Adrien Patoz

Contents

1 Introduction **2**

2 Group theory preliminaries **2**
2.1 Group presentation 2
2.2 Cayley graph 3
2.3 Some proprieties of the Cayley graph 4

3 Quasi-isometries **5**
3.1 Quasi-isometries 5
3.2 Quasi-inverses 5
3.3 Cayley graphs and quasi-isometries 6
3.4 A digression on growth 7

4 Divergence in a group **8**

5 Divergence in groups with a law **12**

1 Introduction

Acknowledgments I would like to thank Alessandro Sisto for his supervision throughout my internship, he taught me a lot about geometric group theory and pointed me to some really interesting papers which led me to understand geometric group theory better.
I would also like to thank Antoine Goldsborough who gave me my internship project and discussed it with me.
Outline of my internship My internship took place in the Heriot-Watt University of Edinburgh, between February and July.
I started my internship knowing very little about geometric group theory, so I read a lot on the subject, mainly "office hours with a geometric group theorist" and the note of the course given by Laura Ciobanu, Alexandre Martin and Alessandro Sisto at the Heriot-Watt University. Later on, Antoine Goldsborough gave me a project to work, namely, proving the theorem
Theorem (Cornelia Drutu, Mark Sapir, '07) *Let G be a non virtually cyclic group satisfying a law, then G has linear divergence.*
geometrically in order to extend the ideas to probabilistic divergence on groups with a probabilistic law.

2 Group theory preliminaries

This chapter is here to collect various group theory notions and facts.

2.1 Group presentation

Definition 1. *Let G be a group and $S \subseteq G$ a subset, then we denote by $\langle S \rangle$ the subgroup generated by S.*
And we denote by $\langle\langle S \rangle\rangle$ the normal subgroup generated by S, namely:

$$\langle\langle S \rangle\rangle := \left\{gsg^{-1}, s \in \langle S \rangle, g \in G\right\}$$

We will also need to define free groups:

Definition 2. *Let A a set, the free group generated by A, noted $F(A)$ is defined as follow:*
Take an abstract copy of A that we call A^{-1} and denote its elements by a^{-1} for $a \in A$.
$F(A)$ is defined as the set of words on $A \sqcup A^{-1}$ up to the equivalence relation $aa^{-1} = a^{-1}a = 1$ and with operation the concatenation.

With these definitions in mind, let us define group presentation:

Definition 3. *Let S be a set and R a set of words in $A \sqcup A^{-1}$, called relators. We say that a group has presentation $\langle S \mid R \rangle$ if it is isomorphic to $F(S)/\langle\langle R \rangle\rangle$.*

For example, $F(S)$ has presentation $\langle S \mid \varnothing \rangle$.
A more interesting example is that $\mathbb{Z}^2$ has presentation $\langle \{a, b\} \mid \{aba^{-1}b^{-1}\} \rangle$. We will prefer the notation $\langle a, b \mid aba^{-1}b^{-1} \rangle$ or $\langle a, b \mid ab = ba \rangle$ for readability purposes.

2.2 Cayley graph

The purpose of this section is to introduce a metric space on which a finitely generated group acts naturally, namely the cayley graph.
In this section, G will always denote a group generated by a finite set $S \subseteq G \backslash \{1\}$ such that $S = S^{-1}$. These two conditions on S are not very deep but will be more convenient for our definitions.
We start by defining a metric on any given graph:

Definition 4. *Let Γ be a graph, recall that Γ is defined by points called vertices and copies of $[0,1]$ connecting pairs of vertices, called edges. Suppose that we assign to each edge e a length $l(e) > 0$, then we can define a pseudo-metric on Γ by doing the following:*
Fix an homeomorphism $\varphi_e : e \to [0,1]$ for every edge e, and define $\rho : \Gamma^2 \to \mathbb{R} \cup \{+\infty\}$ such as:
If x and y are on the same edge e, then $\rho(x,y) = l(e) \cdot |\varphi_e(x) - \varphi_e(y)|$, else $\rho(x,y) = +\infty$, then

$$d(x,y) = \inf_{x=x_0,x_1,\ldots,x_n=y} \sum_{k=0}^{n-1} \rho(x_k, x_{k+1})$$

Is a pseudo-metric on Γ.

The only reason d is not a metric is because we can have $d(x,y) = 0$ with $x \neq y$. However it can't be the case when there is a lower bound on $l(e)$, for instance if we chose $l(e) = 1$ for every edge, then d is a distance and Γ has a metric space structure.
We can now define Cayley graphs:

Definition 5. *The Cayley graph $\mathrm{Cay}(G,S)$ of G with respect to S is the metric graph with:*

1. *vertex set G*
2. *an edge connection g, h if and only if $g^{-1}h \in S$, i.e. if and only if $g = hs$ for some $s \in S$.*
3. *all edges have length 1.*

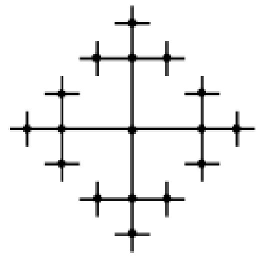
We denote d_S the metric on $\mathrm{Cay}(G,S)$

Fact 6. *This is where our conditions on S become relevant, they imply that there is no edge between a vertex and itself and that if there is an edge between g and h, then there is one between h and g, therefore the Cayley graph $\mathrm{Cay}(G,S)$ can be seen as a non oriented graph.*

Here are some examples of Cayley graphs:

1. $\mathrm{Cay}(\mathbb{Z}, \{\pm 1\})$ is isometric to $\mathbb{R}$
2. Changing the generating set changes the Cayley graph, for example here is $\mathrm{Cay}(\mathbb{Z}, \{\pm 2, \pm 3\})$:

3. $\mathrm{Cay}(\mathbb{Z}/n\mathbb{Z}, \{\pm 1\})$ is a n-gon
4. $\mathrm{Cay}(\mathbb{Z}^2, \{(0, \pm 1), (\pm 1, 0)\})$ is isometric to the grid in $\mathbb{R}^2$
5. $\mathrm{Cay}(F(\{a,b\}), \{a^{\pm 1}, b^{\pm 1}\})$ is a tree and looks like this :



Fact 7. *For* $g, h \in G$, $d_S(g,h) = \min\{n, g^{-1}h = s_1 \dots s_n, s_1, \dots, s_n \in S\}$ *i.e. the word length of* $g^{-1}h$ *in the alphabet* S.

2.3 Some proprieties of the Cayley graph

The goal of this section is to show that the Cayley graph has "nice" geometric proprieties.

Fact 8. G *acts on* $\mathrm{Cay}(G,S)$ *by isometries. Such action extends the action of* G *on itself by left translation.*

Indeed, there is an edge between h_1 and h_2 if and only if there is an edge between gh_1 and gh_2 (because $(gh_1)^{-1}gh_2 = h_1^{-1}h_2$).

Fact 9. $\mathrm{Cay}(G,S)$ *is a geodesic metric space*

A geodesic is the equivalent of a straight line in a general metric space, here are the proper definitions:

Definition 10. *Let* X *be a metric space and* $\alpha : [0,1] \to X$ *a path in* X, *we define the length of* α *noted* $l(\alpha)$ *as*

$$l(\alpha) = \inf_{t_0=0,\dots,t_n=1} \sum_{k=0}^{n-1} d(\alpha(t_k), \alpha(t_{k+1}))$$

By triangular inequality (and an immediate induction), $l(\alpha) \geq d(\alpha(0), \alpha(1))$, note that in $\mathbb{R}^n$ the equality case of this inequality happens when α is a straight line, which brings us to the following definition:

Definition 11. α *is a geodesic if* $l(\alpha) = d(\alpha(0), \alpha(1))$.
The metric space X *is called geodesic if for all* $x, y \in X$, *there exist a geodesic* α *such that* $\alpha(0) = x$ *and* $\alpha(1) = y$.

Back to Cayley graphs, let's look more closely at the action of G on $\mathrm{Cay}(G,S)$.

Proposition 12. *The action of* G *is on* $\mathrm{Cay}(G,S)$ *is proper and cobounded.*

This comes naturally because the graph has a bounded number of edges starting at each vertex. To sum up this introduction, we proved the following theorem:

Theorem 13. *Every finitely presented group acts properly and coboundedly by isometries on a proper, geodesic metric space. An example of such an action is the natural action on a Cayley graph.*

3 Quasi-isometries

3.1 Quasi-isometries

We will want to study groups rather than pairs group /generating set, however Cayley graphs depends on the generating set as shown in examples 1 and 2. In this section we answer the question as to how much the Cayley graph depends on the generating set.
To answer it we need to introduce quasi-isometries:

Definition 14. *Let X, Y be metric spaces and $f : X \to Y$ a map, we say that f is a (K, C) quasi-isometric embedding if for any $x, y \in X$, we have*

$$\frac{d(x, y)}{K} - C \leq d(f(x), f(y)) \leq Kd(x, y) + C$$

The (K, C) quasi-isometric embedding f is a (K, C) quasi-isometry if for any $y \in Y$, there is some $x \in X$ such that $d(f(x), y) \leq C$.

We say that f is a quasi-isometric embedding if it is a (K, C) quasi-isometric embedding for some K, C, same goes for quasi-isometry. As we will see later on, the constants do not really matter. Here are some examples:

1. for $v, b \in \mathbb{R}^2, v \neq 0$, the map $t \mapsto vt + b$ is a quasi-isometric embedding
2. $x \mapsto x^2$ from $\mathbb{R}$ to $\mathbb{R}$ is not a quasi-isometric embedding, the second inequality fails.
3. $\text{Cay}(\mathbb{Z}^2, \{(0, \pm 1), (\pm 1, 0)\})$ can be naturally embedded in $\mathbb{R}^2$, such embedding is a quasi-isometry.

3.2 Quasi-inverses

Definition 15. *Let $f : X \to Y$ be a map between metric spaces, we say that $g : Y \to X$ is a quasi-inverse of f if there is some $D \geq 0$ such that for any $x \in X$, $d_X(g \circ f(x), x) \leq D$ and for any $y \in Y$, $d_Y(y, f \circ g(y)) \leq D$.*

Note that a quasi-inverse is just an inverse "up to bounded error".
Here are some useful proprieties of quasi-isometric embedding and quasi-isometries:

Proposition 16. *1. Composition of quasi-isometric embedding (resp quasi-isometries) is a quasi-isometric embedding (resp quasi-isometry).*

2. Let f be a quasi-isometric embedding, then f is a quasi-isometry if and only if it has a quasi-inverse, moreover, this quasi-inverse is also a quasi-isometry.

3. Being quasi-isometric is an equivalence relation.

Proof. We will only prove point 2 as 1 is a direct consequence of the definitions and " is 1 and 2 combined.
For the implication $\Leftarrow$ we need to show that for any y there is x such that $f(x)$ is close to y, taking $x = g(y)$ is sufficient (where g is the quasi-inverse of f).
Let us prove $\Rightarrow$. Suppose that f is a (K, C) quasi-isometry for some K, C. By definition, for all

$y \in Y$, there is a $x \in X$ such that $d(f(x), y) \leq C$ define $g(y)$ as one of those x.
By definition, $d(f(g(y)), y) \leq C$ for any y, let us bound $d(g(f(x)), x)$:

$$d(g(f(x)), x) \leq Kd(f(g(f(x))), f(x)) + KC \leq 2KC$$

The first inequality is just a re-writing of the definition of quasi-isometric embedding while the second one comes from the fact that $f \circ g$ is C-close to the identity. So g is a quasi-inverse of f, let us prove that it is also a quasi-isometry:
Let D be as in the definition of quasi-inverse, we have that for any $y_1, y_2 \in Y$,

$$d(g(y_1), g(y_2)) \leq Kd(f(g(y_1)), f(g(y_2))) + KC \leq Kd(y_1, y_2) + 2KD + KC$$

The second inequality comes from the fact that $f(g(y_i))$ is D-close to y_i. The proof of the other inequality is similar. □

3.3 Cayley graphs and quasi-isometries

We are now ready to show that "the" Cayley graph of a group is well defined up to quasi-isometry:

Theorem 17. *Let G be a group and S, S' two finite, symmetric generating sets of G. The identity $G \to G$ extends to a quasi-isometry* $\mathrm{Cay}(G, S) \to \mathrm{Cay}(G, S')$.

Proof. Let us first reduce the problem to considering the vertex set of the Cayley graphs.
Consider the composition

$$\mathrm{Cay}(G, S) \overset{\psi}{\to} (G, d_S) \overset{\mathrm{id}}{\to} (G, d_{S'}) \overset{\iota}{\to} \mathrm{Cay}(G, S'),$$

where ψ is any map mapping $x \in \mathrm{Cay}(G, S)$ to $g \in G$ such that $d_S(x, g) \leq \frac{1}{2}$, and ι is just the inclusion.
Note that both ψ and ι are $(1, 1)$ quasi-isometries, so the overall compositionis a quasi-isometry if $(G, d_S) \overset{\mathrm{id}}{\to} (G, d_{S'})$ is.
The identity is surjective so we only need to prove that it is a quasi-isometric embedding, we will show that more strongly it is a bi-Lipschitz map.
Recall that $|g|_S := d_S(1, g)$ is the minimal number of generators from S needed to write g (and similarly for S'), set

$$M = \max\{|x'|_S, |x|_{S'}, x' \in S', x \in S\}$$

Now if $d_S(g, h) = k$, we can write $g^{-1}h = s_1 \ldots s_k$. And we can write each s_i as a product of s_i''s:

$$s_1 \ldots s_k = (s'_{1,1} \ldots s'_{1,M_1}) \ldots (s'_{k,1} \ldots s'_{k,M_k})$$

by definition of M, we have that $M_i \leq M$ for all i, therefore we have the bound:

$$d_{S'}(g, h) = |g^{-1}h|_{S'} \leq \sum_{i=1}^{k} M_i \leq Mk = Md_S(g, h)$$

The inequality $d_S \leq Md_{S'}$ follows from the same argument. □

3.4 A digression on growth

In this section we will introduce a simple to define quasi-isometry invariant.
Let G be generated by finite set S. We define $\beta_{G,S}(n)$ as the cardinality of the ball of radius n in (G, d_S), we call it the *growth function* of G with respect to S.
Here are some examples of growth function (with respect to the standard generating set):

1. $\beta_{\mathbb{Z},\{\pm 1\}}(n) = 2n + 1$
2. $\beta_{\mathbb{Z}^2,S}(n) = 2n^2 + 2n + 1$
3. $\beta_{F_m,S}(n) = 2m(2m-1)^{n-1}$ where F_m is the free group on m generators.

As usual, we are interested in groups, not pairs of group / generating set, therefore we define (a partial order and) an equivalence relation on the collection of functions.
Let $f, g : \mathbb{N} \to \mathbb{N}$, write $f \preceq g$ if there exist $C > 0$ such that for each $n \in \mathbb{N}$ we have

$$f(n) \leq Cg(Cn + C)$$

Further we write $f \asymp g$ if $f \preceq g$ and $g \preceq f$, we will also write $f \prec g$ if $f \preceq g$ but $g \not\preceq f$. $\asymp$ defined as such is an equivalence relation.
Here are some comparisons of functions to bear in mind:

- If $0 < a < b$ then $n^a \prec n^b$
- If $1 < \alpha < \beta$ then $\alpha^n \asymp \beta^n$
- for each $a > 0$, $n^a \prec 2^n$

So, at least we can distinguish exponents of polynomial growth, and we can distinguish polynomial functions from exponential functions.
The definition of $\asymp$ that we gave is enough to ensure that the $\asymp$ equivalence class of the growth function does not depend on the generating set, but the (stronger) following fact is true:

Proposition 18. *The $\asymp$-class of the growth function is a quasi-isometry invariant of groups.*

Proof. Let G, H be two groups generated by S and T respectively, let $f : G \to H$ be a (K, C) quasi-isometric embedding.
The image of the ball $B^G(1, n)$ of radius n is contained in a ball of radius $Kn + C$ in H.
The preimage of any element of H is contained in a ball of radius at most $KC + 1$, so the number of elements in the preimage of any element of H is at most $M > 0$ (a fixed constant).
Combining these two facts, we obtain that

$$\#B^G(1, n) \leq M \cdot \#B^H(1, Kn + C)$$

The other inequality comes from the same ideas used on the quasi-inverse of f. □

Let us mention a theorem due to Gromov on growth function:

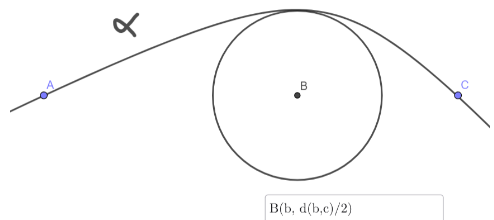
Theorem 19. *A finitely generated group has at most polynomial growth if and only if it is virtually nilpotent.*

In particular, being virtually nilpotent (which is a purely algebraic property) is a quasi-isometry invariant.
From now on, we will almost always say that G has a geometric property if the (any of its) Cayley graph has said property.

4 Divergence in a group

Definition 20. *Let X be a metric space, $a, b, c \in X$ and $\delta > 0$, we call divergence of a and c with respect to b the infimum:*

$$\mathrm{div}_b(a, c; \delta) := \inf\{l(\alpha), \alpha \text{ path from } a \text{ to } c \text{ avoiding } B(b, \delta \cdot \min(d(a, b), d(b, c))\}$$



For a fixed point $b \in X$, we set

$$\mathrm{div}_b(r; \delta) := \sup_{a,c \in D(b,r)} (\mathrm{div}_b(a, c; \delta))$$

And finally let

$$\mathrm{div}(r; \delta) := \sup_{b \in X} (\mathrm{div}_b(r; \delta))$$

The divergence of X.

Here are some examples of divergences that we can already compute:

1. In $\mathbb{R}$, for any $x < y < z$, $\mathrm{div}_y(x, z; 1/2) = \infty$.
2. In $\mathbb{R}^2$, for any $x, y, z \in \mathbb{R}^2$, $\mathrm{div}_y(x, z; 1/2)$ is at most $3d(x, z)$ (take a square with a side being $[x, y]$, the ball can not disconnect it).
3. In an hyperbolic group, the divergence is exponential.

We will come back to 1. later.
We did not give a precise computation in point 2., here is why:
As always we will want to compute the "divergence" of a group rather than a pair group / generating set, so the divergence has to be a quasi-isometry invariant, fortunately it is, up to $\asymp$ equivalence, with that in mind it is natural to say that $\mathbb{R}^2$ has linear divergence, as the computation $d(x, z) \leq \mathrm{div}_y(x, z; 1/2) \leq 3d(x, z)$ gives us that $\mathrm{div}_{\mathbb{R}^2}(r; 1/2) \asymp r$. Using the same idea, the group $\mathbb{Z}^2$ has linear divergence because $\mathrm{Cay}(\mathbb{Z}^2, \{(\pm 1, 0), (0, \pm 1)\}$ can be embedded in $\mathbb{R}^2$.
Point 3. is a bit more tricky to prove, the idea is to split the path in 2 and use the fact that hyperbolic groups are δ-thin for some $\delta > 0$ and iterate this. This example is mainly here to show that there are some groups that have "superlinear" divergence, i.e. some groups G for which $\mathrm{div}_G(n, \delta) \succ n$.
Next is a collection of basic properties that will help us compute some divergences.

Lemma 21. *For any $x, y, z \in X$ and any isometry $f : X \to X$, we have*

$$\operatorname{div}_y(x, z; \delta) = \operatorname{div}_{f(y)}(f(x), f(z); \delta)$$

Proof. If α is a path from x to z avoiding $B(y, \delta r)$, then $f \circ \alpha$ is a path from $f(x)$ to $f(z)$ avoiding the ball $f(B(y, \delta r)) = B(f(y), \delta r)$, taking the infimum of the length of such path, we get that $\operatorname{div}_y(x, z; \delta) \leq \operatorname{div}_{f(y)}(f(x), f(z); \delta)$ the other inequality comes from the fact that f^{-1} is an isometry that sends $(f(x), f(y), f(z))$ to (x, y, z). □

This fact will come in handy as $x \mapsto gx$ is an isometry of the Cayley graph of G for all $g \in G$, which means that we will (almost always) fix the center point to be 1 without loss of generality.

Lemma 22. *If x, y, z are on the same sphere around a point a, then $\operatorname{div}_a(x, z, \delta) \leq \operatorname{div}_a(x, y; \delta) + \operatorname{div}_a(y, z; \delta)$ for all $\delta \geq 0$.*

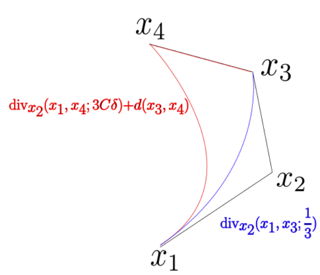
Proof. The concatenation of any path from x to y avoiding $B(a, \delta d(a, x))$ and any path from y to z avoiding $B(a, \delta d(a, y))$ is a path between x and z avoiding $B(a, \delta d(a, x)) = B(a, \delta d(a, z))$. Taking the infimum on those paths gives us the inequality we want. □

Lemma 23. *If x, z are on the same sphere around a point a, and y is any other point, then for all $\delta \geq 0$ we have $\operatorname{div}_a(x, z, \delta') \leq \operatorname{div}_a(x, y; \delta) + \operatorname{div}_a(y, z; \delta)$, where $\delta' = \delta \min\{d(a, x), d(a, z)\} / \max\{d(a, x), d(a, z)\}$.*

Proof. The proof is roughly the same as last lemma, the constant δ changes so that the balls that we want to avoid are exactly the same. □

For the same reason that the triangular inequality is nice, we would like to be able to compare $\operatorname{div}_{x_2}(x_4, x_1; 3C\delta)$ and $\operatorname{div}_{x_2}(x_3, x_1; \delta)$ if x_3 and x_4 are close enough. It is actually possible but requires some conditions:

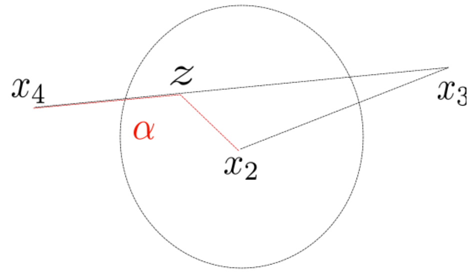
Lemma 24. *Let $\delta \leq 1/3$. For all $x_1, x_2, x_3, x_4 \in G$ satisfying $d(x_2, x_4) \leq C d(x_2, x_3)$ for some $C \geq 1$ we have the following: if $\operatorname{div}_{x_3}(x_2, x_4; \frac{1}{3}) > 4(C+1)d(x_2, x_3)/3$ then $\operatorname{div}_{x_2}(x_4, x_1; 3C\delta) \geq \operatorname{div}_{x_2}(x_3, x_1; \delta) - d(x_3, x_4)$.*



Proof.

Claim 1. *We have that $[x_3, x_4] \cap B\left(x_2, \frac{d(x_2, x_3)}{3}\right) = \emptyset$*

Assume, for a contradiction, that there is a point $z \in [x_3, x_4] \cap B\left(x_2, \frac{d(x_2, x_3)}{3}\right)$). Let $\alpha = [z, x_4] \cup [z, x_2]$, for all $t \in \alpha$, $d(x_3, t) \geq d(x_2, x_3)/3$. Indeed, if $t \in [z, x_4]$ then $d(x_3, t) \geq d(x_3, z) \geq$



$d(x_2, x_3) - d(x_2, z) \geq 2d(x_2, x_3)/3$. If $t \in [z, x_2]$ then $d(x_3, t) \geq d(x_3, z) - d(z, t) \geq d(x_2, x_3)/3$. Thus, $\alpha \cap B\left(x_2, \frac{d(x_2,x_3)}{3}\right) = \emptyset$, so the length of α is at least $\mathrm{div}_{x_3}(x_2, x_4; \frac{1}{3})$, so by assumption, the length of α is strictly greater than $4(C+1)d(x_2, x_3)/3$. On the other hand, it is at most $d(z, x_4) + d(z, x_2) \leq 4d(x_3, x_4)/3 \leq 4(C+1)d(x_2, x_3)$, a contradiction, which proves the claim.
If there is no path from x_4 to x_1 avoiding the ball $B(x_2, 3C\delta \min\{d(x_2, x_4), d(x_2, x_1)\})$ then $\mathrm{div}_{x_2}(x_4, x_1; 3C\delta) = \infty$ and we are done.
Hence assume there is a such path β. We now show that

$$3C\delta \min\{d(x_2, x_4), d(x_2, x_1)\} \geq \delta \min\{d(x_2, x_3), d(x_2, x_1)\}$$

Clearly $3C\delta d(x_2, x_1) \geq \delta \min\{d(x_2, x_3), d(x_2, x_1)\}$, thus is remains to show that $3C\delta d(x_2, x_4) \geq \delta \min\{d(x_2, x_3), d(x_2, x_1)\}$.
If $d(x_2, x_3) \leq d(x_2, x_1)$, then by 1, $d(x_2, x_4) \geq d(x_2, x_3)/3$ and so

$$\delta \min\{d(x_2, x_3), d(x_2, x_1)\} = \delta d(x_2, x_3) \leq 3C\delta d(x_2, x_4)$$

If $d(x_2, x_3) > d(x_2, x_1)$ then

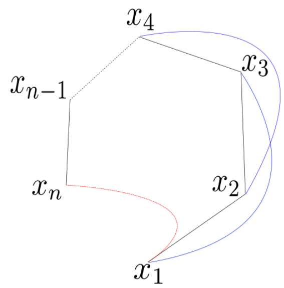
$$\delta \min\{d(x_2, x_3), d(x_2, x_1)\} = \delta d(x_2, x_1) \leq \delta d(x_2, x_4) \leq 3C\delta d(x_2, x_4)$$

Hence, as β avoids $B(x_2, 3C\delta \min\{d(x_2, x_4), d(x_2, x_1)\})$, it avoids $B(x_2, \delta \min\{d(x_2, x_3), d(x_2, x_1)\})$. We also have $d(x_2, x_3)/3 \geq \delta \min\{d(x_2, x_3), d(x_2, x_1)\}$. Hence the path given by $[x_3, x_4] \cup \beta$ avoids the ball $B(x_2, \delta \min\{d(x_2, x_3), d(x_2, x_1)\})$. Therefore, $l(\beta) + d(x_3, x_4) \geq \mathrm{div}_{x_2}(x_3, x_1; \delta)$. As β was chosen to be any path joining x_4 to x_1 avoiding $B(x_2, 3C\delta \min\{d(x_2, x_4), d(x_2, x_1)\})$, this proves the lemma. □

And a quick induction on the number of points gives us

Lemma 25. *Let $n > 3$ and $x_1, \dots, x_n \in G$ such that $2/3d(x_i, x_{i+1}) \leq d(x_1, x_2) \leq 3/2d(x_i, x_{i+1})$ for all $i \in \{1, \dots, n-1\}$ and $\mathrm{div}_{x_{i+1}}(x_i, x_{i+2}; 1/3) > 4d(x_1, x_2) + (n-3)d(x_1, x_2)$ for all $i \in \{1, \dots, n-2\}$, then $\mathrm{div}_{x_2}(x_1, x_n, (15/2)^{n-2}\delta) > \mathrm{div}_{x_2}(x_1, x_3, \delta) - (n-3)d(x_1, x_2)$*

Proof. We will prove this by induction on n, the case $n = 4$ is a consequence of lemma 24.
Suppose the lemma is true for $n > 3$, let $x_1, \dots, x_{n+1} \in G$ satisfying the conditions of the lemma. By induction hypothesis $\mathrm{div}_{x_2}(x_1, x_n, (15/2)^{n-2}\delta) > \mathrm{div}_{x_2}(x_1, x_3, \delta) - (n-3)d(x_1, x_2)$. We want to



apply lemma 24 on $(x_1, x_2, x_3, x_4) = (x_1, x_2, x_n, x_{n+1})$, we have that $d(x_2, x_{n+1}) \leq (1+3/2)d(x_2, x_n)$ and

$$\begin{aligned}
\mathrm{div}_{x_2}(x_1, x_n; 1/3) &> \mathrm{div}_{x_2}(x_1, x_n, (15/2)^{n-2}\delta) \\
&> \mathrm{div}_{x_2}(x_1, x_3, \delta) - (n-3)d(x_1, x_2) \\
&> 4d(x_1, x_2) \\
&\geq 4d(x_n, x_{n+1})
\end{aligned}$$

So we can apply the lemma with $C = 5/2$ and get:

$$\begin{aligned}
\mathrm{div}_{x_2}(x_1, x_{n+1}, (15/2)^{n-1}\delta) &> \mathrm{div}_{x_2}(x_1, x_n, (15/2)^{n-2}\delta) - d(x_n, x_{n+1}) \\
&\geq \mathrm{div}_{x_2}(x_1, x_n, \delta) - (n-3)d(x_1, x_2) - d(x_n, x_{n+1}) \\
&\geq \mathrm{div}_{x_2}(x_1, x_3, \delta) - (n-2)d(x_1, x_2)
\end{aligned}$$

□

Back to example 1., we see that this "one dimensional" case will be something that we will want to avoid, let's define it properly:

Definition 26. *Let X be a locally compact, connected metric space. For any compact $K \subseteq X$, denote $e(X, K)$ the number of connected composants of $X \setminus K$. Then, the number of "ends" of X, denoted $e(X)$ is the supremum*

$$e(X) = \sup_{K \subseteq X\, compact} (e(X, K))$$

Note that the number of ends of a metric space is a quasi-isometry invariant so we can talk about the number of ends of a group.
Basically, the number of ends of a group is the number of "directions" you can go in its Cayley graph, for instance, it is 2 for $\mathbb{Z}$, $+\infty$ for $\mathbb{Z}^2$, 1 for any finite group. We will see that this trichotomy is the general case:

Lemma 27. *A group has either 1, 2 or infinitely many ends.*

As promised, groups that "look like" $\mathbb{Z}$ have to be treated differently:

Lemma 28. *A group that has 2 ends is virtually cyclic.*

This fact is very technical to prove. We want an element x that has infinite order, the way to chose such an element is to take it along a geodesic ray that is fully contained in one end.

5 Divergence in groups with a law

This section will be the main result of my internship, first we need to define what a law is:

Definition 29. *Let $w(x_1, \dots, x_n)$ be a word in $x_1, \dots, x_n$, we say that the group G has law w if for all $g_1, \dots, g_n \in G$, $w(g_1, \dots, g_n) = 1$.*

Here are some examples of groups with a law:

1. $\mathbb{Z}/n\mathbb{Z}$ satisfies the law $n \cdot x = 0$.
2. more generally, Burnside groups are defined as satisfying the law $x^k = 1$ for some k.
3. any abelian group satisfies the law $[a, b](= aba^{-1}b^{-1}) = 1$

We are going to be working with groups with a law, so we would like to have a better understanding of the laws in general, the goal of the next to results is to simplify the law enough to be able to work with them easily.

Lemma 30. *If a group satisfies a law, it satisfies a 2 letters law.*

Proof. Suppose the law is given by a word $w(x_1, \dots, x_n)$ with $n > 2$, let $k = |w|$ the length of w. Then $w'(x_1, \dots, x_{n-1}) := w(x_1, x_2, \dots, x_{n-1}, x_2^k x_1 x_2^k)$ is a $n-1$ letter word. It is non trivial because it has the same number of occurrences of x_1 as $w(x_1, \dots, x_n)$, indeed, let's look at 2 consecutive occurrences of x_1 in $w'(x_1, \dots, x_{n-1})$:

- if there is no letter between them they come from consecutive occurrences of x_1 in w and so do not cancel each other.
- if there is only x_2 between them, then either they come from 2 consecutive x_n in w and cannot cancel each other or they come from a subword $x x_2^{\pm m} x'$ with $x, x' \in \{x_1, x_n\}$ in which case the number of x_2 separating them is either $\pm m$, $\pm k + \pm m$, $\pm k \pm k \pm m$ which cannot be 0 because $k > m + 1 > 0$.
- if there are other letters between these x_1, then they cannot be canceled by adding some x_2 and so the x_1 cannot cancel each other.

and the law given by $w' = 1$ is satisfied. By induction a 2 letters law is satisfied.
If the law only has 1 letter, it is of the form $x^k = 1$, and then $(xy)^k = 1$ is also a satisfied law on the group. □

Remark 31. *The proof of 30 shows that if a group satisfies a law, then it satisfies a law of the form $w(x, y) = x^{\pm n_1} y^{\pm 1} x^{\pm n_2} y^{\pm 1} \dots y^{\pm 1} x^{\pm n_k}$ for some $k \in \mathbb{N}$ and some $n_1, \dots, n_k \in \mathbb{N}$.*

We will now give 2 versions of the theorem shown in the introduction, the first one is weaker than the second one (we just add an hypothesis so that the proof works) but the idea of the proof is interesting enough on its own to be shared.
The hypothesis that we will add is that our group is finitely presented, meaning it is isomorphic to $\langle S|R \rangle$ with both S and R finite. This hypothesis gives us the following lemma:

Lemma 32. *Let G be a finitely presented group, there exists $M \geq 0$ such that for all $x, y \in G$ such that $d(1, x) = d(1, y)$, there exists a path from x to y in $B(1, d(1, x) + M) \setminus B(1, d(1, x) - M)$*

With that in mind, we are ready to prove the weak version of our theorem:

Theorem 33. *Let G be a finitely presented non virtually cyclic group satisfying a law. Then G has linear divergence.*

Proof. Let G be a group with super-linear divergence, let us prove that G cannot satisfy a law. By remark 31 it suffices to prove that for all $n \in \mathbb{N}$ and $k_1, \ldots, k_n \in \mathbb{Z}$ there exists $x, y \in G$ such that $x^{k_1} y x^{k_2} y \ldots y x^{k_n} \neq 1$. Fix $n \in \mathbb{N}$ and $k_1, \ldots, k_n \in \mathbb{Z}$ and let $w(a,b) := a^{k_1} b a^{k_2} b \ldots b a^{k_n}$, let $l(w)$ be the length of w.
By assumption, there exist $x, y \in G$ such that $d(1,x) = d(1,y)$ and $\mathrm{div}_1(x, y; 1/3 \cdot \delta_0) > 2l(w)d(1,x)$.
Let us assume that the following statement is true.
There exists an element $z \in D(1, d(1,x))$ such that $\mathrm{div}_1(x, z; 1/3 \cdot \delta_0) > 2l(w)d(1,x)$ and $\mathrm{div}_1(z, y; 1/3 \cdot \delta_0) > 2l(w)d(1,x)$
In that case, we use lemma 25 on $x_i = w|_i (xz, yz)$ the restriction of w to the i first letters. The condition on the distances is given because by assumption x, y and z are on the same sphere around 1, and we have that

$$\mathrm{div}_{x_{i+1}}(x_i, x_{i+2}; 1/(3 \cdot \delta_0)) = \mathrm{div}_1((w(x,y))_i^{-1}, (w(x,y))_{i+1}; 1/(3 \cdot \delta_0))$$

where $(w(x,y))_i$ denotes the i-th letter of the word $w(x,y)$. Therefore we only need to have that $\mathrm{div}_1(a^{-1}, b; 1/(3 \cdot \delta_0))$ is big enough for every consecutive letters a and b in $w(x,y)$. Given the form of the law we have that consecutive letters in $w(xz, yz)$ are always of the form $a^{\pm 1}, b^{\pm 1}$ with $a, b \in \{x, y, z\}, a \neq b$ and by choice of x, y, z, for any of these possibilities the divergence is big enough to use lemma 25, which gives us that:

$$\mathrm{div}_{(w(xz,yz))_1}(1, w(xz, yz); 1/3) > 0$$

which implies that $w(xz, yz) \neq 1$.
The reason we added the finitely presented condition is actually so that the statement is always true, here is one way to chose such a z:
Using the finitely presented hypothesis, we know that there exists $M > 0$ depending only on G such that there is a path from x to y in $B(1, d(1,x) + M) \setminus B(1, d(1,x) - M)$. Let x and y be such that $d(1,x) = d(1,y)$ is big enough so that $d(1,x) + M \leq 3/2 d(1,x)$ and $d(1,x) \leq 3/2(d(1,x) - M)$.
Let $\alpha : [0,1] \to G$ be a path from x to y in $B(1, d(1,x) + M) \setminus B(1, d(1,x) - M)$, for all $t \in [0,1]$, we have that

$$\mathrm{div}_1(x, \alpha(t); 1/(3 \cdot \delta_0)) + \mathrm{div}_1(\alpha(t), y; 1/(3 \cdot \delta_0)) \geq$$

$$\mathrm{div}_1(x, y; 1/(3 \cdot \delta_0) \cdot (d(1,x) + M)/d(1,x)) \geq \mathrm{div}_1(x, y; 1/(3 \cdot \delta_0)).$$

So by continuity of the divergence there exists $t_0 \in [0,1]$ such that $\mathrm{div}_1(x, \alpha(t_0); 1/(3 \cdot \delta_0)) > l(w)d(1,x)$ and $\mathrm{div}_1(y, \alpha(t_0); 1/(3 \cdot \delta_0)) > l(w)d(1,x)$. Let $z = \alpha(t_0)$.

□

The main idea of the proof is basically to "follow the word" and concatenate divergences, but in order to do that we seem to need to have at least 3 points far away from each other, next we will see how to deal with the situation where every ball has exactly two "clusters", meaning that for every radii n, there exist $x_n, y_n \in D(1,n)$ such that for every point $z \in D(1,n)$, $\mathrm{div}_1(x_n, z; \delta) \leq 2l(w)n$ or $\mathrm{div}_1(y_n, z; \delta) \leq 2l(w)n$.

Lemma 34. *Let G be a non virtually cyclic group, for each point $x \in G$ and every geodesic β passing through x, there exist a path $\alpha \subset G$ such that $\alpha(0) = x$ and for all $t \geq 0$, $d(\alpha(t), \beta) \geq t$.*

Proof. G being non virtually cyclic, there exist a point $y \in \beta$ and a path α_3 in G starting at y such that for a parametrization of β as α_1, α_2 with $\alpha_1(0) = \alpha_2(0) = y$, we have $d(\alpha_i(t), \alpha_j(t)) \geq t$ for all $t \geq 0$ and all $i \neq j \in \{1, 2, 3\}$.
There is an isometry of G that sends y to x, let γ_i the image of α_i.
Let $\beta(t)$ be a parametrization of β at speed 1 and such that $\beta(0) = x$.

Claim 2. *if $d(\gamma_i(t), \beta) \leq t/10$, then $\min(d(\gamma_i(t), \beta(t)), d(\gamma_i(t), \beta(-t))) \leq t/3$.*

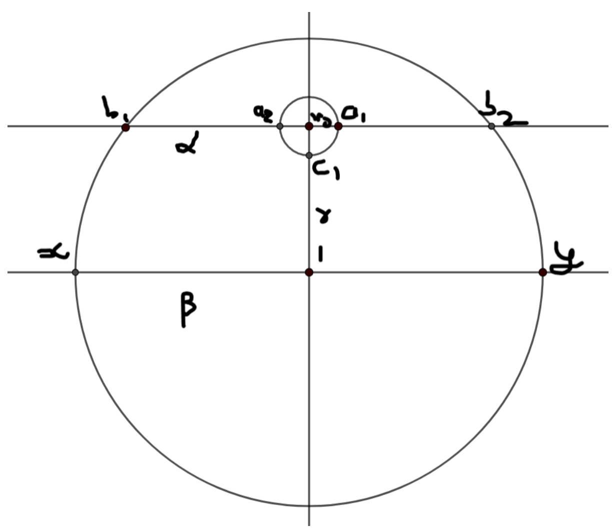
Using the claim, if for every i, $d(\gamma_i(t), \beta) \leq t/10$, then either there is $i \neq j$ such that $d(\gamma_i(t), \beta(t)) \leq t/3$ or there are $i \neq j$ such that $d(\gamma_i(t), \beta(-t))) \leq t/3$, for those i, j, we have that $d(\gamma_i(t), \gamma_j(t)) \leq 2t/3 < t$ which is a contradiction.
Therefore, there is an i such that $d(\gamma_i(t), \beta) \geq t/10$, and up to reparametrization of γ_i, there exist a path such that $d(\alpha(t), \beta) \geq t$ for all $t \geq 0$. □

Theorem 35. *Let G be a non virtually cyclic group satisfying a law. Then G has linear divergence.*

Proof. As mentioned before, if we have 3 points "far away" from each other, we can use a similar proof to theorem 33.
Therefore, let's assume that all $n > 0$, $B(1, n)$ has 2 clusters, i.e. there exists $x_n, y_n \in D(1, n)$ such that for all $z \in D(1, n)$, $\min(\text{div}_1(z, x; \delta), \text{div}_1(z, y; \delta)) \leq 4l(w) \cdot \min(d(z, x), d(z, y))$.
There is a geodesic β such that $D(1, n) \cap \beta_+$ and $D(1, n) \cap \beta_-$ are in two different clusters, for example take any geodesic passing through x and y.
By lemma 34, there exist a geodesic ray γ starting from 1, a point x_0 in $B(1, n) \cap \gamma$ and a ball $B(x_0, r)$ around x that doesn't intersect β or $D(1, n)$.
By assumption, this ball has exactly two clusters as well, similarly to the definition of β, take a geodesic α intersecting $D(x_0, r)$ two times such that $\alpha_- \cap D(x_0, r)$ and $\alpha_+ \cap D(x_0, r)$ are in two different clusters.
Let us define the following points:

$$a_1 = D(x_0, r) \cap \alpha_-, a_2 = D(x_0, r) \cap \alpha_+, b_1 = D(1, n) \cap \alpha_+, b_2 = D(1, n) \cap \alpha_-, \text{and } c_1 = \gamma \cap D(x_0, r)$$

c_1 is in a cluster on $D(x_0, r)$, without loss of generality, let say the cluster of a_2. Same goes for b_2, it is in a cluster from $D(1, n)$, let say the one of y.
Then there is a "short" path φ_1 from c_1 to a_2 avoiding $B(x_0, \delta \cdot r)$ and there is a short path φ_2 from b_2 to y avoiding $B(1, \delta \cdot n) \supseteq B(x_0, \delta \cdot r)$ for δ sufficiently close to 1. The path given by $[a_1, b_2] \cup \varphi_2 \cup [y, 1] \cup [1, c_1] \cup \varphi_1$ links a_1 and a_2 outside of the ball $B(x_0, \delta \cdot r)$, has length $d(a_1, b_2) + l(\varphi_2) + d(y, 1) + d(1, c_1) + l(\varphi_1) \leq n + 2l(w)n + n + n + 2l(w)r \leq r \cdot (3n/r + 2l(w)(1 + n/r)) \leq Cd(x_0, a_1)$ for a fixed C depending only on n which is a contradiction to the divergence $\text{div}_{x_0}(a_1, a_2; \delta)$ being superlinear.
Therefore, we can always find a point z that is far from either x or y divergence-wise. □



b1
a2
a1
b2
α
c1
γ
x
1
y
β