Journal of Convex Analysis
Volume 19 (2012), No. 2, 541–567

c Horizontal Convexity on Carnot Groups

A. Calogero
*Dipartimento di Statistica, Università degli Studi di Milano Bicocca,
Via Bicocca degli Arcimboldi 8, 20126 Milano, Italy
andrea.calogero@unimib.it*

R. Pini
*Dipartimento di Statistica, Università degli Studi di Milano Bicocca,
Via Bicocca degli Arcimboldi 8, 20126 Milano, Italy
rita.pini@unimib.it*

Received: May 03, 2010

Given a real–valued function c defined on the cartesian product of a generic Carnot group $\mathbf{G}$ and the first layer V_1 of its Lie algebra, we introduce a notion of c horizontal convex (c H–convex) function on $\mathbf{G}$ as the supremum of a suitable family of affine functions; this family is defined pointwisely, and depends strictly on the horizontal structure of the group. This abstract approach provides c H–convex functions that, under appropriate assumptions on c, are characterized by the nonemptiness of the c H–subdifferential and, above all, are locally H–semiconvex, thereby admitting horizontal derivatives almost everywhere. It is noteworthy that such functions can be recovered via a Rockafellar technique, starting from a suitable notion of c H–cyclic monotonicity for maps. In the particular case where $c(g,v) = \langle \xi_1(g), v\rangle$, we obtain the well–known weakly H–convex functions introduced by Danielli, Garofalo and Nhieu. Finally, we suggest a possible application to optimal mass transportation.

Keywords: Carnot group, horizontal convexity, c horizontal convexity, c horizontal differential, c horizontal cyclic monotonicity

1991 Mathematics Subject Classification: Primary 52A01; Secondary 22E25

1. Introduction

In $\mathbb{R}^n$ and, more generally, in a Banach space $\mathbf{X}$, the notion of convexity of a function $f : \mathbf{X} \to [-\infty, +\infty]$ can be given in terms of the pointwise supremum of the affine functions $x \mapsto \langle x, y\rangle + \alpha$ lying below the function itself. Among the nice properties enjoyed by proper, lower semicontinuous and convex functions, we recall that they can be characterized by the nonemptiness of the subdifferential ∂f at every point of their domain and, by a well–known result due to Rockafellar, they can be completely recovered by their subgradients. In addition, the multivalued map $x \mapsto \partial f(x)$ benefits from an interesting condition, since it can be characterized as a maximal monotone map [18].

This abstract formulation of convexity is fit for an extension by substituting the affine function $\langle \cdot, y\rangle + \alpha$ with the more general function $c(\cdot, y) + \alpha$, where $c : \mathbf{X} \times \mathbf{X}^* \to \mathbb{R}$. This generalization leads to the definition of c convex function, and the associ-

ISSN 0944-6532 / \$ 2.50 © Heldermann Verlag

ated c subdifferential multivalued map $x \mapsto \partial_c f(x)$ arises in a natural way, together with the notion of c cyclic monotonicity. These concepts date back to a first paper by E. J. Balder [3], and were introduced in order to extend the duality theory to nonconvex optimization problems. In this framework, H. Dietrich [10] investigated several properties of c subdifferentiability and local c subdifferentiability of c convex functions. Subsequently, L. Rüschendorf [19], in connection with the coupling problem, gave a characterization of optimal solutions via generalized subgradients of c convex functions.

As a matter of fact, as can be found in the fundamental paper by W. Gangbo and R. J. McCann [12], a context where c concavity plays a key role is in finding the solution of an optimal mass transportation problem, where $c(x, y)$ denotes the cost per unit mass displaced from x to y ($x, y \in \mathbb{R}^d$). Indeed, the support of the optimal measure on $\mathbb{R}^d \times \mathbb{R}^d$ is contained in the graph of $x \mapsto \partial^c \psi(x)$, where ψ is a c concave function called potential; if $c(x, y) = h(x - y)$, then, under suitable regularity assumptions on h, like strict convexity and superlinearity, a deep result says that this multivalued map $\partial^c \psi$ is essentially single–valued. A remarkable result concerns the finiteness and the regularity of a c concave function; it is noteworthy that a c concave function inherits structure and smoothness from the function c, like locally Lipschitz, local semiconcavity, and local boundedness. This implies that there exists an optimal transport map s defined on $\mathrm{dom}(\nabla\psi) \subset \mathbb{R}^d$ by the formula $s(x) = x - (\nabla h)^{-1}(\nabla\psi(x))$.

In the quite recent literature, in the Heisenberg group $\mathbb{H}$ and, more generally, in Carnot groups, several concepts of convexity have been introduced (see, for instance, [8], or [7]). Among them, the most suitable to many purposes is the so–called weakly H–convexity (H–convexity, in the sequel). An H–convex function u is, essentially, a function that is convex along any horizontal line, a particular horizontal curve. Balogh and Rickly proved that these functions are regular enough, since they are locally Lipschitz continuous with respect to any homogeneous distance (see [4], [14]). In [6], we show that, for real–valued functions on $\mathbb{H}$, H–convexity is equivalent to H–subdifferentiability, i.e. the horizontal subdifferential is nonempty at every point of the domain; in this paper this result will be extended to a generic Carnot group.

Unexpectedly, it turns out (see [5]) that there is an abstract definition in $\mathbb{H}$ of convexity, given in terms of H–affine functions, that is equivalent to the H–convexity. A real–valued function u is (abstractly) H–convex if

$$u(g) = \sup_{(v,\alpha)\in\mathcal{P}_g} (\langle \xi_1(g), v\rangle + \alpha)$$

where ξ_1 is defined in Section 2, and $\mathcal{P}_g$ is the set of pairs $(v, \alpha) \in V_1 \times \mathbb{R}$ such that the H–affine function $g' \mapsto \langle \xi_1(g'), v\rangle + \alpha$ supports u on every horizontal line through g (see [8], p. 320). Let us stress the peculiarity of the set of parameters $(v, \alpha) \in V_1 \times \mathbb{R}$, that depends on the point g. Moreover, let us notice that v belongs to the first layer of the Lie algebra of $\mathbf{G}$ and plays the role of a "subgradient".

This point of view can be extended, by taking a general function $c(g, v)$ instead of $\langle \xi_1(g), v\rangle$; this paper is devoted to the study of the main features of these functions, that will be called c (horizontally) convex (c H–convex, briefly), as well as to the interrelationships with their c H–subdifferentials. We cannot leave unmentioned the papers by L. Ambrosio and S. Rigot [1], [15], where the optimal transport mass is investigated in the framework of particular Carnot groups; in this context they introduce a different notion of c concavity and c superdifferential that does not take into account the "horizontal" structure of a Carnot group.

Our investigation follows some classical steps in convex analysis. The main results mirror similar ones in the context of classical convexity. First of all, in Section 3, we provide the mentioned link between H–convexity of a real–valued function and H–subdifferentiability in a generic Carnot group $\mathbf{G}$ (see Theorem 3.4). In Section 4, we introduce the notions of c H–convexity and c H–subdifferential $\partial_H^c u$ for a proper function $u : \mathbf{G} \to (-\infty, +\infty]$. Theorem 4.5 characterizes real-c H–convexity of a function via its c H–subdifferentiability; this result is not an extension of Theorem 3.4, since it deals with abstract H–convexity.

Section 5 is devoted to the problem of the regularity of a proper c H–convex function, and we try to establish the almost sure single–valuedness of its c H–subdifferential. In Euclidean spaces, semiconvexity turns out to be a fundamental tool for the study of c convex functions (see, for instance, [20], Chapter 10). Semiconvexity can be extended in a natural way in a Carnot group starting from H–convexity, and it gives rise to the notion of H–semiconvexity (see Definition 5.1). Despite their abstract and entangled definition, c H–convex functions prove to be well–behaved whenever c is. One of the most interesting result of the paper is Theorem 5.4: we show that, like in the classical case, our functions are locally H–semiconvex, and therefore they share the regularity of the H–convex functions. This entails that, in the real–valued case, and under measurability assumptions if the step is greater than 2, a c H–convex function u is differentiable almost everywhere along the horizontal directions; furthermore, we get that $\partial_H^c u(g)$ is a singleton for almost every g.

Another relevant issue that shares its aim with classical convexity concerns the connection between c H–convexity of a function on $\mathbf{G}$, and the c H–cyclic monotonicity of a subset of $\mathbf{G} \times V_1$ (see Definition 6.4). The main and more delicate outcome of Section 4 shows that, from a c H–cyclically monotone set and via Rockafellar techniques, it is possible to construct, at least locally, a c H–convex function u such that the graph of $g \mapsto \partial_H^c u(g)$ contains the starting set. In this setting, the analysis of the finiteness of the function plays a critical role.

Finally, inspired by the precious paper by Gangbo and McCann [12], we present a possible application of all these arguments and tools to the optimal mass transportation problem in the Heisenberg group. Despite this application arises in a very particular situation, where the optimal map moves the points only along horizontal segments, we think that our approach could be potentially interesting.

2. Basic notions on Carnot groups

A Carnot group $\mathbf{G}$ 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

$$\mathfrak{g} = V_1 \oplus V_2 \oplus \ldots \oplus V_r,$$
$$[V_1, V_j] = V_{j+1}, \quad j = 1, \ldots r-1,$$
$$[V_1, V_r] = 0.$$

We assume that a scalar product $\langle \cdot, \cdot \rangle$ is given on $\mathfrak{g}$ for which the levels V_j are mutually orthogonal. The first layer V_1 of the Lie algebra plays a key role: we call *horizontal vector fields* its elements, and denote by m its dimension.

We fix an orthonormal basis $X = \{X_1, X_2, \ldots, X_m\}$ of V_1, and we continue to denote by X the corresponding system of left–invariant vector fields on $\mathbf{G}$ defined by $X_i(g) = (L_g)_*(X_i)$, $i = 1, \ldots, m$, where $(L_g)_*$ is the differential of the left translation on $\mathbf{G}$ defined by $L_g(g') = gg'$. The system X defines a basis for the horizontal sub–bundle $\mathcal{H}\mathbf{G}$ of the tangent bundle $\mathcal{T}\mathbf{G}$ (i.e. $\mathcal{H}_g\mathbf{G} = \mathrm{span}\{X_1(g), \ldots, X_m(g)\}$ for every $g \in \mathbf{G}$).

The action of X_i on a function $u : \mathbf{G} \to \mathbb{R}$ is given by

$$X_i u(g) = \lim_{\alpha \to 0} \frac{u(g \exp(\alpha X_i)) - u(g)}{\alpha}.$$

Clearly, $\exp : \mathfrak{g} \to \mathbf{G}$ is the exponential map, 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 $\Delta_\lambda(v_1 + v_2 + \ldots + v_r) = \lambda v_1 + \lambda^2 v_2 + \ldots + \lambda^r v_r$, if $v_i \in V_i$ for every $1 \leq i \leq r$. By means of the exponential map, one lifts these dilations to the family of the automorphisms $\delta_\lambda(g) = \exp(\Delta_\lambda(\xi(g)))$. The homogeneous dimension associated with the dilations $\{\delta_\lambda\}_{\lambda>0}$ is given by $Q = \sum_{i=1}^r i \dim V_i$ that often replaces the topological dimension $N = \sum_{i=1}^r \dim V_i$ in the study of Carnot groups.

The Euclidean distance to the origin $|\cdot|_\mathfrak{g}$ on $\mathfrak{g}$ induces a homogeneous pseudo–norm $\|\cdot\|_\mathfrak{g}$ on $\mathfrak{g}$ defined by $\|v_1 + v_2 + \ldots + v_r\|_\mathfrak{g} = \left(\sum_{i=1}^r |v_i|_\mathfrak{g}^{2r!/i}\right)^{2r!}$. Again, via the exponential map, we lift $\|\cdot\|_\mathfrak{g}$ to a pseudo–norm $\|\cdot\|_\mathbf{G}$, and hence to a pseudo–distance d on $\mathbf{G}$ defining $\|g\|_\mathbf{G} = \|\xi(g)\|_\mathfrak{g}$ and $d(g, g') = \|g^{-1}g'\|_\mathbf{G}$.

Let $\Omega \subset \mathbf{G}$ be an open set, k be a nonnegative integer, and $0 < \alpha \leq 1$. The class $\Gamma^k(\Omega)$ represents the Folland–Stein space of functions having continuous derivatives up to the order k with respect to the horizontal vector fields $X_1, \ldots, X_m$. A function $u : \Omega \to \mathbb{R}$ is said to belong to the class $\Gamma^{0,\alpha}(\Omega)$ if there exists a positive constant C_α such that

$$|u(g) - u(g')| \leq C_\alpha d(g, g')^\alpha,$$

for every g and g' in Ω. A function $f \in \Gamma^1(\Omega)$ belongs to the class $\Gamma^{1,\alpha}(\Omega)$ if, for every $i = 1, \ldots, m$, the horizontal derivative $X_i f$ exists in Ω and $X_i f \in \Gamma^{0,\alpha}(\Omega)$. As usual, we say that u is Lipschitz continuous if $u \in \Gamma^{0,1}$; the symbol $\Gamma^{0,1}_{\mathrm{loc}}(\Omega)$ denotes the class of locally Lipschitz continuous functions on Ω.

Let us recall that the horizontal gradient of a function $u \in \Gamma^1(\Omega)$ at $g \in \Omega$ is the element of V_1

$$\mathbb{X}u(g) = \sum_{i=1}^{m} (X_i u(g)) X_i.$$

The *horizontal plane* H_g associated to $g \in \mathbf{G}$ is given by

$$H_g = L_g\left(\exp(V_1)\right) = \{g' \in \mathbf{G} : g' = gh, \text{ for some } h \in \exp(V_1)\}. \tag{1}$$

Note that $g' \in H_g$ implies that $g \in H_{g'}$ and $g^{-1}g' \in H_e$, where e is the unit element of the group $\mathbf{G}$. If we consider the set H_e, and identify $\mathbf{G}$ with $\mathbb{R}^N$ (remember that N is the topological dimension), it turns out that the set H_e is an iperplane in $\mathbb{R}^N$. Differently, if $g \neq e$, one can show that the horizontal plane H_g is an iperplane in the classical sense (in particular an $\mathbb{R}^N$–convex set, using the subsequent notation) if and only if $\mathbf{G}$ has step 2 (see Example 2.3).

As a matter of fact, the elements of the first layer V_1 of the Lie algebra $\mathfrak{g}$ generate all the vector fields of $\mathfrak{g}$ and consequently, via the exponential map, the points of the horizontal plane H_e play a similar role in $\mathbf{G}$. More precisely, the following structure result holds:

Proposition 2.1 (see [11], Lemma 1.40). *Let $\mathbf{G}$ be a stratified group. Then, there exist $C > 0$ and $R \in \mathbb{N}$ such that any $g \in \mathbf{G}$ can be expressed as $g = h_1 h_2 \cdots h_R$, with suitable $h_i \in H_e$ and $\|h_i\|_{\mathbf{G}} \leq C\|g\|_{\mathbf{G}}$, for every $i = 1, 2, \ldots, R$.*

We recall that a Lipschitz curve $\gamma : [0,T] \to \mathbf{G}$ is said to be horizontal if $\gamma'(\lambda) \in \mathcal{H}_{\gamma(\lambda)}\mathbf{G}$, i.e. $\gamma'(\lambda) = \sum_{i=1}^m a_i(\lambda) X_{\gamma(\lambda)}$, for almost every $\lambda \in [0,T]$. The sub–Riemannian length of a horizontal curve γ is

$$L(\gamma) = \int_0^T \left(\sum_{i=1}^m a_i^2(\lambda) \right)^{1/2} d\lambda;$$

the Carnot–Caratheodory distance d_{CC} from g to g' is

$$d_{CC}(g, g') = \inf\{L(\gamma) : \gamma \text{ is a horizontal curve connecting } g \text{ and } g'\}.$$

A curve γ joining g and g' is a geodesic if it is a length minimizing horizontal curve, i.e. $L(\gamma) = d_{CC}(g, g')$. Another kind of curve connecting two points g and g' arises via their twisted convex combination $\sigma_{g,g'}$ defined by

$$\sigma_{g,g'}(\lambda) = g\delta_\lambda(g^{-1}g'), \quad \lambda \in [0,1]. \tag{2}$$

If $g' \in H_g$, we say that $\sigma_{g,g'}$ is a *horizontal segment*; it is a horizontal curve and, in particular, a geodesic.

We say that $A \subset \mathbb{R}^n$ is $\mathbb{R}^n$–convex if $(1-\lambda)x + \lambda y \in A$, for every x, y in A, and $\lambda \in [0,1]$. Consequently, a function is $\mathbb{R}^n$–convex if $u((1-\lambda)x+\lambda y) \leq (1-\lambda)u(x)+\lambda u(y)$, with x, y, λ as before. An $\mathbb{R}^n$–segment is the $\mathbb{R}^n$–convex hull of two points, and an $\mathbb{R}^n$–plane is the set $\{x \in \mathbb{R}^n : \langle x, a\rangle = b\}$, for some fixed $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$. These notations should be pedant, but it is important in this paper to distinguish the different notions of convexity, plane, segment that we introduce.

Let us explain these arguments with two basic models.

Example 2.2. The Heisenberg group $\mathbb{H}$.

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

$$\begin{gathered} V_1 = \operatorname{span}\{X_1, X_2\}, \qquad X_1 = \partial_x - \tfrac{y}{2}\partial_t, \qquad X_2 = \partial_y + \tfrac{x}{2}\partial_t, \\ V_2 = \operatorname{span}\{T\}, \qquad T = \partial_t. \end{gathered} \tag{3}$$

The bracket $[\cdot,\cdot] : \mathfrak{h} \times \mathfrak{h} \to \mathfrak{h}$ is defined as $[X_1, X_2] = T$, and it vanishes in the other cases; taking into account the action of the bracket, $X * Y$ is defined by the Baker–Campbell–Dynkin–Hausdorff formula

$$X * Y = X + Y + [X, Y]/2. \tag{4}$$

The exponential map $\exp(\alpha X_1 + \beta X_2 + \gamma T) = (\alpha, \beta, \gamma)$ enjoys the property $\exp(X)\exp(Y) = \exp(X * Y)$, for every X and Y in $\mathfrak{g}$; consequently, the law group on $\mathbb{H}$ is

$$gg' = (x, y, t)(x', y', t') = (x + x', y + y', t + t' + (xy' - x'y)/2).$$

The dilation is a family of automorphisms given by $\delta_\lambda(x, y, t) = (\lambda x, \lambda y, \lambda^2 t)$, and hence the homogeneous dimension is 4. Given two points $g = (x, y, t)$ and $g' = (x', y', t')$, the noncommutative twisted convex combination in (2) is

$$\begin{aligned} \sigma_{g,g'}(\lambda) = \ & \big((1-\lambda)x + \lambda x', (1-\lambda)y + \lambda y', \\ & \quad t + \lambda(xy' - x'y)/2 + \lambda^2(t' - t + (x'y - xy')/2)\big). \end{aligned}$$

The horizontal plane H_g is, by (1),

$$H_g = \{(x', y', t') \in \mathbb{H} : \ t' = t + (xy' - x'y)/2, \quad x', y' \in \mathbb{R}\};$$

it is a "real" plane, i.e. an $\mathbb{R}^3$–plane. If we choose g' on the horizontal plane H_g, the curve $\sigma_{g,g'}$ is a horizontal curve and a geodesic that we call, by definition, horizontal segment γ from g to g' : more precisely,

$$\gamma(\lambda) = ((1-\lambda)x + \lambda x', (1-\lambda)y + \lambda y', t + \lambda(xy' - x'y)/2).$$

Note that γ is an $\mathbb{R}^3$–segment lying in $H_g \cap H_{g'}$.

Example 2.3. The Engel group $\mathbb{E}$.

The Engel group $\mathbb{E}$ is a Carnot group of step 3 and, in some sense, is an extension of $\mathbb{H}$: indeed, the Lie algebra $\mathfrak{e} = \mathbb{R}^4 = V_1 \oplus V_2 \oplus V_3$ is defined, using (3), by

$$\begin{gathered} V_1 = \operatorname{span}\{\tilde{X}_1, \tilde{X}_2\} \ \text{ with } \tilde{X}_1 = X_1 - (\tfrac{t}{2} + \tfrac{xy}{12})\partial_s \text{ and } \tilde{X}_2 = X_2 + \tfrac{x^2}{12}\partial_s, \\ V_2 = \operatorname{span}\{\tilde{T}\} \ \text{ with } \tilde{T} = T + \tfrac{x}{2}\partial_s, \\ V_3 = \operatorname{span}\{\tilde{S}\} \ \text{ with } \tilde{S} = \partial_s. \end{gathered}$$

The bracket acts as $[\tilde{X}_1, \tilde{X}_2] = \tilde{T}$, $[\tilde{X}_1, \tilde{T}] = \tilde{S}$, and it vanishes in the other cases. Since, in $\mathfrak{e}$, in the Baker–Campbell–Dynkin–Hausdorff formula (4) there is one more term (precisely $([X,[X,Y]]+[Y,[Y,X]])/12$) and $\exp(\alpha\tilde{X}_1+\beta\tilde{X}_2+\gamma\tilde{T}+\eta\tilde{S}) = (\alpha,\beta,\gamma,\eta)$, the group law in $\mathbb{E}$ becomes

$$gg' = \Big(x+x', y+y', t+t'+(xy'-x'y)/2, s+s'+(xt'-x't)/2+(x-x')(xy'-yx')/12\Big),$$

where $g=(x,y,t,s)$ and $g'=(x',y',t',s')$. The horizontal plane H_g is

$$H_g = \Big\{(x',y',t',s')\in\mathbb{E} :\ t' = t+(xy'-x'y)/2,$$
$$s' = s+(-6t(x'-x)+2x^2y'-2x'xy+yx'^2-xx'y')/12, \text{ with } x,y\in\mathbb{R}\Big\};$$

note that H_g is not an $\mathbb{R}^4$–plane. Clearly, the dilation is given by $\delta_\lambda(x,y,t,s) = (\lambda x, \lambda y, \lambda^2 t, \lambda^3 s)$. If we consider $g'\in H_g$, the horizontal segment γ with endpoints g and g' is defined via (2): γ is a geodesic, lies in $H_g\cap H_{g'}$ and, in general, is not an $\mathbb{R}^4$–segment.

We have seen that, unlike the Euclidean spaces, where the Euclidean distance is the most natural choice, in a Carnot group several distances were introduced for different purposes. However, all these distances ρ are homogeneous, namely, they are left invariant and satisfy the relation $\rho(\delta_r g',\delta_r g) = r\rho(g',g)$ for every $g',g\in\mathbf{G}$, and $r>0$. The distance functions d and d_{CC} are homogeneous, equivalent, and have the same value at the endpoints of a horizontal segment.

Let ρ be any homogeneous distance on $\mathbf{G}$, and let $u:\mathbf{G}\to\mathbb{R}$. We say that u is Pansu differentiable at $g\in\mathbf{G}$ if there exists a $\mathbf{G}$–linear map $Du(g):\mathbf{G}\to\mathbb{R}$, i.e., a group homomorphism that satisfies the relation $Du(g)(\delta_r h) = rDu(g)(h)$ for every $h\in\mathbf{G}$ and $r>0$, such that

$$\lim_{\rho(h,e)\to 0}\frac{|u(gh)-u(g)-Du(g)(h)|}{\rho(h,e)} = 0.$$

We call the map $Du(g)$ the Pansu differential of u at g. An easy computation gives us that if u is Pansu differentiable at g, then

$$Du(g)(h) = \lim_{\lambda\to 0^+}\frac{u(g\delta_\lambda(h))-u(g)}{\lambda}$$

exists for every $h\in\mathbf{G}$. If $u\in\Gamma^1(\mathbf{G})$, then the Pansu differential $Du(g)$ is given by the formula

$$Du(g)(h) = \langle \mathbb{X}u(g), \xi_1(h)\rangle,$$

for every g and h in $\mathbf{G}$ (see [8]).

It is known that a Rademacher–Stefanov type result holds in the Carnot group setting; therefore, a Lipschitz continuous function is differentiable almost everywhere in the horizontal directions. A further result, due to Danielli, Garofalo and Salsa, will play a crucial role in the sequel:

Theorem 2.4 (see [9], Theorem 2.7). *Let Ω be an open subset of $\mathbf{G}$, and $u : \Omega \to \mathbb{R}$, with $u \in \Gamma^{0,1}(\Omega)$. Then there exists a set $E \subset \Omega$ with Haar measure zero such that the Pansu differential $Du(g)$ and the horizontal gradient $\mathbb{X}u(g)$ exist for every $g \in \Omega \setminus E$, and*

$$Du(g)(h) = \langle \mathbb{X}u(g), \xi_1(h) \rangle, \quad \textit{for every } h \in \mathbf{G}.$$

Furthermore, $\mathbb{X}u \in L^\infty(\Omega)$.

Finally, for what concerns classical convex analysis, we will refer to [17]. In particular, we say that a function u defined on a subset Ω of $\mathbf{G}$ is *proper* if $u(g) \neq -\infty$ for every $g \in \Omega$, and $u \not\equiv +\infty$; moreover, if $u(g) \neq +\infty$ for every $g \in \Omega$, then we say that u is *real–valued*. The domain of u, $\mathrm{dom}(u)$, is the subset of Ω where u is finite.

3. H–convexity and H–subdifferentiability

In the last few years, several notions of convexity have been introduced in the framework of Carnot groups, but the most suitable one showed to be the notion of H–convexity. This notion is due to Caffarelli in unpublished works from 1996, and it appeared in the paper [8]; afterwards, several papers have been devoted to the investigations of H–convexity. Among other things, the horizontal Monge–Ampère equation in $\mathbf{G}$ defined by $\det[\mathbb{X}^2u]^*(g) = f(g, u, \mathbb{X}u)$ is (degenerate) elliptic precisely on the class of $u \in \Gamma^2(\mathbf{G})$ which are H–convex (see below for the definition of $[\mathbb{X}^2u]^*$). This section will concern results about H–convex functions in a Carnot group $\mathbf{G}$.

A subset Ω of $\mathbf{G}$ is *H–convex* if it contains every horizontal segment with endpoints in Ω, i.e. $g\delta_\lambda(g^{-1}g') \in \Omega$, for every $g \in \Omega$, $g' \in H_g \cap \Omega$ and $\lambda \in [0, 1]$.

Definition 3.1. Let $\Omega \subset \mathbf{G}$ be H–convex. A function $u : \Omega \to (-\infty, +\infty]$ is *H–convex* if it is $\mathbb{R}$–convex on every horizontal segment, i.e.

$$u(g\delta_\lambda(g^{-1}g')) \leq (1 - \lambda)u(g) + \lambda u(g') \tag{5}$$

for all $g \in \Omega$, $g' \in H_g \cap \Omega$, and $\lambda \in [0, 1]$.

A function $u : \Omega \subset \mathbf{G} \to [-\infty, +\infty)$ is said to be H–concave if $-u$ is H–convex.

It is clear, by the definition, that an $\mathbb{R}^3$–convex function u is H–convex in the Heisenberg group, since every horizontal segment is a particular $\mathbb{R}^3$–segment. This argument can be extended to any Carnot group of step two. On the contrary, if one consider a group $\mathbf{G}$ of step greater than 2 this is no longer true. An enlightening example can be found in [14]: the function $u : \mathbb{E} \to \mathbb{R}$, $u(x, y, t, s) = s$ is not H–convex in the Engel group $\mathbb{E}$, despite it is $\mathbb{R}^4$–convex.

In spite of the notion of H–convexity, that requires a suitable behaviour on the horizontal planes only, H–convex functions enjoy some nice regularity properties. Balogh and Rickly (see [4] if $\mathbf{G} = \mathbb{H}$, and [14]) proved the following result:

Theorem 3.2 (see [14], Theorem 1.4). *Let $\Omega \subset \mathbf{G}$ be an H–convex, open subset. Then every H–convex function $u : \Omega \to \mathbb{R}$, measurable if the step of $\mathbf{G}$ is greater than 2, belongs to $\Gamma^{0,1}_{\mathrm{loc}}(\Omega)$.*

The possibility to remove the measurability assumption in the previous result is an interesting and open question.

In [8], a regular function $u : \Omega \to \mathbb{R}$, where Ω is an open and H–convex subset of $\mathbf{G}$, is characterized in terms of its horizontal gradient $\mathbb{X}u$, and its symmetrized horizontal Hessian $[\mathbb{X}^2 u]^*$. Indeed, if $u \in \Gamma^1(\Omega)$, then u is an H–convex function if and only if

$$u(g') \geq u(g) + \langle \mathbb{X}u(g), \xi_1(g') - \xi_1(g)\rangle, \quad \forall g \in \Omega,\ \forall g' \in H_g \cap \Omega; \tag{6}$$

if $u \in \Gamma^2(\Omega)$, then u is H–convex if and only if $[\mathbb{X}^2 u]^*(g)$ is positive semidefinite for every $g \in \Omega$, where

$$[\mathbb{X}^2 u]^*(g) = \frac{1}{2}\{\mathbb{X}^2 u(g) + \mathbb{X}^2 u(g)^T\}$$

and $\mathbb{X}^2 u(g) = (X_i X_j u(g))_{i,j=1,\dots,m}$ is an $m \times m$ matrix.

It is well known that if $f : \mathbb{R}^n \to \mathbb{R}$ is a differentiable function, then the $\mathbb{R}^n$–convexity of f can be characterized by the monotonicity of the gradient, i.e., $(\nabla f(x) - \nabla f(y)) \cdot (x - y) \geq 0$, for every x, y in the domain (see, for instance, [2], Theorem 2.13). This result can be adapted to the sub–Riemannian setting; indeed, if $u \in \Gamma^1(\Omega)$, then one can easily show from (6) that u is an H–convex function if and only if

$$\langle \mathbb{X}u(g) - \mathbb{X}u(g'), \xi_1(g) - \xi_1(g')\rangle \geq 0, \quad \forall g \in \Omega,\ \forall g' \in H_g \cap \Omega. \tag{7}$$

We say that the set $\{g_i\}_{i=0}^n \subset \Omega$ is an *H–sequence* if, for some $n > 0$ and for every $i = 0, \dots, n-1$, $g_{i+1} \in H_{g_i}$. An H–sequence is closed if $g_0 \in H_{g_n}$; in this case, we usually set $g_{n+1} = g_0$. This notion, that will be fundamental in the next sections, allows us to extend the characterization in (7). Indeed, an easy calculation shows that, when $u \in \Gamma^1(\Omega)$, then u is an H–convex function if and only if

$$\sum_{i=0}^{n} \langle \mathbb{X}u(g_i), \xi_1(g_{i+1})\rangle \leq \sum_{i=0}^{n} \langle \mathbb{X}u(g_i), \xi_1(g_i)\rangle, \tag{8}$$

for every closed H–sequence $\{g_i\}_{i=0}^n \subset \Omega$. This last property will lead to consider the more general notion of c H–cyclic monotonicity in Section 6.

In [8] the authors relate the property of H–convexity of a real–valued function to the nonemptyness of its H–subdifferential. Let us recall that the H–subdifferential of a function $u : \Omega \subset \mathbf{G} \to (-\infty, +\infty]$ at $g_0 \in \Omega$ is defined as

$$\partial_H u(g_0) = \{p \in V_1 :\ u(g) \geq u(g_0) + \langle p, \xi_1(g) - \xi_1(g_0)\rangle,\ \forall g \in H_{g_0} \cap \Omega\}.$$

Moreover, we say that $\partial^H u(g_0)$ is the H–superdifferential of u at g_0 if $\partial^H u(g_0) = -\partial_H(-u)(g_0)$.

A first link between H–subdifferentiability of a function and H–convexity is provided by the following:

Proposition 3.3 (see [8], Proposition 10.5). *Let $u : \Omega \to \mathbb{R}$, where Ω is an open and H–convex subset of $\mathbf{G}$. If $\partial_H u(g) \neq \emptyset$ for every $g \in \Omega$, then u is H–convex.*

The converse of this result, as in the classical case, is more difficult. In [6] we prove that this holds when $\mathbf{G} = \mathbb{H}$. As a matter of fact, next theorem shows that the result can be improved.

Theorem 3.4. *Let $u : \Omega \subset \mathbf{G} \to \mathbb{R}$, where Ω is open and H–convex. Let u be H–convex, and measurable if $r > 2$. Then $\partial_H u(g) \neq \emptyset$ for every $g \in \Omega$.*

In order to prove Theorem 3.4, one can extend in a natural way the proof in [6] from the Heisenberg group to a generic Carnot group, with the additional assumption that u is measurable if $r > 2$. We recall the main tools of such proof and we leave its details to the reader. In this setting, a fundamental role is played by the regularity results for H–convex functions due to Balogh and Rickly (see Theorem 3.2), and the differentiability almost everywhere in the horizontal directions for Lipschitz continuous functions due to Danielli, Garofalo and Salsa (see Theorem 2.4). The assumption of H–convexity and the two results above, together, lead to the inclusion $\mathbb{X}u(g) \in \partial_H u(g)$ a.e. in Ω. A crucial point lies in proving that the graph of the multivalued map $g \mapsto \partial_H u(g)$ is closed, i.e. for every sequence $\{(g_n, p_n)\}_n \subset \Omega \times V_1$ with $p_n \in \partial_H u(g_n)$, such that $g_n \to g_0 \in \Omega$ and $p_n \to p$, then $p \in \partial_H u(g_0)$. In order to do this, we exploit the continuity of the function u, together with the "continuity" of the left translation on the group, that is involved in the definition of the horizontal planes (1); more precisely, given $g_0 \in \Omega$, and $g' \in H_{g_0} \cap \Omega$, for every $g_n \to g_0$ there exists $\{g'_n\}_n$ such that $g'_n \in H_{g_n} \cap \Omega$ and $g'_n \to g'$. The reader can give a look at Lemma 4.1 in [6] to find more details in the case $\mathbf{G} = \mathbb{H}$.

4. *c* H–convexity and *c* H–subdifferential

The class of c convex functions was introduced, to our knowledge, by Dietrich [10], and subsequently exploited by several authors in connection with optimal couplings and optimal mass transportation problems; to get an idea about it, one can read the paper by Rüschendorf [19], or give a look at the book by C. Villani [20]. Briefly, if Ω_1, Ω_2 are two sets, and $c : \Omega_1 \times \Omega_2 \to \mathbb{R}$, then a proper function $f : \Omega_1 \to (-\infty, +\infty]$ is said to be c convex if there exists a set $\mathcal{P} \subset \Omega_2 \times \mathbb{R}$ such that

$$f(x) = \sup_{(y,\alpha)\in\mathcal{P}} (c(x,y) + \alpha), \quad \forall x \in \Omega_1. \tag{9}$$

In the investigation about the properties of c convex functions, a fundamental role is played by the notion of c subdifferential $\partial_c f$ defined as

$$\partial_c f(x) = \{y \in \Omega_2 : \ f(x') \geq f(x) + c(x', y) - c(x, y), \ \forall x' \in \Omega_1\}. \tag{10}$$

In particular cases, for instance if $\Omega_1 = \Omega_2 = \mathbb{R}^n$ and $c(x,y) = \langle x, y\rangle$, one can easily recover some classical notions: in (9) we obtain the abstract notion of convexity, where a convex function is defined as the pointwise supremum of a family of affine functions; in (10) we obtain the notion of subgradient, i.e. the set of coefficients y

such that the affine function $x' \mapsto \langle x' - x, y\rangle + f(x)$ supports the function f at the point x.

In [1], the authors deal with an optimal mass transportation problem in the Heisenberg group, and they are lead to consider the class of c convex functions on $\mathbb{H}$. In particular they prove the existence and the uniqueness of an optimal transport map assuming that the cost function $c : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$ is either the function d^2, or the function d^2_{CC} (see [15] for the more general case of groups of type H). As a matter of fact, the notion of c convexity they work with does not take into account the horizontal structure; more precisely, they say that $f : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$ is c convex if (9) holds, at every $x \in \mathbb{H}$, for a suitable nonempty set $\mathcal{P} \subset \mathbb{H} \times \mathbb{R}$. Consequently, their definition of c subdifferential is exactly as in (10), with $\Omega_1 = \Omega_2 = \mathbb{H}$.

The aim of this paper is the investigation of c convexity from another viewpoint: in Sections 4–6 we provide a different notion of c convexity and c subdifferential, having in mind the horizontal structure of Carnot groups, and we investigate their properties. First of all, we note that in the general situation a c subdifferential is an element of the space Ω_2; taking into account that the H–subdifferential is contained in the first layer V_1, we consider a “cost” function

$$c : \mathbf{G} \times V_1 \to \mathbb{R}.$$

Now, we are in the position to introduce our main definition:

Definition 4.1. We say that a proper function $u : \Omega \subset \mathbf{G} \to (-\infty, +\infty]$ is a *c H–convex function* if for every $g \in \Omega$ we have

$$u(g) = \sup_{(v,\alpha)\in\mathcal{P}_g} (c(g, v) + \alpha),$$

where $\mathcal{P}_g = \{(v, \alpha) \in V_1 \times \mathbb{R} : \; c(g', v) + \alpha \le u(g'),\ \forall g' \in H_g \cap \Omega\}$ is, for every $g \in \Omega$, a nonempty set.

Moreover, we say that u is c H–concave if $-u$ is c H–convex.

We would like to stress the difference between (9) and Definition 4.1: while, in the former case, the index set $\mathcal{P}$ is fixed, in the latter one it depends on the point g. At first sight this difference is a problem: in the classical case where $\Omega_1 = \Omega_2 = \mathbb{R}^n$ and $c(x, y) = \langle x, y\rangle$, the pointwise supremum at every point $x \in \mathbb{R}^n$ of a family of affine functions with parameters in a set $\mathcal{P}_x \subset \mathbb{R}^n \times \mathbb{R}$ depending on x, can be a non convex function. However, if we consider the case $\mathbf{G} = \mathbb{H}$ and

$$c(g, v) = \langle \xi_1(g), v\rangle, \tag{11}$$

the notion of c H–convexity corresponds to the so called “abstract H–convexity” in [5] (see, in particular, Definition 4.3); there, we proved that these functions coincide with the H–convex ones defined in the previous section, at least when they are real–valued. Indeed, the following holds:

Proposition 4.2 (see [5], Theorem 1.1). *If $u : \mathbb{H} \to \mathbb{R}$ and c is as in* (11)*, then u is c H–convex if and only if u is H–convex.*

This is one of the convictive reasons to say that our Definition 4.1 is consistent. In the sequel, we say briefly that a function is $\langle \xi_1(\cdot), \cdot \rangle$ H–convex if it is c H–convex in a generic Carnot group $\mathbf{G}$ with cost function c as in (11).

Notice that, when dealing with c H–convex functions, as well as with c convex functions, one has to face with the possible value $+\infty$; this gives rise to some difficulties when regularity properties are required. The investigation of conditions entailing the finiteness of a c H–convex function will be the topic of Proposition 6.3 and is closely connected with the nonemptiness of the c H–subdifferential.

With further regularity on c, one can hopefully find interesting results about c H–convex functions. To this purpose, in the sequel, according to the context, some assumptions will be taken into consideration:

(**c1**) for every $p \in V_1$, the function $c(\cdot, p)$ belongs to $\Gamma^{1,1}_{\mathrm{loc}}(\mathbf{G})$, with uniform Lipschitz bound on V_1;

(**c2**) let $\Omega \subset \mathbf{G}$; for every $g \in \Omega$ and for all $\{v_n\}_n \subset V_1$ with $\|v_n\|_{\mathfrak{g}} \to +\infty$, there exists $g' \in H_g \cap \Omega$ such that
$$\limsup_n (c(g', v_n) - c(g, v_n)) = +\infty;$$

(**c3**) let $\Omega \subset \mathbf{G}$; for every $g \in \Omega$, the function $\mathbb{X}c(g, \cdot) : V_1 \to V_1$ is one–to–one.

Notice that the function $c : \mathbf{G} \times V_1 \to \mathbb{R}$ defined in (11) fulfills all the properties above.

Let us spend a few words on the role that the above conditions on c will play in the sequel. The regularity of c expressed by (**c1**) will imply some regularity for any real–valued c H–convex function, like the local boundedness and the horizontal differentiability almost everywhere. Condition (**c2**), that represents a sort of horizontal superlinearity of c, will provide a link between the c H–convexity of a function and the nonemptiness of its c H–subdifferential at every point. Condition (**c3**) will be useful when dealing with the connection between the horizontal derivatives of c, of a c H–convex function u, and of its c H–subdifferential.

As in the classical setting, a concept strictly related to the c H–convexity is the following:

Definition 4.3. Let $u : \Omega \to (-\infty, +\infty]$, with $\Omega \subset \mathbf{G}$. The c *H–subdifferential* of u at $g \in \Omega$ is the (possibly empty) set
$$\partial^c_H u(g) = \{p \in V_1 : \ u(g') \geq u(g) + c(g', p) - c(g, p), \ \forall g' \in H_g \cap \Omega\}.$$

In particular, we say that u is c H–subdifferentiable at g_0 if $\partial^c_H u(g_0) \neq \emptyset$. Clearly, the $\langle \xi_1(\cdot), \cdot \rangle$ H–subdifferential of a function coincides with its H–subdifferential.

We will denote by $\partial^c_H u$ the multivalued map $g \mapsto \partial^c_H u(g)$. To this purpose, given a multivalued map $T : \mathbf{G} \to \mathcal{P}(V_1)$, we recall that its domain $\mathrm{dom}(T)$ is the set of points $g \in \mathbf{G}$ for which $T(g)$ is nonempty, and the graph of T is the set $\mathrm{graph}(T) = \{(g, v) \in \mathbf{G} \times V_1 : \ g \in \mathrm{dom}(T), v = T(g)\}$.

Remark 4.4. From the definition of $\partial^c_H u$ we easily get that, if $u(g) = +\infty$, then $\partial^c_H u(g) \neq \emptyset$ if and only if $u(g') = +\infty$ for every $g' \in H_g \cap \Omega$.

Our next aim is to establish some results rephrasing those in Proposition 3.3 and in Theorem 3.4, for the more general case of c H–convexity. Under suitable assumptions on the function c, a characterization of c H–convexity via the nonemptiness of the c H–subdifferential at every point can be given.

In order to prove next theorem, let us supply an extension of the concept of H–Fenchel transform introduced in [5]. Let $u : \Omega \to (-\infty, +\infty]$, with $\Omega \subset \mathbf{G}$. The c H–Fenchel transform of u is the family of functions $\{u_g^c\}_{g\in\Omega}$, where, for every $g \in \Omega$, $u_g^c : V_1 \to [-\infty, +\infty]$ is given by

$$u_g^c(v) = \sup_{g'\in H_g\cap\Omega} \left(c(g', v) - u(g')\right),$$

for every $v \in V_1$. Notice that

$$u_g^c(v) \geq c(g', v) - u(g'), \quad \text{for all } g' \in H_g \cap \Omega. \tag{12}$$

Furthermore, $u_g^c(v) = -\infty$ for some $v \in V_1$ if and only if $u(g') = +\infty$ for every $g' \in H_g \cap \Omega$.

The following theorem holds:

Theorem 4.5. *Let $u : \Omega \subset \mathbf{G} \to \mathbb{R}$. If $\partial_H^c u(g) \neq \emptyset$ for every $g \in \Omega$, then u is c H–convex. Moreover, let us suppose that c satisfies (**c2**) and $c(g, \cdot) : V_1 \to \mathbb{R}$ is continuous, for every $g \in \Omega$; if u is c H–convex, then $\partial_H^c u(g) \neq \emptyset$ for every $g \in \Omega$.*

Proof. Assume that $\partial_H^c u(g) \neq \emptyset$ for every $g \in \Omega$. If $(v, \alpha) \in \mathcal{P}_g$, then

$$u(g) \geq c(g, v) + \alpha. \tag{13}$$

We prove that, for every $g \in \Omega$, the set $\mathcal{P}_g$ is nonempty, and it contains an element (v, α) such that in (13) we have an equality. Notice that $p \in \partial_H^c u(g)$ if and only if

$$u(g) + u_g^c(p) = c(g, p). \tag{14}$$

Indeed,

$$\begin{aligned} p \in \partial_H^c u(g) \quad &\Longleftrightarrow \quad u(g') \geq u(g) + c(g', p) - c(g, p), \quad \forall g' \in H_g \cap \Omega \\ &\Longleftrightarrow \quad c(g, p) - u(g) \geq u_g^c(p). \end{aligned}$$

Taking into account (12), we obtain that (14) holds, and that $(p, -u_g^c(p))$ belongs to $\mathcal{P}_g$. Hence, for every $g \in \Omega$, we have that $u(g) = \sup_{\mathcal{P}_g}\{c(g, p) + \alpha\}$, thereby proving that u is c H–convex.

Conversely, fix $g_0 \in \Omega$. Since u is c H–convex, there exists a sequence $\{(p_n, \alpha_n)\}_n \subset \mathcal{P}_{g_0}$ such that

$$\begin{gathered} c(g, p_n) + \alpha_n \leq u(g), \quad \forall g \in H_{g_0} \cap \Omega \\ c(g_0, p_n) + \alpha_n \to u(g_0), \end{gathered} \tag{15}$$

with

$$c(g_0, p_n) + \alpha_n - u(g_0) > -1/n. \tag{16}$$

Inequalities (15) and (16) give, for every $g \in H_{g_0} \cap \Omega$ and for every n,

$$u(g) > u(g_0) + c(g, p_n) - c(g_0, p_n) - 1/n. \tag{17}$$

Let us first prove that $\{p_n\}_n$ is bounded in V_1. By contradiction, suppose that $\{p_n\}_n$ is unbounded; hence, by (**c2**), there exists $g' \in H_{g_0} \cap \Omega$ such that, by (17),

$$u(g') \geq \limsup_n (u(g_0) + c(g', p_n) - c(g_0, p_n) - 1/n) = +\infty.$$

This contradicts the assumption that u is real–valued. Therefore $\{p_n\}_n$ is bounded in V_1 and we can suppose that $p_n \to p \in V_1$. The continuity of $c(g_0, \cdot)$ and (16) imply that $\alpha_n \to -c(g_0, p) + u(g_0)$ and, consequently,

$$\alpha_n - u(g_0) + c(g_0, p) > -1/n \tag{18}$$

for sufficiently large n. For every $g \in H_{g_0} \cap \Omega$ and large n, (15), (18) and the continuity of $c(g_0, \cdot)$ give

$$\begin{aligned} u(g) &\geq \lim_n (c(g, p_n) + \alpha_n) \\ &\geq \lim_n (c(g, p_n) + u(g_0) - c(g_0, p) - 1/n) \\ &\geq c(g, p) + u(g_0) - c(g_0, p) \end{aligned}$$

This proves that $p \in \partial_H^c u(g_0)$. □

A consequence of the previous result is an extension of Proposition 4.2.

Corollary 4.6. *Let $\Omega \subset \mathbf{G}$ be an H–convex, open set, and let $u : \Omega \to \mathbb{R}$. If the function u is $\langle \xi_1(\cdot), \cdot \rangle$ H–convex, then u is H–convex. If the function u is H–convex, and measurable if $r > 2$, then u is $\langle \xi_1(\cdot), \cdot \rangle$ H–convex.*

Proof. The proof follows from Theorem 3.4, Theorem 4.5 and Proposition 3.3. □

Next two examples show that finiteness is a binding condition for the previous results, that fail when non real–valued functions are involved. Consequently, the investigation about the finiteness of a c H–convex function is critical (see Proposition 6.3).

Example 4.7. Consider the $\mathbb{R}^3$–convex function $u : \mathbb{H} \to (-\infty, +\infty]$ defined by

$$u(x, y, t) = \begin{cases} 0 & t \leq 0 \\ +\infty & t > 0. \end{cases}$$

It is an exercise to show that u is not $\langle \xi_1(\cdot), \cdot \rangle$ H–convex, while it is H–convex.

The previous example shows that the mentioned class of "abstract H–convex" functions and the class of $\langle \xi_1(\cdot), \cdot \rangle$ H–convex functions are coincident only for real–valued functions. In next example, c fulfills assumption (**c2**) together with a stronger concavity requirement, but this seems to be irrelevant.

Example 4.8. Let us consider the function $u : \mathbb{H} \to (-\infty, +\infty]$ defined as follows:

$$u(x, y, t) = \begin{cases} +\infty & \text{if } \max\{x, y\} > 0 \\ 0 & \text{if } \max\{x, y\} \leq 0. \end{cases}$$

This function turns out to be c H–convex, with $c(g, p) = -\langle \xi_1(g) - p, \xi_1(g) - p \rangle$; indeed, tedious computations show that, for every $g \in \mathbb{H}$,

$$\begin{aligned} \mathcal{P}_g = \{((v_1, v_2), \alpha) :& \\ & \alpha \leq 0 \text{ if } v_1 \leq 0 \text{ and } v_2 \leq 0,\ \alpha \leq v_1^2 \text{ if } v_1 > 0 \text{ and } v_2 \leq 0, \\ & \alpha \leq v_2^2 \text{ if } v_1 \leq 0 \text{ and } v_2 > 0,\ \alpha \leq v_1^2 + v_2^2 \text{ if } v_1 > 0 \text{ and } v_2 > 0\}. \end{aligned}$$

However, at any point $g = (x, y, t)$ such that $\max\{x, y\} > 0$, the set $\partial_H^c u(g)$ is empty. We remark that, for every fixed $v \in V_1$, the function c is strictly H–concave.

In general, it is reasonable to detect some properties about c implying the inclusion of the class of the H–convex functions in the class of the c H–convex functions. Next result provides a comparison between H–convexity and c H–convexity for real–valued functions; a similar one in the classical Euclidean case can be found in [13], Proposition 2.4.

Proposition 4.9. *Assume that, for every $p \in V_1$, the function $c(\cdot, p)$ is H–concave in $\mathbf{G}$ and, for every $g \in \Omega$,*

$$\bigcup_{v \in V_1} \partial^H c(g, v) = V_1. \tag{19}$$

Let $u : \Omega \subset \mathbf{G} \to \mathbb{R}$ be an H–convex function on the open, H–convex set Ω; moreover, if $r > 2$, we assume that u is measurable. Then u is c H–convex. In particular, any affine function $\phi(g) = \langle \xi_1(g), v \rangle + \alpha$, with $v \in V_1$ and $\alpha \in \mathbb{R}$, is c H–convex.

Proof. Fix any $g \in \Omega$. From the assumptions on u and from Theorem 3.4, we have that $\partial_H u(g) \neq \emptyset$. Hence there exists $p \in \partial_H u(g)$ such that:

$$u(g') \geq u(g) + \langle p, \xi_1(g') - \xi_1(g) \rangle, \quad \forall g' \in H_g \cap \Omega. \tag{20}$$

From (19), let $v = v(g, p)$ be such that $p \in \partial^H c(g, v)$. By assumption we have that

$$c(g', v) \leq c(g, v) + \langle p, \xi_1(g') - \xi_1(g) \rangle, \quad \forall g' \in H_g \cap \Omega. \tag{21}$$

Inequalities (20) and (21) imply that $v \in \partial_H^c u(g)$. From Theorem 4.5 the thesis follows. □

5. Regularity properties of c H–convex functions

The definition of c H–convexity given in Section 4, owing to its structure, does not highlight any properties of the function; in order to detect some regularity, an accurate analysis is needed.

The problem of the regularity of a c H–convex function has already been solved for real–valued $\langle\xi_1(\cdot),\cdot\rangle$ H–convex functions: indeed, Corollary 4.6 says that a $\langle\xi_1(\cdot),\cdot\rangle$ H–convex function is H–convex and hence, by the result of Balogh and Rickly (see Theorem 3.2), it is locally Lipschitz continuous.

In the classical situation, this investigation goes through the notion of semiconvexity, introduced by Douglis to select unique solutions for the Hamilton–Jacobi equation (see, for example, [20]); under suitable regularity assumptions on c, a c convex function f is locally semiconvex, and therefore it shares all the regularity enjoyed by convex functions (e.g., two derivatives almost everywhere, locally Lipschitz where finite).

Encouraged by these results, we introduce the following definition:

Definition 5.1. Let $\Omega \subset \mathbf{G}$ be H–convex. A function $u : \Omega \to (-\infty, +\infty]$ is *H–semiconvex* (or ℓ *H–semiconvex*) if it is $\mathbb{R}$–semiconvex on every horizontal segment, i.e., there exists a positive constant ℓ such that

$$u(g\delta_\lambda(g^{-1}g')) \leq (1-\lambda)u(g) + \lambda u(g') + \ell\lambda(1-\lambda)\|\xi_1(g^{-1}g')\|_{\mathfrak{g}}^2$$

for all $g \in \Omega$, $g' \in H_g \cap \Omega$, and $\lambda \in [0, 1]$.

We say that u is locally H–semiconvex in Ω if, for every open ball $B \subset \Omega$, u is H–semiconvex on B; here and in the sequel we consider balls arising from the gauge distance d, that are H–convex. Via the equality

$$(1-\lambda)\|\xi_1(g)\|_{\mathfrak{g}}^2 + \lambda\|\xi_1(g')\|_{\mathfrak{g}}^2 - \|\xi_1(g\delta_\lambda(g^{-1}g'))\|_{\mathfrak{g}}^2 = (1-\lambda)\lambda\|\xi_1(g) - \xi_1(g')\|_{\mathfrak{g}}^2,$$

an easy computation shows that u is H–semiconvex if and only the function

$$g \mapsto u(g) + \ell\|\xi_1(g)\|_{\mathfrak{g}}^2$$

is H–convex. Hence, the characterization (7) for H–convex functions in $\Gamma^1(\Omega)$, where Ω is open, gives us that u is ℓ H–semiconvex if and only if

$$\langle \mathbb{X}u(g) + 2\ell\xi_1(g) - \mathbb{X}u(g') - 2\ell\xi_1(g'), \xi_1(g) - \xi_1(g')\rangle \geq 0, \quad \forall g \in \Omega,\ g' \in H_g \cap \Omega. \quad (22)$$

Moreover, if $u \in \Gamma^2(\Omega)$, then u is ℓ H–semiconvex if and only if $[\mathbb{X}^2 u]^* \geq -2\ell I$ within Ω, i.e., $[\mathbb{X}^2 u(g)]^* + 2\ell I$ is positive semidefinite for all $g \in \Omega$.

The following fundamental proposition is the horizontal version of a result in [12] (see Proposition C2); as a matter of fact, our proof is completely different on account of the definition of c H–convexity:

Theorem 5.2. *Let $\psi : \Omega \subset \mathbf{G} \to (-\infty, +\infty]$ be a proper c H–convex function. Assume that* (**c1**) *is satisfied, i.e., for every open ball $B \subset \Omega$ there exists $K_B > 0$ such that*

$$\|\mathbb{X}c(g', v) - \mathbb{X}c(g, v)\|_{\mathfrak{g}} \leq 2K_B d(g', g), \quad \forall g', g \in B \text{ and } \forall v \in V_1. \quad (23)$$

Then, ψ is locally H–semiconvex.

Proof. Fix an open ball $B \subset \Omega$. By the assumptions, for all $g \in B$, $g' \in H_g \cap B$, and $v \in V_1$, we have

$$\|\mathbb{X}c(g',v) - \mathbb{X}c(g,v)\|_{\mathfrak{g}} \leq 2K_B\|g^{-1}g'\|_{\mathbf{G}} = 2K_B\|\xi_1(g') - \xi_1(g)\|_{\mathfrak{g}},$$

since $\xi(g^{-1}g') \in V_1$. In particular,

$$\langle \mathbb{X}c(g,v) - \mathbb{X}c(g',v), \xi_1(g') - \xi_1(g)\rangle \leq 2K_B\langle \xi_1(g') - \xi_1(g), \xi_1(g') - \xi_1(g)\rangle.$$

Hence, by (22), we have that $c(\cdot, v)$ is K_B H–semiconvex in B for every $v \in V_1$.

Since ψ is c H–convex, by definition

$$\psi(g) = \sup_{(v,\alpha)\in\mathcal{P}_g} (c(g,v) + \alpha),$$

where $\mathcal{P}_g = \{(v,\alpha) \in V_1 \times \mathbb{R} : \; c(g',v) + \alpha \leq \psi(g'), \; \forall g' \in H_g \cap \Omega\}$. Let us consider the function $\phi : B \to (-\infty, +\infty]$ defined, for every $g \in B$, by

$$\phi(g) = \psi(g) + K_B\|\xi_1(g)\|^2_{\mathfrak{g}}. \tag{24}$$

We will prove that ϕ is H–convex on B. By contradiction, assume that there exist $g \in B$ and $g' \in H_g \cap B$ such that ϕ is not $\mathbb{R}$–convex along the horizontal segment $\sigma_{g,g'}$; for every $\lambda \in [0,1]$, we denote by g_λ the point $\sigma_{g,g'}(\lambda)$ of such horizontal segment (see (2)). The following three cases can occur:

First case: ϕ is real–valued on $\sigma_{g,g'}$; in this case, there exists $\lambda \in (0,1)$ such that

$$2\epsilon = \phi(g_\lambda) - ((1-\lambda)\phi(g) + \lambda\phi(g')),$$

for some positive ϵ. From the definition (24) of ϕ and the c H–convexity of ψ, there exists $(p,\alpha) \in \mathcal{P}_{g_\lambda}$ such that

$$\phi(g_\lambda) - (c(g_\lambda,p) + \alpha + K_B\|\xi_1(g_\lambda)\|^2_{\mathfrak{g}}) < \epsilon$$

with

$$c(g'',p) + \alpha \leq \psi(g''), \quad \forall g'' \in H_{g_\lambda} \cap \Omega.$$

From

$$c(g,p) + \alpha \leq \psi(g) \quad \text{and} \quad c(g',p) + \alpha \leq \psi(g'),$$

and since $c(\cdot,p) + K_B\|\xi_1(\cdot)\|^2_{\mathfrak{g}}$ is H–convex in B, we get

$$\begin{aligned}
\phi(g_\lambda) &< c(g_\lambda,p) + \alpha + K_B\|\xi_1(g_\lambda)\|^2_{\mathfrak{g}} + \epsilon \\
&\leq (1-\lambda)(c(g,p) + \alpha + K_B\|\xi_1(g)\|^2_{\mathfrak{g}}) + \lambda(c(g',p) + \alpha + K_B\|\xi_1(g')\|^2_{\mathfrak{g}}) + \epsilon \\
&\leq (1-\lambda)(\psi(g) + K_B\|\xi_1(g)\|^2_{\mathfrak{g}}) + \lambda(\psi(g') + K_B\|\xi_1(g')\|^2_{\mathfrak{g}}) + \epsilon \\
&= (1-\lambda)\phi(g) + \lambda\phi(g') + \epsilon \\
&= \phi(g_\lambda) - \epsilon,
\end{aligned}$$

a contradiction.

Second case: ϕ is finite at the endpoints g and g', but $\phi(g_\lambda) = +\infty$ for some $\lambda \in (0,1)$. Then, by definition of ϕ, for infinitely many integers n there exists $(p_n, \alpha_n) \in \mathcal{P}_{g_\lambda}$ such that

$$n < c(g_\lambda, p_n) + \alpha_n + K_B \|\xi_1(g_\lambda)\|_{\mathfrak{g}}^2,$$

with

$$c(g'', p_n) + \alpha_n \leq \psi(g''), \quad \forall g'' \in H_{g_\lambda} \cap \Omega.$$

From

$$c(g, p_n) + \alpha_n \leq \psi(g) \quad \text{and} \quad c(g', p_n) + \alpha_n \leq \psi(g'),$$

and since $c(\cdot, p) + K_B \|\xi_1(\cdot)\|_{\mathfrak{g}}^2$ is H–convex in B, we get

$$\begin{aligned} n &< c(g_\lambda, p_n) + \alpha_n + K_B \|\xi_1(g_\lambda)\|_{\mathfrak{g}}^2 \\ &\leq (1-\lambda)(c(g, p_n) + \alpha_n + K_B \|\xi_1(g)\|_{\mathfrak{g}}^2) + \lambda(c(g', p_n) + \alpha_n + K_B \|\xi_1(g')\|_{\mathfrak{g}}^2) \\ &\leq (1-\lambda)(\psi(g) + K_B \|\xi_1(g)\|_{\mathfrak{g}}^2) + \lambda(\psi(g') + K_B \|\xi_1(g')\|_{\mathfrak{g}}^2) \\ &= (1-\lambda)\phi(g) + \lambda\phi(g') \\ &\leq \max\{\phi(g), \phi(g')\}, \end{aligned}$$

a contradiction.

Third case: If $\phi(g) = +\infty$, or $\phi(g') = +\infty$, then (5) holds for every $\lambda \in [0,1]$.

Hence ϕ is H–convex in B and the thesis follows. $\square$

In the case (11), condition (23) is satisfied with $K_B = 0$ and hence, as a consequence of the previous result, we have that

Remark 5.3. Every $\langle \xi_1(\cdot), \cdot \rangle$ H–convex function $u : \mathbf{G} \to (-\infty, +\infty]$ is H–convex.

We note that Theorem 5.2 is a very general result for proper function. In the next section, we will investigate the problem of the finiteness of a c H–convex function. However, the proposition above and the result by Balogh and Rickly give rise to some interesting regularity for real–valued c H–convex functions:

Corollary 5.4. *Let Ω be an open, H–convex subset of $\mathbf{G}$, and let $\psi : \Omega \to \mathbb{R}$ be a c H–convex function, measurable if $r > 2$. Assume that* (**c1**) *holds. Then,*

(i) ψ *is locally bounded;*

(ii) $\mathbb{X}\psi$ *exists a.e. on* Ω.

Proof. (i) From Theorem 5.2, for every open ball $B \subset \Omega$ there exists $\ell > 0$ such that the function $g \mapsto \Psi(g) = \psi(g) + \ell \|\xi_1(g)\|_{\mathfrak{g}}^2$ is H–convex on B, and, by the assumptions, it is measurable if $r > 2$. From Theorem 3.2, since Ψ is Lipschitz on every ball $B \subset \Omega$, Ψ is bounded on B; this implies the boundedness of ψ on B, for every $B \subset \Omega$.

(ii) From Theorems 3.2 and 2.4, $\mathbb{X}\Psi(g)$ exists for almost all $g \in B$; we conclude that $\mathbb{X}\psi(g)$ exists for almost all $g \in B$. $\square$

In the Euclidean case (see, for instance, Proposition 2.7 in [13]) a connection can be stated between the c subdifferential of a function f, and the gradients ∇c and ∇f; a perfectly symmetrical result holds in our framework.

Proposition 5.5. *Let $c : \mathbf{G} \times V_1 \to \mathbb{R}$ and $u : \Omega \subset \mathbf{G} \to \mathbb{R}$ be such that $\mathbb{X}u(g_0)$ and $\mathbb{X}c(g_0, v)$ exist for every $v \in V_1$ and for some $g_0 \in \mathrm{int}(\Omega)$.*

(*i*) *If $p \in \partial^c_H u(g_0)$, then $p \in (\mathbb{X}c(g_0, \cdot))^{-1} (\mathbb{X}u(g_0))$;*

(*ii*) *if $\partial^c_H u(g_0) \neq \emptyset$ and $\mathbb{X}c(g_0, \cdot) : V_1 \to V_1$ is one–to–one, then*

$$\partial^c_H u(g_0) = \left\{ (\mathbb{X}c(g_0, \cdot))^{-1} (\mathbb{X}u(g_0)) \right\}. \tag{25}$$

Proof. Let $p \in \partial^c_H u(g_0)$. From the definition of c H–subdifferential, for all $g \in H_{g_0} \cap \Omega$, we have that

$$u(g) - c(g, p) \geq u(g_0) - c(g_0, p);$$

in particular, g_0 is a minimum point for the function $g \mapsto u(g) - c(g, p)$ on the plane $H_{g_0} \cap \Omega$. This implies that

$$\mathbb{X}u(g_0) = \mathbb{X}c(g_0, p). \tag{26}$$

If we consider the function $\mathbb{X}c(g_0, \cdot) : V_1 \to V_1$, then (26) implies (*i*). The additional assumption in (*ii*) gives easily (25). □

Under more regularity assumptions on c, Theorem 4.5, Corollary 5.4 and Proposition 5.5 entail the following

Corollary 5.6. *Let Ω be an open, H–convex subset of* $\mathbf{G}$, *and $u : \Omega \to \mathbb{R}$ be a c H–convex function, measurable if $r > 2$. Assume that c fulfills* (**c1**), (**c2**) *and* (**c3**), *and that $c(g, \cdot) : V_1 \to \mathbb{R}$ is continuous, for every $g \in \Omega$.*

Then, for a.e. $g \in \Omega$, $\partial^c_H u(g)$ is a singleton, and $\partial^c_H u(g) = \{(\mathbb{X}c(g, \cdot))^{-1}(\mathbb{X}u(g))\}$.

Exploiting the previous results, a necessary condition for a function to be c H–convex can be given. Let Ω be an open set, and c be a function satisfying the assumptions of Corollary 5.6; we assume, in addition, that $c(\cdot, v) \in \Gamma^2(\Omega)$, for every $v \in V_1$. Consider a c H–convex function $u \in \Gamma^2(\Omega)$; then, from Theorem 4.5 and Proposition 5.5, we get that $\partial^c_H u(g)$ is a singleton and it is given by (25). For every $g_0 \in \Omega$, denote by p_0 the unique c H–subgradient of u at g_0; then, the function

$$\theta_{g_0} : H_{g_0} \cap \Omega \to \mathbb{R}, \quad \theta_{g_0}(g) = u(g) - c(g, p_0),$$

has a minimum at g_0. This implies that $[\mathbb{X}^2 \theta_{g_0}]^*(g_0) \geq 0$. From (25), we obtain a necessary condition for the c H–convexity of u :

$$[\mathbb{X}^2 u]^*(g) \geq [\mathbb{X}^2 c]^* \left(g, (\mathbb{X}c(g, \cdot))^{-1}(\mathbb{X}u(g)) \right), \quad \forall g \in \Omega.$$

In the particular situation where $\mathbf{G} = \mathbb{H}$ and $c(g, v) = -\|\xi_1(g) - v\|^2_{\mathfrak{h}}$, we obtain

$$[\mathbb{X}^2 u]^*(g) \geq -2I, \quad \forall g \in \mathbb{H}.$$

6. *c* H–cyclic monotonicity

In $\mathbb{R}^n$ and, more generally, in Banach spaces $\mathbf{X}$, the graph of the multivalued map defined via the subdifferential ∂f of a function f is a cyclically monotone subset of $\mathbf{X} \times \mathbf{X}^*$, i.e.

$$\sum_{i=0}^{n} \langle x_{i+1}, x_i^* \rangle \leq \sum_{i=0}^{n} \langle x_i, x_i^* \rangle,$$

for every finite sequence $\{(x_i, x_i^*)\}_{i=0}^n \subset \operatorname{graph}(\partial f)$, with $x_{n+1} = x_0$. A cyclically monotone subset in $\mathbf{X} \times \mathbf{X}^*$ is called maximal if it is not a proper subset of another cyclically monotone set in $\mathbf{X} \times \mathbf{X}^*$. In this context, a well–known result due to R. T. Rockafellar [18] says that the maximal cyclically monotone subsets of $\mathbf{X} \times \mathbf{X}^*$ are completely characterized as the graphs of the multivalued maps $x \mapsto \partial f(x)$, where f is a proper lower semicontinuous convex function.

This result was extended to the case of c convex functions $f : \Omega_1 \to (-\infty, +\infty]$ and c cyclically monotone sets $\Gamma \subset \Omega_1 \times \Omega_2$, where Ω_i are very general spaces (see, for instance, [19]). We recall that Γ is said to be c cyclically monotone if for all $\{(x_i, y_i)\}_{i=0}^n \subset \Gamma$, with $x_{n+1} = x_0$,

$$\sum_{i=0}^{n} c(x_{i+1}, y_i) \leq \sum_{i=0}^{n} c(x_i, y_i). \tag{27}$$

We would like to stress that the c subdifferential of f at a point is a (possibly empty) subset of Ω_2 defined in (10).

The aim of this section is to adapt Rockafellar's ideas in [16] to the sub–Riemannian structure of a Carnot group, in the "c case". Two are the main features of our setting. First, the horizontal subdifferential $\partial_H u$, that plays a fundamental role in the study of the horizontal convexity of u, is a subset of V_1. Thereby the graph of the map $g \mapsto \partial_H u(g)$ is a subset of $\mathbf{G} \times V_1$, and this is the main reason why we will introduce the notion of c H–cyclic monotonicity for a subset of $\mathbf{G} \times V_1$. Furthermore, the H–subdifferential of a function at a point carries information about the function only along horizontal segment through the point itself. To this purpose, in [6] we proved that, if u is a real–valued, H–convex function on $\mathbb{H}$, and so H–subdifferentiable, then their H–subgradients are sufficient to "reconstruct" the function. More precisely, using the definition of H–sequence (see Section 3), we proved the following

Theorem 6.1 (see [6], Theorem 6.4). *If $u : \mathbb{H} \to \mathbb{R}$ is an H–convex function, then*

$$u(g) = u(g_0) + \sup_{\mathcal{Q}_g} \left\{ \sum_{i=0}^{n-1} \langle p_i, \xi_1(g_{i+1}) - \xi_1(g_i) \rangle \right\}, \tag{28}$$

where g_0 is fixed, and

$$\mathcal{Q}_g = \big\{ \{(g_i, p_i)\}_{i=0}^n : \{g_i\}_{i=0}^n \ H\text{–sequence}, \ g_n = g, \ p_i \in \partial_H u(g_i) \big\}.$$

In the sequel, we deal with the more general case of c H–convex functions. To begin with, let us investigate the finiteness of a c H–convex function.

First of all, given a subset A of $\mathbf{G}$ and a point $g_0 \in A$, we will consider a particular set of points that has a good behaviour with respect to horizontal displacements from g_0 within A. Let us denote by $\mathcal{H}(g_0, A)$ the subset of A that contains exactly those points that can be reached starting from g_0 and moving along horizontal segments whose endpoints lye in A. More precisely, a point g belongs to $\mathcal{H}(g_0, A)$ if there exists an H–sequence $\{g_i\}_{i=0}^n$ such that $g_n = g$ and $g_i \in A$, for every $i = 1, 2, \ldots, n$. In some cases, this set is a singleton; as an example, if $A = \{(0, 0, t) \in \mathbb{H} : t \in \mathbb{R}\}$ and $g_0 = (0, 0, 0)$, we get $\mathcal{H}(g_0, A) = \{g_0\}$. From Proposition 2.1 we easily get the following

Remark 6.2. If g_0 is an interior point of A, then g_0 is an interior point of $\mathcal{H}(g_0, A)$.

In the next proposition, we prove a sufficient condition for the finiteness of a c H–convex function defined on a set Ω, at least on $\mathcal{H}(g_0, \Omega)$. This result will play a fundamental role in the main theorem of this section. Let us recall that, for a given multivalued map T, $\mathrm{dom}(T) = \{g \in \mathbf{G} : T(g) \neq \emptyset\}$.

Proposition 6.3. *Let* $u : \Omega \subset \mathbf{G} \to (-\infty, +\infty]$, *and let* $g_0 \in \Omega$ *be such that* $u(g_0) < +\infty$ *and* $\partial_H^c u(g_0) \neq \emptyset$. *Then,* u *is real–valued in* $\mathcal{H}(g_0, \mathrm{dom}(\partial_H^c u))$.

Proof. For any $g \in \mathcal{H}(g_0, \mathrm{dom}(\partial_H^c u))$, $g \neq g_0$, there exists an H–sequence $\{g_i\}_{i=0}^n$ with $g_n = g$ and $\partial_H^c u(g_i) \neq \emptyset$, for every $i = 0, 1, \ldots, n$. This implies that $u(g) \neq +\infty$. Indeed, since g_i and g_{i+1} are the endpoints of a horizontal segment, from Remark 4.4 it follows that $u(g_i) = +\infty$ if and only if $u(g_{i+1}) = +\infty$. Since $u(g_0) \neq +\infty$, we get the result. □

Recalling that an H–sequence $\{g_i\}_{i=0}^n$ is closed when $g_n \in H_{g_0}$ (in this case we set $g_{n+1} = g_0$), we give the following natural notion:

Definition 6.4. We say that $\mathcal{R} \subset \mathbf{G} \times V_1$ is a c H*–cyclically monotone set* if, for every sequence $\{(g_i, p_i)\}_{i=0}^n \subset \mathcal{R}$ such that $\{g_i\}_{i=0}^n$ is a closed H–sequence, we have that

$$\sum_{i=0}^{n} c(g_{i+1}, p_i) \leq \sum_{i=0}^{n} c(g_i, p_i). \tag{29}$$

We say that a multivalued map $T : \mathbf{G} \to \mathcal{P}(V_1)$ is a c H*–cyclically monotone map* if $\mathrm{graph}(T)$ is c H–cyclically monotone.

From this definition, we can express in a different way the characterization of H–convex functions in Γ^1 presented in (8):

Remark 6.5. Let $u \in \Gamma^1(\Omega)$. Then u is H–convex if and only if the map $g \mapsto \partial_H u(g)$ has a $\langle \xi_1(\cdot), \cdot \rangle$ H–cyclically monotone graph.

Notice that, for all functions u and c, without any regularity assumptions, the map $g \mapsto \partial_H^c u(g)$ has a c H–cyclically monotone graph. Indeed, if $\{(g_i, p_i)\}_{i=0}^n \subset \mathrm{graph}(\partial_H^c u)$ and $\{g_i\}_{i=0}^n$ is a closed H–sequence, then

$$u(g_{i+1}) - u(g_i) \geq c(g_{i+1}, p_i) - c(g_i, p_i), \quad i = 0, \ldots, n,$$

implies (29).

The following result is the converse of the previous note, and it provides a crucial link between our approach and some possible application in optimal mass transportation problems:

Theorem 6.6. *Let* $T : \mathbf{G} \to \mathcal{P}(V_1)$ *be a c H–cyclically monotone map. Then, for all* $g_0 \in \mathrm{int}(\mathrm{dom}(T))$*, there exists a c H–convex function* $f_{g_0} : \mathcal{H}(g_0, \mathrm{dom}(T)) \to \mathbb{R}$*, such that*

$$T(g) \subset \partial^c_H f_{g_0}(g), \quad \textit{for every } g \in \mathcal{H}(g_0, \mathrm{dom}(T)). \tag{30}$$

Let us first make some comments. The function f_{g_0}, that will be defined in (31), is the c version of the Rockafellar's function (28) in the sub–Riemannian setting.

The reader will infer that the function f_{g_0}, with $g_0 \in \mathrm{int}(\mathrm{dom}(T))$, could be defined, using (31), at every point g linked via a horizontal segment to a point in $\mathcal{H}(g_0, \mathrm{dom}(T))$; however, one cannot guarantee that f_{g_0} is real–valued at g and, above all, that it is c H–convex. Moreover, if $g_0 \in \mathrm{dom}(T)$ and there does not exist any point $g \in H_{g_0} \cap \mathrm{dom}(T)$, different from g_0, using (31), we obtain a trivial function whose domain is $\{g_0\}$ and $f_{g_0}(g_0) = -\infty$.

The domain $\mathcal{H}(g_0, \mathrm{dom}(T))$ of the function f_{g_0} can have a very strange shape. However, from Remark 6.2, if g_0 is an interior point of $\mathrm{dom}(T)$, then g_0 is an interior point of $\mathrm{dom}(f_{g_0})$.

In the sequel, we will denote by $\mathrm{dom}(f_{g_0})$ the set $\mathcal{H}(g_0, \mathrm{dom}(T))$.

Proof of Theorem 6.6. Let us suppose that T is c H–cyclically monotone, and fix $g_0 \in \mathrm{int}(\mathrm{dom}(T))$. For every $g \in \mathcal{H}(g_0, \mathrm{dom}(T))$ we define $\mathcal{Q}_g$ as the set of all sequences $\{(g_i, p_i)\}_{i=0}^n$, where $\{g_i\}_{i=0}^n$ is an H–sequence with starting point g_0, $p_i \in T(g_i)$ for $0 \le i \le n$, and $g_n \in H_g$. Let $f_{g_0} : \mathcal{H}(g_0, \mathrm{dom}(T)) \to (-\infty, +\infty]$ be the function defined by

$$f_{g_0}(g) = \sup_{\mathcal{Q}_g} \left\{ \sum_{i=0}^{n-1} (c(g_{i+1}, p_i) - c(g_i, p_i)) + c(g, p_n) - c(g_n, p_n) \right\}. \tag{31}$$

First of all, since $g_0 \in \mathrm{int}(\mathrm{dom}(T))$, the set $\mathcal{Q}_g$ is nonempty, and then $f_{g_0}(g)$ is greater than $-\infty$. Let us show that f_{g_0} is proper. For every $g_1 \in H_{g_0} \cap \mathcal{H}(g_0, \mathrm{dom}(T))$, we have

$$f_{g_0}(g_0) \ge c(g_1, p_0) - c(g_0, p_0) + c(g_0, p_1) - c(g_1, p_1);$$

if we choose $g_1 = g_0$, then we obtain $f_{g_0}(g_0) \ge 0$. Since T is c H–cyclically monotone, we have that

$$\sum_{i=0}^{n-1} (c(g_{i+1}, p_i) - c(g_i, p_i)) + c(g_0, p_n) - c(g_n, p_n) \le 0$$

for every sequence in $\mathcal{Q}_{g_0}$: clearly this implies $f_{g_0}(g_0) \le 0$. Hence $f_{g_0}(g_0) = 0$ and f_{g_0} is proper.

Next, let us choose $\overline{g} \in \mathcal{H}(g_0, \mathrm{dom}(T))$, and $\overline{p} \in T(\overline{g})$. For every $\alpha < f(\overline{g})$, there exists a sequence $\{(g_i, p_i)\}_{i=0}^n$ in $\mathcal{Q}_{\overline{g}}$ such that

$$\alpha < \sum_{i=0}^{n-1}(c(g_{i+1}, p_i) - c(g_i, p_i)) + c(\overline{g}, p_n) - c(g_n, p_n).$$

Let $g \in H_{\overline{g}} \cap \mathcal{H}(g_0, \mathrm{dom}(T))$. By adding to the sequence above the point $(\overline{g}, \overline{p})$, we obtain a new sequence that belongs to $\mathcal{Q}_g$. Then, by (31), we have

$$\begin{aligned} f_{g_0}(g) \geq\ & \sum_{i=0}^{n-1}(c(g_{i+1}, p_i) - c(g_i, p_i)) + c(\overline{g}, p_n) - c(g_n, p_n) + c(g, \overline{p}) - c(\overline{g}, \overline{p}) \\ >\ & \alpha + c(g, \overline{p}) - c(\overline{g}, \overline{p}). \end{aligned}$$

Since $\alpha < f_{g_0}(\overline{g})$ is arbitrary, we conclude that $\overline{p} \in \partial^c_H f_{g_0}(\overline{g})$. Hence we obtain (30).

From Proposition 6.3, since $\partial^c_H f_{g_0}(g) \neq \emptyset$ for every $g \in \mathcal{H}(g_0, \mathrm{dom}(T))$, and $f_{g_0}(g_0)$ is finite, we can conclude that f_{g_0} is real–valued within $\mathcal{H}(g_0, \mathrm{dom}(T))$. Finally, Theorem 4.5 and the nonemptiness of $\partial^c_H f_{g_0}(g)$ for every $g \in \mathcal{H}(g_0, \mathrm{dom}(T))$ implies that f_{g_0} is c H–convex. □

With some regularity assumptions on the function c, a c H–cyclically monotone multivalued map is, in fact, an a.e. single–valued map in its domain; furthermore, the graph of T coincides, locally, with the graph of the c H–subdifferential of a real–valued c H–convex function. At first sight, this seems to be a local conclusion in $\mathrm{dom}(T)$, but the different functions f_g that we construct on the sets $\mathcal{H}(g, \mathrm{dom}(T))$, with $g \in \mathrm{int}(\mathrm{dom}(T))$, share indeed the same c H–subdifferential. This is the content of the following proposition that provides the sub–Riemmanian version of the results in [12].

Corollary 6.7. *Assume that c satisfies* (**c1**), (**c2**) *and* (**c3**), *and that $c(g, \cdot) : V_1 \to \mathbb{R}$ is continuous, for every $g \in \mathbf{G}$. Let $T : \mathbf{G} \to \mathcal{P}(V_1)$ be a c H–cyclically monotone map, and denote by g_0 an interior point of* $\mathrm{dom}(T)$.

Then there exists a real–valued c H–convex function f_{g_0} with the following properties:

(i) *g_0 is an interior point of* $\mathrm{dom}(f_{g_0})$;
(ii) *for every $g \in \mathrm{dom}(\mathbb{X}f_{g_0})$, $T(g) = \partial^c_H f_{g_0}(g) = \{\mathbb{X}c(g, \cdot)^{-1}(\mathbb{X}f_{g_0}(g))\}$;*
(iii) $\mathrm{dom}(f_{g_0}) \setminus \mathrm{dom}(\mathbb{X}f_{g_0})$ *has null measure, with the additional assumption that f_{g_0} is measurable if $r > 2$.*

Let g_1 be another point in the interior of $\mathrm{dom}(T)$. *Then*

(iv) *if $g \in \mathrm{dom}(f_{g_0}) \cap \mathrm{dom}(f_{g_1})$, we have that $T(g) \subset \partial^c_H f_{g_0}(g) \cap \partial^c_H f_{g_1}(g)$;*
(v) *if $g \in \mathrm{dom}(\mathbb{X}f_{g_0}) \cap \mathrm{dom}(\mathbb{X}f_{g_1})$, we have that*

$$\partial^c_H f_{g_0}(g) = \{\mathbb{X}c(g, \cdot)^{-1}(\mathbb{X}f_{g_0}(g))\} = \{\mathbb{X}c(g, \cdot)^{-1}(\mathbb{X}f_{g_1}(g))\} = \partial^c_H f_{g_1}(g).$$

Proof. Clearly f_{g_0} is defined in Theorem 6.6 and consequently if finite and c H–convex. Remark 6.2 guarantees that (i) holds. From the construction of the function f_{g_0} in the proof of Theorem 6.6, we have that $\partial^c_H f_{g_0}(g) \neq \emptyset$ for every $g \in \mathrm{dom}(f_{g_0})$;

this argument and Proposition 5.5 imply (*ii*). Corollary 5.6 implies (*iii*). The last part of the Corollary follows from the previous implications and Theorem 6.6. □

7. An elementary application to optimal mass transportation in $\mathbb{H}$

Recently, as we mentioned, some papers have been devoted to the study of optimal mass transportation within Carnot groups. Whereas it should be clear to the reader that the focus of this paper is not this one, we would like to show, following timidly the line of the paper by Gangbo and McCann [12], how the tools introduced in the previous sections could be applied, at least if $\mathbf{G} = \mathbb{H}$.

Let (Ω_1, μ) and (Ω_2, ν) be probability spaces, and let us denote by $\Gamma(\mu, \nu)$ the set of the probability measures γ on $\Omega_1 \times \Omega_2$ with marginals μ and ν, i.e. such that $\gamma(A \times \Omega_2) = \mu(A)$ and $\gamma(\Omega_1 \times B) = \nu(B)$, for all μ–measurable sets A and ν–measurable sets B. We say that a map $s : \Omega_1 \to \Omega_2$ pushes μ forward to ν, i.e., $\nu = s_\sharp\mu$, if $\nu(B) = \mu(s^{-1}(B))$ for all ν–measurable sets B.

Monge's problem, formulated in 1781, takes into consideration $\Omega_1 = \Omega_2 = \mathbb{R}^n$, two measures μ and ν on $\mathbb{R}^n$, a cost function $c : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, and

$$\inf_{\{s:\ s_\sharp\mu=\nu\}} \int_{\mathbb{R}^n} c(x, s(x))\, d\mu(x). \tag{32}$$

A function $s_* : \mathbb{R}^n \to \mathbb{R}^n$, which minimizes (32), is called *optimal map.* In 1942, Kantorovich provided a relaxed version of the previous problem, as follows:

$$\inf_{\gamma\in\Gamma(\mu,\nu)} \int_{\mathbb{R}^n\times\mathbb{R}^n} c(x, y)\, d\gamma(x, y). \tag{33}$$

A measure $\gamma_* \in \Gamma(\mu, \nu)$ is an optimal measure if it is a minimum in (33). Since, for every μ such that $s_\sharp\mu = \nu$, the measure $\gamma = (\mathbf{1} \times s)_\sharp\mu$ belongs to $\Gamma(\mu, \nu)$, the change of variable shows that the functional in (33) coincides with the one in (32); this implies that the Kantorovich's infimum encompasses a large class of objects than that of Monge.

Among the other results, Gangbo and McCann proved that for a cost $c(x, y) = h(x-y)$, where h is a strictly convex and superlinear function (here, for simplicity, we assume $h \in C^1$), satisfying a technical condition that they call (H2) (see [12], p. 121), there exists a unique solution for both the Monge and the Kantorovich problems. In particular, if μ and ν are Borel measures on $\mathbb{R}^n$ such that μ is absolutely continuous with respect to the Lebesgue measure, and if the infimum in (33) is finite, then there exists a unique optimal measure $\gamma_* = (\mathbf{1} \times s_*)_\sharp\mu$, where s_* is an optimal map for the Monge's problem that is μ–a.e. defined through a c concave function φ, usually called "potential", via the formula $s_*(x) = x - (\nabla h)^{-1}(\nabla\varphi(x))$ (see Theorems 1.2 and 3.7 in [12]). Here, c convexity, and hence c concavity, are defined as in (9).

The main ingredients of this result can be summarized as follows: if γ_* is optimal, then its support $\mathrm{supp}(\gamma_*)$ is $(-c)$ cyclically monotone (according to (27), with the obvious changes of the sign due to the c concavity of φ). Consequently, there exists a c concave and Rockafellar's function φ such that $\mathrm{supp}(\gamma_*) \subset \mathrm{graph}(\partial^c\varphi)$. Since

φ is locally semiconcave, it is differentiable a.e. where it is finite; in particular, if $x \in \mathrm{dom}(\nabla\varphi)$, then the c superdifferential is a singleton and it is given by $\{x - (\nabla h)^{-1}(\nabla\varphi(x))\}$. Finally, the function s_* defined a.e. by the condition $(x, s_*(x)) \in \mathrm{graph}(\partial^c\varphi)$, provides the optimal map. One moment's reflection shows that the mentioned objects and tools have already been defined in the previous sections in our framework.

Let $(\mathbb{H}, \mu)$ and $(\mathbb{H}, \nu)$ be probability spaces; given a function $c : \mathbb{H} \times V_1 \to \mathbb{R}$, we define the "profit" function $C : \mathbb{H} \times \mathbb{H} \to [-\infty, \infty)$ as follows:

$$C(g, g') = \begin{cases} c(g, \xi_1(g')) & \text{if } (g, g') \in \mathcal{S} \\ -\infty & \text{if } (g, g') \notin \mathcal{S}, \end{cases}$$

where $\mathcal{S}$ denotes the (symmetric) set $\mathcal{S} = \{(g, g') \in \mathbb{H} \times \mathbb{H} : \ g' \in H_g\}$. We study the problem

$$\sup_{\gamma \in \Gamma(\mu,\nu)} \mathcal{C}(\gamma), \quad \text{where } \mathcal{C}(\gamma) = \int_{\mathbb{H}\times\mathbb{H}} C(g, g')\, d\gamma(g, g'). \tag{34}$$

We say that γ_* is optimal if $\mathcal{C}(\gamma_*) \geq \mathcal{C}(\gamma)$, for every $\gamma \in \Gamma(\mu, \nu)$. It is noteworthy that, with this type of profit function, any optimal map s_* moves every points, at least a.e., along their horizontal planes, i.e. $s_*(g) = g \exp v$, for some $v = v(g) \in V_1$. We will denote by S_γ the set $(\mathbf{1} \times \xi_1)(\mathrm{supp}(\gamma)) \subset \mathbb{H} \times V_1$.

The aim of this section is to show that, for our elementary problem (34), it can be reasonably introduced a notion of "potential" on $\mathbb{H}$ that identifies the optimal map. In order to do this, we have the following:

Proposition 7.1. *Let $c : \mathbb{H} \times V_1 \to (-\infty, 0]$ be a continuous function. Let γ_* be an optimal solution for problem* (34)*, with $\mathcal{C}(\gamma_*) > -\infty$, and suppose that $\gamma'_* = (\mathbf{1} \times \xi_1)_\# \gamma_*$ is optimal for*

$$\sup_{\gamma' \in \Gamma(\mu,\nu')} \int_{\mathbb{H}\times V_1} c(g, v)\, d\gamma'(g, v), \tag{35}$$

where $\nu' = (\xi_1)_\sharp \nu$. Then, the set S_{γ_*} is c H–cyclically monotone.

The assumptions of the proposition above deserve some comments. Indeed, for any $\gamma \in \Gamma(\mu, \nu)$, the measure $\gamma' = (\mathbf{1} \times \xi_1)_\sharp \gamma$ is in $\Gamma(\mu, \nu')$. On the contrary, if γ' belongs to $\Gamma(\mu, \nu')$, one cannot infer, in general, the existence of $\gamma \in \Gamma(\mu, \nu)$ such that $(\mathbf{1} \times \xi_1)_\sharp \gamma = \gamma'$. This implies that, if γ_* is optimal for (34), one cannot infer that $(\mathbf{1} \times \xi_1)_\sharp \gamma_*$ is optimal for (35).

Sketch of the proof. First of all notice that, if c is bounded from above, then for any $\gamma \in \Gamma(\mu, \nu)$ such that $\mathcal{C}(\gamma) > -\infty$, $\mathrm{supp}(\gamma) \subset \mathcal{S}$. By the change of variables theorem, we get

$$\mathcal{C}(\gamma) = \int_{\mathcal{S}} C(g, g')\, d\gamma(g, g') = \int_{S_\gamma} c(g, v)\, d\gamma'(g, v),$$

where $\gamma' = (\mathbf{1} \times \xi_1)_\sharp \gamma$.

Let γ_* and γ'_* satisfy the assumptions. By contradiction, assume that S_{γ_*} is not c H–cyclically monotone; then, there exists $\{(g_i^*, p_i^*)\}_0^n \subset \mathrm{supp}(\gamma'_*)$, where $\{g_i^*\}_0^n$ is a closed H–sequence, such that the continuous function $f : (\mathbb{H})^{n+1} \times (V_1)^{n+1} \to \mathbb{R}$

$$f(g_0, g_1, \dots, g_n, p_0, p_1, \dots, p_n) = \sum_{i=0}^{n} (c(g_{i+1}, p_i) - c(g_i, p_i))$$

is positive at $g_i = g_i^*$ and $p_i = p_i^*$. At this step, the proof follows the same line of Theorem 2.3 in [12], showing that γ'_* cannot be an optimal measure for problem (35). □

The result above allows us to connect the tools of the previous sections to the optimal transportation, and to introduce a notion of "potential" in the Heisenberg framework. Since this will be defined via the Rockafellar's function of Theorem 6.6, we must take into account that such theorem provides only local information.

Let $c : \mathbb{H} \times V_1 \to (-\infty, 0]$ be a continuous function satisfying (**c1**), (**c2**) and (**c3**), and γ_* and γ'_* be as in Proposition 7.1. We consider the multivalued map $T_{\gamma_*} : \mathbb{H} \to \mathcal{P}(V_1)$ defined as

$$T_{\gamma_*}(g) = \{v \in V_1 : \ (g, v) \in S_{\gamma_*}\}.$$

Proposition 7.1 guarantees that T_{γ_*} is a c H–cyclically monotone map. From Corollary 6.7, there exists a family of c H–convex functions

$$\Psi = \{\psi_g : \ g \in \mathrm{int}(\mathrm{dom}(T_{\gamma_*}))\}$$

such that $\mathrm{dom}(\psi_g)$ is a subset of $\mathrm{dom}(T_{\gamma_*})$ and contains g as an interior point. Moreover, for every $g \in \mathrm{int}(\mathrm{dom}(T_{\gamma_*}))$ and for a.e. $g' \in \mathrm{dom}(\psi_g)$, there exists $\mathbb{X}\psi_g(g')$ and hence $\partial_H^c \psi_g(g')$ is a singleton. Finally, if g' is in the domain of two functions ψ_{g_1}, ψ_{g_2} in Ψ, then $\partial_H^c \psi_{g_1}(g')$ and $\partial_H^c \psi_{g_2}(g')$ have nonempty intersection.

For these reasons, given a point $g \in \mathrm{int}(\mathrm{dom}(T_{\gamma_*}))$, we define the c H–subdifferential of the family Ψ at g as the set

$$\partial_H^c \Psi(g) = \bigcap_{\{g' \in \mathrm{int}(\mathrm{dom}(T_{\gamma_*})) : \ g \in \mathrm{dom}(\psi_{g'})\}} \partial_H^c \psi_{g'}(g).$$

Since, for every $g' \in \mathrm{int}(\mathrm{dom}(T_{\gamma_*}))$ and for every $g \in \mathrm{dom}(\psi_{g'})$, Theorem 6.6 guarantees that $T_{\gamma_*}(g) \subset \partial_H^c \psi_{g'}(g)$, we have that $\partial_H^c \Psi(g)$ is nonempty. Clearly, for a.e. $g \in \mathrm{int}(\mathrm{dom}(T_{\gamma_*}))$, the set $\partial_H^c \Psi(g)$ is a singleton and it defines a.e. the optimal map s_*. More precisely, if $\partial_H^c \Psi(g)$ is a singleton, then $(g, \xi_1(s_*(g))) \in \mathrm{graph}(\partial_H^c \Psi)$. If we set $\mathbb{X}\Psi(g)$ as $\mathbb{X}\psi_{g'}(g)$, for some g' such that $g \in \mathrm{dom}(\mathbb{X}\psi_{g'})$, the optimal map s_* is given, almost surely, by

$$s_*(g) = g \exp\left((\mathbb{X}c(g, \cdot))^{-1} (\mathbb{X}\Psi(g)) - \xi_1(g) \right).$$

We conclude that the family of functions Ψ plays the role of the "potential" of the problem.

References

[1] L. Ambrosio, S. Rigot: Optimal mass transportation in the Heisenberg group, J. Funct. Anal. 208 (2004) 261–301.

[2] M. Avriel, W. E. Diewert, S. Schaible, I. Zang: Generalized Concavity, Plenum Press, New York (1988).

[3] E. J. Balder: An extension of duality-stability relations to nonconvex optimization problems, SIAM J. Control Optimization 15 (1977) 329–343.

[4] Z. M. Balogh, M. Rickly: Regularity of convex functions on Heisenberg groups, Ann. Sc. Norm. Super. Pisa, Cl. Sci. (5) 2 (2003) 847–868.

[5] A. Calogero, R. Pini: Note on the Fenchel transform in the Heisenberg group, J. Math. Anal. Appl. 368(1) (2010) 69–79.

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

[7] L. Capogna, D. Danielli, S. Pauls, J. T. Tyson: An Introduction to the Heisenberg Group and the Sub-Riemannian Isoperimetric Problem, Birkhäuser, Basel (2007).

[8] D. Danielli, N. Garofalo, D. M. Nhieu: Notion of convexity in Carnot groups, Commun. Anal. Geom. 11 (2003) 263–341.

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

[10] H. Dietrich: Zur c–Konvexität und c–Subdifferenzierbarkeit von Funktionalen, Optimization 19 (1988) 355–371.

[11] G. B. Folland, E. M. Stein: Hardy Spaces on Homogeneous Groups, Princeton University Press, Princeton (1982).

[12] W. Gangbo, R. J. McCann: The geometry of optimal transportation, Acta Math. 177 (1996) 113–161.

[13] C. E. Gutiérrez, T. van Nguyen: On Monge–Ampère type equations arising in optimal transportation problems, Calc. Var. Partial Differ. Equ. 28 (2007) 275–316.

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

[15] S. Rigot: Mass transportation in the groups of type H, Commun. Contemp. Math. 7 (2005) 509–537.

[16] R. T. Rockafellar: Characterization of the subdifferentials of convex functions, Pac. J. Math. 17 (1966) 497–510.

[17] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton (1969).

[18] R. T. Rockafellar: On the maximal monotonicity of subdifferential mappings, Pac. J. Math. 33 (1970) 209–216.

[19] L. Rüschendorf: On c–optimal random variables, Stat. Probab. Lett. 27 (1996) 267–270.

[20] C. Villani: Optimal Transport. Old and New, Springer, Berlin (2009).