NOTES ON GEOMETRIC GROUP THEORY

WENYUAN YANG

Contents

Date: Dec 22 2020.

The basic references to this course are the following

(1) Bowditch, **A course on geometric group theory**, Mathematical Society of Japan, Tokyo, 2006.
http://www.warwick.ac.uk/~masgak/papers/bhb-ggtcourse.pdf
(2) CF Miller III, **Combinatorial group theory**,
http://www.ms.unimelb.edu.au/~cfm/notes/cgt-notes.pdf.

A comprehensive treatment to hyperbolic groups among many other things:

- M. Bridson and A. Haefliger, **Metric spaces of non-positive curvature**, vol. 319, Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 1999.

A undergraduate textbook discussing many interesting examples:

- J.Meier, **Groups, graphs and trees: An introduction to the geometry of infinite groups**, Cambridge University Press, Cambridge, 2008.

Here are more references you might be interested in.
The original monograph of M. Gromov introducing hyperbolic groups:

- M. Gromov, **Hyperbolic groups**, in "Essays in Group Theory" (G. M. Gersten, ed.), MSRI Publ. 8, 1987, pp. 75-263

Three early collaborative works to understand the above Gromov 1987 monograph:

(1) H. Short (editor), **Notes on hyperbolic groups**, Group theory from a geometrical viewpoint, World Scientific Publishing Co., Inc., 1991.
https://www.i2m.univ-amu.fr/~short/Papers/MSRInotes2004.pdf
(2) É. Ghys and P. de la Harpe (editors), **Sur les groupes hyperboliques d'aprs Mikhael Gromov**. Progress in Mathematics, 83. Birkhuser Boston, Inc., Boston, MA, 1990. xii+285 pp. ISBN 0-8176-3508-4
(3) M. Coornaert, T. Delzant and A. Papadopoulos, **Géométrie et théorie des groupes: les groupes hyperboliques de Gromov**, Lecture Notes in Mathematics, vol. 1441, Springer-Verlag, Berlin, 1990, x+165 pp.

1. Review of group theory

1.1. Group and generating set.

Definition 1.1. A group $(G, \cdot)$ is a set G endowed with an operation

$$\cdot : G \times G \to G, (a, b) \to a \cdot b$$

such that the following holds.

(1) $\forall a, b, c \in G, a \cdot (b \cdot c) = (a \cdot b) \cdot c$,
(2) $\exists 1 \in G$: $\forall a \in G$, $a \cdot 1 = 1 \cdot a$.
(3) $\forall a \in G$, $\exists a^{-1} \in G$: $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

In the sequel, we usually omit $\cdot$ in $a \cdot b$ if the operation is clear or understood. By the associative law, it makes no ambiguity to write abc instead of $a \cdot (b \cdot c)$ or $(a \cdot b) \cdot c$.

Examples 1.2. (1) $(\mathbb{Z}^n, +)$ for any integer $n \ge 1$
(2) General Linear groups with matrix multiplication: $GL(n, \mathbb{R})$.
(3) Given a (possibly infinite) set X, the permutation group $Sym(X)$ is the set of all bijections on X, endowed with mapping composition.
(4) Dihedral group $D_{2n} = \langle r, s | s^2 = r^{2n} = 1, srs^{-1} = r^{-1} \rangle$. This group can be visualized as the symmetry group of a regular $(2n)$-polygon: s is the reflection about the axe connecting middle points of the two opposite sides, and r is the rotation about the center of the $2n$-polygon with an angle $\pi/2n$.
(5) Infinite Dihedral group $D_\infty = \langle r, s | s^2 = 1, srs^{-1} = r^{-1} \rangle$. We can think of a regular ∞-polygon as a real line. Consider a group action of D_∞ on the real line.

Definition 1.3. A subset H in a group G is called a *subgroup* if H endowed with the group operation is itself a group. Equivalently, H is a subgroup of G if

(1) $\forall a, b \in H, a \cdot b \in H$
(2) $\forall a \in H$, $\exists a^{-1} \in H$: $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

Note that (1) and (2) imply that the identity 1 lies in H.

Given a subset $X \subset G$, the *subgroup generated by* X, denoted by $\langle X \rangle$, is the minimal subgroup of G containing X. Explicitly, we have

$$\langle X \rangle = \{x_1^{\epsilon_1} x_2^{\epsilon_2} \cdots x_n^{\epsilon_n} : n \in \mathbb{N}, x_i \in X, \epsilon_i \in \{1, -1\}\}.$$

A subset X is called a *generating set* of G if $G = \langle X \rangle$. If X is finite, then G is called *finitely generated*.

Check Examples 1.2 and find out which are finitely generated, and if yes, write a generating set.

Exercise 1.4. (1) *Prove that* $(\mathbb{Q}, +)$ *is not finitely generated.*
(2) *Prove that* $\{r, rsr^{-1}\}$ *is a generating set for* D_∞.

Exercise 1.5. (1) *Suppose that* G *is a finitely generated group. If* $H \subset G$ *is of finite index in* G*, then* H *is finitely generated.*
(2) *Conversely, suppose that* H *is a finite index subgroup of a group* G*. If* H *is finitely generated, then* G *is also finitely generated.*

Exercise 1.6. *Let* N *be a normal subgroup of a group* G*. Suppose that* N *and* G/N *are finitely generated. Then* G *is finitely generated.*

1.2. **Group action.**

Definition 1.7. Let G be a group and X be set. A *group action of* G *on* X is a function

$$G \times X \to X, (g, x) \to g \cdot x$$

such that the following holds.

(1) $\forall x \in X$, $1 \cdot x = x$.
(2) $\forall g, h \in G, (gh) \cdot x = g \cdot (h \cdot x)$.

Usually we say that G acts on X. Similarly, we often omit $\cdot$ in $g \cdot x$, but keep in mind that $gx \in X$, which is not a group element!

Remark. A group can act *trivially* on any set X by just setting $g \cdot x = x$. So we are mainly interested in nontrivial group actions.

Examples 1.8. (1) $\mathbb{Z}$ acts on the real line $\mathbb{R}$: $(n, r) \to n + r$.
(2) $\mathbb{Z}$ acts on the circle $S^1 = \{e^{i\theta} : \theta \in \mathbb{R}\}$: $(n, e^{\theta i}) \to e^{n\theta i}$. Here i is the imaginary unit.
(3) $\mathbb{Z}^n$ acts on $\mathbb{R}^n$.
(4) $GL(n, \mathbb{R})$ acts on $\mathbb{R}^n$.

Recall that $Sym(X)$ is the permutation group of X. We have the following equivalent formulation of a group action.

Lemma 1.9. *A group G acts on a set X if and only if there exists a group homomorphism $G \to Sym(X)$.*

Proof. (=>). Define $\phi : G \to Sym(X)$ as follows. Given $g \in G$, let $\phi(g)(x) = g \cdot x$ for any $x \in X$. Here $g \cdot x$ is given in definition of the group action of G on X.

It is an exercise to verify that $\phi(g)$ is a bijection on X. Moreover, the condition (2) in definition 1.7 is amount to say that ϕ is homomorphism.

(<=). Let $\phi : G \to Sym(X)$ be a group homomorphism. Construct a map $G \times X \to X$: $(g, x) \to \phi(g)(x)$. Then it is easy to see that this map gives a group action of G on X. □

So when we say a group action of G on X, it is same as specifying a group homomorphism from G to $Sym(X)$.

Remark. (1) A trivial group action is to specify a trivial group homomorphism, sending every element in G to the identity in $Sym(X)$.
(2) In general, the group homomorphism $G \to Sym(X)$ may not injective. If it is injective, we call the group action is *faithful*.
(3) In practice, the set X usually comes with extra nice structures, for example, X is a vector space, a topological space, or a metric space, etc. The homomorphic image of G in $Sym(X)$ may preserve these structures. In this case, we say that G acts on X by linear transformations, by homeomorphisms, or by isometries ...

We now recall Cayley's theorem, which essentially says that we should understand groups via group actions on sets with various good structures.

Theorem 1.10. *Every group is a subgroup of the permutation group of a set.*

Proof. Let $X = G$. Clearly the group operation $G \times G \to G$ gives a group action of G on G. Thus, we obtain a homomorphism $G \to Sym(G)$. The injectivity is clear. □

For any $x \in X$, the *orbit* of x under the group action is the set $\{g \cdot x : g \in G\}$. We denote it by $G \cdot x$ or even simply by Gx. The *stabilizer* of x

$$G_x := \{g \in G : g \cdot x = x\}$$

is clearly a group.

Lemma 1.11. *Suppose that G acts on X. Then for any x, there exists a bijection between $\{gG_x : g \in G\}$ and Gx. In particular, if Gx is finite, then $[G : G_x] = |Gx|$.*

Proof. We define a map $\phi : gG_x \to gx$. First, we need to show that this map is well-defined: that is to say, if $gG_x = g'G_x$, then $gx = g'x$. This follows from the definition of G_x.

For any $gx \in Gx$, we have $\phi(gG_x) = gx$. So ϕ is surjective.

To see that ϕ is injective, let $gG_x, g'G_x$ such that $gx = g'x$. Then $g^{-1}g'x = x$ and $g^{-1}g' \in G_x$. Hence $gG_x = g'G_x$. This shows that ϕ is injective. □

Exercise 1.12.
1. *Let H be a subgroup in G. Then $\cap_{g \in G}(gHg^{-1})$ is a normal subgroup in G.*
2. *Let H be a finite index subgroup of G. Then there exists a normal subgroup N of G such that $N \subset H$ and $[G : N] < \infty$. (Hint: construct a group action)*

Theorem 1.13 (M. Hall)**.** *Suppose G is a finitely generated group. Then for any integer $n > 1$, there are only finitely many subgroups H in G such that $[G : H] = n$.*

Proof. Fix n. Let H be a subgroup of index n. Let $X = \{H, g_1H, \cdots, g_{n-1}H\}$ be the set of all H-cosets. Then G acts on X of by left-multiplication. That is, $(g, g_iH) \to gg_iH$. Clearly, the stabilizer of $H \in X$ is $H \subset G$. Put in other words, the subgroup H can be recovered from the action of G on X.

For any set X with n elements, a finitely generated G has finitely many different actions on X. By Lemma 1.2, a group action is the same as a group homomorphism. A homomorphism is determined by the image of a generating set. As G is finitely generated and $Sym(X)$ is finite, there exist only finitely many group homomorphisms.

Consequently, for any $n > 0$, there exist only finitely many H of finite index n. □

1.3. **(Free) abelian groups.** Recall that a group G is called *abelian* if $ab = ba$ for any $a, b \in G$. In this subsection, we study finitely generated abelian group.

Definition 1.14. Let X be a set. The group $A(X) := \bigoplus_{x \in X} \langle x \rangle$ is called the *free abelian group* generated by X. The set X is called a *basis* of $A(X)$.

By definition, we see that there is an injective map $X \to A(X)$ defined by $x \to (0, ...0, x, 0, ...)$ for $x \in X$. Clearly, $A(X)$ is generated by (the image under the injective map) of X.

Let $m \in \mathbb{Z}$ and $a = (n_1x, ..., n_ix, ...) \in A(X)$. We define the scalar multiplication

$$m \cdot a = (mn_1x, ..., mn_ix, ...) \in A(X).$$

A *linear combination* of elements $a_i \in A(X), 1 \leq i \leq n$ is an element in $A(X)$ of the form $\sum_{1 \leq i \leq n} k_i \cdot a_i$ for some $k_i \in \mathbb{Z}, 1 \leq i \leq n$.

Exercise 1.15. (1) *Let Y be a subset in a free abelian group G of finite rank. Then Y is basis of G if and only if $G = \langle Y \rangle$ and any element in G can be written as a unique linear combination of elements in Y.*
(2) *Prove that the group of rational numbers $\mathbb{Q}$ is not free abelian.*

Exercise 1.16. *Prove that $\mathbb{Z}^m \cong \mathbb{Z}^n$ if and only if $m = n$.*

If $|X|$ is finite, then $|X|$ is called the *rank* of $A(X)$. In general, a free abelian group may have different basis. The rank of a free abelian group is well-defined, by Exercise 1.16.

Every abelian group is a quotient of a free abelian group.

Lemma 1.17. *Let X be a subset. For any map of X to an abelian group G, there exists a unique homomorphism ϕ such that*

$$\begin{array}{ccc} X & \to & A(X) \\ & \searrow & \downarrow \\ & & G \end{array}$$

is commutative.

Corollary 1.18. *Every abelian group is a quotient of a free abelian group.*

A free abelian group is characterized by the following universal mapping property in the category of abelian groups.

Lemma 1.19. *Let X be a subset, A be an abelian group and $X \to A$ be a map. Suppose that for any abelian group G and a map $X \to G$, there exists a unique homomorphism $\phi : A \to G$ such that*

$$\begin{array}{ccc} X & \to & A \\ & \searrow & \downarrow \\ & & G \end{array}$$

is commutative. Then $A \cong A(X)$.

Let $H \times K$ be a direct product of groups. For any $A \subset H, B \subset K$, the direct product $A \times B$ is a subgroup of $H \times K$. But conversely, not every subgroup in $H \times K$ arises in this way. For example, the subgroup generated by (a, a) in $\langle a \rangle \bigoplus \langle a \rangle$ cannot be written as a direct product of any subgroup in $\langle a \rangle$.

In free abelian groups, the following theorem says that, up to a change of a basis, every subgroup is a product of subgroups in each summand.

Theorem 1.20. *Let A be a free abelian group of finite rank n, and let H be a subgroup of A. Then there exists a basis $a_1, ..., a_n$ of A, and positive integers $k_1, ..., k_r$ with $0 \leq r \leq n, k_1 | k_2 | ... | k_r$, such that $k_1 a_1, ..., k_r a_r$ is a basis for H. In particular, H is free of rank $r \leq n$.*

Proof. Use induction on the rank of A. If $n = 1$, it is trivial. As all subgroups of $\mathbb{Z}$ are of form $n\mathbb{Z}$ for $n \geq 1$.

Given a basis $X = \{x_1, ..., x_n\}$ of A, every element $a \in A$ can be written as a finite linear combination of x_i. That is, $a = \sum_{1 \leq i \leq n} k_i x_i$, where $k_i \in \mathbb{Z}$.

Assume now that every subgroup of a free abelian group of rank $n - 1$ satisfies the conclusion.

Consider a subset S of integers in $\mathbb{Z}$, which consists of all possible coefficients k_i in linear combinations of any element in H for some basis X. Precisely,

$$S = \{k_1 \in \mathbb{Z} : \exists \text{ a basis } X \text{ of } A \text{ and } k_2, ..., k_n \in \mathbb{Z} \text{ such that } \sum_{1\le i\le n} k_i x_i \in H\}$$

Observe that $0 \in S$ and S is in fact a subgroup of $\mathbb{Z}$. Thus, $S = k_1\mathbb{Z}$ for some $k_1 > 0$.

By definition of S, there exists a basis X such that $y_1 = k_1x_1 + ... + k_nx_n \in H$. As $k_1|k_i$, there exists $d_i \in \mathbb{Z}$ such that $y_1 = k_1(x_1 + d_2x_2 + ...d_nx_n)$. Denote $\hat{x}_1 = x_1 + d_2x_2 + ...d_nx_n$. Then $\hat{X} = \{\hat{x}_1, x_2, ..., x_n\}$ is also a basis of A.

Clearly, we can write $A = A_1 \bigoplus A_2$, where $A_1 = \langle \hat{x}_1\rangle$ and $A_2 = \langle x_2, ..., x_n\rangle$.

For the basis $\hat{X}$, any element $h \in H$ can be written as a linear combination of $\hat{x}_1, x_2, ..., x_n$. As $y_1 = k_1\hat{x}_1$, any $h \in H$ is also a linear combination of $y_1, x_2, ..., x_n$. (But note that $\{y_1, x_2, ..., x_n\}$ may not be a basis of A!) Hence,

$$H = \langle y_1\rangle \bigoplus (H \cap A_2) = \langle k_1\hat{x}_1\rangle \bigoplus (H \cap A_2).$$

Note that A_2 is a free abelian group of rank $n-1$. Apply Induction Assumption to the subgroup $H \cap A_2$ of A_2. There exists a basis $\{\hat{x}_2, ..., \hat{x}_n\}$ of A_2 and $k_2|k_3|...|k_n$ such that $\{k_2\hat{x}_2, ..., k_n\hat{x}_n\}$ is a basis of $H \cap A_2$.

Since $\{\hat{x}_1, \hat{x}_2, ..., \hat{x}_n\}$ is a basis of A, we see $k_1|k_2|...|k_n$ by definition of $S = \langle k_1\rangle$. This finishes the proof. □

We can now classify all finitely generated abelian groups.

Theorem 1.21. *Any finitely generated abelian group is isomorphic to the following form*

$$\mathbb{Z} \bigoplus \mathbb{Z} \bigoplus ... \bigoplus \mathbb{Z} \bigoplus \mathbb{Z}_{k_1} \bigoplus \mathbb{Z}_{k_2} \bigoplus ... \bigoplus \mathbb{Z}_{k_r}$$

where $k_1|k_2|...|k_r$. *They are uniquely determined by the number of summand* $\mathbb{Z}$ *and integers* k_i.

The number of summand $\mathbb{Z}$ is called the *rank* of an abelian group.

Proof. By Lemma 1.17, there exists a free abelian group A of finite rank such that G is a quotient of A. Let $N < A$ be the kernel of the epimorphism. Then by the group isomorphism theorem, we have that $A/N \cong G$.

By Theorem 1.20, there exists a basis $\{a_1, ..., a_n\}$ of A and positive integers $k_1, ..., k_r$ with $0 \le r \le n, k_1|k_2|...|k_r$, such that $N = \langle k_1a_1\rangle + ... + \langle k_ra_r\rangle$.

It is known that for any two group G_1, G_2 and their normal subgroups $N_1 \subset G_1, N_2 \subset H_1$, we have

$$(G_1 \times G_2)/(N_1 \times N_2) \cong G_1/N_1 \times G_2/N_2.$$

The proof is completed by applying finitely many times the above result for A/N. □

Corollary 1.22. *Every subgroup in a finitely generated abelian group is finitely generated.*

Recall that an element is called *torsion element* if it is of finite order.

Corollary 1.23. *If a finitely generated abelian group has no torsion, then it is free abelian.*

Remark. The condition "finitely generated" is necessary, as $\mathbb{Q}$ has no torsion and is not free abelian.

2. Free groups and their subgroups

2.1. **Words and their reduced forms.** Let $\tilde{X}$ be an alphabet set. A *word* w over $\tilde{X}$ is a finite sequence of letters in $\tilde{X}$. We usually write $w = x_1x_2...x_n$, where $x_i \in \tilde{X}$. The *empty word* is the word with an empty sequence of letters. The *length* of a word w is the length of the sequence of letters.

Two words are *equal* if their sequences of letters are identical. Denote by $\mathcal{W}(\tilde{X})$ the set of all words over $\tilde{X}$. Given two words $w, w' \in (\tilde{X})$, the *concatenation* of w and w' is a new word, denoted by ww', which is obtained from w followed by w'.

Given a set X, we take another set X^{-1} such that there exists a bijection $X \to X^{-1} : x \to x^{-1}$. Let $\tilde{X} = X \sqcup X^{-1}$ be the disjoint union of X and X^{-1}. Roughly speaking, the free group $F(X)$ generated by X will be the set of words $\mathcal{W}$ endowed with the operation of word concatenation.

Given a word w, if there exists two consecutive letters of form xx^{-1} or $x^{-1}x$ where $x, x^{-1} \in \tilde{X}$, then we call xx^{-1} or $x^{-1}x$ an *inverse pair* of w. A word w is called *reduced* if w contains no inverse pair. Given a word w, we define an operation on w called a *reduction*, by which we mean deleting an inverse pair xx^{-1} or $x^{-1}x$ to obtain a new word w':

$$w = w_1xx^{-1}w_2 \xrightarrow{reduction} w' = w_1w_2.$$

After a reduction, the length of a word decreases by 2. A finite sequence of reductions

$$w \xrightarrow{reduction\#1} w_1 \xrightarrow{reduction\#2} w_2 ... \xrightarrow{reduction\#n} w_n$$

will be referred to as a reduction process.

Clearly, any word w admits a reduction process to get a reduced word. This reduced word is called a *reduced form* of w. But a word may have different reduction processes to become reduced. For example, $w = xx^{-1}xx^{-1}$. However, we will prove that reduced forms of a word does not depend on the reduction process.

Lemma 2.1. *Any word w has a unique reduced form.*

Proof. We prove the lemma by induction on the length $|w|$ of w. The base cases that $|w| = 1, 2$ are trivial. Now assume that the lemma holds for any word of length $|w| \le n$.

Let w be a word of length of n. Let

$$w \xrightarrow{reduction\#1} w_1 \xrightarrow{reduction\#2} w_2 ... \xrightarrow{reduction\#l} w_l$$

and

$$w \xrightarrow{reduction\#1'} w_1' \xrightarrow{reduction\#2'} w_2' ... \xrightarrow{reduction\#m'} w_m'$$

be any two reduction processes of w such that w_l, w_m' are reduced. We will show that $w_l = w_m'$.

We have the following claim.

Claim. *Suppose that $w_1 \neq w_1'$. Then there are two reductions*

$$w_1 \xrightarrow{reduction\ \#1} \hat{w}$$

and

$$w_1' \xrightarrow{reduction\ \#1'} \hat{w}'$$

such that $\hat{w} = \hat{w}'$.

Proof of Claim. Let xx^{-1} be the inverse pair for the reduction #1, and yy^{-1} the inverse pair for the reduction #1′. We have two cases.

Case 1. The inverse pairs xx^{-1}, yy^{-1} are disjoint in w. In this case, we let reduction a be reduction #1′, and reduction b be reduction #1. Thus, $\hat{w} = \hat{w}'$.

Case 2. The inverse pairs xx^{-1}, yy^{-1} have overlaps. Then either $x^{-1} = y$ or $y^{-1} = x$. In either cases, we have $w_1 = w_1'$. This contradicts the assumption that $w_1 \neq w_1'$. □

We are now ready to complete the proof of Lemma. First, if $w_1 = w_1'$, then $w_l = w_m'$ by applying the induction assumption to $w_1 = w_1'$ of length $n-2$. Otherwise, by the claim, there are two reductions applying to w_1, w_1' respectively such that the obtained words $\hat{w} = \hat{w}'$ are the same.

Note that $\hat{w}$ is of length $n-4$. Applying induction assumption to $\hat{w}$, we see that any reduction process

$$\hat{w} \xrightarrow{reduction\ process} \bar{w}$$

of $\hat{w}$ gives the same reduced form $\bar{w}$.

By the claim, the reduction a together any reduction process $\hat{w} \xrightarrow{reduction\ process} \bar{w}$ gives a reduction process for w_1 to $\bar{w}$. By induction assumption to w_1, we have $w_l = \bar{w}$. By the same reasoning, we have $w_m' = \bar{w}$. This shows that $w_l = w_m' = \bar{w}$. □

2.2. **Construction of free groups by words.** Denote by $F(X)$ the set of all reduced words in $\mathcal{W}(\tilde{X})$. By Lemma 2.1, there is a map

$$\mathcal{W}(\tilde{X}) \to F(X)$$

by sending a word to its reduced form.

We now define the group operation on the set $F(X)$. Let w, w' be two words in $F(X)$. The product $w \cdot w'$ is the reduced form of the word ww'.

Theorem 2.2. *$(F(X), \cdot)$ is a group with a generating set X.*

Proof. It suffices to prove the associative law for the group operation. Let w_1, w_2, w_3 be words in $F(X)$. We want to show $(w_1 \cdot w_2) \cdot w_3 = w_1 \cdot (w_2 \cdot w_3)$. By Lemma 2.1, the reduced form of a word does not depend on the reduction process. Observe that $(w_1 \cdot w_2) \cdot w_3$ and $w_1 \cdot (w_2 \cdot w_3)$ can be viewed as reduced forms of different reduction processes of the word $w_1 w_2 w_3$. The proof is thus completed. □

Let $\iota : X \to F(X)$ be the inclusion of X in $F(X)$. Usually we will not distinguish x and $\iota(x)$ below, as ι is injective.

Lemma 2.3. *For any map of a set X to a group G, there exists a unique homomorphism $\phi : F(X) \to G$ such that*

$$\begin{array}{ccc} X & \to & F(X) \\ & \searrow & \downarrow \\ & & G \end{array}$$

is commutative.

Proof. Let w_1, w_2 be two reduced words in $F(X)$. Without loss of generality, assume that $w_1 = x_1 x_2 ... x_n z_1 z_2 ... z_r$ and $w_2 = z_1^{-1} ... z_r^{-1} y_1 y_2 ... y_m$, where $x_i, y_j, z_k \in \tilde{X} = X \sqcup X^{-1}$ and $x_n \neq y_1^{-1}$. Then $w_1 \cdot w_2 = x_1 x_2 ... x_n y_1 y_2 ... y_m$.

Denote by j the map $X \to G$. Define $\phi(x) = j(x)$ for all $x \in X$ and $\phi(x^{-1}) = j(x)^{-1}$ for $x^{-1} \in X^{-1}$. Define ϕ naturally over other elements in $F(X)$. It is straightforward to verify that $\phi(w_1 \cdot w_2) = \phi(w_1)\phi(w_2)$.

Since a homomorphism of $F(X)$ to G is determined by the value of its restriction over a generating set of $F(X)$, we have that the chosen map $j : X \to G$ determines the uniqueness of ϕ. □

Corollary 2.4. *Every group is a quotient of a free group.*

Proof. Let X be a generating set of G. Let $F(X)$ be the free group generated by X. By Lemma 2.3, we have an epimorphism of $F(X) \to G$. □

Exercise 2.5. *Let X be a set containing only one element. Prove that $F(X) \cong \mathbb{Z}$.*

Analogous to free abelian group, the class of free groups is characterized by the following universal mapping property in GROUP category.

Lemma 2.6. *Let X be a subset, F be a group and $i : X \to F$ be a map. Suppose that for any group G and a map $j : X \to G$, there exists a unique homomorphism $\phi : F \to G$ such that*

$$\begin{array}{ccc} X & \xrightarrow{i} & F \\ & \searrow^{j} & \vdots \phi \\ & & G \end{array}$$

is commutative. Then $F \cong F(X)$.

Proof. By Lemma 2.3 for free group $F(X)$ and $i : X \to F$, there is a unique homomorphism $\varphi : F(X) \to F$ such that $i = \varphi\iota$, where $\iota : X \to F(X)$ is the inclusion map. ie.

$$\begin{array}{ccc} X & \xrightarrow{\iota} & F(X) \\ & \searrow^{i} & \vdots \varphi \\ & & F \end{array} \tag{1}$$

On the other hand, by the assumption to $G = F(X)$ and $\iota : X \to F(X)$, there is a unique homomorphism $\phi : F \to F(X)$ such that we have $\iota = \phi i$.

$$\begin{array}{ccc} X & \xrightarrow{i} & F \\ & \searrow^{\iota} & \vdots \phi \\ & & F(X) \end{array} \tag{2}$$

Thus we obtained $\iota = \phi\varphi\iota$, and the following commutative diagram follows from the above (1)(2).

$$\begin{array}{ccc} X & \xrightarrow{\iota} & F(X) \\ & \searrow^{\iota} & \vdots \phi\varphi \\ & & F(X) \end{array} \tag{3}$$

Note that the identification $Id_{F(X)}$ between $F(X) \to F(X)$ also makes the above diagram (3) commutative. By the uniqueness statement of Lemma 2.3, $\phi\varphi = Id_{F(X)}$.

It is analogous to prove that $\varphi\phi = Id_F$. Hence ϕ or φ is an isomorphism. □

Recall that the commutator subgroup $[G, G]$ of a group G is the subgroup in G generated by the set of all commutators. That is:

$$[G, G] = \langle \{[f, g] := fgf^{-1}g^{-1} : f, g \in G\} \rangle$$

Use universal mapping property of free groups and free abelian groups to prove the following.

Exercise 2.7. *Prove that $F(X)/[F(X), F(X)] \cong A(X)$, where $A(X)$ is the free abelian group generated by X.*

A subset Y is called a *basis* of $F(X)$ if $F(X) \cong F(Y)$. In this case, we often say that $F(X)$ is *freely generated* by X. Use Exercise 1.16 to prove the following.

Exercise 2.8. *If $|X| < \infty$ and Y is a basis of $F(X)$, then $|X| = |Y|$.*

The *rank* of $F(X)$ is defined to be the cardinality of X. By Exercise 2.8, the rank of a free group is well-defined: does not depend on the choice of basis.

When the rank is finite, we usually write $F_n = F(X)$ for $n = |X|$.

Convention. *Since there is a map $\mathcal{W}(\tilde{X}) \to F(X) \to G$ for a generating set X of G, we write $w =_G g$ for a word $w \in \mathcal{W}(\tilde{X}), g \in G$, if the image of w under the map $\mathcal{W}(\tilde{X}) \to G$ is the element g.*

Theorem 2.9. *Let G be a group with a generating set X. Then $G \cong F(X)$ if and only if any non-empty word $w \in \mathcal{W}(\tilde{X})$ with $w =_G 1 \in G$ contains an inverse pair.*

Proof. We have first a surjective map $\mathcal{W}(\tilde{X}) \to F(X) \to G$, where $F(X) \to G$ is the epimorphism given by Lemma 2.3.

=>. let $w \in \mathcal{W}(\tilde{X})$ be a word such that $w =_G 1$. Since $F(X) \cong G$, we have w is mapped to the empty word in $F(X)$. That is to say, the reduced form of W is the empty word. Thus, w contains an inverse pair.

<=. Suppose that $F(X) \to G$ is not injective. Then there exists a non-empty reduced word $w \in F(X)$ such that $w =_G 1$. Then w contains an inverse pair. As w is reduced, this is a contradiction. Hence $F(X) \to G$ is injective. □

Exercise 2.10.
1. *Let Y be a set in the free group $F(X)$ generated by a set X such that $y^{-1} \notin Y$ for any $y \in Y$. If any reduced word w over $\tilde{Y} = Y \sqcup Y^{-1}$ is a reduced word over $\tilde{X} = X \sqcup X^{-1}$, then $\langle Y \rangle \cong F(Y)$.*
2. *Let $S = \{b^n a b^{-n} : n \in \mathbb{Z}\}$ be a set of words in $F(X)$ where $X = \{a, b\}$. Prove that $\langle S \rangle \cong F(S)$.*
3. *Prove that for any set X with $|X| \geq 2$ any $n \geq 1$, $F(X)$ contains a free subgroup of rank n.*

2.3. **Ping-Pong Lemma and free groups in linear groups.** In this subsection, we gives some common practice to construct a free subgroup in concrete groups. We formulate it in Ping-Pong Lemma. Before stating the lemma, we look at the following example.

Lemma 2.11. *The subgroup of $\mathbb{SL}(2, \mathbb{Z})$ generated by the following matrices*

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

is isomorphic to F_2.

Proof. See Proposition 3.7, on page 59 in our reference [8]. □

Exercise 2.12. *The subgroup of* $\mathbb{SL}(2,\mathbb{C})$ *generated by the following matrices*

$$\begin{pmatrix} 1 & a_1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ a_2 & 1 \end{pmatrix}, |a_1| \geq 2, |a_2| \geq 2;$$

is isomorphic to F_2.

Lemma 2.13 (Ping-Pong Lemma)**.** *Suppose that* G *is generated by a set* S*, and* G *acts on a set* X*. Assume, in addition, that for each* $s \in \tilde{S} = S \sqcup S^{-1}$*, there exists a set* $X_s \subset X$ *with the following properties.*

(1) $\forall s \in \tilde{S}$, $s \cdot X_t \subset X_s$, *where* $t \in \tilde{S} \setminus \{s^{-1}\}$.
(2) $\exists o \in X \setminus \cup_{s \in \tilde{S}} X_s$, *such that* $s \cdot o \in X_S$ *for any* $s \in \tilde{S}$.

Then $G \cong F(S)$.

Proof. By Lemma 2.3 and Lemma 1.9, we have the following homomorphism:

$$\iota : F(S) \to G \to Sym(X).$$

Let w be a reduced non-empty word in $F(S)$. Write $w = s_1 s_2 ... s_n$ for $s_i \in \tilde{S}$. By Theorem 2.9, it suffices to show that $g = \iota(s_1)\iota(s_2)...\iota(s_n)$ is not an identity in $Sym(X)$.

We now apply the permutation g to $o \in X$ to get

$$g \cdot o = \iota(s_1)\iota(s_2)...\iota(s_{n-1})\iota(s_n) \cdot o \subset \iota(s_1)\iota(s_2)...\iota(s_{n-1}) X_{s_n} \subset ... \subset X_{s_1}.$$

However, as $o \in X_{s_1}$, we have $g \neq 1 \in Sym(X)$. This shows that $F(S) \cong G$. □

Ping-Pong Lemma has a variety of forms, for instance:

Exercise 2.14. *Let* G *be a group generated by two elements* a, b *of infinite order. Assume that* G *acts on a set* X *with the following properties.*

(1) *There exists non-empty subsets* $A, B \subset X$ *such that* A *is not included in* B.
(2) $a^n(B) \subset A$ *and* $b^n(A) \subset B$ *for all* $n \in \mathbb{Z} \setminus \{0\}$.

Prove that G *is freely generated by* $\{a, b\}$.

We now prove that $\mathbb{SL}(2,\mathbb{R})$ contains many free subgroups.

Proposition 2.15. *Let* $A \in \mathbb{SL}(2,\mathbb{R})$ *with two eigenvalues* λ, λ^{-1} *for* $\lambda > 1$*, and corresponding eigenvectors* $v_\lambda, v_{\lambda^{-1}}$*. Choose* $B \in \mathbb{SL}(2,\mathbb{R})$ *such that* $B\langle v_\lambda \rangle \neq \langle v_\lambda \rangle$, $B\langle v_\lambda \rangle \neq \langle v_{\lambda^{-1}} \rangle$ *and* $B\langle v_{\lambda^{-1}} \rangle \neq \langle v_\lambda \rangle$, $B\langle v_{\lambda^{-1}} \rangle \neq \langle v_{\lambda^{-1}} \rangle$.

Then there exist $N, M > 0$ *depending only on* A, B *such that*

$$F(a, b) = \langle a, b \rangle$$

where $a = A^n, b = BA^mB^{-1}$ *for* $n, > N, m > M$.

Proof. Observe that BAB^{-1} has the same eigenvalues λ, λ^{-1}, but eigenvectors $Bv_\lambda, Bv_{\lambda^{-1}}$ respectively.

Let $\theta \in (0, 2\pi)$ be a (very small) angle. Denote by $X_{v,\theta} \subset \mathbb{R}^2$ the open sector around the line $\langle v_\lambda \rangle$ with angle θ.

We claim the following fact about the dynamics of A on vectors.

Claim. $\forall \theta \in (0, 2\pi), \exists N > 0$ *such that the following holds.*

$$\text{For } \forall n > N, v \in \mathbb{R}^2 \setminus \langle v_\lambda^{-1} \rangle, \text{ we have } A^n v \in X_{v_\lambda, \theta}.$$

and

$$\textit{For } \forall n > N, v \in \mathbb{R}^2 \setminus \langle v_\lambda \rangle, \textit{ we have } A^{-n}v \in X_{v_\lambda^{-1},\theta}.$$

Proof of Claim. Since $\{v_\lambda, v_\lambda^{-1}\}$ is a basis of $\mathbb{R}^2$, the conclusion follows by a simple calculation. □

By the same reasoning, we also have

Claim. $\forall \theta \in (0, 2\pi), \exists M > 0$ *such that the following holds.*

$$\textit{For } \forall m > M, v \in \mathbb{R}^2 \setminus \langle Bv_\lambda^{-1} \rangle, \textit{ we have } BA^mB^{-1}v \in X_{Bv_\lambda,\theta}.$$

and

$$\textit{For } \forall m > M, v \in \mathbb{R}^2 \setminus \langle Bv_\lambda \rangle, \textit{ we have } BA^{-m}B^{-1}v \in X_{Bv_{\lambda^{-1}},\theta}.$$

Denote $a = A^n, b = BA^mB^{-1}, X_a = X_{v_\lambda,\theta}, X_a^{-1} = X_{v_\lambda^{-1},\theta}, X_b = X_{Bv_\lambda,\theta}, X_b^{-1} = X_{Bv_\lambda,\theta}$. Let $S = \{a, b\}$. By the above claims, we obtain the following.

$$\forall s \in \tilde{S},\ s \cdot X_t \subset X_s, \text{ where } t \in \tilde{S} \setminus \{s^{-1}\}.$$

Choose θ small enough such that $X_a \cup X_a^{-1} \cup X_b \cup X_b^{-1} \neq \mathbb{R}^2$. Choose any $o \in \mathbb{R}^2 \setminus \cup_{s\in\tilde{S}} X_s$. By the claims, $s \cdot o \in X_s$. Hence, all conditions of Ping-Pong Lemma are satisfied. We obtain that $F(\{a, b\}) = \langle a, b \rangle$. □

In fact, Jacques Tits proved the following celebrated result in 1972, which is usually called Tits alternative.

Theorem 2.16. *Let G be a finitely generated linear group. Then either G is virtually solvable or contains a free subgroup of rank at least 2.*

Remark. Note that a virtually solvable group does not contain any free group of rank at least 2. This explains the name of Tits alternative.

2.4. **Subgroups of free groups.** We shall give two proofs of the following theorem of Nielsen.

Theorem 2.17. *Any subgroup of a free group is free.*

The first proof is to consider group action on trees, and to use Ping-Pong Lemma. We first introduce a combinatorial formulation of the notion of a graph.

Definition 2.18. A *graph* $\mathcal{G} = (V, E)$ consists of a set V of vertices and a set E of directed edges. For each directed edge $e \in E$, we associate to e the *initial point* $e_- \in V$ and *terminal point* $e_+ \in V$. There is an orientation-reversing map

$$^{-}: E \to E, e \to \bar{e}$$

such that $e \neq \bar{e}$, $e = \bar{\bar{e}}$ and $e_- = (\bar{e})_+, e_+ = (\bar{e})_-$.

An *orientation* of $\mathcal{G}$ picks up exactly one directed edge in $\{e, \bar{e}\}$ for all $e \in E$. Formally, an orientation is a subset in E such that it contains exactly one element in $\{e, \bar{e}\}$ for all $e \in E$

Remark. Clearly, such a map $^-$ has to be bijective. Moreover, $e_+ = (\bar{e})_-$ can be deduced from other conditions: $e_+ = \bar{\bar{e}}_+ = \bar{e}_-$.

Remark. Every combinatorial graph can be geometrically realized by a common graph in the sense of CW-complex. We take the set of points V, and for each pair $(e, \bar{e})$, we take an interval $[0, 1]$ and attach its endpoints to $e_-, e_+ \in V$ respectively. Then we get a CW-complex.

Combinatorially, we define a *path* to be a concatenation of directed edges:

$$\gamma = e_1 e_2 ... e_n, e_i \in E$$

where $(e_i)_+ = (e_{i+1})_-$ for $1 \leq i < n$. The initial point γ_- and terminal point γ_+ of γ are defined as follows:

$$\gamma_- = (e_1)_-, \gamma_+ = (e_n)_+.$$

If $(e_n)_+ = (e_1)_-$, the path γ is called a *circuit* at $(e_1)_-$. By convention, we think of a vertex in $\mathcal{G}$ as a path (or circuit), where there are no edges.

A *backtracking* in γ is a subpath of form $e_i e_{i+1}$ such that $e_i = \bar{e}_{i+1}$. A path without backtracking is called *reduced*. If a path γ contains a backtracking, we can obtain a new path after deleting the backtracking. So any path can be converted to a reduced path by a reduction process. Similarly as Lemma 2.1, we can prove the following.

Lemma 2.19. *The reduced path is independent of the reduction process, and thus is unique.*

A *graph morphism* $\phi : \mathcal{G} \to \mathcal{G}'$ between two graphs $\mathcal{G}, \mathcal{G}'$ is a vertex-to-vertex, edge-to-edge map such that $\phi(e_-) = \phi(e)_-, \phi(e_+) = \phi(e)_+$ and $\phi(\bar{e}) = \overline{\phi(e)}$. It is called a *graph isomorphism* if ϕ is bijective.

By definition, a *tree* is a graph where every reduced circuit is a point. In other words, there exists a unique reduced path between two points.

Now lets consider the free group $F(S)$ over a set S. We define a tree $\mathcal{G}$ for which the vertex set V is all elements in $F(S)$. Two reduced words $W, W' \in F(S)$ are connected by an edge if there exists $s \in \tilde{S}$ such that $W' = Ws$. Formally, the edge set E is defined to be $F(S) \times \tilde{S}$. The map $^-$ sends $(W, s) \in F(S) \times \tilde{S}$ to $(Ws, s^{-1}) \in F(S) \times \tilde{S}$. Such a graph $\mathcal{G}$ is indeed a tree, and $F(S)$ acts on $\mathcal{G}$ by graph isomorphisms.

We shall use Ping-Pong Lemma to prove the following theorem, which implies Theorem 2.17.

Theorem 2.20. *Suppose that G acts on a tree T such that the stabilizer of each vertex is trivial. In other words, G acts on a tree T freely. Then G is a free group.*

Remark. In the proof, we understand the tree T as the geometric realization of the combinatorial notion of a tree. That is to say, we will not distinguish the edge e and $\bar{e}$, and in fact identify them to a single edge.

Proof. **Step 1. Find a fundamental set.** We consider a *fundamental set* X for the action of G on T. Roughly, X will be a connected subset such that it contains exactly one vertex from each orbit Gv for $v \in T$. Precisely, we define such a set in the following inductive way. Fix a basepoint $o \in T$. Let $X_0 = \{o\}$.

Suppose X_i is defined. We are going to define X_{i+1}. Consider a vertex $v \in T \setminus X_i$ which is connected by an edge e to some vertex in X_i. If v is not in the orbit $G \cdot X_i$, then we define X_{i+1} to be the union $X_i \cup e$. Finally, we define $X = \cup_{i \geq 0} X_i$. Then X is a connected set, and contains exactly one vertex from each G-orbit in T.

However, it is important to note that $G \times X$ may not contain all edges in T. In order that $G \times X = T$, we have to include some edges to X.

We denote by E_0 the set of edges e of T such that X contains exactly one endpoint of e. We also denote by e_- the endpoint of e in X, and e_+ the other endpoint of e outdid X. Define $\bar{X} = X \cup \tilde{S}$. Then $\bar{X}$ is still connected and $G \times \bar{X} = T$.

Step 2. Find free basis of G. For each $e \in E_0$, we know that $e_- \in X$ and $e_+ \notin X$. Recall that X contains (exactly) one vertex from each G-orbit in T. Thus, there exist an element $g_e \in G \setminus 1$ and a unique vertex $v \in X$ such that $g_e v = e_+$. The element g_e is unique, otherwise the stabilizer of v is nontrivial. This is a contradiction, since G acts on T freely.

Observe that $g_e^{-1}(e_-) \in T \setminus X$ is connected by the edge $g_e^{-1}(e)$ to $v \in X$. Denote $e' = g_e^{-1}(e)$. Thus, $e \neq e'$ and $e' \in E_0$. By the uniqueness of $g_{e'}$, we also see that $g_{e'} = g_e^{-1}$.

In conclusion, for each $e \in E_0$, there exists a unique $e \neq e' \in E_0$ and a unique $g_e \in G \setminus 1$ such that $g_e^{-1}(e) = e'$. Moreover, $g_{e'} = g_e^{-1}$.

Denote $\tilde{S} = \{g_e : e \in E_0\}$. Note that edges e, e' in E_0 are paired. From each such pair, we choose exactly one edge and denote them by $E_1 \subset E_0$. Define $S = \{g_e : e \in E_1\}$. Obviously, $\tilde{S} = S \cup S^{-1}$.

Step 3. Verify Ping-Pong Lemma. We now prove that $G = F(S)$ by using Ping-Pong Lemma.

For each $e \in E_0$, we define X_e to be the subgraph of T such that for each vertex z in X_v, there exists a (unique) reduced path from o to z containing the edge e. We note that X_e is connected and contains the endpoint e_+ of e. Moreover, $X_{e_1} \cap X_{e_2} = \emptyset$ for $e_1 \neq e_2 \in E_0$, and any path between two points in X_{e_1} and X_{e_2} respectively have to intersect X. This follows from the fact that T is a tree. So if a path γ intersects X_e but $\gamma \cap X = \emptyset$, then γ lies in X_e.

We first verify that $g_e(o) \in X_e$, where $e \in E_0$. We connect o and $g_e^{-1} e_+ \in X$ by a unique reduced path γ in X. Since X is a fundamental set, we have that $g_v \gamma \cap X = \emptyset$. Since $g_e \gamma$ contains the endpoint e_+ of e and $e_+ \in X_e$, we obtain that $g_e o \in g_e \gamma \subset X_e$.

Secondly, we prove that $g_e X_t \subset X_e$ for $t \neq e' \in E_0$. Indeed, for any $z \in X_t$, there exists a unique reduced path α between o, z containing the edge t. We also connect $g_e' o$ and o by a unique reduced path β. Since $g_e' o \in X_{e'}$, we know that β contains the edge e'. Consider the path $\gamma = \beta\alpha$. Since $e' \neq t$, we obtain that γ is a reduced path. The endpoints of the path $g_e \gamma$ are $\{o, g_e z\}$. Since $g_e \gamma$ contains the edge $g_e e' = e$, the endpoint $g_e z$ must lie in X_e by definition of X_e. So we proved that $g_e X_t \subset X_e$.

Therefore, we have verified the conditions of Ping-Pong Lemma 2.13. So $G = F(S)$. □

Exercise 2.21. *Suppose G acts by graph isomorphisms without inversions on a connected graph X such that there exists a finite subgraph K with $G \cdot K = X$. Assume that the edge stabilizers and the vertex stabilizers are finitely generated. Then G is finitely generated.*

2.5. **Fundamental groups of graphs.** The second proof of Theorem 2.17 is to use a combinatorial notion of fundamental groups of a graph.

The *concatenation* $\gamma\gamma'$ of two paths γ, γ' is defined in the obvious way, if $\gamma_+ = \gamma'_-$.

Definition 2.22. Let $\mathcal{G}$ be a graph and $o \in \mathcal{G}$ be a basepoint. Then the *fundamental group* $\pi_1(\mathcal{G}, o)$ of $\mathcal{G}$ consists of all reduced circuits based at o, where the group multiplication is defined by sending two reduced circuits to the reduced form of their concatenation.

The group identity in $\pi_1(\mathcal{G}, o)$ is the just the based point $o \in \mathcal{G}$.

Remark. We can consider an equivalence relation over the set of all circuits based at o: two circuits are *equivalent* if they have the same reduced form. By Lemma 2.19, this is indeed an equivalence relation. Then the fundamental group $\pi_1(\mathcal{G}, o)$ can be also defined as the set of equivalent classes $[\gamma]$ of all circuits based at o, endowed with the multiplication:

$$[\gamma] \times [\gamma'] \to [\gamma\gamma'].$$

It is easy to see that these two definitions give the isomorphic fundamental groups.

A particular important graph is the graph of a rose which consists of one vertex o with all other edges $e \in E$ such that $e_- = e_+ = o$. Topologically, the rose is obtained by attaching a collection of circles to one point.

Here we list a few properties about the fundamental group of a graph.

(1) We fix an orientation on a rose. Then the fundamental group of a rose is isomorphic to the free group over the orientation.
(2) Any graph contains a *spanning* tree which is a tree with the vertex set of the graph. We can collapse a spanning tree to get a rose. It is easy to see that the fundamental group of a graph is isomorphic to that of this rose.

A graph morphism $\phi : \mathcal{G} \to \mathcal{G}'$ naturally defines a homomorphism between the fundamental group as follows:

$$\phi_* : \pi_1(\mathcal{G}, o) \to \pi_1(\mathcal{G}', \phi(o))$$

by sending a reduced circuit γ in $\pi_1(\mathcal{G}, o)$ to the reduced path of $\phi(\gamma)$ in $\pi_1(\mathcal{G}', \phi(o))$.

Given a vertex v in $\mathcal{G}$, consider the star

$$Star_{\mathcal{G}}(v) = \{e \in E(\mathcal{G}) : e_- = v\}.$$

A graph morphism $\phi : \mathcal{G} \to \mathcal{G}'$ naturally induces a graph morphism between the stars of v and $\phi(v)$.

A graph morphism $\phi : \mathcal{G} \to \mathcal{G}'$ is called an *immersion* if for every vertex v, the induced graph morphism between the stars of v and $\phi(v)$

$$Star_{\mathcal{G}}(v) \to Star_{\mathcal{G}'}(\phi(v))$$

is injective. That is, ϕ is locally injective. If, in addition, it is surjective, then ϕ is called a *covering*.

An important consequence of an immersion is the following result.

Lemma 2.23. *An immersion induces an imbedding of fundamental groups. That is, ϕ_* is injective.*

Proof. The proof is clear by definition of an immersion. It indeed follows from the following useful fact. Given any circuit γ' based at $\phi(o)$ in $\mathcal{G}'$, there exists a unique circuit γ based at o in $\mathcal{G}$ such that $\phi(\gamma) = \gamma'$. Moreover, if γ' has no backtracking, then so is γ.

Hence, if $\phi_*(\gamma)$ is the basepoint $\phi(o)$ for a reduced circuit γ based at o in $\mathcal{G}$, then $\gamma = o$. □

Let $\phi : \mathcal{G} \to \mathcal{G}'$ be a graph morphism. We shall make use of an operation called *folding* to convert ϕ to an immersion form a **new graph** to $\mathcal{G}'$.

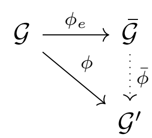
A pair of edges e, e' in $\mathcal{G}$ is called *foldable* if $e_- = (e')_-$, $\bar{e} \neq e'$, and $\phi(e) = \phi(e')$. Given a foldable pair of edges e, e', we can formulate a graph morphism ϕ_e called *folding* and a new graph $\bar{\mathcal{G}}$ as follows

$$\phi_e : \mathcal{G} \to \bar{\mathcal{G}} := \mathcal{G}/\{e = e', \bar{e} = \bar{e}'\}$$

by identifying the edges $e = e'$ and $\bar{e} = \bar{e}'$ respectively.

Observe that such an operation definitely decreases the number of edges and vertices. It is also possible that two loops can be identified. In this case, the fundamental group of new graph $\bar{\mathcal{G}}$ changes.

Moreover, given a foldable pair of edges e, e', we can naturally define a new graph morphism $\bar{\phi} : \bar{\mathcal{G}} \to \mathcal{G}'$ such that the following diagram



is commutative.

We do the above *folding process* for each foldable pair of edges, and finally obtain an immersion from a new graph $\bar{\mathcal{G}}$ to $\mathcal{G}'$. Precisely, we have the following.

Lemma 2.24. *Let $\phi : \mathcal{G} \to \mathcal{G}'$ be a graph morphism. Then there exists a sequence of foldings $\phi_i : \mathcal{G}_i \to \mathcal{G}_{i+1}$ for $0 \le i < n$ and an immersion $\bar{\phi} : \bar{\mathcal{G}} \to \mathcal{G}'$ such that $\phi = \bar{\phi}\phi_n \cdots \phi_0$, where $\mathcal{G}_0 = \mathcal{G}, \mathcal{G}_n = \bar{\mathcal{G}}$.*

An important consequence of the above folding process is that ϕ_* and $\bar{\phi}_*$ have the same image in the fundamental group of $\mathcal{G}'$.

We apply the above theory to subgroups of a free group and to prove Theorem 2.17.

Theorem 2.25 (Nielsen basis)**.** *Let H be a subgroup of a free group $F(S)$. Then H is a free group. Moreover, for any generating set T of H, there exists an algorithm to find a free basis for H.*

Proof. Let H be a subgroup of a free group $F(S)$. Suppose that H is generated by a set $T \subset F(S)$. By the above discussion, there exists a rose $\mathcal{G}'$ with one vertex and $2|S|$ edges whose fundamental group is $F(S)$. Here in fact, we choose an orientation on $\mathcal{G}'$ and then identify $\pi_1(\mathcal{G}')$ as $F(S)$.

Note that T are a set of reduced words. For each word $W \in T$, we associate to W a circuit graph $\mathcal{C}_W$ of $2|T|$ edges with a basepoint o and an orientation such that the clock-wise "label" of $\mathcal{C}_W$ is the word W. It is obvious that there exists a graph morphism $\mathcal{C}_W \to \mathcal{G}'$.

We attach all $\mathcal{C}_W$ at o for $W \in T$ to get a graph $\mathcal{G}$. Then we have a graph morphism $\phi : \mathcal{G} \to \mathcal{G}'$. It is also clear that the image $\phi_*(\pi_1(\mathcal{G}))$ is the subgroup H in $F(S)$. Hence, a consequence of Lemma 2.24 is that any subgroup of a free group is free. Moreover, since the immersion given by Lemma 2.24 induces an injective homomorphism, we can easily obtain a free basis of H by writing down the generating elements of the fundamental group of $\bar{\mathcal{G}}$. □

Theorem 2.26 (Hall)**.** *Let H be a finitely generated subgroup in a free group F of finite rank. For any element $g \in F \setminus H$, there exists a finite index subgroup Γ of F such that $H \subset \Gamma$ and $g \notin \Gamma$.*

Remark. A subgroup with the above property is called *separable*. In other words, a subgroup H is separable in G if it is the intersection of all finite index subgroups of G containing H.

To prove this theorem, we need two additional facts about the covering of a graph.

Lemma 2.27. *Let $\phi : (\mathcal{G}, o) \to (\mathcal{G}', \phi(o))$ be a covering, and γ be a reduced circuit in $\mathcal{G}'$ based at $\phi(o)$. If γ is not in $\phi_*(\pi_1(\mathcal{G}, o))$, then any lift of γ is not a circuit.*

Proof. This follows just from the definition of a covering. □

Lemma 2.28. *Let $\phi : (\mathcal{G}, o) \to (\mathcal{G}', \phi(o))$ be a covering. If $\mathcal{G}$ and $\mathcal{G}'$ are finite, then $\phi_*(\pi_1(\mathcal{G}, o))$ is of finite index in $\pi_1(\mathcal{G}', o)$.*

Proof. Denote by H the subgroup $\phi_*(\pi_1(\mathcal{G}, o))$. We count the right coset Hg where $g \in \pi_1(\mathcal{G}', o)$. Then any lift of the circuit in Hg based at o has the same terminal endpoint. Moreover, if $Hg \neq Hg'$, then the endpoints of corresponding lifts are different. Indeed, if not, we get a circuit and finally we see that $g'g^{-1} \in H$. Since $\mathcal{G}$ is finite, we see that there are only finitely many different right H-cosets. □

We are now in a position to give the Stalling's proof of Theorem 2.26.

Proof of Theorem 2.26. Let $\mathcal{G}'$ be a rose. We have put an orientation on $\mathcal{G}'$, a subset E_0 of edges, such that $\pi_1(\mathcal{G}')$ is identical to $F(E_0)$.

Let H be a finitely generated subgroup in F with a finite generating set T. Given $g \not{i}nH$, we write g as a reduced word W_g over S, and similarly for each $t \in T$ a word W_t. As in the proof of Theorem 2.25, we construct a graph by gluing circuits labeled by W_t for $t \in T$ and use the folding to get an immersion $\phi : (\mathcal{G}, o) \to (\mathcal{G}', \phi(o))$, where $\mathcal{G}$ has the fundamental group H. This naturally induces an orientation E_1 on $\mathcal{G}$. Now we attach a path labeled by W_g at o by following the orientation $\mathcal{G}$. Since $g \notin H$, the endpoint of the path must be different from o, i.e.: the path is not closed. The new graph is still denoted by $\mathcal{G}$ for simplicity. And $\phi : (\mathcal{G}, o) \to (\mathcal{G}', \phi(o))$ is still an immersion.

Denote by V the vertex set of $\mathcal{G}$. For each $e \in E_0$, we have a set of directed edges $\phi_e^{-1}(e)$ in E_1. Since ϕ is an immersion, each edge in $\phi_e^{-1}(e)$ defines an ordered pair of endpoints in V. Thus, each $e \in E_0$ defines a bijective map ι_e on a subset of the vertex set V of $\mathcal{G}$. Similarly, we can define $\iota_{\bar{e}}$ for $e \in E_0$.

Since V is finite, ι_e can be extended to a bijective map of V. (We actually have many choices). Let's denote again by ι_e one such bijective map of V.

It is easy to use these maps $\iota_e, \iota_{\bar{e}}$ for $e \in E_0$ to complete the immersion $\phi : (\mathcal{G}, o) \to (\mathcal{G}', \phi(o))$ to a covering $\tilde{\phi} : (\tilde{\mathcal{G}}, o) \to (\mathcal{G}', \phi(o))$. Precisely,

For each $e \in E_0$, we use ι_e to connect v and $\iota_e(v)$ by a directed edge e, if such an edge was not in $\phi_e^{-1}(e)$. We do similarly for each $\bar{e}$ where $e \in E_0$. It is clear that the such obtained graph $\tilde{\mathcal{G}}$ is a finite covering.

Thus, the fundamental group of $\tilde{\mathcal{G}}$ contains H but not g, since the path labeled by W_g is not closed in $\tilde{\mathcal{G}}$. The proof is complete. □

The following two exercises are consequences of Theorem 2.26.

Exercise 2.29. *A free group F is residually finite: for any $g \neq 1 \in F$, there exists a homorphism $\phi : F \to G$ to a finite group G such that $\phi(g) \neq 1$.*

Using the residual finiteness, we can prove the following.

Exercise 2.30. *Free groups of finite rank are Hopfian: any epimorphism is an isomorphism.*

3. Groups presentation and decision problems

3.1. Group Presentation.

Definition 3.1. Let G be a group, and $X \subset G$ be a subset. The *normal closure* of X, denoted by $\langle\langle X \rangle\rangle$, is the minimal normal subgroup containing X. Equivalently,

$$\begin{aligned} \langle\langle X \rangle\rangle &= \langle \{gxg^{-1} : x \in X, g \in G\} \rangle \\ &= \{(g_1 x_1^{\epsilon_1} g_1^{-1})...(g_n x_n^{\epsilon_n} g_n^{-1}) : n \in \mathbb{N}, x_i \in X, g_i \in G, \epsilon_i \in \{1, -1\}\}. \end{aligned}$$

Definition 3.2. A *presentation* $P = \langle S|R \rangle$ consists of a set S called *generators*, and a set of words R in $\mathcal{W}(\tilde{S})$ called *relators*. A presentation is *finite*, if both S and R are finite.

A group G is *presented by* P if

$$G \cong F(S)/\langle\langle R \rangle\rangle.$$

In this case, we write $G = \langle S|R \rangle$ for simplicity.

A group is *finitely presented* if it is presented by a finite presentation.

Remark.
(1) By the universal mapping property of free groups, any group has a presentation.
(2) Suppose X is finite. Then the normal subgroup $\langle\langle R \rangle\rangle$ is not finitely generated, unless the group G is finite.

Theorem 3.3. *Let $G = \langle S|R \rangle$ and $\phi : S \to H$ be a function, where H is a group. Then there exists a unique homomorphism $\varphi : G \to H$ with $\varphi|_S = \phi$ if and only if $\phi_H(r) = 1 \in H$ for all $r \in R$, where $\phi_H : F(S) \to H$ is the canonical homomorphism such that*

$$\begin{array}{ccc} S & \xrightarrow{\iota} & F(S) \\ & \searrow^{\phi} & \downarrow \phi_H \\ & & H \end{array}$$

Proof. Let $\phi_G : F(S) \to G$ be the epimorphism such that

$$\begin{array}{ccc} S & \xrightarrow{\iota} & F(S) \\ & \searrow^{\iota_G} & \downarrow \phi_G \\ & & G \end{array}$$

where $\iota_G : S \to G$ is the natural inclusion.

=>. Suppose that there exists $\varphi : G \to H$ with $\varphi|_S = \phi$. Then $\varphi\phi_G : F(S) \to H$ is a homomorphism. We claim that $\varphi\phi_G = \phi_H$. Indeed, it suffices to verify that $\varphi\phi_G$ and ϕ_H are identical on the generating set $\iota(S)$: $\varphi\phi_G(\iota(s)) = \phi_H(\iota(s))$ for any $s \in S$. Then $\varphi\iota_G(s) = \varphi(s) = \phi(s) = \phi_H(\iota(s))$. Hence, it follows that $\varphi\phi_G = \phi_H$.

<=. Note that $G \cong F(S)/N$, where $N = \ker(\phi_G) = \langle\langle R \rangle\rangle$. We define a map $\varphi : G \to H$ as follows. Let $g = s_1 s_2..s_n N$, where $s_i \in \tilde{S}$. Define $\varphi(g) = \phi_H(s_1)\phi_H(s_2)...\phi_H(s_n)$.

We verify that this map is well-defined. Assume that $g = s_1 s_2..s_n N = s_1' s_2'..s_m' N$. There exists $w \in N$ such that $s_1 s_2..s_n w = s_1' s_2'..s_m'$. As $\phi_H(r) = 1$ for all $r \in R$, it follows that $\phi_H(w) = 1$. Thus, $\phi_H(s_1)\phi_H(s_2)...\phi_H(s_n) = \phi_H(s_1')\phi_H(s_2')...\phi_H(s_m')$.

We now prove that $\varphi(gg') = \varphi(g)\varphi(g')$. As $\varphi(g)$ does not depend on the choice of representatives of g, we assume that $g = s_1 s_2..s_n N, g' = s_1' s_2'..s_m' N$. Then $\varphi(gg') = \phi_H(s_1)\phi_H(s_2)...\phi_H(s_n)\phi_H(s_1')\phi_H(s_2')...\phi_H(s_m') = \varphi(g)\varphi(g')$. □

Example 3.4. Let $G = \langle a, t | t^{-1}at = a^2 \rangle$. Check the map $\phi(a) = a^2, \phi(t) = t$ extends to a homomorphism of $G \to G$.

Exercise 3.5. (1) *Let $G = \langle S|R \rangle$. Let R' be a set of words in $\langle\langle R \rangle\rangle$. Then*

$$\langle S|R \rangle \cong \langle S|R \cup R' \rangle.$$

(2) *Let $G = \langle S|R \rangle$. Choose a set T and for each $t \in T$, we take a word in $w_t \in \mathcal{W}(\tilde{S})$. Then*

$$\langle S|R \rangle \cong \langle S \cup T|R \cup \{t^{-1}w_t\} \rangle.$$

The modifications given by Exercise 3.5 are called *Tietze transformations*. They gives ways to modify(add/delete generators and relators) a presentation of a group, without changing the group. Moreover, any two presentations of a given group can be transformed to each other by a sequence of Tietze transformations(Thm 2.8, in reference [7]).

Exercise 3.6. *Finish Exercises 2.4, 2.5, 2.6 and 2.7 in reference* [7].

3.2. **Decision Problems.** We are only interested in decision problems for finitely presented groups. Let G be a finitely presented group. It should be noted that if one finite presentation has solvable word problem, then any other finite presentation does so. So we can speak about the solvability of word problem for a finitely presented group.

By Theorem 2.9, it is easy to see that a free group has a solvable word problem. Indeed, given a word w, it is easy to build an algorithm to find an inverse pair in w and cancel it to get a new word with length $|w| - 2$. After finitely many steps, we get a reduced word. If it is an empty word, then w represents the identity and the algorithm return "yes". Otherwise, $w \neq 1$ and the algorithm returns "no".

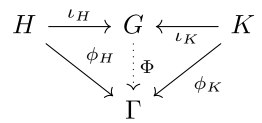
Let $w \in \mathcal{W}(\tilde{X})$ be a word. Write $w = x_1x_2...x_n$ for $x_i \in \tilde{X}$. Let σ be a cyclic permutation of the set $\{1, 2, ..., n\}$(i.e.: some power of the permutation: $1 \to 2, ..., i \to i+1, ..., n \to 1$). A *cyclic permutation* of w is the new word formed by $x_{\sigma(1)}x_{\sigma(2)}...x_{\sigma(n)}$. A word w is called *cyclically reduced* if all permutation of w is reduced. If w is already reduced, it is amount to saying that the first letter of w is not the inverse of its last letter.

Exercise 3.7. *Prove that in a free group, two cyclically reduced words are conjugate if and only if each is a cyclic permutation of the other. Therefore, the conjugacy problem is solvable in free groups.*

4. Groups acting on trees: a brief introduction to Bass-Serre Theory

4.1. **Free products.** In this section, we first introduce a *free product* G of two given groups H, K. The group G is the biggest one among the groups generated by H, K with the property that any such groups are the quotient of G.

Precisely, a group G is called a *free product* of H and K if there exist a pair of homomorphisms $\iota_H : H \to G$ and $\iota_K : K \to G$ such that they are universal in the following diagram:



By the universal property, it is easy to see that $\iota_H : H \to G$ and $\iota_K : K \to G$ are both injective. Moreover, G is unique up to isomorphism, so G must be generated by H and K.

Suppose that $H = \langle S|D\rangle, K = \langle T|E\rangle$ are given by presentations. Then by Theorem 3.3, the group G given by the presentation $\langle S \cup T|D \cup E\rangle$ is the free product of H, K.

Understanding H and K as disjoint alphabet sets, an *alternating word* w is of form $h_1k_1 \cdots h_mk_m$, where $h_i \in H, k_j \in K$. The length of w is the number of letters in word. It is called *reduced* if $h_i \in H \setminus 1, k_j \in K \setminus 1$.

We consider the set Ω of all reduced alternating words $h_1k_1 \cdots h_mk_m$ in H and K. The following result is fundamental in understanding free products.

Theorem 4.1. *[Normal form theorem]*[7, Thm 3.1] *Every element of $G = H \star K$ is equal to a unique alternating expression of the form $h_1k_1 \cdots h_mk_m \in \Omega$*

Sketch of the proof. Since G is genearted by H and K, any element in G can be written as an alternating expression of the form $h_1k_1 \cdots h_mk_m$. To prove the uniqueness, we shall construct a free action of G on the set Ω of all alternating expressions. To that end, we first construct the homomorphisms of H and K into the symmetry group of Ω and then by the above universal property, the homomorphisms of $G \to Sym(\Omega)$ is defined correspondingly.

For each $h \in H$, the associated bijection $\phi_H(h)$ is given by sending $h_1k_1 \cdots h_mk_m$ to $hh_1k_1 \cdots h_mk_m$ with a neccessary modification so that the image is alternating expression. Note that $\phi_H : H \to Sym(\Omega)$ is an injective homomorphism. Similarly, we can define $\phi_K : K \to Sym(\Omega)$ which is also injective. These define a group homormorphism $G \to Sym(\Omega)$.

Consider the empty word $\emptyset$ in Ω. Any alternating expression $h_1k_1 \cdots h_mk_m$ of element g in G maps $\emptyset$ to the alternating word $h_1k_1 \cdots h_mk_m$. This word is nonempty iff $g \neq 1$ This establishes the uniqueness of the statement. □

Corollary 4.2. *If an alternating word $w = h_1k_1 \cdots h_mk_m$ reprensents the identity in G, then it must be not reduced: there exists some i such that $h_i = 1$ or $k_i = 1$.*

In particular, if two reduced words represent the same group element, then they are equal letter by letter.

4.2. **Free products acting on trees.** Let $G = H \star K$ be a free product. We define a graph Γ as follows.

(1) The vertex set V consists of two types H and K: $V = \{gH, gK : g \in G\}$.

(2) The edge set E consists of all group elements in G.
(3) The edge $g \in E = G$ connects gH and gK.

Then G acts on Γ: each element $g \in G$ sends xH to gxH and xK to gxK. The edge relation is preserved. So G acts on Γ by graph isomorphism.

Theorem 4.3 (Bass-Serre Trees for free products)**.** *The graph Γ is a tree so that the degree of vertex of type H (resp. K) equals $\sharp H$ (resp. $\sharp K$).*

Moreover, the action of G on Γ has trivial edge stablizers and vertices stablizers of type H and K conjugated to H and K respectively so that the quotient is an interval.

Proof. By definition of action of G on the graph, there are two different orbits of vertices: $G \cdot H$ and $G \cdot K$. The vertex H is adjacent to hK for $h \in H$. That is to say, the set of edges adjacent to H has one-to-one correspondence with the set of elements in H. Similarly, the edges adjacent to K correspond to the set K. Since G is generated by H and K, the graph Γ is connected.

We now prove that Γ is a tree. Let γ be an immersed loop: there exists no backtracking. Up to a tranlation, we can assume that γ is based at H; the case at K is similar. According to the adjacency, there are an even number of edges in γ, where edges in γ must be of form $H \leftrightarrow hK$ for $h \in H$ or $K \leftrightarrow kH$ for $k \in K$ up to translation. Thus, tracing out the loop γ, we see that the terminal point is the vertex $h_1k_1 \cdots h_nk_nH$ for $h_i \in H$ and $k_i \in K$.

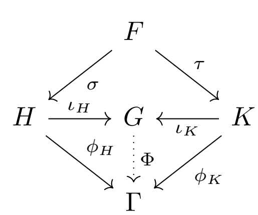
Since γ is a loop, we have the equality $h_1k_1 \cdots h_nk_nH = H$. We obtain that

$$h_1k_1 \cdots h_nk_n = h \tag{4}$$

for some $h \in H$. Since there exists no backtracking, we see that $k_i \neq 1$ for $i < n$ and $h_i \neq 1$ for $i > 1$. It is possible that $k_n = 1$ or $h_1 = 1$, but they cannot happen at the same time. So, up to removing h_1 or k_n from the left side in (4), we obtain a reduced word of length at least 2. But the right side in (4) is a reduced word of length 1. This is a contradiction to the normal form theorem 4.1. The graph Γ is thus a tree. The proof is complete. □

4.3. **Free products with amalgamation.** Now suppose that each of H and K contain a subgroup isomorphic to F: there exists monomorphisms $\sigma : F \to H$ and $\tau : F \to K$. We want to formulate a biggest group G generated by H and K so that $H \cap K = F$ is realized inside G.

Precisely, a group G is called a *free product* of H and K with malgamation over F if there exist a pair of homomorphisms $\iota_H : H \to G$ and $\iota_K : K \to G$ such that they are universal in the following diagram:



By the universal property, it is easy to see that $\iota_H : H \to G$ and $\iota_K : K \to G$ are both injective. Moreover, G is unique up to isomorphism, so G is generated by H and K.

Let $w = h_1k_1 \cdots h_nk_n$ be an alternating word over the alphabet set H and K such that $h_i \in H$ and $k_i \in K$. If it has length strictly bigger than 1 and $h_i \in H \setminus F, k_i \in K \setminus F$, then w is called a *reduced* alternating word. Let Ω be the set of all such that $h_i \in H \setminus F$ and $k_i \in K \setminus F$, unless the alternating word is of length 1.

The following fact is obvious.

Lemma 4.4. *Every element of $G = H \star_F K$ is equal to a reduced alternating expression of the form $h_1k_1 \cdots h_mk_m$. It may not be unqiue, but the length of the alternating word is unique.*

However, such an expression will not be unique, due to the nontrivial intersection F. To obtain a unique normal form, we have to choose a *right coset transversal* T_H and T_K of F in H and K respectively: namely, in each right H-coset, choose a right coset representative. We then consider the set Ω of words concatenating F with the alternating words in T_H and T_K.

Given a reduced alternating form $h_1k_1 \cdots h_mk_m$, we convert the letters from right to left so that they become to be the corresponding right coset representatives. In the final form, we will get a *normal form* $fh'_1k'_1 \cdots h'_mk'_m$ for some $f \in F$. So, any element has a normal form which turns out to be unique. Also note that in this process, the length of a normal form is the same as that of the original one.

Theorem 4.5. *[Normal form theorem]*[7, Thm 3.7] *With the choice of the right coset transversal T_H and T_K as above, every element of $G = H \star_F K$ is equal to a unique normal form $fh_1k_1 \cdots h_mk_m$ with $f \in F, h_i \in T_H$ and $k_i \in T_K$ when present.*

Corollary 4.6. *If an alternating expression $h_1k_1 \cdots h_mk_m$ gives the identity, then it is not reduced: there exists h_i or k_i such that $h_i, k_i \in F$.*

Let $G = H \star_F K$. We define a graph Γ as follows.

(1) The vertex set V consists of two types H and K: $V = \{gH, gK : g \in G\}$.
(2) The edge set E consists of all left F-cosets in G.
(3) The edge $gF \in E$ connects gH and gK.

Again, G acts on Γ: each element $g \in G$ sends xH to gxH and xK to gxK. The edge relation is preserved. So G acts on Γ by graph isomorphism.

We have the same result for free product with amalgamation. The only difference is that the edge stabilizer is a conjugate of F, instead of a trivial group.

Theorem 4.7 (Bass-Serre Tree for amalgamation)**.** *The graph Γ is a tree so that the degree of vertex of type H (resp. K) equals $\sharp H/F$ (resp. $\sharp K/F$).*

Moreover, the action of G on Γ has edge stablizers conjugated to F and vertices stablizers of type H and K conjugated to H and K respectively so that the quotient is an interval.

Proof. The proof is similar to that of Theorem 4.3. We emphasize the differences below.

We now prove that Γ is a tree. Let γ be an immersed loop: there exists no backtracking. Up to a tranlation, we can assume that γ is based at H; the case at K is similar. According to the adjacency, there are an even number of edges in γ, where edges in γ must be of form $H \leftrightarrow hK$ for $h \in H/F$ or $K \leftrightarrow kH$ for $k \in K/F$

up to translation. Tracing out the loop γ, we see that the terminal point is the vertex $h_1k_1\cdots h_nk_nH$ for $h_i \in H$ and $k_i \in K$.

Since γ is a loop, we have the equality $h_1k_1\cdots h_nk_nH = H$. We obtain that

$$h_1k_1\cdots h_nk_n = h \tag{5}$$

for some $h \in H$. Since there exists no backtracking, we see that $k_i \notin 1$ for $i < n$ and $h_i \notin F$ for $i > 1$. It is possible that $k_n \in F$ or $h_1 \in F$, but they cannot happen at the same time. So, combinning h_1 with k_1 or k_n with h_n from the left side in (5) if necessary, we obtain a reduced word of length at least 2. But the right side in (5) is a reduced word of length 1. This is a contradiction to the normal form theorem 4.5. The graph Γ is thus a tree. The proof is complete. □

4.4. **HNN extension.** Let G be a group with two isomorphic subgroups H and K. Let $\tau : H \to K$ be an isomorphism. We want to build a new group $\tilde{G}$ such that $G \subset \tilde{G}$ and H, K become conjugate in $\tilde{G}$. If G is given by a presentation

$$\langle S|\mathcal{R}\rangle.$$

As usual, we request $\tilde{G}$ to be the biggest one with this property. Then the desired group $\tilde{G}$ must have presentation as follows

$$\langle S, t|\mathcal{R}, tht^{-1} = \tau(h), \forall h \in H\rangle,$$

which is called *HNN* extension of G over H, K, denote by $G\star_{H\sim K}$. The new generator t is usually called *stable* letter.

By definition, every element in $\tilde{G}$ can be written as a product of form called *t-expression* as follows:

$$g_0t^{\epsilon_1}g_1\cdots t^{\epsilon_n}g_n$$

where $g_i \in G$ or $\epsilon_i \in \{1, -1\}$. Any subword tht^{-1} for $h \in H$ and $t^{-1}kt$ for $t \in K$ is called *t-pinch* in the above form. A t-expression form without t-pinches is called *reduced.*

A reduced t-expression of an element may not be unique, though different reduced t-expressions have equal length. In order to get a normal form, we choose right coset transversal T_H and T_K of H and K in G respectively.

Theorem 4.8. *[Normal form theorem]*[7, Thm 3.1] *Every element of $\tilde{G} = G\star_{H\sim K}$ is equal to a unique reduced t-expression of the form $g_0t^{\epsilon_1}g_1t^{\epsilon_2}g_2\cdots t^{\epsilon_n}g_n$ with $g_i \in T_H \cup T_K$ when present. If $\epsilon_i = 1$ for $i > 0$, then $g_i \in T_H$; if $\epsilon_i = -1$ for $i > 0$ then $g_i \in K$.*

Corollary 4.9 (Briton's Lemma)**.** *If a t-expression in $\tilde{G} = G\star_{H\sim K}$ represents the trivial element, then it must contain t-pinches.*

Let $\tilde{G} = G\star_{H\sim K}$. We define a graph Γ as follows.

(1) The vertex set V consists of all left cosets of G: $V = \{xG : x \in \tilde{G}\}$.
(2) The edge set E is the set of the 2-tuples $\{x(K, tH) : x \in \tilde{G}\}$.
(3) The edge $x(K, tH) \in E$ connects xG and xtG, and so $xt^{-1}(K, tH)$ connects $xt^{-1}G$ and xG accordingly.

We define an action of $\tilde{G}$ on Γ: each element $g \in \tilde{G}$ sends xG to gxG. The edge relation is preserved. Then, G acts on Γ by graph isomorphism. By definition, G acts transitively on edges so Γ/G is a loop.

Theorem 4.10 (Bass-Serre Tree for HNN extension)**.** *The graph Γ is a tree so that the degree of every vertex is $\sharp G/H + \sharp G/K$. Moreover, the action of G on Γ has edge stablizers conjugated to H (or equivalently K) and vertex stablizers G so that the quotient Γ/G is a loop.*

Proof. The proof is similar to that of Theorem 4.3. In this case, there is only one orbit of vertices and edges.

We now prove that Γ is a tree. According to the adjacency, each edge in Γ issuing from the vertex G must be of the following form:

(1) $G \overset{g(K,tH)}{\longleftrightarrow} gtG$ for some $g \in G$,
(2) $G \overset{gt^{-1}(K,tH)}{\longleftrightarrow} gt^{-1}G$ for some $g \in G$.

Thus, if a loop γ based at G has the terminal point $g_1t^{\epsilon_1}g_2t^{\epsilon_2}\cdots g_nt^{\epsilon_n}G$ where $g_i \in G$ and $\epsilon_i \in \{1,-1\}$. Conversely, any t-epression $g_1t^{\epsilon_1}g_2t^{\epsilon_2}\cdots g_nt^{\epsilon_n}$ labels a connected path between G and $g_1t^{\epsilon_1}g_2t^{\epsilon_2}\cdots g_nt^{\epsilon_n}G$.

Observe that $g_1t^{\epsilon_1}g_2t^{\epsilon_2}$ is given by a backtracking iff $t^{\epsilon_1}g_2t^{\epsilon_2}$ is t-pinch. Indeed, for definiteness, let us consider $\epsilon_1 = 1$. Then if the two edge path as follows has backtracking

$$G \overset{g_1(K,tH)}{\longleftrightarrow} g_1tG \overset{g_1tg_2t^{\epsilon_i}(K,tH)}{\longleftrightarrow} g_1tg_2t^{\epsilon_i}G$$

then the two edges coincide: $g_1K = g_1tg_2t^{\epsilon_i}K, g_1tH = g_1tg_2t^{\epsilon_i}(tH)$. So we must have

$$tg_2t^{\epsilon_i} \in K, g_2t^{\epsilon_i}t \in H$$

which occurs only when $\epsilon_i = -1$ and $g_2 \in H$. In other words, $tg_2t^{\epsilon_i}$ is a t-pinch. The case $\epsilon_1 = -1$ is similar.

Let γ be an immersed loop: there exists no backtracking. Up to a tranlation, we can assume that γ is based at G. Thus, tracing out the loop γ gives the terminal point $g_1t^{\epsilon_1}g_2t^{\epsilon_2}\cdots g_nt^{\epsilon_n}G$ where $g_i \in G$ and $\epsilon_i \in \{1,-1\}$. Since γ is a loop, we have the equality $g_1t^{\epsilon_1}g_2t^{\epsilon_2}\cdots g_nt^{\epsilon_n}G = G$. We obtain that

$$g_1t^{\epsilon_1}g_2t^{\epsilon_2}\cdots g_nt^{\epsilon_n} = g \tag{6}$$

for some $g \in G$.

Since γ has no backtracking, the t-expression $g_1t^{\epsilon_1}g_2t^{\epsilon_2}\cdots g_nt^{\epsilon_n}$ is reduced of length at least 2. The right side in (6) is a reduced word of length 1. This is a contradiction to the normal form theorem 4.8. The graph Γ is thus a tree. The proof is complete. □

Exercise 4.11. *Draw a portion of the Cayley graph of the group $\langle a, t : ta^2t^{-1} = a^3\rangle$.*

4.5. **Abelianization of free products.** Let $G = H \star K$ be a free product. By the universal property, there exists a natural morphism $G \to H \times K$. We denote by N the kernel of this morphism.

We start with a classification of isometries on trees. An element is called *elliptic* if it fixes a vertex. It is called *hyperbolic* if it preserves a unique bi-infinite geodesic and acts by translation.

Lemma 4.12. *Let g be an isometry of a tree T. Then either g fixes a point or g preserves a unique geodesic by translation.*

Proof. Suppose that g does not fix any point. Fix a basepoint $o \in T$. Consider the geodesic segments $[o, go]$ and $[o, g^{-1}]$ both originating from o. Since T is a tree, let b be the branching point of these two geodesics. We then claim that $d(o, m) <$

$d(o, go)/2$. Otherwise, the middle point m of $[o, go]$ coincides with that of $[o, g^{-1}o]$, which is gm. So m is fixed by g: a contradiction. Hence, $d(o, b) < d(o, go)/2$. We now form a geodesic γ preserved by g. Let $\gamma = \cup_{i\in\mathbb{Z}}[g^{-1}m, m]$. It is clear by the the above Claim that γ is a geodesic. The proof is then complete. □

Exercise 4.13. *Prove the uniqueness statement in Lemma 4.12.*

Exercise 4.14. *Any finite group acts on a tree with a global fixed point.*

Corollary 4.15. *All finite subgroups in a free product must be conjuated into H or K.*

Exercise 4.16. *Let $G = H \star_F K$ be a free product of non-trivial groups H, K over F. Using Bass-Serre tree to prove that the center of G is contained in F.*

Lemma 4.17. *N is a free group generated by $S = \{[h, k] : h \in H \setminus 1, k \in K \setminus 1\}$.*

Proof. By Theorem 2.20, we only need to show that N acts freely on the Bass-Serre tree Γ. To prove the freeness of N, any vertex stabilizer of the Bass-Serre Γ is sent to a non-trivial subgroup under the morphism $H \star K \to H \times K$. This implies that the kernel N of this morphism acts freely on Γ. Thus the conclusion follows. □

4.6. **Graph of groups.** Let $\mathcal{G} = (V, E)$ be a finite graph. A *graph of groups* $(\mathcal{G}, G_\star)$ with the underlying graph $\mathcal{G}$ associates to each vertex $v \in V$ a *vertex group* G_v, each edge $e \in E$ an *edge group* G_e and monomorphisms $\partial_e^\pm : G_e \to G_{e_\pm}$.

Fix a spanning tree Γ and denote by S the set of edges not in Γ. The *fundamental group* $\pi_1(\mathcal{G}, G_\star)$ of the graph $(\mathcal{G}, G_\star)$ of groups is the group generated by the union of $\cup_{v\in V} G_v$ with $\{t_e : e \in S\}$ subject to the following relations:

(1) for each $e \in \Gamma^1$ and each $x \in G_e$, $\partial_e^-(x) = \partial_e^+(x)$.
(2) for each $e \in S$ and each $x \in G_e$, $t_e\partial_e^-(x)t_e^{-1} = \partial_e^+(x)$.

Examples 4.18. Using this langauge, a free product $H \star_F K$ with amalgamation is a graph of groups where the graph is one edge with two distinct vertices. The edge group is F and two vertex groups are H and K respectively. The edge morphisms are given by subgroup inclusions.

The HNN extension $H\star_K$ is a graph of groups where the graph is just one loop. The edge group is K and the vertex group is H. The edge morphism $\partial_- e : H \to K$ is given by subgroup inclusion of K into H, and the other $\partial_+ e : H \to K$ is given by conjugation $H \to tHt^{-1}$, where t is the stable letter.

By a similar consruction, the fundamental group $G := \pi_1(\mathcal{G}, G_\star)$ acts on a tree T called *Bass-Serre tree* so that T/G is the underlying graph $\mathcal{G}$. The vertex and edge stabilier in the tree T are isomorphic to the corrresponding vertex and edge stabilier of $\mathcal{G}$. This property entails the construction of T to staisfy:

(1) the vertex set $V(T)$ consists of left cosets of vertex groups in G.
(2) the edge set $E(T)$ consists of left cosets of edge groups in G.

The details can be found in Chapter I.5.3 [12], see Theorem 14.

4.7. **Groups acting on trees.** Assume that G acts by graph isomorphisms on a tree $T = (V, E)$. Without loss of generality, assume that G acts without inversion: no element $g \in G$ sends e to $\bar{e}$, where $\bar{e}$ is the directed edge with opposite orientation of e. If we think of the tree T as a geodesic metric space, then G acts by isometry on T and the action without inversion is equivalent to the assumption that no non-trivial fixes the middle point of some edge.

Furthermore, we can assume that the action of G on T is *minimal*: there exists no proper G-invariant subtree.

Exercise 4.19. *Let G act without inversion on a tree T. Then there exists a unique minimal G-invariant subtree in T.*

In what follows, to each action without inversion of a group G on a tree T, we associate a finite graph of groups defined as above.

Let $\mathcal{G} = (V, E)$ be the quotient graph T/G. Assume that $\mathcal{G}$ is finite. We choose a spanning tree Γ in $\mathcal{G}$ and consider its lift in T denoted by $\tilde{\Gamma}$. For each $v \in \Gamma^0$ and $e \in \Gamma^1$, let G_v be the vertex stabilizer of (its lift) $\tilde{v}$ in T and G_e be the edge stabilizer of (its lift) $\tilde{e}$ in T.

Let $\partial_e^{\pm} : G_e \to G_{e_\pm}$ for $e \in \Gamma^1$ be the inclusion of edge stabilizer into vertex stabilizer.

It remains to define the egde group on the set S of edges not in Γ^1 and the corresponding $\partial_e^{\pm} : G_e \to G_{e_\pm}$. To that end, we choose a lift $\tilde{e}$ for each $e \in S$ so that $\tilde{e}_\pm \cap \Gamma^0 \neq \emptyset$. Assume that $\tilde{e}_\pm \cap \tilde{\Gamma}^0 = \tilde{e}_-$: the initial endpoint of edge $\tilde{e}$ is contained in $\tilde{\Gamma}$ (the lift of Γ) in T. Let G_e be the stabilizer of $\tilde{e}$ in G. Define $\partial_e^- : G_e \to G_{e_-}$ be the edge group inclusion. Since Γ is a spanning tree of the quotient T/G, the vertex set of $\tilde{\Gamma}$ is a fundamental domain of T^0: $G\Gamma^0 = T^0$. Thus, there exists an element $t_e \in G$ such that $t_e \tilde{e}_+ \in \Gamma^0$. Define

$$\partial_e^+ : G_e \to G_{t_e \tilde{e}_+} = t_e G_{\tilde{e}_+} t_e^{-1}$$

which sends $x \in G_e$ to $t_e x t_e^{-1}$.

Recall that, given a finite graph of groups, we can build a Bass-Serre tree T and an action of the its fundamental group on T such that $T/G = \mathcal{G}$, and the vertex and edge stabilier in the tree are isomorphic to the corrresponding vertex and edge stabilier of $\mathcal{G}$. The bulk of the Bass-Serre theory says that these two operations are inverse in the following sense.

Theorem 4.20. *If $\mathcal{G}$ is a graph of groups, then there is a group G, a tree T, and an action G without inversion on T so that*

1. *If $\bar{G}$ is the graph of groups associated with the action G on T, then $\bar{G}$ and G are isomorphic.*
2. *If $G' \curvearrowright T'$ is another action on a tree satisfying (1), then there is an isomorphism $G' \sim G$ so that the two actions become simplicially isomorphic.*

4.8. **End compactifications.** Let Γ be an infinite, connected, locally finite graph. (We shall consider the application to the Cayley graph of a finitely generated group.) We shall define a compactification of Γ firstly introduced by Freudenthal, called *end compactification.*

We start with a few notionsin graph theory.

1. If $S \subset \Gamma^0$, then the induced *spanning graph* on S denoted by $\tilde{S}$ has the vertex set S with all edges in Γ^1 of both endpoints in S.
2. If E is a finite set of edges, we denote by $\Gamma \setminus E$ the graph with vertex set Γ^0 and edge set $\Gamma^1 \setminus E$.
3. Let $\mathcal{C}(E)$ be the set of connected components of $\Gamma \setminus E$.

Two points x, y are called *separated* by E if x, y lie in distinct components $\Gamma \setminus E$. Two components C_1, C_2 in $\mathcal{C}(E)$ are distinct if and only if every path between any two points $x \in C_1, y \in C_2$ intersects E^0.

We can define the end compactification of Γ as follows. Consider the directed system $\mathcal{F}(\Gamma)$ of all finite set of edges in Γ with $E < F$ iff $E \subset F$. There is a natural map from $\mathcal{C}(F)$ to $\mathcal{C}(E)$ induced by inclusions of infinite components. Let the *end boundary* $\partial\Gamma$ be the inverse limit of the directed system $\mathcal{C}(E)$ over all finite set of edges E in Γ. By definition, each point $\xi \in \partial\Gamma$ called *an end* is a collection of infinite components $C_E(\xi)$ of $\Gamma \setminus E$ for every $E \in \mathcal{F}(\Gamma)$ so that $C_E(\xi) \cap C_{E'}(\xi)$ is infinite for any two $E, E' \in \mathcal{F}(\Gamma)$. For any E, the component $C = C_E(\xi)$ of $\mathcal{C}(E)$ is uniquely determined by ξ. By abuse of language, we say that a component C contains ξ if $C = C_E(\xi)$. Any two distinct end $\xi \neq \zeta$ are necessarily *separated* by some E: $C_E(\xi) \neq C_E(\zeta)$.

The end boundary $\partial_{\mathcal{E}}\Gamma$ compactifies Γ in the following way. A sequence of points $x_n \in \Gamma$ converges to an end $\xi \in \partial\Gamma$ if and only if for any $E \in \mathcal{F}(\Gamma)$ we have that x_n lies in all $C_E(\xi)$ but finitely many n.

The end boundary is *visual boundary*: any two distinct points are connected by a bi-infinite geodesic. It is not hard to see that a quasi-isometry between locally finite graphs induces a homeomorphism between their ends boundary. In what follows, we shall put a visual metric on end boundary and shows that a quasi-isometry actually induces a bi-Hölder homeomorphism.

4.9. **Visual metrics on end boundary.** Fix a basepoint $o \in \Gamma$ and $0 < \lambda < 1$. We define a metric ρ_λ on $\partial_{\mathcal{E}}\Gamma$ as follows. Let B_n be the edge set of the induced graph on the vertices with distance $\le n$ to the identity o. The inverse limit of the directed system $\{\mathcal{C}(B_n)\}$ is homeomorphic to $\partial_{\mathcal{E}}\Gamma$.

Visual metric. Let ξ, ζ be two distinct ends. If n is the minimal integer such that ξ and ζ belongs to different components in $\mathcal{C}(B_n)$, then define $\rho_\lambda(\xi, \zeta) = \lambda^n$. Equivalently, n is the maximal integer such that ξ and ζ belongs to the same component in $\mathcal{C}(B_n)$.

By definition, the visual metric is *ultrametric*: for any triple of points $x, y, z \in \partial_{\mathcal{E}}\Gamma$,

$$\rho_\lambda(x, y) \le \max\{\rho_\lambda(x, z), \rho_\lambda(z, y)\}$$

Sublinearly biLipschitz equivalence. By definition, if $\rho_\lambda(\xi, \eta) = \lambda^n$ for $n \ge 1$, then every path from ξ and η is within distance n to identity, and some path from ξ and η is disjoint with B_{n-1}.

Let u be a strictly sublinear nondecreasing positive function on the half line such that $\limsup_r u(2r)/u(r) < \infty$.

A $O(u)$-*sublinearly biLipschitz equivalence (SBE)* between metric spaces (X, o) and (Y, o') is a map $f : X \to Y$ if there exists $c > 1$ and $v = O(u)$ such that

(1) $\forall x_1, x_2 \in X$,

$$\begin{aligned} &\frac{1}{c} d(x_1, x_2) - v(\sup\{d(o, x_1), d(o, x_2)\}) \\ &\le d(f(x_1), f(x_2)) \le \\ &c d(x_1, x_2) + v(\sup\{d(o, x_1), d(o, x_2)\}) \end{aligned}$$

(2) $\forall y \in Y, \exists x \in X,\ d(y, f(x)) \le v(d(x, o))$.

Lemma 4.21. *Sublinearly biLipschitz equivalences induce bi-Hölder homeomorphisms between the end boundaries.*

Proof. Two infinite embedded rays γ, γ' are equivalent if, up to removal of finite subpaths, they belong to the same component of $\Gamma_1 \setminus B$ for every ball B. It is easy to verify that each equivalent class $[\gamma]$ determines a unique end, and vice versa. We thus identify the set $\partial_{\mathcal{E}}\Gamma_1$ of ends as the set of infinite embedded rays up to the equivalence.

Let R be a sublinear function. A *sublinear R-neighborhood* of a subset S based at $o \in X$ is defined as follows

$$N_R(S) := \{x \in X : d(x, S) \le R(d(o, x))\}$$

Claim. *Let γ be an infinite embedded ray. If α is an infinite embedded ray in a sublinear R-neighborhood of γ, then α is equivalent to γ.*

Proof of the Claim. Since γ determines an end, for balls B_n of radius n at identity, γ eventually lies in a component $C_n \in \mathcal{C}(B_n)$ and $C_{n+1} \subset C_n$. By the definition of sublinear neighborhood, the distance of points on $\alpha \setminus B_n$ to γ grows sublinearly in n. Thus, for every sufficiently large n, there exists $m < n$ depending on n and R so that $\alpha \setminus B_m$ lies in C_n. And $m \to \infty$ as $n \to \infty$. Thus, $[\alpha] = [\gamma]$. □

Let $f : \Gamma_1 \to \Gamma_2$ be a $O(u)$-SBE between two Cayley graphs of infinitely-ended groups. Without loss of generality, assume that $f(1) = 1$ by (pre-/post-)composing elements in G_1 and G_2. Then there exists a sublinear function R depending on u with the following properties:

(1) for any geodesic ray γ, the R-neighborhood of $f(\gamma)$ contains at least one infinite embedded ray.
(2) If γ is an infinite path outside B_n between two end $p, q \in \partial_{\mathcal{E}}\Gamma_2$, then there exists a continuous path α between $\xi = \Pi^{-1}(p), \eta = \Pi^{-1}(q)$ so that $d(1, \alpha) > n/c - R(n)$.

The Claim with the property (1) thus implies that $f(\gamma)$ determines a unique end: any two infinite embedded rays in $N_R(f(\gamma))$ are equivalent. By abuse of language, we denote the end by $[f(\gamma)]$. Hence, we defined a map $\Phi : \partial_{\mathcal{E}}\Gamma_1 \to \partial_{\mathcal{E}}\Gamma_2$ by $\Phi([\gamma]) = [f(\gamma)]$. By the definition of $O(u)$-SBE and the Claim, it is readily verified that Φ is a bijection from the end of Γ_1 to the end of Γ_2.

Let $\xi, \eta \in \partial_{\mathcal{E}}\Gamma_1$. Assume that $\rho_\lambda(\xi, \eta) = \lambda^n$ where n is the minimal radius of the ball B centered at 1 so that ξ, η lie in distinct components of $\Gamma_1 \setminus B$. Let $p = \Phi(\xi)$ and $q = \Phi(\eta)$. Let m be the maximal radius of the ball B' such that p, q lie in the same component of $\Gamma_2 \setminus B'$. If γ is a path connecting p to q outside B' but lies in a distance $m - 1$ of 1, then by the Property (2), there exists a continuous path α from ξ to η, and $d(1, \alpha) \ge m/c - R(n)$. By definition of n, we have $d(1, \alpha) \le n$. This implies that $m/c - R(n) \le n$. Therefore, $\rho_\lambda(p, q) \ge \lambda^{cn+R(n)} \ge C\rho_\lambda(\xi, \eta)^\beta$ for some $C, \beta > 0$ and every n. □

Corollary 4.22. *A quasi-isometry between locally finite graphs induces a homeomorphism between end boundaries.*

The end compactification, denoted by $\partial_{\mathcal{E}}G$, of a finitely generated group G is defined to be the end boundary of any Cayley graph of G with respect to finite generating set.

Lemma 4.23. *Let G be a finitely genearted group. If $\sharp\partial_{\mathcal{E}}G \ge 3$, then G acts by homeomorphisms on the end boundary as a convergence group action (see definition 12.1).*

Corollary 4.24 (Hopf)**.** *The end boundary of a finitely generated group contains* $0, 1, 2, \infty$ *points.*

Obviously, a finitely generated group has no end iff it is a finite group.

Exercise 4.25. *If a finitely generated group* G *has two ends, then* G *is virtually cyclic.*

4.10. **Groups with infinitely many ends.** It is easy to see that a free product amalgamated over finite edge groups has infinitely many ends. Similarly, HNN extension over finite edge groups do so. The converse is also true and proved by J. Stallings.

Theorem 4.26 (Stalling)**.** *If a finitely generated group* G *has infinitely many ends, then* G *is either a free amalgamated product* $H \star_F K$ *or a HNN extension* $H_{\star_F}$ *over a finite group.*

By Bass-Serre theory, it admits the following equivalent form using actions on trees.

Corollary 4.27. *A finitely generated group* G *has infinitely many ends iff it admits an isometric action on a simplicial tree with finite edge stabilizers. In other words,* G *is infinitely-ended iff* G *is a graph of groups with finite edge groups.*

We now introduce a compactification of a possibly locally infinite tree T.

Let $\partial_G T$ be the Gromov boundary which consists of asymptotic classes of geodesic rays $[\gamma]$. Denote by T^0 by the set of vertices of T. Then the Bowditch boundary ∂T of T is the union of $\partial_G T \cup T^0$. The end boundary is then $\partial_G T \cup T^0$, where T^∞ is the set of vertices of T with infinite degree. Note that any two points $x, y \in \partial T$ is connected by a unique geodesic.

The topology of $\bar{T} = T^0 \cup \partial T$ is given by a neighborhood system of each point $x \in \partial T \cup T$.

(1) $x \in \partial_G T$: each finite set E of eges in T gives a neighborhood $U(x, E)$ of x which consists of $y \in \partial_E T$ such that $[x, y] \cap E = \emptyset$.
(2) $x \in T^0$: each finite set E of eges in T adjacent to x gives a neighborhood $U(x, E)$ of x which consists of $y \in \partial_E T$ such that $[x, y] \cap E = \{x\}$.

A set S in $\bar{T}$ is claimed to be open iff S contains a neighborhood of each point in S. It is clear that

$$U(x, E_1) \cap U(x, E_2) = U(x, E_1 \cup E_2)$$

which implies that the system of open sets indeed defines a topology on $\bar{T}$. With respect to this topology, it is straightforward to verify that $\{U(x, E)\}_E$ is a neighborhood basis of x: each $U(x, E)$ is open.

Lemma 4.28. *The following statements hold:*

(1) *The topology on* $\bar{T}$ *is compact and the set* T^∞ *is dense in* $\bar{T}$.
(2) *The subspace topology on* $\partial_G T$ *coincides with the Gromov topology.*
(3) *The set of vertices in* T *with finite degree are isolated points in* ∂T.
(4) *If a group* G *acts by isometry on* T*, then the isometry action extends to an action on* $\bar{T}$ *by homeomorphism.*

An infinitely-ended group is called *accessible* if it admits a graph of groups with finite edge groups and finite or one-ended groups. Equivalently, it acts on a *terminal tree* T so that finite edge groups are finite and every vertex groups are either finite or one-ended.

Theorem 4.29 (Dunwoody)**.** *A finitely presented infinitely-ended group G is accessible. Moreover, the end boundary of G is homemorphic to the end boundary of T.*

5. Geometry of finitely generated groups I: Growth of groups

The concept of a Cayley graph of a group is a basic tool to study groups in Geometric Group Theory. We will understand a graph in various ways: as a topological space, a combinatorial object and a metric space.

5.1. **Cayley graphs.** Topologically, a graph $\mathcal{G}$ is a 1-dimensional CW-complex. Start with a set V, we glue a set E of intervals $[0,1]$ to X by boundary maps $f_e, e \in E$. A boundary map $f_e : \partial e := \{0,1\} \to X$ sends two boundary points of e to X. Then the graph $\mathcal{G}$ is the quotient space

$$V \sqcup E/_{\partial e \asymp f_e(\partial e)}$$

obtained from the disjoint union $V \sqcup E$ glued by maps f_e.

In this course, it is helpful to take a combinatorial formulation of the notion of a graph.

Definition 5.1. A *graph* $\mathcal{G}$ consists of a set V of vertices and a set E of directed edges. For each directed edge $e \in E$, we associate to e the *initial point* $e_- \in V$ and *terminal point* $e_+ \in V$. There is an orientation-reversing map

$$\bar{}: E \to E, e \to \bar{e}$$

such that $e \neq \bar{e}$, $e = \bar{\bar{e}}$ and $e_- = (\bar{e})_+, e_+ = (\bar{e})_-$.

Remark. Clearly, such a map $\bar{}$ has to be bijective. Moreover, $e_+ = (\bar{e})_-$ can be deduced from other conditions: $e_+ = \bar{\bar{e}}_+ = \bar{e}_-$.

In topological terms, a path in a graph can be understand as a continuous map from $[0,1]$ to the graph. Combinatorially, we define a *path* to be a concatenation of directed edges:

$$\gamma = e_1 e_2 ... e_n, e_i \in E$$

where $(e_i)_+ = (e_{i+1})_-$ for $1 \le i < n$. The initial point γ_- and terminal point γ_+ of γ are defined as follows:

$$\gamma_- = (e_1)_-, \gamma_+ = (e_n)_+.$$

A path γ is called a *cycle* if $(e_n)_+ = (e_1)_-$. A *backtracking* in γ is a subpath of form $e_i e_{i+1}$ such that $e_i = \bar{e}_{i+1}$.

The *concatenation* $\gamma\gamma'$ of two paths γ, γ' is defined in the obvious way, if $\gamma_+ = \gamma'_-$.

Recall that a set S in a group G is said to be *symmetric*, if $s^{-1} \in S$ for any $s \in S$. Here, we do allow $s = s^{-1}$, that is, s is of order 2.

Definition 5.2. Let G be a group and S be a symmetric generating set without identity. The Cayley graph $\mathscr{G}(G,S)$ of G with respect to S is a graph with the vertex set G and edge set $G \times S$. Define $(g,s)_- = g, (g,s)_+ = gs$, and the map

$$\bar{}: G \times S \to G \times S, (g,s) \to (gs, s^{-1}).$$

It is clear to see that $\bar{}$ satisfies the conditions in definition of a graph.

With S being understood as an alphabet set, we can attach every edge/path a label in S in $\mathscr{G}(G,S)$. Define the label function

$$\mathrm{Lab} : G \times S \to S, (g,s) \to s.$$

The label of a path is defined as the natural concatenation of labels of each edge in the path. Thus, there is a natural bijection:

$$\{ \text{ paths originating from identity } \} \leftrightarrow \{ \text{ all words over } S \},$$

in particular, where the set of paths without backtracking issuing from any fixed point is bijective to the set of all reduced words over S.

A *graph morphism* $\phi : \mathcal{G} \to \mathcal{G}'$ between two graphs $\mathcal{G}, \mathcal{G}'$ is a vertex-to-vertex, edge-to-edge map such that $\phi(e_-) = \phi(e)_-, \phi(e_+) = \phi(e)_+$ and $\phi(\bar{e}) = \overline{\phi(e)}$. Left-multiplication of an element $g \in G$ induces a graph automorphism of $\mathscr{G}(G, S)$:

$$V \to V : x \to gx,$$
$$E \to E : (x, s) \to (gx, s).$$

Note that the automorphism also preserves the labels.

Lemma 5.3. *Let G be a group with a generating set S. Assume that $1 \notin S$ and $S = S^{-1}$. The Cayley graph $\mathscr{G}(G, S)$ is a connected graph: there is a path between any two vertices.*

Proof. It suffices to show that any vertex $g \in G$ can be connected by a path to the identity 1. We can write g as a product of generators $g = s_1 s_2 ... s_n$, where $s_i \in S$. Clearly, the path $\gamma = (1, s_1)(s_1, s_2)(s_1 s_2, s_3)...(s_1 s_2 ... s_{n-1}, s_n)$ is a path between 1 and g. □

Recall that a *tree* is a graph such that any nontrivial cycle has no backtracking.

Lemma 5.4. *The Cayley graph of a free group $F(S)$ with respect to the generating set $\tilde{S}$ is a tree.*

Proof. Let $\mathscr{G}(G, S)$ be the Cayley graph. Let γ be a nontrivial cycle in $\mathscr{G}(G, S)$. Applying a graph automorphism induced by the element γ_-^{-1} in G, we can assume that $\gamma_- = \gamma_+ = 1$. Tracing out the path, we obtain a non-empty word $w \in \mathcal{S}$. By definition of a Cayley graph, the word w represents the identity - the empty word. Thus by Theorem 2.9, w is not reduced, and has to contain an inverse pair. So, γ has a backtracking. We conclude that $\mathscr{G}(G, S)$ is a tree. □

A graph is called *locally finite*, if for every vertex, there are only finitely many vertices adjacent to it. In this course, we are mainly interested in finitely generated groups. If a group is finitely generated, then there are exactly $|S|$ edges originating from each vertex. It is clear that the Cayley graph as a topological space is locally compact if and only if it is locally finite. Thus, for a finitely generated group, the Cayley graph has a good topology to handle with.

5.2. **Word metrics and Growth function.** In this subsection, we put a natural metric called word metric on a group, of which we give two definitions below.

Let G be a group with a generating set S. We first define a norm on a group G.

Definition 5.5. The *word norm* $|g|_S$ of an element $g \in G$ is the shortest length of a word w such that $w =_G g$. Formally,

$$|g|_S = \inf\{|w| : \phi(w) = g, \phi : F(S) \to G\}.$$

We define the *word metric* d_S on G: $d_S(g, h) = |g^{-1}h|_S$ for any $g, h \in G$.

Remark. d_S is indeed a metric, as $|gh| \le |g| + |h|$, and $|g| = 0$ iff $g = 1 \in G$

The word metric can also be obtained by metrizing the Cayley graph $\mathscr{G}(G, S)$. For each edge $e \in \mathscr{G}(G, S)$, we assign the unit length 1 to it. The length of a path is the number of edges in it. We define a distance d_S on $\mathscr{G}(G, S)$: for any two vertices x, y, the distance $d_S(x, y)$ is the infimum of lengths of all paths between x, y. It is easy to see that this definition of word metric d_S is equal to the one defined as above.

A metric d on G is called *proper* if the ball around the identity of any finite radius contains only finitely many elements. That is, $|\{g \in G : d(1, g) \le n\}|$ is finite for any $n > 0$.

A metric d on G is called *left-invariant* if the left multiplication on G is an isometry: $d(gx, gy) = d(x, y)$ for any $g, x, y \in G$.

Lemma 5.6.
1. *Word metric d_S is a left-invariant metric on G. In other words, G acts on itself by isometries.*
2. *Word metric on the Cayley graph is a geodesic metric: the distance of any two vertices can be realized as the length of a path between them.*
3. *If $|S| < \infty$, d_S is proper.*

It should be noted that word metric depends on the choice of a generating set. However, word metrics with respect to different generating sets are bi-Lipschitz.

Lemma 5.7. *Let S, T be two finite generating sets of G. Then thee exists a constant $C \ge 1$ such that*

$$C^{-1} d_T(., .) \le d_S(., .) \le C d_T(., .).$$

Proof. For any generator $s \in S$, there is a word w_s over $\tilde{T}$ such that $w =_G s$. Set

$$C_1 = \max\{|w_s| : s \in S\}.$$

Observe that any word over $\tilde{S}$ can be written as a word over $\tilde{T}$. Thus, for any element $g \in G$, we have $|g|_T \le C_1 |g|_S$.

The other inequality can be obtained in the same way. □

From now on, we assume that G is a group with a finite generating set. As shown above, word metric is a proper metric. So the first information that we want to know about a Cayley graph is to see how the number of elements in a ball increases.

We prepare some notations first. Let $B_S(g, n) = \{h \in G : d_S(g, h) \le n\}$ denote the ball at g of radius n, and $S(g, n) = \{h \in G : d_S(g, h) = n\}$ denote the sphere around g of radius n. We now define a growth function

$$\phi_S : \mathbb{N} \to \mathbb{N}, \phi(n) = |B_S(1, n)|$$

for all $n \ge 0$.

Exercise 5.8. $|S(1, n+m)| \leq |S(1,n)| \cdot |S(1,m)|$ *for any* $n, m \in \mathbb{N}$.

Let $\phi, \varphi : \mathbb{N} \to \mathbb{N}$ be two monotonically non-discreasing functions. We say that ϕ *dominates* φ if there exists $C \geq 1$ such that

$$\varphi(n) \leq C\phi(Cn)$$

for $n > 0$. Denote $\varphi \prec \phi$.

Two functions ϕ, φ are *equivalent* if they dominate each other. Note that $a_0 + a_1 n + ... + a_i n^i$ is equivalent to n^i. All exponential functions like λ^n for $\lambda > 1$ are equivalent to the the standard one e^n. But e^n dominates any polynomial function.

Lemma 5.9. *Let S, T be two finite generating sets of G. Then $\phi_S(n)$ and $\phi_T(n)$ are equivalent.*

Thus, if a group has a polynomial growth function for one generating set, then it has a polynomial growth function for any other generating set (with same degree).

In the sequel, we will not distinguish growth functions that are equivalent. In other words, we are only interested in the equivalent classes of a growth function.

Lemma 5.10. *Let A be an abelian group with finite rank d. Then the growth function of A is equivalent to n^d.*

Proof. By the classification theorem 1.21 of abelian groups, any abelian group of rank d is isomorphism to a direct product of a free abelian group of rank d and a finite group.

Observe that the growth function of $G \times F$ for a finite group F is equivalent to that of G. Thus, it suffices to prove that the growth function of Z^d with respect to the standard generating set is polynomial.

The growth function of Z^2 is $2n^2+2n+1$, which is equivalent to n^2. Use induction on the rank. Assume that Z^{d-1} has growth function n^{d-1}. Let $S = \{s_1, s_2, ...s_d\}$ be the standard generating of Z^d. Any element g of length n can be written as $g = s_1^{n_1} s_1^{n_1} ... s_d^{n_d}$, where $s_i \in S$ and $\sum_{1 \leq i \leq n} n_i = n$. Denote $h = s_1^{n_1} s_1^{n_1} ... s_{d-1}^{n_{d-1}}$. Then h is an element of length $n-1$ in Z^{d-1}. Then $|B(1,n)| \prec n \cdot n^{d-1} = n^d$. □

Lemma 5.11. *Free groups of rank at least 2 has exponential growth function.*

Proof. We calculate the growth function of F_2 with respect to the standard generating set. Let $b_n = |S(1,n)|$. Then $b_0 = 1, b_1 = 4$. Observe that we have $b_{n+1} = 3b_n$ for $n > 0$. Then $\phi(n) = 1 + \sum_{1 \leq i \leq n} b_n = 1 + 4(1 + 3 + ... + 3^{n-1}) = 2 \cdot 3^n - 1$. □

5.3. **Nilpotent groups and Polynomial growth.** Let G be a group. Define inductively the following groups, called *lower central series* of G.

$$G_0 = G, G_1 = [G, G_0], G_2 = [G, G_1], ...G_n = [G, G_{n-1}], ...$$

The group G is called *nilpotent* of degree $n \geq 1$ if $G_n = \{1\}$, but $G_{n-1} \neq \{1\}$ if $n \geq 2$.

Example 5.12. The following group in $GL(2, \mathbb{R})$

$$G = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} : a, b, c \in \mathbb{R} \right\}$$

is a nilpotent group of degree 2.

Exercise 5.13. *Let G be a nilpotent group. Prove that G_i/G_{i+1} is an abelian group.*

We collect some useful results about nilpotent groups.

Lemma 5.14. *A group G is a nilpotent group of degree (at most in the "<=") n if and only if any n-fold commutator is trivial:*

$$[[...[[g_0, g_1], g_2], ...], g_n] = 1, \forall g_i \in G$$

Proof. The direction "=>" is clear. The direction "<=" is given by the following general fact for any group G.

For any $n \geq 1$, the G_n is generated by the set of all possible n-fold commutators. This is proved by an induction on n. When $n = 1$, it is clear. The induction is completed by using the following identity.

$$[ab, c] = a[b, c]a^{-1}[a, c].$$

Note that the inverse of a i-fold commutator is i-fold commutator. Indeed, we write

$$[[...[[g_0, g_1], g_2], ...], g_n]^{-1} = [g_n, [b, g_{n-1}]]$$

where $b := [[...[[g_0, g_1], g_2], ...], g_{n-2}]$ is a $(n-2)$-fold commutator. Note the identity $[a, [b, c]] = [a[b, c]a^{-1}, a^{-1}] = [[aba^{-1}, aca^{-1}], a^{-1}]$. So we obtain that

$$[g_n, [b, g_{n-1}]] = [g_n[b, g_{n-1}]g_n^{-1}, g_n^{-1}]$$

is n-fold commutator.

In conclusion, if a, b are i-fold commutators and $c \in G$, then $[ab, c]$ is a product of $(i + 1)$-fold commutators. □

Lemma 5.15. *Let G be a nilpotent group of degree n.*

1. *Any subgroup in G is nilpotent of degree at most n.*
2. *If G is finitely generated, then $[G, G]$ is a finitely generated nilpotent group of degree at most $n - 1$.*
3. *If G is finitely generated, then any subgroup in G is finitely generated.*

Proof. **(1)**: The statement (1) is clear.

(2):Assume that G is generated by a finite set S. Let X be the set of all m-fold commutator over $S \cup S^{-1}$

$$[[...[[s_0, s_2], s_3], ...], s_m], \forall s_i \in S \cup S^{-1}$$

where $m < n$. Then X is a finite set. We claim that X is a generating set for $[G, G]$. Recall that $[G, G]$ is the group generated by all commutators $\{[f, g], f, g \in G\}$. It suffices to show that every $[f, g]$ can be written as a word over S.

Assume that $f = s_1 s_2 ... s_k$ and $g = t_1 t_2 ... t_l$ for $s_i, t_j \in S \cup S^{-1}$. We apply the identity

$$st = [s, t]ts$$

to $[f, g]$ to get

$$\begin{aligned}
[f, g] &= s_1 s_2 ...(s_k) t_1 t_2 ... t_l s_k^{-1} ... s_2^{-1} s_1^{-1} t_l^{-1} ... t_2^{-1} t_l^{-1} \\
&= s_1 s_2 ...[s_k, t_1] t_1 (s_k) t_2 ... t_l s_k^{-1} ... s_2^{-1} s_1^{-1} t_l^{-1} ... t_2^{-1} t_l^{-1} \\
&= s_1 s_2 ...[s_k, t_1] t_1 [s_k, t_2] t_2 (s_k) ... t_l s_k^{-1} ... s_2^{-1} s_1^{-1} t_l^{-1} ... t_2^{-1} t_l^{-1} \\
&= s_1 s_2 ...[s_k, t_1] t_1 [s_k, t_2] t_2 ... t_l (s_k) s_k^{-1} ... s_2^{-1} s_1^{-1} t_l^{-1} ... t_2^{-1} t_l^{-1}.
\end{aligned}$$

After finitely times, s_k and s_k^{-1} are cancelled in $[f, g]$. We do the same procedure to cancel s_i and s_i^{-1}, t_j and t_j^{-1} in $[f, g]$. In the final form, we have that $[f, g]$ is a product of m-fold commutators where $m < n$, as any i-fold commutator for $i \geq n$ is trivial. This proves that $[G, G]$ is generated by X.

(3). Let H be a subgroup of G. Note that $H \cap G_1$ is normal in H, and we have an injective homomorphism

$$H/(H \cap G_1) \to G/G_1.$$

As G/G_1 is abelian and finitely generated, any subgroup in G/G_1 is abelian and finitely generated by Corollary 1.22. Thus, $H/(H \cap G_1)$ is also finitely generated and abelian.

We use induction on the degree n of a nilpotent group. If $n = 1$, then G is abelian. The conclusion follows from Corollary 1.22. Assume now that the (3) holds for every finitely generated nilpotent group of degree $d < n$.

Let G be of degree n. Observe that $G_1 = [G, G]$ is of degree at most $n - 1$. Thus, $(H \cap G_1)$ is finitely generated. By Exercise 1.5, we see that H is finitely generated. □

Theorem 5.16 (J. Wolf, 1968)**.** *A finitely generated nilpotent group has a polynomial growth function.*

Proof. Let G be a nilpotent group of degree d generated by a finite symmetric set $1 \notin S$. We use induction on n. When $d = 1$, it is true by Lemma 5.10. By Lemma 5.15, $[G, G]$ is of degree at most $d - 1$ and is generated by a finite set X of i-fold commutators for $i < d - 1$.

Denote $S = \{s_1, s_2, \cdots, s_m\}$. We estimate the growth function ϕ of G with respect to S.

Let $g \in S(1, n)$ and write $g = s_{i_1} s_{i_2} \cdots s_{i_n}$, where $s_{i*} \in S$. We apply the identity

$$st = ts[s^{-1}, t^{-1}]$$

to the product presentation of g and turn it into a "standard" presentation as follows:

$$g = s_1^{\epsilon_1} \cdots s_m^{\epsilon_m} h$$

where $|\epsilon_1| + \cdots + |\epsilon_m| = n$ and $h \in [G, G]$.

We now show the following claim.

Claim. *There exists a constant $C > 1$ such that h is a word of length at most Cn^d over X. That is, h can be written as a product of at most Cn^d generators in X.*

Proof of Claim. Let's see the process of $s_{i_1} s_{i_2} ... s_{i_n}$ to $s_1^{\epsilon_1} ... s_m^{\epsilon_m} h$. We first move s_1 to the left. Suppose, in the worst case, that $s_{i_n} = s_1$. To move s_{i_n} to the left, we need at most n swaps of s_{i_n} with other generators s_{i_*}. However, in the process, we produce at most n commutators. We then move s_2 to left. In this step, we have to make at most $2n$ swaps with either s_{i*} or 1-fold commutators produced in previous step.

Observe that after moving $n \geq l \geq 1$ letters to the left, there are at most $n \cdot C_l^0$ letters, $n \cdot C_l^1$ 1-fold commutators, ..., $n \cdot C_l^i$ i-fold commutators, ... and $n \cdot C_l^l$ l-fold commutators.

However, for any $l \geq d$, every l-fold commutator is trivial. Hence, after moving all $n \geq 1$ letters to the left, we actually have a constant $C > 1$ such that there are at most

$$n \cdot C_n^0 + n \cdot C_n^1 + ... + n \cdot C_n^{d-1} \leq Cn^d.$$

i-fold commutators for $i < d$ in the whole process. In conclusion, we transformed $g = s_{i_1} s_{i_2} ... s_{i_n}$ to

$$g = s_1^{\epsilon_1} ... s_m^{\epsilon_m} h$$

where h is a word of length at most Cn^d over X. □

As shown above, $G_1 = [G, G]$ is of degree at most $d-1$. By induction assumption, the growth function φ of G_1 with respect to X is polynomial, say $\varphi(n) = n^k$ for some $k > 0$.

By the claim, $h \in G_1$ is a word of length at most Cn^d. There are at most $\varphi(Cn^d) = Cn^{dk}$ such elements in G_1.

On the other hand, there are at most n^m elements of form $s_1^{\epsilon_1} ... s_m^{\epsilon_m}$. In fact, such type of elements can be seen as elements in a free abelian group of rank m, and thus, the number of such elements of length n is at most n^m.

As a consequence, $|S(1, n)|$ contains at most n^{dk+m} elements. It follows that $\phi(n) = |B(1, n)| < n^{dk+m+1}$. □

Remark. The degree $m + dk + 1$ in the proof is nevertheless optimal. The Bass-Guivarc'h formula by Guivarc'h(1971) and Bass(1972) states that the minimal degree of a growth function of G is

$$d(G) = \sum_{k \geq 0} (k + 1) \ \operatorname{rank}(G_k / G_{k+1})$$

where $G = G_0 \supseteq G_1 \supseteq \ldots$ is the lower central series of G. The rank denotes the rank of an abelian group G_k / G_{k+1}.

In 1981, Gromov proves the converse of Theorem 5.16. Given a property P, we say a group G has *virtually* property P if G contains a finite index subgroup which has property P.

Theorem 5.17 (Gromov, 1981)**.** *Let G be a finitely generated group of polynomial growth, then G is virtually nilpotent: G contains a finite index subgroup which is nilpotent.*

Exercise 5.18. *Find a non-abelian group but which is virtually abelian, and a non-nilpotent group but which is virtually nilpotent.*

Thus, Theorem 5.17 could not be strengthened to get that G is nilpotent.

5.4. **Growth types of groups.** Let G be a group with a finite symmetric generating set $1 \notin S$. Denote $B_n = |B(1, n)|$ the number of elements in the ball $B(1, n)$. Denote $b_n = |S(1, n)|$ the number of elements in the sphere $S(1, n)$.

Definition 5.19. The *growth rate* of G with respect to S is the following limit

$$\delta_G = \lim_{n \to \infty} n^{-1} \ln B_n$$

Remark. The limit exists because of the sub-multiplicative inequality:

$$b_{n+m} \leq b_n b_m.$$

Clearly, $\delta_G = 0$ if G has polynomial growth function. Thus, $\delta_G = 0$ for any nilpotent group. On the other hand, $\delta_G > 0$ for any free group of rank at least 2.

Exercise 5.20. (1) *(Fekete Lemma)Let a_n be a sequence of positive numbers such that $a_{n+m} \leq a_n + a_m$. Then*
$$\limsup_{n\to\infty} n^{-1}a_n = \inf\{a_n/n : n \geq 1\}.$$
(2) *Consider the following growth series*
$$\theta(z) = b_0 + b_1 z + ... + b_n z^n + ..., z \geq 0$$
and
$$\Theta(z) = B_0 + B_1 z + ... + B_n z^n + ..., z \geq 0$$
Prove that $\theta(z)$ and $\Theta(z)$ have the same nature of convergence: $\theta(z)$ is convergent iff $\Theta(z)$ is convergent. Their convergence radius is $e^{-\delta_G}$, and $\theta(z)$ and $\Theta(z)$ both diverge at $e^{-\delta_G}$.

Exercise 5.21. (1) *The growth function of a finitely generated group always dominates that of any finitely generated subgroup.*
(2) *The growth function of a finitely generated group always dominates that of any quotient group.*

Thus, maximal growth functions that a finitely generated group could have are exponential functions.

Definition 5.22. Let G be a finitely generated group. Let $\phi(n)$ be the growth function of G.
(1) (Polynomial growth) G has *polynomial* growth if there exists $d \in \mathbb{N}$ such that $\phi(n) \prec n^d$.
(2) (Exponential growth) G has *exponential* growth if $e^n \prec \phi(n)$.
(3) (Intermediate growth) G has *intermediate* growth if G does not belong to the polynomial and exponential growth types.

Remark. Many classes of groups have either polynomial or exponential growth. For example, there are no groups of intermediate growth in linear groups (Tits alternative), in solvable groups (Milnor, Wolf)... However, Grigorchuk constructed the first group of intermediate growth in 1983, answering a long-standing open question of Milnor about whether there exist groups of intermediate growth.

Exercise 5.23. *A finitely generated group is of exponential growth if and only if the growth rate with respect to some (or any) generating set is positive.*

By Theorem 5.16, virtually nilpotent groups can not contain a free subgroup of rank at least two. On the other hand, since there exists no relation in a free group, a free group of rank at least two can not contain a non-cyclic nilpotent group. In a word, the common intersection of the class of free groups and the class of nilpotent groups are cyclic groups.

The equivalent class of growth function is a quasi-isometric invariant, cf. Lemma 5.9. However, the growth rate depends on the choice of generating sets.

Define the minimal growth rate δ_G of a finitely generated group G as follows
$$\delta_G = inf\{\delta_{G,S}\}$$
where the infimum is taken over all finite generating set S of G.

Exercise 5.24. *Let F be a free group of rank n.*
(1) *Let S be the standard generating set of F. Prove that $\delta_{F,S} = \log(2n-1)$.*

(2) *Let T be a finite generating set of F. Prove that $\delta_{F,T} \geq \log(2n-1)$. Therefore, we see that δ_F is realized by some generating set: $\delta_F = \delta_{F,S} = \log(2n-1)$.*

(3) *By the second statement, explain that the value of the growth rate is not invariant under quasi-isometries.*

(Tips: Use the epimorphism $F \to F/[F,F]$ to prove that T contains a subset T_1 such that $|T_1| = n$ and T_1 freely generates a subgroup. Here you may want to use Exercise 2.30.)

Exercise 5.25. *Let H, K be two groups with finite generating sets S, T respectively. Denote by θ_H and θ_K be the corresponding growth series of H and K. Then with respect to $S \cup T$, the growth function of the free product $\Gamma = H \star K$ is given by*

$$\theta_\Gamma(z) = \frac{\theta_H(z)\theta_K(z)}{1 - (\theta_H(z) - 1)(\theta_K(z) - 1)}$$

6. Geometry of finitely generated groups II: Svarc-Milnor Lemma

6.1. **Length metric spaces.** Let (X, d) be a metric space. Let $p : [a, b] \to X$ be a parameterized continuous path, where $a, b \geq 0$. It is called *rectifiable* if

$$\sup \sum_{0 \leq i \leq n} d(p(t_i), p(t_{i+1})) < \infty \tag{7}$$

over all finite partitions $\{t_0 = a, ..., t_n = b\}$ of $[a, b]$. The *length* $\mathrm{Len}(p)$ of p is defined to be the supremum of the above sum (7) over all possible partitions of $[a, b]$.

Definition 6.1. Let (X, d) be a metric space. We define an induced metric $\bar{d}$ called *length metric* as follows. Let $x, y \in X$ be two points. Then $\bar{d}(x, y)$ is the infimum of lengths of all possible rectifiable paths between x, y.

If $d = \bar{d}$, then (X, d) is called a *length metric* space.

Definition 6.2. A metric space (X, d) is called *proper* if any closed ball $\bar{B}(x, r)$ at $x \in X$ with radius $r \geq 0$ is compact.

Remark. The terminology of a "proper" metric space comes from the continuous map $X \to \mathbb{R}$,

$$x \in X \to d(o, x) \in \mathbb{R}.$$

A metric space is proper if and only if the above map is proper in the *topological* sense: the preimage of a compact set is compact.

Theorem 6.3 (Hopf-Rinow)**.** *Let (X, d) be a length metric space. Then (X, d) is proper if and only if it is a locally compact and complete space.*

Remark. The assumption that X is a length metric space is necessary in the direction "<=". For example, consider a metric space where the distance between two distinct points is 1. This metric space is complete and locally compact (every point is closed and open). But the ball of of radius 1 is not compact, unless it consists of finitely many points.

Proof. The direction => follows by definition. We prove the other direction.

Fix $x \in X$, we define the real number in $[0, \infty]$

$$R = \sup\{r \geq 0 : \bar{B}(x, r) \textit{ is compact}\}.$$

As X is locally compact, we have $R > 0$. We argue by way of contradiction. Assume that $R < \infty$.

We first show that $B = \bar{B}(x, R)$ is compact. We use the following criterion of compactness: A metric space is compact if and only if it is complete and totally bounded.

As B is closed in a complete space, it is complete. We now prove that B is totally bounded. For any $\epsilon > 0$, we need find a finite set S in B such that $B \subset N_\epsilon(S)$.

Note that $\bar{B}(x, R - \epsilon/3)$ is compact by definition of R. By totally boundedness, there exists a finite set S in $\bar{B}(x, R - \epsilon/3)$ such that

$$\bar{B}(x, R - \epsilon/3) \subset N_{\epsilon/3}(S)$$

Since X is a length metric space, we have

$$B = \bar{B}(x, R) \subset N_{\epsilon/2}(\bar{B}(x, R - \epsilon/3) \subset N_\epsilon(S).$$

Thus, B is totally bounded and then compact.

For any $y \in \bar{B}$, there exists an open ball U_y at y such that $\bar{U}_y$ is compact. By compactness of B, we can find finitely many U_y to cover B, and denote by U the finite union of U_y. Then there exists $\epsilon > 0$ such that $N_\epsilon(B) \subset U$. Indeed, consider the continuous function $f(x) = d(x, X \setminus U)$. By compactness of $\bar{B}$, $\epsilon = \inf_{x \in B}\{f(x)\} > 0$ gives the desired constant.

As $\bar{U}$ is closed, we have $\bar{N}_\epsilon(B) = \bar{B}(x, R+\epsilon)$ is compact. This gives a contradiction to the definition of R. This proves that $R = \infty$. □

6.2. **Geodesics.** Let $p_1 : [a,b] \to X, p_2 : [c,d] \to X$ be two paths. We say p_1 and p_2 are *equivalent* if there exists a continuous and monotonically increasing function $\phi : [a,b] \to [c,d]$ such that $p_2(\phi(t)) = p_1(t)$. Clearly, the length of a path does not depend on its parameterization.

Define $\phi : [a,b] \to [0, \mathrm{Len}(p)]$ by $\phi(t) = \mathrm{Len}(p[a,t])$. Then ϕ is continuous and monotonically increasing. We define the *length parameterization* $\bar{p} : [0, \mathrm{Len}(p)] \to X$ of p as follows. Let $s \in [0, \mathrm{Len}(p)]$. There exists $t \in [a,b]$ such that $\mathrm{Len}(p[a,t]) = s$. Set $\bar{p}(s) = p(t)$. Clearly, $\bar{p}(s)$ does not depend on the choice of t. It is also easy to check that $\bar{p}$ is continuous.

Note that

$$\mathrm{Len}(\bar{p}[0,s]) = s$$

where $0 \leq s \leq \mathrm{Len}(p)$.

We often use the *linear parameterization* of p which is defined as

$$\hat{p} : [0,1] \to X, \ \hat{p}(t) = \bar{p}(\mathrm{Len}(p) \cdot t)$$

Definition 6.4. A path p is called a *geodesic* if $\mathrm{Len}(p) = d(p_-, p_+)$. Equivalently, a path is a geodesic if its length parametrization $\bar{p} : [0, \mathrm{Len}(p)] \to X$ is an isometric map.

A metric space is called a *geodesic* metric space if there exists a geodesic between any two points.

We shall prove that a proper length metric space is a geodesic space. In order to do so, we need the following result, which is a special case of Arzela-Ascoli Theorem.

Lemma 6.5. *Let (X,d) be a compact metric space, and $p_n : [0,1] \to X$ a sequence of linearly parameterized paths with uniformly bounded length. Then*

1. *There exists a subsequence p_{n_i} of p_n which uniformly converges to a path $p_\infty : [0,1] \to X$.*
2. *For any $\epsilon > 0$, there exists $N = N(\epsilon) > 0$ such that*
$$Len(p_\infty) \leq Len(p_{n_i}) + \epsilon$$
for all $n_i > N$.

Proof. **(1)**. Since p_n are linearly parameterized paths with uniformly bounded lengths. Thus, there exists $C > 1$ (=supermum of lengths) such that

$$d(p_n(t), p_n(s)) \leq C|t-s| \tag{8}$$

for all $n > 0$.

We choose a countable dense subset A, for example, the rational numbers, in $[0,1]$. Using a diagonal argument, we obtain a subsequence of $\{p_n\}$, still denoted by $\{p_n\}$, which are convergent on A.

We next show that p_n are convergent at any $t \in [0,1]$. As X is complete, it suffices to show that $\{p_n(t)\}$ is a Cauchy sequence. Let $\epsilon > 0$. Choose $a \in A$ such that $|a-t| \le \epsilon/4C$. As $\{p_n(a)\}$ is convergent, there exists $N = N(\epsilon)$ such that $|p_n(a) - p_m(a)| \le \epsilon/2$ for $n, m > N$. Thus,

$$\begin{aligned} d(p_n(t), p_m(t)) &\le d(p_n(t), p_n(a)) + d(p_n(a), p_m(a)) + d(p_m(a), p_m(t)) \\ &\le \epsilon/4 + \epsilon/2 + \epsilon/2 \le \epsilon \end{aligned} \tag{9}$$

for $n, m > N$.

Assume that p_n converges pointwise to $p_\infty : [0,1] \to X$. Thus, by (8), we obtain

$$d(p_\infty(t), p_\infty(s)) \le C|t-s| \tag{10}$$

for $t, s \in [0,1]$.

We now show that $p_n \to p_\infty$ uniformly. Let $\epsilon > 0$. We want to find $N = N(\epsilon)$ such that $d(p_n(t), p_\infty(t)) \le \epsilon$ for any $t \in [0,1]$ and $n > N$.

Let $M = 4C/\epsilon$. We consider the pointwise convergence of $p_n(t)$, where $t = i/M, 0 \le i \le M$. There exists $N > 0$ such that

$$d(p_n(i/M), p_\infty(i/M)) \le \epsilon/2,\ 0 \le i \le M$$

for $n > N$.

Thus for each $t \in [0,1]$, there exists $|t - i/M| \le \epsilon/4C$. Hence,

$$\begin{aligned} d(p_n(t), p_\infty(t)) &\le d(p_n(t), p_n(i/M)) + d(p_n(i/M), p_\infty(i/M)) + d(p_\infty(i/M), p_\infty(t)) \\ &\le \epsilon/4 + \epsilon/2 + \epsilon/2 \le \epsilon \end{aligned} \tag{11}$$

for $n > N$. This proves that $p_n \to p_\infty$ uniformly.

(2). Let $\{0 = t_0, t_1, ..., t_m = 1\}$ be a finite subdivision of $[0,1]$ such that

$$\mathrm{Len}(p_\infty) \le \epsilon/2 + \sum_{0 \le i < m} d(p_\infty(t_i), p_\infty(t_{i+1})).$$

As $p_n \to p_\infty$ uniformly, there exists $N = N(\epsilon/(4m)) > 0$ such that

$$d(p_n(t), p_\infty(t)) \le \epsilon/(4m)$$

for any $t \in [0,1]$.

Thus, we have the following

$$\begin{aligned} \mathrm{Len}(p_\infty) &\le \epsilon/2 + \textstyle\sum_{0 \le i < m} d(p_\infty(t_i), p_\infty(t_{i+1})) \\ &\le \epsilon/2 + \epsilon/(4m) \cdot (2m) + \textstyle\sum_{0 \le i < m} d(p_n(t_i), p_n(t_{i+1})) \\ &\le \epsilon + \mathrm{Len}(p_n). \end{aligned}$$

for any $n > N$. □

We now prove the following.

Theorem 6.6. *A proper length metric space is a geodesic metric space.*

Proof. Let $x, y \in X$. As X is a length metric space, there exists a sequence of rectifiable curves p_n between x, y such that $\mathrm{Len}(p_n) \to d(x,y)$.

Since X is proper, we can assume that p_n are contained in a closed ball in X. By Lemma 6.5, there exists a subsequence of p_n uniformly converging to a path p_∞. Then by the second statement of Lemma 6.5, p_∞ is a geodesic between x, y. □

Remark. The "proper" assumption is necessary: consider the Euclidean plane minus the origin with induced length metric. Here, it is not a geodesic space, because it is not complete. See Theorem 6.3.

On the other hand, there are non-proper length metric space which are geodesic. For instance, any (non-locally finite) graph with each edge length 1 is a geodesic metric space. See [4, Part I] for more examples.

Let $f_n : X \to Y$ be a sequence of functions. Recall that f_n converges *locally uniformly* to f_∞ if $f_n|_K \to f_\infty|_K$ uniformly on every compact set $K \subset X$. Another useful result is the following.

Theorem 6.7. *Let (X, d) be a proper length metric space. Given $o \in X$, let $p_n : [0, \infty)$ be a sequence of length parameterized geodesic rays with the same origin $(p_n)_- = o$. Then there exists a subsequence of p_n which converges locally uniformly to a geodesic ray $p_\infty : [0, \infty)$ with $p_\infty(0) = o$.*

Proof. Consider the sequence of compact balls $\bar{B}(o, n), n \in \mathbb{N}$. We apply Lemma 6.5 to $\bar{B}_(o, n)$. By a diagonal argument, we obtain a subsequence of p_n which converges locally uniformly to a geodesic ray p_∞. □

Exercise 6.8. *Complete the proof of Theorem 6.7.*

6.3. **Quasi-isometries.**

Definition 6.9. Let $\phi : (X, d_X) \to (Y, d_Y)$ be a map between two metric spaces. Given constants $\lambda \ge 1, c > 0$, ϕ is called a (λ, c)-*quasi-isometric embedding map* if the following inequality holds

$$\lambda^{-1} d_X(x, x') - c \le d_Y(\phi(x), \phi(x')) \le \lambda d_X(x, x') + c, \tag{12}$$

for all $x, x' \in X$.

If, in addition, there exists $R > 0$ such that $Y \subset N_R(\phi(X))$, then ϕ is called a (λ, c)-*quasi-isometry*. In this case, we also say that X is *quasi-isometric* to Y.

Remark. When λ, c are clear in context or do not matter, we omit them and just say ϕ is a quasi-isometric embedding or a quasi-isometry. In general, a quasi-isometric embedding ϕ is not continuous.

We see a few simple examples first.

Examples 6.10.
(1) Any two metric spaces of bounded diameter are quasi-isometric.
(2) The inclusion $\mathbb{Z}^n \to \mathbb{R}^n$ is a $(1, 0)$-quasi-isometric embedding map, where $\mathbb{Z}^n$ is equipped with the word metric, and $\mathbb{R}^n$ is equipped the usual metric. In fact, this is a quasi-isometry.

A quasi-isometric embedding map is generally not an injective map or even a continuous map! However, it is injective on a large scale.

Exercise 6.11. *Let $\phi : X \to Y$ be a (λ, c)-quasi-isometric embedding map. Then there exists a constant $C = C(\lambda, c) > 0$ such that if $d_X(x, x') > C$, then*

$$d_Y(f(x), f(x')) > 0.$$

Hence, the quasi-isometry can (only) capture the large scale, coarse geometry of metric spaces.

Let $\phi : X \to Y$ be a quasi-isometric embedding. A map $\psi : Y \to X$ is called a *quasi-inverse* of ϕ if the following hold

$$d_X(\phi \cdot \psi(y), y) \le R, \ \forall y \in Y.$$

for some finite $R > 0$. It is obvious that if a quasi-isometric embedding admits a quasi-inverse, then it must be a quasi-isometry. On the other hand, by definition, a quasi-isometry always admits a quasi-inverse, and any two quasi-inverses are uniformly bounded.

We prove some useful properties about quasi-isometries.

Lemma 6.12. (1) *Let $\phi : X \to Y$ and $\psi : Y \to Z$ be two quasi-isometric embeddings. Then $\psi \cdot \phi : X \to Z$ is a quasi-isometric embedding.*
(2) *Let $\phi : X \to Y$ be a quasi-isometry. Then any quasi-inverse of ϕ is a quasi-isometry $\psi : Y \to X$. So a quasi-isometric embedding is a quasi-isometry if and only if it has a quasi-inverse.*

Proof. We only prove the (2). As $\phi : X \to Y$ is a quasi-isometry, there exists $R > 0$ such that

$$Y \subset N_R(\phi(X)).$$

That is to say, for any $y \in Y$, there exists $x \in X$ such that $d(y, \phi(x)) < R$. By definition of a quasi-isometric map, we have

$$\mathrm{Diam}(\phi^{-1}(y)) \le \lambda c. \tag{13}$$

This gives a (non-canonical) way to define a map $\psi : Y \to X$ by $\psi(y) = x$, where $d(\phi(x), y) \le R$. By definition, ψ is a quasi-inverse of ϕ. We claim that $\psi : Y \to X$ is a quasi-isometry. By (13), we see that $X \subset N_{c\lambda}(\psi(Y))$. Hence, it suffices to verify that ψ satisfies the inequality (12).

By definition of ψ, we have the following

$$d_Y(\phi(\psi(y)), y) = d_Y(\phi(x), y) \le R \tag{14}$$

for any $y \in Y$.

As ϕ is a (λ, c)-quasi-isometric map, we have

$$d_X(\psi(y), \psi(y')) \le \lambda d_Y(\phi(\psi(y)), \phi(\psi(y'))) + c \le \lambda d_Y(y, y') + 2\lambda R + c.$$

Similarly, we now verify the other inequality:

$$d_X(\psi(y), \psi(y')) \ge \lambda^{-1} d_Y(\phi(\psi(y)), \phi(\psi(y'))) - c \ge \lambda^{-1} d_Y(y, y') - 2\lambda^{-1} R - c.$$

Hence, we proved that $\psi : Y \to X$ is a $(\lambda, 2\lambda R + c)$-quasi-isometry. □

We consider the set of all quasi-isometries of X. Two quasi-isometries $\phi, \psi : X \to X$ are called *equivalent* if they differ by a bounded constant: $||\phi - \psi||_\infty < \infty$. Denote by $QI(X)$ the set of equivalent classes of quasi-isometries of X.

Exercise 6.13. *The set $QI(X)$ with the composition operation is a group. Moreover, there exists a homomorphism from the isometry group $Isom(X)$ of X into the group $QI(X)$.*

Exercise 6.14. *Suppose two metric spaces X, Y are quasi-isometric. Then $QI(X)$ is isomorphic to $QI(Y)$ (given by conjugating the isometric actions on X).*

As proven in Lemma 5.7, word metrics with respect to two finite generating sets are bi-Lipichitz. Clearly, a group with word metric is quasi-isometric to its Cayley graph with induced length metric. So two Cayley graphs with respect to different generating sets are quasi-isometric. Here, we can actually construct an equivariant quasi-isometry between them.

Lemma 6.15. *Let G be a finitely generated group with two finite generating sets S, T. Then there exists a quasi-isometry $\phi : \mathscr{G}(G, S) \to \mathscr{G}(G, T)$ such that $\phi(gx) = g\phi(x)$ for any $x \in \mathscr{G}(G, S)$.*

Proof. Let ϕ be the identity map on the vertex set: $\phi(x) = x$ for $x \in G$. We now define the value of ϕ on edges. Each edge e in $\mathscr{G}(G, S)$ can be translated by an element g to an edge adjacent to the identity. We first define the values of ϕ on the set $Z = \{(1, s) : s \in \tilde{S}\}$ of all edges adjacent to 1. Note that $|E| = 2|S|$.

For each $e = (1, s) \in E$, we choose a geodesic p_e in $\mathscr{G}(G, T)$ such that

$$(p_e)_- = 1, (p_e)_+ = s.$$

Then ϕ is defined to map e linearly to p_e.

For other edge $e' \in G \times \tilde{S}$, there exists $g \in G$ such that $ge = e'$ for $e \in Z$. We define $\phi(e') = g\phi(e)$.

Such defined map ϕ clearly satisfies $\phi(gx) = g\phi(x)$ for any $x \in \mathscr{G}(G, S)$. □

About growth functions, we see that quasi-isometric groups have equivalent growth functions. In other words, the growth function is a quasi-isometric invariant.

Exercise 6.16. *Let G, H be two finitely generated groups with finite generating sets S, T respectively. Assume that there exists a quasi-isometric map $\phi : (G, d_S) \to (H, d_T)$. Then the growth function ϕ_H of H dominates the growth function ϕ_G of G. In particular, if ϕ is a quasi-isometry, then $\phi_G \asymp \phi_H$.*

6.4. **Svarc-Milnor Lemma.** In this subsection, we prove a useful lemma, due to Svarc and independently Milnor, which gives lots of interesting quasi-isometries in practice. In some literatures, Svarc-Milnor Lemma is also referred to as the fundamental lemma of GGT. Before stating the lemma, we need prepare several notions.

Let G be a group acting by isometries on a metric space (X, d). That is, there is a homomorphism $G \to Isom(X)$.

Definition 6.17. The action of G on X is called a *proper* action if the following set

$$\{g : gK \cap K \neq \emptyset\}$$

is finite for any compact set K in X. We also say that G acts properly on X.

Remark. The action of G on X is proper if and only if the continuous map

$$G \times X \to X \times X, (g, x) \to (gx, x)$$

is proper in a topological sense. Here, G is equipped with the discrete topology.

Examples 6.18.
1. $\mathbb{Z}^n$ acts properly on $\mathbb{R}^n$.
2. Let M be a topological manifold. Then its fundamental group acts properly on its universal space.
3. Let G act properly on X. Then any subgroup acts properly on X.

Recall that in a proper metric space, any closed ball is compact. Then G acts properly on a proper metric space (X, d) iff the set $\{g : d(x, gx) \leq n\}$ is finite for all $n \geq 0, x \in X$.

Lemma 6.19. *Suppose G acts properly on a proper metric space X. Fix $x \in X$.*

1. *The point stabilizer $G_x = \{g : gx = x\}$ is finite.*

(2) *The orbit Gx with induced subspace topology is a discrete space.*

Remark. The converse is also true, see [11, Theorem 5.3.4] and other characterization of proper actions there.

If $G_x = \{1\}$, then we say G acts *freely* on X.

Proof. It is clear that there exists $r > 0$ such that $B(x, r) \cap B(gx, r) = \emptyset$ for any $g \in G$. Thus, Gx is a discrete space. $\square$

Definition 6.20. The action of G on X is called *co-compact* if there exists a compact subset K in X such that $X = G \cdot K$.

Given two group actions of G on metric spaces X, Y, a map $\phi : X \to Y$ between X and Y is called *G-equivariant* if the following holds

$$\phi(g \cdot x) = g \cdot \phi(x), \forall g \in G, x \in X.$$

Lemma 6.21 (Svarc-Milnor Lemma)**.** *Suppose G acts properly and co-compactly on a proper length space (X, d). Then*

(1) *G is finitely generated by a set S.*
(2) *Fix a basepoint $o \in X$. Then the map*
$$(G, d_S) \to (Go, d),\ g \to go,$$
is a G-equivariant quasi-isometric map.

The same proof of Svarc-Milnor Lemma can prove the following result. Note that the assumption of the proper action is only required to obtain a finite generating set. We call an action of G on a metric space X is *co-bounded* if there exists a bounded set K such that $G \cdot K = X$.

Exercise 6.22. *Suppose G acts by co-boundedly on a proper length space (X, d). Fix a basepoint $o \in X$. Then there exists a (possibly infinite) generating set S of G such that the map*
$$(G, d_S) \to (Go, d),\ g \to go,$$
is a G-equivariant quasi-isometric map.

Corollary 6.23. *Let G be a finitely generated group. Then any finite index subgroup is finitely generated and quasi-isometric to G.*

By the above corollary, G is quasi-isometric to $G \times F$, where F is finite. Here is another way to get quasi-isometric groups.

Exercise 6.24. *Let*
$$1 \to N \to G \to \Gamma \to 1$$
be a group extension, where G is finitely generated. Assume that N is finite. Then G is quasi-isometric to Γ.

Examples 6.25. (1) If $n \neq m$, then $\mathbb{R}^n$ is not quasi-isometric to $\mathbb{R}^m$.
(2) All free groups of finite rank at least two are quasi-isometric.

Exercise 6.26. *Let $n \geq 3$ be an integer. Prove that any two trees with vertices of degree between 3 and n are quasi-isometric.*

6.5. **Quasi-isometry classification of groups.** Gromov's Theorem 5.17 implies that virtually nilpotent groups are preserved by quasi-isometries. In the class of abelian groups, we have the following result.

Theorem 6.27. *Let G be a group quasi-isometric to an abelian group of finite rank. Then G is virtually abelian.*

Proof. The case of rank two can be proven by Bass-Guivarc'h formula (see the remark after Theorem 5.16) and Gromov's Theorem 5.17. The general case is much harder, and out of our scope. □

Let P be a group property. It is called *geometric* if it is preserved by quasi-isometries. That is to say, any group quasi-isometric to a group with property P also has the property P.

For example, the property of being virtual nilpotent, virtual abelian is geometric. In the next section, we will see more quasi-isometric invariant properties of groups.

7. Hyperbolic spaces

Let (X, d) be a proper length space. By Lemma 6.6, it is a geodesic space. For any two points $a, b \in X$, there is at least one geodesic between them. We often denote by $[a, b]$ a choice of a geodesic between a, b, if there is no ambiguity in context.

Let p be a path in (X, d). For any two points a, b on p, we denote by $[a, b]_p$ the subpath of p between a, b.

7.1. **Thin-triangle property.**

In (X, d), a geodesic triangle $\Delta = \Delta(abc)$ consists of three geodesics $[a, b], [b, c], [c, a]$. For $\delta \geq 0$, a point $o \in X$ is called a *δ-center* of Δ, if $d(o, [a, b]) \leq \delta, d(o, [b, c]) \leq \delta, d(o, [a, c]) \leq \delta$

Definition 7.1. We say that (X, d) is a *δ-hyperbolic space* for some $\delta \geq 0$ if every geodesic triangle has a δ-center.

Examples 7.2.
(1) Bounded metric spaces are δ-hyperbolic, where δ can be the diameter.
(2) Trees are 0-δ-hyperbolic.
(3) Real hyperbolic spaces are δ-hyperbolic for $\delta = (\log 3)/2$.

Definition 7.3. Let $a, b, c \in X$ be three points. The *Gromov product* of a, c with respect to b is defined as follows:

$$(a, c)_b = (d(a, b) + d(c, b) - d(a, c))/2.$$

It is straightforward to verify the following:

$$\begin{aligned} d(a, b) &= (a, c)_b + (b, c)_a \\ d(a, c) &= (b, c)_a + (a, b)_c \\ d(b, c) &= (a, b)_c + (a, c)_b \end{aligned}$$

In a hyperbolic space, the Gromov product obtains the following geometric meaning, which roughly measures the distance of the basepoint b to any geodesic between a, c.

Lemma 7.4. *Let a, b, c be three points in a δ-hyperbolic space. Then the following holds*

$$d(b, [a, c]) - 4\delta \leq (a, c)_b \leq d(b, [a, c])$$

where $[a, b]$ denotes some geodesic with endpoints a, b.

Proof. See Bowditch's book [2], Lemma 6.1 and Lemma 6.2. □

Exercise 7.5. *Let $\Delta = \Delta(abc)$ be a geodesic triangle and o be a k-center for $k > 0$. Let $x \in [a, c]$ such that $d(x, c) = (a, b)_c$. Then $d(x, o) \leq 6k$.*

A direct consequence is the stability of geodesics: any two geodesics between same endpoints are uniformly close.

Corollary 7.6. *Let p, q be two geodesics in a δ-hyperbolic space with same endpoints. Then $p \subset N_{4\delta}(q), q \subset N_{4\delta}(p)$.*

Lemma 7.4 allows us to obtain the stability of a slight general path called taut path. A path p is called *t-taut* for some $t \geq 0$ if $\text{Len}(p) \leq d(p_-, p_+) + t$.

Lemma 7.7. *Let p be a t-taut path for $t \geq 0$, and q be a geodesic with same endpoints as p. Then there exists $D = D(t, \delta)$ such that*

$$p \subset N_D(q), q \subset N_D(p).$$

Proof. See Bowditch's book [2], Lemma 6.4. □

Definition 7.8 (Thin-triangle property)**.** A geodesic triangle has δ-*thin* property for some $\delta \geq 0$ if each side lies in the δ-neighbourhood of the other two sides.

Note that any δ-thin triangle has a δ-center. In fact, a hyperbolic space can also be characterized by all geodesic triangles being δ-thin.

Lemma 7.9. *In a δ-hyperbolic space, any geodesic triangle has the 6δ-thin property.*

Proof. See Bowditch's book [2], Lemma 6.5. □

7.2. **Stability of quasi-geodesics.** We first derive the exponential growth of a path outside a ball, which is a consequence of the thin triangle property.

Lemma 7.10 (Exponential divergence)**.** *Let p be a rectifiable path between a, b, and $[a, b]$ denote some geodesic between a, b. Then we have*

$$[a, b] \subset N_D(p)$$

where $D = 6\delta(\log_2 Len(p) + 1) + 2$.

Proof. Let c_0 be the middle point of p dividing p in half. We then divide in half the subpaths $[a, c_0]_p$, and $[c_0, b]_p$ and denote by c_1, c_2 the corresponding middle points. We continue to divide until all subpaths are just of length less than 1. So there are at most $\log_2 \mathrm{Len}(p) + 1$ divisions.

Let $x \in [a, b]$. By Lemma 7.9, we only need at most $log_2 \mathrm{Len}(p) + 1$ "6δ-jumps" from x to a point on a geodesic, which has same endpoints as a subpath in p of length at most 1. Thus, we obtain $d(x, p) \leq 6\delta(\log_2 \mathrm{Len}(p) + 1) + 2$. □

Corollary 7.11. *There exist constants $C_1, C_2 > 0$ with the following property.*

Let $[a, b]$ be a geodesic between a and b, and c be a point in $[a, b]$. Denote $r = \min\{d(c, a), d(c, b)\}$. Then any rectifiable path p with $p_- = a, p_+ = b$ outside the open ball $B(c, r)$ has length bigger than $C_1 e^{C_2 r}$.

Proof. Let $D = 6\delta(\log_2 \mathrm{Len}(p)+1)+2$ is given by Lemma 7.10. Then $[a, b] \subset N_D(p)$. By assumption, $r < D$. Thus, the conclusion follows. □

Definition 7.12. A path p in (X, d) is called a (λ, c)-*quasi-geodesic* for $\lambda \geq 1, c \geq 0$ if the following holds

$$\mathrm{Len}(q) \leq \lambda d(q_-, q_+) + c$$

for any (connected) subpath q of p.

Remark. The length parametrization of p gives a (λ, c)-quasi-isometric embedding of the interval $[0, \mathrm{Len}(p)]$ in X.

Lemma 7.13 (Stability of quasi-geodesics)**.** *For any $\lambda \geq 1, c \geq 0$, there exists $D = D(\delta, \lambda, c) > 0$ with the following property. Let p, q be two (λ, c)-quasi-geodesics in a δ-hyperbolic space (X, d). Then $p \subset N_D(q)$.*

Proof. Without loss generality, assume that q is a geodesic. Let $x \in q$ such that $d(x,p)$ is maximal. Denote $R = d(x,q)$. It suffices to give a bound on R.

Since $d(x,p)$ is maximal, it follows that $p \cap B(x,R) = \emptyset$. If $d(q_-,x) \geq 2R$, choose $y_1 \in [q_-,x]_q$ such that $d(y_1,x) = 2R$. Otherwise, let $y_1 = q_+$(which we call as a degenerated case). Similarly, choose $y_2 \in [x,q_+]_q$ with the same property.

As $d(y_1,p) \leq d(x,p)$, there exists a point $z_1 \in p$ such that $d(y_1,z_1) \leq R$. In the degenerated case, we just take $z_1 = y_1 = q_-$. Similarly, there exists $z_2 \in p$ such that $d(y_2,z_2) \leq R$.

We now compose a path connecting y_1 and y_2: $\bar{p} = [y_1,z_1][z_1,z_2]_p[z_2,y_2]$. Note that $[y_i,z_i]$ intersects trivially $B(x,R)$. Indeed, if not, we arrive a contradiction with $d(y_i,x) = 2R$. In the degenerated case, we have that $[y_i,z_i]$ are trivial paths. Thus, we see that $\bar{p} \cap B(x,R) = \emptyset$. Then by Lemma 7.10, $\mathrm{Len}(\bar{p}) \geq C_1 e^{C_2 R}$ for some constants $C_i > 0$.

On the other hand, $\mathrm{Len}(\bar{p}) \leq 2R + \mathrm{Len}([z_1,z_2]_p) \leq 2R + 6R\lambda + c$. As C_i, λ, c are fixed, we see that R has to be bounded by an uniform constant $D = D(\delta,\lambda,c)$. The proof is complete. □

Exercise 7.14. *Let p be a path in a hyperbolic space. Given a non-discreasing function $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, let p be a path such that Len$(q) \leq f(d(q_-,q_+))$ for any subpath q of p. Assume that f is sub-exponential (i.e.: $\lim_{n\to\infty} \log f(n)/n = 0$). Then p is a quasi-geodesic path.*

7.3. **Boundedness of δ-centers.** A *projection* of a point a to a path p is a point z in p such that $d(a,z) = d(a,p)$. Note that a projection point is usually not unique, but different projection points are uniformly close. This is a useful result we shall prove now.

Lemma 7.15 (Bounded centers)**.** *Let X be a δ-hyperbolic space. For any $k > 0$, there exists a constant $D = D(\delta,k) > 0$ with the following property. Let Δ be a geodesic triangle with vertices a,b,c, and z be a projection of c to the opposite side. Then $d(o,z) \leq D$ for any k-center o.*

Proof. We first show the following.

Claim. *such a projection point z is a D_0-center of Δ for some uniform constant $D_0 > 0$.*

Proof. Connect c to z, and then z to b by geodesics. We have a path $p = [c,z][z,a]$. It is a general result without the hyperbolic assumption that p is a $(3,0)$ quasi-geodesic. Indeed, it suffices to consider a subpath q with $q_- \in [c,z], q_+ \in [z,a]$. As $d(c,[a,z]) = d(a,z)$, we have that $d(q_-,q_+) \geq d(q_-,z)$. Thus, $d(q_+,z) \leq 2d(q_-,q_+)$. It follows that $\mathrm{Len}(q) = d(q_+,z) + d(q_-,z) \leq 3d(q_-,q_+)$.

By Lemma 7.13, there exists a constant $D_0 = D(\delta,3,0) > 0$ such that $d(z,[a,c]) \leq D$ and $d(z,[b,c]) \leq D$. That is, z is a D_0-center of Δ. □

Let o be a k-center of Δ. Let $x \in [a,c]$ such that $d(x,c) = (a,b)_c$. By Exercise 7.5, we have that $d(x,o) \leq 6k$. As z is a D_0-center of Δ, we have $d(x,z) \leq 6D_0$. Hence, $d(o,z) \leq 6(k + D_0)$. □

Corollary 7.16. *Let p be a geodesic and $x \in X$ a point. Then there exists an uniform constant $D > 0$ such that*

$$\mathit{Diam}(\{z \in p : d(x,p) = d(x,z)\}) \leq D.$$

7.4. **Comparison of triangles.** We consider a refined thin-triangle property, which is based on the idea of comparing triangles in various model spaces.

Let Δ be a geodesic triangle with sides p, q, r in a proper length space (X, d_X). Two points x, y in Δ are called *congruent* if they have the same distance to the common endpoint of the two sides where lie x, y respectively.

Assume that (Y, d_Y) is another proper length space such that there exists a geodesic triangle Δ' with sides p', q', r' such that $\text{Len}(r) = \text{Len}(r'), \text{Len}(p) = \text{Len}(p'), \text{Len}(q) = \text{Len}(q')$. Such a geodesic triangle will be referred to as a *companion triangle* of Δ.

Many proper length spaces have this property that every numbers a, b, c with $a \le b + c$ can be realized by lengths of a geodesic triangle, such as Euclidean spaces, classical hyperbolic spaces and trees.

There is a natural bijective map $\phi : \Delta \to \Delta'$ which sends sides of Δ isometrically to those of Δ'. We say that Δ is *δ-thinner* than Δ' for some $\delta \ge 0$ if for any two congruent points $x', y' \in \Delta'$, we have $d_X(\phi^{-1}(x'), \phi^{-1}(y)) \le d_Y(x', y') + \delta$.

Thus, it is easy to prove the following characterization of δ-hyperbolicity.

Exercise 7.17. *Let (X, d) be a δ-hyperbolic space. Then there exists a constant $\delta' > 0$ such that every geodesic triangle is δ'-thinner than a companion geodesic triangle in a tree.*

A proper length space is called *CAT(0)* (resp. *CAT(-1)*) if every geodesic triangle is 0-thinner than its companion triangle in Euclidean (resp. classical hyperbolic) plane. They are important metric spaces in GGT. We refer the reader to [4] for a through discussion about these spaces.

7.5. **Hyperbolicity is a quasi-isometric invariant.** We give another definition of a quasi-geodesic. A *parameterized quasi-geodesic* is a quasi-isometric embedding map of a finite or infinite interval of $\mathbb{R}$ in (X, d).

The following lemma implies that a parameterized quasi-geodesic can be converted to a continuous quasi-geodeisc without essensal lose.

Lemma 7.18. *Let $\phi : I \to X$ be a (λ, c)-quasi-isometric embedding map, where I is a finite or infinite interval in $\mathbb{R}$. Then there exist a (λ', c')-quasi-geodesic p and a constant $D > 0$ such that the following holds*

$$\phi(I) \subset N_D(p), p \subset N_D(\phi(I)).$$

Proof. Consider the set $\phi(\mathbb{Z} \cap I)$ of points in X. We connect consecutive points by geodesic segments and obtain a path p. Then it is straightforward to verify that p is a quasi-geodesic and satisfies the conclusion. $\square$

Then we obtain that hyperbolicity is a quasi-isometric invariant.

Theorem 7.19. *Let X, Y be two proper length spaces. Assume that $\phi : X \to Y$ is a quasi-isometry. If X is hyperbolic, then Y is also hyperbolic.*

So we can see that for $n \ge 2$, $\mathbb{R}^n$ cannot be quasi-isometric to a tree or any other hyperbolic space.

7.6. **Approximation trees in hyperbolic spaces.** In this section, we prove a very useful result, which gives a tree-like picture of any finite set in a hyerbolic space.

Lemma 7.20. *Let X be a hyperbolic space, and F be a finite set. There exists a constant $c = c(|F|)$ and an embedded tree $T \subset X$ with $F \subset T^0$ such that the following holds*

$$d_T(x, y) \leq d_X(x, y) + c.$$

In other words, there exists an injective $(1, c)$-quasi-isometric map $\iota : T \to X$ with $F \subset \iota(T^0)$.

Proof. See Lemma 6.7 in Bowditch's book [2]. The key step is the following.

Claim. *For $t > 0$, let p be a t-taut path and project a point $o \in X$ to $x \in p$. Then $q = [o, x][x, p_-]_p$ is a t'-taut path for some t' depending only on t.*

Proof. We also project o to $y \in [p_-, p_+]$. By Lemma 7.13 and δ-thin property, we can show that $d(x, y) \leq D$ for some constant $D = D(t)$.

Then $\mathrm{Len}(q) = d(o, x) + \mathrm{Len}([x, p_-]_p) \leq d(o, x) + d(p_-, x) + t \leq t + 2D + d(o, y) + d(y, p_-)$.

By the Claim of Lemma 7.15, y is δ-center of the triangle $\Delta(o, y, p_-)$. Thus $d(o, y) + d(y, p_-) \leq 2\delta + d(o, p_-)$. Then we obtain that $\mathrm{Len}(q) \leq d(o, p_-) + t + 2D + 2\delta$. This completes the proof. □

The embedded tree T is constructed as follows. List $F = \{x_0, x_1, ..., x_{n-1}\}$. Let T_1 be a geodesic segment between x_0, x_1. Inductively, the tree T_i is defined by adding a new branch, which is obtained by connecting x_i and a projection point of x_i to T_{i-1}. In each step, every path in T_i is a t-taut path for some $t = t(i)$ given by the Claim. Thus, $T = T_n$ is the tree we wanted. □

The proof of the above Claim generalizes in the following exercise.

Exercise 7.21. *Let p be a (λ, c)-quasi-geodesic in a geodesic metric space X. For a point $x \in X$, let z be a projection point of x to p. Then $[x, z][z, p_+]_p$ is a (λ', c')-quasi-geodesic for $\lambda' \geq 1, c' \geq 0$ depending on λ, c.*

We now come to the original definition of a hyperbolic space due to Gromov.

Lemma 7.22. *Suppose (X, d) is a hyperbolic space. Then there exists $\delta > 0$ such that the following holds*

$$d(x, y) + d(z, w) \leq \max\{d(x, z) + d(y, w), d(x, w) + d(y, z)\} + \delta \tag{15}$$

for any four points $x, y, z, w \in X$.

Equivalently, (15) is amount to saying that

$$(x, y)_w \geq \min\{(x, z)_w, (y, z)_w\} - \delta/2. \tag{16}$$

Proof. The proof is clear by considering the approximation tree of four points. □

Exercise 7.23. *Try to prove the converse of Lemma 7.22: Suppose (X, d) satisfies the (15), then (X, d) is δ-hyperbolic.*

8. Boundary of hyperbolic spaces

In this section, we introduce a visual boundary of a proper hyperbolic space which as a set consists of all (equivalent) geodesic rays. We shall put a compact topology and also a visual metric on this boundary.

8.1. Asymptotic classes of geodesic rays.

Definition 8.1. Let $p, q : [0, \infty) \to X$ be two geodesic rays. We say that p, q are *asymptotic* if there exists $D > 0$ such that $p \subset N_D(q)$.

Clearly, the asymptotic relation is an equivalent relation over all geodesic rays.

In what follows, we use the following estimates:

(1) δ-thin triangle property: any side of a triangle is contained in the δ-neighborhood of the union of the other two sides.
(2) Let p, q are two geodesic segments originating from the same point $p_- = q_- = o$. If $x \in p$ and $y \in q$ have the same distance to o and $d(x, p_+) > d(p_+, q_+)$ then $d(x, y) \le 2\delta$.

Lemma 8.2 (Uniformity of asymptotic rays)**.** *Let p, q be two asymptotic geodesic rays. Then there exist $t_0, s_0 > 0$ such that*

$$p([t_0, \infty)) \subset N_{4\delta}(q([s_0, \infty)))$$

and

$$q([s_0, \infty)) \subset N_{4\delta}(p([t_0, \infty)))$$

If p, q has the same initial point $p_- = q_-$, then we can choose $s_0 = t_0 = 0$: $p \subset N_{4\delta}(q)$ and $q \subset N_{4\delta}(p)$.

Proof. Since p, q are asymptotic, we have $p \subset N_D(q)$ for some $D > 0$. Set

$$t_0 = d(p_-, q_-) + D + 4\delta.$$

First of all, we observe that for any $t \ge t_0$, $p(t) \in N_{2\delta}(q)$. Since $p \subset N_D(q)$, there exists $s \in [0, \infty)$ such that $d(p(2t), q(s)) \le D$. We consider the quadrangle

$$[p_-, q_-] \cdot [q_-, q(s)]_q \cdot [q(s), p(2t)] \cdot [p(2t), p_-]_p.$$

Since a quadrangle has 2δ-thin property, we have

$$p(t) \in N_{2\delta}([p_-, q_-] \cdot [q_-, q(s)]_q \cdot [q(s), p(2t)]).$$

As $d(p(t), p_-), d(p(t), p(2t)) \ge t \ge t_1$, we see that $p(t) \in N_{2\delta}([q_-, q(s)]_q)$. The observation follows.

By the observation, we can choose a sequence of increasing numbers $t_0 < t_1 < \cdots t_n \to \infty$ and $s_0 < s_1 < \cdots s_n \to \infty$ such that

$$d(p(t_n), q(s_n)) \le 2\delta$$

for any $n \ge 0$.

Using again 2δ-thin quadrangles, we conclude that

$$[p(t_n), p(t_{n+1})] \subset N_{4\delta}([q(s_n), q(s_{n+1})])$$

thus $p([t_0, \infty)) \subset N_{2\delta}(q([s_0, \infty)))$ is proved. The other inclusion is analogous. □

Fix a basepoint $o \in X$, let $\partial_o X$ be the set of all equivalent classes of asymptotic geodesic rays issuing at o. We can also think of every point $x \in X$ as the set of all geodesic segments between o and x. Denote $\bar{X}_o = X \cup \partial_o X$.

8.2. **Topological boundary: Gromov boundary.** We shall describe two euivalent topologies on $\bar{X}_o$ such that $\bar{X}_o$ is a compactification of X: X is an open dense subset in $\bar{X}_o$.

1. The first topology defined by neighbourhood base. For every point x in $\bar{X}_o$, we assign a countable neighbourhood base $\mathbb{V}_x$.

For $x \in X$, we let $\mathbb{V}_x$ be the neighbourhood base under the original topology of X. For example, $\mathbb{V}_x = \{B(x, n^{-1}) : n \in \mathbb{N}\}$.

Let $x \in \partial_o^\infty X$ be an equivalent class of geodesic rays. Fix a big integer $k > 12\delta$. Define

$$U_{x,n} = \{y \in \bar{X}_o : \exists c' \in y, \exists c \in x, d(c(kn), c'(kn)) < 4\delta\}$$

for every $n \in \mathbb{N}$.

Let $\mathbb{V}_x = \{U_{x,n} : n \in \mathbb{N}\}$. Given a neighborhood base $\mathbb{V}_x$, we denote by $\tilde{\mathbb{V}}_x$ the generated neighborhood system of x, which consists of all sets of X containing an element in $\mathbb{V}_x$.

We are verifying the following properties of the neighborhood system $\tilde{\mathbb{V}}_x$.

Lemma 8.3. (1) *Let $U, V \in \tilde{\mathbb{V}}_x$. Then $U \cap V \in \tilde{\mathbb{V}}_x$.*
(2) *Let $U \in \tilde{\mathbb{V}}_x$. Then there exists $V \in \tilde{\mathbb{V}}_x$ such that $V \subset U$ and $U \in \tilde{\mathbb{V}}_y$ for any $y \in V$.*

Proof. Note first that $U_{x,n+1} \subset U_{x,n}$. Indeed, if $y \in U_{x,n+1}$ there exists $c' \in y$ and $c \in x$ such that

$$d(c(kn+n), c'(kn+n)) \le 4\delta.$$

Consider the geodesic triangle with vertices $o, c(kn+n), c'(kn+n)$. Since $d(c(kn+n), c(kn)) = k \ge 8\delta$, we have $d(c(kn), c'(kn)) \le 2\delta$ by δ-thin triangle. Thus, $y \in U_{x,n}$. The first statement is then obvious.

Assume that $U = U_{x,n}$. We shall prove that $V = U_{x,n+1}$ is the desired set. If $y \in V$, then there exist $c \in x$ and $\gamma \in y$ such that

$$d(c(kn+k), \gamma(kn+k)) \le 4\delta.$$

If $z \in U_{y,n+1}$, then there exist $\beta \in z$ and $\gamma' \in y$ such that

$$d(\gamma'(kn+k), \beta(kn+k)) < 4\delta.$$

By Lemma 8.2, we have $d(\gamma'(kn+k), \gamma(kn+k)) < 4\delta$ for $[\gamma] = [\gamma']$, so

$$d(\gamma(kn+k), \beta(kn+k)) < 8\delta.$$

Then $d(c(kn+k), \beta(kn+k)) \le 12\delta$. We now apply the thin-triangle property to triangle $(o, c(kn+k), \beta(kn+k))$. Since $k \ge 12\delta \ge d(\gamma(kn+k), \beta(kn+k))$, we obtain $d(c(kn), \beta(kn)) \le 2\delta$. So $U_{y,n+1} \subset U_{x,n}$. Hence, $U_{x,n} \in \tilde{\mathbb{V}}_y$. □

We define a set S in $\bar{X}_o$ to be *open* if $S \in \tilde{\mathbb{V}}_x$ for every $x \in S$. The first statement of Lemma 8.3 implies that such a topology is well-defined. The second statement implies that

Corollary 8.4. *$\mathbb{V}_x$ is a neighborhood base of the topology we constructed.*

Proof. It suffices to prove that every $U \in \mathbb{V}_x$ contains an open subset $x \in Z$. Let Z be the set of $z \in X$ such that $U \in \tilde{\mathbb{V}}_z$. Clearly, $x \in Z$, and in fact, $Z \subset U$. We now prove that Z is open: Z is a neighborhood of every point in it.

Let $z \in Z$. As $U \in \tilde{\mathbb{V}}_z$, by Lemma 8.3, there exists $V \in \tilde{\mathbb{V}}_z$ such that $V \subset U$ and $U \in \tilde{\mathbb{V}}_y$ for any $y \in V$. Thus, $V \subset Z$ by definition of Z. Since $V \in \tilde{\mathbb{V}}_z$, we have $Z \in \tilde{\mathbb{V}}_z$. Hence, Z is an open set. □

2. The second topology defined by sequence convergence. Alternatively, we can define the topology on $\bar{X}_o$ by the following convergence of sequences:

$x_n \to x \in \bar{X}_o$ iff there exists a sequence of geodesic rays (or segments) c_n in (the equivalent class of) x_n originating from o so that c_n converges locally uniformly to a geodesic ray or segment c in (the equivalent class of) x.

Thus, a set S is *closed* iff S contains the limit point of every convergence sequence in S.

Lemma 8.5. *The two topologies on $\bar{X}_o$ coincide.*

Proof. Assume that $\bar{X}_o$ is equipped with the first topology generated by neighborhood systems $\tilde{\mathbb{V}}_x$. We have to verify that $x_n \to x \in \bar{X}_o$ in the first topology is equivalent to the following statement:

there exists a sequence of geodesic rays (or segments) c_n in (the equivalent class of) x_n originating from o so that c_n converges locally uniformly to a geodesic ray or segment c in (the equivalent class of) x.

This equivalence follows easily from the definition of $U_{x,n}$. Thus, (open sets of) the first topology contains (open sets of) the second topology.

Conversely, assume that $\bar{X}_o$ is equipped with the topology generated by convergence $x_n \to x \in \bar{X}_o$. Note that the complement $U^c_{x,n}$ consists of $y \in \bar{X}_o$ such that for any $c' \in y$ and any $c \in x$, we have $d(c(kn), c'(kn)) \geq 4\delta$. Thus, any convergence sequence $y_n \in U^c_{x,n}$ converges locally uniformly to a geodesic ray c' so that $d(c(kn), c'(kn)) \geq 4\delta$ for any $c \in x$. As a consequence, $U^c_{x,n}$ is closed, and $U_{x,n}$ is open in this topology. Thus, (open sets of) the second topology contains (open sets of) the first topology. □

Lemma 8.6. *Let $\bar{X}_o$ be endowed with the first or second topology. Then $\bar{X}_o$ is a compactification of X.*

Proof. The topology on $\bar{X}_o$ is first countable by definition: every point has a countable neighbourhood base. Clearly, X is an open subset in $\bar{X}_o$, and the induced topology on X agrees with the original topology of X. It is also dense: for any point $x \in \partial_o X$, the sequence of points $c(n)$ converges to x.

We now prove that $\bar{X}_o$ is compact. It suffices to prove that $\partial_o X$ is compact. Indeed, let $\mathbb{U}$ be an open cover of $\bar{X}_o$. If $\partial_o X$ is proven to be compact, there exists a finite open sub-cover $\mathbb{V}$ of $\partial_o X$. Note that $X \setminus \cup\mathbb{V}$ is a bounded set. Indeed, if not, there exists a sequence of points $x_i \in X \setminus \cup\mathbb{V}$ such that $d(o, x_i) \to \infty$. Consider the geodesics $[o, x_i]$. By Arzela-Ascoli Lemma 6.5, there exists a subsequence still denoted by x_i, which converges to a geodesic ray c with $c(0) = o$. Thus, $x_i \to c \in \partial_o X$. This gives a contradiction, as x_i lies in the closed set $X \setminus \cup\mathbb{V}$.

As X is a proper space, $X \setminus \cup\mathbb{V}$ is a compact set. Thus, we see that there exists finite a sub-cover of $\bar{X}_o$. Thus, we proved that $\bar{X}_o$ is compact, provided that $\partial_o X$ is compact. We show below that $\partial_o X$ is compact.

Recall that for a first countable metrizable space, compactness is equivalent to sequentially compactness.

By Arzela-Ascoli Lemma 6.5, it is easy to see that $\partial_o X$ is sequentially compact. Since $\partial_o X$ is metrizable (a result proven in Lemma 8.15 below), we see that $\partial_o X$ is compact. □

Let o, o' be two basepoints in X. Define a bijection ι between $\bar{X}_o$ and $\bar{X}_{o'}$ as follows. Let ι be the identification on X. For any $x \in \partial_o X$, let $c : [0, \infty) \to X$

be a geodesic ray in X. Consider the sequence of geodesics $p_n = [o', c(n)]$. By Arzela-Ascoli Lemma 6.5, there exists a subsequence of p_n which converges to a geodesic ray c' with $c'(0) = o'$. Define $\iota(x) = [c'] \in \partial_{o'} X$.

Exercise 8.7. *Verify that $\iota : \bar{X}_o \to \bar{X}_{o'}$ is well-defined and bijective, and prove that ι is a homeomorphism.*

Therefore, the topology on $\bar{X}_o$ does not depend on the choice of the basepoint $o \in X$. In the sequel, we will omit the index o for simplicity. We call ∂X the *Gromov boundary* of the hyperbolic space.

Let $x \in X, y \in \partial X$. If a geodesic p with $p_- = x$ is asymptotic to a geodesic ray in y, we say that p *ends at* y and write $p_+ = y$. A geodesic with endpoints $x, y \in \partial X$ is defined similarly.

The following property is usually referred as visual property of boundary. It is of a particular feature in a hyperbolic space.

Lemma 8.8 (Visibility of boundary)**.** *In a hyperbolic space X, there exists a geodesic between any two distinct $x, y \in \bar{X}$.*

Proof. If both x, y are in X, the conclusion follows as X is a proper length space.

Consider $x \in X, y \in \partial X$. There exists a geodesic p in y. We connect x and $p(n)$ by a geodesic q_n. By Arzela-Ascoli Lemma 6.5, there exists a subsequence of q_n converging to a geodesic ray q_∞. By thin triangle property, we see that q_∞ is equivalent to p.

Now assume that $x, y \in \partial X$. Fix a basepoint o. By the above argument, there exist geodesic rays p, q with $p_- = q_- = o$ and $p_+ = x, q_+ = y$. Connect $p(n)$ and $q(n)$ by a geodesic segment r_n.

We now prove the following claim.

Claim. *There exists a constant $D = D(x, y) > 0$ such that*

$$d(o, r_n) \leq D$$

for all $n > 0$.

Proof. Suppose to the contrary that there exists $n_i \to \infty$ such that $d(o, r_{n_i}) > i$. By the thin-triangle property, we obtain that $d(p(i), q(i)) \leq \delta$ for all $i \geq 0$. This gives a contradiction, as $x \neq y$. □

By Arzela-Ascoli Lemma 6.5 and the Claim, we see that a subsequence of r_n converges to r_∞ with $d(o, r_\infty) \leq D$. It is also clear that each half-ray of r_∞ is asymptotic to either x or y. □

We can now consider a geodesic triangle with vertices at boundary. It is easy to verify the following.

Exercise 8.9. *Let X be a hyperbolic space. There exists $\delta > 0$ such that any geodesic triangle with vertices $x, y, z \in \bar{X}$ has δ-thin triangle property.*

8.3. **Gromov boundary of any hyperbolic metric spaces.** Using sequences, we are able to define Gromov boundary for (possibly non-geodesic, non-proper) hyperbolic metric space.

Let X be a hyperbolic metric space defined by using the four-points condition.

Definition 8.10. A sequence (x_n) in X *converges at infinity* if $(x_i, x_j)_o \to \infty$ as $i, j \to \infty$. Two such sequences $(x_n), (y_n)$ are called *equiavlent* if $(x_i, y_j)_o \to \infty$ as $i, j \to \infty$. The *Gromov boundary* $\partial_s X$ of X is the set of all equiavlent classes of sequences converging at infinity.

See also [4, Chapiter III.H.3] for a discussion about the details.

Exercise 8.11. [4, Ch. III, Lemma 3.13] *If X is a proper geodesic hyperbolic space, there exists a natural bijection from $\partial_s X$ to ∂X.*

8.4. **Visual metric on the boundary.** In this subsection, we shall put a metric on $\partial X = \partial_o X$ to metrize the topology introduced in the previous subsection.

For two distinct $x, y \in \partial X$, we define $d(o, [x, y]) = d(o, A)$, where A is the union of all geodesics between x, y. The following exercise certifies this definition.

Exercise 8.12. *Let A be the union of all geodesics between x, y. Then $A \subset N_{10\delta}(p)$ for any geodesic $p = [x, y]$.*

So the inequality (16) in Lemma 7.22 proves the following.

Lemma 8.13. *There exists $\delta > 0$ such that the following holds*

$$d(o, [x, y]) \ge \min\{d(o, [x, z]), d(o, [y, z])\} - \delta \tag{17}$$

for any distinct $x, y, z \in \bar{X}$ and $o \in X$.

We first define a "quasi-metric" on $\bar{X} = X \cup \partial_o X$ as follows. Fix $a \in (0, 1]$ and a basepoint $o \in X$. For any two distinct $x, y \in \bar{X}$, set

$$\bar{\rho}_a(x, y) = e^{-ad(o,[x,y])}$$

for $\lambda \in (0, 1)$. Then by Lemma 8.13, we have

$$\bar{\rho}_a(x, y) = \bar{\rho}_a(y, x) \tag{18}$$

$$\bar{\rho}_a(x, y) \le K \max\{\bar{\rho}_a(y, z), \bar{\rho}_a(x, z)\} \tag{19}$$

for $x, y, z \in \bar{X}$ and $K = e^{a\delta}$. If $x = y$, we define $\bar{\rho}_a(x, y) = 0$.

It is useful to cite the following general result which holds for any quasi-metric satisfying the condition (18, 19). The proof is elementary and elegant.

Lemma 8.14. *Let $\bar{\rho} : M \times M \to \mathbb{R}_{\ge 0}$ a function satisfying (18, 19) with $1 \le K \le \sqrt{2}$. Then there exists a metric ρ on M such that*

$$\frac{\bar{\rho}(x, y)}{K} \le \rho(x, y) \le \bar{\rho}(x, y)$$

for any $x, y \in M$.

Proof. Define

$$\rho(x, y) = \inf\{\sum_{i=1}^{n} \bar{\rho}(x_i, x_{i+1})\}$$

over all finite sequences $(x_0 = x, x_1, \cdots, x_{n+1} = y)$ in M. Thus, $\rho(x, y)$ is a metric on M satsifying triangle inequality etc. Note that $\bar{\rho}(x, y) \le \rho(x, y)$. It remains to prove $\bar{\rho}(x, y) \le K\rho(x, y)$. We shall prove that

$$\bar{\rho}(x, y) \le K \sum_{i=1}^{n} \bar{\rho}(x_i, x_{i+1})$$

for every finite sequences. We proceed by induction on the length of sequences. The case $n = 1$ follows from (19): $\bar{\rho}(x, y) \le K \max\{\bar{\rho}(y, z), \bar{\rho}(x, z)\}$.

Assume that the case n is verified:

$$\bar{\rho}(x, y) \le K \sum_{i=1}^{n} \bar{\rho}(x_i, x_{i+1}) \tag{20}$$

Now let us prove the case for sequences $(x_0 = x, x_1, \cdots, x_{n+1}, x_{n+2} = y)$ of length $n+1$. Let $R = \sum_{i=1}^{n+1} \bar{\rho}(x_i, x_{i+1})$. The goal is to prove $\bar{\rho}(x, y) \le KR$.

Let p be the maximal index such that

$$\sum_{i=1}^{p} \bar{\rho}(x_i, x_{i+1}) \le R/2$$

It is possible that $\bar{\rho}(x_p, x_{p+1}) \le R$, but we have

$$\sum_{i=p+1}^{n+1} \bar{\rho}(x_i, x_{i+1}) \le R/2$$

Now we have

$$\begin{aligned} \bar{\rho}(x, y) &\le K \max\{\bar{\rho}(x, x_p), \bar{\rho}(x_p, y)\} \\ &\le K \max\{\bar{\rho}(x, x_p), K\bar{\rho}(x_p, x_{p+1}), K\bar{\rho}(x_{p+1}, y)\} \\ &\le K \max\{KR/2, R, K^2R/2\} \\ &\le KR \end{aligned}$$

where the first two inequalities follow from (19), the third one uses indution (20) on n to $\bar{\rho}(x, x_p)$ and $\bar{\rho}(x_p, y)$, and the last one follows by $K^2 \le 2$. The proof is complete. □

Thus, we choose $a \in (0, 1]$ small enough such that $e^{a\delta} \le \sqrt{2}$ (there is a critical value a_0 such that any $a \in (0, a_0]$ works). Then we get a metric ρ_a on $\bar{X}$ by Lemma 8.14 such that

$$\bar{\rho}_a(x, y)/2 \le \rho_a(x, y) \le \bar{\rho}_a(x, y) \tag{21}$$

Lemma 8.15. *The induced topology on ∂X by ρ is the same as the topology defined in previous subsection.*

Proof. It can be verified by showing that every $U \in \mathbb{V}_x$ contains an open ball B at x, and each open ball B at x contains some $U \in \mathbb{V}_x$. □

Remark. The induced topology on X from the metric ρ is discrete, so it may not be same as the original one on X. However, we can choose a maximal net of X, for instance, the vertex set if X is a Cayley graph. Then this construction of visual metric works for the restricted metric on the net. In this situation, the metric topology on the net coincides with discrete topology.

8.5. **Boundary maps induced by quasi-isometries on boundaries.** We shall first prove that a quasi-isometry induces a homeomorphism called boundary map between Gromov boundaries of two hyperbolic spaces. With respect to visual metric, the boundary map satisfy very good properties called quasi-symmetric map.

We start with a few useful facts about how Gromov products transform under quasi-isometries.

Let X be a metric complete space and A be a closed subset. Let $\pi_A : X \to A$ be the projection map sending a point to the closest point on A:

$$d(\pi_A(x), A) = d(x, A).$$

If γ is a geodesic (more generally any quasi-convex subset), then the closest projection map π_γ commutes with quasi-isometry up to finite additive error:

Exercise 8.16. *Let $\phi : X \to Y$ be a (λ, c)-quasi-isometry between two proper geodesic δ-hyperbolic spaces X, Y. Let γ be a geodesic. Prove that there exists a constant $D = D(\lambda, c, \delta)$ such that for any point $x \in X$,*

$$d_H(\phi(\pi_\gamma(x)), \pi_{\phi\gamma}(\phi(x))) \le D$$

where d_H denotes the Hausdorff distance.

Let F be a finite set of points in a δ-hyperbolic space X. An embedded tree $T \subset X$ is called *approximation tree* of F if $F \subset T^0$ are leaves of T (vertices of valence 1) and

$$d_T(x, y) \le d_X(x, y) + c$$

where c depends only on $|F|, \delta$. The approximation tree exists by Lemma 7.20.

The following result says that quasi-isometries preserves the structure of approximation trees.

Proposition 8.17. *Let $\phi : X \to Y$ be a (λ, c)-quasi-isometry between two proper geodesic δ-hyperbolic spaces X, Y. Let F be a finite set in X. There exist a constant D depending on $\lambda, c, \delta, |F|$ and two combinatorially isomorphic approximation trees $T_1 \subset X$ for F and $T_2 \subset Y$ with $\phi F \subset Y$ such that ϕ coarsely preserves their combinatorial structures of T_1, T_2:*

1. *ϕ sends the leaves F of T_1 to leaves ϕF of T_2*
2. *ϕ sends vertices T_1^0 into the D-neighborhood of the corresponding vertices T_2^0.*

Proof. We follow the proof of Lemma 7.20 to construct approximation trees T_1, T_2 for F and $\phi(F)$.

Recall that approximation tree for $F = \{x_0, x_1, \cdots, x_n\}$ is obtained as follows: starting with a geodesic segment between two points x_0, x_1 in F, drop inductively an orthogonal from x_i to the existing subtree T_i. We do the construction for $\phi F = \{\phi x_0, \phi x_1, \cdots, \phi x_n\}$ simutaneously by dropping orthogonals from ϕx_i to the existing subtree T_i'.

By Exercise 8.16, the projections from x_i to T_i is sent by ϕ into a D-neighborhood of the projection of ϕx_i to T_i'. The proof is thus complete. □

By examining the triangles (three points) and quadrangles (four points), we obtain the following corollay from Proposition **??**.

Lemma 8.18. *Let $f : X \to Y$ be a (λ, c)-quasi-isometric embedding between two δ-hyperbolic spaces. Then there exists $C = C(\lambda, c, \delta) > 0$ such that for any $x, y, z, w \in X$ and their images $x' = f(x), y' = f(y), z' = f(z), w' = f(w) \in Y$:*

1. $\lambda^{-1}\langle x, y\rangle_z - C \le \langle x', y'\rangle_{z'} \le \lambda\langle x, y\rangle_z + C$
2. *Let $S_w(x, y, z) = \langle x, y\rangle_w - \langle x, z\rangle_w$ and $S_{w'}(x', y', z') = \langle x', y'\rangle_{w'} - \langle x', z'\rangle_{w'}$. Then*

$$\lambda^{-1}|S_w(x, y, z)| - C \le |S_{w'}(x', y', z')| \le \lambda|S_w(x, y, z)| + C$$

(3) *If* $S_w(x,y,z) > 0$, *then*

$$\lambda^{-1}S_w(x,y,z) - C \le S_{w'}(x',y',z') \le \lambda S_w(x,y,z) + C$$

Theorem 8.19. *Let X, Y be two hyperbolic spaces. Assume that there exists a quasi-isometry between X and Y. Then the Gromov boundary of X is homeomorphic to that of Y.*

Proof. Let ϕ be a quasi-isometry between X and Y. Fix a basepoint $o \in X$. Let ∂X be the set of all equivalent classes of geodesic rays issuing at o, and ∂Y be the set of all equivalent classes of geodesic rays issuing at $\phi(o)$.

First we see that there exists a bijection $\Phi := \partial\phi$ between ∂X and ∂Y. Let p be a geodesic ray ending at $p_\infty \in \partial X$. Then $\phi(p)$ is a quasi-geodesic ray. Apply Arzela-Ascoli Lemma 6.5 to $[\phi(o), \phi(p(n))]$. We obtain a geodesic ray q ending at $q_\infty \in Y_\infty$. Note that $q \subset N_D(\phi(p)), \phi(p) \subset N_D(q)$ for some uniform constant $D > 0$. Define $\Phi(p_\infty) = q_\infty$.

Clearly, Φ is well-defined. Moreover, it is bijective. Let $\Phi(p_\infty) = \Phi(p'_\infty)$ for $p_\infty, p'_\infty \in \partial X$. Then q, q' are asymptotic in Y. Since $\phi(p) \subset N_D(q), \phi(p') \subset N_D(q')$, we see that each of $\phi(p)$ and $\phi(p')$ lies in a uniform neighborhood of the other one. By the quasi-isometry, it follows that p, q are asymptotic. This proves that Φ is injective. Surjectivity is similar by using quasi-inverse of ϕ.

Using visual metrics, it is also easy to see it is a homeomorphism. It suffices to verify that Φ is continuous. We use sequences. Let $x_n \to x \in \partial X$. Then $d(o, [x_n, x]) \to \infty$. Suppose that $\Phi(x_n)$ does not converge to $\Phi(x)$. Then there exists a subsequence of x_n such that $d(\phi(o), [\Phi(x_n), \Phi(x)]) \le L$ for an uniform constant L. This leads to a contradiction with the first statement of Lemma 8.18. So $\Phi(x_n)$ converges to $\Phi(x)$, proving that Φ is a homeomorphism. □

We now procced to prove the boundary maps are quasi-conformal and quasi-symmetric maps.

Definition 8.20. An embedding $f : X \to Y$ is called *quasi-symmetric* if there exists a homeomorphism $\eta : [0,\infty) \to [0,\infty)$ so that

$$|x-a| \le t|x-b| \Leftarrow |f(x) - f(a)| \le \eta(t)|f(x) - f(b)|$$

for all triples (a, b, x) of points in X and for all $t > 0$.

Remark. If

$$f(t) = \begin{cases} \lambda t^{1/\alpha}, \ \forall t \in (0,1); \\ \lambda t^{\alpha}, \ \forall t \ge 1 \end{cases}$$

then f is called (λ, α)*-power quasi-symmetric* .

Theorem 8.21. *Let $\phi : X \to X$ be a quasi-isometry between hyperbolic spaces. Then the induced map $\partial\phi : \partial X \to \partial X$ is a quasi-symmetric map with respect to visual metric.*

Proof. Let $t = \frac{\rho_a(x,y)}{\rho_a(x,z)}$. Since $1/2e^{-a\langle x,y\rangle_w} \le \rho_a(x,y) \le e^{-a\langle x,y\rangle_w}$ by (21) we have

$$t/2 \le e^{-a\langle x,y\rangle_w}/e^{-a\langle x,z\rangle_w} \le 2t.$$

Recall that $S_w(x,y,z) = \langle x,y\rangle_w - \langle x,z\rangle_w$. Thus,

$$-\log 2t \le aS_w(x,y,z) \le -\log t/2. \tag{22}$$

Case 1. $t \le 1/2$. Then $aS_w(x,y,z) \ge -\log 2t \ge 0$ by (22). By Lemma 8.18, there exists $C = C(\lambda, c, \delta)$ such that

$$aS_{w'}(x',y',z') \ge a\lambda^{-1}S_w(x,y,z) - Ca \ge \lambda^{-1}(-\log 2t) - Ca.$$

Hence,

$$\frac{\rho_a(x',y')}{\rho_a(x',z')} \le \frac{2e^{-a\langle x',y'\rangle_{w'}}}{e^{-a\langle x',z'\rangle_{w'}}} \le 2e^{-aS_{w'}(x',y',z')} \le 2e^{Ca}(2t)^{1/\lambda}.$$

Case 2. $t \ge 1/2$. Note that

$$\begin{cases} \log 2t \le |\log t/2| : 1/2 \le t \le 1 \\ \log 2t > |\log t/2| : t \ge 1 \end{cases}.$$

Thus, (22) implies

$$|aS_w(x,y,z)| \le \begin{cases} |\log t/2| : 1/2 \le t \le 1 \\ \log 2t : t \ge 1 \end{cases}.$$

By Lemma 8.18, we have

$$a|S_{w'}(x',y',z')| \le a\lambda|S_w(x,y,z)| + Ca \le \lambda \max\{\log 2t, |\log t/2|\} + Ca$$

for $t \ge 1/2$. Hence,

$$\frac{\rho_a(x',y')}{\rho_a(x',z')} \le \frac{2e^{-a\langle x',y'\rangle_{w'}}}{e^{-a\langle x',z'\rangle_{w'}}} \le 2e^{-aS_{w'}(x',y',z')} \le \begin{cases} 2e^{Ca}(2/t)^{\lambda} : 1/2 \le t \le 1 \\ 2e^{Ca}(2t)^{\lambda} : t \ge 1 \end{cases}.$$

Therefore, $\partial\phi$ is η-quasi-symmetric, where

$$\eta(t) = \begin{cases} 2e^{Ca}(2t)^{1/\lambda} : t \le 1/2 \\ 2e^{Ca}(2/t)^{\lambda} : 1/2 \le t \le 1 \\ 2e^{Ca}(2t)^{\lambda} : t \ge 1 \end{cases}$$

Since the term $2e^{Ca}(2/t)^{\lambda}$ is bounded by a constant C_0 for $t \in [1/2, 1]$, we can choose

$$\eta(t) = \begin{cases} C_0 t^{1/\alpha} : t \le 1 \\ C_0 t^{\alpha} : t \ge 1 \end{cases}$$

so that $\partial\phi$ is power quasi-symmetric. □

Definition 8.22. A homeomorphism $f : X \to Y$ is called *quasi-conformal* if there exists a constant H so that

$$\limsup_{r\to 0} \frac{\sup_{d(x,y)=r} d(f(x), f(y))}{\inf_{d(x,y)=r} d(f(x), f(y))} \le H < \infty$$

forall x in X.

Remark. A homeomorphism $f : \mathbb{R} \to \mathbb{R}$ is quasi-conformal if it is everywhere differentiable with nonzero derivative.

However, for $X = Y = \mathbb{R}^n$ with $n \ge 2$, it is a fundamental fact in the theory of quasiconformal maps that quasi-conformal maps are quasi-symmetric, which appears to be a global condition.

It is much easier to prove quasi-conformality of the boundary map.

Theorem 8.23. *Let $\phi : X \to X$ be a quasi-isometry between hyperbolic spaces. Then the induced map $\partial\phi : \partial X \to \partial X$ is a quasi-conformal map with respect to visual metric.*

8.6. **Gromov boundary of free groups.** Let F be a free group generated by a finite set X. The Gromov boundary of F consists of all infinite reduced words over $\tilde{X}$. This is a susbet of the product $\tilde{X}^{\mathbb{N}}$. We equip X with discrete topology. Then the product $\tilde{X}^{\mathbb{N}}$ is a compact space by Tychonoff's theorem. Then Gromov boundary of F is a closed subset and thus compact.

We can also define a metric on ∂F. Let w_1, w_2 be two infinite words. Define $\rho(w_1, w_2) = e^{-n}$, where n is the maximal length of a common initial subword u of w, w', i.e.: $w_1 = uw_1'$ and $w_2 = uw_2'$.

It is easy to verify that ρ is a metric and the induced topology agrees with the subspace topology from $\tilde{X}^{\mathbb{N}}$.

It can be verified that ∂F is totally disconnected, and has no isolated points. Such a space is homeomorphic to a Cantor set.

9. Hyperbolic groups

9.1. **Finite order elements.**

Definition 9.1. A finitely generated group G is called *hyperbolic* if there exists a finite generating set S such that the Cayley graph $\mathscr{G}(G, S)$ is δ-hyperbolic for some $\delta \geq 0$.

Remark. Since hyperbolicity is a quasi-isometric invariant (see Theorem 7.19), we see that the definition of a hyperbolic group does not depend on the choice of a generating set.

Definition 9.2. Let $L, \lambda, c > 0$. A path p in X is called a *L-local (λ, c)-quasi-geodesic* if every subpath of p with length L is a (λ, c)-quasi-geodesic.

In hyperbolic spaces, a local quasi-geodesic turns out to be a global quasi-geodesic in the following sense. Its proof is not easy, we refer to [4, Page 405] for a proof.

Lemma 9.3 (Local=>Global)**.** *For any $\lambda, c > 0$, there exist $L_0 > 0$ and $\lambda', c' > 0$ with the following property.*

Fix any $L > L_0$. Let γ be a L-local (λ, c)-quasi-geodesic. Then γ is a (λ', c')-quasi-geodesic.

A particular interesting case in what follows is that a L-local geodesic path will be a global quasi-geodesic for large L.

Corollary 9.4. *There exists $L_0, \lambda, c > 0$ such that for any $L > L_0$, any L-local geodesic path is a (λ, c)-quasi-geodesic.*

Theorem 9.5. *In a hyperbolic group, there are only finitely many conjugacy classes of finite order elements.*

Proof. Assume that G is a hyperbolic group. Then there exists a finite generating set S such that $\mathscr{G}(G, S)$ is a δ-hyperbolic space for $\delta > 0$.

Let g be an element in G. We consider all the elements H of G which is conjugate to g. Choose $h \in H$ such that $d(1, h) = d(1, H)$. Then h is a shortest element in H to 1. Let p be a geodesic path between $1, h$. Denote by w the label of p, which is a word in $F(S)$ representing h.

We consider a path q labeled by the word w^n for some n. We can write $q = p_1 p_2 ... p_n$, where each p_i is labeled by a word w, in other words, is a translated copy of p.

We have the following general fact

Claim. *q is a $|w|$-local geodesic.*

Proof. Indeed, if not, there exists $x, y \in q$ such that $\mathrm{Len}([x, y]_q) = |w|$ and $d(x, y) < |w|$. Without loss of generality, we assume that $x \in p_1, y \in p_2$. Let w_1 be the label of the subsegment of p_1 from x to $(p_1)_+$, w_2 the label of the subsegment of p_2 from $(p_1)_+$ to y. Since $|w_1| + |w_2| = |w|$, we see that $w = w_2 w_1$. Then $w = w_1^{-1}(w_1 w_2) w_1$ is conjugate to $w_1 w_2$. Note that $w_1 w_2$ represents an element of length $d(x, y) < |w|$. This gives a contradiction to the choice of w, which has the minimal length among all the words representing elements in H. □

Let L_0 be the constant given by applying Corollary 9.3 to a local geodesic path.

Now assume that g has finite order n. Then by the claim, q is a $|w|$-local geodesic. By Lemma 9.3, we see $|w| < L_0$. If not, q is a global quasi-geodesic by Lemma 9.3. In particular, this implies that $q_- \neq q_+$. But q starts as 1 and ends at $g^n = 1$. So this is a contradiction.

Thus, we have proved that some conjugate h of g lies in the ball $B(1, L_0)$. Since $\mathscr{G}(G,S)$ is proper, this claim implies that any finite order element can be conjugated into a finite ball. This clearly proves the theorem. □

Remark. The conclusion can be strengthened that there are only finitely many conjugacy classes of finite subgroups in a hyperbolic group.

9.2. **Cone types and finite state automaton.** Assume that G is generated by a finite set S. Denote by $\mathscr{G}(G,S)$ the Cayley graph of G with respect to S.

Definition 9.6. For any $g \in G$, the *cone* $\Pi(g)$ at g is the set of elements $u \in G$ such that there exists SOME geodesic $[1, gu]$ in $\mathscr{G}(G,S)$ containing g.

Two cones $\Pi(g), \Pi(g')$ are of *same type* if $\Pi(g) = \Pi(g')$.

Theorem 9.7 (Cannon)**.** *Let G be a hyperbolic group with S a finite generating set. Then there are only finitely many cone types in $\mathscr{G}(G,S)$.*

Proof. For any $g \in G$, consider the set

$$S_g = \{h : d(g, gh) \leq \delta + 1,\ d(1, gh) \leq d(1, g)\}.$$

Then S_g has uniform finite cardinality for any $g \in G$. We shall prove that S_g determines the cone type of g. In other words, we will prove that if $S_g = S_{g'}$, then $\Pi(g) = \Pi(g')$.

Let $u \in \Pi(g)$. Denote $n = d(1, u)$. We prove $u \in \Pi(g')$ by induction on n. Let $n = 1$ and then $u \in S$. Since $u \notin S_g = S_{g'}$, we have $u \in \Pi(g')$.

We assume now that if an element u with $d(1, u) \leq n$ lies in $\Pi(g)$, then $u \in \Pi(g')$.

Let $v \in \Pi(g)$ such that $d(1, v) = n + 1$. Write $v = us$ for some generator $s \in \tilde{S}$ and $u \in \Pi(g)$ with $d(1, u) = n$. By inductive assumption, $u \in \Pi(g')$. Then $d(1, g'u) = d(1, g') + n$.

We claim that $us \in \Pi(g')$. Suppose not. Then $d(1, g'us) < d(1, g'u) + 1$. Choose a geodesic γ between $1, g'us$. Apply the thin-triangle property to $\Delta(1, g'u, g'us)$. Since $d(g'u, g'us) = 1$, there exists a vertex $x \in \gamma$ such that $d(x, g'us) = n$ and $d(x, g') \leq \delta$. Then $d(1, g'^{-1}x) \leq \delta$. Denote $h = g'^{-1}x$. Then $d(h, us) = n$.

Observe that $h \in S_{g'} = S_g$. Since $gus \in \Pi(g)$, we have $d(1, g) + n + 1 = d(1, gus)$. On the other hand, $d(1, gus) = d(1, gh(h^{-1}us)) \leq d(1, gh) + d(h, us) \leq d(1, g) + n$. This gives a contradiction. Thus, $us \in \Pi(g')$. This finishes the proof. □

Definition 9.8. A *finite state automaton* $\mathcal{M}$ is a finite oriented graph $\mathcal{G}$ with edges labeled by an alphabet set S. The vertices V are called *states*. There exists a *start* state $\iota \in V$ and a set $T \subset V$ of *accept* states.

A *language* is a subset of $\mathcal{W}(S)$. Every finite state automaton *defines* a language $L(\mathcal{M})$ which consists of the set of words which labels a path starting at ι and terminating at a state in T. A language L is called *regular* if it is defined by some finite state automaton.

Let G be a group with a finite generating set S. We define a natural language called *geodesic language* which consists of all words labelling geodesics originating at 1 in $\mathscr{G}(G,S)$.

We want to know whether a geodesic language is regular. It is clear that if a geodesic language is regular, then there are finitely many cone types. In fact, the converse is also true.

Lemma 9.9. *Let G be a group with a finite generating set S such that there are finitely many cone types in $\mathscr{G}(G,S)$. Then the set of all words labelling geodesics in $\mathscr{G}(G,S)$ is regular.*

Proof. We define a finite state automaton $\mathcal{M}$. Let $\mathcal{G}$ be a graph such that the vertex set is all cone types $\Pi(g)$ of $\mathscr{G}(G,S)$. Let $\tilde{S} = S \cup S^{-1}$ be the alphabet set. There exists an oriented edge labeled by $s \in \tilde{S}$ from the cone type $\Pi(g)$ to $\Pi(h)$ if $s \in \Pi(g) \cap \tilde{S}$ and $\Pi(gs)$ has the same cone type as $\Pi(h)$, i.e.: $\Pi(gs) = \Pi(h)$.

We need verify that edges are well-defined. Let $\Pi(g) = \Pi(g')$. We need show that $\Pi(gs) = \Pi(g's)$ for any $s \in \tilde{S} \cap \Pi(g)$. Assume that $\Pi(g)$ is connected to $\Pi(h)$ by edge labeled by $s \in \tilde{S}$. Then by definition of edges, $\Pi(gs) = \Pi(h)$. Let $\Pi(g) = \Pi(g')$. We need show that $\Pi(g's) = \Pi(h)$.

Let $u \in \Pi(gs)$. Then there exists SOME geodesic $p = [1, gsu]$ containing gs. As $s \in \Pi(g)$, there exists a geodesic $q = [1, gs]$ contains g. Thus, we can assume that p also contains g, up to a replacement of the segment $[1, gs]_p$ of p with q. This implies that $su \in \Pi(g)$. As $\Pi(g) = \Pi(g')$, there exists a geodesic $[1, g'su]$ containing g'. Similarly, we can assume that $[1, g'su]$ contains $g's$. This implies that $u \in \Pi(g's)$. Thus, $\Pi(gs) \subset \Pi(g's)$.

The case that $\Pi(g's) \subset \Pi(gs)$ is symmetric. Hence, we have proved that $\Pi(gs) = \Pi(g's)$. So the edge relation is well-defined. □

By the above lemma, a direct consequence of Theorem 9.7 is the following result.

Theorem 9.10. *Let G be a hyperbolic group with a finite generating set S. Then the geodesic language is regular.*

We now deduce from the finite state automaton that an infinite hyperbolic group contains an infinite order element.

Lemma 9.11 (Existence of infinite order elements)**.** *Let G be an infinite hyperbolic group. Then there exists an infinite order element in G.*

Proof. Let N be the number of cone types in $\mathscr{G}(G,S)$. By Theorem 9.7, $N < \infty$. Choose $g \in G$ such that $d(1, g) = N + 1$. Such g exists as G is infinite. Connect $1, g$ by a geodesic p. By Theorem 9.10, the word labeling p is a geodesic word that corresponds to a path in the graph $\mathcal{G}$. Since $\mathcal{G}$ has N vertices and p contains $N + 2$ vertices, there exists a subpath q of p that corresponds to a loop in $\mathcal{G}$. In other words, we can write $p = r_1 q r_2$, where q_-, q_+ have the same cone type.

Clearly, any path $r_1 q^n r_2$ gives a path in $\mathcal{G}$ and defines a geodesic word in $\mathscr{G}(G,S)$. Then q^n is a geodesic. Let w be the label of q. So the element given by w is of infinite order. □

Corollary 9.12. *An infinite torsion group is not hyperbolic.*

In a hyperbolic group, an infinite order element is also called *hyperbolic*, whereas a finite order element is called *elliptic*.

9.3. **Growth rate of a hyperbolic group.**

Definition 9.13. The growth function $\phi(n)$ of a language counts the number of words with length equal to n. The growth series is defined as follows

$$\Theta(z) = \sum_{i\geq 0} \phi(i) z^i, z \in \mathbb{R}_{>0}.$$

Remark. The convergence radius of $\Theta(z)$ is the following limit

$$\delta = (\limsup_{n\to\infty} \phi(n)^{1/n})^{-1}.$$

Lemma 9.14. *The growth series $\Theta(z)$ of a regular language $\mathcal{L}$ is rational for $t < \delta$.*

Proof. Let $\mathcal{M}$ be the finite state automaton which produces $\mathcal{L}$. Let n be the number of vertices in $\mathcal{G}$, which are indexed by $\{1, 2, ..., n\}$. We assume that 1 is the start state, and the others $2, ...n$ are accept states.

Define a matrix $A = (a_{ij})_{1\leq i,j\leq n}$, where a_{ij} is the number of edges of $\mathcal{G}$ which starts at i and ends at j.

Let $v = (1, 0...0), w = (1, 1, ..., 1)$ be row vectors. Observe that $vA^i w^T$ is the number of words of length i. Then

$$\Theta(z) = \sum_{i\geq 0} vA^i w^T z^i, z \in \mathbb{R}_{>0}.$$

By Cayley-Hamilton theorem, there exists a polynomial $f(z) = c_0 + c_1 z + ... + c_m z^m$ such that $f(A) = 0$. Define $\bar{f}(z) = c_m + c_{m-1} z + ... + c_0 z^m$. We calculate the coefficient of z^n in $\bar{f}(z)\Theta(z)$ as follows. Observe that for any $n \geq m$,

$$\sum_{0\leq i\leq m} v(c_i A^{n-m+i}) w^T z^n = v(A^{n-m} \sum_{0\leq i\leq m} (c_i A^i)) w^T z^n = 0.$$

Thus, $\bar{f}(z)\Theta(z)$ is a polynomial of degree at most m. So $\Theta(z)$ is a rational function, i.e. the ratio of two polynomials. □

Corollary 9.15. *The convergence radius of $\Theta(z)$ is an algebraic number.*

Proof. Since $\Theta(z)$ is rational, its convergence radius is a root of a polynomial. □

Recall that the *growth rate* of G with respect to S is the following limit

$$\delta_G = \lim_{n\to\infty} n^{-1} \ln b_n,$$

where b_n is the number of elements with length n. By Exercise 5.20, we have $e^{-\delta_G}$ is the convergence radius of the growth series $\Theta(z)$. Thus, we have the following.

Corollary 9.16. *The growth rate of a hyperbolic group with respect a finite generating set is $\log(g)$, where g is an algebraic number.*

10. Word and conjugacy problems of hyperbolic groups

In this section, we consider some algorithms problems in hyperbolic groups. In particular, we shall prove that word problems and conjugacy problems are solvable in hyperbolic groups.

10.1. Hyperbolic groups are finitely presentable.

Theorem 10.1. *Let G be a hyperbolic group with a finite generating set S. Then there exists a finite set R of words in $\mathcal{W}(\tilde{S})$ such that $G \cong \langle S|R\rangle$.*

Proof. Let $\phi : F(S) \to G$ be the natural epimorphism. Let R be the set of all loops in $\mathscr{G}(G, S)$ with length upper bounded by $4D + 2$, where D is constant given by Lemma 7.13[Stability of quasigeodesics]. Since $\mathscr{G}(G, S)$ is locally finite, we see R is a finite set.

For any $w \in \ker(\phi)$, let $p = e_1e_2...e_n$ be a loop based at 1 in $\mathscr{G}(G, S)$ such that $\mathrm{Lab}(p) = w$, where $n = \mathrm{Len}(p)$.

We list all vertices of p as follows: $v_0 = 1, v_1, v_2, ..., v_n = 1$, where $n = \mathrm{Len}(p)$. For each $0 < i < n$, we connect v_i and 1 by a geodesic segment q_i with $(q_i)_- = 1, (q_i)_+ = v_i$. Then we decompose p as a product of n loops $q_ie_iq_{i+1}^{-1}$ as follows:

$$\begin{aligned}\mathrm{Lab}(p) &= \mathrm{Lab}(e_1)\mathrm{Lab}(e_2)\cdots\mathrm{Lab}(e_n)\\ &= (\mathrm{Lab}(e_1)\mathrm{Lab}(q_1)^{-1})(\mathrm{Lab}(q_1)\mathrm{Lab}(e_2)\mathrm{Lab}(q_2)^{-1})\\ &\cdots(\mathrm{Lab}(q_i)\mathrm{Lab}(e_i)\mathrm{Lab}(q_{i+1})^{-1}\cdots\mathrm{Lab}(q_{n-1})\mathrm{Lab}(e_n).\end{aligned}$$

Note that q_i and q_{i+1} are two geodesics such that $d((p_i)_+, (p_{i+1})_+) \le 1$. By Lemma 7.13[Stability of quasigeodesics], for each vertex w_{ij} of q_i with $0 < j < \mathrm{Len}(q_i)$, there exists $w'_{i,j} \in q_{i+1}$ such that $d(w_{ij}, w'_{ij}) \le D$ for an uniform constant $D > 0$. Similarly, we see that the loop $q_ie_i(q_{i+1})^{-1}$ can be decomposed as a product of loops c_{ij} for $1 \le j \le \mathrm{Len}(p_i)$, where

$$c_{ij} := [w_{ij}, w_{ij+1}]_{q_i}[w_{ij+1}, w'_{ij+1}][w'_{ij+1}, w'_{ij}]_{q_{i+1}}[w'_{ij}, w_{ij}].$$

Note that c_{ij} are of length upper bounded by $4D + 2$. Thus, $c_{ij} \in R$.

Therefore, any word $w \in \ker(\phi)$ can be written as a finite product of conjugates of relators in R. This implies by Lemma 3.3 that there exists a homomorphism $G = \langle S| \ker(\phi)\rangle \to \langle S|R\rangle$. By the choice of R as loops in $\mathscr{G}(G, S)$, we have $R \subset \ker(\phi)$. By Lemma 3.3, the homomorphism is an isomorphism. Hence, $G = \langle S|R\rangle$. □

10.2. Solving word problem.

We shall prove that word problems of hyperbolic groups can be solved. In order to do so, we shall prove that there is a special presentation called Dehn presentation for a hyperbolic group. A Dehn presentation by definition can allow to solve word problems in linear time.

Definition 10.2 (Dehn presentation)**.** Let $G \cong \langle S|R\rangle$ be a group presentation. Assume that R contains all cyclic permutations of every $r \in R$.

The $\langle S|R\rangle$ is called a *Dehn presentation* if any $w \in \langle\langle R\rangle\rangle$, there exists a subword u of w and a relator $r \in R$ such that $r = uv^{-1}$ and $\mathrm{Len}(u) > \mathrm{Len}(v)$.

Remark. Groups with a Dehn presentation have solvable word problems:

(1) Step 1: Given a reduced word $w \in F(S)$, compare subwords of w with relators in R. Then there exists a subword u of w and a relator $r \in R$ such that $r = uv$ and $\mathrm{Len}(u) > \mathrm{Len}(v)$.

(2) Step 2: We obtain a new word w_1 by replacing u with v, which has length strictly less than $|w|$.
(3) Step 3: Go to the Step 1.

Theorem 10.3 (Dehn presentation for hyperbolic groups). *A hyperbolic group has a Dehn presentation. Thus, word problem is solvable in hyperbolic groups.*

Proof. Let $\phi : F(S) \to G$ be the natural epimorphism. Let $L_0 > 2$ be the constant given by applying Corollary 9.3 to a local geodesic path.

Let R be the set of words in $\mathcal{W}(\tilde{S})$ labeling loops p such that $p = t^{-1}q$, where t is a geodesic and $\mathrm{Len}(q) \le L_0$. Thus, every p is of length at most $2L_0$. So R is a finite set. Moreover, it must be non-empty. Indeed, we shall show now that any loop in $\mathscr{G}(G, S)$ can be decomposed as a product of loops labeled by words in R.

Let $w \in \ker(\phi)$ be a reduced word in $F(S)$. Then in $\mathscr{G}(G, S)$, there is an embedded loop p such that $\mathrm{Lab}(p) = w$. We argue by induction on length n of w. The case $n = 2$ is trivial as $L_0 \ge 2$: $w \in R$. Now assume that ant word w of length at most $n - 1$ can be written as a finite product of conjugates of words in R. We prove below the case $\mathrm{Len}(w) = n$.

First we observe that p is not a L_0-local geodesic, otherwise by Corollary 9.3, the loop p would be a quasi-geodesic. Thus, there exists a subpath q of p such that $\mathrm{Len}(q) \le L_0$ and q is not a geodesic in $\mathscr{G}(G, S)$. We connect q_-, q_+ by a geodesic segment t. This gives a loop $t^{-1}q$ of length less than $2L_0$, which lies in R by definition.

We write $p = [1, q_-]_p q [q_+, 1]_p$. Thus, $w = u\mathrm{Lab}(q)v$, where $u = \mathrm{Lab}([1, q_-]_p), v = \mathrm{Lab}([q_+, 1]_p)$. Note that $u\mathrm{Lab}(t)v$ is of length at most $n-1$ and $\phi(u\mathrm{Lab}(t)v) = \phi(w)$. Thus, $u\mathrm{Lab}(t)v$ can be written as a finite product of conjugates of $r \in R$. Thus $w = u\mathrm{Lab}(t)vv^{-1}\mathrm{Lab}(t)^{-1}\mathrm{Lab}(q)v = (u\mathrm{Lab}(t)v)(v^{-1}\mathrm{Lab}(t^{-1}q)v)$. This proves that w is a finite product of conjugates of $r \in R$. Thus, $\langle S|R\rangle$ is a finite presentation of G.

Moreover, it is a Dehn presentation by the arguments above: any word w contains a subword $\mathrm{Lab}(p)$, which is a subword of a loop $pt^{-1} \in R$. □

10.3. **Solving conjugacy problem.** Solving conjugacy problem depends on the following result.

Lemma 10.4 (Bounding conjugators). *Assume that $g = fhf^{-1}$ for some $g, h, f \in G$. Then there exist a constant $D = D(|g|, |h|)$ and $f' \in G$ such that $d(1, f') \le D$ and $g = f'hf'^{-1}$.*

Remark. Here D, as shown in the proof, is a linear function of $|g|$ and $|h|$.

Proof. We draw a quadrangle $rptq$ for the points $1, g, f, fh = gf$, where $r = [1, g], p = [g, gf], t = [fh, f], q = [1, f]$. Note that p, q are of same length, and has same label representing f.

Let $x \in p$ such that $d(g, x) \ge |g|+|h|+2\delta$ and $d(x, gf) \ge |g|+|h|+2\delta$. Note that $rptq$ has 2δ-thin property. Then there exists $z \in q$ such that $d(x, z) \le 2\delta$. We obtain by triangle inequality that $d(g, x) + 2\delta + |g| \ge d(1, z) \ge d(g, x) - 2\delta - |g| \ge d(1, z)$.

Let $y \in q$ such that $d(1, y) = d(g, x)$. Thus, $d(1, y)+2\delta+|g| \ge d(1, z) \ge d(1, y) - 2\delta - |g| \ge d(1, z)$. It follows that $d(y, z) \le 2\delta + |g|$. This gives $d(x, y) \le 4\delta + |g|$.

We consider the vertices $\{x_0, x_1, ..., x_n\}$ on p such that $d(g, x_i) \ge |g| + |h| + 2\delta$ and $d(x_i, gf) \ge |g| + |h| + 2\delta$, where

$$n = |f| - 2|g| - 2|h| - 4\delta.$$

By the above argument, for each $x_i \in p$, there exists $y_i \in q$ such that $d(x_i, y_i) \leq 4\delta + |g|$ and $d(1, y_i) = d(g, x_i)$.

Let $N = |B(1, 4\delta + |g|)|$. Assume that $|f| > 2|g| + 2|h| + 4\delta + N + 1$. Then $n > N+1$. There exists $1 \leq i < j \leq n$ such that $x_i^{-1} y_i = x_j^{-1} y_j = k \in B(1, 4\delta+|g|)$. Clearly, we have $gf_1 = f_1 k$ and $kf_2 = f_2 h$ for $f_1 = \mathrm{Lab}([1, y_i]_q)$, $f_2 = \mathrm{Lab}([y_j, f]_q)$. Thus, $gf_1 f_2 = f_1 f_2 h$.

Denote $f' = f_1 f_2$. Note that $|f'| \leq |f_1| + |f_2| < |f|$. So we find a new element f' with smaller length than f such that $gf' = f'h$. Inductively, the length of f' gets decreased until $|f'| \leq 2|g|+2|h|+4\delta+N+1$. Thus, we set $D = 2|g|+2|h|+4\delta+N+1$. This gives the desired constant. □

Let $g, h \in G$. By Lemma 10.4, if g, h are conjugate, then there exists $f \in G$ with length bounded by a linear function $D = D(|g|, |h|)$ such that $g = fhf^{-1}$ for some $f \in G$. So to solve conjugacy problem, it suffices to list a finite number of f with length bounded by D, and use word algorithm to check whether $fhf^{-1}g^{-1}$ is the identity.

11. Subgroups in hyperbolic groups

11.1. **Quasi-convex subgroups.** In a metric space, a (quasi-)convex subspace usually share good properties of the ambient space. We start with a quasi-convexity notion in the Cayley graph of a group.

Definition 11.1. Let G be a group with a finite generating set S. A subgroup H is called σ*-quasi-convex* with respect to S in G if for any geodesic p with $p_-, p_+ \in H$, we have $p \subset N_\sigma(H)$ in $\mathscr{G}(G, S)$.

Remark. Quasi-convexity depends on the choice of a generating set. For example, $Z = \langle ab \rangle$ in $Z \times Z = \langle a, b : ab = ba \rangle$ is quasi-convex with respect to $\{a, b, ab\}$, but not quasi-convex with respect to $\{a, b\}$.

Lemma 11.2 (Quasi-convex subgroups are finitely generated)**.** *Let G be a group with a finite generating set S. Then a quasi-convex subgroup is finitely generated.*

Proof. Let H be a σ-quasi-convex subgroup in G with respect to S. For any $h \in H$, let p be a geodesic between $1, h$ in $\mathscr{G}(G, S)$. Write $p = e_1 e_2 ... e_n$, where $n = \mathrm{Len}(p)$ and e_i are edges labeled by $s_i \in S$. Then $\mathrm{Lab}(p) = s_1 s_2 ... s_n$ is a word over $\tilde{S}$ representing h.

Since H is σ-quasi-convex, there exists $h_i \in H$ such that $d_S((e_i)_+, h_i) < \sigma$ for all $1 \le i \le n$. Here we choose $h_n = h$. Thus, $d_S(h_i, h_{i+1}) \le 2\sigma + 1$ and then $h_i^{-1} h_{i+1} \in B(1, 2\sigma + 1)$. Observe that $h = h_n = h_1 (h_1^{-1} h_2) ... (h_{n-1}^{-1} h_n)$. Denote $T = H \cap B(1, 2\sigma + 1)$, which is a finite set. Thus H is generated by T. □

The main result in this subsection is the following lemma, saying that quasi-convexity is stable under taking intersection.

Lemma 11.3 (Stability of quasi-convexity under intersection)**.** *Let H, K be quasi-convex subgroups in a group G. Then $H \cap K$ is quasi-convex and thus finitely generated.*

Proof. Assume that H, K are σ-quasi-convex for $\sigma > 0$. Let $c \in H \cap K$ and p be a geodesic between $1, c$ in $\mathscr{G}(G, S)$. Then $p \subset N_\sigma(H) \cap N_\sigma(K)$. We aim to find a constant $\sigma' > 0$ such that $p \subset N_{\sigma'}(H \cap K)$.

Let $x \in p$. We consider the set Q of paths q in $\mathscr{G}(G, S)$ such that $q_- = x, q_+ \in H \cap K$ and $q \subset N_\sigma(H) \cap N_\sigma(K)$. The set Q is non-empty, as $[x, c]_p$ lies in Q.

Denote $\sigma' = |B(1, 2\sigma)|^2 + 1$. We claim that there exists some path q in Q such that $\mathrm{Len}(q) \le \delta'$. This clearly implies that $x \in N_{\sigma'}(H \cap K)$.

Let $q \in Q$ such that $\mathrm{Len}(q) \ge \sigma'$. We are going to make a "shortcut" on q to get a new $q' \in Q$ such that $\mathrm{Len}(q') \le \mathrm{Len}(q) - 1$. This will finish the proof of our claim.

Let $v_1, v_2, ... v_n$ be the set of vertices of q, where $n > \sigma'$. Since $q \subset N_\sigma(H) \cap N_\sigma(K)$, there exists h_i, k_i such that $d(v_i, h_i), d(v_i, k_i) \le \sigma$ for $1 \le i \le n$. Note that $n > \sigma' > |B(1, 2\sigma)|^2$. By pigeonhole principle, there exist some $1 \le i < j \le n$ such that $v_i^{-1} h_i = v_j^{-1} h_j$ and $v_i^{-1} k_i = v_j^{-1} k_j$. It also follows that $k_i^{-1} h_i = k_j^{-1} h_j$. Note that $t := k_i k_j^{-1} = h_i h_j^{-1} \in H \cap K$.

We define $q' = [x, v_i]_q (t \cdot [v_j, q_-]_q)$, where $t \cdot [v_j, q_-]_q$ is a translated copy by t of subsegment $[v_j, q_-]_q$ in q. Thus, every vertex on q' lies in $N_\sigma(H) \cap N_\sigma(K)$. Since $t \in H \cap K$, the segment $t[v_j, q_-]_q$ starts at v_i and ends at $tq_+ \in H \cap K$. This shows that $q' \in Q$. As $j > i$, we have $\mathrm{Len}(q') < \mathrm{Len}(q)$. □

The idea of the proof of Lemma 11.3 can be used to prove the following stronger lemma.

Lemma 11.4. *Let H, K be two subgroups in a finitely generated group G. Then for any $\sigma > 0$ there exists $\sigma' > 0$ such that $N_\sigma(H) \cap N_\sigma(K) \subset N_{\sigma'}(H \cap K)$.*

Sketch of Proof. For a pair of $a, b \in B(1, \sigma)$, we define $G_{a,b} = \{g \in G : ga \in H, gb \in K\}$. For each $G_{a,b}$, we choose an element $g_{ab} \in G_{a,b}$ such that $d(1, g_{ab}) = d(1, G_{a,b})$. Let F be the union of finitely many g_{ab}, where $a, b \in B(1, \sigma)$. Define $\sigma' = \max\{d(f, H \cap K) : f \in F\}$. It can be verified that σ' is the desired constant. □

Form now on, we are interested in understanding quasi-convex subgroups in hyperbolic groups. As remarked above, the quasi-convexity of a subgroup, in general, depends on the choice of the generating set. However, the notion of quasi-convexity in hyperbolic groups is independent of the choice of generating sets in the following sense.

Lemma 11.5. *Let H be a subgroup in a hyperbolic group G. Let S, T be any two finite generating sets of G. Then H is σ_S-quasi-convex with respect to S if and only if H is σ_T-quasi-convex with respect to T*

Proof. Clearly, it is a consequence of the stability of quasi-geodesics, see Lemma 7.13. □

Hence, in what follows, we can speak of the quasi-convexity of a subgroup in a hyperbolic group without ambiguity.

We introduce another class of subgroups called undistorted subgroups, which turns out to be equivalent to the class of quasi-convex subgroups in hyperbolic groups.

Definition 11.6. Let H be a finitely generated group in a finitely generated group. We say that H is *undistorted* in G if there exist finite generating sets S, T for G, H respectively such that the inclusion $(H, d_T) \to (G, d_S)$ is a quasi-isometric map.

We now prove that quasi-convex subgroups in a hyperbolic group are the same as undistorted groups.

Lemma 11.7 (Quasi-convexity $\Leftrightarrow$ Undistortedness)**.** *Let G be a hyperbolic group with a finite generating set S. Then a subgroup H is quasi-convex if and only if for some (or any) finite generating set T of H, the inclusion $(H, d_T) \to (G, d_G)$ is a quasi-isometric map.*

Proof. =>. We continue to use the notions in the proof of Lemma 11.2. By construction, we have
$$d_T(1, h) \le n \le d_S(1, h).$$
We consider the finite generating set $S \cup T$ for G. Then clearly $d_{S\cup T}(1, h) \le d_T(1, h)$. Since $d_{S\cup T}$ is bi-Lipchitz to d_S, there exists $\lambda > 1$ such that $d_S(1, h) \le \lambda d_T(1, h)$. This proves that $(H, d_T) \to (G, d_S)$ is a $(\lambda, 0)$-quasi-isometric embedding map.

<=. Without loss of generality, assume that $T \subset S$. Then the Cayley graph of H with respect to T is a subgraph of $\mathscr{G}(G, S)$. And the natural inclusion $\iota : (H, d_T) \to (G, d_S)$ is a quasi-isometric embedding map.

Let $h \in H$ and a geodesic p between $1, h$ in the Cayley graph of H with respect to T. Then p is a quasi-geodesic in $\mathscr{G}(G, S)$. By Lemma 7.13, any geodesic q in

$\mathscr{G}(G, S)$ between $1, h$ is contained in a σ-neighbourhood of p for some uniform σ. Since the vertex set of p lies in H, this implies that $q \subset N_{\sigma+1}(H)$. The proof is complete. □

Remark. The => direction does not need the assumption that G is hyperbolic, but which is necessary in the <= direction: the inclusion of Z into $Z \times Z$ is a quasi-isometric map (no matter which generating sets of $Z, Z \times Z$ are chosen), but Z may not be quasi-convex with respect to some generating set of $Z \times Z$.

Corollary 11.8. *Let H be a quasi-convex subgroup in a hyperbolic group. Then H is a hyperbolic group.*

Proof. The proof is same as the proof of Theorem 7.19. Let T, S be generating sets of H and G respectively. Without loss of generality, assume that $T \subset S$. Let $\iota : H \to G$ be the inclusion. By Lemma 11.7, ι is a quasi-isometric map.

Let $\Delta = \Delta(a, b, c)$ be a geodesic triangle in (H, T), where $a, b, c \in H$. As $\mathscr{G}(H, T)$ is a subgraph of $\mathscr{G}(G, S)$ and ι is a quasi-isometric map, Δ is a quasi-geodesic triangle in $\mathscr{G}(G, S)$. By the δ-thin triangle property in $\mathscr{G}(G, S)$, there exists $o \in G$ such that o is a δ-center of Δ. Without loss of generality, we can assume that $o \in \Delta$. Again since ι is a quasi-isometric map, $o \in H$ is δ'-center of Δ for some δ' in $\mathscr{G}(H, T)$. □

Exercise 11.9. *Let H be a finitely generated subgroup in an abelian group G of finite rank. Then H is undistorted in G.*

In free groups, the quasi-convexity of a subgroup is equivalent to the finitely generatedness.

Exercise 11.10. *Prove that any finitely generated subgroup in a free group of finite rank is quasi-convex. Thus, the intersection of any two finitely generated subgroups is finitely generated.*

11.2. **Hyperbolic elements in a hyperbolic group.** Recall that an infinite order element in a hyperbolic group is called hyperbolic. In this subsection, we study hyperbolic elements in details.

Theorem 11.11. *Let g be an arbitrary element in a hyperbolic group G. Then the centralizer $C(g) = \{c \in G : cg = gc\}$ is σ-quasi-convex for some σ depending on g.*

Proof. Let $c \in C(g)$. Then $gc = cg$. We draw a quadrangle $rptq$ for the points $1, g, gc = cg, c$, where $r = [1, g], p = [g, gc], t = [cg, c], q = [1, c]$. For any $x \in q$, we want to bound the distance $d(x, C(g))$ by a uniform constant.

Let $y \in p$ such that $d(p_-, y) = d(1, x)$. By the thin-triangle property, there exists a constant $D_1 = D(|g|)$ such that $d(x, y) \leq D_1$. Observe that $[p_-, y]_p$ and $[1, x]_q$ have the same label representing the element $x \in G$.

Let $h = x^{-1}y$ be the element represented by the label on $[x, y]$. Thus, we have g, h are conjugate: $g = xhx^{-1}$. By Lemma 10.4, there exists a constant $D_2 = D(|g|, |h|)$ and an element $u \in G$ such that $d(1, u) \leq D_2$ and $g = u^{-1}hu$. This implies that $gxu = xug$ and then $xu \in C(g)$. Moreover, $d(x, xu) \leq D_2$, where D_2 is an uniform constant depending only on g. □

Before going on, we need the following useful result about quasi-convex subsets.

Lemma 11.12. *Let Q be a σ-quasi-convex subset in a δ-hyperbolic space X. There exists a constant $\lambda = \lambda(\sigma, \delta), c = c(\sigma, \delta) > 0$ with the following property.*

Let $x \in X$ and $y \in Q$ a projection point of x. Then for any $z \in Q$, the path $p = [x, y][y, z]$ is a (λ, c)-quasi-geodesic.

Proof. This can be proven similarly as Exercise 7.21, which is a special case: a quasi-geodesic is a quasi-convex subspace.

Let $w \in [y, z]$ be a shortest point to x. Then $d(x, w) \le d(x, y)$. On the other hand, since Q is σ-quasi-convex, $[y, z] \subset N_\sigma(Q)$. Then there exists $w' \in Q$ such that $d(w, w') \le \sigma$. Since $d(x, y) = d(x, Q)$, it follows that $d(x, y) \le d(x, w) + \sigma$. Thus, $d(x, w) \le d(x, y) \le d(x, w) + \sigma$.

Since w is a δ-center of $\Delta(x, y, w)$, we have $d(x, w) + d(y, w) \le d(x, y) + 2\delta$. Then $d(y, w) \le 2\delta + \sigma$. Note that w is also a δ-center of $\Delta(x, w, z)$. Then $d(z, w) + d(w, z) \le d(x, z) + 2\delta$.

We calculate the length of $p = [x, y][y, z]$: $d(x, y) + d(y, z) \le d(x, w) + 2d(y, w) + d(w, z) \le 6\delta + 2\sigma + d(x, z)$. This shows that p is a $(1, 6\delta + 2\sigma)$-quasi-geodesic. □

We are now ready to prove the following result.

Theorem 11.13. *Let H be an infinite quasi-convex subgroup in a hyperbolic group G. Then H is of finite index in its normalizer $N(H) = \{g \in G : gH = Hg\}$. In particular, $N(H)$ is also quasi-convex.*

Proof. We want to find an uniform constant $D > 0$ such that for any $g \in N(H)$, $d(H, Hg) \le D$. If this is proven, we see that any coset Hg in $N(H)$ contains a representative element $g' \in Hg$ such that $d(1, g') \le D$. This will complete the proof.

Let σ be the quasi-convexity constant for H. Note that gH is also σ-quasi-convex for any $g \in G$. For any $x, y \in G$, we project x, y to $z, w \in gH$ respectively. We claim that

Claim. *If $d(z, w)$ is sufficiently large, then the concatenated path $p = [x, z][z, w][w, y]$ is a quasi-geodesic.*

Proof. By Lemma 11.12: p is a L-local (λ, c)-quasi-geodesic, where $L = d(u, v)$ and λ, c depends only on σ. Thus, by Lemma 9.3, p is a quasi-geodesic, if L is sufficiently large. □

For any $g \in N(H)$, we consider the σ-quasi-convex subset $Hg = gH$. Let $y \in H$. Then $d(y, Hg) \le d(1, g)$. Let $z, w \in gH$ be projections points of $1, y$ respectively. Note that $d(z, w) > d(1, y) - 2d(1, g)$. Since H is infinite, we can choose $d(1, y)$ to be sufficiently large such that $d(z, w)$ is also very large. It follows by the Claim that $p = [1, z][z, w][w, y]$ is a quasi-geodesic. By Lemma 7.13, there exists an uniform constant D such that $p \subset N_D([1, y])$. Since H is σ-quasi-convex, $p \subset N_{D+\sigma}(H)$.

Since $z, w \in gH$ and gH is σ-quasi-convex, we see that $[z, w] \subset N_\sigma(gH)$. It follows that $Hg = gH$ contains an element g' such that $d(g', H) \le D + 2\sigma$. □

Remark. The infiniteness of H cannot be dropped: consider a hyperbolic group which is the direct product of a hyperbolic group G with a finite group F.

An important fact about hyperbolic elements is that the subgroup generated by a hyperbolic group is quasi-convex or quasi-isometrically embedded.

Lemma 11.14. *Let g be a hyperbolic element in a hyperbolic group G. Then $n \in \mathbb{Z} \to g^n \in G$ is a quasi-isometric embedding map of $\mathbb{Z} \to G$. Moreover, $\langle g \rangle$ is of finite index in $C(g)$.*

Proof. Denote $H = \langle g \rangle$. We consider the centralizer $C(g)$ of g. By Lemma 11.11, $C(g)$ is quasi-convex and thus finitely generated by a finite set T. Observe that the center K of $C(g)$ is the intersection of centralizers of all $t \in T$ in $C(g)$. Since it is a finite intersection of quasi-convex subgroups, the center K is finitely generated and abelian.

Note that $H \subset K$. By Exercise 11.9, the natural inclusion of H into K is a quasi-isometric map. And $K \to G$ is also a quasi-isometric map. This implies that the inclusion of H into G is a quasi-isometric map also.

Since $C(g) \subset N(H)$, it follows by Lemma 11.13 that H is of finite index in $C(g)$. □

A direct corollary is the following result, saying that existence of abelian groups of rank at least two is an obstruction to hyperbolicity.

Corollary 11.15. *A hyperbolic group cannot contain a subgroup isomorphic to $Z \times Z$.*

Exercise 11.16. *Prove that $SL(n, \mathbb{Z})$ for $n \geq 3$ is not a hyperbolic group.*

Furthermore, we can give another obstruction result to hyperbolicity of a group.

Corollary 11.17. *A hyperbolic group cannot contain Baumslag-Solitar groups*

$$B(m, n) = \langle a, t | a^n = ta^m t^{-1} \rangle$$

for $m, n \in \mathbb{Z} \setminus 0$ and $|m| \neq |n|$.

Proof. Without loss of generality $|m| < |n|$. We first note that the following identity holds

$$t^l a^{m^l} t^{-l} = a^{n^l}$$

for any $l \in \mathbb{N}$. By Lemma 11.14, we have $|n|\lambda + c > d(1, g^n) > \lambda^{-1}|n| - c$ for any $n \in \mathbb{Z}$ and some $\lambda, c > 0$. Thus $2l \cdot d(1, t) + \lambda |m|^l + c > d(1, t^l a^{m^l} t^{-l}) = d(1, a^{n^l}) > \lambda^{-1}|n|^l - c$ holds for all $l > 0$. This gives a contradiction, as $|m| < |n|$. The proof is complete. □

Definition 11.18. Let G be a group with a finite generating set S. The *translation length* of an element $g \in G$ is the following limit

$$\tau(g) = \lim_{n \to \infty} d_S(1, g^n)/n.$$

Exercise 11.19. *Verify the following properties of translation length.*

1. *Prove that the limit $\tau(g)$ exists.*
2. *$\tau(g) = \tau(hgh^{-1})$ for any $h \in G$.*
3. *$\tau(g^n) = n\tau(g)$ for any $n \in \mathbb{N}$.*

We prove now that every hyperbolic element almost leaves invariant an uniform quasi-geodesic (called quasi-axe).

Lemma 11.20 (Quasi-axe)**.** *There exist constants $\lambda, c, C \geq 0$ with the following property. For any hyperbolic element g in G, there exists a (λ, c)-quasi-geodesic p in such that $p \subset N_D(g^n p)$ for any $n \in \mathbb{Z}$.*

Proof. Let L_0, λ, c be given by Corollary 9.3. Denote by h an element of minimal length in the conjugacy class of g. First note that the conclusion holds under conjugation of $g \in G$. We will prove the lemma for h.

Assume that $d(1, h) > L_0$. Fix a geodesic $q = [1, h]$. As in proof of Theorem 9.5, the concatenated bi-infinite path $p = ...(h^{-n}q)...(h^{-1}q)q(hq)...(h^n q)...$ is a $\mathrm{Len}(q)$-local geodesic. By Lemma 9.3, p is a (λ, c)-quasi-geodesic.

We now consider the case that $d(1, h) \le L_0$, for which by Lemma 11.14, the path $p = [h^{-n-1}, h^{-}n]...[h^{-1}, 1][1, h]...[h^n, h^{n+1}]$ is a quasi-geodesic. Note that there are only finitely many such elements h with $d(1, h) \le L_0$. By increasing by finite times the above λ, c, we see that p in the both cases are (λ, c)-quasi-geodesics.

By construction of p, it follows that $p \subset N_D(h^n p)$ for any $n \in \mathbb{Z}$ and some finite $D > 0$. Similarly as Lemma 8.2, we can argue by thin-triangle property to get an uniform constant C depending only on λ, c such that $p \subset N_D(g^n p)$ for any $n \in \mathbb{Z}$. □

Finally, we consider the spectrum of translation lengths of all hyperbolic elements in a hyperbolic group.

Theorem 11.21 (Discrete Rational Translation Lengths)**.** *Let G be a hyperbolic group. Then the set of translation lengths of hyperbolic elements is a discrete set of positive rational numbers in $\mathbb{R}$. Moreover, the denominator of each translation length is uniformly bounded by a constant.*

Proof. Let $Q = \{\tau_g : g \in G\}$. We prove the following statements.

(1). The set Q is a discrete set in $\mathbb{R}$. Since the translation length is a conjugacy invariant, without loss of generality, we assume that g is a minimal element in its conjugacy class. By Lemma 11.20, $n|g| \le \lambda d(1, g^n) + c$ for some uniform constant $\lambda, c > 0$. This implies that $|g| \le \lambda\tau(g)$. Hence, $\{\tau(g) \le r\}$ for any $r > 0$ is a finite set.

(2). Each $\tau \in Q$ is a rational number of form a/b, where b is less than an uniform constant. We consider the path $p = [1, g][g, g^2]...[g^{n-1}, g^n]....$ By Lemma 11.20, p is a (λ, c)-quasi-geodesic for some uniform constant $\lambda, c > 0$.

We fix some N very large, which will be computed below. Connect $1, g^N$ by a geodesic q. By Lemma 7.13, there are an uniform constant D and $v_i \in q$ such that $d(g^i, v_i) \le D$ for $1 \le i \le N$. Denote by $t_i = v_i^{-1} g^i \in B(1, D)$.

Let M be the number of cone types in $\mathscr{G}(G, S)$. By Theorem 9.7, $M < \infty$. We choose $N > M \cdot |B(1, D)| + 1$. By pigonhole principle, there are $1 \le i < j \le N$ such that $t = t_i = t_j$ and v_i, v_j has the same cone types.

Let h be the element represented by $[v_i, v_j]_q$. Then $h = tg^{j-i}t^{-1}$. Since v_i, v_j has the same cone type, we see that $[1, h][h, h^2]...[h^n, h^{n+1}]$ is a geodesic. Thus, $\tau(h) = |h|$ and then $\tau(g^{j-i}) = |h|$. This implies that $\tau(g) = |h|/(j - i)$, where $j - i \le N$. Hence, $\tau(g)$ is a rational number, whose denominator is uniformly upper bounded. □

12. Convergence groups: Tits alternative in hyperbolic groups

In this section, we study hyperbolic groups and their subgroups using the group action on Gromov boundary called *convergence group action*.

12.1. **Convergence group actions and classification of elements.** We introduce now a notion of a convergence group action. We refer the reader to [1] for more details about convergence group actions.

Definition 12.1. Assume that a (discrete) group G acts by homeomorphisms on a compact metrizable space M. Assume that $|M| \ge 3$. Then G is called *a convergence group* on M if the induced action of G on the space of distinct triples

$$\Theta^3(M) := \{(x, y, z) \in M^3 : x \ne y \ne z \ne x\}$$

is proper: the following set

$$\{g \in G : gK \cap K \ne \emptyset\}$$

is finite for any compact set K in $\Theta^3(M)$.

Corollary 12.2. *Assume that G acts as a convergence group on M. Then every element of infinite order in G has at most two fixed points in M.*

Proof. Let g be an infinite order element with distinct fixed points a, b, c. Consider the compact set $K = \{(a, b, c)\} \subset \Theta^3(M)$. Since $g^n(a, b, c) = (a, b, c)$ for all $n \in \mathbb{Z}$, this contradicts to the properness of G on $\Theta^3(M)$. □

Let $g \in G$ be an element in a convergence group action of G on M. It is called

(1) *elliptic* if g is of finite order,
(2) *parabolic* if it has infinite order and a unique fixed point in M,
(3) *hyperbolic* if it has infinite order and two fixed points in M.

By Corollary 12.2, any element in a convergence group can be classified into mutually exclusive classes of elliptic, parabolic and hyperbolic elements.

The convergence group action has the following convergence property for infinite sequences.

Definition 12.3. A group G is said to have the *convergence property* if for any sequence $\{g_n\}$ of G, there exists a subsequence $\{g_{n_i}\}$ (called *collasping sequence*) and two points $a, b \in M$ such that g_{n_i} converges to b locally uniformly on $M \setminus \{a\}$.

Theorem 12.4 (Convergence property of sequences)**.** *Let G be acting on a compact metrizable space M as a convergence group. Then any infinite set $\{g_n : n \in \mathbb{N}\}$ has convergence property.*

Remark. The converse of Theorem 12.4 is also true and much eaiser (Exercise 12.9): the convergence property characterizes the convergence group action.

Lemma 12.5 (Collapsing property)**.** *Assume that there exists a sequence g_n with the following property. Let $x_n \to x, y_n \to y, z_n \to z$ for distinct x, y, z, and $g_n(x_n, y_n, z_n) \to (a, b, c)$. Then at least two of a, b, c are the same.*

Proof. Suppose, to the contrary, that a, b, c are distinct. We can then choose disjoint compact neighbourhoods $U_a, U_b, U_c, U_x, U_y, U_z$ of a, b, c, x, y, z respectively. Then $K = U_a \times U_b \times U_c$ and $L = U_x \times U_y \times U_z$ are compact in $\theta^3(M)$. Since $g_n(x_n, y_n, z_n) \to (a, b, c)$, the set

$$\{g : gL \cap K \ne \emptyset\}$$

is infinite. This is a contraction, as G acts properly on $\Theta^3(M)$. Thus, at least two of a, b, c are the same. □

Lemma 12.6 (Democracy 2:1)**.** *Let $g_n \in G$ such that for some $(x_n, y_n, z_n) \to (x, y, z) \in \Theta^3(M)$, we have $g_n(x_n, y_n, z_n) \to (a, b, b)$ where $a \neq b$. If $w_n \to w \neq x$, then $g_n w_n \to b$.*

In particular, g_n converges to b locally uniformly in $M \setminus x$.

Proof. Since $w \neq x$, either $(x_n, y_n, w_n) \to (x, y, w) \in \Theta^3(M)$ or $(x_n, z_n, w_n) \to (x, z, w) \in \Theta^3(M)$. In both cases, we can apply Lemma 12.5 and see that $g_n w_n \to \{a, b\}$. If $g_n w_n \to b$, then the proof is done. We assume that $g_n w_n \to a$ and deduce a contradcition.

Since $|M| \geq 3$, there exists $c \in M$ such that $c \notin \{a, b\}$. Consider $u_n = g_n^{-1}(c)$. Since M is compact, assume that $u_n \to u$, up to a passenage of a subsequence of u_n. Using Lemma 12.5 with $g_n(x_n, y_n, u_n) \to (a, b, c)$, we have $u \in \{x, y\}$. Similarly, $g_n(x_n, z_n, u_n) \to (a, b, c)$ implies $u \in \{x, z\}$. Thus, $u = x$.

So $(u_n, w_n) \to (x, w)$ and $(g_n u_n, g_n w_n) \to (c, a)$. Since $(y_n, z_n) \to (y, z)$ and $(g_n y_n, g_n z_n) \to (b, b)$, we would obtain contradiction with Lemma 12.5 for $y \neq z$, and a, b, c are distinct. Therefore, we proved that $g_n w_n \to b$.

It remains to show that g_n converges to b locally uniformly in $M \setminus x$. Let K be a compact set in $M \setminus x$. Suppose that g_n does not converge uniformly to b in K. We put a metric d on M. Then there exist $\epsilon > 0$, a sequence $w_i \in K$ and a subsequence g_{ni} such that $d(g_{ni}(w_i), b) \geq \epsilon$. That is to say, $g_{ni}(w_i)$ does not converge to b. Without loss of generality, assume that $w_i \to w \in K$. Then this gives a contradiction. Thus, g_n converges to b over K. □

The second and third paragraph of the above proof proves the following. A

Corollary 12.7. *Assume that for $(x_n, y_n) \to (x, y)$ for $x \neq y$, $g_n(x_n, y_n) \to (a, a)$ and for $(u_n, v_n) \to (u, v)$, $g_n(u_n, v_n) \to (b, b)$. If $\{x, y\} \cap \{u, v\} = \emptyset$, then $a = b$.*

Proof. Since $|M| \geq 3$, take $c \notin \{a, b\}$ and consider $w_n = g_n^{-1}(c)$. Apply Lemma 12.5 to $(\{x_n, y_n\}, \{u_n, v_n\}, w_n)$ for the contradiction. □

Remark. Fix $\{g_n\}$. A point a is *voted* by a pair of distinct points (x, y) for $x \neq y$ if for some $(x_n, y_n) \to (x, y)$ we have $g_n(x_n, y_n) \to (a, a)$. This corollary could be reformulated as follows: if a point is voted by two pairs of distinct points (x, y) and (u, v), then $\{x, y\} \cap \{u, v\} \neq \emptyset$. Equivalently, a point cannot be voted by two disjoint pairs of points.

Lemma 12.8 (Democracy 3:0). *Let $g_n \in G$ such that for some $(x_n, y_n, z_n) \to (x, y, z) \in \Theta^3(M)$, we have $g_n(x_n, y_n, z_n) \to (b, b, b)$. If $g_n^{-1} c \to a$ for some $c \neq b \in M$, then g_n converges to b locally uniformly in $M \setminus a$.*

Proof. Apply Lemma 12.6 to $(g_n^{-1} c, u_n, v_n) \to (c, u, v) \in \Theta^3(M)$, where u_n, v_n are chosen from (x_n, y_n, z_n). □

We are now ready to give the proof of Theorem 12.4.

Proof of Theorem 12.4. Since M is compact and contains at least 3 points, there exists a subsequence g_{ni} and $(x, y, z) \in \Theta^3(M)$ such that $g_{ni}(x, y, z) \to (a, b, c)$.

By Lemma 12.5, without loss of generality, at least two of $\{a, b, c\}$ are the same. We are thus lead to consider the following two cases:

Case 1. $a \neq b = c$. By Lemma 12.6, g_{ni} converges to b locally uniformly in $M \setminus x$.

Case 2. $a = b = c$. Choose $d \neq b$ and pass to a subsequence g_{ni} so that the sequence $u_{ni} = g_{ni}^{-1} d \to x \in M$. By Lemma 12.8, g_{ni} converges to b locally uniformly in $M \setminus x$.

Renaming $a := x$ finishes the proof. □

Exercise 12.9. *Prove the converse of Theorem 12.4 holds: If any infinite set $\{g_n : n \in \mathbb{N}\}$ in G has the convergence property, then G is a convergence group action.*

Exercise 12.10. *Let G be acting on a compact metrizable space M as a convergence group. Then any infinite set $\{g_n : n \in \mathbb{N}\}$ in G contains a subsequence g_{ni} and points $a, b \in M$ so that*

(1) *g_{ni} converges to b locally uniformly in $M \setminus a$, and*
(2) *g_{ni}^{-1} converges to a locally uniformly in $M \setminus b$.*

12.2. **Proper actions on hyperbolic spaces induce convergence actions.** Let (X, d) be a proper length hyperbolic space. For any fixed point $o \in X$, the Gromov boundary $X_{o,\infty}$ gives a compact topology to the set of asymptotic classes of geodesic rays originating at o.

Moreover, we can put a visual metric ρ_o on $X_{o,\infty}$. It should be noted that the topology of $X_{o,\infty}$ does not depend on the choice of $o \in X$ (cf. Exercise 8.7), but the visual metric ρ_o does depend on o. Recall that for any two $x, y \in X_\infty$,

$$\rho_o(x, y) \asymp \lambda^{d(o,[x,y])},$$

for some $0 < \lambda < 1$.

We consider a group action of G on X by isometries. By Theorem 8.19, every isometry induces a homeomorphism of X_∞. So we have the following.

Lemma 12.11. *Assume that G acts by isometries on a proper length hyperbolic space (X, d). Then G acts by homeomorphisms on X_∞.*

Theorem 12.12. *Assume that G acts properly on a proper length hyperbolic space (X, d). If X_∞ contains at least 3 points, then G acts on X_∞ as a convergence group.*

Proof. Denote $M = X_\infty$. Fix a basepoint $o \in X$. We consider the visual metric ρ_o on M. Let K be a compact set in $\Theta^3(M)$. Define

$$\epsilon = \inf_{(x,y,z) \in K} \{\min\{\rho_o(x, y), \rho_o(y, z), \rho_o(z, x)\}\}.$$

Since K is compact, we have $\epsilon > 0$.

Recall that for visual metric, we have

$$\rho_o(x, y) \asymp \lambda^{d(o,[x,y])}$$

for some $0 < \lambda < 1$. Thus, there exists some uniform constant $L = L(\epsilon)$ such that

$$d(o, [x, y]), d(o, [y, z]), d(o, [z, x]) \leq L$$

for any $(x, y, z) \in K$. In other words, o is a L-center for any geodesic triangle $\Delta(x, y, z)$ with $(x, y, z) \in K$. Clearly, go is a L-center for any geodesic triangle $\Delta(gx, gy, gz)$ with $(x, y, z) \in K$.

Recall that for any $L > 0$ there exists $D = D(L)$ such that any two L-centers of a geodesic triangle $\Delta(x, y, z)$ have bounded distance by D. Here $x, y, x \in X \cup X_\infty$.

We now argue by way of contradiction to prove that G acts properly on $\theta^3(M)$. Assume that there exists infinitely many g_n such that $g_n K \cap K \neq \emptyset$. Thus, we can choose $(x_n, y_n, z_n) \in K$ and $(g_n x_n, g_n y_n, g_n z_n) \in K$. By the above argument, o is a L-center for all $\Delta(x_n, y_n, z_n)$. It follows that $g_n o$ is a L-center for all $\Delta(g_n x_n, g_n y_n, g_n z_n)$.

Note that o is also a L-center for every $\Delta(g_n x_n, g_n y_n, g_n z_n)$, as $(g_n x_n, g_n y_n, g_n z_n) \in K$. Thus, $d(o, g_n o) \le D$ for all n. This is a contradiction, as G acts properly on X. The proof is complete. $\square$

12.3. **Dynamics of hyperbolic elements.** We prove that a hyperbolic element g has the following simple dynamics on M. Denote by a, b the two distinct fixed points of g in M. We shall call $g_+ := b$ the *attracting* point of g if there exists a sequence of elements g_n in $\{g^n : n \ge 0\}$ such that $g_n x \to b$ for some $x \in M \setminus \{a, b\}$. The other fixed point $g_- := a$ is called *repelling*.

Lemma 12.13 (Repelling-Attracting dynamics)**.** *Let g be a hyperbolic element in a convergence group G with two fixed points $\{a, b\}$. Assume that b is an attracting fixed point.*

Then there exists a sequence of elements g_n in $\{g^n : n \ge 0\}$ such that for any $x_n \to x \in M \setminus \{a\}$, we have $g_n x_n \to b$. In other words, g_n converges to b locally uniformly in $M \setminus a$.

Proof. Since b is an attracting fixed point, there exists a sequence of elements g_n in $\{g^n : n \ge 0\}$ such that $g_n y \to b$ for some $y \in X \setminus \{a, b\}$. We apply Lemma 12.6 to $g_n(a, y, b) \to (a, b, b)$, and conclude that for $x_n \to x \in M \setminus \{a\}$, we have $g_n x_n \to b$. $\square$

A corollary is that in the definition of attracting points, one can replace "some" by "any": $g_n x \to b$ for any $x \in M \setminus \{a, b\}$.

We now note the following fact saying that the role of repelling, attracting points are reversed by taking the inverse of g.

Lemma 12.14. *If g_-, g_+ are repelling, attracting points of a hyperbolic element g. Then g_+, g_- are repelling, attracting points of a hyperbolic element g^{-1}.*

Proof. Let $x \in M \setminus \{g_-, g_+\}$. Then by compactness of M, there exists a subsequence of elements g^{-n} for $n > 0$ such that $g^{-n} x \to y$. By Lemma 12.5, $y \in \{g_-, g_+\}$. If $y = g_-$, then g_- is a contracting point for g^{-1} by definition.

We now consider $y = g_+$ and then $g^{-n} x \to g_+$. We apply Lemma 12.13, we see that $x = g^n(g^{-n} x) \to g_+$. This is a contradiction, completing the proof. $\square$

12.4. **Tits alternative in hyperbolic groups.** We apply the general theory of a convergence group action to a hyperbolic group on its Gromov boundary.

Recall that a finitely generated group G is hyperbolic if the Cayley graph $\mathscr{G}(G, S)$ is hyperbolic for some finite generating set S.

For different generating sets of G, the corresponding Gromov boundaries are homeomorphic (cf.Theorem 8.19). Thus, Gromov boundary ∂G of a hyperbolic group G can be defined to be that of the Cayley graph $\mathscr{G}(G, S)$.

Since G acts freely on $\mathscr{G}(G, S)$, by Theorem 12.12, G acts as a convergence group on its Gromov boundary.

In a hyperbolic group, an infinite order element is called hyperbolic by definition. The following lemma says that infinite order elements are hyperbolic with respect to the convergence action on the Gromov boundary. Therefore, two notions of hyperbolic elements are identical in hyperbolic groups.

Lemma 12.15. *Let g be a hyperbolic element (by definition which is infinite order element) in a hyperbolic group G. Then g has exactly two fixed points.*

In other words, with respect to the convergence action on Gromov boundary, there exists no parabolic elements in a hyperbolic group.

Proof. By Lemma 11.14, $n \in \mathbb{Z} \to g^n \in G$ is a (λ, c)-quasi-isometric map. We connect g^n, g^{n+1} for each $n \in \mathbb{Z}$ and get a quasi-geodesic path p. See lemma 7.18.

By the construction, there exists $R > 0$ in such that $p \subset N_R(g^n p)$ for any $n \in \mathbb{Z}$. We shall construct a bi-infinite geodesic q which lies in an uniform neighbourhood of p. Then the two directions of q defines two points in ∂G, which are fixed by g.

We fix a basepoint $o \in p$. Choose $x_n \in [o, p_-]_p, y_n \in [o, p_+]_p$ such that $d(o, x_n) = d(o, y_n) = n$ for any $n > 0$. We connect x_n, y_n be a geodesic q_n. By Lemma 7.13,

$$q_n \subset N_D([x_n, y_n]_p), [x_n, y_n]_p \subset N_D(q_n)$$

for an uniform constant $D = D(\lambda, c) > 0$.

Since $d(o, q_n) \leq D$, we apply Ascoli-Arzela Lamma to q_n and find a subsequence which converges to a bi-infinite geodesic q. Moreover, $p \subset N_D(q), q \subset N_D(p)$. Thus, $q \subset N_{R+D}(g^n q)$ for any $n \in \mathbb{Z}$.

Let q_-, q_+ be the two endpoints of q in ∂G. Then $g(q_-) = q_-, g(q_+) = q_+$. The proof is completed by Corollary 12.2. □

Let $g \in G$ be a hyperbolic element. By Lemma 12.13, we denote by g_-, g_+ the repelling and attracting fixed points of g.

Lemma 12.16. *Let g, h be two hyperbolic elements in G. Then either $Fix(g) \cap Fix(h) = \emptyset$ or $Fix(g) = Fix(h)$.*

Proof. Without loss of generality, assume that $g_+ = h_+$. (Other cases are similar, by taking inverses of h or g.) We shall prove that $g_- = h_-$.

We connect g^n, g^{n+1} for each $n \in \mathbb{Z}$ and get a quasi-geodesic path p_g. Similarly, we construct a quasi-geodesic path p_h for h. By the proof of Lemma 12.15, there are two geodesics q_g, q_h such that $q_g \subset N_D(p_g), q_h \subset N_D(p_h)$. Moreover, q_g and q_h end at $g_+ = h_+$. In other words q_g and q_h are asymptotic at $g_+ = h_+$.

By thin triangle property, it is easy exercise that q_g, q_h are eventually lying in an uniform 4δ-neighbourhood of each other. Now there exists a big number $N > 0$ with the following property. For any $n > N$, there exists m such that $d(g^n, h^m) \leq 2D + 4\delta$. So $d(1, g^{-n}h^m) \leq 2D + 4\delta$ for any $n \geq N$. Thus, there exist $n_1 > n_2 > N, m_1 > m_2 > 0$ such that

$$g^{-n_1}h^{m_1} = g^{-n_2}h^{m_2}.$$

It follows that $g^k = h^l$, where $k = n_2 - n_1, l = m_2 - m_1$. Since g^k, h^l have the same fixed points as g, h respectively, it follows that $\{g_-, g_+\} = \{h_-, h_+\}$. □

We are now ready to prove an analogue of Tits alternative in the setting of hyperbolic groups. Recall that a group is *elementary* if it is finite or contains a finite index cyclic group.

Theorem 12.17. *Let H be a subgroup in a hyperbolic group G. Then either H is elementary or H contains a free subgroup of rank at least 2.*

We present two different proofs. The first proof generalizes to any convergence actions by a ping-pong argument.

Proof 1 (using convergence group property). Suppose that H is infinite. Let $h \in H$ be an infinite order element, which exists by Lemma 9.11. Let q be a geodesic between h_-, h_+. By a similar argument as in the proof of Lemma 12.15, there exists $D > 0$ such that $\langle h \rangle$ and q are lying in D-neighbourhood of each other. We are going to analyze the following two cases.

Case I: every element $g \in H$ preserves the endpoints h_-, h_+. Then gq also connects h_-, h_+. Thus, gp, p are contained in a 4δ-neighborhood of each other. It can proven that $g\langle h \rangle = g'\langle h \rangle$ for some g' with $d(1, g') \le D + 4\delta$. This implies that $[H : \langle h \rangle] < \infty$.

Case II: there exists $g \in H$ which does not preserves h_-, h_+. Since ghg^{-1} is a hyperbolic element with fixed points gh_-, gh_+, it follows by Lemma 12.16 that $g\{h_-, h_+\} \cap \{h_-, h_+\} = \emptyset$.

Choose U_-, U_+, V_-, V_+ be disjoint neighbourhoods of h_-, h_+, gh_-, gh_+ respectively. By Lemma 12.13, there is a subsequence h^{ni} of h^n such that $h^{ni} \to h_+$ locally uniformly in $\partial G \setminus \{h_-\}$. Hence, there exist $n, m > 0$ such that

$$a(M \setminus U_-) \subset U_+, a^{-1}(M \setminus U_+) \subset U_-,$$

and

$$b(M \setminus V_-) \subset V_+, b^{-1}(M \setminus V_+) \subset V_-,$$

where $a = h^n, b = gh^m g^{-1}$. Since U_-, U_+, V_-, V_+ are disjoint, we see that

$$a(X) \subset U_+, X = U_+, V_-, V_+;\ a^{-1}(X) \subset U_-, X = U_-, V_-, V_+,$$

and

$$b(X) \subset V_+, X = V_+, U_-, U_+;\ b^{-1}(X) \subset V_-, X = V_-, U_-, U_+.$$

By Ping-Pong Lemma 2.13, we see that $\langle a, b \rangle$ is a free group. □

Proof 2 (using local-global quasi-geodesics). The second proof proceeds similarly as the first proof, but differs at the argument in the Case II proving that $\langle h^m, gh^m g^{-1} \rangle$ is a free group of rank 2 for $m \gg 0$.

Fix $m > 0$ and Denote $T_m := \{a^m, b^m, a^{-m}, b^{-m}\}$. The goal is to prove that the natural map $\pi : \mathbb{F}(T_m) \to \langle T_m \rangle < G$ is injective for $m \gg 0$. Let $W = t_1 t_2 \cdots t_n$ be a reduced non-empty word over T_m where $t_i \in T_m$. Let γ_W be the path labeled by W in the Cayley graph $\mathscr{G}(G, S)$.

Claim. *There exist $\lambda, c, m_0 > 0$ such that for any $m > m_0$, γ_W is (λ, c)-quasi-geodesic.*

Proof of the Claim. Define

$$L := \min\{d(1, t) : t \in T_m\}$$

It is clear that $L \to \infty$ as $m \to \infty$. To prove quasi-geodesicity of γ_W, it suffices to prove that γ_W is a L-local (γ', c')-quasi-geodesic for some $\lambda', c' > 0$.

Let $\alpha := \{h^n : n \in \mathbb{Z}\}, \beta := \{gh^n g^{-1} : n \in \mathbb{Z}\}$ be the quasi-axis of h, ghg^{-1} respectively. To verify the local quasi-geodesicity, we are reduced to consider quasi-geodesicity of four paths labeled by $a^{\pm m} b^{\pm n}$.

Note that $a^m, a^{-m} \in \alpha$ and $b^m, b^{-m} \in \beta$. Since quasi-geodesics α, β have disjoint endpoints at the Gromov boundary, there exists $D > 0$ depending on quasigeodesicity constant of α, β such that

$$diam(\mathrm{Pr}_\alpha(\beta)) \le D, diam(\mathrm{Pr}_\beta(\alpha)) \le D.$$

It is easy exercise that there are λ', c' depending on D such that $a^{\pm m} b^{\pm n}$ labels a (γ', c')-quasi-geodesic for any $n, m > 0$.

Therefore, by taking sufficiently large m and so L, γ_W is a L-local (γ', c')-quasi-geodesic. By Lemma 9.3, γ_W is a (γ, c)-quasi-geodesic. □

By the claim, the injectivity of π follows: indeed, if $\pi(W) = 1$ in G, then the endpoints of γ_W are the identity so γ_W has length at most c. However, γ_W contains at least one geodesic labeled by a letter in T_m so $Len(\gamma_W) \ge L$. Choose m large enough such that $L > c$. We thus obtain a contradiction. Hence, $\langle T_m \rangle$ is a free group of rank 2. □

13. Subgroups in convergence groups

All the results established in previous section are special cases of the general theory of subgroups in convergence group action.

13.1. **South-North dynamics of hyperbolic elements.**

Lemma 13.1. *A hyperbolic element h acts properly and co-compactly outside its fixed points.*

Proof. Let U_-, U_+ be disjoint open neighborhoods of h_-, h_+ respectively. By the property of a convergence group, there exists n such that $h^{ni}U_+ \subset U_+$ for any $i \ge 1$. Similarly, there exists m such that $h^{-mi}U_- \subset U_-$ for any $i \ge 1$. By taking mn, we can say that there exists n such that $h^{ni}U_+ \subset U_+$ for $h^{-ni}U_- \subset U_-$ for any $i \ge 1$.

For simplicity, up to take a finite power of h, we can assume that $h^iU_+ \subset U_+$ and $h^{-i}U_- \subset U_-$ for any $i \ge 1$.

Define $\dot{U}_+ = (X_\infty \setminus h\bar{U}_-) \cap U_+, \dot{U}_- = (X_\infty \setminus h^{-1}\bar{U}_+) \cap U_-$, which are open sets. Then we claim that $X_\infty \setminus \{\dot{U}_- \cup \dot{U}_+\}$ is a compact fundamental domain of $\langle h \rangle$.

First, $h\dot{U}_+ \subset \dot{U}_+$, since $hU_+ \subset U_+$ and $h\bar{U}_-^c \subset \bar{U}_-^c$. Similarly, $h^{-1}\dot{U}_- \subset \dot{U}_-$.

Secondly, $h\dot{U}_- \cap \dot{U}_+ = \emptyset$, since $h\dot{U}_- \cap \dot{U}_+ \subset hU_- \cap \dot{U}_+ = \emptyset$. Similarly, $h^{-1}\dot{U}_+ \cap \dot{U}_- = \emptyset$.

Consequently, we obtain the following by using these two observations:

$$\cap\{h^i(\dot{U}_+ \cup \dot{U}_-) : i \in \mathbb{Z}\} = \{h^i\dot{U}_+ : i \in \mathbb{Z}\} \cap \{h^i\dot{U}_- : i \in \mathbb{Z}\}.$$

By convergence property, it follows that

$$\cap\{h^i\dot{U}_+ : i \in \mathbb{Z}\} = h_+, \cap\{h^i\dot{U}_- : i \in \mathbb{Z}\} = h_-.$$

The proof is complete. □

Corollary 13.2. *The stabilizer of the two fixed points of a hyperbolic element $h \in G$ contains $\langle h \rangle$ as a finite index subgroup.*

The following is a generalization of Lemma 12.16 to convergence group actions.

Lemma 13.3. *Let h be a hyperbolic element and p any element in G. If $p(h_-) = h_-$, then $p(h_+) = h_+$.*

In particular, a hyperbolic element can not share only one fixed point with a parabolic or hyperbolic element.

Proof. Note that $p(\partial G \setminus \{h_-, h_+\}) = \partial G \setminus \{h_-, p(h_+)\}$ and $p(h_+) \ne h_-$, for h is a homeomorphism on ∂G.

Assume to the contrary that $p(h_+) \ne h_+$. Then $z := p^{-1}(h_+) \in \partial G \setminus \{h_-, h_+\}$. Thus, $p(\partial G \setminus \{h_-, z\}) = \partial G \setminus \{h_-, h_+\}$.

For any n, we have $h^{-n}(z) \in \partial G \setminus \{h_-, h_+\}$ and $z \ne h^{-n}(z)$. The equality $p(\partial G \setminus \{h_-, z\}) = \partial G \setminus \{h_-, h_+\}$ implies $ph^{-n}(z) \in \partial G \setminus \{h_-, h_+\}$. Let K be a compact fundamental domain for the action of $\langle h \rangle$ on $\partial G \setminus \{h_-, h_+\}$. Then there exists m such that $h^m ph^{-n}(z) \in K$. Note that $h^m ph^{-n}(h_-) = h_-$ and $h^m ph^{-n}(h_+) = h^m(ph_+) \to h_+$ for $p(h_+) \ne h_-$.

Noting that $h_\pm \notin K$, the convergence property for $h^m ph^{-n}$ shows that $\{h^m ph^{-n}\}$ is a finite set. That is $h^r p = ph^t$. Then $ph^t(z) = h^r p(z) = h^r(h_+) = h_+ = p(z)$. Thus, $h^t(z) = z$. This is a contradiction, as h has only two fixed points h_-, h_+. □

13.2. **Producing loxodromic elements.** We now give a helpful criteria to decide whether an element is loxodromic.

Lemma 13.4. [1, Lemma 2.4] *Let $f \in G$ be an element of infinite order. If there exists an open subset $U \subset M$ such that $f\overline{U} \subset U$, then f is loxodromic.*

Proof. By the assumption, we have $f^{n+1}(\overline{U}) \subset f^n(\overline{U})$ for $n \geq 1$. Let $K = \bigcap_{0 \leq n \leq \infty} f^n(\overline{U})$. Since M is compact, we obtain that K is nonempty. Moreover, it is easy to see that $K = fK$.

On the other hand, let $V := M \setminus \overline{U}$. Then we have $f^{-1}\overline{V} \subset V$. Let $L = \bigcap_{0 \leq n \leq \infty} f^{-n}(\overline{V})$. Arguing as above, we have L is nonempty and is f-invariant.

Note that K and L are disjoint closed f-invariant sets. By the convergence property, we obtain that K and L consist of a single point, i.e., that K and L are the fixed points of f. □

The following lemma is a direct corollary of Lemma 13.4.

Lemma 13.5. [1, Lemma 2.5] *Let $\{g_n\}$ be a subsequence of G converging to b locally uniformly on $M \setminus \{a\}$. If $a \neq b$, then g_n is loxodromic for all sufficiently large n.*

Using Lemma 13.4, one can produce many loxodromic elements.

Lemma 13.6. *Suppose $f \in G$ is a loxodromic element with distinct fixed points $p, q \in M$. Let $g \in \Gamma$ be an element which does not keep invariant the set $\{p, q\}$. Then $f^n g$ or $f^{-n} g$ are loxodromic for all sufficiently large n. Moreover, p and q are not the fixed points of $f^n g$ and $f^{-n} g$.*

Proof. Since g does not preserve the set $\{p, q\}$, without loss of generality, suppose that $z = g(p) \notin \{p, q\}$. Assuming further that p is the attracting point of the sequence $\{f^n\}$. Otherwise, up to a change of notations, we replace f^n by f^{-n}.

As $z \notin \{p, q\}$, we can take a small open neighborhood U_p of p such that $q \notin U_p$ and $p, q \notin g\overline{U}_p$. Then the convergence property of $\{f^n\}$ gives the inclusion $f^n g \overline{U}_p \subset U_p$ for all sufficiently large n. By Lemma 13.4, this implies that $f^n g$ is loxodromic. □

13.3. **Limit sets of subgroups.** In the sequel, we always assume that G is a convergence group acting on M. In this setting, we introduce two equivalent definitions of a limit set for a subgroup. The notation ΘM denotes the space of distinct ordered triples of points of M.

Definition 13.7 (LS1)**.** The *limit set* $\Lambda_1(\Gamma)$ of a subgroup $\Gamma \subset G$ is the set of accumulation points of all Γ-orbits in M.

The large diagonal $\triangle M := M^3 \setminus \Theta^3 M$ is the set of triples with at least two same entries, i.e., $(x, x, y) \in \triangle M$. By sending each triple $(x, x, y) \in \triangle M$ to $x \in M$, we give the union $\Theta^3 M \sqcup M$ the quotient topology of the product topology M^3. Thus $\Theta^3 M \sqcup M$ is compact, and the subspace topology on M coincides with its original compact topology.

The following lemma says that a convergence group also has the convergence property on the compact space $\Theta^3 M \sqcup M$.

Lemma 13.8. [1, Proposition 1.8] *If G has the convergence property, then so does G on $\Theta^3 M \sqcup M$: for any sequence $\{g_n\}$ of G, there exists a subsequence*

$\{g_{n_i}\}$ and two points $a, b \in M$ such that g_{n_i} converges to b locally uniformly on $(\Theta^3 M \sqcup M) \setminus \{a\}$.

As above, M can be thought of as a boundary to compactify the triple space $\Theta^3 M$. By Definition 12.1, a convergence group acts properly discontinuously on $\Theta^3 M$. Inspired by a definition of limit sets in Kleinian groups, we give another definition of limit set in convergence groups as follows.

Definition 13.9 (LS2)**.** The *limit set* $\Lambda_2(\Gamma)$ of a subgroup $\Gamma \subset G$ is the set of accumulation points of a (or any) Γ-orbit in $\Theta^3 M$.

Remark. It is easily seen that the qualifier "any" in the (LS2) definition is justified by Lemma 13.8. Indeed, if z is the limit point of a sequence $g_n x$ for some $x \in \Theta^3 M$, by Lemma 13.8, (some subsequence of) g_n converges to z locally uniformly on $(\Theta^3 M \sqcup M) \setminus \{a\}$ for some $a \in M$. Hence,we have $g_n y \to z$ for any $y \in \Theta^3 M$. Thus, the limit set $\Lambda_2(\Gamma)$ is independent of the choice of the orbit in $\Theta^3 M$.

The following equivalence follows directly from the definitions and the above remark. The proof is left as exercise (using similar arguments as in the remark).

Lemma 13.10. *Let Γ be a subgroup of a convergence group G. Then $\Lambda_1(\Gamma) = \Lambda_2(\Gamma)$.*

Exercise 13.11. *Give a proof of Lemma 13.10.*

In future, we will use $\Lambda(\Gamma)$ to denote the limit set of Γ.

13.4. **Conical points and parabolic points.**

Definition 13.12. A point $z \in M$ is called *conical* if there exists $g_n \in G$ and $(a, b) \in M^2$ with $a \neq b$ such that

$$g_n(z, w) \to (a, b)$$

for any $w \neq z \in M$.

Corollary 13.13. *A conical point $z \in \Lambda G$ is a limit point of G: it is an accumulation point of some G-orbit.*

Proof. By the convergence property, there exists a subsequence of g_n still denoted by g_n such that

$$g_n \to x$$

locally uniformly on $M \setminus y$. Thus, $x = b$ and $y = z$. By Exercise 12.10, we have $g_n^{-1} \to z$ locally uniformly on $M \setminus b$. Thus, z is a limit point of G. □

The following result generalizes Lemma 13.3: the fixed points of a hyperbolic hyperbolic are conical.

Proposition 13.14. [1, Proposition 3.2] *In a convergence group, a conical point cannot be fixed by a parabolic point.*

Proof. Suppose that z is a conical point. Thus, there exists $g_n \in G$ and (a, b) for $a \neq b \in M$ such that $g_n(z, w) \to (a, b)$ for any $w \in M \setminus z$. Consequently, By Lemma 12.6, $g_n \to b$ on $M \setminus z$ and $g_n^{-1} \to z$ on $M \setminus b$.

Fix n so that $g_n^{-1} b \neq z$ (passing to subsequence of g_n, this holds for any n). Thus, $g_m^{-1} g_n \to z$ on $M \setminus g_n^{-1} b$. By Lemma 13.5, $g_m^{-1} g_n$ is hyperbolic for sufficiently large m.

Let $z = p(z)$ be a fixed point of a parabolic element p. Consider $\gamma_n := g_n p g_n^{-1}$. Then $\{\gamma_n\}$ is an infinite set. Indeed, if it was finite, we would see that $g_m^{-1} g_n p = g_m^{-1} g_n p$ for $m \neq n$: a parabolic element is commute with a loxodromic element. This is impossible by Lemma 13.3.

Therefore, γ_n contains a collapsing sequence still denoted by γ_n. Now $g_n(z) \to a$ and $\gamma_n g_n(z) = g_n p(z) = g_n(z) \to a$. And $g_n(w) \to b$ and $\gamma_n g_n(w) = g_n p(w) \to b$ for $p(w) \neq z$. By the convergence property of γ_n (cf. Lemma 12.6), we conclude that either $\gamma_n \to a$ on $M \setminus b$ or $\gamma_n \to b$ on $M \setminus a$. By Lemma 13.5, γ_n is a loxodromic element. However, $\gamma_n = g_n p g_n^{-1}$ is conjugate to a parabolic element p. This is a contradiction. The proof is complete. □

13.5. **Properties of limit sets.** In this section, we aim to give some further characterization of limit sets of non-elementary subgroups. Note that a subgroup of a convergence group is *non-elementary* if its limit set contains more than two points.

Lemma 13.15. *If $|\Lambda(\Gamma)| \geq 2$, then $\Lambda(\Gamma)$ is the minimal one among the family of Γ-invariant closed subsets in M of cardinality at least two.*

Proof. It is easy to see that the limit set $\Lambda(\Gamma)$ is a Γ-invariant closed subset. So it suffices to show that for any Γ-invariant closed subset $N \subset M$ with $|N| \geq 2$, we have $\Lambda(\Gamma) \subset N$.

Let x, y be distinct points in N. We shall show that any point $p \in \Lambda(\Gamma)$ is an accumulation point of the orbit Γx or Γy.

Since $p \in \Lambda(\Gamma)$, there exists a sequence $\{\gamma_n\}$ and a point $z \in M$ such that $\gamma_n(z) \to p$. By the convergence property, we assume that γ_n converges to b locally compactly on $M \setminus \{a\}$ for two points $a, b \in M$.

If $b = p$, then we are done with a choice of $\{x, y\}$, say $x \in M \setminus \{a\}$, such that $\gamma_n(x) \to p$. In other words, p is an accumulation point of the orbit Γx.

If $b \neq p$, then we must have $p = a$, as $\gamma_n(z) \to p$. By the convergence property, the inverse γ_n^{-1} converges to a locally compactly on $M \setminus \{b\}$. Hence we are reduced to a similar case as "$b = p$". Therefore, we obtain that every limit point of Γ is an accumulation point of some Γ-orbit in N. The proof is complete. □

If we consider a non-elementary subgroup, the cardinality restriction in Lemma 13.15 will not be necessary.

The following result characterizes the limit set as the minimal, group invariant, closed subset.

Theorem 13.16. *If $|\Lambda(\Gamma)| > 2$, then $\Lambda(\Gamma)$ is the minimal Γ-invariant closed subset in M. In particular, Γ acts properly discontinuous on the complement $M \setminus \Lambda\Gamma$.*

In fact, Theorem 13.16 is a direct consequence of Lemma 13.15 and the following Lemma 13.17.

Lemma 13.17. *If $|\Lambda(\Gamma)| > 2$, then Γ can not fix a point in M.*

Proof of Lemma 13.17. Suppose, to the contrary, that Γ fixes a point $p \in M$. By Proposition 13.14 or Lemma 13.3, a loxodromic element can not share only one fixed point with a parabolic or loxodromic element. Then Γ can not contain both loxodromic and parabolic elements.

Case I. If Γ contains a loxodromic element f, then all loxodromic elements in Γ also fix the two fixed point of f. Moreover, the other elliptic elements in Γ also keep invariant these two points. Otherwise, by Lemma 13.6, we can produce loxodromic elements which do not fix the fixed points of f. Hence Γ keeps invariant the two fixed point of f. Then it follows that that $\langle f\rangle$ is a finite index subgroup in Γ. This proves that $\Lambda(\Gamma)$ consists of two points, a contradiction with the assumption that $|\Lambda(\Gamma)| > 2$. Hence there are no loxodromic elements in Γ.

Case II. We are now able to assume that Γ consists of parabolic and elliptic elements. Let $q \in \Lambda(\Gamma) \setminus \{p\}$. Then there exists a sequence γ_n and a point $z \in M$ such that $\gamma_n(z)$ converges to the limit point q. By the convergence property, we assume that $\gamma_n \to z$ locally uniformly on $M \setminus \{w\}$. Thus we have two cases as follows.

First if $z = q$, then $w = p$, as p is fixed by Γ. Since $p \neq q$, by Lemma 13.5, we obtain that γ_n have to be loxodromic elements for all sufficiently large n.

Otherwise, we assume $z \neq q$. It follows that $w = q$. By the convergence property, we have the inverse $\gamma_n^{-1} \to w$ locally uniformly on $M \setminus \{z\}$. As the first case, we have that γ_n are loxodromic elements for all sufficiently large n.

Summarizing the above two cases, we got a contradiction with the just proved fact that Γ can not contain loxodromic elements. Therefore, we conclude that Γ could not fix a point in M. □

We now draw some corollaries to Theorem 13.16, which describe several well-known properties on the limit sets of non-elementary subgroups.

Note that the set of accumulation points of a Γ-orbit in M gives a Γ-invariant closed set. So for a non-elementary subgroup, the (LS1) defintion of its limit set can be strengthened as the following form.

Corollary 13.18. *If* $|\Lambda(\Gamma)| > 2$*, then* $\Lambda(\Gamma)$ *is the set of accumulation points of a (or any)* Γ*-orbit in* M*.*

By Corollary 13.18, any Γ-orbit of a limit point is dense in the limit set $\Lambda(\Gamma)$. This shows that limit sets are perfect.

Corollary 13.19. *If* $|\Lambda(\Gamma)| > 2$*, then* $\Lambda(\Gamma)$ *has no isolated point, i.e. it is a perfect set and contains uncountable many points.*

If $\Lambda(\Gamma) \neq M$, then we take a Γ-orbit of a point in $M \setminus \Lambda(\Gamma)$. Then the set of accumulation points of this orbit is the limit set of Γ. This proves the following.

Corollary 13.20. *If* $|\Lambda(\Gamma)| > 2$ *and* $\Lambda(\Gamma) \neq M$*, then* $\Lambda(\Gamma)$ *has no interior points, i.e. it is nowhere dense in* M*.*

13.6. Classification of subgroups in convergence groups.

Theorem 13.21. *Let* H *be a subgroup of a convergence group* G *acting on* M*. If* $|\Lambda(H)| < \infty$*, then exactly one of the following cases holds:*

(1) *H is finite,*
(2) *H consists of elliptic and parabolic elements. In this case, $\Lambda(H)$ is one point and H may be a torsion group.*
(3) *H contains a finite index subgroup generated by a hyperbolic element h. In this case $\Lambda(H)$ consists of two points.*

If $|\Lambda(H)| \geq 3$*, then* H *contains a free subgroup of rank 2.*

Proof. By the convergence property, $|\Lambda(H)| \leq 2$.

Suppose that H is not a finite group. The proof of Lemma 13.17 shows the second and third assertions. It remains to consider $|\Lambda(H)| > 2$.

If $|\Lambda(H)| \geq 2$, then there exist two loxodromic elements h, k such that their fixed points are disjoint. Using ping-pong lemma as in the proof of Theorem 12.17, we see that $\langle h^m, k^n \rangle$ is a free group of rank 2 for $m, n \gg 0$. $\square$

14. Introduction to relatively hyperbolic groups

There are various equivalent formulations of relative hyperbolicity. We start with the definition of geometrically finite actions on hyperbolic spaces.

14.1. **Cusp-uniform and geometrically finite actions.** Assume that G admits a proper action on a proper hyperbolic space (X, d). Let ΛG be the limit set of the induced convergence action of G on the Gromov boundary of X (Definition 13.7). Equivalently, ΛG also concides the set of the accumulation points of Gx for any $x \in X$.

Definition 14.1 (Bounded parabolic point)**.** A point $p \in M$ in a convergence group is called *parabolic* if its stabilizer G_p is infinite and $p = \Lambda G_p$. If G_p acts co-compactly on $\Lambda G \setminus p$, then p is called a *bounded parabolic* point.

Remark. The fixed point of a parabolic element must be a parabolic point: its stabilizer is infinite, and $\sharp \Lambda G_p = 1$ follows Theorem 13.21. However, the stabilizer of a parabolic point may not necessarily contain a parabolic element. In fact, it may be an infinite torsion group.

Proposition 14.2 (Tukia [13])**.** *In a convergence group, a conical point cannot be a parabolic point.*

Proof. Let z be a parabolic point. If G_z contains a parabolic element then the conclusion follows from Proposition 13.14. Let us now assume that G_z is an infinite torsion group: every element is ellptic. This case is proved by Tukia [13]. □

Definition 14.3. A function $h : X \to \mathbb{R}$ is called a *coarse horofunction* at $p \in \partial X$ if there exist constants $C_1, C_2 > 0$ such that for any $x, a \in X$ with $d(a, [x, p]) \le C_1$, we have

$$|h(x) - h(a) - d(x, a)| \le C_2$$

Given a constant $K > 0$, a *K-horoball B_p* centered at p contains $h^{-1}((-\infty, c])$ and is contained in $h^{-1}((-\infty, c + K])$ for some constant c.

Remark. A coarse horofunction is a quasi-perturbution of a horofunction defined in Definition 17.4. Namely, for any coarse horofunction h, choose a geodesic ray γ ending at p such that $h(\gamma(0)) = 0$. Then

$$|h - b_\gamma|_\infty = \sup_{x \in X} |h(x) - b_\gamma(x)| < \infty$$

where $b_\gamma(x) := \lim_{t\to\infty}(d(x, \gamma(t)) - d(\gamma(0), \gamma(t)))$.

Exercise 14.4. *If X is a hyperbolic space, then a K-horoball B_p at p is σ-quasiconvex, where σ depends on K and the coarse horofunction h at p (indeed the constants C_1, C_2 there). Thus, a horoball itself is a hyperbolic space.*

Moreover, the Gromov boundary of B_p consists of only one point.

Let Π be the set of bounded parabolic points in the limit set ΛG.

Definition 14.5 (Cusp-uniform action)**.** If there exists a G-invariant family of disjoint horoballs B_p centered at $p \in \Pi$ so that the action of G on their complements $X \setminus \cup_{p\in\Pi} B_p$ is co-compact, then the action of G on X is called *geometrically finite.*

Recall the definition of conical points in Definition 13.12 is defined using convergence action. It admits the following geometric interpretation which explains the name.

Lemma 14.6. *Let q be a limit point of $G \curvearrowright X$. If there exists a sequence of elements g_n such that* $\sup\{d(g_n o, \gamma)\} < \infty$ *for every geodesic ray at q, then q is a conical point.*

Conversely, if q is a conical point, then there exists a sequence of elements g_n such that $\sup\{d(g_n o, \gamma)\} < \infty$ *for every geodesic ray at q.*

Lemma 14.7. *Assume that $G \curvearrowright X$ is geometrically finite. Then Π/G is finite and every limit point of G is either bounded parabolic or conical.*

Proof. Let p be a parabolic point and B_p the horoball at p. Since horoballs at different points are disjoint, we have $gB_p = B_p$ for any $g \in G_p$ and $gB_p \cap B_p = \emptyset$ for any $g \in G \setminus G_p$. Thus, G_p acts co-compactly on the boundary ∂B. Indeed, by assumption, G acts co-compactly on $X \setminus \cup_{p\in\Pi} B_p$. Then the natural orbital map gives a topological embedding of $\partial B_p/G_p$ as a closed subset into $X \setminus \cup_{p\in\Pi} B_p/G$. Thus, $\partial B_p/G_p$ is compact and the claim follows. As a consquence, Π/G is is a finite set.

Let q be a non-parabolic point in ΛG. Since $G \curvearrowright X \setminus \cup_{p\in\Pi} B_p$ is compact, let K be compact set K so that $GK = X \setminus \cup_{p\in\Pi} B_p$.

Let $o \in K$ be a basepoint. Consider a geodesic ray γ ending at q and issuing from o. There exists a sequence of elements g_n such that

$$d(g_n o, \gamma) \le D := diam(K) < \infty$$

This implies that q is conical. □

Definition 14.8 (Geometrically finite action)**.** If the limit set of a convergence action $G \curvearrowright M$ consists of bounded parabolic points and conical points, the action of G on M is called *geometrically finite*.

Proposition 14.9. [3, Prop 6.13] *Assume that G acts properly on a hyperbolic space X so that every limit point is either bounded parabolic or conical. Then $G \curvearrowright X$ is geometrically finite.*

Lemma 14.10. *Assume that G acts properly on a hyperbolic space X so that every limit point is either bounded parabolic or conical. Then there exists a G-invariant horoball at each bounded parabolic point.*

Let $\mathbb{H}$ denote the set of maximal parabolic subgroups of the cusp-uniform action of G on X. We also say that the pair $(G, \mathbb{H})$ is relatively hyperbolic. The collection $\mathbb{H}$ is often referred as a *peripheral structure* of G, and each element of $\mathbb{H}$ a *peripheral subgroup* of G.

The following theorem of Yaman gives a topological chracterization of relatively hperbolic groups.

Theorem 14.11 (Yaman)**.** *If a group G admits a geometrically finite convergence action on a compact metrizable space M, then there exists a hyperbolic space X and a cusp-uniform action of G on X such that the Gromov boundary is G-equivariant homemorphic to the limit set of G in M.*

14.2. **Combinatorial horoballs.** Let Γ be a connected graph, equipped with the combinatoral metric d. We define a graph $\mathcal{H}(\Gamma) = \Gamma \times \mathbb{N}_{\ge 0}$ called a *combinatorial horoball* over (horosphere) Γ so that the following holds:

(1) For $k \ge 0$, we connect (v, k) and $(v, k+1)$ by a *vertial edge* of length 1.

(2) For $k \geq 0$, we add a horozontal edge of length 1 between every pair of vertices $(u,k),(v,k) \in \Gamma \times \{k\}$ if $d(u,v) \leq 2^k$.

We consider a class of special paths in $\mathcal{H}(\Gamma)$ between any two points $x = (u,k), y = (v,l)$ as follows:

Let p_x be the vertical ray from $x = (u,k)$ consisting of vertical edges $(u,i) \leftrightarrow (u,i+1)$ where $i \geq k$. Let p_y be the vertical ray from $y = (v,l)$. There exists a minimal $m \geq k, l$ such that (u,m) is connected to (v,m) by a horozontal edge. The *special path* from x to y is then obtained by connecting $(u,m) \in p_x$ to $(v,m) \in p_y$ by one horozontal edge. The subpath of a special path is special.

By construction, it is easy to see that the triangle with special paths as sides satisfies the thin triangle property: any side is contained in the 1-neighborhood of the other two sides. With more effort (Exercise 14.13), we can show the following.

Lemma 14.12. *The graph $\mathcal{H}(\Gamma)$ is a δ-hyperbolic space where δ is a universal constant, where special paths are uniformly quasi-geodesics. The Gromov boundary of $\mathcal{H}(\Gamma)$ consists of only one point.*

Remark. Each sublevel set (i.e.: $\Gamma \times [N, \infty)$) of a combinatorial horoball is K-horoball for a uniform constant K, since the height function $(v,i) \mapsto i$ is a coarse horofunction.

If H acts by isometry on Γ, then H acts by isometry on $\mathcal{H}(\Gamma)$. In particular, any finitely generated group acts properly on the combinatorial horoballs over its Cayley graphs. However the action is "elemetrary", since the group has a global fixed point at the boundary.

Exercise 14.13. *Let X be a geodesic metric space. Assume that we can choose a family of special paths $p_{x,y}$ from x to y such that*

(1) *$p_{x,y} = p_{y,x}$ and subpaths of $p_{x,y}$ are special paths between their endpoints.*
(2) *The triangle obtained by special paths has (uniform) thin-triangle property.*

Then X is a hyperbolic space, where special paths are uniform quasi-geodesics.

Remark (on the proof). A proof can be given as follows: using the triangle property (2), first prove analogue of Lemma 7.10, where $[a,b]$ is replaced by a special path between a, b, and then use it to prove analogue of Lemma 7.13: the specical path follows travel quasi-geodesics.

14.3. **Augmented spaces.** We consider a finitely generated group G with a finite collection of finitely generated subgroups $\mathbb{H} = \{H_i\}_{i \in I}$. Fix a finite generating set S which contains the generating set of each $H \in \mathbb{H}$. For each H-coset gH, we atatch a copy $g\mathcal{H}(\Gamma)$ of a combinatorial horoball over its Cayley graph of gH. The resulted space is called *augmented space* $X(G, \mathbb{H})$.

Definition 14.14. The pair $(G, \mathbb{H})$ is called *relatively hyperbolic* if the augmented space $X(G, \mathbb{H})$ is hyperbolic.

The following result is clear by definition of cusp-uniform action.

Lemma 14.15. *If the augmented space is a hyperbolic space, then the action of G on $X(G, \mathbb{H})$ is cusp-uniform.*

15. Farb's and Osin's definition

In this susbection, we consider a countable group G with a collection of subgroups $\mathbb{H} = \{H_i\}_{i\in I}$. We refer the reader to [5] and [10] for more details, where the latter considers finitely generated group but the latter allows even non-countable groups.

15.1. **Coned-off and relative Cayley graphs.** We shall put a metric d_G on a group G, which is *proper* if any bounded set is finite, and *left invariant* if $d_G(gx_1, gx_2) = d_G(x_1, x_2)$ for any $g, x_1, x_2 \in G$. For given $g \in G$, we define the norm $|g|_{d_G}$ with respect to d_G to be the distance $d_G(1, g)$.

It is well-known that a group is countable if and only if it admits a proper left invariant metric. Indeed, any countable group could be embedded into 2-generated groups. Let d_G be some proper, left invariant metric on G.

A subset X of G is a *relative generating set* for $(G, \mathbb{H})$ if G is generated by the set $(\cup_{i\in I} H_i) \cup X$ in the traditional sense.

Definition 15.1 (Relatively Cayley graphs)**.** Fixing a relative generating set X for $(G, \mathbb{H})$, the constructed Cayley graph $\mathscr{G}(G, X \cup \mathcal{H})$ is called the *relative Cayley graph* of G with respect to $\mathbb{H}$.

The ramified version of of relative Cayley graphs is the following coned-off cayley graphs introduced by Farb.

Definition 15.2 (Coned-off Cayley graphs)**.** Fix a relative generating set X for $(G, \mathbb{H})$ and consider the Cayley graph $\mathcal{G}(G, X)$ of G with respect to X. Each left coset gH for $g \in G, H \in \mathbb{H}$ is associated to a cone point c_{gH} and a half edge from each element in gH to the cone point $c(gH)$ is added. The resulted graph $\hat{\mathcal{G}}(G, X \cup \mathcal{H})$ is called the *Coned-off Cayley graph* of G with respect to $\mathbb{H}$.

It is clear that $\mathscr{G}(G, X \cup \mathcal{H})$ is quasi-isometric to $\hat{\mathcal{G}}(G, X \cup \mathcal{H})$. From now on, we assume that $(G, \mathbb{H})$ has a finite relative generating set X.

15.2. **BCP conditions.** The following condition was introduced by Farb [5].

Definition 15.3. (Bounded coset penetration) The pair $(G, \mathbb{H})$ is said to satisfy the *bounded coset penetration property with respect to* d_G (or BCP property with respect to d_G for short) if, for any $\lambda \ge 1$, $c \ge 0$, there exists a constant $a = a(\lambda, c, d_G)$ such that the following conditions hold. Let p, q be (λ, c)-quasigeodesics without backtracking in $\mathscr{G}(G, X \cup \mathcal{H})$ such that $p_- = q_-$, $p_+ = q_+$.

1) Suppose that s is an H_i-component of p for some $H_i \in \mathbb{H}$, such that $d_G(s_-, s_+) > a$. Then there exists an H_i-component t of q such that t is connected to s.

2) Suppose that s and t are connected H_i-components of p and q respectively, for some $H_i \in \mathbb{H}$. Then $d_G(s_-, t_-) < a$ and $d_G(s_+, t_+) < a$.

The following corollary is immediate by an elementary argument.

Corollary 15.4. *BCP property of* $(G, \mathbb{H})$ *is independent of the choice of left invariant proper metrics.*

In view of Corollary 15.4, we shall not mention explicitly proper left invariant metrics when saying the BCP property of $(G, \mathbb{H})$.

Definition 15.5. (Farb Definition) A countable group G is *hyperbolic relative to* $\mathbb{H}$ *in the sense of Farb* if the Caylay graph $\mathscr{G}(G, X \cup \mathcal{H})$ is hyperbolic and the pair $(G, \mathbb{H})$ satisfies the BCP property.

15.3. **Osin's Definition.** The following terminoldge is introduced by Osin [10] in relative Cayley graphs.

Definition 15.6. Let p, q be paths in $\mathscr{G}(G, X \cup \mathcal{H})$. A subpath s of p is called an *H_i-component*, if s is the maximal subpath of p such that s is labeled by letters from H_i.

Two H_i-components s, t of p, q respectively are called *connected* if there exists a path c in $\mathscr{G}(G, X \cup \mathcal{H})$ such that $c_- = s_-, c_+ = t_-$ and c is labeled by letters from H_i. An H_i-component s of p is *isolated* if no other H_i-component of p is connected to s.

We say a path p *without backtracking* by meaning that all H_i-components of p are isolated. A vertex u of p is *nonphase* if there is an H_i-component s of p such that u is a vertex of s but $u \neq s_-, u \neq s_+$. Other vertices of p are called *phase*.

As the notion of relative generating sets, we can define in a similar fashion the relative presentations and (relative) Dehn functions of G with respect to $\mathbb{H}$. We refer the reader to [10] for precise definitions.

We now give the first definition of relative hyperbolicity due to Osin [10]. Note that the full version of Osin's definition applies to general groups without assuming the finiteness of $\mathbb{H}$.

Definition 15.7. (Osin Definition) A countable group G is *hyperbolic relative to $\mathbb{H}$ in the sense of Osin* if G is finitely presented with respect to $\mathbb{H}$ and the relative Dehn function of G with respect to $\mathbb{H}$ is linear.

The following lemma plays an important role in Osin's approach [10] to relative hyperbolicity. The finite subset Ω and constant κ below depend on the choice of finite relative presentations of G with respect to $\mathbb{H}$. In our later use of Lemma 15.8, when saying there exists κ, Ω such that the inequality (23) below holds in $\mathscr{G}(G, X \cup \mathcal{H})$, we have implicitly chosen a finite relative presentation of G with respect to $\mathbb{H}$.

Lemma 15.8. [10, Lemma 2.27] *Suppose $(G, \mathbb{H})$ is relatively hyperbolic in the sense of Osin and X is a finite relative generating set for $(G, \mathbb{H})$. Then there exists $\kappa \geq 1$ and a finite subset $\Omega \subset G$ such that the following holds. Let c be a cycle in $\mathscr{G}(G, X \cup \mathcal{H})$ with a set of isolated H_i-components $S = \{s_1, \dots, s_k\}$ of c for some $i \in I$, Then*

$$\sum_{s \in S} d_{\Omega_i}(s_-, s_+) \leq \kappa Len(c), \tag{23}$$

where $\Omega_i := \Omega \cap H_i$.

Remark. By the definition of d_{Ω_i}, if $d_{\Omega_i}(g, h) < \infty$ for $g, h \in G$, then there exists a path p labeled by letters from Ω_i in this new Cayley graph $\mathscr{G}(G, X \cup \Omega \cup \mathcal{H})$ such that $p_- = g, p_+ = h$.

Using Lemma 15.8, the following lemma can be proven exactly as Proposition 3.15 in [10]. The finite set Ω below is given by Lemma 15.8.

Lemma 15.9. [10] *Suppose $(G, \mathbb{H})$ is relatively hyperbolic in the sense of Osin and X is a finite relative generating set for $(G, \mathbb{H})$. For any $\lambda \geq 1, c \geq 0$, there exists a constant $\epsilon = \epsilon(\lambda, c) > 0$ such that the following holds. Let p, q be (λ, c)-quasigeodesics without backtracking in $\mathscr{G}(G, X \cup \mathcal{H})$ such that $p_- = q_-, p_+ = q_+$.*

Then for any phase vertex u of p(resp. q), there exists a phase vertex v of q(resp.p) such that $d_{X\cup\Omega}(u,v) < \epsilon$.

The following lemma is well-known in the theory of relatively hyperbolic groups.

Lemma 15.10. [10] *Suppose $(G,\mathbb{H})$ is relatively hyperbolic in the sense of Osin. Then the following statements hold for any $g \in G$ and $H_i, H_j \in \mathbb{H}$,*
1) If $H_i^g \cap H_i$ is infinite, then $g \in H_i$,
2) If $i \neq j$, then $H_i^g \cap H_j$ is finite.

The following lemma states that for a given finite relative generating set, we can always find a finite subset Σ such that $(G,\mathbb{H})$ satisfies BCP property with respect to d_Σ.

Lemma 15.11. *Suppose $(G,\mathbb{H})$ is relatively hyperbolic in the sense of Osin and X is a finite relative generating set for $(G,\mathbb{H})$. Then there exists a finite set $\Sigma \subset G$ such that then $(G,\mathbb{H})$ satisfies BCP property with respect to d_Σ.*

Proof. Let Ω be the finite set given by Lemma 15.8 for $\mathscr{G}(G, X\cup\mathcal{H})$. We take a new finite relative generating set $\hat{X} := X \cup \Omega$. Using Lemma 15.8 again, we obtain a finite set Σ and constant $\mu > 1$ such that the inequality (23) holds in $\mathscr{G}(G, \hat{X}\cup\mathcal{H})$.

We now verify BCP property 1). Let p, q be (λ, c)-quasigeodesics without backtracking in $\mathscr{G}(G, X\cup\mathcal{H})$. Since $\hat{X}$ is finite, the embedding $\mathscr{G}(G, X\cup\mathcal{H}) \hookrightarrow \mathscr{G}(G, \hat{X}\cup\mathcal{H})$ is a quasi-isometry. Regarded as paths in $\mathscr{G}(G, \hat{X}\cup\mathcal{H})$, p, q are (λ', c')-quasigeodesics without backtracking in $\mathscr{G}(G, \hat{X}\cup\mathcal{H})$, for some constants $\lambda' \geq 1, c' \geq 0$ depending on $\hat{X}$.

Let $\epsilon = \epsilon(\lambda, c)$ be the constant given by Lemma 15.9. Set

$$a = \mu(\lambda' + 1)(2\epsilon + 1) + c'\mu.$$

We claim that a is the desired constant for the BCP property of $(G,\mathbb{H})$. If not, we suppose there exists an H_i-component s of p such that $d_\Sigma(s_-, s_+) > a$ and no H_i-component of q is connected to s.

By Lemma 15.9, there exist phase vertices u, v of q such that $d_{X\cup\Omega}(s_-, u) < \epsilon, d_{X\cup\Omega}(s_-, v) < \epsilon$. Thus by regarding p, q as paths in $\mathscr{G}(G, \hat{X}\cup\mathcal{H})$, there exist paths l and r labeled by letters from Ω such that $l_- = e_-, l_+ = u, r_- = e_+$, and $r_+ = v$. We consider the cycle $c := er[u,v]_q^{-1}l^{-1}$ in $\mathscr{G}(G, \hat{X}\cup\mathcal{H})$, where $[u,v]_q$ denotes the subpath of q between u and v. Since $[u,v]_q$ is a (λ', c')-quasigeodesic, we compute $\mathrm{Len}(c)$ by the triangle inequality and have

$$\mathrm{Len}(c) \leq (\lambda' + 1)(2\epsilon + 1) + c'.$$

Obviously e is an isolated H_i-component of c. Using Lemma 15.8 for the cycle c in $\mathscr{G}(G, \hat{X}\cup\mathcal{H})$, we have $d_\Sigma(e_-, e_+) < \mu\mathrm{Len}(c) < a$. This is a contradiction.

Therefore, BCP property 1) is verified with respect to d_Σ. BCP property 2) can be proven in a similar way. □

We conclude this subsection with the following theorem which is proven in [10] for finitely generated relatively hyperbolic groups.

Theorem 15.12. *The pair $(G,\mathbb{H})$ is relatively hyperbolic in the sense of Farb if and only if it is relatively hyperbolic in the sense of Osin.*

Proof. By Corollary **??**, BCP property of $(G, \mathbb{H})$ follows from Lemma 15.11. The hyperbolicity of relative Cayley graph $\mathscr{G}(G, X \cup \mathcal{H})$ is proven in [10, Corollary 2.54]. Thus, $(G, \mathbb{H})$ is relatively hyperbolic in the sense of Farb.

The sufficient part is proven in the appendix of Osin [10] for finitely generated relatively hyperbolic groups. We remark that the only argument involved to use word metrics with respect to finite generating sets is in the proof of Lemma 6.12 in [10]. But Osin's argument also works for any proper left invariant metric. Hence, Osin's proof is through for the countable case. $\square$

16. Relatively quasiconvex subgroups

In this section, we shall introduce the notion of a relatively quasiconvex subgroup. This is analogous to quasiconvexity of subgroups in hyperbolic groups. The main result is that relatively quasiconvex subgroups are relatively hyperbolic with respect to induced peripheral subgroups.

Definition 16.1. [6] Suppose $(G, \mathbb{H})$ is relatively hyperbolic and d is some proper left invariant metric on G. A subgroup Γ of G is called *relatively σ-quasiconvex* with respect to $\mathbb{H}$ if there exists a constant $\sigma = \sigma(d) > 0$ such that the following condition holds. Let p be an arbitrary geodesic path in $\mathscr{G}(G, X \cup \mathcal{H})$ such that $p_-, p_+ \in \Gamma$. Then for any vertex $v \in p$, there exists a vertex $w \in \Gamma$ such that $d(u, w) < \sigma$.

Corollary 16.2. [6] *Relative quasiconvexity is independent of the choice of proper left invariant metrics.*

In fact, when proving relative quasiconvexity, we usually verify the relative quasiconvexity with respect to some partial distance function, as indicated in the following corollary.

Corollary 16.3. *Suppose $(G, \mathbb{H})$ is relatively hyperbolic and Γ is a subgroup of G. Let $\mathcal{A} \subset G$ be a finite set and $d_{\mathcal{A}}$ the partial distance function with respect to $\mathcal{A}$. If there exists a constant $\sigma = \sigma(d_{\mathcal{A}}) > 0$ such that for any geodesic p with endpoints at Γ, the vertex set of p lies in σ-neighborhood of Γ with respect to $d_{\mathcal{A}}$. Then Γ is relatively quasiconvex.*

The following result says that there exists a quasi-isometric map between a relatively quasiconvex subgroup to the ambient relatively hyperbolic group.

Lemma 16.4. *Suppose $(G, \mathbb{H})$ is relatively hyperbolic. Let $\Gamma < G$ be relatively σ-quasiconvex. Then Γ is finitely generated by a finite subset $Y \subset G$ with respect to a finite collection of subgroups*

$$\mathbb{K} = \{H_i^g \cap \Gamma : |g|_d < \sigma, i \in I, \sharp\, H_i^g \cap \Gamma = \infty\}. \tag{24}$$

Moreover, X can be chosen such that $Y \subset X$ and there is a Γ-equivariant quasi-isometric map $\iota : \mathscr{G}(\Gamma, Y \cup \mathcal{K}) \to \mathscr{G}(G, X \cup \mathcal{H})$.

Proof. The argument is inspired by the one of [10, Lemma 4.14].

For any $\gamma \in \Gamma$, we take a geodesic p in $\mathscr{G}(G, X \cup \mathcal{H})$ with endpoints 1 and γ. Suppose the length of p is n. Let $g_0 = 1, g_1, \ldots, g_n = \gamma$ be the consecutive vertices of p. By the definition of relative quasiconvexity, for each vertex g_i of p, there exists an element γ_i in Γ such that $d(g_i, \gamma_i) < \sigma$.

Denote by x_i the element $\gamma_i^{-1} g_i$, and by e_{i+1} the edge of p going from g_i to g_{i+1}. Obviously we have $\gamma_{i+1} = \gamma_i x_i \mathrm{Lab}(e_{i+1}) x_{i+1}^{-1}$.

Set $\kappa = \max\{|x|_d : x \in X\}$. Then κ is finite, as X is finite. Let $Z_0 = \{\gamma \in \Gamma : |\gamma|_d \le 2\sigma + \kappa\}$ and $Z_{x,y,i} = \{xhy^{-1} : \ h \in H_i\} \cap \Gamma$. Since the metric d is proper, the set $B_\sigma := \{g \in G : |g|_d \le \sigma\}$ is finite.

For simplifying notations, we define sets

$$\Pi = \{(x, y, i) : x, y \in B_\sigma, i \in I\}$$

and

$$\Xi = \{(x, y, i) : \sharp\, Z_{x,y,i} = \infty, x, y \in B_\sigma, i \in I\}.$$

If e_{i+1} is an edge labeled by a letter from X, then the element $x_i \mathrm{Lab}(e_{i+1}) x_{i+1}^{-1}$ belongs to Z_0. If e_{i+1} is an edge labeled by a letter from H_k, then $x_i \mathrm{Lab}(e_{i+1}) x_{i+1}^{-1}$ belongs to $Z_{x_i, x_{i+1}, k}$. By the construction, we obtain that the subgroup Γ is also generated by the set

$$Z := Z_0 \cup \left(\cup_{(x,y,i) \in \Pi} Z_{x,y,i} \right).$$

For each $(x, y, i) \in \Pi$, if $Z_{x,y,i}$ is nonempty, then we take an element of the form $x h_i y^{-1} \in Z_{x,y,i}$ for some $h_i \in H_i$. Denote by Z_1 the union of all such elements $\bigcup_{(x,y,i) \in \Pi} x h_i y^{-1}$. Note that $Z_1 \subset Z$. Then we have that Γ is generated by the set

$$\hat{Z} := Y \cup \left(\cup_{(z,z,i) \in \Xi} Z_{z,z,i} \right),$$

where $Y := Z_0 \cup Z_1 \cup \left(\bigcup_{(z,z,i) \in \Pi \setminus \Xi} Z_{z,z,i} \right)$. Indeed, for each triple $(x, y, i) \in \Pi$, we have

$$Z_{x,y,i} = Z_{x,x,i} \cdot x h_i y^{-1}, \text{ where } x h_i y^{-1} \in Z_1.$$

On the other direction, it is obvious that $\hat{Z} \subset Z$.

Let $\hat{X} = X \cup Y \cup B_\sigma$. By the above construction, we define a Γ-equivariant map ϕ from $\mathscr{G}(\Gamma, Z)$ to $\mathscr{G}(G, \hat{X} \cup \mathcal{H})$ as follows. For each vertex $\gamma \in V(\mathscr{G}(\Gamma, Z))$, $\phi(\gamma) = \gamma$. For each edge $[\gamma, s] \in E(\mathscr{G}(\Gamma, Z))$, if $s \in Z_0$, then $\phi([\gamma, s]) = [\gamma, s]$; if $s \in Z_{x,y,i}$ for some $(x, y, i) \in \Xi$, then $s = x t y^{-1}$ for some $t \in H_i$ and we set $\phi([\gamma, s]) = [\gamma, x][\gamma x, t][\gamma x t, y^{-1}]$.

For any $\gamma_1, \gamma_2 \in V(\mathscr{G}(\Gamma, Z))$, it is easy to see that $d_{\hat{X} \cup \mathcal{H}}(\gamma_1, \gamma_2) < 3 d_Z(\gamma_1, \gamma_2)$. For the other direction, we take a geodesic q in $\mathscr{G}(G, X \cup \mathcal{H})$ with endpoints γ_1, γ_2.

Since $\hat{X}$ is finite, there exist constants $\lambda \geq 1, c \geq 0$ depending only on $\hat{X}$, such that the graph embedding $\mathscr{G}(G, X \cup \mathcal{H}) \hookrightarrow \mathscr{G}(G, \hat{X} \cup \mathcal{H})$ is a G-equivariant (λ, c)-quasi-isometry. Thus, q is a (λ, c)-quasigeodesic in $\mathscr{G}(G, \hat{X} \cup \mathcal{H})$, i.e.

$$d_{X \cup \mathcal{H}}(\gamma_1, \gamma_2) < \lambda d_{\hat{X} \cup \mathcal{H}}(\gamma_1, \gamma_2) + c.$$

Since q is a geodesic in $\mathscr{G}(G, X \cup \mathcal{H})$ ending at Γ, we can apply the above analysis to q and obtain that $d_Z(\gamma_1, \gamma_2) < d_{X \cup \mathcal{H}}(\gamma_1, \gamma_2)$. Then we have

$$d_Z(\gamma_1, \gamma_2) < \lambda d_{\hat{X} \cup \mathcal{H}}(\gamma_1, \gamma_2) + c.$$

Therefore, ϕ is a Γ-equivariant quasi-isometric map.

We now claim the subgraph embedding $\imath : \mathscr{G}(\Gamma, \hat{Z}) \hookrightarrow \mathscr{G}(\Gamma, Z)$ is a Γ-equivariant $(2, 0)$-quasi-isometry. This is due to the following observation: every element of Z can be expressed as a word of $\hat{Z}$ of length at most 2.

Finally, we obtain a Γ-equivariant quasi-isometric map $\iota := \phi \cdot \imath$ from $\mathscr{G}(\Gamma, Y \cup \mathcal{K})$ to $\mathscr{G}(G, \hat{X} \cup \mathcal{H})$. □

Remark. Eliminating redundant entries of $\mathbb{K}$ such that all entries of $\mathbb{K}$ are non-conjugate in Γ, we keep the same notation $\mathbb{K}$ for the reduced collection. It is easy to see the construction of the quasi-isometric map $\iota : \mathscr{G}(\Gamma, Y \cup \mathcal{K}) \to \mathscr{G}(G, X \cup \mathcal{H})$ works for the reduced $\mathbb{K}$.

In the following of this subsection, we assume the Γ-equivariant quasi-isometric map $\iota : \mathscr{G}(\Gamma, Y \cup \mathcal{K}) \to \mathscr{G}(G, X \cup \mathcal{H})$ is the one constructed in Lemma 16.4. In particular X is the suitable chosen relative generating set such that $Y \subset X$.

Lemma 16.5. *Suppose $(G, \mathbb{H})$ is relatively hyperbolic. Let $\Gamma < G$ be relatively σ-quasiconvex. Then the quasi-isometric map $\iota : \mathscr{G}(\Gamma, Y \cup \mathcal{K}) \to \mathscr{G}(G, X \cup \mathcal{H})$ sends* *distinct peripheral* $\mathbb{K}$*-cosets of* Γ *to a d-distance* σ *from distinct peripheral* $\mathbb{H}$*-cosets of* G*.*

Proof. Taking into account Lemma 16.4 and Remark 16, we suppose all entries of $\mathbb{K}$ are non-conjugate. We continue the notations in the proof of Lemma 16.4.

By the construction of ϕ, we can see the map ϕ sends the subset $gZ_{x,x,i}$ to a uniform d-distance σ from the peripheral coset gxH_i of G for each $(x, x, i) \in \Xi$ and $g \in G$. Here σ is the quasiconvex constant associated to Γ. Observe that $\imath : \mathscr{G}(\Gamma, \hat{Z}) \hookrightarrow \mathscr{G}(\Gamma, Z)$ is an embedding. Therefore, we have the quasi-isometric map $\iota = \phi \cdot \imath$ maps each peripheral $\mathbb{K}$-coset to a uniform distance from a peripheral $\mathbb{H}$-coset.

We now show the "injectivity" of ι on $\mathbb{K}$-cosets. Let $\gamma H_i^g \cap \Gamma$, $\gamma' H_{i'}^{g'} \cap \Gamma$ be distinct peripheral $\mathbb{K}$-cosets of Γ, where $\gamma, \gamma' \in \Gamma$ and $H_i^g \cap \Gamma, H_{i'}^{g'} \cap \Gamma \in \mathbb{K}$.

Using Lemma 15.10, it is easy to deduce that if $\gamma(H_i^g \cap \Gamma)\gamma^{-1} \cap (H_{i'}^{g'} \cap \Gamma)$ is infinite, then $i = i'$ and $\gamma \in H_i^g \cap \Gamma$.

It is seen from the above discussion that there is a uniform constant $\sigma > 0$, such that $\iota(\gamma H_i^g \cap \Gamma) \subset N_\sigma(\gamma g H_i)$ and $\iota(\gamma' H_{i'}^{g'} \cap \Gamma) \subset N_\sigma(\gamma' g' H_{i'})$. It suffices to show that $\gamma g H_i \neq \gamma' g' H_{i'}$.

Without loss of generality, we assume that $i = i'$. Suppose, to the contrary, that $\gamma g H_i = \gamma' g' H_i$. Then we have $\gamma g = \gamma' g' h$ for some $h \in H_i$. It follows that $\gamma g H_i g^{-1} \gamma^{-1} = \gamma' g' H_i g'^{-1} \gamma'^{-1}$. This implies that $H_i^g \cap \Gamma$ is conjugate to $H_i^{g'} \cap \Gamma$ in Γ, i.e. $H_i^g \cap \Gamma = (H_i^{g'} \cap \Gamma)^{\gamma^{-1}\gamma'}$. Since any two entries of $\mathbb{K}$ are non-conjugate in Γ, we have $H_i^g \cap \Gamma = H_i^{g'} \cap \Gamma$. As a consquence, we have $\gamma^{-1}\gamma' \in H_i^g \cap \Gamma$, as $H_i^g \cap \Gamma \in \mathbb{K}$ is infinite. This is a contradiction, since we assumed $\gamma H_i^g \cap \Gamma \neq \gamma' H_{i'}^{g'} \cap \Gamma$.

Therefore, ι sends distinct peripheral $\mathbb{K}$-cosets of Γ to a uniform distance from distinct peripheral $\mathbb{H}$-cosets of G. □

Before proceeding to prove the relative hyperbolicity of relatively quasiconvex subgroups, we need justify the finite collection $\mathbb{K}$ in (24) as a set of representatives of Γ-conjugacy classes of $\hat{\mathbb{K}}$ in (25).

Lemma 16.6. [9] *Suppose* $(G, \mathbb{H})$ *is relatively hyperbolic. Let* $\Gamma < G$ *be relatively* σ*-quasiconvex. Then the following collection of subgroups of* Γ

$$\hat{\mathbb{K}} = \{H_i^g \cap \Gamma : \sharp\, H_i^g \cap \Gamma = \infty, g \in G, i \in I\}. \tag{25}$$

consists of finitely many Γ*-conjugacy classes. In particular,* $\mathbb{K}$ *is a set of representatives of* Γ*-conjugacy classes of* $\hat{\mathbb{K}}$*.*

Proof. This is proven by adapting an argument of Martinez-Pedroza [9, Proposition 1.5] with our formulation of BCP property 15.3. We refer the reader to [9] for the details. □

We are ready to show the relative hyperbolicity of $(\Gamma, \mathbb{K})$. Using notations in the proof of Lemma 16.4, we recall that $\mathbb{K} = \{Z_{x,x,i} : (x, x, i) \in \Xi\}$.

Lemma 16.7. *Suppose* $(G, \mathbb{H})$ *is relatively hyperbolic. If* $\Gamma < G$ *is relatively* σ*-quasiconvex, then* $(\Gamma, \mathbb{K})$ *is relatively hyperbolic.*

Proof. Recall that ι is the Γ-equivariant quasi-isometric map from $\mathscr{G}(\Gamma, Y \cup \mathcal{K})$ to $\mathscr{G}(G, X \cup \mathcal{H})$. In particular we assumed $Y \subset X$.

We shall prove the relative hyperbolicity of Γ using Farb's definition. First, it is straightforward to verify that $\mathscr{G}(\Gamma, Y \cup \mathcal{K})$ has the thin-triangle property, using the quasi-isometric map ι and the hyperbolicity of $\mathscr{G}(G, X \cup \mathcal{H})$.

Let d_G be a proper left invariant metric on G. Denote by d_Γ the restriction of d_G on Γ. Obviously d_Γ is a proper left invariant metric on Γ. We are going to verify BCP property 1) with respect to d_Γ, for the pair $(\Gamma, \mathbb{K})$. The verification of BCP property 2) is similar.

Let $[\gamma, s]$ be an edge of $\mathscr{G}(\Gamma, Y \cup \mathcal{K})$, where $s \in Z_{x,x,i}$ for some $(x, x, i) \in \Xi$. By the construction of ι, $[\gamma, s]$ is mapped by ι to the concatenated path $[\gamma, x][\gamma x, t][\gamma z t, x^{-1}]$, which clearly contains an H_i-component $[\gamma x, t]$. Note that $|x|_d \le \sigma$. To simplify notations, we reindex $\mathbb{K} = \{K_j\}_{j \in J}$.

Given $\lambda \ge 1$ and $c \ge 0$, we consider two (λ, c)-quasigeodesics p, q without backtracking in $\mathscr{G}(\Gamma, Y \cup \mathcal{K})$ such that $p_- = q_-$, $p_+ = q_+$. By Lemma 16.5, as p, q are assumed to have no backtracking, the paths $\hat{p} = \iota(p)$, $\hat{q} = \iota(q)$ in $\mathscr{G}(G, X \cup \mathcal{H})$ also have no backtracking. Moreover, for each H_i-component $\hat{s}$ of $\hat{p}$(resp. $\hat{q}$), there is a K_j-component s of p(resp. q) such that $\hat{s} \subset \iota(s)$.

Note that paths $\hat{p}, \hat{q}$ are (λ', c')-quasigeodesic without backtracking in $\mathscr{G}(G, X \cup \mathcal{H})$ for some $\lambda' \ge 1, c' \ge 1$. By BCP property of $(G, \mathbb{H})$, we have the constant $\hat{a} = a(\lambda', c', d_G)$. Set $a = \hat{a} + 2\sigma$, where σ is the quasiconvex constant of Γ. Let s be a K_j-component of p for some $j \in J$. We claim that if $d_\Gamma(s_-, s_+) > a$, then there is a K_j-component t of q connected to s.

By the property of the map ι, there exists an H_i-component $\hat{s}$ of $\hat{p}$ such that the following hold
$$d_G(\hat{s}_-, \iota(s)_-) \le \sigma,\ d_G(\hat{s}_+, \iota(s)_+) \le \sigma.$$
Thus, we have $d_G(\hat{s}_-, \hat{s}_+) > \hat{a}$. Using BCP property 1) of $(G, \mathbb{H})$, there exists an H_i-component $\hat{t}$ of $\hat{q}$, that is connected to $\hat{s}$. By the construction of ι, there is a K_k-component t of q for some $k \in J$ such that $\hat{t} \subset \iota(t)$.

Since $\hat{s}$ and $\hat{t}$ are connected as H_i-components, endpoints of $\hat{s}$ and $\hat{t}$ belong to the same H_i-coset. By Lemma 16.5, it follows that $k = j$. Furthermore, endpoints of s and t must belong to the same K_j-coset. Hence s and t are connected in $\mathscr{G}(\Gamma, Y \cup \mathcal{K})$. Therefore, it is verified that $(\Gamma, \mathbb{K})$ satisfies BCP property 1). $\square$

17. Boundaries of relative hyperbolic groups

17.1. **Topological characterization of hyperbolic groups.** We say that a convergence group action is a *uniform convergence group* if it acts properly discontinuously and cocompactly on the space of distinct triples.

Theorem 17.1 (Bowditch)**.** *A convergence group is uniform if and only if it is hyperbolic.*

Theorem 17.2 (Tukia)**.** *A convergence group is uniform if and only if every limit points are conical.*

Theorem 17.3. *A subgroup H is quasiconvex in a hyperbolic group if and only if every limit points of H are conical.*

17.2. **Horofunction (Buseman) boundary.** Let X be a metric space. Fix a basepoint $o \in X$. Consider the family of nomalized distance functions at each point $x \in X$:

$$\mathcal{X} = \{y \mapsto b_x(y) : x \in X\}$$

where $b_x(y) := d(x,y) - d(o,y)$. So we can identify X with $\mathcal{X}$ by sending x to a 1-lipschitz function $b_x(\cdot)$.

Let $Lip_1(X,o)$ be the set of 1-lipschitz functions on X vanishing at the basepoint o. By the following exercise, $\mathcal{X}$ is precompact in $Lip_1(X,o)$ endowed with uniform convergence topology on bounded balls equivalently, pointwise convergence). Note that $\mathcal{X}$ with the subspace topology is homeomorphic to X with metric topology. So X admits a topological embedding into $Lip_1(X,o)$.

Definition 17.4. The *horofunction compactification* $\bar{X}$ of X is the toplogical closure of $\mathcal{X}$ in $Lip_1(X,o)$, and the *horofunction boundary* denoted by $\partial_b X$ is $\bar{X} \setminus X$. The latter is the set of limiting functions $b_\xi(y)$ of $\mathcal{X}$ which are called *horofunctions*. Namely, each point $\xi \in \partial_b X$ is associated with 1-lipschitz horofunction as follows:

$$b_\xi(y) = \lim_{x_n \to \xi} b_{x_n}(y).$$

In plain words, $\partial_b X$ can be understood either as a set of points to compactify X or as the set of horofunctions in the space of 1-lipschitz functions

Exercise 17.5. *Let $\mathcal{F}$ be a family of L-lipschitz functions on a metric space X. Prove that $F(x) := \inf\{f(x) : f \in \mathcal{F}\}$ is L-lipschitz if F is finite on one point.*

Corollary 17.6 (McShane)**.** *If f is a L-lipschtz function on a subset A of a metric space, then there exists an L-lipschitz extension of f to X.*

Proof. The extension is given by $F(x) = \inf_{a \in A}\{f(a) + L \cdot d(a,x)\}$. □

Lemma 17.7. *Let X be a separable metric space. Each isometry g of X extends by homeomorphism to horofunction compactification $\partial_b X \cup X$: the isometry on X extends to horofunction boundary $\partial g : \partial_b X \to \partial_b X$ defined by*

$$\partial g : \xi \mapsto \partial g(\xi) := b_\xi(g^{-1}y) - b_\xi(g^{-1}o)$$

where $b_\xi(y)$ is the horofunction associated to ξ.

Proof. If $x_n \to \xi$ for $x_n \in X$, then $b_{x_n}(y) \to b_\xi(y)$ for any $y \in X$. Noting

$$b_{gx_n}(y) = d(y, gx_n) - d(o, gx_n) = d(g^{-1}y, x_n) - d(g^{-1}o, x_n)$$

we see that $b_{gx_n}(y) \to b_\xi(g^{-1}y) - b_\xi(g^{-1}o)$. Thus, g extends to a homeomorphism on $\partial_b X \cup X$. □

Exercise 17.8. *Let X be a hyperbolic space. There exists a continuous and surjective map from the horofunction boundary $\partial_b X$ to Gromov boundary ∂X.*

17.3. **Floyd boundary.**

References

1. B. Bowditch, *Convergence groups and configuration spaces*, Geometric Group Theory Down Under (J. Cossey, C.F. Miller, W.D. Neumann, M. Shapiro, eds.), pp. 23–54, de Gruyter, 1999.
2. ______, *A course on geometric group theory*, Mathematical Society of Japan, Tokyo, 2006.
3. ______, *Relatively hyperbolic groups*, Int. J. Algebra Comput. (2012), no. 22, p1250016.
4. M. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, vol. 319, Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 1999.
5. B. Farb, *Relatively hyperbolic groups*, Geom. Funct. Anal. **8** (1998), no. 5, 810–840.
6. G. Hruska, *Relative hyperbolicity and relative quasiconvexity for countable groups*, Algebr. Geom. Topol. **10** (2010), 1807–1856.
7. CF Miller III, *Combinatorial group theory*, http://www.ms.unimelb.edu.au/ cfm/notes/cgt-notes.pdf.
8. J.Meier, *Groups, graphs and trees: An introduction to the geometry of infinite groups*, Cambridge University Press, Cambridge, 2008.
9. E. Martinez-Pedroza, *Combination of quasiconvex subgroups of relatively hyperbolic groups*, Groups, Geometry, and Dynamics **3** (2009), 317–342.
10. D. Osin, *Relatively hyperbolic groups: intrinsic geometry, algebraic properties, and algorithmic problems*, vol. 179, Mem. Amer. Math. Soc., 2006.
11. John G. Ratcliffe, *Foundations of hyperbolic manifolds*, Graduate Texts in Mathematics, vol. 149, Springer-Verlag, New York, 1994. MR 1299730
12. Jean-Pierre Serre, *Trees*, Springer-Verlag, Berlin-New York, 1980, Translated from the French by John Stillwell. MR 607504
13. P. Tukia, *Conical limit points and uniform convergence groups*, J. Reine. Angew. Math. **501** (1998), 71–98.