Sharp quantitative nonembeddability of the Heisenberg group into superreflexive Banach spaces

Tim Austin* Brown
Assaf Naor† Courant Institute
Romain Tessera E.N.S. Lyon

Abstract

Let $\mathbb{H}$ denote the discrete Heisenberg group, equipped with a word metric d_W associated to some finite symmetric generating set. We show that if $(X, \|\cdot\|)$ is a p-convex Banach space then for any Lipschitz function $f : \mathbb{H} \to X$ there exist $x, y \in \mathbb{H}$ with $d_W(x, y)$ arbitrarily large and

$$\frac{\|f(x) - f(y)\|}{d_W(x, y)} \lesssim \left(\frac{\log\log d_W(x, y)}{\log d_W(x, y)}\right)^{1/p}. \tag{1}$$

We also show that any embedding into X of a ball of radius $R \geqslant 4$ in $\mathbb{H}$ incurs bi-Lipschitz distortion that grows at least as a constant multiple of

$$\left(\frac{\log R}{\log\log R}\right)^{1/p}. \tag{2}$$

Both (1) and (2) are sharp up to the iterated logarithm terms. When X is Hilbert space we obtain a representation-theoretic proof yielding bounds corresponding to (1) and (2) which are sharp up to a universal constant.

1 Introduction

Let $\mathbb{H} \stackrel{\text{def}}{=} \langle a, b \,|\, aba^{-1}b^{-1} \text{ is central}\rangle$ denote the discrete Heisenberg group, with canonical generators $a, b \in \mathbb{H}$. We let $d_W(\cdot, \cdot)$ denote the left-invariant word metric on $\mathbb{H}$ associated to the symmetric generating set $S \stackrel{\text{def}}{=} \{a, b, a^{-1}, b^{-1}\}$.

A Banach space $(X, \|\cdot\|_X)$ is *superreflexive* if it admits an equivalent uniformly convex norm, i.e., a norm $\|\cdot\|$ satisfying $\alpha\|x\|_X \leqslant \|x\| \leqslant \beta\|x\|_X$ for some $\alpha, \beta > 0$ and all $x \in X$, such that for all $\varepsilon \in (0, 1)$ there exists $\delta > 0$ for which we have

$$\|x\| = \|y\| = 1 \ \wedge \ \|x - y\| = \varepsilon \implies \|x + y\| \leqslant 2 - \delta. \tag{3}$$

Here we prove the following result:

Theorem 1.1. *Let $(X, \|\cdot\|_X)$ be a superreflexive Banach space. Then there exist $c, C > 0$ such that for every $f : \mathbb{H} \to X$ which is 1-Lipschitz with respect to the metric d_W, there are $x, y \in X$ with $d_W(x, y)$ arbitrarily large and*

$$\frac{\|f(x) - f(y)\|_X}{d_W(x, y)} \leqslant \frac{C}{(\log d_W(x, y))^c}.$$

*Partially supported by a fellowship from Microsoft Corporation.
†Supported by NSF grants CCF-0635078 and CCF-0832795, BSF grant 2006009, and the Packard Foundation.

The fact that $\mathbb{H}$ does not admit a bi-Lipschitz embedding into any superreflexive Banach space was proved in [18, 7]. These proofs use an argument of Semmes [25], based on a natural extension of Pansu's differentiability theorem [22].

A natural way to quantify the extent to which $\mathbb{H}$ does not admit a bi-Lipschitz embedding into $(X, \|\cdot\|_X)$ is via Gromov's notion [13, Sec. 7.3] of *compression rate*, defined for a Lipschitz function $f : \mathbb{H} \to X$ as the largest function $\omega_f : (0,\infty) \to [0,\infty)$ such that for all $x, y \in \mathbb{H}$ we have $\|f(x) - f(y)\|_X \geqslant \omega_f\left(d_W(x,y)\right)$. The fact that $\mathbb{H}$ does not admit a bi-Lipschitz embedding into a superreflexive Banach space X means that $\liminf_{t\to\infty} \omega_f(t)/t = 0$ for all Lipschitz functions $f : \mathbb{H} \to X$. The differentiability-based proof of this nonembeddability result involves a limiting argument that does not give information on the rate at which $\omega_f(t)/t$ vanishes. Theorem 1.1 supplies such information, via an approach which is different from the arguments in [18, 7].

Cheeger and Kleiner proved [6] that $\mathbb{H}$ does not admit a bi-Lipschitz embedding into L_1. In [8] it was shown that there exists $c > 0$ such that for any Lipschitz function $f : \mathbb{H} \to L_1$ we have $\omega_f(t)/t \leqslant 1/(\log t)^c$ for arbitrarily large t. This result covers Theorem 1.1 when the superreflexive Banach space X admits a bi-Lipschitz embedding into L_1: such spaces include L_p for $p \in (1,2]$. Theorem 1.1 is new even for spaces such as L_p for $p \in (2,\infty)$, which do not admit a bi-Lipschitz embedding into L_1 (see [3]). Moreover, our method yields sharp results, while the constant c obtained in [8] is far from sharp.

In order to state our sharp version of Theorem 1.1, we recall the following important theorem of Pisier [23, Thm. 3.1]: if X is superreflexive then it admits an equivalent norm $\|\cdot\|$ for which there exist $p \geqslant 2$ and $K > 0$ satisfying the following improvement of (3):

$$\forall x, y \in X, \quad \left\|\frac{x+y}{2}\right\|^p \leqslant \frac{\|x\|^p + \|y\|^p}{2} - \frac{1}{K^p}\left\|\frac{x-y}{2}\right\|^p. \tag{4}$$

A Banach space admitting an equivalent norm satisfying (4) is said to be p-convex. If $(X, \|\cdot\|)$ satisfies (4) then the infimum over those $K > 0$ satisfying (4) is denoted $K_p(X)$. For concreteness, when $p \in (1,2]$ we have $K_2(L_p) \leqslant 1/\sqrt{p-1}$ and for $p \geqslant 2$ we have $K_p(L_p) \leqslant 1$ (see [2]).

The following theorem is a refinement of Theorem 1.1.

Theorem 1.2. *Assume that the Banach space $(X, \|\cdot\|)$ satisfies (4). Let $f : \mathbb{H} \to X$ be a 1-Lipschitz function. Then for every $t \geqslant 3$ there exists an integer $t \leqslant n \leqslant t^2$ such that*[1]

$$\frac{\omega_f(n)}{n} \lesssim K_p(X)\left(\frac{\log\log n}{\log n}\right)^{1/p}. \tag{5}$$

The estimate (5) is sharp up to the iterated logarithm term. Indeed, L_p is p-convex when $p \in [2,\infty)$, and in [26, 27] it was shown that there exists $f : \mathbb{H} \to L_p$ satisfying

$$\frac{\omega_f(n)}{n} \gtrsim \frac{1}{(\log n)^{1/p}\log\log n}$$

for all $n \geqslant 3$ (we refer to [27] for a more refined result of this type).

Our proof of Theorem 1.2 circumvents the difficulties involved with proving quantitative variants of differentiability results by avoiding the need to reason about arbitrary Lipschitz mappings.

[1] In (5), and in the rest of this paper, the notation $\lesssim, \gtrsim$ denotes the corresponding inequalities up to a universal multiplicative factor. The notation $A \asymp B$ stands for $A \lesssim B \wedge B \lesssim A$.

Instead, we start by using a simple result from [21] which reduces the problem to equivariant mappings. Specifically, in [21, Thm. 9.1] it is shown that if X satisfies (4) and $f : \mathbb{H} \to X$ is 1-Lipschitz, then there exists a Banach space Y that also satisfies (4), with $K_p(Y) = K_p(X)$ (in fact, Y is finitely representable in $\ell_p(X)$), an action π of $\mathbb{H}$ on Y by linear isometric automorphisms, and a 1-cocycle $F : \mathbb{H} \to Y$ (i.e., $F(xy) = \pi(x)F(y) + F(x)$ for all $x, y \in \mathbb{H}$) with $\omega_F = \omega_f$. Thus, in proving Theorem 1.2 it suffices to assume that f itself is a 1-cocycle. We note that if X is Hilbert space then Y is also Hilbert space; this is an older result of Gromov (see [10]). More generally, when $X = L_p$ then it is shown in [21] that we can take $Y = L_p$.

Having reduced the problem to 1-cocycles, our starting point is a (non-quantitative) proof, explained in Section 2, showing that if X is an ergodic Banach space, then for every 1-cocycle $f : \mathbb{H} \to X$ we have $\liminf_{t\to\infty} \omega_f(t)/t = 0$. It turns out that the ideas of this proof, which crucially use the fact that f is a 1-cocycle, can be (nontrivially) adapted to yield Theorem 1.2.

Recall that X is ergodic if for every linear isometry $T : X \to X$ and every $x \in X$ the sequence $\left\{\frac{1}{n}\sum_{j=0}^{n-1} T^j x\right\}_{n=1}^{\infty}$ converges in norm. Reflexive spaces, and hence also superreflexive spaces, are ergodic (see [11, p. 662]). If a Banach space X has the property that all Banach spaces that are finitely representable in X are ergodic, then X must be superreflexive [4]. Thus, when using the reduction to 1-cocyles based on the result of [21], the class of Banach spaces to which it naturally applies is the class of superreflexive spaces.

1.1 Bi-Lipschitz distortion of balls

For $R \geqslant 1$ let $B_R = \{x \in \mathbb{H} : \ d_W(e, x) \leqslant R\}$ denote the ball of radius R centered at the identity element $e \in \mathbb{H}$. The bi-Lipschitz distortion of (B_R, d_W) in $(X, \|\cdot\|)$, denoted $c_X(B_R)$, is the infimum over those $D \geqslant 1$ such that there exists $f : B_R \to X$ satisfying

$$\forall\, x, y \in B_R, \quad d_W(x, y) \leqslant \|f(x) - f(y)\| \leqslant D d_W(x, y). \tag{6}$$

Another way to measure the extent to which $\mathbb{H}$ does not admit a bi-Lipschitz embedding into X is via the rate at which $c_X(B_R)$ grows to ∞ with R. In [8] it was shown that

$$c_{L_1}(B_R) \gtrsim (\log R)^c \tag{7}$$

for some universal constant $c > 0$. This result is of importance due to an application to theoretical computer science; see [9, 8, 20] for a detailed discussion. Evaluating the supremum over those $c > 0$ satisfying (7) remains an important open problem. Theorem 1.2 implies the following sharp bound on the bi-Lipschitz distortion of B_R into a p-convex Banach space:

Theorem 1.3. *If a Banach space $(X, \|\cdot\|)$ satisfies (4) then for every $R \geqslant 4$ we have*

$$c_X(B_R) \gtrsim \frac{1}{K_p(X)} \left(\frac{\log R}{\log\log R}\right)^{1/p}.$$

Thus in particular for $p \in (1, 2]$ we have $c_{L_p}(B_R) \gtrsim \sqrt{p-1} \cdot (\log R)^{\frac{1}{2}-o(1)}$ and for $p \geqslant 2$ we have $c_{L_p}(B_R) \gtrsim (\log R)^{\frac{1}{p}-o(1)}$. Theorem 1.3 is a formal consequence of Theorem 1.2. The simple deduction of Theorem 1.3 from Theorem 1.2 is presented in Section 6. It follows from the results of [1, 24] (see the explanation in [15, 27]) that for every $p \geqslant 2$ we have $c_{L_p}(B_R) \lesssim (\log R)^{1/p}$. Thus Theorem 1.3 is sharp up to iterated logarithms.

1.2 The case of Hilbert space

In section 7 we prove the following Poincaré-type inequality for functions on $\mathbb{H}$ taking value in Hilbert space:

Theorem 1.4. *For every $f : \mathbb{H} \to L_2$ and every $R \in \mathbb{N}$ we have*

$$\sum_{x \in B_R} \sum_{k=1}^{R^2} \frac{\left\| f(xc^k) - f(x) \right\|_2^2}{k^2} \lesssim \sum_{x \in B_{22R}} \left(\|f(xa) - f(x)\|_2^2 + \|f(xb) - f(x)\|_2^2 \right). \tag{8}$$

This result has the following two sharp consequences. First, assume that $\theta : (0, \infty) \to [0, \infty)$ is nondecreasing, and that $\theta \leqslant \omega_f$ for some 1-Lipschitz $f : \mathbb{H} \to L_2$. Then since $|B_{22R}| \asymp |B_R|$ and $d_W(c^k, e) \asymp \sqrt{k}$ for all $k \in \mathbb{N}$, inequality (8) implies that

$$\frac{1}{2} \int_1^\infty \frac{\theta(t)^2}{t^3} dt = \int_1^\infty \frac{\theta(\sqrt{s})^2}{s^2} ds \lesssim \sum_{k=1}^\infty \frac{\theta\left(\sqrt{k}\right)^2}{k^2} \lesssim 1. \tag{9}$$

Combined with [26, Theorem 1], we obtain

Corollary 1.5. *A nondecreasing function $\theta : (0, \infty) \to [0, \infty)$ satisfies $\theta \leqslant \omega_f$ for some Lipschitz function $f : \mathbb{H} \to L_2$ if and only if*

$$\int_1^\infty \left(\frac{\theta(t)}{t} \right)^2 \frac{dt}{t} < \infty. \tag{10}$$

A second corollary of Theorem 1.4 yields a sharp bound (up to universal constants) on $c_{L_2}(B_R)$. Indeed, fix $R \geqslant 2$ and assume that $f : B_R \to L_2$ satisfies $d_W(x, y) \leqslant \|f(x) - f(y)\| \leqslant D d_W(x, y)$ for all $x, y \in B_R$. Let $f^* : \mathbb{H} \to L_2$ have Lipschitz constant at most $2D$ and coincide with f on $B_{R/2}$ (see equation (41) for an explicit formula defining such an extension f^*). It follows from (8) applied to f^* that $D^2 \gtrsim \sum_{k=1}^R \frac{1}{k} \gtrsim \log R$. Thus $c_{L_2}(B_R) \gtrsim \sqrt{\log R}$. In conjunction with the previously quoted upper bound on $c_{L_2}(B_R)$, we have

Corollary 1.6. *For every $R \geqslant 2$ we have $c_{L_2}(B_R) \asymp \sqrt{\log R}$.*

Roughly speaking, the proof of Theorem 1.4 proceeds via a reduction to the case of 1-cocycles corresponding to the irreducible representations of $\mathbb{H}$ (see Section 7). But actually, since the representation theory of the continuous Heisenberg group is simpler than the representation theory of the discrete Heisenberg group $\mathbb{H}$, we first apply a discretization argument which reduces Theorem 1.4 to an inequality on the real Heisenberg group. Then an averaging argument reduces the proof to an inequality on cocycles. Every unitary representation of the continuous Heisenberg group decomposes as a direct integral of irreducibles, and cocycles themselves can be decomposed accordingly. Since the desired inequality involves a sum of squares of norms, it suffices to prove it for cocycles corresponding to irreducible representations (that is, for each direct integrand separately). The computation for irreducible representations is carried out in Section 7.1.

2 Sublinear growth of Heisenberg cocycles in ergodic spaces

Write $c \stackrel{\text{def}}{=} [a,b] = aba^{-1}b^{-1}$. Thus c lies in the center of $\mathbb{H}$ and for every $n \in \mathbb{N}$ we have $d_W\left(c^{n^2}, e_{\mathbb{H}}\right) = 4n$ (in fact $c^{n^2} = [a^n, b^n] = a^n b^n a^{-n} b^{-n}$).

Let $(X, \|\cdot\|)$ be a Banach space and $\pi : G \to \mathrm{Aut}(X)$ be an action of $\mathbb{H}$ on X by linear isometric automorphisms. In addition let $f \in Z^1(\pi)$ be a 1-cocycle, so $f : \mathbb{H} \to X$ and for all $x, y \in \mathbb{H}$ we have $f(xy) = \pi(x)f(y) + f(x)$. We assume in what follows that f is 1-Lipschitz, or equivalently that $\max\{\|f(a)\|, \|f(b)\|\} = 1$.

In this section we quickly show that if X is an ergodic Banach space then $\liminf_{t\to\infty} \omega_f(t)/t = 0$, but without obtaining any quantitative bounds.

If X is ergodic then the operator on X defined by

$$Px \stackrel{\text{def}}{=} \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \pi(c)^n x$$

is a contraction onto the subspace $X_0 \subseteq X$ of $\pi(c)$-invariant vectors, and since $Px = x$ for any $x \in X_0$ it follows at once that P is idempotent. Also, since c is central in $\mathbb{H}$, the projection P commutes with $\pi(g)$ for all $g \in \mathbb{H}$, and hence the maps $g \mapsto Pf(g)$ and $g \mapsto (I-P)f(g)$ are both still members of $Z^1(\pi)$. Since P and $I - P$ are bounded, these cocycles are both still Lipschitz functions from $\mathbb{H}$ to X.

We complete the proof by showing that for any $\varepsilon > 0$ we have

$$\left\| f\left(c^{N^2}\right) \right\| \leqslant 2\varepsilon d_W\left(e_{\mathbb{H}}, c^{N^2}\right) \leqslant 8\varepsilon N \tag{11}$$

for all sufficiently large N. To prove this we consider the two cocycles Pf and $(I-P)f$ separately. On the one hand, Pf takes values among the $\pi(c)$-invariant vectors, and hence the cocycle identity implies that

$$Pf\left(c^N\right) = \sum_{n=0}^{N-1} \pi(c)^n Pf(c) = NPf(c),$$

and therefore $\left\|Pf\left(c^N\right)\right\| = N\|Pf(c)\|$. However, $\left\|Pf\left(c^N\right)\right\| \leqslant d_W\left(e_{\mathbb{H}}, c^N\right) \lesssim \sqrt{N}$, so these relations are compatible only if $Pf(c) = 0$.

On the other hand, let $\widetilde{f} \stackrel{\text{def}}{=} (I-P)f$ and for each $K \geqslant 1$,

$$v_K \stackrel{\text{def}}{=} -\frac{1}{K} \sum_{k=1}^{K} f\left(c^k\right).$$

Observe from the cocycle identity and the centrality of c that

$$\begin{aligned} -\pi(g)v_K + \widetilde{f}(g) = \frac{1}{K} \sum_{k=1}^{K} \left(\pi(g)\widetilde{f}\left(c^k\right) + \widetilde{f}(g)\right) &= \frac{1}{K} \sum_{k=1}^{K} \widetilde{f}\left(gc^k\right) \\ &= \frac{1}{K} \sum_{k=1}^{K} \widetilde{f}\left(c^k g\right) = \frac{1}{K} \sum_{k=1}^{K} \pi\left(c^k\right) \widetilde{f}(g) - v_K. \end{aligned} \tag{12}$$

Re-arranging (12) gives

$$\widetilde{f}(g) = \pi(g)v_K - v_K + \frac{1}{K}\sum_{k=1}^{K} \pi(c)^k \widetilde{f}(g). \tag{13}$$

For any fixed $g \in \mathbb{H}$ the last term of this right-hand side of (13) converges to $P(I-P)f(g) = 0$ (using again that X is ergodic), and so in particular once K is sufficiently large we obtain that for all $g \in \mathbb{H}$ we have,

$$\max\left\{\left\|\widetilde{f}\left(a^{\pm 1}\right) - \left(\pi\left(a^{\pm 1}\right)v_K - v_K\right)\right\|, \ \left\|\widetilde{f}\left(b^{\pm 1}\right) - \left(\pi\left(b^{\pm 1}\right)v_K - v_K\right)\right\|\right\} \leqslant \varepsilon.$$

Having obtained this approximation to $\widetilde{f}$ by a coboundary, let $c^{N^2} = s_1 s_2 \cdots s_{4N}$ be an expression for c^{N^2} as a word in S, and observe from another appeal to the cocycle identity that

$$\widetilde{f}\left(c^{N^2}\right) = \sum_{i=0}^{4N-1} \pi\left(s_1 s_2 \cdots s_i\right)\widetilde{f}\left(s_{i+1}\right) = \sum_{i=0}^{4N-1} \pi\left(s_1 s_2 \cdots s_i\right)\left(\pi\left(s_{i+1}\right)v_K - v_K\right) + R_N$$
$$= \pi(s_1 s_2 \cdots s_{4N})v_K - v_K + R_N$$

for some remainder R_N which is a sum of $4N$ terms all of norm at most ε. Since the action π is isometric and we may let N grow independently of K we obtain

$$\left\|\widetilde{f}\left(c^{N^2}\right)\right\| \leqslant 2\|v_K\| + \|R_N\| \leqslant 8\varepsilon N$$

for all sufficiently large N. Since ε was arbitrary and $\widetilde{f}\left(c^{N^2}\right) = f\left(c^{N^2}\right)$ by our analysis of Pf above, this completes the proof of (11).

3 A uniform convexity lemma for ergodic averages

We prove here a simple lemma on the behavior of ergodic averages in p-convex Banach spaces.

Lemma 3.1. *Assume that $(X, \|\cdot\|)$ satisfies (4). Fix $z \in X$ and an operator $T : X \to X$ with $\|T\| \leqslant 1$. For every integer $n \geqslant 0$ denote*

$$s_n \stackrel{\text{def}}{=} \frac{1}{2^n}\sum_{j=0}^{2^n-1} T^j z.$$

Then for every $\ell \in \mathbb{N}$ we have:

$$\sum_{i=0}^{\infty} \frac{1}{2^\ell}\sum_{j=0}^{2^\ell-1} \left\|s_{(i+1)\ell} - T^{j2^{i\ell}} s_{i\ell}\right\|^p \leqslant (2K)^p \|z\|^p. \tag{14}$$

Proof. A consequence of (4) is that for every $x_1, \ldots, x_n \in X$ we have:

$$\frac{1}{n}\sum_{i=1}^{n}\left\|x_i - \frac{1}{n}\sum_{j=1}^{n} x_j\right\|^p \leqslant (2K)^p\left(\frac{1}{n}\sum_{i=1}^{n}\|x_i\|^p - \left\|\frac{1}{n}\sum_{i=1}^{n} x_i\right\|^p\right). \tag{15}$$

For the derivation of (15) from (4) see [19, Lem. 3.1].

Due to the identity

$$s_{(i+1)\ell} = \frac{1}{2^\ell} \sum_{j=0}^{2^\ell - 1} T^{j2^{i\ell}} s_{i\ell},$$

inequality (15) implies that:

$$\frac{1}{2^\ell} \sum_{j=0}^{2^\ell-1} \left\| T^{j2^{i\ell}} s_{i\ell} - s_{(i+1)\ell} \right\|^p \leqslant (2K)^p \left(\frac{1}{2^\ell} \sum_{j=0}^{2^\ell-1} \left\| T^{j2^{i\ell}} s_{i\ell} \right\|^p - \left\| s_{(i+1)\ell} \right\|^p \right)$$
$$\leqslant (2K)^p \left(\|s_{i\ell}\|^p - \left\| s_{(i+1)\ell} \right\|^p \right). \quad (16)$$

The desired inequality (14) now follows by summing (16) over $i \in \{0, 1, \ldots\}$. □

4 Estimates for Heisenberg cocycles

Let $\pi : \mathbb{H} \to \mathrm{Aut}(X)$ and $f \in Z^1(\pi)$ be as in Section 2. For every $n \in \mathbb{N}$ define a linear operator $P_n : X \to X$ by

$$P_n \stackrel{\mathrm{def}}{=} \frac{1}{2^n} \sum_{j=0}^{2^n-1} \pi(c)^j.$$

Thus $\|P_n\| \leqslant 1$.

Lemma 4.1. *Assume that $(X, \|\cdot\|)$ satisfies (4). Then for every $\ell, k, m \in \mathbb{N}$ there exist integers $i \in [k+1, k+m]$ and $j \in [0, 2^\ell - 1]$ satisfying for all $n \in \mathbb{N}$,*

$$\left\| \pi\left(c^{-j2^{i\ell}}\right) P_{(i+1)\ell} f\left(c^{n^2}\right) - P_{i\ell} f\left(c^{n^2}\right) \right\| \leqslant \frac{16Kn}{m^{1/p}}. \quad (17)$$

Proof. Consider the Banach space $Y = X \oplus X$, equipped with the norm

$$\|(x, y)\|_Y = \left(\|x\|^p + \|y\|^p\right)^{1/p}.$$

We also define $T : Y \to Y$ by $T(x, y) = (\pi(c)x, \pi(c)y)$. Then $\|T\| \leqslant 1$. Since $(Y, \|\cdot\|_Y)$ satisfies (4) we may apply Lemma 3.1 to $z = (f(a), f(b)) \in Y$, obtaining the following estimate:

$$(4K)^p \geqslant \sum_{i=k+1}^{k+m} \frac{1}{2^\ell} \sum_{j=0}^{2^\ell-1} \left(\left\| P_{(i+1)\ell} f(a) - \pi\left(c^{j2^{i\ell}}\right) P_{i\ell} f(a) \right\|^p + \left\| P_{(i+1)\ell} f(b) - \pi\left(c^{j2^{i\ell}}\right) P_{i\ell} f(b) \right\|^p \right)$$
$$= \sum_{i=k+1}^{k+m} \frac{1}{2^\ell} \sum_{j=0}^{2^\ell-1} \left(\left\| \pi\left(c^{-j2^{i\ell}}\right) P_{(i+1)\ell} f(a) - P_{i\ell} f(a) \right\|^p + \left\| \pi\left(c^{-j2^{i\ell}}\right) P_{(i+1)\ell} f(b) - P_{i\ell} f(b) \right\|^p \right)$$
$$\geqslant m \min_{\substack{k+1 \leqslant i \leqslant k+m \\ 0 \leqslant j \leqslant 2^\ell - 1}} \left(\left\| \pi\left(c^{-j2^{i\ell}}\right) P_{(i+1)\ell} f(a) - P_{i\ell} f(a) \right\|^p + \left\| \pi\left(c^{-j2^{i\ell}}\right) P_{(i+1)\ell} f(b) - P_{i\ell} f(b) \right\|^p \right).$$

It follows that there exist integers $i \in [k+1, k+m]$, $j \in [0, 2^\ell - 1]$ such that

$$\max\left\{ \left\| \pi\left(c^{-j2^{i\ell}}\right) P_{(i+1)\ell} f(a) - P_{i\ell} f(a) \right\|, \left\| \pi\left(c^{-j2^{i\ell}}\right) P_{(i+1)\ell} f(b) - P_{i\ell} f(b) \right\| \right\} \leqslant \frac{4K}{m^{1/p}}. \quad (18)$$

Consider the operator

$$Q_n \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=0}^{n-1}\pi(a)^i.$$

The cocycle identity implies that $f(a^n) = nQ_nf(a)$. Thus

$$\pi\left(c^{-j2^{i\ell}}\right)P_{(i+1)\ell}f(a^n) - P_{i\ell}f(a^n) = n\left(\pi\left(c^{-j2^{i\ell}}\right)P_{(i+1)\ell} - P_{i\ell}\right)Q_nf(a)$$
$$= nQ_n\left(\pi\left(c^{-j2^{i\ell}}\right)P_{(i+1)\ell} - P_{i\ell}\right)f(a), \quad (19)$$

where the last equality in (19) holds since c is in the center of $\mathbb{H}$, and therefore Q_n commutes with all of $\{P_r\}_{r=0}^{\infty}$. Since $\|Q_n\| \leqslant 1$, it follows from (19) and (18) that

$$\left\|\pi\left(c^{-j2^{i\ell}}\right)P_{(i+1)\ell}f(a^n) - P_{i\ell}f(a^n)\right\| \leqslant \frac{4Kn}{m^{1/p}}. \quad (20)$$

Since $f(a^{-n}) = -\pi(a)^{-n}f(a^n)$, and $\pi(a)$ commutes with $\pi(c)$, and hence with all of $\{P_r\}_{r=0}^{\infty}$, it follows that also

$$\left\|\pi\left(c^{-j2^{i\ell}}\right)P_{(i+1)\ell}f(a^{-n}) - P_{i\ell}f(a^{-n})\right\| \leqslant \frac{4Kn}{m^{1/p}}. \quad (21)$$

An identical argument implies the analogous bounds with a replaced by b:

$$\left\|\pi\left(c^{-j2^{i\ell}}\right)P_{(i+1)\ell}f(b^n) - P_{i\ell}f(b^n)\right\| \leqslant \frac{4Kn}{m^{1/p}} \quad (22)$$

and

$$\left\|\pi\left(c^{-j2^{i\ell}}\right)P_{(i+1)\ell}f(b^{-n}) - P_{i\ell}f(b^{-n})\right\| \leqslant \frac{4Kn}{m^{1/p}}. \quad (23)$$

The cocycle identity implies that for all $n \in \mathbb{N}$,

$$f\left(c^{n^2}\right) = f\left([a^n, b^n]\right) = \pi(a^nb^na^{-n})f(b^{-n}) + \pi(a^nb^n)f(a^{-n}) + \pi(a^n)f(b^n) + f(a^n).$$

Thus, using (20), (21), (22) and (23), we conclude the validity of (17). □

Lemma 4.2. *For every $m, n \in \mathbb{N}$ we have*

$$\left\|P_mf\left(c^{n^2}\right)\right\| \lesssim \frac{n^{5/3}}{2^{m/3}}. \quad (24)$$

Proof. Note that for every $k \in \mathbb{N}$ we have

$$P_m - \pi\left(c^k\right)P_m = \frac{1}{2^m}\sum_{j=0}^{k-1}\pi(c)^j - \frac{1}{2^m}\sum_{j=2^m}^{2^m+k-1}\pi(c)^j.$$

Thus,

$$\left\|P_m - \pi\left(c^k\right)P_m\right\| \leqslant \frac{2k}{2^m}. \quad (25)$$

The cocycle identity implies that

$$f\left(c^{(kn)^2}\right) = \sum_{j=0}^{k^2-1}\pi\left(c^{jn^2}\right)f\left(c^{n^2}\right). \quad (26)$$

Using the fact that f is 1-Lipschitz, $\|P_m\| \leqslant 1$ and $d_W\left(e_{\mathbb{H}}, c^{(kn)^2}\right) \leqslant 4kn$, we deduce from (26) that

$$4kn \geqslant \left\|P_m f\left(c^{(kn)^2}\right)\right\| \geqslant \sum_{j=0}^{k^2-1}\left(\left\|P_m f\left(c^{n^2}\right)\right\| - \left\|P_m - \pi\left(c^{jn^2}\right)P_m\right\| \cdot \left\|f\left(c^{n^2}\right)\right\|\right)$$

$$\overset{(25)}{\geqslant} k^2\left\|P_m f\left(c^{n^2}\right)\right\| - \sum_{j=0}^{k^2-1}\frac{2jn^2}{2^m}\cdot 4n \geqslant k^2\left\|P_m f\left(c^{n^2}\right)\right\| - \frac{4n^3k^4}{2^m}.$$

Thus,

$$\left\|P_m f\left(c^{n^2}\right)\right\| \leqslant \frac{4n}{k} + \frac{4n^3k^2}{2^m}. \tag{27}$$

Choosing $k = \left\lceil (2^{m-1}/n^2)^{1/3}\right\rceil$ in (27) (roughly the optimal choice of k), we obtain (24). □

Lemma 4.3. *For every $m, n \in \mathbb{N}$ we have:*

$$\left\|f\left(c^{n^2}\right) - P_m f\left(c^{n^2}\right)\right\| \leqslant 2^{m/3}n^{1/3}. \tag{28}$$

Proof. In this proof the relation to Section 2 becomes clear. Define $\widetilde{f} : \mathbb{H} \to X$ by

$$\widetilde{f}(h) \overset{\text{def}}{=} f(h) - P_m f(h) = (I - P_m)f(h).$$

Note that $\widetilde{f} \in Z^1(\pi)$. Fix an integer $k \geqslant 1$ that will be determined later. Consider the vector $v \in X$ defined by

$$v \overset{\text{def}}{=} -\frac{1}{k}\sum_{j=0}^{k-1}\widetilde{f}\left(c^j\right).$$

Then

$$\|v\| \lesssim \frac{1}{k}\sum_{j=0}^{k-1}\sqrt{j} \lesssim \sqrt{k}. \tag{29}$$

Since c is in the center of $\mathbb{H}$, we have the following identity for every $h \in \mathbb{H}$:

$$-\pi(h)v + \widetilde{f}(h) = \frac{1}{k}\sum_{j=0}^{k-1}\left(\pi(h)\widetilde{f}\left(c^j\right) + \widetilde{f}(h)\right) = \frac{1}{k}\sum_{j=0}^{k-1}\widetilde{f}\left(hc^j\right) = \frac{1}{k}\sum_{j=0}^{k-1}\widetilde{f}\left(c^j h\right)$$

$$= \frac{1}{k}\sum_{j=0}^{k-1}\left(\pi\left(c^j\right)\widetilde{f}(h) + \widetilde{f}\left(c^j\right)\right) = \frac{1}{k}\sum_{j=0}^{k-1}\pi\left(c^j\right)\widetilde{f}(h) - v. \tag{30}$$

Note that

$$\frac{1}{k}\sum_{j=0}^{k-1}\pi\left(c^j\right)\widetilde{f}(h) = \frac{1}{k}\sum_{j=0}^{k-1}\left(\pi\left(c^j\right)f(h) - \frac{1}{2^m}\sum_{i=0}^{2^m-1}\pi\left(c^{j+i}\right)f(h)\right)$$

$$= \frac{1}{2^m}\sum_{i=0}^{2^m-1}\left(\frac{1}{k}\sum_{j=0}^{k-1}\pi\left(c^j\right) - \frac{1}{k}\sum_{j=i}^{i+k-1}\pi\left(c^j\right)\right)f(h).$$

Hence,

$$\left\|\frac{1}{k}\sum_{j=0}^{k-1}\pi\left(c^j\right)\widetilde{f}\,(h)\right\| \leqslant \frac{d_W(h,e_{\mathbb{H}})}{2^m}\sum_{i=0}^{2^m-1}\frac{2i}{k} \leqslant \frac{2^m}{k}d_W(h,e_{\mathbb{H}}). \tag{31}$$

Combining (30) and (31), we see that $\widetilde{f}$ is close to a coboundary in the following sense:

$$\left\|\widetilde{f}(h)-(\pi(h)v-v)\right\| \leqslant \frac{2^m}{k}d_W(h,e_{\mathbb{H}}). \tag{32}$$

If we now write $c^{n^2}=h_1h_2\cdots h_{4n}$ for $h_1,\ldots,h_{4n}\in\left\{a,a^{-1},b,b^{-1}\right\}$, then the cocycle identity for $\widetilde{f}$ implies the following bound:

$$\begin{aligned}
\left\|\widetilde{f}\left(c^{n^2}\right)\right\| &= \left\|\sum_{i=0}^{4n-1}\pi(h_1\cdots h_{4n-i-1})\widetilde{f}(h_{4n-i})\right\| \\
&\overset{(32)}{\leqslant} \left\|\sum_{i=0}^{4n-1}\pi(h_1\cdots h_{4n-i-1})(\pi(h_{4n-i})v-v)\right\|+\frac{4n2^m}{k} \\
&= \left\|\pi\left(c^{n^2}\right)v-v\right\|+\frac{4n2^m}{k} \\
&\overset{(29)}{\lesssim} \sqrt{k}+\frac{n2^m}{k}.
\end{aligned} \tag{33}$$

The optimal choice for k in (33) is $k\asymp n^{2/3}2^{2m/3}$. For this choice of k, (33) becomes the desired bound (28). □

5 Proof of Theorem 1.2

As explained in the introduction, using [21, Thm. 9.1] we may assume without loss of generality that $f\in Z^1(\pi)$ for some action π of $\mathbb{H}$ on X by linear isometric automorphisms. We may also assume that $t\geqslant 8^p$. Let m be the largest integer such that

$$m^m \leqslant \left(\frac{t}{4}\right)^{p/3}. \tag{34}$$

Having defined m, let k be the smallest integer such that

$$m^{\frac{3}{2p}+\frac{3(k+1)}{p}} \geqslant t, \tag{35}$$

and set

$$\ell \overset{\text{def}}{=} \left\lceil \frac{6}{p}\log_2 m\right\rceil. \tag{36}$$

By Lemma 4.1 there exist integers $i\in[k+1,k+m]$ and $j\in[0,2^\ell-1]$ satisfying for all $n\in\mathbb{N}$,

$$\left\|\pi\left(c^{-j2^{i\ell}}\right)P_{(i+1)\ell}f\left(c^{n^2}\right)-P_{i\ell}f\left(c^{n^2}\right)\right\| \leqslant \frac{16Kn}{m^{1/p}}. \tag{37}$$

Choose

$$n \overset{\text{def}}{=} \frac{1}{4}\left\lceil m^{\frac{3}{2p}}2^{\frac{i\ell}{2}}\right\rceil, \tag{38}$$

We may write

$$f\left(c^{n^2}\right) = \pi\left(c^{-j2^{i\ell}}\right) P_{(i+1)\ell} f\left(c^{n^2}\right) + \left(P_{i\ell} f\left(c^{n^2}\right) - \pi\left(c^{-j2^{i\ell}}\right) P_{(i+1)\ell} f\left(c^{n^2}\right)\right) + \left(f\left(c^{n^2}\right) - P_{i\ell} f\left(c^{n^2}\right)\right).$$

Hence, by Lemma 4.2, inequality (37), and Lemma 4.3, we obtain the following bound:

$$\omega_f(4n) = \omega_f\left(d_W\left(c^{n^2}, e_{\mathbb{H}}\right)\right) \leqslant \left\|f\left(c^{n^2}\right)\right\| \lesssim \frac{n^{5/3}}{2^{(i+1)\ell/3}} + \frac{8Kn}{m^{1/p}} + 2^{i\ell/3} n^{1/3} \overset{(38)\wedge(36)}{\lesssim} \frac{Kn}{m^{1/p}}. \tag{39}$$

Observe that

$$4n \overset{(38)}{\geqslant} m^{\frac{3}{2p}} 2^{\frac{i\ell}{2}} \geqslant m^{\frac{3}{2p}} 2^{\frac{(k+1)\ell}{2}} \overset{(36)}{\geqslant} m^{\frac{3}{2p}+\frac{3(k+1)}{p}} \overset{(35)}{\geqslant} t.$$

At the same time,

$$4n \overset{(38)}{\leqslant} 2m^{\frac{3}{2p}} 2^{\frac{i\ell}{2}} \leqslant 2m^{\frac{3}{2p}} 2^{\frac{(k+m)\ell}{2}} \overset{(36)}{\leqslant} 4m^{\frac{3}{2p}+\frac{3k}{p}} \cdot m^{\frac{3m}{p}} \overset{(35)}{<} 4tm^{\frac{3m}{p}} \overset{(34)}{\leqslant} t^2.$$

Hence $t \leqslant 4n \leqslant t^2$. The definition of m implies that $m \gtrsim \frac{p}{3}\frac{\log n}{\log\log n}$, and therefore (39) becomes:

$$\frac{\omega_f(4n)}{n} \lesssim K\left(\frac{\log\log n}{\log n}\right)^{1/p}.$$

The proof of Theorem 1.2 is complete. □

6 Deduction of Theorem 1.3 from Theorem 1.2

Fix $R \geqslant 4$ and a function $f : B_R \to X$ satisfying

$$\forall\, x, y \in B_R, \quad d_W(x,y) \leqslant \|f(x) - f(y)\| \leqslant D d_W(x,y). \tag{40}$$

Our goal is to bound D from below. Without loss of generality assume that $f(e) = 0$. Define $f^* : \mathbb{H} \to X$ by

$$f^*(x) = \begin{cases} f(x) & x \in B_{R/2}, \\ 2\left(1 - \frac{d_W(x,e)}{R}\right) f(x) & x \in B_R \setminus B_{R/2}, \\ 0 & x \in \mathbb{H} \setminus B_R. \end{cases} \tag{41}$$

Then f^* is $2D$-Lipschitz and coincides with f on $B_{R/2}$. Let $\mathcal{N} \subseteq \mathbb{H}$ be a maximal $3R$-separated subset of $\mathbb{H}$. Thus the function $f^{**} : \mathbb{H} \to X$ given by $f^{**}(x) \overset{\text{def}}{=} \sum_{y\in\mathcal{N}} f^*(y^{-1}x)$ (only one summand is nonzero for any given x) is also $2D$-Lipschitz.

Fix a free ultrafilter $\mathscr{U}$ on $\mathbb{N}$. Consider the semi-normed space $Y = (\ell_\infty(\mathbb{H}, X), \|\cdot\|_Y)$, where

$$\|\psi\|_Y \overset{\text{def}}{=} \lim_{M\to\mathscr{U}} \left(\frac{1}{|B_M|}\sum_{z\in B_M} \|\psi(z)\|^p\right)^{1/p}.$$

Note that since X satisfies (4), so does Y. Y is a semi-normed space rather than a normed space, so we should formally deal below with the quotient $Y/\{f \in \ell_\infty(\mathbb{H}, X) : \|f\|_Y = 0\}$, but we will

ignore this inessential formality in what follows. (Complete details are as in the proof of Theorem 9.1 in [21]. Alternatively one can note that our proof of Theorem 1.2 carries over without change to the class of semi-normed spaces.)

Define $F : \mathbb{H} \to Y$ by $F(x)(z) \stackrel{\text{def}}{=} f^{**}(zx) - f^{**}(z)$. This is well defined since the metric d_W is left-invariant, and therefore $\|F(x)\|_Y \leqslant 2Dd_W(x, e_{\mathbb{H}})$ for all $x \in \mathbb{H}$. Moreover, by left-invariance, F is $2D$-Lipschitz. Theorem 1.2 therefore implies that there exist $x, y \in \mathbb{H}$ such that $\sqrt{R/4} \leqslant d_W(x, y) \leqslant R/4$ and

$$\lim_{M \to \mathscr{U}} \left(\frac{1}{|B_M|} \sum_{z \in B_M} \left(\frac{\|f^{**}(zx) - f^{**}(zy)\|}{d_W(x, y)} \right)^p \right)^{1/p} \lesssim DK_p(X) \left(\frac{\log\log R}{\log R} \right)^{1/p}. \tag{42}$$

Fix an integer $M > 2d_W(x, e) + 4R$ and write $m = M - 2d_W(x, e) - 4R$. Since $\mathcal{N}$ is a maximal $3R$-separated subset of $\mathbb{H}$, we have $B_m x \subseteq \bigcup_{w \in \mathcal{M}} wB_{3R}$, where $\mathcal{M} \stackrel{\text{def}}{=} \{w \in \mathcal{N} : \ wB_{3R} \cap B_m x \neq \emptyset\}$. Hence, since for $r \geqslant 1$ we have $|B_r| \asymp r^4$, we can bound the cardinality of $\mathcal{M}$ as follows:

$$|\mathcal{M}| \gtrsim \left(\frac{M - 2d_W(x, e) - 4R}{3R} \right)^4. \tag{43}$$

If $w \in \mathcal{M}$ then there exists $z \in B_m$ and $g \in B_{3R}$ such that $zx = wg$. Hence, for every $h \in B_{R/4}$ we have

$$d_W(whx^{-1}, e) = d_W(zxg^{-1}hx^{-1}, e) \leqslant d_W(z, e) + 2d_W(x, e) + d_W(g, e) + d_W(h, e) < M.$$

Thus the sets $\{wB_{R/4}x^{-1}\}_{w \in \mathcal{M}}$ are disjoint and contained in B_M. Moreover, if $w \in \mathcal{M}$ and $z \in wB_{R/4}x^{-1}$ then $d_W(zx, w) \leqslant R/4$, and hence also $d_W(zy, w) \leqslant d_W(zy, zx) + d_W(zx, w) \leqslant R/2$. By the definition of f^{**}, this implies that $f^{**}(zx) = f(w^{-1}zx)$ and $f^{**}(zy) = f(w^{-1}zy)$. Hence,

$$\sum_{z \in B_M} \left(\frac{\|f^{**}(zx) - f^{**}(zy)\|}{d_W(x, y)} \right)^p \geqslant \sum_{w \in \mathcal{M}} \sum_{z \in wB_{R/4}x^{-1}} \left(\frac{\left\|f(w^{-1}zx) - (w^{-1}zy)\right\|}{d_W(x, y)} \right)^p$$

$$\stackrel{(40)}{\geqslant} |\mathcal{M}| \cdot |B_{R/4}| \stackrel{(43)}{\gtrsim} \left(1 - \frac{2d_W(x, e) + 4R}{M}\right)^4 |B_M|. \tag{44}$$

Theorem 1.3 now follows from (42) and (44). □

7 Embeddings into Hilbert space

In this section, we prove Theorem 1.4. We will deduce it from an inequality on cocycles for the real Heisenberg group. We switch to the real Heisenberg group because its representation theory is simpler. However, this comes at the cost of adding a (straightforward) discretization step to the proof. The upshot is that we obtain as a byproduct a smooth Poincaré inequality on $\mathbb{H}(\mathbb{R})$ of independent interest; see Theorem 7.5.

The real Heisenberg group $\mathbb{H}(\mathbb{R})$ is defined as the matrix group

$$\mathbb{H}(\mathbb{R}) \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} 1 & u & w \\ 0 & 1 & v \\ 0 & 0 & 1 \end{pmatrix} : \ u, v, w \in \mathbb{R} \right\}.$$

The discrete Heisenberg group $\mathbb{H}$ sits inside $\mathbb{H}(\mathbb{R})$ as the cocompact discrete subgroup consisting of unipotent matrices with integer coefficients. We equip the group $\mathbb{H}(\mathbb{R})$ with the word metric $d_{S_{\mathbb{R}}}$ associated with the compact symmetric generating set $S_{\mathbb{R}} = \{a^u, b^v, c^w; |u|, |v|, |w| \leqslant 1\}$, where

$$a^u = \begin{pmatrix} 1 & u & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ b^v = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & v \\ 0 & 0 & 1 \end{pmatrix},\ c^w = \begin{pmatrix} 1 & 0 & w \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let μ denote a Haar measure on $\mathbb{H}(\mathbb{R})$, which coincides with Lebesgue measure under the natural identification of $\mathbb{H}(\mathbb{R})$ with $\mathbb{R}^3$.

Theorem 7.1. *For every continuous unitary representation π of $\mathbb{H}(\mathbb{R})$, any continuous cocycle $\gamma \in Z^1(\pi)$ satisfies the inequality*

$$\int_1^\infty \frac{\left\|\gamma\left(c^t\right)\right\|^2}{t^2} dt \lesssim \int_{-1}^1 \left(\|\gamma(a^u)\|^2 + \|\gamma(b^u)\|^2\right) du. \tag{45}$$

Section 7.1 is devoted to the proof of Theorem 7.1. It is clearly enough to check inequality (45) when the representation π is irreducible. The proof therefore boils down to a quantitative study of 1-cocycles with values in an irreducible representation of $\mathbb{H}(\mathbb{R})$. In the next three subsections we deduce Theorem 1.4 from Theorem 7.1 by a succession of reductions. Finally, in the last subsection, we state a smooth Poincaré inequality on the real Heisenberg group, whose proof, being very similar to the discrete one, is explained in a few sentences.

7.1 Proof of Theorem 7.1

By the Stone-von Neumann theorem (see for example [12, Ch. 2]), irreducible representations of $\mathbb{H}(\mathbb{R})$ are of two types: those that factor through the center, and, for every $\lambda \in \mathbb{R} \smallsetminus \{0\}$, the representation π_λ on $L_2(\mathbb{R})$ satisfying

$$\forall\, h \in L_2(\mathbb{R}), \quad \pi_\lambda(a^u b^v c^w)(h)(x) \stackrel{\text{def}}{=} e^{2\pi i \lambda v x} h(x+u) e^{2\pi i \lambda w}. \tag{46}$$

Note that if a nontrivial irreducible representation π factors through the center then any 1-cocycle $\gamma \in Z^1(\pi)$ must vanish on the center. Indeed, $\gamma(c^w)$ is invariant under $\pi(\mathbb{H})$ for every $w \in \mathbb{R}$, which, since the representation is supposed to be irreducible and nontrivial, implies that $\gamma(c^w) = 0$. Therefore, in proving Theorem 7.1 we may assume that $\pi = \pi_\lambda$ for some $\lambda \neq 0$.

By [14, Thm. 7] all 1-cocycles $\gamma \in Z^1(\pi_\lambda)$ can be approximated uniformly on compact sets by coboundaries. Hence, it is enough to consider the case where γ is of the form $\gamma(x) = \pi_\lambda(x)h - h$, for some $h \in L_2(\mathbb{R})$. We may assume that $\|h\| = 1$. By the definition (46), for every $w \in \mathbb{R}$ we have $\|\gamma(c^w)\|^2 = 4\sin^2(\pi\lambda w)$, from which we deduce that

$$\int_1^\infty \frac{\|\gamma(c^w)\|^2}{w^2} dw \lesssim \int_1^\infty \frac{\sin^2(\pi\lambda w)}{w^2} dw = |\lambda| \int_{|\lambda|}^\infty \frac{\sin^2(\pi w)}{w^2} dw \lesssim \min\{|\lambda|, 1\}. \tag{47}$$

Also,

$$\begin{aligned} \int_{-1}^1 \left(\|\gamma(a^u)\|^2 + \|\gamma(b^u)\|^2\right) du &= \int_{-1}^1 \int_{\mathbb{R}} \left(|h(x+u) - h(x)|^2 + 2|h(x)|^2 \left(1 - \cos(2\pi\lambda u x)\right)\right) dx du \\ &\asymp \int_{-1}^1 \int_{\mathbb{R}} |h(x+u) - h(x)|^2 dx du + \int_{\mathbb{R}} |h(x)|^2 \min\{\lambda^2 x^2, 1\} dx. \end{aligned} \tag{48}$$

Since $\mathbb{R} = (\mathbb{R} \setminus [-|u|,|u|]) \cup (u + (\mathbb{R} \setminus [-|u|,|u|])) \cup (-u + (\mathbb{R} \setminus [-|u|,|u|]))$ for every $u \in \mathbb{R}$, we can bound $\|h\|^2 = 1$ from above as follows.

$$\begin{aligned}
1 &\leqslant \int_{\mathbb{R}\setminus[-|u|,|u|]} \left(|h(x+u)|^2 + |h(x-u)|^2 + |h(x)|^2\right) dx \\
&\lesssim \int_{\mathbb{R}} \left(|h(x+u) - h(x)|^2 + |h(x) - h(x-u)|^2\right) dx + \int_{\mathbb{R}\setminus[-|u|,|u|]} |h(x)|^2 dx \\
&\leqslant \int_{\mathbb{R}} \left(|h(x+u) - h(x)|^2 + |h(x) - h(x-u)|^2\right) dx \\
&\quad + \frac{1}{\min\{\lambda^2 u^2, 1\}} \int_{\mathbb{R}} |h(x)|^2 \min\{\lambda^2 x^2, 1\} dx. \qquad (49)
\end{aligned}$$

Write $k = \left\lceil 1/\sqrt{|\lambda|} \right\rceil$. By applying (49) with $u = kv$, and integrating over $v \in [-1,-1/2] \cup [1/2,1]$, we see that

$$\begin{aligned}
\int_1^\infty \frac{\|\gamma(c^w)\|^2}{w^2} dw &\overset{(47)}{\lesssim} \min\{|\lambda|, 1\} \\
&\overset{(49)}{\lesssim} \min\{|\lambda|,1\} \int_{-1}^1 \int_{\mathbb{R}} |h(x+kv) - h(x)|^2 dx dv \\
&\quad + \frac{\min\{|\lambda|,1\}}{\min\left\{\lambda^2 \left\lceil 1/\sqrt{|\lambda|}\right\rceil^2, 1\right\}} \int_{\mathbb{R}} |h(x)|^2 \min\{\lambda^2 x^2, 1\} dx \\
&\lesssim \min\{|\lambda|,1\} k \sum_{j=1}^k \int_{-1}^1 \int_{\mathbb{R}} |h(x+jv) - h(x+(j-1)v)|^2 dx dv \\
&\quad + \int_{\mathbb{R}} |h(x)|^2 \min\{\lambda^2 x^2, 1\} dx \\
&= \min\{|\lambda|,1\} k^2 \int_{-1}^1 \int_{\mathbb{R}} |h(x+u) - h(x)|^2 dx du + \int_{\mathbb{R}} |h(x)|^2 \min\{\lambda^2 x^2, 1\} dx \\
&\overset{(48)}{\lesssim} \int_{-1}^1 \left(\|\gamma(a^u)\|^2 + \|\gamma(b^u)\|^2\right) du.
\end{aligned}$$

The proof of Theorem 7.1 is complete. □

7.2 Reduction to finitely supported functions

Claim 7.2. *Inequality (8) is a consequence of the following statement. For every finitely supported $\phi : \mathbb{H} \to L_2$, we have*

$$\sum_{x\in\mathbb{H}} \sum_{k=1}^\infty \frac{\left\|\phi(xc^k) - \phi(x)\right\|_2^2}{k^2} \lesssim \sum_{x\in\mathbb{H}} \left(\|\phi(xa) - \phi(x)\|_2^2 + \|\phi(xb) - \phi(x)\|_2^2\right). \qquad (50)$$

Proof. Fix $R \in \mathbb{N}$ and $f : \mathbb{H} \to L_2$. Note that since (8) is not sensitive to adding a constant to the function f, we can assume without loss of generality that the average of f over B_{7R} is zero.

Define a cutoff function $\xi : \mathbb{H} \to [0,1]$ by

$$\xi(x) \stackrel{\text{def}}{=} \begin{cases} 1 & x \in B_{5R}, \\ 6 - \frac{d_W(x,e)}{R} & x \in B_{6R} \setminus B_{5R}, \\ 0 & x \in \mathbb{H} \setminus B_{6R}, \end{cases}$$

and let $\phi \stackrel{\text{def}}{=} \xi f$. Then ϕ is supported on B_{6R}. Since ξ is $1/R$-Lipschitz and takes values in $[0,1]$, for all $x \in \mathbb{H}$ and $s \in S$,

$$\begin{aligned} \|\phi(x) - \phi(xs)\|_2^2 &\lesssim |\xi(x) - \xi(xs)|^2 \cdot \|f(x)\|_2^2 + |\xi(xs)|^2 \cdot \|f(x) - f(xs)\|_2^2 \\ &\leqslant \frac{1}{R^2}\|f(x)\|^2 + \|f(x) - f(xs)\|_2^2. \end{aligned} \tag{51}$$

Note that if $k \in \{1, \ldots, R^2\}$ then $d_W(e, c^k) \leqslant 4R$, and hence for $x \in B_R$ we have $xc^k \in B_{5R}$. Therefore, an application of (50) to ϕ yields the estimate

$$\begin{aligned} \sum_{x \in B_R} \sum_{k=1}^{R^2} \frac{\|f(xc^k) - f(x)\|_2^2}{k^2} &\leqslant \sum_{x \in \mathbb{H}} \sum_{k=1}^{\infty} \frac{\|\phi(xc^k) - \phi(x)\|_2^2}{k^2} \lesssim \sum_{x \in \mathbb{H}} \max_{s \in S} \|\phi(xs) - \phi(x)\|_2^2 \\ &= \sum_{x \in B_{7R}} \max_{s \in S} \|\phi(xs) - \phi(x)\|_2^2 \stackrel{(51)}{\leqslant} \frac{1}{R^2} \sum_{x \in B_{7R}} \|f(x)\|_2^2 + \sum_{x \in B_{7R}} \max_{s \in S} \|f(xs) - f(x)\|_2^2. \end{aligned} \tag{52}$$

By [17, Thm. 2.2] (a discrete version of the classical Heisenberg local Poincaré inequality [16]),

$$\frac{1}{R^2} \sum_{x \in B_{7R}} \|f(x)\|_2^2 \lesssim \sum_{x \in B_{22R}} \left(\|f(xa) - f(x)\|_2^2 + \|f(xb) - f(x)\|_2^2 \right), \tag{53}$$

where we used the fact that the average of f on B_{7R} vanishes. The desired inequality (8) is now a consequence of (52) and (53). □

7.3 Reduction to an inequality on the real Heisenberg group

Claim 7.3. *Inequality (50) is a consequence of the following statement. For every continuous and compactly supported function $\psi : \mathbb{H}(\mathbb{R}) \to L_2$, we have*

$$\int_{\mathbb{H}(\mathbb{R})} \int_1^{\infty} \frac{\left\|\psi(xc^t) - \psi(x)\right\|_2^2}{t^2} dt d\mu(x) \lesssim \int_{\mathbb{H}(\mathbb{R})} \left(\sup_{s \in S_{\mathbb{R}}} \|\psi(xs) - \psi(x)\|_2^2 \right) d\mu(x). \tag{54}$$

Proof. For $r > 0$ let $B_r^{\mathbb{R}} \subseteq \mathbb{H}(\mathbb{R})$ denote the ball of radius r with respect to the metric $d_{S_{\mathbb{R}}}$. Note that $\mathbb{H}(\mathbb{R}) = \bigcup_{g \in \mathbb{H}} g B_2^{\mathbb{R}}$. Let $\sigma : \mathbb{H}(\mathbb{R}) \to [0,1]$ be a continuous nonnegative function, which equals 1 on $B_2^{\mathbb{R}}$ and 0 outside of $B_3^{\mathbb{R}}$. Let $\tilde{\sigma} = \sum_{g \in \mathbb{H}} \sigma_g$, where $\sigma_g(x) = \sigma(g^{-1}x)$. For all $x \in \mathbb{H}(\mathbb{R})$ we have $1 \leqslant \tilde{\sigma}(x) \leqslant C$ for some $C \in (0, \infty)$. Writing $\beta = \sigma/\tilde{\sigma}$ and $\beta_g(x) = \beta(g^{-1}x)$, we see that $\{\beta_g\}_{g \in \mathbb{H}}$ is a continuous partition of unity for $\mathbb{H}(\mathbb{R})$ satisfying

$$\sup_{x \in \mathbb{H}(\mathbb{R})} \sum_{g \in \mathbb{H}} \sup_{s \in S_{\mathbb{R}}} |\beta_g(xs) - \beta_g(x)| < \infty. \tag{55}$$

Throughout the ensuing argument we will use repeatedly the fact that for every $x \in \mathbb{H}(\mathbb{R})$ the number of elements $g \in \mathbb{H}$ for which $\beta_g(x) > 0$ is bounded by a constant independent of x, and that $\sum_{g\in\mathbb{H}} \mathbf{1}_{gB_2} \asymp \mathbf{1}_{\mathbb{H}(\mathbb{R})}$.

Let $\phi : \mathbb{H} \to L_2$ be a finitely supported function on the discrete Heisenberg group $\mathbb{H}$. Define a function on the real Heisenberg group $\mathbb{H}(\mathbb{R})$ by

$$\psi(x) = \sum_{g\in\mathbb{H}} \phi(g)\beta_g(x).$$

Then ψ is compactly supported. We will eventually apply (54) to ψ, but before doing so we will need some preparatory estimates.

The metric $d_{S_\mathbb{R}}$ restricted to $\mathbb{H} \subseteq \mathbb{H}(\mathbb{R})$ is bi-Lipschitz equivalent to d_W [5, Thm. 8.3.19]. It follows that for all $g_0 \in \mathbb{H}$, if $x \in g_0 B_3^{\mathbb{R}}$ then the sum $\psi(x) - \phi(g_0) = \sum_{g\in\mathbb{H}} \beta_g(x)(\phi(g) - \phi(g_0))$ is supported on elements of the form $g = g_0 h$, with $d_W(h, e) \leqslant K$, for some universal constant $K \in \mathbb{N}$. Thus, using (55) we see that for all $x \in g_0 B_2^{\mathbb{R}}$,

$$\sup_{s\in S_\mathbb{R}} \|\psi(xs) - \psi(x)\|_2^2 \lesssim \sum_{h\in B_K} \|\phi(g_0 h) - \phi(g_0)\|_2^2. \tag{56}$$

Integrating (56) over $g_0 B_2^{\mathbb{R}}$ gives the following inequality:

$$\int_{g_0 B_2^{\mathbb{R}}} \left(\sup_{s\in S_\mathbb{R}} \|\psi(xs) - \psi(x)\|_2^2 \right) d\mu(x) \lesssim \sum_{h\in B_K} \|\phi(g_0 h) - \phi(g_0)\|_2^2. \tag{57}$$

By summing (57) over $g_0 \in \mathbb{H}$ we see that

$$\int_{\mathbb{H}(\mathbb{R})} \left(\sup_{s\in S_\mathbb{R}} \|\psi(xs) - \psi(x)\|_2^2 \right) d\mu(x) \lesssim \sum_{g\in\mathbb{H}} \max_{s\in S} \|\phi(g) - \phi(gs)\|_2^2, \tag{58}$$

where we used the bound

$$\sum_{g_0\in\mathbb{H}} \sum_{h\in B_K} \|\phi(g_0 h) - \phi(g_0)\|_2^2 \lesssim \sum_{g\in\mathbb{H}} \max_{s\in S} \|\phi(g) - \phi(gs)\|_2^2, \tag{59}$$

which follows by writing each $h \in B_K$ as a product of at most K elements of S, and using the triangle inequality.

In order to deduce from (54) a corresponding bound on ϕ, we need to bound ϕ in terms of ψ. To this end, note that for all $g_0 \in \mathbb{H}$,

$$\psi(g_0) = \phi(g_0) + \sum_{g\in\mathbb{H}} \beta(g^{-1}g_0)(\phi(g) - \phi(g_0)) = \phi(g_0) + \sum_{h\in B_K} \beta(h^{-1})(\phi(g_0 h) - \phi(g_0)).$$

It follows that for all $g_0, g_1 \in \mathbb{H}$ we have

$$\|\phi(g_0) - \phi(g_1)\|_2^2 \lesssim \|\psi(g_0) - \psi(g_1)\|_2^2 + \sum_{h\in B_K} \left(\|\phi(g_0) - \phi(g_0 h)\|_2^2 + \|\phi(g_1) - \phi(g_1 h)\|_2^2 \right). \tag{60}$$

If $x, y \in \mathbb{H}(\mathbb{R})$ satisfy $\max\{d_{S_\mathbb{R}}(x, g_0), d_{S_\mathbb{R}}(y, g_1)\} \leqslant 2$, then using (56) we deduce from (60) that

$$\|\phi(g_0) - \phi(g_1)\|_2^2 \lesssim \|\psi(x) - \psi(y)\|_2^2 + \sum_{h\in B_K} \left(\|\phi(g_0) - \phi(g_0 h)\|_2^2 + \|\phi(g_1) - \phi(g_1 h)\|_2^2 \right). \tag{61}$$

Fix $g_0 \in \mathbb{H}$, $x \in g_0 B_1^{\mathbb{R}}$, $k \in \mathbb{N}$ and $t \in [k, k+1]$. Writing $g_1 = g_0 c^k$ and $y = xc^t$, we have $d_{S_{\mathbb{R}}}(x, g_0) \leqslant 1$ and $d_{S_{\mathbb{R}}}(y, g_1) = d_{S_{\mathbb{R}}}(c^{-k} g_0^{-1} x c^t, e) = d_{S_{\mathbb{R}}}(g_0^{-1} x c^{t-k}, e) \leqslant d_{S_{\mathbb{R}}}(g_0^{-1} x, e) + d_{S_{\mathbb{R}}}(c^{t-k}, e) \leqslant 2$. We may therefore apply (61) and deduce that for all $g_0 \in \mathbb{H}$, $x \in g_0 B_1^{\mathbb{R}}$ and $k \in \mathbb{N}$,

$$\|\phi(g_0 c^k) - \phi(g_0)\|_2^2 \lesssim \|\psi(xc^t) - \psi(x)\|_2^2 + \sum_{h \in B_K} \left(\|\phi(g_0) - \phi(g_0 h)\|_2^2 + \|\phi(g_0 c^k) - \phi(g_0 c^k h)\|_2^2 \right). \quad (62)$$

Integrating (62) over $x \in g_0 B_1^{\mathbb{R}}$ and $t \in [k, k+1]$, we see that

$$\begin{aligned} \frac{\|\phi(g_0 c^k) - \phi(g_0)\|_2^2}{k^2} \lesssim \int_{g_0 B_1^{\mathbb{R}}} \int_k^{k+1} \frac{\|\psi(xc^t) - \psi(x)\|_2^2}{t^2} dt d\mu(x) \\ + \frac{1}{k^2} \sum_{h \in B_K} \left(\|\phi(g_0) - \phi(g_0 h)\|_2^2 + \|\phi(g_0 c^k) - \phi(g_0 c^k h)\|_2^2 \right). \quad (63) \end{aligned}$$

Since $g_0 B_1^{\mathbb{R}}$ and $g_0' B_1^{\mathbb{R}}$ intersect at a set of measure zero if $g_0 \neq g_0'$, by summing (63) over $g_0 \in \mathbb{H}$ and $k \in \mathbb{N}$, we see that

$$\begin{aligned} &\sum_{g_0 \in \mathbb{H}} \sum_{k=1}^{\infty} \frac{\|\phi(g_0 c^k) - \phi(g_0)\|_2^2}{k^2} \\ &\lesssim \int_{\mathbb{H}(\mathbb{R})} \int_1^{\infty} \frac{\|\psi(xc^t) - \psi(x)\|_2^2}{t^2} dt d\mu(x) + \sum_{g_0 \in \mathbb{H}} \sum_{h \in B_K} \|\phi(g_0) - \phi(g_0 h)\|_2^2 \\ &\overset{(59)}{\lesssim} \int_{\mathbb{H}(\mathbb{R})} \int_1^{\infty} \frac{\|\psi(xc^t) - \psi(x)\|_2^2}{t^2} dt d\mu(x) + \sum_{g \in \mathbb{H}} \max_{s \in S} \|\phi(g) - \phi(gs)\|_2^2 \\ &\overset{(54)}{\lesssim} \int_{\mathbb{H}(\mathbb{R})} \left(\sup_{s \in S_{\mathbb{R}}} \|\psi(xs) - \psi(x)\|_2^2 \right) d\mu(x) + \sum_{g \in \mathbb{H}} \max_{s \in S} \|\phi(g) - \phi(gs)\|_2^2 \\ &\overset{(58)}{\lesssim} \sum_{g \in \mathbb{H}} \max_{s \in S} \|\phi(g) - \phi(gs)\|_2^2. \end{aligned}$$

□

7.4 Reduction to a 1-cocycle on $\mathbb{H}(\mathbb{R})$

Claim 7.4. *Inequality (54) follows from Theorem 7.1.*

Proof. Let $\psi : \mathbb{H}(\mathbb{R}) \to L_2$ be continuous and supported in $B_r^{\mathbb{R}}$ for some $r \geqslant 1$. Take a maximal family $\{x_i B_{10r}^{\mathbb{R}}\}_{i=1}^{\infty}$ of disjoint balls of radius $10r$. Define $\varphi : \mathbb{H}(\mathbb{R}) \to L_2$ by $\varphi(g) = \sum_{i=1}^{\infty} \psi(x_i^{-1} g)$. Note that since $d_{S_{\mathbb{R}}}(c^{r^2}, e) \leqslant 4r$, for each $i \in \mathbb{N}$ we have

$$\int_{\mathbb{H}(\mathbb{R})} \int_1^{\infty} \frac{\left\|\psi(xc^t) - \psi(x)\right\|_2^2}{t^2} dt d\mu(x) = \int_{x_i B_{5r}^{\mathbb{R}}} \int_1^{r^2} \frac{\left\|\varphi(xc^t) - \varphi(x)\right\|_2^2}{t^2} dt d\mu(x). \quad (64)$$

Similarly,

$$\int_{\mathbb{H}(\mathbb{R})} \left(\sup_{s \in S_{\mathbb{R}}} \|\psi(xs) - \psi(x)\|_2^2 \right) d\mu(x) = \int_{x_i B_{2r}^{\mathbb{R}}} \left(\sup_{s \in S_{\mathbb{R}}} \|\varphi(xs) - \varphi(x)\|_2^2 \right) d\mu(x). \quad (65)$$

Let X be the space of all finitely supported complex-valued functions on $\mathbb{H}(\mathbb{R})$. We denote by π the action of $\mathbb{H}(\mathbb{R})$ on X given by $\pi(x)\delta_y \stackrel{\text{def}}{=} \delta_{xy}$, where $\delta_z : \mathbb{H}(\mathbb{R}) \to \mathbb{C}$ denotes the function which equals 1 at $z \in \mathbb{H}(\mathbb{R})$ and equals 0 elsewhere. Let $\mathscr{U}$ be a free ultrafilter on $\mathbb{N}$. Define a scalar product $[\cdot,\cdot]$ on X by:

$$[\delta_x, \delta_y] \stackrel{\text{def}}{=} \lim_{n\to\mathscr{U}} \frac{1}{\mu(B_n^{\mathbb{R}})} \int_{B_n^{\mathbb{R}}} \langle \varphi(zx), \varphi(zy)\rangle d\mu(z), \tag{66}$$

where $\langle\cdot,\cdot\rangle$ denotes the scalar product on L_2.

Since $\{B_n\}_{n=1}^{\infty}$ is a Følner sequence for $\mathbb{H}(\mathbb{R})$, a limit along a free ultrafilter of averages over B_n when $n \to \infty$ is an invariant mean on $\mathbb{H}(\mathbb{R})$. It follows that the scalar product defined in (66) is $\pi(\mathbb{H}(\mathbb{R}))$-invariant. Thus π is a unitary representation of $\mathbb{H}(\mathbb{R})$ (formally we should first pass to the completion of the quotient of X by the subspace consisting of norm-zero elements, but we will ignore this inessential point in what follows). We note that π is also continuous in the strong operator topology. Indeed, since ψ is continuous and compactly supported, φ is uniformly continuous. Thus, writing $\|f\|_X^2 \stackrel{\text{def}}{=} [f,f]$ for $f \in X$, we have for every $y \in \mathbb{H}(\mathbb{R})$,

$$\begin{aligned}\lim_{x\to e} \|\pi(x)\delta_y - \delta_y\|_X^2 &= \lim_{x\to e}\lim_{n\to\mathscr{U}} \frac{1}{\mu(B_n^{\mathbb{R}})}\int_{B_n^{\mathbb{R}}} \|\varphi(gxy)-\varphi(gy)\|_2^2 d\mu(g)\\ &\leqslant \lim_{x\to e}\sup_{g\in\mathbb{H}(\mathbb{R})} \|\varphi(gxy)-\varphi(gy)\|_2^2 = 0,\end{aligned} \tag{67}$$

implying the strong continuity of π.

Let $\gamma : \mathbb{H}(\mathbb{R}) \to X$ be given by $\gamma(x) = \delta_x - \delta_e$. Then $\pi \in Z^1(\pi)$ and for all $x \in X$,

$$\|\gamma(x)\|_X^2 = \lim_{n\to\mathscr{U}} \frac{1}{\mu(B_n^{\mathbb{R}})}\int_{B_n^{\mathbb{R}}} \|\varphi(gx)-\varphi(g)\|_2^2 d\mu(g). \tag{68}$$

Arguing as in (67), the uniform continuity of φ and (68) imply that γ is a continuous 1-cocycle.

Fix $n > 100r$ large enough so as to ensure that we have $\mu(B_n \smallsetminus B_{n-100r}) \leqslant \mu(B_n)/2$. Define $I = \{i \in \mathbb{N} :\ x_i B_{5r}^{\mathbb{R}} \subseteq B_n\}$ and write $\Omega = \bigcup_{i\in I} x_i B_{5r}^{\mathbb{R}}$ and $\Omega' = \bigcup_{i\in I} x_i B_{10r}^{\mathbb{R}}$. By the maximality of $\left\{x_i B_{10r}^{\mathbb{R}}\right\}_{i=1}^{\infty}$ we have $\Omega' \supseteq B_{n-100r}$. Hence $\mu(\Omega') \geqslant \mu(B_n)/2$. Since $\mathbb{H}(\mathbb{R})$ is doubling, $\mu(\Omega') \lesssim \mu(\Omega)$, and therefore $|I|\mu(B_{5r}^{\mathbb{R}}) \geqslant \mu(\Omega) \gtrsim \mu(B_n)$. Note that if $g \in x_i B_{5r}^{\mathbb{R}}$ for some $i \in \mathbb{N}$ then for every $t \in [1, r^2]$ we have $gc^t \in x_i B_{10r}$. Hence,

$$\begin{aligned}\frac{1}{\mu(B_n^{\mathbb{R}})}\int_{B_n^{\mathbb{R}}} \|\varphi(gc^t)-\varphi(g)\|_2^2 d\mu(g) &\geqslant \frac{1}{\mu(B_n^{\mathbb{R}})}\sum_{i\in I}\int_{x_i B_{5r}^{\mathbb{R}}} \|\varphi(gc^t)-\varphi(g)\|_2^2 d\mu(g)\\ &= \frac{|I|}{\mu(B_n^{\mathbb{R}})}\int_{B_{5r}^{\mathbb{R}}} \|\psi(xc^t)-\psi(x)\|_2^2 d\mu(x) \gtrsim \frac{1}{\mu(B_{5r}^{\mathbb{R}})}\int_{B_{5r}^{\mathbb{R}}} \|\psi(xc^t)-\psi(x)\|_2^2 d\mu(x).\end{aligned} \tag{69}$$

It follows from (68) and (69) that

$$\begin{aligned}\frac{1}{\mu(B_{5r}^{\mathbb{R}})}\int_{B_{5r}^{\mathbb{R}}}\int_1^{r^2} \frac{\|\psi(xc^t)-\psi(x)\|_2^2}{t^2} d\mu(x) &\lesssim \int_1^{r^2} \frac{\|\gamma(c^t)\|_X^2}{t^2} dt \overset{(45)}{\lesssim} \sup_{s\in S_{\mathbb{R}}} \|\gamma(s)\|_X^2\\ &\overset{(68)}{=} \sup_{s\in S_{\mathbb{R}}}\lim_{n\to\mathscr{U}} \frac{1}{\mu(B_n^{\mathbb{R}})}\int_{B_n^{\mathbb{R}}} \|\varphi(gs)-\varphi(g)\|_2^2 d\mu(g).\end{aligned} \tag{70}$$

Let $J \subseteq \mathbb{N}$ denote the set of $i \in \mathbb{N}$ such that $B_n \cap x_i B_{2r} \neq \emptyset$. Then $|J|\mu(B^{\mathbb{R}}_{2r}) \leqslant \mu(B^{\mathbb{R}}_{n+4r})$. It follows that for all $s \in S_{\mathbb{R}}$,

$$\begin{aligned}\int_{B^{\mathbb{R}}_n} \|\varphi(gs)-\varphi(g)\|_2^2 d\mu(g) &\leqslant \sum_{i\in J} \int_{x_i B_{2r}} \|\varphi(gs)-\varphi(g)\|_2^2 d\mu(g) \\ &= |J| \int_{B_{2r}} \|\psi(xs)-\psi(x)\|_2^2 d\mu(x) \leqslant \frac{\mu(B^{\mathbb{R}}_{n+4r})}{\mu(B^{\mathbb{R}}_{2r})} \int_{B_{2r}} \|\psi(xs)-\psi(x)\|_2^2 d\mu(x). \quad (71)\end{aligned}$$

Substituting (71) into (70) we conclude that

$$\begin{aligned}\int_{\mathbb{H}(\mathbb{R})} \int_1^\infty \frac{\left\|\psi(xc^t)-\psi(x)\right\|_2^2}{t^2} dt d\mu(x) &\overset{(64)}{=} \int_{B^{\mathbb{R}}_{5r}} \int_1^{r^2} \frac{\|\psi(xc^t)-\psi(x)\|_2^2}{t^2} d\mu(x) \\ &\overset{(70)\wedge(71)}{\lesssim} \frac{\mu(B^{\mathbb{R}}_{5r})}{\mu(B^{\mathbb{R}}_{2r})} \sup_{s\in S_{\mathbb{R}}} \lim_{n\to\mathscr{U}} \frac{\mu(B^{\mathbb{R}}_{n+4r})}{\mu(B^{\mathbb{R}}_n)} \int_{B_{2r}} \|\psi(xs)-\psi(x)\|_2^2 d\mu(x) \\ &\overset{(65)}{\lesssim} \int_{\mathbb{H}(\mathbb{R})} \left(\sup_{s\in S_{\mathbb{R}}} \|\psi(xs)-\psi(x)\|_2^2 \right) d\mu(x)\end{aligned}$$

This completes the proof of Claim 7.4, and therefore also the proof of Theorem 1.4. □

7.5 A smooth Poincaré inequality on $\mathbb{H}(\mathbb{R})$

Equip $\mathbb{H}(\mathbb{R})$ with the left-invariant Riemannian metric given by $du^2 + dv^2 + (dw - udv)^2$. In what follows, given a smooth function $f : \mathbb{H}(\mathbb{R}) \to \mathbb{R}$ we let $\nabla_{\mathbb{H}} f$ denote its gradient with respect to this Riemannian structure.

Theorem 7.5. *For every smooth function $f : \mathbb{H}(\mathbb{R}) \to \mathbb{R}$, and all $R > 0$,*

$$\int_{B^{\mathbb{R}}_R} \int_1^{R^2} \frac{\left|f(xc^t)-f(x)\right|^2}{t^2} dt d\mu(x) \lesssim \int_{B^{\mathbb{R}}_{CR}} |\nabla_{\mathbb{H}} f(x)|^2 \, d\mu(x),$$

where $C > 0$ is a universal constant.

The proof of this Poincaré inequality can be obtained from Theorem 7.1 in a way similar, and actually even shorter than its discrete counterpart. Indeed, the discretization step of Claim 7.3 is not needed here. The other difference lies in the first step, where instead of the discrete Poincaré inequality (53), we use the following smooth version, which is due to [16]. For all $R > 0$, and for all smooth functions $f : \mathbb{H}(\mathbb{R}) \to \mathbb{R}$ whose integral over $B^{\mathbb{R}}_R$ is zero,

$$\frac{1}{R^2} \int_{B^{\mathbb{R}}_R} |f(g)|^2 d\mu(g) \lesssim \int_{B^{\mathbb{R}}_{cR}} |\nabla_{\mathbb{H}} f(g)|^2 d\mu(g),$$

where $c > 0$ is a universal constant.

References

[1] P. Assouad. Plongements Lipschitziens dans $\mathbf{R}^n$. *Bull. Soc. Math. France*, 111(4):429–448, 1983.

[2] K. Ball, E. A. Carlen, and E. H. Lieb. Sharp uniform convexity and smoothness inequalities for trace norms. *Invent. Math.*, 115(3):463–482, 1994.

[3] Y. Benyamini and J. Lindenstrauss. *Geometric nonlinear functional analysis. Vol. 1*, volume 48 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2000.

[4] A. Brunel and L. Sucheston. Sur quelques conditions équivalentes à la super-réflexivité dans les espaces de Banach. *C. R. Acad. Sci. Paris Sér. A-B*, 275:A993–A994, 1972.

[5] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

[6] J. Cheeger and B. Kleiner. Differentiating maps into L^1 and the geometry of BV functions. To appear in Ann. Math., preprint available at `http://arxiv.org/abs/math/0611954`, 2006.

[7] J. Cheeger and B. Kleiner. On the differentiability of Lipschtz maps from metric measure spaces into banach spaces. In *Inspired by S.S. Chern, A Memorial volume in honor of a great mathematician*, volume 11 of *Nankai tracts in Mathematics*, pages 129–152. World Scientific, Singapore, 2006.

[8] J. Cheeger, B. Kleiner, and A. Naor. Compression bounds for Lipschitz maps from the Heisenberg group to L_1. Preprint, 2009. `http://arxiv.org/abs/0910.2026`.

[9] J. Cheeger, B. Kleiner, and A. Naor. A $(\log n)^{\Omega(1)}$ integrality gap for the Sparsest Cut SDP. In *Proceedings of 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2009)*, pages 555–564, 2009.

[10] Y. de Cornulier, R. Tessera, and A. Valette. Isometric group actions on Hilbert spaces: growth of cocycles. *Geom. Funct. Anal.*, 17(3):770–792, 2007.

[11] N. Dunford and J. T. Schwartz. *Linear operators. Part I.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[12] G. B. Folland. *Harmonic analysis in phase space*, volume 122 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1989.

[13] M. Gromov. Asymptotic invariants of infinite groups. In *Geometric group theory, Vol. 2 (Sussex, 1991)*, volume 182 of *London Math. Soc. Lecture Note Ser.*, pages 1–295. Cambridge Univ. Press, Cambridge, 1993.

[14] A. Guichardet. Sur la cohomologie des groupes topologiques. II. *Bull. Sci. Math. (2)*, 96:305–332, 1972.

[15] A. Gupta, R. Krauthgamer, and J. R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *44th Symposium on Foundations of Computer Science (FOCS 2003)*, pages 534–543. IEEE Computer Society, 2003.

[16] D. Jerison. The Poincaré inequality for vector fields satisfying Hörmander's condition. *Duke Math. J.*, 53(2):503–523, 1986.

[17] B. Kleiner. A new proof of Gromov's theorem on groups of polynomial growth. *J. Amer. Math. Soc.*, 23(4):815–829 (electronic), 2010.

[18] J. R. Lee and A. Naor. L_p metrics on the Heisenberg group and the Goemans-Linial conjecture. In *Proceedings of 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006)*, pages 99–108, 2006. Available at `http://www.cims.nyu.edu/~naor/homepage%20files/L_pHGL.pdf`.

[19] M. Mendel and A. Naor. Towards a calculus for non-linear spectral gaps. In *SODA '10: Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, 2010. Available at `http://arxiv.org/PS_cache/arxiv/pdf/0910/0910.2041v1.pdf`.

[20] A. Naor. L_1 embeddings of the Heisenberg group and fast estimation of graph isoperimetry. To appear in *Proceedings of the International Congress of Mathematicians, (Hyderabad, 2010)*, 2010.

[21] A. Naor and Y. Peres. L_p compression, traveling salesmen, and stable walks. To appear in Duke Math. J., preprint available at `http://arxiv.org/abs/0904.4728`, 2009.

[22] P. Pansu. Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un. *Ann. of Math. (2)*, 129(1):1–60, 1989.

[23] G. Pisier. Martingales with values in uniformly convex spaces. *Israel J. Math.*, 20(3-4):326–350, 1975.

[24] S. Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the Fifteenth Annual Symposium on Computational Geometry (Miami Beach, FL, 1999)*, pages 300–306 (electronic), New York, 1999. ACM.

[25] S. Semmes. On the nonexistence of bi-Lipschitz parameterizations and geometric problems about A_∞-weights. *Rev. Mat. Iberoamericana*, 12(2):337–410, 1996.

[26] R. Tessera. Asymptotic isoperimetry on groups and uniform embeddings into Banach spaces. To appear in Comment. Math. Helv., preprint available at `http://arxiv.org/abs/math/0603138`, 2006.

[27] R. Tessera. Quantitative property A, Poincaré inequalities, L^p-compression and L^p-distortion for metric measure spaces. *Geom. Dedicata*, 136:203–220, 2008.