SCUOLA NORMALE SUPERIORE DI PISA
Perfezionamento in Matematica
XVIII ciclo, anni 2003-2005

PhD Thesis

Federica Dragoni

CARNOT-CARATHÉODORY METRICS AND VISCOSITY SOLUTIONS

Advisor: Prof. Italo Capuzzo Dolcetta

Contents

Introduction.

Sub-Riemannian geometries have many interesting applications in very different settings, as optimal control theory ([17, 26, 27]), calculus of variations ([2, 4]) or stochastic differential equations ([16]). Many physical phenomena seem to induce in a natural way an associated sub-Riemannian structure, for example, one can think of Berry's phase problem, a swimming microorganism (studied in [75]), the optimal control in laser-induced population transfer (see [20]) and the perceptual completion in the visual cortex ([35]). Sub-Riemannian geometries (known also as Carnot-Carathéodory spaces) arise whenever there are privileged and prohibited paths. In fact, the main difference between these geometries and the Riemannian geometries is the need to move along some prescribed vector fields. A Sub-Riemannian metric is indeed a Riemannian metric defined only on a subbundle of the tangent bundle to the manifold. More precisely, let $\mathcal{X}=\{X_1,...,X_m\}$ be a family of vector fields, defined on a n-dimensional manifold M (with in general $m\leq n$), a sub-Riemannian metric is a Riemannian metric $\langle\ ,\ \rangle$ defined on the fibers of $\mathcal{H}:=\mathrm{Span}(\mathcal{X})\subset TM$. A subbundle $\mathcal{H}$ is usually called distribution.
For sake of simplicity, from now on we will always assume that $M=\mathbb{R}^n$. This in particular implies that the tanget space at any point x is equal to $\mathbb{R}^n$ and the tangent bundle is isomorphic to $\mathbb{R}^{2n}$. An admissible (or $\mathcal{X}$-horizontal) curve is any absolutely continuous curve $\gamma:[0,T]\to\mathbb{R}^n$, such that

$$\dot\gamma(t)\in\mathrm{Span}(X_1(\gamma(t)),...,X_m(\gamma(t))),\quad \text{a.e. } t\in[0,T].$$

Since the Riemannian metric $\langle\ ,\ \rangle$ is defined along the fibers of $\mathcal{H}$, for the horizontal curves and, only for the horizontal curves, we can introduce a length-functional as follows:

$$l(\gamma) = \int_0^T \left\langle \dot{\gamma}(t), \dot{\gamma}(t) \right\rangle^{\frac{1}{2}} dt.$$

A sub-Riemannian (or Carnot-Carathéodory) distance on $\mathbb{R}^n$ can be defined, for any $x, y \in \mathbb{R}^n$, as

$$d(x, y) = \inf\{l(\gamma) \mid \gamma \, \text{admissible curve joining} \, x \, \text{to} \, y\}. \tag{1}$$

It is obvious that, whenever there are not admissible curves joining x to y, then $d(x, y) = +\infty$. For this purpose the Hörmander condition is introduced. In fact, in Carnot-Carathéodory spaces satisfying the Hörmander condition, it is always possible to join two given points by an admissible curve (Chow's Theorem, [17, 75]). Then the associated Carnot-Carathéodory distance is finite. The Hörmander condition (known also as bracket generating condition) is satisfied if the Lie algebra associated to the distribution $\mathcal{H} = \mathrm{Span}(\mathcal{X})$ spans the whole tangent space at any point of the manifold (that in this particular case is at any point equal to $\mathbb{R}^n$). We recall that the bracket between two vector fields X and Y is the vector field defined as $[X, Y] = XY - YX$, acting by derivation on smooth real functions. The Lie algebra $\mathcal{L}(\mathcal{X})$ associated to $\mathcal{X} = \{X_1, ..., X_m\}$ is the set of all the brackets between elements of $\mathcal{X}$, so the Hörmander condition holds, if and only if,

$$\mathrm{Span}\big(\mathcal{L}(X_1(x), ..., X_m(x))\big) = T_x M = \mathbb{R}^n, \quad \text{for any } x \in \mathbb{R}^n.$$

The first chapter is dedicated to the study of sub-Riemannian geometries and topological and metric implications of the Hörmander condition.

In the second chapter, we are interested in solving some first-order nonlinear partial differential equations (PDEs) related to the Hörmander condition. Therefore we will introduce the theory of viscosity solutions for continuous and discontinuous functions. Later we concentrate on the two particular nonlinear PDEs: The eikonal equation and an evolution Hamilton-Jacobi equation.

We first solve the generalized eikonal equation:

$$H_0(x, Du) = 1, \tag{2}$$

with vanishing condition at some fixed point $y \in \mathbb{R}^n$ and where $H_0(x,p)$ is the geometrical Hamiltonian defined by

$$H_0(x,p) = |\sigma(x)p|, \tag{3}$$

with $\sigma(x)$ Hörmander-matrix (i.e. a $m \times n$ real-valued matrix with smooth coefficiente and such that its rows satisfy the Hörmander condition).
Later we consider an evolution Hamilton-Jacobi equation of the form:

$$u_t + \Phi(H_0(x,Du)) = 0 \tag{4}$$

where Φ is a suitable positive and convex function and $H_0(x,p)$ satisfies the structural assumpiton (2). The main model for the PDE (4) is

$$u_t + \frac{1}{\alpha}|\sigma(x)Du|^\alpha = 0, \quad \text{with } \alpha > 1. \tag{5}$$

We solve both the previous PDEs by using the Hörmander condition and suitable representative formulas. In order to solve the eikonal equation, fixed the point $y \in \mathbb{R}^n$, we define the minimal-time function

$$d(x,y) = \inf_{X(\cdot)\in F_{x,y}} T(X(\cdot)), \tag{6}$$

where $F_{x,y}$ is the set of all the trajectories joining x to y in a time $T(X(\cdot))$ which are solutions of the differential inclusion

$$\dot{X}(t) \in \partial H_0(X(t),0) = \sigma^T(X(t))\overline{B_1(0)},$$

with $\overline{B_1(0)}$ unit Euclidean ball in $\mathbb{R}^m$ centered at the origin.
It is possible to show under very weak assumptions that a minimal-time distance is a generalized distance (i.e. a distance not always symmetric) solving a Dynamical Programming Principle:

$$d(x,y) = \inf_{X(\cdot)\in F_{x,y}} [t + d(X(t),y)], \quad \text{for any } 0 < t < d(x,y).$$

For our particular Hamiltonian in (3), the minimal-time distance turns out to be equivalent to the Carnot-Carathéodory distance associated to the matrix $\sigma(x)$. By using the Dynamical Programming Principle, we prove that $u(x) = d(x,y)$ solves in the viscosity sense the corresponding eikonal equations on

$\mathbb{R}^n \setminus \{y\}$. By a sub-Riemanninan generalization of the Rademacher's Theorem (see [75, 77, 80]), the viscosity result implies that the Carnot-Carathéodory distance is a almost everywhere solution, too (at least in Carnot groups). To solve the eikonal equation is the key-point in order to solve the Cauchy problem for Eq. (4) with lower semicontinuous initial data $g(x)$. To get the existence of a viscosity solution for this class of PDEs, we use a suitable representative formula: the metric Hopf-Lax formula given by

$$u(t,x) = \inf_{y \in \mathbb{R}^n} \left[g(y) + t\Phi^* \left(\frac{d(x,y)}{t} \right) \right], \tag{7}$$

where Φ^* is the Legendre-Fenchel transform:

$$\Phi^*(t) = \sup_{s \geq 0} \{st - \Phi(s)\}, \quad \text{for any } t > 0.$$

When $H_0(x, Du) = |Du|$, (7) is exactly the classic Hopf-Lax formula studied in [7, 8, 47] in the continuous viscosity setting, and in [3] in the semicontinuous case. The metric Hopf-Lax formula was first introduced by Lu-Manfredi-Stroffolini [73] in the particular case of the Heisenberg group and then generalized to general metric spaces by Capuzzo Dolcetta and Ishii in [27] (see also [26, 25]).
We investigate some properties of function (7) and show that, whenever $g : \mathbb{R}^n \to \mathbb{R}$ is lower semicontinuous, $u(t,x)$ lower converges to g as $t \to 0^+$, and it is locally Lipschitz continuous in x w.r.t the metric $d(x,y)$. Moreover, if the function g is bounded, the infimum in (7) is a minimum and the metric Hopf-Lax function is locally Lipschitz continuous in t (in the Euclidean sense) and non decreasing in t. The main result proved in the second chapter is that $u(t,x)$ given in (7) is a viscosity solution of the Cauchy problem for Eq. (4) with lower semicontinuous initial data $g(x)$.

The third chapter is devoted to investigating the metric inf-convolution. The metric inf-convolution of some function $g(x)$ is defined as the Hopf-Lax function whenever $\Phi(t) = \frac{1}{2}t^2$, i.e.

$$g_t(x) = \inf_{y \in \mathbb{R}^n} \left[g(y) + \frac{d(x,y)^2}{2t} \right]. \tag{8}$$

We show that functions (8) give a monotonously Lipschitz continuous approximation of the function $g(x)$, as $t \to 0^+$ (assuming that $g(x)$ is lower semicontinuous and bounded in $\mathbb{R}^n$), similarly to the known Euclidean case (see [7, 24]). In particular we are interested in the limiting behavior of the logarithm-transform for the solutions of some subelliptic heat problems:

$$\begin{cases} w_t^\varepsilon - \varepsilon \displaystyle\sum_{i,j=1}^n a_{i,j}(x) \frac{\partial^2 w^\varepsilon}{\partial x_i \partial x_j} = 0, & x \in \mathbb{R}^n, t > 0, \\ \qquad\qquad w^\varepsilon(0,x) = \mathrm{e}^{-\frac{g(x)}{2\varepsilon}}, & x \in \mathbb{R}^n, \end{cases}$$

where $A(x) = \big(a_{i,j}(x)\big)_{i,j=1}^n = \sigma^t(x)\sigma(x)$ and $\sigma(x)$ is a $m \times n$-Hörmander matrix (with $m \leq n$). It is well-known that, if $\sigma(x)$ is a Hörmander-matrix, then the second-order differential operator $L\,u = \sum_{i,j=1}^n a_{i,j}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}$ is hypoelliptic, which means that, whenever f is smooth, the solutions of $L\,u = f$ are smooth, too. Let $d(x,y)$ be the Carnot-Carathéodory distance associated to the Hörmander-matrix $\sigma(x)$, then the logarithm-transform of the solutions w^ε converge to the Carnot-Carathéodory inf-convolution defined by (8). More precisely, we show that

$$\lim_{\varepsilon \to 0^+} -2\varepsilon \log w^\varepsilon(t,x) = g_t(x),$$

for any $g : \mathbb{R}^n \to \mathbb{R}$ bounded and continuous. In order to prove the previous limit, we use the integral representation (by heat kernel) of the solutions w^ε and then we apply the Large Deviation Principle ([94]). The difficulty is to verify the applicability of the Large Deviation Principle in the hypoelliptic case. In particular, we need to generalize the result proved by Varadhan in the uniformly elliptic setting in [93] to the subelliptic case. We give a new proof which uses methods of measure theory and covers any known results extending them up to the Hörmander-case.

For the metric Hopf-Lax function and the metric inf-convolution, many problems are still open. For example it would be very useful to find some horizontal $C^{1,1}$-approximation for continuous (or even semicontinuous) functions, by using both the metric inf-convolutions and the corresponding metric sup-convolutions, as it happens in the Euclidean case (see e.g. [24]).

Unfortunatelly nothing is at the moment known about the semiconcavity or/and the semiconvexity for the metric inf/sup-convolutions functions. It is even not known is the Carnot-Carathéodry distance or a power of the distance could be semicovex or semiconcave in some suitable sense.
To find an horizontal $C^{1,1}$-regularization by metric inf/sup-convolutions would lead to many interesting applications in the study of uniqueness and regularity for nonlinear first-order and second-order PDEs related to the Hörmander condition.

Acknowledgements. I want to express my gratitude towards my advisor Prof. Italo Capuzzo Dolcetta and to Prof. Juan Manfredi who helped to correct the first version of this thesis and suggested improvements.
I also like to thank Prof. Luigi Ambrosio for the many useful suggestions, in particular concerning the proof of the Large Deviation Principle.

Chapter 1

Sub-Riemannian geometries.

In this first chapter we introduce and study a particular kind of degenerate Riemannian geometries: the so called sub-Riemannian geometries.
The main characteristic of these geometries is that there are non-admissible curves. Roughly speaking, one can think of a sub-Riemannian geometry as a Riemannian metric on a manifold, with some constraints on the direction of the motion. In fact, only the curves whose velocity belongs to a some given subbundle of the tangent bundle are admissible.

1.1 Basic definitions and main properties.

1.1.1 Historical introduction and Dido's problem.

To give a clear and complete historical overview of sub-Riemannian geometries is impossible. In fact, these geometries have been developed in the last hundred years in many different setting and by using various names. Hence it is very hard to go back its origin to some particular work or author: we just quote some of the most significant steps. A key result for the sub-Riemannian geometries is given by the Chow's Theorem, proved in the end of thirties, independently by Chow's [30] and by Rashevskii [82]. This result ensures that, under the bracket generating condition, it is always possible to join any pair of points by a horizontal curve. So these two authors can be considered as the "fathers" of the sub-Riemannian theory.
Nevertheless, Carathéodory have already got a version of Chow-Rashevskii's

result in 1909 for corank-one distributions (see [28]).
Indeed Carathéodory proved the reverse implication: the author showed that in a connected manifold endowed with an analytic corank-one distribution if there exist two points that cannot be connected by a horizontal curve, then the distribution is integrable. We recall that an integrable distribution is the opposite case to the bracket generating condition. In fact, a distribution is bracket generating if and only if the associated Lie algebra spans all the tangent space, at any point. A distribution is instead integrable if and only if at any point there exists a hypersurface tangent everywhere to the distribution. Therefore the distributions satisfying the bracket generating condition are called completely non-integrable.
Carathéodory's result is considered the first theorem in the sub-Riemannian theory. Carathéodory applied this theorem to Carnot's work about thermodynamics, getting the existence of integrating factors for the Pfaffian equation. Many years after, Gromov and others started to use the name of Carnot-Carathéodory metrics to indicate the sub-Riemannian case. One can also find Sub-Riemannian metrics with the name of singular Riemannian metrics (see, for example, [21, 53]) or nonholonomic metrics ([95]).
We want also to recall that sub-Riemannian theory is linked to the hypoelliptic PDEs theory. The origin of this theory can be traced back to a work of Hörmander published in 1967 ([55]) and so usually the bracket generating condition is called the Hörmander condition. Nevertheless in the seventies hypoellipticity has been used in many works of Stein ([88]) and others, without giving it a specific name.

The most famous example of a sub-Riemannian geometry is the Heisenberg group. Next we present a very famous problem for the calculus of variation: the isoperimetric problem, known also ad Dido's problem. This minimizing problem in the plane can be in fact used to introduce the 1-dimensional Heisenberg group.

First let us tell something about this Phoenician myth of the foundation of Carthage (in [18], you can find one of the most complete treatment of this myth). The myth of Dido has been made famous by Virgil in the Aeneid

(IV) but it is much more ancient: in fact the first reference to this myth can be found in the literary work of Timeo di Tauromenio (*Storie*, IV-V century B.C.).
Elissa (that is the Phoenician name of Dido) was the daughter of the Phoenician king Tiro. Her brother Pygmalion, after the death of their father, killed her husband who was a rich and powerful priest of the God Melkart. So she decided to leave with some followers and docked at the African North coast. There she bought from Jarbas, the king of Messitania, as much land as could be contained by an ox-skin. Dido cut the ox-skin in many thin strips and then she stringed them together in order to get a very long strip. Using this strip and the African coast she bounded her future kingdom: Carthage (i.e. Qart Hashdat that, in Phoenician, means " new city"). Dido's great idea consists in understanding that the biggest area can be obtained using an arc of a circle.

Dido's problem is the first example of a minimizing problem of calculus of variations. Its dual formulation is the well-known isoperimetric problem. Next we show that the isoperimetric problem in the plane can be easily solved by the introduction of the (1-dimensional) Heisenberg group.
The isoperimetric problem can be easily rewritten in the following way.
Let $\gamma : [0,T] \to \mathbb{R}^2$ be a curve on the plain, that for sake of simplicity we assume smooth closed and bounded, which we write as $\gamma(t) = (x(t), y(t))$. In this simple case, the "perimeter" of the set delimited by the curve γ is the Euclidean length of the curve, which is

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

with $\|\dot{\gamma}(t)\| = \sqrt{\dot{x}^2(t) + \dot{y}^2(t)}$.
To solve the isoperemetric problem means to minimize the previous length-function with the constraint of a constant area.
Therefore, we fix a constant $C > 0$: the area of the domain in the plane, delimited by the curve γ is given by the Stokes' formula:

$$(A) \qquad \frac{1}{2}\int_0^T (x(t)\dot{y}(t) - y(t)\dot{x}(t))dt = C.$$

The isoperemetric problem in the plane can be expressed by the minimization problem:

$$\min \{l(\gamma) \mid \gamma \text{ closed rectifiable curve, satisfying } (A)\} \tag{1.1}$$

The idea is now to define a "third dimension" z by lifting the constraint of the constant area. This leads to a new constraint:

$$\dot{z} = \frac{1}{2}(-y\dot{x} + x\dot{y}). \tag{1.2}$$

Note that the third dimension z is not a free dimension (i.e. it does not lead to an independent variable) but it is related to x and y by Eq. (1.2). In other words this means that only the curves whose derivatives satisfy (1.2) will be considered. Any curve $\gamma : [0,T] \to \mathbb{R}^3$ satisfying (1.2) is called admissible (or horizontal).
As we will see better in Sec.1.3, the kernel of the 1-form $\eta(x,y,z) := dz - \frac{1}{2}(xdy - ydx)$ gives rise to a sub-Riemannian geometry called the 1-dimensional Heisenberg group.
We know, as Dido knew almost three-thousand years ago, that the solutions of the isoperimetric problem in the plane are the arcs of circle. Using the Heisenberg group we can easily verify this result (see Remark 1.117).

Proposition 1.1. *The solutions of the isoperimetric problem* (1.1) *are the projections on the two first-components of the geodesics (that are the horizontal curves with minimal length) of the 1-dimensional Heisenberg group.*

By looking at this example, we can sum up the following general idea: the original constrained minimization problem can be written as a minimization problem without any constraint but living in a higher-dimension space, where constraints on the geometry have been introduced.
This is one of the most useful application of the theory of sub-Riemannian geometries to the calculus of variations.

More information on the links between the Dido-isoperimetric problem and the Heisenberg group can be found in [2] or also in [75], Sections I.1.1-I.1.3.

1.1.2 Some new development: the visual cortex.

Recently Citti and Sarti introduced a very interesting new application (see [33, 31, 32, 34, 35]). In particular by using of a sub-Riemannian geometry, the so called roto-traslation geometry, they got important mathematical as well as numerical results in the study of the image completion.
Let us briefly introduce the phenomenon: look at the picture in Fig. 1.1. It is possible to discern two different fishes even though the two fishes are not fully portrayed. This happens since the human visual cortex completes automatically the internal objects and in this way new contours arise (called *apparent* or also *subjective* contours).

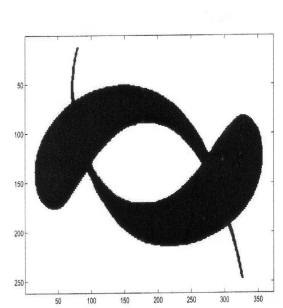


Figure 1.1: `The two Kanizsa's fishes.`

The perceptual completion can be *modal* (to extend contours as in Fig. 1.2) or *amodal* (to reconstract the shape of partially occluded objects as in Fig. 1.3).

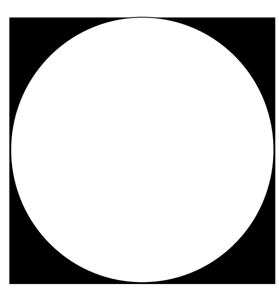


Figure 1.2: `Kanizsa's triangle: example of modal completion.`

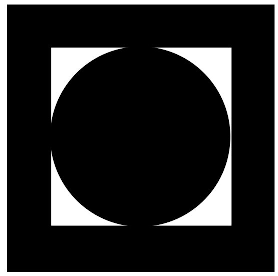


Figure 1.3: `Example of amodal completion.`

Mathematical models of this phenomenon have to take into account many different facts. All the classical models involve the minimizing of the elastic-functional, which is defined by

$$E(\gamma) = \int_\gamma \left(1 + k^2\right) d\gamma, \tag{1.3}$$

where $\gamma(t) = (x(t), y(t))$ is a curve in the plane and k indicates the curvature of γ, i.e.

$$-k(t) := \frac{\ddot{y}(t)\dot{x}(t) - \ddot{x}(t)\dot{y}(t)}{(\dot{x}^2(t) + \dot{y}^2(t))^{\frac{3}{2}}}.$$

The functional E is very difficult to study, in particular numerically, since it depends on the curvature which is a second-order differential operator.
Without giving details on the biologic models for the visaul cortex, we like just to recall that the orientation sensitive simple cells induce a fibration of orientations. So the natural space to study this phenomenon seems to be the 3-dimensional image-orientation manifold. Increasing the dimension of the space with the introduction of a suitable third component, it is possible to reduce the order of the function to study. Hence, we consider as third-component the "orientation" $\theta(t)$, which leads to:

$$\dot{x}(t) = \cos\theta(t) \qquad \dot{y}(t) = \sin\theta(t).$$

where we assumed the curve γ parametrized by length-arc.
We can so associate to any curve $\gamma(t) = (x(t), y(t))$ a 3-dimensional curve $\widehat{\gamma}(t) = (x(t), y(t), \theta(t))$ where $\dot{\theta}(t) = k(t)$.
The corresponding length-functional can be written as

$$l(\widehat{\gamma}) = \int_{\widehat{\gamma}} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{\theta}^2}\, d\widehat{\gamma} \tag{1.4}$$

It is possible to show that solving the minimization problem (1.3) is equivalent to the minimization problem (1.4), assuming the following condition on the admissible curves:

$$\dot{\hat{\gamma}}(t) = (\cos\theta(t), \sin\theta(t), k(t)). \tag{1.5}$$

This means that one can study a model for image completion by looking at the geodesics of the roto-traslation (sub-Riemannian) geometry.

1.1.3 Riemannian metrics.

Before introducing the sub-Riemannian geometries, it is useful to recall some basis notions about Riemannian metrics. For more details we refer to [22, 46, 48].

Definition 1.2. *Let M be a generic set and suppose that there exist $\{U_\alpha\}_\alpha$ open subsets of $\mathbb{R}^n$ and $\{x_\alpha : U_\alpha \to M\}_\alpha$ injective maps, such that*

(i) $\bigcup_\alpha x_\alpha(U_\alpha) = M$.

(ii) *For any α, β such that $x_\alpha(U_\alpha) \cap x_\beta(U_\beta) = W \neq \emptyset$, then $x_\alpha^{-1}(W)$ and $x_\beta^{-1}(W)$ are open sets in $\mathbb{R}^n$ and $x_\beta^{-1} \circ x_\alpha$ are smooth maps.*

(iii) *The family $\{(U_\alpha, x_\alpha)\}$ is maximal w.r.t. the assumptions (i)-(ii).*

Then we say that $\{(U_\alpha, x_\alpha)\}$ is a smooth structure *on the set M.*

Remark 1.3 (Induced topology)**.** *A smooth structure defined on a set M induces in a natural way a topology on the set. In fact, we can say that $A \subset M$ is an open subset of M whenever $x_\alpha^{-1}(A \cap x_\alpha(U_\alpha))$ is an open subset of $\mathbb{R}^n$, for any α. M endowed with this topology is a paracompact, Hausdorff topological space (see [46]). Moreover, for any α and U_α, $x_\alpha(U_\alpha)$ is an open set and x_α is a continuous map.*

To introduce the notion of curves on a smooth manifold, we first need to recall the following definitions.

Definition 1.4. *Let M be a smooth manifold, a* local parametrization *(or* system of coordinates*) at a point $p \in M$ is a map y defined on some neighborhood U of p, such that y is a diffeomorfism on its range.*

Definition 1.5. *Let M_1 and M_2 be smooth manifolds with dimension n and m respectively. A map $\phi : M_1 \to M_2$ is* differentiable *if and only if for any $p \in M_1$ and for any local parametrization (V, y) at $\phi(p) \in M_2$, there exists a local parametrization (U, x) at $p \in M_1$ such that*

(i) $\phi(x(U)) \subset y(V)$,

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

Next we introduce the notion of a curve on a smooth manifold.

Definition 1.6. *Let M be a smooth manifold and I a real interval, then any differentiable 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 following real functional

$$\begin{aligned} \dot{\gamma}(0) : \mathcal{D} &\longrightarrow \mathbb{R} \\ f &\longmapsto \dot{\gamma}(0)f := \frac{d(f \circ \gamma)}{dt}\Big|_{t=0} \end{aligned} \tag{1.6}$$

is a *tangent vector* at the point $p \in M$.

Definition 1.7. *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.*

$$T_pM := \{\dot{\gamma}(0) | \, \dot{\gamma}(0) \text{ satisfying } (1.6) \text{ and } \gamma(0) = p\}. \tag{1.7}$$

Remark 1.8. *If M is a n-dimensional smooth manifold, then for any point $p \in M$ the tangent space T_pM is a n-dimensional space.*

Definition 1.9. *Let M be a smooth manifold, we call* tangent bundle *of M the following $2n$-dimensional vector space:*

$$TM := \{(p, v) | \, p \in M, \ v \in T_pM\}. \tag{1.8}$$

Let M_1 and M_2 be two different smooth manifolds and let $\phi : M_1 \to M_2$ be a differentiable map between the two manifolds. By definition, for any $p \in M_1$ and $v \in T_pM_1$ there exists a curve $\gamma : I \to M_1$ such that $\gamma(0) = p$ and $\dot{\gamma}(0) = v$, so $\beta := \phi \circ \gamma$ defines a curve on the smooth manifold M_2

with $\beta(0) = \phi(p) \in M_2$. We can so define a linear map between the two corresponding tangent bundles TM_1 and TM_2 by considering

$$\begin{aligned} d\phi_p : T_pM_1 &\longrightarrow T_{\phi(p)}M_2 \\ v &\longmapsto d\phi_p(v) := \dot{\beta}(0) \end{aligned} \quad . \tag{1.9}$$

The previous map is well-defined, since it does not depend on the chosen curve. More details on this remark can be found in [46].

Definition 1.10. *Let M_1 and M_2 be two smooth manifolds and $\phi : M_1 \to M_2$ a differentiable map, the linear map $d\phi_p$ defined by* (1.9) *is called* differential *of ϕ at the point p.*

We now introduce the notion of a Riemannian metric, by using a system of local coordinates.

Definition 1.11 (Riemannian metric)**.** *A* Riemannian metric *on a smooth manifold M is an application from a point $p \in M$ to an inner product $\langle\ ,\ \rangle_p$ defined on the tangent space T_pM, which "changes in a differentiable way", i.e. such that, given a local coordinate map $\boldsymbol{x} : U \subset \mathbb{R}^n \to M$ around $p \in M$ and given a point $q = \boldsymbol{x}(x_1, ..., x_n) \in \boldsymbol{x}(U)$, then, set $\frac{\partial}{\partial x_i}(q) = d\boldsymbol{x}_q(0, ..., 1, ..., 0)$, the following function*

$$g_{i\,j}(x_1, ..., x_n) := \left\langle \frac{\partial}{\partial x_i}(q), \frac{\partial}{\partial x_j}(q) \right\rangle_q ,$$

is differentiable on whole U.

The function $g_{i\,j}$ is called *local representation of the Riemannian metric* w.r.t. the system of local coordinates $\boldsymbol{x} : U \subset \mathbb{R}^n \to M$.
Note also that a smooth manifold with a Riemannian metric defined on is usually called Riemannian manifold.

Remark 1.12. *In general we indicate an inner product simply as $\langle\ ,\ \rangle$ (i.e. omitting the base-point).*

Remark 1.13. *To characterize a Riemannian metric, we can equivalently require that for any couple of differential vector fields X, Y defined on the smooth manifold M the function $\langle X, Y\rangle$ is (locally) differentiable on M.*

The following theorem is one of the main result on the theory of Riemannian manifolds.

Theorem 1.14 ([46])**.** *On any Hausdorff, smooth manifold with a countable basis, it is always possible to define a Riemannian metric on.*

In order to introduce a notion of geodesics, we need first to define the length of curves.

Definition 1.15. *Let* $\big(M, \langle\ ,\ \rangle\big)$ *be a Riemannian manifold and* $\gamma : [0,T] \to M$ *an absolutely continuous curve, we call* length of the curve γ *the following real functional*

$$l(\gamma) := \int_0^T \big\langle \dot\gamma(t), \dot\gamma(t) \big\rangle^{\frac{1}{2}} dt. \tag{1.10}$$

Definition 1.16. *Let* $\big(M, \langle\ ,\ \rangle\big)$ *be a Riemannian manifold and* $p, q \in M$*, the* (Riemannian) distance *between these two points is defined as*

$$d(p,q) := \inf\{l(\gamma) \,|\, \gamma \text{ a.c. curve, joining } p \text{ to } q\}. \tag{1.11}$$

The geodesics are usually defined as the curves with vanishing acceleration. To make formal this definition, we have to say what we mean by acceleration, and so we need to introduce the so called Levi-Civita connection and covariant derivatives. Nevertheless, it is possible to show that all the curves with minimum-length (i.e. curves realizing the distance (1.11)) are geodesics. The inverse claim is not true. In fact, if one thinks of a sphere and two its no-antipodal points, there are two different arcs of maximum-circle joining them. Since any maximum-circle has vanishing acceleration, there are two different geodesics joining these two points but only one realizes the distance. Usually, the curves, minimizing the length are called *minimizing geodesics*. By geodesics we will always refer to these minimizing curves.

Theorem 1.17 (Existence and uniqueness)**.** *Let* M *be a Riemannian manifold and* $p_0 \in M$*, for any* $\varepsilon > 0$*, there exists a neighborhood* 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, then there exists at least a geodesic joining any pair of points.*

In order to get some global existence-result, the assumption of completeness is necessary. In fact, if one looks at $\mathbb{R}\setminus\{0\}$ with the Euclidean metric there are not geodesics joining two opposite points p and $-p$.

1.1.4 Carnot-Carathéodory metrics and the Hörmander condition.

The main difference between Riemannian and sub-Riemannian geometries is that in the sub-Riemannian case not every curve is admissible. In this section we want to introduce a rigorous mathematical defintion of sub-Riemannian geometry and study their main properties.

Definition 1.18. *Let M be a n-dimensional smooth manifold and $r \leq n$, a r-dimensional* distribution $\mathcal{H}$ *is a subbundle of the tangent bundle, i.e.* $\mathcal{H} := \{(p,v) \,|\, p \in M,\ v \in \mathcal{H}(p)\}$*, where $\mathcal{H}(p)$ is a r-dimensional subspace of the tangent space at the point p.*

Remark 1.19. *Note that sometimes the dimension of the distribution (which can be called also rank) can depend on the point p (see Example 1.22). Nevertheless the main sub-Riemannian geometries are associated to rank-constant distributions. E.g. the Heisenberg group and more in general any Carnot group (see Sec.1.3 and Sec.1.2).*

Definition 1.20 (Sub-Riemannian metric)**.** *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 fibers of the subbundle $\mathcal{H}$.*

Definition 1.21 (Sub-Riemannian geometry)**.** *A* sub-Riemannian geometry *is a tern $\big(M, \mathcal{H}, \langle\ ,\ \rangle\big)$, where M is a smooth manifold, $\mathcal{H}$ is a distribution, and $\langle\ ,\ \rangle$ is a Riemannian metric defined on $\mathcal{H}$.*

From now on, we indicate by $X_1, ..., X_m$ the vector fields spanning the distribution $\mathcal{H}$, i.e.

$$\mathcal{H}(p) = \mathrm{Span}(X_1(p), ..., X_m(p)),$$

at any point $p \in M$.

Example 1.22. *The Grušin plane is the sub-Riemannian geometry defined on* $\mathbb{R}^2$ *by the distribution spanned by the two 2-dimensional vector fields* $X(p) = (1,0)^t$ *and* $Y(p) = (0, x_1)^t$*, with* $p = (x_1, x_2) \in \mathbb{R}^2$*, endowed with the real Euclidean metric. Note that* $r(p) = 1$*, at the origin, while* $r(p) = 2$*, otherwise.*

It is possible to introduce a weak-linear-independent condition which includes both the rank-constant distributions and the Grušin-type spaces (which generalizes Example 1.22), assuming that for any $p \in M$ there exist $1 \leq r(p) \leq m$ and $1 \leq j_1 < ... < j_{r(p)} \leq m$ such that

$$\operatorname{rank}\{X_{j_1}(p), ..., X_{j_{r(p)}}(p)\} = r(p), \text{ and } X_j(p) = 0, \quad \forall\, j \notin \{j_1, ..., j_{r(p)}\}. \tag{1.12}$$

The fact that $\langle\ ,\ \rangle$ is defined only on $\mathcal{H}$ implies that we can define a notion of length only for a particular class of curves.

Definition 1.23 (Horizontal curves)**.** *Let* $\big(M, \mathcal{H}, \langle\ ,\ \rangle\big)$ *be a sub-Riemannian geometry and* $\gamma : [0,T] \to M$ *an absolutely continuous, we say that* γ *is a* horizontal *(or also admissible) 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 $h : [0,T] \to \mathbb{R}^m$ *such that*

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

where $h(t) = (h_1(t), \dots, h_m(t))$.

For the horizontal curves and only for those, it is possible to define a length-functional as

$$l(\gamma) := \int_0^T \|\dot\gamma(t)\|\, dt, \tag{1.13}$$

where $\|\dot\gamma(t)\| = \langle \dot\gamma(t), \dot\gamma(t)\rangle^{\frac{1}{2}}$.

Exactly as in the Riemannian case, by using the length for the horizontal curves we can introduce a notion of distance on the manifold.

Definition 1.24 (Carnot-Carathéodory distance)**.** *We call* sub-Riemannian distance *or also* Carnot-Carathéodory distance *the function* $d: M \times M \to [0, +\infty]$*, defined by*

$$d(p,q) := \inf\{l(\gamma) | \, \gamma \textit{ horizontal curve joining } p \textit{ to } q\}. \tag{1.14}$$

Theorem 1.25. *The function* $d(p,q)$ *defined by* (1.14) *induces a distance on the whole manifold* M*.*

Proof. The symmetry is easy to show. In fact for any horizontal curve γ which joins p to q in a time T, we can define the inverse curve $\widetilde{\gamma}(t) := \gamma(T-t)$. The curve $\widetilde{\gamma}$ is a horizontal curve joining q to p. Moreover $l(\gamma) = l(\widetilde{\gamma})$.

To show the triangular inequality, let us fix two points $p, q \in M$. First, we consider the case $d(p,q) < +\infty$. Given a third point $z \in M$, we can always assume that $d(p,z)d(z,q) < +\infty$. In fact, whenever $d(p,z) = +\infty$ or $d(z,q) = +\infty$, the triangular inequality is trivially satisfied. Then, given two horizontal curves γ_1 and γ_2 joining respectively p to z and z to q, we can consider the attached path $\gamma := \gamma_1 \vee \gamma_2$.
It is immediate to note that γ is still an horizontal curve. Moreover, γ joins p to q and $l(\gamma) \leq l(\gamma_1) + l(\gamma_2)$. Therefore

$$d(p,q) \leq d(p,z) + d(z,q). \tag{1.15}$$

The remaining case $d(p,q) = +\infty$ is trivail to prove. One has just to remark that, for any third point $z \in M$, then $d(p,z) = +\infty$ or $d(z,q) = +\infty$. In fact, if we assume that $d(p,z)d(z,q) < +\infty$, then there exist two horizontal curves γ_1 and γ_2, both having a finite length, joining respectively p to z and z to q. So the attached curve γ (defined as above) is an horizontal curve joining p to q and with $l(\gamma)$, which implies $d(p,q) \leq l(\gamma) < +\infty$, which leads to a contradiction.

Note that, since the length-functional is non-negative, then it $d(p,q) \geq 0$ for any $p, q \in M$. We only remain to prove that $d(p,q) = 0$ if and only if $p = q$. This result is a consequence of the local Euclidean estimate for Carnot-Caratéodory distances, which we are going to prove in the next subsection (see Lemma 1.44 and Remark 1.45). Therefore, so far we omit the proof of this property. □

Nevertheless it could happen that there is not any horizontal curve joining two given points. In this case the sub-Riemannian distance between these points is infinite. So we need to introduce a condition ensuring that the distance between two points is always finite.
At this purpose, let us recall that given two vector fields X, Y defined on some manifold M, the bracket between X and Y is the vector field defined as $[X, Y] := XY - YX$, i.e. actioning on smooth functions $f : M \to \mathbb{R}$ by derivation as

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

Example 1.26. *Let be $M = \mathbb{R}^2$, we set $p = (x, y) \in \mathbb{R}^2$, and consider the two vector fields $X(x, y) = (1, 0)^t$ and $Y(x, y) = (0, x)^t$. Then*

$$[X, Y](f) = x\frac{\partial^2}{\partial x \partial y} f(x, y) - \frac{\partial}{\partial y} f(x, y) - x\frac{\partial^2}{\partial y \partial x} f(x, y) = -\frac{\partial}{\partial y} f(x, y).$$

By induction, a k-length bracket is a vector field defined as $[Z_i, Z_j^{(k-1)}]$, where $Z_i \in \{X, Y\}$ and $Z_j^{(k-1)}$ is a bracket between X and Y with length less or equal to $k - 1$. For example the 2-length-brackets between two vector fields X and Y are given by $[X, [X, Y]], [X, [Y, X]], [Y, [X, Y]], [Y, [Y, X]]$.

Note that $[X, Y] = -[Y, X]$, hence $[X, Y] = [Y, X]$ if and only if $[X, Y] = 0$, i.e. if and only if the two vector fields commute.

Let us consider a family of vector fields $\mathcal{X} = \{X_1, ..., X_m\}$ spanning some distribution $\mathcal{H} \subset TM$, the associated Lie algebra is the set of all the brackets between the vector fields of the family, i.e.

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

Definition 1.27. *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.*

Definition 1.28 (Hörmander condition)**.** *We say that a sub-Riemannian geometry satisfies the* Hörmander condition *if and only if the associated distribution is bracket generating.*

Theorem 1.29 (Chow's Theorem)**.** *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 two given points of M.*

From Chow's Theorem it follows that, whenever a sub-Riemannian geometry satisfies the Hörmander condition, the distance defined by (1.14) is finite at any pair of points. In the Subsection 1.1.6, we are going to give more details on Chow's Theorem and show an easy proof in the particular case of the Heisenberg group.
To introduce the notion of step of a bracket generating distribution, let us introduce the following notion. Let $\mathcal{H}$ be a family of vector fields, we write:

$$\begin{aligned}
&\mathcal{L}^1 := \mathrm{Span}(\{Z = X \,|\, X \in \mathcal{H}\}),\\
&\mathcal{L}^2 := \mathrm{Span}(\{Z = [X,Y] \,|\, X, Y \in \mathcal{L}^1\}),\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&\mathcal{L}^i = \mathrm{Span}(\{Z = [X,Y] \,|\, X \in \mathcal{H}, Y \in \mathcal{L}^{i-1}\}).
\end{aligned} \tag{1.16}$$

Moreover, we indicate by $\mathcal{L}^i(p)$ the vector space corresponding to $\mathcal{L}^i$ evaluated at the point $p \in M$.

Definition 1.30 (Step of a distribution)**.** *Let $\mathcal{X} = \{X_1, ..., X_m\}$ be a family of vector fields defined on a smooth manifold M and $\mathcal{H}$ the distribution generated by $X_1, ..., X_m$. 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*

$$\bigcup_{i=1}^{k(p)} \mathcal{L}^i(p) = T_pM.$$

Example 1.31. *Any Riemannian geometry is bracket generating with step equal to 1 at any point. In fact $\big(M, \mathcal{H}, \langle\ ,\ \rangle\big)$ is a Riemannian manifold if and only if $\mathcal{H} = TM$ (which means $\mathcal{L}^1(p) = T_pM$ for any $p \in M$).*

Example 1.32. *The Heisenberg group is associated to a bracket generating with step equal to 2 at any point (see Definition 1.106).*

Example 1.33. *The Grušin plane is associated to a bracket generating with step 2 at the origin, and with step 1 otherwise (see Definition 1.22).*

The reverse implication of Chow's Theorem ("finite distance implies the Hörmander condition") holds only for analytic distributions but it is in general false for smooth distributions, see the following counter-example.

Example 1.34. *Look at the sub-Riemannian metric generated by the 2-dimensional vector fields* $X = (1,0)^t$ *and* $Y = (0, a(x))^t$, *with* $a \in C^\infty(\mathbb{R})$ *such that* $a(x) = 0$, *if* $x \leq 0$, *and* $a(x) > 0$, *if* $x > 0$, *so that* X, Y *are smooth vector fields. The corresponding sub-Riemannian distance is finite. Nevertheless, the associated distribution is not bracket generating. In fact, if* $x \leq 0$, *then* $Y = (0,0)^t$ *and so* $\mathrm{Span}(\mathcal{L}(X,Y)(p)) = \mathrm{Span}(X(p)) \neq \mathbb{R}^2$.

Example 1.35. *Let* X *and* Y *be as in the Example 1.34, but with* $a(x) = 1$ *if* $x \geq 0$, *and* $a(x) = 0$ *if* $x < 0$. *In this case* $a \notin C^\infty(\mathbb{R})$, *however, we can use this example in order to investigate the previous one.*
On the half-plane $x < 0$, *we can move only in one direction, then the spanned distribution is not bracket generating.*
Nevertheless, it is easy to write explicitly the associated Carnot-Carathéodory distance, that is

$$d((x,y),(x',y')) = \begin{cases} \sqrt{|x-y|^2 + |x'-y'|^2}, & x_1 \geq 0 \; x' \geq 0, \\ |x| + |x'| + |y-y'|, & x < 0 \; x' < 0, \\ |x| + \sqrt{|x'|^2 + |y-y'|^2}, & x < 0 \; x' \geq 0, \\ |x'| + \sqrt{|x|^2 + |y-y'|^2}, & x \geq 0 \; x' < 0. \end{cases}$$

It is immediate to note that $d(x,y)$ *is a finite distance.*

Remark 1.36 (Involutive distributions)**.** *The opposite case to the bracket generating distributions is given by the so called involutive distributions. We recall that a distribution* $\mathcal{H}$ *is said involutive if and only if* $[X,Y] \in \mathcal{H}$, *for any* $X, Y \in \mathcal{H}$. *By the Frobenius Theorem (see [50]), it is known that, if* M *is a* n*-dimensional smooth manifold and* $\mathcal{H}$ *is a* r*-dimensional involutive distribution defined on* M, *then, given* $p \in M$, *the set of all the horizontal curves through a fixed point* p, *is an immerse* r*-dimensional submanifold,*

called leaf. *Therefore, if $q \in M$ does not belong to the lift of p, there is not any horizontal curve joining p to q. So, in this case, $d(p,q) = +\infty$.*
In particular, if $r < n$ such a point q always exists.

1.1.5 Notions equivalent to the Carnot-Carathéodory distance.

Next we want to intorduce a notion of distance equivalent to the Carnot-Carathéodory distance. This new distance is very used in control theory and so it is often called *control distance* (or also *minimal time distance*). Let $\mathcal{X} = \{X_1, ..., X_m\}$ be smooth vector fields generating a distribution $\mathcal{H}$, we recall that an absolutely continuous curve $\gamma : [0,T] \to M$ is horizontal if and only if

$$\dot{\gamma}(t) = \sum_{i=1}^{m} h_i(t) X_i(\gamma(t)), \quad \text{a.e. } t \in [0,T], \tag{1.17}$$

for suitable $h_i(t)$ measurable functions. We can think of (1.17) as a control system, which we know to be well-posed whenever $h_i \in L^1([0,T])$. We like to recall that in control theory the functions h_i are usually called *control functions* while the solutions of (1.17) are called *control paths*.

Remark 1.37. *If the vector fields $X_1, ..., X_m$ are linearly independent at any point, then the coordinates h_i are unique. Moreover the uniqueness still holds if we assume the weak-linearly-independent-condition* (1.12).

Under assumption (1.12), we can define the length of a horizontal curve $\gamma : [0,T] \to M$, by using the local coordinates h_i, i.e.

$$l(\gamma) = \int_0^T \left(h_1^2(t) + ... + h_m^2(t)\right)^{\frac{1}{2}} dt. \tag{1.18}$$

Remark 1.38. *In general, the control functions h_i can be non-unique. In such a case, we define the length-functional taking the infimum of* (1.18) *over all the admissible coordinates h_i. More precisely, $h_1, ..., h_m$ are admissible coordinates if and only if they satisfy* (1.17) *and, moreover, $h_1, ..., h_m \in L^1([0,T])$ (otherwise the corresponding integral is equal to $+\infty$). This implies*

$$l(\gamma) = \inf\left\{ \int_0^T \left((h_1(t))^2 + ... + (h_m(t))^2\right)^{\frac{1}{2}} dt \,\middle|\, h_i \in L^1([0,T]) \text{ satisfying } (1.17) \right\}.$$

We so can re-write the sub-Riemannian distance defined in (1.14) by using the local expression (1.18).
Next we introduce a definition of sub-Riemannian distance as suitable minimal time function.

Definition 1.39. *We say that a horizontal curve $\gamma : [0,T] \to M$ is* subunit *if and only if the coordinates given by* (1.17) *are such that*

$$\sum_{i=1}^{m} h_i^2(t) \leq 1, \quad a.e.\ t \in [0,T].$$

Definition 1.40. *For any $x, y \in M$, we look at the following minimal time function:*

$$\widehat{d}(x,y) = \inf\{T \,|\, \exists\, \gamma : [0,T] \to M \text{ a.c. subunit hor. curve joining } x \text{ to } y\}. \tag{1.19}$$

Proposition 1.41 ([76], Proposition 1.1.10)**.** *The function $\widehat{d}(x,y)$ defines a distance on the smooth manifold M.*

We want to show that the minimal time distance $\widehat{d}(x,y)$ is equivalent to the Carnot-Carathódory distance. This is a consequence of the following result.

Theorem 1.42 ([76])**.** *Let γ be a horizontal curve and, let us assume for sake of simplicity that $T = 1$. If $h = (h_1, ..., h_m)^t$ are the local coordinates of $\dot{\gamma}$ w.r.t. $X_1, ..., X_m$, then for any $1 \leq p \leq +\infty$ we can define a distance on M by*

$$d^p(x,y) = \inf\{l_p(\gamma) \,|\, \gamma : [0,1] \to M \text{ a.c. hor. curve joining } x \text{ to } y\}, \tag{1.20}$$

with

$$l_p(\gamma) := \|h\|_p = \begin{cases} \left(\displaystyle\int_0^1 |h(t)|^p dt\right)^{\frac{1}{p}}, & \text{if} \quad 1 \leq p < +\infty, \\ \underset{t\in[0,1]}{\text{ess. sup}}\, |h(t)|, & \text{if} \quad p = +\infty, \end{cases}$$

where the essential supremum of $h(t)$ is defined by

$$\underset{t\in[0,1]}{\text{ess. sup}}\, |h(t)| = \inf\left\{M \geq 0 \,|\, |h(t)| \leq M,\ a.e.\, t \in [0,1]\right\}.$$

Then $d^p(x,y) = \widehat{d}(x,y)$ for any $x, y \in M$ and for any $1 \leq p \leq +\infty$.

Proof. For sake of completeness, since the equivalence between the Carnot-Carathédory distance and the control distance is a key point for the study of nonlinear PDEs related to the Hörmander condition, we quote give a proof of this theorem.
First note that by the Hölder inequality it follows that

$$\|h\|_1 \leq \|h\|_p \leq \|h\|_\infty ,$$

for any $h \in [L^\infty(0,1)]^m$ and $1 \leq p \leq +\infty$. Hence

$$d^1(x,y) \leq d^p(x,y) \leq d^\infty(x,y). \tag{1.21}$$

The second step is to prove that

$$\widehat{d}(x,y) = d^\infty(x,y). \tag{1.22}$$

Let $\gamma : [0,T] \to M$ be a subunit horizontal curve such that $\gamma(0) = x$ and $\gamma(T) = y$ defined by

$$\dot{\gamma}(t) = \sum_{i=1}^{m} h_i(t) X_i(\gamma(t)), \tag{1.23}$$

where $h = (h_1, ..., h_m)^t$ and $\|h\|_\infty \leq 1$.
We consider the rescaled curve $\widetilde{\gamma} : [0,1] \to M$ defined by $\widetilde{\gamma}(t) := \gamma(T\,t)$, which is still horizontal and such that

$$\widetilde{h}(t) = \left\|\widetilde{h}\right\|_\infty \leq T.$$

Then $d^\infty(x,y) \leq \widehat{d}(x,y)$.
To prove the reverse inequality one can proceed similarly.
Hence (1.22) is verified.

The third step consists in proving that

$$d^\infty(x,y) \leq d^1(x,y). \tag{1.24}$$

This step is the most difficult one. Let $\gamma : [0,1] \to M$ be a horizontal curve such that $\gamma(0) = x$ and $\gamma(1) = y$ and let $h = (h_1, ..., h_m)$ be defined by (1.23). We need to build a new horizontal curve $\widetilde{\gamma}$ such that $l_\infty(\widetilde{\gamma}) \leq \|h\|_1$.

Suppose that $\|h\|_1 > 0$ and let $\phi : [0,1] \to [0,1]$ be the following absolutely continuous function:

$$\phi(t) = \frac{1}{\|h\|_1} \int_0^t |h(\tau)| d\tau,$$

ϕ is non decreasing and its "inverse function" $\psi : [0,1] \to [0,1]$ defined as

$$\psi(s) = \inf\{t \in [0,1] | \, \phi(t) = s\}$$

is a monotone function, too. So it is differentiable a.e. $s \in [0,1]$.
We want to prove that

$$\dot{\phi}(\psi(s))\dot{\psi}(s) = 1, \quad \text{a.e. } s \in [0,1]. \tag{1.25}$$

Therefore we define

$$B = \{t \in [0,1] \,|\, \phi \,\text{is not differentiable in}\, t\}$$

and

$$D = \{t \in [0,1] \,|\, \phi \,\text{is not differentiable in}\, \psi(t)\}.$$

Since ϕ is an absolutely continuous function, it is such that vanishing-measure sets go into vanishing-measure sets, i.e. $|B| = 0$ implies $|\phi(B)| = 0$.
Moreover $D \subset \phi(B)$, so $|D| = 0$ and then we have (1.25).
So we can define $\widetilde{\gamma}(s) := \gamma(\psi(s))$. Now, let be

$$E = \{t \in [0,1] \,|\, \gamma \,\text{is not differentiable in}\, \psi(t)\},$$

γ is absolutely continuous, then the previous argument shows also that $|E| = 0$. Hence

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

If $|h(\psi(s))| \neq 0$, it is trivial to remark that

$$\dot{\psi}(s) = \frac{1}{\dot{\phi}}(\psi(s)) = \frac{\|h\|_1}{|h(s)|}.$$

We can define

$$\widetilde{h}_i(s) = \begin{cases} \|h\|_1 \dfrac{h_i(\psi(s))}{|h(\psi(s))|}, & \text{if } |h(\psi(s))| \neq 0, \\ 0, & \text{if } |h(\psi(s))| = 0. \end{cases}$$

Setting

$$\dot{\widetilde{\gamma}}(s) = \sum_{i=1}^{m} \widetilde{h}_i(s) X_i(\widetilde{\gamma}(s)), \quad \text{a.e. } s \in [0,1],$$

we get a horizontal curve such that

$$\left\| \widetilde{h} \right\|_{\infty} \leq \|h\|_1 ,$$

so that (1.24) holds. Hence, by estimate (1.21), we can conclude

$$d^{\infty}(x,y) = d^1(x,y) = d^p(x,y),$$

for any $1 \leq p \leq +\infty$, and, therefore, by (1.22), we get $\widehat{d}(x,y) = d^p(x,y)$, for any $x, y \in M$. □

Corollary 1.43. *The Carnot-Carathéodory distance $d(x,y)$ defined by* (1.14) *is equivalent to the minimal time distance $\widehat{d}(x,y)$ defined by* (1.19).

Proof. The Theorem 1.42 with $p = 1$ implies that $d^1(x,y) = \widehat{d}(x,y)$. We need only to recall that the Euclidean norm of $\mathbb{R}^m$ is equivalent to the norm defined as $|(h_1, ..., h_m)| = |h_1| + ... + |h_m|$. This implies that $d^1(x,y)$ is equivalent to $d(x,y)$ and then $\widehat{d}(x,y)$ is so. □

To conclude this subsection, we show an Euclidean estimate from below for the sub-Riemannian distance $\widehat{d}(x,y)$. Note that, since the distances $\widehat{d}(x,y)$ and $d(x,y)$ are equivalent, the same estimate holds for $d(x,y)$.

Lemma 1.44. *Let M be a smooth manifold and $\mathcal{X} = \{X_1, ..., X_m\}$ smooth vector fields generating a distribution $\mathcal{H}$ and satisfying the Hörmander condition with step equal to $k \geq 1$. If $\widehat{d}(x,y)$ is the minimal time distance defined by* (1.19)*, then for any $K \subset M$ compact there exists a constant $C > 0$ such that*

$$C|x-y| \leq \widehat{d}(x,y), \quad \text{for any } x, y \in K. \tag{1.26}$$

Proof. Let us fix a compact set K and choose $0 < \varepsilon << 1$ such that

$$K_{\varepsilon} = \left\{ z \in M \,\middle|\, \min_{x \in K} |z - x| \leq \varepsilon \right\} \subset\subset M.$$

At any point $x \in M$ we can define a $n \times m$-matrix by

$$A(x) := [X_1(x), ..., X_m(x)],$$

and

$$M = \sup_{x \in K_\varepsilon} \|A(x)\|,$$

where by $\| \quad \|$ we indicate the usual norm of matrices.
Fix $x, y \in K$ and let $\gamma : [0, T] \to M$ be a subunit horizontal curve such that $\gamma(0) = x$ and $\gamma(T) = y$.
Note that, by the Hörmander condition, such a curve γ always exists. Then we set $r = \min\{\varepsilon, |x - y|\}$ so that $TM \geq r$ (one can find a detailed proof of this claim in [76], Lemma 1.1.8).
Therefore, if we choose $r = \varepsilon$ and consider the diameter D of K defined as

$$D := \sup\{|x - y| \mid x, y \in K\},$$

we get

$$T \geq \frac{\varepsilon}{M} \geq \frac{\varepsilon}{MD}|x - y|. \tag{1.27}$$

While, if we choose $r = |x - y|$, we find

$$T \geq \frac{|x - y|}{M}. \tag{1.28}$$

Since γ is an arbitrary subunit horizontal curve joining x to y, from (1.27) and (1.28) it follows that

$$\widehat{d}(x, y) \geq \min\left\{\frac{1}{M}, \frac{\varepsilon}{MD}\right\}|x - y|.$$

Passing to the limit, as $\varepsilon \to 0^+$ we get estimate (1.26). □

Remark 1.45. *From the previous estimate, it follows that the sub-Riemannian distance $d(x, y)$ is positive definite, so this remark concludes the proof of Proposition 1.25.*

In Sec.1.1.7 we will show that there exists also an Euclidean estimate from above but it is an Hölder-estimate while the estimate from below is Lipschitz-estimate.

1.1.6 Chow's Theorem.

Chow's Theorem is the main result for bracket generating distributions. In this subsection we want to give a sketch of the non-trivial proof of this result. Nevertheless, before proving the theorem in the general case, we like to quote a very simple and nice proof by Gromov in [51], which holds in the particular case of the 1-dimensional Heisenberg group. So we need to introduce briefly the 1-dimensional Heisenberg group (more details will be given in Sec.1.3). We call (1-dimensional) Heisenberg group the sub-Riemannian geometry defined on $\mathbb{R}^3$ by the distribution $\mathcal{H}$, associated to the vector fields

$$X = \begin{pmatrix} 1 \\ 0 \\ -\frac{y}{2} \end{pmatrix}, \quad Y = \begin{pmatrix} 0 \\ 1 \\ \frac{x}{2} \end{pmatrix}, \quad \text{for any } \ (x, y, z) \in \mathbb{R}^3,$$

and endowed with the standerd Euclidean metric (on $\mathbb{R}^2$).
Note that, if we consider the 1-form

$$\eta := dz - \frac{1}{2}(xdy - ydx),$$

then $\mathcal{H} = \ker(\eta)$.
We indicate the n-dimensional Hiesenberg group by $\mathbb{H}^n$ and by $\mathbb{H}^1$ the 1-dimensional Hiesenberg group.
Note that a curve $\gamma : [0, T] \to \mathbb{R}^3$ is $\mathbb{H}^1$-horizontal, if and only if, $\eta(\gamma(t)) = 0$, a.e. $t \in [0, T]$. This definition is called the canonical definition of 1-dimensional Heisenberg group.
There is also another definition: the polarized Heisenberg group. In Sec.1.3.3 we will show that these two definitions are indeed equivalent.
Gromov uses this second definition but we prefer to rewrite the Gromov's proof using the canonical definition.
Next we rewrite Chow's Theorem in the particular case of the 1-dimensional Heisenberg group.

Theorem 1.46. *Given two points in* $\mathbb{R}^3$*, there exists an absolutely continuous* $\mathbb{H}^1$*-horizontal curve joining them.*

Proof. 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 $\widetilde{\gamma}(t) = (x(t), y(t))$ be a plane curve joining (x_1, y_1) to (x_2, y_2). For sake

of simplicity we assume $T = 1$.
Remark that we can look only at the absolutely continuous curves with constant curvature, i.e. we can assume that

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

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

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

Obviously γ is an absolutely continuous curve in $\mathbb{R}^3$. Moreover, since $z(0) = z_1$ and $z(1) = z_1 - C$, choosing $C = z_1 - z_2$, then γ joins p to q.
In order to conclude the proof, we need only to observe that, for a.e. $t \in [0, 1]$, it holds $\eta(\gamma(t)) = 0$ and so γ is a $\mathbb{H}^1$-horizontal curve. $\square$

To prove Theorem 1.29 in the general case is more difficult.
The general result was proved, almost contemporaneously but indipendently, by Rashevsky in [82] (1938) and by Chow in [30] (1939). There are many different proofs of this result. We choose to briefly sketch the proof given in [17], by using the point-of-view of control theory. First we need to recall some definitions. From now on we indicate by $B_R^d(p)$ the open ball centered at p with radius R, w.r.t. the metric $d(x, y)$, i.e.

$$B_R^d(p) = \{q \in M \mid d(p, q) < R\}.$$

Definition 1.47. *Let M be a connected sub-Riemannian manifold and $p \in M$, the* accessible set *A_p is the set of all the points of M joined to p by a horizontal curve.*

Definition 1.48. *An* immersed submanifold *of a manifold M is a subset $A \subset M$ endowed with a manifold structure such that*

(i) the inclusion map $i : A \to M$ is an immersion,

(ii) any continuous map $f : P \to M$, where P is a manifold, is continuous when we consider the restriction map $f : P \to A$, where A is endowed with its manifold topology.

To prove Chow's Theorem is equivalent to show that, under the Hörmander condition, the accessible set of any points of the manifold coincides with the manifold itself. The key is the following property for the accessible sets.

Theorem 1.49 (Sussmann-Stefan's Theorem, [17]). *Let M be a connected smooth manifold, then for any $p \in M$ the accessible set A_p is an immersed submanifold.*

By using Sussemann-Stefan's Theorem, Chow's Theorem follows immediately.

Proof of Theorem 1.29. Fix $p \in M$ and look at the accessible set A_p.
Note that $q \in A_p$ if and only if $d(p,q) < +\infty$, so we can write A_p as union of the open balls $B^d_R(p)$. Then A_p is an open subset of M.
Moreover by Sussemann-Stefan's Theorem we know that A_p is an immersed submanifold of M, therefore the vectors fields $X_1, ...X_m$ can be seen also as tangent bundles to the immersed submanifold, i.e.

$$X_1, ...X_m \in TA_p.$$

Remember that, if some vector fields belong to a tangent space, then also their brackets belong to it. So

$$[X_i, X_j], [[X_i, X_j], X_k], ... \in TA_p.$$

By the bracket generating condition, we have that $TA_p = TM$, which in particular implies that the two spaces have the same dimension. Since A_p is an open immersed submanifold of M, whenever it has the same dimension of the manifold, it must coincide with a connected component of M. M is connected, so $A_p = M$ and this conclude the proof of Chow's Theorem. $\square$

1.1.7 Relationship between the Carnot-Carathéodory distance and the Euclidean distance.

Now we want to study the relationship between any sub-Riemannian distance satisfying the Hörmander condition and the Euclidean distance.
First, we show that both of these distances induce on $\mathbb{R}^n$ the same topology. Then we prove the main (local) estimates for Carnot-Carathéodory distances.

One can find a proof of this topological result in [75] (Theorem 2.3, Sec.I.2.5). There, the Author gets the equivalence of the two topologies, directly from the Ball-box Theorem (Theorem 2.10, pages 29-30), fixing a neighborhood basis in the sub-Riemannian topology and building a suitable neighborhood basis in the original topology of the manifold. Nevertheless, we choose to follow the approach given in [17].
So look at the control system (1.17) and let p be an initial point and U_{pT} an open neighborhood of the origin in $L^1([0,T],\mathbb{R}^m)$.

Definition 1.50. *We call* end-point map *the function $E_p : U_{pT} \to M$, defined as $h \longmapsto x_h(T)$, where x_h a solution of* (1.17)*, w.r.t. the control function h.*

We quote the following result without giving any proof.

Theorem 1.51 (End-point mapping Theorem, [17])**.** *Let M be a smooth manifold, $\mathcal{H}$ a bracket generating distribution and $d(x,y)$ the associated sub-Riemannian distance. Then the end-point map is open.*

By the End-point mapping Theorem, we can deduce the following result.

Theorem 1.52. *Let M, $\mathcal{H}$ and $d(x,y)$ be as in Theorem 1.51, then $d(x,y)$ induces on M the original topology defined on the manifold.*

Proof. Note that $B_R^d(p)$ is the image of the ball $B_R(0)$ in L^1 under the end-point map, then by the End-point mapping Theorem $B_R^d(p)$ is an open set in M.
In order to prove the inverse result, we need to fix a point $p \in M$ and a neighborhood U of p. Since the end-point map E_p is continuous at 0, so

there exists $R > 0$ such that E_p maps the ball $B_R(0) \subset L^1$ into U. Hence any neighborhood U contains a ball $B^d_R(p)$ and this concludes the proof. □

The previous theorem applied in $\mathbb{R}^n$ endowed with the Euclidean topology implies the compactness of the sub-Riemannian balls, whenever the associated distribution is bracket generating.

Corollary 1.53. *Let $\mathcal{H}$ be a bracket generating distribution defined on $\mathbb{R}^n$ and $d(x, y)$ the associated sub-Riemannian distance. Then a closed d-ball*

$$B^d_R(x) := \{y \in \mathbb{R}^n \mid d(x, y) \leq R\}, \quad \textit{for some } R > 0 \textit{ and } x \in \mathbb{R}^n,$$

is compact in the n-dimensional Euclidean space.

Proof. Let τ_1 be the topology induced by the metric $d(x, y)$ and τ_2 the Euclidean topology. They are equivalent, if and only if, for any fixed $x \in \mathbb{R}^n$, there exist $\mathcal{B}_1$ and $\mathcal{B}_2$ neighborhood-basis, w.r.t. τ_1 and τ_2, respectively, such that for all $B \in \mathcal{B}_1$ there exist $U, V \in \mathcal{B}_2$ with $U \subset B \subset V$.
We can set $\mathcal{B}_1 = \{B^d_\varepsilon(x) \mid \varepsilon > 0\}$ and $\mathcal{B}_2 = \{B_\varepsilon(x) \mid \varepsilon > 0\}$, where we indicate by $B^d_\varepsilon(x)$ the ball w.r.t. the sub-Riemannian metric $d(x, y)$, and by $B_\varepsilon(x)$ the usual Euclidean ball, both of them centered at x, with radius ε. Then, the equivalency of the two topologies implies that, for any $x \in \mathbb{R}^n$ and $\varepsilon > 0$, there exist $r, R > 0$ such that

$$B_r(x) \subset B^d_\varepsilon(x) \subset B_R(x).$$

Therefore the d-balls are bounded sets, then $\overline{B^d_\varepsilon(x)}$ are compact w.r.t. the Euclidean topology, for any $\varepsilon > 0$. □

Remark 1.54. *Remember that the two metrics induce the same topology but they are not equivalent. In fact fixed $y \in \mathbb{R}^n$, we get that the radii depend on x and y. Hence, the two metrics in general are not equivalent.*

Remark 1.55. *Note that the Hörmander condition is very important in order to get the equivalence of the two topologies. In Example 1.35, we introduced a finite sub-Riemannian distance, which does not satisfy the Hörmander condition. It is very easy to see that such a distance is discontinuous w.r.t. the Euclidean topology. In fact, for any $x < 0$*

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

We give now the main estimates for Carnot-Carathéodory distances satisfying the Hörmander condition.

Theorem 1.56. *Let $d(x, y)$ be a sub-Riemannian distance defined on a smooth manifold M 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*

$$C_1|x-y| \leq d(x,y) \leq C_2|x-y|^{\frac{1}{k}}, \tag{1.29}$$

for any $x, y \in K$.

Proof. The estimate from below is given in Lemma 1.44.
We remain to show the estimate from above using the Hörmander condition and introdusing a new distance equivalent to the Carnot-Carathéodory distance.
Let us recall that, thanks to the Hörmander condition, any absolutely continuous curve γ satisfies

$$\dot{\gamma}(t) = \sum_{i=1}^{n} h_i(t) Y_i(\gamma(t)), \quad \text{a.e. } t, \tag{1.30}$$

where h_i are measurable functions and Y_i is a basis of the tangent bundle TM, where only brackets between the elements $X_1, ..., X_m$ and with length less or equal to k, appear (for more details see e.g. [17]).
Let $C(\delta)$ be the set of all the absolutely continuous curves $\gamma : [0,1] \to M$ such that (1.30) holds with $|h_i(t)| < \delta^s$ for $i = 1, ..., n$ and $Y_i \in \mathcal{L}^s$ (i.e. Y_i is a bracket with length equal to s). Then we define a distance ρ by

$$\rho(x,y) = \inf\{\delta > 0 | \, \exists\, \gamma \in C(\delta) \, : \, \gamma(0) = x, \gamma(1) = y\}.$$

Fixed a compact K and two points $x, y \in K$, there exists an absolutely continuous curve $\gamma : [0,1] \to M$, joining x to y and such that $|\dot{\gamma}(t)| \leq C|x-y|$, for suitable constant $C = C(K)$ and for a.e. $t \in [0,1]$.
From (1.30), it follows that

$$|h_i(t)| \leq C_1|\dot{\gamma}(t)| \leq C_2|x-y| = C_2\big(|x-y|^{\frac{1}{s}}\big)^s,$$

whenever $Y_i \in \mathcal{L}^s$. Since $0 < s \leq k$, we have proved that

$$\rho(x,y) \leq C|x-y|^{\frac{1}{k}}.$$

The distance ρ is equivalent to the Carnot-Carathéodory distance $d(x,y)$ (see [78], Theorem 4) and this concludes the proof. □

Remark 1.57. *Estimate* (1.29) *implies that the two metrics are equivalent whenever $k = 1$, that is the Riemannian case.*

1.1.8 Sub-Riemannian geodesics.

In this subsection we are going to intorduce and study sub-Riemannian geodesics.

Definition 1.58 (Geodesic)**.** *A* minimizing geodesic *(or simply* geodesic*) between two points x and y is any absolutely continuous horizontal curve which realizes the distance* (1.14)*.*

As we have briefly remarked in the section about the Dido's problem and the perceptual visual completion, sub-Riemannian geodesics can be used in order to solve minimization problems of the calculus of variations.

Exactly as in the Riemannian case, the geodesics can be found also minimizing the energy E among all the horizontal curves instead of the length-functional (1.13). Let us give more details on this point.

Definition 1.59. *Let $\gamma : [0,T] \to M$ be an absolutely continuous horizontal curve, we call* energy *of γ the following functional:*

$$E(\gamma) = \int_0^T \frac{1}{2} \|\dot{\gamma}(t)\|^2 \, dt = \frac{1}{2} \int_0^T \big\langle \dot{\gamma}(t), \dot{\gamma}(t) \big\rangle dt. \tag{1.31}$$

By the Cauchy-Schwartz inequality, it follows that

$$\int fg \leq \sqrt{\int f^2} \sqrt{\int g^2},$$

with the identity true if and only if $f = cg$ for some constant c.
If we apply the previous inequality to the functions $f = \|\dot{\gamma}\|$ and $g = 1$, we get that, for any horizontal curve $\gamma : [0, T] \to M$, it holds

$$l(\gamma) = \int_0^T \|\dot{\gamma}(t)\| \, dt \le \sqrt{T} \sqrt{\int_0^T \|\dot{\gamma}(t)\|^2 \, dt} = \sqrt{T}\sqrt{2E(\gamma)}.$$

From this remark, we find the following result.

Proposition 1.60 ([75])**.** *A curve γ minimizes the energy-functional E among all the horizontal curves joining q to p in a time T if and only if γ minimizes the length-functional among all the horizontal curves joining q to p, parameterized by the constant-time $c = d(p,q)/T$.*

Since the length of a curve does not depend on the chosen parametrization, we can always restrict to looking at the infimum only among the curves with the previous constant parametrization. Therefore, by Proposition (1.60) it follows that minimizing the energy or minimizing the length is equivalent.

We now introduce equations to compute the geodesics in sub-Riemannian geometries. At this purpose, we need to associate a Hamiltonian to a given sub-Riemannian geometry.
We define the map $\beta : T^*M \to TM$ acting as $p(\beta_x(q)) = \langle\langle p, q \rangle\rangle_x$, for any $p, q \in T^*M$ and $x \in M$, where by $\langle\langle \ , \ \rangle\rangle$ we indicate the cometric associated to the sub-Riemannian metric $\big\langle \ , \ \big\rangle$

Remark 1.61 ([75])**.** *The map β associated to a sub-Riemannian geometry $\big(M, \mathcal{H}, \big\langle \ , \ \big\rangle\big)$ is uniquely defined by the following two conditions:*

1. *$Im(\beta_x) = \mathcal{H}_x$,*
2. *$p(v) = \langle \beta_x(p), v \rangle$ for any $v \in \mathcal{H}_x$ and any $p \in T^*M$.*

Definition 1.62 (Hamiltonian)**.** *The quadratic form*

$$H(x,p) = \frac{1}{2} \langle p, p \rangle_x \tag{1.32}$$

is called sub-Riemannian Hamiltonian *(or also* kinetic energy*).*

Suppose that γ is a horizontal curve, then there exists $p \in T^*M$ such that $\dot{\gamma}(t) = \beta_{\gamma(t)}(p)$ and

$$\frac{1}{2}\|\dot{\gamma}\|^2 = H(\gamma(t), p). \tag{1.33}$$

By the uniqueness of the map β, we get the following characterization for sub-Riemannian geometries.

Proposition 1.63 ([75], Proposition 1.10)**.** *Any sub-Riemannian geometry is uniquely determinate by its Hamiltonian. Conversely, any non-negative quadratic Hamiltonian with constant rank $k \geq 1$ defines a unique sub-Riemannian structure associated to a k-rank-constant distribution.*

To get the equations of geodesics, we introduce the momentum.

Definition 1.64. *Let X be a vector field defined on a smooth manifold M, we call* momentum *of X the following functional defined on the cotangent space:*

$$P_X(x, p) = p(X(x)), \tag{1.34}$$

for any $p \in T^*M$ and $x \in M$.

Let $\{X_\alpha\}_\alpha$ be a family of vector fields spanning the distribution $\mathcal{H}$. If we look at some system of local coordinates x_i defined on M, then we know that it is possible to express the vector fields X_α as

$$X_\alpha(x) = \sum_{i=1}^{n} X_\alpha^i(x) \frac{\partial}{\partial x_i}.$$

The corresponding momentum can be written as

$$P_{X_\alpha}(x, p) = \sum_{i=1}^{n} X_\alpha^i(x) P_{\frac{\partial}{\partial x_i}},$$

where $P_{\frac{\partial}{\partial x_i}}$ are the momenta w.r.t. the coordinates vector fields.
Set $p_i = P_{\partial/\partial x_i}$, (x_i, p_i) is a system of local coordinates on the cotangent bundle T^*M, usually called *canonic coordinates.*
We define the matrix $(g_{\alpha\,\delta})_{\alpha,\delta}$ as

$$g_{\alpha\,\delta} = \langle X_\alpha(x), X_\delta(x)\rangle_x$$

and we indicate by $g^{\alpha\,\delta}$ the coefficients of the inverse matrix.
We get so the following explicit formulation for the Hamiltonian H in (1.32):

$$H(x,p) = \frac{1}{2}\sum_{\alpha,\,\delta} g^{\alpha\,\delta}(x)P_\alpha(x,p)P_\delta(x,p). \tag{1.35}$$

Remark 1.65. *In particular, if $\{X_\alpha\}_\alpha$ is an orthonormal system w.r.t. the sub-Riemannian metric $\langle\ ,\ \rangle$, then* (1.35) *can be simplified as*

$$H(x,p) = \frac{1}{2}\sum_{\alpha} P_\alpha^2(x,p), \tag{1.36}$$

or equivalently, using the canonic coordinates, as

$$H(x,p) = \frac{1}{2}\sum_{\alpha} \langle X_\alpha(x), p\rangle^2 . \tag{1.37}$$

By (1.37) we get the following system of first-order differential equations (defined on the cotangent bundle T^*M):

$$\begin{cases} \dot{x}_i = \dfrac{\partial H}{\partial p_i}, \\ \dot{p}_i = -\dfrac{\partial H}{\partial x_i}. \end{cases} \tag{1.38}$$

Definition 1.66. *The Hamiltonian system* (1.38) *is called* equations of the normal geodesics.

Theorem 1.67. *Let $\zeta(t) = \big(\gamma(t), p(t)\big)$ be a solution of the Hamiltonian system* (1.38)*. Then, any short enough arc of $\gamma(t)$ is a geodesic.*

Definition 1.68 (Normal geodesics)**.** *Let $\zeta(t) = \big(\gamma(t), p(t)\big)$ be a solution of the Hamiltonian system* (1.38)*, $\gamma(t)$ is called* (sub-Riemannian) normal geodesic.

Exactly as in the Riemannian case, not all the normal geodesics are (minimizing) geodesics.
Nevertheless in the Riemannian geometries, all the (minimizing) geodesics satisfy the equations of normal geodesics, while in the sub-Riemannian case, there exists geodesics that are not normal geodesics.

Remark 1.69 (Singular geodesics)**.** *In the sub-Riemannian case, there are geodesics that are not normal geodesics, i.e. that cannot be found as solutions of the Hamiltonian system* (1.38)*. Note that, if $\zeta(t)$ solves the system* (1.38)*, then $H(\gamma(t), p(t))$ is constant. Any geodesic $\gamma(t)$ solving*

$$H(\gamma(t), p(t)) = 0,$$

is called singular geodesic*. Such a geodesic does not solves the system* (1.38)*. (Recall that by $p(t)$ we indicate the dual variable of $\gamma(t)$).*

Singular geodesics are usually more difficult to study than normal geodesics. Nevertheless, in the case of contact distributions, all the geodesics are normal. We recall that a distribution $\mathcal{H}$ is a *contact distribution* if and only if it is defined by the kernel of a single one-form η with the property that the restriction on any vector space $\mathcal{H}(p)$, $d\eta\big|_{\mathcal{H}(p)}$, is symplectic (i.e. non degenerate) for any point $p \in M$. E.g. the Heisenberg group is a contact distribution. So the Heisenberg group does not have any singular geodesic. Next we give two examples of distributions admitting singular geodesics.

Example 1.70. *Let us assume $M = \mathbb{R}^3$, we define $\mathcal{H}$ as the distribution spanned at any point $(x, y, z) \in \mathbb{R}^3$ by $X_1 = (1, 0, 0)^t$ and $X_2 = (0, 1-x, x^2)^t$. We look at the horizontal curve $\gamma : [0, T] \to \mathbb{R}^3$ defined as $\gamma(t) = (0, t, 0)$. The associated Hamiltonian is*

$$H((x, y, z), (p_1, p_2, p_3)) = -\frac{1}{2}\big(p_1^2 + (p_2(1-x) + p_3x^2)^2\big).$$

So it is easy to verify that γ does not satisfy the equation of normal geodesics. Nevertheless, if the interval is small enough, then γ is a geodesic (see [72], Sec.2.3).

Example 1.71 (Martinet distribution)**.** *Let be $M = \mathbb{R}^3$ and $X_1 = (1, 0, -y^2)^t$ and $X_2 = (0, 1, 0)$ for $(x, y, z) \in \mathbb{R}^3$. The distribution $\mathcal{H}$ spanned by the vector fields X_1 and X_2 is known as* Martinet distribution*. It is a bracket generating distribution with step equal to 3 (in fact, $[X_1, X_2] = (1, 0, -2y)^t$ and $[[X_1, X_2], X_2] = (0, 0, -2)^t$). The existence of singular geodesics in the Martinet distribution is given in [75], Theorem III.3.4.*

To conclude this subsection, we recall a result of local and global existence for sub-Riemannian geodesics (see e.g. [75]).

Theorem 1.72. *Let M be a smooth manifold and $\mathcal{H}$ a bracket generating distribution. Then*

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

In the Riemannian case, it is also well-known that geodesics are locally unique. Some uniqueness results about sub-Riemannian geodesics can be found in [89] or also in [75]. Nevertheless in sub-Riemannian geometries the geodesics are in general not unique, even locally. We will show that in the next subsection.

1.1.9 The Grušin plane.

In this subsection we study in detail the easiest example of a sub-Riemannian geometry with non-constant-rank: the Grušin plane.

Definition 1.73. *Sub-Riemannian geometries of Grušin-type are defined on $\mathbb{R}^n$ (for $n \geq 2$) by vector fields given, for any $(x, y) \in \mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^k$, as*

$$X_1(x,y) = \partial_{x_1}, ..., X_m(x,y) = \partial_{x_m},\ Y_1(x,y) = |x|^\alpha \partial_{y_1}, ..., Y_k(x,y) = |x|^\alpha \partial_{y_k},$$

with $1 \leq m \leq n-1$, $k = n-m$ and $\alpha > 0$, endowed with the m-dimensional Euclidean metric.

The Grušin plain $\mathbb{G}_2$ corresponds to the case $n = 2$ and $m = \alpha = 1$. More precisely, $\mathbb{G}_2$ is the sub-Riemannian geometry induced on $\mathbb{R}^2$ by the vector fields

$$X = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 \\ x \end{pmatrix}, \quad \text{for } (x,y) \in \mathbb{R}^2.$$

Note that, at the origin $(0,0)$, $\mathrm{Span}(X,Y) = \mathrm{Span}(X) \neq \mathbb{R}^2$.
Nevertheless $[X,Y] = (0,1)^t$, so $\mathrm{Span}(X,Y,[X,Y]) = \mathbb{R}^2$ at any point $(x,y) \in \mathbb{R}^2$. Therefore $\mathbb{G}_2$ satisfies the Hörmander condition with step 2.
The associated sub-Riemannian metric can be explicitly written as

$$ds^2 = dx^2 + \frac{dy^2}{x^2}, \tag{1.39}$$

whenever $x \neq 0$ (see [17]).
A curve across the y-axis has a finite length if and only if its velocity vector is parallel to the x-axis.
In fact, the length of a horizontal curve $\gamma : [0,T] \to \mathbb{R}^2$ w.r.t. the Grušin metric (1.39) can be written as

$$l(\gamma) = \int_0^T \left(\dot{x}(t)^2 + \frac{\dot{y}(t)^2}{x(t)^2} \right)^{1/2} dt, \tag{1.40}$$

with $\gamma(t) = (x(t), y(t))$.
Therefore if $x(t_0) = 0$ for some $t_0 \in [0,T]$, we get $l(\gamma) < +\infty$ if and only if $\dot{y}(t_0) = 0$, i.e. if and only if the velocity vector $\dot{\gamma}(t_0)$ is parallel to the x-axis.

By (1.39) we can define a family of dilations w.r.t. the Grušin metric.

Definition 1.74 (Dilations)**.** *For any $\lambda \in \mathbb{R}$, we define a family of dilations $\delta_\lambda : G_2 \to G_2$ as*

$$\delta_\lambda(x,y) = (\lambda x, \lambda^2 y).$$

Remark 1.75. *For general sub-Riemannian geometries of Grušin-type (Definition 1.73), a family of dilations is given by $\delta_\lambda(x,y) = (\lambda x, \lambda^{\alpha+1} y)$ for $(x,y) \in \mathbb{R}^m \times \mathbb{R}^k$.*

We indicate by $d_{\mathbb{G}_2}(x,y)$ the sub-Riemannian distance defined minimizing (1.40) over all the $\mathbb{G}_2$-horizontal curves, so it is easy to verify that, for any $(x,y), (x',y') \in G_2$ and $\lambda \in \mathbb{R}$, it holds

$$d_{\mathbb{G}_2}\big(\delta_\lambda(x,y), \delta_\lambda(x',y')\big) = |\lambda| d_{\mathbb{G}_2}\big((x,y),(x',y')\big). \tag{1.41}$$

By (1.41) we can deduce deduce the following "exact" estimate.

Theorem 1.76 ([17]). *Let $\mathbb{G}_2$ be the Grušin plane and $d_{\mathbb{G}_2}$ the associated sub-Riemannian distance then, for any $(x, y) \in \mathbb{R}^2$,*

$$\frac{1}{2}\big(|x| + |y|^{\frac{1}{2}}\big) \leq d_{\mathbb{G}_2}\big((0,0),(x,y)\big) \leq 3\big(|x| + |y|^{\frac{1}{2}}\big). \tag{1.42}$$

Remark 1.77. *The norm $\|(x, y)\| = |x| + |y|^{\frac{1}{2}}$ is usually called the homogeneous norm of the Grušin plane.*
Estimate (1.42) *tells that the homogenous norm is equivalent to the Carnot-Carathéodory norm $\|(x, y)\|_C = d_{\mathbb{G}_2}((0,0),(x,y))$.*

Remark 1.78. *Form estimate* (1.42) *it follows that the structure of the Grušin plane is not isotropic. In fact, the ball centered at the origin and with radius $r > 0$ is equivalent to the Euclidean rectangle $(-r, r) \times (-r^2, r^2)$.*

to conclude the study of the Grušin plane, we compute explicitly the geodesics starting from the origin.

Theorem 1.79. *The geodesics starting from the origin in the Grušin plane can be parameterized as*

$$\begin{cases} x(t) = \dfrac{a}{b}\sin(b\,t), \\ y(t) = \dfrac{a^2}{b}\left(\dfrac{t}{2} - \dfrac{\sin(2b\,t)}{4b}\right), \end{cases} \tag{1.43}$$

for any $a \in \mathbb{R}$ and $b \neq 0$, and by

$$\begin{cases} x(t) = at, \\ y(t) = 0, \end{cases} \tag{1.44}$$

for any $a \in \mathbb{R}$.

Proof. Using (1.37) we can write the Hamiltonian associated to $\mathbb{G}_2$:

$$H(x, y, p_1, p_2) = p_1^2 + x^2 p_2^2.$$

Then we get the following equations for the normal geodesics:

$$\begin{cases} \dot{x} = p_1, \\ \dot{y} = x^2 p_2, \end{cases} \tag{1.45}$$

and

$$\begin{cases} \dot{p_1} = -xp_2^2, \\ \dot{p_2} = 0, \end{cases} \tag{1.46}$$

with initial data

$$\begin{cases} x(0) = 0 \\ y(0) = 0 \end{cases} \quad \begin{cases} p_1(0) = a \\ p_2(0) = b \end{cases} \tag{1.47}$$

for any $a, b \in \mathbb{R}$.
$\dot{p_2} = 0$ implies $p_2 = b$. By replacing this in the first equation of system (1.46), we get

$$\dot{p_1} = -b^2 x. \tag{1.48}$$

By deriving (1.48) and using the first equation in (1.45) we find:

$$\ddot{p_1} = -b^2 p_1. \tag{1.49}$$

Assume $b \neq 0$, the solutions of (1.49) are

$$p_1(t) = A\cos(b\,t) + B\sin(b\,t),$$

with A and B suitable real-constants.
By substituting the expression for $p_1(t)$ in (1.48), we get

$$x(t) = -\frac{\dot{p_1}(t)}{b^2} = \frac{A\sin(b\,t) - B\cos(b\,t)}{b}.$$

Assumptions (1.47) imply that $A = a$ and $B = 0$, so

$$x(t) = \frac{a}{b}\sin(b\,t). \tag{1.50}$$

Moreover, by (1.45) we can write

$$\dot{y}(t) = bx^2(t) = \frac{a^2}{b}\sin^2(b\,t). \tag{1.51}$$

By integratining with $y(0) = 0$, we can conclude

$$y(t) = \frac{a^2}{b}\left(\frac{t}{2} - \frac{\sin(2b\,t)}{4b}\right). \tag{1.52}$$

It remains only to solve the case $b = 0$, which means $\dot{y} = 0$, with initial condition $y(0) = 0$: hence $y(t) = 0$.
Similarly $p_1(t) = a$, which implies $\dot{x}(t) = a$ and $x(0) = 0$, then $x(t) = at$. $\square$

Remark 1.80. *Note that all the curves parameterized by* (1.43) *are horizontal. In fact*

$$\dot{\gamma}(t) = a\cos(b\,t)X(\gamma(t)) + a\sin(b\,t)Y(\gamma(t)).$$

Remark 1.81. *Using geodesics* (1.43)*, we can parametrize the unit ball of the Grušin plane. If we assume that the geodesics are all parameterized in the unit interval* $[0,1]$*, then* $l(\gamma) = 1$ *whenever* $\|X\gamma(t)\|^2 = a^2 = 1$*. So the unit ball can be written as*

$$\begin{cases} x(t) = \pm\,\dfrac{\sin(b\,t)}{b}, \\ y(t) = \dfrac{2bt - \sin(2b\,t)}{4b^2}, \end{cases}$$

with $0 < b \leq 2\pi$ *and* $0 \leq t < 1$*.*

Next we show that geodesics starting from the origin are not locally unique in the Grušin plane.
Look at a point $(0,h) \in \mathbb{R}^2$ with $h \neq 0$, we write the geodesics joining the origin to that point. As usually, we assume that the geodesics are parameterized in $[0,1]$. Therefore

$$0 = x(1) = \frac{a}{b}\sin(b)$$

and

$$h = y(1) = \frac{a}{b}\left(\frac{1}{2} - \frac{\sin(2b\,t)}{4b}\right).$$

Hence we find $b = k\pi$ and $a = \pm\sqrt{2k\pi h}$ for $k \in \mathbb{N}$, which implies

$$\gamma_k^{\pm}(t) = \begin{cases} x_k(t) = \pm\sqrt{\dfrac{2h}{k\pi}}\sin(k\pi\,t), \\ y_k(t) = h\left(t - \dfrac{\sin(2k\pi\,t)}{2k\pi}\right). \end{cases} \tag{1.53}$$

All the curves parameterized by (1.53) (see Fig. 1.4), satisfy the equations of normal geodesics in $\mathbb{G}_2$.

Note that $l(\gamma_k^{\pm}) = \sqrt{2k\pi h}$, so $\gamma_k^{\pm}(t)$ minimize the length-functional only when $k = 1$. This means that there exist two different (minimizing) geodesics joining the origin to the point $(0,h)$, for $h \neq 0$, i.e. γ_1^+ and γ_1^-. This behavior is very different from the Riemannian case.

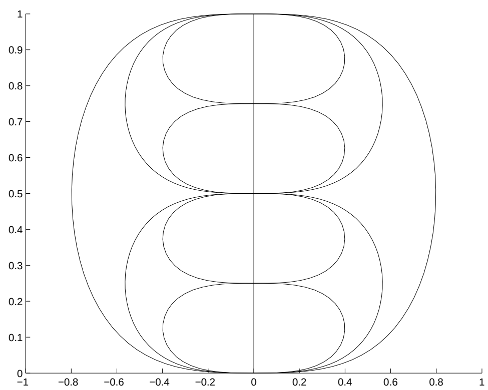


Figure 1.4: Some of the curves given in (1.53): the case $k = 1, 2, 4$.

1.2 Carnot groups.

1.2.1 Nilpotent Lie groups.

Carnot groups are particular nilpotent Lie groups. So we first need to recall some basic notions about Lie groups.
The literature about Lie groups is very large, for example, one can see [1, 29, 79]. In particular we refer to a paper of J. Heinonen [52], where a very short but clear treatment on the Carnot groups can be found. We start giving the definition of Lie group.

Definition 1.82. *A* Lie group *is a smooth manifold with a group structure such that the inner operation and the inversion are smooth maps.*

We indicate by $\mathbb{G}$ a Lie group and by $\cdot : \mathbb{G} \times \mathbb{G} \to \mathbb{G}$, $(g, h) \longmapsto gh$ and $i : \mathbb{G} \to \mathbb{G}$, $g \longmapsto g^{-1}$ the inner operation and the inversion, respectively. Note that, when we require the smoothness of the inversion, we mean that it have to be smooth w.r.t. the product-topology induced by $\mathbb{G}$ on $\mathbb{G} \times \mathbb{G}$.

Definition 1.83. *Let $\mathbb{G}$ be a Lie group, the associated* Lie algebra g *is the set of all the left-invariant vector fields defined on $\mathbb{G}$.*

To get a Lie algebra according to the classic definition, we need to show that g is a vector space where an antisymmetric bilinear form, satisfying the Jacobi identity, is defined on. We recall that a bilinear form [,] defined on a vector space V satisfies the Jacobi identity, if and only if, for any $X, Y, Z \in V$ the following identity holds:

$$[[X,Y],Z]+[[Y,Z],X]+[[Z,X],Y]=0. \tag{1.54}$$

Proposition 1.84. *The set* g *of all the left-invariant vector fields (see Definition 1.83) is a Lie algebra.*

Proof. Let e be the unit-element of the group $\mathbb{G}$, we can identify the Lie algebra g with the tangent space at the point e, that we indicate by $T_{\mathrm{e}}\mathbb{G}$, associating to a vector $X \in T_{\mathrm{e}}\mathbb{G}$, the left-invariant vector field $X_g := L_g X$, where L_g is the left translation w.r.t. the element $g \in \mathbb{G}$. Therefore we can see the Lie algebra g like a vector space and $\dim \mathrm{g} = \dim T_{\mathrm{e}}\mathbb{G} = \dim \mathbb{G}$.
It is trivial to check that the bracket $[X,Y] = XY - YX$ defines on $T_{\mathrm{e}}\mathbb{G}$ an antisymmetric bilinear form, which satisfies the Jacobi identity (1.54).
So we can conclude that g has the structure of a Lie algebra. ☐

Proposition 1.85. *If $X \in$ g, 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*

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

Definition 1.86. *Let $\mathbb{G}$ be a Lie group, the* exponential map *is the smooth function given by*

$$\begin{aligned} \exp : \mathrm{g} &\longrightarrow G \\ X &\longmapsto \Phi_X(1). \end{aligned} \tag{1.55}$$

We recall some basic properties of the exponential map.

Theorem 1.87 ([91])**.** *Let $\mathbb{G}$ be a Lie group and* g *the associated Lie algebra, the exponential map defined by* (1.55) *is an analytic function. Moreover, it is a diffeomorphism in a neighborhood of the origin.*

Remark 1.88. *The exponential map satisfies the following properties:*

1. $\exp\big((t+s)X\big) = \exp tX \exp sX$,

2. $\exp(-tX) = (\exp tX)^{-1}$,

3. $\exp(X+Y) \neq \exp X + \exp Y$.

The next result is very useful, since gives a formula in order to explicitly compute many exponential maps.

Proposition 1.89 (Campbell-Hausdorff formula, [52])**.** *For any* $X, Y \in \mathrm{g}$ *the following formula holds:*

$$\exp X \exp Y = \exp\left(X + Y + \frac{1}{2}[X,Y] + R(X,Y)\right), \tag{1.56}$$

where $R(X,Y)$ *is a remainder polynomial for a nilpotent Lie group, written by all the possible non-vanishing* k*-brackets of* X *and* Y *with* $k \geq 2$*.*

Example 1.90 (Heisenberg group)**.** *The Heisenberg group is a Lie group and the associated distribution is bracket generating with step 2.*
Therefore $R(X,Y) = 0$ *for any* $X, Y \in \mathrm{g}$*, and so formula* (1.56) *can be rewritten as*

$$\exp X \exp Y = \exp\left(X + Y + \frac{1}{2}XY - \frac{1}{2}YX\right). \tag{1.57}$$

Example 1.91 (Group of matrices)**.** *If* $\mathbb{G}$ *is a group of matrices, the exponential map is the usual exponential matrix.*

Definition 1.92 (Nilpotent groups)**.** *A group* $\mathbb{G}$ *is* nilpotent *if and only if its central series defined as*

$$\begin{cases} \mathbb{G}^{(1)} = \mathbb{G}, \\ \mathbb{G}^{(i+1)} = [\mathbb{G}, \mathbb{G}^{(i)}] = \{ghg^{-1}h^{-1} \mid g \in \mathbb{G}, h \in \mathbb{G}^{(i)}\}, \ i \geq 1, \end{cases} \tag{1.58}$$

is finite, i.e. if and only if there exists $k \in \mathbb{N}$ *such that*

$$\mathbb{G}^{(k+1)} = \{0\} \neq \mathbb{G}^{(k)}.$$

Moreover, in this case we say that $\mathbb{G}$ *is nilpotent with step equal to* k*.*

Remark 1.93. *A group is nilpotent with step 1 if and only if it is an Abelian group.*

Definition 1.94 (Nilpotent Lie algebra)**.** *A Lie algebra* g *is nilpotent with step equal to k if and only if, setting*

$$\begin{cases} \mathrm{g}^{(1)} = \mathrm{g}, \\ \mathrm{g}^{(i+1)} = [\mathrm{g}, \mathrm{g}^{(i)}] = \{[X,Y] |\, X \in \mathrm{g}, Y \in \mathrm{g}^{(i)}\},\ i \geq 1, \end{cases} \tag{1.59}$$

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

Remark 1.95. *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 1.96 (Nilpotent groups and sub-Riemannian geometries)**.** *Remark that 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.*

The following result is one of the main properties of the exponential map.

Proposition 1.97 ([52])**.** *Let* $\mathbb{G}$ *be a Lie group, nilpotent and simply connected, then the exponential map* $\exp : \mathrm{g} \to \mathbb{G}$ *is a diffeomorphism.*
Conversely, if g *is a Lie algebra with finite dimension, then there exists a unique Lie group* $\mathbb{G}$*, simply connected, such that* g *is its associated Lie algebra.*

Remark 1.98. *If* $\mathbb{G}$ *is a Lie group, nilpotent and simply connected, then it is isomorphic to a closed unipotent subgroup of* $GL(n,\mathbb{R})$*, where* $GL(n,\mathbb{R})$ *is the set of matrices with eigenvalues equal to 1. Then, there exists a system of local coordinates where we can express it is an upper triangular matrix with all the diagonal-elements equal to 1.*

Using the Campbell-Hausdorff formula and the exponential map, we get an easy interpretation for Lie groups.

Corollary 1.99. *Let* $\mathbb{G}$ *be a Lie group, simply connected and nilpotent, then by the exponential map we can identify* $\mathbb{G}$ *with the associated Lie algebra* g *and, consequently, with* $\mathbb{R}^n$ *endowed with the polynomial operation, given by the Campbell-Hausdorff formula, i.e.* $X * Y := X + Y + \frac{1}{2}[X,Y] + R(X,Y)$.

We are so able to give a geometric characterization for Abelian groups.

Remark 1.100. *A group $\mathbb{G}$ is Abelian if and only if it is isomorphic to $\mathbb{R}^n$ endowed with the usual (Euclidean) structure of group.*

The coordinates defined by using the exponential map are usually called *canonic coordinates.*

1.2.2 Calculus on Carnot groups.

In this subsection we study Carnot groups, refering to [52].

Definition 1.101 (Carnot group)**.** *A* Carnot group *is a Lie group, nilpotent and simply connected, whose Lie algebra* g *admits a stratification, i.e. there exist $V_1, ...V_k$ vector spaces such that*

$$\mathrm{g} = V_1 \oplus ... \oplus V_k.$$

Moreover g *is endowed with the Carnot-Carathéodory metric associated to V_1.*

A Lie group as in Definition 1.101 is also called *homogeneous group* or *stratified group.* Moreover, we say that V_1 generates the Lie algebra g, while the vector spaces V_i for $i \geq 2$ are called *slices* of g.
We now introduce a family of dilations on any Carnot group. Dilations imply that any Carnot group is self-similar.

Definition 1.102. *For $\lambda > 0$, a family of* dilations *on* g *is a family of smooth maps $\delta_\lambda : \mathrm{g} \to \mathrm{g}$ defined as dilations*

$$\delta_\lambda(X) = \lambda^i X, \quad \textit{whenever } X \in V_i. \tag{1.60}$$

Applying the exponential map, we can rewritten the dilations given by (1.60) as automorphisms of $\mathbb{G}$, by setting $\widetilde{\delta_\lambda} = \exp^{-1} \delta_\lambda \exp$. For sake of simplicity, we still indicate $\widetilde{\delta_\lambda}$ simply by δ_λ.
Note that we will indicate by $d_C(x, y)$ the Carnot-Carathéodory distance associated to a Carnot group $\mathbb{G}$.
Since a Carnot-Carathéodory distance is defined minimizing the length-functional over all the V_1-horizontal curves (Definition 1.14), it follows immediately that

$$d_C(\delta_\lambda(x), \delta_\lambda(y)) = \lambda d_C(x, y), \tag{1.61}$$

for any $x, y \in \mathbb{G}$ and $\lambda > 0$.
Let $\mathbb{G}$ be a Carnot group and $d_C(x,y)$ the associated Carnot-Carathéodory distance. Using the family of dilations δ_λ, we can define a norm on $\mathbb{G}$, setting

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

Moreover it is immediate to verify that:

$$\|x\|_C = 0 \quad \Longleftrightarrow \quad x = 0,$$

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

It is possible to introduce a new norm on $\mathbb{G}$ equivalent to the Carnot-Carathéodory norm $\| \quad \|_C$ but easier to compute, using the stratification-property characterizing a Carnot group. In fact, we can write any point $x \in \mathbb{G}$ as $x = (x_1, ..., x_k)$ where $x_i \in V_i$.

Definition 1.103. *Let $\mathbb{G}$ be a Carnot group with stratification $V_1, ..., V_k$, the* homogeneous norm *is defined as*

$$\|x\| := \left(\sum_{i=1}^{k} |x_i|^{\frac{2k!}{i}} \right)^{\frac{1}{2k!}}, \tag{1.62}$$

where $|x_i|$ is the usual r-dimensional Euclidean norm defined on the space vector V_i with $r = \mathrm{dim} V_i$.

Example 1.104 (Homogeneous norm in the Heisenberg group)**.** *Since the n-dimensional Heisenberg group $\mathbb{H}^n$ is a Carnot group with $k = 2$, by* (1.62) *we get*

$$\|(z,t)\| = \left(|z|^4 + t^2\right)^{\frac{1}{4}}, \quad z \in \mathbb{R}^{2n},\ t \in \mathbb{R}. \tag{1.63}$$

Note that

$$\|\delta_\lambda(x)\| = \lambda \|x\| , \tag{1.64}$$

and

$$\left\|x^{-1}\right\| = \|x\| , \tag{1.65}$$

where by x^{-1} we indicate the inverse element in the group $\mathbb{G}$.
Using the homogeneous norm, we can define a left-invariant distance on $\mathbb{G}$, simply by taking for any $x, y \in \mathbb{G}$

$$d(x,y) = \left\|x^{-1}y\right\| . \tag{1.66}$$

The function (1.66) is a distance on $\mathbb{G}$ and moreover it is left-invariant.
In fact, for any $h \in \mathbb{G}$

$$d(hx, hy) = \left\|(hx)^{-1}hy\right\| = \left\|x^{-1}h^{-1}hy\right\| = \left\|x^{-1}y\right\| = d(x, y).$$

W.r.t. the dilations, the new distance defined in (1.66) has the same behavior as the original Carnot-Carathéodory distance, i.e. for any $x, y \in \mathbb{G}$

$$d(\delta_\lambda(x), \delta_\lambda(y)) = \lambda d(x, y). \tag{1.67}$$

We conclude by remarking that all the distances defined on a Carnot group and satisfying (1.67) are equivalent.
The advantage of using the homogeneous distance instead of the Carnot-Carathéodory distance is that the homogeneous distance is much simpler to compute.
The main problem is that in some cases the homogeneous distance does not satisfy a triangular inequality and so it does not induce a metric space on $\mathbb{G}$.

Note that in the particular case of the Heisenberg group the homogeneous norm (1.63) does satisfy the standard triangular inequality.

Remark 1.105 (Non-Euclidean nature of Carnot groups)**.** *There is a very famous result by Stephen Semmes ([85]), which shows that a bi-Lipschitz embedding of the Heisenberg group into an Euclidean space does not exist. The non-Euclidean behavior of the Carnot spaces is related to the non-commutativity of the associated Lie algebras. In fact, if we assume that a bi-Lipschitz emending exists, then the corresponding differential is an injective application between the two corresponding tangent spaces.*
By the identification of a tangent space with the Lie algebra and using the exponential map, we would find an isomorphism between a non-Abelian group and an Abelian one, which gives the contradiction.

1.3 The Heisenberg group.

The Heisenberg group is the most famous and the most studied sub-Riemannian geometry.

The 1-dimensional Heisenberg algebra is the Lie algebra spanned by a basis X, Y, Z satisfying the bracket-relations:

$$[X, Y] = Z, \quad [X, Z] = [Y, Z] = 0.$$

The origin of the name goes back to the fact that Heisenberg wrote down the previous bracket-relations in his works about quantum mechanics.
x Some information on "Heisenberg's quantum mechanics theory" can be found in [75] (Chapter 13).

Next we give a formal definition of the n-dimensional Heisenberg group, which is usually indicate by $\mathbb{H}^n$.

Definition 1.106. $\mathbb{H}^n$ *is the* $2n+1$*-dimensional Carnot group defined on* $\mathbb{R}^{2n+1} = \mathbb{C}^n \times \mathbb{R}$*, by the following composition law*

$$(z, t) \cdot (z', t') = \left(z + z', t + t' + \frac{1}{2} \mathrm{Im} \left\langle z, z' \right\rangle \right),$$

where $\mathrm{Im} \langle z, z' \rangle$ *is the imaginary part of the standard hermitian inner product of* $\mathbb{C}^n$ *between* z *and* z'*.*

In the next example we show that the operation defined above is non-Abelian.

Example 1.107. *Note that* $(1, 0, 0) \cdot (0, 1, 0) = (1, 1, 1/2)$*, while* $(0, 1, 0) \cdot (1, 0, 0) = (1, 1, -1/2)$*.*

The Heisenberg group is a bracket generating sub-Riemannian geometry with step 2. Next we give two different definitions of this particular Carnot group and we show that the two given definitions are indeed equivalent.

1.3.1 The polarized Heisenberg group.

Definition 1.108 (The polarized Heisenberg group). *We call* 1-dimensional polarized Heisenberg group *the sub-Riemannian geometry spanned on* $\mathbb{R}^3$ *by the vector fields*

$$X = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 \\ 1 \\ -x \end{pmatrix} \tag{1.68}$$

endowed with the usual 2-dimensional Euclidean metric.

The bracket between X and Y gives the third dimension of the space: for any smooth function $f : \mathbb{R}^3 \to \mathbb{R}$,

$$Zf := [X,Y]f = X(Yf) - Y(Xf) = f_{xy} - f_z - xf_{xz} - f_{yx} + xf_{zx} = -f_z,$$

so that $Z = (0,0,-1)^t$.
This means that $\text{Span}\{X(p), Y(p), Z(p)\} = \mathbb{R}^3$, for any $p = (x,y,z) \in \mathbb{R}^3$, which means that the distribution spanned by the vector fields (1.68) is bracket generating with step 2.

Remark 1.109 ([17]). *The associated sub-Riemannian metric can be written as*

$$ds^2 = dx^2 + dy^2 + \frac{dz^2}{x^2}. \tag{1.69}$$

By the family of dilations defined on this Carnot group, it is possible to deduce the following estimate for the Carnot-Carathéodory.

Theorem 1.110. *Let* $d(x,y)$ *be the Carnot-Carathéodory distance defined by* (1.69)*, then for any* $(x,y,z) \in \mathbb{R}^3$

$$\frac{1}{3}(|x| + |y| + |z|^{\frac{1}{2}}) \leq d((0,0,0),(x,y,z)) \leq 4(|x| + |y| + |z|^{\frac{1}{2}}). \tag{1.70}$$

For a proof of the theorem see [17].

Remark 1.111. *Estimate* (1.70) *shows that the sub-Riemannian ball centered at the origin with radius* $r > 0$ *is equivalent to the Euclidean hyper-rectangle* $(-r,r) \times (-r,r) \times (-r^2, r^2)$.
This means that the Heisenberg geometry is an anisotropic geometry.

By using the geodesic equations (1.38), we can compute the geodesics starting from the origin. Applying (1.37), we find that the associated Hamiltonian is given by

$$H(x,y,z,p_1,p_2,p_3) = \frac{1}{2}\big(p_1^2 + (p_2 - xp_3)^2\big).$$

Recall that in the Heisenberg group there are not singular geodesics, the geodesics in the polarized Heisenberg group are solutions of

$$\begin{cases} \dot{x} = p_1 \\ \dot{y} = (p_2 - xp_3) \\ \dot{z} = -x(p_2 - xp_3) \end{cases} \tag{1.71}$$

and

$$\begin{cases} \dot{p_1} = p_3(p_2 - xp_3) \\ \dot{p_2} = 0 \\ \dot{p_3} = 0 \end{cases} \tag{1.72}$$

with initial data

$$\begin{cases} x(0) = 0 \\ y(0) = 0 \\ z(0) = 0 \end{cases} \qquad \begin{cases} p_1(0) = a \\ p_2(0) = b \\ p_3(0) = c \end{cases} \tag{1.73}$$

for $a, b, c \in \mathbb{R}$.
By (1.72) we get

$$p_2 = b, \quad p_3 = c.$$

So the other equations can be rewritten as

$$\begin{cases} \dot{x} = p_1, \\ \dot{y} = b - cx, \\ \dot{z} = -x(b - cx), \\ \dot{p_1} = c(b - cx). \end{cases} \tag{1.74}$$

Deriving the first-equation in (1.74) and replacing that in the last equation, we find:

$$\ddot{x} + c^2 x - bc = 0. \tag{1.75}$$

Assuming $c \neq 0$, the solutions of (1.75) with initial condition $x(0) = 0$ and $\dot{x}(0) = p_1(0) = a$ are

$$x(t) = \frac{a}{c}\sin(c\,t) + \frac{b}{c}(1 - \cos(c\,t)). \tag{1.76}$$

From (1.76), it follows a first-order equation for the second-component of the geodesics, more precisely

$$\dot{y} = b - cx = b\cos(c\,t) - a\sin(c\,t).$$

By integrating with $y(0) = 0$, we get

$$y(t) = \frac{b}{c}\sin(c\,t) - \frac{a}{c}(1 - \cos(c\,t)).$$

By the knowledge of $x(t)$, we can also deduce an equation for the third-component of the geodesics, i.e.

$$\begin{aligned}\dot{z} = -x(b - cx) &= -\frac{1}{c}(a\sin(c\,t) + b - b\cos(c\,t))(b - a\sin(c\,t) - b + b\cos(c\,t)) \\ &= \frac{a^2}{c}\sin^2(c\,t) + \frac{b^2}{c}\cos^2(c\,t) - \frac{ab}{c}\sin(2c\,t) + \frac{ab}{c}\sin(c\,t) - \frac{b^2}{c}\cos(c\,t).\end{aligned}$$

By integrating the previous equation with $z(0) = 0$, we find

$$z(t) = \frac{a^2 + b^2}{2c}t - \frac{a^2 - b^2}{4c^2}\sin(2c\,t) + \frac{ab}{2c^2}\cos(2c\,t) - \frac{ab}{c^2}\cos(c\,t) - \frac{b^2}{c^2}\sin(c\,t) + \frac{ab}{2c^2}.$$

We remain to solve the case $c = 0$. If $c = 0$ the system (1.74) can be rewritten as

$$\begin{cases} \dot{x} = p_1, \\ \dot{y} = b, \\ \dot{z} = -bx, \\ \dot{p_1} = 0. \end{cases}$$

It is easy to see that $p_1(t) = a$ and so

$$\begin{cases} x(t) = at, \\ y(t) = bt. \end{cases}$$

The third-component is deduced integrating $\dot{z} = -abt$, i.e.

$$z(t) = -\frac{ab}{2}t^2.$$

To sum up, we have proved the following result.

Theorem 1.112. *The geodesics of the polarized Heisenberg group starting from the origin are curves* $\gamma(t) = (\gamma_1(t), \gamma_2(t), \gamma_3(t))$ *parameterized by*

$$\begin{cases} \gamma_1(t) = \dfrac{a}{c}\sin(c\,t) + \dfrac{b}{c}(1-\cos(c\,t)), \\ \gamma_2(t) = \dfrac{b}{c}\sin(c\,t) - \dfrac{a}{c}(1-\cos(c\,t)), \\ \gamma_3(t) = \dfrac{a^2+b^2}{2c}t - \dfrac{a^2-b^2}{4c^2}\sin(2c\,t) + \dfrac{ab}{2c^2}\cos(2c\,t) - \dfrac{ab}{c^2}\cos(c\,t) - \dfrac{b^2}{c^2}\sin(c\,t) + \dfrac{ab}{2c^2}, \end{cases} \tag{1.77}$$

for any $a, b \in \mathbb{R}$ *and* $c \neq 0$, *and*

$$\begin{cases} \gamma_1(t) = at, \\ \gamma_2(t) = bt, \\ \gamma_3(t) = -\dfrac{ab}{2}t^2, \end{cases} \tag{1.78}$$

for any $a, b \in \mathbb{R}$.

Remark 1.113. *The geodesics given in Theorem 1.112 are horizontal: in fact* $\dot{\gamma_3}(t) = -\gamma_1(t)\dot{\gamma_2}(t)$ *for any* $t > 0$.

1.3.2 The canonical Heisenberg group.

Definition 1.114 (Canonical Heisenberg group)**.** *We call* 1-dimensional exponential Heisenberg group *or also* canonical Heisenberg group, *the sub-Riemannian geometry induced by the vector fields*

$$X = \begin{pmatrix} 1 \\ 0 \\ -\frac{y}{2} \end{pmatrix} \quad Y = \begin{pmatrix} 0 \\ 1 \\ \frac{x}{2} \end{pmatrix} \tag{1.79}$$

endowed with the usual 2-dimensional Euclidean metric.

Also in this case it is easy to note that the distribution is bracket generating with step 2, since $Z = [X, Y] = (0, 0, 1)^t$.

Also for the canonical Heisenberg group we compute the geodesics. The associated Hamiltonian is

$$H(x, y, z, p_1, p_2, p_3) = \left(p_1 - \frac{y}{2}p_3\right)^2 + \left(p_2 + \frac{x}{2}p_3\right)^2. \tag{1.80}$$

The equations of the geodesics are

$$\begin{cases} \dot{x} = p_1 - \dfrac{y}{2}p_3, \\ \dot{y} = p_2 + \dfrac{x}{2}p_3, \\ \dot{z} = \dfrac{x}{2}p_2 - \dfrac{y}{2}p_1 + \dfrac{1}{4}(x^2+y^2)p_3, \end{cases} \tag{1.81}$$

and

$$\begin{cases} \dot{p_1} = -\dfrac{p_3}{2}\left(p_2 + \dfrac{x}{2}p_3\right), \\ \dot{p_2} = \dfrac{p_3}{2}\left(p_1 - \dfrac{y}{2}p_3\right), \\ \dot{p_3} = 0, \end{cases} \tag{1.82}$$

with initial data (1.73).

It is immediate that $p_3 = c$, hence

$$\begin{cases} \dot{p_1} = -\dfrac{c}{2}\left(p_2 + \dfrac{c}{2}x\right) = -\dfrac{c}{2}\dot{y}, \\ \dot{p_2} = \dfrac{c}{2}\left(p_1 - \dfrac{c}{2}y\right) = \dfrac{c}{2}\dot{x}. \end{cases} \tag{1.83}$$

By deriving (1.81) and using (1.83), we can deduce

$$\begin{cases} \ddot{x} + c\dot{y} = 0, \\ \ddot{y} - c\dot{x} = 0. \end{cases} \tag{1.84}$$

We assume $c \neq 0$ and, for sake of simplicity, we set $p = \dot{x}$ and $q = \dot{y}$. So we have to study the first-order system

$$\begin{cases} \dot{p} + cq = 0, \\ \dot{q} - cp = 0. \end{cases} \tag{1.85}$$

By the first-equation, we find $q = -\dot{p}/c$, which implies $\dot{q} = -\ddot{p}/c$. Replacing this in the second-equation of (1.85), we find

$$\ddot{p} + c^2 p = 0. \tag{1.86}$$

The solutions of (1.86) are

$$p(t) = k_1 \sin(c\,t) + k_2 \cos(c\,t),$$

with k_1 and k_2 suitable constants. Since $q = -\dot{p}/c$, then

$$q(t) = -k_1 \cos(c\,t) + k_2 \sin(c\,t).$$

Using the initial data (1.73), it follows that $a = p_1(0) = \dot{x}(0) = p(0)$ and $b = q(0)$, so $k_1 = -b$ and $k_2 = a$.
To sum up, the solutions of system (1.85) are given by

$$\begin{cases} \dot{x}(t) = -b\sin(c\,t) + a\cos(c\,t), \\ \dot{y}(t) = b\cos(c\,t) + a\sin(c\,t). \end{cases} \tag{1.87}$$

Simply integrating, we can conclude

$$\begin{cases} x(t) = \dfrac{a}{c}\sin(c\,t) - \dfrac{b}{c}(1 - \cos(c\,t)), \\ y(t) = \dfrac{b}{c}\sin(c\,t) + \dfrac{a}{c}(1 - \cos(c\,t)). \end{cases} \tag{1.88}$$

To find the third-component of the normal geodesics is not difficult. In fact, by (1.88) we can write

$$\dot{z}(t) = \frac{a^2 + b^2}{2c}(1 - \cos(c\,t)), \tag{1.89}$$

with $z(0) = 0$, which means

$$z(t) = \frac{a^2 + b^2}{2c^2}(ct - \sin(c\,t)). \tag{1.90}$$

It remains only to solve (1.83) when $c = 0$. In this case the two first-components are $x(t) = at$ and $y(t) = bt$. So the equation for the third-component can be written as

$$\dot{z} = ab\frac{t}{2} - ab\frac{t}{2} = 0,$$

with vanishing initial condition $z(0) = 0$, which means $z(t) = 0$. Therefore if $c = 0$ we have found that the usual Euclidean lines are geodesics.
In short, we have proved the following result.

Theorem 1.115. *The geodesics of the canonical Heisenberg group starting from the origin can be parameterized by*

$$\begin{cases} \gamma_1(t) = \dfrac{a}{c}\sin(c\,t) - \dfrac{b}{c}(1-\cos(c\,t)), \\ \gamma_2(t) = \dfrac{b}{c}\sin(c\,t) + \dfrac{a}{c}(1-\cos(c\,t)), \\ \gamma_3(t) = \dfrac{a^2+b^2}{2c^2}(ct - \sin(c\,t)), \end{cases} \tag{1.91}$$

whenever $c \neq 0$, *and by*

$$\begin{cases} \gamma_1(t) = at, \\ \gamma_2(t) = bt, \\ \gamma_2(t) = 0, \end{cases} \tag{1.92}$$

for any $a, b \in \mathbb{R}$.

Remark 1.116. *It is immediate that all the previous curves are horizontal and they are parameterized by arc-length whenever* $a^2 + b^2 = 1$.

By the previous remark, we can give a parametrization for the unit ball in the Heisenberg group. Looking at the geodesics parametrized by arc-length, we can set $a = \cos\phi$ and $b = \sin\phi$. The unit ball in the Heisenberg group can be written as

$$\begin{cases} x(t) = \dfrac{\cos\phi}{c}\sin(c\,t) - \dfrac{\sin\phi}{c}(1-\cos(c\,t)), \\ y(t) = \dfrac{\sin\phi}{c}\sin(c\,t) + \dfrac{\cos\phi}{c}(1-\cos(c\,t)), \\ z(t) = \dfrac{ct - \sin(c\,t)}{2c^2}, \end{cases}$$

with $0 \leq t < 1$, $0 \leq \phi \leq 2\pi$ and $-2\pi \leq c \leq 2\pi$.

Remark 1.117. *To show Proposition 1.1, we only need to observe that the projections of the Heisenberg geodesics* (1.91) *are circles, whenever they are parameterized by arc-length (Fig. 1.6).*

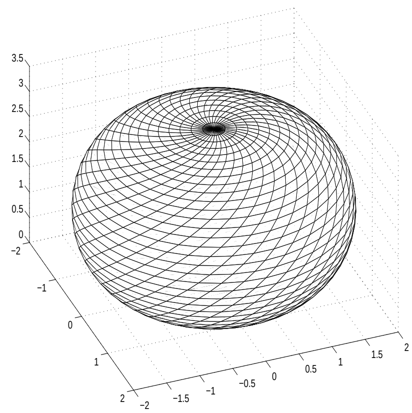


Figure 1.5: `Unit ball of the canonical Heisenberg group.`

1.3.3 Equivalence of the two definitions.

In this subsection we want to show that the polarized Heisenberg group and the canonical Heisenberg group are the same Lie group, i.e. there exists an isomorphism between which respects the topology. Later we conclude studying the behavior of the geodesics via this isomorphism.

To write the isomorphism, we are going to use the exponential maps. Recall that the exponential map goes from a Lie algebra to the original Lie group. Moreover, for any Lie group $\mathbb{G}$, the Lie algebra is isomorphic to the tangent space at the unit element, so the exponential map can be seen as an application from the tangent space $T_e\mathbb{G}$ to $\mathbb{G}$.
Since $\mathbb{G} = \mathbb{R}^n$, the tangent space at any point is the same $\mathbb{R}^n$. Therefore, the exponential map can be interpreted as a change of coordinates in $\mathbb{R}^n$.
Let be $n = 3$, we indicate by $T_e\mathbb{R}^3$, $\mathbb{R}^3$ as tangent space at some point of $\mathbb{R}^3$ (note that in this case the point is meanless).
Fixed a basis X, Y, Z for the tangent space $T_e\mathbb{R}^3$, we consider the function of coordinates defined as

$$\theta(\boldsymbol{v}) := (\alpha, \beta, \gamma),$$

for any $\boldsymbol{v} = \alpha X + \beta Y + \gamma Z \in T_e\mathbb{R}^3$.
Note that $\theta(\boldsymbol{v})$ is an element of $\mathbb{R}^3$ read w.r.t. the standard basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_2$.

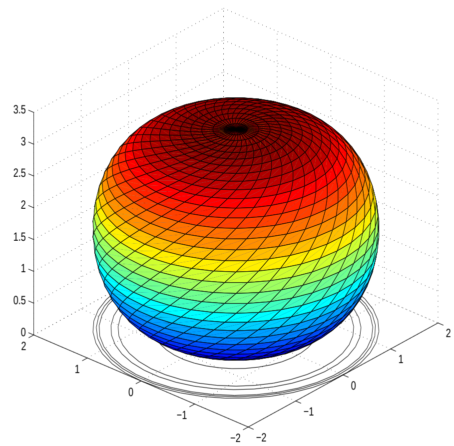


Figure 1.6: `Heisenberg ball and the corresponding projections on the horizontal plane.`

In short, we have

$$\begin{tikzcd} \mathbb{R}^3 \arrow[dd, bend right=40, "\widehat{\exp}"'] \\ T_e\mathbb{R}^3 \arrow[u, "\theta"'] \arrow[d, "\exp"] \\ \mathbb{R}^3 \end{tikzcd}$$

where exp is the usual exponential map.
When we interpret the exponential map as a change of coordinates in $\mathbb{R}^3$, we are indeed looking at the map $\widehat{\exp} := \exp \circ \theta^{-1}$.
Since θ is a linear application between two vector spaces, then $\widehat{\exp}$ has the same regularity-properties as the original exponential map: in particular, it is analytic. Moreover, this map is a diffeomorphism near the origin (Theorem 1.87). From now on, we identify the two notations and we will use exp to mean $\widehat{\exp}$.
We now indicate by $\mathbb{G}_1$ the canonical Heisenberg group (Definition 1.114) and by $\mathbb{G}_2$ the polarized Heisenberg group (Definition 1.108) and we use the indexes 1 and 2 in order to express the corresponding objects, for example $\exp_i$ is the exponential map of the group $\mathbb{G}_i$, for $i = 1, 2$.
Note that $\mathfrak{g}_1 = \mathfrak{g}_2 = \mathbb{R}^3$, so we can define the following application F between

the two groups:

$$
\begin{tikzcd}
& \mathbb{R}^3 \arrow[dl, "\exp_1"'] \arrow[dr, "\exp_2"] & \\
G_1 = \mathbb{R}^3 \arrow[rr, "F"] & & G_2 = \mathbb{R}^3
\end{tikzcd}
$$

We know that $F := \exp_2 \circ \exp_1^{-1}$ is a local diffeomorphism. Now we want to show explicitly that it is a isomorphism (which implies directly that teh function respects the topology). Therefore we only need to prove that:

- F is invertible,
- F is an homomorphism, i.e. for any $p, q \in G_1$,

  $$F(\pi_1(p,q)) = \pi_2(F(p), F(q)),$$

  where π_i is the composition law for the group $\mathbb{G}_i$ with $i = 1, 2$.

The first step consists in computing the two exponential maps.
Remember that the basis associated to $\mathbb{G}_1$ is given by the vector fields

$$X = \begin{pmatrix} 1 \\ 0 \\ -\frac{y}{2} \end{pmatrix}, \quad Y = \begin{pmatrix} 0 \\ 1 \\ \frac{x}{2} \end{pmatrix}, \quad Z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{1.93}$$

To write the exponential map $\exp_1$, we need to calculate the vector flux in the origin at the time 1. Therefore, fixed $\alpha, \beta, \delta \in \mathbb{R}$, we must solve the following first-order differential system:

$$\begin{cases} \dot{\gamma_1}(t) = \alpha, \\ \dot{\gamma_2}(t) = \beta, \\ \dot{\gamma_3}(t) = \dfrac{\beta\gamma_1}{2} - \dfrac{\alpha\gamma_2}{2} + \delta, \end{cases} \tag{1.94}$$

with initial data $\gamma_i(0) = 0$ for $i = 1, 2, 3$.
By the two first-equations of (1.94), we get trivially $\gamma_1(t) = \alpha t$ and $\gamma_2(t) = \beta t$. So the third-equation becomes

$$\dot{\gamma_3}(t) = \frac{\alpha\beta}{2}t - \frac{\alpha\beta}{2}t + \delta = \delta.$$

Then the vector flux is $\gamma(t) = (\alpha t, \beta t, \delta t)$.
Hence the exponential map is

$$\exp_1(\alpha, \beta, \delta) = (\alpha, \beta, \delta),$$

for any $(\alpha, \beta, \delta) \in \mathbb{R}^3$.
In particular $\exp_1$ is invertible and $\exp_1^{-1} = \exp_1$.

Remark 1.118. *Note that starting from Definition 1.68 the corresponding exponential map is the identity. This is the reason in order to choose the exponential Heisenberg group as canonical definition.*

As above, we calculate the exponential map of the polarized Heisenberg group. We recall that the basis associated to $\mathbb{G}_2$ is given by

$$X = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 \\ 1 \\ -x \end{pmatrix}, \quad Z = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}. \tag{1.95}$$

So in this case the system for the vector flux is

$$\begin{cases} \dot{\gamma}_1(t) = \alpha, \\ \dot{\gamma}_2(t) = \beta, \\ \dot{\gamma}_3(t) = -\beta\gamma_1 - \delta. \end{cases} \tag{1.96}$$

Solving the system, we find

$$\exp_2(\alpha, \beta, \delta) = \left(\alpha, \beta, -\delta - \frac{\alpha\beta}{2}\right).$$

Also in this case, the exponential map is invertible and it is equal to its inverse (i.e. $\exp_2^{-1} = \exp_2$).

Therefore we can now explicitly write the isomorphism $F = \exp_2 \circ \exp_1^{-1}$, which is

$$F(\alpha, \beta, \delta) = \left(\alpha, \beta, -\delta - \frac{\alpha\beta}{2}\right).$$

Note that F is invertible and coincides with its inverse, too.

We remain only to prove that F is an homomorphism from $\mathbb{G}_1$ to $\mathbb{G}_2$. We need to find explicitly the two composition laws, by using the Campbell-Hausdorff formula (1.57). First we calculate π_1.
Let be $p = (p_1, p_2, p_3)$ and $q = (q_1, q_2, q_3)$, we set $X_p = \exp_1^{-1} p$ and $X_q = \exp_1^{-1} q$, then

$$X_p = p_1 X_1 + p_2 Y_1 + p_3 Z_1$$

and analogously

$$X_q = q_1 X_1 + q_2 Y_1 + q_3 Z_1.$$

Applying formula (1.57), we get

$$\begin{aligned}
\pi_1(p,q) &= \exp_1 X_p \exp_1 X_q = \exp_1\left(X_p + X_q + \frac{1}{2}[X_p, X_q]\right)\\
&= \exp_1\left(p_1X_1 + p_2Y_1 + p_3Z_1 + q_1X_1 + q_2Y_1 + q_3Z_1 + \frac{p_1q_2}{2}Z_1 - \frac{p_2q_1}{2}Z_1\right)\\
&= \exp_1\left((p_1+q_1)X_1 + (p_2+q_2)Y_1 + \left(p_3 + q_3 + \frac{p_1q_2}{2} - \frac{p_2q_1}{2}\right)Z_1\right),
\end{aligned}$$

i.e.

$$\pi_1(p,q) = \left(p_1 + q_1, p_2 + q_2, p_3 + q_3 + \frac{1}{2}(p_1q_2 - p_2q_1)\right). \tag{1.97}$$

Remark 1.119. *Note that the composition law* (1.97) *is exactly the law given in Definition 1.106.*

To find the composition law of $\mathbb{G}_2$, we start from $X_p = \exp_2^{-1} p$ and $X_q = \exp_2^{-1} q$, which gives

$$X_p = p_1X_2 + p_2Y_2 + \left(-p_3 - \frac{p_1p_2}{2}\right)Z_2$$

and

$$X_q = q_1X_2 + q_2Y_2 + \left(-q_3 - \frac{q_1q_2}{2}\right)Z_2.$$

In this case, formula (1.57) gives

$$\begin{aligned}
\pi_2(p,q) &= \exp_2 X_p \exp_2 X_q = \exp_2\left(X_p + X_q + \frac{1}{2}[X_p, X_q]\right) = \exp_2\Big(p_1X_2 + p_2Y_2\\
&\quad + \left(-p_3 - \frac{p_1p_2}{2}\right)Z_2 + q_1X_2 + q_2Y_2 + \left(-q_3 - \frac{q_1q_2}{2}\right)Z_2 + \frac{p_1q_2}{2}Z_2 - \frac{p_2q_1}{2}Z_2\Big)\\
&= \exp_2\left((p_1+q_1)X_2 + (p_2+q_2)Y_2 + \left(-p_3 - \frac{p_1p_2}{2} - q_3 - \frac{q_1q_2}{2} + \frac{p_1q_2}{2} - \frac{p_2q_1}{2}\right)Z_2\right)\\
&= \left(p_1 + q_1, p_2 + q_2, p_3 + q_3 + \frac{p_1p_2}{2} + \frac{q_1q_2}{2} - \frac{p_1q_2}{2} + \frac{p_2q_1}{2} - \frac{1}{2}(p_1+q_1)(p_2+q_2)\right).
\end{aligned}$$

Hence

$$\pi_2(p,q) = (p_1+q_1, p_2+q_2, p_3+q_3-p_1q_2). \tag{1.98}$$

By (1.97) and (1.98), it is easy to check that F is an homomorphism between $\mathbb{G}_1$ and $\mathbb{G}_2$, in fact:

$$\begin{aligned} F\big(\pi_1(p,q)\big) &= F\left(p_1+q_1, p_2+q_2, p_3+q_3+\frac{1}{2}(p_1q_2-p_2q_1)\right) \\ &= \left(p_1+q_1, p_2+q_2, -p_3-q_3-\frac{1}{2}(p_1q_2-p_2q_1)-\frac{1}{2}(p_1+q_1)(p_2+q_2)\right) \\ &= \left(p_1+q_1, p_2+q_2, -p_3-q_3-\frac{p_1p_2}{2}-\frac{q_1q_2}{2}-p_1q_2\right). \end{aligned}$$

On the other side,

$$\begin{aligned} \pi_2\big(F(p),F(q)\big) &= \pi_2\left(\left(p_1,p_2,-p_3-\frac{p_1p_2}{2}\right),\left(q_1,q_2,-q_3-\frac{q_1q_2}{2}\right)\right) \\ &= \left(p_1+q_1, p_2+q_2, -p_3-q_3-\frac{p_1p_2}{2}-\frac{q_1q_2}{2}-p_1q_2\right). \end{aligned}$$

Therefore

$$F\big(\pi_1(p,q)\big) = \pi_2\big(F(p),F(q)\big)$$

and so F is the isomorphism between $\mathbb{G}_1$ and $\mathbb{G}_2$ which we are looking for.

We conclude showing that, by the application of F, the geodesics (1.91) go into the geodesics (1.76) and, analogously, the geodesics (1.92) goes into the geodesics (1.78).

It is immediate to note that F fixes the first-component and the second-component of any curve. Nevertheless the two first-components of the corresponding geodesics are not the same. In fact, the parameter c depends on the orientation of the vector field Z. So, if we want to compute the range of geodesics associated to the same parameters, we must start by geodesic-equations (1.91), rewritten w.r.t. $-c$, i.e.

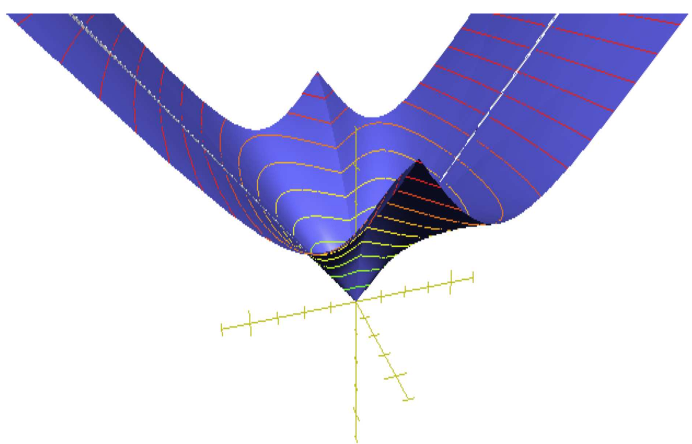
$$\gamma(t) = \begin{cases} \gamma_1(t) = \dfrac{a}{c}\sin ct + \dfrac{b}{c}(1-\cos ct), \\ \gamma_2(t) = \dfrac{b}{c}\sin ct - \dfrac{a}{c}(1-\cos ct), \\ \gamma_3(t) = -\dfrac{a^2+b^2}{2c^2}(ct-\sin ct). \end{cases}$$

Now the two first-components coincide with the corresponding ones in (1.76). Set $\widehat{\gamma}(t)$ the image by F of the geodesic $\gamma(t)$, it remains only to calculate the image of the third-component, i.e.

$$
\begin{aligned}
&\widehat{\gamma_3}(t) = -\gamma_3(t) - \frac{1}{2}\gamma_1(t)\gamma_2(t) = \frac{a^2+b^2}{2c}t - \frac{a^2+b^2}{2c^2}\sin(c\,t) \\
&-\frac{1}{2c^2}(a\sin(c\,t)+b-b\cos(c\,t))(b\sin(c\,t)-a+a\cos(c\,t)) = \frac{a^2+b^2}{2}t - \frac{a^2+b^2}{2c^2}\sin(c\,t) \\
&-\frac{1}{2c^2}\left(ab\sin^2(c\,t) - a^2\sin(c\,t) + (a^2-b^2)\sin(c\,t)\cos(c\,t) + b^2\sin(c\,t) - ab + ab\cos(c\,t) + ab\right. \\
&\left.-ab\cos^2(c\,t)\right) = \frac{a^2+b^2}{2c}t - \frac{a^2-b^2}{4c^2}\sin(2c\,t) + \frac{ab}{2c^2}\cos(2c\,t) - \frac{ab}{c^2}\cos(c\,t) - \frac{b^2}{c^2}\sin(c\,t) + \frac{ab}{2c^2},
\end{aligned}
$$

Hence $\widehat{\gamma_3}$ is exactly the third-component of (1.77) and this verifies the case $c \neq 0$. The case $c = 0$ is trivial. In fact, the image of the third-component of (1.92) is simply $\widehat{\gamma_3}(t) = \gamma_3(t) + \frac{1}{2}\gamma_1(t)\gamma_2(t) = 0 - \frac{ab}{2}t^2$. We have so proved that geodesics go into geodesics, as it is natural to expect.

By using the equation of geodesics for the canonical Heisenberg group, it is also possible to draw the graph of the Carnot-Carathéodory distance from the origin. Since the Heisenberg group is invariant by rotation in the (x, y)-plane, we can assume $x = y$. In this way the graph of the distance $d((0,0,0),(x,x,z))$ can be expressed by $\sqrt{\frac{2\,z\,c^2}{c-\sin c}}$ where the parameter c and the variables x and z are related by the equation $x = \sqrt{\frac{4\,z\,(1-\cos c)}{c-\sin c}}$.



Chapter 2

Viscosity solutions and metric Hopf-Lax formula.

2.1 An introduction to the theory of viscosity solutions.

Let us start by a simple example: the 1-dimensional eikonal equation. We consider the following Dirichlet problem:

$$\begin{cases} |u'(x)| = 1, \quad x \in (-1,1), \\ u(-1) = u(1) = 0. \end{cases} \tag{2.1}$$

It is easy to note that there are not classical solutions for Dirichlet problem (2.1). In fact, given a function $u \in C^1([-1,1])$ such that $u(-1) = u(1) = 0$, by Rolle's Theorem there exists a point $x_0 \in (-1,1)$ such that

$$u'(x_0) = \frac{u(1) - u(-1)}{2} = 0.$$

So the euquation is not satisfied at the point x_0. Moreover by continuity of $u'(x)$, $|u'(x)| < 1$ in some open neighborhood of x_0. The condition $u \in C^1$ is clearly too strong to get general existence results for nonlinear PDEs.
Therefore the idea of looking at weaker notions of solutions for nonlinear PDEs as the eikonal equation.
For example, if we look at the function $u(x) = -|x| + 1$ (which is continuous

in the interval $(-1, 1)$ but not differentiable at the point $x_0 = 0$), we can note that u satisfies the eikonal equation at any point exept the point $x_0 = 0$. Then the idea of looking at functions differentiable "except in some points", which satisfy the PDE just where the derivatives exist. By the Rademacher's Theorem, we know that locally Lipschitz continuous functions are differentiable almost everywhere. Then we call *almost everywhere solution* for a first-order PDE, any locally Lipschitz continuous function satisfying the equation at almost points.
Going back to the eikonal problem (2.1), we can note that any saw-tooth function (known also as Rademacher functions)

$$u_k(x) = \begin{cases} x + 1 - \dfrac{i}{2^{k-1}}, & \text{if } x \in \left[-1 + \dfrac{i}{2^{k-1}}, -1 + \dfrac{2i+1}{2^k}\right), \\ -x - 1 + \dfrac{i+1}{2^{k-1}}, & \text{if } x \in \left[-1 + \dfrac{2i+1}{2^k}, -1 + \dfrac{i+1}{2^{k-1}}\right), \end{cases} \tag{2.2}$$

with $i = 0, 1, ..., 2^k - 1$ and $k \in \mathbb{N}$, is an almost everywhere solution.

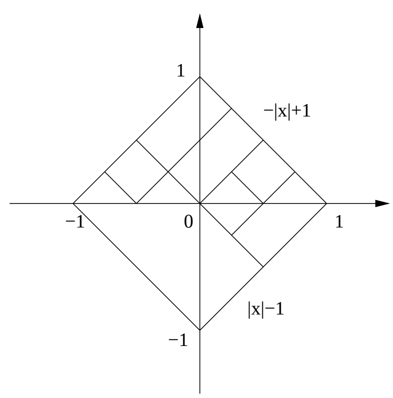


Figure 2.1: `Some different almost everywhere solutions of the eikonal problem` (2.1).

This example shows that almost everywhere solutions usually exist but are not unique. Moreover, as $k \to +\infty$, $u_k(x)$ converges uniformly to the zero function that satisfies the eikonal equation nowhere. So almost everywhere solutions are not stable w.r.t. uniform convergence.

We are going to introduce a different notion of weak solutions which leads to existence, uniqueness and stability for a large class of nonlinear PDEs: the so called viscosity solutions.
Viscosity solutions for Hamilton-Jacobi first-order PDEs have been introduced in the early '80 by M.G. Crandall and P.L. Lions in [39] (see [40] for the English version). In the following years, many authors have been worked to develop this new theory, hence many results for existence, uniqueness and regularity for first-order and second-order nonlinear PDEs have been proved. Here we will concentrate in particular on first-order equations, for more datails on the second-order case, we suggest to look at [23, 38].

2.1.1 Viscosity solutions for continuous functions.

We now introduce the classical theory of viscosity solutions for continuous functions. This theory has been introduced by M.G. Crandall and P.L. Lions in [39, 40] and then developed by H. Ishii, G. Barles, L.C. Evans and many others. In particular, we suggest to see [7, 11, 47, 71] for a very useful and clear treatment of viscosity theory for the Hamilton-Jacobi equations.
We begin giving the main definitions.

Definition 2.1. *Let $\Omega \subset \mathbb{R}^n$ be an open set and $u : \Omega \to \mathbb{R}$ a continuous function and let us consider the first-order PDE*

$$H(x, u(x), Du(x)) = 0, \quad x \in \Omega. \tag{2.3}$$

(i) *We say that u is a* viscosity subsolution *of* (2.3) *at a point $x_0 \in \Omega$ if and only if for any test function $\varphi \in C^1(\Omega)$ such that $u - \varphi$ has a local maximum at x_0, then*

$$H(x_0, u(x_0), D\varphi(x_0)) \leq 0; \tag{2.4}$$

(ii) *We say that u is a* viscosity supersolution *of* (2.3) *at a point $x_0 \in \Omega$, if and only if, for any test function $\varphi \in C^1(\Omega)$ such that $u - \varphi$ has a local minimum at x_0, then*

$$H(x_0, u(x_0), D\varphi(x_0)) \geq 0; \tag{2.5}$$

(iii) *We say that u is a* viscosity solution *in the open set Ω if u is a viscosity subsolution and a viscosity supersolution at any point $x_0 \in \Omega$.*

Remark 2.2. *An analogous definition holds for second-order PDEs.*

Remark 2.3. *If $u(x) - \varphi(x)$ has a local maximum (resp. minimum) at x_0, we can assume that the maximum (resp. minimum) is zero and moreover, by adding a suitable quadratic perturbation, we can suppose that the maximum (resp. minimum) is strict (see [36], Lemma 1.1).*

Example 2.4. *We want to show that $u(x) = |x|$ is the unique viscosity solution of the eikonal equation*

$$-|Du(x)| + 1 = 0, \quad x \in (-1, 1). \tag{2.6}$$

For the uniqueness see [7].
Note that, whenever $x_0 \neq 0$, x_0 is an extremal-point for $u - \varphi$ where the function is differentiable, then $D(\varphi - u)(x_0) = 0$, which means

$$D\varphi(x_0) = Du(x_0) = 1.$$

The problem occurs only at the point $x_0 = 0$, where the modulus is not differentiable. Let us first consider the subsolution property, i.e. x_0 local (strict) maximum point and assume $u(0) = \varphi(0)$; then $-|x| \leq u(x) - u(0) \leq \varphi(x) - \varphi(0)$, near 0, which implies

$$1 \geq \frac{\varphi(x) - \varphi(0)}{x}, \quad \textit{for } x < 0 \quad \textit{and} \quad \frac{\varphi(x) - \varphi(0)}{x} \geq -1, \quad \textit{for } x > 0.$$

Since $\varphi \in C^1$, passing to the limit as $x \to 0$, we can conclude $|\varphi'(0)| \leq 1$. Therefore the subsolution condition (2.4) is satisfied at any point $x_0 \in (-1, 1)$. To verify the supersolution condition, we assume that $u - \varphi$ attends a local maximum (equal to 0) at the point 0. Proceeding as above we find $D^-\varphi(0) \geq 1$ and $D^+\varphi(0) \leq -1$, which means $D^-\varphi(0) \neq D^+\varphi(0)$. Therefore there cannot be C^1-functions touching $u(x) = -|x| + 1$ from below, which means the condition (2.5) is trivially verified.

Example 2.5. *Similarly to the previous example, one can show that $v(x) = -u(x) = -|x|$ is a viscosity solution of*

$$|Du(x)| - 1 = 0, \quad x \in (-1, 1). \tag{2.7}$$

Remark 2.6. *Note that while $u(x) = |x|$ does not satisfy Eq. (2.7) in the viscosity sense. In fact, the function $\varphi(x) = -x^2+1$ is a C^1 function touching u from above at 0. Nevertheless $-|\varphi'(0)|+1 = 1 > 0$, therefore the subsolution property is not satisfied.*
In fact, the equations $H(x, u, Du) = 0$ and $-H(x, u, Du) = 0$ are not in general equivalent in the viscosity sense. It is very easy to show that, if $u(x)$ is a viscosity solution of $H(x, u, Du) = 0$, then $v(x) = -u(x)$ is a viscosity solution of $-H(x, -v, -Dv) = 0$.

In the next propositions we show that the definition of viscosity solutions is consistent with the notion of classical solutions: in fact classical solutions are viscosity solutions, too, and C^1-viscosity solutions are also classic solutions.

Proposition 2.7. *Let Ω be an open set and $u \in C^1(\Omega)$. Then u is a viscosity solution if and only if u is a classical solution.*

Proof. To prove that C^1-viscosity solutions are also classical solutions is trivial. In fact, we can choose, as test function, u itself and, therefore, it satisfies the equation in the classical sense.
In order to get the inverse implication, we can observe that, for any test function φ, if x_0 is an extremal point of $u - \varphi$ where both of the functions are differentiable, then $Du(x_0) = D\varphi(x_0)$. So if u satisfies the equation, any test function does so and, therefore, u is a viscosity solution. ☐

Remark 2.8. *For second-order PDEs the consistence with classical solutions is not always true. In this case, one has to assume that the equation is non-increasing in the second-order variables (e.g. elliptic or degenerate elliptic PDEs). Moreover, to ensure uniqueness at least of classical solutions, usually we assume also a monotone-increasing dependence w.r.t. u.*

The previous proposition can be written punctually, as follows.

Remark 2.9. *If u is a viscosity solution at a point where it is differentiable, then u satisfies the equation in the classical sense at that point (see [47], Theorem 10.1.1 for a proof of this result).*

Viscosity solutions are also consistent with almost everywhere solutions, in the sense that any locally Lipschitz continuous viscosity solution satisfies the PDE almost everywhere.

Proposition 2.10. *Let Ω be an open set and u be a viscosity solution in Ω. If u is Lipschitz in Ω (briefly $u \in \mathrm{Lip}(\Omega)$), then u is an almost everywhere solution, too.*

Proof. The result follows immediately by Remark 2.9 and the Rademacher's Theorem. □

Remark 2.11. *The reverse implication of Proposition 2.10 is in general not true, as we have seen in the case of the eikonal equation.*

Now we want to investigate the stability properties for viscosity solutions.

Proposition 2.12. *Let Ω be an open set and $u \in C(\Omega)$ and suppose x_0 is a maximum point (resp. minimum point) for $u - \varphi$ in $\overline{B_r(x_0)} \subset \Omega$, for some $\varphi \in C^1(\Omega)$. If $u_n \in C(\Omega)$ are such that*

$$\lim_{n\to\infty} u_n(x) = u(x),$$

for any $x \in \Omega$, let x_n be a sequence of minimum points for $u_n - \varphi$ in $\overline{B_r(x_0)}$, then $x_n \to x$, as $n \to \infty$, and moreover

$$\lim_{n\to\infty} u_n(x_n) = u(x_0).$$

From the previous result the stability follows imediately.

Proposition 2.13. *Let H_n and H be continuous in all the variables and such that $H_n \to H$, as $n \to \infty$. Let u_n be viscosity solutions of $H_n(x, u_n, Du_n) = 0$ and*

$$u(x) = \lim_{n\to+\infty} u_n(x).$$

If u is continuous, then u is a viscosity solution of $H(x, u, Du) = 0$.

The proofs of the two previous results are here omitted and we will give those later directly in the semicontinuous case (see Proposition 2.23 and Proposition 2.25).

Another important property of viscosity solutions is the behavior w.r.t. the operations of infimum and supremum. Let us start by the easy case of only two (or a finite number of) continuous functions.

Proposition 2.14.

1. *Let $u, v \in C(\Omega)$ be two viscosity subsolutions of $H(x, u, Du) = 0$, then $u \vee v$ is a viscosity subsolution of the same equation.*

2. *Let $u, v \in C(\Omega)$ be two viscosity supersolutions of $H(x, u, Du) = 0$, then $u \wedge v$ is a viscosity supersolution of the same equation.*

The proof of the previous result is pretty easy (e.g. [7], Proposition 2.1). The previous result can be generalized to inifinitly many functions.

Proposition 2.15.

1. *Let $\mathcal{F}$ be a family of viscosity subsolutions of $H(x, u, Du) = 0$, and define*
$$u(x) := \sup_{v \in \mathcal{F}} v(x).$$
If u is continuous, then u is a viscosity subsolution of the equation.

2. *Let $\mathcal{F}$ be a family of viscosity supersolutions of $H(x, v, Dv) = 0$, and define*
$$u(x) := \inf_{v \in \mathcal{F}} v(x).$$
If u is continuous, then it is a viscosity supersolutions of the same equation.

We now omit the proof since we are going to give that directly for lower semicontinuous viscosity solutions (see Proposition 2.26).
The behavior w.r.t. to the infimum (and the supremum respectively) is the key point to get existence of viscosity solutions by the so called Perron's method [81]. In 1987 H. Ishii used for the first time the *Perron's method* to solve nonlinear first-order equations (*Perron's method for Hamilton-Jacobi equations*, Duke Math. J. 55). This method had been introduced in 1923 by Oskar Perron in order to find solutions for the Laplace equation and consists in building a solution as the supremum of a suitable family of viscosity

subsolutions. Since the supremum of viscosity subsolutions is a viscosity subsolution, one has just to prove that it is a viscosity supersolution, too.
The Perron's method can be sketched, as follows:

Theorem 2.16 (Perron's Method). *Let us assume that*

1. *Comparison principle for the equation holds, i.e. given u viscosity subsolution and v viscosity supersolution satisfying the same boundary condition, then $u \leq v$.*

2. *Suppose that there exist $\underline{u}$ and $\overline{u}$ which are, respectively, a viscosity subsolution and a viscosity supersolution, satisfying the same boundary condition.*

We define

$$W(x) = \sup\{w(x) \,|\, \underline{u} \leq w \leq \overline{u} \text{ and } w \text{ viscosity subsolution}\},$$

Then $W(x)$ is a viscosity solution of the same equation, which satisfies the same boundary condition satisfied by $\underline{u}$ and $\overline{u}$.

The Perron's method is one of the most used method in order to get existence of solutions in the classical sense as well as in the viscosity sense. However we quote the result without giving any proof since we will apply a different method to prove existence for our class of PDEs, by the use of a suitable representative formula, the so called Hopf-Lax function (we will give many datails on this formula in Sec. 2.3).
Another method to get existence for nonlinear PDEs is the so called method of the characteristics. We refer to [47] for more details on this method.
A last method for existence of viscosity solution we like to quote is the so called vanishing viscosity approximation, which consists in adding to the equation a "viscosity term", i.e. a term of the form $-\varepsilon\Delta u$, for any $\varepsilon > 0$. This leads often to an elliptic or parabolic PDE, which is usually easier to solve than the starting Hamilton-Jacobi equation. Several times it is even possible to get classical solutions for the family of approximating PDEs. The limit of those classical solutions solves in the viscosity sense the starting Hamilton-Jacobi equation. The name "viscosity solutions" comes exactly

from this idea. We will go back on the vanishing viscosity approximation in the last section. More information can be found in [71] (Sec. 1.4 and Sec. 8), [47] (Sec. II.7.3), [36] (Theorem 3.1.) or [13, 87, 40]. See also [9] for a similar method.

We conclude quoting some results conerning uniqueness. Whenever the Hamiltonian satisfies some a coercive assumption or a Lipschitz assumption w.r.t. the gradient-variable, uniqueness for the the Dirichlet problem or the Cauchy problem can be found in [7, 11, 37, 41, 47, 71] and, only for the evolution equation, in [36, 58], too. In particular, we like to quote a pretty recent by A. Cutrí and G. Da Lio ([42]), which holds for the Cauchy problem and Hamiltonians satisfying a very general Hörmander condition (which means for the kind of degenerate Hamiltonians we are going to treat in this Thesis, see Sec. 2.4.2).

2.1.2 Discontinuous viscosity solutions.

There are different ways to define a notation of viscosity solutions, for discontinuous functions. By Weirstrass Theorem, we know that in $\mathbb{R}^n$ any continuous function attends maximum and minimum in a closed ball. This means that it is sufficient to require lower semicontinuity for viscosity supersolution (where we test at minimum points) and upper semicontinuous for viscosity subsolutions (where instead we test at maximum points). If the function is not continuous we can consider the right semicontinuous regularization of the function to test the supersolution and subsolution conditions. This leads to the following notion.

Definition 2.17. *Let $\Omega \subset \mathbb{R}^n$ be an open set and $u : \Omega \to \mathbb{R}$ be locally bounded function.*

1. *We say that u is a (discontinuous) viscosity subsolution of $H(x,u,Du) = 0$ if, defined the* upper semicontinuous envelop *of u*

   $$u^*(x) := \inf\{v(x) | v \in C(\Omega) \text{ and } v \geq u \text{ in } \Omega\} = \limsup_{r\to 0^+}\{u(y) |\, |y-x| \leq r\},$$

   u^* is a viscosity subsolution of the equation.

2. *We say that u is a (discontinuous) viscosity supersolution of $H(x,u,Du)=0$ if, defined the* lower semicontinuous envelop *of u*

$$u_*(x) := \sup\{v(x) | v \in C(\Omega) \text{ and } v \leq u \text{ in } \Omega\} = \liminf_{r\to 0^+}\{u(y) |\, |y-x| \leq r\},$$

u^* is a viscosity supersolution of the equation.

3. *We say that u is a (discontinuous) viscosity solution of $H(x,u,Du)=0$ if u^* is a viscosity subsolution of the equation and u^* is a viscosity supersolution of the equation.*

Remark 2.18. *Note that u^* and u_* are locally bounded and upper and lower semicontinuous, respectively. Moreover $u_* \leq u \leq u^*$.*

This notion of discontinuous viscosity solutions is used in [57, 56, 12] to get existence and uniqueness for a pretty large class of Hamilton-Jacobi equations.
In this thesis we istead follow a different notion of discontinuous viscosity solutions, which has been introduced by Barron and Jensen in [14, 15] and then applied to Dirichlet problem by Barles in [10] and to the Cauchy problem for bounded equations in infinite-dimensional spaces by Ishii in [59]. Some objections concerning the use of upper and lower semicontinuous envelopes to solve problems with discontinuous initial data can be found in [14]. We only remark that we are going to look at lower semicontinuous initial data. The easiest idea is so to apply directly the standard definition but testing only at the minimum points. For Hamiltonian which are convex w.r.t. the gradient, there is a very consistent way to do this. Let us give more details.

Definition 2.19. *Let Ω be an open set and u be a lower semicontinuous (briefly $u \in LSC(\Omega)$), we say that u is a* lower semicontinuous viscosity solution *(shorthly LSC-viscosity solutions) for Eq.* (2.3) *at some point $x_0 \in \Omega$ if and only if for any test function $\varphi \in C^1(\Omega)$ such that $u-\varphi$ attends a local minimum at x_0, then*

$$H(x_0, \varphi(x_0), D\varphi(x_0)) = 0. \tag{2.8}$$

Remark 2.20. *Note that Definition 2.19 differs from the standard viscosity supersolution requirement because of the identity instead of the inequality.*

Therefore Definition 2.19 is also called strong viscosity supersolution condition.

LSC-viscosity solutions satisfy almost all of the properties that we have seen for continuous viscosity solutions. Moreover Definition 2.19 is consistent with the standard definition of viscosity solutions (Definition 2.1).
A very nice proof of this result can be found in [10] (Theorem 2.1).
In [14] the authors have proved the consistence for evolution Hamitlon-Jacobi equations of the form

$$u_t(t,x) + H(x,u,Du) = 0, \tag{2.9}$$

where by Du we indicate the gradient w.r.t. x, by assumeing that the Hamiltonian is Lipschitz w.r.t. the gradient. We will show that this assumption is indeed not necessary for the consistence.

Theorem 2.21. *Let be $\Omega = [0,+\infty) \times \mathbb{R}^n$ and let $H(x,z,p) = H(x,p)$ be continuous in both of the variables and convex w.r.t. p. We assume that for $R > 0$ there exists a family of modulus of continuity $\omega_R(\cdot)$ and a continuous function $f : [0,+\infty) \to [0,+\infty)$ such that*

$$|H(x,p) - H(y,p)| \leq \omega_R(|x-y|)f(|p|) \quad \forall\, x,y \in B_R(0). \tag{2.10}$$

Then, if $u \in C(\Omega)$ is a LSC-viscosity solution of Eq. (2.9), then u is a standard viscosity solution (Definition 2.1).

Proof. It is immediate that if $u \in C(\Omega)$ is a strong viscosity supersolution (i.e. Definition 2.19 holds) then u is in particular a viscosity supersolution. So we need only to check that it is a viscosity subsolution, too, by using the convexity of H w.r.t. the gradient.
So let $\varphi \in C^1(\Omega)$ be such that $u-\varphi$ attends a local maximum at (t_0,x_0), we have to prove that

$$\varphi_t(t_0,x_0) + H(x_0, D\varphi(t_0,x_0)) \leq 0. \tag{2.11}$$

Define $\phi = u - \varphi$, we apply Theorem 1.1 in [14]: for any $\varepsilon > 0$ there exist $\psi \in C^\infty(\Omega)$, α_k and (t_k,x_k), with $k = 1, ..., N < +\infty$, such that

(i) $\alpha_k \geq 0$, for any $k = 1, ..., N$ and $\sum_{k=1}^N \alpha_k = 1$,

(ii) $\phi - \psi$ attends a minimum equal to 0 at any point (t_k, x_k), for $k = 1, ..., N$.

(iii) $(t_k, x_k) \in B_{o_\varepsilon(1)\varepsilon^{1/2}}(s, y)$ for some $(s, y) \in B_{o_\varepsilon(1)}(t_0, x_0)$, where by $o_\varepsilon(1)$ we indicate a quantity going to 0 as $\varepsilon \to 0^+$.

(iv) $\sum_{k=1}^N \alpha_k \psi_t(t_k, x_k) = 0 = \sum_{k=1}^N \alpha_k D\psi(t_k, x_k)$.

Since (t_k, x_k) are minimum points for $u - \varphi - \psi$ and u satisfies the equation, then $\widetilde{\varphi} = \varphi + \psi$ is a test function touching from above at (t_k, x_k), hence

$$\varphi_t(t_k, x_k) + \psi_t(t_k, x_k) + H(x_k, D\varphi(t_k, x_k) + D\psi(t_k, x_k)) = 0, \text{ for any } k = 1, .., N.$$

We add and subtract $\pm\varphi_t(s, y)$, $\pm H(y, D\widetilde{\varphi}(t_k, x_k))$ and $\pm H(y, D\varphi(s, y) + D\psi(t_k, x_k))$, and we estimate the corresponding differences. By property (iii) and using that $\varphi \in C^1(\Omega)$, we get

$$\varphi_t(t_k, x_k) - \varphi_t(s, y) \geq -o_\varepsilon(1)\varepsilon^{\frac{1}{2}} = -o_\varepsilon(1).$$

Moreover

$$|H(x_k, D\widetilde{\varphi}(t_k, x_k)) - H(y, D\widetilde{\varphi}(t_k, x_k))| \leq \omega(|x_k - y|) f(|D\widetilde{\varphi}(t_k, x_k)|).$$

By property (iii) we know that $(t_k, x_k) \in B_{o_\varepsilon(1)}(s, y)$ and $(s, y) \in B_{o_\varepsilon(1)}(t_0, x_0)$; hence $(t_k, x_k) \in K$ compact subset of $[0, +\infty) \times \mathbb{R}^n$ for any $k = 1, ..., N$. Since f and $D\widetilde{\varphi}$ are continuous, then $f(|D\widetilde{\varphi}(t_k, x_k)|) \leq C$ for any $k = 1, ..., N$, which implies

$$|H(x_k, D\widetilde{\varphi}(t_k, x_k)) - H(y, D\widetilde{\varphi}(t_k, x_k))| \leq C\omega(|x_k - y|) = o_\varepsilon(1).$$

We remain to estimate the last term, using the convexity of $H(x, p)$ w.r.t. the gradient which implies that the Hamiltonian is locally Lipschitz continuous w.r.t. the gradient. Note that $(t_k, x_k), (s, y) \in \overline{B_r(t_0, x_0)}$, so by the local Lipschitz continuity of H, we find

$$\begin{aligned} |H(y, D\varphi(t_k, x_k) + D\psi(t_k, x_k) - H(y, D\varphi(s, y) + D\psi(t_k, x_k)| \\ \leq C|D\varphi(t_x, x_k) - D\varphi(s, y)| = o_\varepsilon(1). \end{aligned}$$

To sum up, we have found the following estimate

$$o_\varepsilon(1) \geq \varphi_t(s,y) + \psi_t(t_k,x_k) + H(y, D\varphi(s,y) + D\psi(t_k,x_k)).$$

Multiplying by $\alpha_k \geq 0$ and summing over $k = 1,...,N$, we can use the convexity of H w.r.t. p, so that

$$\sum_{k=1}^N \alpha_k o_\varepsilon(1) \geq \sum_{k=1}^N \alpha_k \varphi_t(s,y) + \sum_{k=1}^N \alpha_k \psi_t(t_k,x_k) \\ + H\left(y, \sum_{k=1}^N \alpha_k D\varphi(s,y) + \sum_{k=1}^N \alpha_k D\psi(t_k,x_k)\right).$$

Using property (iv), we get

$$o_\varepsilon(1) \geq \varphi_t(s,y) + H(y, D\varphi(s,y));$$

using that $\varphi \in C^1(\Omega)$ and $(s,y) \in B_{o_\varepsilon(1)}(t_0,x_0)$, passing to the limit as $\varepsilon \to 0^+$, we find (2.11). □

Remark 2.22. *The Hamiltonian $H(x,p) = \frac{1}{\alpha}|\sigma(x)p|^\alpha$ for $\alpha \geq 1$ satisfies all the assumptions of Theorem 2.21 hold whenever $\sigma(x)$ has smooth coefficients.*

We now show some properties for LSC-viscosity solutions.

Proposition 2.23. *Let u be a lower semicontinuous function, $\varphi \in C^1$ and x_0 a minimum point for $u - \varphi$ in $\overline{B_r(x_0)}$. Let u_n be a sequence of lower semicontinuous functions such that*

$$\liminf_{n\to\infty} u_n(x) = u(x).$$

If x_n is a sequence of minimum points for $u_n - \varphi$ in $\overline{B_r(x_0)}$, then $x_n \to x$ as $n \to \infty$ and moreover

$$\liminf_{n\to\infty} u_n(x_n) = u(x_0).$$

Proof. Without loss of generality, we can assume that the minimum in $\overline{B_r(x_0)}$ is strict. Moreover $\overline{B_r(x_0)}$ is a compact set and so there exists a subsequence of x_n (which we still indicate by x_n) which converges to some point $y \in$

$\overline{B_r(x_0)}$.
We need only to check that $y = x_0$. Note that

$$u_n(x_n) - \varphi(x_n) \le u_n(x) - \varphi(x), \quad \forall\, x \in \overline{B_r(x_0)}.$$

Since u_n is lower semicontinuous for any $n \in \mathbb{N}$ and $x_n \to y$, passing to the lower limit in the previous inequality, we find

$$u(y) - \varphi(y) \le \liminf_{n\to+\infty} u_n(x_n) - \varphi(y) \le u(x) - \varphi(x). \tag{2.12}$$

Therefore y is a minimum point for $u - \varphi$ in $\overline{B_r(x_0)}$. Since we have assumed that the minimum is strict, then $y = x_0$.
Moreover, writing (2.12) at the point $x = x_0$, we get

$$u(x_0) - \varphi(x_0) \le \liminf_{n\to+\infty} u_n(x_n) - \varphi(x_0) \le u(x_0) - \varphi(x_0).$$

We can so conclude that

$$\liminf_{n\to+\infty} u_n(x_n) = u(x_0).$$

□

Remark 2.24. *Without loss of generality we can assume $u_n(x_n) = \varphi(x_n)$ for any n, which implies $u(x_0) = \varphi(x_0)$.*

The previous lemma can be used to show the stability of LSC-viscosity solutions.

Proposition 2.25. *Let $H_n(x, z, p)$ and $H(x, z, p)$ be continuous Hamiltonians and assume that $H_n(x, z, p) \to H(x, z, p)$ as $n \to \infty$. Let u_n be LSC-viscosity solutions of $H_n(x, u_n, Du_n) = 0$. Define*

$$u(x) = \liminf_{n\to+\infty} u_n(x).$$

If u is a lower semicontinuous function, then u is a LSC-viscosity solution for the limit-equation $H(x, u, Du) = 0$.

Proof. Let $\varphi \in C^1$ be such that $u - \varphi$ attends a local strict minimum at x_0, which means there exists $r > 0$ such that

$$u(x_0) - \varphi(x_0) < u(x) - \varphi(x), \quad \forall\, x \in \overline{B_r(x_0)} \setminus \{x_0\}.$$

Since u_n are lower semicontinuous functions, there is a sequence of points x_n where $u_n - \varphi$ attends minimum in $\overline{B_r(x_0)}$.
By Proposition 2.23 $x_n \to x_0$; hence there exists $\overline{n} \in \mathbb{N}$ such that for any $n > \overline{n}$, $x_n \in B_r(x_0)$ and

$$\liminf_{n\to+\infty} u_n(x_n) = u(x_0).$$

We recall that $\varphi(x_n) = u_n(x_n)$ and $\varphi(x_0) = u(x_0)$ (see Remark 2.24): u_n are LSC-viscosity solutions of $H_n(x, u_n, Du_n) = 0$, so

$$H_n(x_n, \varphi(x_n), D\varphi(x_n)) = 0.$$

Moreover $\varphi \in C^1$ and $H(x, z, p)$ continuous; hence passing to the limit as $n \to \infty$, we get

$$H(x, \varphi(x_0), D\varphi(x_0)) = 0.$$

$\square$

Next we study the behavior of LSC-viscosity solutions w.r.t. the infimum and the supremum.

Proposition 2.26. *Let $v \in \mathcal{F}$ be a family of LSC-viscosity solutions of $H(x, v, Dv) = 0$ and $u(x) = \inf_{v\in\mathcal{F}} v(x)$. If u is a lower semicontinuous function, then it is a LSC-viscosity solution of the same equation.*

Proof. Let $\varphi \in C^1$ be such that $u - \varphi$ has a local strict minimum at x_0, i.e. there exists $r > 0$ such that

$$u(x_0) - \varphi(x_0) < u(x) - \varphi(x), \quad \forall\, x \in \overline{B_r(x_0)} \backslash \{x_0\}.$$

Since $u(x_0) = \inf\{v(x_0) \,|\, v \in \mathcal{F}\}$, for any $n \in \mathbb{N}$ there exists $v_n \in \mathcal{F}$ such that

$$v_n(x_0) < u(x_0) + \frac{1}{n}.$$

Let x_n be a sequence of minimum points for $v_n - \varphi$ in $\overline{B_r(x_0)}$, i.e.

$$v_n(x_n) - \varphi(x_n) \le v_n(x) - \varphi(x), \quad \forall\, x \in \overline{B_r(x_0)}, \tag{2.13}$$

we know that (up to a subsequence) $x_n \to y$ for some point $y \in \overline{B_r(x_0)}$. Writing (2.13) at the point $x = x_0$, we find

$$u(x_n) - \varphi(x_n) \leq v_n(x_n) - \varphi(x_n) \leq v_n(x_0) - \varphi(x_0) < u(x_0) + \frac{1}{n} - \varphi(x_0).$$

Passing to the lower limit as $n \to +\infty$, and using the lower semicontinuity of u and the continuity of test function φ, we conclude

$$u(y) - \varphi(y) \leq u(x_0) - \varphi(x_0).$$

Therefore y is a minimum point of $u - \varphi$ in $\overline{B_r(x_0)}$. Since x_0 is a strict pint of minimum, then $x_0 = y$.
We now use that v_n are LSC-viscosity solutions (with $\varphi(x_n) = u_n(x_n)$), i.e.

$$H(x_n, \varphi(x_n), D\varphi(x_n)) = 0.$$

Passing to the limit as $n \to +\infty$ and, recalling that $\varphi(x_0) = u(x_0)$, we find

$$H(x_0, u(x_0), D\varphi(x_0)) = 0.$$

□

We are going to look in particular at the Cauchy problem for an evolution Hamilton-Jacobi equation:

$$\begin{cases} u_t + H(t, x, u, Du) = 0, & \text{in } \mathbb{R}^n \times (0, +\infty), \\ u = g, & \text{in } \mathbb{R}^n \times \{0\}, \end{cases} \tag{2.14}$$

with a lower semicontinuous initial data g.
For sake of simplicity, we prefer to re-write Definition 2.19 explicitly in the evolution case

Definition 2.27. *Let Ω be an open subset of $\mathbb{R}^n \times (0, +\infty)$ and $u \in LSC(\Omega)$, u is a* LSC-viscosity solution *for the Hamilton-Jacobi equation* (2.9), *at the point (t_0, x_0) if and only if, for any test function $\varphi \in C^1(\Omega)$ such that $u - \varphi$ attends a local minimum at (t_0, x_0),*

$$\frac{\partial \varphi}{\partial t}(t_0, x_0) + H(t_0, x_0, u(t_0, x_0), D\varphi(t_0, x_0)) = 0. \tag{2.15}$$

In [14] the authors have proved that upper semicontinuous viscosity solutions (defined testing the subsolution-property at the maximum points) for a terminal-time Cauchy problem are unique, under suitable assumptions on the Hamiltonians. In particular the Hamiltonian is assumed to be globally Lipschitz w.r.t. the gradient. By this result, the uniqueness for LSC-viscosity solutions of (2.14) follows immediately.
We like to quote also [10], where a uniqueness result for the stationary equation is proved, under almost the same conditions on the Hamiltonian. In the same paper, it is possible to find a counter-example for the uniqueness in the case of non-convex Hamiltonians. In particular the author explicitly builds two different solutions for the following Cauchy problem:

$$\begin{cases} \dfrac{\partial u}{\partial t} + (x-t)\left|\dfrac{\partial u}{\partial x}\right| = 0, \quad \text{in } (0,+\infty)\times\mathbb{R} \\ u(0,x) = \begin{cases} 1, & x\in(\alpha,\beta) \\ -1, & \text{otherwise} \end{cases} \end{cases}$$

with $\alpha, \beta \in \mathbb{R}$ constants chosen in a suitable way.
The existence for LSC-viscosity solutions for the Cauchy problem (2.14) has been proved in [3] in the simple case when $H = H(p)$, by using the classic Hopf-Lax formula.

To generalize the previous result we need first to solve the associated eikonal equation.

2.2 The generalized eikonal equation.

In this section we study a generalized eikonal equation, with Hamiltonian of Hörmander's type.
Generalized solutions and, in particular, viscosity solutions of eikonal equations have been studied by P.L. Lions in [71] for convex geometrical Hamiltonian, and by A. Siconolfi in [86] in the non convex case.
The easiest example of eikonal equation is $|Du| = 1$: the viscosity solutions are Euclidean distances (see Example 2.4 for the 1-dimensional case).
For more general eikonal equations, it can be very difficult to write explicitly

the viscosity solutions. Nevertheless, under suitable assumptions, we can build the viscosity solutions as minimal-time functions. We will in particular show that these functions satisfy the definition of generalized distance.

Definition 2.28. *A* generalized distance *is a non negative function* $d : \mathbb{R}^n \times \mathbb{R}^n \to [0, +\infty)$ *satisfying*

$$d(x,y) \geq 0, \ \forall\, x,y \in \mathbb{R}^n, \quad d(y,y) = 0, \ \forall\, y \in \mathbb{R}^n, \tag{2.16}$$

$$d(x,y) \leq d(x,z) + d(z,y), \quad \forall\, x,y,z \in \mathbb{R}^n. \tag{2.17}$$

In the previous definition we do not require neither the symmetry nor the non-degenery (i.e. $d(x,y) = 0$ implies that $x = y$). E.g. the degenerate distance associated to a Finsler metric is a generalized distance in the sense of Definition 2.16 (see [43] for some details on Finsler metrics).
It is trivial that any finite distance is a generalized distance, too.

We now look at an eikonal equation

$$H_0(x, Du(x)) = 1, \tag{2.18}$$

where H_0 is a *geometrical Hamiltonian*, i.e. $H_0 : \mathbb{R}^{2n} \longrightarrow [0, +\infty)$ continuous in both of the variables, convex and positively homogeneous of degree 1, w.r.t. the gradient-variable.
Moreover we assume that there exists a smooth $m \times n$-matrix $\sigma(x)$ satisfying the Hörmander condition and such that

$$\sigma^t(x)\overline{B_1(0)} \subset \partial H_0(x, 0), \quad \text{for any } x \in \mathbb{R}^n \tag{2.19}$$

where $\sigma^t(x)$ is the transpose matrix of $\sigma(x)$, $\partial H_0(x, 0)$ is the subgradient of the convex function $p \longmapsto H_0(x, p)$ at the point $(x, 0)$, and $\overline{B_1(0)}$ is the closed Euclidean unit ball in $\mathbb{R}^m$ (with $m \leq n$).

Remark 2.29. *For details on the Hörmander condition we refer to Sec. 1.1. We recall that a matrix* $\sigma(x) = [X_1(x), ..., X_m(x)]^t$ *satisfies the Hörmander condition if and only if the associated Lie algebra* $\mathcal{L}\big(\sigma\big)(x) = \mathcal{L}\big(X_1, ..., X_m\big)(x)$ *is equal to* $\mathbb{R}^n$ *at any* $x \in \mathbb{R}^n$.

Under the Hörmander condition and assumprion (2.19), fixed a point $y \in \mathbb{R}^n$, we can build a generalized distance which is a viscosity solution of the vanishing Dirichlet eikonal problem in $\mathbb{R}^n \backslash \{y\}$, i.e.

$$\begin{cases} H(x, Du(x)) = 1, & \mathbb{R}^n \backslash \{y\}, \\ u(y) = 0. \end{cases}$$

We look at the differential inclusion

$$\dot{X}(t) \in \partial H_0(X(t), 0), \quad t \in (0, +\infty). \tag{2.20}$$

We recall that a solution of (2.20) is an absolutely continuous function $X : (0, +\infty) \to \mathbb{R}^n$ satisfying (2.20) almost everywhere (see [6, 5]).
We indicate by $F_{x,y}$ the set of all the solutions $X(\cdot)$ of the differential inclusion (2.20), joining x to y in some finite time, i.e. $X(0) = x$ and $X(T) = y$, for some $0 \le T = T(X(\cdot)) < +\infty$.
First note that assumption (2.19) implies that $F_{x,y} \neq \emptyset$, for any couple of points $x, y \in \mathbb{R}^n$. In fact, the set of all the solutions of (2.20) includes in particular any solution of the following family of control systems

$$\begin{cases} \dot{X}(t) = \sigma^t(X(t))\alpha(t), & t \in (0, +\infty), \\ X(0) = x, \end{cases} \tag{2.21}$$

where the control α is a t-measurable function, with $|\alpha(t)| \le 1$ for a.e. $t > 0$. So, by Chow's Theorem 1.29, we know that there exists a solution of (2.21), joining x to y in a finite time, and then $F_{x,y} \neq \emptyset$.
We define

$$d(x, y) := \inf_{X(\cdot) \in F_{x,y}} T(X(\cdot)), \tag{2.22}$$

and show that the previous minimal-time function is a generalized distance.

Lemma 2.30. *Let $H_0(x, p)$ be a geometrical Hamiltonian satisfying* (2.19)*, the minimal-time function defined by* (2.22) *is a generalized distance and induces on $\mathbb{R}^n$ the Euclidean topology.*

Proof. We have already remarked that the Hörmander condition implies that $d(x, y)$ is finite in any couple of points. Also property (2.16) is immediate. We must only check the triangle inequality (2.17). Let be

$$E_{x,y} := \{T = T(X(\cdot)) \,|\, X(\cdot) \in F_{x,y}\}.$$

For any $T_1 \in E_{x,y}$ and $T_2 \in E_{y,z}$, for $i = 1, 2$ we indicate by $X_i(\cdot)$ the trajectory w.r.t. T_i, and look at the path defined as

$$X(t) := \begin{cases} X_1(t), & 0 \le t \le T_1, \\ X_2(t - T_1), & T_1 \le t \le T_1 + T_2. \end{cases}$$

Then the path $X(\cdot)$ satisfies the differential inclusion (2.20) with $X(0) = x$ and $X(T_1 + T_2) = z$. So $X(\cdot) \in F_{x,z}$, which means that $T_1 + T_2 \in E_{x,z}$. Then $d(x, z) \le T_1 + T_2$ and, taking the infimum over $E_{x,y}$ and $E_{y,z}$, we can conclude that

$$d(x, z) \le d(x, y) + d(y, z).$$

By the Hörmander-inclusion (2.19) $d(x, y)$ induces on $\mathbb{R}^n$ the Euclidean topology. The proof of this claim is very simlar to prove Theorem 1.52. □

Remark 2.31. *In general $d(x, y)$ is not symmetric. In fact, let be $X(\cdot) \in F_{x,y}$, then the inverse path $\widetilde{X}(t) := \widetilde{X}(T - t)$ may not satisfy* (2.20).

Remark 2.32. *The topological property tells that $d(x, y)$ is continuous in $\mathbb{R}^n$, for any fixed y. This fact is very important in order to use $d(x, y)$ for solving a Hamilton-Jacobi-Cauchy problem (see Sec. 2.4.2).*

Example 2.33. *Without the Hörmander condition the topology induced by the distance may be different from the original one. E.g. the distance introduced in Example 1.35) is finite but it is not continuous w.r.t. the Euclidean topology defined on $\mathbb{R}^2$, and in fact the Hörmander condition is not satisfied. The Hörmander condition implies always the topological property above (see Theorem 1.52).*

To prove that $d(x, y)$ is a viscosity solution of the eikonal equation (2.18) in $\mathbb{R}^n \setminus \{y\}$, we proceed as in [7], showing an associated Dynamical Programming Principle.

Lemma 2.34 (Dynamical Programming Principle)**.** *Under the assumptions of Lemma 2.30*

$$d(x, y) = \inf_{X(\cdot) \in F_{x,y}} [t + d(X(t), y)], \tag{2.23}$$

for any $x, y \in \mathbb{R}^n$ and $0 \le t \le d(x, y)$.

Proof. First we prove that

$$d(x, y) \leq \inf_{X(\cdot)\in F_{x,y}} [t + d(X(t), y)]. \tag{2.24}$$

Let be $y \in \mathbb{R}^n$ and be $X(\cdot) \in F_{x,y}$, we fix $z = X(t)$.
Since $d(z, y) = \inf_{X(\cdot)\in F_{z,y}} T(X(\cdot))$, for any $\varepsilon > 0$ there exists $\tilde{X}(\cdot) \in F_{z,y}$ such that

$$d(z, y) > T(\tilde{X}(\cdot)) - \varepsilon.$$

We define

$$\overline{X}(s) := \begin{cases} X(s), & 0 < s \leq t, \\ \tilde{X}(s - t), & t < s. \end{cases}$$

Note that $\overline{X}(\cdot) \in F_{x,y}$. So for any $\varepsilon > 0$

$$d(x, y) \leq T(\overline{X}(\cdot)) = t + T(\tilde{X}(\cdot)) < t + d(z, y) + \varepsilon.$$

Passing to the limit as $\varepsilon \to 0^+$, we find

$$d(x, y) \leq t + d(z, y) = t + d(X(t), y).$$

Taking the infimum over $X(\cdot) \in F_{x,y}$, we get (2.24).
We remain to prove the reverse inequality, i.e.

$$\inf_{X(\cdot)\in F_{x,y}} [t + d(X(t), y)] \leq d(x, y). \tag{2.25}$$

Fix $y \in \mathbb{R}^n$ and remark that for any $X(\cdot) \in F_{x,y}$ and for any $0 \leq t \leq d(x, y) \leq T(X(\cdot))$

$$T(X(\cdot)) \geq t + d(X(t), y).$$

Taking the infimum over $X(\cdot) \in F_{x,y}$, we get (2.25) and therefore identity (2.23). $\square$

Next we present some examples of geometrical Hamiltonians satisfying our assumptions.

Note that any geometrical Hamiltonian $H_0(x, p) \geq |\sigma(x)p|$ is such that (2.19) holds. In fact, whenever $H_1(x, p) \leq H_2(x, p)$ and $H_1(x, 0) = H_2(x, 0) = 0$, then $\partial H_1(x, 0) \subset \partial H_2(x, 0)$. Using this remark we can build a lot of interesting examples.

Example 2.35. *The Hamiltonian $H_0(x,p)=|\sigma(x)p|_m+\lambda|p|_n$ is continuous in all the variables, positively homogeneous of degree 1 and convex w.r.t. p and satisfying (2.19), for any $\lambda\geq 0$. More in general, given an invertible $n\times n$-matrix $A(x)$, those assumptions are satisfied by any Hamiltonian like $H_0(x,p)=|\sigma(x)p|_m+\lambda|A(x)p|_n$.*

Example 2.36. *For $\varepsilon>0$, the Hamiltonian*

$$H_0(x,p)=\big(|\sigma(x)p|_m^2+\varepsilon^2|p|_n^2\big)^{\frac{1}{2}},$$

verifies our assumptions. In fact, it is continuous w.r.t. x and p, positively homogeneous w.r.t. p and satisfying (2.19), since $H_0(x,p)\geq|\sigma(x)p|_m$. We have only to prove that it is convex w.r.t. p.
So let be $\lambda\in[0,1]$ and look at

$$\begin{aligned}H_0(x,\lambda p+(1-\lambda)q)=&[\lambda^2(|\sigma(x)p|_m^2+\varepsilon^2|p|_n^2)+(1-\lambda)^2(|\sigma(x)q|_m^2+\varepsilon^2|q|_n^2)\\&+2\lambda(1-\lambda)(\big\langle\sigma(x)p,\sigma(x)q\big\rangle_m+\varepsilon^2\big\langle p,q\big\rangle_n)]^{\frac{1}{2}}\\=&[\lambda^2H_0(x,p)^2+(1-\lambda)^2H_0(x,q)^2+2\lambda(1-\lambda)I]^{\frac{1}{2}},\end{aligned}$$

with $I=\big\langle\sigma(x)p,\sigma(x)q\big\rangle_m+\varepsilon^2\big\langle p,q\big\rangle_n$. Note that

$$\begin{aligned}&2\lambda H_0(x,p)(1-\lambda)H_0(x,q)=\\&2\lambda(1-\lambda)[|\sigma(x)p|_m^2|\sigma(x)q|_m^2+\varepsilon^2|\sigma(x)p|_m^2|q|_n^2+\varepsilon^4|p|_n^2|q|_n^2+\varepsilon^2|\sigma(x)q|_m^2|p|_n^2]^{\frac{1}{2}}\\&\geq 2\lambda(1-\lambda)[|\sigma(x)p|_m^2|\sigma(x)q|_m^2+\varepsilon^4|p|_n^2|q|_n^2+2\varepsilon^2|\sigma(x)p|_m|\sigma(x)q|_m|p|_n|q|_n]^{\frac{1}{2}},\end{aligned}$$

since $A^2+B^2\geq 2AB$, for any $A,B\geq 0$ (setting $A=\varepsilon|\sigma(x)p|_m|q|_n$ and $B=\varepsilon|\sigma(x)q|_m|p|_n$).
Applying the Cauchy-Schwartz inequality, we get

$$2\lambda(1-\lambda)H_0(x,p)H_0(x,q)=2\lambda(1-\lambda)[(|\sigma(x)p|_m|\sigma(x)q|_m+\varepsilon^2|p|_n|q|_n)^2]^{\frac{1}{2}}\geq I,$$

which gives

$$H_0(x,\lambda p+(1-\lambda)q)\leq\lambda H_0(x,p)+(1-\lambda)H_0(x,q).$$

To calculate the subdifferential of the previous examples is not trivial. Looking at Example 2.35, we can be sure that $\sigma^t(x)\overline{B_1^m(0)}\cup\lambda\overline{B_1^n(0)}\subset\partial H_0(x,0)$ but we are not able to get the identity.
While in the next example is easy to calculate the subdifferential, nevertheless we get a Hamiltonian not really interesting.

Example 2.37. *Look at* $\partial H_0(x,0) = \sigma^t(x)\overline{B_1^m(0)} \oplus \lambda\overline{\big(B_1^n(0) \cap \mathrm{Ker}(\sigma(x))\big)}$, *if we indicate by* $\pi_{\mathrm{Ker}(\sigma)}$ *the projection-map on the* $\mathrm{Ker}(\sigma(x))$, *the associated Hamiltonian is*

$$H_0(x,p) = |\sigma(x)p|_m + \lambda|\pi_{\mathrm{Ker}(\sigma)}(p)|_n = |\sigma(x)p|_m + \lambda|p - \sigma^t(x)p|_n.$$

In order to prove, that we proceed exactly as in Lemma 2.38.

The main model of the generalized eikonal equation that we are going to study is the *horizontal eikonal equation*:

$$|\sigma(x)Du(x)| = 1, \quad x \in \mathbb{R}^n, \tag{2.26}$$

where $\sigma(x)$ is a smooth $m \times n$-matrix satisfying the Hörmander condition. We show that in this particular case the minimal-time distance defined by (2.22) coincides with the Carnot-Carathéodory distance associated to the matrix $\sigma(x)$. We indicate by $|\ \cdot\ |_m$ the Euclidean norm in $\mathbb{R}^m$ and by $\big\langle \cdot, \cdot \big\rangle_m$ the inner product in $\mathbb{R}^m$.

Lemma 2.38. *Let be* $H_0(x,p) = |\sigma(x)p|_m$ *where* $\sigma(x)$ *is a smooth* $m \times n$-*matrix satisfying the Hörmander condition with rank equal to* m, *then*

$$\sigma^t(x)\overline{B_1(0)} = \partial H_0(x,0),$$

for any $x \in \mathbb{R}^n$. *Therefore* $d(x,y) = d_\sigma(x,y)$ *for any* $x, y \in \mathbb{R}^n$.

Proof. Let be $p \in \sigma^t(x)\overline{B_1(0)}$, then there exists $\alpha \in \overline{B_1(0)}$ such that $p = \sigma^t(x)\alpha$. By the Cauchy-Schwartz inequality, for any $q \in \mathbb{R}^n$ we get

$$\big\langle p, q \big\rangle_n = \big\langle \sigma^t(x)\alpha, q \big\rangle_n = \big\langle \alpha, \sigma(x)q \big\rangle_m \le |\alpha|_m\, |\sigma(x)q|_m \le |\sigma(x)q|_m = H_0(x,q).$$

Hence $p \in \partial H(x,0)$ and so we can conclude that $\sigma^t(x)\overline{B_1(0)} \subset \partial H(x,0)$. In order to prove the reverse inequality, we fix x and omit to write the dependence on it. Since $\mathrm{rank}(\sigma) = m$, we can write $\mathbb{R}^n = \mathrm{Ker}(\sigma) \oplus \mathrm{Im}(\sigma^t)$. If $v \in \partial H_0(x,0)$, then

$$\big\langle v, p \big\rangle_n \le H_0(x,0) = |\sigma(x)\, p|_m, \quad \forall\, p \in \mathbb{R}^n.$$

Choosing $p \in \mathrm{Ker}(\sigma)$ we get $\big\langle v, p \big\rangle_n \le 0$, that implies $v \in \mathrm{Im}(\sigma^t)$. So there exists $w \in \mathbb{R}^m$ such that $v = \sigma^t w$. Hence

$$\big\langle \sigma^t(x)w, p \big\rangle_n = \big\langle w, \sigma(x)\, p \big\rangle_m \le |\sigma\, p|_m, \quad \forall\, p \in \mathbb{R}^n.$$

Since $\mathrm{rank}(\sigma) = m$, there exists a $\overline{p} \in \mathbb{R}^n$ such that $\sigma(x)\,\overline{p} = w$, so

$$\langle w, w \rangle_m = |w|_m^2 \le |w|_m,$$

that implies $w \in \overline{B_1(0)} \subset \mathbb{R}^m$.
Therefore $v \in \sigma^t(x)\overline{B_1(0)}$, which concludes $\partial H_0(x,0) \subset \sigma^t(x)\overline{B_1(0)}$. □

Corollary 2.39. *Under assumptions of Lemma 2.38, the minimal-time function* (2.22) *coincides with the Carnot-Carathéodory distance associated to the matrix $\sigma(x)$ (see Sec. 1.1.5, Definition 1.40)*

We are going to prove that the Carnot-Carathéodory distance $u(x) = d_\sigma(x,y)$ is a viscosity solution (and then an almost everywhere solution, too) of the horizontal eikonal problem

$$\begin{cases} |\sigma(x)Du(x)| = 1, & \text{in } \mathbb{R}^n \backslash \{y\}, \\ u(y) = 0, \end{cases} \tag{2.27}$$

for any fixed $y \in \mathbb{R}^n$.
We now give one of the main results of this section.

Theorem 2.40 ([44])**.** *Let $\sigma(x)$ be a smooth $m \times n$-matrix as in Lemma 2.38, then the associated Carnot-Carathéodory distance $d_\sigma(x,y)$ is a viscosity solution of the eikonal problem* (2.27)*, for any fixed $y \in \mathbb{R}^n$.*

Proof. In order to prove the theorem, we use the expression of $d_\sigma(x,y)$ as minimal-time function and the Dynamical Programming Priciple (Lemma 2.34).
We define $u(x) = d(x,y)$ and prove first that u is a viscosity subsolution in $\mathbb{R}^n \backslash \{y\}$.
At this purpose, let be $x \neq y$ and $\varphi \in C^1(\mathbb{R}^n)$ be such that $u - \varphi$ has a local maximum at x, i.e. there exists $R > 0$ such that

$$\varphi(x) - \varphi(z) \le d(x,y) - d(z,y), \quad \forall\, z \in B_R(x).$$

Let be $\alpha \in \overline{B_1(0)}$ and let $X_\alpha(\cdot)$ be a solution of a control system with constant control α, i.e.

$$\begin{cases} \dot{X}_\alpha(t) = \sigma^t(X_\alpha(t))\alpha \\ X_\alpha(0) = x \end{cases}$$

Note that, since $\sigma \in C^\infty$ and $X_\alpha(\cdot) \in C^\infty$, then in particular $\dot{X}_\alpha(0) = \sigma^t(x)\alpha$. For time t small enough, then $X_\alpha(t) \in B_R(x)$. Hence

$$\varphi(x) - \varphi(X_\alpha(t)) \leq d(x,y) - d(X_\alpha(t),y) \leq t + d(X_\alpha(t),y) - d(X_\alpha(t),y) = t,$$

which implies

$$\frac{\varphi(x) - \varphi(X_\alpha(t))}{t} \leq 1.$$

Since $X_\alpha(\cdot)$ is smooth, we can pass to the limit as $t \to 0^+$, and get

$$-\big\langle D\varphi(x), \dot{X}_\alpha(0)\big\rangle_n = -\big\langle D\varphi(x), \sigma^t(x)\alpha\big\rangle_n = \big\langle \sigma(x)D\varphi(x), -\alpha\big\rangle_m \leq 1, \tag{2.28}$$

for any $|\alpha|_m \leq 1$. Taking the infimum for $\alpha \in \overline{B_1(0)}$, we can conclude $|\sigma(x)D\varphi(x)|_m \leq 1$.

We remain to prove that u is a viscosity supersolution, too.
Fixed $x \neq y$, by Dynamical Programming Principle we know that for any $\varepsilon > 0$ there exists $\overline{X}_\varepsilon(\cdot) \in F_{x,y}$ such that

$$d(x,y) > d(\overline{X}_\varepsilon(t), y) + t - \varepsilon t. \tag{2.29}$$

Let $\varphi \in C^1(\mathbb{R}^n)$ be such that $u - \varphi$ has a local minimum at x, i.e. there exists $R > 0$ such that

$$d(x,y) - d(z,y) \leq \varphi(x) - \varphi(z), \quad \forall\, z \in B_R(x).$$

$\overline{X}_\varepsilon$ is absolutely continuous, so for t small enough, $\overline{X}_\varepsilon(t) \in B_R(x)$, which implies

$$d(x,y) - d(\overline{X}_\varepsilon(t), y) \leq \varphi(x) - \varphi(\overline{X}_\varepsilon(t)). \tag{2.30}$$

By using (2.29) in (2.30), we find

$$\frac{\varphi(x) - \varphi(\overline{X}_\varepsilon(t))}{t} \geq 1 - \varepsilon. \tag{2.31}$$

Note that in general $\overline{X}_\varepsilon$ is not differentiable so in order to conclude, we cannot pass directly to the limit as $t \to 0^+$.
Nevertheless by the absolutely continuity we can write

$$\begin{aligned} \frac{\varphi(x) - \varphi(\overline{X}_\varepsilon(t))}{t} &= -\frac{1}{t}\int_0^t \big\langle D\varphi\big(\overline{X}_\varepsilon(s)\big), \dot{\overline{X}}_\varepsilon(s)\big\rangle_n ds \\ &= -\frac{1}{t}\int_0^t \big\langle D\varphi\big(\overline{X}_\varepsilon(s)\big), \sigma^t(\overline{X}_\varepsilon(s))\overline{\alpha}(s)\big\rangle_n ds \end{aligned}$$

Now we add and subtract $\pm\big\langle D\varphi\big(\overline{X}_\varepsilon(s)\big), \sigma^t(x)\overline{\alpha}(s)\big\rangle_n$ and $\pm\big\langle D\varphi(x), \sigma^t(x)\overline{\alpha}(s)\big\rangle_n$. Since the coefficients of $\sigma(x)$ are smooth and $\varphi \in C^1$, and the absolutely continuity of $\overline{X}_\varepsilon(s)$, for $0 < t << 1$ we have

$$-\frac{1}{t}\int_0^t \big\langle D\varphi\big(\overline{X}_\varepsilon(s)\big), \sigma^t\big(\overline{X}_\varepsilon(s)\big)\overline{\alpha}(s)\big\rangle_n ds \le -\frac{1}{t}\int_0^t \big\langle D\varphi(x), \sigma^t(x)\overline{\alpha}(s)\big\rangle_n ds + o(1)$$
$$= -\frac{1}{t}\int_0^t \big\langle \sigma(x)D\varphi(x), \overline{\alpha}(s)\big\rangle_m ds + o(1) \le |\sigma(x)D\varphi(x)|_m + o(1), \quad (2.32)$$

(in fact $|\overline{\alpha}(s)|_m = 1$ a.e. $s \in [0,T]$).
From (2.31) and (2.32), it follows that

$$1 - \varepsilon \le |\sigma(x)D\varphi(x)|_m + o(1)$$

Passing to the limit as $t \to 0^+$, we find

$$|\sigma(x)D\varphi(x)|_m \ge 1 - \varepsilon.$$

Passing then to the limit as $\varepsilon \to 0^+$, we can conclude $|\sigma(x)D\varphi(x)|_m \ge 1$. □

Corollary 2.41. *Note that $\sigma(x)Du$ is the horizontal gradient Xu, made w.r.t. the sub-Riemannian geometry induced by the lines of $\sigma(x)$.*
So whenever Rademacher's Theorem holds, we get that the Carnot-Carathéodory distance is an almost everywhere solution too, which means that for any fixed $y \in \mathbb{R}^n$,

$$|Xd_\sigma(x,y)| = 1, \quad a.e.\ x \in \mathbb{R}^n.$$

Remark 2.42. *In order to get the a.e. existence of the horizontal gradient Xu starting from the local d-Lipschitz continuity, we need to use a suitable sub-Riemannian generalization of Rademacher's Theorem. Rademacher's Theorem was generalized by Pansu in 1989 ([80]) to the Carnot groups. In [77] Monti and Serra-Cassano proved the same result under more general assumptions on the vector fields Rademacher's Theorem holds also in Carnor-Carathéodory spaces. In particular they require that the vector fields have the following structure:*

$$X_j(x) = \partial_j + \sum_{i=m+1}^{n} \alpha_{i\,j}(x)\partial_i, \quad \text{for any } j = 1, ..., m, \qquad (2.33)$$

which includes any Carnot groups but also Grušin-type spaces.

By using the same techniques applied in the proof of Theorem 2.40, we can get the existence whenever the subdifferential of the Hamiltonian has the form

$$\partial H_0(x,0) = \sigma^t(x)\overline{B_1^m(0)} + \lambda A^t(x)\overline{B_1^n(0)},$$

with $A(x)$ invertible $n \times n$ matrix.

Moreover there exists a paper by Stefania Bortoletto ([19]) where a Dynamical Programming Principle is used in order to prove that the minimal-time function is an viscosity solution for the associated eikonal equation. That work covers our case whenever the subdifferential of the Hamiltonian is a global Lipschitz multi-function w.r.t. the space variable.

2.3 The Hopf-Lax function.

The Hopf-Lax formula has been introduced by Hopf, in 1965, in order to solve a Cauchy problem for nonlinear first-order equations, in the context of the almost everywhere solutions.
As we remarked in Sec. 2.1.1, the Hopf-Lax formula is also very useful to prove existence of viscosity solutions for a Cauchy-Hamilton-Jacobi problems as (2.14). This method is in particular very used in optimal control theory. where the authors show that the formulas, introduced by Hopf, are viscosity solutions of problem (2.14), in the simple case when $H = H(p)$.

2.3.1 Optimal control theory and Hopf-Lax formula.

Let us start looking at a Hamilton-Jacobi-Bellman equation, which mans to consider Hamiltonians of the form

$$H(x,p) = \sup_{a\in\mathcal{A}}[-F(x,a)\cdot p - L(x,a)], \tag{2.34}$$

where $\mathcal{A} \subset \mathbb{R}^m$ is a closed set, $F : \mathbb{R}^n \times \mathcal{A} \to \mathbb{R}^n$ and $L : \mathbb{R}^n \times \mathcal{A} \to \mathbb{R}$ are Lipschitz continuous functions in the variable x, uniformly w.r.t. $a \in \mathcal{A}$.

Remark 2.43. *Since the Hamiltonian* (2.34) *is defined as envelope of functions affine w.r.t. the gradient-variable p, then $H(x,p)$ is convex w.r.t. p.*

Starting from $H(x,p)$ as in (2.34), we can associate to this a functional in the following way: for any fixed function $g : \mathbb{R}^n \to \mathbb{R}$,

$$J(t,x;\alpha) := \int_0^t L(y(s,x;\alpha),\alpha)ds + g(y(t,x;\alpha)), \tag{2.35}$$

where $y(t,x;\alpha)$ are the solutions of the family of control systems

$$\begin{cases} \dot{y}(t) = F(y(t),\alpha(t)), \\ y(0) = x, \end{cases}$$

with $\alpha : [0,+\infty) \to \mathcal{A}$ measurable function. Note that $\alpha(t)$ is usually called control function.

Definition 2.44. *The* value function *associated to the functional J is*

$$v(t,x) = \inf_{\alpha \in \mathcal{A}} J(t,x;\alpha). \tag{2.36}$$

We are interested in studying the value function (2.36) and show that this is a viscosity solution for the Cauchy problem (2.14) with $H(x,p)$ defined as in (2.34).
Let us recall the following result of existence and uniqueness.

Theorem 2.45. *Let $H(x,p)$ be as in* (2.34) *and let us assume that the initial data g is bounded and uniformly continuous. Then the value function* (2.36) *is a viscosity solution for the Cauchy problem* (2.14)*. Moreover, $v(t,x)$ is the unique continuous and bounded viscosity solution.*

For a proof of the previous result we refer to [7] (Proposition I.3.5 and Theorem I.3.7). We just point out that the following semigroup identity is key, i.e. one needs first to prove that for any $x \in \mathbb{R}^n$ and $0 < \tau \leq t$

$$v(t,x) = \inf_{\alpha \in \mathcal{A}} \left\{ \int_0^\tau L(y(s,x;\alpha),\alpha(s))ds + v(y(\tau,x;\alpha),t-\tau) \right\}.$$

We now consider a general Hamiltonian $H = H(p)$, depending only on the gradient variable and convex. We can represent $H(p)$ as (2.34) by using the Legendre-Fenchel transform (see Appendix A). The Legendre-Fenchel transform is defined as

$$H^*(p) = \sup_{q \in \mathbb{R}^n} [q \cdot p - H(q)]. \tag{2.37}$$

Whenever the Hamiltonian is convex, the Legendre-Fenchel transform is involutive, which means $H^{**}(p) = H(p)$. This implies

$$H(p) = \sup_{q \in \mathbb{R}^n} [q \cdot p - H^*(q)]. \tag{2.38}$$

By using (2.37) we can associate to $H(p)$ the following functional

$$J(t, x; \alpha) = \int_0^t H^*(y(s, x; \alpha))ds + g(y(t, x; \alpha)),$$

where $y(t, x; \alpha)$ are solutions of the family of control systems

$$\begin{cases} \dot{y}(t) = \alpha(t), \\ y(0) = x, \end{cases}$$

i.e.

$$y(t, x; \alpha) = x + \int_0^t \alpha(s)ds.$$

To sum up we get the following (formal) representative formula, for the solution of the Cauchy problem (2.14)

$$v(t, x) = \inf_{\alpha \in \mathcal{A}} \left\{ \int_0^t H^*(y(s, x; \alpha))ds + g(y(t, x; \alpha)) \right\}. \tag{2.39}$$

The difficulty in using directly formula (2.39) is that the formula is given by an infumum in an infinite dimensional space.

Definition 2.46. *We call* Hopf-Lax function *the following marginal function:*

$$u(t, x) = \inf_{y \in \mathbb{R}^n} \left[g(y) + tH^* \left(\frac{x - y}{t} \right) \right]. \tag{2.40}$$

M. Bardi and L.C. Evans in [8] starting from a work by E. Hopf [54] proved that the value function (2.39) coincides with the Hopf-Lax function (2.40) and solves the Cauchy problem (2.14) in the viscosity sense, whenever the Hamiltonian depends only on the gradient and it is convex. We briefly mention that in [54] Hopf started by remarking that the solution of the Cauchy problem with an affine initial data $g(x) = \alpha \cdot x + \beta$, (with $\alpha \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$) is given by the affine function

$$u(t, x) = g(x) - tH(Dg(x)).$$

So the idea is that whenever we start from a generic initial data g, we can look for affine solutions of the form

$$v^{y,z}(t,x) = g(z) + y \cdot (x-z) - tH(y), \tag{2.41}$$

varying $y, z \in \mathbb{R}^n$.
For any y and z chosen, the function $v^{y,z}$ satisfies the evolutive equation $u_t + H(Du) = 0$ but not the initial data $g(x)$. Hopf's idea is to look as candidate solutions at the envelopes of the affine functions (2.41), i.e.

$$u_1(t,x) = \inf_{z\in\mathbb{R}^n} \sup_{y\in\mathbb{R}^n} v^{y,z}(t,x) \tag{2.42}$$

$$u_2(t,x) = \sup_{y\in\mathbb{R}^n} \inf_{z\in\mathbb{R}^n} v^{y,z}(t,x) \tag{2.43}$$

and he proved that u_1 and u_2 are almost everywhere solutions (so called L-solutions to indicate that they are Lipschitz continuous) for the Cauchy problem (2.14) under suitable assumptions on the Hamiltonian and on the initial datum (see Theorems 5a and 5b).
Using the formulas suggested by Hopf Bardi and Evans have proved the following results.

Theorem 2.47 ([8], Theorem 2.1)**.** *If the Hamiltonian $H = H(p)$ is a convex function and g is uniformly Lipschitz continuous, then the function u_1 defined by* (2.42) *is the unique uniformly continuous viscosity solution for the Cauchy problem* (2.14).

Proof. We give only a sketch of this proof. First we can note (see [8], Lemma 2.1) that, set

$$x(t) = x - \int_0^t z(s)ds,$$

for any $z \in L^1 = L^1([0,t];\mathbb{R}^n)$ the unique uniformly continuous viscosity solution of (2.14) is given by

$$\widehat{u}(t,x) = \inf_{z\in L^1} \left\{ \int_0^t H^*(z(s))ds + g(z(t)) \right\} = u(t,x). \tag{2.44}$$

To get the result, it is sufficient to show that $u_1(t,x) = \widehat{u}(t,x)$.
We remark that the function defined in (2.42) can be rewritten as

$$u_1(t,x) = \inf_{z\in\mathbb{R}^n} \left[g(z) + tH^*\left(\frac{x-z}{t}\right)\right]. \tag{2.45}$$

Choosing $z = x(t)$ and by Jensen's inequality, it follows that

$$\widehat{u}(t,x) \geq u_1(t,x).$$

On the other side, set $z(s) = \frac{x-z}{t}$ for $0 \leq s \leq t$, we get

$$\widehat{u}(t,x) \leq \inf_{z\in\mathbb{R}^n} \left[tH^* \left(\frac{x-z}{t} \right) + g(z) \right] = u_1(t,x).$$

Therefore we can conclude $u_1 = \widehat{u} = u$. □

The proof of the next result is a bit more complicate and so we omit that and refer directly to [8].

Theorem 2.48 ([8], Theorem 3.1). *Assuming that the Hamiltonian $H(x,z,p) = H(p)$ is continuous and g is uniformly Lipschitz continuous and convex, then the function u_2 defined by* (2.43) *is the unique uniformly continuous viscosity solution for the Cauchy problem* (2.14).

From the two previous theorems, it is possible to deduce the following result.

Theorem 2.49 ([8], Proposition 4.1). *Let $H(x,z,p) = H(p)$ be a continuous and convex Hamiltonian and g a uniformly Lipschitz and convex initial datum, then $u_1 = u_2 =: u$ is the unique uniformly continuous viscosity solution for the Cauchy problem* (2.14).
Moreover, if we assume

$$|D^2u| \in L^\infty(\mathbb{R}^n),$$

then u is a classic solution of (2.14) *and its second-order derivatives D^2u, Du_t and u_{tt} are bounded in $[0,+\infty) \times \mathbb{R}^{n+1}$.*

We conclude remarking that the Hopf-Lax function (2.40) gives a representative formula whenever the Hamiltonian depends only on the gradient and the initial datum is continuous and bounded (see [7, 47]), or also if the initial datum is merely lower semicontinuous (see [3]).
Our aim is to generalize this last result to Hamiltonians depending also on the space (see Sec. 2.4.2).

2.3.2 Some properties of the Euclidean Hopf-Lax function.

In this subsection we recall some properties of the classic Hopf-Lax function (2.40), refering mainly to [3, 47]. Since $|x-y|$ is the Euclidean distance between x and y, sometimes we call (2.40) Euclidean Hopf-Lax function.

First we remark that, for bounded and lower semicontinuous data g (briefly $g \in BLSC(\mathbb{R}^n)$), the infimum (2.40) is indeed a minimum.

Lemma 2.50. *Let be $g \in BLSC(\mathbb{R}^n)$ (i.e. bounded and lower semicontinuous), then for any $x \in \mathbb{R}^n$ and $t > 0$ the infimum in (2.40) is a minimum and it is attended inside a closed ball centered at x and with radius $R(t)$, depending only on the Hamiltonian. Moreover $R(t)$ is a non-decreasing function for $t > 0$.*

So far we omit the proof and give this directly for the general metric Hopf-Lax function (Lemma 2.63).
The idea to get most of the properties of the Hopf-Lax function is to play in a suitable way with the point where the minimum is attended. Next we will skip the proofs when they are similar to the corresponding metric result, showing those directly in the general case.
First by simply choosing $y = x$ in (2.40) we can observe that

$$u(t,x) \leq g(x),$$

for any $x \in \mathbb{R}^n$ and $t \geq 0$.
Moreover for any $g \in LSC(\mathbb{R}^n)$ (i.e. lower semicontinuous)

$$u(0,x) = g(x), \quad \forall\, x \in \mathbb{R}^n,$$

and, if we assume that there exists some constant $C > 0$ such that

$$g(x) \geq -C(1+|x|),$$

then an analogously superlinear-estimate holds for the Hopf-Lax function: More precisely, there exists a constant $C' > 0$ such that

$$u(t,x) \geq -C'(1+t+|x|).$$

For a proof of the previous results we refer to [3] (Theorem 5.2).

Then, using the geodesics, that in the Euclidean case are the straight lines, it is possible to prove the following functional identity.

Lemma 2.51 ([47], Lemma 3.3.1)**.** *For any $x \in \mathbb{R}^n$ and $0 \leq s < t$, the Hopf-Lax function $u(t,x)$ defined by* (2.40) *satisfies*

$$u(t,x) = \min_{y \in \mathbb{R}^n} \left\{ u(s,y) + (t-s) H^* \left(\frac{x-y}{t-s} \right) \right\}. \tag{2.46}$$

Remark 2.52. *From the functional identity* (2.46) *it follows that the Hopf-Lax function is non-increasing in $t > 0$.*

Moreover identity (2.46) is very useful in order to prove the local Lipschitz continuity of $u(t,x)$ in time (while the same regularity in space much easier to get).

Proposition 2.53. *Let be $g \in BC(\mathbb{R}^n)$ (i.e. bounded and continuous in $\mathbb{R}^n$) and let us assume that the Hamiltonian $H(x,z,p) = H(p)$ is convex, then the Hopf-Lax function* (2.40) *is locally Lipschitz continuous in $[0,+\infty) \times \mathbb{R}^n$.*

Proof. We first show the local Lipschitz continuity in space.
By Lemma 2.50 we can choose a point $\overline{y}$ where the Hopf-Lax function $u(t,y)$ attends the minimum, then

$$\begin{aligned} u(t,x) - u(t,y) &\leq g(\overline{y}) + tH^* \left(\frac{x-\overline{y}}{t} \right) - g(\overline{y}) - tH^* \left(\frac{y-\overline{y}}{t} \right) \\ &= t \left[H^* \left(\frac{x-\overline{y}}{t} \right) - H^* \left(\frac{y-\overline{y}}{t} \right) \right]. \end{aligned}$$

Since $H^*(p)$ is convex, we can use that convex functions are always local Lipschitz continuity to get the same regularity for the Hopf-Lax function.
The main problem is that the point $\overline{y}$ depends on t. Nevertheless, using the non decreasing-property of the radius $R(t)$ (see Lemma 2.50), it is possible to show that, fixed a compact subset of $(0,+\infty) \times \mathbb{R}^n$, there exists a compact $K \subset \mathbb{R}^n$ which does not depend on t and such that $\frac{x-\overline{y}}{t}$ and $\frac{y-\overline{y}}{t}$ belong to K

(we will give more details on this claim proving the analogous result in the general metric setting). Therefore

$$u(t,x) - u(t,y) \leq tLip_{loc}(H^*) \left| \frac{x - \overline{y} - y + \overline{y}}{t} \right| = Lip_{loc}(H^*)|x - y|.$$

Swapping x and y, we get the local Lipschitz continuity in space. In order to get the same regularity in time, we use the functional identity proved in Lemma 2.51 and the local Lipschitz continuity w.r.t. x. $\square$

Without using the local Lipschitz continuity of the convex function $H(p)$, it is possible to get a global Lipschitz continuity starting from Lipschitz initial data. The proof of this result is very easy and seems very close to the previous one. Nevertheless, it is not applicable to the general metric case.

Proposition 2.54. *Let be $g \in \mathrm{Lip}(\mathbb{R}^n)$, then the Hopf-Lax function* (2.40) *is locally Lipschitz continuous in $[0, +\infty) \times \mathbb{R}^n$.*

Proof. As above, we prove only the Lipschitz continuity in space. In fact, to get the same regularity in time we only need to use Lemma 2.51.
To get the Lipschitz continuity w.r.t. x, we choose a point $\overline{y} \in \mathbb{R}^n$ where the minimum in $u(t,y)$ is attended, so that

$$u(t,x) - u(t,y) \leq g(z) + tH^*\left(\frac{x-z}{t}\right) - g(\overline{y}) - tH^*\left(\frac{y-\overline{y}}{t}\right),$$

for any $z \in \mathbb{R}^n$.
Set $z = x - y + \overline{y}$ and using the Lipschitz continuity of g, we get

$$u(t,x) - u(t,y) \leq g(x-y+\overline{y}) - g(\overline{y}) \leq \mathrm{Lip}(g)|-\overline{y}+x-y+\overline{y}| = \mathrm{Lip}(g)|x-y|.$$

Swapping x and y we can conclude. $\square$

Remark 2.55. *Proposition 2.54 still holds if the Hopf-Lax function is defined starting from a generic function $\Phi = H^*$ not necessary convex.*

The previous proof cannot be applied to a non Euclidean setting. In fact to delete the unknown point $\overline{y}$, the commutativity of the operation plays an important role.

We next recall a well-known link between the Hopf-Lax function and problems of calculus of variations, using essentially the geodesic-properties of the Euclidean space. We define

$$v(t,x)=\inf\left\{\int_0^t H^*(\dot\gamma(s))ds+g(y)\,\middle|\,\gamma:(0,t)\to\mathbb{R}^n\,\text{a.c. with}\,\gamma(0)=x\right\}. \tag{2.47}$$

Proposition 2.56 ([47], Theorem 3.3.4)**.** *If $x\in\mathbb{R}^n$ and $t>0$, then the solution of the minimization problem* (2.47) *is the Hopf-Lax function* (2.40).

We conclude the section recalling that for the Hopf-Lax function (2.40) several properties of semiconcavity hold (see for example Lemma 3.3.3 and Lemma 3.3.4 in [47] or [24] Sec. 1.6). To generalize these properties to the Carnot-Carathéodory case is not immediately. In fact, first one would need to define a suitable notion of convexity (e.g. see [73] for a definition of horizontal convexity) and investigate the property of the Carnot-Carathéodory distance w.r.t. this notion.

In the next section we define the metric Hopf-Lax function and we stuady the main properties of this formula

2.4 The metric Hopf-Lax function.

Now we want to define a metric Hopf-Lax formula associated to a generalized distance $d(x,y)$ (see Definition 2.28).

Definition 2.57. *Let be $g:\mathbb{R}^n\to\mathbb{R}$ and $\Phi:[0,+\infty)\longrightarrow[0+\infty)$ be convex, non-decreasing and such that $\Phi(0)=0$. The* metric Hopf-Lax function *(associated to the generalized distance $d(x,y)$ and the function $g(x)$) is defined as*

$$u(t,x)=\inf_{y\in\mathbb{R}^n}\left[g(y)+t\Phi^*\left(\frac{d(x,y)}{t}\right)\right], \tag{2.48}$$

where Φ^* is the Legendre-Fenchel transform of Φ, i.e.

$$\Phi^*(t)=\sup_{s\geq 0}\{ts-\Phi(s)\}.$$

Remark 2.58. *Note that the Legendre-Fenchel transform of* Φ *is a positive function* $\Phi^* : [0,+\infty) \to [0,+\infty]$ *convex, non decreasing and such that* $\Phi^*(0) = 0$ *(see Appendix A, for more details).*

2.4.1 Properties of the metric Hopf-Lax function.

We begin starting from an easy upper bound.

Remark 2.59. *Set* $y = x$ *in* (2.48)*, we get* $u(t,x) \leq g(x)$ *for any* $t > 0$*. Then we can deduce the behavior for large times of* $u(t,x)/t$*, in fact*

$$\limsup_{t\to+\infty} \frac{1}{t} u(t,x) \leq 0.$$

From now on, we look at lower semicontinuous data g.
Moreover we indicate simply by $d(x)$ the generalized distance from the origin of $\mathbb{R}^n$ to a point x.
The next properties are very important in order to get the existence-result proved in Sec. 2.4.2. To get these we need to require a d-superlinear estimate for the datum g, i.e. we assume that there exists a constant $C > 0$ such that

$$g(x) \geq -C(1 + d(x)), \quad x \in \mathbb{R}^n. \tag{2.49}$$

Lemma 2.60. *Let* $d(x,y)$ *be a generalized distance inducing the Euclidean topology on* $\mathbb{R}^n$ *and* $g \in LSC(\mathbb{R}^n)$ *be such that* (2.49) *holds. The metric Hopf-Lax function* (2.48) *weak lower converges to* g *in the sense of Barles-Perthame (see [3, 11]) i.e.*

$$\liminf_{(t,x)\to(0^+,\overline{x})} u(t,x) = \inf\left\{ \liminf_{n\to\infty} u(t_n,x_n) \,|\, (t_n,x_n) \to (0^+,\overline{x}) \right\} = g(\overline{x}). \tag{2.50}$$

Proof. We begin remarking that the Hopf-Lax function (2.48) can be rewritten as

$$u(t,x) = \inf_{y\in\mathbb{R}^n} \left[g(y) + \sup_{p\geq 0} [pd(x,y) - t\Phi(p)] \right]. \tag{2.51}$$

Using that Φ is non decreasing, it is easy to see that for any $r \geq 0$

$$\begin{aligned} \sup_{p\geq 0} [d(x,y)\,p - t\Phi(p)] &\geq \max_{p\in[0,r]} [d(x,y)\,p - t\Phi(p)] \\ &\geq d(x,y) \max_{p\in[0,r]} p - t \max_{p\in[0,r]} \Phi(p) = d(x,y)\,r - t\Phi(r). \end{aligned} \tag{2.52}$$

So (2.51) becomes

$$u(t,x) \geq \inf_{y\in\mathbb{R}^n} \left[g(y) + r\,d(x,y) - t\Phi(r)\right], \quad \forall\, r \geq 0. \tag{2.53}$$

We have assumed that the initial datum g satisfies the d-superlinear estimate (2.49) from below; therefore using the triangle inequality

$$d(0,y) \leq d(0,x) + d(x,y),$$

we get

$$\begin{aligned} u(t,x) &\geq \inf_{y\in\mathbb{R}^n} \left[-C(1+d(y)) + rd(x,y) - t\Phi(r)\right] \\ &\geq \inf_{y\in\mathbb{R}^n} \left[(r-C)\,d(x,y) - C - C\,d(x) - t\Phi(r)\right], \quad \forall\, r \geq 0. \end{aligned} \tag{2.54}$$

Moreover g is lower semicontinuous, so for any fixed $\varepsilon > 0$ there exists $\delta > 0$ such that if $d(\overline{x}, y) < 2\delta$ then

$$g(y) \geq g(\overline{x}) - \varepsilon. \tag{2.55}$$

We can first choose $\overline{r} > C$ such that

$$\big(\overline{r} - C\big)\delta - C(d(\overline{x}) + \delta) - C \geq g(\overline{x})$$

and then $\tau > 0$ such that

$$\tau\Phi(\overline{r}) \leq \varepsilon.$$

Therefore for $0 < t \leq \tau$ and $x \in B^d(\overline{x}, \delta)$ we get the following estimate:

$$\begin{aligned} &\big(\overline{r} - C\big)d(x,y) - C - C\,d(x) - t\Phi(\overline{r}) \\ &\qquad \geq \big(\overline{r} - C\big)\delta - C - C\,d(\overline{x}) - C\delta - t\Phi(\overline{r}) \\ &\qquad\qquad \geq g(\overline{x}) - \varepsilon, \quad \forall\, y \in \mathbb{R}^n \backslash B^d_{2\delta}(\overline{x}). \end{aligned}$$

Setting $B = B^d_{2\delta}(\overline{x})$, we can conclude

$$\inf_{y\in\mathbb{R}^n\backslash B} \left[g(y) + t\Phi^*\left(\frac{d(x,y)}{t}\right)\right] \geq g(\overline{x}) - \varepsilon.$$

It remains to estimate the infimum inside the open ball B.
If we write (2.52) with $r = 0$, we get

$$g(y) + t\Phi^*\left(\frac{d(x,y)}{t}\right) \geq g(y) - t\Phi(0) = g(y),$$

since $\Phi(0) = 0$.
Moreover we know that $y \in B$, i.e. $d(\overline{x}, y) < 2\delta$, then

$$g(y) \geq g(\overline{x}) - \varepsilon.$$

Therefore we can conclude that for any $x \in B^d_\delta(\overline{x})$ and $0 \leq t \leq \tau$

$$u(t, x) \geq g(\overline{x}) - \varepsilon.$$

This means that

$$\liminf_{(t,x)\to(0^+,\overline{x})} u(t, x) \geq g(\overline{x}). \tag{2.56}$$

To check the reverse inequality is not difficult. In fact, if we look at a sequence $(t_n, \overline{x})$ with $t_n \to 0^+$, we have

$$\liminf_{(t,x)\to(0^+,\overline{x})} u(t, x) \leq \liminf_{n\to\infty} u(t_n, \overline{x}). \tag{2.57}$$

Since $\Phi \geq 0$, we note that

$$\begin{aligned} u(t_n, \overline{x}) = \inf_{y\in\mathbb{R}^n} \sup_{p\geq 0} \left[g(y) + d(\overline{x}, y)p - t_n\Phi(p)\right] &\leq \inf_{y\in\mathbb{R}^n} \sup_{p\geq 0} \left[g(y) + d(\overline{x}, y)p\right] \\ &\leq \sup_{p\geq 0} \left[g(\overline{x}) + d(\overline{x}, \overline{x})p\right] = g(\overline{x}). \end{aligned}$$

So $u(t, \overline{x}) \leq g(\overline{x})$ and then we can conclude that (2.57) holds. □

Lemma 2.61. *Under the assumptions of Lemma 2.60, there exists a constant* $C' > 0$ *such that*

$$u(t, x) \geq -C'(1 + d(x) + t). \tag{2.58}$$

Proof. In order to prove this lemma, we need only to write the estimate (2.54) with $r = C > 0$, i.e.

$$u(t, x) \quad \geq \quad \inf_{y\in\mathbb{R}^n} [-C\, d(x) - C - t\Phi(C)] \quad = \quad -C\, d(x) \ - \ C \ - \ t\Phi(C),$$

which gives (2.58) with $C' = \max\{C, \Phi(C)\}$. □

Lemma 2.62. *Under the assumptions of Lemma 2.60, the metric Hopf-Lax function* (2.48) *is lower semicontinuous in* $[0, +\infty) \times \mathbb{R}^n$.

Proof. In order to prove that u is a lower semicontinuous function, we have to show that its sublevel sets are closed.
We consider a sequence (t_k, x_k) such that $u(t_k, x_k) \leq \gamma$ for some $\gamma \in \mathbb{R}$. Assuming that $(t_k, x_k) \to (t, x)$, we have to prove that $u(t, x) \leq \gamma$.
By Lemma 2.60 and using that g is lower semicontinuous, we can look only at $t > 0$.
Let $\{y_k^n\}$ be a minimizing sequence for the Hopf-Lax function $u(t_k, x_k)$.
Since g and Φ^* are lower semicontinuous functions and applying (2.54) with $r = 1 + C$, we get the following inequality-chain

$$\gamma \geq u(t_k, x_k) = \liminf_{n\to\infty} \left[g(y_k^n) + t_k \Phi^* \left(\frac{d(x_k, y_k^n)}{t_k} \right) \right]$$
$$\geq \liminf_{n\to\infty} d(x_k, y_k^n) - C - C\, d(x_k) - t_k \Phi(C). \quad (2.59)$$

We want to show that the sequence $\{y_k^n\}$ converges to some point y_k, as $n \to +\infty$ for any fixed $k > 0$. Set $K := C + Cd(x_k) + t_k\Phi(C)$, by (2.59) and using the definition of lower limit, we find that for any $\varepsilon > 0$ there exists $\overline{n}(\varepsilon) \in \mathbb{N}$ such that

$$d(x_k, y_k^n) \leq K + \gamma + \varepsilon, \quad \forall\, n \geq \overline{n}(\varepsilon).$$

We can choose $\varepsilon = 1$, then we have $y_k^n \in \overline{B}_k = \overline{B_R^d(x_k)}$ with $R := K + \gamma + 1$ for any $n \geq \overline{n}(1)$. We have assumed that $d(x, y)$ induces on $\mathbb{R}^n$ the Euclidean topology, hence by Corollary 1.53 for any $k > 0$ fixed the closed d-ball $\overline{B}_k$ is a compact set. This implies that $\{y_k^n\}_n$ attends a subsequence convergent to some point $y_k \in \overline{B}_k$.
Remember that by assumptions $x_k \to x$ and $t_k \to t$. Then, there exists $\overline{k} \in \mathbb{N}$ such that for any $k > \overline{k}$ we have $x_k \in \overline{B_1^d(x)}$ and $t_k < t + 1$. Therefore $y_k \in \overline{B_R^d(x)}$, where $R = 1 + \gamma + C(d(x) + 1) + C + \Phi(C)(t + 1) + 1$, for $k > \overline{k}$. Then there exists a subsequence, which we still indicate by y_k, and a point $y \in \overline{B_R^d(x)}$ such that $y_k \to y$, as $k \to +\infty$. In order to conclude the proof, we need only to use the lower semicontinuity of g, Φ^* and $d(x, y)$.
In fact

$$\gamma \geq u(t_x, x_k) = \liminf_{k\to+\infty} \left[g(y_k) + t_k \Phi^* \left(\frac{d(x_k, y_k)}{t_k} \right) \right] \geq g(y) + t\Phi^* \left(\frac{d(x, y)}{t} \right)$$
$$\geq \inf_{y\in\mathbb{R}^n} \left[g(y) + t\Phi^* \left(\frac{d(x, y)}{t} \right) \right] = u(t, x),$$

so that the sublevels of u are closed sets. $\square$

Next we investigate the Lipschitz properties of the metric Hopf-Lax function (2.48). First, we show that for any datum g, bounded and lower semicontinuous, the infimum in formula (2.48) is a minimum.

Lemma 2.63. *Let $d(x,y)$ be a generalized distance and $g \in BLSC(\mathbb{R}^n)$. Then for any $x \in \mathbb{R}^n$ and $t > 0$ the infimum in (2.48) is a minimum. Moreover, it is attended in a closed d-ball centered at x and with radius $R(t)$ depending on Φ and g and non decreasing w.r.t. $t > 0$.*

Proof. We need only to prove that, for any $x \in \mathbb{R}^n$ and $t > 0$ fixed, there exists a radius $R(t)$ large enough so that

$$g(y) + t\Phi^*\left(\frac{d(x,y)}{t}\right) \geq \|g\|_\infty, \quad \forall\, y \in \mathbb{R}^n \backslash \overline{B^d_{R(t)}(x)}. \tag{2.60}$$

Then, since

$$u(t,x) \leq g(x) \leq \|g\|_\infty,$$

the infimum is attended inside the closed d-ball $\overline{B^d_{R(t)}(x)}$ and so, by the lower semicontinuity of the function $f(y) = g(y) + t\Phi^*\left(\frac{d(x,y)}{t}\right)$, it is a minimum. In order to prove (2.60), remark that $\Phi^*(\tau)$ is convex, then there exists a supporting line $m\tau + q$. Moreover $\Phi^*(0) = 0$ which implies $q \leq 0$ and, by the non-decreasing property of Φ^*, we can also assume $m > 0$. Chosen

$$R(t) = \frac{2\|g\|_\infty - tq}{m},$$

for any $y \in \mathbb{R}^n \backslash \overline{B^d_{R(t)}(x)}$ we have

$$g(y) + t\Phi^*\left(\frac{d(x,y)}{t}\right) \geq g(y) + md(x,y) + tq \geq -\|g\|_\infty + 2\|g\|_\infty = \|g\|_\infty.$$

Since $q \leq 0$ and $m > 0$, we can conclude that $R(t)$ is non-decreasing w.r.t. $t > 0$. $\square$

Remark 2.64. *When the convex function Φ is a power-function, i.e. $\Phi(t) = \frac{1}{\alpha}t^\alpha$ with $\alpha \geq 1$, then $\Phi^*(t) = \frac{1}{\beta}t^\beta$ with $\beta = \frac{\alpha}{\alpha-1}$ if $\alpha > 1$, while if $\alpha = 1$*

$$\Phi^*(t) = \begin{cases} 0, & 0 \leq t \leq 1, \\ +\infty, & t > 1. \end{cases}$$

So for (convex) power-functions, the Hopf-Lax function is given by

$$u(t,x) = \inf_{y\in\mathbb{R}^n}\left[g(y) + \frac{1}{\beta}\frac{d(x,y)^\beta}{t^{\beta-1}}\right], \tag{2.61}$$

whenever $\alpha > 1$*, and by*

$$u(t,x) = \inf\left\{g(y)\,|\,d(x,y) < t\right\}, \tag{2.62}$$

whenever $\alpha = 1$*.*
Moreover by simple calculations it is possible to show that infimum (2.61) *and infimum* (2.62) *are attended in the closed* d*-ball centered at* x *and with radius* $R(t) = (2\beta)^{\frac{1}{\beta}}t^{\frac{\beta-1}{\beta}}\|g\|_\infty^{\frac{1}{\beta}}$ *and* $R(t) = t$*, respectively.*

Lemma 2.63 is very useful in order to prove the local Lipschitz continuity for the metric Hopf-Lax function.
Recall that a function $f:\mathbb{R}^n \to \mathbb{R}$ is d-Lipschitz continuous w.r.t. a non symmetric distance $d(x,y)$, if and only if, there exists $C > 0$ such that

$$|f(x) - f(y)| \le C\max\{d(x,y), d(y,x)\}.$$

Proposition 2.65. *For any generalized distance* $d(x,y)$*, fixed* $t > 0$*, the metric Hopf-Lax function* (2.48) *is locally* d*-Lipschitz continuous w.r.t.* x*.*

Proof. By Lemma 2.63, we can choose $\overline{y}$ such that

$$u(t,y) = g(\overline{y}) + t\Phi^*\left(\frac{d(y,\overline{y})}{t}\right).$$

Remark that Φ^* is convex and so locally Lipschitz continuous. Hence, for any $K \subset \mathbb{R}^n$ compact, there exists a constant $C(K) > 0$ such that

$$u(t,x) - u(t,y) \le C(K)|d(x,\overline{y}) - d(y,\overline{y})|, \quad \forall\, x,y \in K. \tag{2.63}$$

From the triangle inequality (2.17), it follows that

$$|d(x,\overline{y}) - d(y,\overline{y})| \le \max\{d(x,y), d(y,x)\}.$$

So (2.63) becomes

$$u(t,x) - u(t,y) \le C(K)\max\{d(x,y), d(y,x)\}.$$

Swapping x with y we conclude the proof. $\square$

In order to prove the local Lipschitz continuity w.r.t. t, we need to use the geodesics. Therefore the next results hold in any length-space.
Nevertheless we look only at the Carnot-Carathéodory distances, so that we do not need to define the length of curves by the variation, inside general metric spaces.
We briefly recall that we call *geodesics* any absolutely continuous horizontal curves which realize the minimum (1.14) (see Chapter 1, Definition 1.24).
In particular, for Carnot-Carathéodory distances satisfying the Hörmander condition, $(\mathbb{R}^n, d)$ is a length spaces, which means for any $x, y \in \mathbb{R}^n$ there exists a geodesic γ joining x to y and such that $l(\gamma) = d(x, y)$ (see Theorems 1.72). As in any length space we can assume that the geodesic $\gamma : [0, T] \longrightarrow \mathbb{R}^n$ is parameterized by arc-length, so $l(\gamma) = T$ and $d(\gamma(t), \gamma(s)) = |t - s|$ for $s, t \in [0, l(\gamma)]$ (see [89], Lemma 3.3).
In order to show the local Lipschitz continuity in t, we proceed as in [47]. First we prove a suitable functional identity.

Lemma 2.66 (Functional identity)**.** *Let $d(x, y)$ be a Carnot-Carathéodory distance satisfying the Hörmander condition and $g \in BLSC(\mathbb{R}^n)$, then for any $0 \le s < t$ the Hopf-Lax function* (2.48) *satisfies*

$$u(t, x) = \inf_{y \in \mathbb{R}^n} \left[u(s, y) + (t - s)\Phi^* \left(\frac{d(x, y)}{t - s} \right) \right]. \tag{2.64}$$

Proof. By the usual triangle inequality (2.17) and using the non decreasing-property and the convexity of Φ^*, we get

$$\begin{aligned} \Phi^* \left(\frac{d(x, z)}{t} \right) \le \Phi^* \left(\frac{d(x, y) + d(y, z)}{t} \right) &= \Phi^* \left(\frac{t - s}{t} \frac{d(x, y)}{t - s} + \frac{s}{t} \frac{d(y, z)}{s} \right) \\ &\le \left(1 - \frac{s}{t} \right) \Phi^* \left(\frac{d(x, y)}{t - s} \right) + \frac{s}{t} \Phi^* \left(\frac{d(y, z)}{s} \right), \end{aligned}$$

for any $x, y, z \in \mathbb{R}^n$.
Fixed x, then for any y we choose a minimum point z for $u(s, y)$ (that exists, by Lemma 2.63). Using such a point z, we get

$$u(t, x) \le g(z) + t\Phi^* \left(\frac{d(x, z)}{t} \right) \le u(s, y) + (t - s)\Phi^* \left(\frac{d(x, y)}{t - s} \right).$$

Taking the infimum over $y \in \mathbb{R}^n$, we find the following inequality

$$u(t,x) \le \inf_{y\in\mathbb{R}^n} \left[u(s,y) + (t-s)\Phi^*\left(\frac{d(x,y)}{t-s}\right)\right].$$

In order to prove the reverse inequality, we choose a minimum point w for $u(t,x)$. Let be $T = d(x,w)$, there exists $\gamma : [0,T] \to \mathbb{R}^n$ such that $\gamma(0) = x$, $\gamma(T) = w$ and $d(\gamma(s),\gamma(t)) = t - s$, for any $0 \le s \le t \le T$.
We define $\overline{y} := \gamma(\frac{T(t-s)}{t})$ so that

$$\frac{d(x,\overline{y})}{t-s} = \frac{d(x,w)}{t} = \frac{d(\overline{y},w)}{s},$$

getting the other required inequality. In fact,

$$\inf_{y\in\mathbb{R}^n} \left[u(s,y) + (t-s)\Phi^*\left(\frac{d(x,y)}{t-s}\right)\right] \le u(s,\overline{y}) + (t-s)\Phi^*\left(\frac{d(x,\overline{y})}{t-s}\right) \le g(w)$$
$$+ s\Phi^*\left(\frac{d(\overline{y},w)}{s}\right) + (t-s)\Phi^*\left(\frac{d(x,\overline{y})}{t-s}\right) = g(w) + t\Phi^*\left(\frac{d(x,w)}{t}\right) = u(t,x).$$

□

Remark 2.67. *Choosing $y = x$ in* (2.64)*, by Lemma 2.66 we deduce that the metric Hopf-Lax function* (2.48) *is non-increasing in t.*

Proposition 2.68. *Let be $g \in BLSC(\mathbb{R}^N)$ and let us assume that $t\Phi^*\left(\frac{1}{t}\right)$ is convex and non-increasing for $t > 0$. Then the Hopf-Lax function* (2.48) *associated to a Carnot-Carathéodory distance satisfying the Hörmander condition is locally Lipschitz continuous in $t > 0$.*

Proof. Since $u(t,x)$ is non increasing in t, for any $0 \le s \le t$

$$u(t,x) - u(s,x) \le 0.$$

So we only need to check the estimate from below.
Choosing a minimum point $\overline{y} = \overline{y}(t)$ for $u(t,x)$, then for any $T_1 \le s \le t \le T_2$ we find

$$0 \ge u(t,x) - u(s,x) \ge t\Phi^*\left(\frac{d(x,\overline{y})}{t}\right) - s\Phi^*\left(\frac{d(x,\overline{y})}{s}\right) = \lambda.$$

Fix t and let s be free.
If $d(x,\overline{y}) = 0$, then $u(t,x) = u(s,x)$ and so we have concluded.
Therefore we can assume $d(x,\overline{y}) \neq 0$ and set

$$\tau = \frac{t}{d(x,\overline{y})} \quad \text{and} \quad \sigma = \frac{s}{d(x,\overline{y})},$$

so that

$$\lambda = d(x,\overline{y}) \left[\tau\Phi^* \left(\frac{1}{\tau}\right) - \sigma\Phi^* \left(\frac{1}{\sigma}\right)\right].$$

Using the local Lipschitz continuity, which is true for any convex function, then for any $\widetilde{T} > 0$ there exists $C = C(\widetilde{T}) > 0$ such that

$$\lambda \geq -Cd(x,\overline{y})(\tau - \sigma) = -C(t-s), \quad \widetilde{T} \leq \sigma \leq \tau.$$

If we choose $\widetilde{T} = \frac{T_1}{R(T_2)}$, whenever $s,t \in [T_1, T_2]$, then $\sigma, \tau \in [\widetilde{T}, +\infty)$ (by Lemma 2.63 and the non decreasing-property of the radius $R(t)$,) so that

$$0 \geq u(t,x) - u(s,x) \geq -C\left(\frac{T_1}{R(T_2)}\right)(t-s),$$

for any $T_1 \leq s \leq t \leq T_2$, and this concludes the proof. □

Remark 2.69. *If we look at $\Phi(t) = \frac{1}{\alpha}t^\alpha$ with $\alpha > 1$, by Remark 2.64, it is immediate that $t\Phi^*\left(\frac{1}{t}\right)$ is convex and non-increasing. More in general, this property holds whenever Φ is strictly convex and there exists $(\Phi^*)''(t)$. In fact, from the dual formula for the first-order derivatives of strictly convex functions the non-increasing property follows trivially. Moreover*

$$\left(t\Phi^*\left(\frac{1}{t}\right)\right)'' = (\Phi^*)''\left(\frac{1}{t}\right)\frac{1}{t^2},$$

since Φ^* is strictly convex, then

$$\left(t\Phi^*\left(\frac{1}{t}\right)\right)'' > 0.$$

An example of this kind of functions is $\Phi(t) = \mathrm{e}^t - 1$.
The case $\Phi(t) = t$ instead does not satisfy both the requirements.

In order to conclude the study of the properties of the metric Hopf-Lax function, we point out a link with a problem of calculus of variations, exactly as known in the Euclidean case (see Proposition 2.56).
Therefore we consider the minimization problem

$$v(t,x) = \inf\left\{ \int_0^t \Phi^*(|\dot\gamma(s)|)ds + g(\gamma(t)) \,\middle|\, \gamma \text{ a.c., horizontal, with } \gamma(0) = x \right\}. \tag{2.65}$$

Proposition 2.70. *Let be $g \in LSC(\mathbb{R}^n)$ and $d(x,y)$ be a Carnot-Carathéodory distance satisfying the Hörmander condition then the infimum* (2.65) *coincides with the metric Hopf-Lax function* (2.48).

Proof. By Jensen's inequality, it follows immediately that

$$\Phi^*\left(\frac{1}{t}\int_0^t |\dot\gamma(s)|ds\right) \le \frac{1}{t}\int_0^t \Phi^*(|\dot\gamma(s)|)ds.$$

So for $t > 0$ and for all the a.c. horizontal curve $\gamma : [0,t] \to \mathbb{R}^n$ joining x to a point $y \in \mathbb{R}^n$

$$g(y) + t\Phi^*\left(\frac{d(x,y)}{t}\right) \le g(y) + t\Phi^*\left(\frac{l(\gamma)}{t}\right) \le g(y) + \int_0^t \Phi^*(|\dot\gamma(s)|)ds.$$

Taking the infimum over y, we get $u(t,x) \le v(t,x)$.

To prove the reverse inequality, we must use the length-structure of the space.
Fix $t > 0$ and $y \in \mathbb{R}^n$, there exists a geodesic γ parameterized by arc-length and joining x to y.
Set $T = d(x,y)$ and define $\widetilde\gamma(s) := \gamma(\frac{Ts}{t})$, then $|\dot{\widetilde\gamma}(s)| = \frac{T}{t}|\dot\gamma(\overline{s})| = \frac{T}{t}$, which implies

$$\int_0^t \Phi^*(|\dot{\widetilde\gamma}(s)|)ds = \int_0^t \Phi^*\left(\frac{T}{t}\right)ds = t\Phi^*\left(\frac{d(x,y)}{t}\right).$$

Adding $g(y)$, we get

$$v(t,x) \le g(y) + t\Phi^*\left(\frac{d(x,y)}{t}\right).$$

We can conclude by taking the infimum over $y \in \mathbb{R}^n$. $\square$

Remark 2.71. *By Lemma 2.63 and Proposition 2.70, for a Carnot-Carathéodory distance satisfying the Hörmander condition and* $g \in BLSC(\mathbb{R}^n)$*, it is possible to express the value function of the minimization problem* (2.65) *as a minimum in* $\mathbb{R}^n$.

Remark 2.72. *It is not clear how to use the properties of the datum* g *in order to get some regularity results for the Hopf-Lax function; in fact the ideas used in [47] in the Euclidean case do not work in this is more general case. In particular, we do not know if starting from a global d-Lipschitz continuous datum* g*, then the Hopf-Lax function has the same regularity.*

The Euclidean Hopf-Lax function is called inf-convolution of the function g whenever $H(t) = \frac{1}{2}t^2$ and it is used to approximate the function g as $t \to 0^+$. Analogously, when $\Phi(t) = \frac{1}{2}t^2$ we call the metric Hopf-Lax function (2.48) metric inf-convolution with kernel the distance $d(x, y)$. We will study this special case in Chapter 3, looking in particular at Carnot-Carathéodory distances satisfying the Hörmander condition.

In Remark 2.64 we have explicitly calculate the metric Hopf-Lax function for the convex power-functions $\Phi(t) = \frac{1}{\alpha}t^\alpha$, $\alpha > 1$.
We conclude this subsection giving another example.

Example 2.73. *Let us consider* $\Phi(t) = \mathrm{e}^t - 1$ *for* $t \leq 0$*. It is immediate that* Φ *is a convex, positive, non-decreasing function, with* $\Phi(0) = 0$*. In order to write the associated metric Hopf-Lax function, we need only to calculate the Legendre-Fenchel transform* $\Phi^*(t)$*, i.e. the supremum over* $s \in [0, +\infty)$ *of the function*

$$f(s) = st - \mathrm{e}^s + 1,$$

for any $t \geq 0$ *fixed. First, we can note that*

$$f(0) = 0 \quad and \quad \lim_{s \to +\infty} f(s) = -\infty.$$

Moreover $f'(s) = t - \mathrm{e}^s$ *and so there exists* $\overline{s} > 0$ *such that* $f'(\overline{s}) = 0$ *if and only if* $t > 1$*. This implies* $\sup_{s \geq 0} f(s) = 0$ *whenever* $0 \leq t \leq 1$*, and* $\overline{s} = \ln t$ *for* $t > 1$*. Then for* $t > 1$

$$\sup_{s \geq 0} f(s) = f(\ln t) = t \ln t - t + 1.$$

To sum up, the Hopf-Lax function is given by (2.48), *where*

$$\Phi^*(t) = \begin{cases} 0, & 0 \le t \le 1, \\ t \ln t - t + 1, & t > 1. \end{cases}$$

2.4.2 The Hopf-Lax solution for the Cauchy problem.

In this subsection we consider the Cauchy problem (2.14) for a Hamilton-Jacobi equation, where the Hamiltonian depends on both the space-variable and the gradient-variable, i.e.

$$\begin{cases} u_t + H(x, Du) = 0, & \text{in } \mathbb{R}^n \times (0, +\infty), \\ u = g, & \text{in } \mathbb{R}^n \times \{0\}. \end{cases} \tag{2.66}$$

with $g \in LSC(\mathbb{R}^n)$ and $H(x,p)$ Hamiltonian with the form

$$H(x, p) = \Phi(H_0(x, p)),$$

where H_0 is a geometrical Hamiltonian and Φ is a convex function.
More precisely, we assume that

(H1) $H_0 : \mathbb{R}^{2n} \longrightarrow [0, +\infty)$ is continuous w.r.t. both of the variables, convex and positively homogeneous of degree 1 w.r.t. p.

(H2) $\Phi : [0, +\infty) \longrightarrow [0 + \infty)$ is convex, C^1, non-decreasing, with $\Phi(0) = 0$ and $\lim_{t \to 0^+} \Phi'(t) = 0$.

The model is

$$\Phi(H_0(x, p)) = \frac{1}{\alpha} |\sigma(x) p|^\alpha, \tag{2.67}$$

with $\alpha > 1$ and $\sigma(x)$ $m \times n$-matrix of Hörmander-type.

Remark 2.74. *The assumption* $\lim_{t \to 0^+} \Phi'(t) = 0$ *implies that* Φ^* *is strictly increasing (we will give more details on this claim in Appendix A).*

In this subsection we want to prove that the metric Hopf-Lax function (2.48) is a lower semicontinuous viscosity solution of the Hamilton-Jacobi-Cauchy problem (2.66).

Theorem 2.75 ([44]). *Assume that $d(x,y)$ is a generalized distance inducing on $\mathbb{R}^n$ the Euclidean topology. Assume also that, for any y, $x \mapsto d(x,y)$ is a viscosity solution of the eikonal equation $H_0(x, Du(x)) = 1$ in $\mathbb{R}^n \setminus \{y\}$. Let $g \in LSC(\mathbb{R}^n)$ be such that (2.49) holds and Φ satisfying $(H2)$, then the metric Hopf-Lax function (2.48) is a lower semicontinuous viscosity solution of the Cauchy problem (2.66) and moreover the estimate (2.58) holds.*

Proof. By Lemma 2.62 and Lemma 2.60, we know that $u(t,x)$ is lower semicontinuous in $[0,+\infty) \times \mathbb{R}^n$ and u assumes the initial datum g (in the lower semicontinuous sense). Moreover by Lemma 2.61 estimate (2.61) holds. In order to prove the theorem, it remains only to check that u satisfies Definition 2.27.
We show that u is a LSC-viscosity solution because it is infimum of LSC-viscosity solutions of (2.9). So we must prove that, fixed $y \in \mathbb{R}^n$, the following function

$$v^y(t,x) := g(y) + t\Phi^*\left(\frac{d(x,y)}{t}\right) \tag{2.68}$$

is a LSC-viscosity solution of the evolutive Hamilton-Jacobi equation (2.9).
First we need to introduce a strictly convex approximation of the convex function Φ. To understand why we must proceed so, one can look at an heuristic proof. In fact, we can compute the derivatives of v^y assuming at the everything is smooth. Then we can replace the derivatives in (2.9) and search for the necessary conditions, so that v^y satisfying the equation.
Set $\tau = \frac{d(x,y)}{t} > 0$, then

$$v^y_t(t,x) = \Phi^*(\tau) - \tau\big(\Phi^*\big)'(\tau); \quad Dv^y(t,x) = \big(\Phi^*\big)'(\tau)Dd(x,y).$$

Using the homogeneous property of the eikonal Hamiltonian $H_0(x,p)$, we find

$$0 = \Phi^*(\tau) - \tau\big(\Phi^*\big)'(\tau) + \Phi\Big(\big(\Phi^*\big)'(\tau)H_0(x, Dd(x,y))\Big).$$

Since $d(x,y)$ is a solution of (2.18), we can write $H_0(x, Dd(x,y)) = 1$ and so the function v^y is a solution of (2.9) if and only if the following dual formula holds:

$$-\Phi\Big(\big(\Phi^*\big)'(\tau)\Big) = \Phi^*(\tau) - \tau\big(\Phi^*\big)'(\tau). \tag{2.69}$$

Formula (2.69) is always true when Φ is a strictly convex function.
So we introduce a strictly convex approximation Φ_δ of the original convex

function Φ. Set

$$\Phi_\delta(s) := \Phi(s) + \frac{\delta}{2}s^2, \quad \text{for } \delta > 0,$$

we prove that v_δ^y, defined replacing Φ_δ^* to Φ^* in (2.68), is a lower semicontinuous viscosity solution of the corresponding Hamilton-Jacobi equation, that is

$$v_t(t,x) + \Phi_\delta\big(H_0(x, Dv(t,x)\big) = 0, \text{ in } (0,+\infty)\times\mathbb{R}^n. \tag{2.70}$$

It is immediate that v_δ^y is lower semicontinuous.
Let $\varphi \in C^1$ be such that $v_\delta^y - \varphi$ has local minimum (of 0) at (t_0, x_0), i.e. there exists $r > 0$ and $0 < \overline{t} < t_0$ such that for any $x \in B_r(x_0)$ and $t_0 - \overline{t} < t < t_0 + \overline{t}$

$$v_\delta^y(t,x) - \varphi(t,x) \geq v_\delta^y(t_0,x_0) - \varphi(t_0,x_0) = 0. \tag{2.71}$$

Writing (2.71) at $x = x_0$, we get that $T(t) := v_\delta^y(t,x_0) - \varphi(t,x_0)$ has a local minimum at $t = t_0$. Moreover $T \in C^1$, then $\dot{T}(t_0) = 0$, i.e.

$$\varphi_t(t_0,x_0) = \big(v_\delta^y(t,x)\big)_t(t_0,x_0) = \Phi_\delta^*\left(\frac{d(x_0,y)}{t_0}\right) - t_0\big(\Phi_\delta^*\big)'\left(\frac{d(x_0,y)}{t_0}\right)\frac{d(x_0,y)}{t_0^2}. \tag{2.72}$$

Since Φ_δ is strictly convex, the dual formula (2.69) holds. Writing this formula at the point $s = \frac{d(x_0,y)}{t_0}$, (2.72) becomes

$$\varphi_t(t_0,x_0) = -\Phi_\delta\left((\Phi_\delta^*)'\left(\frac{d(x_0,y)}{t_0}\right)\right). \tag{2.73}$$

We assume $x_0 \neq y$ and check the requirement for viscosity solutions using (2.73), i.e. we need to verify that

$$\big(\Phi_\delta^*\big)'\left(\frac{d(x_0,y)}{t_0}\right) = H_0(x_0, D\varphi(t_0,x_0)). \tag{2.74}$$

Since $x_0 \neq y$, we can use the fact that $d(x,y)$ is a viscosity solution of the associated eikonal equation (2.18) at the point x_0.
So, fixed $t = t_0 > 0$, by (2.71) we have that for any $x \in B_r(x_0)$

$$\Phi_\delta^*\left(\frac{d(x,y)}{t_0}\right) - \frac{1}{t_0}\varphi(t_0,x) \geq \Phi_\delta^*\left(\frac{d(x_0,y)}{t_0}\right) - \frac{1}{t_0}\varphi(t_0,x_0) = 0. \tag{2.75}$$

Note that in the previous inequality we have replaced the starting test function φ by the translated function $\widetilde{\varphi} = \varphi - t_0 g(y)$. In fact the equation depends only on the derivatives of the solutions so testing by a function or by a constant translation of the same funtion gives the same result. For sake of simplicity, we call the translated test function still φ. Since $\Phi^*_\delta(t)$ is strictly increasing for any $t \geq 0$ (see Proposition A.6), then it is invertible and its inverse function is still strictly increasing. Therefore by (2.75) we get

$$(\Phi^*_\delta)^{-1}\left(\Phi^*_\delta\left(\frac{d(x,y)}{t_0}\right)\right) \geq (\Phi^*_\delta)^{-1}\left(\frac{\varphi(t_0,x)}{t_0}\right),$$

or equivalently

$$d(x,y) - t_0\big(\Phi^*_\delta\big)^{-1}\left(\frac{\varphi(t_0,x)}{t_0}\right) \geq 0, \tag{2.76}$$

where we have extended $(\Phi^*)^{-1}$ by setting $(\Phi^*)^{-1}(t) = 0$ for any $t \leq 0$.
Set $k(x) := d(x,y) - t_0\big(\Phi^*_\delta\big)^{-1}\big(\frac{\varphi(t_0,x)}{t_0}\big)$, (2.76) implies that $k(x)$ has a local minimum at x_0. Let be

$$\psi(x) := t_0\big(\Phi^*_\delta\big)^{-1}\left(\frac{\varphi(t_0,x)}{t_0}\right),$$

we can use ψ as test function for the eikonal viscosity solution $d(x,y)$, at the point x_0 (in fact $\psi \in C^1$), so that

$$H_0(x_0, D\psi(x_0)) = 1. \tag{2.77}$$

By (2.75) we can remark that $\varphi(t_0,x_0) = t_0\Phi^*\left(\frac{d(x_0,y)}{t_0}\right) > 0$. Since $x_0 \neq y$, we can deduce that $\varphi(t_0,x) > 0$ near x_0, so that $\big(\Phi^*_\delta\big)^{-1}(s)$ is strictly positive and we are able to write

$$D\psi(x_0) = t_0 D\left[\big(\Phi^*_\delta\big)^{-1}\left(\frac{\varphi(t_0,x)}{t_0}\right)\right]\Bigg|_{x=x_0} = \left[\big(\Phi^*_\delta\big)'\left(\frac{d(x_0,y)}{t_0}\right)\right]^{-1} D\varphi(t_0,x_0). \tag{2.78}$$

Replacing (2.78) in (2.77), we get

$$H_0\left(x_0, \left[\big(\Phi^*_\delta\big)'\left(\frac{d(x_0,y)}{t_0}\right)\right]^{-1} D\varphi(t_0,x_0)\right) = 1.$$

Since $H_0(x,p)$ is positively homogeneous w.r.t. p, we get (2.74).
Now we remain to check identity (2.70) whenever $x_0 = y$. By (2.73) it is trivial that $\varphi_t(t_0,x_0) = 0$ so, using the homogeneous property of H_0, we have

to prove that $D\varphi(t_0, x_0) = 0$.
In such a particular case, (2.71) at t_0 becomes

$$g(x_0) + t_0 \Phi_\delta^* \left(\frac{d(x, x_0)}{t_0} \right) - \varphi(t_0, x) \geq g(x_0) - \varphi(t_0, x_0) = 0.$$

We can translate φ of a constant $g(x_0)$ so that $\widetilde{\varphi}(x_0) = g(x_0) - \varphi(t_0, x_0) = 0$. Moreover subtracting a suitable quadratic perturbation, we can assume that the test-function is strictly positive in $B_r(x_0) \backslash \{x_0\}$. In fact, if we look at

$$\psi(x) = \widetilde{\varphi}(t_0, x) - C|x - x_0|^2,$$

with $C > \max_{B_r(x_0)} \widetilde{\varphi}(t_0, x)$, then $\psi(x_0) = 0$ and $\psi(x) \leq 0$, so ψ attends maximum at x_0 and moreover $\psi \in C^1$. Hence $D\psi(x_0) = 0$ but $D\psi(x_0) = D\varphi(t_0, x_0)$, so we can conclude that $-H_0(x_0, D\varphi(t_0, x_0)) = 0 = \varphi_t(t_0, x_0)$.
Therefore v_δ^y is a LSC-viscosity solution of the equation (2.70).

Now remark that v^y is lower semicontinuous and is a pointwise-limit of LSC-viscosity solutions of (2.70), in fact $v_\delta^y(t, x) \to v^y(t, x)$, as $\delta \to 0^+$, for any $(t, x) \in (0, +\infty) \times \mathbb{R}^n$.
Set $H_\delta(x, p) = \Phi_\delta \circ H_0(x, p)$, it is immediate that $H_\delta \to H$, as $\delta \to 0^+$. Therefore, by Proposition 2.25, v^y is a lower semicontinuous viscosity solution of (2.9).

Recall that the metric Hopf-Lax function (2.48) is lower semicontinuous (see Lemma 2.62). Then, since it is infimum of LSC-viscosity solutions of (2.9), by Proposition 2.26 we can conclude that it is a LSC-viscosity solution of the Hamilton-Jacobi equation (2.9). □

Example 2.76. *Examples of positive convex functions, satisfying all the assumptions (H2) are $\Phi(t) = \frac{1}{\alpha} t^\alpha$ (with $\alpha > 1$) and $\Phi(t) = \mathrm{e}^t - t - 1$. While the functions $\Phi(t) = t$ and $\Phi(t) = \mathrm{e}^t - 1$ do not satisfy the assumptions since $\lim_{t \to 0^+} \Phi'(t) = 1$, in both the cases. For some details on these examples see Appendix A (Examples A.8, A.13, A.9 and A.11, respectively).*

Using the eikonal solution built in Sec. 2.2 (Theorem 2.40), Theorem 2.75 gives the following general result for the existence.

Theorem 2.77 ([44]). *Let be $H(x,p) = \Phi(|\sigma(x)p|)$, with $\sigma(x)$ $m \times n$ Hörmander-matrix with smooth coefficients and Φ satisfying assumptions $(H2)$. If $g \in LSC(\mathbb{R}^n)$ is such that (2.49) holds, then the Hopf-Lax function (2.48) is a viscosity solution of the Hamilton-Jacobi-Cauchy problem (2.66) and moreover the d-superlinear estimate (2.58) holds.*

2.4.3 Examples and applications.

We show that, whenever the initial datum g is continuous, the metric Hopf-Lax function is so.

Proposition 2.78. *If $g \in C(\mathbb{R}^n)$, then the Hopf-Lax function defined in (2.48) is a continuous function in $[0,+\infty) \times \mathbb{R}^n$.*

Proof. By Lemma (2.62) we know that the Hopf-Lax function $u(t,x)$ is lower semicontinuous in $[0,+\infty) \times \mathbb{R}^n$. So we only need to show that $u(t,x)$ is also upper semicontinuous, i.e. we want to prove that its upperlevel sets are closed.
Fixed $\gamma \in \mathbb{R}$, and let (t_x, x_k) be a sequence in the γ-upperlevel.
We must check that, if $(t_k, x_k) \to (t,x)$, as $k \to +\infty$, then $u(t,x) \geq \gamma$.
As in the proof of Lemma 2.62, we can assume $t > 0$. By definition (2.48), for any $y \in \mathbb{R}^n$

$$u(t_k, x_k) \leq g(y) + t_k \Phi^* \left(\frac{d(x_k, y)}{t_k} \right). \tag{2.79}$$

The second member of (2.79) is continuous, so passing to the upper-limit, we get

$$\limsup_{k \to +\infty} u(t_k, x_k) \leq g(y) + t\Phi^* \left(\frac{d(x,y)}{t} \right). \tag{2.80}$$

Taking the infimum of (2.80) over $y \in \mathbb{R}^n$, we can conclude

$$\gamma \leq \limsup_{k \to +\infty} u(t_k, x_k) \leq u(t,x).$$

$\square$

Remark 2.79. *By Theorem 2.21 it follows that, whenever the initial datum g is continuous, the metric Hopf-Lax function is a classic viscosity solution, following the usual definition of Crandall and Lions.*

Now we show that when the Hamiltonian depends only on the gradient, we find back the known (Euclidean) Hopf-Lax formula.
In fact, if we consider the Hamilton-Jacobi equation

$$u_t + H(|Du|) = 0,$$

with $H(p)$ continuous and convex, then we can set $\Phi = H$.
The associated eikonal equation is $|Du| = 1$. We have already observed that the (unique) viscosity solution of the previous eikonal equation (with condition $u(y) = 0$) is the Euclidean distance $d(x, y) = |x - y|$ (see Example 2.4).
Therefore in this case we can write formula (2.48) as

$$u(x, y) = \inf_{y \in \mathbb{R}^n} \left[g(y) + H^* \left(\frac{|x - y|}{t} \right) \right],$$

finding back the classic Hopf-Lax formula.

We apply now formula (2.48) to the model (2.67).
By Lemma 2.38, we know that the minimal-time function (2.22) is the Carnot-Carathéodory distance $d(x, y)$ associated to the matrix $\sigma(x)$.
In particular, whenever $\sigma(x)$ is a $n \times n$-matrix (invertible and symmetric), we get a Riemannian distance.

Now we want to pay particular attention to $m \times n$- matrix of Hörmander-type. In this case we can observe that $\sigma(x)Du$ is a different way to write the horizontal gradient Xu, where $\mathcal{X}u = (X_1 u, \dots, X_m u)^T$ and $\sigma(x)$ is the matrix with rows $X_i(x)$ for $i = 1, \dots, m$.
By Remark 2.64 we know that the metric Hopf-Lax function for the model (2.67) (with $\alpha > 1$) is given by (2.61).
To sum up, the Hopf-Lax function (2.61) solves (2.67) in the viscosity sense, whenever $\alpha > 1$, and the initial datum g is lower semicontinuous in $\mathbb{R}^n$ and satisfies (2.49).

Moreover by Propositions 2.65 and Proposition 2.68 we know that there exist u_t and Xu for a.e. $t > 0$ and a.e. $x \in \mathbb{R}^n$, whenever

$g \in BLSC(\mathbb{R}^n)$ and the Rademacher's Theorem holds (see Remark 2.42 for more details on this point). In fact,by using the local Lipschitz properties, we get that the Hopf-Lax viscosity solution of (2.66) satisfies the equation a.e. $(t,x) \in (0,+\infty)\times\mathbb{R}^n$ and so it is an almost everywhere solution, too.

Finally, we look at the model problem (2.67) with $\alpha > 1$ in a special sub-Riemannian case: the 1-dimensional Heisenberg group.
We recall that the matrix $\sigma(x)$ associated to the Heisenberg group can be written as

$$\sigma^t(x,y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{y}{2} & \frac{x}{2} \end{pmatrix}.$$

We show that, for the Heisenberg group, the metric Hopf-Lax function (2.61) coincides with the formula proved by Manfredi and Stroffolini in [74].
First note that

$$\sigma(x,y)Du = \Big(u_x - \frac{y}{2}u_z, u_y + \frac{x}{2}u_z, u_z\Big)^t = Xu,$$

where Xu is the usual horizontal gradient for the Heisenberg group $\mathbb{H}^1$.
So, in this case, the explicit expression for the Cauchy problem (2.66) is

$$\begin{cases} u_t + \dfrac{1}{\alpha}\left(\big(u_x - \frac{y}{2}u_z\big)^2 + \big(u_y - \frac{x}{2}u_z\big)^2\right)^{\frac{\alpha}{2}} = 0, \text{ in } \mathbb{R}^n \times (0,+\infty), \\ u = g, \quad \text{in } \mathbb{R}^n \times \{0\}. \end{cases}$$

Since $\mathbb{H}^1$ is a Carnot group, there is a group of dilations which are usually indicated by δ_λ, for $\lambda > 0$ (see Sec. 1.2).
We want only to recall that, in the particular case of the Heisenberg group $\mathbb{H}^1$, the dilations are

$$\delta_\lambda(x,y,z) = (\lambda x, \lambda y, \lambda^2 z)$$

and the operation is defined by

$$(x,y,z)\cdot(x',y',z') = \left(x+x', y+y', z+z' + \frac{1}{2}(xy' - x'y)\right)$$

and so the inverse element is $(x,y,z)^{-1} = (-x,-y,-z)$.

Then we get

$$\frac{1}{t}d_{\mathbb{H}^1}\big((x,y,z),(x_1,y_1,z_1)\big) =$$
$$= d_{\mathbb{H}^1}\big(\delta_{\frac{1}{t}}(x,y,z),\delta_{\frac{1}{t}}(x_1,y_1,z_1)\big) = d_{\mathbb{H}^1}\left(\left(\frac{x}{t},\frac{y}{t},\frac{z}{t^2}\right),\left(\frac{x_1}{t},\frac{y_1}{t},\frac{z_1}{t^2}\right)\right). \quad (2.81)$$

Moreover, we know that

$$d_{\mathbb{H}^1}\big((x,y,z),(x_1,y_1,z_1)\big) = |(x,y,z)^{-1}(x_1,y_1,z_1)|_{\mathbb{H}^1},$$

where the Heisenberg norm is defined as $|(x,y,z)|_{\mathbb{H}^1} := d_{\mathbb{H}^1}((0,0,0),(x,y,z))$. By this remark, (2.81) gives us

$$\frac{1}{t}d_{\mathbb{H}^1}\big((x,y,z),(x_1,y_1,z_1)\big) = \left|\left(\frac{x_1-x}{t},\frac{y_1-y}{t},\frac{z_1-z+\frac{1}{2}(x_1y-xy_1)}{t^2}\right)\right|_{\mathbb{H}^1}.$$

To sum up, in the case of the 1-dimensional Heisenberg group, we can write the Carnot-Carathéodory Hopf-Lax function (2.61), as

$$u(x_1,y_1,z_1,t) =$$
$$= \inf_{\mathbb{R}^3}\left[g(x,y,z)+\frac{t}{\beta}\left(\left|\left(\frac{x_1-x}{t},\frac{y_1-y}{t},\frac{z_1-z+\frac{1}{2}(x_1y-xy_1)}{t^2}\right)\right|_{\mathbb{H}^1}\right)^{\beta}\right]$$

getting the same formula found by Manfredi and Stroffolini.

Chapter 3

Carnot-Carathéodory inf-convolutions.

3.1 Definition and basic properties.

The approximation by inf-convolutions is known, in semigroup theory, as Yosida's regularization and, in optimization theory, as Moreau's regularization. It very useful in order to prove many results in nonlinear analysis. For example, we want to recall that the inf-convolutions are used by Barron and Jensen in [14], in order to get the uniqueness of (upper semicontinuous) v. solution of terminal-time Cauchy problems for evolutive Hamilton-Jacobi equations, under a (global) Lipschitz assumption on the Hamiltonian, w.r.t. the gradient.

We begin by defining the inf-convolutions in the classical setting and quoting some related properties.

Definition 3.1. *For any $g : \mathbb{R}^n \to \mathbb{R}$, the inf-convolutions is the Hopf-Lax function* (2.40) *with $H(p) = \frac{1}{2}|p|^2 = H^*(p)$, that is*

$$g_t(x) = \inf_{y \in \mathbb{R}^n} \left[g(y) + \frac{|x-y|^2}{2t} \right]. \tag{3.1}$$

Sometime, speaking about function (3.1), we say Euclidean inf-convolution, since $|x - y|$ means the Euclidean distance between x and y.

For function (3.1) many properties holds. We recall only the main ones. We assume that the function g is bounded and lower semicontinuous (see [7], Lemma V.5.6).

Lemma 3.2. *Let be $g \in BLSC(\mathbb{R}^n)$ and $t > 0$, then*

(i) *g_t is locally Lipschitz continuous in $\mathbb{R}^n$.*

(ii) *For any $x \in \mathbb{R}^n$, $g_t(x) \to g(x)$ as $t \to 0^+$.*

(iii) *For any $x \in \mathbb{R}^n$ and $t > 0$, the infimum in (3.1) is a minimum and it is attended in a closed ball centered at x and with radius $R(t) = t(2\|g\|_\infty)^{\frac{1}{2}}$.*

(iv) *Let $M_t(x)$ be the set of all the minimum-points of $g_t(x)$, then the subdifferential $\partial g_t(x)$ is empty whenever $M_t(x)$ is not a singleton while, whenever $M_t(x) = \{y_t\}$, then $\partial g_t(x) = \left\{2\frac{x-y_t}{t^2}\right\}$.*

Remark 3.3. *The results proved in [7] are more general than the previous lemma. In fact, the authors study an approximation by inf-convolutions w.r.t. both the space-variable and the time-variable. More precisely, for any $\varepsilon, C > 0$, the approximating function is defined as*

$$g_\varepsilon(x,t) = \inf_{y \in \mathbb{R}^n} \left[g(y) + \mathrm{e}^{-Ct} \frac{|x-y|^2}{\varepsilon^2} \right].$$

Analogously, it is possible to study a more general definition of inf-convolution, adding a quadratic perturbation in the time-variable, too, i.e.

$$g_\varepsilon(x,t) = \inf \left\{ g(y) + \mathrm{e}^{-Ct} \frac{|x-y|^2}{\varepsilon^2} + \frac{|t-s|^2}{\varepsilon^2} \,\middle|\, y \in \mathbb{R}^n, s \in (T_1, T_2) \right\},$$

given $0 < T_1 < T_2$.
Both the previous marginal-functions are Lipschitz approximations of the function g. Nevertheless, we are interested in studying only the simple version defined by formula (3.1).

Remark 3.4 (sup-convolutions)**.** *For any function $g:\mathbb{R}^n \to \mathbb{R}^n$, it is possible to define the* sup-convolution*, that is*

$$g^t(x) = \sup_{y\in\mathbb{R}^n}\left[g(y) - \frac{|x-y|^2}{2t}\right]. \tag{3.2}$$

For the (Euclidean) sup-convolution, one can show the analogous properties of the ones given in Lemma (3.2) *(see for example [24]).*

A first generalization of the Euclidean inf-convolution and sup-convolution has been given by J.M. Lasry and P.L. Lions in [66]. There, the authors are interested in finding a method to approximate bounded and uniformly continuous functions, defined in an infinite-dimensional Hilbert space $\mathbb{H}$. In particular, setting $|\cdot|_{\mathbb{H}}$ the norm in the Hilbert space $\mathbb{H}$, it is possible to prove that for any $t>0$

$$\underline{g}_t(x) = \sup_{z\in\mathbb{H}}\inf_{y\in\mathbb{H}}\left[g(y) + \frac{1}{2t}|z-y|^2_{\mathbb{H}} - \frac{1}{t}|z-x|^2_{\mathbb{H}}\right]$$

and

$$\overline{g}_t(x) = \inf_{z\in\mathbb{H}}\sup_{y\in\mathbb{H}}\left[g(y) - \frac{1}{2t}|z-y|^2_{\mathbb{H}} + \frac{1}{t}|z-x|^2_{\mathbb{H}}\right]$$

are elements of $BC^{1,1}(\mathbb{H})$ satisfying

$$\sup_{\mathbb{H}}|\nabla g_t|_{\mathbb{H}} \le C_t \sup_{\mathbb{H}}|g|_{\mathbb{H}},$$

$$\sup_{\mathbb{H}}|\nabla g_t(x) - \nabla g_t(y)|_{\mathbb{H}}|x-y|^{-1}_{\mathbb{H}} \le C_t \sup_{\mathbb{H}}|g|_{\mathbb{H}},$$

$$\inf_{\mathbb{H}} g \le g_t \le \sup_{\mathbb{H}} g,$$

$$\sup_{\mathbb{H}}|\nabla g_t|_{\mathbb{H}} \le \sup_{z\neq y}|g(x)-g(y)|_{\mathbb{H}}|x-y|^{-1}_{\mathbb{H}} \le +\infty,$$

where g_t is equal to $\underline{g}_t$ and $\overline{g}_t$, respectively, and $C_t>0$.
As $t\to 0^+$, both these approximations converge uniformly to g on $\mathbb{H}$ and moreover $\underline{g}_t \le g \le \overline{g}_t$. Finally, in [66], it is remarked that, if one is interested in getting Lipschitz continuous regularizations (and not $C^{1,1}$ regularizations), then it is sufficient to look at formulas (3.1) and (3.2) with $|\cdot| = |\cdot|_{\mathbb{H}}$.

We are interested in generalizing the inf-convolutions to bounded and lower semicontinuous functions in finite-dimensional metric spaces. In particular, we want to look at $\mathbb{R}^n$ endowed with a Carnot-Carathéodory distance (satisfying the Hörmander condition).
As we have already remarked, the Carnot-Carathéodory distances satisfying the Hörmander condition are finite distances on $\mathbb{R}^n$.

To define a suitable metric approximations in this setting we need only to apply the metric Hopf-Lax formula (2.48) with $\Phi(s) = \frac{1}{2}s^2 = \Phi^*(s)$.

Definition 3.5. *Let $d(x, y)$ be a Carnot-Carathéodory distance satisfying the Hörmander condition, for any $g : \mathbb{R}^n \to \mathbb{R}$ and $t > 0$, we call* Carnot-Carathéodory inf-convolution *(or simply* Carnot-Carathéodory inf-convolution*) the function $g_t : \mathbb{R}^n \to \mathbb{R}$, defined as*

$$g_t(x) = \inf_{y \in \mathbb{R}^n} \left[g(y) + \frac{d(x, y)^2}{2t} \right]. \tag{3.3}$$

Remark 3.6. *Note that function* (3.3) *can be defined also by starting form a generalized distance $d(x, y)$ (and many of the properties still wiill be still true). In the general metric case,* (3.3) *is called* metric inf-convolution.

By applying the properties proved in Sec. 2.4.1 for the metric Hopf-Lax formula, we can sum-up the following theorem.

Theorem 3.7. *Let be $g \in LSC(\mathbb{R}^n)$, then the Carnot-Carathéodory inf-convolution g_t satisfies the following properties:*

(i) *$g_t \leq g$ for any $t > 0$.*

(ii) *g_t is locally d-Lipschitz w.r.t. x for $t > 0$. Hence g_t is locally (Euclidean) Hölder continuous with exponent $1/k$ where k is the step of the Carnot-Carathéodory distance $d(x, y)$.*

(iii) *g_t lower-converges to g, as $t \to 0^+$.*

(iv) *If $g(x) \leq -C(1 + d(0, x))$ for some constant $C > 0$, then*

$$g_t(x) \leq -C'(1 + t + d(0, x)),$$

for any $x \in \mathbb{R}^n$ and $t > 0$, where $C' = \max\{C, \frac{1}{2}C^2\}$.

(v) *If $g \in BLSC(\mathbb{R}^n)$, then the infimum in* (3.3) *is attended at a point which belongs to the closed Carnot-Carathéodory ball centered at x and with radius $R(t) = t^{\frac{1}{2}} \|g\|_{\infty}^{\frac{1}{2}}$. Moreover g_t is locally Lipschitz continuous w.r.t. $t > 0$.*

Proof. Property (i) has been proved in Remark 2.59.
Moreover by Proposition 2.65 g_t is d-Lipschitz continuous w.r.t. x. Then, by using the estimate (1.29) for Carnot-Carathéodory distances satisfying the Hörmander condition (see Chapter 1, Theorem 1.56), we can conclude property (ii).
Property (iii) is given by Lemma 2.60 and property (iv) by Lemma 2.61. Property (v) follows from Lemma 2.63 and Proposition 2.68. □

To sum up, we want to point out that Carnot-Carathéodory inf-convolutions given by (3.3) are Lipschitz functions (in the suitable metric) and monotone approximations of the original bounded and lower semicontinuous function.
Moreover, whenever the Rademacher's Theorem holds (see 2.42), the Carnot-Carathéodory inf-convolutions are almost everywhere horizontal differentiable. More precisely, set $u(t,x) = g_t(x)$, u_t and $Xu = \sigma(x)D_x u$ exist a.e. $t > 0$ and $x \in \mathbb{R}^n$.

3.2 Inf-convolutions and logarithms of heat kernels.

In this section we generalize to the Carnot-Carathéodory inf-convolutions a convergence-result known in the corresponding Euclidean case. We follow an idea recently introduced by I. Capuzzo Dolcetta [26].
More precisely the inf-convolution can be approximated by the logarithms of suitable integral convolutions of the datum, i.e.

$$\lim_{\varepsilon\to 0^+} -2\varepsilon \log\left((4\pi\varepsilon t)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4\varepsilon t}} e^{-\frac{g(y)}{2\varepsilon}} dy\right) = \inf_{y\in\mathbb{R}^n}\left[g(y) + \frac{|x-y|^2}{2t}\right].$$

In other words, this means to investigate the limiting behavior of solutions w^ε of subelliptic heat equations (with conductivity $\varepsilon > 0$), and prove that $-2\varepsilon \log w^\varepsilon$ converges, as $\varepsilon \to 0^+$, to the Carnot-Carathéodory inf-convolutions of the initial datum.

3.2.1 The Euclidean approximation and the Large Deviation Principle.

Here we recall in details the result of convergence proved by Capuzzo Dolcetta in [26]. We start looking at the Cauchy problem for the following Hamilton-Jacobi equation

$$\begin{cases} u_t + \frac{1}{2}|Du|^2 = 0, & x \in \mathbb{R}^n, t > 0, \\ u(x,0) = g(x), & x \in \mathbb{R}^n. \end{cases} \tag{3.4}$$

The idea is to regularize the problem (3.4) in a suitable way in order to get a family of Cauchy problems for heat equations: in fact we know that heat equations are solved by integral convolutions of the fundamental solutions and the initial datum.
The first step consists in adding a (so called viscosity) term $-\varepsilon\Delta\, u$ with $\varepsilon > 0$.
We so find a family of Cauchy problems for second-order PDEs with initial data g:

$$\begin{cases} u_t^\varepsilon - \varepsilon\Delta u^\varepsilon + \frac{1}{2}|Du^\varepsilon|^2 = 0, & x \in \mathbb{R}^n, t > 0, \\ u^\varepsilon(x,0) = g(x), & x \in \mathbb{R}^n. \end{cases} \tag{3.5}$$

Using the Hopf-Cole transform $w^\varepsilon := e^{-\frac{u^\varepsilon}{2\varepsilon}}$, it is possible to linearize the previous nonlinear equations. In fact, computing the corresponding derivatives, we find that, whenever u^ε solves problem (3.5), its Hopf-Cole transform w^ε solves the following Cauchy problems for the heat equation (with conductivity $\varepsilon > 0$):

$$\begin{cases} w_t^\varepsilon - \varepsilon\Delta w^\varepsilon = 0, & x \in \mathbb{R}^n,\ t > 0, \\ w^\varepsilon(x,0) = g^\varepsilon(x), & x \in \mathbb{R}^n. \end{cases} \tag{3.6}$$

It is well-known that the solutions of problems (3.6) are given by convolutions of the fundamental solution and the initial datum $g^\varepsilon(x)$ (see for example [47], Section 2.3), i.e.

$$w^{\varepsilon}(t,x) = (4\pi\varepsilon t)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4\varepsilon t}} e^{-\frac{g(y)}{2\varepsilon}} dy.$$

So by applying the anti-transform of Hopf-Cole and using the explicit formula for the solutions of the heat problems (3.6), we can deduce the following logarithmic representative formula for the solutions of the Cauchy problems (3.5):

$$u^{\varepsilon}(t,x) = -2\varepsilon \log\left((4\pi\varepsilon t)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4\varepsilon t}} e^{-\frac{g(y)}{2\varepsilon}} dy \right). \tag{3.7}$$

It is natural to expect that as $\varepsilon \to 0^+$ the solutions $u^{\varepsilon}(t,x)$ given in (3.7) converge to the unique solution of the starting Cauchy problem (3.4).
By the Hopf-Lax formula (see [47] or [7]) we know that the (unique) solution of the Cauchy problem (3.4) is the inf-convolution of the initial datum. In [26], Capuzzo Dolcetta shows directly that

$$\lim_{\varepsilon \to 0^+} u^{\varepsilon}(t,x) = g_t(x), \tag{3.8}$$

where $g_t(x)$ is defined by (3.1) and $g \in BC(\mathbb{R}^n)$.
This means that the (Euclidean) inf-convolution can be seen as limit of logarithms of suitable integral convolutions.

Limit (3.8) follows trivially from a Laplace-Varadhan's type result, which is an application of the Large Deviation Principle ([94]).
Let us give some details.

Definition 3.8 (Large Deviation Principle)**.** *Let P^{ε} be a family of probability measures defined on the Borel sets of some complete and separable metric space X. We say that the family of probability measures P^{ε} satisfies the* Large Deviation Principle *if there exists a function (called* rate function*) $I : X \to [0, +\infty]$ such that*

(i) $I \in LSC(X)$,

(ii) *for any $k < +\infty$, the sublevel set $\{x \in X \mid I(x) \le k\}$ is compact,*

(iii) *for any $A \subset X$ open set, it holds*

$$\liminf_{\varepsilon \to 0^+} \varepsilon \log P^{\varepsilon}(A) \ge - \inf_{x \in A} I(x),$$

(iv) *for any $C \subset X$ closed set, it holds*

$$\limsup_{\varepsilon\to 0^+} \varepsilon \log P^\varepsilon(C) \le -\inf_{x\in C} I(x).$$

Theorem 3.9. *Let X be a complete and separable metric space and P^ε be a family of probability measures satisfying the Large Deviation Principle, then for any $F \in BC(X)$*

$$\lim_{\varepsilon\to 0^+} \varepsilon \log\left(\int_X \exp\left[\frac{F(x)}{\varepsilon}\right] dP^\varepsilon\right) = \sup_{x\in X}[F(x) - I(x)]. \tag{3.9}$$

Proof. Let be $F \in BC(\mathbb{R}^n)$ and $I(x) \ge 0$, then

$$\sup_{x\in X}[F(x) - I(x)] \le \sum_{x\in X} F(x) < +\infty.$$

So by the definition of supremum, for any $\delta > 0$ there exists a point $y \in X$ such that

$$F(y) - I(y) \ge \sup_{x\in X}[F(x) - I(x)] - \frac{\delta}{2}.$$

$F(x)$ is a continuous function, then we can find a neighborhood U of y such that $F(x) \ge F(y) - \frac{\delta}{2}$, which implies

$$\begin{aligned}\liminf_{\varepsilon\to 0^+} \varepsilon \log\left(\int_X \exp\left[\frac{F(x)}{\varepsilon}\right] dP^\varepsilon\right) &\ge \liminf_{\varepsilon\to 0^+} \varepsilon \log\left(\int_U \exp\left[\frac{F(x)}{\varepsilon}\right] dP^\varepsilon\right)\\ &\ge \liminf_{\varepsilon\to 0^+} \varepsilon \log\left(\exp\left[\frac{F(y)}{\varepsilon} - \frac{\delta}{2\varepsilon}\right] P^\varepsilon(U)\right).\end{aligned}$$

By applying property (iv) of the Large Deviation Principle (Definition 3.8), we find

$$\begin{aligned}\liminf_{\varepsilon\to 0^+} \varepsilon \log\left(\int_X \exp\left[\frac{F(x)}{\varepsilon}\right] dP^\varepsilon\right) &\ge -\inf_{x\in U} I(x) + F(y) - \frac{\delta}{2}\\ &\ge F(y) - I(y) - \frac{\delta}{2} \ge \sup_{x\in X}[F(x) - I(x)] - \delta.\end{aligned}$$

Passing to the limit as $\delta \to 0^+$, we get

$$\liminf_{\varepsilon\to 0^+} \varepsilon \log\left(\int_X \exp\left[\frac{F(x)}{\varepsilon}\right] dP^\varepsilon\right) \ge \sup_{x\in X}[F(x) - I(x)].$$

We remain to show the reverse inequality for the upper limit.
Since X is a separable space and $F \in BC(X)$, then for any $\delta > 0$ there exists a finite number of closed sets $K_1, ..., K_{N_\delta}$ covering X and such that

$$\max_{x,y \in K_i} |F(x) - F(y)| < \delta.$$

Hence

$$\int_X \exp\left[\frac{F(x)}{\varepsilon}\right] dP^\varepsilon \leq \sum_{i=1}^{N_\delta} \int_{K_i} \exp\left[\frac{F(x)}{\varepsilon}\right] dP^\varepsilon \leq \sum_{i=1}^{N_\delta} \int_{K_i} \exp\left[\frac{F_i + \delta}{\varepsilon}\right] dP^\varepsilon,$$

where $F_i = \inf_{x \in K_i} F(x) > -\infty$. Therefore

$$\begin{aligned} &\limsup_{\varepsilon \to 0^+} \varepsilon \log\left(\int_X \exp\left[\frac{F(x)}{\varepsilon}\right] dP^\varepsilon\right) \\ &\qquad \leq \sup_{i=1,\dots,N_\delta} \limsup_{\varepsilon \to 0^+} \varepsilon \log\left(\int_X \exp\left[\frac{F_i + \delta}{\varepsilon}\right] dP^\varepsilon\right) \\ &\qquad\qquad = \sup_{i=1,\dots,N_\delta} \limsup_{\varepsilon \to 0^+} [F_i + \delta + \varepsilon \log P^\varepsilon(K_i)]. \end{aligned}$$

Applying property (iii) of the Large Deviation Principle

$$\begin{aligned} \sup_{i=1,\dots,N_\delta} \limsup_{\varepsilon \to 0^+} [F_i + \delta + \varepsilon \log P^\varepsilon(K_i)] &\leq \sup_{i=1,\dots,N_\delta} [-\inf_{x \in K_i} I(x) + F_i + \delta] \\ &\leq \sup_{i=1,\dots,N_\delta} \sup_{x \in K_i} [F(x) - I(x)] + \delta \leq \sup_{x \in X} [F(x) - I(x)] + \delta \end{aligned}$$

passing to the limit as $\delta \to 0^+$, we get

$$\limsup_{\varepsilon \to 0^+} \varepsilon \log\left(\int_X \exp\left[\frac{F(x)}{\varepsilon}\right] dP^\varepsilon\right) \leq \sup_{x \in X} [F(x) - I(x)].$$

By the two found inequalities we can deduce

$$\begin{aligned} &\sup_{x \in X} [F(x) - I(x)] \leq \liminf_{\varepsilon \to 0^+} \varepsilon \log\left(\int_X \exp\left[\frac{F(x)}{\varepsilon}\right] dP^\varepsilon\right) \\ &\leq \limsup_{\varepsilon \to 0^+} \varepsilon \log\left(\int_X \exp\left[\frac{F(x)}{\varepsilon}\right] dP^\varepsilon\right) \leq \sup_{x \in X} [F(x) - I(x)], \end{aligned}$$

which concludes the proof. $\square$

If we can apply the limit (3.9) to the function $F = -g/2$, starting form the gaussian measures

$$dP^{\varepsilon}_{t,x} = (4\pi\varepsilon t)^{-\frac{n}{2}} \mathrm{e}^{-\frac{|x-\cdot|^2}{4\varepsilon t}} d\mathcal{L}^n,$$

with rate function

$$I_{t,x}(y) = \frac{|x-y|^2}{4t},$$

then we get exactly the convergence result (3.8).
So we need only to show the applicability of the Large Deviation Principle. Properties (i) and (ii) are trivially satisfied, then we must only verify the estimates for the upper and lower limits in the closed and the open sets, respectively. These are not trivial to prove. In the next subsection we recall briefly a paper of Varadhan [93] published in 1967, where the author proved the applicability of the Large Deviation Principle but only up to the Riemannian case.

3.2.2 Applicability of the Large Deviation Principle: the proof of Varadhan for the Riemannian case.

To show properties (iii) and (iv) of Definition 3.8 is difficult even in the simple Euclidean case. The difficulties arrise whenever the sets are unbounded.
Before recalling the ideas in the proof of Varadhan, we like to give an easy proof for bounded sets, in the Euclidean case.

Lemma 3.10. *Let $B \subset \mathbb{R}^n$ be a Borel and bounded set, then*

$$\lim_{\varepsilon\to 0^+} \varepsilon \log \left[(4\pi\varepsilon t)^{-\frac{n}{2}} \int_B \mathrm{e}^{-\frac{|x-y|^2}{4\varepsilon t}} dy \right] = -\inf_{y\in B} \frac{|x-y|^2}{4t}.$$

Proof. In order to prove the limit, we use the convergence in the bounded sets of L^q-norms to the L^∞-norm, as $q \to +\infty$.
Therefore, set $q = 1/\varepsilon$ and note that

$$\left(\int_B \mathrm{e}^{-\frac{|x-y|^2}{4\varepsilon t}} dy \right)^{\varepsilon} = \left\| \mathrm{e}^{-\frac{|x-\cdot|^2}{4t}} \right\|_{q,\,B}.$$

For any B Borel and bounded set, it is well-known (see, for example, [70]) that

$$\lim_{q\to+\infty} \left\| \mathrm{e}^{-\frac{|x-\cdot|^2}{4t}} \right\|_{q,\,B} = \left\| \mathrm{e}^{-\frac{|x-\cdot|^2}{4t}} \right\|_{+\infty,\,B} = \sup_{y\in B} \mathrm{e}^{-\frac{|x-y|^2}{4t}}.$$

By the continuity of the logarithm function, it follows that

$$\lim_{\varepsilon\to 0^+} \log\left(\int_B \mathrm{e}^{-\frac{|x-y|^2}{4\varepsilon t}}dy\right)^{\varepsilon} = \lim_{q\to+\infty} \log\left\|\mathrm{e}^{-\frac{|x-\cdot|^2}{4t}}\right\|_{q,\,B} = -\inf_{y\in B}\frac{|x-y|^2}{4t}.$$

Moreover, it is easy to show that $\lim_{\varepsilon\to 0^+}\varepsilon\log(4\pi\varepsilon t)^{-\frac{n}{2}} = 0$ and this remark concludes the proof. □

Next we quote the result proved by Varadhan in [93], recalling briefly the main ideas. Later investigating the Carnot-Carathéodory case, we will show how it could be possible to generalize these ideas in the more general case. Nevertheless we will give a very different but simpler proof, following the remarks from the proof of Lemma 3.10.

Let us define

$$L^{\varepsilon}f = \varepsilon\sum_{i,j=1}^{n} a_{i,j}(x)\frac{\partial^2 f}{\partial x_i\partial x_j},$$

and we indicate by $p^{\varepsilon}(t,x,y)$ the heat kernel associated to the parabolic equation

$$\frac{\partial p^{\varepsilon}}{\partial t} = L^{\varepsilon}p^{\varepsilon}. \tag{3.10}$$

Moreover, we indicate by L and $p(t,x,y)$ the operator L^{ε} and the heat kernel $p^{\varepsilon}(t,x,y)$ when $\varepsilon = 1$. So in particular $L^{\varepsilon} = \varepsilon L$.
We assume that L is an *uniformly elliptic* operator, which means that for any $x \in \mathbb{R}^n$ $A(x) := (a_{i,j}(x))_{i,j=1,\dots,n}$ is a symmetric and positive definite $n\times n$-matrix, satisfying the following uniformly elliptic condition, i.e.

$$C_1\sum_{i=1}^{n}\xi_i^2 \le \sum_{i,j=1}^{n} a_{i,j}(x)\,\xi_i\cdot\xi_j \le C_2\sum_{i=1}^{n}\xi_i^2, \tag{3.11}$$

for suitable positive and finite constants C_1 and C_2.
We also assume that the coefficients satisfy an uniform Hölder-condition:

$$|a_{i,j}(x) - a_{i,j}(y)| \le M|x-y|^h,$$

for some $M > 0$ and $0 < h \le 1$.

Next, we need to define the family of probability measures and the

distance associated to the equation (3.10).
So, for any $B \subset \mathbb{R}^n$ Borel set, we define

$$P^{\varepsilon}_{t,x}(B) = \int_B p^{\varepsilon}(t,x,y)dy. \tag{3.12}$$

Then, let $\Gamma(x,y)$ be the set of all the curves joining x to y in some finite time T, for any $\gamma \in \Gamma(x,y)$, the length can be defined as

$$l(\gamma) = \int_0^T \sqrt{\left[\dot{\gamma}(t)\right]^T A^{-1}(\gamma(t))\, \dot{\gamma}(t)}\, dt. \tag{3.13}$$

Remark 3.11. *Note that for any x the inverse matrix $A^{-1}(x)$ exists since we have assumed that $A(x)$ is a positive definite matrix.*

The distance associated to the operator $L = \sum_{i,j=1}^m a_{i,j}(x)\frac{\partial^2}{\partial x_i \partial x_j}$ can be written as

$$d(x,y) = \inf\{l(\gamma)\,|\, \gamma \in \Gamma(x,y)\}. \tag{3.14}$$

Whenever L is an uniformly elliptic operator, then the associated distance is a Riemannian distance.
We now quote the result of Varadhan.

Theorem 3.12. *Let be $x \in \mathbb{R}^n$ and $t > 0$. Let $P^{\varepsilon}_{t,x}$ be the probability measures defined by* (3.12) *and $d(x,y)$ be distance defined in* (3.14)*, then*

(i) *for any open set $A \subset \mathbb{R}^n$*

$$\liminf_{\varepsilon \to 0^+} \varepsilon \log P^{\varepsilon}_{t,x}(A) \geq -\inf_{y \in A} \frac{d(x,y)^2}{4t},$$

(ii) *for any closed set $C \subset \mathbb{R}^n$*

$$\limsup_{\varepsilon \to 0^+} \varepsilon \log P^{\varepsilon}_{t,x}(C) \leq -\inf_{y \in C} \frac{d(x,y)^2}{4t}.$$

Sketch of the proof [92].
The idea of the proof of Vardhan is to use the asymptotic behavior of $-4tp(t,x,y)$, as $t \to 0^+$.
We need first to go from the behavior of the heat kernel $p^{\varepsilon}(t,x,y)$ for small conductivity ε to the behavior of the heat kernel $p(t,x,y)$ for small times. Then the next result holds.

Lemma 3.13. *For any $\varepsilon, t > 0$ and $x, y \in \mathbb{R}^n$, we have*

$$p^\varepsilon(t, x, y) = p(\varepsilon t, x, y).$$

Proof. The proof follows form the uniqueness of solutions for the following Cauchy problems. In fact, if the coefficients $a_{i,j}(x)$ do not depend on the time-variable, for any $x \in \mathbb{R}^n$ fixed the function $w^\varepsilon(t, x, y) := p(\varepsilon t, x, y)$ satisfies

$$\begin{cases} \dfrac{\partial w^\varepsilon}{\partial t} - \varepsilon \displaystyle\sum_{i,j+1}^{n} a_{i,j}(x) \frac{\partial^2 f}{\partial y_i \partial y_j} w^\varepsilon = 0, \quad y \in \mathbb{R}^n, \ t > 0, \\ w^\varepsilon(0, x, y) = \delta_x(y), \quad y \in \mathbb{R}^n. \end{cases} \tag{3.15}$$

So, by the uniqueness for Cauchy problems (3.15), we get the identity $p^\varepsilon(t, x, y) = p(\varepsilon t, x, y)$. ☐

The next step consists in using the theory of Markov processes.
In fact, it is possible to show that the following semigroup property (also known as Chapman-Kolmogorov equation) holds:

$$p^\varepsilon(t + s, x, y) = \int_{\mathbb{R}^n} p^\varepsilon(t, x, z) p^\varepsilon(s, z, y) dz,$$

whenever p^ε is the heat kernel associated to an uniformly elliptic operator.
So, fixed a strictly increasing N-partition $t_1, ..., t_N$ of $[0, T]$, one can define the probability of a cylinder $A = \{\gamma \in \Gamma(x, y) \,|\, \gamma(t_1), ..., \gamma(t_N) \in B\}$ as

$$P^\varepsilon_{t,x}(A) = \int_B p^\varepsilon(t_1, x, y_1) p^\varepsilon(t_2 - t_1, y_1, y_2) ... p^\varepsilon(t_N - t_{N-1}, y_{N-1}, y_N) dy_1 ... dy_N.$$

The Markov's property let us show a link between the distance (3.14) and the energy-functional

$$I(\gamma) = \frac{1}{2} \int_0^T \left[\dot\gamma(t)\right]^T A^{-1}(\gamma(t)) \, \dot\gamma(t) \, dt. \tag{3.16}$$

The functional (3.16) depends on the parametrization chosen for the curve γ but it is minimum whenever the velocity is constant.
Instead the length-functional (3.13) does not depend on the chosen parametrization: so to minimize this, we can look only at parameterizations with constant velocity.

Lemma 3.14. *Let $d(x,y)$ be defined by* (3.14) *and I be the energy-functional* (3.16)*, then for any $0 \leq \alpha < \beta \leq T$*

$$\inf\{I(\gamma)\,|\,\gamma \in \Gamma(x,y),\ with\,\gamma(\alpha) = x,\ \gamma(\beta) = y\} = \frac{d(x,y)^2}{2(\beta-\alpha)}.$$

Proof. Let $\gamma(t)$ be a curve joining x to y in a time $\beta - \alpha > 0$ and set $\dot\gamma(t) = (\alpha_1(t), ..., \alpha_n(t))$. We have already remarked that it is possible to look only at $\gamma(t)$ such that $\sum_{i=1}^n \alpha_i(t)^2 = C$, for some $C > 0$.
This makes easy to calculate the corresponding functionals $l(\gamma)$ and $I(\gamma)$, i.e.

$$l(\gamma) = (\beta-\alpha)C, \quad I(\gamma) = \frac{1}{2}(\beta-\alpha)C^2.$$

Then $I(\gamma) = \frac{l(\gamma)^2}{2(\beta-\alpha)}$: by taking the infimum, we conclude the proof. □

By induction, it is easy to generalize the previous lemma to a generic N-partition of $[0,T]$.

Lemma 3.15. *Let $d(x,y)$ be the distance defined by* (3.14) *and I be the energy-functional* (3.16) *then, for any N-partition of $[0,T]$ $0 \leq t_1 < t_2 < ... < t_N \leq T$,*
it holds

$$\inf\{I(\gamma)\,|\,\gamma\ continuous\ curve\ :\ \gamma(t_i) = x_i,\ i = 1,..,N\} = \frac{1}{2}\sum_{i=1}^{N}\frac{d(x_{i+1},x_i)^2}{t_{j+1}-t_j}.$$

To conclude the sketch of this proof, we recall the following pointwise-limit of $-2tp(t,x,y)$ as $t \to 0^+$.

Theorem 3.16 ([92])**.** *Let $p(t,x,y)$ be the heat kernel associated to the parabolic equation* (3.10) *with $\varepsilon = 1$ and $d(x,y)$ be the Riemannian distance defined by* (3.14)*, then for any x and y punctually,*

$$\lim_{\tau\to 0^+} 4\tau \log p(\tau,x,y) = -d(x,y)^2. \tag{3.17}$$

Moreover the previous convergence is uniform whenever $d(x,y)$ is bounded.

Lemma 3.15 and Theorem 3.16 are used in order to prove the estimate for the upper limit in closed and bounded sets.

Then, Varadhan investigates the limiting behavior outside large balls (see [93], Lemma 3.2) and shows that this limit is small.
This leads to the estimate for the upper limit in closed set (Theorem 3.3) and then we deduce the corresponding estimate for the lower limit in open sets (Lemma 3.4. and Theorem 3.5).
In this chapter we study the Carnot-Carathéodory distance and we generalize the result of Varadhan to hypoellitic operators.

3.3 Carnot-Carathéodory inf-convolutions and ultraparabolic equations.

3.3.1 Heat kernels for hypoelliptic operators.

Before giving a representative formula for the solutions of (3.28), we recall what a hypoelliptic operator is.
It is well-known that the solutions (in distributional sense) of an uniformly elliptic equation $Lu = f$ are smooth whenever f is so. Nevertheless, the previous property holds for more general operators.

Definition 3.17. *A differential operator L is* hypoelliptic *(or also* subelliptic*) if and only if the solutions of $Lu = f$ are smooth, whenever $f \in C^\infty$.*

The main example of hypoelliptic operators are the sums of squares of Hörmander vector fields, that are

$$L = \sum_{i=1}^{m} X_i^2, \tag{3.18}$$

where $X_1, ..., X_m$ are smooth vector fields, satisfying the Hörmander condition (see [55], Theorem 1.1).

The theory of hypoelliptic operators has been developed in many different setting. In particular we are interested in studying the heat kernel associated to the subelliptic heat equation

$$\frac{\partial u}{\partial t} = Lu. \tag{3.19}$$

Definition 3.18. *Whenever the operator L is hypoelliptic, the equation* (3.19) *is usually called* ultraparabolic.

It is well-known that the heat kernel $p(t,x,y)$ for the heat equation $u_t = \Delta u$ in $\mathbb{R}^n$, is given by the following formula:

$$p(t,x,y) = \frac{1}{(4\pi t)^{\frac{n}{2}}} \exp\left(-\frac{|x-y|^2}{4t}\right). \tag{3.20}$$

Moreover, when L is uniformly elliptic, the equation (3.19) is parabolic and so there exists an associated heat kernel. In general, it is not possible to write an explicit formulation of the elliptic heat kernels but they satisfy almost the same properties than (3.20).
From now on we indicate an heat kernel (in both the uniformly elliptic case and the hypoelliptic one) by $p(t,x,y)$.
Remark that any parabolic equations can be seen as an heat equation on a Riemannian manifold $M = (\mathbb{R}^n, d)$, where $d(x,y)$ is a suitable associated Riemannian distance. So by Li-Yau's Theorem ([69]), we know that in particular the following exponential estimates hold

$$\frac{c}{\mathcal{L}^n(B^d(x,\sqrt{t}))} \exp\left(-\frac{d(x,y)^2}{ct}\right) \leq p(t,x,y) \leq \frac{C}{\mathcal{L}^n(B^d(x,\sqrt{t}))} \exp\left(-\frac{d(x,y)^2}{Ct}\right), \tag{3.21}$$

with $c, C > 0$ suitable constants.

All the previous remarks about the uniformly elliptic heat kernels are still true in the hypoelliptic case.
Next we introduce the definition of heat kernels and then we show exponential estimates similar to (3.21) for the hypoelliptic case (for details see [60]).
Let μ be a smooth, non-vanishing, measure on $\mathbb{R}^n$.
We indicate by H_t the semigroup generated by L w.r.t. the measure μ, i.e. H_t is the unique continuous family of operators defined on $L^1(d\mu)$ and such that, for any $s, t > 0$, all the following properties hold:

(i) $H_t \cdot H_s = H_{t+s}$,

(ii) H_t is semi-positive definite (i.e. $f \geq 0$ implies $H_t f \geq 0$),

(iii) $H_t 1 = 1$,

(iv) $\lim_{t\to 0^+} \|t^{-1}(H_t f - f) - H_0 f\|_{L^\infty(\mathbb{R}^n)} = 0$, for any $f \in C^2(\mathbb{R}^n)$.

The heat kernel $p(t,x,y)$ can be defined as the unique function such that

$$H_t f(x) = \int_M p(t,x,y) f(y) d\mu(y), \quad \text{for any}\, t>0 \text{ and } f \in L^1(d\mu),$$

i.e. $p(t,x,y)$ is the fundamental solution of the operator $L^x = \sum_{i=1}^m X_i^2(x)$, that is the unique solution of

$$L^x\, p(t,x,y) = \delta_{(0,y)}(t,x)$$

in $[0,+\infty) \times \mathbb{R}^n$ in the distribution sense.

An explicit formula for the heat kernel in the Heisenberg group is given in [49]. For the general hypoelliptic case, one can find many details about the properties of the fundamental solution in [62, 61]. There exists also a more recent paper [65] by E. Lanconelli and A. Pascucci in the general case where the operator depends on both the space and the time.

Remark 3.19. *Recall that any hypoelleptic fundamental solution is smooth outside the diagonal, then the corresponding heat kernel is smooth for any $t > 0$ and $x, y \in \mathbb{R}^n$.*

We are now interested in giving some exponential estimates for the heat kernel of sum of squares of Hörmander vector fields.
Since we look at the usual Lebesgue measure on $\mathbb{R}^n$, from now to on by μ we will always mean the Lebesgue measure $\mathcal{L}^n$. Nevertheless, most of the next results hold also for general smooth non-vanishing measures defined on connected manifolds.

Then, let $d(x,y)$ be the Carnot-Carathéodory distance induced by $X_1, ..., X_m$ on $\mathbb{R}^n$, we indicate by $B^d(x,r)$ the open ball in the metric $d(x,y)$ and by $B(x,r)$ the usual Euclidean ball, both centered at x and with radius equal to $r > 0$.
In [84] there is remarked that, by the Campbell-Hausdorff formula (1.56),

if $k \geq 1$ is the step of the distribution associated to $X_1, ..., X_m$, then there exists $c > 0$ such that for any $r > 0$ and $x \in \mathbb{R}^n$ the following inclusion holds:

$$B(x, cr^k) \subset B^d(x, r). \tag{3.22}$$

Exponential estimates for hypoelliptic heat kernels in the case of the horizontal Lapalcian have been proved first by A. Sánchez-Calle in [84] for $0 < t \leq 1$ and $d(x, y) < t^{\frac{1}{2}}$, and then generalized by D.S. Jerison and A. Sánchez-Calle in [60] for $0 < t \leq 1$ and $x, y \in K$ compact, and, finally, by S. Kusuoka and D. Stroock in [63], for $0 < t \leq 1$ and any $x, y \in \mathbb{R}^n$.
We recall the most general result by S. Kusuoka and D. Stroock, which tells that there exists $M \in [1, +\infty)$ such that

$$\frac{1}{M\mu(B^d(x, \sqrt{t}))} \mathrm{e}^{-M\frac{d(x,y)^2}{t}} \leq p(t, x, y) \leq \frac{M}{\mu(B^d(x, \sqrt{t}))} \mathrm{e}^{-\frac{d(x,y)^2}{Mt}}. \tag{3.23}$$

By the previous estimate and the inclusion (3.22), we can deduce an upper-bound for hypoelliptic heat kernels, which will turn out to be very useful to generalize the result of Varadhan, by using methods of measure theory.

Corollary 3.20. *Let $p(t, x, y)$ be the heat kernel associated to the hypoelliptic operator* (3.18) *and $k \geq 1$ be the step of the corresponding distribution. Then there exist two constants $C_1, C_2 > 0$ such that*

$$p(t, x, y) \leq C_1 t^{-\frac{kn}{2}} \mathrm{e}^{-C_2\frac{d(x,y)^2}{t}}, \tag{3.24}$$

for $0 < t \leq 1$ and $x, y \in \mathbb{R}^n$.

Proof. By the inclusion (3.22) and the monotonicity of the Lebesgue measure, we can deduce

$$\mu(B(x, ct^{\frac{k}{2}})) \leq \mu(B^d(x, t^{\frac{1}{2}})).$$

Therefore the upper bound given in (3.23) implies (3.24) with $C_1 = \frac{M}{\omega_n c^n}$, where ω_n is the measure of the unit Euclidean ball in $\mathbb{R}^n$, and $C_2 = 1/M$. $\square$

3.3.2 Limiting behavior of solutions of subelliptic heat equations.

The aim of this section is to generalize the known Euclidean convergence to the Carnot-Carathéodory case, following the ideas introduced by Capuzzo Dolcetta in [26], and recalled in Sec. 3.2.1.
Therefore we start looking at the Cauchy problem for the Hamilton-Jacobi equation:

$$\begin{cases} u_t + \dfrac{1}{2}|\sigma(x)Du|^2 = 0, & x \in \mathbb{R}^n,\ t > 0, \\ u(0,x) = g(x), & x \in \mathbb{R}^n. \end{cases} \tag{3.25}$$

with $g \in BC(\mathbb{R}^n)$ and $\sigma(x)$ $m \times n$-matrix (with $m \leq n$), satisfying the Hörmander condition.

In Theorem 2.77 we have proved that the metric Hopf-Lax function (2.48) is a LSC-viscosity solution (and an almost everywhere solution, too) for the Hamilton-Jacobi-Cauchy problem (3.25). In this particular case, the Hopf-Lax function coincides with the Carnot-Carathéodory inf-convolution. Moreover $g \in BC(\mathbb{R}^n)$, by applying Proposition 2.78 and Remark 2.79, the Carnot-Carathéodory inf-convolutions is a viscosity solution in the classical viscosity sense introduced by Crandall and Lions.
Moreover, whenever the initial data is continuous in $\mathbb{R}^n$, the comparison principle proved in [42] implies that the Carnot-Carathéodory inf-convolution (3.3) is the unique viscosity solution of the Cauchy problem (3.25).

Next we approximate problem (3.25) by second-order PDEs. Let us define

$$A(x) := \sigma^t(x)\sigma(x) = (a_{i,j}(x))_{i,j},$$

for $i, j = 1, ..., n$. We introduce the following second-order operator:

$$Lu = \sum_{i,j=1}^{n} a_{i,j}(x)\frac{\partial^2 u}{\partial x_i \partial x_j}. \tag{3.26}$$

Since $\sigma(x)$ is a $m \times n$-matrix of Hörmander-type, the operator L is hypoelliptic and, it is usually known as *horizontal Laplacian*.

By adding a second-order term $-\varepsilon Lu$, we can approximate (3.25) by the following family of Cauchy problems:

$$\begin{cases} u_t^\varepsilon - \varepsilon \displaystyle\sum_{i,j=1}^n a_{i,j}(x) \frac{\partial^2 u^\varepsilon}{\partial x_i \partial x_j} + \frac{1}{2} |\sigma(x) D u^\varepsilon|^2 = 0, \quad x \in \mathbb{R}^n,\ t > 0, \\ \qquad\qquad\qquad\qquad u^\varepsilon(x,0) = g(x), \quad x \in \mathbb{R}^n. \end{cases} \tag{3.27}$$

Using the Hopf-Cole transform $w^\varepsilon = \mathrm{e}^{-\frac{u^\varepsilon}{2\varepsilon}}$, we can linearize the Cauchy problems (3.27). We first compute the derivatives of the Hopf-Cole transform:

$$\begin{aligned} w_t^\varepsilon - \varepsilon \sum_{i,j=1}^n a_{i,j}(x) \frac{\partial^2 w^\varepsilon}{\partial x_i \partial x_j} = \\ = -\frac{w^\varepsilon}{2\varepsilon} \left(u_t^\varepsilon - \varepsilon \sum_{i,j=1}^n a_{i,j}(x) \frac{\partial^2 u^\varepsilon}{\partial x_i \partial x_j} + \frac{1}{2} \sum_{i,j=1}^n a_{i,j}(x) \frac{\partial u^\varepsilon}{\partial x_i} \frac{\partial u^\varepsilon}{\partial x_j} \right). \end{aligned}$$

Note that

$$|\sigma(x) D u^\varepsilon|^2 = \sigma(x) D u^\varepsilon \cdot \sigma(x) D u^\varepsilon = \sigma^t(x) \sigma(x) D u^\varepsilon \cdot D u^\varepsilon$$

and using that $A(x) = \sigma^t(x)\sigma(x)$, we can deduce

$$|\sigma(x) D u^\varepsilon|^2 = A(x) D u^\varepsilon \cdot D u^\varepsilon = \sum_{i,j=1}^n a_{i,j}(x) \frac{\partial u^\varepsilon}{\partial x_i} \frac{\partial u^\varepsilon}{\partial x_j}.$$

Therefore, if u^ε is a solution of the Cauchy problem (3.27), then its Hopf-Cole transform w^ε solves the following family of Cauchy problems for ultraparabolic second-order PDEs:

$$\begin{cases} w_t^\varepsilon - \varepsilon \displaystyle\sum_{i,j=1}^n a_{i,j}(x) \frac{\partial^2 w^\varepsilon}{\partial x_i \partial x_j} = 0, \quad x \in \mathbb{R}^n,\ t > 0, \\ \quad w^\varepsilon(x,0) = \mathrm{e}^{-\frac{g(x)}{2\varepsilon}} := g^\varepsilon(x), \quad x \in \mathbb{R}^n. \end{cases} \tag{3.28}$$

By the theory for ultraparabolic equations recalled in the previous subsection, we know that for any $\varepsilon > 0$ there exists an heat kernel $p^\varepsilon(t,x,y)$ associated to the hypoelliptic operator L. Then we can write the solution w^ε of the Cauchy problem (3.28) as

$$w^\varepsilon(t,x) = \int_{\mathbb{R}^n} p^\varepsilon(t,x,y) e^{-\frac{g(y)}{2\varepsilon}} dy. \tag{3.29}$$

Applying the anti-transform of Hopf-Cole, we formally get the following solutions for the approximating Cauchy problems (3.27):

$$u^\varepsilon(t,x) = -2\varepsilon \log \left(\int_{\mathbb{R}^n} p^\varepsilon(t,x,y) e^{-\frac{g(y)}{2\varepsilon}} dy \right). \tag{3.30}$$

One can expect that, as $\varepsilon \to 0^+$, the previous formalas converge to the unique solution of the original Hamilton-Jacobi-Cauchy problem (3.25), i.e. to the Carnot-Carathéodory inf-convolution of $g(x)$.
Exactly as in the Euclidean case, we are going to show this converge by applying the Theorem 3.9.
The difficulty consists again in proving the properties for the lower limit and the upper limit in Definition 3.8.
To this purpose, we fix $x \in \mathbb{R}^n$ and $t > 0$ and define the following probability measures

$$P^\varepsilon_{t,x}(B) := \int_B p^\varepsilon(t,x,y)dy, \tag{3.31}$$

for any $B \subset \mathbb{R}^n$ Borel set.

We want to show that the family of probability measures (3.31) satisfies the Large Deviation Principle with rate function

$$I_{t,x}(y) = \frac{d(x,y)^2}{4t}. \tag{3.32}$$

So we need to generalize the result proved by Varadhan in [93] for uniformly elliptic operators (see Sec. 3.2.2).

We first give an idea how to generalize to the hypoelliptic case the key-steps in the proof of the Varadhan.

We begin remarking that Lemma 3.13 still holds since the solution of the Cauchy problem (3.15) is unique also starting from hypoelliptic operators (see for example [65]).
Moreover, the Markov's property is true in the hypoelliptic case, too.

We then define the horizontal objects that we need. Let $\gamma : [0,T] \to \mathbb{R}^n$ be a σ-horizontal curve, we know that for $i = 1, ..., m$ there exist measurable functions h_i such that

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

So we indicate by $\Gamma(x, y)$ the set of all the absolutely continuous horizontal curves joining x to y in a finite time T. The length-functional (3.13) and the energy-functional (3.16) act on $\Gamma(x, y)$, respectively, as

$$l(\gamma) = \int_0^T \sqrt{h_1^2(t) + ... + h_m^2(t)} dt,$$

and

$$I(\gamma) = \frac{1}{2} \int_0^T [h_1^2(t) + ... + h_m^2(t)] dt.$$

Then the distance defined by (3.14) is the Carnot-Carathéodory distance associated to the matrix $\sigma(x)$.

Remark that Lemma 3.14 and Lemma 3.15 hold in the Riemannian case as well as in the sub-Riemannian case. In fact in both the cases minimizing the energy-functional or the length-functional is exactly the same thing (see [75]).

Theorem 3.16 is the main result in order to get the estimates for the upper limit and the lower limit. That result has been generalized to the hypoelliptic case first by Leandre in [67, 68], using probabilistic methods, and then by S. Kusuoka and D. Stroock in [64], by a different and simpler analytic proof. We quote the result below.

Theorem 3.21 ([67, 68, 64]). *Let $p(t, x, y)$ be the heat kernel of an hypoelliptic operator L and $d(x, y)$ be the associated sub-Riemannian distance, then for any $x, y \in \mathbb{R}^n$*

$$\lim_{\tau \to 0^+} 4\tau \log p(\tau, x, y) = -d(x, y)^2. \tag{3.33}$$

Moreover the previous limit is uniform whenever $d(x, y)$ is bounded.

Using Lemma 3.13 and Theorem 3.21, we are able to show the applicability of the Large Deviation Principle in the hypoelliptic case.
The proof of Varadhan seems to be applicable also in the sub-Riemannian case. Nevertheless that proof is very technical and probabilistic. We want to give a different and much shorter proof, using the same idea as one in Lemma 3.10 and standard methods of measure theory.
So we are going to prove first the limit in bounded sets and then use that result to investigate the limiting behavior in unbounded sets.

Lemma 3.22. *Let $p^\varepsilon(t,x,y)$ be the heat kernel associated to the hypoelliptic operator $L^\varepsilon u = \varepsilon \sum_{i,j=1}^n a_{i,j}(x)\frac{\partial^2 u}{\partial x_j \partial x_i}$, where $A(x) = (a_{i,j}(x))_{i,j=1}^n = \sigma^t(x)\sigma(x)$ and $\sigma(x)$ is a $m \times n$-matrix satisfying the Hörmander condition. If $P^\varepsilon_{t,x}$ and $I_{t,x}$ are the family of probability measures and the rate function defined in* (3.31) *and* (3.32)*, then for any $B \subset \mathbb{R}^n$ bounded and Borel set*

$$\lim_{\varepsilon\to 0^+}[P^\varepsilon_{t,x}(B)]^\varepsilon = \mathrm{e}^{-\inf_{y\in B} I_{t,x}(y)}. \tag{3.34}$$

Proof. We fix $t>0$, $x \in \mathbb{R}^n$ and a bounded and Borel set $B \in \mathbb{R}^n$.
Since the exponential function is continuous, we can write the locally uniform limit (3.33) as

$$\lim_{\tau\to 0^+} p(\tau,x,y)^\tau = \mathrm{e}^{-\frac{d(x,y)^2}{4}}. \tag{3.35}$$

Moreover

$$\left(\int_B p(\tau,x,y)dy\right)^\tau = \left(\int_B [p(\tau,x,y)^\tau]^{\frac{1}{\tau}} dy\right)^\tau,$$

so we set $q = \frac{1}{\tau}$ and define

$$f_q(y) := p\left(\frac{1}{q},x,y\right)^{\frac{1}{q}}.$$

We can then write

$$\left(\int_B [p(\tau,x,y)^\tau]^{\frac{1}{\tau}}\, dy\right)^\tau = \left(\int_B \left[p\left(\frac{1}{q},x,y\right)^{\frac{1}{q}}\right]^q dy\right)^{\frac{1}{q}} = \|f_q\|_{q,B},$$

where by $\| \quad \|_{q,B}$ we indicate the usual L^q-norm in the set B with $q \geq 1$.
By Lemma 3.13 and setting $\tau = \varepsilon t$, we easily get

$$\lim_{\varepsilon\to 0^+}[P^\varepsilon_{t,x}(B)]^\varepsilon = \lim_{\tau\to 0^+}\left(\int_B p(\tau,x,y)dy\right)^{\frac{\tau}{t}} = \lim_{q\to+\infty}\|f_q\|^{\frac{1}{t}}_{q,B}. \tag{3.36}$$

Note that, as $\varepsilon \to 0^+$, i.e. as $q \to +\infty$, the L^q-norm converges to the corresponding L^∞-norm in bounded sets. Hence by limit (3.35) we can deduce that

$$\lim_{q \to +\infty} f_q(y) = f(y) := \mathrm{e}^{-\frac{d(x,y)^2}{4}}, \qquad \text{local uniformly.}$$

By the triangle inequality for the L^q-norm, it follows that

$$\|f_q\|_{q,B} \leq \left(\int_B |f_q(y) - f(y)|^q dy \right)^{\frac{1}{q}} + \left(\int_B |f(y)|^q dy \right)^{\frac{1}{q}}. \tag{3.37}$$

Note that, if B is bounded, then

$$\lim_{q \to +\infty} \left(\int_B |f(y)|^q dy \right)^{\frac{1}{q}} = \|f\|_{\infty,B} = \sup_{y \in B} |f(y)| = \mathrm{e}^{-\inf_{y \in B} \frac{d(x,y)^2}{4}}.$$

It remains to show that the first term of (3.37) goes to zero as $q \to +\infty$. Using the locally uniform limit (3.35), this is immediate to prove. In fact,

$$0 \leq \left(\int_B |f_q(y) - f(y)|^q dy \right)^{\frac{1}{q}} \leq \|f_q - f\|_{\infty,B} \left[\mathcal{L}^n(B)\right]^{\frac{1}{q}} \longrightarrow 0,$$

To sum up, we have proved that

$$\lim_{q \to +\infty} \|f_q\|_{q,B}^{\frac{1}{t}} \leq \mathrm{e}^{-\inf_{y \in B} \frac{d(x,y)^2}{4t}}.$$

In order to get the reverse inequality, we proceed analogously remarking that

$$\|f_q\|_{q,B} \geq \left(\int_B |f_q(y) - f(y)|^q dy \right)^{\frac{1}{q}} - \left(\int_B |f(y)|^q dy \right)^{\frac{1}{q}}. \tag{3.38}$$

From that we deduce

$$\lim_{q \to +\infty} \|f_q\|_{q,B}^{\frac{1}{t}} \geq \mathrm{e}^{-\inf_{y \in B} \frac{d(x,y)^2}{4t}},$$

getting so limit (3.34). □

It is known that for general measures the convergence in bounded sets gives directly the corresponding estimate for the lower limit in the opens sets. The corresponding estimate for the upper limit in closed sets is instead much more difficult to prove. Also in our case it is very easy to study the limiting behavior in open unbounded sets, as we state in the following theorem.

Theorem 3.23. *Let $p^\varepsilon(t,x,y)$ be the heat kernel associated to the hypoelliptic operator $L^\varepsilon = \varepsilon \sum_{i,j=1}^n a_{i,j}(x)\frac{\partial^2}{\partial x_j \partial x_i}$, with $A(x) = (a_{i,j}(x))_{i,j=1}^n = \sigma^t(x)\sigma(x)$ and $\sigma(x)$ $m \times n$-matrix satisfying the Hörmander condition. If $P^\varepsilon_{t,x}$ and $I_{t,x}$ are the family of probability measures and the rate function defined in* (3.31) *and* (3.32)*, then for any $A \subset \mathbb{R}^n$ open set*

$$\liminf_{\varepsilon\to 0^+}[P^\varepsilon_{t,x}(A)]^\varepsilon \geq \mathrm{e}^{-\inf_{y\in A} I_{t,x}(y)}. \tag{3.39}$$

Proof. Let $A \subset \mathbb{R}^n$ be an open set, we want to prove that (3.39) holds.
We show this by approximation. In fact, if we set

$$A_R := A \cap B_R(0),$$

then A_R is a bounded open set and so we can apply the convergence result proved in Lemma 3.22. Therefore

$$\liminf_{\varepsilon\to 0^+}[P^\varepsilon_{t,x}(A)]^\varepsilon \geq \liminf_{\varepsilon\to 0^+}[P^\varepsilon_{t,x}(A_R)]^\varepsilon = \mathrm{e}^{-\inf_{y\in A_R}\frac{d(x,y)^2}{4t}}.$$

Taking the supremum over $R>0$ and using the semicontinuity of the characteristic function of open sets, we can immediately conclude that

$$\liminf_{\varepsilon\to 0^+}[P^\varepsilon_{t,x}(A)]^\varepsilon \geq \sup_{R>0}\mathrm{e}^{-\inf_{y\in A_R}\frac{d(x,y)^2}{4t}} \geq \mathrm{e}^{-\inf_{y\in A}\frac{d(x,y)^2}{4t}}.$$

□

To get the corresponding estimate for the upper limit in closed unbounded sets, one needs in general some equi-tight property.

Definition 3.24. *We recall that a family of probability measures P^ε is equi-tight, if and only if, for any $\delta > 0$ there exists a compact set K, such that $P^\varepsilon(\mathbb{R}^n \backslash K) < \delta$.*

It is easy to see that the probability power-measures $\mu_\varepsilon = P^\varepsilon_{t,x}$ are equi-tight. Nevertheless the powers $[P^\varepsilon_{t,x}]^\varepsilon$ are not so. Therefore we need to investigate the limiting behavior μ_ε outside "large balls".
To this purpose, it is useful the following "upper limit version" of De l'Hôpital Theorem.

Lemma 3.25. *Let f and g be two real-valued differentiable functions such that*

$$\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} g(x) = 0$$

(or equivalently $= \pm\infty$), then

$$\limsup_{x\to 0^+} \frac{f(x)}{g(x)} \leq \limsup_{x\to 0^+} \frac{f'(x)}{g'(x)}. \tag{3.40}$$

Proof. Let be $\delta > 0$ and $x \in (0, \delta)$, by the Cauchy Theorem, there exist two points $\xi(x), \zeta(x) \in (0, \delta)$ such that

$$\frac{f(x)}{g(x)} = \frac{f'(\xi(x))}{g'(\zeta(x))}.$$

Therefore

$$\begin{aligned}\limsup_{x\to 0^+} \frac{f(x)}{g(x)} = \inf_{\delta>0} \sup_{x\in(0,\delta)} \frac{f(x)}{g(x)} &= \inf_{\delta>0} \sup_{x\in(0,\delta)} \frac{f'(\xi(x))}{g'(\zeta(x))} \\ &\leq \inf_{\delta>0} \sup_{x\in(0,\delta)} \frac{f'(x)}{g'(x)} = \limsup_{x\to 0^+} \frac{f'(x)}{g'(x)}.\end{aligned}$$

□

The following lemma gives an equi-upper limit which plays the role of an equi-tight property.

Lemma 3.26. *For any $\delta \in (0, 1)$, there exists $R_\delta > 0$ such that*

$$\limsup_{\tau\to 0^+} \left(\int_{\mathbb{R}^n \setminus \overline{B^d_{R_\delta}(x)}} p(\tau, x, y) dy \right)^\tau < \delta.$$

Proof. Set $B_R^- = \mathbb{R}^n \setminus \overline{B}^d_R(x)$ with $R > 0$, we use the exponential estimate (3.24), so that

$$\limsup_{\tau\to 0^+} \left(\int_{B_R^-} p(\tau, x, y) dy \right)^\tau \leq \limsup_{\tau\to 0^+} C_1^\tau \tau^{-\frac{nk}{2}\tau} \left(\int_{B_R^-} \mathrm{e}^{-C_2 \frac{d(x,y)^2}{\tau}} dy \right)^\tau.$$

It is trivial that

$$\lim_{\tau\to 0^+} (C_1 \tau^{-\frac{nk}{2}})^\tau = 1.$$

Hence, it remains only to estimate

$$L_R = \limsup_{\tau\to 0^+}\left(\int_{B_R^-} \mathrm{e}^{-C_2\frac{d(x,y)^2}{\tau}}dy\right)^\tau.$$

By using the continuity of the logarithm function, we can look at $\log L_R$ instead of L_R and apply Lemma 3.25 to the limit, so that we find

$$\begin{aligned}\log L_R = \limsup_{\tau\to 0^+}\frac{\log\int_{B_R^-} \mathrm{e}^{-C_2\frac{d(x,y)^2}{\tau}}dy}{\frac{1}{\tau}} &\leq \limsup_{\tau\to 0^+} -\tau^2\frac{\int_{B_R^-} \mathrm{e}^{-C_2\frac{d(x,y)^2}{\tau}}C_2\frac{d(x,y)^2}{\tau^2}dy}{\int_{B_R^-} \mathrm{e}^{-C_2\frac{d(x,y)^2}{\tau}}dy}\\ &= \limsup_{\tau\to 0^+} -C_2\frac{\int_{B_R^-} \mathrm{e}^{-C_2\frac{d(x,y)^2}{\tau}}d(x,y)^2dy}{\int_{B_R^-} \mathrm{e}^{-C_2\frac{d(x,y)^2}{\tau}}dy}.\end{aligned}$$

Since $y\in\mathbb{R}^n\backslash\overline{B}_R^d(x)$, then $d(x,y)\geq R$. Therefore we get

$$\log L_R \leq \limsup_{\tau\to 0^+} -C_2R^2\frac{\int_{B_R^-} \mathrm{e}^{-C_2\frac{d(x,y)^2}{\tau}}dy}{\int_{B_R^-} \mathrm{e}^{-C_2\frac{d(x,y)^2}{\tau}}dy} = -C_2R^2.$$

We can conclude that for any $R>0$

$$\limsup_{\tau\to 0^+}\left(\int_{B_R^-} p(\tau,x,y)dy\right)^\tau \leq \mathrm{e}^{-C_2R^2}.$$

Now let us take $0<\delta<1$, we can choose $R_\delta > \sqrt{\frac{-\log\delta}{C_2}}$ such that $\mathrm{e}^{-C_2R_\delta^2}<\delta$ and this concludes the proof. □

By the previous lemma, it is not difficult to show the upper limiting estimate in closed unbounded sets.

Theorem 3.27. *Let $p^\varepsilon(t,x,y)$ be the heat kernel associated to the hypoelliptic operator $L^\varepsilon = \varepsilon\sum_{i,j=1}^n a_{i,j}(x)\frac{\partial^2}{\partial x_j\partial x_i}$, with $A(x) = (a_{i,j}(x))_{i,j=1}^n = \sigma^t(x)\sigma(x)$ and $\sigma(x)$ $m\times n$-matrix satisfying the Hörmander condition. If $P_{t,x}^\varepsilon$ and $I_{t,x}$ are the family of probability measures and the rate function defined in* (3.31) *and* (3.32)*, then for any $C\subset\mathbb{R}^n$ closed set*

$$\limsup_{\varepsilon\to 0^+}[P_{t,x}^\varepsilon(C)]^\varepsilon \leq \mathrm{e}^{-\inf_{y\in C} I_{t,x}(y)}. \tag{3.41}$$

Proof. Let C be a closed set; by using Lemma 3.13, we can take $\tau = \varepsilon t$. Then proving (3.41) is the same as proving that

$$\limsup_{\tau \to 0^+} \left(\int_C p(\tau, x, y) dy \right)^{\tau} \leq e^{-\inf_{y \in C} \frac{d(x,y)^2}{4}}. \tag{3.42}$$

Then we fix $\delta \in (0, 1)$ and choose $R_\delta > 0$ as in Lemma 3.26.
Since $\tau \in (0, 1)$, we can decompose $C = C_\delta \cup C_\delta$ with $C_\delta = C \cap \overline{B}^d_{R_\delta}(x)$ and $C_\delta = C \backslash \overline{B}^d_{R_\delta}(x)$. This gives

$$\begin{aligned} &\limsup_{\tau \to 0^+} \left(\int_C p(\tau, x, y) dy \right)^{\tau} \\ \leq &\limsup_{\tau \to 0^+} \left(\int_{C_\delta} p(\tau, x, y) dy \right)^{\tau} + \limsup_{\tau \to 0^+} \left(\int_{C_\delta^-} p(\tau, x, y) dy \right)^{\tau}. \end{aligned}$$

By applying Lemma 3.22 in C_δ and Lemma 3.26 in C_δ^-, we can deduce the following estimates:

$$\limsup_{\tau \to 0^+} \left(\int_C p(\tau, x, y) dy \right)^{\tau} \leq e^{-\inf_{y \in C_\delta} \frac{d(x,y)^2}{4}} + \delta \leq e^{-\inf_{y \in C} \frac{d(x,y)^2}{4}} + \delta.$$

Passing to the limit in the previous inequality as $\delta \to 0^+$, we find exactly (3.42). □

Using both Theorems 3.23 and 3.27, we can finally give the following result.

Theorem 3.28 ([45]). *Let be $g \in BC(\mathbb{R}^n)$ and $d(x, y)$ be the Carnot-Carathéodory distance associated to the Hörmander-matrix $\sigma(x)$. If $p^\varepsilon(t, x, y)$ is the heat kernel associated to the hypoelliptic operator $L^\varepsilon = \varepsilon \sum_{i,j=1}^n a_{i,j}(x) \frac{\partial^2}{\partial x_j \partial x_i}$, with $A(x) = (a_{i,j}(x))_{i,j=1}^n = \sigma^t(x)\sigma(x)$ and $\sigma(x)$ Hörmander-matrix, then*

$$\lim_{\varepsilon \to 0^+} -2\varepsilon \log \int_{\mathbb{R}^n} p^\varepsilon(t, x, y) e^{-\frac{g(y)}{2\varepsilon}} dy = \inf_{y \in \mathbb{R}^n} \left[g(y) + \frac{d(x, y)^2}{2t} \right]. \tag{3.43}$$

Proof. In order to prove the result, we need only to apply Theorem 3.9 to the continuous and bounded function $F = -g/2$, w.r.t. the family of probability measures defined by (3.31) and the rate function (3.32).

Properties (i) and (ii) hold, since $d(x, y)$ is a Carnot-Carathéodory distance satisfying the Hörmander condition and so $d(x, y)$ induces on $\mathbb{R}^n$ the Euclidean topology (Theorem 1.52). This means that $d(x, y)$ is continuous and the sublevels are compact sets.

In order to get properties (iii) and (iv), it remains only to apply the logarithm function (which is continuous and non decreasing) to the limiting estimates given by Theorem 3.23 and Theorem 3.27. $\square$

We conclude the section remarking that Theorem 3.28 generalizes the Euclidean result proved by I. Capuzzo Dolcetta to the hypoelliptic case.
Moreover, Theorem 3.28 let us also find back the known Euclidean and Riemannian result of Varadhan by using a new analytic proof.

Appendix A

The Legendre-Fenchel transform.

In this appendix we want to study some properties of the Legendre-Fenchel transform, looking in particular at positive and convex functions. This allows us to point out the meaning of assumptions $(H2)$ in order to prove Theorem 2.75.

Definition A.1. *Let $F : \mathbb{R} \to \mathbb{R}$ be a real-valued function, then the Legendre-Fenchel transform (or LF transform) $F^* : \mathbb{R} \to \mathbb{R}$ is defined as*

$$F^*(t) = \sup_{s \in \mathbb{R}} \{st - F(s)\}. \tag{A.1}$$

We like to recall that in physics, sometimes, formula (A.1) is called simply Legendre transform but this is not very exact. In fact, Legendre studied such a formula using a very different formulation, which holds only for differentiable (or, at least, convex) functions. More precisely, for differentiable functions we can equivalently define formula (A.1) by setting

$$F^*(t) = tx(t) - F(x(t)), \tag{A.2}$$

where $x(t)$ is found solving $F'(t) = t$.
Nevertheless the main results in the non-differentiable and non-convex case were developed by Fenchel by using exactly formula (A.1), hence the name.

The main property of the LF transform is given in the following trivial result.

Remark A.2. *The LF transform is a convex function.*

We indicate by $F^{**}(t)$ the double LF transform $(F^*)^*(t)$.
Note that $F^{**}(t) = F(t)$ if and only if $F(t)$ is convex. In such a case one can say that the LF transform is involutive.

For the LF transform, many other properties hold. One can find more information in [83, 90]. We want only to recall that the LF transform can been defined also for functions $F : \mathbb{R}^n \to \mathbb{R}$, simply by using the inner product $s \cdot t$ in formula (A.1), but we are not interested in this kind of generalization. Indeed we want to investigate the behavior of formula (A.1) whenever $F(t)$ is defined on some real result.
So let us take $F : [a, b] \to \mathbb{R}$ with $[a, b] \subset \mathbb{R}$, it is possible to extend the function $F(t)$ by setting $F(t) = +\infty$, for any $t \in \mathbb{R}\backslash[a, b]$. This implies that, whenever $F(t)$ is convex in $[a, b]$, the extended function $\overline{F}$ is convex in the whole $\mathbb{R}$. Moreover, the LF transform $\overline{F}^*(t)$ vanishes in $\mathbb{R}\backslash[a, b]$ and so we can look only at the restriction of $\overline{F}^*$ on $[a, b]$. If we apply the previous remarks to a function $\Phi(t)$ defined only for non-negative numbers, we find the definition of $\Phi^*(t)$ used in studying the Hopf-Lax function (2.48). More precisely we give the next definition.

Definition A.3. *For any function $\Phi : [0, +\infty) \to [0, +\infty)$, we call LF transform of $\Phi(t)$ the function $\Phi^* : [0, +\infty) \to \mathbb{R}$ defined by*

$$\Phi^*(t) = \sup_{s \geq 0}\{st - \Phi(s)\}. \tag{A.3}$$

Next we show some properties for the LF transform, which hold whenever $\Phi(t)$ satisfies assumptions $(H2)$. In particular we have used these properties in order to prove Theorem 2.75.

Lemma A.4. *Let $\Phi : [0, +\infty) \to [0, +\infty)$ be non-decreasing, convex and such that $\Phi(0) = 0$, then the LF transform $\Phi^*(t)$ defined by formula* (A.3) *is convex, non-negative and non decreasing with $\Phi^*(0) = 0$.*

Proof. We know that any LF transform is convex.
We can note that $\Phi(0)$ implies that $\Phi^*(t) \geq 0$ and $\Phi^*(0) = 0$.
So we can conclude remarking that, punctually,

$$t_1 s - \Phi(s) \leq t_2 s - \Phi(s),$$

for any $s \geq 0$ and $t_1 \leq t_2$; hence the non decreasing property follows. $\square$

Now we remain to explain the requirement on the derivative $\Phi'(t)$ at the origin. To this purpose, we introduce a different way in order to extend the original function $\Phi(t)$.

Remark A.5. *For functions as in* $(H2)$*, we can define the extended function to the whole* $\mathbb{R}$ *by setting*

$$\overline{\Phi}(t) = \begin{cases} \Phi(t), & t \geq 0, \\ 0, & t < 0. \end{cases}$$

$\overline{\Phi}$ *is convex in the whole* $\mathbb{R}$*, whenever* $\Phi(t)$ *is convex in* $[0, +\infty)$*. Then, the LF transform* $\overline{\Phi}^*$ *is the same than* (A.3) *in* $[0, +\infty)$*, and it is eqaul to* $+\infty$ *otherwise.*

The extended function $\overline{\Phi}$ given in Remark A.5 is a continuous function, since $\Phi \in C([0, +\infty))$ and $\Phi(0) = 0$.
Moreover, starting from a function $\Phi \in C^1(([0, +\infty))$, the extended function $\overline{\Phi}$ is C^1 if and only if $\lim_{t \to 0^+} \Phi'(t) = 0$.
Form now on, by $\overline{\Phi}$ we will always mean the extended function given in Remark A.5 and we write $\Phi'(0)$ meaning $\lim_{t \to 0^+} \Phi'(t)$.
The next proposition gives the key-property of the LF transform for strictly convex functions satisfying assumptions $(H2)$. Such a property has been applied to the strictly convex regularization $\Phi_\delta(t)$, in order to prove that the metric Hopf-Lax function solves, in the viscosity sense, the Cauchy problem (2.66) (Theorem 2.75).

Proposition A.6. *Assume that* $\Phi : [0, +\infty) \to [0, +\infty)$ *is* C^1*, non decreasing, strictly convex, such that* $\Phi(0) = 0$ *and* $\Phi'(0) = 0$*, then the LF transform* $\Phi^*(t)$*, defined by formula* (A.3)*, is strictly increasing in* $[0, +\infty)$*.*

Proof. We need only to show that for any $t > 0$, we have $\Phi^*(t) > 0$.
Hence this implies the proposition by using the other assumptions on $\Phi(t)$.
First we recall that $\overline{\Phi} \in C^1(\mathbb{R})$. Let us define $\psi(s) = ts - \Phi(s)$ and look at $t > 0$, we can note that $\psi \in C^1(\mathbb{R})$ and $\psi(0) = 0$.
Moreover $\Phi'(0) = 0$ implies that $\psi'(0) = t > 0$. Since $\Phi(t)$ is strictly convex, then the function is superlinear, so that $\lim_{s\to+\infty} \psi(s) = -\infty$. Therefore we can conclude that for any fixed $t > 0$,

$$\Phi^*(t) = \sup_{s\geq 0} \psi(s) > 0.$$

□

From the previous proposition the next result follows immediately.

Corollary A.7. *Under assumptions of Proposition A.6, then the LF transform $\Phi^*(t)$ is invertible in $[0, +\infty)$ and its inverse is strictly increasing, too.*

Proof. By Proposition A.6 $\Phi^*(t)$ is invertible. We need only to remark that since $\Phi^*(t)$ is strictly increasing, its inverse function is so: in fact

$$((\Phi^*)^{-1})'(t) = ((\Phi^*)')^{-1}(t) > 0.$$

□

We want to conclude pointing out the behavior of the LF transform (A.1) at the non-differentiability points, in order to better understand the rule of assumption $\Phi'(0) = 0$ in geting a LF transform strictly monotone (and so invertible).
First we remark that, whenever a convex function $\Phi(t)$ is differentiable at some point x, then the subdifferential $\partial\Phi(x)$ is a single point. This means that there exists a unique supporting line at the point x (ses Fig. A.1).

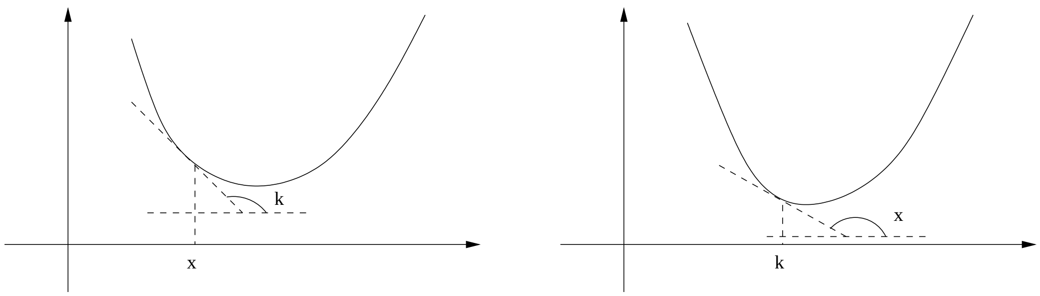


Figure A.1: `LF transform at a point of differentiability.`

Instead, whenever x is a point where the function is not differentiable, there are infinite supporting lines at x. Let be l and k, respectively, the minimum-slope and the maximum-slope of the supporting lines at x, then the LF transform is a straight-line from l to k, with slope exactly equal to x (see Fig. A.2).

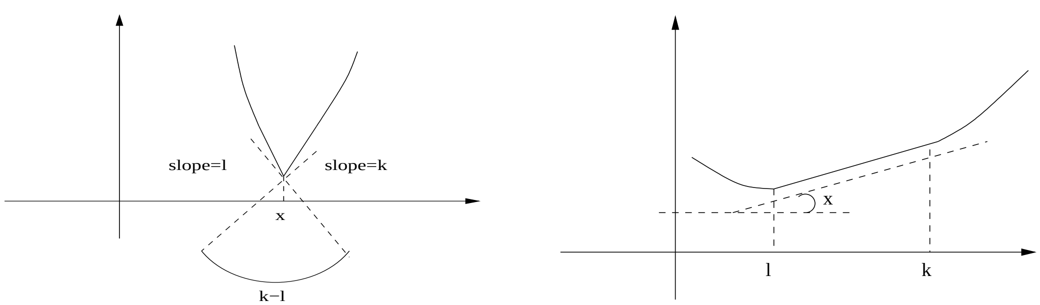
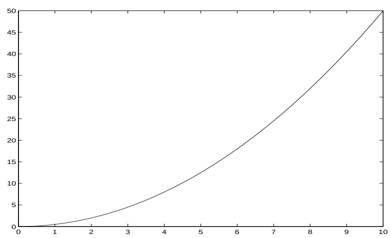


Figure A.2: `LF transform at a non-differentiability point.`

Note that in our case the function $\Phi(t)$ can be not differentiable only at the point $x = 0$. Then for non-negative points the LF transform is an horizontal straight-line exactly in $[0, k]$, with $k = \lim_{t\to 0^+} \Phi'(t) \geq 0$. Hence, if we assume $k = 0$, $\Phi^*(t)$ does not have any horizontal straight-line piece and therefore $\Phi^*(t)$ is strictly increasing.
We conclude giving some examples and some counter-examples.

Example A.8. *For any strictly convex power* $\Phi(t) = \frac{1}{\alpha}t^\alpha$ *with* $\alpha > 1$ *it is immediate that all the previous assumptions are satisfied. In the following pictures (Fig. A.3 and Fig. A.4), there are two different examples of strictly* *convex powers, on the left-hand side, and the corresponding LF transforms on the right-hand side.*



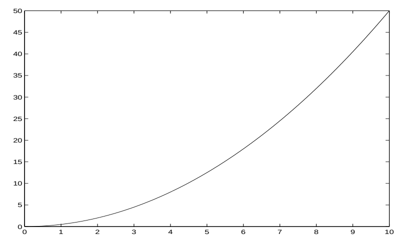


Figure A.3: `Case` $\alpha = 2$.

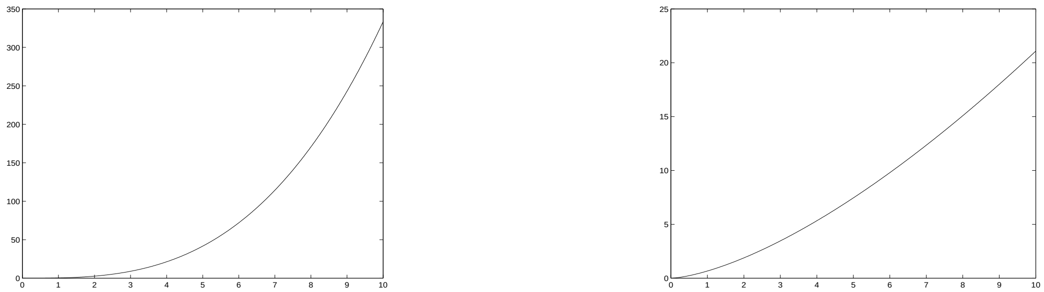


Figure A.4: Case $\alpha = 3$.

Note that in both these cases the LF transform is strictly increasing.

Example A.9. *The linear function* $\Phi(t) = t$ *does not satisfy all the previous assumptions. In fact, the function is not strictly convex and* $\Phi'(t) = 1$ *for any* $t \in \mathbb{R}$.
Therefore the corresponding LF transform vanishes from 0 *to* 1*, and it is not invertible (see Fig. A.5).*

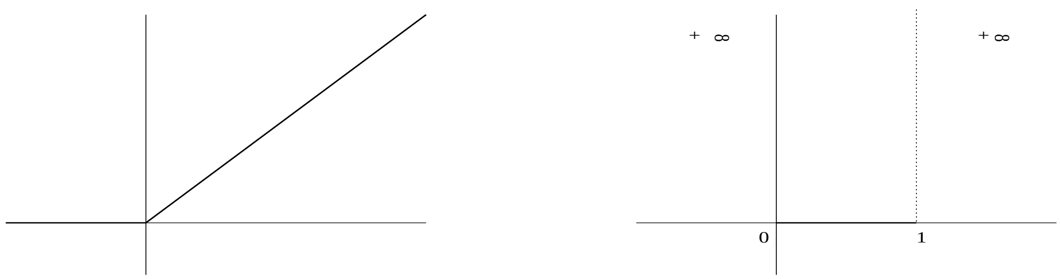


Figure A.5: The function $\Phi(t) = t$ (on the left-hand side) and the corresponding LF transform (on the right-hand side).

Remark A.10. *If we look at the strictly convex approximations of* $\Phi(t) = t$ *given in Theorem 2.75, that is* $\Phi_\delta(t) = t + \frac{\delta}{2}t^2$*, we get* $\Phi'(0) = \lim_{t \to 0^+} 1 + \delta t = 1$ *for any* $\delta > 0$*. So assumptions* $(H2)$ *are not satisfied.*

Example A.11. *A strictly convex function, satisfying all the assumptions except* $\Phi'(0) = 0$ *is the exponential function translated in the origin, that is* $\Phi(t) = \mathrm{e}^t - 1$*. Also in this case we get* $\Phi'(0) = \lim_{t \to 0^+} \mathrm{e}^t = 1$ *and so, as one*

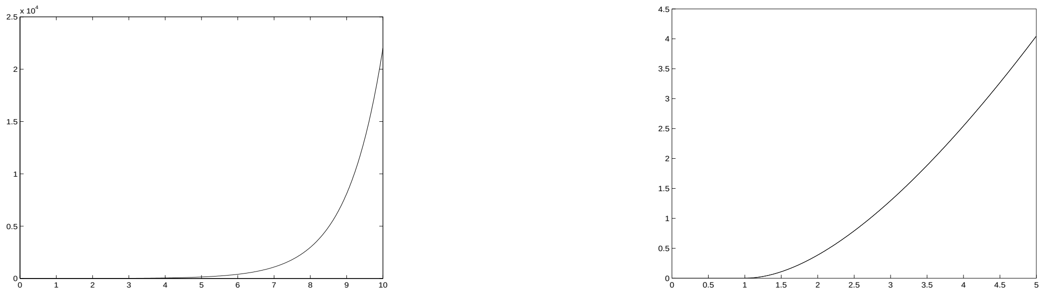


Figure A.6: `The translated exponential function (on the left-hand side) and its LF transform (on the right-hand side).`

can see in Fig. A.6, the LF transform is not invertible in the interval $[0,1]$.

Remark A.12. *For the functions given in Examples A.9 and A.11 the proof of Theorem 2.75 does not hold.*

Example A.13. *We conclude giving an example of a non-standard function satisfying* $(H2)$. *So, let* $\Phi(t) = \mathrm{e}^t - t - 1$, *it is trivial that* $\Phi(0) = 0$ *and* $\Phi'(0) = 1 - 1 = 0$. *Moreover* $\Phi'(t) = \mathrm{e}^t - 1 \geq 0$ *and* $\Phi''(t) = \mathrm{e}^t \geq 1 > 0$ *for any* $t \geq 0$, *then* $\Phi(t)$ *is non decreasing and strictly convex.*

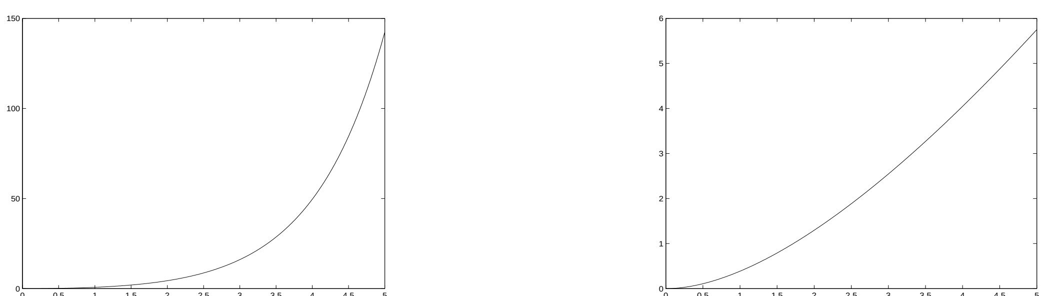


Figure A.7: $\Phi(t) = \mathrm{e}^t - t - 1$ `(on the left-hand side) and its LF transform (on the right-hand side).`

Bibliography

[1] J.F. Adams. *Lectures on Lie Groups*. Benjamin, New York, 1969.

[2] A. Agrachev, J.P. Gauthier. *On the Dido problem and plane isoperimetric problems*. Acta Appl. Math. 57, N.3 (1999), 287-338.

[3] O. Alvarez, E.N. Barron, H. Ishii. *Hopf-Lax formula for semicontinuous data*. Indiana University Math. J. 48, N.3 (1999), 993-1035.

[4] L. Ambrosio, S. Rigot. *Optimal mass transportation in Heinsenberg group*. J. Funct. Anal. 208, N.2 (2004), 261-301.

[5] J.P. Aubin, A. Cellina. *Differential Inclusions: Set-Valued Maps and Viability Theory*. Springer, Berlin, 1984.

[6] J.P. Aubin, H. Frankowska. *Set-Valued Analysis*. Birkhäuser, Boston, 1990.

[7] M. Bardi, I. Capuzzo Dolcetta. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Birkhäuser, Boston, 1997.

[8] M. Bardi, L.C. Evans. *On Hopf's formulas for solutions of Hamilton-Jacobi equations*. Nonlinear Analysis Theory Methods and Applications 8, N.II (1984), 1373-1381.

[9] G. Barles. *Existence results for first-order Hamilton-Jacobi equations*. Inst. H. Poincaré Ann. Non Linéaire 1 (1984), 325-340.

[10] G. Barles. *Discontinuous viscosity solutions of first-order Hamilton-Jacobi equations: a guided visit*. Nonlinear Anal. 20, N.9 (1993), 1123-1134.

[11] G. Barles. *Solutions de Viscosité des Équations de Hamilton-Jacobi.* Springer-Verlag, Berlin, 1994.

[12] G. Barles, B. Perthame. *Disontinuous solutions of deterministic optimal stopping time problems.* Math. Mod. Num. Analysis 21, N.4 (1987), 557-579.

[13] G. Barles, B. Perthame. *Exit time problems in optimal control and vanishing viscosity method.* SIAM J. Control Opt. 26, II (1988), 1133-1148.

[14] E.N. Barron, R. Jensen. *Semicontinuous viscosity solutions for Hamilton-Jacobi equations with convex Hamiltonians.* Comm. PDE 15, N.12 (1990), 1173-1742.

[15] E.N. Barron, R. Jensen. *Optimal control and semicontinuous viscosity solutions.* Proc. Am. Math. Soc. 133, N.2 (1991), 397-402.

[16] F. Baudoin. *An introduction to the geometry of stochastic flows.* Imperial College press, London, 2004.

[17] A. Bellaïche. *The tangent space in Sub-Riemmanian geometry*, in *Sub-Riemannian Geometry*, A. Bellaïche and J.J. Risler Editors, Basel, Birkhauser, 1996.

[18] P. Bono, M.V. Tessitore. *Il mito di Didone.* Bruno Mondadori, Milano, 1998.

[19] S. Bortoletto. *The Bellman equation for constrained deterministic optimal control problems.* Diff. Int. Eq. 6, N.4 (1993), 905-924.

[20] U. Boscain, G. Charlot, J.-P. Gauthier, S. Guérin, H.-R. Jauslin. *Optimal control in laser-induced population transfer for two- and three-level quantum systems.* J. Math. Physics 43, N.5 (2002).

[21] B.W. Brockett. *Control Theory and Singular Riemannian Geometry*, New directions in applied mathematics. Springer, 1980.

[22] H. Busemann. *The Geometry of Geodesics*. Academic Press, New York, 1955.

[23] X. Cabré, L.A. Caffarelli. *Fully Nonlinear Elliptic Equations*. AMS, Colloquium Publications 45, Providence, 1995.

[24] P. Cannarsa, C. Sinestrari. *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*. Birkhäuser, Boston, 2004.

[25] I. Capuzzo Dolcetta. *The Hopf solution of Hamilton-Jacobi equations*. Elliptic and Parabolic Pb., 343-351, World Sci., River Edge, NJ, 2002.

[26] I. Capuzzo Dolcetta. *Representations of solutions of Hamilton-Jacobi equations*. Progr. Nonlinear Diff. Eq. Appl. 54, Birkhäuser, Basel, 2003.

[27] I. Capuzzo Dolcetta, H. Ishii. *Hopf formulas for state dependent Hamilton-Jacobi equations*. preprint.

[28] C. Carathéodoty. *Untersuchungen über die grundlagen der thermodynamik*. Math. Ann. 67, N.3 (1909), 355-386.

[29] C. Chevalley. *Theory of Lie Groups*. Princeton University Press, Princeton, 1946.

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

[31] G. Citti, M. Manfredini. *Blow-up in non-homogeneous Lie groups and rectifiability*. Huston J. Math. 31, N.2 (2005), 333-353.

[32] G. Citti, M. Manfredini. *Uniform estimates of the fundamental solution for a family of hypoelliptic operators,*. Potential Analysis 25 (2006), 147-164.

[33] G. Citti, M. Manfredini, A. Sarti. *From neural oscillations to variational problems in the visual cortex*. J. Physiology 97 (2003), 87-385.

[34] G. Citti, M. Manfredini, A. Sarti. *Neuronal oscillation in the visual cortex: Gamma-convergence to the Riemannian Mumford-Shah functional*. SIAM J. Math. Analysis 35, N.6 (2004), 1394-1419.

[35] G. Citti, A. Sarti. *A cortical based model of perceptual completion in the roto-traslation space*. Journal of Math. Imaging and Vision 24 (2006), 307-326.

[36] M.G. Crandall, L.C. Evans, P.L. Lions *Some properties of viscosity solutions of Hamilton-Jacobi equations*. Trans. Am. Mat. Soc. 282 (1984), 487-502.

[37] M.G. Crandall, H. Ishii, P.L. Lions. *Uniqueness of viscosity solutions of Hamilton-Jacobi equations revisited*. J. Math. Soc. Japan 39, N.4 (1987), 581-596.

[38] M.G. Crandall, H. Ishii, P.L. Lions. *User's guide to viscosity solutions of second order partial differential equations*. Bull. Am. Math. Soc. 27, N.1 (1992), 1-67.

[39] M.G. Crandall, P.L. Lions. *Condition d'unicité pour les solutions generalisées des équations de Hamilton-Jacobi du premier ordre*. C. R. Acad. Sci. Paris Sér. I Math 292 (1981), 183-186.

[40] M.G. Crandall, P.L. Lions. *Viscosity solutions of Hamiltonian-Jacobi equations*. Trans. Am. Math. Soc. 277 (1983), 1-42.

[41] M.G. Crandall, P.L. Lions. *On existence and uniqueness of solutions of Hamilton-Jacobi equations*. Nonlinear Analysis Theory Appl. 10, N.4 (1986), 353-370.

[42] A. Cutrì, F. Da Lio. *Comparison and uniqueness results for noncoercitive first-order Hamilton-Jacobi equations*. ESAIM-COCV 13, N.3 (2007), 484-502.

[43] A. Davini. *Smooth approximation of weak Finsler metrics*. Diff. Int. Eq. 18, N.5 (2005), 509-530.

[44] F. Dragoni. *Metric Hopf-Lax formula with semicontinuous data*. Discrete and Continuous Dynamical Systems 17, N. 4 (2007), 713-729.

[45] F. Dragoni. *Limiting behavior of solutions of subelliptic heat equations*. tNonlinear Diff. Equations and Applications 14, N.3-4 (2008), 429-443..

[46] M.P. Do Carmo. *Riemannian Geometry.* Birkhäuser, Basel, 1992.

[47] L.C. Evans. *Partial Differential Equations.* Am. Math. Soc., Providence RI, 1988.

[48] S. Gallot, D. Hulin, J. Lafontaine. *Riemannian Geometry.* Springer-Verlag, Berlin, 1990.

[49] B. Gaveau. *Principe de moindre action, propagation de la chaleur et estimées sous elliptiques sur certain groupes nilpotents.* Acta Math. 139 (1977), 95-153.

[50] G. Gentili, F. Podestà, E. Vesentini. *Lezioni di Geometria Differenziale.* Bollati Boringhieri, Torino, 1995.

[51] M. Gromov. *Carnot-Carathéodory spaces seen from within*, in *Sub-Riemannian Geometry*, A. Bellaïche and J.J. Risler Editors, Basel, Birkhauser, 1996.

[52] J. Heinonen. *Calculus on Carnot Groups.* Fall School in Analysis (Jyväskylä, 1994), 1-31.

[53] R. Hermann. *Geodesics of singular Riemannian metrics.* Bull. Am. Math. Soc. 79 (1973), 780-782.

[54] E. Hopf. *Generalized solutions of non-linear equations of first order.* J. Math. Mech. 14 (1965), 951-973.

[55] L. Hörmander. *Hypoelliptic second order differential operators.* Acta Math. 119 (1967), 147-171.

[56] H. Ishii. *Hamilton-Jacobi equations.* Bull. Fac. Sci. Eng. Chuo Univ. 28 (1985), 33-77.

[57] H. Ishii. *Perron's method for Hamilton-Jacobi equations.* Duke Math. J. 55, N.2 (1987), 369-384.

[58] H. Ishii. *Comparison results for Hamilton-Jacobi equations without growth condition on solutions from above.*Appl. Anal. 67 (1997), 357-372.

[59] H. Ishii. *A generalization of a theorem of Barron and Jensen and a comparison theorem for lower semicontinuous viscosity solutions*. Proc. Roy. Soc. Edimburg Sect. A 131, N.1 (2001), 137-151.

[60] D.S. Jerison, A. Sánchez-Calle. *Estimates for the heat kernel for a sum of squares of vector fields.* Indiana Univ. Math. J. 35, N.4 (1986), 835-854.

[61] A.E. Kogoj, E. Lanconelli. *X-elliptic operators and X-controll distances.* Ricerche di Matematica Supplemento 49 (2000), 223-243.

[62] A.E. Kogoj, E. Lanconelli. *One-side Liouville theorems for a class of hypoelliptic ultraparabolic equations.* Geom. An. PDE several complex variables 368 (2005), 305-312.

[63] S. Kusuoka, D. Stroock. *Applications of Malliavin calculus, part III.* J. Fac. Sci. Univ. Tokyo Sect. 1A Math. 34, N.2 (1987), 391-442.

[64] S. Kusuoka, D. Stroock. *Long time estimates for the heat kernel associated with a uniformly suelliptic symmetric second order operator.* Ann. Math. 127, N.1 (1998), 165-189.

[65] E. Lanconelli, A. Pascucci. *On the fundamental solution for hypoelliptic second order partial differential equations with non-negative characteristic form*. Ricerche Mat. 48, N.1 (1999), 81-106.

[66] J.M. Lasry, P.L. Lions. *A remark on regularization in Hilbert spaces.* Israel J. of Math. 55, N.3 (1986), 257-266.

[67] R. Léandre. *Majoration en temps petit de la densité d'une diffusion dégénérée*. Prob. Theory Related Fields 74, N.2 (1987), 289-294.

[68] R. Léandre. *Minoration en temps petit de la densité d'une diffusion dégénérée*. Journal of Functional Analysis 74, N.2 (1987), 399-414.

[69] P. Li, S.T. Yau. *On the parabolic kernel of the Schrödinger operator.* Acta Math. 156, N.3-4 (1986), 153-201.

[70] E.H. Lieb, M. Loss. *Analysis.* Am. Math. Soc., Providence, 1997.

[71] P.L. Lions. *Generalized Solutions of Hamilton-Jacobi Equations.* Research Notes in Mathematics 69, Pitman, 1982.

[72] W. Liu, H.J. Sussmann. *Shortest paths for sub-Riemannian metrics on rank-two distributions.* Mem. Am. Math. Soc. 118 (1995), 0-104.

[73] G. Lu, J.J. Manfredi, B. Stroffolini. *Convex functions on the Heisenberg group.* Calc. Var. 19 (2004), 1-22.

[74] J.J. Manfredi, B. Stroffolini. *A version of the Hopf-Lax formula in the Heisenberg group.* Comm. in Part. Diff. Eq. 27, N.5-6 (2002), 1139-1159.

[75] R. Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Application.* Am. Math. Soc., United States of America, 2002.

[76] R. Monti. *Distances, boundaries and surface measures in Carnot-Carathéodory spaces.* PDh Thesis, University of Trento, 1994. `www.unitn.it/dipartiment/mate/download/tesi_dott.doc`.

[77] R. Monti, F. Serra Cassano. *Surface measures in Carnot-Carathéodory spaces.* Calc. Var. Part. Diff. Eq. 13, N.3 (2001), 339-376.

[78] A. Nagel, E.M. Stein, S. Wainger. *Balls and metrics defined by vector fields I: basic properties.* Acta Math. 155 (1985), 103-147.

[79] K. Nomizu. *Lie Gruops and Differential Geometry.* Math. Soc. of Japan, Tokyo, 1956.

[80] P. Pansu. *Metriques de Carnot-Carathéodory et quasi isometries des espaces symmetriques de rang un.* Ann. Math. 129 (1989), 1-60.

[81] O. Perron. *Eine neue Behandlung der Randwertufgabe für $\Delta u = 0$.* Math. Z. 18 (1923), 42-54.

[82] P.K. Rashevsky. *Any two points of a totally nonholonomic space may be connected by an admissible line.* Uch. Zap. Ped. Inst. Im. Liebknechta (Russian), Ser. Phys. Math. 2 (1938), 83-94.

[83] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, 1997.

[84] A. Sánchez-Calle. *Fundamental solutions and geometry of the sum of squares of vector fields.* Invent. Math. 78 (1984), 143-160.

[85] S. Semmes. *On the nonexistence of bi-Lipschitz parametrizations and geometric about A_∞-weights.* Revista Matematíca Iberoamericana 12, N.2 (1996) 337-410.

[86] A. Siconolfi. *Metric character of Hamilton-Jacobi equations.* Trans. Am. Math. Soc. 355, N.5 (2003), 1987-2009.

[87] P.E. Souganidis. *Existence of viscosity solutions of Hamilton-Jacobi equations.* J. Diff. Eq. 56 (1985), 345-390.

[88] E.M. Stein. *Singular integrals and differentiability properties of functions.* Princeton University Press, Princeton, 1970.

[89] R.S. Strichartz. *Sub-Riemannian geometry.* Diff. Geom. 24 (1986), 221-263.

[90] H. Touchette. Legendre-Fenchel transform in a nutshell. `www.maths.qmul.ac.uk/ ht/archive1/lfth2.pdf`.

[91] V.S. Varadarajan. *Lie Groups, Lie Algebras, and Their Rappresentations.* Springer-Verlag, New York, 1984.

[92] S.R.S. Varadhan. *On the behavior of the fundamental solution of the heat equation with variable coefficients.* Comm. Pure Appl. Math. 20 (1967), 431-455.

[93] S.R.S. Varadhan. *Diffusion Processes in a Small Time Interval.* Comm. Pure Appl. Math. 20 (1967), 659-685.

[94] S.R.S. Varadhan. *Large Deviations and Applications.* Soc. for Industrial and Applied Math., Philadelphia PA, 1984.

[95] A.M. Vershik, V. Ya Gershkovich. *Nonholonomic dynamical systems, geometry of distributions and variational problems*, in *Dynamical System VII.* V.I. Arnold, S.P. Novikov, Encyclopedia of Math. Sciences 16, Springer, 1994.