J. reine angew. Math. **750** (2019), 241–297
DOI 10.1515/crelle-2016-0051

Journal für die reine und angewandte Mathematik

Besicovitch Covering Property on graded groups and applications to measure differentiation

By *Enrico Le Donne* at Jyväskylä and *Séverine Rigot* at Nice

---

Abstract. We give a complete answer to which homogeneous groups admit homogeneous distances for which the Besicovitch Covering Property (BCP) holds. In particular, we prove that a stratified group admits homogeneous distances for which BCP holds if and only if the group has step 1 or 2. These results are obtained as consequences of a more general study of homogeneous quasi-distances on graded groups. Namely, we prove that a positively graded group admits continuous homogeneous quasi-distances satisfying BCP if and only if any two different layers of the associated positive grading of its Lie algebra commute. The validity of BCP has several consequences. Its connections with the theory of differentiation of measures is one of the main motivations of the present paper. As a consequence of our results, we get for instance that a stratified group can be equipped with some homogeneous distance so that the differentiation theorem holds for each locally finite Borel measure if and only if the group has step 1 or 2. The techniques developed in this paper allow also us to prove that sub-Riemannian distances on stratified groups of step 2 or higher never satisfy BCP. Using blow-up techniques this is shown to imply that on a sub-Riemannian manifold the differentiation theorem does not hold for some locally finite Borel measure.

Contents

The first-named author acknowledges the support of the Academy of Finland project no. 288501. The second-named author is partially supported by ANR grants ANR-12-BS01-0014-01 and ANR-15-CE40-0018.

1. Introduction

Covering theorems are known to be among the fundamental tools in analysis and geometry. They reflect, in a certain sense, the geometry of the space and are commonly used to establish connections between local and global properties. All covering theorems are based on the same principle: from an arbitrary cover of a set, one tries to extract a subcover that is as disjointed as possible. Among classical covering theorems, the Besicovitch Covering Property (BCP), in which we are interested in the present paper, originates from works of A. Besicovitch in connection with the theory of differentiation of measures in Euclidean spaces ([4,5], see also [11, Section 2.8], [31] and Section 6).

The development of analysis and geometry on abstract metric spaces leads naturally to the question of the validity of suitable covering theorems on non-Euclidean spaces. Graded groups provide a natural framework for many developments. A graded group is a Lie group equipped with an appropriate family of dilations. A homogeneous quasi-distance on a graded group is a left-invariant quasi-distance that is one-homogeneous with respect to the family of dilations. Graded groups equipped with homogeneous quasi-distances naturally generalize finite-dimensional normed vector spaces. Due to the presence of translations and dilations, they provide a setting where many aspects of classical analysis and geometry can be carried out. Beyond such a priori considerations, these spaces form an important framework because of their occurrences in many settings. There is a characterization of positively graduable Lie groups as connected locally compact groups admitting contractive automorphisms ([36], see also Theorem 2.8). The interest of E. Siebert for groups admitting contractive automorphisms was motivated by phenomenon appearing in probability theory on groups. Such groups have more generally been considered by various authors. In particular, homogeneous groups equipped with homogeneous quasi-distances as considered in [12] (see also Definition 2.21) fit within this framework. They have been considered mainly in connection with their applications in harmonic analysis, complex analysis of several variables, and study of some non-elliptic differential operators. We refer to [12] and the references therein for a detailed presentation of these aspects. A class of homogeneous groups equipped with homogeneous distances of particular interest are stratified groups equipped with sub-Riemannian distances. They are also known as Carnot groups according to a terminology due to P. Pansu. See Section 7 where we use the more explicit terminology sub-Riemannian Carnot groups. One of their occurrences is as metric tangent spaces to sub-Riemannian manifolds where they play in some sense the role Euclidean spaces play in Riemannian geometry (see e.g. [3, 26]). These structures have also connections with optimal control theory (see e.g. [1, 27]).

In the present paper, we give a complete answer to which graded groups admit continuous homogeneous quasi-distances for which BCP holds, see Theorem 1.2. Characterizations of homogeneous and stratified groups admitting homogeneous distances satisfying BCP follow as particular cases of our results on graded groups, see Corollary 1.3 and Corollary 1.4. We also complete previous results about the non-validity of BCP for sub-Riemannian distances on stratified groups of step ≥ 2, see Theorem 1.9. Finally, we give applications to measure differentiation which is one of the main motivations for this paper, see Theorem 1.5, Theorem 1.6, Corollary 1.7, and Theorem 1.10.

To explain these results, we first recall the Besicovitch Covering Property (BCP) in the general quasi-metric setting. We refer to Section 3 for a more detailed discussion about this covering property. See also Section 2.5 for our conventions about quasi-metric spaces, which are the classical ones. A quasi-metric space (X, d) satisfies BCP if there exists a constant $N \geq 1$ such that the following holds. For any bounded set $A \subset X$ and any family $\mathcal{B}$ of balls such that each point of A is the center of some ball of $\mathcal{B}$, there is a finite or countable subfamily $\mathcal{F} \subset \mathcal{B}$ such that the balls in $\mathcal{F}$ cover A, and every point in X belongs to at most N balls in $\mathcal{F}$.

We briefly recall now the definitions of graded groups and homogeneous quasi-distances. We refer to Section 2 for a complete presentation. A graded group is a simply connected Lie group G whose Lie algebra $\mathfrak{g}$ is endowed with a positive grading $\mathfrak{g} = \bigoplus_{t \in (0,+\infty)} V_t$ where $[V_s, V_t] \subset V_{s+t}$ for all $s, t > 0$. At the level of the Lie algebra, the associated dilation of factor λ is defined as the unique linear map $\delta_\lambda : \mathfrak{g} \to \mathfrak{g}$ such that $\delta_\lambda(X) = \lambda^t X$ for all $X \in V_t$. Denoting also by δ_λ the unique Lie group homomorphism induced by this Lie algebra homomorphism, a quasi-distance d on G is said to be homogeneous if it is left-invariant and one-homogeneous with respect to the associated family of dilations (where the latter means $d(\delta_\lambda(p), \delta_\lambda(q)) = \lambda d(p, q)$ for all $p, q \in G$ and all $\lambda > 0$).

To state the main results of this paper and for later convenience, we introduce the following definition that will be shown to algebraically characterize the validity of BCP for some homogeneous quasi-distances.

Definition 1.1 (Graded groups with commuting different layers). Let G be a graded group and let $\bigoplus_{t>0} V_t$ be the associated positive grading of its Lie algebra. We say that G has *commuting different layers* if $[V_t, V_s] = \{0\}$ for all $t, s > 0$ such that $t \neq s$.

Our main results read as follows.

Theorem 1.2. *Let G be a graded group. There exist continuous homogeneous quasi-distances on G for which BCP holds if and only if G has commuting different layers.*

Homogeneous groups are those graded groups that can be equipped with homogeneous distances, i.e., homogeneous quasi-distances that satisfy the triangle inequality. Equivalently, all layers V_t with $t < 1$ of the associated positive grading of their Lie algebra are $\{0\}$, see Definition 2.21 and Proposition 2.22. For such groups, we prove that homogeneous distances are continuous, see Corollary 2.28, and we get the following corollary.

Corollary 1.3. *Let G be a homogeneous group. There exist homogeneous distances on G for which BCP holds if and only if G has commuting different layers.*

An important class of homogeneous groups are stratified groups, which are those for which the degree-one layer of the associated positive grading generates the Lie algebra, see Definitions 2.4 and 2.5. A stratified group has commuting different layers if and only if it has (nilpotency) step 1 or 2. For such groups, Corollary 1.3 reads as follows.

Corollary 1.4. *Let G be a stratified group. There exist homogeneous distances on G for which BCP holds if and only if G is of step ≤ 2.*

One of the main motivations for studying the validity of BCP on graded groups is its connection with the theory of differentiation of measures. If μ is a locally finite Borel measure on a metric space (X, d), we say that the differentiation theorem holds on (X, d) for μ if

$$\lim_{r \downarrow 0^+} \frac{1}{\mu(B_d(p,r))} \int_{B_d(p,r)} f(q)\, d\mu(q) = f(p)$$

for μ-almost every $p \in X$ and all $f \in L^1_{\mathrm{loc}}(\mu)$. For homogeneous groups equipped with homogeneous distances, we prove the following characterization.

Theorem 1.5. *Let G be a homogeneous group and let d be a homogeneous distance on G. The differentiation theorem holds on (G, d) for all locally finite Borel measures if and only if (G, d) satisfies BCP.*

This characterization is a consequence of a characterization of the validity of the differentiation theorem for all locally finite Borel measures on metric spaces due to D. Preiss, taking the additional structure into account, namely, using left-translations and dilations, see Section 6. Together with Corollary 1.3 and Corollary 1.4, we get the following results.

Theorem 1.6. *Let G be a homogeneous group. There exists some homogeneous distance d on G such that the differentiation theorem holds on (G, d) for all locally finite Borel measures if and only if G has commuting different layers.*

Corollary 1.7. *Let G be a stratified group. There exists some homogeneous distance d on G such that the differentiation theorem holds on (G, d) for all locally finite Borel measures if and only if G is of step ≤ 2.*

To put our results in perspective, let us first recall that it is well known since the works of Besicovitch in the 1940s that BCP holds in Euclidean spaces, and more generally in finite-dimensional normed vector spaces. It is also known that the Riemannian distance on a Riemannian manifold of class C^2 satisfies a property that generalizes BCP, see [11, Section 2.8] and Sections 6 and 7. On the contrary, it is also well known that BCP does not hold on infinite-dimensional normed vector spaces. Until recently only few results were known for graded groups equipped with homogeneous quasi-distances. It was proved independently and at the same time in [34] and [17] that BCP does not hold on the stratified Heisenberg groups equipped with the Cygan–Korányi distance. Later it was proved in [33] that BCP also fails for sub-Riemannian distances on stratified groups under some regularity assumptions on the sub-Riemannian distance. After these negative answers about the validity of BCP, it was commonly believed that there would probably not exist homogeneous quasi-distances satisfying BCP on graded groups.

Let us now recall that any two homogeneous distances on a homogeneous group, and more generally any two homogeneous quasi-distances on a graded group, are biLipschitz equivalent. However, it turns out that the validity of BCP is not stable under a biLipschitz change of quasi-distance.

Theorem 1.8 ([20, Theorem 1.6]). *Let (X, d) be a metric space. Assume that there exists an accumulation point in (X, d). Then, for all $0 < c < 1$, there exists a distance d_c on X such that $c\, d \leq d_c \leq d$ and for which BCP does not hold.*

See also [31, Theorem 3] from which Theorem 1.8 is inspired. Notice that Theorem 1.8 can be extended to quasi-distances. As a consequence the non-validity of BCP for some homogeneous quasi-distances cannot give any hint towards the existence or non-existence of some other homogeneous quasi-distance satisfying BCP. Since for many purposes the choice of a specific quasi-distance up to biLipschitz equivalence does not really matter, the question of the existence of some homogeneous quasi-distance for which BCP holds on a graded group remained meaningful.

The present paper follows two previous papers, [20] and [21]. The existence of some homogeneous distances that satisfy BCP on the stratified Heisenberg groups is proved in [20]. On the contrary it is proved in [21] that natural analogues of these distances on stratified groups of step ≥ 3 do not satisfy BCP. These two cases strongly suggested that the structure of the dilations, which comes from the structure of the grading of the Lie algebra, plays a crucial role. Theorem 1.2 characterizes precisely in which sense the structure of the grading plays a role for our purposes.

Let us now say few words about the proof of Theorem 1.2. Among the simplest examples of positively graduable groups are the Abelian ones, the Heisenberg groups, and free-nilpotent groups of step 2, see Example 2.10, Example 2.11, and Section 4.1. They play a key role in our proof of Theorem 1.2. A first step is indeed the study of the validity or non-validity of BCP for homogeneous quasi-distances relatively to various possible positive gradings on these groups.

To prove the existence of homogeneous quasi-distances satisfying BCP on graded groups with commuting different layers, we first prove that Hebisch–Sikora's quasi-distances satisfy BCP on stratified free-nilpotent groups of step 2, see Theorem 4.5. Hebisch–Sikora's quasi-distances are those homogeneous quasi-distances whose unit ball centered at the identity is a Euclidean ball, see Examples 2.23 and 2.36. Theorem 4.5 extends [20, Theorem 1.14] to stratified free-nilpotent groups of step 2 and any rank $r \geq 2$. Theorem 1.14 in [20] gives indeed the conclusion for the stratified first Heisenberg group, i.e., the stratified free-nilpotent group of step 2 and rank 2. The proof of Theorem 4.5 is in spirit inspired by the proof of this previous result but requires a slightly different approach.

To prove the non-existence of continuous homogeneous quasi-distances satisfying BCP on graded groups for which there exists two different layers that do not commute, we first consider the case of the non-standard Heisenberg groups, see Theorem 5.6. A non-standard Heisenberg group is the first Heisenberg group viewed as a graded group whose Lie algebra is endowed with a positive grading that is not a stratification, see Example 2.11. As already mentioned, the fact that there exists no continuous homogeneous quasi-distance satisfying BCP on non-standard Heisenberg groups was suggested by [21, Theorem 1.6]. However, we stress that the proof of Theorem 5.6 is not a technical modification of the arguments in [21]. It requires indeed a completely new approach, see Section 5.

In both cases, the general conclusion, see Theorems 4.1 and 5.2, follows using structure properties of graded groups, using submetries, that plays a central role here, and using some constructions on metric spaces that preserve the validity of BCP. These tools are given in Sections 2 and 3.

In Section 2 we establish preliminary results about graded groups. In Sections 2.1 and 2.2 we fix the definitions and the terminology we shall use throughout the paper for graded and stratified Lie algebras and Lie groups and for the associated families of dilations. Section 2.3 is devoted to various examples and to the description of some constructions on graded groups for later use. In particular, the Heisenberg groups and various positive gradings of their Lie algebra to be used later in the paper are given in Example 2.11. In Section 2.4, we prove some structure properties for graded groups. Proposition 2.15 gives a description of graded groups with commuting different layers. Proposition 2.18 explains how every graded group with some different layers not commuting gives rise a non-standard Heisenberg group. Section 2.5 is devoted to homogeneous quasi-distances. The meaning of the terminology homogeneous groups that we use in this paper is in particular given in Definition 2.21. We stress that working with quasi-distances rather than with distances naturally occurs in applications but may lead to topological issues. In our setting, we prove that homogeneous quasi-distances on graded groups induce the manifold topology, see Proposition 2.26. As a consequence, homogeneous distances on homogeneous groups are continuous, see Corollary 2.28. These results seem not to have been previously noticed in the literature and may be of independent interest. On the contrary we stress that homogeneous quasi-distances may or may not be continuous and Proposition 2.29 characterizes continuous homogeneous quasi-distances.

In Section 3 we first recall general facts about the Besicovitch Covering Property (BCP) and one of its variants which we call the Weak Besicovitch Covering Property (WBCP). We remark that these two variants are not equivalent in general metric spaces, see Example 3.4. However, in our setting, and more generally on doubling metric spaces, BCP and WBCP are equivalent, see Proposition 3.7. It turns out that working with WBCP is for our purposes technically more convenient. Next, we consider some constructions that preserve the validity of (W)BCP. In particular, products of metric spaces are considered in Theorem 3.16. The role of surjective morphisms of Lie algebra and submetries is given in Propositions 3.20 and 3.21. Results in this section will be used together with the structure properties proved in Section 2.4 to deduce Theorem 1.2 from the particular cases mentioned above.

Sections 4 and 5 are devoted to the proof of Theorem 1.2 together with Corollaries 1.3 and 1.4. In Section 4 we prove the existence of continuous homogeneous quasi-distances for which BCP holds on graded groups with commuting different layers, following the scheme already described above, see Theorem 4.1. In Section 5 we consider more general quasi-distances, called self-similar, which are only required to be one-homogeneous with respect to some dilation, see Definition 5.1. We prove that continuous self-similar quasi-distances do not satisfy BCP on graded groups for which there exist two different layers of the associated positive grading that do not commute, see Theorem 5.2. Self-similar, rather than homogeneous, quasi-distances may occur naturally. We stress that in this case, additional topological issues have to be taken into account, see Section 5.2.

In Section 6 we give applications to measure differentiation on graded groups. We prove Theorem 1.5, using the notion of σ-finite dimensionality.

In Section 7 we consider sub-Riemannian distances on stratified groups. We complete the results of [33] with the following general negative answer. We refer to Definition 7.1 for the definition of sub-Riemannian Carnot groups.

Theorem 1.9. *Let (G, d) be a sub-Riemannian Carnot group of step ≥ 2. Then BCP does not hold on (G, d).*

The proof of this result is independent of Theorem 1.2 but uses some techniques developed in Section 3, in particular Proposition 3.21. Using the fact that sub-Riemannian Carnot groups appear as metric tangent spaces to sub-Riemannian manifolds, we get the following consequence about measure differentiation on sub-Riemannian manifolds. We refer to

Section 7.2 for the definition of sub-Riemannian manifolds, which in this paper does not include the Riemannian ones.

Theorem 1.10. *Let M be a sub-Riemannian manifold and let d be its sub-Riemannian distance. Then there exists some locally finite Borel measure for which the differentiation theorem on (M, d) does not hold.*

Acknowledgement. The authors would like to thank Tapio Rajala for fruitful conversations and improving feedback.

2. Preliminaries on graded groups

2.1. Graded and stratified Lie algebras and Lie groups. All Lie algebras considered here are over $\mathbb{R}$ and finite dimensional.

Definition 2.1 (Positively graduable Lie algebras). A *positive grading* of a Lie algebra $\mathfrak{g}$ is a family $(V_t)_{t\in(0,+\infty)}$ of vector subspaces of $\mathfrak{g}$, where all but finitely many of the vector subspaces V_t are $\{0\}$ such that

$$\mathfrak{g} = \bigoplus_{t\in(0,+\infty)} V_t$$

and where $[V_s, V_t] \subset V_{s+t}$ for all $s, t > 0$. Here $[V, W] := \operatorname{span}\{[X, Y] : X \in V,\ Y \in W\}$. We say that a Lie algebra is *positively graduable* if it admits a positive grading.

A positively graduable Lie algebra may admit several positive gradings that are not isomorphic, see for instance Example 2.11 (and Definition 2.16 for the definition of isomorphisms of graded Lie algebras). We will use the terminology "graded Lie algebra" when considering a positively graduable Lie algebra equipped with a given positive grading as stated in the following definition.

Definition 2.2 (Graded Lie algebras). We say that a Lie algebra is *graded* when it is positively graduable and endowed with a positive grading called the *associated positive grading*.

Given a positive grading $\mathfrak{g} = \bigoplus_{t>0} V_t$ and given $t \in (0, +\infty)$, the subspace V_t is called the *degree-t layer* of the grading.

Recall that, for a Lie algebra $\mathfrak{g}$, the lower central series is defined inductively by $\mathfrak{g}^{(1)} = \mathfrak{g}$, $\mathfrak{g}^{(k+1)} = [\mathfrak{g}, \mathfrak{g}^{(k)}]$. A Lie algebra $\mathfrak{g}$ is called *nilpotent* if $\mathfrak{g}^{(s+1)} = \{0\}$ for some integer $s \geq 1$. We say that $\mathfrak{g}$ is *nilpotent of step s* if $\mathfrak{g}^{(s+1)} = \{0\}$ but $\mathfrak{g}^{(s)} \neq \{0\}$. Positively graduable Lie algebras are nilpotent. However, nilpotent Lie algebras that are not positively graduable do exist. Regarding this last statements and more properties, see [9, 12, 13, 19].

Definition 2.3 (Stratifiable Lie algebras). A *stratification* of step s of a Lie algebra $\mathfrak{g}$ is a direct-sum decomposition $\mathfrak{g} = V_1 \oplus V_2 \oplus \cdots \oplus V_s$ for some integer $s \geq 1$ where $V_s \neq \{0\}$, $[V_1, V_j] = V_{j+1}$ for all integers $j \in \{1, \ldots, s\}$, and where we have set $V_{s+1} := \{0\}$. We say that a Lie algebra is *stratifiable* if it admits a stratification.

Equivalently, a stratifiable Lie algebra $\mathfrak{g}$ is a positively graduable Lie algebra that admits a positive grading whose degree-one layer generates $\mathfrak{g}$ as a Lie algebra. A stratification is uniquely determined by its degree-one layer V_1. Moreover, one has $\mathfrak{g} = V_1 \oplus [\mathfrak{g}, \mathfrak{g}]$. Recall that the rank of a nilpotent Lie algebra is defined as $\dim \mathfrak{g} - \dim[\mathfrak{g}, \mathfrak{g}]$. For a stratified Lie algebra it coincides with the dimension of the degree-one layer of any of its stratification. However, we stress that an arbitrary vector subspace V of a stratifiable Lie algebra $\mathfrak{g}$ that is in direct sum with $[\mathfrak{g}, \mathfrak{g}]$, i.e., satisfies $\mathfrak{g} = V \oplus [\mathfrak{g}, \mathfrak{g}]$, may not generate a stratification, see Example 2.12. Note also that a positive grading of a stratifiable Lie algebra may not be a stratification, see Example 2.11.

Any two stratifications of a Lie algebra are isomorphic, see [19]. In particular, they have equal step that we will call the *step of the stratifiable Lie algebra*. Note that a stratifiable Lie algebra of step s is nilpotent of step s. When we fix a given stratification of a stratifiable Lie algebra, we will use the terminology "stratified Lie algebra" as stated in the following definition.

Definition 2.4 (Stratified Lie algebras). We say that a Lie algebra is *stratified* when it is stratifiable and endowed with a stratification called the *associated stratification*.

Definition 2.5 (Positively graduable, graded, stratifiable, stratified groups). We say that a Lie group G is a *positively graduable* (respectively *graded*, *stratifiable*, *stratified*) *group* if G is a connected and simply connected Lie group whose Lie algebra is positively graduable (respectively graded, stratifiable, stratified).

For the sake of completeness, we will give in Theorem 2.8 below a characterization of positively graduable groups in terms of existence of a contractive group automorphism. This characterization is due to E. Siebert.

2.2. Dilations on graded algebras and graded groups.

Definition 2.6 (Dilations on graded Lie algebras). Let $\mathfrak{g}$ be a graded Lie algebra with associated positive grading $\mathfrak{g} = \bigoplus_{t>0} V_t$. For $\lambda > 0$, we define the *associated dilation of factor* λ as the unique linear map $\delta_\lambda : \mathfrak{g} \to \mathfrak{g}$ such that $\delta_\lambda(X) = \lambda^t X$ for all $X \in V_t$.

Dilations $\delta_\lambda : \mathfrak{g} \to \mathfrak{g}$ are Lie algebra automorphisms. Moreover, the family of dilations $(\delta_\lambda)_{\lambda>0}$ is a one-parameter group of Lie algebra automorphisms, i.e., $\delta_\lambda \circ \delta_\eta = \delta_{\lambda\eta}$ for all $\lambda, \eta > 0$.

Throughout this paper, given a Lie group homomorphism $\varphi : G \to H$, we will denote by $\varphi_* : \mathfrak{g} \to \mathfrak{h}$ the associated Lie algebra homomorphism. Recall that, if G is simply connected, given a Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$, there exists a unique Lie group homomorphism $\varphi : G \to H$ such that $\varphi_* = \phi$ (see [37, Theorem 3.27]). This allows to define dilations on graded groups as stated in the following definition.

Definition 2.7 (Dilations on graded groups). Let G be a graded group with Lie algebra $\mathfrak{g}$. Let $\delta_\lambda : \mathfrak{g} \to \mathfrak{g}$ be the associated dilation of factor $\lambda > 0$. The *associated dilation of factor* λ *on* G is the unique Lie group automorphism, also denoted by $\delta_\lambda : G \to G$, such that $(\delta_\lambda)_* = \delta_\lambda$.

For technical simplicity, we keep the same notation for both dilations on the Lie algebra $\mathfrak{g}$ and the group G. There will be no ambiguity here. Indeed, graded groups being nilpotent and simply connected, the exponential map $\exp : \mathfrak{g} \to G$ is a diffeomorphism (see [7, Theorem 1.2.1] or [12, Proposition 1.2]) and one has $\delta_\lambda \circ \exp = \exp \circ \delta_\lambda$ (see [37, Theorem 3.27]), hence dilations on $\mathfrak{g}$ and dilations on G coincide in exponential coordinates.

For the sake of completeness, we give now an equivalent characterization of positively graduable groups due to Siebert. If G is a topological group with identity e and $\tau : G \to G$ is a group automorphism, we say that τ is *contractive* if, for all $g \in G$, one has

$$\lim_{k\to\infty} \tau^k(g) = e.$$

We say that G is *contractible* if G admits a contractive automorphism.

For graded groups, associated dilations of factor $\lambda \in (0,1)$ are contractive automorphisms. Hence positively graduable groups are contractible. Conversely, E. Siebert proved (see Theorem 2.8 below) that if G is a connected locally compact group and $\tau : G \to G$ is a contractive automorphism, then G is a simply connected Lie group and τ induces a positive grading on the Lie algebra $\mathfrak{g}$ of G. Note however that τ itself may not be a dilation associated to the induced grading.

Theorem 2.8 ([36, Corollary 2.4]). *A topological group G is a positively graduable Lie group if and only if G is a connected locally compact contractible group.*

2.3. Examples. We first introduce the definition of basis adapted to a positive grading for later use.

Definition 2.9. Let $\mathfrak{g}$ be a graded Lie algebra with associated positive grading given by $\mathfrak{g} = \bigoplus_{t>0} V_t$. Let $n := \dim \mathfrak{g}$ and $0 < t_1 < \cdots < t_l$ be such that $V_{t_i} \neq \{0\}$ for all $i = 1, \dots, l$ and $V_t = \{0\}$ for all $t \notin \{t_1, \dots, t_l\}$. We say that a basis $(X_1, \dots, X_n)$ of $\mathfrak{g}$ is *adapted to the positive grading* if $(X_{m_{i-1}+1}, \dots, X_{m_i})$ is a basis of V_{t_i} for all $1 \le i \le l$. Here $m_0 = 0$ and $m_i - m_{i-1} = \dim V_{t_i}$.

Example 2.10 (Abelian Lie algebras and Lie groups). Abelian Lie algebras are stratifiable Lie algebras of step 1 and any direct-sum decomposition is a positive grading.

In particular, if $\mathfrak{g}$ is an Abelian n-dimensional Lie algebra, the trivial direct-sum decomposition $\mathfrak{g} = \bigoplus_{t>0} V_t$, where $V_1 = \mathfrak{g}$ and $V_t = \{0\}$ for all $t \neq 1$ gives the stratification. The connected and simply connected Lie group with Lie algebra $\mathfrak{g}$ can be identified with $\mathbb{R}^n$ equipped with the Abelian group law. Associated dilations are the usual multiplication by a scalar positive real number given by $x \mapsto \lambda x$ for $\lambda > 0$.

More generally, for any real numbers $0 < d_1 \le \cdots \le d_n$, the maps

$$(x_1, \dots, x_n) \mapsto (\lambda^{d_1} x_1, \dots, \lambda^{d_n} x_n),$$

define a family of dilations on the Abelian group $\mathbb{R}^n$ associated to some positive grading of its Lie algebra.

Example 2.11 (Heisenberg Lie algebras and Lie groups). The n-th Heisenberg Lie algebra $\mathfrak{h}_n$ is the $(2n+1)$-dimensional Lie algebra that admits a basis $(X_1, \dots, X_n, Y_1, \dots, Y_n, Z)$

where the only non-trivial bracket relations are $[X_j, Y_j] = Z$ for all $1 \leq j \leq n$. We call such a basis a *standard basis* of $\mathfrak{h}_n$.

The n-th Heisenberg group $\mathbb{H}^n$ is the connected and simply connected Lie group whose Lie algebra is $\mathfrak{h}_n$. Using exponential coordinates of the first kind, we write $p \in \mathbb{H}^n$ as

$$p = \exp\left(\left(\sum_{j=1}^{n} x_j X_j + y_j Y_j\right) + zZ\right)$$

and we identify p with $(x_1, \dots, x_n, y_1, \dots, y_n, z)$. Using the Baker–Campbell–Hausdorff formula, the group law is given by

$$(x_1, \dots, x_n, y_1, \dots, y_n, z) \cdot (x'_1, \dots, x'_n, y'_1, \dots, y'_n, z')$$
$$= (x_1 + x'_1, \dots, x_n + x'_n, y_1 + y'_1, \dots, y_n + y'_n, z + z' + \frac{1}{2}\sum_{j=1}^{n}(x_j y'_j - y_j x'_j)).$$

Heisenberg Lie algebras are stratifiable of step 2. Using a standard basis,

$$\mathfrak{h}_n = V_1 \oplus V_2$$

is a stratification, where

$$V_1 := \operatorname{span}\{X_j, Y_j : 1 \leq j \leq n\}, \quad V_2 := \operatorname{span} Z.$$

Dilations associated to this stratification are given by

$$(x_1, \dots, x_n, y_1, \dots, y_n, z) \mapsto (\lambda x_1, \dots, \lambda x_n, \lambda y_1, \dots, \lambda y_n, \lambda^2 z).$$

Heisenberg Lie algebras also admit positive gradings that are not stratifications. We will in particular consider in this paper such gradings on the first Heisenberg Lie algebra. Namely, for $\alpha \in (1, +\infty)$, we call *non-standard Heisenberg Lie algebra of exponent* α the first Heisenberg Lie algebra equipped with the following *non-standard grading of exponent* α:

$$\mathfrak{h}_1 = W_1 \oplus W_\alpha \oplus W_{\alpha+1},$$

where

$$W_1 := \operatorname{span}\{X_1\}, \quad W_\alpha := \operatorname{span}\{Y_1\}, \quad W_{\alpha+1} := \operatorname{span}\{Z\},$$

and where (X_1, Y_1, Z) is a standard basis of $\mathfrak{h}_1$. Note that up to isomorphisms of graded Lie algebras (see Definition 2.16) and up to powers (see Example 2.14), these non-standard gradings give all the possible positive gradings of $\mathfrak{h}_1$ that are not a stratification. Dilations associated to the non-standard grading of exponent α are given by

$$(x_1, y_1, z) \mapsto (\lambda x_1, \lambda^\alpha y_1, \lambda^{\alpha+1} z).$$

We will use the terminology *non-standard Heisenberg group (of exponent* α*)* when considering the first Heisenberg group as a graded group whose Lie algebra is endowed with the non-standard grading of exponent α.

Example 2.12. There are examples of stratifiable Lie algebras $\mathfrak{g}$ for which one can find a subspace V in direct sum with $[\mathfrak{g}, \mathfrak{g}]$ but that does not generate a stratification. One can for instance consider the stratifiable Lie algebra $\mathfrak{g}$ of step 3 generated by e_1, e_2 and e_3 and with the relation $[e_2, e_3] = 0$. If $V := \operatorname{span}\{e_1, e_2 + [e_1, e_2], e_3\}$, one has $\mathfrak{g} = V \oplus [\mathfrak{g}, \mathfrak{g}]$ but V does not generate a stratification of $\mathfrak{g}$, see [19].

Example 2.13 (Direct product of graded Lie groups). The direct product of graded groups is positively graduable. If G and H are graded groups with associated positive grading of their Lie algebras given by $\mathfrak{g} = \bigoplus_{t>0} V_t$ and $\mathfrak{h} = \bigoplus_{t>0} W_t$, respectively, then $\bigoplus_{t>0}(V_t \oplus W_t)$ is a positive grading of the Lie algebra of $G \times H$. This can be extended to the direct product of finitely many graded groups in the obvious way. In the rest of this paper, we will always consider the direct product of graded groups as graded groups with associated positive grading given by the above mentioned grading.

Example 2.14 (Power of graded Lie algebras and of graded groups). Let $\mathfrak{g}$ be a graded Lie algebra with associated positive grading $\mathfrak{g} = \bigoplus_{s>0} V_s$ and let $t > 0$ be a real positive number. Then $\mathfrak{g} = \bigoplus_{s>0} W_s$, where $W_{ts} := V_s$ is a positive grading of $\mathfrak{g}$, which we call the *t-power* of the initial positive grading.

For $\lambda > 0$, let δ_λ denote the dilation of factor λ associated to the initial positive grading and $\tilde{\delta}_\lambda$ denote the dilation of factor λ associated to the grading of its t-power. In exponential coordinates of the first kind associated to a basis adapted to these gradings, we have

$$\delta_\lambda(x_1, \dots, x_n) = (\lambda^{s_1} x_1, \dots, \lambda^{s_n} x_n),$$
$$\tilde{\delta}_\lambda(x_1, \dots, x_n) = (\lambda^{ts_1} x_1, \dots, \lambda^{ts_n} x_n)$$

for some $0 < s_1 \le \dots \le s_n$.

If G is a graded group, we call *t-power of G* the group G considered as the graded group whose Lie algebra is endowed with the t-power of the initial positive grading.

2.4. Structure of graded algebras and graded groups. We give in this subsection some results about the structure of graded algebras and groups to be used later in this paper. They may be more generally of independent interest.

First, we consider graded groups with commuting different layers, see Definition 1.1. Notice that for such graded groups, a layer of the positive grading may not commute with itself.

Proposition 2.15. *Let G be a graded group with commuting different layers. Then G is the direct product of powers of stratified groups of step ≤ 2.*

See Example 2.14 for the definition of powers of a graded group, Example 2.13 for the definition of direct product graded groups and Definition 2.5 for the definition of stratified groups.

Proof. Let $\mathfrak{g}$ denote the Lie algebra of G. Let $0 < t_1 < \dots < t_m$ be such that $V_{t_k} \neq \{0\}$ for all $k = 1, \dots, m$ and $V_t = \{0\}$ for all $t \notin \{t_1, \dots, t_m\}$. We have $\mathfrak{g} = V_{t_1} \oplus V_{t_2} \oplus \dots \oplus V_{t_m}$.

If $[V_{t_1}, V_{t_1}] = \{0\}$, then $[V_{t_1}, \mathfrak{g}] = \{0\}$ since $[V_{t_1}, V_s] = \{0\}$ for all $s \neq t_1$. It follows that $\mathfrak{h} := V_{t_1}$ and $\mathfrak{h}' := V_{t_2} \oplus \dots \oplus V_{t_m}$ are ideals of $\mathfrak{g}$. Hence (see [30, p. 388]) G is the direct product of $\exp(\mathfrak{h})$ and $\exp(\mathfrak{h}')$. Moreover, $\exp(\mathfrak{h})$ is the t_1-power of an Abelian stratified group.

If $[V_{t_1}, V_{t_1}] \neq \{0\}$, we set $\mathfrak{h} := V_{t_1} \oplus [V_{t_1}, V_{t_1}] \subseteq V_{t_1} \oplus V_{2t_1}$. Then we consider V'_{2t_1} any complement of $[V_{t_1}, V_{t_1}]$ in V_{2t_1}, i.e., $[V_{t_1}, V_{t_1}] \oplus V'_{2t_1} = V_{2t_1}$. We set $V'_{t_1} = \{0\}$, $V'_t = V_t$ for $t \neq t_1, 2t_1$ and $\mathfrak{h}' := \bigoplus_{t>0} V'_t$. We have $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}'$. We prove now that G is the direct product of $\exp(\mathfrak{h})$ and $\exp(\mathfrak{h}')$. As before, to get the conclusion, we prove that both $\mathfrak{h}$ and $\mathfrak{h}'$ are ideals of $\mathfrak{g}$.

To show that $\mathfrak{h}$ is an ideal, take $X \in \mathfrak{h}$ and $Y \in \mathfrak{g}$. It is enough to consider the following four cases. First, assume $X \in V_{t_1}$ and $Y \in V_{t_1}$. Then $[X, Y] \in \mathfrak{h}$ by the definition of $\mathfrak{h}$. Second, assume $X \in V_{t_1}$ and $Y \in V_t$, with $t \neq t_1$. Then $[X, Y] = 0$ since G has commuting different layers. Third, assume $X = [X_a, X_b]$ with $X_a, X_b \in V_{t_1}$ and $Y \in V_{2t_1}$. By Jacobi's identity and since $2t_1 \neq t_1$ and G has commuting different layers, we get

$$[[X_a, X_b], Y] = [[X_a, Y], X_b] + [[Y, X_b], X_a] = 0.$$

By bilinearity of the Lie bracket, it follows that $[X, Y] = 0 \in \mathfrak{h}$ for all $X \in [V_{t_1}, V_{t_1}]$ and $Y \in V_{2t_1}$. Finally, assume $X \in [V_{t_1}, V_{t_1}]$ and $Y \in V_t$ with $t \neq 2t_1$. Then $[X, Y] = 0$ since G has commuting different layers. All together it follows that $[X, Y] \in \mathfrak{h}$ for all $X \in \mathfrak{h}$ and all $Y \in \mathfrak{g}$ hence $\mathfrak{h}$ is an ideal.

To show that $\mathfrak{h}'$ is an ideal, we note that if $X \in V_t$ with $t > t_1$ and $Y \in V_s$ for some $s \geq t_1$, we have $t + s > 2t_1$ and hence

$$[X, Y] \in \bigoplus_{l>2t_1} V_l = \bigoplus_{l>2t_1} V_l' \subset \mathfrak{h}'.$$

Since $\mathfrak{h}' \subset \bigoplus_{t>t_1} V_t$ and by bilinearity of the Lie bracket, it follows that $[\mathfrak{h}', \mathfrak{g}] \subset \mathfrak{h}'$ hence $\mathfrak{h}'$ is an ideal. It follows that G is the direct product of $\exp(\mathfrak{h})$ and $\exp(\mathfrak{h}')$. Moreover, $\exp(\mathfrak{h})$ is the t_1-power of a stratified group of step 2.

Finally, arguing by induction on the dimension of $\mathfrak{g}$, we get the conclusion. □

We will next consider the case where there exist two different layers of the positive associated grading that do not commute. We first introduce the notions of morphisms of graded Lie algebras and of graded subalgebras.

Definition 2.16 (Morphism of graded Lie algebras). Let $\mathfrak{g} = \bigoplus_{t>0} V_t$, $\mathfrak{h} = \bigoplus_{t>0} W_t$ be graded Lie algebras. We say that $\phi : \mathfrak{g} \to \mathfrak{h}$ is a *morphism* (respectively *isomorphism*) *of graded Lie algebras* if ϕ is a Lie algebra homomorphism (respectively isomorphism) such that $\phi(V_t) \subset W_t$ for all $t > 0$.

Let $\mathfrak{g} = \bigoplus_{t>0} V_t$ be a graded Lie algebra with associated dilations $(\delta_\lambda)_{\lambda>0}$. We say that a Lie subalgebra $\hat{\mathfrak{g}}$ of $\mathfrak{g}$ is *homogeneous* if $\delta_\lambda(\hat{\mathfrak{g}}) = \hat{\mathfrak{g}}$ for all $\lambda > 0$. If $\hat{\mathfrak{g}}$ is a homogeneous Lie subalgebra of $\mathfrak{g}$, then $\bigoplus_{t>0}(V_t \cap \hat{\mathfrak{g}})$ is a positive grading of $\hat{\mathfrak{g}}$ called the *induced positive grading* and associated dilations are the restriction of δ_λ to $\hat{\mathfrak{g}}$.

Definition 2.17 (Graded subalgebra). Let $\mathfrak{g}$ be a graded Lie algebra. We say that $\hat{\mathfrak{g}}$ is a *graded subalgebra of the graded algebra* $\mathfrak{g}$ if $\hat{\mathfrak{g}}$ is a homogeneous Lie subalgebra of $\mathfrak{g}$ endowed with the induced positive grading.

Proposition 2.18. *Let $\mathfrak{g} = \bigoplus_{t>0} V_t$ be a graded Lie algebra. Assume that, for some $t < s$, $[V_t, V_s] \neq \{0\}$. Then there exist a graded subalgebra $\hat{\mathfrak{g}}$ of $\mathfrak{g}$ and a surjective morphism of graded Lie algebras from $\hat{\mathfrak{g}}$ to $\mathfrak{h}$ where $\mathfrak{h}$ is the t-power of the non-standard Heisenberg Lie algebra of exponent s/t.*

See Example 2.11 for the definition of non-standard Heisenberg Lie algebras and Example 2.14 for the definition of the t-power of a graded Lie algebra.

Proof. Let $X_1 \in V_t$ and $X_2 \in V_s$ be such that $[X_1, X_2] \neq 0$. Let $\hat{\mathfrak{g}}$ denote the Lie subalgebra of $\mathfrak{g}$ generated by X_1 and X_2. We have $\delta_\lambda(\hat{\mathfrak{g}}) = \hat{\mathfrak{g}}$ for all $\lambda > 0$, where δ_λ are the associated dilations on $\mathfrak{g}$. Hence $\hat{\mathfrak{g}}$ is homogeneous. We endow it with the induced positive grading $\hat{\mathfrak{g}} = \bigoplus_{u>0} \hat{V}_u$, where $\hat{V}_u := V_u \cap \hat{\mathfrak{g}}$ to make it a graded subalgebra of $\mathfrak{g}$. We have

$$\hat{\mathfrak{g}} = \hat{V}_t \oplus \hat{V}_s \oplus \hat{V}_{t+s} \oplus \Big(\bigoplus_{u>t+s} \hat{V}_u \Big)$$

with

$$\hat{V}_t = \operatorname{span}\{X_1\}, \quad \hat{V}_s = \operatorname{span}\{X_2\}, \quad \hat{V}_{t+s} = \operatorname{span}\{X_3\},$$

where $X_3 := [X_1, X_2]$. Let $n := \dim \hat{\mathfrak{g}}$ and $(X_4, \dots, X_n)$ be a basis of $\bigoplus_{u>t+s} \hat{V}_u$ such that $(X_1, \dots, X_n)$ is a basis of $\hat{\mathfrak{g}}$ adapted to its positive grading (see Definition 2.9).

Let $\mathfrak{h} := W_t \oplus W_s \oplus W_{s+t}$ be the t-power of the non-standard Heisenberg Lie algebra of exponent s/t (see Example 2.11 and Example 2.14). Let $Y_1 \neq 0 \in W_t$, $Y_2 \neq 0 \in W_s$, and set $Y_3 := [Y_1, Y_2]$. Then $W_{s+t} = \operatorname{span}\{Y_3\}$.

Let $\phi : \hat{g} \to \mathfrak{h}$ be the linear map defined by $\phi(X_i) := Y_i$ for $i = 1, 2, 3$ and $\phi(X_i) := 0$ for $i \geq 4$. It can easily be checked that ϕ is a Lie algebra homomorphism. It can also easily be checked that $\phi(\hat{V}_u) = W_u$ for all $u > 0$. Here we set $W_u := \{0\}$ for $u \notin \{t, s, s+t\}$. Hence ϕ is a surjective morphism of graded Lie algebras. □

Remark 2.19. To conclude this subsection, let us mention the following general fact. For any graded Lie algebra $\mathfrak{g}$, there exist a positive grading of a free-nilpotent Lie algebra $\mathfrak{f}$ and a surjective morphism $\phi : \mathfrak{f} \to \mathfrak{g}$ of graded Lie algebras. We shall use this fact in the simple case of stratified Lie algebras of step 2 where all what we need can be easily constructed by hand, see the proof of Theorem 4.24. We refer to [6, Chapter II], [16], [35], [32], [8] for more details on the subject.

2.5. Homogeneous quasi-distances. Given a nonempty set X, we say that a map $d : X \times X \to [0, +\infty)$ is a quasi-distance on X if it is symmetric, $d(p, q) = 0$ if and only if $p = q$, and there exists a constant $C \geq 1$ such that $d(p, q) \leq C(d(p, p') + d(p', q))$ for all $p, p', q \in X$ (quasi-triangle inequality with multiplicative constant C). We call (X, d) a quasi-metric space. When speaking of a ball B in (X, d), it will be understood that B is a set of the form $B = B_d(p, r)$ for some $p \in X$ and some $r > 0$, where $B_d(p, r) := \{q \in X : d(q, p) \leq r\}$. When d satisfies the triangle inequality, i.e., the quasi-triangle inequality with a multiplicative constant $C = 1$, then d is a distance on X.

Definition 2.20 (Homogeneous quasi-distances on graded groups). Let G be a graded group with associated dilations $(\delta_\lambda)_{\lambda>0}$. We say that a quasi-distance d on G is *homogeneous* if d is left-invariant, i.e., $d(p \cdot q, p \cdot q') = d(q, q')$ for all $p, q, q' \in G$, and one-homogeneous with respect to all dilations $(\delta_\lambda)_{\lambda>0}$, i.e., $d(\delta_\lambda(p), \delta_\lambda(q)) = \lambda d(p, q)$ for all $p, q \in G$ and all $\lambda > 0$.

Note that, in this definition, we do not require any topological property, and in particular any continuity property, of a homogeneous quasi-distance with respect to the manifold topology on the group. We will discuss these topological issues below, see Proposition 2.26, Corollary 2.28, and Proposition 2.29. Let us stress that we will consider in Section 5 a more general class of quasi-distances, called self-similar quasi-distances in the present paper. Additional topological issues occur for self-similar quasi-distances. For the sake of clarity, we devote the present subsection to homogeneous quasi-distances. We postpone the discussion about topological properties of self-similar quasi-distances to Section 5, see especially Section 5.2.

Homogeneous quasi-distances on arbitrary graded groups do exist. One can for instance follow the arguments in [12, Chapter 1]. Note however that our terminology is slightly different from the terminology adopted for graded groups in [12]. See also Example 2.25 below for another construction of homogeneous quasi-distances on arbitrary graded groups.

On the other hand, homogeneous distances do exist if and only if, for all $t < 1$, degree-t layers of the associated positive grading are $\{0\}$. These groups are called homogeneous in [15] and we will follow here this terminology.

Definition 2.21 (Homogeneous groups). We say that G is a homogeneous group if G is a graded group whose associated positive grading $\bigoplus_{t>0} V_t$ of its Lie algebra is such that $V_t = \{0\}$ for all $t \in (0, 1)$.

As already mentioned, we have the following proposition.

Proposition 2.22. *Let G be a graded group. There exists a homogeneous distance on G if and only if G is a homogeneous group.*

Proof. If some degree-t layer of the positive grading is non-trivial for some $t < 1$, a map $d : G \times G \to [0, +\infty)$ that is left-invariant and one-homogeneous with respect to some non-trivial associated dilation cannot satisfy the triangle inequality (i.e., the quasi-triangle inequality with a multiplicative constant $C = 1$). On the other hand, W. Hebisch and A. Sikora proved in [15] the existence of homogeneous distances on homogeneous groups. □

Homogeneous distances considered by Hebisch and Sikora play a central role in Section 4 and we describe them below.

Example 2.23 (Hebisch and Sikora's homogeneous distances on homogeneous groups). Let G be a homogeneous group with identity e, associated positive grading of its Lie algebra given by $\mathfrak{g} = \bigoplus_{t>0} V_t$ and associated dilations $(\delta_\lambda)_{\lambda>0}$. Let $n := \dim \mathfrak{g}$ and let $(X_1, \dots, X_n)$ be a basis of $\mathfrak{g}$ adapted to the positive grading (see Definition 2.9). Using exponential coordinates of the first kind, we identify $p \in G$ with $(p_1, \dots, p_n)$ where $p = \exp(p_1 X_1 + \cdots + p_n X_n)$. For $R > 0$, we set

$$A_R := \left\{ p \in G : \sum_{i=1}^n p_i^2 \le R^2 \right\}$$

and

$$d_R(p, q) := \inf\{\lambda > 0 : \delta_{1/\lambda}(p^{-1} \cdot q) \in A_R\}.$$

Hebisch and Sikora proved in [15] that there exists $R^* > 0$ such that for all $0 < R < R^*$, d_R defines a homogeneous distance on G.

Example 2.24 (Homogeneous quasi-distances on powers of graded groups). If G is a graded group, d a homogeneous quasi-distance on G and $t \in (0, +\infty)$, then $d^{1/t}$ is a homogeneous quasi-distance on its t-power (see Example 2.14 for the definition of powers of graded groups). Notice that when G is a homogeneous group, d a homogeneous distance and $t > 1$, then $(G, d^{1/t})$ is a snowflake of (G, d).

Example 2.25 (Existence of homogeneous quasi-distances on graded groups). Let G be a graded group with associated positive grading $\bigoplus_{s>0} V_s$ of its Lie algebra. All t-powers of G where $t \min\{s > 0 : V_s \neq \{0\}\} \geq 1$ are homogeneous groups. If d is a homogeneous distance on such a t-power of G, it follows from Example 2.24 that d^t is a homogeneous quasi-distance on G.

A quasi-distance d on a set X induces a topology on X declaring a set O to be open if and only if for all $x \in O$ there exists $r > 0$ such that $B_d(x, r) \subset O$. Equivalently a set F is said to be closed if and only if for all sequences (x_k) of points in F such that $d(x_k, x)$ converges to 0 for some $x \in X$, we have $x \in F$. In the following proposition, we prove that on a graded group the topology induced by a homogeneous quasi-distance and the manifold topology coincide.

Proposition 2.26. *The topology induced by a homogeneous quasi-distance on a graded group coincides with the manifold topology of the group. Moreover, a set is relatively compact if and only if it is bounded with respect to the homogeneous quasi-distance.*

Proof. Let G be a graded group with identity e, associated positive grading of its Lie algebra given by $\mathfrak{g} = \bigoplus_{t>0} V_t$ and associated dilations $(\delta_\lambda)_{\lambda>0}$. Let d be a homogeneous quasi-distance satisfying the quasi-triangle inequality with multiplicative constant C.

To prove that the topology $\mathcal{T}_d$ induced by d and the manifold topology $\mathcal{T}_m$ coincide, it is enough to show that $d(e, p)$ goes to 0 if and only if p converges to e. Here, and in the rest of this proof, the latter convergence (and more generally the convergence of some sequence of points) means convergence with respect to the manifold topology on G.

First, we show that the quasi-distance $d(e, \cdot)$ from e is continuous at e with respect to $\mathcal{T}_m$. Since graded groups are nilpotent and simply connected, we can consider exponential coordinates of second kind with respect to a suitable choice of basis of $\mathfrak{g}$, see [7]. Namely, first consider a basis $(X_1, \dots, X_n)$ adapted to the grading of $\mathfrak{g}$, i.e., for all $i = 1, \dots, n$, there exists $d_i > 0$ such that $X_i \in V_{d_i}$, and consequently, $\delta_\lambda(X_i) = \lambda^{d_i} X_i$. Second, consider the diffeomorphism $p \mapsto (P_1(p), \dots, P_n(p))$ from G onto $\mathbb{R}^n$ such that for all $p \in G$,

$$p = \exp(P_1(p)X_1) \cdots \exp(P_n(p)X_n).$$

In particular, $P_i(e) = 0$ for all $i = 1, \dots, n$. Then, using the quasi-triangle inequality, the left-invariance and the homogeneity of the quasi-distance, we get

$$\begin{aligned} d(e, p) &= d(e, \exp(P_1(p)X_1) \cdots \exp(P_n(p)X_n)) \\ &\leq \sum_{i=1}^{n} C^i d(e, \exp(P_i(p)X_i)) \\ &\leq \sum_{i=1}^{n} C^i d(e, \exp(\delta_{|P_i(p)|^{1/d_i}}(\operatorname{sgn}(P_i(p))X_i)) \\ &= \sum_{i=1}^{n} C^i |P_i(p)|^{1/d_i} d(e, \operatorname{sgn}(P_i(p)) \exp(X_i)), \end{aligned}$$

where $\mathrm{sgn}(P_i(p))$ denotes the sign of $P_i(p)$. The last upper bound goes to 0 when p converges to e and this proves the claim.

Second, we show that if (p_k) is a sequence for which $d(e, p_k)$ goes to 0, then p_k converges to e. We set

$$\|p\| := \sum_{i=1}^{n} |P_i(p)|. \tag{2.27}$$

Arguing by contradiction, up to a subsequence, there would exist $\varepsilon > 0$ such that $\|p_k\| > \varepsilon$ for all k. Since the map $\lambda \mapsto \|\delta_\lambda(q)\|$ is continuous with respect to $\mathcal{T}_m$ for all $q \in G$, we get that, for all k, one can find $\lambda_k \in (0, 1)$ such that $\|\delta_{\lambda_k}(p_k)\| = \varepsilon$. By compactness with respect to $\mathcal{T}_m$ of $\{p \in G : \|p\| = \varepsilon\}$, up to a subsequence, we would get that $\delta_{\lambda_k}(p_k)$ converges to some $q \in G$ with $\|q\| = \varepsilon$. In particular, $q \neq e$ and hence $d(e, q) > 0$. However,

$$\begin{aligned} 0 < d(e,q) &\leq C(d(e, \delta_{\lambda_k}(p_k)) + d(\delta_{\lambda_k}(p_k), q)) \\ &= C(\lambda_k d(e, p_k) + d(e, q^{-1} \cdot \delta_{\lambda_k}(p_k))) \\ &\leq C(d(e, p_k) + d(e, q^{-1} \cdot \delta_{\lambda_k}(p_k))). \end{aligned}$$

By the continuity of $d(e, \cdot)$ at e with respect to $\mathcal{T}_m$ and since $q^{-1} \cdot \delta_{\lambda_k}(p_k)$ converges to e, we have that $d(e, q^{-1} \cdot \delta_{\lambda_k}(p_k))$ goes to 0. Hence the last upper bound in the above inequalities goes to 0. This gives a contradiction and proves the claim. All together we get that both topologies coincide.

Relative compactness with respect to the manifold topology is equivalent to boundedness with respect to $\|\cdot\|$. Hence, to show that relative compactness is equivalent to boundedness with respect to the quasi-distance, we show that boundedness with respect to d and boundedness with respect to $\|\cdot\|$ are equivalent. By contradiction, assume that one can find a sequence (p_k) such that, for some $M > 0$, $d(e, p_k) \leq M$ for all k, but $\|p_k\|$ goes to $+\infty$. Arguing as above, one can find a positive sequence (λ_k) converging to 0 such that $\|\delta_{\lambda_k}(p_k)\| = 1$ for all k. On the other hand, since $d(e, \delta_{\lambda_k}(p_k)) = \lambda_k d(e, p_k) \leq M\lambda_k$, $d(e, \delta_{\lambda_k}(p_k))$ goes to 0. As shown before, this implies that $\delta_{\lambda_k}(p_k)$ converges to e and gives a contradiction. One shows in a similar way that boundedness with respect to $\|\cdot\|$ implies boundedness with respect to the quasi-distance. □

In the rest of this paper, when not specified, all topological properties on graded groups will be understood as defined with respect to the manifold topology, or, equivalently in view of Proposition 2.26, with respect to the topology induced by any homogeneous quasi-distance.

If d is a distance on a set X, then d is continuous on $X \times X$ with respect to the topology it induces and subsets of the form $\{y \in X : d(y, x) < r\}$, respectively $\{y \in X : d(y, x) \leq r\}$, are open, respectively closed, with respect to the topology induced by d. In particular, in view of Proposition 2.26, we have the following consequence.

Corollary 2.28. *Every homogeneous distance on a homogeneous group G is continuous on $G \times G$ with respect to the manifold topology.*

On the contrary, we stress that, if d is a quasi-distance on a set X, subsets of the form $\{y \in X : d(y, x) < r\}$, respectively $\{y \in X : d(y, x) \leq r\}$, may not be open, respectively closed, for the topology induced by d. Moreover, even when one of these properties holds, the quasi-distance may not be continuous on $X \times X$ with respect to the topology it induces, see Remark 2.30 below. In particular, in view of Proposition 2.26, a homogeneous quasi-distance on a graded group G may not be continuous on $G \times G$ with respect to the manifold topology. The first part of the proof of Proposition 2.26 only says that the quasi-distance $d(e, \cdot)$ from e is continuous at e (or equivalently by left-invariance, the quasi-distance $d(q, \cdot)$ from q is continuous at q for all $q \in G$). In other terms, it only says that a ball $B_d(p, r)$ contains its center p in its interior. We show in the next proposition that a homogeneous quasi-distance on a graded group G is continuous on $G \times G$ if and only if its spheres are closed, or equivalently its unit sphere centered at e is closed .

Proposition 2.29 (Continuity of homogeneous quasi-distances). *A homogeneous quasi-distance d on a graded group G is continuous on $G \times G$ if and only if its unit sphere centered at e is closed.*

Proof. First, note that a homogeneous quasi-distance $d : G \times G \to [0, +\infty)$ is continuous on $G \times G$ if and only if the quasi-distance $d(e, \cdot)$ from e is continuous on G. If $d(e, \cdot)$ is continuous on G, its unit sphere $S_d(e, 1) := \{p \in G : d(e, p) = 1\}$ is obviously closed. Conversely assume that $S_d(e, 1)$ is closed. We already know from the proof of Proposition 2.26 that $d(e, \cdot)$ is continuous at e. To prove the continuity of $d(e, \cdot)$ at an arbitrary point p with $p \neq e$, it is sufficient to show that for any sequence (p_k) such that $d(p_k, p)$ goes to 0, one can extract a subsequence whose quasi-distance to e converges to $d(e, p)$. Set $\lambda_k := d(e, p_k)$. We have $\lambda_k \leq C(d(e, p)+d(p, p_k))$ hence the sequence (λ_k) is bounded. Up to a subsequence, one can thus assume that λ_k converges to some $\lambda \geq 0$. If $\lambda = 0$, then $\lambda_k = d(e, p_k)$ goes to 0, hence p_k converges to e (see Proposition 2.26) and we would have $p = e$. Since we are considering $p \neq e$, we thus have $\lambda > 0$. Then $\|\delta_{1/\lambda_k}(p_k) - \delta_{1/\lambda}(p)\|$ goes to 0 (remember (2.27) for the definition of $\| \cdot \|$), i.e., $\delta_{1/\lambda_k}(p_k)$ converges to $\delta_{1/\lambda}(p)$. As $\delta_{1/\lambda_k}(p_k) \in S_d(e, 1)$ and $S_d(e, 1)$ is closed, we get $\delta_{1/\lambda}(p) \in S_d(e, 1)$. Hence $\lambda = d(e, p)$ which concludes the proof. □

Remark 2.30. In view of Proposition 2.29, it is easy to construct examples of homogeneous quasi-distances that are not continuous. In such a case, the unit ball centered at the identity may or may not be closed. For instance, in $\mathbb{R}^2$ equipped with its trivial Abelian stratification, consider the homogeneous quasi-distance d_1 whose unit ball centered at the origin is the union of the Euclidean closed unit disk centered at the origin with the interval $[-2, 2] \times \{0\}$ (see Example 2.31 below for a characterization of homogeneous quasi-distances on graded groups in terms of their unit ball). Its unit sphere at the origin is the union of two points $\{(-2, 0), (2, 0)\}$ with the set $\{(x, y) \in \mathbb{R}^2 : x^2+y^2 = 1\}\setminus\{(1, 0), (-1, 0)\}$, which is not closed. In this example, the unit ball (and hence any ball) is (are) closed and sets of the form $\{p \in \mathbb{R}^2 : d_1(0, p) < r\}$ are not open. One can also consider the homogeneous quasi-distance d_2 whose unit ball centered at the origin is the Euclidean closed unit disk centered at the origin minus the segments $[-1, -1/2) \cup (1/2, 1]$, which is not closed. Its unit sphere at the origin is the union of two points $\{(-1/2, 0), (1/2, 0)\}$ with the set $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \setminus \{(1, 0), (-1, 0)\}$, which is not closed as well. However, sets of the form $\{p \in \mathbb{R}^2 : d_2(0, p) < r\}$ are open. In both examples, the quasi-distance from the origin is not continuous at points $p \in \mathbb{R}^* \times \{0\}$.

We conclude this section with a characterization of homogeneous quasi-distances by means of their unit ball. Together with Proposition 2.29, we get a way to construct continuous quasi-distances on graded groups. Finally, we also recall a characterization of homogeneous distances by means of their unit ball.

Example 2.31 (Characterization of homogeneous quasi-distances by means of their unit ball). Let G be a graded group with identity e and with associated dilations $(\delta_\lambda)_{\lambda>0}$. Let d be a homogeneous quasi-distance on G. By Proposition 2.26, e belongs to the interior of $B_d(e,1)$ and $B_d(e,1)$ is relatively compact. By left-invariance, $B_d(e,1)$ is symmetric, i.e., $p \in B_d(e,1)$ implies $p^{-1} \in B_d(e,1)$. Finally, it follows from the homogeneity of d that, for all $p \in G$, the set $\{\lambda > 0 : \delta_{1/\lambda}(p) \in B_d(e,1)\}$ is a closed subinterval of $(0,+\infty)$ (for the relative topology on $(0,+\infty)$).

Conversely, assume that K is a subset of G that contains e in its interior, K is relatively compact, symmetric, and such that the set $\{\lambda > 0 : \delta_{1/\lambda}(p) \in K\}$ is a closed subinterval of $(0,+\infty)$ for all $p \in G$. Then

$$d(p,q) := \inf\{\lambda > 0 : \delta_{1/\lambda}(p^{-1} \cdot q) \in K\}$$

defines a homogeneous quasi-distance on G. It is the homogeneous quasi-distance whose unit ball centered at e is the set K.

For the sake of completeness, we give below a detailed proof of this claim. Although the general scheme of the proof is a classical one, we stress that some of the arguments use the topological properties proved in Proposition 2.26. For $p \in G$, we set

$$I(p) := \{\lambda > 0 : \delta_{1/\lambda}(p) \in K\}$$

and

$$\rho(p) := \inf\{\lambda > 0 : \delta_{1/\lambda}(p) \in K\}.$$

Note that since K contains e in its interior and $\delta_{1/\lambda}(p)$ converges to e when λ goes to $+\infty$, we have $I(p) \neq \emptyset$ and $\rho(p) < +\infty$ for all $p \in G$. Obviously, one has $\rho(e) = 0$. Conversely let $p \in G$ with $p \neq e$. Then $\|\delta_{1/\lambda}(p)\|$ goes to $+\infty$ when λ goes to 0 (remember (2.27) for the definition of $\|\cdot\|$). Since K is relatively compact, and hence bounded with respect to $\|\cdot\|$, it follows that $\delta_{1/\lambda}(p) \notin K$ for all $\lambda > 0$ small enough. Hence $\rho(p) \neq 0$ and consequently, we get that

$$\rho(p) = 0 \quad \text{if and only if} \quad p = e. \tag{2.32}$$

Next, since K is symmetric, we have $I(p) = I(p^{-1})$. Hence

$$\rho(p) = \rho(p^{-1}). \tag{2.33}$$

Third, since $\delta_\lambda \circ \delta_\eta = \delta_{\lambda\eta}$ for all $\lambda, \eta > 0$, one has $I(\delta_\lambda(p)) = \lambda I(p)$ for all $p \in G$ and all $\lambda > 0$. Hence

$$\rho(\delta_\lambda(p)) = \lambda \rho(p). \tag{2.34}$$

Finally, the fact ρ satisfies the quasi-triangle inequality, i.e., there exists some constant $C > 0$ such that

$$\rho(p \cdot q) \leq C\,(\rho(p) + \rho(q)) \quad \text{for all } p, q \in G, \tag{2.35}$$

follows from the fact that ρ is bi-Lipschitz equivalent to any homogeneous quasi-norm on G. Indeed, let d_0 be a homogeneous quasi-distance on G (remember that homogeneous quasi-distances on graded groups do exist, see Example 2.25) and set $\rho_0(p) := d_0(e, p)$. Since the topology induced by d_0 and the manifold topology coincide (see Proposition 2.26) and since e belongs to the interior of K, there exists $\eta > 0$ such that $p \in K$ as soon as $\rho_0(p) < \eta$. By homogeneity, it follows that $2\rho_0(p)/\eta \in I(p)$ for all $p \neq e$. Hence $\rho(p) \leq 2\rho_0(p)/\eta$ for all $p \in G$. On the other hand, since K is relatively compact, it is bounded with respect to ρ_0 by Proposition 2.26. Hence one can find $M > 0$ such that $\rho_0(p) \leq M$ for all $p \in K$. It follows that for $p \neq e$, $\delta_{2M/\rho_0(p)}(p) \notin K$ and hence $\rho_0(p)/2M \notin I(p)$. By assumption, $I(p)$ is closed subinterval of $(0, +\infty)$. Moreover, since K contains e in its interior and $\delta_{1/\lambda}(p)$ converges to e when λ goes to $+\infty$, $I(p)$ is an unbounded closed subinterval of $(0, +\infty)$, i.e.,

$$I(p) = [\rho(p), +\infty)$$

for all $p \in G$ with $p \neq e$. It follows that $\rho_0(p)/2M \leq \rho(p)$ for all $p \in G$. All together we get that one can find a constant $L > 0$ such that

$$L^{-1}\rho_0(p) \leq \rho(p) \leq L\, \rho_0(p)$$

for all $p \in G$. Then the fact that ρ_0 satisfies the quasi-triangle inequality implies that ρ satisfies the quasi-triangle inequality as well. This proves (2.35).

All together (2.32)–(2.35) imply that d is a homogeneous quasi-distance on G. Finally, one has $\rho(p) \leq 1$ if and only if $1 \in I(p)$, i.e., $p \in K$, hence $K = B_d(e, 1)$.

Example 2.36 (Construction of continuous homogeneous quasi-distances on graded groups). The characterization given in Example 2.31 together with Proposition 2.29 gives an effective way to construct continuous homogeneous quasi-distances on arbitrary graded groups. In particular, one can extend Hebisch and Sikora's construction of Example 2.23. Namely, following the notations of Example 2.23, we get that, for all $R > 0$, d_R induces a homogeneous quasi-distance on an arbitrary graded group. It is the homogeneous quasi-distance whose unit ball centered at the identity is a Euclidean ball of radius R, using exponential coordinates of the first kind relative to some basis of the Lie algebra adapted to the positive grading. Moreover, its unit sphere centered at the identity is a Euclidean sphere of radius R hence is closed. It follows that d_R is continuous.

Example 2.37 (Characterization of homogeneous distances on homogeneous groups). We recall here a characterization, already contained in a slightly different form in [15], of homogeneous distances on homogeneous groups in terms of their unit ball. Let G be a homogeneous group with associated dilations $(\delta_\lambda)_{\lambda>0}$ and with identity element e. If d is a homogeneous distance on G, then e belongs to the interior of $B_d(e, 1)$, $B_d(e, 1)$ is compact and symmetric. Since d satisfies the quasi-triangle inequality with a multiplicative constant $C = 1$, we have $\delta_\lambda(p) \cdot \delta_{1-\lambda}(q) \in B_d(e, 1)$ for all $p, q \in B_d(e, 1)$ and all $\lambda \in [0, 1]$.

Conversely, assume that K is a subset of G that contains e in its interior, K is compact, symmetric and such that $\delta_\lambda(p) \cdot \delta_{1-\lambda}(q) \in K$ for all $p, q \in K$ and all $\lambda \in [0, 1]$. Then

$$d(p, q) := \inf\{\lambda > 0 : \delta_{1/\lambda}(p^{-1} \cdot q) \in K\}$$

defines a homogeneous distance on G. It is the homogeneous distance whose unit ball centered at e is the set K. We refer to [15] for the proof of the fact that d satisfies the quasi-triangle inequality with a multiplicative constant $C = 1$ and to Example 2.31 for all other properties that must be satisfied by a homogeneous distance.

3. Besicovitch Covering Property

3.1. BCP and WBCP. Recall from the introduction the definition of the Besicovitch Covering Property in the general quasi-metric setting. See Section 2.5 for the definition and our conventions about quasi-metric spaces.

Definition 3.1 (Besicovitch Covering Property). Let (X, d) be a quasi-metric space. We say that (X, d) satisfies the *Besicovitch Covering Property* (BCP) if there exists a constant $N \geq 1$ such that the following holds. Let A be a bounded subset of X and let $\mathcal{B}$ be a family of balls such that each point of A is the center of some ball of $\mathcal{B}$; then there is a finite or countable subfamily $\mathcal{F} \subset \mathcal{B}$ such that the balls in $\mathcal{F}$ cover A, and every point in X belongs to at most N balls in $\mathcal{F}$, that is,

$$\mathbb{1}_A \leq \sum_{B \in \mathcal{F}} \mathbb{1}_B \leq N,$$

where $\mathbb{1}_A$ denotes the characteristic function of the set A.

The Besicovitch Covering Property originates from works of Besicovitch in connection with the theory of differentiation of measures in Euclidean spaces ([4, 5], see also Section 6). Finite-dimensional normed vector spaces satisfy BCP (see [11, Chapter 2.8]) whereas infinite-dimensional normed vector spaces do not satisfy BCP.

Definition 3.1 for BCP is a common and classical one, even though one can find various variants in the literature. In the Euclidean setting, these variants are equivalent. One of them, called in the present paper the Weak Besicovitch Covering Property (WBCP), see Definition 3.3, turns out to be equivalent to BCP in our setting of graded groups equipped with homogeneous quasi-distances, and more generally for doubling quasi-metric spaces, see Proposition 3.7. For our purposes, working with WBCP is actually technically more convenient.

For the sake of completeness, we discuss in more details in the rest of this section the relationships between BCP and WBCP, first pointing out that BCP and WBCP may happen to be not equivalent for general metric spaces, see Example 3.4. This might be of independent interest, and, to our knowledge, cannot be found explicitly written in the literature. We will next prove that for doubling quasi-metric spaces, and hence for graded groups equipped with homogeneous quasi-distances, BCP and WBCP are equivalent. We first introduce some convenient terminology.

Definition 3.2 (Family of Besicovitch balls). Let (X, d) be a quasi-metric space. We say that a family $\mathcal{B} := \{B = B_d(x_B, r_B)\}$ of balls in (X, d) is a *family of Besicovitch balls* if $\mathcal{B}$ is a finite family of balls such that, for all B, $B' \in \mathcal{B}$ with $B \neq B'$, one has $x_B \notin B'$, and for which $\bigcap_{B \in \mathcal{B}} B \neq \emptyset$.

Definition 3.3 (Weak BCP). Let (X, d) be a quasi-metric space. We say that (X, d) satisfies the *Weak Besicovitch Covering Property* (WBCP) if there exists a constant $Q \geq 1$ such that $\operatorname{Card} \mathcal{B} \leq Q$ for every family $\mathcal{B}$ of Besicovitch balls in (X, d).

If (X, d) satisfies BCP, then (X, d) satisfies WBCP. One can indeed take $Q = N$, where N is given by Definition 3.1. Conversely, as already mentioned, WBCP is in general strictly weaker than BCP, as the following example shows.

Example 3.4. Here is an example of a metric space that does not satisfy BCP and for which Card $\mathcal{B} = 1$ for every family $\mathcal{B}$ of Besicovitch balls. Let $X = \{x_1, x_2, \dots\}$ be a countable set of points. Let us define $d : X \times X \to [0, +\infty)$ as follows. We set $d(x_i, x_i) = 0$ for all $i \geq 1$ and

$$d(x_i, x_j) = 1 - \frac{1}{\max(i, j)} \quad \text{for } i \neq j.$$

We first check that d defines a distance on X. The fact that $d(x, y) = 0$ if and only if $x = y$ and $d(x, y) = d(y, x)$ are obvious from the definition. To prove the triangle inequality, let $j < i$ and $k \geq 1$ be fixed. If $k < i$, we have

$$d(x_i, x_j) = 1 - \frac{1}{i} \leq 1 - \frac{1}{i} + 1 - \frac{1}{\max(j, k)} = d(x_i, x_k) + d(x_k, x_j).$$

If $i < k$, then $k \geq 3$, hence $i/(i+1) \leq 1 \leq k/2$ and so $1 - 1/i \leq 2(1 - 1/k)$. It follows that

$$d(x_i, x_j) = 1 - \frac{1}{i} \leq 2\left(1 - \frac{1}{k}\right) = d(x_i, x_k) + d(x_k, x_j).$$

Hence d satisfies the triangle inequality.

We claim that BCP does not hold in (X, d). Indeed, set

$$r_i := 1 - \frac{1}{i} \quad \text{for } i = 1, 2, \dots.$$

Consider $A := \{x_i : i \geq 2\}$ and the family $\mathcal{B} := \{B_d(x_i, r_i) : i \geq 2\}$. Since $d(x_j, x_i) = r_i$ for all $j < i$ and $d(x_j, x_i) = r_j > r_i$ for $j > i$, we have

$$B_d(x_i, r_i) = \{x_1, \dots, x_i\} \quad \text{for } i = 2, 3, \dots.$$

It follows that for any subfamily $\mathcal{F} \subset \mathcal{B}$ whose balls cover the set A, we have

$$\sup\{i \geq 2 : B_d(x_i, r_i) \in \mathcal{F}\} = +\infty,$$

that is, Card $\mathcal{F} = +\infty$. On the other hand, $x_1 \in \bigcap_{B \in \mathcal{F}} B$. In particular, x_1 belongs to infinitely many balls in $\mathcal{F}$ which shows that (X, d) does not satisfy BCP.

Let us now check that Card $\mathcal{B}$ is equal to 1 for every family $\mathcal{B}$ of Besicovitch balls and hence (X, d) satisfies WBCP. By contradiction, assume that $\{B_d(x_{i_l}, \rho_{i_l})\}_{l=1}^k$ is a family of Besicovitch balls with $k \geq 2$. Assume with no loss of generality that $i_1 < i_2 < \dots < i_k$. We have $d(x_{i_j}, x_{i_k}) = r_{i_k}$ and $x_{i_j} \notin B_d(x_{i_k}, \rho_{i_k})$ for all $j = 1, \dots, k-1$, so $r_{i_k} > \rho_{i_k}$. It follows that $d(x_{i_k}, x_l) = \max(r_{i_k}, r_l) \geq r_{i_k} > \rho_{i_k}$ for all $l \neq i_k$ and thus

$$B_d(x_{i_k}, \rho_{i_k}) = \{x_{i_k}\} \subset X \setminus \bigcup_{j=1}^{k-1} B(x_{i_j}, \rho_{i_j}),$$

which contradicts the fact that $\bigcap_{l=1}^k B(x_{i_l}, \rho_{i_l}) \neq \emptyset$.

Remark 3.5. If (X, d) satisfies WBCP, then it satisfies a weak form of BCP that can be stated as follows. There is a constant $N \geq 1$ such that the following holds. Let A be a bounded subset of X. Let $\mathcal{B}$ be a family of balls such that each point of A is the center of some ball of $\mathcal{B}$ and such that either $\sup\{r_B : B \in \mathcal{B}\} = +\infty$ or $B \in \mathcal{B} \mapsto r_B$ attains only an isolated set of values in $(0, +\infty)$. Then there is a finite or countable subfamily $\mathcal{F} \subset \mathcal{B}$ such that the balls in $\mathcal{F}$ cover A, and every point in X belongs to at most N balls in $\mathcal{F}$ (see [31]). Note that in Example 3.4, the number 1 is an accumulation point of the set $\{r_i : i \geq 2\}$ in $(0, +\infty)$.

As already mentioned, for doubling quasi-metric spaces, BCP and WBCP are equivalent. Let us recall the definition of doubling quasi-metric spaces.

Definition 3.6 (Doubling quasi-metric space). A quasi-metric space (X, d) is said to be *doubling* if there is a constant $C \geq 1$ such that for each $r > 0$, each ball in (X, d) with radius $2r$ can be covered by a family of at most C balls of radius r.

Proposition 3.7. *Let (X, d) be a doubling quasi-metric space. Then (X, d) satisfies BCP if and only if (X, d) satisfies WBCP.*

As a classical fact, graded groups equipped with homogeneous quasi-distances are doubling. The next corollary hence follows.

Corollary 3.8. *Let G be a graded group and let d be a homogeneous quasi-distance on G. Then (G, d) satisfies BCP if and only if (G, d) satisfies WBCP.*

For the sake of completeness, we give below a proof of Proposition 3.7. This proof follows closely the arguments of the proof of [25, Theorem 2.7] about the validity of BCP in Euclidean spaces, using the following well-known property of doubling quasi-metric spaces (see e.g. [24] for more details about doubling (quasi-)metric spaces).

Remark 3.9. If (X, d) be a doubling quasi-metric space, then there are constants $c \geq 1$ and $s \geq 0$ such that if $x \in X$, $r > 0$ and $\lambda \geq 1$, the cardinality of every set in $B_d(x, \lambda r)$ whose points are at least r apart is at most $c\lambda^s$.

Proof of Proposition 3.7. Since any quasi-metric space satisfying BCP also satisfies WBCP, we only need to prove that (X, d) satisfies BCP when (X, d) is a doubling quasi-metric space satisfying WBCP.

Let A be a bounded subset of X and let $\mathscr{B}$ be a family of balls such that each point of A is the center of some ball of $\mathscr{B}$. For each $x \in A$ choose one ball $B_d(x, r(x))$ in $\mathscr{B}$. As A is bounded, we claim that we may assume that

$$M_1 := \sup_{x \in A} r(x) < +\infty.$$

Indeed, otherwise pick some point x in A with $r(x) \geq \operatorname{diam} A$. Then $\mathscr{F} := \{B_d(x, r(x))\}$ is obviously a subfamily of $\mathscr{B}$ which shows that BCP holds in (X, d).

Choose $x_1 \in A$ with $r(x_1) \geq M_1/2$ and then inductively

$$x_{j+1} \in A \setminus \bigcup_{i=1}^{j} B_d(x_i, r(x_i)) \quad \text{with } r(x_{j+1}) \geq \frac{M_1}{2}$$

as long as possible. Since A is bounded and points x_i are at least $M_1/2$ apart, it follows from Remark 3.9 that the process terminates and we get a finite sequence $x_1, \dots, x_{k_1}$.

Let

$$M_2 := \sup\left\{ r(x) : x \in A \setminus \bigcup_{i=1}^{k_1} B_d(x_i, r(x_i)) \right\},$$

and choose

$$x_{k_1+1} \in A \setminus \bigcup_{i=1}^{k_1} B_d(x_i, r(x_i)) \quad \text{with } r(x_{k_1+1}) \geq \frac{M_2}{2}$$

and again inductively

$$x_{j+1} \in A \setminus \bigcup_{i=1}^{j} B_d(x_i, r(x_i)) \quad \text{with } r(x_{j+1}) \geq \frac{M_2}{2}$$

as long as possible.

Continuing this process, we get a finite or infinite increasing sequence of integers

$$0 = k_0 < k_1 < k_2 < \cdots,$$

a decreasing sequence of positive numbers M_i with $2M_{i+1} \leq M_i$, and a sequence of balls $B_i := B_d(x_i, r(x_i)) \in \mathcal{B}$ with the following properties. If $I_j := \{k_{j-1}+1, \dots, k_j\}$ for $j = 1, 2, \dots$, then

$$\frac{M_j}{2} \leq r(x_i) \leq M_j \qquad \text{for } i \in I_j, \tag{3.10}$$

$$x_{j+1} \in A \setminus \bigcup_{i=1}^{j} B_i \qquad \text{for } j = 1, 2, \dots,, \tag{3.11}$$

$$x_i \in A \setminus \bigcup_{m \neq k} \bigcup_{j \in I_m} B_j \quad \text{for } i \in I_k. \tag{3.12}$$

The first two properties (3.10) and (3.11) follow from the construction. To prove (3.12), let $m \neq k$, $i \in I_k$ and $j \in I_m$. If $m < k$, then $x_i \notin B_j$ by (3.11). If $k < m$, then by construction $r(x_j) < r(x_i)$, and $x_j \notin B_i$ by (3.11), and so $x_i \notin B_j$.

Let us now check that this subfamily of balls satisfies the conditions for BCP to hold. If the sequence $k_0, k_1, \dots$ is finite, it follows immediately from the construction that the balls B_i cover A. If the sequence is infinite, then M_j converges to 0, (3.10) implies that $r(x_i)$ converges to 0, and it follows as well from the construction that

$$A \subset \bigcup_{i=1}^{+\infty} B_i.$$

To verify the other property for the validity of BCP, assume that a point $x \in X$ belongs to p balls $B_{m_1}, \dots, B_{m_p}$. Since WBCP holds in (X, d), we have by (3.12) that the indices m_i can belong to at most Q different blocks I_j, where Q is given by Definition 3.3, that is,

$$\operatorname{Card}\{j : I_j \cap \{m_1, \dots, m_p\} \neq \emptyset\} \leq Q.$$

To conclude, let us check that

$$\operatorname{Card}(I_j \cap \{m_1, \dots, m_p\}) \leq M \quad \text{for } j = 1, 2, \dots$$

for some constant M depending only on the doubling constant of (X, d). Let j be fixed. The points x_l, $l \in I_j \cap \{m_1, \dots, m_p\}$, are at least $M_j/2$ apart by (3.10) and (3.11) and are all contained in $B_d(x, M_j)$, and so the claim follows from Remark 3.9. □

Remark 3.13. Note that the subfamily constructed in the previous proof satisfies the following additional property: $B_d(x_i, r(x_i)/4) \cap B_d(x_j, r(x_j)/4) = \emptyset$ for all $i \neq j$. Indeed, let $i < j$. Then $r(x_j) \leq 2r(x_i)$ and $x_j \notin B_i$ by (3.11), hence

$$d(x_j, x_i) > r(x_i) > \frac{r(x_i)}{4} + \frac{r(x_j)}{4}.$$

3.2. Preserving BCP. Let us first recall that the validity of (W)BCP is not stable under a biLipschitz change of (quasi-)distance, see Theorem 1.8. More generally, the validity of (W)BCP might not be stable under natural operations on quasi-metric spaces. See for instance the example before Theorem 3.16 about product of quasi-metric spaces. The fact that (W)BCP holds on a quasi-metric space (X, d) depends indeed on the precise shape of balls in (X, d). We give in this section cases where the validity of (W)BCP is preserved, to be used later. This might be more generally of independent interest.

We begin with the following simple remark. If d_1 and d_2 are two quasi-distances on a space X such that any ball with respect to d_2 is a ball with respect to d_1, with the same center but possibly with a different radius, then the validity of (W)BCP in (X, d_1) implies the validity of (W)BCP in (X, d_2).

For instance if (W)BCP holds in (X, d), then, for any $s > 0$, d^s defines a quasi-distance on X and (W)BCP holds on (X, d^s). Note that it is well known that a metric space (X, d) and its snowflakes (X, d^s), $0 < s < 1$, have for many other purposes significantly different properties. For graded groups, we get the following proposition, to be used later.

Proposition 3.14. *Let G be a graded group and let d be a homogeneous quasi-distance on G. Let $t > 0$. If BCP holds on (G, d), then BCP holds on the t-power of G equipped with the homogeneous quasi-distance $d^{1/t}$.*

See Example 2.14 for the definition of the t-power of a graded group and Example 2.24 for homogeneous quasi-distances on t-powers.

Another simple remark is the fact that a subset of a quasi-metric space that satisfies (W)BCP also satisfies (W)BCP when equipped with the restricted quasi-distance. We state it below for later reference.

Proposition 3.15. *Let (X, d_X) be a quasi-metric space. Let $Y \subset X$. If BCP (respectively WBCP) holds on (X, d_X), then BCP (respectively WBCP) holds on (Y, d_Y), where d_Y denotes the quasi-distance d_X restricted to Y.*

Given two quasi-metric spaces (X, d_X) and (Y, d_Y), there are many ways to define quasi-distances on $X \times Y$. If (X, d_X) and (Y, d_Y) both satisfy WBCP, then WBCP may fail for classical choices of quasi-distances on $X \times Y$, as the following example shows. We know that, for $s \geq 1$, the set $\mathbb{R}$ equipped with the snowflake distance $d_s(x, x') := |x' - x|^{1/s}$ satisfies WBCP. Let $s > 1$, $r \geq 1$, and let $d_{s,r}$ be the distance on $\mathbb{R} \times \mathbb{R}$ given by

$$d_{s,r}((x, y), (x', y')) := (d_1(x, x')^r + d_s(y, y')^r)^{1/r}.$$

Then, if $r \in [1, s)$, WBCP does not hold on $(\mathbb{R} \times \mathbb{R}, d_{s,r})$. Indeed, $d_{s,r}$ is a left-invariant distance on $\mathbb{R} \times \mathbb{R}$ (equipped with the Abelian group law) and is one-homogeneous with respect to the dilations $\delta_\lambda(x, y) := (\lambda x, \lambda^s y)$. Its unit ball centered at the origin is given by $B_{d_{s,r}}(0, 1) = \{(x, y) \in \mathbb{R} \times \mathbb{R} : |x|^r + |y|^{r/s} \leq 1\}$. It follows from [21, Lemma 3.2] that if WBCP holds in $(\mathbb{R} \times \mathbb{R}, d_{s,r})$, one would have $r \geq s$.

However, the following theorem, to be used later, shows that one can always find a quasi-distance satisfying WBCP on a product of quasi-metric spaces that satisfy WBCP.

Theorem 3.16. *Let (X, d_X) and (Y, d_Y) be two metric spaces. Assume that WBCP holds in (X, d_X) and in (Y, d_Y). Then $X \times Y$ equipped with the max distance*

$$d_{X\times Y}((x, y), (x', y')) := \max(d_X(x, x'), d_Y(y, y'))$$

satisfies WBCP.

Proof. Let $Q \in \mathbb{N}$ be such that Card $\mathcal{F} \leq Q$ for any family $\mathcal{F}$ of Besicovitch balls in (X, d_X) or in (Y, d_Y). Let $\mathcal{B} := \{B_{d_{X\times Y}}(p_i, r_i)\}_{i=1}^N$ be a family of Besicovitch balls in $(X \times Y, d_{X\times Y})$. Let $i, j \in \{1, \dots, N\}$, $i \neq j$, and let $p_i := (x_i, y_i)$ and $p_j := (x_j, y_j)$. By the definition of families of Besicovitch balls and by the definition of $d_{X\times Y}$, we have

$$d_{X\times Y}(p_i, p_j) = \max(d_X(x_i, x_j), d_Y(y_i, y_j)) > \max(r_i, r_j)$$

hence $d_X(x_i, x_j) > \max(r_i, r_j)$ or $d_Y(y_i, y_j) > \max(r_i, r_j)$. In other terms, for any pair of indices (i, j) with $i \neq j$, we have

$$x_i \notin B_{d_X}(x_j, r_j) \quad \text{and} \quad x_j \notin B_{d_X}(x_i, r_i) \tag{3.17}$$

or

$$y_i \notin B_{d_Y}(y_j, r_j) \quad \text{and} \quad y_j \notin B_{d_Y}(y_i, r_i). \tag{3.18}$$

Let us consider the graph Γ with N vertices $i = 1, \dots, N$ and where i is connected to j if and only if $i \neq j$ and (3.17) holds. Then, for any complete subgraph γ of Γ (a complete graph is a graph where any two vertices are connected), $\{B_{d_X}(x_i, r_i)\}_{i\in\gamma}$ is a family of Besicovitch balls in (X, d_X). Let Γ' be the complementary graph of Γ, that is, the graph with the same vertices as Γ and where two vertices are connected in Γ' if and only if they are not connected in Γ. Since (3.18) holds whenever (3.17) does not, $\{B_{d_Y}(y_i, r_i)\}_{i\in\gamma'}$ is a family of Besicovitch balls in (Y, d_Y) for any complete subgraph γ' of Γ'.

As a special case of Ramsey's theorem stated in the language of graph theory, there exists a function $f(k, l)$ such that for any given graph Γ with $N \geq f(k, l)$ vertices, then either Γ contains a complete subgraph of order k or its complementary graph Γ' contains a complete subgraph of order l (the order of a complete graph is the number of its vertices). An upper bounded for $f(k, k)$ for $k \geq 3$ has been proved by P. Erdős and G. Szekeres ([10]), namely $f(k, k) < 4^{k-1}$.

Going back to the family of Besicovitch balls $\mathcal{B}$, it follows that if the numbers N of balls in $\mathcal{B}$ is larger than 4^Q (we may assume with no loss of generality that $Q \geq 2$), there would exist either a family of Besicovitch balls in (X, d_X) with cardinality $Q + 1$ or a family of Besicovitch balls in (Y, d_Y) with cardinality $Q + 1$. This contradicts the fact that by assumption any family of Besicovitch balls in (X, d_X) or in (Y, d_Y) has cardinality at most Q. Hence $(X \times Y, d_{X\times Y})$ satisfies WBCP. □

Submetries, also known as metric submersions, will play a important role in our arguments. They are indeed well adapted tools for our purposes. We first recall the definition.

Definition 3.19 (Submetry). Let (X, d_X) and (Y, d_Y) be quasi-metric spaces. We say that $\pi : X \to Y$ is a *submetry* if π is a surjective map such that $\pi(B_{d_X}(p,r)) = B_{d_Y}(\pi(p), r)$ for all $p \in X$ and all $r > 0$.

We recall the following property of submetries related to WBCP.

Proposition 3.20 ([21, Proposition 2.7]). *Let (X, d_X) and (Y, d_Y) be quasi-metric spaces. Assume that there exists a submetry from (X, d_X) onto (Y, d_Y). If (X, d_X) satisfies WBCP, then (Y, d_Y) satisfies WBCP.*

Proposition 3.20 will be used in the proof of our main results together with Proposition 3.21 below.

Proposition 3.21. *Let $\hat{G}$ and G be graded groups with graded Lie algebra $\hat{\mathfrak{g}}$ and $\mathfrak{g}$, respectively. Assume that there exists a surjective morphism of graded Lie algebras $\phi : \hat{\mathfrak{g}} \to \mathfrak{g}$. Let $\varphi : \hat{G} \to G$ denote the unique Lie group homomorphism such that $\varphi_* = \phi$ and let $\hat{d}$ be a homogeneous distance, respectively a continuous homogeneous quasi-distance, on $\hat{G}$. Then*

$$d(p,q) := \hat{d}(\varphi^{-1}(\{p\}), \varphi^{-1}(\{q\}))$$

defines a homogeneous distance, respectively a continuous homogeneous quasi-distance, on G and $\varphi : (\hat{G}, \hat{d}) \to (G, d)$ is a submetry.

We stress that continuity of the quasi-distance $\hat{d}$, which means global continuity on $\hat{G} \times \hat{G}$, is necessary in order to get that d is a quasi-distance on G, which turns out to be continuous on $G \times G$ as well, and also in order to get that $\varphi : (\hat{G}, \hat{d}) \to (G, d)$ is a submetry. Indeed, consider the homogeneous quasi-distance d_2 on $\mathbb{R}^2$ given in Remark 2.30, which is not globally continuous. The projection of $B_{d_2}(0,1)$ onto the x-axis is the open segment $I := (-1, 1)$. If $x \in \mathbb{R}^*$, we have $\{\lambda > 0 : x/\lambda \in I\} = (|x|, +\infty)$ which is not a closed subinterval of $(0, +\infty)$. It follows from Example 2.31 that I is not the unit ball of some homogeneous quasi-distance on $\mathbb{R}$. However, when $\hat{d}$ is a homogeneous distance, recall that $\hat{d}$ is continuous on $\hat{G} \times \hat{G}$ (see Corollary 2.28).

Proof of Proposition 3.21. First, we prove that d defines a quasi-distance on G and that $\varphi : (\hat{G}, \hat{d}) \to (G, d)$ is a submetry. By [21, Proposition 2.8], it is sufficient to prove that for all $p, q \in G$ and all $\hat{p} \in \varphi^{-1}(\{p\})$, one can find $\hat{q} \in \varphi^{-1}(\{q\})$ such that $d(p,q) = \hat{d}(\hat{p}, \hat{q})$. Set $\hat{K} := \operatorname{Ker}\varphi$. We have $\varphi^{-1}(\{p\}) = \hat{K} \cdot \hat{p}$ and $\hat{k} \cdot \varphi^{-1}(\{q\}) = \varphi^{-1}(\{q\})$ for all $\hat{k} \in \hat{K}$. By left-invariance of $\hat{d}$, it follows that

$$\hat{d}(\hat{p}, \varphi^{-1}(\{q\})) = \hat{d}(\hat{k} \cdot \hat{p}, \hat{k} \cdot \varphi^{-1}(\{q\}) = \hat{d}(\hat{k} \cdot \hat{p}, \varphi^{-1}(\{q\})).$$

In other words, the function $\hat{p}' \in \varphi^{-1}(\{p\}) \mapsto \hat{d}(\hat{p}', \varphi^{-1}(\{q\}))$ is constant. Hence, by the definition of d, we get

$$d(p,q) = \hat{d}(\hat{p}, \varphi^{-1}(\{q\})) = \inf_{\hat{q} \in \varphi^{-1}(\{q\})} \hat{d}(\hat{p}, \hat{q}).$$

Bounded sets with respect to $\hat{d}$ are relatively compact (see Proposition 2.26), the set $\varphi^{-1}(\{q\})$ is closed and $\hat{d}$ is assumed to be continuous, hence one can find $\hat{q} \in \varphi^{-1}(\{q\})$ such that

$$\hat{d}(\hat{p}, \hat{q}) = \inf_{\hat{q}' \in \varphi^{-1}(\{q\})} \hat{d}(\hat{p}, \hat{q}') = d(p, q).$$

This proves that d defines a quasi-distance on G and that $\varphi : (\hat{G}, \hat{d}) \to (G, d)$ is a submetry.

Note that if $\hat{d}$ satisfies the quasi-triangle inequality with multiplicative constant C, then d satisfies the quasi-triangle inequality with the same multiplicative constant (see the proof of [21, Proposition 2.8]). In particular, if $\hat{d}$ is a distance on $\hat{G}$, then d is a distance on G.

Next, one can easily check that d is homogeneous. This follows from the fact that $\hat{d}$ is a homogeneous quasi-distance together with the fact that ϕ is a surjective morphism of graded Lie algebra.

To conclude the proof, it remains to prove that d is globally continuous on $G \times G$. Since d is left-invariant, it is sufficient to prove that $d(e, \cdot)$ is globally continuous on G. Here e denotes the identity in G and below $\hat{e}$ will denote the identity in $\hat{G}$. Let $p \in G$ and let (p_k) be a sequence converging to p. First, let $\hat{p}_k \in \varphi^{-1}(\{p_k\})$ be such that $d(e, p_k) = \hat{d}(\hat{e}, \hat{p}_k)$. Since the sequence (p_k) is relatively compact, it is bounded with respect to d (see Proposition 2.26). Hence $(\hat{p}_k)$ is bounded with respect to $\hat{d}$ and, once again by Proposition 2.26, relatively compact. Up to a subsequence, one can thus assume that $\hat{p}_k$ converges to some $\hat{p} \in \varphi^{-1}(\{p\})$. Since $\hat{d}$ is continuous, it follows that $\hat{d}(\hat{e}, \hat{p}_k)$ converges to $\hat{d}(\hat{e}, \hat{p})$ and one gets

$$d(e, p) \leq \hat{d}(\hat{e}, \hat{p}) = \lim_{k \to +\infty} \hat{d}(\hat{e}, \hat{p}_k) = \lim_{k \to +\infty} d(e, p_k).$$

Next, let $\hat{p}' \in \varphi^{-1}(\{p\})$ be such that $d(e, p) = \hat{d}(\hat{e}, \hat{p}')$. Since $\varphi = \exp \circ \phi \circ \exp^{-1}$ is an open map (see [37, p. 104]), one can find a sequence $\hat{p}'_k \in \varphi^{-1}(\{p_k\})$ converging to $\hat{p}'$. Since $\hat{d}$ is assumed to be continuous, $\hat{d}(\hat{e}, \hat{p}'_k)$ goes to $\hat{d}(\hat{e}, \hat{p}')$. On the other hand, we have

$$d(e, p_k) \leq \hat{d}(\hat{e}, \hat{p}'_k).$$

Hence

$$\lim_{k \to +\infty} d(e, p_k) \leq \hat{d}(\hat{e}, \hat{p}') = d(e, p).$$

All together we finally get that $\lim_{k \to +\infty} d(e, p_k) = d(e, p)$ which concludes the proof. □

4. Graded groups with commuting different layers

In this section we consider graded groups with commuting different layers, see Definition 1.1, and we prove the following results.

Theorem 4.1. *Let G be a graded group with commuting different layers. There exist continuous homogeneous quasi-distances on G for which BCP holds.*

Corollary 4.2. *Let G be a homogeneous group with commuting different layers. There exist homogeneous distances on G for which BCP holds.*

The proof of Theorem 4.1 and Corollary 4.2 is divided into three steps. First, we consider stratified free-nilpotent Lie groups of step 2. We prove that for such groups, some homogeneous (quasi-)distances that satisfy BCP are those whose unit ball centered at the origin coincides with a Euclidean ball centered at the origin in exponential coordinates of the first kind associated to a choice of basis of the Lie algebra, see Theorem 4.5. Next, we prove the existence of homogeneous distances satisfying BCP for stratified groups of step 2, see Theorem 4.24. These homogeneous distances are induced by homogeneous distances satisfying BCP on stratified free-nilpotent Lie groups of step 2 via submetries. Finally, the general case follows from Theorem 4.24 together with the structure property given by Proposition 2.15, and Theorem 3.16, see Section 4.3.

4.1. Free-nilpotent groups of step 2. Let $r \geq 2$ be an integer. We denote by $\mathbb{F}_{r2}$ the stratified free-nilpotent Lie group of step 2 and rank r whose Lie algebra $\mathfrak{f}_{r2}$ is endowed with a given stratification $\mathfrak{f}_{r2} = V \oplus W$, where $[V, V] = W$ and where

$$\dim V = r, \quad \dim W = \frac{r(r-1)}{2}.$$

We set $n := \dim \mathfrak{f}_{r2}$. We fix a basis $(X_1, \dots, X_r)$ of V and we set $X_{ij} := [X_i, X_j]$. Then $(X_{ij})_{1 \leq i < j \leq r}$ is a basis of W. Using exponential coordinates of the first kind associated to the basis $(X_1, \dots, X_r, (X_{ij})_{1\leq i<j\leq r})$ adapted to the given stratification of $\mathfrak{f}_{r2}$, we write $p \in \mathbb{F}_{r2}$ as

$$p = \exp\left(\sum_{i=1}^{r} p_i X_i + \sum_{1\leq i<j\leq r} p_{ij} X_{ij}\right)$$

and we identify p with $(p_1, \dots, p_r, (p_{ij})_{1\leq i<j\leq r}) = [v_p, w_p]$, where $v_p := (p_1, \dots, p_r)$ and $w_p := (p_{ij})_{1\leq i<j\leq r}$.

The group law is given by $v_{p\cdot q} = v_p + w_q$ and $w_{p\cdot q} = ((p \cdot q)_{ij})_{1\leq i<j\leq r}$, where

$$(p \cdot q)_{ij} = p_{ij} + q_{ij} + \frac{1}{2}(p_i q_j - q_i p_j) \tag{4.3}$$

for $1 \leq i < j \leq r$. The identity element is the origin.

The associated dilations are given by

$$\delta_\lambda(p) = (\lambda v_p, \lambda^2 w_p).$$

In this section we denote by $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ the Euclidean norm and scalar product in $\mathbb{R}^n$, $\mathbb{R}^r$ and $\mathbb{R}^{n-r}$ with respect to our choice of coordinates. Equivalently, we equip $\mathfrak{f}_{r2}$ with a Euclidean structure for which our chosen basis $((X_1, \dots, X_r), (X_{ij})_{1\leq i<j\leq r})$ is orthonormal and we consider on V and W the induced Euclidean structures.

For $R > 0$, we consider the homogeneous quasi-distance d on $\mathbb{F}_{r2}$ whose unit ball centered at the origin is given by

$$B_d(0, 1) := \{p \in \mathbb{F}_{r2} : \|p\|^2 \leq R^2\}. \tag{4.4}$$

Such a quasi-distance is well defined and is continuous, see Example 2.36 (we drop here the index R for simplicity of notations). Recall also that it follows from [15] (see also Example 2.23) that d is a distance whenever $R < R^*$ for some $R^* > 0$.

This subsection is devoted to the proof of the validity of BCP on $\mathbb{F}_{r2}$ equipped with such a quasi-distance.

Theorem 4.5. *Let $R > 0$ be fixed. Let d be the homogeneous quasi-distance on $\mathbb{F}_{r2}$ whose unit ball centered at the origin is given by* (4.4). *Then BCP holds on* $(\mathbb{F}_{r2}, d)$.

From now on in this subsection, we let $R > 0$ be fixed and d denotes the homogeneous quasi-distance on $\mathbb{F}_{r2}$ whose unit ball centered at the origin is given by (4.4). We set $B := B_d(0,1)$. We begin with a series of remarks for later use.

First, for $p \in \mathbb{F}_{r2}$, we define the function $A_p : \mathbb{F}_{r2} \to \mathbb{R}$ by

$$A_p(q) := \|q\|^2 - 2\langle p, q\rangle + \sum_{1 \le i < j \le r} (p_{ij} - q_{ij})(p_i q_j - q_i p_j) + \frac{1}{4}(p_i q_j - q_i p_j)^2. \tag{4.6}$$

Lemma 4.7. *Let $p \in \partial B$. Then $q \in B_d(p,1)$ if and only if $A_p(q) \le 0$.*

Proof. Let $p \in \partial B$. We have $\|p\|^2 = \|v_p\|^2 + \|w_p\|^2 = R^2$ and $q \in B_d(p,1)$ if and only if $\|p^{-1} \cdot q\|^2 = \|v_{p^{-1}\cdot q}\|^2 + \|w_{p^{-1}\cdot q}\|^2 \le R^2$. Then the lemma follows from the specific form of the group law given in (4.3). □

Next, we denote by $\angle(\cdot,\cdot)$ the (non-oriented) angle $\in [0,\pi]$ between two vectors in a Euclidean space. In the proof of Theorem 4.1, we are going to look at points p, q for which the angles $\angle(v_p, v_q)$ and $\angle(w_p, w_q)$ are small. These two angles are dilation invariant. Namely, we have

$$\angle(v_{\delta_\lambda(p)}, v_{\delta_\lambda(q)}) = \angle(v_p, v_q) \quad \text{and} \quad \angle(w_{\delta_\lambda(p)}, w_{\delta_\lambda(q)}) = \angle(w_p, w_q) \tag{4.8}$$

for all $p, q \in \mathbb{F}_{r2}$ and all $\lambda > 0$. Note that, on the contrary, the angle $\angle(p,q)$ is not dilation invariant.

Lemma 4.9. *For all $\varepsilon > 0$, there exists a constant $\delta > 0$ such that, for all $p, q \in \mathbb{F}_{r2}$, if $\angle(v_p, v_q) < \delta$ and $\angle(w_p, w_q) < \delta$, then*

$$|p_i q_j - q_i p_j| \le \varepsilon \|v_p\| \|v_q\| \quad \textit{for all } 1 \le i < j \le r, \tag{4.10}$$

$$\langle v_p, v_q\rangle \ge (1-\varepsilon)\|v_p\|\|v_q\|, \tag{4.11}$$

$$\langle w_p, w_q\rangle \ge (1-\varepsilon)\|w_p\|\|w_q\|. \tag{4.12}$$

Proof. To prove (4.10), note that $|p_i q_j - q_i p_j|$ represents the area of the planar quadrilateral generated by the two-dimensional vectors (p_i, p_j) and (q_i, q_j). For fixed length of these vectors, this area goes to zero when the angle between the two vectors goes to zero. This angle is smaller than $\angle(v_p, v_q)$, and the length of the two vectors is bounded by $\|v_p\|$ and $\|v_q\|$, respectively. This implies (4.10). To prove (4.11) and (4.12), note that if a and b are two vectors in a Euclidean space such that $\angle(a,b) < \pi/2$, then $\langle a, b\rangle$ is positive and represents the product between $\|a\|$ and the norm of the orthogonal projection of b onto a. In particular, if $\angle(a,b)$ goes to 0, this product converges to $\|a\|\|b\|$. □

Lemma 4.13. *Let $\varepsilon > 0$ and $p, q \in B$ be such that* (4.10)*,* (4.11) *and* (4.12) *hold. Let $A_p(q)$ be defined by* (4.6)*. Then we have*

$$\begin{aligned} A_p(q) &\le \|v_q\|^2 + \|w_q\|^2 - 2(1-\varepsilon)\|v_p\|\|v_q\| - 2(1-\varepsilon)\|w_p\|\|w_q\| \\ &\quad + 2r^2 R\varepsilon \|v_p\|\|v_q\| + \frac{r^2}{4}\varepsilon^2 R^2 \|v_p\|\|v_q\| \end{aligned} \tag{4.14}$$

$$\begin{aligned} &\le \|v_q\|^2 + \|w_q\|^2 - 2(1-\varepsilon)\|v_p\|\|v_q\| - 2(1-\varepsilon)\|w_p\|\|w_q\| \\ &\quad + 2r^2 R^2 \varepsilon \|v_q\| + \frac{r^2}{4}\varepsilon^2 R^3 \|v_q\|. \end{aligned} \tag{4.15}$$

Proof. By the definition of $A_p(q)$, we have

$$A_p(q) = \|v_q\|^2 + \|w_q\|^2 - 2\langle v_p, v_q\rangle - 2\langle w_p, w_q\rangle + \sum_{1\le i<j\le r} (p_{ij} - q_{ij})(p_i q_j - q_i p_j) + \frac{1}{4}(p_i q_j - q_i p_j)^2.$$

Then (4.14) and (4.15) follow from (4.10), (4.11) and (4.12) together with the following simple observations. First, $\dim W \le r^2$, which is used to bound the number of terms in the sum. Second, since $p, q \in B$, we have $\|v_p\|, \|w_p\|, \|v_q\|, \|w_q\| \le R$, which is used to bound some of the terms. □

Next, we will consider in the proof of Theorem 4.5 the following parabolic regions. For $a > 0$, we set

$$\mathcal{P}_a := \{p \in \mathbb{F}_{r2} : R\|w_p\| > a\|v_p\|^2\}.$$

These regions, as well as their complement, are invariant under dilations. Namely,

$$\delta_\lambda(\mathcal{P}_a) = \mathcal{P}_a \quad \text{and} \quad \delta_\lambda(\mathcal{P}_a^c) = \mathcal{P}_a^c \tag{4.16}$$

for all $a > 0$ and all $\lambda > 0$.

Lemma 4.17. *We have the following bounds.*

(i) *Let $p \in B$. If $p \notin \mathcal{P}_a$. Then*

$$\|w_p\| \le \frac{R}{2a}\left(\sqrt{1+4a^2} - 1\right). \tag{4.18}$$

(ii) *Let $p \in \partial B$. If $p \in \mathcal{P}_a$, then*

$$\|w_p\| \ge \frac{R}{2a}\left(\sqrt{1+4a^2} - 1\right). \tag{4.19}$$

(iii) *Let $p \in \partial B$. If $p \notin \mathcal{P}_a$, then*

$$\|v_p\| \ge \frac{R}{a}\sqrt{\frac{\sqrt{1+4a^2}-1}{2}}. \tag{4.20}$$

Proof. (i) From the assumptions, we have

$$\|w_p\|^2 + \frac{R}{a}\|w_p\| \le \|w_p\|^2 + \|v_p\|^2 \le R^2.$$

Hence $a\|w_p\|^2 + R\|w_p\| - aR^2 \le 0$ and $\|w_p\| \ge 0$, which is equivalent to

$$0 \le \|w_p\| \le \frac{R}{2a}\left(-1 + \sqrt{1+4a^2}\right)$$

and gives (4.18).

(ii) From the assumptions, we have

$$\|w_p\|^2 + \frac{R}{a}\|w_p\| > \|w_p\|^2 + \|v_p\|^2 = R^2.$$

Hence $a\|w_p\|^2 + R\|w_p\| - aR^2 > 0$ and $\|w_p\| \geq 0$, which is equivalent to

$$\|w_p\| > \frac{R}{2a}\big(-1 + \sqrt{1+4a^2}\big)$$

and implies (4.19).

(iii) From the assumptions, we have

$$\frac{a^2}{R^2}\|v_p\|^4 + \|v_p\|^2 \geq \|w_p\|^2 + \|v_p\|^2 = R^2.$$

Hence $a^2\|v_p\|^4 + R^2\|v_p\|^2 - R^4 \geq 0$, which is equivalent to

$$\|v_p\|^2 \geq \frac{R^2}{2a^2}\big(-1 + \sqrt{1+4a^2}\big)$$

and gives (4.20). □

To prove Theorem 4.5, we are going to partition $\mathbb{F}_{r2}$ into three disjointed regions,

$$\mathbb{F}_{r2} = \mathcal{P}_{a'} \sqcup (\mathcal{P}_a \setminus \mathcal{P}_{a'}) \sqcup \mathcal{P}_a^c,$$

for some suitable choice of $0 < a < a'$.

The next three lemmas show that, for a suitable choice of $0 < a < a'$, if p and q are two points for which the angles $\angle(v_p, v_q)$ and $\angle(w_p, w_q)$ are small and both belong to one of these regions, then either $q \in B_d(p, d(0,p))$ or $p \in B_d(q, d(0,q))$. We first consider the case where $p, q \in \mathcal{P}_a^c$.

Lemma 4.21. *Let $a = 0.9$. There exists a constant $\delta > 0$ such that, if $p, q \in \mathcal{P}_a^c$ are such that $\angle(v_p, v_q) < \delta$ and $\angle(w_p, w_q) < \delta$, then $q \in B_d(p, d(0,p))$ or $p \in B_d(q, d(0,q))$.*

Proof. The value a has been chosen in such a way that

$$1 + \frac{\sqrt{1+4a^2}-1}{2} - 2\sqrt{\frac{\sqrt{1+4a^2}-1}{2a^2}} < 0.$$

Hence, we can fix some $\varepsilon > 0$ so that

$$1 + \frac{\sqrt{1+4a^2}-1}{2} - 2(1-\varepsilon)\sqrt{\frac{\sqrt{1+4a^2}-1}{2a^2}} + 2r^2R\varepsilon + \frac{r^2}{4}\varepsilon^2R^2 < 0.$$

Let $p, q \in \mathcal{P}_a^c$. Assume that $\angle(v_p, v_q) < \delta$ and $\angle(w_p, w_q) < \delta$ where $\delta > 0$ is given by Lemma 4.9, i.e., is such that (4.10), (4.11) and (4.12) hold for our choice of ε. Using dilations together with (4.8) and (4.16), and exchanging the role of p and q if necessary, one can assume with no loss of generality that $p \in \partial B$ and $q \in B$.

Then let us prove that $q \in B_d(p, 1)$. By Lemma 4.7, this is equivalent to $A_p(q) \leq 0$. By (4.15), we have

$$A_p(q) \leq \|v_q\|^2 + \|w_q\|^2 - 2(1-\varepsilon)\|v_p\|\|v_q\| - 2(1-\varepsilon)\|w_p\|\|w_q\|$$
$$+ 2r^2R^2\varepsilon\|v_q\| + \frac{r^2}{4}\varepsilon^2R^3\|v_q\|.$$

To bound the first term, we use that $\|v_q\| \leq R$ since $q \in B$. To bound the second one, we use that $q \notin \mathcal{P}_a$ both through (4.18) and the fact that $\|w_q\| < \frac{a}{R}\|v_q\|^2 \leq a\|v_q\|$. In the third term, we use that $p \notin \mathcal{P}_a$ through (4.20). Since the fourth term is not positive, we get

$$\begin{aligned} A_p(q) &\le R\|v_q\| + \frac{R}{2a}(\sqrt{1+4a^2}-1)a\|v_q\| - 2(1-\varepsilon)\frac{R}{a}\sqrt{\frac{\sqrt{1+4a^2}-1}{2}}\|v_q\| \\ &\quad + 2r^2R^2\varepsilon\|v_q\| + \frac{r^2}{4}\varepsilon^2R^3\|v_q\| \\ &= R\|v_q\|\left(1 + \frac{\sqrt{1+4a^2}-1}{2} - 2(1-\varepsilon)\sqrt{\frac{\sqrt{1+4a^2}-1}{2a^2}} + 2r^2R\varepsilon + \frac{r^2}{4}\varepsilon^2R^2\right) \\ &\le 0 \end{aligned}$$

by the choice of ε. □

Next, we consider the case where $p, q \in \mathcal{P}_{a'}$.

Lemma 4.22. *Let $a' = 1.9$. There exists a constant $\delta > 0$ such that if $p, q \in \mathcal{P}_{a'}$ are such that $\angle(v_p, v_q) < \delta$ and $\angle(w_p, w_q) < \delta$, then $q \in B_d(p, d(0,p))$ or $p \in B_d(q, d(0,q))$.*

Proof. The value a' has been chosen in such a way that

$$\frac{1}{a'} + 1 - \frac{\sqrt{1+4a'^2}-1}{a'} < 0.$$

Hence, we can fix some $\varepsilon > 0$ so that

$$\frac{1}{a'} + 1 - (1-\varepsilon)\frac{\sqrt{1+4a'^2}-1}{a'} < 0 \quad \text{and} \quad -2(1-\varepsilon) + 2r^2R\varepsilon + \frac{r^2}{4}\varepsilon^2R^2 < 0.$$

Let $p, q \in \mathcal{P}_{a'}$. Assume that $\angle(v_p, v_q) < \delta$ and $\angle(w_p, w_q) < \delta$ where $\delta > 0$ is given by Lemma 4.9, i.e., is such that (4.10), (4.11) and (4.12) hold for our choice of ε. Arguing as in the proof of Lemma 4.21, one can assume with no loss of generality that $p \in \partial B$ and $q \in B$.

Let us prove that $q \in B_d(p, 1)$. By Lemma 4.7, this is equivalent to $A_p(q) \le 0$. By (4.14), we have

$$\begin{aligned} A_p(q) &\le \|v_q\|^2 + \|w_q\|^2 - 2(1-\varepsilon)\|v_p\|\|v_q\| - 2(1-\varepsilon)\|w_p\|\|w_q\| \\ &\quad + 2r^2R\varepsilon\|v_p\|\|v_q\| + \frac{r^2}{4}\varepsilon^2R^2\|v_p\|\|v_q\|. \end{aligned}$$

To bound the first term, we use that $q \in \mathcal{P}_{a'}$, i.e., $\|v_q\|^2 < \frac{R}{a'}\|w_q\|$. To bound the second one, we use that $\|w_q\| \le R$ since $q \in B$. In the fourth term, we use that $p \in \mathcal{P}_{a'}$ through (4.19). This gives

$$\begin{aligned} A_p(q) &\le \frac{R}{a'}\|w_q\| + R\|w_q\| - 2(1-\varepsilon)\|v_p\|\|v_q\| - 2(1-\varepsilon)\frac{R}{2a'}\big(\sqrt{1+4a'^2}-1\big)\|w_q\| \\ &\quad + 2r^2R\varepsilon\|v_p\|\|v_q\| + \frac{r^2}{4}\varepsilon^2R^2\|v_p\|\|v_q\| \\ &= R\|w_q\|\left(\frac{1}{a'} + 1 - (1-\varepsilon)\frac{\sqrt{1+4a'^2}-1}{a'}\right) \\ &\quad + \|v_p\|\|v_q\|\left(-2(1-\varepsilon) + 2r^2R\varepsilon + \frac{r^2}{4}\varepsilon^2R^2\right) \\ &\le 0 \end{aligned}$$

by the choice of ε. □

Finally, we consider the case where $p, q \in \mathcal{P}_a \setminus \mathcal{P}_{a'}$.

Lemma 4.23. *Let $a = 0.9$ and $a' = 1.9$. There exists a constant $\delta > 0$ such that if $p, q \in \mathcal{P}_a \setminus \mathcal{P}_{a'}$ are such that $\angle(v_p, v_q) < \delta$ and $\angle(w_p, w_q) < \delta$, then $q \in B_d(p, d(0, p))$ or $p \in B_d(q, d(0, q))$.*

Proof. The values of a and a' have been chosen in such a way that

$$\frac{1}{2} + \frac{\sqrt{1+4a'^2}-1}{4} - \frac{2}{a'}\sqrt{\frac{\sqrt{1+4a'^2}-1}{2}} < 0 \quad \text{and} \quad \frac{1}{2a} + \frac{1}{2} - \frac{\sqrt{4a^2+1}-1}{a} < 0.$$

Hence, we can fix some $\varepsilon > 0$ so that

$$\frac{1}{2} + \frac{\sqrt{1+4a'^2}-1}{4} - 2(1-\varepsilon)\frac{1}{a'}\sqrt{\frac{\sqrt{1+4a'^2}-1}{2}} + 2r^2R\varepsilon + \frac{r^2}{4}\varepsilon^2R^2 < 0$$

and

$$\frac{1}{2a} + \frac{1}{2} - (1-\varepsilon)\frac{\sqrt{4a^2+1}-1}{a} < 0.$$

Let $p, q \in \mathcal{P}_a \setminus \mathcal{P}_{a'}$. Assume that $\angle(v_p, v_q) < \delta$ and $\angle(w_p, w_q) < \delta$ where $\delta > 0$ is given by Lemma 4.9, i.e., is such that (4.10), (4.11) and (4.12) hold for our choice of ε. Arguing as in the proof of Lemma 4.21, one can assume with no loss of generality that $p \in \partial B$ and $q \in B$.

Let us prove that $q \in B_d(p, 1)$. By Lemma 4.7, this is equivalent to $A_p(q) \le 0$. By (4.15), we have

$$A_p(q) \le \frac{1}{2}\|v_q\|^2 + \frac{1}{2}\|v_q\|^2 + \frac{1}{2}\|w_q\|^2 + \frac{1}{2}\|w_q\|^2 - 2(1-\varepsilon)\|v_p\|\|v_q\| - 2(1-\varepsilon)\|w_p\|\|w_q\| + 2r^2R^2\varepsilon\|v_q\| + \frac{r^2}{4}\varepsilon^2R^3\|v_q\|.$$

To bound the first term, we use that $q \in \mathcal{P}_a$, i.e., $\|v_q\|^2 < \frac{R}{a}\|w_q\|$. To bound the second term, we use that $q \in B$, hence $\|v_q\| \le R$. To bound the third term, we use that $q \notin \mathcal{P}_{a'}$ through both (4.18) and the fact that $\|w_q\| < \frac{a'}{R}\|v_q\|^2 \le a'\|v_q\|$. In the fourth one, we use that $q \in B$, hence $\|w_q\| \le R$. In the fifth term, we use that $p \notin \mathcal{P}_{a'}$ through (4.20) and in the sixth one we use that $p \in \mathcal{P}_a$ through (4.19). This gives

$$\begin{aligned}
A_p(q) &\le \frac{R}{2a}\|w_q\| + \frac{R}{2}\|v_q\| + \frac{R}{4a'}(\sqrt{1+4a'^2}-1)a'\|v_q\| + \frac{R}{2}\|w_q\| \\
&\quad - 2(1-\varepsilon)\frac{R}{a'}\sqrt{\frac{\sqrt{1+4a'^2}-1}{2}}\|v_q\| - 2(1-\varepsilon)\frac{R}{2a}(\sqrt{4a^2+1}-1)\|w_q\| \\
&\quad + 2r^2R^2\varepsilon\|v_q\| + \frac{r^2}{4}\varepsilon^2R^3\|v_q\| \\
&= R\|v_q\|\left(\frac{1}{2} + \frac{\sqrt{1+4a'^2}-1}{4} - 2(1-\varepsilon)\frac{1}{a'}\sqrt{\frac{\sqrt{1+4a'^2}-1}{2}} + 2r^2R\varepsilon + \frac{r^2}{4}\varepsilon^2R^2\right) \\
&\quad + R\|w_q\|\left(\frac{1}{2a} + \frac{1}{2} - (1-\varepsilon)\frac{\sqrt{4a^2+1}-1}{a}\right) \\
&\le 0
\end{aligned}$$

by the choice of ε. □

We are now going to conclude the proof of Theorem 4.5.

Proof of Theorem 4.5. Let $a = 0.9$, $a' = 1.9$ and let $\delta > 0$ be small enough so that Lemmas 4.21, 4.22 and 4.23 hold. Let N be an upper bound of the maximum number of vectors in a $\max(r, n-r)$-dimensional Euclidean space that pairwise make an angle larger than $\delta/2$. Such a bound exists and is finite by compactness of finite-dimensional Euclidean spheres.

We are going to prove that a family of Besicovitch balls in $(\mathbb{F}_{r2}, d)$ cannot have a cardinality larger than $3N^2$. This will imply that WBCP, and hence BCP by Corollary 3.8, holds in $(\mathbb{F}_{r2}, d)$. Let $\{B_d(p_i, r_i)\}_{i\in I}$ be a family of Besicovitch balls. Using left-translations and shrinking balls if necessary, one can assume with no loss of generality that $0 \in \bigcap_{i\in I} B_d(p_i, r_i)$ and $r_i = d(0, p_i)$ for all $i \in I$. If $\operatorname{Card} I > 3N^2$, by the pigeonhole principle one can find $I_1 \subseteq I$ with $\operatorname{Card} I_1 \geq \operatorname{Card} I/N > 3N$ such that $\angle(v_{p_i}, v_{p_j}) < \delta$ for all $i, j \in I_1$. Then, once again by the pigeonhole principle, one can find $I_2 \subseteq I_1$ with $\operatorname{Card} I_2 \geq \operatorname{Card} I_1/N > 3$ such that $\angle(w_{p_i}, w_{p_j}) < \delta$ for all $i, j \in I_2$. Finally, there exists at least two distinct points p_i and p_j with $i, j \in I_2$ that both belong either to $\mathcal{P}_{a'}$, or $\mathcal{P}_a \setminus \mathcal{P}_{a'}$ or $\mathcal{P}_a^c$. Then Lemmas 4.21, 4.22 and 4.23, lead to a contradiction since by the definition of a family of Besicovitch balls, we have $p_i \notin B_d(p_j, r_j) = B_d(p_j, d(0, p_j))$ for all $i \neq j \in I$. □

4.2. Stratified groups of step 2. This subsection is devoted to the proof of Corollary 4.2 in the case of stratified groups of step 2 as restated below.

Theorem 4.24. *Let G be a stratified group of step 2. There exist homogeneous distances on G for which BCP holds.*

Proof. Let G be a stratified group of step 2 and rank r whose Lie algebra is endowed with a given stratification $\mathfrak{g} = V_1 \oplus V_2$. Let $(Y_1, \ldots, Y_r)$ be a basis of V_1. Let $\mathbb{F}_{r2}$ be the free-nilpotent Lie group of step 2 and rank r. With the conventions and notations of Section 4.1, let $\phi : \mathfrak{f}_{r2} \to \mathfrak{g}$ denote the unique morphism of graded Lie algebras such that $\phi(X_i) = Y_i$ for $i = 1, \ldots, r$, which is surjective. Let $\varphi : \mathbb{F}_{r2} \to G$ denote the unique Lie group homomorphism such that $\varphi_* = \phi$. Let d be a homogeneous distance on $\mathbb{F}_{r2}$ for which BCP holds. Such distances exist by Theorem 4.5 and [15] (see also Example 2.23). Recall also that homogeneous distances are continuous. It follows from Proposition 3.21 that

$$d_G(p, q) := d(\varphi^{-1}(\{p\}), \varphi^{-1}(\{q\})) \tag{4.25}$$

defines a homogeneous distance on G and

$$\varphi : (\mathbb{F}_{r2}, d) \to (G, d_G)$$

is a submetry. To conclude the proof, the fact that WBCP holds on (G, d_G) follows from Proposition 3.20. Hence BCP holds on (G, d_G) by Corollary 3.8. □

Remark 4.26. More generally, arguing as in the proof of Theorem 4.24, one gets that, if d is a continuous homogeneous quasi-distance satisfying BCP on $\mathbb{F}_{r2}$, then (4.25) defines a continuous homogeneous quasi-distance for which BCP holds on the stratified group G of step 2 and rank r. In particular, all homogeneous quasi-distances on $\mathbb{F}_{r2}$ whose unit ball centered at the origin are given by (4.4) are continuous and hence induce on G continuous homogeneous quasi-distances for which BCP holds.

Remark 4.27. It can be checked that if d is a homogeneous quasi-distance on $\mathbb{F}_{r2}$ whose unit ball centered at the origin is given by (4.4) for some $R > 0$, then the unit ball centered at the origin for the quasi-distance d_G given by (4.25) on the stratified group G of step 2 can be described as a Euclidean ball centered at the origin in exponential coordinates of the first kind relative to a suitable choice of basis of $\mathfrak{g}$ adapted to its stratification. More precisely, one can find a basis $(Z_1, \dots, Z_r)$ of V_1 and a basis $(Z_{r+1}, \dots, Z_n)$ of V_2 such that, in exponential coordinates of the first kind relative to the basis $(Z_1, \dots, Z_n)$ of $\mathfrak{g}$, we have

$$B_{d_G}(0,1) = \left\{ p \in G : \sum_{i=1}^{n} p_i^2 \leq R^2 \right\}.$$

One can reasonably expect that for any choice of basis of $\mathfrak{g}$ adapted to its stratification, homogeneous quasi-distances on G whose unit ball centered at the origin is a Euclidean ball in exponential coordinates of the first kind relative to the chosen basis satisfy BCP. This would require technical modifications of our arguments and we do not wish here to go further about these technicalities.

4.3. Arbitrary groups with commuting different layers. We conclude in this section the proof of Theorem 4.1 and Corollary 4.2.

Proof of Theorem 4.1 *and Corollary* 4.2. Let G be a graded group with commuting different layers. By Proposition 2.15, G can be written as a direct product of powers of stratified groups of step ≤ 2.

For an Abelian Lie group with trivial associated positive grading (i.e., stratified of step 1), the Euclidean distance (and more generally any distance induced by a norm) is a homogeneous distance that satisfies BCP. For a stratified group of step 2, we know by Theorem 4.24 that there exist homogeneous distances that satisfy BCP. Next, if d is a homogeneous distance for which BCP holds on a graded group, then $d^{1/t}$ is a homogeneous quasi-distance on its t-power (see Example 2.24) that satisfies BCP by Proposition 3.14. In addition, since d is a distance, it is continuous (see Corollary 2.28), and hence $d^{1/t}$ is a continuous quasi-distance.

Hence, on each factor of the decomposition of G, there exist continuous homogeneous quasi-distances for which BCP, and hence WBCP, holds. Then Theorem 4.1 follows from Theorem 3.16.

Note that if G is a homogeneous group, all t-powers in its decomposition as a direct product are t-powers with $t \geq 1$. This implies the existence of homogeneous distances on G for which BCP holds and proves Corollary 4.2. □

5. Graded groups with two different layers not commuting

In this section we consider graded groups for which there are two different layers of the associated positive grading of their Lie algebra that do not commute, and we consider a more general class of quasi-distances defined as follows.

Definition 5.1 (Self-similar quasi-distances on graded groups). Let G be a graded group with associated dilations $(\delta_\lambda)_{\lambda>0}$. We say that a quasi-distance d on G is *self-similar* if it is left-invariant and one-homogeneous with respect to some non-trivial dilation, i.e., if there exists $\lambda > 0$, $\lambda \neq 1$, such that

$$d(\delta_\lambda(p), \delta_\lambda(q)) = \lambda d(p,q)$$

for all $p, q \in G$.

Note that $d(\delta_\lambda(p), \delta_\lambda(q)) = \lambda d(p,q)$ implies $d(\delta_{\lambda^k}(p), \delta_{\lambda^k}(q)) = \lambda^k d(p,q)$ for all $k \in \mathbb{Z}$. In particular, the previous definition is equivalent to the existence of some $0 < \lambda < 1$ such that $d(\delta_{\lambda^k}(p), \delta_{\lambda^k}(q)) = \lambda^k d(p,q)$ for all $p, q \in G$ and all $k \in \mathbb{Z}$.

We prove the following result.

Theorem 5.2. *Let G be a graded group and let $\bigoplus_{t>0} V_t$ be the associated positive grading of its Lie algebra. Assume that $[V_t, V_s] \neq \{0\}$ for some $t \neq s$. Let d be a self-similar quasi-distance on G that is continuous with respect to the manifold topology. Then WBCP, and hence BCP, does not hold in (G, d).*

Remark 5.3. Although we will not use it here, it can be noticed that self-similar quasi-distances on graded groups are doubling, hence BCP and WBCP are equivalent in this context (see Proposition 3.7).

Since homogeneous quasi-distances are in particular self-similar, we get the following corollary.

Corollary 5.4. *Let G be a graded group whose associated positive grading of its Lie algebra is given by $\bigoplus_{t>0} V_t$. Assume that $[V_t, V_s] \neq \{0\}$ for some $t \neq s$. Let d be a continuous homogeneous quasi-distance on G. Then BCP does not hold in (G, d).*

Homogeneous distances on homogeneous groups are continuous with respect to the manifold topology, recall Corollary 2.28. Hence, in such a case, one can drop the continuity assumption and we get the following corollary.

Corollary 5.5. *Let G be a homogeneous group whose associated positive grading of its Lie algebra is given by $\bigoplus_{t\geq 1} V_t$. Assume that $[V_t, V_s] \neq \{0\}$ for some $t \neq s$. Let d be a homogeneous distance on G. Then BCP does not hold in (G, d).*

The proof of Theorem 5.2 is divided into two steps. First, we prove that there does not exist continuous self-similar quasi-distances that satisfy WBCP on the non-standard Heisenberg groups, see Theorem 5.6. Next, we deduce Theorem 5.2 from Theorem 5.6 together with Proposition 2.18 and the use of submetries via a generalization of Proposition 3.21 to continuous self-similar distances, see Proposition 5.26. Let us stress that one of the main differences between homogeneous and self-similar quasi-distances are topological issues that we will explain in Section 5.3.

5.1. Non-standard Heisenberg groups. This subsection is devoted to the proof of Theorem 5.6 below. To simplify notations, we denote here by $\mathfrak{h}$ the first Heisenberg Lie algebra, by (X, Y, Z) a standard basis of $\mathfrak{h}$, and by $\mathbb{H}$ the first Heisenberg group. Recall from Example 2.11 that for $\alpha > 1$, the group $\mathbb{H}$ is called the non-standard Heisenberg group of exponent α when considered as a graded group whose Lie algebra is endowed with the non-standard grading of exponent α, i.e., the grading given by

$$\mathfrak{h} = W_1 \oplus W_\alpha \oplus W_{\alpha+1},$$

where $W_1 := \operatorname{span}\{X\}$, $W_\alpha := \operatorname{span}\{Y\}$, $W_{\alpha+1} := \operatorname{span}\{Z\}$, and where the only non-trivial bracket relation is $[X, Y] = Z$.

Theorem 5.6. *Let $\alpha > 1$. There exists no continuous self-similar quasi-distances for which WBCP holds on the non-standard Heisenberg group of exponent α.*

In this statement, continuity of self-similar quasi-distances means continuity with respect to the manifold topology. We refer to Section 5.3 for the study of topological properties of self-similar quasi-distances.

From now on in this subsection, we fix $\alpha > 1$. Following Example 2.11, we use exponential coordinates of the first kind, we write $p \in \mathbb{H}$ as $p = \exp(xX + yY + zZ)$ and we identify p with (x, y, z). Recall that dilations $(\delta_\lambda)_{\lambda>0}$ relative to the non-standard grading of exponent α are given by

$$\delta_\lambda(x, y, z) = (\lambda x, \lambda^\alpha y, \lambda^{\alpha+1} z). \tag{5.7}$$

To prove Theorem 5.6, we argue by contradiction. We let d be a self-similar quasi-distance on the non-standard Heisenberg group of exponent α. Hence d is left-invariant and, for some fixed $0 < \rho < 1$, we have

$$d(\delta_{\rho^k}(p), \delta_{\rho^k}(q)) = \rho^k\, d(p, q) \tag{5.8}$$

for all $p, q \in \mathbb{H}$ and all $k \in \mathbb{Z}$.

Next, we assume that d is continuous on $\mathbb{H} \times \mathbb{H}$ with respect to the manifold topology. We set $B := B_d(0, 1)$. The continuity of d implies in particular that B is closed and that its boundary ∂B is given by $\partial B = \{p \in \mathbb{H} : d(0, p) = 1\}$.

Finally, arguing by contradiction, we assume that WBCP holds on $(\mathbb{H}, d)$.

To get a contradiction, we first prove a series of lemmas, Lemma 5.9 to Lemma 5.22 below. The final conclusion will follow from these lemmas and is given at the end of this subsection.

First, the assumption about the validity of WBCP has the following consequence.

Lemma 5.9. *For all $p \in B$ and all $\overline{\lambda} > 0$, there exists a constant $0 < \lambda < \overline{\lambda}$ such that $p \cdot \delta_\lambda(p^{-1}) \in B$.*

Proof. We first prove that for all points $p \in \partial B$, there exist arbitrarily large values of j such that $d(p, \delta_{\rho^j}(p)) \leq 1$. By contradiction, assume that one can find $p \in \partial B$ and $k \geq 0$ such that $d(p, \delta_{\rho^j}(p)) > 1$ for all $j \geq k$. For $l \geq 0$, set $r_l := \rho^{lk}$ and $q_l := \delta_{r_l}(p)$. We have $d(0, q_l) = r_l$ by (5.8) hence $0 \in \bigcap_{l \geq 0} B_d(q_l, r_l)$. We also have

$$d(q_j, q_l) = d(\delta_{r_j}(p), \delta_{r_l}(p)) = r_j\, d(p, \delta_{\rho^{(l-j)k}}(p)) > r_j = \max(r_j, r_l)$$

for all $0 \leq j < l$. Hence $q_j \notin B_d(q_l, r_l)$ for all $j \neq l$. It follows that for all finite set $J \subset \mathbb{N}$, $\{B_d(q_j, r_j)\}_{j \in J}$ is a family of Besicovitch balls, which contradicts the validity of WBCP in $(\mathbb{H}, d)$.

To conclude the proof, let $p \in \partial B$. Then $p^{-1} \in \partial B$ by left-invariance of d and it follows that $d(p^{-1}, \delta_{\rho^j}(p^{-1})) \leq 1$ for arbitrarily large values of j. Hence, by left-invariance of d, we get that $p \cdot \delta_\lambda(p^{-1}) \in B$ for arbitrarily small values of λ. If p belongs to the interior of B, the claim follows from the continuity of the map $\lambda \mapsto p \cdot \delta_\lambda(p^{-1})$. □

As a consequence of Lemma 5.9, we prove in Lemma 5.10 a geometric property of the unit ball B. Namely, starting at a point $p \in \partial B$ with $x_p \neq 0$, there is segment that is all contained in B. This segment is a part of the flow line of the vector field $-X$ starting at $p \in \partial B$.

Lemma 5.10. *For all $p = (x_p, y_p, z_p) \in \partial B$ with $x_p \neq 0$, the segment*

$$\hat{\sigma}_p := \left\{ \left((1-t)x_p, y_p, z_p + t\,\frac{x_p y_p}{2} \right) : t \in [0,1] \right\}$$

is contained in B.

Proof. The segment $\hat{\sigma}_p$ is a part of the flow line of the vector field $-X$. For technical convenience, we will thus use in this proof exponential coordinates of the second kind in which X is a constant vector field. Namely, for $p \in \mathbb{H}$, we write

$$p = \exp(x_3(p)Z) \cdot \exp(x_2(p)Y) \cdot \exp(x_1(p)X)$$

and we identify p with $[x_1(p), x_2(p), x_3(p)]$. The relation between exponential coordinates of the second kind $[x_1(p), x_2(p), x_3(p)]$ and exponential coordinates of the first kind (x_p, y_p, z_p) when p is written as $\exp(x_p X + y_p Y + z_p Z)$ (which are used elsewhere in this subsection) is given by $x_1(p) = x_p$, $x_2(p) = y_p$ and $x_3(p) = z_p + x_p y_p/2$. In exponential coordinates of the second kind, the group law is given by

$$\begin{cases} x_1(p \cdot p') = x_1(p) + x_1(p'), \\ x_2(p \cdot p') = x_2(p) + x_2(p'), \\ x_3(p \cdot p') = x_3(p) + x_3(p') + x_1(p)\,x_2(p'), \end{cases}$$

the vector field X is the constant vector field $X = \partial_{x_1}$, and the segment $\hat{\sigma}_p$ writes as

$$\hat{\sigma}_p = \{[(1-t)x_1(p), x_2(p), x_3(p)] : t \in [0,1]\}.$$

Let $\theta \in (0, \pi/2)$ be fixed. For $p = [x_1(p), x_2(p), x_3(p)]$ with $x_1(p) > 0$, we define $C_{p,\theta}$ as the portion of the half-cone with vertex p, axis $\{\exp(-x_1 X) : x_1 \geq 0\}$ and aperture 2θ contained in the half space $\{x_1 \geq 0\}$. If $p = [x_1(p), x_2(p), x_3(p)]$ with $x_1(p) < 0$, then $C_{p,\theta}$ is defined as the portion of the half-cone with vertex p, axis $\{\exp(x_1 X); x_1 \geq 0\}$ and aperture 2θ contained in the half space $\{x_1 \leq 0\}$. Namely, if $x_1(p) > 0$,

$$\begin{aligned} C_{p,\theta} := \{[x_1, x_2, x_3] \in \mathbb{H} : 0 \leq x_1 \leq x_1(p) \text{ and} \\ ((x_2 - x_2(p))^2 + (x_3 - x_3(p))^2)^{1/2} \leq (x_1(p) - x_1)\tan\theta\} \end{aligned}$$

and, if $x_1(p) < 0$,

$$\begin{aligned} C_{p,\theta} := \{[x_1, x_2, x_3] \in \mathbb{H} : x_1(p) \leq x_1 \leq 0 \text{ and} \\ ((x_2 - x_2(p))^2 + (x_3 - x_3(p))^2)^{1/2} \leq (x_1 - x_1(p))\tan\theta\}. \end{aligned}$$

For $p \in \mathbb{H}$, we set
$$\gamma_p(\lambda) := p \cdot \delta_\lambda(p^{-1}).$$
Noting that, in exponential coordinates of the second kind,
$$p^{-1} = [-x_1(p), -x_2(p), -x_3(p) + x_1(p)x_2(p)],$$
we have
$$\begin{cases} x_1(\gamma_p(\lambda)) = (1-\lambda)x_1(p), \\ x_2(\gamma_p(\lambda)) = (1-\lambda^\alpha)x_2(p), \\ x_3(\gamma_p(\lambda)) = (1-\lambda^{\alpha+1})x_3(p) + (\lambda^{\alpha+1} - \lambda^\alpha)x_1(p)x_2(p). \end{cases} \tag{5.11}$$
If $x_1(p) \neq 0$, we get
$$\frac{d}{d\lambda}\gamma_p(\lambda)_{|\lambda=0} = -x_1(p)X$$
hence $\gamma_p(\lambda) \in C_{p,\theta}$ for all $\lambda \geq 0$ small enough. Since $C_{q,\theta} \subset C_{p,\theta}$ for all $q \in C_{p,\theta}$ with $x_1(q) \neq 0$, it follows that, for all $q \in C_{p,\theta}$ with $x_1(q) \neq 0$,
$$\gamma_q(\lambda) \in C_{p,\theta} \quad \text{for all } \lambda \geq 0 \text{ small enough.} \tag{5.12}$$

For $p \in \mathbb{H}$ with $x_1(p) \neq 0$ and $\theta \in (0, \pi/2)$ fixed, one can find $L_{p,\theta} > 0$ such that all curves $(\gamma_q(\lambda))_{\lambda \in [0,1]}$ for $q \in C_{p,\theta}$ are $L_{p,\theta}$-Lipschitz. This can be easily checked from the explicit expression of $\gamma_q(\lambda)$, see (5.11).

For $t \in [0, 1]$, we define $E_{p,\theta}(t)$ as the intersection of $C_{p,\theta}$ with the two-dimensional plane $\{[x_1, x_2, x_3] \in \mathbb{H}; x_1 = (1-t)x_1(p)\}$,
$$E_{p,\theta}(t) := C_{p,\theta} \cap \{[x_1, x_2, x_3] \in \mathbb{H} : x_1 = (1-t)x_1(p)\}.$$

Let $p \in \partial B$ with $x_1(p) \neq 0$, $s \in (0, 1)$ and $\theta \in (0, \pi/2)$ be fixed. We first prove that
$$E_{p,\theta}(s) \cap B \neq \emptyset.$$
We set $L := (1-s)^{-1}L_{p,\theta}$ and
$$I := \{t \in [0, s] : \text{there exists } \gamma : [0, t] \to C_{p,\theta} \ L\text{-Lipschitz}, \ \gamma(0) = p, \ \gamma(t) \in E_{p,\theta}(t) \cap B\},$$
and we actually prove that $s \in I$ from which the claim follows.

First, $0 \in I$ hence I is nonempty. Second, $C_{p,\theta}$ and $E_{p,\theta}(t) \cap B$ being closed, I is closed by the Ascoli–Arzelà Theorem. Hence, we have $\sup I \in I$. By contradiction, assume that $t := \sup I < s$. Since $t \in I$, one can find a L-Lipschitz curve $\gamma : [0, t] \to C_{p,\theta}$ such that $\gamma(0) = p$ and $\gamma(t) \in E_{p,\theta}(t) \cap B$. Set $q := \gamma(t)$. Since $q \in C_{p,\theta}$ with $x_1(q) \neq 0$, it follows from (5.12) that $\gamma_q(\lambda) \in C_{p,\theta}$ for all $\lambda \geq 0$ small enough. Since $q \in B$, it follows from Lemma 5.9 that $\gamma_q(\lambda) \in B$ for arbitrarily small positive values of λ. Hence, one can find $\overline{\lambda} > 0$ such that $\gamma_q(\lambda) \in C_{p,\theta}$ for all $0 \leq \lambda \leq \overline{\lambda}$ and $\gamma_q(\overline{\lambda}) \in B$. It follows that one can find $\eta > 0$ such that the curve c defined on $[0, t+\eta]$ by
$$c(u) := \begin{cases} \gamma(u) & \text{if } u \in [0, t], \\ \gamma_q\left(\dfrac{u-t}{1-t}\right) & \text{if } u \in [t, t+\eta], \end{cases}$$

satisfies $c([0,t+\eta]) \subset C_{p,\theta}$ and $c(t+\eta) \in B$. Since γ_q is $L_{p,\theta}$-Lipschitz, the curve c is $(1-t)^{-1}L_{p,\theta}$-Lipschitz, and hence L-Lipschitz, on $[t,t+\eta]$. Finally, by the definition of γ_q (recall (5.11)) and since $x_1(q) = (1-t)x_1(p)$, we have

$$x_1(c(t+\eta)) = \left(1 - \frac{\eta}{1-t}\right)x_1(q) = (1-(t+\eta))\,x_1(p)$$

and hence $c(t+\eta) \in E_{p,\theta}(t+\eta)$. This shows that $t+\eta \in I$, which gives a contradiction.

Hence, for all points $p \in \partial B$ with $x_1(p) \neq 0$ and $s \in (0,1)$, we have $E_{p,\theta}(s) \cap B \neq \emptyset$ for all $\theta \in (0,\pi/2)$. Letting $\theta \downarrow 0$ and since B is closed, it follows that

$$[(1-s)x_1(p), x_2(p), x_3(p)] \in B$$

for all $s \in (0,1)$. Using once again the fact that B is closed, we finally get that the closed segment $\hat{\sigma}_p$ is contained in B as wanted. □

Lemma 5.13 to Lemma 5.20 below are successive consequences of Lemma 5.10. In addition to Lemma 5.10, the only properties used to prove these lemmas are the left-invariance of the quasi-distance d and topological properties of the unit ball B.

Lemma 5.13. *For all $p = (x_p, y_p, z_p) \in \partial B$ with $x_p \neq 0$, the segment*

$$\sigma_p := \left\{\left((1-t)x_p, y_p, z_p - t\frac{x_p y_p}{2}\right) : t \in [0,1]\right\}$$

is contained in B.

Proof. By the left-invariance of d, we have $q^{-1} \in B$ for all $q \in B$. Then it follows from Lemma 5.10 that for $p \in \partial B$ with $x_p \neq 0$, $\sigma_p = (\hat{\sigma}_{p^{-1}})^{-1} \subset B$. □

Given $p = (0, y_p, z_p) \in \mathbb{H}$ and $w > 0$, we set

$$D(p,w) := \{(-t, y_p, z_p - tw + u) : t \geq 0,\ u \geq 0\}.$$

It is the two-dimensional region in the plane $\{y = y_p\}$ above the half-line starting at p with direction $(-1, 0, -w)$.

Lemma 5.14. *Let $\overline{y} > 0$ be such that $(0, \overline{y}, z) \in B$ for some $z > 0$. Furthermore, set $\overline{z} := \max\{z > 0 : (0,\overline{y},z) \in B\}$ and $\overline{p} := (0,\overline{y},\overline{z})$. Then $\overline{p} \in \partial B$ and for all $0 < w < \overline{y}/2$, we have*

$$D(\overline{p}, w) \cap B = \{\overline{p}\}. \tag{5.15}$$

Proof. By contradiction, assume that there is some point $q \in D(\overline{p}, w) \cap B$ with $q \neq \overline{p}$. Then, by the definition of $D(\overline{p}, w)$, we have $q = (-t, \overline{y}, \overline{z} - tw + u)$ for some $u, t \geq 0$. Since $q \neq \overline{p}$ and by the choice of $\overline{z}$, we have $t > 0$ and hence $x_q = -t \neq 0$. By Lemma 5.13, it follows that $\sigma_q \subset B$. In particular, the end point $(0, \overline{y}, \overline{z} - tw + u + t\overline{y}/2)$ of σ_q belongs to B. By the choice of $\overline{z}$, we must then have $-tw + u + t\overline{y}/2 \leq 0$, and hence

$$t\left(w - \frac{\overline{y}}{2}\right) \geq u \geq 0.$$

This contradicts the assumption $0 < w < \overline{y}/2$ and concludes the proof. □

Given $p = (0, y_p, z_p) \in \mathbb{H}$ with $y_p, z_p > 0$ and $v > 0$, we set

$$S(p, v) := \{(0, (1-s)y_p, z_p + sv + u) : s \in [0, 1],\ u \geq 0\}.$$

It is the two-dimensional region in the plane $\{x = 0\}$ above the segment from the z-axis to p with slope $-v/y_p$.

Lemma 5.16. *There exists a point $p = (0, y_p, z_p) \in \partial B$ with $y_p, z_p > 0$ and $v > 0$ such that*

$$S(p, v) \cap B = \{p\}$$

and such that, for all $0 \leq y \leq y_p$, $(0, y, z) \in B$ for some $z > 0$.

Proof. The only property of the unit ball B used in this proof is the fact that B is a compact neighborhood of the origin. Let $p_0 = (0, y_0, z_0)$ with $y_0, z_0 > 0$ be a point in the interior of B such that $(0, y, z) \in B$ for all $0 \leq y \leq y_0$ and all $0 \leq z \leq z_0$. For $q = (0, y_q, z_q)$ with $y_q, z_q > 0$, let

$$I(q) := \{(0, y, z_q) : 0 \leq y \leq y_q\}$$

denote the horizontal segment from the z-axis to q and let

$$I^+(q) := \{(0, y, z_q + u) : 0 \leq y \leq y_q,\ u > 0\}$$

denote the two-dimensional infinite rectangular strip strictly above $I(q)$ in the plane $\{x = 0\}$. Set

$$\overline{z} := \inf\{z > z_0 : I^+((0, y_0, z)) \cap B = \emptyset\}.$$

Since B is bounded and p_0 belongs to the interior of B, we have $z_0 < \overline{z} < +\infty$. We set $\overline{q} := (0, y_0, \overline{z})$. We have $I^+(\overline{q}) \cap B = \emptyset$ and, since B is closed, $I(\overline{q}) \cap B \neq \emptyset$.

If there is some point $p = (0, y_p, \overline{z}) \in I(\overline{q}) \cap B$ with $y_p > 0$, then, for any $v > 0$, we have $S(p, v) \setminus \{p\} \subset I^+(\overline{q})$ hence $S(p, v) \cap B = \{p\}$. Note that $p \in \partial B$.

Otherwise $I(\overline{q}) \cap B = \{(0, 0, \overline{z})\}$. For $q = (0, y_0, z_q)$ with $z_q > 0$, let

$$J(q) := \{(0, (1-t)y_0, z_q + t(\overline{z} - z_0) : t \in [0, 1]\}$$

denote the segment in the plane $\{x = 0\}$ from the z-axis to q with slope $-(\overline{z} - z_0)/y_0$ and let

$$J^+(q) := \{(0, (1-t)y_0, z_q + t(\overline{z} - z_0) + u : t \in [0, 1],\ u > 0\}$$

the two-dimensional region strictly above $J(q)$ in the plane $\{x = 0\}$. Set

$$\hat{z} := \inf\{z > z_0 : J^+((0, y_0, z)) \cap B = \emptyset\}.$$

Arguing as above, we have $z_0 < \hat{z} \leq \overline{z}$. We set $\hat{q} := (0, y_0, \hat{z})$. We have $J^+(\hat{q}) \cap B = \emptyset$ and $J(\hat{q}) \cap B \neq \emptyset$. Since $\hat{z} - z_0 > 0$, $J(\hat{q})$ meets the z-axis at $(0, 0, \overline{z} + \hat{z} - z_0)$ which belongs to $I^+(\overline{q})$. Hence it cannot belong to B and there is some point $p = (0, y_p, z_p) \in J(\hat{q}) \cap B$ with $y_p > 0$. Note that $z_p > 0$ and that p belongs to ∂B. Then for any $v > (\overline{z} - z_0)y_p/y_0$, we have $S(p, v) \setminus \{p\} \subset J^+(\hat{q})$ hence $S(p, v) \cap B = \{p\}$.

Finally, in both cases, we have $y_p \leq y_0$. Then, for all $0 \leq y \leq y_p$, we have $0 \leq y \leq y_0$ and, by the choice of p_0, we get that $(0, y, z) \in B$ for all $0 \leq z \leq z_0$. In particular, $(0, y, z) \in B$ for some $z > 0$, which concludes the proof. □

Lemma 5.17 below is a consequence of Lemma 5.14 together with Lemma 5.16. Given $p = (0, y_p, z_p) \in \mathbb{H}$ with $y_p, z_p > 0$ and $v > 0$, we set

$$R(p,v) := \left\{ \left(-t, (1-s)y_p, z_p + sv - \frac{t(1-s)y_p}{4} + u \right) : u, t \geq 0,\ s \in [0,1) \right\}.$$

It is the three-dimensional region obtained by the following union. For $s \in [0,1]$, let

$$p_s := (0, (1-s)y_p, z_p + sv).$$

Note that the set $\{p_s : s \in [0,1]\}$ is the lower boundary of $S(p,v)$. Then, for $s \in [0,1)$, the set $D(p_s, (1-s)y_p/4)$ is the intersection of $R(p,v)$ with the plane $\{y = (1-s)y_p\}$ and we have

$$R(p,v) = \bigcup_{s \in [0,1)} D(p_s, (1-s)\tfrac{y_p}{4}).$$

Lemma 5.17. *There exists a point $p = (0, y_p, z_p) \in \partial B$ with $y_p, z_p > 0$ and $v > 0$ such that $R(p,v) \cap B = \{p\}$.*

Proof. Let $p = (0, y_p, z_p) \in \partial B$ with $y_p, z_p > 0$ and $v > 0$ be given by Lemma 5.16. Since $S(p,v) \cap B = \{p\}$, we have $z_p = \max\{z > 0 : (0, y_p, z) \in B\}$. Then it follows from Lemma 5.14 that

$$D(p, \tfrac{y_p}{4}) \cap B = \{p\}. \tag{5.18}$$

Recall that $D(p, y_p/4)$ is the intersection of $R(p,v)$ with the plane $\{y = y_p\}$.

For $s \in (0,1)$, set $q_s := (0, (1-s)y_p, z_s)$, where $z_s := \max\{z > 0 : (0, (1-s)y_p, z) \in B\}$. Note that z_s is well defined since for all $s \in (0,1)$, we have $(0, (1-s)y_p, z) \in B$ for some $z > 0$ by Lemma 5.16. We have $q_s \in \partial B$ and it follows from Lemma 5.14 that

$$D(q_s, (1-s)\tfrac{y_p}{4}) \cap B = \{q_s\}.$$

On the other hand, $q_s \notin S(p,v)$ since $S(p,v) \cap B = \{p\}$ and hence $z_s < z_p + sv$. It follows that, for all $s \in (0,1)$,

$$D(p_s, (1-s)\tfrac{y_p}{4}) \cap B \subset (D(q_s, (1-s)\tfrac{y_p}{4}) \cap B) \setminus \{q_s\},$$

where $p_s := (0, (1-s)y_p, z_p + sv)$. Hence,

$$D(p_s, (1-s)\tfrac{y_p}{4}) \cap B = \emptyset. \tag{5.19}$$

Recalling that $D(p_s, (1-s)y_p/4)$ is the intersection of $R(p,v)$ with the plane $\{y = (1-s)y_p\}$, the lemma finally follows from (5.18) and (5.19). □

Lemma 5.20. *There exists a point $q = (0, y_q, z_q) \in \partial B$ with $y_q, z_q < 0$ and $v > 0$ such that $\hat{R}(q,v) \cap B_d(q,1) = \{0\}$, where*

$$\hat{R}(q,v) := \left\{ \left(-t, sy_q, sv + \frac{t(3-s)y_q}{4} + u \right) : u, t \geq 0,\ s \in [0,1) \right\}. \tag{5.21}$$

Proof. Let $p = (0, y_p, z_p) \in \partial B$ with $y_p, z_p > 0$ and $v > 0$ be given by Lemma 5.17 and set $q := p^{-1}$. By left-invariance of d, we have

$$q \cdot R(q^{-1}, v) \cap B_d(q,1) = p^{-1} \cdot (R(p,v) \cap B) = \{0\}.$$

Noting that $q \cdot R(q^{-1}, v) = p^{-1} \cdot R(p,v) = \hat{R}(q,v)$, we get the required conclusion. □

The next lemma gives a geometric property of dilations of the region $\hat{R}(q,v)$.

Lemma 5.22. *Let $q=(0,y_q,z_q)\in\mathbb{H}$ with $y_q,z_q<0$, $v>0$ and let $\hat{R}(q,v)$ be given by* (5.21)*. Then, for all $\hat{q}=(x_{\hat{q}},y_{\hat{q}},z_{\hat{q}})$ with $x_{\hat{q}},y_{\hat{q}}<0$, there exists $\hat{\lambda}>0$ such that, for all $\lambda>\hat{\lambda}$, $\hat{q}\in\delta_\lambda(\hat{R}(q,v))$.*

Proof. Let $\hat{q}=(x_{\hat{q}},y_{\hat{q}},z_{\hat{q}})$ with $x_{\hat{q}},y_{\hat{q}}<0$ be given. To prove that $\hat{q}\in\delta_\lambda(\hat{R}(q,v))$, we have to find $t\ge 0$, $u\ge 0$ and $s\in[0,1)$ such that

$$
\begin{cases}
\hat{x}_q=-t\lambda,\\
\hat{y}_q=\lambda^\alpha s y_q,\\
\hat{z}_q=\lambda^{\alpha+1}\Big(sv+\dfrac{t(3-s)y_q}{4}+u\Big).
\end{cases}
\tag{5.23}
$$

From the first equation, we get $t=-\hat{x}_q/\lambda>0$ since $\hat{x}_q<0$. From the second equation, we get $s=\lambda^{-\alpha}\hat{y}_q y_q^{-1}$. We have $s>0$ since $\hat{y}_q<0$ and $y_q<0$. We also have $s<1$ for all $\lambda>0$ large enough. The third equation gives

$$
\begin{aligned}
u&=-\frac{t(3-s)y_q}{4}-sv+\lambda^{-\alpha-1}\hat{z}_q\\
&=\frac{\hat{x}_q y_q}{4\lambda}\Big(3-\lambda^{-\alpha}\frac{\hat{y}_q}{y_q}\Big)-\lambda^{-\alpha}\frac{\hat{y}_q}{y_q}v+\lambda^{-\alpha-1}\hat{z}_q\\
&=\frac{3\hat{x}_q y_q}{4}\frac{1}{\lambda}+o\Big(\frac{1}{\lambda}\Big).
\end{aligned}
$$

It follows that $u>0$ for $\lambda>0$ large enough. All together, we get that, for all $\lambda>0$ large enough, one can find $t\ge 0$, $u\ge 0$ and $s\in[0,1)$ such that (5.23) holds as wanted. □

We are now going to conclude the proof of Theorem 5.6.

Proof of Theorem 5.6. Recall that we are arguing by contradiction. We consider a continuous self-similar quasi-distance d on the non-standard Heisenberg group of exponent α and we are assuming that d satisfies WBCP. To get a contradiction, we are going to construct with the help of Lemma 5.20 and Lemma 5.22 families of Besicovitch balls with arbitrarily large cardinality.

Let us choose a point $q_1=(x_1,y_1,z_1)$ with $x_1,y_1<0$ and set $r_1:=d(0,q_1)$. By induction assume that $q_1=(x_1,y_1,z_1),\dots,q_m=(x_m,y_m,z_m)$ have already been chosen so that $x_i,y_i<0$ for all $i=1,\dots,m$ and so that $\{B_d(q_i,r_i)\}_{i=1}^m$ is a family of Besicovitch balls where $r_i:=d(0,q_i)$.

Let $q=(0,y_q,z_q)\in\partial B$ with $y_q,z_q<0$ and $v>0$ be given by Lemma 5.20. For all $k\ge 1$, we have

$$\delta_{\rho^{-k}}(\hat{R}(q,v))\cap B_d(\delta_{\rho^{-k}}(q),\rho^{-k})=\delta_{\rho^{-k}}(\hat{R}(q,v)\cap B_d(q,1))=\{0\}.$$

On the other hand, it follows from Lemma 5.22 that $q_i\in\delta_{\rho^{-k}}(\hat{R}(q,v))$ for all $i=1,\dots,m$ and all $k\ge 1$ large enough. Hence $q_i\notin B_d(\delta_{\rho^{-k}}(q),\rho^{-k})$ for all $i=1,\dots,m$ and all $k\ge 1$ large enough.

Next, by continuity of d with respect to the manifold topology, we have

$$\lim_{k\to+\infty} d(\delta_{\rho^k}(q_i), q) = d(0, q)$$

for all $i = 1, \dots, m$. It follows that

$$\lim_{k\to+\infty} d(q_i, \delta_{\rho^{-k}}(q)) = \lim_{k\to+\infty} \rho^{-k} d(\delta_{\rho^k}(q_i), q) = +\infty$$

for all $i = 1, \dots, m$. Hence we can also choose $k \geq 1$ large enough so that

$$d(q_i, \delta_{\rho^k}(q)) > r_i$$

for all $i = 1, \dots, m$.

All together, we have proved that we can find an integer $k \geq 1$ large enough so that the balls $B_d(q_1, r_1), \dots, B_d(q_m, r_m), B_d(\delta_{\rho^{-k}}(q), \rho^{-k})$ form a family of Besicovitch balls. We have

$$\delta_{\rho^{-k}}(q) = (0, \rho^{-k\alpha} y_q, \rho^{-k(\alpha+1)} z_q)$$

with $\rho^{-k\alpha} y_q < 0$, and $\rho^{-k} = d(0, \delta_{\rho^{-k}}(q))$. Then, using Remark 5.24 below, we can choose $q_{m+1} = (x_{m+1}, y_{m+1}, z_{m+1})$ with $x_{m+1}, y_{m+1} < 0$ close enough to $\delta_{\rho^{-k}}(q)$ so that, setting $r_{m+1} := d(0, q_{m+1})$, the family $\{B_d(q_i, r_i)\}_{i=1}^{m+1}$ is a family of Besicovitch balls. □

Remark 5.24 (Being a family of Besicovitch balls is an open condition). Let us assume that $\{B_d(q_i, r_i)\}_{i=1}^m$ is a family of Besicovitch balls, where $r_i := d(e, q_i)$ in a graded group G with identity e and equipped with a continuous self-similar quasi-distance d. One can find $U_1, \dots, U_m$ open neighborhoods of $q_1, \dots, q_m$, respectively, such that, for all points $(q'_1, \dots, q'_m) \in U_1 \times \dots \times U_m$, $\{B_d(q'_i, r'_i)\}_{i=1}^m$ is a family of Besicovitch balls. Here we have set $r'_i := d(e, q'_i)$. Indeed, we have $d(q_i, q_j) - d(e, q_i) > 0$ for all $i \neq j$. By the continuity of d on $G \times G$ with respect to the manifold topology, one can find $U_1, \dots, U_m$ open neighborhoods of $q_1, \dots, q_m$, respectively, such that, if $(q'_1, \dots, q'_m) \in U_1 \times \dots \times U_m$, then $d(q'_i, q'_j) - d(e, q'_i) > 0$ for all $i \neq j$. Hence $\{B_d(q'_i, r'_i)\}_{i=1}^m$ is a family of Besicovitch balls.

5.2. Topological properties of self-similar distances. As already mentioned, one of the main differences between self-similar and homogeneous quasi-distances are their topological properties. One cannot extend Proposition 2.26 to self-similar quasi-distances. There are indeed examples of self-similar distances on homogeneous groups such that the distance from the identity e is not continuous at e with respect to the manifold topology. Hence the topology induced by such self-similar distances does not coincide with the manifold topology.

One such example is the following. We consider $\mathbb{R}$ equipped with the usual addition as a group law and the dilations $\delta_\lambda(x) := \lambda x$. We take $(v_i)_{i\in I}$ a basis of $\mathbb{R}$ viewed as a vector space over $\mathbb{Q}$ and we choose it in such a way that some sequence v_{i_j} converges to 0 as j goes to $+\infty$ for the usual topology of $\mathbb{R}$. For $x \in \mathbb{R}$, we write $x = \sum x_i v_i$, where $x_i \in \mathbb{Q}$ and all but finitely many of the points x_i are 0 and we consider the left-invariant distance d such that $d(0, x) = \sum |x_i|$. This distance is $\mathbb{Q}$-homogeneous, i.e., $d(\delta_q(x), \delta_q(y)) = q d(x, y)$ for all x, $y \in \mathbb{R}$ and all $q \in \mathbb{Q}$, and hence self-similar. We have $d(0, v_i) = 1$ for all $i \in I$. In particular, $d(0, v_{i_j}) = 1$ for all j whereas v_{i_j} converges to 0 as j goes to $+\infty$ for the usual topology of $\mathbb{R}$. Hence $d(0, \cdot)$ is not continuous at 0 with respect to the manifold topology on $\mathbb{R}$.

However, with the additional assumption of the continuity of $d(e, \cdot)$ at e with respect to the manifold topology, one can extend Proposition 2.26 in the following way.

Proposition 5.25. *Let G be a graded group with identity e. Let d be a self-similar quasi-distance on G. Assume that $d(e,\cdot)$ is continuous at e with respect to the manifold topology. Then the topology induced by d coincides with the manifold topology. Moreover, a set is relatively compact if and only if it is bounded with respect to d.*

Proof. First, the fact that $d(e,\cdot)$ is assumed to be continuous at e with respect to the manifold topology $\mathcal{T}_m$ implies that $\mathcal{T}_d \subset \mathcal{T}_m$ where $\mathcal{T}_d$ denotes the topology induced by d.

Then the proof can be completed with technical modifications of the proof of Proposition 2.26 that we briefly sketch below. In the rest of this proof, the convergence of some sequence of points means convergence with respect to the manifold topology.

Let $0 < \lambda < 1$ be such that $d(\delta_\lambda(p), \delta_\lambda(q)) = \lambda d(p,q)$ for all $p, q \in G$. Let (p_k) be a sequence such that $d(e,p_k)$ goes to 0 and let us prove that p_k converges to e. Using the conventions and notations introduced in the proof of Proposition 2.26, we argue by contradiction and assume that, up to a subsequence, there exists $\varepsilon > 0$ such that $\|p_k\| > \varepsilon$ for all k (see (2.27) for the definition of $\|\cdot\|$). For each fixed k, $\|\delta_{\lambda^l}(p_k)\|$ converges to 0 when l goes to $+\infty$. Hence one can find a sequence of integers $l_k \geq 1$ such that, for all k,

$$\|\delta_{\lambda^{l_k}}(p_k)\| \leq \varepsilon \leq \|\delta_{\lambda^{l_k-1}}(p_k)\|.$$

Then

$$\|\delta_{\lambda^{l_k}}(p_k)\| = \sum_{i=1}^n \lambda^{l_k d_i}|P_i(p)| \geq \lambda^{\overline{\alpha}} \sum_{i=1}^n \lambda^{(l_k-1)d_i}|P_i(p)| = \lambda^{\overline{\alpha}}\|\delta_{\lambda^{l_k-1}}(p_k)\| \geq \varepsilon\lambda^{\overline{\alpha}},$$

where $\overline{\alpha} := \max_{1\leq i\leq n} d_i$. Hence

$$\varepsilon\lambda^{\overline{\alpha}} \leq \|\delta_{\lambda^{l_k}}(p_k)\| \leq \varepsilon$$

for all k. By compactness with respect to the manifold topology of $\{p \in G : \varepsilon\lambda^{\overline{\alpha}} \leq \|p\| \leq \varepsilon\}$, we get that, up to a subsequence, $\delta_{\lambda^{l_k}}(p_k)$ converges to some $p \in G$ such that $\varepsilon\lambda^{\overline{\alpha}} \leq \|p\| \leq \varepsilon$. In particular, $p \neq e$ and $d(e,p) > 0$. On the other hand, we have

$$\begin{aligned} 0 < d(e,p) &\leq C(d(e,\delta_{\lambda^{l_k}}(p_k)) + d(\delta_{\lambda^{l_k}}(p_k), p)) \\ &= C(\lambda^{l_k} d(e,p_k) + d(e, p^{-1}\cdot\delta_{\lambda^{l_k}}(p_k))) \\ &\leq C(d(e,p_k) + d(e, p^{-1}\cdot\delta_{\lambda^{l_k}}(p_k))). \end{aligned}$$

Since $d(e,p_k)$ converges to 0 and since $d(e,\cdot)$ is assumed to be continuous at e with respect to the manifold topology and $p^{-1}\cdot\delta_{\lambda^{l_k}}(p_k)$ converges to e, we get that the last upper bound in the above inequalities goes to 0, which gives a contradiction. It follows $\mathcal{T}_m \subset \mathcal{T}_d$ and, all together, we get that $\mathcal{T}_m = \mathcal{T}_d$.

The proof that a set is relatively compact if and only if it is bounded with respect to d can be achieved with similar technical modifications of the arguments of the proof of Proposition 2.26. □

Proposition 5.25 implies the following generalization of Proposition 3.21 to self-similar quasi-distances. In the statement and the proof of Proposition 5.26 continuity means continuity with respect to the manifold topology.

Proposition 5.26. *Let $\hat{G}$ and G be graded groups with graded Lie algebra $\hat{\mathfrak{g}}$ and $\mathfrak{g}$, respectively. Assume that there exists a surjective morphism of graded Lie algebras $\phi : \hat{\mathfrak{g}} \to \mathfrak{g}$. Let $\varphi : \hat{G} \to G$ denote the unique Lie group homomorphism such that $\varphi_* = \phi$ and let $\hat{d}$ be a continuous self-similar quasi-distance on $\hat{G}$. Then*

$$d(p,q) := \hat{d}(\varphi^{-1}(\{p\}), \varphi^{-1}(\{q\}))$$

defines a continuous self-similar quasi-distance on G and $\varphi : (\hat{G}, \hat{d}) \to (G, d)$ is a submetry.

Proof. One proves that d is a self-similar quasi-distance on G and that

$$\varphi : (\hat{G}, \hat{d}) \to (G, d)$$

is a submetry with the same arguments as in the proof of Proposition 3.21. To prove that d is continuous on $G \times G$, we first prove that $d(e, \cdot)$ is continuous at e. Here e denotes the identity in G and below $\hat{e}$ will denote the identity in $\hat{G}$. Let (p_k) be a sequence converging to e (with respect to the manifold topology). Since φ is an open map and since $\hat{e} \in \varphi^{-1}(\{e\})$, one can find a sequence $\hat{p}_k \in \varphi^{-1}(\{p_k\})$ converging to $\hat{e}$. We have $d(e, p_k) \leq \hat{d}(\hat{e}, \hat{p}_k)$. Since $\hat{d}$ is assumed to be continuous, $\hat{d}(\hat{e}, \hat{p}_k)$ converges to 0. It follows that $d(e, p_k)$ converges to 0 as well and hence $d(e, \cdot)$ is continuous at e. The proof can now be completed using Proposition 5.25 and following the arguments of the proof of Proposition 3.21. □

5.3. Arbitrary graded groups with two different layers not commuting. We conclude in this subsection the proof of Theorem 5.2.

Proof of Theorem 5.2. Let G be a graded group with associated positive grading $\mathfrak{g} = \bigoplus_{t>0} V_t$ of its Lie algebra. Let $t < s$ be such that $[V_t, V_s] \neq \{0\}$. Let d be a continuous self-similar quasi-distance on G. We argue by contradiction and assume that WBCP holds on (G, d).

By Proposition 2.18, there exists a graded subalgebra $\hat{\mathfrak{g}}$ of $\mathfrak{g}$ and a surjective morphism of graded Lie algebras from $\hat{\mathfrak{g}}$ onto $\mathfrak{h}$, where $\mathfrak{h}$ is the t-power of the non-standard Heisenberg Lie algebra of exponent s/t.

We denote by $\hat{G} := \exp(\hat{\mathfrak{g}})$ the graded group whose Lie algebra $\hat{\mathfrak{g}}$ is endowed with the positive grading induced by the given positive grading of $\mathfrak{g}$. The restriction of d to $\hat{G}$, still denoted by d, is a continuous self-similar quasi-distance on $\hat{G}$. Since we assume that WBCP holds on (G, d), WBCP also holds on $(\hat{G}, d)$ by Proposition 3.15.

Next, it follows from Proposition 5.26 that there exists a continuous self-similar quasi-distance d_H on the t-power, denoted by H, of the non-standard Heisenberg group of exponent s/t, and a submetry from $(\hat{G}, d)$ onto (H, d_H). Since WBCP holds on $(\hat{G}, d)$, WBCP holds on (H, d_H) by Proposition 3.20. Finally, we get that WBCP would hold for the continuous self-similar quasi-distance $(d_H)^t$ on the non-standard Heisenberg group of exponent s/t. This contradicts Theorem 5.6 and concludes the proof. □

6. Differentiation of measures

In this section, we give applications to differentiation of measures. In particular, we prove Theorem 1.5.

Following the terminology of [31], if (X, d) is a metric space, we say that d is *finite dimensional* on a subset $Y \subset X$ if there exist $C^* \in [1, +\infty)$ and $r^* \in (0, +\infty]$ such that $\operatorname{Card} \mathcal{B} \leq C^*$ for every family $\mathcal{B} = \{B = B_d(x_B, r_B)\}$ of Besicovitch balls in (X, d) such that $x_B \in Y$ and $r_B < r^*$ for all $B \in \mathcal{B}$ (see Definition 3.2 for the definition of families of Besicovitch balls). If we need to specify the constants C^* and r^*, we say that d is finite dimensional on Y with constants C^* and r^*. We say d is *σ-finite dimensional* if X can be written as a countable union of subsets on which d is finite dimensional. Note that WBCP holds on (X, d) if and only if d is finite dimensional on X for some constant $C^* \in [1, +\infty)$ and with $r^* = +\infty$.

For homogeneous distances on homogeneous groups, we first prove the following equivalence between σ-finite dimensionality and validity of (W)BCP.

Proposition 6.1. *Let G be a homogeneous group and let d be a homogeneous distance on G. Then BCP holds on (G, d) if and only if d is σ-finite dimensional.*

Proof. If BCP holds on (G, d), then WBCP holds as well and it follows from the definitions that d is finite dimensional on G. To prove the converse, let m denote the Hausdorff dimension of (G, d) and let μ denote the m-dimensional Hausdorff measure on (G, d). It is well known that μ is a Haar measure on G. Since the distance d from which this m-dimensional Hausdorff measure is constructed is homogeneous, μ is moreover m-homogeneous with respect to the associated dilations $(\delta_\lambda)_{\lambda>0}$ on G, i.e., $\mu(\delta_\lambda(A)) = \lambda^m \mu(A)$ for all $A \subset G$ and all $\lambda > 0$. In particular, μ is a doubling outer measure on (G, d). It follows that for every subset $A \subset G$, μ-a.e. point $p \in A$ is a μ-density point for A, i.e.,

$$\lim_{r \downarrow 0} \frac{\mu(A \cap B_d(p, r))}{\mu(B_d(p, r))} = 1.$$

Assume that d is σ-finite dimensional, i.e., $G = \bigcup_{n \in \mathbb{N}} G_n$, where d is finite dimensional on each G_n. Then one can find $n \in \mathbb{N}$ such that $\mu(G_n) > 0$ and we set $A := G_n$. Let $C^* \in [1, +\infty)$ and $r^* \in (0, +\infty]$ be such that d is finite dimensional on A with constant C^* and r^*. Since $\mu(A) > 0$, one can find a point $p \in A$ that is a μ-density point for A. Up to a translation, one can assume that $p = e$ where e denotes the identity in G. Next, by homogeneity, for every $\lambda \geq 1$, d is finite dimensional on $\delta_\lambda(A)$ with constants C^* and r^*.

Let us prove that WBCP, and hence BCP by Corollary 3.8, holds in (G, d). We consider $\{B_d(p_i, r_i)\}_{i=1}^k$ a family of Besicovitch balls in (G, d). Up to a dilation, one can assume with no loss of generality that $r_i < r^*$ for all $i = 1, \dots, k$. Up to a translation, one can further assume that $e \in \bigcap_{i=1}^k B_d(p_i, r_i)$. Shrinking balls if necessary, one can also assume that, for all $i = 1, \dots, k$, $r_i = d(e, p_i)$.

By Remark 5.24 (recall that homogeneous distances on a homogeneous group are continuous, see Corollary 2.28), one can find a constant $\varepsilon^* > 0$ such that $\{B_d(q_i, s_i)\}_{i=1}^k$ is a family of Besicovitch balls in (G, d) as soon as $d(p_i, q_i) \leq \varepsilon^*$ for all $i = 1, \dots, k$ and where $s_i := d(e, q_i)$. Moreover, one can choose ε^* small enough, namely $\varepsilon^* < r^* - \max_{1 \leq i \leq k} r_i$, so that $d(e, q) < r^*$ as soon as $d(q, p_i) \leq \varepsilon^*$ for some $i \in \{1, \dots, k\}$.

Fix $R \geq 1$ such that $R \geq 2 \max_{1 \leq i \leq k} r_i$ and $\varepsilon > 0$ such that $\varepsilon \leq \min(\varepsilon^*, \max_{1 \leq i \leq k} r_i)$. Let $r < 1$ be such that

$$\frac{\mu(A \cap B_d(e, r))}{\mu(B_d(e, r))} > 1 - \left(\frac{\varepsilon}{R}\right)^m.$$

Then we claim that $B_d(p_i,\varepsilon)\cap\delta_{R/r}(A)\neq\emptyset$ for all $i=1,\dots,k$. Indeed, arguing by contradiction, assume that $B_d(p_i,\varepsilon)\cap\delta_{R/r}(A)=\emptyset$. By the choice of R and ε, we have the inclusion $B_d(p_i,\varepsilon)\subset B_d(e,R)$. Using the left-invariance and the homogeneity of μ, it follows that

$$\begin{aligned}1-\left(\frac{\varepsilon}{R}\right)^m &< \frac{\mu(A\cap B_d(e,r))}{\mu(B_d(e,r))}\\ &= \frac{\mu(\delta_{R/r}(A)\cap B_d(e,R))}{\mu(B_d(e,R))}\\ &\leq \frac{\mu(B_d(e,R)\setminus B_d(p_i,\varepsilon))}{\mu(B_d(e,R))}\\ &= \frac{\mu(B_d(e,R))-\mu(B_d(p_i,\varepsilon))}{\mu(B_d(e,R))}\\ &= 1-\left(\frac{\varepsilon}{R}\right)^m,\end{aligned}$$

which gives a contradiction.

Hence one can find $q_i\in B_d(p_i,\varepsilon)\cap\delta_{R/r}(A)$ for all $i=1,\dots,k$. Since $\varepsilon\leq\varepsilon^*$, we get that $\{B_d(q_i,s_i)\}_{i=1}^k$ is a family of Besicovitch balls in (G,d) with $q_i\in\delta_{R/r}(A)$ and $s_i:=d(e,q_i)<r^*$ for all $i=1,\dots,k$. By the choice of R and r, we have $R/r\geq 1$ and it follows that $k\leq C^*$. Hence WBCP holds in (G,d). □

Recall now that if (X,d) is a metric space and μ is a locally finite Borel measure on X, we say that *the differentiation theorem holds on* (X,d) *for the measure* μ if

$$\lim_{r\downarrow 0+}\frac{1}{\mu(B_d(p,r))}\int_{B_d(p,r)} f(q)\,d\mu(q)=f(p)$$

for μ-almost every $p\in X$ and all $f\in L^1_{\mathrm{loc}}(\mu)$.

The connection between σ-finite dimensionality and measure differentiation in the general metric setting is given by the following result due to D. Preiss.

Theorem 6.2 ([31]). *Let (X,d) be a complete separable metric space. The differentiation theorem holds on (X,d) for all locally finite Borel measures if and only if d is σ-finite dimensional.*

The proof of Theorem 1.5 can now be completed using Theorem 6.2 and Proposition 6.1.

7. Sub-Riemannian distances

In this section we consider sub-Riemannian distances on stratified groups, and more generally on sub-Riemannian manifolds. We prove Theorem 1.9 and Theorem 1.10. The proofs of these results are independent of the main result in this paper, namely independent of Theorem 1.2, but use some of the techniques developed here, in particular Proposition 3.21.

7.1. No BCP in sub-Riemannian Carnot groups. Let G be a stratified group with associated stratification of its Lie algebra given by $\mathfrak{g}=V_1\oplus\cdots\oplus V_s$. We say that an absolutely continuous curve $\gamma:I\to G$ is *horizontal* if one has $\dot\gamma(s)\in\operatorname{span}\{X(\gamma(s));\ X\in V_1\}$ for a.e. $s \in I$. For a given scalar product inducing a norm $\|\cdot\|$ on the first layer V_1 of the stratification, we define the length, with respect to $\|\cdot\|$, of a horizontal curve γ by

$$l_{\|\cdot\|}(\gamma) := \int_I \|\dot\gamma(s)\|\, ds.$$

We say that d is a *sub-Riemannian distance* on G if there exists a norm $\|\cdot\|$ induced by a scalar product on V_1 such that

$$d(p,q) = \inf\{l_{\|\cdot\|}(\gamma) : \gamma \text{ horizontal curve from } p \text{ to } q\}.$$

It is well known that such a d defines a homogeneous distance on G.

Definition 7.1 (Sub-Riemannian Carnot groups). We say that (G,d) is a *sub-Riemannian Carnot group* if G is a stratified group equipped with a sub-Riemannian distance d.

It is proved in [33] that BCP does not hold on sub-Riemannian Carnot groups of step at least 2 under some assumptions on the regularity of length-minimizing curves and of the sub-Riemannian distance (see [33, Theorem 1]). This result applies in particular to sub-Riemannian distances on the stratified first Heisenberg group as we recall now.

Theorem 7.2 ([33]). *Let d be a sub-Riemannian distance on the stratified first Heisenberg group $\mathbb{H}^1$. Then BCP does not hold on $(\mathbb{H}^1, d)$.*

We extend in Theorem 1.9 the results of [33] to any sub-Riemannian Carnot group of step ≥ 2 without any further regularity assumptions.

Theorem 1.9 will be proved by showing that from any such sub-Riemannian Carnot group, there exists a surjective morphism onto the stratified n-th Heisenberg group for some $n \in \mathbb{N}^*$. Moreover, the distance on the n-th Heisenberg group induced by this morphism is a sub-Riemannian distance (see Proposition 3.21 for the definition of this induced distance). Then the conclusion will follow from Theorem 7.2 and Lemma 7.4 below. Lemma 7.4 gives the existence of an isometric copy of a sub-Riemannian first Heisenberg group inside every sub-Riemannian n-th Heisenberg group. We refer to Example 2.11 for the definition of the n-th Heisenberg Lie algebra $\mathfrak{h}_n$ and the stratified n-th Heisenberg group $\mathbb{H}^n$. We first prove the following preliminary result.

Proposition 7.3. *Let (G,d) be a step 2 sub-Riemannian Carnot group with associated stratification of its Lie algebra given by $\mathfrak{g} = V_1 \oplus V_2$. Assume that $\dim V_2 = 1$ and that V_2 is the center of $\mathfrak{g}$. Then $\mathfrak{g}$ is the n-th Heisenberg Lie algebra $\mathfrak{h}_n$. Moreover, there exist positive real numbers $a_1,\dots,a_n$ and a standard basis $(X_1,\dots,X_n,Y_1,\dots,Y_n,Z)$ of $\mathfrak{h}_n$ such that d is the sub-Riemannian distance with respect to the norm induced by the scalar product on V_1 for which $(a_i X_i, a_i Y_i)_{1\leq i\leq n}$ is orthonormal.*

Proof. Let $Z \in V_2$ be fixed so that $V_2 = \mathbb{R}Z$. Since V_2 is the center of $\mathfrak{g}$, the Lie bracket restricted to $V_1 \times V_1$ can be identified with a non-degenerate skew-symmetric bilinear form on V_1. Then the following facts follow from classical results of linear algebra. First, $\dim V_1$ is even, $\dim V_1 = 2n$ for some $n \in \mathbb{N}^*$. Second, considering V_1 equipped with the scalar product with respect to which the sub-Riemannian d is defined, one can find an orthonormal basis $(\tilde{X}_1, \dots, \tilde{X}_n, \tilde{Y}_1, \dots, \tilde{Y}_n)$ of V_1 such that $[X_i, Y_j] = \delta_{ij}\alpha_i Z$ for some positive real numbers $\alpha_1, \dots, \alpha_n$. We set $X_i := (\sqrt{\alpha_i})^{-1}\tilde{X}_i$. Then $(X_1, \dots, X_n, Y_1, \dots, Y_n, Z)$ is a basis of $\mathfrak{g}$ with the only non-trivial bracket relations $[X_i, Y_i] = Z$ for $1 \le i \le n$. It follows that $\mathfrak{g}$ is the n-th Heisenberg Lie algebra $\mathfrak{h}_n$ and $(X_1, \dots, X_n, Y_1, \dots, Y_n, Z)$ is a standard basis of $\mathfrak{h}_n$. Moreover, setting $a_i := \sqrt{\alpha_i}$, we get that the sub-Riemannian distance d is the sub-Riemannian distance for which $(a_i X_i, a_i Y_i)_{1\le i\le n}$ is orthonormal. □

Lemma 7.4 below relies on classical results from sub-Riemannian geometry and more specifically on a classical study of length-minimizing curves in the sub-Riemannian Heisenberg groups. In particular, length-minimizing curves can be explicitly computed using Pontryagin's maximum principle and the fact that there are no strictly abnormal curves. Also each horizontal curve can be recovered from its projection on the first layer of the stratification via the lift of the spanned area in each coordinate plane.

Lemma 7.4. *Let $(X_1, \dots, X_n, Y_1, \dots, Y_n, Z)$ be a standard basis of $\mathfrak{h}_n$. Further, let $0 < a_1 \le \dots \le a_n$ be positive real numbers and let d denote the sub-Riemannian distance on $\mathbb{H}^n$ for which $(a_i X_i, a_i Y_i)_{1\le i\le n}$ is orthonormal. Then $H := \exp(\mathrm{span}\{X_n, Y_n, Z\})$ is geodesically closed and the distance d restricted to H is the sub-Riemannian distance on H for which $(a_n X_n, a_n Y_n)$ is orthonormal.*

Proof. In the statement H is geodesically closed means that for any two points $p, q \in H$ there exists a length-minimizing curve, i.e., a horizontal curve γ such that $d(p,q) = l_{\|\cdot\|}(\gamma)$, joining p and q and whose image is contained in H. Here $\|\cdot\|$ denotes the norm induced by the scalar product for which $(a_i X_i, a_i Y_i)_{1\le i\le n}$ is orthonormal. Then the fact that the distance d restricted to H is the sub-Riemannian distance on H for which $(a_n X_n, a_n Y_n)$ is orthonormal follows immediately.

We use in this proof exponential coordinates $(x_1, \dots, x_n, y_1, \dots, y_n, z)$ of the first kind with respect to the basis $(a_1 X_1, \dots, a_n X_n, a_1 Y_1, \dots, a_n Y_n, Z)$ and V_1 denotes the first layer of the stratification, $V_1 = \mathrm{span}\{X_i, Y_i : 1 \le i \le n\}$. By left-invariance, it is enough to show that given any point $p \in H$, there exists a length-minimizing curve joining 0 to p whose image is contained in H. As already said, the proof below relies on classical results from sub-Riemannian geometry. We refer to, e.g., [28] or [23] for more details.

Let us first consider $p = \exp(Z) = (0, \dots, 0, 1)$. Writing down the normal geodesic equation, one can see that if $\gamma : I \to \mathbb{H}^n$ is a length-minimizing curve from 0 to p with $\gamma(s) = (x_1(s), \dots, x_n(s), y_1(s), \dots, y_n(s), z(s))$, then its projections $s \mapsto (x_i(s), y_i(s))$ on each (x_i, y_i)-plane is a circle passing through the origin. Moreover, if we denote by (v_i, v_{n+i}) the center of each one of these circles, we have

$$l_{\|\cdot\|}(\gamma) = 2\pi \left(\sum_{i=1}^{2n} v_i^2 \right)^{1/2}.$$

On the other hand, since γ is a horizontal curve, the last coordinate of its end point is equal to $\sum_{i=1}^{n} a_i^2 \pi (v_i^2 + v_{n+i}^2)$. Hence we must have

$$\sum_{i=1}^{n} a_i^2 \pi (v_i^2 + v_{n+i}^2) = 1.$$

Notice that the minimum of $\sqrt{\sum_{i=1}^{2n} v_i^2}$ on the ellipsoid $\{\sum_{i=1}^{n} a_i^2 \pi (v_i^2 + v_{n+i}^2) = 1\}$ is attained on the smallest axis of the ellipsoid. It follows that the horizontal curve γ contained in H and whose projection on the (x_n, y_n)-plane is a circle passing through the origin and centered at $(1/a_n\sqrt{\pi}, 0)$ is a length-minimizing curve between 0 and p.

Next, from the above length-minimizing curve between 0 and $\exp(Z)$ and using dilations and rotations in H around the z-axis, it can be seen that one can join the origin to any point in $H \setminus \exp(V_1)$ with a length-minimizing curve that stays in H.

Finally, it can be checked that for $p \in H \cap \exp(V_1)$, the segment from 0 to p is the length minimizing curve from 0 to p and that this segment stays obviously inside H. □

We now conclude this section with the proof of the non-validity of BCP on sub-Riemannian Carnot groups of step ≥ 2.

Proof of Theorem 1.9. We argue by contradiction and we assume that (G, d) is a sub-Riemannian Carnot group of step ≥ 2 on which BCP holds. We prove that it is enough to consider the case of a sub-Riemannian distance on the stratified first Heisenberg group.

To prove this claim, we will perform a series of quotients. At the level of the Lie algebras, these quotients are surjective morphisms of stratified Lie algebras. Moreover, if we start from a sub-Riemannian distance, the distance induced by each one of these morphisms (as in Proposition 3.21), and for which BCP also holds by Propositions 3.20 and 3.21, will still be a sub-Riemannian distance. Indeed, if the stratification of the source Lie algebra is given by $V_1 \oplus \cdots \oplus V_s$, we will only consider morphisms of stratified Lie algebras whose kernel is contained either in $V_2 \oplus \cdots \oplus V_s$ or in the center of the source Lie algebra.

By a first quotient we can assume that the step of G is exactly 2. Note that in view of Corollary 1.3, we could have skipped this first step. However this first quotient gives a proof of Theorem 1.9 that is independent of our other results about more general graded groups. By a second quotient we can assume that the second layer of G is one-dimensional. By a third quotient we can assume that the center of the Lie algebra of G is the second layer of its stratification.

By Proposition 7.3 we know that the Lie algebra of the group G is the n-th Heisenberg Lie algebra. We also know that there exist a standard basis $(X_1, \dots, X_n, Y_1, \dots, Y_n, Z)$ of $\mathfrak{h}_n$ and positive real numbers $a_1, \dots, a_n$ such that d is the sub-Riemannian distance for which $(a_i X_i, a_i Y_i)_{1 \leq i \leq n}$ is orthonormal. Up to reordering the numbers a_j, it follows from Lemma 7.4 that there is a stratified subgroup H of G isomorphic to the stratified first Heisenberg group and such that the sub-Riemannian distance d restricted to H is a sub-Riemannian distance on H. Since we started from a sub-Riemannian distance on G for which BCP holds, it follows from Proposition 3.15 that BCP should hold for some sub-Riemannian distance on the stratified first Heisenberg group. This contradicts Theorem 7.2 and concludes the proof. □

7.2. Differentiation of measures on sub-Riemannian manifolds. In this subsection we prove that sub-Riemannian distances on sub-Riemannian manifolds are not σ-finite dimensional, see Theorem 7.5. Then Theorem 1.10 follows from Theorem 7.5 and Theorem 6.2. We refer to Section 6 for the definition of σ-finite dimensionality and its connection with measure differentiation. The proof of Theorem 7.5 follows from the fact that at regular points, the metric tangent space to a sub-Riemannian manifold is isometric to a sub-Riemannian Carnot group of step ≥ 2 together with Theorem 1.9.

To state our result, we first recall well-known facts about sub-Riemannian manifolds (see e.g. [27]). A sub-Riemannian manifold is a smooth Riemannian n-manifold (M, g) equipped with a smooth distribution Δ of k-planes, where $k < n$. We also assume that M is connected and that the distribution satisfies Hörmander's condition. Namely, for all $p \in M$, we assume that, if $X_1, \dots, X_k$ is a local basis of vector fields for the distribution near p, these vector fields, along with all their commutators, span $T_p M$. We denote by $V_i(p)$ the subspace of $T_p M$ spanned by all the commutators of the X_j's of order $\leq i$ evaluated at p. Then the distribution satisfies Hörmander's condition if, for all $p \in M$, there is some i such that $V_i(p) = T_p M$. Next, an absolutely continuous curve in M is said to be horizontal if it is a.e. tangent to the distribution Δ. Then the sub-Riemannian distance d between two points $p, q \in M$ is defined by

$$d(p,q) := \inf\{\text{length}_g(\gamma) : \gamma \text{ horizontal curve from } p \text{ to } q\}.$$

We say that $p \in M$ is a regular point if, for each i, $\dim V_i(p)$ remains constant near p. Regular points form an open dense set in M.

Our main result in this section reads as follows.

Theorem 7.5. *Let M be a sub-Riemannian manifold and let d be its sub-Riemannian distance. Then d is not σ-finite dimensional.*

Notice that our definition of sub-Riemannian manifolds does not includes Riemannian ones. We recall that it is known that the Riemannian distance on a Riemannian manifold of class C^2 is σ-finite dimensional (see [11, Section 2.8]).

Sub-Riemannian manifolds equipped with their sub-Riemannian distance include sub-Riemannian Carnot groups of step ≥ 2. Besides, as said before, the following fact plays a crucial role for our purposes.

Theorem 7.6 ([3, Theorem 7.36] and [26, Theorem 1]). *A sub-Riemannian manifold equipped with its sub-Riemannian distance admits a metric tangent space in Gromov's sense at every point. At regular points, this space is isometric to a sub-Riemannian Carnot group of step ≥ 2.*

A metric tangent space in Gromov's sense at a point p in a metric space (X, d) is defined as a Hausdorff limit of some sequence of pointed metric spaces $(X, d/\lambda_i, p)$ with $\lambda_i \to 0$ as $i \to +\infty$. We refer to [14] for more details about metric tangent spaces.

Proof of Theorem 7.5. We first prove that one only needs to consider the case where all points in M are regular and moreover, for each i, $\dim V_i(p)$ remains constant in M. Indeed, since regular points form a nonempty open set, one can find a connected open set O in M such that, for each i, $\dim V_i(p)$ remains constant in O. We equip O with the distribution of k-planes which is the restriction to O of the given distribution on M, thus making O a sub-Riemannian manifold. We denote by d_O the associated sub-Riemannian distance on O. By definition we have for all $p, q \in O$,

$$d_O(p,q) = \inf\{\text{length}_g(\gamma) : \gamma \text{ horizontal curve from } p \text{ to } q \text{ in } O\}.$$

In particular, we have $d(p,q) \leq d_O(p,q)$ for all p, $q \in O$, with possibly strict inequality.

However, it can easily be checked that, for all $p \in O$ and all $r < \operatorname{dist}(p, M \setminus O)/3$, one has $B_d(p,r) = B_{d_O}(p,r)$. It follows that if d is σ-finite dimensional on M, then d_O is σ-finite dimensional on O as well. Indeed, assume that M can be written as $M = \bigcup_{n\in\mathbb{N}} M_n$, where d is finite dimensional on each M_n with constants C_n^* and r_n^*. Set

$$O_k := \left\{p \in O : \operatorname{dist}(p, M \setminus O) > \frac{1}{k}\right\}.$$

Then $B_d(p,r) = B_{d_O}(p,r)$ for all $p \in O_k$ and all $r < 1/3k$. It follows that, for each k and n, d_O is finite dimensional on $O_k \cap M_n$ with constants C_n^* and $\min(r_n^*, 1/3k)$. Since $O = \bigcup_{k,n\in\mathbb{N}}(O_k \cap M_n)$, this shows that d_O is σ-finite dimensional on O.

Hence, from now on in this proof, let us assume that the distribution on M is such that, for each i, $\dim V_i(p)$ remains constant in M and let $\dim V_i$ denote its constant value. Such distributions are called generic in [26]. Arguing by contradiction we also assume that the sub-Riemannian distance d on M is σ-finite dimensional, $M = \bigcup_{n\in\mathbb{N}} M_n$, where d is finite dimensional on each M_n.

We let $m := \sum_i i(\dim V_i - \dim V_{i-1})$ denote the Hausdorff dimension of (M,d) (see [26, Theorem 2]) and let μ denote the m-dimensional Hausdorff measure in (M,d). It is proved in [26] that μ is doubling on each compact set. This implies in particular that for any subset $A \subset M$, μ-a.e. point $p \in A$ is a μ-density point for A. On the other hand since $\mu(M) > 0$ and $M = \bigcup_{n\in\mathbb{N}} M_n$, one can find $n \in \mathbb{N}$ such that $\mu(M_n) > 0$, and hence one can find some μ-density point $p \in M_n$ for M_n. Then [18, Proposition 3.1] and Theorem 7.6 imply that (M_n, d) admits a metric tangent space in Gromov's sense at p and this space is isometric to a sub-Riemannian Carnot group (G, d_∞) of step ≥ 2. In particular, by the definition of a metric tangent space, there exist a sequence (λ_i) with $\lambda_i \to 0$ as $i \to +\infty$ and maps $\phi_i : G \to M_n$ such that, for all $p, q \in G$,

$$\frac{1}{\lambda_i} d(\phi_i(p), \phi_i(q)) \xrightarrow[i\to+\infty]{} d_\infty(p,q).$$

Since d is finite dimensional on M_n, Lemma 7.7 below implies that d_∞ is a finite-dimensional sub-Riemannian distance on G. This contradicts Proposition 6.1 and Theorem 1.9 and concludes the proof. □

Lemma 7.7. *Let (X,d) and (X_∞, d_∞) be metric spaces. Assume that there exist a sequence (λ_i) with $\lambda_i \to 0$ as $i \to +\infty$ and maps $\phi_i : X_\infty \to X$ such that for all $x, y \in X_\infty$,*

$$\frac{1}{\lambda_i} d(\phi_i(x), \phi_i(y)) \xrightarrow[i\to+\infty]{} d_\infty(x,y).$$

Assume that d is finite dimensional on X. Then d_∞ is finite dimensional on X_∞.

Proof. Assume that d is finite dimensional on X with constant $C^* \in [1,+\infty)$ and $r^* \in [0,+\infty]$. Let $\{B_{d_\infty}(x_1,r_1), \dots, B_{d_\infty}(x_N,r_N)\}$, $N \in \mathbb{N}$, be a family of Besicovitch balls in (X_∞, d_∞) with $r_l < r^*$ for all $l = 1,\dots,N$. Let $x_0 \in \bigcap_{1\leq l\leq N} B_{d_\infty}(x_l, r_l)$. Let $\varepsilon > 0$ be small enough so that, for all $l,k = 1,\dots,N$ with $l \neq k$, $d_\infty(x_k, x_l) > \max(r_l, r_k) + 2\varepsilon$ and so that $r_l + \varepsilon < r^*$ for all $l = 1,\dots,N$. Let i be large enough so that $\lambda_i \leq 1$ and

$$\left|\frac{1}{\lambda_i} d(\phi_i(x_l), \phi_i(x_k)) - d_\infty(x_l, x_k)\right| < \varepsilon$$

for all $l, k = 0, \ldots, N$. Then for all $l, k = 1, \ldots, N$ with $l \neq k$, we have

$$\frac{1}{\lambda_i} d(\phi_i(x_l), \phi_i(x_k)) > d_\infty(x_l, x_k) - \varepsilon > \max(r_l, r_k) + \varepsilon$$

and, for all $l = 1, \ldots, N$,

$$\frac{1}{\lambda_i} d(\phi_i(x_0), \phi_i(x_l)) < d_\infty(x_0, x_l) + \varepsilon \leq r_l + \varepsilon.$$

Hence $\{B_d(\phi_i(x_1), \lambda_i(r_1 + \varepsilon)), \ldots, B_d(\phi_i(x_N), \lambda_i(r_N + \varepsilon))\}$ is a family of Besicovitch balls in (X, d) with radii $< r^*$. It follows that $N \leq C^*$ which proves that d_∞ is finite dimensional on X_∞. □

Remark 7.8. One can relax the hypothesis of constant rank for the distribution Δ and the proof of Theorem 7.5 can be generalized provided there exists a regular point $p \in M$ such that $\dim \Delta_p < \dim M$.

8. Final remarks

8.1. Assouad's embedding theorem and BCP. Assouad's embedding theorem for snowflakes of doubling metric spaces has the following consequence in connection with the Besicovitch Covering Property.

Proposition 8.1. *Let (X, d) be a doubling metric space. Then there exists a continuous quasi-distance ρ on X that is biLipschitz equivalent to d and such that (X, ρ) satisfies BCP.*

Proof. By Assouad's embedding theorem (see [2]), the snowflaked metric space $(X, d^{1/2})$ admits a biLipschitz embedding $F : (X, d^{1/2}) \to \mathbb{R}^n$ into some Euclidean space. Since Euclidean spaces satisfy BCP, it follows that $F(X)$ equipped with the restriction of the Euclidean distance satisfies BCP (see Proposition 3.15). Set

$$\rho(x, y) := \|F(x) - F(y)\|^2$$

for $x, y \in X$. This defines a continuous quasi-distance on X that satisfies BCP. Finally, if L is the biLipschitz constant of the embedding F, then ρ is L^2-biLipschitz equivalent to the distance d. □

In particular, if (G, d) is a homogeneous group equipped with a homogeneous distance, Proposition 8.1 gives the existence of a continuous quasi-distance biLipschitz equivalent to d and for which BCP holds. In case G is a homogeneous group for which there are two different layers of the grading that do not commute, it follows from Theorem 1.2 that such a quasi-distance cannot be homogeneous, i.e., cannot be both left-invariant and one-homogeneous with respect to the associated dilations. What is not known is whether one can find one such ρ with one of these properties. It is not known either whether one can find a distance, rather than a quasi-distance, biLipschitz equivalent to d and for which BCP holds.

Notice that there are examples of non-doubling metric spaces satisfying BCP. Hence the doubling assumption in Proposition 8.1 is not a necessary condition to get the existence of biLipschitz equivalent (quasi-)distances satisfying BCP.

8.2. Metric spaces of negative type. For the stratified Heisenberg groups $\mathbb{H}^n$, J. Lee and A. Naor proved in [22] that the homogeneous distance d such that

$$d(0,p) = \left(\left(\sum_{j=1}^{n} x_j^2 + y_j^2 \right)^2 + \sqrt{\left(\sum_{j=1}^{n} x_j^2 + y_j^2 \right)^2 + 16z^2} \right)^{1/4}$$

in exponential coordinates of the first kind relative to a standard basis of $\mathfrak{h}_n$ is of negative type (see Example 2.11 for the definition of standard basis of $\mathfrak{h}_n$). Up to a multiplicative constant, this distance turns out to coincide with the Hebisch–Sikora's distance d_R when $R = 2$, namely, $d_2 = 8^{-1/4} d$ (see Example 2.23 for the definition of d_R).

Recall that a metric space (X,d) is said to be of negative type if $(X, \sqrt{d})$ is isometric to a subset of a Hilbert space. One could wonder whether the validity of BCP and the property of being of negative type may have some connections. One can for instance wonder whether Corollary 1.3 could give some hints towards the existence of homogeneous distances of negative type on homogeneous groups with commuting different layers, such as stratified groups of step 2. One can also wonder whether Corollary 1.3 could give some hints towards the nonexistence of homogeneous distances of negative type on a homogeneous group for which there are two different layers of the grading that do not commute, such as stratified groups of step 3 or higher. Unfortunately, it turns out that one can find subsets in a Hilbert space that, when equipped with the restriction of the Hilbert norm, are doubling metric spaces for which BCP does not hold. Hence it is not clear whether one can easily find connections between the validity of BCP and the property of being of negative type.

8.3. Finite topological dimension and an open problem. It is simple to show that if a metric space X satisfies BCP, then X has finite topological dimension. Here, topological dimension means Lebesgue covering dimension. Indeed, let N be the constant in the definition of BCP, see Definition 3.1. Without loss of generality we may assume that X is bounded. Then, given an open cover $\mathcal{U}$ of X, we shall prove that there is a refinement of $\mathcal{U}$ with multiplicity at most N, deducing that the topological dimension is at most $N-1$. For every point a in X take a ball of small enough radius so that it is included in one element of $\mathcal{U}$. By the BCP this family of balls has a subfamily with multiplicity N covering the space X.

If a metric space is assumed to be separable, then by the last argument, together with Remark 3.5, we have that WBCP implies finite topological dimension. With a longer argument, J. Nagata showed in [29] that also WBCP implies finite topological dimension, even if the space is not separable. Moreover, because of Lemma 7.7 a space with WBCP has the property that also the topological dimensions of its tangents are uniformly finite. Doubling spaces have such a property.

To our knowledge, there is no known example of a doubling metric space (X,d) for which there is no distances biLipschitz equivalent to d satisfying BCP.

References

[1] *A. A. Agrachev* and *Y. L. Sachkov*, Control theory from the geometric viewpoint, Encyclopaedia Math. Sci. **87**, Springer, Berlin 2004.

[2] *P. Assouad*, Plongements Lipschitziens dans $\mathbb{R}^n$, Bull. Soc. Math. France **111** (1983), no. 4, 429–448.

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

[4] *A. S. Besicovitch*, A general form of the covering principle and relative differentiation of additive functions, Proc. Cambridge Philos. Soc. **41** (1945), 103–110.

[5] *A. S. Besicovitch*, A general form of the covering principle and relative differentiation of additive functions. II, Proc. Cambridge Philos. Soc. **42** (1946), 1–10.

[6] *N. Bourbaki*, Elements of mathematics. Lie groups and Lie algebras. Chapters 1–3, Springer, Berlin 1998.

[7] *L. J. Corwin* and *F. P. Greenleaf*, Representations of nilpotent Lie groups and their applications. Part I: Basic theory and examples, Cambridge Stud. Adv. Math. **18**, Cambridge University Press, Cambridge 1990.

[8] *C. Drutu* and *M. Kapovich*, Lectures on geometric group theory, manuscript 2011.

[9] *J. L. Dyer*, A nilpotent Lie algebra with nilpotent automorphism group, Bull. Amer. Math. Soc. **76** (1970), 52–56.

[10] *P. Erdős* and *G. Szekeres*, A combinatorial problem in geometry, Compos. Math. **2** (1935), 463–470.

[11] *H. Federer*, Geometric measure theory, Grundlehren Math. Wiss. **153**, Springer, New York 1969.

[12] *G. B. Folland* and *E. M. Stein*, Hardy spaces on homogeneous groups, Math. Notes **28**, Princeton University Press, Princeton 1982.

[13] *R. Goodman*, Filtrations and asymptotic automorphisms on nilpotent Lie groups, J. Differential Geom. **12** (1977), no. 2, 183–196.

[14] *M. Gromov*, Metric structures for Riemannian and non-Riemannian spaces, Progr. Mat. **152**, Birkhäuser, Boston 1999.

[15] *W. Hebisch* and *A. Sikora*, A smooth subadditive homogeneous norm on a homogeneous group, Studia Math. **96** (1990), no. 3, 231–236.

[16] *N. Jacobson*, Lie algebras, Dover Publications, New York 1979.

[17] *A. Korányi* and *H. M. Reimann*, Foundations for the theory of quasiconformal mappings on the Heisenberg group, Adv. Math. **111** (1995), no. 1, 1–87.

[18] *E. Le Donne*, Metric spaces with unique tangents, Ann. Acad. Sci. Fenn. Math. **36** (2011), no. 2, 683–694.

[19] *E. Le Donne*, A primer of Carnot groups: Homogenous groups, CC spaces, and regularity of their isometries, preprint 2016, `https://arxiv.org/abs/1604.08579`.

[20] *E. Le Donne* and *S. Rigot*, Besicovitch covering property for homogeneous distances on the Heisenberg group, preprint 2014, `https://arxiv.org/abs/1406.1484`; to appear in J. Eur. Math. Soc. (JEMS).

[21] *E. Le Donne* and *S. Rigot*, Remarks about the Besicovitch covering property in Carnot groups of step 3 and higher, Proc. Amer. Math. Soc. **144** (2016), no. 5, 2003–2013.

[22] *J. R. Lee* and *A. Naor*, L^p metrics on the Heisenberg group and the Goemans–Linial conjecture, in: 47th Annual IEEE symposium on foundations of computer science (FOCS '06), IEEE Press, Piscataway (2006), 99–108.

[23] *A. Lerario* and *L. Rizzi*, How many geodesics join two points on a contact sub-Riemannian manifold?, preprint 2014, `https://arxiv.org/abs/1405.4294`.

[24] *J. Luukkainen* and *E. Saksman*, Every complete doubling metric space carries a doubling measure, Proc. Amer. Math. Soc. **126** (1998), no. 2, 531–534.

[25] *P. Mattila*, Geometry of sets and measures in Euclidean spaces. Fractals and rectifiability, Cambridge Stud. Adv. Math. **44**, Cambridge University Press, Cambridge 1995.

[26] *J. Mitchell*, On Carnot–Carathéodory metrics, J. Differential Geom. **21** (1985), no. 1, 35–45.

[27] *R. Montgomery*, A tour of subriemannian geometries, their geodesics and applications, Math. Surveys Monogr. **91**, American Mathematical Society, Providence 2002.

[28] *R. Monti*, Some properties of Carnot–Carathéodory balls in the Heisenberg group, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **11** (2000), no. 3, 155–167.

[29] *J. Nagata*, On a special metric and dimension, Fund. Math. **55** (1964), 181–194.

[30] *L. S. Pontryagin*, Topological groups, Gordon and Breach Science Publishers, New York 1966.

[31] *D. Preiss*, Dimension of metrics and differentiation of measures, in: General topology and its relations to modern analysis and algebra V (Prague 1981), Sigma Ser. Pure Math. **3**, Heldermann, Berlin (1983), 565–568.

[32] *C. Reutenauer*, Free Lie algebras, London Math. Soc. Monogr. (N. S.) **7**, Clarendon Press, Oxford 1993.

[33] *S. Rigot*, Counter example to the Besicovitch covering property for some Carnot groups equipped with their Carnot–Carathéodory metric, Math. Z. **248** (2004), no. 4, 827–848.

[34] *E. Sawyer* and *R. L. Wheeden*, Weighted inequalities for fractional integrals on Euclidean and homogeneous spaces, Amer. J. Math. **114** (1992), no. 4, 813–874.

[35] *J.-P. Serre*, Lie algebras and Lie groups, Lecture Notes in Math. **1500**, Springer, Berlin 2006.

[36] *E. Siebert*, Contractive automorphisms on locally compact groups, Math. Z. **191** (1986), no. 1, 73–90.
[37] *F. W. Warner*, Foundations of differentiable manifolds and Lie groups. Reprint, Grad. Texts in Math. **94**, Springer, New York 1983.

Enrico Le Donne, Department of Mathematics and Statistics, University of Jyväskylä,
P.O. Box 35, FI-40014, Finland
https://orcid.org/0000-0002-4415-9916
e-mail: ledonne@msri.org

Séverine Rigot, Université Côte d'Azur, CNRS, LJAD, France
e-mail: rigot@unice.fr

Eingegangen 1. Juli 2016, in revidierter Fassung 16. August 2016