Journal of Convex Analysis
Volume 12 (2005), No. 1, 187–196

Geodetically Convex Sets in the Heisenberg Group

Roberto Monti
Mathematisches Institut, Universität Bern,
Sidlerstrasse 5, 3012 Bern, Switzerland
Permanent Address: Universita di Padova, Dipartimento di Matematica
Pura e Applicata, Via Belzoni 7, 35131 Padova, Italy
monti@math.unipd.it

Matthieu Rickly
Mathematisches Institut, Universität Bern,
Sidlerstrasse 5, 3012 Bern, Switzerland
matthieu.rickly@math-stat.unibe.ch

In memory of Matthias Studer

Received August 2, 2003

We prove that the geodetic envelope of a subset of the Heisenberg group containing three points not lying on the same geodesic is the whole group. As a corollary, we obtain that a function on the group which is convex along geodesics must be constant.

1. Introduction

Recently, several notions of convexity have been introduced and studied in Heisenberg groups and in more general Carnot groups. A weak and a strong definition of convex function are discussed in [4]: roughly speaking, the function is required to be convex along the integral curves of the left invariant horizontal vector fields. A different approach has been proposed in [9], where the notion of convexity in the viscosity sense has been transposed from the Euclidean to the sub–Riemannian setting. Relations between convexity in the viscosity sense and horizontal convexity as well as sharp regularity properties of convex functions in Heisenberg groups are discussed in [9] and [3].

The main motivation for the study of convexity in sub–Riemannian geometries is the development of a theory of fully non–linear sub–elliptic PDE's. For instance, it has been recently proved in [6] and [7] that convex functions in the Heisenberg group naturally define a Monge–Ampère measure satisfying a comparison principle.

In this paper we discuss a definition of convex set of geometric type. The Heisenberg group $\mathbb{H} \cong \mathbb{R}^3$ can be endowed with a left invariant metric d which is known as Carnot–Carathéodory metric. The resulting metric space $(\mathbb{H}, d)$ is geodesic: every pair of points can be connected by at least one geodesic. As in Euclidean space, a subset $A \subset \mathbb{H}$ can be defined to be *geodetically convex* if the image of any geodesic connecting two elements of A is contained in A. We describe the family of geodetically convex sets showing that, differently from the Euclidean case, it is very poor. Indeed, the geodetic convex envelope of a set consisting of merely two points can be the whole group:

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

Theorem 1.1. *Let $A = \{(x, y, t_1), (x, y, t_2)\} \subseteq \mathbb{H} \cong \mathbb{R}^3$ where $t_1 \neq t_2$. Then the smallest geodetically convex set containing A is $\mathbb{H}$.*

This is due to the fact that pairs of points with the same projection on the (x, y)–plane admit an infinite number of geodesics connecting them. The proof of Theorem 1.1 is given in §3.1. A consequence of Theorem 1.1 is the following

Theorem 1.2. *Let $p_1, p_2, p_3 \in \mathbb{H}$ be three points not lying on the same geodesic. Then the smallest geodetically convex set containing $\{p_1, p_2, p_3\}$ is $\mathbb{H}$.*

Thus, the only geodetically convex subsets of $\mathbb{H}$ are the empty set, points, arcs of geodesics and $\mathbb{H}$. The proof of Theorem 1.2 is given in §3.2.

The lack of geodetically convex sets has its counterpart in the lack of geodetically convex functions on $\mathbb{H}$. A function $u : \mathbb{H} \to \mathbb{R}$ is said to be *geodetically convex* if for any $p_0, p_1 \in \mathbb{H}$ and any geodesic $\gamma : [0, d(p_0, p_1)] \to \mathbb{H} \cong \mathbb{R}^3$ parameterized by arc–length connecting p_0 and p_1, the function $t \mapsto u(\gamma(t))$ is convex in the usual sense.

Corollary 1.3. *If $u : \mathbb{H} \to \mathbb{R}$ is geodetically convex, then u is constant.*

The proof is given in §3.3.

2. Basic facts concerning geodesics

We recall the definition and the properties of geodesics in the Heisenberg group needed in the sequel. In the following,

$$\mathbb{H} \cong \mathbb{R}^3 = \{(x, y, t) \mid x,\ y,\ t \in \mathbb{R}\}$$

denotes the Heisenberg group with the group law

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

One can check that the unit element is $0 \in \mathbb{R}^3$ and that the inverse of $p = (x, y, t)$ is $p^{-1} = (-x, -y, -t)$. The t–axis $Z = \{(0, 0, t) \mid t \in \mathbb{R}\}$ is the center of the group. In the setup of the Heisenberg group, Euclidean translations and dilations are replaced by the *left translations* $l_p : \mathbb{H} \to \mathbb{H}$, $p \in \mathbb{H}$, defined by

$$l_p(q) = p * q,$$

and by the *anisotropic dilations* $\delta_r : \mathbb{H} \to \mathbb{H}$, $r > 0$, defined by

$$\delta_r(x, y, t) = (rx, ry, r^2 t).$$

Clearly, $(\delta_r)_{r>0}$ is a group of automorphisms of $\mathbb{H}$.

The differential structure on $\mathbb{H}$ is determined by the left invariant vector fields

$$X = \frac{\partial}{\partial x} + 2y\frac{\partial}{\partial t}, \quad Y = \frac{\partial}{\partial y} - 2x\frac{\partial}{\partial t} \quad \text{and} \quad T = \frac{\partial}{\partial t},$$

where X and Y are the so–called *horizontal vector fields*. The pointwise linear span of X and Y forms a 2–dimensional bundle on $\mathbb{R}^3$ called *horizontal bundle* or *horizontal* *distribution.* The only non–trivial bracket relation is $[X,Y] = -4T$. Thus the distribution is bracket–generating.

A Lipschitz curve $\gamma : [0,1] \to \mathbb{H}$ is said to be *horizontal* if

$$\dot\gamma(t) \in \mathrm{span}_{\mathbb{R}}\{X(\gamma(t)),\, Y(\gamma(t))\}, \quad \dot\gamma(t) = h_1(t)X(\gamma(t)) + h_2(t)Y(\gamma(t)),$$

for almost every $t \in [0,1]$. The length of a horizontal curve $\gamma : [0,1] \to \mathbb{H}$ is defined to be

$$L(\gamma) = \int_0^1 (h_1^2(t) + h_2^2(t))^{\frac{1}{2}}\, dt.$$

Given $p_1,\, p_2 \in \mathbb{H}$, define

$$d(p_1, p_2) := \inf\{L(\gamma) \mid \gamma \text{ is a horizontal curve connecting } p_1 \text{ and } p_2\}.$$

It is well–known that d is a metric on $\mathbb{H}$, the so–called *Carnot–Carathéodory distance*, which is left–invariant, i.e. $d(l_p(p_1), l_p(p_2)) = d(p_1, p_2)$ for all $p,\, p_1,\, p_2 \in \mathbb{H}$, homogeneous, i.e. $d(\delta_r(p_1), \delta_r(p_2)) = rd(p_1, p_2)$ for all $p_1,\, p_2 \in \mathbb{H}$ and $r > 0$, and which induces the Euclidean topology on $\mathbb{R}^3$.

Definition 2.1. A *geodesic* connecting two points $p_0, p_1 \in \mathbb{H}$ is a length minimizing horizontal curve $\gamma : [0,T] \to \mathbb{H}$, i.e. a curve such that $\gamma(0) = p_0$, $\gamma(T) = p_1$ and $L(\gamma) = d(p_0, p_1)$.

It can be shown that $(\mathbb{H},\, d)$ is complete and that any two points can be connected by a (not necessarily unique) geodesic. Geodesics can be computed explicitly and we refer, for instance, to [5], [8], [2], [10] or [1] for a discussion of the problem. Precisely, geodesics starting from the origin $0 \in \mathbb{H}$ are smooth curves $\gamma = (\gamma_1, \gamma_1, \gamma_3)$ of the form

$$\begin{cases} \gamma_1(t) = \dfrac{\alpha\sin(\varphi t) + \beta(1-\cos(\varphi t))}{\varphi} \\ \gamma_2(t) = \dfrac{\beta\sin(\varphi t) - \alpha(1-\cos(\varphi t))}{\varphi} \\ \gamma_3(t) = 2\dfrac{\varphi t - \sin(\varphi t)}{\varphi^2}. \end{cases} \tag{1}$$

The real parameters α, β, φ specify the geodesic. It is useful to work with geodesics parameterized by arc–length. The condition ensuring arc–length parametrization is $\alpha^2 + \beta^2 = 1$ and the curve must be consequently defined on an interval $[0, L]$ where $L = L(\gamma)$. With this choice $t \geq 0$ is the arc–length parameter. In the case $\varphi = 0$, the formulae (1) must be understood in the limit sense and the curve γ is a geodesic on $[0, L]$ for any $0 \leq L < +\infty$. If $\varphi \neq 0$ the curve γ in (1) is length minimizing on $[0, L]$ if and only if $L \leq 2\pi/|\varphi|$. For $L > 2\pi/|\varphi|$ the curve $\gamma : [0, L] \to \mathbb{H}$ is not a geodesic anymore.

Geodesics starting from an arbitrary point can be recovered from (1) by left translations. Note that isometries of $(\mathbb{H}, d)$ and dilations transform geodesics into geodesics.

In the following proposition, we list some known properties of geodesics that can be derived from (1) and will be used in the sequel.

Proposition 2.2.

(*i*) *For any* $p \in \mathbb{H} \setminus Z$ *there exists a unique geodesic connecting* 0 *and* p.

(*ii*) *For any* $p \in Z$, $p \neq 0$, *and for any pair* $(\alpha, \beta) \in \mathbb{R}^2$ *with* $\alpha^2 + \beta^2 = 1$, *there exists a unique geodesic* γ *connecting 0 and* p *such that* $\dot\gamma(0) = \alpha\partial_x + \beta\partial_y = (\alpha, \beta, 0)$. *Moreover, the union of all the images of geodesics connecting 0 and* p *is the boundary of a convex open set which is invariant with respect to rotations of* $\mathbb{R}^3$ *that fix* Z.

(*iii*) *The image of the geodesic connecting the point* $p = (x, y, t)$ *with the point* $p^* = (-x, -y, t)$ *is the line segment* $[p^*, p]$.

(*iv*) *For* $\varphi \neq 0$ *the projection onto the* (x, y)*–plane of the geodesic* γ *in* (1) *is an arc of circle with radius* $1/|\varphi|$.

(*v*) *A geodesic* $\gamma : [0, L] \to \mathbb{H}$ *with parameter* $\varphi \in \mathbb{R} \setminus \{0\}$ *and* $L < 2\pi/|\varphi|$ *–respectively* $\varphi = 0$ *and* $L \geq 0$*– can be uniquely extended on* $[0, 2\pi/|\varphi|]$ *–respectively on* $[0, \widetilde{L}]$ *for any* $\widetilde{L} \geq L$.

(*vi*) *The mapping* $\Phi : \{(\alpha, \beta, \varphi, t) \mid \alpha^2 + \beta^2 = 1,\ \varphi \in \mathbb{R},\ t \in (0, \frac{2\pi}{|\varphi|})\} \to \mathbb{H} \setminus Z$ *given by*

$$\Phi(\alpha, \beta, \varphi, t) = \left(\frac{\alpha \sin\varphi t + \beta(1 - \cos\varphi t)}{\varphi}, \frac{\beta \sin\varphi t - \alpha(1 - \cos\varphi t)}{\varphi}, 2\frac{\varphi t - \sin\varphi t}{\varphi^2} \right)$$

is a homeomorphism.

Let us now state some definitions and preliminary results that will be needed in the proofs of Theorem 1.1 and 1.2.

Definition 2.3. We say that a set $C \subseteq \mathbb{H}$ is *geodetically convex* if for all $p_0, p_1 \in C$ and all geodesics $\gamma : [0, L] \to \mathbb{H}$ with $L = d(p_0, p_1)$, $\gamma(0) = p_0$ and $\gamma(L) = p_1$ we have $\gamma([0, L]) \subseteq C$. The *geodetic convex envelope* $\mathcal{C}(A)$ of $A \subseteq \mathbb{H}$ is the smallest geodetically convex subset of $\mathbb{H}$ containing A.

Definition 2.4. For $p_0, p_1 \in \mathbb{H}$, $\Gamma(p_0, p_1)$ denotes the set of images of geodesics connecting p_0 and p_1. Given $A \subseteq \mathbb{H}$ we define $\mathcal{G}(A) := \bigcup_{p_0, p_1 \in A} \Gamma(p_0, p_1)$, $\mathcal{G}^0(A) := A$ and $\mathcal{G}^{n+1}(A) := \mathcal{G}(\mathcal{G}^n(A))$ for all $n \in \mathbb{N}_0$.

Lemma 2.5. *For* $A \subseteq \mathbb{H}$ *we have* $\mathcal{C}(A) = \bigcup_{n \in \mathbb{N}_0} \mathcal{G}^n(A)$.

Proof. Using the fact that $\mathcal{G}^n(A) \subseteq \mathcal{G}^{n+1}(A)$ for all $n \in \mathbb{N}_0$, one easily checks that $\bigcup_{n \in \mathbb{N}_0} \mathcal{G}^n(A)$ is geodetically convex and contains A. This gives $\mathcal{C}(A) \subseteq \bigcup_{n \in \mathbb{N}_0} \mathcal{G}^n(A)$. On the other hand $A \subseteq \mathcal{C}(A)$, and if $\mathcal{G}^n(A) \subseteq \mathcal{C}(A)$ for some $n \in \mathbb{N}_0$, then $\mathcal{G}^{n+1}(A) \subseteq \mathcal{C}(A)$. $\square$

In the following $\mathcal{R}$ denotes the set of rotations of $\mathbb{R}^3$ that fix the center Z. Precisely,

$$\mathcal{R} = \left\{ R_\vartheta = \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) & 0 \\ -\sin(\vartheta) & \cos(\vartheta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \,\middle|\, \vartheta \in [0, 2\pi) \right\}. \tag{2}$$

The rotations R_ϑ are isometries of the Heisenberg group endowed with the Carnot–Carathéodory metric. The map $G : \mathbb{H} \to \mathbb{H}$ given by

$$G(x, y, t) = (-x, y, -t)$$

is also an isometry.

Lemma 2.6. *Let* $A \subseteq \mathbb{H}$*,* $p \in \mathbb{H}$*,* $r > 0$ *and* $R \in \mathcal{R}$*. Then* $\mathcal{C}(l_p(A)) = l_p(\mathcal{C}(A))$*,* $\mathcal{C}(\delta_r(A)) = \delta_r(\mathcal{C}(A))$*,* $\mathcal{C}(R(A)) = R(\mathcal{C}(A))$ *and* $\mathcal{C}(G(A)) = G(\mathcal{C}(A))$*.*

Proof. We prove the statement for R. Since isometries of $(\mathbb{H}, d)$ map geodesics to geodesics, it follows easily by induction that

$$R(\mathcal{G}^n(A)) = \mathcal{G}^n(R(A)), \quad n \in \mathbb{N}_0.$$

Moreover, by Lemma 2.5

$$R(\mathcal{C}(A)) = R\left(\bigcup_{n\in\mathbb{N}_0} \mathcal{G}^n(A)\right) = \bigcup_{n\in\mathbb{N}_0} R(\mathcal{G}^n(A)) = \bigcup_{n\in\mathbb{N}_0} \mathcal{G}^n(R(A)) = \mathcal{C}(R(A)).$$

□

Lemma 2.7. *We have* $\{(0,0,t) \mid -1 \leq t \leq 1\} \subseteq \mathcal{G}^2(\{(0,0,-1),(0,0,1)\})$*.*

Proof. By Proposition 2.2 (*ii*), the intersection of $\mathcal{G}^1(\{(0,0,-1),(0,0,1)\})$ with the plane

$$\Pi(\tau) := \{(x,y,t) \in \mathbb{H} \mid t = \tau\}, \ \tau \in (-1,1),$$

is a circle of radius $r(\tau) > 0$ centered at $(0,0,\tau)$. Pick any point p on the circle and denote by p^* the reflection of p with respect to $(0,0,\tau)$ in $\Pi(\tau)$. $(0,0,\tau) \in \mathcal{G}^2(\{(0,0,-1),(0,0,1)\})$ now follows from Proposition 2.2 (*iii*). □

3. Proof of the main results

3.1. Proof of Theorem 1.1

By Lemma 2.6, it is enough to show that $\mathcal{C}(A) = \mathbb{H}$ with $A = \{(0,0,-1),\ (0,0,1)\}$. The proof is divided in three steps. First, we show that rotations that fix Z and reflections with respect to the (x,y)–plane map $\mathcal{C}(A)$ onto itself. Second, we prove that

$$\left\{(x,y,t) \in \mathbb{R}^3 \mid 0 \leq |t| < h\left(\sqrt{x^2+y^2}\right)\right\} \subseteq \mathcal{C}(A) \tag{3}$$

for some non–increasing function $h : [0,+\infty) \to [0,+\infty]$. The last step consists in showing that h is nowhere finite.

1. Let $\vartheta \in [0, 2\pi)$ and denote by $R_\vartheta \in \mathcal{R}$ the rotation around Z with angle ϑ, as in (2). By Lemma 2.6, we have

$$R_\vartheta(\mathcal{C}(A)) = \mathcal{C}(R_\vartheta(A)) = \mathcal{C}(A).$$

We denote by $S : \mathbb{H} \to \mathbb{H}$, $S(x,y,t) = (x,y,-t)$, the reflection with respect to the (x,y)–plane. We claim that $S(\mathcal{C}(A)) = \mathcal{C}(A)$. By rotational symmetry of $\mathcal{C}(A)$, $S(\mathcal{C}(A)) = G(\mathcal{C}(A))$, where $G(x,y,t) = (-x,y,-t)$. By Lemma 2.6

$$S(\mathcal{C}(A)) = G(\mathcal{C}(A)) = \mathcal{C}(G(A)) = \mathcal{C}(A).$$

2. We claim that for any pair of points $(x,y,-t),(x,y,t) \in \mathcal{C}(A)$ we have

$$\{(x,y,t') \mid 0 \le |t'| \le |t|\} \subseteq \mathcal{C}(A).$$

Notice first that $\mathcal{C}(\{(x,y,-t),(x,y,t)\}) \subseteq \mathcal{C}(A)$ and

$$\mathcal{C}(\{(x,y,-t),(x,y,t)\}) = \mathcal{C}(l_p \circ \delta_r(\{(0,0,-1),(0,0,1)\}))$$

with $p=(x,y,0)$ and $r=\sqrt{|t|}$.

By Lemma 2.6 and Lemma 2.7 we obtain

$$\begin{aligned}\{(x,y,t') \mid 0 \le |t'| \le |t|\} &= l_p \circ \delta_r(\{(0,0,t') \mid 0 \le |t'| \le 1\}) \\ &\subseteq l_p \circ \delta_r(\mathcal{C}(\{(0,0,-1),(0,0,1)\})) \\ &= \mathcal{C}(l_p \circ \delta_r(\{(0,0,-1),(0,0,1)\})) \\ &= \mathcal{C}(\{(x,y,-t),(x,y,t)\}).\end{aligned}$$

Now define the function $h:[0,+\infty) \to [0,+\infty]$ by letting

$$h(r) := \sup\{t \ge 0 \mid (r,0,t) \in \mathcal{C}(A)\}.$$

By the above considerations we have (3).

We claim that h is non–increasing. Otherwise we can find $0 \le r_1 < r_2 < +\infty$ such that $0 \le h(r_1) < h(r_2)$. Pick $h(r_1) < t < h(r_2)$. Then by Proposition 2.2 (*iii*) the line segment $[(-r_2,0,t),(r_2,0,t)]$ is contained in $\mathcal{C}(A)$, which contradicts $h(r_1) < t$.

3. Our goal in this last step is to prove that h is nowhere finite, which concludes the proof of Theorem 1.1. *Assume by contradiction that there exists* $0 < r < +\infty$ *with* $0 \le h(r) < +\infty$. Without loss of generality we can also assume $h(r) > 0$. Indeed, if h is nowhere positive and finite, then

$$r_0 := \inf\{r \ge 0 \mid h(r) = 0\} < +\infty.$$

It is easy to see that $r_0 > 0$ and thus $Z \subseteq \mathcal{C}(A)$. For any $t>0$, $\mathcal{G}^2(\{(0,0,-t),(0,0,t)\}) \subseteq \mathcal{C}(A)$ contains a Carnot-Carathéodory ball centered at the origin whose radius tends to $+\infty$ as t tends to $+\infty$ (note $\delta_r(\mathcal{G}^2(B)) = \mathcal{G}^2(\delta_r(B))$ for any set $B \subseteq \mathbb{H}$ and any $r>0$), which contradicts $h(r)=0$ for $r > r_0$.

Let $p^- = (r,0,-h(r))$, $p^+ = (r,0,h(r))$. By Proposition 2.2 (*ii*) (modulo left translation), there exists a geodesic $\gamma:[0,d(p^-,p^+)] \to \mathbb{H}$ connecting p^- and p^+ with

$$\dot\gamma(0) = \frac{1}{\sqrt{2}}\big(X(p^-)+Y(p^-)\big) = \frac{1}{\sqrt{2}}(\partial_x + (\partial_y - 2r\partial_t)) = \frac{1}{\sqrt{2}}(1,1,-2r).$$

Hence, for some $s>0$ we have $\gamma(s) < -h(r)$ and $\gamma_1^2(s)+\gamma_2^2(s) > r^2$. But then the same is true for the geodesic $\widetilde{\gamma}:[0,d((r,0,-t),(r,0,t))] \to \mathbb{H}$ defined by

$$\widetilde{\gamma} = l_p \circ \delta_\lambda \circ l_{p^{-1}} \circ \gamma, \qquad \lambda = \sqrt{t/h(r)},\ p=(r,0,0),$$

connecting $(r,0,-t)$ and $(r,0,t)$, provided that $t \in (0,h(r))$ is sufficiently close to $h(r)$. This contradicts the fact that h is non–increasing. $\square$

3.2. Proof of Theorem 1.2

In the following, the map $P : \mathbb{H} \to \mathbb{H}$ defined by $P(x, y, t) = (x, y, 0)$ is the orthogonal projection onto the (x, y)–plane.

The proof of Theorem 1.2 relies upon Theorem 1.1. We show that $\mathcal{C}(\{p_1, p_2, p_3\})$ contains a pair of points lying on the same vertical line. The following lemma is a key step towards this goal.

Lemma 3.1. *Let $A \subseteq \mathbb{H}$. Suppose there exist $q \in \mathbb{R}^2$, a neighbourhood U of q in $\mathbb{R}^2$ and a continuous function $f : U \to \mathbb{R}$ such that*

$$\widetilde{A} := \{(x, y, t) \in A \mid (x, y) \in U\} = \{(x, y, f(x, y)) \mid (x, y) \in U\}.$$

Then A is not geodetically convex.

Proof. Without loss of generality we can assume that $q = (0, 0)$. *Suppose by contradiction that A is geodetically convex.* Note that, by Proposition 2.2 (i) (modulo left translation), for a given pair of points in $\widetilde{A}$ there is a unique geodesic connecting them.

1. Choose $r > 0$ such that $B := \{(x, y) \mid x^2 + y^2 < r^2\} \subset\subset U$ and define $g : \partial B \to \mathbb{R}$ by

$$g(x, y) := f(x, y) - f(-x, -y).$$

Since $g(x, y) = -g(-x, -y)$, the continuity of g implies $g(x, y) = 0$ for some $(x, y) \in \partial B$, i.e. $f(x, y) = f(-x, -y)$ for a pair of points (x, y), $(-x, -y)$ in ∂B. Since A is geodetically convex, $\{(sx, sy, f(x, y)) \mid s \in [-1, 1]\} \subseteq \widetilde{A}$. Hence $f(sx, sy) = f(x, y)$ for all $s \in [-1, 1]$. By Lemma 2.6 we may assume that $x > 0$, $y = 0$ and $f(0) = 0$.

2. Let $v \in S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} \subseteq \mathbb{R}^2$. Then the map $s \mapsto f(sv)$ from $[-r, r]$ to $\mathbb{R}$ is monotonic. Otherwise, by continuity of f, we could find $s_1 < s_2$ in $[-r, r]$ such that $f(s_1 v) = f(s_2 v)$, but either $f(s_i v) > \min\{f(sv) \mid s \in [s_1, s_2]\}$ or $f(s_i v) < \max\{f(sv) \mid s \in [s_1, s_2]\}$, $i = 1, 2$, which is not possible because the image of the geodesic connecting $(s_1 v, f(s_1 v))$ and $(s_2 v, f(s_2 v))$ is a line segment contained in $\widetilde{A}$.

3. The uniform continuity of f on compact subsets of U, the well–known estimate $d(p_1, p_2) \le C_K |p_1 - p_2|^{1/2}$ on compact subsets K of $\mathbb{H} \cong \mathbb{R}^3$ and the fact that the length of a geodesic is actually the Euclidean length of its projection onto the (x, y)–plane imply that the projection of the image of a geodesic connecting $(x_1, y_1, f(x_1, y_1))$ and $(x_2, y_2, f(x_2, y_2))$ is contained in an arbitrary small neighbourhood of (x_i, y_i), $i = 1, 2$, provided (x_1, y_1) and (x_2, y_2) are chosen sufficiently close to each other.

4. For $0 < \epsilon < \pi/2$, let

$$\begin{aligned}
p_0 &:= \big(r/2(\cos(-\epsilon), \sin(-\epsilon)), f(r/2(\cos(-\epsilon), \sin(-\epsilon)))\big),\\
p_1 &:= \big(r/2(\cos(\epsilon), \sin(\epsilon)), f(r/2(\cos(\epsilon), \sin(\epsilon)))\big),\\
q_0 &:= \big(r/2(\cos(\pi - \epsilon), \sin(\pi - \epsilon)), f(r/2(\cos(\pi - \epsilon), \sin(\pi - \epsilon)))\big),\\
q_1 &:= \big(r/2(\cos(\pi + \epsilon), \sin(\pi + \epsilon)), f(r/2(\cos(\pi + \epsilon), \sin(\pi + \epsilon)))\big).
\end{aligned}$$

Let $\gamma^p : [0, d(p_0, p_1)] \to \mathbb{H}$ and $\gamma^q : [0, d(q_0, q_1)] \to \mathbb{H}$ be the geodesics satisfying $\gamma^p(0) = p_0$, $\gamma^p(d(p_0, p_1)) = p_1$, $\gamma^q(0) = q_0$ and $\gamma^q(d(q_0, q_1)) = q_1$. If $\epsilon > 0$ is chosen small enough, then

(cf. **3.**)

$$P \circ \gamma^p([0, d(p_0, p_1)]) \subseteq (\{(x, y) \in B \mid 0 < x < r\}), \qquad \text{and}$$
$$P \circ \gamma^q([0, d(q_0, q_1)]) \subseteq (\{(x, y) \in B \mid -r < x < 0\}).$$

5. The horizontal plane at $(x, 0, 0)$ is spanned by the vectors ∂_x and $\partial_y - 2x\partial_t$. Now notice that there exists $s_p \in (0, d(p_0, p_1))$ with the properties

$$\gamma_2^p(s_p) = \gamma_3^p(s_p) = 0, \; \gamma_1^p(s_p) > 0 \quad \text{and} \quad \dot{\gamma}_2^p(s_p) > 0.$$

Indeed, γ^p must cross the (x, t)–plane, $\gamma^p([0, d(p_0, p_1)]) \subseteq \widetilde{A}$, $f \equiv 0$ on $\{(x, 0) \mid |x| \leq r\}$ (cf. **1.**) and $P \circ \gamma^p([0, d(p_0, p_1)])$ is a line segment or an arc of circle by Proposition 2.2. It follows that

$$\dot{\gamma}_3^p(s_p) = 2\dot{\gamma}_1^p(s_p)\gamma_2^p(s_p) - 2\dot{\gamma}_2^p(s_p)\gamma_1^p(s_p) = -2\dot{\gamma}_2^p(s_p)\gamma_1^p(s_p) < 0.$$

Similarly, there exists $s_q \in (0, d(q_0, q_1))$ with the properties

$$\gamma_2^q(s_q) = \gamma_3^q(s_q) = 0, \; \gamma_1^q(s_q) < 0, \; \dot{\gamma}_2^q(s_q) < 0 \quad \text{and} \quad \dot{\gamma}_3^q(s_q) < 0.$$

In particular, since $\gamma^p([0, d(p_0, p_1)])$, $\gamma^q([0, d(q_0, q_1)]) \subseteq \widetilde{A}$ and f is continuous, we can find $0 < \delta < \epsilon$ and $0 < \widetilde{\rho_1}, \widetilde{\rho_2} < r$ such that

$$f(\widetilde{\rho_1}(\cos(\delta), \sin(\delta))) < 0 \quad \text{and} \quad f(\widetilde{\rho_2}(\cos(\pi + \delta), \sin(\pi + \delta))) < 0.$$

Now, since $f(0) = 0$ and f is continuous on $\{\rho(\cos(\delta), \sin(\delta)) \mid -r \leq \rho \leq r\}$, there exist $0 < \rho_1, \rho_2 < r$ such that $t = f(\rho_1(\cos(\delta), \sin(\delta))) = f(\rho_2(\cos(\pi+\delta), \sin(\pi+\delta))) < 0$. Since the image of the geodesic connecting $(\rho_1(\cos(\delta), \sin(\delta)), t)$ and $(\rho_2(\cos(\pi+\delta), \sin(\pi+\delta)), t)$ is a line segment contained in $\widetilde{A}$, we get $(0, 0, t) \in \widetilde{A}$, a contradiction. $\square$

Proof of Theorem 1.2. Given three points $p_1, p_2, p_3 \in \mathbb{H}$ not lying on the same geodesic, we have to show that $\mathcal{C}(\{p_1, p_2, p_3\}) = \mathbb{H}$. Without loss of generality, we can assume $p_1 = 0$.

1. We claim that there exist two points $q_1, q_2 \in \mathcal{C}(\{p_1, p_2, p_3\})$ such that $q_1 \neq q_2$ and $P(q_1) = P(q_2)$. Then Theorem 1.2 follows from Theorem 1.1. *Assume by contradiction that no such pair of points exists.* Then there is always a *unique* geodesic connecting any given two points in $\mathcal{C}(\{p_1, p_2, p_3\})$.

2. Consider the geodesic $\kappa : [0, d(p_2, p_3)] \to \mathbb{H}$ satisfying $\kappa(0) = p_2$ and $\kappa(d(p_2, p_3)) = p_3$. For $\sigma \in [0, d(p_2, p_3)]$ let $\gamma_\sigma : [0, d(p_1, \kappa(\sigma))] \to \mathbb{H}$ be the unique geodesic such that $\gamma_\sigma(0) = p_1$ and $\gamma_\sigma(d(p_1, \kappa(\sigma))) = \kappa(\sigma)$. We show that if $\sigma < \tau$, then $\gamma_\sigma \cap \gamma_\tau = \{p_1\}$. If the intersection is larger, let $t_1 := \max\{t \in [0, d(p_1, \kappa(\sigma))] \mid \gamma_\sigma(t) \in \gamma_\tau\}$ and let t_2 be the unique element in $[0, d(p_1, \kappa(\tau))]$ with $\gamma_\tau(t_2) = \gamma_\sigma(t_1)$. By uniqueness of geodesics, $t_1 = t_2$ and $\gamma_\sigma|[0, t_1] = \gamma_\tau|[0, t_2]$. It then follows from Proposition 2.2 (v) that either $t_1 = d(p_1, \kappa(\sigma))$ or $t_2 = d(p_1, \kappa(\tau))$ and hence $\gamma_\sigma \subseteq \gamma_\tau$ or $\gamma_\tau \subseteq \gamma_\sigma$. Consider for instance the case $\gamma_\sigma \subseteq \gamma_\tau$. Clearly, $\kappa([\sigma, \tau]) \subseteq \gamma_\tau$. By Proposition 2.2 (v), it follows easily that $\kappa(0) = p_2 \in \gamma_\tau$ and thus $\gamma_0 \cup \kappa([0, \tau]) \subseteq \gamma_\tau$. By our assumption on $\mathcal{C}(\{p_1, p_2, p_3\})$, some extension $\tilde{\gamma}_\tau$ of γ_τ must contain $\gamma_0 \cup \kappa$. Consequently $p_1, p_2, p_3 \in \tilde{\gamma}_\tau$, a contradiction.

3. Consider the open set

$$U = \{(\sigma, s) \in \mathbb{R}^2 \mid \sigma \in (0, d(p_2, p_3)),\ s \in (0, d(p_1, \kappa(\sigma)))\},$$

and the mapping $F : U \to \mathbb{R}^2$ given by

$$F(\sigma, s) := P(\gamma_\sigma(s)).$$

By **2.**, F is injective. Moreover, by Proposition 2.2 (vi), $(\sigma, s) \to \gamma_\sigma(s)$ is continuous, because the endpoint $\kappa(\sigma)$ varies continuously. By the theorem on the invariance of domains – see e.g. Proposition 7.4 in A. Dold, "Lectures on Algebraic Topology", Springer, 1972 – the mapping F is open. In particular the set $V := F(U)$ is open and the inverse mapping $F^{-1} : V \to U$ is continuous. But then so is the function $f : V \to \mathbb{R}$ defined by

$$f(x, y) := g(F^{-1}(x, y)),$$

where $g : U \to \mathbb{R}$ is the third component of $(\sigma, s) \mapsto \gamma_\sigma(s)$. We have

$$\{(x, y, t) \in \mathcal{C}(\{p_1, p_2, p_3\}) \mid (x, y) \in V\} = \{(x, y, f(x, y)) \mid (x, y) \in V\},$$

and by Lemma 3.1 the set $\mathcal{C}(\{p_1, p_2, p_3\})$ cannot be geodetically convex. This contradiction concludes the proof. $\square$

3.3. Proof of Corollary 1.3

We have to show that a function $u : \mathbb{H} \to \mathbb{R}$ which is convex along geodesics is constant. First we show that u must be constant on the vertical axis Z. *Assume by contradiction this is not true.*

Case 1: There exist three distinct points $(0, 0, t_1), (0, 0, t_2), (0, 0, t_3) \in Z$ such that $u(0, 0, t_1) \le u(0, 0, t_2) < u(0, 0, t_3)$. The set $C := \{p \in \mathbb{H} \mid u(p) < u(0, 0, t_3)\}$ is geodetically convex, because u is convex on geodesics. Moreover $(0, 0, t_1), (0, 0, t_2) \in C$ and by Theorem 1.1 it follows that $C = \mathbb{H}$, contradicting $(0, 0, t_3) \notin C$.

Case 2: u assumes exactly two values on the vertical axis (say 0 and 1), $u(0, 0, t) = 0$ for some $t \in \mathbb{R}$ and $u(p) = 1$ for any $p \neq (0, 0, t)$ on the vertical axis (otherwise we are in Case 1). Consider two distinct geodesics γ and κ connecting $(0, 0, t)$ and $(0, 0, -t)$ (we can assume $t \neq 0$ since geodesics are preserved by left translations). We have $\gamma \cap \kappa = \{(0, 0, -t), (0, 0, t)\}$. By convexity of u on γ and κ, we can find $p \in \gamma \setminus (\gamma \cap \kappa)$ and $q \in \kappa \setminus (\gamma \cap \kappa)$ with $u(p)$, $u(q) < 1$. The set $C := \{p' \in \mathbb{H} \mid u(p') < 1\}$ is geodetically convex and contains $(0, 0, t)$, p and q. Since these points do not lie on the same geodesic, Theorem 1.2 gives $C = \mathbb{H}$ which contradicts $(0, 0, t') \notin C$ when $t' \neq t$.

By left translation, the previous argument shows that u must be constant on any vertical line. Suppose now we could find two vertical lines v_1 and v_2 and $c_1 < c_2$, such that c_i, $i = 1, 2$, is the value of u restricted to v_i. But then, if we choose two points on v_1 sufficiently far apart, the union of images of geodesics connecting these two points will intersect v_2, which is impossible since by geodetic convexity we must have $u \le c_1$ on this union. $\square$

References

[1] L. Ambrosio, S. Rigot: Optimal mass transportation in the Heisenberg group, J. Funct. Anal. 203 (2004) 261–301.

[2] A. Bellaïche: The tangent space in sub-Riemannian geometry, in: Sub-Riemannian Geometry, Prog. Math. 144, Birkhäuser, Basel (1996) 1–78.

[3] Z. M. Balogh, M. Rickly: Regularity of convex functions on Heisenberg groups, Ann. Scuola. Norm. Sup. Pisa Cl. Sci. (5) Vol. II (2003) 847–868.

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

[5] B. Gaveau: Principe de moindre action, propagation de la chaleur et estimées sous elliptiques sur certains groupes nilpotents, Acta Math. 139 (1977) 95–153.

[6] C. E. Gutiérrez, A. Montanari: Maximum and comparison principles for convex functions on the Heisenberg group, Commun. Partial Differ. Equations, to appear.

[7] N. Garofalo, F. Tournier: New properties of convex functions in the Heisenberg group, Trans. Am. Math. Soc., to appear.

[8] A. Korányi: Geometric properties of Heisenberg–type groups, Adv. Math. 56 (1985) 28–38.

[9] G. Lu, J. Manfredi, B. Stroffolini: Convex functions on the Heisenberg group, Calc. Var. Partial Differ. Equ. 19 (2003) 1–22.

[10] R. Monti: Some properties of Carnot-Carathéodory balls in the Heisenberg group, Rend. Mat. Acc. Lincei, s. 9, v. 11 (2000) 155–167.