HOMOGENEOUS LENGTH FUNCTIONS ON GROUPS

D.H.J. POLYMATH

Abstract. A pseudo-length function defined on an arbitrary group $G = (G, \cdot, e, (\,)^{-1})$ is a map $\ell : G \to [0, +\infty)$ obeying $\ell(e) = 0$, the symmetry property $\ell(x^{-1}) = \ell(x)$, and the triangle inequality $\ell(xy) \leqslant \ell(x) + \ell(y)$ for all $x, y \in G$. We consider pseudo-length functions which saturate the triangle inequality whenever $x = y$, or equivalently those that are *homogeneous* in the sense that $\ell(x^n) = n\,\ell(x)$ for all $n \in \mathbb{N}$. We show that this implies that $\ell([x, y]) = 0$ for all $x, y \in G$. This leads to a classification of such pseudo-length functions as pullbacks from embeddings into a Banach space. We also obtain a quantitative version of our main result which allows for defects in the triangle inequality or the homogeneity property.

1. Introduction

Let $G = (G, \cdot, e, (\,)^{-1})$ be a group (written multiplicatively, with identity element e). A *pseudo-length function* on G is a map $\ell : G \to [0, +\infty)$ that obeys the properties

- $\ell(e) = 0$,
- $\ell(x^{-1}) = \ell(x)$,
- $\ell(xy) \leqslant \ell(x) + \ell(y)$

for all $x, y \in G$. If in addition we have $\ell(x) > 0$ for all $x \in G \backslash \{e\}$, we say that ℓ is a *length function*. By setting $d(x, y) := \ell(x^{-1}y)$, it is easy to see that pseudo-length functions (resp. length functions) are in bijection with left-invariant pseudometrics (resp. left-invariant metrics) on G.

From the axioms of a pseudo-length function it is clear that one has the upper bound

$$\ell(x^n) \leqslant |n|\,\ell(x)$$

for all $x \in G$ and $n \in \mathbb{Z}$. Let us say that a pseudo-length function $\ell : G \to [0, +\infty)$ is *homogeneous* if equality is always attained here, thus one has

$$\ell(x^n) = |n|\,\ell(x) \tag{1.1}$$

for all $x \in G$ and any $n \in \mathbb{Z}$. Using the axioms of a pseudo-length function, it is not difficult to show that the homogeneity condition (1.1) is equivalent to the triangle inequality holding with equality whenever $x = y$ (i.e., that (1.1) holds for $n = 2$); see [7, Lemma 1].

If one has a real or complex Banach space $\mathbb{B} = (\mathbb{B}, \|\,\|)$, and $\phi : G \to \mathbb{B}$ is any homomorphism from G to $\mathbb{B}$ (viewing the latter as a group in additive notation), then the function $\ell : G \to [0, +\infty)$ defined by $\ell(x) := \|\phi(x)\|$ is easily verified to be a homogeneous pseudo-length function. Furthermore,

2010 *Mathematics Subject Classification.* 20F12, 20F65.

Key words and phrases. homogeneous length function, pseudo-length function, quasimorphism, Banach space embedding.

if ϕ is injective, then ℓ will in fact be a homogeneous length function. For instance, the function

$$\ell((n,m)) := |n + \sqrt{2}m|$$

is a length function on $\mathbb{Z}^2$, where in this case $\mathbb{B} := \mathbb{R}$ and $\phi(n,m) := n+\sqrt{2}m$. On the other hand, one can easily locate many length functions that are not homogeneous, for instance by taking the square root of the length function just constructed.

The main result of this paper is that such Banach space constructions are in fact the *only* way to generate homogeneous (pseudo-)length functions.

Theorem 1.2 (Classification of homogeneous length functions)**.** *Given a group G, let $\ell : G \to [0, +\infty)$ be a homogeneous pseudo-length function. Then there exist a real Banach space $\mathbb{B} = (\mathbb{B}, \| \|)$ and a group homomorphism $\phi : G \to \mathbb{B}$ such that $\ell(x) = \|\phi(x)\|$ for all $x \in G$. Furthermore, if ℓ is a length function, one can take ϕ to be injective, i.e., an isometric embedding.*

We will derive Theorem 1.2 from a more quantitative result bounding the pseudo-length of a commutator

$$[x,y] := xyx^{-1}y^{-1}; \tag{1.3}$$

see Proposition 2.1 below. Our arguments will be elementary, relying on directly applying the axioms of a homogeneous length function to various carefully chosen words of x and y, and repeatedly taking an asymptotic limit $n \to \infty$ to dispose of various error terms that arise in the estimates obtained in this fashion.

As one quick corollary of Theorem 1.2, we will obtain the following characterization of the groups that admit homogeneous length functions.

Corollary 1.4. *A group admits a homogeneous length function if and only if it is abelian and torsion-free.*

1.1. **Examples and approaches.** We now make a few remarks to indicate the nontriviality of Theorem 1.2. Corollary 1.4 asserts in particular that there are no non-abelian groups with homogeneous length functions. Whether or not such a striking geometric rigidity phenomenon holds was previously unknown to experts. Moreover, the corollary fails to hold if one or more of the precise conditions in the theorem are weakened. For instance, such pseudo-length functions indeed exist (i) on non-abelian monoids, and (ii) on all balls of finite radius in free groups. We explain these two cases further in Section 4.

Given these cases, one could *a priori* ask if every non-abelian group admits a homogeneous length function. This is not hard to disprove; here are two examples.

Example 1.5 (Nilpotent groups)**.** If G is a nilpotent group of nilpotency class 2 (such as the Heisenberg group), then one has the identity $[x,y]^{n^2} = [x^n, y^n]$ for all $x, y \in G$ and natural numbers n (since the map $(g,h) \mapsto [g,h]$ is now a bihomomorphism : $G \times G \to [G,G] \subset Z(G)$). If $[x,y]$ is non-trivial, then any homogeneous length function on G would assign a linearly growing quantity to the right-hand side and a quadratically growing quantity to the left-hand side, which is absurd; thus such groups cannot admit homogeneous length functions. The claim then also follows for nilpotent groups of higher nilpotency class, since they contain subgroups of nilpotency class 2.[1]

Example 1.6 (Connected Lie groups)**.** As we explain in Remark 2.8, a homogeneous length function in fact induces a *bi-invariant* metric. Thus, if G is a connected Lie group with a homogeneous length function, then by [13, Lemma 7.5], $G \cong K \times \mathbb{R}^n$ for some compact Lie group K and integer $n \geqslant 0$. By (1.1), K cannot have torsion elements, hence must be trivial. But then G is abelian.

Prior to Corollary 1.4, the above examples left open the question of whether *any* non-abelian group admits a homogeneous length function. One may as well consider groups generated by two non-commuting elements. As a prototypical example, let $\mathbf{F}_2$ be the free group on two generators a, b. The word length function on $\mathbf{F}_2$ is a length function, but it is not homogeneous, since for instance the word length of $(bab^{-1})^n = ba^nb^{-1}$ is $n+2$, which is not a linear function of n. (It is however the case that the word length of x^n has linear *growth* in n for any non-trivial x.) Similarly for the Levenshtein distance (edit distance) on $\mathbf{F}_2$.

Our initial attempts to construct homogeneous length functions on $\mathbf{F}_2$ all failed in light of our main result. However, many of these methods apply under minor weakening of the hypotheses, such as working with monoids rather than groups, or weakening homogeneity. Results in these cases are discussed further in Section 4.

1.2. **Further motivations.** We conclude this part with some motivations from functional analysis and probability, or more precisely the study of Banach space embeddings. If G is an additive subgroup of a Banach space $\mathbb{B}$, then clearly the norm on $\mathbb{B}$ restricts to a homogeneous length function on G. In [4, 7] one can find several equivalent conditions for a given length function on a given group to arise in this way (studied in the broader context of additive mappings and separation theorems in functional analysis); see also [15, Theorem 2.10(II)] for an alternate proof. These conditions are summarized in the following:

Proposition 1.7 (see [10, Proposition 3.9])**.** *Given a group G equipped with a length function $\ell : G \to [0, +\infty)$, the following are equivalent:*

1. *There exists a Banach space $\mathbb{B}$ and a group homomorphism $G \to \mathbb{B}$ that is an isometric embedding.*
2. *G is abelian and $d_G(x, y) := \ell(x^{-1}y)$ is a translation-invariant metric for which G is 'normed', i.e., $d_G(e, x^n) = |n|\, d_G(e, x)$ for each $x \in G$ and $n \in \mathbb{Z}$. In particular, G is torsion-free.*
3. *G is $\{2\}$-normed (i.e., $\ell(x^2) = 2\ell(x)$ for each $x \in G$) and is* weakly commutative*, i.e., for all $x, y \in G$ there exists $n = n(x, y) \in \mathbb{N}$ such that $(xy)^{2^n} = x^{2^n} y^{2^n}$.*
4. *G is $\{2\}$-normed and amenable.*

[1] One can also show that solvable non-abelian groups cannot admit homogeneous length functions; see the discussion on lamplighter groups in the comments to `terrytao.wordpress.com/2017/12/16/`.

In view of these equivalences, it is natural to try and characterize the groups possessing a norm. This question is answered by Corollary 1.4, which shows these are precisely the abelian torsion-free groups. Thus in assertion (2), the hypothesis of admitting a 'norm' (i.e., homogeneous length function) is equivalent to that of being abelian and torsion-free; and hence one may use either set of data to infer the remaining assertions in Proposition 1.7.

Groups and semigroups with translation-invariant metrics also naturally arise in probability theory, with the most important 'normed' (i.e., homogeneous) examples being Banach spaces [12]. Notice however that in certain fundamental stochastic settings, formulating – and proving – results does not require the full Banach space structure. In this vein, a general variant of the Hoffmann-Jørgensen inequality was shown in [9] in arbitrary metric semigroups – including Banach spaces as well as (non-abelian) compact Lie groups. Similarly in [10], the authors transferred the (sharp) Khinchin–Kahane inequality from Banach spaces to abelian groups G equipped with a homogeneous length function. To explore extensions of these results to the non-abelian setting – e.g., of Lie groups with left-invariant metrics – we need to first understand if such objects exist. As explained above, this question was not answered in the literature; but it is now settled by our main result.

Finally, there may also be a relation to the Ribe program [14], which aims to reformulate aspects of Banach space theory in purely metric terms. Indeed, from Corollary 1.4 we see that a metric space X is isometric to an additive subgroup of a Banach space if and only if there is a group structure on X which makes the metric left-invariant and the length function $\ell(x) := d(1, x)$ homogeneous.

Acknowledgements

This project is an online collaboration that originated from a blog post at `terrytao.wordpress.com/2017/12/16`, and is the fourteenth of the "Polymath" projects [8]. Further information about this "Polymath14" project, including a full list of participants, may be found at

`michaelnielsen.org/polymath1/index.php?title=linear_norm`

We also thank Michal Doucha for useful references and comments (and in particular in bringing the paper [15] to our attention).

2. Key proposition

The key proposition used to prove Theorem 1.2 is the following estimate, which can treat a somewhat more general class of functions than homogeneous pseudo-length functions, in which the symmetry hypothesis is dropped and one allows for an error in the triangle inequality and the homogeneity property (which is now only claimed for $n = 2$).

Proposition 2.1. *Let $G = (G, \cdot)$ be a group, let $c, c' \in \mathbb{R}$, and let $\ell : G \to \mathbb{R}$ be a function obeying the following axioms:*

(i) *For any $x, y \in G$, one has*

$$\ell(xy) \leqslant \ell(x) + \ell(y) + c. \tag{2.2}$$

(ii) *For any* $x \in G$, *one has*

$$\ell(x^2) \geqslant 2\ell(x) - c'. \tag{2.3}$$

Then for any $x, y \in G$, *one has*

$$\ell([x,y]) \leqslant 4c + 5c', \tag{2.4}$$

where the commutator $[x,y]$ *was defined in* (1.3).

Notably, we neither assume symmetry $\ell(x^{-1}) = \ell(x)$, not even up to a constant, nor $\ell(e) = 0$; we also allow ℓ to take on negative values. The reader may however wish to restrict attention to length functions (and set $c = c' = 0$) for a first reading of the arguments below. The factors of $4, 5$ are probably not optimal here, but the crucial feature of the bound (2.4) for our application is that the right-hand side vanishes when $c = c' = 0$.

For the remainder of this section, let G, c, c', and ℓ satisfy the hypotheses of the proposition. Our task is to establish the bound (2.4). We shall now use (2.2) and (2.3) repeatedly to establish a number of further inequalities relating the pseudo-lengths $\ell(x)$ of various elements x of G, culminating in (2.4). Many of our inequalities will involve terms that depend on an auxiliary parameter n, but we will be able to eliminate several of them by the device of passing to the limit $n \to \infty$. It is because of this device that we are able to obtain a bound (2.4) whose right-hand side is completely uniform in x and y.

From (2.2) and induction we have the approximate upper homogeneity bound

$$\ell(x^n) \leqslant n\ell(x) + (n-1)c \tag{2.5}$$

for any natural number $n \geqslant 1$. Similarly, from (2.3) and induction one has the lower homogeneity bound

$$\ell(x^n) \geqslant n\ell(x) - (n-1)c'$$

whenever n is a power of two. It is convenient to rearrange this latter inequality as

$$\ell(x) \leqslant \frac{\ell(x^n)}{n} + \frac{n-1}{n}c'. \tag{2.6}$$

This inequality (particularly in the asymptotic limit $n \to \infty$) will be the principal means by which the hypothesis (2.3) is employed.

We remark that by further use of (2.5) one can also obtain a similar estimate to (2.6) for natural numbers n that are not powers of two[2], but the powers of two will suffice for the arguments that follow.

Lemma 2.7 (Approximate conjugation invariance)**.** *For any* $x, y \in G$, *one has*

$$\ell(yxy^{-1}) \leqslant \ell(x) + c + c'.$$

Remark 2.8. Setting $c = c' = 0$, we conclude that any homogeneous pseudo-length function is conjugation invariant, and thus determines a bi-invariant metric on G. It should not be surprising that this observation is used in the proof of Theorem 1.2, since it is a simple consequence of that

[2]For instance, this argument shows that if a pseudo-length function obeys the doubling condition $\ell(x^2) = 2\ell(x)$ for each $x \in G$, then it is homogeneous; see also [7, Lemma 1].

theorem. Observe also that the zero locus of such a map $\ell : G \to [0, +\infty)$ is a normal subgroup of G; thus ℓ factors through a homogeneous length function : $\overline{\ell} : G/\ell^{-1}(0) \to [0, +\infty)$.

Proof of Lemma 2.7. From (2.6) (with x replaced by yxy^{-1}) one has

$$\ell(yxy^{-1}) \leqslant \frac{\ell(yx^n y^{-1})}{n} + \frac{n-1}{n} c'$$

whenever n is a power of two. On the other hand, from (2.5) and (2.2) one has

$$\ell(yx^n y^{-1}) \leqslant \ell(y) + n\ell(x) + \ell(y^{-1}) + (n+1)c$$

and thus

$$\ell(yxy^{-1}) \leqslant \ell(x) + c + c' + \frac{\ell(y) + \ell(y^{-1}) + c - c'}{n}.$$

Sending $n \to \infty$, we obtain the claim. □

Lemma 2.9 (Splitting lemma)**.** *Let $x, y, z, w \in G$ be such that x is conjugate to both wy and zw^{-1}. Then one has*

$$\ell(x) \leqslant \frac{\ell(y) + \ell(z)}{2} + \frac{3}{2}(c + c'). \tag{2.10}$$

Proof. If we write $x = swys^{-1} = tzw^{-1}t^{-1}$ for some $s, t \in G$, then from (2.6) we have

$$\begin{aligned} \ell(x) &\leqslant \frac{\ell(x^n x^n)}{2n} + \frac{2n-1}{2n} c' \\ &= \frac{\ell(s(wy)^n s^{-1} t(zw^{-1})^n t^{-1})}{2n} + \frac{2n-1}{2n} c' \end{aligned}$$

whenever n is a power of two. From Lemma 2.7 and (2.2) one has

$$\begin{aligned} \ell((wy)^{k+1} s^{-1} t(zw^{-1})^{k+1}) &= \ell(wy(wy)^k s^{-1} t(zw^{-1})^k zw^{-1}) \\ &\leqslant \ell(y(wy)^k s^{-1} t(zw^{-1})^k z) + c + c' \\ &\leqslant \ell((wy)^k s^{-1} t(zw^{-1})^k) + \ell(y) + \ell(z) + 3c + c' \end{aligned}$$

for any $k \geqslant 0$, and hence by induction

$$\ell((wy)^n s^{-1} t(zw^{-1})^n) \leqslant \ell(s^{-1}t) + n(\ell(y) + \ell(z) + 3c + c').$$

Inserting this into the previous bound for $\ell(x)$ (via two applications of (2.2)), we conclude that

$$\ell(x) \leqslant \frac{\ell(y) + \ell(z) + 3c + c'}{2} + \frac{\ell(s) + \ell(s^{-1}t) + \ell(t^{-1}) + 2c}{2n} + \frac{2n-1}{2n} c';$$

sending $n \to \infty$, we obtain the claim. □

Corollary 2.11. *If $x, y \in G$, let $f = f_{x,y} : \mathbb{Z}^2 \to \mathbb{R}$ denote the function*

$$f(m, k) := \ell(x^m [x, y]^k).$$

Then for any $m, k \in \mathbb{Z}$, we have

$$f(m, k) \leqslant \frac{f(m-1, k) + f(m+1, k-1)}{2} + 2(c + c'). \tag{2.12}$$

Proof. Observe that $x^m[x,y]^k$ is conjugate to both $x(x^{m-1}[x,y]^k)$ and to $(y^{-1}x^m[x,y]^{k-1}xy)x^{-1}$, hence by (2.10) one has

$$\ell(x^m[x,y]^k) \leqslant \frac{\ell(x^{m-1}[x,y]^k) + \ell(y^{-1}x^m[x,y]^{k-1}xy)}{2} + \frac{3}{2}(c+c').$$

Since $y^{-1}x^m[x,y]^{k-1}xy$ is conjugate to $x^{m+1}[x,y]^{k-1}$, the claim now follows from Lemma 2.7. $\square$

We now prove Proposition 2.1. Let $x, y \in G$. We can write the inequality (2.12) in probabilistic form as

$$f(m,k) \leqslant \mathbf{E} f\left(\left(m, k-\frac{1}{2}\right) + Y\left(1, -\frac{1}{2}\right)\right) + 2(c+c')$$

where $Y = \pm 1$ is a Bernoulli random variable that equals 1 or -1 with equal probability. The key point here is the drift of $\left(0, -\frac{1}{2}\right)$ in the right-hand side. Iterating this inequality, we see that if we let n be a large power of 2, then we have

$$f(0,n) \leqslant \mathbf{E} f\left((Y_1 + \cdots + Y_{2n})\left(1, -\frac{1}{2}\right)\right) + 4(c+c')n,$$

where $Y_1, \dots, Y_{2n}$ are independent copies of Y (so in particular $Y_1 + \cdots + Y_{2n}$ is an even integer).

From (2.2) and (2.5) one has the inequality

$$\begin{aligned} f(m,k) \quad &\leqslant \quad |m| \left(\max(\ell(x), \ell(x^{-1})) + c\right) \\ &\quad\quad + |k| \left(\max(\ell([x,y]), \ell([x,y]^{-1})) + c\right) + \ell(e) \end{aligned}$$

(the final term being needed to deal with the $m = k = 0$ case). We conclude that

$$f\left((Y_1 + \cdots + Y_{2n})\left(1, -\frac{1}{2}\right)\right) \leqslant A|Y_1 + \cdots + Y_{2n}| + \ell(e)$$

where A is a quantity independent of n; more explicitly, one can take

$$A := \max\left(\ell(x), \ell(x^{-1})\right) + \frac{1}{2}\max\left(\ell([x,y]), \ell([x,y]^{-1})\right) + \frac{3}{2}c.$$

Taking expectations, since the random variable $Y_1 + \cdots + Y_{2n}$ has mean zero and variance $2n$, we see from the Cauchy–Schwarz inequality (or Jensen's inequality) that

$$\mathbf{E}|Y_1 + \cdots + Y_{2n}| \leqslant \left(\mathbf{E}|Y_1 + \cdots + Y_{2n}|^2\right)^{1/2} = \sqrt{2n}$$

and hence

$$f(0,n) \leqslant A\sqrt{2n} + \ell(e) + 4(c+c')n.$$

But from (2.6) we have

$$\ell([x,y]) \leqslant \frac{f(0,n)}{n} + \frac{n-1}{n}c'.$$

Combining these two bounds and sending $n \to \infty$, we obtain Proposition 2.1.

Remark 2.13. As discussed in the footnote preceding Lemma 2.7, if $c = 0$ then the condition $c' = 0$, or equivalently $\ell(x^2) = 2\ell(x)$ for all $x \in G$, is in turn equivalent to $\ell(x^k) = k\ell(x)$ for all $x \in G$, for any $k \geqslant 2$. One can thus ask if (2.3) can be replaced by a similar condition of the form

$$\ell(x^k) \geqslant k\ell(x) - c'$$

for some $k \geqslant 2$. We claim that such a condition essentially implies (2.3). Indeed, using the triangle inequality for $x^k = (x^2)^{\lfloor k/2 \rfloor} x^\delta$ where $\delta = 0, 1$, we obtain:

$$k\ell(x) - c' \leqslant \ell(x^k) \leqslant O(k)\ell(x^2) + c\,O(k),$$

which in turn implies $\ell(x^2) \geqslant O(1)\ell(x) - O(1)c - c'/k$.

Remark 2.14. One can deduce a 'local' version of Proposition 2.1 as follows: notice that the constants c, c' can be described in terms of ℓ from (2.2) and (2.3), to yield

$$\ell([x,y]) \leqslant 4\left(\sup_{z,w \in G} \ell(zw) - \ell(z) - \ell(w)\right) + 5\left(\sup_{z \in G} 2\ell(z) - \ell(z^2)\right), \quad (2.15)$$

for any group G and function $\ell : G \to \mathbb{R}$, and any $x, y \in G$. (Both sides are zero when G is a Banach space and ℓ is the norm, so equality is obtained in that case.) Now it is enough to consider $G = \langle x, y \rangle$ without loss of generality, which may lead to a better bound on $\ell([x,y])$ than taking the suprema over all of G. Notice also that at least one of the constants c, c' must be non-negative, from (2.2) with $x = y$ and (2.3):

$$\max\{c, c'\} \geqslant \max\left\{\ell(e^2) - 2\ell(e), 2\ell(e) - \ell(e^2)\right\} = \max\{-\ell(e), \ell(e)\} \geqslant 0. \quad (2.16)$$

In fact, this reasoning and our results imply that the only way to get $c = c' = 0$ on the right-hand side of (2.15) is when ℓ arises from pulling back the norm of a Banach space $\mathbb{B}$ along a group homomorphism $G \to \mathbb{B}$, or equivalently along a group homomorphism from the torsion-free abelianization of G to $\mathbb{B}$.

3. Proof of main theorem

With Proposition 2.1 in hand, it is not difficult to conclude the proof of Theorem 1.2. Suppose that G is a group with a homogeneous pseudo-length function $\ell : G \to [0, +\infty)$. Applying Proposition 2.1 with $c = c' = 0$, we conclude that $\ell([x,y]) = 0$ for all $x, y \in G$, thus by the triangle inequality ℓ vanishes on the commutator subgroup $[G,G]$, and therefore factors through the abelianization $G_{\rm ab} := G/[G,G]$ of G. (This already establishes part of one implication of Corollary 1.4.) Factoring out by $[G,G]$ like this, we may now assume without loss of generality that G is abelian. To reflect this, we now use additive notation for G, thus for instance $\ell(nx) = |n|\ell(x)$ for each $x \in G$ and $n \in \mathbb{Z}$, and one can also view G as a module over the integers $\mathbb{Z}$.

At this point we repeat the arguments in [10, Theorem B], which treated the case when G was separable, though it turns out that this separability hypothesis is unnecessary.

If x is a torsion element of G (thus $nx = 0$ for some n), then the homogeneity condition forces $\ell(x) = 0$. Thus ℓ vanishes on the torsion subgroup of G; factoring out by this subgroup, we may thus assume without loss of generality that G is not only abelian, but is also torsion-free.

We can view G as a subgroup of the $\mathbb{Q}$-vector space $G \otimes_{\mathbb{Z}} \mathbb{Q}$, the elements of which can be formally expressed as $\frac{1}{n}x$ for natural numbers n and elements $x \in G$ (with two such expressions $\frac{1}{n}x$, $\frac{1}{m}y$ identified if and only if $mx = ny$, and the $\mathbb{Q}$-vector space operations defined in the obvious fashion); the fact that this is well defined as a $\mathbb{Q}$-vector space follows from the hypotheses that G is abelian and torsion-free. We can then define a seminorm $\| \|_{\mathbb{Q}} :$ $G \otimes_{\mathbb{Z}} \mathbb{Q} \to [0, +\infty)$ by setting

$$\left\| \frac{1}{n} x \right\|_{\mathbb{Q}} := \frac{1}{n} \ell(x)$$

for any $x \in G$ and natural number n; the linear growth condition ensures that $\| \|_{\mathbb{Q}}$ is well-defined. It is not difficult to verify the homogeneity condition

$$\|qx\|_{\mathbb{Q}} = |q| \|x\|_{\mathbb{Q}}$$

and the triangle inequality

$$\|x + y\|_{\mathbb{Q}} \leqslant \|x\|_{\mathbb{Q}} + \|y\|_{\mathbb{Q}}$$

for all $x, y \in G \otimes_{\mathbb{Z}} \mathbb{Q}$ and $q \in \mathbb{Q}$, so that $\| \|_{\mathbb{Q}}$ is indeed a seminorm over the $\mathbb{Q}$-vector space $G \otimes_{\mathbb{Z}} \mathbb{Q}$.

The $\mathbb{Q}$-vector space $G \otimes_{\mathbb{Z}} \mathbb{Q}$ may be embedded in turn into the $\mathbb{R}$-vector space $G \otimes_{\mathbb{Z}} \mathbb{R} = (G \otimes_{\mathbb{Z}} \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{R}$, elements of which take the form $\alpha_1 x_1 + \cdots + \alpha_k x_k$ for some $x_1, \ldots, x_k \in G \otimes_{\mathbb{Z}} \mathbb{Q}$ and $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$. We identify two such elements $\alpha_1 x_1 + \cdots + \alpha_k x_k$, $\beta_1 y_1 + \cdots + \beta_l y_l$ if and only if there are $\mathbb{Q}$-linearly independent elements $z_1, \ldots, z_m \in G \otimes_{\mathbb{Z}} \mathbb{Q}$ and coefficients $a_{in}, b_{jn} \in \mathbb{Q}$ for $i = 1, \ldots, k$, $j = 1, \ldots, l$, $n = 1, \ldots, m$ such that $x_i = \sum_{n=1}^m a_{in} z_n$ for each $i = 1, \ldots, k$, $y_j = \sum_{n=1}^m b_{jn} z_n$ for each $j = 1, \ldots, l$, and $\sum_{i=1}^k \alpha_i a_{in} = \sum_{j=1}^l \beta_j b_{jn}$ for each $n = 1, \ldots, m$; one can check that this criterion does not depend on the precise choice of linearly independent elements $z_1, \ldots, z_m$, so long as their span contains all of the $x_1, \ldots, x_k, y_1, \ldots, y_l$.[3] It is easy to verify[4] that $G \otimes_{\mathbb{Z}} \mathbb{R}$ has the structure of an $\mathbb{R}$-vector space that contains $G \otimes_{\mathbb{Z}} \mathbb{Q}$ (and hence G). We can extend the seminorm $\| \|_{\mathbb{Q}}$ to $G \otimes_{\mathbb{Z}} \mathbb{R}$ by setting

$$\|\alpha_1 x_1 + \cdots + \alpha_k x_k\|_{\mathbb{R}} := \lim_{(q_1, \ldots, q_k) \to (\alpha_1, \ldots, \alpha_k)} \|q_1 x_1 + \cdots + q_k x_k\|_{\mathbb{Q}}$$

[3]Stated slightly differently, let $z_1, \ldots, z_m$ denote a $\mathbb{Q}$-basis of the $\mathbb{Q}$-span of the x_i, y_j, and define the vector $\boldsymbol{z} := (z_1, \ldots, z_m)^T$ by abuse of notation; similarly define $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\alpha}, \boldsymbol{\beta}$. Then $\boldsymbol{x} = A\boldsymbol{z}$, $\boldsymbol{y} = B\boldsymbol{z}$ for matrices $A \in \mathbb{Q}^{k \times m}, B \in \mathbb{Q}^{l \times m}$, and we identify $\boldsymbol{\alpha}^T \boldsymbol{x} = \boldsymbol{\beta}^T \boldsymbol{y}$ if and only if $\boldsymbol{\alpha}^T A = \boldsymbol{\beta}^T B$.

[4]One subtle point is that if G is already embedded inside an existing $\mathbb{R}$-vector space V, then the vector space $G \otimes_{\mathbb{Z}} \mathbb{R}$ constructed here is not necessarily a subspace of that vector space (although there will be a canonical linear map from $G \otimes_{\mathbb{Z}} \mathbb{R}$ to V). For instance, if $G = \langle 1, \sqrt{2} \rangle \subset \mathbb{R}$, then while G is contained in a one-dimensional $\mathbb{R}$-vector space, $G \otimes_{\mathbb{Z}} \mathbb{R}$ is in fact two-dimensional (with the generators 1 and $\sqrt{2}$ of G forming a $\mathbb{R}$-basis of this space); the point is that the vector space operations on $G \otimes_{\mathbb{Z}} \mathbb{R}$ do not agree in this case with those of the ambient vector space $\mathbb{R}$, for instance in $G \otimes_{\mathbb{Z}} \mathbb{R}$ the second generator $\sqrt{2}$ is not equal to the scalar multiple of the first generator 1 by $\sqrt{2}$. See also Remark 3.1 below. To avoid confusion, we recommend to the reader that they view G as an abstract group that is not already embedded into an existing $\mathbb{R}$-vector space.

where $(q_1, \dots, q_k)$ is restricted to $\mathbb{Q}^k$. For any given $\alpha_1, \dots, \alpha_k, x_1, \dots, x_k$ it is easy to see using the triangle inequality and homogeneity that the limit exists. We still need to verify that the seminorm is well-defined, in the sense that

$$\|\alpha_1 x_1 + \cdots + \alpha_k x_k\|_{\mathbb{R}} = \|\beta_1 y_1 + \cdots + \beta_l y_l\|_{\mathbb{R}}$$

whenever $\alpha_1 x_1 + \cdots + \alpha_k x_k = \beta_1 y_1 + \cdots + \beta_l y_l$. By the triangle inequality it suffices to verify this at the origin, or more precisely that

$$\lim_{(q_1,\dots,q_k)\to(\alpha_1,\dots,\alpha_k)} \|q_1 x_1 + \cdots + q_k x_k\|_{\mathbb{Q}} = 0$$

whenever $x_1, \dots, x_k \in G \otimes_{\mathbb{Z}} \mathbb{Q}$ and $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ are such that $\alpha_1 x_1 + \cdots + \alpha_k x_k = 0$. To prove this one can induct on k, the case $k = 0$ being trivial. If one of the $x_1, \dots, x_k$ is a $\mathbb{Q}$-linear combination of the others, then we can eliminate it and invoke the induction hypothesis for $k-1$, so we may assume without loss of generality that the $x_1, \dots, x_k$ are linearly independent over $\mathbb{Q}$. But then, by construction of $G \otimes_{\mathbb{Z}} \mathbb{R}$, the condition $\alpha_1 x_1 + \cdots + \alpha_k x_k = 0$ forces $\alpha_1 = \cdots = \alpha_k = 0$, and the claim now follows from the triangle inequality and homogeneity.

By construction, $\| \|_{\mathbb{R}} : G \otimes_{\mathbb{Z}} \mathbb{R} \to [0, +\infty)$ obeys the homogeneity condition

$$\|\alpha x\|_{\mathbb{R}} = |\alpha| \|x\|_{\mathbb{R}}$$

and the triangle inequality

$$\|x + y\|_{\mathbb{R}} \leqslant \|x\|_{\mathbb{R}} + \|y\|_{\mathbb{R}}$$

for all $x, y \in G \otimes_{\mathbb{Z}} \mathbb{R}$ and $\alpha \in \mathbb{R}$; thus $\| \|_{\mathbb{R}}$ is a seminorm on $G \otimes_{\mathbb{Z}} \mathbb{R}$. By the usual metric completion device of taking formal Cauchy sequences and factoring out the elements that converge to zero, we can find a real Banach space $\mathbb{B} = (\mathbb{B}, \| \|_{\mathbb{B}})$ and an $\mathbb{R}$-linear map $\phi : G \otimes_{\mathbb{Z}} \mathbb{R} \to \mathbb{B}$ such that $\|x\|_{\mathbb{R}} = \|\phi(x)\|_{\mathbb{B}}$ for each $x \in G \otimes_{\mathbb{Z}} \mathbb{R}$. It is now clear that ϕ restricts to a homomorphism from G to $\mathbb{B}$ such that $\ell(x) = \|\phi(x)\|_{\mathbb{B}}$ for each $x \in G$. Finally, if ℓ is a length function, then $\|\phi(x)\|_{\mathbb{B}}$ must be non-zero for each $x \neq e$, which implies that the kernel of ϕ is trivial and hence ϕ is injective.

This concludes the proof of Theorem 1.2. Since the homomorphism $\phi : G \to \mathbb{B}$ can only be injective for abelian torsion-free G, we obtain the "only if" portion of Corollary 1.4. Conversely, if a group G is abelian and torsion-free, by the above constructions it embeds into a real vector space $\mathbb{B} := G \otimes_{\mathbb{Z}} \mathbb{R}$; now by Zorn's lemma $\mathbb{B}$ has a norm (e.g., consider the ℓ^1 norm with respect to a Hamel basis of $\mathbb{B}$), which restricts to the desired homogeneous length function on G.

Remark 3.1. Purely algebraically, $G \otimes_{\mathbb{Z}} \mathbb{R}$ is the smallest $\mathbb{R}$-vector space 'enveloping' the group G in the sense that it contains a homomorphic image of G such that any homomorphism from G to a vector space V extends uniquely to a linear map from $G \otimes_{\mathbb{Z}} \mathbb{R}$ to V. If G embeds into a vector space, this universal map $G \to G \otimes_{\mathbb{Z}} \mathbb{R}$ is not necessarily the smallest-*dimensional* vector space embedding of G, since for example $\mathbb{R}$ contains copies of $\mathbb{Z}^n$ for any n.

Similarly, the space $\mathbb{B}$ from the above proof is the smallest Banach space containing an isometric copy of G, in the sense that any group homomorphism $f : G \to V$ that satisfies $\|f(g)\| \leqslant \ell(g)$ extends uniquely to an $\mathbb{R}$-linear map $\mathbb{B} \to V$ of norm at most 1. The existence of such a universal $\mathbb{B}$ equipped with a map $G \to \mathbb{B}$ is a purely formal statement that could also be derived abstractly from the adjoint functor theorem. The nontrivial part of our result is that the universal map $G \to \mathbb{B}$ constructed in the proof is actually an *isometry*.

It is worth noting that our $\mathbb{B}$ may not necessarily coincide with the algebraic enveloping vector space $G \otimes_{\mathbb{Z}} \mathbb{R}$. Returning to the example from the introduction, the length function $\ell(n, m) = |n + \sqrt{2}m|$ on $\mathbb{Z}^2$ has $G \otimes_{\mathbb{Z}} \mathbb{R} \cong \mathbb{R}^2$ but $\mathbb{B} \cong \mathbb{R}$.

Remark 3.2. If one is interested in the actual 'smallest' object containing an abelian torsion-free group G, then we mention two such objects (in the appropriate categories). First, $G \otimes_{\mathbb{Z}} \mathbb{Q}$ is indeed the smallest $\mathbb{Q}$-vector space containing G. Similarly, as observed in [10], one can directly take the metric completion of $G \otimes_{\mathbb{Z}} \mathbb{Q}$ without first passing to $G \otimes_{\mathbb{Z}} \mathbb{R}$. This constructs the unique smallest Banach space containing G.

Remark 3.3. The above arguments also show that homogeneous pseudo-length functions on G are in bijection with seminorms on the real vector space $G_{\mathrm{ab},0} \otimes_{\mathbb{Z}} \mathbb{R}$, where $G_{\mathrm{ab},0}$ denotes the torsion-free abelianization of G.

4. Further remarks and results

One could ask if the homogeneity condition can be relaxed to an asymptotic homogeneity condition

$$\lim_{n\to\infty} \frac{\ell(x^n)}{n} = \ell(x), \quad \text{for each} \quad x \in G.$$

In fact this condition is equivalent to homogeneity, since for all positive integers $m \in \mathbb{N}$, the sequence $\{\ell(x^{mn})/(mn) : n \geqslant 1\}$ would converge as $n \to \infty$ to both $\ell(x)$ and $\ell(x^m)/m$. In particular, if $\ell : G \to [0, +\infty)$ is a pseudo-length function and homogeneity is replaced by the condition that the homogenization of ℓ is positive,

$$\ell_{\mathrm{hom}}(g) := \lim_{n\to\infty} \frac{\ell(g^n)}{n} > 0, \qquad \forall g \neq e,$$

then one can replace ℓ by ℓ_{hom} and use Theorem 1.2 as a homogeneous length function on G, to conclude that G embeds into a Banach space as above. (This condition was studied in [15, Theorem 2.10(III)] in the special case when G is already known to be abelian.) In particular, such (pseudo-)length functions cannot exist on non-abelian groups.

However, if we weaken any of several conditions in the above formulations, then examples of such structures do, in fact, exist. Here we mention several such cases and discuss other related problems.

4.1. **Monoids and embeddings.** Our first weakening is to replace 'groups' by the more primitive structures 'monoids' or 'semigroups'. In this case, Robert Young (private communication) described to us non-abelian monoids with homogeneous, bi-invariant length functions: consider the free monoid $\mathrm{FMon}(X)$ on any alphabet X of size at least 2, with the 'edit' distance $d(v, w)$ between strings $v, w \in \mathrm{FMon}(X)$ being the least number of single generator insertions and deletions to get from v to w. The triangle inequality and positivity are easily verified, while homogeneity of the corresponding length function $\ell(x) := d(e,x)$ is trivial. Moreover, the metric $d(\cdot,\cdot)$ turns out to be bi-invariant:

$$d(gxh, gyh) = d(x,y) \quad \text{for all} \quad g,h,x,y \in \mathrm{FMon}(X).$$

This specializes to left- and right-invariance upon taking $g \in X$ and $h = e$, or $h \in X$ and $g = e$, respectively.

Note moreover that $\mathrm{FMon}(X)$ embeds into $\mathrm{FGp}(X)$ (the free group generated by X and X^{-1}, the collection of symbols defined to be inverses of elements of X), and in particular, is cancellative. While this trivially addresses the embeddability issue, notice that a more refined version of embeddability fails. Namely, by our main theorem, $\mathrm{FMon}(X)$ does not embed into any group in the category $\mathcal{C}_{\text{bi}-\text{inv, hom}}$ with cancellative semigroups with homogeneous bi-invariant metrics as objects and isometric semigroup maps as morphisms. Thus, one may reasonably ask what is a sufficiently small category in which the embeddability works.

Proposition 4.1. *Let $\mathcal{C}_{\text{bi}-\text{inv}}$ denote the category whose objects are cancellative semigroups with bi-invariant metrics, and morphisms are isometric semigroup maps. Then* $\mathrm{FMon}(X)$ *embeds isometrically into* $\mathrm{FGp}(X)$ *in* $\mathcal{C}_{\text{bi}-\text{inv}}$.

Proof. From above, $\mathrm{FMon}(X)$ is an object of $\mathcal{C}_{\text{bi}-\text{inv}}$; denote the metric by d_{FM}. One can check that $d_{FM}(w,w')$ equals the difference $l(w)+l(w')$ (the sum of the lengths in X) minus twice the length of the longest common (possibly non-contiguous) substring in w, w'.

We next claim $\mathrm{FGp}(X)$ is also an object of $\mathcal{C}_{\text{bi}-\text{inv}}$. Namely, for a word $w = x_1 x_2 \cdots x_m$ in the free group, we consider *non-crossing matchings* in w, i.e., sets M of pairs of letters in $\{1, 2, \dots m\}$ such that the following hold.

- If $(i,j) \in M$, then $i < j$ and $x_j = x_i^{-1}$.
- If $(i,j) \in M$ and $(k,l) \in M$, then either $(i,j) = (k,l)$ or i, j, k and l are distinct.
- If $i < k < j < l$ and $(i,j) \in M$, then $(k,l) \notin M$.

Given a matching M as above, consider the set $U = U(M)$ of indices k, $1 \leqslant k \leqslant m$ which are not part of a pair in M. Define the *deficiency* of the matching M as the cardinality of the set $U(M)$, and define the length $\ell(w)$ of the word w as the infimum of the deficiency over all non-crossing matchings in w. This length was previously studied in [6], including checking that it is well-defined on all of $\mathrm{FGp}(X)$; moreover, $\ell(w)$ equals the smallest number of conjugates of elements in $X \sqcup X^{-1}$ whose product is w. Now define $d_{FG}(w,w') := \ell(w^{-1}w')$. It is easy to see that ℓ is a conjugacy invariant length function.

We claim that $d_{FG} \equiv d_{FM}$ on $\mathrm{FMon}(X)$, which proves the result. It is easy to show that if two words in $\mathrm{FMon}(X)$ differ by a single insertion or deletion, then their distance in $\mathrm{FGp}(X)$ is at most one, hence exactly one. In the other direction, we claim that a non-crossing matching on $w^{-1}w'$, with w and w' containing only positive generators (in X), is just a 'rainbow', i.e. nested arches with one end in w^{-1} and the other in w'. But then $d_{FG}(w,w')$ equals $l(w) + l(w')$ minus twice the length of a common substring, which is maximal by the minimality of the deficiency. Hence $d_{FG}(w,w') = d_{FM}(w,w')$, completing the proof. $\square$

4.2. **Quasi-pseudo-length functions via quasimorphisms.** We see that Proposition 2.1 is also of interest for $c > 0$, as follows. A *quasi-morphism* on a group G is a map $f : G \to \mathbb{R}$ whose *defect* is bounded,

$$D(f) := \sup_{x,y \in G} |f(xy) - f(x) - f(y)| < +\infty.$$

Let us use the term *quasi-pseudo-length function* for a function $\ell : G \to [0, +\infty)$ which satisfies the relaxed triangle inequality (2.2) for some c. Then every quasi-morphism as above induces a quasi-pseudo-length function by setting

$$\ell(x) := |f(x)|, \tag{4.2}$$

where we can take $c = D(f)$.

A quasi-morphism is *homogeneous* if $f(x^n) = nf(x)$ for $n \in \mathbb{N}$. Replacing any quasi-morphism f by $\lim_{n\to\infty} f(x^n)/n$ results in a homogeneous quasi-morphism. (Note that this limit exists by the de Bruijn–Erdös generalization [3, Theorem 23] of Fekete's subadditive lemma.) A homogeneous quasi-morphism defines a quasi-pseudo-length function that is homogeneous.[5] In this case, Proposition 2.1 makes a rather trivial statement: a homogeneous quasi-morphism is bounded on commutators,

$$|f([x,y])| \leqslant 4D(f).$$

In fact, as observed in [1, Lemme 1.1], one can improve the constant 4 to 3 in this case.

Nevertheless, quasi-morphisms can be utilized to construct interesting quasi-pseudo-length functions, for example satisfying homogeneity on specific commutators. The following quasi-morphism is due to Brooks [5, Section 2]. For a given word w in the free group $\mathbf{F}_2$, written in reduced form, let $f_w : \mathbf{F}_2 \to \mathbb{R}$ be the function which assigns to every other $g \in \mathbf{F}_2$, also written in reduced form, the maximum number of times such that w occurs in g without overlaps, minus the analogous number of times that w^{-1} can maximally occur in g. Since $f_w(w^n) = nf_w(w)$, homogenizing this quasi-morphism and taking the absolute value results in a quasi-pseudo-length function that is homogeneous on the powers of w. For example with w being the commutator of the generators of $\mathbf{F}_2$, we see that although the quasi-pseudo-length function must be bounded on commutators by Proposition 2.1, it can nevertheless grow linearly on the powers of a fixed commutator.

[5]Of course, homogeneous quasi-morphisms must satisfy

$$\sup_{n\geqslant 1} |f(e) - f(g^n) - f(g^{-n})| = \sup_{n\geqslant 1} |f(e) - n(f(g) + f(g^{-1}))| < \infty,$$

whence $f(g^{-1}) = -f(g)$ for all $g \in G$. On the other hand, if $f \not\equiv 0$ then the corresponding homogeneous quasi-pseudo-length function $\ell := |f|$ never has bounded defect, since if $f(g) \neq 0$ then $\inf_{n\geqslant 0} \ell(e) - \ell(g^n) - \ell(g^{-n}) = \inf_{n\geqslant 0} -2n\ell(g) = -\infty$. So in a sense, the two concepts are 'morally different'.

4.3. **Finite balls in free groups.** From Proposition 2.1 and a standard compactness argument, we can establish the following local version of the theorem.

Theorem 4.3. *For any $\varepsilon > 0$ there exists $R \geqslant 4$ with the following property: if a, b are two elements of a group G, $B_{a,b}(R) \subset G$ is the collection of all words in a, b, a^{-1}, b^{-1} of length at most R (so in particular $B_{a,b}(R)$ contains $[a, b]$), and the map $\ell : B_{a,b}(R) \to [0, +\infty)$ is a "local pseudo-length function" which obeys the triangle inequality*

$$\ell(xy) \leqslant \ell(x) + \ell(y) \tag{4.4}$$

whenever $x, y, xy \in B_{a,b}(R)$, with equality when $x = y$, then one has

$$\ell([a, b]) \leqslant \varepsilon(\ell(a) + \ell(b)).$$

Proof. By pulling back to the free group $\mathbf{F}_2$ generated by a and b, we may assume without loss of generality that $G = \mathbf{F}_2$. Without loss of generality we may also normalize $\ell(a) + \ell(b) = 1$. If the claim failed, then one could find a sequence $R_n \to \infty$ and local pseudo-length functions $\ell_n : B_{a,b}(R_n) \to [0, +\infty)$ such that $\ell_n(a) + \ell_n(b) = 1$, but that $\ell_n([a, b]) \geqslant \varepsilon$. By the Arzela–Ascoli theorem, we can pass to a subsequence that converges pointwise to a homogeneous pseudo-length function $\ell : G \to [0, +\infty)$ such that $\ell([a, b]) \geqslant \varepsilon$, which contradicts Proposition 2.1. □

Remark 4.5. By carefully going through the arguments in the previous section, choosing n to be various small powers of R instead of sending n to infinity, one can extract an explicit value of R of the form $R = C\varepsilon^{-A}$ for some absolute constants $C, A > 0$; we leave the details to the interested reader.

On the other hand, for any finite R one can construct local pseudo-length functions $\ell : B(0, R) \to [0, +\infty)$ such that $\ell(x) > 0$ for all $x \in B(0, R) \backslash \{e\}$. One construction is as follows. Any two matrices $U_a, U_b \in SO(3)$ define a representation $x \mapsto U_x$ of the free group $\mathbf{F}_2$ in the obvious fashion. Every U_x is then a rotation around some axis in $\mathbb{R}^3$ by some angle $0 \leqslant \theta \leqslant \pi$ in one of the two directions around that axis; if U_a and U_b are sufficiently close to the identity, then the angle of rotation of U_x will be at most $\pi/2$ for all $x \in B(0, R)$. We set $\ell(x)$ to be equal to this angle of rotation. Also, if U_a, U_b are chosen generically, the representation will be faithful (this follows from the dominance of word maps on simple Lie groups such as $SO(3)$, see [2]), so $\ell(x) > 0$ for any non-identity x. From the triangle inequality for angles we thus have (4.4) whenever $x, y, xy \in B(0, R)$, with equality when $x = y$. Note that as one sends $R \to \infty$, the local pseudo-length functions constructed here converge to zero pointwise, so in the limit we do not get any counterexample to the main theorem.

4.4. **Quasi-norms and commutator lengths.** Our main result is a classification theorem for homogeneous pseudo-length functions. We will see next that even in the case of homogeneous quasi-length functions, i.e., where we replace $\ell(xy) \leqslant \ell(x) + \ell(y)$ by $\ell(xy) \leqslant \ell(x) + \ell(y) + c$ for a fixed constant $c \geqslant 0$, we get non-trivial restrictions on these functions which give the main result in the case when $c = 0$. Further, we see that the restrictions we obtain are essentially the strongest possible.

Recall that the *commutator length* $\mathrm{cl}(g)$ of a word in $[G, G]$ is the length k of the shortest expression $g = [a_1, b_1] \cdot [a_2, b_2] \cdots [a_k, b_k]$ of g as a product of commutators. The *stable commutator length* is defined as $\lim_{n\to\infty} \mathrm{cl}(g^n)/n$, where the limit exists by sub-additivity of the function $n \mapsto \mathrm{cl}(g^n)$.

Then Proposition 2.1, together with (2.2), (2.6), easily imply the following estimates:

Proposition 4.6. *Let ℓ, c, c' be as in Proposition 2.1. Then for $g \in [G, G]\backslash\{e\}$, $\ell(g) \leqslant 5(c + c')\,\mathrm{cl}(g) - c$ and $\ell(g) \leqslant 5(c + c')\,\mathrm{scl}(g) + c'$.*

Observe that in various cases where we do not have a (quasi-)homogeneous quasi-norm, we get much weaker bounds. For example, consider the word $[a^k, b^m]$ in the free group $\mathbf{F}_2$, generated by a and b, for some integers k and m.

- The norm of such an element with respect to the word metric is $2(|k| + |m|)$.
- If we have a length function ℓ which is symmetric and conjugation-invariant (but not assumed to be homogeneous), then we have the bound $\ell([a^k, b^m]) \leqslant 2 \min(|k|\,\ell(a), |m|\,\ell(b))$. Furthermore, there are conjugation-invariant length functions for which these inequalities hold with equality.

  Namely, fix positive real numbers $\ell(a)$ and $\ell(b)$; now given a word $w = x_1 \cdots x_m$, consider the set $U = U(M)$ of indices k, $1 \leqslant k \leqslant m$ which are not part of a pair in a matching M; here M is defined as in the proof of Proposition 4.1. Now define the deficiency of M as the sum $\sum_{i\in U} \ell(x_i)$, and as in *loc. cit.*, define $\ell(w)$ to be the infimum of the deficiency over all non-crossing matchings in w. Then ℓ is a conjugation-invariant length function. Further, $\ell([a^k, b^m]) \geqslant 2 \min(|k|\,\ell(a), |m|\,\ell(b))$ as, for any matching M for $w = [a^k, b^m]$, if some pair (i, j) corresponds to letters a and a^{-1}, then no pair corresponds to letters b and b^{-1} and conversely. Further, it is easy to find a matching for which the deficiency is $\min(|k|\,\ell(a), |m|\,\ell(b))$. On the other hand, ℓ is not homogeneous; for instance, $\ell([a, b]) = 2$ and $\ell([a, b]^3) = 4$. Similarly, we have $\ell([a^k, b^k]) = 2|k|$ and $\ell([a^k, b^k]^3) \leqslant 4|k|$, which demonstrates that $2\ell(x) - \ell(x^2)$ is unbounded (as must be the case, according to (2.15)).
- On the other hand, the function ℓ associating to each word the length of its cyclically reduced form is homogeneous but not a quasi-length function. For this we have $\ell([a^k, b^m]) = 2(|k| + |m|)$.

Observe that all of the bounds on $\ell([a^k, b^m])$ here become unbounded as $k, m \to \infty$. This should be compared with Proposition 2.1, which establishes a bound $\ell([a^k, b^m]) \leqslant 4c + 5c'$ that is uniform in k and m for any function ℓ satisfying the hypotheses of that proposition.

Now we show that the bound in Proposition 2.1 is essentially sharp for every finitely generated group G. By assumption, also the abelianization $G_{\mathrm{ab}} := G/[G, G]$ is finitely generated. For simplicity of notation, assume G_{ab} is torsion-free. Let x_1, x_2, $\ldots$, x_k be the generators of $G_{\mathrm{ab}} \cong \mathbb{Z}^k$ and let $\varphi : G \to G_{\mathrm{ab}}$ be the abelianization map. Let $\alpha_i : G \to \mathbb{Z}$ for $i = 1, \ldots, k$ be the homomorphisms such that for all $g \in G$,

$$\varphi(g) = \sum_{i=1}^{k} \alpha_i(g) x_i.$$

Choose and fix elements $y_i \in G$ such that $\varphi(y_i) = x_i$. We define a function $\psi : G \to \mathbb{R}$ by

$$\psi(g) = \mathrm{cl}(g y_1^{-\alpha_1(g)} y_2^{-\alpha_2(g)} \cdots y_k^{-\alpha_k(g)}).$$

Repeatedly using relations of the form $\mathrm{cl}(ba) = \mathrm{cl}(ba[a^{-1}, b^{-1}]) \leqslant \mathrm{cl}(ab) + 1$, we see that

$$\psi(g_1 g_2) \leqslant \psi(g_1) + \psi(g_2) + (2k + 1),$$

i.e., ψ is a quasi-pseudo-length function. Further, by similar arguments we see that

$$|\psi((xy)^n) - \psi(x^n y^n)| \leqslant n,$$

from which we deduce that on replacing ψ by its homogenization (via [3, Theorem 23]), we still get a quasi-pseudo-length function.

Therefore, we can define functions satisfying Proposition 2.1 which are homogeneous by considering a length function $\ell_{G_{\mathrm{ab}}}$ on G_{ab} and taking the linear combination

$$\ell(g) := \ell_{G_{\mathrm{ab}}}(\varphi(g)) + \frac{c}{2k + 1} \psi(g).$$

This gives a homogeneous function satisfying the hypotheses of Proposition 2.1 while being strictly positive on all nontrivial elements.

References

[1] C. Bavard, *Longueur stable des commutateurs*, L'Enseignement Mathématique, 37(1–2):109–150, 1991.

[2] A. Borel, *On free subgroups of semisimple groups*, L'Enseignement Mathématique, 29(1–2):151–164, 1983.

[3] N.G. de Bruijn and P. Erdös, *Some linear and some quadratic recursion formulas. II*, Indagationes Mathematicae, 55:152–163, 1952.

[4] F. Cabello Sánchez and J.M.F. Castillo, *Banach space techniques underpinning a theory for nearly additive mappings*, Dissertationes Mathematicae (Rozprawy Matematyczne), 404:1–73, 2002.

[5] K. Fujiwara, *Quasi-homomorphisms on mapping class groups*, Handbook of Teichmüller theory. Vol. II, 241–269, IRMA Lect. Math. Theor. Phys., 13, Eur. Math. Soc., Zürich, 2009.

[6] S. Gadgil, *Watson–Crick pairing, the Heisenberg group and Milnor invariants*, Journal of Mathematical Biology 59(1), Article:123, 2009.

[7] Z. Gajda and Z. Kominek, *On separation theorems for subadditive and superadditive functionals*, Studia Mathematica, 100(1):25–38, 1991.

[8] T. Gowers and M. Nielsen, *Massively collaborative mathematics*, Nature, 461(7266):879–881, 2009.

[9] A. Khare and B. Rajaratnam, *The Hoffmann-Jørgensen inequality in metric semigroups*, Annals of Probability, 45(6A):4101–4111, 2017.

[10] A. Khare and B. Rajaratnam, *The Khinchin–Kahane inequality and Banach space embeddings for metric groups*, preprint (arXiv:1610.03037), 2016.

[11] S. Kobayashi, *Hyperbolic Manifolds and Holomorphic Mappings*, World Scientific, 1970.

[12] M. Ledoux and M. Talagrand, *Probability in Banach Spaces (Isoperimetry and Processes)*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer-Verlag, Berlin-New York, 1991.
[13] J. Milnor, *Curvatures of left invariant metrics on lie groups*, Advances in Mathematics, 21(3):293–329, 1976.
[14] A. Naor, *An introduction to the Ribe program*, Japan Journal of Mathematics, 7:167–233, 2012.
[15] P. Niemiec, *Universal valued Abelian groups*, Advances in Mathematics, 235:398–449, 2013.

HTTP://MICHAELNIELSEN.ORG/POLYMATH1/INDEX.PHP