UNIVERSITÄT
BERN

The Hopf-Lax formula in Carnot groups: a control theoretic approach

Z. M. Balogh, A. Calogero, and R. Pini

Research Report 2012-08

14.08.2012

Mathematics Institute
University of Bern
Sidlerstrasse 5
CH-3012 Bern
Switzerland

www.math.unibe.ch

The Hopf–Lax formula in Carnot groups: a control theoretic approach

Z. M. Balogh[*], A. Calogero[†], R. Pini[†]

July 9, 2012

Abstract

The purpose of this paper is to bring a new light on the state–dependent Hamilton–Jacobi equation and its connection with the Hopf–Lax formula in the framework of a Carnot group $(\mathbf{G}, \circ)$. The equation we shall consider is of the form

$$\begin{cases} u_t + \Psi(X_1 u, \ldots, X_m u) = 0 & (x,t) \in \mathbf{G} \times (0, \infty) \\ u(x, 0) = g(x) & x \in \mathbf{G}, \end{cases}$$

where $X_1, \ldots, X_m$ are a basis of the first layer of the Lie algebra of the group $\mathbf{G}$, and $\Psi : \mathbb{R}^m \to \mathbb{R}$ is a superlinear, convex function. The main result shows that the unique viscosity solution of the Hamilton–Jacobi equation can be given by the Hopf–Lax formula

$$u(x,t) = \inf_{y \in \mathbf{G}} \left\{ t \Psi^{\mathbf{G}} \left(\delta_{\frac{1}{t}} (y^{-1} \circ x) \right) + g(y) \right\},$$

where $\Psi^{\mathbf{G}} : \mathbf{G} \to \mathbb{R}$ is the $\mathbf{G}$–Legendre–Fenchel transform of Ψ, defined by a control theoretical approach. We recover, as special cases some known results: the classical Hopf–Lax formula in the Euclidean spaces $\mathbb{R}^n$ showing that $\Psi^{\mathbb{R}^n}$ is the Legendre–Fenchel trasnsform Ψ^* of Ψ; moreover, we recover the result by Manfredi and Stroffolini in the Heisenberg group for particular Hamiltonian function Ψ. In this paper we follow an optimal control problem approach and we obtain several properties for the value functions u and $\Psi^{\mathbf{G}}$: in particular we prove a precise estimate for the horizontal gradient of the solution u, two existence results of the optimal control for the optimal problems and we show that $\Psi^{\mathbf{G}}$ is convex.

Key words: Hamilton–Jacobi equation, Legendre–Fenchel transform, Carnot groups, viscosity solution, Hopf–Lax formula, optimal control problem

MSC: Primary: 35R03 Secondary: 49L25; 26B25; 22E30; 53C17

---

[*]Institute of Mathematics, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland (zoltan.balogh@math.unibe.ch), Research supported by the Swiss National Science Foundation, the European Science Foundation Project HCAA and the FP7 EU Commission Project CG-DICE.

[†]Dipartimento di Matematica e Applicazioni, Università degli Studi di Milano Bicocca, Via Cozzi 53, 20126 Milano, Italy (andrea.calogero@unimib.it, rita.pini@unimib.it)

Contents

1 Introduction

Hamilton–Jacobi equations play a major role in the theory of optimal mass transportation and logarithmic Sobolev inequalities and have been studied recently in the general context of geodesic metric measure spaces by Lott–Villani [28], Balogh–Engulatov–Hunziker–Maasalo [6], Ambrosio–Gigli–Savaré [3], [4] and Ambrosio and Di Marino [5] . In these works it was shown that even in the general setting of geodesic spaces the solution to the Hamilton–Jacobi equation can be expressed as an inf–convolution akin to the classical Hopf–Lax formula [26, 27].

As our starting point, let us recall (see e.g. [22]) that in the classical Euclidean space the solution of the state-independent Hamilton–Jacobi equation

$$\begin{cases} u_t(x,t) + H(Du(x,t)) = 0 & (x,t) \in \mathbb{R}^n \times (0,T) \\ u(x,0) = g(x) & x \in \mathbb{R}^n \end{cases} \tag{1}$$

under appropriate regularity assumptions: convexity and superlinearity of $H : \mathbb{R}^n \to \mathbb{R}$ and Lipschitz continuity of $g : \mathbb{R}^n \to \mathbb{R}$, can be represented by the Hopf–Lax formula

$$u(x,t) = \inf_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \tag{2}$$

that gives the unique viscosity solution of (1). Here and in the sequel, for a function $u : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, by abuse of notation we set $Du(x,t) = \nabla_x u(x,t)$. The function $L = H^* : \mathbb{R}^n \to \mathbb{R}$ is the Legendre–Fenchel transform of H given by

$$L(p) = \sup_{q \in \mathbb{R}^n} \{p \cdot q - H(q)\} \tag{3}$$

($x \cdot y$ will denote here and in the sequel the usual inner product in Euclidean spaces). For more general state–dependent Hamiltonians $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ solutions to the equation

$$u_t + H(x, Du) = 0$$

have no elegant representations similar to (2). In certain cases, however, (2) could be recovered by defining an appropriate sub–Riemannian structure on the space $\mathbb{R}^n$. For example, consider in $\mathbb{R}^3$ the equation

$$\begin{cases} u_t + \Phi\left(\left|\left(u_{x_1} - \dfrac{x_2}{2}u_{x_3},\ u_{x_2} + \dfrac{x_1}{2}u_{x_3}\right)\right|\right) = 0 & (x,t) \in \mathbb{R}^3 \times (0,\infty) \\ u(x,0) = g(x) & x \in \mathbb{R}^3 \end{cases} \tag{4}$$

where $|\cdot|$ will denote from now on the Euclidean norm, and $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$ is a superlinear and strictly convex increasing function. Manfredi and Stroffolini (see [29]) used the sub–Riemannian structure of the first Heisenberg group $I\!H$ to prove that the unique viscosity solution of (4) is given by

$$u(x,t) = \inf_{y \in I\!H} \left\{ t\Phi^*\left(\frac{d_{CC}(x,y)}{t}\right) + g(y) \right\}, \tag{5}$$

where d_{CC} is the sub–Riemannian (or Carnot–Carathéodory) metric in the first Heisenberg group, and Φ^* is the usual (one–dimensional) Legendre–Fenchel transform of Φ.

The essential observation of [29] was that the use of the sub–Riemannian metric in the Heisenberg group is given by the left invariant vector fields

$$X_1 = \partial_{x_1} - \frac{x_2}{2}\partial_{x_3}, \quad \text{and} \quad X_2 = \partial_{x_2} + \frac{x_1}{2}\partial_{x_3},$$

which appear in (4). In fact, one can easily see that in (4) the sub–Riemannian *metric gradient* (X_1u, X_2u) of u appears, and it is composed with the one–dimensional function Φ. This observation can be used to make the connection to the Hopf–Lax type formula (5), where the solution appears in terms of the sub–Riemannian metric of the Heisenberg group.

The above observation was generalized to more general sub–Riemannian geometries defined in terms of Hörmander vector fields by Dragoni in [20]. Moreover, an even more general version of this result is valid in the setting of geodesic metric spaces as shown in [28],[6], [3], [4],[5] .

Let observe that by the form of the equation (4) a strong condition of *homogeneity* is assumed on the Hamiltonian H. This is imposed by the fact that H is a composition by the metric gradient with the one–dimensional function Φ. Moreover, generalizations in geodesic metric settings ([28], [6],[3], [4],[5]) are also only valid under a similar homogeneity assumption.

A simple equation of non–homogeneous type such as

$$u_t + (X_1u)^2 + (X_2u)^4 = 0$$

is not covered by the aforementioned results.

The purpose of this paper is to go beyond the assumption of homogeneity and to study general *non–homogeneous* Hamilton–Jacobi equations in the Heisenberg, and more general Carnot groups. The equations we shall consider are of the form

$$u_t + \Psi(X_1u, \dots, X_mu) = 0,$$

where $m \leq n$, and $X_1, \dots, X_m$ are first order linear operators

$$X_i = \sum_{j=1}^{n} q_{j,i} \partial_{x_j}, \text{ for } i = 1, \dots, m,$$

with smooth (but non constant) coefficients $q_{j,i} : \mathbb{R}^n \to \mathbb{R}$. We shall also assume that $\Psi : \mathbb{R}^m \to \mathbb{R}$ is a superlinear, convex function.

We can view the differential operators X_i as vector fields over $\mathbb{R}^n$ and consider the system $\mathbb{X} = (X_1, \dots, X_m)$. In this way, for a function $u : \mathbb{R}^n \times [0, \infty) \to \mathbb{R}$ we denote by

$$\mathbb{X}u(x,t) := (X_1 u(x,t), \dots, X_m u(x,t)) \in \mathbb{R}^m \tag{6}$$

its $\mathbb{X}$–gradient, or horizontal gradient, at the point $(x,t) \in \mathbb{R}^n \times (0, \infty)$. Using this notation we shall consider the Hamilton–Jacobi boundary value problem:

$$\begin{cases} u_t(x,t) + \Psi(\mathbb{X}u(x,t)) = 0 & (x,t) \in \mathbb{R}^n \times (0,\infty), \\ u(x,0) = g(x) & x \in \mathbb{R}^n. \end{cases} \tag{7}$$

This setting allows us to consider state–dependent Hamiltonians of the form $H(x, Du(x,t)) = \Psi(\mathbb{X}u(x,t))$ since the coefficients of the vector fields X_i are in general state–dependent. Choosing $m = n$ and $X_i = \partial_{x_i}, i = 1, \dots, n$, we recover the classical case (1).

The main assumption in this paper is that the system of vector fields $\mathbb{X}$ forms the basis of the *first layer* V_1 in the Lie algebra $\mathfrak{g}$ of a Carnot group $(\mathbf{G}, \circ)$ with a family of isotropic dilations $\delta_t : \mathbf{G} \to \mathbf{G}$ for $t > 0$. We refer to [11] for the notation and terminology related to Carnot groups but we will review this material in the next section. Below we formulate the main result of this paper stated later as Theorem 3.4:

Theorem 1.1 *Assume that $\mathbb{X}$ is a basis of the first layer for a Carnot group $(\mathbf{G}, \circ)$ with isotropic dilations δ_t. Assume also that $\Psi : \mathbb{R}^m \to \mathbb{R}$ is convex and superlinear, and the boundary data $g : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous (in the Euclidean metric) and bounded.*

Then the unique viscosity solution to (7) is given by the Hopf–Lax type formula

$$u(x,t) = \inf_{y \in \mathbf{G}} \left\{ t\Psi^{\mathbf{G}} \left(\delta_{\frac{1}{t}}(y^{-1} \circ x) \right) + g(y) \right\}, \tag{8}$$

where $\Psi^{\mathbf{G}} : \mathbf{G} \to \mathbb{R}$ is the $\mathbf{G}$–Legendre–Fenchel transform of Ψ defined below.

The definition of $\Psi^{\mathbf{G}}$ is achieved in terms of an optimal control problem related to the system $\mathbb{X}$ of the group $\mathbf{G}$. Consider first the control dynamics $f^{\mathbf{G}} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ given by

$$f^{\mathbf{G}}(w, \alpha) = \sum_{i=1}^{m} \alpha_i X_i(w),$$

where X_i are viewed as vector fields, and the associated control problem

$$\begin{cases} \dot{w}(s) = f^{\mathbf{G}}(w(s), \alpha(s)) \\ w(0) = e, \end{cases} \tag{9}$$

where e denotes the unit element of $\mathbf{G}$. It turns out that, under our assumptions on $\mathbb{X}$, for every choice of the control function α there is a unique response, i.e. a solution to (9) denoted by w_α. Given $x \in \mathbb{R}^n$ we denote by $\mathcal{F}_x^{\mathbf{G}}(\mathbb{R}^m)$ the set of control functions steering e to x in unit time:

$$\mathcal{F}_x^{\mathbf{G}}(\mathbb{R}^m) := \{\alpha : [0,1] \to \mathbb{R}^m \text{ measurable} : \ w_\alpha(1) = x\}.$$

Observe that, since $\mathbb{X}$ is bracket generating, by Chow's theorem [10], it follows that our system is controllable, in particular $\mathcal{F}_x^{\mathbf{G}}(\mathbb{R}^m) \neq \emptyset$ for each $x \in \mathbb{R}^n$.

Denoting by Ψ^* the usual Legendre–Fenchel transform of Ψ we can define the **G**–Legendre–Fenchel transform by

$$\Psi^{\mathbf{G}}(x) := \inf_{\alpha \in \mathcal{F}_x^{\mathbf{G}}(\mathbb{R}^m)} \int_0^1 \Psi^*(\alpha(s))ds. \tag{10}$$

We prove that $\Psi^{\mathbf{G}} : \mathbf{G} \to \mathbb{R}$ is convex with respect to the group operations (see Definition 2.1 and Proposition 3.1) and hence admits a.e. the horizontal gradient $\mathbb{X}\Psi^{\mathbf{G}}$. In addition, by using Jensen's inequality, one can deduce that in the Euclidean case when $n = m$ and $X_i = \partial_{x_i}$, $i = 1, \dots, n$, it follows that $\Psi^{\mathbf{G}} = \Psi^*$ and thus (8) reduces to the classical Hopf–Lax formula (2). Similarly, we shall show that in the homogeneous case $\Psi^{\mathbf{G}}$ can be expressed in terms of the Carnot–Carathéodory distance on $\mathbf{G}$ and the results in [29] and [20] can be recovered.

In a general Carnot group $\mathbf{G}$, our version of the **G**–Legendre–Fenchel transform provides an extension of the notion of Ψ^* since the restriction of $\Psi^{\mathbf{G}}$ to $\exp(V_1)$ is reduced to Ψ^*.

In fact we can show that for functions g satisfying $g(\pi(x)) \leq g(x)$ for all $x \in \mathbb{R}^n$, where $\pi : \mathbb{R}^n \to \mathbb{R}^m \subset \mathbb{R}^n$ is the canonical projection $\pi(x_1, \dots, x_m, \dots, x_n) = (x_1, \dots, x_m)$; the Hopf–Lax formula reduces to a finite–dimensional optimal problem.

Corollary 1.1 *Suppose that $g(x) \geq \tilde{g}(x)$ for all $x \in \mathbb{R}^n$, where $\tilde{g} = g \circ \pi$. Then we obtain*

$$u(x,t) = \inf_{q \in \mathbb{R}^m} \left\{ t\Psi^*\left(\frac{\pi(x) - q}{t}\right) + \tilde{g}(q) \right\};$$

in particular, $x \to u(x,t)$ depends only on the first m components of x.

To comment on the proof of Theorem 1.1, we should mention that our control theoretic approach differs essentially from the techniques used in [29] or [20], and it is more related to the paper by Bardi and Evans [9]. The strategy to prove this result is long and contains several interesting results; let us spend a few words to describe the two steps of such strategy.

In our approach we use the well–known fact from control theory, that the value function associated to an optimal control problem is the solution of a related Hamilton–Jacobi–Bellman equation ([24], [22]). However, a direct application of this general principle to our situation faces technical difficulties.

Our first step it to consider a superlinear function ψ, (here the convexity is not required), and the **free endpoint optimal control problem**

$$\inf_{\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)} \left\{ \int_0^t \psi(\alpha(s))ds + g(w_\alpha(t)) \right\} \tag{11}$$

where $t > 0$ and $x \in \mathbf{G}$ are given, and $\overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)$ is the set of measurable control functions $\alpha : [0,t] \to \mathbb{R}^m$ such that there is a unique response to (9), where we replace, in (9), $f^{\mathbf{G}}$ with $-f^{\mathbf{G}}$ and e with x.

For the value function u of this optimal control problem (11), we construct the associated Hamilton–Jacobi–Bellman equation and classical arguments guarantee that u is Euclidean Lipschitz, and, in particular, its horizontal gradient $\mathbb{X}u$ is defined a.e. These results imply, that the Hamilton–Jacobi boundary value problem associated to (11) is

$$\begin{cases} u_t(x,t) + \psi^*(\mathbb{X}u(x,t)) = 0 & (x,t) \in \mathbf{G} \times (0,T) \\ u(x,0) = g(x) & x \in \mathbf{G} \end{cases}$$

and that its unique viscosity solution is precisely the value function u of the problem (11) (see Theorem 3.2). In this context, in Proposition 2.3 we provide a careful estimate of $\mathbb{X}u$.

The second step of our strategy is to consider a **fixed endpoint control problem** and to introduce the related value function function $\mathcal{L}(x,t)$ as

$$\mathcal{L}(x,t) := \inf_{\alpha \in \mathcal{F}^{\mathbf{G}}_{e,0,x,t}(\mathbb{R}^m)} \left\{ \int_0^t \psi(\alpha(s))ds \right\} \tag{12}$$

where $t > 0$ and $x \in \mathbf{G}$ are given, and $\mathcal{F}^{\mathbf{G}}_{e,0,x,t}(\mathbb{R}^m)$ is the set of measurable control functions $\alpha : [0,t] \to \mathbb{R}^m$ such that there is a unique response to (9) such that $w(t) = x$. Clearly $\Psi^{\mathbf{G}}(x) = \mathcal{L}(x,1)$ and the investigation of some properties of $\mathcal{L}$ gives, under the assumption that ψ is convex and superlinear, all the tools to prove the Hopf–Lax formula in Theorem 1.1.

The paper is organized as follows. In Section 2 we review preliminary notions about Carnot groups, and study the value functions u and $\mathcal{L}$ for the optimal control problems (11) and (12) respectively. In Section 3 we show that the function u is the unique viscosity solution for a Hamilton–Jacobi–Bellman equation; we introduce the $\mathbf{G}$–Legendre–Fenchel transform (10) and prove our main result, Theorem 1.1. In Section 4, we indicate how to deduce, from Theorem 1.1, the previous versions of Hopf–Lax formula (5) in [29] and the general case studied in [20]. In doing so we observe that, in general, it is hard to compute $\Psi^{\mathbf{G}}(x)$ in (10) explicitly, since the associated optimal control problem is complicated to solve. Qualitative information about $\Psi^{\mathbf{G}}$ can, nevertheless be obtained using Pontryagin's Minimun Principle. In Section 4 we discuss examples using Pontryagin's theorem. In the Appendix we prove an existence result for optimal control problems (12). Section 5 is devoted to final remarks and open questions.

Acknowledgement: We thank Martino Bardi and Juan Manfredi for stimulating conversations on the subject of this paper.

2 Optimal Control Problems in Carnot groups

2.1 Carnot groups

Several recent books are devoted to the study of Carnot groups: in this paper we refer to [11], using the same notations. A Carnot group $(\mathbf{G}, \circ)$ of step r is a connected, simply connected, nilpotent Lie group whose Lie algebra $\mathfrak{g}$ of left–invariant vector fields admits a stratification, i.e. there exist non zero subspaces $\{V_j\}_1^r$ such that

$$\mathfrak{g} = \bigoplus_{i=1}^r V_i, \qquad [V_1, V_j] = V_{j+1} \neq 0 \quad \text{for } j = 1, \dots r-1, \qquad [V_1, V_r] = 0.$$

We assume that a scalar product is given on $\mathfrak{g}$ for which the subspaces V_j are mutually orthogonal. The first layer V_1 of the Lie algebra plays a key role: its elements are called horizontal vectors. We denote by n_i the dimension of the vector space V_i and, in particular, we let $m = n_1$. Let $X = \{X_1, X_2, \ldots, X_n\}$ be an orthonormal basis of $\mathfrak{g}$ such that for every j (with $1 \le j \le r$) the set

$$\{X_i;\ \text{with } n_1 + \ldots + n_{j-1} < i \le n_1 + \ldots + n_j\}$$

is a basis for V_j (here we put $n_0 = 0$ and $N_j = n_0 + \ldots + n_j$ and hence $N_r = n$). With an abuse of notations, we keep on denoting by X the corresponding system of left–invariant vector fields on $\mathbf{G}$ defined by $X_i(x) = (L_x)_*(X_i),\ i = 1, \ldots, m$, where $(L_x)_*$ is the differential of the left translation on $\mathbf{G}$ defined by $L_x(y) = x \circ y$ (in the sequel we will drop the circle by writing xy instead of $x \circ y$). The system $\mathbb{X} = (X_1, X_2, \ldots, X_m)$ defines a basis for the horizontal sub–bundle $\mathcal{H}\mathbf{G}$ of the tangent bundle $\mathcal{T}\mathbf{G}$, i.e. $\mathcal{H}_x\mathbf{G} = \operatorname{span}\{X_1(x), \ldots, X_m(x)\}$ for every $x \in \mathbf{G}$.

The exponential map $\exp : \mathfrak{g} \to \mathbf{G}$ is defined by $\exp(x_1X_1 + \ldots + x_nX_n) = (x_1, \ldots, x_n)$, and it is a global diffeomorphism; we denote by $\xi = (\xi_1, \xi_2, \ldots, \xi_r)$ the inverse of exp, where $\xi_j : \mathbf{G} \to V_j$. A natural family of non–isotropic dilations on $\mathfrak{g}$ associated with its grading is given by $\tilde{\delta}_\lambda(v_1 + v_2 + \ldots + v_r) = \lambda v_1 + \lambda^2 v_2 + \ldots + \lambda^r v_r$, if $v_i \in V_i$, $1 \le i \le r$. By means of the exponential map, one lifts these dilations to the family of the automorphisms of $\mathbf{G}$ $\delta_\lambda(x) = \exp(\tilde{\delta}_\lambda(\xi(x)))$, i.e.

$$\delta_\lambda(x) = (\lambda^{\sigma_1}x_1, \ldots, \lambda^{\sigma_n}x_n),$$

with $x \in \mathbb{R}^n$, where $\sigma_i = j$ for $N_{j-1} < i \le N_j$. The homogeneous dimension associated with the dilations δ_λ is given by $Q = \sum_{i=1}^r \sigma_i n_i$ that often replaces the topological dimension n in the study of Carnot groups. The Baker–Campbell–Dynkin–Hausdorff formula for the bracket relations in the Lie algebra defines the group law on $\mathbf{G}$. More precisely, the group law is given by

$$xy = (x_1 + y_1, \ldots, x_m + y_m, x_{m+1} + y_{m+1} + Q_{m+1}(x, y), \ldots, x_n + y_n + Q_n(x, y)), \tag{13}$$

for every $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$; for every $m < i \le n$ with $N_{j-1} < i \le N_j$ and $2 \le j \le r$, the function $Q_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a homogeneous polynomial of degree σ_i with respect to the dilation δ_λ of the group $\mathbf{G}$. Moreover Q_i depends only on x_k, y_h with $1 \le k, h \le N_{j-1}$, and is a sum of terms each of which contains a factor of the type $(x_ky_h - x_hy_k)$ with $1 \le k, h \le N_{j-1}$. We denote by $e = 0 \in \mathbb{R}^n$ the null element in $\mathbf{G}$.

Starting from the law (13) of the group $\mathbf{G}$, we can obtain an expression for the system X in $\mathfrak{g}$ via the polynomials $\{Q_i\}$; indeed, for every $j = 1, \ldots, r$ and $N_{j-1} < i \le N_j$, we get that

$$X_i(x) = \partial_i + \sum_{s=N_j+1}^{n} q_{s,i}(x)\partial_s, \qquad \text{with} \quad q_{s,i}(x) = \frac{\partial Q_s}{\partial y_i}(x, 0). \tag{14}$$

Our convention is that $X_i = \partial_i$ for $j = r$ and we set $q_{s,i}(x) = 0$ if $s \le N_j$. It is not difficult to prove that for every $s > N_j$, the mentioned homogeneity of Q_s implies that the function $q_{s,i}$ is a homogeneous polynomial of degree $\sigma_s - \sigma_i$ with respect to the dilation δ_λ; in particular

$$q_{s,i}(\delta_\lambda(x)) = \lambda^{\sigma_s - 1}q_{s,i}(x), \qquad \forall x \in \mathbf{G},\ 1 \le i \le m. \tag{15}$$

For more details, see, e.g., [25] and [11].

We recall that a Lipschitz curve $\gamma = (\gamma_1, \ldots, \gamma_n) : [0,T] \to \mathbf{G}$ is said to be horizontal if $\dot\gamma(t) \in \mathcal{H}_{\gamma(t)}\mathbf{G}$, i.e., $\dot\gamma(t) = \sum_{i=1}^m a_i(t) X_i(\gamma(t))$, for almost every $t \in [0,T]$ (we denote $\dot{} = \frac{d}{dt}$). We will set

$$|\dot\gamma(t)|_{\mathcal{H}} = \sqrt{\sum_{i=1}^m |a_i(t)|^2}. \tag{16}$$

In particular, a horizontal curve is said to be sub–unit if $|\dot\gamma(t)|_{\mathcal{H}} \le 1$ for a.e. $t \in [0,T]$. For a general, non-horizontal curve γ we have

$$\begin{aligned} \dot\gamma(t) &= \sum_{i=1}^n \dot\gamma_i(t)\partial_i \\ &= \sum_{i=1}^m \dot\gamma_i(t) X_i + \sum_{l=m+1}^n \left(\dot\gamma_l(t) - \sum_{i=1}^m \dot\gamma_i(t) q_{l,i}(\gamma(t)) \right) \partial_l; \end{aligned}$$

hence a necessary and sufficient condition for γ to be horizontal is that

$$\dot\gamma_l(t) = \sum_{i=1}^m \dot\gamma_i(t) q_{l,i}(\gamma(t)) \quad \text{a.e.,} \qquad \text{for } l = m+1, \ldots, n. \tag{17}$$

Note that, given a function $\alpha : [0,T] \to \mathbb{R}^m$ in $L^1([0,T])$ and a point $x \in \mathbb{R}^n$, the equations in (17) imply that there exists a unique horizontal curve $w : [0,T] \to \mathbf{G}$ such that $w(0) = x$ and $(\dot w_1, \ldots, \dot w_m) = \alpha$ a.e. in $[0,T]$.

2.2 The value function u for the free endpoint OCP

We will now study the possibility to optimally controlling the solution w of the ordinary differential equation

$$\begin{cases} \dot w(s) = f(w(s), \alpha(s)) & \text{a.e. } s \in (t_0, t) \\ w(t_0) = x, \\ (w(t), t) \in S, \end{cases} \tag{18}$$

where the initial point (x, t_0) is a fixed point in $\mathbb{R}^n \times [0, \infty)$, and the final set S, that is usually called target set, tht is a fixed subset of $\mathbb{R}^n \times [t_0, \infty)$. We refer to the books [24, 8] as sources for this theory. The function $\alpha : [t_0, t] \to \mathbb{R}^m$ appearing in (18) is called the control strategy and the dynamics of the control problem is $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$. Clearly, we are looking for a function $w : [t_0, t] \to \mathbb{R}^n$ that starts at time t_0 from the point x and arrives at time t into the target set S.

Let A be a fixed subset of $\mathbb{R}^m$. Here and in the sequel, we denote by $\mathcal{F}_{x,t_0,\cdot,t}(A)$ the set of the admissible controls, that is measurable functions $\alpha : [t_0, t] \to A$ such that there exists a unique function $w : [t_0, t] \to \mathbb{R}^n$ that satisfies the dynamics, the initial and the final conditions in (18). Such function w is called the trajectory associated to the control α. Usually, the set A is called control set. If in (18) we replace f with $-f$, i.e. the dynamics is $\dot w = -f(w, \alpha)$, we denote by $\overline{\mathcal{F}}_{x,t_0,\cdot,t}(A)$ the set of the admissible controls.

Now let us consider the functions $\psi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ (the running cost) and $g : \mathbb{R}^m \to \mathbb{R}$ (the final cost). We are interested in the optimal control problem

$$\begin{cases} \displaystyle\inf_{\alpha \in \overline{\mathcal{F}}_{x,t_0,\cdot,t}(A)} J_{x,t}(\alpha) \\ \displaystyle J_{x,t}(\alpha) = \int_{t_0}^t \psi(w(s), \alpha(s))\, ds + g(w(t)) \end{cases} \tag{19}$$

Note that the functions ψ and f do not depend directly on the time s; we say that the problem is autonomous. In this paper we work with optimal controls where the final time t is fixed,

- when the target set is $\mathbb{R}^n \times \{t\}$ we say that (19) is a free endpoint problem, and the set of admissible controls is denoted, as before, by $\mathcal{F}_{x,t_0,\cdot,t}(A)$ and $\overline{\mathcal{F}}_{x,t_0,\cdot,t}(A)$;
- when the target set is $\{(y,t)\}$, for a given $y \in \mathbb{R}^n$, we say that (19) is a fixed endpoint problem, and the set of admissible controls is denoted by $\mathcal{F}_{x,t_0,y,t}(A)$ and $\overline{\mathcal{F}}_{x,t_0,y,t}(A)$.

For a free endpoint problem, the method of dynamic programming investigates problem (19) by studying the value function $u : \mathbb{R}^n \times [0,T] \to [-\infty, +\infty]$ defined by

$$u(x,t) = \inf_{\alpha \in \overline{\mathcal{F}}_{x,t_0,\cdot,t}(A)} J_{x,t}(\alpha).$$

Clearly, if $\overline{\mathcal{F}}_{x,t_0,\cdot,t}(A) \neq \emptyset$, then $u(x,t) < \infty$; if ψ and g are bounded, then $u > -\infty$. If there exists a control $\alpha^*_{x,t} \in \overline{\mathcal{F}}_{x,t_0,\cdot,t}(A)$ such that $u(x,t) = J_{x,t}(\alpha^*_{x,t})$, we say that $\alpha^*_{x,t}$ is optimal for the problem (19).

Now, let us consider a Carnot group $\mathbf{G}$. Equation (17) is a necessary and a sufficient for a curve $\gamma : [0,T] \to \mathbf{G}$ to be horizontal. To interpret this condition as a dynamics for an optimal control problem, let us define the function $f^{\mathbf{G}} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ by

$$f^{\mathbf{G}}(x,v) = \begin{pmatrix} v_1 \\ \dots \\ v_m \\ \displaystyle\sum_{i=1}^m v_i q_{m+1,i}(x) \\ \dots \\ \displaystyle\sum_{i=1}^m v_i q_{n,i}(x) \end{pmatrix}, \tag{20}$$

for every $x = (x_1, \dots, x_n) \in \mathbf{G}$ and $v = (v_1, \dots, v_m) \in \mathbb{R}^m$. In order to emphasize that the dynamics of the next control problem is in relation with the structure of the Carnot group $\mathbf{G}$, we add the apex "$\mathbf{G}$" to notation of the set of admissible controls. Hence let us consider the problem

$$\begin{cases} \displaystyle\inf_{\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)} J_{x,t}(\alpha) \\ J_{x,t}(\alpha) = \displaystyle\int_0^t \psi(\alpha(s))ds + g(w(t)) \\ \dot{w} = -f^{\mathbf{G}}(w,\alpha) \\ w(0) = x \end{cases} \tag{21}$$

where $\psi : \mathbb{R}^m \to \mathbb{R}$, $g : \mathbf{G} \to \mathbb{R}$, and the points t, T $(0 \le t \le T)$, and $x \in \mathbf{G}$ are fixed. Notice that, for every control in the class $\overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)$ or in the class $\mathcal{F}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)$, the associated trajectory is a horizontal curve in $\mathbf{G}$ that starts at time 0 from the point $x \in \mathbf{G}$.

From a geometric point of view, our aim is to minimize the functional $J_{x,t}$ along all the possible horizontal curves w in $\mathbf{G}$ starting from x, where the "cost"of such curves is determined via ψ by the horizontal velocity of w at every point and via g by the final point $w(t)$.

From the optimal control point of view, we remark that, with regard to problem (21), due to the lack of compactness of the control set we cannot apply standard results of optimal control theory to study the value function of the problem. However, we will show that under suitable assumptions, in order to minimize $J_{x,t}$, it is possible to restrict our attention to the set of controls whose values lie in a fixed compact set. More precisely, in this paper we will consider the following properties for the function $\psi : \mathbb{R}^m \to \mathbb{R}$:

(H1) ψ is convex;

(H2) there exist $l_0 \geq 0$ and a function $l : [0, \infty) \to [0, \infty)$ such that

$$\lim_{r\to\infty} \frac{l(r)}{r} = \infty \quad \text{and} \quad \psi(u) \geq l(|u|) - l_0, \qquad \forall u \in \mathbb{R}^m.$$

Concerning (H1) and (H2), straightforward computations show that ψ satisfies (H1) and (H2) if and only if ψ satisfies (H1) and it is superlinear. Notice that (H2) implies that there exists M_0 such that

$$l(v) > v, \qquad \forall v \geq M_0. \tag{22}$$

Our first result is the following:

Theorem 2.1 *Let $g : \mathbb{R}^n \to \mathbb{R}$ be Euclidean locally Lipschitz and bounded, and $\psi : \mathbb{R}^m \to \mathbb{R}$ be a function satisfying* (H2)*. Then, for any positive T and R, there exists $\mu = \mu(R, T) > 0$ such that for the value function $u : B_R \times [0, T] \to \mathbb{R}$ defined as*

$$u(x, t) = \inf_{\alpha\in\overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)} J_{x,t}(\alpha), \tag{23}$$

we have

$$u(x, t) = \inf_{\alpha\in\overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(B_\mu)} J_{x,t}(\alpha), \qquad \forall (x, t) \in B_R \times [0, T],$$

where $B_\mu = \{v \in \mathbb{R}^m : \ |v| \leq \mu(R, T)\}$.

The idea of the proof of this theorem can be found in [16] (see Theorem 7.4.6): we first prove that it is possible to restrict our attention only on controls in $\overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)$ such that $\|\alpha\|_1 \leq M$, where M depends on the fixed T and R. In the second step, we prove that we can further restrict our study only on bounded control in $\overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(B_\mu)$, as stated in the theorem.

Let us start with a preliminary result.

Proposition 2.1 *Let $\mathbf{G}$ be a Carnot group and let T, R and M be fixed positive constants. Then there exists $R^* = R^*(R, M)$, $R^* \geq M$, such that, for $\alpha : [0, t] \to \mathbb{R}^m$ with $0 \leq t \leq T$ and $\|\alpha\|_1 \leq M$, and for $x \in \mathbb{R}^n$ with $|x| \leq R$, we have*

$$|w(s)| \leq R^*, \qquad s \in [0, t],$$

where w is the trajectory associated to α with $w(0) = x$.

Proof: Let us prove the assertion by induction on the step of $\mathbf{G}$. Let i be such that $1 \le i \le N_j$, with $j = 1$; then

$$|w_i(s)| \le |x_i| + \int_0^s |\alpha_i(v)| dv \le R + M = R_1. \tag{24}$$

Suppose that, for every i satisfying $N_{j-1} < i \le N_j$, and $j \le J$, there exists $R_j \in \mathbb{R}$ such that the following inequality holds:

$$|w_i(s)| \le R_j, \qquad s \in [0,t].$$

Let us consider i, $N_j < i \le N_{j+1}$: taking into account that $q_{i,k}$ is a polynomial of degree at most j, that depends only on $x_k,\ y_h$ with $1 \le k,h \le N_j$, by the induction assumption we have that

$$\begin{aligned} |w_i(s)| &\le |x_i| + \sum_{k=1}^m \int_0^s |\alpha_k(v)| \cdot |q_{i,k}(w(v),0)| dv \\ &\le R + \tilde{Q}^*_{j+1} \sum_{k=1}^m \int_0^s |\alpha_k(v)| (\max\{|w_i(v)| :\ 1 \le i \le N_j\})^j dv \\ &\le R + m\tilde{Q}^*_{j+1} R_j^j M =: R_{j+1}, \end{aligned} \tag{25}$$

where $\tilde{Q}^*_j$ is a constant that depends on the polynomials $\{q_{i,k}\}_{N_j<i\le N_{j+1}, 1\le k\le m}$. This concludes the proof. □

As mentioned, let us prove that it is possible to restrict our attention only on L^1–bounded controls.

Proposition 2.2 *Under the assumptions of Theorem 2.1, there exists a positive constant $M = M(T)$ with $M > 1$ such that, for every $(x,t) \in B_R \times [0,T]$, we have*

$$\inf_{\{\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m):\ \|\alpha\|_1 \le M\}} J_{x,t}(\alpha) = \inf_{\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)} J_{x,t}(\alpha).$$

Proof: Fix (x,t). Let us denote $\lambda_{x,t} = \inf_{\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)} J_{x,t}(\alpha)$; clearly $\lambda_{x,t}$ is finite, since $t\psi(0)+g(x) = J_{x,t}(0) \ge \lambda_{x,t}$. Let $\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)$: we first prove that $\alpha \in L^1([0,t])$. Assume that this is not the case: then, for $K > 0$, denote by $I_K := \{s \in [0,t] : |\alpha(s)| > K\} \subset [0,t]$. Clearly, I_K is of positive length measure and $\int_{I_K} |\alpha(s)| ds = \infty$. Choosing $K > M_0$ we obtain from (H2) and (22), that

$$\begin{aligned} J_{x,t}(\alpha) &\ge \int_{I_K} \psi(\alpha(s)) ds + \inf g \\ &\ge \int_{I_K} [l(|\alpha(s)|) - l_0] ds + \inf g \\ &> \int_{I_K} [|\alpha(s)| - l_0] ds + \inf g \end{aligned}$$

This implies that $J_{x,t}(\alpha) = \infty$, which is a contradiction. Hence $\inf_\alpha J_{x,t}(\alpha)$ does not change if it is taken over the controls $\alpha \in L^1([0,t]) \cap \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)$.

Let $\{\alpha^k\}_{k\geq 0} \subset L^1([0,t])$ be a minimizing sequence, i.e., $\lim_{k\to\infty} J_{x,t}(\alpha^k) = \lambda_{x,t}$. For M_0 as in (22), we get that

$$\begin{aligned}\|\alpha^k\|_1 &= \int_{\{s:\ |\alpha^k(s)|\leq M_0\}} |\alpha^k(s)|ds + \int_{\{s:\ |\alpha^k(s)|>M_0\}} |\alpha^k(s)|ds \\ &\leq tM_0 + \int_0^t l(|\alpha^k(s)|)ds. \end{aligned} \tag{26}$$

Moreover, by (H2),

$$\begin{aligned}\int_0^t l(|\alpha^k(s)|)ds &\leq \int_0^t \psi(\alpha^k(s))ds + l_0 t \\ &\leq J_{x,t}(\alpha^k) + l_0 t - \inf g \end{aligned} \tag{27}$$

Let k be sufficiently large in order that $J_{x,t}(\alpha^k) - 1 \leq \lambda_{x,t} \leq J_{x,t}(\alpha^k)$. Note that, since $\lambda_{x,t} \leq t\psi(0) + g(x)$, we have $\lambda_{x,t} \leq T\psi(0) + \sup g$. Relations (26) and (27) imply that

$$\begin{aligned}\|\alpha^k\|_1 &\leq t(M_0 + l_0) + J_{x,t}(\alpha^k) - \inf g \\ &\leq T(M_0 + l_0 + \psi(0)) + \sup g + 1 - \inf g := M.\end{aligned}$$

Hence the infimum of $J_{x,t}(\alpha)$ does not change if it is taken only over the controls $\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)$ with $\|\alpha\|_1 \leq M$, where

$$M = T(M_0 + l_0 + \psi(0)) + \mathrm{var}(g) - 1. \tag{28}$$

$\square$

The previous two propositions are used in order to prove Theorem 2.1.

Proof of Theorem 2.1: Let $\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)$ be a control with $\|\alpha\|_1 \leq M$, where M is as in Proposition 2.2. Let us define, for every positive μ, the sets $I_\mu = \{s \in [0,t] :\ |\alpha(s)| > \mu\}$ and $I_\mu^C = [0,t] \setminus I_\mu$, and the function

$$\alpha^\mu(s) = \begin{cases} \alpha(s) & \text{if } s \notin I_\mu \\ 0 & \text{if } s \in I_\mu \end{cases}$$

Clearly $\|\alpha^\mu\|_1 \leq M$; moreover $\alpha^\mu \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)$ since there are no conditions on the values of the associated trajectory w^μ at the final and fixed time t. Clearly $w^\mu(0) = x$.

First of all, let us suppose that $g = 0$. Then

$$\begin{aligned}\int_0^t \psi(\alpha^\mu(s))ds - \int_0^t \psi(\alpha(s))ds &= -\int_{I_\mu} \psi(\alpha(s))ds \\ &\leq \int_{I_\mu} \left(-l(|\alpha(s)|) + l_0\right) ds.\end{aligned} \tag{29}$$

Hence, if $\mu \geq \max(M_0, l_0)$, we obtain $\int_0^t \psi(\alpha^\mu(s))ds < \int_0^t \psi(\alpha(s))ds$ and the proof, in the case $g = 0$, is finished.

In the case that the final payoff g is different from zero, we have to work some more: we will show that there exists a constant $C = C(R,T)$, such that

$$|w^\mu(s) - w(s)| \leq C \int_{I_\mu} |\alpha(v)|dv, \qquad s \in [0,t]. \tag{30}$$

Let us prove the assertion by induction on the step j. Let first, i be such that $1 \leq i \leq N_1$, then

$$\begin{aligned} |w_i^\mu(s) - w_i(s)| &\leq \int_0^t |f_i^{\mathbf{G}}(w^\mu, \alpha^\mu) - f_i^{\mathbf{G}}(w, \alpha)| dv \\ &= \int_{I_\mu} |0 - \alpha_i(v)| dv + \int_{I_\mu^C} |\alpha_i(v) - \alpha_i(v)| dv \\ &\leq \int_{I_\mu} |\alpha(v)| dv. \end{aligned}$$

Hence, with $C_1 = \sqrt{m}$, we have

$$\sqrt{\sum_{i=1}^m |w_i^\mu(s) - w_i(s)|^2} \leq C_1 \int_{I_\mu} |\alpha(v)| dv.$$

Suppose that, for j with $1 \leq j < r$ there exists $C_j \in \mathbb{R}$ with $C_j = C_j(R, T)$, such that the following inequality holds:

$$\sqrt{\sum_{i=1}^{N_j} |w_i^\mu(s) - w_i(s)|^2} \leq C_j \int_{I_\mu} |\alpha(v)| dv, \qquad s \in [0, t]. \tag{31}$$

Let us consider i, $N_j < i \leq N_{j+1}$:

$$\begin{aligned} |w_i^\mu(s) - w_i(s)| &\leq \int_0^t |f_i^{\mathbf{G}}(w^\mu, \alpha^\mu) - f_i^{\mathbf{G}}(w, \alpha)| dv \\ &\leq \sum_{k=1}^m \left\{ \int_{I_\mu} |0 - \alpha_k(v) q_{i,k}(w(v))| \, dv + \right. \\ &\qquad \left. + \int_{I_\mu^C} |\alpha_k(v)| \, |q_{i,k}(w^\mu(v)) - q_{i,k}(w(v))| \, dv \right\}. \end{aligned}$$

Taking into account that $q_{i,k}$ is a polynomial of degree at most j, that depends only on x_k, y_h with $1 \leq k, h \leq N_j$, and that $\|\alpha\|_1 \leq M$ and $\|\alpha^\mu\|_1 \leq M$, denote by $\tilde{Q}_{j+1}$ a constant such that both the following inequalities are satisfied for all i, $N_j < i \leq N_{j+1}$, and k, $1 \leq k \leq m$:

$$|q_{i,k}(w(v))| \leq \tilde{Q}_{j+1} (R^*)^j, \tag{32}$$

$$|q_{i,k}(w^\mu(v)) - q_{i,k}(w(v))| \leq \tilde{Q}_{j+1} (R^*)^{j-1} \sqrt{\sum_{p=1}^{N_j} |w_p^\mu(v) - w_p(v)|^2}, \tag{33}$$

where R^* is as in Proposition 2.1. By the induction assumption (31), the inequality $M \leq R^*$

and the previous two inequalities we obtain

$$\begin{aligned}
|w_i^\mu(s) - w_i(s)| &\leq \tilde{Q}_{j+1} \sum_{k=1}^{m} \Bigg\{ (R^*)^j \int_{I_\mu} |\alpha_k(v)|\, dv + \\
&\qquad\qquad + (R^*)^{j-1} \int_{I_\mu^C} |\alpha_k(v)| \sqrt{\sum_{p=1}^{N_j} |w_p^\mu(v) - w_p(v)|^2 dv} \Bigg\} \\
&\leq \tilde{Q}_{j+1} (R^*)^{j-1} \sum_{k=1}^{m} \Bigg\{ R^* \int_{I_\mu} |\alpha(v)|\, dv + C_j \int_0^t |\alpha(v)| \int_{I_\mu} |\alpha(u)| du dv \Bigg\} \\
&\leq \tilde{Q}_{j+1} (R^*)^j m\, (1 + C_j) \int_{I_\mu} |\alpha(v)|\, dv,
\end{aligned}$$

using $M \leq R^*$ (see the proof of Proposition 2.1). Hence it is clear that this last inequality and (31) imply

$$\sqrt{\sum_{i=1}^{N_{j+1}} |w_i^\mu(s) - w_i(s)|^2} \leq C_{j+1} \int_{I_\mu} |\alpha(v)| dv, \qquad s \in [0, t].$$

where

$$C_{j+1}^2 = n_{j+1} \tilde{Q}_{j+1}^2 (R^*)^{2j} m^2 (1 + C_j)^2 + C_j^2, \tag{34}$$

depends on R. Moreover, $C_{j+1} \geq C_j$. This proves (30) with $C = C_r$.

Let K_g be the Lipschitz constant of g in B_{R^*}. Then

$$|g(w^\mu(t)) - g(w(t))| \leq K_g C \int_{I_\mu} |\alpha(s)| ds. \tag{35}$$

To conclude the proof, we use (29) and (35) to obtain

$$\begin{aligned}
J_{x,t}(\alpha^\mu) - J_{x,t}(\alpha) &= \int_0^t \psi(\alpha^\mu(s)) ds + g(w^\mu(t)) - \int_0^t \psi(\alpha(s)) ds - g(w(t)) \\
&\leq \int_{I_\mu} (-l(|\alpha(s)|) + l_0 + K_g C |\alpha(s)|) ds.
\end{aligned}$$

Let us choose $\mu = \mu(R, T)$ in such a way that $l(v) > l_0 + K_g C v$ for all $v > \mu$. Then $J_{x,t}(\alpha^\mu) < J_{x,t}(\alpha)$ for all controls α with $\|\alpha\|_1 \leq M$. Since the bound $\|\alpha\|_1 \leq M$ is satisfied along a minimizing sequence, we have proved that the infimum of the functional remains the same if the functional is restricted over the controls with $\|\alpha\|_\infty \leq \mu$. □

Let us summarize briefly our results. Let g and ψ be as in Theorem 2.1. Denote by R and T two fixed positive constants. We have seen that the constant μ can be chosen in such a way that the function l defined in (H2) satisfies

$$l(v) > l_0 + K_g C v \tag{36}$$

for all $v > \mu$. Here K_g denotes the Lipschitz constant of the function g on the set B_{R^*}, where $R^* = R^*(R, M)$ is as in Proposition 2.1, $M = M(T)$ is as in Proposition 2.2 and $C = C(R, T)$ as follows from the proof of Theorem 2.1.

Let (x,t) be a point in $B_R \times [0,T] \subset \mathbb{R}^n \times \mathbb{R}$: in order to study problem (21), we know, by the mentioned theorem, that

$$\inf_{\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)} J_{x,t}(\alpha) = \inf_{\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(B_\mu)} J_{x,t}(\alpha).$$

Since we restrict our attention on the set of controls $\alpha : [0,t] \to B_\mu \subset \mathbb{R}^m$ and $\|\alpha\|_1 \leq T\mu = M$, Proposition 2.1 guarantees that there exists a constant $R^* = R^*(R,T,M)$ such that the trajectories associated to each of these controls, i.e. the horizontal curves w, lie within B_{R^*}. In conclusion we reduced the problem to a compact control set $B_\mu \subset \mathbb{R}^m$ and and to a compact set $B_{R^*} \subset \mathbb{R}^n$ of the possibile values of the trajectories. Hence, we can consider the functions ψ, g and f bounded. In addition (20) implies that, for every x, y in B_{R^*}, v in B_μ,

$$|f^{\mathbf{G}}(x,v) - f^{\mathbf{G}}(y,v)| = \left\| \begin{pmatrix} 0 \\ \dots \\ 0 \\ \sum_{i=1}^m v_i(q_{m+1,i}(x) - q_{m+1,i}(y)) \\ \dots \\ \sum_{i=1}^m v_i(q_{n,i}(x) - q_{n,i}(y)) \end{pmatrix} \right\| \leq \text{Const}|x-y|,$$

since the functions $q_{j,p}$ are polynomials on the compact set B_{R^*}.

These arguments give us two important consequences: the first is the existence of an optimal control for the problem (21) when ψ is convex:

Corollary 2.1 *Let* $\mathbf{G}$ *be a Carnot group. Let* T *and* R *be fixed, with* $|x| \leq R$. *Let* ψ *be a function satisfying* (H1) *and* (H2), *and* g *be locally Lipschitz and bounded. Then there exists an optimal control for problem (21).*

The proof of this corollary is an easy application of Theorem 5.2.1 in [12]. We have only to mention that $\overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)$ is nonempty since the null function is an admissible control and that, from the convexity of ψ is convex, the sets

$$F(y) = \{v \in \mathbb{R}^{n+1} : \ (v_1, \dots, v_n) = -f^{\mathbf{G}}(y,a), \ v_{n+1} \geq \psi(a), \quad a \in B_\mu\}$$

are convex.

The second consequence of Theorem 2.1 is that we are able to guarantee that the value function u is Lipschitz and admits a.e. the horizontal gradient $\mathbb{X}u$ defined in (6). For the sake of convenience, we will denote by the same symbol $\mathbb{X}u$ both the $\mathbb{R}^m$–vector in (6) and the element of V_1

$$\mathbb{X}u(x,t) = \sum_{i=1}^m (X_i u(x,t)) X_i(x),$$

where $X_i u(x,t)$ denotes the action of X_i on the function $u(x,t)$ at x given by

$$X_i u(x,t) = \lim_{\lambda \to 0} \frac{u(x \ \exp(\lambda X_i), t) - u(x,t)}{\lambda}.$$

Proposition 2.3 *Suppose that the assumptions of Theorem 2.1 are satisfied, and let T and R be fixed. Let $u : B_R \times [0,T] \to \mathbb{R}$ be the value function defined in (23). Then*

$$|\mathbb{X}u(x,t)| \le K_g \tilde{C}, \tag{37}$$

for a.e. $x \in B_R$ and every $t \in [0,T]$, where K_g is the Lipschitz constant of g on B_{R^*} and $\tilde{C} = \tilde{C}(R,T)$ is defined in (93).

Taking into account Theorem 2.1, the proof of the existence of an upper bound of $|\mathbb{X}u(x,t)|$ is an immediate consequence of well known results (see for example [22], subsection 10.3.3) that guarantee that u is Lipschitz in the Euclidean metric for $(x,t) \in B_R \times [0,T]$. Hence for a.e. $x \in B_R$ and for all $t \in [0,T]$ we have a uniform upper bound for $|Du(x,t)|$ and therefore there exists a constant K such that

$$|\mathbb{X}u(x,t)| \le K,$$

for a.e. $x \in B_R$ and every $t \in [0,T]$. The proof of the precise estimate in (37) will be deferred to the Appendix.

The previous proposition provides a really sharp estimate of the horizontal gradient of the value function; we will discuss this result for the Heisenberg group $I\!H$ and the Engel group $\mathbb{E}$ (see Example 4.2 and Example 4.3 below).

2.3 The value function $\mathcal{L}$ for the fixed endpoint OCP

In the next stage in our passage to the Hopf-Lax formula we shall consider an optimal control problem with fixed endpoints as follows. Let $x, y \in \mathbf{G}$ and s, t be fixed, with $0 \le s \le t$, and let us consider the problem

$$\begin{cases} \inf\limits_{\alpha \in \mathcal{F}^{\mathbf{G}}_{x,s,y,t}(\mathbb{R}^m)} \displaystyle\int_s^t \psi(\alpha(u))du \\ \dot{w} = f^{\mathbf{G}}(w,\alpha) \\ w(s) = x \\ w(t) = y \end{cases} \tag{38}$$

where $\psi : \mathbb{R}^m \to \mathbb{R}$. We recall that the set $\mathcal{F}^{\mathbf{G}}_{x,s,y,t}(\mathbb{R}^m)$ contains all the admissible controls, i.e., measurable functions $\alpha : [s,t] \to \mathbb{R}^m$ such that there exists a unique horizontal curve $w : [s,t] \to \mathbf{G}$ (the trajectory) satisfying the conditions $w(s) = x,\ w(t) = y$. The Hörmander condition on the vector fields implies, if $0 \le s < t$, that $\mathcal{F}^{\mathbf{G}}_{x,s,y,t}(\mathbb{R}^m) \ne \emptyset$. Note that, if $x \ne y$, then $\mathcal{F}_{x,t,y,t}(\mathbb{R}^m) = \emptyset$; hence, in all this subsection, we consider $s < t$.

We emphasize that, with respect to the previous subsection, in the dynamics we have dropped the minus sign and hence the "bar" to the set $\mathcal{F}$.

Let us define $L(x,s,y,t)$ as the value function for (38), i.e.

$$L(x,s,y,t) = \inf_{\alpha \in \mathcal{F}^{\mathbf{G}}_{x,s,y,t}(\mathbb{R}^m)} \int_s^t \psi(\alpha(u))du$$

and the value function $\mathcal{L} : \mathbf{G} \times [0,\infty) \to (-\infty, +\infty]$ as

$$\mathcal{L}(x,t) = \begin{cases} L(e,0,x,t) & \text{if } t > 0 \\ 0 & \text{if } (x,t) = (e,0) \\ +\infty & \text{if } t = 0,\ x \ne e. \end{cases} \tag{39}$$

Since ψ depends only on α, it is easy to see that

$$L(x,s,y,t) = L(e,0,x^{-1}y,t-s) = \mathcal{L}(x^{-1}y,t-s).$$

The purpose of this subsection is to provide some useful properties of the function $\mathcal{L}$.

First of all, if we consider a horizontal curve $\gamma : [0,t] \to \mathbf{G}$ and a point $x \in \mathbf{G}$, it is well known that the left translation L_x on $\mathbf{G}$ of γ, i.e. the curve $x\gamma : [0,t] \to \mathbf{G}$, is horizontal. Moreover, if $\alpha^\sharp \in \mathcal{F}^{\mathbf{G}}_{e,0,x,t}(\mathbb{R}^m)$ and $\tilde{\alpha} \in \mathcal{F}^{\mathbf{G}}_{e,0,y,s}(\mathbb{R}^m)$, then the control $\alpha : [0,t+s] \to \mathbb{R}^m$ given by

$$\alpha(u) = \begin{cases} \alpha^\sharp(u) & 0 \le u \le t \\ \tilde{\alpha}(u-t) & t < u \le t+s, \end{cases}$$

is in $\mathcal{F}^{\mathbf{G}}_{e,0,xy,t+s}(\mathbb{R}^m)$. Hence we have:

Remark 2.1

$$\mathcal{L}(x,t) + \mathcal{L}(y,s) \ge \mathcal{L}(xy,t+s), \qquad \forall t,s > 0,\ x,y \in \mathbf{G}. \tag{40}$$

We need more technical arguments to prove the following:

Proposition 2.4 *Let $\mathcal{L}$ be as in (39). Then*

$$\mathcal{L}(x,ts) = t\mathcal{L}(\delta_{\frac{1}{t}}(x),s), \qquad \forall t > 0,\ s > 0,\ x \in \mathbf{G} \tag{41}$$

Proof: Let α be in $\mathcal{F}^{\mathbf{G}}_{e,0,x,ts}(\mathbb{R}^m)$ and w the associated trajectory. Let us consider $\alpha^\sharp : [0,s] \to \mathbb{R}^m$ defined by $\alpha^\sharp(u) = \alpha(tu)$, for $u \in [0,s]$. Denote by $w^\sharp$ the associated trajectory. Let us prove that

$$w^\sharp(u) = \delta_{\frac{1}{t}}(w(tu)), \qquad u \in [0,s]. \tag{42}$$

For every i with $1 \le i \le N_1 = m$, and $u \in [0,s]$, we have that

$$w_i^\sharp(u) = \int_0^u \alpha_i(tv)dv = \frac{1}{t}\int_0^{tu} \alpha_i(v)dv = \frac{1}{t}w_i(tu).$$

We shall proceed by induction. Now let j and i be such that $2 \le j \le r$, and $N_{j-1} < i \le N_j$. Since the polynomial $q_{i,k}(y)$ depends only on the first N_{j-1} components of $y = (y_1,\dots,y_n)$, we have

$$w_i^\sharp(u) = \sum_{k=1}^m \int_0^u \alpha_k^\sharp(v) q_{i,k}(w^\sharp(v))dv = \sum_{k=1}^m \int_0^u \alpha_k(tv) q_{i,k}\left(\delta_{1/t}(w(tv))\right)dv.$$

Since $\sigma_i = j$ and $\sigma_k = 1$, using (15), we obtain

$$w_i^\sharp(u) = \frac{1}{t}\sum_{k=1}^m \int_0^{tu} \alpha_k(v) q_{i,k}\left(\delta_{1/t}(w(v))\right)dv = \frac{1}{t^j}\sum_{k=1}^m \int_0^{tu} \alpha_k(v) q_{i,k}(w(v))dv = \frac{1}{t^j}w_i(tu).$$

Hence we get (42) and, in particular, $w^\sharp(s) = \delta_{\frac{1}{t}}(x)$, i.e. $\alpha^\sharp \in \mathcal{F}^{\mathbf{G}}_{e,0,\delta_{1/t}(x),s}(\mathbb{R}^m)$. Finally, the equality

$$\int_0^{ts} \psi(\alpha(u))du = t\int_0^s \psi(\alpha^\sharp(u))du$$

implies the inequality:

$$t\mathcal{L}(\delta_{\frac{1}{t}}(x),s) \le \mathcal{L}(x,ts).$$

The opposite inequality follows in the same manner. □

The previous Remark 2.1 and Proposition 2.4 give an interesting property of the function $\mathcal{L}$. Let us recall the notion of convexity in a Carnot group:

Definition 2.1 *(see [18], Definition 3.2). Let* $\mathbf{G}$ *be a Carnot group. A function* $f : \mathbf{G} \to \mathbb{R}$ *is said to be convex if for every* $x, y \in \mathbf{G}$ *and for every* $\lambda \in [0,1]$*, the following inequality is satisfied*

$$f(x\delta_\lambda(x^{-1}y)) \leq f(x) + \lambda(f(y) - f(x)). \tag{43}$$

We note that if we consider a function defined on a subset of $\mathbb{R}^m$, it is clear that the notion of convexity is the classical one.

If we only require that (43) holds for y that belong to $L_x(\exp V_1)$ in Definition 2.1, we obtain a more general notion of convexity: called weak H–convexity (see Definition 5.5 in [18]). In this framework, Balogh and Rickly proved that a weakly H–convex function, measurable if $r > 2$, is regular enough, since it is locally Lipschitz continuous with respect to any homogeneous distance (see [7], [30]). Moreover, it is known that a Rademacher–Stepanov type result holds in the Carnot group setting; therefore, it turns out that such weakly H–convex functions are differentiable almost everywhere with respect to the horizontal directions (see for example [19]).

In [18] the authors relate the property of weak H–convexity of a real–valued function to the nonemptyness of its H–subdifferential. Let us recall that the H–subdifferential of a function $f : \Omega \subset \mathbf{G} \to (-\infty, +\infty]$ at $x \in \Omega$ is defined as

$$\partial_H f(x) = \{p \in V_1 : \ f(x\exp(q)) \geq f(x) + p \cdot q, \ \forall q \in V_1 : \ x\exp(q) \in \Omega\}.$$

It is easy to show (see [18], Proposition 10.5) that if $\partial_H f(x) \neq \emptyset$ for every $x \in \Omega$, then f is weakly H–convex. The converse of this result, as in the classical case, is more difficult. Calogero and Pini prove that this holds when $\mathbf{G}$ is the Heisenberg group $I\!H$ (see [14]); in [15] they improved this result to a generic Carnot gruop with an additional assumption of measurability of f if $r > 2$.

It turns out that our function $\mathcal{L}$ is group convex in the sense of the above definition. Indeed, for every $x, y \in \mathbf{G}, \ t > 0$, and $\lambda \in [0,1]$ we have, from (40) and (41),

$$(1-\lambda)\mathcal{L}(x,t) + \lambda\mathcal{L}(y,t) = \mathcal{L}(\delta_{1-\lambda}(x), (1-\lambda)t) + \mathcal{L}(\delta_\lambda(y), \lambda t) \geq \mathcal{L}(\delta_{1-\lambda}(x)\delta_\lambda(y), t).$$

Since $\delta_{1-\lambda}(x)\delta_\lambda(y) = x\delta_\lambda(x^{-1}y)$, we obtain that $x \mapsto \mathcal{L}(x,t)$ is convex, and hence continuous, measurable and weakly H–convex. More precisely we have the following result:

Theorem 2.2 *Let* $t > 0$ *be fixed. The function* $\mathcal{L}(\cdot, t)$ *is convex in* $\mathbf{G}$ *and is locally Lipschitz continuous. The horizontal gradient* $\mathbb{X}\mathcal{L}(x,t)$ *exists for a.e.* $x \in \mathbf{G}$.

The function $\mathcal{L}$ has an additional symmetry property that we are going to use in the proof of our Hopf–Lax formula:

Proposition 2.5 *For every* $t > 0$ *and* $x, y \in \mathbf{G}$ *we have*

$$\inf_{\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,y,t}(\mathbb{R}^m)} \int_0^t \psi(\alpha(u))du = \inf_{\alpha \in \mathcal{F}^{\mathbf{G}}_{y,0,x,t}(\mathbb{R}^m)} \int_0^t \psi(\alpha(u))du \tag{44}$$

i.e.

$$\mathcal{L}(y^{-1}x, t) = \inf_{\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,y,t}(\mathbb{R}^m)} \int_0^t \psi(\alpha(u))du.$$

Proof: For every admissible control $\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,y,t}(\mathbb{R}^m)$ with associated trajectory w, let us define the functions $\tilde{\alpha}(s) = \alpha(t-s)$, $\tilde{w}(s) = w(t-s)$ for every $s \in [0,t]$. It is easy to see that $\tilde{\alpha} \in \mathcal{F}^{\mathbf{G}}_{y,0,x,t}(\mathbb{R}^m)$ with associated trajectory $\tilde{w}$. Clearly a reverse argument holds: hence the two classes of functions are one–to–one. Moreover

$$\int_0^t \psi(\alpha(v))dv = \int_0^t \psi(\tilde{\alpha}(s))dv$$

and (44) holds true. □

3 The Hamilton–Jacobi–Bellman equation in Carnot groups

This section is devoted to the proof of our main result. We shall use the notations introduced in the previous section. We shall need the following classical result (see [22], [8], [2]):

Theorem 3.1 *Let ψ, f and g be bounded, Lipschitz functions w.r.t. w, uniformly w.r.t. $a \in A$. Let $T > 0$ and the target set $S = \mathbb{R}^n \times \{t\}$, with $0 < t < T$, be fixed. Let the control set $A \subset \mathbb{R}^m$ be compact. Then the value function*

$$u(x,t) = \inf_{\alpha \in \overline{\mathcal{F}}_{x,0,\cdot,t}(A)} \int_0^t \psi(w(s), \alpha(s))ds + g(w(t))$$

is the unique viscosity solution of the problem

$$\begin{cases} u_t(x,t) + H(x, Du(x,t)) = 0 & (x,t) \in \mathbb{R}^n \times (0,T) \\ u(x,0) = g(x) & x \in \mathbb{R}^n \end{cases} \tag{45}$$

where $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is

$$H(x,v) = \max_{a \in A} \left(v \cdot f(x,a) - \psi(x,a) \right). \tag{46}$$

3.1 The value function u as viscosity solution

This subsection is devoted to prove our first main result (Theorem 3.2).

Let $\mathbf{G}$ be a Carnot group, and T and R be fixed. By Theorem 2.1 we have that the value function $u : B_R \times [0,T] \to \mathbb{R}$ defined in (23) satisfies

$$u(x,t) = \inf_{\alpha \in \overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(B_\mu)} J_{x,t}(\alpha),$$

for all $(x,t) \in B_R \times [0,T]$. Hence, since we are able to restrict our attention to the controls $\overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(B_\mu)$, Theorem 3.1 implies that our value function u is the unique viscosity solution of the problem (45) where the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, by (46), is defined as

$$H(x,v) = \max_{a \in B_\mu} \left(v \cdot f^{\mathbf{G}}(x,a) - \psi(a) \right). \tag{47}$$

Again by Theorem 2.1 and Proposition 2.3, we know that our value function u admits the gradient $Du(x,t)$ and hence the horizontal gradient $\mathbb{X}u(x,t)$, for all $t \in [0,T]$ and a.e.

$x \in B_R$. Let $v \in \mathbb{R}^n$ be such that $v = Du(x,t)$ with $(x,t) \in B_R \times [0,T]$. From the definition of $f^{\mathbf{G}}$ in (20), we have

$$\begin{aligned}
\sup_{a\in\mathbb{R}^m} \left(v \cdot f^{\mathbf{G}}(x,a) - \psi(a)\right) &= \sup_{a\in\mathbb{R}^m} \left[\sum_{j=1}^m a_j \left(v_j + \sum_{i=m+1}^n v_i q_{i,j}(x)\right) - \psi(a)\right] \\
&= \sup_{a\in\mathbb{R}^m} \left[\sum_{j=1}^m a_j X_j u(x,t) - \psi(a)\right] \\
&= \psi^* \left(\mathbb{X}u(x,t)\right), \qquad (48)
\end{aligned}$$

where ψ^* denotes the Legendre–Fenchel transform of ψ. Let us recall that for $f : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$, its Legendre–Fenchel transform $f^* : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$, is given by

$$f^*(p) = \sup_{x\in\mathbb{R}^m} (p \cdot x - f(x)).$$

We just remind to the reader some properties that will be useful in the sequel. The function f^* is convex; moreover, if f is superlinear and continuous, then f^* is real valued and superlinear. Finally, if f is a lower semicontinuous and convex function, then $f^{**} = f$ (see e.g. [31]).

In general, if in (47) we take the supremum over a subset of $\mathbb{R}^m$, we do not get the Legendre–Fenchel transform of ψ. However, since we are able to show that $\mathbb{X}u$ is confined within a bounded set, if $(x,t) \in B_R \times [0,T]$, then the supremum in the previous equality is always realized in a fixed compact set.

At this point we are able to state our first main result

Theorem 3.2 *Let* $\mathbf{G}$ *be a Carnot group. Let* g *be Euclidean locally Lipschitz and bounded, and let* ψ *satisfy (H2). Then the value function* u *of the problem*

$$\begin{cases} \inf\limits_{\alpha\in\overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(\mathbb{R}^m)} \displaystyle\int_0^t \psi(\alpha(s))ds + g(w(t)) \\ \dot{w}(s) = -f^{\mathbf{G}}(w(s), \alpha(s)) \\ w(0) = x \end{cases} \qquad (49)$$

is the unique viscosity solution of

$$\begin{cases} u_t(x,t) + \psi^*(\mathbb{X}u(x,t)) = 0 & (x,t) \in \mathbf{G} \times (0,T) \\ u(x,0) = g(x) & x \in \mathbf{G} \end{cases} \qquad (50)$$

We remark that the function ψ^* that appears in (50) is a Legendre–Fenchel transform that essentially works on the first layer V_1 of the Lie algebra $\mathfrak{g}$ of the Carnot group $\mathbf{G}$, i.e. ψ works on $\mathbb{R}^m$. A different notion of Fenchel transform on the Heisenberg group $I\!H$ can be found in [13]; via this transform it is possible to characterize the weakly H–convex functions on the $I\!H$ by extending a result well-known in the Euclidean context. However, both these notions of Fenchel transform take into account the subriemannian structure of $\mathbf{G}$, but at the moment it is not clear to us their connection.

Proof of Theorem 3.2: Let us fix T and R. Theorem 2.1 implies that for every $\tilde{\mu} \geq \mu(T,R)$ we have

$$u(x,t) = \inf_{\alpha\in\overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(B_\mu)} J_{x,t}(\alpha) = \inf_{\alpha\in\overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(B_{\tilde{\mu}})} J_{x,t}(\alpha),$$

for all $(x,t) \in B_R \times [0,T]$. Hence, by Theorem 3.1, u is the unique viscosity solution for the problem (45) where the Hamiltonian $\tilde{H} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, by (46), is defined as

$$\tilde{H}(x,v) = \max_{a \in B_{\tilde{\mu}}} \left(v \cdot f^{\mathbf{G}}(x,a) - \psi(a) \right). \tag{51}$$

Proposition 2.3 implies that there exists a constant K such that $|\mathbb{X}u(x,t)| \leq K$ for $(x,t) \in B_R \times [0,T]$.

Now, let us consider $\tilde{\mu}$ such that $\tilde{\mu} \geq \mu$ and

$$\max_{a \in B_{\tilde{\mu}}} (v \cdot a - \psi(a)) = \sup_{a \in \mathbb{R}^m} (v \cdot a - \psi(a)), \qquad \forall |v| \leq K. \tag{52}$$

If $(x,t) \in B_R \times [0,T]$ is such that $\mathbb{X}u(x,t)$ exists and if $v = Du(x,t)$, then we obtain by (51) and (52)

$$\begin{aligned}
\tilde{H}(x,v) &= \max_{a \in B_{\tilde{\mu}}} \left(Du(x,t) \cdot f^{\mathbf{G}}(x,a) - \psi(a) \right) \\
&= \max_{a \in B_{\tilde{\mu}}} (\mathbb{X}u(x,t) \cdot a - \psi(a)) \\
&= \sup_{a \in \mathbb{R}^m} (\mathbb{X}u(x,t) \cdot a - \psi(a)) \\
&= \psi^*(\mathbb{X}u(x,t)).
\end{aligned}$$

This concludes the proof. □

3.2 The Legendre–Fenchel transform $\Psi^{\mathbf{G}}$ and the Hopf–Lax formula

Let us recall that the Hopf–Lax formula provides the viscosity solution of an initial value problem for the state–independent Hamilton–Jacobi equation, reducing the computation of u to a finite dimensional optimization problem:

Theorem 3.3 (see Theorem 3 p. 601 in [22]) *Let $\Psi : \mathbb{R}^n \to \mathbb{R}$ be convex and superlinear. Let g be Lipschitz and bounded. Let $T > 0$ be fixed. Then the unique viscosity solution for*

$$\begin{cases} u_t(x,t) + \Psi(Du(x,t)) = 0 & (x,t) \in \mathbb{R}^n \times (0,T] \\ u(x,0) = g(x) & x \in \mathbb{R}^n \end{cases} \tag{53}$$

is given by the Hopf–Lax formula

$$u(x,t) = \min_{y \in \mathbb{R}^n} \left\{ t\Psi^* \left(\frac{x-y}{t} \right) + g(y) \right\}, \tag{54}$$

for every $(x,t) \in \mathbb{R}^n \times (0,\infty)$

Bardi and Evans in [9] gave a proof of the above theorem using an optimal control method, by considering the value function

$$u(x,t) = \inf \left\{ \int_0^t \Psi^*(\alpha(s))ds + g(w(t)) : \ w(s) = x - \int_0^s \alpha(v)dv \right\};$$

we note that in this optimal control problem the dynamics is $\dot{w} = -\alpha$. Our approach follows this idea.

We shall start with the following:

Definition 3.1 *Let* $\mathbf{G}$ *be a Carnot group such that* m *is the dimension of the first layer* V_1 *of its Lie algebra* $\mathfrak{g}$. *Let* $\Psi : \mathbb{R}^m \to \mathbb{R}$. *We define the* $\mathbf{G}$*–Legendre–Fenchel transform* $\Psi^{\mathbf{G}} : \mathbf{G} \to [-\infty, \infty]$ *by*

$$\Psi^{\mathbf{G}}(x) = \inf_{\alpha \in \mathcal{F}_x^{\mathbf{G}}(\mathbb{R}^m)} \int_0^1 \Psi^*(\alpha(s))ds, \tag{55}$$

where $\mathcal{F}_x^{\mathbf{G}}(\mathbb{R}^m) = \mathcal{F}_{e,0,x,1}^{\mathbf{G}}(\mathbb{R}^m)$.

As in (10), Ψ^* is the usual Legendre–Fenchel transform of Ψ. As in the previous subsections, the set $\mathcal{F}_{e,0,x,1}^{\mathbf{G}}(\mathbb{R}^m)$, denoted by $\mathcal{F}_x^{\mathbf{G}}(\mathbb{R}^m)$ in the sequel, contains all the measurable functions $\alpha : [0,1] \to \mathbb{R}^m$ such that there exists a unique continuous function $w : [0,1] \to \mathbf{G}$ solution of the dynamics $\dot{w} = f^{\mathbf{G}}(w, \alpha)$, that represents the condition for the horizontality of w in the group $\mathbf{G}$, satisfying the conditions $w(0) = e,\ w(1) = x$. In general, our results will be stated under the assumption that Ψ is convex and superlinear; this implies that Ψ^* is real–valued. Moreover, the Hörmander condition on the vector fields implies that $\mathcal{F}_x^{\mathbf{G}}(\mathbb{R}^m) \neq \emptyset$. These two considerations guarantee that $\Psi^{\mathbf{G}}$ is real–valued. Clearly, if in (38) we put $\psi = \Psi^*$, we have

$$\Psi^{\mathbf{G}}(x) = \mathcal{L}(x, 1) \tag{56}$$

and the properties for the function $\mathcal{L}$ that we obtain in subsection 2.3 are inherited by the function $\Psi^{\mathbf{G}}$. In particular

Proposition 3.1 *The function* $\Psi^{\mathbf{G}}$ *is convex in* $\mathbf{G}$ *and the horizontal gradient* $\mathbb{X}\Psi^{\mathbf{G}}$ *exists a.e.*

Our next statement was announced in the introduction as Theorem 1.1:

Theorem 3.4 *Let* $\mathbf{G}$ *be a Carnot group. Let* g *be Euclidean locally Lipschitz and bounded. Let* Ψ *satisfy* (H1) *and* (H2). *Then the unique viscosity solution of the problem*

$$\begin{cases} u_t(x,t) + \Psi(\mathbb{X}u(x,t)) = 0 & (x,t) \in \mathbf{G} \times (0,T) \\ u(x,0) = g(x) & x \in \mathbf{G} \end{cases} \tag{57}$$

is given, when $t > 0$ *and* $x \in \mathbf{G}$, *by the Hopf–Lax formula*

$$u(x,t) = \inf_{y \in \mathbf{G}} \left\{ t\Psi^{\mathbf{G}}\left(\delta_{\frac{1}{t}}(y^{-1}x)\right) + g(y) \right\}. \tag{58}$$

We will see in Proposition 4.4 that, in some particular Carnot groups (one of these is the Heisenberg group), the function $\Psi^{\mathbf{G}}$ goes to $+\infty$ when $|x| \to \infty$: in these particular Carnot groups, the infimum in (58) is indeed a minimum.

Proof of Theorem 3.4: Set $\psi = \Psi^*$ and $f = f^{\mathbf{G}}$, in particular in the definition of the function $\mathcal{L}$. We know, by Theorem 3.2, that

$$u(x,t) = \inf_{\alpha \in \overline{\mathcal{F}}_{x,0,\cdot,t}^{\mathbf{G}}(\mathbb{R}^m)} \int_0^t \Psi^*(\alpha(s))ds + g(w(t))$$

is the unique viscosity solution for (57). Hence, taking into account Proposition 2.4 and Proposition 2.5,

$$\begin{aligned} u(x,t) &= \inf_{y \in \mathbf{G}} \left\{ \inf_{\alpha \in \overline{\mathcal{F}}_{x,0,y,t}^{\mathbf{G}}(\mathbb{R}^m)} \left\{ \int_0^t \Psi^*(\alpha(s))ds \right\} + g(y) \right\} \\ &= \inf_{y \in \mathbf{G}} \left\{ \mathcal{L}(y^{-1}x, t) + g(y) \right\} \\ &= \inf_{y \in \mathbf{G}} \left\{ t\mathcal{L}\left(\delta_{\frac{1}{t}}(y^{-1}x), 1\right) + g(y) \right\} \end{aligned}$$

The statement follows from (56). $\square$

A calculation similar to the previous proof, with the same assumptions, gives that for every $x \in \mathbf{G},\ s \geq 0,\ h > 0$,

$$u(x, s+h) = \inf_{y \in \mathbf{G}} \left\{ h \Psi^{\mathbf{G}} \left(\delta_{\frac{1}{h}} \left(y^{-1} x \right) \right) + u(y, s) \right\}.$$

4 Applications and examples

In this section, we will show that Theorem 3.4 implies the classical Hopf–Lax formula, Theorem 3.1, the Hopf-Lax formulae by Manfredi e Stroffolini [29], and by Dragoni [20].

4.1 The Hopf–Lax formula in Euclidean spaces

Let $\mathbf{G} = \mathbb{R}^n$ and consider a function $\Psi : \mathbb{R}^n \to \mathbb{R}$ convex and superlinear. For simplicity, we assume that Ψ^* is C^1. Clearly $\mathbb{R}^n$ is a Carnot group of step 1 ($n = m$), where the horizontal vector fields are $X_i = \partial_i$, for $1 \leq i \leq n$. In this situation all the Lipschitz curves are admissible trajectories, since the dynamics $\dot{w} = \alpha$ is given by the function $f^{\mathbb{R}^n} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ defined by $f^{\mathbb{R}^n}(x, a) = a$ for every x and a in $\mathbb{R}^n$.

In this context, we are interested in problem (57):

$$\begin{cases} u_t(x,t) + \Psi(Du(x,t)) = 0 & (x,t) \in \mathbb{R}^n \times (0,T), \\ u(x,0) = g(x) & x \in \mathbb{R}^n. \end{cases}$$

Clearly the function $\Psi^{\mathbb{R}^n}$ in (55) is defined, for every $x \in \mathbb{R}^n$, by

$$\Psi^{\mathbb{R}^n}(x) = \inf \left\{ \int_0^1 \Psi^*(\dot{w}(s)) ds : \ w(0) = e,\ w(1) = x \right\},$$

and it coincides with $\Psi^*(x)$, from the following Proposition 4.1 Therefore, applying (58), we obtain

$$u(x,t) = \inf_{y \in \mathbb{R}^n} \left\{ t \Psi^* \left(\frac{x-y}{t} \right) + g(y) \right\}$$

that coincides with the classical Hopf–Lax formula (54).

4.2 The Hopf–Lax formula for homogeneous Hamiltonians

The Heisenberg group $\mathbb{H}$ is the Lie group whose Lie algebra $\mathfrak{h}$ admits a stratification of step 2; in particular $\mathfrak{h} = \mathbb{R}^3 = V_1 \oplus V_2$, with

$$\begin{array}{lll} V_1 = \operatorname{span}\{X_1, X_2\} & \text{with } X_1 = \partial_{x_1} - \frac{x_2}{2}\partial_{x_3} & \text{and } X_2 = \partial_{x_2} + \frac{x_1}{2}\partial_{x_3}, \\ V_2 = \operatorname{span}\{X_3\} & \text{with } X_3 = \partial_{x_3}. & \end{array} \tag{59}$$

The bracket $[\cdot,\cdot] : \mathfrak{h} \times \mathfrak{h} \to \mathfrak{h}$ is defined as $[X_1, X_2] = X_3$, and it vanishes for all the other basis vectors. Hence $\mathbb{X}u = (u_{x_1} - \frac{x_2}{2}u_{x_3}, u_{x_2} + \frac{x_1}{2}u_{x_3})$. Taking into account the action of the bracket, $X * Y$ is defined by the Baker–Campbell–Dynkin–Hausdorff formula

$$X * Y = X + Y + [X,Y]/2. \tag{60}$$

The group law is defined by the relation $\exp(X)\exp(Y) = \exp(X * Y)$, for every X and Y in $\mathfrak{g}$; consequently, the law group on $\mathbb{H}$ is given by

$$(x_1, x_2, x_3)(x'_1, x'_2, x'_3) = (x_1 + x'_1, x_2 + x'_2, x_3 + x'_3 + (x_1 x'_2 - x'_1 x_2)/2).$$

In this situation one has clearly that $Q_3(x,y) = (x_1y_2 - x_2y_1)/2$; hence (14) implies $q_{3,1}(x) = -x_2/2$ and $q_{3,2}(x) = x_1/2$. The dilation is a family of automorphisms given by $\delta_\lambda(x_1,x_2,x_3) = (\lambda x_1, \lambda x_2, \lambda^2 x_3)$, and hence the homogeneous dimension is 4. Relation (17) shows that $\gamma = (\gamma_1,\gamma_2,\gamma_3) : \mathbb{R} \to I\!\!H$ is a horizontal curve if and only if

$$\dot\gamma_3 = (\gamma_1\dot\gamma_2 - \gamma_2\dot\gamma_1)/2.$$

In the Heisenberg setting the dynamics (20) is $\dot w = -f^{I\!\!H}(w,v)$ where $f^{I\!\!H} : \mathbb{R}^3 \times \mathbb{R}^2 \to \mathbb{R}^3$ is given by

$$f^{I\!\!H}(x,v) = \begin{pmatrix} v_1 \\ v_2 \\ (v_2x_1 - v_1x_2)/2 \end{pmatrix}, \qquad \forall v \in \mathbb{R}^2,\ x \in \mathbb{R}^3. \tag{61}$$

Manfredi and Stroffolini proved the following result

Theorem 4.1 *(see Theorem 3 in [29]) Let $\Phi : [0,\infty) \to [0,\infty)$ be an increasing, strictly convex function such that*

$$\lim_{s\to\infty} \frac{\Phi(s)}{s} = \infty, \qquad \lim_{s\to 0} \frac{\Phi(s)}{s} = 0.$$

Let g be a bounded and Lipschitz continuous function. Then the unique viscosity solution for the problem in the Heisenberg group

$$\begin{cases} u_t(x,t) + \Phi(|\mathbb{X}u(x,t)|) = 0 & (x,t) \in I\!\!H \times (0,T) \\ u(x,0) = g(x) & x \in I\!\!H \end{cases} \tag{62}$$

is given by

$$u(x,t) = \min_{y\in I\!\!H} \left\{ t\Phi^*\left(\frac{d_{CC}(x,y)}{t}\right) + g(y) \right\} \tag{63}$$

Let us recall the definition of the Carnot–Carathéodory distance d_{CC} from x to y in a general Carnot group $\mathbf{G}$. Using our previous notations we define

$$d_{CC}(x,y) = \inf_{\alpha\in\mathcal{F}^{\mathbf{G}}_{x,0,y,t}(\mathbb{R}^m)} \int_0^t |\alpha(s)|ds. \tag{64}$$

In the above expression we are minimizing the length-functional $\ell(\gamma) = \displaystyle\int |\dot\gamma(s)|_{\mathcal{H}}\, ds$ (see (16)) over all the horizontal curves γ with endpoints x and y (see, for instance, [10]). An equivalent definition of the distance $d_{CC}(x,y)$ can be given using the sub–unit curve. Notice that for any horizontal curve $\gamma : [0,t] \to \mathbf{G}$ such that $\gamma(0) = x$, $\gamma(t) = y$, and for any $t' > 0$, the curve $\tilde\gamma : [0,t'] \to \mathbf{G}$ defined by $\tilde\gamma(s) = \gamma(ts/t')$ is still horizontal, with endpoints x and y and $\ell(\tilde\gamma) = \ell(\gamma)$; moreover,

$$\dot{\tilde\gamma}(s) = \frac{t}{t'}\dot\gamma\left(\frac{t}{t'}s\right), \tag{65}$$

for all $s \in [0,t']$. In particular, any horizontal curve can be reparametrized in order to be defined over any fixed interval, preserving both length and endpoints. Furthermore, to any horizontal curve $\gamma : [0,t] \to \mathbf{G}$ from x to y one can associate a horizontal curve $\tilde\gamma : [0,\ell(\gamma)] \to \mathbf{G}$ by letting $\tilde\gamma(s) = \gamma(v)$ if $\tau(v) = \displaystyle\int_0^v |\dot\gamma(u)|_{\mathcal{H}}\, du = s$. This curve satisfies the condition $|\dot{\hat{\gamma}}(s)|_{\mathcal{H}} = 1$ for a.e. $s \in [0, \ell(\gamma)]$, thereby it is parametrized by arc length (for details, see [10], p.346).

Chow's theorem [10] guarantees that any two points in $\mathbf{G}$ can be joined by a horizontal curve: the curves that realize the infimum in (64) will be called geodesics. Suppose, in particular, that γ is a geodesic from x to y parametrized by arc length, i.e., $|\gamma|_{\mathcal{H}} = 1$ a.e. and $\ell(\gamma) = d_{CC}(x, y)$. If we consider the geodesic $\hat{\gamma} : [0, t] \to \mathbf{G}$, $\hat{\gamma}(s) = \gamma(d_{CC}(x, y)s/t)$, from (65) we have that

$$|\dot{\hat{\gamma}}(s)|_{\mathcal{H}} = \left| \frac{d_{CC}(x, y)}{t} \dot{\gamma}(d_{CC}(x, y)s/t) \right|_{\mathcal{H}} = \frac{d_{CC}(x, y)}{t}. \tag{66}$$

Now we are in the position to show that Theorem 4.1 is a straightforward consequence of Theorem 3.2. Noticing that $(\Phi(|\cdot|))^*(s) = \Phi^*(|s|)$, for every $s \geq 0$, the $I\!H$–Legendre–Fenchel transform $\Psi^{I\!H}$ of the function $\Psi(\cdot) = \Phi(|\cdot|)$ is defined, for every $x \in I\!H$, by

$$\Psi^{I\!H}(x) = \inf_{\alpha \in \mathcal{F}_x^{I\!H}(\mathbb{R}^2)} \int_0^1 \Phi^*(|\alpha(s)|) ds, \tag{67}$$

where,

$$\mathcal{F}_x^{I\!H}(\mathbb{R}^2) = \{\alpha : \ \dot{w} = f^{I\!H}(w, \alpha), \ w(0) = e, \ w(1) = x\}. \tag{68}$$

Let $w^* : [0, 1] \to I\!H$ be a geodesic such that $w(0) = e, \ w(1) = x$, and let α^* be its control. Using (66) with $t = 1$ and (67) we have

$$\Psi^{I\!H}(x) \leq \int_0^1 \Phi^*(|\alpha^*(s)|) ds = \Phi^*(d_{CC}(e, x)).$$

Since Φ^* is a convex and increasing function, the Jensen inequality implies that, for every admissible control α,

$$\Phi^*(d_{CC}(e, x)) = \Phi^* \left(| \int_0^1 \alpha(s) ds | \right) \leq \Phi^* \left(\int_0^1 |\alpha(s)| ds \right) \leq \int_0^1 \Phi^*(|\alpha(s)|) ds,$$

and thus

$$\Psi^{I\!H}(x) = \Phi^*(d_{CC}(e, x)).$$

Consequently we obtain the Hopf–Lax formula (63).

It is worthwhile mentioning also the paper by Dragoni [20], where it is proved that, for Carnot–Carathéodory metrics d_{CC} which satisfy the Hörmander condition, the Hopf–Lax function is a viscosity solution of the Cauchy problem for a Hamilton–Jacobi equation for a state–dependent Hamiltonian. In particular, in Theorem 4 [20], she considers the following problem

$$\begin{cases} u_t(x, t) + \Phi(|\sigma(x) Du(x, t)|) = 0 & (x, t) \in \mathbf{G} \times (0, T) \\ u(x, 0) = g(x) & x \in \mathbf{G}. \end{cases} \tag{69}$$

Under the assumptions that $\Phi : [0, +\infty) \to [0, +\infty)$ is differentiable, convex, not decreasing with $\Phi(0) = 0$, and $\sigma(x)$ is an $m \times n$ matrix with C^∞ coefficients satisfying the Hörmander condition, she proves that, for a lower semicontinuous function g suitably bounded from below, the Hopf–Lax function

$$u(x, t) = \inf_{y \in \mathbb{R}^n} \left(t\Phi^* \left(\frac{d_{CC}(x, y)}{t} \right) + g(y) \right)$$

is a viscosity solution for problem (69). The model example is $H(x,Du) = \frac{1}{\alpha}|\sigma(x)Du|^{\alpha}$, where $\alpha > 1$. In a Carnot group, if g is bounded and lower semicontinuous, u_t and $\mathbb{X}u(x,t) = \sigma(x)Du(x,t)$ exist for almost every $t > 0$ and $x \in \mathbb{R}^n$, and the Hamiltonian has the peculiar form $H(x,Du) = \frac{1}{\alpha}|\mathbb{X}u|^{\alpha}$. Following the same line of the proof of Theorem 4.1, we can show that this result in the Carnot setting is again a consequence of Theorem 3.2.

4.3 Comparison between the functions $\Psi^{\mathbf{G}}$ and Ψ^*

We have seen that in the Euclidean space $\mathbb{R}^m$ a fundamental property of the function $\Psi^{\mathbb{R}^m}$ is that it coincides with Ψ^*; now, if we consider a generic Carnot group $\mathbf{G}$, the function $\Psi^{\mathbf{G}}$ inherits this property on $\exp(V_1)$, i.e., by an abuse of notation,

$$\Psi^{\mathbf{G}}\big|_{\exp(V_1)} = \Psi^*.$$

The following result contains this idea and it will be useful in order to study the function $\Psi^{\mathbf{G}}$:

Proposition 4.1 *Let Ψ be as in* (H1) *and* (H2)*, and such that Ψ^* is C^1. Let us consider the projection $\pi : \mathbf{G} \to \mathbb{R}^m$ defined by $\pi(x_1,\dots,x_n) = (x_1,\dots,x_m)$. Then*

$$\Psi^{\mathbf{G}}(x) = \Psi^*(\pi(x)), \qquad \forall x \in \exp(V_1) \subset \mathbf{G} \tag{70}$$

$$\Psi^{\mathbf{G}}(x) \geq \Psi^*(\pi(x)), \qquad \forall x \in \mathbf{G} \tag{71}$$

Proof: Let $x = (x_1,\dots,x_m,0,\dots,0) \in \exp(V_1)$, and consider the curve $w^\sharp : [0,1] \to \mathbf{G}$ defined by $w^\sharp(s) = (x_1 s,\dots,x_m s,0,\dots,0)$; clearly, $\alpha^\sharp = (x_1,\dots,x_m)$. Taking into account the Euler's necessary condition of the Calculus of Variations (see e.g. Theorem 1, p. 116 in [22]) and since Ψ^* is convex, one argues that the first infimum below is realized by the function $\alpha^\sharp$; therefore we have that

$$\begin{aligned}
\Psi^*(\pi(x)) &= \int_0^1 \Psi^*(\alpha^\sharp(s))\,ds \\
&= \inf_\alpha \left\{ \int_0^1 \Psi^*(\alpha(s))ds : \ w(0) = e,\ w(1) = x,\ \dot{w} = (\alpha_1,\dots,\alpha_m,0,\dots,0) \right\} \\
&\leq \inf_\alpha \left\{ \int_0^1 \Psi^*(\alpha(s))ds : w(0) = e,\ w(1) = x,\ \dot{w} = f^{\mathbf{G}}(w,\alpha) \right\} \\
&= \Psi^{\mathbf{G}}(x)
\end{aligned}$$

Hence $\Psi^{\mathbf{G}}(x) \geq \Psi^*(\pi(x))$. Since, from the definition,

$$\Psi^{\mathbf{G}}(x) \leq \int_0^1 \Psi^*(\alpha^\sharp(s))ds,$$

we obtain (70). Now, let $x \in \mathbf{G}$; recalling that $\xi_1 : \mathbf{G} \to \mathfrak{g}$ is defined as $\xi_1 = (\exp|_{V_1})^{-1}$, we know that

$$\begin{aligned}
\Psi^{\mathbf{G}}(x) &= \inf_\alpha \left\{ \int_0^1 \Psi^*(\alpha(s))ds : \ w(0) = e,\ w(1) = x,\ \dot{w} = f^{\mathbf{G}}(w,\alpha) \right\} \\
&\geq \inf_\alpha \left\{ \int_0^1 \Psi^*(\alpha(s))ds : \ w(0) = e,\ w(1) = \xi_1(x),\ \dot{w} = (\alpha_1,\dots,\alpha_m,0,\dots,0) \right\} \\
&= \Psi^*(\pi(x)).
\end{aligned}$$

□

As an application of the proposition above, we show that for particular cost functions g the Hopf–Lax formula reduces to a finite–dimensional optimization problem as formulated in the introduction.

Corollary 4.1 *Suppose that $g(x) \geq \tilde{g}(x)$ for all $x \in \mathbb{R}^n$ where $\tilde{g} = g \circ \pi$ and $\pi : \mathbb{R}^n \to \mathbb{R}^m$ is the canonical projection. Then, under the assumptions of Proposition 4.1, we have that*

$$u(x,t) = \inf_{q \in \mathbb{R}^m} \left\{ t\Psi^* \left(\frac{\pi(x) - q}{t} \right) + \tilde{g}(q) \right\};$$

in particular, $x \to u(x,t)$ depends only on the first m components of x.

Proof: Taking into account (70) and (71), we have

$$\begin{aligned} u(x,t) &= \inf_{y \in \mathbf{G}} \left\{ t\Psi^{\mathbf{G}} \left(\delta_{\frac{1}{t}}(y^{-1}x) \right) + g(y) \right\} \\ &= \inf_{y \in x \exp(V_1)} \left\{ t\Psi^{\mathbf{G}} \left(\pi \left(\delta_{\frac{1}{t}}(y^{-1}x) \right) \right) + g(y) \right\} \\ &= \inf_{q \in \mathbb{R}^m} \left\{ t\Psi^* \left(\frac{\pi(x) - q}{t} \right) + \tilde{g}(q) \right\}, \end{aligned}$$

thereby proving the assertion. □

4.4 The value function $\mathcal{L}$ and Pontryagin's maximum principle

In subsection 2.3 we introduced and studied the properties of the value function $\mathcal{L} : \mathbf{G} \times [0, \infty) \to \mathbb{R}$. As we have seen, the $\mathbf{G}$–Legendre–Fenchel transform $\Psi^{\mathbf{G}}$ inherits these properties and it is the main tool in the proof of Theorem 3.4. The definition and properties of the functions $\mathcal{L}$ and $\Psi^{\mathbf{G}}$ are obtained without a reference to existence of the optimal controls of the associated control problems. However for more precise results, an existence result is needed.

Concerning the properties of the functions L, $\mathcal{L}$ and $\Psi^{\mathbf{G}}$, we shall consider again the optimal control problem (38)

$$\inf_{\alpha \in \mathcal{F}^{\mathbf{G}}_{e,0,x,t}(\mathbb{R}^m)} J_{e,0,x,t}(\alpha), \qquad \text{with} \quad J_{e,0,x,t}(\alpha) = \int_0^t \psi(\alpha(s))ds, \tag{72}$$

where $x \in \mathbf{G}$, and $t > 0$ are fixed. We shall prove the following result:

Theorem 4.2 *Let $\mathbf{G}$ be a Carnot group. Let t and R be fixed positive constants, and $x \in B(0,R)$. Let us suppose that ψ satisfies conditions* (H1) *and* (H2)*. Then there exists α^* minimizing $J_{e,0,x,t}$.*

The proof of this result is based on standard techniques but for the sake of completeness we give the full details in the Appendix (see section 6.2). In this subsection we focus our attention on necessary conditions of optimality for the fixed endpoint optimal problem (72), assuming that ψ is a C^1 function satisfying (H1) and (H2). The main result we use is the celebrated Pontryagin's Theorem. Introducing the Hamiltonian function

$$\mathcal{H} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}, \qquad \mathcal{H}(x,a,p_0,p) = p_0\psi(a) + p \cdot f^{\mathbf{G}}(x,a),$$

equation (20) gives

$$\mathcal{H}(x,a,p_0,p) = p_0\psi(a) + \sum_{j=1}^{m} a_j \left(p_j + \sum_{i=m+1}^{n} p_i q_{i,j}(x) \right). \tag{73}$$

The result of Pontryagin guarantees that necessary conditions for α^* (with associated trajectory w^*) to be an optimal control is the existence of a nonpositive constant p_0^*, of a constant c and of a continuous function $p^* = (p_1^*, \dots, p_n^*) : [0,t] \to \mathbb{R}^n$ such that in $[0,t]$ we have

$$\begin{aligned}
(p_0^*, p^*) &\neq 0, \\
\dot{p}^* &= -\nabla_x \mathcal{H}(w^*, \alpha^*, p_0^*, p^*), \\
\mathcal{H}(w^*, \alpha^*, p_0^*, p^*) &= \min_{a \in \mathbb{R}^m} \mathcal{H}(w^*, a, p_0^*, p^*), \qquad (74) \\
\mathcal{H}(w^*, \alpha^*, p_0^*, p^*) &= c. \qquad (75)
\end{aligned}$$

A control α^* with trajectory w^* that satisfies the previous necessary conditions is called extremal. Usually the pair (p_0^*, p^*) is called costate. See, for example, [21] and [24] for more details. There are two distinct possibilities for the constant p_0^* :

a. if $p_0^* \neq 0$ we say that the trajectory w^* is normal. Without loss of generality we can assume that $p_0^* = -1$ in this case;

b. if $p_0^* = 0$ we say that the trajectory w^* is abnormal. Then $\mathcal{H}$ does not depend on ψ and, in this case, the Pontryagin's principle is less useful. In particular, (73) and (74) imply that, in $[0,t]$, we have

$$\mathcal{H}(w^*, \alpha^*, 0, p^*) = \min_{a \in \mathbb{R}^m} \sum_{j=1}^{m} a_j \left(p_j^* + \sum_{i=m+1}^{n} p_i^* q_{i,j}(w^*) \right).$$

Since the minimum does exist, the linearity of the right hand side of the previous expression with respect to a implies that, within $[0,t]$,

$$p_j^* + \sum_{i=m+1}^{n} p_i^* q_{i,j}(w^*) = 0, \qquad \text{for } j = 1, \dots, m. \tag{76}$$

Notice that, for an extremal pair (α^*, w^*), there may exist several non zero vectors (p_0, p); and it can happen that the pair (α^*, w^*) is at the same time, both normal and abnormal. A strictly abnormal (normal) extremal is an abnormal (normal) extremal which is not also normal (abnormal). For more details see, for example, the paper by Sussmann in [10] and the book [2] (in particular, Theorem 12.10).

The normality gives further useful properties for the optimal control. In this context the convexity of ψ plays a fundamental role and, in order to prove these properties, we recall some well–known notions and results in Convex Analysis (see [31]). For every function $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}$, the subdifferential of f at a point $x_0 \in \Omega$ is defined as the set

$$\partial f(x_0) = \{ p \in \mathbb{R}^m : \ f(x) \geq f(x_0) + p \cdot (x - x_0), \quad \forall x \in \Omega \}.$$

A function f is convex if and only if $\partial f(x) \neq \emptyset$ for all $x \in \Omega$; moreover, if f is convex and C^1, then $\partial f(x) = \{\nabla f(x)\}$. Finally, a useful relationship between the subgradient ∂f and the Legendre–Fenchel trasform f^* is that $p \in \partial f(x)$ if and only if $f(x) + f^*(p) = x \cdot p$.

Now let us state the mentioned property for normal controls:

Proposition 4.2 *Let* $\mathbf{G}$ *be a Carnot group and let* $\psi \in C^1$ *satisfy* (H1), (H2). *Let* α^* *be a normal optimal control for the fixed endpoint problem (72). Then the* $\mathbb{R}^m$*–curve* α^* *is contained in a level curve of the function*

$$a \mapsto \psi^*(\nabla\psi(a)). \tag{77}$$

Proof: Let α^* be a normal optimal control that minimizes $J_{e,0,x,t}$ and assume that $p_0^* = -1$. Using (73) Pontryagin's Minimum Principle (74) implies that

$$\nabla\psi(\alpha^*(s)) - v(s) = 0,$$

for $s \in [0, t]$, where $v = (v_1, \dots, v_m) : [0, t] \to \mathbb{R}^m$ is defined by

$$v_j(s) = p_j^*(s) + \sum_{i=m+1}^{n} p_i^*(s) q_{i,j}(w^*(s)), \qquad \forall s \in [0, t].$$

Since ψ is convex and regular, $v(s) \in \partial\psi(\alpha^*(s)) = \{\nabla\psi(\alpha^*(s))\}$. Moreover we have that

$$\psi(\alpha^*(s)) + \psi^*(v(s)) = \alpha^*(s) \cdot v(s), \qquad \forall s \in [0, t].$$

Relation (75) implies that, in $[0, t]$,

$$\psi(\alpha^*) - \alpha^* \cdot v = c.$$

Hence the previous two equalities give

$$-\psi^*(v(s)) = \psi(\alpha^*(s)) - \alpha^*(s) \cdot v(s) = c, \qquad \forall s \in [0, t].$$

Since $v(s) = \nabla\psi(\alpha^*(s))$ the statement of the proposition follows. $\square$

The above proposition gives a necessary condition for a control α^* to be optimal and normal. A consequence of this is that in the case of homogeneous Hamiltonians of subsection 4.2, where the running cost of the optimal problem is $\psi = \Psi^*$, with $\Psi(\cdot) = \Phi(|\cdot|)$ (see (67)), the optimal control α^* has a good behavior: more precisely we have that, for a fixed $x \in I\!H$,

$$\Psi^{I\!H}(x) = \inf_{\alpha \in \mathcal{F}_x^{I\!H}(\mathbb{R}^2)} \int_0^1 \Psi^*(\alpha) ds = \int_0^1 \Psi^*(\alpha^*) ds$$

and we can conclude that $\Psi^*(\alpha^*(s))$ is constant on $[0, 1]$. To see this, note that $(\Phi(|\cdot|))^*(a) = \Phi^*(|a|)$ and $\nabla\Psi^*(a) = \frac{(\Phi^*)'(|a|)}{|a|} a$ for all $a \in \mathbb{R}^2$. The necessary condition of optimality (77) requires that the set $\{\alpha^*(s) : \ s \in [0, 1]\}$ is contained in a level set of the function

$$a \mapsto \psi^*(\nabla\psi(a)) = \Psi(\nabla\Psi^*(a)) = \Phi(\left|(\Phi^*)'(|a|)\right|);$$

since such function is radial, we obtain that $|\alpha^*(s)|$ is constant on $[0, 1]$. This implies that $\Psi^{I\!H}(x) = \Phi^*(|\alpha^*(s)|)$, for all $s \in [0, 1]$ and the expression of $\Psi^{I\!H}$ does not involve an integral.

This previous situation is a particular one: in general $\Psi(\alpha^*(s))$ is not constant on $s \in [0, 1]$. This entails, in particular, that the expression of the function $\Psi^{\mathbf{G}}$ may involve an integral.

In order to discuss some examples that prove the previous assertion, we have to provide a necessary condition for the normality of an optimal control:

Proposition 4.3 *Let* $\mathbf{G}$ *be a Carnot group of step 2. Let* ψ *be in* C^1 *and satisfying* (H1) *and* (H2)*. Let* α^* *be optimal for the problem* (72) *(where* $\overline{x} \neq e$*), with associated trajectory* w^**, and costate* (p_0^*, p^*)*. Suppose that* α^* *is abnormal. Then*

$$\sum_{k=1}^{m} \sum_{l=m+1}^{n} c_l \alpha_k^* \frac{\partial q_{l,k}}{\partial x_i}(w^*) = 0, \qquad \text{for } i = 1, \ldots, m, \tag{78}$$

where c_i *is constant for every* $i = m+1, \ldots, n$*. In particular, in the Heisenberg group* $\mathbb{H}$*, every extremal is normal.*

Proof: Taking into account (73) and the properties of the polynomials $q_{k,l}$, the Pontryagin Principle implies

$$\dot{p}_i^* = -\sum_{k=1}^{m} \sum_{l=N_j+1}^{n} p_l^* \alpha_k^* \frac{\partial q_{l,k}}{\partial x_i}(w^*), \qquad \text{for } 1 \le j \le r, \ N_{j-1} < i \le N_j \tag{79}$$

$$p_0^* \frac{\partial \psi}{\partial a_k}(\alpha^*) = -p_k^* - \sum_{l=m+1}^{n} p_l^* q_{l,k}(w^*), \qquad \text{for } k = 1, \ldots, m \tag{80}$$

In particular, for i satisfying $N_{r-1} < i \le N_r$, (79) implies that $p_i^* = c_i$ with c_i constant. Suppose that α^* is abnormal, i.e. $p_0^* = 0$. Since $\mathbf{G}$ is of step $r = 2$, the polynomials $q_{l,k}$ in (79) contain only linear terms of x_s with $1 \le s \le m$. Hence, by (79), we have

$$\dot{p}_i^* = -\sum_{k=1}^{m} \sum_{l=m+1}^{n} c_l \alpha_k^* \frac{\partial q_{l,k}}{\partial x_i}(w^*), \qquad \text{for } i = 1, \ldots, m.$$

Differentiating w.r.t. the time, (80) implies that

$$\dot{p}_i^* = -\sum_{l=m+1}^{n} \sum_{k=1}^{m} c_l \frac{\partial q_{l,i}}{\partial x_k}(w^*) \dot{w}_k^*, \qquad \text{for } i = 1, \ldots, m.$$

Since $\dfrac{\partial q_{l,i}}{\partial x_k}(w^*) = -\dfrac{\partial q_{l,k}}{\partial x_i}(w^*)$, using the previous two equalities we obtain (78).

In the case $\mathbf{G} = \mathbb{H}$, we recall that $q_{3,1} = -x_2/2$ and $q_{3,2} = x_1/2$: hence condition (76) is

$$\begin{aligned} p_1(s) + c_3 q_{3,1}(w^*(s)) &= p_1(s) - c_3 w_2^*/2 = 0, \\ p_2(s) + c_3 q_{3,2}(w^*(s)) &= p_2(s) + c_3 w_1^*/2 = 0, \end{aligned} \tag{81}$$

for $s \in [0,1]$, and condition (78) is

$$\begin{aligned} c_3 \left(\alpha_1^* \frac{\partial q_{3,1}}{\partial x_1}(w^*) + \alpha_2^* \frac{\partial q_{3,2}}{\partial x_1}(w^*) \right) &= \frac{c_3 \alpha_2^*}{2} = 0 \\ c_3 \left(\alpha_1^* \frac{\partial q_{3,1}}{\partial x_2}(w^*) + \alpha_2^* \frac{\partial q_{3,2}}{\partial x_2}(w^*) \right) &= -\frac{c_3 \alpha_1^*}{2} = 0. \end{aligned}$$

Since the final point of the trajectory w^* is different from the initial point e, then the control α^* cannot be identically 0 on $[0,1]$. Hence from the previous equalities we get that $c_3 = 0$. From (81) we obtain $p_1^* = p_2^* = 0$, i.e. $(p_0^*, p^*) = 0$. This is impossible, and hence $\mathbb{H}$ does not admit abnormal optimal controls. $\square$

We remark that, in the previous proposition, there are two reasons for which in $\mathbb{H}$ every extremal is normal: the step of the group is 2, and the dimension of the second layer V_2 is exactly 1. The next example shows that there exists Carnot group of step 2 that admits normal and abnormal optimal control for the problem (72)

Example 4.1 Consider the following vector fields

$$
\begin{aligned}
X_1(x) &= \partial_{x_1} - (x_2 + 2x_3)\partial_{x_4} - 3x_2\partial_{x_5}, \\
X_2(x) &= \partial_{x_2} + x_1\partial_{x_4} + 3x_1\partial_{x_5}, \\
X_3(x) &= \partial_{x_3} + 2x_1\partial_{x_4}, \\
X_4(x) &= \partial_{x_4}, \\
X_5(x) &= \partial_{x_5},
\end{aligned}
$$

for $x \in \mathbb{R}^5$. A direct calculation shows that the non vanishing bracket operations are only $[X_1, X_2] = 2X_4 + 6X_5$ and $[X_1, X_3] = 4X_4$. Hence the Lie algebra $\mathfrak{g}$ admits a decomposition of the type $\mathfrak{g} = V_1 \oplus [V_1, V_1]$, where $V_1 = \mathrm{span}\{X_1, X_2, X_3\}$. The Baker–Campbell–Dynkin–Hausdorff formula (60) gives us the law group on $\mathbf{G}$: for every $x, y \in \mathbb{R}^5$ we have

$$xy = \big(x_1+y_1, x_2+y_2, x_2+y_3, x_4+y_4+(x_1y_2-x_2y_1)+2(x_1y_3-x_3y_1), x_5+y_5+3(x_1y_2-x_2y_1)\big).$$

A curve γ is horizontal in $\mathbf{G}$ if

$$
\begin{aligned}
\dot\gamma_4 &= -(\gamma_2 + 2\gamma_3)\dot\gamma_1 + \gamma_1\dot\gamma_2 + 2\gamma_1\dot\gamma_3, \\
\dot\gamma_5 &= -3\gamma_2\dot\gamma_1 + 3\gamma_1\dot\gamma_2.
\end{aligned}
$$

Let us consider a convex, superlinear function $\psi \in C^1(\mathbb{R}^3)$ such that $\nabla\psi(0,1,1) \neq 0$ and suppose that we are interested in calculating

$$\Psi^{\mathbf{G}}(\overline{x}) = \inf_{\alpha \in \mathcal{F}^{\mathbf{G}}_{e,0,\overline{x},1}(\mathbb{R}^3)} \int_0^1 \psi(\alpha)ds,$$

with the fixed endpoint $\overline{x} = (0,1,1,0,0)$. The Hamiltonian in this case is

$$\mathcal{H} = p_0\psi(a) + p_1a_1 + p_2a_2 + p_3a_3 + p_4[-(x_2+2x_3)a_1 + x_1a_2 + 2x_1a_3] + p_5(-3x_2a_1 + 3x_1a_2).$$

Condition (79) and (80) imply, in $[0,1]$,

$$
\begin{aligned}
&\dot p_1 = -p_4\alpha_2 - 2p_4\alpha_3 - 3p_5\alpha_2 \\
&\dot p_2 = p_4\alpha_1 + 3p_5\alpha_1 \\
&\dot p_3 = 2p_4\alpha_1 \\
&p_4 = c_4 \\
&p_5 = c_5 \\
&p_0\psi_{\alpha_1}(\alpha) + p_1 - p_4(w_2 + 2w_3) - 3p_5w_2 = 0 \\
&p_0\psi_{\alpha_2}(\alpha) + p_2 + p_4w_1 + 3p_5w_1 = 0 \\
&p_0\psi_{\alpha_3}(\alpha) + p_3 + 2p_4w_1 = 0,
\end{aligned}
$$

where c_4 and c_5 are constants. Since the fixed endpoint $\overline{x}$ belongs to $\exp(V_1)$, we know (see (70)) that the optimal control is $\alpha^*(s) = (0,1,1)$, for every $s \in [0,1]$, and the associated optimal trajectory is $w^*(s) = (0,s,s,0,0)$. Such optimal control is abnormal since the costate $\mathbf{p} = (0,0,0,0,\beta,-\beta)$, with $\beta \neq 0$, solves the necessary condition.

At the same time, the above control is also normal, since the costate $\mathbf{p} = (-1, c_1, c_2, c_3, \beta, -\beta)$, with $\beta \in \mathbb{R}$ and (c_1, c_2, c_3) chosen to satisfy the equation

$$-\nabla\psi(0,1,1) + \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \mathbf{0},$$

solves the necessary condition.

In what follows we shall consider the example mentioned in the introduction in greater detail.

Example 4.2 Consider the Hamilton–Jacobi boundary value problem in the Heisenberg group $I\!H$

$$\begin{cases} u_t + \frac{1}{4}\left(u_{x_1} - \frac{1}{2}x_2 u_{x_3}\right)^2 + \frac{3}{4^{4/3}}\left(u_{x_2} + \frac{1}{2}x_1 u_{x_3}\right)^{4/3} = 0 & (x,t) \in I\!H \times (0,T) \\ u(x,0) = g(x) & x \in I\!H \end{cases} \tag{82}$$

Suppose that g satisfies our usual assumption. Our Theorem 3.4 guarantees that the unique viscosity solution is given by the Hopf–Lax formula (58), i.e.

$$u(x,t) = \inf_{y \in I\!H} \left\{ t\Psi^{I\!H}\left(\delta_{\frac{1}{t}}(y^{-1}x)\right) + g(y) \right\}$$

with $\Psi(a_1, a_2) = \frac{1}{4}a_1^2 + \frac{3}{4^{4/3}}a_2^{4/3}$. A direct calculation gives $\Psi^*(a) = a_1^2 + a_2^4$, and thus

$$\Psi^{I\!H}(x) = \inf_{\alpha \in \mathcal{F}_x^{I\!H}(\mathbb{R}^2)} \int_0^1 (\alpha_1^2 + \alpha_2^4) ds.$$

First of all, let us show that the expression of $\Psi^{I\!H}(x)$ involves an integral. We know, see Theorem 4.2, that for every fixed $x \in I\!H$ there exists an optimal control $\alpha^* = \alpha^*(x)$. Moreover such optimal control is normal (see Proposition 4.3) and satisfies condition (77) of Remark 4.2 with $\psi = \Psi^*$. Hence, in $[0,1]$, we have

$$\psi^*(\nabla\psi(\alpha^*)) = \Psi\Big((\nabla(\Psi^*))(\alpha^*)\Big) = \frac{1}{4}(2\alpha_1^*)^2 + \frac{3}{4^{4/3}}(4(\alpha_2^*)^3)^{4/3} = (\alpha_1^*)^2 + 3(\alpha_2^*)^4 = C.$$

where C is a constant that depends on x. This condition implies that, in general, $\Psi^*(\alpha^*(s))$ is not constant on $[0,1]$. More precisely, $\Psi^*(\alpha^*(s))$ is constant if and only if the final point of the associated trajectory lies on the horizontal plane $\exp(V_1) = \{(x_1, x_2, 0) : \ x_i \in \mathbb{R}\} \subset I\!H$. Moreover, in this particular case of the point x, Proposition 4.1 implies that

$$\Psi^{I\!H}(x_1, x_2, 0) = \Psi^*(x_1, x_2) = x_1^2 + x_2^4.$$

For a general Lipschitz function g, an explicit expression of the viscosity solution u is not possible. However, we know that u is regular and Proposition 2.3 gives us a good estimate of its horizontal gradient $\mathbb{X}u(x,t)$. More precisely, let us fix two positive R and T. For our choice of the function ψ, by (H2) and (22), we have $l_0 = 0$ and $M_0 = 1 + \epsilon$, for every $\epsilon > 0$; hence, by (28), we have that

$$M = T(1+\epsilon) + \mathrm{var}(g) + 1. \tag{83}$$

In the Heisenberg group we have $q_{3,1}(x_1, x_2, x_3) = -x_2/2$ and $q_{3,2}(x_1, x_2, x_3) = x_1/2$: hence the constants $\tilde{Q}_2^*$ and $\tilde{Q}_2$ that appear in (25) and in (32)-(33) are $\tilde{Q}_2^* = Q_2^* = 1/2$. The definition of R_1 and R_2 in (24) and (25) give us that R^* in Proposition 2.1 is

$$R^* = \sqrt{R_1^2 + R_1^2 + R_2^2} = \sqrt{2(R+M)^2 + (R + (R+M)M)^2}$$

Recalling that $\tilde{C}_1 = 0$ for definition, using (90) we obtain $\tilde{C}_2 = 3/2R^*$. By the definition of $\tilde{C}$ in (93), Proposition 2.3 gives the estimate for the horizontal gradient of the value function u :

$$|\mathbb{X}u(x,t)| \le K_g\sqrt{2+2\tilde{C}_2^2} = K_g\sqrt{2+9(R+M)^2+\frac{9}{2}(R+(R+M)M)^2},$$

for a.e. $x \in B_R$ and every $t \in [0,T]$, with M as in (83).

Finally, let us consider the particular case of the initial condition $g(x) = x_1 + x_2$ for the problem (82). Note that g is unbounded. Still, the solution is given by the Hopf–Lax formula (58) and, since g does not depend on x_3, from Corollary 4.1 we get

$$u(x,t) = x_1 + x_2 - t\left(\frac{1}{4} + \frac{3}{4^{4/3}}\right).$$

The next example, in some sense, is a generalization of the previous one considering a Carnot group of step three where we will find normal and abnormal controls:

Example 4.3 The Engel group $\mathbb{E}$ is a Carnot group of step 3 and it can be seen as an extension of the Heisenberg group $I\!\!H$: indeed, the Lie algebra $\mathfrak{e} = \mathbb{R}^4 = V_1 \oplus V_2 \oplus V_3$ is defined, using (59), by

$$\begin{array}{lll} V_1 = \operatorname{span}\{\tilde{X}_1, \tilde{X}_2\} & \text{with } \tilde{X}_1 = X_1 - (\frac{x_3}{2} + \frac{x_1x_2}{12})\partial_{x_4} & \text{and } \tilde{X}_2 = X_2 + \frac{x_1^2}{12}\partial_{x_4}, \\ V_2 = \operatorname{span}\{\tilde{X}_3\} & \text{with } \tilde{X}_3 = X_3 + \frac{x_1}{2}\partial_{x_4}, & \\ V_3 = \operatorname{span}\{\tilde{X}_4\} & \text{with } \tilde{X}_4 = \partial_{x_4}. & \end{array}$$

The bracket acts as $[\tilde{X}_1, \tilde{X}_2] = \tilde{X}_3$, $[\tilde{X}_1, \tilde{X}_3] = \tilde{X}_4$, and it vanishes for the other vectors. The horizontal gradient of a function u is

$$\mathbb{X}u = \Big(u_{x_1} - \frac{x_2}{2}u_{x_3} - \frac{6x_3 + x_1x_2}{12}u_{x_4}, u_{x_2} + \frac{x_1}{2}u_{x_3} + \frac{x_1^2}{12}u_{x_4}\Big).$$

Taking into account the action of the bracket, and since in $\mathfrak{e}$ in the Baker–Campbell–Dynkin–Hausdorff formula (60) there is one more term (precisely $([X,[X,Y]] + [Y,[Y,X]])/12$), the group law in $\mathbb{E}$ becomes

$$xx' = \Big(x_1 + x_1', x_2 + x_2', x_3 + x_3' + Q_3(x,x'), x_4 + x_4' + Q_4(x,x')\Big),$$

where $Q_3(x,x') = (x_1x_2' - x_1'x_2)/2$ coincides with the $I\!\!H$ situation and

$$Q_4(x,x') = (x_1x_3' - x_1'x_3)/2 + (x_1 - x_1')(x_1x_2' - x_2x_1')/12$$

for $x = (x_1,x_2,x_3,x_4)$ and $x' = (x_1',x_2',x_3',x_4')$. Hence in $I\!\!H$ and $\mathbb{E}$ the polynomials $q_{3,1}$ and $q_{3,2}$ coincide: moreover in $\mathbb{E}$ we have $q_{4,1}(x) = -x_3/2 - x_1x_2/12$, $q_{4,2}(x) = x_1^2/12$ and $q_{4,3}(x) = x_1/2$. An easy computation shows that $\gamma = (\gamma_1,\gamma_2,\gamma_3,\gamma_4) : \mathbb{R} \to \mathbb{E}$ is a horizontal curve if

$$\dot\gamma_3 = \frac{1}{2}(\gamma_1\dot\gamma_2 - \gamma_2\dot\gamma_1), \qquad \dot\gamma_4 = \frac{1}{12}(\gamma_1^2\dot\gamma_2 - \gamma_1\gamma_2\dot\gamma_1 - 6\gamma_3\dot\gamma_1).$$

Hence the dynamics is defined via the function $f^{\mathbb{E}} : \mathbb{R}^4 \times \mathbb{R}^2 \to \mathbb{R}^4$ as

$$f^{\mathbb{E}}(x,v) = \begin{pmatrix} v_1 \\ v_2 \\ (v_2x_1 - v_1x_2)/2 \\ (v_2x_1^2 - v_1x_1x_2 - 6v_1x_3)/12 \end{pmatrix}, \qquad \forall v \in \mathbb{R}^2,\ x \in \mathbb{R}^4.$$

Theorem 3.4 guarantees that, with assumption on g,

$$u(x,t) = \inf_{y\in\mathbb{E}} \left\{ t\Psi^{\mathbb{E}}\left(\delta_{\frac{1}{t}}(y^{-1}x)\right) + g(y)\right\},$$

with $\Psi(a_1,a_2) = \frac{1}{4}a_1^2 + \frac{3}{4^{4/3}}a_2^{4/3}$, is the unique viscosity solution for

$$\begin{cases} u_t + \dfrac{1}{4}\left(u_{x_1} - \dfrac{x_2}{2}u_{x_3} - \dfrac{6x_3+x_1x_2}{12}u_{x_4}\right)^2 + \dfrac{3}{4^{4/3}}\left(u_{x_2} + \dfrac{x_1}{2}u_{x_3} + \dfrac{x_1^2}{12}u_{x_4}\right)^{4/3} = 0 \\ u(x,0) = g(x) \end{cases}$$

Clearly

$$\Psi^{\mathbb{E}}(x) = \inf_{\alpha\in\mathcal{F}_x^{\mathbb{E}}(\mathbb{R}^2)} \int_0^1 (\alpha_1^2 + \alpha_2^4)ds.$$

Consider $\overline{x}\in\mathbb{E}$ fixed endpoint. For the previous optimal control problem, the Hamiltonian is given by, with $p = (p_1,p_2,p_2,p_4)$,

$$\begin{aligned} \mathcal{H} &= p_0\Psi^*(a) + p\cdot f^{\mathbb{E}}(x,a) \\ &= p_0(a_1^2+a_2^4) + p_1a_1 + p_2a_2 + p_3(x_1a_2 - x_2a_1)/2 + p_4(x_1^2a_2 - x_1x_2a_1 - 6x_3a_1)/12 \end{aligned}$$

and the necessary condition of Pontryagin gives

$$\begin{aligned} &\dot{p}_1^* = -p_3^*\alpha_2^*/2 + w_2^*p_4^*\alpha_1^*/12 - w_1^*p_4^*\alpha_2^*/6 \\ &\dot{p}_2^* = p_3^*\alpha_1^*/2 + w_1^*p_4^*\alpha_1^*/12 \\ &\dot{p}_3^* = p_4^*\alpha_1^*/2 \\ &p_4^* = c_4 \\ &2p_0^*\alpha_1^* + p_1^* - p_3^*w_2^*/2 + p_4^*(-w_1^*w_2^* - 6w_3^*)/12 = 0 \\ &4p_0^*(\alpha_2^*)^3 + p_2^* + p_3^*w_1^*/2 + p_4^*(w_1^*)^2/12 = 0 \end{aligned}$$

where c_4 is a constant. Now, if we consider the particular endpoint $\overline{x} = (0,1,0,0)\in\exp(V_1)$, we know (see (70)) that the optimal control is $\alpha^*(s) = (0,1)$, for every $s\in[0,1]$, and the associated optimal trajectory is $w^*(s) = (0,s,0,0)$. For every constant $\beta\in\mathbb{R}$, let us consider $(p_0,p) = (\beta,0,-4\beta,0,c_4)$. An easy computation shows that the couple (α^*,w^*) is extremal with costate (p_0,p). Hence such extremal is normal and abnormal.

Since in $\mathbb{E}$ the constants $\tilde{Q}_j^*$ and $\tilde{Q}_j$ that appear in (25) and in (32)-(33) are $\tilde{Q}_2^* = Q_2^* = 1/2$, $\tilde{Q}_3^* = Q_3^* = 7/12$, hence R^* in Proposition 2.1 is

$$R^* = \sqrt{2(R+M)^2 + (R+(R+M)M)^2 + [R + 7/6(R+(R+M)M)^2M]^2},$$

with M as in (83). Proposition 2.3 gives the estimate for the horizontal gradient of the value function u :

$$|\mathbb{X}u(x,t)| \le K_g\sqrt{2 + 2\tilde{C}_2^2 + 2\tilde{C}_3^2}, \qquad \text{a.e. in } B_R\times[0,T]$$

where, by (93), $\tilde{C}_2 = 3R^*/2$ and $\tilde{C}_3 = 7(R^*)^2/12\left(1 + 2\sqrt{1+9(R^*)^2/4}\right)$.

We conclude this section with a result concerning the behavior of the functions $\mathcal{L}$ and $\Psi^{\mathbf{G}}$ for large $|x|$:

Proposition 4.4 *Let* $\mathbf{G}$ *be a Carnot group and let* ψ *satisfy (H1), (H2), and* $\psi \in C^1$. *Let* $t > 0$ *be fixed. Suppose that, for every* $x \in \mathbf{G}$, *the optimal control of the problem*

$$\inf_{\alpha \in \mathcal{F}^{\mathbf{G}}_{e,0,x,t}(\mathbb{R}^m)} \int_0^t \psi(\alpha(s))ds$$

is normal. Then

$$\lim_{|x|\to\infty} \mathcal{L}(x,t) = +\infty. \tag{84}$$

In the proof the following lemma will be used:

Lemma 4.1 *Let* $\psi \in C^1(\mathbb{R}^m)$ *satisfy* (H1), (H2). *For* $h > 0$ *set* $\inf_{|x|\geq h} \dfrac{\psi(x)}{|x|} = C'_h$. *Then* $\inf_{|x|>h} |\nabla\psi(x)| \geq C'_h$ *and* $\lim_{h\to\infty} C'_h = +\infty$.

Proof: From the assumption, for any nonzero $x \in \mathbb{R}^m$, $\psi(0) - \psi(x) \geq \nabla\psi(x)\cdot(-x)$. This implies that

$$|\nabla\psi(x)| \geq \nabla\psi(x)\cdot\frac{x}{|x|} \geq \frac{\psi(x)-\psi(0)}{|x|}.$$

To prove the assertion, by contradiction suppose that $C'_h \leq K$, for some $K \in \mathbb{R}$ and for every h. Then, by taking a suitable subsequence, there exists $\{x_h\}_h$, $|x_h| \geq h$, such that $\frac{\psi(x_h)}{|x_h|} < K + \frac{1}{h}$, contradicting the superlinearity of ψ. $\square$

Proof of Proposition 4.1: For the sake of simplicity, we prove this result only in the Heisenberg group. Let $\{x^h\}_{h\geq 0} \subset I\!\!H$ be a sequence, such that $|x^h| \to \infty$ for $h \to \infty$. For every h, let us consider the optimal control $\alpha^h \in \mathcal{F}^{I\!\!H}_{e,0,x^h,t}(\mathbb{R}^2)$, and denote by w^h the associated trajectory; if there exists a constant A such that $|\alpha^h(s)| \leq A$, for every $s \in [0,t]$ and h, then we have that

$$\begin{aligned}
|w_i^h(s)| &\leq \int_0^s |\alpha_i^h(v)|\,dv \leq tA, \qquad \forall s \in [0,t], \quad i = 1,2,\\
|w_3^h(t)| &\leq \frac{1}{2}\int_0^t |\alpha_2^h(s)w_1^h(s) - \alpha_1^h(s)w_2^h(s)|\,ds \leq A^2t^2,
\end{aligned}$$

implying that $|x^h| = |w^h(t)| \leq tA\sqrt{2 + A^2t^2}$; this contradicts $|x^h| \to \infty$. Then, there exists a sequence $\{K_h\}$, $K_h \to +\infty$, and $S_h \subset [0,t]$, $|S_h| > 0$, such that $|\alpha^h(s)| \geq K_h$ for every $s \in S_h$. Since α_h is a normal optimal strategy, from (77) there exists C_h such that

$$\psi^*(\nabla\psi(\alpha_h(s))) = C_h, \quad \forall s \in [0,t].$$

Set $C'_h = \inf_{|x|\geq K_h} \frac{\psi(x)}{|x|}$. The previous lemma implies $C'_h \to \infty$. Define $C''_h = \inf_{|z|\geq C'_h} \psi^*(z)$; from the superlinearity of ψ^*, we argue in particular that $C''_h \to \infty$. Therefore,

$$\psi^*(\nabla\psi(\alpha_h(s))) \geq C''_h, \qquad \forall s \in S_h.$$

Notice that, since $\psi^*(\nabla\psi(\alpha_h(s))) = C_h$ for every $s \in [0,t]$, we have that $C_h \geq C''_h$ and $S_h = [0,t]$, and $|\alpha_h(s)| > K_h$ for every $s \in [0,t]$. Define $C'''_h = \inf_{|x|\geq K_h} \psi(x)$; then $C'''_h \to +\infty$, because ψ is superlinear. In particular, $\psi(\alpha_h(s)) > C'''_h$ for every $s \in [0,t]$ and for every h, which implies (84). $\square$

5 Final remarks and open questions

Question 1: Considering the Hamilton–Jacobi boundary value problem in (7) we raise the question about the possibility to obtain an Hopf–Lax type formula only with the assumption that the system $\mathbb{X} = (X_1, \ldots, X_m)$ satisfies a Hörmander condition without the additional assumption of an underlying Carnot group structure.

We think that with the methods of this paper it would be possible to address this problem. An even more general situation would be to consider Hamilton-Jacobi equations where the associated control problem contains a drift term. Such a problem has been recently investigated by Agrachev and Lee in [1]. It is interesting to see that in this case additional assumptions are needed for the validity of the Hopf-Lax formula.

Question 2: It would interesting to develop a game theoretic approach to Hopf-Lax formulas in the setting of Carnot groups or more general sub-Riemannian geometries.

Hopf–Lax formulas that we consider in this paper are of the type *inf–sup*. More precisely, if in (2) we insert the definition of L in (3), we obtain

$$u(x,t) = \inf_{y\in\mathbb{R}^n} \left\{ g(y) + t \sup_{q\in\mathbb{R}^n} \left(\frac{(x-y)\cdot q}{t} - H(q) \right) \right\}.$$

It is clear that in the previous expression, in general, if we change the order of the *inf* and of the *sup*, we obtain a different function. Hopf–Lax formulas with *sup–inf* can be obtained using the theory of differential games; see [9], [23].

Question 3: It would be interesting to study the relation between the Legendre-Fenchel transform $\Psi^{\mathbf{G}}$ introduced in Section 3 of our paper and the one defined by Calogero and Pini in [13] in the case of the first Heisenberg group $\mathbf{G} = I\!H$ as

$$\Psi_x^*(v) = \sup_{\{y=(y_1,y_2,y_3)\in x\exp(V_1)\}} \left(v\cdot(y_1,y_2) - \Psi(y) \right).$$

It is clear that this definition can be extended to the general Carnot setting but the relationship between this notion and $\Psi^{\mathbf{G}}$ is not clear to us.

Question 4: A natural question is to consider a weaker regularity assumption on g, by replacing the Lipschitz continuity in the Euclidean sense with the assumption of Lipschitz continuity with respect the d_{CC} metric. In this line of investigation is it possible to obtain a Lipschitz regularity of the value function u with respect the d_{CC} metric (see Proposition 2.3)?

6 Appendix

6.1 Lipschitz estimate of the value function u

In this subsection we provide the proof of Proposition 2.3 concerning a precise estimate of the norm of the horizontal gradient of the value function.

Proof of Proposition 2.3: Let us consider x and $\hat{x}$ in $\mathbf{G}$ with $|x| \le R$, $|\hat{x}| \le R$, and suppose that $\mathbb{X}u(x,t)$ exists. We know that

$$u(x,t) = \inf_{\alpha\in\overline{\mathcal{F}}^{\mathbf{G}}_{x,0,\cdot,t}(B_\mu)} J_{x,t}(\alpha), \qquad \text{and} \qquad u(\hat{x},t) = \inf_{\alpha\in\overline{\mathcal{F}}^{\mathbf{G}}_{\hat{x},0,\cdot,t}(B_\mu)} J_{\hat{x},t}(\alpha).$$

Fix $\epsilon > 0$, and let α be a control such that

$$u(\hat{x}, t) + \epsilon \geq \int_0^t \psi(\alpha(s))\, ds + g(\hat{w}(t)),$$

where $\hat{w}$ is the trajectory associated to α with $\hat{w}(0) = \hat{x}$.

Clearly, we get that

$$u(x, t) \leq \int_0^t \psi(\alpha(s))\, ds + g(w(t)),$$

where w is the trajectory associated to α with $w(0) = x$. This implies that

$$u(x, t) - u(\hat{x}, t) \leq g(w(t)) - g(\hat{w}(t)) + \epsilon.$$

In order to exploit the Lipschitz property of g, one estimates $|w(t) - \hat{w}(t)|$.

Let i be such that $1 \leq i \leq N_j$, with $j = 1$; then, from $\dot{w} = -f^{\mathbf{G}}(w, \alpha)$,

$$|w_i(s) - \hat{w}_i(s)| = \left| x_i - \int_0^s \alpha_i(v) dv - \hat{x}_i + \int_0^s \alpha_i(v) dv \right| = |x_i - \hat{x}_i| \qquad s \in [0, t]. \tag{85}$$

Let us consider an integer p, with $1 \leq p \leq m$. If $\lambda \in \mathbb{R}^+$ and $\hat{x} = x \exp(\lambda X_p)$, where X_p is an element of the basis of $\mathcal{H}\mathbf{G}$, we obtain

$$\lim_{\lambda \to 0^+} \frac{|w_i(s) - \hat{w}_i^p(s, \lambda)|}{\lambda} = \begin{cases} 0 & \text{if } i \neq p \\ 1 & \text{if } i = p \end{cases} \tag{86}$$

where $\hat{w}^p$ is the trajectory associated to the control α and the initial point $\hat{x}$. Let us consider i, $N_j < i \leq N_{j+1}$ with $j \geq 1$. Using the same arguments and notations of the proof of Theorem 2.1, for every $s \in [0, t]$ we have, from (33) and from the inequality $\|\alpha\|_1 \leq R^*$,

$$\begin{aligned} |w_i(s) - \hat{w}_i(s, \lambda)| &\leq |x_i - \hat{x}_i| + \sum_{k=1}^m \int_0^s |\alpha_k(v)|\, |q_{i,k}(w(v)) - q_{i,k}(\hat{w}(v, \lambda))|\, dv \\ &\leq |x_i - \hat{x}_i| + m\tilde{Q}_{j+1}(R^*)^j \sup_{v \in [0,t]} \sqrt{\sum_{l=1}^{N_j} |w_l(v) - \hat{w}_l(v, \lambda)|^2}. \end{aligned} \tag{87}$$

Since $\hat{x} = x \exp(\lambda X_p)$, we have $\hat{x}_i = x_i + Q_i(x, \exp(\lambda X_p))$; we note that the Taylor's formula of Q_i is

$$Q_i(x, \exp(\lambda X_p)) = Q_i(x, 0) + q_{i,p}(x, 0)\lambda + r_{i,p}(x, \lambda)$$

with $r_{i,p}(x, \lambda)/\lambda \to 0$ for $\lambda \to 0^+$. We recall that $Q_i(x, 0) = 0$. Hence, by (86) and (87), for every $s \in [0, t]$ we have

$$\lim_{\lambda \to 0^+} \frac{|w_i(s) - \hat{w}_i^p(s, \lambda)|}{\lambda} \leq \tag{88}$$

$$\begin{aligned} &\leq |q_{i,p}(x)| + m\tilde{Q}_{j+1}(R^*)^j \lim_{\lambda \to 0} \frac{1}{\lambda} \sqrt{\lambda^2 + \sum_{l=N_1+1}^{N_j} \sup_{v \in [0,t]} \left| w_l(v) - \hat{w}_l^p(v, \lambda) \right|^2} \\ &\leq \tilde{Q}_{j+1}(R^*)^j + m\tilde{Q}_{j+1}(R^*)^j \sqrt{1 + \sum_{l=N_1+1}^{N_j} \lim_{\lambda \to 0} \sup_{v \in [0,t]} \frac{\left| w_l(v) - \hat{w}_l^p(v, \lambda) \right|^2}{\lambda^2}} \\ &\leq \tilde{Q}_{j+1}(R^*)^j \left(1 + m \sqrt{1 + \sum_{d=1}^{j-1} \sum_{l=N_d+1}^{N_{d+1}} \lim_{\lambda \to 0} \sup_{v \in [0,t]} \frac{\left| w_l(v) - \hat{w}_l^p(v, \lambda) \right|^2}{\lambda^2}} \right). \end{aligned} \tag{89}$$

Set $\tilde{C}_1 = 0$, and for every j with $2 \le j \le r$, we define

$$\tilde{C}_j = \tilde{Q}_j (R^*)^{j-1} \left(1 + m \sqrt{1 + \sum_{d=1}^{j-1} n_d \tilde{C}_d^2} \right). \tag{90}$$

We will show that for every i, $N_j < i \le N_{j+1}$ with $j \ge 1$, we have

$$\lim_{\lambda \to 0^+} \sup_{v \in [0,t]} \frac{|w_i(v) - \hat{w}_i^p(v,\lambda)|}{\lambda} \le \tilde{C}_{j+1}. \tag{91}$$

Let us prove the assertion by induction on the step j. For $J = 1$ and for every i, with $m = N_1 < i \le N_2$ we obtain, by (89),

$$\lim_{\lambda \to 0} \sup_{v \in [0,t]} \frac{|w_i(v) - \hat{w}_i^p(v,\lambda)|}{\lambda} \le \tilde{Q}_2 R^* (1+m) = \tilde{C}_2. \tag{92}$$

Now, let us suppose that (91) holds for every i and j, with $N_j < i \le N_{j+1},\ 1 \le j \le J < r-1$, Consider i such that $N_{J+1} < i \le N_{J+2}$: we obtain by (89)

$$\lim_{\lambda \to 0} \sup_{v \in [0,t]} \frac{|w_i(s) - \hat{w}_i^p(s,\lambda)|}{\lambda} \quad \le \quad \tilde{Q}_{J+2} (R^*)^{J+1} \left(1 + m \sqrt{1 + \sum_{d=0}^{J} n_{d+1} \tilde{C}_{d+1}^2} \right) = \tilde{C}_{J+2}.$$

Hence (91) holds. Consequently, for a.e. $x \in B_{R^*}$ we have

$$\begin{aligned}
|\mathbb{X}u(x,t)|^2 &= \sum_{p=1}^m |X_p u(x,t)|^2 \\
&= \sum_{p=1}^m \left| \lim_{\lambda \to 0} \frac{u(x \exp(\lambda X_p), t) - u(x,t)}{\lambda} \right|^2 \\
&\le K_g^2 \sum_{p=1}^m \sum_{i=1}^n \lim_{\lambda \to 0} \frac{|w_i(t) - \hat{w}_i^p(t)|^2}{\lambda^2} \\
&\le K_g^2 \left(m + m \sum_{d=1}^r n_d \tilde{C}_d^2 \right) \qquad (93) \\
&= K_g^2 \tilde{C}^2.
\end{aligned}$$

□

6.2 Existence of the optimal control for the fixed endpoint OCP

This subsection is devoted to the proof of Theorem 4.2 concerning the existence of an optimal control for the problem

$$\inf_{\alpha \in \mathcal{F}^{\mathbf{G}}_{e,0,x,t}(\mathbb{R}^m)} J_{e,0,x,t}(\alpha), \qquad \text{where} \quad J_{e,0,x,t}(\alpha) = \int_0^t \psi(\alpha(s))ds, \tag{94}$$

where $x \in \mathbf{G}$ and $t > 0$ are fixed. The proof follows the idea due to Cesari [17] contained in Theorem 4.1 of [24] that provides sufficient conditions for unbounded control sets. For completeness, let us recall this result:

Theorem 6.1 (see Theorem 4.1 in [24]) *Let* $\psi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R},\ g : \mathbb{R}^n \times [0, \infty) \to \mathbb{R}$ *and* $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$. *Let* $x \in \mathbb{R}^n$ *be fixed, and* A *be a closed subset of* $\mathbb{R}^m$. *Let us consider the problem*

$$\inf_{\alpha \in \mathcal{F}_{x,0,S}(A)} \int_0^t \psi(w(s), \alpha(s))ds + g(w(t), t), \tag{95}$$

where S, *the target set, is a compact subset of* $\mathbb{R}^n \times (0, \infty)$, *and the nonempty set of admissible controls is* $\mathcal{F}_{x,0,S}(A) = \{\alpha : [0, t] \to A \text{ measurable};\ \dot{w} = f(w, \alpha),\ w(0) = x,\ (w(t), t) \in S\}$. *Suppose that the function* f *is continuous, and there exist positive constants* C_1 *and* C_2 *such that, for all* $x, y \in \mathbb{R}^n$ *and* $a \in A$,

$$|f(x, a)| \le C_1(1 + |x| + |a|), \tag{96}$$

$$|f(x, a) - f(y, a)| \le C_2|x - y|(1 + |a|). \tag{97}$$

Assume that the functions ψ *and* g *are continuous, and* ψ *satisfies* (H2). *Finally, suppose that the set*

$$F(y) = \{v \in \mathbb{R}^{n+1} :\ (v_1, \ldots, v_n) = f(y, a),\ v_{n+1} \ge \psi(y, a),\ a \in A\} \tag{98}$$

is convex for every $y \in \mathbb{R}^n$. *Then there exists an optimal control for problem (95).*

In the case of problem (94) the difficulties arise from the dynamics $\dot{w} = f^{\mathbf{G}}(w, \alpha)$ that does not satisfy the growth conditions (96) and (97). On the other hand, in our problem the trajectory has the initial point $(e, 0)$ and the final point (x, t) fixed, i.e. the target set is $S = \{(x, t)\}$: this will simplify some notations and results with respect to the arguments in [24]. The line of our proof is similar to the proof of Theorem 95: this line consists of six lemmata and a final conclusion. Some of these lemmata are very technical and in our situation they present really few differences with respect to the presentation in [24]; hence we refer to this text for more details on the proof.

We start with a technical lemma:

Lemma 6.1 (see Lemma 5.1, Chapter III, in [24]) *Given* $\nu > 0,\ \mu > 0$ *there exists* $\delta > 0$ *such that for every set* $I \subset [0, t]$, *with* $|I| < \delta$ *and* $\int_0^t l(|\alpha(s)|)ds \le \nu$ *il follows that* $\int_I |\alpha(s)|ds < \mu$.

Proof: See the mentioned lemma. □

Despite f does not satisfy the growth condition (97), we can get an upper estimate of the distance from the origin of the points of the trajectories for L^1–controls using the "stratification" of the group $\mathbf{G}$:

Proposition 6.1 *Let* $\mathbf{G}$ *be a Carnot group. Let* t *and* R *be fixed positive constants. Let* $x \in B(0, R)$ *and let* ψ *satisfy* (H2). *For every* M_1, *such that*

$$\{\alpha \in \mathcal{F}^{\mathbf{G}}_{e,0,x,t}(\mathbb{R}^m) :\ ||\alpha||_1 \le M_1\} \ne \emptyset,$$

there exists a constant R_1^* *such that, for every control* $\alpha \in \mathcal{F}^{\mathbf{G}}_{e,0,x,t}(\mathbb{R}^m)$ *with* $\|\alpha\|_1 \le M_1$, *we have*

$$|w(s)| \le R_1^*, \qquad s \in [0, t],$$

where w *is the trajectory associated to* α *with* $w(0) = e,\ w(t) = x$.

Proof: The proof follows the line of the proof of Proposition 2.1. $\square$

The superlinearity of ψ allows us to restrict our attention to a subset $\mathcal{H}^\nu$ of $\mathcal{F}^{\mathbf{G}}_{e,0,x,t}(\mathbb{R}^m)$: this subset is an equiabsolutely continuous set (see [24] for the relevant definitions) and hence it contains a minimizing sequence that converges to a function w^*, candidate to be the optimal trajectory. This is the object of the following result:

Lemma 6.2 (a new version of Lemma 5.2, Chapter III, in [24]) *Let ψ be as in* (H2)*. Let t and R be fixed positive constants, and $x \in B(0,R)$. Then there exists a minimizing sequence of controls $\{\alpha^r\}_r \subset L^1([0,t]) \cap \mathcal{F}^{\mathbf{G}}_{e,0,x,t}(\mathbb{R}^m)$ such that w^r converges uniformly on $[0,t]$ to a limit w^* as $r \to \infty$, where w^r is the trajectory associated to α^r defined via the dynamics and satisfying $w^r(0) = e,\ w^r(t) = x$. Moreover, w^* is absolutely continuous.*

Proof: Let $\{\alpha^r\}_r \subset \mathcal{F}^{\mathbf{G}}_{e,0,x,t}(\mathbb{R}^m)$ be a minimizing sequence. Then there exists a constant γ such that $J_{e,0,x,t}(\alpha^r) \le \gamma$ for every r. Hence

$$\int_0^t l(|\alpha^r(s)|)ds \le \int_0^t [\psi(\alpha^r(s)) + l_0]ds \le \gamma + l_0 t.$$

Set $\nu = \gamma + l_0 t$ and let us define $\mathcal{H}^\nu$ as the set of the trajectories w, with $w(0) = e$, $w(t) = x$, and associated to a control $\alpha \in \mathcal{F}_{e,0,x,t}$ with $\int_0^t l(|\alpha(s)|)ds \le \nu$. We show that $\mathcal{H}^\nu$ is an equiabsolutely continuous set. Let M_0 be as in (22): we obtain

$$\begin{aligned}\|\alpha\|_1 &= \int_{\{s:\ |\alpha(s)|\le M_0\}} |\alpha(s)|ds + \int_{\{s:\ |\alpha(s)|> M_0\}} |\alpha(s)|ds \\ &\le tM_0 + \int_0^t l(|\alpha(s)|)ds.\end{aligned}$$

If $w \in \mathcal{H}^\nu$, its control α satisfies $\|\alpha\|_1 \le tM_0 + \nu = M_1$. Proposition 6.1 implies that $|w(s)| \le R_1^*$ in $[0,t]$; hence $\mathcal{H}^\nu$ is uniformly bounded. Moreover, for every $0 \le s \le s' \le t$, we have

$$|w_i(s') - w_i(s)| \le \int_s^{s'} |\alpha_i(v)|dv \le \int_s^{s'} |\alpha(v)|dv, \qquad \text{for } i = 1, \ldots, m.$$

Let us consider i and j, with $N_j < i \le N_{j+1}$ and $1 < j \le r$. Defining $\tilde{Q}_j$ as in the proof of Theorem 2.1, we have

$$\begin{aligned}|w_i(s') - w_i(s)| &\le \int_s^{s'} |f_i(v)|dv \\ &\le \sum_{k=1}^m \int_s^{s'} |\alpha_k(v) q_{i,k}(w(v))|dv \\ &\le m\tilde{Q}_{j+1}(R^*)^j \int_s^{s'} |\alpha(v)|dv.\end{aligned}$$

Hence there exists $C > 0$ such that $|w(s') - w(s)| \le C \int_s^{s'} |\alpha(v)|dv$. If $I \subset [0,t]$ is the union of k disjoint intervals $[s_i, s_i'],\ i = 1, \ldots, k$, we obtain

$$\sum_{i=1}^k |w(s_i') - w(s_i)| \le C \int_I |\alpha(v)|dv.$$

Given $\epsilon > 0$ let $\mu = \frac{\epsilon}{2C}$, and take δ as in Lemma 6.1 small enough such that $C\delta \leq \frac{\epsilon}{2}$. If I is such that $|I| < \delta$, then

$$\sum_{i=1}^{k} |w(s_i') - w(s_i)| \leq C \int_I |\alpha(v)| dv \leq C\mu = \frac{\epsilon}{2}.$$

Thus $\mathcal{H}^\nu$ is equiabsolutely continuous. The proof is completed using Ascoli's Theorem as in Lemma 5.2 in [24]. □

The previous proposition implies that w^* is absolutely continuous and $\dot{w}^*$ is defined a.e. Hence, if we consider its associated control $\alpha^* = (\dot{w}_1^*, \ldots, \dot{w}_m^*)$, α^* is a candidate to be the optimal control that minimizes $J_{e,0,x,t}$, but we are not able to guarantee that

$$\lim_{r\to\infty} \int_0^t \psi(\alpha^r(v)) dv = \int_0^t \psi(\alpha^*(v)) dv.$$

The aim of the next three lemmata is to prove that this is actually true.

Lemma 6.3 (see Lemma 5.3, Chapter III, in [24]) *Let the assumptions of the previous lemma be satisfied. Set*

$$Z^r(s) = \int_0^s \psi(\alpha^r(v)) dv, \qquad r = 1, 2, \ldots. \tag{99}$$

Then the sequence $\{\alpha^r\}$ in Lemma 6.2 can be chosen in such a way that $Z^r(s)$ converges pointwisely on $[0,t]$ to a limit $Z^*(s)$ as $r \to \infty$. Moreover $\mathcal{Z}^*(s) = Z^*(s) + l_0 s$ is a monotone function on $[0,t]$.

Proof: Let us give the idea of the proof (for details, see the mentioned lemma in [24]). Is is easy to show that there exists a constant C such that

$$0 \leq \int_0^s (\psi(\alpha^r(v)) + l_0) dv \leq C$$

for every $r = 1, 2, \ldots$ and $s \in [0,t]$. Now, if we consider the function

$$\mathcal{Z}^r(s) = Z^r(s) + l_0 s, \tag{100}$$

then, for every fixed r, we obtain that $\mathcal{Z}^r$ is monotone on $[0,t]$ and $0 \leq \mathcal{Z}^r \leq C$. By Helly's Theorem, there exists a subsequence of $\{\mathcal{Z}^r\}$ which converges pointwisely to a limit $\mathcal{Z}^*$, i.e.

$$\lim_{r\to\infty} \mathcal{Z}^r(s) = \mathcal{Z}^*(s), \tag{101}$$

for $s \in [0,t]$. Taking $Z^*(s) = \mathcal{Z}^*(s) - l_0 s$ we get the assertion. □

Now, let us consider for every $x \in \mathbb{R}^n$ the set $F(x)$ defined in (98). Note that the convexity of ψ implies that this set is convex, for every x; this property is crucial in order to prove Lemma 6.5. The proof of the next lemma is very technical and does not present considerable differences with respect to the lemma in [24]: we need essentially to bypass the lack of assumption (96) concerning the growth of $f^{\mathbf{G}}$.

Lemma 6.4 (see Lemma 5.4, Chapter III, in [24]) *Let the assumptions of Theorem 4.2 be satisfied. Let $x \in \mathbf{G}$ be fixed. Let*

$$\tilde{\mathcal{O}}_\nu = \bigcup_{|x'-x|<\nu} F(x')$$

and C_ν be the convex hull of $\tilde{\mathcal{O}}_\nu$. Then

$$F(x) = \bigcap_{\nu>0} \overline{C_\nu}.$$

Proof: As we mentioned, we refer to [24] for the details. Let us spend a few words on the lack of the growth condition (96) on $f^{\mathbf{G}}$. Note that in our case, where the function $f^{\mathbf{G}}$ represents the dynamics for a horizontal curve in a Carnot group $\mathbf{G}$, we have, by (20), for every $y \in \mathbf{G}$

$$|f^{\mathbf{G}}(y,\alpha)| \le \sqrt{\sum_{j=1}^{m} \alpha_j^2 + \sum_{j=m+1}^{r} \sum_{i=1}^{m} \alpha_i^2 q_{j,i}^2(y)} \le |\alpha| \sqrt{1 + \sum_{j=m+1}^{r} \sum_{i=1}^{m} q_{j,i}^2(y)}.$$

Now it is clear that, if $x \in \mathbf{G}$ is fixed and $x' \in \mathbf{G}$ is such that $|x'-x| < \nu$, using the continuity of the functions $q_{j,i}$ we obtain that $|f^{\mathbf{G}}(x',\alpha)| \le |\alpha| C'$ for some positive constant $C' = C'(x,\nu)$. In this situation the right hand side of the inequality

$$0 \le \frac{|f^{\mathbf{G}}(x',\alpha)|}{\psi(\alpha)} \le \frac{|\alpha| C'}{l(|\alpha|) - l_0}$$

goes to 0 if $|\alpha| \to \infty$. $\square$

Since w^* is absolutely continuous, and $Z^*(s) + l_0 s$ is monotone, the derivatives $\dot{w}^*(s)$ and $\dot{Z}^*(s)$ exist for almost all $s \in [0,t]$.

In the next lemma we are going to prove that such limit function satisfies the differential inclusion (102), at every point where the derivative exists.

Lemma 6.5 (see Lemma 5.5, Chapter III, in [24]) *Let the assumptions of Theorem 4.2 be satisfied. Let $\{w^r\}$ and $\{Z^r\}$ be as in Lemma 6.3. Then*

$$(\dot{w}^*(s), \dot{Z}^*(s)) \in F(w^*(s)) \tag{102}$$

for a.e. $s \in [0,t]$.

Proof: See the mentioned lemma, recalling that in our situation $F(x)$ is convex for every x, and apply the previous crucial Lemma 6.4. $\square$

The next and last lemma proves essentially a selection property starting from the inclusion (102) and allows us to construct the optimal control α^*. More precisely

Lemma 6.6 (see Lemma 5.6 , Chapter III, in [24]) *Let the assumptions of Theorem 4.2 be satisfied. Then there exist an integrable function α^* and a measurable, non negative function v^* such that $\dot{w}^*(s) = f^{\mathbf{G}}(w^*(s), \alpha^*(s))$ and $\dot{Z}^*(s) = \psi(\alpha^*(s)) + v^*(s)$ for almost every $s \in [0,t]$.*

Proof: See the mentioned lemma. $\square$

We are now in the position to prove the existence of the optimal control for (38).

Proof of Theorem 4.2: Let α^* be as in Lemma 6.6; it is measurable. The same lemma guarantees that the function w^* is indeed the associated trajectory to the dynamics. Lemma 6.2 provides us the information that

$$w^*(0) = \lim_{r\to\infty} w^r(0) = e, \quad \text{and} \quad w^*(t) = \lim_{r\to\infty} w^r(t) = x.$$

Hence $\alpha^* \in \mathcal{F}^{\mathbf{G}}_{e,0,x,t}(\mathbb{R}^m)$. By the previous lemmata

$$\begin{aligned}
\inf_{\alpha\in\mathcal{F}^{\mathbf{G}}_{e,0,t,x}(\mathbb{R}^m)} J_{e,0,x,t}(\alpha) &= \lim_{r\to\infty} J_{e,0,x,t}(\alpha^r)\\
\text{(see (99))} \quad &= \lim_{r\to\infty} (Z^r(t) - Z^r(0))\\
\text{(see (100))} \quad &= \lim_{r\to\infty} (\mathcal{Z}^r(t) - \mathcal{Z}^r(0) - l_0 t)\\
\text{(see (101))} \quad &= \mathcal{Z}^*(t) - \mathcal{Z}^*(0) - l_0 t\\
\text{(since } \mathcal{Z}^* \text{ is monotone)} \quad &= \int_0^t \dot{\mathcal{Z}}^*(s)ds - l_0 t\\
\text{(see definition of } Z^*) \quad &= \int_0^t \dot{Z}^*(s)ds\\
\text{(see Lemma 6.6)} \quad &= \int_0^t [\psi(\alpha^*(s)) + v^*(s)]ds\\
\text{(since } v^* \geq 0) \quad &\geq \int_0^t \psi(\alpha^*(s))ds\\
&= J_{e,0,x,t}(\alpha^*).
\end{aligned}$$

Hence α^* is optimal. $\square$

References

[1] A. Agrachev and P.W.Y. Lee. Continuity of optimal control costs and its application to weak KAM theory. *Calc. Var. Partial Differential Equations*, 39(1-2):213–232, 2010.

[2] A. Agrachev and Y. Sachkov. *Control Theory from the Geometric Viewpoint.* Springer, 2004.

[3] L. Ambrosio, N. Gigli, and G. Savaré. Calculus and heat flow on metric measure spaces and applications to spaces with Ricci curvature bounded below. *Preprint*, 2011.

[4] L. Ambrosio, N. Gigli, and G. Savaré. Density of lipschitz functions and equivalence of weak gradients in metric measure spaces. *Preprint*, 2011.

[5] L. Ambrosio and S. Di Marino. Equivalent definitions of BV space and total variation in metric measure spaces. *Preprint*, 2012.

[6] Z.M. Balogh, A. Engulatov, L. Hunziker, and O.E. Maasalo. Functional inequalities and Hamilton-Jacobi equations in geodesic spaces. *Potential Anal.*, 36:317–337, 2012.

[7] Z.M. Balogh and M. Rickly. Regularity of convex functions on Heisenberg groups. *Ann. Scuola Norm. Sup. Pisa Cl. Sci*, 2:847–868, 2003.

[8] M. Bardi and I. Capuzzo Dolcetta. *Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations.* Birkhäuser, Boston, 1997.

[9] M. Bardi and L.C. Evans. On Hopf's formulas for solutions of Hamilton–Jacobi equations. *Nonlinear Anal.*, 8:1373–1381, 1984.

[10] A. Bellaïche and J.-J. Risler (eds.). *Sub–Riemannian Geometry*, volume 144. Progress in Mathematics, Birkhäuser, 1996.

[11] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni. *Stratified Lie Groups and Potential Theory for their Sub–Laplacians.* Springer, 2007.

[12] A. Bressan and B. Piccoli. *Introduction to Mathematical Theory of Control.* American Institute of Mathematical Sciences, 1975.

[13] A. Calogero and R. Pini. Note on the Fenchel transform in the Heisenberg group. *J. Math. Anal. Appl.*, 1:69–79, 2010.

[14] A. Calogero and R. Pini. Horizontal normal map on the Heisenberg group. *J. Nonlinear Convex Anal.*, 12(2):287–307, 2011.

[15] A. Calogero and R. Pini. c horizontal convexity on Carnot groups. *J. Convex Anal.*, 19:541–567, 2012.

[16] P. Cannarsa and C. Sineastri. *Semiconcave Function, Hamilton–Jacobi Equation, and Optimal Control.* Birkhäuser, Boston, MA, 2004.

[17] L. Cesari. Existence theorems for weak and usual optimal solutions in Lagrange problems with unilateral constraints I. *Trans. Amer. Math. Soc.*, 124:369–412, 1966.

[18] D. Danielli, N. Garofalo, and D.M. Nhieu. Notions of convexity in Carnot groups. *Comm. Anal. Geom.*, 11:263–341, 2003.

[19] D. Danielli, N. Garofalo, and S. Salsa. Variational inequalities with lack of ellipticity. I. Optimal interior regularity and non–degeneracy of the free boundary. *Indiana Univ. Math. J.*, 52:361–398, 2003.

[20] F. Dragoni. Metric Hopf–Lax formula with semicontinuous data. *Discrete Contin. Dyn. Syst.*, 4:713–729, 2007.

[21] L.C. Evans. *An Introduction to Mathematical Optimal Control Theory.* available in the web site of Prof. L.C. Evans (Berkley University).

[22] L.C. Evans. *Partial Differential Equations.* American Mathematical Society, 2010.

[23] L.C. Evans and P.E. Souganidis. Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations. *Indiana Univ. Math. J.*, 33(5):773–797, 1984.

[24] W.H. Fleming and R.W. Rishel. *Deterministic and Stochastic Optimal Control.* Springer–Verlag, 1975.

[25] G.B. Folland and E.M. Stein. *Hardy Spaces on Homogeneous Groups.* Princeton University Press, 1982.

[26] E. Hopf. The partial differential equation $u_t + uu_x = \mu u_{xx}$. *Comm. Pure Appl. Math.*, 3:201–230, 1950.

[27] P.D. Lax. Hyperbolic systems of conservation laws II. *Comm. pure appl. Math.*, 10:537–566, 1957.

[28] J. Lott and C. Villani. Hamilton–Jacobi semigroup on length spaces and applications. *J. Math. Pures Appl.*, 88:219–229, 2007.

[29] J.J. Manfredi and B. Stroffolini. A version of the Hopf–Lax formula in the Heisenberg group. *Communications in Partial Differential Equations*, 27:1139–1159, 2002.

[30] M. Rickly. First order regularity of convex functions on Carnot groups. *J. Geom. Anal.*, 16:679–702, 2006.

[31] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, 1969.