Acta Math., 229 (2022), 55–200
DOI: 10.4310/ACTA.2022.v229.n1.a2

Foliated corona decompositions

by

Assaf Naor
Princeton University
Princeton, NJ, U.S.A.

Robert Young
New York University
New York, NY, U.S.A.

Dedicated to the memory of Louis Nirenberg.

Contents

1. Introduction 56
1.1. Geometric implications 59
1.1.1. Embeddings 60
1.1.2. Aspects of the Ribe program 62
1.1.3. Dimension reduction 69
1.1.4. Permanence of compression rates of groups 72
1.2. Decomposing surfaces into approximately ruled pieces . . . 73
1.2.1. Fractal Venetian blinds abound 75
1.2.2. A maximally bumpy surface 81
1.3. Roadmap 84
2. Preliminaries 84
2.1. The Heisenberg group 85
2.2. Intrinsic graphs and intrinsic Lipschitz graphs 87
2.3. Characteristic curves 90
2.4. Automorphisms and characteristic curves 93
2.5. Measures on lines and the kinematic formula 95
2.6. Vertical perimeter and parametric vertical perimeter 96
3. Constructing surfaces and embeddings 98
3.1. Obtaining an embedding from an intrinsic graph 104
3.2. Constructing a bumpy intrinsic graph 108
3.2.1. The horizontal perimeter of Γ_{ψ_i} 110
3.2.2. The vertical perimeter of Γ_{ψ_i} 117

A. N. was supported by the BSF, the Packard Foundation and the Simons Foundation. R. Y. was supported by NSF grant 1612061 and the Sloan Foundation. The research that is presented here was conducted under the auspices of the Simons Algorithms and Geometry (A&G) Think Tank.

4. Pseudoquads and foliated patchworks 121
5. Foliated corona decompositions 130
6. Vertical perimeter and foliated corona decompositions 132
6.1. Vertical perimeter for graphs with foliated corona decompositions 132
7. The subdivision algorithm: constructing a foliated corona decomposition 137
8. Extended non-monotonicity 141
8.1. Extended non-monotonicity in $\mathbb{R}$ 141
8.2. Extended non-monotonicity in $\mathbb{H}$ 144
9. The kinematic formula and the proof of Proposition 7.4 148
10. Outline of proof of Proposition 7.2 155
11. Extended-monotone sets are close to half-spaces 157
11.1. Stability of locally monotone sets 159
11.2. Locally monotone sets are bounded by rectifiable ruled surfaces 161
11.3. Stability of extended monotone sets 170
12. L_1 bounds and characteristic curves on monotone intrinsic graphs 175
12.1. Bounding the tails of f 177
12.2. Constructing the approximating plane 181
12.3. Bounding $\|f-F\|_{L_1(10Q)}$ 183
12.4. Characteristic curves are close to lines 186
Acknowledgements 190
Appendix A. On the implicit dependence on p in [58] 190
Added in proof 195
References 196

1. Introduction

Since our main theorem (Theorem 1.1 below) can be stated without the need to recall any specialized background, we will start by formulating it. After doing so, we will explain its significance and context, as well as geometric applications that answer longstanding open questions. We will then describe our main conceptual contribution, called a *foliated corona decomposition*, which is a new structural methodology that we introduce in the *proof* of this theorem; see Remark 1.2 and mainly §1.2 for an overview.

For a smooth function $f:\mathbb{R}^3\to\mathbb{R}$ define $\mathsf{X}f,\mathsf{Y}f:\mathbb{R}^3\to\mathbb{R}$ by setting, for $h=(x,y,z)\in\mathbb{R}^3$,

$$\mathsf{X}f(h)\stackrel{\mathrm{def}}{=}\frac{\partial f}{\partial x}(h)+\frac{1}{2}y\frac{\partial f}{\partial z}(h)\quad\text{and}\quad \mathsf{Y}f(h)\stackrel{\mathrm{def}}{=}\frac{\partial f}{\partial y}(h)-\frac{1}{2}x\frac{\partial f}{\partial z}(h). \tag{1.1}$$

Also, for $t\in(0,\infty)$, define $D_{\mathsf{v}}^t f:\mathbb{R}^3\to\mathbb{R}$ by setting, for $h=(x,y,z)\in\mathbb{R}^3$,

$$D_{\mathsf{v}}^t(h) \stackrel{\mathrm{def}}{=} \frac{f(x,y,z+t)-f(h)}{\sqrt{t}}. \tag{1.2}$$

THEOREM 1.1. *Every compactly supported smooth function $f:\mathbb{R}^3\to\mathbb{R}$ satisfies*([1])

$$\left(\int_0^\infty \left(\int_{\mathbb{R}^3} |D_{\mathsf{v}}^t f(h)|\,dh\right)^4 \frac{dt}{t}\right)^{1/4} \lesssim \int_{\mathbb{R}^3} (|\mathsf{X}f(h)|+|\mathsf{Y}f(h)|)\,dh. \tag{1.3}$$

Also, one cannot replace the $L_4(dt/t)$ norm above by an $L_q(dt/t)$ norm for any $0<q<4$.

The second assertion (sharpness) of Theorem 1.1 resolves negatively the conjecture of [58] that (1.3) holds with the $L_4(dt/t)$ norm in the left-hand side replaced by the $L_2(dt/t)$ norm. Notwithstanding the optimality of (1.3), it should be noted that it was previously unknown whether such a bound holds true merely for *some* finite exponent, namely that there exists $0<p<\infty$ such that, in the setting of Theorem 1.1, we have

$$\left(\int_0^\infty \left(\int_{\mathbb{R}^3} |D_{\mathsf{v}}^t f(h)|\,dh\right)^p \frac{dt}{t}\right)^{1/p} \lesssim \int_{\mathbb{R}^3} (|\mathsf{X}f(h)|+|\mathsf{Y}f(h)|)\,dh. \tag{1.4}$$

It is simple to justify (see [91, Remark 4]) that if (1.4) holds, then the analogous bound holds for any larger exponent $P>p$.

Remark 1.2. To briefly indicate what goes into Theorem 1.1, we first note that the functional inequality (1.3) is equivalent to a certain isoperimetric-type inequality (see (1.31)) for sufficiently smooth surfaces in $\mathbb{R}^3$. By [91], it turns out that it suffices to prove this isoperimetric-type inequality for a more restricted class of surfaces (intrinsic Lipschitz graphs; see §2.2). Such surfaces can still be very complicated, as one can see in Figure 1. However, notice that the example in Figure 1 has an anisotropic texture, with features of many different scales that line up along a 1-dimensional foliation.

We prove the desired isoperimetric-type inequality by showing that the texture of *any* intrinsic Lipschitz graph can be encoded as a foliated corona decomposition, which is a multi-scale hierarchical partition of the surface. The pieces of this decomposition are roughly rectangular regions that mimic the dimensions and orientation of the features of the surface. Crucially, we can control the number and size of these pieces. The desired inequality holds locally on each piece up to suitably controlled error, and the full inequality is obtained by summing the resulting estimates. This process is illustrated in Figures 2 and 3, and a more detailed overview can be found in §1.2.

([1]) We will use throughout the following (standard) asymptotic notation. For $a,b\in(0,\infty)$, the notation $a\lesssim b$ and $b\gtrsim a$ mean that $a\leqslant Cb$ for some universal constant $C\in(0,\infty)$. The notation $a\asymp b$ stands for $(a\lesssim b)\wedge(b\lesssim a)$. If we need to allow for dependence on parameters, we indicate this by subscripts. For example, in the presence of an auxiliary parameter q, the notation $a\lesssim_q b$ means that $a\leqslant C(q)b$, where $C(q)\in(0,\infty)$ is allowed to depend only on q, and analogously for the notation $a\gtrsim_q b$ and $a\asymp_q b$.

Figure 1. An example of an intrinsic Lipschitz graph.

In contrast to Theorem 1.1, we have the following theorem, the case $p=2$ of which is due to [4] and the case $p\in(1,2]$ of which is due (via a different proof) to [58].

THEOREM 1.3. *Let $f\colon\mathbb{R}^3\to\mathbb{R}$ be smooth and compactly supported. Then, for every $p\in(1,2]$,*

$$\left(\int_0^\infty\left(\int_{\mathbb{R}^3}|D_{\mathsf{v}}^t f(h)|^p\,dh\right)^{2/p}\frac{dt}{t}\right)^{1/2}\lesssim\frac{1}{\sqrt{p-1}}\left(\int_{\mathbb{R}^3}(|\mathsf{X}f(h)|^p+|\mathsf{Y}f(h)|^p)\,dh\right)^{1/p}. \qquad (1.5)$$

See [58] for a variant of Theorem 1.3 when $p>2$. The pertinent point of comparison to (1.3) is as $p\to1^+$, namely there is a jump discontinuity at the endpoint $p=1$.

It should be noted that the dependence on p in the right-hand side of (1.5) is not specified in [58], but one obtains (1.5) in the form stated above by tracking the dependence on p in the proof of [58]; we explain how to do so in Appendix A below. We conjecture that the following bound holds, which is better than (1.5) only in terms of the dependence on p; its geometric ramifications will be derived later (see Remark 1.15), at which point it will become clear why we need to record an explicit (power-type) dependence as $p\to1^+$ in (1.5), rather than using the implicit $\lesssim_p$ notation as done in [58].

Conjecture 1.4. In the setting of Theorem 1.3, we have

$$\left(\int_0^\infty\left(\int_{\mathbb{R}^3}|D_{\mathsf{v}}^t f(h)|^p\,dh\right)^{2/p}\frac{dt}{t}\right)^{1/2}\lesssim\frac{1}{\sqrt[4]{p-1}}\left(\int_{\mathbb{R}^3}(|\mathsf{X}f(h)|^p+|\mathsf{Y}f(h)|^p)\,dh\right)^{1/p}. \qquad (1.6)$$

Another key point of comparison between Theorem 1.1 and the literature is with its higher-dimensional counterpart due to [91]. For a smooth function $f\colon\mathbb{R}^5\to\mathbb{R}$, denote in analogy to (1.1) and (1.2) for every $h=(x_1,y_1,x_2,y_2,z)\in\mathbb{R}^5$ and $t\in(0,\infty)$,

$$\mathsf{X}_1f(h)\stackrel{\mathrm{def}}{=}\frac{\partial f}{\partial x_1}(h)-\frac{1}{2}y_1\frac{\partial f}{\partial z}(h),\qquad \mathsf{X}_2f(h)\stackrel{\mathrm{def}}{=}\frac{\partial f}{\partial x_2}(h)-\frac{1}{2}y_2\frac{\partial f}{\partial z}(h),$$

$$\mathsf{Y}_1f(h)\stackrel{\mathrm{def}}{=}\frac{\partial f}{\partial y_1}(h)+\frac{1}{2}x_1\frac{\partial f}{\partial z}(h),\qquad \mathsf{Y}_2f(h)\stackrel{\mathrm{def}}{=}\frac{\partial f}{\partial y_2}(h)+\frac{1}{2}x_2\frac{\partial f}{\partial z}(h),$$

and

$$D_{\mathsf{v}}^t(h)\stackrel{\mathrm{def}}{=}\frac{f(x_1,y_1,x_2,y_2,z+t)-f(h)}{\sqrt{t}}.$$

We then have the following theorem (it holds with $\mathbb{R}^5$ replaced mutatis mutandis by $\mathbb{R}^{2k+1}$ for every $k\geqslant 2$; we are focusing only on $\mathbb{R}^5$, because the crucial qualitative difference that we establish here is between dimension 3 and all the larger odd dimensions).

THEOREM 1.5. *Let $f\colon\mathbb{R}^5\to\mathbb{R}$ is smooth and compactly supported. Then, for every $p\in[1,2]$,*

$$\begin{aligned}&\left(\int_0^\infty\left(\int_{\mathbb{R}^5}|D_{\mathsf{v}}^t f(h)|^p\,dh\right)^{2/p}\frac{dt}{t}\right)^{1/2}\\ &\quad\lesssim\left(\int_{\mathbb{R}^5}(|\mathsf{X}_1f(h)|^p+|\mathsf{Y}_1f(h)|^p+|\mathsf{X}_2f(h)|^p+|\mathsf{Y}_2f(h)|^p)\,dh\right)^{1/p}.\end{aligned} \qquad (1.7)$$

The case $p=2$ of Theorem 1.5 is from [4], and in the range $p\in(1,2]$ the bound (1.7) but with $\lesssim$ replaced by $\lesssim_p$ is from [58]. The case $p=1$ of Theorem 1.5 is from [91]. Inequality (1.7) as stated above, i.e., with the right-hand side multiplied by a universal constant rather than a constant that depends on p as in [58], follows by interpolating between the cases $p=1$ and $p=2$ of [91] and [4], respectively. Indeed, (1.7) asserts the boundedness of a linear operator, the $L_2(L_p)$ norms in the left-hand side of (1.7) are an interpolation family by classical interpolation theory [11], and the Sobolev $W^{1,p}$ norms in the right-hand side of (1.7) are an interpolation family by [5, Theorem 8.8].

1.1. Geometric implications

Let $\mathbb{H}$ be the 3-dimensional Heisenberg group with real coefficients. As a set, $\mathbb{H}$ is identified with $\mathbb{R}^3$, and the group structure on $\mathbb{H}$ is given by

$$gh\stackrel{\mathrm{def}}{=}\left(x+\chi,y+\upsilon,z+\zeta+\tfrac{1}{2}(x\upsilon-y\chi)\right)\quad\text{for all } g=(x,y,z),h=(\chi,\upsilon,\zeta)\in\mathbb{R}^3. \qquad (1.8)$$

The identity element is $\mathbf{0}=(0,0,0)$ and the inverse of $g=(x,y,z)$ is $g^{-1}=(-x,-y,-z)$. The center of $\mathbb{H}$ is $\{0\}\times\{0\}\times\mathbb{R}$ and if we let $\mathbb{H}_{\mathbb{Z}}$ be the discrete subgroup of $\mathbb{H}$ that is generated by $(1,0,0)$ and $(0,1,0)$, then we have

$$\mathbb{H}_{\mathbb{Z}}=\left\{\left(x,y,z+\tfrac{1}{2}xy\right):x,y,z\in\mathbb{Z}\right\}\subseteq\mathbb{Z}\times\mathbb{Z}\times\tfrac{1}{2}\mathbb{Z}.$$

Let $d_W:\mathbb{H}_{\mathbb{Z}}\times\mathbb{H}_{\mathbb{Z}}\to\mathbb{N}\cup\{0\}$ be the left-invariant word metric on $\mathbb{H}_{\mathbb{Z}}$ that is induced by the symmetric set of generators $\{(-1,0,0),(1,0,0),(0,-1,0),(0,1,0)\}$. It is well known (and elementary to verify) that, for every $g=(x,y,z),h=(\chi,\upsilon,\zeta)\in\mathbb{H}_{\mathbb{Z}}$, we have

$$d_W(g,h)\asymp|x-\chi|+|y-\upsilon|+\sqrt{|2z-2\zeta-x\upsilon+y\chi|}. \tag{1.9}$$

In fact, an exact formula for $d_W(g,h)$, which directly implies (1.9), is derived in [13]. For every $n\in\mathbb{N}$, denote the word-ball of radius n centered at the identity element by

$$\mathcal{B}_n\stackrel{\mathrm{def}}{=}\{g\in\mathbb{H}_{\mathbb{Z}}:d_W(g,\mathbf{0})\leqslant n\}. \tag{1.10}$$

1.1.1. Embeddings

Recall that a metric space (M,d_M) is said to admit a bi-Lipschitz embedding into a Banach space $(X,\|\cdot\|_X)$ if there exist $D\in[1,\infty)$ and $\phi: M\to X$ such that

$$d_M(x,y)\leqslant\|\phi(x)-\phi(y)\|_X\leqslant Dd_M(x,y)\quad\text{for all } x,y\in M. \tag{1.11}$$

The infimum over those $D\in[1,\infty)$ for which this holds is called the *X-distortion* of M and is denoted by $\mathsf{c}_X(M)$. If no such D exists, then one writes $\mathsf{c}_X(M)=\infty$.

Theorem 1.6 below is a sharp asymptotic evaluation of $\mathsf{c}_{\ell_1}(\mathcal{B}_n)$. It answers a question posed in [63], [20]–[23], [81], [96], [58]; these references ask for the asymptotic evaluation of $\mathsf{c}_{\ell_1}(\mathcal{B}_n)$, but most of them also conjecture that $\mathsf{c}_{\ell_1}(\mathcal{B}_n)\asymp\sqrt{\log n}$, so Theorem 1.6 constitutes both a resolution of an open problem, and an unexpected answer. The fact that $\lim_{n\to\infty}\mathsf{c}_{\ell_1}(\mathcal{B}_n)=\infty$ is due to [20], the previously best known upper bound [3] was $\mathsf{c}_{\ell_1}(\mathcal{B}_n)\lesssim\sqrt{\log n}$ and the previously best-known lower bound [23] was $\mathsf{c}_{\ell_1}(\mathcal{B}_n)\geqslant(\log n)^\delta$ for some positive but very small universal constant δ; thus both the upper and the lower bounds of Theorem 1.6 are new.

THEOREM 1.6. *$\mathsf{c}_{\ell_1}(\mathcal{B}_n)\asymp\sqrt[4]{\log n}$ for every integer $n\geqslant 2$.*

In contrast, the word-ball of radius $n\geqslant 2$ in the 5-dimensional Heisenberg group has ℓ_1-distortion of order $\sqrt{\log n}$; this was proved in [91] using Theorem 1.5.

The statement of Theorem 1.6 has two parts. While the lower bound

$$\mathsf{c}_{\ell_1}(\mathcal{B}_n)\gtrsim\sqrt[4]{\log n}$$

is framed above as a "negative result" (impossibility of embedding), it encapsulates a "positive result," namely the aforementioned new structural information on surfaces in $\mathbb{H}$, to which most of this article is devoted. The upper bound

$$\mathsf{c}_{\ell_1}(\mathcal{B}_n) \lesssim \sqrt[4]{\log n}$$

is a "positive result," namely a new geometric realization of $\mathcal{B}_n$, but we will soon see that it has ramifications for counterexamples to natural geometric questions.

The estimate (1.3) of Theorem 1.1 implies the lower bound $\mathsf{c}_{\ell_1}(\mathcal{B}_n) \gtrsim \sqrt[4]{\log n}$. In fact, such *vertical-versus-horizontal Poincaré inequalities* were originally envisaged as obstructions to embeddings of $\mathcal{B}_n$ into various spaces; see [4], [86], [58], [90], and most pertinently §3 of [91], where we treated such matters in greater generality than what is needed here; in particular, for any $p \geqslant 1$, if every compactly supported smooth function $f: \mathbb{R}^3 \to \mathbb{R}$ satisfies the inequality

$$\left(\int_0^\infty \left(\int_{\mathbb{R}^3} |D_{\mathsf{v}}^t f(h)| \, dh \right)^p \frac{dt}{t} \right)^{1/p} \lesssim \int_{\mathbb{R}^3} (|\mathsf{X} f(h)| + |\mathsf{Y} f(h)|) \, dh, \tag{1.12}$$

then by [91, § 3] and the reasoning in [91, § 1.3] we have $\mathsf{c}_{\ell_1}(\mathcal{B}_n) \gtrsim (\log n)^{1/p}$.

Thus, $\mathsf{c}_{\ell_1}(\mathcal{B}_n) \gtrsim \sqrt[4]{\log n}$, since Theorem 1.1 asserts that (1.12) holds for $p=4$. This also demonstrates that the matching upper bound $\mathsf{c}_{\ell_1}(\mathcal{B}_n) \lesssim \sqrt[4]{\log n}$ of Theorem 1.6 implies the second assertion of Theorem 1.1, namely the optimality of the $L_4(dt/t)$ norm in the left-hand side of (1.3). Here we prove the following more refined embedding statement which we formulate as a separate theorem because it has further noteworthy applications.

THEOREM 1.7. *For every $\vartheta \geqslant \frac{1}{4}$ and every integer $n \geqslant 2$ there exists $\phi = \phi_{n,\vartheta}: \mathbb{H}_{\mathbb{Z}} \to \ell_1$ with respect to which every two points $g=(x,y,z), h=(\chi,\upsilon,\zeta) \in \mathbb{H}_{\mathbb{Z}}$ with $d_W(g,h) \leqslant 2n$ satisfy*

$$\|\phi(g)-\phi(h)\|_{\ell_1} \asymp |x-\chi| + |y-\upsilon| + \frac{\sqrt{|2z-2\zeta-x\upsilon+y\chi|}}{(\log n)^\vartheta}. \tag{1.13}$$

By (1.9) and the case $\vartheta = \frac{1}{4}$ of Theorem 1.7, the following weakening of (1.13) holds.

$$\frac{d_W(g,h)}{\sqrt[4]{\log n}} \lesssim \|\phi(g)-\phi(h)\|_{\ell_1} \lesssim d_W(g,h) \quad \text{for all } g,h \in \mathcal{B}_n.$$

So, the upper bound $\mathsf{c}_{\ell_1}(\mathcal{B}_n) \lesssim \sqrt[4]{\log n}$ of Theorem 1.6 follows from Theorem 1.7. However, Theorem 1.7 is of further use thanks to the following embedding result of [57]. At present, the fact that both our embedding and that of [57] yield the same expression (up to universal constant factors) for the metric in the image seems to be a fortunate and consequential coincidence; it would be valuable, if possible, to explain conceptually why those formulas coincided (e.g. is this inevitable due to underlying symmetries?).

Theorem 1.8. *For any $p>2$, any $\vartheta\geqslant 1/p$ and any integer $n\geqslant 2$ there is*

$$\psi=\psi_{n,p,\vartheta}:\mathbb{H}_{\mathbb{Z}}\longrightarrow \ell_p$$

such that every $g=(x,y,z), h=(\chi,\upsilon,\zeta)\in\mathbb{H}_{\mathbb{Z}}$ with $d_W(g,h)\leqslant 2n$ satisfy

$$\|\psi(g)-\psi(h)\|_{\ell_p}\asymp |x-\chi|+|y-\upsilon|+\frac{\sqrt{|2z-2\zeta-x\upsilon+y\chi|}}{(\log n)^{\vartheta}}. \tag{1.14}$$

Theorem 1.8 is not formulated explicitly in [57], but it is a direct consequence of [57, Lemma 3.1] combined with the finite-determinacy theorem of [93], which together imply that, for every $\varepsilon\in(0,\frac{1}{2}]$, there exists an embedding $\sigma=\sigma_{\varepsilon,p}:\mathbb{H}\to\ell_p$ for which every $g=(x,y,z), h=(\chi,\upsilon,\zeta)\in\mathbb{H}_{\mathbb{Z}}$ satisfy

$$\|\sigma(g)-\sigma(h)\|_{\ell_p}\asymp |x-\chi|^{1-\varepsilon}+|y-\upsilon|^{1-\varepsilon}+\varepsilon^{1/p}|2z-2\zeta-x\upsilon+y\chi|^{(1-\varepsilon)/2}. \tag{1.15}$$

(Without reference to [93], Lemma 3.1 in [57] asserts the existence of such an embedding into L_p rather than into ℓ_p.) To derive Theorem 1.8 from (1.15), let $\pi:\mathbb{H}\to\mathbb{R}$ be the map that is given by setting $\pi(x,y,z)=(x,y)$ for $(x,y,z)\in\mathbb{H}$ and choose

$$\varepsilon=\frac{1}{\log n}\quad\text{and}\quad \psi=\frac{\sigma}{(\log n)^{\vartheta-1/p}}\oplus\pi:\mathbb{H}_{\mathbb{Z}}\longrightarrow\ell_p\oplus\mathbb{R}^2\cong\ell_p. \tag{1.16}$$

1.1.2. Aspects of the Ribe program

Inspired by a fundamental rigidity theorem of [102] and first put forth in [16], the Ribe program is a web of conjectures and analogies whose goal is to transfer linear phenomena in the geometry of Banach spaces to questions about metric spaces, where Lipschitz mappings take the role of bounded linear operators; see e.g. the surveys [51], [82], [7], [94], [84]. We will next explain how the above results answer natural questions in this area.

Theorem 1.9 below follows from Theorems 1.7 and 1.8, and from [4], [58]. It answers a longstanding question in metric embedding theory; even though (to the best of our knowledge) this question never appeared in published(2) texts, it was a folklore open problem. To briefly explain the context, the classical work [50] (together with a differentiation argument of [68]) implies that, for $1\leqslant p<r<q<\infty$, if a Banach space X admits a bi-Lipschitz embedding into both L_p and L_q, then X also admits a bi-Lipschitz

(2) We have seen it appear in writing only in grant proposals, and it was posed verbally among experts. In particular, we are indebted to Gideon Schechtman for valuable discussions on this matter over the years.

embedding into L_r. The case $r=2$ of this statement is that if X embeds into L_p for two finite values of p that lie on both sides of 2, then X must embed into (hence, by [31], be linearly isomorphic to) a Hilbert space; a different proof of the latter statement, as a special case of a much more general phenomenon, follows from [54]. In light of these facts about the geometry of Banach spaces, one is naturally led to ask if a metric space M that embeds bi-Lipschitzly into L_p for two finite values of p that lie on both sides of 2 must admit a bi-Lipschitz embedding into a Hilbert space.

Theorem 1.9. *For any $2<p\leqslant 4$ there is a metric space M that admits a bi-Lipschitz embedding into ℓ_1 and into ℓ_r for all $r\geqslant p$, yet M does not admit a bi-Lipschitz embedding into L_q for any $1<q<p$. More generally, M does not admit a bi-Lipschitz embedding into a Banach space whose modulus of uniform convexity has power type q, for $2\leqslant q<p$.*

For the statement of Theorem 1.9, recall that a Banach space $(X,\|\cdot\|)$ has modulus of uniform convexity of power type q if there is $C>0$ such that the sharpened triangle inequality $\|x+y\|\leqslant 2-C\|x-y\|^q$ holds for any unit vectors $x,y\in X$. By [24], [42], for $1<q<\infty$ any $L_q(\mu)$ space has modulus of uniform convexity of power type $\max\{q,2\}$.

Proof of Theorem 1.9 assuming Theorems 1.7 and 1.8. For every $n\in\mathbb{N}$, we define $M_n=\phi_{n,\vartheta}(\mathcal{B}_n)\subseteq\ell_1$, where $\phi_{n,\vartheta}$ is as in Theorem 1.7 applied with $\vartheta=1/p\geqslant\frac14$.

By considering the union of sufficiently widely-spaced translations in ℓ_1 of the finite sets $\{M_n\}_{n=1}^\infty$, we see that there is $M\subseteq\ell_1$ such that $\sup_{n\in\mathbb{N}}\mathsf{c}_M(M_n)<\infty$.

For every $r\geqslant p$, consider $\psi_{n,r,\vartheta}(\mathcal{B}_n)\subseteq\ell_r$, where $\psi_{n,r,\vartheta}$ is as in Theorem 1.8. Theorems 1.7 and 1.8 show that $\psi_{n,r,\vartheta}(\mathcal{B}_n)$ is bi-Lipschitz equivalent with $O(1)$ distortion to M_n. Hence, by considering a suitable union of translations in ℓ_r of the finite sets $\{\psi_{n,r,\vartheta}(\mathcal{B}_n)\}_{n=1}^\infty$, we see that $\mathsf{c}_{\ell_r}(M)<\infty$. Let X be a Banach space whose modulus of uniform convexity has power type q, for $2\leqslant q<p$. By [58], we have

$$(\log n)^{1/q}\lesssim_X \mathsf{c}_X(\mathcal{B}_n)\lesssim(\log n)^{\vartheta}\mathsf{c}_X(M_n)=(\log n)^{1/p}\mathsf{c}_X(M_n),$$

where the penultimate step holds because, due to (1.13), M_n and $\mathcal{B}_n$ are bi-Lipschitz equivalent with distortion $O((\log n)^{\vartheta})$. Therefore, since $q<p$,

$$\mathsf{c}_X(M_n)\gtrsim_X(\log n)^{1/q-1/p}\xrightarrow[n\to\infty]{}\infty.$$

Hence, $\mathsf{c}_X(M)=\infty$, as required. For future reference we record in passing that we obtained the following bound when $X=L_q$ and $1<q<p$:

$$\mathsf{c}_{L_q}(M_n)\gtrsim_q(\log n)^{1/\max\{q,2\}-1/p}. \tag{1.17}$$

Note that the bound in [4], which is asymptotically weaker than that of [58], suffices for the qualitative conclusion $\mathsf{c}_X(M)=\infty$ of Theorem 1.9. The above estimates seem to be the best that one could achieve using available methods; it would be very interesting to determine the optimal behavior, e.g. if an n-point metric space W embeds with $O(1)$ distortion into ℓ_1 and also into ℓ_p for some $p>2$, how large can $\mathsf{c}_{\ell_2}(W)$ be? □

Remark 1.10. With more care, it is possible to ensure that the metric space M of Theorem 1.9 is a left-invariant metric $\delta=\delta_p$ on $\mathbb{H}_{\mathbb{Z}}$; see Theorem 3.2. Concretely, for $p=4$, the metric δ_4 can be taken to satisfy the following bounds for any $(a,b,c)\in\mathbb{H}_{\mathbb{Z}}$ with $|c|\geqslant 3$:

$$\delta_4(\mathbf{0},(a,b,c))\asymp |a|+|b|+\frac{\sqrt{|c|}}{\sqrt[4]{\log|c|}\cdot(\log\log|c|)^2}.$$

By the reasoning in [87, §9], since $\mathbb{H}_{\mathbb{Z}}$ is amenable, it follows that $(\mathbb{H}_{\mathbb{Z}},\delta)$ admits a bi-Lipschitz embedding into L_1 and L_r, for all $r\geqslant p$, which is also equivariant (with respect to an action of $\mathbb{H}_{\mathbb{Z}}$ on, respectively, L_1 and L_p by affine isometries); we did not investigate if this holds for equivariant embeddings into the sequence spaces ℓ_1 and ℓ_r.

The natural question how the embeddability of a group into L_p depends on p was also studied in the literature; see [18], [25], and especially the recent solution of this question in [69], where it is proved that the phenomenon of Theorem 1.9 does not hold for equivariant coarse embeddings (namely, for such embeddings the corresponding set of p is always an interval). Note that for coarse embeddings that need not be equivariant, the statement of [69] was previously known as a direct consequence of [74, Remark 5.10] (from here, using [87], one gets the full equivariant statement of [69] for amenable groups). Theorem 3.2 shows that the situation is markedly different if one considers bi-Lipschitz embeddings rather than coarse embeddings.

The following question arises naturally from Theorem 1.9 and seems quite difficult.

Question 1.11. For a metric space M, how complicated can the following set be?

$$\{1\leqslant p<\infty:\mathsf{c}_{L_p}(M)<\infty\}.$$

Theorem 1.9 leaves the possibility that there is better behavior in the reflexive range, i.e., that if a metric space M embeds bi-Lipschitzly into ℓ_p and ℓ_q for $1<q<2<p<\infty$, then M embeds bi-Lipschitzly into a Hilbert space. If true, this would be an excellent theorem, but due to Theorem 1.9 we speculate that the answer is negative. A substantial new idea seems to be needed here. Less ambitiously, does the above assumption (even allowing $q=1$) imply that M embeds into a Hilbert space with finite average distortion (see [85] for the relevant definition)? Does this imply that every n-point subset of M embeds into a Hilbert space with bi-Lipschitz distortion $o(\log n)$, i.e., asymptotically better than the distortion that is guaranteed by the general embedding theorem of [15]?

The above reasoning also leads to Theorem 1.12 below, which answers another natural question arising in the Ribe program, on the factorization of Lipschitz functions.

We first briefly make preparatory observations that will be also useful elsewhere. Recall that for $K\in\mathbb{N}$ a metric space X is said to be K*-doubling* if, for every $r>0$, any ball $B\subseteq X$ of radius r can be covered by K balls of radius $\frac{1}{2}r$. A metric space X is said to be *doubling* if it is K-doubling for some $K\in\mathbb{N}$. The metric space M of Theorem 1.9 can be taken to be doubling. Indeed, fix $p>2$ and $n\in\mathbb{N}$. As in the proof of Theorem 1.9, write $\vartheta=1/\min\{p,4\}$. It was shown in [57] that $\psi_{n,p,\vartheta}(\mathbb{H}_\mathbb{Z})$ is a $O(1)$-doubling subset of ℓ_p. Let $S\subseteq\ell_p$ be the disjoint union of translates in ℓ_p of the finite sets $\{\psi_{n,p,\vartheta}(\mathcal{B}_n)\}_{n=1}^\infty$ that are sufficiently widely-spaced so as to ensure that S is a doubling subset of ℓ_p, and $\sup_{n\in\mathbb{N}}\mathsf{c}_S(M_n)<\infty$. As in the proof of Theorem 1.9, using Theorem 1.7 we get an embedding $\varphi\colon S\to\ell_1$ satisfying

$$\|\varphi(x)-\varphi(y)\|_{\ell_1}\asymp\|x-y\|_{\ell_p}$$

for all $x,y\in S$. Thus, $\varphi(S)=M$ is a doubling subset of ℓ_1.

Since S is doubling, by [62] we can extend φ to a Lipschitz function $f\colon\ell_p\to\ell_1$. If there were Lipschitz mappings $g\colon\ell_p\to\ell_2$ and $h\colon g(\ell_p)\to\ell_1$ such that $f=h\circ g$, then it would follow that, for all $x,y\in S$,

$$\|x-y\|_{\ell_p}\asymp\|\varphi(x)-\varphi(y)\|_{\ell_1}=\|h(g(x))-h(g(y))\|_{\ell_1}\lesssim\|g(x)-g(y)\|_{\ell_2}\lesssim\|x-y\|_{\ell_p}.$$

Therefore, $g\circ\varphi^{-1}$ would be a bi-Lipschitz embedding of M into ℓ_2, which we proved above was impossible. We thus arrive at the following statement.

THEOREM 1.12. *For any $2<p<\infty$ there is a Lipschitz mapping $f\colon\ell_p\to\ell_1$ that cannot be factored through a subset of a Hilbert space using Lipschitz mappings. Namely, there do not exist Lipschitz mappings $g\colon\ell_p\to\ell_2$ and $h\colon g(\ell_p)\to\ell_1$ such that $f=h\circ g$. More generally, f cannot be factored using Lipschitz mappings through a subset of a Banach space whose modulus of uniform convexity has power type q for $2\leqslant q<\min\{4,p\}$.*

By [66, Theorem 5.2], for $p\geqslant 2$ any linear operator from ℓ_p to ℓ_1 factors through ℓ_2 (the factorization is via linear operators, though by [46] this is equivalent to factorization using Lipschitz functions as above). Theorem 1.12 demonstrates that there is no analogue of this factorization phenomenon for Lipschitz mappings.

Such investigations arose in the Ribe program in the seminal work [45], which had a major influence on the subsequent fruitful efforts by many mathematicians in search of metric analogues of the extension and factorization paradigm of [72]. This search is itself intimately intertwined with the search for metric theories of type and cotype.

We refer to the survey [73] for an exposition of the powerful and deep theory of type and cotype of Banach spaces; it suffices to say here that one can define linear invariants of Banach spaces that are called *type* 2 and *cotype* 2, such that L_p has type 2 if $2\leqslant p<\infty$ and cotype 2 if $1\leqslant p\leqslant 2$, and such that the following extension and factorization phenomenon [72] holds.

Suppose that Y is a Banach space of type 2 and that Z is a Banach space of cotype 2. Let X be a linear subspace of Y, and let $\tau: X\to Z$ be a bounded linear operator. Then, there exist a bounded linear operator $T: Y\to Z$ that extends τ, a Hilbert space H and bounded linear operators $A: Y\to H$ and $B: A(Y)\to Z$ with $T=BA$.

Paper [45] raised the question of when the analogous statement holds in the metric setting. Namely, now Y and Z are metric spaces, X is an arbitrary subset of Y, $f: X\to Z$ is a Lipschitz mapping, and we ask for the same extension and factorization through a Hilbert space H, i.e., to establish the existence of Lipschitz mappings $F: Y\to Z$, $\alpha: Y\to H$, and $\beta: \alpha(Y)\to Z$, such that the following diagram commutes:

$$\begin{tikzcd} Y \arrow[r, "\alpha"] \arrow[dr, "F"] & \alpha(Y) \arrow[r, hook, "\subseteq"] \arrow[d, "\beta"] & H \\ X \arrow[u, hook, "\subseteq"] \arrow[r, "f"] & Z. & \end{tikzcd} \tag{1.18}$$

An implicit but central part of this endeavor encompasses the important issue of how to define useful notions of type 2 and cotype 2 for metric spaces so that, at the very least, ℓ_p has type 2 for $2\leqslant p<\infty$ and cotype 2 for $1\leqslant p\leqslant 2$. Clearly (1.18) has two components. The first is if f admits the Lipschitz extension F. The second is if F can be factored through a subset of a Hilbert space. While these questions come hand-in-hand in the linear theory of [72] (see also [98]), they are different issues in the metric setting.

The main focus of [45] was the Lipschitz extension problem, so it highlighted the first component above. At the time, the metric version of the extension problem was a bold and speculative question, but [6] introduced metric notions of type 2 and cotype 2 and obtained a powerful extension result for maps from spaces of Markov type 2 to spaces of Markov cotype 2. Combined with [88], this provides a quite satisfactory understanding of the extension component of (1.18) when the target space is ℓ_p, $1<p<2$. However, this understanding is currently confined to the reflexive range, and the question remains a major open problem when the target space is ℓ_1 (see [52], [77] for a partial negative answer, and [67] for an intriguing algorithmic reformulation).

In contrast to the achievement of [6], Theorem 1.12 demonstrates that there is *no* way to define notions of type 2 and cotype 2 for metric spaces so that any map from a space of type 2 to a space of cotype 2 factors through Hilbert space and such that ℓ_p has type 2 when $2<p<\infty$ and cotype 2 when $p=1$. Though this resolves the factorization question when the target is ℓ_1, it remains a fascinating open problem to see if a factorization theory analogous to [6] can be developed when the target is ℓ_q for $1<q<2$.

It is instructive to examine the dual interpretation of Theorem 1.12. Just as the dual formulation of the linear factorization and extension problems was key to [72], duality also plays an important role in the non-linear theory. The duality lemma that was found in [6] for Lipschitz extension(3) does not shed light on Lipschitz factorization, but the factorization issue was broached in [32], [19]. One can deduce from [19] the following factorization criterion. Given $\Phi>0$, metric spaces (X,d_X) and (Z,d_Z), and $f\colon X\to Z$, there exist a Hilbert space H and a factorization $f=\beta\circ\alpha$ for some Lipschitz mappings

$$\alpha\colon X\longrightarrow H \quad \text{and} \quad \beta\colon \alpha(X)\longrightarrow Z$$

with $\|\alpha\|_{\mathrm{Lip}}\|\beta\|_{\mathrm{Lip}}\leqslant\Phi$ if and only if, for all $n\in\mathbb{N}$ and $x_1,...,x_n\in X$, any two symmetric stochastic matrices $\mathsf{A}=(a_{ij}),\mathsf{B}=(b_{ij})\in\mathsf{M}_n(\mathbb{R})$ such that $\mathsf{A}-\mathsf{B}$ is positive semidefinite satisfy the following quadratic inequality:

$$\sum_{i=1}^n\sum_{j=1}^n a_{ij}d_Z(f(x_i),f(x_j))^2\leqslant\Phi^2\sum_{i=1}^n\sum_{j=1}^n b_{ij}d_X(x_i,x_j)^2. \tag{1.19}$$

Theorem 1.12 yields the first example of a Lipschitz mapping $f\colon\ell_p\to\ell_1$ for $2<p<\infty$ that fails to satisfy (1.19) for any $\Phi>0$, despite the fact that if f were a linear operator, then by [72] it would automatically satisfy (1.19) with $\Phi\lesssim_p\|f\|_{\mathrm{Lip}}$.

Remark 1.13. Another counterexample to the non-linear version of [72] arises from an embedding of the Laakso graphs into a non-classical Banach space. Let $\{\Lambda_n\}_{n=1}^\infty$ be the Laakso graphs [55], [56], indexed so that $|\Lambda_n|=n$; these are series-parallel (hence planar) graphs that are $O(1)$-doubling when equipped with their shortest-path metric.

On one hand, the Laakso graphs do not admit a bi-Lipschitz embedding into a Hilbert space. In fact, by [55], [59], we have $\mathsf{c}_{\ell_2}(\Lambda_n)\gtrsim\sqrt{\log n}$ (this is sharp by the general embedding theorem of [101]). Moreover, by [75], for every uniformly convex Banach space X we have $\lim_{n\to\infty}\mathsf{c}_X(\Lambda_n)=\infty$.

On the other hand, by [41], we have $\sup_{n\in\mathbb{N}}\mathsf{c}_{\ell_1}(\Lambda_n)<\infty$, and by [47], we have $\sup_{n\in\mathbb{N}}\mathsf{c}_Y(\Lambda_n)<\infty$ when Y is a Banach space that is not reflexive. By considering

(3) Quoting what [6] says about this crucial duality step: "*This lemma is a variant of one used by Maurey. A related lemma was found earlier by Johnson, Lindenstrauss and Schechtman: their result actually characterises extensions which factor through subsets of Hilbert space, a problem much closer to Maurey's argument. Their lemma provided much of the stimulus for the present work.*" Unfortunately, it seems that the work of Johnson, Lindenstrauss and Schechtman that is mentioned in [6] was never published.

translates of the images of the embeddings in ℓ_1 that are sufficiently widely spaced, we obtain a doubling subset $\Lambda\subseteq\ell_1$ such that $\mathsf{c}_Y(\Lambda)<\infty$ for any non-reflexive Banach space Y and $\mathsf{c}_X(\Lambda)=\infty$ for any uniformly convex Banach space X.

By [44], there exists a Banach space $\mathbb{J}$ that has type 2, yet $\mathbb{J}$ is not reflexive; a different construction of such a Banach space was found in [100]. So, Λ embeds bi-Lipschitzly into both the cotype-2 space ℓ_1 and the type-2 space $\mathbb{J}$, yet not into a Hilbert space. This is impossible in the linear setting; by [54] a Banach space of type 2 and cotype 2 is isomorphic to a Hilbert space (this is a far reaching generalization of the aforementioned consequence of [50] that motivates Theorem 1.9). This reasoning also produces a stronger asymptotic estimate than (1.17), since $\mathsf{c}_{\ell_2}(\Lambda_n)\gtrsim\sqrt{\log n}$, but it cannot shed light on the ℓ_p setting of (1.17), because it relies precisely on the non-reflexivity of $\mathbb{J}$ (through the use of [47]) to deduce that $\sup_{n\in\mathbb{N}}\mathsf{c}_{\mathbb{J}}(\Lambda_n)<\infty$.

The Laakso graphs also lead to a counterexample to the metric version of [72]. Let $\varphi\colon\Lambda\to\mathbb{J}$ be a bilipschitz embedding. Since Λ is a doubling subset of ℓ_1, one can use [62] to construct a Lipschitz map $f\colon\ell_1\to\mathbb{J}$ that extends φ. As above, f cannot factor through a Hilbert space (or even through any uniformly convex Banach space X) by Lipschitz maps, because such a factorization would produce a bilipschitz embedding of Λ into a Hilbert space (respinto X).

This discussion shows that if one is allowed to replace ℓ_p in Theorems 1.9 and 1.12 by non-classical (indeed, "exotic" and hard to come by) Banach spaces such as $\mathbb{J}$, then it is possible to demonstrate the failure of the metric space version of [72] and its important precursor [54] using well-known examples.

Part of the impetus for the search for definitions of metric space notions of type 2 and cotype 2 was the hope of obtaining a metric version of the theorem of [54], but it was well known to experts that the metric definitions of type 2 and cotype 2 found over the past decades are not suitable for this purpose (see e.g. the discussion in [30]). The above discussion demonstrates conclusively that it is impossible to define metric space notions of type 2 and cotype 2 that are bi-Lipschitz invariant, pass to subsets, coincide for Banach spaces with type 2 and cotype 2, and for which [54] holds for doubling metric spaces, i.e., any doubling space that has both type 2 and cotype 2 admits a bi-Lipschitz embedding into a Hilbert space (the corresponding statement with Λ replaced by a metric space that is not doubling follows by using [16] instead of the Laakso graphs in the above reasoning; in fact, by [9], the infinite binary tree embeds bilipschitzly into both ℓ_1 and $\mathbb{J}$, but not into a Hilbert space). Theorem 1.9 shows that this is so, even if one restricts attention to subsets of ℓ_p for $p>2$.

1.1.3. Dimension reduction

By a highly influential lemma of [45], any finite subset S of a Hilbert space embeds with bi-Lipschitz distortion $O(1)$ into a k-dimensional Hilbert space for $k\lesssim\log|S|$; see [84] for an indication of the significance of this statement. The question whether this phenomenon holds with Hilbert space replaced by ℓ_1 was a prominent open problem until it was resolved negatively in [17], where it was shown that, for arbitrarily large $n\in\mathbb{N}$, there is an n-point subset D_n of ℓ_1 such that, if D_n embeds with bi-Lipschitz distortion $O(1)$ into ℓ_1^k, then necessarily $k\geqslant n^c$ for some universal constant $c>0$. In [60] it was shown that D_n can be taken to be $O(1)$-doubling, and in [89] it was shown that ℓ_1^k can be replaced by an arbitrary k-dimensional subspace of the Schatten–von Neumann trace class S_1; both of these enhancements hold without changing the conclusion (other than perhaps values of universal constants).

The examples $\{D_n\}_{n=1}^\infty$ of [17] are the diamond graphs [92], while their aforementioned doubling counterparts in [60] are the Laakso graphs $\{\Lambda_n\}_{n=1}^\infty$ that we discussed in Remark 1.13. By [75] and [47], we have

$$\sup_{n\in\mathbb{N}} \mathsf{c}_X(D_n)=\sup_{n\in\mathbb{N}} \mathsf{c}_X(\Lambda_n)=\infty$$

for every uniformly convex Banach space X. In fact, by [47] the converse of this statement holds true (though we do not need it below), namely X admits an equivalent uniformly convex norm if and only if $\sup_{n\in\mathbb{N}}\mathsf{c}_X(D_n)=\infty$ or $\sup_{n\in\mathbb{N}}\mathsf{c}_X(\Lambda_n)=\infty$. Theorem 1.14 below obtains new examples that demonstrate the failure of dimension reduction in ℓ_1 à la [45], which are qualitatively different than the previously known examples, since our examples do admit a bi-Lipschitz embedding into a uniformly convex Banach space (specifically, into ℓ_p for any $p>2$). At present, this comes with a worse lower bound on the target dimension, but see Remark 1.15 below which explains how Conjecture 1.4 would remedy this (for the very same example that we consider in Theorem 1.14).

THEOREM 1.14. *There is a universal constant $c>0$ with the following property. For all $n\in\mathbb{N}$ and $2<p\leqslant 4$ there exists a $O(1)$-doubling subset $\mathcal{H}_n=\mathcal{H}_n(p)$ of ℓ_1 with $|\mathcal{H}_n|\leqslant n$ such that*

$$\mathsf{c}_{\ell_q}(\mathcal{H}_n)\lesssim 1 \quad \text{for all } q\geqslant p,$$

and for every $D\geqslant 1$, if X is a finite-dimensional subspace of the Schatten–von Neumann trace class S_1 for which $\mathsf{c}_X(\mathcal{H}_n)\leqslant D$, then necessarily

$$\dim(X)\geqslant \exp\Big(\frac{c}{D^2}(\log n)^{1-2/p}\Big). \tag{1.20}$$

In the statement of Theorem 1.14, recall that for $p\geqslant 1$ the Schatten–von Neumann trace class S_p is the Banach space of all the compact operators $T\colon \ell_2\to\ell_2$ that satisfy

$$\|T\|_{\mathsf{S}_p}\stackrel{\mathrm{def}}{=}(\mathbf{Trace}[(T^*T)^{p/2}])^{1/p}<\infty.$$

Note that ℓ_p is the subspace of S_p consisting of the diagonal operators. Thus, the dimension reduction lower bound (1.20) holds in particular for any subspace X of ℓ_1.

The proof of Theorem 1.14 is short (modulo previously stated results and the available literature), so we present the quick derivation now instead of postponing it to a later section; it mimics the reasoning of [61] while combining it with [58], Theorems 1.7 and 1.8, as well as structural information on subspaces of S_1 from [89].

Proof of Theorem 1.14. By (1.9), we have $|\mathcal{B}_m|\asymp m^4$ for all $m\in\mathbb{N}$. So, fix $m\in\mathbb{N}$ with $m\asymp\sqrt[4]{n}$ such that $n\lesssim|\mathcal{B}_m|\leqslant n$. Using the mapping $\phi_{m,1/p}\colon\mathbb{H}_{\mathbb{Z}}\to\ell_1$ of Theorem 1.7, define

$$\mathcal{H}_n\stackrel{\mathrm{def}}{=}\phi_{m,1/p}(\mathcal{B}_m).$$

By combining Theorems 1.7 and 1.8, we indeed have $\mathsf{c}_{\ell_q}(\mathcal{H}_n)\lesssim 1$ for all $q\geqslant p$.

Let X be a finite-dimensional subspace of S_1. Fix $1<r\leqslant 2$ whose value will be specified later so as to optimize the ensuing reasoning. By [89, Theorem 12], we have([4])

$$\mathsf{c}_{\mathsf{S}_r}(X)\leqslant\dim(X)^{1-1/r}.$$

Hence, if $\mathsf{c}_X(\mathcal{H}_n)\leqslant D$, then, since $\mathsf{c}_{\mathcal{H}_n}(\mathcal{B}_m)\lesssim(\log n)^{1/p}$ by Theorem 1.7, we have

$$\mathsf{c}_{\mathsf{S}_r}(\mathcal{B}_m)\lesssim(\log n)^{1/p}\mathsf{c}_{\mathsf{S}_r}(\mathcal{H}_n)\leqslant(\log n)^{1/p}D\mathsf{c}_{\mathsf{S}_r}(X)\leqslant(\log n)^{1/p}D\dim(X)^{1-1/r}.$$

At the same time, by [58], we have([5])

$$\mathsf{c}_{\mathsf{S}_r}(\mathcal{B}_m)\gtrsim\sqrt{(r-1)\log n},$$

so we conclude that

$$\inf_{1<r\leqslant 2}\frac{\dim(X)^{1-1/r}}{\sqrt{r-1}}\gtrsim\frac{(\log n)^{1/2-1/p}}{D}.$$

This gives the desired bound (1.20) by choosing $r-1\asymp 1/\log(\dim(X))$. □

([4]) If one only wishes to rule out embeddings into low-dimensional subspaces of ℓ_1 rather than of S_1, then it suffices to use here [65, Theorem 1.2], which yields an embedding into ℓ_r rather than S_r.

([5]) As in the discussion before Conjecture 1.4, the dependence on r in this estimate is not stated in [58], while it is crucial for us here; a justification why the reasoning in [58] implies this appears in Appendix A.

Remark 1.15. By substituting (1.6) into the reasoning of [58], a positive resolution of Conjecture 1.4 would imply that, for every $r\in(1,2]$ and $n\in\mathbb{N}$,

$$\mathsf{c}_{\ell_r}(\mathcal{B}_n)\gtrsim\sqrt[4]{r-1}\cdot\sqrt{\log n}. \tag{1.21}$$

An incorporation of this improved distortion lower bound into the above proof of Theorem 1.14 (while using [65] in place of [89], since we are in the simpler ℓ_p setting) would imply that, for any finite-dimensional subspace X of ℓ_1, if $\mathsf{c}_X(\mathcal{H}_n(p))\leqslant D$, then the following improvement over (1.20) holds true:

$$\dim(X)\geqslant\exp\Big(\frac{c}{D^4}(\log n)^{2-4/p}\Big). \tag{1.22}$$

Notably, for $p=4$ this would be an improvement from

$$\dim(X)\geqslant\exp\Big(\frac{c}{D^2}\sqrt{\log n}\Big)$$

to

$$\dim(X)\geqslant n^{c/D^4}, \tag{1.23}$$

namely a power-type dimension reduction lower bound as in [17]. Understanding what is the correct behavior as $p\to 2^+$ remains an intriguing open question; some deterioration of the lower bound as in (1.20) or (1.22) must occur because by [45] logarithmic dimension reduction is possible for finite subsets of a Hilbert space.

Another question that this discussion obviously raises is if (1.21) could be enhanced to

$$\mathsf{c}_{\mathsf{S}_r}(\mathcal{B}_n)\gtrsim\sqrt[4]{r-1}\cdot\sqrt{\log n}. \tag{1.24}$$

If so, then (1.23) would hold when X is a subspace of S_1 rather than ℓ_1. More substantially, this would resolve a difficult open question (see the discussion following Question 13 in [91]) by showing that $\mathbb{H}_{\mathbb{Z}}$ does not admit a bi-Lipschitz embedding into S_1. In fact, for the latter conclusion it would suffice to establish the weaker property

$$\lim_{n\to\infty}\mathsf{c}_{\mathsf{S}_{1+1/\log n}}(\mathcal{B}_n)=\infty. \tag{1.25}$$

Indeed, by [89] we have $\mathsf{c}_{\mathsf{S}_1}(\mathcal{B}_n)\gtrsim\mathsf{c}_{\mathsf{S}_r}(\mathcal{B}_n)$ when $r=1+1/\log n$. Due to its significant consequences, we expect that proving (1.25), and all the more so its stronger version (1.24), would require a major and conceptually new idea.

We end this discussion on dimension reduction by noting that [104] shows that one could embed $\mathcal{B}_n$ with optimal distortion (up to universal constant factors) into a Euclidean space of dimension $O(1)$. Theorem 1.14 shows that this fails badly if one aims for optimal ℓ_1-distortion embedding of $\mathcal{B}_n$ into a bounded-dimensional subspace of ℓ_1.

1.1.4. Permanence of compression rates of groups

Suppose that (M, d_M) is a metric and $(X, \|\cdot\|_X)$ is a Banach space. The *compression rate* of a Lipschitz mapping $f\colon M\to X$ is the non-decreasing function $\omega_f\colon [0,\infty)\to[0,\infty)$ that is defined [39] by

$$\omega_f(s) \stackrel{\mathrm{def}}{=} \inf_{\substack{x,y\in M\\ d_M(x,y)\geqslant s}} \|f(x)-f(y)\|_X \quad \text{for all } s\geqslant 0. \tag{1.26}$$

Equivalently, ω_f is the largest non-decreasing function from $[0,\infty)$ to $[0,\infty)$ such that

$$\|f(x)-f(y)\|_X \geqslant \omega_f(d_M(x,y)) \quad \text{for all } x,y\in M.$$

There is a great deal of interest in determining the largest possible compression rate of 1-Lipschitz mappings from a finitely generated group G (equipped with a word metric that is induced by some finite generating set) to certain Banach spaces, notable and useful examples of which are Hilbert space and L_1. The literature on this topic is too extensive to discuss here, and we only mention that a substantial part of it is devoted to understanding the extent to which compression rates are preserved under various group operations (e.g. various semidirect products). Theorem 1.16 below provides a new example of the lack of such permanence which does not seem to be accessible using previously available methods. It leverages the fact that we establish here a marked difference between the L_1 embeddability of Heisenberg groups of dimension 3 and dimension 5.

Theorem 1.16. *There exists a finitely group G that has two finitely generated normal subgroups $H, K\lhd G$ such that the following properties hold true.*

(1) *Any $h\in H$ and $k\in K$ commute.*

(2) *$H\cap K$ is the center of G.*

(3) *H and K are isomorphic.*

(4) *H and K are undistorted in G; in fact, they admit generating sets S_H and S_K such that $S_H\cup S_K$ generates G and the word metric on G that is induced by $S_H\cup S_K$ restricts to the word metrics on H and K that are induced by S_H and S_K, respectively.*

(5) *The L_1 compression of G is asymptotically smaller than that of H (hence also of $K\cong H$). Concretely, there exists a Lipschitz mapping $f\colon H\to \ell_1$ that satisfies*

$$\omega_f(s) \gtrsim \frac{s}{\sqrt[4]{\log s}\cdot(\log\log s)^2} \quad \textit{for all } s\geqslant 3, \tag{1.27}$$

yet for any Lipschitz mapping $F\colon G\to L_1$ there are arbitrarily large $s\geqslant 4$ for which

$$\omega_F(s) \leqslant \frac{s}{\sqrt{(\log s)\log\log s}}. \tag{1.28}$$

Proof. Let $G_{\mathbb{R}}$ be the 5-dimensional Heisenberg group, i.e., $\mathbb{R}^5$ with the group operation

$$(x_1,y_1,x_2,y_2,z)(x_1',y_1',x_2',y_2',z') \\ = \big(x_1+x_1',y_1+y_1',x_2+x_2',y_2+y_2',z+z'+\tfrac{1}{2}(x_1y_1'+x_2y_2'-y_1x_1'-y_2x_2')\big)$$

for $(x_1,y_1,x_2,y_2,z),(x_1',y_1',x_2',y_2',z')\in\mathbb{R}^5$. Let G be the 5-dimensional integer Heisenberg group, which is the subgroup

$$G=\big\{\big(x_1,y_1,x_2,y_2,z+\tfrac{1}{2}(x_1y_1+x_2y_2)\big):x_1,x_2,y_1,y_2,z\in\mathbb{Z}\big\}.$$

The subgroups H and K are natural copies of $\mathbb{H}_{\mathbb{Z}}$ in G, namely

$$H=\{(x_1,y_1,x_2,y_2,z)\in G:x_2=y_2=0\},$$
$$K=\{(x_1,y_1,x_2,y_2,z)\in G:x_1=y_1=0\}.$$

One directly checks the first four assertions of Theorem 1.16. The bound (1.27) follows by considering the mapping $f:\mathbb{H}_{\mathbb{Z}}\to\ell_1(\ell_1)\cong\ell_1$ that is given by

$$f\stackrel{\mathrm{def}}{=}\bigoplus_{n=1}^{\infty}\frac{1}{n^2}\phi_{2^{2^n},1/4},$$

where the mappings that are being concatenated are those of Theorem 1.7. The final assertion (1.28) of Theorem 1.16 follows from [91, Theorem 9]. □

Remark 1.17. The term $\log\log s$ in (1.27) and (1.28) can be improved slightly; for (1.27) this follows by examining the above proof, and for (1.28) this is explained by [91, Theorem 9]. However, some unbounded lower-order correction is necessary in (1.27) for the specific groups that we used in the proof of Theorem 1.16; see Remark 3.3.

Obviously, Theorem 1.16 raises the question if a similar phenomenon could occur for embeddings into a Hilbert space rather than into L_1. Also, in Theorem 1.16 the compression rate of the subgroups H and K grows roughly (suppressing lower-order factors) like $s/\sqrt[4]{\log s}$ as $s\to\infty$, while the compression rate of G grows slower than $s/\sqrt{\log s}$. What are the possible asymptotic profiles of the compression rates that exhibit such phenomena?

1.2. Decomposing surfaces into approximately ruled pieces

In the previous sections, we discussed consequences of Theorem 1.1 (and the refined version of its second part in Theorem 1.7). In this section, we will give an overview of the concepts involved in the proof of Theorem 1.1, especially our main contribution, which is a new way to describe the structure of surfaces in $\mathbb{H}$.

The statement of Theorem 1.1 is in terms of smooth functions $f:\mathbb{H}\to\mathbb{R}$, but the main bound (1.3) has an equivalent formulation in terms of surfaces in $\mathbb{H}$; see (1.32) below. We will prove it by showing that surfaces in $\mathbb{H}$ admit a multi-scale hierarchical decomposition into pieces that are close to ruled surfaces (unions of horizontal lines) and that most of these pieces (in a quantitative sense) are long and narrow, giving the decomposition the appearance of a Venetian blind with many narrow slats; see Figures 2 and 3 for examples. For reasons that will be clarified soon, we call the above structure a *foliated corona decomposition*. This decomposition is conceptually central to this work, and the most involved part of this paper is to formulate this decomposition, prove its existence, and demonstrate its utility for the aforementioned applications (more are forthcoming).

The defining feature of this decomposition is that its pieces, which we call *pseudoquads*, have widely varying aspect ratios. Each pseudoquad is roughly rectangular, and we define the *aspect ratio* of a pseudoquad to be its width divided by its height; long, narrow rectangles have large aspect ratios, while tall, skinny rectangles have small aspect ratios. The fact that the pieces of the decomposition (the slats of the Venetian blind) can have unbounded aspect ratios allows the decomposition to have additional symmetries and ultimately leads to the exponent 4 in Theorem 1.1.

Specifically, in order to work with long, narrow pieces, we must prove results on the geometry of $\mathbb{H}$ that are invariant not only under the usual scaling automorphisms, but also under automorphisms that stretch and shear $\mathbb{H}$. The resulting automorphism-invariant bounds allow us to produce a decomposition that is likewise invariant under rescaling, stretching, and shearing. Furthermore, the overlap of the pieces of our decomposition is controlled by a coercive quantity that scales like the fourth power of the aspect ratio under automorphisms. This leads to a new *weighted Carleson packing condition* in which overlaps are normalized by the fourth power of the aspect ratio; this condition leads directly to the exponent 4 in the bound (1.3) of Theorem 1.1.

Proving the optimality of Theorem 1.1 entails finding a surface for which (1.32) is sharp. Part of the construction of such a surface can be seen in Figure 3. The surface in the figure can be viewed as a surface with a foliated corona decomposition for which the weighted Carleson packing condition is sharp. For this reason, it is pedagogically beneficial to describe that construction after describing foliated corona decompositions. In truth, the general decomposition methodology and the construction that demonstrates its optimality are intertwined: limitations of such a construction indicate what decomposition to look for. We therefore suggest to also consider the alternative route of first examining the construction of the specific (sharp) example prior to considering the task of decomposing general surfaces; the proofs in the rest of this article follow the latter ("reverse") route as this leads to a more gradual introduction of notation and concepts.

The ensuing considerations belong firmly to the setting of the continuous Heisenberg group and its Carnot–Carathéodory geometry. They therefore assume some familiarity with notions from that setting; the pertinent background appears in §2 below.

1.2.1. Fractal Venetian blinds abound

In what follows, for any $s>0$ the Hausdorff measure $\mathcal{H}^s$ on $\mathbb{H}$ will be with respect to the Carnot–Carathéodory metric d on $\mathbb{H}$. We denote the standard generators of $\mathbb{H}$ by $X=(1,0,0), Y=(0,1,0)$, and $Z=(0,0,1)$.

For $\Omega\subseteq\mathbb{H}$ and $a\in\mathbb{R}$, consider the symmetric difference

$$\mathsf{D}_a\Omega \stackrel{\mathrm{def}}{=} \Omega\triangle\Omega Z^{2^{-2a}} = \big(\Omega\setminus\Omega Z^{2^{-2a}}\big)\cup\big(\Omega Z^{2^{-2a}}\setminus\Omega\big) \tag{1.29}$$

If $\Omega, U\subseteq\mathbb{H}$ are measurable, then, following [58], [91], we define $\bar{\mathsf{v}}_U(\Omega):\mathbb{R}\to\mathbb{R}$ by

$$\bar{\mathsf{v}}_U(\Omega)(a) \stackrel{\mathrm{def}}{=} 2^a\mathcal{H}^4(U\cap\mathsf{D}_a\Omega) = 2^a\int_U \big|\mathbf{1}_\Omega(h)-\mathbf{1}_\Omega(hZ^{-2^{-2a}})\big|\,d\mathcal{H}^4(h) \quad \text{for all } a\in\mathbb{R}. \tag{1.30}$$

Thus, $\bar{\mathsf{v}}_U(\Omega)(a)$ is a (normalized) measurement of the amount that Ω changes within U when translated up and down by the specified (Carnot–Carathéodory) distance 2^{-a}.

By [91, Lemma 38], in order to prove the first part of Theorem 1.1, namely inequality (1.3) for any compactly supported smooth function $f:\mathbb{H}\to\mathbb{R}$, it suffices to prove that every measurable subset $\Omega\subseteq\mathbb{H}$ satisfies the following isoperimetric-type inequality:(6)

$$\|\bar{\mathsf{v}}_{\mathbb{H}}(\Omega)\|_{L_4(\mathbb{R})} \lesssim \mathcal{H}^3(\partial\Omega). \tag{1.31}$$

This amounts in essence to an application of the coarea formula (e.g. [1]).

A central step of [91] is a further reduction of (1.31) to the special case that Ω is (a piece of) an *intrinsic Lipschitz epigraph*. An intrinsic Lipschitz epigraph Γ^+ is a region of $\mathbb{H}$ that is bounded by an intrinsic Lipschitz graph Γ. The notion of an intrinsic

(6) The exponent 4 is not important here, i.e., [91] shows that, for any $q\geqslant 1$, if

$$\|\bar{\mathsf{v}}_{\mathbb{H}}(\Omega)\|_{L_q(\mathbb{R})} \lesssim \mathcal{H}^3(\partial\Omega)$$

holds for every measurable subset $\Omega\subseteq\mathbb{H}$, then

$$\bigg(\int_0^\infty\bigg(\int_{\mathbb{R}^3}|D_{\mathsf{v}}^t f(h)|\,dh\bigg)^q\frac{dt}{t}\bigg)^{1/q} \lesssim \int_{\mathbb{R}^3}(|\mathsf{X}f(h)|+|\mathsf{Y}f(h)|)\,dh$$

holds for every compactly supported smooth function $f:\mathbb{R}^3\to\mathbb{R}$.

Lipschitz graph was introduced in [36] and all of the relevant background is explained in §2.2 below. The intrinsic Lipschitz condition is parameterized by an intrinsic Lipschitz constant $\lambda\in(0,1)$. By combining [91, Proposition 55, Theorem 57 and Lemma 58] (see the deduction on [91, p. 232]) it follows that, to prove (1.31), it suffices to show that for every $0<\lambda<1$ the vertical perimeter of any intrinsic λ-Lipschitz epigraph $\Gamma^+\subseteq\mathbb{H}$ satisfies the growth bound

$$\|\bar{\mathrm{v}}_{B_r(\mathbf{0})}(\Gamma^+)\|_{L_4(\mathbb{R})} \lesssim_\lambda r^3 \quad \text{for all } r>0, \tag{1.32}$$

where $B_r(\mathbf{0})$ is the (Carnot–Carathéodory) ball of radius r centered at $\mathbf{0}=(0,0,0)$.([7])

The structural information that underlies the reduction of (1.31) to (1.32) is that, for any $0<\lambda<1$, any (sufficiently nice; see [91] for precise assumptions) surface in $\mathbb{H}$ has a multi-scale hierarchical decomposition into pieces that are close to intrinsic λ-Lipschitz graphs, and moreover that decomposition has controlled overlap in the sense that it satisfies a $O(1)$-Carleson packing condition. As such, this decomposition is an intrinsic Heisenberg analog of the *corona decompositions* that were introduced and developed for subsets of Euclidean space in [26], and have since led to a variety of powerful applications in harmonic analysis (see also the monograph [27]).

The corona decomposition of [91] is in some respects a Heisenberg variant of a "vanilla" corona decomposition. Like corona decompositions in $\mathbb{R}^n$, it is a hierarchical partition of a surface into pieces of bounded aspect ratio, and the Carleson packing condition governing overlaps of pieces depends only on the diameter of the pieces. Nevertheless, there are key differences, including the fact that the proof in [91] relies on a new "stopping rule" (based on the quantitative non-monotonicity of [23]) that yields, in fact, a different proof of the existence of corona decompositions even in Euclidean space (though, for less general sets than those that [26] treats). In addition, while "vanilla" Euclidean corona decompositions cover a surface in $\mathbb{R}^n$ by pieces that are approximately graphs of Lipschitz functions, the approximating graphs in [91] are intrinsic Lipschitz, like the surface depicted in Figure 1. While Lipschitz graphs in Euclidean space vary slowly in all directions, intrinsic Lipschitz graphs vary slowly in horizontal directions but can vary quickly in vertical directions and can have Hausdorff dimension 2.5 with respect to the Euclidean metric [53]. This can make these graphs difficult to analyze, and even after the decomposition step of [91], the challenge of establishing estimates such as (1.32) remains.

In [91], we addressed this challenge for the 5-dimensional Heisenberg group $\mathbb{H}^5$, but our techniques do not shed light on the 3-dimensional setting of Theorem 1.1. An

([7]) Also this deduction in [91] does not rely on the specific value of the exponent 4. Namely, for any $q\geqslant 1$, if for every $0<\lambda<1$, every intrinsic λ-Lipschitz epigraph $\Gamma^+\subseteq\mathbb{H}$ satisfies $\|\bar{\mathrm{v}}_{B_r(\mathbf{0})}(\Gamma^+)\|_{L_q(\mathbb{R})}\lesssim_\lambda r^3$ for every $r>0$, then $\|\bar{\mathrm{v}}_{\mathbb{H}}(\Omega)\|_{L_q(\mathbb{R})}\lesssim\mathcal{H}^3(\partial\Omega)$ holds for every measurable subset $\Omega\subseteq\mathbb{H}$.

intrinsic Lipschitz graph in $\mathbb{H}^5$ is the intrinsic graph of a function ψ that is defined on a 4-dimensional vertical hyperplane V_0. An inspection of the intrinsic Lipschitz condition shows that the restriction of ψ to any coset of $\mathbb{H}$ that is contained in V_0 is Lipschitz with respect to the Carnot–Carathéodory metric on $\mathbb{H}$. In [91], we applied a representation-theoretic functional inequality of [4] to each of these restrictions, yielding a bound on the vertical variation of ψ. The desired control on the vertical perimeter of intrinsic Lipschitz graphs in $\mathbb{H}^5$ followed by integrating this bound over the cosets of $\mathbb{H}$ in V_0.

In the 3-dimensional setting of the present work, the intrinsic graph Γ in (1.32) corresponds to an intrinsic Lipschitz function $\psi\colon V_0 \to \mathbb{R}$, where V_0 is a 2-dimensional vertical plane in $\mathbb{H}$. For concreteness, assume in what follows that $V_0 = \{(x, 0, z) : x, z \in \mathbb{R}\}$ is the xz-plane. The reasoning of [91] is irrelevant to proving (1.32): one cannot restrict ψ to cosets of a lower-dimensional Heisenberg group, as there is no such group!

Our strategy here is therefore entirely different from that of [91]. We will prove (1.32) by finding a new structural description of intrinsic Lipschitz graphs in $\mathbb{H}$. Specifically, we will prove that they admit a hierarchical family of partitions into pieces that are approximately ruled surfaces and bound the total error of these approximations.

We call this description of Γ a *foliated corona decomposition.* It is a sequence of nested partitions of Γ into approximately rectangular regions, called *pseudoquads*, of varying heights and widths. On each pseudoquad, Γ is close to a vertical plane, and these vertical planes can be glued together to form a collection of ruled surfaces such that at most locations and scales, Γ is approximated by one of the ruled surfaces; see Remark 7.6. Furthermore, the decomposition satisfies a new *weighted* variant of the classical *Carleson packing condition.* Namely, we bound the weighted sum of the measures of the pseudoquads in the decomposition, where the measure of each pseudoquad is normalized by the fourth power of its aspect ratio. We will see that the occurrence of the fourth power here is *dictated* by the requirement that this decomposition should be invariant under certain automorphisms of $\mathbb{H}$ (scaling, stretch, and shear automorphisms).

THEOREM 1.18. *Any intrinsic Lipschitz graph in $\mathbb{H}$ has a foliated corona decomposition.*

The above description of foliated corona decompositions and the statement of Theorem 1.18 clearly lack rigorous definitions, but they convey the essence of what is achieved here. The necessary technical matters are treated in §5 below, where a precise formulation of Theorem 1.18 appears as Theorem 5.2. The justification that Theorem 1.18 can be used to achieve our goal (1.32) is carried out in §6 below; the groundwork of constructing a foliated corona decomposition makes this deduction quite mechanical.

We will next cover a few technical details necessary to describe foliated corona decompositions and the subdivision mechanism that produces them. Recall that $V_0 \subseteq \mathbb{H}$ is the xz-plane. Fix $0<\lambda<1$ and let Γ be an intrinsic λ-Lipschitz graph that is the intrinsic graph of $\psi\colon V_0 \to \mathbb{R}$. That is, $\Gamma = \Psi(V_0)$, where $\Psi(v) = vY^{\psi(v)}$ for all $v \in V_0$. The function ψ satisfies the intrinsic Lipschitz condition (Definition 2.2); the non-linear nature of this condition is the source of subtleties that ensue (and the reason why basic questions on the rectifiability properties of intrinsic Lipschitz graphs remain open; see e.g. [29]).

For any $p \in \Gamma$, there is a horizontal curve γ contained in Γ that passes through p, so Γ is the union of all such curves. It is often convenient to work in V_0 instead of Γ. To this end, let $\Pi\colon \mathbb{H} \to V_0$ be the projection to V_0, so $\Pi(\Psi(v)) = v$ for $v \in V_0$. The projected curve $\Pi \circ \gamma$ is a curve in V_0 which we call a *characteristic curve*; see §2.3 for a detailed discussion. Parameterize γ so that $\Pi(\gamma(t)) = (t, 0, g(t))$ for some continuous function g. This function is a solution of the differential equation $g'(t) = -\psi(t, 0, g(t))$, and conversely, each solution gives a characteristic curve. If Γ is a vertical plane, then $\psi(x, 0, z) = ax + b$ for some $a, b \in \mathbb{R}$, in which case the characteristic curves are parallel parabolas.

Since horizontal curves pass through every point of Γ, there is a characteristic curve through every point of V_0, so one can reconstruct Γ from its set of characteristic curves. Note that the characteristic curve through p is not necessarily unique: when ψ is not smooth, these curves can split and rejoin [12]. When ψ is smooth, the characteristic curves foliate V_0, so there is a coordinate system on V_0 such that the foliation forms one set of coordinate lines. However, it is difficult to use this coordinate system to study the geometry of Γ because the distance between two characteristic curves can vary wildly. Foliated corona decompositions provide a way to overcome this difficulty.

A *pseudoquad* for Γ is a region in V_0 that is bounded by characteristic curves above and below, and by vertical line segments on either side. We call a pseudoquad Q *rectilinear* if its top and bottom boundaries approximate two parallel parabolas; if the top and bottom boundaries of Q are exactly two parallel parabolas, we call Q a *parabolic rectangle*. Parabolic rectangles are the projections to V_0 of rectangles in $\mathbb{H}$ bounded by two horizontal line segments and two vertical line segments. The width $\delta_x(Q)$ and height $\delta_z(Q)$ of such a pseudoquad are defined to be, respectively, the width and height of its approximating parabolic rectangle; see §4. The aspect ratio of Q is

$$\alpha(Q) = \frac{\delta_x(Q)}{\sqrt{\delta_z(Q)}}.$$

Let $Q_0 \subseteq V_0$ be a rectilinear pseudoquad. A foliated corona decomposition for Γ with root at Q_0 is a sequence of nested partitions of Q_0 into rectilinear pseudoquads. We construct such a decomposition using the following *subdivision algorithm* which, importantly, outputs pseudoquads that can be divided into two sets $\mathcal{V}_{\mathsf{V}}$ and $\mathcal{V}_{\mathsf{H}}$, called, respectively, the *vertically cut* pseudoquads and *horizontally cut* pseudoquads. The algorithm repeatedly cuts pseudoquads into halves. Let Q be a pseudoquad in the decomposition. If $\Psi(Q)$ is a region in Γ that is sufficiently close to a vertical plane V_Q and if the characteristic curves through Q are close to characteristic curves for V_Q, then cut Q in half along one of the characteristic curves of Γ. In this case, say that Q is horizontally cut and add it to $\mathcal{V}_{\mathsf{H}}$. Otherwise, cut Q in half along a vertical line through its center, say that Q is vertically cut, and add it to $\mathcal{V}_{\mathsf{V}}$. By applying this procedure iteratively, we obtain a sequence of nested partitions of Q_0; see Figure 2.

A crucial part of the algorithm is the mechanism determining whether to cut the pseudoquad horizontally or vertically. We stated qualitatively how this step depends on the geometry of $\Psi(Q)$, but we implement it quantitatively by introducing a coercive quantity called *R-extended non-monotonicity.* This is a family of measures $\Omega^P_{\Gamma^+,R}$ on the vertical plane V_0, parameterized by $R>0$; see §8. These are inspired by the quantitative non-monotonicity of [23], but there are key differences. For instance, while the non-monotonicity of Γ on a subset $U\subseteq\mathbb{H}$ measures how lines intersect Γ inside U, the R-extended non-monotonicity of Γ on a subset $W\subseteq V_0$ measures how lines intersect Γ inside an R-neighborhood of $\Psi(W)$. We refer to §8 for the details, in particular to Lemma 9.2, which shows that, for any measurable $U\subseteq V_0$,

$$\sum_{i\in\mathbb{Z}}\Omega^P_{\Gamma^+,2^{-i}}(U)\lesssim_\lambda |U|, \tag{1.33}$$

where $|U|$ is the area of U and λ is the intrinsic Lipschitz constant of ψ.

Analogously to [23], extended non-monotonicity is coercive in the following sense. Let $U=[0,1]\times\{0\}\times[0,1]\subseteq V_0$ and for $r>0$, let rU be the square of side r concentric with U. There is a universal constant $r>1$ such that if δ is sufficiently small, R is sufficiently large, $\psi(0)$ is bounded, and $\Omega_{\Gamma^+,R}(rU)<\delta$, then $\Psi(U)$ is close to a vertical plane and the characteristic curves that pass through U are close to characteristic curves of that vertical plane (i.e., parabolas). The proof of this geometric statement (whose precise formulation appears as Proposition 7.2) is the most technically involved part of this work; it is outlined in §10 and carried out in §11 and §12.

By translation, rescaling, and applying a shear automorphism, a similar coercive property applies to any pseudoquad of aspect ratio 1, but for the subdivision algorithm, we need a coercive property for pseudoquads of arbitrary aspect ratio. If Q is a pseudoquad of aspect ratio $\alpha(Q)$, the *stretch automorphism* $s(x,y,z)=(\alpha(Q)^{-1}x,\alpha(Q)y,z)$ sends Q to a pseudoquad of aspect ratio 1. The extended non-monotonicity of $s(Q)$ scales like $\alpha(Q)^4$, so if the extended non-monotonicity of Γ^+ on Q is at most $\delta|Q|/\alpha(Q)^4$, then $\Psi(Q)$ is close to a vertical plane and the characteristic curves that pass through Q are close to characteristic curves of that vertical plane.

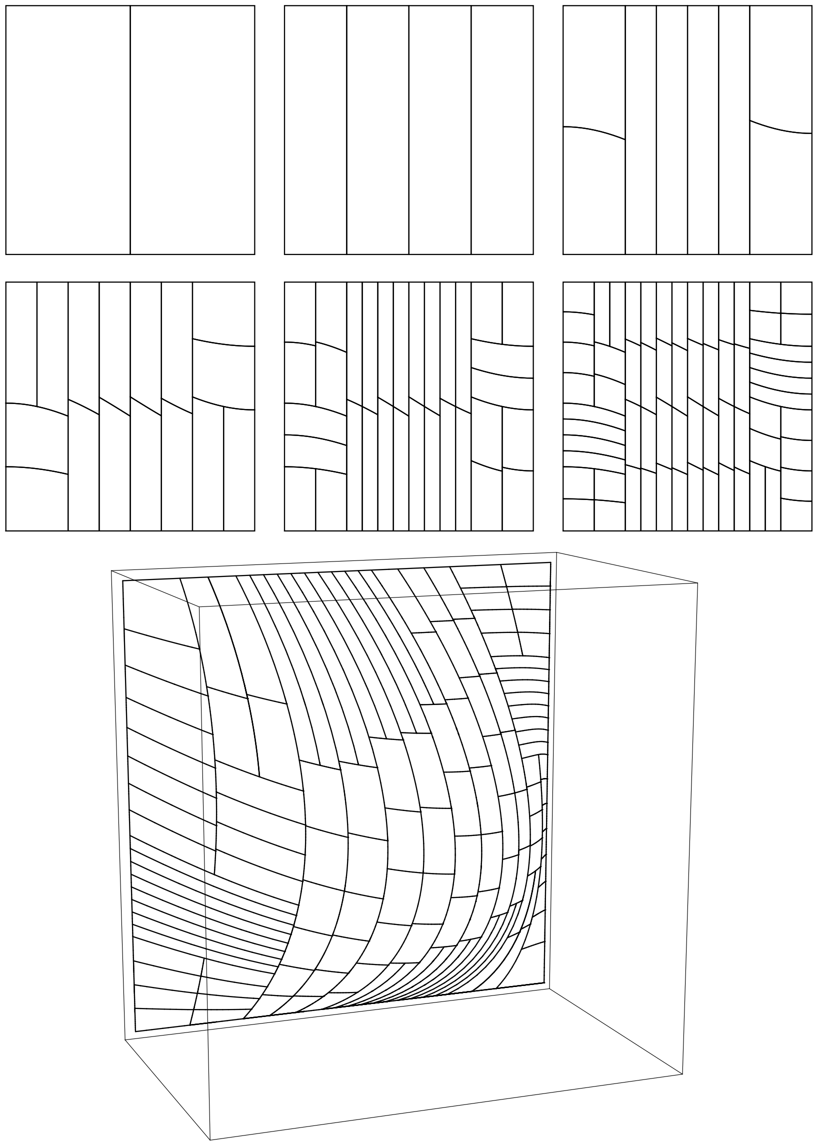

Figure 2. Stages in the construction of a foliated corona decomposition for a bump function as in the top row of Figure 3 below. The aspect ratio of the regions in the decomposition varies widely. On the sides, where the surface is close to a vertical plane, the aspect ratio is large and the regions are short and wide; near the top and bottom, where it is further from a plane, the regions are tall and narrow.

Therefore, in the subdivision algorithm above, there is $\delta>0$ such that we cut Q horizontally if and only if the extended non-monotonicity of Γ^+ on Q is at most $\delta|Q|/\alpha(Q)^4$. This criterion, combined with (1.33), leads to a crucial bound on the total pseudoquads that have been vertically cut by the subdivision algorithm. Specifically, if Q is a pseudoquad of the decomposition and $\mathcal{D}_{\mathsf{V}}(Q)$ is the set of vertically cut pseudoquads Q' in the decomposition that are contained in Q, then

$$\sum_{Q'\in\mathcal{D}_{\mathsf{V}}(Q)} \frac{|Q'|}{\alpha(Q')^4} \lesssim_\lambda |Q|. \tag{1.34}$$

The condition (1.34) is the aforementioned weighted Carleson packing condition, and the L_4 norm that appears in Theorem 1.1 arises directly from the exponent 4 in (1.34).

Thus, the L_4 norm in Theorem 1.1 is ultimately dictated by having to prove a coercive property for intrinsic Lipschitz graphs that is invariant under stretch automorphisms. This stretch-invariance has multiple effects. On one hand, stretch-invariance means that it suffices to prove the coercive property for pseudoquads of aspect ratio 1; indeed, it is enough to consider pseudoquads that approximate the unit square. On the other hand, it induces a substantial complication in the proofs: since the intrinsic Lipschitz constant is not invariant under stretch automorphisms, the coercivity must be independent of the intrinsic Lipschitz constant.

1.2.2. A maximally bumpy surface

The optimality part of Theorem 1.1 corresponds to constructing (in §3) an intrinsic Lipschitz graph for which the $L_4(\mathbb{R})$ norm in (1.32) cannot be replaced by the $L_q(\mathbb{R})$ norm for any $0<q<4$. Theorem 1.7 is deduced in §3.1 by analyzing this construction; the level sets of the resulting embedding into L_1 are a superposition of certain random rotations, scalings and translations of this surface.

We will show that, for any sufficiently small $\varepsilon>0$, there are intrinsic Lipschitz surfaces in $\mathbb{H}$ of bounded (Heisenberg) perimeter that are ε-far from planes at ε^{-4} different scales, many more than the ε^{-2} different scales that are possible (by [91]) for such surfaces in the 5-dimensional Heisenberg group $\mathbb{H}^5$ (or, for that matter, in $\mathbb{R}^n$, by the Jones travelling salesman theorem [48] and the higher-dimensional analogues thereof [26]).

We construct these surfaces by adding bumps to a vertical plane. While surfaces that demonstrate that the bound of [91] for $\mathbb{H}^5$ is optimal can be constructed by adding round bumps with equal width and height, it is more natural in $\mathbb{H}$ to add oblong bumps with width (horizontal size) w, depth d (size perpendicular to the surface), and height h (vertical size). The automorphisms of the Heisenberg group preserve the ratio dw/h, so we can construct a family of bump functions by applying automorphisms to a prototype bump with $d=w=h=1$. The resulting bumps have $h=dw$, and we define the aspect ratio α of such a bump to be

$$\alpha=\frac{w}{\sqrt{h}}=\frac{w}{\sqrt{dw}}=\sqrt{\frac{w}{d}}.$$

A horizontal curve connecting one side of the bump to its other side has slope roughly $d/w=\alpha^{-2}$, so adding a layer of bumps with aspect ratio $\alpha\geqslant 1$ to a surface multiplies its perimeter by roughly $1+\alpha^{-4}$. Thus, we can start with a unit square, then add ε^{-4} layers of bumps of width $\varepsilon^{-1}r_i$, depth εr_i, and height r_i^2, for $r_1\gg\ldots\gg r_{\varepsilon^{-4}}$. These bumps all have aspect ratio ε^{-1}, so the resulting surface Σ has bounded perimeter, and for any $x\in\Sigma$, the intersections $B_{r_i}(x)\cap\Sigma$ are each εr_i-far away from any plane. So, Σ is ε-far from planes at ε^{-4} different scales. The implementation of this strategy in §3 is in essence an example of a foliated corona decomposition. At each stage, we use the characteristic curves of the surface that was obtained in the previous stage to guide us where to glue the next layer of bumps. Figure 3 shows a sketch of the construction.

It is highly informative to examine why this construction does not work in $\mathbb{H}^5$. Bumps on a surface in $\mathbb{H}^5$ have five dimensions, which we denote w_1, w_2, d_1, d_2, and h, so that h is vertical, the other four dimensions are horizontal, and d_2 is normal to the surface. The automorphisms of $\mathbb{H}^5$ preserve the ratios $d_1w_1/(d_2w_2)$, d_1w_1/h, and d_2w_2/h. If β is a bump with $d_1w_1=d_2w_2=h$ and $d_2\leqslant w_2$, then the slopes of β in the three horizontal directions are roughly d_2/w_1, d_2/w_2, and d_2/d_1. So, adding β to a vertical rectangle with dimensions $w_1\times w_2\times d_1\times h$ increases the volume of the rectangle by a factor of roughly

$$\nu(w_1,w_2,d_1,d_2,h)\stackrel{\mathrm{def}}{=}1+\max\left\{\frac{d_2^2}{w_1^2},\frac{d_2^2}{w_2^2},\frac{d_2^2}{d_1^2}\right\},$$

and the resulting bump is roughly $(d_2/\sqrt{h})$-far from a 4-dimensional hyperplane at scale $\sqrt{h}$. If $d_2/\sqrt{h}=\varepsilon$, then $d_1w_1=h=\varepsilon^{-2}d_2^2$, and

$$\nu(w_1,w_2,d_1,d_2,h)\geqslant 1+\frac{d_2^2}{\max\{d_1^2,w_1^2\}}\geqslant 1+\frac{d_2^2}{d_1w_1}=1+\varepsilon^2.$$

Hence, this construction results, at best, in a surface that is ε-far from planes at ε^{-2} different scales. One may also consider bumps where d_1w_1, d_2w_2, and h are not proportional, such as bumps with $d_1=w_1=w_2=d_2^{-1}=r\gg 1=h$. This is more subtle than it might initially seem. Indeed, because the d_1- and w_1-directions do not commute, there are no $r\times r\times r\times r^{-1}\times 1$ boxes in $\mathbb{H}^5$ that stay close to horizontal. Consequently, a bump of these dimensions behaves similarly to a collection of smaller bumps with $d_1w_1=d_2w_2=h$, which are governed by the previous reasoning.

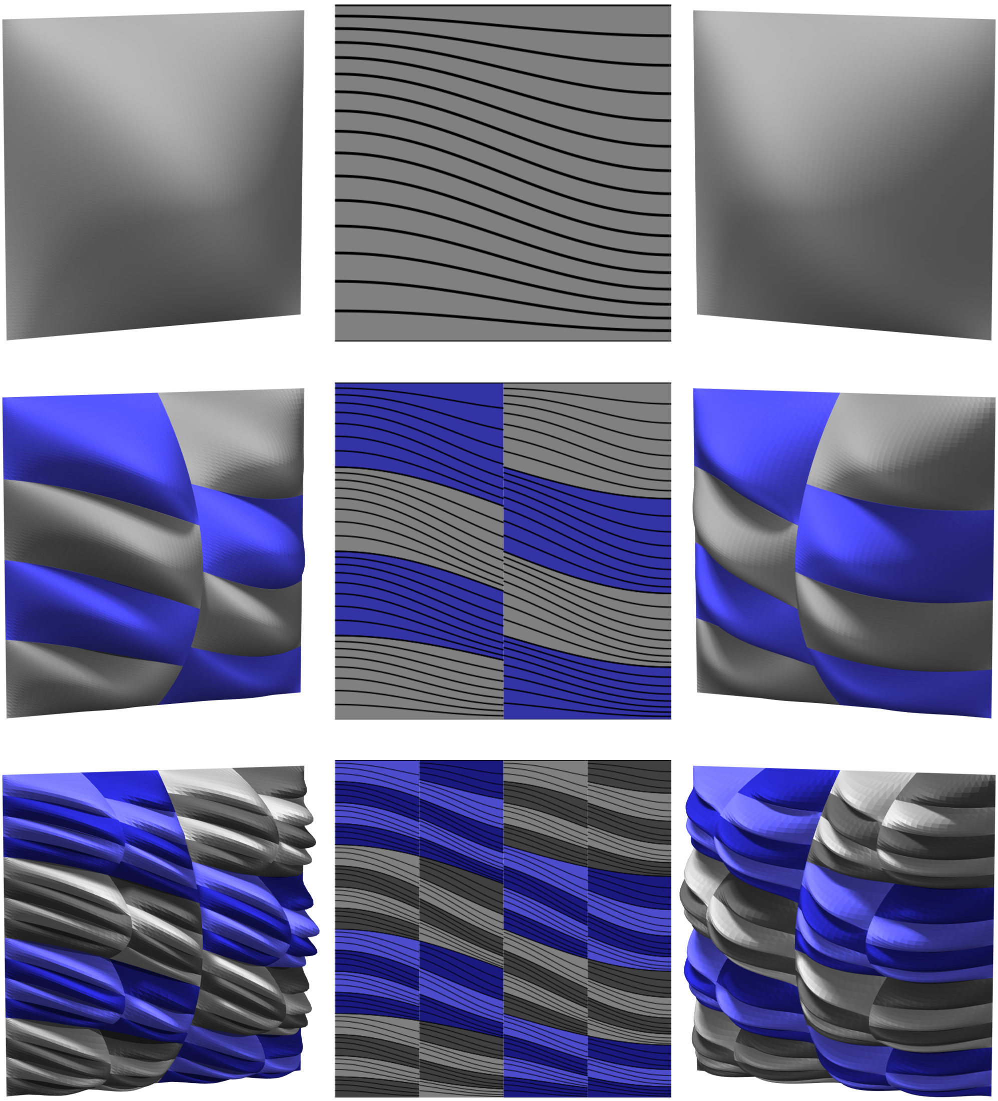

Figure 3. The first three steps of the construction of a maximally rough surface in $\mathbb{H}$. The left and right columns show the same surface from two different angles. The center column shows a projection of the surface to the plane, with characteristic curves marked. Since the second derivatives of these curves are small, the Heisenberg area of the surface is bounded, but the surface can be made ε-far from a plane at ε^{-4} different scales — much more than what is possible in $\mathbb{H}^5$.

1.3. Roadmap

In §2, we present notation for working with the Heisenberg group and some definitions and results related to intrinsic graphs and characteristic curves. In §3, we construct an intrinsic graph with large vertical perimeter and use it to construct the embeddings used in Theorem 1.7 and its consequences.

The rest of the paper is devoted to defining and constructing foliated corona decompositions and using them to prove equation (1.32) bounding the vertical perimeter of an intrinsic Lipschitz graph. In §4, we define a rectilinear foliated patchwork, which decomposes an intrinsic Lipschitz graph into rectilinear pseudoquads, and in §5, we define the weighted Carleson packing condition required for such a patchwork to be a foliated corona decomposition. Then, in §6, we show that an intrinsic Lipschitz graph that admits a foliated corona decomposition satisfies equation (1.32).

It remains to show that every intrinsic Lipschitz graph admits a foliated corona decomposition. We produce foliated corona decompositions by the subdivision algorithm described in §7. The fact that the patchworks produced by this algorithm satisfy the weighted Carleson packing condition relies on careful analysis of a coercive quantity, the extended parametric non-monotonicity, defined in §8. When this coercive quantity is small, the graph satisfies strong geometric bounds, detailed in Proposition 7.2. Assuming Proposition 7.2, we prove the weighted Carleson condition in §9. In §10, we outline the proof of Proposition 7.2, and in §11 and §12 we prove it.

2. Preliminaries

Most of this section presents initial facts about the Heisenberg group that will be used throughout what follows. However, we will start by briefly setting notation for measure theoretical boundaries and interiors that are best described in greater generality (though they will be applied below only to either the Heisenberg group or the real line).

Let $(\mathbb{M}, d_{\mathbb{M}}, \mu)$ be a non-degenerate metric measure space, i.e., $(\mathbb{M}, d_{\mathbb{M}})$ is a metric space and μ is a Borel measure on $\mathbb{M}$ such that $\mu(B_{\mathbb{M}}(x,r))>0$ for all $x\in\mathbb{M}$ and $r>0$, where $B_{\mathbb{M}}(x,r)=\{y\in\mathbb{M}: d_{\mathbb{M}}(x,y)\leqslant r\}$ is the closed $d_{\mathbb{M}}$-ball of radius r centered at x.

Given a subset $S\subseteq\mathbb{M}$, we define the *measure-theoretic support* $\mathrm{supp}_\mu(S)$ of S to be the usual measure-theoretic support of the indicator function $\mathbf{1}_S:\mathbb{M}\to\{0,1\}$, namely

$$\mathrm{supp}_\mu(S) \stackrel{\mathrm{def}}{=} \bigcap_{r>0}\{x\in\mathbb{M}: \mu(B_{\mathbb{M}}(x,r)\cap S)>0\}. \tag{2.1}$$

The *measure-theoretic boundary* of S is defined as

$$\partial_\mu S \stackrel{\mathrm{def}}{=} \mathrm{supp}_\mu(S)\cap\mathrm{supp}_\mu(\mathbb{M}\setminus S) = \bigcap_{r>0}\left\{x\in\mathbb{M}: 0<\frac{\mu(B_{\mathbb{M}}(x,r)\cap S)}{\mu(B_{\mathbb{M}}(x,r))}<1\right\}. \tag{2.2}$$

The *measure-theoretic interior* of S is defined as

$$\mathrm{int}_\mu(S) \stackrel{\mathrm{def}}{=} \mathbb{M}\setminus \mathrm{supp}_\mu(\mathbb{M}\setminus S) = \bigcup_{r>0}\{x\in\mathbb{M} : \mu(B_{\mathbb{M}}(x,r)\setminus S)=0\}. \tag{2.3}$$

These definitions are non-standard; other works define the measure-theoretic boundary as the set of points where the density of S is not 0 or 1. The advantage of our definition is that one may check that $\mathrm{int}_\mu(S)$ is open in $\mathbb{M}$ and its (topological) boundary $\partial\,\mathrm{int}_\mu(S)$ is contained in $\partial_\mu S$. The sets $\mathrm{int}_\mu(S)$, $\mathrm{int}_\mu(\mathbb{M}\setminus S)$, and $\partial_\mu S$ are disjoint and their union is $\mathbb{M}$, i.e.,

$$\mathbb{M} = \mathrm{int}_\mu(S) \sqcup \mathrm{int}_\mu(\mathbb{M}\setminus S) \sqcup \partial_\mu S. \tag{2.4}$$

2.1. The Heisenberg group

Here we summarize basic notation and terminology related to the Heisenberg group.

Throughout what follows, $\|\cdot\|: \mathbb{R}^3\to\mathbb{R}$ will denote the Euclidean norm on $\mathbb{R}^3$, namely $\|(a,b,c)\|=\sqrt{a^2+b^2+c^2}$ for all $a,b,c\in\mathbb{R}$. Let

$$X \stackrel{\mathrm{def}}{=} (1,0,0), \quad Y \stackrel{\mathrm{def}}{=} (0,1,0), \quad Z \stackrel{\mathrm{def}}{=} (0,0,1)$$

be the standard basis of $\mathbb{R}^3$, and let $x,y,z: \mathbb{R}^3\to\mathbb{R}$ be the coordinate functions. Namely, for $u=(a,b,c)\in\mathbb{R}^3$, we set $x(u)=a$, $y(u)=b$, and $z(u)=c$. With this notation, the Heisenberg group operation (1.8) can be written as

$$uv = u+v+\tfrac{1}{2}(x(u)y(v)-y(u)x(v))Z \quad \text{for all } u,v\in\mathbb{H}=\mathbb{R}^3. \tag{2.5}$$

The linear span of a set of vectors $S\subseteq\mathbb{R}^3$ will be denoted $\langle S\rangle$. The plane $\mathsf{H}\stackrel{\mathrm{def}}{=}\langle X,Y\rangle$ is called the space of *horizontal vectors*. Let $\pi: \mathbb{R}^3\to\mathsf{H}$ be the orthogonal projection. A *horizontal line* in $\mathbb{H}$ is a coset of the form $w\langle h\rangle\subseteq\mathbb{H}$ for some $w\in\mathbb{H}$ and $h\in\mathsf{H}$.

The union of the horizontal lines passing through a point $u\in\mathbb{H}$ is the plane $u\mathsf{H}$, which we denote H_u and call the *horizontal plane centered at* u. Every plane $P\subseteq\mathbb{R}^3$ either contains a coset of $\langle Z\rangle$ (a *vertical line*), in which case we call P a *vertical plane*, or can be written $P=\mathsf{H}_u$ for some unique $u\in\mathbb{H}$.

If $I\subseteq\mathbb{R}$ is an interval and $\gamma: I\to\mathbb{H}$ is a curve such that $x\circ\gamma, y\circ\gamma, z\circ\gamma: I\to\mathbb{R}$ are Lipschitz, then $\gamma'(t)$ is defined for almost all $t\in I$. One then says that γ is a *horizontal curve* if γ is tangent to $\mathsf{H}_{\gamma(t)}$ at $\gamma(t)$ for almost all $t\in I$, i.e., for almost all $t\in I$ we have

$$\frac{d}{ds}(\gamma(t)^{-1}\gamma(s))\bigg|_{s=t} \in \mathsf{H}.$$

Note that horizontality is left-invariant; if γ is a horizontal curve and $g\in\mathbb{H}$, then $g\cdot\gamma$ is also a horizontal curve. If $\gamma(t)=(\gamma_x(t),\gamma_y(t),\gamma_z(t))$, then this requirement is equivalent to the differential equation

$$2\gamma_z'(t)=\gamma_x(t)\gamma_y'(t)-\gamma_y(t)\gamma_x'(t).$$

Define

$$\ell(\gamma)\stackrel{\mathrm{def}}{=}\int_I\|\pi(\gamma'(t))\|\,dt.$$

The sub-Riemannian or Carnot–Carathéodory metric $d:\mathbb{H}\times\mathbb{H}\to[0,\infty)$ is defined by letting $d(v,w)$ be the infimum of $\ell(\gamma)$ over all horizontal curves γ connecting $v\in\mathbb{H}$ to $w\in\mathbb{H}$. This metric is left-invariant, i.e., $d(ga,gb)=d(a,b)$ for all $a,b,g\in\mathbb{H}$.

If γ is a horizontal curve connecting v to w, then $\pi\circ\gamma$ is a curve in $\mathbb{R}^2$ of the same length connecting $\pi(v)$ to $\pi(w)$, so $d(v,w)\geqslant\|\pi(v)-\pi(w)\|$. Consequently, any horizontal line in $\mathbb{H}$ is a geodesic. Also, d satisfies (e.g. [10], [40], [80]) the ball-box inequality

$$d(\mathbf{0},h)\leqslant|x|+|y|+4\sqrt{|z|}\leqslant 2d(\mathbf{0},h)+4\cdot\frac{d(\mathbf{0},h)}{\sqrt{2\pi}}\leqslant 4d(\mathbf{0},h)\quad\text{for all } h=(x,y,z)\in\mathbb{H}. \tag{2.6}$$

For $h\in\mathbb{H}$ and $r\geqslant0$ we let $B_r(h)=\{g\in\mathbb{H}:d(g,h)\leqslant r\}=hB_r(\mathbf{0})$ denote the closed ball of radius r centered at h with respect to the sub-Riemannian metric d on $\mathbb{H}$. Throughout what follows, we will not use this notation for balls with respect to any other metric.

For $\sigma>0$ denote by $\mathcal{H}^\sigma$ the σ-dimensional Hausdorff measure that d induces on $\mathbb{H}$. Thus $\mathcal{H}^4$ is the Lebesgue measure on $\mathbb{R}^3$, which is also the Haar measure on $\mathbb{H}$. Given a measurable subset $E\subseteq\mathbb{H}$, the associated *perimeter measure* that is induced by d will be denoted by $\mathrm{Per}_E(\cdot)$; we refer to [35] for background on this fundamental notion, noting only that there exists $\eta>0$ such that if $E\subseteq\mathbb{H}$ has a piecewise smooth boundary, then

$$\mathrm{Per}_E(U)=\eta\mathcal{H}^3(U\cap\partial E)$$

for every open subset $U\subseteq\mathbb{H}$.

It is also beneficial to describe the group operation on $\mathbb{H}$ in terms of a symplectic form. Let $\omega_{\mathbb{R}^2}:\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}$ be the standard symplectic form, i.e.,

$$\omega_{\mathbb{R}^2}((a,b),(\alpha,\beta))\stackrel{\mathrm{def}}{=}a\beta-b\alpha=\det\begin{pmatrix}a & b\\ \alpha & \beta\end{pmatrix}\quad\text{for all } (a,b),(\alpha,\beta)\in\mathbb{R}^2.$$

Under this notation, (2.5) can be written as follows:

$$uv=u+v+\tfrac12\omega_{\mathbb{R}^2}(\pi(u),\pi(v))Z\quad\text{for all } u,v\in\mathbb{H}. \tag{2.7}$$

This lets us define automorphisms of $\mathbb{H}$. Let $A:\mathbb{R}^2\to\mathbb{R}^2$ be an invertible linear map with determinant $J\in\mathbb{R}\setminus\{0\}$, so that $\omega_{\mathbb{R}^2}(A(v),A(w))=J\omega_{\mathbb{R}^2}(v,w)$ for any $v,w\in\mathbb{R}^2$. It follows from (2.7) that the map $\tilde{A}:\mathbb{H}\to\mathbb{H}$ that is defined by

$$\tilde{A}(x,y,z)\stackrel{\mathrm{def}}{=}(A(x,y),Jz) \quad \text{for all } (x,y,z)\in\mathbb{H} \tag{2.8}$$

is an automorphism of $\mathbb{H}$ which, since $\tilde{A}(\mathsf{H})=\mathsf{H}$, sends horizontal curves to horizontal curves and is thus Lipschitz with respect to the sub-Riemannian metric on $\mathbb{H}$. If A is an orthogonal matrix, then $\tilde{A}$ is an isometry. As a notable special case, for $a,b>0$, we define

$$s_{a,b}(x,y,z)\stackrel{\mathrm{def}}{=}(ax,by,abz) \quad \text{for all } (x,y,z)\in\mathbb{H}, \tag{2.9}$$

which we call a *stretch map*. When $a=b=t$, $s_{t,t}$ is the usual scaling automorphism of $\mathbb{H}$, which scales the sub-Riemannian metric on $\mathbb{H}$ by a factor of t. For simplicity, in what follows we will sometimes write $s_{t,t}=s_t$.

2.2. Intrinsic graphs and intrinsic Lipschitz graphs

Throughout what follows, we denote the xz-plane by V_0, namely

$$V_0\stackrel{\mathrm{def}}{=}\{(x,y,z)\in\mathbb{H}: y=0\}=\mathbb{R}\times\{0\}\times\mathbb{R}\subseteq\mathbb{H}.$$

Note that the restriction of $\mathcal{H}^3$ to V_0 is proportional to the Lebesgue measure on V_0.

Fix $U\subseteq V_0$. The intrinsic graph of a function $\psi: U\to\mathbb{R}$ is defined in [36] to be

$$\Gamma_\psi\stackrel{\mathrm{def}}{=}\{vY^{\psi(v)}: v\in U\}=\left\{\left(x(v),\psi(v),z(v)+\tfrac{1}{2}x(v)\psi(v)\right): v\in U\right\}\subseteq\mathbb{H}, \tag{2.10}$$

where in (2.10), as well as throughout what follows, it is convenient to use the exponential notation

$$u^t=tu=(tx(u),ty(u),tz(u))$$

for $u\in\mathbb{H}$ and $t\in\mathbb{R}$. Observe that any coset of $\langle Y\rangle$ that passes through U intersects Γ_ψ in exactly one point. We will also use the following notation for the intrinsic epigraph of ψ:

$$\Gamma_\psi^+\stackrel{\mathrm{def}}{=}\{vY^t: (v,t)\in U\times(\psi(v),\infty)\}.$$

Suppose that $U\subseteq V_0$ is an open subset of V_0 and that $g: U\to\mathbb{R}$ is smooth. For every $\psi: U\to\mathbb{R}$ define a function $\partial_\psi g: U\to\mathbb{R}$ by

$$\partial_\psi g\stackrel{\mathrm{def}}{=}\frac{\partial g}{\partial x}-\psi\frac{\partial g}{\partial z}. \tag{2.11}$$

If ψ is smooth, then we define the *horizontal derivative* of ψ to be the function

$$\partial_\psi \psi = \frac{\partial \psi}{\partial x} - \psi \frac{\partial \psi}{\partial z}. \tag{2.12}$$

Let $v \in U$ and let $p \stackrel{\mathrm{def}}{=} vY^{\psi(v)} \in \Gamma_\psi$. One can interpret $\partial_\psi \psi$ by considering the horizontal plane H_p. This plane locally intersects Γ_ψ in a curve, and the tangent vector of this curve at p is given by $X + \partial_\psi \psi(v) Y$. The horizontal derivative also determines the slope of the intrinsic tangent plane to Γ_ψ, where the slope of a vertical plane is the slope of its projection to H. As $r \to 0$, rescalings of the intersections $B_r(p) \cap \Gamma_\psi$ converge to a vertical tangent plane with slope $\partial_\psi \psi(v)$.

The following proposition is part of [2, Theorem 1.2]. It expresses the area $\mathcal{H}^3(\Gamma_\psi)$ of Γ_ψ, namely the 3-dimensional Hausdorff measure (with respect to the sub-Riemannian metric) of Γ_ψ in terms of $\partial_\psi \psi$.

Proposition 2.1. ([2]) *There exists a constant $c > 0$ such that if $U \subseteq V_0$ is an open set and $\psi: U \to \mathbb{R}$ is smooth, then*

$$\mathcal{H}^3(\Gamma_\psi) \asymp \mathcal{S}^3(\Gamma_\psi) = c \int_U \sqrt{1 + (\partial_\psi \psi)^2}\, dw \asymp \mathcal{H}^3(U) + \|\partial_\psi \psi\|_{L_1(U)}, \tag{2.13}$$

where $\mathcal{S}^3$ is the 3-dimensional spherical Hausdorff measure on $\mathbb{H}$.

Recent work [49] has shown that the spherical Hausdorff measure and the Hausdorff measure on Γ_ψ are equal up to a multiplicative constant, so the first equivalence in (2.13) can be replaced by an equality up to a constant factor.

For $\lambda \in (0,1)$, define the double cone

$$\mathrm{Cone}_\lambda \stackrel{\mathrm{def}}{=} \{h \in \mathbb{H} : |y(h)| > \lambda d(\mathbf{0}, h)\}.$$

This is a cone centered on the horizontal line $\langle Y \rangle$ which is scale-invariant, i.e.,

$$s_{t,t}(\mathrm{Cone}_\lambda) = \mathrm{Cone}_\lambda \quad \text{for all } t > 0.$$

The intersection $\mathsf{H} \cap \mathrm{Cone}_\lambda$ is a double cone in H with angle depending on λ. Specifically,

$$\begin{aligned} \mathsf{H} \cap \mathrm{Cone}_\lambda &= \{(x, y, 0) \in \mathbb{H} : |y| > \lambda \sqrt{x^2 + y^2}\} \\ &= \left\{(x, y, 0) \in \mathbb{H} : |y| > \frac{\lambda}{\sqrt{1 - \lambda^2}} |x|\right\}. \end{aligned} \tag{2.14}$$

Definition 2.2. Let $U \subseteq V_0$ and let $\Gamma \subseteq \mathbb{H}$ be an intrinsic graph over U. For any $\lambda \in (0,1)$, we say that Γ is an *intrinsic λ-Lipschitz graph* if $(h\, \mathrm{Cone}_\lambda(V_0)) \cap \Gamma = \varnothing$ for every $h \in \Gamma$. Equivalently, for every $p, q \in \Gamma$,

$$|y(q) - y(p)| \leqslant \lambda d(p, q).$$

We say that Γ is an *intrinsic Lipschitz graph* if it is intrinsic λ-Lipschitz for some $\lambda \in (0,1)$. If $\Gamma = \Gamma_\psi$ for some $\psi: U \to \mathbb{R}$, then we say that ψ is an *intrinsic Lipschitz function*.

Definition 2.2 gives the same class of intrinsic Lipschitz graphs as the definition introduced in [36], but it gives different classes of intrinsic λ-Lipschitz graphs; see §3.2 of [103] for a proof that the definitions are equivalent.

The following simple bound will be convenient later.

LEMMA 2.3. *Let $0\leqslant\lambda\leqslant 1$ and let $\Gamma=\Gamma_\psi$ be an intrinsic λ-Lipschitz graph of a function $\psi: U\subseteq V_0\to\mathbb{R}$. Let $v,w\in U$ and write $p=vY^{\psi(v)}\in\Gamma$ and $q=wY^{\psi(w)}\in\Gamma$. Then,*

$$|y(p)-y(q)|=|\psi(v)-\psi(w)|\leqslant\frac{2}{1-\lambda}d(p,q\langle Y\rangle).$$

Proof. Denote $m=d(p,w\langle Y\rangle)$. Let $c\in w\langle Y\rangle$ be a point such that $d(p,c)=m$. By the intrinsic Lipschitz condition,

$$|y(c)-y(q)|\leqslant m+|y(p)-y(q)|\leqslant m+\lambda d(p,q)\leqslant m+\lambda(m+|y(c)-y(q)|).$$

This simplifies to give

$$|y(c)-y(q)|\leqslant\frac{1+\lambda}{1-\lambda}m.$$

Hence,

$$|y(p)-y(q)|\leqslant|y(p)-y(c)|+|y(c)-y(q)|\leqslant\frac{2m}{1-\lambda}. \qquad\square$$

Intrinsic Lipschitz graphs satisfy the following version of Rademacher's differentiation theorem due to [38, Theorem 4.29].

THEOREM 2.4. ([38]) *Let $0<\lambda<1$, let $U\subseteq V_0$ be an open set, and let $f: U\to\mathbb{R}$ be a function such that $\Gamma_\psi\subseteq\mathbb{H}$ is an intrinsic λ-Lipschitz graph. Then, for almost every $p\in U$, Γ_ψ has an intrinsic tangent plane at $pY^{\psi(p)}$ whose slope satisfies*

$$|\partial_\psi\psi(p)|\leqslant\frac{\lambda}{\sqrt{1-\lambda^2}}. \tag{2.15}$$

We note that [38, Theorem 4.29] is concerned with the (almost everywhere) existential statement of horizontal derivatives. The upper bound in (2.15) follows from (2.14) and the fact that the intrinsic tangent plane at $pY^{\psi(p)}$ is disjoint from $p\,\mathrm{Cone}_\lambda$ (see also Lemma 2.7). This bound on the horizontal derivatives of an intrinsic Lipschitz graph leads to a bound on the perimeter measure. The following result follows from [37, Theorem 4.1], which proves a similar bound on the Hausdorff measure of Γ, and the results of [35], which imply that the Hausdorff measure of Γ and the perimeter measure of Γ^+ differ by at most a multiplicative constant. Let $\Pi:\mathbb{H}\to V_0$ be the natural (non-linear) projection to V_0 along cosets of $\langle Y\rangle$, i.e., $\Pi(v)=vY^{-y(v)}$ for every $v\in\mathbb{H}$. Equivalently,

$$\Pi(x,y,z)\stackrel{\mathrm{def}}{=}\big(x,0,z-\tfrac{1}{2}xy\big)\quad\text{for all }(x,y,z)\in\mathbb{H}. \tag{2.16}$$

LEMMA 2.5. ([37]) *Fix* $\lambda\in(0,1)$. *Let* $\psi: V_0\to\mathbb{R}$ *be* λ*-intrinsic Lipschitz. The perimeter measure* $\mathrm{Per}_{\Gamma_\psi^+}$ *satisfies the following equivalence for measurable subsets* $A\subseteq\Gamma_\psi$:

$$\mathrm{Per}_{\Gamma_\psi^+}(A)\asymp_\lambda |\Pi(A)|,$$

where here, and henceforth, $|\cdot|$ *denotes the Haar measure on* V_0, *normalized to coincide with the usual* 2*-dimensional area measure in* $\mathbb{R}^3$.

2.3. Characteristic curves

Let $U\subseteq V_0$ be an open set and let $\psi: U\to\mathbb{R}$ be a continuous function. The differential operator ∂_ψ given in (2.11) defines a vector field on V_0 that is continuous and has x-coordinate 1, so by the Peano existence theorem, there is at least one flow line of ∂_ψ through every point of U, defined on an interval. These flow lines are the graphs of functions $g: I\to\mathbb{R}$ satisfying

$$g'(t)+\psi(t,0,g(t))=0 \quad \text{for all } t\in I. \tag{2.17}$$

We call these flow lines *characteristic curves* of Γ_ψ.

The solution to (2.17) guaranteed by the Peano existence theorem is only local, but when ψ is intrinsic Lipschitz, we can define g on all of $\mathbb{R}$. Indeed, by the Peano existence theorem, if $S_r=[-1,1]\times\{0\}\times[-r,r]$ and $\sup_{q\in S_r}|\psi(q)|\leqslant r$, then there exists a $g: (-1,1)\to[-r,r]$ that solves (2.17) with initial condition $g(0)=0$. Let $(x,0,z)\in S_r$. By Lemma 2.3 with $v=\mathbf{0}$, $w=(x,0,z)$, there is some $C=C_\psi>0$ such that

$$|\psi(x,0,z)|\leqslant|\psi(\mathbf{0})|+\frac{2}{1-\lambda}d(Y^{\psi(\mathbf{0})},(x,0,z))\leqslant C+C|x|+C\sqrt{|z|}\leqslant 2C+C\sqrt{r}.$$

If r is sufficiently large, then $\sup_{q\in S_r}|\psi(q)|\leqslant r$, so (2.17) can be solved on $(-1,1)$. More generally, for any x_0 and z_0, there is a $g: (x_0-1,x_0+1)\to\mathbb{R}$ that solves (2.17) with initial condition $g(x_0)=z_0$. By patching together such solutions, we obtain a global solution to (2.17).

In this section, we will show that the characteristic curves of Γ_ψ are the projections of horizontal curves in Γ_ψ, and use them to describe Γ_ψ. In the next section, we will describe how characteristic curves transform under automorphisms of $\mathbb{H}$; later, we will use these curves to describe how horizontal lines intersect an intrinsic Lipschitz graph.

LEMMA 2.6. *Let* $\Gamma=\Gamma_\psi$. *The characteristic curves of* Γ *are exactly the projections* (*under* Π) *of horizontal curves* $\phi: I\to\Gamma$ *such that* $x(\phi(t))=t$ *for every* $t\in I$.

Because characteristic curves can branch and rejoin (see [12] for such examples), there are intrinsic Lipschitz graphs with horizontal curves whose x-coordinate is not monotone. Thus the condition $x(\phi(t))=t$ of Lemma 2.6 cannot be dropped.

Proof. First, we claim that if ϕ is a horizontal curve in Γ with $x(\phi(t))=t$, then $\Pi\circ\phi$ is a characteristic curve of Γ. Write $\Gamma=\Gamma_\psi$ and let $\phi\colon I\to\Gamma$ be a horizontal curve of the form $\phi(t)=X^tY^{f(t)}Z^{g(t)}$. Then f and g are Lipschitz, $\Pi(\phi(t))=(t,0,g(t))$, and, since $\phi(t)\in\Gamma$, we have $f(t)=\psi(t,0,g(t))$. Since ϕ is horizontal,

$$\frac{d}{du}\phi(t)^{-1}\phi(t+u)\bigg|_{u=0}\in\mathsf{H}$$

for almost every $t\in I$. Observe that

$$\begin{aligned}\phi(t)^{-1}\phi(t+u)&=(X^tY^{f(t)}Z^{g(t)})^{-1}(X^{t+u}Y^{f(t+u)}Z^{g(t+u)})\\&=X^uY^{f(t+u)-f(t)}Z^{g(t+u)-g(t)+uf(t)}.\end{aligned}$$

Since f and g are Lipschitz, the following identity holds almost everywhere:

$$\frac{d}{du}\phi(t)^{-1}\phi(t+u)\bigg|_{u=0}=X+f'(t)Y+(g'(t)+f(t))Z. \tag{2.18}$$

That is, g satisfies (2.17).

Conversely, suppose that g is a solution of (2.17) and let $f(t)=\psi(t,0,g(t))$. By [12, Theorems 1.1 and 1.2], f is Lipschitz. Therefore, $\phi(t)=X^tY^{f(t)}Z^{g(t)}$ is a Lipschitz curve in Γ such that $\Pi(\phi(t))=(t,0,g(t))$ and such that ϕ satisfies (2.18) almost everywhere. In combination with (2.17), this implies that ϕ is horizontal. □

If ψ is smooth, the characteristic curves of Γ_ψ foliate U. If ψ is merely intrinsic Lipschitz, characteristic curves can branch and rejoin, but if two characteristic curves pass through the same point, then they are tangent at that point; see [12, Figure 1] for an example of this phenomenon.

Characteristic curves satisfy bounds based on the intrinsic Lipschitz constant of Γ.

Lemma 2.7. *Fix $\lambda\in(0,1)$ and denote*

$$L\stackrel{\text{def}}{=}\frac{\lambda}{\sqrt{1-\lambda^2}}.$$

Let $\Gamma=\Gamma_\psi$ be an intrinsic λ-Lipschitz graph over an open set and let $\gamma\colon I\to V_0$ be a characteristic curve for Γ parameterized so that $x(\gamma(t))=t$ for all $t\in I$. Then,

$$|\psi(\gamma(s))-\psi(\gamma(t))|\leqslant L|s-t|\quad\text{for all } s,t\in I. \tag{2.19}$$

Also, if we denote $g(t)=z(\gamma(t))$, then

$$|g(t)-g(s)-g'(s)\cdot(t-s)|\leqslant\tfrac{1}{2}L(t-s)^2\quad\text{for all } s,t\in I. \tag{2.20}$$

Proof. Since γ is characteristic, the curve $\phi(t)=\gamma(t)\cdot Y^{\psi(\gamma(t))}$ is horizontal. The intrinsic Lipschitz condition implies that

$$\frac{|y(\phi(t+\delta))-y(\phi(t))|}{d(\phi(t),\phi(t+\delta))} \leqslant \lambda \quad \text{for all } \delta \in \mathbb{R}\setminus\{0\}. \tag{2.21}$$

By Pansu's theorem [95], for almost every $t\in I$, there is a vector $h_t\in \mathsf{H}$ such that

$$\lim_{\delta\to 0}\frac{d(\phi(t)h_t^\delta,\phi(t+\delta))}{\delta}=0.$$

Indeed, $h_t=(1,m,0)$, where $m=(\psi\circ\gamma)'(t)$. Then

$$\liminf_{\delta\to 0}\frac{|y(\phi(t+\delta))-y(\phi(t))|}{d(\phi(t),\phi(t+\delta))} \geqslant \liminf_{\delta\to 0}\frac{|\delta m|-d(\phi(t)h_t^\delta,\phi(t+\delta))}{\delta\|h_t\|+d(\phi(t)h_t^\delta,\phi(t+\delta))}=\frac{|m|}{\sqrt{1+m^2}}.$$

By (2.21), it follows that $|m|/\sqrt{1+m^2}\leqslant\lambda$, so, for almost every $t\in I$,

$$|(\psi\circ\gamma)'(t)|=|m|\leqslant L. \tag{2.22}$$

This implies (2.19). By (2.17), $g'(t)=-\psi(\gamma(t))$, so it follows from (2.22) that $|g''(t)|\leqslant L$ for almost every $t\in I$. The remaining bound (2.20) is therefore justified as follows:

$$|g(t)-(g(s)+g'(s)\cdot(t-s))|=\left|\int_s^t (t-u)g''(u)\,du\right|\leqslant L\left|\int_s^t (t-u)\,du\right|=L\frac{(t-s)^2}{2}. \qquad \square$$

Since there is a characteristic curve through every point $p\in U$ and the derivative of such a curve at p is $-\psi(p)$, an intrinsic graph Γ can be reconstructed from its characteristic curves. Indeed, one way to construct intrinsic Lipschitz graphs is to construct a foliation of V_0 by C_1 curves $\{z=g_\alpha(x)\}$, $\alpha\in A$ such that $\mathrm{Lip}(g'_\alpha)\lesssim 1$ for every $\alpha\in A$. Each such curve lifts to a horizontal curve, and one can show that the union of these lifts is an intrinsic Lipschitz graph. (This is how the graphs in Figure 3 were constructed.)

For illustration, we consider planes in $\mathbb{H}$. A vertical plane V that is not orthogonal to V_0 is an intrinsic graph over V_0. The horizontal curves in V are parallel lines; let L be one such line. The image $\Pi(L)$ is a parabola in V_0, and the characteristic curves of V are the parabolas parallel to $\Pi(L)$. The second derivative of these parabolas depends on the angle between V and V_0.

Let $v\in\mathbb{H}$. The horizontal plane H_v centered at v is not an intrinsic graph, but the horizontal line $v\langle Y\rangle$ divides H_v into two intrinsic graphs. The horizontal lines in H_v all pass through v, and their projections to V_0 are parabolas through $\Pi(v)$. Since they all intersect at v, their projections are all tangent at $\Pi(v)$. These parabolas foliate the complement in V_0 of the vertical line through $\Pi(v)$. They have unboundedly large second derivatives, so the two halves of H_v are locally intrinsic Lipschitz graphs, but not globally.

2.4. Automorphisms and characteristic curves

Recall that any invertible linear map $A:\mathbb{R}^2\to\mathbb{R}^2$ induces an automorphism $\tilde{A}$ of $\mathbb{H}$ as in (2.8). We are particularly interested in the case that Y is an eigenvector of A. In this case, $\tilde{A}(\langle Y\rangle)=\langle Y\rangle$, so $\tilde{A}$ sends cosets of $\langle Y\rangle$ to cosets of $\langle Y\rangle$. A set Γ is an intrinsic graph if and only if it intersects each coset of $\langle Y\rangle$ at most once, so $\tilde{A}$ sends intrinsic graphs to intrinsic graphs.

One family of maps with this property are the stretch maps

$$s_{a,b}(x,y,z)=(ax,by,abz)$$

defined in (2.9). To construct a second family of maps with the above property, let $b\in\mathbb{R}$ and consider the linear map $A_b(x,y)=(x,y+bx)$, which is a shear of the plane $\mathbb{R}^2$. The induced map $\tilde{A}_b$ is an automorphism of $\mathbb{H}$ given by the formula

$$\tilde{A}_b(x,y,z)=(x,y+bx,z) \quad \text{for all } (x,y,z)\in\mathbb{H},$$

and we call such maps *shear maps.* (Note that these are different from the shear maps considered in [106].)

Let $\Pi:\mathbb{H}\to V_0$ be as in (2.16), i.e., the projection to V_0 along cosets of $\langle Y\rangle$. The maps above preserve cosets of $\langle Y\rangle$, so composed with Π they induce maps from V_0 to V_0.

LEMMA 2.8. *Fix $h=(x_0,y_0,z_0)\in\mathbb{H}$ and $v=(x,0,z)\in V_0$. For any $a,b,t\in\mathbb{R}$ we have*

$$\begin{aligned}
\Pi(s_{a,b}(vY^t)) &= s_{a,b}(v)=(ax,0,abz),\\
\Pi(\tilde{A}_b(vY^t)) &= \left(x,0,z-\tfrac{1}{2}bx^2\right),\\
\Pi(hvY^t) &= \left(x+x_0,0,z+z_0-xy_0-\tfrac{1}{2}x_0y_0\right).
\end{aligned}$$

Proof. $\Pi(gY^t)=\Pi(g)$ for all $g\in\mathbb{H}$ and $t\in\mathbb{R}$. Since $s_{a,b}$ and $\tilde{A}_b$ are homomorphisms,

$$\begin{aligned}
\Pi(s_{a,b}(vY^t)) &= \Pi(s_{a,b}(v)Y^{bt})=s_{a,b}(v)=(ax,0,abz),\\
\Pi(\tilde{A}_b(vY^t)) &= \Pi(\tilde{A}_b(v)Y^t)=(x,bx,z)Y^{-bx}=\left(x,0,z-\tfrac{1}{2}bx^2\right).
\end{aligned}$$

Finally,

$$\begin{aligned}
\Pi(hvY^t)=\Pi(hv) &= \left(x_0+x,y_0,z_0+z-\tfrac{1}{2}xy_0\right)Y^{-y_0}\\
&= \left(x_0+x,0,z_0+z-\tfrac{1}{2}xy_0-\tfrac{1}{2}(x_0+x)y_0\right). \qquad \square
\end{aligned}$$

We next describe how these maps affect characteristic curves and intrinsic graphs.

Lemma 2.9. *Let $U \subseteq V_0$ and $\psi: U \to \mathbb{R}$ be a continuous function. Write $\Gamma = \Gamma_\psi$. Let $C = \{(x, 0, z) \in V_0 : z = g(x)\}$ be a characteristic curve of Γ. Let $q: \mathbb{H} \to \mathbb{H}$ be a stretch map, shear map, or left translation, and let $\hat{q}: V_0 \to V_0$, $\hat{q}(v) = \Pi(q(v))$, be the map that q induces on V_0. Then, $q(\Gamma)$ is the intrinsic graph of a function $\hat{\psi}: \hat{q}(U) \to \mathbb{R}$, and $\hat{q}(C)$ is a characteristic curve of $q(\Gamma)$. Also, the following statements hold.*

- *If $a, b \in \mathbb{R} \setminus \{0\}$ and $q = s_{a,b}$, then $\hat{\psi}(\hat{q}(v)) = b\psi(v)$ for all $v \in U$.*
- *If $b \in \mathbb{R}$ and $q = \tilde{A}_b$, then $\hat{\psi}(\hat{q}(v)) = \psi(v) + bx(v)$ for all $v \in U$.*
- *If $h \in \mathbb{H}$ and $q(p) = hp$ for all $p \in \mathbb{H}$, then $\hat{\psi}(\hat{q}(v)) = \psi(v) + y(h)$ for all $v \in U$.*

Proof. Any coset of $\langle Y \rangle$ intersects $q(\Gamma)$ at most once, so $q(\Gamma)$ is an intrinsic graph with domain $\Pi(q(\Gamma)) = \hat{q}(\Gamma)$.

Let $\gamma \subseteq \Gamma$ be the horizontal curve such that $\Pi(\gamma) = C$. Then, $q(\gamma)$ is a horizontal curve in $q(\Gamma)$. For all $g \in \mathbb{H}$ and $t \in \mathbb{R}$, we have $\Pi(gY^t) = \Pi(g)$. Consequently, we have $\Pi(q(\gamma)) = \Pi(q(C)) = \hat{q}(C)$, and $\hat{q}(C)$ is characteristic for $q(\Gamma)$.

For any $v \in U$, we have $q(vY^{\psi(v)}) \in q(\Gamma)$, and since $q(\Gamma)$ is an intrinsic graph, we must have $q(vY^{\psi(v)}) = \hat{q}(v) Y^{\hat{\psi}(\hat{q}(v))}$. The claimed expressions for $\hat{\psi}$ follow directly. □

Observe that if $q: \mathbb{H} \to \mathbb{H}$ preserves cosets of $\langle Y \rangle$, then

$$q(\Pi(p)) \in q(p\langle Y \rangle) = q(p)\langle Y \rangle, \tag{2.23}$$

so $\Pi \circ q = \Pi \circ q \circ \Pi$. In particular, if q_1 and q_2 are stretch maps, shear maps, or left translations, then

$$\hat{q_1} \circ \hat{q_2} = \Pi \circ q_1 \circ \Pi \circ q_2 = \Pi \circ q_1 \circ q_2 = \widehat{q_1 \circ q_2}.$$

Consequently, if $a, b, c \in \mathbb{R}$ and $q(v) = Y^b Z^{-c} \tilde{A}_{2a}(v)$ for all $v \in \mathbb{H}$, then

$$\hat{q}(x, 0, z) = (x, 0, z - ax^2 - bx - c).$$

That is, for any quadratic function f, there is a map $q: \mathbb{H} \to \mathbb{H}$ so that the characteristic curves of $q(\Gamma)$ are the characteristic curves of Γ translated by f.

Finally, stretch maps and shear maps send intrinsic Lipschitz graphs to intrinsic Lipschitz graphs (with a possible change in the Lipschitz constant).

Lemma 2.10. *Let Γ be an intrinsic Lipschitz graph, and let $a, b \in \mathbb{R} \setminus \{0\}$. Then, $s_{a,b}(\Gamma)$ and $\tilde{A}_b(\Gamma)$ are intrinsic Lipschitz graphs, with an intrinsic Lipschitz constant depending on a and b, and the intrinsic Lipschitz constant of Γ.*

Proof. Let $q = s_{a,b}$ or $q = \tilde{A}_b$. As Γ is an intrinsic Lipschitz graph, there is a scale-invariant double cone $C \subseteq \mathbb{H}$ containing a neighborhood of Y such that $pC \cap \Gamma = \varnothing$ for all

$p\in\Gamma$. The image $q(C)$ is a scale-invariant double cone containing a neighborhood of Y. Since

$$\bigcap_{\lambda\in(0,1)} \mathrm{Cone}_\lambda = \langle Y\rangle\setminus\{\mathbf{0}\},$$

there is a $0<\lambda<1$ such that $\mathrm{Cone}_\lambda\subseteq q(C)$. For all $p\in\Gamma$,

$$q(p)\,\mathrm{Cone}_\lambda\cap q(\Gamma)\subseteq q(p)q(C)\cap q(\Gamma)=q(pC\cap\Gamma)=\varnothing,$$

so $q(\Gamma)$ is intrinsic λ-Lipschitz. □

2.5. Measures on lines and the kinematic formula

Let $\mathcal{L}$ be the space of horizontal lines in $\mathbb{H}$. For $U\subseteq\mathbb{H}$, denote the set of horizontal lines that intersect U by

$$\mathcal{L}(U)\stackrel{\mathrm{def}}{=}\{L\in\mathcal{L}: L\cap U\neq\varnothing\}.$$

Let $\mathcal{N}$ be the unique (up to constants) measure on $\mathcal{L}$ that is invariant under the action of the isometry group of $\mathbb{H}$. Scalings of horizontal lines are horizontal lines, so scaling automorphisms of $\mathbb{H}$ act on $\mathcal{L}$, and $\mathcal{L}(s_{t,t}(M))=t^3\mathcal{L}(M)$ for all $t>0$. Henceforth $\mathcal{N}$ will be normalized so that $\mathcal{N}(\mathcal{L}(B_r(x)))=r^3$ for every $r>0$ and $x\in\mathbb{H}$.

The Heisenberg group satisfies the following kinematic formula, which we record here for ease of later use (see [79] or [23, equation (6.1)]). There exists a constant $c>0$ such that, for any finite-perimeter set $E\subseteq\mathbb{H}$ and any open subset $U\subseteq\mathbb{H}$,

$$\mathrm{Per}_E(U)=c\int_{\mathcal{L}}\mathrm{Per}_{E\cap L}(U\cap L)\,d\mathcal{N}(L). \tag{2.24}$$

Consider also the set $\mathcal{L}^{\#}\stackrel{\mathrm{def}}{=}\{(L,p):L\in\mathcal{L}\text{ and }p\in L\}$ of *pointed horizontal lines*. Associate with each measurable subset $K\subseteq\mathcal{L}^{\#}$ the following two quantities:

$$\int_{\mathcal{L}}\mathcal{H}^1(\{p\in L:(L,p)\in K\})\,d\mathcal{N}(L) \tag{2.25}$$

and

$$\int_0^{2\pi}\int_{\mathbb{H}}\mathbf{1}_K(p\langle\cos(\theta)X+\sin(\theta)Y\rangle,p)\,d\mathcal{H}^4(p)\,d\theta. \tag{2.26}$$

Both of the expressions in (2.25) and (2.26) define measures on $\mathcal{L}^{\#}$ that are invariant under the isometry group of $\mathbb{H}$, which acts transitively on $\mathcal{L}^{\#}$. Therefore, they are proportional, and there is a constant $C>0$ such that, for every measurable $K\subseteq\mathcal{L}^{\#}$,

$$\int_{\mathcal{L}}\mathcal{H}^1(K_L)\,d\mathcal{N}(L)=C\int_0^{2\pi}\int_{\mathbb{H}}\mathbf{1}_K(L_{p,\theta},p)\,d\mathcal{H}^4(p)\,d\theta, \tag{2.27}$$

where we use the following notation for every $L\in\mathcal{L}$, $p\in\mathbb{H}$, and $\theta\in[0,2\pi]$:

$$K_L\stackrel{\mathrm{def}}{=}\{p\in L:(L,p)\in K\}\subseteq L\quad\text{and}\quad L_{p,\theta}\stackrel{\mathrm{def}}{=}p\langle\cos(\theta)X+\sin(\theta)Y\rangle\in\mathcal{L}. \tag{2.28}$$

2.6. Vertical perimeter and parametric vertical perimeter

Given a measurable subset $E\subseteq V_0$, a measurable function $\psi\colon V_0\to\mathbb{R}$, and (a scale) $a\in\mathbb{R}$, we define the (*normalized*) *parametric vertical perimeter at scale* a of ψ on E by

$$\bar{\mathsf{v}}^P_{E,\psi}(a)\stackrel{\mathrm{def}}{=}\frac{1}{2^{-a}}\int_E |\psi(v)-\psi(vZ^{-2^{-2a}})|\,d\mathcal{H}^3(v). \tag{2.29}$$

This notion relates to the usual vertical perimeter (1.30) of the epigraph of ψ as follows.

Lemma 2.11. (Parametric vertical perimeter versus vertical perimeter of epigraph) *For any measurable subset $E\subseteq V_0$, any measurable function $\psi\colon V_0\to\mathbb{R}$, and any $a\in\mathbb{R}$,*

$$\bar{\mathsf{v}}^P_{E,\psi}(a)=\bar{\mathsf{v}}_{\Pi^{-1}(E)}(\Gamma^+_\psi)(a).$$

Proof. Recalling (1.29), for $\Omega\subseteq\mathbb{H}$ and $a\in\mathbb{R}$ we denote $\mathsf{D}_a\Omega=\Omega\triangle\Omega Z^{2^{-2a}}$. Then,

$$\begin{aligned}\mathsf{D}_a\Gamma^+_\psi=&\{vY^t : v\in V_0 \text{ and } \psi(v)<t\leqslant\psi(vZ^{-2^{-2a}})\}\\ &\cup\{vY^t : v\in V_0 \text{ and } \psi(vZ^{-2^{-2a}})<t\leqslant\psi(v)\},\end{aligned}$$

since, by definition,

$$\Gamma^+_\psi Z^{2^{-2a}}=\{vY^t : v\in V_0 \text{ and } \psi(vZ^{-2^{-2a}})<t\}.$$

Therefore,

$$\begin{aligned}\bar{\mathsf{v}}_{\Pi^{-1}(E)}(\Gamma^+_\psi)(a)&=\frac{\mathcal{H}^4(\Pi^{-1}(E)\cap\mathsf{D}_a\Gamma^+_\psi)}{2^{-a}}\\ &=\frac{1}{2^{-a}}\int_E|\psi(v)-\psi(vZ^{-2^{-2a}})|\,d\mathcal{H}^3(v)=\bar{\mathsf{v}}^P_{E,\psi}(a),\end{aligned}$$

where the second equality uses the fact that the map

$$(x,y,z)\longmapsto(x,0,z)\cdot Y^y=\big(x,y,z+\tfrac12 xy\big)$$

has constant Jacobian 1. □

An advantage of the parametric vertical perimeter is that it increases or decreases by a constant factor under a stretch map or a shear map, as computed in the following.

Lemma 2.12. *Let $\psi\colon V_0\to\mathbb{R}$ and $E\subseteq V_0$ be measurable. Let $q\colon\mathbb{H}\to\mathbb{H}$, $\hat{q}\colon V_0\to V_0$, and $\hat{\psi}\colon V_0\to\mathbb{R}$ be as in Lemma 2.9, i.e., q is a stretch map or a shear map, $\hat{q}$ is the map induced on V_0, and $\hat{\psi}$ is the function such that $q(\Gamma_\psi)=\Gamma_{\hat{\psi}}$. Then, for all $t\in\mathbb{R}$, we have*

- *If $a,b\in\mathbb{R}\setminus\{0\}$ and $q=s_{a,b}$, then*

$$\bar{\mathsf{v}}^P_{\hat{q}(E),\hat{\psi}}(t)=|ab|^{3/2}\cdot\bar{\mathsf{v}}^P_{E,\psi}(t+\log_2\sqrt{|ab|}\,).$$

- *If $b\in\mathbb{R}\setminus\{0\}$ and $q=\tilde{A}_b$, then*

$$\bar{\mathsf{v}}^P_{\hat{q}(E),\hat{\psi}}(t)=\bar{\mathsf{v}}^P_{E,\psi}(t).$$

Proof. If $q=s_{a,b}$ for some $a,b\in\mathbb{R}\setminus\{0\}$, then

$$\hat{q}(x,0,z)=(ax,0,abz) \quad\text{and}\quad \hat{\psi}(\hat{q}(v))=b\psi(v)$$

for every $v=(x,0,z)\in V_0$. So,

$$\begin{aligned}\bar{\mathsf{v}}^P_{\hat{q}(E),\hat{\psi}}(t)&=2^t\int_{\hat{q}(E)}|\hat{\psi}(v)-\hat{\psi}(vZ^{-2^{-2t}})|\,d\mathcal{H}^3(v)\\&=2^t|b|\int_{\hat{q}(E)}|\psi(\hat{q}^{-1}(v))-\psi(\hat{q}^{-1}(v)Z^{-(ab)^{-1}2^{-2t}})|\,d\mathcal{H}^3(v)\\&=2^ta^2b^2\int_E|\psi(v)-\psi(vZ^{-(ab)^{-1}2^{-2t}})|\,d\mathcal{H}^3(v)\\&=|ab|^{3/2}\cdot\bar{\mathsf{v}}^P_{E,\psi}(t+\log_2\sqrt{|ab|}\,).\end{aligned}$$

Next, if $q=\tilde{A}_b$ for some $b\in\mathbb{R}\setminus\{0\}$, then $\hat{\psi}(\hat{q}(v))=\psi(v)+bx(v)$ for all $v=(x,0,z)\in E$, and by Lemma 2.8 we have

$$\hat{q}(v)=\Pi(q(x,0,z))=\big(x,0,z-\tfrac{1}{2}bx^2\big).$$

So,

$$\hat{\psi}(vZ^{-2t})=\psi(\hat{q}^{-1}(vZ^{-2^{-2t}}))+bx(\hat{q}^{-1}(vZ^{-2^{-2t}}))=\psi(\hat{q}^{-1}(v)Z^{-2^{-2t}})+bx(v),$$

and hence

$$\begin{aligned}\bar{\mathsf{v}}^P_{\hat{q}(E),\hat{\psi}}(t)&=2^t\int_{\hat{q}(E)}|\psi(\hat{q}^{-1}(v))-\psi(\hat{q}^{-1}(v)Z^{-2^{-2t}})|\,d\mathcal{H}^3(v)\\&=2^t\int_E|\psi(v)-\psi(vZ^{-2^{-2t}})|\,d\mathcal{H}^3(v)=\bar{\mathsf{v}}^P_{E,\psi}(t).\end{aligned}$$ □

We end this section by recording a straightforward a-priori upper bound on $\bar{\mathsf{v}}^P_{E,\psi}(a)$.

LEMMA 2.13. *Suppose that $E\subseteq V_0$ is measurable and $\psi\colon V_0\to\mathbb{R}$ is smooth. Then,*

$$\bar{\mathsf{v}}^P_{E,\psi}(a)\leqslant\min\left\{2^{a+1}\|\psi\|_{L_\infty(V_0)},2^{-a}\left\|\frac{\partial\psi}{\partial z}\right\|_{L_\infty(V_0)}\right\}\mathcal{H}^3(E)\quad\textit{for all } a\in\mathbb{R}.$$

Proof. For all $v=(x,0,z)\in E$, we (trivially) have

$$\begin{aligned}|\psi(v)-\psi(vZ^{-2^{-2a}})|&=|\psi(x,0,z)-\psi(x,0,z-2^{-2a})|\leqslant 2\|\psi\|_{L_\infty(V_0)},\\|\psi(v)-\psi(vZ^{-2^{-2a}})|&=|\psi(x,0,z)-\psi(x,0,z-2^{-2a})|\leqslant 2^{-2a}\left\|\frac{\partial\psi}{\partial z}\right\|_{L_\infty(V_0)}.\end{aligned}$$

Recalling the definition (2.29), we obtain the desired inequality by integrating over E. □

3. Constructing surfaces and embeddings

In this section, we will prove Proposition 3.4, following the reasoning sketched in §1.2.2, to construct surfaces that are α-far from planes at α^{-4} different scales. We use these surfaces to prove the following theorem.

THEOREM 3.1. *For any $k>1$, there is a left-invariant metric $\Delta=\Delta_k:\mathbb{H}\times\mathbb{H}\to[0,\infty)$ on $\mathbb{H}$ and a measure space $(\mathcal{S},\mu)$ such that $(\mathbb{H},\Delta)$ embeds isometrically in $L_1(\mu)$ and such that for any $h=(a,b,c)\in\mathbb{H}$ we have*

$$|a|+|b|\lesssim\Delta(\mathbf{0},h)\lesssim|a|+|b|+\frac{\min\{\sqrt{|c|},k\}}{\sqrt[4]{\log k}}. \tag{3.1}$$

If moreover $1\leqslant|c|\leqslant k^2$, then, in fact

$$\Delta(\mathbf{0},h)\asymp|a|+|b|+\frac{\sqrt{|c|}}{\sqrt[4]{\log k}}. \tag{3.2}$$

We will prove Theorem 3.1 in §3.1 after deriving two of its applications, and stating Proposition 3.4. The first application of Theorem 3.1 is the proof of Theorem 1.7.

Proof of Theorem 1.7 assuming Theorem 3.1. Letting Δ and $(\mathcal{S},\mu)$ be as in Theorem 3.1, fix $\xi:\mathbb{H}\to L_1(\mu)$ such that $\|\xi(g)-\xi(h)\|_{L_1(\mu)}=\Delta(g,h)$ for all $g,h\in\mathbb{H}$. Also, using [3], fix $m\in\mathbb{N}$ and $\varphi:\mathbb{H}\to\mathbb{R}^m$ such that $\|\varphi(g)-\varphi(h)\|_{\ell_1^m}\asymp\sqrt{d(g,h)}$ for all $g,h\in\mathbb{H}$.

Suppose that $\vartheta\geqslant\frac{1}{4}$. Consider the function $\tau:\mathbb{H}\to L_1(\mu)\oplus\mathbb{R}^2\oplus\mathbb{R}^m\cong L_1(\nu)$ (for a suitable measure ν) that is given by

$$\tau\stackrel{\mathrm{def}}{=}\frac{\xi}{(\log k)^{\vartheta-1/4}}\oplus\pi\oplus\frac{\varphi}{(\log k)^{\vartheta}}. \tag{3.3}$$

Since Δ is left-invariant, every $g=(x,y,z),h=(\chi,\upsilon,\zeta)\in\mathbb{H}$ with $1\leqslant d(g,h)\leqslant k$ satisfy

$$\|\tau(g)-\tau(h)\|_{L_1(\nu)}\asymp|x-\chi|+|y-\upsilon|+\frac{\sqrt{|2z-2\zeta-x\upsilon+y\chi|}}{(\log k)^{\vartheta}}, \tag{3.4}$$

using (2.6) and Theorem 3.1. While (3.4) would hold even without the third component of τ in (3.3), thanks to that component $\tau(\mathbb{H}_\mathbb{Z})$ is a locally-finite subset of $L_1(\nu)$. Every finite subset of $L_1(\nu)$ embeds with distortion $O(1)$ in ℓ_1 (by approximating by simple functions), so by [93], it follows that $\tau(\mathbb{H}_\mathbb{Z})$ admits a bi-Lipschitz embedding into ℓ_1 of distortion $O(1)$. As the word metric d_W on $\mathbb{H}_\mathbb{Z}$ is bounded above and below by universal constant multiples of d, this gives Theorem 1.7 provided k is a large enough universal constant multiple of n. $\square$

A second application of Theorem 3.1 is to construct a left-invariant metric on $\mathbb{H}_{\mathbb{Z}}$ with the properties of Theorem 1.9, at the cost of losing an iterated logarithm in the associated distortion bounds that we derived in the proof of Theorem 1.9. While the power of the iterated logarithm can be improved by taking more care in the ensuing reasoning, some unbounded lower-order loss *must* be incurred here; see Remark 3.3.

THEOREM 3.2. *For any $2<p\leqslant 4$ there is a left-invariant metric $\delta=\delta_p$ on $\mathbb{H}_{\mathbb{Z}}$ that admits a bi-Lipschitz embedding into both ℓ_1 and ℓ_q for all $q\geqslant p$, yet not into any Banach space whose modulus of uniform convexity has power-type r for $2\leqslant r<p$ (in particular, $(\mathbb{H}_{\mathbb{Z}},\delta)$ does not admit a bi-Lipschitz embedding into a Hilbert space or ℓ_s for $1<s<p$). Moreover, if we denote $\vartheta=1/p$, then for every $h=(a,b,c)\in\mathbb{H}_{\mathbb{Z}}$ with $|c|\geqslant 3$ we have*

$$\delta(\mathbf{0},h)\asymp |a|+|b|+\frac{\sqrt{|c|}}{(\log|c|)^{\vartheta}(\log\log|c|)^2}. \tag{3.5}$$

Proof. Define a left-invariant metric $\delta\colon\mathbb{H}_{\mathbb{Z}}\times\mathbb{H}_{\mathbb{Z}}\to[0,\infty)$ as a superposition of the metrics $\{\Delta_k\}_{k>0}$ of Theorem 3.1, by setting for every $h=(a,b,c)\in\mathbb{H}_{\mathbb{Z}}$,

$$\delta(\mathbf{0},h)\stackrel{\mathrm{def}}{=}\sum_{n=1}^{\infty}\frac{1}{n^2e^{(4\vartheta-1)n}}\Delta_{e^{e^{4n}}}(\mathbf{0},h). \tag{3.6}$$

We will first verify (3.5), which in particular implies that the sum defining δ converges, and hence by Theorem 3.1 we would know that δ is indeed a left-invariant metric on $\mathbb{H}_{\mathbb{Z}}$, and that $(\mathbb{H}_{\mathbb{Z}},\delta)$ admits an isometric embedding into $\ell_1(L_1(\mu))$. By [93], it follows from this that $(\mathbb{H}_{\mathbb{Z}},\delta)$ also admits a bi-Lipschitz embedding into the sequence space ℓ_1.

Fix $h=(a,b,c)\in\mathbb{H}_{\mathbb{Z}}$ with $|c|\geqslant e^{e^4}$, and choose $m=m(c)\in\mathbb{N}$ such that

$$e^{e^{4m}}\leqslant\sqrt{|c|}<e^{e^{4(m+1)}}. \tag{3.7}$$

Then,

$$\begin{aligned}\delta(\mathbf{0},h)&\stackrel{(3.1)}{\lesssim}|a|+|b|+\sum_{n=1}^{\infty}\frac{\min\{\sqrt{|c|},e^{e^{4n}}\}}{n^2e^{4\vartheta n}}\lesssim|a|+|b|+\sum_{n=1}^{m}\frac{e^{e^{4n}}}{n^2e^{4\vartheta n}}+\sum_{n=m+1}^{\infty}\frac{\sqrt{|c|}}{n^2e^{4\vartheta n}}\\&\asymp|a|+|b|+\frac{e^{e^{4m}}}{m^2e^{4\vartheta m}}+\frac{\sqrt{|c|}}{m^2e^{4\vartheta m}}\stackrel{(3.7)}{\lesssim}|a|+|b|+\frac{\sqrt{|c|}}{(\log|c|)^{\vartheta}(\log\log|c|)^2}.\end{aligned}$$

Conversely, since the sum in (3.6) is at least its summands for $n=1$ and $n=m+1$,

$$\delta(\mathbf{0},h)\gtrsim|a|+|b|+\frac{\sqrt{|c|}}{(m+1)^2e^{4\vartheta(m+1)}}\stackrel{(3.7)}{\gtrsim}|a|+|b|+\frac{\sqrt{|c|}}{(\log|c|)^{\vartheta}(\log\log|c|)^2}.$$

This is (3.5) if $|c| \geqslant e^{e^4}$, but then (3.5) follows formally in the remaining range $3 \leqslant |c| < e^{e^4}$ (simply use the triangle inequality to reduce the upper bound to the case of large enough $|c|$ that we just proved, and take only the $n=1$ summand in (3.6) for the lower bound).

By contrasting (3.5) with (1.9), we see that, for every integer $n \geqslant 3$,

$$\mathsf{c}_{(\mathcal{B}_n,\delta)}(\mathcal{B}_n, d_W) \lesssim (\log n)^{\vartheta} (\log\log n)^2. \tag{3.8}$$

At the same time, if $2 \leqslant r < p$ and X is a Banach space whose modulus of uniform convexity has power-type r, then by [58] we have

$$\mathsf{c}_X(\mathcal{B}_n, d_W) \gtrsim_X (\log n)^{1/r}. \tag{3.9}$$

By combining (3.8) and (3.9), we deduce that

$$\mathsf{c}_X(\mathcal{B}_n, \delta) \gtrsim_X \frac{(\log n)^{1/r-\vartheta}}{(\log\log n)^2} = \frac{(\log n)^{1/r-1/p}}{(\log\log n)^2} \xrightarrow[n\to\infty]{} \infty.$$

Consequently, $(\mathbb{H}_{\mathbb{Z}}, \delta)$ does not admit a bi-Lipschitz embedding into X.

It remains to show that $(\mathbb{H}_{\mathbb{Z}}, \delta)$ admits a bi-Lipschitz embedding into ℓ_q for any $q \geqslant p$. As before, finite subsets of L_q embed with distortion $O(1)$ in ℓ_p (by approximating by simple functions). Thus, due to [93], since $(\mathbb{H}_{\mathbb{Z}}, \delta)$ is locally finite, it suffices to show that $(\mathbb{H}_{\mathbb{Z}}, \delta)$ admits a bi-Lipschitz embedding into L_q. By [57, Lemma 3.1], for any $0 < \varepsilon < \frac{1}{2}$, there exists a left-invariant metric ρ_ε on $\mathbb{H}_{\mathbb{Z}}$ such that $(\mathbb{H}_{\mathbb{Z}}, \rho_\varepsilon)$ embeds isometrically into L_q, and

$$\rho_\varepsilon(\mathbf{0}, h) \asymp |a|^{1-\varepsilon} + |b|^{1-\varepsilon} + \varepsilon^{1/q} |c|^{(1-\varepsilon)/2} \quad \text{for all } h = (a, b, c) \in \mathbb{H}_{\mathbb{Z}}. \tag{3.10}$$

Define a left-invariant metric $\rho\colon \mathbb{H}_{\mathbb{Z}} \times \mathbb{H}_{\mathbb{Z}} \to [0, \infty)$ by setting, for every $h = (a, b, c) \in \mathbb{H}_{\mathbb{Z}}$,

$$\rho(\mathbf{0}, h) \stackrel{\mathrm{def}}{=} \bigg(|a|^q + |b|^q + \sum_{n=1}^{\infty} \frac{1}{n^{2q} e^{n(q\vartheta-1)}} \rho_{2e^{-n}}(\mathbf{0}, h)^q \bigg)^{1/q}.$$

By design, $(\mathbb{H}_{\mathbb{Z}}, \rho)$ embeds isometrically into $\ell_q(L_q)$. So, the proof of Theorem 3.2 will be complete if we show that $\delta(\mathbf{0}, h) \asymp \rho(\mathbf{0}, h)$ for all $h = (a, b, c) \in \mathbb{H}_{\mathbb{Z}}$ with, say, $|c| \geqslant 300$. To see this, by combining (3.5) and (3.10) it suffices to show that

$$\bigg(\sum_{n=1}^{\infty} \frac{1}{n^{2q} e^{nq\vartheta} |c|^{qe^{-n}}} \bigg)^{1/q} \asymp \frac{1}{(\log|c|)^{\vartheta} (\log\log|c|)^2}. \tag{3.11}$$

Fix $s=s(c)\in\mathbb{N}$ such that $2e^s\leqslant\log|c|<2e^{s+1}$ (this is possible because $|c|\geqslant 300$). Then,

$$\begin{aligned}&\bigg(\sum_{n=1}^{\infty}\frac{1}{n^{2q}e^{nq\vartheta}|c|^{qe^{-n}}}\bigg)^{1/q}\\ &\lesssim\bigg(\sum_{k=0}^{s-1}\frac{1}{(s-k)^{2q}e^{(s-k)q\vartheta}|c|^{qe^{-(s-k)}}}\bigg)^{1/q}+\bigg(\sum_{n=s+1}^{\infty}\frac{1}{n^{2q}e^{nq\vartheta}}\bigg)^{1/q}\\ &\asymp\frac{1}{e^{s\vartheta}}\bigg(\sum_{k=0}^{s-1}\frac{e^{qk\vartheta}}{(s-k)^{2q}(|c|^{e^{-s}})^{qe^k}}\bigg)^{1/q}+\frac{1}{s^2e^{s\vartheta}}\asymp\frac{1}{s^2e^{s\vartheta}}\asymp\frac{1}{(\log|c|)^{\vartheta}(\log\log|c|)^2},\end{aligned}$$

where the final step holds by our choice of s, and the penultimate step holds as $|c|^{e^{-s}}\geqslant 2e$ by our choice of s, and therefore the sum in question is dominated by its $k=0$ summand. This proves half of the equivalence (3.11), and the remaining direction of (3.11) follows by bounding from below the sum in the left-hand side of (3.11) by its $n=s$ summand. □

Remark 3.3. It is evident from the above proof of Theorem 3.2 that the power 2 of $\log\log|c|$ in (3.5) can be improved to any fixed power that is strictly larger than 1. However, the lower-order term cannot be removed altogether. Specifically, suppose that $\mathfrak{d}$ is a left-invariant metric on $\mathbb{H}_{\mathbb{Z}}$ such that every $h=(a,b,c)\in\mathbb{H}_{\mathbb{Z}}$ with $|c|\geqslant 3$ satisfies

$$\mathfrak{d}(\mathbf{0},h)\asymp|a|+|b|+\frac{\sqrt{|c|}}{\sqrt[4]{\log|c|}}. \tag{3.12}$$

We claim that neither ℓ_1 nor ℓ_4 contains a bi-Lipschitz copy of $(\mathbb{H}_{\mathbb{Z}},\mathfrak{d})$. In fact, we will next show that for every integer $n\geqslant 3$ the word-ball $\mathcal{B}_n\subseteq\mathbb{H}_{\mathbb{Z}}$ satisfies the distortion bounds

$$\sqrt[4]{\log\log n}\lesssim\mathsf{c}_{\ell_1}(\mathcal{B}_n,\mathfrak{d})\lesssim\log\log n \tag{3.13}$$

and

$$\mathsf{c}_{\ell_4}(\mathcal{B}_n,\mathfrak{d})\asymp\sqrt[4]{\log\log n}. \tag{3.14}$$

We conjecture that the first inequality in (3.13) is sharp.

To prove (3.13), by substituting Theorem 1.1 into [91, Lemma 33], and then substituting the resulting inequality into [91, Lemma 30], we get that there is a universal constant $\kappa\geqslant 5$ such that, for every integer $n\geqslant 3$, every function $f\colon\mathbb{H}_{\mathbb{Z}}\to\ell_1$ satisfies

$$\bigg(\sum_{c=1}^{n^2}\frac{1}{c^3}\bigg(\sum_{h\in\mathcal{B}_n}\|f(hZ^c)-f(h)\|_{\ell_1}\bigg)^4\bigg)^{1/4}\lesssim\sum_{h\in\mathcal{B}_{\kappa n}}(\|f(hX)-f(h)\|_{\ell_1}+\|f(hY)-f(h)\|_{\ell_1}). \tag{3.15}$$

Suppose that $D\geqslant 1$ is such that

$$\mathfrak{d}(g,h)\leqslant\|f(g)-f(h)\|_{\ell_1}\leqslant D\mathfrak{d}(g,h)\quad\text{for all } g,h\in\mathcal{B}_{2\kappa n}.$$

Then, by (3.12) and (3.15), we have

$$D \gtrsim \bigg(\sum_{c=3}^{n^2} \frac{1}{c^3} \bigg(\frac{\sqrt{c}}{\sqrt[4]{\log c}} \bigg)^4 \bigg)^{1/4} = \bigg(\sum_{c=3}^{n^2} \frac{1}{c \log c} \bigg)^{1/4} \asymp \sqrt[4]{\log \log n}. \tag{3.16}$$

This proves the first inequality in (3.13). For the second inequality in (3.13) consider the sum

$$\mathfrak{d}_{1,n} = \sum_{j=0}^{5\lceil \log\log n\rceil} \Delta_{2^{2^j}}$$

of metrics from Theorem 3.1. Then, by Theorem 3.1, the metric space $(\mathbb{H}_\mathbb{Z}, \mathfrak{d}_{1,n})$ embeds isometrically into ℓ_1 and $\mathfrak{d} \lesssim \mathfrak{d}_{1,n} \lesssim (\log\log n)\mathfrak{d}$ on $\mathcal{B}_n \times \mathcal{B}_n$.

The proof of (3.14) is analogous. For the lower bound on $\mathsf{c}_{\ell_4}(\mathcal{B}_n, \mathfrak{d})$, use (the case $q{=}4$ of) [58, Theorem 1.1] to get the following estimate for any function $f\colon \mathbb{H}_\mathbb{Z} \to \ell_4$:

$$\sum_{c=1}^{n^2} \frac{1}{c^3} \bigg(\sum_{h\in\mathcal{B}_n} \|f(hZ^c) - f(h)\|_{\ell_4} \bigg)^4 \lesssim \sum_{h\in\mathcal{B}_{21n}} (\|f(hX)-f(h)\|_{\ell_4}^2 + \|f(hY)-f(h)\|_{\ell_4}^4).$$

With this inequality at hand, the desired lower bound follows as in (3.16). For the upper bound on $\mathsf{c}_{\ell_4}(\mathcal{B}_n, \mathfrak{d})$, use the following metric on $\mathbb{H}_\mathbb{Z}$ which embeds isometrically into ℓ_4:

$$\mathfrak{d}_{4,n} = \bigg(\sum_{j=0}^{5\lceil \log\log n\rceil} \Delta_{2^{2^j}}^4 \bigg)^{1/4}.$$

The above reasoning also shows mutatis mutandis that an unbounded lower-order factor loss is needed in the compression bound (1.27) of Theorem 1.16. Specifically, there is no mapping $f\colon \mathbb{H}_\mathbb{Z} \to \ell_1$ that is Lipschitz with respect to the word metric on $\mathbb{H}_\mathbb{Z}$ and whose compression rate (recall (1.26)) satisfies $\omega_f(s) \gtrsim s/\sqrt[4]{\log s}$ when $s \geqslant 2$. It would be worthwhile to obtain a characterization of the possible compression rates of embeddings of $\mathbb{H}_\mathbb{Z}$ into ℓ_1 in the spirit of [91, Theorem 9], but this would require more work. Specifically, one would need to replace the use in [91] of [105, Corollary 5] by a better embedding of $\mathbb{H}_\mathbb{Z}$ into ℓ_1; we expect that the existence of such an embedding could be deduced using the ideas of the present section, but we did not attempt to carry this out.

The main ingredient in the proof of Theorem 3.1 is the following proposition, which is proved in §3.2. It constructs a function $\psi\colon V_0 \to \mathbb{R}$ whose intrinsic graph has small horizontal perimeter but large vertical perimeter due to bumps at many different scales. Here and throughout the rest of this section, we denote the unit square in V_0 by U, i.e.,

$$U \stackrel{\mathrm{def}}{=} [0,1] \times \{0\} \times [0,1] \subseteq V_0.$$

PROPOSITION 3.4. *There are universal constants $\rho, R, r\in\mathbb{R}$ with $R>r$ and $\rho>2^{2(R-r)}$ such that, for any $\alpha\in\mathbb{N}$, there is a smooth function $\psi: V_0\to\mathbb{R}$ that has the following properties.*

(1) *ψ is periodic with respect to the integer lattice $\mathbb{Z}\times\{0\}\times\mathbb{Z}$ of V_0.*

(2) *$\|\partial_\psi\psi\|_{L_2(U)}\lesssim 1$.*

(3) *$\|\psi\|_{L_\infty(V_0)}\leqslant 1/\alpha^2$.*

(4) *$\bar{\mathsf{v}}^P_{U,\psi}(a)\gtrsim 1/\alpha$ for any integer $0\leqslant n<\alpha^4$ and any $a\in I+\log_2(\alpha\rho^n)$, where $I=[r,R]$. Hence,*

$$\|\bar{\mathsf{v}}^P_{U,\psi}\|_{L_1([\log_2(\alpha\rho^n)+r,\log_2(\alpha\rho^n)+R])}\gtrsim\frac{1}{\alpha},$$

(5) *For any $q>0$, we have*

$$\|\bar{\mathsf{v}}^P_{U,\psi}\|_{L_q(\mathbb{R})}\gtrsim\alpha^{4/q-1}. \tag{3.17}$$

(6) *$\bar{\mathsf{v}}^P_{U,\psi}(a)\lesssim\min\{1/\alpha, 2^a/\alpha^2\}$ for any $a\in\mathbb{R}$.*

By Proposition 2.1, the second assertion of Proposition 3.4 implies that $\mathcal{H}^3(\partial E)\lesssim 1$, where E is the epigraph of the restriction of ψ to the unit square $U\subseteq V_0$. In combination with Proposition 3.4 (5), since α can be arbitrarily large, this shows that the $L_4(\mathbb{R})$ norm in (1.31) cannot be replaced by $L_q(\mathbb{R})$ for any $q\in(0,4)$; as explained in the introduction, this also implies the optimality of Theorem 1.1. Furthermore, since the L_q-variant of (1.31) is a consequence of the L_q-variant of (1.32), Proposition 3.4 also implies that, for any $q\in(0,4)$, there is $\lambda\in(0,1)$ such that, for any $c>0$, there is an intrinsic λ-Lipschitz graph Γ satisfying

$$\|\bar{\mathsf{v}}_{B_1(\mathbf{0})}(\Gamma^+)\|_{L_q(\mathbb{R})}\geqslant c.$$

We expect that the construction in §3.2 can be modified to produce an intrinsic Lipschitz graph directly (for instance, by stopping the construction early in regions where $\partial_\psi\psi$ gets too large), but this is not needed here, so we leave the details to future work.

Proposition 3.4 (5) follows directly from Proposition 3.4 (4). Indeed, since $\rho>2^{2(R-r)}$, the intervals $\{[\log_2(\alpha\rho^n)+r,\log_2(\alpha\rho^n)+R]\}_{n\in\mathbb{Z}}$ are disjoint. Consequently,

$$\begin{aligned}\|\bar{\mathsf{v}}^P_{U,\psi}\|^q_{L_q(\mathbb{R})}&\geqslant\sum_{n=0}^{\alpha^4-1}\|\bar{\mathsf{v}}^P_{U,\psi}\|^q_{L_q([\log_2(\alpha\rho^n)+r,\log_2(\alpha\rho^n)+R])}\\&\geqslant\sum_{n=0}^{\alpha^4-1}\frac{1}{(R-r)^{q-1}}\|\bar{\mathsf{v}}^P_{U,\psi}\|^q_{L_1([\log_2(\alpha\rho^n)+r,\log_2(\alpha\rho^n)+R])}\gtrsim\alpha^{4-q},\end{aligned} \tag{3.18}$$

where the penultimate step is an application of Jensen's inequality and the final step holds because $R-r>0$ is a constant and, by Proposition 3.4 (4), each of the summands is at least a universal constant multiple of α^{-q}.

3.1. Obtaining an embedding from an intrinsic graph

Here we show how Theorem 3.1 follows from Proposition 3.4. Let $\rho, r, R>0$ be the universal constants of Proposition 3.4. Without loss of generality, we may take $k>8$. Let $\alpha\in\mathbb{N}$ be the unique integer satisfying

$$\sqrt[4]{\log_\rho\left(\tfrac{1}{8}k\right)}\leqslant \alpha < 1+\sqrt[4]{\log_\rho\left(\tfrac{1}{8}k\right)}. \tag{3.19}$$

Let $\psi=\psi_\alpha$ be the function produced by Proposition 3.4. Write $\Gamma=\Gamma_\psi$ and $\Gamma^+=\Gamma^+_\psi$. Denote by $A\subseteq V_0\cap\mathbb{H}_\mathbb{Z}$ the discrete subgroup that is generated by X and Z, so that as a subset of $\mathbb{R}^3$ we have $A=\mathbb{Z}\times\{0\}\times\mathbb{Z}$. For every $p\in\mathbb{H}$ define, for all $h_1,h_2\in\mathbb{H}$,

$$\lambda_p(h_1,h_2)\stackrel{\mathrm{def}}{=}|\mathbf{1}_{p^{-1}\Gamma^+}(h_1)-\mathbf{1}_{p^{-1}\Gamma^+}(h_2)|=\begin{cases}1, & \text{if } |\{ph_1,ph_2\}\cap\Gamma^+|=1,\\ 0, & \text{otherwise.}\end{cases}$$

By the A-periodicity of ψ, we have $a\Gamma=\Gamma$ and $\lambda_{ap}(h_1,h_2)=\lambda_p(h_1,h_2)$ for all $a\in A$ and $p,h_1,h_2\in\mathbb{H}$. We can therefore define λ_p also when p is an equivalence class in the quotient $A\backslash\mathbb{H}$. Consider the following fundamental domain for A:

$$P\stackrel{\mathrm{def}}{=}\{X^aZ^cY^b: a,c\in[0,1) \text{ and } b\in\mathbb{R}\}=\left\{\left(a,b,c+\tfrac{1}{2}ab\right):(a,b,c)\in[0,1)\times\mathbb{R}\times[0,1)\right\}.$$

We may define $l:\mathbb{H}\times\mathbb{H}\to[0,\infty)$ by

$$l(h_1,h_2)\stackrel{\mathrm{def}}{=}\int_{A\backslash\mathbb{H}}\lambda_p(h_1,h_2)\,d\mathcal{H}^4(p)=\int_P\lambda_p(h_1,h_2)\,d\mathcal{H}^4(p).$$

Since $\mathbb{H}$ is a unimodular group (namely, one directly checks that the Lebesgue measure $\mathcal{H}^4$ is a bi-invariant Haar measure on $\mathbb{H}$), and $\lambda_p(gh_1,gh_2)=\lambda_{pg}(h_1,h_2)$, we have

$$l(gh_1,gh_2)=l(h_1,h_2)\quad\text{for all } g,h_1,h_2\in\mathbb{H},$$

i.e., l is a left-invariant semi-metric on $\mathbb{H}$.

LEMMA 3.5. *For every $a\in\mathbb{R}$ we have $l(\mathbf{0},Z^{2^{-2a}})=2^{-a}\cdot\bar{\mathrm{v}}^P_{U,\psi}(a)$.*

Proof. For $v\in V_0$ and $b\in\mathbb{R}$, we have $vY^b\in\Gamma^+$ if and only if $b>\psi(v)$. So, for any $c>0$,

$$\lambda_{vY^b}(\mathbf{0},Z^c)=\lambda_{\mathbf{0}}(vY^b,vZ^cY^b)=\begin{cases}1, & \text{if } \psi(v)<b\leqslant\psi(vZ^c) \text{ or } \psi(vZ^c)<b\leqslant\psi(v).\\ 0, & \text{otherwise.}\end{cases}$$

Consequently,

$$\int_\mathbb{R}\lambda_{vY^b}(\mathbf{0},Z^c)\,db=|\psi(vZ^c)-\psi(v)|.$$

Therefore, fixing $a\in\mathbb{R}$ and denoting $c=2^{-2a}$, we see that

$$\begin{aligned}l(\mathbf{0},Z^{2^{-2a}})&=\int_P\lambda_p(\mathbf{0},Z^c)\,dp=\int_\mathbb{R}\int_U\lambda_{vY^b}(\mathbf{0},Z^c)\,dv\,db\\&=\int_U|\psi(vZ^c)-\psi(v)|\,dv=2^{-a}\cdot\bar{\mathrm{v}}^P_{U,\psi}(a).\end{aligned}$$

□

For every $\theta\in[0,2\pi)$ let $R_\theta:\mathbb{H}\to\mathbb{H}$ be the rotation around the z-axis by angle θ. Define the following left-invariant semi-metric on $\mathbb{H}$, which is also (by design) invariant under the family of $\{R_\theta:\theta\in[0,2\pi)\}$ automorphisms of $\mathbb{H}$:

$$M(h_1,h_2)\stackrel{\mathrm{def}}{=}\int_0^{2\pi} l(R_\theta(h_1),R_\theta(h_2))\,d\theta \quad \text{for all } h_1,h_2\in\mathbb{H}.$$

LEMMA 3.6. *For every $w\in\mathsf{H}$ we have $M(\mathbf{0},w)\lesssim\|w\|$.*

Proof. By the rotation-invariance of M, it suffices to show that $M(\mathbf{0},X^t)\lesssim|t|$ for all t. In fact, by the left-invariance of M and the triangle inequality, it suffices to prove that $M(X^t,X^{-t})\lesssim t$ for $0<t<\frac14$.

Let $L_0=\langle X\rangle\subseteq\mathbb{H}$ be the x-axis. Recall that $L_{p,\theta}=pR_\theta(L_0)$ for $p\in\mathbb{H}$ and $\theta\in[0,2\pi)$. The map $(p,\theta)\mapsto(L_{p,\theta},p)$ is a bijection between $\mathbb{H}\times[0,\pi)$ and the set of pointed lines $\mathcal{L}^\#=\{(L,p):L\in\mathcal{L}\text{ and }p\in L\}$.

By the above definitions, we have

$$M(X^{-t},X^t)=\int_0^{2\pi}\int_P\lambda_p(R_\theta(X^{-t}),R_\theta(X^t))\,d\mathcal{H}^4(p)\,d\theta.$$

Let $K\subseteq P\times[0,2\pi)$ be the set of pairs (p,θ) such that $L_{p,\theta}$ intersects Γ transversally, i.e., $L_{p,\theta}$ crosses the tangent plane of Γ at every intersection. Since Γ is smooth, the complement of K has measure zero.

Let $U'=B_8(\mathbf{0})\cap V_0$. Let $(p,\theta)\in K$ be such that $\lambda_p(R_\theta(X^{-t}),R_\theta(X^t))\neq0$. Then, the line segment from $pR_\theta(X^{-t})$ to $pR_\theta(X^t)$ crosses Γ at some point $g\in\Gamma$; we claim that $\Pi(g)\in U'$.

By Proposition 3.4 (3), we have $\|\psi\|_{L_\infty(V_0)}\leqslant1$, so $|y(g)|\leqslant1$ and $|y(p)|\leqslant|y(g)|+t\leqslant2$. Since $p\in P$, there are $a,b,c\in\mathbb{R}$ such that $p=X^aZ^cY^b$, and these satisfy $|a|\leqslant1$, $|c|\leqslant1$, and $|b|=|y(p)|\leqslant2$. By (2.6),

$$d(\mathbf{0},g)\leqslant|a|+4\sqrt{|c|}+|b|\leqslant7$$

and

$$d(\mathbf{0},\Pi(g))\leqslant d(\mathbf{0},g)+|y(g)|\leqslant8,$$

so $\Pi(g)\in U'$.

Let $\Gamma(U')=\Gamma\cap\Pi^{-1}(U')=\Gamma_\psi|_{U'}$ and, for $L\in\mathcal{L}$, let

$$I_L=\{p\in L:d(p,L\cap\Gamma(U'))\leqslant t\}.$$

We have seen above that, if $(p,\theta)\in K$ and $\lambda_p(R_\theta(X^{-t}),R_\theta(X^t))\neq0$, then there is some $g\in L_{p,\theta}\cap\Gamma(U')$ such that $d(p,g)\leqslant t$. That is, $p\in I_{L_{p,\theta}}$. Furthermore, if L intersects Γ transversally, then

$$\mathcal{H}^1(I_L)\leqslant 2t|L\cap\Gamma(U')|=2t\,\mathrm{Per}_{\Gamma^+\cap L}(\Pi^{-1}(U')).$$

Hence,

$$M(X^{-t},X^t)=\int_0^{2\pi}\int_P \lambda_p(R_\theta(X^{-t}),R_\theta(X^t))\,d\mathcal{H}^4(p)\,d\theta \overset{(2.27)}{\lesssim} \int_{\mathcal{L}} \mathcal{H}^1(I_L)\,d\mathcal{N}(L)$$
$$\lesssim t\int_{\mathcal{L}} \mathrm{Per}_{\Gamma^+\cap L}(\Pi^{-1}(U'))\,d\mathcal{N}(L) \overset{(2.24)}{\asymp} t\,\mathrm{Per}_{\Gamma^+}(\Pi^{-1}(U'))\lesssim t,$$

where $\mathrm{Per}_{\Gamma^+}(\Pi^{-1}(U'))\lesssim 1$, by Proposition 2.1 and Proposition 3.4 (2). □

Next, define a left-invariant semi-metric Λ on $\mathbb{H}$ by

$$\Lambda(h_1,h_2)\overset{\mathrm{def}}{=}\int_{r-\log_2\rho}^{R+\log_2\rho} 2^a M(s_{2^{-a}}(h_1),s_{2^{-a}}(h_2))\,da \quad \text{for all } h_1,h_2\in\mathbb{H}.$$

Lemma 3.7. *For all $c>0$ we have*

$$\Lambda(\mathbf{0},Z^c)=\Lambda(\mathbf{0},Z^{-c})\lesssim \min\left\{\frac{\sqrt{c}}{\alpha},\frac{1}{\alpha^2}\right\}.$$

Also, for all

$$\frac{1}{\alpha^2\rho^{2\alpha^4}}\leqslant c\leqslant \frac{1}{\alpha^2},$$

we have

$$\Lambda(\mathbf{0},Z^c)=\Lambda(\mathbf{0},Z^{-c})\gtrsim \frac{\sqrt{c}}{\alpha}.$$

Proof. Write $c=2^{-2t}$ for some $t\in\mathbb{R}$. Since Λ is a left-invariant metric,

$$\Lambda(\mathbf{0},Z^c)=\Lambda(\mathbf{0},Z^{-c}).$$

By Lemma 3.5, we have the following identity:

$$\Lambda(\mathbf{0},Z^c)=2\pi\int_{r-\log_2\rho}^{R+\log_2\rho} 2^a l(\mathbf{0},Z^{2^{-2(t+a)}})\,da=2\pi 2^{-t}\int_{r-\log_2\rho}^{R+\log_2\rho} \bar{\mathrm{v}}_{U,\psi}^P(t+a)\,da. \tag{3.20}$$

So,

$$\Lambda(\mathbf{0},Z^c)\lesssim \min\left\{\frac{\sqrt{c}}{\alpha},\frac{1}{\alpha^2}\right\}$$

for $c\in(0,\infty)$, by (3.20) and the final assertion of Proposition 3.4.

If

$$\frac{1}{\alpha^2\rho^{2\alpha^4}}\leqslant c\leqslant \frac{1}{\alpha^2},$$

then $t\in[\log_2(\alpha\rho^n),\log_2(\alpha\rho^{n+1})]$ for some integer $0\leqslant n<\alpha^4$. Hence,

$$[t-\log_2\rho+r,t+\log_2\rho+R]\supseteq[\log_2(\alpha\rho^n)+r,\log_2(\alpha\rho^n)+R],$$

so (3.20) implies that

$$\Lambda(\mathbf{0},Z^c)\geqslant 2\pi 2^{-t}\|\bar{\mathrm{v}}_{U,\psi}^P\|_{L_1([\log_2(\alpha\rho^n)+r,\log_2(\alpha\rho^n)+R])}\gtrsim \frac{\sqrt{c}}{\alpha},$$

where the final step is the third assertion of Proposition 3.4 (and the definition of t). □

LEMMA 3.8. *$\Lambda(h_1,h_2)\lesssim d(h_1,h_2)$ for all $h_1,h_2\in\mathbb{H}$.*

Proof. By Lemma 3.6, we have $M(\mathbf{0},X^t)\lesssim|t|$ for any $t\in\mathbb{R}$, so

$$\Lambda(\mathbf{0},X^t)=\int_{r-\log_2\rho}^{R+\log_2\rho} 2^a M(\mathbf{0},X^{2^{-a}t})\,da\lesssim t(R-r+2\log_2\rho)\lesssim|t|.$$

Therefore, also $\Lambda(\mathbf{0},Y^t)=\Lambda(\mathbf{0},X^t)\lesssim|t|$, by the rotation-invariance of Λ. Since Λ is left-invariant, it suffices to show that $\Lambda(\mathbf{0},h)\lesssim d(\mathbf{0},h)$ for all $h\in\mathbb{H}$. Any $h\in\mathbb{H}$ can be written as $h=X^aY^b[X^c,Y^c]$ for $a,b,c\in\mathbb{R}$ satisfying $|a|,|b|,|c|\lesssim d(\mathbf{0},h)$, so

$$\Lambda(\mathbf{0},h)\leqslant\Lambda(\mathbf{0},X^a)+\Lambda(\mathbf{0},Y^b)+2\Lambda(\mathbf{0},X^c)+2\Lambda(\mathbf{0},Y^c)\lesssim d(\mathbf{0},h). \qquad\square$$

Proof of Theorem 3.1. Define a semi-metric Δ on $\mathbb{H}$ by setting, for every $h_1,h_2\in\mathbb{H}$,

$$\Delta(h_1,h_2)\overset{\text{def}}{=}k\alpha\Lambda(s_{1/k\alpha}(h_1),s_{1/k\alpha}(h_2))+\sqrt{(x(h_1)-x(h_2))^2+(y(h_1)-y(h_2))^2}. \tag{3.21}$$

Observe that $(\mathbb{H},\Delta)$ embeds isometrically into L_1, because Λ is an integral of so-called cut semi-metrics (see e.g. [28, 4.1] for the definition). Such semi-metrics embed isometrically into $\mathbb{R}$, so an integral of cut semi-metrics embeds isometrically in L_1. By construction, Λ is both left-invariant and invariant under the rotations $\{R_\theta:\theta\in[0,2\pi]\}$.

Suppose that $v=(a,b,c)\in\mathbb{H}$ and let $w=(a,b,0)$ be such that $w\in\mathsf{H}$ and $v=wZ^c$. By Lemma 3.6 and the second part of Lemma 3.7, we have

$$|a|+|b|\lesssim\Delta(\mathbf{0},v)\leqslant\Delta(\mathbf{0},w)+\Delta(\mathbf{0},Z^c)\lesssim|a|+|b|+\frac{\min\{\sqrt{|c|},k\}}{\alpha}.$$

Recalling that $\alpha\asymp\sqrt[4]{\log k}$ is given in (3.19), this establishes (3.1).

To prove Theorem 3.1, it therefore remains to establish (3.2), i.e.,

$$1\leqslant|c|\leqslant k^2\quad\text{implies}\quad\Delta(\mathbf{0},v)\gtrsim|a|+|b|+\frac{\sqrt{|c|}}{\alpha}. \tag{3.22}$$

By Lemma 3.8, there is $L>0$ such that $\Delta(h_1,h_2)\leqslant Ld(h_1,h_2)$ for any $h_1,h_2\in\mathbb{H}$. By the first part of Lemma 3.7, there is $C>0$ such that $\Delta(\mathbf{0},Z^c)\geqslant C\sqrt{c}/\alpha$ for all $1\leqslant c\leqslant k^2$. On one hand, if $\|w\|\geqslant C\sqrt{|c|}/2L\alpha$, then $\Delta(\mathbf{0},v)\geqslant\|w\|\asymp|a|+|b|+\sqrt{|c|}/\alpha$. On the other hand, if $\|w\|<C\sqrt{|c|}/2L\alpha$, then

$$\Delta(\mathbf{0},v)\geqslant\Delta(\mathbf{0},Z^c)-\Delta(\mathbf{0},w)\geqslant\frac{C\sqrt{c}}{\alpha}-L\|w\|\geqslant\frac{C\sqrt{|c|}}{2\alpha}\gtrsim|a|+|b|+\frac{\sqrt{|c|}}{\alpha}.$$

In either case, (3.22) holds. $\square$

3.2. Constructing a bumpy intrinsic graph

In this section, we prove Proposition 3.4. We start with a brief overview of our strategy. As sketched in §1.2.2, we will prove Proposition 3.4 by constructing a smooth function $\psi: V_0 \to \mathbb{R}$ whose intrinsic graph is roughly α^{-1}-far from a vertical plane at α^4 different scales. Specifically, for a suitable choice of universal constant $\rho>1$, we will construct ψ as a sum $\psi=\sum_{i=0}^{\alpha^4-1} \beta_i$. Each of the summands $\beta_i: V_0 \to \mathbb{R}$ will itself be a sum of smooth bump functions of amplitude $\|\beta_i\|_{L_\infty(V_0)} \asymp \alpha^{-2}\rho^{-i}$ that are supported on regions whose width (x-coordinate) is ρ^{-i} and whose height (z-coordinate) is roughly $\alpha^{-2}\rho^{-2i}$; their aspect ratio is therefore roughly

$$\frac{\rho^{-i}}{\sqrt{\alpha^{-2}\rho^{-2i}}} \asymp \alpha.$$

These regions cover V_0 and have disjoint interiors. We will see that the bumpiness of β_i at scale $\alpha^{-2}\rho^{-2i}$ implies the desired lower bounds on $\bar{\mathrm{v}}^P_{U,\psi}(t)$ when t is near $\log_2(\alpha\rho^i)$.

In order to ensure that $\|\partial_\psi \psi\|_{L_2(U)}$ is bounded, we construct β_i iteratively. For $i \in \mathbb{N}$, we set $\psi_i=\sum_{j=0}^{i-1} \beta_j$ and align the long axis of the bump functions making up β_i with the characteristic curves of Γ_{ψ_i}. This ensures that the characteristic curves of Γ_ψ cross the bumps from left to right. Since $\partial_\psi f$ measures the change in $f: V_0 \to \mathbb{R}$ along the characteristic curves of Γ_ψ and each bump has amplitude roughly $\alpha^{-2}\rho^{-i}$ and width ρ^{-i}, we have $|\partial_\psi \beta_i| \lesssim \alpha^{-2}\rho^{-i}/\rho^{-i} \asymp \alpha^{-2}$.

This iterative procedure is one of the motivations for the definition of a foliated corona decomposition. A foliated corona decomposition of an arbitrary intrinsic graph Γ can be viewed as a sequence of partitions of V_0 into regions as above, where the pieces of the partition are aligned with the characteristic curves of Γ. One can use these partitions to reconstruct Γ as a sum of perturbations, just as we constructed ψ as a sum of bump functions. Theorem 1.18 then states that any intrinsic Lipschitz graph can be constructed by such a process.

This construction also demonstrates the importance of the aspect ratio. If the construction is modified so that the bump functions making up β_i are supported on regions of aspect ratio α_i, then $\|\partial_\psi \beta_i\|_{L_2(U)} \asymp \alpha_i^{-2}$. If the scales of the bump functions are sufficiently separated, then $\{\partial_\psi \beta_i\}_{i \geqslant 0}$ are roughly orthogonal in $L_2(U)$ and

$$\|\partial_\psi \psi\|^2_{L_2(U)} \asymp \sum_{i \geqslant 0} \|\partial_\psi \beta_i\|^2_{L_2(U)} \asymp \sum_{i \geqslant 0} \alpha_i^{-4}.$$

For ψ to be intrinsic λ-Lipschitz, we must have $\|\partial_\psi \psi\|^2_{L_2(U)} \lesssim_\lambda 1$, which necessitates that

$$\sum_i \alpha_{i \geqslant 0}^{-4} \lesssim_\lambda 1.$$

This motivates the $\alpha(Q)^{-4}$ factor in the weighted Carleson condition (1.34).

We next set some notation in preparation for the proof of Proposition 3.4. If the function $\psi: V_0 \to \mathbb{R}$ is smooth, then the vector field

$$M_\psi \stackrel{\mathrm{def}}{=} \frac{\partial}{\partial x} - \psi \frac{\partial}{\partial z}$$

corresponding to ∂_ψ is smooth (recall the definitions in §2.2). The flow lines of M_ψ are the characteristic curves of Γ_ψ, which foliate V_0 (recall the terminology in §2.3). For $s \in \mathbb{R}$, let $\Phi(\psi)_s: V_0 \to V_0$ be the flow of M_ψ, so that $\Phi(\psi)_0 = \mathrm{id}_{V_0}$ and such that for any $v \in V_0$, the curve $s \mapsto \Phi(\psi)_s(v)$ is a characteristic curve of Γ_ψ.

Denote $\psi_0 \equiv 0$ and let $\Gamma_0 = \Gamma_{\psi_0} = V_0$. This function and graph are periodic with respect to $\mathbb{Z} \times \{0\} \times \mathbb{Z}$ and ψ_0 is zero on ∂U. Suppose that $i \geqslant 0$ and that $\psi_i: V_0 \to \mathbb{R}$ is smooth, periodic with respect to $\mathbb{Z} \times \{0\} \times \mathbb{Z}$, and zero on ∂U. We construct $\psi_{i+1}: V_0 \to \mathbb{R}$ as follows. Let

$$G_i \stackrel{\mathrm{def}}{=} \{(m\rho^{-i}, 0, n\alpha^{-2}\rho^{-2i}) : m, n \in \mathbb{Z}\} \subseteq V_0. \tag{3.23}$$

Label the points in G_i arbitrarily as $v_{i,1}, v_{i,2}, \dots$, and note that the points $U \cap \{v_{i,1}, v_{i,2}, \dots\}$ form a $\rho^i \times \alpha^2 \rho^{2i}$ grid in U. For each $j \in \mathbb{N}$ and $s, t \in \mathbb{R}$, define

$$R_{i,j}(s,t) \stackrel{\mathrm{def}}{=} \Phi(\psi_i)_s(v_{i,j} Z^t) \in V_0. \tag{3.24}$$

Each $R_{i,j}$ is a diffeomorphism from $\mathbb{R}^2$ to V_0. For any $s_0, t_0 \in \mathbb{R}$, the image $R_{i,j}(s_0 \times \mathbb{R})$ is a vertical line and $R_{i,j}(\mathbb{R} \times t_0)$ is a characteristic curve of Γ_{ψ_i}. Using the terminology of foliated patchworks that we will introduce in §4, the map $R_{i,j}$ sends rectangles in V_0 to pseudoquads of Γ_{ψ_i} (regions in V_0 that are bounded by characteristic curves of Γ_{ψ_i} above and below, and by vertical line segments on either side). Denote

$$Q_{i,j} \stackrel{\mathrm{def}}{=} R_{i,j}([0, \rho^{-i}] \times [0, \alpha^{-2}\rho^{-2i}]) \subseteq V_0. \tag{3.25}$$

Thus, $Q_{i,j}$ is a pseudoquad whose lower-left corner is $v_{i,j}$. The sets $Q_{i,1}, Q_{i,2}, \dots$ cover V_0 and have disjoint interiors. They are obtained by cutting V_0 into vertical strips of width ρ^{-i}, then cutting each vertical strip along characteristic curves separated by $\alpha^{-2}\rho^{-2i}$.

Since ψ_i is zero on ∂U, the top and bottom edges of U are characteristic curves of Γ_{ψ_i}. The bottom boundary of each $Q_{i,0}$ and the top boundary of $Q_{i,\alpha^2\rho^{2i}-1}$ thus lie in ∂U, and the $Q_{i,j}$'s partition U (up to overlap on boundaries). In particular, the resulting partition of V_0 is periodic with respect to $\mathbb{Z} \times \{0\} \times \mathbb{Z}$.

Note, however, that the $Q_{i,j}$'s from one step in this construction generally do not partition the $Q_{i,j}$'s from another step. One can modify the construction so that the partitions in each step are nested, as in Figure 2, but it requires some additional care.

Let $\beta\colon V_0\to\mathbb{R}$ be a smooth function supported on the unit square U such that β is not identically zero and its partial derivatives of order at most 2 are all in the interval $[-1,1]$. Fix also $\alpha,\rho\in\mathbb{N}$ with $\rho>1$. Define $\beta_{i,j}\colon V_0\to\mathbb{R}$ by setting it to be zero on $V_0\setminus Q_{i,j}$ and, for all $R_{i,j}(s,t)\in Q_{i,j}$,

$$\beta_{i,j}(R_{i,j}(s,t))\stackrel{\mathrm{def}}{=}\alpha^{-2}\rho^{-i}\beta(\rho^i s,0,\alpha^2\rho^{2i}t). \tag{3.26}$$

Thus, $\beta_{i,j}$ is a bump function supported on $Q_{i,j}$. Write

$$\beta_i\stackrel{\mathrm{def}}{=}\sum_{j=1}^{\infty}\beta_{i,j}, \tag{3.27}$$

and

$$\psi_{i+1}\stackrel{\mathrm{def}}{=}\psi_i+\beta_i. \tag{3.28}$$

Since $Q_{i,1},Q_{i,2},\ldots$ have disjoint interiors,

$$\|\psi_{i+1}\|_{L_\infty(V_0)}\leqslant\|\psi_i\|_{L_\infty(V_0)}+\alpha^{-2}\rho^{-i},$$

so by induction we have

$$\|\psi_i\|_{L_\infty(V_0)}\leqslant\frac{\alpha^{-2}}{\rho-1}\leqslant\alpha^{-2}. \tag{3.29}$$

Since the $Q_{i,j}$'s form a periodic partition of V_0, ψ_{i+1} is periodic. Since ∂U is contained in the boundaries of the $Q_{i,j}$, we have $\psi_{i+1}|_{\partial U}=\psi_i|_{\partial U}=0$.

Thus, by induction, for any integer $i\geqslant 0$, ψ_i satisfies the first and third assertions (periodicity and L_∞ boundedness) of Proposition 3.4. We will show that if ρ is large enough (depending only on β), then $\psi=\psi_{\alpha^4}$ satisfies the remaining assertions of Proposition 3.4, namely, the stated upper bounds on $\partial_\psi\psi$ and lower bounds on $\bar{\mathrm{v}}^P_{U,\psi}(a)$.

3.2.1. The horizontal perimeter of Γ_{ψ_i}

In this section, we prove the second assertion of Proposition 3.4 by bounding $\|\partial_{\psi_i}\psi_i\|_{L_2(U)}$. This bound, combined with Proposition 2.1, gives an upper bound on $\mathcal{H}^3(\Gamma_{\psi_i|_U})$.

Write for simplicity $\partial_i\stackrel{\mathrm{def}}{=}\partial_{\psi_i}$ and let $D_i\stackrel{\mathrm{def}}{=}\partial_{i+1}\psi_{i+1}-\partial_i\psi_i$. For $f,g\in L_2(U)$ we write

$$\langle f,g\rangle_U\stackrel{\mathrm{def}}{=}\int_U fg\,d\mathcal{H}^3.$$

Lemma 3.9. *For every $\rho\geqslant 5$ and $\alpha\geqslant 1$,*

$$\|D_i\|_{L_\infty(V_0)}\lesssim\alpha^{-2}\quad\textit{for all } i\in\mathbb{N},$$

and

$$|\langle D_m,D_n\rangle_U|\lesssim\alpha^{-4}\rho^{m-n}\quad\textit{for all } m,n\in\mathbb{N}.$$

Note that Lemma 3.9 implies that, for every $i\in\mathbb{N}$,

$$\|\partial_{\psi_i}\psi_i\|_{L_2(U)} \lesssim \frac{\sqrt{i}}{\alpha^2}. \tag{3.30}$$

Thus, $\|\partial_{\psi_i}\psi_i\|_{L_2(U)}\lesssim 1$ for $i\lesssim\alpha^4$, i.e., the second assertion of Proposition 3.4 holds true. To deduce (3.30) from Lemma 3.9, write

$$\partial_{\psi_i}\psi_i = \sum_{n=0}^{i-1} D_n, \tag{3.31}$$

and expand the squares to get

$$\begin{aligned}\|\partial_{\psi_i}\psi_i\|_{L_2(U)}^2 &= \sum_{n=0}^{i-1}\|D_n\|_{L_2(U)}^2 + 2\sum_{m=0}^{i-1}\sum_{n=m+1}^{i-1}\langle D_m, D_n\rangle \\ &\lesssim \sum_{n=0}^{i-1}\alpha^{-4} + \sum_{m=0}^{i-1}\sum_{k=1}^{\infty}\alpha^{-4}\rho^{-k} \asymp i\alpha^{-4},\end{aligned}$$

where the penultimate step is Lemma 3.9 and the final step holds because $\rho\geqslant 2$.

Fix an integer $i\geqslant 0$ and note that

$$D_i = \partial_{i+1}\psi_{i+1} - \partial_i\psi_i = (\partial_{i+1}-\partial_i)\psi_{i+1} + \partial_i\beta_i = -\beta_i\frac{\partial\psi_{i+1}}{\partial z} + \partial_i\beta_i. \tag{3.32}$$

We will prove Lemma 3.9 by bounding the terms in the right-hand side of (3.32) separately. To this end, it will be convenient to define as follows a system of *flow coordinates* on $Q_{i,j}$.

Fix $i\in\mathbb{N}\cup\{0\}$ and $j\in\mathbb{N}$. Write for simplicity $(x_0,0,z_0)=v_{i,j}$, $Q=Q_{i,j}$, and $R=R_{i,j}$. Denote $R^{-1}=(s,t)\colon Q\to\mathbb{R}^2$ and let $(x,0,z)\colon Q\to\mathbb{R}^2$ be the standard coordinate system. Then, s and t are functions of x and z and, conversely, x and z are functions of s and t. Recalling the differential equation (2.17) for characteristic curves, we have

$$x = x_0 + s \quad\text{and}\quad z = z_0 + t - \int_0^s \psi_i(R(\sigma,t))\,d\sigma.$$

Consequently,

$$\begin{pmatrix} \dfrac{\partial x}{\partial s} & \dfrac{\partial x}{\partial t} \\[2mm] \dfrac{\partial z}{\partial s} & \dfrac{\partial z}{\partial t}\end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\psi_i & 1-\displaystyle\int_0^s \frac{\partial\psi_i}{\partial t}(R(\sigma,t))\,d\sigma\end{pmatrix}, \tag{3.33}$$

where, for every $f\colon\mathbb{R}^2\to\mathbb{R}$, the partial derivatives $\partial f/\partial s$ and $\partial f/\partial t$ denote $\partial_s[f\circ R]$ and $\partial_t[f\circ R]$, respectively. In particular, it follows that

$$\frac{\partial s}{\partial z} = 0 \quad\text{and}\quad \frac{\partial z}{\partial t}\cdot\frac{\partial t}{\partial z} = 1.$$

Also,

$$\frac{\partial}{\partial s}=\frac{\partial}{\partial x}-\psi_i\frac{\partial}{\partial z}=\partial_i, \tag{3.34}$$

so $\partial/\partial s$ does not depend on j.

Observe that, by the definition of β_i, for all $s,t\in[0,\rho^{-i}]\times[0,\alpha^{-2}\rho^{-2i}]$, we have

$$\beta_i(R_{i,j}(s,t))=\alpha^{-2}\rho^{-i}\beta(\rho^i s,0,\alpha^2\rho^{2i}t).$$

It follows that, for any $m,n\in\mathbb{N}\cup\{0\}$, we have

$$\left\|\frac{\partial^m}{\partial s^m}\frac{\partial^n}{\partial t^n}\beta_i\right\|_{L_\infty(Q_{i,j})}=\alpha^{-2}\rho^{-i}\rho^{mi}(\alpha^2\rho^{2i})^n\left\|\frac{\partial^m}{\partial x^m}\frac{\partial^n}{\partial z^n}\beta\right\|_{L_\infty(U)}. \tag{3.35}$$

This is especially useful when $m+n\leqslant 2$, since in this case

$$\left\|\frac{\partial^m}{\partial x^m}\frac{\partial^n}{\partial z^n}\beta\right\|_{L_\infty(U)}\leqslant 1.$$

Thus,

$$\left\|\frac{\partial\beta_i}{\partial t}\right\|_{L_\infty(Q_{i,j})}\leqslant\rho^i\quad\text{and}\quad\left\|\frac{\partial^2\beta_i}{\partial t^2}\right\|_{L_\infty(Q_{i,j})}\leqslant\alpha^2\rho^{3i}. \tag{3.36}$$

Furthermore, since $\{Q_{i,j}\}_{j=1}^\infty$ cover V_0,

$$\|\partial_i\beta_i\|_{L_\infty(V_0)}=\max_{j\in\mathbb{N}}\|\partial_i\beta_i\|_{L_\infty(Q_{i,j})}\overset{(3.34)}{=}\max_{j\in\mathbb{N}}\left\|\frac{\partial\beta_i}{\partial s}\right\|_{L_\infty(Q_{i,j})}\overset{(3.35)}{\leqslant}\alpha^{-2}. \tag{3.37}$$

The following lemma obtains bounds on vertical derivatives that will be used later.

Lemma 3.10. *If $\rho\geqslant 8$, then for all $i\in\mathbb{N}\cup\{0\}$ we have*

$$\left\|\frac{\partial\psi_i}{\partial z}\right\|_{L_\infty(V_0)}\leqslant 2\rho^{i-1} \tag{3.38}$$

and

$$\left\|\frac{\partial^2\psi_i}{\partial z^2}\right\|_{L_\infty(V_0)}\leqslant 2\alpha^2\rho^{3i-3}. \tag{3.39}$$

Furthermore, if (s,t) are the above flow coordinates on $Q_{i,j}$ for some $j\in\mathbb{N}$, then the following bound holds pointwise on $Q_{i,j}$:

$$\frac{3}{4}<e^{-2\rho^{-1}}\leqslant\frac{\partial t}{\partial z}=\left(\frac{\partial z}{\partial t}\right)^{-1}\leqslant e^{2\rho^{-1}}<\frac{4}{3}. \tag{3.40}$$

Proof. Denote, for every integer $i \geqslant 0$,

$$m_i \stackrel{\text{def}}{=} \left\| \frac{\partial \psi_i}{\partial z} \right\|_{L_\infty(V_0)} \quad \text{and} \quad \mu_i \stackrel{\text{def}}{=} \left\| \frac{\partial^2 \psi_i}{\partial z^2} \right\|_{L_\infty(V_0)}. \tag{3.41}$$

Thus, $m_0=\mu_0=0$. Fix $j\in\mathbb{N}$ and let (s,t) be the flow coordinates on $Q_{i,j}$. We will first use the above identities to deduce bounds on vertical derivatives of t in terms of m_i, μ_i, and then bootstrap these bounds to deduce the desired bounds on m_i, μ_i themselves.

By (3.33), the following identity holds pointwise on $Q_{i,j}$:

$$\frac{\partial}{\partial s}\frac{\partial z}{\partial t} = -\frac{\partial \psi_i}{\partial t} = -\frac{\partial \psi_i}{\partial z}\frac{\partial z}{\partial t} - \frac{\partial \psi_i}{\partial x}\frac{\partial x}{\partial t} = -\frac{\partial \psi_i}{\partial z}\frac{\partial z}{\partial t}.$$

Consequently,

$$\frac{\partial}{\partial s}\left(\log \frac{\partial z}{\partial t}\right) = -\frac{\partial \psi_i}{\partial z}.$$

Since $\frac{\partial z}{\partial t}=1$ when $s=0$, we integrate to get the identity

$$\frac{\partial z}{\partial t} = \exp\Big(-\int_0^s \frac{\partial \psi_i}{\partial z}(R_{i,j}(\sigma,t))\, d\sigma\Big). \tag{3.42}$$

By differentiating (3.42), we also get

$$\frac{\partial^2 z}{\partial t^2} = -\frac{\partial z}{\partial t}\int_0^s \frac{\partial^2 \psi_i}{\partial z^2}(R_{i,j}(\sigma,t))\frac{\partial z}{\partial t}(R_{i,j}(\sigma,t))\, d\sigma. \tag{3.43}$$

For points in $Q_{i,j}$, we have $|s|\leqslant \rho^{-i}$, so it follows from (3.42) that

$$\left|\log \frac{\partial z}{\partial t}\right| \leqslant \rho^{-i} m_i,$$

i.e.,

$$e^{-\rho^{-i} m_i} \leqslant \frac{\partial z}{\partial t} \leqslant e^{\rho^{-i} m_i}. \tag{3.44}$$

By substituting (3.44) into (3.43), we deduce that

$$\left|\frac{\partial^2 z}{\partial t^2}\right| \leqslant \rho^{-i} e^{2\rho^{-i} m_i}\mu_i. \tag{3.45}$$

Since

$$\frac{\partial z}{\partial t}\cdot\frac{\partial t}{\partial z} = 1,$$

it follows from (3.44) that

$$e^{-\rho^{-i} m_i} \leqslant \frac{\partial t}{\partial z} \leqslant e^{\rho^{-i} m_i}, \tag{3.46}$$

and also

$$\left|\frac{\partial^2 t}{\partial z^2}\right| = \left|-\left(\frac{\partial z}{\partial t}\right)^{-3}\frac{\partial^2 z}{\partial t^2}\right| \overset{\substack{(3.44)\\(3.45)}}{\leqslant} \rho^{-i}e^{5\rho^{-i}m_i}\mu_i. \tag{3.47}$$

The bounds (3.46) and (3.47) on the vertical derivatives of the flow coordinate t are in terms of the bounds m_i and μ_i on the vertical derivatives of ψ_i, but they imply as follows unconditional bounds on m_i and μ_i (hence also, by (3.46) and (3.47) once more, unconditional bounds on the vertical derivatives of t). Firstly, observe that

$$\left\|\frac{\partial\beta_i}{\partial z}\right\|_{L_\infty(Q_{i,j})} \leqslant \left\|\frac{\partial\beta_i}{\partial t}\right\|_{L_\infty(Q_{i,j})}\left\|\frac{\partial t}{\partial z}\right\|_{L_\infty(Q_{i,j})} \overset{\substack{(3.36)\\(3.46)}}{\leqslant} \rho^i e^{\rho^{-i}m_i}$$

and

$$\begin{aligned}\left\|\frac{\partial^2\beta_i}{\partial z^2}\right\|_{L_\infty(Q_{i,j})} &= \left\|\frac{\partial}{\partial z}\frac{\partial t}{\partial z}\frac{\partial\beta_i}{\partial t}\right\|_{L_\infty(Q_{i,j})} = \left\|\frac{\partial^2 t}{\partial z^2}\frac{\partial\beta_i}{\partial t} + \left(\frac{\partial t}{\partial z}\right)^2\frac{\partial^2\beta_i}{\partial t^2}\right\|_{L_\infty(Q_{i,j})}\\ &\overset{\substack{(3.36)\\(3.46)\\(3.47)}}{\leqslant} \rho^{-i}e^{5\rho^{-i}m_i}\mu_i\rho^i + e^{2\rho^{-i}m_i}\alpha^2\rho^{3i} = e^{5\rho^{-i}m_i}\mu_i + e^{2\rho^{-i}m_i}\alpha^2\rho^{3i}.\end{aligned}$$

Since $\{Q_{i,j}\}_{j=1}^\infty$ cover V_0, it follows that

$$\left\|\frac{\partial\beta_i}{\partial z}\right\|_{L_\infty(V_0)} \leqslant \rho^i e^{\rho^{-i}m_i} \quad\text{and}\quad \left\|\frac{\partial^2\beta_i}{\partial z^2}\right\|_{L_\infty(V_0)} \leqslant e^{5\rho^{-i}m_i}\mu_i + e^{2\rho^{-i}m_i}\alpha^2\rho^{3i}.$$

Since, by (3.28), we have

$$\frac{\partial\psi_{i+1}}{\partial z} = \frac{\partial\psi_i}{\partial z} + \frac{\partial\beta_i}{\partial z} \quad\text{and}\quad \frac{\partial^2\psi_{i+1}}{\partial z^2} = \frac{\partial^2\psi_i}{\partial z^2} + \frac{\partial^2\beta_i}{\partial z^2},$$

we deduce that

$$m_{i+1} \leqslant m_i + \rho^i e^{\rho^{-i}m_i} \quad\text{and}\quad \mu_{i+1} \leqslant \mu_i + e^{5\rho^{-i}m_i}\mu_i + e^{2\rho^{-i}m_i}\alpha^2\rho^{3i}. \tag{3.48}$$

By induction, we suppose that (3.38) and (3.39) hold for some integer $i \geqslant 0$, that is,

$$m_i \leqslant 2\rho^{i-1} \quad\text{and}\quad \mu_i \leqslant 2\alpha^2\rho^{3i-3}. \tag{3.49}$$

Since $\rho \geqslant 8$, it follows that

$$m_{i+1} \overset{\substack{(3.48)\\(3.49)}}{\leqslant} \left(\frac{2}{\rho} + e^{2\rho^{-1}}\right)\rho^i \leqslant \left(\frac{1}{4} + \sqrt[4]{e}\right)\rho^i \leqslant 2\rho^i.$$

Thus (3.38) holds for all integers $i \geqslant 0$. Likewise,

$$\mu_{i+1} \overset{\substack{(3.48)\\(3.49)}}{\leqslant} \left(\frac{2}{\rho^3} + \frac{2e^{10\rho^{-1}}}{\rho^3} + e^{4\rho^{-1}}\right)\alpha^2\rho^{3i} \leqslant \left(\frac{2}{8^3} + \frac{2e^{5/4}}{8^3} + \sqrt{e}\right)\alpha^2\rho^{3i} \leqslant 2\alpha^2\rho^{3i},$$

so (3.39) also holds for all integers $i \geqslant 0$. The remaining assertion (3.40) follows by substituting the above bound on m_i into (3.46). □

Next, we will use the bounds of Lemma 3.10 to bound $\{D_i\}_{i=0}^{\infty}$ and their derivatives.

LEMMA 3.11. *Suppose that $\rho\geqslant 8$. For every integer $i\geqslant 0$ we have*

$$\|D_i\|_{L_\infty(V_0)} \leqslant 3\alpha^{-2}, \tag{3.50}$$

$$\left\|\frac{\partial D_i}{\partial z}\right\|_{L_\infty(V_0)} \leqslant 6\rho^{2i}, \tag{3.51}$$

$$\|\partial_i D_i\|_{L_\infty(V_0)} = 5\alpha^{-2}\rho^i. \tag{3.52}$$

Proof. Fix $j\in\mathbb{N}$. Let (s,t) be the flow coordinates on $Q_{i,j}$. By (3.32) and (3.34), we have

$$D_i = -\beta_i\frac{\partial\psi_{i+1}}{\partial z}+\frac{\partial\beta_i}{\partial s}. \tag{3.53}$$

Therefore, by Lemma 3.10 and (3.35), we have

$$\|D_i\|_{L_\infty(Q_{i,j})} \leqslant \alpha^{-2}\rho^{-i}\cdot 2\rho^i+\alpha^{-2} = 3\alpha^{-2}.$$

This proves (3.50), because $\{Q_{i,j}\}_{j=1}^{\infty}$ cover V_0.

Next, we consider $\partial D_i/\partial z$. By differentiating (3.53), we see that

$$\frac{\partial D_i}{\partial z} = -\frac{\partial t}{\partial z}\cdot\frac{\partial\beta_i}{\partial t}\cdot\frac{\partial\psi_{i+1}}{\partial z}-\beta_i\frac{\partial^2\psi_{i+1}}{\partial z^2}+\frac{\partial t}{\partial z}\cdot\frac{\partial^2\beta_i}{\partial s\partial t}.$$

Hence, by Lemma 3.10 and (3.35), we see that

$$\left\|\frac{\partial D_i}{\partial z}\right\|_{L_\infty(Q_{i,j})} \leqslant \frac{4}{3}\cdot\rho^i\cdot 2\rho^i+\alpha^{-2}\rho^{-i}\cdot 2\alpha^2\rho^{3i}+\frac{4}{3}\cdot\rho^{2i} = 6\rho^{2i}.$$

As before, this proves (3.51), because $\{Q_{i,j}\}_{j=1}^{\infty}$ cover V_0.

Finally, we consider $\partial_i D_i$. Note first that, for any $m\in\mathbb{N}$,

$$\left\|\frac{\partial(\partial_m\psi_m)}{\partial z}\right\|_\infty \overset{(3.31)}{\leqslant} \sum_{n=0}^{m-1}\left\|\frac{\partial D_n}{\partial z}\right\|_\infty \overset{(3.51)}{\leqslant} 6\frac{\rho^{2m}-1}{\rho^2-1} \leqslant 7\rho^{2m-2}, \tag{3.54}$$

where we used the assumption $\rho\geqslant 8$. Recalling (3.32) and (3.34), we have

$$\partial_i D_i = \frac{\partial}{\partial s}\left(-\beta_i\frac{\partial\psi_{i+1}}{\partial z}+\frac{\partial\beta_i}{\partial s}\right) = -\frac{\partial\beta_i}{\partial s}\cdot\frac{\partial\psi_{i+1}}{\partial z}-\beta_i\frac{\partial}{\partial s}\left(\frac{\partial\psi_{i+1}}{\partial z}\right)+\frac{\partial^2\beta_i}{\partial s^2}.$$

Using Lemma 3.10 and (3.35), it follows that

$$\|\partial_i D_i\|_{L_\infty(Q_{i,j})} \leqslant 3\alpha^{-2}\rho^i+\alpha^{-2}\rho^{-i}\left\|\frac{\partial}{\partial s}\frac{\partial\psi_{i+1}}{\partial z}\right\|_{L_\infty(Q_{i,j})}. \tag{3.55}$$

To bound the last term in (3.55), we first calculate the Lie bracket

$$\left[\frac{\partial}{\partial z},\frac{\partial}{\partial s}\right]=\left[\frac{\partial}{\partial z},\frac{\partial}{\partial x}-\psi_i\frac{\partial}{\partial z}\right]=-\frac{\partial\psi_i}{\partial z}\cdot\frac{\partial}{\partial z}.$$

This implies that

$$\begin{aligned}\frac{\partial}{\partial s}\frac{\partial\psi_{i+1}}{\partial z}&=\frac{\partial}{\partial z}\left(\frac{\partial\psi_{i+1}}{\partial s}\right)+\frac{\partial\psi_i}{\partial z}\cdot\frac{\partial\psi_{i+1}}{\partial z}\\&=\frac{\partial}{\partial z}\left(\partial_i\psi_i+\frac{\partial\beta_i}{\partial s}\right)+\frac{\partial\psi_i}{\partial z}\cdot\frac{\partial\psi_{i+1}}{\partial z}=\frac{\partial(\partial_i\psi_i)}{\partial z}+\frac{\partial t}{\partial z}\cdot\frac{\partial^2\beta_i}{\partial s\partial t}+\frac{\partial\psi_i}{\partial z}\cdot\frac{\partial\psi_{i+1}}{\partial z}.\end{aligned}$$

Therefore, by Lemma 3.10, (3.35), (3.51), and (3.54), we conclude that, since $\rho\geqslant 8$,

$$\left\|\frac{\partial}{\partial s}\frac{\partial\psi_{i+1}}{\partial z}\right\|_{L_\infty(Q_{i,j})}\leqslant 7\rho^{2i-2}+\frac{4}{3}\cdot\rho^{2i}+2\rho^{i-1}\cdot 2\rho^i\leqslant 2\rho^{2i}.$$

Due to (3.55), this implies the final desired bound (3.52) of Lemma 3.11. □

The first assertion (3.50) of Lemma 3.11 gives the first assertion of Lemma 3.9. To prove the second assertion of Lemma 3.9, we first bound the variation of D_m on each of the pseudoquads $\{Q_{n,j}\}_{j=1}^\infty$ when $n\geqslant m$.

Lemma 3.12. *Fix two integers $n\geqslant m\geqslant 0$. For any $j\in\mathbb{N}$ and any $w,w'\in Q_{n,j}$, we have*

$$|D_m(w)-D_m(w')|\lesssim\alpha^{-2}\rho^{m-n}.$$

Proof. Let $R=R_{n,j}$ and let $(s,t),(s',t')\in[0,\rho^{-n}]\times[0,\alpha^{-2}\rho^{-2n}]$ be such that

$$R(s,t)=w\quad\text{and}\quad R(s',t')=w'.$$

With respect to flow coordinates on $Q_{n,j}$, we have

$$\frac{\partial D_m}{\partial s}=\partial_n D_m=\partial_m D_m+(\psi_m-\psi_n)\frac{\partial D_m}{\partial z}.$$

Since $\|\psi_m-\psi_n\|_{L_\infty(V_0)}\leqslant\alpha^{-2}\rho^{-m}+\alpha^{-2}\rho^{-n}\leqslant 2\alpha^{-2}\rho^{-m}$, using Lemma 3.11 we get that

$$\left\|\frac{\partial D_m}{\partial s}\right\|_{L_\infty(Q_{n,j})}\leqslant 5\alpha^{-2}\rho^m+2\alpha^{-2}\rho^{-m}\cdot 6\rho^{2m}=17\alpha^{-2}\rho^m.$$

Hence, using Lemmas 3.10 and 3.11, we conclude that

$$\begin{aligned}|D_m(w)-D_m(w')|&\leqslant\left\|\frac{\partial D_m}{\partial s}\right\|_{L_\infty(Q_{n,j})}|s-s'|+\left\|\frac{\partial D_m}{\partial z}\right\|_{L_\infty(Q_{n,j})}\left\|\frac{\partial z}{\partial t}\right\|_{L_\infty(Q_{n,j})}|t-t'|\\&\leqslant 17\alpha^{-2}\rho^{m-n}+6\rho^{2m}\cdot\frac{4}{3}\cdot\alpha^{-2}\rho^{-2n}\lesssim\alpha^{-2}\rho^{m-n}.\end{aligned}$$

□

Prior to proving Proposition 3.4, we record a quick consequence of Green's theorem.

LEMMA 3.13. *Let $M\subseteq V_0$ be a region bounded by a simple piecewise-smooth closed curve and let $f\colon V_0\to\mathbb{R}$ be a smooth function. Then,*

$$\int_M \partial_f f\,dw=\int_{\partial M}\left(\frac{f^2}{2},f\right)\cdot d\mathbf{r}.$$

In particular, if $g\colon V_0\to\mathbb{R}$ is another smooth function such that $f=g$ on ∂M, then

$$\int_M \partial_f f\,dw=\int_M \partial_g g\,dw.$$

Proof. Since

$$\nabla\times\left(\frac{f^2}{2},f\right)=\frac{\partial f}{\partial x}-f\frac{\partial f}{\partial z}=\partial_f f,$$

the lemma follows from Green's theorem. □

Proof of Lemma 3.9. The first assertion of Lemma 3.9 was proved in Lemma 3.11, so here we treat its second assertion, namely that $\{D_n\}_{n=0}^{\infty}$ are almost-orthogonal.

Fix $m,n\in\mathbb{N}\cup\{0\}$ with $n\geqslant m$ and $j\in\mathbb{N}$. As $\psi_{n+1}-\psi_n=\beta_n=0$ on $\partial Q_{n,j}$, Lemma 3.13 implies that $\int_{Q_{n,j}} D_n(w)\,dw=0$. So, fixing an arbitrary basepoint $w_0\in Q$, we have

$$\left|\int_{Q_{n,j}} D_m(w)D_n(w)\,dw\right|=\left|\int_{Q_{n,j}} (D_m(w)-D_m(w_0))D_n(w)\,dw\right|\lesssim\alpha^{-4}\rho^{m-n}\mathcal{H}^3(Q_{n,j}),$$

where in the final step we used (3.50) and Lemma 3.12. Hence, $|\langle D_m,D_n\rangle_U|$ is at most

$$\sum_{\substack{j\in\mathbb{N}\\ Q_{n,j}\subseteq U}}\left|\int_{Q_{n,j}} D_m(w)D_n(w)\,dw\right|\leqslant\sum_{\substack{j\in\mathbb{N}\\ Q_{n,j}\subseteq U}}\mathcal{H}^3(Q_{n,j})\alpha^{-4}\rho^{m-n}\asymp\alpha^{-4}\rho^{m-n}. \qquad\square$$

3.2.2. The vertical perimeter of Γ_{ψ_i}

Here we will complete the proof of Proposition 3.4.

Define $\phi\colon V_0\to\mathbb{R}$ to be the A-periodic extension of $\beta|_U$, i.e.,

$$\phi(x,0,z)\stackrel{\mathrm{def}}{=}\beta(\{x\},0,\{z\})$$

for $(x,0,z)\in V_0$, where $\{a\}=a-\lfloor a\rfloor$ is the fractional part of $a\in\mathbb{R}$. Because the function

$$\bar{\mathsf{v}}^P_{U,\phi}\colon\mathbb{R}\longrightarrow[0,\infty)$$

is continuous and not identically zero, there exist $\eta, R, r\in\mathbb{R}$ with $r<R$ such that

$$\bar{\mathsf{v}}^P_{U,\phi}(a)\geqslant\eta>0\quad\text{for all } a\in I\overset{\mathrm{def}}{=}[r,R].\tag{3.56}$$

We will show that if $\rho\in\mathbb{N}$ is large enough (depending only on the initial choice of bump function β), then the conclusion of Proposition 3.4 holds for the above interval I. To this end, we will first establish the following pointwise bound on the vertical perimeter of each of the perturbations $\{\beta_i\}_{i=0}^{\infty}$ in terms of the vertical perimeter of ϕ.

LEMMA 3.14. *Suppose that $\rho>5$. For every $i\in\mathbb{N}\cup\{0\}$ and $a\in\mathbb{R}$ we have*

$$\bar{\mathsf{v}}^P_{U,\beta_i}(a)\geqslant\frac{1}{2\alpha}\bar{\mathsf{v}}^P_{U,\phi}(a-\log_2(\alpha\rho^i))-\frac{3\rho^{i-1}}{2^a}.$$

In particular, if $\rho\geqslant 12/2^r\eta$ and $a\in I+\log_2(\alpha\rho^i)$, then

$$\bar{\mathsf{v}}^P_{U,\beta_i}(a)\geqslant\frac{\eta}{2\alpha}-\frac{3\rho^{i-1}}{2^{r+\log_2(\alpha\rho^i)}}\geqslant\frac{\eta}{4\alpha}.$$

Proof of Proposition 3.4 assuming Lemma 3.14. Fix an integer $\rho\geqslant\max\{12/2^r\eta,8\}$ that will be specified later and let $\psi=\psi_{\alpha^4}$. The first three assertions of Proposition 3.4 were established in the construction of ψ and in the discussion after Lemma 3.9. We will establish the last three by showing that

$$\bar{\mathsf{v}}^P_{U,\psi}(a)\lesssim\min\left\{\frac{1}{\alpha},\frac{2^a}{\alpha^2}\right\}\quad\text{for all } a\in\mathbb{R},\tag{3.57}$$

and

$$\bar{\mathsf{v}}^P_{U,\psi}(a)\gtrsim\frac{1}{\alpha}\quad\text{for all } a\in\bigcup_{n=0}^{\alpha^4-1}(I+\log_2(\alpha\rho^n)).\tag{3.58}$$

For every $i\in\mathbb{N}\cup\{0\}$, by the definition of β_i and by (3.36) and (3.40), we have

$$\|\beta_i\|_{L_\infty(V_0)}\leqslant\alpha^{-2}\rho^{-i}\quad\text{and}\quad\left\|\frac{\partial\beta_i}{\partial z}\right\|_{L_\infty(V_0)}\leqslant 2\rho^i.$$

Due to Lemma 2.13, for every $a\in\mathbb{R}$ we have

$$\bar{\mathsf{v}}^P_{U,\beta_i}(a)\leqslant\min\{2^{a+1}\alpha^{-2}\rho^{-i},2^{-a+1}\rho^i\}=2\alpha^{-1}2^{-|a-\log_2(\alpha\rho^i)|}.\tag{3.59}$$

Consequently,

$$\bar{\mathsf{v}}^P_{U,\psi_{\alpha^4}}(a)=\bar{\mathsf{v}}^P_{U,\sum_{i=0}^{\alpha^4}\beta_i}(a)\leqslant\sum_{i=0}^{\alpha^4}\bar{\mathsf{v}}^P_{U,\beta_i}(a)\overset{(3.59)}{\leqslant}\sum_{i=0}^{\infty}2\alpha^{-1}2^{-|a-\log_2(\alpha\rho^i)|}\lesssim\alpha^{-1}.$$

This proves (3.57), because by Lemma 2.13 we also have

$$\bar{\mathsf{v}}^P_{U,\psi_{\alpha^4}}(a) \lesssim 2^a \|\psi_{\alpha^4}\|_{L_\infty(V_0)} \overset{(3.29)}{\leqslant} 2^a\alpha^{-2}.$$

It remains to prove (3.58), as we saw in (3.18) that this implies the remaining assertions of Proposition 3.4. Fix $n\in\{0,\dots,\alpha^4-1\}$ and $a\in I+\log_2(\alpha\rho^n)$, so that

$$\bar{\mathsf{v}}^P_{U,\beta_n}(a) > \frac{\eta}{4\alpha}$$

by Lemma 3.14. Let $s=\max\{|r|,|R|\}$, so that $|a-\log_2(\alpha\rho^n)|\leqslant s$. It follows from (3.59) that

$$\bar{\mathsf{v}}^P_{U,\beta_i}(a) \leqslant 2\alpha^{-1}2^{-|\log_2(\alpha\rho^n)-\log_2(\alpha\rho^i)|+|a-\log_2(\alpha\rho^n)|} \leqslant 2\alpha^{-1}\rho^{-|n-i|}2^s \tag{3.60}$$

for any $i\in\mathbb{N}\cup\{0\}$. Hence, by combining Lemma 3.14 and (3.60), we conclude that

$$\begin{aligned}\bar{\mathsf{v}}^P_{U,\psi_{\alpha^4}}(a) = \bar{\mathsf{v}}^P_{U,\sum_{i=0}^{\alpha^4}\beta_i}(a) &\geqslant \bar{\mathsf{v}}^P_{U,\beta_n}(a) - \sum_{i=0}^{n-1}\bar{\mathsf{v}}^P_{U,\beta_i}(a) - \sum_{i=n+1}^{\alpha^4}\bar{\mathsf{v}}^P_{U,\beta_i}(a)\\ &\geqslant \frac{\eta}{4\alpha} - 2\sum_{k=1}^{\infty}2\alpha^{-1}\rho^{-k}2^s \geqslant \frac{\eta}{4\alpha}-\frac{5}{\alpha\rho}2^s.\end{aligned}$$

Choosing

$$\rho \overset{\mathrm{def}}{=} \left\lceil \max\left\{8, \frac{12}{2^r\eta}, \frac{40\cdot 2^s}{\eta}\right\}\right\rceil,$$

this completes the proof of Proposition 3.4. □

Proof of Lemma 3.14. We will start by introducing some (convenient, though ad hoc) notation and making some preliminary observations. For $i\in\mathbb{N}\cup\{0\}$ define a (discontinuous in the first variable) map $\mathfrak{S}_i:\mathbb{R}^2\to\mathbb{R}^2$ as follows. If $s\in\mathbb{R}$, then let $m\in\mathbb{Z}$ be the unique integer such that $s\in[m\rho^{-i},(m+1)\rho^{-i})$, and set, for every $t\in\mathbb{R}$,

$$\mathfrak{S}_i(s,t) \overset{\mathrm{def}}{=} \Phi(\psi_i)_{s-m\rho^{-i}}(m\rho^{-i},0,t),$$

where we recall the notation $\Phi(\cdot)_\cdot(\cdot)$ for characteristic curves that we set at the start of §3.2. Note that, by design, $x(\mathfrak{S}_i(s,t))=s$. Observe also that the lines $\mathbb{R}\times\{0\}\times\{0\}$ and $\mathbb{R}\times\{0\}\times\{1\}$ are characteristic curves for Γ_{ψ_i}, since ψ_i vanishes on those lines. Hence, $\mathfrak{S}_i(s,0)=(s,0)$ and $\mathfrak{S}_i(s,1)=(s,1)$ for all $s\in[0,1]$. As $x(\mathfrak{S}(s,t))=s$ for all $t\in\mathbb{R}$, by the continuity of $\mathfrak{S}_i$ in the second variable, this implies that $\mathfrak{S}_i(s,[0,1])=\{s\}\times\{0\}\times[0,1]$. So,

$$\mathfrak{S}_i([0,1]^2)=U. \tag{3.61}$$

The mapping $\mathfrak{S}_i$ is related as follows to the mappings $R_{i,1}, R_{i,2}, \ldots$ that are given in (3.24). Suppose as above that $s\in[m\rho^{-i},(m+1)\rho^{-i})$ for some $m\in\mathbb{Z}$, and fix $t\in\mathbb{R}$ and $n\in\mathbb{Z}$. Recalling that $v_{i,1}, v_{i,2}, \ldots$ is an enumeration of the points in the grid G_i that is given in (3.23), let $j\in\mathbb{N}$ be the index for which $v_{i,j}=(m\rho^{-i},0,n\alpha^{-2}\rho^{-2i})$. Then,

$$\mathfrak{S}_i(s,t)=R_{i,j}(s-m\rho^{-i},t-n\alpha^{-2}\rho^{-2i}).$$

Recalling the definition (3.25) of the pseudo-quad $Q_{i,j}$, this implies that

$$\overline{\mathfrak{S}_i([m\rho^{-i},(m+1)\rho^{-i})\times[n\alpha^{-2}\rho^{-2i},(n+1)\alpha^{-2}\rho^{-2i}])}=Q_{i,j}.$$

Also, recalling the definitions (3.26) and (3.27), it follows that if we define $\phi_i\colon V_0\to\mathbb{R}$ by

$$\phi_i(s,0,t)\overset{\mathrm{def}}{=}\alpha^{-2}\rho^{-i}\phi(\rho^i s,0,\alpha^2\rho^{2i}t)\quad\text{for all }(s,t)\in\mathbb{R}^2,\tag{3.62}$$

then

$$\beta_i(\mathfrak{S}_i(s,t))=\phi_i(s,0,t)\quad\text{for all }(s,t)\in\mathbb{R}^2.\tag{3.63}$$

Fix $i\in\mathbb{N}\cup\{0\}$, $a\in\mathbb{R}$, and $(x,0,z)\in U$. Now let $s=s(x,z), t=t(x,z), t'=t'(x,z,a)\in\mathbb{R}$ satisfy

$$\mathfrak{S}_i(s,t)=(x,0,z)\quad\text{and}\quad\mathfrak{S}_i(s,t')=(x,0,z-2^{-2a}).\tag{3.64}$$

Due to (3.40), we have $e^{-2\rho^{-1}}2^{-2a}\leqslant t-t'\leqslant e^{2\rho^{-1}}2^{-2a}$. Hence,

$$|t'-(t-2^{-2a})|\leqslant(e^{2\rho^{-1}}-1)2^{-2a}\leqslant 3\rho^{-1}2^{-2a}.\tag{3.65}$$

Now,

$$\begin{aligned}
&|\beta_i(x,0,z)-\beta_i(x,0,z-2^{-2a})|\\
&\overset{\substack{(3.63)\\(3.64)}}{=}|\phi_i(s,0,t)-\phi_i(s,0,t-2^{-2a})-\phi_i(s,0,t')+\phi_i(s,0,t-2^{-2a})|\\
&\geqslant |\phi_i(s,0,t)-\phi_i(s,0,t-2^{-2a})|-|\phi_i(s,0,t')-\phi_i(s,0,t-2^{-2a})|\\
&\overset{(3.62)}{\geqslant}|\phi_i(s,0,t)-\phi_i(s,0,t-2^{-2a})|-\rho^i|t'-(t-2^{-2a})|\\
&\overset{(3.65)}{\geqslant}|\phi_i(s,0,t)-\phi_i(s,0,t-2^{-2a})|-3\rho^{i-1}2^{-2a}.
\end{aligned}$$

In other words, we established the following pointwise estimate for the vertical difference quotients that occur in the definition (2.29) of (parameterized) vertical perimeter:

$$\frac{|\beta_i(x,0,z)-\beta_i(x,0,z-2^{-2a})|}{2^{-a}}\geqslant\frac{|\phi_i(s,0,t)-\phi_i(s,0,t-2^{-2a})|}{2^{-a}}-\frac{3\rho^{i-1}}{2^a}.$$

By integrating this inequality over U, we get

$$\begin{aligned}\bar{\mathsf{v}}^P_{U,\beta_i}(a) &\overset{(2.29)}{\geqslant} \int_0^1\int_0^1 \frac{|\phi_i(s(x,z),0,t(x,z))-\phi_i(s(x,z),0,t(x,z)-2^{-2a})|}{2^{-a}}\,dx\,dz-\frac{3\rho^{i-1}}{2^a}\\ &\overset{(3.33)}{=} \int_{\mathfrak{S}_i^{-1}(U)} \frac{|\phi_i(s,0,t)-\phi_i(s,0,t-2^{-2a})|}{2^{-a}}\left|\frac{\partial z}{\partial t}(s,t)\right|\,ds\,dt-\frac{3\rho^{i-1}}{2^a}\\ &\overset{(3.40)\,(3.61)}{\geqslant} \frac{1}{2}\int_0^1\int_0^1 \frac{|\phi_i(s,0,t)-\phi_i(s,0,t-2^{-2a})|}{2^{-a}}\,ds\,dt-\frac{3\rho^{i-1}}{2^a}.\end{aligned}$$

It therefore remains to note the following identity.

$$\int_0^1\int_0^1 \frac{|\phi_i(s,0,t)-\phi_i(s,0,t-2^{-2a})|}{2^{-a}}\,ds\,dt$$

$$= \int_0^{\rho^i}\int_0^{\alpha^2\rho^{2i}} \frac{\alpha^{-4}\rho^{-4i}|\phi(\sigma,0,\tau)-\phi(\sigma,0,\tau-\alpha^2\rho^{2i}2^{-2a})|}{2^{-a}}\,d\sigma\,d\tau \tag{3.66}$$

$$= \frac{1}{\alpha}\int_U \frac{|\phi(v)-\phi(vZ^{-\alpha^2\rho^{2i}2^{-2a}})|}{\alpha\rho^i 2^{-a}}\,dv \tag{3.67}$$

$$= \frac{1}{\alpha}\bar{\mathsf{v}}^P_{U,\phi}(a-\log_2(\alpha\rho^i)), \tag{3.68}$$

where (3.66) uses the definition (3.62) and the change of variables $(s,t)=(\rho^{-i}\sigma,\alpha^{-2}\rho^{-2i}\tau)$, (3.67) holds by the periodicity of ϕ, and (3.68) is a restatement of the definition (2.29). $\square$

4. Pseudoquads and foliated patchworks

Let Γ be the intrinsic Lipschitz graph of $f\colon V_0\to\mathbb{R}$. A *pseudoquad* Q is a region of V_0 bounded by two vertical lines and two characteristic curves of Γ, i.e., a region of the form

$$Q=\{(x,0,z)\in V_0 : x\in I \text{ and } g_1(x)\leqslant z\leqslant g_2(x)\},$$

where $I=[a,b]\subseteq\mathbb{R}$ is a closed, bounded interval and $g_1,g_2\colon\mathbb{R}\to\mathbb{R}$ are functions whose graphs are characteristic. We say that I is the *base* of Q, and we call g_1 and g_2 the *lower* and *upper bounds* of Q, respectively. The width of the pseudoquad Q is just the length $\ell(I)=b-a$ of its base $I=[a,b]$. But, the height of Q is not always well behaved, since characteristic curves can join and split. We therefore introduce *rectilinear pseudoquads*, which approximate projections of rectangles in vertical planes. If Γ is a vertical plane, its characteristic curves are a family of parallel parabolas; conversely, any pseudoquad bounded by two parallel parabolas is the projection of a rectangle in $\mathbb{H}$ (a loop composed of two parallel horizontal lines and two vertical lines) to V_0. Thus, if

$$R=\{(x,0,z)\in V_0 : x\in I \text{ and } h_1(x)\leqslant z\leqslant h_2(x)\},$$

where $h_1, h_2: \mathbb{R} \to \mathbb{R}$ are quadratic functions that differ by a constant, then we call R a *parabolic rectangle* with *width*

$$\delta_x(R) \stackrel{\text{def}}{=} \ell(I)$$

and *height*

$$\delta_z(R) \stackrel{\text{def}}{=} h_2 - h_1.$$

For $r>0$ and an interval I, let rI be the scaling of I around its center by a factor of r, i.e.,

$$rI \stackrel{\text{def}}{=} \left[\tfrac{1}{2}(a+b) - \tfrac{1}{2} r\ell(I), \tfrac{1}{2}(a+b) + \tfrac{1}{2} r\ell(I)\right].$$

For $\rho>0$, let

$$\begin{aligned} \rho R &\stackrel{\text{def}}{=} \{(x,0,z) \in V_0 : x \in \rho I \text{ and } z \in \rho^2 [h_1(x), h_2(x)]\} \\ &= \left\{(x,0,z) \in V_0 : x \in \rho I \text{ and } \left|z - \tfrac{1}{2}(h_1(x)+h_2(x))\right| \leqslant \tfrac{1}{2}\rho^2 \delta_z(R)\right\}. \end{aligned} \tag{4.1}$$

For $0<\mu\leqslant\frac{1}{32}$, a *$\mu$-rectilinear pseudoquad* is a pair (Q,R), where Q is a pseudoquad and R is a parabolic rectangle with the same base I as Q such that, if g_1 and g_2 (resp. h_1 and h_2) are the lower and upper bounds of Q (resp. R), then

$$\max\{\|g_1 - h_1\|_{L_\infty(4I)}, \|g_2 - h_2\|_{L_\infty(4I)}\} \leqslant \mu \delta_z(R). \tag{4.2}$$

We will frequently refer to a μ-rectilinear pseudoquad (Q,R) as simply Q, but we define its width and height to be the width and height of the associated parabolic rectangle, i.e., $\delta_x(Q)=\delta_x(R)$ and $\delta_z(Q)=\delta_z(R)$. Likewise, for $\rho\geqslant 1$, we define $\rho Q=\rho R$. Note that Q need not be contained in $1Q=R$, but the following lemma holds.

LEMMA 4.1. *Let Q be a μ-rectilinear pseudoquad. Then $Q\subseteq 2Q$. In fact, for every $t\in\mathbb{R}$,*

$$QZ^{t\delta_z(Q)} \subseteq \sqrt{2|t|+2}\cdot Q.$$

Proof. Let R, g_1, g_2, h_1, h_2 be as above. Let $m_g=\frac{1}{2}(g_1+g_2)$ and $m_h=\frac{1}{2}(h_1+h_2)$. Fix $(x,0,z)\in Q$, so that $g_1(x)\leqslant z\leqslant g_2(x)$. For $i\in\{1,2\}$, we have

$$|m_h(x) - g_i(x)| \leqslant |m_h(x) - h_i(x)| + |h_i(x) - g_i(x)| \leqslant \tfrac{1}{2}\delta_z(Q) + \mu\delta_z(Q) \leqslant \delta_z(Q),$$

so

$$|m_h(x) - (z + t\delta_z(Q))| \leqslant (1+|t|)\delta_z(Q).$$

Therefore,

$$(x,0,z+\delta_z(Q)) \in \sqrt{2|t|+2}\cdot Q. \qquad \square$$

Continuing with the above notation, define the *aspect ratio* of Q to be

$$\alpha(Q) \stackrel{\text{def}}{=} \frac{\delta_x(Q)}{\sqrt{\delta_z(Q)}}. \tag{4.3}$$

We use a square root here because the distance in the Heisenberg metric between the top and bottom of Q is proportional to $\sqrt{\delta_z(Q)}$; thus this aspect ratio is invariant under the Heisenberg scaling. Let $|Q|$ be the Lebesgue measure of Q as a subset of $V_0 \cong \mathbb{R}^2$.

The following lemma is a direct consequence of Lemma 2.9.

LEMMA 4.2. *Let $a, b \in \mathbb{R} \setminus \{0\}$ and let $g = q \circ \rho_h \circ s_{a,b}: \mathbb{H} \to \mathbb{H}$ be a composition of a shear map q, a left-translation by $h \in \mathbb{H}$, and a stretch map $s_{a,b}$. Let $\hat{g}: V_0 \to V_0$ be the map induced on V_0, i.e., $\hat{g}(x) = \Pi(g(x))$ for all $x \in V_0$. Suppose that (Q, R) is a μ-rectilinear pseudoquad for an intrinsic graph Γ. Then, $(\widehat{Q}, \widehat{R}) = (\hat{g}(Q), \hat{g}(R))$ is a μ-rectilinear pseudoquad for the intrinsic graph $\hat{g}(\Gamma)$, with the following parameters*:

$$\delta_x(\widehat{Q}) = |a| \delta_x(Q), \quad \delta_z(\widehat{Q}) = |ab| \delta_z(Q), \quad |\widehat{Q}| = |a^2 b| \cdot |Q|, \quad \alpha(\widehat{Q}) = \sqrt{\frac{|a|}{|b|}} \cdot \alpha(Q).$$

Remark 4.3. For any μ-rectilinear pseudoquad (Q, R), there is a transformation of $\mathbb{H}$ that sends R to a square in V_0, and Q to an approximation of the square. That is, if $a, b, c, d, x_0, w \in \mathbb{R}$ are such that

$$R = \{(x, 0, z) \in V_0 : |x - x_0| \leqslant w \text{ and } |ax^2 + bx + c - z| \leqslant d\},$$
$$h(v) = s_{w^{-1}, wd^{-1}}(X^{-x_0} Y^b Z^{-c} \tilde{A}_{2a}(v)),$$

and $\hat{h} = \Pi \circ h$, then, by the remarks after Lemma 2.9, $\hat{h}(R) = [-1, 1] \times \{0\} \times [-1, 1]$.

By Lemma 4.2, $(\hat{h}(Q), \hat{h}(R))$ is μ-rectilinear, so if $\hat{g}_1$ and $\hat{g}_2$ are the lower and upper bounds of $\hat{h}(Q)$, then $|\hat{g}_1(t) + 1| < 2\mu$ and $|\hat{g}_2(t) - 1| < 2\mu$ for all $t \in [-4, 4]$.

We will prove Theorem 1.18 by constructing a collection of nested partitions of V_0 into pseudoquads. We will describe these partitions by associating a rectilinear pseudoquad with each vertex of a rooted tree. Let (T, v_0) be a rooted tree with vertex set $\mathcal{V}(T)$. For $v \in \mathcal{V}(T)$, we let $\mathcal{C}(v) = \mathcal{C}^1(v)$ denote the set of *children* of v and inductively, for $n \geqslant 2$, let

$$\mathcal{C}^n(v) = \bigcup_{w \in \mathcal{C}^{n-1}(v)} \mathcal{C}(w)$$

be the set of nth generation descendants of v. Let $\mathcal{D}(v) = \bigcup_{n=0}^{\infty} \mathcal{C}^n(v)$, where $\mathcal{C}^0(v) = \{v\}$. For $v \in \mathcal{V}(T) \setminus \{v_0\}$, there is a unique *parent* vertex w such that $v \in \mathcal{C}(w)$, and we denote this vertex by $\mathcal{P}(v)$. If $w \in \mathcal{D}(v)$, we say that w is a *descendant* of v or that v is an *ancestor* of w and write $w \leqslant v$. This is a partial order with maximal element v_0.

Definition 4.4. (Rectilinear foliated patchwork) If Q is a μ-rectilinear pseudoquad, a *μ-rectilinear foliated patchwork* for Q is a complete rooted binary tree (Δ, v_0) (i.e., every vertex has exactly two children) such that every vertex $v\in\mathcal{V}(\Delta)$ is associated with a μ-rectilinear pseudoquad (Q_v, R_v) with $Q_{v_0}=Q$. Each vertex $v\in\mathcal{V}(\Delta)$ is either *vertically cut* or *horizontally cut* in the following sense.

Let w and w' be the children of v, let $I=[a,b]$ be the base of Q_v, and let g_1 and g_2 (resp. h_1 and h_2) be the lower and upper bounds of Q_v (resp. R_v).

(1) If v is vertically cut, then Q_w and $Q_{w'}$ are the left and right halves of Q_v, separated by the vertical line $x=\frac{1}{2}(a+b)$. That is,

$$\begin{aligned} Q_w &= \bigl\{(x,0,z)\in V_0 : a\leqslant x\leqslant \tfrac{1}{2}(a+b) \text{ and } g_1(x)\leqslant z\leqslant g_2(x)\bigr\},\\ Q_{w'} &= \bigl\{(x,0,z)\in V_0 : \tfrac{1}{2}(a+b)\leqslant x\leqslant b \text{ and } g_1(x)\leqslant z\leqslant g_2(x)\bigr\}. \end{aligned}$$

Similarly,

$$R_w = \bigl(\bigl[a,\tfrac{1}{2}(a+b)\bigr]\times\{0\}\times\mathbb{R}\bigr)\cap R_v \quad \text{and} \quad R_{w'} = \bigl(\bigl[\tfrac{1}{2}(a+b),b\bigr]\times\{0\}\times\mathbb{R}\bigr)\cap R_v.$$

We therefore have

$$\delta_x(Q_w)=\delta_x(Q_{w'})=\tfrac{1}{2}\delta_x(Q_v) \quad \text{and} \quad \delta_z(Q_w)=\delta_z(Q_{w'})=\delta_z(Q_v).$$

(2) If v is horizontally cut, then Q_w and $Q_{w'}$ are the top and bottom halves of Q_v, separated by a characteristic curve. That is, there is a function $c:\mathbb{R}\to\mathbb{R}$ whose graph is characteristic, a quadratic function $k:\mathbb{R}\to\mathbb{R}$, and $d\in(0,\infty)$ such that

$$\begin{aligned} Q_w &= \{(x,0,z)\in V_0 : a\leqslant x\leqslant b \text{ and } g_1(x)\leqslant z\leqslant c(x)\},\\ Q_{w'} &= \{(x,0,z)\in V_0 : a\leqslant x\leqslant b \text{ and } c(x)\leqslant z\leqslant g_2(x)\},\\ R_w &= \{(x,0,z)\in V_0 : a\leqslant x\leqslant b \text{ and } k(x)-d\leqslant z\leqslant k(x)\},\\ R_{w'} &= \{(x,0,z)\in V_0 : a\leqslant x\leqslant b \text{ and } k(x)\leqslant z\leqslant k(x)+d\}. \end{aligned}$$

Then, $\delta_x(Q_w)=\delta_x(Q_{w'})=\delta_x(Q_v)$ and $\delta_z(Q_w)=\delta_z(Q_{w'})=d$. Furthermore, Q_w and $Q_{w'}$ are assumed to be μ-rectilinear. Thus,

$$\max\{\|(k-d)-g_1\|_{L_\infty(4I)}, \|k-c\|_{L_\infty(4I)}, \|(k+d)-g_2\|_{L_\infty(4I)}\}\leqslant \mu d. \tag{4.4}$$

In either case, $Q_v=Q_w\cup Q_{w'}$ and $Q_w, Q_{w'}$ have disjoint interiors. Let $\mathcal{V}_{\mathsf{V}}(\Delta)\subseteq\mathcal{V}(\Delta)$ be the set of vertically cut vertices, and let $\mathcal{V}_{\mathsf{H}}(\Delta)\subseteq\mathcal{V}(\Delta)$ be the set of horizontally cut vertices.

It follows from the above definition that $v \leqslant w$ if and only if $Q_v \subseteq Q_w$. Furthermore, if the interior of Q_v intersects Q_w, then either $v \leqslant w$ or $w \leqslant v$.

LEMMA 4.5. *For every $\varepsilon > 0$ there exists $0 < \mu = \mu(\varepsilon) \leqslant \frac{1}{32}$ such that if Q is a μ-rectilinear pseudoquad, then*

$$(1-\varepsilon)\delta_x(Q)\delta_z(Q) \leqslant |Q| \leqslant (1+\varepsilon)\delta_x(Q)\delta_z(Q). \tag{4.5}$$

If Q is horizontally or vertically cut as in Definition 4.4 and Q' is a child of Q, then

$$\left(\tfrac{1}{2}-\varepsilon\right)|Q| \leqslant |Q'| \leqslant \left(\tfrac{1}{2}+\varepsilon\right)|Q|. \tag{4.6}$$

If Q is vertically cut, then $\delta_x(Q') = \frac{1}{2}\delta_x(Q)$, $\delta_z(Q') = \delta_z(Q)$, and $\alpha(Q') = \frac{1}{2}\alpha(Q)$. If Q is horizontally cut, then $\delta_x(Q') = \delta_x(Q)$, and

$$\left(\tfrac{1}{2}-2\mu\right)\delta_z(Q) \leqslant \delta_z(Q') \leqslant \left(\tfrac{1}{2}+2\mu\right)\delta_z(Q). \tag{4.7}$$

Finally,

$$(\sqrt{2}-\varepsilon)\alpha(Q) \leqslant \alpha(Q') \leqslant (\sqrt{2}+\varepsilon)\alpha(Q). \tag{4.8}$$

When $\varepsilon = \frac{1}{4}$, we can take here $\mu = \frac{1}{32}$.

Proof. Suppose that $\mu \leqslant \frac{1}{8}\varepsilon$. Let (Q, R) be a μ-rectilinear pseudoquad. Suppose that g_1 and g_2 (resp. h_1 and h_2) be the lower and upper bounds of Q (resp. R), and let I be the base of Q. Then $|R| = \delta_x(Q)\delta_z(Q)$ and

$$\big||Q| - \delta_x(Q)\delta_z(Q)\big| = \big||Q| - |R|\big| \leqslant \int_I |g_1 - h_1|\,dx + \int_I |g_2 - h_2|\,dx \leqslant 2\mu\delta_x(Q)\delta_z(Q), \tag{4.9}$$

so (4.5) is satisfied.

Let Q' be a child of Q. If Q is vertically cut, then the formulas for $\delta_x(Q')$, $\delta_z(Q')$, and $\alpha(Q')$ follow from Definition 4.4. As Q' is μ-rectilinear, (4.9) implies that

$$\big||Q'| - \tfrac{1}{2}|Q|\big| \leqslant \big||Q'| - \delta_x(Q')\delta_z(Q')| + \tfrac{1}{2}|\delta_x(Q)\delta_z(Q) - |Q|\big| \leqslant 2\mu\delta_x(Q)\delta_z(Q) \leqslant 4\mu|Q|,$$

so Q satisfies (4.6) if Q is vertically cut.

If Q is horizontally cut, then we have $\delta_x(Q') = \delta_x(Q)$ by Definition 4.4. Let c, k, and $d = \delta_z(Q')$ be as in Definition 4.4, and let $t \in I$. Since $\|g_i - h_i\|_{L_\infty(I)} \leqslant \mu\delta_z(Q)$ for $i \in \{1, 2\}$ and $\delta_z(Q) = h_2 - h_1$, we have

$$(1-2\mu)\delta_z(Q) \leqslant g_2(t) - g_1(t) \leqslant (1+2\mu)\delta_z(Q).$$

By (4.4),

$$(1-\mu)\cdot 2d \leqslant g_2(t) - g_1(t) \leqslant (1+\mu)\cdot 2d.$$

Then,

$$d \leqslant \frac{1+2\mu}{2-2\mu}\delta_z(Q) < \delta_z(Q),$$

so

$$|2d-\delta_z(Q)| \leqslant |2d-(g_2(t)-g_1(t))|+|(g_2(t)-g_1(t))-\delta_z(Q)| \leqslant 4\mu\delta_z(Q),$$

proving (4.7). This directly implies equation (4.8), and the horizontally cut case of (4.6) follows from the above calculation and (4.9). □

The following two lemmas will be helpful later.

LEMMA 4.6. *For any quadratic function $q:\mathbb{R}\to\mathbb{R}$ and any $t\in\mathbb{R}$,*

$$|q(t)| \leqslant (1+t+2t^2)\|q\|_{L_\infty([-1,1])}.$$

Proof. One only needs to note that, since q is quadratic, for any $t\in\mathbb{R}$ we have

$$q(t) = q(0)+t\cdot\frac{q(1)-q(-1)}{2}+t^2\cdot\frac{q(-1)-2q(0)+q(1)}{2}. \qquad \square$$

LEMMA 4.7. *For every $r\geqslant 2$ there is $\mu=\mu(r)>0$ such that, if Δ is a μ-rectilinear foliated patchwork and $v,w\in\mathcal{V}(\Delta)$ satisfy $w\leqslant v$, then $rQ_w\subseteq rQ_v$.*

Proof. It suffices to consider the case that $w\in\mathcal{C}(v)$. If v is vertically cut, this holds vacuously, so suppose that v is horizontally cut. Let g_1 and g_2 (resp. h_1 and h_2) be the lower and upper bounds of Q_v (resp. R_v), and let I be their base. Denote $m_v=\frac{1}{2}(h_1+h_2)$. Then, rR_v is bounded by $m_v\pm\frac{1}{2}r^2\delta_z(Q_v)$.

Let c, k, and $d=\delta_z(Q_w)$ be as in Definition 4.4. By Lemma 4.5, we have $d\leqslant\frac{3}{4}\delta_z(Q_v)$. Then,

$$\|(k-d)-h_1\|_{L_\infty(4I)} \leqslant \|k-d-g_1\|_{L_\infty(4I)}+\|g_1-h_1\|_{L_\infty(4I)} \leqslant \mu d+\mu\delta_z(Q_v) \leqslant 2\mu\delta_z(Q_v).$$

Likewise, $\|(k+d)-h_2\|_{L_\infty(4I)}\leqslant 2\mu\delta_z(Q_v)$. By Lemma 4.6, since k, h_1, and h_2 are quadratic functions, if μ is at most a sufficiciently small universal constant multiple of r^{-2}, then

$$\max\{\|(k-d)-h_1\|_{L_\infty(rI)}, \|(k+d)-h_2\|_{L_\infty(rI)}\} \leqslant \tfrac{1}{16}\delta_z(Q_v).$$

By the triangle inequality,

$$\|k-m_v\|_{L_\infty(rI)} \leqslant \tfrac{1}{16}\delta_z(Q_v).$$

Suppose that Q_w is the lower half of Q_v, so that Q_w is bounded by g_1 and c, and R_w is bounded by $k-d$ and k. Let $m_w=k-\frac{1}{2}d$, so that rQ_w is bounded by $m_w\pm\frac{1}{2}r^2d$. For $x\in rI$,

$$|m_v(x)-m_w(x)| \leqslant \tfrac{1}{16}\delta_z(Q_v)+\tfrac{1}{2}d \leqslant \tfrac{7}{16}\delta_z(Q_v) \leqslant \tfrac{1}{2}r^2\delta_z(Q_v)-\tfrac{1}{2}r^2d,$$

so

$$\left[m_w(x)-\tfrac{1}{2}r^2d, m_w(x)+\tfrac{1}{2}r^2d\right] \subseteq \left[m_v(x)-\tfrac{1}{2}r^2\delta_z(Q_v), m_v(x)+\tfrac{1}{2}r^2\delta_z(Q_v)\right].$$

That is, $rQ_w \subseteq rQ_v$. The case that Q_w is the upper half of Q_v is treated analogously. □

Let Δ be a μ-rectilinear foliated patchwork for a μ-rectilinear pseudoquad Q. For every subset of vertices $S \subseteq \mathcal{V}(\Delta)$, define the *weight* of S to be

$$W(S) \stackrel{\text{def}}{=} \sum_{w\in S} \frac{|Q_w|}{\alpha(Q_w)^4} \stackrel{(4.3)}{=} \sum_{w\in S} \frac{\delta_z(Q_w)^2}{\delta_x(Q_w)^4}|Q_w|. \tag{4.10}$$

We will use this to define a weighted Carleson condition which is a variant of the Carleson packing condition that is used in the theory of uniform rectifiability [27].

Definition 4.8. (Weighted Carleson packing condition) Let Δ be a μ-rectilinear foliated patchwork for a μ-rectilinear pseudoquad Q. We say that Δ satisfies a *weighted Carleson packing condition* or that Δ is *weighted Carleson* with constant $C \in (0,\infty)$ if every $v \in \mathcal{V}(\Delta)$ satisfies

$$W(\mathcal{D}(v)\cap \mathcal{V}_{\mathsf{V}}(\Delta)) \leqslant C|Q_v|, \tag{4.11}$$

where we recall that $\mathcal{D}(v)$ are the descendants of v and $\mathcal{V}_{\mathsf{V}}(\Delta)$ are the vertically cut vertices.

Remark 4.9. Vertical cuts increase W and horizontal cuts decrease it. More precisely, suppose that $v, w \in \mathcal{V}(\Delta)$ and w is a child of v. If v is vertically cut, then by Lemma 4.5 (with $\varepsilon = \frac{1}{4}$),

$$W(\{w\}) = \alpha(Q_w)^{-4}|Q_w| = 16\alpha(Q_v)^{-4}|Q_w| \geqslant 16\alpha(Q_v)^{-4}\cdot\left(\tfrac{1}{2}-\varepsilon\right)|Q_v| \geqslant 4W(\{v\}). \tag{4.12}$$

When $\varepsilon \to 0^+$, $W(\{w\})$ approaches $8W(\{v\})$. If v is horizontally cut, then by Lemma 4.5 (with $\varepsilon = \frac{1}{4}$),

$$W(\{w\}) = \alpha(Q_w)^{-4}|Q_w| \leqslant (\sqrt{2}-\varepsilon)^{-4}\left(\tfrac{1}{2}+\varepsilon\right)W(\{v\}) \leqslant \tfrac{3}{7}W(\{v\}), \tag{4.13}$$

and $W(\{w\})$ approaches $\frac{1}{8}W(\{v\})$ when $\varepsilon \to 0^+$.

The next lemma implies that, even though only $\mathcal{V}_{\mathsf{V}}(\Delta)$ appears in (4.11), this condition formally implies bounds on $\mathcal{V}_{\mathsf{H}}(\Delta)$ as well.

LEMMA 4.10. *Let Δ be a $\frac{1}{32}$-rectilinear foliated patchwork for Q with*

$$W(\mathcal{V}_{\mathsf{V}}(\Delta)) < \infty,$$

and let v_0 be the root of Δ. Then,

$$W(\mathcal{V}_{\mathsf{V}}(\Delta)) \lesssim W(\mathcal{V}_{\mathsf{H}}(\Delta)) \lesssim W(\mathcal{V}_{\mathsf{V}}(\Delta)) + \alpha(Q)^{-4}|Q|.$$

Proof. Let $\mathcal{T}_{\mathsf{H}}$ (resp. $\mathcal{T}_{\mathsf{V}}$) be the set of connected components of the subgraph of Δ spanned by $\mathcal{V}_{\mathsf{H}}(\Delta)$ (resp. $\mathcal{V}_{\mathsf{V}}(\Delta)$). Let $T\in\mathcal{T}_{\mathsf{H}}$ and let $M(T)$ be the maximal vertex of T. Each vertex of T is horizontally cut, so by (4.13), we have $W(\mathcal{C}(v))\leqslant\frac{6}{7}W(\{v\})$ for all $v\in\mathcal{V}(T)$. Therefore, $W(\mathcal{V}(T))\asymp W(\{M(T)\})$, because

$$W(\mathcal{V}(T))=\sum_{n=0}^{\infty}W(\mathcal{C}^n(M(T))\cap\mathcal{V}(T))\leqslant\sum_{n=0}^{\infty}\left(\frac{6}{7}\right)^n W(\{M(T)\})\lesssim W(\{M(T)\}).$$

Hence, if we denote $S_M=\{M(T):T\in\mathcal{T}_{\mathsf{H}}\}$, then $W(\mathcal{V}_{\mathsf{H}}(\Delta))\asymp W(S_M)$.

Now, take $T\in\mathcal{T}_{\mathsf{V}}$. By (4.12), we have $W(\{w\})\geqslant 4W(\{v\})$ for all $v\in\mathcal{V}(T)$ and $w\in\mathcal{C}(v)$. As $W(\mathcal{V}(T))<\infty$, it follows that T must be finite. Let

$$m(T)=\{w\in\mathcal{V}(T):\mathcal{C}(w)\not\subseteq\mathcal{V}(T)\}$$

be the lower boundary of T, and let $S_m=\bigcup_{T\in\mathcal{T}_{\mathsf{V}}}m(T)$.

For all $v\in\mathcal{V}(T)$, let $A(v)=\{w\in\mathcal{V}(T):w\geqslant v\}$ be the set of ancestors of v in T. By (4.12),

$$W(A(v))\leqslant\sum_{n=0}^{|A(v)|-1}4^{-n}W(\{v\})\leqslant 2W(\{v\}).$$

Every vertex of T is an ancestor of a leaf, so it follows that

$$W(\mathcal{V}(T))\leqslant W\left(\bigcup_{v\in m(T)}A(v)\right)\leqslant\sum_{v\in m(T)}W(A(v))\leqslant 2W(m(T))\leqslant 2W(\mathcal{V}(T)).$$

Therefore, $W(\mathcal{V}_{\mathsf{V}}(\Delta))\asymp W(S_m)$.

If $v\in S_M$ and $v\neq v_0$, then $\mathcal{P}(v)$ is horizontally cut and has a vertically cut child, so $\mathcal{P}(v)\in S_m$. In fact, $\mathcal{P}(S_M\setminus\{v_0\})=S_m$. Since $W(\{v\})\asymp W(\{\mathcal{P}(v)\})$ for all v and since $\mathcal{P}$ is a two-to-one map, it follows that $W(S_M\setminus\{v_0\})\asymp W(S_m)$. Therefore,

$$W(\mathcal{V}_{\mathsf{H}}(\Delta))\asymp W(S_M)\lesssim W(S_m)+W(\{v_0\})\asymp W(\mathcal{V}_{\mathsf{V}}(\Delta))+\alpha(Q)^{-4}|Q|,$$

and

$$W(\mathcal{V}_{\mathsf{V}}(\Delta))\asymp W(S_m)\asymp W(S_M\setminus\{v_0\})\leqslant W(S_M). \qquad \square$$

Suppose that $\Delta=(Q_v)_{v\in\mathcal{V}(\Delta)}$ is a μ-rectilinear foliated patchwork for $\Gamma=\Gamma_f$. For $\sigma>0$, a set of σ*-approximating planes* for Δ is a collection of vertical planes $(P_v)_{v\in\mathcal{V}_{\mathsf{H}}(\Delta)}$ such that, for every $v\in\mathcal{V}_{\mathsf{H}}(\Delta)$, if $f_v\colon V_0\to\mathbb{R}$ is the affine function such that $\Gamma_{f_v}=P_v$, then

$$\frac{\|f_v-f\|_{L_1(10Q_v)}}{|Q_v|}\leqslant\sigma\frac{\delta_z(Q_v)}{\delta_x(Q_v)}. \tag{4.14}$$

The following lemma verifies that the choice of right-hand side in (4.14) produces a condition that is invariant under stretch automorphisms and shear automorphisms.

LEMMA 4.11. *Let* $\Delta=(Q_v)_{v\in\mathcal{V}(\Delta)}$ *be a* μ*-rectilinear foliated patchwork for an intrinsic Lipschitz graph* $\Gamma=\Gamma_f$ *with a set* $(P_v)_{v\in\mathcal{V}_{\mathsf{H}}(\Delta)}$ *of* σ*-approximating planes and let* $r\colon\mathbb{H}\to\mathbb{H}$ *be a left translation, a stretch automorphism, or a shear map. Moreover, let* $\hat{r}=\Pi\circ r\colon V_0\to V_0$ *be the map induced on* V_0*. Then,*

$$\Delta'=((\hat{r}(Q_v),\hat{r}(R_v)))_{v\in\mathcal{V}(\Delta)}$$

is a μ*-rectilinear foliated patchwork for* $r(\Gamma)$ *and*

$$(r(P_v))_{v\in\mathcal{V}_{\mathsf{H}}(\Delta)}$$

is a set of σ*-approximating planes for* Δ'*.*

Proof. By Lemmas 2.10 and 4.2, $r(\Gamma)$ is an intrinsic Lipschitz graph and the elements of Δ' are μ-rectlinear pseudoquads for $r(\Gamma)$. It is straightforward to check that Definition 4.4 holds for Δ'. Let $v\in\mathcal{V}_{\mathsf{H}}(\Delta)$ and let $f_v\colon V_0\to\mathbb{R}$ be the affine function such that $P_v=\Gamma_{f_v}$. By Lemma 2.9, there are functions $\hat{f}$ and $\hat{f}_v$ such that

$$r(\Gamma)=\Gamma_{\hat{f}}\quad\text{and}\quad r(P_v)=\Gamma_{\hat{f}_v}.$$

If r is a left translation or a shear map and $w\in 10Q_v$, then $\hat{r}(w)\in 10\hat{r}(Q_v)$ and

$$|\hat{f}(\hat{r}(w))-\hat{f}_v(\hat{r}(w))|=|f(w)-f_v(w)|.$$

In this case, $\delta_x(Q_v)=\delta_x(\hat{r}(Q_v))$ and $\delta_z(Q_v)=\delta_z(\hat{r}(Q_v))$, so if P_v is a σ-approximating plane for Q_v, then $r(P_v)$ is a σ-approximating plane for $\hat{r}(Q_v)$.

If $r=s_{a,b}$ for some $a,b\in\mathbb{R}\setminus\{0\}$, then we have $\hat{r}=r|_{V_0}$, $\hat{r}(10Q_v)=10\hat{r}(Q_v)$, and for any $w\in 10Q_v$,

$$|\hat{f}(\hat{r}(w))-\hat{f}_v(\hat{r}(w))|=|b|\cdot|f(w)-f_v(w)|.$$

In this case,

$$\delta_x(\hat{r}(Q_v))=|a|\delta_x(Q_v)\quad\text{and}\quad\delta_z(\hat{r}(Q_v))=|ab|\delta_z(Q_v),$$

so, by (4.14),

$$\begin{aligned}\frac{\|\hat{f}_v-\hat{f}\|_{L_1(10\hat{r}(Q_v))}}{|\hat{r}(Q_v)|}&=\frac{|a^2b^2|\,\|f_v-f\|_{L_1(10Q_v)}}{|a^2b|\cdot|Q_v|}\\&\leqslant|b|\sigma\frac{\delta_z(Q_v)}{\delta_x(Q_v)}=\sigma\frac{\delta_z(\hat{r}(Q_v))}{\delta_x(\hat{r}(Q_v))}.\end{aligned}$$

□

5. Foliated corona decompositions

An intrinsic graph that admits rectilinear foliated patchworks that satisfy a weighted Carleson condition and have approximating planes is said to have a foliated corona decomposition.

Definition 5.1. Fix $0<\mu_0\leqslant\frac{1}{32}$ and $D:\mathbb{R}^+\times\mathbb{R}^+\to\mathbb{R}^+$. We say that an intrinsic Lipschitz graph Γ has a (D,μ_0)-*foliated corona decomposition* if for every $0<\mu\leqslant\mu_0$, every $\sigma>0$ and every μ-rectilinear pseudoquad $Q\subseteq V_0$, there is a μ-rectilinear foliated patchwork Δ for Q such that Δ is $D(\mu,\sigma)$-weighted Carleson and has a set of σ-approximating planes.

The following theorem is a more precise formulation of Theorem 1.18.

THEOREM 5.2. *For every $0<\lambda<1$ there is a function $D_\lambda:\mathbb{R}^+\times\mathbb{R}^+\to\mathbb{R}^+$ such that for any $0<\mu_0\leqslant\frac{1}{32}$, any intrinsic λ-Lipschitz graph has a (D_λ,μ_0)-foliated corona decomposition.*

Definition 5.1 requires the root of the foliated patchwork to be μ-rectilinear; the next lemma shows that intrinsic Lipschitz graphs have many μ-rectilinear pseudoquads.

LEMMA 5.3. *Let $\mu_0>0$, let $0<\lambda<1$, and let $\Gamma=\Gamma_f$ be an intrinsic λ-Lipschitz graph. There is an $\alpha_0>0$ with the following property. Let Q be a pseudoquad for Γ, let v be a point in the lower boundary of Q and suppose that vZ^s is in the upper boundary. Let $r=\delta_x(Q)$. If $r/\sqrt{s}\leqslant\alpha_0$, then there is a parabolic rectangle R such that (Q,R) is μ_0-rectilinear.*

Proof. Denote

$$L\stackrel{\mathrm{def}}{=}\frac{\lambda}{\sqrt{1-\lambda^2}}\quad\text{and}\quad\alpha_0\stackrel{\mathrm{def}}{=}\min\left\{\sqrt{\frac{\mu_0}{16L}},\frac{\mu_0(1-\lambda)}{24}\right\}.$$

Let $g_1,g_2:\mathbb{R}\to\mathbb{R}$ be the lower and upper bounds of Q, and let I be its base. After a translation, we may suppose that $v=\mathbf{0}$ and $f(v)=0$. Then, $I\subseteq[-r,r]$, $g_1(0)=0$, $g_2(0)=s$, and $g_1'(0)=-f(\mathbf{0})=0$. Let $R=I\times[0,s]$; we claim that (Q,R) is a μ_0-rectilinear pseudoquad.

It suffices to show that, for all $t\in[-4r,4r]$ and $i\in\{1,2\}$, we have $|g_i(t)-g_i(0)|\leqslant\mu_0 s$. By Lemma 2.7, for all $t\in[-4r,4r]$, we have

$$|g_i(t)-(g_i(0)+tg_i'(0))|\leqslant 8r^2L\leqslant\tfrac12\mu_0 s.$$

In particular, $|g_1(t)|\leqslant\mu_0 s$. Lemma 2.3 implies that

$$|g_2'(0)|=|f(Z^s)-f(\mathbf{0})|\leqslant\frac{3}{1-\lambda}d(\mathbf{0},Z^s)=\frac{3\sqrt{s}}{1-\lambda}\leqslant\frac{\mu_0}{8}\cdot\frac{s}{r},$$

so, if $|t|\leqslant 4r$, then

$$|g_2(t)-g_2(0)|\leqslant\frac{\mu_0}{8}\cdot\frac{s}{r}\cdot 4r+\frac{\mu_0 s}{2}\leqslant\mu_0 s.\qquad\square$$

Corollary 5.4. *Continuing with the setting and notation of Lemma 5.3, any $\frac{1}{32}$-rectilinear pseudoquad Q such that $\alpha(Q)\leqslant\frac{1}{2}\alpha_0$ is μ_0-rectilinear.*

Proof. Let v be in the lower boundary of Q. Then there is an $s\geqslant\left(1-\frac{1}{16}\right)\delta_z(Q)$ such that vZ^s is in the upper boundary. If $\alpha(Q)\leqslant\frac{1}{2}\alpha_0$, then

$$\delta_x(Q)\leqslant \tfrac{1}{2}\alpha_0\sqrt{\delta_z(Q)}\leqslant \alpha_0\sqrt{s},$$

so Lemma 5.3 implies that Q is μ_0-rectilinear. □

The following lemma shows that the choice of μ_0 is not important; we can increase μ_0 at the cost of an increase in D.

Lemma 5.5. *For any $\lambda>0$ and $0<\mu_0<\mu_0'\leqslant\frac{1}{32}$, and any $D:\mathbb{R}^+\times\mathbb{R}^+\to\mathbb{R}^+$, there exists $D':\mathbb{R}^+\times\mathbb{R}^+\to\mathbb{R}^+$ such that, if Γ is an intrinsic λ-Lipschitz graph that has a (D,μ_0)-foliated corona decomposition, then Γ also has a (D',μ_0')-foliated corona decomposition.*

Proof. Fix $0<\mu<\mu_0'$ and $0<\sigma<1$. Let $\alpha_0>0$ be as in Lemma 5.3. Suppose that we are given a μ-rectilinear pseudoquad Q. We wish to construct a rectilinear foliated patchwork for Q with a set of σ-approximating planes. If $\alpha(Q)<\frac{1}{2}\alpha_0$, then, by Corollary 5.4, Q is μ_0-rectilinear, and since Γ admits a (D,μ_0)-foliated corona decomposition, the desired patchwork and set of approximating planes for Q exist.

We thus suppose that $\alpha(Q)\geqslant\frac{1}{2}\alpha_0$ and denote

$$i_0=\left\lceil \log_2\frac{\alpha(Q)}{\alpha_0}\right\rceil+1.$$

We will construct a rectilinear foliated patchwork for Q by first cutting Q vertically i_0 times into pseudoquads $P_1,...,P_{2^{i_0}}$ of width $2^{-i_0}\delta_x(Q_0)$, height $\delta_z(Q)$, and aspect ratio

$$\alpha(P_i)=\frac{\delta_x(P_i)}{\sqrt{\delta_z(P_i)}}=2^{-i_0}\alpha(Q)<\frac{\alpha_0}{2}.$$

By Corollary 5.4, each P_i is μ_0-rectilinear, and thus admits a $D(\mu,\sigma)$-weighted Carleson rectilinear foliated patchwork and a set of σ-approximating planes. Combining these patchworks, we obtain a rectilinear foliated patchwork Δ' for Q_0. Let v_0 be its root vertex. Note that, for any $0\leqslant m\leqslant i_0$ and any $w\in\mathcal{C}^m(v_0)$, we have $\alpha(Q_w)=2^{-m}\alpha(Q_0)$.

It remains to check that Δ' is weighted Carleson. Let $p_1,...,p_{2^{i_0}}\in\mathcal{V}(\Delta')$ be the vertices such that $P_j=Q_{p_j}$. If $v\in\mathcal{V}(\Delta')$ and $v\leqslant p_j$ for some j, then Q_v satisfies the weighted Carleson condition (4.11) with constant at most $D(\mu,\sigma)$.

Otherwise, v is an ancestor of some p_j and $Q_v=P_a\cup...\cup P_b$ for some $a\leqslant b$. For each $w\in\mathcal{V}(\Delta')$, let $A(w)$ be the set of ancestors of w. Every ancestor of p_j except possibly p_j

itself is vertically cut, so by (4.12), the weight of $\mathcal{P}^k(p_j)$ decays exponentially. Thus

$$W(A(p_j))=\sum_{k=0}^{i_0} W(\{\mathcal{P}^k(p_j)\}) \overset{(4.12)}{\leqslant} \sum_{k=0}^{i_0} 4^{-k} W(\{p_j\}) \leqslant 2W(\{p_j\}).$$

For each $w \in \mathcal{V}(\Delta')$, let

$$\mathcal{D}_{\mathsf{V}}(w)=\mathcal{D}(w)\cap \mathcal{V}_{\mathsf{V}}(\Delta')$$

be the set of vertically cut descendants of w. As every element of $\mathcal{D}_{\mathsf{V}}(v)$ is a descendant or an ancestor of some p_j with $a\leqslant j\leqslant b$,

$$\begin{aligned} W(\mathcal{D}_{\mathsf{V}}(v)) &\leqslant \sum_{j=a}^{b}[W(\mathcal{D}_{\mathsf{V}}(p_j))+W(A(p_j))] \leqslant \sum_{j=a}^{b}[D(\mu,\sigma)|P_j|+2W(\{p_j\})] \\ &= D(\mu,\sigma)|Q_v|+2\cdot\alpha(P_a)^{-4}|Q_v| \leqslant D(\mu,\sigma)|Q_v|+\alpha_0^{-4}|Q_v|. \end{aligned}$$

Therefore, Δ' is $(D(\mu,\sigma)+\alpha_0^{-4})$-weighted Carleson. □

6. Vertical perimeter and foliated corona decompositions

In this section we will assume Theorem 5.2 and prove the following theorem, which bounds the vertical perimeter of half-spaces bounded by intrinsic Lipschitz graphs.

THEOREM 6.1. *For any $0<\lambda<1$ and $r>0$, if Γ is an intrinsic λ-Lipschitz graph, then*

$$\|\bar{\mathsf{v}}_{B_r(\mathbf{0})}(\Gamma^+)\|_{L_4(\mathbb{R})} \lesssim_\lambda r^3.$$

This coincides with the bound (1.32) needed in §1.2.1. Combined with the reduction from arbitrary sets to intrinsic Lipschitz graphs described in that section, this completes the proof of Theorem 1.1.

6.1. Vertical perimeter for graphs with foliated corona decompositions

Theorem 6.1 is a consequence of the following lemma.

LEMMA 6.2. *Suppose that $f: V_0\to\Gamma$ is intrinsic Lipschitz and denote $\Gamma=\Gamma_f$. Fix $\sigma>0$. Let $Q\subseteq V_0$ be a $\frac{1}{32}$-rectilinear pseudoquad. Let Δ be a $\frac{1}{32}$-rectilinear foliated patchwork for Q and let $(P_v)_{v\in\mathcal{V}_{\mathsf{H}}(\Delta)}$ be a set of σ-approximating planes. Denoting $t_0=-\log_4 \delta_z(Q)$, we have*

$$\|\bar{\mathsf{v}}^P_{Q,f}\|_{L_4([t_0,\infty))} \lesssim \sigma|Q|^{3/4}W(\mathcal{V}(\Delta))^{1/4}. \tag{6.1}$$

Note that while the intrinsic Lipschitz constant of f appears in Theorem 6.1, it does not appear in (6.1). Indeed, this bound is invariant under scalings and stretch automorphisms; if Γ, Q, and $(\Delta,(Q_v)_{v\in\mathcal{V}(\Delta)})$ are as in Lemma 6.2, $a,b>0$, $s=s_{a,b}$, and

$$\hat{s}=\Pi\circ s|_{V_0}=s|_{V_0},$$

then, by Lemma 4.11, $\hat{s}(Q)$ is a pseudoquad in $s(\Gamma)=\Gamma_{\hat{f}}$, where $\hat{f}=bf\circ\hat{s}^{-1}$. Furthermore, $\Delta'=(\hat{s}(Q_v))_{v\in\mathcal{V}(\Delta)}$ is a foliated patchwork for $\hat{s}(Q)$ and $(s(P_v))_{v\in\mathcal{V}_{\mathsf{H}}(\Delta)}$ is a set of σ-approximating planes.

By Lemma 4.2, $\alpha(\hat{s}(Q_v))=\sqrt{a/b}\,\alpha(Q_v)$ and $|\hat{s}(Q_v)|=a^2b|Q_v|$, so

$$W(\mathcal{V}(\Delta'))=b^3W(\mathcal{V}(\Delta))$$

and

$$|\hat{s}(Q)|^{3/4}W(\mathcal{V}(\Delta'))^{1/4}=(a^2b)^{3/4}|Q|^{3/4}b^{3/4}W(\mathcal{V}(\Delta))^{1/4}=(ab)^{3/2}|Q|^{3/4}W(\mathcal{V}(\Delta))^{1/4}.$$

If (6.1) holds for f and Q, then, by Lemma 2.12,

$$\begin{aligned}\|\bar{\mathsf{v}}^P_{\hat{s}(Q),\hat{f}}\|_{L_4([t_0-\log_4(ab),\infty))}&=(ab)^{3/2}\|\bar{\mathsf{v}}^P_{Q,f}\|_{L_4([t_0,\infty))}\\&\lesssim\sigma(ab)^{3/2}|Q|^{3/4}W(\mathcal{V}(\Delta))^{1/4}=\sigma|\hat{s}(Q)|^{3/4}W(\mathcal{V}(\Delta'))^{1/4}.\end{aligned}$$

That is, (6.1) holds for $\hat{f}$ and $s(Q)$.

To prove Lemma 6.2, we will need some lemmas on partitions and coherent sets. A collection $\{Q_1,...,Q_n\}$ of pseudoquads is a *partition* of Q if $Q=\bigcup_{i=1}^n Q_i$ and if the Q_i overlap only along their boundaries. A *coherent subtree* of T is a connected subtree such that, for every $v\in T$, either all children of v are contained in T or none of them are. A *coherent subset* of $\mathcal{V}(\Delta)$ is the vertex set of a coherent subtree.

LEMMA 6.3. *Let Δ be a rectilinear foliated patchwork for Q and suppose that $S\subseteq\mathcal{V}(\Delta)$ is coherent. Let $M=\max S$ be the maximal element of S and let $\min S$ be the set of minimal elements of S. Denote*

$$F_1=F_1(S)\stackrel{\text{def}}{=}\{p\in Q_M: \textit{there are infinitely many } v\in S \textit{ such that } p\in Q_v\}.$$

Then,

$$Q_M=F_1\bigcup\Big(\bigcup_{w\in\min S}Q_w\Big). \tag{6.2}$$

The interiors of $\{Q_w:w\in\min S\}$ are pairwise disjoint and disjoint from F_1. If S is finite, then $\min S$ is a partition of Q_M.

Proof. Let $v\in\min S$ and let $p\in\operatorname{int} Q_v$. If $u\in S$ and $p\in Q_u$, then either $u<v$ or $v\leqslant u$. The first is impossible by the minimality of v, so $v\leqslant u$. It follows that there are only finitely many $w\in S$ such that $p\in Q_w$ and no such w is minimal except v. That is, $\operatorname{int} Q_v$ is disjoint from F_1, and if $u\in\min S$ and $u\neq v$, then $\operatorname{int} Q_v$ is disjoint from $\operatorname{int} Q_u$.

If $p\in Q_M\setminus F_1$, then the set $\{v\in S:p\in Q_v\}$ is finite, and thus has a minimal element v_0. Let w be a child of v_0 such that $p\in Q_w$. The minimality of v_0 implies that $w\notin S$, so $v\in\min(S)$ by the coherence of S. This implies (6.2). $\square$

LEMMA 6.4. *Fix $0<\mu\leqslant\frac{1}{32}$ and let $(\Delta,(Q_v)_{v\in\Delta})$ be a μ-rectilinear foliated patchwork for Q with $W(\mathcal{V}_{\mathsf{V}}(\Delta))<\infty$. For any $0<\sigma\leqslant\delta_z(Q)$, denote $S_\sigma=\{v\in\mathcal{V}(\Delta):\delta_z(Q_v)\geqslant\sigma\}$ and let $F_\sigma=\min S_\sigma$. Then, $\{Q_v\}_{v\in F_\sigma}$ is a partition of Q into horizontally cut pseudoquads such that $\sigma\leqslant\delta_z(Q_v)<4\sigma$ for all $v\in F_\sigma$.*

Proof. By Definition 4.4 and Lemma 4.5, the height of every pseudoquad of Δ is equal to the height of its sibling and at most the height of its parent. Therefore, S_σ is coherent. If $v\in S_\sigma$, then

$$W(\{v\})=\alpha(Q_v)^{-4}|Q_v|\asymp\delta_z(Q_v)^3\delta_x(Q_v)^{-3}\geqslant\sigma^3\delta_x(Q)^{-3},$$

which is bounded away from zero, so Lemma 4.10 implies that S_σ is finite. By Lemma 6.3, F_σ partitions Q.

Suppose that $v\in F_\sigma$ and let $w\in\mathcal{C}(v)$. By the minimality of v, we have $v\in S_\sigma$ and $w\notin S_\sigma$, so $\delta_z(Q_v)\geqslant\sigma>\delta_z(Q_w)$. Since $\delta_z(Q_w)<\delta_z(Q_v)$, v is horizontally cut. Furthermore, by Lemma 4.5, $\sigma>\delta_z(Q_w)\geqslant\frac{1}{4}\delta_z(Q_v)$, so v is a horizontally cut pseudoquad such that $\sigma\leqslant\delta_z(Q_v)<4\sigma$, as desired. $\square$

We will use these partitions to decompose the parametric vertical perimeter of f and prove Lemma 6.2.

Proof of Lemma 6.2. By the remarks after Lemma 6.2, condition (6.1) is invariant under scaling, so we may rescale so that $\delta_z(Q)=1$. Let Δ be a $\frac{1}{32}$-rectilinear foliated patchwork for Q and let $(P_v)_{v\in\mathcal{V}_{\mathsf{H}}(\Delta)}$ be a set of σ-approximating planes. Without loss of generality, we suppose that $W(\mathcal{V}(\Delta))<\infty$. For each $v\in\mathcal{V}_{\mathsf{H}}(\Delta)$, let $f_v: V_0\to\mathbb{R}$ be the affine function such that $\Gamma_{f_v}=P_v$. For $i\in\mathbb{N}\cup\{0\}$, let $C_i=F_{2^{-2i-1}}\subseteq\mathcal{V}_{\mathsf{H}}(\Delta)$ be as in Lemma 6.4, so that $\{Q_v\}_{v\in C_i}$ is a partition of Q into horizontally-cut pseudoquads with heights in $[2^{-2i-1},2^{-2i+1})$. No vertex of Δ appears in more than one of the C_i's.

We start by bounding $\bar{\mathsf{v}}^P_{Q_v,f}(t)$ from above for each $v\in C_i$ for a fixed $i\in\mathbb{N}\cup\{0\}$. Then we have $2^{-2i}\leqslant 2\delta_z(Q_v)$, so Lemma 4.1 implies that $Z^{-2^{-2t}}Q_v\subseteq 10Q_v$ for any $t\in[i,i+1]$.

Therefore, since f_v is constant on vertical lines,

$$\begin{aligned}
\bar{\mathsf{v}}^P_{Q_v,f}(t) &= 2^t \int_{Q_v} |f(w)-f(wZ^{-2^{-2t}})|\,dw \\
&\leqslant 2^t \int_{Q_v} (|f(w)-f_v(w)|+|f_v(wZ^{-2^{-2t}})-f(wZ^{-2^{-2t}})|)\,dw \\
&= 2^t(\|f-f_v\|_{L_1(Q_v)}+\|f-f_v\|_{L_1(Z^{-2^{-2t}}Q_v)}) \\
&\overset{(4.14)}{\leqslant} 2^{t+1}|Q_v|\sigma\frac{\delta_z(Q_v)}{\delta_x(Q_v)} \overset{(4.5)}{\asymp} \sigma\delta_z(Q_v)^{3/2} \asymp \sigma\alpha(Q_v)^{-1}|Q_v|.
\end{aligned}$$

Since $\{Q_v\}_{v\in C_i}$ is a partition of Q, we have $\bar{\mathsf{v}}^P_{Q,f}(t)=\sum_{v\in C_i}\bar{\mathsf{v}}^P_{Q_v,f}(t)$ for all $t\in\mathbb{R}$. Thus,

$$\|\bar{\mathsf{v}}^P_{Q,f}\|_{L_4([i,i+1))} \leqslant \sum_{v\in C_i}\|\bar{\mathsf{v}}^P_{Q_v,f}\|_{L_4([i,i+1])} \lesssim \sum_{v\in C_i}\sigma\alpha(Q_v)^{-1}|Q_v|. \tag{6.3}$$

Consequently,

$$\begin{aligned}
\|\bar{\mathsf{v}}^P_{Q,f}\|^4_{L_4([0,\infty))} = \sum_{i=0}^{\infty}\|\bar{\mathsf{v}}^P_{Q,f}\|^4_{L_4([i,i+1))} &\overset{(6.3)}{\lesssim} \sigma^4\sum_{i=0}^{\infty}\bigg(\sum_{v\in C_i}\alpha(Q_v)^{-1}|Q_v|\bigg)^4 \\
&\leqslant \sigma^4\sum_{i=0}^{\infty}\bigg(\sum_{v\in C_i}|Q_v|\bigg)^3\bigg(\sum_{v\in C_i}\alpha(Q_v)^{-4}|Q_v|\bigg) \\
&\overset{(4.10)}{=} \sigma^4|Q|^3\sum_{i=0}^{\infty}W(C_i) \leqslant \sigma^4|Q|^3W(\mathcal{V}(\Delta)),
\end{aligned}$$

where the third step is an application of Hölder's inequality. □

Finally, we use Lemmas 2.11 and 2.12 to prove Theorem 6.1.

Proof of Theorem 6.1. After scaling, it suffices to prove the theorem in the case that $r=1$, i.e., that if Γ is the intrinsic graph of an intrinsic λ-Lipschitz function $f: V_0\to\mathbb{R}$, then

$$\|\bar{\mathsf{v}}_{B_1}(\Gamma^+)\|_{L_4(\mathbb{R})} \lesssim_\lambda 1. \tag{6.4}$$

By the definition (1.30),

$$\bar{\mathsf{v}}_{B_1}(\Gamma^+)(t) = 2^t|B_1\cap(\Gamma^+\triangle\Gamma^+Z^{2^{-2t}})| \leqslant 2^t|B_1| \lesssim 2^t \quad \text{for all } t\in\mathbb{R}.$$

Hence,

$$\|\bar{\mathsf{v}}_{B_1}(\Gamma^+)\|_{L_4((-\infty,a])} = \bigg(\int_{-\infty}^a \bar{\mathsf{v}}_{B_1}(\Gamma^+)(t)^4\,dt\bigg)^{1/4} \lesssim \bigg(\int_{-\infty}^a 2^{4t}\,dt\bigg)^{1/4} \lesssim 2^a \quad \text{for all } a\in\mathbb{R},$$

and therefore we have the following simple a-priori bound.

$$\|\bar{\mathrm{v}}_{B_1}(\Gamma^+)\|_{L_4(\mathbb{R})} \lesssim 2^a + \|\bar{\mathrm{v}}_{B_1}(\Gamma^+)\|_{L_4((a,\infty))} \quad \text{for all } a\in\mathbb{R}. \tag{6.5}$$

We will first treat the (trivial) case $B_5\cap\Gamma=\varnothing$, so that either $B_5\subseteq\Gamma^+$ or $B_5\subseteq\Gamma^-$. Without loss of generality, suppose that $B_5\subseteq\Gamma^+$. This implies that $B_1\subseteq\Gamma^+\cap Z^{2^{-2t}}\Gamma^+$ for any $t\geqslant 0$, so $\bar{\mathrm{v}}_{B_1}(\Gamma^+)(t)=0$, and therefore in this case (6.4) follows from the case $a=0$ of (6.5).

We may thus suppose, from now on, that $B_5\cap\Gamma\neq\varnothing$. Fix any point $p\in B_5\cap\Gamma$. Then, $d(p,\langle Y\rangle)\leqslant 5$ and $p=vY^{f(v)}$ for some $v\in V_0$ with $|f(v)|\leqslant 5$, so by Lemma 2.3, we have

$$|f(\mathbf{0})| \leqslant |f(v)| + |f(v)-f(\mathbf{0})| \leqslant 5 + \frac{2}{1-\lambda} d(p,\langle Y\rangle) \lesssim \frac{1}{1-\lambda}.$$

Likewise, for any $t\in\mathbb{R}$,

$$|f(Z^t)| \leqslant |f(\mathbf{0})| + \frac{2}{1-\lambda} d(\mathbf{0},Z^t) \lesssim 1 + \frac{\sqrt{|t|}}{1-\lambda}.$$

For $t\in\mathbb{R}$, let $g_t\colon\mathbb{R}\to\mathbb{R}$ be a function such that $g_t(0)=t$ and the graph of g_t is characteristic for f. By (2.17), $g_t'(0)=-f(Z^t)$, so by Lemma 2.7 and the estimate above,

$$\max_{x\in[-1,1]} |g_t(x)-t| \leqslant |g_t'(0)| + \frac{\lambda}{2\sqrt{1-\lambda^2}} = |f(Z^t)| + \frac{\lambda}{2\sqrt{1-\lambda^2}} \lesssim_\lambda 1+\sqrt{|t|}. \tag{6.6}$$

The right-hand side of (6.6) grows slower than $|t|$ as $|t|\to\infty$, so there is $t_0=t_0(\lambda)>1$ such that the pseudoquad Q that is bounded by the lines $x=\pm 1$ and $z=g_{\pm t_0}(x)$ is $\frac{1}{32}$-rectilinear and contains the projection $\Pi(B_1)$.

Theorem 5.2 applied with the choice of parameters $\mu_0=\frac{1}{32}$ and $\sigma=1$ shows that Q has a foliated patchwork Δ and a set of 1-approximating planes that satisfy

$$W(\mathcal{V}_{\mathsf{V}}(\Delta)) \lesssim_\lambda |Q| \lesssim_\lambda 1.$$

By Lemma 4.10, this implies that

$$W(\mathcal{V}(\Delta)) \lesssim W(\mathcal{V}_{\mathsf{V}}(\Delta)) + \alpha(Q)^{-4}|Q| \lesssim_\lambda 1. \tag{6.7}$$

By Lemmas 2.11 and 6.2, we conclude as follows:

$$\begin{aligned}
\|\bar{\mathrm{v}}_{B_1}(\Gamma^+)\|_{L_4(\mathbb{R})} &\lesssim \frac{1}{\sqrt{\delta_z(Q)}} + \|\bar{\mathrm{v}}_{B_1}(\Gamma^+)\|_{L_4([-\log_4\delta_z(Q),\infty))} \\
&\leqslant \frac{1}{\sqrt{\delta_z(Q)}} + \|\bar{\mathrm{v}}^P_{Q,f}\|_{L_4([-\log_4\delta_z(Q),\infty))} \\
&\lesssim \frac{1}{\sqrt{\delta_z(Q)}} + |Q|^{3/4} W(\mathcal{V}(\Delta))^{1/4} \lesssim_\lambda 1,
\end{aligned}$$

where the first step is an application of (6.5) with $a=-\log_4\delta_z(Q)$, the second step is an application of Lemma 2.11 because $Q\supseteq\Pi(B_1)$, the third step is an application of Lemma 6.2, and the final step holds due to (6.7) and because $|Q|\asymp\delta_z(Q)\asymp_\lambda 1$. $\square$

7. The subdivision algorithm: constructing a foliated corona decomposition

In this section, we will formulate an iterative subdivision algorithm (Lemma 7.3 below) and prove that, given certain propositions on the geometry of pseudoquads, this algorithm produces a foliated corona decomposition. In the following sections, we will prove these geometric propositions. Together, these arguments establish Theorem 5.2.

Fix $\lambda, \sigma \in (0,1)$. Let $f: V_0 \to \mathbb{R}$, and suppose that $\Gamma = \Gamma_f$ is an intrinsic λ-Lipschitz graph. Let $0 < \mu \leqslant \frac{1}{32}$. To show that Γ admits a foliated corona decomposition, we must show that, for any μ-rectilinear pseudoquad Q, there is a μ-rectilinear foliated patchwork Δ for Q which has a set of σ-approximating planes and such that Δ is weighted Carleson.

In order to describe the subdivision algorithm that produces Δ, we will introduce the *R-extended parametric normalized non-monotonicity* of Γ, denoted by $\Omega^P_{\Gamma^+,R}$, which is a measure on V_0 with density based on how horizontal line segments of length at most $R>0$ intersect Γ. If Γ is a plane, for instance, then $\Omega^P_{\Gamma^+,R}=0$, while $\Omega^P_{\Gamma^+,R}$ has positive density when Γ is bumpy at scale R.

This is in the spirit of the quantitative non-monotonicity used in [23] and [91], but it counts different segments, and, like the parametric vertical perimeter, it is defined in terms of the function f. We will give a full definition in §8 and discuss the relationship between extended non-monotonicity and quantitative non-monotonicity in Remarks 8.4 and 10.2. In §9, we will show that there is $c>0$ depending on the intrinsic Lipschitz constant of Γ such that the following kinematic formula (inequality) holds for every measurable subset $U \subseteq V_0$:

$$\sum_{i \in \mathbb{Z}} \Omega^P_{\Gamma^+, 2^{-i}}(U) \leqslant c|U|. \tag{7.1}$$

Definition 7.1. Suppose that $\eta, r, R > 0$ and let Q be a $\frac{1}{4}$-rectilinear pseudoquad. We say that Γ is *(η, R)-paramonotone* on rQ if it satisfies the following bound:

$$\frac{\Omega^P_{\Gamma^+, R\delta_x(Q)}(rQ)}{|Q|} \leqslant \frac{\eta}{\alpha(Q)^4}. \tag{7.2}$$

This condition is invariant under scalings, stretch maps, and shear maps; see the discussion immediately after the proof of Lemma 8.8 below.

One of the main results of [23] was that for small $\eta>0$, any η-monotone set is close to a plane in $\mathbb{H}$; this is a "stability version" of the characterization of monotone sets in [21]. The following proposition, which we will prove in §§10–12, states not only that paramonotone pseudoquads are close to vertical planes in $\mathbb{H}$, but also that their characteristic curves are close to the characteristic curves of their approximating planes.

PROPOSITION 7.2. *There is a universal constant $r>1$ such that, for any $\sigma>0$ and any $0<\zeta\leqslant\frac{1}{32}$, there are $\eta, R>0$ such that, if $\Gamma=\Gamma_f$ is the intrinsic Lipschitz graph of $f\colon V_0\to\mathbb{R}$, and if Q is a $\frac{1}{32}$-rectilinear pseudoquad for Γ such that Γ is (η,R)-paramonotone on rQ, then the folowing statements hold.*

(1) *There is a vertical plane $P\subseteq\mathbb{H}$ (a σ-approximating plane) and an affine function $F\colon V_0\to\mathbb{R}$ such that P is the intrinsic graph of F and*

$$\frac{\|F-f\|_{L_1(10Q)}}{|Q|}\leqslant\sigma\frac{\delta_z(Q)}{\delta_x(Q)}. \tag{7.3}$$

(2) *Let $u\in 4Q$ and let $g_\Gamma, g_P\colon\mathbb{R}\to\mathbb{R}$ be such that $\{z=g_\Gamma(x)\}$ (resp. $\{z=g_P(x)\}$) is a characteristic curve for Γ (resp. P) that passes through u. Then,*

$$\|g_P-g_\Gamma\|_{L_\infty(4I)}\leqslant\zeta\delta_z(Q).$$

It is important to observe that the bounds in Proposition 7.2 do not depend on the intrinsic Lipschitz constant of f. Indeed, this proposition holds when Γ is merely the intrinsic graph of a continuous function. This is important because paramonotonicity is invariant under stretch automorphisms; a bound that depended on the intrinsic Lipschitz constant of Γ would not be invariant.

Proposition 7.2 allows us to construct a μ-rectilinear foliated patchwork and a collection of σ-approximating planes by recursively subdividing Q according to a greedy algorithm.

LEMMA 7.3. *Let r be as in Proposition 7.2. Fix $0<\mu\leqslant\frac{1}{32}$ and $\sigma>0$. There are $\eta, R>0$ with the following property. Let Γ be an intrinsic Lipschitz graph and let Q be a μ-rectilinear pseudoquad. There is a μ-rectilinear foliated patchwork Δ for Q such that, for all $v\in\mathcal{V}(\Delta)$, Q_v is horizontally cut if and only if Γ is (η,R)-paramonotone on rQ, and Δ admits a set of σ-approximating planes.*

Proof. Let r, η, and R be positive constants such that Proposition 7.2 is satisfied with $\zeta=\frac{1}{4}\mu$.

We construct Δ by a greedy algorithm. Denote the root vertex of Δ by v_0 and let $Q_{v_0}=Q$; by assumption, it is μ-rectilinear. Suppose by induction that we have already constructed a μ-rectilinear pseudoquad (Q_v,R_v). Let $v\in\mathcal{V}(\Delta)$ be a vertex with children w and w'. Let $I=[a,b]$ be the base of Q_v and let $g_1,g_2\colon\mathbb{R}\to\mathbb{R}$ be its lower and upper bounds, respectively.

Suppose that Γ is not (η,R)-paramonotone on rQ_v. The vertical line $x=\frac{1}{2}(a+b)$ cuts Q_v and R_v vertically into two halves. Let Q_w and $Q_{w'}$ be the halves of Q_v and let

R_w and $R_{w'}$ be the halves of R_v. Since (Q_v, R_v) is μ-rectilinear, (Q_w, R_w) and $(Q_{w'}, R_{w'})$ are both μ-rectilinear.

Now suppose that Γ is (η, R)-paramonotone on rQ_v. Proposition 7.2 states that there is a σ-approximating plane P for Q_v such that, for every $u\in 4Q_v$, any characteristic curve of Γ that passes through u is $\zeta\delta_z(Q)$-close to the characteristic curve of P that passes through u. For $i\in\{1,2\}$, let $u_i=\big(\frac{1}{2}(a+b), g_i\big(\frac{1}{2}(a+b)\big)\big)$, and let m be the midpoint of u_1 and u_2.

Let $g_3\colon\mathbb{R}\to\mathbb{R}$ be a function whose graph is a characteristic curve for Γ that passes through m. Let Q_w and $Q_{w'}$ be the pseudoquads with base I that are bounded by the graphs of g_1, g_3, and g_2.

The characteristic curves of P that pass through u_1, u_2, and m are parallel evenly-spaced parabolas; let $h_1, h_2, h_3\colon V_0\to\mathbb{R}$ be the corresponding quadratic functions and let $d=h_2-h_3=h_3-h_1$ be the constant distance between them. Let R_w and $R_{w'}$ be the parabolic rectangles with base I that are bounded by these three parabolas. By Proposition 7.2, we have $\|g_i-h_i\|_{L_\infty(4I)}\leqslant\zeta\delta_z(Q)$ for $i\in\{1,2,3\}$. In particular, every $x\in I$ satisfies

$$\begin{aligned}|\delta_z(Q)-2d| &\leqslant |\delta_z(Q)-(g_2(x)-g_1(x))|+|g_2(x)-g_3(x)-d|+|g_3(x)-g_1(x)-d|\\ &\leqslant 3\zeta\delta_z(Q),\end{aligned}$$

so $d\geqslant\frac{1}{4}\delta_z(Q)$ and

$$\|g_i-h_i\|_{L_\infty(4I)}\leqslant 4\zeta d=\mu d$$

for $i\in\{1,2,3\}$. That is, (Q_w, R_w) and $(Q_{w'}, R_{w'})$ are μ-rectilinear and satisfy Definition 4.4 with $k=h_3$. We construct the desired rectilinear foliated patchwork by repeating this process for every vertex of Δ. □

Pseudoquads that are not paramonotone contribute to the non-monotonicity of Γ, so, as in [91], the total number and size of these pseudoquads is bounded by the measure of Γ. In §9, we will use an argument based on the Vitali covering lemma to prove that rectilinear foliated patchworks constructed using Lemma 7.3 satisfy a weighted Carleson condition, as stated in the following proposition.

PROPOSITION 7.4. *Let $r>1$ and $0<\mu\leqslant\frac{1}{32r^2}$. Let $\eta, R>0$ and let $0<\lambda<1$. Let Γ be an intrinsic λ-Lipschitz graph, let Δ be a μ-rectilinear foliated patchwork for Γ, and suppose that, for all $v\in\mathcal{V}(\Delta)$, the pseudoquad Q_v is horizontally cut if and only if Γ is (η, R)-paramonotone on rQ_v. Let $W\colon 2^{\mathcal{V}(\Delta)}\to[0,\infty]$ be as in (4.10). Then, for any $v\in\mathcal{V}(\Delta)$,*

$$W(\{w\in\mathcal{V}_{\mathsf{V}}(\Delta):w\leqslant v\})\lesssim_{\eta,r,R,\lambda}|Q_v|. \tag{7.4}$$

With these tools at hand, Theorem 5.2 follows directly.

Proof of Theorem 5.2 assuming Propositions 7.2 and 7.4. Let r be as in Proposition 7.2 and write $\mu_0=1/32r^2$. Fix $0<\mu\leqslant\mu_0$ and $\sigma>0$, and let η and R be as in Lemma 7.3. Since Γ is an intrinsic λ-Lipschitz graph, Lemma 7.3 produces a μ-rectilinear foliated patchwork Δ rooted at Q with a set of σ-approximating planes. By Proposition 7.4, this patchwork is weighted Carleson with a constant depending on η, r, R, σ, and λ. Since $r>1$ is universal and η and R depend only on μ and σ, we obtain Theorem 5.2 by using Lemma 5.5 to increase $\mu_0=1/32r^2$ to $\mu_0=\frac{1}{32}$. $\square$

Observe in passing that since in the above proof the patchwork that established Theorem 5.2 was obtained from Proposition 7.2, we actually derived the following more nuanced formulation of Theorem 5.2; it is worthwhile to state it explicitly here because this is how it will be used in forthcoming work of the second named author.

THEOREM 7.5. *For every $0<\lambda<1$ there is a function $D_\lambda:\mathbb{R}^+\times\mathbb{R}^+\to\mathbb{R}^+$, and for every $0<\mu\leqslant\frac{1}{32}$ and $\sigma>0$ there are $\eta=\eta(\mu,\sigma)$, $R=R(\mu,\sigma)>0$ with the following properties. Suppose that $\Gamma\subseteq\mathbb{H}$ is an intrinsic λ-Lipschitz graph over V_0 and $Q\subseteq V_0$ is a μ-rectilinear pseudoquad for Γ. Then, there is a μ-rectilinear foliated patchwork Δ for Q such that Δ is $D_\lambda(\mu,\sigma)$-weighted Carleson and has a set of σ-approximating planes. Moreover, for all vertices $v\in\mathcal{V}(\Delta)$, the associated pseudoquad Q_v is horizontally cut if and only if Γ is (η,R)-paramonotone on rQ, where $r>1$ is the universal constant in Proposition 7.2.*

Remark 7.6. While the results in this paper rely only on approximating a intrinsic Lipschitz graph by vertical planes to bound its vertical perimeter, Theorem 7.5 allows one to glue vertical planes together to approximate an intrinsic Lipschitz graph by ruled surfaces. Indeed, with notation as in Theorem 7.5, let $F\subseteq\mathcal{V}(\Delta)$ be a finite coherent subset such that every vertex in F is horizontally cut. Let v_F be the maximal element of F and let $m(F)$ be the set of minimal elements of F. Then, $\{Q_v:v\in m(F)\}$ is a partition of Q_{v_F} into a stack of pseudoquads $Q_1,...,Q_k$ that are vertically adjacent. The characteristic curves bounding these pseudoquads can be approximated by parabolas, denoted $h_0,...,h_k$, and the μ-rectilinearity of Δ implies that these parabolas do not intersect inside Q_{v_F}; see the proof of Lemma 9.5. We can then construct a foliation of Q_{v_F} by parabolas by linearly interpolating between the h_i's. Since any parabola is the projection of a horizontal line to V_0, this foliation is the set of characteristic curves of a ruled surface $\Sigma\subseteq\mathbb{H}$. By passing to a limit, one can construct a ruled surface corresponding to any coherent subset of horizontally cut vertices. This procedure is roughly analogous to the method used in [26] to approximate stopping-time regions in uniformly rectifiable sets in $\mathbb{R}^n$ by Lipschitz graphs. In our setting, we can use linear interpolation instead of using a partition of unity as in [26] or [91].

By Proposition 7.2, Σ approximates Γ and the characteristic curves of Σ approximate the characteristic curves of Σ inside Q_{v_F} (with accuracy depending on the heights of the Q_i's). In fact, if $v\in\mathcal{V}_{\mathsf{H}}(\Delta)$ is a vertex such that every descendant of v is horizontally cut (i.e., $\mathcal{D}(v)\subseteq\mathcal{V}_{\mathsf{H}}(\Delta)$), then Σ coincides with Γ over Q_v. We omit the details of these approximations because they are not needed in the current work, but complete details will be given in forthcoming work of the second named author where they will be used to analyze intrinsic Lipschitz functions.

We will prove Propositions 7.2 and 7.4 in the following sections. Specifically, in §8, we will define extended non-monotonicity and extended parametric normalized non-monotonicity and prove some of their basic properties. In §9, we will prove that Proposition 7.2 implies Proposition 7.4. Finally, in §§10–12, we will prove Proposition 7.2.

8. Extended non-monotonicity

8.1. Extended non-monotonicity in $\mathbb{R}$

In this section, we define the *extended non-monotonicity* and *extended parameterized non-monotonicity* of a set $E\subseteq\mathbb{H}$. Like the quantitative non-monotonicity that was defined in [23] and the horizontal width that was defined in [33], these measure how horizontal lines intersect ∂E.

We first define these quantities on subsets of lines, then define them on subsets of $\mathbb{H}$ by integrating over the space of horizontal lines. Let $\mathcal{L}$ be the space of horizontal lines in $\mathbb{H}$. Let $\mathcal{N}$ be the Haar measure on $\mathcal{L}$, normalized so that the measure of the set of lines that intersect the ball of radius r is equal to r^3.

Recall that a measurable subset $S\subseteq\mathbb{R}$ is *monotone* [21] if its indicator function is a monotone function (i.e., S is equal to either $\varnothing$, $\mathbb{R}$, or some ray). For a measurable set $U\subseteq\mathbb{R}$, we define the *non-monotonicity* of S on U by

$$\mathrm{NM}_S(U)\stackrel{\mathrm{def}}{=}\inf\{\mathcal{H}^1(U\cap(M\triangle S)):M \text{ is monotone}\},$$

where, as usual, $M\triangle S=(M\setminus S)\cup(S\setminus M)$ is the symmetric difference of M and S.

For $S\subseteq\mathbb{R}$, we say that S has *finite perimeter* if $\partial_{\mathcal{H}^1}S$ is a finite set, where we recall the notation (2.2) for measure theoretical boundary, which in the present setting becomes

$$\partial_{\mathcal{H}^1}S\stackrel{\mathrm{def}}{=}\{x\in\mathbb{R}:0<\mathcal{H}^1((x-\varepsilon,x+\varepsilon)\cap S)<2\varepsilon \text{ for all } \varepsilon>0\}.$$

If $S\subseteq\mathbb{R}$ is a set of finite perimeter, then there is a unique collection of disjoint closed intervals of positive length $\mathcal{I}(S)=\{I_1(S),I_2(S),\dots\}$ such that $S\triangle\bigcup\mathcal{I}(S)$ has measure zero. For any $R>0$, we define as follows a point measure $\omega_{S,R}$ supported on the boundaries of the intervals in $\mathcal{I}(S)$ of length at most R:

$$\omega_{S,R} \stackrel{\mathrm{def}}{=} \sum_{\substack{I\in\mathcal{I}(S)\\ \mathcal{H}^1(I)\leqslant R}} \mathcal{H}^1(I)\cdot(\delta_{\min I}+\delta_{\max I}).$$

Let

$$\widehat{\omega}_{S,R}=\tfrac{1}{2}(\omega_{S,R}+\omega_{\mathbb{R}\setminus S,R}).$$

These measures are inspired by analogous measures $\{\widehat{w}_i\}_{i\in\mathbb{Z}}$ used in [23]. It was shown in [23] that, if $\delta>0$ is sufficiently small, then the non-monotonicity of S at scale δ^i is bounded in terms of a measure $\widehat{w}_i$ that counts the set of endpoints of intervals in S or $\mathbb{R}\setminus S$ of length between δ^i and δ^{i+1}. The main difference between $\widehat{w}_i$ and $\widehat{\omega}_{S,\delta^i}$ is that $\widehat{w}_i$ ignores intervals of length less than δ^{i+1}, but $\widehat{\omega}_{S,\delta^i}$ weights them by their lengths.

For $U\subseteq\mathbb{R}$, we call $\widehat{\omega}_{S,R}(U)$ the *R-extended non-monotonicity* of S on U. (We will typically use this notation when $R>\operatorname{diam} U$.) We use the term "extended" here because it depends not only on $S\cap U$, but also on the behavior of S outside U. For example, let $U=[a,b]$ and suppose that $S\subseteq\mathbb{R}$ is a set with locally finite perimeter. If $\widehat{\omega}_{S,R}(U)=0$ for all $R>0$, then there can be no finite-length interval in $\mathcal{I}(S)$ or $\mathcal{I}(\mathbb{R}\setminus S)$ with a boundary point in U. That is, $U\cap\partial_{\mathcal{H}^1}S$ is empty or, up to a measure-zero set, $S=[c,\infty)$ or $S=(-\infty,c]$ for some $c\in[a,b]$. Similarly, when $S=[a,b]$ and $R>b-a$, if $\widehat{\omega}_{S,R}(U)$ is much smaller than $b-a$, then either $U\cap\partial_{\mathcal{H}^1}S$ is almost empty, or U is almost monotone on an R-neighborhood of S. This follows from the following two lemmas. The first lemma is based on the bounds in [23, Proposition 4.25] and [33, Lemma 3.4].

LEMMA 8.1. *Let $a,b\in\mathbb{R}$, let $U\subseteq[a,b]$, and let $R\geqslant b-a$. For any finite-perimeter set $S\subseteq\mathbb{R}$,*

$$\mathrm{NM}_S(U)\leqslant \operatorname{diam}((a,b)\cap\partial_{\mathcal{H}^1}S)\leqslant\widehat{\omega}_{S,R}((a,b)).$$

Proof. Let $\delta=\operatorname{diam}((a,b)\cap\partial_{\mathcal{H}^1}S)$. Consider the following set of closed intervals:

$$\mathbf{J}\stackrel{\mathrm{def}}{=}\{I\in\mathcal{I}(S)\cup\mathcal{I}(\mathbb{R}\setminus S): I\cap(a,b)\neq\varnothing\}.$$

This set is finite, so we may label its elements $J_1,...,J_n$ in increasing order. After changing S on a measure-zero subset, the interiors of the J_i's are alternately contained in S and disjoint from S. If $n=1$, then $\mathrm{NM}_S(U)=0$ and $\delta=0$, so we suppose that $n\geqslant 2$. Then,

$$\delta=\min(J_n)-\max(J_1)=\sum_{i=2}^{n-1}\mathcal{H}^1(J_i)$$

and

$$\widehat{\omega}_{S,R}((a,b)) \geqslant 2\sum_{m=2}^{n-1}\mathcal{H}^1(J_m) \geqslant \delta.$$

Regardless of whether J_1 and J_n are in or out of S, there is a monotone subset $M \subseteq \mathbb{R}$ such that $\mathbf{1}_M$ agrees with $\mathbf{1}_S$ on J_1 and J_n. Then,

$$\mathrm{NM}_S(U) \leqslant \mathcal{H}^1(U \cap (S \triangle M)) \leqslant \mathcal{H}^1([a,b] \setminus (J_1 \cup J_n)) = \delta. \qquad \square$$

A similar reasoning gives the following lower bound. Recall that $\mathrm{supp}_{\mathcal{H}^1}$ and $\mathrm{int}_{\mathcal{H}^1}$ denote measure-theoretic support and interior, see (2.1)–(2.3).

Lemma 8.2. *Fix $R, a, b \in \mathbb{R}$ with $a<b$ and $R \geqslant b-a$. Let $S \subseteq \mathbb{R}$ have locally finite perimeter such that $a, b \in \mathrm{supp}_{\mathcal{H}^1}(\mathbb{R}\setminus S)$. For any closed interval $I \subseteq [a,b]$, either*

$$I \subseteq \mathrm{int}_{\mathcal{H}^1} S \quad \text{or} \quad \widehat{\omega}_{S,R}(I) \geqslant \tfrac{1}{2}\mathcal{H}^1(S \cap I).$$

Proof. Suppose $I \not\subseteq \mathrm{int}_{\mathcal{H}^1} S$. Let $I_1, ..., I_n$ be the intervals in $\mathcal{I}(S)$ that intersect I. By assumption, each of the intervals $I_1, ..., I_n$ has at least one endpoint in I. Furthermore, since $a, b \in \mathrm{supp}_{\mathcal{H}^1}(\mathbb{R}\setminus S)$, we have $I_j \subseteq [a,b]$ for all $j \in \{1, ..., n\}$. In particular, $\max_{j\in\{1,...,n\}} \ell(I_j) \leqslant R$. Up to a null set, we have $S \cap I \subseteq \bigcup_{j=1}^n I_j$, so

$$\widehat{\omega}_{S,R}(I) \geqslant \sum_{j=1}^{n} \frac{\mathcal{H}^1(I_j)}{2} \geqslant \frac{\mathcal{H}^1(S \cap I)}{2}. \qquad \square$$

These lemmas yield the following description of sets with small extended non-monotonicity, which states that points in their measure theoretic boundary must be either very close to each other, or very far from each other.

Proposition 8.3. *Let $S \subseteq \mathbb{R}$ be a set with locally finite perimeter and fix $c, d \in \mathbb{R}$ with $c<d$. Let $R \geqslant d-c$ and suppose that $0<\varepsilon<\frac{1}{8}(d-c)$ and $\widehat{\omega}_{S,R}((c,d))<\varepsilon$. Then,*

$$\mathrm{diam}((t-R, t+R) \cap \partial_{\mathcal{H}^1} S) < \varepsilon \quad \textit{for all } t \in [c+4\varepsilon, d-4\varepsilon] \cap \partial_{\mathcal{H}^1} S. \tag{8.1}$$

Proof. Fix $t \in [c+4\varepsilon, d-4\varepsilon] \cap \partial_{\mathcal{H}^1} S$. We will prove that this implies that

$$(t-R, t+R) \cap \partial_{\mathcal{H}^1} S \subseteq (c,d). \tag{8.2}$$

Equation (8.1) is a consequence of the inclusion (8.2), since, by Lemma 8.1,

$$\mathrm{diam}((t-R, t+R) \cap \partial_{\mathcal{H}^1} S) \overset{(8.2)}{\leqslant} \mathrm{diam}((c,d) \cap \partial_{\mathcal{H}^1} S) \leqslant \widehat{\omega}_{S,R}((c,d)) < \varepsilon.$$

Suppose by way of contradiction that (8.2) fails. So, there is

$$u\in(t-R,t+R)\cap\partial_{\mathcal{H}^1}S$$

with $u\geqslant d$ or $u\leqslant c$. We will treat only the case $u\geqslant d$, since the case $u\leqslant c$ is analogous. Lemma 8.2 applied with $[a,b]=[t,u]$ and $I=[t,d]$, gives

$$\tfrac{1}{2}\mathcal{H}^1(S\cap[t,d])\leqslant\widehat{\omega}_{S,R}([t,d])<\varepsilon.$$

If we replace S by $\mathbb{R}\setminus S$, the Lemma 8.2 gives

$$\frac{\mathcal{H}^1([t,d]\setminus S)}{2}\leqslant\widehat{\omega}_{S,R}([t,d])<\varepsilon.$$

So $d-t<4\varepsilon$, which contradicts the choice of t. □

Remark 8.4. Despite the name "extended non-monotonicity", there is no direct comparison between the extended non-monotonicity of S on U and the non-monotonicity of S on a neighborhood of U. For example, if $R>0$, $0<\varepsilon<1$, and $S=[-\varepsilon,\varepsilon]\cup[R,\infty)$, then $\mathrm{NM}_S(\mathbb{R})=4\varepsilon$, but $\widehat{\omega}_{S,R}$ is the point measure

$$\widehat{\omega}_{S,R}=\varepsilon\delta_{-\varepsilon}+\big(\tfrac{1}{2}\varepsilon+\tfrac{1}{2}R\big)\delta_\varepsilon+\tfrac{1}{2}R\delta_R,$$

so $\widehat{\omega}_{S,R}([-1,1])$ is large, despite S having small non-monotonicity. Conversely, for any $T\subseteq\mathbb{R}$ that contains $[-1,1]$, the boundary $\partial_{\mathcal{H}^1}T$ is disjoint from $(-1,1)$, so

$$\widehat{\omega}_{T,R}((-1,1))=0,$$

regardless of the behavior of T on the rest of $\mathbb{R}$.

8.2. Extended non-monotonicity in $\mathbb{H}$

We have defined NM and $\widehat{\omega}$ for subsets of $\mathbb{R}$, but the same definitions are valid for subsets of any line $L\in\mathcal{L}$. This lets us define the non-monotonicity of a subset of $\mathbb{H}$ by integrating over horizontal lines.

When $U,E\subseteq\mathbb{H}$ are measurable sets, we define the non-monotonicity of E on U by

$$\mathrm{NM}_E(U)\stackrel{\mathrm{def}}{=}\int_{\mathcal{L}}\mathrm{NM}_{E\cap L}(U\cap L)\,d\mathcal{N}(L).$$

(Note that this definition differs from the definition in [23]. Specifically, in [23], this was only defined in the case that $U=B_r(x)$ for some $r\in(0,\infty)$ and $x\in\mathbb{H}$, and was normalized by a factor of r^{-3} to make it scale-invariant.)

Definition 8.5. Fix $R>0$. Let $E\subseteq\mathbb{H}$ be a set with finite perimeter. By the kinematic formula (§2.5), for almost every $L\in\mathcal{L}$, the intersection $E\cap L$ is a set with finite perimeter, and we define, for $U\subseteq\mathbb{H}$,

$$\widehat{\omega}_{E,R}(U,L)\stackrel{\text{def}}{=}\widehat{\omega}_{E\cap L,R}(U\cap L). \tag{8.3}$$

We then define a measure $\mathrm{ENM}_{E,R}$ on $\mathbb{H}$ by setting

$$\mathrm{ENM}_{E,R}(U)\stackrel{\text{def}}{=}\int_{\mathcal{L}}\widehat{\omega}_{E,R}(U,L)\,d\mathcal{N}(L).$$

We call $\mathrm{ENM}_{E,R}(U)$ the *R-extended non-monotonicity* of E on U, and for $\nu>0$ we say that E is *(ν,R)-extended monotone* on U if $\mathrm{ENM}_{E,R}(U)\leqslant\nu$. Like $\widehat{\omega}_{S,R}(\cdot)$, $\mathrm{ENM}_{E,R}(U)$ depends on the behavior of E in an R-neighborhood of U. If $R\leqslant R'$, then

$$\mathrm{ENM}_{E,R}\leqslant\mathrm{ENM}_{E,R'}\,.$$

When we say that a subset $U\subseteq\mathbb{H}$ is *convex*, we will always mean that it is convex as a subset of the vector space $\mathbb{R}^3$. For every $g\in\mathbb{H}$, the map $v\mapsto gv$ is an affine map from $\mathbb{H}$ to itself, so convexity is preserved by left multiplication.

LEMMA 8.6. *Let $U\subseteq\mathbb{H}$ be a measurable bounded set and let $K\subseteq U$ be convex. Let $E\subseteq\mathbb{H}$ be a finite-perimeter set. Then, for every $R>\operatorname{diam}U$, we have*

$$\mathrm{NM}_E(K)\leqslant\mathrm{ENM}_{E,R}(U).$$

Proof. Let $L\in\mathcal{L}$ be a horizontal line. By convexity, the intersection $I=L\cap K$ is an interval and $\ell(I)\leqslant\operatorname{diam}U$. By Lemma 8.1,

$$\mathrm{NM}_{E\cap L}(I)\leqslant\widehat{\omega}_{E\cap L,R}(I)\leqslant\widehat{\omega}_{E,R}(U,L).$$

Integrating both sides of this inequality with respect to $\mathcal{N}$ yields the desired bound. □

We will also define a parametric version of extended non-monotonicity that is better adapted to intrinsic Lipschitz graphs. This is based on a different measure on the space of horizontal lines, denoted $\mathcal{N}_P$, which we next describe.

Let $W_0=\{x=0\}$ be the yz-plane and let $\mathcal{L}_P$ be the set of horizontal lines that are not parallel to W_0. Each $L\in\mathcal{L}_P$ intersects W_0 in a single point $w(L)$, called the *intercept* of L, and has a unique *slope* $m(L)\in\mathbb{R}$ such that $L=w(L)\cdot\langle X+m(L)Y\rangle$.

The map $(m,w)\colon\mathcal{L}_P\to\mathbb{R}\times W_0$ is a bijection, and we define $\mathcal{N}_P$ to be the pullback of the Lebesgue measure on $\mathbb{R}\times W_0$ under this bijection. This measure is preserved by shear maps and translations. If $a,b>0$ and if $L_{(0,y,z),m}$ is the line with slope m and intercept $(0,y,z)$, then $s_{a,b}(L_{(0,y,z),m})=L_{(0,by,abz),mb/a}$, so, for any measurable set $A\subseteq\mathcal{L}_P$,

$$\mathcal{N}_P(s_{a,b}(A))=b^3\mathcal{N}_P(A). \tag{8.4}$$

Let $E\subseteq\mathbb{H}$. For any $R>0$, any $U\subseteq V_0$, and any $L\in\mathcal{L}_P$, we define

$$\widehat{\omega}^P_{E,R}(U,L)\stackrel{\mathrm{def}}{=}\widehat{\omega}_{x(E\cap L),R}(x(\Pi^{-1}(U)\cap L)). \tag{8.5}$$

This is similar to $\widehat{\omega}_{E,R}(\Pi^{-1}(U),L)$ in (8.3), but the projection to the x-coordinate that appears in (8.5) changes the measures and lengths involved by a constant factor.

When E is a finite-perimeter subset of $\mathbb{H}$, we define a measure $\Omega^P_{E,R}$ on V_0, by setting for any measurable subset $U\subseteq V_0$,

$$\Omega^P_{E,R}(U)\stackrel{\mathrm{def}}{=}\frac{1}{R}\int_{\mathcal{L}_P}\widehat{\omega}^P_{E,R}(U,L)\,d\mathcal{N}_P(L). \tag{8.6}$$

We call $\Omega^P_{E,R}(U)$ the *R-extended parametric normalized non-monotonicity* of E on U. Now, note that the definition (8.6) includes an R^{-1} factor that does not appear in Definition 8.5; we will see that this normalization allows for the kinematic formula (1.33) to hold.

In general, the measure $\Omega^P_{E,R}$ is not necessarily locally finite. Indeed, if $B\subseteq\mathbb{H}$ is a ball, then the set of lines that pass through B has infinite $\mathcal{N}_P$-measure. But when Γ is an intrinsic λ-Lipschitz graph, any line with sufficiently large slope intersects Γ exactly once. If $E=\Gamma^+$ and $L\in\mathcal{L}_P$ is a line such that $L\cap E$ is non-monotone, then L has bounded slope; it follows that $\Omega^P_{\Gamma^+,R}(K)$ is finite for any compact $K\subseteq V_0$. Furthermore, $\Omega^P_{\Gamma^+,R}$ is bounded below by $\mathrm{ENM}_{\Gamma^+,R}$.

LEMMA 8.7. *Let $R>0$ and $x\in\mathbb{H}$. Suppose that E is a finite-perimeter subset of $\mathbb{H}$, and let $U\subseteq\mathbb{H}$ be measurable. Then,*

$$\mathrm{ENM}_{E,R}(U)\lesssim R\Omega^P_{E,R}(\Pi(U)).$$

Proof. Let $L\in\mathcal{L}_P$ and let $m=m(L)$ be the slope of L, so that the restriction $x|_L$ shrinks lengths by a factor of $\phi(m)=\sqrt{1+m^2}$. Then,

$$\widehat{\omega}^P_{E,R}(\Pi(U),L)=\frac{\widehat{\omega}_{E,\phi(m)\cdot R}(\Pi^{-1}(\Pi(U)),L)}{\phi(m)}\geqslant\frac{\widehat{\omega}_{E,R}(U,L)}{\phi(m)}.$$

For $w\in W_0$ and $m\in\mathbb{R}$, let $L_{w,m}$ be the line $L_{w,m}=w\cdot\langle X+mY\rangle\in\mathcal{L}_P$. Then, it follows that

$$R\Omega^P_{E,R}(\Pi(U))=\int_{W_0}\int_{\mathbb{R}}\widehat{\omega}^P_{E,R}(\Pi(U),L_{w,m})\,dm\,dw\geqslant\int_{W_0}\int_{\mathbb{R}}\frac{\widehat{\omega}_{E,R}(U,L_{w,m})}{\sqrt{1+m^2}}\,dm\,dw.$$

For $\theta\in\mathbb{R}$, let $R_\theta:\mathbb{H}\to\mathbb{H}$ be the rotation by angle θ around the z-axis. Since $\mathcal{N}$ is invariant under translations and rotations, there is $c>0$ such that, for any measurable $f:\mathcal{L}\to\mathbb{R}$,

$$\int_{\mathcal{L}}f(M)\,d\mathcal{N}(M)=c\int_{W_0}\int_{-\pi/2}^{\pi/2}f(R_\theta(L_{g,0}))\,d\theta\,dg.$$

Any line in $\mathcal{L}_P$ can be written as $R_\theta(L_{g,0})$ for some $\theta\in\mathbb{R}$ and $g\in W_0$. Specifically, for $w\in W_0$ and $m\in\mathbb{R}$, let

$$\theta(m)=\arctan m$$

and let $g_m(w)$ be the W_0-intercept of $R_{-\theta(m)}(L_{w,m}))$, so that $L_{w,m}=R_{\theta(m)}(L_{g_m(w),0})$. Writing g_m in coordinates as $g_m=(0,b_m,c_m)$, its Jacobian is

$$J_{g_m}(y,z)=\det\begin{pmatrix}\dfrac{db_m}{dy} & \dfrac{db_m}{dz}\\ \dfrac{dc_m}{dy} & \dfrac{dc_m}{dz}\end{pmatrix}=\det\begin{pmatrix}\cos(\arctan m) & 0\\ \dfrac{dc_m}{dy} & 1\end{pmatrix}=\frac{1}{\sqrt{1+m^2}}.$$

Consequently,

$$\begin{aligned}\int_{W_0}\int_{-\pi/2}^{\pi/2} f(R_\theta(L_{g,0}))\,d\theta\,dg&=\int_{W_0}\int_{\mathbb{R}} f(L_{w,m})\frac{d\theta}{dm}J_{g_m}(w)\,dm\,dw\\&=\int_{W_0}\int_{\mathbb{R}}\frac{f(L_{w,m})}{(1+m^2)^{3/2}}\,dm\,dw.\end{aligned}$$

Thus,

$$\begin{aligned}\mathrm{ENM}_{E,R}(U)=\int_{\mathcal{L}}\widehat{\omega}_{E,R}(U,L)\,d\mathcal{N}(L)&=c\int_{W_0}\int_{\mathbb{R}}\frac{\widehat{\omega}_{E,R}(U,L_{w,m})}{(1+m^2)^{3/2}}\,dm\,dw\\&\leqslant cR\Omega^P_{E,R}(\Pi(U)),\end{aligned}$$

as desired. $\square$

One advantage of Ω^P over ENM is that Ω^P scales nicely under automorphisms.

Lemma 8.8. *Fix $a,b\in\mathbb{R}\setminus\{0\}$ and let*

$$g=q\circ\rho_h\circ s_{a,b}\colon\mathbb{H}\longrightarrow\mathbb{H}$$

be a composition of a shear map q, a left-translation by $h\in\mathbb{H}$, and a stretch map $s_{a,b}$. Let $\hat{g}\colon V_0\to V_0$ be the map induced on V_0, i.e., $\hat{g}(x)=\Pi(g(x))$ for all $x\in V_0$. Let $E\subseteq\mathbb{H}$ be a set with finite perimeter. For any measurable $U\subseteq V_0$ and any $R>0$, if $\Omega^P_{E,R}(U)$ is finite, then

$$\Omega^P_{g(E),|a|R}(\hat{g}(U))=|b|^3\Omega^P_{E,R}(U) \tag{8.7}$$

and

$$\frac{\Omega^P_{g(E),|a|R}(\hat{g}(U))}{|\hat{g}(U)|}=\frac{b^2}{a^2}\cdot\frac{\Omega^P_{E,R}(U)}{|U|}. \tag{8.8}$$

In particular, if g is a composition of a scaling, shear, and translation, i.e., when $a=b$ above, then g preserves the density of $\Omega^P_{E,R}$.

Proof. The identity (8.7) is verified by computing as follows, using (8.4):

$$\begin{aligned}\Omega^P_{g(E),|a|R}(\hat g(U)) &= \frac{1}{|a|R}\int_{\mathcal{L}_P} \widehat{\omega}^P_{g(E),|a|R}(\hat g(U),L)\,d\mathcal{N}_P(L)\\ &= \frac{|b|^3}{|a|R}\int_{\mathcal{L}_P} \widehat{\omega}^P_{g(E),|a|R}(\hat g(U),g(L))\,d\mathcal{N}_P(L)\\ &= \frac{|b|^3}{|a|R}\int_{\mathcal{L}_P} |a|\widehat{\omega}^P_{E,R}(U,L)\,d\mathcal{N}_P(L)\\ &= |b|^3\Omega^P_{E,R}(U),\end{aligned}$$

By Lemma 2.8, we have $|\hat g(U)|=a^2|b|\cdot|U|$, which implies (8.8). □

Suppose that Q is a pseudoquad for an intrinsic Lipschitz graph $\Gamma\subseteq\mathbb{H}$, and that g is as in Lemma 8.8. If Γ is (η,R)-paramonotone on rQ as in Definition 7.1, then the density of $\Omega^P_{\Gamma^+,R\delta_x(Q)}$ is bounded as follows:

$$\frac{\Omega^P_{\Gamma^+,R\delta_x(Q)}(rQ)}{|Q|}\leqslant\frac{\eta}{\alpha(Q)^4}.$$

Let $\widehat{Q}=\hat g(Q)$ and $\widehat{\Gamma}=g(\Gamma)$. Then (8.8) and Lemma 4.2 imply that

$$\frac{\Omega^P_{\widehat{\Gamma}^+,R\delta_x(\widehat{Q})}(r\widehat{Q})}{|\widehat{Q}|}\leqslant\frac{\eta b^2}{a^2\alpha(Q)^4}=\frac{\eta}{\alpha(\widehat{Q})^4},$$

so $\widehat{\Gamma}$ is (η,R)-paramonotone on $r\widehat{Q}$ if and only if Γ is (η,R)-paramonotone on rQ.

In particular, it follows from Lemma 8.7 that if Γ is (η,R)-paramonotone on rQ, then

$$\mathrm{ENM}_{\Gamma^+,R\delta_x(Q)}(\Pi^{-1}(rQ))\lesssim R\delta_x(Q)\Omega^P_{\Gamma^+,R\delta_x(Q)}(rQ)\leqslant\frac{\delta_x(Q)|Q|}{\alpha(Q)^4}\eta R\asymp_Q\eta R. \tag{8.9}$$

9. The kinematic formula and the proof of Proposition 7.4

In this section, we prove Proposition 7.4 using two lemmas. The first bounds the total weight of the vertically cut descendants of a vertex $v\in\mathcal{V}(\Delta)$ in terms of Ω^P.

LEMMA 9.1. *Let r, η, R, λ, Γ, and Δ be as in Proposition 7.4. Then, for any $v\in\mathcal{V}(\Delta)$,*

$$W(\{w\in\mathcal{V}_{\mathsf{V}}(\Delta):w\leqslant v\})\lesssim_{\eta,r,R}\sum_{i=0}^{\infty}\Omega^P_{\Gamma^+,2^{-i}R\delta_x(Q_v)}(rQ_v). \tag{9.1}$$

The second is a kinematic formula bounding Ω^P in terms of Lebesgue measure on V_0.

LEMMA 9.2. *Let $0<\lambda<1$ and let Γ be an intrinsic λ-Lipschitz graph. For any measurable set $U\subseteq V_0$,*

$$\sum_{i\in\mathbb{Z}} \Omega^P_{\Gamma^+,2^{-i}}(U) \lesssim_\lambda |U|. \tag{9.2}$$

Proposition 7.4 follows from Lemmas 9.1 and 9.2.

Proof of Proposition 7.4 assuming Lemmas 9.1 and 9.2. Let us fix $v\in\mathcal{V}(\Delta)$ and denote $\delta=\delta_x(Q_v)$. Due to Lemma 9.1,

$$W(\{w\in\mathcal{V}_{\mathsf{V}}(\Delta): w\leqslant v\}) \lesssim_{\eta,r,R} \sum_{i=0}^{\infty} \Omega^P_{\Gamma^+,2^{-i}R\delta}(rQ_v).$$

Let k be the integer such that $2^{k-1}\leqslant R\delta<2^k$. Then,

$$\sum_{i=0}^{\infty} \Omega^P_{\Gamma^+,2^{-i}R\delta}(rQ_v) \leqslant \sum_{i=0}^{\infty} 2\Omega^P_{\Gamma^+,2^{-i+k}}(rQ_v) \overset{(9.2)}{\lesssim_\lambda} |rQ_v|.$$

Thus,

$$W(\{w\in\mathcal{V}_{\mathsf{V}}(\Delta): w\leqslant v\}) \lesssim_{\eta,r,\lambda,R} |Q_v|. \qquad \square$$

We first establish Lemma 9.1, which we prove using an argument based on the Vitali covering lemma. The first step is to construct partitions of Q into pseudoquads with dyadic widths. As in Lemma 6.4, we construct these partitions from coherent subtrees.

LEMMA 9.3. *Let $0<\mu\leqslant\frac{1}{32}$ and let $(\Delta,(Q_v)_{v\in\Delta})$ be a μ-rectilinear foliated patchwork for a μ-rectilinear pseudoquad Q. Fix $j\in\mathbb{N}\cup\{0\}$. For $v\in\mathcal{V}(\Delta)$, let $\mathcal{D}_{\mathsf{V}}(v)\subseteq\mathcal{V}(\Delta)$ denote the set of vertically cut descendants of v, and let*

$$F_j(v) \stackrel{\text{def}}{=} \{w\in\mathcal{D}_{\mathsf{V}}(v):\delta_x(Q_w)=2^{-j}\delta_x(Q)\}.$$

Then, for any $w,w'\in F_j(v)$, if $w\neq w'$, then Q_w and $Q_{w'}$ have disjoint interiors.

Proof. Let $\mathcal{D}(v)$ be the set of descendants of v, and let

$$R_j=\{w\in\mathcal{D}(v):\delta_x(Q_w)\geqslant 2^{-j}\delta_x(Q_v)\}.$$

By Lemma 4.5, this is a coherent set and $F_j(v)=\min R_j$, so Lemma 6.3 implies that $F_j(v)$ consists of pseudoquads with disjoint interiors. $\square$

Let v_0 be the root of Δ (so $Q=Q_{v_0}$). For each $j\in\mathbb{N}\cup\{0\}$ we write $F_j=F_j(v_0)$. Denote $I=x(Q)$ and $l_j=2^{-j}\ell(I)=2^{-j}\delta_x(Q)$. Let $I_{j,1},...,I_{j,2^j}$ be the partition of I into 2^j intervals of length l_j such that, for any $v\in\mathcal{V}(\Delta)$, there are $j,m\in\mathbb{N}\cup\{0\}$ such that $x(Q_v)=I_{j,m}$. We partition F_j into columns as follows:

$$F_{j,m} \stackrel{\text{def}}{=} \{w\in F_j: x(Q_w)=I_{j,m}\} \quad \text{for all } m\in\{1,...,2^j\}. \tag{9.3}$$

Each column satisfies the following version of the Vitali covering lemma.

Lemma 9.4. *For each $j\in\mathbb{N}\cup\{0\}$ and $m\in\{1,...,2^j\}$, there is a (possibly finite) sequence of vertices $D_{j,m}=\{v_1,v_2,...\}\subseteq F_{j,m}$ such that $rQ_{v_1},rQ_{v_2},...$ are pairwise disjoint and*

$$W(D_{j,m})\asymp_r W(F_{j,m}).$$

We prove Lemma 9.4 using the following expansion property.

Lemma 9.5. *Let $r>1$ and let $0<\mu\leqslant 1/32r^2$. Let Δ be a μ-rectilinear foliated patchwork. Let $v,w\in\mathcal{V}(\Delta)$ be two vertices such that $x(Q_v)=x(Q_w)$, and suppose that $rQ_v\cap rQ_w$ is non-empty. If $\delta_z(Q_v)\geqslant\delta_z(Q_w)$, then $Q_w\subseteq 3rQ_v$, and if $\delta_z(Q_w)\geqslant\delta_z(Q_v)$, then $Q_w\subseteq 3rQ_v$.*

Proof. Write $I=[-1,1]$. By rescaling and translating, we may suppose without loss of generality that $x(Q_v)=x(Q_w)=I$. Moreover, we may suppose that Q_v is vertically below Q_w. We first construct a stack of pseudoquads of width at least 2 that connects Q_v and Q_w.

For $u\in\mathcal{V}(\Delta)$, let $A(u)=\{t\in\mathcal{V}(\Delta):t\geqslant u\}$ be the set of ancestors of u. If $u\neq v_0$, let $S(u)$ be the sibling of u and let $P(u)$ be the parent of u. Let

$$J=A(v)\cup A(w)\cup S(A(v)\cup A(w)).$$

Since $A(v)\cup A(w)$ spans a connected subtree of Δ, so does J, and J is a coherent subset of $\mathcal{V}(\Delta)$. Furthermore, J is finite, so $K=\min J$ is a partition of Q.

If $u\in K$, then u is either an ancestor of v or w, or a sibling of such an ancestor. In either case, $\delta_x(Q_u)\geqslant 2$, and the base of Q_u either contains I or its interior is disjoint from I. Let $K'=\{u\in K:I\subseteq x(Q_u)\}$. For each $u\in K'$, Q_u intersects the z-axis in an interval. We denote the elements of K' by $u_1,...,u_n$, in order of increasing z-coordinate. These pseudoquads form a stack; each pseudoquad Q_{u_i} is vertically adjacent to $Q_{u_{i+1}}$. We suppose that $u_a=v$ and $u_b=w$, with $a<b$.

Rectilinearity implies that the boundaries of the Q_{u_i}'s have similar slopes. For each $i\in\{1,...,n\}$, let g_i be the lower bound of Q_{u_i} and let g_{i+1} be its upper bound. These may be defined on different domains, but all of their domains contain I. For each $i\in\{1,...,n\}$, let R_{u_i} be the parabolic rectangle associated with Q_{u_i} and let $d_i=\delta_z(Q_{u_i})$, so that there are quadratic functions $h_i:\mathbb{R}\to\mathbb{R}$ satisfying

$$\big\|g_i-\big(h_i-\tfrac12 d_i\big)\big\|_{L_\infty(I)}\leqslant\mu d_i\quad\text{and}\quad\big\|g_{i+1}-\big(h_i+\tfrac12 d_i\big)\big\|_{L_\infty(I)}\leqslant\mu d_i.$$

Then,

$$\big\|\big(h_{i+1}-\tfrac12 d_{i+1}\big)-\big(h_i+\tfrac12 d_i\big)\big\|_{L_\infty(I)}\leqslant\mu(d_i+d_{i+1}).\tag{9.4}$$

Hence, for any $i,j\in\{1,\dots,n\}$ with $i<j$,

$$\Big\|h_j-h_i-\sum_{k=i}^{j-1}\frac{d_k+d_{k+1}}{2}\Big\|_{L_\infty(I)}\leqslant\mu\sum_{k=i}^{j-1}(d_k+d_{k+1}).$$

Since

$$h_j-h_i-\sum_{k=i}^{j-1}\frac{d_k+d_{k+1}}{2}$$

is quadratic, by Lemma 4.6 it follows that

$$\Big\|h_j-h_i-\sum_{k=i}^{j-1}\frac{d_k+d_{k+1}}{2}\Big\|_{L_\infty([-r,r])}\leqslant 4r^2\mu\sum_{k=i}^{j-1}(d_k+d_{k+1})\leqslant\sum_{k=i}^{j-1}\frac{d_i+d_{i+1}}{8}.$$

Denoting

$$D=\sum_{k=a}^{b-1}\frac{d_k+d_{k+1}}{2},$$

it follows that, for all $x\in[-r,r]$, we have

$$\tfrac{3}{4}D\leqslant h_b(x)-h_a(x)\leqslant\tfrac{5}{4}D. \tag{9.5}$$

Suppose that $\delta_z(Q_w)\leqslant\delta_z(Q_v)$. For each $i\in\{1,\dots,n\}$, the definition (4.1) of rQ states

$$rQ_{u_i}=\big\{(x,z)\in V_0: x\in[-r,r] \text{ and } |z-h_i(x)|\leqslant\tfrac{1}{2}r^2d_i\big\}.$$

Since $\delta_z(Q_w)\leqslant\delta_z(Q_v)$ and rQ_v intersects rQ_w, there is $t\in[-r,r]$ such that

$$h_b(t)-h_a(t)\leqslant\tfrac{1}{2}r^2(d_b+d_a)=\tfrac{1}{2}r^2(\delta_z(Q_v)+\delta_z(Q_w))\leqslant r^2\delta_z(Q_v),$$

and thus by (9.5) we have $D\leqslant\frac{4}{3}r^2\delta_z(Q_v)$.

Let $(x,z)\in Q_w$. By (9.4),

$$z\in[g_b(x),g_{b+1}(x)]\subseteq\big[h_b(x)-\tfrac{3}{4}\delta_z(Q_w),h_b(x)+\tfrac{3}{4}\delta_z(Q_w)\big],$$

so

$$|h_a(x)-z|\leqslant|h_a(x)-h_b(x)|+|h_b(x)-z|\leqslant\tfrac{5}{4}D+\tfrac{3}{4}\delta_z(Q_w)\leqslant\tfrac{1}{2}(3r)^2\delta_z(Q_v),$$

where the penultimate step uses (9.5), and the final step uses the upper bound on D that we derived above and the assumption $\delta_z(Q_w)\leqslant\delta_z(Q_v)$. It follows $(x,z)\in 3rQ_v$, and thus $Q_w\subseteq 3rQ_v$. If $\delta_z(Q_v)\leqslant\delta_z(Q_w)$, then the analogous reasoning shows $Q_v\subseteq 3rQ_w$. □

Proof of Lemma 9.4. Similarly to the proof of the Vitali covering lemma, we define inductively a sequence $S_0, S_1 ...$, of subsets of $F_{j,m}$ as follows. Let $S_0=\varnothing$. For each $i\in\mathbb{N}$, let v_i be an element of $F_{j,m}\setminus\bigcup_{k=0}^{i-1} S_i$ that maximizes $\delta_z(Q_{v_i})$. Define

$$S_i=\left\{w\in F_{j,m}\setminus\bigcup_{k=0}^{i-1} S_k : rQ_{v_i}\cap rQ_w\neq\varnothing\right\}.$$

If $\bigcup_{k=1}^{i} S_k=F_{j,m}$, then we stop. By construction, $rQ_{v_1}, rQ_{v_2}, ...$ are disjoint. We will show that the set $D_{j,m}=\{v_1, v_2, ...\}\subseteq F_{j,m}$ satisfies the desired properties.

We first claim that $F_{j,m}=S_1\cup S_2\cup ...$, where this holds by construction if there are only finitely many v_i's. So, suppose that there are infinitely many v_i's and let $w\in F_{j,m}$. There are only finitely many elements of $F_{j,m}$ with height greater than $\delta_z(Q_w)$, so there is $i\in\mathbb{N}$ such that $\delta_z(Q_{v_i})<\delta_z(Q_w)$. By the maximality of $\delta_z(Q_{v_i})$, this implies that $w\in S_1\cup...\cup S_{i-1}$.

We next show that

$$W(D_{j,m})\asymp_r W(F_{j,m}).$$

As $D_{j,m}\subseteq F_{j,m}$, we have $W(D_{j,m})\leqslant W(F_{j,m})$. Conversely, if $w\in S_i$, then rQ_w intersects rQ_{v_i} and $\delta_z(Q_w)\leqslant\delta_z(Q_{v_i})$, so Lemma 9.5 implies that $Q_w\subseteq 3rQ_{v_i}$. Since the elements of $F_{j,m}$ are pairwise disjoint (Lemma 9.3) pseudoquads of the same width, we have $\alpha(Q_w)\geqslant\alpha(Q_{v_i})$ and

$$\begin{aligned}W(S_i)&=\sum_{w\in S_i}\alpha(Q_w)^{-4}|Q_w|\leqslant\alpha(Q_{v_i})^{-4}\sum_{w\in S_i}|Q_w|\\&=\alpha(Q_{v_i})^{-4}\left|\bigcup_{w\in S_i}Q_w\right|\leqslant\alpha(Q_{v_i})^{-4}|3rQ_{v_i}|\asymp r^3W(\{v_i\}).\end{aligned}$$

By summing this bound over j, we conclude

$$W(F_{j,m})=W(S_1)+W(S_2)+...\lesssim r^3\big(W(\{v_1\})+W(\{v_2\})+...\big)=r^3W(D_{j,m}).\qquad\square$$

We are now ready to prove Lemma 9.1.

Proof of Lemma 9.1. It suffices to treat the case where v is the root of Δ, so $Q_v=Q$. Fix $j\in\mathbb{N}\cup\{0\}$ and $m\in\{1,...,2^j\}$. Let $F_{j,m}$ and $D_{j,m}$ be as in Lemma 9.4.

Since, by definition, $F_{j,m}$ consists only of vertices that are vertically cut, by hypothesis, Γ is not (η,R)-paramonotone on Q_w for each $w\in F_{j,m}$, i.e.,

$$\Omega^P_{\Gamma^+,Rl_j}(rQ_w)>\eta\alpha(Q_w)^{-4}|Q_w|=\eta W(\{w\})\quad\text{for all } w\in F_{j,m}.$$

Let $S_m = rI_{j,m} \times \{0\} \times \mathbb{R} \subseteq V_0$. The sets $\{rQ_w\}_{w\in D_{j,m}}$ are disjoint subsets of $S_m \cap rQ$, so

$$W(F_{j,m}) \asymp_r W(D_{j,m}) \leqslant \eta^{-1} \sum_{w\in D_{j,m}} \Omega^P_{\Gamma^+, Rl_j}(rQ_w) \leqslant \eta^{-1} \Omega^P_{\Gamma^+, Rl_j}(S_m \cap rQ).$$

By summing this bound over $m\in\{1, ..., 2^j\}$, we get

$$W(F_j) \lesssim_r \sum_{m=1}^{2^j} \eta^{-1} \Omega^P_{\Gamma^+, Rl_j}(S_m \cap rQ) \lesssim_r \eta^{-1} \Omega^P_{\Gamma^+, Rl_j}(rQ), \tag{9.6}$$

where the last step holds because the scaled intervals $rI_{j,1}, ..., rI_{j,2^j}$ have bounded overlap (depending on r). By summing this bound over j, we conclude as follows:

$$W(\mathcal{V}_{\mathsf{V}}(\Delta)) = \sum_{j=0}^{\infty} W(F_j) \lesssim_r \eta^{-1} \sum_{j=0}^{\infty} \Omega^P_{\Gamma^+, R2^{-j}\delta_x(Q)}(rQ). \qquad \square$$

Next, we prove Lemma 9.2 using the following kinematic formula for intrinsic Lipschitz graphs. Recall (§2.1) that for a measurable subset $E \subseteq \mathbb{H}$, we let Per_E denote the perimeter measure of E; this measure is supported on ∂E, and when E is bounded by an intrinsic Lipschitz graph, it differs from 3-dimensional Hausdorff measure on ∂E by at most a multiplicative constant. For any horizontal line $L\in\mathcal{L}$, let $\partial_{\mathcal{H}^1|_L} E$ be the measure-theoretic boundary of E in L and let $\mathrm{Per}_{E,L}$ be the counting measure on $\partial_{\mathcal{H}^1|_L} E$.

LEMMA 9.6. *Fix $0<\lambda<1$. Let $\psi: V_0 \to \mathbb{R}$ be intrinsic λ-Lipschitz, and let $\Gamma = \Gamma_\psi$ be its intrinsic graph. Let $U \subseteq V_0$ be a measurable set. For almost every $L \in \mathcal{L}_P$, the intersection $L \cap \Gamma^+$ has locally finite perimeter. If $\mathcal{M} \subseteq \mathcal{L}_P$ is the set of lines that intersect Γ at least twice, then*

$$\int_{\mathcal{M}} \mathrm{Per}_{\Gamma^+, L}(\Pi^{-1}(U))\, d\mathcal{N}_P(L) \lesssim_\lambda |U|. \tag{9.7}$$

Proof. The measures $\mathcal{N}_P$ and $\mathcal{N}$ are absolutely continuous with respect to each other. Indeed, for each $m>0$, if $D \subseteq \mathcal{L}_P$ is a set of lines with slopes that lie in $[-m, m]$, then $\mathcal{N}_P(D) \asymp_m \mathcal{N}(D)$. By (2.24), there is $c>0$ such that, for any measurable $A \subseteq \mathbb{H}$,

$$\mathrm{Per}_{\Gamma^+}(A) = c \int_{\mathcal{L}} \mathrm{Per}_{\Gamma^+, L}(A)\, d\mathcal{N}(L).$$

Since Γ^+ has locally finite perimeter, this implies that, for almost every line $L \in \mathcal{L}_P$, the intersection $L \cap \Gamma^+$ has locally finite perimeter. For $L \in \mathcal{L}_P$ let $m(L)$ be the slope of L as in §8. Suppose that $p \in L \cap \Gamma$. By (2.14), if $|m(L)| > \lambda/\sqrt{1-\lambda^2}$, then $L \subseteq p \cdot \mathrm{Cone}_\lambda$ and thus L intersects Γ exactly once. Consequently, $|m(M)| \leqslant \lambda/\sqrt{1-\lambda^2}$ for every $M \in \mathcal{M}$, and hence $\mathcal{N}_P(D) \asymp_\lambda \mathcal{N}_P(D)$ for every measurable $D \subseteq \mathcal{M}$. So, by (2.24) and Lemma 2.5,

$$\int_{\mathcal{M}} \mathrm{Per}_{\Gamma^+, L}(\Pi^{-1}(U))\, d\mathcal{N}_P(L) \lesssim_\lambda \mathrm{Per}_{\Gamma^+}(\Pi^{-1}(U)) \asymp_\lambda |U|. \qquad \square$$

Proof of Lemma 9.2. For a finite-perimeter set $S\subseteq\mathbb{R}$ and $R>0$, let $\mathcal{I}(S)$ and

$$\widehat{\omega}_{S,R}=\tfrac{1}{2}(\omega_{S,R}+\omega_{\mathbb{R}\setminus S,R})$$

be as in §8.1. Divide $\mathcal{I}(S)$ according to the length of the intervals as follows:

$$C_j(S)\stackrel{\mathrm{def}}{=}\{I\in\mathcal{I}(S):2^{-j-1}<|I|\leqslant 2^{-j}\}\quad\text{for all } j\in\mathbb{Z}.$$

Let $\mathcal{E}_j(S)\subseteq\mathbb{R}$ be the set of endpoints of the intervals in $C_j(S)$. Let $\lambda_{S,j}$ be the counting measure on $\mathcal{E}_j(S)$ and let

$$\hat{\lambda}_j(S)\stackrel{\mathrm{def}}{=}\tfrac{1}{2}(\lambda_{S,j}+\lambda_{\mathbb{R}\setminus S,j}).$$

Then,

$$\sum_{j\in\mathbb{Z}}\hat{\lambda}_{S,j}\leqslant\mathrm{Per}_S\,.$$

(This is not necessarily an equality as the left-hand side is influenced only by bounded intervals while the right-hand side could have a contribution from rays.)

For each $k\in\mathbb{Z}$, the measure $\widehat{\omega}_{S,2^{-k}}$ is a point measure supported on the set

$$\bigcup_{j=k}^{\infty}(\mathcal{E}_j(S)\cup\mathcal{E}_j(\mathbb{R}\setminus S)),$$

that weights each point according to the lengths of the intervals it bounds. In particular,

$$\mathrm{supp}(\widehat{\omega}_{S,2^{-k}}-\widehat{\omega}_{S,2^{-k-1}})\subseteq\mathcal{E}_k(S)\cup\mathcal{E}_k(\mathbb{R}\setminus S)$$

and

$$2^{-k-2}\leqslant\widehat{\omega}_{S,2^{-k}}(p)-\widehat{\omega}_{S,2^{-k-1}}(p)\leqslant 2^{-k}\quad\text{for all } p\in\mathcal{E}_k(S)\cup\mathcal{E}_k(\mathbb{R}\setminus S).$$

Consequently, if we denote

$$\widehat{\kappa}_{S,k}\stackrel{\mathrm{def}}{=}2^k(\widehat{\omega}_{S,2^{-k}}-\widehat{\omega}_{S,2^{-k-1}}),$$

then $\widehat{\kappa}_{S,j}\asymp\hat{\lambda}_{S,j}$ and

$$\sum_{j\in\mathbb{Z}}\widehat{\kappa}_{S,j}=\sum_{j\in\mathbb{Z}}2^{j+1}\widehat{\omega}_{S,2^{-j}}-\sum_{j\in\mathbb{Z}}2^{j}\widehat{\omega}_{S,2^{-j}}=\sum_{j\in\mathbb{Z}}2^{j}\widehat{\omega}_{S,2^{-j}}.$$

It follows that

$$\sum_{j\in\mathbb{Z}}2^{j}\widehat{\omega}_{S,2^{-j}}\asymp\sum_{j\in\mathbb{Z}}\hat{\lambda}_{S,j}\leqslant\mathrm{Per}_S\,.\tag{9.8}$$

For every measurable $E \subseteq \mathbb{H}$ and $U \subseteq V_0$, and every $L \in \mathcal{L}_P$, we have

$$\sum_{j\in\mathbb{Z}} 2^j \widehat{\omega}^P_{E,2^{-j}}(U,L) \overset{(8.5)}{\underset{(9.8)}{\lesssim}} \mathrm{Per}_{x(E\cap L)}(x(\Pi^{-1}(U)\cap L)) = \mathrm{Per}_{E,L}(\Pi^{-1}(U)). \tag{9.9}$$

Let $\mathcal{M} \subseteq \mathcal{L}_P$ be the set of lines that intersect Γ at least twice. If $L \in \mathcal{L}_P \setminus \mathcal{M}$, then $\mathcal{I}(L\cap\Gamma^+)$ consists of infinite rays, so $\widehat{\omega}_{\Gamma^+,R}(U,L)=0$ for any $U \subseteq V_0$. Thus,

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \Omega^P_{\Gamma^+,2^{-j}}(U) &\overset{(8.6)}{=} \sum_{j\in\mathbb{Z}} \int_{\mathcal{M}} 2^j \widehat{\omega}^P_{\Gamma^+,2^{-j}}(U,L)\, d\mathcal{N}_P(L) \\ &\overset{(9.9)}{\lesssim} \int_{\mathcal{M}} \mathrm{Per}_{\Gamma^+,L}(\Pi^{-1}(U))\, d\mathcal{N}_P(L) \overset{(9.7)}{\lesssim_\lambda} |U|. \qquad \square\end{aligned}$$

10. Outline of proof of Proposition 7.2

The rest of this paper is dedicated to the proof of Proposition 7.2. This is the longest part of the proof of Theorem 5.2, and we will divide it into two pieces.

In the first step (§11), we prove the following Proposition 10.1, which is a stability result for extended-monotone sets (Definition 8.5). For every $r>0$ and $h\in\mathbb{H}$, let $\overline{B}_r(h) \subseteq \mathbb{H}$ be the convex hull of $B_r(h)$ (as a subset of $\mathbb{R}^3$); when h is omitted, we take it to be $\mathbf{0}$. The convex hull of B_r with respect to the horizontal lines or with respect to all lines in $\mathbb{R}^3$ is the same, and $\overline{B}_r \subseteq B_{2r}$.

Proposition 10.1. *Let $E \subseteq \mathbb{H}$ be a measurable set. For any $\varepsilon>0$, there are $\nu, R>0$ such that, if $E \subseteq \mathbb{H}$ is (ν', R')-extended monotone on $\overline{B}_1$ for some $\nu', R'>0$ that satisfy $R' \geqslant R$ and $\nu' R' \leqslant \nu R$, then there is a plane $P \subseteq \mathbb{H}$ such that*

$$|\overline{B}_1 \cap (P^+ \triangle E)| < \varepsilon.$$

If Γ is an intrinsic Lipschitz graph and $E=\Gamma^+$, then we can take P to be a vertical plane.

Proposition 10.1 is in the spirit of the stability theorem for monotone sets that was proved in [23], though here we do not need to obtain an explicit dependence of ν and R on ε (in [23] it was important to get power-type dependence). The lack of explicit dependence lets us use a compactness argument that was not available in the context of [23]. At the same time, [23, Theorem 4.3] states that, if the non-monotonicity of E is small on the unit ball B_1, then there is a smaller ball B_{ε^3} on which E is $O(\varepsilon)$-close to a plane, while Proposition 10.1 assumes a stronger hypothesis, namely that $\mathrm{ENM}_{E,R}(\overline{B}_1) < \nu$, and obtains the stronger conclusion that E is close to a plane on the same ball $\overline{B}_1$.

Remark 10.2. The stronger conclusion above is crucial for the covering argument that we used in §9 because of the delicacy of the Vitali-type argument used in Lemma 9.4. We use Lemma 9.4 to show that, if Δ is a μ-rectilinear foliated patchwork, with $0<\mu<1/32r^2$, and $F\subseteq\mathcal{V}(\Delta)$ is a collection of vertices corresponding to pseudoquads of the same width, then there is a large subset G of these pseudoquads such that, if $Q, Q'\in G$, then rQ is disjoint from rQ'. Lemma 9.4 only holds when $\mu=O(r^{-2})$. If μr^2 is too large, then a μ-rectilinear foliated patchwork could contain arbitrarily many vertically cut pseudoquads $Q_1, ..., Q_n$ of equal height and width such that $rQ_1, ..., rQ_n$ all intersect.

We do not see how a modified subdivision algorithm that uses monotonicity instead of paramonotonicity can ensure that the conditions of Lemma 9.4 are satisfied. For example, consider a modification of the subdivision algorithm in §7 that produces a patchwork Δ by cutting a pseudoquad Q horizontally or vertically depending on whether Γ is η-monotone (rather than paramonotone) on rQ for some $r>0$. Note that [23, Theorem 4.3] implies that, if Γ is sufficiently monotone on rQ, then Γ is $O(r^{-1/3})$-close to a plane on Q. Indeed, there are sets that have zero non-monotonicity on rQ, but are only $\varepsilon(r)$-close to a plane on Q, where $\varepsilon(r)\to 0$ as $r\to\infty$. It follows that this modified algorithm can, at best, produce $\mu(r)$-rectilinear foliated patchworks, where $\mu(r)\to 0$ as $r\to\infty$. In particular, as $\mu(r)$ depends on r, we cannot choose μ so that $\mu<1/32r^2$.

Consequently, we cannot prove the weighted Carleson condition for this modified algorithm. The weighted Carleson condition bounds the number of vertically cut pseudoquads based on the total non-monotonicity of Γ, but without Lemma 9.4, a small amount of non-monotonicity can lead to many vertically cut pseudoquads. That is, if $Q_1, ..., Q_n$ are pseudoquads in the patchwork such that $rQ_1, ..., rQ_n$ all intersect, then non-monotonicity on the intersection $rQ_1\cap...\cap rQ_n$ could force the algorithm to cut all of the Q_i's vertically.

Using extended non-monotonicity rather than non-monotonicity lets us avoid this problem. The fact that r is a universal constant in Proposition 7.2 means that, for any μ, there is a subdivision algorithm that produces a μ-rectilinear foliated patchwork by cutting each pseudoquad Q based on whether Γ is $(\eta(\mu), R(\mu))$-paramonotone on rQ. In particular, we can choose $\mu<1/32r^2$, so that Lemma 9.4 applies.

In the second step, we prove parts (1) and (2) of Proposition 7.2. By Remark 4.3, after a stretch, shear, and translation, we may suppose that Q is a rectilinear pseudoquad for Γ that is close to $[-1,1]^2$ and Γ is (η, R)-paramonotone on rQ. For any given c, if R is sufficiently large, η is sufficiently small, and $\Pi(\overline{B}_c)\subseteq rQ$, then, by Lemma 8.7, Γ^+ has small extended non-monotonicity on $\overline{B}_c$, so Γ^+ is close to a half-space P^+ on $\overline{B}_c$.

Note that, even though Γ^+ is close to a half-space P^+ on $\overline{B}_c$, it does not immediately follow that the corresponding intrinsic Lipschitz function f is L_1-close to an affine function. Using Remark 4.3 to normalize Q stretches Γ and changes its intrinsic Lipschitz constant. Consequently, even though f is close to an affine function on most of Q, it may still take on large values on the rest of Q. To show that this does not happen, we must introduce new methods based on analyzing the characteristic curves of Γ.

For example, a key step in the proof of part (1) of Proposition 7.2 is to show that $\|f\|_{L_1(Q)}$ is bounded. Since f is intrinsic Lipschitz, $\|f\|_{L_1(Q)}<\infty$, but we need a bound independent of the intrinsic Lipschitz constant. We obtain such a bound by studying how lines intersect the characteristic curves. Since Q is μ-rectilinear, the top and bottom boundaries of Q are characteristic curves that are close to the top and bottom edges of $[-1,1]^2$. If L is a horizontal line such that $\Pi(L)$ crosses $[-1,1]^2$ from top to bottom, then $\Pi(L)$ must also cross the top and bottom boundaries of Q. At these intersection points, the slope of $\Pi(L)$ is less than the slope of the boundary, so the corresponding points of L lie in Γ^+. If $\Gamma^+\cap L$ is close to monotone, then most of the interval between these points lies in Γ^+ and therefore, f is bounded on $Q\cap\Pi(L)$. By integrating over a family of lines that all cross the top and bottom boundaries, we obtain the desired L_1 bound. Similar arguments based on characteristic curves lead to part (2) of Proposition 7.2, which completes the proof of Proposition 7.2.

11. Extended-monotone sets are close to half-spaces

In this section, we will prove Proposition 10.1 by studying limits of (ε,R)-extended monotone sets. Let $U\subseteq\mathbb{H}$ be measurable and let $E_1,E_2,\ldots\subseteq\mathbb{H}$ be a sequence of measurable sets such that E_i is $(1/i,i)$-extended monotone on U. By passing to a subsequence, we may suppose that $\mathbf{1}_{E_i}$ converges weakly to a function $f\in L_\infty(\mathbb{H})$ taking values in $[0,1]$. We call f a *U-LEM (limit of extended monotones) function.*

One difficulty of studying f is that it need not take values only in $\{0,1\}$. Indeed, the extended monotonicity $\mathrm{ENM}_{E_i,i}(\overline{B}_1)$ only depends on the intersection of E_i with lines that pass through $\overline{B}_1$. These lines do not cover all of $\mathbb{H}$, so there are regions of $\mathbb{H}$ where f can take on arbitrary values.

Nevertheless, in §11.1, we will show that, after changing f on a measure-zero set, $f(\overline{B}_1)\subseteq\{0,1\}$. This will follow from the fact that, by Lemma 8.6,

$$\lim_{i\to\infty}\mathrm{NM}_{E_i}(\overline{B}_1)=0.$$

We will show that a sequence of sets with non-monotonicity going to zero on $\overline{B}_1$ converges to a subset which is monotone on $\overline{B}_1$. If U is an open set, a subset $E\subseteq\mathbb{H}$ is said to be *monotone* on U if $\mathrm{NM}_E(U)=0$.

Then, in §11.2, we will use techniques from [21] and [23] to characterize sets such that $\mathrm{NM}_F(\overline{B}_1)=0$. A set that is monotone on $\overline{B}_1$ need not be a half-space, but we will show that if F is such a set, then the measure-theoretic boundary $\partial_{\mathcal{H}^4}F$ is a union of horizontal lines that has an *approximate tangent plane* at every point. That is, for any $g\in\partial_{\mathcal{H}^4}F$, the blowups $g\cdot s_{n,n}(g^{-1}\partial_{\mathcal{H}^4}F)$ converge in the Hausdorff metric to a plane T_g as $n\to\infty$. In fact, at all but countably many points $g\in\partial_{\mathcal{H}^4}F$, there is a unique horizontal line L_g through g that is contained in $\partial_{\mathcal{H}^4}F$, and T_g is the vertical plane containing L_g; in this case, g has an approximate tangent subgroup in the sense of [71]. At the remaining points, T_g is the horizontal plane centered at g.

Finally, in §11.3, we prove Proposition 10.1. The proof is somewhat involved, but, as an illustration, we consider the case that $f=\mathbf{1}_E$, where E is precisely ∞-extended monotone on $\overline{B}_1$. That is, for every line L, either $\overline{B}_1\cap\partial(L\cap E)=\varnothing$ or $L\cap E$ is a monotone subset of L.

We first claim that, for every point $b\in\overline{B}_1\cap\partial_{\mathcal{H}^4}E$, if the approximate tangent plane T_b is vertical and H_b is the horizontal plane centered at b, then

$$\mathsf{H}_b\cap\partial_{\mathcal{H}^4}E=\mathsf{H}_b\cap T_b.$$

Let $T_b^{\pm}$ be the two half-spaces bounded by T_b, labeled so that $T_b^{+}\cap B_r(b)$ approximates $E\cap B_r(b)$ at small scales. Let $L_b=\mathsf{H}_b\cap T_b$ be the horizontal line in $\partial_{\mathcal{H}^4}E$ that passes through b and let L be a line through b that intersects T_b transversally. Then $E\cap L$ is a monotone set with $b\in\partial_{\mathcal{H}^1}(E\cap L)$, so

$$T_b^{+}\cap L\subseteq E\cap L \quad\text{and}\quad T_b^{-}\cap L\subseteq L\setminus E.$$

This holds for every horizontal line through b except L_b, so L_b cuts H_b into two half-planes $P_\pm=T_b^{\pm}\cap\mathsf{H}_b$ such that $P_+\subseteq E$ and $P_-\subseteq\mathsf{H}_b\setminus E$.

When $b'\in L_b$ is close to b, the plane $\mathsf{H}_{b'}$ intersects H_b along L_b and the angle between the two planes is small. As above, there are two half-planes $P'_\pm=T_{b'}^{\pm}\cap\mathsf{H}_{b'}$ such that $P'_+\subseteq E$ and $P'_-\subseteq\mathsf{H}_{b'}\setminus E$. As b' varies over points close to b, the half-plane P'_+ varies over half-planes close to P_+. Therefore P_+ is in the interior of E, P_- is in the exterior, and

$$\mathsf{H}_b\cap\partial_{\mathcal{H}^4}E=L_b.$$

Suppose that L_1 and L_2 are two lines in $\partial_{\mathcal{H}^4}E$ that intersect $\overline{B}_1$, and suppose by way of contradiction that they are not coplanar. By the hyperboloid lemma [21, Lemma 2.4] (see Lemma 11.1), for any point $q\in L_1$ except possibly a single point, there is a horizontal line M that connects q to a point r in L_2. Then $r\in\mathsf{H}_q\cap\partial_{\mathcal{H}^4}E=L_1$, so L_1 and L_2 intersect and are thus coplanar; this is a contradiction. It follows that $\overline{B}_1\cap\partial_{\mathcal{H}^4}E$ is contained in a plane. The proof of Proposition 10.1 runs along the same lines, but it takes some further technical work to apply the weaker hypothesis that f is merely an LEM function.

One of the key tools in the proof is the following "hyperboloid lemma", which is stated as [21, Lemma 2.4]. A pair of horizontal lines $L_1, L_2 \in \mathcal{L}$ are said to be *skew* if L_1 and L_2 are disjoint and the projections $\pi(L_1), \pi(L_2) \subseteq \mathsf{H} \cong \mathbb{R}^2$ are not parallel.

LEMMA 11.1. (Cheeger–Kleiner hyperboloid lemma [21]) *For any $L_1, L_2 \in \mathcal{L}$, the following statements hold.*

(1) *Suppose that the projections $\pi(L_1)$ and $\pi(L_2)$ are parallel but $\pi(L_1) \neq \pi(L_2)$. Then, every point in L_1 can be joined to L_2 by a unique line. In fact, there is a unique fiber $\pi^{-1}(p)$ such that every line joining L_1 to L_2 passes through $\pi^{-1}(p)$. Conversely, for every $a \in \pi^{-1}(p)$, there is a unique line joining L_1 to L_2 that passes through a.*

(2) *If L_1 and L_2 are skew, then there is a hyperbola $S \subseteq \mathsf{H}$ with asymptotes $\pi(L_1)$ and $\pi(L_2)$ such that every tangent line of S has a unique horizontal lift that intersects L_1 and L_2. If $p \in \mathsf{H}$ is the intersection between $\pi(L_1)$ and $\pi(L_2)$, and $a \in L_1$ is such that $\pi(a) \neq p$, then there is a unique horizontal line that connects a to a point in L_2.*

11.1. Stability of locally monotone sets

We begin the proof of Proposition 10.1 by using a compactness argument to prove the following lemma. Throughout what follows, given a measure space $(\mathcal{S}, \Sigma, \mu)$ and a measurable subset $\Omega \in \Sigma$ with $\mu(\Omega) > 0$, we use the (standard) notation $⨍_\Omega$ to denote the averaging operator on Ω, i.e.,

$$⨍_\Omega f \, d\mu \stackrel{\text{def}}{=} \frac{1}{\mu(\Omega)} \int_\Omega f \, d\mu \quad \text{for all } f \in L_1(\Omega, \mu).$$

LEMMA 11.2. *Let $U \subseteq \mathbb{H}$ be a bounded open set and let $E_1, E_2, ... \subseteq \mathbb{H}$ be a sequence of measurable sets such that $\mathrm{NM}_{E_i}(U) < 1/i$ for every $i \in \mathbb{N}$. There is a subsequence $(E_{i_j})_{j \in \mathbb{N}}$ and a set $F \subseteq U$ that is monotone on U such that*

$$\lim_{j \to \infty} |(E_{i_j} \cap U) \triangle F| = 0.$$

It follows that, for any $\varepsilon > 0$, there is a $\delta > 0$ such that if $E \subseteq \mathbb{H}$ is a measurable set and $\mathrm{NM}_E(U) < \delta$, then there is a set $F \subseteq U$ such that $|(E \cap U) \triangle F| < \varepsilon$ and F is monotone on U.

Proof. After passing to a subsequence, we may suppose that the characteristic functions $\mathbf{1}_{E_i}$ converge weakly to a function $f \in L_\infty(U)$ taking values in $[0, 1]$. We claim that f is a characteristic function.

By [23, Theorem 4.3] (see also [91, Theorem 63]), for every $\varepsilon>0$ there are $c(\varepsilon)>0$ and $\delta(\varepsilon)>0$ such that if $p\in\mathbb{H}$, $\alpha>0$, and $\mathrm{NM}_E(B_\alpha(p))<\delta(\varepsilon)\alpha^{-3}$, then there is a half-space P^+ such that

$$\fint_{B_{c(\varepsilon)\alpha}(p)} |\mathbf{1}_{P^+}(h)-\mathbf{1}_E(h)|\,d\mathcal{H}^4(h)<\varepsilon. \tag{11.1}$$

(The hypothesis in [23] is that $\mathrm{NM}_E(B_\alpha(p))<\delta(\varepsilon)$, but our definition of $\mathrm{NM}_E(B_\alpha(p))$ differs from the definition in [23] by a normalization factor.)

By the Lebesgue density theorem, for almost every point $p\in U$, we have

$$\lim_{s\to 0}\fint_{B_s(p)} |f(h)-f(p)|\,d\mathcal{H}^4(h)=0. \tag{11.2}$$

Let p be such a point and let $r>0$ be such that $B_r(p)\subseteq U$. By (11.1), for any $0<s<r$, any $\varepsilon>0$, and any sufficiently large $i\in\mathbb{N}$ (depending on s,ε), there is a half-space Q_i^+ with

$$\fint_{B_{c(\varepsilon)s}(p)} |\mathbf{1}_{Q_i^+}(h)-\mathbf{1}_{E_i}(h)|\,d\mathcal{H}^4(h)<\varepsilon.$$

Choose a half-space Q^+ such that, for infinitely many $i\in\mathbb{N}$, we have

$$\fint_{B_{c(\varepsilon)s}(p)} |\mathbf{1}_{Q^+}(h)-\mathbf{1}_{E_i}(h)|\,d\mathcal{H}^4(h)<2\varepsilon.$$

Then,

$$\fint_{B_{c(\varepsilon)s}(p)} |\mathbf{1}_{Q^+}(h)-f(h)|\,d\mathcal{H}^4(h)<3\varepsilon. \tag{11.3}$$

Since the function $(x\in[0,1])\mapsto x(1-x)$ is non-negative and 1-Lipschitz,

$$\fint_{B_{c(\varepsilon)s}(p)} f(h)(1-f(h))\,d\mathcal{H}^4(h)\leqslant 3\varepsilon+\fint_{B_{c(\varepsilon)s}(p)} \mathbf{1}_{Q^+}(h)(1-\mathbf{1}_{Q^+}(h))\,d\mathcal{H}^4(h)=3\varepsilon.$$

This holds for all $0<s<r$, so

$$\lim_{s\to 0}\fint_{B_s(p)} f(h)(1-f(h))\,d\mathcal{H}^4(h)=0.$$

By (11.2), this implies $f(p)(1-f(p))=0$, and hence $f(p)\in\{0,1\}$.

Thus, f is equivalent to a characteristic function on U. Let $F=f^{-1}(1)$. By weak convergence,

$$\lim_{i\to\infty}|U\cap(E_i\triangle F)|=0.$$

For any $i\in\mathbb{N}$,

$$\begin{aligned}\mathrm{NM}_F(U)&=\int_{\mathcal{L}}\mathrm{NM}_{F\cap L}(U\cap L)\,d\mathcal{N}(L)\\&\leqslant\int_{\mathcal{L}}(\mathrm{NM}_{E_i\cap L}(U\cap L)+\mathcal{H}^1(U\cap L\cap(E_i\triangle F)))\,d\mathcal{N}(L)\\&\lesssim\mathrm{NM}_{E_i}(U)+|U\cap(E_i\triangle F)|.\end{aligned}$$

Both terms on the right-hand side go to zero as $i \to \infty$, so

$$\mathrm{NM}_F(U) = 0,$$

i.e., F is monotone on U. □

Corollary 11.3. *Let $U \subseteq \mathbb{H}$ be a convex bounded open set and let $f\colon \mathbb{H} \to [0,1]$ be a U-LEM function. There is a monotone set $E \subseteq U$ such that $f|_U = \mathbf{1}_E$ up to a measure-zero set.*

Proof. Suppose that $E_1, E_2, \ldots \subseteq \mathbb{H}$ are measurable, E_i is $(1/i, i)$-extended monotone on U for all $i \in \mathbb{N}$, and $\mathbf{1}_{E_i}$ converges weakly to f. By Lemma 8.6, for $i > \operatorname{diam} U$ we have

$$\mathrm{NM}_{E_i}(U) \leqslant \mathrm{ENM}_{E_i, i}(U) \leqslant \frac{1}{i}.$$

So, by Lemma 11.2, $f|_U = \mathbf{1}_F$ for some set $F \subseteq U$ that is monotone on U. □

11.2. Locally monotone sets are bounded by rectifiable ruled surfaces

Here we will describe sets that are monotone on an open subset of $\mathbb{H}$, which we call *locally monotone sets*. Note that a locally monotone set need not be a half-space; see [23, Example 9.1]. Regardless, we use the techniques developed in [21] and [23] to describe such sets.

Proposition 11.4. *Let $E \subseteq \mathbb{H}$ be a measurable set that is monotone on a convex open set $U \subseteq \mathbb{H}$. Then, the following statements hold.*

(1) *$U \cap \partial_{\mathcal{H}^4} E$ has empty interior.*

(2) *For every $p \in U \cap \partial_{\mathcal{H}^4} E$, there is a horizontal line L through p with $U \cap L \subseteq \partial_{\mathcal{H}^4} E$. If this line is not unique, then $U \cap \mathsf{H}_p \subseteq \partial_{\mathcal{H}^4} E$, and we call p a* characteristic point.

(3) *$\partial_{\mathcal{H}^4} E$ has an approximate tangent plane T_p at every $p \in U \cap \partial_{\mathcal{H}^4} E$. The plane T_p is horizontal if and only if p is a characteristic point, and there are only countably many characteristic points in U.*

(4) *If T_p is vertical, then it divides $\mathbb{H}$ into two half-spaces T_p^+ and T_p^- such that the following holds. For $\varepsilon, t > 0$, let*

$$W^{\pm}_{\varepsilon,t} = \{v \in T_p^{\pm} \cap \overline{B}_t(p) : d(v, T_p) > \varepsilon t\}$$

For any $0 < \varepsilon < \frac{1}{10}$, there is $r > 0$ such that, if $0 < \alpha < r$, then

$$W^{+}_{\varepsilon,\alpha} \subseteq \operatorname{int}_{\mathcal{H}^4}(E) \quad \textit{and} \quad W^{-}_{\varepsilon,\alpha} \subseteq \operatorname{int}_{\mathcal{H}^4}(\mathbb{H} \setminus E).$$

We rely on the following proposition and lemmas, which adapt results from [21].

Proposition 11.5. (Generalization of [21, Proposition 5.8]) *Let $E \subseteq \mathbb{H}$ be a measurable set that is monotone on a convex open set $U \subseteq \mathbb{H}$. Let L be a horizontal line and let $p, q \in L$ be points such that $p \neq q$ and the segment $[p, q] \subseteq L$ is contained in U. We choose the linear order on L so that $p<q$. Suppose that $q \in \mathrm{int}_{\mathcal{H}^4}(E)$.*

(1) *If $p \in \mathrm{supp}_{\mathcal{H}^4}(E)$ and $r \in L \cap U$ satisfies $p<r<q$, then $r \in \mathrm{int}_{\mathcal{H}^4}(E)$.*

(2) *If $p \in \mathrm{supp}_{\mathcal{H}^4}(\mathbb{H} \backslash E)$ and $r \in L \cap U$ satisfies $p<q<r$, then $r \in \mathrm{int}_{\mathcal{H}^4}(E)$.*

Proof. [21, Proposition 5.8] proves this result in the case that $U = \mathbb{H}$, generalizing [21, Proposition 4.6], which proves it when E is *precisely monotone* (i.e., $M \cap E$ and $M \cap E^c$ are connected sets for every horizontal line M). The reasoning in [21, Proposition 5.8] only uses the fact that, for almost every line segment S in a small neighborhood of $[p, \max\{q, r\}]$, the intersection $S \cap E$ is monotone. This holds here, so the conclusion of Proposition 5.8 holds here as well. For completeness, we will sketch the argument of [21].

For any $x \in \mathbb{H}$ and $v_1, v_2 \in \mathsf{H}$, let $\gamma_{x,v_1,v_2} : [0, 2] \to \mathbb{H}$ be the broken geodesic

$$\gamma_{x,v_1,v_2}(t) = \begin{cases} x v_1^t, & t \in [0, 1], \\ x v_1 v_2^{t-1}, & t \in [1, 2]. \end{cases}$$

In case (1), we have $p<r<q$ with $p \in \mathrm{supp}_{\mathcal{H}^4}(E)$ and $q \in \mathrm{int}_{\mathcal{H}^4}(E)$. Given an $\varepsilon>0$, one considers the paths γ_{x,v_1,v_2}, where $x \in B_\varepsilon(p) \cap E$ and $v_1, v_2 \in \mathsf{H}$ satisfy

$$\|v_i - (p^{-1} r)^{1/2}\| < \varepsilon.$$

Then $\gamma_{x,v_1,v_2}(0)$ is close to p, $\gamma_{x,v_1,v_2}(2)$ is close to r, and γ_{x,v_1,v_2} lies in a small neighborhood of $[p, q]$. Further, for any x, we can vary v_1 and v_2 so that $\gamma_{x,v_1,v_2}(2) = x v_1 v_2$ covers a neighborhood of r.

Suppose that E is precisely monotone and that $q \in \mathrm{int}(E)$. Let x, v_1, and v_2 be as above, and let $\lambda_1(t) = x v_1^t$ and $\lambda_2(t) = x v_1 v_2^t$ be the two segments of γ_{x,v_1,v_2}. These are two lines that are close to L, so there are $t_1, t_2 > 1$ such that $\lambda_i(t_i)$ is close to q. Since $q \in \mathrm{int}(E)$, if ε is sufficiently small, then $\lambda_i(t_i) \in E$. Since $\lambda_1(0) = x \in E$ and $\lambda_1(t_1) \in E$, precise monotonicity implies $\lambda_1(1) \in E$, and since $\lambda_2(0) = \lambda_1(1) \in E$ and $\lambda_2(t_2) \in E$, we have $\lambda_2(1) = x v_1 v_2 \in E$. If we fix x and let v_1 and v_2 vary, then $x v_1 v_2$ covers a neighborhood of r, so $r \in \mathrm{int}(E)$.

In our case, E is not precisely monotone and $q \in \mathrm{int}_{\mathcal{H}^4}(E)$, but the reasoning above still holds for almost every triple (x, v_1, v_2). As $\mathcal{H}^4(B_\varepsilon(p) \cap E) > 0$, there is $x \in B_\varepsilon(p) \cap E)$ such that $x v_1 v_2 \in E$ for almost every pair (v_1, v_2). Therefore, $r \in \mathrm{int}_{\mathcal{H}^4}(E)$.

In case (2), we have $p<q<r$, with $p \in \mathrm{supp}_{\mathcal{H}^4}(\mathbb{H} \backslash E)$ and $q \in \mathrm{int}_{\mathcal{H}^4}(E)$. Let $s \in L \cap U$ be such that $p<q<r<s$, and consider γ_{x,v_1,v_2} such that

$$x \in B_\varepsilon(p) \backslash E, \quad \|v_1 - p^{-1} s\| < \varepsilon, \quad \text{and} \quad \|v_2 - s^{-1} r\| < \varepsilon.$$

That is, γ_{x,v_1,v_2} is a path from a neighborhood of p to a neighborhood of s to a neighborhood of r. Again, for any x, we can vary v_1 and v_2 so that $\gamma_{x,v_1,v_2}(2)=xv_1v_2$ covers a neighborhood of r. If ε is sufficiently small, we have $\gamma_{x,v_1,v_2}([0,2])\subseteq U$.

Suppose again that E is precisely monotone and that $q\in\mathrm{int}(E)$. Let

$$\lambda_1(t)=xv_1^t \quad\text{and}\quad \lambda_2(t)=xv_1v_2^t.$$

Since λ_1 and λ_2 are both close to L, if ε is sufficiently small, there are $t_1\in(0,1)$ and $t_2>1$ such that $\lambda_i(t_i)$ is close to q and $\lambda_i(t_i)\in E$. Since $\lambda_1(0)=x\notin E$ and $\lambda_1(t_1)\in E$, we have $\lambda_1(1)\in E$, and since $\lambda_2(0)=\lambda_1(1)\in E$ and $\lambda_2(t_2)\in E$, we have $\lambda_2(1)=xv_1v_2\in E$. For any fixed x, as v_1 and v_2 vary, xv_1v_2 covers a neighborhood of r.

Again, when E is not precisely monotone and $q\in\mathrm{int}_{\mathcal{H}^4}(E)$, the reasoning above fails for a null set of triples (x,v_1,v_2). Since $p\in\mathrm{supp}_{\mathcal{H}^4}(\mathbb{H}\setminus E)$, there is an $x\in B_\varepsilon(p)\cap(\mathbb{H}\setminus E)$ such that $xv_1v_2\in E$ for all but a measure-zero set of pairs (v_1,v_2), so $r\in\mathrm{int}_{\mathcal{H}^4}(E)$. □

By Proposition 11.5 and the proof of [21, Lemma 4.8], we get the following lemma.

Lemma 11.6. (Generalization of [21, Lemma 4.8]) *Let $E\subseteq\mathbb{H}$ be a measurable set that is monotone on a convex open set $U\subseteq\mathbb{H}$. If L is a horizontal line such that $L\cap U$ contains at least two points of $\partial_{\mathcal{H}^4}E$, then $L\cap U\subseteq\partial_{\mathcal{H}^4}E$.*

Proof. Let $I=L\cap U$. Let $p,q\in I\cap\partial_{\mathcal{H}^4}E$ be distinct points. Choose the linear order on L so that $p<q$. Let $r\in I$ be such that $q<r$. By part (1) of Proposition 11.5, if $r\in\mathrm{int}_{\mathcal{H}^4}(E)$, then $q\in\mathrm{int}_{\mathcal{H}^4}(E)$, which is a contradiction. Likewise, if $r\in\mathrm{int}_{\mathcal{H}^4}(\mathbb{H}\setminus E)$, then $q\in\mathrm{int}_{\mathcal{H}^4}(\mathbb{H}\setminus E)$, which is a contradiction, so $r\in\partial_{\mathcal{H}^4}E$. Thus $[q,\infty)\cap I\subseteq\partial_{\mathcal{H}^4}E$. By symmetry, $I\setminus(p,q)=I\cap\big((-\infty,p]\cup[q,\infty)\big)\subseteq\partial_{\mathcal{H}^4}E$ for any distinct points $p,q\in I\cap\partial_{\mathcal{H}^4}E$. Let $r,s\in I\cap[q,\infty)$ be such that $r<s$. Then $r,s\in I\cap\partial_{\mathcal{H}^4}E$, so $I\setminus(r,s)\subseteq\partial_{\mathcal{H}^4}E$. Since (r,s) and (p,q) are disjoint, $I\subseteq\partial_{\mathcal{H}^4}E$. □

Likewise, the following lemma is based on the proof of [21, Lemma 4.9].

Lemma 11.7. (Generalization of [21, Lemma 4.9]) *Let $E\subseteq\mathbb{H}$ be a measurable set that is monotone on a convex open set $U\subseteq\mathbb{H}$. For every $p\in U\cap\partial_{\mathcal{H}^4}E$, there is a horizontal line L such that $p\in L$ and $L\cap U\subseteq\partial_{\mathcal{H}^4}E$.*

Proof. Let $B\subseteq U$ be a ball centered at p and let H_p be the horizontal plane centered at p. Let $B'=B\setminus\{p\}$. Suppose by way of contradiction that $\mathsf{H}_p\cap B'\cap\partial_{\mathcal{H}^4}E=\varnothing$. Since $\mathsf{H}_p\cap B'$ is connected, we have $\mathsf{H}_p\cap B'\subseteq\mathrm{int}_{\mathcal{H}^4}(E)$ or $\mathsf{H}_p\cap B'\subseteq\mathrm{int}_{\mathcal{H}^4}(\mathbb{H}\setminus E)$. Without loss of generality, we assume that $\mathsf{H}_p\cap B'\subseteq\mathrm{int}_{\mathcal{H}^4}(E)$.

Let M be a line through p and let $q,r\in M\cap B$ be two points on opposite sides of p. Then $q,r\in\mathrm{int}_{\mathcal{H}^4}(E)$, so, by part (1) of Proposition 11.5, we have $p\in\mathrm{int}_{\mathcal{H}^4}(E)$. This is a contradiction, so there exists some point q lying in $\mathsf{H}_p\cap B'\cap\partial_{\mathcal{H}^4}E$. Let L be the line containing p and q. Then, by Lemma 11.6, $L\cap U\subseteq\partial_{\mathcal{H}^4}E$, as desired. □

The fact that $U\cap\partial_{\mathcal{H}^4}E$ has empty interior also follows from the techniques of [21].

Lemma 11.8. *If E and U are as in Lemma 11.7, then $U\cap\partial_{\mathcal{H}^4}E$ has empty interior.*

Proof. The measure-theoretic version of [21, Lemma 4.12], whose proof appears in (part (4) of) the proof of [21, Theorem 5.1], asserts that, if $F\subseteq\mathbb{H}$ is monotone on $\mathbb{H}$, then $\partial_{\mathcal{H}^4}F\neq\mathbb{H}$. That proof relies on the monotonicity of a configuration of line segments, and it directly shows that there is a large enough universal constant $r>0$ such that this configuration lies in the ball $B_r(\mathbf{0})$. Consequently, if $B_r(\mathbf{0})\subseteq U$, then there is a point $p\in B_r(\mathbf{0})$ such that $p\notin\partial_{\mathcal{H}^4}E$. By rescaling and translation, this is true with $B_r(\mathbf{0})$ replaced by an arbitrary ball, and thus $\mathrm{int}_{\mathcal{H}^4}(E)\cup\mathrm{int}_{\mathcal{H}^4}(\mathbb{R}\setminus E)$ is dense in U. □

Lemma 11.8 proves part (1) of Proposition 11.4. Lemmas 11.6 and 11.7 imply the first half of part (2) of Proposition 11.4. Before proving the rest of Proposition 11.4, we make the following definition.

Definition 11.9. Let $U\subseteq\mathbb{H}$ be a convex open set and let $A\subseteq\mathbb{H}$. We say that A is *U-ruled* if, for all $L\in\mathcal{L}$, if $L\cap U$ intersects A in two points, then $L\cap U\subseteq A$. We call such a line L a *U-ruling* of A.

Lemmas 11.6–11.8 imply that $U\cap\partial_{\mathcal{H}^4}E$ is U-ruled and has empty interior. We will prove the rest of Proposition 11.4 by studying lines in the boundary of such a set. The following lemma is based on Step B3 in [23, §8.2], which shows that the boundary of a monotone set cannot contain skew lines.

Lemma 11.10. *Let M_1 be the line $\langle X\rangle$ and let M_2 be the line $Z\langle Y\rangle$. There exists $r_0>1$ such that any $\overline{B}_{r_0}$-ruled set containing $(M_1\cup M_2)\cap\overline{B}_{r_0}$ has non-empty interior.*

Proof. Let r_0 be large enough that $[-2,2]^3\subseteq\overline{B}_{r_0}$. Let E be a $\overline{B}_{r_0}$-ruled set with $\overline{B}_{r_0}$-rulings $M_1,M_2\in\mathcal{L}$. By Lemma 11.1, there is a hyperbola $S\subseteq\mathsf{H}$, asymptotic to the x-axis and the y-axis, such that every tangent line of S has a unique horizontal lift that intersects M_1 and M_2. Indeed, for every $t\neq 0$, the points $X^t\in M_1$ and $ZY^{2/t}\in M_2$ are connected by a horizontal line

$$L_t(u)\stackrel{\mathrm{def}}{=}X^t\left(-t,\frac{2}{t},0\right)^u=\left((1-u)t,\frac{2u}{t},u\right)\quad\text{for all } u\in\mathbb{R}.$$

For $t\in[-2,-1]\cup[1,2]$ and $u\in[0,1]$, the point $L_t(u)$ lies on a horizontal line segment connecting two points in E, so $L_t(u)\in E$. The resulting family of points

$$S\stackrel{\mathrm{def}}{=}\{L_t(u):t\in[-2,-1]\cup[1,2],u\in[0,1]\}\subseteq E$$

consists of two disjoint embedded surfaces.

Let

$$w \overset{\text{def}}{=} L_{\sqrt{2}}\left(\frac{1}{2}\right) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, \frac{1}{2}\right),$$

and let $w'=s_{-1,-1}(w)=L_{-\sqrt{2}}\left(\frac{1}{2}\right)$. Let M be the horizontal line from w to w'. Then M intersects S twice, at w and w', so $M\cap\overline{B}_{r_0}\subseteq E$. One calculates

$$\frac{d}{dt}L_t(u)\bigg|_{(t,u)=(\sqrt{2},1/2)} = (1-u,-2ut^{-2},0)|_{(t,u)=(\sqrt{2},1/2)} = \left(\frac{1}{2},-\frac{1}{2},0\right)$$

$$\frac{d}{du}L_t(u)\bigg|_{(t,u)=(\sqrt{2},1/2)} = (-t,2t^{-1},1)|_{(t,u)=(\sqrt{2},1/2)} = (-\sqrt{2},\sqrt{2},1),$$

so M intersects S transversally at w and w'. By transversality, any horizontal line M' close to M intersects S near w and w', so $M'\cap\overline{B}_{r_0}\subseteq E$. These lines cover a neighborhood of M, so E contains a non-empty open set. □

As shown in the next lemma, for any pair of skew lines, there is an automorphism of $\mathbb{H}$ that sends them to M_1 and M_2. The next lemma uses this fact to show that nearby skew lines in $\partial_{\mathcal{H}^4}E$ must have nearly parallel projections. For $\phi\in\mathbb{R}$, let $R_\phi:\mathbb{H}\to\mathbb{H}$ be the rotation by angle ϕ around the z-axis.

Lemma 11.11. *Let r_0 be as in Lemma 11.10. Let $L_1,L_2\in\mathcal{L}$ be skew lines and let $p\in\mathsf{H}$ be the intersection of $\pi(L_1)$ and $\pi(L_2)$. Suppose that the angle between $\pi(L_1)$ and $\pi(L_2)$ is $\theta\in\left(0,\frac{1}{2}\pi\right)$. For $i\in\{1,2\}$, let $q_i\in\mathbb{H}$ be the point where $\pi^{-1}(p)$ intersects L_i. Suppose that*

$$d(q_1,q_2)\leqslant\frac{\sqrt{\theta}}{r_0\sqrt{2}}. \tag{11.4}$$

If $L_1,L_2\in\mathcal{L}$ are $\overline{B}_1(q_1)$-rulings of an $\overline{B}_1(q_1)$-ruled set S, then S has non-empty interior.

Proof. After applying a translation and rotation, and possibly replacing S by $s_{1,-1}(S)$, we may suppose that $q_1=\mathbf{0}$, $q_2=Z^h$ for some $h>0$, and that $\pi(L_1)$ and $\pi(L_2)$ form angles of $\frac{1}{2}\theta$ with the x-axis. (We cannot control which line forms a positive angle with the x-axis and which line forms a negative angle.) Let $t=\tan\frac{1}{2}\theta\in(0,1)$ so that the lines

$$\pi(s_{\sqrt{t},1/\sqrt{t}}(L_1)) \quad\text{and}\quad \pi(s_{\sqrt{t},1/\sqrt{t}}(L_2))$$

are perpendicular. There is an angle $\phi=\pm\frac{1}{4}\pi$ such that, if

$$f\overset{\text{def}}{=}R_\phi\circ s_{1/\sqrt{h},1/\sqrt{h}}\circ s_{\sqrt{t},1/\sqrt{t}},$$

then $f(L_1)=M_1$ and $f(L_2)=M_2$, where M_1 and M_2 are the lines in Lemma 11.10. Now, by the ball-box inequality and our hypothesis on $d(q_1,q_2)$, we have

$$\operatorname{Lip}(f^{-1})=\frac{\sqrt{h}}{\sqrt{t}}\overset{(2.6)}{\leqslant}\frac{d(q_1,q_2)}{\sqrt{\tan\frac{1}{2}\theta}}\leqslant\frac{d(q_1,q_2)}{\sqrt{\frac{1}{2}\theta}}\overset{(11.4)}{\leqslant}\frac{1}{r_0}.$$

Thus, $f^{-1}(\overline{B}_{r_0}) \subseteq \overline{B}_{r_0 \operatorname{Lip}(f^{-1})} \subseteq \overline{B}_1$, or $\overline{B}_{r_0} \subseteq f(\overline{B}_1)$. Since $f(S)$ is a $f(\overline{B}_1)$-ruled set, and M_1 and M_2 are $f(\overline{B}_1)$-rulings of $f(S)$, by Lemma 11.10, $f(S)$ has non-empty interior, and thus S has non-empty interior. □

It follows from Lemmas 11.8 and 11.11 that two lines in $\partial_{\mathcal{H}^4} E$ with different angles must either intersect or stay at least a definite distance apart. In the terminology of [23], every pair of rulings of $\partial_{\mathcal{H}^4} E$ must form a degenerate initial condition.

Lemma 11.12. *For any $\varepsilon > 0$, there is $\delta > 0$ such that, if S is a $\overline{B}_1$-ruled set with empty interior, and L_1 and L_2 are $\overline{B}_1$-rulings of S that intersect $\overline{B}_\delta$ and such that*

$$\angle(\pi(L_1), \pi(L_2)) > \varepsilon,$$

then L_1 and L_2 intersect.

Proof. We suppose that $0 < \varepsilon < 1$ and take

$$\delta = \frac{\varepsilon^{3/2}}{100 r_0} \leqslant \frac{1}{100},$$

where r_0 is as in Lemma 11.10.

Let $p \in \mathsf{H}$ be the intersection of the projections $\pi(L_1)$ and $\pi(L_2)$. Since $\pi(\overline{B}_\delta)$ is the ball B_δ^{H} of radius δ in H, the projections intersect B_δ^{H} and form an angle of at least ε, so

$$\|p\| \leqslant \frac{\delta}{\sin \frac{1}{2}\varepsilon} \leqslant \frac{4\delta}{\varepsilon} < \frac{1}{4}.$$

For $i \in \{1, 2\}$, let $q_i = \pi^{-1}(p) \cap L_i$. By assumption, L_1 and L_2 intersect $\overline{B}_\delta \subseteq B_{2\delta}$, so if $b_i \in L_i \cap \overline{B}_\delta$, then

$$\begin{aligned} d(\mathbf{0}, q_i) &\leqslant d(\mathbf{0}, b_i) + d(b_i, q_i) = d(\mathbf{0}, b_i) + \|\pi(b_i) - \pi(q_i)\| \\ &\leqslant d(\mathbf{0}, b_i) + \|\pi(b_i)\| + \|p\| \leqslant 3\delta + \|p\|. \end{aligned}$$

In particular, $d(\mathbf{0}, q_i) \leqslant \frac{1}{2}$. Hence, $\overline{B}_{1/2}(q_1) \subseteq \overline{B}_1$, so S is a $\overline{B}_{1/2}(q_1)$-ruled set. Further,

$$d(q_1, q_2) \leqslant 2\|p\| + 6\delta < \frac{20\delta}{\varepsilon} \leqslant \frac{\sqrt{\varepsilon}}{5 r_0}.$$

Because S has empty interior, Lemma 11.11 implies that L_1 and L_2 cannot be skew lines, and must therefore intersect. □

The next lemma completes the proof of part (2) of Proposition 11.4.

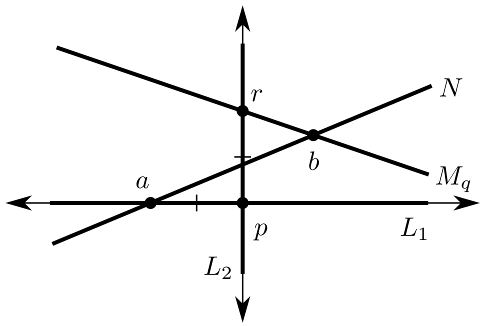

Figure 4. If line M_q intersects the y-axis L_2 but not the x-axis L_1, there must be a line N intersecting L_1 and M_q as seen above. Lines above are projected to H by π.

Lemma 11.13. *Suppose that U is a convex open set and that $E\subseteq\mathbb{H}$ is monotone on U. Let $p\in U$, and let L_1 and L_2 be two distinct U-rulings of $\partial_{\mathcal{H}^4}E$ that intersect at p. Then, $U\cap\mathsf{H}_p\subseteq\partial_{\mathcal{H}^4}E$, and there is a neighborhood A containing p such that*

$$A\cap\partial_{\mathcal{H}^4}E=A\cap\mathsf{H}_p,$$

where we recall that H_p denotes the horizontal plane through p.

Proof. Since U is convex, $\partial_{\mathcal{H}^4}E$ is U-ruled. After translating and applying an automorphism, we may suppose that $p=\mathbf{0}$, and that L_1 and L_2 are the x-axis and y-axis, respectively. Set $\varepsilon=\frac{1}{40}$ and let $\delta>0$ satisfy Lemma 11.12. Suppose that $\overline{B}_\delta\subseteq U$.

Fix $q\in B_{\delta/8}\cap\partial_{\mathcal{H}^4}E$. By Lemma 11.7, $\partial_{\mathcal{H}^4}E$ has a U-ruling M_q that passes through q. We will show that M_q intersects both L_1 and L_2 and that any such line passes through p.

For any horizontal line L, let $\bar{L}=\pi(L)$. Either $\angle(\bar{L}_1,\overline{M}_q)\geqslant\frac{\pi}{4}$ or $\angle(\bar{L}_2,\overline{M}_q)\geqslant\frac{\pi}{4}$. Therefore, by Lemma 11.12, M_q intersects either L_1 or L_2. Suppose by way of contradiction that M_q intersects L_2 but not L_1. By Lemma 11.12, this implies that $\angle(\bar{L}_1,\overline{M}_q)\leqslant\varepsilon$. Let r be the intersection of $\overline{M}_q$ with $\bar{L}_2$ and let $t=d(p,r)>0$ (see Figure 4). Straightforward trigonometry shows that $t<\frac{1}{2}\delta$.

Let $a=pX^{-t}\in L_1$. By Lemma 11.1, there is a unique point $b\in M_q$ such that there is a horizontal line N that passes through a and b. Indeed, since r, p, a, and b are the vertices of a quadrilateral Q in $\mathbb{H}$ whose sides are horizontal lines, the projection $\pi(Q)$ has zero signed area. Since the triangle $\triangle\pi(p)\pi(r)\pi(a)$ has area $\frac{1}{2}t^2$, the triangle $\triangle\pi(b)\pi(r)\pi(a)$ must also have area $\frac{1}{2}t^2$, so $\pi(b)$ is the intersection of $\overline{M}_q$ with the line $\langle X+Y\rangle$. Because $\overline{M}_q$ has slope between $-\varepsilon$ and ε, this implies that $|\pi(b)-(t,t)|\leqslant 4\varepsilon t\leqslant\frac{1}{10}t$. In particular,

$$d(r,b)=|\pi(r)-\pi(b)|\leqslant 2t,\quad \angle(\bar{L}_1,\overline{N})>\varepsilon,\quad\text{and}\quad \angle(\bar{L}_2,\overline{N})>\varepsilon.$$

Then,

$$d(p,b)\leqslant d(p,r)+d(r,b)\leqslant 3t<\delta,$$

so $b \in \overline{B}_\delta$.

Since $a, b \in U \cap \partial_{\mathcal{H}^4} E$, N is a U-ruling of $\partial_{\mathcal{H}^4} E$. By Lemma 11.12 and the fact that $\angle(\bar{L}_2, \overline{N}) > \varepsilon$, N intersects L_2. That is, L_1, L_2, and N are three distinct lines in $\mathbb{H}$ that intersect pairwise. If three distinct lines intersect pairwise, then they must all intersect at the same point. Otherwise, their projections to H would contain a non-degenerate triangle that lifts to a horizontal closed curve in $\mathbb{H}$, but this is impossible since the signed area of the projection of a horizontal closed curve must vanish. But L_1 intersects N at a and intersects L_2 at p, where $d(p, a) = t > 0$ by construction. This is a contradiction, so M_q intersects L_1 and L_2. Since M_q, L_1, and L_2 are distinct lines that intersect pairwise, M_q must intersect L_1 and L_2 at p.

Hence, every point $q \in B_{\delta/2} \cap \partial_{\mathcal{H}^4} E$ lies on the horizontal plane H_p through p. The measure-theoretic boundary of E disconnects $B_{\delta/2}$, so

$$B_{\delta/2} \cap \partial_{\mathcal{H}^4} E = B_{\delta/2} \cap \mathsf{H}_p.$$

Consequently, any line L through p intersects $U \cap \partial_{\mathcal{H}^4} E$ in at least two points, so $U \cap L \subseteq \partial_{\mathcal{H}^4} E$. The union of all such lines is H_p, so $U \cap \mathsf{H}_p \subseteq \partial_{\mathcal{H}^4} E$ □

Finally, we prove parts (3) and (4) of Proposition 11.4.

Proof of parts (3) *and* (4) *of Proposition* 11.4. Due to Lemma 11.13, if p is a characteristic point, then we have that $\partial_{\mathcal{H}^4} E$ has a horizontal approximate tangent plane at p. Lemma 11.13 also implies that, if p is a characteristic point, then there is a ball B such that B contains no characteristic points other than p. That is, the characteristic points form a discrete subset of $\mathbb{H}$; since $\mathbb{H}$ is separable, there are only countably many characteristic points.

Let $p \in U \cap \partial_{\mathcal{H}^4} E$ be a non-characteristic point such that there is a unique line L through p. Let V be the vertical plane that contains L. Fix $0 < \varepsilon < \frac{1}{10}$. We claim that there is $r > 0$ such that, if $0 < \alpha \leqslant r$, then $\overline{B}_\alpha(p) \cap \partial_{\mathcal{H}^4} E$ is contained in the $\varepsilon\alpha$-neighborhood of V.

We translate, rotate, and rescale so that $p = \mathbf{0}$, L is the x-axis, and $\overline{B}_1$ is a subset of U that contains no characteristic points. Then, $V = V_0$ is the xz-plane. Let $\Pi\colon \mathbb{H} \to V_0$ be the projection to V_0 along cosets of $\langle Y \rangle$, as in §2.2, so that $\Pi(x, y, z) = \left(x, 0, z - \frac{1}{2} xy\right)$.

For each point $s \in \overline{B}_1 \cap \partial_{\mathcal{H}^4} E$, there is a unique U-ruling M_s passing through s. By Lemma 11.12, there is $\delta \in (0,1)$ such that $\angle(M_s, L) < \frac{\varepsilon^2}{200}$ for every $s \in \overline{B}_\delta \cap \partial_{\mathcal{H}^4} E$. Let $r = \min\left\{\delta, \frac{1}{80}\varepsilon\right\}$ and let $0 < \alpha \leqslant r$. Let $q \in \overline{B}_\alpha \cap \partial_\mu E$ and suppose by way of contradiction that $d(q, V_0) = |y(q)| > \varepsilon\alpha$. Without loss of generality, we may suppose that $y(q) > \varepsilon\alpha$.

Let $m \in \mathbb{R}$ be the slope of $\pi(M_q)$, so that $M_q = q \cdot \langle X + mY \rangle$. Let $\gamma(t) = q \cdot (X + mY)^t$ parameterize M_q. Then,

$$|m| = |\sin \angle(M_s, L)| < \frac{1}{200} \varepsilon^2.$$

Since $q\in\overline{B}_\alpha\subseteq B_{2\alpha}$, we have $\Pi(q)\in B_{4\alpha}$, and thus $|z(\Pi(q))|\leqslant 16\alpha^2$. By (2.17), for all $t\in\mathbb{R}$,

$$\frac{d}{dt}z(\Pi(\gamma(t)))=-y(\gamma(t))=-y(q)-mt.$$

Consequently,

$$z(\Pi(\gamma(t)))=z(q)-y(q)t-\tfrac{1}{2}mt^2 \quad \text{for all } t\in\mathbb{R}.$$

Letting $s=20\alpha/\varepsilon$, it follows that

$$z(\Pi(\gamma(s)))\leqslant 16\alpha^2-\alpha\varepsilon s+\tfrac{1}{200}\varepsilon^2\cdot\tfrac{1}{2}s^2\leqslant -3\alpha^2$$

and

$$z(\Pi(\gamma(-s)))\geqslant -16\alpha^2+\alpha\varepsilon s-\tfrac{1}{200}\varepsilon^2\cdot\tfrac{1}{2}s^2\geqslant 3\alpha^2.$$

So, there is t with $|t|<s\leqslant\frac{1}{4}$ and $z(\Pi(\gamma(t)))=0$, i.e., $\Pi(\gamma(t))\in L$. The coset $N=\gamma(t)\langle Y\rangle$ is thus a horizontal line that intersects M_q at $\gamma(t)$ and intersects L at $\Pi(\gamma(t))$. Since

$$d(\mathbf{0},\gamma(t))\leqslant d(\mathbf{0},q)+|t|\leqslant 2\alpha+\tfrac{1}{4}\leqslant\tfrac{1}{2}$$

and

$$d(\mathbf{0},\Pi(\gamma(t)))\leqslant 2d(\mathbf{0},\gamma(t))\leqslant 1,$$

$\gamma(t)$ and $\Pi(\gamma(t))$ belong to $\overline{B}_1\cap\partial_{\mathcal{H}^4}E$, so $N\cap\overline{B}_1\subseteq\partial_{\mathcal{H}^4}E$. Then, M_q and N are distinct U-rulings of $\partial_{\mathcal{H}^4}E$ passing through $\gamma(t)$, which contradicts the fact that there are no characteristic points in $\overline{B}_1$. Therefore, $d(q,V_0)\leqslant\varepsilon\alpha$ for all $q\in\overline{B}_\alpha\cap\partial_{\mathcal{H}^4}E$.

Let $T_p=V_0$, and let T_p^+ and T_p^- be the corresponding half-spaces. The argument above shows that, for any $0<\alpha\leqslant r$, the sets $W^\pm_{\varepsilon,\alpha}$ are disjoint from $\partial_{\mathcal{H}^4}E$, so each set is contained in either $\operatorname{int}_{\mathcal{H}^4}(E)$ or $\operatorname{int}_{\mathcal{H}^4}(\mathbb{H}\setminus E)$.

Consider $W^+_{\varepsilon,r}$ and $W^-_{\varepsilon,r}$. Every line sufficiently close to the y-axis intersects both of these sets, so if both are contained in $\operatorname{int}_{\mathcal{H}^4}(E)$, then by Proposition 11.5, $p\in\operatorname{int}_{\mathcal{H}^4}(E)$ as well. Likewise, if both are contained in $\operatorname{int}_{\mathcal{H}^4}(\mathbb{H}\setminus E)$, then $p\in\operatorname{int}_{\mathcal{H}^4}(\mathbb{H}\setminus E)$. Either of these conclusions is a contradiction, so one of $W^+_{\varepsilon,r}$ and $W^-_{\varepsilon,r}$ is contained in $\operatorname{int}_{\mathcal{H}^4}(E)$ and the other is contained in $\operatorname{int}_{\mathcal{H}^4}(E)$. If necessary, we switch T_p^+ and T_p^- so that

$$W^+_{\varepsilon,r}\subseteq\operatorname{int}_{\mathcal{H}^4}(E).$$

We claim that $W^+_{\varepsilon,\alpha}\subseteq\operatorname{int}_{\mathcal{H}^4}(E)$ for every $\alpha\in(0,r]$. Fix $0<\beta\leqslant r$ with $\frac{1}{2}\beta<\alpha<\beta$. Then, $W^+_{\varepsilon,\alpha}$ intersects $W^+_{\varepsilon,\beta}$, so if $W^+_{\varepsilon,\beta}\subseteq\operatorname{int}_{\mathcal{H}^4}(E)$, then $W^+_{\varepsilon,\alpha}\subseteq\operatorname{int}_{\mathcal{H}^4}(E)$ as well. By induction, $W^+_{\varepsilon,\alpha}\subseteq\operatorname{int}_{\mathcal{H}^4}(E)$ for $0<\alpha\leqslant r$. Likewise, $W^-_{\varepsilon,\alpha}\subseteq\operatorname{int}_{\mathcal{H}^4}(\mathbb{H}\setminus E)$ for $0<\alpha\leqslant r$. $\square$

11.3. Stability of extended monotone sets

Here we prove Proposition 10.1. We show that there are $\nu>0$ and $R>0$ such that, if E is a set that is (ν,R)-extended monotone on $\overline{B}_1$, then E is close to a half-space on $\overline{B}_1$. If $R'\geqslant R$ and $\nu' R'\leqslant \nu R$, then (ν',R')-extended monotonicity implies (ν,R)-extended monotonicity, so this implies the full proposition.

To prove this, it suffices to show that, if f is a $\overline{B}_1$-LEM function, then $f|_{\overline{B}_1}$ is the characteristic function of a half-space. Suppose that f is a weak limit of a sequence $(\mathbf{1}_{E_i})_i$, where $E_1,E_2,...\subseteq\mathbb{H}$ are sets such that E_i is $(1/i,i)$-extended monotone on $\overline{B}_1$. By Corollary 11.3, $f|_{\overline{B}_1}$ is the characteristic function of a locally monotone subset $F\subseteq\overline{B}_1$, but this result only uses the fact that each E_i is $(1/i)$-monotone on $\overline{B}_1$. In this section, we improve Corollary 11.3 by using the stronger hypothesis that the E_i's are extended monotone sets.

The first issue is that $\mathrm{ENM}_{E_i,R}(\overline{B}_1)$ only depends on the intersection of E_i with lines through $\overline{B}_1$. These lines do not cover all of $\mathbb{H}$, so a $\overline{B}_1$-LEM function need not take values in $\{0,1\}$ outside $\overline{B}_1$. The following lemma shows that it takes values in $\{0,1\}$ on lines that intersect the boundary of F transversally. For $p\in\mathbb{H}$ and $V\in\mathsf{H}$ a horizontal vector, the coset $p\langle V\rangle$ is a horizontal line. Let

$$p\langle V\rangle^+=\{pV^t:t>0\}\quad\text{and}\quad p\langle V\rangle^-=\{pV^t:t<0\}.$$

LEMMA 11.14. *Let f be a $\overline{B}_1$-LEM function and let $F=f^{-1}(1)\cap\overline{B}_1$ be the corresponding locally monotone set. Let $p\in\overline{B}_1\cap\partial_{\mathcal{H}^4}F$ be a point with a vertical approximate tangent plane T_p and let $V\in\mathsf{H}_p$ be a horizontal vector pointing into T_p^+. Then,*

$$p\langle V\rangle^+\subseteq\mathrm{int}_{\mathcal{H}^4}(f^{-1}(1))\quad\textit{and}\quad p\langle V\rangle^-\subseteq\mathrm{int}_{\mathcal{H}^4}(f^{-1}(0)).\tag{11.5}$$

Proof. Let $E_i\subseteq\mathbb{H}$ be a sequence of sets such that E_i is $(1/i,i)$-monotone on $\overline{B}_1$ and $\mathbf{1}_{E_i}$ converges weakly to f. Let $L=p\langle V\rangle$, $L^\pm=p\langle V\rangle^\pm$, and $\theta=\angle(V,T_p)$. Let $\varepsilon=\frac{1}{20}\theta$ and let $W^\pm_{\varepsilon,t}$ be as in Proposition 11.4. For $t>0$, $L^\pm$ intersects $W^\pm_{\varepsilon,t}$ in an interval of length at least $\frac{1}{2}t$.

Fix $t>0$ and let $q=pV^t$. For the first inclusion in (11.5), the goal is to demonstrate that $q\in\mathrm{int}_{\mathcal{H}^4}(f^{-1}(1))$. Let $0<\alpha<\frac{1}{2}t$ be a radius such that $\overline{B}_\alpha(p)\subseteq\overline{B}_1$, $W^+_{\varepsilon,\alpha}\subseteq F$ up to a null set, and $W^-_{\varepsilon,\alpha}\subseteq\mathbb{H}\setminus F$ up to a null set. For any $\delta>0$, let $\mathcal{K}_\delta\subseteq\mathcal{L}$ be the set of lines of the form $q'\langle V'\rangle$, where $q'\in\overline{B}_\delta(q)$ and $V'\in\mathsf{H}$ is a horizontal vector such that $\angle(V,V')<\delta$. For $K\in\mathcal{K}_\delta$, let $K^\pm=K\cap T_p^\pm$.

Since the lines $\mathcal{K}_\delta$ are all close to L, there is a δ depending on θ and α such that $0<\delta<\min\{\varepsilon,\alpha\}$ and every line $K\in\mathcal{K}_\delta$ intersects both $W^+_{\varepsilon,\alpha}$ and $W^-_{\varepsilon,\alpha}$ in intervals of length at least $\frac{1}{4}\alpha$. We claim that

$$\lim_{i\to\infty}\mathcal{H}^4((\mathbb{H}\setminus E_i)\cap\overline{B}_\delta(q))=0,$$

and thus that $f=1$ almost everywhere on $\overline{B}_\delta(q)$.

For each $i\in\mathbb{N}$ define

$$\mathcal{T}_i \stackrel{\mathrm{def}}{=} \left\{K\in\mathcal{K}_\delta : \mathcal{H}^1(K\cap\overline{B}_1\cap(E_i\triangle F)) < \tfrac{1}{8}\alpha\right\}.$$

By Fubini's theorem, for any measurable subset $A\subseteq\mathbb{H}$ and any horizontal vector $M\in\mathsf{H}$ that is not parallel to T_p, we have

$$\int_{T_p} \mathcal{H}^1(b\langle M\rangle\cap A)\sin(\angle(M,T_p))\,d\mathcal{H}^3(b) \asymp \mathcal{H}^4(A). \tag{11.6}$$

Therefore, $\lim_{i\to\infty}\mathcal{N}(\mathcal{T}_i)=\mathcal{N}(\mathcal{K}_\delta)$, and for almost every $K\in\mathcal{T}_i$,

$$\mathcal{H}^1(K^+\cap F\cap\overline{B}_\alpha(p)) \geqslant \mathcal{H}^1(K^+\cap W^+_{\varepsilon,\alpha}) > \tfrac{1}{4}\alpha.$$

By the definition of $\mathcal{T}_i$, this implies that

$$\mathcal{H}^1(K^+\cap E_i\cap\overline{B}_\alpha(p)) > \tfrac{1}{8}\alpha, \tag{11.7}$$

and likewise

$$\mathcal{H}^1(K^-\cap E_i^c\cap\overline{B}_\alpha(p)) > \tfrac{1}{8}\alpha. \tag{11.8}$$

Let

$$\mathcal{S}_i \stackrel{\mathrm{def}}{=} \{K\in\mathcal{T}_i : \mathcal{H}^1(K\cap\overline{B}_\delta(q)\cap(\mathbb{H}\setminus E_i)) > 0\}.$$

Suppose that $i\geqslant d(p,q)+2\delta+2\alpha$ and $K\in\mathcal{S}_i$. By (11.7), (11.8), and the definition of $\mathcal{S}_i$, there are disjoint intervals

$$I_1 = K^-\cap\overline{B}_\alpha(p), \quad I_2 = K^+\cap\overline{B}_\alpha(p), \quad \text{and} \quad I_3 = K\cap\overline{B}_\delta(q)$$

such that the following conditions hold:

- I_2 is between I_1 and I_3;
- $I_1\cup I_2\cup I_3$ has diameter at most i;
- $\mathcal{H}^1(I_1\cap(\mathbb{H}\setminus E_i))>\frac{1}{8}\alpha$;
- $\mathcal{H}^1(I_2\cap E_i)>\frac{1}{8}\alpha$;
- $\mathcal{H}^1(I_3\cap(\mathbb{H}\setminus E_i))>0$.

Lemma 8.2 implies that

$$\widehat{\omega}_{E_i,i}(\overline{B}_1,K) \geqslant \widehat{\omega}_{E_i,i}(\overline{B}_\alpha(p),K) \geqslant \tfrac{1}{2}\mathcal{H}^1(E_i\cap I_2) \geqslant \tfrac{1}{16}\alpha.$$

Hence,

$$\frac{\alpha}{16}\mathcal{N}(\mathcal{S}_i) \leqslant \int_{\mathcal{L}} \widehat{\omega}_{E_i,i}(\overline{B}_1,K)\,d\mathcal{N}(K) = \mathrm{ENM}_{E_i,i}(\overline{B}_1) \leqslant \frac{1}{i},$$

and so $\lim_{i\to\infty}\mathcal{N}(\mathcal{S}_i)=0$.

Let

$$\mathcal{R}_i \stackrel{\mathrm{def}}{=} \{K\in\mathcal{K}_\delta : \mathcal{H}^1(K\cap\overline{B}_\delta(q)\cap(\mathbb{H}\setminus E_i))>0\}.$$

Then $\mathcal{N}(\mathcal{R}_i)\leqslant\mathcal{N}(\mathcal{S}_i)+\mathcal{N}(\mathcal{K}_\delta\setminus\mathcal{T}_i)$, and so $\lim_{i\to\infty}\mathcal{N}(\mathcal{R}_i)=0$. By (11.6),

$$\mathcal{H}^4(\overline{B}_\delta(q)\cap(\mathbb{H}\setminus E_i))\asymp_\delta \int_{\mathcal{K}_\delta}\mathcal{H}^1(K\cap\overline{B}_\delta(q)\cap(\mathbb{H}\setminus E_i))\,d\mathcal{N}(K)\leqslant\int_{\mathcal{R}_i}2\delta\,d\mathcal{N}(K),$$

where the last inequality follows from the fact that $\mathcal{H}^1(K\cap\overline{B}_\delta(q))\leqslant 2\delta$ for any horizontal line K. We therefore conclude as follows:

$$\lim_{i\to\infty}\mathcal{H}^4(\overline{B}_\delta(q)\cap(\mathbb{H}\setminus E_i))\leqslant\lim_{i\to\infty}2\delta\mathcal{N}(\mathcal{R}_i)=0. \qquad \square$$

By Lemma 11.7, $\overline{B}_1\cap\partial_{\mathcal{H}^4}F$ is a union of line segments. Extended monotonicity implies that these line segments can be extended to lines.

Lemma 11.15. *Let f be a $\overline{B}_1$-LEM function and let $F=f^{-1}(1)\cap\overline{B}_1$ be the corresponding locally monotone set. Let L be a horizontal line. If an open subinterval $I\subseteq L$ is contained in $\overline{B}_1\cap\partial_{\mathcal{H}^4}F$, then $L\subseteq\partial_{\mathcal{H}^4}F$.*

Proof. By Proposition 11.4, $\partial_{\mathcal{H}^4}F$ has at most countably many characteristic points. Let $p\in I$ be non-characteristic. Then, the vertical plane T_p containing L is the approximate tangent plane to $\partial_{\mathcal{H}^4}F$ at p. Recalling that H_p is the horizontal plane centered at p, every horizontal line through p, other than L itself, intersects $\partial_{\mathcal{H}^4}F$ transversally at p, so by Lemma 11.14, we have

$$T_p^+\cap\mathsf{H}_p\subseteq\mathrm{int}_{\mathcal{H}^4}(F)\quad\text{and}\quad T_p^-\cap\mathsf{H}_p\subseteq\mathrm{int}_{\mathcal{H}^4}(\mathbb{H}\setminus F).$$

Since L lies in the closures of $T_p^+\cap\mathsf{H}_p$ and $T_p^-\cap\mathsf{H}_p$, we have

$$L\subseteq\mathrm{supp}_{\mathcal{H}^4}(F)\cap\mathrm{supp}_{\mathcal{H}^4}(\mathbb{H}\setminus F)=\partial_{\mathcal{H}^4}F. \qquad \square$$

Finally, we show that if $\overline{B}_1\cap\partial_{\mathcal{H}^4}F$ is non-planar, then we can construct an arrangement of lines that leads to a contradiction.

Lemma 11.16. *Let f be a $\overline{B}_1$-LEM function. There is a plane $Q\subseteq\mathbb{H}$ such that $f|_{\overline{B}_1}=\mathbf{1}_{Q^+}$ outside of a null set. In fact, the same holds true in a larger set. Let*

$$S\stackrel{\mathrm{def}}{=}(Q\cap\overline{B}_1)\mathsf{H} \tag{11.9}$$

be the union of the horizontal lines intersecting $Q\cap\overline{B}_1$. Then, we have $f|_S=\mathbf{1}_{Q^+}$ outside of a null set.

Proof. Let $F=f^{-1}(1)\cap\overline{B}_1$ be the locally monotone set corresponding to f and suppose, by way of contradiction, that $\overline{B}_1\cap\partial_{\mathcal{H}^4}F$ is non-planar. By part (2) of Proposition 11.4 and by Lemma 11.15, for every point $p\in\overline{B}_1\cap\partial_{\mathcal{H}^4}F$, there is a horizontal line M_p through p such that $M_p\subseteq\partial_{\mathcal{H}^4}F$.

Reasoning as in [21, Lemma 4.11] shows that there are two $\overline{B}_1$-rulings of F that satisfy one of the cases of Lemma 11.1, i.e., they are a pair of skew lines or a pair of lines with distinct parallel projections. Indeed, suppose that J and K are $\overline{B}_1$-rulings of F with parallel projections. If $\pi(J)\neq\pi(K)$, we are done; otherwise, J and K are contained in a vertical plane V. Let L be a $\overline{B}_1$-ruling of F not in V, which exists by the assumed non-planarity. Then, L is skew to J or K, or parallel to V with a distinct projection. It remains to treat the case when any two $\overline{B}_1$-rulings of F have non-parallel projections. Let J and K be two such rulings. If J and K are disjoint, we are done, so we suppose J and K intersect at a point p and are thus contained in the horizontal plane H_p centered at p. If L is a $\overline{B}_1$-ruling of F that is not contained in H_p (it exists by assumed non-planarity), then L intersects H_p at a single point other than p, so L is skew to either J or K, as desired.

This shows that there are two $\overline{B}_1$-rulings L_1 and L_2 of F that are skew or have distinct parallel projections. Let $I=L_1\cap\overline{B}_1$ and let $p\in I$ be a non-characteristic point such that $\pi(p)\notin\pi(L_2)$. By Lemma 11.1, there is a horizontal line M that goes through p and intersects L_2 at q. This line is not equal to L_1, so it intersects $\partial_{\mathcal{H}^4}F$ transversally at p. By Lemma 11.14, this implies that $q\in\operatorname{int}_{\mathcal{H}^4}(f^{-1}(0))$ or $q\in\operatorname{int}_{\mathcal{H}^4}(f^{-1}(1))$, but $q\in L_2\subseteq\partial_{\mathcal{H}^4}F$, which is a contradiction. Therefore, $\overline{B}_1\cap\partial_{\mathcal{H}^4}F$ is planar, and there is a plane Q such that $F\cap\overline{B}_1=Q^+\cap\overline{B}_1$ up to a null set. Since f takes values in $\{0,1\}$ inside $\overline{B}_1$, this implies the first part of Lemma 11.16.

With S as in (11.9), take $w\in Q^+\cap S$. Then w lies on a horizontal line that intersects $Q\cap\overline{B}_1$ transversally, and Lemma 11.14 implies that $w\in\operatorname{int}_{\mathcal{H}^4}(f^{-1}(1))$. It follows that $f=1$ almost everywhere in $Q^+\cap S$ and likewise that $f=0$ almost everywhere in $Q^-\cap S$. □

The second part of Proposition 10.1 states that extended monotone intrinsic graphs are close to vertical planes. This follows from the fact that neighborhoods of the center of a horizontal plane cannot be approximated by intrinsic graphs.

LEMMA 11.17. *Let V_0 be the xz-plane and let $E_1,E_2,\ldots\subseteq\mathbb{H}$ be a sequence of intrinsic graphs over V_0 such that E_i^+ is $(1/i,i)$-extended monotone on $\overline{B}_1$ and $\mathbf{1}_{E_i^+}$ converges weakly to a function $f\in L_\infty(\mathbb{H})$ as $i\to\infty$. There is a vertical plane $Q\subseteq\mathbb{H}$ such that $f|_{\overline{B}_1}=\mathbf{1}_{Q^+}$ outside of a null set. Furthermore, if S is as in (11.9), then $f|_S=\mathbf{1}_{Q^+}$ outside of a null set.*

Proof. For any intrinsic graph Γ and any $g\in\Gamma^+$, we have $gY^t\in\Gamma^+$ for every $t>0$.

As $\mathcal{H}^4$ is right-invariant, this implies that, for any measurable set $U\subseteq\mathbb{N}$ and any $i\in\mathbb{N}$,

$$\mathcal{H}^4(U\cap E_i^+)\leqslant\mathcal{H}^4(U\cap E_i^+Y^t).$$

Therefore,

$$\int_U f\,d\mathcal{H}^4\leqslant\int_{UY^t} f\,d\mathcal{H}^4.$$

Consequently,

$$f(g)\leqslant f(gY^t)\quad\text{for almost every }(g,t)\in\mathbb{H}\times(0,\infty).\tag{11.10}$$

If f is almost-surely constant on $\overline{B}_1$, we can take Q to be a vertical plane that does not intersect $\overline{B}_1$. We thus suppose that $f|_{\overline{B}_1}$ is not almost-surely constant. By Lemma 11.16, there is a plane Q that satisfies $f|_S=\mathbf{1}_{Q^+}$ outside of a null set, where S is given in (11.9).

Suppose for contradiction that Q is horizontal. Let $c\in\mathbb{H}$ be such that $Q=\mathsf{H}_c=c\mathsf{H}$ and let $p\in Q\cap\operatorname{int}(\overline{B}_1)$ be such that $x(p)\neq x(c)$. Let L be the horizontal line from c to p and let $V=(x_V,y_V,0)$ be the horizontal vector such that $p=cV$. Set

$$q=cV^{-1}=c(-x_V,-y_V,0).$$

We claim that there is $\varepsilon>0$ such that $\{pY^{\pm\varepsilon},qY^{\pm\varepsilon}\}\subseteq S$. Choose $\varepsilon>0$ so that $pY^t\in\overline{B}_1$ and $r_t=c(x_V,y_V+t,0)\in\overline{B}_1\cap Q$ for all $t\in[-2\varepsilon,2\varepsilon]$. Then,

$$\begin{aligned}r_t\big(-2x_V,-2y_V-\tfrac{3}{2}t,0\big)&=c(x_V,y_V+t,0)\big(-2x_V,-2y_V-\tfrac{3}{2}t,0\big)\\&=c\big(-x_V,-y_V-\tfrac{1}{2}t,\tfrac{1}{4}x_Vt\big)=qY^{-t/2}.\end{aligned}$$

It follows that $qY^{-t/2}\in r_t\mathsf{H}\subseteq S$. In particular, $qY^{\pm\varepsilon}\in S$. At the same time, pY^{ε} and $pY^{-\varepsilon}$ are on opposite sides of Q; equation (11.10) implies that $pY^{\varepsilon}\in Q^+$ and $pY^{-\varepsilon}\in Q^-$. Likewise, $qY^{\pm\varepsilon}\in Q^{\pm}$. But since c is between p and q, the points pY^{ε} and qY^{ε} are on opposite sides of Q, which is a contradiction. Therefore, Q is a vertical plane. □

Proof of Proposition 10.1. If the first part of the proposition were false, then there would exist $\varepsilon>0$ and a sequence of measurable sets $(E_i)_{i=1}^{\infty}$ such that, for any $i\in\mathbb{N}$, the set E_i is $(1/i,i)$-extended monotone on $\overline{B}_1$ and $|\overline{B}_1\cap(P^+\triangle E_i)|>\varepsilon$ for every plane $P\subseteq\mathbb{H}$. There is a subsequence $(E_{i(j)})_{j=1}^{\infty}$ whose characteristic functions converge weakly to a $\overline{B}_1$-LEM function f. By Lemma 11.16, there is a plane $Q\subseteq\mathbb{H}$ such that $f=\mathbf{1}_{Q^+}$ almost everywhere on $\overline{B}_1$. Then,

$$\lim_{j\to\infty}|\overline{B}_1\cap(Q^+\triangle E_{i(j)})|=0,$$

which is a contradiction.

Similarly, if the second part of the proposition were false, then there would exist $\varepsilon>0$ and a sequence of intrinsic graphs $(E_i)_{i=1}^{\infty}$ over V_0 such that, for any $i\in\mathbb{N}$, the epigraph E_i^+ is $(1/i,i)$-extended monotone on $\overline{B}_1$ and $|\overline{B}_1\cap(P^+\triangle E_i^+)|>\varepsilon$ for every vertical plane $P\subseteq\mathbb{H}$. Passing to a subsequence, we may suppose that the indicators $\mathbf{1}_{E_i^+}$ converge weakly to a $\overline{B}_1$-LEM function f. By Lemma 11.17, there is a vertical plane $Q\subseteq\mathbb{H}$ such that $f=\mathbf{1}_{Q^+}$ almost everywhere on $\overline{B}_1$. Then,

$$\lim_{i\to\infty}|\overline{B}_1\cap(Q^+\triangle E_i^+)|=0,$$

which is a contradiction. □

12. L_1 bounds and characteristic curves on monotone intrinsic graphs

Here we complete the proof of Proposition 7.2, which obtains L_1 bounds for paramonotone pseudoquads and bounds their characteristic curves.

Fix $0<\mu\leqslant\frac{1}{32}$ and a μ-rectilinear pseudoquad Q in an intrinsic Lipschitz graph $\Gamma=\Gamma_f$. Suppose that Γ is (η,R)-paramonotone on rQ. By Remark 4.3, we can normalize Q and Γ so that the corresponding parabolic rectangle is the square $[-1,1]\times\{0\}\times[-1,1]$; by Lemma 8.8 and the discussion immediately after its proof, the normalized pseudoquad remains paramonotone. So, it suffices to prove Proposition 7.2 for such pseudoquads.

For $t>0$, denote

$$D_t=[-t,t]\times\{0\}\times[-t^2,t^2]\subseteq V_0.$$

By our choice of normalization, we have $tQ=D_t$. Furthermore, $D_t\subseteq B_{5t}$ and $\Pi(\overline{B}_t)\subseteq D_t$. We will proceed in several steps.

(1) First, we will prove in Lemma 12.2 that there is a universal constant $\kappa>0$ such that $\|f\|_{L_1(Q)}\leqslant\kappa$ when η is sufficiently small. This relies on Lemma 12.1 that bounds the tails of f in regions that are bounded above and below by supercharacteristic curves (projections of horizontal curves in $\Gamma\cup\Gamma^+$).

(2) Next, we will show that Γ is close to a plane on a ball around the origin. Since $\|f\|_{L_1(Q)}\leqslant\kappa$, the intersections $\Gamma^+\cap B_\kappa$ and $\Gamma^-\cap B_\kappa$ both have positive measure. For any $r>0$, we have $\Pi(\overline{B}_r)\subseteq rQ$, so $\mathrm{ENM}_{\Gamma^+,R}(\overline{B}_r)\lesssim\eta R$. When ηR is sufficiently small, and r and R are sufficiently large, Proposition 10.1 implies that there is a vertical plane P that intersects B_κ and approximates Γ on B_r, i.e.,

$$\mathcal{H}^4((\Gamma^+\triangle P^+)\cap\overline{B}_r)<\varepsilon.$$

Furthermore, since $\|f\|_{L_1(Q)}\leqslant\kappa$, the slope and y-intercept of $\pi(P)$ are both at most some universal constant.

We then apply an automorphism that sends P to V_0. Since the slope and y-intercept of P are bounded, there is a universal constant $c>0$ and a map $q:\mathbb{H}\to\mathbb{H}$ (a composition of a left translation in the y–direction and a shear) such that $q(P)=V_0$ and

$$B_{c^{-1}s-c}\subseteq q(B_s)\subseteq B_{cs+c}$$

for all $s>c^2$. We let

$$\widehat{\Gamma}=q(\Gamma)\quad\text{and}\quad \widehat{Q}=\hat{q}(Q)=\Pi(q(Q)),$$

and let $\hat{f}$ be such that $\widehat{\Gamma}=\Gamma_{\hat{f}}$. Since q preserves $\mathcal{H}^4$, we have

$$\mathcal{H}^4((\widehat{\Gamma}^+\triangle V_0^+)\cap\overline{B}_{c^{-1}r-c})<\varepsilon. \tag{12.1}$$

This inequality controls $\hat{f}$ on $V_0\cap\overline{B}_{c^{-1}r-c}$, and we choose r large enough that

$$11\widehat{Q}\subseteq\overline{B}_{c^{-1}r-c}.$$

(3) By (12.1),

$$\int_{10\widehat{Q}}\min\{1,|\hat{f}(p)|\}\,d\mathcal{H}^3(p)\leqslant\varepsilon,$$

so a bound on the tails of $\hat{f}$ would lead to a bound on $\|\hat{f}\|_{L_1(10\widehat{Q})}$. We bound the tails in Lemma 12.4, by finding supercharacteristic curves above and below $10\widehat{Q}$, then applying Lemma 12.1 again. This implies that $\|\hat{f}\|_{L_1(10\widehat{Q})}\lesssim\varepsilon$ when η is sufficiently small, which proves the first part of Proposition 7.2.

(4) Finally, we bound the characteristic curves of $\widehat{\Gamma}$ in Lemma 12.6, by showing that, if $\widehat{\Gamma}$ contains characteristic curves that are not nearly parallel to the x-axis, then either $\|\hat{f}\|_{L_1}$ is bounded away from zero, or $\Omega^P_{\Gamma^+,R}$ is bounded away from zero. This completes the proof of Proposition 7.2.

We will use the following notation for horizontal lines. Every horizontal line in $\mathcal{L}_P$ can be written uniquely as follows for some $w=(0,y_0,z_0)\in\mathbb{H}$ and $m\in\mathbb{R}$:

$$L_{w,m}\stackrel{\mathrm{def}}{=}w\langle X+mY\rangle.$$

Let $\rho_{L_{w,m}}:\mathbb{R}\to L_{w,m}$ be the following parametrization, so that $x(\rho_L(t))=t$ for all $t\in\mathbb{R}$:

$$\rho_{L_{w,m}}(t)\stackrel{\mathrm{def}}{=}w(X+mY)^t\quad\text{for all } t\in\mathbb{R}.$$

For every $x\in\mathbb{R}$ define

$$g_{L_{w,m}}(x)\stackrel{\mathrm{def}}{=}z(\Pi(\rho_{L_{w,m}}(x)))=-\tfrac{1}{2}mx^2-y_0x+z_0. \tag{12.2}$$

Note that, since $L_{w,m}$ is horizontal, we have $y(\rho_{L_{w,m}}(x))=-g'_{L_{w,m}}(x)$.

12.1. Bounding the tails of f

We start by showing that, if Q is a rectilinear pseudoquad for $\Gamma=\Gamma_f$ such that Γ^+ is (η, R)-paramonotone on rQ, as in Proposition 7.2, and Q is normalized so that the corresponding parabolic rectangle is a 2×2 square, as in Remark 4.3, then there is a universal constant κ such that $\|f\|_{L_1(Q)}\leqslant\kappa$ when r and R are sufficiently large and η is sufficiently small.

This step relies on the following lemma, which will also be used in step 3. A *supercharacteristic curve* (resp. *subcharacteristic curve*) for Γ is the projection $\Pi(\gamma)$ of a horizontal curve $\gamma: I\to\mathbb{H}$ such that $x(\gamma(t))=t$ for all $t\in I$ and $\gamma(I)\subseteq\Gamma\cup\Gamma^+$ (resp. $\gamma(I)\subseteq\Gamma\cup\Gamma^-$).

Such a curve can be written as a graph of the form $\{z=g(x)\}\subseteq V_0$. By the argument of Lemma 2.6, g is differentiable almost everywhere and satisfies $g'(x)=y(\gamma(x))$ for almost every $x\in I$. Since g is locally Lipschitz, we have

$$g(x)=g(x_0)+\int_{x_0}^{x} g'(t)\,dt=g(x_0)+\int_{x_0}^{x} y(\gamma(t))\,dt \quad \text{for all } x, x_0\in I,$$

and therefore $g'(x)=y(\gamma(x))$ for *every* $x\in I$. In particular, $g'(x)\leqslant -f(x,0,g(x))$ for all $x\in I$. We then say that g is a *function with supercharacteristic graph.*

LEMMA 12.1. *Let $g_1, g_2: [-2,2]\to\mathbb{R}$ be functions with supercharacteristic graphs such that* $\sup g_1([-2,2])<\inf g_2([-2,2])$. *For* $0\leqslant r\leqslant 2$, *let*

$$U_r=\{(x,0,z)\in V_0: |x|\leqslant r \textit{ and } g_1(x)\leqslant z\leqslant g_2(x)\}.$$

Denoting $H=\max\{\|g_1\|_{L_\infty([-2,2])}, \|g_2\|_{L_\infty([-2,2])}\}$, *for any* $t\geqslant 8H$ *we have*

$$|\{v\in U_1: f(v)\geqslant t\}|\lesssim \frac{1}{t^2}\Omega^P_{\Gamma^+,4}(U_2). \tag{12.3}$$

Likewise, if $g_1, g_2: [-2,2]\to\mathbb{R}$ *have subcharacteristic graphs, and* U_r *and* H *are as above, then for any* $t\geqslant 8H$ *we have*

$$|\{v\in U_1: f(v)\leqslant -t\}|\lesssim \frac{1}{t^2}\Omega^P_{\Gamma^+,4}(U_2).$$

Once we prove Lemma 12.1, we will apply it to the case that Q approximates $[-1,1]^2$ and g_1 and g_2 are the lower and upper bounds of Q.

Proof. Fix $t\geqslant 8H$ and $y_0, m, z_0\in\mathbb{R}$ such that

$$\left|y_0-\tfrac{1}{2}t\right|<\tfrac{1}{12}t \quad\text{and}\quad |m|<\tfrac{1}{12}t.$$

Let $L=L_{(0,y_0,z_0),m}$. For any $s\in[-2,2]$ we have

$$\left|g_L'(s)+\tfrac{1}{2}t\right|=\left|y(\rho_L(s))-\tfrac{1}{2}t\right|<\tfrac{1}{4}t, \tag{12.4}$$

so $-\frac{3}{4}t<g_L'(s)<-\frac{1}{4}t$ on $[-2,2]$.

We claim that for almost every such L we have

$$\widehat{\omega}^P_{\Gamma^+,4}(U_2,L)\geqslant\tfrac{1}{2}\mathcal{H}^1(x(\Gamma^-\cap L\cap\Pi^{-1}(U_1))). \tag{12.5}$$

By (12.4), we have

$$g_L(-2)=g_L(s)-\int_{-2}^s g_L'(u)\,du>-H+(s+2)\frac{t}{4}\geqslant -H+2H=H,$$

and

$$g_L(2)=g_L(s)+\int_s^2 g_L'(u)\,du<H-(2-s)\frac{t}{4}\leqslant H-2H=-H.$$

Hence, $\Pi(L)$ crosses U_2 negatively (from top to bottom), as depicted in Figure 12.1. The curve $\Pi(L)$ only intersects the top and bottom of U_2, not the sides, so we say that $\Pi(L)$ is transverse to the boundary of U_2 if $\Pi(L)$ intersects the top and bottom boundaries transversally; that is, if $g_L(u)=g_i(u)$ for some $u\in[-2,2]$ and $i=1,2$, then $g_L'(u)\neq g_i'(u)$.

Suppose that $\Pi(L)$ is transverse to the boundary of U_2 and that $L\cap\Gamma^+$ has finite perimeter; these are true for almost every L. If $\Pi(L)$ does not intersect U_1, then the right-hand side of (12.5) is zero and the inequality holds trivially. We thus suppose in addition that L intersects U_1. In this case, there is some $s\in[-1,1]$ such that $|g_L(s)|\leqslant H$.

Fix $i\in\{1,2\}$ and suppose that $\Pi(L)$ crosses the graph of g_i negatively at $(u,0,g_L(u))$. Let $v=\rho_L(u)$ be the point on L over the intersection. Then

$$g_L(u)=g_i(u)\quad\text{and}\quad g_L'(u)<g_i'(u).$$

Since the graph of g_i is supercharacteristic, $f(u,0,g_i(u))\leqslant -g_i'(u)$, and therefore

$$y(v)=-g_L'(u)>-g_i'(u)\geqslant f(u,0,g_i(u))=f(\Pi(v)).$$

That is, $v\in\Gamma^+$.

Since $\Pi(L)$ is transverse to the boundary of U_2, the intersection $\Pi(L)\cap U_2$ consists of a collection of intervals. Let $[a_1,b_1],...,[a_n,b_n]\subseteq\mathbb{R}$ be the disjoint intervals such that

$$x(\Pi(L)\cap U_2)=[a_1,b_1]\cup...\cup[a_n,b_n],$$

and these intervals are in ascending order. The projection $\Pi(L)$ does not intersect the left or right boundary of U_2, so $\Pi(L)$ crosses the graph of g_1 or g_2 at each a_i or b_i. Since

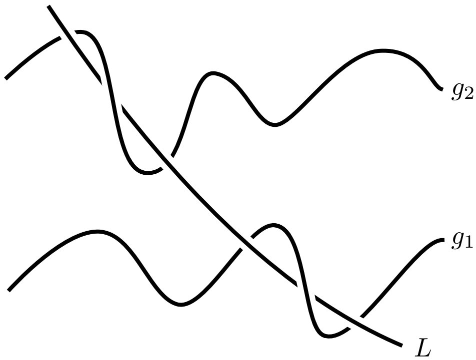

Figure 5. Two characteristic curves g_1 and g_2 and a horizontal line L, projected to V_0; the positive y-axis points toward the reader. Since $\Pi(L)$ crosses U_2 negatively, the segments of L at the first and last crossings lie in Γ^+, and therefore the size of the intersection $L\cap\Gamma^-$ is bounded by $\widehat{\omega}^P_{\Gamma^+,4}(U_2, LZ^t)$.

g_L is decreasing and $\sup g_1([-2,2])<\inf g_2([-2,2])$, the crossings of g_2 all have smaller x-coordinate than the crossings of g_1.

Consider $S=x(L\cap\Gamma^+)$. Since $\Pi(L)$ crosses the graph of g_2 negatively at a_1 and crosses the graph of g_1 negatively at b_n, the argument above implies that $a_1, b_n\in S$. Furthermore, for each $i\in\{1,...,n\}$, one of following three cases holds.

(1) $\Pi(L)$ crosses the graph of g_2 negatively at a_i and positively (from bottom to top) at b_i.

(2) $\Pi(L)$ crosses the graph of g_2 negatively at a_i and crosses the graph of g_1 negatively at b_i.

(3) $\Pi(L)$ crosses the graph of g_1 positively at a_i and negatively at b_i.

In each case, $a_i\in S$ or $b_i\in S$. By Lemma 8.2 (applied with $[a,b]=[a_1,b_n]$),

$$\widehat{\omega}_{S,4}([a_i,b_i])=\widehat{\omega}_{\mathbb{R}\setminus S,4}([a_i,b_i])\geqslant\frac{1}{2}\mathcal{H}^1(x(\Gamma^-\cap L)\cap[a_i,b_i]).$$

Summing over $i\in\{1,...,n\}$, we find that

$$\widehat{\omega}_{S,4}\left(\bigcup_{i=1}^n[a_i,b_i]\right)=\widehat{\omega}^P_{\Gamma^+,4}(U_2,L)\geqslant\frac{1}{2}\mathcal{H}^1(x(\Gamma^-\cap L\cap\Pi^{-1}(U_2))). \tag{12.6}$$

This proves (12.5).

Next, let

$$A=U_1\cap f^{-1}([t,\infty)).$$

By (12.4), $y(\rho_L(s))<t$ for all $s\in[-2,2]$, so if $\Pi(\rho_L(s))\in A$, then $\rho_L(s)\in\Gamma^-$. So, by (12.5),

$$\tfrac{1}{2}\mathcal{H}^1(x(\Pi(L)\cap A))\leqslant\tfrac{1}{2}\mathcal{H}^1(x(\Gamma^-\cap L\cap\Pi^{-1}(U_1)))\leqslant\widehat{\omega}^P_{\Gamma^+,4}(U_2,L).$$

By Fubini's Theorem, for any y_0 and m as above,

$$\frac{1}{2}|A| = \frac{1}{2}\int_{\mathbb{R}} \mathcal{H}^1(x(L_{(0,y_0,z_0),m} \cap A))\, dz_0 \leqslant \int_{\mathbb{R}} \widehat{\omega}^P_{\Gamma^+,4}(U_2, L_{(0,y_0,z_0),m})\, dz_0.$$

Therefore, recalling the definition (8.6) of Ω^P, we have

$$\begin{aligned}\Omega^P_{\Gamma^+,4}(U_2) &= \frac{1}{4}\int_{\mathcal{L}} \widehat{\omega}^P_{\Gamma^+,4}(U_2, L)\, d\mathcal{N}_P(L)\\ &\geqslant \frac{1}{4}\int_{-t/12}^{t/12}\int_{5t/12}^{7t/12}\int_{\mathbb{R}} \widehat{\omega}^P_{\Gamma^+,4}(U_2, L_{(0,y_0,z_0),m})\, dz_0\, dy_0\, dm\\ &\geqslant \frac{1}{8}\int_{-t/12}^{t/12}\int_{5t/12}^{7t/12} |A|\, dy_0\, dm = \frac{t^2}{288}|A|.\end{aligned}$$

That is,

$$|\{v \in U_1 : f(v) \geqslant t\}| \lesssim \frac{1}{t^2}\Omega^P_{\Gamma^+,4}(U_2).$$

This proves (12.3).

We can show that

$$|\{v \in U_1 : f(v) \leqslant -t\}| \lesssim \frac{1}{t^2}\Omega^P_{\Gamma^+,4}(U_2).$$

when g_1 and g_2 have subcharacteristic graphs by either applying a similar argument or by replacing Γ, U_r, etc. by $s_{1,-1}(\Gamma)$, $s_{1,-1}(U_r)$, etc. □

The desired bound on $\|f\|_{L_1(Q)}$ follows by integrating (12.3) with respect to t.

LEMMA 12.2. *Let $f\colon V_0 \to \mathbb{R}$ be a continuous function and let Γ be its intrinsic graph. Let $(Q, [-1,1]\times\{0\}\times[-1,1])$ be a $\frac{1}{32}$-rectilinear pseudoquad for Γ. Suppose that $\Omega^P_{\Gamma^+,4}(2Q) \leqslant 1$. There is a universal constant $\kappa > 0$ such that $\|f\|_{L_1(Q)} \leqslant \kappa$.*

Proof. Let g_1 and g_2 be the lower and upper bounds of Q and for $0 \leqslant r \leqslant 2$, let U_r be as in Lemma 12.1. Then $Q = U_1$ and $U_2 \subseteq 2Q$. Let $H = 2$. Since the graphs of g_1 and g_2 are supercharacteristic and $U_2 \subseteq 2Q$, Lemma 12.1 implies that, for any $t \geqslant 16$,

$$|\{v \in Q : f(v) \geqslant t\}| \lesssim t^{-2}\Omega^P_{\Gamma^+,4}(U_2) \leqslant t^{-2}\Omega^P_{\Gamma^+,4}(2Q) \leqslant t^{-2}.$$

Since the graphs of g_1 and g_2 are also subcharacteristic, for any $t \geqslant 16$ we also have

$$|\{v \in U_1 : f(v) \leqslant -t\}| \lesssim t^{-2}.$$

Then

$$\|f\|_{L_1(Q)} = \int_0^\infty \big|\{v \in Q : |f(v)| \geqslant t\}\big|\, dt \lesssim 16|Q| + \int_{16}^\infty t^{-2}\, dt \lesssim 1.$$ □

12.2. Constructing the approximating plane

Now we will use Lemma 12.2 and the results of §11 to show that, if Q is a paramonotone pseudoquad for Γ_f, then f is close on Q to an affine function with bounded coefficients.

LEMMA 12.3. *Let $\kappa>0$ be the constant in Lemma 12.2, and let $C=4\kappa$. For any $0<\varepsilon<1$ and $r\geqslant 2\kappa+6$, there are $0<\eta<\frac{1}{2}$ and $R>0$ with the following property.*

Let $\Gamma=\Gamma_f$ be an intrinsic graph such that $(Q,[-1,1]\times\{0\}\times[-1,1])$ is a $\frac{1}{32}$-rectilinear pseudoquad for Γ. Let g_1 and g_2 be the lower and upper bounds of Q, respectively. If Q is (η,R)-paramonotone on rQ, then there is a vertical plane $P\subseteq\mathbb{H}$ such that

$$\mathcal{H}^4(\overline{B}_r\cap(P^+\triangle\Gamma^+))<\varepsilon. \tag{12.7}$$

Moreover, P is the graph of an an affine function $F\colon V_0\to\mathbb{R}$ of the form $F(w)=a+bx(w)$, whose coefficients satisfy $\max\{|a|,|b|\}\leqslant C$.

Proof. We have $\delta_x(Q)=2$ and $\alpha(Q)=\sqrt{2}$. Also, $2\leqslant|Q|\leqslant 6$. Hence, recalling (7.2), if Γ is (η,R)-paramonotone on rQ, then assuming $R\geqslant 2$ and $\eta R<1$, we have

$$\Omega^P_{\Gamma^+,4}(2Q)\leqslant\tfrac{1}{2}R\Omega^P_{\Gamma^+,2R}(rQ)\leqslant\tfrac{1}{2}R\eta\alpha(Q)^{-4}|Q|\leqslant R\eta<1,$$

so by Lemma 12.2 we have $\|f\|_{L_1(Q)}<\kappa$.

Since $\Pi(\overline{B}_r)\subseteq rQ$, (8.9) implies that

$$\mathrm{ENM}_{\Gamma^+,2R}(\overline{B}_r)\lesssim\eta R.$$

By Proposition 10.1, when R is sufficiently large and ηR is sufficiently small, there is a half-space P^+ bounded by a vertical plane such that

$$\mathcal{H}^4(\overline{B}_r\cap(P^+\triangle\Gamma^+))<\varepsilon.$$

If necessary, we may rotate P infinitesimally around the z-axis so that it is not perpendicular to V_0. Then P is the graph of an affine function $F\colon V_0\to\mathbb{R}$. Let $a,b\in\mathbb{R}$ be such that $F(w)=a+bx(w)$ for all $w\in V_0$.

For all $w\in V_0$, let $\bar{f}(w)$ (resp. $\overline{F}(w)$) be the element of $[-2\kappa,2\kappa]$ that is closest to $f(w)$ (resp. $F(w)$). As $r\geqslant 2\kappa+6$, the intrinsic graphs of $\overline{F}$ and $\bar{f}$ over Q both lie in $\overline{B}_r$. Therefore,

$$\|\overline{F}-\bar{f}\|_{L_1(Q)}\leqslant\mathcal{H}^4(\overline{B}_r\cap(\Gamma^+_{\bar{f}}\triangle\Gamma^+_{\overline{F}}))\leqslant\mathcal{H}^4(\overline{B}_r\cap(\Gamma^+_f\triangle\Gamma^+_F))\leqslant\varepsilon,$$

and thus

$$\|\overline{F}\|_{L_1(Q)}\leqslant\varepsilon+\|\bar{f}\|_{L_1(Q)}\leqslant\varepsilon+\|f\|_{L_1(Q)}\leqslant 2\kappa. \tag{12.8}$$

The map F is affine, and $[-1,1]\times\{0\}\times[-\frac{1}{2},\frac{1}{2}]\subseteq Q$, so $|\{q\in Q:|F(q)|>2\kappa\}|>1$ if $|a|>2\kappa$ or $|b|>4\kappa$, which implies that $\|\overline{F}\|_{L_1(Q)}>2\kappa$, in contradiction to (12.8). So,

$$\max\{|a|,|b|\}\leqslant 4\kappa. \qquad\square$$

We will next use Lemma 12.3 to construct a new intrinsic Lipschitz graph $\widehat{\Gamma}$ that is close to V_0 on a ball around $\mathbf{0}$. Let $0<\varepsilon<1$ and $r>0$ be numbers to be chosen later. Let η, R, C, Γ, f, and Q be as in Lemma 12.3, so that there is a vertical plane P approximating Q that is the graph of an affine function $F(w)=a+bx(w)$ with $\max\{|a|,|b|\}\leqslant C$.

Let $q=q_{a,b}\colon \mathbb{H}\to\mathbb{H}$ be the map given by

$$q(x,y,z)\stackrel{\mathrm{def}}{=}Y^{-a}(x,y-bx,z)=\big(x,y-a-bx,z+\tfrac{1}{2}ax\big)\quad\text{for all }(x,y,z)\in\mathbb{H}.$$

This is a shear map that preserves the x-coordinate and sends P to V_0. Let $\hat{q}\colon V_0\to V_0$ be the map that q induces on V_0, i.e.,

$$\hat{q}(x,0,z)=\Pi(q(x,0,z))=\big(x,0,z+ax+\tfrac{1}{2}bx^2\big)\quad\text{for all }x,z\in\mathbb{R}. \tag{12.9}$$

Let $\widehat{\Gamma}=q(\Gamma)$ and $\widehat{Q}=\hat{q}(Q)$. By Lemma 2.9, $\widehat{Q}$ is a pseudoquad for $\widehat{\Gamma}$ that contains $\mathbf{0}$ and $\widehat{\Gamma}=\Gamma_{\hat{f}}$, where

$$\hat{f}(v)=f(\hat{q}^{-1}(v))-a-bx(v)=f(\hat{q}^{-1}(v))-F(\hat{q}^{-1}(v)).$$

Since $a,b\in[-C,C]$, there is a universal constant $c>0$ such that, for all $s>c^2$,

$$D_{c^{-1}s-c}\subseteq\hat{q}(D_s)=s\widehat{Q}\subseteq D_{cs+c}, \tag{12.10}$$

where we recall $D_s=[-s,s]\times\{0\}\times[-s^2,s^2]$, and

$$B_{c^{-1}s-c}\subseteq q(B_s)\subseteq B_{cs+c}. \tag{12.11}$$

Bounds on Γ and Q correspond directly to bounds on $\widehat{\Gamma}$ and $\widehat{Q}$. For example, shear maps preserve $\mathcal{H}^4$, so

$$\mathcal{H}^4(B_{c^{-1}r-c}\cap(V_0^+\triangle\widehat{\Gamma}^+))\leqslant\mathcal{H}^4(q(B_r)\cap(V_0^+\triangle\widehat{\Gamma}^+))=\mathcal{H}^4(B_r\cap(P^+\triangle\Gamma^+))<\varepsilon. \tag{12.12}$$

In particular, when r is sufficiently large, we have

$$\big\|\min\{|f-F|,\tfrac{1}{2}r\}\big\|_{L_1(10Q)}\leqslant\mathcal{H}^4(B_r\cap(P^+\triangle\Gamma^+))<\varepsilon. \tag{12.13}$$

Maps induced by shears preserve the Lebesgue measure $\mathcal{H}^3$ on V_0, so by (12.10),

$$\|f-F\|_{L_1(10Q)}=\|\hat{f}\|_{L_1(10\widehat{Q})}\leqslant\|\hat{f}\|_{L_1(D_{11c})}, \tag{12.14}$$

and by Lemma 8.8, $\widehat{\Gamma}$ is (η,R)-paramonotone on $r\widehat{Q}$.

12.3. Bounding $\|f-F\|_{L_1(10Q)}$

Next, we bound $\|f-F\|_{L_1(10Q)}$. Lemmas 12.2 and 12.3, together with (12.13) imply that $\|f-F\|_{L_1(Q)}$ and $\left\|\min\{|f-F|, \frac{1}{2}r\}\right\|_{L_1(10Q)}$ can be made arbitrarily small. It remains to show that $|f-F|$ does not have large tails on $10Q$. We previously used Lemma 12.1 to bound the tails of f on Q, but this used the fact that Q is bounded above and below by characteristic curves. We will have to do more work to find supercharacteristic curves above and below $10Q$. In fact, we will show the following bound on $\hat{f}$, then use (12.14) to show a similar bound on $|f-F|$.

LEMMA 12.4. *For any $\delta>0$, there is $\beta=\beta(\delta)>0$ with the following property. Let $\widehat{\Gamma}=\Gamma_{\hat{f}}$ be an intrinsic Lipschitz graph. Let $\tau>0$ and suppose that*

$$\mathcal{H}^4(B_{144\tau}\cap(\widehat{\Gamma}^+\triangle V_0^+))<\beta\tau^4, \tag{12.15}$$

and that the density of $\Omega^P_{\widehat{\Gamma}^+,48\tau}$ on $D_{24\tau}$ is bounded by

$$\tau^{-3}\Omega^P_{\widehat{\Gamma}^+,48\tau}(D_{24\tau})<\beta.$$

Then, $\|\hat{f}\|_{L_1(D_{8\tau})}\leqslant\delta\tau^4$.

Proof. Recall that, by Lemma 8.8, the density of $\Omega^P_{\widehat{\Gamma}^+,48\tau}$ is invariant under scaling, so, after rescaling, it is enough to treat the case $\tau=1$. Let

$$\mathcal{U}\stackrel{\mathrm{def}}{=}\left\{L_{(0,y_0,z_0),m}: z_0\in[200,201],\ y_0\in[1,2],\ \text{and } m\in\left[-\tfrac{1}{20}y_0,-\tfrac{1}{21}y_0\right]\right\}.$$

We claim that there is some $L\in\mathcal{U}$ such that the segment $\Pi(\rho_L([-16,16]))$ is a supercharacteristic curve above D_8. A similar construction will produce a second supercharacteristic curve below D_8, so we can use Lemma 12.1 to bound $\hat{f}$ from above.

We clip $\hat{f}$ between -24 and 24 and call the result h; that is, for all $w\in V_0$, let $h(w)$ be the element of $[-24,24]$ that is closest to $\hat{f}(w)$. For $L\in\mathcal{L}_P$ and $t\in\mathbb{R}$, let

$$h_L(t)=h(\Pi(\rho_L(t))).$$

Define

$$\begin{aligned}
\mathcal{U}_1&\stackrel{\mathrm{def}}{=}\{L\in\mathcal{U}:\Pi(\rho_L([-16,16]))\text{ is supercharacteristic}\},\\
\mathcal{U}_2&\stackrel{\mathrm{def}}{=}\left\{L\in\mathcal{U}:\int_{-24}^{24}|h_L(t)|\,dt>\frac{1}{24}\right\},\\
\mathcal{U}_3&\stackrel{\mathrm{def}}{=}\{L\in\mathcal{U}:\widehat{\omega}^P_{\widehat{\Gamma}^+,48}(D_{24},L)\geqslant 1\}.
\end{aligned}$$

We claim that almost every $L\in\mathcal{U}$ is contained in $\mathcal{U}_1\cup\mathcal{U}_2\cup\mathcal{U}_3$.

Let $L\in\mathcal{U}$ and suppose that $x(L\cap\widehat{\Gamma}^+)$ is a subset of $\mathbb{R}$ with locally finite perimeter. This is true for almost every L. Suppose that $L\notin\mathcal{U}_1\cup\mathcal{U}_2$. Then, $\Pi(\rho_L([-16,16]))$ is not supercharacteristic, so there is some $a\in[-16,16]$ such that $\rho_L(a)\in\widehat{\Gamma}^-$. Let p be the intersection point of L with V_0; by our choice of parameters, $x(p)\in[20,21]$. Also, since $m<-\frac{1}{24}$, we have $y(\rho_L(t))>\frac{1}{24}$ for $t\leqslant 19$. Since $L\notin\mathcal{U}_2$, there are $b_1\in[16,17]$ and $b_2\in[18,19]$ such that, for $i\in\{1,2\}$, we have $h_L(b_i)\leqslant\frac{1}{24}<y(\rho_L(b_i))$, and thus $\rho_L(b_i)\in\widehat{\Gamma}^+$. Similarly, $y(\rho_L(t))<-\frac{1}{24}$ for all $t\geqslant 22$, so there is $c\in[22,23]$ such that $h_L(c)>y(\rho_L(c))$ and $\rho_L(c)\in\widehat{\Gamma}^-$. There is an element of $\partial_{\mathcal{H}^1}x(L\cap\widehat{\Gamma}^+)$ in (a,b_1) and another in (b_2,c). Since $a,b_1,b_2,c\in[-24,24]$, Lemma 8.1 implies that

$$\widehat{\omega}^P_{\widehat{\Gamma}^+,48}(D_{24},L)\geqslant b_2-b_1\geqslant 1$$

and thus $L\in\mathcal{U}_3$.

Therefore, $\mathcal{U}_1\cup\mathcal{U}_2\cup\mathcal{U}_3$ contains all of $\mathcal{U}$ except a null set. We will next show that $\mathcal{N}_P(\mathcal{U}_2)$ and $\mathcal{N}_P(\mathcal{U}_3)$ are bounded by multiples of β.

Suppose $L=L_{(0,y_0,z_0),m}$. As in (12.2), let

$$g_L(t)=z(\Pi(\rho_L(t)))=-\tfrac{1}{2}mt^2-y_0t+z_0.$$

For every $t\in[-24,24]$, we have

$$|g_L(t)-200|\leqslant 1+\frac{m}{2}t^2+y_0|t|\leqslant 1+\frac{24^2}{20}+48\leqslant 100, \tag{12.16}$$

so $\Pi(\rho_L([-24,24]))\subseteq D_{24}$. Furthermore, $D_{24}\subseteq B_{120}$, so for all $v\in D_{24}$ and $t\in[-24,24]$, we have $vY^t\in B_{144}$. Thus,

$$\|h\|_{L_1(D_{24})}\leqslant\mathcal{H}^4(B_{144}\cap(\widehat{\Gamma}^+\triangle V_0^+))<\beta. \tag{12.17}$$

Therefore, for any $y_0\in[1,2]$ and $m\in\left[-\frac{1}{20}y_0,-\frac{1}{21}y_0\right]$,

$$\int_{200}^{201}\int_{-24}^{24}|h_L(t)|\,dt\,dz_0\leqslant\|h\|_{L_1(D_{24})}<\beta.$$

It follows that $\{z_0\in[200,201]:L_{(0,y_0,z_0),m}\in\mathcal{U}_2\}$ has measure at most 24β, and thus

$$\mathcal{N}_P(\mathcal{U}_2)\leqslant\int_1^2\int_{-y_0/20}^{-y_0/21}24\beta\,dm\,dy_0\leqslant 24\beta.$$

To bound $\mathcal{N}_P(\mathcal{U}_3)$, observe that

$$\mathcal{N}_P(\mathcal{U}_3)\leqslant\int_{\mathcal{L}_P}\widehat{\omega}^P_{\widehat{\Gamma}^+,48}(D_{24},L)\,d\mathcal{N}_P(L)=48\Omega^P_{\widehat{\Gamma}^+,48}(D_{24})<48\beta.$$

It follows that, if β is sufficiently small, then

$$\mathcal{N}_P(\mathcal{U}_1) \geqslant \mathcal{N}_P(\mathcal{U}) - \mathcal{N}_P(\mathcal{U}_2) - \mathcal{N}_P(\mathcal{U}_3) > 0.$$

Therefore, $\mathcal{U}_1$ is non-empty. That is, there exists a line $L \in \mathcal{U}$ with parametrization ρ_L such that $S_2 = \Pi(L) \cap \{-16 \leqslant x \leqslant 16\}$ is a supercharacteristic curve. By (12.16), S_2 is above D_8 and $S_2 \subseteq D_{24}$. By symmetry, there also exists a line L' and a supercharacteristic curve $S_1 = \Pi(L') \cap \{-16 \leqslant x \leqslant 16\}$ that lies below D_8 and satisfies $S_1 \subseteq D_{24}$.

By Lemma 12.1 applied to a rescaling of $\widehat{\Gamma}$, there is some $C > 24$ such that, for any $t > C$,

$$|\{v \in D_8 : \hat{f}(v) \geqslant t\}| \lesssim t^{-2} \Omega^P_{\widehat{\Gamma}^+, 48}(D_{24}) \leqslant t^{-2}\beta.$$

Applying another symmetry, the analogous reasoning shows that, for any $t > C$,

$$|\{v \in D_8 : \hat{f}(v) \leqslant -t\}| \lesssim t^{-2}\beta.$$

Then, for all sufficiently small β,

$$\begin{aligned} \|\hat{f}\|_{L_1(D_8)} &= \|h\|_{L_1(D_8)} + \int_{24}^{\infty} \big|\{v \in D_8 : |\hat{f}(v)| \geqslant t\}\big|\, dt \\ &\lesssim \|h\|_{L_1(D_8)} + C\big|\{v \in D_8 : |\hat{f}(v)| \geqslant 24\}\big| + \int_C^{\infty} t^{-2}\beta\, dt \\ &\leqslant \beta + C\big|\{v \in D_8 : |\hat{f}(v)| \geqslant 24\}\big| + \beta, \end{aligned}$$

where we use the fact that $C > 24$ to go from the first line to the second. But

$$\big|\{v \in D_8 : |\hat{f}(v)| \geqslant 24\}\big| = \big|\{v \in D_8 : |h(v)| = 24\}\big| \leqslant \tfrac{1}{24}\|h\|_{L_1(D_8)} \leqslant \tfrac{1}{24}\beta,$$

so $\|\hat{f}\|_{L_1(D_8)} \lesssim \beta$. This proves Lemma 12.4, for β at most a constant multiple of δ. □

We will use the following corollary in the proof of Proposition 7.2.

Corollary 12.5. *Let c be the universal constant in (12.10)–(12.14), and let κ be the universal constant in Lemma 12.2. Denote*

$$\tau \overset{\text{def}}{=} \tfrac{1}{8}\max\{100, 11c\} \quad \textit{and} \quad r \overset{\text{def}}{=} \max\{2\kappa + 6, 144c\tau + c^2\}.$$

For any $\lambda > 0$, there are $\eta, R > 0$ with the following property. Let $\Gamma = \Gamma_f$ be an intrinsic Lipschitz graph and $(Q, [-1,1] \times \{0\} \times [-1,1])$ a $\frac{1}{32}$-rectilinear pseudoquad for Γ. Suppose that Γ is (η, R)-paramonotone on rQ, and P, F, and $\widehat{\Gamma} = \Gamma_{\hat{f}}$ are as in Lemma 12.3 and the remarks immediately after its proof. Then,

$$\|F - f\|_{L_1(10Q)} \leqslant \lambda |Q| \quad \text{and} \quad \|\hat{f}\|_{L_1(D_{100})} \leqslant \lambda.$$

Proof. Set $\delta=\lambda\tau^{-4}$. Let $\beta=\beta(\delta)$ be as in Lemma 12.4. By Lemma 12.3, there are η_0 and R_0 such that, if Γ is (η_0, R_0)-paramonotone on rQ, then

$$\mathcal{H}^4(\overline{B}_{144c\tau+c^2}\cap(P^+\triangle\Gamma^+))<\beta\tau^4.$$

By (12.12), this implies that

$$\mathcal{H}^4(\overline{B}_{144\tau}\cap(V_0^+\triangle\widehat{\Gamma}^+))<\beta\tau^4.$$

We take $R>R_0$ and $\eta R<\eta_0 R_0$ such that (η, R)-paramonotonicity implies (η_0, R_0)-paramonotonicity. Then, by (12.10) and the paramonotonicity of Q,

$$\begin{aligned}\Omega^P_{\widehat{\Gamma}^+,48\tau}(D_{24\tau}) &\leqslant \frac{R\delta_x(\widehat{Q})}{48\tau}\Omega^P_{\widehat{\Gamma}^+,R\delta_x(\widehat{Q})}(D_{24\tau})\\ &\leqslant \frac{R}{24\tau}\Omega^P_{\Gamma^+,R\delta_x(Q)}(D_{24c\tau+c^2})\leqslant \frac{R}{24\tau}|Q|\eta\alpha(Q)^{-4}\lesssim\eta.\end{aligned}$$

If η is sufficiently small, then Lemma 12.4 implies that

$$\|\hat{f}\|_{L_1(D_{\max\{100,11c\}})}<\lambda.$$

By (12.14), this implies that $\|F-f\|_{L_1(10Q)}\leqslant\lambda|Q|$. $\square$

12.4. Characteristic curves are close to lines

Finally, in this section we will show that the characteristic curves of $\widehat{\Gamma}$ are close to horizontal lines and prove Proposition 7.2. The key argument is that when characteristic curves fail to be horizontal, configurations like those in Figure 6 produce non-monotonicity.

Lemma 12.6. *For any $A>0$, there are $\delta=\delta(A), \theta=\theta(A)>0$ with the following property. Let $\widehat{\Gamma}=\Gamma_{\hat{f}}$ be an intrinsic Lipschitz graph. Suppose that*

$$\Omega^P_{\widehat{\Gamma}^+,16}(D_8)<\theta \quad\text{and}\quad \|\hat{f}\|_{L_1(D_8)}<\delta.$$

Let $\gamma:\mathbb{R}\to V_0$ be a characteristic curve through $\mathbf{0}$ and write $\gamma(t)=(t,0,g(t))$ for $t\in\mathbb{R}$. Then, $|g(t)|<A$ for all $t\in[-1,1]$.

Proof. We may suppose that $0<A<1$. Choose

$$\delta=\frac{A^2}{96}\quad\text{and}\quad\theta=\frac{A^3}{10^5}.$$

Our goal is to show that, if $\|\hat{f}\|_{L_1(D_8)}<\delta$ and if there is $t_0\in[-1,1]$ with $|g(t_0)|\geqslant A$, then $\Omega^P_{\widehat{\Gamma}^+,16}(D_8)\geqslant\theta$. After applying a symmetry, we may suppose that $t_0>0$ and that $g(t_0)\leqslant -A$, as in Figure 6.

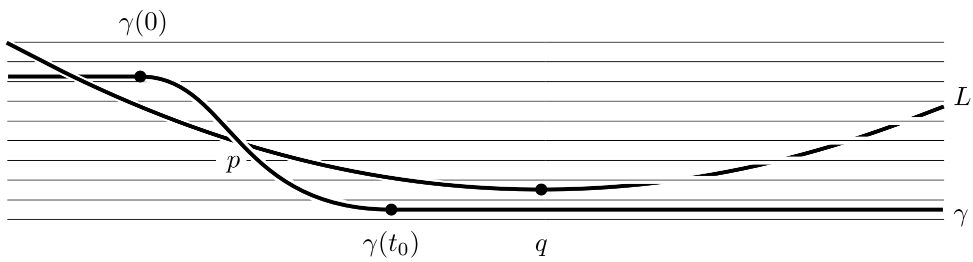

Figure 6. A characteristic curve γ and a horizontal line L, projected to V_0. The projection of L crosses γ positively at p, so L passes behind $\widehat{\Gamma}$ at p, and L intersects V_0 (shown as parallel horizontal lines) at q. If $\hat{f}$ is zero away from γ, then L intersects $\widehat{\Gamma}$ at least three times (twice near p and once at q) and the contribution to $\widehat{\omega}^P$ is at least $\frac{1}{2}(x(q)-x(p))$.

Take $z_0\in\left(-\frac{1}{2}A,0\right)$, $y_0\in\left[\frac{1}{4}A,\frac{1}{2}A\right]$, $m\in\left[-\frac{1}{5}y_0,-\frac{1}{6}y_0\right]$, $w=(0,y_0,z_0)$. Let $L=L_{w,m}$. Suppose that $\Pi(L)$ and γ intersect transversally and $L\cap\widehat{\Gamma}^-$ has finite perimeter; these hold for almost every tuple (y_0,z_0,m). We will show that, if

$$\int_0^8 |\hat{f}(t,0,g_L(t))|\,dt<\frac{A}{24}, \tag{12.18}$$

then $\widehat{\omega}^P_{\widehat{\Gamma}^+,16}(D_8,L)\geqslant 1$, where $g_L=z(\Pi(\rho_L))$.

Suppose that (12.18) holds. For $t\in[-8,8]$, we have

$$|g_L(t)|\leqslant|z_0|+\tfrac{1}{2}|m|t^2+|y_0t|<1+\tfrac{64}{20}+4<64,$$

so $\Pi(\rho_L([-8,8]))\subseteq D_8$. The graphs of g_L and g intersect as depicted in Figure 6. That is, $g_L(0)=z_0<g(0)$, g_L is decreasing on $[0,5]$, and $g_L(0)-g_L(1)=\frac{1}{2}m+y_0<\frac{1}{2}A$, so

$$g_L(t_0)\geqslant g_L(1)>g_L(0)-\tfrac{1}{2}A>-A\geqslant g(t_0).$$

It follows that the graph of g_L crosses γ positively at some point $p=(a,0,g(a))$, where $a\in[0,t_0]$. Since g is characteristic,

$$\hat{f}(a,0,g(a))=-g'(a)>-g_L'(a)=y(\rho_L(a)),$$

so $\rho_L(a)\in\widehat{\Gamma}^-$.

Let q be the point where L intersects V_0. Then $x(q)=-y_0/m\in[5,6]$. Since $m\leqslant-\frac{1}{24}A$, we have $y(\rho_L(t))\geqslant\frac{1}{24}A$ for $t\leqslant 4$ and $y(\rho_L(t))\leqslant-\frac{1}{24}A$ for $t\geqslant 7$. By (12.18), there are $b_1\in[1,2]$ and $b_2\in[3,4]$ such that

$$\hat{f}(b_i,0,g_L(b_i))<\tfrac{1}{24}A\leqslant y(\rho_L(b_i)).$$

This implies that $\rho_L(b_i)\in\widehat{\Gamma}^+$. Similarly, there is $c\in[7,8]$ such that $y(\rho_L(c))<\hat{f}(c,0,g_L(c))$ and thus $\rho_L(c)\in\widehat{\Gamma}^-$. There is an element of $\partial_{\mathcal{H}^1}(x(L\cap\widehat{\Gamma}^+))$ in (a,b_1) and another in (b_2,c), and, by Lemma 8.1,

$$\widehat{\omega}^P_{\widehat{\Gamma}^+,16}(D_8,L)\geqslant b_2-b_1\geqslant 1,$$

as desired.

Therefore, for almost every (m,y_0,z_0) as above, regardless of whether (12.18) holds,

$$\widehat{\omega}^P_{\widehat{\Gamma}^+,16}(D_8,L)+\frac{24}{A}\int_0^8|\hat{f}(t,0,g_L(t))|\,dt\geqslant 1, \tag{12.19}$$

since we showed that at least one of the summands on the left-hand side of (12.19) is at least 1. By integrating (12.19) with respect to z_0, we see that for almost every (m,y_0) that satisfy $y_0\in[\frac{A}{4},\frac{A}{2}]$ and $m\in[-\frac{y_0}{5},-\frac{y_0}{6}]$, we have

$$\begin{aligned}\int_{-A/2}^0\widehat{\omega}^P_{\widehat{\Gamma}^+,16}(D_8,L)\,dz_0&\geqslant\frac{A}{2}-\frac{24}{A}\int_{-A/2}^0\int_0^8|\hat{f}(x,0,g_L(x))|\,dx\,dz_0\\&\geqslant\frac{A}{2}-\frac{24}{A}\|\hat{f}\|_{L_1(D_8)}\geqslant\frac{A}{2}-\frac{24\delta}{A}=\frac{A}{4}.\end{aligned}$$

By integrating this bound over m and y_0 as above, we conclude as follows:

$$\Omega^P_{\widehat{\Gamma}^+,16}(D_8)\geqslant\frac{1}{16}\int_{A/4}^{A/2}\int_{-y_0/5}^{-y_0/6}\int_{-A/2}^0\widehat{\omega}^P_{\widehat{\Gamma}^+,16}(D_8,L)\,dz_0\,dm\,dy_0\geqslant\frac{A^3}{10^5}. \qquad\square$$

Part (2) of Proposition 7.2 follows from Lemma 12.6.

COROLLARY 12.7. *For every $0<\zeta<1$ there are $\delta=\delta(\zeta)>0$ and $\theta=\theta(\zeta)>0$ with the following property. Let $\widehat{\Gamma}=\Gamma_{\hat{f}}$ be an intrinsic Lipschitz graph such that*

$$\Omega^P_{\widehat{\Gamma}^+,128}(D_{100})<\theta\quad\textit{and}\quad\|\hat{f}\|_{L_1(D_{100})}<\delta.$$

Let $\widehat{Q}$ be a pseudoquad for $\widehat{\Gamma}$ with $x(\widehat{Q})=[-1,1]$ such that $\mathbf{0}\in\widehat{Q}$ and $\delta_z(\widehat{Q})=2$. For $u\in 4\widehat{Q}$, if $g_u:\mathbb{R}\to\mathbb{R}$ is such that $\{z=g_u(x)\}$ is a characteristic curve for $\widehat{\Gamma}$ that passes through u, then

$$\|g-z(u)\|_{L_\infty([-4,4])}\leqslant\zeta.$$

That is, $\widehat{Q}$ satisfies part (2) of Proposition 7.2 for $P=V_0$.

Proof. For $p\in V_0$ and $t>0$, denote $D_t(p)=pD_t$. Let $A=\frac{1}{64}\zeta$ and let $\delta,\theta>0$ be constants satisfying Lemma 12.6 for this choice of A.

Let $p\in D_{36}$ be such that $D_{64}(p)\subseteq D_{100}$. Then,

$$\Omega^P_{\widehat{\Gamma}^+,8\cdot 16}(D_{8^2}(p))<\theta\quad\text{and}\quad\|\hat{f}\|_{L_1(D_{8^2}(p))}<\delta,$$

so by Lemma 8.8, the rescaling $s_{1/8,1/8}(p^{-1}\widehat{\Gamma})$ satisfies Lemma 12.6. Hence, if

$$\gamma=\{z=g_p(x)\}$$

is a characteristic curve for $\widehat{\Gamma}$ that passes through p, then

$$\|g_p-z(p)\|_{L_\infty([x(p)-8,x(p)+8])}\leqslant 64A=\zeta.$$

Let g_1 and g_2 be the lower and upper bounds of $\widehat{Q}$, respectively. Then $g_1(0)\in[-3,0]$ and $g_2(0)\in[0,3]$, so $\|g_1-g_1(0)\|_{L_\infty([-8,8])}\leqslant\zeta$ and $\|g_2-g_2(0)\|_{L_\infty([-8,8])}\leqslant\zeta$. Therefore, $4\widehat{Q}\subseteq D_{36}$. If $u\in 4\widehat{Q}$ and $\{z=g_u(x)\}$ is a characteristic curve, then

$$\|g_u-z(u)\|_{L_\infty([-4,4])}\leqslant\|g_u-z(u)\|_{L_\infty([x(u)-8,x(u)+8])}\leqslant\zeta. \qquad \square$$

Finally, we combine the results of this section to prove Proposition 7.2.

Proof of Proposition 7.2. By Lemmas 2.9 and 8.8, if Q is a pseudoquad of Γ and h is a composition of a shear map, a translation, and a stretch map, then Q and Γ satisfy Proposition 7.2 if and only if $\hat{h}(Q)=\Pi(h(Q))$ and $h(\Gamma)$ do. So, by Remark 4.3, it suffices to prove Proposition 7.2 for rectilinear pseudoquads of the form

$$(Q,[-1,1]\times\{0\}\times[-1,1]).$$

Let r be as in Corollary 12.5; we may suppose $r>100$. Let $\delta=\delta(\zeta),\theta=\theta(\zeta)>0$ as in Corollary 12.7. Then we can choose $R_0=R_0(\lambda,\zeta)>0$ and $\eta_0=\eta_0(\lambda,\zeta)>0$ so that, if Γ is (η_0,R_0)-paramonotone on rQ and P, F, and $\widehat{\Gamma}=\Gamma_{\hat{f}}$ are as above, then

$$\|F-f\|_{L_1(10Q)}\leqslant\lambda|Q| \quad \text{and} \quad \|\hat{f}\|_{L_1(D_{100})}\leqslant\delta. \tag{12.20}$$

Denote $R=\max\{R_0,128\}$ and $\eta=\min\{\theta/R,\eta_0R_0/R\}$. Since $R\geqslant R_0$, $\eta R\leqslant\eta_0R_0$, and Γ is (η,R)-paramonotone on rQ, it is also (η,R)-paramonotone, so Q satisfies (12.20), which implies part (1) of Proposition 7.2. Furthermore,

$$\Omega^P_{\widehat{\Gamma}+,128}(D_{100})\leqslant\frac{128}{R}\Omega^P_{\widehat{\Gamma}+,R}(rQ)\leqslant\frac{1}{128}R|Q|\alpha(Q)^{-4}\eta<\theta.$$

Thus, $\widehat{\Gamma}$ satisfies the hypotheses of Corollary 12.7, and so $\widehat{Q}$ satisfies part (2) of Proposition 7.2. As $\widehat{Q}$ is the image of Q under a shear map, part (2) of Proposition 7.2 holds for Q as well. $\square$

Acknowledgements

We thank Alexandros Eskenazis for a discussion that led to Remark 1.13. We are also grateful to the anonymous referees for their careful reading of this article and their many helpful corrections and suggestions.

Our former colleague Louis Nirenberg passed away as this project was being completed. Over the years, he made significant efforts (partially in collaboration with A. N.) to answer the question that we resolve here, though in hindsight those attempts were doomed to fail because they aimed to prove (1.4) with $p=2$, which we now know does not hold. His deep mathematical insights, his contagious joie de vivre, and his kindness are dearly missed.

Appendix A. On the implicit dependence on p in [58]

A version of Theorem 1.3 was stated in [58] with an implicit dependence on the exponent p. In this appendix, we explain how the arguments in [58] can be used to derive the explicit dependence on p that we needed in §1.1.3.

Let $(E,\|\cdot\|_E)$ be a Banach space and fix $q\in[2,\infty]$. The q-uniform convexity constant of X, denoted $K_q(E)$, is defined [6], [8] as the infimum over $K\in(0,\infty]$ such that

$$\left(\|x\|_E^q+\frac{1}{K^q}\|y\|_E^q\right)^{1/q}\leqslant\left(\frac{1}{2}\|x+y\|_E^q+\frac{1}{2}\|x-y\|_E^q\right)^{1/q}\quad\text{for all } x,y\in E. \tag{A.1}$$

Setting $x=0$ in (A.1) shows that necessarily $K\geqslant 1$. By convexity, (A.1) always holds when $K=\infty$ or when $q=\infty$ and $K=1$. Thus, (A.1) quantifies the extent to which the norm $\|\cdot\|_E$ is strictly convex. An equivalent (but somewhat less convenient to work with) formulation of this fact (see [34], [8]) is that $K_q(E)$ is bounded above and below by universal constant multiples of the infimum over those $C>0$ such that the sharpened triangle inequality

$$\|u+v\|_E\leqslant 2-C^{-q}\|u-v\|_E^q$$

holds for any two unit vectors $u,v\in E$.

Theorem 1.3 is the special case $E=\mathbb{R}$, $q=2$, and $1<p\leqslant 2$ of the following theorem.

THEOREM A.1. *For any $p>1$ and $q\geqslant 2$, if $(E,\|\cdot\|_E)$ is a Banach space with*

$$K_q(E)<\infty,$$

then every smooth and compactly supported function $f:\mathbb{H}\to E$ satisfies

$$\begin{aligned}\left(\int_0^\infty \|D_{\mathsf{v}}^t f\|_{L_p(\mathcal{H}^4;E)}^{\max\{p,q\}}\frac{dt}{t}\right)^{1/\max\{p,q\}}&\\ \lesssim \max\{(p-1)^{1/q-1},K_q(E)\}\|\nabla_{\mathbb{H}} f\|_{L_p(\mathcal{H}^4;\ell_p^2(E))},&\end{aligned}\tag{A.2}$$

where we use the (standard) notation

$$\nabla_{\mathbb{H}} f \stackrel{\mathrm{def}}{=} (\mathsf{X}f, \mathsf{Y}f) \in E \times E$$

for the horizontal gradient.

Theorem A.1 is due to [58], except that it is stated there with a factor that depends in an unspecified way on p, q, and E in place of the quantity

$$\max\left\{K_q(E), \frac{1}{(p-1)^{1-1/q}}\right\}.$$

This is because the proof of [58] uses the vector-valued Littlewood–Paley–Stein inequality of [70], for which explicit bounds on the relevant constants were not available in the literature at the time when [58] was written. However, such bounds were subsequently derived in [43] (using in part an argument of [58] itself), so we will next briefly explain how to obtain Theorem A.1 by incorporating this input into [58].

Let $\{h_t\}_{t>0}$ and $\{p_t\}_{t>0}$ be the heat and Poisson kernels on $\mathbb{R}$, respectively, i.e.,

$$h_t(s) \stackrel{\mathrm{def}}{=} \frac{1}{2\sqrt{\pi t}} e^{-s^2/4t} \quad \text{and} \quad p_t(s) \stackrel{\mathrm{def}}{=} \frac{t}{\pi(s^2+t^2)} \quad \text{for all } s>0.$$

It will be convenient to denote the time derivatives $\partial h_t/\partial t$ and $\partial p_t/\partial t$ by $\dot{h}_t$ and $\dot{p}_t$, respectively, i.e.,

$$\dot{h}_t(s) = \frac{s^2-2t}{8\sqrt{\pi} t^{5/2}} e^{-s^2/4t} \quad \text{and} \quad \dot{p}_t(s) = \frac{s^2-t^2}{\pi(s^2+t^2)^2} \quad \text{for all } s>0.$$

By a straightforward evaluation of the integral in (A.3) below, one checks the following standard identity (semigroup subordination; see e.g. [14, §4.4]):

$$\dot{p}_t(s) = \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{e^{-t^2/4u}}{\sqrt{u}} \dot{h}_u(s)\, du \quad \text{for all } s>0. \tag{A.3}$$

Fix $\phi \in L_q(\mathbb{R}; E)$ and $p \geqslant 1$. The following bound holds for any $t>0$:

$$\begin{aligned} \|t\dot{p}_t * \phi\|^p_{L_q(\mathbb{R},E)} &= 2^p \left\| \int_0^\infty \frac{te^{-t^2/4u}}{2u\sqrt{\pi u}} u\dot{h}_u * \phi\, du \right\|^p_{L_q(\mathbb{R},E)} \\ &\leqslant \frac{2^{p-1}t}{\sqrt{\pi}} \int_0^\infty u^{-3/2} e^{-t^2/4u} \|u\dot{h}_u * \phi\|^p_{L_q(\mathbb{R};E)}\, du. \end{aligned} \tag{A.4}$$

The first step of (A.4) is the representation (A.3), and the second step of (A.4) is Jensen's inequality, because

$$\int_0^\infty t \exp \frac{te^{-t^2/4u}}{2u\sqrt{\pi u}}\, du = 1.$$

Integration of (A.4) gives

$$\begin{aligned}\int_0^\infty \|t\dot{p}_t*\phi\|_{L_q(\mathbb{R};E)}^p\frac{dt}{t} &\leqslant \frac{2^{p-1}}{\sqrt{\pi}}\int_0^\infty\Big(\int_0^\infty e^{-t^2/4u}\,dt\Big)u^{-3/2}\|u\dot{h}_u*\phi\|_{L_q(\mathbb{R};E)}^p\,du\\ &= 2^{p-1}\int_0^\infty \|u\dot{h}_u*\phi\|_{L_q(\mathbb{R};E)}^p\frac{du}{u}.\end{aligned}\tag{A.5}$$

Now, if $q\geqslant 2$ and $K_q(E)<\infty$, then it was proved(8) in [43] that

$$\bigg(\int_0^\infty \|t\dot{h}_t*\phi\|_{L_q(\mathbb{R};E)}^q\frac{dt}{t}\bigg)^{1/q}\lesssim K_q(E)\|\phi\|_{L_q(\mathbb{R};E)}.\tag{A.6}$$

In combination with (A.5), we therefore see that also

$$\bigg(\int_0^\infty \|t\dot{p}_t*\phi\|_{L_q(\mathbb{R};E)}^q\frac{dt}{t}\bigg)^{1/q}\lesssim K_q(E)\|\phi\|_{L_q(\mathbb{R};E)}.\tag{A.7}$$

Remark A.2. The reason why we passed from the vector-valued Littlewood–Paley–Stein inequality (A.6) for the heat semigroup to its counterpart (A.7) for the Poisson semigroup is that at the time when [57] was written this was known (with $K_q(E)$ in (A.7) replaced by an unspecified constant factor) for the Poisson semigroup due to [70], while the validity of (A.6) was an open question. For this reason, [57] worked with the Poisson semigroup, so it is simplest to use (A.7) when we refer below to steps in [57]. However, one could repeat the reasoning of [57] mutatis mutandis while working directly with the heat semigroup and using (A.6). The above subordination argument is standard, but we included the quick derivation to verify that the constants are universal.

---

(8) Paper [43] states (A.6) with the factor $K_q(E)$ in the right-hand side replaced by a parameter $\mathfrak{m}_q(E)$ that is called [99] the martingale cotype-q constant of E. There is no need to state the definition of $\mathfrak{m}_q(E)$ here, because it will not have a role in the ensuing discussion; it suffices to recall that, by the martingale inequality of [97], we have

$$\mathfrak{m}_q(E)\lesssim K_q(E).$$

So, (A.5) is a formal consequence of [43], but the above formulation is essentially (namely, up to $O(1)$-renorming) equivalent to that of [43]. For the reverse direction, use the fact that there is a norm $|||\cdot|||$ on E that satisfies $\|x\|_E\asymp|||x|||$ for all $x\in E$ and such that

$$K_q(E,|||\cdot|||)\lesssim\mathfrak{m}_q(E).$$

This renorming statement is essentially due to the deep work [97], except that it is derived in [97] with the weaker property

$$\|x\|_E\leqslant|||x|||\lesssim\mathfrak{m}_q(E)\|x\|_E.$$

The existence of such a norm which is $O(1)$-equivalent to $\|\cdot\|_E$ follows by combining [64] and [76], though we checked (details omitted) that one could adapt the reasoning in [97], so as to obtain a proof of this fact which avoids any reference to the non-linear considerations of [64] and [76]. Alternatively, Gilles Pisier has recently showed us (private communication) a derivation of this $O(1)$-renorming result from the *statement* of [97, Theorem 3.1].

The case $p=q$ of Theorem A.1 follows by substituting (A.7) into [58]. Specifically, we are asserting that the implicit constant in [58, Theorem 2.1] is $O(K_q(E))$ when $p=q$. To check this, note that in the proof of [58, Theorem 2.1] the only loss of a factor that is not a universal constant occurs in [58, equation (18)], which is an instantiation of [58, inequality (15)]; the latter inequality is the same as (A.7) when $p=q$, except that the constant factor in the right-hand side is now specified to be $O(K_q(E))$.

The case $p>q$ of Theorem A.1 follows from the case $p=q$. When $p>q$, we have $K_q(E)\geqslant K_p(E)$ (for justification of this monotonicity, see [8] or [78, §6.2]) and

$$(p-1)^{1-1/q}\leqslant (p-1)^{1-1/p}$$

(since $p>q\geqslant 2$), so the constant on the right-hand side of (A.2) increases as q decreases. We thus suppose from now that $1<p<q$.

For $M>1$, let $\beta_M:\mathbb{H}\to[0,1]$ be a smooth bump function that is $O(1)$-Lipschitz (with respect to the Carnot–Carathéodory metric d), satisfies $\beta_M(h)=1$ for all $h\in B_M$, and has $\operatorname{supp}(\beta_M)\subseteq B_{M+1}$.

For a smooth compactly supported $f:\mathbb{H}\to E$, consider $F_M:\mathbb{H}\to L_p(\mathcal{H}^4;E)$ given by

$$F_M(h)(g)\stackrel{\mathrm{def}}{=}\beta_M(h)f(gh)\quad\text{for all } g,h\in\mathbb{H}. \tag{A.8}$$

We have $(q-1)^{1/q-1}\leqslant 1\leqslant K_q(E)$, so the case $p=q$ of Theorem A.1 with E replaced by $L_p(\mathcal{H}^4;E)$ gives

$$\begin{aligned}\bigg(\int_0^\infty \|D_{\mathsf{v}}^t F_M\|^q_{L_q(\mathcal{H}^4;L_p(\mathcal{H}^4;E))}\frac{dt}{t}\bigg)^{1/q}&\\ \lesssim K_q(L_p(\mathcal{H}^4;E))\|\nabla_{\mathbb{H}}F_M\|_{L_q(\mathcal{H}^4;\ell_q^2(L_p(\mathcal{H}^4;E)))}&\\ \lesssim \max\{(p-1)^{1/q-1},K_q(E)\}\|\nabla_{\mathbb{H}}F_M\|_{L_q(\mathcal{H}^4;\ell_q^2(L_p(\mathcal{H}^4;E)))},&\end{aligned}\tag{A.9}$$

where the last step uses the fact that, by [83, inequality (4.4)],(9) we have

$$K_q(L_p(\mathcal{H}^4;E))\lesssim \max\{(p-1)^{1/q-1},K_q(E)\}. \tag{A.10}$$

To bound the final term in (A.9) from above, note that by the left invariance of $\nabla_{\mathbb{H}}$,

$$\nabla_{\mathbb{H}}F_M(h)(g)=(\mathsf{X}\beta_M(h)f(gh),\mathsf{Y}\beta_M(h)f(gh))+\beta_M(h)\nabla_{\mathbb{H}}f(gh).$$

Hence, for all $h\in\mathbb{H}$,

$$\begin{aligned}&\|\nabla_{\mathbb{H}}F_M(h)\|_{\ell_q^2(L_p(\mathcal{H}^4;E))}\\ &\qquad\lesssim \|f\|_{L_\infty(\mathcal{H}^4;E)}\mathbf{1}_{B_{M+1}(\mathbf{0})\setminus B_M(\mathbf{0})}(h)+\|\nabla_{\mathbb{H}}f\|_{L_p(\mathcal{H}^4;E)}\mathbf{1}_{B_{M+1}(\mathbf{0})}(h).\end{aligned}$$

(9) Formally, [83, inequality (4.4)] is the dual of (A.10); see [8, Lemma 5] for the relevant duality.

So,

$$\|\nabla_{\mathbb{H}} F_M\|_{L_q(\mathcal{H}^4;\ell_q^2(L_p(\mathcal{H}^4;E)))} \lesssim M^{3/q}\|f\|_{L_\infty(\mathcal{H}^4;E)}+M^{4/q}\|\nabla_{\mathbb{H}} f\|_{L_p(\mathcal{H}^4;E)}. \tag{A.11}$$

In order to bound the left-hand side of (A.9) from below, note that, by (2.6), if $0<t<\frac{1}{16}M^2$ and $h\in B_{M-4\sqrt{t}}(\mathbf{0})$, then $hZ^t\in B_M$, and therefore $\beta(h)=\beta(hZ^t)=1$. Hence,

$$\|D_{\mathsf{v}}^t F_M(h)\|_{L_p(\mathcal{H}^4;E)}=\|D_{\mathsf{v}}^t f\|_{L_p(\mathcal{H}^4;E)} \quad \text{for all } h\in B_{M-4\sqrt{t}}(\mathbf{0}).$$

Consequently,

$$\begin{aligned}\|D_{\mathsf{v}}^t F_M\|_{L_q(\mathcal{H}^4;L_p(\mathcal{H}^4;E))} &\geqslant \mathcal{H}^4(B_{M-4\sqrt{t}}(\mathbf{0}))^{1/q}\|D_{\mathsf{v}}^t f\|_{L_p(\mathcal{H}^4;E)}\\ &\asymp (M-4\sqrt{t})^{4/q}\|D_{\mathsf{v}}^t f\|_{L_p(\mathcal{H}^4;E)}.\end{aligned}$$

Hence, for every $0<T<\frac{1}{4}M$, we have

$$\left(\int_0^{T^2}\|D_{\mathsf{v}}^t F_M\|^q_{L_q(\mathcal{H}^4;L_p(\mathcal{H}^4;E))}\frac{dt}{t}\right)^{1/q} \gtrsim (M-4T)^{4/q}\left(\int_0^{T^2}\|D_{\mathsf{v}}^t f\|^q_{L_p(\mathcal{H}^4;E)}\frac{dt}{t}\right)^{1/q}.$$

Combining this with (A.9) and (A.11), letting $M\to\infty$ and then $T\to\infty$, gives Theorem A.1.

Remark A.3. In the setting of the proof of Theorem A.1, the Hardy–Littlewood–Stein (Poisson semigroup) $\mathcal{G}$-function of a function $\phi\in L_q(\mathbb{R};E)$ is the function

$$\mathcal{G}_q(\phi)\colon \mathbb{R}\longrightarrow\mathbb{R}$$

that is defined by

$$\mathcal{G}_q(\phi)(x)\stackrel{\mathrm{def}}{=}\left(\int_0^\infty \|t\dot{p}_t*\phi(x)\|_E^q\frac{dt}{t}\right)^{1/q} \quad \text{for all } x\in E. \tag{A.12}$$

By [70], if $K_q(E)<\infty$, then for every $1<p<\infty$,

$$\|\mathcal{G}_q(\phi)\|_{L_p(\mathbb{R})}\lesssim_{p,q,K_q(E)}\|\phi\|_{L_p(\mathbb{R};E)}. \tag{A.13}$$

If the implicit constant in (A.13) were

$$O\left(\max\left\{K_q(E),\frac{1}{(p-1)^{1-1/q}}\right\}\right)$$

for $1<p<q$ (this is so when $p\geqslant q$, by (A.7) and Jensen's inequality), then Theorem A.1 would follow by direct substitution into [58] without the need to consider the above

averaging argument using the auxiliary function F_M in (A.8). However, it seems that the interpolation argument [70] does not yield this dependence. Determining the optimal dependence on p, q, and $K_q(E)$ in the $\mathcal{G}$-function bound (A.13) remains an interesting open question.

The same question for the heat semigroup variant of (A.13), i.e., with $\dot{p}_t$ replaced by $\dot{h}_t$ in (A.12), is a bigger mystery. That such an inequality for the vector-valued heat semigroup Hardy–Littlewood–Stein $\mathcal{G}$-function holds with any dependence on p, q, and $K_q(E)$ was established recently in [107], but as $p\to 1^+$ the dependence of [107] seems sub-optimal. Obtaining the analogue of (A.6) for the n-dimensional heat semigroup (in which case ϕ is a mapping from $\mathbb{R}^n$ to E) would be very interesting. In [107], this is achieved with a constant that is independent of n but has a much worse dependence on $K_q(E)$.

A substitution of Theorem A.1 into the reasoning of [58] yields the following restatement of the non-embedding result of [58], with explicit dependence on $K_q(E)$.

Theorem A.4. *For $q\geqslant 2$, if E is a Banach space with $K_q(E)<\infty$, then for every $n\in\mathbb{N}$, the word-ball in $\mathbb{H}$ of radius n has E-distortion*

$$\mathsf{c}_E(\mathcal{B}_n)\gtrsim \frac{(\log n)^{1/q}}{K_q(E)}.$$

Since by [8], the Schatten–von Neumann trace class S_r has $K_2(\mathsf{S}_r)=\sqrt{r-1}$ when $1<r\leqslant 2$, Theorem A.4 implies the lower bound on $\mathsf{c}_{\mathsf{S}_r}(\mathcal{B}_n)$ that we used in §1.1.3 (recall that the behavior as $r\to 1^+$ was important for that application). This also shows that the following question about a possible strengthening of Theorem A.4 would imply the distortion lower bound (1.24) that we asked about in §1.1.3. In fact, a positive answer to this question would be a remarkable geometric result, which, as we explained in §1.1.3, would have strong implications; at present, we do not have sufficient evidence to conjecture that the answer is indeed positive in such great generality.

Question A.5. Can the conclusion of Theorem A.4 be improved to

$$\mathsf{c}_E(\mathcal{B}_n)\gtrsim \left(\frac{\log n}{K_q(E)}\right)^{1/q}\ ?$$

Added in proof

We recently learned from Q. Xu that he resolved many of the questions on the growth-rate of the optimal constants in vector-valued Littlewood–Paley-–Stein inequalities that we raised in Remark A.3. This will appear in Xu's forthcoming work [108]. See also his forthcoming work [109] for the evaluation of the order of magnitude of the constants in the classical (real-valued) Littlewood–Paley inequalities.

References

[1] Ambrosio, L., Some fine properties of sets of finite perimeter in Ahlfors regular metric measure spaces. *Adv. Math.*, 159 (2001), 51–67.

[2] Ambrosio, L., Serra Cassano, F. & Vittone, D., Intrinsic regular hypersurfaces in Heisenberg groups. *J. Geom. Anal.*, 16 (2006), 187–232.

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

[4] Austin, T., Naor, A. & Tessera, R., Sharp quantitative nonembeddability of the Heisenberg group into superreflexive Banach spaces. *Groups Geom. Dyn.*, 7 (2013), 497–522.

[5] Badr, N., Real interpolation of Sobolev spaces. *Math. Scand.*, 105 (2009), 235–264.

[6] Ball, K., Markov chains, Riesz transforms and Lipschitz maps. *Geom. Funct. Anal.*, 2 (1992), 137–172.

[7] — The Ribe programme. *Astérisque*, 352 (2013), Exp. No. 1047, 147–159, Séminaire Bourbaki. Vol. 2011/2012. Exposés 1043–1058.

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

[9] Baudier, F., Metrical characterization of super-reflexivity and linear type of Banach spaces. *Arch. Math.* (*Basel*), 89 (2007), 419–429.

[10] Bellaïche, A. & Risler, J.-J. (Eds.), *Sub-Riemannian Geometry.* Progr. Math., 144. Birkhäuser, Basel, 1996.

[11] Bergh, J. & Löfström, J., *Interpolation Spaces. An Introduction.* Grundlehren der Mathematischen Wissenschaften, 223. Springer, Berlin–New York, 1976.

[12] Bigolin, F., Caravenna, L. & Serra Cassano, F., Intrinsic Lipschitz graphs in Heisenberg groups and continuous solutions of a balance equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 32 (2015), 925–963.

[13] Blachère, S., Word distance on the discrete Heisenberg group. *Colloq. Math.*, 95 (2003), 21–36.

[14] Bochner, S., *Harmonic Analysis and the Theory of Probability.* Univ. of California Press, Berkeley–Los Angeles, 1955.

[15] Bourgain, J., On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.*, 52 (1985), 46–52.

[16] — The metrical interpretation of superreflexivity in Banach spaces. *Israel J. Math.*, 56 (1986), 222–230.

[17] Brinkman, B. & Charikar, M., On the impossibility of dimension reduction in l_1. *J. ACM*, 52 (2005), 766–788.

[18] Chatterji, I., Druţu, C. & Haglund, F., Kazhdan and Haagerup properties from the median viewpoint. *Adv. Math.*, 225 (2010), 882–921.

[19] Chávez-Domínguez, J. A., Lipschitz factorization through subsets of Hilbert space. *J. Math. Anal. Appl.*, 418 (2014), 344–356.

[20] Cheeger, J. & Kleiner, B., Differentiating maps into L^1, and the geometry of BV functions. *Ann. of Math.*, 171 (2010), 1347–1385.

[21] — Metric differentiation, monotonicity and maps to L^1. *Invent. Math.*, 182 (2010), 335–370.

[22] Cheeger, J., Kleiner, B. & Naor, A., A $(\log n)^{\Omega(1)}$ integrality gap for the sparsest cut SDP, in *Proc. Symposium on Foundations of Computer Science* (FOCS 2009), pp. 555–564. IEEE Computer Soc., Los Alamitos, CA, 2009.

[23] — Compression bounds for Lipschitz maps from the Heisenberg group to L_1. *Acta Math.*, 207 (2011), 291–373.

[24] Clarkson, J. A., Uniformly convex spaces. *Trans. Amer. Math. Soc.*, 40 (1936), 396–414.
[25] Czuroń, A., Property $F\ell_q$ implies property $F\ell_p$ for $1<p<q<\infty$. *Adv. Math.*, 307 (2017), 715–726.
[26] David, G. & Semmes, S., Singular integrals and rectifiable sets in $\mathbf{R}^n$: Beyond Lipschitz graphs. *Astérisque*, 193 (1991), 152.
[27] — *Analysis of and on Uniformly Rectifiable Sets*. Math. Surveys Monogr., 38. Amer. Math. Soc., Providence, RI, 1993.
[28] Deza, M. M. & Laurent, M., *Geometry of Cuts and Metrics*. Algorithms Combin., 15. Springer, Berlin–Heidelberg, 1997.
[29] Di Donato, D., Fässler, K. & Orponen, T., Metric rectifiability of $\mathbb{H}$-regular surfaces with Hölder continuous horizontal normal. To appear in *Int. Math. Res. Not. IMRN*. https://doi.org/10.1093/imrn/rnab227.
[30] Ding, J., Lee, J. R. & Peres, Y., Markov type and threshold embeddings. *Geom. Funct. Anal.*, 23 (2013), 1207–1229.
[31] Enflo, P., Uniform structures and square roots in topological groups. I, II. *Israel J. Math.*, 8 (1970), 230–252, 253–272.
[32] Farmer, J. D. & Johnson, W. B., Lipschitz p-summing operators. *Proc. Amer. Math. Soc.*, 137 (2009), 2989–2995.
[33] Fässler, K., Orponen, T. & Rigot, S., Semmes surfaces and intrinsic Lipschitz graphs in the Heisenberg group. *Trans. Amer. Math. Soc.*, 373 (2020), 5957–5996.
[34] Figiel, T., On the moduli of convexity and smoothness. *Studia Math.*, 56 (1976), 121–155.
[35] Franchi, B., Serapioni, R. & Serra Cassano, F., Rectifiability and perimeter in the Heisenberg group. *Math. Ann.*, 321 (2001), 479–531.
[36] — Intrinsic Lipschitz graphs in Heisenberg groups. *J. Nonlinear Convex Anal.*, 7 (2006), 423–441.
[37] — Regular submanifolds, graphs and area formula in Heisenberg groups. *Adv. Math.*, 211 (2007), 152–203.
[38] — Differentiability of intrinsic Lipschitz functions within Heisenberg groups. *J. Geom. Anal.*, 21 (2011), 1044–1084.
[39] Gromov, M., Asymptotic invariants of infinite groups, in *Geometric Group Theory*, Vol. 2 (Sussex, 1991), London Math. Soc. Lecture Note Ser., 182, pp. 1–295. Cambridge Univ. Press, Cambridge, 1993.
[40] — Carnot–Carathéodory spaces seen from within, in *Sub-Riemannian Geometry*, Progr. Math., 144, pp. 79–323. Birkhäuser, Basel, 1996.
[41] Gupta, A., Newman, I., Rabinovich, Y. & Sinclair, A., Cuts, trees and l_1-embeddings of graphs. *Combinatorica*, 24 (2004), 233–269.
[42] Hanner, O., On the uniform convexity of L^p and l^p. *Ark. Mat.*, 3 (1956), 239–244.
[43] Hytönen, T. & Naor, A., Heat flow and quantitative differentiation. *J. Eur. Math. Soc.* (*JEMS*), 21 (2019), 3415–3466.
[44] James, R. C., Nonreflexive spaces of type 2. *Israel J. Math.*, 30 (1978), 1–13.
[45] Johnson, W. B. & Lindenstrauss, J., Extensions of Lipschitz mappings into a Hilbert space, in *Conference in Modern Analysis and Probability* (New Haven, CT, 1982), Contemp. Math., 26, pp. 189–206. Amer. Math. Soc., Providence, RI, 1984.
[46] Johnson, W. B., Maurey, B. & Schechtman, G., Non-linear factorization of linear operators. *Bull. Lond. Math. Soc.*, 41 (2009), 663–668.
[47] Johnson, W. B. & Schechtman, G., Diamond graphs and super-reflexivity. *J. Topol. Anal.*, 1 (2009), 177–189.
[48] Jones, P. W., Rectifiable sets and the traveling salesman problem. *Invent. Math.*, 102 (1990), 1–15.

[49] Julia, A., Nicolussi Golo, S. & Vittone, D., Area of intrinsic graphs and coarea formula in Carnot groups. *Math. Z.*, 301 (2022), 1369–1406.

[50] Kadec, M. I. & Pełczyński, A., Bases, lacunary sequences and complemented subspaces in the spaces L_p. *Studia Math.*, 21 (1961/62), 161–176.

[51] Kalton, N. J., The nonlinear geometry of Banach spaces. *Rev. Mat. Complut.*, 21 (2008), 7–60.

[52] — The uniform structure of Banach spaces. *Math. Ann.*, 354 (2012), 1247–1288.

[53] Kirchheim, B. & Serra Cassano, F., Rectifiability and parameterization of intrinsic regular surfaces in the Heisenberg group. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 3 (2004), 871–896.

[54] Kwapień, S., Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients. *Studia Math.*, 44 (1972), 583–595.

[55] Laakso, T. J., Ahlfors Q-regular spaces with arbitrary $Q>1$ admitting weak Poincaré inequality. *Geom. Funct. Anal.*, 10 (2000), 111–123.

[56] — Plane with A_∞-weighted metric not bi-Lipschitz embeddable to $\mathbb{R}^N$. *Bull. London Math. Soc.*, 34 (2002), 667–676.

[57] Lafforgue, V. & Naor, A., A doubling subset of L_p for $p>2$ that is inherently infinite dimensional. *Geom. Dedicata*, 172 (2014), 387–398.

[58] — Vertical versus horizontal Poincaré inequalities on the Heisenberg group. *Israel J. Math.*, 203 (2014), 309–339.

[59] Lang, U. & Plaut, C., Bilipschitz embeddings of metric spaces into space forms. *Geom. Dedicata*, 87 (2001), 285–307.

[60] Lee, J. R., Mendel, M. & Naor, A., Metric structures in L_1: dimension, snowflakes, and average distortion. *European J. Combin.*, 26 (2005), 1180–1190.

[61] Lee, J. R. & Naor, A., Embedding the diamond graph in L_p and dimension reduction in L_1. *Geom. Funct. Anal.*, 14 (2004), 745–747.

[62] — Extending Lipschitz functions via random metric partitions. *Invent. Math.*, 160 (2005), 59–95.

[63] — L_p metrics on the Heisenberg group and the Goemans–Linial conjecture, in *Proc. Symposium on Foundations of Computer Science* (FOCS 2006), pp. 99–108. IEEE Computer Soc., Los Alamitos, CA, 2006.

[64] Lee, J. R., Naor, A. & Peres, Y., Trees and Markov convexity. *Geom. Funct. Anal.*, 18 (2009), 1609–1659.

[65] Lewis, D. R. & Tomczak-Jaegermann, N., Hilbertian and complemented finite-dimensional subspaces of Banach lattices and unitary ideals. *J. Funct. Anal.*, 35 (1980), 165–190.

[66] Lindenstrauss, J. & Pełczyński, A., Absolutely summing operators in L_p-spaces and their applications. *Studia Math.*, 29 (1968), 275–326.

[67] Makarychev, K. & Makarychev, Y., Metric extension operators, vertex sparsifiers and Lipschitz extendability. *Israel J. Math.*, 212 (2016), 913–959.

[68] Mankiewicz, P., On Lipschitz mappings between Fréchet spaces. *Studia Math.*, 41 (1972), 225–241.

[69] Marrakchi, A. & de la Salle, M., Isometric actions on L_p-spaces: dependence on the value of p. Preprint, 2020. arXiv:2001.02490[math.GR].

[70] Martínez, T., Torrea, J. L. & Xu, Q., Vector-valued Littlewood–Paley–Stein theory for semigroups. *Adv. Math.*, 203 (2006), 430–475.

[71] Mattila, P., Serapioni, R. & Serra Cassano, F., Characterizations of intrinsic rectifiability in Heisenberg groups. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 9 (2010), 687–723.

[72] Maurey, B., *Théorèmes de factorisation pour les opérateurs linéaires à valeurs dans les espaces L^p*. Astérisque, 11. Société Mathématique de France, Paris, 1974.
[73] — Type, cotype and K-convexity, in *Handbook of the Geometry of Banach Spaces*, Vol. 2, pp. 1299–1332. North-Holland, Amsterdam, 2003.
[74] Mendel, M. & Naor, A., Euclidean quotients of finite metric spaces. *Adv. Math.*, 189 (2004), 451–494.
[75] — Markov convexity and local rigidity of distorted metrics, in *Computational Geometry* (SCG 2008), pp. 49–58. ACM, New York, 2008.
[76] — Markov convexity and local rigidity of distorted metrics. *J. Eur. Math. Soc.* (*JEMS*), 15 (2013), 287–337.
[77] — Spectral calculus and Lipschitz extension for barycentric metric spaces. *Anal. Geom. Metr. Spaces*, 1 (2013), 163–199.
[78] — Nonlinear spectral calculus and super-expanders. *Publ. Math. Inst. Hautes Études Sci.*, 119 (2014), 1–95.
[79] Montefalcone, F., Some relations among volume, intrinsic perimeter and one-dimensional restrictions of BV functions in Carnot groups. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 4 (2005), 79–128.
[80] Montgomery, R., *A Tour of Subriemannian Geometries*, *their Geodesics and Applications*. Math. Surveys Monogr., 91. Amer. Math. Soc., Providence, RI, 2002.
[81] Naor, A., L_1 embeddings of the Heisenberg group and fast estimation of graph isoperimetry, in *Proc. International Congress of Mathematicians* (ICM 2010), Vol. III, pp. 1549–1575. Hindustan Book Agency, New Delhi, 2010.
[82] — An introduction to the Ribe program. *Jpn. J. Math.*, 7 (2012), 167–233.
[83] — Comparison of metric spectral gaps. *Anal. Geom. Metr. Spaces*, 2 (2014), 1–52.
[84] — Metric dimension reduction: A snapshot of the Ribe program, in *Proc. International Congress of Mathematicians* (ICM 2018), Vol. I, pp. 759–838. World Sci., Hackensack, NJ, 2018.
[85] — An average John theorem. *Geom. Topol.*, 25 (2021), 1631–1717.
[86] Naor, A. & Neiman, O., Assouad's theorem with dimension independent of the snowflaking. *Rev. Mat. Iberoam.*, 28 (2012), 1123–1142.
[87] Naor, A. & Peres, Y., L_p compression, traveling salesmen, and stable walks. *Duke Math. J.*, 157 (2011), 53–108.
[88] Naor, A., Peres, Y., Schramm, O. & Sheffield, S., Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces. *Duke Math. J.*, 134 (2006), 165–197.
[89] Naor, A., Pisier, G. & Schechtman, G., Impossibility of dimension reduction in the nuclear norm [extended abstract], in *Proc. Symposium on Discrete Algorithms*, pp. 1345–1352. SIAM, Philadelphia, PA, 2018.
[90] Naor, A. & Young, R., The integrality gap of the Goemans–Linial SDP relaxation for sparsest cut is at least a constant multiple of $\sqrt{\log n}$, in *Proc. Symposium on Theory of Computing* (STOC 2017), pp. 564–575. ACM, New York, 2017.
[91] — Vertical perimeter versus horizontal perimeter. *Ann. of Math.*, 188 (2018), 171–279.
[92] Newman, I. & Rabinovich, Y., A lower bound on the distortion of embedding planar metrics into Euclidean space. *Discrete Comput. Geom.*, 29 (2003), 77–81.
[93] Ostrovskii, M. I., Embeddability of locally finite metric spaces into Banach spaces is finitely determined. *Proc. Amer. Math. Soc.*, 140 (2012), 2721–2730.
[94] — *Metric Embeddings*. De Gruyter Studies in Mathematics, 49. De Gruyter, Berlin, 2013.
[95] Pansu, P., Métriques de Carnot–Carathéodory et quasiisométries des espaces symétriques de rang un. *Ann. of Math.*, 129 (1989), 1–60.

[96] — Difficulté d'approximation (d'après Khot, Kindler, Mossel, O'Donnell,...). *Astérisque*, 352 (2013), Exp. No. 1045, 83–120, Séminaire Bourbaki. Vol. 2011/2012. Exposés 1043–1058.
[97] Pisier, G., Martingales with values in uniformly convex spaces. *Israel J. Math.*, 20 (1975), 326–350.
[98] — *Factorization of Linear Operators and Geometry of Banach Spaces*. CBMS Regional Conference Series in Mathematics, 60. Amer. Math. Soc., Providence, RI, 1986.
[99] — Probabilistic methods in the geometry of Banach spaces, in *Probability and Analysis* (Varenna, 1985), Lecture Notes in Math., 1206, pp. 167–241. Springer, Berlin, 1986.
[100] Pisier, G. & Xu, Q. H., Random series in the real interpolation spaces between the spaces v_p, in *Geometrical Aspects of Functional Analysis* (1985/86), Lecture Notes in Math., 1267, pp. 185–209. Springer, Berlin, 1987.
[101] Rao, S., Small distortion and volume preserving embeddings for planar and Euclidean metrics, in *Proc. Symposium on Computational Geometry* (Miami Beach, FL, 1999), pp. 300–306. ACM, New York, 1999.
[102] Ribe, M., On uniformly homeomorphic normed spaces. *Ark. Mat.*, 14 (1976), 237–244.
[103] Rigot, S., Quantitative notions of rectifiability in the Heisenberg groups. Preprint, 2019. arXiv:1904.06904[math.MG].
[104] Tao, T., Embedding the Heisenberg group into a bounded-dimensional Euclidean space with optimal distortion. *Rev. Mat. Iberoam.*, 37 (2021), 1–44.
[105] Tessera, R., Quantitative property A, Poincaré inequalities, L^p-compression and L^p-distortion for metric measure spaces. *Geom. Dedicata*, 136 (2008), 203–220.
[106] Xie, X., Some examples of quasiisometries of nilpotent Lie groups. *J. Reine Angew. Math.*, 718 (2016), 25–38.
[107] Xu, Q., Vector-valued Littlewood–Paley–Stein theory for semigroups II. *Int. Math. Res. Not. IMRN*, 21 (2020), 7769–7791.
[108] — Holomorphic functional calculus and vector-valued Littlewood–Paley–Stein theory for semigroups. Preprint, 2021. arXiv:2105.12175[math.FA].
[109] — Optimal orders of the best constants in the Littlewood–Paley inequalities. *J. Funct. Anal.*, 283 (2022), Article No. 109570.

Assaf Naor
Princeton University
Department of Mathematics
Fine Hall, Washington Road
Princeton, NJ 08544-1000
U.S.A.
naor@math.princeton.edu

Robert Young
New York University
Courant Institute of Mathematical Sciences
251 Mercer Street
New York, NY 10012
U.S.A.
ryoung@cims.nyu.edu

Received May 13, 2020
Received in revised form April 19, 2021