Subdifferential and Properties of Convex Functions with respect to Vector Fields

Martino Bardi*
Dipartimento di Matematica
Università di Padova
via Trieste 63
35121 Padova, Italy
(`bardi@math.unipd.it`).

Federica Dragoni
School of Mathematics,
Cardiff University
Senghennydd Road
Cardiff CF24 2AG Wales, UK
(`DragoniF@cardiff.ac.uk`).

Abstract

We study properties of functions convex with respect to a given family $\mathcal{X}$ of vector fields, a notion that appears natural in Carnot-Carathéodory metric spaces. We define a suitable subdifferential and show that a continuous function is $\mathcal{X}$-convex if and only if such subdifferential is nonempty at every point. For vector fields of Carnot type we deduce from this property that a generalized Fenchel transform is involutive and a weak form of Jensen inequality. Finally we introduce and compare several notions of $\mathcal{X}$-affine functions and show their connections with $\mathcal{X}$-convexity.

Keywords: convex functions in Carnot groups, Carnot-Carathéodory metric spaces, subdifferential, Legendre-Fenchel transform, convex duality, Jensen inequality.

1 Introduction

Classical convex analysis was successfully extended to Riemannian manifolds by means of the notion of geodesic convexity. This concept can be defined in more general sub-Riemannian contexts. However, in the simplest example of such geometry, the Heisenberg group, Monti and Rickly [25] proved that all geodetically convex functions are constant, so this property is too restrictive. A notion of *horizontal convexity* in the Heisenberg group, that seems to have been first conceived by Caffarelli, was introduced and studied independently by Lu, Manfredi, and Stroffolini [20] and by Danielli, Garofalo, and Nhieu [15] (in more general Carnot groups and with the name of weak H-convexity). It uses convex combinations built by the group operation and dilations. Lu, Manfredi, and Stroffolini [20, 19] introduced also the notion of convexity in viscosity sense. It requires a stratification of the Lie algebra associated to the Carnot group, the choice of a basis of the first layer, that is the horizontal subspace, formed by left-invariant vector fields $\mathcal{X} = \{X_1, \dots, X_m\}$, and uses the Hessian matrix $D^2_{\mathcal{X}}u$ associated to these fields. These papers stimulated intensive work by several authors concerning the equivalence of these notions and the regularity properties of horizontally convex functions in stratified Lie groups, see, e.g., [2, 17, 18, 26, 21, 16, 11, 9, 10, 22] and the survey in the book [8] or [4] for more references.

In the paper [4] we considered the context of more general Carnot-Carathéodory spaces without the algebraic structure of Carnot groups. More precisely, we are given a finite family of vector fields on $\mathbb{R}^n$, $\mathcal{X} = \{X_1, \dots, X_m\}$ and the C-C metric

$$d(x,y) := \inf\{T \geq 0 \,|\, \exists\, \gamma \text{ admissible in } [0,T] \text{ with } \gamma(0) = x,\ \gamma(T) = y\}, \tag{1}$$

---

*This author is partially supported by the Fondazione CaRiPaRo Project "Nonlinear Partial Differential Equations: models, analysis, and control-theoretic problems", the MIUR project PRIN "Viscosity, geometric, and control methods for nonlinear diffusive models", and the European Project Marie Curie ITN "SADCO - Sensitivity Analysis for Deterministic Controller Design".

where a curve γ is admissible if it is absolutely continuous in $[0,T]$ and for some measurable functions $\alpha_i(t)$ with $\sum_{i=1}^m \alpha_i^2(t) = 1$

$$\dot{\gamma}(t) = \sum_{i=1}^m \alpha_i(t) X_i(\gamma(t)), \quad \text{a.e. } t \in [0,T]. \tag{2}$$

Our notion of convexity is based on the $\mathcal{X}$-lines, that are solutions of the previous system with constant α_i, i.e.,

$$\dot{x}(t) = \sum_{i=1}^m \alpha_i X_i(x(t)), \tag{3}$$

for some $\alpha \in \mathbb{R}^m$. Given $\Omega \subset \mathbb{R}^n$ open set, we say that a function $u : \Omega \to \mathbb{R}$ is *$\mathcal{X}$-convex* if $u \circ x_\alpha$ is convex for any $\mathcal{X}$-line x_α contained in Ω. Clearly this reduces to the classical convexity if $\mathcal{X}$ is the canonical basis of $\mathbb{R}^n$. In [4] we showed that if $X_1, \dots, X_m$ are the generators of a Carnot group this notion is also equivalent to the horizontal convexity defined in [15, 20]. Moreover we proved a characterization in terms of the inequality $D^2_{\mathcal{X}} u \geq 0$ in the viscosity sense, and a local Lipschitz estimate for $\mathcal{X}$-semiconvex functions in terms of the C-C distance:

$$|u(x) - u(y)| \leq L d(x,y), \quad \forall\, x, y \in \Omega_1, \tag{4}$$

where the constant L depends on the open set Ω_1 such that $\overline{\Omega}_1 \subset \Omega$. Further estimates for $\mathcal{X}$-convex functions can be found in the very recent paper of Magnani and Scienza [23].

In this paper we continue the study of $\mathcal{X}$-convex functions introduced in [4].

In Section 2 we define the $\mathcal{X}$-plane through a point x, $\mathbb{V}_x$, that corresponds to the horizontal space in Carnot groups, and define its parametrization Φ_x by the time-1 map of the flow of (3). We compute explicitly these objects in some important examples, in particular *vector fields of Carnot-type*, i.e., of the form

$$X_j = \frac{\partial}{\partial x_j} + \sum_{i=m+1}^n a_{ij}(x) \frac{\partial}{\partial x_i}, \quad j = 1, \dots, m. \tag{5}$$

In Section 3, following suitable motivations, we define the $\mathcal{X}$-subdifferential of u at x, $\partial_{\mathcal{X}} u(x)$, as the set of $p \in \mathbb{R}^m$ such that

$$u(y) \geq u(x) + p \cdot \Phi_x^{-1}(y), \quad \forall\, y \in \Omega \cap \mathbb{V}_x,$$

and prove that a continuous function u is $\mathcal{X}$-convex if and only if $\partial_{\mathcal{X}} u(x) \neq \emptyset$ for all x. Results of this kind were proved by Calogero and Pini in the Heisenberg group [10] and by Magnani and Scienza in Carnot groups [22].

In Section 4 we give two applications of the preceding result to fields of Carnot type. The first concerns a generalized Legendre-Fenchel transform of u and states that it is involutive if and only if u is $\mathcal{X}$-convex. This is a form of convex duality that extends a result obtained by Calogero and Pini in the Heisenberg group [9]. Next we prove some weak versions of Jensen integral inequality for $\mathcal{X}$-convex functions. The main result is that, if $\Omega = \Omega_1 \times \Omega_2$, $\Omega_1 \subseteq \mathbb{R}^m$ is convex, $\Omega_2 \subseteq \mathbb{R}^{n-m}$, μ_i is a finite measure on Ω_i, $i = 1, 2$, then

$$⨍_{\Omega} u \, d\mu_1 \times d\mu_2 \geq ⨍_{\Omega_2} u \left(⨍_{\Omega_1} \mathrm{y}^1 d\mu_1 \,,\, \mathrm{x}^2 \right) d\mu_2. \tag{6}$$

In Section 5 we name *$\mathcal{X}$-affine* a function u such that u and $-u$ are both $\mathcal{X}$-convex, and show some properties of these functions. If the fields are of Carnot type we prove that the weak Jensen inequality (6) is an equality, and therefore it is sharp. If, in addition, the C-C metric satisfies

$$d(x,y) < +\infty \quad \forall\, x, y,$$

we show that u is $\mathcal{X}$-affine if and only if there are $\beta \in \mathbb{R}$ and $p \in \mathbb{R}^m$ such that

$$u(x) = \beta + p \cdot \pi_m(x),$$

where π_m is the projection to the first m coordinates, a property that corresponds to being *horizontally affine* in Carnot groups [15, 9]. We also show that for Carnot-type fields a continuous function is $\mathcal{X}$-convex if and only if it can be represented as an envelope of horizontally affine functions.

Finally, let us mention that our initial motivation in the study of $\mathcal{X}$-convex functions is their role in the theory of nonlinear partial differential equations elliptic with respect to the derivatives X_iX_ju, and therefore degenerate elliptic with respect to the Euclidean derivatives if $m < n$. In particular, equations of Monge-Ampère type involving vector fields $X_1, \dots, X_m$ are well-posed in the viscosity sense among $\mathcal{X}$-convex functions [5, 6], and we showed in [4] that estimates like (4) are very useful in the study of these equations.

2 Preliminaries

2.1 Definitions and notations

Throughout the paper we are given a family of vector fields $\mathcal{X} = \{X_1, \dots, X_m\}$, $X_i : \mathbb{R}^n \to \mathbb{R}^n$, $i = 1, \dots, m$, $m \leq n$, at least of class C^1. We denote with $\sigma(x)$ the $n \times m$-matrix whose columns are the coefficients of the vector fields and call $\mathcal{X}$-line associated to the vector $\alpha \in \mathbb{R}^m$ a curve $x_\alpha : I \to \mathbb{R}^n$ solving the ODE

$$\dot{x}(t) = \sigma(x(t))\alpha, \tag{7}$$

where $I \subseteq \mathbb{R}$ is the maximal interval of existence of the solution. We are also given

$$\Omega \subseteq \mathbb{R}^n \quad \text{open set}$$

and we denote with I_{max} the maximal interval such that $x_\alpha(t)$ remains in Ω.

Definition 2.1. *We say that a function $u : \Omega \to \mathbb{R}$ is $\mathcal{X}$-convex if, for any $\alpha \in \mathbb{R}^m$, $u \circ x_\alpha$ is convex in I_{max}, where x_α is the $\mathcal{X}$-line defined by* (7).

We also say that Ω is $\mathcal{X}$-convex if, for all $x, y \in \Omega$ any $\mathcal{X}$-segment joining x to y is contained in Ω. (A $\mathcal{X}$-segment between two points x and y is the piece of the $\mathcal{X}$-line joining the two points, exactly as in the Euclidean case.) In this case (up to some reparametrization) we can assume that $x_\alpha(0) = x$ and $x_\alpha(1) = y$ and require that $u \circ x_\alpha(t)$ is convex on the interval $[0, 1]$.

If Ω is not $\mathcal{X}$-convex, Definition 2.1 requires that $u \circ x_\alpha(t)$ is convex in all connected components of the pre-image $x_\alpha^{-1}(\Omega)$. These connected components are disjoint open intervals. In this way we often do not need to impose any assumption on the domain of the function.

We recall that $\mathcal{X}$-convexity implies some regularity properties. Exactly as in the case of classical convex functions one can prove that $\mathcal{X}$-convex functions have locally bounded first derivatives (in the viscosity sense) in the directions of the vector fields and they are Lipschitz continuous with respect to the C-C distance (1). When such distance is continuous in the usual topology, all $\mathcal{X}$-convex functions are continuous, the first derivatives in the directions of the vector fields exist and they are in L^∞. Moreover, if the vector fields satisfy the Hörmander condition (e.g., in any Carnot group), then $\mathcal{X}$-convex functions are Hölder continuous of exponent $1/k$, where k is the step of the Hörmander condition. We refer to Section 6 in [4] for more details on these regularity properties.

Definition 2.2. *We call* $\mathcal{X}$-plane *associated to a point x the set of all the points that one can reach from x through a $\mathcal{X}$-line, i.e.*

$$\mathbb{V}_x := \big\{ y \in \mathbb{R}^n \,|\, \exists\, \alpha \in \mathbb{R}^m \text{ such that } x_\alpha(0) = x,\, x_\alpha(1) = y \big\}. \tag{8}$$

Roughly speaking the $\mathcal{X}$-plane associated to the point x is the union of all the $\mathcal{X}$-lines starting from the point x. In the particular case of vector fields that are generators of a Carnot group, $\mathbb{V}_x$ is the so-called horizontal space (see, e.g., [15]). We denote also

$$\mathcal{X}_x := \operatorname{Span}\big(X_1(x), \dots, X_m(x)\big) \subseteq \mathbb{R}^n$$

Note that $\mathbb{V}_x$ is a subset of $\mathbb{R}^n$ as manifold, while $\mathcal{X}_x$ is a set of "velocities", i.e. elements of the tangent space. If $\mathbb{V}_x$ is a subspace of $\mathbb{R}^n$, then $\mathbb{V}_x$ and $\mathcal{X}_x$ have the same dimension and they can be identified if necessary.

The function we define next gives a parametrization of $\mathbb{V}_x$ and will be extensively used in the paper.

Definition 2.3. The time-1 map of the flow (7) defining the $\mathcal{X}$-lines $x_\alpha(\cdot)$ is

$$\begin{array}{rl} \Phi_x : \mathbb{R}^m & \to \mathbb{V}_x \\ \alpha & \mapsto y = x_\alpha(1) \end{array} . \tag{9}$$

Next result collects some elementary properties of this function.

Lemma 2.1. *1. Φ_x is surjective, so Φ_x is invertible if and only if it is injective.*

2. If Φ_x is injective then $X_1, \dots, X_m$ are linearly independent at x.

3. If the vector fields are C^k, then Φ_x is C^k in both α and x. (

4. If Φ_x^{-1} exists, it takes any $\mathcal{X}$-lines starting at the point x into a Euclidean line starting at the origin of $\mathbb{R}^m$.

Remark 2.1. *i)* By the Lemma, if Φ_x is locally invertible and the fields are C^1 the set $\mathbb{V}_x$ is an m-dimensional submanifold of $\mathbb{R}^n$ with charts given by suitable restrictions of Φ_x.

ii) The converse of property 2 is not always true: the single vector field

$$X(x^1, x^2, x^3) = \big(x^2, -x^1, 1 - (x^1)^2 - (x^2)^2\big)$$

on $\mathbb{R}^3$ is linearly independent because it is never zero. If we consider a point x on the unit cylinder around the x^3-axis, the $\mathcal{X}$-lines from that point are unit circles on the cylinder. In this case you can reach a point y antipodal to x at the time 1 by moving with a starting velocity α but also with starting velocity $-\alpha$. (In the same way we can always reach any two points on the circle by two different $\mathcal{X}$-lines). Nevertheless Φ_x is locally invertible around 0.

iii) Property 4 holds because if $y = x_\alpha(t)$ with $x_\alpha(0) = x$, then $y = x_{t\alpha}(1)$.

Before turning to the examples we introduce some more notations for the case $m < n$, to which we are mostly interested. We indicate by $\mathrm{x}^1 \in \mathbb{R}^m$ and by $\mathrm{x}^2 \in \mathbb{R}^{n-m}$, respectively, the first m components and the last $n-m$ components of a point $x \in \mathbb{R}^n$, i.e.,

$$x = (\mathrm{x}^1, \mathrm{x}^2) \in \mathbb{R}^m \times \mathbb{R}^{n-m}.$$

Finally, π_m indicates the projection on the first m components

$$\pi_m x = \pi_m(x) := \mathrm{x}^1.$$

2.2 Examples of $\mathcal{X}$-planes.

We now compute $\mathbb{V}_x$ and Φ_x for different families of vector fields. In particular we want to show that Φ_x is invertible in the case of Carnot-type vector fields. On the other hand Φ_x is not even locally invertible in any Grušin-type space. Let us start with a very easy model.

Example 2.1 (Linearly independent constant vector fields)**.** Suppose $X_i^j(x) = 0$ for $j \neq i$ and $X_i^i(x) = 1$ for $i = 1, \dots, m$, $j = 1, \dots, n$ and $m < n$. In this case the $\mathcal{X}$-lines are Euclidean lines where the last $n - m$ components are constant and the $\mathcal{X}$-plane is

$$\mathbb{V}_x = \big\{(\mathrm{y}^1, \mathrm{y}^2) \in \mathbb{R}^n \,|\, \mathrm{y}^2 = \mathrm{x}^2\big\} = \mathbb{R}^m \times \{\mathrm{y}^2 = \mathrm{x}^2\}.$$

Moreover Φ_x is invertible and

$$\Phi_x^{-1}(\mathrm{y}^1, \mathrm{y}^2) = \mathrm{y}^1 - \mathrm{x}^1 = \pi_m(y - x).$$

The case of linearly independent constant vector fields is the easiest example of Carnot-type vector fields, that we introduce next.

Definition 2.4 (Carnot-type vector fields.)**.** *We say that* $X_1, \dots, X_m$, $m < n$, *are* Carnot-type vector fields *if the* $n \times m$ *matrix associated to them has the following form:*

$$\sigma(x) = \begin{pmatrix} Id_{m\times m} \\ A(x^1, \dots, x^m) \end{pmatrix}, \tag{10}$$

where the matrix $A(x^1, \dots, x^m)$ *is a* $(n - m) \times m$ C^1 *matrix depending only on the first* m *components of* x*.*

Interesting examples of Carnot-type vector fields are the generators of the Heisenberg group and of any other Carnot group, but in general no structure of Lie group is required (e.g., the Martinet distribution does not generate a Carnot group but it is associated to Carnot-type vector fields). See [24] and [8] for more details on these sub-Riemannian examples. Moreover Carnot-type vector fields are not required to satisfy the Hörmander condition.

Lemma 2.2. *If* $X_1, \dots, X_m$ *are of Carnot-type,then for any* $x \in \mathbb{R}^n$ *the function* Φ_x *defined by* (9) *is invertible and*

$$\Phi_x^{-1}(y) = \pi_m(y - x), \quad \text{for any } y \in \mathbb{R}^n.$$

Moreover there are C^1 *functions* $C^j : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ *such that the* $\mathcal{X}$*-plane can be written as*

$$\mathbb{V}_x = \big\{y \in \mathbb{R}^n \,|\, y^j = x^j + C^j(\mathrm{y}^1, \mathrm{x}^1), j = m + 1, \dots, n\big\}. \tag{11}$$

Proof. By (10) we get the following ODE for the $\mathcal{X}$-lines:

$$\dot{x}_\alpha(t) = \begin{cases} \qquad\qquad \alpha_1 \\ \qquad\qquad \vdots \\ \qquad\qquad \alpha_m \\ \sigma^1(x_\alpha^1(t), \dots, x_\alpha^m(t))\, \alpha \\ \vdots \\ \sigma^{n-m}(x_\alpha^1(t), \dots, x_\alpha^m(t))\, \alpha \end{cases}$$

where by $\sigma^i(x)$ we indicate the rows of $\sigma(x)$, i.e. $[\sigma^1(x), \dots, \sigma^{n-m}(x)]^t = A(\mathrm{x}^1)$. (Note that σ^i is a $1 \times m$ matrix while $\alpha \in \mathbb{R}^m$ is interpreted as a $m \times 1$ matrix, hence $\sigma^i\,\alpha$ is a well defined scalar.) By integrating the previous equation we get

$$(x_\alpha(t))^j = \begin{cases} x^j + \alpha_j\, t \qquad j = 1, \dots, m \\ x^j + \displaystyle\int_0^t \sigma^j(\alpha_1\, s + x^1, \dots, \alpha_m\, s + x^m)\, \alpha\, ds \qquad j = m + 1, \dots, n \end{cases} \tag{12}$$

which implies that $\Phi_x(\alpha) = x_\alpha(1)$ is invertible with

$$\Phi_x^{-1}(y) = \pi_m(y - x) = \mathrm{y}^1 - \mathrm{x}^1,$$

and the representation (11) holds with

$$C^j(\mathrm{y}^1,\mathrm{x}^1) = \int_0^1 \sigma^j\big((y^1-x^1)\,s+x^1,\dots,(y^m-x^m)\,s+x^m\big)\,\pi_m(y-x)\,ds$$

for $j = m+1,\dots,n$. □

Example 2.2. By computing the corresponding $\mathcal{X}$-lines, we can find the expression of $C^i(\cdot)$ in the following subcases of Carnot-type vector fields.

1. Linearly independent vector fields: $C^i(\mathrm{y}^1,\mathrm{x}^1) = 0$ for any $i = 1,\dots,n-m$ (see Example 2.1).
2. Heisenberg group (see e.g. [24] or [8] for a definition): $\mathbb{V}_x$ is the horizontal plane through x and
$$C^1(\mathrm{y}^1,\mathrm{x}^1) = \frac{\mathrm{y}^1_1\cdot\mathrm{x}^1_2 - \mathrm{y}^1_2\cdot\mathrm{x}^1_1}{2},$$
where by $\cdot$ we indicate the standard inner product in $\mathbb{R}^m$, for $\mathrm{y}^1 = (\mathrm{y}^1_1,\mathrm{y}^1_2)\in\mathbb{R}^d\times\mathbb{R}^d$, $\mathrm{x}^1 = (\mathrm{x}^1_1,\mathrm{x}^1_2)\in\mathbb{R}^d\times\mathbb{R}^d$ (note that in this case $m = 2d$ and $n = 2d+1$ with $d\geq 1$); therefore $C^1 = 0$ if and only if $x*y = y*x$ where $*$ is the law defined in the Heisenberg group (in fact this implies $\mathrm{y}^1_1\cdot\mathrm{x}^1_2 - \mathrm{y}^1_2\cdot\mathrm{x}^1_1 = 0$).
3. Martinet distribution (see [24] for a definition and some properties): $C^1(y^1,y^2,x^1,x^2) = -\frac{(y^2-x^2)^2}{3} - (y^2-x^2) - (x^2)^2$ (in this case $m = 2$ and $n = 3$).

Carnot-type vector fields are not the only family of vector fields where Φ_x is invertible for any x. In the next example we study the case of the rototranslation geometry which is a very well-know sub-Riemannian geometry, recently studied as a model for the visual cortex (see [14] and also [12, 13]).

Example 2.3 (The Rototraslation geometry)**.** The rototraslation geometry is the geometry defined on $\mathbb{R}^3$ by the vector fields

$$X_1(x^1,x^2,x^3) = \begin{pmatrix}\cos x^3\\ \sin x^3\\ 0\end{pmatrix} \quad\text{and}\quad X_2(x^1,x^2,x^3) = \begin{pmatrix}0\\0\\1\end{pmatrix}.$$

The $\mathcal{X}$-lines can be computed by solving

$$\dot{x}_\alpha(t) = \begin{cases}\dot{x}^1_\alpha(t) = \alpha_1\,\cos\big(x^3_\alpha(t)\big)\\ \dot{x}^2_\alpha(t) = \alpha_1\,\sin\big(x^3_\alpha(t)\big)\\ \dot{x}^3_\alpha(t) = \alpha_2.\end{cases}$$

If we assume $\alpha_2\neq 0$ we get

$$y = x_\alpha(1) = \begin{cases} y^1 = x^1 + \dfrac{\alpha_1}{\alpha_2}\big(\sin(\alpha_2+x^3) - \sin x^3\big)\\ \quad\;\, = x^1 + \dfrac{\alpha_1}{\alpha_2}\sin\alpha_2\,\cos x^3 + \dfrac{\alpha_1}{\alpha_2}\sin x^3(\cos\alpha_2 - 1)\\ y^2 = x^2 + \dfrac{\alpha_1}{\alpha_2}\big(\cos x^3 - \cos(\alpha_2+x^3)\big)\\ \quad\;\, = x^2 + \dfrac{\alpha_1}{\alpha_2}\sin\alpha_2\,\sin x^3 + \dfrac{\alpha_1}{\alpha_2}\cos x^3\,(1-\cos\alpha_2)\\ y^3 = x^3 + \alpha_2\end{cases}$$

while if $\alpha_2 = 0$ the $\mathcal{X}$-lines in $t = 1$ assume the form:

$$y = x_\alpha(1) = \begin{cases} y^1 = x^1 + \alpha_1 \cos x^3 \\ y^2 = x^2 + \alpha_1 \sin x^3 \\ y^3 = x^3 + \alpha_2 \end{cases}$$

Using the $\mathcal{X}$-lines, we can write the set $\mathbb{V}_x$ as

$$\begin{aligned} &\text{If } x^3 \neq k\,\frac{\pi}{2},\ k \in \mathbb{Z}, \quad \mathbb{V}_x = \left\{ (y^1, y^2, y^3) \in \mathbb{R}^3 \,|\, \frac{y^1 - x^1}{\cos x^3} = \frac{y^2 - x^2}{\sin x^3} \right\}, \\ &\text{If } x^3 = k\,\pi,\ k \in \mathbb{Z}, \quad \mathbb{V}_x = \left\{ (y^1, y^2, y^3) \in \mathbb{R}^3 \,|\, y^2 = x^2 \right\}, \\ &\text{If } x^3 = \frac{\pi}{2} + k\,\pi,\ k \in \mathbb{Z}, \quad \mathbb{V}_x = \left\{ (y^1, y^2, y^3) \in \mathbb{R}^3 \,|\, y^1 = x^1 \right\}. \end{aligned}$$

Moroever Φ_x is invertible on $\mathbb{V}_x$ and $\Phi^{-1}_{(x^1,x^2,x^3)}(y^1, y^2, y^3) = (\alpha_1, \alpha_2)$ with $\alpha_2 = y^3 - x^3$ while

$$\begin{aligned} \alpha_1 &= \frac{y^1 - x^1}{\cos x^3} \left(= \frac{y^2 - x^2}{\sin x^3} \right), && \text{if } x^3 \neq k\,\frac{\pi}{2}, \\ \alpha_1 &= \frac{y^1 - x^1}{\cos x^3}, && \text{if } x^3 = k\,\pi, \\ \alpha_1 &= \frac{y^2 - x^2}{\sin x^3}, && \text{if } x^3 = \frac{\pi}{2} + k\,\pi. \end{aligned}$$

It is obvious that whenever $X_1, \dots, X_m$ are not linearly independent at some point x, then the corresponding Φ_x cannot be invertible. One of the main example of this is given by Grušin spaces, see, e.g., [7] and the next example.

Example 2.4 (The Grušin plane)**.** Consider the vector fields on $\mathbb{R}^2$

$$X_1(x^1, x^2) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad X_2(x^1, x^2) = \begin{pmatrix} 0 \\ x^1 \end{pmatrix}.$$

They are not linearly independent at points of the line $x^1 = 0$. Therefore Φ_x cannot be injective. The $\mathcal{X}$-lines can be found by solving

$$\dot{x}^1_\alpha(t) = \alpha_1 \quad \text{and} \quad \dot{x}^2_\alpha(t) = \alpha_2\, x^1_\alpha(t),$$

which gives $x^1_\alpha(1) = x^1 + \alpha_1$ and $x^2_\alpha(1) = x^2 + \alpha_2\, x^1 + \frac{\alpha_1\,\alpha_2}{2}$. To find an expression for $\mathbb{V}_x$ we can remark that $\alpha_1 = y^1 - x^1$ implies

$$y^2 = x^2 + \alpha_2 x^1 + \alpha_2 \frac{y^1 - x^1}{2} = x^2 + \alpha_2 \frac{x^1 + y^1}{2}.$$

This means that, whenever $y^1 \neq -x^1$, then α_2 can be uniquely determinated, otherwise it cannot. Moreover

$$\mathbb{V}_x = \left\{ (y^1, y^2) \in \mathbb{R}^2 \,|\, y^1 \neq -x^1 \right\} \cup \left\{ (y^1, y^2) \in \mathbb{R}^2 \,|\, y^1 = -x^1 \text{ and } y^2 = x^2 \right\} =: \mathbb{V}^1_x \cup \mathbb{V}^2_x.$$

Then the restriction of Φ_x to $\mathbb{V}^1_x$ is injective but the restriction to $\mathbb{V}^2_x$ is not.

To our knowledge most of the results proved in this paper (in particular the characterization of convex functions by a nonempty subdifferential) are open for Grušin spaces, although the results proved in [4] (i.e., the viscosity characterization for $\mathcal{X}$-convex functions and their local intrinsic Lipschitz continuity and the corresponding bounds for the intrinsic gradient) apply also to the case of Grušin spaces.

3 $\mathcal{X}$-subdifferential and $\mathcal{X}$-convex functions.

Definition 3.1. *Given $u : \Omega \to \mathbb{R}$, we denote with δ_i the i-th vector of the canonical basis of $\mathbb{R}^m$ and with $x_{\delta_i}(t)$ the corresponding $\mathcal{X}$-line starting from x at $t = 0$. The* $\mathcal{X}$-partial derivatives *(or* derivatives along the vector fields*) of u are*

$$X_i u(x) := \lim_{t\to 0} \frac{u(x_{\delta_i}(t)) - u(x)}{t}, \quad \text{for } i = 1, \dots, m.$$

The $\mathcal{X}$-gradient at the point x is

$$\nabla_{\mathcal{X}} u(x) := \sum_{i=1}^{m} X_i u(x) X_i(x).$$

In the case of a Carnot-Carathéodory space the $\mathcal{X}$-gradient coincides with the usual horizontal gradient (see [8] for some definitions). For later use we will identify the $\mathcal{X}$-gradient with the corresponding coordinate-vector w.r.t. the basis $X_1, \dots, X_m$, i.e.

$$\nabla_{\mathcal{X}} u(x) \in \mathcal{X}_x \quad \longleftrightarrow \quad D_{\mathcal{X}} u(x) = (X_1 u(x), \dots, X_m u(x))^t \in \mathbb{R}^m.$$

Definition 3.2. *We say that a function $u : \Omega \to \mathbb{R}$ is* $\mathcal{X}$-directionally differentiable *at a point $x \in \Omega$ if there exists $p \in \mathbb{R}^m$ (depending on the point x) such that*

$$\lim_{t\to 0} \frac{u(x_\alpha(t)) - u(x)}{t} = p \cdot \alpha, \quad \forall\, \alpha \in \mathbb{R}^m. \tag{13}$$

Note that if such a p exists, then it is unique and $p = D_{\mathcal{X}} u(x)$.

The following lemma states the existence of a supporting $\mathcal{X}$-hyperplane for the graph of u at the points of $\mathcal{X}$-directional differentiability of u.

Lemma 3.1. *If Φ_x is invertible and $u : \Omega \to \mathbb{R}$ is $\mathcal{X}$-convex and $\mathcal{X}$-directionally differentiable at a point $x \in \Omega$, then*

$$u(x) + D_{\mathcal{X}} u(x) \cdot \alpha \le u(\Phi_x(\alpha)), \quad \forall\, \alpha \in \mathbb{R}^m,$$

or, equivalently,

$$u(x) + D_{\mathcal{X}} u(x) \cdot \Phi_x^{-1}(y) \le u(y), \quad \forall\, y \in \mathbb{V}_x \cap \Omega.$$

Proof. For sake of simplicity we take $\Omega = \mathbb{R}^n$. For $y \in \mathbb{V}_x$ we take $\alpha \in \mathbb{R}^m$ and $x_\alpha : [0,1] \to \Omega$ such that $x_\alpha(0) = x$ and $x_\alpha(1) = y$. By definition of $\mathcal{X}$-convexity and writing $t = (1-t)0 + t1$, we find

$$u(x_\alpha(t)) \le (1-t)\, u(x_\alpha(0)) + t\, u(x_\alpha(1)) = u(x) + t\big(u(y) - u(x)\big), \quad \forall\, y \in \mathbb{V}_x,$$

which implies

$$\frac{u(x_\alpha(t)) - u(x)}{t} \le u(y) - u(x), \quad \forall\, y \in \mathbb{V}_x.$$

Passing to the limit as $t \to 0$ and using (13), we conclude

$$D_{\mathcal{X}} u(x) \cdot \alpha \le u(y) - u(x),$$

which proves the lemma. □

Motivated by Lemma 3.1 and by the classical definition of Euclidean subdifferential we introduce a notion of subdifferential along the vector fields for non-smooth functions. For $p, q \in \mathcal{X}_x$ denote

$$\langle p, q \rangle_{\mathcal{X}} := \sum_{i=1}^{m} p_i\, q_i \quad \text{for } p = \sum_{i=1}^{m} p_i X_i(x),\ q = \sum_{i=1}^{m} q_i X_i(x).$$

Definition 3.3. *Assume* $\Phi_x : \mathbb{R}^m \to \mathbb{V}_x$ *(defined in* (9)*) is invertible for any fixed* $x \in \Omega$*. The* $\mathcal{X}$-subdifferential of $u : \Omega \to \mathbb{R}$ at x *is the set*

$$\partial_{\mathcal{X}} u(x) := \left\{ p \in \mathcal{X}_x \,|\, u(x) + \left\langle p, \Psi_x(y) \right\rangle_{\mathcal{X}} \leq u(y),\, \forall\, y \in \Omega \cap \mathbb{V}_x \right\},$$

where $\Psi_x(y) = \sum_{i=1}^m \left(\Phi_x^{-1}(y)\right)_i X_i(x) \in \mathcal{X}_x$.

Remark 3.1. If the function Φ_x is not invertible (e.g. in the Grušin case) we can generalize the previous definition and call $\mathcal{X}$-subdifferential the set

$$\partial_{\mathcal{X}} u(x) := \left\{ p \in \mathcal{X}_x \,|\, u(y) \geq u(x) + \left\langle p, \Theta_x^y \right\rangle_{\mathcal{X}}, \forall\, \Theta_x^y \text{ and } \forall\, y \in \Omega \cap \mathbb{V}_x \right\},$$

where $\Theta_x^y := \sum_{i=1}^m \left(\eta_x^y\right)_i X_i(x) \in \mathcal{X}_x$ for any $\eta_x^y \in \Phi_x^{-1}(y)$, Φ_x^{-1} being the pre-image of Φ_x at the point y. Note that Lemma 3.1 is still true using this more general definition.

Remark 3.2. We can always identify any element in $\mathcal{X}_x$ by its coordinate vector w.r.t. the given family of vector fields. Using this identification, we can re-write the $\mathcal{X}$-subdifferential simply as a subset of $\mathbb{R}^m$, i.e.

$$\partial_{\mathcal{X}} u(x) \leftrightarrow \widetilde{\partial_{\mathcal{X}}} u(x) := \left\{ p \in \mathbb{R}^m \,|\, u(y) \geq u(x) + p \cdot \Phi_x^{-1}(y), \forall\, y \in \Omega \cap \mathbb{V}_x \right\}. \tag{14}$$

We will usually work with this set $\widetilde{\partial_{\mathcal{X}}} u(x)$ instead of the original $\mathcal{X}$-subdifferential.

Remark 3.3. If $X_1, \ldots, X_m$ are vector fields of Carnot-type, then by Lemma 2.2

$$\widetilde{\partial_{\mathcal{X}}} u(x) = \{ p \in \mathbb{R}^m \,|\, u(y) \geq u(x) + p \cdot \pi_m(y - x), \forall\, y \in \Omega \cap \mathbb{V}_x \}.$$

In this case $p \in \widetilde{\partial_{\mathcal{X}}} u(x)$ is the slope of a hyperplane supporting the restriction of u to $\mathbb{V}_x$. Moreover this notion of $\mathcal{X}$-subdifferential extends to general vector fields the notion of horizontal subdifferential introduced in Carnot groups in [15] (Definition 3.1) and studied later in [10] in the case of the Heisenberg group (see also [22]).

The main result of this section is the following.

Theorem 3.1. *Assume that* $X_1, \ldots, X_m$ *are linearly independent and* $u : \Omega \to \mathbb{R}$ *is continuous and* $\mathcal{X}$*-convex. Then* $\partial_{\mathcal{X}} u(x) \neq \emptyset$ *for all* $x \in \Omega$.

In view of Remark 3.2 we will write $\partial_{\mathcal{X}} u(x)$ instead of $\widetilde{\partial_{\mathcal{X}}} u(x)$, with a slight abuse of notation.

Remark 3.4. The result does not assume the Hörmander condition, so it generalizes to a very large class of vector fields what proved in [10] for the Heisenberg group and in [22] in Carnot groups. The case of Grušin spaces remains open since in this case the vector fields are not linearly independent at the origin.

Remark 3.5. If the vector fields $X_1, \ldots, X_m$ satisfy the Hörmander condition, we can remove the continuity assumption on u, requiring that u is upper semicontinuous on $\overline{\Omega}$ and locally bounded (see [4], Theorem 6.1)

Before proving the result we want to show that linearly independent vector fields imply that the associated Φ_x is locally invertible around 0, for any fixed x (i.e. the inverse Φ_x^{-1} exists for $y \in \mathbb{V}_x$ near x) and next we show that for $\mathcal{X}$-convex function the notion of $\mathcal{X}$-subdifferential can be written locally.

Lemma 3.2. *The map* Φ_x *has the Jacobian matrix such that* $D\Phi_x(0) = \sigma(x)$*. In particular, if* $X_1, \ldots, X_m$ *are linearly independent at* x*, then* Φ_x *is locally invertible at 0 for any fixed* x.

Proof. Recall that $\Phi_x(t\,\alpha) = x_\alpha(t)$, so taking the derivative in time, we get

$$\dot{x}_\alpha(t) = \frac{d}{dt}\Phi_x(t\,\alpha) = D\Phi_x(t\,\alpha)\,\alpha.$$

Moreover by the definition of $\mathcal{X}$-lines we know that

$$\dot{x}_\alpha(t) = \sigma(x_\alpha(t))\,\alpha = \sigma(\Phi_x(t\,\alpha))\,\alpha,$$

which means

$$D\Phi_x(t\,\alpha)\,\alpha = \sigma(\Phi_x(t\,\alpha))\,\alpha, \quad \forall\,\alpha \in \mathbb{R}^m.$$

If $t = 0$ and using $\Phi_x(0) = x$, we have $D\Phi_x(0)\ \alpha = \sigma(x)\ \alpha$ for any $\alpha \in \mathbb{R}^m$ which implies $D\Phi_x(0) = \sigma(x)$. The last statement follows easily, since Φ_x is surjective by definition. $\square$

Lemma 3.3. *For $u : \Omega \to \mathbb{R}$ $\mathcal{X}$-convex, consider the following local definition of subdifferential*

$$\widehat{\partial}_{\mathcal{X}} u(x) := \left\{ p \in \mathbb{R}^m \,|\, u(y) \geq u(x) + p\cdot\Phi_x^{-1}(y), \forall\, y \in \Omega \cap \mathbb{V}_x \cap B_R(x) \right\},$$

for some $R > 0$. Then $\partial_{\mathcal{X}} u(x) = \widehat{\partial}_{\mathcal{X}} u(x)$.

Proof. It is obvious that $\partial_{\mathcal{X}} u(x) \subset \widehat{\partial}_{\mathcal{X}} u(x)$, so we have to show only the reverse inclusion. Let us fix some ball $B_R(x)$ with radius R and centered at x and let us consider a point $z \in \big(\Omega \cap \mathbb{V}_x\big)\backslash B_R(x)$. Since $z \in \mathbb{V}_x$ there exist $\alpha \in \mathbb{R}^m$ and $x_\alpha(\cdot)$ $\mathcal{X}$-line such that $z = x_\alpha(1)$ and $x_\alpha(0) = x$; moreover $u(x_\alpha(t))$ is convex in t. Since the $\mathcal{X}$-lines are continuous, we can find λ close enough to 0 such that $y := x_\alpha(\lambda\,0 + (1-\lambda)\,1) \in \Omega \cap B_R(x)$. By the local definition of $\mathcal{X}$-subdifferential and using that $y \in \Omega \cap B_R(x) \cap \mathbb{V}_x$, $p \in \widehat{\partial}_{\mathcal{X}} u(x)$ implies

$$u(y) \geq u(x) + p\cdot\Phi_x^{-1}(y).$$

To conclude we apply the convexity of $u \circ x_\alpha$ and remark that $\Phi_x^{-1} \circ x_\alpha(t)$ is linear in t (in fact $\Phi_x^{-1} \circ x_\alpha(t) = \Phi_x^{-1} \circ \Phi_x(\alpha\,t) = \alpha\,t$). Therefore

$$\lambda\,u(x) + (1-\lambda)\,u(z) \geq u(y) \geq u(x) + p\cdot\big(\lambda\,\Phi_x^{-1}(x) + (1-\lambda)\,\Phi_x^{-1}(z)\big),$$

which gives, by using $\Phi_x^{-1}(x) = 0$ and dividing by $1-\lambda$,

$$u(z) \geq u(x) + p\cdot\Phi_x^{-1}(z),$$

and this implies $p \in \partial_{\mathcal{X}} u(x)$. $\square$

The previous lemma implies that the invertibility assumption on Φ_x can be replaced by local invertibility, which holds as soon as $X_1, \ldots, X_m$ are linearly independent.

Next we show that $\partial_{\mathcal{X}} u$ is an upper semicontinuous set-valued map.

Lemma 3.4. *Assume Φ_x is locally invertible for all $x \in \Omega$ and $u : \Omega \to \mathbb{R}$ is continuous and $\mathcal{X}$-convex. Then the $\mathcal{X}$-subdifferential map of u is closed, i.e.,*

$$x_n \to x,\ p_n \in \partial_{\mathcal{X}} u(x_n) \ \text{ and } \ p_n \to p \quad \Longrightarrow \quad p \in \partial_{\mathcal{X}} u(x).$$

Proof. By Lemma 3.3 we can assume $\Omega = \mathbb{R}^n$ and Φ_x invertible everywhere. By definition of $\mathcal{X}$-subdifferential, if $p_n \in \partial_{\mathcal{X}} u(x_n)$ then

$$u(y) \geq u(x_n) + p_n\cdot\Phi_{x_n}^{-1}(y), \quad \forall\, y \in \mathbb{V}_{x_n}. \tag{15}$$

The idea is to pass to the limit in the previous inequality. We first show that " $\lim_n \mathbb{V}_{x_n} = \mathbb{V}_x$ " which means that $y \in \mathbb{V}_x$ if and only if $y = \lim_{n\to+\infty} y_n$ with $y_n \in \mathbb{V}_{x_n}$. In fact, for any $y_n \in \mathbb{V}_{x_n}$ with $x_n \to x$, by the continuity of the $\mathcal{X}$-lines w.r.t. the initial condition, we have $y_n = x_\alpha(1; x_n) \to x_\alpha(1; x) = y \in \mathbb{V}_x$, as $n \to +\infty$. Viceversa, let us consider $y \in \mathbb{V}_x$, then there exists $\alpha \in \mathbb{R}^m$ such that $y = x_\alpha(1; x)$. For any $x_n \to x$, we look at $y_n = x_\alpha(1; x_n)$ and we get $y_n \to y$. Therefore we can pass to the limit as $n \to +\infty$ in (15) and use the continuity of $u(x)$ and the continuity of $\Phi_x^{-1}(y)$ in (x, y) to find

$$u(y) \geq u(x) + p \cdot \Phi_x^{-1}(y), \quad \forall\, y \in \mathbb{V}_x,$$

i.e. $p \in \partial_{\mathcal{X}} u(x)$. □

We are now ready to prove the main result of this section.

Proof of Theorem 3.1. By Lemma 3.2 Φ_x is locally invertible around 0 and by Lemma 3.3 we can assume that Φ_x is globally invertible, without loss of generality. Lemma 3.4 says that $\partial_{\mathcal{X}} u(\cdot)$ is closed, so we only have to find suitable sequences $x_n \to x$ and $p_n \in \partial_{\mathcal{X}} u(x_n)$ such that $|p_n| \leq L$ for some constant $L > 0$. We use the property of continuous functions that the set of points where there exist test functions touching from above is dense in the domain, and then apply Proposition 6.1 in [4] to such test functions.

We first build the sequence of approximating points. Fix a point $x_0 \in \Omega$ and look at the function $\varphi_\varepsilon(x) := \frac{|x - x_0|^2}{2\,\varepsilon}$ (where we set $\varepsilon = \frac{1}{n}$). Clearly $\varphi_\varepsilon \in C^\infty$; so if we fix a closed ball $\overline{B_R}(x_0)$ there is a maximum point x_ε for $u - \varphi_\varepsilon$. Since u is continuous, it is bounded on any closed ball and this implies that $x_\varepsilon \to x_0$ as $\varepsilon \to 0^+$ and moreover $x_\varepsilon \in B_R(x_0)$ for $\varepsilon > 0$ sufficiently small (see [3], Lemma II.1.8 for more details). By subtracting from φ_ε the constant $(\varphi_\varepsilon - u)(x_\varepsilon)$ we get a test function, that we still call φ_ε, touching u from above at the point x_ε.

Now we show that the $\mathcal{X}$-gradient of a test function touching from above a $\mathcal{X}$-convex function u is in the $\mathcal{X}$-subgradient of the function u at the touching point, i.e.,

$$p_\varepsilon := D_{\mathcal{X}} \varphi_\varepsilon(x_\varepsilon) \in \partial_{\mathcal{X}} u(x_\varepsilon). \tag{16}$$

If we prove the claim (16) then we are done, because $|p_\varepsilon| \leq L$ on $B_R(x_0)$ by Proposition 6.1 of [4], with $L > 0$ constant which may depend on R and x_0 but not on ε. Let us recall that for $\mathcal{X}$-convex functions the $\mathcal{X}$-subgradient can be defined locally. We assume by contradiction that $p_\varepsilon \notin \partial_{\mathcal{X}} u(x_\varepsilon)$, i.e.,

$$\exists\, z \in \mathbb{V}_{x_\varepsilon} \cap B_R(x_0)\ :\ u(z) < u(x_\varepsilon) + p_\varepsilon \cdot \Phi_{x_\varepsilon}^{-1}(z). \tag{17}$$

Then there exists $\alpha \in \mathbb{R}^m$ and x_α $\mathcal{X}$-line such that $x_\alpha(0) = x_\varepsilon$ and $x_\alpha(1) = z$. Let us consider the functions $u_\alpha := u \circ x_\alpha$ and $\psi_\alpha := \varphi_\varepsilon \circ x_\alpha$. The assumption (17) can be written as

$$u_\alpha(1) < u_\alpha(0) + p_\varepsilon \cdot \alpha. \tag{18}$$

We set $r_\alpha := p_\varepsilon \cdot \alpha = \psi'_\alpha(0)$; the strict inequality in (18) means that there is $\delta > 0$ such that

$$u_\alpha(1) - u_\alpha(0) < r_\alpha - \delta.$$

Since u_α is convex the slope of the corresponding secant line is non decreasing, so for $t > 0$

$$\frac{u_\alpha(-t) - u_\alpha(0)}{-t - 0} = \frac{u_\alpha(0) - u_\alpha(-t)}{t} \leq \frac{u_\alpha(1) - u_\alpha(0)}{1} < r_\alpha - \delta$$

and then

$$u_\alpha(-t) > u_\alpha(0) - t r_\alpha + t\delta.$$

Since ψ_α is touching u_α from above at 0, i.e. $\psi_\alpha \geq u_\alpha$ near 0 and $\psi_\alpha(0) = u_\alpha(0)$,

$$\psi_\alpha(-t) > \psi_\alpha(0) - t r_\alpha + t\delta. \tag{19}$$

Moreover ψ_α is C^1, so we can write its Taylor's expansion of order 1, i.e.,

$$\psi_\alpha(-t) = \psi_\alpha(0) - t\,\psi'_\alpha(0) + o(t) = \psi_\alpha(0) - t\, r_\alpha + o(t) \quad \text{as } t \to 0+\,. \tag{20}$$

Putting together (19) and (20) we find $o(t) > t\delta$, which gives the desired contradiction and concludes the proof. □

The converse of Theorem 3.1 is easier to show and it was proved in [15] (Proposition 10.5) in the case of Carnot groups.

Theorem 3.2. *If $u : \Omega \to \mathbb{R}$ has $\partial_{\mathcal{X}} u(x) \neq \emptyset$ for all $x \in \Omega$, then u is $\mathcal{X}$-convex on Ω.*

Proof. We fix $x_0 \in \Omega$, $\alpha \in \mathbb{R}^m$, and must show that $u \circ x_\alpha$ is convex, i.e.

$$u\big(x_\alpha\big(\lambda\, t_1 + (1-\lambda)\, t_2\big)\big) \leq \lambda u(x_\alpha(t_1)) + (1-\lambda) u(x_\alpha(t_2))$$

for any $\lambda \in (0,1)$ and for any $t_1, t_2 \in I_{max}$. Let us define

$$\begin{aligned} x_\lambda &:= x_\alpha\big(\lambda\, t_1 + (1-\lambda)\, t_2\big),\\ x_1 &:= x_\alpha(t_1),\\ x_2 &:= x_\alpha(t_2). \end{aligned}$$

Note that, since x_1, x_2, x_λ belong to the same $\mathcal{X}$-line starting from the point x_0, then $x_1, x_2 \in \mathbb{V}_{x_\lambda}$. Since the $\mathcal{X}$-subdifferential is everywhere non empty, there exists $p_\lambda \in \mathcal{X}_{x_\lambda}$ such that

$$u(y) \geq u(x_\lambda) + \big\langle p_\lambda, \Phi^{-1}_{x_\lambda}(y)\big\rangle_{\mathcal{X}}, \quad \forall\, y \in \mathbb{V}_{x_\lambda}.$$

We write this inequality for $y = x_1$ and $y = x_2$ and combine the two inequalities to get

$$\begin{aligned} &\lambda u(x_1) + (1-\lambda) u(x_2) \geq \\ &\qquad \lambda u(x_\lambda) + \lambda \big\langle p_\lambda, \Phi^{-1}_{x_\lambda}(x_1)\big\rangle_{\mathcal{X}} + (1-\lambda) u(x_\lambda) + (1-\lambda)\big\langle p_\lambda, \Phi^{-1}_{x_\lambda}(x_2)\big\rangle_{\mathcal{X}} \\ &\qquad\qquad = u(x_\lambda) + \big\langle p_\lambda, \lambda \Phi^{-1}_{x_\lambda}(x_1) + (1-\lambda)\Phi^{-1}_{x_\lambda}(x_2)\big\rangle_{\mathcal{X}}. \end{aligned} \tag{21}$$

To conclude it remains to prove that $\lambda \Phi^{-1}_{x_\lambda}(x_1) + (1-\lambda)\Phi^{-1}_{x_\lambda}(x_2) = 0$. Let us write $t_\lambda = \lambda\, t_1 + (1-\lambda)\, t_2$, so $x_\lambda = x_\alpha(t_\lambda)$. We first need to reparametrize the $\mathcal{X}$-line $x_\alpha(t)$ so that the starting point is x_λ, namely,

$$\widetilde{x}_\alpha(s) := x_\alpha(s + t_\lambda), \quad s \in \mathbb{R}.$$

Then

$$\begin{aligned} x_1 = x_\alpha(t_1) = \widetilde{x}_\alpha(t_1 - t_\lambda) = \widetilde{x}_{(t_1 - t_\lambda)\alpha}(1) \quad &\Rightarrow \quad \Phi^{-1}_{x_\lambda}(x_1) = (t_1 - t_\lambda)\,\alpha;\\ x_2 = x_\alpha(t_2) = \widetilde{x}_\alpha(t_2 - t_\lambda) = \widetilde{x}_{(t_2 - t_\lambda)\alpha}(1) \quad &\Rightarrow \quad \Phi^{-1}_{x_\lambda}(x_2) = (t_2 - t_\lambda)\,\alpha; \end{aligned}$$

Hence

$$\lambda \Phi^{-1}_{x_\lambda}(x_1) + (1-\lambda)\Phi^{-1}_{x_\lambda}(x_2) = \alpha\big[\lambda\,(t_1 - t_\lambda) + (1-\lambda)\,(t_2 - t_\lambda)\big] = 0,$$

which concludes the proof. □

We conclude the section by looking at the $\mathcal{X}$-subdifferential of $\mathcal{X}$-directionally differentiable functions (see Definition 3.2).

Proposition 3.1. *If $u : \Omega \to \mathbb{R}$ is $\mathcal{X}$-directionally differentiable at some point $x \in \Omega$, then*

$$\partial_{\mathcal{X}} u(x) \neq \emptyset \quad \Rightarrow \quad \partial_{\mathcal{X}} u(x) = \{D_{\mathcal{X}} u(x)\}\,.$$

Proof. Let us assume that there exists $q \in \partial_{\mathcal{X}} u(x)$. Then, for $\mathcal{X}$-lines with $x_\alpha(0) = x$,

$$u(x_\alpha(t)) \geq u(x) + q \cdot \Phi_x^{-1}(x_\alpha(t)), \quad t \in I_{max} \text{ and } \forall\, \alpha \in \mathbb{R}^m.$$

Recall that $x_\alpha(t) = x_{t\alpha}(1)$, so $\Phi_x^{-1}(x_\alpha(t)) = t\alpha$, which means

$$q \cdot \alpha\, t \leq u(x_\alpha(t)) - u(x), \quad t \in I_{max} \text{ and } \forall\, \alpha \in \mathbb{R}^m.$$

Therefore

$$\begin{aligned} q\cdot\alpha &\le \frac{u(x_\alpha(t))-u(x)}{t} \qquad \forall\, t>0,\\ q\cdot\alpha &\ge \frac{u(x_\alpha(t))-u(x)}{t} \qquad \forall\, t<0. \end{aligned}$$

Taking the limits as $t\to 0^+$ and $t\to 0^-$ we can conclude

$$q\cdot\alpha \le p\cdot\alpha \le q\cdot\alpha \quad\Rightarrow\quad q=p.$$

where $p=D_{\mathcal{X}}u(x)$. $\square$

In the case of $\mathcal{X}$-convex functions, we know that the $\mathcal{X}$-subdifferential is always nonempty, so the previous result can be rewritten as follows.

Corollary 3.1. *If $u:\Omega\to\mathbb{R}$ is $\mathcal{X}$-directionally differentiable and $\mathcal{X}$-convex, then $\partial_{\mathcal{X}}u(x)=\{D_{\mathcal{X}}u(x)\}$ at any $x\in\Omega$.*

4 Two applications to Carnot-type vector fields.

Throughout this section we assume that $X_1,\dots,X_m$ are Carnot-type vector fields, see Definition 2.4.

4.1 Fenchel transform

The next definition extends to vector fields of Carnot-type the notion of Legendre-Fenchel transform introduced by Calogero and Pini [9] in the Heisenberg group.

Definition 4.1. *The Fenchel transform of $u:\Omega\to\mathbb{R}$ is the family of functions $\{u_x^*\}_{x\in\mathbb{R}^n}$ where*

$$u_x^*(p) := \sup_{y\in\mathbb{V}_x\cap\Omega}[p\cdot\pi_m y - u(y)], \qquad p\in\mathbb{R}^m.$$

Note that $u_x^*:\mathbb{R}^m\to\mathbb{R}\cup\{+\infty\}$ is convex. Then it has the classical Legendre-Fenchel transform, or convex conjugate,

$$(u_x^*)^*(q) := \sup_{p\in\mathbb{R}^m}[q\cdot p - u_x^*(p)], \qquad q\in\mathbb{R}^m.$$

The next result states that this iterated transform is involutive if computed at $q=\pi_m x$. It extends to general Carnot-type fields one of the main results found in [9] for the Heisenberg group.

Theorem 4.1. *i) For any function $u:\Omega\to\mathbb{R}$ and for all $x\in\Omega$*

$$(u_x^*)^*(\pi_m x)\le u(x) \tag{22}$$

ii) if u is continuous then $(u_x^*)^*(\pi_m x)=u(x)$ if and only if $\partial_{\mathcal{X}}u(x)\neq\emptyset$;
iii) a continuous function u is $\mathcal{X}$-convex if and only if

$$(u_x^*)^*(\pi_m x)=u(x) \qquad \forall\, x\in\Omega. \tag{23}$$

Proof. i) By definition of u_x^*, for all $p\in\mathbb{R}^m$ and $y\in\mathbb{V}_x\cap\Omega$

$$u_x^*(p)+u(y)\ge p\cdot\pi_m y.$$

Then

$$u(y)\ge \sup_{p\in\mathbb{R}^m}[p\cdot\pi_m y - u_x^*(p)] = (u_x^*)^*(\pi_m y), \qquad \forall y\in\mathbb{V}_x\cap\Omega,$$

and by choosing $y = x$ we get (22).

ii) The definitions give

$$(u_x^*)^*(\pi_m x) = \sup_{p\in\mathbb{R}^m} \{p \cdot \pi_m x - \sup_{y\in\mathbb{V}_x\cap\Omega} [p \cdot \pi_m y - u(y)]\}$$

$$= \sup_{p\in\mathbb{R}^m} \inf_{y\in\mathbb{V}_x\cap\Omega} [p \cdot \pi_m(x - y) + u(y) - u(x)] + u(x).$$

By Remark 3.3 , if $\partial_{\mathcal{X}} u(x) \neq \emptyset$ there is $\bar{p} \in \mathbb{R}^m$ such that $\bar{p} \cdot \pi_m(x - y) + u(y) - u(x) \geq 0$ for all $y \in \mathbb{V}_x \cap \Omega$. Then $(u_x^*)^*(\pi_m x) \geq u(x)$.

Conversely, $(u_x^*)^*(\pi_m x) \geq u(x)$ implies $\sup_{p\in\mathbb{R}^m} \inf_{y\in\mathbb{V}_x\cap\Omega}[p \cdot \pi_m(x - y) + u(y) - u(x)] \geq 0$. Then for all $\varepsilon > 0$ there exists $p_\varepsilon \in \mathbb{R}^m$ such that

$$u(y) - u(x) \geq p_\varepsilon \cdot \pi_m(y - x) - \varepsilon \qquad \forall y \in \mathbb{V}_x \cap \Omega. \tag{24}$$

We claim that, for some C, $|p_\varepsilon| \leq C$ for all $\varepsilon \in]0, 1]$. Then there is a sequence $\varepsilon_k \to 0$ such that $p_{\varepsilon_k} \to \bar{p}$. By passing to the limit in (24) we get

$$u(y) - u(x) \geq \bar{p} \cdot \pi_m(y - x) \qquad \forall y \in \mathbb{V}_x \cap \Omega,$$

thus $\partial_{\mathcal{X}} u(x) \neq \emptyset$.

To prove the claim we choose $y = \Phi_x(\alpha)$ in (24), so that $\pi_m(x - y) = \alpha$ by Lemma 2.2. For $|\alpha| = 1$ the continuity of u implies that $u(y)$ is bounded. Then there is C such that

$$C \geq p_\varepsilon \cdot \alpha \qquad \forall\, |\alpha| = 1,\ \varepsilon \in]0, 1],$$

and so $|p_\varepsilon| \leq C$ for all $\varepsilon \in]0, 1]$.

iii) By Theorems 3.1 and 3.2 a continuous function $u : \Omega \to \mathbb{R}$ is $\mathcal{X}$-convex if and only if $\partial_{\mathcal{X}} u(x) \neq \emptyset$ for all $x \in \Omega$. Then the conclusion follows from *ii)* . □

4.2 A Jensen-type inequality

In this section μ is a given positive and finite measure on Ω.

Notation. If $u : \Omega \to \mathbb{R}$ and $0 < \mu(\Omega) < +\infty$ we set

$$⨍_\Omega u \, d\mu := \frac{1}{\mu(\Omega)} \int_\Omega u \, d\mu.$$

We briefly recall that for standard convex functions in the Euclidean setting Jensen's inequality states

$$⨍_\Omega u(y) \, d\mu(y) \geq u\left(⨍_\Omega y \, d\mu(y)\right). \tag{25}$$

Its proof is based on integrating the inequality

$$u(y) \geq u(x) + p \cdot (y - x), \quad \forall\, y \in \Omega,$$

where p is any element of the classical subdifferential of u at x. By Theorem 3.1 we can apply the same idea to (continuous) $\mathcal{X}$-functions, and get the following inequality

$$\int_{\mathbb{V}_x\cap\Omega} u \, d\mu_x \geq u(x) + p_x \cdot \int_{\mathbb{V}_x\cap\Omega} \Phi_x^{-1}(y) \, d\mu_x(y), \tag{26}$$

where $p_x \in \partial_{\mathcal{X}} u(x)$ and μ_x is the renormalized projection of μ on $\mathbb{V}_x$. In the classical case the integral on the right hand side vanishes if we choose $x = x_b = ⨍_\Omega y \, d\,\mu$ because $\Phi_x^{-1}(y) = y - x$ and $\mathbb{V}_x = \mathbb{R}^n$. In the general case of (26), instead, the measure μ_x depends on x and the integrals are on the lower-dimensional sets $\mathbb{V}_x$. Nevertheless, in the case of Carnot-type vector fields, the special structure of $\mathbb{V}_x$ allows to derive a weak version of Jensen's inequality.

Theorem 4.2. *Let $X_1, \dots, X_m$ be Carnot-type vector fields and μ be a measure with $0 < \mu(\Omega) < +\infty$. Assume that $\Omega = \Omega_1 \times \Omega_2$ with $\Omega_1 \subset \mathbb{R}^m$ and $\Omega_2 \subset \mathbb{R}^{n-m}$, $\mu = \mu_1 \times \mu_2$ with $\mu_1 = \mu\big|_{\Omega_1}$ and $\mu_2 = \mu\big|_{\Omega_2}$, and $\fint_{\Omega_1} \mathrm{y}^1 d\mu_1 \in \Omega_1$ (e.g., Ω_1 is convex). If $u : \mathbb{R}^n \to \mathbb{R}$ is a continuous $\mathcal{X}$-convex function, then*

$$\fint_{\Omega} u \, d\mu \geq \fint_{\Omega_2} u \left(\fint_{\Omega_1} \mathrm{y}^1 d\mu_1 \ , \ \mathrm{y}^2 \right) d\mu_2, \tag{27}$$

with $y = (\mathrm{y}^1, \mathrm{y}^2) \in \Omega_1 \times \Omega_2$ and where by y^1 and y^2 we mean respectively the first m components and the last $n - m$ components of y.

Proof. By Lemma 2.2 $y = (\mathrm{y}^1, \mathrm{y}^2) \in \mathbb{V}_x$ if and only if $\mathrm{y}^2 = \mathrm{x}^2 + C(\mathrm{y}^1, \mathrm{x}^1)$. Then by Theorem 3.1 there exists p_x such that

$$u(\mathrm{y}^1, \mathrm{y}^2) = u(\mathrm{y}^1, \mathrm{x}^2 + C(\mathrm{y}^1, \mathrm{x}^1)) \geq u(\mathrm{x}^1, \mathrm{x}^2) + p_x \cdot (\mathrm{y}^1 - \mathrm{x}^1), \quad \forall \, \mathrm{y}^1 \in \Omega_1.$$

We choose $\mathrm{x}^1 = \mathrm{x}_b^1 := \fint_{\Omega_1} \mathrm{y}^1 \, d\mu_1$ and integrate the previous inequality in $d\mu_1(\mathrm{y}^1)$ to get

$$\fint_{\Omega_1} u(\mathrm{y}^1, \mathrm{x}^2 + C(\mathrm{y}^1, \mathrm{x}_b^1)) \, d\,\mu_1(\mathrm{y}^1) \geq u(\mathrm{x}_b^1, \mathrm{x}^2), \quad \forall \, \mathrm{x}^2 \in \Omega_2.$$

By integrating the previous inequality in $d\mu_2(\mathrm{x}^2)$ we find

$$\fint_{\Omega} u(\mathrm{y}^1, \mathrm{x}^2 + C(\mathrm{y}^1, \mathrm{x}_b^1)) \, d\mu(\mathrm{y}^1, \mathrm{x}^2) \geq \fint_{\Omega_2} u \left(\fint_{\Omega_1} \mathrm{y}^1 d\mu_1(\mathrm{y}^1) \ , \ \mathrm{x}^2 \right) d\mu_2(\mathrm{x}^2). \tag{28}$$

Now we define the following change of variables from $\mathbb{R}^n$ into itself:

$$(\mathrm{y}^1, \mathrm{y}^2) = T(\mathrm{y}^1, \mathrm{x}^2) = (\mathrm{y}^1, \mathrm{x}^2 + C(\mathrm{y}^1, \mathrm{x}_b^1)). \tag{29}$$

(Remember that now x_b^1 is fixed.) Since the function $C(\cdot)$ depends only on the first m components of y, then

$$J\,T(y) = \begin{pmatrix} Id_{m \times n} & \mathbf{0}_{m \times (n-m)} \\ \star & Id_{(n-m) \times (n-m)} \end{pmatrix}.$$

We can observe that $J\,T(y)$ is a triangular matrix where all the coefficients on the diagonal are equal to 1; so $|\det J\,T| = 1$. Therefore the inequality (28) can be rewritten as (27). $\square$

Remark 4.1. Note that in general the inequality (27) cannot be improved. Take for instance $m = 1$ and $n = 2$, i.e. $X_1(x^1, x^2) = (1, 0)^T$ on $\mathbb{R}^2$, $\Omega = [a, b] \times [c, d]$, $d\,\mu = \frac{1}{|\Omega|} \, d\,x^1 \, d\,x^2$ where $d\,x^1$ and $d\,x^2$ are standard Lebesgue measures, and $u(x^1, x^2) = x^1 - (x^2)^2$.

Example 4.1. Let $X_1, \dots, X_m$ be linearly independent and constant (see Example 2.1), and Ω and μ be as in the last theorem. Then the proof of (27) is easier and can be generalised to any Radon measure by using the Disintegration Theorem (see, e.g., Theorem 2.28 in [1]).

An immediate consequence of this theorem is the following Jensen inequality for functions that do not exceed their restriction to the first m components, e.g., $u(\mathrm{x}^1, \mathrm{x}^2) \geq u(\mathrm{x}^1, \mathrm{x}_o^2)$ for some $\mathrm{x}_o^2 \in \mathbb{R}^{n-m}$.

Corollary 4.1. *Under the assumptions of Theorem 4.2 suppose also that $u(\mathrm{x}^1, \mathrm{x}^2) \geq \tilde{u}(\mathrm{x}^1)$ for all $(\mathrm{x}^1, \mathrm{x}^2) \in \Omega$, for some $\tilde{u} : \mathbb{R}^m \to \mathbb{R}$. Then*

$$\fint_{\Omega} u \, d\mu \geq \tilde{u} \left(\fint_{\Omega_1} \mathrm{y}^1 d\mu_1 \right).$$

To show that our inequality is optimal in the case of Carnot-type vector fields, we introduce and study in the next section affine functions with respect to $\mathcal{X}$ and we verify that for them the inequality (28) holds as identity.

5 $\mathcal{X}$-affine functions.

In this section we first introduce a notion of $\mathcal{X}$-superdifferential dual of the $\mathcal{X}$-subdifferential, prove some regularity of functions that have both these objects nonempty at some point, and then introduce three notions of affine functions with respect to the vector fields, compare them and study their properties.

5.1 General vector fields

Definition 5.1. *The* $\mathcal{X}$-superdifferential of $u:\Omega\to\mathbb{R}$ at x *is the set*

$$\overline{\partial_{\mathcal{X}}}u(x) := -\partial_{\mathcal{X}}(-u)(x).$$

All the properties proved for the $\mathcal{X}$-subdifferential are still true if we consider the $\mathcal{X}$-superdifferential, in particular we can still use the identification with the corresponding subset of $\mathbb{R}^m$, i.e.

$$\overline{\partial_{\mathcal{X}}}u(x) \leftrightarrow \big\{p\in\mathbb{R}^m \,|\, u(y)\le u(x)+p\cdot\Phi_x^{-1}(y), \forall\, y\in\Omega\cap\mathbb{V}_x\big\} =: \overline{\widetilde{\partial}_{\mathcal{X}}}u(x), \tag{30}$$

(see Remark 3.2) and the local expression proved in Lemma 3.3. Of course all the properties of the $\mathcal{X}$-subdifferential proved in the previous sections have an analogue for the $\mathcal{X}$-superdifferential. In particular, $\overline{\partial_{\mathcal{X}}}u(x)\neq\emptyset$ for all $x\in\Omega$ if the function u is $\mathcal{X}$-concave, i.e., $-u$ is $\mathcal{X}$-convex.

We next prove some additional useful properties for the $\mathcal{X}$-subdifferential and the $\mathcal{X}$-superdifferential.

Proposition 5.1. *If* $u:\Omega\to\mathbb{R}$ *is continuous, then*

- **(i)** $\partial_{\mathcal{X}}u(x)$ *and* $\overline{\partial_{\mathcal{X}}}u(x)$ *are closed and convex subsets of* $\mathbb{R}^m$ *(by using the identifications* (14) *and* (30)*);*
- **(ii)** *if* $\partial_{\mathcal{X}}u(x)\neq\emptyset$ *and* $\overline{\partial_{\mathcal{X}}}u(x)\neq\emptyset$ *at* $x\in\Omega$*, then* u *is* $\mathcal{X}$*-directionally differentiable at* x *and*

$$\partial_{\mathcal{X}}u(x)=\overline{\partial_{\mathcal{X}}}u(x)=\big\{D_{\mathcal{X}}u(x)\big\};$$

- **(iii)** *if* $\partial_{\mathcal{X}}u(x)\neq\emptyset$ *and* $\overline{\partial_{\mathcal{X}}}u(x)\neq\emptyset$ *for all* $x\in\Omega$*, then* $D_{\mathcal{X}}u$ *is a continuous fumction on* Ω*; if, in addition,* $u\in C^{\alpha}_{loc}(\Omega)$ *then* $D_{\mathcal{X}}u$ *is also locally* α*-Hölder continuous.*

Proof. Note that $\partial_{\mathcal{X}}u(x)$ and $\overline{\partial_{\mathcal{X}}}u(x)$ are closed by Lemma 3.4. To show the convexity of the $\mathcal{X}$-subdifferential take $p_1,p_2\in\partial_{\mathcal{X}}u(x)$ and look at

$$p_\lambda := \lambda\, p_1+(1-\lambda)p_2,\quad \lambda\in(0,1).$$

By definition

$$u(y)\ge u(x)+p_i\cdot\Phi_x^{-1}(y),\quad y\in\mathbb{V}_x\cap\Omega,\quad 1=1,2.$$

Taking a convex combination we get

$$u(y)\ge u(x)+\big[\lambda\, p_1+(1-\lambda)\,p_2\big]\cdot\Phi_x^{-1}(y)=u(x)+p_\lambda\cdot\Phi_x^{-1}(y),$$

which means that $p_\lambda\in\partial_{\mathcal{X}}u(x)$.

Next we prove (ii). We first prove that for any $p\in\partial_{\mathcal{X}}u(x)$ and $q\in\overline{\partial_{\mathcal{X}}}u(x)$ we have $p=q$; then it is easy to conclude. By the definitions

$$\begin{aligned}(I):\quad & p\in\partial_{\mathcal{X}}u(x)\ \Rightarrow\ u(y)\ge u(x)+p\cdot\Phi_x^{-1}(y),\quad y\in\mathbb{V}_x\cap\Omega,\\ (II):\quad & q\in\overline{\partial_{\mathcal{X}}}u(x)\ \Rightarrow\ u(y)\le u(x)+q\cdot\Phi_x^{-1}(y),\quad y\in\mathbb{V}_x\cap\Omega.\end{aligned}$$

Recall that whenever $y\in\mathbb{V}_x$ there exists $\alpha\in\mathbb{R}^m$ such that $y=x_\alpha(1)$ and $\Phi_x^{-1}(y)=\alpha$; hence $(II)-(I)$ implies

$$(q-p)\cdot\alpha\le 0,\quad \forall\,\alpha\in\mathbb{R}^m.$$

This is possible only if $p=q$ and we have proved that $\partial_{\mathcal{X}}u(x)=\overline{\partial_{\mathcal{X}}}u(x)=\{p\}$. It remains to show the limit (13) (see Definition 3.2). If $y\in\mathbb{V}_x$ we can write $y=x_\alpha(t)$ for some $\alpha\in\mathbb{R}^m$, $t\in\mathbb{R}$ and $x_\alpha(\cdot)$ $\mathcal{X}$-line. In this case $\Phi_x^{-1}(y)=\alpha\, t$. Combining now (I) and (II) with $p=q$, we find, for all t small enough,

$$\frac{u(x_\alpha(t))-u(x)}{t}=p\cdot\alpha,\quad \forall\alpha\in\mathbb{R}^m.$$

So by taking the limit as $t\to 0^+$ we can conclude that Definition 3.2 is satisfied and $p=D_{\mathcal{X}}u(x)$.

Finally, (iii) follows from the last identity, by recalling that the $\mathcal{X}$-lines depend locally in a Lipschitz way on the initial position x and u is continuous. ☐

The following definition is a very general geometric notion of affine function which holds w.r.t. any given family of vector fields.

Definition 5.2. *A continuous function $A:\Omega\to\mathbb{R}$ is $\mathcal{X}$-affine if $A\circ x_\alpha$ is (Euclidean) affine for any $x_\alpha(\cdot)$ $\mathcal{X}$-line contained in Ω.*

Note that A $\mathcal{X}$-affine means that it is at the same time $\mathcal{X}$-convex and $\mathcal{X}$-concave.

For $\mathcal{X}$-affine functions, we can give a PDE characterization which follows directly from the characterization of $\mathcal{X}$-convex and $\mathcal{X}$-concave functions in terms of matrix inequalities in viscosity sense proved in [4], Theorem 3.1.

Proposition 5.2. *If $X_1,\dots,X_m$ are C^2-vector fields and $A:\Omega\to\mathbb{R}$ is continuous, then A is $\mathcal{X}$-affine if and only if $-D^2_{\mathcal{X}}A(x)\le 0$ and $-D^2_{\mathcal{X}}A(x)\ge 0$, i.e.*

$$a^TD^2_{\mathcal{X}}A(x)a=0,\quad \forall\, a\in\mathbb{R}^m,$$

in the viscosity sense.

Lemma 5.1. *Any $\mathcal{X}$-affine function A is $\mathcal{X}$-directionally differentiable and $D_{\mathcal{X}}A$ is continuous.*

Proof. Since A is $\mathcal{X}$-convex and $\mathcal{X}$-concave, Theorem 3.1 implies that $\partial_{\mathcal{X}}A(x)\neq\emptyset$ and $\overline{\partial_{\mathcal{X}}}A(x)\neq\emptyset$ at any point $x\in\Omega$. We can so directly conclude applying Proposition 5.1, properties (ii) and (iii). ☐

Looking at our definition of $\mathcal{X}$-subdifferential, there is another very natural definition for affine functions.

Definition 5.3. *Assume that Φ_x defined in (9) is invertible (e.g. in the case of linearly independent vector fields). We say that a continuous function $A:\Omega\to\mathbb{R}$ is Φ-affine if for any $x_0\in\Omega$ there exist $\beta_{x_0}\in\mathbb{R}$ and $p_{x_0}\in\mathcal{X}_{x_0}=\mathrm{Span}\big(X_1(x_0),\dots,X_m(x_0)\big)$ such that*

$$A(x)=\beta_{x_0}+\big\langle p_{x_0},\Psi_{x_0}(x)\big\rangle_{\mathcal{X}},\quad \forall\, x\in\Omega\cap\mathbb{V}_{x_0},\tag{31}$$

with $\langle\cdot,\cdot\rangle_{\mathcal{X}}$ and $\Psi_x(\cdot)$ defined as in Definition 3.3.

Remark 5.1. By using the same identification between $p_x\in\mathcal{X}_x$ and its coordinate vector $(p_x^1,\dots,p_x^m)\in\mathbb{R}^m$, i.e. $p_x=\sum_{i=1}^m p_x^iX_i(x)$, we can say that a continuous function $A:\Omega\to\mathbb{R}$ is Φ-affine if and only if for any $x_0\in\Omega$ there exist $\beta_{x_0}\in\mathbb{R}$ and $p_{x_0}\in\mathbb{R}^m$ such that

$$A(x)=\beta_{x_0}+p_{x_0}\cdot\Phi_{x_0}^{-1}(x),\quad \forall\, x\in\Omega\cap\mathbb{V}_{x_0}.\tag{32}$$

As we did in the case of the $\mathcal{X}$-subdifferential, for sake of simplicity we will use definition (32) instead of the original definition (31).

Remark 5.2. Since $x_0\in\mathbb{V}_{x_0}$ and $\Phi_{x_0}^{-1}(x_0)=0$, $\beta_{x_0}=A(x_0)$.

Under suitable assumptions, the two definitions of affine functions introduced above are indeed equivalent.

Proposition 5.3. *Let $X_1, \dots, X_m$ be such that the associated function Φ_x defined in* (9) *is invertible and $A : \Omega \to \mathbb{R}$ be continuous. Then A is Φ-affine if and only if it is $\mathcal{X}$-affine.*

Proof. For sake of simplicity we take $\Omega = \mathbb{R}^n$. We first show that the representation formula (32) implies that A is $\mathcal{X}$-affine. Given $\alpha \in \mathbb{R}^m$, $x_0 \in \Omega$ and $x_\alpha(\cdot)$ corresponding $\mathcal{X}$-line, we need to prove that if A is given by (32) then $A \circ x_\alpha(t)$ is (Euclidean) affine in t. In fact,

$$A \circ x_\alpha(t) = \beta_{x_0} + p_{x_0} \cdot \Phi^{-1}_{x_0}(x_\alpha(t)) = \beta_0 + t\,(p_0 \cdot \alpha).$$

since $x_0 = x_\alpha(0)$ and $x_\alpha(t) = x_{t\,\alpha}(1)$, so $\Phi^{-1}_{x_0}(x_\alpha(t)) = x_0$. Then $A \circ x_\alpha$ is affine in t.

For the reverse implication we use Lemma 5.1. Let us fix $x_0 \in \Omega$ and consider a $\mathcal{X}$-line $x_\alpha(\cdot)$ with starting point x_0, then $A \circ x_\alpha(t)$ is affine in t. This implies that there exists a constant $\lambda \in \mathbb{R}$ such that

$$\lambda = \frac{d}{dt} A \circ x_\alpha(t) = D_{\mathcal{X}} A(x_\alpha(t)) \cdot \alpha;$$

taking in particular $t = 0$, we get $\lambda = D_{\mathcal{X}} A(x_0) \cdot \alpha$. Therefore we can conclude that for any $x \in \mathbb{V}_{x_0}$ (i.e. $x = x_\alpha(t)$ for some t)

$$A(x) = A(x_0) + \big(D_{\mathcal{X}} A(x_0) \cdot \alpha\big)\, t = A(x_0) + D_{\mathcal{X}} A(x_0) \cdot \Phi^{-1}_{x_0}(x),$$

with $\beta_{x_0} = A(x_0)$ and $p_{x_0} = D_{\mathcal{X}} A(x_0)$ (we have used $\alpha\, t = \Phi^{-1}_{x_0}(x)$ whenever $x = x_\alpha(t)$). ☐

5.2 Carnot-type vector fields

In the particular case of Carnot groups a different notion of affine functions was previously introduced and studied: the H-affine functions [15]. This notion can be generalized to any family of Carnot-type vector fields since one only needs $\Phi^{-1}_{x_0}(x) = \pi_m(x - x_0) = \pi_m(x) - \pi_m(x_0)$.

Definition 5.4. *Let $\mathcal{X} = \{X_1, \dots, X_m\}$ be a family of Carnot-type vector fields. A continuous function $A : \Omega \to \mathbb{R}$ is* horizontally affine *if there exist two constants $\beta \in \mathbb{R}$ and $p \in \mathbb{R}^m$ such that*

$$A(x) = \beta + p \cdot \pi_m(x), \quad \forall\, x \in \Omega. \tag{33}$$

This definition allows to give a nice characterization of any $\mathcal{X}$-convex function as a suitable envelope of horizontally affine ones, reminiscent of the Euclidean case. This property was called *abstract convexity* in [9] for the Heisenberg group.

Proposition 5.4. *For Carnot-type vector fields, $u \in C(\Omega)$ is $\mathcal{X}$-convex if and only if for all $x \in \Omega$*

$$u(x) = \max\{A(x) : A \text{ horizontally affine, } A(y) \le u(y) \ \ \forall\, y \in \mathbb{V}_x \cap \Omega\}. \tag{34}$$

Proof. Assume (34) holds, fix $x \in \Omega$, and let $\bar{A}$ be the affine functions where the maximum is attained. If we rewrite $\bar{A}(y) = u(x) + \bar{p} \cdot \pi_m(y - x)$ we get

$$u(y) \ge u(x) + \bar{p} \cdot \pi_m(y - x) \qquad \forall\, y \in \mathbb{V}_x \cap \Omega, \tag{35}$$

so $\partial_{\mathcal{X}} u(x) \ne \emptyset$. By Theorem 3.2 we can conclude that u is $\mathcal{X}$-convex.

Viceversa, if $u \in C(\Omega)$ is $\mathcal{X}$-convex by Theorem 3.1 for any x there is $\bar{p}$ such that (35) holds. We define $\bar{A}(y) := u(x) + \bar{p} \cdot \pi_m(y - x)$ so that $u(x) = \bar{A}(x)$ and $u \ge A$ on $\mathbb{V}_x \cap \Omega$. Then (34) holds. ☐

Next we compare the last definition with the preceding notions of generalized affine function. Clearly horizontally affine functions are Euclidean affine. Moreover, any horizontally affine function is Φ-affine (and so also $\mathcal{X}$-affine). In fact, it is sufficient to choose $p_{x_0} = p$ and $\beta_{x_0} = -p \cdot \pi_m(x_0) + \beta$, for any x_0.

However, the reverse implication is not always true. For example we can consider $X_1(x^1, x^2) = (1, 0)^t$ on $\mathbb{R}^2$ (so $(x^1, x^2) \in \mathbb{V}_{(x_0^1, x_0^2)}$ if and only if $x^2 = x_0^2$) and look at

$$A(x^1, x^2) = \beta + p\, x^2 x^1 = \beta + (p\, x^2) \cdot \pi_1(x^1, x^2).$$

It is easy to check that A is Φ-affine but it is not horizontal affine.

The reason why the equivalence fails in the previous example is that different $\mathcal{X}$-planes are disjoint and this is related to the existence of points that cannot be connected by an admissible trajectory (2). We next give a sufficient condition for the equivalence between the functions given by (33) and those given by (31). Define the Carnot-Carathéodory distance in Ω associated to $X_1, \dots, X_m$ as

$$d_\Omega(x,y) := \inf \{T \geq 0 \,|\, \exists\, \gamma : [0,T] \to \Omega \text{ admissible, } \gamma(0) = x,\ \gamma(T) = y\}, \tag{36}$$

with the definition of admissible curve given in the Introduction. In the next theorem we assume that

$$d_\Omega(x,y) < +\infty \quad \forall\, x, y \in \Omega. \tag{37}$$

We recall that the last condition holds in any open set Ω if the vector fields are smooth and satisfy the *Hörmander condition* in Ω, i.e., they and their iterated Lie brackets generate the whole space $\mathbb{R}^n$ at every $x \in \Omega$.

Lemma 5.2. *If A is horizontally affine, then $p_{x_0} = p_x$ for any $x \in \mathbb{V}_{x_0}$.*

Proof. By Remark 5.2 we can write

$$A(x) = A(x_0) + p_{x_0} \cdot \pi_m(x - x_0), \quad \forall\, x \in \mathbb{V}_{x_0}. \tag{38}$$

Note that $x \in \mathbb{V}_{x_0}$ implies $x_0 \in \mathbb{V}_x$, by a simple reparametrization for the $\mathcal{X}$-line joining x_0 to x. Therefore $A(x_0) = A(x) + p_x \cdot \pi_m(x_0 - x)$. Adding up the last two identities gives $0 = (p_{x_0} - p_x) \cdot \pi_m(x - x_0)$, for any $x \in \mathbb{V}_{x_0}$, and this means $p_{x_0} = p_x$ for any $x \in \mathbb{V}_x$. ☐

Theorem 5.1. *Let $\mathcal{X}$ be a family of Carnot-type vector fields satisfying* (37)*. Then a function $A : \Omega \to \mathbb{R}$ is horizontally affine if and only if it is Φ-affine (and hence $\mathcal{X}$-affine).*

Proof. We only have to prove that the class of functions satisfying (38) verifies also condition (33). Let x_0, x be any pair of points in Ω; by the assumption (37) there exists a horizontal curve joining x_0 to x, that is, an absolutely continuous curve solving

$$\dot{\gamma}(t) = \sum_{i=1}^{m} \alpha_i(t) X_i(\gamma(t)), \quad \gamma(t) \in \Omega,$$

for a suitable m-valued bounded measurable function $\alpha(t) = (\alpha_1(t), \dots, \alpha_m(t))$. It is known that the horizontal curve $\gamma(t)$ can be uniformly approximated by piecewise $\mathcal{X}$-lines, by approximating the measurable function $\alpha(t)$ by piecewise constant functions. Then there is a sequence $x^{(k)} \in \Omega$ such that $|x^{(k)} - x| < 1/k$ and $x^{(k)}$ is the endpoint of a concatenation of $\mathcal{X}$-lines. Denote with x_j, $j = 0, \dots, N_k$, the vertices of the polygonal and set $x^{(k)} := x_{N_k}$. Since $x_{j+1} \in \mathbb{V}_{x_j}$, Lemma 5.2 and (38) give

$$p_{x_{j+1}} = p_{x_j}, \quad A(x_{j+1}) = A(x_j) + p_{x_j} \cdot \pi_m(x_{j+1} - x_j), \quad j = 0, \dots, N_k - 1.$$

By iterating we get

$$p_{x^{(k)}} = p_{x_0}, \quad A(x^{(k)}) = A(x_{N_k}) = A(x_0) + p_{x_0} \cdot \pi_m(x^{(k)} - x_0). \tag{39}$$

Now observe that A Φ-affine implies it is $\mathcal{X}$-affine, by Proposition 5.3. Then Lemma 5.1 gives $p_x = D_{\mathcal{X}} A(x)$ continuous in x. Therefore letting $k \to \infty$ in (39) we obtain $p_x = p_{x_0} =: p$ independent of x and $A(x) = A(x_0) + p \cdot \pi_m(x - x_0)$, which completes the proof. ☐

We finally verify that for Carnot-type vector fields the Jensen-type inequality (28) holds as identity for Φ-affine functions, and therefore also for $\mathcal{X}$-affine functions.

Proposition 5.5. *Assume* $X_1, \ldots, X_m$ *are Carnot-type vector fields,* $\Omega = \Omega_1 \times \Omega_2 \subset \mathbb{R}^m \times \mathbb{R}^{n-m}$ *open, and* $\mu = \mu_1 \times \mu_2$ *where* μ_i *is a measure on* Ω_i *such that* $0 < \mu_i(\Omega_i) < +\infty$*; let* $A : \Omega \to \mathbb{R}$ *be a* Φ*-affine function and assume* $\fint_{\Omega_1} \mathrm{y}^1 d\mu_1 \in \Omega_1$*. Then*

$$\fint_{\Omega} A \, d\mu = \fint_{\Omega_2} A\left(\fint_{\Omega_1} \mathrm{y}^1 d\mu_1 \, , \, \mathrm{y}^2 \right) d\mu_2. \tag{40}$$

Proof. The particular structure of $\mathbb{V}_x$ in the case of Carnot-type vector fields found in Lemma 2.2 gives

$$A(\mathrm{y}^1, C(\mathrm{y}^1, \mathrm{x}^1) + \mathrm{x}^2) = \beta_{(\mathrm{x}^1,\mathrm{x}^2)} + p_{(\mathrm{x}^1,\mathrm{x}^2)} \cdot (\mathrm{y}^1 - \mathrm{x}^1),$$

where $x = (\mathrm{x}^1, \mathrm{x}^2) \in \Omega_1 \times \Omega_2$ and $y = (\mathrm{y}^1, \mathrm{y}^2)$. We will use it for $x = (\mathrm{x}^1_b, \mathrm{x}^2)$ with $\mathrm{x}^1_b = \fint_{\Omega_1} \mathrm{y}^1 d\mu_1$. By the definition of $T(\cdot)$ given in (29) and the fact that $|\det T| = 1$ (see the proof of Theorem 4.2 for more details), for Φ-affine functions the left-hand side of (40) is equal to

$$\begin{aligned} \fint_{\Omega} A \, d\,\mu = \fint_{\Omega} & A(\mathrm{y}^1, C(\mathrm{y}^1, \mathrm{x}^1_b) + \mathrm{y}^2) \, d\mu \\ & = \fint \left(\beta_{(\mathrm{x}^1_b,\mathrm{y}^2)} + p_{(\mathrm{x}^1_b,\mathrm{y}^2)} \cdot (\mathrm{y}^1 - \mathrm{x}^1_b) \right) d\mu = \fint_{\Omega_2} \beta_{(\mathrm{x}^1_b,\mathrm{y}^2)} \, d\mu_2. \end{aligned}$$

The right-hand side of (40) is given by

$$\fint_{\Omega_2} \left(\beta_{(\mathrm{x}^1_b,\mathrm{y}^2)} + p_{(\mathrm{x}^1_b,\mathrm{y}^2)} \cdot \left(\fint_{\Omega_1} \mathrm{y}^1 \, d\mu_1 - \mathrm{x}^1_b \right) \right) d\mu_2 = \fint_{\Omega_2} \beta_{(\mathrm{x}^1_b,\mathrm{y}^2)} \, d\mu_2,$$

thus the identity is verified. ☐

References

[1] Ambrosio, L.; Fusco, N.; Pallara D.; *Functions of bounded variation and free discontinuity problems.* Clarendon Press, New York, 2000.

[2] Balogh, Z.; Rickly, M.; *Regularity of convex functions on Heisenberg groups.* Ann. Sc. Norm. Super. Pisa Cl. Sci. (5), 2 (2003), no. 4, 847–868.

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

[4] Bardi, M.; Dragoni, F; *Convexity and semiconvexity along vector fields.* Calc. Var. Partial Differential Equations 42 (2011), 405–427.

[5] Bardi, M.; Mannucci, P.; *Comparison principles for subelliptic equations of Monge-Ampère type.* Boll. Unione Mat. Ital. 9, no.1 (2008), 2, 489–495.

[6] Bardi, M.; Mannucci, P.; *Comparison principles and Dirichlet problem for fully nonlinear degenerate equations of Monge-Ampère type.* Forum Math. to appear.

[7] Bellaïche, A.; Risler, J.J. (Editors); *Sub-Riemannian geometry.* Progress in Mathematics 144, Birkhäuser Verlag, Basel, 1996.

[8] Bonfiglioli, A.; Lanconelli, E.; Uguzzoni, F.; *Stratified Lie Groups and Potential Theory for their Sub-Laplacians.* Springer, Berlin 2007.

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

[10] Calogero, A.; Pini, R.; *Horizontal normal map on the Heisenberg group.* J. Nonlinear Convex Analysis 12 (2011), 287–307.

[11] Capogna, L.; Pauls, S.D.; Tyson, J.T.; *Convexity and horizontal second fundamental forms for hypersurfaces in Carnot groups.* Trans. Amer. Math. Soc. 362 (2010), 4045–4062

[12] G. Citti, M. Manfredini, A. Sarti. *From neural oscillations to variational problems in the visual cortex.* J. Physiology 97 (2003), 87-385.

[13] G. Citti, M. Manfredini, A. Sarti. *Neuronal oscillation in the visual cortex: Gamma-convergence to the Riemannian Mumford-Shah functional.* SIAM J. Math. Analysis 35 (2004), 1394-1419.

[14] G. Citti, A. Sarti. *A cortical based model of perceptual completion in the roto-traslation space.* J. Math. Imaging Vision 24 (2006), 307–326.

[15] Danielli, D.; Garofalo, N.; Nhieu, D.M.; *Notions of convexity in Carnot groups.* Comm. Anal. Geom. 11 (2003), 263–341.

[16] Garofalo, N.; Tournier, F.; *New properties of convex functions in the Heisenberg group.* Trans. Amer. Math. Soc. 358 , no. 5, (2006), 2011–2055 (electronic).

[17] Gutiérrez, C.E.; Montanari, A.; *Maximum and comparison principles for convex functions on the Heisenberg group.* Comm. Partial Differential Equations 29, no. 9-10, (2004),1305–1334.

[18] Gutiérrez, C.E.; Montanari, A.; *On the second order derivatives of convex functions on the Heisenberg group.* Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 3 (2004), no. 2, 349–366.

[19] Juutinen, P.; Lu, G.; Manfredi, J.J.; Stroffolini, B.; *Convex functions on Carnot groups.* Rev. Mat. Iberoam. 23 (2007), no. 1, 191–200.

[20] Lu, G.; Manfredi, J.J.; Stroffolini, B.; *Convex functions in the Heisenberg group.* Calc. Var. Partial Differential Equations 19 (2004), no. 1, 1–22.

[21] Magnani, V.; *Lipschitz continuity, Aleksandrov theorem and characterizations for H-convex functions.* Math. Ann. 334 (2006), no. 1, 199–233.

[22] Magnani, V.; Scienza, M.; *Characterization of differentiability for h-convex functions in stratified groups.* Ann. Scuola Norm. Sup. Pisa Cl. Sci. to appear.

[23] Magnani, V.; Scienza, M.; *Regularity estimates for convex functions in Carnot-Carathéodory spaces.* preprint 2012

[24] Montgomery, R.; *A tour of subriemannian geometries, their geodesics and applications.* Mathematical Surveys and Monographs 91, American Mathematical Society, Providence, RI, 2002.

[25] Monti, R.; Rickly, M.; *Geodetically convex sets in the Heisenberg group.* J. Convex Anal. 12 (2005), no. 1, 187–196.

[26] Wang, C.; *Viscosity convex functions on Carnot groups.* Proc. Amer. Math. Soc. 133 (2005), no. 4, 1247–1253.