{t,x,\nu}(T)(\omega)), \tag{1}$$

where $\xi^{t,x,\nu}$ is the solution of the controlled dynamics

$$\begin{cases} d\xi^{t,x,\nu}(s) = \sqrt{2}\sigma^T(\xi^{t,x,\nu}(s)) \circ dB^{\nu}(s), & s \in (t,T], \\ dB^{\nu}(s) = \nu(s) dB_m(s), & s \in (t,T], \\ \xi^{t,x,\nu}(t) = x, \end{cases} \tag{2}$$

with N dimension of the space and m ($m < N$) dimension of the distribution $\mathcal{H}$ at every points and controls

$$\mathcal{A} = \{\nu : [t,T] \to Sym(m) \mid \nu \geq 0,\ I_m - \nu^2 \geq 0,\ Tr(I_m - \nu^2) = 1\},$$

and $\sigma(x) = [X_1(x), \dots, X_m(x)]^T$ the matrix associated to the sub-Riemannian geometry in which is embedded the hypersurface, i.e. $X_1, \dots, X_m$ vector fields spanning the distribution $\mathcal{H}$. Similarly

$$V^{\varepsilon}(t,x) = \inf_{\nu \in \mathcal{A}_1} ess \sup_{\omega \in \Omega} g(\xi_{\varepsilon}^{t,x,\nu_1}(T)(\omega)), \tag{3}$$

where $\xi_{\varepsilon}^{t,x,\nu_1}$ is the solution of a dynamics similar to (2) adapted to the Riemannian approximation, and the controls are given by

$$\mathcal{A}_1 = \{\nu_1 : [t,T] \to Sym(N) \mid \nu \geq 0,\ I_N - \nu^2 \geq 0,\ Tr(I_N - \nu^2) = 1\}.$$

Since solving the minimization problem (1) is particularly difficult due to the presence of an essential supremum, we consider the L^p approximation of value function (1) and (3) (see Definition 5.24 for further details). Therefore we are interested in finding the optimal controls for the L^p approximation in the Heisenberg group and the approximated Heisenberg group and then in using them to show the convergence between V^ε and V for a suitable $\varepsilon = \varepsilon(p)$. This result implies that, even though the comparison principle is mostly still open, the solution found in [8] is the same solution in [22].

The thesis is organized as follows:

- In Chapter 1 we introduce the sub-Riemannian geometries and their properties looking in particular to the Riemannian approximation.
- In Chapter 2 we recall some stochastic tools, as the Brownian motion and the stochastic integral, in both Itô and Stratonovich forms.
- In Chapter 3 we introduce the Euclidean mean curvature flow, the definition of viscosity solution and we state shortly the result found in [27] which connects this evolution with an associated stochastic optimal control problem.
- In Chapter 4 we define the sub-Riemannian mean curvature flow and the Riemannian approximated mean curvature flow.
- In Chapter 5 we prove that the V^ε as defined in (3) is the viscosity solution of approximated Riemannian mean curvature flow. This result is contained in the preprint [35].
- In Chapter 6 we find the asymptotic behaviour, for large p of the p-optimal control for the stochastic dynamic associated to the Heisenberg group. This result is contained in a published article in *Nonlinear Analysis*,

in collaboration with N. Dirr and F. Dragoni ([21]). According to our knowledge, this result was open even in the Euclidean case.

- In Chapter 7 we generalise the results given in Chapter 6 and find the asymptotic behaviour of the p-optimal control for the stochastic dynamics associated to the approximated Heisenberg group. This result is contained in the preprint [25] in collaboration with F. Dragoni.

- In Chapter 8 we obtain a limsup estimate for V^ε. We also briefly sketch how to use the results contained in Chapter 7 to prove the corresponding liminf estimate. This estimate is more technical and an improvement of the results in [25] may be necessary. This part is still a work in progress in collaboration of F. Dragoni.

- In the Conclusion we sum up all the results found and we propose a short plan of possible future research. At the end there is a short Appendix containing some real analysis inequalities, used in the proofs of the main results.

Chapter 1

Sub-Riemannian manifolds

In this chapter we introduce the main definitions related to sub-Riemannian geometry. This type of geometries has found some applications in IT and neurogeometry (see [13]) and they are geometries in which not all the curves on a manifold are admissible.

1.1 Riemannian manifold

A *Riemannian manifold* is a smooth manifold endowed with a metric which allows us to measure the length of a curve on a manifold. We will start recalling the definitions of topology and continuous function (for further details, see e.g. [27]).

Definition 1.1. *Let X be a set and $\tau \subset \mathcal{P}(X)$ a collection of subset of X which respects the following axioms:*

- $X, \emptyset \in \tau$.
- *If $\{A_n\} \in \tau$ is a finite or countable collection of sets, then $\cup A_n \in \tau$.*
- *If $A_1, A_2 \in \tau$ then $A_1 \cap A_2 \in \tau$.*

Then (X, τ) *is called* topological space *and the elements of* τ *are called* open sets*. The complement of an open set is called* closed set*.*

Example 1.2. *Let* X *be a set. We take* $A \subset X$ *and define* $\tau = \{X, A, X \setminus A, \emptyset\}$*.* (X, τ) *is a topological space. We observe that* $A, X \setminus A$ *are open and closed by definition.*

Definition 1.3. *Let* (X, τ_1) *,* (Y, τ_2) *be two topological spaces and* $f : X \to Y$ *a function. The function* f *is called continuous if and only if*

$$A \in \tau_2 \Rightarrow f^{-1}(A) \in \tau_1.$$

Definition 1.4. *Let* (X, τ) *be a topological space and* $x \in X$*. A subset* $\mathcal{I}(x) \subset X$ *is called neighbourhood of* x *if there exists an open set* U *(i.e.* $U \in \tau$*) such that*

$$x \in U \subset \mathcal{I}(x).$$

Definition 1.5. *Let* X *be a topological space, this space is called an* Hausdorff space *if and only if*

$$\forall\ x, y \in X \ \text{ s.t. } x \neq y \ \exists\ \mathcal{I}(x) \ \text{ and } \mathcal{I}(y) \quad \text{s.t. } \mathcal{I}(x) \cap \mathcal{I}(y) = \emptyset$$

where $\mathcal{I}(x)$ *is a neighbourhood of* x *and* $\mathcal{I}(y)$ *is a neighbourhood of* y*.*

Example 1.6. *Every topological space* X *with the discrete topology is Hausdorff because the points are open sets.*

Counterexample 1.7. *Every topological space* X *with the trivial topology and at least two elements is not Hausdorff. In fact, considering two different elements* $x, y \in X$ *it is not possible to find two distinct neighbourhoods by the definition of trivial topology.*

Counterexample 1.8. $\mathbb{R}$ *with the topology*

$$\tau = \{A \subset \mathbb{R} |\ A = (a, +\infty) \ \text{with } a \in \overline{\mathbb{R}}\} \quad \text{where } \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$$

is not an Hausdorff space. We remark that the $\mathbb{R}$ *and* $\emptyset$ *are in the topology due to the definition of* $\overline{\mathbb{R}}$.

We now introduce the definition of the distance on a topological space.

Definition 1.9. *Let* X *be a non empty set, then the* distance *is a function*

$$d : X \times X \to [0, \infty)$$

such that it holds true:

- $d(x, y) = 0 \Leftrightarrow x = y$,
- $d(x, y) = d(y, x) \quad \forall\ x, y \in X$,
- $d(x, y) \leq d(x, z) + d(z, y) \quad \forall\ x, y, z \in X$.

Remark 1.10. *Every metric space* (X, d) *is Hausdorff because, assuming that* $x \neq y$, *we can write* $\mathcal{I}(x) = B(x, |\frac{x-y}{2}|)$ *and and* $\mathcal{I}(y) = B(y, |\frac{x-y}{2}|)$ *and so* $\mathcal{I}(x) \cap \mathcal{I}(y) = \emptyset$ *where* $B(x, r) = \{y \in X \mid d(x, y) < r\}$.

The definition of *homeomorphism* is crucial to understand how we can work with topological spaces without transformations which can break or cut them.

Definition 1.11. *Let* (X, τ_1) *and* (Y, τ_2) *be two topological spaces, a function* $f : X \to Y$ *is called* homeomorphism *if and only if:*

- f *is a bijective function,*
- f *is a continuous function,*
- f^{-1} *is a continuous function.*

Example 1.12. *Let* (X, τ) *be a topological space, then the identity on* X *is trivially an homeomorphism.*

Example 1.13. *Given the two intervals (0,1) and* (a, b) *where* $a < b$ *and* $a, b \in \mathbb{R}$ *Let* $((0, 1), d)$, $((a, b), d)$ *be two metric spaces with* d *the standard Euclidean distance in* $\mathbb{R}$ *and* $(0, 1)$ *and* $(a, b) \subset \mathbb{R}$ *then these spaces are homeomorphic by the following function*

$$f : (0, 1) \to (a, b), \quad t \to at + b(1 - t).$$

These are some properties which are preserved by homeomorphism between two topological spaces, such as being *connected* and the *compactness.*

Definition 1.14. *Let* (X, τ) *be a topological space. This space is called* connected *if and only if the only open and closed sets are* X *and the empty set.*

Example 1.15. $\mathbb{R}^N$ *is a connected space with Euclidean topology.*

Counterexample 1.16. *Let* (X, τ) *be a topological space where* τ *is the discrete topology. Then this space is not connected.*

Definition 1.17. *Let* (X, τ) *be a topological space. This space is called* path connected *if and only if for every two points* $x, y \in X$ *there exists a continuous function* $\alpha : [0, 1] \to X$ *such that* $\alpha(0) = x$ *and* $\alpha(1) = y$.

Remark 1.18. *If a topological space is path connected then it is connected.*

Definition 1.19. *Let* (X, τ) *be a topological space. This space is called* compact *if for every arbitrary collection of open subsets* $\{U_\alpha\}_{\alpha \in A}$ *of* X *where* $A \subset \mathbb{N}$ *such that*

$$X = \bigcup_{\alpha \in A} U_\alpha$$

there is a finite subset $J \subset A$ *of elements such that*

$$X = \bigcup_{i \in J} U_i.$$

Now we are ready to give the definition of *manifold.*

Definition 1.20. *A set M is called* N-dimensional topological manifold *if it is a Hausdorff space and every point $x \in M$ has a neighbourhood which is homeomorphic to an open subset of the Euclidean space $\mathbb{R}^N$.*

We introduce now the definition of homotopy and simply connected manifold.

Definition 1.21. *Let $f, g : X \to Y$ be two continuous functions from a topological space to an other one. The* homotopy *is a continuous (in both arguments) function H such that*

$$H : X \times [0,1] \to Y, \quad H(x,0) = f(x), \quad H(x,1) = g(x), \quad \forall x \in X.$$

Definition 1.22. *A topological space X is called* simply connected *if and only if X is path connected and for all $p, q \in X$ and for every pair of arcs (i.e continuous function from $[0,1]$ to X) which start from p and ended to q there exists an homotopy which transforms the first arc in the second one.*

Example 1.23. *$\mathbb{R}^N$ with $N \in \mathbb{N}$ and S^N with $N \geq 2$ are simply connected.*

Counterexample 1.24. *S^1 and $\mathbb{R}^2 \setminus \{(0,0)\}$ are not simply connected.*

Let us give now the definition of smooth manifold and diffeomorphism.

Definition 1.25. *Let M be a topological space, $U \subset M$ and $V \subset \mathbb{R}^N$ an open set with respect to the standard topology of $\mathbb{R}^N$. An homeomorphism $x : U \to V$ such that $x(u) = (x_1(u), \ldots, x_n(u))$ is called a* coordinate system *on U. Let $A \subset \mathbb{N}$. $\{U_\alpha, x_\alpha\}_{\alpha \in A}$ is called* chart *and x^{-1} is called a* parametrization*. An* atlas *on M is a collection of charts $\{U_\alpha, x_\alpha\}_{\alpha \in A}$ such that $\cup_{\alpha \in A} U_\alpha = M$ and the homeomorphisms $x_\beta \circ x_\alpha : x_\alpha^{-1}(U_\beta \cap U_\alpha) \to x_\beta(U_\beta \cap U_\alpha)$ are the transitions maps.*

Definition 1.26. *A* smooth manifold *M is an* N-dimensional topological manifold *such that it admits an atlas $\{U_\alpha, x_\alpha\}_{\alpha \in A}$, $x_\alpha : U_\alpha \to \mathbb{R}^N$, $N \in \mathbb{N}$ and all transition maps are C^∞ diffeomorphisms.*

Example 1.27. *S^1 is a smooth manifold. We can introduce two charts $\phi_N : \mathbb{R} \to S^1$ and $\phi_S : \mathbb{R} \to S^1$ such that*

$$\phi_N : u_N \to \left(\frac{2u_N}{u_N^2 + 1}, \frac{u_N^2 - 1}{u_N^2 + 1} \right)$$

and

$$\phi_S : u_S \to \left(\frac{2u_S}{u^2 + 1}, \frac{1 - u_S^2}{u_S^2 + 1} \right).$$

We have $V_N = \mathbb{R}$, $U_N = S^1 \setminus \{(0,1)\}$, $U_S = S^1 \setminus \{(1,0)\}$ and $U_N \cap U_S = S^1 \setminus \{(0,1),(1,0)\}$, and so we obtain

$$\phi_N^{-1} \circ \phi_S : \mathbb{R} \setminus \{0\} \to \mathbb{R} \setminus \{0\}, \quad u_S \to u_N = \frac{1}{u_S}.$$

Definition 1.28. *Let M_1, M_2 be smooth manifolds with dimension N and m. A map $\phi : M_1 \to M_2$ is* smooth *if and only if for any $p \in M_1$ and for any parametrization (V, y) at $\phi(p) \in M_2$, there exists a local parametrization (U,x) at $p \in M_1$ such that:*

- *$\phi(x(U)) \subset y(V)$,*
- *$y^{-1} \circ \phi \circ x : U \subset \mathbb{R}^N \to V \subset \mathbb{R}^m$ is a smooth map.*

Example 1.29. *The map $\exp : \mathbb{R} \to \mathbb{R}^+$ is a smooth map, in fact we can take $(\mathbb{R}, Id)$ and $(\mathbb{R}^+, Id)$ like charts and we can see that $\exp(\mathbb{R}) = \mathbb{R}^+$ and $Id^{-1} \circ \exp \circ Id = exp$ is a smooth map.*

Example 1.30. *The stereographic projection from $S^2 \setminus \{(0,0,1)\}$ to $\mathbb{R}^2$, i.e.*

$$\phi : S^2 \setminus \{(0,0,1)\} \to \mathbb{R}^2, \quad (x,y,z) \to \left(\frac{x}{1-z}, \frac{y}{1-z} \right)$$

is a smooth function.

Definition 1.31. *Let $f : M_1 \to M_2$, f is a* diffeomorphism *if and only if f is a bijective map and f and f^{-1} are smooth.*

Example 1.32. *It is possible to prove that the stereographic projection ϕ is a diffeomorphism, in fact the invertible map is*

$$\phi^{-1} : \mathbb{R}^2 \to S^2 \backslash \{(0,0,1)\}, \quad (x,y) \to \left(\frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2}, \frac{-1+x^2}{1+x^2+y^2} \right),$$

which is smooth.

Counterexample 1.33. *The function $f : \mathbb{R} \to \mathbb{R}, \quad t \to t^3$ is differentiable but not a diffeomorphism because the inverse (i.e. $f^{-1} : \mathbb{R} \to \mathbb{R}, \quad t \to t^{\frac{1}{3}}$) is not a differentiable map.*

Absolute continuity and curves on a manifold

Definition 1.34 (Absolute continuity)**.** *Let $I \subset \mathbb{R}$, a function $f : I \to \mathbb{R}$ is* absolutely continuous *on I if for every positive number ε, there is a positive number δ such that whenever a finite sequence of pairwise disjoint sub-intervals $(x_k, y_k)_{k=1}^N$ of I with x_k, $y_k \in I$*

$$\sum_{k=1}^{N} |y_k - x_k| < \delta \Rightarrow \sum_{k=1}^{N} |f(y_k) - f(x_k)| < \varepsilon.$$

Theorem 1.35 (Lebesgue's theorem for absolutely continuous functions)**.** *The following conditions for a function $f : [a,b] \to \mathbb{R}$ are equivalent:*

- *f is absolutely continuous,*
- *f has a derivative f' almost everywhere, the derivative is Lebesgue integrable and*

  $$f(x) = f(a) + \int_a^x f'(t)dt,$$

  for all $x \in [a,b]$.
- *There exists a Lebesgue integrable function g on $[a,b]$ such that*

  $$f(x) = f(a) + \int_a^x g(t)dt,$$

  for all $x \in [a,b]$. As consequence, it holds true $g = f'$ almost everywhere.

Definition 1.36. *Let us consider a function $\gamma : [0,T] \to M$. We say that γ is absolutely continuous if there exists a partition $0 = t_0 < \cdots < t_n = T$ and charts (U_i, ϕ_i) for $1 \leq i \leq n$ such that $\gamma([t_{i-1}, t_i]) \subset U_i$ and $\phi_i \circ \gamma|_{[t_{i-1},t_i]}$ is absolutely continuous.*

Definition 1.37. *Let M be a smooth manifold and $I \subset \mathbb{R}$ a real interval, then any absolutely continuous map $\gamma : [0,T] \to M$ is a* curve *on M. Let be $\gamma(0) = p \in M$ and $\mathcal{D}$ be the set of all the functions which are differentiable at the point p, then the tangent vector is defined as*

$$\dot{\gamma}(0) : \mathcal{D} \to \mathbb{R}, \quad f \to \dot{\gamma}(0) := \left.\frac{d(f \circ \gamma)}{dt}\right|_{t=0}. \tag{1.1}$$

Example 1.38. *Given S^2, an example of a curve $\gamma : [0, 2\pi] \to S^2$ is*

$$\gamma(t) = (\cos(t), \sin(t), 0).$$

Example 1.39. *Let $\mathbb{T}^2 = S^1 \times S^1$ be a 2-dimensional torus, this object can be written in a 3-dimensional space as*

$$\begin{cases} x(\theta, \phi) = (1 + \cos\theta)\cos\phi, \\ y(\theta, \phi) = (1 + \cos\theta)\sin\phi, \\ z(\theta, \phi) = \sin\theta \end{cases}$$

with $\theta, \phi \in [0, 2\pi]$. Then a curve on the torus may be

$$\gamma : [0, \pi] \to \mathbb{T}^2 \ , \ \gamma(t) = (x(t), y(t), z(t)) = (2\cos t, 2\sin t, 0).$$

Counterexample 1.40. *Given $\mathbb{R}$, the function $\gamma : [0,1] \to \mathbb{R}$*

$$\gamma(t) = \begin{cases} 1, & \text{if } x \in \mathbb{Q} \cap [0,1], \\ 0, & \text{if } \ x \in \mathbb{R} \setminus (\mathbb{Q} \cap [0,1]), \end{cases}$$

does not satisfy the Definition 1.37 because it is not continuous.

Definition 1.41. *Let us consider the vector space V over $\mathbb{R}$, the dual space V^* is defined as the set of all the bounded and linear maps $\phi : V \to \mathbb{R}$.*

Definition 1.42. *We call* tangent space *to the smooth manifold M at a point p the set of all the tangent vectors at the point p, i.e.*

$$T_pM = \{\dot\gamma(0) \mid \dot\gamma(0) \text{ satisfying } (1.3) \text{ and } \gamma(0) = p\}.$$

Let M be a smooth manifold, we call tangent bundle *of M the following $2N$-dimensional vector bundle*

$$TM = \{(p, v) \mid p \in M,\ v \in T_pM\}.$$

Remark 1.43 (Curves on a sphere)**.** *Let $\gamma : [0, T] \to \mathbb{R}^3$ be a curve. We wonder when such a curve is a curve on a sphere. For this purpose we can parametrize the sphere using the polar coordinates as*

$$\begin{cases} x = \cos u \sin v, \\ y = \sin u \sin v, \\ z = \cos v, \end{cases}$$

with $(u, v) \in [0, 2\pi) \times [0, \pi]$. From the notions of differential geometry in $\mathbb{R}^3$ we know that the tangent space at a point $p \in S^2$ is a plane which is tangent in p. The plane is generated by (having called $\phi(u, v)$ the parametrization of the sphere)

$$\phi_u = (\sin(u)\sin(v), \cos(u)\sin(v), 0),$$

$$\phi_v = (\cos(u)\cos(v), \sin(u)\cos(v), -\sin(v)),$$

and then, called $\gamma(t) = (x(t), y(t), z(t))$ a curve, we have that the curve is on the sphere if and only if

$$\begin{cases} \dot x(t) = -a\sin(u(t))\sin(v(t)) + b\cos(u(t))\cos(v(t)), \\ \dot y(t) = a\cos(u(t))\sin(v(t)) + b\sin(u(t))\cos(v(t)), \\ \dot z(t) = -b\sin(v(t)), \end{cases}$$

with $a, b \in \mathbb{R}$.

Example 1.44. $T\mathbb{R}^N = \cup_{p\in\mathbb{R}^N} T_p\mathbb{R}^N = \mathbb{R}^{2N}$.

Example 1.45. $TS^2 = \cup_p T_pS^2$ *that means that the tangent bundle of* S^2 *is, roughly speaking, the union of all planes which are tangent to the sphere.*

Definition 1.46. *Let* M_1, M_2 *be two different smooth manifolds,* $\phi : M_1 \to M_2$ *a differentiable map,* $p \in M_1$ *and* $v \in T_pM_1$. *Let us consider any curve* $\gamma : I \to M_1$ *such that* $\gamma(0) = p$ *and* $\dot\gamma(0) = v$. *Let* $\beta : I \to M_2$ *such that* $\beta := \phi \circ \gamma$. *The* differential *of* ϕ *at the point* p *is the linear map* $d\phi_p$ *defined as*

$$d\phi_p : T_pM_1 \to T_{\phi(p)}M_2, \quad v \mapsto d\phi_p(v) := \dot\beta(0).$$

These curves can be measured is some way, but before to do that, we have to introduce a *metric* which allows us to define an inner product in every point

Definition 1.47. *A* Riemannian metric $< , >$ *on a smooth manifold* M *is a smoothly chosen inner product* $< , >_p$: $T_pM \times T_pM \to \mathbb{R}$ *on each of the tangent spaces* T_pM *of* M. *In other words, for each* $p \in M$*:*

- $< u, v >_p = < v, u >_p$ *for all* $u, v \in T_pM$,
- $< u, u >_p \geq 0$ *for all* $u \in T_pM$,
- $< u, u >_p = 0$ *if and only if* $u = 0$.

Sometimes in the thesis we will use the symbol g_p instead of $< , >_p$.

Definition 1.48. *Let* $(M, < , >)$ *be a Riemannian manifold and* $\gamma : [0, T] \to M$ *absolutely continuous curve, we call* length of the curve M *the real functional*

$$l(\gamma) = \int_0^T < \dot\gamma(t), \dot\gamma(t) >_{\gamma(t)}^{\frac{1}{2}} dt.$$

Let $(M, < , >)$ *be a Riemannian manifold and* $p, q \in M$, *then the* Riemannian distance *between these two points is defined as*

$$d(p, q) = \inf\{l(\gamma) | \gamma \text{ a.c. curve on } M \text{ joining } p \text{ to } q\}. \tag{1.2}$$

Remark 1.49. *Let us remark that d as defined in (1.2) is a distance because*

- *$d(x,y) = 0 \Rightarrow x = y$ for all $x, y \in M$ is true. In fact, we obtain by definition*

$$\int_0^T < \dot{\gamma(t)}, \dot{\gamma(t)} >_{\gamma(t)}^{\frac{1}{2}} dt = 0 \Rightarrow \dot{\gamma(t)} = 0 \Rightarrow \gamma(t) = C \in \mathbb{R},$$

  where C is a constant.

- *$d(x,y) = d(y,x)$ for all $x, y \in M$ is true. In fact it is only necessary to remark that we can define a curve $\gamma(t) = \omega(1-t)$, where ω is a curve which connects x to y.*

- *$d(x,y) \leq d(x,z) + d(z,y)$ for all $x, y, z \in M$ is true. It follows from the fact that*

$$l(\gamma) \leq l(\alpha) + l(\beta),$$

  where α a curve which connects x to z and β a curve which connects z to y and γ the concatenation of α and β. To finish to prove the inequality we take the infimum before on the right hand side and then on the left one.

Example 1.50. *In $\mathbb{R}^N$ with standard Euclidean metric the distance is given by*

$$d(p,q) = \sqrt{|x_1 - y_1|^2 + \ldots |x_n - y_n|^2},$$

where $p = (x_1, \ldots, x_n)$ and $q = (y_1, \ldots, y_n)$, because the shortest path between two points is represented by the line which joins p and q.

Example 1.51. *Let S^1 be the standard Euclidean sphere in $\mathbb{R}^2$ and $< , >$ the Riemannian metric induced by the standard inner product in $\mathbb{R}^2$ (i.e. the restriction on $< , >$ on the vectors tangent along S^1), then the curve defined on this manifold will be*

$$\gamma : [0, 2\pi] \to S^1 \quad s.t \;\; \gamma(\theta) = (\cos(\theta), \sin(\theta)),$$

and

$$\dot{\gamma}(\theta) = (-\sin(\theta), \cos(\theta)),$$

and then $< \dot{\gamma}, \dot{\gamma} >= 1$ *we have that the distance between* $q = q(\theta_1)$ *and* $p = p(\theta_2)$ *is given by*

$$d(q,p) = |\theta_1 - \theta_2|.$$

Definition 1.52. *Let* M *be a smooth manifold, a* vector field *is a function*

$$F : M \to TM, \quad p \to (p, v)$$

such that the composition $\mathcal{P} \circ F = Id$ *where*

$$\mathcal{P} : TM \to M, \quad (p, v) \to p.$$

Example 1.53. *Given* S^1 *and* TS^1*, a vector field is, for* $t \in [0, 2\pi]$

$$F : S^1 \to TS^1, \quad p = (\cos(t), \sin(t)) \to (p, (-\sin(t), \cos(t))^T).$$

Definition 1.54. *Let* M *be a smooth manifold, we call* $\chi(M)$ *the union of all vector fields, we define a* covariant derivative *as the function*

$$\nabla : \chi(M) \times \chi(M) \to \chi(M), \quad (X, Y) \to \nabla_X Y$$

such that, given λ_1 *and* λ_2 *scalar values and* f_1*,* f_2 *smooth functions, the following properties are satisfied*

- $\nabla_X(\lambda_1 Y_1 + \lambda_2 Y_2) = \lambda_1 \nabla_X Y_1 + \lambda_2 \nabla_X Y_2$,
- $\nabla_{f_1 X_1 + f_2 X_2} Y = f_1 \nabla_{X_1} Y + f_2 \nabla_{X_2} y$,
- $\nabla_X(fY) = f \nabla_X Y + Y \nabla_X f$.

Counterexample 1.55. *Let the function be defined as*

$$\nabla : \chi(M) \times \chi(M) \to \chi(M), \quad (X, Y) \to \nabla_X Y = X + Y$$

then it is not a covariant derivative, in fact

$$\nabla_X(\lambda_1 Y_1 + \lambda_2 Y_2) = X + (\lambda_1 Y_1 + \lambda_2 Y_2) \neq \lambda_1 \nabla_X Y_1 + \lambda_2 \nabla_X Y_2.$$

1.2 Sub-Riemannian geometry

In Riemannian geometry all the curves on the manifold are admissible, this is due to the fact that we use as tangent bundle TM in the definition of the admissible curves. Unluckily in some applications we "lose" some curves (see e.g. [13], [14]) and so we have to introduce a "weaker" version of geometry which however satisfies some regularity conditions which allow us to connect always two points with a curve.

Definition 1.56. *Let M be a N-dimensional smooth manifold, we define for every point p a subspace of T_pM called $\mathcal{H}(p)$. We define the associated* distribution *as the subbundle of the tangent bundle given by*

$$\mathcal{H} = \{(p, v) \mid p \in M \quad v \in \mathcal{H}(p)\}.$$

Example 1.57. *For a Riemannian manifold it holds true $\mathcal{H} = TM$.*

Example 1.58. *[1-dimensional Heisenberg group] One classical example of distribution is the distribution associated to the Heisenberg group (which we will indicate with $\mathbb{H}^1$), which is given by the following vector fields on $\mathbb{R}^3$*

$$X(x, y, z) = \frac{\partial}{\partial x} - \frac{y}{2}\frac{\partial}{\partial z}, \quad Y(x, y, z) = \frac{\partial}{\partial y} + \frac{x}{2}\frac{\partial}{\partial z}.$$

In this case $\mathcal{H} = span(X, Y)$.

Definition 1.59. *Let M be a smooth manifold and $\mathcal{H} \subset TM$ a distribution, a* sub-Riemannian metric *on M is a Riemannian metric defined on the subbundle $\mathcal{H}$.*

Definition 1.60. *Let M be a smooth manifold and $\mathcal{H} \subset TM$ a distribution and g a Riemannian metric of M defined on the subbundle $\mathcal{H}$. A sub-Riemannian geometry is the triple $(M, \mathcal{H}, g)$.*

Remark 1.61. *All Riemannian manifolds are sub-Riemannian manifolds.*

Example 1.62. *Let* $\mathbb{R}^2$ *with the following vector fields*

$$X_1(x,y) = \frac{\partial}{\partial x}, \qquad X_2(x,y) = x\frac{\partial}{\partial y}$$

and the standard Euclidean metric induced on $\mathcal{H}_{(x,y)}$. $(\mathbb{R}^2, \mathcal{H}, g)$ *generates a sub-Riemannian geometry called the* Grušin plane.

Next we give the definition of admissible curves which are in this setting called horizontal.

Definition 1.63. *Let* $(M, \mathcal{H}, g)$ *be a sub-Riemannian geometry and* $\gamma : [0,T] \to M$ *an absolutely continuous curve,* γ *is an* horizontal curve *if and only if*

$$\dot{\gamma}(t) \in \mathcal{H}_{\gamma(t)}, \quad \text{for a.e. } t \in [0,T],$$

or, equivalently, if there exists a measurable function $h : [0,T] \to \mathbb{R}^N$ *such that*

$$\dot{\gamma}(t) = \sum_{i=1}^{m} h_i(t) X_i(\gamma(t)), \quad \text{for a.e. } t \in [0,T],$$

where $h(t) = (h_1(t), \ldots, h_m(t))$ *and* $X_1, \ldots, X_m$ *are some vector fields spanning the distribution* $\mathcal{H}$.

Remark 1.64. *We define a length-functional as*

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

where $\|\dot{\gamma}(t)\|_{\gamma(t)} = g_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))^{\frac{1}{2}}$. *If* $X_1, \ldots, X_m$ *are orthonormal respect to sub-Riemannian metric and*

$$\dot{\gamma}(t) = \sum_{i=1}^{r} h_i(t) X_i(\gamma(t))$$

then, by definition of orthonormality we have

$$l(\gamma) = \int_0^T \|\dot{\gamma}(t)\|_{\gamma(t)} dt = \int_0^T \|h(t)\| dt,$$

where the norm of h *is a standard Euclidean norm.*

Remark 1.65. *Let $X_1, \dots, X_m$ be orthogonal vector fields w.r.t. the sub-Riemannian metric and let γ and η two admissible curves such that*

$$\dot{\gamma}(t) = \sum_{i=1}^{m} \alpha_i(t) X_i(\gamma(t)) \ , \quad \dot{\eta}(t) = \sum_{i=1}^{m} \beta_i(t) X_i(\eta(t)).$$

then we have

$$g(\dot{\gamma}(t), \dot{\eta}(t)) = < \alpha(t), \beta(t) >,$$

where $< \ , \ >$ is the standard Euclidean inner product.

Example 1.66. *On a Riemannian manifold all curves are horizontal curves.*

Example 1.67. *Let $\mathbb{H}^1$ be the Heisenberg group of dimension 1, we can construct a horizontal curve in this way: let $p = (x_1, y_1, z_1)$ and $q = (x_2, y_2, z_2)$ be two given points of $\mathbb{R}^3$. Let $\overline{\gamma}(t) = (x(t), y(t))$ be a plane joining (x_1, y_1) to (x_2, y_2). We assume $T = 1$, we can look only absolutely continuous curves with constant curvature, i.e. we can assume that*

$$\int_{\overline{\gamma}} x dy = \int_0^1 x(t)\dot{y}(t)dt = \frac{1}{2}\int_0^1 (x(t)\dot{y}(t) - y(t)\dot{x}(t))dt = C,$$

for some $C \in \mathbb{R}$. Then we can define a curve in $\mathbb{R}^3$, setting $\gamma(t) = (x(t), y(t), z(t))$ where the third coordinate is given by

$$z(t) = z_1 + \frac{1}{2}\int_0^t (x(s)\dot{y}(s) - y(s)\dot{x}(s))ds.$$

We observe that γ is absolutely continuous, $z(0) = z_1$, $z(1) = z_1 - C$ and choosing $C = z_1 - z_2$, then M joins p to q. Hence the curve will be in the form

$$\gamma(t) = (x(t), y(t), z(t)) = \left(x(t), y(t), z_1 + \frac{1}{2}\int_0^t x(s)\dot{y}(s) - y(s)\dot{x}(s)ds\right)$$

and so, differentiating the curve γ we obtain

$$\dot{\gamma}(t) = \left(\dot{x}(t), \dot{y}(t), \frac{1}{2}(x(s)\dot{y}(s) - y(s)\dot{x}(s))\right) = X_1(\gamma(t)) + X_2(\gamma(t)).$$

Hence, by definition, the curve is horizontal.

Counterexample 1.68. *We observe that* $\gamma : [0,T] \to \mathbb{H}^1$ *s.t.* $\gamma(t) = (t,t,t)$ *is not an horizontal curve, in fact we obtain that* $X_1(\gamma(t)) = (1,0,-\frac{t}{2})$ *and* $X_2(\gamma(t)) = (0,1,\frac{t}{2})$ *and then we have to verify the definition of horizontal. Then, considering* $a(t), b(t) \in L^1([0,T])$ *we obtain*

$$\dot{\gamma}(t) = (1,1,1) = \left(a(t), b(t), (b(t) - a(t))\frac{t}{2}\right).$$

One sees immediately that such $a(t)$, $b(t)$ *cannot exist.*

Definition 1.69. *We call* sub-Riemannian distance *(or* Carnot-Carathéodory distance*) the function* $d : M \times M \to [0, +\infty]$ *defined by*

$$d(p,q) = \inf\{l(\gamma) \mid \gamma \text{ horizontal curve joining } p \text{ to } q\}.$$

Definition 1.70. *A* (minimizing) geodesic *between two points* x *and* y *is any a.c. horizontal curve which minimizes the length.*

Definition 1.71. *Let* M *be a manifold and* X,Y *two vector fields defined on this manifold and* $f : M \to \mathbb{R}$ *a smooth function, then we define the bracket between* X *and* Y *as the vector field defined as*

$$[X,Y](f) = X(Y(f)) - Y(X(f)).$$

Let us consider $\mathcal{X} = \{X_1, \dots, X_m\}$ *spanning some distribution* $\mathcal{H} \subset TM$*, we define the k-bracket as*

$$\mathcal{L}^{(k)} = \{[X,Y] \mid X \in \mathcal{L}^{(k-1)} \quad Y \in \mathcal{L}^{(1)}\}$$

with $k \geq 2$, $i_j \in \{1, \dots, m\}$ *and* $\mathcal{L}^{(1)} = \mathcal{X}$*. The associated Lie algebra is the set of all brackets between the vector fields of the family*

$$\mathcal{L}(\mathcal{X}) := \{[X_i, X_j^{(k)}] \mid X_j^{(k)} \text{ } k\text{-length bracket of } X_1, \dots X_m \; k \in \mathbb{N}\} = \bigcup_{k \geq 1} \mathcal{L}^{(k)}.$$

The definition of Hörmander condition is crucial in order to work with PDEs in sub-Riemannian setting, because it is a (weak) regularity condition.

Definition 1.72 (Hörmander condition)**.** *Let M be a smooth manifold and $\mathcal{H}$ a distribution defined on M. We say that the distribution is* bracket generating *if and only if, at any point, the Lie algebra $\mathcal{L}(\mathcal{X})$ spans the whole tangent space of M at point p. We say that a sub-Riemannian geometry satisfies the* Hörmander condition *if and only if the associated distribution is bracket generating.*

Example 1.73. *Let us consider the Grušin plane as in the Example 1.62. We can compute $[X, Y](f)$ with f smooth function*

$$[X, Y](f) = \frac{\partial}{\partial x_1}\left(x_1 \frac{\partial}{\partial x_2} f\right) - x_1 \frac{\partial}{\partial x_2}\left(\frac{\partial}{\partial x_1} f\right) = \frac{\partial}{\partial x_2} f,$$

and so the commutator spans a new vector field: $Z(x_1, x_2) = (0, 1)^T$. X and Z generate $T_x\mathbb{R}^2 = \mathbb{R}^2$ for all $x \in \mathbb{R}^2$.

Counterexample 1.74. *Let $\mathbb{R}^2$ be the manifold and let the vector field $X(x_1, x_2) = (0, 1)^T$ be the distribution. Trivially it does not satisfy the ő rmander condition.*

We now generalize the definition of gradient, divergence, Laplacian and Hessian for the sub-Riemannian case.

Definition 1.75. *Let $u : M \to \mathbb{R}$, $v : M \to \mathbb{R}^m$ be a differentiable function and $\mathcal{X} = \{X_1, \ldots, X_m\}$ a collection of vector fields on M satisfying the Hörmander condition which spans a distribution $\mathcal{H}$*

- *We define the* horizontal gradient *as*

$$\nabla_0 u = (X_1 u, \ldots, X_m u)^T.$$

- *We define the* horizontal divergence *as*

$$div_{\mathcal{X}} v(x) = (X_1 v_1) + \cdots + (X_m v_m).$$

- *We define the* horizontal Laplacian *as*

$$\Delta_0 u = X_1(X_1 u) + \cdots + X_m(X_m u).$$

- *We define the* horizontal Hessian *as the matrix*

$$\mathcal{X}^2 u = (X_i X_j u)_{i,j=1,\dots,m}.$$

Definition 1.76. *Let $\mathcal{X} = \{X_1, \dots, X_m\}$ be a family of vector fields defined on a smooth manifold M and $\mathcal{H}$ the distribution generated by $X_1, \dots, X_m$. Given $p \in M$ we call* step *of the distribution $\mathcal{H}$ at the point p, and we indicate by $k(p)$, the smallest natural number such that*

$$\bigcup_{i=1}^{k(p)} \mathcal{L}^i(p) = T_p M, \quad \forall p \in M,$$

where $\mathcal{L}^1 = span(\{Z = X \,|X \in \mathcal{H}\})$, $\mathcal{L}^i = span(\{Z = [X,Y] \mid X \in \mathcal{H}$ and $Y \in \mathcal{L}^{i-1}\})$.

Example 1.77. *We call the* 1-dimensional exponential Heisenberg group *or canonical Heisenberg group as in the Example 1.58 endowed with the metric g such that $X(x_1, x_2, x_3), Y(x_1, x_2, x_3)$ are orthonormal between each other. Computing as the previous example we can observe that $Z(x_1, x_2, x_3) = [X, Y](x_1, x_2, x_3) = (0, 0, 1)^T$ and so we can write explicitly the subbundle as*

$$\mathcal{H} = \{((x_1, x_2, x_3), (v_1, v_2, v_3)^T) \in \mathbb{R}^3 \times T_p\mathbb{R}^3 |$$
$$v = \alpha X(x) + \beta Y(x) = \big(\alpha, \beta, -\frac{x_2}{2}\alpha + \beta\frac{x_1}{2}\big)^T\},$$

with $\alpha, \beta \in \mathbb{R}$ or, by a Cartesian representation

$$v_3 = -\frac{x_2}{2}v_1 + \frac{x_1}{2}v_2.$$

We can compute the gradient

$$\nabla_0 f = (Xf, Yf)^T = \left(f_x - \frac{x_2}{2}f_{x_3}, f_{x_2} + \frac{x_1}{2}f_{x_3}\right)^T$$

and the horizontal Laplacian is given by

$$\Delta_0 f = f_{x_1x_1} - x_2 f_{x_1x_2} + x_1 f_{x_3x_2} + f_{x_2x_2}\left(\frac{x_1^2}{4} - \frac{x_2^2}{4}\right) f_{x_3x_3}.$$

Example 1.78. *Let $\mathbb{R}^2 \times S^1$, we consider the vector fields: $X_1(x_1, x_2, \theta) = (cos\theta, sin\theta, 0)^T$, $X_2(x_1, x_2, \theta) = (0, 0, 1)^T$ with the metric which induces the standard Euclidean metric on $\mathbb{R}^2 \times S^1$. This vector fields generates the* roto-traslation group *which is closely related to Citti-Sarti visual cortex model (see for further details [13, 14, 16]). We can observe that the distribution has step equal to 2, in fact we have that $X_3(x_1, x_2, \theta) = [X_1, X_2] = (-\sin\theta, \cos\theta, 0)^T$. The subbundle is given by*

$$\mathcal{H} = \{((x_1, x_2, \theta), (v_1, v_2, v_3)^T) \in (\mathbb{R}^2 \times S^1) \times T_p(\mathbb{R}^2 \times S^1)\},$$

with

$$v = \alpha X_1 + \beta X_2 = (\alpha \sin\theta, \alpha \cos\theta, \beta)^T,$$

or by a Cartesian representation

$$v_1 = \tan\theta v_2,$$

and the gradient is given by

$$\nabla_0 f = (X_1 f, X_2 f)^T = (\cos\theta f_{x_1} + \sin\theta f_{x_2}, f_\theta)^T,$$

and the horizontal Laplacian is

$$\Delta_0 f = \cos^2\theta f_{x_1x_1} + 2\sin\theta\cos\theta f_{x_1x_2} + \sin^2\theta f_{x_2x_2} + f_{\theta\theta}.$$

Let us conclude this subsection by recalling a weak regularity condition related to sub-Riemannian geometry.

Theorem 1.79. *([12])[Chow] Let M be a smooth manifold and $\mathcal{H}$ a bracket generating distribution (see Definition 1.72) defined on M. If M is connected, then there exists a horizontal curve joining any two given points of M.*

1.3 Geodesics in different geometries

We wonder now how to find geodesics in different spaces, this question is important to understand better the space in which we work; we will start from the Euclidean case in order to arrive to the sub-Riemannian one.

Euclidean spaces

This case is well-known. In Euclidean spaces geodesics exist, they are unique and they are smooth (they are the segment of straight line which connects p and q).

Riemannian spaces

In this case the situation start to be different because, in some cases, we start to lose the uniqueness of geodesics (in some cases can be even infinite). We state now a theorem of local existence for the geodesics.

Theorem 1.80. *([24], Theorem 1.17) Let M be a Riemannian manifold and $p_0 \in M$, for any $\varepsilon > 0$, there exists a neighbourhood U of p_0 such that for any $p \in U$ there exists a unique (minimizing) geodesic, joining p_0 to p with length less or equal to ε. Moreover, if the Riemannian manifold M is complete (i.e. every maximal geodesic is defined on $\mathbb{R}$), then there exists at least a geodesic joining any pair of points.*

Remark 1.81. *This is not a (global) uniqueness theorem, for example we can take S^1 and we can observe that there are at least two geodesics which connects north pole and south pole.*

Sub-Riemannian space

The sub-Riemannian space is the case where we "lose" many curves and so the result depends on the distribution. We can state the following theorem.

Theorem 1.82. *([42], Theorem 1.6.3 and 1.6.4) Let M be a smooth manifold and $\mathcal{H}$ a bracket generating distribution. Then:*

- *local existence: for any $p \in M$ there exists a neighbourhood U of p such that, for any $q \in U$, there exists a geodesic joining p and q.*

- *global existence: if moreover M is connected and complete w.r.t. the sub-Riemannian metric induced by $\mathcal{H}$, for any pair of points $p, q \in M$ there exists a geodesic joining p and q.*

Unfortunately, uniqueness does not hold in this spaces but there are some results on regularity in [43].

Theorem 1.83. *([24], Theorem 1.115) The geodesics of the canonical Heisenberg group (as stated in Example 1.58) starting from the origin can be parametrized by*

$$\begin{cases} \gamma_1(t) = \frac{a}{c}\sin(ct) - \frac{b}{c}(1-\cos(ct)), \\ \gamma_2(t) = \frac{b}{c}\sin(ct) + \frac{a}{c}(1-\cos(ct)), \\ \gamma_3(t) = \frac{a^2+b^2}{2c^2}(ct - \sin(ct)) \end{cases}$$

if $c \neq 0$, and by

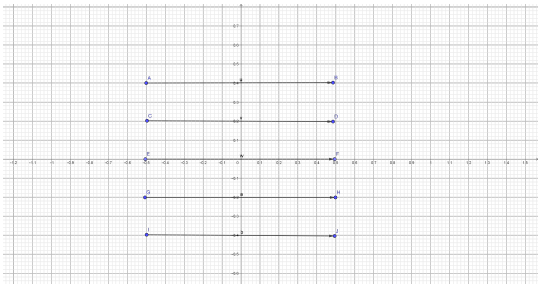
$$\begin{cases} \gamma_1(t) = at, \\ \gamma_2(t) = bt, \\ \gamma_3(t) = 0 \end{cases}$$

if $c = 0$ for any $a, b \in \mathbb{R}$.

1.4 Geometries without Hörmander condition

We consider Hörmander as a (weak) regularity condition which allows us to obtain T_pM in every single point of our manifold by means of commutators. We can consider now some examples in which Chow's Theorem does not hold due to the lack of the Hörmander condition.

Example 1.84. *We consider $\mathbb{R}^2$ with the vector field $X(x, y) = (1, 0)^T$. We can sketch this geometric object in this way*



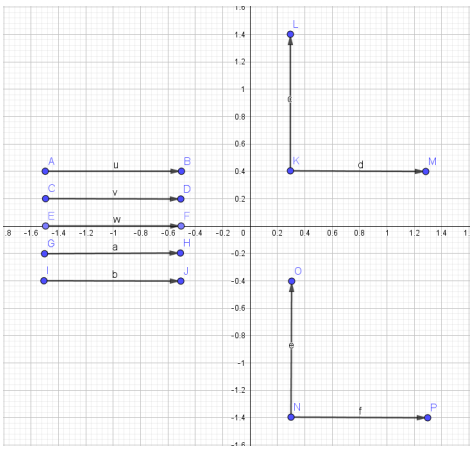
We observe that the Chow's Theorem in this case does not hold, in fact if we consider, for instance, $x = (0, 1)$ *and* $y = (3, 4)$ *we see immediately that these does not exist any geodesics which connects these two points. For this reason we have* $d(x, y) = +\infty$.

However, the fact that the Hörmander condition does not hold does not imply that distance between two points is infinity, the following example is important for this reason.

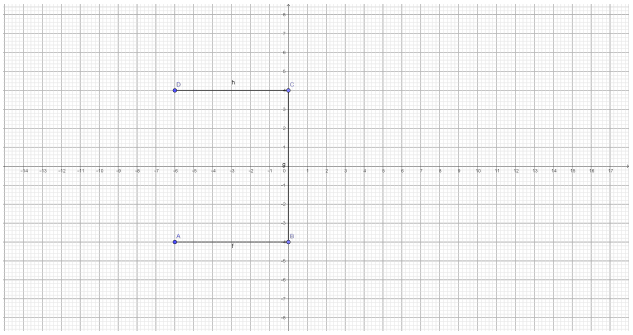
Example 1.85. *Let* $\mathbb{R}^2$ *be with the vector fields* $X(x, y) = (1, 0)^T$ *and* $Y(x, y) = (0, a(x))^T$ *where*

$$a(x) = \begin{cases} 1, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0 \end{cases}$$

and so we have



and as consequence we obtain a particular situation: the left part of plane has not a "regular" geometry while the other part is locally like a Riemannian manifold. We observe that there exists always a geodesics between two points



and the distance is described by the function

$$d((x,y),(x',y')) = \begin{cases} \sqrt{|x-y|^2+|x'-y'|^2}, & x \geq 0, \quad x' \geq 0, \\ |x|+|x'|+|y-y'|, & x < 0, \quad x' < 0, \\ |x|+\sqrt{|x'|^2+|y-y'|^2}, & x < 0, \quad x' \geq 0, \\ |x'|+\sqrt{|x|^2+|y-y'|^2}, & x \geq 0, \quad x < 0. \end{cases}$$

However, this distance is not continuous w.r.t. the Euclidean metric (i.e $\lim_{(x,y)\to(x',y')} d(x,y) \neq 0$*), in fact if we consider two points with* $x' < 0$ *and* $x < 0$ *we have*

$$\lim_{(x,y)\to(x',y')} |x|+|x'|+|y-y'| = 2|x| \neq 0.$$

The following Proposition is crucial to show that the distance d is continuous.

Proposition 1.86. *([24]) Let* $d(x,y)$ *be a sub-Riemannian distance defined on a smooth manifold and satisfying the Hörmander condition with step* k*. Then, for any compact* $K \subset M$*, there exist two constants* $C_1 = C_1(K) > 0$ *and* $C_2 = C_2(K) > 0$ *such that*

$$C_1|x-y| \leq d(x,y) \leq C_2|x-y|^{\frac{1}{k}},$$

for any $x, y \in K$*.*

Remark 1.87. *Then it follows from elementary theorems on limits that* $\lim_{x\to y} d(x,y) = 0$ *w.r.t the Euclidean metric.*

Lie groups and Lie algebras

The idea of *Lie group* is important because it connects geometry with algebra and so we can work with manifolds with algebraic properties (for further informations see [1], [29], [37]).

Definition 1.88. *Let* $\mathbb{G}$ *be a smooth manifold.* $\mathbb{G}$ *is a* Lie group *if* $\mathbb{G}$ *is endowed by a group structure and the following functions*

$$* : \mathbb{G} \times \mathbb{G} \to \mathbb{G}, \qquad (x,y) \to x * y,$$

$$^{-1} : \mathbb{G} \to \mathbb{G}, \qquad x \to x^{-1}$$

are smooth.

Example 1.89. *The group of* $SO(2)$ *of matrices (with the operation of multiplication as composition)*

$$SO(2,\mathbb{R}) = \left\{ \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} : \phi \in [0, 2\pi) \right\}$$

is a Lie group.

We recall now shortly the definition of Lie algebra.

Definition 1.90. *A* Lie algebra *is a vector space* g *over a field* $\mathbb{F}$ *(generally* $\mathbb{R}$*) with and operation* $[\ ,\ ] : g \times g \to g$ *called* Lie bracket *such that*

- *It is bilinear, i.e.* $[\alpha x + \beta y, z] = \alpha[x,z] + \beta[y,z]$ *for all* $x, y, z \in g$ *and* $\alpha, \beta \in \mathbb{F}$,
- *It is skew symmetric, i.e.* $[x,y] = -[y,x]$ *for all* $x, y \in g$,
- *It holds Jacobi identity i.e.* $[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0$ *for all* $x, y, z \in g$.

Definition 1.91. *Let $a \in \mathbb{G}$, then we define the* left translation *as $\tau_a : \mathbb{G} \to \mathbb{G}$ $\tau_a(x) = a * x$. The* Lie algebra *g of $\mathbb{G}$ is the set of the vector fields X, which are left invariant, i.e.*

$$(Xf)(\tau_a(x)) = X(f \circ \tau_a)(x)$$

for all $x, a \in \mathbb{G}$ and for all $f \in C^\infty(\mathbb{G})$.

Proposition 1.92. *([1]) The set g of all the left-invariant vector fields is a Lie algebra with $[\ ,\ ]$ the commutator for vector fields.*

Proposition 1.93. *([1]) If $X \in g$, then there exists a unique 1-parameter subgroup $\Phi_X(t)$, called vector flux, defined on $\mathbb{G}$, for all $t \in \mathbb{R}$ by the system*

$$\begin{cases} \frac{d}{dt}\Phi_X(t)|_{t=0} = X(\Phi_X(t)), \\ \Phi_X(0) = e. \end{cases}$$

Definition 1.94. *Let $\mathbb{G}$ be a Lie group , the exponential map is the smooth function given by*

$$exp : g \to \mathbb{G}, \quad X \to \Phi_X(1) \tag{1.3}$$

where $\Phi_X : (\mathbb{R}, +) \to (\mathbb{G}, *)$ is as in Proposition 1.93.

Example 1.95. *The circle group T is the multiplicative group of all complex numbers with absolute value 1, i.e*

$$T = \{z \in \mathbb{C} : |z| = 1\}.$$

It is not difficult to observe that the inverse and the composition are smooth function and, for this reason, T is a Lie group. The exponential map is given by

$$\theta \to z = e^{i\theta} = cos\theta + isin\theta$$

in fact if $\theta = 0$ we have $e = 1$, if we derive we have the vector field of multiplicative group.

Definition 1.96. *A group $\mathbb{G}$ is* nilpotent *if and only if its central series defined as*

$$\begin{cases} \mathbb{G}^{(1)} = \mathbb{G}, \\ \mathbb{G}^{(i+1)} = [\mathbb{G}, \mathbb{G}^{(i)}] = \{ghg^{-1}h^{-1} \mid g \in \mathbb{G} \quad h \in \mathbb{G}^{(i)}\}, \quad i \geq 1 \end{cases}$$

is finite, i.e. that there exists a $k \in \mathbb{N}$ such that

$$\mathbb{G}^{(k+1)} = \{e\} \neq \mathbb{G}^{(k)}.$$

Example 1.97. *An abelian group is nilpotent with $k = 1$.*

Example 1.98. *The direct product of two nilpotent groups is nilpotent.*

Counterexample 1.99. *The symmetric group $\mathcal{S}_3$ is not nilpotent.*

Definition 1.100. *A Lie algebra g is* nilpotent *with step equal to $k \in \mathbb{N}$ if and only if, setting*

$$\begin{cases} g^{(1)} = g, \\ g^{(i+1)} = [g, g^{(i)}] = \{[X, Y] \mid X \in g, Y \in g^{(i)}\}, \quad i \geq 1 \end{cases}$$

there exists $k \in \mathbb{N}$ such that $g^{(k+1)} = \{0\} \neq g^{(k)}$.

Example 1.101. *The Lie algebra defined by $(\mathbb{R}^N, +)$ is abelian and so it is nilpotent.*

Example 1.102. *The vector fields of Heisenberg group are a nilpotent algebra, in fact we have $g^{(1)} = \{X, Y\}$ such that $[X, Y] = Z$, $g^{(2)} = \{Z\}$ and $g^{(3)} = \{0\}$.*

Counterexample 1.103. *Let the rototranslation algebra be given by the following $g = \{U_1, U_2, U_3\}$ defined by*

$$[U_1, U_2] = U_3, \quad [U_2, U_3] = U_1, \quad [U_1, U_3] = 0.$$

It is possible to observe that $[g, g] = \{U_1, U_3\}$, $[g, [g, g]] = \{U_1, U_3\}$ and so this Lie algebra is not nilpotent.

Remark 1.104. *Given a Lie group $\mathbb{G}$, the associated Lie algebra is nilpotent with step equal to k if and only if $\mathbb{G}$ is so.*

Remark 1.105. *A Lie group is nilpotent with step k if and only if the associated Lie algebra has a bracket generating sub-Riemannian structure with step k.*

1.5 Carnot group

Definition 1.106. *A* Carnot group *is a Lie group, nilpotent and simply connected, whose Lie algebra g admits a stratification i.e. there exist $V_1, \dots, V_k$ vector spaces such that*

$$g = V_1 \oplus \cdots \oplus V_k.$$

The value $k \in \mathbb{N}$ is called step. Let us recall that with $\oplus$ we indicate the direct sum *of two vector spaces i.e. called A,B two vector spaces the direct sum is given by $A + B = \{a + b \mid a \in A, b \in B\}$ such that $A \cap B = \{0\}$.*

Example 1.107. *The Heisenberg group is a Carnot group which can be written as*

$$\mathcal{H} = V_1 \oplus V_2$$

where $V_1 = span\{X_1, X_2\}$, $V_2 = span\{X_3\}$ where X_1, X_2 and X_3 are defined as in Example 1.58.

Counterexample 1.108. *The Grušin plane (see Example 1.62) is not any Carnot group because there is not a structure of group associated to it.*

Carnot groups have good properties due to the stratification of its algebra. This becomes clear from the definition of dilatation.

Definition 1.109. *Let us consider a Carnot group $\mathbb{G}$ and its g associated Lie algebra. For $\lambda > 0$ we define a family of* dilatations *on g as a family of smooth*

maps $\delta_\lambda : g \to g$ *defined as*

$$\delta_\lambda(X) = \lambda^i X, \quad \textit{if} \quad X \in V_i.$$

Example 1.110. *In* $\mathbb{H}^1$ *the dilatation is*

$$(X_1, X_2, X_3) \to (\lambda X_1, \lambda X_2, \lambda^2 X_3)$$

where $\lambda > 0$ *and* X_1, X_2 *and* X_3 *are defined as in Example 1.58.*

Definition 1.111. *Let us consider a Carnot group* $\mathbb{G}$ *and its Lie algebra* g*. We define the dilations for the Carnot group* $\mathbb{G}$*, by setting* $\overline{\delta} = \exp^{-1} \delta \exp$*, where* δ *is as in Definition 1.109. For sake of simplicity, we will write* δ *instead of* $\overline{\delta}$*.*

Remark 1.112. *Since the Carnot-Carathéodory distance is defined minimizing the length functional over all the horizontal curves it follows immediately that*

$$d_C(\delta_\lambda(x), \delta_\lambda(y)) = \lambda d_C(x, y)$$

for any $x, y \in \mathbb{G}$ *and* $\lambda > 0$*.*

Definition 1.113. *Let us consider* $\mathbb{G}$ *a Carnot group and* d_C *the associated Carnot-Carathéodory distance. Hence the norm associated to the distance is given by*

$$\|x\|_C = d_C(0, x).$$

Proposition 1.114 ([43], Chapter 1, Section 7)**.** *Let us consider* $\mathbb{G}$ *a Carnot group,* g *its Lie algebra and* δ *the related dilatation. Then it holds true*

- $\delta_{\mu\lambda}(x) = \delta_\mu(\delta_\lambda(x))$ *for* μ, $\lambda > 0$ *and* $x \in \mathbb{G}$,
- $\delta_\lambda(xy) = \delta_\lambda(x)\delta_\lambda(y)$ *for* $\lambda > 0$ *and* $x, y \in \mathbb{G}$.

1.6 Riemannian approximation

Let $\mathcal{H} = \{X_1, \ldots, X_m\}$ be a distribution of left-invariant vector fields on $\mathbb{R}^N$. It is possible to extend this orthonormal frame to a new one by completing it with $N-m$ vector fields of the form $\varepsilon X_{m+1}, \ldots \varepsilon X_N$ with $\varepsilon > 0$ such that the vector fields $X_1(x), \ldots, X_m(x), X_{m+1}(x), \ldots, X_N(x)$ are linearly independent among each other for all $x \in \mathbb{R}^N$. This is called *Riemannian approximation* because it allows us to work in a Riemannian space rather than a sub-Riemannian one. We define $\mathcal{H}^\varepsilon = \{X_1, \ldots, X_m, X^\varepsilon_{m+1} = \varepsilon X_{m+1}, \ldots, X^\varepsilon_N = \varepsilon X_N\}$ as the *Riemannian approximation* of $\mathcal{H}$. It is possible to define on this manifold a Riemannian metric g_ε and, consequently, a distance d_ε. We define g_ε in such a way that the vector fields $X^\varepsilon_1, \ldots X^\varepsilon_N$ are orthonormal, i.e.

$$g_\varepsilon(v, w) = \sum_{i=1}^{N} \alpha^i \beta^i,$$

where $v = \sum_{i=1}^N \alpha^i X^\varepsilon_i$ and $w = \sum_{i=1}^N \beta^i X^\varepsilon_i$. To develop better some computations we define the following two matrices.

Definition 1.115. *We write*

$$\sigma(x) = [X_1(x), \ldots, X_m(x)]^T \tag{1.4}$$

for the $m \times N$ matrix given by the vector fields of the sub-Riemannian geometry and

$$\sigma_\varepsilon(x) = [X_1(x), \ldots X_m(x), \varepsilon X_{m+1}(x) \ldots, \varepsilon X_N(x)]^T, \tag{1.5}$$

for its Riemannian approximation. Furthermore, we observe that $\det(\sigma_\varepsilon(x)) \neq 0$ *since $X_1(x), \ldots X_N(x)$ are linearly independent among each other for all $x \in \mathbb{R}^N$.*

Example 1.116 (Riemannian approximation of $\mathbb{H}^1$)**.** *Let $x = (x_1, x_2, x_3) \in \mathbb{R}^3$. We know that the one dimensional Heisenberg group is given by the manifold $\mathbb{R}^3$ and the vector fields given by the Example 1.58. The matrix $\sigma(x)$ is given*

by

$$\sigma(x) = \begin{bmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \end{bmatrix}. \tag{1.6}$$

We observe that the Riemannian approximation is given by X_1, X_2 *and* $X_3^\varepsilon = \varepsilon X_3 = (0,0,\varepsilon)^T$ *where* $X_3 = [X_1, X_2] = (0,0,1)^T$ *and so* $\sigma_\varepsilon(x)$ *is given by*

$$\sigma_\varepsilon(x) = \begin{bmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \\ 0 & 0 & \varepsilon \end{bmatrix}.$$

It is possible to see that

$$(\mathbb{G}, d_\varepsilon) \to (\mathbb{G}, d_c) \quad \text{as} \quad \varepsilon \to 0,$$

in the Gromov-Hausdorff sense (it will be explained in detail later in this chapter), where $\mathbb{G}$ is a Carnot group (see [8] and [36]). We are interested in the convergence of the distance of the approximated Riemannian geometry and the Heisenberg group. The problem of convergence is discussed in [43] and it is very technical, so we will focus on the geodesics in the Heisenberg group and its approximation with a similar approach of [24]. We recall the following theorem.

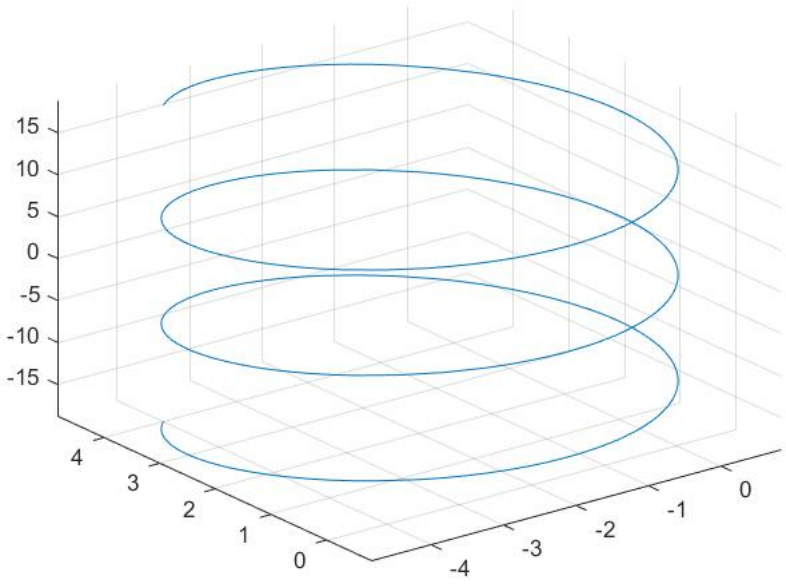
Theorem 1.117. *([24], Theorem 1.115) The geodesics of the canonical Heisenberg group starting from the origin can be parametrized as*

$$\begin{cases} \gamma_1(t) = \frac{a}{c}\sin(ct) - \frac{b}{c}(1-\cos(ct)), \\ \gamma_2(t) = \frac{b}{c}\sin(ct) + \frac{a}{c}(1-\cos(ct)), \\ \gamma_3(t) = \frac{a^2+b^2}{2c^2}(ct - \sin(ct)) \end{cases}$$

if $c \neq 0$ *and by*

$$\begin{cases} \gamma_1(t) = at, \\ \gamma_2(t) = bt, \\ \gamma_3(t) = 0 \end{cases}$$

otherwise.



The curve of Theorem 1.117 for $a = 2$, $b = 2$, $c = 1$

We wonder, using an approach based on the Hamiltonian and Hamilton's equations, if we can prove the convergence of $d_\varepsilon \to d$ locally uniformly. To prove it we have to find the explicit form of the geodesics for the approximated Heisenberg group.

Theorem 1.118. *The geodesics of the approximated Heisenberg group starting from the origin can be written as*

$$\begin{cases} \gamma_1^\varepsilon(t) = \frac{a}{c}\sin(ct) - \frac{b}{c}(1-\cos(ct)), \\ \gamma_2^\varepsilon(t) = \frac{b}{c}\sin(ct) + \frac{a}{c}(1-\cos(ct)), \\ \gamma_3^\varepsilon(t) = \frac{a^2+b^2}{2c^2}(ct - \sin(ct)) + \frac{c}{4}\varepsilon^2 t, \end{cases}$$

if $c \neq 0$ *and*

$$\begin{cases} \gamma_1^\varepsilon(t) = at, \\ \gamma_2^\varepsilon(t) = bt, \\ \gamma_3^\varepsilon(t) = 0. \end{cases}$$

Proof. Firstly, we define the Hamiltonian in the approximated case as

$$H(x_1, x_2, x_3, p_1, p_2, p_3) = \frac{1}{2}\left(\left(p_1 - \frac{x_2}{2}p_3\right)^2 + \left(p_2 + \frac{x_1}{2}p_3\right)^2 + \varepsilon^2 p_3^2\right)$$

then, using the Hamilton's equations we have

$$\begin{cases} \dot{x}_1 = p_1 - \frac{x_2}{2}p_3, \\ \dot{x}_2 = p_2 + \frac{x_1}{2}p_3, \\ \dot{x}_3 = -\frac{x_1}{2}p_2 - \frac{x_2}{2}p_1 + \frac{1}{4}(x_1^2 + x_2^2 + \varepsilon^2)p_3, \end{cases} \tag{1.7}$$

and

$$\begin{cases} \dot{p_1} = -\frac{p_3}{2}\left(p_2 + \frac{x_1}{2}p_3\right), \\ \dot{p_2} = \frac{p_3}{2}\left(p_1 - \frac{x_2}{2}p_3\right), \\ \dot{p_3} = 0 \end{cases}$$

It is immediate to observe that $p_3 = c$ with $c \in \mathbb{R}$ and so

$$\begin{cases} \dot{p_1} = -\frac{c}{2}\left(p_2 + \frac{c}{2}x_1\right) = -\frac{c}{2}\dot{x}_2, \\ \dot{p_2} = \frac{c}{2}\left(p_1 - \frac{c}{2}x_2\right) = \frac{c}{2}\dot{x}_1. \end{cases} \tag{1.8}$$

differentiating (1.7) we obtain

$$\begin{cases} \ddot{x}_1 = \dot{p}_1 - \frac{\dot{x}_2}{2}c \\ \ddot{x}_2 = \dot{p}_2 + \frac{\dot{x}_1}{2}c \end{cases}$$

and using (1.8) we obtain the system

$$\begin{cases} \ddot{x}_1 = -\frac{\dot{x}_2}{2}c - \frac{\dot{x}_2}{2}c = -c\dot{x}_2, \\ \ddot{x}_2 = \frac{\dot{x}_1}{2}c + \frac{\dot{x}_1}{2}c = c\dot{x}_1, \end{cases}$$

hence

$$\begin{cases} \ddot{x}_1 + c\dot{x}_2 = 0, \\ \ddot{x}_2 - c\dot{x}_1 = 0. \end{cases}$$

Now we assume $c \neq 0$ and we write $m = \dot{x}$ and $n = \dot{y}$ obtaining in this way

$$\begin{cases} \dot{m} + cn = 0, \\ \dot{n} - cm = 0, \end{cases}$$

and differentiating the first equation

$$\ddot{m} + c^2 m = 0,$$

and so the solution is

$$m(t) = k_1 \sin(ct) + k_2 \cos(ct).$$

Thus

$$n(t) = -k_1 \cos(ct) + k_2 \sin(ct).$$

Then considering $\dot{x}(0) = a$ and $\dot{y}(0) = b$ we write

$$\begin{cases} \dot{x}_1(t) = -b \sin(ct) + a \cos(ct), \\ \dot{x}_2(t) = b \cos(ct) + a \sin(ct), \end{cases}$$

and, integrating we find

$$\begin{cases} x_1(t) = \frac{a}{c} \sin(ct) - \frac{b}{c}(1 - \cos(ct)), \\ x_2(t) = \frac{b}{c} \sin(ct) + \frac{a}{c}(1 - \cos(ct)), \end{cases}$$

which implies

$$\dot{x}_3(t) = \frac{a^2 + b^2}{2c}(1 - \cos(ct)) + \frac{c}{4}\varepsilon^2.$$

Hence we conclude integrating the previous formula

$$x_3(t) = \frac{a^2 + b^2}{2c^2}(ct - \sin(ct)) + \frac{c}{4}\varepsilon^2 t.$$

Meanwhile if $c = 0$ we have that $x_1(t) = at$, $x_2(t) = bt$, which gives

$$\dot{x}_3(t) = \frac{abt}{2} - \frac{bat}{2} = 0.$$

So the geodesics in this case have the form

$$\begin{cases} x_1(t) = at, \\ x_2(t) = bt, \\ x_3(t) = 0. \end{cases}$$

□

It is immediate to obtain the following corollary.

Corollary 1.119. *Let γ_ε be a geodesic in the approximated Heisenberg group starting from the origin and γ a geodesic of 1-dimensional Heisenberg group starting from the origin then we obtain*

$$\lim_{\varepsilon\to 0}\gamma_\varepsilon(t)=\gamma(t)$$

pointwise in t.

Proof. For $c=0$ is trivial the claim, for $c\neq 0$ we have to observe that

$$\lim_{\varepsilon\to 0}\gamma_3^\varepsilon(t)=\gamma_3(t).$$

and that $\gamma_1^\varepsilon=\gamma_1$ and $\gamma_2^\varepsilon=\gamma_2$. □

1.7 Notions of convergence between metric spaces

Now we need to estimate the difference between different metric spaces and their shapes, for this reason we have to introduce the notion of *Hausdorff distance.*

Definition 1.120. *Let (E,d) be a metric space and $E_1,E_2\subset E$, then the* Hausdorff distance *between E_1, E_2 is given by*

$$Haus_E(E_1,E_2)=\inf\{\varepsilon>0|E_1\subset(E_2)_\varepsilon,\quad E_2\subset(E_1)_\varepsilon\}$$

where $(E_i)_\varepsilon=\{z\in E|d(z,E_i)\leq\varepsilon\}$ and $d(z,E_i)=\inf_{x\in E_i}d(z,x)$ or, equivalently,

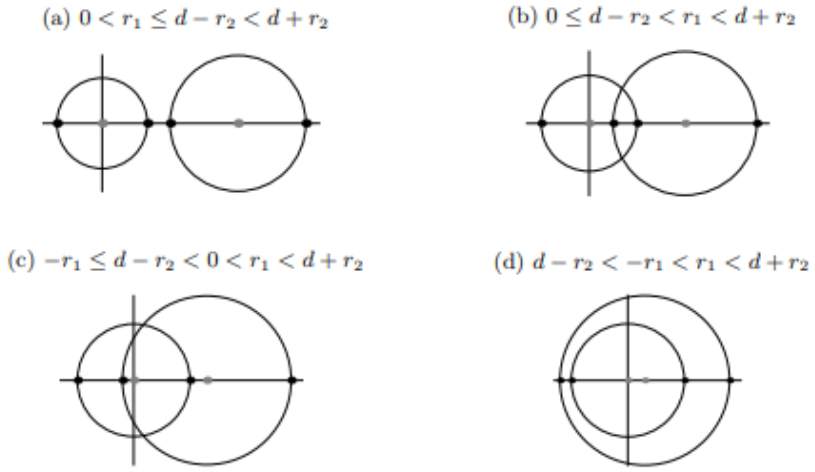
$$Haus_E(E_1,E_2)=\max\{\sup_{x\in E_1}\inf_{y\in E_2}d(x,y),\sup_{y\in E_2}\inf_{x\in E_1}d(x,y)\}.$$

The following propositions are some examples which are useful to understand how the Hausdorff distance works.

Proposition 1.121 ([40], Proposition 1, p.248)**.** *The Hausdorff distance between two circles $S_1, S_2 \subset \mathbb{R}^2$ is*

$$Haus_{\mathbb{R}^2}(S_1, S_2) = d(c_1, c_2) + |r_2 - r_1|,$$

where c_1, c_2 are the centers and r_1, r_2 the radius and d the standard Euclidean distance.



All the possible position between two circles (images taken from [40]).

Proposition 1.122 ([40], Proposition 2, p.249)**.** *Consider two hyperspheres $S_i^{N-1} = \{T \in \mathbb{R}^N, d(c_i, T) = r_i\} \subset \mathbb{R}^N$, $N \geq 3$, $i = 1, 2$. The Hausdorff distance between these two hyperspheres is*

$$Haus_{\mathbb{R}^N}(S_1^{N-1}, S_2^{N-1}) = d(c_1, c_2) + |r_2 - r_1|.$$

where d is the standard Euclidean distance.

Proposition 1.123 ([40], Proposition 3, p.251)**.** *The Hausdorff distance between the segment L and a circle S_1 in the plane is determinated by the following expression*

$$Haus_{\mathbb{R}^2}(L, S_1) = \max\{|d(t_1, c) - r|, |d(t_2, c) - r|, f(c, c', r, t_1, t_2)\}$$

where

$$f(c,c',r,t_1,t_2)=\begin{cases} d(c,c')+r, & \frac{\overline{t_1c}}{\|t_1t_2\|}\frac{\overline{t_1t_2}}{\|t_1t_2\|}\in[0,1],\\ \min\{d(c,t_1),d(c,t_2)\}+r, & otherwise,\end{cases}$$

with t_1, t_2 are the two endpoints of the segment L, c is the centre of the circle S_1 and c' is the orthogonal projection the circle on L.

Remark 1.124. *From the Propositions 1.122, 1.123 we deduce immediately that*

$$d(A,B)\neq Haus_Z(A,B).$$

where $d(A,B)=\inf\{d_E(x,y)|x\in A,y\in B\}$ and d_E is the standard Euclidean distance.

Now we have to introduce some technical definitions.

Definition 1.125. *Let $\phi : M\to N$ be a smooth map between (smooth) manifolds M and N, and suppose $f : N\to\mathbb{R}$ is a smooth function on N. Then the pullback of f by ϕ is the smooth function $\phi\circ f$ on M defined by $(\phi\circ f)(x)=f(\phi(x))$. To indicate a pullback we will use the symbol $f^*\phi$.*

Definition 1.126. *Let (M,g), (N,h) be two Riemannian manifolds. The* isometric embedding *is a smooth embedding $f : M\to N$ which preserves the metric in the sense that g is equal to the pullback of h by f, i.e. $g=f^*h$, i.e*

$$\forall v,w\in T_xM \text{ it holds } g(v,w)=h(df(v),df(w)).$$

We can introduce the definition of *Gromov-Hausdorff distance* which measure in a better way the difference of shapes between two sets using the Hausdorff distance and isometric embeddings.

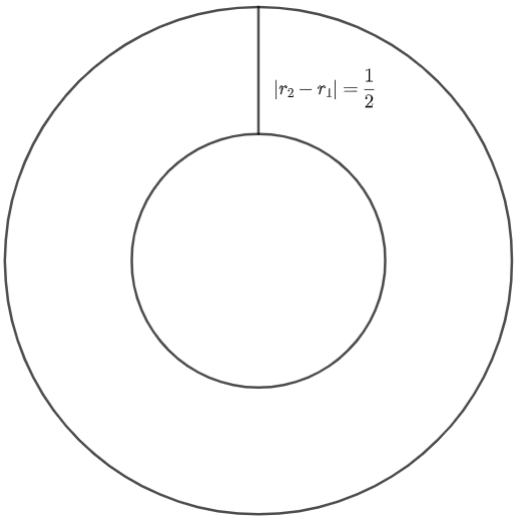
Definition 1.127. *Let (X,d_X), (Y,d_Y) be two metric spaces, then the* Gromov-Hausdorff distance *between X, Y is given by*

$$d_{GH}(X,Y)=\inf_{f,g,Z} Haus_Z(f(X),g(Y))$$

where the infimum is taken over all metric spaces Z and $f : X \to Z$, $g : Y \to Z$ are isometric embeddings.

Example 1.128. *Let (S^1, d_{S^1}) and $(\frac{1}{2}S^1, d_{\frac{1}{2}S^1})$ where d_{S^1} is the geodesics distance on S^1 and $\frac{1}{2}S^1$ rescaled of $\frac{1}{2}$ with $d_{\frac{1}{2}S^1}$ the associated geodesics distance. It is clear from the Proposition 1.121 that the infimum is obtained when one circle is concentric in the other ($d(c_1, c_2) = 0$) so, taking an isometric embedding which make coincide these two centres, we have*

$$d_{GH}\left(S^1, \frac{1}{2}S^1\right) = \frac{1}{2}.$$



Graphic representation of Example 1.128.

Using this distance we can define a notion of convergence between metric spaces.

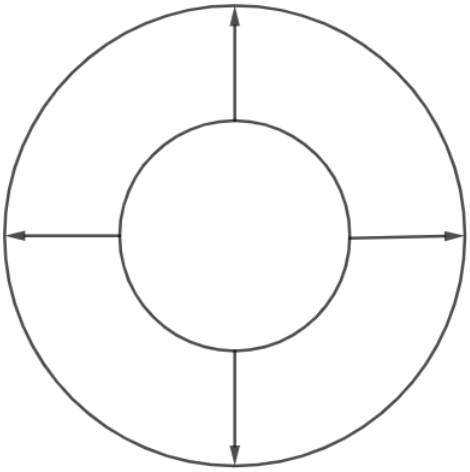
Definition 1.129. *A sequence of compact metric spaces $(X_n)_{n\in\mathbb{N}}$* Gromov-Hausdorff converges *to a compact metric space X if $d_{GH}(X_n, X) \to 0$ as $n \to \infty$.*

Example 1.130. *Following the Example 1.128 it is straightforward to show that $(\frac{n}{n+1}S^1, d_{\frac{n}{n+1}S^1}) \to^{G-H} (S^1, d_{S^1})$ as $n \to \infty$. Hence we have that the*

Gromov-Hausorff distance is

$$d_{GH}\left(S^1, \frac{n}{n+1}S^1\right) = \frac{1}{n+1},$$

which converges trivially to zero as $n \to \infty$.



Graphic representation of Example 1.130.

Unluckily the definition of Gromov-Hausdorff convergence works only with compact metric spaces, for this reason we have to introduce a further generalization.

Definition 1.131. *Let us consider the metric space* (X, d). *We define the ball of radius* $r > 0$ *and center* $x \in X$

$$B_X(x, r) = \{y \in X | d(x, y) < r\}.$$

Definition 1.132 (Pointed Gromov-Hausdorff convergence)**.** *A metric space is called* proper *if all the closed balls are compact. A sequence of pointed proper length spaces* (X_n, d_n, x_n) *converges in Gromov-Hausdorff to* (X, d, x) *if the sequence of the balls* $\overline{B}_{X_n}(x_n, r)$ *converges in Gromov-Hausdorff sense to* $\overline{B}_X(x, r)$.

Now we have to focus on the compact metric space; we wonder if it possible to write them without using the standard topological definition but in a simpler way. For this reason we have to introduce the following definitions.

Definition 1.133. *A metric space (X, d) is called* complete *if every Cauchy sequence is convergent.*

Definition 1.134. *A metric space (X, d) is* totally bounded *if and only if for every real number $\varepsilon > 0$ there exists a finite collection of open balls in X of radius ε whose union contains X.*

Example 1.135. *The definition of completeness and totally boundness are independent between each other, for example we can take $\mathbb{Q} \cap [0, 1]$ with the standard Euclidean metric: it is totally bounded but it is not complete (because $\mathbb{Q}$ is not complete). On the converse, $\mathbb{R}$ is complete with standard Euclidean distance but it is not totally bounded.*

The Example 1.135 suggests us that the compactness will be achieved if it possible to complete that interval (i.e. putting the the irrational numbers in the holes of the interval), for this reason we can state the following theorem.

Theorem 1.136. *Let (X, d) be a finite dimension metric space, then*

(X, d) is a compact space $\Leftrightarrow$ (X, d) is totally bounded and complete metric space.

Definition 1.137. *Let (X, d) be a metric space and $\gamma : [a, b] \to X$ a curve. We define the length of the curve*

$$L(\gamma) = \sup \sum_{i=1}^{m} d(\gamma(t_{i-1}), \gamma(t_i))dt.$$

where the supremum is taken over all $m \in \mathbb{N}$ and all sequences $t_0 \leq t_1 \leq \cdots \leq t_m \subset [a, b]$. If $L(\gamma) < \infty$ then the curve is rectifiable.

Definition 1.138. *Let (X, d) be a metric space, the* length metric associated with d *is the function $d' : X \times X \to [0, \infty]$ defined by*

$$d'(x, y) = \inf L(\gamma),$$

where γ is rectificable and $\gamma(0) = x$ and $\gamma(1) = y$. We call (X,d) a length space *if $d = d'$. Furthermore we say that (X, d) is a* proper length space *if it is a length space where all the closed balls are also compact.*

Definition 1.139. *Let d, d' be two different metrics on a metrizable space X, we say that they induce the same topology if and only of for any $x \in M$, $r > 0$ there exists positive numbers r_1, r_2 such that $B^{(d)}(x, r_2) \subset B^{(d')}(x, r)$ and $B^{(d')}(x, r_1) \subset B^{(d)}(x, r_2)$ where $B^{(d)}(x, r)$ is the open ball with centre x, radius r induced by the distance d.*

Example 1.140. *Let $\mathbb{R}^N$ be the standard set of the real and consider the following distances*

$$d(x, y) = \sqrt{|x_1 - y_1|^2 + \cdots + |x_n - y_n|^2},$$

$$d'(x, y) = |x_1 - y_1| + \cdots + |x_n - y_n|.$$

Applying the Definition 1.139 it is not hard to see that they generates the same topology.

The next proposition is crucial to obtain the convergence in the sense of Gromov-Hausdorff.

Proposition 1.141 ([9], Proposition 2.8)**.** *Let X be a set equipped with a family of metrics $(d_t)_{t>0}$ generating a common topology. For K compact in X, let us define*

$$\omega_K(\varepsilon) := \sup_{x,y \in K, \varepsilon \geq 0} \left(d_\varepsilon(x, y) - d_{t+\varepsilon}(x, y) \right).$$

Assume that

- *for each $t \geq 0$, (X, d_t) is a proper length space (see Definition 1.138).*

- *for fixed $x, y \in X$, the function $t \to d_t(x, y)$ is non increasing.*
- *for each compact set K in X, $\omega_K(\varepsilon) \to 0$ as $\varepsilon \to 0$.*

Then (X, d_t) converges in the sense of pointed Gromov-Hausdorff to (X, d_0) .

1.8 Heisenberg approximation and the convergence of the associated distances

In this section we will prove that the following theorem holds true.

Theorem 1.142. *Let $(\mathbb{H}^1, d)$ be the 1-dim Heisenberg group and $(\mathbb{R}^3, d_\varepsilon)$ its Riemannian approximation, then*

$$(\mathbb{R}^3, d_\varepsilon) \to^{\varepsilon \to 0} (\mathbb{H}^1, d)$$

in the sense of pointed Gromov-Hausdorff convergence.

The distance in this case can be written explicitly using Theorems 1.117 and 1.118 and so we have

$$d_\varepsilon(0,x) = \int_0^1 \sqrt{(-b\sin(ct)+a\cos(ct))^2+(b\cos(ct)+a\sin(ct))^2+\left(\frac{a^2+b^2}{2c}(1-\cos(ct))+\frac{c}{4}\varepsilon^2\right)^2}dt = \int_0^1 \sqrt{a^2+b^2+\left(\frac{a^2+b^2}{2c}\right)^2(1-\cos(ct))^2+\frac{a^2+b^2}{4}\varepsilon^2+\frac{c^2}{16}\varepsilon^4}dt,$$

and

$$d(0,x) = \int_0^1 \sqrt{(-b\sin(ct)+a\cos(ct))^2+(b\cos(ct)+a\sin(ct))^2+\left(\frac{a^2+b^2}{2c}(1-\cos(ct))\right)^2}dt = \int_0^1 \sqrt{a^2+b^2+\left(\frac{a^2+b^2}{2c}\right)^2(1-\cos(ct))^2}dt,$$

if $c \neq 0$ and

$$d_\varepsilon(0,x) = d(0,x) = \int_0^1 \sqrt{a^2+b^2}dt,$$

if $c = 0$, where $a, b, c \in \mathbb{R}$ are all constants. We observe that

Lemma 1.143. *([23], Lemma 2.3) Let γ be an horizontal curve with velocity $\alpha^{\gamma}(s)$ such that $\gamma(0) = x$ and $\gamma(1) = y$, then for all $z \in \mathbb{R}^N$ we have that $\overline{\gamma} = z \circ \gamma$ is an horizontal curve and $\alpha^{\overline{\gamma}}(s) = \alpha^{\gamma}(s)$ and $\overline{\gamma}(0) = z \circ x$ and $\overline{\gamma}(1) = z \circ y$.*

It is immediate to prove the following result.

Lemma 1.144. *Let us consider $\mathbb{G}$ a Carnot group and d the Carnot-Carathéodory distance defined on it. Hence it holds true that*

$$d(x, y) = d(0, x^{-1} \circ y) \quad \forall x, y \in \mathbb{G}.$$

We can easily check that the distance converges pointwise.

Theorem 1.145. *Let us consider $(\mathbb{H}, d)$ and its Riemannian approximation. It holds true*

$$d_{\varepsilon}(0, x) \to d(0, x)$$

pointwise as $\varepsilon \to 0$.

Proof. Lemma 1.144 allows us to reduce to the case of the curves starting from the origin.
If $c = 0$ the thesis is trivial. If $c \neq 0$, observing that the cosine is a bounded function and that we can suppose $\varepsilon \leq 1$. We have that

$$d_{\varepsilon}(0, x) \leq \int_0^1 \sqrt{a^2 + b^2 + \left(\frac{a^2 + b^2}{c}\right)^2 + \frac{a^2 + b^2}{4} + \frac{c^2}{16}} dt.$$

The right hand side is constant (hence $L^1([0, 1])$) and, as consequence, we can apply the Dominated Convergence Theorem and obtain the thesis. □

We conclude the section with the following result.

Theorem 1.146. *Let us consider $(\mathbb{H}^1, d)$ and its Riemannian approximation. It holds true*

$$d_{\varepsilon}(0, x) \to d(0, x),$$

locally uniformly as $\varepsilon \to 0$.

Proof. We can write, using the identity $(a^2 - b^2) = (a + b)(a - b)$ and that, using the inequality

$$\left(\int_0^1 f(x)dx \right)^2 \leq \int_0^1 f^2(x)dx,$$

then we have $|d_\varepsilon^2(0,x) - d^2(0,x)| \leq \int_0^1 \frac{a^2+b^2}{4}\varepsilon^2 + \frac{c^2}{16}\varepsilon^4 dt$ then we obtain

$$|d_\varepsilon(0,x) - d(0,x)| \leq \frac{\int_0^1 \frac{a^2+b^2}{4}\varepsilon^2 + \frac{c^2}{16}\varepsilon^4 dt}{|d_\varepsilon(0,x) + d(0,x)|}.$$

By this estimate above it is possible to observe that we obtain $|d_\varepsilon(0,x) + d(0,x)| \leq Mt$ and so we obtain that, in a compact set

$$\sup_{x \in K \subset \mathbb{R}^3} |d_\varepsilon(0,x) - d(0,x)| \leq \varepsilon^2 \frac{\frac{a^2+b^2}{4} + \frac{c^2}{16}\varepsilon^2}{M},$$

which gives immediately the uniform convergence. □

Chapter 2

Stochastic background: Brownian motion and SDE

In this thesis we will use some tools related to stochastic control theory and, for this reason, it is necessary to give a short introduction about some of the most used topics in stochastic analysis. In this chapter we will start from the axiomatic theory started by Kolmogorov at the beginning of XX century (for further details see [5] and [39]) and then we will continue introducing stochastic integrals and the definition of stochastic differential equation (SDE).

2.1 Basic notions of probability

Definition 2.1. *Let Ω be a set, a* σ-algebra *(which we indicate with $\mathcal{F}$) is a collection of subsets of Ω such that:*

- $\Omega \in \mathcal{F}$,
- *Let $A \subset \Omega$ be a subset such that $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$,*
- *Let $A_1, A_2 \cdots \in \mathcal{F}$ then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.*

We indicate the *power set of* Ω with the symbol $\mathcal{P}(\Omega)$.

Definition 2.2. *Let $\mathcal{S} \subset \mathcal{P}(\Omega)$ be a collection of subsets of Ω, we call the generated σ-algebra the smallest σ-algebra containing S, i.e.*

$$\mathcal{F}_{\mathcal{S}} = \cap\{\mathcal{F} | \mathcal{F} \ \sigma\text{-algebra s.t } \mathcal{S} \subset \mathcal{F}\}.$$

Definition 2.3. *Let Ω be a set and $\mathcal{F}$ a σ-algebra; a σ-additive measure* measure *is a function $\mu : \mathcal{F} \to [0, \infty)$ such that:*

- $\mu(\emptyset) = 0$,
- *If $\{E_n\}_{n\in I}$ is a collection of countable disjoint subsets, then, for $E := \cup_{n\in I} E_n$,*

$$\mu(E) = \mu\left(\sum_{n\in I} E_n\right).$$

Due to the fact that we are working on these spaces, we have to introduce functions which allow to maintain the measurability.

Definition 2.4. *Let us consider $\Omega \subset \mathbb{R}^N$, $\mathcal{F}$ a σ-algebra defined on Ω and μ a measure. A* measure space *is the triple $(\Omega, \mathcal{F}, \mu)$. Given $(X, \mathcal{F}_X, \mu_X)$ and $(Y, \mathcal{F}_Y, \mu_Y)$ two measure spaces, a function $f : X \to Y$ is called measurable if for all $A \in \mathcal{F}_Y$, $f^{-1}(A) \in \mathcal{F}_X$.*

Example 2.5. *Let $([0, T], \mathcal{F}, \mu)$ be a triple such that $\mathcal{F}$ is the collection of subsets generated by the countable union of open intervals and μ the measure defined as*

$$\mu([a, b]) = b - a$$

then $([0, T], \mathcal{F}, \mu)$ is the measure space with $\mu([0, T]) = T < \infty$. This is the Lebesgue measure on $[a, b]$.

The probability space and the random variable are basic definitions to work in probability theory and these concepts are related to measure theory.

Definition 2.6. *Let us consider $(\Omega, \mathcal{F})$ and $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$. We say that a function $f : \Omega \to \mathbb{R}^N$ is $\mathcal{F}|\mathcal{B}$-measurable if, for all $E \in \mathcal{B}(\mathbb{R}^N)$ it holds $f^{-1}(E) \in \mathcal{F}$.*

Definition 2.7. *A measure space $(\Omega, \mathcal{F}, \mu)$ such that $\mu(\Omega) = 1$ is called* probability space*. We denote the measure of this space with by symbol $\mathbb{P}$. Let $E \in \mathcal{F}$ be an* event*, then E happens* almost surely *if and only if $\mathbb{P}[E] = 1$.*

Definition 2.8. *Let $(\Omega, \mathcal{F}, \mu)$ be a probability space. A* random variable *is a function $\xi : \Omega \to \mathbb{R}^N$ such that is $\mathcal{F}$-measurable. We define the induced probability measure as*

$$\mu_\xi(B) = \mathbb{P}(\xi^{-1}(B))$$

where $B \subset \mathbb{R}^N$ is a Borel measurable set. The measure with respect to this random variable is called distribution *of ξ.*

Example 2.9. *The simplest example to introduce a random variable is* coin toss*: in fact there are two sides for a coin which we will call 0 and 1, so the random variable ξ can assume only these two values. We suppose that the game is fair, so the σ-algebra will be given by $\mathcal{P}(\{0,1\}) = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$ the probability measure will be given by*

$$\mathbb{P}(\xi = 1) = \mathbb{P}(\xi = 0) = \frac{1}{2}.$$

Example 2.10 (Normal distribution)**.** *Let $\sigma, \mu > 0$ be real numbers, we define the normal distribution on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ by the Lebesgue density*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

i.e. $\mu(A) = \int_A f d\mu$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. It is possible to check that all the properties of measures are satisfied and, in particular

$$\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1.$$

We introduce now two standard notions in probability theory.

Definition 2.11. *Let us consider $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space and $\xi : \Omega \to \mathbb{R}^N$ a random variable. If $\int_\Omega |\xi(\omega)| d\mathbb{P}(\omega) < \infty$, where $\omega \in \Omega$, then we define the* expectation *as*

$$\mathbb{E}[\xi] := \int_\Omega \xi(\omega) d\mathbb{P}(\omega) = \int_{\mathbb{R}^N} x d\mu_\xi(x).$$

Definition 2.12. *Let us consider* $(\Omega, \mathcal{F}, \mathbb{P})$ *a probability space and* $\xi : \Omega \to \mathbb{R}^N$ *a random variable. If* $\int_\Omega |\xi(\omega)|^2 d\mathbb{P}(\omega) < \infty$ *where* $\omega \in \Omega$ *we define the* variance *of the random variable* X *as*

$$Var(\xi) := \mathbb{E}[\xi - \mathbb{E}[\xi]]^2.$$

Definition 2.13. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space. Two events* $A, B \in \mathcal{F}$ *are called* independent *if and only if*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \quad \forall A, B \in \mathcal{F}.$$

A collection $\mathcal{A} = \{\mathcal{A}_i | i \in I\}$ *of families* $\mathcal{A}_i$ *of measurable sets is independent if and only if*

$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \ldots \mathbb{P}(A_{i_k}),$$

for all choices $A_{i_1} \in \mathcal{A}_1, \ldots, \ A_{i_k} \in \mathcal{A}_k$.
A collection of random variables $\{\xi_i : i \in I\}$ *is independent if the collection generated by the* σ*-algebras* $\mathcal{F}_{\xi_i}$ *is independent.*

We conclude this brief section stating some standard definition about the convergence of random variables, which are necessary to define Brownian motion and the Itô integral.

Definition 2.14. *Let us consider a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and let* $\{\xi_n\}_{n\in\mathbb{N}}$ *be a sequence of random variables.*

- *The sequence* $\{\xi_n\}_{n\in\mathbb{N}}$ *converges* in L^2 *if and only if*

$$\lim_{n\to\infty} \mathbb{E}[|\xi_n - \xi|^2] = 0.$$

- *The sequence* $\{\xi_n\}_{n\in\mathbb{N}}$ *converges* in probability *if and only if for all* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}(|\xi_n - \xi| > \varepsilon) = 0.$$

- *The sequence* $\{\xi_n\}_{n\in\mathbb{N}}$ *converges* almost surely *if and only if*

$$\mathbb{P}\left(\lim_{n\to\infty}\xi_n = \xi\right) = 1.$$

Definition 2.15. *A* (continuous-time) stochastic process *is a parametrized collection of random variables*

$$\{\xi(t)\}_{t\geq 0}$$

defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ *and assuming values on* $\mathbb{R}^N$.

Notation: We will write $\xi(t) = \xi(t,\omega)$ if ω is fixed.

2.2 Brownian Motion

We introduce now the idea of a Brownian motion. This stochastic process was observed for the first time by the biologist Robert Brown during his experiments and then it was studied in details by Albert Einstein in 1905 and Norbert Wiener. The study of this stochastic process is crucial in many applications in applied science and in pure mathematics.

Definition 2.16. *Let us consider a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. *A real valued stochastic process* $\{B(t)\}_{t\geq 0}$ *is called* Brownian motion *starting at* $x \in \mathbb{R}$ *if the following properties hold (we recall that, for sake of simplicity, we will consider* $B(t) = B(t,\omega)$*):*

- $B(0) = x$ *a.s.*
- *the process has independent increments i.e. for all time* $0 \leq t_1 \leq \cdots \leq t_n$ *the increments* $B(t_n) - B(t_{n-1})$, $\ldots$, $B(t_2) - B(t_1)$ *are independent random variables (as in Definition 2.13).*
- *Given* $0 < s < t$, $B(t) - B(s)$ *has a normal distribution with expectation zero and variance* $\sqrt{t-s}$.

$\{B(t)\}_{t\geq 0}$ *is called* standard Brownian motion *if* $B(0) = 0$.

Remark 2.17. *It is possible to prove that the function* $t \to B(t)$ *is continuous almost surely due to the properties of the Brownian Motion and Kolmogorov continuity theorem (see [51], p.51 for further remarks).*

Theorem 2.18 ([41], Theorem 1.4, p.23)**.** *There exists a standard Brownian motion in* $\mathbb{R}^N$.

Sketch of the proof. The idea is to construct the right joint distribution of Brownian motion step by step on the finite sets of the dyadic numbers, i.e.

$$\mathcal{D}_n = \left\{ \frac{k}{2^n} : 0 \leq k \leq 2^n \right\}.$$

We interpolate linearly the values on $\mathcal{D}_n$ and check that the uniform limit of these continuous functions exists and is a Brownian motion. To do this let us define $\mathcal{D} = \cup_{n=0}^{\infty} \mathcal{D}_n$ and let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space on which a collection $\{\xi(t) : t \in \mathcal{D}\}$ of independent, standard normally distributed random variables can be defined. Let $B(0) = 0$ and $B(1) = \xi(1)$. For each $n \in \mathbb{N}$ we define the random variables $B(d)$, $d \in \mathcal{D}_n$ such that

- for all $r < s < t$ in $\mathcal{D}_n$ the random variable $B(t) - B(s)$ is normally distributed with mean zero and variance $t - s$, and is independent of $B(s) - B(r)$,
- the vectors $\{B(d) : d \in \mathcal{D}_n\}$ and $\{\xi(t) : t \in \mathcal{D} \setminus \mathcal{D}_n\}$ are independent.

Formally, we obtain

$$F_0(t) = \begin{cases} 0, & \text{for } t = 0, \\ \xi(1), & \text{for } t = 1, \\ \text{linear interpolation between} \end{cases}$$

and

$$F_n(t) = \begin{cases} 2^{-(n+1)/2}\xi(t), & \text{for } t \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}, \\ 0, & \text{for } t \in \mathcal{D}_{n-1}, \\ \text{linear between consecutive points in } \mathcal{D}_n. \end{cases}$$

Hence these functions are continuous in $[0,1]$ by construction and we can define for all n and $d \in \mathcal{D}_n$

$$B(d) = \sum_{i=0}^{n} F_i(d) = \sum_{i=0}^{\infty} F_i(d).$$

Then, by some technicalities, it is possible to prove that it converges to Brownian motion (in law). □

From the standard Brownian motion we can obtain other Brownian motions "zooming" the standard one or going backward w.r.t the time.

Lemma 2.19 ([41], Lemma 1.7)**.** *Let us consider $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. Suppose that $\{B(t)\}_{t\geq 0}$ is a Brownian motion and let $a > 0$. Then the process $\{\xi(t)\}_{t\geq 0}$ defined by $\xi(t) = \frac{1}{a}B(a^2 t)$ is a Brownian motion.*

Theorem 2.20 ([41], Theorem 1.9)**.** *Let us consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that $\{B(t)\}_{t\geq 0}$ is a standard Brownian motion, then the process $\{\xi(t)\}_{t\geq 0}$ defined by*

$$\xi(t) = \begin{cases} 0, & \text{for } t = 0, \\ tB(\frac{1}{t}), & \text{for } t > 0 \end{cases}$$

is also a standard Brownian motion.

Corollary 2.21 ([41], Corollary 1.11)**.** *[Law of large numbers] Let us consider $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space and $\{B(t)\}_{t\geq 0}$ a standard Brownian motion. Then it holds almost surely $\lim_{t\to\infty} \frac{B(t)}{t} = 0$.*

Theorem 2.22 ([41], Theorem 1.12)**.** *Let us consider a probability space $(\Omega, \mathbb{F}, \mathbb{P})$ and a standard Brownian motion $\{B(t)\}_{t\geq 0}$. There exists $C > 0$ real*

constant such that for all $h > 0$ sufficiently small and for all $0 \leq t \leq 1 - h$, such that

$$|B(t+h) - B(t)| \leq C\sqrt{h \log(1/h)}$$

almost surely.

Remark 2.23. *An immediate consequence of Theorem 2.22 it that the Brownian motion upper bounded in time.*

An interesting fact of the Brownian motion is related to its regularity. In fact we can see that this stochastic process is not differentiable anywhere.

Theorem 2.24 ([41], Theorem 1.13)**.** *Let us consider a probability space $(\Omega, \mathbb{F}, \mathbb{P})$ and a standard Brownian motion $\{B(t)\}_{t\geq 0}$. Then for every constant $c < \sqrt{2}$ and for every $\varepsilon > 0$ there exist $0 < h < \varepsilon$ and $t \in [0, 1-h]$ such that*

$$|B(t+h) - B(t)| \geq c\sqrt{h \log(1/h)},$$

almost surely.

Remark 2.25. *From Theorem 2.24 follows that $B(t)$ is not differentiable.*

Theorem 2.26 (Modulus of continuity)**.** *Let us consider a probability space $(\Omega, \mathbb{F}, \mathbb{P})$ and a standard Brownian motion $\{B(t)\}_{t\geq 0}$. It holds*

$$\limsup_{h\to 0} \sup_{0\leq t\leq 1-h} \frac{|B(t+h) - B(t)|}{\sqrt{2h \log(1/h)}} = 1,$$

almost surely.

The Theorems 2.22, 2.24, 2.26 allow us to understand better regularity of the Brownian motions, in particular we have to introduce the following definition.

Definition 2.27. *Let $f : I \to \mathbb{R}$ be a function, the function is* α-Hölder continuous *(where $0 < \alpha < 1$) if for $C \geq 0$ it holds true*

$$|f(x) - f(y)| \leq C|x - y|^{\alpha},$$

for all $x, y \in I$.

By Theorem 2.26 we can prove the following result.

Theorem 2.28 ([41], Corollary 1.20)**.** *The Brownian motion is locally everywhere Hölder with exponent $\alpha < 1/2$.*

Remark 2.29. *The Brownian motion is not Hölder for $\alpha = \frac{1}{2}$.*

We cannot expect more regularity than the fact the Brownian motion is Hölder due to the following theorem.

Theorem 2.30 ([41], Theorem 1.30)**.** *The Brownian motion is nowhere differentiable almost surely.*

2.3 Itô Integral and Itô formula

Now we want to compute the stochastic integral

$$\int_a^b f(t,\omega)dB(t,\omega),$$

where $f : [a,b] \times \Omega \to \mathbb{R}$ is a measurable function. For sake of simplicity we will write the integral as

$$\int_a^b f(t)dB(t).$$

Unfortunately the Riemann definition of integral leads to some problems, as we can show in the following example.

Example 2.31. *We use the definition of Riemann integral to compute the integral*

$$\int_0^T B(s)dB(s)$$

as counterexample. Firstly we observe that

$$\mathbb{E}\left[\int_0^T B(s)dB(s)\right] = \lim_{k\to 0}\mathbb{E}\left[\sum_k B(t_k)(B(t_{k+1}) - B(t_k))\right] = 0$$

due to the property of Brownian motion. However, using the definition of Riemann integral we have and assuming that chain rule holds, i.e. $B(t)$ is differentiable

$$\int_0^T B(s)dB(s) = \frac{B(t)^2}{2}$$

and, recalling that $\mathbb{E}\big[\frac{B(t)^2}{2}\big] = \frac{t}{2}$, there is a contradiction. Hence we cannot use the definition of Riemann integral to compute the integral.

Roughly speaking, we have to approximate the integral in a different way: we consider the following sum such that

$$\sum_i f(t_i^*)[B(t_{i+1}) - B(t_i)] \to \int_a^b f(t)dB(t).$$

The convergence of the sum is in a weak sense and $t_i^* \in [t_i, t_{i+1}]$. If $t_i^* = t_i$ then the integral is called *Itô* integral, meanwhile if $t_i^* = \frac{t_{i+1}-t_i}{2}$ then is called *Stratonovich* integral. These two integrals give two different results but there are some law which allows us to change one integral into the other.

Example 2.32. *Now we compute explicitly the integral*

$$\int_0^T B(s)dB(s),$$

following the definition of Itô integral. We can write

$$B(t_k) = \frac{1}{2}(B(t_{k+1}) + B(t_k)) - \frac{1}{2}(B(t_{k+1}) - B(t_k)),$$

and we have

$$\begin{aligned}\sum_{k\le n} B(t_k)(B(t_{k+1}) - B(t_k)) &= \sum_{k\le n}\frac{1}{2}(B(t_{k+1})^2 - B(t_k)^2) - \sum_{k\le n}\frac{1}{2}(B(t_{k+1}) - B(t_k))^2 \\ &= \frac{1}{2}B(t_k)^2 - \sum_{k\le n}\frac{1}{2}(B(t_{k+1}) - B(t_k))^2.\end{aligned}$$

The second term is composed by elements whose expectation is equal to $\frac{\Delta t}{2}$, where Δt is the length of the interval $[t_k, t_{k+1}]$ and their variance $\frac{\Delta t^2}{2}$. Hence the sums are $\frac{n\Delta t}{2}$ and $\frac{n\Delta t^2}{2}$ and so

$$\sum_{t_n<T}\frac{1}{2}(B(t_{k+1}) - B(t_k))^2 \to \frac{T}{2} \text{ as } \Delta t \to 0.$$

The result in this case is

$$\int_0^T B(s)dB(s) = \frac{1}{2}B(T) - \frac{T}{2},$$

meanwhile using the Stratonovich definition we have

$$\sum_{k\leq n} \frac{B(t_{k+1}) + B(t_k)}{2}(B(t_{k+1}) - B(t_k)) = \sum_{k\leq n} \frac{(B(t_{k+1})^2 - B(t_k)^2)}{2}.$$

Computing the series above, we obtain

$$\int_0^T B(s) \circ dB(s) = \frac{B(T)^2}{2}.$$

Definition 2.33. *Let $B(t) = B(t,\omega)$ be a N-dimensional Brownian motion, then we define $\mathcal{F}_t = \mathcal{F}_t^{(n)}$ to be a σ-algebra generated by the random variables $B(s)$ with $s \leq t$. In other words, $\mathcal{F}_t$ is the smallest σ-algebra containing all the set in the form*

$$\{\omega : B(t_1) \in F_1, \ldots, B(t_k) \in F_k\}$$

where $t_j \leq t$ and $F_j \subset \mathbb{R}^N$ are Borel sets, $j \leq k = 1, \ldots$
$\mathcal{F}_t$ is increasing if and only if for $s < t$ then $\mathcal{F}_s \subset \mathcal{F}_t$ and $\mathcal{F}_t \subset \mathcal{F}$.

Definition 2.34. *Let $\{\mathcal{F}_t\}_{t\geq 0}$ be an increasing family of σ-algebras of subsets of Ω. A process $f(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}^n$ is called $\mathcal{F}_t$-adapted if for each $t \geq 0$ the function*

$$\omega \to f(t,\omega)$$

is $\mathcal{F}_t$-measurable.

Definition 2.35. *Let $\mathcal{V} = \mathcal{V}(S,T)$ be the class of functions*

$$f(t,\omega) : [0,\infty[\times\Omega \to \mathbb{R}$$

such that

- *$(t,\omega) \to f(t,\omega)$ is $\mathcal{B}\times\mathcal{F}$-measurable, where $\mathcal{B}$ denotes the Borel σ-algebra on $[0,\infty[$,*

- $f(t,\omega)$ *is* $\mathcal{F}_t$*-adapted,*
- $\mathbb{E}[\int_S^T f(t,\omega)^2 dt] < \infty$.

Definition 2.36. *A function* $\phi \in \mathcal{V}$ *is called* elementary *if it has the form*

$$\phi(t,\omega) = \sum_j e_j(\omega)\chi_{[t_j,t_{j+1}[}(t).$$

Note that, since $\phi \in \mathcal{V}$*, each function* e_j *must to be* $\mathcal{F}_{t_j}$*-measurable.*

Definition 2.37. *The Itô integral for an elementary function is defined as*

$$\int_S^T \phi(t,\omega)dB(t) = \sum_{j\geq 0} e_j(\omega)[B(t_{j+1}) - B(j_j)](\omega).$$

We consider ω fixed for the next definitions and remarks.

Theorem 2.38 (Itô isometry for elementary functions, [45], p.26)**.** *If* $\phi(t,\omega)$ *is bounded and elementary then*

$$\mathbb{E}\left[\left(\int_S^T \phi(t,\omega)dB(t,\omega)\right)^2\right] = \mathbb{E}\left[\int_S^T \phi(t,\omega)^2 dt\right].$$

Now we can give a formal definition for the Itô integral.

Definition 2.39. *Let* $f \in \mathcal{V}(S,T)$*. Then the Itô integral of* f *is defined by*

$$\int_S^T \phi_n(t,\omega)dB(t,\omega) \to^{L^2(\mathbb{P})} \int_S^T f(t,\omega)dB(t,\omega), \tag{2.1}$$

where $\{\phi_n\}$ *is a sequence of elementary functions such that*

$$\mathbb{E}\left[\int_S^T |(f(t,\omega) - \phi_n(t,\omega))|^2 dt\right] \to 0. \tag{2.2}$$

Remark 2.40. *The limit* (2.1) *exists and does not depend on the actual choice of* $\{\phi_n\}$ *as soon as* (2.2) *holds true.*

Using the Definition 2.39 one can prove the following theorem which generalizes Theorem 2.38.

Theorem 2.41 (Itô isometry, [45], Corollary 3.1.7)**.** *For all* $f \in \mathcal{V}(S,T)$

$$\mathbb{E}\left[\left(\int_S^T f(t,\omega)dB(t,\omega)\right)^2\right] = \mathbb{E}\left[\int_S^T f^2(t,\omega)dt\right].$$

Corollary 2.42 ([45], Corollary 3.1.8)**.** *If* $f(t) \in \mathcal{V}(S,T)$ *and* $f_n(t,\omega) \in \mathcal{V}(S,T)$ *for* $n = 1, 2\ldots$ *and* $\mathbb{E}[\int_S^T (f_n(t) - f(t))^2 dt] \to 0$ *as* $n \to \infty$ *then*

$$\int_S^T f_n(t,\omega)dB(t,\omega) \to^{L^2(\mathbb{P})} \int_S^T f(t,\omega)dB(t,\omega),$$

as $n \to \infty$ *where by* $L^2(\mathbb{P})$ *convergence we denote the standard* L^2 *convergence w.r.t. the probability measure* $\mathbb{P}$*.*

It is interesting to observe that this theorem holds, which allows us to work with the integral in a similar way to the Riemann integral

Theorem 2.43 ([45], Theorem 3.2.1)**.** *Let* $f, g \in \mathcal{V}$ *and* $0 \le S \le U \le T$*, then we have*

- $\int_S^T f dB(t) = \int_S^U f dB(t) + \int_U^T f dB(t)$,
- $\int_S^T (cf + g)dB(t) = c\int_S^T f dB(t) + \int_S^T g dB(t)$,
- $\mathbb{E}\left[\int_S^T f dB(t)\right] = 0$,
- $\int_S^T f dB(t)$ *is* $\mathcal{F}_T$*-measurable.*

Definition 2.44. *Let us consider* $\{\xi(t)\}_{t\ge 0}$ *a real valued stochastic process defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*. Then the* quadratic variation *is the quantity*

$$[\xi]_t = \lim_{\|P\|\to 0} \sum_{k=1}^{n} (\xi(t_k) - \xi(t_{k-1}))^2,$$

where P *ranges over partitions of the interval* $[0,t]$*.*

Theorem 2.45. *([45]) The following relation between the Stratonovich integral and Itô integral holds true*

$$\int_S^T f \circ dB(t) = \int_S^T f dB(t) + \frac{1}{2}\int_S^T d < f, B(t) >,$$

where the last term on the right hand side is defined as the following $L^2(\mathbb{P})$ *limit*

$$\int_S^T d < f, B(t) > :=^{L^2(\mathbb{P})} \lim_{N\to\infty} \sum_{i=1}^{N} (f(t_{i+1}) - f(t_i))(B(t_{i+1}) - B(t_i)).$$

Definition 2.46. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and let* $\xi : \Omega \to \mathbb{R}^N$ *be a random variable such that* $\mathbb{E}[|\xi|] < \infty$. *If* $\mathcal{J} \subset \mathcal{F}$ *is a* σ*-algebra then the* conditional expectation *of* ξ *given by* $\mathcal{J}$, *denoted by* $\mathbb{E}[\xi|\mathcal{J}]$ *is defined as follows:*

- $\mathbb{E}[\xi|\mathcal{J}]$ *is* $\mathcal{H}$*-measurable,*
- $\int_H \mathbb{E}[\xi|\mathcal{J}]d\mathbb{P} = \int_J \xi d\mathbb{P}$ *, for all* $J \in \mathcal{J}$.

Definition 2.47. *A filtration is a family* $\mathcal{M} = \{\mathcal{M}_t\}_{t\geq 0}$ *of* σ*-algebra* $\mathcal{M}_t \subset \mathcal{F}$ *such that*

$$0 \leq s < t \Rightarrow \mathcal{M}_s \subset \mathcal{M}_t.$$

An N*-dimensional stochastic process* $\{\xi(t)\}_{t\geq 0}$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$ *is called a* martingale *with respect to a filtration* $\{\mathcal{M}_t\}_{t\geq 0}$ *if*

- $\xi(t)$ *is* $\mathcal{M}_t$*-measurable for all* t,
- $\mathbb{E}[|\xi(t)|] < \infty$ *for all* t,
- $\mathbb{E}[\xi(s)|\mathcal{M}_t] = \xi(t)$ *for all* $s \geq t$.

Theorem 2.48 ([45], Theorem 3.2.5)**.** *Let* $f \in \mathcal{V}(0, T)$. *Then there exists a* t*-continuous version of*

$$\int_0^t f(s)dB(s); \quad 0 \leq t \leq T$$

i.e there exists a t*-continuous stochastic process* $J(t)$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$ *such that*

$$\mathbb{P}\left[J(t) = \int_0^t f dB(s)\right] = 1, \quad \text{for all } 0 \leq t \leq T \ .$$

Definition 2.49. *Let $\mathcal{V}_{\mathcal{J}} = \mathcal{V}_{\mathcal{J}}(S,T)$ be the class of functions*

$$f(t,\omega) : [0,\infty[\times\Omega \to \mathbb{R}$$

such that

- *$(t,\omega) \to f(t,\omega)$ is $\mathcal{B}\times\mathcal{F}$-measurable, where $\mathcal{B}$ denotes the Borel σ-algebra on $[0,\infty[$,*
- *There exists an increasing family of σ-algebras $\mathcal{H}_t$; $t \geq 0$ such that:*
  i) $B(t)$ is a martingale (see Definition 2.47) with respect to $\mathcal{J}_t$,
  ii) $f(t)$ is $\mathcal{H}_t$-adapted,
- *$\mathbb{E}[\int_S^T f(t,\omega)^2 dt] < \infty$.*

Let $B = (B_1, \dots B_n)$ be an n-dimensional Brownian motion, we denote by $\mathcal{V}_{\mathcal{J}}^{m\times n}(S,T)$ the set of $m \times n$ matrices $v = [v_{ij}(t,\omega)]$, where $v_{ij} \in \mathcal{V}_{\mathcal{J}}$. If $v \in \mathcal{V}_{\mathcal{J}}^{m\times n}(S,T)$, using matrix notation, we define

$$\int_S^T v dB := \int_S^T \begin{bmatrix} v_{11} & \dots & v_{1n} \\ . & & . \\ . & & . \\ v_{m1} & \dots & v_{mn} \end{bmatrix} \begin{bmatrix} dB_1 \\ . \\ . \\ dB_n \end{bmatrix}$$

as the $m\times 1$ matrix whose i-th component is the sum of (extended) 1-dimensional Itô integrals:

$$\sum_{j=1}^{n} \int_S^T v_{ij}(s) dB_j(s).$$

Definition 2.50. *We denote with $\mathcal{W}_{\mathcal{J}}(S,T)$ the class of processes $f(t,\omega) \in \mathbb{R}$ satisfying the following conditions:*

- *$(t,\omega) \to f(t,\omega)$ is $\mathcal{B}\times\mathcal{F}$-measurable, where $\mathcal{B}$ denotes the Borel σ-algebra on $[0,\infty)$,*
- *There exists an increasing family of σ-algebras $\mathcal{J}_t$; $t \geq 0$ such that:*
  i) $B(t,\omega)$ is a martingale with respect to $\mathcal{J}_t$,
  ii) $f(t,\omega)$ is $\mathcal{J}_t$-adapted,

- $\mathcal{P}\left[\int_S^T f(t,\omega)^2 ds < \infty\right]$.

We put $\mathcal{W}_{\mathcal{J}} = \bigcap_{T>0} \mathcal{W}_{\mathcal{J}}(0,T)$.

Definition 2.51. *Let $B(t)$ be a 1-dimensional Brownian motion $(\Omega, \mathcal{F}, \mathbb{P})$ An Itô process is a stochastic process $\xi(t)$ on $(\Omega, \mathcal{F}, \mathbb{P})$ of form*

$$\xi(t) = \xi(0) + \int_0^t b(\xi(s))ds + \int_0^t \sigma(\xi(s))dB(s)$$

where $\sigma \in \mathcal{W}_{\mathcal{J}}$ so that

$$\mathbb{P}\left[\int_0^t \sigma(s)^2 ds < \infty \text{ for all } t \geq 0\right] = 1.$$

We also assume that b is $\mathcal{J}_t$-adapted and

$$\mathbb{P}\left[\int_0^t |b(s)| ds < \infty \text{ for all } t \geq 0\right] = 1.$$

Remark 2.52. *Equivalently an Itô process can be written as*

$$\begin{cases} d\xi(t) = bdt + \sigma dB(t), \\ \xi(0) = x \end{cases}$$

and this is a first example of stochastic differential equation.

Unfortunately, the computation rules of classical analysis cannot be extended to stochastic differential but we can found different rules in case of Itô integral which can allows us to compute explicitly the value of the differential.

Theorem 2.53 (1-dimensional Itô formula, [45], Theorem 4.1.2)**.** *Let $\xi(t)$ be an Itô process given by*

$$d\xi(t) = bdt + \sigma dB(t).$$

Let $g(t,x) \in C^2([0,\infty) \times \mathbb{R})$, then

$$\zeta(t) = g(t, \xi(t))$$

is again an Itô process and

$$d\zeta(t) = \frac{\partial g}{\partial t}(t,\xi(t))dt + \frac{\partial g}{\partial x}(t,\xi(t))d\xi(t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,\xi(t))(d\xi(t))^2$$

where $(d\xi(t))^2 = d\xi(t) \cdot d\xi(t)$ *is computed according to the rules*

$$dt \cdot dt = 0, \quad dB(t) \cdot dt = 0, \quad dB(t) \cdot dB(t) = dt.$$

Example 2.54. *We want to estimate the integral*

$$\int_0^t B(s)dB(s).$$

We choose $\xi(t) = B(t)$ *and* $g(t,x) = \frac{1}{2}x^2$*, then*

$$\zeta(t) = g(t, B(t)) = \frac{1}{2}B(t)^2$$

is an Itô process. then we have

$$\begin{aligned} d\zeta(t) &= \frac{\partial g}{\partial t}dt + \frac{\partial g}{\partial x}dB(t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(dB(t))^2 \\ &= B(t)dB(t) + \frac{1}{2}(dB(t))^2 = B(t)dB_t + \frac{1}{2}dt \end{aligned}$$

or, equivalently

$$\frac{1}{2}B(t)^2 = \int_0^t B(s)dB(s) + \frac{1}{2}t.$$

Integration by parts still holds, as stated in the following theorem.

Theorem 2.55 ([45], Theorem 4.1.5)**.** *Suppose* $f(s,\omega) = f(s)$ *and so* f *is continuous and of bounded variation in* $[0,t]$

$$\int_0^t f(s)dB(s) = f(t)B(t) - \int_0^t B(s)df(s).$$

Let $B(t) = (B_1(t), \ldots, B_m(t))$ be a m-dimensional Brownian motion. If $u_i(t)$ and $v_{ij}(t,\omega)$ satisfy the conditions given in the definition of 1-dimensional Itô process, then we can write the n-dimensional Itô process as

$$\begin{cases} d\xi_1 = b_1 dt + \sigma_{11}dB_1 + \cdots + \sigma_{1m}dB_m, \\ \ldots \\ \ldots \\ d\xi_n = b_n dt + \sigma_{n1}dB_1 + \cdots + \sigma_{nm}dB_m, \end{cases}$$

or, in matrix notation, in a simpler way as

$$d\xi(t) = bdt + \sigma dB(t).$$

Theorem 2.56 ([45], Theorem 4.2.1)**.** *Let*

$$d\xi(t) = bdt + \sigma dB(t)$$

be an n-dimensional Itô process and let $g(t,\omega) = (g_1(t,\omega), \ldots, g_p(t,\omega))$ *be a* C^2 *map from* $[0,\infty) \times \mathbb{R}^N$ *into* $\mathbb{R}^p$*, then the process*

$$\zeta(t,\omega) = g(t,\xi(t))$$

is again an Itô process whose component Y_k *is given by*

$$d\zeta_k = \frac{\partial g_k}{\partial t}(t,\xi)dt + \sum_{i=1}^{n} \frac{\partial g_k}{\partial x_i}(t,\xi(t))d\xi_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 g_k}{\partial x_i \partial x_j} d\xi_i d\xi_j$$

where the product $d\xi_i \cdot d\xi_j$ *following the rules* $dB_i \cdot dB_j = \delta_{ij}dt$*,* $dB_i \cdot dt = dt \cdot dB_i = dt \cdot dt = 0$*.*

2.4 Stochastic differential equation

In many cases it is difficult to express explicitly the solution of an SDE but the following theorem states that under certain conditions there is the existence and the uniqueness of the solution for SDE.

Theorem 2.57 ([45], Theorem 5.2.1)**.** *Let* $T > 0$ *and* $b : [0,T] \times \mathbb{R}^N \to \mathbb{R}^N$*,* $\sigma : [0,T] \times \mathbb{R}^N \to \mathbb{R}^{n\times m}$ *be measurable functions satisfying*

$$|b(t,x)| + |\sigma(t,x)| \leq C(1+|x|), \quad x \in \mathbb{R}^N, \quad t \in [0,T]$$

for some constant C *(where* $|\sigma|^2 = \sum |\sigma_{ij}|^2$*) and such that*

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq D|x-y|, \quad x,y \in \mathbb{R}^N, \quad t \in [0,T]$$

for some constant D*. Let* ζ *be a random variable which is independent from* σ*-algebra* $\mathcal{F}_\infty^{(m)}$ *generated by* $B(s)$*,* $s \geq 0$ *and such that*

$$\mathbb{E}[|\zeta|^2] < \infty.$$

Then the stochastic differential equation (SDE)

$$\begin{cases} d\xi(t) = b(t,\xi(t))dt + \sigma(t,\xi(t))dB(t), \quad 0 \le t \le T, \\ \xi(0) = \zeta \end{cases}$$

has a unique t-continuous solution $\xi(t)$ *with the property that* $X_t(\omega)$ *is adapted to the filtration* $\mathcal{F}_t^Z$ *generated by* ζ *and* $B(s)$, $s \le t$ *and*

$$\mathbb{E}\left[\int_0^T |\xi(t)|^2 dt\right] < \infty.$$

At this point it is important to define the difference between *strong* and *weak* solutions. This distinction is important because allows us to work with a weaker notion option of solution which is often used in stochastic control theory.

Definition 2.58. *Let* $\{B(t)\}_{t\ge 0}$ *be a standard Brownian motion on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *with an admissible filtration* $\mathcal{F} = \{\mathcal{F}_t\}_{t\ge 0}$. *A* strong solution *of the stochastic differential equation*

$$d\xi(t) = b(t,\xi(t))dt + \sigma(t,\xi(t))dB(t)$$

with the initial condition $x \in \mathbb{R}$ *is an adapted process* $\xi(t) = \xi^x(t)$ *with continuous paths such that, for all* $t \ge 0$,

$$\xi(t) = x + \int_0^t b(s,\xi(s))ds + \int_0^t \sigma(s,\xi(s))dB(s) \quad \textit{almost surely.} \tag{2.3}$$

Definition 2.59. *A* weak solution *of the stochastic differential equation with the initial condition* x *is a continuous stochastic process* $\xi(t)$ *defined on some probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *such that for some Brownian motion* $B(t)$ *defined on this probability space and some admissible filtration* $\mathcal{F}$ *the process* $\xi(t)$ *is adapted and satisfies the Equation* (2.3).

Chapter 3

Mean Curvature Flow and viscosity solutions

A hypersurface embedded in $\mathbb{R}^N$ may evolve in many different ways. One of the most studied type of evolutions is according to its geometrical properties. One of the most studied among these evolutions is the *mean curvature flow* due to its applications not only in theoretical and applied maths but also in physics and neuroscience. The main examples are the reconstruction of images using the model of Citti-Sarti and the study of the visual cortex (see [13], [14]). We say that a hypersurface evolves by MCF if it contracts in the direction of its normal direction with a normal velocity proportional to its mean curvature.

In this chapter we will introduce the ideas of evolution by mean curvature flow and of viscosity solution. In the end, we will state shortly a theorem which connects stochastic optimal control and Euclidean mean curvature flow.

3.1 A general survey

We recall some standard definitions from [28] in order to introduce the general equation in the Euclidean setting.

Definition 3.1 (Evolution by mean curvature flow)**.** *A family of smooth embedded hypersurfaces $\{\Gamma_t\}_{t\in I}$ in $\mathbb{R}^{N+1}$ moves by mean curvature if, for any smooth curve $y(t):[0,T]\to\Gamma_t$ such that $y(t)\in\Gamma_t$ it holds*

$$\frac{\partial y(t)}{\partial t} = -\vec{k}(y(t)) \tag{3.1}$$

for $y(t)\in\Gamma_t$ and $t\in I$, $I\subset\mathbb{R}$ an open interval. Here $\vec{k}(y)$ is the mean curvature vector at $y\in\Gamma$, i.e.

$$\vec{k}(y) = k(y)n_E(y),$$

where $k(y)$ is the mean curvature at the point $y\in\Gamma$ (defined as the divergence of the Euclidean normal) and $n_E(y)$ the outer (Euclidean) normal vector.

From now we can consider the family of smooth embeddings $F(\ ,t):\Gamma^N\to\mathbb{R}^{N+1}$ with $\Gamma_t = F(\Gamma_0,t)$ where Γ_0 is an N-dimensional manifold. We can rewrite the Definition 3.1 as

$$\frac{\partial F}{\partial t}(x,t) = -\vec{k}(F(x,t)), \tag{3.2}$$

for $x\in\Gamma_0$ and $t\in I$.

Example 3.2. *Let us consider an hyperplane in $\mathbb{R}^N$ parametrized as $a_1x_1+\cdots+a_Nx_N=0$ where $a_1,\ldots,a_N\in\mathbb{R}$. This implies that the equation (3.2) is given by*

$$\frac{\partial F}{\partial t}(x,t) = \vec{0},$$

because trivially the curvature of the plane is $\vec{k}=\vec{0}$.

Example 3.3. *Let us consider a sphere of radius r with the centre in the origin of $\mathbb{R}^N$. As the mean curvature is everywhere equal to N/r and since we choose the pointing inwards unit normal, the evolution equation is defined, considering $F(x,t)=r(t)F_0(x)$*

$$r'(t)F_0(x) = \frac{\partial}{\partial t}F(x,t) = k(x,t)n_E(x,t) = -NF_0(x)/r(t),$$

which is an ODE that can be integrated elementary (using the technique of separation of variables) and we get $r(t) = \sqrt{r^2 - 2Nt}$. *We have a singularity at the time* $T_{max} = r^2/(2N)$ *then we can write* $r(t) = \sqrt{2N(T_{max} - t)}$. *This is the most simple example of* homotetic solution *(as stated in Definition 3.5).*

Example 3.4. *Let us consider a cylinder* $S^{N-m}(\mathbb{R}) \times \mathbb{R}^m$. *Similarly to the Example 3.3 we have*

$$r'(t) = -\frac{(N-m)}{r(t)}$$

so we have as solution $r(t) = \sqrt{r_0^2 - 2(N-m)t}$.

Definition 3.5. *We call the solution for mean curvature flow in the form*

$$\Gamma_t = \lambda(t)\Gamma_0$$

homotetic solution.

Example 3.6. *The simplest example of a homotetic solution of mean curvature flow is given by* shrinking spheres *(see Example 3.3 for further details), in fact if the radius is* r, *these satisfies*

$$\Gamma_t = \sqrt{r^2 - 2Nt}\Gamma_0$$

for all $t \in (-\infty, \frac{1}{2N})$.

Not all the surfaces evolve smoothly, unfortunately there are some manifolds which may develop singularities such as the dumbbell in $\mathbb{R}^3$. This fact changes the topological property of the object which we are studying, for instance from a connected manifold we can have two separate connected manifolds (see [28], [31] for further details).

Theorem 3.7 ([28], Theorem 2.5)**.** *For every smoothly embedded, convex and compact hypersurface* Γ_0 *the solution of mean curvature flow remains smoothly embedded, compact until disappears into a point in finite time. In this process the solution become asymptotically round.*

Proposition 3.8 (Sphere comparison, [28], Proposition 3.3.)**.** *Let $\{\Gamma_t\}_{t\geq 0}$ be a solution of mean curvature flow, then*

$$\Gamma_0 \subset B_r(x_0) \Rightarrow \Gamma_k \subset B_{\sqrt{r^2-2Nk}}(x_0),$$

and

$$\Gamma_0 \cap B_r(x_0) = \emptyset \Rightarrow \Gamma_k \cap B_{\sqrt{r^2-2Nk}}(x_0) = \emptyset,$$

for all $t \in [0,T)$ with $k \leq \frac{\rho^2}{2N}$.

Definition 3.9. *We say that a solution of mean curvature flow $\{\Gamma_t\}_{t<t_0}$ reaches $x_0 \in \mathbb{R}^N$ at the time t_0 if there exists a sequence (x_j, t_j) such that for $t_n \to t_0$ we have that $x_j \in \Gamma_{t_j}$ and $x_j \to x_0$.*

Corollary 3.10 ([28], Corollary 3.6)**.** *Let $\{\Gamma_t\}_{t<t_0}$ be a solution of mean curvature flow which reaches $x_0 \in \mathbb{R}^N$ at the time t_0, then for all $t < t_0$,*

$$d(\Gamma_t, x_0) \leq \sqrt{2n(t_0 - t)}. \tag{3.3}$$

Theorem 3.11 ([28], Proposition 3.7)**.** *Let $\{\Gamma_t\}_{t>0}$ be a solution of mean curvature flow. If for $0 \leq \beta \leq N$ and some $\varepsilon > 0$ the initial hypersurface satisfies*

$$\Gamma_0 \subset \{x \in \mathbb{R}^N \mid (N-1-\beta)x_N^2 \geq |\overline{x}| - \varepsilon^2\},$$

where $\overline{x} = (x_1, \ldots, x_{N-1})$ then

$$\Gamma_t \subset \{x \in \mathbb{R}^N \mid (N-1-\beta)x_N^2 \geq |\overline{x}| - \varepsilon^2 + 2\beta t\},$$

for $t < \frac{\varepsilon^2}{2\beta}$ as long as the solution stays smooth for this time.

Remark 3.12 ([28])**.** *For $\beta = N$ and $\beta = N-1$ the Theorem 3.11 describes comparison with shrinking spheres and cylinders. For $0 < \beta < N-1$ the Theorem 3.11 implies that a singularity has to form at the latest at time $\frac{\varepsilon^2}{2\beta}$ since the solution of mean curvature flow will then be forced to lie inside a cone with the vertex at the origin.*

3.2 Introduction to viscosity solution

In this subsection we introduce the definition of viscosity solutions, which are a generalization of a classic solution for certain type of PDEs. In particular they are well suited for fully nonlinear PDEs. An important equation in PDE theory is the *eikonal equation*, which represents some wave propagation phenomena. The simplest example of eikonal equation is given by

$$\begin{cases} |u'(x)| = 1, \quad x \in (-1,1), \\ u(-1) = u(1) = 0. \end{cases}$$

Unfortunately it is possible to see, using Rolle's theorem, that it is impossible to solve this equation finding a classical solution. To show this explicitly we remark that, since the initial conditions are $u(-1) = u(1) = 0$ then there will exists a point $y \in (-1,1)$ such that $u'(y) = 0$. Hence there will be a subset $(-a,a) \subset (-1,1)$ (with $0 < a < 1$) such that $|u'(x)| < 1$ for $x \in (-a,a)$.

In order to solve this kind of equation we need a new class of solutions which are called *viscosity solutions* and were introduced by M.G.Crandall, L.C.Evans and P.L.Lions in 1984 (see [17], [18], [19], [20]). We give the general definition of viscosity solution for a second order differential equation.

Definition 3.13. *Let $F(x,u,Du,D^2u) = 0$ be a partial differential equation of second order and let $u : \Omega \subset \mathbb{R}^N \to \mathbb{R}$ a continuous function*

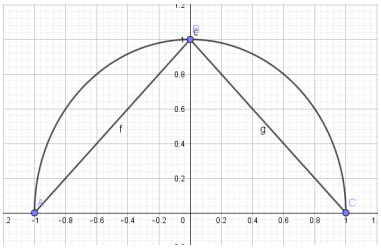
- *u is a* viscosity subsolution *in $x_0 \in \Omega$ of $F(x,u,Du,D^2u) = 0$ if for any $\phi \in C^2(\Omega)$ such that $u - \phi$ has a local maximum in x_0 we have*

$$F(x,u,D\phi,D^2\phi) \leq 0.$$

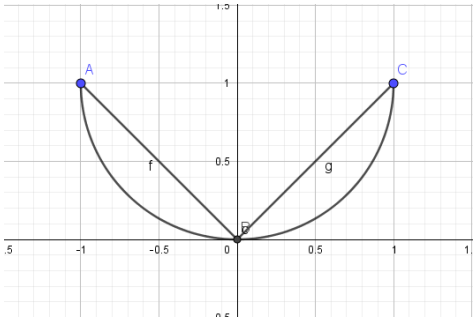
- *u is a* viscosity supersolution *in $x_0 \in \Omega$ of $F(x,u,Du,D^2u) = 0$ if for any $\phi \in C^2(\Omega)$ such that $u - \phi$ has a local minimum in x_0 we have*

$$F(x,u,D\phi,D^2\phi) \geq 0.$$

- *u is a* viscosity solution *if and only if it is (viscosity) supersolution and subsolution.*



Graphic representation of a viscosity subsolution. The test function touches above the viscosity solution



Graphic representation of a viscosity supersolution. The test function touches below the viscosity solution

The definition of viscosity solution is particularly useful in order to deal with a particular type of equations, called degenerate elliptic.

Definition 3.14. *Let $F(x, z, p, M)$ be a partial differential equation of order two, this equation is* degenerate elliptic *if and only if*

$$F(x, z, p, M) \leq F(x, z, p, \tilde{M}) \quad \text{if} \quad M \geq \tilde{M} \text{ (in the sense of matrices).}$$

So we will consider from this point degenerate elliptic equations. The following propositions can be proved easily using the property of viscosity solutions (see [17], [18], [19], [20], for further remarks).

Proposition 3.15. *Let us consider a degenerate elliptic equation $F(x, u, Du, D^2u) = 0$ as in Definition 3.14. A classical solution of the equation is a viscosity solution.*

Proposition 3.16. *Let $\Omega \subset \mathbb{R}^N$ an open set. Let $u \in C^2(\Omega)$ be a viscosity solution of the degenerate elliptic equation $F(x, u, Du, D^2u) = 0$, then it is a classical solution.*

Proposition 3.17. *Let u be a viscosity solution of the equation $F(x, u, Du, D^2u) = 0$ in an open domain then $v(x) = -u(x)$ is the viscosity solution equation of $-F(x, -v, -Dv, -D^2v) = 0$.*

We use viscosity solutions in particular because they have the stability property under certain conditions, as stated in the following theorem.

Theorem 3.18. *Let F_ε be continuous and converging (locally uniformly) to a function F as $\varepsilon \to 0$ and let u_ε be a viscosity solutions of*

$$F_\varepsilon(x, u_\varepsilon, Du_\varepsilon, D^2u_\varepsilon) = 0$$

such that $u_\varepsilon \to u$ (locally uniformly) as $\varepsilon \to 0^+$. Then u is a viscosity solution of

$$F(x, u, Du, D^2u) = 0.$$

In the end, we state a theorem of existence for viscosity solution.

Theorem 3.19. *([46]) [Perron's methods] Let $F(x, u, Du, D^2u) = 0$ be defined on an open set with a given condition on the boundary. We assume:*

- *If $\underline{u}$ is a subsolution and $\overline{u}$ is a supersolution which satisfies both the condition on the boundary then we have*

$$\underline{u} \leq \overline{u} \quad \text{on } \overline{\Omega},$$

- *There exist $\underline{u}$ subsolution and $\overline{u}$ supersolution which satisfies the boundary condition.*

We define

$$W(x) = \sup\{\underline{u}(x) \leq w(x) \leq \overline{u}(x) | \quad w \text{ is a subsolution}\}.$$

Then W is a viscosity solution of $F(x,u,Du,D^2u)=0$ which satisfies the boundary conditions solved by $\underline{u}$ and $\overline{u}$.

3.3 Level set equation: Euclidean case

We now focus our attention on the hypersurfaces which can be expressed as level set equations, i.e.

$$\Gamma_t = \{(x,t) \in \mathbb{R}^N \times (0,\infty) |\ u(x,t) = 0\}.$$

We will obtain the equation of the evolution by mean curvature flow in this case following [31]. We assume now that (heuristically) all the hypersurfaces Γ_t are smooth and their spatial gradient does not vanish in some open region called O of $(0,\infty)\times\mathbb{R}^N$. Let $n_E = n_E(x,t)$ the unit normal vector at the point x and at time t. Hence the point $x \in \Gamma_t \cap O$ for a fixed $t \geq 0$ evolves w.r.t. the ODE

$$\begin{cases} \dot{x}(s) = -[div(n_E)n_E](x(s),s), \quad s \geq t, \\ x(t) = x. \end{cases}$$

Since $x(s) \in \Gamma_s$ (with $s \geq t$) we obtain $u(x(s),s) = 0$ $(s > t)$ and then, differentiating, we have

$$0 = \frac{d}{ds}u(x(s),s) = -[(Du \cdot n_E)div(n_E)](x(s),s) + \frac{\partial u}{\partial t}(x(s),s).$$

Setting $s = t$ we obtain

$$\frac{\partial u}{\partial t} = (Du \cdot n_E)div(n_E) \ \text{ at } (x,t).$$

Then, recalling that for the level set it holds true $n_E = \frac{Du}{|Du|}$ and considering $g : \mathbb{R}^N \to \mathbb{R}$ smooth function which represents the starting hypersurface Γ_0 we obtain finally

$$\begin{cases} \frac{\partial u}{\partial t} = (Du \cdot n_E)div(n_E) = (\delta_{ij} - \frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}/|Du|^2)\frac{\partial^2 u}{\partial x_i \partial x_j}, \quad (x,t) \in \mathbb{R}^N \times (0,\infty), \\ u(x,0) = g(x), \quad on\ \mathbb{R}^N \times \{t=0\}. \end{cases} \tag{3.4}$$

With the help of the theory developed by Lions, Evans and Spruck ([31]) in 1991 we can define viscosity solutions for this equation in this way.

Definition 3.20. *([31]) A function* $u \in C(\mathbb{R}^N \times [0,\infty)) \cap L^\infty(\mathbb{R}^N \times [0,\infty))$ *is* weak subsolution *of* (3.4) *if for each* $\phi \in C^\infty(\mathbb{R}^{N+1})$ *such that* $u - \phi$ *has a local maximum at a point* $(x_0, t_0) \in \mathbb{R}^N \times (0,\infty)$ *it holds*

$$\begin{cases} \frac{\partial \phi}{\partial t} \leq (\delta_{ij} - \frac{\partial \phi}{\partial x_i}\frac{\partial \phi}{\partial x_j}/|D\phi|^2)\frac{\partial^2 \phi}{\partial x_i \partial x_j}, \quad at \ \ (x_0,t_0) \ \ if \qquad D\phi(x_0,t_0) \neq 0, \\ \frac{\partial \phi}{\partial t} \leq (\delta_{ij} - \eta_i\eta_j)\frac{\partial^2 \phi}{\partial x_i \partial x_j}, \quad for \ \ \eta \in \mathbb{R}^N \ with \ \ |\eta| \leq 1, \quad D\phi(x_0,t_0) = 0. \end{cases}$$

Definition 3.21. *([31]) A function* $u \in C(\mathbb{R}^N \times [0,\infty)) \cap L^\infty(\mathbb{R}^N \times [0,\infty))$ *is* weak supersolution *of* (3.4) *if for each* $\phi \in C^\infty(\mathbb{R}^{N+1})$ *such that* $u - \phi$ *has a local minimum at a point* $(x_0, t_0) \in \mathbb{R}^N \times (0,\infty)$ *it holds*

$$\begin{cases} \frac{\partial \phi}{\partial t} \geq (\delta_{ij} - \frac{\partial \phi}{\partial x_i}\frac{\partial \phi}{\partial x_j}/|D\phi|^2)\frac{\partial^2 \phi}{\partial x_i \partial x_j}, \quad at \ \ (x_0,t_0) \ \ if \qquad D\phi(x_0,t_0) \neq 0, \\ \frac{\partial \phi}{\partial t} \geq (\delta_{ij} - \eta_i\eta_j)\frac{\partial \phi}{\partial x_i \partial x_j}, \quad for \ \ \eta \in \mathbb{R}^N \ with \ |\eta| \leq 1, \quad D\phi(x_0,t_0) = 0. \end{cases}$$

Definition 3.22. *A function* $u \in C(\mathbb{R}^N \times [0,\infty)) \cap L^\infty(\mathbb{R}^N \times [0,\infty))$ *is a weak solution if* u *is both subsolution and supersolution.*

Another way to define a solution is given by Giga, Chen and Goto ([11]). Before introducing this definition, we have to define what are the upper and lower envelopes of a function.

Definition 3.23. *Let* X *be a metric spaces and* $L \subset X$*. Let* $h : L \to \mathbb{R} \cup \{+\infty, -\infty\}$ *be a function then the* upper semicontinuous envelope u^* *(resp. lower) is defined by*

$$h^*(z) = \lim_{r \to 0} \sup\{h(\xi) : \xi \in B_r(z) \cap L\}$$

and

$$h_*(z) = \lim_{r \to 0} \inf\{h(\xi) : \xi \in B_r(z) \cap L\}$$

where $z \in \overline{L}$*.*

Remark 3.24. *If a function $h(z)$ is continuous at the point z then $h(z) = h_*(z) = h^*(z)$.*

Definition 3.25. *Let Ω be open set in $\mathbb{R}^N$ and $T > 0$. Let $\mathcal{O}$ be an open set in $(0,T) \times \Omega$. Let $F : [0,T] \times \Omega \times \mathbb{R} \times \mathbb{R}^N \times Sym(N) \to \mathbb{R}$ continuous. Let $z = (t,x)$ then*

- *A function $u : \mathcal{O} \to \mathbb{R} \cup \{-\infty\}$ is a* viscosity subsolution *of*

$$\frac{\partial u}{\partial t} + F(z, u, Du, D^2u) = 0$$

in $\mathcal{O}$ if

1. *$u^*(z) < \infty$ for $z \in \mathcal{O}$,*
2. *If $(\phi, \overline{z}) \in C^2(\mathcal{O}) \times \mathcal{O}$ satisfies*

$$\max_{\mathcal{O}}(u^* - \phi) = (u^* - \phi)(\overline{z})$$

then

$$\frac{\partial \phi}{\partial t}(\overline{z}) + F(\overline{z}, u^*(\overline{z}), D\phi(\overline{z}), D^2\phi(\overline{z})) \leq 0.$$

- *A function $u : \mathcal{O} \to \mathbb{R} \cup \{+\infty\}$ is a* viscosity supersolution *in $\mathcal{O}$ if*

1. *$u_*(z) > -\infty$ for $z \in \mathcal{O}$,*
2. *If $(\phi, \overline{z}) \in C^2(\mathcal{O}) \times \mathcal{O}$ satisfies*

$$\min_{\mathcal{O}}(u_* - \phi) = (u_* - \phi)(\overline{z})$$

then

$$\frac{\partial \phi}{\partial t}(\overline{z}) + F(\overline{z}, u_*(\overline{z}), D\phi(\overline{z}), D^2\phi(\overline{z}))) \geq 0.$$

In this case F is continuous but, unfortunately, in mean curvature flow equation there are some points of discontinuity (which represents the singularities) so we have to adapt the Definition 3.25.

Definition 3.26. *We suppose that the function F is not continuous and $z = (t,x)$.*

- *Assume that F is lower semicontinuous in $W = [0,T] \times \Omega \times \mathbb{R} \times \mathbb{R}^N \times Sym(N)$ with values in $\mathbb{R} \cup \{-\infty\}$. A subsolution of*

$$\frac{\partial u}{\partial t} + F(z, u, Du, D^2u) = 0,$$

is defined as in the Definition 3.25.

- *Assume that F is upper semicontinuous in W with values in $\mathbb{R} \cup \{+\infty\}$. A supersolution of*

$$\frac{\partial u}{\partial t} + F(z, u, Du, D^2u) = 0,$$

is defined as in Definition 3.25.

- *Assume that F is defined only in a dense subset of W and that $F_* < \infty$, $F^* > -\infty$ in W. If u is a subsolution of*

$$\frac{\partial u}{\partial t} + F_*(z, u, Du, D^2u) = 0,$$

in $\mathcal{O}$, then u is called a subsolution of $u_t + F(z, u, Du, D^2u) = 0$ in $\mathcal{O}$. If u is a supersolution of

$$\frac{\partial u}{\partial t} + F^*(z, u, Du, D^2u) = 0,$$

in $\mathcal{O}$, then u is called a supersolution of $u_t + F(z, u, Du, D^2u) = 0$ in $\mathcal{O}$.

This definition allows us to write a new form of viscosity solution using the envelopes, in order to avoid some problems related to the continuity.

Remark 3.27. *We can write the equation of mean curvature flow as*

$$\frac{\partial u}{\partial t} + H(x, Du, D^2u) = 0,$$

where

$$H(x, p, S) = -Tr(S) + \left\langle \frac{Sp}{|p|}, \frac{p}{|p|} \right\rangle.$$

If $|Du| \neq 0$ *then we observe that*

$$H^*(x, Du, D^2u) = H(x, Du, D^2u), \quad H_*(x, Du, D^2u) = H(x, Du, D^2u).$$

If $|Du| = 0$ *then we observe that, after writing* $a = \frac{p}{|p|}$*, it holds*

$$\max_{|a|=1} < Sa, a >= \lambda_{max}(S),$$

$$\min_{|a|=1} < Sa, a >= \lambda_{min}(S),$$

and so we have

$$H^*(x, Du, D^2u) = -Tr(D^2u) + \lambda_{max}(D^2u),$$

and

$$H_*(x, Du, D^2u) = -Tr(D^2u) + \lambda_{min}(D^2u).$$

Using the Remark 3.27, we can adapt the definition of viscosity solution for mean curvature flow by level set.

Definition 3.28.

- *A* viscosity subsolution *(according to the Definition of [11]) at the point* (x_0, t_0) *of a mean curvature flow is a continuous function* u *such that if for any* $\phi \in C^2(\mathbb{R}^N \times (0, +\infty))$ *such that* $u - \phi$ *has a maximum in* (x_0, t_0) *then*

$$\begin{cases} \frac{\partial \phi}{\partial t} - \Delta\phi + \left\langle \frac{D^2\phi D\phi}{|D\phi|}, \frac{D\phi}{|D\phi|} \right\rangle \leq 0, & \text{if } |Du| \neq 0, \\ \frac{\partial \phi}{\partial t} - \Delta\phi + \lambda_{min}(D^2u) \leq 0, & \text{if } |Du| = 0. \end{cases}$$

- *A* viscosity supersolution *(according to the Definition of [11]) at the point* (x_0, t_0) *of a mean curvature flow is a continuous function* u *such that if for any* $\phi \in C^2(\mathbb{R}^N \times (0, +\infty))$ *such that* $u - \phi$ *has a minimum in* (x_0, t_0) *then*

$$\begin{cases} \frac{\partial \phi}{\partial t} - \Delta\phi + \left\langle \frac{D^2\phi D\phi}{|D\phi|}, \frac{D\phi}{|D\phi|} \right\rangle \geq 0, & \text{if } |Du| \neq 0, \\ \frac{\partial \phi}{\partial t} - \Delta\phi + \lambda_{max}(D^2u) \geq 0, & \text{if } |Du| = 0. \end{cases}$$

- u *is a* viscosity solution *(by Giga) if and only if is a subsolution and a supersolution.*

3.4 A stochastic approach to mean curvature

In this section we state briefly a result obtained independently by Cardaliaguet, Buckdahn and Quincampoix [27] and Soner and Touzi [49] in 2002 which connects the Euclidean mean curvature flow with stochastic optimal control theory. We define the following class of controls

$$\mathcal{V} = \{\nu \in Sym(N) \mid \nu \geq 0,\ I_N - \nu^2 \geq 0 \text{ and } Tr(I_N - \nu^2) = 1\}.$$

Hence, the set of admissible controls $\mathcal{V}$ is the set of projection matrix of codimension one. Now let $B_N(s)$ and N-dimensional $\{\mathcal{F}_s\}$-Brownian motion on some complete stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s\}_{s\geq 0})$. We denote with $\mathcal{A} = \mathcal{A}(\Omega, \mathcal{F}, \mathbb{P}, B)$ the set of all $\mathcal{V}$-valued $\{\mathcal{F}_s\}$-progressively measurable processes ν. A process $\nu \in \mathcal{A}$ is called admissible control. Then the following theorem holds.

Theorem 3.29. *([27], Theorem 1.1) Let $g : \mathbb{R}^N \to \mathbb{R}$ be a bounded uniformly continuous function. Let $T > 0$ be fixed and, for any initial condition $(t, x) \in [0, T] \times \mathbb{R}^N$ we write*

$$V(t,x) = \inf_{\nu\in\mathcal{A}} \left(ess \sup_{\omega\in\Omega} g(\xi^{t,x,\nu}(T))(\omega) \right),$$

where $\xi^{t,x,\nu}$ is the solution to

$$\begin{cases} d\xi^{t,x,\nu}(s) = \sqrt{2}\nu(s)dB_N(s), & s \in [0,t), \\ \xi^{t,x,\nu}(t) = x. \end{cases}$$

Then $V : [0, T] \times \mathbb{R}^N \to \mathbb{R}$ is the viscosity solution of the equation of the Euclidean mean curvature flow

$$\begin{cases} -\frac{\partial}{\partial t}V - \Delta V + \left\langle D^2V \frac{DV}{|DV|}, \frac{DV}{|DV|} \right\rangle = 0, & \text{in } (0,T) \times \mathbb{R}^N, \\ V(T,x) = g(x), & x \in \mathbb{R}^N. \end{cases}$$

In particular, the function V is continuous.

Roughly speaking, the controls given in $\mathcal{A}$ constrain the stochastic process in a subspace of Euclidean space $\mathbb{R}^N$ and the value function (which solves in viscosity sense the evolution by mean curvature flow) consists in the terminal cost with an essential supremum in the space of probability. Hence the problem to find a viscosity solution of evolution by mean curvature flow becomes a problem of minimization of a value function associated a stochastic dynamics.

We can give the intuition about the dynamic following the idea of [52]: the Brownian motion moving on the tangent plane of a generic manifold, moves away from the manifold in the normal direction with a velocity equal to the half of its mean curvature. Then, going back in time, this means that if the Brownian motion (multiplied by $\sqrt{2}$) starts on the solution of the mean curvature flow, called $\Gamma(t)$, and diffuses on the tangent plane at all times, then it will arrive $\Gamma(0)$ at time t. On the other part, the Brownian motion which is moving on some other plane would go away from the manifold, and we could never guarantee that it would return with probability one. For this reason we choose the class of controls $\mathcal{A}$.

Chapter 4

Horizontal Mean Curvature Flow

In this chapter we generalize the definition of the mean curvature flow to a generic sub-Riemannian setting. In this setting the equation associated to the evolution is not defined in some points called *characteristic points*, i.e. points in which the horizontal normal (defined as the renormalized projection of Euclidean normal w.r.t. the distribution of sub-Riemannian geometry) is not defined.

To avoid this issue, we will use the *Riemannian approximation* (as introduced in Chapter 1) which allows us to recover a Riemannian structure depending on a small $\varepsilon > 0$ and the approach of Chen, Giga and Goto for viscosity solution as seen in Chapter 3.

4.1 Horizontal mean curvature flow

Geometric definition

We recall briefly the definitions of normal in a Riemannian and a sub-Riemannian setting.

Definition 4.1. *Let Γ be a C^1 hypersurface embedded in $\mathbb{R}^N$, then the Euclidean normal at the point x is the vector $n_E(x)$ of length one such that*

$$< n_E(x), v >= 0 \quad \textit{for all} \quad v \in T_x\Gamma.$$

This Euclidean normal can be reprojected on the horizontal distribution, obtaining the horizontal normal.

Definition 4.2. *Let Γ be a C^1 hypersurface embedded in $\mathbb{R}^N$, then the* horizontal normal *at the point $x \in \Gamma$ is the renormalized projection of the Euclidean normal on the horizontal space, i.e.*

$$n_0(x) = \frac{pr_{\mathcal{H}} n_E(x)}{|pr_{\mathcal{H}} n_E(x)|_g},$$

where $n_E(x)$ is the Euclidean normal, $\mathcal{H}$ the distribution of sub-Riemannian geometry and g the associated metric.

Example 4.3 (Horizontal normal in $\mathbb{H}^1$)**.** *Let us consider the Euclidean normal $n_E(x) = (n_1(x), n_2(x), n_3(x))^T$ to a surface Γ, where $x = (x_1, x_2, x_3) \in \Gamma \subset \mathbb{R}^3$, then the projection to the horizontal space can be computed as*

$$pr_{\mathcal{H}} n_E(x) = \begin{bmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \end{bmatrix} \begin{bmatrix} n_1(x) \\ n_2(x) \\ n_3(x) \end{bmatrix} = \begin{bmatrix} n_1(x) - \frac{x_2}{2} n_3(x) \\ n_2(x) + \frac{x_1}{2} n_3(x) \end{bmatrix} = \begin{bmatrix} \alpha_1(x) \\ \alpha_2(x) \end{bmatrix},$$

w.r.t. the basis generated by X_1 and X_2. The horizontal normal can be written as

$$n_0(x) = \frac{pr_{\mathcal{H}} n_E(x)}{|pr_{\mathcal{H}} n_E(x)|_g} = \frac{(n_1(x) - \frac{x_2}{2} n_3(x), n_2(x) + \frac{x_1}{2} n_3(x))^T}{\sqrt{(n_1(x) - \frac{x_2}{2} n_3(x))^2 + (n_2(x) + \frac{x_1}{2} n_3(x))^2}}.$$

The definition of horizontal normal leads to some technical difficulty because at some points the horizontal normal is not defined meanwhile the Euclidean one exists. These points are called characteristic.

Definition 4.4. *Let $(\mathbb{R}^N, \mathcal{H}, g)$ be a sub-Riemannian geometry and $\Gamma \subset \mathbb{R}^N$ a hypersurface, then we define* characteristic points *the points where the Euclidean normal is perpendicular to the horizontal space, i.e.*

$$char(\Gamma) = \{x \in \mathbb{R}^N | \mathcal{H}T_x\Gamma = \mathcal{H}_x\mathbb{R}^N\},$$

where $\mathcal{H}T_x\Gamma$ represents the intersection of the Euclidean tangent space with the horizontal space at the point $x \in \Gamma$.

Another important definitions are mean curvature and horizontal mean curvature.

Definition 4.5. *Let $(\mathbb{R}^N, \mathcal{H}, g)$ be a sub-Riemannian geometry, $\Gamma \subset \mathbb{R}^N$ a hypersurface. Let $(\mathbb{R}^N, \mathcal{H}_\varepsilon, g_\varepsilon)$ be a Riemannian approximation of the sub-Riemannian geometry.*

- *The* mean curvature *is defined as the divergence of the Euclidean normal, i.e.*

  $$k(x) = div(n_E(x)),$$

  where $x \in \Gamma$ and $n_E(x)$ the Euclidean normal at $x \in \Gamma$.

- *The* horizontal mean curvature *is defined as the horizontal divergence of the horizontal normal, i.e.*

  $$k_0(x) = div_{\mathcal{H}} n_0(x).$$

  where $x \in \Gamma$, $n_0(x)$ the horizontal normal at $x \in \Gamma$ and $div_{\mathcal{H}}$ the horizontal divergence as in Definition 1.75 (i.e. the horizontal divergence is computed w.r.t. the family of vector fields $\mathcal{X}$ which spans $\mathcal{H}$).

Remark 4.6. *It is not possible to define the horizontal mean curvature in characteristic points since the horizontal normal is not defined.*

Formulation of the HMCF with level set

In this subsection we consider a manifold $\Gamma = \{x \in \mathbb{R}^N |\ u(x) = 0\}$ which is C^1.

Definition 4.7. *Let us consider* $\Gamma = \{x \in \mathbb{R}^N |\ u(x) = 0\} \subset \mathbb{R}^N$ *a* C^1 *hypersurface, the Euclidean normal is defined as (if* $|\nabla u(x)| \neq 0$*)*

$$n(x) = \frac{\nabla u(x)}{|\nabla u(x)|}. \tag{4.1}$$

We can project the normal (4.1) *in the sub-Riemannian setting, however in some points this projection is not defined, even if we reparametrize the manifold.*

Definition 4.8. *Let* $(\mathbb{R}^N, \mathcal{H}, g)$ *be a sub-Riemannian geometry and* $\Gamma = \{x \in \mathbb{R}^N \mid u(x) = 0\}$ *a* C^1 *hypersurface, then the horizontal normal for level sets is defined as*

$$n_0(x) := \frac{\nabla_0 u(x)}{|\nabla_0 u(x)|_g} = \left(\frac{X_1 u(x)}{\sqrt{\sum_{i=1}^m (X_i u(x))^2}}, \ldots, \frac{X_m u(x)}{\sqrt{\sum_{i=1}^m (X_i u(x))^2}} \right)^T.$$

Remark 4.9. *Let* $(\mathbb{R}^N, \mathcal{H}, g)$ *be a sub-Riemannian geometry, and* $\Gamma = \{x \in \mathbb{R}^N | u(x) = 0\}$ *a* C^1 *hypersurface, it holds true*

$$x \in \Gamma \quad \textit{characteristic point} \Leftrightarrow \nabla_0 u(x) = 0.$$

Example 4.10 (Horizontal normal for level set case in $\mathbb{H}^1$)**.** *Let us consider now the level set* $\Gamma = \{x \in \mathbb{H}^1 \mid u(x) = 0\}$*. We assume that* $u : \mathbb{R}^3 \to \mathbb{R}$ *is differentiable. Hence, the horizontal normal can be expressed as*

$$n_0(x) = \frac{\nabla_0 u(x)}{|\nabla_0 u(x)|_g} = \frac{(u_{x_1}(x) - \frac{x_2}{2} u_{x_3}(x), u_{x_2}(x) + \frac{x_1}{2} u_{x_3}(x))^T}{\sqrt{(u_{x_1}(x)) - \frac{x_2}{2} u_{x_3}(x))^2 + (u_{x_2}(x) + \frac{x_1}{2} u_{x_3}(x))^2}}.$$

Definition 4.11. *Let us consider* $\Gamma = \{x \in \mathbb{R}^N | u(x) = 0\}$ *a* $C^{1,1}$ *hypersurface, then the* mean curvature *in the point* $x \in \Gamma$ *is defined as the divergence of the normal i.e.*

$$k_E(x) = \sum_{i=1}^N \frac{\partial}{\partial x_i} \left(\frac{\nabla u(x)}{|\nabla u(x)|} \right).$$

This operator can be projected in the sub-Riemannian space (except at the characteristic points due to the fact that the horizontal normal is not defined).

Definition 4.12. *Let $(\mathbb{R}^N, \mathcal{H}, g)$ be a sub-Riemannian geometry, and $\Gamma = \{x \in \mathbb{R}^N |\ u(x) = 0\}$ a C^1 hypersurface, we define the* horizontal mean curvature *at $x \in \Gamma$ as*

$$k_0(x) = \sum_{i=1}^{m} X_i \left(\frac{X_i u(x)}{\sqrt{\sum_{i=1}^{m} (X_i u(x))^2}} \right).$$

Now it is possible to define the equation of evolution by horizontal mean curvature flow.

Definition 4.13. *Let Γ_t be a family of smooth hypersurfaces in a sub-Riemannian geometry. We say that Γ_t is an* evolution by horizontal mean curvature flow *of Γ_0 if and only if $\Gamma = \Gamma_0$ and for any smooth horizontal curve $x : [0, T] \to \mathbb{R}^N$ such that $x(t) \in \Gamma_t$ for all $t \in [0, T]$, the horizontal normal velocity v_0 (as defined in the left part of (4.2)) in is equal to minus the horizontal mean curvature, i.e.*

$$v_0(x(t)) := g_{x(t)}(\dot{x}(t), n_0(x(t))) = -k_0(x(t)), \tag{4.2}$$

where $n_0(x(t))$ and $k_0(x(t))$ as defined in Definitions 4.2 and 4.5.

Remark 4.14. *The Equation (4.2) is never defined at characteristic points.*

Furthermore, recalling [22], we know that the associated PDE to the evolution by horizontal mean curvature flow of a level set $\Gamma_t = \{(t, x) | u(t, x) = 0\}$ is given by

$$u_t = Tr((\mathcal{X}^2 u)^*) - \left\langle (\mathcal{X}^2 u)^* \frac{\mathcal{X}u}{|\mathcal{X}u|}, \frac{\mathcal{X}u}{|\mathcal{X}u|} \right\rangle = \Delta_0 u - \Delta_{0,\infty} u. \tag{4.3}$$

Remark 4.15. *We recall that $\Delta_{0,\infty}$ is the generalization in sub-Riemannian setting of the infinite horizontal Laplacian. In the Euclidean setting, the infinity Laplace equation was derived as the limit as $p \to \infty$ of the Euler-Lagrange equation*

$$div(|Du|^{p-2} Du) = |Du|^{p-2} \Delta u + (p-2)|Du|^{p-4} < D^2 u Du, Du >= 0.$$

Then, dividing all for $(p-2)|Du|^{p-2}$ *and taking* $p \to \infty$ *we obtain the infinite Laplacian in the Euclidean setting.*

4.2 Approximated Riemannian mean curvature flow

Let us start generalizing the definition of horizontal normal and horizontal mean curvature to the Riemannian approximation. In this way we will avoid the problems generated by characteristic points.

Definition 4.16. *Let* $(\mathbb{R}^N, \mathcal{H}, g)$ *be a sub-Riemannian geometry,* $\Gamma \subset \mathbb{R}^n$ *a hypersurface. Let us consider* $(\mathbb{R}^N, \mathcal{H}^\varepsilon, g_\varepsilon)$ *a Riemannian approximation of the sub-Riemannian geometry. Then the* approximated Riemannian normal *is given by*

$$n_\varepsilon(x) = \frac{pr_{\mathcal{H}^\varepsilon} n_E(x)}{|pr_{\mathcal{H}^\varepsilon} n_E(x)|_{g_\varepsilon}} \quad \text{with } x \in \Gamma,$$

where $n_E(x)$ *is the Euclidean normal,* $\mathcal{H}^\varepsilon$ *is Riemannian approximation of the distribution of sub-Riemannian geometry and* g_ε *is the associated metric.*

Example 4.17 (Approximated Riemannian normal for $\mathbb{H}^1$)**.** *We compute the approximated Riemannian normal for the 1-dimensional Heisenberg group. We compute*

$$pr_{\mathcal{H}^\varepsilon} n_E(x) = \begin{bmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \\ 0 & 0 & \varepsilon \end{bmatrix} \begin{bmatrix} n_1(x) \\ n_2(x) \\ n_3(x) \end{bmatrix} = \begin{bmatrix} n_1(x) - \frac{x_2}{2} n_3(x) \\ n_2(x) + \frac{x_1}{2} n_3(x) \\ \varepsilon n_3(x) \end{bmatrix} = \begin{bmatrix} \alpha_1(x) \\ \alpha_2(x) \\ \varepsilon \alpha_3(x) \end{bmatrix},$$

and so the approximated Riemannian normal can be written, due to the orthogonality of the vector fields X_1, X_2, εX_3

$$\begin{aligned} n_\varepsilon(x) &= \frac{pr_{\mathcal{H}_\varepsilon} n_E(x)}{|pr_{\mathcal{H}_\varepsilon} n_E(x)|_{g_\varepsilon}} \\ &= \frac{(n_1(x) - \frac{x_2}{2} n_3(x), n_2(x) + \frac{x_1}{2} n_3(x), \varepsilon n_3(x))^T}{\sqrt{(n_1(x) - \frac{x_2}{2} n_3(x))^2 + (n_2(x) + \frac{x_1}{2} n_3(x))^2 + (\varepsilon n_3(x))^2}}. \end{aligned}$$

Definition 4.18. *The* approximated Riemannian mean curvature *is defined as*

$$k_\varepsilon(x) = div_{\mathcal{H}^\varepsilon} n_\varepsilon(x),$$

where $x \in \Gamma$, $n_\varepsilon(x)$ the approximated Riemannian normal at $x \in \Gamma$ and $div_{\mathcal{H}^\varepsilon}$ is as in Definition 1.75 (i.e. the horizontal divergence is computed w.r.t. to family of vector fields $\mathcal{X}_\varepsilon$ which span $\mathcal{H}_\varepsilon$).

The previous two definition may be adapted for the level set equation as follows.

Definition 4.19. *Let $(\mathbb{R}^N, \mathcal{H}, g)$ and $\Gamma = \{x \in \mathbb{R}^N \mid u(x) = 0\}$ a C^1 hypersurface. Let us consider $(\mathbb{R}^N, \mathcal{H}_\varepsilon, g_\varepsilon)$ the Riemannian approximation of the sub-Riemannian geometry. Then the approximated Riemannian normal can be expressed as*

$$n_\varepsilon(x) = \left(\frac{X_1^\varepsilon u(x)}{\sqrt{\sum_{i=1}^N (X_i^\varepsilon u(x))^2}}, \dots, \frac{X_N^\varepsilon u(x)}{\sqrt{\sum_{i=1}^N (X_i^\varepsilon u(x))^2}} \right)^T.$$

Definition 4.20. *Let $(\mathbb{R}^N, \mathcal{H}, g)$ be a sub-Riemannian geometry, and $\Gamma = \{x \in \mathbb{R}^N \mid u(x) = 0\}$ a C^1 hypersurface and let us consider $(\mathbb{R}^N, \mathcal{H}_\varepsilon, g_\varepsilon)$ the Riemannian approximation of the sub-Riemannian geometry. We define the approximated Riemannian mean curvature as*

$$k_\varepsilon(x) = \sum_{i=1}^N X_i^\varepsilon \left(\frac{X_i^\varepsilon u(x)}{\sqrt{\sum_{i=1}^N (X_i^\varepsilon u(x))^2}} \right).$$

Example 4.21. *Let us consider $\Gamma = \{x \in \mathbb{R}^3 | u(x) = 0\}$. The approximated Riemannian normal for the 1-dimensional Heisenberg group is given by*

$$\begin{aligned} n_\varepsilon(x) &= \frac{\nabla_\varepsilon u(x)}{|\nabla_\varepsilon u(x)|_{g_\varepsilon}} \\ &= \frac{(u_{x_1}(x) - \frac{x_2}{2} u_{x_3}(x), u_{x_2}(x) + \frac{x_1}{2} u_{x_3}(x), \varepsilon u_{x_3}(x))^T}{\sqrt{(u_{x_1}(x)) - \frac{x_2}{2} u_{x_3}(x))^2 + (u_{x_2}(x) + \frac{x_1}{2} u_{x_3}(x))^2 + \varepsilon^2 u_{x_3}^2(x)}}. \end{aligned}$$

We give the definition of approximated Riemannian mean curvature flow.

Definition 4.22. *Let $\mathcal{H}^\varepsilon$ be a Riemannian approximation of $\mathcal{H}$ and g_ε the associated Riemannian metric, we say that Γ_t is an evolution by the Riemannian mean curvature flow of the hypersurface Γ_0 if and only if $\Gamma = \Gamma_0$ and for any smooth curve $x_\varepsilon : [0,T] \to \mathbb{R}^N$ such that $x_\varepsilon(t) \in \Gamma_t$ for all $t \in [0,T]$, the approximated Riemannian normal velocity is equal to minus the Riemannian curvature, i.e.*

$$v_\varepsilon(x_\varepsilon(t)) = g_{\varepsilon_{x_\varepsilon(t)}}(\dot{x}_\varepsilon(t), n_\varepsilon(x_\varepsilon(t))) = -k_\varepsilon(x_\varepsilon(t)),$$

where n_ε and k_ε are the Riemannian approximated external normal and the Riemannian approximated curvature of Γ_t and the tangent vector to curve is given by

$$\dot{x}_\varepsilon(t) := \sum_{i=1}^{N} \alpha_i^\varepsilon(t) X_i^\varepsilon(x_\varepsilon(t)) = \sigma_\varepsilon^T(x_\varepsilon(t))\alpha^\varepsilon(t) \quad \forall t \in [0,T],$$

where σ_ε is defined as (1.5).

In order to find heuristically the PDE associated to approximated mean curvature flow we will proceed as did in [22]. Let $\Gamma_t = \{u(x,t) = c\}$ be a level set where the function u is smooth, then the approximated horizontal normal of Γ_t) is given by

$$n_\varepsilon(x_\varepsilon(t)) = \frac{\sum_{i=1}^N (X_i^\varepsilon u(x_\varepsilon(t)))X_i^\varepsilon(x_\varepsilon(t))}{\sqrt{\sum_{i=1}^N (X_i^\varepsilon u(x_\varepsilon(t)))^2}}.$$

Since $x(t)$ is horizontal and smooth then we have to use the property that $\{X_i^\varepsilon\}_{i=1}^N$ is orthonormal w.r.t. g_ε

$$g_{\varepsilon_{x(t)}}\left(\dot{x_\varepsilon}(t), \frac{\sum_{i=1}^N (X_i^\varepsilon u)X_i^\varepsilon(x)}{\sqrt{\sum_{i=1}^N (X_i^\varepsilon u)^2}}\right) = \left\langle \alpha^\varepsilon(t), \frac{\mathcal{X}_\varepsilon u}{|\mathcal{X}_\varepsilon u|} \right\rangle_N,$$

where $\mathcal{X}_\varepsilon u = (X_1 u, \ldots, X_m u, X_{m+1}^\varepsilon u, \ldots, X_N^\varepsilon u)$ are the components of the approximated Riemannian gradient, then we have that the approximated

horizontal normal velocity can be written as

$$g_{\varepsilon_{x_\varepsilon(t)}}(\dot{x}_\varepsilon(t), n_\varepsilon(x_\varepsilon(t))) = \left\langle \alpha^\varepsilon(t), \frac{\mathcal{X}_\varepsilon u}{|\mathcal{X}_\varepsilon u|} \right\rangle_N$$
$$= \left\langle \alpha^\varepsilon(t), \frac{\sigma_\varepsilon(x)Du}{|\sigma_\varepsilon(x)Du|} \right\rangle_N = |\mathcal{X}_\varepsilon u|^{-1} \langle \dot{x}_\varepsilon(t), Du \rangle_N . \tag{4.4}$$

We observe now that $x_\varepsilon(t) \in \Gamma_t$ if and only if $u(t, x_\varepsilon(t)) = c$. Hence, we can take the derivative in time and then we have

$$u_t(t, x_\varepsilon(t)) = - < \dot{x}_\varepsilon(t), Du(t, x_\varepsilon(t)) >_N$$
$$= -|\mathcal{X}_\varepsilon u| < \dot{x}_\varepsilon(t), n_\varepsilon(t) >_{g_\varepsilon} = |\mathcal{X}_\varepsilon u| k_\varepsilon(x_\varepsilon(t)).$$

We recall that $k_\varepsilon(x_\varepsilon(t)) = \sum_{i=1}^N X_i^\varepsilon \left(\frac{X_i^\varepsilon u(x_\varepsilon(t))}{|\mathcal{X}_\varepsilon u|} \right)$ then we have, observing that

$$\sum_{i=1}^N X_i^\varepsilon \left(\frac{X_i^\varepsilon u)}{|\mathcal{X}_\varepsilon u|} \right) = \sum_{i=1}^N \frac{X_i^\varepsilon X_i^\varepsilon u |\mathcal{X}_\varepsilon u| - X_i^\varepsilon u \left(\frac{1}{2} \frac{\sum_{j=1}^N (X_i^\varepsilon (X_j^\varepsilon u) + X_j^\varepsilon (X_i^\varepsilon u)) X_i^\varepsilon u}{|\mathcal{X}_\varepsilon u|} \right)}{|\mathcal{X}_\varepsilon u|^2},$$

hence

$$k_\varepsilon = |\mathcal{X}_\varepsilon u|^{-1} \left(Tr((\mathcal{X}_\varepsilon^2 u)^*) - \left\langle (\mathcal{X}_\varepsilon^2 u)^* \frac{\mathcal{X}_\varepsilon u}{|\mathcal{X}_\varepsilon u|}, \frac{\mathcal{X}_\varepsilon u}{|\mathcal{X}_\varepsilon u|} \right\rangle \right),$$

where

$$(\mathcal{X}_\varepsilon^2 u)_{ij}^* = \frac{X_i^\varepsilon (X_j^\varepsilon u) + X_j^\varepsilon (X_i^\varepsilon u)}{2}. \tag{4.5}$$

and finally we have

$$u_t = Tr((\mathcal{X}_\varepsilon^2 u)^*) - \left\langle (\mathcal{X}_\varepsilon^2 u)^* \frac{\mathcal{X}_\varepsilon u}{|\mathcal{X}_\varepsilon u|}, \frac{\mathcal{X}_\varepsilon u}{|\mathcal{X}_\varepsilon u|} \right\rangle = \Delta_\varepsilon u - \Delta_{0,\infty,\varepsilon} u, \tag{4.6}$$

where Δ_ε is the Laplace-Beltrami operator of the approximated Riemannian geometry and $\Delta_{0,\infty,\varepsilon}$ the infinite approximated Riemannian Laplacian. We observe now that we can write the equation (4.6) as

$$u_t + F_\varepsilon(x, Du, D^2 u) = 0, \tag{4.7}$$

with

$$F_\varepsilon(x, p, S) = -Tr(\sigma_\varepsilon(x) S \sigma_\varepsilon^T(x) + A_\varepsilon(x, p))$$
$$+ \left\langle \left(\sigma_\varepsilon(x) S \sigma_\varepsilon^T(x) + A_\varepsilon(x, p) \right) \frac{\sigma_\varepsilon(x)p}{|\sigma_\varepsilon(x)p|}, \frac{\sigma_\varepsilon(x)p}{|\sigma_\varepsilon(x)p|} \right\rangle, \tag{4.8}$$

and we can take the equation found by Dirr, Dragoni and Von Renesse in [22]

$$F(x,p,S) = -Tr(\sigma(x)S\sigma^T(x) + A(x,p)) + \left\langle \left(\sigma(x)S\sigma^T(x) + A(x,p)\right)\frac{\sigma(x)p}{|\sigma(x)p|}, \frac{\sigma(x)p}{|\sigma(x)p|}\right\rangle,$$

where

$$A_\varepsilon(x,p) = \frac{1}{2} < \nabla_{X_i^\varepsilon} X_j^\varepsilon(x) + \nabla_{X_j^\varepsilon} X_i^\varepsilon(x), p >$$

and

$$A(x,p) = \frac{1}{2} < \nabla_{X_i} X_j(x) + \nabla_{X_j} X_i(x), p > .$$

We observe that the equation $F(x,p,S)$ is well defined and continuous outside the characteristic points. We define $\mathcal{V} = \{(x,p) \in \Gamma \times T_x\Gamma \mid \sigma(x)p = 0\}$ (see Definition 4.4). In this way we observe that the definition of F is

$$F : (\mathbb{R}^{2N} \setminus \mathcal{V}) \times Sym(N) \to \mathbb{R}$$

while the definition of F_ε, due to $det(\sigma_\varepsilon(x)) \neq 0$, is

$$F_\varepsilon : \mathbb{R}^N \times (\mathbb{R}^N \setminus \{0\}) \times Sym(N) \to \mathbb{R}.$$

We can compute the envelopes that are for F_ε

$$F_\varepsilon^*(x,p,S) = \begin{cases} -Tr(\overline{S}_\varepsilon) + \left\langle \overline{S}_\varepsilon \frac{\sigma_\varepsilon(x)p}{|\sigma_\varepsilon(x)p|}, \frac{\sigma_\varepsilon(x)p}{|\sigma_\varepsilon(x)p|}\right\rangle, & |p| \neq 0, \\ -Tr(\overline{S}_\varepsilon) + \lambda_{max}(\overline{S}_\varepsilon), & |p| = 0, \end{cases}$$

and

$$F_{\varepsilon *}(x,p,S) = \begin{cases} -Tr(\overline{S}_\varepsilon) + \left\langle \overline{S}_\varepsilon \frac{\sigma_\varepsilon(x)p}{|\sigma_\varepsilon(x)p|}, \frac{\sigma_\varepsilon(x)p}{|\sigma_\varepsilon(x)p|}\right\rangle, & |p| \neq 0, \\ -Tr(\overline{S}_\varepsilon) + \lambda_{min}(\overline{S}_\varepsilon), & |p| = 0, \end{cases}$$

where $\overline{S}_\varepsilon = \sigma_\varepsilon(x)S\sigma_\varepsilon^T(x) + A_\varepsilon(x,p)$ with λ_{max} and λ_{min} the maximum and the minimum eigenvalues of the matrix $\overline{S_\varepsilon}$ and

$$F^*(x,p,S) = \begin{cases} -Tr(\overline{S}) + \left\langle \overline{S} \frac{\sigma(x)p}{|\sigma(x)p|}, \frac{\sigma(x)p}{|\sigma(x)p|}\right\rangle, & |\sigma(x)p| \neq 0, \\ -Tr(\overline{S}) + \lambda_{max}(\overline{S}), & |\sigma(x)p| = 0, \end{cases}$$

and

$$F_*(x,p,S) = \begin{cases} -Tr(\overline{S}) + \left\langle \overline{S}\frac{\sigma(x)p}{|\sigma(x)p|}, \frac{\sigma(x)p}{|\sigma(x)p|} \right\rangle, & |\sigma(x)p| \neq 0, \\ -Tr(\overline{S}) + \lambda_{min}(\overline{S}), & |\sigma(x)p| = 0, \end{cases}$$

where $\overline{S} = \sigma(x)S\sigma^T(x) + A(x,p)$ with λ_{max} and λ_{min} the maximum and the minimum eigenvalues of the matrix $\overline{S}$. In order to compute upper/lower we use the remarks made in [27], proof Theorem 1.1.

Remark 4.23. *We observe that $|\sigma_\varepsilon(x)p| > |\sigma(x)p|$. We can write them explicitly as*

$$|\sigma_\varepsilon(x)p| = \sqrt{\sum_{i=1}^{N}(A_i^\varepsilon(x)p)^2},$$

$$|\sigma(x)p| = \sqrt{\sum_{i=1}^{m}(A_i(x)p)^2}$$

where A_i are the row of the matrices and then the thesis is trivial.

Taking $\overline{S}_\varepsilon = (\mathcal{X}_\varepsilon^2 u)^*$ and $\sigma_\varepsilon(x)p = \mathcal{X}_\varepsilon u$ we can extend the definition of viscosity solution for approximated horizontal mean curvature flow.

Definition 4.24. *Let $\Gamma_0 = \{x \in \mathbb{R}^N \mid u_0^\varepsilon(x) = 0\}$ be hypersurface in $\mathbb{R}^N$. We say that $\Gamma^\varepsilon(t) = \{x \in \mathbb{R}^N \mid u^\varepsilon(t,x) = 0\}$ is a generalized evolution by approximated horizontal mean curvature flow if $u^\varepsilon(t,x)$ satisfies the initial condition $u^\varepsilon(0,x) = u_0^\varepsilon(x)$ and it is a viscosity solution of the approximated horizontal mean curvature flow in the sense of Giga, that means u^ε is a continuous function and*

- *for any $\phi \in C^2((0,+\infty)\times\mathbb{R}^N)$ such that $u^\varepsilon - \phi$ has a local minimum at (t_0,x_0) then*

$$\begin{cases} \phi_t - \Delta_\varepsilon\phi + \Delta_{0,\infty,\varepsilon}\phi \geq 0, & \text{at } (t_0,x_0) \text{ if } D\phi(t_0,x_0) \neq 0, \\ \phi_t - \Delta_\varepsilon\phi + \lambda_{max}((\mathcal{X}_\varepsilon^2\phi)^*) \geq 0, & \text{at } (t_0,x_0) \text{ if } D\phi(t_0,x_0) = 0 \end{cases}.$$

- *for any* $\phi \in C^2((0,+\infty)\times\mathbb{R}^N)$ *such that* $u^\varepsilon - \phi$ *has a local maximum at* (t_0,x_0) *then*

$$\begin{cases} \phi_t - \Delta_\varepsilon\phi + \Delta_{0,\infty,\varepsilon}\phi \le 0, & \text{at } (t_0,x_0) \text{ if } D\phi(t_0,x_0)\neq 0, \\ \phi_t - \Delta_\varepsilon\phi + \lambda_{min}((\mathcal{X}^2_\varepsilon\phi)^*) \le 0, & \text{at } (t_0,x_0) \text{ if } D\phi(t_0,x_0)= 0. \end{cases}$$

4.3 Remarks on $\mathbb{H}^1$

In this subsection we introduce some properties of the Heisenberg group. These remarks allow us to simplify the problem of the evolution by horizontal and approximated Riemannian mean curvature flow.

Remark 4.25. *We observe that the drift term in the approximated Heisenberg group is equal to zero, in fact we can take as vector fields which give the distribution* $X_1^\varepsilon(x) = (1,0,-\frac{x_2}{2})^T$, $X_2^\varepsilon(x) = (0,1,\frac{x_1}{2})^T$, $X_3^\varepsilon(x) = (0,0,\varepsilon)^T$, *where* $x=(x_1,x_2,x_3)\in\mathbb{R}^3$, *and then we observe that*

$$X_1^\varepsilon X_1^\varepsilon(x) = X_2^\varepsilon X_2^\varepsilon(x) = X_3^\varepsilon X_1^\varepsilon(x) = \\ X_1^\varepsilon X_3^\varepsilon(x) = X_2^\varepsilon X_3^\varepsilon(x) = X_3^\varepsilon X_2^\varepsilon(x) = X_3^\varepsilon X_3^\varepsilon(x) = 0$$

and

$$X_1X_2(x) = -\frac{1}{2}\frac{\partial}{\partial x_3}, \quad X_2X_1(x) = \frac{1}{2}\frac{\partial}{\partial x_3}.$$

Hence the drift term is $\sum_{i,j=1}^3 \nabla_{X_i}X_j = -\frac{1}{2}\frac{\partial}{\partial x_3} + \frac{1}{2}\frac{\partial}{\partial x_3} = 0$, *in this way we can observe that the Stratonovich integral coincides with the Itô one.*

We make another interesting remark about the fact that σ and σ_ε are globally Lipschitz.

Lemma 4.26. *The matrices* σ_ε *and* σ *are globally Lipschitz, in particular the Lipschitz constant is* $\frac{1}{2}$.

Proof. Recalling the matrix norm $\|A\| = \sum_{i=1}^{3}\sum_{j=1}^{3}|a_{ij}|$, we obtain by computation

$$\begin{aligned}\|\sigma_\varepsilon(x) - \sigma_\varepsilon(x')\| &= \left|\left|\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{x_2}{2} & \frac{x_1}{2} & \varepsilon \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{x_2'}{2} & \frac{x_1'}{2} & \varepsilon \end{bmatrix}\right|\right| \\ &\leq \frac{1}{2}(|x_1 - x_1'| + |x_2 - x_2'| + |x_3 - x_3'|)\end{aligned}$$

and

$$\begin{aligned}\|\sigma(x) - \sigma(x')\| &= \left|\left|\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{x_2}{2} & \frac{x_1}{2} \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{x_2'}{2} & \frac{x_1'}{2} \end{bmatrix}\right|\right| \\ &\leq \frac{1}{2}(|x_1 - x_1'| + |x_2 - x_2'| + |x_3 - x_3'|).\end{aligned}$$

□

Unfortunately, the matrices σ and σ_ε are not globally bounded but only locally.

Lemma 4.27. *Let $x \in B_R(0) \subset \mathbb{R}^3$ where $R > 0$, then $\sigma(x)$ and $\sigma_\varepsilon(x)$ are locally bounded.*

Proof. Recalling the matrix norm $\|A\| = \sum_{i=1}^{3}\sum_{j=1}^{3}|a_{ij}|$ and that $x \in B_R(0)$ then we have

$$|x_i| \leq R, \qquad i = 1, 2, 3$$

and so we have

$$\|\sigma(x)\| = \left|\left|\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{x_2}{2} & \frac{x_1}{2} \end{bmatrix}\right|\right| \leq 2 + R$$

and

$$\|\sigma(x)\| = \left|\left|\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{x_2}{2} & \frac{x_1}{2} & \varepsilon \end{bmatrix}\right|\right| \leq 2 + R + \varepsilon.$$

□

It is possible to show that compact surfaces in the Heisenberg group have at least one characteristic points (see [22]), as, for example, in the case of the Euclidean sphere which has two points.

Example 4.28. *The classical Euclidean sphere has two points which are characteristic. In order to find them we recall that,* $S^2 := \{(x_1, x_2, x_3) \in \mathbb{H}^1 \mid u(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 1 = 0\}$*, then*

$$\begin{aligned} char(\Sigma) &= \left\{ x = (x_1, x_2, x_3) \in S^2 \;\middle|\; \left(u_{x_1}(x) - \frac{x_2}{2} u_{x_3}(x)\right)^2 + \left(u_{x_2}(x) + \frac{x_1}{2} u_{x_3}(x)\right)^2 = 0 \right\} \\ &= \left\{ (x_1, x_2, x_3) \in S^2 \;\middle|\; (x_1^2 + x_2^2)(4 + x_3^2) = 0 \right\}. \end{aligned}$$

Hence the characteristic points in this case are $(0, 0, 1)$ *and* $(0, 0, -1)$*, i.e. the north and south pole.*

Remark 4.29. *We observe that the Euclidean sphere is a regular surface in Euclidean setting, in fact called* $S^2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1\}$ *the surface we have that the gradient is*

$$\nabla u = (2x_1, 2x_2, 2x_3)$$

which cannot be equal to zero because $(0, 0, 0)$ *does not belong to the sphere.*

Research part

The next chapters are the research part of this thesis. The choice of maintaining the chapters self-contained is in order to simplify the reading of the results obtained. However, there are some references to the previous chapters, whenever this may be useful.
We recall briefly the structure of this second part of the thesis:

- In Chapter 5, we find the stochastic representation of the solution of the approximated Riemannian mean curvature flow, see preprint [35].
- In Chapter 6 we find the asymptotic behaviour for large p of the p-optimal controls for the stochastic dynamics associated to the Heisenberg group, see [21] (published paper).
- In Chapter 7 we find the asymptotic behaviour for large p of the p-optimal controls for the stochastic dynamics associated to the Riemannian approximated Heisenberg group, see preprint [25].
- In Chapter 8 we obtain a limsup estimate for the function V^ε, with respect to the function V, and we give some ideas on how we intend to approach the corresponding liminf estimate, necessary to conclude the convergence. This is a work in progress.

Chapter 5

Stochastic representation of Approximated MCF

5.1 Introduction

The evolution by mean curvature flow (MCF) has been studied extensively and it has many applications in image processing and neurogeometry (see e.g. [13]). We say that a hypersurface evolves by MCF if it contracts in the normal direction with normal velocity proportional to its mean curvature see e.g. [28] for further details. It is well-known that this evolution may develop singularities in finite time in the Euclidean and Riemannian setting (as in the case of the dumbbell, see [28] for further details). To deal with such a singularities, many generalised approaches to study this evolution have been developed. In particular in 1991, Chen, Giga and Goto [11] and, independently Evans and Spruck [31] introduced the so called level set approach, which consists in studying the evolving hypersurfaces as level sets of (viscosity) solutions of suitable associated nonlinear PDEs. In this paper we are interested in a degenerate version of such an evolution, namely evolution by horizontal mean curvature flow (HMCF) and its Riemannian approximation: we consider a hypersurface embedded in a sub-Riemannian structure (Carnot-type geometry),

then the evolution contracts in the direction of the so called horizontal normal proportionally to its horizontal curvature (see Section 2 for details). We consider the level set approach which is now associated to a parabolic PDE far more degenerate than in the standard case.

The following stochastic approach was developed by Cardaliaguet, Quincampoix and Buckdahn in [27] and contemporaneously but independently by Soner and Touzi [49] for the standard (Euclidean) case and generalised then by Dirr, Dragoni and von Renesse in [22] for the case of HMCF, that there is a connection between this equation and a suitable stochastic optimal control problem. In the Euclidean setting the dynamic can be expressed using the definition of the Itô integral while in the sub-Riemannian case we have to use the definition of the Stratonovich integral. Roughly speaking, in the last case the dynamic is far more complex because we have a deterministic part (related to first order derivatives induced by the chosen geometry) and a stochastic one (related to some second order derivatives induced by the chosen geometry). However, as in the case of the Heisenberg group, sometimes it is possible to find some simplification of this dynamic as remarked in [21], making the dynamic similar to the Euclidean one. It is well known that it is possible to generalize this equation using a Riemannian approximation, as e.g. in [13].

The aim of the chapter is to find a stochastic representation of the viscosity solution of approximated Riemannian mean curvature flow, generalizing the result obtained by [22].

5.2 Preliminaries

We now briefly recall some basic geometrical definitions which will be key for defining the evolution by HMCF (for further details see Chapter 1). For more definitions and properties on sub-Riemannian geometries we refer to [42]

and also [4] for the particular case of Carnot groups.

Definition 5.1. *Let M be a N-dimensional smooth manifold, we can define for every point p a subspace of T_xM called $\mathcal{H}_x$. We define the* distribution *as $\mathcal{H} = \{(x, v)|\ p \in M,\ v \in \mathcal{H}_x\}$.*

Definition 5.2. *Let M be a manifold and X,Y two vector fields defined on this manifold and $f : M \to \mathbb{R}$ a smooth function, then we define the* (Lie) bracket *between X and Y as $[X, Y](f) = XY(f) - YX(f)$.*
Let us consider $\mathcal{X} = \{X_1, \dots, X_m\}$ spanning some distribution $\mathcal{H} \subset TM$, we define the k-bracket as $\mathcal{L}^{(k)} = \{[X, Y]|X \in \mathcal{L}^{(k-1)}\ \ Y \in \mathcal{L}^{(1)}\}$ with $i_j \in \{1, \dots, m\}$ and $\mathcal{L}^{(1)} = \mathcal{X}$. The associated Lie algebra is the set of all brackets between the vector fields of the family

$$\mathcal{L}(\mathcal{X}) := \{[X_i, X_j^{(k)}]|X_j^{(k)}\ \textit{k-length bracket of } X_1, \dots X_m\ k \in \mathbb{N}\}.$$

The definition of Hörmander condition is crucial in order to work with PDEs in sub-Riemannian setting, because it allows us to recover the whole tangent space for every point.

Definition 5.3 (Hörmander condition)**.** *Let M be a smooth manifold and $\mathcal{H}$ a distribution defined on M. We say that the distribution is* bracket generating *if and only if, at any point, the Lie algebra $\mathcal{L}(\mathcal{X})$ spans the whole tangent space. We say that a sub-Riemannian geometry satisfies the* Hörmander condition *if and only if the associated distribution is bracket generating.*

Definition 5.4. *Let M be a smooth manifold and $\mathcal{H} = span\{X_1, \dots, X_m\} \subset TM$ a distribution and g a Riemannian metric of M defined on the subbundle $\mathcal{H}$. A* sub-Riemannian geometry *is the triple $(M, \mathcal{H}, g)$.*

Definition 5.5. *Let $(M, \mathcal{H}, g)$ be a sub-Riemannian geometry and $\gamma : [0, T] \to M$ an absolutely continuous curve, we say that γ is an* horizontal curve *if and only if*

$$\dot{\gamma}(t) \in \mathcal{H}_{\gamma(t)}, \quad \textit{for a.e. } t \in [0, T],$$

or, equivalently, if there exists a measurable function $h : [0,T] \to \mathbb{R}^N$ *such that*

$$\dot{\gamma}(t) = \sum_{i=1}^{m} h_i(t) X_i(\gamma(t)), \quad \text{for a.e. } t \in [0,T],$$

where $h(t) = (h_1(t), \dots, h_m(t))$ *and* $X_1, \dots X_m$ *are some vector fields spanning the distribution* $\mathcal{H}$*.*

Example 5.6 (The Heisenberg group)**.** *The most significant sub-Riemannian geometry is the so called Heisenberg group. For a formal definition of the Heisenberg group and the connection between its structure as non commutative Lie group and its manifold structure we refer to [4]. Here we simply introduce the 1-dimensional Heisenberg group as the sub-Riemannian structure induced on* $\mathbb{R}^3$ *by the vector fields*

$$X_1(x) = \begin{pmatrix} 1 \\ 0 \\ -\frac{x_2}{2} \end{pmatrix} \quad \text{and} \quad X_2 = \begin{pmatrix} 0 \\ 1 \\ \frac{x_1}{2} \end{pmatrix}, \quad \forall\, x = (x_1, x_2, x_3) \in \mathbb{R}^3.$$

We observe that the associated matrix is given by

$$\sigma(x) = \begin{bmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \end{bmatrix}. \tag{5.1}$$

The introduced vector fields satisfy the Hörmander condition with step 2: in fact $[X_1, X_2](x) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ *for any* $x \in \mathbb{R}^3$*.*

The Hörmander condition is crucial to state the following theorem.

Theorem 5.7. *([12])[Chow] Let* M *be a smooth manifold and* $\mathcal{H}$ *a bracket generating distribution defined on* M*. If* M *is connected, then there exists a horizontal curve joining any two given points of* M*.*

Carnot type geometries

From now on we consider only the case where the starting topological manifold M is the Euclidean $\mathbb{R}^N$. Moreover, in this paper we will concentrate on sub-Riemannian geometries with a particular structure: the so called Carnot-type geometries.

Definition 5.8. *Let us consider $(M, \mathcal{H}, g)$ a sub-Riemannian geometry. We say that $X_1, \dots, X_m$, $m < N$, are Carnot-type vector fields if the coefficients of X_i are 0 for $j \in \{1, \dots, m\} \setminus \{i\}$, the i-component is equal to 1 and the other $m - N$ components are polynomial in x.*

The previous structure allows us to consider an easy and explicit Riemannian approximation. Nevertheless the approach apply also to the case where this additional structure is not fulfilled. This structure applies to a large class of geometries. The Heisenberg group introduced in Example 5.6 is obviously a Carnot-type geometry. Carnot groups (see [4] for definitions and properties) are a very important class of sub-Riemannian geometries with in addition a non commutative Lie group structure associated.

For later use we also introduce the matrix associated to the vector fields $X_1, \dots, X_m$, which is the $N \times m$ matrix defined as

$$\sigma(x) = [X_1(x), \dots, X_m(x)]^T.$$

Example 5.9. *In the case of the Heisenberg group introduced in Example 5.6, the matrix σ is given by*

$$\sigma(x) = \begin{bmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \end{bmatrix}, \quad \forall\, x = (x_1, x_2, x_3) \in \mathbb{R}^3.$$

In general, for Carnot-type geometries, the matrix σ assumes the following structure:

$$\sigma(x) = \begin{bmatrix} I_{m \times m} & A(x_1, \dots x_m) \end{bmatrix} \tag{5.2}$$

where the matrix $A(x_1, \dots, x_m)$ is a $(N-m) \times m$ matrix depending only on the first m components of x.

We now want to introduce the Riemannian approximation, which will be crucial for our results.

Let us consider a distribution $\mathcal{H}$ spanned by the Carnot-type vector fields $\{X_1, \dots, X_m\}$ defined on $\mathbb{R}^N$ with $m < N$ and satisfying the Hörmander condition. It is possible to complete the distribution $\mathcal{H}$ by adding $N-m$ vector fields $X_{m+1}, \dots, X_N$ in order to construct an orthonormal basis for all $x \in \mathbb{R}^N$, i.e.

$$\text{Span}\big(X_1(x), \dots, X_m(x), X_{m+1}(x), \dots, X_N(x)\big) = T_x\mathbb{R}^N \equiv \mathbb{R}^N, \ \forall\, x \in \mathbb{R}^N.$$

The geometry induced, for all $\varepsilon > 0$, by the distribution

$$\mathcal{H}_\varepsilon(x) = span\{X_1(x), \dots, X_m(x), \varepsilon X_{m+1}(x), \dots, \varepsilon X_N(x)\}, \quad \forall x \in \mathbb{R}^N$$

is called *Riemannian approximation* of our starting sub-Riemannian topology. The associated matrix is now

$$\sigma_\varepsilon(x) = [X_1(x), \dots X_m(x), \varepsilon X_{m+1}(x) \dots, \varepsilon X_N(x)]^T. \tag{5.3}$$

Note that $\det(\sigma_\varepsilon(x)) \neq 0$.

Note that, in the case of Carnot-type geometries, we can always choose

$$X_i(x) = e_i, \quad \forall i = m+1, \dots, N \quad \forall x \in \mathbb{R}^N,$$

where by e_i we indicate the standard Euclidean unit vector with 1 at the i-th component.

Example 5.10 (Riemannian approximation of $\mathbb{H}^1$)**.** *In the case of the Heisenberg group introduced in Example 5.6, the matrix associated to the Riemannian approximation is for every point $x = (x_1, x_2, x_3)$ given by*

$$\sigma_\varepsilon(x) = \begin{bmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \\ 0 & 0 & \varepsilon \end{bmatrix}.$$

This technique is called Riemannian approximation since, as $\varepsilon \to 0^+$, then the geometry induced by Riemannian approximation converges, in sense of Gromov-Hausdorff (see [36] for further details), to the original sub-Riemannian geometry (as shown, as example, in [13]).

5.3 Horizontal mean curvature evolution

Given a smooth hypersurface Γ, we indicate by $n_E(x)$ the standard (Euclidean) normal to the hypersurface Γ at the point x. Since the family of vector fields $\mathcal{X}_\varepsilon = \{X_1, \dots, X_m, \varepsilon X_{m+1}, \dots, \varepsilon X_N\}$ span the whole of $\mathbb{R}^N$ at any point of Γ, then $n_E(x)$ can be written w.r.t. such a basis, i.e.

$$n_E(x) = \frac{\sum_{i=1}^N \alpha_i(x) X_i^\varepsilon(x)}{\sqrt{\sum_{i=1}^N \alpha_i^2(x)}}.$$

where X_i^ε are the elements of $\mathcal{X}_\varepsilon$. The following definitions will be key for this paper (see Chapter 4 for further details).

Definition 5.11. *Given a smooth hypersurface Γ, the* horizontal normal *is the renormalized projection of the Euclidean normal on the horizontal space $\mathcal{H}_x$, i.e.*

$$n_0(x) := \frac{\alpha_1(x)X_1(x) + \dots \alpha_m(x)X_m(x)}{\sqrt{\alpha_1^2(x) + \dots + \alpha_m^2(x)}} \in \mathcal{H}_x \subset \mathbb{R}^N.$$

With an abuse of notation we will often indicate by $n_0(x)$ the associated m-valued vector

$$n_0(x) = \frac{(\alpha_1(x), \dots, \alpha_m(x))}{\sqrt{\alpha_1^2(x) + \dots + \alpha_m^2(x)}} \in \mathbb{R}^m. \tag{5.4}$$

The main difference between the horizontal normal and a standard normal is that the first may not exist even for smooth hypersurfaces. In fact at some points the horizontal normal is not defined while the Euclidean one exists. These points are called *characteristic points*.

Definition 5.12. *Given a smooth hypersurface* Γ*, characteristic points occur whenever* $n_E(x)$ *is orthogonal to the horizontal plane* $\mathcal{H}_x$*, then its projection on such a subspace vanishes, i.e.*

$$\alpha_1^2(x) + \cdots + \alpha_m^2(x) = 0.$$

Note that these points do not exist in the associated Riemannian approximation, in fact whenever Γ is smooth the normal is defined at any point, which means

$$\sum_{i=1}^{N} \alpha_i^2(x) \neq 0, \quad \forall x \in \Gamma.$$

We recall that, for every smooth hypersurface, the *mean curvature* at the point $x \in \Gamma$ is defined as the Euclidean divergence of the Euclidean normal at that point. Similarly, for every smooth hypersurface, we can now introduce the horizontal mean curvature.

Definition 5.13. *Given a smooth hypersurface* Γ *and a non characteristic point* $x \in \Gamma$*, the* horizontal mean curvature *is defined as the horizontal divergence of the horizontal normal, i.e.* $k_0(x) = div_{\mathcal{H}} n_0(x)$*, where* $n_0(x)$ *is the* m*-valued vector associated to the horizontal normal (see* (7.3)*) while* $div_{\mathcal{H}}$ *is the divergence w.r.t. the vector fields* $X_1, \ldots, X_m$*, i.e.*

$$k_0(x) = X_1 \left(\frac{\alpha_1(x)}{\sqrt{\sum_{i=1}^m \alpha_i^2(x)}} \right) + \cdots + X_m \left(\frac{\alpha_m(x)}{\sqrt{\sum_{i=1}^m \alpha_i^2(x)}} \right).$$

Obviously the horizontal mean curvature is never defined at characteristic points, since there the horizontal normal does not exist.

Definition 5.14. *Let* Γ_t *be a family of smooth hypersurfaces in* $\mathbb{R}^N$*. We say that* Γ_t *is an* evolution by horizontal mean curvature flow *of* Γ *if and only if* $\Gamma_0 = \Gamma$ *and for any smooth horizontal curve* $\gamma : [0, T] \to \mathbb{R}^N$ *such that* $\gamma(t) \in \Gamma_t$ *for all* $t \in [0, T]$*, the horizontal normal velocity* v_0 *is equal to minus the horizontal mean curvature, i.e.*

$$v_0(\gamma(t)) := -k_0(\gamma(t)) n_0(\gamma(t)), \tag{5.5}$$

where $n_0(\gamma(t))$ and $k_0(\gamma(t))$ as respectively the horizontal normal and the horizontal mean curvature defined by Definitions 5.11 and 5.13 at the point $\gamma(t)$.

Note that Definition 5.14 is never defined at characteristic points.

In this subsection we consider a level set manifold Γ which is smooth. We now compute the horizontal normal and the horizontal curvature for smooth hypersurface expressed as zero level set, i.e.

$$\Gamma = \left\{x \in \mathbb{R}^N | u(x) = 0\right\},$$

for some smooth function $u : \mathbb{R}^N \to \mathbb{R}$. Then the Euclidean normal is simply $n_E(x) = \frac{\nabla u(x)}{|\nabla u(x)|}$, which implies that the horizontal normal can be expressed as

$$n_0(x) = \left(\frac{X_1 u(x)}{\sqrt{\sum_{i=1}^m (X_i u(x))^2}}, \ldots, \frac{X_m u(x)}{\sqrt{\sum_{i=1}^m (X_i u(x))^2}} \right). \tag{5.6}$$

Note that $(X_1 u, \ldots, X_m u) \in \mathbb{R}^m$ is the so called horizontal gradient.
Similarly we can then write the horizontal mean curvature as

$$k_0(x) = \sum_{i=1}^m X_i \left(\frac{X_i u(x)}{\sqrt{\sum_{i=1}^m (X_i u(x))^2}} \right). \tag{5.7}$$

Let $\Gamma_t = \{(x,t) | u(x,t) = 0\}$ where u is C^2. Applying (5.6) and (5.7) to the Definition 5.14 we obtain that u solves the following PDE, which is

$$u_t = Tr((\mathcal{X}^2 u)^*) - \left\langle (\mathcal{X}^2 u)^* \frac{\mathcal{X} u}{|\mathcal{X} u|}, \frac{\mathcal{X} u}{|\mathcal{X} u|} \right\rangle \tag{5.8}$$

where $\mathcal{X} u$ is the so called horizontal gradient, that is

$$\mathcal{X} u := (X_1 u, \ldots, X_m u)^T$$

and $(\mathcal{X}^2 u)^*$ is the symmetric horizontal Hessian, that is

$$((\mathcal{X}^2 u)^*)_{ij} := \frac{X_i(X_j u) + X_j(X_i u)}{2}.$$

This equation follows from the definition of horizontal mean curvature flow and that the horizontal normal can be expressed as

$$\sum_{i=1}^{N} X_i\left(\frac{X_i u}{|\mathcal{X}u|}\right)=\sum_{i=1}^{N}\frac{X_iX_iu|\mathcal{X}u|-X_iu\big(\frac{1}{2}\frac{\sum_{j=1}^{N}(X_i(X_ju)+X_j(X_iu))X_iu}{|\mathcal{X}u|}\big)}{|\mathcal{X}u|^2}$$

hence

$$k_\varepsilon=|\mathcal{X}u|^{-1}\left(Tr((\mathcal{X}^2u)^*)-\left\langle(\mathcal{X}^2u)^*\frac{\mathcal{X}u}{|\mathcal{X}u|},\frac{\mathcal{X}u}{|\mathcal{X}u|}\right\rangle\right).$$

We consider the equation found by Dirr, Dragoni and Von Renesse in [22], i.e.

$$u_t+F(x,Du,D^2u)=0$$

where

$$\begin{aligned}F(x,p,S)=&-Tr(\sigma(x)S\sigma^T(x)+A(x,p))\\&+\left\langle\big(\sigma(x)S\sigma^T(x)+A(x,p)\big)\frac{\sigma(x)p}{|\sigma(x)p|},\frac{\sigma(x)p}{|\sigma(x)p|}\right\rangle\end{aligned}\tag{5.9}$$

where

$$A(x,p)=\frac{1}{2}<\nabla_{X_i}X_j(x)+\nabla_{X_j}X_i(x),p>.$$

We observe that the equation $F(x,p,S)$ is well defined and continuous outside the characteristic points and we define $\mathcal{V}=\{(x,p)\in\Gamma\times T_x\Gamma|\ \sigma(x)p=0\}$. In this way we observe that

$$F:(\mathbb{R}^{2N}\setminus\mathcal{V})\times Sym(N)\to\mathbb{R}.$$

We remark that the function F has some points in which is discontinuous, As consequence, in order to work with viscosity solutions, we have to compute the upper and lower envelops of this function (for further details about envelopes see Chapter 3).

Definition 5.15. *Let us consider a locally bounded function* $u:\mathbb{R}\times[0,T]\to\mathbb{R}$.

- *The* upper semicontinuous envelope *is defined as*

$$\begin{aligned}u^*(t,x):=&\inf\{v(t,x)|\ v\ \textit{cont. and}\ \ v\geq u\}\\=&\limsup_{r\to0^+}\{u(s,y)||y-x|\leq r,|t-s|\leq r\}.\end{aligned}$$

- *The* lower semicontinuous envelope *is defined as*

$$\begin{aligned} u_*(t,x) &:= \sup\{u(t,x)|\ u \text{ cont. and } v \leq u\} \\ &= \liminf_{r\to 0^+}\{u(s,y)||y-x| \leq r, |t-s| \leq r\}. \end{aligned}$$

Remark 5.16. *If the function* $u : \mathbb{R}^N \times [0,T] \to \mathbb{R}$ *is continuous then it holds true*

$$u_*(t,x) = u(t,x) = u^*(t,x), \quad \text{for all} \quad (t,x) \in [0,T] \times \mathbb{R}^N.$$

Remark 5.17. *Applying the Definition 5.15 to the function* F *as defined in* (5.9) *we obtain*

$$F^*(x,p,S) = \begin{cases} -Tr(\overline{S}) + \left\langle \overline{S}\frac{\sigma(x)p}{|\sigma(x)p|}, \frac{\sigma(x)p}{|\sigma(x)p|} \right\rangle, & |\sigma(x)p| \neq 0, \\ -Tr(\overline{S}) + \lambda_{max}(\overline{S}), & |\sigma(x)p| = 0, \end{cases}$$

and

$$F_*(x,p,S) = \begin{cases} -Tr(\overline{S}) + \left\langle \overline{S}\frac{\sigma(x)p}{|\sigma(x)p|}, \frac{\sigma(x)p}{|\sigma(x)p|} \right\rangle, & |\sigma(x)p| \neq 0, \\ -Tr(\overline{S}) + \lambda_{min}(\overline{S}), & |\sigma(x)p| = 0, \end{cases}$$

where $\overline{S} = \sigma(x)S\sigma^T(x) + A(x,p)$ *with* λ_{max} *and* λ_{min} *the maximum and the minimum eigenvalues of the matrix* $\overline{S}$*. In order to compute upper/lower we use the remarks made in [27], proof Theorem 1.1.*

5.4 Approximated Riemannian mean curvature flow

The Equation (5.8) can be approximated to a Riemannian mean curvature flow using the Riemannian approximation (as seen in Chapter 1). This leads the following generalizations of the definitions of horizontal normal and horizontal divergence (see Chapter 4 for further details).

Definition 5.18. *Given a smooth hypersurface* Γ*, the* approximated Riemannian normal *is the renormalized projection of the Euclidean normal on the*

horizontal space $\mathcal{H}^\varepsilon_x$*, i.e.*

$$n_\varepsilon(x) := \frac{\sum_{i=1}^m \alpha_i(x)X_i(x) + \varepsilon\sum_{i=m+1}^N \alpha_i(x)X_i(x)}{\sqrt{\alpha_1^2(x)+\cdots+\alpha_m^2(x)+\varepsilon^2\alpha_{m+1}^2(x)+\cdots+\varepsilon^2\alpha_N^2(x)}} \in \mathcal{H}_x \subset \mathbb{R}^N.$$

With an abuse of notation, we will often indicate by $n_\varepsilon(x)$ *the associated* N*-valued vector*

$$n_\varepsilon(x) = \frac{(\alpha_1(x),\dots,\alpha_m(x),\varepsilon\alpha_{m+1}(x),\dots,\varepsilon\alpha_N(x))^T}{\sqrt{\alpha_1^2(x)+\cdots+\alpha_m^2(x)+\varepsilon^2\alpha_{m+1}^2(x)+\cdots+\varepsilon^2\alpha_N^2(x)}} \in \mathbb{R}^N. \quad (5.10)$$

Definition 5.19. *Given a smooth hypersurface* Γ *and a point* $x \in \Gamma$*, the* approximated Riemannian mean curvature *is defined as the horizontal divergence of the approximated Riemannian normal, i.e.* $k_\varepsilon(x) = div_{\mathcal{H}^\varepsilon} n_\varepsilon(x)$*, where* $n_\varepsilon(x)$ *is the* N*-valued vector associated to the horizontal normal (see* (5.10)*) while* $div_{\mathcal{H}}$ *is the divergence w.r.t. the vector fields* $X_1,\dots,X_m,\varepsilon X_{m+1},\dots,\varepsilon X_N$*, i.e.*

$$\begin{aligned} k_\varepsilon(x) = &\sum_{i=1}^m X_i\left(\frac{\alpha_i(x)}{\sqrt{\sum_{j=1}^m \alpha_j^2(x)+\varepsilon^2\sum_{k=m+1}^N \alpha_k^2(x)}}\right) \\ &+\varepsilon\sum_{i=m+1}^N X_i\left(\frac{\varepsilon\alpha_i(x)}{\sqrt{\sum_{j=1}^m \alpha_j^2(x)+\varepsilon^2\sum_{k=m+1}^N \alpha_k^2(x)}}\right). \end{aligned} \quad (5.11)$$

Remark 5.20. *In this setting we do not have characteristic points on the hypersurface* Γ*.*

We define now the approximated Riemannian mean curvature flow.

Definition 5.21. *Let* Γ_t *be a family of smooth hypersurfaces in* $\mathbb{R}^N$*. We say that* Γ_t *is an* evolution by approximated Riemannian mean curvature flow *of* Γ *if and only if* $\Gamma_0 = \Gamma$ *and for any smooth horizontal curve* $\gamma_\varepsilon : [0,T] \to \mathbb{R}^N$ *such that* $\gamma(t) \in \Gamma_t$ *for all* $t \in [0,T]$*, the horizontal normal velocity* v_ε *is equal to minus the horizontal mean curvature, i.e.*

$$v_\varepsilon(\gamma(t)) := -k_\varepsilon(\gamma_\varepsilon(t))n_\varepsilon(\gamma_\varepsilon(t)),$$

where $n_\varepsilon(x(t))$ *and* $k_\varepsilon(x(t))$ *as respectively the horizontal normal and the horizontal mean curvature defined by Definitions 5.18 and 5.19.*

Developing all the computations following the example of [22] (see also Chapter 4, Section 4.2 of this thesis) we obtain the following partial differential equation

$$u_t = Tr((\mathcal{X}_\varepsilon^2 u)^*) - \left\langle (\mathcal{X}_\varepsilon^2 u)^* \frac{\mathcal{X}_\varepsilon u}{|\mathcal{X}_\varepsilon u|}, \frac{\mathcal{X}_\varepsilon u}{|\mathcal{X}_\varepsilon u|} \right\rangle = \Delta_\varepsilon u - \Delta_{0,\infty,\varepsilon} u, \tag{5.12}$$

where

$$(\mathcal{X}_\varepsilon^2 u)^*_{ij} = \frac{X_i^\varepsilon (X_j^\varepsilon u) + X_j^\varepsilon (X_i^\varepsilon u)}{2}. \tag{5.13}$$

We observe now that we can write the Equation (5.12) as

$$u_t + F_\varepsilon(x, Du, D^2 u) = 0, \tag{5.14}$$

with

$$\begin{aligned} F_\varepsilon(x,p,S) &= -Tr(\sigma_\varepsilon(x) S \sigma_\varepsilon^T(x) + A_\varepsilon(x,p)) \\ &\quad + \left\langle \left(\sigma_\varepsilon(x) S \sigma_\varepsilon^T(x) + A_\varepsilon(x,p)\right) \frac{\sigma_\varepsilon(x)p}{|\sigma_\varepsilon(x)p|}, \frac{\sigma_\varepsilon(x)p}{|\sigma_\varepsilon(x)p|} \right\rangle \end{aligned} \tag{5.15}$$

with

$$(A_\varepsilon)_{ij}(x,p) = \frac{1}{2}\langle \nabla_{X_i^\varepsilon} X_j^\varepsilon + \nabla_{X_j^\varepsilon} X_i^\varepsilon, p \rangle.$$

Remark 5.22. *Applying the Definition 5.15 to the function F_ε as defined in* (5.9) *we obtain*

$$F^*(x,p,S) = \begin{cases} -Tr(\overline{S}_\varepsilon) + \left\langle \overline{S}_\varepsilon \frac{\sigma_\varepsilon(x)p}{|\sigma_\varepsilon(x)p|}, \frac{\sigma_\varepsilon(x)p}{|\sigma_\varepsilon(x)p|} \right\rangle, & |p| \neq 0, \\ -Tr(\overline{S}_\varepsilon) + \lambda_{max}(\overline{S}_\varepsilon), & |p| = 0, \end{cases}$$

and

$$F_*(x,p,S) = \begin{cases} -Tr(\overline{S}_\varepsilon) + \left\langle \overline{S}_\varepsilon \frac{\sigma_\varepsilon(x)p}{|\sigma_\varepsilon(x)p|}, \frac{\sigma_\varepsilon(x)p}{|\sigma_\varepsilon(x)p|} \right\rangle, & |p| \neq 0, \\ -Tr(\overline{S}_\varepsilon) + \lambda_{min}(\overline{S}_\varepsilon), & |p| = 0, \end{cases}$$

where $\overline{S}_\varepsilon = \sigma_\varepsilon(x) S \sigma_\varepsilon^T(x) + A_\varepsilon(x,p)$ with λ_{max} and λ_{min} the maximum and the minimum eigenvalues of the matrix $\overline{S}_\varepsilon$. In order to compute upper/lower we use the remarks made in [27], proof Theorem 1.1.

The approximated Riemannian stochastic control problem

Let us consider a family of smooth vector fields $\mathcal{X} = \{X_1, \dots X_m\}$ and its Riemannian approximation $\mathcal{X}_\varepsilon = \{X_1, \dots, X_m, \varepsilon X_{m+1}, \dots, \varepsilon X_N\}$.

Definition 5.23. *We define the* horizontal Brownian motion *the solution of the process*

$$d\xi = \sum_{i=1}^{m} X_i(\xi) \circ dB_m^i,$$

where B_m is a m-dimensional Brownian motion, $\circ$ the Stratonovich differential and X_i the vector fields of $\mathcal{X}$ which span the distribution $\mathcal{H}$. We define the Riemannian approximated horizontal Brownian motion *as*

$$d\xi_\varepsilon = \sum_{i=1}^{N} X_i^\varepsilon(\xi_\varepsilon) \circ dB_N^i$$

where B_N is an N-dimensional Brownian motion and X_i^ε the vector fields of $\mathcal{X}_\varepsilon$ which span the distribution $\mathcal{H}_\varepsilon$.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space, B_j is a j-dimensional Brownian motion adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ with $j = m, N$, we recall that a *predictable* process is a time-continuous stochastic process $\{\xi(t)\}_{t\geq 0}$ defined on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, measurable with respect to the σ-algebra generated by all left-continuous adapted process. Given a smooth function $g : \mathbb{R}^N \to \mathbb{R}$ (which parametrizes the starting hypersurface at time $t = 0$) we introduce $\xi^{t,x,\nu}$ the solution of the stochastic dynamic

$$\begin{cases} d\xi^{t,x,\nu}(s) = \sqrt{2}\sigma^T(\xi^{t,x,\nu}(s)) \circ dB_m^\nu(s), \qquad s \in (t, T], \\ dB_m^\nu(s) = \nu(s) dB_m(s), \\ \xi^{t,x,\nu}(t) = x, \end{cases} \tag{5.16}$$

where the matrix σ is defined in (5.2), $\circ$ represents the differential in the sense of Stratonovich and

$$\mathcal{A} = \left\{\nu : [t, T] \to Sym(m) \text{ predictable } | \nu \geq 0,\ I_m - \nu^2 \geq 0,\ Tr(I_m - \nu^2) = 1\right\} \tag{5.17}$$

and the function $V : [0,T] \times \mathbb{R}^N \to \mathbb{R}$ defined as

$$V(t,x) := \inf_{\nu \in \mathcal{A}} ess \sup_{\omega \in \Omega} g(\xi^{t,x,\nu}(T)(\omega)). \tag{5.18}$$

Similarly, for $\varepsilon > 0$ fixed, we introduce $\xi_\varepsilon^{t,x,\nu_1}$ the solution of

$$\begin{cases} d\xi_\varepsilon^{t,x,\nu_1}(s) = \sqrt{2}\sigma_\varepsilon^T(\xi_\varepsilon^{t,x,\nu_1}(s)) \circ dB_N^{\nu_1}(s), \qquad s \in (t,T], \\ dB_N^{\nu_1}(s) = \nu_1(s) dB_N(s), \\ \xi_\varepsilon^{t,x,\nu_1}(t) = x, \end{cases} \tag{5.19}$$

where σ_ε is the matrix defined in (5.3) and

$$\mathcal{A}_1 = \left\{ \nu_1 : [t,T] \to Sym(N) \text{ predictable} \mid \nu_1 \geq 0,\ I_N - \nu_1^2 \geq 0,\ Tr(I_N - \nu_1^2) = 1 \right\} \tag{5.20}$$

and the function $V^\varepsilon : [0,T] \times \mathbb{R}^N \to \mathbb{R}$ defined by

$$V^\varepsilon(t,x) := \inf_{\nu \in \mathcal{A}_1} ess \sup_{\omega \in \Omega} g(\xi_\varepsilon^{t,x,\nu_1}(T)(\omega)). \tag{5.21}$$

It is possible to show that the function V solves in the viscosity sense respectively the level-set equation for the evolution by HMCF (see [22]).

Note also that the sets of controls (5.17) and (5.20) can be rewritten respectively as

$$\mathcal{A} = \{\nu^2 |\ \nu \in \mathcal{A}\} = Co\{I_m - a \otimes a |\ a \in \mathbb{R}^m,\ \ |a| = 1\},$$

and

$$\mathcal{A}_1 = \{\nu_1^2 |\ \nu_1 \in \mathcal{A}_1\} = Co\{I_N - \overline{a} \otimes \overline{a} |\ \overline{a} \in \mathbb{R}^N,\ \ |\overline{a}| = 1\},$$

where Co is the convex hull (see [27] for more details).
Next we introduce the p-regularising approximation of the functions V and V^ε.

Definition 5.24. *For $p > 1$, the p-value function associated to the value function* (5.18) *is defined as*

$$V_p(t,x) := \inf_{\nu \in \mathcal{A}} \mathbb{E}[|g(\xi^{t,x,\nu})(T)(\omega)|^p]^{\frac{1}{p}}, \tag{5.22}$$

and, similarly, we can introduce the following ε-p-regularising function, that is the p-value function associated to the value function (5.21),

$$V_p^\varepsilon(t,x) := \inf_{\nu_1\in\mathcal{A}} \mathbb{E}[|g(\xi_\varepsilon^{t,x,\nu_1})(T)(\omega)|^p]^{\frac{1}{p}}. \tag{5.23}$$

Definition 5.25. *We define the Hamiltonian associated to the horizontal stochastic optimal control problem* (5.16) *the function*

$$H(x,p,S) = \sup_{\nu\in\mathcal{A}}\Big[- Tr(\sigma(x)S\sigma^T(x)\nu^2(s)) + \sum_{i,j=1}^m (\nu^2(s))_{ij}\big\langle \nabla_{X_i}X_j(x), p\big\rangle\Big].$$

where σ is defined as in (5.2), *$p\in\mathbb{R}^N$ and $S\in Sym(N)$.*

Definition 5.26. *We define the Hamiltonian associated to the approximated stochastic optimal control problem* (5.19) *the function*

$$H_\varepsilon(x,p,S) = \sup_{\nu_1\in\mathcal{A}_1}\Big[- Tr(\sigma_\varepsilon(x)S\sigma_\varepsilon^T(x)\nu_1^2(s)) + \sum_{i,j=1}^N (\nu_1^2(s))_{ij}\big\langle \nabla_{X_i^\varepsilon}X_j^\varepsilon(x), p\big\rangle\Big].$$

where σ_ε is defined as in (5.6), *$p\in\mathbb{R}^N$ and $S\in Sym(N)$.*

Remark 5.27. *The function V_p solves in viscosity sense PDE:*

$$\begin{cases} -(V_p) + H_p(x, DV_p, D^2V_p) = 0, \quad t\in[0,T),\ x\in\mathbb{R}^N, \\ V_p(T,x) = g(x), \qquad\qquad\qquad\qquad x\in\mathbb{R}^N \end{cases} \tag{5.24}$$

where

$$H_p(x,q,M) := \sup_{\nu\in\mathcal{A}}\Big[- (p-1)r^{-1}Tr[\nu\nu^T qq^T] + Tr[\nu\nu^T M]\Big], \tag{5.25}$$

(see [27] for further details).

Remark 5.28. *Similarly to Remark 5.27, for $\varepsilon>0$ and $p>1$ fixed, the function V_p^ε solves in the viscosity sense the PDE*

$$\begin{cases} -(V_p^\varepsilon) + H_p^\varepsilon(x, DV_p^\varepsilon, D^2V_p^\varepsilon) = 0, \quad t\in[0,T),\ x\in\mathbb{R}^N, \\ V_p^\varepsilon(T,x) = g(x), \qquad\qquad\qquad\qquad x\in\mathbb{R}^N \end{cases} \tag{5.26}$$

where

$$H_p^\varepsilon(x,r,q,M) := H_p(x,r,q_\varepsilon,M_\varepsilon) = \sup_{\nu\in\mathcal{A}_1}\left[-(p-1)r^{-1}Tr[\nu\nu^T q_\varepsilon q_\varepsilon^T]+Tr[\nu\nu^T M_\varepsilon]\right], \tag{5.27}$$

where $\mathcal{A}_1$ *is given in* (5.20) *and, for all* $q\in\mathbb{R}^N$ *and* $M=(M_{ij})_{i,j=1}^N \in Sym(N)$,

$$q_\varepsilon := \begin{bmatrix} q_1 \\ \dots \\ q_m \\ \varepsilon q_{m+1} \\ \dots \\ \varepsilon q_N \end{bmatrix}$$

and

$$M_\varepsilon := \begin{bmatrix} M_{11} & \dots & M_{1m} & M_{1(m+1)} & \dots & \varepsilon M_{1N} \\ & & & \vdots & & \\ M_{m1} & \dots & M_{mm} & \varepsilon M_{(m+1)m} & \dots & \varepsilon M_{Nm} \\ \varepsilon M_{(m+1)1} & \dots & \varepsilon M_{(m+1)m} & \varepsilon^2 M_{(m+1)(m+1)} & \dots & \varepsilon^2 M_{(m+1)N} \\ & & & \vdots & & \\ \varepsilon M_{1N} & \dots & \varepsilon M_{mN} & \varepsilon^2 M_{(m+1)N} & \dots & \varepsilon^2 M_{NN} \end{bmatrix}.$$

5.5 V^ε as viscosity solution

In this section we will prove the main result of this paper, but before doing it, we have to introduce some technical lemmas.

Lemma 5.29 (Comparison Principle)**.** *Let us consider* $0<\varepsilon<1$ *fixed. Let* g_1, g_2 *be continuous functions on* $[0,T]\times\mathbb{R}^N$ *with* $g_1\le g_2$ *and* $V_i^\varepsilon(t,x)$ *for* $i=1,2$ *as defined in* (5.21) *with terminal costs* g_i *then it holds true*

$$V_1^\varepsilon(t,x)\le V_2^\varepsilon(t,x) \text{ on } [0,T]\times\mathbb{R}^N.$$

Proof. It follows from the assumption $g_1\le g_2$ and from the properties of infimum and essential supremum. □

Lemma 5.30. *Let us consider $0 < \varepsilon < 1$ fixed. Let g be a bounded and uniformly continuous function on $[0,T] \times \mathbb{R}^N$ and let $V^\varepsilon(t,x)$ be defined as in (5.21) with g as terminal cost. Let us consider $\phi : \mathbb{R} \to \mathbb{R}$ continuous and strictly increasing. Then*

$$\phi(V_g^\varepsilon(t,x)) = V_{\phi(g)}^\varepsilon(t,x).$$

Proof. Since ϕ is an increasing and continuous function, we remark that $\phi(\inf A) = \inf \phi(A)$ where $A \subset \mathbb{R}$. Then, for every measurable function $f : \Omega \to \mathbb{R}$ it is easy to see that

$$\phi(ess\sup f) = ess\sup(\phi(f)),$$

and so we can conclude the proof. □

Remark 5.31. *Lemmas 5.29 and 5.30 allow us to conclude that the set $\{V(t,x) \leq 0\}$ depends only on the set $\{g(x) \leq 0\}$ and not on the specific form of g. Furthermore we will show that $V^\varepsilon(t,x)$ solves (in the viscosity sense) the level set equation for the evolution by horizontal mean curvature flow for a fixed $0 < \varepsilon < 1$.*

We state now the main theorem of the paper.

Theorem 5.32. *Let us consider $0 < \varepsilon < 1$ fixed. Let $g : \mathbb{R}^N \to \mathbb{R}$ be globally bounded and Lipschitz function, $T > 0$ and*

$$\sigma_\varepsilon(x) = [X_1(x), ..., X_m(x), \varepsilon E_{m+1}(x), \ldots, \varepsilon E_N(x)]^T,$$

a $N \times N$ matrix obtained from the Riemannian approximation of the $m \times N$ Hörmander matrix $\sigma(x) = [X_1(x), ..., X_m(x)]^T$ with $m \leq N$ and smooth coefficients and $E_i = (0, \ldots, 0, 1, 0 \ldots, 0)^T$ where 1 is in the i-th component. Assuming that σ_ε and $\nu_\varepsilon(x) = \sum_{i=1}^N \nabla_{X_i^\varepsilon} X_j^\varepsilon(x)$ are Lipschitz (in order to have non-explosion for the solution of the SDE), then the value function $V^\varepsilon(t,x)$ defined by (5.21) is a bounded lower semicontinuous viscosity solution of the

level set equation for the evolution by approximated Riemannian mean curvature flow, with terminal condition $V^\varepsilon(T,x) = g(x)$.

Remark 5.33. *$V^\varepsilon(t,x)$ is a lower semicontinuous function.*

In order to prove the Theorem 5.32 we have to introduce the half-relaxed upper-limit.

Definition 5.34. *We define the* relaxed half-relaxed upper-limit *of* $V_p^\varepsilon(t,x)$

$$V^{\sharp,\varepsilon}(t,x) := \limsup_{(s,y)\to(t,x)\quad p\to\infty} V_p^\varepsilon(s,y).$$

This lemma allows to use the definition of upper half-relaxed limit instead of the definition of upper envelope.

Lemma 5.35. *Let us consider* $0 < \varepsilon < 1$ *fixed. It holds true*

$$V^{\sharp,\varepsilon}(t,x) = V^{*,\varepsilon}(t,x) \quad \textit{for all} \quad (t,x) \in [0,T]\times\mathbb{R}^N$$

where they are defined as in Definitions 5.15 and 5.34.

Proof. We observe that $V^{\sharp,\varepsilon} \geq V^\varepsilon$ and $V^{\sharp,\varepsilon}$ is upper semicontinuous function. Then, since $V^{*,\varepsilon}$ is the smallest upper envelope it holds $V^{\sharp,\varepsilon} \geq V^{*,\varepsilon}$. On the other hand, recalling that $V_p^\varepsilon(t,x) \leq V^\varepsilon(t,x)$ for any t,x, and $p > 1$ and $\varepsilon > 0$ fixed, then taking the lim sup in t,x and p we obtain that $V^{\sharp,\varepsilon} \leq V^{*,\varepsilon}$ and as consequence the result follows. □

Another important observation is related to the L^p-norm related to $V^\varepsilon(t,x)$, i.e. $V_p^\varepsilon(t,x)$ as in Definition 5.24.

We obtain the following result for $0 < \varepsilon < 1$ fixed.

Lemma 5.36. *Let us consider* $0 < \varepsilon < 1$ *fixed. Under the assumptions of Theorem 5.32, we have*

$$V^\varepsilon(t,x) = \lim_{p\to\infty} V_p^\varepsilon(t,x) \quad \textit{for all} \quad (t,x) \in [0,T]\times\mathbb{R}^N,$$

as pointwise convergence.

Proof. As the L^p norm are bounded by essential supremum and increasing we obtain immediately for each fixed control and $\varepsilon > 0$

$$V^\varepsilon(t,x) \geq V_p^\varepsilon(t,x).$$

The other inequality will be proved as in [22]. Let us consider $q \geq 1$, then by the property of the infimum we can find a control ν_q such that

$$\left(\mathbb{E}[g^p(\xi_\varepsilon^{t,x,\nu_{1,q}}(T))]\right)^{\frac{1}{q}} \leq V_q^\varepsilon(t,x) + \frac{1}{q}.$$

The controlled SDE (5.19) has a drift part which depends on the control only through ν_1^2 (we recall by assumption that $\varepsilon > 0$ is fixed) and our control set is convex in ν_1^2. Proceeding as [22], we obtain that there exists a probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}, B_N, \nu_1)$ such that for a subsequence q_k the process $\xi_\varepsilon^{t,x,\nu_{1,q_k}}$ converges weakly to ξ^{t,x,ν_1} and so for any fixed $\overline{q} \geq 1$

$$\lim_{k\to\infty}\left(\mathbb{E}[g^q(\xi_\varepsilon^{t,x,\nu_{1,q_k}}(T))]\right)^{\frac{1}{q}} = \left(\mathbb{E}[g^{\overline{q}}(\xi_\varepsilon^{t,x,\nu_1}(T))]\right)^{\frac{1}{\overline{q}}}.$$

Since the L^q is non decreasing in q

$$\left(\mathbb{E}[g^q(\xi_\varepsilon^{t,x,\nu_1}(T))]\right)^{\frac{1}{\overline{q}}} \leq \lim_{q\to\infty} V_q^\varepsilon(t,x).$$

Finally, using the convergence of L^q norm to L^∞ we obtain

$$V^\varepsilon(t,x) \leq \lim_{q\to\infty} V_q^\varepsilon(t,x).$$

□

In order to prove that V^ε is a viscosity solution of approximated Riemannian mean curvature flow we have to recall a further lemma.

Lemma 5.37 ([27])**.** *Let $S \in Sym(N)$ such that the space of the eigenvectors associated to the maximum eigenvalue is of the dimension one. Then, $S \to \lambda_{max}(S)$ is C^1 in a neighbourhood of S. Moreover, $D\lambda_{max}(S)(H) =< Ha, a >$, for any $a \in \mathbb{R}^m$ eigenvector associated to $\lambda_{max}(S)$ and $|a| = 1$.*

The Theorem 5.32 is the consequence of the following theorem.

Theorem 5.38. *Let us consider $0 < \varepsilon < 1$ fixed. Let $g : \mathbb{R}^N \to \mathbb{R}$ be a globally bounded and Lipschitz function, $T > 0$ and $\sigma_\varepsilon(x)$ a Riemannian approximation of the $m \times N$-Hörmander matrix $\sigma(x)$. Since the comparison principle holds (see [3]), then the value function $V^\varepsilon(t,x)$ is the unique continuous viscosity solution of approximated Riemannian mean curvature flow, satisfying $V^\varepsilon(T,x) = g(x)$.*

Proof. We divide this proof in two steps: we prove that $V^\varepsilon(t,x)$ is a viscosity supersolution and $V^{\varepsilon,\sharp}(t,x)$ is a viscosity subsolution.

- *V^ε is a viscosity supersolution*: Let us consider $\phi \in C^1([0,T]; C^2(\mathbb{R}^N))$ such that $V^\varepsilon - \phi$ has a local minimum at (t,x). Two cases are possible: if $\mathcal{X}_\varepsilon \phi(t,x) \neq (0,\dots,0)$ we have to verify that

$$-\phi_t(t,x) - \Delta_\varepsilon \phi(t,x) + \Delta_{\varepsilon,\infty}\phi(t,x) \geq 0,$$

  where the equation is given as in (5.12).
  If $\mathcal{X}_\varepsilon \phi(t,x) = (0,\dots,0)$ we have to verify that

$$-\phi_t(t,x) - \Delta_\varepsilon \phi(t,x) + \lambda_{max}((\mathcal{X}_\varepsilon^2 \phi)^*(t,x)) \geq 0,$$

  where $(\mathcal{X}_\varepsilon^2\phi)^*$ is defined as (5.13).
  For any $p > 1$ there exists a sequence (t_p, x_p) such that $V_p^\varepsilon - \phi$ has a local minimum at (t_p, x_p) and $(t_p, x_p) \to (t,x)$ a $p \to \infty$. In fact, we can always assume that (t,x) is a strict minimum in some $B_R(t,x)$ (to obtain this it is sufficient to substitute a generic test function ϕ with the test function $\phi + |x - x_p|^4$). Set $K = \overline{B_{\frac{R}{2}}(t,x)}$, the sequence of minimum points (t_p, x_p) converge to some $(\bar{t}, \bar{x}) \in K$. As V^ε is the limit of V_p^ε as $p \to \infty$ (see Lemma 5.36) and lower semicontinuous, therefore by a standard argument yields that $(\bar{t}, \bar{x})$ is a minimum, hence it equals (t,x). Then it holds true

$$-\phi_t(t_p,x_p) + H_\varepsilon(x_p, (p-1)V_p^{-1} D\phi (D\phi)^T + D^2\phi)(t_p,x_p) \geq 0.$$

If $\sigma_\varepsilon(x)D\phi(t,x) \neq 0$, we can write the Hamiltonian in a more explicit way. Set

$$S_1 = (p-1)V_p^{-1}(\mathcal{X}_\varepsilon\phi(t_p,x_p))(\mathcal{X}_\varepsilon\phi(t_p,x_p))^T,$$

and

$$S_2 = (\mathcal{X}_\varepsilon^2\phi)^*(t_p,x_p),$$

then

$$\begin{aligned} H_\varepsilon(x_p,S_1,S_2) &= -Tr(S_1+S_2) + \lambda_{max}(S_1+S_2) \\ &= -Tr(S_1) - Tr(S_2) + \lambda_{max}(S_1+S_2) \\ &= -(p-1)(V_p^\varepsilon)^{-1}(t_p,x_p)|\mathcal{X}_\varepsilon\phi(t_p,x_p)|^2 \\ &\quad - \Delta_\varepsilon\phi(t_p,x_p) + \lambda_{max}(S_1+S_2), \end{aligned} \tag{5.28}$$

since the trace operator is linear and $Tr((\mathcal{X}_\varepsilon\phi(x_p))(\mathcal{X}_\varepsilon\phi(x_p))^T = |\mathcal{X}_\varepsilon\phi(x_p)|^2$. Now we use the Lemma 5.37 in order to expand λ_{max}. We consider the matrix

$$S = \frac{\mathcal{X}_\varepsilon\phi(t,x)(\mathcal{X}_\varepsilon\phi(t,x))^T}{V^\varepsilon(t,x)},$$

for which $\lambda_{max}(S) = \frac{|\mathcal{X}_\varepsilon\phi(t,x)|^2}{V^\varepsilon(t,x)}$ and where $a = \frac{\mathcal{X}_\varepsilon\phi(t,x)}{|\mathcal{X}_\varepsilon\phi(t,x)|}$ since $\mathcal{X}_\varepsilon\phi(t,x) \neq 0$ (see [27] for further remarks). Let us consider

$$S_p = \frac{(\mathcal{X}_\varepsilon\phi(t_p,x_p))(\mathcal{X}_\varepsilon\phi(t_p,x_p))^T}{V_p^\varepsilon(t_p,x_p)},$$

it is immediate to observe that S_p converges to S as $p \to \infty$. By Taylor's formula we know that there exists a $\theta_p \in (0,1)$ such that

$$\begin{aligned} \lambda_{max}\left(S_p + \frac{(\mathcal{X}_\varepsilon^2\phi)^*(t_p,x_p)}{p-1}\right) &= \lambda_{max}(S_p) \\ &+ \frac{1}{p-1}D\lambda_{max}\left(S_p + \frac{\theta_p}{p-1}(\mathcal{X}_\varepsilon^2\phi)^*(t_p,x_p)\right)(\mathcal{X}_\varepsilon^2\phi)^*(t_p,x_p). \end{aligned}$$

Using the fact that λ_{max} is C^1 in a neighbourhood of S and $S_p \to S$ to get

$$\begin{aligned} \lambda_{max}\left(S_p + \frac{(\mathcal{X}_\varepsilon^2\phi)^*(t_p,x_p)}{p-1}\right) &= \lambda_{max}(S_p) \\ &+ \frac{1}{p-1}D\lambda_{max}(S)(\mathcal{X}_\varepsilon^2\phi)^*(t_p,x_p) + o\left(\frac{1}{p}\right), \end{aligned}$$

where $po(1/p) \to 0$ when $p \to \infty$. As consequence we obtain

$$\lambda_{max}\left(S_p + \frac{(\mathcal{X}_\varepsilon^2\phi)^*(t_p,x_p)}{p-1}\right) = \lambda_{max}(S_p) + \frac{< (\mathcal{X}_\varepsilon^2\phi)^*(t_p,x_p)\mathcal{X}_\varepsilon\phi(t,x), \mathcal{X}_\varepsilon\phi(t,x) >}{(p-1)|(\mathcal{X}_\varepsilon\phi)(t,x)|^2},$$

then, expanding the p-Hamiltonian (5.28) we obtain immediately the inequality, If $\mathcal{X}_\varepsilon\phi(t,x) = 0$ then we use the subadditivity of $S \to \lambda_{max}(S)$ and remark that, since V_p^ε is supersolution

$$\begin{aligned}
0 &\leq -\phi_t + H_\varepsilon(x_p, D\phi, (p-1)(V_p^\varepsilon)^{-1}D\phi(D\phi)^T + D^2\phi) \\
&\leq -\phi_t - (p-1)(V_p^\varepsilon)^{-1}|\mathcal{X}_\varepsilon\phi|^2 - Tr((\mathcal{X}_\varepsilon^2\phi)^*) \\
&+ \lambda_{max}((p-1)(V_p^\varepsilon)^{-1}\mathcal{X}_\varepsilon\phi(\mathcal{X}_\varepsilon\phi)^T + (\mathcal{X}_\varepsilon^2\phi)^*) \\
&\leq -\phi_t - (p-1)(V_p^\varepsilon)^{-1}|\mathcal{X}_\varepsilon\phi|^2 - Tr((\mathcal{X}_\varepsilon^2\phi)^*) \\
&+ (p-1)(V_p^\varepsilon)^{-1}|\mathcal{X}_\varepsilon\phi|^2 + \lambda_{max}(\mathcal{X}_\varepsilon^2\phi)^* \\
&= -\phi_t - Tr((\mathcal{X}_\varepsilon^2\phi)^*) + \lambda_{max}(\mathcal{X}_\varepsilon^2\phi)^*.
\end{aligned}$$

In the end, we can conclude now that V^ε is a supersolution.

- $V^{*,\varepsilon}$ *is the subsolution*: As consequence of Lemma 5.35 we can write $V^{*,\varepsilon} = V^{\sharp,\varepsilon}$. Let $\phi \in C^1([0,T]; C^2(\mathbb{R}^N))$ such that $V^{\sharp,\varepsilon} - \phi$ has a strict maximum at (t_0, x_0). Let us consider a sequence of maximum points of $V_p^\varepsilon - \phi$, we can find a subsequence converging to (t,x). Since V_p^ε is the solution of

$$\begin{cases} -(V_p)_t + H_\varepsilon(x, DV_p^\varepsilon, (p-1)(V_p^\varepsilon)^{-1}DV_p^\varepsilon(DV_p^\varepsilon)^T + D^2V_p^\varepsilon) = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad x \in \mathbb{R}^N,\ t \in [0,T), \\ \qquad\qquad\qquad\qquad\qquad V_p^\varepsilon(T,x) = g(x), \qquad x \in \mathbb{R}^N, \end{cases} \tag{5.29}$$

then we have that

$$0 \leq -\phi_t + H_\varepsilon(x, (p-1)(V_p^\varepsilon)^{-1}D\phi(D\phi)^T + D^2\phi) \tag{5.30}$$

at the point (t_p, x_p). We define for any $z > 0$, $x, d \in \mathbb{R}^N$and any $N \times N$ symmetric matrix S

$$H_p^\varepsilon(x,z,d,S) = -\frac{(p-1)}{z}|\sigma_\varepsilon(x)d|^2 - Tr(\sigma_\varepsilon^T(x)S\sigma_\varepsilon(x) + A_\varepsilon(x,d)) \\ +\lambda_{max}\left(\frac{(p-1)}{z}(\sigma_\varepsilon(x)d)(\sigma_\varepsilon(x)d)^T + \sigma_\varepsilon^T(x)S\sigma_\varepsilon(x) + A_\varepsilon(x,d)\right)$$

and

$$(H_\varepsilon)^*(x,d,S) = \begin{cases} -Tr(\sigma_\varepsilon^T(x)S\sigma_\varepsilon(x) + A_\varepsilon(x,d)) \\ \qquad +\left\langle (\sigma_\varepsilon^T(x)S\sigma_\varepsilon(x) + A_\varepsilon(x,d))\frac{\sigma_\varepsilon(x)d}{|\sigma_\varepsilon(x)d|}, \frac{\sigma_\varepsilon(x)d}{|\sigma_\varepsilon(x)d|}\right\rangle, \\ \qquad\qquad\qquad\qquad\qquad\qquad |d| \neq 0, \\ -Tr(\sigma_\varepsilon^T(x)S\sigma_\varepsilon(x) + A_\varepsilon(x,d)) \\ \qquad +\lambda_{max}(\sigma_\varepsilon^T(x)S\sigma_\varepsilon(x) + A_\varepsilon(x,d)), \qquad |d| = 0, \end{cases}$$

and, as stated in [22], we can observe that

$$H_p^\varepsilon(x,z,d,S) \geq (H^\varepsilon)^*(x,d,S).$$

We remark that for $|d| = 0$ is immediate, for $|d| \neq 0$ we observe that

$$\lambda_{max}\left(\frac{(p-1)}{z}(\sigma_\varepsilon(x)d)(\sigma_\varepsilon(x)d)^T + \sigma_\varepsilon^T(x)S\sigma_\varepsilon(x) + A_\varepsilon(x,p)\right) \\ \geq \frac{(p-1)}{z}|\sigma_\varepsilon(x)d|^2 + \lambda_{max}(\sigma_\varepsilon^T(x)S\sigma_\varepsilon(x) + A_\varepsilon(x,p))$$

and, called $\overline{S}_\varepsilon = \sigma_\varepsilon^T(x)S\sigma_\varepsilon(x) + A_\varepsilon(x,p)$

$$\lambda_{max}(\overline{S}_\varepsilon) = \max_{|a|=1} < \overline{S}_\varepsilon a, a >,$$

we obtain immediately the inequality. Let us consider $\varepsilon > 0$, set $z = \phi^{-1}(t_p, x_p) > 0$, $d = D\phi(t_p, x_p)$, $S = D^2\phi(t_p, x_p)$, then taking the limsup of (5.30) we obtain for $p \to \infty$ and recalling that, by definition, $(H_\varepsilon)^* \geq (H_\varepsilon)_*$ we obtain

$$0 \geq \phi_t + (H_\varepsilon)_*(x, D\phi, D^2\phi)$$

at (t, x). The result follows immediately.

□

Now, in order to prove the main theorem of this section, we need a further lemma.

Lemma 5.39. *Let us consider* $0 < \varepsilon < 1$ *fixed. For any* $x \in \mathbb{R}^N$, $V^{\varepsilon,\sharp}(T,x) \leq g(x)$.

Proof. By contradiction, we assume that it is not true and that there exists a point x_0 such that $V^{\varepsilon,\sharp}(T,x) \geq g(x_0) + \delta$, for $\delta > 0$ sufficiently small. We use as test function

$$\phi(t,x) = \alpha(T-t) + \beta|x - x_0|^2 + g(x_0) + \frac{\delta}{2},$$

with $\alpha > -C\beta$, with C a constant depending just on the data of the problem and the point x_0 and $\beta > 1$ sufficiently large. We remark that

$$\phi_t(t,x) = \alpha, \qquad D\phi(t,x) = 2\beta(x - x_0), \qquad D^2\phi(t,x) = 2\beta Id.$$

We can find a sequence $(t_k, x_k) \to (T, x_0)$ and $p_k \to \infty$ as $k \to \infty$ such that $V^{\varepsilon}_{p_k} - \phi$ has a positive local maximum at some point (s_k, y_k), for any $k > 1$. To obtain the contradiction we use the fact that $V^{\varepsilon}_{p_k}$ is solution of the Equation (5.29) in order to obtain $\alpha + C\beta \leq 0$. We observe that the functions V^{ε}_{p} are bounded uniformly in p and ε is fixed so, by the growth of $|x - x_0|$, the maximum points are such that $y_k \in \overline{B_R(x_0)} =: K$ with R independent of k. In the point (s_k, y_k) it holds true

$$\begin{aligned} 0 &\geq \alpha - H_\varepsilon(y_k, (p-1)\phi^{-1}D\phi(D\phi)^T + D^2\phi) \\ &\geq \alpha - 2\beta Tr(\sigma_\varepsilon(y_k)\sigma^T_\varepsilon(y_k) + A_\varepsilon(y_k, y_k - x_0)) \\ &+ 2\beta\lambda_{min}(\sigma_\varepsilon(y_k)\sigma^T_\varepsilon(y_k) + A_\varepsilon(y_k, y_k - x_0)). \end{aligned}$$

Then recalling that there is a compact set K such that $y_k \in K$ for all k, by continuity, we get $0 \geq \alpha + C\beta$, with

$$C = -\max_{x\in K} Tr(\sigma_\varepsilon(x)\sigma_\varepsilon^T(x)) - \max_{x\in K} A_\varepsilon(x, x - x_0)$$
$$+ \min_{k\in K} \lambda_{min}(\sigma_\varepsilon(x)\sigma_\varepsilon^T(x)) + \min_{x\in K} \lambda_{min}(A_\varepsilon(x, x - x_0))$$

with such estimate we obtain the contradiction, i.e. the thesis. □

Corollary 5.40. *Let us consider $0 < \varepsilon < 1$ fixed. Let $g : \mathbb{R}^N \to \mathbb{R}$ be bounded and Hölder continuous, $T > 0$ and $\sigma_\varepsilon(x)$ a $N \times N$-Hörmander matrix like in Theorem 5.32. Since the comparison principle holds (see [3]), then the value function $V^\varepsilon(t, x)$ is the unique continuous viscosity solution of the level set equation* (5.12)*, satisfying $V^\varepsilon(T, x) = g(x)$.*

Proof. We have already shown that $V^{\varepsilon,*}(t, x) = V^{\varepsilon,\#}(t, x)$ is a viscosity subsolution while $V_*^\varepsilon(t, x) = V^\varepsilon(t, x)$ is a viscosity supersolution of (5.12) with initial condition g. For Lemma 5.39 we know that $V^{\varepsilon,\#}(t, x) \leq g(x)$ and $V(T, x) = g(x)$ so, by comparison principle, it holds $V^{\varepsilon,\#}(t, x) \leq V^\varepsilon(t, x)$. By definition of lim sup we have $V^{\varepsilon,\#}(t, x) \geq V^\varepsilon(t, x)$ i.e. $V^\varepsilon(t, x)$ is upper semicontinuous. Since $V^\varepsilon(t, x)$ is also lower semicontinuous we can conclude immediately stating that $V^\varepsilon(t, x)$ is continuous. □

Chapter 6

Optimal control for the p-Hamiltonian in $\mathbb{H}^1$

6.1 Introduction

The evolution by mean curvature flow is a geometrical degenerate PDE broadly used in mathematics, see e.g. [13], [28] and references in them for an overview on the subject. Roughly speaking this describes the motion of a hypersurface contracting in the normal direction with (normal) velocity equal to the mean curvature at that point. Unfortunately, even smooth surfaces evolving by mean curvature flow can develop singularities in finite time, so a weak notion for this evolution is necessary. The notion that we are considering here follows a nonlinear PDE-approach, based on Chen-Giga-Goto [11] and Evans-Spruck [31]. Roughly speaking, the idea consists in associating a PDE to a smooth hypersurface evolving such that the function which solves this PDE has level sets which evolve by mean curvature flow. Then one can define the solutions of the "generalized evolution by mean curvature flow" as the zero-level sets of the viscosity solution of this PDE. This, so called, level set approach requires to solve (in the viscosity sense) a degenerate parabolic PDE. In the last decades this evolution has been generalised to the case of

sub-Riemannian geometries, i.e. to the study of hypersurfaces evolving by the so-called horizontal mean curvature flow (see [8, 22] and others). This is partially motivated by the sub-Riemannian modelling of the visual cortex applied for example to the study of image processing, developed by Citti-Sarti and al. (see e.g. [13]).

Sub-Riemannian geometries are degenerate manifolds where the Riemannian inner product is defined just on a sub-bundle of the tangent bundle. To be more precise, we will consider $X_1, \dots, X_m$ smooth vector fields on $\mathbb{R}^n$ and a Riemannian inner product defined on the distribution $\mathcal{H}$ generated by such vector fields. Then it is possible to define intrinsic derivatives of any order by taking the derivatives along the vector fields $X_1, \dots, X_m$. That allows us to write differential operators like Laplacian, infinite-Laplacian etc, using intrinsic derivatives. In particular we can write the level-set equation associated to evolution by horizontal mean curvature flow. Even if different authors have studied this geometrical evolutions, many questions remain open due to the high degeneracy of the associated PDEs (see e.g. [8, 22, 32]). To keep the computation easier and more explicit, in this paper we will focus only on the specific case of the Heisenberg group, which is the main model for a sub-Riemannian geometry. Still the approach works in the general case of Carnot-type vector fields, which in particular includes all Carnot groups.

A connection between certain stochastic control problems and a large class of geometric evolution equations, including the (Euclidean) evolution by mean curvature flow, has been found by Buckdahn, Cardaliaguet and Quincampoix in [27] and Soner and Touzi in [49, 52]. The control, loosely speaking, constrains the increments of the stochastic process to a lower dimensional subspace of $\mathbb{R}^N$, while the cost functional consists only of the terminal cost but involves an essential supremum over the probability space. It turns out that the value function solves the level set equation associated with the geometric

evolution. Moreover, one can show that the set of points from which the initial hypersurface can be reached almost surely in a given time by choosing an appropriate control coincides with the set evolving by mean curvature flow. This stochastic approach generalizes very naturally to sub-Riemannian geometries by using an intrinsic Brownian motion associated with the sub-Riemannian geometry.

This approach can be used to obtain certain existence results in general sub-Riemannian manifolds. In particular, the value function may be used for defining a generalized flow. More precisely, the value function V associated to this stochastic control problem is defined as the infimum, over the admissible controls, of the essential supremum of the final cost g (at some fixed terminal time $T > t$), for the controlled path $\xi^{t,x,\nu}$ starting from x at the time t. We can show that $u(t,x) := V(T-t,x)$ is a viscosity solution of the level set equation of the evolution by horizontal mean curvature flow. So $\Gamma(t) = \{x \in \mathbb{R}^N \mid u(t,x) = 0\}$ is a generalized evolution by horizontal mean curvature flow in general sub-Riemannian manifolds. This approach has been successfully used to study the evolution by horizontal mean curvature flow in general sub-Riemannian geometries by two of the authors, together with Max von Renesse in [22]. In that paper, following the approach in [27], the authors introduce a suitable p-regularising stochastic optimal control problem, which does not degenerate when the (horizontal) gradient vanishes. The value functions u_p associated to the p-problem do not converge as $p \to +\infty$, but their $p-$th roots $u_p^{\frac{1}{p}}$ do, in a similar way as the L^p-seminorms of a measurable function converge to the essential supremum. This limit of $u_p^{\frac{1}{p}}$ can be shown to solve in the viscosity sense the level set equation for the horizontal mean curvature flow in general sub-Riemannian geometries.

The aim of this chapter is to understand better, at least at a formal level, the asymptotic behaviour of the optimal controls of these approximating control problems.

For stochastic control problems, the optimal control is of feedback form. This means that there exist a function, depending on the value function and its derivatives, which selects the optimal control depending on the state of the system. This function is obtained by a point-wise optimization over the control space which, in the simplest case, is just the Legendre transform connecting Lagrangian and Hamiltonian, for details see e.g. [33].

For standard control problems, it is possible to define a forward-backward system of stochastic ODEs which yields the path associated with the optimal control, see e.g [10]. This approach does not need the derivatives of the value function, the price to pay is that we have a system which is forward-backward. In principle this would allow to reduce questions regarding the convergence of value functions to convergence of a family of systems of stochastic ODEs.

Unfortunately, in the case here, the value functions does not converge, only their p-th roots do. Re-writing everything depending on the p-th root of the value function would lead to a system of ODEs that still depends on the (in principle unknown) value function. The good news is, however, that the convergence of the value function has been shown in [22], but without rate. Therefore our approach of studying the convergence of the Hamiltonian in p nevertheless is able to shed some light on the behaviour of the approximating optimal controlled paths near characteristic points of the limit problem.

This result gives also an idea on the structure of the optimal controls for the p-problem, which is crucial for showing the convergence of the stochastic approach for the Riemannian approximation to the value function solving the level-set equation in the horizontal case. These results will be contained in two follow-up papers in preparation.

6.2 Preliminary

In order to prove our main result, we have to recall briefly the basic definitions about sub-Riemannian geometries. For further remarks and definitions

on this topic we refer to [4] and [42] (and also Chapter 1).
Let M be a N-dimensional smooth manifold, we recall that a *distribution* is a subbundle of the tangent bundle, i.e. as vector space

$$\mathcal{H} := \{(x, v) | x \in M \quad v \in \mathcal{H}_x\},$$

where $\mathcal{H}_x$ is a subspace of the tangent space T_xM at every point $x \in M$.
Given two vector fields X,Y defined on a manifold M, we can consider the bracket between X and Y, that is the vector field acting on the smooth functions $f : M \to \mathbb{R}$ as

$$[X, Y](f) = XY(f) - YX(f).$$

Let us now consider a family of vector fields $\mathcal{X} := \{X_1, \dots, X_m\}$, we define the set of all the k-brackets of $\mathcal{X}$ as

$$\mathcal{L}^{(k)}(\mathcal{X}) := \{[X, Y] | X \in \mathcal{L}^{(k-1)}(\mathcal{X}), \quad Y \in \mathcal{L}^{(1)}(\mathcal{X})\},$$

with $\mathcal{L}^{(1)}(\mathcal{X}) = \mathcal{X}$. The associated Lie algebra is the set of all brakets between the vector fields of the family

$$\mathcal{L}(\mathcal{X}) := \{[X_i, X_j^{(k)}] | X_j^{(k)} \in \mathcal{L}^{(k)}(\mathcal{X}),\ k \in \mathbb{N}\}.$$

We can now recall the Hörmander condition.

Definition 6.1 (Hörmander condition)**.** *Let M be a smooth manifold and $\mathcal{H}$ a distribution defined on M and let $\mathcal{X}$ be a family fo vector fields spanning $\mathcal{H}$. We say that the distribution is* bracket generating *if and only if, at any point, the Lie algebra $\mathcal{L}(\mathcal{X})$ spans the whole tangent space at that point. Moreover we say that the family of vector fields $\mathcal{X}$ satisfy the* Hörmander condition *if and only if there exists $r \in \mathbb{N}$ such that $T_xM = \bigcup_{k=1}^{r} Span\big(\mathcal{L}^{(k)}(\mathcal{X})(x)\big)$; in this case the natural number r is called step of the group. The space $\mathcal{H}_x = Span\big(X_1, \dots, X_m\big)(x)$ is usually called* horizontal space *at the point x.*

In this setting not all the curves on the manifolds will be admissible.

Definition 6.2. *Let M be a smooth manifold and $\mathcal{H}$ a bracket generating distribution defined on M and generating by a family of vector fields $\mathcal{X} = \{X_1, \dots, X_m\}$. Consider an absolutely continuous curve $\gamma : [0, T] \to M$, we say that γ is a* horizontal curve *if and only if*

$$\dot{\gamma}(t) \in \mathcal{H}_{\gamma(t)}, \quad \text{for a.e } t \in [0, T]$$

or, equivalently, if there exists a measurable function $h : [0, T] \to \mathbb{R}^N$ such that

$$\dot{\gamma}(t) = \sum_{i=1}^{m} h_i(t) X_i(\gamma(t)), \quad \text{for a.e } t \in [0, T],$$

where $h(t) = (h_1(t), \dots, h_m(t))$.

In the next example we will introduce the most significant and famous model in this setting: the Heisenberg group.

Example 6.3 (The Heisenberg group)**.** *For a formal definition of the Heisenberg group and the connection between its structure as non commutative Lie group and its manifold structure we refer to [4]. Here we simply introduce the 1-dimensional Heisenberg group as the geometries induced on $\mathbb{R}^3$ by the vector fields*

$$X_1(x) = \begin{pmatrix} 1 \\ 0 \\ -\frac{x_2}{2} \end{pmatrix} \quad \text{and} \quad X_2 = \begin{pmatrix} 0 \\ 1 \\ \frac{x_1}{2} \end{pmatrix}, \quad \forall\, x = (x_1, x_2, x_3) \in \mathbb{R}^3.$$

The horizontal space in this case is given by $H_x = Span\big(X_1(x), X_2(x)\big)$; so in particular at the origin the horizontal space is the plane $x_3 = 0$.
Note that the above vector fields satisfy the Hörmander condition with step 2: in fact $[X_1, X_2](x) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ for any $x \in \mathbb{R}^3$.

From now on we consider only the case where the starting topological manifold M is the Euclidean $\mathbb{R}^N$.

For later use we also introduce the matrix associated to the vector fields $X_1, \dots, X_m$, which is the $N \times m$ matrix defined as

$$\sigma(x) = [X_1(x), \dots, X_m(x)]^T.$$

Example 6.4. *In the case of the Heisenberg group introduced in the Example 6.3, the matrix σ is given by*

$$\sigma(x) = \begin{bmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \end{bmatrix}, \quad \forall\, x = (x_1, x_2, x_3) \in \mathbb{R}^3. \tag{6.1}$$

Moreover, in this paper we will concentrate on a sub-Riemannian geometries with a particular structure: the so called Carnot-type geometries. In general, for Carnot-type geometries, the matrix σ assumes the following structure:

$$\sigma(x) = \begin{bmatrix} I_{m\times m} & A(x_1, \dots x_m) \end{bmatrix},$$

where the matrix $A(x_1, \dots, x_m)$ is a $(N-m) \times m$ depending only on the first m components of x.

All Carnot groups are Carnot-type geometries (see e.g. [4] for definitions and examples of Carnot groups). The previous assumption on the structure of the vector fields will allow us to consider an easy and explicit form fort the Riemannian approximation. Nevertheless the approach introduced apply also to the case where this additional structure is not fulfilled.

6.3 Horizontal mean curvature flow

Given a smooth hypersurface Γ on $\mathbb{R}^N$ (or more in general on a N-dimensional manifold M), we indicate by $n_E(x)$ the standard (Euclidean) normal to Γ at the point x. We now consider the vector fields $X_1, \dots, X_m$ introduced in the previous section (i.e. spanning a bracket-generating distribution $\mathcal{H}_x$), and we look at the unit vector obtained projecting the Euclidean normal on the distribution $\mathcal{H}_x$ generated by $X_1, \dots, X_m$, see the following definition (see Chapter 4 for further details).

Definition 6.5 (Horizontal normal)**.** *Given a smooth hypersurface* Γ *on* $\mathbb{R}^N$ *and a family of vector fields* $X_1, \dots, X_m$ *satisfying the Hörmander condition, the* horizontal normal $n_0(x)$ *is the renormalized projection of the Euclidean normal* $n_E(x)$ *on the horizontal space* $\mathcal{H}_x = Span\big(X_1(x), \dots, X_m(x)\big)$.

Since $n_0(x) \in \mathrm{Span}\big(X_1(x), \dots, X_m(x)\big)$ and the Riemannian inner product is introduced in such a way $X_1, \dots, X_m$ are orthonormal, then, whenever the projection of $n_E(x)$ onto $\mathcal{H}_x$ does not vanish, there exists $h_1, \dots, h_m$ measurable functions such that

$$n_0(x) = \frac{h_1(x)X_1(x) + \cdots + h_m(x)X_m(x)}{\sqrt{h_1^2(x) + \cdots + h_m^2(x)}} \in \mathcal{H}_x \subset \mathbb{R}^N, \quad x \in \Gamma.$$

With an abuse of notation we sometimes identify n_0 with the associated m-valued vector

$$n_0(x) \to \left(\frac{h_1(x)}{\sqrt{h_1^2(x) + \cdots + h_m^2(x)}}, \dots, \frac{h_m(x)}{\sqrt{h_1^2(x) + \cdots + h_m^2(x)}} \right)^T \in \mathbb{R}^m. \tag{6.2}$$

The main difference between the standard normal and the horizontal normal is that the second may not exist even for smooth hypersurfaces. In fact whenever the Euclidean normal is orthogonal to the horizontal plane $\mathcal{H}_x$, then the horizontal normal cannot be introduced. The points where this happens are called characteristic points, see the definition below.

Definition 6.6 (Characteristic points)**.** *Given a smooth hypersurface* Γ *on* $\mathbb{R}^N$*, the characteristic points occur whenever* $n_E(x)$ *is orthogonal to the horizontal plane* $\mathcal{H}_x$*, then its projection on such a subspace vanishes, i.e.*

$$h_1^2(x) + \cdots + h_m^2(x) = 0.$$

We recall that for every smooth hypersurface the *mean curvature* at a point is defined as the divergence of the Euclidean normal at that point. Similarly, for every smooth hypersurface, we can now introduce the horizontal mean curvature.

Definition 6.7 (Horizontal mean curvature)**.** *Given a smooth hypersurface* Γ *and a non characteristic point* $x \in \Gamma$*, the* horizontal mean curvature *is defined as the horizontal divergence of the horizontal normal, i.e.* $k_0(x) = div_{\mathcal{H}} n_0(x)$*, where* $n_0(x)$ *is the m-valued vector associated to the horizontal normal defined in* (7.3)*, while* $div_{\mathcal{H}}$ *is the divergence w.r.t. the vector fields* $X_1, \dots, X_m$*, i.e.*

$$k_0(x) = X_1 \left(\frac{h_1(x)}{\sqrt{\sum_{i=1}^m h_i^2(x)}} \right) + \cdots + X_m \left(\frac{h_m(x)}{\sqrt{\sum_{i=1}^m h_i^2(x)}} \right) \quad x \in \partial\Gamma. \quad (6.3)$$

Obviously the horizontal mean curvature is never defined at characteristic points, since there the horizontal normal does not exist.

We can finally introduce the main definition of this section.

Definition 6.8 (Evolution by mean curvature flow)**.** *Let* Γ_t *be a family of smooth hypersurfaces in* $\mathbb{R}^N$*, depending on a time parameter* $t \geq 0$*. We say that* Γ_t *is an* evolution by horizontal mean curvature flow *of some hypersurface* Γ *if and only if* $\Gamma_0 = \Gamma$ *and for any smooth horizontal curve* $\gamma : [0, T] \to \mathbb{R}^N$ *such that* $\gamma(t) \in \Gamma_t$ *for all* $t \in [0, T]$*, the horizontal normal velocity* v_0 *is equal to minus the horizontal mean curvature, i.e.*

$$v_0(\gamma(t)) := -k_0(\gamma(t)) n_0(\gamma(t)), \quad t \in [0, T], \quad (6.4)$$

where n_0 *and* k_0 *as respectively the horizontal normal and the horizontal mean curvature introduced in Definitions 6.5 and 6.7.*

We recall that (6.4) is not defined at characteristic points, then the need to develop a generalised notion of evolution by mean curvature flow which can deal with characteristic points and general singularities, as we will do in the next section following some very well-known approaches, already used to deal with singularities in the Euclidean case.

6.4 The level-set equation: a stochastic approach.

In this section we introduce the level set equation for the (generalised) evolution by horizontal mean curvature flow, and a stochastic representation for the viscosity solutions of that equation.

The level set approach for the Euclidean evolution was introduced by Evans and Spruck in [31] and Chen, Giga and Goto in [11]. We briefly recall this approach directly for the horizontal evolution introduced in (6.4), for more details see [22] and [8]. The basic idea starts by parametrising all (smooth) hypersurface involved as zero level sets, i.e.

$$\Gamma = \Gamma_0 = \big\{x \in \mathbb{R}^N | u_0(x) = 0\big\} \quad \text{and} \quad \Gamma_t = \big\{x \in \mathbb{R}^N | u(t,x) = 0\big\},$$

for some smooth function $u : [0, +\infty) \times \mathbb{R}^N \to \mathbb{R}$. From now on, we indicate the points $x \in \Gamma_t$ as $x(t)$. Then the Euclidean normal is simply $n_E(x(t)) = \frac{\nabla u(t,x(t))}{|\nabla u(t,x(t))|}$, where the gradient is done only w.r.t. the space variable x, for all $x(t) \in \Gamma_t$. This implies that the horizontal normal (at non-characteristic points) can be expressed as

$$n_0(x(t)) = \left(\frac{X_1 u(t,x(t))}{\sqrt{\sum_{i=1}^m (X_i u(t,x(t)))^2}}, \dots, \frac{X_m u(t,x(t))}{\sqrt{\sum_{i=1}^m (X_i u(t,x(t)))^2}} \right) \in \mathbb{R}^m, \tag{6.5}$$

(where above we have simply identifies the normal with its coordinate vector in $\mathbb{R}^m$). Note that $D_{\mathcal{X}} u = (X_1 u, \dots, X_m u) \in \mathbb{R}^m$ is the so called horizontal gradient, then $n_0 = \frac{D_{\mathcal{X}} u}{|D_{\mathcal{X}} u|}$ and by $|\cdot|$ we indicate the standard Euclidean norm in $\mathbb{R}^m$. Similarly we can then write the horizontal mean curvature given in (6.3) as

$$k_0(x(t)) = \sum_{i=1}^m X_i \left(\frac{X_i u(t,x(x))}{\sqrt{\sum_{i=1}^m (X_i u(t,x(t)))^2}} \right). \tag{6.6}$$

Applying (6.5) and (6.6) to the Definition 6.8, we obtain that, whenever Γ_t satisfies the evolution (6.4) and $|D_{\mathcal{X}} u| > 0$, then u solves the following PDE:

$$u_t = Tr((D^2_{\mathcal{X}} u)^*) - \left\langle (D^2_{\mathcal{X}} u)^* \frac{D_{\mathcal{X}} u}{|D_{\mathcal{X}} u|}, \frac{D_{\mathcal{X}} u}{|D_{\mathcal{X}} u|} \right\rangle, \tag{6.7}$$

where $(D^2_{\mathcal{X}}u)^*$ is the symmetric horizontal Hessian, that is

$$((D^2_{\mathcal{X}}u)^*)_{ij} := \frac{X_i(X_ju) + X_j(X_iu)}{2}.$$

(see Chapter 4 for further remarks about the derivation of the PDE (6.7)).

Remark 6.9. *We remark that, given $u\colon \mathbb{R}^3 \to \mathbb{R}$ smooth function, then the symmetrized horizontal Hessian $(D^2_{\mathcal{X}}u)^*$ has only second order derivatives (see Subsection 4.3, Chapter 4 for further remarks).*

Equation (6.7) was introduced and studied in [8] and [31]. Different approaches lead to different ways to interpret the singularity $|D_{\mathcal{X}}u|$, which happens even if the surface is smooth in the Euclidean sense. To see this, consider the unit sphere in three dimensions centred at $(0, 0, 1)$. In the Heisenberg geometry, the horizontal gradient to any level set function vanishes in the point $(0, 0, 1)$. This example will be used for some numerical illustrations in the final chapter.

Equation (6.7) is very degenerate and in general the solutions will need to be interpreted in the sense of the viscosity solutions (see [17], for a definition and properties).

Uniqueness of viscosity solutions is in full generality an open problem due to the presence of points where $|D_{\mathcal{X}}u|$ vanishes, i.e. characteristic points. In fact these points makes difficult to find a general comparison principle for horizontal mean curvature flow. However, in some cases it is possible to have it under certain conditions (see e.g. [8], [32]).

Here we concentrate on the stochastic approach initiated independently by Cardaliaguet, Buckdahn and Quincampoix in [27] and by Soner and Touzi in [49]. The same approach was later generalised by Dirr, Dragoni and von Renesse in [22] to cover the horizontal case considered in this paper. Roughly speaking the idea consists in expressing the viscosity solution of the level set equation as value function of suitable associated stochastic controlled systems. This is made more precise in the following result.

Theorem 6.10 ([22]). *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space and B be an m-dimensional Brownian motion adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$.*

Let $g : \mathbb{R}^N \to \mathbb{R}$ be a bounded and Hölder function. Let us consider $T > 0$. For any $(t, x) \in [0, T] \times \mathbb{R}^N$, we define

$$V(t,x) = \inf_{v\in\mathcal{A}} ess\sup_{\Omega} g(\xi^{t,x,\nu}(T)), \tag{6.8}$$

where

$$\mathcal{A} = \{\nu \in Sym(m) |\ \nu \geq 0 \quad I_m - \nu^2 \geq 0, \quad Tr(I_m - \nu^2) = 1\}, \tag{6.9}$$

and $\xi^{t,x,\nu}$ are the solution of the stochastic controlled dynamics

$$\begin{cases} d\xi^{t,x,\nu}(s) = \sqrt{2}\sigma^T(\xi^{t,x,\nu}(s)) \circ dB^\nu(s), & s \in (t, T], \\ dB^\nu(s) = \nu(s) dB_m(s), & s \in (t, T], \\ \xi^{t,x,\nu}(t) = x. \end{cases} \tag{6.10}$$

Let also us assume that the matrix σ is an $m \times N$ Hörmander matrix with smooth coefficients and that σ and $\sum_{i=1}^m \nabla_{X_i} X_j$ are Lipschitz functions. Then the value function V defined in (6.8) *is a viscosity solution of the level set equation* (6.7)*. Here $\circ$ denote the Stratonovich differential.*

The approach to prove the theorem above is classical and follows the ideas introduce in [27] for the standard (Euclidean) evolution, which means that require to first consider a more regular problem known as p-regularizing problem, which we introduce in the next section.

At characteristic points, (i.e. $|D_{\mathcal{X}} u| = 0$) the approach in [22] yields a discontinuous nonlinearity (different for sub and supersolutions): for subsolutions we get that u_t equals the minimal eigenvalue of the horizontal Hessian, while for supersolutions we get the maximal eigenvalue. The corresponding control would be a projection on the respective eigenspace. One result here (see Remark 6.24) is a refinement of this conclusion: the optimal control is not unique if both eigenvalues are equal, otherwise it is the projection on the eigenspace of the *maximal* eigenvalue.

6.5 The p-regularizing problem.

To show directly that the value function V defined in (6.8) is a viscosity solution on Equation (6.7) is extremely hard, the two main difficulties being that the PDE is highly degenerate and the value function is a L^∞-norm. The idea from [27] is to consider the value function associated to L^p-norm approximating (at leats on bounded sets) (6.8), which we indicate by V_p, and show that this new value function solves in the viscosity sense the corresponding PDE. Then we can recover the result given in Theorem 6.10 by a limit-argument as $p \to +\infty$. Note that the p-problem associated to the new value function V_p is far more regular than the level set equation (6.7) i and in fact the associated p-Hamiltonian has no points of discontinuity. For all $1 < p < +\infty$ we define the p-value function as

$$V_p(t,x) := \inf_{\nu \in \mathcal{A}} \mathbb{E}[g^p(\xi^{t,x,\nu}(T))]^{\frac{1}{p}}. \tag{6.11}$$

where $\xi^{t,x,\nu}$ and $\mathcal{A}$ are defined as in Theorem 6.10. Then one can show (see [22]) that V_p solves in the viscosity sense

$$\begin{cases} -(V_p)_t + H_p(V_p, DV_p, D^2V_p) = 0, & x \in \mathbb{R}^n,\ t \in [0,T), \\ V_p(T,x) = g(x), & x \in \mathbb{R}^n, \end{cases} \tag{6.12}$$

where the p-Hamiltonian H_p is defined as

$$H_p(r,q,M) := \sup_{\nu \in \mathcal{A}} \Big[-(p-1) r^{-1} Tr[\nu\nu^T qq^T] + Tr[\nu\nu^T M] \Big], \tag{6.13}$$

and $q \in \mathbb{R}^m$ and $M \in Sym(m)$ where σ is the matrix in (6.1).

The aim of this article is to find information on the structure of the optimal control for the p-Hamiltonian associated to the p-Hamilton-Jacobi equation regularising the the level set equation for the evolution by horizontal mean curvature flow.

For sake of simplicity, we now introduce the following function

$$h_p(r,q,M,\nu) := -(p-1) r^{-1} Tr[\nu\nu^T qq^T] + Tr[\nu\nu^T M], \tag{6.14}$$

so that the p-Hamiltonian can be rewritten as

$$H_p(r, q, M) = \sup_{\nu \in \mathcal{A}} h_p(r, q, M, \nu). \tag{6.15}$$

Remark 6.11. *Without the Remark 6.9, the function in* (6.14) *would have another term which will depends on ν and a second vector d. This term may lead some additional difficulties to search the optimal control.*

As proved in [49], finding the infimum of the optimal controls for the solution V_p is equivalent to optimise the supremum of the Hamiltonian as defined in (6.13). Then we will concentrate in finding the structure of the optimal controls giving the supremum in (6.15), for p large enough.
We conclude this section with the following two remarks, which will be very useful for the later results.

Remark 6.12. *It is possible (see [27]) to rewrite the set of admissible controls as*

$$\mathcal{A} = Co\{\nu = I_m - a \otimes a, \mid a \in \mathbb{R}^m, \quad |a| = 1\} \tag{6.16}$$

where with Co we mean the convex hull.

Remark 6.13. *Note that it is possible to consider $r > 0$. As we are not interested in the value function itself, but only in its level sets, we can use as initial datum a positive function, e.g. $u_0(x) = 1 + \tanh(dist(x, \Sigma_0))$. By the comparison principle, this will remain positive. Note that the level of interest here will be the 1-level set, not the zero level set, which is empty. For comparison principles for a large set of hypersurfaces in the Heisenberg group we refer to [32].*

6.6 Optimal control for the p-Hamiltonian in $\mathbb{H}^1$

In this section we will prove the main results of the paper. We will focus on the optimal control for the p-Hamiltonian defined in (6.13) and, in order to keep

the computations easier, we will consider only the case of the 1-dimensional Heisenberg group, introduced in Example 6.3. We will divide our investigation in two separate cases: the case $q \neq 0$ and the case $q = 0$. Remember that the case $q = 0$ corresponds to the case when the horizontal gradient vanishes, which is associated to the characteristic points introduced in Definition 8.11. Let us recall that, in the case of the 1-dimensional Heisenberg group, $N = 3$ and $m = 2$.

Next we introduce the main idea: note that it is possible to express any admissible control as $\nu = I_2 - a \otimes a$ with $|a| = 1$ (see (6.16)). Moreover a generic unit vector a can be expressed by rotating any given fixed direction. Therefore, fixed an initial direction, to maximise the supremum in (6.13) on the set of all admissible controls $\mathcal{A}$ can be reduced to maximise the same function h_p among the rotational angle $\theta \in [0, 2\pi)$, which will be much easier. Since this can be done starting from any fixed direction, we will rotate exactly the direction which we know to be associated to the optimal problem for the limit case $p = +\infty$, that is $\frac{q}{|q|}$, (remember q is the gradient variable). For more details on the optimal control in the case $p = +\infty$ we refer to [22]. This of course cannot be done whenever $q = 0$, then we will treat that case later, rotating a different starting direction.

Case $q \neq 0$: non-characteristic points.

Now let us fix the variables r,q,M with $q \neq 0$. For sake of simplicity we also assume that M is a diagonal matrix, i.e. there exist $\lambda_1, \lambda_2 \in \mathbb{R}$ such that

$$M = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}. \tag{6.17}$$

In the next remark we highlight that the assumption above on M is indeed not restrictive.

Remark 6.14. *This optimisation is taken at a fixed point x and depending on variable r, q, and M. When writing these quantities explicitly, i.e. by specifying*

their entries, we implicitly refer to a coordinate system on $\mathbb{R}^2$. *We may choose an orthonormal coordinate system at this point. For simplicity we choose it in such a way that the symmetric matrix* M *is diagonal. Note that for* h_p *as in* (6.14) *it holds that*

$$h_p(r, q, M, \nu) = h_p(r, \mathcal{O}q, \mathcal{O}M\mathcal{O}^T, \mathcal{O}\nu)$$

for any orthonormal matrix $\mathcal{O}$, *i.e. such that* $\mathcal{O}\mathcal{O}^T = I$ *For two vectors* v_1 *and* v_2 *we have*

$$\mathcal{O}v_1(\mathcal{O}v_2)^T = \mathcal{O}(v_1 v_2^T)\mathcal{O}^T,$$

and for two matrices A *and* B *we have*

$$Tr\big((\mathcal{O}A\mathcal{O}^T)(\mathcal{O}B\mathcal{O}^T)\big) = Tr(\mathcal{O}(AB)\mathcal{O}^T) = Tr(\mathcal{O}^T\mathcal{O}(AB)) = Tr(AB).$$

We now introduce the following unit vector in polar coordinates:

$$n_\infty := \frac{q}{|q|} = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}, \tag{6.18}$$

for a suitable associated angle $\alpha \in [0, 2\pi)$ fixed. We now express any admissible control $\nu \in \mathcal{A}$ as

$$\nu = \nu_\theta \in \mathcal{A} \iff \nu_\theta = I_2 - n_\theta \otimes n_\theta, \tag{6.19}$$

where

$$n_\theta := R_\theta n_\infty = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} = \begin{bmatrix} \cos(\theta+\alpha) \\ \sin(\theta+\alpha) \end{bmatrix}, \quad \text{for } \theta \in [0, 2\pi). \tag{6.20}$$

Using (6.19), we can rewrite all the admissible control $\nu \in \mathcal{A}$ as

$$\nu_\theta := I_2 - n_\theta \otimes n_\theta = \begin{bmatrix} \sin^2(\theta+\alpha) & -\sin(\theta+\alpha)\cos(\theta+\alpha) \\ -\sin(\theta+\alpha)\cos(\theta+\alpha) & \cos^2(\theta+\alpha) \end{bmatrix}. \tag{6.21}$$

Moreover $\nu_\theta = \nu_{\theta+\pi}$, then we can restrict our attention to $\theta \in [0, \pi)$.

It is also easy to check that ν_θ is indeed an admissible control-matrix, in fact

it is a symmetric projection matrix, in fact: $\nu_\theta = \nu_\theta^T$, which trivially implies $\nu_\theta \nu_\theta^T = \nu_\theta^2$. Moreover for any generic vector $q \in \mathbb{R}^2$, we have

$$\nu_\theta^2 q = \nu_\theta(q - n_\theta < n_\theta, q >) = \nu_\theta q - < n_\theta, q > (n_\theta - n_\theta < n_\theta, n_\theta >) = \nu_\theta q.$$

Given a diagonal matrix M as in (6.17), and using that ν_θ is symmetric and a projection matrix, then

$$Tr(\nu_\theta \nu_\theta^T M) = Tr(\nu_\theta^2 M) = Tr(\nu_\theta M) = \lambda_1(\nu_\theta)_{11} + \lambda_2(\nu_\theta)_{22}, \tag{6.22}$$

where by $(\nu_\theta)_{ij}$ we indicate the coefficient in position (i, j) of the matrix ν_θ.

The next remarks will be useful for the later proofs of the main results.

Remark 6.15. *Given any admissible control ν_θ, expressed by* (6.21)*, for every given vector $q \in \mathbb{R}^2$, we have*

$$\nu_\theta q = q - n_\theta < n_\theta, q > .$$

Remark 6.16. *For all $q \in \mathbb{R}^2$, we deduce*

$$Tr(\nu_\theta \nu_\theta^T q q^T) = Tr(\nu_\theta q \otimes \nu_\theta q) = |\nu_\theta q|^2. \tag{6.23}$$

Remark 6.17. *Given any $q \in \mathbb{R}^2$ and by using* (6.23)*, we have*

$$|\nu_\theta q|^2 = |q|^2 - 2 < q, n_\theta >< q, n_\theta > + < q, n_\theta >^2 = |q|^2 - < q, n_\theta >^2 .$$

We can now find the structure of the optimal controls for large p.

Theorem 6.18. *Let us consider the p-Hamiltonian H_p introduced in* (6.13)*. Fixed r, q and M and assume that $q \neq 0$ and that M is a diagonal matrix as in* (6.17)*. Then, for large p, the optimal control is $\overline{\nu} = \nu_{\overline{\theta}}$, where ν_θ is defined in* (6.21) *and*

$$\overline{\theta} = \overline{C}\,\frac{1}{p} + O\left(\frac{1}{p^2}\right), \tag{6.24}$$

with $\overline{C} = \frac{C_1}{C_2}$, $C_1 = (\lambda_1 - \lambda_2)\sin 2\alpha$, α is defined in (6.18)*, and $C_2 = 2r^{-1}|q|^2 > 0$.*

Proof. By (6.22) and Remark 6.17 we can rewrite the function h_p introduced in (6.14) as

$$h_p(r,q,M,\nu_\theta) = -(p-1)r^{-1}(|q|^2 - < q, n_\theta >^2) \tag{6.25}$$

$$+ \lambda_1(\nu_\theta)_{11} + \lambda_2(\nu_\theta)_{22}, \tag{6.26}$$

where ν_θ is any admissible control expressed as in (6.21). We observe that the first term in (6.25) can be written as

$$-(p-1)r^{-1}(|q|^2 - < n_\theta, q >^2) = -(p-1)r^{-1}|q|^2(1 - < n_\theta, n_\infty >^2),$$

where n_∞ is the vector introduced in (6.18) (remember that n_∞ depends on the fixed vector q). Then h_p becomes

$$h_p(x,r,q,M,\nu_\theta) = (p-1)r^{-1}|q|^2(1 - < n_\theta, n_\infty >^2) + \lambda_1 \sin^2(\alpha+\theta) + \lambda_2 \cos^2(\alpha+\theta).$$

Recalling the definitions of n_∞ and n_θ, given respectively in (6.18) and in (6.20), we compute

$$\begin{aligned} < n_\theta, n_\infty >^2 &= (\cos(\alpha+\theta)\cos\alpha + \sin(\alpha+\theta)\sin\alpha)^2 \\ &= (\cos^2\alpha\cos\theta - \sin\alpha\cos\alpha\sin\theta + \sin\alpha\cos\alpha\sin\theta + \sin^2\alpha\cos\theta)^2 \\ &= (\sin^2\alpha + \cos^2\alpha)^2\cos^2\theta = \cos^2\theta. \end{aligned}$$

Thus $1 - < n_\theta, n_\infty >^2 = 1 - \cos^2\theta = \sin^2\theta$, and the function h_p simplify as below:

$$h_p(x,r,q,M,\nu_\theta) = -(p-1)r^{-1}|q|^2\sin^2\theta + \lambda_1\sin^2(\theta+\alpha) + \lambda_2\cos^2(\theta+\alpha). \tag{6.27}$$

We need to find the supremum of h_p among all the admissible controls, i.e. among all $\theta \in [0,\pi)$, hence we look at the stationary points.
For sake of simplicity, fixed x, r, q, M, we introduce the following notation:

$$f_p(\theta) := h_p(x,r,q,M,\nu_\theta).$$

Taking the derivative, we find

$$\begin{aligned} f_p'(\theta) =& -2(p-1)r^{-1}|q|^2 \sin\theta\cos\theta + 2(\lambda_1-\lambda_2)\cos(\alpha+\theta)\sin(\alpha+\theta) \\ =& -(p-1)r^{-1}|q|^2 \sin 2\theta + (\lambda_1-\lambda_2)\sin(2\theta+2\alpha). \end{aligned}$$

Note that for large p the stationary points occur for θ near 0 or θ near $\frac{\pi}{2}$. Looking at the function f_p we can see that for $\theta \approx 0$ we get the maximum while for $\theta \approx \frac{\pi}{2}$ we select the minimum. To find the zero of f_p' we introduce a suitable linearisation for the derivative function f_p'. The previous remark suggests us the following ansatz:

$$\theta = \frac{\beta}{p}.$$

We are now going to use the Taylor expansion of $\cos\frac{\beta}{p}$ and $\sin\frac{\beta}{p}$, as $p \to +\infty$,

$$\cos\left(\frac{\beta}{p}\right) = 1 + O\left(\frac{1}{p^2}\right) \quad \text{and} \quad \sin\left(\frac{\beta}{p}\right) = \theta + O\left(\frac{1}{p^3}\right).$$

Using the ansatz $\theta = \frac{\beta}{p}$ and the above Taylor's expansions, f_p' can be rewritten as

$$\begin{aligned} f_p'(\theta) =& -(p-1)r^{-1}|q|^2\left(\theta + O\left(1/p^3\right)\right) \\ &+ (\lambda_1-\lambda_2)\left(\sin 2\alpha\left(1+O\left(1/p^2\right)\right) + \cos 2\alpha\left(\theta + O\left(1/p^3\right)\right)\right) \\ =& -C(p-1)\theta + (\lambda_1-\lambda_2)\sin 2\alpha + +\theta\cos 2\alpha + O\left(\frac{1}{p^2}\right), \end{aligned}$$

where $C = r^{-1}|q|^2$. Then, for p large, $f_p'(\theta_p) = 0$ if and only if

$$\theta_p = \frac{(\lambda_1-\lambda_2)\sin 2\alpha}{C(p-1)-\cos 2\alpha} + O\left(\frac{1}{p^2}\right).$$

Now, we set $C_1 = (\lambda_1-\lambda_2)\sin 2\alpha$, $C_2 = C = r^{-1}|q|^2$ and $C_3 = -2C - 2\cos 2\alpha$, then we can rewrite θ_p in the following more compact form:

$$\theta_p = \frac{C_1}{C_2 p + C_3} = \frac{C_1}{C_3(\frac{C_2}{C_3}p+1)} = \frac{C_1}{C_3}\frac{1}{\frac{C_2}{C_3}p+1}. \tag{6.28}$$

Finally set $x = \frac{1}{\frac{C_2}{C_3}p}$, to conclude we need to need only to apply the Taylor's expansion, near $x = 0$, for the function $\frac{1}{\frac{1}{x}+1} = \frac{x}{x+1}$ (that is $\frac{x}{x+1} = x + O(x^2)$),

then (6.28) can be rewritten as

$$\theta_p = \frac{C_1}{C_3}\left[\frac{C_3}{C_2 p} + O\left(\frac{1}{p^2}\right)\right] = \overline{C}\,\frac{1}{p} + O\left(\frac{1}{p^2}\right),$$

with $\overline{C} = \frac{C_1}{C_2}$. So $\beta = \overline{C} + O\left(\frac{1}{p}\right)$.

□

Let us now make a few remarks on the previous result on some special cases.

Remark 6.19 (Case $\lambda_1 = \lambda_2$.)**.** *Note that λ_1 and λ_2 are the eigenvalues of the matrix of the second order derivatives. Whenever $\lambda_1 = \lambda_2 =: \lambda$, then*

$$f_p(\theta) = -(p-1)r^{-1}|q|^2 \sin^2\theta + \lambda,$$

the obviously, for all $p > 1$, the maximum is attained for $\theta = 0$, This implies that in this case, at non characteristic points (i.e. $|q| \neq 0$), the optimal control for the p-problem is actually the same of the optimal control for the limit problem $p = +\infty$, i.e.

$$\overline{\nu} = I_2 - n_0 \otimes n_0,$$

where n_0 is the horizontal normal, see [22] for more details for the case $p = +\infty$.

Remark 6.20 (Case $\alpha = 0$ or $\alpha = \frac{\pi}{2}$.)**.** *If $\alpha \in \left\{0, \frac{\pi}{2}\right\}$ again the maximum is attained in $\theta = 0$. Remember that the case $\alpha = 0$ corresponds to the case case where the horizontal gradient in $\mathbb{R}^2$ point is $(1,0)^t$, then the horizontal normal points in the direction of the first vector field X_1, while in the case $\alpha = \frac{\pi}{2}$ the horizontal normal points in the direction of the second vector field X_2.*

To conclude this section let give the general result, removing the additional assumption of M diagonal.

Theorem 6.21. *Let us consider the p-Hamiltonian H_p introduced in* (6.13). *Fixed r, q and M and assume that $q \neq 0$. Then, for large p, the optimal control is $\overline{\nu} = \nu_{\overline{\theta}}$, where ν_θ is defined in* (6.21) *and*

$$\overline{\theta} = \overline{C}\,\frac{1}{p} + O\left(\frac{1}{p^2}\right), \tag{6.29}$$

with the constant $\overline{C}$ depends only on the variable r, q and the eigenvalues of the matrix M. Note that the expansion is only valid for $p|q|^2$ large. For $p|q|^2 = O(\lambda_1)$ see Section 7.

Proof. This follows immediately by Theorem 6.18 and Remark 6.14. □

Case $q = 0$: characteristic points.

In the case that $q = 0$ the function h_p is simplified and it does not depend on p. In fact it has the form

$$h_p(r, 0, M, \nu) = h(M, \nu) = Tr[\nu\nu^T M].$$

Furthermore we observe that we can generate a general admissible control starting from the rotation of a generic unit vector. For sake of simplicity (see Remark 6.23 later) we fix as starting vector

$$n_\infty := \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

and we rotate it by a rotation matrix $R(\theta)$ as follows:

$$n_\theta := R(\theta)n_\infty = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix},$$

where $\theta \in [0, 2\pi)$. We define the admissible control as

$$\nu_\theta := I_2 - n_\theta \otimes n_\theta = \begin{bmatrix} \sin^2\theta & -\sin\theta\cos\theta \\ -\sin\theta\cos\theta & \cos^2\theta \end{bmatrix}. \tag{6.30}$$

Then, by recalling that ν_θ is a projection matrix we obtain immediately

$$h_p(r, 0, M, \nu_\theta) = \lambda_1 \sin^2\theta + \lambda_2 \cos^2\theta.$$

Theorem 6.22. *Let us consider the p-Hamiltonian introduced as in* (6.13)*, ν_θ defined as in* (6.21) *and fixed r, q and M. Let us denote by λ_1, λ_2 the eigenvalues of M. Assume that $q = 0$. Then, for all $p > 1$, the optimal control for the p-Hamiltonian is independent on p and it is given by $\overline{\nu} = \nu_{\overline{\theta}}$ where the control is defined in* (6.30) *and*

$$\overline{\theta} = \frac{\pi}{2} \quad \textit{whenever } \lambda_1 - \lambda_2 > 0,$$

$$\overline{\theta} = 0 \quad \textit{whenever } \lambda_1 - \lambda_2 < 0,$$

while for $\lambda_1 = \lambda_2 =: \lambda$, $h_p(r, 0, M, \nu_\theta) = \lambda$ is constant so all possible angle θ are associated to optimal controls.

.

Proof. First recall that by Remark 6.14 we can assume that M has the diagonal form given in (6.17). For fixed r and m and denoting $h_p(r, 0, M, \nu_\theta) =: f(\theta)$, we look at the stationary points for the case $\lambda_1 \neq \lambda_2$. Taking the derivatives, we have

$$\begin{aligned} f'(\theta) = 2\lambda_1 \sin\theta\cos\theta - 2\lambda_2 \sin\theta\cos\theta &= 2(\lambda_1 - \lambda_2)\sin\theta\cos\theta \\ &= (\lambda_1 - \lambda_2)\sin 2\theta. \end{aligned} \tag{6.31}$$

If $\lambda_1 = \lambda_2$ we obtain that $f'(\theta) = 0$ for all $\theta \in [0, 2\pi)$, i.e. the function f is constant. If $\lambda_1 \neq \lambda_2$ we note that

$$\sin 2\theta = 0 \Longleftrightarrow \theta = \frac{k\pi}{2} \quad \text{with } k \in \{0, 1, 2, 3\}. \tag{6.32}$$

In order to find the maximum values for $\lambda_1 \neq \lambda_2$ we can easily compute the second derivatives, that is

$$f''(\theta) = 2(\lambda_1 - \lambda_2)\cos 2\theta. \tag{6.33}$$

so the stationary points in $[0, \pi)$ are $\theta = 0$ and $\theta = \frac{\pi}{2}$. Then

1. for $\lambda_1 - \lambda_2 > 0$ it holds true $f''(\frac{\pi}{2}) = -2(\lambda_1 - \lambda_2) < 0$.

2. for $\lambda_1 - \lambda_2 < 0$ it holds true $f''(0) = -2(\lambda_1 - \lambda_2) < 0$,

and this concludes the proof.

□

Remark 6.23. *An easy computation shows that the optimal control is independent on the choice of the staring vector n_∞. In fact, choosing a generic unit vector $n_\infty = (\cos\alpha, \sin\alpha)^T$ we would have got the control in the form given in* (6.21) *and the same results found in Theorem 6.22 for the angle $\theta + \alpha$, then the optimal control would be exactly the same.*

Remark 6.24. *It is instructive to compare this with the formula derived in [22], i.e. $\nu = I_2 - a \otimes a$ where $|a| = 1$. There, in characteristic points, we get a projection on the on the eigenspace of the minimal eigenvalue for subsolutions and on that of the maximal eigenvalue for supersoutions. Here we see: The optimal control in a characteristic point is not unique if both eigenvalues are equal, otherwise it is the projection on the eigenspace of the* maximal *eigenvalue.*

6.7 Numerical computations and illustrations

In this section we give some computed examples for the optimal control to illustrate the behaviour of the controlled random walk, in particular in the vicinity of critical points.

Consider the π-periodic function f_p from (6.27). First note that q and p appear only as $p|q|^2$, so f_p can be expressed as a function of $p|q|^2$. This means that in the following graphs the limit $p \to \infty$ corresponds to a scaling in the $|q|$-direction. Moreover $|q| \to 0$ means moving towards a characteristic point, $p \to \infty$ away from it. As r^{-1} appears only multiplying p, we see that local convergence does not depend on the choice of the level set function. This is to be expected, as the limit evolution is geometric. Moreover, note that r, the

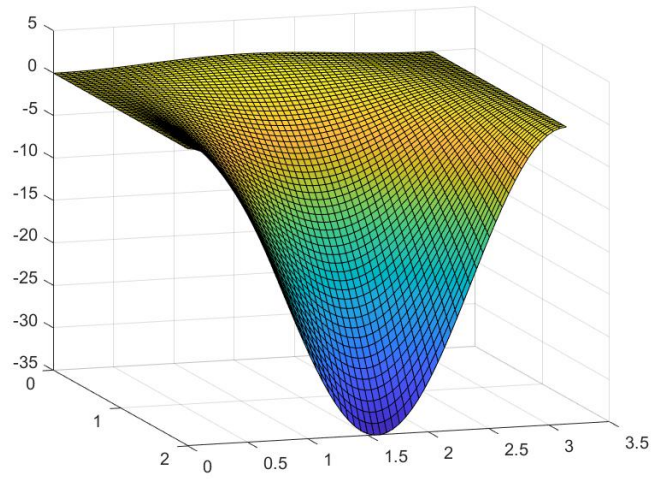


Figure 6.1: $f_{10}(\theta)$ for $p = 10$ and $\alpha = 0$.

value function, is constant on a level set. First let us consider Figure 6.1, which plots f_p for $p = 10$, $\alpha = 0$, $\lambda_1 = 1$, $\lambda_2 = 0$ and $r = 1$. The q_1-axis points in the direction of eigenvector corresponding to the eigenvalue λ_1. Geometrically this means that the projected horizontal eigenvector points exactly in the direction of the largest eigenvalue of the horizontal Hessian. In this case we see that for a distinct value of q the maximum jumps from 0 to $\pi/2$, corresponding to either maximizing the second term in f_p, i.e. making $|\sin(\theta)| = 1$, or maximizing the first term, making $\sin(\theta) = 0$.

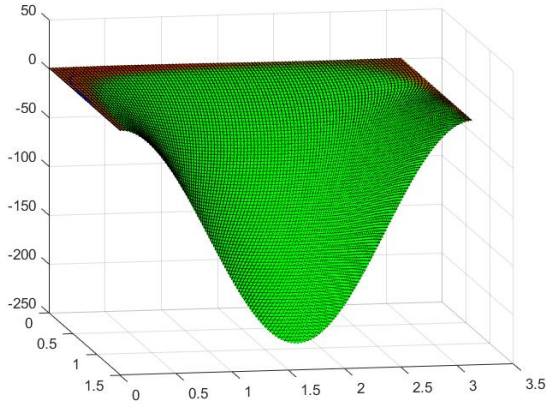


Figure 6.2: $f_{100}(\theta)$ for $p = 100$ and $\alpha = 0$. We remark that the function goes faster to $-\infty$.

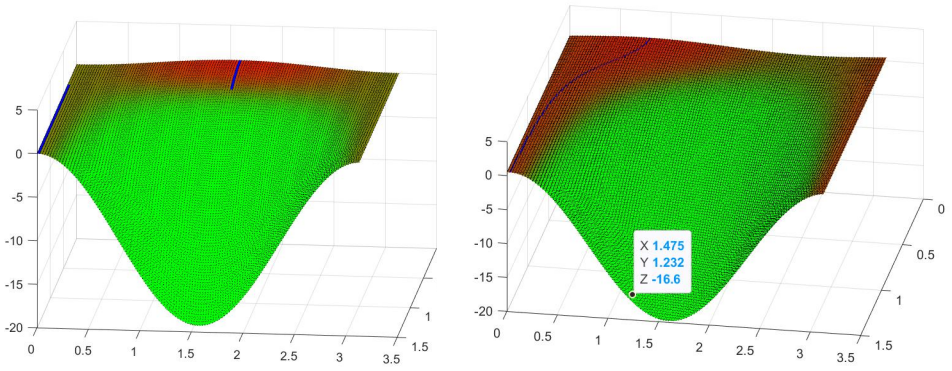


Figure 6.3: On the left $f_{10}(\theta)$ (i.e. $p = 10$) for $\alpha = 0$, on the right for $\alpha = \pi/4$, the branch of maximizing θ is in both cases in blue. Note that the maximising angle is discontinuous in the left picture and continuous on the right picture.

For $\alpha \neq 0$, however, there is a continuous branch of maximizing angles. The situation is illustrated in Figure 6.2.

It turns out that the case on the left, where the q-vector points in the direction of the eigenvector with the largest eigenvalue of the horizontal Hessian, is the only case where such a singularity occurs. Here, in the limit $p \to \infty$,

for the critical point, the optimal control is projection on the eigenspace for the maximal eigenvalue, i.e. parallel to q, while immediately away from the critical point the optima control projects *orthogonally* to q.

We see in Figure 6.3 that this vortex-like discontinuity occurs for $\alpha = \pi$ (equivalent to $\alpha = 0$) for small (non vanishing) $|q|$.

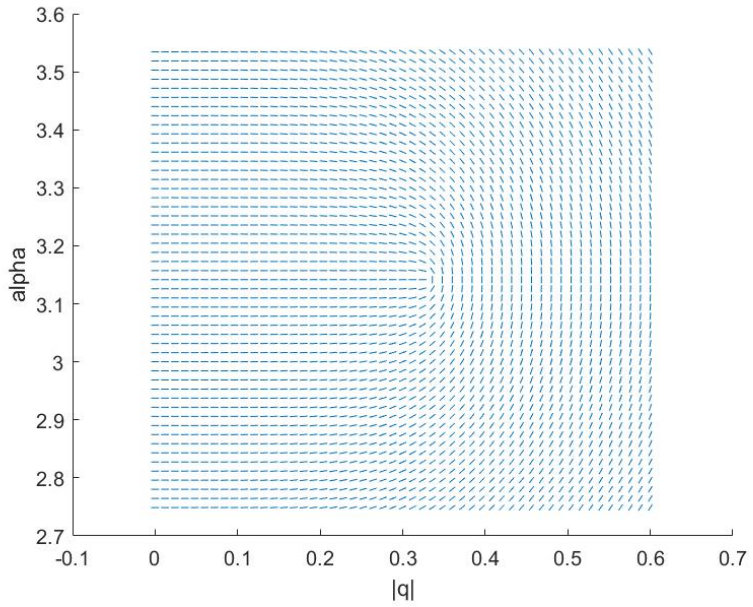


Figure 6.4: Direction of optimal control (i.e. allowed direction of motion of the controlled process) for $p = 10$, $M = diag(1, 0)$. Note the singularity when $|q|^2 p = \lambda_1 - \lambda_2 = 1$. In this regime, the asymptotic expansion of Section 6 is not valid.

For $\lambda_1 = \lambda_2$, this discontinuity is moved into the characteristic point, see below the case of the unit sphere (see Figure 6.5).

Finally, in order to illustrate the convergence of the optimal control, we plot in Figure 6.4 for the same parameters an entire period for both $p = 10$ and $p = 50$.

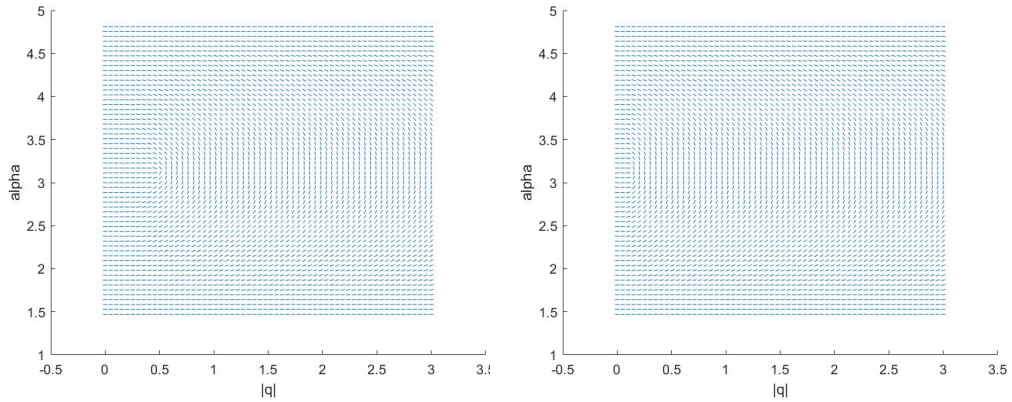


Figure 6.5: Direction of optimal control (i.e. allowed direction of motion of the controlled process) for $M = diag(1, 0)$, i.e. $\lambda_1 = 1$ and $\lambda_2 = 0$. On the left picture: $p = 5$; while on the right picture: $p = 30$.

Let us apply this to two specific surfaces, the unit sphere centred at $(0, 0, 1)$ and and ellipsoid with same center, but given by $2x^2 + y^2 + (z - 1)^2 = 1$. Due to the lower symmetry, the singularity of the control field for the ellipsoid is moved away from the characteristic point and clearly visible. We have plotted the horizontal normal as dashed line and the control field as solid line, both in a 3-dimensional perspective and the projection of the vectors on the $x - y$-plane.

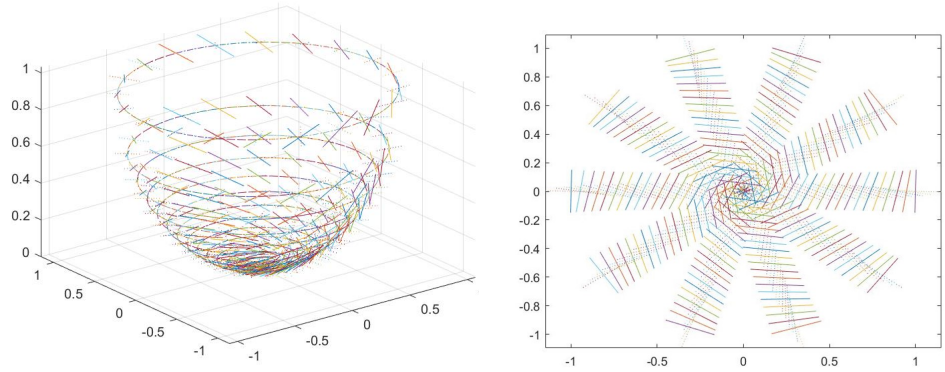


Figure 6.6: Direction of optimal control (solid) and horizontal normal (dashed) for the unit sphere with characteristic point $(0, 0, 0)$. On the left these vectors are represented in 3D, while on the right we see their 2D-projection on the $x - y$-plane (for $p = 5$).

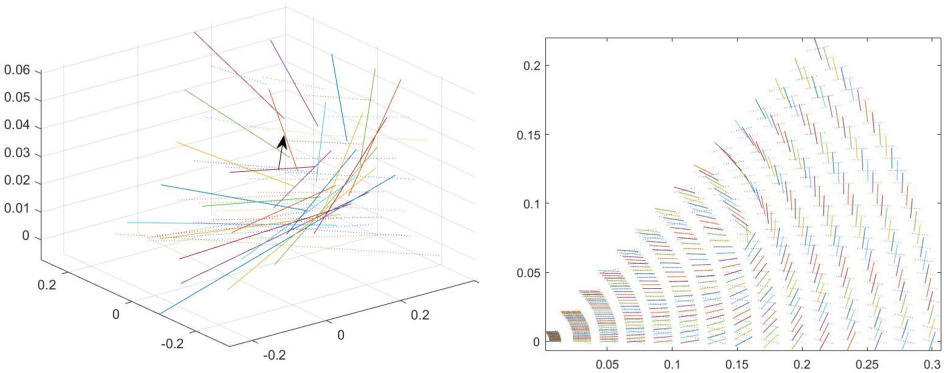


Figure 6.7: Direction of optimal control (solid) and horizontal normal (dashed) for the ellipsoid $2x^2 + y^2 + (z - 1)^2 = 1$ with characteristic point $(0, 0, 0)$. On the left these vectors are represented in 3D, while on the right we see their 2D-projection on the $x - y$-plane (for $p = 5$). The arrow points out the singularity. At this point, the horizontal normal points in the direction of an eigenvector of the horizontal Hessian.

Chapter 7

Optimal controls for the $\mathbb{H}^1$ approximated

7.1 Introduction

The evolution by mean curvature flow (MCF) has been studied extensively and it has many applications in image processing, see e.g. [13]. We say that a hypersurface evolves by MCF if it contracts in the normal direction with normal velocity proportional to its mean curvature. It is well-known that this evolution may develop singularities in finite time. To deal with such a singularities, many generalised approaches to study this evolution have been developed. In 1991, Chen, Giga and Goto [11] and, independently Evans and Spruck [31] introduced the *level set approach*, which consists in studying the evolving hypersurfaces as level sets of (viscosity) solutions of suitable associated nonlinear PDEs. In this paper we are interested in a degenerate version of this evolution, called evolution by horizontal mean curvature flow (HMCF): we consider a hypersurface embedded in a sub-Riemannian geometry, then the evolution contracts in the direction of the so called horizontal normal proportionally to its horizontal curvature. We consider the level set approach which is now associated to a parabolic PDE, which is far more degenerate than

in the standard case. The approach here is to interpret the solution of the level set equations as value function of suitable associated stochastic control problem: this approach has been developed by Cardaliaguet, Quincampoix and Buckdahn in [27] and contemporaneously but independently by Soner and Touzi [49] for the standard (Euclidean) case and generalised then by Dirr, Dragoni and von Renesse in [22] for the horizontal mean curvature flow in general sub-Riemannian structures. The horizontal mean curvature flow is a generalisation of the Euclidean mean curvature evolution to hypersurfaces embedded in sub-Riemannian structure. This evolution has been studied with different approaches by several authors. In particular in this chapter we focus on the stochastic approach in [22] together with the Riemannian approximation approach used e.g. in [13], [14]. In this chapter we will study the optimal control of approximated Riemannian problem and we will look the relation which exists between these controls and the sub-Riemannian ones in the Heisenberg group.

7.2 Preliminaries

The Heisenberg group

We briefly recall some basic geometrical definitions, which are key for the purpose of the paper. For more definitions and properties on sub-Riemannian geometries we refer to [4], and to [42] for the specific case of Carnot groups (see Chapter 1 for further details). We start giving the definition of the Heisenberg group.

Definition 7.1 (Heisenberg group)**.** *The 1-dimensional Heisenberg group is the sub-Riemannian structure induced on* $\mathbb{R}^3$ *by the vector fields*

$$X_1(x) = \begin{pmatrix} 1 \\ 0 \\ -\frac{x_2}{2} \end{pmatrix} \quad \textit{and} \quad X_2(x) = \begin{pmatrix} 0 \\ 1 \\ \frac{x_1}{2} \end{pmatrix}, \quad \forall\, x = (x_1, x_2, x_3) \in \mathbb{R}^3.$$

For a formal definition of the Heisenberg group and the connection between its structure as non-commutative Lie group and its manifold structure we refer to [4].

For later use we also introduce the matrix associated to the vector fields

$$\sigma(x) = \begin{bmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \end{bmatrix}, \quad \forall\, x = (x_1, x_2, x_3) \in \mathbb{R}^3. \tag{7.1}$$

We now introduce the Riemannian approximation of the Heisenberg group, which is crucial for our results (see [4] for further details). It is possible to complete the distribution $\mathcal{H}$ by adding a vector field X_3, so that X_1, X_2 and X_3 together generate the whole $\mathbb{R}^3$ at any points. The geometric structure induced on the (Euclidean) $\mathbb{R}^3$ by the vector fields X_1, X_2 and εX_3 for all $\varepsilon > 0$ is called *Riemannian approximation* of the Heisenberg group: in this case

$$X_3(x) = (0, 0, 1)^T, \quad \forall\, x \in \mathbb{R}^3$$

and the associated matrix is

$$\sigma_\varepsilon(x) = \begin{bmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \\ 0 & 0 & \varepsilon \end{bmatrix}, \quad \forall\, x = (x_1, x_2, x_3) \in \mathbb{R}^3. \tag{7.2}$$

This technique is called Riemannian approximation since, as $\varepsilon \to 0^+$, then the geometry induced by Riemannian approximation converges, in sense of Gromov-Hausdorff (see [36] for further details), to the original sub-Riemannian geometry (as showed, as example, in [13]).

Evolution by horizontal mean curvature evolution

Given a smooth hypersurface $\Gamma \subset \mathbb{R}^3$, we indicate by $n_E(x)$ the standard Euclidean normal to Γ at the point x (for further details see Chapter 4).

Definition 7.2. *Given a smooth hypersurface* Γ*, the* horizontal normal *for the Heisenberg group is the renormalized projection of the Euclidean normal on the horizontal space, i.e.*

$$n_0(x) = \frac{\alpha_1(x)X_1(x) + \alpha_2(x)X_2(x)}{\sqrt{\alpha_1^2(x) + \alpha_2^2(x)}}. \tag{7.3}$$

The main difference between the horizontal normal and the standard Euclidean normal is that the first may not exist even for smooth hypersurfaces: in fact at some points the horizontal normal is not defined meanwhile the Euclidean one exists. These points are called *characteristic points* (see e.g. [22] for further details). All compact hypersurfaces in the Heisenberg group have at least one characteristic points, so such points are unavoidable.

We recall that for every smooth hypersurface the *mean curvature* is defined as the divergence of the Euclidean normal. Similarly, for every smooth hypersurface, we can now introduce the horizontal mean curvature as the horizontal divergence of the horizontal normal.

Obviously the horizontal mean curvature is never defined at characteristic points, since there the horizontal normal does not exist. It is possible to adapt the definition of mean curvature flow as stated in [28] to the case of Carnot-type geometry (for further information we refer to [22]). The evolution by horizontal mean curvature flow (briefly HMCF) describes the motion of a hypersurface contracting with normal velocity proportional to the horizontal mean curvature. Exactly as for the corresponding Euclidean motion, this evolution develops singularities and needs to be interpreted also at points where the horizontal mean curvature is not defined. For sake of simplicity in this paper we focus only on the case of the 1-dimensional Heisenberg group, so all definitions and notions are written adapted to this simpler case. Consider a hypersurface Γ parametrized as zero level set, i.e.

$$\Gamma = \big\{x \in \mathbb{R}^3 | g(x) = 0\big\},$$

for some continuous function $g : \mathbb{R}^3 \to \mathbb{R}$. We say for $t > 0$ a hypersurface Γ_t is a generalised evolution by horizontal mean curvature flow of Γ if $\Gamma_0 = \Gamma$ and Γ_t can be parametrized as the zero-level set of a (continuous) function $V : [0, +\infty) \times \mathbb{R}^3 \to \mathbb{R}$ which solves in the viscosity sense

$$V_t = Tr((\mathcal{X}^2 V)^*) - \left\langle (\mathcal{X}^2 V)^* \frac{\mathcal{X}V}{|\mathcal{X}V|}, \frac{\mathcal{X}V}{|\mathcal{X}V|} \right\rangle, \tag{7.4}$$

where $\mathcal{X}V$ is the so called horizontal gradient, that is

$$\mathcal{X}V := (X_1 V, X_2 V)^T$$

and $(\mathcal{X}^2 V)^*$ is the symmetrized horizontal Hessian, that is

$$((\mathcal{X}^2 V)^*)_{ij} := \frac{X_i(X_j V) + X_j(X_i V)}{2} \quad \text{with } i, j = 1, 2.$$

Similarly one can introduce the same evolution and the corresponding level set approximation for the Riemannian approximation of the Heisenberg group: in that case one has to replace in the PDE (7.4), the horizontal gradient and the symmetrized horizontal Hessian, respectively, by

$$\mathcal{X}^\varepsilon V := (X_1 V, X_2 V, \varepsilon X_3 V)^T$$

and

$$((\mathcal{X}^{\varepsilon 2} V)^*)_{ij} := \frac{X_i^\varepsilon(X_j^\varepsilon V) + X_j^\varepsilon(X_i^\varepsilon V)}{2} \quad \text{with } i, j = 1, 2, 3,$$

with $X_1^\varepsilon = X_1$, $X_2^\varepsilon = X_2$ and $X_3^\varepsilon = \varepsilon X_3$, for every $\varepsilon > 0$.

Associated stochastic control problem

As proved in [27] [48], [49] for the Euclidean and Riemannian cases and in [22] for the horizontal (sub-Riemannian) case, it is possible to connect the evolution by mean curvature flow to the value function of suitable associated stochastic optimal controlled dynamics. This allow to find a stochastic representation for the viscosity solution of horizontal mean curvature. Next we recall the main ideas of such approach. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered

probability space, B_i is an i-dimensional Brownian motion adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ with $i = 2, 3$, we recall that a *predictable* variable is a time-continuous stochastic process $\{\xi_t\}_{t\geq 0}$ defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, measurable with respect to the σ-algebra generated by all left-continuous adapted process. Given a smooth function $g : \mathbb{R}^3 \to \mathbb{R}$ (which parametrizes the starting hypersurface at time $t = 0$) we introduce the function $V : [0, T] \times \mathbb{R}^3 \to \mathbb{R}$ defined as

$$V(t, x) := \inf_{\nu\in\mathcal{A}} ess\sup_{\omega\in\Omega} g(\xi^{t,x,\nu}(T)(\omega)), \tag{7.5}$$

and $\xi^{t,x,\nu}$ is the solution of the stochastic dynamic

$$\begin{cases} d\xi^{t,x,\nu}(s) = \sqrt{2}\sigma^T(\xi^{t,x,\nu}(s))\nu(s)dB_2(s), \qquad s \in (t, T], \\ \xi^{t,x,\nu}(t) = x, \end{cases}$$

where the matrix σ is defined in (7.1) and

$$\mathcal{A} = \big\{\nu : [t, T] \to Sym(2) | \nu \geq 0,\ I_2 - \nu^2 \geq 0,\ Tr(I_2 - \nu^2) = 1\big\}. \tag{7.6}$$

Similarly, for $\varepsilon > 0$ we introduce the function $V^\varepsilon : [0, T] \times \mathbb{R}^3 \to \mathbb{R}$ defined by

$$V^\varepsilon(t, x) := \inf_{\nu\in\mathcal{A}_1} ess\sup_{\omega\in\Omega} g(\xi_\varepsilon^{t,x,\nu_1}(T)(\omega)), \tag{7.7}$$

where $\xi_\varepsilon^{t,x,\nu_1}$ is the solution of

$$\begin{cases} d\xi_\varepsilon^{t,x,\nu_1}(s) = \sqrt{2}\sigma_\varepsilon^T(\xi_\varepsilon^{t,x,\nu_1}(s))\nu_1(s)dB_3(s), \qquad s \in (t, T], \\ \xi_\varepsilon^{t,x,\nu_1}(t) = x, \end{cases}$$

σ_ε is the matrix defined in (7.2) and

$$\mathcal{A}_1 = \big\{\nu_1 : [t, T] \to Sym(3) | \nu \geq 0,\ I_3 - \nu^2 \geq 0,\ Tr(I_3 - \nu^2) = 1\big\}. \tag{7.8}$$

It is possible to show that the functions V and V^ε solve in the viscosity sense respectively the level-set equation for the evolution by HMCF and the level set

equation for the Riemannian approximation of the HMCF (see [22, 26]). Note also that the sets of controls (7.6) and (7.8) can be rewritten respectively as

$$\mathcal{A} = \{\nu^2|\ \nu \in \mathcal{A}\} = Co\{I_2 - a \otimes a|\ a \in \mathbb{R}^2\ \ |a| = 1\},$$

and

$$\mathcal{A}_1 = \{\nu_1^2|\ \nu_1 \in \mathcal{A}\} = Co\{I_3 - \overline{a} \otimes \overline{a}|\ \overline{a} \in \mathbb{R}^3\ \ |\overline{a}| = 1\},$$

see [27] for more details.

Next we introduce the p-regularising approximation of the functions V and V^ε.

Definition 7.3. *For $p > 1$, the p-value function associated to the value function* (7.5) *is defined as*

$$V_p(t,x) := \inf_{\nu \in \mathcal{A}} \mathbb{E}[|g(\xi^{t,x,\nu})(T)(\omega)|^p]^{\frac{1}{p}}.$$

and similarly we can introduce the following ε-p-regularising function, that is the p-value function associated to the value function (7.7),

$$V_p^\varepsilon(t,x) := \inf_{\nu_1 \in \mathcal{A}} \mathbb{E}[|g(\xi_\varepsilon^{t,x,\nu_1})(T)(\omega)|^p]^{\frac{1}{p}}.$$

The function V_p solves in viscosity sense the following PDE:

$$\begin{cases} -(V_p) + H_p(x, DV_p, D^2V_p) = 0, & t \in [0,T),\ x \in \mathbb{R}^3, \\ V_p(T,x) = g(x), & x \in \mathbb{R}^3 \end{cases} \tag{7.9}$$

where

$$H_p(x,q,M) := \sup_{\nu \in \mathcal{A}} \Big[-(p-1) r^{-1} Tr[\nu\nu^T qq^T] + Tr[\nu\nu^T M] \Big], \tag{7.10}$$

see [27]. Similarly for $\varepsilon > 0$ and $p > 1$ fixed, V_p^ε solves in the viscosity sense

$$\begin{cases} -(V_p^\varepsilon) + H_p^\varepsilon(x, DV_p^\varepsilon, D^2V_p^\varepsilon) = 0, & t \in [0,T),\ x \in \mathbb{R}^3, \\ V_p^\varepsilon(T,x) = g(x), & x \in \mathbb{R}^3 \end{cases} \tag{7.11}$$

where

$$H_p^\varepsilon(x,r,q,M) := H_p(x,r,q_\varepsilon,M_\varepsilon) = \sup_{\nu\in\mathcal{A}_1}\left[-(p-1)r^{-1}Tr[\nu\nu^T q_\varepsilon q_\varepsilon^T]+Tr[\nu\nu^T M_\varepsilon]\right], \tag{7.12}$$

where $\mathcal{A}_1$ is given in (7.8) and, for all $q = (q_1,q_2,q_3)^T \in \mathbb{R}^3$ and $M = (M_{ij})_{i,j=1}^3 \in Sym(3)$,

$$q_\varepsilon := \begin{bmatrix} q_1 \\ q_2 \\ \varepsilon q_3 \end{bmatrix} \quad \text{and} \quad M_\varepsilon := \begin{bmatrix} M_{11} & M_{12} & \varepsilon M_{13} \\ M_{12} & M_{22} & \varepsilon M_{23} \\ \varepsilon M_{13} & \varepsilon M_{23} & \varepsilon^2 M_{33} \end{bmatrix}. \tag{7.13}$$

7.3 Optimal controls for the ε-p-regularised mean curvature flow

In this section we want to study the ε-p-Hamiltonian (7.12), which is an approximation of the Hamiltonian studied in [26, 21]. The aim of this paper is to find the asymptotic structure for the optimal controls of the ε-p-Hamiltonian H_p^ε, for large p and small ε. For sake of simplicity we focus on the following specific subsequence

$$\varepsilon = \varepsilon(p) = \frac{1}{e^p \ln(p) p^2 (p-1)}. \tag{7.14}$$

Remark 7.4. *The reason of such a choice for $\varepsilon(p)$ is that we want to use the asymptotic results proved in this paper to show the convergence of the Riemannian approximating value function V^ε to the value function V solving the horizontal mean curvature flow, see [26]), where the same exact subsequence is introduced. For all the results (both in this paper but also in [26]) this particular choice of $\varepsilon(p)$ is not relevant and the proofs work for any $\varepsilon = \varepsilon(p)$ as soon as*

$$\lim_{p\to\infty} \varepsilon(p)^2 (p-1) = 0.$$

We remark that $\varepsilon(p)$ in (7.14) satisfies this condition.

We now introduce the function

$$h_p^\varepsilon(x,r,q,M,\nu) := h_p(x,r,q_\varepsilon,M_\varepsilon,\nu) = -(p-1)r^{-1}Tr[\nu\nu^T q_\varepsilon q_\varepsilon^T] + Tr[\nu\nu^T M_\varepsilon], \tag{7.15}$$

where q_ε and M_ε are defined in (7.13). The ε-p-Hamiltonian introduced in (7.12) can be rewritten as

$$H_p^\varepsilon(x,r,q,M) = \sup_{\nu\in\mathcal{A}_1} h_p^\varepsilon(x,r,q,M,\nu). \tag{7.16}$$

Remark 7.5. *Trivially, for q_1, q_2, $q_3 \in \mathbb{R}$ and $\varepsilon > 0$*

$$|q|^2 = q_1^2 + q_2^2 + q_3^3 \neq 0 \quad \Longleftrightarrow \quad |q_\varepsilon|^2 = q_1^2 + q_2^2 + \varepsilon^2 q_3^2 \neq 0.$$

To find the asymptotic structure of the controls we need to consider the following three different cases:

Case A: $q_1^2 + q_2^2 \neq 0$,

Case B: $q_1^2 + q_2^2 = 0$ and $q_3 \neq 0$,

Case C: $q_1^2 + q_2^2 + q_3^2 = 0$.

Note that the case B corresponds to the case of characteristic points for the horizontal problem, which are not degenerate for the correspondent Riemannian approximation; see [22] for more details on the optimal controls for the horizontal problem and the behaviour at characteristic points.

Case A: $q_1^2 + q_2^2 \neq 0$.

Let us recall that, given any generic vector $q_\infty = (q_1, q_2, 0)^T \in \mathbb{R}^3$, we can associate to it an unit vector

$$n_\infty := \frac{q_\infty}{|q_\infty|} = \begin{bmatrix} \frac{q_1}{\sqrt{q_1^2+q_2^2}} \\ \frac{q_2}{\sqrt{q_1^2+q_2^2}} \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{bmatrix}, \tag{7.17}$$

for some suitable $\alpha \in [0, 2\pi)$.

Since q is the variable associated to the gradient, the vector n_∞ corresponds to the optimal control for the case $p = +\infty$ and $\varepsilon = 0$ (see [21] and [22] for more details on this claim). In order to express a generic admissible control, we can rotate the vector n_∞, thus we can write any generic unit vector in $\mathbb{R}^3$ as

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{bmatrix} = \begin{bmatrix} \cos(\theta+\alpha) \\ \cos\phi\sin(\theta+\alpha) \\ \sin\phi\sin(\theta+\alpha) \end{bmatrix} =: n_{\phi,\theta}. \tag{7.18}$$

Note that we use the notation $n_{\phi,\theta}$ to highlight the dependence on the rotational angles θ and ϕ, while α depends on the gradient variable via (7.17). Using the approach first developed in [22] for the horizontal case, we are ready to prove a first asymptotic result for the optimal controls of H_p^ε.

Theorem 7.6. *Fixed $p > 1$ and $\varepsilon = \varepsilon(p)$ as in (7.14), and consider the Hamiltonian H_p^ε introduced in (7.16) with $r \in \mathbb{R}$, $q = (q_1, q_2, q_3)^T \in \mathbb{R}^3$ and $M \in Sym(3)$. Assume that $q_1^2 + q_2^2 \neq 0$. Then, for p sufficiently large, the optimal control of $H_p^{\varepsilon(p)}$ is given by $\overline{\nu}_p = \nu_{\overline{\theta}_p, \overline{\phi}_p}$, with*

$$\begin{cases} \overline{\theta}_p = \frac{C}{p} + O\left(\frac{1}{p^2}\right), \\ \overline{\phi}_p = O\left(\frac{1}{p^2}\right), \end{cases} \tag{7.19}$$

where C is a constant depending only on r, q, and M.

Proof. Without loss of generality, we can assume that the matrix M is diagonal, i.e. there exist $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ such that

$$M = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \tag{7.20}$$

see Remark 6.1, for details on this claim. This trivially implies that M_ε is diagonal as well (see (7.13)). Recall that the admissible controls are projection

matrices, thus by using the previous rotation argument to express all unit vectors in $\mathbb{R}^3$ by (7.18), we can deduce that

$$\nu \in \mathcal{A}_1 \quad \Longleftrightarrow \quad \nu_{\phi,\theta} = I_3 - n_{\phi,\theta} \otimes n_{\phi,\theta} \tag{7.21}$$

for some $\theta \in [0, \pi)$ and $\phi \in [0, 2\pi)$, which implies

$$(\nu_{\theta,\phi})_{11} = 1 - \cos^2(\theta + \alpha), \qquad (\nu_{\theta,\phi})_{22} = 1 - (\cos\phi \sin(\theta + \alpha))^2,$$

$$(\nu_{\theta,\phi})_{33} = 1 - (\sin\phi \sin(\theta + \alpha))^2,$$

where by $(\nu_{\theta,\phi})_{ij}$ we indicate the coefficient in position (i, j) of the matrix $\nu_{\theta,\phi}$. Moreover, for any generic vector $q \in \mathbb{R}^3$, we have

$$(\nu_{\theta,\phi})^2 q = \nu_{\theta,\phi}(q - n_{\theta,\phi} < n_{\theta,\phi}, q >) = \nu_{\theta,\phi} q.$$

By using the diagonal matrix structure of the matrix M_ε, and that $\nu_{\theta,\phi}$ is a symmetric projection matrix, we can easily deduce

$$\begin{aligned} Tr(\nu_{\theta,\phi}\nu_{\theta,\phi}^T M_\varepsilon) &= Tr((\nu_{\theta,\phi})^2 M_\varepsilon) = Tr(\nu_{\theta,\phi} M_\varepsilon) \\ &= \lambda_1(\nu_{\theta,\phi})_{11} + \lambda_2(\nu_{\theta,\phi})_{22} + \varepsilon^2\lambda_3(\nu_{\theta,\phi})_{33}. \end{aligned} \tag{7.22}$$

Let us recall that the vector q_ε can be written as

$$q_\varepsilon = \begin{bmatrix} \cos\beta\cos\alpha \\ \cos\beta\sin\alpha \\ \varepsilon\sin\beta \end{bmatrix} |q| = n_\infty^\varepsilon |q|, \tag{7.23}$$

where α is defined by (7.17) and β is associated to the third component $\frac{q_3}{|q|}$. Note also that

$$\begin{aligned} < n_\infty^\varepsilon, n_{\phi,\theta} > &= \varepsilon \sin\beta \sin\phi \sin(\theta + \alpha) + \cos\beta \sin\alpha \cos\phi \sin(\theta + \alpha) \\ &\quad + \cos\beta \cos\alpha \cos(\theta + \alpha), \end{aligned} \tag{7.24}$$

Recalling that (7.15) can be rewritten as

$$h_p(r, q, M, \nu_\theta) = -(p-1)r^{-1}(|q_\varepsilon| - < q_\varepsilon, n_{\theta,\phi} >^2) + \lambda_1(\nu_{\theta,\phi})_{11} + \lambda_2(\nu_{\theta,\phi})_{22} + \varepsilon^2\lambda_3(\nu_{\theta,\phi})_{33},$$

which implies, recalling (7.23) and (7.24)

$$h_p^\varepsilon(r,q,M,\nu_{\phi,\theta}) = -(p-1)r^{-1}|q|^2[(\cos^2\beta+\varepsilon^2\sin^2\beta)$$
$$-[\varepsilon(\sin\beta\sin\phi\sin(\theta+\alpha))+\cos\beta\sin\alpha\cos\phi\sin(\theta+\alpha)+\cos\beta\cos\alpha\cos(\theta+\alpha)]^2$$
$$+\lambda_1(1-\cos^2(\theta+\alpha))+\lambda_2(1-(\sin(\theta+\alpha)\cos\phi)^2)+\varepsilon^2\lambda_3(1-(\sin\phi\sin(\theta+\alpha))^2).$$

Following the approach in [21] we compute the stationary points which are indeed the points realising the supremum (see [21] for more details on this claim). Let us define the function $f_p^\varepsilon(\theta,\phi) := h_p^\varepsilon(r,q,M,\nu_{\theta,\phi})$, for r, q, M fixed. In order to computing the gradient we have to compute the partial derivatives of $f_p^\varepsilon(\theta,\phi)$:

$$\frac{\partial f_p^\varepsilon(\theta,\phi)}{\partial\theta} = 2(p-1)r^{-1}|q|^2\Big(\big(\varepsilon\sin\beta\sin\phi\sin(\theta+\alpha)+\cos\beta\sin\alpha\cos\phi\sin(\theta+\alpha)$$
$$+\cos\beta\cos\alpha\cos(\theta+\alpha)\big)\big(\varepsilon\sin\beta\sin\phi\cos(\theta+\alpha)+\cos\beta\sin\alpha\cos\phi\cos(\theta+\alpha)$$
$$-\cos\beta\cos\alpha\sin(\theta+\alpha)\big)\Big)+2(-\lambda_1+\lambda_2\cos^2\phi+\lambda_3\varepsilon^2\sin^2\phi)\cos(\theta+\alpha)\sin(\theta+\alpha),$$

and

$$\frac{\partial f_p^\varepsilon(\theta,\phi)}{\partial\phi} = 2(p-1)r^{-1}|q|^2\Big(\big(\varepsilon\sin\beta\sin\phi\sin(\theta+\alpha)+\cos\beta\sin\alpha\cos\phi\sin(\theta+\alpha)$$
$$+\cos\beta\cos\alpha\cos(\theta+\alpha)\big)\big(\varepsilon\sin\beta\cos\phi\sin(\theta+\alpha)-\cos\beta\sin\alpha\sin\phi\sin(\theta+\alpha)\big)\Big)$$
$$+2(-\lambda_2+\varepsilon^2\lambda_3)\sin\phi\cos\phi\sin^2(\theta+\alpha).$$

Since we are looking for the supremum of the Hamiltonian (7.16) as $p\to\infty$, we need

$$\lim_{p\to\infty}\Big(-pr^{-1}|q|^2\big(|n_\infty^\varepsilon|^2-<n_\infty^\varepsilon,n_{\phi,\theta}>^2\big)\Big)>-\infty,$$

which, by using $r^{-1}|q|^2>0$, implies

$$|n_\infty^\varepsilon|^2 =< n_\infty^\varepsilon, n_{\phi,\theta}>^2,$$

i.e.

$$\cos^2\beta+\varepsilon^2\sin^2\beta = (\cos\beta\cos\alpha\cos(\theta+\alpha)+\cos\beta\sin\alpha\cos\phi\sin(\theta+\alpha))^2$$
$$+2\varepsilon(\cos\beta\cos\alpha\cos(\theta+\alpha)+\cos\beta\sin\alpha\cos\phi\sin(\theta+\alpha))(\sin\beta\sin\phi\sin(\theta+\alpha))$$
$$+\varepsilon^2(\sin\beta\sin\phi\sin(\theta+\alpha))^2.$$

Taking $\varepsilon \to 0^+$, we can deduce

$$\cos^2\beta = (\cos\beta\cos\alpha\cos(\theta+\alpha) + \cos\beta\sin\alpha\cos\phi\sin(\theta+\alpha))^2,$$

which is verified for $\phi, \theta = 2k\pi$ and $k \in \mathbb{N}$. Since $\theta \in [0,\pi)$ and $\phi \in [0, 2\pi)$, we can conclude $\theta, \phi = 0$. This justifies the following ansatz:

$$\theta = \frac{\gamma}{p} \text{ and } \phi = \frac{\delta}{p}, \text{ where } \gamma \text{ and } \delta \text{ are parameters independent from } p. \tag{7.25}$$

For further discussion on the previous ansatz see also [21]. This simplifies the computation to find the stationary points for H_p^ε. Recall the well-known Taylor's expansion of cosine and sine for $\frac{\delta}{p}$ and $\frac{\gamma}{p}+\alpha$, centred respectively at 0 and α, i.e.

$$\begin{aligned}
&\cos\left(\frac{\delta}{p}\right) = 1 + O\left(\frac{1}{p^2}\right), \quad \sin\left(\frac{\delta}{p}\right) = \frac{\delta}{p} + O\left(\frac{1}{p^2}\right),\\
&\cos\left(\frac{\gamma}{p}+\alpha\right) = \cos\alpha - \sin\alpha\frac{\gamma}{p} + O\left(\frac{1}{p^2}\right),\\
&\sin\left(\frac{\gamma}{p}+\alpha\right) = \sin\alpha + \cos\alpha\frac{\gamma}{p} + O\left(\frac{1}{p^2}\right).
\end{aligned}$$

Using the $\varepsilon = \varepsilon(p)$ given in (7.14), we can drop the dependence on the parameter ε and introduce the notation:

$$\begin{aligned}
\overline{f}_p(\theta,\phi) = f_p^{\varepsilon(p)}(\theta,\phi) = -(p-1)r^{-1}|q|^2\Bigg[\Bigg(\cos^2\beta + \frac{\sin^2\beta}{\left(e^p p^2(p-1)\ln p\right)^2}\Bigg)\\
-\left(\frac{\sin\beta\sin\theta\sin(\theta+\alpha)}{e^p p^2(p-1)\ln p} + \cos\beta\sin\alpha\cos\phi\sin(\theta+\alpha) + \cos\beta\cos\alpha\cos(\theta+\alpha)\right)\Bigg]^2\\
+\lambda_1(1-\cos^2(\theta+\alpha)) + \lambda_2(1-(\sin(\theta+\alpha)\cos\phi)^2) + \frac{\lambda_3(1-(\sin\phi\sin(\theta+\alpha))^2}{\left(e^p p^2(p-1)\ln p\right)^2},
\end{aligned}$$

The linearised partial derivative of $\overline{f}_p$ w.r.t. θ can be written as

$$
\begin{aligned}
\frac{\partial \overline{f}_p(\theta,\phi)}{\partial\theta} &= 2(p-1)r^{-1}|q|^2\Bigg(\Bigg(\frac{1}{e^p\ln(p)p(p-1)}\Bigg(\sin\beta\Bigg(\frac{\delta}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&\Bigg(\sin\alpha+\cos\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)+\cos\beta\sin\alpha\Bigg(1+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&\Bigg(\sin\alpha+\cos\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)+\cos\beta\cos\alpha(\cos\alpha-\sin\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&\Bigg(\frac{1}{e^p\ln(p)p(p-1)}\Bigg(\sin\beta\Bigg(\frac{\delta}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\Bigg(\cos\alpha-\sin\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&+\cos\beta\sin\alpha\Bigg(1+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\Bigg(\cos\alpha-\sin\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&-\cos\beta\cos\alpha\Bigg(\sin\alpha+\cos\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\Bigg)\\
&+2\lambda_1\Bigg(-\Bigg(\cos\alpha-\sin\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\Bigg(\sin\alpha+\cos\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&+2\lambda_2\Bigg(-(\sin\alpha+\cos\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\Bigg(1+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&\Bigg(-\Bigg(\cos\alpha-\sin\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\Bigg(1+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&+\frac{2\lambda_3}{(e^p\ln(p)(p(p-1))^2}\Bigg(\Bigg(\frac{\delta}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\Bigg(\sin\alpha+\cos\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&\Bigg(\Bigg(-\frac{\delta}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\Bigg(\Bigg(-\frac{\delta}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\Bigg(\cos\alpha-\sin\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&=2(p-1)r^{-1}|q|^2\Bigg(\cos\beta+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\Bigg(\cos\beta\cos2\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&+2\lambda_1\Bigg(-\cos\alpha\sin\alpha+\cos2\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&+2\lambda_2\Bigg(\cos\alpha\sin\alpha-\cos2\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\Bigg)\\
&=2(p-1)r^{-1}|q|^2\cos^2\beta\cos2\alpha\frac{\gamma}{p}-2\lambda_1\cos\alpha\sin\alpha+2\lambda_1\cos2\alpha\frac{\gamma}{p}\\
&+2\lambda_2\cos\alpha\sin\alpha-2\lambda_2\cos2\alpha\frac{\gamma}{p}+O\Bigg(\frac{1}{p^2}\Bigg)\\
&=\Bigg(2(p-1)r^{-1}|q|^2\cos^2\beta\cos2\alpha+2\lambda_1\cos2\alpha-2\lambda_2\cos2\alpha\Bigg)\frac{\gamma}{p}\\
&+2(\lambda_2-\lambda_1)\cos\alpha\sin\alpha+O\Bigg(\frac{1}{p^2}\Bigg).
\end{aligned}
$$

Similarly one can compute the linearised partial derivative w.r.t. ϕ, that is

$$\begin{aligned}
&\frac{\partial \overline{f}_p(\theta,\phi)}{\partial \phi}\\
&= 2(p-1)r^{-1}|q|^2\bigg(\frac{1}{e^p\ln(p)p^2(p-1)}\bigg(\sin\beta\sin\alpha\frac{\delta}{p}+O\bigg(\frac{1}{p^2}\bigg)\bigg)+\cos\beta+O\bigg(\frac{1}{p^2}\bigg)\bigg)\\
&\bigg[\frac{1}{e^p\ln(p)p^2(p-1)}\bigg(\sin\beta\sin\alpha+\sin\beta\cos\alpha\frac{\gamma}{p}+O\bigg(\frac{1}{p^2}\bigg)\bigg)-\cos\beta\sin^2\alpha\frac{\delta}{p}+O\bigg(\frac{1}{p^2}\bigg)\bigg]\\
&+2\lambda_2\bigg(-\sin^2\alpha\frac{\delta}{p}+O\bigg(\frac{1}{p^2}\bigg)\bigg)+\frac{2\lambda_3}{(e^p\ln(p)p^2(p-1))^2}\bigg(\sin^2\alpha\frac{\delta}{p}+O\bigg(\frac{1}{p^2}\bigg)\bigg)\\
&= 2(p-1)r^{-1}|q|^2\bigg(\cos\beta+O\bigg(\frac{1}{p^2}\bigg)\bigg)\bigg(-\cos\beta\sin^2\alpha\frac{\delta}{p}+O\bigg(\frac{1}{p^2}\bigg)\bigg)\\
&+2\lambda_2\bigg(-\sin^2\alpha\frac{\delta}{p}+O\bigg(\frac{1}{p^2}\bigg)\bigg)+O\bigg(\frac{1}{p^2}\bigg)\\
&= -2(p-1)r^{-1}|q|^2\cos^2\beta\sin^2\alpha\frac{\delta}{p}-2\lambda_2\sin^2\alpha\frac{\delta}{p}+O\bigg(\frac{1}{p^2}\bigg)\\
&= \bigg(-2(p-1)r^{-1}|q|^2\cos^2\beta\sin^2\alpha-2\lambda_2\sin^2\alpha\bigg)\frac{\delta}{p}+O\bigg(\frac{1}{p^2}\bigg).
\end{aligned}$$

Hence the stationary points can be found as solutions of the system

$$\begin{cases}
\bigg(2(p-1)r^{-1}|q|^2\cos^2\beta\cos2\alpha+2\lambda_1\cos2\alpha-2\lambda_2\cos2\alpha\bigg)\frac{\gamma}{p}\\
\quad+2(\lambda_2-\lambda_1)\cos\alpha\sin\alpha+O\bigg(\frac{1}{p^2}\bigg)=0,\\
\bigg(2(p-1)r^{-1}|q|^2\cos^2\beta\sin^2\alpha-2\lambda_2\sin^2\alpha\bigg)\frac{\delta}{p}+O\bigg(\frac{1}{p^2}\bigg)=0,
\end{cases}$$

which gives, recalling the ansatz (7.25)

$$\begin{cases}
\theta_p=\frac{(\lambda_2-\lambda_1)\sin2\alpha}{2r^{-1}|q|^2(p-1)\cos^2\beta\cos2\alpha+2(\lambda_1-\lambda_2)\cos2\alpha}+O\bigg(\frac{1}{p^2}\bigg),\\
\phi_p=O\bigg(\frac{1}{p^2}\bigg).
\end{cases}$$

Set $A=(\lambda_2-\lambda_1)\sin2\alpha$, $B=2r^{-1}|q|^2\cos^2\beta\cos2\alpha$ and $D=2(\lambda_1-\lambda_2)\cos2\alpha-2r^{-1}|q|^2\cos^2\beta\cos2\alpha$, and applying $\frac{A}{Bp+D}=\frac{A}{D}\bigg(\frac{1}{\frac{Bp}{D}+1}\bigg)$, we can set $x=\frac{D}{B\,p}$. Applying the Taylor's expansion for $\frac{x}{x+1}$ near 0, that is $\frac{x}{x+1}=x+O(x^2)$, we conclude

$$\frac{\gamma}{p}=\frac{A}{D}\bigg[\frac{D}{Bp}+O\bigg(\frac{1}{p^2}\bigg)\bigg]=\frac{A}{Bp}+O\bigg(\frac{1}{p^2}\bigg).$$

Finally by taking $C=\frac{A}{B}=\frac{(\lambda_2-\lambda_1)\sin2\alpha}{2r^{-1}|q|^2\cos^2\beta\cos2\alpha}$, we get (7.19). □

Corollary 7.7. *Let us consider the ε-p-Hamiltonian H_p^ε introduced in* (7.16) *and fix $r\in\mathbb{R}$, $q=(q_1,q_2,q_3)^T\in\mathbb{R}^3$ and a matrix $M\in Sym(3)$. Assume that*

$q_1^2 + q_2^2 \neq 0$ *and* $\varepsilon = \varepsilon(p)$ *is given in* (7.14). *Then, for* p *sufficiently large, the optimal control* $\overline{\nu}_p$ *is given by*

$$\overline{\nu}_p = \begin{bmatrix} \sin^2\alpha - \frac{C}{p}\sin 2\alpha + O\left(\frac{1}{p^2}\right) & \frac{1}{2}\sin 2\alpha + \frac{C}{p}\cos 2\alpha + O\left(\frac{1}{p^2}\right) & O\left(\frac{1}{p^2}\right) \\ \frac{1}{2}\sin 2\alpha + \frac{C}{p}\cos 2\alpha + O\left(\frac{1}{p^2}\right) & \cos^2\alpha + \frac{C}{p}\sin 2\alpha + O\left(\frac{1}{p^2}\right) & O\left(\frac{1}{p^2}\right) \\ O\left(\frac{1}{p^2}\right) & O\left(\frac{1}{p^2}\right) & 1 - O\left(\frac{1}{p^2}\right) \end{bmatrix}, \tag{7.26}$$

where α *is the direction associated to the vector* q *by* (7.23).

Proof. Let us consider a generic function $f : \mathbb{R} \to \mathbb{R}$ such that $f(p) = O\left(\frac{1}{p^2}\right)$ as $p \to +\infty$, then by using standard theory of limits, it holds true that, $\lim_{p\to\infty} \frac{\sin(f(p))}{f(p)} = 1$, which means that it is possible to expand the sine as $\sin\left(O\left(\frac{1}{p^2}\right)\right) = O\left(\frac{1}{p^2}\right)$, and, in a similar way, $\cos\left(O\left(\frac{1}{p^2}\right)\right) = 1 - O\left(\frac{1}{p^2}\right)$. By using the standard trigonometric identities for $\sin(\theta + \alpha)$ and $\cos(\theta + \alpha)$ we obtain that

$$\cos\left(\frac{C}{p} + O\left(\frac{1}{p^2}\right)\right) = \cos\left(\frac{C}{p}\right) + O\left(\frac{1}{p^2}\right),$$

and

$$\sin\left(\frac{C}{p} + O\left(\frac{1}{p^2}\right)\right) = \sin\left(\frac{C}{p}\right) + O\left(\frac{1}{p^2}\right).$$

This implies that

$$
\begin{aligned}
&n_{\theta_p,\phi_p} = \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1-O\left(\frac{1}{p^2}\right) & O\left(\frac{1}{p^2}\right) \\ 0 & O\left(\frac{1}{p^2}\right) & 1-O\left(\frac{1}{p^2}\right) \end{bmatrix} \begin{bmatrix} \cos\left(\frac{C}{p}+O\left(\frac{1}{p^2}\right)\right) & -\sin\left(\frac{C}{p}+O\left(\frac{1}{p^2}\right)\right) & 0 \\ \sin\left(\frac{C}{p}+O\left(\frac{1}{p^2}\right)\right) & \cos\left(\frac{C}{p}+O\left(\frac{1}{p^2}\right)\right) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1-O\left(\frac{1}{p^2}\right) & O\left(\frac{1}{p^2}\right) \\ 0 & O\left(\frac{1}{p^2}\right) & 1-O\left(\frac{1}{p^2}\right) \end{bmatrix} \begin{bmatrix} \cos\left(\frac{C}{p}\right)+O\left(\frac{1}{p^2}\right) & -\sin\left(\frac{C}{p}\right)+O\left(\frac{1}{p^2}\right) & 0 \\ \sin\left(\frac{C}{p}\right)+O\left(\frac{1}{p^2}\right) & \cos\left(\frac{C}{p}\right)+O\left(\frac{1}{p^2}\right) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} \cos\left(\frac{C}{p}\right)+O\left(\frac{1}{p^2}\right) & -\sin\left(\frac{C}{p}\right)+O\left(\frac{1}{p^2}\right) & 0 \\ \sin\left(\frac{C}{p}\right)+O\left(\frac{1}{p^2}\right) & \cos\left(\frac{C}{p}\right)+O\left(\frac{1}{p^2}\right) & O\left(\frac{1}{p^2}\right) \\ O\left(\frac{1}{p^2}\right) & O\left(\frac{1}{p^2}\right) & 1-O\left(\frac{1}{p^2}\right) \end{bmatrix} \begin{bmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} \cos\left(\alpha+\frac{C}{p}\right)+O\left(\frac{1}{p^2}\right) \\ \sin\left(\alpha+\frac{C}{p}\right)+O\left(\frac{1}{p^2}\right) \\ O\left(\frac{1}{p^2}\right) \end{bmatrix} = \begin{bmatrix} \cos\alpha-\frac{C}{p}\sin\alpha+O\left(\frac{1}{p^2}\right) \\ \sin\alpha+\frac{C}{p}\cos\alpha+O\left(\frac{1}{p^2}\right) \\ O\left(\frac{1}{p^2}\right) \end{bmatrix}, \qquad (7.27)
\end{aligned}
$$

which concludes the proof by simply computing the corresponding control in the form introduced in (7.21). □

We next investigate the case $q_1^2+q_2^2=0$, which corresponds to the characteristic case for the limit problem.

Case B: $q_1^2+q_2^2=0$ and $q_3\neq 0$.

We now look at the case where $q_1^2+q_2^2=0$ but $|q|\neq 0$, i.e. q is not a degenerate point for the approximated ε-problem (as $p\to\infty$ and for $\varepsilon>0$ fixed) but it is characteristic for the correspondent horizontal problem (as $p\to+\infty$ and $\varepsilon=0$). For more details on the case of characteristic points we refer to [21]. The difficulty here is that now we cannot anymore use as starting unit vector $\frac{q_\infty}{|q_\infty|}$ (in fact $|q_\infty|^2=0$, then the unit vector is not defined). We

instead rotate a generic 3-dimensional unit vector such that the projection on $x_3 = 0$ does not vanish, i.e.

$$n_\alpha = \begin{bmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{bmatrix}, \quad \text{for some generic } \alpha \in [0, 2\pi). \tag{7.28}$$

Similarly to the previous case, we can write all the unit vectors in $\mathbb{R}^3$ as rotation of n_∞, i.e.

$$n_{\phi,\theta} = \begin{bmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\phi\cos(\theta+\alpha) \\ \sin(\theta+\alpha) \\ -\sin\phi\cos(\theta+\alpha) \end{bmatrix}, \tag{7.29}$$

for $\phi \in [0, \pi)$ and $\theta \in [0, 2\pi)$. Note that now α is an additional parameter for the problem even if we do no highlight this dependence in the notation and it does not depend on any fixed variables for h_p^ε, i.e. it is independent of r, q, M. Then all admissible controls in $\mathcal{A}$ can be written as

$$\nu_{\theta,\phi} = I_3 - n_{\theta,\phi} \otimes n_{\theta,\phi} = \begin{bmatrix} 1-(n_{\theta,\phi})_1^2 & -(n_{\theta,\phi})_1(n_{\theta,\phi})_2 & -(n_{\theta,\phi})_1(n_{\theta,\phi})_3 \\ -(n_{\theta,\phi})_1(n_{\theta,\phi})_2 & 1-(n_{\theta,\phi})_2^2 & (n_{\theta,\phi})_2(n_{\theta,\phi})_3 \\ -(n_{\theta,\phi})_1(n_{\theta,\phi})_3 & -(n_{\theta,\phi})_2(n_{\theta,\phi})_3 & 1-(n_{\theta,\phi})_3^2 \end{bmatrix}, \tag{7.30}$$

for $\theta \in [0, \pi)$, $\phi \in [0, \pi)$ (here we have used that $\nu_{\theta,\phi} = \nu_{\theta+\pi,\phi}$ in order to restrict to $\theta \in [0, \pi)$ since a priori $\theta \in [0, 2\,\pi)$).

Remark 7.8. *The choice of using a generic vector n_α is coherent with the fact that, at characteristic points, the optimal control is given by*

$$\overline{\nu} = \begin{bmatrix} \sin^2\alpha & -\sin\alpha\cos\alpha \\ -\sin\alpha\cos\alpha & \cos^2\alpha \end{bmatrix},$$

for all $\alpha \in [0, 2\pi)$, i.e. every admissible control is optimal, see [22] for further details on this point.

We are now ready to prove the following result.

Theorem 7.9. *Let us consider the Hamiltonian H_p^ε introduced in* (7.16) *and fix $r \in \mathbb{R}$, $q = (0, 0, q_3)^T \in \mathbb{R}^3$ with $q_3 \neq 0$ and M in the diagonal form given in* (7.20). *We assume that $\varepsilon = \varepsilon(p)$ satisfies* (7.14). *Then the optimal control $\overline{\nu}$ can be asymptotically expressed, for p large enough, as follows*

1. *If $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$, we have*
$$\overline{\nu} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{7.31}$$
2. *If $\lambda_1 > \lambda_2$ and $\lambda_2 < 0$, we have*
$$\overline{\nu} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{7.32}$$
3. *If $\lambda_1 < \lambda_2$ and $\lambda_1 < 0$, we have*
$$\overline{\nu} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{7.33}$$
4. *Finally, if $\lambda_1 = \lambda_2 = \lambda < 0$,*
$$\overline{\nu} = \begin{bmatrix} \sin^2(\theta+\alpha) & -\sin(\theta+\alpha)\cos(\theta+\alpha) & 0 \\ -\sin\beta\cos(\theta+\alpha) & \cos^2(\theta+\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{7.34}$$
Note that in the case $\lambda_1 = \lambda_2 = \lambda < 0$, the optimal controls are not unique (see later Remark 7.10).

Proof. By assumption (7.20), we deduce $M_\varepsilon = diag(\lambda_1, \lambda_2, \varepsilon^2\lambda_3)$ for some $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, which implies

$$\begin{aligned} h_p^\varepsilon(r, q, M, \nu_{\theta,\phi}) = &-(p-1)r^{-1}\varepsilon^2 q_3^2(1 - \sin^2\phi\cos^2(\theta+\alpha)) \\ &+ \lambda_1(\nu_{\theta,\phi})_{11} + \lambda_2(\nu_{\theta,\phi})_{22} + \varepsilon^2\lambda_3(\nu_{\theta,\phi})_{33}, \end{aligned} \tag{7.35}$$

where h_p^ε is defined as in (7.15). By using $(\nu_{\theta,\phi})_{11} = 1 - \cos^2\phi\cos^2(\theta+\alpha)$, $(\nu_{\theta,\phi})_{22} = 1 - \sin^2(\theta+\alpha)$ and $(\nu_{\theta,\phi})_{33} = 1 - \sin^2\phi\cos^2(\theta+\alpha)$, we can deduce

$$\begin{aligned} h_p^\varepsilon(r,q,M,\nu_{\theta,\phi}) &= \big(\lambda_3 - (p-1)r^{-1}q_3^2\big)\varepsilon^2(1-\sin^2\phi\cos^2(\theta+\alpha)) \\ &\quad + \lambda_1(1-\cos^2\phi\cos^2(\theta+\alpha)) + \lambda_2(1-\sin^2(\theta+\alpha)). \end{aligned} \tag{7.36}$$

Fix α and ε and let us define $f_p^\varepsilon(\theta,\phi) := h_p^\varepsilon(r,q,M,\nu_{\phi,\theta})$. To find the stationary points we need to compute the partial derivatives:

$$\begin{aligned} \frac{\partial f_p^\varepsilon(\theta,\phi)}{\partial\theta} &= 2\big(\lambda_3 - (p-1)r^{-1}q_3^2\big)\varepsilon^2\sin^2\phi\sin(\theta+\alpha)\cos(\theta+\alpha) \\ &\quad + 2\lambda_1\cos(\theta+\alpha)\sin(\theta+\alpha)\cos^2\phi - 2\lambda_2\sin(\theta+\alpha)\cos(\theta+\alpha) \\ &= \left(\big(\lambda_3 - (p-1)r^{-1}q_3^2\big)\varepsilon^2\sin^2\phi + \lambda_1\cos^2\phi - \lambda_2\right)\sin 2(\theta+\alpha), \end{aligned}$$

and

$$\begin{aligned} \frac{\partial f_p^\varepsilon(\theta,\phi)}{\partial\phi} &= -2\big(\lambda_3 - (p-1)r^{-1}q_3^2\big)\varepsilon^2\cos\phi\sin\phi\cos^2(\theta+\alpha) \\ &\quad + 2\lambda_1\cos\phi\sin\phi\cos^2(\theta+\alpha) \\ &= \left(\big((p-1)r^{-1}q_3^2 - \lambda_3\big)\varepsilon^2 + \lambda_1\right)\sin(2\phi)\cos^2(\theta+\alpha). \end{aligned}$$

To find the stationary points is equivalent to solve the system $\frac{\partial f_p^\varepsilon(\theta,\phi)}{\partial\theta} = 0$ and $\frac{\partial f_p^\varepsilon(\theta,\phi)}{\partial\phi} = 0$. We first focus on the first term in $\frac{\partial f_p^\varepsilon(\theta,\phi)}{\partial\phi}$. By using assumption (7.14), one can easily show that, for p large enough,

$$(p-1)r^{-1}\varepsilon^2(p) + \lambda_1 - \varepsilon^2(p)\lambda_3 \neq 0.$$

In fact $\lim_{p\to\infty}((p-1)r^{-1}\varepsilon^2(p) + \lambda_1 - \varepsilon^2(p)\lambda_3) = \lambda_1$, and this implies

(i): if $\lambda_1 > 0$, there exists $\overline{p} > 1$ such that for all $p > \overline{p}$ we have

$$(p-1)r^{-1}\varepsilon^2(p) + \lambda_1 - \varepsilon^2(p)\lambda_3 > 0.$$

(ii): if $\lambda_1 < 0$, there exists $\overline{p} > 1$ such that for all $p > \overline{p}$ we have

$$(p-1)r^{-1}\varepsilon^2(p) + \lambda_1 - \varepsilon^2(p)\lambda_3 < 0.$$

(iii): if $\lambda_1 = 0$, then $\big(\lambda_3 - (p-1)q_3^2 r^{-1}\big)\varepsilon^2 + \lambda_1 = \big(\lambda_3 - (p-1)r^{-1}q_3^2\big)\varepsilon^2$, hence there exists $\overline{p} > 1$ such that $\lambda_3 - (p-1)q_3^2 r^{-1} < 0$. Therefore the term is strictly negative, i.e. in particular it is non vanishing.

Using the above remark, the system for stationary points can be simplified as follows

$$\begin{cases} \Big(\big(\lambda_3 - (p-1)r^{-1}q_3^2\big)\varepsilon^2 \sin^2\phi + \lambda_1 \cos^2\phi - \lambda_2\Big)\sin(\theta+\alpha)\cos(\theta+\alpha) = 0, \\ \sin(2\phi)\cos^2(\theta+\alpha) = 0. \end{cases} \tag{7.37}$$

The solutions of (7.37) are:

(1): If $\cos(\theta+\alpha) = 0$, then $\theta + \alpha = \frac{\pi}{2} + k\pi$ for $k \in \mathbb{N}$, which in our domain means $\theta_0 = \frac{\pi}{2} - \alpha$. Note that in this case the second equation of (7.37) is satisfied for all values $\phi \in [0, \pi)$. Hence the stationary points are:

$$P_1 = \left(\frac{\pi}{2} - \alpha, \phi\right), \quad \text{for all} \ \ \phi \in [0, \pi).$$

(2): If $\sin(\theta+\alpha) = 0$, then $\theta + \alpha = k\pi$ for $k \in \mathbb{N}$, which in the given domain implies $\theta_0 = -\alpha$. Recall also that $\sin(\theta+\alpha) = 0 \Rightarrow \cos^2(\theta+\alpha) = 1$. To solve the second equation in (7.37) we need $\sin(2\phi) = 0$, which means $\phi = 0$ and $\phi = \frac{\pi}{2}$. Hence the other stationary points are

$$P_2 = (-\alpha, 0) \quad \text{and} \quad P_3 = \left(-\alpha, \frac{\pi}{2}\right).$$

(3): If $\varepsilon^2(p)(-(p-1)r^{-1}q_3^2 + \lambda_3) - \lambda_1)\sin^2\phi + \lambda_1 - \lambda_2 = 0$.

(3)-(a): Suppose that $\lambda_1 \neq 0$, this implies

$$\sin^2\phi = \frac{\lambda_1 - \lambda_2}{(p-1)r^{-1}\varepsilon^2 q_3^2 + \lambda_1 - \varepsilon^2\lambda_3}. \tag{7.38}$$

Let us recall that we have chosen $\varepsilon = \varepsilon(p)$ such that $\varepsilon^2(p)(p-1) \to 0$, as $p \to \infty$ (see assumption (7.14)), that implies, for p large enough, that

$$\sin^2\phi \approx \frac{\lambda_1 - \lambda_2}{\lambda_1} = 1 - \frac{\lambda_2}{\lambda_1}.$$

By definition of sine, it has to be $0 \leq 1 - \frac{\lambda_2}{\lambda_1} \leq 1$ that means

$$-1 \leq -\frac{\lambda_2}{\lambda_1} \leq 0 \text{ which implies } \lambda_1 \geq \lambda_2 \geq 0 \text{ with } \lambda_1 \neq 0$$

otherwise the equation (7.38) is impossible at least for large p. Nevertheless also in the case when the identity (7.38) gives some solutions, we get back points already considered in case (1) (for non trivial solutions $\phi \neq 0$) or in case (2) (for the trivial solution $\phi = 0$). Thus this subcase of case (3) can be ignored.

(3)-(b): Suppose that $\lambda_1 = 0$, then we get $\varepsilon^2(-(p-1)r^{-1}q_3^2+\lambda_3))\sin^2\phi - \lambda_2 = 0$. We observe that, whenever $\varepsilon = \varepsilon(p)$ satisfying (7.14), we have

$$\lim_{p\to\infty}\left(\varepsilon^2(p)(-(p-1)r^{-1}q_3^2+\lambda_3))\sin^2\phi - \lambda_2\right) = -\lambda_2.$$

Hence, in this case, stationary points can be found only if $\lambda_1 = \lambda_2 = 0$, which implies $\varepsilon^2(-(p-1)r^{-1}q_3^2+\lambda_3)\sin^2\phi = 0$. For p large enough $-(p-1)r^{-1}q_3^2+\lambda_3 < 0$, so the previous identity is satisfied only for $\sin^2\phi = 0$, i.e. (in our given domain) $\phi = 0$. Therefore in the case $\lambda_1 = \lambda_2 = 0$, we find also the following stationary points:

$$P_4 = (\theta, 0), \quad \text{for all } \theta \in [0, \pi).$$

Hence, computing the values of the function h_p^ε at the stationary points found in (i)-(iii), we have to consider

$$L_1(p) := f_p^\varepsilon\left(\frac{\pi}{2} - \alpha, \phi\right) = \lambda_1 + \varepsilon^2(-(p-1)r^{-1}q_3^2+\lambda_3)\cos^2\phi,$$

$$L_2(p) := f_p^\varepsilon(-\alpha, 0) = \lambda_2 + \varepsilon^2(-(p-1)r^{-1}q_3^2+\lambda_3),$$

$$L_3(p) := f_p^\varepsilon\left(-\alpha, \frac{\pi}{2}\right) = \lambda_1 + \lambda_2,$$

$$L_4(p) := f_p^\varepsilon(\theta, 0) = \varepsilon^2(-(p-1)r^{-1}q_3^2+\lambda_3), \quad \text{for the case } \lambda_1 = \lambda_2 = 0.$$

For sake of simplicity, let us denote $\overline{f}_p(\theta,\phi) = f_p^{\varepsilon(p)}(\theta,\phi)$ and $C(p,\lambda_3,q_3,r) = \varepsilon^2(p)(-(p-1)r^{-1}q_3+\lambda_3)$, remark that

$$\lim_{p\to\infty} C(p,\lambda_3,q_3,r) = 0. \tag{7.39}$$

Note also that

$$L_1(p) - L_2(p) = \lambda_1 - \lambda_2 - C(p,q_3,\lambda_3,r)\sin^2\phi,$$

and

$$\lim_{p\to\infty}(L_1(p) - L_2(p)) = \lambda_1 - \lambda_2. \tag{7.40}$$

Depending on the sign of $\lambda_1 - \lambda_2$ we can deduce the sign of $L_1(p) - L_2(p)$ for p large enough.

Case A: Assume $\lambda_1 - \lambda_2 > 0$, then there exists $\overline{p} > 1$ such that, for all $p > \overline{p}$, we have

$$L_1(p) > L_2(p).$$

Hence it remains to check the sign of $L_3(p) - L_1(p)$. We observe that

$$L_3(p) - L_1(p) = \lambda_2 - C(p,\lambda_3,q_3,r)\cos^2\phi$$

and $\lim_{p\to\infty}(L_3(p) - L_1(p)) = \lambda_2$. As consequence we have to check the following three subcases:

A.1: If $\lambda_2 > 0$, then there exists $\overline{p} > 1$ such that for all $p > \overline{p}$, we have $L_3(p) > L_1(p)$, which, together with $L_1(p) > L_2(p)$, implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_3(p).$$

Hence the maximum is attained at the point $\left(-\alpha, \frac{\pi}{2}\right)$, the associated unit vector is $(0,0,-1)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

A.2: If $\lambda_2 < 0$, there exists $\overline{p} > 1$ such that for all $p > \overline{p}$, we have $L_3(p) < L_1(p)$, which, together with $L_1(p) > L_2(p)$, implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_1(p).$$

Hence the maximum is attained at the point $\left(\frac{\pi}{2} - \alpha, \phi\right)$, the associated unit vector is $(0, 1, 0)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

A.3: If $\lambda_2 = 0$ we cannot deduce anything from the limit, but, recalling that $L_1(p) < \lambda_1$, $L_2(p) < \lambda_2$ and that, since $p > 1$, $q_3 \neq 0$, and $r > 0$ (see [21] for details on $r > 0$), we can deduce $C(p, \lambda_3, q_3, r) < 0$. that implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_3(p).$$

Hence the maximum is attained at the points $\left(-\alpha, \frac{\pi}{2}\right)$, the associated unit vector is $(0, 0, -1)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Case B: Assume $\lambda_1 - \lambda_2 < 0$, there exists $\overline{p} > 1$ such that, for all $p > \overline{p}$, we have $L_1(p) < L_2(p)$. Similarly the previous case, it remains to estimate

$$L_3(p) - L_2(p) = \lambda_1 + C(p, \lambda_3, q_3, r).$$

By using limit (7.39), we know that

$$\lim_{p \to \infty} (L_3(p) - L_2(p)) = \lambda_1.$$

Therefore we need to consider the following three subcases.

B.1: If $\lambda_1 < 0$, there exists $\overline{p} > 1$ such that, for all $p > \overline{p}$, we have $L_3(p) < L_2(p)$, which implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_2(p).$$

Hence the maximum is attained at the points $\big(-\alpha, 0\big)$, the associated unit vectors is $(1,0,0)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

B.2: If $\lambda_1 > 0$, there exists $\overline{p} > 1$ such that, for all $p > \overline{p}$, we have $L_3(p) > L_2(p)$ which implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_3(p).$$

Hence the maximum is attained at the point $\big(-\alpha, \frac{\pi}{2}\big)$, the associated unit vector is $(0,0,-1)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

B.3: If $\lambda_1 = 0$ we cannot deduce anything from the limit but, recalling that $L_1(p) < \lambda_1$ and $L_2(p) < \lambda_2$, we deduce $C(p, \lambda_3, q_3, r) < 0$, that implies $L_1(p) < L_2(p) < L_3$, i.e.

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_3(p).$$

Hence the maximum is attained at the point $\big(-\alpha, \frac{\pi}{2}\big)$, the associated unit vector is $(0,0,-1)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Case C: Assume $\lambda_1 = \lambda_2 =: \lambda$. In this case we can rewrite the values $L_i(p)$ for $i = 1, 2, 3, 4$, as follows

$$L_1(p) = f_p^\varepsilon\left(\frac{\pi}{2} - \alpha, \phi\right) = \lambda + C(\lambda_3, q_3, r, p),$$
$$L_2(p) = f_p^\varepsilon(\theta, 0) = \lambda + C(\lambda_3, q_3, r, p),$$
$$L_3(p) = f_p^\varepsilon\left(-\alpha, \frac{\pi}{2}\right) = 2\lambda,$$
$$L_4(p) = f_p^\varepsilon(\theta, 0) = C(\lambda_3, q_3, r, p), \quad \text{for the case } \lambda = 0.$$

We remark, recalling that $C(\lambda_3, q_3, r, p) < 0$ for all $p > \overline{p}$, and that

$$L_1(p) = L_2(p) < \lambda \quad \text{for all } p > \overline{p}.$$

Hence we can divide this case in three subcases.

C.1: If $\lambda > 0$, for p large enough, we have

$$L_2(p) = L_1(p) < L_3(p),$$

which implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_3(p).$$

Hence the maximum is attained at the point $\left(-\alpha, \frac{\pi}{2}\right)$, the associated unit vector is $(0, 0, -1)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

C.2: If $\lambda = 0$, for p large enough, we have

$$L_4(p) = L_1(p) < L_3(p),$$

which implies

$$\max\{L_1(p), L_3(p), L_4(p)\} = L_3(p).$$

Hence the maximum is attained at the point $\left(-\alpha, \frac{\pi}{2}\right)$, the associated unit vector is $(0,0,-1)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

C.3: If $\lambda < 0$, for a p large enough, we have

$$L_3(p) < L_1(p) = L_2(p),$$

which implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_1(p) = L_2(p).$$

Then the maximum is attained at the point $(\theta, 0)$ which is associated to the unit vector $(\cos(\theta+\alpha), \sin(\theta+\alpha), 0)^T$, so the optimal control is given by

$$\begin{aligned} \overline{\nu} &= I_3 - \begin{bmatrix} \cos(\theta+\alpha) \\ \sin(\theta+\alpha) \\ 0 \end{bmatrix} \otimes \begin{bmatrix} \cos(\theta+\alpha) \\ \sin(\theta+\alpha) \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \sin^2(\theta+\alpha) & -\cos(\theta+\alpha)\sin(\theta+\alpha) & 0 \\ -\cos(\theta+\alpha)\sin(\theta+\alpha) & \cos^2(\theta+\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

The maximum could also be attained at the point $\left(\frac{\pi}{2} - \alpha, \phi\right)$, which is just a particular case of the case above for $\theta = \frac{\pi}{2} - \alpha$.

□

Remark 7.10. *Note that, even if the optimal controls for the case $\lambda_1 = \lambda_2 = \lambda$ are not unique, the value of the function* (7.35) *associated for all controls* (7.31)

does not depend on the angle $\theta+\alpha$; in fact:

$$
\begin{aligned}
&h_p^{\varepsilon}(r,q,M,\overline{\nu})\\
&=\varepsilon^2\Big(\lambda_3-(p-1)r^{-1}q_3^2\big(1-\sin^2\phi\cos^2(\theta+\alpha)\big)\Big)+\lambda\sin^2(\theta+\alpha)+\lambda\cos^2(\theta+\alpha)\\
&=\varepsilon^2\Big(\lambda_3-(p-1)r^{-1}q_3^2\big(1-\sin^2\phi\cos^2(\theta+\alpha)\big)\Big)+\lambda.
\end{aligned}
$$

Case C: $q_1^2+q_2^2+q_3^2=0$

It remains to consider the case of $|q|=0$, i.e. $q_1=q_2=q_3=0$. Under this assumption (8.23) can be rewritten as

$$H_p^{\varepsilon}(r,0,M,\nu)=H_p^{\varepsilon}(M,\nu)=\sup_{\nu\in\mathcal{A}_1}Tr[\nu_1\nu_1^TM_{\varepsilon}]. \tag{7.41}$$

Similarly to case B, we know from [22] that also in this case any admissible control is optimal in the limit case, so we will use a similar strategy for the proof.

Theorem 7.11. *Let us consider the Hamiltonian H_p^{ε} introduced in* (7.41) *and fix and $M\in Sym(3)$ in the diagonal form given in* (7.20)*. We assume that $\varepsilon=\varepsilon(p)$ is given in* (7.14)*. Then the optimal control for the Hamiltonian H_p^{ε} can be asymptotically expressed, for large p, as*

1. *If one of the following holds true:*

   a) $\lambda_1>0$, $\lambda_2>0$,

   b) $\lambda_1>0$, $\lambda_2=0$ *and* $\lambda_3\le 0$,

   c) $\lambda_2>0$, $\lambda_1=0$ *and* $\lambda_3\le 0$,

   d) $\lambda_1=\lambda_2:=\lambda\le 0$ *and* $\lambda_3\ge 0$,

   then the p-optimal control is

$$\overline{\nu}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}.$$

2. *If one of the following holds true:*

   a) $\lambda_1 > \lambda_2$ *and* $\lambda_2 < 0$,

   b) $\lambda_1 > 0$, $\lambda_2 = 0$ *and* $\lambda_3 \geq 0$,

   then the p-optimal control is

$$\overline{\nu} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

3. *If one of the following holds true:*

   a) $\lambda_1 < \lambda_2$ *and* $\lambda_1 < 0$,

   b) $\lambda_2 > 0$, $\lambda_1 = 0$ *and* $\lambda_3 \geq 0$,

   then the p-optimal control is

$$\overline{\nu} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

4. *If* $\lambda_1 = \lambda_2 =: \lambda \leq 0$ *and* $\lambda_3 \leq 0$, *then the p-optimal control is given by*

$$\overline{\nu} = \begin{bmatrix} \sin^2(\theta+\alpha) & -\sin(\theta+\alpha)\cos(\theta+\alpha) & 0 \\ -\sin\beta\cos(\theta+\alpha) & \cos^2(\theta+\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{7.42}$$

   Note that in this last case the optimal controls are not unique, see Remark 7.12 later, for more details on this case.

Proof. We consider, as in the Case B, the vector $n_{\phi,\theta}$ as defined in (7.29) (where $\theta \in [0,\pi)$ and $\phi \in [0,\pi)$) and its associated general admissible control $\nu_{\theta,\phi}$. By assumption (7.20), we deduce $M_\varepsilon = diag(\lambda_1, \lambda_2, \varepsilon^2\lambda_3)$ for some λ_1, λ_2, $\lambda_3 \in \mathbb{R}$, that implies recalling (7.41)

$$h_p^\varepsilon(r,q,M,\nu_{\theta,\phi}) := \lambda_1(\nu_{\theta,\phi})_{11} + \lambda_2(\nu_{\theta,\phi})_{22} + \varepsilon^2\lambda_3(\nu_{\theta,\phi})_{33}. \tag{7.43}$$

Recalling $(\nu_{\theta,\phi})_{11}$, $(\nu_{\theta,\phi})_{22}$ and $(\nu_{\theta,\phi})_{33}$ as in Theorem 7.9 we can rewrite (7.43) explicitly as

$$\begin{aligned} h_p^\varepsilon(r,q,M,\nu_{\theta,\phi}) &:= \lambda_1(1-\cos^2\phi\cos^2(\theta+\alpha)) \\ &+ \lambda_2(1-\sin^2(\theta+\alpha)) + \varepsilon^2\lambda_3(1-\sin^2\phi\cos^2(\theta+\alpha)). \end{aligned} \tag{7.44}$$

Let us fix α and ε and recall the notation $f_p^\varepsilon(\theta,\phi) = h_p^\varepsilon(r,q,M,\nu_{\phi,\theta})$, to find the stationary points we compute the partial derivatives: proceeding as Case B we obtain

$$\begin{aligned} \frac{\partial f_p^\varepsilon}{\partial\theta}(\theta,\phi) &= 2\lambda_1\cos^2\phi\sin(\theta+\alpha)\cos(\theta+\alpha) \\ &\quad - 2\lambda_2\sin(\theta+\alpha)\cos(\theta+\alpha) + 2\varepsilon^2\lambda_3\sin^2\phi\cos(\theta+\alpha)\sin(\theta+\alpha) \\ &= \left(\lambda_1\cos^2\phi - \lambda_2 + \varepsilon^2\lambda_3\sin^2\phi\right)\sin 2(\theta+\alpha) \end{aligned} \tag{7.45}$$

and

$$\begin{aligned} \frac{\partial f_p^\varepsilon}{\partial\phi}(\theta,\phi) &= 2\lambda_1\cos\phi\sin\phi\cos^2(\theta+\alpha) - 2\varepsilon^2\lambda_3\cos\phi\sin\phi\cos^2(\theta+\alpha) \\ &= \left(\lambda_1 - \varepsilon^2\lambda_3\right)\sin(2\phi)\cos^2(\theta+\alpha). \end{aligned} \tag{7.46}$$

Let us remark that, choosing $\varepsilon = \varepsilon(p)$ such that $\varepsilon^2(p)(p-1) \to 0$ as $p \to \infty$, we have

$$\lim_{p\to\infty}(\lambda_1 - \varepsilon^2(p)\lambda_3) = \lambda_1.$$

This leads to

1. If $\lambda_1 > 0$ then, there will exist a $\overline{p} > 1$ such that for all $p > \overline{p}$ we have
$$\lambda_1 - \varepsilon^2(p)\lambda_3 > 0.$$

2. If $\lambda_1 < 0$ then, there will exist a $\overline{p} > 1$ such that for all $p > \overline{p}$ we have
$$\lambda_1 - \varepsilon^2(p)\lambda_3 < 0.$$

3. If $\lambda_1 = 0$ there will exist at last one $\overline{p} > 1$ such that $\varepsilon^2(\overline{p})\lambda_3 > 0$ if $\lambda_3 > 0$ (or $\varepsilon^2(\overline{p})\lambda_3 < 0$ if $\lambda_3 < 0$), then for all $p > \overline{p}$ we have
$$\varepsilon^2(\overline{p})\lambda_3 \neq 0.$$

Hence, in order to find the stationary points, we have to find the zeros of the following system (recalling (7.45) and (7.46))

$$\begin{cases} 2\Big(\lambda_1 \cos^2\phi - \lambda_2 + \varepsilon^2\lambda_3 \sin^2\phi\Big)\sin(\theta+\alpha)\cos(\theta+\alpha) = 0, \\ \sin(2\phi)\cos^2(\theta+\alpha) = 0. \end{cases}$$

To do it we will do the following remarks

(i) Points $P_1 = \big(\frac{\pi}{2} - \alpha, \phi\big)$ for all $\phi \in [0,\pi)$, $P_2 = (-\alpha, 0)$, $P_3 = \big(-\alpha, \frac{\pi}{2}\big)$ are obtained as in Theorem 7.9.

(ii) If $(\varepsilon^2(p)\lambda_3 - \lambda_1)\sin^2\phi + \lambda_1 - \lambda_2 = 0$.

(ii)-(a) We suppose that $\lambda_1 \neq 0$. This allows us to write

$$\sin^2\phi = \frac{\lambda_1 - \lambda_2}{\lambda_1 - \varepsilon^2(p)\lambda_3}. \tag{7.47}$$

Proceeding as Theorem 7.9, as $p \to \infty$ (see assumption (7.14)), we obtain, for p large enough, that

$$\sin^2\phi \approx \frac{\lambda_1 - \lambda_2}{\lambda_1} = 1 - \frac{\lambda_2}{\lambda_1}.$$

And, recalling the same computation in Theorem 7.9, we have

$$\lambda_1 \geq \lambda_2 \geq 0 \text{ with } \lambda_1 \neq 0$$

otherwise the equation (7.47) is impossible. Nevertheless also in the case when the identity (7.47) gives solution (that one can find explicitly inverting in the corresponding domain $\sin^2$) we get back points already considered in case (i) (for non trivial solutions ϕ) or in case (ii) (for the trivial solution). Thus this subcase of case (iii) can be ignored.

(iii)-(b) We suppose now that $\lambda_1 = 0$, then we get $\varepsilon^2(p)\lambda_3 \sin^2\phi - \lambda_2 = 0$. We observe

$$\lim_{p\to\infty}(\varepsilon^2(p)\lambda_3 \sin^2\phi - \lambda_2) = -\lambda_2.$$

Hence we deduce that in order to find stationary points, we obtain $\lambda_1 = \lambda_2 = 0$, then

$$(\varepsilon^2(p)\lambda_3)\sin^2\phi = 0 \Leftrightarrow \sin^2\phi = 0 \Leftrightarrow \phi = 0. \tag{7.48}$$

Therefore in the case $\lambda_1 = \lambda_2 = 0$, we find also the following stationary points:

$$P_4 = (\theta, 0), \quad \text{for all } \theta \in [0, \pi).$$

Hence we obtain

$$\begin{aligned}
L_1(p) :=& f_p^\varepsilon\left(\frac{\pi}{2} - \alpha, \phi\right) = \lambda_1 + \varepsilon^2\lambda_3 \sin^2\phi,\\
L_2(p) :=& f_p^\varepsilon(-\alpha, 0) = \lambda_2 + \varepsilon^2\lambda_3,\\
L_3(p) :=& f_p^\varepsilon\left(-\alpha, \frac{\pi}{2}\right) = \lambda_1 + \lambda_2,\\
L_4(p) :=& f_p^\varepsilon(\theta, 0) = \varepsilon^2\lambda_3, \quad \text{in the case } \lambda_1 = \lambda_2 = 0.
\end{aligned}$$

Let us define $\overline{f}_p(\theta, \phi) := f_p^{\varepsilon(p)}(\theta, \phi)$ and $C(p, \lambda_3) := \varepsilon^2(p)\lambda_3$. We remark also that the sign of $C(p, \lambda_3)$ depends on λ_3 and that

$$\lim_{p\to\infty} C(p, \lambda_3) = 0. \tag{7.49}$$

Let us remark that

$$L_1(p) - L_2(p) = \lambda_1 - \lambda_2 - C(p, \lambda_3)\sin^2\phi$$

and

$$\lim_{p\to\infty}(L_1(p) - L_2(p)) = \lambda_1 - \lambda_2. \tag{7.50}$$

Hence we obtain three subcases which have to be studied

Case A: Assume $\lambda_1 - \lambda_2 > 0$: there will exist a $\overline{p} > 1$ such that $p > \overline{p}$

$$L_1(p) > L_2(p).$$

Hence we have to check now the sign of

$$L_3(p) - L_1(p) = \lambda_2 - C(p, \lambda_3) \sin^2 \phi,$$

and

$$\lim_{p \to \infty} (L_3(p) - L_1(p)) = \lambda_2,$$

hence we have three more subcases.

A.1 If $\lambda_2 > 0$ then as in Theorem 7.9 (**A.1**) we have that the maximum is attained at the point $\left(-\alpha, \frac{\pi}{2}\right)$, then the associated unit vector is $(0, 0, -1)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

A.2 If $\lambda_2 < 0$ then as in Theorem 7.9 (**A.2**) we have that the maximum is attained at the point $\left(\frac{\pi}{2} - \alpha, \phi\right)$, then the associated unit vector is and $(0, -1, 0)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

A.3 If $\lambda_2 = 0$ then we have $\lambda_1 > 0$ and we cannot deduce anything from the limit but we have to check the sign of $C(\lambda_3, p)$.

i. If $\lambda_3 > 0$ then, for a p large enough, it holds true

$$L_2(p) < L_3(p) < L_1(p)$$

which implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_1(p).$$

Hence the maximum is attained at the point $\left(\frac{\pi}{2}-\alpha,\phi\right)$, then the associated unit vector is $(0,-1,0)^T$ and the optimal control is given by

$$\overline{\nu}=I_3-\begin{bmatrix}0\\-1\\0\end{bmatrix}\otimes\begin{bmatrix}0\\-1\\0\end{bmatrix}=\begin{bmatrix}1&0&0\\0&0&0\\0&0&1\end{bmatrix}.$$

ii. If $\lambda_3<0$ then, for a p large enough, it holds true

$$L_2(p)<L_1(p)<L_3(p)$$

which implies

$$\max\{L_1(p),L_2,L_3(p)\}=L_3(p).$$

Hence the maximum is attained at the point $\left(-\alpha,\frac{\pi}{2}\right)$, then the associated unit vector is $(0,0,-1)^T$ and the optimal control is given by

$$\overline{\nu}=I_3-\begin{bmatrix}0\\0\\-1\end{bmatrix}\otimes\begin{bmatrix}0\\0\\-1\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}.$$

iii. If $\lambda_3=0$ then, for a p large enough,, it holds true

$$L_2(p)<L_3(p)=L_1(p)$$

which implies

$$\max\{L_1(p),L_2(p),L_3(p)\}=L_1(p)=L_3(p).$$

Hence the maximum is attained at the point $\left(\frac{\pi}{2}-\alpha,\phi\right)$, then the associated unit vector is $(0,-1,0)^T$ and the optimal control is given by

$$\overline{\nu}=I_3-\begin{bmatrix}0\\-1\\0\end{bmatrix}\otimes\begin{bmatrix}0\\-1\\0\end{bmatrix}=\begin{bmatrix}1&0&0\\0&0&0\\0&0&1\end{bmatrix}$$

or at the point $\big(-\alpha, \frac{\pi}{2}\big)$, then the associated unit vectors are $(0,0,-1)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Case B: Assume $\lambda_1 - \lambda_2 < 0$, there exists a $\overline{p} > 1$ such that $p > \overline{p}$

$$L_1(p) < L_2(p)$$

Hence we have to check now the sign of

$$L_3(p) - L_2(p) = \lambda_1 - C(p, \lambda_3).$$

We remark that

$$\lim_{p\to\infty} (L_3(p) - L_2(p)) = \lambda_1$$

so we have to check the different subcases related to the sign of λ_1.

B.1 If $\lambda_1 < 0$, as in Theorem 7.9 (**B.1**) we have that the maximum is attained at the point $\big(-\alpha, 0\big)$, then the associated unit vector is $(1,0,0)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

B.2 If $\lambda_1 > 0$ then as in Theorem 7.9 (**B.2**) we have that the maximum is attained at the point $\big(-\alpha, \frac{\pi}{2}\big)$, then the associated unit vector is $(0,0,-1)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

B.3 If $\lambda_1 = 0$ then $\lambda_2 > 0$ and we cannot apply the theorem of the previous points, hence we have to study the sign of $C(p, \lambda_3)$ which depends on λ_3

i. If $\lambda_3 > 0$, then, for a p large enough,

$$L_3(p) < L_1(p) < L_2(p)$$

which implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_2(p).$$

Hence the maximum is attained at the point $\big(-\alpha, 0\big)$, then the associated unit vector is $(1,0,0)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

ii. If $\lambda_3 < 0$ then, for a p large enough,

$$L_2(p) < L_1(p) < L_3(p)$$

which implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_3(p).$$

Hence the maximum is attained at the point $\big(-\alpha, \frac{\pi}{2}\big)$, then the associated unit vector is $(0,0,-1)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

iii. If $\lambda_3 = 0$ then, for a p large enough,

$$L_1(p) < L_2(p) = L_3(p)$$

which implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_2(p) = L_3(p).$$

Hence the maximum is attained in the point $\big(-\alpha, 0\big)$, then the associated unit vector is $(1,0,0)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix}1\\0\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}0 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$$

or at the point $\big(-\alpha, \frac{\pi}{2}\big)$, then the associated unit vectors is $(0,0,-1)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix}0\\0\\-1\end{bmatrix} \otimes \begin{bmatrix}0\\0\\-1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 0\end{bmatrix}.$$

Case C: Assume $\lambda_1 = \lambda_2$. Recalling the stationary points found in **Case A** and **Case B** and the associated values function, we have

$$L_1(p) := \overline{f}_p\left(\frac{\pi}{2} - \alpha, \phi\right) = \lambda + C(p,\lambda_3)\sin^2\phi$$

$$L_2(p) := \overline{f}_p(\theta, 0) = \lambda(2 - \cos^2(\theta+\alpha) - \sin^2(\theta+\alpha)) + C(p,\lambda_3) = \lambda + C(p,\lambda_3),$$

$$L_3(p) := \overline{f}_p\left(-\alpha, \frac{\pi}{2}\right) = 2\lambda,$$

$$L_4(p) := f_p^{\varepsilon}(\theta, 0) = C(p,\lambda_3), \quad \text{in the case } \lambda_1 = \lambda_2 = 0.$$

Let us remark that $C(p,\lambda_3) < 0$ if $\lambda_3 < 0$ and $C(p,\lambda_3) > 0$ if $\lambda_3 > 0$ and that the case $(-\alpha, 0)$ is considered in the case $(\theta, 0)$.

C.1 If $\lambda > 0$ then, as in Theorem 7.9 (**C.1**), we have that the maximum is attained at the point the maximum is attained at the point $\big(-\alpha, \frac{\pi}{2}\big)$ then the associated unit vector is $(0,0,-1)^T$ and the

optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

C.2 If $\lambda < 0$ then we obtain that for a large p

i. If $\lambda_3 > 0$

$$L_1(p) > L_2(p) > L_3(p)$$

which implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_1(p).$$

Then the maximum is atteined at $\left(\frac{\pi}{2} - \alpha, \phi\right)$, with associated unit vector $(0, -1, 0)^T$ and then the associated optimal control is

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

ii. If $\lambda_3 < 0$

$$L_2(p) > L_1(p) > L_3(p)$$

which implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = A_1(p)$$

Then the maximum is attained at $(\theta, 0)$ which is associated to the unit vector $(\cos(\theta + \alpha), \sin(\theta + \alpha), 0)^T$, hence the optimal

control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} \cos(\theta+\alpha) \\ \sin(\theta+\alpha) \\ 0 \end{bmatrix} \otimes \begin{bmatrix} \cos(\theta+\alpha) \\ \sin(\theta+\alpha) \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} \sin^2(\theta+\alpha) & -\sin(\theta+\alpha)\cos(\theta+\alpha) & 0 \\ -\sin(\theta+\alpha)\cos(\theta+\alpha) & \cos^2(\theta+\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

iii. If $\lambda_3 = 0$

$$L_1(p) = L_2(p) > L_3(p)$$

which implies

$$\max\{L_1(p), L_2(p), L_3(p)\} = L_1(p) = L_2(p)$$

Then the maximum is attained in $(\theta, 0)$ which is associated to the unit vector $(\cos(\theta+\alpha), \sin(\theta+\alpha), 0)^T$, hence the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} \cos(\theta+\alpha) \\ \sin(\theta+\alpha) \\ 0 \end{bmatrix} \otimes \begin{bmatrix} \cos(\theta+\alpha) \\ \sin(\theta+\alpha) \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} \sin^2(\theta+\alpha) & -\sin(\theta+\alpha)\cos(\theta+\alpha) & 0 \\ -\sin(\theta+\alpha)\cos(\theta+\alpha) & \cos^2(\theta+\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The maximum can be attained also at $\left(\frac{\pi}{2} - \alpha, \phi\right)$ and the associated optimal control may be seen as a particular case of the found optimal control in this case.

C.3 If $\lambda = 0$ then we have for a large p

i. If $\lambda_3 > 0$

$$L_4(p) < L_1(p) < L_3(p)$$

which implies

$$\max\{L_1(p), L_3(p), L_4(p)\} = L_3(p).$$

The maximum is attained at the point $\big(-\alpha, \frac{\pi}{2}\big)$, then the associated unit vector is $(0, 0, -1)^T$ and the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

ii. If $\lambda_3 < 0$

$$L_4(p) > L_1(p) > L_3(p)$$

which implies

$$\max\{L_1(p), L_3(p), L_4(p)\} = L_4(p).$$

Then the maximum is attained at $(\theta, 0)$ which is associated to the unit vector $(\cos(\theta+\alpha), \sin(\theta+\alpha), 0)^T$, hence the optimal control is given by

$$\overline{\nu} = I_3 - \begin{bmatrix} \cos(\theta+\alpha) \\ \sin(\theta+\alpha) \\ 0 \end{bmatrix} \otimes \begin{bmatrix} \cos(\theta+\alpha) \\ \sin(\theta+\alpha) \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} \sin^2(\theta+\alpha) & -\sin(\theta+\alpha)\cos(\theta+\alpha) & 0 \\ -\sin(\theta+\alpha)\cos(\theta+\alpha) & \cos^2(\theta+\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

iii. If $\lambda_3 = 0$

$$L_1(p) = L_3(p) = L_4(p)$$

which implies

$$\max\{L_1(p), L_3, L_4(p)\} = L_1(p) = L_3(p) = L_4(p)$$

Hence the maximum is attained at every stationary points that we have found before and so the non unique optimal control is given by

$$\begin{aligned}\overline{\nu} &= I_3 - \begin{bmatrix}\cos(\theta+\alpha)\\ \sin(\theta+\alpha)\\ 0\end{bmatrix} \otimes \begin{bmatrix}\cos(\theta+\alpha)\\ \sin(\theta+\alpha)\\ 0\end{bmatrix} \\ &= \begin{bmatrix}\sin^2(\theta+\alpha) & -\sin(\theta+\alpha)\cos(\theta+\alpha) & 0\\ -\sin(\theta+\alpha)\cos(\theta+\alpha) & \cos^2(\theta+\alpha) & 0\\ 0 & 0 & 1\end{bmatrix}.\end{aligned}$$

The maximum can be attained also at $\left(\frac{\pi}{2}-\alpha,\phi\right)$ and $\left(-\alpha,\frac{\pi}{2}\right)$ and the associated optimal controls may be seen as a particular case of the found optimal control in this case.

Remark 7.12. *Proceeding as in Remark 7.10, even if the control with $\lambda_1 = \lambda_2 = \lambda$ is not unique, the value of the function* (7.43) *associated for all controls* (7.42) *does not depend on the angle $\theta+\alpha$; in fact:*

$$h_p^{\varepsilon}(r,q,M,\overline{\nu}) = \varepsilon^2\lambda_3 + \lambda\sin^2(\theta+\alpha) + \lambda\cos^2(\theta+\alpha) = \varepsilon^2\lambda_3 + \lambda.$$

□

Chapter 8

Convergence of Riemannian approximation for stochastic representation

8.1 Introduction

The evolution by mean curvature flow (MCF) has been studied extensively and it has many applications in image processing (see e.g. [13]) Roughly speaking we say that a hypersurface evolves by MCF if it contracts in the normal direction with normal velocity proportional to its mean curvature see e.g. [28] for more details. It is well-known that this evolution may develop singularities in finite time. To deal with such a singularities, many generalised approaches to study this evolution have been developed. In particular in 1991, Chen,Giga and Goto [11] and, independently Evans and Spruck [31] introduced the so called level set approach, which consists in studying the evolving hypersurfaces as level sets of (viscosity) solutions of suitable associated nonlinear PDEs. In this paper we are interested in a degenerate version of such evolution, namely evolution by horizontal mean curvature flow (HMCF):

we consider a hypersurface embedded in a sub-Riemannian structure, then the evolution contracts in the direction of the so called horizontal normal proportionally to its horizontal curvature (see Section 2 for details). We consider the level set approach which is now associated to a parabolic PDE, far more degenerate than in the standard case. Here we are going to study a Riemannian approximations similarly to what is done in [8], still the approach is very different since we interpret the solution of the level set equations as value function of suitable associated stochastic control problem, following the approach developed by Caradaileguet, Quincampoix and Buckdahn in [27] and contemporaneously but independently by Soner and Touzi [49] for the standard (Euclidean) case and generalised then by Dirr, Dragoni and von Renesse in [22] for the case of HMCF. We will show that under suitable conditions that there exists a subsequence of $\varepsilon(p)$ with p positive parameter such that it holds

$$\limsup_{p\to\infty} V_p^{\varepsilon(p)}(t,x) \leq V(t,x) \quad \text{where } \varepsilon(p) \to 0 \text{ as } p \to \infty.$$

where $V_p^{\varepsilon} : [0,T] \times \mathbb{R}^N \to \mathbb{R}$ is the solution of approximated mean curvature flow and $V : [0,T] \times \mathbb{R}^N \to \mathbb{R}$ the solution of horizontal mean curvature flow. Then we will give an idea to approach in the Section 8.5 the inequality

$$\liminf_{p\to\infty} V_p^{\varepsilon(p)}(t,x) \geq V(t,x) \quad \text{where } \varepsilon(p) \to 0 \text{ as } p \to \infty.$$

From these estimates we will obtain that, even though the comparison principle is still an open question, the solution found by Citti and Capogna [8] are exactly the same found by Buckdahn, Cardalieguet and Quincampoix [27] and by Soner and Touzi [48] which was generalized by Dirr, Dragoni, Von Renesse [22].

8.2 Preliminaries

We briefly recall some basic geometrical definitions which will be key for defining the evolution by HMCF. For more definitions and properties on sub-Riemannian geometries we refer to [42], and also to [4] for the particular case of Carnot groups (see Chapter 1 for further details).

Definition 8.1. *Let M be a N-dimensional smooth manifold, we can define for every point p a subspace of T_pM called $\mathcal{H}_p$. We define the* distribution *associated to M as the subbundle of the tangent bundle $\mathcal{H} = \{(p, v) |\ p \in M,\ v \in \mathcal{H}_x\}$.*

Definition 8.2. *Let M be a manifold, X,Y two vector fields defined on M and $f : M \to \mathbb{R}$ a smooth function, the* braket *between X and Y is the vector fields acting on smooth functions as $[X, Y](f) = XY(f) - YX(f)$.*
Let us consider $\mathcal{X} = \{X_1, \dots, X_m\}$ spanning some distribution $\mathcal{H} \subset TM$, we define the k-bracket as $\mathcal{L}^{(k)} = \{[X, Y] \mid X \in \mathcal{L}^{(k-1)}, \quad Y \in \mathcal{L}^{(1)}\}$ with $i_j \in \{1, \dots, m\}$ and $\mathcal{L}^{(1)} = \mathcal{X}$. The associated Lie algebra is the set of all brackets between elements of $\mathcal{X}$, i.e.

$$\mathcal{L}(\mathcal{X}) := \left\{ [X_i, X_j^{(k)}] \mid X_j^{(k)} \text{ k-length bracket of } X_1, \dots, X_m,\ k \in \mathbb{N} \right\}.$$

Definition 8.3 (Hörmander condition)**.** *Let M be a smooth manifold and $\mathcal{H}$ a distribution defined on M and $\mathcal{X}$ a family of vector fields on M generating the distribution $\mathcal{H}$. We say that the distribution is* bracket generating *, at any point, the Lie algebra $\mathcal{L}(\mathcal{X})$ spans the whole tangent space. We say that a sub-Riemannian geometry satisfies the* Hörmander condition *if and only if the associated distribution is braket generating.*

Definition 8.4. *Let M be a smooth manifold and $\mathcal{H} = span\{X_1, \dots, X_m\} \subset TM$ a distribution and g a Riemannian metric of M defined on the subbundle $\mathcal{H}$. A* sub-Riemannian geometry *is the triple $(M, \mathcal{H}, g)$.*

Definition 8.5. *Let $(M, \mathcal{H}, g)$ be a sub-Riemannian geometry and $\gamma : [0, T] \to M$ an absolutely continuous curve, we say that γ is an* horizontal curve *if and only if*

$$\dot{\gamma}(t) \in \mathcal{H}_{\gamma(t)}, \quad \text{for a.e. } t \in [0, T],$$

or, equivalently, if there exists a measurable function $h : [0, T] \to \mathbb{R}^N$ such that

$$\dot{\gamma}(t) = \sum_{i=1}^{m} h_i(t) X_i(\gamma(t)), \quad \text{for a.e. } t \in [0, T],$$

where $h(t) = (h_1(t), \dots, h_m(t))$ *and* $X_1, \dots, X_m$ *are vector fields spanning the distribution* $\mathcal{H}$*.*

Example 8.6 (The Heisenberg group)**.** *For a formal definition of the Heisenberg group and the connection between its structure as non commutative Lie group and its manifold structure we refer to [4]. Here we simply introduce the 1-dimensional Heisenberg group as the sub-Riemannian structure induced on* $\mathbb{R}^3$ *by the vector fields*

$$X_1(x) = \begin{pmatrix} 1 \\ 0 \\ -\frac{x_2}{2} \end{pmatrix} \quad \text{and} \quad X_2 = \begin{pmatrix} 0 \\ 1 \\ \frac{x_1}{2} \end{pmatrix}, \quad \forall\, x = (x_1, x_2, x_3) \in \mathbb{R}^3.$$

The introduced vector fields satisfy the Hörmander condition with step 2: in fact $[X_1, X_2](x) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ *for any* $x \in \mathbb{R}^3$*.*

From now on we consider only the case where the starting topological manifold M is the Euclidean $\mathbb{R}^N$. Moreover, in this paper we will concentrate on a sub-Riemannian geometries with a particular structure: the so called Carnot-type geometries.

Definition 8.7. *Let us consider a sub-Riemannian geometry* $(\mathbb{R}^N, \mathcal{H}, <\ ,\ >_{\mathcal{H}})$ *such that* $\mathcal{H}$ *is generated by* m *vector fields* $X_1, \dots, X_m$*. We say that* $X_1, \dots, X_m$*,* $m < N$*, are Carnot-type vector fields if the coefficients of* X_i *are 0 for* $j \in \{1, \dots, m\} \setminus \{i\}$*, the* i*-component is equal to 1 and the other* $N - m$ *components are polynomial in* $x_1, \dots, x_m$*.*

The previous structure allows us to consider an easy and explicit Riemannian approximation. Nevertheless the approach of this paper applies also to the case where this additional structure is not fulfilled. This structure applies to a large class of geometries. The Heisenberg group introduced in Example

8.6 is obviously a Carnot-type geometry but more in general all Carnot groups are Carnot-type geometries (see [4] for definitions and properties).

For later use we also introduce the matrix associated to the vector fields $X_1, \dots, X_m$, which is the $N \times m$ matrix defined as

$$\sigma(x) = [X_1(x), \dots, X_m(x)]^T. \tag{8.1}$$

For Carnot-type geometries for later we introduce the matrix σ associated to the vector fields, i.e. $\sigma = [X_1, \dots, X_m]$ that takes form

$$\sigma(x) = \begin{bmatrix} I_{m\times m} & A(x_1, \dots, x_m) \end{bmatrix}, \tag{8.2}$$

where the matrix $A(x_1, \dots, x_m)$ is a $(N-m) \times m$ depending only on the first m components of x.

Example 8.8. *In the case of the Heisenberg group introduced in Example 8.6, the matrix σ is given by*

$$\sigma(x) = \begin{bmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \end{bmatrix}, \quad \forall\, x = (x_1, x_2, x_3) \in \mathbb{R}^3.$$

We now want to introduce the Riemannian approximation, which will be crucial for our results.

Let us consider a family $\mathcal{X} = \{X_1, \dots, X_m\}$ of Carnot-type vector fields defined on $\mathbb{R}^N$, with $m < N$, satisfying the Hörmander condition. It is possible to complete the distribution $\mathcal{H}$ by adding $N - m$ vector fields $X_{m+1}, \dots, X_N$, i.e.

$$\mathrm{Span}\big(X_1(x), \dots, X_m(x), X_{m+1}(x), \dots, X_N(x)\big) = T_x\mathbb{R}^N \equiv \mathbb{R}^N,\ \forall\, x \in \mathbb{R}^N.$$

The geometry induced, for all $\varepsilon > 0$, by the distribution

$$\mathcal{H}_\varepsilon = \{X_1, \dots, X_m, \varepsilon X_{m+1}, \dots, \varepsilon X_N\}$$

is called *Riemannian approximation* of the sub-Riemannian topology induced on $\mathbb{R}^N$ by $\mathcal{H}$. The associated matrix is

$$\sigma_\varepsilon(x) = [X_1(x), \dots X_m(x), \varepsilon X_{m+1}(x) \dots, \varepsilon X_N(x)]^T. \tag{8.3}$$

Note that $\det(\sigma_\varepsilon(x)) \neq 0$.
In the case of Carnot-type geometries, then we can always choose

$$X_i(x) = e_i, \quad \forall i = m+1, \dots, N, \quad \forall x \in \mathbb{R}^N,$$

where by e_i we indicate the standard Euclidean unit vector for the i-th component.

Example 8.9 (Riemannian approximation of $\mathbb{H}^1$)**.** *In the case of the Heisenberg group introduced in Example 8.6, the matrix associated to the Riemannian approximation is*

$$\sigma_\varepsilon(x) = \begin{bmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \\ 0 & 0 & \varepsilon \end{bmatrix} \qquad \forall x = (x_1, x_2, x_3) \in \mathbb{R}^3.$$

This technique is called Riemannian approximation since, as $\varepsilon \to 0^+$, then the geometry induced by Riemannian approximation converges, in sense of Gromov-Hausdorff (see [36] for further details), to the original sub-Riemannian geometry (as showed, as example, in [13]).

8.3 Horizontal mean curvature evolution

Given a smooth hypersurface Γ, we indicate by n_E the standard (Euclidean) normal to Γ at the point x. Since the vector fields
$X_1(x), \dots, X_m(x), X_{m+1}(x), \dots, X_N(x)$ span the whole of $\mathbb{R}^N$ at any point, then n_E can be written as

$$n_E(x) = \frac{\sum_{i=1}^N \alpha_i(x) X_i(x)}{\sqrt{\sum_{i=1}^N \alpha_i^2(x)}}. \tag{8.4}$$

The following definitions will be key for this paper.

Definition 8.10. *Given a smooth hypersurface Γ, the* horizontal normal *is the renormalized projection of the Euclidean normal on the horizontal space*

$\mathcal{H}_x$, *i.e.*

$$n_0(x) := \frac{\alpha_1(x)X_1(x) + \cdots + \alpha_m(x)X_m(x)}{\sqrt{\alpha_1^2(x) + \cdots + \alpha_m^2(x)}} \in \mathcal{H}_x \subset \mathbb{R}^N.$$

With an abuse of notation we will often indicate by n_0 *the associated* m*-valued vector*

$$n_0(x) = \frac{(\alpha_1(x), \ldots, \alpha_m(x))^T}{\sqrt{\alpha_1^2(x) + \cdots + \alpha_m^2(x)}} \in \mathbb{R}^m. \tag{8.5}$$

The main difference between the horizontal normal and a standard normal is that the first may not exist even for smooth hypersurfaces. In fact at some points the horizontal normal is not defined meanwhile the Euclidean one exists. These points are called characteristic points.

Definition 8.11. *Given a smooth hypersurface* Γ*, characteristic points occur whenever* n_E *is orthogonal to the horizontal plane* $\mathcal{H}_x$*, then its projection on such a subspace vanishes, i.e.*

$$\alpha_1^2(x) + \cdots + \alpha_m^2(x) = 0,$$

where $\alpha_i(x)$ *are the coordinates of* $n_E(x)$ *w.r.t. the basis* $X_1, \ldots, X_N$ *as in* (8.4)

Note that characteristic points do not exist in the associated Riemannian approximation, in fact, whenever Γ is smooth, the normal is defined at any point, which means

$$\sum_{i=1}^{N} \alpha_i^2(x) \neq 0, \quad \forall x \in \Gamma.$$

(further details in Chapter 4). We recall that for every smooth hypersurface the *mean curvature* is defined as the divergence of the Euclidean normal. Similarly, for every smooth hypersurface, we can now introduce the horizontal mean curvature.

Definition 8.12. *Given a smooth hypersurface* Γ *and a non characteristic point* $x \in \Gamma$*, the* horizontal mean curvature *is defined as the horizontal divergence*

of the horizontal normal, i.e. $k_0(x) = div_{\mathcal{H}} n_0(x)$, *where* $n_0(x)$ *is given* (8.5) *while* $div_{\mathcal{H}}$ *is the divergence w.r.t. the vector fields* $X_1, \dots, X_m$, *i.e.*

$$k_0(x) = X_1 \left(\frac{\alpha_1(x)}{\sqrt{\sum_{i=1}^m \alpha_i^2(x)}} \right) + \cdots + X_m \left(\frac{\alpha_m(x)}{\sqrt{\sum_{i=1}^m \alpha_i^2(x)}} \right).$$

Obviously the horizontal mean curvature is never defined at characteristic points, since there the horizontal normal does not exist.

Definition 8.13. *Let* Γ_t *be a family of smooth hypersurfaces in* $\mathbb{R}^N$ *depending on a positive parameter* $t > 0$. *We say that* Γ_t *is an* evolution by horizontal mean curvature flow *of* Γ *if and only if* $\Gamma_0 = \Gamma$ *and for any smooth horizontal curve* $\gamma : [0, T] \to \mathbb{R}^N$ *such that* $\gamma(t) \in \Gamma_t$ *for all* $t \in [0, T]$, *the horizontal normal velocity* v_0 *is equal to minus the horizontal mean curvature, i.e.*

$$v_0(\gamma(t)) := -k_0(\gamma(t)) n_0(\gamma(t)), \tag{8.6}$$

where $n_0(x(t))$ *and* $k_0(x(t))$ *as respectively the horizontal normal and the horizontal mean curvature defined by Definitions 8.10 and 8.12.*

Note that (8.6) is never defined at characteristic points.

In this subsection we consider a smooth hypersurface Γ parametrized as zero level set

$$\Gamma = \big\{ x \in \mathbb{R}^N | u(x) = 0 \big\},$$

for some smooth function $u : \mathbb{R}^N \to \mathbb{R}$. Then the Euclidean normal is simply $n_E(x) = \frac{\nabla u(x)}{|\nabla u(x)|}$, which implies that the horizontal normal can be expressed as

$$n_0(x) = \left(\frac{X_1 u(x)}{\sqrt{\sum_{i=1}^m (X_i u(x))^2}}, \dots, \frac{X_m u(x)}{\sqrt{\sum_{i=1}^m (X_i u(x))^2}} \right). \tag{8.7}$$

Similarly we can then write the horizontal mean curvature (HMCF) as

$$k_0(x) = \sum_{i=1}^m X_i \left(\frac{X_i u(x)}{\sqrt{\sum_{i=1}^m (X_i u(x))^2}} \right). \tag{8.8}$$

Let us consider

$$\Gamma = \{ (x, t) \in \mathbb{R}^N \times (0, \infty) |\ u(x, t) = 0 \}. \tag{8.9}$$

Applying (8.8) to Definition 8.13 we obtain that u solves the following PDE

$$u_t = Tr((\mathcal{X}^2 u)^*) - \left\langle (\mathcal{X}^2 u)^* \frac{\mathcal{X}u}{|\mathcal{X}u|}, \frac{\mathcal{X}u}{|\mathcal{X}u|} \right\rangle, \tag{8.10}$$

where $<\ ,\ >$ is the standard Euclidean inner product, $\mathcal{X}u$ the horizontal gradient

$$\mathcal{X}u := (X_1 u, \ldots, X_m u)^T$$

and $(\mathcal{X}^2 u)^*$ is the symmetrized horizontal Hessian, that is

$$((\mathcal{X}^2 u)^*)_{ij} := \frac{X_i(X_j u) + X_j(X_i u)}{2} \quad \text{where } i, j = 1, \ldots, m.$$

Definition 8.14. *Given* $\Gamma_0 = \{x \in \mathbb{R}^N | u(x) = 0\}$ *we say that* $\Gamma_t = \{(x,t) \in \mathbb{R}^N \times (0,\infty) | u(x,t) = 0\}$ *is a (generalized) evolution by HMCF iff* $\Gamma_0 = \Gamma$ *and* u *is an solution of* (8.10) *(see [22] for further details).*

Equation (8.10) can be approximated to a Riemannian mean curvature flow using the Riemannian approximation as seen in Section 8.2. This leads the following generalizations of horizontal normal and horizontal divergence.

Definition 8.15. *Given a smooth hypersurface* Γ*, the* approximated Riemannian normal *is*

$$n_\varepsilon(x) := \frac{\sum_{i=1}^m \alpha_i(x) X_i(x) + \varepsilon \sum_{i=m+1}^N \alpha_i(x) X_i(x)}{\sqrt{\alpha_1^2(x) + \cdots + \alpha_m^2(x) + \varepsilon^2 \alpha_{m+1}^2(x) + \cdots + \varepsilon^2 \alpha_N^2(x)}} \in \mathcal{H}_x \subset \mathbb{R}^N.$$

Definition 8.16. *The* approximated Riemannian mean curvature *is defined the divergence w.r.t.* $X_1, \ldots, X_m, \varepsilon X_{m+1}, \ldots, \varepsilon X_N$ *on* n_ε.

Definition 8.15 allows us to introduce the approximated Riemannian mean curvature flow. For sake of simplicity we write that directly for hypersurface in level set formulation for Γ_t.

Definition 8.17. *We say that* $\Gamma_t = \{(x,t) \in \mathbb{R}^N \times (0,\infty) | u(x,t) = 0\}$ *is a (generalized) evolution by approximated Riemannian mean curvature flow if and only if* $\Gamma_0 = \Gamma$ *and is viscosity solution of*

$$u_t = Tr((\mathcal{X}_\varepsilon^2 u)^*) - \left\langle (\mathcal{X}_\varepsilon^2 u)^* \frac{\mathcal{X}_\varepsilon u}{|\mathcal{X}_\varepsilon u|}, \frac{\mathcal{X}_\varepsilon u}{|\mathcal{X}_\varepsilon u|} \right\rangle = \Delta_\varepsilon u - \Delta_{0,\infty,\varepsilon} u, \tag{8.11}$$

where

$$\mathcal{X}_\varepsilon u = (X_1 u, \dots, X_m u, \varepsilon X_{m+1} u, \dots, \varepsilon X_N u)$$

and

$$(\mathcal{X}_\varepsilon^2 u)^*_{ij} = \frac{X_i^\varepsilon (X_j^\varepsilon u) + X_j^\varepsilon (X_i^\varepsilon u)}{2}. \tag{8.12}$$

The approximated Riemannian stochastic control problem

Let us consider a family of smooth vector fields $\mathcal{H} = \{X_1, \dots X_m\}$ and its Riemannian approximation $\mathcal{H}_\varepsilon = \{X_1, \dots, X_m, \varepsilon X_{m+1}, \dots, \varepsilon X_N\}$.

Definition 8.18. *We define the* horizontal Brownian motion *the solution of the process*

$$d\xi = \sum_{i=1}^{m} X_i(\xi) \circ dB_m^i,$$

where B_m is a m-dimensional Brownian motion and $\circ$ the Stratonovich differential. We define the Riemannian approximated horizontal Brownian motion *as*

$$d\xi^\varepsilon = \sum_{i=1}^{N} X_i^\varepsilon(\xi^\varepsilon) \circ dB_N^i$$

where B_N is an N-dimensional Brownian motion.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space, B_i is a i-dimensional Brownian motion adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ with $i = m, N$, we recall that a *predictable* variable is a time-continuous stochastic process $\{\xi_t\}_{t\geq 0}$ defined on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, measurable with respect to the σ-algebra generated by all left-continuous adapted process. Given a smooth function $g : \mathbb{R}^N \to \mathbb{R}$ (which parametrizes the starting hypersurface at time $t = 0$) we introduce the function $V : [0, T] \times \mathbb{R}^N \to \mathbb{R}$ defined as

$$V(t, x) := \inf_{\nu \in \mathcal{A}} ess \sup_{\omega \in \Omega} g(\xi^{t,x,\nu}(T)(\omega)), \tag{8.13}$$

and $\xi^{t,x,\nu}$ is the solution of the stochastic dynamic

$$\begin{cases} d\xi^{t,x,\nu}(s) = \sqrt{2}\sigma^T(\xi^{t,x,\nu}(s)) \circ dB_m^{\nu}(s), \qquad s \in (t,T], \\ dB_m^{\nu}(s) = \nu(s) dB_m(s), \\ \xi^{t,x,\nu}(t) = x, \end{cases} \tag{8.14}$$

where the matrix σ is defined in (8.1) and

$$\mathcal{A} = \big\{\nu : [t,T] \to Sym(m) | \nu \geq 0,\ I_m - \nu^2 \geq 0,\ Tr(I_m - \nu^2) = 1\big\}. \tag{8.15}$$

Let us observe that it is possible to rewrite (8.14) explicitly

$$\begin{cases} d\xi^{t,x,\nu}(s) = \sqrt{2}\sigma^T(\xi^{t,x,\nu}(s)) dB_m(s) + \sum_{i,j=1}^m (\nu^2(s))_{ij} \nabla_{X_i} X_j(\xi^{t,x,\nu}(s)),\ s \in (t,T], \\ \xi^{t,x,\nu}(t) = x, \end{cases} \tag{8.16}$$

i.e. the SDE can be written as a sum of a *stochastic term* and a *drift term*. Similarly, for $\varepsilon > 0$ we introduce the function $V^{\varepsilon} : [0,T] \times \mathbb{R}^N \to \mathbb{R}$ defined by

$$V^{\varepsilon}(t,x) := \inf_{\nu \in \mathcal{A}_1} ess \sup_{\omega \in \Omega} g(\xi_{\varepsilon}^{t,x,\nu_1}(T)(\omega)), \tag{8.17}$$

where $\xi_{\varepsilon}^{t,x,\nu_1}$ is the solution of

$$\begin{cases} d\xi_{\varepsilon}^{t,x,\nu_1}(s) = \sqrt{2}\sigma_{\varepsilon}^T(\xi_{\varepsilon}^{t,x,\nu_1}(s)) \circ dB_N^{\nu_1}(s), \qquad s \in (t,T], \\ dB_N^{\nu_1}(s) = \nu_1(s) dB_N(s), \\ \xi_{\varepsilon}^{t,x,\nu_1}(t) = x, \end{cases} \tag{8.18}$$

σ_{ε} is the matrix defined in (8.3) and

$$\mathcal{A}_1 = \big\{\nu_1 : [t,T] \to Sym(N) |\ \nu_1 \geq 0,\ I_N - \nu_1^2 \geq 0,\ Tr(I_N - \nu_1^2) = 1\big\}. \tag{8.19}$$

It is possible to rewrite (8.18) explicitly as did with (8.14). It is possible to show that the functions V and V^{ε} solve in the viscosity sense for equation for the evolution by HMCF as introduced in Definition 8.14 and the level set equation for the Riemannian approximation of the HMCF as introduced in Definition 8.17 (see [22, 35] for further details).

Note also that the sets of controls (8.15) and (8.19) can be rewritten respectively as

$$\mathcal{A} = \{\nu^2|\ \nu \in \mathcal{A}\} = Co\{I_m - a \otimes a|\ a \in \mathbb{R}^m,\ \ |a| = 1\},$$

and

$$\mathcal{A}_1 = \{\nu_1^2|\ \nu_1 \in \mathcal{A}_1\} = Co\{I_N - \overline{a} \otimes \overline{a}|\ \overline{a} \in \mathbb{R}^N,\ \ |\overline{a}| = 1\},$$

where Co is the convex hull (see [27] for more details).

Next we introduce the p-regularising approximation of the functions V and V^ε.

Definition 8.19. *For $p > 1$, the p-value function associated to the value function* (8.13) *is defined as*

$$V_p(t,x) := \inf_{\nu \in \mathcal{A}} \mathbb{E}[|g(\xi^{t,x,\nu})(T)(\omega)|^p]^{\frac{1}{p}}, \tag{8.20}$$

and, similarly, we can introduce the following ε-p-regularising function, that is

$$V_p^\varepsilon(t,x) := \inf_{\nu_1 \in \mathcal{A}} \mathbb{E}[|g(\xi_\varepsilon^{t,x,\nu_1})(T)(\omega)|^p]^{\frac{1}{p}}. \tag{8.21}$$

Remark 8.20. *The function V_p solves in viscosity sense the following PDE:*

$$\begin{cases} -(V_p) + H_p(x, DV_p, D^2V_p) = 0, & t \in [0,T),\ x \in \mathbb{R}^3, \\ V_p(T,x) = g(x), & x \in \mathbb{R}^3 \end{cases}$$

where

$$H_p(x,q,M) := \sup_{\nu \in \mathcal{A}} \Big[-(p-1) r^{-1} Tr[\nu\nu^T qq^T] + Tr[\nu\nu^T M] \Big], \tag{8.22}$$

see [27]. Similarly for $\varepsilon > 0$ and $p > 1$ fixed, V_p^ε solves in the viscosity sense

$$\begin{cases} -(V_p^\varepsilon) + H_p^\varepsilon(x, DV_p^\varepsilon, D^2V_p^\varepsilon) = 0, & t \in [0,T),\ x \in \mathbb{R}^N, \\ V_p^\varepsilon(T,x) = g(x), & x \in \mathbb{R}^N \end{cases}$$

where

$$H_p^\varepsilon(x,r,q,M) := H_p(x,r,q_\varepsilon,M_\varepsilon) = \sup_{\nu\in\mathcal{A}_1}\left[-(p-1)r^{-1}Tr[\nu\nu^T q_\varepsilon q_\varepsilon^T]+Tr[\nu\nu^T M_\varepsilon]\right], \tag{8.23}$$

$\mathcal{A}_1$ *is given in* (8.19) *and, for all* $q\in\mathbb{R}^N$ *and* $M=(M_{ij})_{i,j=1}^N \in Sym(N)$,

$$q_\varepsilon := \begin{bmatrix} q_1 \\ \dots \\ q_m \\ \varepsilon q_{m+1} \\ \dots \\ \varepsilon q_N \end{bmatrix} \tag{8.24}$$

and

$$M_\varepsilon := \begin{bmatrix} M_{11} & \dots & M_{1m} & \varepsilon M_{1(m+1)} & \dots & \varepsilon M_{1N} \\ & & & \vdots & & \\ M_{m1} & \dots & M_{mm} & \varepsilon M_{(m+1)m} & \dots & \varepsilon M_{Nm} \\ \varepsilon M_{(m+1)1} & \dots & \varepsilon M_{(m+1)m} & \varepsilon^2 M_{(m+1)(m+1)} & \dots & \varepsilon^2 M_{(m+1)N} \\ & & & \vdots & & \\ \varepsilon M_{1N} & \dots & \varepsilon M_{mN} & \varepsilon^2 M_{(m+1)N} & \dots & \varepsilon^2 M_{NN} \end{bmatrix}.$$

8.4 Some properties of σ_ε

Next we show some properties for the matrices σ and σ_ε defined as in (8.1) and (8.3).

We will consider the following assumptions.

The hypersurface Γ is smooth. (A1)

The matrix σ is σ_ε are globally Lipschitz and bounded. (A2)

$\nabla_{X_i}X_j$ and $\nabla_{X_i^\varepsilon}X_j^\varepsilon$ are globally Lipschitz and bounded for all $i,j\in\{1,\dots,N\}$ (A3)

We recall that the structure of σ_ε is given by

$$\sigma_\varepsilon(x) = \begin{bmatrix} I_{m\times m} & A(x) \\ 0_{m\times(N-m)} & \varepsilon I_{(N-m)\times(N-m)} \end{bmatrix}, \tag{8.25}$$

where the matrix $A(x)$ is composed by smooth and globally bounded coefficients and that the norm of a $m \times N$ matrix A is

$$\|A\| := \sum_{j=1}^{m}\sum_{i=1}^{N} |a_{ij}|. \tag{8.26}$$

For all σ associated to Carnot-type vector fields geometry (see [4] for further details), we introduce the following extended $N \times N$ matrix

$$\overline{\sigma}(x) := \begin{bmatrix} I_{m\times m} & A(x) \\ 0_{m\times(N-m)} & 0_{(N-m)\times(N-m)} \end{bmatrix}. \tag{8.27}$$

We start proving the following technical lemmas.

Lemma 8.21. *Given σ bounded and associated to a Carnot type geometry as in* (8.1) *and σ_ε defined as in* (8.25)*, then, for all $\varepsilon > 0$*

$$\|\sigma_\varepsilon\| \le \|\sigma\| + \varepsilon(N-m). \tag{8.28}$$

Proof. We recall that the matrix σ_ε has components

$$(\sigma_\varepsilon)_{ij}(x) = \begin{cases} \delta_{ij}, & \text{for } \ i,j = 1,\dots,m, \\ a_{ij}(x), & \text{for } i = m\dots N, \text{ and } j = 1,\dots,m, \\ 0, & \text{for } i = 1\dots m, \text{ and } j = m,\dots,N, \\ \varepsilon\delta_{ij}, & \text{for } \ i,j = m+1,\dots,N, \end{cases} \tag{8.29}$$

By assumption that σ is bounded, i.e. there exists a $C_\sigma > 0$ such that

$$\|\sigma\| \le C_\sigma,$$

hence by (8.29) we deduce

$$\|\sigma_\varepsilon\| \le \|\sigma\| + \varepsilon(N-m) \le C_\sigma + \varepsilon(N-m).$$

□

Lemma 8.22. *Given the matrices σ and σ_ε defined as in* (8.2) *and* (8.25)*, if σ is globally Lipschitz, then σ_ε is globally Lipschitz and*

$$Lip(\sigma_\varepsilon) = Lip(\sigma).$$

Proof. For any $x, y \in \mathbb{R}^N$

$$\begin{aligned}\|\sigma_\varepsilon(x)-\sigma_\varepsilon(y)\| &= \left\| \begin{bmatrix} I_{m\times m} & A(x) \\ 0_{(N-m)\times m} & \varepsilon I_{(N-m)\times(N-m)} \end{bmatrix} - \begin{bmatrix} I_{m\times m} & A(y) \\ 0_{(N-m)\times m} & \varepsilon I_{(N-m)\times(N-m)} \end{bmatrix} \right\| \\ &= \left\| \begin{bmatrix} 0_{m\times m} & A(x)-A(y) \\ 0_{m\times(N-m)} & 0_{(N-m)\times(N-m)} \end{bmatrix} \right\| = \|A(x)-A(y)\| \le Lip(A)|x-y|,\end{aligned}$$

where $Lip(A) = Lip(\sigma)$ and so $Lip(\sigma_\varepsilon) = Lip(\sigma)$. □

We conclude with an estimate about the difference between σ_ε and $\overline{\sigma}$.

Lemma 8.23. *Given σ_ε and $\overline{\sigma}$ defined as in* (8.25) *and* (8.27) *, then*

$$\|\sigma_\varepsilon(x) - \overline{\sigma}(x)\| = \varepsilon(N-m), \qquad \textit{for all } \; x \in \mathbb{R}^N.$$

Proof. We can deduce immediately that

$$\begin{aligned}\|\sigma_\varepsilon(x) - \overline{\sigma}(x)\| &= \left\| \begin{bmatrix} 0_{m\times m} & 0_{m\times(N-m)} \\ 0_{(N-m)\times m} & \varepsilon I_{(N-m)\times(N-m)} \end{bmatrix} \right\| \\ &= \varepsilon \|I_{(N-m)\times(N-m)}\| = \varepsilon(N-m).\end{aligned}$$

□

We recall that it holds true $\|A\| = \|A^T\|$, where $A \in M_{N,N}$. Therefore it is possible to apply the properties of Lemmas 8.21, 8.22 and 8.23 to the transpose of σ and σ_ε.

Convergence in the case of vanishing drift

In this section we prove the convergence of V^ε to V up to a subsequence $\varepsilon = \varepsilon(p) \to 0$ depending on p.

Theorem 8.24. *Assume (A1) and (A2), and let $g : \mathbb{R}^N \to \mathbb{R}$ be a globally Lipschitz function and consider V_p as defined in (8.20) with $\xi^{t,x,\nu}$ solution of (8.14) and, for $\varepsilon > 0$, V_p^ε defined in (8.21) with $\xi_\varepsilon^{t,x,\tilde{\nu}}$ solution of (8.18). Then, there exists a sequence $\varepsilon(p) \to 0$, as $p \to \infty$, such that*

$$\limsup_{p\to\infty} V_p^{\varepsilon(p)}(t,x) \le V(t,x).$$

Before proving the theorem, we need the following lemmas.

Lemma 8.25. *Let $\nu \in \mathcal{A}$ be an admissible control, then*

$$\nu_1(s) := \begin{bmatrix} \nu(s) & 0_{m\times(N-m)} \\ 0_{(N-m)\times m} & I_{(N-m)\times(N-m))} \end{bmatrix}, \ s \in [0,T]. \tag{8.30}$$

is an admissible control for $\mathcal{A}_1$.

Proof. Since $\nu \in Sym(m)$, trivially $\nu_1 \in Sym(N)$ and $\nu_1 \ge 0$. It remains to check that $I_N - \nu_1^2 \ge 0$ and $Tr(I_N - \nu_1^2) = 1$. At this purpose, we explicitly compute

$$I_N - \nu_1^2 = \begin{bmatrix} I_m - \nu^2 & 0_{m\times(N-m)} \\ 0_{m\times(N-m)} & 0_{(N-m)\times(N-m)} \end{bmatrix}.$$

Thus, recalling that $I_m - \nu^2 \ge 0$ and $Tr(I_m - \nu^2) = 1$, the result follows. □

Lemma 8.26. *Let us consider a family $\mathcal{F}$ of vectors $v \in \mathbb{R}^N$ such that the elements $v \in \mathcal{F}$ are uniformly bounded by a positive constant, i.e. $|v| \le C \ \ \forall v \in \mathcal{F}$, then the family of matrices*

$$A_v := I_n - v \otimes v$$

is uniformly bounded by a constant depending only on C and N.

Proof. Recall that $(v \otimes v)_{ij} = v_i v_j$ and $\sum_{i=1}^N v_i^2 \le C$ which implies that

$$|v_i| \le \sqrt{C} \ \text{ for all } \ i = 1, \dots, N.$$

Thus $|(v \otimes v)_{ij}| = |v_i||v_j| \leq C$ which implies $\|v \otimes v\| = \sum_{i,j=1}^{N} |v_i||v_j| \leq N^2 C$. Hence we can conclude that

$$\|I_N - v \otimes v\| \leq \|I_N\| + \|v \otimes v\| \leq N + N^2 C.$$

□

Remark 8.27. *We can apply Lemma 8.26 to all the admissible controls in $\mathcal{A}_1$ and $\mathcal{A}$. Note that in this case $|v| \leq 1$ which implies $\|\nu\| \leq N + N^2$.*

Lemma 8.28. *Given a control $\nu \in \mathcal{A}$ and the corresponding $\nu_1 \in \mathcal{A}_1$ defined by (8.30), we consider $\xi^{t,x,\nu}$ solution of*

$$\begin{cases} d\xi^{t,x,\nu}(s) = \sqrt{2}\sigma^T(\xi^{t,x,\nu}(s))\nu(s)dB_m(s), \quad s \in (t,T], \\ \xi^{t,x,\nu}(t) = x, \end{cases}$$

where σ is as in (8.2) and $\overline{\xi}^{t,x,\nu_1}$ solves

$$\begin{cases} d\overline{\xi}^{t,x,\nu_1}(s) = \sqrt{2}\overline{\sigma}^T(\overline{\xi}^{t,x,\nu_1}(s))\nu_1(s)dB_N(s), \quad s \in (t,T], \\ \overline{\xi}^{t,x,\nu_1}(t) = x, \end{cases}$$

where $\overline{\sigma}$ is defined in (8.27). Then $\xi^{t,x,\nu} = \overline{\xi}^{t,x,\nu_1}$ for a.e. $s \in (t,T]$.

Proof. First, we write the trajectory solution of the two dynamics in integral form, i.e.

$$\xi^{t,x,\nu}(s) = x + \sqrt{2}\int_t^s \sigma^T(\xi^{t,x,\nu}(r))\nu(r)dB_m(r),$$

and

$$\overline{\xi}^{t,x,\nu_1}(s) = x + \sqrt{2}\int_t^s \overline{\sigma}^T(\overline{\xi}^{t,x,\nu_1}(r))\nu_1(r)dB_N(r).$$

Note that

$$\begin{aligned} \overline{\sigma}^T(\overline{\xi}^{t,x,\nu_1}(s))\nu_1(s) &= \begin{bmatrix} I_{m\times m} & 0_{m\times(N-m)} \\ A(\overline{\xi}^{t,x,\nu_1}(s)) & 0_{(N-m)\times m} \end{bmatrix} \begin{bmatrix} \nu & 0_{m\times(N-m)} \\ 0_{(N-m)\times n} & I_{(N-m)} \end{bmatrix} \\ &= \begin{bmatrix} \nu & 0_{m\times(N-m)} \\ A(\overline{\xi}^{t,x,\nu_1}(s))\nu & 0_{(N-m)\times(N-m)} \end{bmatrix}, \end{aligned}$$

which implies

$$\overline{\sigma}^T(\overline{\xi}^{t,x,\nu_1}(s))\nu_1(s)dB_N(s) = \sigma^T(\overline{\xi}^{t,x,\nu_1}(s))\nu(s)dB_m(s). \tag{8.31}$$

So, by using (8.31), we obtain

$$\xi^{t,x,\nu}(s) - \overline{\xi}^{t,x,\nu_1}(s) = \sqrt{2}\int_t^s \Big(\overline{\sigma}^T(\xi^{t,x,\nu}(r)) - \overline{\sigma}^T(\xi^{t,x,\nu_1}(r))\Big)\nu_1(r)dB_N(r),$$

which implies

$$\mathbb{E}[|\xi^{t,x,\nu}(s)-\overline{\xi}^{t,x,\nu_1}(s)|^2] = 2\mathbb{E}\left[\left\| \int_t^s \Big(\overline{\sigma}^T(\xi^{t,x,\nu}(r))-\overline{\sigma}^T(\xi^{t,x,\nu_1}(r))\Big)\nu_1(r)dB_N(r)\right\|^2\right]. \tag{8.32}$$

Applying Itô's isometry to the right side of (8.32) we get

$$\mathbb{E}[|\xi^{t,x,\nu}(s)-\overline{\xi}^{t,x,\nu_1}(s)|^2] = 2\mathbb{E}\left[\int_t^s \left\|\Big(\overline{\sigma}^T(\xi^{t,x,\nu}(r))-\overline{\sigma}^T(\xi^{t,x,\nu_1}(r))\Big)\nu_1(r)\right\|^2 dr\right].$$

Then, by using that σ is globally Lipschitz (by assumption (A2)) and ν_1 is bounded (by Lemma 8.26), we deduce

$$\begin{aligned}\mathbb{E}[|\xi^{t,x,\nu}(s) - \overline{\xi}^{t,x,\nu_1}(s)|^2] &\leq 2(Lip(\sigma)C)^2\mathbb{E}\left[\int_t^s |\xi^{t,x,\nu}(r) - \overline{\xi}^{t,x,\nu_1}(r)|^2 dr\right] \\ &= 2(Lip(\sigma)C)^2\left[\int_t^s \mathbb{E}[|\xi^{t,x,\nu}(r) - \overline{\xi}^{t,x,\nu_1}(r)|^2]dr\right],\end{aligned}$$

where $C > 0$ is a real number. Applying Gronwall's Lemma with $u(s) := \mathbb{E}[|\xi^{t,x,\nu}(s) - \overline{\xi}^{t,x,\nu_1}(s)|^2]$ and observing that the inequality is in the form $u(s) \leq C\int_t^s u(\tau)d\tau + 0$ we obtain (see Appendix A for further details)

$$\mathbb{E}[|\xi^{t,x,\nu}(s) - \xi^{t,x,\nu_1}(s)|^2] \leq 0,$$

which implies that $\xi^{t,x,\nu} = \xi^{t,x,\nu_1}$ for a.e. $s \in [t,T]$. □

We are able to prove now Theorem 8.24.

Notation: In the proof we write $\mathbb{E}[|f|^p]^{\frac{1}{p}}$ meaning $(\int_\Omega |f|^p)^{\frac{1}{p}}$.

Proof of Theorem 8.24. Let ν_p be the optimal control realizing the infimum in $V_p(t,x)$. We define the associated control $\nu_1 := \nu_1(\nu_p)$ as in (8.30) (omitting the

dependence on p in the notation). By Lemma 8.28 we know that $\xi^{t,x,\nu_p} = \overline{\xi}^{t,x,\nu_1}$. Using the Minkwoski's Inequality for L^p norms and the assumption that g is globally Lipschitz we obtain

$$\begin{aligned}
&V_p^\varepsilon(t,x) - V_p(t,x) = \mathbb{E}[|g(\xi_\varepsilon^{t,x,\nu_1})(T)|^p]^{\frac{1}{p}} - \mathbb{E}[|g(\xi^{t,x,\nu_p})(T)|^p]^{\frac{1}{p}} \\
&= \mathbb{E}[|g(\xi_\varepsilon^{t,x,\nu_1})(T)|^p]^{\frac{1}{p}} - \mathbb{E}[|g(\overline{\xi}^{t,x,\nu_1})(T)|^p]^{\frac{1}{p}} \leq \mathbb{E}[|g^p(\xi_\varepsilon^{t,x,\nu_1})(T) - g^p(\overline{\xi}^{t,x,\nu_1})(T)|^{\frac{1}{p}}] \\
&\leq Lip(g)\mathbb{E}[|\xi_\varepsilon^{t,x,\nu_1}(T) - \overline{\xi}^{t,x,\nu_1}(T)|^p]^{\frac{1}{p}}.
\end{aligned}$$

For sake of simplicity, we write $\xi_\varepsilon := \xi_\varepsilon^{t,x,\nu_1}$ and $\overline{\xi} := \overline{\xi}^{t,x,\nu_1}$. We recall that, by applying the Itô's Lemma to the differentiable function $G(x) = |x|^p$, we have

$$\begin{aligned}
&\mathbb{E}[|\xi_\varepsilon(T) - \overline{\xi}(T)|^p] \\
&= \mathbb{E}\left[\int_t^T p|\xi_\varepsilon(\tau) - \overline{\xi}(\tau)|^{p-2}(\xi_\varepsilon(\tau) - \overline{\xi}(\tau))\Big(\sigma_\varepsilon^T(\xi_\varepsilon(\tau))\nu_1(\tau) - \overline{\sigma}^T(\overline{\xi}(\tau))\nu_1(\tau)\Big)dB_N(\tau)\right] \\
&+\mathbb{E}\left[\int_t^T p(p-1)|\xi_\varepsilon(\tau) - \overline{\xi}(\tau)|^{p-2} Tr\Bigg(\Big((\sigma_\varepsilon^T(\xi_\varepsilon(\tau)) - \overline{\sigma}^T(\overline{\xi}(\tau)))\nu_1(\tau)\Big)\right. \\
&\left.\Big((\sigma_\varepsilon^T(\xi_\varepsilon(\tau)) - \overline{\sigma}^T(\overline{\xi}(\tau)))\nu_1(\tau)\Big)^T\Bigg)d\tau\right]. \qquad (8.33)
\end{aligned}$$

The first term on the right side of (8.33) vanishes because the expectation of an Itô integral is always zero. For the second term we recall that trivially $Tr(SS^T) = \|S\|^2$, and that

$$\|\sigma_\varepsilon^T(\xi_\varepsilon(\tau)) - \overline{\sigma}^T(\overline{\xi}(\tau))\|^2 = \|\sigma_\varepsilon^T(\xi_\varepsilon(\tau)) \pm \sigma_\varepsilon^T(\overline{\xi}(\tau)) - \overline{\sigma}^T(\overline{\xi}(\tau))\|^2. \qquad (8.34)$$

Applying the triangular inequality and the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ with $a = \|\sigma_\varepsilon^T(\xi_\varepsilon(\tau)) - \sigma_\varepsilon^T(\overline{\xi}(\tau))\|$, $b = \|\sigma_\varepsilon^T(\overline{\xi}(\tau)) - \overline{\sigma}^T(\overline{\xi}(\tau))\|$ we obtain

$$\begin{aligned}
&\|\sigma_\varepsilon^T(\xi_\varepsilon(\tau)) \pm \sigma_\varepsilon^T(\overline{\xi}(\tau)) - \overline{\sigma}^T(\overline{\xi}(\tau))\|^2 \\
&\leq 2\|\sigma_\varepsilon^T(\xi_\varepsilon(\tau)) - \sigma_\varepsilon^T(\overline{\xi}(\tau))\|^2 + 2\|\sigma_\varepsilon^T(\overline{\xi}(\tau)) - \overline{\sigma}^T(\overline{\xi}(\tau))\|^2.
\end{aligned}$$

Recalling that, by assumption, σ is a globally Lipschitz matrix and by applying Lemma 8.23 we have

$$\|\sigma_\varepsilon^T(\xi_\varepsilon(\tau)) - \overline{\sigma}^T(\overline{\xi}(\tau))\|^2 \leq 2Lip(\sigma)^2|\xi_\varepsilon(\tau) - \overline{\xi}(\tau)|^2 + 2\varepsilon^2(N-m)^2. \qquad (8.35)$$

Thus, by combining the estimates (8.33) and (8.35) we deduce

$$\begin{aligned}
&\mathbb{E}[|\xi_\varepsilon(T)-\xi_\varepsilon(T)|^p]\\
&\leq \mathbb{E}\bigg[\int_t^T p(p-1)C|\xi_\varepsilon(\tau)-\overline{\xi}(\tau)|^{p-2}(2Lip(\sigma)^2|\xi_\varepsilon(\tau)-\overline{\xi}(\tau)|^2+2\varepsilon^2(N-m)^2)d\tau\bigg]\\
&=\mathbb{E}\bigg[\int_t^T 2p(p-1)Lip(\sigma)^2C|\xi_\varepsilon(\tau)-\overline{\xi}(\tau)|^p+2\varepsilon^2(N-m)^2Cp(p-1)|\xi_\varepsilon(\tau)-\overline{\xi}(\tau)|^{p-2}d\tau\bigg].
\end{aligned} \tag{8.36}$$

Set $C_1 := C_1(m,N) = 2(N-m)^2C$, applying Young's inequality with exponents $\alpha = \frac{p}{p-2}$ and $\beta = \frac{p}{2}$ we find

$$\varepsilon^2p(p-1)C_1|\xi_\varepsilon(\tau)-\overline{\xi}(\tau)|^{p-2} \leq \frac{(\varepsilon^2p(p-1)C_1)^{\frac{p}{2}}}{\frac{p}{2}} + \frac{p-2}{p}|\xi_\varepsilon(\tau)-\overline{\xi}(\tau)|^p. \tag{8.37}$$

The estimate (8.37) implies that the inequality in (8.36) can be rewritten as

$$\begin{aligned}
&\mathbb{E}[|\xi_\varepsilon(T)-\xi_\varepsilon(T)|^p]\\
&\leq \mathbb{E}\bigg[\int_t^T \bigg(2Lip(\sigma)^2Cp(p-1)+\frac{p-2}{p}\bigg)C|\xi_\varepsilon(\tau)-\overline{\xi}(\tau)|^p+2\frac{(\varepsilon^2 2p(p-1)C_1)^{\frac{p}{2}}}{p}d\tau\bigg]\\
&=\bigg(2Lip(\sigma)^2Cp(p-1)+\frac{p-2}{p}\bigg)\mathbb{E}\bigg[\int_t^T|\xi_\varepsilon(\tau)-\overline{\xi}(\tau)|^p d\tau\bigg]+2\frac{(\varepsilon^2p(p-1)C_1)^{\frac{p}{2}}}{p}(T-t).
\end{aligned} \tag{8.38}$$

Applying Gronwall's Lemma to (8.38) we obtain

$$\begin{aligned}
\mathbb{E}[|\xi_\varepsilon(T)-\overline{\xi}(T)|^p] &\leq \frac{2(\varepsilon^2p(p-1)C_1)^{\frac{p}{2}}}{p}(T-t)e^{\int_t^T 2Lip(\sigma)^2Cp(p-1)+\frac{p-2}{p}d\tau}\\
&= \frac{(2(\varepsilon^2p(p-1)C_1)^{\frac{p}{2}}}{p}(T-t)e^{2Lip(\sigma)^2Cp(p-1)+\frac{p-2}{p})(T-t)}.
\end{aligned} \tag{8.39}$$

We now recall that $0 \leq T-t \leq T$. Then, considering (8.39), taking the p-root and recalling that the expectation of a constant is the constant itself, we obtain

$$\mathbb{E}[|\xi_\varepsilon(T)-\overline{\xi}(T)|^p]^{\frac{1}{p}} \leq 2^{\frac{1}{p}}\frac{(\varepsilon^2p(p-1)C_1)^{\frac{1}{2}}}{p^{\frac{1}{p}}}e^{(2Lip(\sigma)^2C(p-1)+\frac{p-2}{p^2})T}T^{\frac{1}{p}},$$

Note that $\lim_{p\to\infty}T^{\frac{1}{p}}=1$, $\lim_{p\to\infty}e^{\frac{p-2}{p^2}}=0$, $\lim_{p\to\infty}2^{\frac{1}{p}}=1$ and $\lim_{p\to\infty}p^{\frac{1}{p}}=1$. Thus, taking for example $\varepsilon(p)=\frac{1}{e^p\sqrt{p(p-1)\ln p}}$, the result follows immediately.

□

Remark 8.29. *Theorem 8.24 holds true for every choice of $\varepsilon(p)$ such that*

$$\lim_{p\to\infty} \varepsilon(p)\frac{e^p}{\sqrt{p(p-1)}} = 0.$$

8.5 The liminf estimate: an ongoing project

In this section we sketch briefly the technical difficulties related to obtain the missing estimate, i.e.

$$\liminf_{p\to\infty} V_p^{\varepsilon(p)}(t,x) \geq V(t,x).$$

This estimate is crucial to obtain the convergence result we aimed to. We consider ν_p as the optimal control for the function V_p^ε found in the Chapter 7, i.e.

$$\nu_p = \begin{bmatrix} \sin^2\alpha - \frac{C}{p}\sin 2\alpha + O\left(\frac{1}{p^2}\right) & \frac{1}{2}\sin 2\alpha + \frac{C}{p}\cos 2\alpha + O\left(\frac{1}{p^2}\right) & O\left(\frac{1}{p^2}\right) \\ \frac{1}{2}\sin 2\alpha + \frac{C}{p}\cos 2\alpha + O\left(\frac{1}{p^2}\right) & \cos^2\alpha + \frac{C}{p}\sin 2\alpha + O\left(\frac{1}{p^2}\right) & O\left(\frac{1}{p^2}\right) \\ O\left(\frac{1}{p^2}\right) & O\left(\frac{1}{p^2}\right) & 1 - O\left(\frac{1}{p^2}\right) \end{bmatrix},$$

where $\alpha \in [0,\pi)$ and $C \in \mathbb{R}$ fixed while we define $\overline{\nu}_p$ as an admissible control for horizontal case obtained, roughly speaking, by a "projection" of ν_p, i.e.

$$\overline{\nu}_p = \begin{bmatrix} \sin^2\alpha - \frac{C}{p}\sin 2\alpha & \frac{1}{2}\sin 2\alpha + \frac{C}{p}\cos 2\alpha & 0 \\ \frac{1}{2}\sin 2\alpha + \frac{C}{p}\cos 2\alpha & \cos^2\alpha + \frac{C}{p}\sin 2\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We remark immediately that $\|\overline{\nu}_p - \nu_p\| = O\big(\frac{1}{p^2}\big)$ where the norm $\|\ \|$ is defined as (8.26).

Assumption: In Carnot group it holds true $\|\overline{\nu}_p - \nu_p\| = O\big(\frac{1}{p^2}\big)$. This extends the result obtained in Chapter 7 for the Heisenberg group.

We obtain, by using Lemma 8.28, the Minkwoski's inequality and the assumption that g is globally Lipschitz, that

$$\begin{aligned}
&V_p(t,x) - V_p^\varepsilon(t,x) = \mathbb{E}[|g(\xi^{t,x,\overline{\nu}_p}(T)|^p]^{\frac{1}{p}} - \mathbb{E}[|g(\xi_\varepsilon^{t,x,\nu_p}(T))|^p]^{\frac{1}{p}} \\
&\quad \leq \mathbb{E}[|g(\overline{\xi}^{t,x,\nu_1})(T))|^p]^{\frac{1}{p}} - \mathbb{E}[|g(\xi_\varepsilon^{t,x,\nu_p}(T))|^p]^{\frac{1}{p}} \leq \mathbb{E}[|g(\overline{\xi}^{t,x,\nu_1}(T)) - g(\xi_\varepsilon^{t,x,\nu_p}(T))|^p]^{\frac{1}{p}} \\
&\quad \leq Lip(g)\mathbb{E}[|\overline{\xi}^{t,x,\nu_1}(T) - \xi_\varepsilon^{t,x,\nu_p}(T)|^p]^{\frac{1}{p}}. \qquad (8.40)
\end{aligned}$$

We define the associated control $\nu_1 := \nu_1(\nu_p)$ as in (8.30) (omitting the dependence on p in the notation). For sake of simplicity, we write $\overline{\xi} := \overline{\xi}^{t,x,\nu_1}$ and $\xi_\varepsilon := \xi_\varepsilon^{t,x,\nu_p}$. Applying the Itô's Lemma to the differentiable function $G(x) = |x|^p$ we obtain

$$\begin{aligned}
&\mathbb{E}[|\overline{\xi}(T) - \xi_\varepsilon(T)|^p] \\
&= \mathbb{E}\bigg[\int_t^T \bigg(p|\overline{\xi}(\tau) - \xi_\varepsilon(\tau)|^{p-2}(\overline{\xi}(\tau) - \xi_\varepsilon(\tau))(\overline{\sigma}^T(\overline{\xi}(\tau))\nu_1(\tau) - \sigma_\varepsilon^T(\xi_\varepsilon(\tau))\nu_p(\tau))dB_N(\tau)\bigg)\bigg] \\
&+ \mathbb{E}\bigg[\bigg(\int_t^T p(p-1)|\overline{\xi}(\tau) - \xi_\varepsilon(\tau)|^{p-2} Tr\bigg((\overline{\sigma}^T(\overline{\xi}(\tau))\nu_1(\tau) - \sigma_\varepsilon^T(\xi_\varepsilon(\tau))\nu_p(\tau)) \\
&((\overline{\sigma}^T(\overline{\xi}(\tau))\nu_1(\tau) - \sigma_\varepsilon^T(\xi_\varepsilon(\tau))\nu_p(\tau))^T\bigg)d\tau\bigg]. \qquad (8.41)
\end{aligned}$$

The first term vanishes because the expectation of an Itô integral is always zero. For the second term we recall that it holds true $Tr(SS^T) = \|S\|^2$. Then, summing and subtracting the quantity $\overline{\sigma}^T(\overline{\xi}(\tau))\nu_p(\tau)$ and applying the trivial inequality $(a+b)^2 \leq 2a^2 + 2b^2$, we obtain

$$\begin{aligned}
&\|\overline{\sigma}^T(\overline{\xi}(\tau))(\nu_1(\tau) - \nu_p(\tau)) + (\overline{\sigma}^T(\overline{\xi}(\tau)) - \sigma_\varepsilon^T(\xi_\varepsilon(\tau)))\nu_p(\tau)\|^2 \\
&\quad \leq 2\|\overline{\sigma}^T(\overline{\xi}(\tau))(\nu_1(\tau) - \nu_p(\tau))\|^2 + 2\|(\overline{\sigma}^T(\overline{\xi}(\tau)) - \sigma_\varepsilon^T(\xi_\varepsilon(\tau)))\nu_p(\tau)\|^2. \qquad (8.42)
\end{aligned}$$

Recalling that $\overline{\sigma}$ is globally bounded and applying Lemma 8.26, we get

$$\begin{aligned}
&\|\overline{\sigma}^T(\overline{\xi}(\tau))(\nu_1(\tau) - \nu_p(\tau)) + (\overline{\sigma}^T(\overline{\xi}(\tau)) - \sigma_\varepsilon^T(\xi_\varepsilon(\tau)))\nu_p(\tau)\|^2 \\
&\quad \leq 2\|\nu_1(\tau) - \nu_p(\tau)\|^2 C_\sigma^2 + 2\|(\overline{\sigma}^T(\overline{\xi}(\tau)) \pm (\overline{\sigma}^T(\xi_\varepsilon(\tau)) - \sigma_\varepsilon^T(\xi_\varepsilon(\tau)))\|^2 (N + N^2)^2,
\end{aligned}$$

We define now $M_1 := N + N^2$. Recalling that σ is globally Lipschitz, applying

Lemma 8.23, we get

$$\begin{aligned}&\|\overline{\sigma}^T(\overline{\xi}(\tau))(\nu_1(\tau)-\nu_p(\tau))+(\overline{\sigma}^T(\overline{\xi}(\tau))-\sigma_\varepsilon^T(\xi_\varepsilon(\tau)))\nu_p(\tau)\|^2\\ \leq&2\|\nu_1(\tau)-\nu_p(\tau)\|^2C_\sigma^2+2(2\varepsilon^2(n-m)^2+2Lip(\sigma)|\overline{\xi}(\tau)-\xi_\varepsilon(\tau)|^2)M_1^2.\end{aligned} \tag{8.43}$$

Using (8.43) we obtain that (8.41) can be estimated as

$$\begin{aligned}&\mathbb{E}[|\overline{\xi}(T)-\xi_\varepsilon(T)|^p]\\ \leq&\mathbb{E}\bigg[\int_t^T p(p-1)|\overline{\xi}(\tau)-\xi_\varepsilon(\tau)|^{p-2}(2\|\nu_1-\nu_p\|^2C_\sigma^2\\ &+2(2\varepsilon^2(N-m)^2+2Lip(\sigma)^2|\overline{\xi}(\tau)-\xi_\varepsilon(\tau)|^2)M_1^2d\tau\bigg]\\ =&\,\mathbb{E}\bigg[\int_t^T 4p(p-1)Lip(\sigma)^2M_1^2|\overline{\xi}(\tau)-\xi_\varepsilon(\tau)|^p\\ &+(2C_\sigma^2\|\nu_1-\nu_p\|^2+4\varepsilon^2(N-m)^2)p(p-1)|\overline{\xi}(\tau)-\xi_\varepsilon(\tau)|^{p-2}\bigg].\end{aligned} \tag{8.44}$$

Applying Young's inequality with exponents $\alpha=\frac{p}{2}$ and $\beta=\frac{p}{p-2}$, we obtain

$$\begin{aligned}&(2C_\sigma^2\|\nu_1-\nu_p\|^2+4\varepsilon^2(N-m)^2p(p-1))|\overline{\xi}(\tau)-\xi_\varepsilon(\tau)|^{p-2}\\ &\leq\frac{((2C_\sigma^2\|\nu_1-\nu_p\|^2+4\varepsilon^2(N-m)^2)p(p-1))^{\frac{p}{2}}}{\frac{p}{2}}+\frac{p-2}{p}|\overline{\xi}(\tau)-\xi_\varepsilon(\tau)|^p\end{aligned}$$

and so, we can rewrite (8.44) as

$$\begin{aligned}&\mathbb{E}[|\overline{\xi}(T)-\xi_\varepsilon(T)|^p]\\ &\leq\mathbb{E}\bigg[\int_t^T\left(4p(p-1)Lip(\sigma)^2M_1^2+\frac{p-2}{p}\right)|\overline{\xi}(\tau)-\xi_\varepsilon(\tau)|^p\\ &+\frac{(2C_\sigma^2\|\nu_1-\nu_p\|^2+4\varepsilon^2(N-m)^2)^{\frac{p}{2}}p(p-1)}{\frac{p}{2}}d\tau\bigg]\\ &=\left(4p(p-1)Lip(\sigma)^2M_1+\frac{p-2}{p}\right)\int_t^T\mathbb{E}\|\overline{\xi}(\tau)-\xi_\varepsilon(\tau)|^p]d\tau\\ &+\frac{((2C_\sigma^2\|\nu_1-\nu_p\|^2+4\varepsilon^2(N-m)^2)p(p-1))^{\frac{p}{2}}}{\frac{p}{2}}(T-t).\end{aligned} \tag{8.45}$$

We have assumed that the difference between controls is independent from τ. Applying Gronwall's Lemma to (8.45) and recalling that the expectation of a

constant is the constant itself, we have

$$\begin{aligned}
&\mathbb{E}[|\overline{\xi}(T)-\xi_\varepsilon(T)|^p] \\
&\quad\le \frac{(2C_\sigma^2\|\nu_1-\nu_p\|^2+4\varepsilon^2(N-m)^2p(p-1))^{\frac{p}{2}}}{\frac{p}{2}}(T-t)e^{\int_t^T 4p(p-1)Lip(\sigma)^2M_1^2+\frac{p-2}{p}d\tau} \\
&\quad\le \frac{(2C_\sigma^2\|\nu_1-\nu_p\|^2+4\varepsilon^2(N-m)^2p(p-1))^{\frac{p}{2}}}{\frac{p}{2}}(T-t)e^{\left(4p(p-1)Lip(\sigma)^2M_1^2+\frac{p-2}{p}\right)(T-t)}.
\end{aligned}$$

Taking the p-root of the estimate above, we conclude

$$\begin{aligned}
&\mathbb{E}[|\overline{\xi}(T)-\xi_\varepsilon(T)|^p]^{\frac{1}{p}} \\
&\quad\le \frac{(2C_\sigma^2\|\nu_1-\nu_p\|^2+4\varepsilon^2(N-m)^2p(p-1))^{\frac{1}{2}}}{\left(\frac{p}{2}\right)^{\frac{1}{p}}}(T-t)e^{\left((p-1)Lip(\sigma)^2M_1^2+\frac{p-2}{p^2}\right)(T-t)}.
\end{aligned} \tag{8.46}$$

So, $\|\nu_1-\nu_p\|=O\left(\frac{1}{p^2}\right)$ seems to do not converge to 0 fast enough as $p\to\infty$ to balance the exponential term in (8.46).

To control the term $\sim e^{Cp}$ we need to find a better estimate for ν_p so that $\|\nu_1-\nu_p\|=O\left(\frac{1}{e^p p^2}\right)$. This needs to improve the results in Chapter 7 by considering a Taylor's expansion in ε and p simultaneously, which should allow us more flexibility in the choice of $\varepsilon(p)$. This will be the object of future research.

Conclusion

Here we state shortly the results obtained in the thesis and the future plan of research. To sum up, the main results of this thesis are:

- In Chapter 5 we generalize the results in [22], obtaining the stochastic representation of the solution of the approximated Riemannian mean curvature flow.

- In Chapter 6 we found the asymptomatical behaviour, for large p, for the p-optimal controls for the stochastic dynamics associated to the horizontal mean curvature flow, in the 1-dimensional Heisenberg group. Furthermore we develop some numerical simulations, in order to understand better the behaviour of the optimal controls. This result was open even in the Euclidean case

- In Chapter 7 we found the asymptomatical behaviour, for large p, for the p-optimal controls for the stochastic dynamics associated to the approximated Riemannian mean curvature flow for the 1-dimensional Heisenberg group.

- In Chapter 8 we prove the following lim sup inequality:

$$\limsup_{p\to\infty} V_p^{\varepsilon(p)}(t,x) \leq V(t,x),$$

by using real analysis tools (see Appendix for further details) for any choice of $\varepsilon(p)$ such that $\lim_{p\to\infty} \frac{\varepsilon(p)e^p}{\sqrt{p(p-1)}} = 0$. In the last section of the

chapter we explain some of the technical difficulties we have to overcome to obtain the corresponding lim inf inequality:

$$\liminf_{p\to\infty} V_p^{\varepsilon(p)}(t,x) \geq V(t,x).$$

Roughly speaking the asymptotic behaviour for V_p^ε found in Chapter 7 seems to give us not enough information to control the constants coming from the application of the Gronwall's Lemma. Therefore we plan to improve the result in Chapter 7 by using an asymptotic expansion in ε and p, simultaneously in order to find a better $\varepsilon(p)$ to conclude the minimizing liminf estimate.

Appendix

In this appendix we will recall some standard results about classical real analysis, in particular about the L^p spaces and some inequalities which are often used in this thesis (for further details, see [47]). We start with a very well known inequality which has been used to find the estimates for the convergence results.

Lemma 8.30 (Gronwall's Lemma)**.** *Let $I = [a, b] \subset \mathbb{R}$ where $a < b$ and let α, β and u be real-valued functions defined in I. Assume that β and u are continuous and that the negative part of α is integrable on every closed and bounded subinterval of I*

- *If β is non negative and if u satisfies the integral inequality*

$$u(t) \leq \alpha(t) + \int_a^t \beta(s)u(s)ds, \quad \forall t \in I,$$

then

$$u(t) \leq \alpha(t) + \int_a^t \alpha(s)\beta(s)exp\left(\int_s^t \beta(r)dr\right).$$

- *If, in addition, the function α is a non decreasing function then*

$$u(t) \leq \alpha(t)exp\left(\int_a^t \beta(s)ds\right).$$

Next some basic definitions and properties of measure spaces and L^p spaces.

Definition 8.31. *Let $(\Omega, \mathcal{B}, \mu)$ be a measure space and $f : \Omega \to \mathbb{R}$ be a measurable function.*

- *If* $0 < p < \infty$ *we define the* L^p norm *of the function* f *as*

$$\|f\|_p = \left(\int_\Omega |f|^p \right)^{\frac{1}{p}}.$$

- *If* $p = \infty$ *we define the* L^∞ *norm as*

$$\|f\|_\infty = \inf\{C \geq 0 | |f(x)| \leq C \quad \textit{almost everywhere}\}.$$

Theorem 8.32 (Jensen's inequality)**.** *Let* $(\Omega, \mathcal{B}, \mu)$ *be a meausure space such that* $\mu(\Omega) = 1$*, if* g *is an integrable function on* Ω *and* ϕ *a convex function, then it holds*

$$\phi\left(\int_\Omega g \, d\mu \right) \leq \int_\Omega \phi \circ g \, d\mu.$$

Theorem 8.33 (Young inequality)**.** *Let* $a, b, \in \mathbb{R}$ *and* $p, q > 1$ *such that* $\frac{1}{p} + \frac{1}{q} = 1$

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Theorem 8.34 (Hölder inequality)**.** *Let* $(X, \mathcal{B}, \mu)$ *be a meausure space and* $1 \leq p, q \leq \infty$ *such that*

$$\frac{1}{p} + \frac{1}{q} = 1$$

and let $f \in L^p(X)$ *and* $g \in L^q(X)$ *then we have that* $fg \in L^1(X)$ *and*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Theorem 8.35 (Minkwoski inequality)**.** *Let* $f, g \in L^p$*, then* $f + g \in L^p$ *and*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

The Hölder inequality gives us an important corollary about the spaces of finite measures.

Corollary 8.36. *Let* $(\Omega, \mathcal{B}, \mu)$ *be a finite measure space. Then, for every* $1 \leq p < q \leq \infty$ *it holds true*

$$L^q(\Omega, \mathcal{B}, \mu) \subset L^p(\Omega, \mathcal{B}, \mu).$$

Theorem 8.37. *For all $1 \leq p \leq \infty$ and positive measure μ we have that $L^p(\Omega, \mathcal{B}, \mu)$ is a complete metric space.*

Theorem 8.38. *If $1 \leq p \leq \infty$ and $\{f_n\}$ is a Cauchy sequence in $L^p(\Omega, \mathcal{B}, \mu)$ with limit to f, then there exists a subsequence which convergence pointwise and almost everywhere to f.*

To conclude it is possible to prove that the space $L^2(\Omega, \mathcal{B}, \mu)$ with the inner product

$$(f, g) \to \int_\Omega fg \, d\mu$$

is an Hilbert space (i.e. a space with the norm inducted by the inner product which is complete).

Acknowledgement

I would like to say thanks to my supervisors F. Dragoni and N. Dirr who taught me how to do research and guided me during these years. I would thank also my mother Oriana, my father Cesare, my sister Ramona, my grandfathers and my grandmothers and the rest of my family, colleagues of research and my friends.

Thank you all!

Bibliography

[1] J.F. Adams. *Lectures on Lie Groups.* University of Chicago Press, 1982

[2] S. Angenent. *Shrinking Doughnuts*. Nonlinear diffusion equations and their equilibrium states. 1989

[3] D. Azagra, J. Ferrera, B. Sanz. *Viscosity solutions to second order partial differential equations on Riemannian manifolds.* Journal of Differential Equations, Vol 245, Issue 2, 2008

[4] A. Bonfiglioli, E. Lanconelli and F. Uguzzoni. *Stratified Lie Group and Potential Theory for their Sub-Laplacian.* Springer, 2007.

[5] L. Breiman. *Probability.* Classics in Applied Mathematics SIAM, 1992

[6] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2011

[7] R. Buckdahn, P. Cardaliaguet and M. Quincampoix. *A representation formula for the mean curvature motion.* SIAM Journal of Mathematical Analysis, vol 33, Issue 4, pp. 827–846, 2001

[8] L. Capogna and G. Citti. *Generalized mean curvature flow in Carnot groups*. Comm. Partial Differential Equations, vol. 34, Issue 8, pp. 937-956. 2009

[9] L. Capogna, D. Danielli, S.D. Pauls, J,Tyson. *An Introduction to the Heisenberg Group and the Sub-Riemannian Isoperimetric Problem.* Birkhauser,2007

[10] R. Carmona. *Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications*, SIAM Financial Mathematics, 2016

[11] Y. Chen, Y. Giga, S. Goto. *Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations* J. Differential Geom. Vol 33, Issue 3, 749-786, 1991.

[12] W.L. Chow. *Über systeme von linearen partiellen differentialgleichungen erster ordnung*. Math Ann. vol 117, 1939

[13] G. Citti, B. Franceschiello, G. Sanguinetti and A. Sarti. *Sub-Riemannian mean curvature flow for image processing*. SIAM Journal Imaging Sci., vol 9, Issue 1, pp. 212–237, 2016

[14] G. Citti, A. Sarti. *A Cortical Based Model of Perceptual Completion in the Roto-Translation Space.* Journal of Mathematical Imaging and Vision, vol 24, pp. 307–326, 2006

[15] G. Citti, A. Sarti. *Subjective surfaces and Riemannian mean curvature flow of graphs.* Acta Math Univ. Comenianae, Vol 70, p 85-103, 2001

[16] G. Citti, B. Franceschiello, A. Mashtakov, A. Sarti *Geometrical optical illusion via sub-Riemannian geodesics in the roto-translation group*, Differential Geometry and its Applications, Vol 65, pp. 55-77, 2019

[17] M.G. Crandall, L.C. Evans, P.L. Lions. *Some properties of viscosity solutions of Hamilton-Jacobi equations*. Transactions of the American Mathematical Society, 1984

[18] M.G. Crandall, H. Ishii, P.L. Lions. *Uniqueness of viscosity solutions of Hamilton-Jacobi equations revisited*. J.Math.Soc. Japan,1987

[19] M.G. Crandall, P.L. Lions. *Viscosity solutions of Hamiltonian-Jacobi equations*. Trans. Am. Math. Soc. vol 277, 1983

[20] M.G. Crandall, P.L. Lions: *On existence and uniqueness of solutions of Hamilton-Jacobi equation*. Nonlinear Analysis Theory Appl. 10, 1986

[21] N. Dirr, F. Dragoni, R. Grande. *Asymptotics for optimal controls for horizontal mean curvature flow.* Nonlinear Analysis (accepted paper) 2020

[22] N. Dirr, F. Dragoni and M. von Renesse. *Evolution by mean curvature flow in sub-Riemannian geometries*. Communications on Pure and Applied Mathematics, vol 9, Issue 2, pp.307-326, 2010

[23] N.Dirr, F.Dragoni, P.Mannucci and C.Marchi, *Stochastic homogenization for functionals with anisotropic rescaling and noncoercive Hamilton–Jacobi equations.* SIAM Journal on Mathematical Analysis 50 (5) , pp. 5198-5242, 2018

[24] F. Dragoni. *Carnot-Carathéodory metrics and viscosity solution*. PhD thesis, 2005

[25] F. Dragoni, R. Grande. *Optimal control approach to Riemannian approximation of sub-Riemannian mean curvature flow* preprint, 2021

[26] F. Dragoni, R. Grande *Riemannian approximation for horizontal mean curvature flow: a stochastic approach*. preprint, 2021

[27] M.P. Do Carmo. *Riemannian Geometry.* Birkhauser, 1992

[28] K. Ecker. *Regularity Theory for Mean Curvature Flow*. Birkhauser, 2004

[29] K. Erdmann, M.J. Wildon. *Introduction to Lie Algebras.* Springer, 2006

[30] L.C. Evans. *Partial Differential Equations.* Am. Math. Soc. Providence, 1988

[31] L.C. Evans and J. Spruck. *Motion of level sets by mean curvature*. Journal of Differential Geometry, vol 33, Issue 3 pp.635-681, 1991

[32] F. Ferrari, Q. Liu, J. Manfredi. *On the horizontal mean curvature flow for axisymmetric surfaces in the Heisenberg group*, Communication in contemporary mathematics, Vol. 16, 2014

[33] W.H. Fleming and H.M. Soner. *C*ontrolled Markov Processes and Viscosity Solutions. Springer, 2006

[34] Y. Giga. *Surface Evolution Equation, a level set approach*. Birkhauser Verlag. 2006

[35] R. Grande. *A stochastic representation for the solution of approximated mean curvature flow*. preprint, 2021

[36] M. Gromov. *Metric Structures for Riemannian and non-Riemannian spaces.* Birkhauser, 1999.

[37] B.C. Hall. *Lie groups, Lie algebras and Representations: an elementary introduction.* Springer, 2003

[38] L. Hörmander. *Hypoelliptic second order differential operator*. Acta Math. 119, 1967

[39] J. Jacod, P. Protter. *Probability essentials.* Springer-Verlag, 2003

[40] T. Marosevic. *The Hausdorff distance between some sets of points.* Mathematical communication Vol 23, Issue 2, pp. 247–257, 2018

[41] P. Morters, Y. Peres. *Brownian motion.* Cambridge University Press, 2010

[42] R. Montgomery. *A Tour of Sub-Riemannian Geometries, their Geodesics and Applications.* Math. Surv. and Monographs 91 AMS, 2000

[43] R.Monti. *Distances, boundaries and surface measures in Carnot-Carathéodory spaces*

[44] R. Monti. *The regularity problem for sub-Riemannian geodesics* . Springer INDAM Series, 2013

[45] B. Oksendal. *Stochastic differential equations: An introduction with applications.* Springer, 2010

[46] O. Perron. *Eine neue Behandlung der Randwetufgabe fur* $\Delta u = 0$. Math Z. vol 18, 1923

[47] W. Rudin. *Real and complex analysis*, third edition. McGraw-Hill International Edition.

[48] H.M. Soner and N. Touzi. *A stochastic representation for the level set equations.* Comm. Part.Diff. Eq., vol 27, Issue 9, pp. 2031–205, 2002

[49] H.M. Soner and N. Touzi. *A stochastic representation for mean curvature type geometric flows*. Ann. Probability, Vol 31, pag 1145-1165, 2003

[50] P.E. Souganidis, N.K. Yip. *Uniqueness of motion by mean curvature flow perturbed by stochastic noise.* Annales de l'I.H.P. Analyse non lineaire, Vol 21, pp. 1-23, 2004

[51] D. W. Stroock, S. R. Srinivasa Varadhan, *Multidimensional Diffusion Processes.* Springer, Berlin. ISBN 978-3-662-22201-0. 1997

[52] N. Touzi. *Stochastic Control Problems, Viscosity Solutions, and Application to Finance.* Scuola Normale Superiore di Pisa. Quaderni, Pisa, 2004