A First Course in Geometric and Analytic Group Theory

Andreas Thom

February 26, 2026

Contents

Introduction

These lecture notes are based on a course taught in the winter term 2025 and summer term 2026. The goal is to introduce geometric and analytic methods in group theory and to show how algebraic questions can be translated into problems about graphs, cell complexes, and large-scale geometry, and how these viewpoints connect to more analytic tools. Throughout, the emphasis is on concrete examples and on techniques that lead to effective proofs and algorithms.

The first part develops the geometric dictionary: free groups via reduced words and Schreier graphs; subgroup methods via core graphs and Stallings folding; and the passage from generators and relations to van Kampen diagrams, Dehn functions and small cancellation. We then introduce Cayley graphs and metric geometry (growth, quasi-isometries, coarse invariants) and use this framework to discuss amenability, growth restrictions and related structural results (including themes around growth gaps).

The second part develops analytic viewpoints: unitary representations and convolution operators; random walks, entropy and the Poisson boundary; operator-algebraic tools such as the reduced group C^*-algebra; and a brief introduction to approximation paradigms for groups (sofic and hyperlinear).

The text is written for readers with a first exposure to group theory; familiarity with basic graph theory and elementary topology is helpful but not required beyond what is developed here.

These notes were prepared from lectures and are based on notes taken by Karl Vincent Kosse. Any errors or inaccuracies are the author's responsibility.

Part I

Geometric methods

Chapter 1

Free groups

Free groups are the basic building blocks of combinatorial group theory. They admit a simple combinatorial model, have a universal property, and enjoy strong algebraic properties such as residual finiteness. In this chapter we develop the theory of free groups from scratch and illustrate it with concrete examples.

1.1 Free groups

In this section, we start out by constructing the free group on a set S from reduced words in the alphabet $S \cup S^{-1}$ and use this model to prove the universal property and the normal form.

1.1.1 Definitions

Definition 1.1.1. Let S be a set and let $S^{-1} := \{s^{-1} \mid s \in S\}$ be a disjoint copy, so $S \cap S^{-1} = \varnothing$ and set $(x^{-1})^{-1} = x$. The alphabet is defined to be $S \cup S^{-1}$. A word in the alphabet is defined to be a finite sequence $x_1 \ldots x_m$ with $x_i \in S \cup S^{-1}$. We allow $m = 0$, which corresponds to the empty word ε. The length of the word is defined to be $|w| = m$. For words $x_1 \ldots x_m$ and $y_1 \ldots y_n$, their *concatenation* is defined to be

$$x_1 \ldots x_m y_1 \ldots y_n.$$

If $S = \{x, y\}$, then examples of words in the alphabet include $x, x^{-1}, xyxy, xxx, xxx^{-1}$. In order to define a group, we would like xxx^{-1} to be equivalent to x.

Definition 1.1.2. A word $w = x_1 x_2 \ldots x_m$ is called reduced if there is no $1 \leq i < m$ such that $x_{i+1} = x_i^{-1}$.

The words xyx^{-1} and xxy are reduced, while $xx^{-1}y$ is not reduced. In the best possible world, there would be a suitable equivalence relation on words such that each equivalence class of words would contain a unique reduced representative. We will see that this is indeed the case.

Definition 1.1.3. A word w is an elementary contraction of w', written $w' \downarrow w$, if $w' = y_1 x x^{-1} y_2$ and $w = y_1 y_2$ for words y_1, y_2 and a letter x. We also write $w \uparrow w'$ and call w' an elementary expansion of w. Two words w, w' are called equivalent if and only if there exists a finite sequence $w_1, \ldots, w_n$ of words with $w_1 = w$, $w_n = w'$ and for all $1 \leq i \leq n-1$ we have either $w_i \downarrow w_{i+1}$ *or* $w_i \uparrow w_{i+1}$.

It is obvious that this defines an equivalence relations on the set of words.

Definition 1.1.4. We define $\mathcal{F}_S$ to be the set of words in the alphabet $S \cup S^{-1}$ modulo equivalence. We denote the equivalence class of a word w by $[w]$.

Theorem 1.1.5. *Concatenation induces a product $\mathcal{F}_S \times \mathcal{F}_S \to \mathcal{F}_S$ that turns $\mathcal{F}_S$ into a group.*

Proof. If $w \downarrow w'$ or $w \uparrow w'$, then for any word y we have $yw \downarrow yw'$ or $yw \uparrow yw'$ and similarly for concatenation on the right. Thus concatenation respects the equivalence relation and we can set $[w_1] \cdot [w_2] \coloneqq [w_1 w_2]$, where associativity of concatenation implies associativity of the product. The class $[\varepsilon]$ acts as the identity element.

For a word $w = x_1 \dots x_m$, we set $w^{-1} \coloneqq x_m^{-1} \dots x_1^{-1}$. Then ww^{-1} contracts to ε, so $[w^{-1}]$ is the inverse of $[w]$. This shows that $\mathcal{F}_S$ is a group. □

Proposition 1.1.6. *Each equivalence class contains a unique reduced word.*

In order to prove this, we need the following lemma.

Lemma 1.1.7. *Let w_1, w_2, w_3 be words s.t. $w_1 \uparrow w_2 \downarrow w_3$. Then, either there exists w_2' such that $w_1 \downarrow w_2' \uparrow w_3$ or $w_1 = w_3$.*

Proof. Let $w_1 = y_1 y_2$ and write $w_2 = y_1 x x^{-1} y_2$ since $w_1 \uparrow w_2$. There are several cases to consider.

Case 1: The contraction $w_2 \downarrow w_3$ happens inside y_1 or y_2, for example $y_1 = a_1 y y^{-1} a_2$. So

$$w_1 = a_1 y y^{-1} a_2 y_2 \uparrow a_1 y y^{-1} a_2 x x^{-1} y_2 \downarrow a_1 a_2 x x^{-1} y_2.$$

But then $w_1 \downarrow a_1 a_2 y_2 \uparrow a_1 a_2 x x^{-1} y_2$. Similary, when the contraction happens in y_2.

Case 2: There is complete overlap of expansion and contraction, but then $w_1 = w_3$.

Case 3: There is overlap of one letter in expansion and contraction, i.e. for $y_2 = x y_2'$ we have

$$w_1 = y_1 y_2 \uparrow y_1 x x^{-1} y_2 = y_1 x x^{-1} x y_2' \downarrow y_1 x y_2' = y_1 y_2 = w_1$$

and similarly for $y_1 = y_1' x^{-1}$. This yields $w_1 = w_3$ again. □

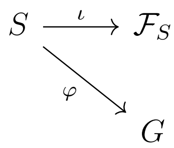
Proof of Proposition 1.1.6. Suppose for contradiction that there are two distinct reduced words u and v in the same equivalence class. Indeed, suppose there is a sequence $u = w_0, w_1, \dots, w_n = v$ where each step is an expansion or a contraction. If a peak $w_{i-1} \uparrow w_i \downarrow w_{i+1}$ occurs, Lemma 1.1.7 replaces this length-two segment by either $w_{i-1} = w_{i+1}$ and is shortening the sequence or by $w_{i-1} \downarrow w_i' \uparrow w_{i+1}$, which moves the contraction to the left of the expansion and reduces the number of expansions done before contractions. Repeating this finitely many times yields a sequence with no peaks, so all contractions occur first and all expansions occur last. Since u is reduced it admits no contraction, we conclude $u = v$. □

Remark 1.1.8. To check whether two words are equivalent, reduce each one by successive elementary contractions until no further contraction is possible. By the Proposition 1.1.6 this yields a unique reduced word, so two words are equivalent exactly when their reduced forms coincide. In particular, $\iota \colon S \cup S^{-1} \to \mathcal{F}_S$ is injective because each generator maps to a distinct reduced word of length one. Throughout of rest of these notes, we will identify S with its image in $\mathcal{F}_S$ via ι and $[w]$ with the unique reduced representative of the equivalence class.

Definition 1.1.9. For $w \in \mathcal{F}_S$, let $|w|$ denote the length of the reduced representative of w.

The most useful property of free groups is their universal property.

Theorem 1.1.10. *Let S be a set, G a group, $\varphi : S \to G$ a map. Then, there exists a unique homomorphism $\overline{\varphi}$ that extends φ. In other words, the following diagram commutes:*



Proof. Let $w = x_1^{\varepsilon_1} x_2^{\varepsilon_2} \dots x_n^{\varepsilon_n}$ with $x_i \in S$ and $\varepsilon_i \in \{\pm 1\}$. Define

$$\overline{\varphi}\left([w]\right) := \varphi(x_1)^{\varepsilon_1} \cdots \varphi(x_n)^{\varepsilon_n},$$

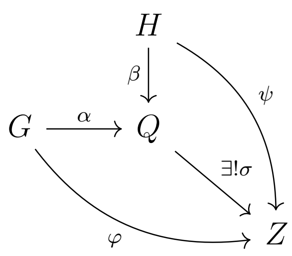
with the product taken in G. Elementary contractions and expansions do not change this product, so $\overline{\varphi}$ is well-defined. By construction it is a homomorphism and satisfies $\overline{\varphi} \circ \iota = \varphi$. Uniqueness follows because S is obviously a generating set of $\mathcal{F}_S$. $\square$

Let $\mathcal{F}$ be a free group. We say that a subset $S \subset \mathcal{F}$ is a *free basis* of $\mathcal{F}$ if the map $\iota : S \to \mathcal{F}$ satisfies the universal property of Theorem 1.1.10. In this case, we have $\mathcal{F} \cong \mathcal{F}_S$. The cardinality of a free basis is called the *rank* of the free group. We will see later that the rank is well-defined.

1.1.2 Free products

Free products provide the coproduct in the category of groups. They will reappear later when we discuss amalgams and HNN extensions in Chapter 2.

Definition 1.1.11. Let G, H be groups. The free product of G and H is a group Q together with homomorphisms $\alpha : G \to Q$, $\beta : H \to Q$ such that for every pair of homomorphisms $\varphi : G \to Z$, $\psi : H \to Z$ there exists a unique $\sigma : Q \to Z$ with $\varphi = \sigma \circ \alpha$ and $\psi = \sigma \circ \beta$. In other words, the following commutes:



We use the notation $G * H$, $\iota_G : G \to G * H, \iota_H : H \to G * H$.

This is just the coproduct in the category of groups.

Lemma 1.1.12. *The free group functor sends coproducts of sets to free products of groups: $\mathcal{F}_{S \sqcup S'} \cong \mathcal{F}_S * \mathcal{F}_{S'}$.*

Proof. For any group Z,

$$\mathrm{Hom}\left(\mathcal{F}_{S \sqcup S'}, Z\right) \cong \mathrm{Fun}\left(S \sqcup S', Z\right) \cong \mathrm{Fun}(S, Z) \times \mathrm{Fun}(S', Z).$$

By adjunction this identifies with $\mathrm{Hom}(\mathcal{F}_S, Z) \times \mathrm{Hom}(\mathcal{F}_{S'}, Z)$, which is the universal property of the free product. $\square$

1.1.3 Concrete examples of free groups

The universal property makes it easy to define homomorphisms from free groups into matrix groups. The issue in constructing concrete free groups is injectivity. The so-called ping-pong lemma gives a clean criterion. We record it and then apply it to a concrete family of matrices.

Theorem 1.1.13. *Let X be a set and let $g, h \in \mathrm{Sym}(X)$. Suppose there are nonempty disjoint subsets $U, V \subset X$ such that for every $n \in \mathbb{Z} \setminus \{0\}$ one has*

$$g^n(V) \subset U, \qquad h^n(U) \subset V.$$

Then the natural homomorphism $\varphi : \mathcal{F}_2 \to \mathrm{Sym}(X)$ with $\varphi(x) = g$ and $\varphi(y) = h$ is injective.

Proof. Let w be a reduced word in the alphabet $\{x^{\pm}, y^{\pm}\}$. If the first and the last letter of w is $x^{\pm 1}$, then successive applications of the hypotheses show that $\varphi(w)(V) \subset U$ and hence $\varphi(w) \neq 1$. If w does not satisfy the assumption, then $x^k w x^{-k}$ for k large enough does. Again, $\varphi(w) \neq 1$ follows. Hence no non-trivial reduced word represents the identity, and φ is injective. □

Now, we consider a natural family of matrices. For $t \in \mathbb{R}$ define

$$A_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \qquad B_t = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix},$$

and let $\varphi_t : \mathcal{F}_2 \to SL_2(\mathbb{R})$ be the homomorphism with $\varphi_t(x) = A_t$ and $\varphi_t(y) = B_t$.

Theorem 1.1.14 (Klein)**.** *If $|t| \geq 2$, then φ_t is injective.*

Proof. Let $X = \mathbb{R}^2$ and set

$$U = \{(a, b) \in \mathbb{R}^2 : |a| > |b|\}, \qquad V = \{(a, b) \in \mathbb{R}^2 : |a| < |b|\}.$$

For $n \neq 0$ we have

$$A_t^n(a, b) = (a + nt\, b, b), \qquad B_t^n(a, b) = (a, b + nt\, a).$$

If $(a, b) \in V$, then $|a + nt\, b| \geq (|t||n| - 1)|b| \geq |b|$, so $A_t^n(V) \subset U$. If $(a, b) \in U$, then $|b + nt\, a| \geq (|t||n| - 1)|a| \geq |a|$, so $B_t^n(U) \subset V$. Hence, Theorem 1.1.13 applies and we conclude that φ_t is injective. □

The previous theorem has the following immediate corollary.

Corollary 1.1.15. *The matrices A_2, B_2 lie in $SL_2(\mathbb{Z})$ and generate a free subgroup. Hence $\mathcal{F}_2$ embeds in $SL_2(\mathbb{Z})$.*

The situation for $t \in (-2, 2)$ is far from trivial and subject of ongoing studies, see for example [29]. There are many positive results towards the following conjecture, but we will not record them here.

Conjecture 1.1.16 (Kim–Koberda [29])**.** *For $t \in \mathbb{Q} \cap (-2, 2)$, the map φ_t is not injective.*

However, the following proposition shows that injectivity is a generic property.

Proposition 1.1.17. *The set* $E := \{t \in \mathbb{C} \mid \varphi_t$ *is not injective*$\}$ *is a set of algebraic numbers.*

Proof. Let $\mathcal{F}_2 = \langle a, b\rangle$ be the free group, and for each t let $\varphi_t : \mathcal{F}_2 \to \mathrm{SL}_2(\mathbb{C})$ be the homomorphism defined by $\varphi_t(a) = A_t$ and $\varphi_t(b) = B_t$. Then φ_t is injective if and only if $\ker(\varphi_t) = \{1\}$.

Fix a non-trivial reduced word $w \in \mathcal{F}_2 \setminus \{1\}$. Since

$$A_t^{\pm 1} = \begin{pmatrix} 1 & \pm t \\ 0 & 1 \end{pmatrix}, \qquad B_t^{\pm 1} = \begin{pmatrix} 1 & 0 \\ \pm t & 1 \end{pmatrix},$$

the matrix entries of $\varphi_t(w)$ are polynomials in t with integer coefficients. Write $\varphi_t(w) = (p_{ij}(t))_{i,j} \in \mathrm{GL}_2(\mathbb{Z}[t])$.

We claim that for fixed $w \neq 1$, at least one polynomial among $p_{11}(t) - 1$, $p_{12}(t)$, $p_{21}(t)$, $p_{22}(t) - 1$ is non-zero. Indeed, if all four were the zero polynomial, then $\varphi_t(w) = 1_2$ for every t, in particular for $t = 2$, contradicting Theorem 1.1.14. Hence the set

$$Z_w := \{t \in \mathbb{C} : \varphi_t(w) = 1_2\}$$

is finite and consists of algebraic numbers, since it is contained in the zero set of a non-zero polynomial with integer coefficients in one variable. This finishes the proof. □

1.2 Residual finiteness of free groups

Residual finiteness is an important algebraic property of groups. Intuitively, it means that the group can be approximated by finite groups. In this section we define residual finiteness and show that free groups are residually finite. As an application, we prove that finitely generated residually finite groups are Hopfian.

1.2.1 Definitions

Definition 1.2.1. A group G is called residually finite if for all non-trivial $g \in G$ there exists a finite group H and a homomorphism $\varphi : G \to H$ such that $\varphi(g) \neq \mathbb{1}_H$.

Example 1.2.2. Finite groups are residually finite. The infinite cyclic group $\mathbb{Z} \cong \mathcal{F}_{\{x\}}$ is residually finite via the quotients $\psi_n : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$. The free group $\mathcal{F}_2$ is residually finite as we will prove shortly by extending the previous argument to the group $\mathrm{SL}_2(\mathbb{Z})$. It is obvious that residual finiteness is inherited by subgroups, so any subgroup of a residually finite group is residually finite. In particular, free groups of rank 2 is residually finite, begin a subgroup of $\mathrm{SL}_2(\mathbb{Z})$ by Corollary 1.1.15.

Lemma 1.2.3. $\mathrm{SL}_2(\mathbb{Z})$ *is residually finite.*

Proof. The natural ring homomorphism $\psi_n : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ induces a group homomorphism $\psi_n' : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ to a finite group. Obviously, this yields finite quotients separating elements of $\mathrm{SL}_2(\mathbb{Z})$. □

We could argue that arbitrary free groups are residually finite by combining the previous two results. However, we will give a more combinatorial proof that will be useful later on.

1.2.2 Residual finiteness of free groups

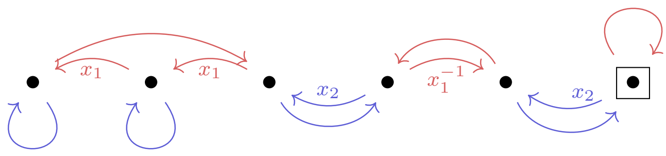
Theorem 1.2.4. *Free groups are residually finite.*

Proof. We give a combinatorial proof. Let S be a set and consider the free group $\mathcal{F}_S$. The idea is to realize a reduced word as a non-trivial permutation by building an S-labelled graph on which the word acts nontrivially.

We explain the idea using a simple example. Let $w \in \mathcal{F}_2$ be a reduced word, for instance $w = x_1^2 x_2 x_1^{-1} x_2$. The aim is to find a natural number n and a homomorphism $\varphi : \mathcal{F}_2 \to \mathrm{Sym}(n)$ such that $\varphi(w) \neq \mathbb{1}_n$, i.e. there exists $x \in \{1, \dots, n\}$ such that $\varphi(w)(x) \neq x$. In order to do so, we need to find $\sigma = \varphi(x_1)$ and $\tau = \varphi(x_2)$ such that $\varphi(w)(x) \neq x$. We have

$$\varphi(w)(x) = \sigma^2 \tau \sigma^{-1} \tau(x).$$

The permutations σ and τ can be visualized as follows:



The algorithm to construct such a diagram is as follows. Start with a vertex and follow the letters of w one by one, drawing edges labelled by the corresponding generator and add a new vertex each time. Starting the right-most vertex x and following all letters of w, you end up at the left-most vertex. Finally, complete the diagram by adding edges such that each vertex has exactly one incoming and one outgoing edge for each generator – this can be done canonically by adding edges to form cycles as shown in the diagram.

The resulting diagram determines σ and τ such that $w(x) \neq x$ and taking $n = |w| + 1$ is enough by construction. It is clear that this procedure can be applied to any reduced word in $\mathcal{F}_S$ for any S, so free groups are residually finite. □

The resulting diagram is what we will call S-labelled Schreier graph, see Section 1.3, and it will be instrumental in the study of free groups far beyond this application.

1.2.3 Hopfian groups

As an application of residual finiteness, we show that finitely generated residually finite groups are Hopfian. This applies in particular to free groups of finite rank.

Definition 1.2.5. A group G is called *Hopfian* if every surjective endomorphism $\pi : G \to G$ is injective.

Theorem 1.2.6. *Every finitely generated residually finite group is Hopfian.*

Proof. Let G be finitely generated and residually finite, and let $\pi : G \to G$ be surjective. For $n \in \mathbb{N}$ let NS_n be the set of normal subgroups of index n. A finitely generated group has only finitely many subgroups of a given finite index, hence Σ_n is finite. Since π is surjective, the assignment $N \mapsto \pi^{-1}(N)$ defines a map from Σ_n to itself. And obviously $\pi^{-1}(N_1) = \pi^{-1}(N_2)$ implies $N_1 = N_2$, so this map is injective and therefore bijective.

Let $H_n := \bigcap_{N\in\Sigma_n} N$. Then H_n has finite index and, by the bijectivity that we just established, $\pi^{-1}(H_n) = H_n$ for all n. Residual finiteness of G gives $\bigcap_{n\geq 1} H_n = \{\mathbb{1}\}$ and hence

$$\ker(\pi) = \bigcap_{n\geq 1} \pi^{-1}(H_n) = \bigcap_{n\geq 1} H_n = \{\mathbb{1}\},$$

so π is injective. □

Corollary 1.2.7. *Let $S = \{s_1, \dots, s_n\} \subset \mathcal{F}_n$ be a generating set of $\mathcal{F}_n$. Then S is a free basis of $\mathcal{F}_n$.*

Proof. Free groups are residually finite, so Theorem 1.2.6 implies that $\mathcal{F}_n$ is Hopfian. Let $\varphi : \mathcal{F}_n \to \mathcal{F}_n$ be the endomorphism with $\varphi(x_i) = s_i$. Since S generates $\mathcal{F}_n$, the map φ is surjective, hence injective. Therefore φ is an automorphism and $S = \varphi(\{x_1, \dots, x_n\})$ is a free basis. □

1.3 Schreier graphs

Schreier graphs are a combinatorial tool to study subgroups of free groups. They provide a graphical representation of the coset space of a subgroup, encoding the action of the free group on this space. This section introduces Schreier graphs and establishes the Schreier correspondence, which relates subgroups of free groups to pointed connected Schreier graphs.

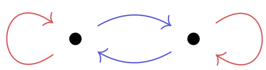
1.3.1 Definitions

Definition 1.3.1. Let S be a set. A S-labelled Schreier graph (or S-Schreier graph) is a directed S-labelled graph (V, E), i.e. it consists of a set of vertices V, a set of edges E together with maps $E \xrightarrow{d,t} V \times V$, $e \mapsto (d(e), t(e))$ and $E \xrightarrow{\ell} S$ sending e to its label, such that for all $x \in V$ and $\sigma \in S$ there exists a unique edge $e \in E$ with $d(e) = x$, $\ell(e) = \sigma$ and a unique edge $e \in E$ with $t(e) = x$, $\ell(e) = \sigma$. We call $d(e)$ the domain and $t(e)$ the target of the edge e.

A pointed S-Schreier graph is a S-Schreier graph together with a distinguished vertex $x \in V$. A S-Schreier graph is called connected if any two vertices can be joined by an undirected path.

The study of disconnected Schreier graphs reduces to the connected case by considering connected components. For finite S-Schreier graphs directed connectivity is in fact equivalent to undirected connectivity.

Example 1.3.2. Let $S = \{\color{red}{\bullet}, \color{blue}{\bullet}\}$.



There is an obvious notion of isomorphism between Schreier graphs, which respects the labelling and the distinguished vertex in the pointed case.

1.3.2 Schreier correspondence

Let's recall some basic notions about group actions.

Definition 1.3.3. Let G be a group, X a set. An action of G on X is a map $m : G \times X \to X$ such that for all $g, h \in G$, $m(g, m(h, x)) = m(gh, x)$ and $m(g, -)$ is a bijection for all $g \in G$. Equivalently, $m_- : G \to \mathrm{Sym}(X)$ is a homomorphism. We write $G \curvearrowright X$ and say that G acts on X.

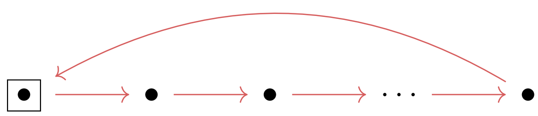
Definition 1.3.4. We say that $G \curvearrowright X$ is transitive if and only if for all $x, y \in X$ there exists $g \in G$ with $gx = y$. If $x \in X$ is a chosen point, then we call $(G \curvearrowright X, x)$ a pointed G-action without requiring any special property of $x \in X$ with respect to the action. Two pointed actions $(G \curvearrowright X, x)$ and $(G \curvearrowright Y, y)$ are called isomorphic if and only if there exists a bijection $\varphi : X \to Y$ such that $\varphi(gz) = g\varphi(z)$ r all $g \in G$, $z \in X$ and $\varphi(x) = y$ fo.

Now, we establish a correspondence between subgroups of free groups and pointed connected Schreier graphs. This correspondence is fundamental in combinatorial group theory and will be used extensively in later sections.

Definition 1.3.5. Let G be a group. Then $\mathrm{Sub}(G) = \{H \leq G\}$ denotes the set of subgroups.

Theorem 1.3.6. *Let S be a set. There are natural bijections between $\mathrm{Sub}(\mathcal{F}_S)$, the set of isomorphism classes of pointed connected S-labelled Schreier graphs, and the set of isomorphism classes of pointed transitive $\mathcal{F}_S$-actions.*

Example 1.3.7. The most basic example is given by $S = \{\bullet\}$. Then $\mathcal{F}_S = \mathcal{F}_1 = \mathbb{Z}$ and $\mathrm{Sub}(\mathbb{Z}) = \{H \leq \mathbb{Z}\} = \{n\mathbb{Z} \mid n \in \mathbb{N}\} = \mathbb{N}$. The transitive pointed $\mathbb{Z}$-actions are given by $\mathbb{Z} \curvearrowright \mathbb{Z}/n\mathbb{Z}$ with distinguished point $0 + n\mathbb{Z}$. The corresponding Schreier graphs are



with n vertices.

Proof of Theorem 1.3.6. We explain how the action and Schreier graph viewpoints encode the same subgroup data. Given $H \leq \mathcal{F}_S$, there is a pointed transitive $\mathcal{F}_S$-action $\mathcal{F}_S \curvearrowright \mathcal{F}_S/H$ with distinguished point H. To reverse this, recall that the stabilizer of some $x \in X$ is defined as $\mathrm{Stab}(x) := \{g \in G \mid gx = x\}$. Then, given a pointed transitive $\mathcal{F}_S$-action $(G \curvearrowright X, x)$, assign to it $\mathrm{Stab}(x)$.

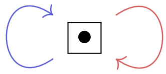
Now, we compute the compositions. For $H \leq \mathcal{F}_S$, $\mathrm{Stab}_{\mathcal{F}_S \curvearrowright \mathcal{F}_S/H}(H) = \{g \in \mathcal{F}_S \mid gH = H\} = H$ so these two constructions are inverse to each other.

Now, suppose $\mathcal{F}_S \curvearrowright X$ is an action. Define an S-Schreier graph by setting $V = X$, $E = S \times X$ and $d, t \colon E \to X$, $\ell : E \to S$ by $d(\sigma, x) = x$, $t(\sigma, x) = \sigma x$ and $\ell(\sigma, x) = \sigma$ for all $x \in X$ and $\sigma \in S$. Conversely, given a pointed S-Schreier graph (G, x_0), define an $\mathcal{F}_S$-action on V by

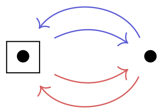
$$\sigma_1^{\varepsilon_1} \dots \sigma_n^{\varepsilon_n} x := t\left(\sigma_n^{\varepsilon_n}, t\left(\sigma_{n-1}^{\varepsilon_{n-1}}, \dots t\left(\sigma_1^{\varepsilon_1}, x\right) \dots\right)\right)$$

for all reduced words $\sigma_1^{\varepsilon_1} \dots \sigma_n^{\varepsilon_n} \in \mathcal{F}_S$, $x \in V$. It is straightforward to check that these two constructions are inverse to each other. It is easy to check that isomorphic actions yield isomorphic Schreier graphs and vice versa, so the bijections descend to isomorphism classes. Moroever, an S-Schreier graph is connected if and only if the corresponding action is transitive. □

Example 1.3.8. Let $S = \{\bullet, \bullet\}$. For $\mathcal{F}_2 \leq \mathcal{F}_2$, we obtain the trivial action $\mathcal{F}_2 \curvearrowright \{*\}$ and the corresponding Schreier graph is:



For $H := \langle a^2, ab, b^2 \rangle \leq \mathcal{F}_2$, observe that H is the kernel of $\varphi : \mathcal{F}_2 \to \mathbb{Z}/2\mathbb{Z}$, $\varphi(a) = \varphi(b) = 1$, so that $\mathcal{F}_2/H = \{H, aH\}$ and the corresponding Schreier graph looks as follows:



1.4 Applications

1.4.1 Definitions

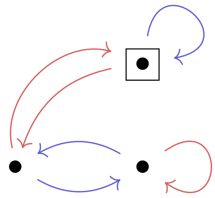
Let S be a set and let (G, x_0) be a pointed S-Schreier graph. A *based loop* is an undirected path that starts and ends at x_0. We say that a loop is *backtracking* if its label contains a subword ss^{-1} or $s^{-1}s$ for some $s \in S$. Two based loops are homotopic if one can be obtained from the other by inserting or deleting backtracking segments; this corresponds to expansion and contraction of words according to Definition 1.1.3. The homotopy classes form a group under concatenation; each class has a unique representative with no backtracking. This fact can be seen by observing that any backtracking can be removed without affecting the homotopy class, and that concatenation of reduced loops followed by reduction yields a unique reduced representative, see also Proposition 1.1.6. We call this group the fundamental group $\pi_1(G, x_0)$. Note by definition reading labels this Group embeds naturally in $\mathcal{F}_S$.

Remark 1.4.1. In algebraic topology language this is just the covering space theory for a bouquet of circles. Schreier graphs are coverings of the wedge of $|S|$ circles, whose fundamental group naturally the free group on S. This interplay between the group theoretic and geometric notions is the so-called Galois correspondence for covering spaces.

Lemma 1.4.2. *Let S be a set. The fundamental group of a pointed S-Schreier graph is the subgroup corresponding to that graph in Theorem 1.3.6.*

Proof. Let (G, x_0) be a pointed connected S-Schreier graph and let $H \leq \mathcal{F}_S$ be the corresponding subgroup. We need to show that $\pi_1(G, x_0) = H$ as subgroups of $\mathcal{F}_S$. By construction of the action from the Schreier graph, an element $g \in \mathcal{F}_S$ lies in H if and only if $gx_0 = x_0$. By construction of the action from the Schreier graph, this is equivalent to the existence of a based loop at x_0 with label g. Finally, by definition of the fundamental group, this is equivalent to $g \in \pi_1(G, x_0)$. Hence $\pi_1(G, x_0) = H$. □

For instance, with $S = \{\textcolor{red}{\bullet}, \textcolor{blue}{\bullet}\}$ and



we obtain the subgroup $H = \langle a^2, ab, ba, b^2\rangle \leq \mathcal{F}_2$. This illustrates how the fundamental group construction recovers the subgroup associated to a Schreier graph.

1.4.2 Nielsen–Schreier Theorem

Lemma 1.4.3. *Every connected graph contains a spanning tree.*

Proof. This can be proved by induction on the number of edges. If there are no cycles, the graph is already a tree. Otherwise, remove an edge from a cycle; this does not disconnect the graph. Repeat until no cycles remain. For infinite graphs, use Zorn's lemma to find a maximal acyclic subgraph, which will be a spanning tree. □

Theorem 1.4.4 (Nielsen–Schreier)**.** *Every subgroup of a free group is free.*

Proof. Let T be a spanning tree of a connected S-Schreier graph G. Every reduced loop is uniquely determined by the sequence of oriented edges it uses outside T. Hence $\pi_1(G)$ is free on the set of edges not in the spanning tree. □

In a more careful analysis, one can even quantify their rank in terms of index. First of all, let us define $\mathrm{rk}(G) = \min\{|S| \mid S \subset G, \langle S\rangle = G\}$. This is called the rank of G.

Lemma 1.4.5. *Let S be a finite set. We have that* $\mathrm{rk}(\mathcal{F}_S) = |S|$.

Proof. The inequality $\mathrm{rk}(\mathcal{F}_S) \leq |S|$ is clear by definition. For the other direction, the free group functor is adjoint to the forgetful functor, so there is in particular a bijection

$$H^S := \mathrm{Fun}(S, H) \cong \mathrm{Hom}(\mathcal{F}_S, H).$$

Thus for finite H,

$$|H|^{|S|} = |H^S| = |\mathrm{Hom}(\mathcal{F}_S, H)| \leq |H|^{\mathrm{rk}(\mathcal{F}_S)}.$$

Therefore $|S| \leq \mathrm{rk}(\mathcal{F}_S)$. This proves the claim. □

Given a Schreier graph corresponding to a subgroup $H \leqslant \mathcal{F}_S$. By construction, the number of vertices of the associated S-Schreier graph is equal to the index $[\mathcal{F}_S : H]$. Hence, a spanning tree has $[\mathcal{F}_S : H]$ vertices and $[\mathcal{F}_S : H] - 1$ edges. Since the total number of edges is $|S|[\mathcal{F}_S : H]$, the number of edges not in the spanning tree is $|S|[\mathcal{F}_S : H] - ([\mathcal{F}_S : H] - 1) = (|S| - 1)\,[\mathcal{F}_S : H] + 1$. Therefore , we have the following result on the rank of subgroups of finite index in free groups.

Theorem 1.4.6. *Let S be a set and $H \leqslant \mathcal{F}_S$ be of finite index. Then*

$$\mathrm{rk}(H) - 1 = [\mathcal{F}_S : H](\mathrm{rk}(\mathcal{F}_S) - 1).$$

Corollary 1.4.7. *Let G be a finitely generated group and $H \leqslant G$ of finite index. Then, H is finitely generated and*

$$\mathrm{rk}(H) - 1 \leq [G : H](\mathrm{rk}(G) - 1).$$

Proof. Let S be a generating set of G with $|S| = \mathrm{rk}(G)$. Then, there is a surjection $\varphi : \mathcal{F}_S \to G$. Let $H' = \varphi^{-1}(H)$. Then, $H' \leqslant \mathcal{F}_S$ is of finite index $[\mathcal{F}_S : H'] = [G : H]$. By Theorem 1.4.6, H' is finitely generated and

$$\mathrm{rk}(H') - 1 = [\mathcal{F}_S : H'](\mathrm{rk}(\mathcal{F}_S) - 1) = [G : H](\mathrm{rk}(G) - 1).$$

Since φ restricts to a surjection $H' \to H$, we have $\mathrm{rk}(H) \leq \mathrm{rk}(H')$, which proves the claim. □

The study of Schreier graphs goes far beyond the Nielsen–Schreier theorem and the rank formula. However, the study of subgroups of infinite index is more subtle and requires additional tools, which we introduce next.

1.4.3 Stallings' pruning and folding

See [50] for the original reference. Let (G, x) be a pointed S-Schreier graph. The vertex core of (G, x) is defined to be

$$\{v \in V \mid \exists \text{ a non-backtracking based loop containing } v\}.$$

That is, the vertex core of G consists of those vertices that actually contribute to the fundamental group $\pi(G, x)$ in a non-trivial manner. We set

$$\mathrm{core}(G, x) \coloneqq \text{induced subgraph on the vertex-core of } G.$$

Note that $\mathrm{core}(G, x)$ is again a pointed S-labelled graph with the same basepoint as G, however, it is usually not a Schreier graph anymore, as some vertices may have less than one incoming or outgoing edge for some label $s \in S$. Geometrically, $\mathrm{core}(G, x)$ is obtained from G by pruning all leaves and induced subtrees that are not part of any non-backtracking based loop.

Definition 1.4.8. A partial pointed S-labelled Schreier Graph is a directed pointed S-labelled graph with at most one outgoing and one ingoing edge labelled $s \in S$.

It is straighforward to extend the notion of fundamental group to partial Schreier graphs. Note that pruning leaves or induced subtrees does not change the fundamental group, so $\pi_1(\mathrm{core}(G, x)) = \pi_1(G, x)$.

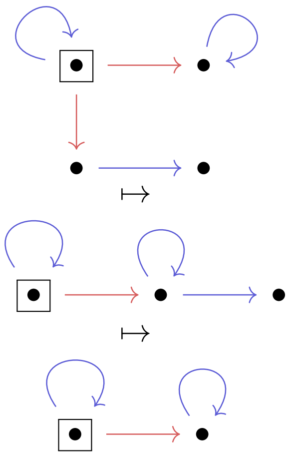
There is a second important operation on S-labelled graphs, called folding. Let (G, x) be a pointed connected S-labelled graph. We now construct a partial S-Schreier graph from (G, x) by identifying redundant edges. The resulting folding functor is central for algorithmic subgroup problems. This construction $(G, x) \mapsto \mathrm{fold}(G, x)$ is most easily characterized by its functoriality properties. Indeed, it is left adjoint to the forgetful functor from partial S-Schreier graphs to S-labelled graphs. Indeed, for every pointed S-labelled graph (G, x) and every partial pointed S-Schreier graph (G', x'), every morphism

$g : (G,x) \to (G',x')$ factors uniquely through a morphism $\bar{g} : \mathrm{fold}(G,x) \to (G',x')$ as follows:

$$\begin{array}{ccc} (G,x) & \longrightarrow & \mathrm{fold}(G,x) \\ & \searrow^{g} & \downarrow{\exists! \bar{g}} \\ & & (G',x') \end{array}$$

$$\mathrm{map}\big((G,x), \mathrm{for}(G',x')\big) \cong \mathrm{map}\big(\mathrm{fold}(G,x),(G',x')\big). \quad .$$

Here, we write for for the forgetful functor from partial S-Schreier graphs to S-labelled graphs. A concrete construction in order to produce $\mathrm{fold}(G,x)$ out of (G,x) can be described as repeatedly folding conflicts, i.e. identifying edges with the same label and the same initial or terminal vertex. For example, given the following S-labelled graph with $S = \{\bullet, \bullet\}$:



It is not entirely clear that this inductive process yields a well-defined result. However, one can show that any such sequence of foldings results in a partial S-Schreier graph that satisfies the universal property of the free forgetful adjunction; hence it is unique up to isomorphism.

1.4.4 Subgroup membership problem

An immediate application of Stallings folding is to solve the subgroup membership problem in free groups. Let $w_1, \ldots, w_n \in \mathcal{F}_S$ and $w \in \mathcal{F}_S$. Can we algorithmically decide whether $w \in \langle w_1, \ldots, w_n \rangle$? The subgroup membership problem is generally undecidable for arbitrary finitely presented groups, but for free groups we can provide an algorithmic solution using Stallings folding. We answer this algorithmically by encoding the generators as loops and folding to a finite core graph. Let $S(w_i)$ be the Schreier graph that forms a cycle spanned by the letters in the reduced word of w_i.

If (G_1, x_1) and (G_2, x_2) are pointed S-labelled graphs, write

$$(G_1 \vee G_2, x) := \big((G_1, x_1) \coprod (G_2, x_2)\big)/\sim,$$

where $\sim$ identifies the points x_1 and x_2. The resulting graph is again a pointed S-labelled graph with distinguished vertex x the image of x_1, x_2. However, it is usually not a S-Schreier graph anymore, as some vertices may have more than one outgoing or ingoing edge for some label $s \in S$. We now apply Stallings folding to obtain a finite partial pointed S-Schreier graph:

$$S(w_1, \ldots, w_n) := \text{fold}\left(\bigvee_{i=1}^{n} S\left(w_i\right)\right).$$

Theorem 1.4.9. *Let $w_1, \ldots, w_n, w \in \mathcal{F}_S$. Then,*

$$\pi_1\left(S(w_1, \ldots, w_n), x\right) = \langle w_1, \ldots, w_n \rangle \leq \mathcal{F}_S..$$

We have that $w \in \langle w_1, \ldots, w_n \rangle$ if and only if $w \in \pi_1(S(w_1, \ldots, w_n), x)$, i.e. if and only if there exists a based loop at x with label w in $S(w_1, \ldots, w_n)$.

Proof. By construction of the folding functor as left adjoint to the forgetful functor, there is a natural map

$$\bigvee_{i=1}^{n} S(w_i) \to S(w_1, \ldots, w_n).$$

By the inductive constructionf of folding, any based loop in $S(w_1, \ldots, w_n)$ lifts to a based loop in $\bigvee_{i=1}^{n} S(w_i)$. Hence, any element of $\pi_1(S(w_1, \ldots, w_n))$ can be represented by a word in the w_i. The converse inclusion is clear by construction, so we obtain the first claim.

The second claim follows directly from the first: w lies in $\langle w_1, \ldots, w_n \rangle$ if and only if it lies in $\pi_1(S(w_1, \ldots, w_n))$, which is equivalent to the existence of a based loop at x with label w in $S(w_1, \ldots, w_n)$. □

1.4.5 Theorems of Hall and Houghton

In order to show the strength of the Schreier graph techniques, we present three important results about subgroups of free groups.

Definition 1.4.10. Let G be a group and $H \leqslant G$ be a subgroup. We say that H is a *free factor* of G if there exists $H' \leqslant G$ such that $H * H' = G$, i.e. more precisely, the natural map $H * H' \to G$ is an isomorphism.

In the abelian setting, we have the following phenomenon: Consider the group $\mathbb{Z}^2$ and the subgroup $N = \{(2n, 0) \mid n \in \mathbb{Z}\} \leq \mathbb{Z}^2$ Does there exist $N' \leqslant \mathbb{Z}^2$ such that $\mathbb{Z}^2 = N \oplus N'$? The answer is obviously negative. However, there is $N' \leqslant \mathbb{Z}^2$ with $N \oplus N' \leqslant \mathbb{Z}^2$ of finite index. Hall's theorem is a non-abelian analogue of this fact for free groups.

Theorem 1.4.11 (M. Hall Jr.)**.** *Let S be a set and $H \leqslant \mathcal{F}_S$ a finitely generated subgroup. Then, H is a free factor in a finite index subgroup $K \leqslant \mathcal{F}_S$.*

Proof. We start by considering the S-Schreier graph corresponding to H. Since H is finitely generated, the core (G, x) of this graph is finite. Since the number of missing outgoing edges with label s is equal to the number of missing incoming edges with label s, we can complete G to a finite S-Schreier graph G' by connecting missing outgoing arrows to missing incoming arrows. Let E be the set of new edges. Then, $\pi_1(G') = \pi_1(G) * \mathcal{F}_E$. Since $\pi_1(G) = H$, we have that H is a free factor of the finite index subgroup $\pi_1(G') \leqslant \mathcal{F}_S$. □

The second result is about intersections of finitely generated subgroups of free groups.

Theorem 1.4.12 (Houghton)**.** *Let S be a set, $H_1, H_2 \leqslant \mathcal{F}_S$ finitely generated. Then $H_1 \cap H_2$ is also finitely generated and*

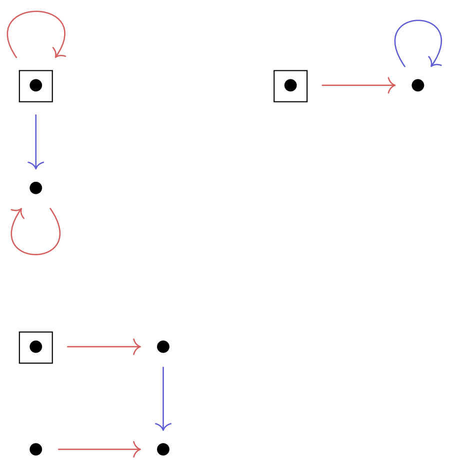
$$\mathrm{rk}(H_1 \cap H_2) - 1 \leq 2 \cdot (\mathrm{rk}(H_1) - 1)(\mathrm{rk}(H_2) - 1).$$

It was long conjectured by Hanna Neumann that the factor 2 is not necessary. And finally, this was proven independently by Joel Friedman [15] and Igor Mineyev [38] around 2010.

Example 1.4.13. Let G_1, G_2 be 2-generated groups and fix surjections $\alpha_i \colon \mathcal{F}_2 \to G_i$. Set $H_1 = \ker(\alpha_1), H_2 = \ker(\alpha_2)$. Now, $\mathrm{rk} H_i - 1 = [\mathcal{F}_2 : H_i](\mathrm{rk}(\mathcal{F}_2) - 1) = |G_i|$. Then $H_1 \cap H_2 = \ker(\alpha_1 \times \alpha_2 : \mathcal{F}_2 \to G_1 \times G_2)$ and we obtain $\mathrm{rk}(H_1 \cap H_2) - 1 = |\operatorname{Im}(\alpha_1 \times \alpha_2)| \leq |G_1||G_2| = (\mathrm{rk}(H_1) - 1)(\mathrm{rk}(H_2) - 1)$.

The proof of Houghton's theorem uses the product of Schreier graphs. Let G, G' be pointed and connected partial S-Schreier graphs. We define $G \times G'$ to be a partial S-Schreier graph on the vertex $V(G) \times V(G')$ with an edge $(x, x') \xrightarrow{a} (y, y')$ if and only if $x \xrightarrow{a} y$ and $x' \xrightarrow{a} y'$.

Example 1.4.14. The upper left graph is G_1, the upper right graph is G_2. Their product $G_1 \times G_2$ is the lower graph.



Lemma 1.4.15. *Let G, G' be connected pointed partial S-Schreier graphs. Then*

$$\pi_1(G \times G') = \pi_1(G) \cap \pi_1(G').$$

Proof. The inclusion $\subseteq$ is clear, since any loop in $G \times G'$ gets projected to a loop in each of the factors. The other inclusion $\supseteq$ follows from the fact that if an element lies in both fundamental groups, there exist loops in each graph with the same labelling. By construction of the product graph, these loops lift to a loop in the product graph with the same labelling. Hence, any element in the intersection of the fundamental groups lifts to a loop in the product graph, proving the claim. $\square$

Let G_1, G_2 be connected pointed partial S-Schreier graphs and assume that $G_1 = \mathrm{core}(G_1)$ and $G_2 = \mathrm{core}(G_2)$. Define $P := \mathrm{core}(G_1 \times G_2) \subset G_1 \times G_2$.

Now for some connected pointed partial S-Schreier graph G with $G = \mathrm{core}(G)$ and $z \in V(G)$, we define $\deg_S(z)$ to be the number of edges incident to z. Also, we set $\xi(z) := \deg_S(z) - 2 \geq 0$.

Lemma 1.4.16. $2(\mathrm{rk}(\pi_1(G)) - 1) = \sum_{z \in V} \xi(z)$.

Proof. The proof is a direct computation using the definition of the rank of the fundamental group and the handshaking lemma:

$$\begin{aligned}\sum_{z \in V} \xi(z) &= \sum_{z \in V} \deg_S(z) - 2 \\ &= 2|E(G)| - 2|V(G)| \\ &= 2\left(|E(G)| - |V(G)|\right) \\ &= 2(\mathrm{rk}(\pi_1(G)) - 1).\end{aligned}$$

□

Proof of Theorem 1.4.12. Let $x \in V(G_1), y \in V(G_2)$. Set $t(x,y) := |S_x \cap T_y|$, where

$$\begin{aligned} S_x &:= \{s \in S^{\pm 1} \mid s \text{ labels an oriented outgoing edge at } x\}, \\ T_y &:= \{s \in S^{\pm 1} \mid s \text{ labels an oriented outgoing edge at } y\}. \end{aligned}$$

Now, we have $\deg_S(x,y) \leq t(x,y) \leq \min\{\deg_S(x), \deg_S(y)\}$ and the following chain of inequalities:

$$t(x,y) - 2 \leq \min\{\xi_{G_1}(x), \xi_{G_2}(y)\} \leq \xi_{G_1}(x)\xi_{G_2}(y).$$

Using the rank formula via vertex degrees, we obtain:

$$\begin{aligned} 2(\mathrm{rk}(\pi_1(P)) - 1) &\leq \sum_{x,y} \max\{t(x,y) - 2, 0\} \\ &\leq \sum_{x,y} \xi_{G_1}(x)\xi_{G_2}(y) \\ &= \left(\sum_x \xi_{G_1}(x)\right)\left(\sum_y \xi_{G_2}(y)\right) \\ &= 2\left(\mathrm{rk}\left(\pi_1(G_1)\right) - 1\right) \cdot 2\left(\mathrm{rk}\left(\pi_1(G_2)\right) - 1\right). \end{aligned}$$

Hence, we obtain

$$\mathrm{rk}(\pi_1(P)) - 1 \leq 2 \cdot (\mathrm{rk}(\pi_1(G_1)) - 1)(\mathrm{rk}(\pi_1(G_2)) - 1).$$

Using the subgroup intersection via product Schreier graph lemma, we conclude that

$$\mathrm{rk}(\pi_1(G_1) \cap \pi_1(G_2)) - 1 \leq 2 \cdot (\mathrm{rk}(\pi_1(G_1)) - 1)(\mathrm{rk}(\pi_1(G_2)) - 1).$$

This completes the proof. □

Chapter 2

Groups with generators and relations

We introduce presentations and the word problem, then develop diagram methods (van Kampen diagrams, curvature, and small cancellation). These tools are applied to Dehn's algorithm and to constructions such as HNN extensions and amalgamated free products, which we use in the examples.

2.1 Generators and relations

2.1.1 Groups defined by generators and relations

We encode a group by generators and relations, and interpret relations through the normal closure in the free group. Let S be a set and $\mathcal{F}_S$ be the free group on the set S. Now, given $R \subset \mathcal{F}_S$, a set of relations, define:

$$\langle S \mid R \rangle \coloneqq \mathcal{F}_S / \langle\!\langle R \rangle\!\rangle \quad \text{with} \quad \langle\!\langle R \rangle\!\rangle = \bigcap_{\substack{H \lhd \mathcal{F}_S \\ R \subset H}} H.$$

We call $\langle\!\langle R \rangle\!\rangle$ the normal closure of R or the normal subgroup generated by R.

Lemma 2.1.1. *Let S be a set and $R \subset \mathcal{F}_S$. Then,*

$$\langle\!\langle R \rangle\!\rangle = \left\{ \prod_{i=1}^{k} g_i r_i^{\epsilon_i} g_i^{-1} \mid k \in \mathbb{N}, g_1, \ldots, g_k \in \mathcal{F}_S, r_1, \ldots, r_k \in R, \epsilon_1, \ldots, \epsilon_k \in \{\pm 1\} \right\}.$$

Proof. Let N be the set on the right-hand side. Every normal subgroup $H \lhd \mathcal{F}_S$ that contains R also contains all conjugates $g r^{\pm 1} g^{-1}$ and hence every finite product of such conjugates, so $N \subseteq H$ for all such H. Therefore $N \subseteq \bigcap_{H \lhd \mathcal{F}_S,\, R \subseteq H} H = \langle\!\langle R \rangle\!\rangle$.

Conversely, N is a normal subgroup: it contains R, is closed under products and inverses by construction, and is closed under conjugation because $g\big(\prod_i h_i r_i^{\epsilon_i} h_i^{-1}\big) g^{-1}$ is again a product of conjugates of elements of R. Hence $\langle\!\langle R \rangle\!\rangle \subseteq N$. □

Note that there exists a homomorphism $\pi : \mathcal{F}_S \to \langle S \mid R \rangle$. Clearly, $\pi(r) = \mathbb{1} \in \langle S \mid R \rangle$ for all $r \in \langle R \rangle$.

Example 2.1.2. Consider $S = \{x, y\}$ and $R = \{xyx^{-1}y^{-1}\}$. We claim that the group

$$G \coloneqq \langle x, y \mid xyx^{-1}y^{-1} \rangle$$

is isomorphic to $\mathbb{Z}^2$. Then $\pi(xyx^{-1}y^{-1}) = 1$ and hence $\pi(x)\pi(y) = \pi(y)\pi(x)$. In particular, every element of G can be written as $\pi(x)^a\pi(y)^b$ for some $a, b \in \mathbb{Z}$. Moreover the obvious surjection $\sigma : \mathcal{F}_2 \to \mathbb{Z}^2$ given by $x \mapsto (1, 0), y \mapsto (0, 1)$ factors through π. This implies that the induced map $\overline{\sigma} : G \to \mathbb{Z}^2$ is an isomorphism.

A group defined by generators and relations comes with a universal property similar to Theorem 1.1.10.

Lemma 2.1.3. *Let S be a set, $R \subset \mathcal{F}_S$ and G a group. A map $\varphi : S \to G$ extends to a homomorphism $\overline{\varphi} : \langle S \mid R\rangle \to G$ if and only if the extension $\varphi' : \mathcal{F}_S \to G$ satisfies $\varphi'(r) = \mathbb{1}_G$ for all $r \in R$.*

Proof. The homomorphismφ' is unique and φ' factors through the quotient $\langle S \mid R\rangle$ if and only if $\varphi'|_{\langle\langle R\rangle\rangle}$ is trivial and if and only if $\varphi'(r) = \mathbb{1} \in G$ for all $r \in R$. □

Definition 2.1.4. A group is called *finitely presented* if and only if there exists a finite set S and a finite $R \subset \mathcal{F}_S$ such that $G \cong \langle S \mid R\rangle$.

Despite the apparent simplicity of the definition of a group by generators and relations, such groups can be extremely complicated. We now turn to the algorithmic aspects of such groups.

2.1.2 Word problem

Given S finite and $R \subset \mathcal{F}_S$, the word problem asks whether a word $w \in \mathcal{F}_S$ represents the identity in the quotient $\langle S|R\rangle$. This simple formulation already hides deep algorithmic phenomena. It was shown by Novikov and Boone in the 1950s that there exist finitely presented groups with undecidable word problem; see [33] for the construction. We first give a criterion for decidability of the word problem in terms of the Dehn function.

Definition 2.1.5. Assume S and R are finite. For $w \in \langle\langle R\rangle\rangle$ define

$$\mathrm{Area}(w) := \min\{k \in \mathbb{N} \mid w = \prod_{i=1}^{k} g_i r_i^{\epsilon_i} g_i^{-1}\}.$$

For $n \in \mathbb{N}$ set $B_S(n) := \{w \in \mathcal{F}_S \mid |w|_S \leq n\}$. Set

$$\delta(n) := \max\{\mathrm{Area}(w) \mid w \in B_S(n) \cap \langle\langle R\rangle\rangle\}.$$

This is the *Dehn function* of the presentation.

Theorem 2.1.6. *For a finitely presented group, the word problem is decidable if and only if the Dehn function δ is computable.*

Idea of proof: If δ is computable, then given w compute $\delta(|w|)$ and enumerate all products of at most $\delta(|w|)$ conjugates of relators. Comparing in the free group decides whether $w \in \langle\langle R\rangle\rangle$. This part of the proof is incomplete as yet as we have a priori no control on the length of the conjugating elements g_i in the products $\prod_{i=1}^{k} g_i r_i^{\epsilon_i} g_i^{-1}$. In order to make this argument work, one needs to use van Kampen diagrams to show that there exists a bound on the lengths of the conjugating elements depending only on $|w|$ and $\delta(|w|)$. We will complete the argument after introducing van Kampen diagrams in the next section.

Conversely, if the word problem is decidable, then for each w one can search for the least k such that $w = \prod_{i=1}^{k} g_i r_i^{\epsilon_i} g_i^{-1}$ by enumerating such products and checking equality in the free group. Since there are finitely many words of length $\leq n$, this allows one to compute $\delta(n)$. □

Example 2.1.7. For $\mathbb{Z}^2 = \langle x, y \mid [x, y]\rangle$ consider

$$w_n = [x^n, y^n] = x^n y^n x^{-n} y^{-n}.$$

The word length of w_n is $4n$. The standard $n \times n$ square diagram shows $\mathrm{Area}(w_n) \leq n^2$, and each relator contributes one square in such a diagram, so $\mathrm{Area}(w_n) \leq n^2$ and one can show that this is actually an equality. Hence, in this case the Dehn function of $\mathbb{Z}^2$ is quadratic.

Behind to preceding idea of proof of Theorem 2.1.6 is the following observation.

Lemma 2.1.8. *Let S be a finite set and $R \subset \mathcal{F}_S$ be finite. Then, there is an enumeration of the elements in $\langle\langle R\rangle\rangle$.*

Proof. Fix $k \in \mathbb{N}$. There are finitely many words in $\mathcal{F}_S$ of length at most k, hence finitely many choices of $g_i \in B_S(k)$, $r_i \in R$, and $\epsilon_i \in \{\pm 1\}$ for products of length at most k. List all products

$$\prod_{i=1}^{m} g_i r_i^{\epsilon_i} g_i^{-1} \qquad (1 \leq m \leq k),$$

and then let k increase. Every element of $\langle\langle R\rangle\rangle$ has some representation of this form, so the resulting process enumerates $\langle\langle R\rangle\rangle$. □

The obstruction in the word problem is that this enumeration only lists trivial words. One needs either a procedure to enumerate the complement of $\langle\langle R\rangle\rangle$ or an effective bound on how long one must search before a representation as a product of conjugates appears.

So a possible algorithm would just run through an enumeration of $\langle\langle R\rangle\rangle$ and its complement in parallel. No matter if $w \in \mathcal{F}_S$ is trivial or not, this algorithm would stop in finite time and decide the question. This strategy can implemented for example if the group is residually finite.

Theorem 2.1.9. *Let S be a finite set and $R \subset \mathcal{F}_S$ finite. If $\langle S|R\rangle$ is residually finite, then the word problem is decidable.*

Proof. Let $w \in \mathcal{F}_S$. Run two procedures in parallel. First, enumerate $\langle\langle R\rangle\rangle$ and stop if w appears. This detects $w = 1$ in the quotient. Second, for each $n \in \mathbb{N}$ enumerate S-tuples in S_n, producing a homomorphisms $\varphi_\sigma : \mathcal{F}_S \to S_n$. Whenever $\varphi_\sigma(r) = 1 \in S_n$ for all $r \in R$, compute $\varphi_\sigma(w)$ and stop if it is non-trivial. By Cayley's theorem this enumerates all finite quotients. Thus, if $\langle S|R\rangle$ is residually finite, every non-trivial element is detected by some S-tuple in S_n for some $n \in \mathbb{N}$ □

Recall that we also proved that the subgroup membership problem is decidable for $\mathcal{F}_S$, see Theorem 1.4.9. However, this was only for finitely generated subgroups and a normal subgroup is finitely generated only if it is of finite index.

2.2 Diagram methods and Dehn's algorithm

In this section we develop combinatorial topology tools to study groups defined by generators and relations. The main tools are van Kampen diagrams, curvature, and small cancellation theory. As an application, we present Dehn's algorithm to solve the word problem in certain groups.

2.2.1 van Kampen diagrams

We take a naive combinatorial approach to 2-complexes. A more thorough treatment can be found in [33]. Our main object of study is the van Kampen diagram, which encodes relations in a group presentation geometrically.

Definition 2.2.1. Let S be a finite set and $R \subset \mathcal{F}_S$. A van Kampen diagram over the presentation $\langle S|R\rangle$ is a pointed finite oriented planar S-labeled graph together with 2-cells attached along the boundary of each bounded face reading relators in $R^{\pm 1}$. We assume that the base point lies on the boundary of the outer face.

Each van Kampen diagram D has a number of vertices V_D, edges E_D, and faces F_D. We define the Euler characteristic as follows: $\chi(D) \coloneqq V_D - E_D + F_D$. Our first aim is to show that van Kampen diagrams always have Euler characteristic 1.

In order to uniformize our language, we speak about cells in dimension $0, 1, 2$ rather than vertices, edges, and faces. A 0-cell is a vertex, a 1-cell is an edge, and a 2-cell is a face.

Definition 2.2.2. A *free face* is a cell f that lies in the boundary of a unique higher-dimensional cell. In a van Kampen diagram, this means either a free 1-face: an edge contained in the boundary of a unique 2-cell, or a free 0-face: a vertex of degree 1. An *elementary collapse* removes a free face f together with the unique cell containing it.

Lemma 2.2.3. *If $X \downarrow X'$ is an elementary collapse, then $\chi(X) = \chi(X')$.*

Proof. In a 2-complex an elementary collapse removes one edge and one face, or one vertex and one edge. In each case the quantity $V_X - E_X + F_X$ is unchanged. □

Lemma 2.2.4. *Any van Kampen diagram that is not a point contains a free face and hence admits an elementary collapse.*

Proof. If the diagram has no 2-cells, it is a finite graph. Since it cannot contain any cycles, it is a tree and therefore has vertices of degree 1, which are free 0-faces.

Assume the diagram has at least one 2-cell. The outer boundary cycle of a van Kampen diagram is nonempty, and unless there is a free 0-cell, each edge on that boundary is contained in exactly one 2-cell. Hence any boundary edge is a free 1-face. □

Corollary 2.2.5. *If D is a van Kampen diagram, then $\chi(D) = 1$.*

Proof. Repeatedly collapse free faces until a point remains. Euler characteristic is preserved by Lemma 2.2.3, so $\chi(D) = 1$. □

Definition 2.2.6. Let D be a van Kampen diagram over the presentation $\langle S|R\rangle$. The boundary label of the diagram is the sequence of labels read counterclockwise along the outer boundary cycle starting at the base point. After reduction, this defines a word in $\mathcal{F}_S$ called the *boundary word* of the diagram.

Lemma 2.2.7. *A word $w \in \mathcal{F}_S$ is the boundary word of some van Kampen diagram over the presentation $\langle S|R\rangle$ if and only if $w \in \langle\langle R\rangle\rangle$.*

Proof. Read the boundary word and push it across the diagram face by face. Each 2-cell contributes a conjugate of a relator, while adjacent inverse labels cancel. Hence the boundary word is a product of conjugates of relators.

Conversely, given a product of conjugates of relators, build a van Kampen diagram by attaching 2-cells for each conjugate of a relator along its boundary, and then identifying edges with inverse labels. □

We are now able to give a complete proof of Theorem 2.1.6 using van Kampen diagrams.

Completion of proof of Theorem 2.1.6. Assume that the Dehn function δ is computable. Since δ is computable, we can compute $\delta(|w|)$. Given a word $w \in \mathcal{F}_S$, we want to decide whether w represents the identity in $\langle S|R\rangle$. Note that by Lemma 2.2.7, w represents the identity if and only if there exists a van Kampen diagram D over $\langle S|R\rangle$ with boundary word $w \in \mathcal{F}_S$. After basic reductions, we can assume that the sequence of labels around the boundary of D reads exactly w without any cancellations. With this additional condition, the total number of edges on the boundary of D is exactly $|w|$ and thus the total number of edges in D is bounded above by $1/2|w|+1/2\sum_{f\in F_D}|\partial f| \leq 1/2(|w|+\max_{r\in R}|r|\cdot\delta(|w|))$. Hence, the size of the diagram is controlled by the number of faces and the length of the relators. If such a van Kampen diagram D exists, it can be found by a finite enumerative search. And, if we find such a diagram, then w represents the identity; otherwise it does not. This completes the proof. □

Note that, in general, this does not mean that we can effectively construct the required van Kampen diagrams. The number of possible diagrams grows extremely fast with the number of faces. On top of that, a computable function may grow very fast, making the search infeasible in practice.

2.2.2 Combinatorial Gauss–Bonnet

In this section we develop a combinatorial version of the Gauss–Bonnet theorem for angled 2-complexes. This will allow us to control the structure of van Kampen diagrams.

Definition 2.2.8. An *angled* van Kampen diagram is a van Kampen diagram together with an assigned angle to each *corner*, i.e., a pair consisting of an incident vertex and 2-cell. We denote the angle at corner c by $\measuredangle c \in [0, \pi)$ and the set of corners of a 2-cell f by $\mathrm{C}(f)$. For a face f, its perimeter $|\partial f|$ counts the number of boundary edges.

Definition 2.2.9. For a 2-cell f set

$$\mathrm{Curv}(f) := \sum_{c\in\mathrm{C}(f)} \measuredangle c - (|\partial f| - 2)\pi.$$

For a vertex v, let $\mathrm{link}(v)$ be the graph obtained by intersecting a small sphere around v with the van Kampen diagram; its edges correspond to corners at v. Define

$$\mathrm{Curv}(v) := \pi(2 - \chi(\mathrm{link}(v))) - \sum_{e\in E(\mathrm{link}(v))} \measuredangle e.$$

Theorem 2.2.10 (Gauss-Bonnet)**.** *For an angled van Kampen diagram,*

$$\sum_f \mathrm{Curv}(f) + \sum_v \mathrm{Curv}(v) = 2\pi.$$

Proof. Let Ψ be the total angle sum over all corners. Then

$$\sum_f \mathrm{Curv}(f) = \Psi - \sum_f (|\partial f| - 2)\pi.$$

For each vertex v, the graph link(v) has deg(v) vertices and $|C(v)|$ edges, where deg(v) is the degree of v in X and $|C(v)|$ is the number of corners at v. Thus $\chi(\text{link}(v)) = \deg(v) - |C(v)|$. Hence, we obtain

$$\sum_v \text{Curv}(v) = \sum_v \pi(2 - \chi(\text{link}(v))) - \Psi = 2\pi V_X - \pi \sum_v \deg(v) + \pi \sum_v |C(v)| - \Psi.$$

Since $\sum_v \deg(v) = 2E_X$ and $\sum_v |C(v)| = \sum_f |\partial f|$, we obtain

$$\sum_f \text{Curv}(f) + \sum_v \text{Curv}(v) = 2\pi V_X - 2\pi E_X + 2\pi F_X = 2\pi.$$

This concludes the proof using Corollary 2.2.5. □

2.2.3 Small cancellation and shells

We assume $\langle S|R\rangle$ is a finite presentation in which all relators are cyclically reduced and R is symmetrized, i.e. closed under cyclic conjugation and inverses. If that is not the case, we can replace R by the symmetrized set of cyclically reduced conjugates of elements of R without changing the group. For example, the relation $aba^{-1}b^{-1}$ can be replaced by the symmetrized set:

$$\{aba^{-1}b^{-1},\ bab^{-1}a^{-1},\ a^{-1}b^{-1}ab,\ b^{-1}a^{-1}ba,$$
$$ab^{-1}a^{-1}b,\ b^{-1}aba^{-1},\ a^{-1}bab^{-1},\ b^{-1}a^{-1}ba\}.$$

Definition 2.2.11. A *piece* is a non-trivial word that occurs as a prefix of two distinct relators in R. The presentation satisfies the $C'(\lambda)$ small cancellation condition if for any piece p contained in a relator r one has

$$|p| < \lambda|r|.$$

Definition 2.2.12. A *cancellable pair* in a van Kampen diagram is a pair of adjacent 2-cells that share a path labeled by a word and its inverse; removing both cells and the shared path yields another diagram. A diagram is *reduced* if it contains no cancellable pair.

An *arc* is a maximal connected subgraph whose vertices all have degree ≤ 2. An edge is *internal* if it lies in the boundary of two 2-cells; an *internal arc* is an arc consisting of internal edges. An *arc-reduced* diagram is obtained from a reduced diagram by replacing each arc by a single oriented edge labeled by its word. In an arc-reduced diagram there are no vertices of degree 2.

Lemma 2.2.13. *In a reduced van Kampen diagram over a symmetrized presentation, the label of every internal arc is a piece.*

Proof. An internal arc is shared by two distinct 2-cells. Reading along the boundaries of those cells gives two occurrences of the arc label in relators from $R^{\pm 1}$. If those occurrences were inverse in the same position, the cells would form a cancellable pair, contradicting reducedness. Hence the label is a piece. □

Definition 2.2.14. Let D be an arc-reduced van Kampen diagram homeomorphic to a disc. A vertex is *interior* if it lies in the interior of D, otherwise it is *exterior*. A 2-cell is *internal* if all of its edges are internal, and *external* otherwise.

A 2-cell f is an *i-shell* if its boundary decomposes as

$$\partial f = pq,$$

where p is a single external arc and q is a path of i internal arcs.

Lemma 2.2.15. *In a $C'(1/6)$ presentation, every internal 2-cell in an arc-reduced diagram has at least 7 sides.*

Proof. Each side of an internal 2-cell is an internal arc and thus labels a piece by Lemma 2.2.13. By $C'(1/6)$ each such arc has length strictly less than $|r|/6$, where r is the relator labeling the cell. Hence at least 7 arcs are required to obtain total boundary length $|r|$. □

Proposition 2.2.16. *Let D be an arc-reduced van Kampen diagram homeomorphic to a disc over a $C'(1/6)$ presentation with at least two 2-cells. Then D contains an i-shell with $i \in \{1, 2, 3\}$.*

Proof. Assign angles so that every vertex has zero curvature: for an interior vertex v set each incident corner angle to $2\pi/\deg(v)$, and for an exterior vertex set each incident corner angle to $\pi/(\deg(v) - 1)$. After freely reducing the boundary word, we may assume D has no degree 1 vertices. Since D is arc-reduced, it has no degree 2 vertices, so $\deg(v) \geq 3$, and interior angles are at most $2\pi/3$ while exterior angles are at most $\pi/2$.

By Lemma 2.2.15, every internal 2-cell has at least 7 sides and thus negative curvature, since

$$\mathrm{Curv}(f) \leq k(2\pi/3) - (k-2)\pi = (6-k)(\pi/3) \leq -\pi/3.$$

Consider an external 2-cell that is not a shell. Then its boundary meets the exterior in at least two disjoint arcs, so it has at least four exterior corners. Hence its angle sum is at most $4(\pi/2) + (k-4)(2\pi/3)$ for k sides, which is at most $(k-2)\pi$ for $k \geq 4$, so its curvature is non-positive as well.

Therefore all positive curvature must come from external cells whose boundary meets the exterior in exactly one arc. Such a cell is an i-shell. If $i \geq 4$, then the cell has $k = i + 1 \geq 5$ sides with two exterior corners, so its angle sum is at most

$$2(\pi/2) + (k-2)(2\pi/3) = (2k-1)(\pi/3) \leq (k-2)\pi$$

and its curvature is non-positive. Since $\chi(D) = 1$, total curvature is $2\pi > 0$ by Theorem 2.2.10, so there must exist an i-shell with $i \in \{1, 2, 3\}$. □

Corollary 2.2.17. *Let D be a arc-reduced van Kampen diagram over a $C'(1/6)$ presentation. Then the boundary label of D contains a subword that is more than half of a cyclic permutation of some relator in $R^{\pm 1}$.*

Proof. If D has a single 2-cell, then the boundary word is a relator and the conclusion is immediate. Otherwise, Proposition 2.2.16 gives an i-shell with $i \leq 3$. The internal path is the concatenation of at most 3 pieces, each of length $< |r|/6$, hence has length $< |r|/2$. Therefore the external arc of that shell has length $> |r|/2$ and lies on the boundary of D. □

We will see in the next section that this corollary implies the existence of a naive solution to the word problem in $C'(1/6)$ groups.

2.2.4 Dehn's algorithm

We now describe Dehn's algorithm and, using the diagram methods above, prove a small-cancellation criterion for when it succeeds. In the following let S be finite and $R \subset \mathcal{F}_S$ be a finite set of relations. Assume that R is symmetrized. Let us look at the most naive algorithm: S finite, $R \subset \mathcal{F}_S$ finite. Start with $w \in \mathcal{F}_S$. The algorithm has only one step that is applied as long as possible: If $w = w_1 u w_2$ and there is $r \in R^{\pm 1}$ with $r = uv$ and $|u| > |v|$, then replace u by v^{-1} to obtain $w' := w_1 v^{-1} w_2$, and freely reduce. We call that operation a Dehn reduction. Note that w and w' represent the same group element since $uv^{-1} = r \in \langle\langle R \rangle\rangle$ and $|w'| < |w|$.

Definition 2.2.18. We say that the algorithm works for a presentation if every trivial word admits repeated Dehn reductions that terminate at the empty word.

Note that if the Dehn algorithm works for a presentation, then the word problem is decidable since one can just run the algorithm and see if it terminates at the empty word; in fact we obtain $\delta(n) \leq n$.

Definition 2.2.19. A word is *Dehn-reduced* if it is freely reduced and contains no subword longer than half of any cyclic permutation of a relator from $R^{\pm 1}$.

Definition 2.2.20. Let D be a van Kampen diagram. A vertex is *semi-exterior* if its link is not connected. Equivalently, the vertex is not contained in the boundary of a single disc component of D.

Theorem 2.2.21 (Greendlinger, see [17])**.** *If $\langle S|R \rangle$ satisfies $C'(1/6)$, then Dehn's algorithm solves the word problem. That is, a word $w \in \mathcal{F}_S$ represents the identity in $\langle S|R \rangle$ if and only if it is Dehn-reduced and non-trivial.*

Proof. Let w be a freely reduced word representing the identity and let D be a reduced van Kampen diagram with boundary label w. If D is a disc, Corollary 2.2.17 yields a subword of w longer than half a relator; Dehn's algorithm shortens w. Iterating and inducting on the number of faces shows that w reduces to the empty word.

If D is not a disc, then it contains a semi-exterior vertex, and D decomposes into disc components meeting along a tree. At least one disc component contributes a boundary subword of w; applying Corollary 2.2.17 to that component again yields a Dehn reduction. Thus the same induction applies. The converse is immediate since each Dehn reduction preserves the group element. □

Apart from solving the word problem, $C'(1/6)$ presentations are quite useful since many questions about the group can be answered easily. As an example, we show that such groups are infinite.

Theorem 2.2.22. *If $\langle S|R \rangle$ satisfies $C'(1/6)$ and $|S| \geq 2$ and R finite, then the group is infinite and contains an element of infinite order.*

Proof. Let us assume that R is symmetrized and that all relators are cyclically reduced. We consider the set F of prefixes of elements $r \in R$ of length greater $|r|/2$. Consider the set $\mathcal{L}$ of reduced words in $S \cup S^{-1}$ without any subwords in that set. By $C'(1/6)$-cancellation condition, such words are Dehn-reduced and hence non-trivial by Theorem 2.2.21.

We claim that any word $w \in \mathcal{L}$ can be extended by some letter from $S \cup S^{-1}$ to a longer word in $\mathcal{L}$. Let $x \in S \cup S^{-1}$ be the last letter in w. Now, suppose there are two distinct letters $a \neq b$ with $a \neq x^{-1}$, $b \neq x^{-1}$ such that $wa \notin \mathcal{L}$ and $wb \notin \mathcal{L}$. Since $w \in \mathcal{L}$, any forbidden subword in wa (resp. wb) must end at the last letter. Hence there exist suffixes v_a of wa and v_b of wb in F. Choose $r_a, r_b \in \mathcal{R}$ with v_a a prefix of r_a, v_b a prefix of r_b and

$$|v_a| > \tfrac{1}{2}|r_a|, \qquad |v_b| > \tfrac{1}{2}|r_b|.$$

Let $\ell := \min\{|v_a|, |v_b|\}$ and let p be the common prefix of v_a and v_b of length $\ell - 1$. It follows that p is freely reduced and a piece of both r_a and r_b. Let $L := \min\{|r_a|, |r_b|\}$. Then $|v_a| \geq \ell$ and $|v_b| \geq \ell$ give $\ell > \frac{1}{2}L$, hence

$$|p| = \ell - 1 > \tfrac{1}{2}L - 1.$$

If $L = 2$, then $\ell > \frac{1}{2}L = 1$ forces $\ell \geq 2$, hence $|p| \geq 1$, contradicting $C'(1/6)$ because $|p| < L/6 = 1/3$. If $L \geq 3$, then $\frac{1}{2}L - 1 \geq \frac{1}{6}L$, so $|p| > \frac{1}{6}L$, again contradicting $C'(1/6)$. Since $2|S| - 2 \geq 1$, there is always at least one allowed letter a with $wa \in \mathcal{L}$. Hence $\mathcal{L}$ contains words of arbitrarily large length.

Let $M := \max\{|f| : f \in F\}$, which is finite since R is assumed to be finite. If there exists arbitrarily long words in $\mathcal{L}$, we can find a word w that contains a subword v of length M twice, i.e. $w = w_1 v w_2 v w_3$ for some words w_1, w_2, w_3. Since vw_2v does not contain a forbidden subword and $|v| = M$, we conclude that $(vw_2)^n$ does not contain any forbidden subword for any $n \in \mathbb{N}$. Hence, the group contains an element of infinite order. □

From a more abstract point of view, the set of Dehn-reduced words forms a regular language since it is characterized by a finite set of forbidden subwords and can thus be recognized by a finite state automaton. Hence, $C'(1/6)$ groups are an example of groups with a regular set of normal forms. Moreover, the argument in the proof of Theorem 2.2.22 is essentially an application of the pumping lemma for regular languages.

2.3 Examples

2.3.1 HNN-construction

This is the Higman–Neumann–Neumann construction. This construction allows one to build new groups from old ones by forcing two isomorphic subgroups to be conjugate. It is a fundamental construction in combinatorial group theory and has many applications, for example in the proof of the Higman embedding theorem.

Definition 2.3.1. Given $G, H_1, H_2 \leqslant G$ where $H_1 \cong H_2$ via α. Then define

$$G*_\alpha = \mathrm{HNN}(G, H_1, H_2, \alpha) := \left\langle G, t \mid th_1t^{-1} = \alpha(h_1) \text{ for all } h_1 \in H_1 \right\rangle.$$

Essentially, we introduce a new letter t in order to force the two subgroups to be conjugate. In practice, one includes the relators of G together with the conjugacy relations, as in the next proposition.

It holds that $BS(n,m) = \mathrm{HNN}(\mathbb{Z}, n\mathbb{Z}, m\mathbb{Z}, n \mapsto m)$. However, this is no surprise; it is pretty much by definition. We will see below, via Britton's Lemma (Theorem 2.3.3), that the base group G and the stable letter t embed in the HNN extension.

Proposition 2.3.2. *Let $G = \langle X \mid R \rangle$ and assume $H_1, H_2 \leqslant G$ are finitely generated by $S \subset H_1$ and $\alpha(S) \subset H_2$, where $\alpha : H_1 \xrightarrow{\sim} H_2$. Then*

$$G*_\alpha = \left\langle X \cup \{t\} \,\middle|\, R \cup \{t\widetilde{s}t^{-1}\widetilde{\alpha(s)}^{-1} \mid s \in S\} \right\rangle$$

where $\widetilde{s} \in \mathcal{F}_X$ is a lift of $s \in H_1$.

Proof. The relations of G are included by construction, and the extra relations force conjugacy between H_1 and H_2. Any homomorphism out of $G*_\alpha$ is determined by its restriction to G and the image of t, so the stated presentation follows from the universal property of the HNN extension. □

Theorem 2.3.3 (Britton)**.** *Let $G*_\alpha$ be an HNN-extension as above. A word*

$$w = g_0 t^{\epsilon_1} g_1 \cdots t^{\epsilon_n} g_n$$

with $g_i \in G$ and $\epsilon_i \in \{\pm 1\}$ represents a non-trivial element if it is reduced and contains no pinch*, i.e. no subword of the form tat^{-1} with $a \in H_1$ or $t^{-1}bt$ with $b \in H_2$.*

Proof. Assume w is reduced and has no pinch. Suppose $w = 1$ in $G*_\alpha$ and let D be a reduced van Kampen diagram with boundary label w. Each HNN relator has the form $tut^{-1}v^{-1}$ (or its inverse), so each such 2-cell has two edges labeled by $t^{\pm 1}$. A *t–band* is a maximal chain of 2-cells where consecutive cells share a t–edge. Since D is reduced, a t–band cannot start and end on the same edge; it must run from one boundary t–edge to another.

Follow the t–band beginning at the first $t^{\pm 1}$ in the boundary word. Reading the side labels of the band shows that the subword between these two boundary t–edges represents an element of H_1 (if the band starts with t and ends with t^{-1}) or of H_2 (if it starts with t^{-1} and ends with t). Hence the boundary word contains a pinch, contradicting the assumption. Therefore $w \neq \mathbb{1}$. □

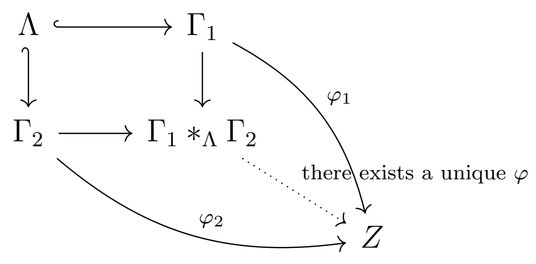
Corollary 2.3.4. *The natural map $G \hookrightarrow G*_\alpha$ is injective, and the subgroup $\langle t \rangle$ is infinite cyclic.*

Proof. A non-trivial element of G is already in Britton-reduced form, so it cannot become trivial in $G*_\alpha$. Similarly, t^n is Britton-reduced for $n \neq 0$, hence non-trivial. □

2.3.2 Amalgamated free products

Amalgamated free products generalize free products by gluing along a common subgroup. They appear naturally when studying graphs of groups and fundamental groups of spaces with identified subspaces.

Definition 2.3.5. Consider groups Γ_1, Γ_2 both with a subgroup isomorphic to Λ. Then $\Gamma_1 *_\Lambda \Gamma_2$ is the universal group "containing" Γ_1, Γ_2 with Λ identified.



$\Gamma_1 * \Gamma_2$ is also called the pushout of the diagram

$$\begin{array}{ccc} \Lambda & \hookrightarrow & \Gamma_1 \\ \downarrow & & \\ \Gamma_2 & & \end{array}$$

$$\Gamma_1 *_\Lambda \Gamma_2 = \left\langle X_1 \cup X_2 \mid R_1 \cup R_2 \cup \{\widetilde{\iota_1(\lambda)}\widetilde{\iota_2(\lambda)}^{-1} \mid \lambda \in \Lambda\} \right\rangle.$$

Proposition 2.3.6. *Let $G = G_1 *_H G_2$ be an amalgamated free product. Choose right coset representatives S_i for H in G_i with $1 \in S_i$. Every element of G can be written uniquely as*

$$hs_1 s_2 \cdots s_n,$$

where $h \in H$, each $s_j \in S_{\varepsilon_j} \setminus \{1\}$, and consecutive s_j, s_{j+1} come from different factors. Such a word is called reduced. *A reduced word with $n \geq 1$ represents a non-trivial element.*

Proof. Consider the set of reduced words and let G_1 and G_2 act by left multiplication followed by reduction. These actions agree on H and generate a transitive action of G. The stabilizer of the empty word is H, which gives the normal form and uniqueness. If a reduced word with $n \geq 1$ represented the identity, it would fix the empty word, contradicting uniqueness. □

Corollary 2.3.7. *Let $g_1 \in G_1 \setminus H$ and $g_2 \in G_2 \setminus H$ be such that $\langle g_1 \rangle$ and $\langle g_2 \rangle$ are infinite and intersect H trivially. Then g_1 and g_2 generate a free subgroup of G.*

Proof. Any non-trivial reduced word in $g_1^{\pm 1}$ and $g_2^{\pm 1}$ is reduced in the sense of Proposition 2.3.6, hence non-trivial. □

2.3.3 Solvable and nilpotent groups

Definition 2.3.8. Let $G^{(0)} := G$ and $G^{(k+1)} := [G^{(k)}, G^{(k)}]$ be the derived series. The group G is called *solvable* if $G^{(n)} = \{\mathbb{1}\}$ for some n. The smallest such n is the *derived length.*

Definition 2.3.9. Define the lower central series by $\gamma_1(G) := G$ and $\gamma_{k+1}(G) := [G, \gamma_k(G)]$. The group G is *nilpotent* if $\gamma_n(G) = \{\mathbb{1}\}$ for some n, and the smallest such n is the *nilpotency class.*

Example 2.3.10. The affine group of the line

$$\mathrm{Aff}(\mathbb{R}) = \{x \mapsto ax + b \mid a \in \mathbb{R}_{>0},\ b \in \mathbb{R}\}$$

is solvable of derived length 2. The subgroup of translations is normal and abelian, and the quotient by that subgroup is isomorphic to $(\mathbb{R}_{>0}, \cdot)$, which is abelian as well. Restricting to $a \in \langle p \rangle$ for some $p > 1$ and $b \in \mathbb{Z}[1/p]$ produces the discrete affine (ax+b) group that also admits the same two-step derived series.

Example 2.3.11. The (discrete) Heisenberg group

$$H = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \middle| \ x, y, z \in \mathbb{Z} \right\}$$

is nilpotent of class 2. Writing

$$X = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad Y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad Z = [X, Y] = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

the group is generated by X, Y with center $\langle Z \rangle$, so it is a central extension of $\mathbb{Z}^2$ by $\mathbb{Z}$. This matrix representation exhibits H as a subgroup of $GL_3(\mathbb{Z})$ and makes it manifestly nilpotent with the lower central series terminating at $\langle Z \rangle$.

2.3.4 Baumslag–Solitar groups

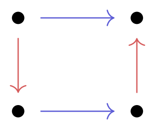
We now illustrate the preceding definitions with one-relator groups and HNN extensions. See [3] for the original reference on Baumslag–Solitar groups and [33] for more details on one-relator groups. Let us take a look at the definition:

$$BS(n, m) \coloneqq \left\langle a, b \mid ba^n b^{-1} a^{-m} \right\rangle .$$

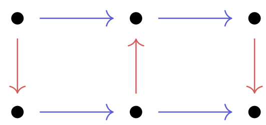
So in $BS(n, m)$ a^n and a^m are conjugate via b. They fall into the general class of one-relator groups. Its clear that $BS(1, 1) = \mathbb{Z}^2$.

$$BS(1, -1) = \left\langle a, b \mid bab^{-1}a \right\rangle .$$

Picture of the relator:



Glue this with its inverse to form:



This is a van Kampen diagram, so its boundary word is also in $\langle\langle R \rangle\rangle$. This is $ab^2a^{-1}b^{-2}$. Therefore, $N \coloneqq \langle a, b^2 \rangle \leqslant BS(1, -1)$ is abelian and normal in $BS(1, -1)$. Indeed, we have the following conjugation relations:

$$\begin{aligned} a \cdot a \cdot a^{-1} &= a \in N \\ a \cdot b^2 \cdot a^{-1} &= b^2 \in N \\ b \cdot a \cdot b^{-1} &= a^{-1} \in N \\ b \cdot b^2 \cdot b^{-1} &= b^2 \in N. \end{aligned}$$

Now, what could be the quotient? $BS(1, -1)/N = \mathbb{Z}/2\mathbb{Z}$. This is because only b survives. Now it could also be trivial, but you can make sure that $b \notin N$. So we can try to recover $BS(1, -1)$ as a semidirect product. It turns out that

$$BS(1, -1) = \mathbb{Z} \ltimes_\alpha \mathbb{Z}$$

with $\alpha(a)(b) = (-1)^a b$.

$BS(2,1) = \langle a, b \mid bab^{-1} = a^2 \rangle$. Let us find some nice subgroups. $\langle a \rangle$ is certainly not normal, since:

$$b^{-1} \langle a \rangle b \nleq \langle a \rangle .$$

This is because $b^{-1}ab$ is a sort of square root of a, since it squares to a. Still, let us assume that $\langle a \rangle = \mathbb{Z}$. Let us take $\alpha : \mathbb{Z} \to \mathrm{Aut}(\mathbb{Z}[\frac{1}{2}])$, $\alpha(1) = \times 2$. Take a look at $\mathbb{Z} \ltimes_\alpha \mathbb{Z}[\frac{1}{2}]$. Define $BS(1,2) \to \mathbb{Z} \ltimes_\alpha \mathbb{Z}[\frac{1}{2}]$, $a \mapsto (0,1)$, $b \mapsto (1,0)$. Now

$$bab^{-1} \mapsto (1,0)(0,1)(-1,0) = (0, \alpha_1(1)) = (0,2) = a^2.$$

So this is a well defined homomorphism (since it respects the relators). This is surjective: We map to $(1,0)$ already, and we map to $(0, \frac{1}{2})$ via $b^{-1}ab$:

$$b^{-1}ab \mapsto (-1,0)(0,1)(1,0) = (0, \alpha_{-1}(1)) = (0, \frac{1}{2}).$$

Injectivity can also be shown. So in fact: $BS(2,1) = \mathbb{Z} \ltimes_\alpha \mathbb{Z}[\frac{1}{2}]$. In particulat, $BS(2,1)$ is residually finite since $\mathbb{Z}[\frac{1}{2}]$ is residually finite and $\mathbb{Z}$ is residually finite.

Lastly, let us take a look at $BS(2,3) = \langle a, b \mid ba^2b^{-1} = a^3 \rangle$. It turns out, that this is actually not residually finite!

Lemma 2.3.12. *$BS(2,3)$ is not residually finite.*

Proof. Assume that $\varphi : BS(2,3) \to H$ with H finite. Let $x = \varphi(a)$ and $y = \varphi(b)$, and set $n := \mathrm{ord}(x)$. Since x^2 and x^3 are conjugate, they have the same order. Writing x as $[1]$ in $\mathbb{Z}/n\mathbb{Z}$, this forces $2 \nmid n$ and $3 \nmid n$, hence n is odd. Therefore $x \in \langle x^2 \rangle$. Using $a^2 = b^{-1}a^3b$ we obtain

$$[yx^2y^{-1}, x] = \mathbb{1} \implies [yxy^{-1}, x] = \mathbb{1},$$

so every finite quotient kills the commutator $[bab^{-1}, a]$.

It remains to see that $[bab^{-1}, a] \neq 1$ in $BS(2,3)$. The group

$$BS(2,3) = \langle a, b \mid ba^2b^{-1} = a^3 \rangle$$

is the HNN extension of $\mathbb{Z} = \langle a \rangle$ with associated subgroups $2\mathbb{Z}$ and $3\mathbb{Z}$, so Britton's lemma (Theorem 2.3.3) gives normal forms and solves the word problem. The word

$$[bab^{-1}, a] = bab^{-1}aba^{-1}b^{-1}a^{-1}$$

has no pinch: the exponents of a between $b^{\pm 1}$ are ± 1, which are in neither $2\mathbb{Z}$ nor $3\mathbb{Z}$. Hence the word is Britton-reduced and represents a non-trivial element. Therefore $BS(2,3)$ is not residually finite. □

2.3.5 Baumslag–Gersten group

The Baumslag–Gersten group is an HNN extension of $BS(1,2)$ and provides a classical example with extreme isoperimetric behavior; see [2].

Definition 2.3.13. The *Baumslag–Gersten group* is

$$BG := \langle a, b, t \mid bab^{-1} = a^2,\ tat^{-1} = b \rangle.$$

Lemma 2.3.14. *Every finite quotient of BG is cyclic.*

Proof. Let $\varphi : BG \to F$ be a homomorphism to a finite group, and write $\bar{a}, \bar{b}, \bar{t}$ for the images. Since $\bar{b} = \bar{t}\bar{a}\bar{t}^{-1}$, the elements $\bar{a}$ and $\bar{b}$ are conjugate and have the same order, say n. The relation $\bar{b}\bar{a}\bar{b}^{-1} = \bar{a}^2$ shows that conjugation by $\bar{b}$ restricts to the automorphism $x \mapsto x^2$ of the cyclic group $\langle \bar{a} \rangle$, so n is odd and the multiplicative order $\mathrm{ord}_n(2)$ divides $\mathrm{ord}(\bar{b}) = n$.

If $n > 1$, then n is odd, so Euler's totient $\varphi(n)$ is even and $\mathrm{ord}_n(2)$ divides $\varphi(n)$. Hence $\mathrm{ord}_n(2)$ is even. An even integer cannot divide the odd integer n, so we must have $n = 1$. Hence $\bar{a} = \bar{b} = \mathbb{1}$, and F is generated by $\bar{t}$, so F is cyclic. □

2.3.6 Higman's group

We now turn to Higman's group, which has no non-trivial finite quotients; see [27].

Theorem 2.3.15 (Higman). *$\langle a_1, a_2, a_3, a_4 \mid a_{i+1} a_i a_{i+1}^{-1} = a_i^2 \rangle$ for $i \in \mathbb{Z}/4\mathbb{Z}$ is infinite and has no finite quotients.*

Proof. Write $a = a_1$, $b = a_2$, $c = a_3$, $d = a_4$. The finite-quotient argument follows Tao's proof. We first show there are no non-trivial finite quotients. Let $\varphi : G \to F$ with F finite, and still denote the images by a, b, c, d. From $bab^{-1} = a^2$ we obtain by induction

$$b^n a b^{-n} = a^{2^n} \qquad (n \geq 1).$$

If $b^n = \mathbb{1}$ then $a = a^{2^n}$, hence $a^{2^n - 1} = \mathbb{1}$ and

$$\mathrm{ord}(a) \mid 2^{\mathrm{ord}(b)} - 1.$$

If a prime p divides $\mathrm{ord}(a)$, then $2^{\mathrm{ord}(b)} \equiv 1 \pmod{p}$, so $\mathrm{ord}(b)$ is divisible by the multiplicative order of 2 modulo p, which is at most $p - 1$. Thus $\mathrm{ord}(b)$ has a prime divisor strictly smaller than p. Cyclically applying the same argument to b, c, d, a gives an infinite descent in primes, which is impossible. Hence $\mathrm{ord}(a) = \mathrm{ord}(b) = \mathrm{ord}(c) = \mathrm{ord}(d) = 1$, and every homomorphism to a finite group is trivial.

To see that G is infinite, we construct it using amalgamated free products. Let

$$G_1 = \langle a, b, c \mid bab^{-1} = a^2,\ cbc^{-1} = b^2 \rangle.$$

This is the amalgamated free product of $\langle a, b \mid bab^{-1} = a^2 \rangle$ and $\langle b, c \mid cbc^{-1} = b^2 \rangle$ along $\langle b \rangle$. By Britton's lemma, $\langle b \rangle$ is infinite cyclic and intersects neither $\langle a \rangle$ nor $\langle c \rangle$, so Corollary 2.3.7 shows that a and c generate a free subgroup $H = \langle a, c \rangle$ in G_1.

Similarly,

$$G_2 = \langle c, d, a \mid dcd^{-1} = c^2,\ ada^{-1} = d^2 \rangle$$

contains the same free subgroup $H = \langle a, c \rangle$. The Higman group is the amalgamated free product $G_1 *_H G_2$, and another application of Corollary 2.3.7 shows that b and d generate a free subgroup. In particular G is infinite. □

2.3.7 Abels' group

Theorem 2.3.16 (Abels). *There exists a solvable finitely generated group that is not residually finite.*

Proof. Fix a prime p and set $R = \mathbb{Z}[1/p]$. Let

$$G = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & p^k & b \\ 0 & 0 & 1 \end{pmatrix} \,\middle|\, a, b, c \in R,\ k \in \mathbb{Z} \right\}.$$

This is a solvable (hence amenable) subgroup of $GL_3(R)$ and is finitely generated, for instance by

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

As a finitely generated linear group over characteristic 0, G is residually finite.

The center consists of matrices with $a = b = 0$, $k = 0$, and $c \in R$, so $Z(G) \cong R$ via the upper right entry. Let

$$Z = \left\{ \begin{pmatrix} 1 & 0 & z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \,\middle|\, z \in \mathbb{Z} \right\} \leq Z(G).$$

Conjugation by $D = \mathrm{diag}(p, 1, 1)$ sends $a \mapsto pa$ and $c \mapsto pc$, so it defines an endomorphism $\varphi : G \to G$ with $\varphi(Z) = pZ \subsetneq Z$. Hence φ induces a surjective endomorphism $\bar{\varphi} : G/Z \to G/Z$. On the central element represented by $c = 1/p$ we have $\bar{\varphi}(1/p + Z) = 1 + Z = 0$, so $\bar{\varphi}$ is not injective. Thus G/Z is not Hopfian, and by Theorem 1.2.6 it is not residually finite. Since G is solvable, so is G/Z. □

Chapter 3

Cayley graphs and metric geometry

This chapter introduces Cayley graphs and growth, then uses quasi-isometries and hyperbolicity to compare large-scale geometry of groups.

3.1 Cayley graphs and growth

3.1.1 Cayley graphs and word metric

Definition 3.1.1. Let G be a group and $S \subseteq G$ a finite generating set. The Cayley graph $\mathrm{Cay}(G, S)$ has vertex set G and edges

$$g \longleftrightarrow gs \qquad (g \in G,\ s \in S \cup S^{-1}),$$

viewed as an undirected graph. The associated path metric is denoted d_S, and

$$|g|_S \coloneqq d_S(\mathbb{1}, g)$$

is the word length. Then $d_S(g, h) = |g^{-1}h|_S$.

3.1.2 Growth functions

Definition 3.1.2. For $n \in \mathbb{N}$ let $B_S(n) \coloneqq \{g \in G \mid |g|_S \leq n\}$. The growth function of (G, S) is

$$\gamma_{G,S}(n) \coloneqq |B_S(n)|.$$

Definition 3.1.3. For functions $f, g : \mathbb{N} \to \mathbb{N}$ write $f \preceq g$ if there exists $C \geq 1$ such that

$$f(n) \leq C\, g(Cn) \qquad \text{for all } n \geq 1,$$

and write $f \simeq g$ if $f \preceq g$ and $g \preceq f$. The equivalence class of $\gamma_{G,S}$ is called the *growth type* of G.

Proposition 3.1.4. *If S and T are finite generating sets of G, then $\gamma_{G,S} \simeq \gamma_{G,T}$.*

Proof. Let $C = \max\{|t|_S \mid t \in T\}$. Then $|g|_S \leq C|g|_T$ for all $g \in G$, so $B_T(n) \subseteq B_S(Cn)$ and hence $\gamma_{G,T}(n) \leq \gamma_{G,S}(Cn)$. The reverse inequality follows by symmetry, hence $\gamma_{G,S} \simeq \gamma_{G,T}$. □

Definition 3.1.5. Let G be finitely generated and let γ_G denote the growth function with respect to any finite generating set. We say G has *polynomial growth* if $\gamma_G(n) \preceq n^d$ for some d, and *exponential growth* if there exist $a > 1$ and $C > 0$ such that $\gamma_G(n) \geq Ca^n$ for all $n \geq 1$. We say G has *subexponential growth* if $\limsup_{n\to\infty} \gamma_G(n)^{1/n} = 1$; for finitely generated groups the limit exists [37]. We say G has *intermediate growth* if it has subexponential growth but not polynomial growth. These notions depend only on the growth type by Proposition 3.1.4.

Example 3.1.6. $G = \mathbb{Z}^d$ with the standard generators. Then $B_S(n) = \{v \in \mathbb{Z}^d \mid \|v\|_1 \leq n\}$, so there exist constants $c_1, c_2 > 0$ with

$$c_1 n^d \leq \gamma_{G,S}(n) \leq c_2 n^d.$$

Example 3.1.7. $G = \mathcal{F}_k$ with free basis S. The number of reduced words of length n is $2k(2k-1)^{n-1}$, so

$$\gamma_{G,S}(n) = 1 + \sum_{j=1}^{n} 2k(2k-1)^{j-1}.$$

In particular there exist constants $a_1, a_2 > 0$ with $a_1(2k-1)^n \leq \gamma_{G,S}(n) \leq a_2(2k-1)^n$.

Example 3.1.8. Let H be the discrete Heisenberg group

$$H = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \;\middle|\; x, y, z \in \mathbb{Z} \right\}.$$

Set

$$X = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad Y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad Z = [X, Y] = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Every element is uniquely $X^a Y^b Z^c$ with $a, b, c \in \mathbb{Z}$, and $YX = XYZ$. Let $S = \{X, Y\}$ and let a word w have N_X letters $X^{\pm 1}$ and N_Y letters $Y^{\pm 1}$. To move all X's to the left and all Y's to the right, each swap changes the Z–exponent by at most 1, so $|c| \leq N_X N_Y$. If $|w| \leq n$, then $N_X + N_Y \leq n$, hence $|a|, |b| \leq n$ and $|c| \leq n^2/4$. Thus $|B_S(n)| \preceq n^4$.

For the lower bound, fix $a, b \in \{0, \dots, \lfloor n/2 \rfloor\}$ and consider words consisting of a copies of X and b copies of Y. The relation $YX = XYZ$ shows that swapping adjacent Y and X increases the Z–exponent by 1, so for each $c \in \{0, \dots, ab\}$ there is such a word with value $X^a Y^b Z^c$. Its length is $a + b \leq n$, so $|B_S(n)| \succeq \sum_{a,b=0}^{\lfloor n/2 \rfloor} (ab + 1) \simeq n^4$. Hence $\gamma_H(n) \simeq n^4$.

Example 3.1.9. Fix a prime p and set

$$G = \left\{ \begin{pmatrix} p^k & a \\ 0 & 1 \end{pmatrix} \;\middle|\; k \in \mathbb{Z},\ a \in \mathbb{Z}[1/p] \right\} \leq GL_2(\mathbb{Z}[1/p]).$$

Let

$$D = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Then $D^k = \mathrm{diag}(p^k, 1)$ and $D^m U D^{-m} = \left(\begin{smallmatrix} 1 & p^m \\ 0 & 1 \end{smallmatrix}\right)$, so the subgroup of unipotent matrices $\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$ with $x \in \mathbb{Z}[1/p]$ is generated by conjugates of U. For any $a \in \mathbb{Z}[1/p]$ and $k \in \mathbb{Z}$ we can write

$$\begin{pmatrix} p^k & a \\ 0 & 1 \end{pmatrix} = D^k \begin{pmatrix} 1 & a/p^k \\ 0 & 1 \end{pmatrix},$$

so $G = \langle D, U \rangle$.

To see that G has exponential growth, write elements in semidirect-product form (k, x) with multiplication $(k, x)(\ell, y) = (k + \ell, x + p^k y)$, where $D = (1, 0)$ and $U = (0, 1)$. For $\varepsilon_i \in \{0, 1\}$ consider

$$w(\varepsilon) = U^{\varepsilon_0} D\, U^{\varepsilon_1} D \cdots U^{\varepsilon_{n-1}} D.$$

A short computation gives

$$w(\varepsilon) = \left(n,\ \sum_{i=0}^{n-1} \varepsilon_i p^i \right),$$

so distinct binary sequences yield distinct elements. Hence $\gamma_G(2n) \geq 2^n$, and G has exponential growth.

3.1.3 Nilpotent groups have polynomial growth

Proposition 3.1.10. *Every finitely generated nilpotent group has polynomial growth.*

Outline of proof: Let S be a finite generating set of G and let $\gamma_1(G) = G \geq \gamma_2(G) \geq \cdots \geq \gamma_{c+1}(G) = \{\mathbb{1}\}$ be the lower central series. If $c = 1$ then G is abelian and hence has polynomial growth by Example 3.1.6. Assume $c > 1$ and set $H = \gamma_c(G)$. Since G/H is nilpotent of class $c - 1$, the induction hypothesis gives a polynomial upper bound

$$|\pi(B_S(n))| \leq P(n)$$

for the projection $\pi : G \to G/H$.

The subgroup H is central and finitely generated abelian. Indeed, each quotient $\gamma_k(G) / \gamma_{k+1}(G)$ is finitely generated abelian. Choose a generating set $T = \{t_1, \ldots, t_m\}$ of H and extend it to a generating set of G adapted to the lower central series (for instance by taking lifts of generators for each abelian quotient). Now every element $g \in B_S(n)$ can be written by the collecting process as

$$g = t_1^{\alpha_1} \cdots t_m^{\alpha_m} \cdot r,$$

where r lies in a fixed set of coset representatives for H and therefore belongs to $\pi(B_S(n))$. Each exponent α_i is bounded in absolute value by a polynomial in n because commuting a generator past another contributes an element of a higher term in the central series. More precisely, rewriting a word in S of length n into the collected form only produces central commutators of weight at most c, so the number of occurrences of any t_i is bounded by Cn^c for some constant C depending on S and the chosen ordering of commutators. Thus

$$|B_S(n)| \leq |B_T(Cn^c)| \cdot |\pi(B_S(n))|,$$

where $B_T(Cn^c)$ is the ball of radius Cn^c in the abelian group H and therefore has size bounded by a polynomial. Combining with $|\pi(B_S(n))| \leq P(n)$ yields a polynomial upper bound $\tilde{P}(n)$ for $|B_S(n)|$, completing the induction. □

3.1.4 Tao's argument for unitary groups of polynomial growth

The following lemma is a simple instance of the Tits alternative that can be proved more directly using an argument of Tao. It exemplifies how growth conditions can be frequently very powerful and have an unexpected interplay between algebraic and geometric properties.

Lemma 3.1.11. *Let $G \subset U(n)$ be finitely generated and of polynomial growth. Then G is virtually abelian.*

Proof. For $A \in M_n(\mathbb{C})$ let $\|A\|_{\text{op}}$ be the operator norm on $\mathbb{C}^n$. For $g, h \in U(n)$ we set $d(g, h) := \|g - h\|_{\text{op}}$, which gives a bi-invariant metric.

If Γ contains a central element z which is not a scalar multiple of $\mathbb{1}$, then Γ lies in the centralizer $Z_{U(n)}(z)$. By the spectral theorem, z is diagonalizable with at least two distinct eigenvalues and corresponding multiplicities $n_1, \ldots, n_k$ with $k \geq 2$, $\sum_i n_i = n$; then

$$Z_{U(n)}(z) \;\cong\; U(n_1) \times \cdots \times U(n_k)$$

via block decomposition. Let $\pi_i : Z_{U(n)}(z) \to U(n_i)$ be the projections. Each $\pi_i(\Gamma)$ is finitely generated and has polynomial growth, hence by induction on n is virtually abelian. This implies that Γ is virtually abelian as well. From now on, we may assume that the only central elements of Γ are scalar multiples of $\mathbb{1}$.

Fix a small parameter $\varepsilon > 0$ to be chosen later. Define $\Gamma_\varepsilon := \langle \{g \in \Gamma : \|g - \mathbb{1}\|_{\text{op}} \leq \varepsilon\} \rangle$. We claim that G_ε has finite index in Γ. Indeed, by compactness any ε-separated subset of $U(n)$ is finite. Let $X \subset \Gamma$ be a maximal ε-separated subset. For each $g \in \Gamma$, there exists $x \in X$ with $d(g, x) \leq \varepsilon$, so $\|x^{-1}g - \mathbb{1}\|_{\text{op}} \leq \varepsilon$, thus $g \in x\Gamma_\varepsilon$. Hence $[\Gamma : \Gamma_\varepsilon] \leq |X|$.

Since polynomial growth is invariant under finite-index subgroups, Γ_ε also has polynomial growth. Being finitely generated, there exists a finite symmetric generating set $S \subset \Gamma_\varepsilon$ with $\|s - \mathbb{1}\|_{\text{op}} \leq \varepsilon$ for all $s \in S$.

Assume Γ_ε is not abelian, thus S contains a non-scalar element. Choose $h_1 \in S$ which is not scalar, and set $\delta_1 := \|h_1 - \mathbb{1}\|_{\text{op}} \in (0, \varepsilon]$.

For $g \in S$, consider the commutators $[g, h_1] = gh_1g^{-1}h_1^{-1}$. Using bi-invariance of $\|\cdot\|_{\text{op}}$ and the expansion $gh_1g^{-1}h_1^{-1} - \mathbb{1} = (g - \mathbb{1})(h_1 - \mathbb{1}) - (h_1 - \mathbb{1})(g - \mathbb{1})$, we get

$$\|[g, h_1] - \mathbb{1}\|_{\text{op}} \;\leq\; 2\varepsilon\delta_1.$$

Also, $\det([g, h_1]) = 1$ for all $g \in S$. If every $[g, h_1]$ were scalar, then for ε small (say $\varepsilon < 2\sin(\pi/n)$, which makes small scalars of $U(n)$ equal to $\mathbb{1}$), we would have $[g, h_1] = \mathbb{1}$ for all $g \in S$, so h_1 is central – contradiction.

Therefore there exists $g_1 \in S$ with $h_2 := [g_1, h_1]$ non-scalar. Set $\delta_2 := \|h_2 - \mathbb{1}\|_{\text{op}}$; then $0 < \delta_2 \leq 2\varepsilon\delta_1$. Iterating this procedure, we obtain non-scalar elements $g_{k+1}, h_{k+1} = [g_k, h_k]$, so that

$$\delta_{k+1} := \|h_{k+1} - \mathbb{1}\|_{\text{op}} \;\leq\; 2\varepsilon\delta_k.$$

Thus δ_k decreases geometrically $\delta_k \;\leq\; (2\varepsilon)^{k-1}\delta_1$.

Consider the word length in Γ_ε with respect to S. If $h_1 \in \langle S \rangle$ and $h_{k+1} = [g_k, h_k]$ with $g_k \in S$, then

$$\ell_S(h_k) \;\leq\; 3 \cdot 2^{k-1}.$$

Indeed, we have $\ell_S(h_{k+1}) \leq 2 + 2\ell_S(h_k)$, so by induction $\ell_S(h_k) \leq 3 \cdot 2^{k-1} - 2$.

Choose ε small enough that $2\varepsilon \leq 1/10$. Set $\eta := 2\varepsilon$ and consider the sequence $(h_k)_k$. Then $\|h_k - \mathbb{1}\|_{\text{op}} \leq \varepsilon \leq \eta$ and $\|h_{k+1} - \mathbb{1}\|_{\text{op}} \leq \eta\|h_k - \mathbb{1}\|_{\text{op}}$.

We claim that for $0 < \eta \leq 1/10$, the following holds: all words

$$h_1^{i_1} h_2^{i_2} \cdots h_m^{i_m} \qquad (0 \leq i_1, \ldots, i_m \leq M)$$

are pairwise distinct, where $M := \lfloor 1/(10\eta) \rfloor$. Indeed, suppose two words are equal. By comparing at the first differing index t and using the lower bound $\|h_t^r - \mathbb{1}\|_{\text{op}} \geq |r|\delta_t/2$

against the upper bound from the right-hand side, which is at most $4m\delta_{t+1} \leq m\eta\delta_t \leq \delta_t/4$ for small η and m, we reach a contradiction.

Thus we have at least

$$N_m := (M+1)^m \;\geq\; c_0^m\,\varepsilon^{-m}$$

distinct elements in Γ_ε for some constant $c_0 > 0$. Now, any word $h_1^{i_1}\cdots h_m^{i_m}$ has S-length at most

$$M\sum_{k=1}^{m} 3\cdot 2^{k-1} \;\leq\; 3M(2^m-1) \;\leq\; 3M2^m.$$

So all N_m distinct elements lie in the ball $B_S(R_m)$ with $R_m := 3M2^m$. Polynomial growth gives $|B_S(R)| \leq CR^d$ for constants C, d. Thus

$$N_m \;\leq\; C(3M2^m)^d \;\leq\; C'\varepsilon^{-d}2^{md}.$$

Using $N_m \geq c_0^m\varepsilon^{-m}$, we get

$$c_0^m\,\varepsilon^{-m} \;\ll\; \varepsilon^{-d}2^{md}.$$

If $\varepsilon < c_0/2^d$, then

$$\left(\frac{c_0}{2^d}\right)^m \;\ll\; \varepsilon^{-d+m},$$

gives a contradiction for large m. Therefore Γ_ε is abelian provided ε is small enough, and since $[\Gamma : \Gamma_\varepsilon] < \infty$, the original group Γ is virtually abelian. □

3.2 Quasi-Isometry

3.2.1 Definitions

Quasi-isometries formalize the idea that large-scale geometry is insensitive to bounded distortion. They are the natural notion of equivalence in geometric group theory and will connect algebraic properties of groups with the geometry of their Cayley graphs.

Definition 3.2.1. $(X, d), (Y, \delta)$ are metric spaces and $f : X \to Y$. Then

- f is called Lipschitz if there exists C such that $\delta(f(x), f(y)) \leq Cd(x, y)$ for all $x, y \in X$.
- f is called a bilipschitz embedding if there exists C such that

$$\frac{1}{C}d(x,y) \leq \delta(f(x), f(y)) \leq C\cdot d(x,y).$$

- f is called a bilipschitz equivalence if f is bilipschitz and surjective. Equivalently, f is Lipschitz, invertible, and f^{-1} is also Lipschitz.

Example 3.2.2. Γ finitely generated group, $S, S' \subset \Gamma$ finite generating sets. Then (as always):

$$d_S(g,h) = l_S(h^{-1}g) = \min\{k \in \mathbb{N} \mid h^{-1}g = s_1 \ldots s_k, s_i \in S\}.$$

Claim: $\mathrm{id}_\Gamma : (\Gamma, d_S) \to (\Gamma, d_{S'})$ is a bilipschitz equivalence.

Since S' is finite, $m := \max\{l_S(s') \mid s' \in S'\}$ exists. Because of the triangle-inequality: $m\cdot l_{S'}(g) \geq l_S(g)$.

Cay(Γ, S) is also a metric space with the path metric d'_S. The canonical map

$$(\Gamma, d_S) \to (\mathrm{Cay}(\Gamma, S), d'_S)$$

is a bilipschitz embedding with $C = 1$.

Definition 3.2.3. Let $f : (X, d) \to (Y, \delta)$.

- f is a quasi-isometric embedding if there exist $b, c > 0$ such that
$$\frac{1}{c} d(x, y) - b \leq \delta\left(f(x), f(y)\right) \leq c \cdot d(x, y) + b.$$
- $f' : X \to Y$ has finite distance to f if there exists C such that
$$\delta\left(f(x), f'(x)\right) \leq C.$$
- f is called a quasi-isometry if f is a quasi-isometric embedding and $\operatorname{Im} f$ is C-dense for some C. Equivalently, for all $y \in Y$ there exists $x \in X$ with $\delta\left(f(x), y\right) \leq C$.
- f is quasi-Lipschitz if there exist $b, c > 0$ such that
$$\delta\left(f(x), f(y)\right) \leq c \cdot d(x, y) + b.$$

Proposition 3.2.4. *(X, d) and (Y, δ) are metric spaces, $f : X \to Y$. TFAE:*

(i) f is quasi-isometry.

(ii) there exists $g : Y \to X$ such that f and g are quasi-Lipschitz and $f \circ g \sim \mathrm{id}_Y$ and $g \circ f \sim \mathrm{id}_X$

Proof. 2. $\implies$ 1. $\delta\left(f(x), f(y)\right) \leq C \cdot d\left(x, y\right) + b$.

Let $x, y \in X$. $g \circ f \sim \mathrm{id}_X \implies g$ is C-dense for some C. Thus there exist $x' \in Y$ and $y' \in Y$ such that
$$d(x, g(x')) \leq C, \qquad d(y, g(y')) \leq C.$$
Thus
$$\begin{aligned} d(x, y) &\leq d\left(g(x'), g(y')\right) + 2C \\ &\leq C \cdot \delta\left(x', y'\right) + b + 2C \\ \Rightarrow & \frac{1}{C} d(x, y) - \frac{b + 2C}{C} \leq \delta\left(x', y'\right). \end{aligned}$$
We need to estimate $\delta\left(x', f(x)\right)$ and $\delta\left(y', f(y)\right)$. But
$$\delta\left(x', f\left(g\left(x'\right)\right)\right) \leq C,$$
so
$$\begin{aligned} \delta\left(x', f(x)\right) &\leq \delta\left(x', f\left(g\left(x'\right)\right)\right) + \delta\left(f\left(g\left(x'\right)\right), f(x)\right) \\ &\leq C + C \cdot d\left(g\left(x'\right), x\right) + b \leq C'. \end{aligned}$$
Thus
$$\frac{1}{C} d\left(x, y\right) - C'' \leq \delta\left(f(x), f(y)\right).$$
1. $\implies$ 2. If f is a quasi-isometry and C-dense, define $g : Y \to X$.
$$g(y) \coloneqq x$$

for some $x \in X$ s.t. $\delta(f(x), y) \leq C$. (the image of f is C-dense in Y).

f is quasi-Lipschitz. Now for g: Let $y, y' \in Y$:

$$d(g(y), g(y')) = d(x, x')$$

for some x, x' s.t. $\delta(f(x), y) \leq C$ and $\delta(f(x'), y') \leq C$.

$$\frac{1}{C} d(x, x') - b \leq \delta(f(x), f(x'))$$

since f is quasi-isometry. That means

$$\begin{aligned} d(g(y), g(y')) \leq d(x, x') &\leq C \cdot \delta(f(x), f(x')) + C \cdot b \\ &\leq C \cdot d(y, y') + C. \end{aligned}$$

Thus, g is quasi-Lipschitz.

$f \circ g \sim \mathrm{id}_Y$. By definition $g(y) = x$ satisfies

$$\delta(f(g(y)), y) = \delta(f(x), y) \leq C.$$

Now lastly $g \circ f \sim \mathrm{id}_X$. Let $x \in X$. Then $g(f(x)) = x'$ with $d(f(x'), f(x)) \leq C$.

$$d(x, g(f(x))) = d(x, x') \leq C \cdot d(f(x), f(x')) + b \cdot C \leq C'.$$

□

Proposition 3.2.5. *Let G, H be finitely generated groups with word metrics and growth functions $\gamma_{G,S}, \gamma_{H,T}$. If G and H are quasi-isometric, then $\gamma_{G,S} \simeq \gamma_{H,T}$.*

Proof. Let $f : G \to H$ be a (c, b)–quasi-isometry and let $g : H \to G$ be a quasi-inverse with $d_G(g(f(x)), x) \leq D$ and $d_H(f(g(y)), y) \leq D$. After composing f with a left translation we may assume $f(\mathbb{1}) = \mathbb{1}$. Then the coarse Lipschitz bounds give

$$f(B_S(n)) \subseteq B_T(cn + b), \qquad g(B_T(n)) \subseteq B_S(cn + b).$$

If $g(y) = g(y')$, then

$$d_H(y, y') \leq d_H(y, f(g(y))) + d_H(f(g(y')), y') \leq 2D,$$

using that $f(g(y))$ is within D of y and similarly for y'. Hence each fiber of g has cardinality at most $M := |B_T(2D)|$. When we count $|B_T(n)|$ we can therefore bound

$$\gamma_{H,T}(n) = |B_T(n)| \leq M \cdot |\{x \in G : \ g(f(x)) \in B_T(n)\}| \leq M \cdot \gamma_{G,S}(cn + b),$$

where the last inequality uses $g(B_T(n)) \subseteq B_S(cn + b)$. Since $cn + b \leq (c + b)n$ for $n \geq 1$ we arrive at

$$\gamma_{H,T}(n) \leq M \, \gamma_{G,S}((c + b)n) \qquad (n \geq 1).$$

Reversing the roles of f and g yields the reverse inequality with possibly different constants M', c', b', so $\gamma_{G,S} \preceq \gamma_{H,T}$ and conversely $\gamma_{H,T} \preceq \gamma_{G,S}$. Therefore $\gamma_{G,S} \simeq \gamma_{H,T}$. □

Corollary 3.2.6. *The groups $\mathbb{Z}^d$ are not quasi-isometric for different d.*

Proof. This follows directly from Example 3.1.6. □

3.2.2 Finite index and regular trees

Proposition 3.2.7. *The* 3*–regular tree and the* 6*–regular tree are quasi-isometric.*

Proof. Let T_3 be the 3–regular tree and fix a base vertex o. Let V_0 be the set of vertices at even distance from o; this is an index-2 subset of the vertex set. Connect $x, y \in V_0$ by an edge if $d_{T_3}(x, y) = 2$ and call the resulting graph T_0, the index-2 subtree obtained by keeping even vertices. Each $x \in V_0$ has 6 such neighbors, so T_0 is a 6–regular tree. The inclusion $T_0 \to T_3$ is 2-Lipschitz and its image is 1-dense, hence it is a quasi-isometry. □

See Figure 3.1 for a schematic picture.

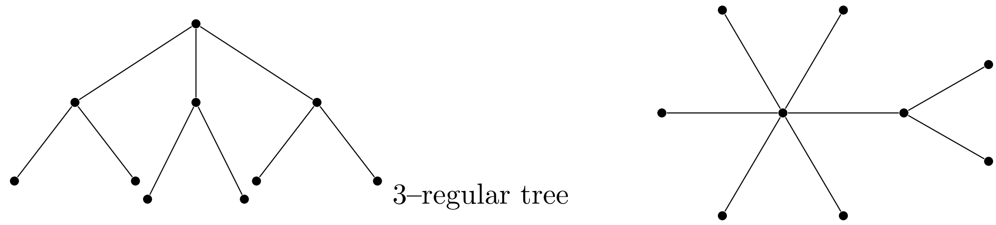


Figure 3.1: Left: a portion of the 3–regular tree. Right: the index-2 subtree on even vertices of T_3, with edges recording paths of length 2, yielding a 6–regular tree.

Proposition 3.2.8. *Let G be finitely generated and let $H \leq G$ have finite index. Then the inclusion $H \hookrightarrow G$ is a quasi-isometry.*

Proof. Fix a finite generating set S of G and a transversal $R \subseteq G$ for the right cosets G/H. Write each $g \in G$ uniquely as $g = hr$ with $h \in H$, $r \in R$, and define $p(g) := h$. Then $d_S(g, p(g)) = |r|_S \leq R_0$ where $R_0 := \max_{r\in R} |r|_S$, so p is at bounded distance from the inclusion.

Let T be a finite generating set for H and set

$$F := \{rsr'^{-1} \mid r, r' \in R,\ s \in S\} \subseteq H.$$

The set F is finite, so $L := \max_{f\in F} |f|_T$ is finite. If $g = hr$ and $gs = h'r'$ with $r, r' \in R$, then $h' = h(rsr'^{-1})$, hence $d_T(p(gs), p(g)) \leq L$. Thus p is Lipschitz, and by the quasi-inverse criterion the inclusion is a quasi-isometry. □

Corollary 3.2.9. *Any finitely generated group containing a nilpotent subgroup of finite index has polynomial growth.*

Proof. Let $H \leq G$ be nilpotent of finite index. By Proposition 3.2.8 the inclusion $H \hookrightarrow G$ is a quasi-isometry, and Proposition 3.2.5 shows that quasi-isometric groups have equivalent growth type. Since H is nilpotent it has polynomial growth by Proposition 3.1.10, hence G does as well. □

Corollary 3.2.10. *All finitely generated free groups of rank at least* 2 *are quasi-isometric.*

Proof. By the Nielsen–Schreier formula, an index d subgroup of $\mathcal{F}_n$ is free of rank $1 + d(n-1)$. Taking $n = 2$ shows that for every $m \geq 2$ there is an index $m-1$ subgroup of $\mathcal{F}_2$ isomorphic to $\mathcal{F}_m$. Apply Proposition 3.2.8. □

Example 3.2.11.
- $(\mathbb{Z}^2, d_S) =_{QI} (\mathbb{R}^2, d_{\text{st}})$. More generally, $\mathbb{Z}^d$ and $\mathbb{Z}^{d'}$ are not quasi-isometric for $d \neq d'$ by Proposition 3.2.5.
- The 3–regular and 6–regular trees are quasi-isometric by Proposition 3.2.7.
- Any two nonabelian finitely generated free groups are quasi-isometric by Corollary 3.2.10.
- $G =_{QI} \{*\}$ for G finite.
- $\Gamma \times H =_{QI} \Gamma$ for H finite.

3.3 Gromov Hyperbolicity

3.3.1 Definitions

We now shift to large-scale geometry. The goal is to relate isoperimetric inequalities in a combinatorial complex to hyperbolicity of its 1-skeleton, culminating in Papasoglu's bigon criterion. We will apply these notions to Cayley graphs as geodesic metric spaces.

Definition 3.3.1. Let $x, y \in (X, d)$ and $L = d(x, y)$. A geodesic from x to y is an isometry

$$\varphi : [0, L] \to X,\ \varphi(0) = x, \varphi(L) = y.$$

Definition 3.3.2. Let $\delta > 0, x, y \in (X, d)$ and $L = d(x, y)$ and $\varphi : [0, L] \to X$ a geodesic.

$$B_\delta \coloneqq \{z \in X \mid \exists\, \lambda \in [0, L] \text{ with } d(z, \varphi(\lambda)) < \delta\}.$$

Definition 3.3.3. A metric space (X, d) is called δ-hyperbolic if and only if for all $x, y, z \in X$ and geodesics $\varphi_{xy}, \varphi_{yz}, \varphi_{xz}$, the image of one of the geodesics is contained in the union of δ-neighbourhoods of the other two geodesics. That means, triangles are δ-*thin.*

(X, d) is called (Gromov-)hyperbolic if and only if there exists δ such that it is δ-hyperbolic.

Lemma 3.3.4. *Gromov hyperbolicity is invariant under quasi-isometries between geodesic metric spaces. In particular, if X and Y are geodesic metric spaces and X is quasi-isometric to Y, then X is Gromov hyperbolic if and only if Y is.*

Definition 3.3.5. A finitely generated group G is (word-)hyperbolic if, for some (equivalently any) finite generating set S, the Cayley graph $\text{Cay}(G, S)$ (viewed as a geodesic metric space) is Gromov hyperbolic.

Example 3.3.6. (i) (X, d) bounded, i.e. $\sup_{x,y} d(x, y) \leq C$. Then (X, d) is C-hyperbolic.

(ii) Let (X, d) be a simplicial tree. Then X is easily seen to be 0-hyperbolic. In particular, Cayley graphs of free groups are regular trees (see Figure 3.1), hence free groups of rank ≥ 2 are Gromov hyperbolic.

(iii) $(\mathbb{Z}^2, d)$ where d is the Manhattan metric is not hyperbolic.

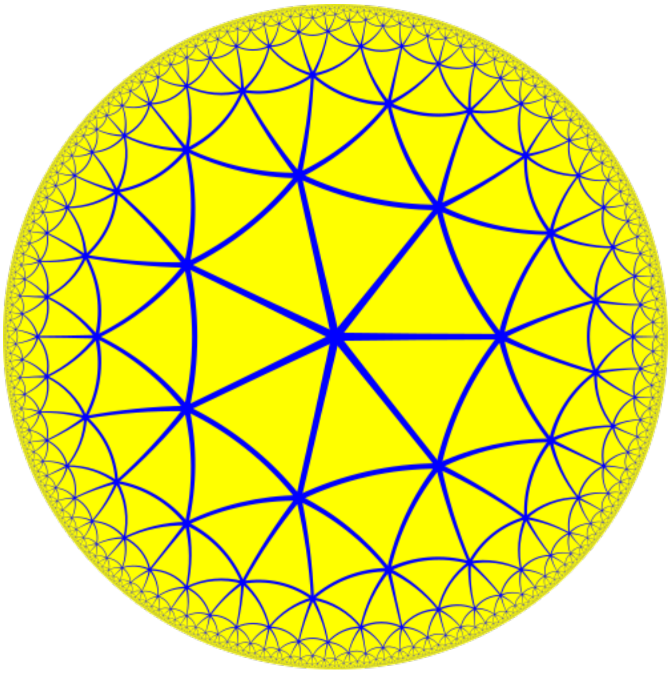


Figure 3.2: The regular tiling $\{3,7\}$ of $\mathbb{H}^2$ (Wikipedia).

3.3.2 The $(3,7)$-triangle complex

We begin with a concrete 2-complex whose local geometry enforces negative curvature. This provides a clean model for the later isoperimetric argument.

Definition 3.3.7. Let X be a connected, 2-dimensional simplicial complex such that

(i) X is simply connected,

(ii) for every vertex $x \in X^{(0)}$ the link $\mathrm{Lk}_X(x)$ is a 7–cycle.

Let $\Gamma := X^{(1)}$ be the 1-skeleton with the graph metric (each edge has length 1).

Definition 3.3.8. Let γ be a *simple* closed edge loop in Γ (i.e. an embedded simplicial cycle). Since X is planar, γ separates X into a bounded and an unbounded component. Define $\mathrm{Fill}(\gamma)$ to be the (finite) subcomplex consisting of all triangles lying in the bounded component, together with their faces. Then $\mathrm{Fill}(\gamma)$ is a finite triangulated topological disc with boundary cycle γ. Write $\mathrm{Area}(\gamma)$ for the number of triangles in $\mathrm{Fill}(\gamma)$.

Remark 3.3.9. Local finiteness ensures $\mathrm{Fill}(\gamma)$ is finite for every simple loop γ. We will only use fillings of simple loops, so no disc diagram formalism is needed.

Linear isoperimetric inequality. We now estimate the area of a filling by combinatorial Euler characteristic and boundary length. Fix a simple closed edge loop γ of length $L := |\gamma|$ and set $R := \mathrm{Fill}(\gamma)$. Let

$$F = \#\{\text{triangles of } R\}, \quad E = \#\{\text{edges of } R\}, \quad V = \#\{\text{vertices of } R\}.$$

Let V_{int} be the number of vertices of R not lying on the boundary cycle $\partial R = \gamma$. Since γ is a simplicial cycle with L edges, it has L boundary vertices.

Lemma 3.3.10. *For $R = \mathrm{Fill}(\gamma)$ one has*

$$F = 2V_{\mathrm{int}} + L - 2.$$

Proof. Let $E_\partial = L$ be the number of boundary edges and let $E_{\text{int}} = E - E_\partial$. Counting triangle edges gives

$$3F = 2E_{\text{int}} + E_\partial = 2(E - L) + L = 2E - L,$$

hence $E = (3F + L)/2$. Euler's formula for a disc is $V - E + F = 1$, and $V = V_{\text{int}} + L$, so

$$(V_{\text{int}} + L) - \frac{3F + L}{2} + F = 1,$$

which rearranges to $F = 2V_{\text{int}} + L - 2$. □

For a vertex $v \in R^{(0)}$ let $t(v)$ be the number of triangles of R incident to v. Then $3F = \sum_{v \in R^{(0)}} t(v)$ because each triangle has three corners.

Lemma 3.3.11. *One has*

$$\sum_{v \in R^{(0)}} t(v) \geq 7V_{\text{int}} + L.$$

Proof. If v is an interior vertex of R, then every triangle incident to v in X lies in R; otherwise an outside triangle would share an edge with a triangle in R, forcing that edge (and hence v) to lie on the boundary of R. Thus $t(v) = 7$ for every interior vertex.

If v is a boundary vertex, then it is an endpoint of a boundary edge, and every boundary edge lies in exactly one triangle of R. Hence each boundary vertex belongs to at least one triangle, i.e. $t(v) \geq 1$. There are L boundary vertices. □

Theorem 3.3.12. *In a planar* $(3, 7)$ *triangle tiling, every simple closed edge loop* γ *of length* L *satisfies*

$$\text{Area}(\gamma) \leq 5L - 14,$$

and in particular $\text{Area}(\gamma) \leq 5L$ *for all* $L \geq 1$.

Proof. By the corner bound, $3F = \sum t(v) \geq 7V_{\text{int}} + L$. Using $F = 2V_{\text{int}} + L - 2$, we obtain

$$3(2V_{\text{int}} + L - 2) \geq 7V_{\text{int}} + L,$$

hence $V_{\text{int}} \leq 2L - 6$, and therefore

$$F = 2V_{\text{int}} + L - 2 \leq 2(2L - 6) + L - 2 = 5L - 14.$$

□

Remark 3.3.13. Any closed edge loop in Γ can be reduced by canceling backtracks and then decomposed into a union of simple loops. Applying Theorem 3.3.12 to each component shows that Γ satisfies a linear isoperimetric inequality for arbitrary null-homotopic loops, with the same constant $K = 5$.

We will see in the next section that a linear isoperimetric inequality implies Gromov hyperbolicity. Let's record the following two corollaries already here:

Corollary 3.3.14. *In a planar* $(3, 7)$ *triangle tiling one may take* $K = 5$ *and* $M = 3$ *(maximum face length) in Proposition 3.3.26, hence every geodesic bigon has width* ≤ 386.

Proof. The linear isoperimetric inequality holds with $K = 5$, and each face is a triangle so $M = 3$. Apply Proposition 3.3.26. □

Corollary 3.3.15. *The* 1*-skeleton* Γ *of a planar* $(3, 7)$ *triangle tiling is Gromov hyperbolic.*

Proof. By the previous corollary all geodesic bigons are uniformly thin. Apply Theorem 3.3.27. □

3.3.3 The $(2, 3, 7)$ triangle group

Definition 3.3.16. The $(2, 3, 7)$ triangle group is the orientation-preserving subgroup of index 2 in the Coxeter reflection group

$$[2,3,7] \;=\; \Big\langle s_0, s_1, s_2 \;\Big|\; s_i^2 = 1,\; (s_0s_1)^2 = (s_1s_2)^3 = (s_2s_0)^7 = 1 \Big\rangle,$$

generated by $a = s_0s_1$ and $b = s_1s_2$ with presentation

$$\Delta(2,3,7) \;=\; \langle a, b \mid a^2 = b^3 = (ab)^7 = 1\rangle.$$

The reflections s_0, s_1, s_2 arise from a hyperbolic triangle with angles $\pi/2, \pi/3, \pi/7$, and the group acts on the regular tiling $\{3, 7\}$ of $\mathbb{H}^2$ by symmetries (see Figure 3.2). Its barycentric subdivision gives the $(3, 7)$-triangle complex from Section 3.3.2, and the dual graph of that subdivision is the Cayley graph of $[2, 3, 7]$ with generators $\{s_0, s_1, s_2\}$. In particular $\Delta(2, 3, 7)$ is quasi-isometric to the $(3, 7)$-triangle complex and $\mathbb{H}^2$.

3.3.4 Hyperbolic groups and Dehn presentations

Definition 3.3.17. Let $\langle S|R\rangle$ be a finite presentation and let $\mathrm{Area}(w)$ be the minimal number of 2-cells in a van Kampen diagram for a null word w. The presentation satisfies a *linear isoperimetric inequality* if there exists $K \geq 1$ such that $\mathrm{Area}(w) \leq K|w|$ for all null words w.

Definition 3.3.18. A finite presentation $\langle S|R\rangle$ is called *Dehn* if every non-trivial freely reduced null word contains a subword that is more than half of some cyclic conjugate of a relator in $R^{\pm 1}$.

Proposition 3.3.19. *A finite presentation is Dehn if and only if the usual Dehn algorithm (replace a subword longer than half of a defining relator by the complementary shorter piece and then freely reduce) reduces every null word to the empty word. In that case the algorithm runs in linear time.*

Proof. If the presentation is Dehn and w is a non-trivial null word, then w contains such a subword u. Replacing u by the complementary shorter piece strictly reduces length, so by induction on $|w|$ the algorithm terminates at the empty word. Each step shortens the word, so the running time is $O(|w|)$. Conversely, if the algorithm always reduces null words to the empty word, then any non-trivial null word must admit a Dehn reduction at the first step; otherwise the algorithm would halt on a non-trivial word. Hence the presentation is Dehn. □

Corollary 3.3.20. *A Dehn presentation satisfies a linear isoperimetric inequality.*

Proof. Each Dehn reduction replaces a subword by a complementary piece coming from a relator, so one 2-cell accounts for each step. Since each step strictly shortens the word, the number of steps is at most $|w|$, hence $\mathrm{Area}(w) \leq |w|$. □

Definition 3.3.21. A *geodesic bigon* in a graph Γ is a pair of geodesics p, q with the same endpoints. Its *width* is

$$\mathrm{width}(p, q) := \max_{v \in p} \mathrm{dist}_\Gamma(v, q).$$

We call the bigon *δ–thin* if $\mathrm{width}(p, q) \leq \delta$.

Let $\langle S|R\rangle$ be a finite presentation and set

$$M \coloneqq \max\{|r| : r \in R\}.$$

Assume a linear isoperimetric inequality:
(LI) There exists $K \geq 1$ such that for every null word w one has $\mathrm{Area}(w) \leq K|w|$.

Lemma 3.3.22. *Let (p, q) be a geodesic bigon in Γ with common endpoints. Set*

$$R \coloneqq \max_{v \in p} \mathrm{dist}_\Gamma(v, q), \qquad m \coloneqq \lfloor R/2 \rfloor.$$

Choose $u \in p$ with $\mathrm{dist}_\Gamma(u, q) = R$. Then there exists a geodesic subpath $L \subset p$ of length $|L| = m$ such that

$$\mathrm{dist}_\Gamma(v, q) \geq m \qquad \text{for every vertex } v \in L.$$

Moreover, if a, b are the endpoints of L, then

$$m \leq \mathrm{dist}_\Gamma(a, q) \leq R \quad \text{and} \quad m \leq \mathrm{dist}_\Gamma(b, q) \leq R.$$

Proof. Along a graph geodesic, the function $f(v) \coloneqq \mathrm{dist}_\Gamma(v, q)$ is 1-Lipschitz. Fix $u \in p$ with $f(u) = R$. Since p has endpoints on q, we have $f = 0$ at both endpoints of p, hence p has length at least R. Thus at least one of the two subpaths of p starting at u and going to an endpoint has length $\geq m$. Let L be the subpath of length m starting at u and going in that direction.

Every vertex v on L satisfies $\mathrm{dist}_p(u, v) \leq m$, hence by the 1-Lipschitz property

$$f(v) \geq f(u) - \mathrm{dist}_p(u, v) \geq R - m \geq m.$$

The endpoint bounds follow from $f(v) \geq m$ on L and $f(v) \leq R$ on all of p. □

Fix $L \subset p$ as above with endpoints a, b. Choose geodesics α, β in Γ of length m from a and b to q; let their endpoints on q be a', b'. Let σ be the subpath of q between a' and b'. Since q is geodesic,

$$|\sigma| = \mathrm{dist}_\Gamma(a', b') \leq |\alpha| + |L| + |\beta| = 3m. \tag{3.1}$$

Let

$$\eta \coloneqq L \cdot \beta \cdot \sigma^{-1} \cdot \alpha^{-1}.$$

Then η is a null-homotopic loop of length $|\eta| \leq 6m$. Let D be a reduced van Kampen diagram for η. Inside D we consider *combinatorial parallels* to the side L.

Let $L_D \subset \partial D$ be the boundary segment corresponding to L. For a vertex $v \in D^{(0)}$ define

$$d(v) \coloneqq \mathrm{dist}_{D^{(1)}}(v, L_D).$$

For $t \geq 0$, let S_t be the full subcomplex spanned by $\{v \in D^{(0)} : d(v) \leq t\}$, the combinatorial t–neighbourhood of L_D inside D.

Lemma 3.3.23. *Every vertex of the boundary segment $\sigma \subset \partial D$ satisfies $d(\cdot) \geq m$. Equivalently, $S_t \cap \sigma = \varnothing$ for all $t < m$.*

Proof. For any $x \in \sigma$ and any $y \in L_D$, every path in $D^{(1)}$ from y to x maps to a path in the Cayley graph, so

$$\mathrm{dist}_{D^{(1)}}(y,x) \geq \mathrm{dist}_\Gamma(y,x) \geq \mathrm{dist}_\Gamma(y,q) \geq m,$$

since $x \in \sigma \subset q$ and every $y \in L$ satisfies $\mathrm{dist}_\Gamma(y,q) \geq m$. Thus $d(x) = \mathrm{dist}_{D^{(1)}}(x, L_D) \geq m$. □

Lemma 3.3.24. *For each integer $t \in \{0,1,\dots,\lfloor m/2 \rfloor\}$ there exists an edge path $\ell_t \subset D^{(1)}$ such that*

(i) ℓ_t connects the t–th vertex on the α–side (measured from L) to the t–th vertex on the β–side,

(ii) every vertex on ℓ_t satisfies $d = t$,

(iii) for different t the paths ℓ_t are edge–disjoint,

(iv) $|\ell_t| \geq m - 2t$.

Proof. Fix $t < m$. By the previous lemma, the neighbourhood S_t does not meet σ. Thus S_t is a connected subcomplex containing L_D, and its boundary inside the disc D has an inner boundary component separating L_D from σ. This inner boundary necessarily meets both sides α and β, and contains a path connecting them; call it ℓ_t. By construction, ℓ_t lies in the level set $d = t$, giving (i) and (ii).

If $t \neq t'$, then d takes different values, so the level sets are disjoint; hence no edge can lie in both ℓ_t and $\ell_{t'}$. This gives (iii).

For (iv), suppose $|\ell_t| < m - 2t$. Then there is a path from a to b inside $D^{(1)}$ of length

$$t + |\ell_t| + t < t + (m-2t) + t = m,$$

obtained by going t steps along α from a to ℓ_t, then along ℓ_t, then t steps back along β to b. This yields a path in Γ from a to b shorter than m, contradicting that $L \subset p$ is a geodesic segment of length m. □

Lemma 3.3.25. *Let D be as above. Then*

$$\mathrm{Area}(\eta) \geq \frac{m^2}{2M} - \frac{6m}{M}.$$

Proof. Let $E(D)$ be the number of edges and $F(D)$ the number of 2-cells of D. By the previous lemma, the paths ℓ_t are edge–disjoint, so

$$E(D) \geq \sum_{t=0}^{\lfloor m/2 \rfloor} |\ell_t| \geq \sum_{t=0}^{\lfloor m/2 \rfloor} (m-2t) \geq \frac{m^2}{4}.$$

The sum of the perimeters of all 2-cells is at most $MF(D)$ and equals $2E_{\mathrm{int}}(D) + P$, where $P := |\partial D| = |\eta|$. Hence

$$MF(D) \geq 2E_{\mathrm{int}}(D) + P = 2E(D) - P,$$

so $F(D) \geq (2E(D) - P)/M$. Using $E(D) \geq m^2/4$ and $P = |\eta| \leq 6m$ gives

$$F(D) \geq \frac{m^2}{2M} - \frac{6m}{M}.$$

□

Proposition 3.3.26. *Assume (LI) with constant K. Then every geodesic bigon (p, q) in Γ is uniformly thin:*

$$\text{width}(p, q) \leq 24KM + 26.$$

Proof. Let (p, q) be a geodesic bigon and let $R = \text{width}(p, q)$. If $R \leq 2$ we are done. Otherwise let $m = \lfloor R/2 \rfloor \geq 1$ and build the loop η as above. Then $|\eta| \leq 6m$ and $\text{Area}(\eta) \geq \frac{m^2}{2M} - \frac{6m}{M}$.

By (LI), $\text{Area}(\eta) \leq K|\eta| \leq 6Km$. Hence

$$\frac{m^2}{2M} - \frac{6m}{M} \leq 6Km \quad \Longrightarrow \quad m \leq 12KM + 12.$$

Since $m = \lfloor R/2 \rfloor$, we have $R \leq 2m + 2 \leq 24KM + 26$. □

Theorem 3.3.27 (Papasoglu)**.** *Let Γ be a geodesic metric graph. Then the following are equivalent:*

(i) Γ is Gromov hyperbolic (uniformly thin geodesic triangles),

(ii) there exists $\delta \geq 0$ such that every geodesic bigon in Γ has width at most δ.

Proof. See [41]. □

Corollary 3.3.28. *If a finitely presented group admits a linear isoperimetric inequality, then its Cayley graph is Gromov hyperbolic.*

Proof. Apply Proposition 3.3.26 and then Theorem 3.3.27. □

Definition 3.3.29. A path is called k–local geodesic if every subpath of length at most k is geodesic.

Lemma 3.3.30. *Let Γ be δ–hyperbolic and let $k > 8\delta$. Then any k–local geodesic is a $(2, 4\delta)$–quasi-geodesic. In particular, any closed k–local geodesic has length at most $4\delta + k$.*

Proof. Let γ be a k–local geodesic and let x, y be points on γ with subpath length L. If $L \leq k$ then $d(x, y) = L$. Otherwise subdivide the subpath into segments of length in $[k/2, k]$. Each segment is geodesic, and δ–thinness of geodesic polygons shows that consecutive segments can backtrack by at most 2δ. Summing gives $d(x, y) \geq L/2 - 2\delta$, hence $L \leq 2d(x, y) + 4\delta$. If γ is closed, take $x = y$ to obtain $L \leq 4\delta$, and combine with the case $L \leq k$ to get the stated bound. □

Definition 3.3.31. Let G be generated by a finite set S and let $L \in \mathbb{N}$. The *short-loop presentation* of length L is

$$\langle S \mid R_L \rangle, \qquad R_L := \{w \in \mathcal{F}_S \mid |w|_S \leq L,\ w = \mathbb{1} \text{ in } G\}.$$

Theorem 3.3.32. *Let G be δ–hyperbolic with respect to a finite generating set S. Then there exists $L = L(\delta, S)$ such that the short-loop presentation $\langle S \mid R_L \rangle$ is Dehn. Consequently, G admits a linear isoperimetric inequality (for this presentation) and Dehn's algorithm runs in linear time.*

Proof. Choose $k > 8\delta$ and set $L = 2k$. Since there are finitely many words of length $\leq L$, the presentation is finite. Let w be a freely reduced null word. If $|w| \leq L$, then $w \in R_L$. If $|w| > L$, then the loop labeled by w cannot be k–local geodesic by Lemma 3.3.30, since any closed k–local geodesic has length $< 2k$. Hence w contains a subword u of length k that is not geodesic. Let v be a geodesic word between the endpoints of u, so $|v| < k$. Then $r \coloneqq uv^{-1}$ is a relator in R_L and u is more than half of r. Thus the presentation is Dehn. The final statement follows from Proposition 3.3.19 and Corollary 3.3.20. □

Remark 3.3.33. Linear isoperimetry depends on the chosen presentation. It does not force the original relators to form a Dehn presentation, so the naive Dehn algorithm can fail for a given presentation. The short-loop presentation above shows that once one adds all null words of length $\leq L$, one obtains a Dehn presentation and hence a correct, linear-time Dehn algorithm.

Chapter 4

Amenability

Amenability can be approached via Følner sets, invariant means, and paradoxical decompositions. We present these equivalent viewpoints and give the contrapositive construction using Philip Hall's marriage lemma.

4.1 Amenability and Tarski's alternative

4.1.1 Følner condition

The Følner condition captures the idea that finite subsets can have arbitrarily small boundary relative to their size.

Definition 4.1.1. Let G be a discrete group. For a finite set $S \subseteq G$ and $\varepsilon > 0$, a finite nonempty set $F \subseteq G$ is called (S, ε)-*Følner* if

$$\frac{|SF \triangle F|}{|F|} \leq \varepsilon.$$

Equivalently, for fixed finite S one may require

$$|SF| \leq (1 + \varepsilon)|F|.$$

The group G is called *amenable* if for every finite $S \subseteq G$ and every $\varepsilon > 0$ there exists a (S, ε)-Følner set F.

Lemma 4.1.2. *Let $G = \langle \Sigma \rangle$ be finitely generated with Σ finite. Assume that for every $\varepsilon > 0$ there exists a finite nonempty $F \subseteq G$ with*

$$\frac{|\Sigma F \triangle F|}{|F|} \leq \varepsilon.$$

Then G satisfies the Følner condition for every finite set $S \subseteq G$.

Proof. Fix a finite $S \subseteq G$. Choose $m \in \mathbb{N}$ with $S \subseteq \Sigma^m$. Write $\partial_T(F) := TF \setminus F$ for the (right) boundary with respect to a finite set T. One checks

$$\partial_{\Sigma^m}(F) \subseteq \bigcup_{k=0}^{m-1} \Sigma^k \, \partial_\Sigma(F),$$

hence

$$|\partial_{\Sigma^m}(F)| \le \Big(\sum_{k=0}^{m-1} |\Sigma|^k\Big)\, |\partial_\Sigma(F)|.$$

Choosing F with $|\partial_\Sigma(F)|/|F|$ sufficiently small forces $|\partial_{\Sigma^m}(F)|/|F|$ small, and thus also

$$\frac{|SF \triangle F|}{|F|} \le \frac{|\Sigma^m F \triangle F|}{|F|}.$$

□

Example 4.1.3.

- Every finite group is amenable: take $F = G$.
- $G = \mathbb{Z}^n$ is amenable: for $F_N = [-N, N]^n \cap \mathbb{Z}^n$ one has $|\Sigma F_N \triangle F_N|/|F_N| \to 0$ for any finite generating set Σ.
- Free groups F_k for $k \ge 2$ are not amenable: their Cayley graphs are regular trees with positive isoperimetric constant.

4.1.2 Tarski alternative

Definition 4.1.4. Let G be a group. A *paradoxical decomposition* of G consists of finite families of pairwise disjoint subsets

$$A_1, \dots, A_m, \qquad B_1, \dots, B_n \subseteq G$$

and group elements $g_1, \dots, g_m$, $h_1, \dots, h_n \in G$ such that

$$G = \bigsqcup_{i=1}^m A_i \sqcup \bigsqcup_{j=1}^n B_j, \qquad G = \bigsqcup_{i=1}^m g_i A_i, \qquad G = \bigsqcup_{j=1}^n h_j B_j.$$

If such a decomposition exists, we say that G is *paradoxical*.

Definition 4.1.5. A *mean* on $\ell^\infty(G)$ is a linear functional $m : \ell^\infty(G) \to \mathbb{R}$ such that $m(f) \ge 0$ if $f \ge 0$ pointwise and $m(\mathbf{1}) = 1$. It is *left invariant* if $m(g \cdot f) = m(f)$ where $(g \cdot f)(x) = f(g^{-1}x)$.

Definition 4.1.6. A *finitely additive probability measure* is a map $\mu : \mathcal{P}(G) \to [0, 1]$ with $\mu(G) = 1$ and $\mu(A \sqcup B) = \mu(A) + \mu(B)$ for disjoint A, B. It is *left invariant* if $\mu(gA) = \mu(A)$.

Theorem 4.1.7 (Tarski)**.** *For a discrete group G the following are equivalent:*

(i) G is amenable (Følner condition).

(ii) There exists a left G-invariant mean on $\ell^\infty(G)$.

(iii) There exists a left G-invariant finitely additive probability measure on $\mathcal{P}(G)$.

(iv) G is not paradoxical.

Proof. Assume (1). For a finite Følner set F define $m_F(f) = \frac{1}{|F|}\sum_{x \in F} f(x)$. Taking a weak-$*$ accumulation point along a Følner net yields a left invariant mean, giving (2). Given (2), define $\mu(A) := m(\mathbf{1}_A)$ to obtain (3). If (3) held and G were paradoxical, finite additivity and invariance would force $1 = \mu(G) = 2$, a contradiction, so (3) implies (4). The implication $4 \Rightarrow 1$ is proved in the next subsection. □

See [42, 52] for detailed proofs and further equivalent formulations.

Remark 4.1.8. There exists a linear functional $\alpha : \ell^\infty(\mathbb{N}) \to \mathbb{R}$ such that

$$\alpha(f) \in [\liminf_{n\to\infty} f(n), \limsup_{n\to\infty} f(n)].$$

This follows from Hahn–Banach by extending the limit functional on convergent bounded sequences.

4.1.3 Hall's marriage lemma and the contrapositive $4 \Rightarrow 1$

Definition 4.1.9. A bipartite graph $\Gamma = (X \sqcup Y, E)$ is *locally finite* if every vertex has finite degree. For $W \subseteq X$ we denote by $N(W) \subseteq Y$ its neighborhood.

Lemma 4.1.10 (P. Hall)**.** *Let $\Gamma = (X \sqcup Y, E)$ be a* finite *bipartite graph. Then there exists a matching saturating all vertices in X if and only if for every subset $W \subseteq X$ one has*

$$|N(W)| \geq |W|.$$

Proof. Necessity. If M is an X-saturating matching and $W \subseteq X$, then each $x \in W$ is matched to a distinct vertex of Y adjacent to x. Hence $|N(W)| \geq |W|$.

Sufficiency. We prove the contrapositive in the standard alternating-paths form. Assume there is no X-saturating matching, and let M be a matching of maximum size. Choose an unmatched vertex $u \in X$. Consider all M-alternating paths starting at u (i.e. paths whose edges alternately lie outside M and inside M). Let $W \subseteq X$ be the set of vertices of X that occur on these paths (including u), and let $Z \subseteq Y$ be the set of vertices of Y that occur on these paths.

Every vertex $z \in Z$ must be matched by M to some vertex in W; otherwise an alternating path would end at an unmatched vertex of Y, and flipping membership of edges along that path would augment M, contradicting maximality. Therefore $|W| \geq |Z| + 1$, where the extra 1 accounts for the unmatched vertex $u \in W$.

On the other hand, every neighbor of any $v \in W$ lies in Z: if w is a neighbor of v, then either $vw \in M$ and removing that matched edge from an alternating path to v yields an alternating path to w, or $vw \notin M$ and adding vw to an alternating path to v yields an alternating path to w. Hence $Z = N(W)$, and thus $|W| \geq |N(W)| + 1$, so Hall's condition is violated. □

Lemma 4.1.11 (M. Hall Jr.)**.** *Let $\Gamma = (X \sqcup Y, E)$ be a bipartite graph such that every vertex in X has finite degree. Assume that for every finite subset $W \subseteq X$ one has*

$$|N(W)| \geq |W|.$$

Then Γ admits a matching saturating all vertices of X.

Proof. This is a compactness argument. □

Lemma 4.1.12 (Schröder–Bernstein)**.** *If there exist injections $f : X \to Y$ and $g : Y \to X$, then there exists a bijection $h : X \to Y$.*

Corollary 4.1.13. *Let $\Gamma = (X \sqcup Y, E)$ be a locally finite bipartite graph. Assume that for every finite $W \subseteq X$ and every finite $Z \subseteq Y$ one has*

$$|N(W)| \geq |W| \qquad \text{and} \qquad |N(Z)| \geq |Z|,$$

where $N(W) \subseteq Y$ and $N(Z) \subseteq X$ are neighborhoods. Then there exists a bijection $\psi : X \to Y$ with $(x, \psi(x)) \in E$ for all $x \in X$.

Proof. By Lemma 4.1.11 there is an injection $f : X \to Y$ with $f(x) \in N(x)$ for all $x \in X$. Applying the same lemma to the opposite bipartite graph yields an injection $g : Y \to X$ with $g(y)$ adjacent to y. By Lemma 4.1.12 there exists a bijection $\psi : X \to Y$. In the standard Schröder–Bernstein construction, ψ agrees with f on a subset of X and with g^{-1} on the complement, so every pair $(x, \psi(x))$ is an edge of Γ. □

Proof of Theorem 4.1.7 (4 ⇒ 1). Assume G is *not* amenable. Then there exist a finite $S \subseteq G$ and $\delta > 0$ such that

$$|SF| \geq (1+\delta)|F| \qquad \text{for all finite } F \subseteq G.$$

Enlarging S if necessary, we may assume $\mathbb{1} \in S$. Choose $n \in \mathbb{N}$ with $(1+\delta)^n \geq 2$ and set $T := S^n$. Then

$$|TF| \geq 2|F| \qquad \text{for all finite } F \subseteq G.$$

Let $X := G \times \{1, 2\}$ and $Y := G$. Join $(g, i) \in X$ to $y \in Y$ if $y = tg$ for some $t \in T$. This graph is locally finite because both T and $\{1, 2\}$ are finite.

For a finite $A \subseteq X$, write $A = (A_1 \times \{1\}) \sqcup (A_2 \times \{2\})$ with $A_1, A_2 \subseteq G$ finite. Then

$$N(A) = TA_1 \cup TA_2 = T(A_1 \cup A_2),$$

so

$$|N(A)| = |T(A_1 \cup A_2)| \geq 2|A_1 \cup A_2| \geq |A_1| + |A_2| = |A|.$$

For a finite $Z \subseteq Y$ one has

$$N(Z) = T^{-1}Z \times \{1, 2\},$$

so $|N(Z)| = 2|T^{-1}Z| \geq 2|Z| \geq |Z|$ since $\mathbb{1} \in T$. By Corollary 4.1.13, there exists a bijection $\psi : X \to Y$ with $(x, \psi(x)) \in E$ for all x. Let $\theta := \psi^{-1} : Y \to X$. For $i \in \{1, 2\}$ and $t \in T$ define

$$A_t^{(i)} := \{g \in G : \ \theta(g) = (t^{-1}g, i)\}.$$

These sets are pairwise disjoint and $\bigsqcup_{i=1}^{2} \bigsqcup_{t \in T} A_t^{(i)} = G$. Moreover, for each i the sets $t^{-1}A_t^{(i)}$ are disjoint and

$$G = \bigsqcup_{t \in T} t^{-1}A_t^{(i)},$$

because θ is bijective onto X and hence hits every $(h, i) \in G \times \{i\}$. This gives a paradoxical decomposition of G with translators in T^{-1}. Thus $\neg(1) \Rightarrow \neg(4)$, hence $4 \Rightarrow 1$. This completes the proof of Theorem 4.1.7. □

4.1.4 Inheritance properties

Amenability is stable under subgroups, quotients, extensions, and directed unions.

Proposition 4.1.14. *If $H \leq G$ and G is amenable, then H is amenable.*

Proof. Let $S \subseteq H$ be finite and $\varepsilon > 0$. Choose a finite (S, ε)-Følner set $F \subseteq G$. Decompose F into left cosets of H:

$$F = \bigsqcup_{i=1}^{m} F_i, \qquad F_i := F \cap Hg_i.$$

Since $S \subseteq H$, each $SF_i \subseteq Hg_i$, and hence

$$|SF \triangle F| = \sum_{i=1}^{m} |SF_i \triangle F_i|.$$

Thus for some i one has $|SF_i \triangle F_i| \leq \varepsilon |F_i|$. Then $F_i g_i^{-1} \subseteq H$ is (S, ε)-Følner in H. □

Remark 4.1.15. The original problem of von Neumann asked whether non-amenability is characterized by the existence of a subgroup isomorphic to $\mathcal{F}_2$:

$$G \text{ amenable} \iff \mathcal{F}_2 \not\leq G.$$

Adyan showed this is false. Day also asked whether every amenable group is elementary amenable, and Grigorchuk's group of intermediate growth gives a negative answer; see [9, 18]. These two questions are often grouped together under the name Day–von Neumann problems.

Proposition 4.1.16. *If $G = \bigcup_{i \in I} G_i$ is a directed union of amenable subgroups, then G is amenable.*

Proof. Let $S \subseteq G$ be finite and $\varepsilon > 0$. Choose $i \in I$ with $S \subseteq G_i$. Since G_i is amenable, there is a finite (S, ε)-Følner set $F \subseteq G_i \subseteq G$. □

Proposition 4.1.17. *Consider a short exact sequence*

$$1 \longrightarrow N \longrightarrow G \xrightarrow{\pi} Q \longrightarrow 1.$$

Then G is amenable if and only if both N and Q are amenable.

Proof. If G is amenable, then N is amenable by Proposition 4.1.14. To see that Q is amenable, let m be a left G-invariant mean on $\ell^\infty(G)$. For $f \in \ell^\infty(Q)$ define $m_Q(f) := m(f \circ \pi)$. If $q \in Q$ and $g \in G$ with $\pi(g) = q$, then $(q \cdot f) \circ \pi = g \cdot (f \circ \pi)$, so m_Q is Q-invariant.

Conversely, assume N and Q are amenable. Let m_N and m_Q be left invariant means on $\ell^\infty(N)$ and $\ell^\infty(Q)$. For $f \in \ell^\infty(G)$ define

$$\overline{f}(q) := m_N\big(n \mapsto f(gn)\big), \qquad q = \pi(g).$$

If $g' = gn_0$ with $n_0 \in N$, then $m_N(n \mapsto f(g'n)) = m_N(n \mapsto f(gn))$ by left invariance of m_N, so $\overline{f}$ is well-defined. Set $m_G(f) := m_Q(\overline{f})$. For $h \in G$ and $q = \pi(g)$ we have

$$\overline{(h \cdot f)}(q) = m_N\big(n \mapsto f(h^{-1}gn)\big) = \overline{f}(\pi(h)^{-1}q),$$

so $\overline{(h \cdot f)} = \pi(h) \cdot \overline{f}$. Hence $m_G(h \cdot f) = m_Q(\pi(h) \cdot \overline{f}) = m_Q(\overline{f})$. Thus m_G is a left G-invariant mean, and G is amenable. □

Definition 4.1.18. Let $\mathcal{P}$ be a group property (for example *abelian, nilpotent*, or *solvable*). A group G is called *virtually* $\mathcal{P}$ if it contains a finite-index subgroup that satisfies $\mathcal{P}$.

Proposition 4.1.19. *Every solvable group is amenable.*

Proof. Every finitely generated abelian group is amenable (finite groups and $\mathbb{Z}^n$ are amenable), and any abelian group is a directed union of finitely generated subgroups, hence amenable by Proposition 4.1.16. If G is solvable of derived length d, then $G^{(1)}$ is solvable of length $d-1$ and $G/G^{(1)}$ is abelian. By induction and Proposition 4.1.17, G is amenable. □

4.2 Elementary amenable groups

4.2.1 Definition and basic examples

Definition 4.2.1. The class of *elementary amenable groups*, denoted EG, is the smallest class of groups that contains all finite and all countable abelian groups and is closed under taking subgroups, quotients, extensions, and countable directed unions.

Elementary amenable groups are built from the most basic amenable groups using the standard permanence operations. This makes EG a tractable subclass of amenable groups, and the next subsection organizes these closure steps via a transfinite hierarchy. Later we will see that EG enjoys a strong growth dichotomy.

Theorem 4.2.2 (Day)**.** *Every solvable group is elementary amenable. In particular, every nilpotent group is elementary amenable.*

Proof. If G is solvable, then the derived series has abelian quotients, so G is built by successive extensions of abelian groups. Hence $G \in \mathrm{EG}$. A nilpotent group is solvable, so it lies in EG as well. □

4.2.2 Transfinite construction

Before describing the transfinite hierarchy of elementary amenable groups we recall the ordinal arithmetic that underlies the induction. Recall that a totally ordered set is called well-ordered if every subset has a minimal element. Ordinals are well-ordered sets up to isomorphism, i.e., they encode the order type of every well-ordering. Each ordinal is either a successor $\alpha+1$, obtained by adjoining a maximum to a previous ordinal, or a limit ordinal, which is the union of all smaller ordinals. It is a basic fact that for all ordinals α, β, either α is a beginning of β or vice versa. The smallest limit ordinal is ω, the order type of the natural numbers. Every ordinal smaller than ω_1, the first uncountable ordinal, is countable, so examples we meet inside this text include ω itself, finite successors like $\omega + 1$, and ordinal arithmetic combinations such as $\omega \cdot 2$ or ω^2, as well as more exotic countable ordinals like ε_0.

Let EG_0 be the class of countable finite and abelian groups. For an ordinal α, define EG_α by

$$\mathrm{EG}_\alpha = \bigcup_{\beta<\alpha} \mathrm{EG}_\beta \quad \text{if } \alpha \text{ is a limit ordinal,}$$

and if α is a successor, let EG_α be the class of groups obtained from $\mathrm{EG}_{\alpha-1}$ by a single step of taking a countable directed union or an extension by a finite or abelian group. The *construction rank* of $G \in \mathrm{EG}$ is the least α with $G \in \mathrm{EG}_\alpha$; it is always 0 or a successor ordinal.

Lemma 4.2.3. *For every ordinal α, the class EG_α is closed under subgroups and quotients.*

Proof. Argue by transfinite induction. The case $\alpha = 0$ is clear. For a limit ordinal, the claim follows from the induction hypothesis. If α is a successor and $G \in \mathrm{EG}_\alpha$, then G is either a directed union of groups in $\mathrm{EG}_{\alpha-1}$ or an extension of a group in $\mathrm{EG}_{\alpha-1}$ by a finite or abelian group. Subgroups and quotients of G are then built from subgroups and quotients in $\mathrm{EG}_{\alpha-1}$, hence lie in EG_α. □

Theorem 4.2.4 (Chou)**.** *The class* EG *is the smallest class of groups containing all finite and abelian groups and closed under countable directed unions and extensions by finite or countable abelian groups, in fact we have*

$$\mathrm{EG} = \bigcup_{\alpha<\omega_1} \mathrm{EG}_\alpha.$$

Proof. Let $\mathcal{C} = \bigcup_{\alpha<\omega_1} \mathrm{EG}_\alpha$. By construction $\mathcal{C}$ is closed under directed unions and extensions by finite or abelian groups, and by Lemma 4.2.3 it is closed under subgroups and quotients. It remains to be shown that $\mathcal{C}$ is closed under arbitrary extensions and countable directed unions. We denote by $\mathrm{EG}_{\alpha,\beta}$ the class of extensions of groups in EG_α by groups in EG_β. By construction $\mathrm{EG}_{\alpha,0} \subset \mathrm{EG}_{\alpha+1} \subseteq \mathcal{C}$ for all α. A straigtforward induction argument shows that $\mathrm{EG}_{\alpha,\beta} \subseteq \mathrm{E}_{\alpha+\beta+1} \subseteq \mathcal{C}$ for all ordinals α, β. Let G be a group that is a countable directed union of groups in $\mathcal{C}$, say $G = \cup_{i\in I} G_i$ with $G_i \in \mathrm{EG}_{\alpha_i}$. Since I is countable $\alpha := \sup_i \alpha_i < \omega_1$ and hence $G_i \in \mathrm{EG}_\alpha$ for all $i \in I$ and thus $G \in \mathrm{EG}_{\alpha+1}$. This finally implies $\mathrm{EG} \subseteq \mathcal{C}$ by definition of EG, while $\mathcal{C} \subseteq \mathrm{EG}$ by definition of the transfinite construction. □

For $G \in \mathrm{EG}$ we write $\mathrm{rk}(G)$ for its construction rank:

$$\mathrm{rk}(G) = \min\{\alpha : G \in \mathrm{EG}_\alpha\}.$$

It is clear from the definition that $\mathrm{rk}(G)$ is always a countable ordinal. Moreover, every countable ordinal arises as the construction rank of some elementary amenable group, see [6].

4.2.3 First applications

Proposition 4.2.5. *Every finitely generated simple elementary amenable group is finite.*

Proof. Let G be finitely generated and simple, and let α be minimal with $G \in \mathrm{EG}_\alpha$. Then α is not a limit ordinal. If $\alpha > 0$, the construction of EG_α writes G as a directed union of groups in $\mathrm{EG}_{\alpha-1}$, so finite generation forces G to lie in one of them, contradicting minimality. Hence $\alpha = 0$ and G is finite. □

Theorem 4.2.6 (Chou)**.** *Every torsion elementary amenable group is locally finite.*

Proof. Proceed by transfinite induction on the construction rank. If $G \in \mathrm{EG}_\alpha$ with α a limit ordinal, then $G \in \mathrm{EG}_\beta$ for some $\beta < \alpha$ and the result follows by induction. If α is a successor, then G is either a directed union of groups in $\mathrm{EG}_{\alpha-1}$ or an extension

$$1 \longrightarrow K \longrightarrow G \longrightarrow F \longrightarrow 1$$

with $K \in \mathrm{EG}_{\alpha-1}$ and F finite or abelian. In the union case, every finitely generated subgroup of G lies in some $\mathrm{EG}_{\alpha-1}$ and is finite by the induction hypothesis. In the extension case, a finitely generated subgroup $H \leq G$ maps to a finite subgroup of F, so $H \cap K$ has finite index in H; since $H \cap K$ is finite by induction, H is finite. Thus G is locally finite. □

4.2.4 Growth dichotomy

We use the growth terminology from Definition 3.1.5. For a finitely generated group G we write γ_G for its growth function.

We first establish a growth dichotomy for elementary amenable groups in Theorem 4.2.7. The lemmas below provide the kernel and linear-algebra inputs for that argument, and Milnor–Wolf will then follow as a corollary.

Theorem 4.2.7 (Chou)**.** *Every elementary amenable group has either polynomial or exponential growth. Moreover, if it has polynomial growth, then it is virtually nilpotent.*

The proof is by induction on construction rank, reducing to abelian extensions. The key steps are Milnor's lemma on finite generation of kernels in subexponential growth and Tits's criterion for virtual nilpotence.

The following more classical theorem of Milnor–Wolf is a direct consequence:

Theorem 4.2.8 (Milnor–Wolf)**.** *Every solvable group has either polynomial or exponential growth. Moreover, if it has polynomial growth, then it is virtually nilpotent.*

Proof of Theorem 4.2.8. By Theorem 4.2.2, solvable groups are elementary amenable. The dichotomy therefore follows from Theorem 4.2.7. See [37, 55] for the original proofs. □

Let's provide the necessary tools in the following lemmas.

Lemma 4.2.9 (Tits)**.** *Let $g \in GL_n(\mathbb{Z})$.*

(i) If g is semisimple and all eigenvalues have absolute value 1, then g has finite order.

(ii) If g has an eigenvalue with absolute value at least 2, then there exists $a \in \mathbb{Z}^n$ such that for each $m \geq 0$ the elements

$$\sum_{i=0}^{m} \varepsilon_i g^i(a), \qquad \varepsilon_i \in \{0,1\},$$

are pairwise distinct.

Proof. We give two proofs, one geometric and one algebraic. *Geometric proof:* Over $\mathbb{C}$ one may diagonalize g. Since all eigenvalues have modulus 1, g acts by isometries for some norm, so orbits in $\mathbb{C}^n$ are bounded. The orbit of any $v \in \mathbb{Z}^n$ is then a bounded discrete set, hence finite. Thus g has finite order.

Algebraic proof: The eigenvalues of g are algebraic integers, since they are roots of the monic characteristic polynomial with coefficients in $\mathbb{Z}$. Any Galois conjugate of an eigenvalue is again an eigenvalue, so all conjugates have modulus 1. By Kronecker's theorem (Theorem A.2.5) each eigenvalue is a root of unity. Since g is semisimple, it is diagonalizable over $\mathbb{C}$ with diagonal entries roots of unity. If N is the least common multiple of their orders, then $g^N = \mathrm{id}$, so g has finite order.

(2) Let λ be an eigenvalue with $|\lambda| \geq 2$. Take a non-zero Jordan block for λ with basis $e_1, \dots, e_k$ such that $g(e_1) = \lambda e_1$ and $g(e_i) = \lambda e_i + e_{i-1}$ for $i > 1$. Let $\gamma : \mathbb{C}^n \to \mathbb{C}$ read off the coefficient of e_k. Then $\gamma \circ g = \lambda\,\gamma$, so $\gamma(g^i(v)) = \lambda^i \gamma(v)$ for every v. Choose $a \in \mathbb{Z}^n$ with $\gamma(a) \neq 0$. Using $\gamma(g^i(a)) = \lambda^i \gamma(a)$ we obtain

$$\gamma\left(\sum_{i=0}^{m} \varepsilon_i g^i(a)\right) = \left(\sum_{i=0}^{m} \varepsilon_i \lambda^i\right) \gamma(a).$$

Distinct binary choices of $(\varepsilon_i)_{i=1}^m$ give distinct sums because $|\lambda| \geq 2$, hence the elements are distinct. $\square$

Lemma 4.2.10. *Let $g \in GL_n(\mathbb{Z})$. Then either g is semisimple or there exists a non-trivial g-invariant subgroup $A \lneq \mathbb{Z}^n$ of smaller rank.*

Proof. If g is not semisimple, let $\lambda_1, \ldots, \lambda_m$ be its distinct eigenvalues and set

$$p(x) \coloneqq \prod_{i=1}^{m}(x - \lambda_i), \qquad h \coloneqq p(g).$$

Each λ_i is an algebraic integer. The coefficients of p are elementary symmetric polynomials in the λ_i, hence algebraic integers that also lie in $\mathbb{Q}$; by Proposition A.2.4 they are integers, so $p \in \mathbb{Z}[x]$ and h is an integer matrix that commutes with g.

In Jordan normal form over $\mathbb{C}$, p vanishes on the eigenvalues and has simple roots. On a Jordan block $J = \lambda I + N$ with $N \neq 0$ we have $p(J) = p'(\lambda)N +$ (higher powers of N), so $p(J)$ is nilpotent and non-zero whenever the block size exceeds 1. Thus h is nilpotent and non-zero, because g is not semisimple, some block has size > 1. Set $A \coloneqq \ker h \leq \mathbb{Z}^n$. Since h commutes with g, A is g-invariant. Nilpotency of h forces $\ker h \neq \{0\}$, while non-zeroness of h guarantees $\ker h \lneq \mathbb{Z}^n$ of smaller rank. $\square$

Lemma 4.2.11 (Milnor). *Let $1 \to K \to G \xrightarrow{\psi} A \to 0$ be exact with G finitely generated and A abelian. If G has subexponential growth, then K is finitely generated. In particular, $[G, G]$ is finitely generated for every group G of subexponential growth.*

Proof. We follow [37]. By induction on the free rank of A it suffices to treat $A = \mathbb{Z}$. Choose $g \in G$ with $\psi(g) = 1$ and elements $e_1, \ldots, e_n \in K$ such that $G = \langle g, e_1, \ldots, e_n\rangle$. For $m \in \mathbb{Z}$ set

$$g_{m,i} \coloneqq g^m e_i g^{-m}.$$

We claim that $K = \langle g_{m,i} \mid m \in \mathbb{Z},\ 1 \leq i \leq n\rangle$. Indeed, write $k \in K$ as a word in $g^{\pm 1}$ and $e_i^{\pm 1}$ and induct on word length, pushing powers of g to the right. Since $\psi(k) = 0$, the total exponent of g vanishes, so one obtains an expression in the $g_{m,i}$.

Fix i and consider the words

$$g_{0,i}^{\varepsilon_0} g_{1,i}^{\varepsilon_1} \cdots g_{m,i}^{\varepsilon_m}, \qquad \varepsilon_j \in \{0, 1\}.$$

There are 2^{m+1} such words, and after cancellation each has length at most $3(m + 1)$ in the generating set $\{g, e_1, \ldots, e_n\}$. Indeed, writing $g_{j,i} = g^j e_i g^{-j}$, the product

$$g_{0,i}^{\varepsilon_0} g_{1,i}^{\varepsilon_1} \cdots g_{m,i}^{\varepsilon_m}$$

is a word in $g^{\pm 1}$ and e_i in which, between successive occurrences of e_i, the g-powers telescope: $g^{-j}g^{j+1} = g$. Hence, after free reduction, each factor contributes at most one letter e_i and at most two letters $g^{\pm 1}$, giving the bound $\leq 3(m + 1)$.

Since G has subexponential growth, for large m we have $2^{m+1} > \gamma_G(3(m + 1))$. Thus the 2^{m+1} words above cannot all represent distinct elements of the ball of radius $3(m+1)$, so two distinct choices of $(\varepsilon_0, \ldots, \varepsilon_m)$ yield the same group element. Let t be the maximal index with $\varepsilon_t \neq \varepsilon'_t$ for two such distinct choices $(\varepsilon_0, \ldots, \varepsilon_m)$ and $(\varepsilon'_0, \ldots, \varepsilon'_m)$. Then the corresponding words have the same suffix $g_{t+1,i}^{\varepsilon_{t+1}} \cdots g_{m,i}^{\varepsilon_m}$, so cancelling this common tail and rearranging yields a relation expressing $g_{t,i}$ as a word in the earlier conjugates $g_{0,i}, \ldots, g_{t-1,i}$. In particular, there exists $m_i < m$ such that

$$g_{m_i+1,i} \in \langle g_{0,i}, \ldots, g_{m_i,i} \rangle.$$

Conjugating by powers of g then shows recursively that every $g_{m,i}$ with $m \geq 0$ lies in $\langle g_{0,i}, \ldots, g_{m_i,i} \rangle$. For $m < 0$ one argues analogously, applying the same counting argument to the generator g^{-1} (equivalently, considering the conjugates $g^{-m} e_i g^m$). Hence all $g_{m,i}$ lie in the subgroup generated by the set $\{g_{j,i} : |j| \leq m_i\}$. Therefore K is finitely generated. □

Lemma 4.2.12 (Tits)**.** *Let* $1 \to K \to G \xrightarrow{\psi} A \to 1$ *be exact with* G *finitely generated and* A *abelian. If* G *has subexponential growth and* K *is virtually nilpotent, then* G *is virtually nilpotent.*

Proof. We follow [51]. By Lemma 4.2.11, K is finitely generated. Replacing K by a finite-index characteristic nilpotent subgroup and G by the subgroup it generates together with a lift of A, we may assume K is nilpotent. Indeed, since K is virtually nilpotent there exists a nilpotent subgroup $K' \leq K$ of finite index. Because K is finitely generated, for $m = [K : K']$ there are only finitely many subgroups of K of index m, so the orbit $\{\varphi(K') \mid \varphi \in \mathrm{Aut}(K)\}$ is finite. Hence

$$K_{\mathrm{char}} \coloneqq \bigcap_{\varphi \in \mathrm{Aut}(K)} \varphi(K')$$

is a finite-index characteristic subgroup of K, contained in K' and therefore nilpotent. Since $K \triangleleft G$ and K_{char} is characteristic in K, we also have $K_{\mathrm{char}} \triangleleft G$. Replacing K by K_{char} and G by the subgroup generated by K together with a lift of A changes G only up to finite index, so it suffices to prove virtual nilpotence for these replacements. By induction on the free rank of A it suffices to treat $A = \mathbb{Z}$. Let $g \in G$ map to a generator of A.

Choose a central series (so $K_i/K_{i-1} \leq Z(K/K_{i-1})$ for all i, which exists since K is nilpotent)

$$\{1\} = K_0 \triangleleft K_1 \triangleleft \cdots \triangleleft K_r = K$$

and refine it so that each factor K_i/K_{i-1} is either finite or free abelian and has no non-trivial g-invariant subgroup of smaller rank. Concretely, start with any central series of K (for example the upper central series). Each quotient K_i/K_{i-1} is finitely generated abelian, hence a direct sum of a finite torsion subgroup and a free part. For each free part, viewed as a $\mathbb{Z}$-lattice with the automorphism induced by conjugation by g, insert a non-trivial g-invariant subgroup of minimal positive rank; iterating yields a finite refinement in which the free factors have no proper non-trivial g-invariant subgroup of smaller rank. On each infinite factor K_i/K_{i-1} the conjugation action of g lies in $GL_n(\mathbb{Z})$ and is semisimple by Lemma 4.2.10. If for some factor g had an eigenvalue with $|\lambda| > 1$, then Lemma 4.2.9(2) would produce exponentially many distinct elements of bounded word length, contradicting subexponential growth. Hence all eigenvalues have modulus 1, and Lemma 4.2.9(1) implies g acts with finite order on each factor. Thus there is $N \geq 1$ such that g^N acts trivially on every K_i/K_{i-1}. Equivalently, for each i we have

$$[g^N, K_i] \subseteq K_{i-1},$$

so the image of g^N in G/K_{i-1} centralizes the image of K_i. Since $(K_i/K_{i-1}) \leq Z(K/K_{i-1})$ by centrality of the series, it follows that

$$K_i/K_{i-1} \leq Z(\langle K, g^N \rangle / K_{i-1}).$$

Consequently the chain

$$\{1\} = K_0 \lhd K_1 \lhd \cdots \lhd K_r = K \lhd \langle K, g^N \rangle$$

is a central series for $\langle K, g^N \rangle$ (the last quotient $\langle K, g^N \rangle / K$ is cyclic, hence abelian), so $\langle K, g^N \rangle$ is nilpotent. Therefore G is virtually nilpotent. □

Proof of Theorem 4.2.7. We follow [6]. Let G be a finitely generated elementary amenable group. If G has exponential growth there is nothing to prove, so assume G has subexponential growth. We show that G is virtually nilpotent, hence has polynomial growth. Proceed by induction on the construction rank of G. The base case is clear. Suppose $\mathrm{rk}(G) = \alpha + 1$. Then there exists $L \lhd G$ with $\mathrm{rk}(L) \leq \alpha$ and G/L finite or abelian. By Lemma 4.2.11, the subgroup L is finitely generated. It cannot have exponential growth, hence by induction L is virtually nilpotent. If G/L is finite, then G is virtually nilpotent. If G/L is abelian, apply Lemma 4.2.12 to conclude that G is virtually nilpotent. Thus G has polynomial growth. □

Famously, Gromov has extended Theorem 4.2.7 under the stronger hypothesis of polynomial growth, see Theorem 7.4.5 for a proof.

4.3 Subexponential growth

4.3.1 Definitions and examples

Recall from Definition 3.1.5 that G has subexponential growth if $\limsup_{n\to\infty} \gamma_G(n)^{1/n} = 1$. Subexponential growth forces amenability and leads to the first examples of intermediate growth. We also record Grigorchuk's theorem explicitly.

Theorem 4.3.1 (Følner)**.** *Let G be finitely generated. If G has subexponential growth, then G is amenable.*

Proof. Let S be a finite symmetric generating set and $B(n) = B_S(n)$. Fix $\varepsilon > 0$. If $|B(n+1)| > (1+\varepsilon)|B(n)|$ for all $n \geq n_0$, then

$$|B(n)| \geq (1+\varepsilon)^{n-n_0} |B(n_0)|,$$

so G has exponential growth, a contradiction. Hence there exists n with

$$|B(n+1)| \leq (1+\varepsilon)|B(n)|.$$

Since $SB(n) = B(n+1)$, the set $B(n)$ is (S, ε)-Følner. By the finitely generated case of the Følner condition, G is amenable. □

See [42, 26] for further discussion.

Theorem 4.3.2 (Grigorchuk)**.** *There exists a finitely generated group of intermediate growth, i.e. subexponential but not polynomial.*

Remark 4.3.3. By Theorem 4.3.1 these groups are amenable. They show that intermediate growth is possible and provide counterexamples to Day's question about elementary amenability; see [18, 26, 21].

4.3.2 The Gap Conjecture

Recall the growth function and comparison $\preceq$ from Definition 3.1.5, and write $f \prec g$ if $f \preceq g$ but $f \not\simeq g$. The relation $\gamma_{G,S}(n) \prec \exp(\sqrt{n})$ is a statement about growth type, not a pointwise inequality.

Conjecture 4.3.4. *Let G be finitely generated. If $\gamma_{G,S}(n) \prec \exp(\sqrt{n})$, then G has polynomial growth, equivalently, is virtually nilpotent by Gromov's theorem 7.4.5.*

The conjecture is open in general; see [20]. Grigorchuk shows that it suffices to prove it for residually finite groups [20]. Known cases include residually nilpotent groups (Gap(1/2) holds; see [20] for references to work of Grigorchuk and Lubotzky–Mann) and residually solvable groups, where Wilson proved a gap with exponent 1/6 and Eberhard–Maini improved this to Gap(β) for all $\beta < 1/4$ [20, 11].

4.4 Amenability and geometry

We now return to Cayley graphs of finitely generated groups. Fix a finite symmetric generating set S for G and write $\Gamma = \mathrm{Cay}(G, S)$. We record two geometric rigidity phenomena due to Whyte [54] and state them directly in this setting.

4.4.1 Trees inside non-amenable Cayley graphs

Definition 4.4.1. Let G be finitely generated. A (left) action of a group H on G is called *translation-like* if it is free and for each $h \in H$ there exists $R_h < \infty$ with

$$d_S(g, h \cdot g) \leq R_h \qquad (\forall g \in G).$$

Definition 4.4.2. Let $\Gamma = \mathrm{Cay}(G, S)$. A spanning subgraph $T \subseteq \Gamma$ is called a *bounded-length* 4*-regular treeing* if every connected component of T is a 4-regular tree and each edge of T is realized by a path of length at most L in Γ for some fixed L.

Theorem 4.4.3 (Whyte)**.** *Let G be finitely generated. The following are equivalent:*

(i) G is non-amenable.

(ii) Γ admits a translation-like action of the free group $\mathcal{F}_2$.

(iii) Equivalently, Γ contains a bounded-length 4-regular treeing.

Proof. Assume G is non-amenable. Then there exists $\varepsilon > 0$ with $|SF| \geq (1+\varepsilon)|F|$ for every finite $F \subseteq G$. Choose n with $(1+\varepsilon)^n \geq 2$ and set $T = S^n$, so $|TF| \geq 2|F|$ for all finite F. As in the proof of Theorem 4.1.7 ($4 \Rightarrow 1$), consider the bipartite graph with left side $G \times \{1, 2\}$ and right side G, where (g, i) is adjacent to tg for $t \in T$. Hall's condition holds on both sides, hence by Corollary 4.1.13 there is a bijection

$$\theta : G \times \{1, 2\} \to G, \qquad \theta(g, i) = t_i(g)g, \quad t_i(g) \in T.$$

The maps $f_i(g) \coloneqq \theta(g, i)$ are injective with disjoint images and move points by at most n in the word metric. Thus G admits a translation-like action of the free semigroup on two generators. Whyte shows how to symmetrize this action (using the marriage lemma to construct bounded-distance inverses) to obtain a free action of $\mathcal{F}_2$ by bijections of bounded displacement. The associated graph with edges g–ag and g–bg (for free generators a, b) is a bounded-length 4-regular treeing of Γ. Conversely, such a treeing canonically defines a translation-like action of $\mathcal{F}_2$ by moving along its edges.

Conversely, if Γ contains such a treeing T, then T is bilipschitz to the 4-regular tree, so there is $c > 0$ with $|\partial_T A| \geq c|A|$ for all finite $A \subseteq G$, where $\partial_T A$ denotes the set of edges of T with exactly one endpoint in A. If G were amenable, there would be Følner sets $F \subseteq G$ for S, i.e. finite F with $|SF| \leq (1+\varepsilon)|F|$ for arbitrarily small ε. Let L be the treeing length and set $F' = B_S(L)F$. Each edge of $\partial_T F'$ is realized by a path of length at most L in Γ that crosses the S-boundary of $B_S(L-1)F$, so bounded degree gives $|\partial_T F'| \leq C_L\,|SB_S(L-1)F \setminus B_S(L-1)F|$ and $|F'| \geq c_L|F|$ for constants depending only on L. Taking ε small forces $|\partial_T F'|/|F'|$ small, contradicting the lower bound for T. Hence G is non-amenable. □

Remark 4.4.4. In particular, for non-amenable G the Cayley graph contains a regular tree quasi-isometrically embedded at bounded distance from a spanning subgraph. This is a geometric replacement for the (false) subgroup version of the von Neumann conjecture.

Conjecture 4.4.5 (Benjamini)**.** *Let G be a finitely generated group of exponential growth. Then some (equivalently any) Cayley graph of G contains a quasi-isometrically embedded regular tree of degree ≥ 3.*

Remark 4.4.6. Conjecture 4.4.5 is known for large classes (e.g. connected Lie groups and finitely generated solvable groups of exponential growth admit quasi-isometrically embedded free semigroups, hence trees), and it holds for all *non-amenable* groups by Theorem 4.4.3; see [8]. Related tree results for non-amenable graphs appear in [5].

Conjecture 4.4.7 (Lyons)**.** *Every finitely generated group of exponential growth admits a spanning subtree of its Cayley graph with exponential growth (hence containing large binary subtrees). This would not by itself give a quasi-isometric embedding of a regular tree, but it would show that some strong tree-like subgraphs exist in complete generality; see [36].*

4.4.2 Bilipschitz rigidity for non-amenable groups

A quasi-isometry need not be close to a bijection: for example $f : \mathbb{Z} \to \mathbb{Z}$, $f(n) = 2n$, is a quasi-isometry but cannot be at bounded distance from any bijection. Non-amenability rules out this phenomenon.

Theorem 4.4.8 (Whyte)**.** *Let G, H be finitely generated non-amenable groups with Cayley graphs Γ_G, Γ_H, and let $f : \Gamma_G \to \Gamma_H$ be a quasi-isometry. Then there exists a bilipschitz equivalence $g : \Gamma_G \to \Gamma_H$ such that*

$$\sup_{x \in \Gamma_G} d\big(f(x), g(x)\big) < \infty.$$

In particular, quasi-isometric non-amenable groups have Cayley graphs that are bilipschitz equivalent (up to bounded distance).

Proof. Write $\Gamma_H = \mathrm{Cay}(H, S_H)$. Let f be a (λ, C)-quasi-isometry. Since f is coarsely onto, there exists R_0 such that every vertex of Γ_H lies within R_0 of $f(\Gamma_G)$. Since Γ_G has bounded degree, the fiber size of f is uniformly bounded: there exists M such that $|f^{-1}(y)| \leq M$ for all y.

Non-amenability of H yields $\varepsilon > 0$ with $|S_H F| \geq (1+\varepsilon)|F|$ for all finite $F \subseteq H$, hence $|B_{S_H}(n)F| \geq (1+\varepsilon)^n|F|$ for all $n \geq 1$. Choose n with $(1+\varepsilon)^n \geq M$ and set $R = R_0 + n$. Consider the bipartite graph with left side G and right side H, where $g \in G$ is joined to $h \in H$ if $d_H\big(f(g), h\big) \leq R$. For a finite $A \subseteq G$ the neighborhood $N(A)$ contains the R-neighborhood of $f(A)$, so

$$|N(A)| \geq |B_{S_H}(n)f(A)| \geq M\,|f(A)| \geq |A|.$$

Hall's condition holds, so by Lemma 4.1.11 there is an injection $\iota : G \to H$ with $d_H\big(f(g), \iota(g)\big) \leq R$. Applying the same argument to a quasi-inverse of f yields an injection $H \to G$. By Lemma 4.1.12 there exists a bijection $g : G \to H$ at uniformly bounded distance from f.

Since g is a bijective quasi-isometry between graphs with integer distances, the additive constants can be absorbed into the multiplicative constant, so g is bilipschitz. This yields the conclusion. □

Chapter 5

Gap Conjecture

5.1 Introduction

This chapter is an expanded and largely self-contained exposition of techniques originating in Grigorchuk's work on Hilbert–Poincaré series and graded algebras associated with groups. The aim is to prove the following result:

Theorem 5.1.1 (Grigorchuk)**.** *Let G be finitely generated and residually p-finite. If $\gamma_G(n) \prec e^{\sqrt{n}}$, then G is virtually nilpotent.*

We follow Grigorchuk's original approach [19] with some simplifications and streamlining. Grigorchuk conjectured that the preceding theorem holds without the residual p-finiteness assumption; this is known as the *Gap Conjecture*, see also Section 4.3.2.

5.2 Lie-algebraic techniques

Fix a prime p and write $\mathbb{F}_p$ for the field with p elements. Let G be a group.

5.2.1 Group algebra, augmentation ideal, and associated graded algebra

Let $\mathbb{F}_p[G]$ be the group algebra. The *augmentation map* $\varepsilon : \mathbb{F}_p[G] \to \mathbb{F}_p$ is defined by $\varepsilon(\sum a_g g) = \sum a_g$. Its kernel

$$\Delta(G) := \ker(\varepsilon)$$

is the *augmentation ideal.* The powers $\Delta(G)^n$ form a decreasing filtration of $\mathbb{F}_p[G]$, and we define the associated graded algebra

$$\mathscr{A}_G := \bigoplus_{n\geq 0} A_n(G), \qquad A_0(G) := \mathbb{F}_p, \quad A_n(G) := \Delta(G)^n/\Delta(G)^{n+1} \ (n \geq 1),$$

with multiplication induced from $\mathbb{F}_p[G]$.

Definition 5.2.1. The *Hilbert–Poincaré series* of $\mathscr{A}_G$ is

$$f_G(t) := \sum_{n=0}^{\infty} a_n(G)\, t^n, \qquad a_n(G) := \dim_{\mathbb{F}_p} A_n(G). \tag{5.1}$$

The series exists as a formal power series as soon as $a_n(G) < \infty$ for all n. A necessary and sufficient condition for (5.1) to define a power series with positive radius of convergence is that $a_1(G) < \infty$:

$$d := a_1(G) = \dim_{\mathbb{F}_p} \Delta(G)/\Delta(G)^2 < \infty. \tag{5.2}$$

Note that $\Delta(G)/\Delta(G)^2$ is a quotient of a direct sum of copies of $\mathbb{F}_p$ indexed by a generating set of G, so $d < \infty$ holds for finitely generated groups.

5.2.2 Zassenhaus filtration

Define the *lower p-central series* or *Zassenhaus series* $(G_n)_{n\geq 1}$ by $G_1 := G$ and

$$G_n := [G, G_{n-1}]\; G^p_{\lceil n/p \rceil} \qquad (n \geq 2), \tag{5.3}$$

where $\lceil n/p \rceil$ is the smallest integer $\geq n/p$. Equivalently (cf. standard references on dimension subgroups),

$$G_n = \prod_{jp^i \geq n} \gamma_j(G)^{p^i}, \tag{5.4}$$

where $\gamma_j(G)$ is the lower central series ($\gamma_1(G) = G$, $\gamma_{j+1}(G) = [G, \gamma_j(G)]$).

Assume throughout that

$$|G/G_2| = |G/[G,G]G^p| < \infty, \tag{5.5}$$

equivalently $d = \dim_{\mathbb{F}_p} H_1(G; \mathbb{F}_p) < \infty$, which matches (5.2).

A group G is said to be *residually p-finite* if there exists a descending sequence of normal subgroups of p-power index with trivial intersection. In the present context, one standard sufficient condition is

$$\bigcap_{n\geq 1} G_n = \{1\},$$

together with (5.5), ensuring that the natural maps $G \to G/G_n$ approximate G by finite p-groups.

5.2.3 Associated Lie p-algebra

Definition 5.2.2. Let $\Bbbk$ be a field of characteristic $p > 0$.

A *restricted Lie p-algebra* (or *restricted Lie algebra*) is a Lie algebra $\mathscr{L}$ over $\Bbbk$ together with a map (the *p-operation*) $(-)^{[p]} : \mathscr{L} \to \mathscr{L}$, $x \mapsto x^{[p]}$, such that for all $x, y \in \mathscr{L}$ and $a \in \Bbbk$,

$$(ax)^{[p]} = a^p x^{[p]}, \qquad \mathrm{ad}(x^{[p]}) = \mathrm{ad}(x)^p, \qquad (x+y)^{[p]} = x^{[p]} + y^{[p]} + \sum_{i=1}^{p-1} s_i(x,y).$$

Here the terms $s_i(x,y)$ are the standard universal Lie polynomials (in iterated brackets of x and y), characterized by the requirement that

$$(\mathrm{ad}(x+y))^p = (\mathrm{ad}\, x)^p + (\mathrm{ad}\, y)^p + \sum_{i=1}^{p-1} \mathrm{ad}\big(s_i(x,y)\big).$$

A *graded restricted Lie p-algebra* is a restricted Lie p-algebra together with an $\mathbb{N}$-grading $\mathscr{L} = \bigoplus_{n\geq 1} \mathscr{L}_n$ such that $[\mathscr{L}_m, \mathscr{L}_n] \subseteq \mathscr{L}_{m+n}$ and $x \in \mathscr{L}_n \Rightarrow x^{[p]} \in \mathscr{L}_{pn}$. (More generally, for a grading by an abelian group A written additively, the grading compatibility is $\deg(x^{[p]}) = p \deg(x)$.)

Set

$$L_n(G) := G_n/G_{n+1}, \qquad \mathscr{L}_G := \bigoplus_{n\geq 1} L_n(G).$$

The commutator induces a graded Lie bracket $[\,\cdot,\cdot\,] : L_m(G) \times L_n(G) \to L_{m+n}(G)$, and the p-power map induces a restricted operation $x \mapsto x^{[p]}$ sending $L_n(G) \to L_{pn}(G)$. Thus $\mathscr{L}_G$ is a graded restricted Lie p-algebra.

Define

$$b_n(G) := \dim_{\mathbb{F}_p} L_n(G) = \log_p |G_n/G_{n+1}| \,. \tag{5.6}$$

We use the following external structural theorem that relates the graded algebra $\mathscr{A}_G$ to the graded restricted Lie p-algebra $\mathscr{L}_G$.

Theorem 5.2.3 (Quillen)**.** *There is a natural isomorphism of graded algebras*

$$\mathscr{A}_G \cong U(\mathscr{L}_G),$$

the universal enveloping algebra of the restricted Lie p-algebra $\mathscr{L}_G$.

A standard consequence due to Jennings is the identity

$$f_G(t) = \sum_{n\geq 0} a_n(G)t^n = \prod_{n=1}^{\infty} \left(\frac{1-t^{pn}}{1-t^n}\right)^{b_n(G)}. \tag{5.7}$$

Proof of (5.7). In a graded restricted Lie algebra, the PBW theorem implies that the Hilbert series of $U(\mathscr{L}_G)$ is the Euler product (5.7) with exponents $b_n(G)$. Since $U(\mathscr{L}_G) \cong \mathscr{A}_G$ by Theorem 5.2.3, the same identity holds for $f_G(t)$. □

5.2.4 Lie-theoretic structure

Lemma 5.2.4. *$\mathscr{L}_G$ is generated as a restricted Lie p-algebra by its degree-one component $L_1(G)$.*

Proof. Let $\mathscr{L}^*$ be the restricted Lie subalgebra generated by $L_1(G)$. We prove by induction on n that $L_n(G) \subset \mathscr{L}^*$.

For $n = 1$ this is true. Assume $L_{n-1}(G) \subset \mathscr{L}^*$. By (5.3), every element of G_n is a product of commutators $[g,h]$ with $g \in G$, $h \in G_{n-1}$, and p-th powers of elements in $G_{\lceil n/p\rceil}$. Modulo G_{n+1} this yields

$$L_n(G) \subset [L_1(G), L_{n-1}(G)] + (L_{\lceil n/p\rceil}(G))^{[p]}.$$

By the inductive hypothesis, $L_{n-1}(G) \subset \mathscr{L}^*$, so the bracket term lies in $\mathscr{L}^*$. Also, $\lceil n/p\rceil < n$ for $n \geq 2$, so $L_{\lceil n/p\rceil}(G) \subset \mathscr{L}^*$, and applying the restricted power $[p]$ remains in $\mathscr{L}^*$. Hence $L_n(G) \subset \mathscr{L}^*$. □

Lemma 5.2.5. *Let $\mathscr{L}_* \subset \mathscr{L}_G$ be the graded Lie subalgebra generated (as an ordinary Lie algebra) by $L_1(G)$. Then $[\mathscr{L}_G, \mathscr{L}_G] \subset \mathscr{L}_*$.*

Proof. In a restricted Lie algebra one has the identity

$$[x, y^{[p]}] = (\operatorname{ad} y)^p(x) \qquad (x, y \in \mathscr{L}_G).$$

Since $(\operatorname{ad} y)^p(x)$ is an iterated Lie bracket of x with y, it belongs to the Lie subalgebra generated by x and y. In particular, if $x \in L_1(G)$ and $y \in \mathscr{L}_G$, then $[x, y^{[p]}]$ lies in the ordinary Lie subalgebra generated by $L_1(G)$, i.e. in $\mathscr{L}_*$.

By Lemma 5.2.4, $\mathscr{L}_G$ is generated as a restricted Lie algebra by $L_1(G)$, hence any element of $\mathscr{L}_G$ is obtained from $L_1(G)$ using Lie brackets and $[p]$-powers. Using the identity above, any bracket of two such elements can be rewritten as a linear combination of iterated brackets involving only elements of degree 1 (no restricted powers remain inside brackets), hence lies in the ordinary Lie subalgebra generated by $L_1(G)$, i.e. in $\mathscr{L}_*$. □

Lemma 5.2.6. *Let $\mathscr{L} = \bigoplus_{n\geq 1} \mathscr{L}_n$ be an $\mathbb{N}$-graded restricted Lie algebra over $\mathbb{F}_p$ generated as a restricted Lie algebra by $\mathscr{L}_1$. Write $[\mathscr{L}, \mathscr{L}]$ for its commutator ideal. Then for every $n \geq 2$ one has:*

(a) If $p \nmid n$, then $\mathscr{L}_n \subseteq [\mathscr{L}, \mathscr{L}]_n$.

(b) If $p \mid n$, then

$$\mathscr{L}_n \subseteq [\mathscr{L}, \mathscr{L}]_n \; + \; \left\langle x^{[p]} : x \in \mathscr{L}_{n/p} \right\rangle_{\mathbb{F}_p}.$$

Proof. Since $\mathscr{L}$ is generated by $\mathscr{L}_1$ as a restricted Lie algebra, every homogeneous element of degree n is an $\mathbb{F}_p$-linear combination of *restricted Lie monomials* in elements of $\mathscr{L}_1$. Such a monomial is built from $\mathscr{L}_1$ by iterating the Lie bracket (which adds degrees) and the operation $[p]$ (which multiplies degrees by p).

If $p \nmid n$, then no monomial of total degree n can have the operation $[p]$ as its last step (indeed applying $[p]$ at the last step forces the total degree to be divisible by p). Hence every monomial of degree n uses a bracket at some stage, so it lies in the commutator ideal. This proves (a).

If $p \mid n$, then a monomial of degree n either has a bracket as its last step (hence lies in $[\mathscr{L}, \mathscr{L}]_n$), or it has the operation $[p]$ as its last step. In the latter case it equals $y^{[p]}$ for some homogeneous monomial y of degree n/p, hence belongs to the $\mathbb{F}_p$-span of $\{x^{[p]} : x \in \mathscr{L}_{n/p}\}$. This proves (b). □

Lemma 5.2.7. *Assume there exists N with $b_N(G) = 0$ (equivalently $L_N(G) = 0$). Then:*

(a) The ordinary graded Lie subalgebra $\mathscr{L}_* \subseteq \mathscr{L}_G$ generated by $L_1(G)$ is concentrated in degrees $< N$, hence finite-dimensional. In particular, $(\mathscr{L}_*)_n = 0$ for all $n \geq N$.

(b) The commutator ideal satisfies $[\mathscr{L}_G, \mathscr{L}_G]_n = 0$ for all $n \geq N$. Consequently, the graded ideal $\mathscr{L}_{\geq N} := \bigoplus_{n\geq N} L_n(G)$ is central (hence abelian).

(c) For every $n \geq N$ one has:

$$L_n(G) = 0 \;\; \text{if} \;\; p \nmid n, \qquad \text{and} \qquad L_n(G) = \left\langle x^{[p]} : x \in L_{n/p}(G) \right\rangle_{\mathbb{F}_p} \;\; \text{if} \;\; p \mid n.$$

In particular, for all sufficiently large n one has $b_n(G) = 0$ unless $n = mp^k$ with $1 \leq m \leq N-1$ and $p \nmid m$.

(d) For each fixed m with $1 \leq m \leq N-1$ and $p \nmid m$, the sequence $k \mapsto b_{mp^k}(G)$ is eventually non-increasing, hence bounded.

Proof. Write $L_n := L_n(G)$ and $\mathscr{L} := \mathscr{L}_G$.

(a) Let $\mathscr{L}_*$ be the ordinary graded Lie subalgebra generated by L_1. Since $\mathscr{L}_*$ is generated in degree 1, one has $(\mathscr{L}_*)_n = [L_1, (\mathscr{L}_*)_{n-1}]$ for all $n \geq 2$ (every Lie monomial of degree n can be written as a bracket of a degree-1 element with a degree-$(n-1)$ Lie monomial). Because $L_N = 0$, we have $(\mathscr{L}_*)_N \subseteq L_N = 0$, hence inductively $(\mathscr{L}_*)_n = 0$ for all $n \geq N$. Therefore $\mathscr{L}_*$ is concentrated in degrees $< N$ and is finite-dimensional.

(b) By Lemma 5.2.5 one has $[\mathscr{L}, \mathscr{L}] \subseteq \mathscr{L}_*$. Taking homogeneous degree-n parts and using (a) gives $[\mathscr{L}, \mathscr{L}]_n = 0$ for all $n \geq N$. Now if $x \in L_n$ with $n \geq N$ and $y \in L_m$ is arbitrary, then $[y, x] \in [\mathscr{L}, \mathscr{L}]_{m+n} = 0$ because $m + n \geq N$. Hence $\mathscr{L}_{\geq N}$ is central, in particular abelian.

(c) Apply Lemma 5.2.6 to the restricted Lie algebra $\mathscr{L}$ (generated by L_1 by Lemma 5.2.4). If $n \geq N$ and $p \nmid n$, then Lemma 5.2.6(a) gives $L_n \subseteq [\mathscr{L}, \mathscr{L}]_n = 0$, so $L_n = 0$.

If $n \geq N$ and $p \mid n$, then Lemma 5.2.6(b) gives

$$L_n \subseteq [\mathscr{L}, \mathscr{L}]_n + \left\langle x^{[p]} : x \in L_{n/p} \right\rangle_{\mathbb{F}_p} = \left\langle x^{[p]} : x \in L_{n/p} \right\rangle_{\mathbb{F}_p},$$

again since $[\mathscr{L}, \mathscr{L}]_n = 0$ for $n \geq N$. This proves the displayed dichotomy in (c). Iterating it shows: if $n \geq N$ and $n = mp^k$ with $p \nmid m$, then $L_n = 0$ whenever $m \geq N$, so for all sufficiently large n the only possible non-zero degrees lie on the finitely many rays mp^k with $1 \leq m \leq N-1$ and $p \nmid m$.

(d) Fix such an m and choose k_0 so that $mp^{k_0} \geq N$. For $k \geq k_0$ we have $L_{mp^k} \subseteq \mathscr{L}_{\geq N}$, which is abelian by (b). On an abelian restricted Lie algebra the operation $[p]$ is $\mathbb{F}_p$-additive (the Jacobson formula has only commutator correction terms, which vanish in the abelian case), and over $\mathbb{F}_p$ it is $\mathbb{F}_p$-linear. Thus the map

$$[p] : L_{mp^k} \longrightarrow L_{mp^{k+1}}, \qquad x \mapsto x^{[p]},$$

is $\mathbb{F}_p$-linear for all $k \geq k_0$. By (c) we also have $L_{mp^{k+1}} = \langle x^{[p]} : x \in L_{mp^k} \rangle_{\mathbb{F}_p} = \operatorname{im}([p])$ for all $k \geq k_0$. Hence $b_{mp^{k+1}} = \dim \operatorname{im}([p]) \leq \dim L_{mp^k} = b_{mp^k}$ for all $k \geq k_0$, so the sequence is eventually non-increasing and bounded. □

5.3 Main results

5.3.1 Preparation

Apart from Quillen's theorem and Jennings formula, we need two other external theorems. The first is a characterization of p-adic analytic pro-p groups due to Lazard [32].

Theorem 5.3.1. *Let $\widehat{G}$ be a finitely generated pro-p group. If $\log_p \left| \widehat{G} : \widehat{G}^{p^n} \right| = O(n)$, then $\widehat{G}$ is p-adic analytic and admits a faithful finite-dimensional linear representation over $\mathbb{Q}_p$.*

The second result is the Tits' alternative for linear groups [51]. We managed to avoid it so far, but now we are going to use it.

Theorem 5.3.2. *A finitely generated linear group over a field is either virtually solvable or contains a nonabelian free subgroup.*

5.3.2 Reduction to linear groups

With the subsequent preparations we are able to show that any finitely generated group satisfying the growth condition of Theorem 5.1.1 must be linear, more specifically, a subgroup of $\mathrm{GL}_m(\mathbb{Z}_p)$.

Theorem 5.3.3. *Let G be a residually p-finite group satisfying* (5.5). *Consider the following assertions:*

1) $a_n(G) \prec e^{\sqrt{n}}$.

2) There exists n such that $b_n(G) = 0$.

3) The indices $\rho_n := |G : G_n|$ have power growth.

4) The limit

$$r := \limsup_{n\to\infty} \frac{1}{n} \log_p \left|G : G^{p^n}\right|$$

is finite.

5) G admits a faithful representation by matrices over $\mathbb{Z}_p$.

Then, the following implications hold:

$$(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5).$$

Proof. (1) $\Rightarrow$ (2) : We prove that if $b_n(G) \geq 1$ for all n, then $a_n(G) \succeq e^{\sqrt{n}}$. This argument proceeds directly from the Jennings product formula. Indeed, assume $b_n \geq 1$ for all n. Then by (5.7),

$$f_G(t) \geq \prod_{n\geq 1} \frac{1-t^{pn}}{1-t^n} =: \frac{\Phi(t)}{\Phi(t^p)}, \qquad \Phi(t) := \prod_{n\geq 1}(1-t^n)^{-1}.$$

Write $\Phi(t) = \sum_{n\geq 0} p(n)t^n$ where $p(n)$ is the partition function. It is a classical result due to Hardy–Ramanujan (see Appendix A.3) that $\log p(n) \asymp \Theta(\sqrt{n})$. To extract a lower bound on coefficients, rewrite

$$\frac{\Phi(t)}{\Phi(t^p)} = \prod_{n\geq 1} \frac{1-t^{pn}}{1-t^n} = \prod_{n\geq 1}\left(1+t^n+\cdots+t^{(p-1)n}\right).$$

All factors have nonnegative coefficients and constant term 1, hence the inequality $f_G(t) \geq \Phi(t)/\Phi(t^p)$ holds coefficientwise, so

$$a_n(G) \;\geq\; [t^n]\frac{\Phi(t)}{\Phi(t^p)}.$$

Moreover, since $1+t^n+\cdots+t^{(p-1)n} \geq 1+t^n$ coefficientwise, we get

$$\frac{\Phi(t)}{\Phi(t^p)} \;\geq\; \prod_{n\geq 1}(1+t^n)$$

coefficientwise. The right-hand side is the generating function for partitions into *distinct* parts, whose coefficients satisfy $[t^n]\prod_{m\geq 1}(1+t^m) \;\succeq\; e^{\sqrt{n}}$, consequently $a_n(G) \;\succeq\; e^{\sqrt{n}}$.

(2) $\Rightarrow$ (3) : Assume $b_N(G) = 0$ for some N. By Lemma 5.2.7(c), for all sufficiently large indices $b_n(G)$ can be non-zero only on the finitely many p-rays mp^k with $1 \leq m \leq N-1$ and $p \nmid m$. By Lemma 5.2.7(d), along each such ray the values $b_{mp^k}(G)$ are eventually bounded by a constant. Let

$$B := \max_{1\leq j\leq N-1} b_j(G), \qquad M := \#\{m \in \{1,\ldots,N-1\} : p \nmid m \text{ and } b_m(G) \neq 0\}.$$

Then for every $n \geq 1$,

$$\sum_{i=1}^{n} b_i(G) \leq \sum_{i=1}^{N-1} b_i(G) \; + \sum_{\substack{1\leq m\leq N-1 \\ p\nmid m}} \; \sum_{k:\, mp^k \leq n} b_{mp^k}(G) \leq \sum_{i=1}^{N-1} b_i(G) \; + \; M \cdot B \cdot \big(1 + \lfloor \log_p n \rfloor\big).$$

Thus $\sum_{i=1}^{n} b_i(G) = O(\log n)$. Since

$$|G : G_{n+1}| = \prod_{i=1}^{n} |G_i/G_{i+1}| = p^{\sum_{i=1}^{n} b_i(G)},$$

it follows that $|G : G_{n+1}| \leq p^{C \log_p n} = n^C$ for some $C > 0$, i.e. $\rho_n = |G : G_n|$ has power growth.

(4) $\Rightarrow$ (5) : Let $\widehat{G}$ be the pro-p completion of G and note that G embeds into $\widehat{G}$, since G is residually p-finite. Since G is dense in $\widehat{G}$ and $\widehat{G}/\widehat{G}^{p^n}$ is a finite p-group, the natural map $G \to \widehat{G}/\widehat{G}^{p^n}$ is surjective, hence $\left|\widehat{G} : \widehat{G}^{p^n}\right| \leq \left|G : G^{p^n}\right|$. Thus assumption (4) implies $\log_p \left|\widehat{G} : \widehat{G}^{p^n}\right| = O(n)$. By Theorem 5.3.1, $\widehat{G}$ embeds into $\mathrm{GL}_m(\mathbb{Q}_p)$ and is p-adic analytic. Being compact, after conjugation it is contained in $\mathrm{GL}_m(\mathbb{Z}_p)$ (standard structure theory of compact subgroups of $\mathrm{GL}_m(\mathbb{Q}_p)$; cf. [48]). ☐

5.3.3 Proof of Theorem 5.1.1

Fix a finite generating set $S = \{s_1, \dots, s_m\}$ of G. Recall the growth function $\gamma_{G,S}$ from the chapter on Cayley graphs and growth (see Definition 3.1.5 and Proposition 3.1.4). Denote the associated word length by $|g|_S$. Thus $\gamma_{G,S}(n) = \#\{g \in G : |g|_S \leq n\}$.

Lemma 5.3.4. *For any generating set S of G,*

$$a_n(G) \leq \gamma_{G,S}(n) \qquad (n \geq 1).$$

Proof. Every element of $\Delta(G)^n$ is an $\mathbb{F}_p$-linear combination of products

$$(s_1 - 1) \cdots (s_n - 1) \qquad (s_i \in S \cup S^{-1}).$$

Expanding out each factor shows that such a product is a linear combination of group elements of the form $s_{i_1} \cdots s_{i_k}$ with $k \leq n$, hence supported inside the ball $\{g \in G : |g|_S \leq n\}$. Therefore $\Delta(G)^n$ is contained in the $\mathbb{F}_p$-span of $\{g \in G : |g|_S \leq n\}$, so

$$\dim_{\mathbb{F}_p} \Delta(G)^n \leq \gamma_{G,S}(n).$$

Finally,

$$a_n(G) = \dim_{\mathbb{F}_p}\big(\Delta(G)^n/\Delta(G)^{n+1}\big) \leq \dim_{\mathbb{F}_p} \Delta(G)^n \leq \gamma_{G,S}(n).$$

☐

Proof of Theorem 5.1.1. Since G is finitely generated, condition (5.5) holds. By Lemma 5.3.4, we have $a_n(G) \leq \gamma_{G,S}(n)$ for all n. Since $\gamma_G(n) \prec e^{\sqrt{n}}$ by hypothesis, and growth type is independent of the generating set by Proposition 3.1.4, it follows that $\gamma_{G,S}(n) \prec e^{\sqrt{n}}$ and hence $a_n(G) \prec e^{\sqrt{n}}$. By Theorem 5.3.3, condition (1) implies condition (5), so G admits a faithful representation by matrices over $\mathbb{Z}_p$.

The growth bound excludes the existence of non-abelian free subgroups and thus, by Theorem 5.3.2, implies that G is virtually solvable. Finally, by Theorem 4.2.8, a finitely generated virtually solvable group of subexponential growth is virtually nilpotent. ☐

Part II

Analytic methods

Chapter 6

Random walks and entropy

In this chapter we study random walks on groups and their connections to entropy and harmonic functions.

6.1 Random walks and the Poisson boundary

6.1.1 Definitions

Definition 6.1.1. Let X be a countable set. We write $\mathcal{P}(X)$ for the set of *probability measures* on X, i.e. functions $p : X \to [0,1]$ such that $\sum_{x\in X} p(x) = 1$. The *support* of p is $\operatorname{supp}(p) := \{x \in X : p(x) > 0\}$.

Definition 6.1.2. Let G be a countable group.

(i) For $\mu, \nu \in \mathcal{P}(G)$ their *convolution* is the probability measure $\mu * \nu \in \mathcal{P}(G)$ defined by

$$(\mu * \nu)(x) := \sum_{y\in G} \mu(y)\, \nu(y^{-1}x) \qquad (x \in G).$$

Inductively, μ^{*n} denotes the n-fold convolution with $\mu^{*1} = \mu$.

(ii) For $g \in G$ and $\nu \in \mathcal{P}(G)$ the *left translate* $g\nu \in \mathcal{P}(G)$ is defined by

$$(g\nu)(x) := \nu(g^{-1}x) \qquad (x \in G).$$

(iii) A probability measure $\mu \in \mathcal{P}(G)$ is called *symmetric* if $\mu(g) = \mu(g^{-1})$ for all $g \in G$. It is called *nondegenerate* if $\operatorname{supp}(\mu)$ generates G as a semigroup.

Fix a countable group G and a probability measure $\mu \in \mathcal{P}(G)$ whose support generates G as a semigroup. The μ-random walk on G is the random process

$$X_0 = e, \qquad X_{n+1} = X_n\, S_{n+1},$$

where $(S_1, S_2, \dots)$ are i.i.d. G-valued random variables with law μ.

Proposition 6.1.3. *Let G be a countable group and let $\mu \in \mathcal{P}(G)$. Let $(S_n)_{n\geq 1}$ be i.i.d. G-valued random variables with law μ and define $X_0 := e$ and $X_n := S_1 \cdots S_n$. Then the distribution of X_n is μ^{*n}.*

Proof. We prove by induction on n. For $n = 1$ the law of S_1 is μ. Assume the claim for n. For any $x \in G$ we have, using independence of S_{n+1} and X_n,

$$\mathbb{P}[X_{n+1} = x] = \sum_{y \in G} \mathbb{P}[X_n = y]\, \mathbb{P}[S_{n+1} = y^{-1}x] = (\mu^{*n} * \mu)(x) = \mu^{*(n+1)}(x).$$

□

The averaging operator behind the walk is the so-called *Markov operator* P acting on bounded functions $f : G \to \mathbb{C}$ by

$$Pf(x) := \sum_{s \in G} \mu(s)\, f(xs) \qquad (x \in G).$$

Then, we have

$$(P^n f)(e) = \sum_{x \in G} \mu^{*n}(x)\, f(x) = \mathbb{E}[f(X_n)].$$

6.1.2 Harmonic functions and the Poisson boundary

A function $f \in \ell^\infty(G)$ is called *μ-harmonic* if $Pf = f$. This is exactly the condition that $(f(X_n))_{n \geq 0}$ is a bounded martingale with respect to the natural filtration of the random walk:

$$\mathbb{E}\big[f(X_{n+1}) \mid X_0, \ldots, X_n\big] = Pf(X_n) = f(X_n).$$

By the martingale convergence theorem, $f(X_n)$ converges almost surely to a limit random variable f_∞. In this sense, bounded μ-harmonic functions are precisely those whose values along the random walk stabilize *at infinity*.

We now explain a functional-analytic construction of the Poisson boundary due to Prunaru, [45]. Unlike the probabilistic approach via path space, the construction proceeds by taking suitable limits of iterates of the Markov operator and then applying Gelfand theory.

For $g \in G$ and $f \in \ell^\infty(G)$ we write

$$(L_g f)(x) := f(g^{-1}x) \qquad (x \in G),$$

so that $L : G \to \mathrm{Isom}(\ell^\infty(G))$ is the left-translation action. Since the Markov operator P acts on the right, it satisfies the intertwining relation

$$L_g \circ P = P \circ L_g \qquad (g \in G). \tag{6.1}$$

Lemma 6.1.4. *The Markov operator $P : \ell^\infty(G) \to \ell^\infty(G)$ is unital, positive, and contractive. Moreover, for every $f \in \ell^\infty(G)$ one has the pointwise Schwarz inequality*

$$|Pf|^2 \leq P(|f|^2).$$

Proof. Unitality and positivity are immediate from the definition. Contractivity follows from

$$|Pf(x)| \leq \sum_s \mu(s)\, |f(xs)| \leq \|f\|_\infty \sum_s \mu(s) = \|f\|_\infty.$$

For the Schwarz inequality, fix $x \in G$. By Cauchy–Schwarz in $\ell^2(G)$,

$$|Pf(x)|^2 = \left| \sum_s \mu(s) f(xs) \right|^2 \leq \left(\sum_s \mu(s) \right) \left(\sum_s \mu(s) |f(xs)|^2 \right) = P(|f|^2)(x).$$

□

Denote the space of bounded μ-harmonic functions by

$$\mathcal{H}(G,\mu) := \{f \in \ell^\infty(G) : Pf = f\}.$$

Definition 6.1.5. Let $\mathcal{C}(P)$ be the set of all $f \in \ell^\infty(G)$ such that, for every $x \in G$, the sequence $(P^n f(x))_{n\geq 0}$ converges. For $f \in \mathcal{C}(P)$ define $\pi(f) \in \ell^\infty(G)$ by

$$\pi(f)(x) := \lim_{n\to\infty} P^n f(x) \qquad (x \in G).$$

Lemma 6.1.6. *The map $\pi : \mathcal{C}(P) \to \ell^\infty(G)$ is linear, unital, positive, G-equivariant, contractive and satisfies $P \circ \pi = \pi$ and $\pi \circ P = \pi$. In particular, $\pi(f) \in \mathcal{H}(G,\mu)$ for every $f \in \mathcal{C}(P)$.*

Proof. Linearity and unitality are clear from the definition. Positivity and contractivity follow since each P^n is positive and contractive, and pointwise limits preserve order and sup bounds. The intertwining relation (6.1) implies $L_g(P^n f) = P^n(L_g f)$ for all n. Taking pointwise limits yields $\pi(L_g f) = L_g \pi(f)$. Finally, from $P^{n+1}f = P^n(Pf)$ we get $\pi(Pf) = \pi(f)$. Continuity of P with respect to pointwise convergence on bounded sets gives $P\pi(f) = \pi(f)$. □

The key point is that even though $P^n f$ need not converge for an arbitrary bounded f, it does converge on the C^*-algebra generated by harmonic functions.

Definition 6.1.7. Let $A \subseteq \ell^\infty(G)$ be the (norm) closed $*$-subalgebra generated by $\mathcal{H}(G,\mu)$.

Lemma 6.1.8. *One has $A \subseteq \mathcal{C}(P)$. In particular, π restricts to a unital positive contraction $\pi : A \to \mathcal{H}(G,\mu)$.*

Proof. Let $\mathcal{C}(P)_0 := \{f \in \mathcal{C}(P) : \pi(f) = 0\}$. If $f \in \mathcal{C}(P)_0$ is nonnegative and $g \in \ell^\infty(G)$, then $fg \in \mathcal{C}(P)_0$. Indeed, it suffices to treat $g \geq 0$. For every n and every $x \in G$,

$$(P^n(fg))(x) \leq \|g\|_\infty\,(P^n f)(x),$$

so $P^n(fg)(x) \to 0$ pointwise.

If $h \in \mathcal{H}(G,\mu)$, then $|h|^2 \in \mathcal{C}(P)$. By Lemma 6.1.4, $P(|h|^2) \geq |h|^2$. Iterating yields a pointwise increasing sequence $P^n(|h|^2)$ bounded above by $\|h\|_\infty^2$, hence it converges pointwise. Moreover, the limit $\pi(|h|^2)$ is μ-harmonic by Lemma 6.1.6, so

$$\pi(|h|^2) - |h|^2 \in \mathcal{C}(P)_0 \quad \text{and is nonnegative.}$$

We know that

$$\big(\pi(|h|^2) - |h|^2\big) g \in \mathcal{C}(P)_0 \qquad (g \in \ell^\infty(G)). \tag{6.2}$$

If $h_1, h_2 \in \mathcal{H}(G,\mu)$, then $h_1h_2 \in \mathcal{C}(P)$. Indeed, by the polarization identity one can write

$$h_1h_2 = \frac{1}{4}\sum_{m=0}^{3} i^m\,|h_1 + i^m\overline{h_2}|^2,$$

and each $h_1 + i^m\overline{h_2}$ is harmonic. Hence h_1h_2 is a linear combination of functions of the form $|h|^2$ handled above. Moreover, applying (6.2) to each term yields

$$\big(\pi(h_1h_2) - h_1h_2\big) g \in \mathcal{C}(P)_0 \qquad (g \in \ell^\infty(G)). \tag{6.3}$$

Let $J \lhd A$ be the closed ideal (in A) generated by all elements of the form

$$h_1h_2 - \pi(h_1h_2) \qquad (h_1, h_2 \in \mathcal{H}(G,\mu)).$$

We claim that if $h_1, \ldots, h_k \in \mathcal{H}(G,\mu)$ and $g := h_1 \cdots h_k$, then $g \in \mathcal{C}(P)$ and $g - \pi(g) \in J$. For $k = 1$ this is clear. For $k = 2$, we have shown $g \in \mathcal{C}(P)$ and (6.3) with $g = 1$ shows $g - \pi(g) \in J$. Assume the claim holds up to $k - 1$. Write

$$h_1 \cdots h_k = \big(h_1h_2 - \pi(h_1h_2)\big)h_3 \cdots h_k \ + \ \pi(h_1h_2)h_3 \cdots h_k.$$

The first term lies in J by definition, and it also lies in $\mathcal{C}(P)$ by (6.3). The second term is a product of $k-1$ harmonic functions, hence lies in $\mathcal{C}(P)$ by induction, and differs from its π-image by an element of J. It follows that $h_1 \cdots h_k \in \mathcal{C}(P)$ and $h_1 \cdots h_k - \pi(h_1 \cdots h_k) \in J$.

Now the $*$-algebra generated by $\mathcal{H}(G,\mu)$ consists of finite linear combinations of such products, so it is contained in $\mathcal{C}(P)$. Taking the norm closure yields $A \subseteq \mathcal{C}(P)$. Moreover, for every a in the algebraic $*$-subalgebra generated by $\mathcal{H}(G,\mu)$ we have $a - \pi(a) \in J$. Since J is norm closed and $\pi : A \to \mathcal{H}(G,\mu)$ is contractive, the same holds for all $a \in A$ by approximation. □

Lemma 6.1.9. *With notation as above,* $\ker(\pi|_A) = J$. *In particular,* $\ker(\pi|_A)$ *is a closed ideal of* A.

Proof. By definition, $J \subseteq A$ and J is generated by elements of the form $h_1h_2 - \pi(h_1h_2)$. Each such generator lies in $\ker(\pi|_A)$ since π is idempotent on A. Hence $J \subseteq \ker(\pi|_A)$.

Conversely, if $f \in \ker(\pi|_A)$, then by the last sentence of the proof of Lemma 6.1.8 we have $f - \pi(f) = f \in J$. □

Definition 6.1.10. Let $\Pi_\mu(G)$ be the maximal ideal space of the commutative C^*-algebra A/J. We call $\Pi_\mu(G)$ the *Poisson space* of (G,μ).

The Gelfand transform yields an isometric $*$-isomorphism $A/J \cong C(\Pi_\mu(G))$. Composing its inverse with π produces a canonical identification between $C(\Pi_\mu(G))$ and the harmonic functions. Thus, there exists a compact G-space $\Pi_\mu(G)$ and a G-equivariant unital linear isometric isomorphism

$$\Gamma : C(\Pi_\mu(G)) \longrightarrow \mathcal{H}(G,\mu) \subseteq \ell^\infty(G).$$

Moreover, the product in $C(\Pi_\mu(G))$ can be computed in $\ell^\infty(G)$ by the formula

$$\Gamma(\varphi\psi)(g) = \lim_{n\to\infty} P^n(\Gamma(\varphi)\Gamma(\psi))(g) \qquad (\varphi, \psi \in C(\Pi_\mu(G)),\ g \in G). \tag{6.4}$$

Remark 6.1.11. It is remarkable and not obvious from (6.4) that it can be used to define an associative product on $\mathcal{H}(G,\mu)$.

The following corollary is the analogue of the Poisson integral representation of harmonic functions on the unit disk, and is the reason for the name *Poisson boundary.*

Corollary 6.1.12. *Let* ν *be the Borel probability measure on* $\Pi_\mu(G)$ *corresponding to the state*

$$\varphi \mapsto (\Gamma\varphi)(e) \qquad (\varphi \in C(\Pi_\mu(G))).$$

Then for every $\varphi \in C(\Pi_\mu(G))$ *and every* $g \in G$ *one has*

$$(\Gamma\varphi)(g) = \int_{\Pi_\mu(G)} \varphi(g \cdot x)\, d\nu(x).$$

Proof. Since Γ is G-equivariant, we have $(\Gamma\varphi)(g) = (L_{g^{-1}}\Gamma\varphi)(e) = (\Gamma(L_{g^{-1}}\varphi))(e)$. By definition of ν, this equals $\int(L_{g^{-1}}\varphi)\,d\nu = \int \varphi(g\cdot x)\,d\nu(x)$. □

Example 6.1.13. For the simple random walk on $\mathbb{Z}$, a function $f : \mathbb{Z} \to \mathbb{C}$ is harmonic if and only if it satisfies the mean value property

$$f(n) = \tfrac{1}{2}f(n-1) + \tfrac{1}{2}f(n+1) \qquad (\forall n \in \mathbb{Z}).$$

Hence

$$f(n+1) - f(n) = f(n) - f(n-1),$$

so the first differences are constant. Hence $f(n) = an + b$ for constants a, b. If f is bounded on $\mathbb{Z}$, then $a = 0$, so f is constant.

Example 6.1.14. Let $\mathcal{F}_2 = \langle a, b\rangle$ with generating set $S = \{a^{\pm 1}, b^{\pm 1}\}$ and let μ be the simple symmetric random walk measure Define $f : \mathcal{F}_2 \to [0,1]$ by

$$f(g) := \begin{cases} \frac{1}{4} & g = \mathbb{1} \\ 1 - \frac{3}{4}\cdot 3^{-|g|}, & \text{if the reduced word of } g \text{ begins with } a, \\ \frac{1}{4}\cdot 3^{-|g|}, & \text{otherwise.} \end{cases}$$

Then f is bounded and not constant.

We claim that f is μ-harmonic. Indeed, for every $g \neq e$ there is exactly one neighbour g^- with $|g^-| = |g| - 1$ and three neighbours g^+ with $|g^+| = |g| + 1$. Moreover, when one moves away from the origin, one stays either always in the a-branch of the tree or always outside it. Hence the harmonicity condition reduces to

$$f(g) = \frac{1}{4}f(g^-) + \frac{3}{4}f(g^+) \qquad (g \neq e).$$

On a fixed branch, if f_n denotes the value of f on words of length n, this is the recursion

$$4f_n = f_{n-1} + 3f_{n+1}.$$

The functions $n \mapsto 1 - \frac{3}{4}\cdot 3^{-n}$ and $n \mapsto \frac{1}{4}\cdot 3^{-n}$ satisfy this recursion. At the origin,

$$f(e) = \frac{1}{4}\big(f(a) + f(a^{-1}) + f(b) + f(b^{-1})\big) = \frac{1}{4}\Big(\frac{3}{4} + 3\cdot\frac{1}{12}\Big) = \frac{1}{4},$$

so f is μ-harmonic everywhere.

6.2 Entropy of random walks

An important numerical invariant of a random walk is its asymptotic entropy, which is a measure of the behaviour of typical trajectories of the walk. It is closely related to the Poisson boundary and harmonic functions, as we will see in the next chapter.

6.2.1 Definitions

For a probability measure p on a countable set, its entropy is

$$H(p) := -\sum_x p(x)\log p(x) = \sum_x p(x)\log\frac{1}{p(x)}$$

(with the convention $0\log 0 := 0$). For the μ-random walk, set $H_n := H(\mu^{*n})$.

Lemma 6.2.1. *Let G be a countable group and let $p, q \in \mathcal{P}(G)$. Assume that $H(p) < \infty$ and $H(q) < \infty$. Then $H(p*q) \leq H(p) + H(q)$.*

Proof. We give a direct proof using convexity of $t \mapsto t\log t$. Define weights on $G \times G$ by $a_{x,y} := p(x)\,q(y)$ for $x, y \in G$. For $z \in G$, group these weights according to the product $z = xy$ and set

$$p_z := (p*q)(z) = \sum_{xy=z} a_{x,y}.$$

For each fixed z, the convexity gives

$$\sum_{xy=z} a_{x,y}\log a_{x,y} \leq \Big(\sum_{xy=z} a_{x,y}\Big)\log\Big(\sum_{xy=z} a_{x,y}\Big).$$

Summing over z yields

$$\sum_{x,y\in G} a_{x,y}\log a_{x,y} \leq \sum_{z\in G} p_z\log p_z.$$

Multiplying by -1 and using $a_{x,y} = p(x)q(y)$, we obtain

$$\begin{aligned} H(p*q) &= -\sum_{z\in G} p_z\log p_z \leq -\sum_{x,y\in G} p(x)q(y)\log\big(p(x)q(y)\big) \\ &= -\sum_{x\in G} p(x)\log p(x)\sum_{y\in G} q(y) - \sum_{y\in G} q(y)\log q(y)\sum_{x\in G} p(x) \\ &= H(p) + H(q). \end{aligned}$$

□

Corollary 6.2.2. *Assume that $H(\mu) < \infty$. Then $H_{m+n} \leq H_m + H_n$ for all $m, n \geq 1$.*

Proof. Apply Lemma 6.2.1 to $p = \mu^{*m}$ and $q = \mu^{*n}$. □

Proposition 6.2.3. *Assume that $H(\mu) < \infty$. Then, we have*

$$\lim_{n\to\infty}\frac{1}{n}H(\mu^{*n}) = \inf_{n\geq 1}\frac{1}{n}H(\mu^{*n}).$$

Proof. By Corollary 6.2.2, the sequence $(H_n)_{n\geq 1}$ is subadditive. By Fekete's lemma for subadditive sequences, the limit of H_n/n exists and equals $\inf_{n\geq 1} H_n/n$. □

Definition 6.2.4. We call

$$h(\mu) := \lim_{n\to\infty}\frac{1}{n}H(\mu^{*n})$$

the *asymptotic entropy* of (G, μ).

6.2.2 The free group on two generators

The aim of this section is to compute the asymptotic entropy of the simple random walk on the free group on two generators in complete detail. Let $\mathcal{F}_2 = \langle a, b\rangle$ be the free group on two generators, with symmetric generating set $S = \{a, a^{-1}, b, b^{-1}\}$. Let $\mu \in \mathcal{P}(\mathcal{F}_2)$ be the simple random walk measure, i.e. $\mu(s) = \frac{1}{4}$ for $s \in S$ and $\mu(g) = 0$ otherwise. Let $X_0 = \mathbb{1}$ and $X_{n+1} = X_n S_{n+1}$ where $(S_n)_{n\geq 1}$ are i.i.d. with law μ. Let $|g|$ denote word length of $g \in \mathcal{F}_2$ with respect to S and set $D_n := |X_n|$.

The size of th spheres in $\mathcal{F}_2$ of radius d equals $s_d = 4 \cdot 3^{d-1}$. Moreover, for each $n \geq 0$ and each $g \in \mathcal{F}_2$, the value $\mu^{*n}(g)$ depends only on $d = |g|$. Equivalently, if one defines $p_{n,d} := \mathbb{P}(D_n = d)$, then for every g with $|g| = d$,

$$\mu^{*n}(g) = \frac{p_{n,d}}{s_d}.$$

Lemma 6.2.5. *The process $(D_n)_{n\geq 0}$ is a Markov chain on $\mathbb{N}$ with transitions*

$$\mathbb{P}(D_{n+1} = 1 \mid D_n = 0) = 1, \quad \mathbb{P}(D_{n+1} = d{+}1 \mid D_n = d) = \frac{3}{4}, \quad \mathbb{P}(D_{n+1} = d{-}1 \mid D_n = d) = \frac{1}{4}$$

for $d \geq 1$. Consequently, with $p_{0,0} = 1$ and $p_{0,d} = 0$ for $d \geq 1$,

$$\begin{aligned} p_{n+1,0} &= \frac{1}{4}\, p_{n,1}, \\ p_{n+1,1} &= p_{n,0} + \frac{1}{4}\, p_{n,2}, \\ p_{n+1,d} &= \frac{3}{4}\, p_{n,d-1} + \frac{1}{4}\, p_{n,d+1} \qquad (d \geq 2). \end{aligned}$$

Proof. If $D_n = 0$, then $X_n = e$ and the next step necessarily moves to distance 1. If $D_n = d \geq 1$, then the reduced word for X_n has a last letter; among the 4 generators, exactly one is its inverse and shortens the reduced word (probability 1/4), while the other three extend it (probability 3/4). The recursion is obtained by conditioning on D_n. □

We are now going to compute the entropy $H_n = H(\mu^{*n})$ in terms of the distribution of D_n. Define the entropy of the radial law by

$$H(p_{n,\bullet}) := -\sum_{d\geq 0} p_{n,d} \log p_{n,d}.$$

Proposition 6.2.6. *For every $n \geq 0$,*

$$H(\mu^{*n}) = H(p_{n,\bullet}) + \sum_{d\geq 0} p_{n,d} \log s_d.$$

In particular, using $s_0 = 1$ and $s_d = 4 \cdot 3^{d-1}$ for $d \geq 1$,

$$H(\mu^{*n}) = H(p_{n,\bullet}) + \log 3\, \mathbb{E}[D_n] + (\log 4 - \log 3)\,(1 - p_{n,0}).$$

Proof. For $|g| = d$ one has $\mu^{*n}(g) = p_{n,d}/s_d$. Hence

$$\begin{aligned} H(\mu^{*n}) &= -\sum_{g\in\mathcal{F}_2} \mu^{*n}(g)\log\mu^{*n}(g) \\ &= -\sum_{d\geq 0}\sum_{|g|=d} \frac{p_{n,d}}{s_d}\log\Big(\frac{p_{n,d}}{s_d}\Big) \\ &= -\sum_{d\geq 0} p_{n,d}\big(\log p_{n,d} - \log s_d\big) \\ &= -\sum_{d\geq 0} p_{n,d}\log p_{n,d} + \sum_{d\geq 0} p_{n,d}\log s_d \\ &= H(p_{n,\bullet}) + \sum_{d\geq 0} p_{n,d}\log s_d. \end{aligned}$$

For $d \geq 1$,

$$\log s_d = \log(4\cdot 3^{d-1}) = \log 4 + (d-1)\log 3 = d\log 3 + (\log 4 - \log 3),$$

and $\log s_0 = 0$. Therefore

$$\sum_{d\geq 0} p_{n,d}\log s_d = \sum_{d\geq 1} p_{n,d}\big(d\log 3 + (\log 4 - \log 3)\big) = \log 3\,\mathbb{E}[D_n] + (\log 4 - \log 3)(1 - p_{n,0}).$$

□

It remains to understand the growth of $\mathbb{E}[D_n]$, the so-called drift of the random walk, and $p_{n,0}$ as $n \to \infty$.

Lemma 6.2.7. *For every $n \geq 0$,*

$$\mathbb{E}[D_{n+1}] = \mathbb{E}[D_n] + \frac{1}{2} + \frac{1}{2}\,p_{n,0}.$$

Equivalently,

$$\mathbb{E}[D_n] = \frac{n}{2} + \frac{1}{2}\sum_{k=0}^{n-1} p_{k,0}.$$

Proof. If $D_n = 0$ then $D_{n+1} = 1$, hence $\mathbb{E}[D_{n+1} - D_n \mid D_n = 0] = 1$. If $D_n = d \geq 1$, then $D_{n+1} = d + 1$ with probability $3/4$ and $D_{n+1} = d - 1$ with probability $1/4$, so

$$\mathbb{E}[D_{n+1} - D_n \mid D_n = d] = \frac{3}{4}\cdot 1 + \frac{1}{4}\cdot(-1) = \frac{1}{2}.$$

Taking expectations yields

$$\mathbb{E}[D_{n+1} - D_n] = 1\cdot p_{n,0} + \frac{1}{2}\cdot(1 - p_{n,0}) = \frac{1}{2} + \frac{1}{2}\,p_{n,0}.$$

Summing from 0 to $n - 1$ gives the second identity. □

Lemma 6.2.8. *We have $\mathbb{P}(\text{the chain } (D_n) \text{ ever hits } 0 \mid D_0 = 1) = \frac{1}{3}$.*

Proof. Let $h(d) := \mathbb{P}(\text{hit 0 eventually} \mid D_0 = d)$. Then $h(0) = 1$ and for $d \geq 1$,

$$h(d) = \frac{1}{4}\,h(d-1) + \frac{3}{4}\,h(d+1).$$

Seeking solutions of the form $h(d) = \lambda^d$ gives

$$\lambda = \frac{1}{4} + \frac{3}{4}\,\lambda^2 \quad\iff\quad 3\lambda^2 - 4\lambda + 1 = 0 \quad\iff\quad (\lambda - 1)(3\lambda - 1) = 0.$$

Thus $\lambda \in \{1, \frac{1}{3}\}$. The condition $h(d) \to 0$ as $d \to \infty$ forces $h(d) = (\frac{1}{3})^d$. Hence $h(1) = \frac{1}{3}$. □

Lemma 6.2.9. *Let $N := \#\{n \geq 0 : \ D_n = 0\}$. Then $\mathbb{E}[N] = \frac{3}{2}$, and consequently*

$$\sum_{n=0}^{\infty} p_{n,0} = \frac{3}{2}.$$

Proof. Starting from $D_0 = 0$, there is one visit at time 0. After each visit to 0, the chain moves to 1 at the next step. From distance 1, the probability to ever return to 0 equals $r = \frac{1}{3}$ by Lemma 6.2.8. By the strong Markov property, after each return the process renews. Hence the number of additional visits to 0 after time 0 is geometric with success parameter $1-r$, so its expectation equals $r/(1-r) = 1/2$. Therefore $\mathbb{E}[N] = 1+1/2 = 3/2$.

Also, note that

$$N = \sum_{n=0}^{\infty} \mathbf{1}_{\{D_n=0\}}, \qquad \text{so} \qquad \mathbb{E}[N] = \sum_{n=0}^{\infty} \mathbb{P}(D_n = 0) = \sum_{n=0}^{\infty} p_{n,0}.$$

□

Theorem 6.2.10. *For the simple random walk on $\mathcal{F}_2$,*

$$\lim_{n\to\infty} \frac{1}{n} H(\mu^{*n}) = \frac{1}{2}\log 3.$$

More precisely,

$$H(\mu^{*n}) = \frac{1}{2}(\log 3)\, n + O(\log n).$$

Proof. First, since $D_n \in \{0, 1, \ldots, n\}$, the support of $(p_{n,d})_{d\geq 0}$ has size at most $n+1$, hence

$$0 \leq H(p_{n,\bullet}) \leq \log(n+1).$$

Next, by Lemma 6.2.7 and Lemma 6.2.9,

$$\mathbb{E}[D_n] = \frac{n}{2} + \frac{1}{2}\sum_{k=0}^{n-1} p_{k,0} = \frac{n}{2} + \frac{1}{2}\Big(\sum_{k=0}^{\infty} p_{k,0} - \sum_{k=n}^{\infty} p_{k,0}\Big) = \frac{n}{2} + \frac{3}{4} + o(1).$$

Insert this into Proposition 6.2.6:

$$H(\mu^{*n}) = H(p_{n,\bullet}) + \log 3\,\mathbb{E}[D_n] + (\log 4 - \log 3)\,(1 - p_{n,0}).$$

The last term is $O(1)$ because $0 \leq 1 - p_{n,0} \leq 1$, and $H(p_{n,\bullet}) = O(\log n)$. Thus

$$H(\mu^{*n}) = \log 3\left(\frac{n}{2} + \frac{3}{4} + o(1)\right) + O(\log n) = \frac{1}{2}(\log 3)\, n + O(\log n),$$

and dividing by n gives the stated limit. □

6.2.3 Concavity modulus

Following [40, §3], for probability measures p, q on a countable set X, we define the *concavity defect*

$$\delta(p,q) := H\left(\frac{p+q}{2}\right) - \frac{H(p)+H(q)}{2} \;\geq\; 0,$$

where the inequality follows from concavity of entropy.

Lemma 6.2.11. *Let X be a countable set, let p, q be probability measures on X, and let $f : X \to \mathbb{R}_{\geq 0}$. Then*

$$\sum_x f(x)\,|p(x)-q(x)| \;\leq\; \Big(8\delta(p,q)\Big)^{1/2} \cdot \left(\sum_x f(x)^2\,(p(x)+q(x))\right)^{1/2}.$$

Proof. We follow [40, §3], with some minor modifications. Set $m := \frac{p+q}{2}$. For each $x \in X$ with $p(x)+q(x) > 0$ define

$$t(x) := \frac{p(x)-q(x)}{p(x)+q(x)} \in [-1,1].$$

Then

$$\begin{aligned}\delta(p,q) &= \frac{1}{2}\sum_{x\in X}\Big(p(x)\log\frac{p(x)}{m(x)} + q(x)\log\frac{q(x)}{m(x)}\Big)\\ &= \frac{1}{2}\sum_{x\in X}\Big(p(x)\log\frac{2p(x)}{p(x)+q(x)} + q(x)\log\frac{2q(x)}{p(x)+q(x)}\Big),\end{aligned}$$

with the convention that terms with $p(x)=q(x)=0$ are 0. Writing $p(x) = \frac{p(x)+q(x)}{2}\big(1+t(x)\big)$ and $q(x) = \frac{p(x)+q(x)}{2}\big(1-t(x)\big)$, the summand becomes

$$\frac{p(x)+q(x)}{4}\Big(\big(1+t(x)\big)\log\big(1+t(x)\big) + \big(1-t(x)\big)\log\big(1-t(x)\big)\Big).$$

Consider the function

$$\varphi(t) := (1+t)\log(1+t) + (1-t)\log(1-t) - \frac{t^2}{2}, \qquad t \in (-1,1).$$

A direct computation gives

$$\varphi''(t) = \frac{1}{1+t} + \frac{1}{1-t} - 1 = \frac{1+t^2}{1-t^2} \;\geq\; 0,$$

and also $\varphi(0) = 0$, $\varphi'(0) = 0$. Hence $\varphi(t) \geq 0$ for all $t \in (-1,1)$, i.e.

$$(1+t)\log(1+t) + (1-t)\log(1-t) \;\geq\; \frac{t^2}{2}.$$

By continuity this inequality also holds at $t = \pm 1$. Substituting back yields the pointwise estimate (for all x with $p(x)+q(x) > 0$)

$$\frac{(p(x)-q(x))^2}{p(x)+q(x)} \;\leq\; 4\left(p(x)\log\frac{2p(x)}{p(x)+q(x)} + q(x)\log\frac{2q(x)}{p(x)+q(x)}\right).$$

Summing over x gives

$$\sum_{x\in X} \frac{(p(x)-q(x))^2}{p(x)+q(x)} \;\le\; 8\,\delta(p,q).$$

Now apply Cauchy–Schwarz:

$$\begin{aligned}
\sum_x f(x)|p(x)-q(x)| &= \sum_x \Big(f(x)\sqrt{p(x)+q(x)}\Big)\cdot \frac{|p(x)-q(x)|}{\sqrt{p(x)+q(x)}} \\
&\le \left(\sum_x f(x)^2\,(p(x)+q(x))\right)^{1/2} \left(\sum_x \frac{(p(x)-q(x))^2}{p(x)+q(x)}\right)^{1/2} \\
&\le \Big(8\delta(p,q)\Big)^{1/2}\cdot\left(\sum_x f(x)^2\,(p(x)+q(x))\right)^{1/2}.
\end{aligned}$$

□

Lemma 6.2.12. *For any probability measures μ,ν on a group G and any $g\in G$,*

$$H(\mu*\nu)-H(\nu) \;\ge\; 2\min\{\mu(e),\mu(g)\}\,\delta(\nu,g\nu).$$

Proof. For $h\in G$ we view $h\nu$ as a translate of ν in the sense of the preceding subsection. Since left translation is a bijection of G, we have $H(h\nu)=H(\nu)$ for all h.

Set $\alpha:=\min\{\mu(e),\mu(g)\}$. Define a subprobability measure $\widetilde{\mu}:=\mu-\alpha\,\delta_e-\alpha\,\delta_g$. Then $\widetilde{\mu}\ge 0$ and $\|\widetilde{\mu}\|_1=1-2\alpha$. If $\alpha<\frac{1}{2}$, set $\mu':=\widetilde{\mu}/(1-2\alpha)\in\mathcal{P}(G)$; if $\alpha=\frac{1}{2}$ we may choose any $\mu'\in\mathcal{P}(G)$. Convolving with ν gives the convex decomposition

$$\mu*\nu=\alpha\,\nu+\alpha\,(g\nu)+(1-2\alpha)\,(\mu'*\nu).$$

By concavity of H,

$$\begin{aligned}
H(\mu*\nu) &\ge 2\alpha\,H\left(\frac{\nu+g\nu}{2}\right)+(1-2\alpha)\,H(\mu'*\nu) \\
&\ge 2\alpha\,H\left(\frac{\nu+g\nu}{2}\right)+(1-2\alpha)\,H(\nu),
\end{aligned}$$

where the second inequality uses again concavity of H applied to the convex combination $\mu'*\nu=\sum_{h\in G}\mu'(h)\,(h\nu)$ and the fact that all $H(h\nu)$ equal $H(\nu)$. Subtracting $H(\nu)$ and using $H(g\nu)=H(\nu)$ yields

$$H(\mu*\nu)-H(\nu)\ge 2\alpha\left(H\left(\frac{\nu+g\nu}{2}\right)-\frac{H(\nu)+H(g\nu)}{2}\right)=2\alpha\,\delta(\nu,g\nu),$$

as claimed. □

6.3 The Liouville property

Definition 6.3.1. We say that (G,μ) has the *Liouville property* if every bounded μ-harmonic function is constant. Equivalently, $\Pi_\mu(G)=\{*\}$.

Theorem 6.3.2. *Let G be countable and let $\mu\in\mathcal{P}(G)$ be non-degenerate and assume that $H(\mu)<\infty$. If $h(\mu)=0$, then (G,μ) has the Liouville property.*

Proof. Assume $h(\mu) = 0$. Fix $g \in G$. Choose $m \geq 1$ such that $\mu^{*m}(g) > 0$. Set

$$\alpha_g := \min\{\mu^{*m}(e), \mu^{*m}(g)\}.$$

If $\alpha_g = 0$, replace μ by the *lazy* measure $\mu' := \frac{1}{2}(\mu + \delta_e)$. Then $H(\mu') < \infty$, $h(\mu') = 0$, and $\mu'^{*m}(e) > 0$ as well as $\mu'^{*m}(g) > 0$; moreover, μ-harmonic and μ'-harmonic functions coincide. Thus we may assume $\alpha_g > 0$.

Since $h(\mu) = 0$, we have $H_n = o(n)$. We claim that

$$\liminf_{n\to\infty} \big(H_{n+m} - H_n\big) = 0. \tag{6.5}$$

Indeed, if $H_{n+m} - H_n \geq \varepsilon$ for all $n \geq N$, then iterating gives $H_{N+km} \geq H_N + k\varepsilon$, hence

$$\frac{H_{N+km}}{N+km} \geq \frac{k\varepsilon}{N+km} \xrightarrow[k\to\infty]{} \frac{\varepsilon}{m},$$

contradicting $H_n/n \to 0$.

Choose a sequence $n_k \to \infty$ with $H_{n_k+m} - H_{n_k} \to 0$. Apply Lemma 6.2.12 with μ replaced by μ^{*m} and $\nu := \mu^{*n_k}$:

$$H_{n_k+m} - H_{n_k} \geq 2\alpha_g\, \delta\big(\mu^{*n_k}, g\mu^{*n_k}\big).$$

Hence $\delta(\mu^{*n_k}, g\mu^{*n_k}) \to 0$. Now Lemma 6.2.11 applied with $f \equiv 1$ gives

$$\|g * \mu^{*n_k} - \mu^{*n_k}\|_1 = \sum_x \big|\mu^{*n_k}(g^{-1}x) - \mu^{*n_k}(x)\big| \longrightarrow 0.$$

Let $f \in \ell^\infty(G)$ be μ-harmonic. Then $f = f * \mu^{*n_k}$ for every k. Therefore

$$|f(g) - f(e)| = \big|(f * (g * \mu^{*n_k}))(e) - (f * \mu^{*n_k})(e)\big| \leq \|f\|_\infty \, \|g * \mu^{*n_k} - \mu^{*n_k}\|_1 \xrightarrow[k\to\infty]{} 0.$$

Thus $f(g) = f(e)$. Since $g \in G$ was arbitrary, f is constant. □

Remark 6.3.3. Conversely, if (G, μ) has the Liouville property and $H(\mu) < \infty$, then $h(\mu) = 0$ (see [28]).

6.3.1 The Poisson boundary and amenability

In this section we explain how the Liouville property is related to amenability of groups.

Definition 6.3.4. Let $G \curvearrowright X$ be a continuous action on a compact Hausdorff space X. A Borel probability measure ν on X is called *μ-stationary* if

$$\nu = \sum_{s\in G} \mu(s)\, s_*\nu.$$

Proposition 6.3.5. *Assume that (G, μ) has the Liouville property. Then for every compact G-space X, every μ-stationary probability measure on X is G-invariant.*

Proof. Let ν be μ-stationary on X, and let $f \in C(X)$. Define $h_f \in \ell^\infty(G)$ by

$$h_f(g) := \int_X f(g \cdot x)\, d\nu(x).$$

Then h_f is μ-harmonic: for every $g \in G$,

$$\begin{aligned}(Ph_f)(g) &= \sum_s \mu(s)\, h_f(gs) = \sum_s \mu(s) \int_X f(gs \cdot x)\, d\nu(x) \\ &= \int_X f(g \cdot x)\, d\Big(\sum_s \mu(s)\, s_* \nu\Big)(x) = \int_X f(g \cdot x)\, d\nu(x) = h_f(g).\end{aligned}$$

Since (G, μ) has the Liouville property, h_f is constant. Evaluating at $g = e$ and $g \in G$ gives

$$\int_X f(g \cdot x)\, d\nu(x) = \int_X f(x)\, d\nu(x) \qquad (g \in G),$$

so ν is G-invariant. □

Lemma 6.3.6. *Let $\mu \in \mathcal{P}(G)$. Assume that for every $g \in G$ one has $\|g * \mu^{*n} - \mu^{*n}\|_1 \to 0$ as $n \to \infty$. Then (G, μ) has the Liouville property.*

Proof. Let $f \in \mathcal{H}(G, \mu)$. Since f is harmonic, $f = f * \mu^{*n}$ for all n. Hence, for every $g \in G$,

$$\begin{aligned}|f(g) - f(e)| &= \Big|\sum_x f(x)\, (g * \mu^{*n})(x) - \sum_x f(x)\, \mu^{*n}(x)\Big| \\ &\le \|f\|_\infty\, \|g * \mu^{*n} - \mu^{*n}\|_1 \xrightarrow[n\to\infty]{} 0.\end{aligned}$$

Thus $f(g) = f(e)$ for all g, i.e. f is constant. □

Theorem 6.3.7. *Let G be a countable group. If G is amenable, then there exists a probability measure $\mu \in \mathcal{P}(G)$ such that*

$$\|g * \mu^{*n} - \mu^{*n}\|_1 \longrightarrow 0 \qquad (n \to \infty)$$

for every $g \in G$.

Proof. Fix an increasing exhaustion by finite sets $\{e\} =: S_0 \subseteq S_1 \subseteq \cdots \subseteq G$ with $\bigcup_m S_m = G$. Choose a sequence of weights $(\tau_m)_{m\ge 0}$ with $\tau_m > 0$ and $\sum_m \tau_m = 1$.

By amenability, for every finite set $F \subseteq G$ and every $\varepsilon > 0$ there exists a finitely supported $\theta \in \mathcal{P}(G)$ such that

$$\|h * \theta - \theta\|_1 < \varepsilon \qquad (h \in F).$$

We construct inductively a sequence $(\alpha_m)_{m\ge 0}$ of finitely supported probability measures and a sequence $(n_m)_{m\ge 1}$ of positive integers with the following properties:

(i) $S_m \subseteq \operatorname{supp}(\alpha_m)$ for all m;

(ii) for $m \ge 1$ one has

$$\|h * \alpha_m - \alpha_m\|_1 < \frac{1}{m} \qquad \big(h \in S_m \cup (\operatorname{supp} \alpha_{m-1})^{n_m}\big). \tag{6.6}$$

(iii) for $m \ge 1$ one has

$$S_m \cup (\operatorname{supp} \alpha_{m-1})^{n_m} \subseteq \operatorname{supp}(\alpha_m). \tag{6.7}$$

Start with any finitely supported α_0 with $S_0 \subseteq \operatorname{supp}(\alpha_0)$. Assume α_{m-1} has been chosen. Pick $n_m \geq 1$ so large that

$$\Big(\sum_{j=0}^{m-1} \tau_j\Big)^{n_m} < \frac{1}{m}. \tag{6.8}$$

Increasing n_m if necessary, we may assume that $(n_m)_{m\geq 1}$ is nondecreasing. Amenability yields for the finite set $F_m := S_m \cup (\operatorname{supp}\alpha_{m-1})^{n_m}$ and $\varepsilon := 1/m$ a finitely supported θ with $\|h * \theta - \theta\|_1 < 1/m$ for all $h \in F_m$. Let $E_m := F_m \cup \operatorname{supp}(\theta)$ and let η_m be the uniform probability measure on E_m. Choose $\varepsilon_m \in (0,1)$ so small that

$$\varepsilon_m \|h * \eta_m - \eta_m\|_1 < \frac{1}{m} - \|h * \theta - \theta\|_1 \qquad (h \in F_m).$$

Then the convex combination $\alpha_m := (1-\varepsilon_m)\,\theta + \varepsilon_m\,\eta_m$ still satisfies (6.6) for all $h \in F_m$, and by construction we have (6.7).

Define $\mu := \sum_{m\geq 0} \tau_m\,\alpha_m \in \mathcal{P}(G)$ and fix $m \geq 1$ and $g \in S_{m-1}$. We set $\ell := n_m$. We claim that if $k = (k_1, \dots, k_\ell) \in \mathbb{N}^\ell \setminus \{0, 1, \dots, m-1\}^\ell$ and we write

$$\theta_k := \alpha_{k_1} * \cdots * \alpha_{k_\ell},$$

then

$$\|g * \theta_k - \theta_k\|_1 \leq \frac{2}{m}.$$

$$\theta_1 := \alpha_{k_1} * \cdots * \alpha_{k_{j-1}}, \qquad \beta := \alpha_{k_j}, \qquad \theta_2 := \alpha_{k_{j+1}} * \cdots * \alpha_{k_\ell},$$

so that $\theta_k = \theta_1 * \beta * \theta_2$. Since left translation interacts with convolution by

$$g * (\nu * \rho) = (g * \nu) * \rho \qquad (\nu, \rho \in \mathcal{P}(G)),$$

we have $g * \theta_k = (g * \theta_1) * \beta * \theta_2$. Consequently, using ℓ^1-contractivity of convolution,

$$\|g * \theta_k - \theta_k\|_1 = \big\|\big((g * \theta_1) * \beta - \theta_1 * \beta\big) * \theta_2\big\|_1 \leq \|(g * \theta_1) * \beta - \theta_1 * \beta\|_1.$$

We now estimate $\|(g * \theta_1) * \beta - \theta_1 * \beta\|_1$. If $j = 1$, then $\theta_1 = \delta_e$ and the expression reduces to $\|g * \beta - \beta\|_1$, which is $< 1/m$ because $g \in S_{m-1} \subseteq S_{k_1}$ and (6.6) applies to α_{k_1}. Assume $j \geq 2$. Since $k_i < m$ for $i < j$ and the supports are increasing by (6.7), we have $\operatorname{supp}(\alpha_{k_i}) \subseteq \operatorname{supp}(\alpha_{m-1})$ for all $i < j$. Hence

$$\operatorname{supp}(\theta_1) \subseteq (\operatorname{supp}\alpha_{m-1})^{j-1} \subseteq (\operatorname{supp}\alpha_{m-1})^{\ell} \subseteq (\operatorname{supp}\alpha_{m-1})^{n_{k_j}} \subseteq (\operatorname{supp}\alpha_{k_j-1})^{n_{k_j}}.$$

Also $g \in S_{m-1} \subseteq \operatorname{supp}(\alpha_{m-1})$ by (6.7), so $\operatorname{supp}(g * \theta_1) \subseteq (\operatorname{supp}\alpha_{m-1})^{\ell} \subseteq (\operatorname{supp}\alpha_{k_j-1})^{n_{k_j}}$.

Now observe that

$$\theta_1 * \beta = \sum_{h \in \operatorname{supp}(\theta_1)} \theta_1(h)\,(h * \beta), \qquad (g * \theta_1) * \beta = \sum_{h \in \operatorname{supp}(g*\theta_1)} (g * \theta_1)(h)\,(h * \beta).$$

Therefore, by the triangle inequality and (6.6) applied to $\beta = \alpha_{k_j}$ (note that $k_j \geq m$),

$$\|\beta - \theta_1 * \beta\|_1 \leq \sum_{h \in \operatorname{supp}(\theta_1)} \theta_1(h)\,\|\beta - h * \beta\|_1 < \frac{1}{m},$$

and similarly $\|\beta - (g * \theta_1) * \beta\|_1 < 1/m$. Hence

$$\|(g * \theta_1) * \beta - \theta_1 * \beta\|_1 \leq \|(g * \theta_1) * \beta - \beta\|_1 + \|\beta - \theta_1 * \beta\|_1 < \frac{2}{m}.$$

This proves our claim.

We now aim at proving $\|g * \mu^{*\ell} - \mu^{*\ell}\|_1 < \frac{4}{m}$. Write

$$\mu^{*\ell} = \sum_{k \in \mathbb{N}^\ell} \tau_{k_1} \cdots \tau_{k_\ell}\, \theta_k.$$

Split the sum as $\mu^{*\ell} = \nu_1 + \nu_2$, where ν_1 is the partial sum over $k \in \{0, 1, \ldots, m-1\}^\ell$ and ν_2 is the complementary sum. Then

$$\|g * \nu_1 - \nu_1\|_1 \le \|g * \nu_1\|_1 + \|\nu_1\|_1 = 2\,\|\nu_1\|_1 = 2\Big(\sum_{j=0}^{m-1} \tau_j\Big)^\ell < \frac{2}{m}$$

by (6.8). On the other hand, Claim 1 gives

$$\|g * \nu_2 - \nu_2\|_1 \le \sum_{k \notin \{0,\ldots,m-1\}^\ell} \tau_{k_1} \cdots \tau_{k_\ell}\, \|g * \theta_k - \theta_k\|_1 \le \frac{2}{m}.$$

Hence $\|g * \mu^{*\ell} - \mu^{*\ell}\|_1 \le \|g * \nu_1 - \nu_1\|_1 + \|g * \nu_2 - \nu_2\|_1 < 4/m$, proving our claim.

Thus, for every $g \in S_{m-1}$ we have $\|g * \mu^{*n_m} - \mu^{*n_m}\|_1 \to 0$ as $m \to \infty$. Moreover, for any $n \ge n_m$,

$$\|g * \mu^{*n} - \mu^{*n}\|_1 = \|(g * \mu^{*n_m} - \mu^{*n_m}) * \mu^{*(n-n_m)}\|_1 \le \|g * \mu^{*n_m} - \mu^{*n_m}\|_1,$$

so the same convergence holds along the full sequence $n \to \infty$. In particular, $\|g * \mu^{*n_m} - \mu^{*n_m}\|_1 \to 0$ along the subsequence n_m. Since $\bigcup_m S_m = G$, this proves the theorem. □

Theorem 6.3.8. *For a countable group G, the following are equivalent:*

(i) G is amenable;

(ii) there exists $\mu \in \mathcal{P}(G)$ such that (G, μ) has the Liouville property.

Proof. (2)⇒(1): assume that (G, μ) has the Liouville property. Consider the left translation action of G on the Stone–Čech compactification βG. The compact convex set of probability measures on βG contains μ-stationary measures e.g. weak-$*$ accumulation points of the Cesàro averages $\frac{1}{N}\sum_{n=1}^N \mu^{*n} * \delta_x$. By Proposition 6.3.5, every μ-stationary measure is G-invariant. An invariant probability measure on βG induces a left-invariant mean on $C(\beta G) \cong \ell^\infty(G)$. Thus G is amenable.

(1)⇒(2): by Theorem 6.3.7, if G is amenable then there exists $\mu \in \mathcal{P}(G)$ such that $\|g * \mu^{*n} - \mu^{*n}\|_1 \to 0$ for every $g \in G$. Moreover, in the construction of μ in the proof of Theorem 6.3.7 one has $S_m \subseteq \operatorname{supp}(\alpha_m)$ for all m and $\tau_m > 0$ for all m, hence $\operatorname{supp}(\mu) = G$. By Lemma 6.3.6, (G, μ) has the Liouville property. □

6.3.2 Slow entropy growth

Definition 6.3.9. A finitely generated group G has *slow entropy growth* if there exists a finitely supported symmetric nondegenerate probability measure μ with $\mu(e) > 0$ such that

$$\liminf_{n\to\infty} n\big(H(\mu^{*(n+1)}) - H(\mu^{*n})\big) \;<\; \infty.$$

Lemma 6.3.10. *If G has polynomial growth, then G has slow entropy growth for any symmetric finitely supported measure $\mu \in \mathcal{P}(G)$.*

Proof. Set $H_n := H(\mu^{*n})$. Let p be a probability measure with finite support A. If X is an A-valued random variable with law p, then by concavity of log,

$$H(p) = \mathbb{E}\Big[\log \frac{1}{p(X)}\Big] \le \log\Big(\mathbb{E}\Big[\frac{1}{p(X)}\Big]\Big) = \log\Big(\sum_{x\in A} p(x)\,\frac{1}{p(x)}\Big) = \log|A|.$$

Applying this to $p = \mu^{*n}$ gives $H_n \le \log|\operatorname{supp}(\mu^{*n})|$.

Fix a finite generating set S of G and choose $R \ge 1$ such that $\operatorname{supp}(\mu) \subseteq B_S(R)$. Then every product of n elements from $\operatorname{supp}(\mu)$ lies in $B_S(nR)$, hence

$$\operatorname{supp}(\mu^{*n}) \subseteq B_S(nR).$$

Since G has polynomial growth, there exist constants $C, d > 0$ such that $|B_S(r)| \le Cr^d$ for all $r \ge 1$. Therefore

$$H_n \le \log|B_S(nR)| \le \log(C(nR)^d) = \log C + d\log R + d\log n \le C_0 + d\log n$$

for a constant C_0 independent of n.

We claim

$$\liminf_{n\to\infty} n\big(H_{n+1} - H_n\big) \le d,$$

which in particular implies that the liminf in Definition 6.3.9 is finite.

Indeed, suppose for contradiction that there exists $\varepsilon > 0$ and N such that

$$H_{n+1} - H_n \ge \frac{d+\varepsilon}{n} \qquad \text{for all } n \ge N.$$

Then for $n > N$ we can sum the increments to obtain

$$H_n \ge H_N + (d+\varepsilon)\sum_{k=N}^{n-1}\frac{1}{k}.$$

Using the standard estimate $\sum_{k=N}^{n-1}\frac{1}{k} \ge \log\frac{n}{N}$, we get

$$H_n \ge (d+\varepsilon)\log n + \text{const},$$

contradicting the upper bound $H_n \le C_0 + d\log n$ for large n. This proves the claim. ☐

Chapter 7

Non-commutative harmonic analysis

7.1 Definitions

Let $\mathcal{H}$ be a complex vector space. An *inner product* on $\mathcal{H}$ is a map $\langle\cdot,\cdot\rangle : \mathcal{H}\times\mathcal{H}\to\mathbb{C}$ which is linear in the first variable, conjugate-linear in the second, and satisfies

$$\langle\xi,\xi\rangle \geq 0 \text{ with equality iff } \xi = 0, \qquad \langle\xi,\eta\rangle = \overline{\langle\eta,\xi\rangle}.$$

The induced norm is $\|\xi\| := \langle\xi,\xi\rangle^{1/2}$. The pair $(\mathcal{H},\langle\cdot,\cdot\rangle)$ is a *Hilbert space* if it is complete for this norm. We write $\mathcal{B}(\mathcal{H})$ for the algebra of bounded linear operators $T : \mathcal{H}\to\mathcal{H}$. For $T\in\mathcal{B}(\mathcal{H})$ define

$$\|T\| := \sup\{\|T\xi\| : \xi\in\mathcal{H},\ \|\xi\| = 1\}.$$

Then $\|T\xi\| \leq \|T\|\,\|\xi\|$ for all $\xi\in\mathcal{H}$, and $\|ST\| \leq \|S\|\,\|T\|$ for all $S,T\in\mathcal{B}(\mathcal{H})$. The *strong operator topology* (SOT) on $\mathcal{B}(\mathcal{H})$ is the topology of pointwise convergence on $\mathcal{H}$:

$$T_i \to T \text{ strongly} \iff \|T_i\xi - T\xi\| \to 0 \text{ for all } \xi\in\mathcal{H}.$$

The *weak operator topology* (WOT) on $\mathcal{B}(\mathcal{H})$ is the topology of pointwise convergence of matrix coefficients:

$$T_i \to T \text{ weakly} \iff \langle T_i\xi,\eta\rangle \to \langle T\xi,\eta\rangle \text{ for all } \xi,\eta\in\mathcal{H}.$$

In particular, strong convergence implies weak convergence. Let G be a countable group. A *unitary representation* of G on a Hilbert space $\mathcal{H}$ is a group homomorphism

$$\pi : G \to \mathcal{U}(\mathcal{H}) := \{U\in\mathcal{B}(\mathcal{H}) : U^*U = UU^* = \mathbb{1}\}.$$

For a countable group G, define $\lambda : G\to\mathcal{U}(\ell^2(G))$ by

$$(\lambda(g)f)(x) := f(g^{-1}x) \qquad (g,x\in G,\ f\in\ell^2(G)).$$

This gives a unitary representation, called the *left-regular representation.*

Definition 7.1.1. Let G be a countable group and let $\lambda : G\to\mathcal{U}(\ell^2(G))$ be the left-regular representation. The *reduced group C^*-algebra $C^*_r(G)$* is

$$C^*_r(G) := \overline{\operatorname{span}}^{\|\cdot\|}\{\lambda(g) : g\in G\} \subseteq \mathcal{B}(\ell^2(G)).$$

Equivalently, $C^*_r(G)$ is the smallest C^*-subalgebra of $\mathcal{B}(\ell^2(G))$ containing $\lambda(G)$.

Let G be a countable group. For $f, g \in \ell^1(G)$ we define the convolution $f * g \in \ell^1(G)$ by

$$(f * g)(x) := \sum_{y \in G} f(y)\, g(y^{-1}x) \qquad (x \in G).$$

For $f \in \ell^1(G)$ we define the (left) convolution operator $\lambda(f)$ on $\ell^2(G)$ by

$$(\lambda(f)\xi)(x) := \sum_{y \in G} f(y)\, \xi(y^{-1}x) \qquad (x \in G,\ \xi \in \ell^2(G)).$$

The next lemma records the basic C^*-algebraic properties of this assignment.

Lemma 7.1.2. *Let G be a countable group. For every $f \in \ell^1(G)$, the operator $\lambda(f)$ is bounded and belongs to $C_r^*(G)$. Moreover, for all $f, g \in \ell^1(G)$ one has $\lambda(f * g) = \lambda(f)\lambda(g)$ and $\|\lambda(f)\| \leq \|f\|_1$. In particular, if $\mu \in \mathcal{P}(G)$ is a probability measure (so $\mu \in \ell^1(G)$ with $\|\mu\|_1 = 1$), then $T_\mu := \lambda(\mu)$ is a contraction and*

$$T_{\mu * \nu} = T_\mu T_\nu \qquad (\mu, \nu \in \mathcal{P}(G)).$$

Proof. For $f \in \ell^1(G)$ and $\xi \in \ell^2(G)$, the formula

$$(\lambda(f)\xi)(x) = \sum_{y \in G} f(y)\, \xi(y^{-1}x)$$

defines the usual left-convolution operator. Young's inequality gives $\|\lambda(f)\xi\|_2 \leq \|f\|_1 \|\xi\|_2$, hence $\|\lambda(f)\| \leq \|f\|_1$ and the defining series for $\lambda(f)$ converges in operator norm.

To see that $\lambda(f) \in C_r^*(G)$, approximate f in $\ell^1(G)$ by finitely supported functions f_n. Then $\|\lambda(f_n) - \lambda(f)\| \leq \|f_n - f\|_1 \to 0$ by the norm estimate above, and each $\lambda(f_n)$ belongs to the linear span of $\{\lambda(g) : g \in G\}$. Since $C_r^*(G)$ is the operator-norm closure of that span, it follows that $\lambda(f) \in C_r^*(G)$.

To check multiplicativity, compute for $\xi \in \ell^2(G)$ and $x \in G$:

$$\begin{aligned}
\big(\lambda(f)\lambda(g)\xi\big)(x) &= \sum_{y \in G} f(y)\, (\lambda(g)\xi)(y^{-1}x) = \sum_{y \in G} f(y) \sum_{z \in G} g(z)\, \xi(z^{-1}y^{-1}x) \\
&= \sum_{w \in G} \left(\sum_{y \in G} f(y)\, g(y^{-1}w) \right) \xi(w^{-1}x) = \sum_{w \in G} (f * g)(w)\, \xi(w^{-1}x) \\
&= \big(\lambda(f * g)\xi\big)(x).
\end{aligned}$$

Thus $\lambda(f * g) = \lambda(f)\lambda(g)$. The final statements follow by specializing to $f = \mu$, $g = \nu$. □

Definition 7.1.3. Let G be a countable group and let $\delta_e \in \ell^2(G)$ denote the delta function at the identity. The *canonical trace* on $C_r^*(G)$ is the linear functional $\tau : C_r^*(G) \to \mathbb{C}$ defined by

$$\tau(a) := \langle a\, \delta_e, \delta_e \rangle \qquad (a \in C_r^*(G)).$$

Lemma 7.1.4. *Let G be a countable group and let τ be the canonical trace on $C_r^*(G)$. Then τ is positive. Moreover, if $a \in C_r^*(G)$ satisfies $a \geq 0$ and $\tau(a) = 0$, then $a = 0$.*

Proposition 7.1.5. *Let G be a countable group. Then τ is a tracial state on $C_r^*(G)$ and satisfies $\tau(\lambda(f)) = f(e)$ for all $f \in \ell^1(G)$. In particular, for $\mu \in \mathcal{P}(G)$ and $n \geq 1$ one has*

$$\tau(T_\mu^n) = \mu^{*n}(e),$$

so the return probability at time n of the μ-random walk equals the trace of T_μ^n.

Proof. Linearity is immediate from the definition. By Lemma 7.1.4, τ is positive. Also $\tau(\mathbb{1}) = \|\delta_e\|_2^2 = 1$, so τ is a state.

For $f \in \ell^1(G)$,

$$\tau(\lambda(f)) = \langle \lambda(f)\delta_e, \delta_e \rangle = (\lambda(f)\delta_e)(e) = \sum_{x \in G} f(x)\, \delta_e(x^{-1}e) = f(e).$$

To see that τ is a trace, it suffices by density to check it on the span of $\{\lambda(g) : g \in G\}$. For $g, h \in G$,

$$\tau(\lambda(g)\lambda(h)) = \tau(\lambda(gh)) = \delta_{gh,e} = \delta_{hg,e} = \tau(\lambda(hg)) = \tau(\lambda(h)\lambda(g)),$$

and bilinearity gives $\tau(ab) = \tau(ba)$ on the dense $*$-subalgebra, hence on $C_r^*(G)$ by continuity.

Finally, by the previous lemma $T_\mu^n = \lambda(\mu)^n = \lambda(\mu^{*n})$, so

$$\tau(T_\mu^n) = \tau(\lambda(\mu^{*n})) = \mu^{*n}(e).$$

□

7.2 Almost invariant vectors and amenability

Definition 7.2.1. Let G be finitely generated and $\pi : G \to \mathcal{U}(\mathcal{H})$ a unitary representation. We say that π has *almost invariant vectors* if there exists a sequence of unit vectors $(\xi_n)_{n \geq 1} \subset \mathcal{H}$ such that for some (equivalently, any) finite generating set S of G one has

$$\max_{s \in S} \|\pi(s)\xi_n - \xi_n\| \longrightarrow 0 \qquad (n \to \infty).$$

We now relate amenability of G to almost invariant vectors for the left-regular representation and to the spectral behavior of the Markov operator of a random walk. The following two lemmas show that amenability of G is equivalent to the existence of almost invariant vectors for λ.

Lemma 7.2.2. *Let G be an infinite finitely generated amenable group and let $\lambda : G \to \mathcal{U}(\ell^2(G))$ be the left-regular representation. Then λ has almost invariant vectors.*

Proof. Fix a finite generating set S of G. Since G is amenable, there exists a Følner sequence $(F_n)_{n \geq 1}$ for S, meaning

$$\frac{|sF_n \triangle F_n|}{|F_n|} \longrightarrow 0 \qquad (n \to \infty)$$

for every $s \in S$. Set $\xi_n := |F_n|^{-1/2}\, \mathbf{1}_{F_n} \in \ell^2(G)$. Then $\|\xi_n\|_2 = 1$ and, for $s \in S$,

$$\|\lambda(s)\xi_n - \xi_n\|_2^2 = \frac{\|\mathbf{1}_{sF_n} - \mathbf{1}_{F_n}\|_2^2}{|F_n|} = \frac{|sF_n \triangle F_n|}{|F_n|} \longrightarrow 0.$$

Hence λ has almost invariant vectors. □

Lemma 7.2.3. *Let G be finitely generated. Assume that the left-regular representation $\lambda : G \to \mathcal{U}(\ell^2(G))$ has almost invariant vectors. Then G is amenable.*

Proof. Fix a finite generating set S. Let (ξ_n) be unit vectors in $\ell^2(G)$ with $\max_{s\in S}\|\lambda(s)\xi_n - \xi_n\|_2 \to 0$. Set $\eta_n := |\xi_n|$ pointwise. Then $\|\eta_n\|_2 = \|\xi_n\|_2 = 1$ and, since $\lambda(s)$ acts by permuting coordinates,

$$\|\lambda(s)\eta_n - \eta_n\|_2^2 = \sum_{x\in G}\big|\eta_n(s^{-1}x) - \eta_n(x)\big|^2 \le \sum_{x\in G}\big|\xi_n(s^{-1}x) - \xi_n(x)\big|^2 = \|\lambda(s)\xi_n - \xi_n\|_2^2.$$

Hence also $\max_{s\in S}\|\lambda(s)\eta_n - \eta_n\|_2 \to 0$. Define $p_n \in \ell^1(G)$ by $p_n(x) := \eta_n(x)^2$. Then $p_n \ge 0$ and $\sum_x p_n(x) = \|\eta_n\|_2^2 = 1$, so $p_n \in \mathcal{P}(G)$. Moreover, for $s \in S$ we estimate using Cauchy–Schwarz:

$$\begin{aligned}\|sp_n - p_n\|_1 &= \sum_{x\in G}\big|p_n(s^{-1}x) - p_n(x)\big| = \sum_x |\eta_n(s^{-1}x) - \eta_n(x)|\,\big(\eta_n(s^{-1}x) + \eta_n(x)\big)\\ &\le \Big(\sum_x |\eta_n(s^{-1}x) - \eta_n(x)|^2\Big)^{1/2}\Big(\sum_x \big(\eta_n(s^{-1}x) + \eta_n(x)\big)^2\Big)^{1/2}\\ &\le \|\lambda(s)\eta_n - \eta_n\|_2 \cdot 2\|\eta_n\|_2 = 2\|\lambda(s)\eta_n - \eta_n\|_2.\end{aligned}$$

Thus $\max_{s\in S}\|sp_n - p_n\|_1 \to 0$. For each n define a (countably additive) probability measure on G by

$$m_n(A) := \sum_{g\in G}\mathbf{1}_A(g)\,|\xi_n(g)|^2 = \sum_{g\in A} p_n(g) \qquad (A \subseteq G).$$

Equivalently, define a mean $M_n : \ell^\infty(G) \to \mathbb{R}$ by

$$M_n(f) := \sum_{g\in G} f(g)\,p_n(g) \qquad (f \in \ell^\infty(G)).$$

Then M_n is positive and $M_n(\mathbf{1}) = 1$. Moreover, for $s \in S$ and $f \in \ell^\infty(G)$,

$$\big|M_n(s\cdot f) - M_n(f)\big| = \Big|\sum_{g\in G} f(g)\,\big(p_n(s^{-1}g) - p_n(g)\big)\Big| \le \|f\|_\infty\,\|sp_n - p_n\|_1.$$

Since $\max_{s\in S}\|sp_n - p_n\|_1 \to 0$, the means M_n are asymptotically left S-invariant. By weak-$*$ compactness of the unit ball of $(\ell^\infty(G))^*$, there exists a weak-$*$ accumulation point M of (M_n). The above inequality implies that M is left S-invariant, hence left G-invariant since S generates G. By Theorem 4.1.7, the existence of a left G-invariant mean on $\ell^\infty(G)$ is equivalent to amenability. Therefore G is amenable. □

Definition 7.2.4. Let G be finitely generated and let $\mu \in \mathcal{P}(G)$ be finitely supported. The *Markov operator* (or averaging operator) of μ is

$$M_\mu := T_\mu = \lambda(\mu) = \sum_{g\in G}\mu(g)\,\lambda(g) \in C_r^*(G) \subseteq \mathcal{B}(\ell^2(G)).$$

If μ is symmetric then M_μ is self-adjoint.

Lemma 7.2.5. *Let G be finitely generated and let $\mu \in \mathcal{P}(G)$ be finitely supported and symmetric, with* supp(μ) *generating G. Then $\|M_\mu\| \le 1$. Moreover, the following are equivalent:*

(i) $\|M_\mu\| = 1$.

(ii) The left-regular representation λ has almost invariant vectors.

Proof. Since $\|\lambda(\mu)\| \le \|\mu\|_1 = 1$ by Lemma 7.1.2, we have $\|M_\mu\| \le 1$.

Assume first that λ has almost invariant vectors for some finite symmetric generating set S. Since $\mathrm{supp}(\mu)$ is finite, it is contained in some finite symmetric generating set (replace S by $S \cup \mathrm{supp}(\mu)$ if needed). Take unit vectors ξ_n with $\max_{s\in\mathrm{supp}(\mu)} \|\lambda(s)\xi_n - \xi_n\|_2 \to 0$. Then

$$\|M_\mu\xi_n - \xi_n\|_2 = \Big\|\sum_s \mu(s)\big(\lambda(s)\xi_n - \xi_n\big)\Big\|_2 \le \sum_s \mu(s)\,\|\lambda(s)\xi_n - \xi_n\|_2 \longrightarrow 0.$$

Since $\|M_\mu\| \le 1$ and $\|M_\mu\xi_n\|_2 \to 1$, it follows that $\|M_\mu\| = 1$.

Conversely, assume $\|M_\mu\| = 1$. Since μ is symmetric, M_μ is self-adjoint. Therefore $\|M_\mu\| = \sup_{\|\xi\|=1}\langle M_\mu\xi, \xi\rangle$. Pick unit vectors ξ_n with $\langle M_\mu\xi_n, \xi_n\rangle \to 1$. For each $s \in \mathrm{supp}(\mu)$ we have

$$\|\lambda(s)\xi - \xi\|_2^2 = 2 - 2\Re\langle\lambda(s)\xi, \xi\rangle.$$

Since μ is symmetric, $\langle M_\mu\xi, \xi\rangle = \sum_s \mu(s)\Re\langle\lambda(s)\xi, \xi\rangle$, and hence

$$1 - \langle M_\mu\xi, \xi\rangle = \frac{1}{2}\sum_s \mu(s)\,\|\lambda(s)\xi - \xi\|_2^2.$$

Applying this to $\xi = \xi_n$ shows that the μ-average of $\|\lambda(s)\xi_n - \xi_n\|_2^2$ tends to 0. Since $\mathrm{supp}(\mu)$ is finite and $\min_{s\in\mathrm{supp}(\mu)} \mu(s) > 0$, it follows that

$$\max_{s\in\mathrm{supp}(\mu)} \|\lambda(s)\xi_n - \xi_n\|_2 \longrightarrow 0.$$

Thus λ has almost invariant vectors. □

Theorem 7.2.6 (Kesten's criterion)**.** *Let G be a finitely generated group. The following are equivalent:*

(i) G is amenable.

(ii) The left-regular representation $\lambda : G \to \mathcal{U}(\ell^2(G))$ has almost invariant vectors.

(iii) There exists a finitely supported symmetric probability measure $\mu \in \mathcal{P}(G)$ with $\mathrm{supp}(\mu)$ generating G such that the Markov operator M_μ satisfies $\|M_\mu\| = 1$.

Proof. (1)⇒(2) is Lemma 7.2.2. (2)⇒(1) is Lemma 7.2.3. (2)⇒(3): choose any finitely supported symmetric probability measure μ whose support generates G (for example, the uniform measure on a finite symmetric generating set). Then Lemma 7.2.5 gives $\|M_\mu\| = 1$. (3)⇒(2) is also Lemma 7.2.5. □

We can also interpret the above in terms of random walks. If μ is as in the theorem, then by Proposition 7.1.5 we have

$$\mu^{*n}(e) = \tau(M_\mu^n) \qquad (n \ge 1).$$

Lemma 7.2.7. *Let G be finitely generated and let $\mu \in \mathcal{P}(G)$ be finitely supported and symmetric. Let $M_\mu \in C_r^*(G)$ be the associated Markov operator. We have*

$$\mu^{*2n}(e) \le \|M_\mu\|^{2n} \quad \textit{for all } n \ge 1 \textit{ and } \quad \lim_{n\to\infty}\big(\mu^{*2n}(e)\big)^{1/2n} = \|M_\mu\|.$$

Proof. Since μ is symmetric, M_μ is self-adjoint. By Proposition 7.1.5,

$$\mu^{*2n}(e) = \tau(M_\mu^{2n}).$$

For any positive operator A in a C^*-algebra and any state φ, one has $0 \le \varphi(A) \le \|A\|$. Applying this with $A = M_\mu^{2n}$ and $\varphi = \tau$ gives

$$\mu^{*2n}(e) = \tau(M_\mu^{2n}) \le \|M_\mu^{2n}\| = \|M_\mu\|^{2n}.$$

Let $B := C^*(M_\mu) \subseteq C_r^*(G)$ be the unital abelian C^*-subalgebra generated by M_μ. By functional calculus, $B \cong C(\sigma(M_\mu))$, where $\sigma(M_\mu) \subset \mathbb{R}$ is the spectrum of M_μ.

The restriction of the canonical trace τ to B is a positive linear functional with $\tau(\mathbb{1}) = 1$; hence it corresponds to a probability measure ν on $\sigma(M_\mu)$ such that

$$\tau(f(M_\mu)) = \int_{\sigma(M_\mu)} f(t)\, d\nu(t) \qquad (f \in C(\sigma(M_\mu))).$$

Since τ is faithful on $C_r^*(G)$ by Lemma 7.1.4, it is faithful on B, and therefore ν has full support on $\sigma(M_\mu)$.

Applying this to $f(t) = t^{2n}$ gives

$$\tau(M_\mu^{2n}) = \int t^{2n}\, d\nu(t).$$

By Proposition 7.1.5, $\tau(M_\mu^{2n}) = \mu^{*2n}(e)$. Hence

$$\mu^{*2n}(e) = \int t^{2n}\, d\nu(t).$$

Since $|t| \le \|M_\mu\|$ on $\sigma(M_\mu)$, we get

$$\big(\mu^{*2n}(e)\big)^{1/2n} \le \|M_\mu\|.$$

For the reverse inequality, fix $\varepsilon > 0$. By definition of the norm of a self-adjoint element, $\|M_\mu\| = \max\{|t| : t \in \sigma(M_\mu)\}$. Since ν has full support, the set $\{t \in \sigma(M_\mu) : |t| > \|M_\mu\| - \varepsilon\}$ has positive ν-measure. Therefore

$$\mu^{*2n}(e) = \int t^{2n}\, d\nu(t) \ge (\|M_\mu\| - \varepsilon)^{2n}\, \nu\big(\{|t| > \|M_\mu\| - \varepsilon\}\big),$$

and taking $2n$-th roots and letting $n \to \infty$ yields

$$\liminf_{n\to\infty} \big(\mu^{*2n}(e)\big)^{1/2n} \ge \|M_\mu\| - \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we obtain the result. □

7.3 Kazhdan's property (T)

7.3.1 Definitions and basic consequences

Definition 7.3.1. A countable group G has *property (T)* if every unitary representation $\pi : G \to \mathcal{U}(\mathcal{H})$ with almost invariant vectors has a non-zero invariant vector.

We recall that the left-regular representation $\lambda : G \to \mathcal{U}(\ell^2(G))$ was defined in Section 7.1.

Lemma 7.3.2. *If G is infinite amenable, then G does not have property (T).*

Proof. Fix a finite generating set S of G and let $\lambda : G \to \mathcal{U}(\ell^2(G))$ be the left-regular representation. By Lemma 7.2.2 (proved in Section 7.2), the representation λ has almost invariant vectors.

On the other hand, an invariant vector for λ is a function $f \in \ell^2(G)$ with $f(s^{-1}x) = f(x)$ for all $s \in S$ and $x \in G$; since S generates G, this forces f to be constant. As G is infinite, the only constant vector in $\ell^2(G)$ is 0, so λ has no non-zero invariant vectors. Therefore G cannot have property (T). □

Lemma 7.3.3. *If a countable group G has property* (T)*, then G is finitely generated.*

Proof. Assume towards a contradiction that G is not finitely generated. Since G is countable, we may write it as an increasing union of finitely generated subgroups

$$G = \bigcup_{i\geq 1} G_i, \qquad G_1 \leq G_2 \leq \cdots ,$$

with $G_i \neq G$ for all i. In particular, $[G : G_i] = \infty$ for all i (otherwise G would be a finite-index extension of the finitely generated group G_i and hence finitely generated).

For each i, consider the quasi-regular representation ρ_i of G on $\ell^2(G/G_i)$,

$$(\rho_i(g)f)(xG_i) := f(g^{-1}xG_i).$$

Let $\delta_{G_i} \in \ell^2(G/G_i)$ be the delta function at the trivial coset G_i. Then δ_{G_i} is fixed by G_i.

Let $\rho := \bigoplus_{i\geq 1} \rho_i$ on $\mathcal{H} := \bigoplus_{i\geq 1} \ell^2(G/G_i)$ and let $\xi_i \in \mathcal{H}$ be the unit vector supported in the i-th summand equal to δ_{G_i}. For every finite subset $K \subseteq G$, choose i such that $K \subseteq G_i$. Then $\rho(g)\xi_i = \xi_i$ for all $g \in K$. Hence (ξ_i) is a sequence of almost invariant vectors for ρ.

On the other hand, $\ell^2(G/G_i)$ has a non-zero G-invariant vector if and only if $[G : G_i] < \infty$ (it must be a non-zero constant function), which never happens. Therefore ρ has no non-zero invariant vector. This contradicts property (T). Thus G is finitely generated. □

7.3.2 Cocycles, coboundaries and harmonic functions

Definition 7.3.4. Let $\pi : G \to \mathcal{U}(\mathcal{H})$ be unitary. A map $b : G \to \mathcal{H}$ is a *1-cocycle* if

$$b(gx) = b(g) + \pi(g)b(x) \qquad (g, x \in G).$$

A *1-coboundary* is $b(g) = \xi - \pi(g)\xi$ for some $\xi \in \mathcal{H}$. An *approximate coboundary* is a cocycle b for which there exists a sequence (ξ_n) in $\mathcal{H}$ such that for every finite $K \subseteq G$ one has

$$\max_{g\in K} \left\| b(g) - (\xi_n - \pi(g)\xi_n) \right\| \longrightarrow 0 \qquad (n \to \infty).$$

Equivalently, it is a limit of coboundaries for the topology of uniform convergence on finite subsets of G. Denote by $Z^1(G, \pi)$ the space of 1-cocycles, by $B^1(G, \pi)$ the space of 1-coboundaries, and by $\overline{B^1}(G, \pi)$ the space of approximate coboundaries. The *reduced cohomology* is

$$H^1_{\mathrm{red}}(G, \pi) := Z^1(G, \pi)/\overline{B^1}(G, \pi).$$

Fix a finitely supported symmetric probability measure μ on G with $\mu(e) > 0$ and $\operatorname{supp}(\mu)$ generating G.

Definition 7.3.5. A cocycle $b \in Z^1(G,\pi)$ is *μ-harmonic* if

$$\sum_{x\in G} \mu(x)\, b(x) = 0,$$

equivalently $\sum_x \mu(x)\, b(gx) = b(g)$ for all $g \in G$.

Proposition 7.3.6. *Every class in $H^1_{\mathrm{red}}(G,\pi)$ has a unique μ-harmonic representative.*

Proof. Set $E := \operatorname{supp}(\mu)$, which is finite and generates G. Define an inner product on $Z^1(G,\pi)$ by

$$\langle b, c\rangle_{Z^1} := \sum_{x\in G} \mu(x)\, \langle b(x), c(x)\rangle = \sum_{x\in E} \mu(x)\, \langle b(x), c(x)\rangle,$$

so that $\|b\|^2_{Z^1} := \langle b, b\rangle_{Z^1}$. For $\xi \in \mathcal{H}$ let $\partial\xi \in B^1(G,\pi)$ denote the coboundary $\partial\xi(g) := \xi - \pi(g)\xi$. Fix $b \in Z^1(G,\pi)$ and compute, using unitarity of π and symmetry of μ,

$$\begin{aligned}
\langle b, \partial\xi\rangle_{Z^1} &= \sum_{x\in G} \mu(x)\, \langle b(x), \xi - \pi(x)\xi\rangle \\
&= \Big\langle \sum_x \mu(x)\, b(x), \xi\Big\rangle \; - \sum_x \mu(x)\, \langle b(x), \pi(x)\xi\rangle \\
&= \Big\langle \sum_x \mu(x)\, b(x), \xi\Big\rangle \; - \sum_x \mu(x)\, \langle \pi(x^{-1}) b(x), \xi\rangle \\
&= \Big\langle \sum_x \mu(x)\, b(x), \xi\Big\rangle \; - \sum_y \mu(y)\, \langle \pi(y) b(y^{-1}), \xi\rangle.
\end{aligned}$$

Here we substituted $y = x^{-1}$ and used $\mu(y) = \mu(y^{-1})$. Since b is a cocycle, evaluating the cocycle identity at $e = yy^{-1}$ gives $b(y^{-1}) = -\pi(y^{-1}) b(y)$, hence $\pi(y) b(y^{-1}) = -b(y)$. Therefore

$$\langle b, \partial\xi\rangle_{Z^1} = 2\Big\langle \sum_{x\in G} \mu(x)\, b(x), \xi\Big\rangle.$$

Consequently,

$$b \perp B^1(G,\pi) \quad\Longleftrightarrow\quad \sum_{x\in G} \mu(x)\, b(x) = 0.$$

The right-hand condition is precisely the μ-harmonicity of b.

Let $\overline{B^1}(G,\pi)$ be the closure of $B^1(G,\pi)$ inside the Hilbert space $Z^1(G,\pi)$. Then $\overline{B^1}(G,\pi)$ is a closed subspace, so every $b \in Z^1(G,\pi)$ has a unique orthogonal decomposition

$$b = b_0 + b_1, \qquad b_0 \in \overline{B^1}(G,\pi), \quad b_1 \in \overline{B^1}(G,\pi)^\perp.$$

By Step 2, b_1 is μ-harmonic. Also $b - b_1 = b_0 \in \overline{B^1}(G,\pi)$, so b and b_1 represent the same class in $\overline{H}^1(G,\pi)$. If b_1' is another μ-harmonic cocycle representing the same class, then $b_1 - b_1' \in \overline{B^1}(G,\pi)$ and $b_1 - b_1' \in \overline{B^1}(G,\pi)^\perp$, hence $b_1 = b_1'$. Finally, by definition, b_1 is the orthogonal projection of b onto $\overline{B^1}(G,\pi)^\perp$. □

7.3.3 The Delorme–Guichardet theorem

We now recast the first cohomology language in terms of fixed points for affine actions. See [10, 24].

Definition 7.3.7. Let $\mathcal{H}$ be a Hilbert space. An *affine isometric action* of a group G on $\mathcal{H}$ is a homomorphism

$$\alpha : G \to \mathrm{Isom}(\mathcal{H}),$$

where $\mathrm{Isom}(\mathcal{H})$ denotes the group of affine isometries of $\mathcal{H}$.

Lemma 7.3.8. *Let α be an affine isometric action of G on a Hilbert space $\mathcal{H}$. Then for every $g \in G$ there exist unique elements $\pi(g) \in \mathcal{U}(\mathcal{H})$ and $b(g) \in \mathcal{H}$ such that*

$$\alpha(g)\xi = \pi(g)\xi + b(g) \qquad (g \in G,\ \xi \in \mathcal{H}).$$

The map $\pi : G \to \mathcal{U}(\mathcal{H})$ is a unitary representation and $b : G \to \mathcal{H}$ is a 1-cocycle for π. Conversely, given a unitary representation π and a 1-cocycle b, the above formula defines an affine isometric action.

Proof. Fix $g \in G$. Since $\alpha(g)$ is an affine isometry, the map $\xi \mapsto \alpha(g)\xi - \alpha(g)0$ is a surjective linear isometry of $\mathcal{H}$, hence a unitary operator. Define

$$\pi(g)\xi := \alpha(g)\xi - \alpha(g)0 \quad \text{and} \quad b(g) := \alpha(g)0.$$

Then $\alpha(g)\xi = \pi(g)\xi + b(g)$ and uniqueness is clear.

The identity $\alpha(gh) = \alpha(g)\alpha(h)$ implies

$$\pi(gh)\xi + b(gh) = \alpha(gh)\xi = \alpha(g)(\pi(h)\xi + b(h)) = \pi(g)\pi(h)\xi + \pi(g)b(h) + b(g),$$

so $\pi(gh) = \pi(g)\pi(h)$ and $b(gh) = b(g) + \pi(g)b(h)$. This is exactly the cocycle identity. The converse direction is immediate from the cocycle identity. □

Lemma 7.3.9. *Let $\alpha(g)\xi = \pi(g)\xi + b(g)$ be an affine isometric action. Then α has a fixed point if and only if b is a coboundary.*

Proof. If ξ_0 is a fixed point, then $\pi(g)\xi_0 + b(g) = \xi_0$ for all g, i.e. $b(g) = \xi_0 - \pi(g)\xi_0$, so b is a coboundary. Conversely, if $b(g) = \xi_0 - \pi(g)\xi_0$, then $\alpha(g)\xi_0 = \xi_0$ for all g. □

Let $\pi : G \to \mathcal{U}(\mathcal{H})$ be unitary and let $b : G \to \mathcal{H}$ be a 1-cocycle. Set $\psi : G \to \mathbb{R}$ by $\psi(g) := \|b(g)\|^2$.

Lemma 7.3.10. *With notation as above, the function ψ is conditionally negative definite (see Appendix A.1).*

Proof. Note that $b(e) = 0$ by the cocycle identity, hence $\psi(e) = 0$. Also $b(g^{-1}) = -\pi(g^{-1})b(g)$, so $\psi(g^{-1}) = \|b(g^{-1})\|^2 = \|b(g)\|^2 = \psi(g)$.

Using the cocycle identity and unitarity of π, one checks that for all $g, h \in G$,

$$\|b(g^{-1}h)\|^2 = \|b(g) - b(h)\|^2.$$

Therefore, if $\sum_i c_i = 0$ then

$$\begin{aligned}\sum_{i,j} c_i c_j\, \psi(g_i^{-1} g_j) &= \sum_{i,j} c_i c_j\, \|b(g_i) - b(g_j)\|^2 \\ &= -2\Big\| \sum_i c_i\, b(g_i) \Big\|^2 \le 0.\end{aligned}$$

This is exactly the defining inequality for conditional negative definiteness. □

Theorem 7.3.11 (Delorme–Guichardet)**.** *If a countable group G has property* (T)*, then every affine isometric action of G on a Hilbert space has a fixed point.*

Proof. We prove the contrapositive.

Assume there exists an affine isometric action α of G on $\mathcal{H}$ without a fixed point. Write $\alpha(g)\xi = \pi(g)\xi + b(g)$. By Lemma 7.3.9, b is not a coboundary. In particular the orbit $\{b(g) : g \in G\} = \{\alpha(g)0 : g \in G\}$ is unbounded (otherwise the orbit would be bounded and hence have a circumcenter fixed by α). Set $\psi(g) := \|b(g)\|^2$, which is unbounded. By Lemma 7.3.10, ψ is conditionally negative definite. By Theorem A.1.6 in Appendix A.1, for each $t > 0$ the function $\varphi_t(g) := \exp(-t\psi(g))$ is positive definite. Let $(\pi_t, \mathcal{H}_t, \xi_t)$ be the corresponding GNS construction (Appendix A.1, Theorem A.1.2), so $\|\xi_t\| = 1$ and

$$\langle \pi_t(g)\xi_t, \xi_t\rangle = \varphi_t(g).$$

For any finite $F \subseteq G$ we have $\max_{g\in F}(1 - \Re\varphi_t(g)) \to 0$ as $t \downarrow 0$, hence

$$\max_{g\in F} \|\pi_t(g)\xi_t - \xi_t\|^2 = 2\max_{g\in F}(1 - \Re\varphi_t(g)) \longrightarrow 0 \qquad (t \downarrow 0),$$

so π_t has almost invariant vectors. On the other hand, since ψ is unbounded there exists a sequence g_n with $\varphi_t(g_n) \to 0$. If π_t had a non-zero invariant vector, then the invariant subspace would be non-zero; since ξ_t is cyclic, its orthogonal projection onto the invariant subspace would be non-zero, and hence $\varphi_t(g) = \langle \pi_t(g)\xi_t, \xi_t\rangle$ would be bounded below by a positive constant for all g, contradicting $\varphi_t(g_n) \to 0$. Therefore π_t has no non-zero invariant vector. Therefore G cannot have property (T). □

7.3.4 Non-trivial reduced cohomology

Here, we prove that failure of (T) yields a unitary representation with non-zero reduced first cohomology.

Theorem 7.3.12. *Let G be finitely generated and not have property (T). Then there exists a unitary representation π with $H^1_{\mathrm{red}}(G, \pi) \neq 0$.*

Proof. We follow the argument of Ozawa, see [40, Appendix A]. Fix a non-degenerate finitely supported symmetric probability measure μ on G with $\mu(e) > 0$. Since G does not have property (T), there exists a unitary representation $\pi : G \to \mathcal{U}(\mathcal{H})$ with almost invariant vectors but no non-zero invariant vectors. If (ξ_n) is almost invariant, then

$$\sum_{x\in G} \mu(x)\, \|\pi(x)\xi_n - \xi_n\| \longrightarrow 0.$$

The absence of invariant vectors implies the existence of a sequence $\varepsilon_n \downarrow 0$ such that

$$\sum_{x\in G} \mu(x)\, \|\xi_n - \pi(x)\xi_n\|^2 = 2\big(1 - \langle T\xi_n, \xi_n\rangle\big) \in (2\varepsilon_n, 4\varepsilon_n).$$

Fix a free ultrafilter $\mathcal{U}$ on $\mathbb{N}$ and form the ultrapower Hilbert space $\mathcal{H}^{\mathcal{U}}$ with the induced unitary representation $\pi^{\mathcal{U}}$. Define $b : G \to \mathcal{H}^{\mathcal{U}}$ by

$$b(x) := \big(\varepsilon_n^{-1/2}(\xi_n - \pi(x)\xi_n)\big)_{n\to\mathcal{U}}.$$

This is a 1-cocycle for $\pi^{\mathcal{U}}$ (the cocycle identity holds coordinatewise). Moreover,

$$\sum_{x\in G}\mu(x)\,\|b(x)\|^2 = \lim_{n\to\mathcal{U}}\varepsilon_n^{-1}\sum_{x\in G}\mu(x)\,\|\xi_n-\pi(x)\xi_n\|^2\in[2,4],$$

so b is non-zero. Finally,

$$\Big\|\sum_{x\in G}\mu(x)\,b(x)\Big\| = \lim_{n\to\mathcal{U}}\varepsilon_n^{-1/2}\,\|(1-T)\xi_n\|\le\lim_{n\to\mathcal{U}}\varepsilon_n^{-1/2}\cdot 2\varepsilon_n=0,$$

so b is μ-harmonic. By Proposition 7.3.6, a non-zero μ-harmonic cocycle defines a non-trivial class in reduced cohomology. Hence $[b]\neq 0$ in $\overline{H}^1(G,\pi^{\mathcal{U}})$. □

In fact we obtain the following characterization of property (T).

Corollary 7.3.13. *Let G be a countable group. The following are equivalent:*

(i) G has property (T).

(ii) $H^1(G,\pi)=0$ for every unitary representation π of G.

(iii) $H^1_{\mathrm{red}}(G,\pi)=0$ for every unitary representation π of G.

Proof. We first prove (1)⇒(2). Let $\pi: G\to\mathcal{U}(\mathcal{H})$ be unitary and let $b\in Z^1(G,\pi)$. Consider the associated affine isometric action α of G on $\mathcal{H}$ given by

$$\alpha(g)\xi := \pi(g)\xi + b(g)\qquad (g\in G,\ \xi\in\mathcal{H}).$$

Then α has a fixed point ξ_0 if and only if b is a coboundary, since $\alpha(g)\xi_0=\xi_0$ for all g is equivalent to $b(g)=\xi_0-\pi(g)\xi_0$.

Now assume that G has property (T). By Theorem 7.3.11, the action α has a fixed point. Hence $b\in B^1(G,\pi)$, and therefore $H^1(G,\pi)=Z^1(G,\pi)/B^1(G,\pi)=0$.

The implication (2)⇒(3) is immediate since $B^1(G,\pi)\subseteq\overline{B^1}(G,\pi)$. Finally, the implication (3)⇒(1) follows from Theorem 7.3.12. □

7.4 Ozawa's proof of Gromov's theorem

7.4.1 A spectral averaging lemma for harmonic cocycles

Definition 7.4.1. A unitary representation π is *weakly mixing* if it contains no non-zero finite-dimensional invariant subspace (equivalently, $\pi\otimes\overline{\pi}$ has no non-zero invariant vector).

Lemma 7.4.2. *Let $\pi: G\to\mathcal{U}(\mathcal{H})$ be weakly mixing, and let $b: G\to\mathcal{H}$ be a μ-harmonic cocycle. Then*

$$\frac{1}{n}\left\|\sum_{x\in G}\mu^{*n}(x)\,\big(b(x)\otimes\overline{b(x)}\big)\right\|\xrightarrow[n\to\infty]{}0.$$

In particular, for every $\xi\in\mathcal{H}$,

$$\frac{1}{n}\sum_{x\in G}\mu^{*n}(x)\,|\langle b(x),\xi\rangle|^2\xrightarrow[n\to\infty]{}0.$$

Proof. Set $\rho := \pi \otimes \overline{\pi}$, a unitary representation of G on $\mathcal{H} \otimes \overline{\mathcal{H}}$. Consider the operator

$$T := \sum_{g \in G} \mu(g)\, \rho(g) \in B(\mathcal{H} \otimes \overline{\mathcal{H}}).$$

It is a contraction since it is a convex combination of unitaries.

Define

$$\zeta := \sum_{g \in G} \mu(g)\, \big(b(g) \otimes \overline{b(g)}\big) \qquad \text{and} \qquad S_n := \sum_{x \in G} \mu^{*n}(x)\, \big(b(x) \otimes \overline{b(x)}\big) \;\; (n \geq 1).$$

We claim that for every $n \geq 1$,

$$S_n = (1 + T + \cdots + T^{n-1})\zeta. \tag{7.1}$$

To prove this, we first show the recursion $S_{n+1} = S_n + T^n\zeta$. Using $\mu^{*(n+1)} = \mu^{*n} * \mu$ we compute

$$S_{n+1} = \sum_{x \in G} \sum_{g \in G} \mu^{*n}(x)\mu(g)\, \big(b(xg) \otimes \overline{b(xg)}\big).$$

By the cocycle identity $b(xg) = b(x) + \pi(x)b(g)$, hence

$$b(xg) \otimes \overline{b(xg)} = b(x) \otimes \overline{b(x)} + \rho(x)\big(b(g) \otimes \overline{b(g)}\big) + b(x) \otimes \overline{\pi(x)b(g)} + \pi(x)b(g) \otimes \overline{b(x)}.$$

Summing over g with weights $\mu(g)$, the two cross terms vanish because b is μ-harmonic: evaluating harmonicity at e gives $\sum_g \mu(g)b(g) = 0$, hence

$$\sum_g \mu(g)\, \overline{\pi(x)b(g)} = \overline{\pi(x) \sum_g \mu(g)b(g)} = 0 \quad \text{and} \quad \sum_g \mu(g)\, \pi(x)b(g) = \pi(x) \sum_g \mu(g)b(g) = 0.$$

Therefore

$$\begin{aligned} S_{n+1} &= \sum_x \mu^{*n}(x)\, \big(b(x) \otimes \overline{b(x)}\big) \; + \sum_x \mu^{*n}(x)\, \rho(x) \sum_g \mu(g)\, \big(b(g) \otimes \overline{b(g)}\big) \\ &= S_n + \sum_x \mu^{*n}(x)\, \rho(x)\, \zeta. \end{aligned}$$

Finally, since $T^n = \sum_x \mu^{*n}(x)\, \rho(x)$, we get $S_{n+1} = S_n + T^n\zeta$. Iterating from $S_1 = \zeta$ yields (7.1).

Now apply the von Neumann's mean ergodic theorem to the contraction T: the Cesàro averages $\frac{1}{n}\sum_{k=0}^{n-1} T^k$ converge in SOT to the orthogonal projection onto $\mathrm{Fix}(T) = \ker(I - T)$. We claim $\mathrm{Fix}(T) = \{0\}$. Indeed, if $T\eta = \eta$, then

$$\|\eta\| = \|T\eta\| \leq \sum_g \mu(g)\, \|\rho(g)\eta\| = \|\eta\|,$$

so equality holds in the triangle inequality. This forces $\rho(g)\eta$ to be independent of g for all g with $\mu(g) > 0$, hence $\rho(g)\eta = \eta$ for those g. Thus η is invariant under the subgroup generated by $\mathrm{supp}(\mu)$. Since π is weakly mixing, it has no non-zero invariant vectors on any non-trivial subgroup, and in particular $\rho = \pi \otimes \overline{\pi}$ has no non-zero invariant vectors; therefore $\eta = 0$.

Consequently,

$$\frac{1}{n}\|S_n\| = \left\| \frac{1}{n}\sum_{k=0}^{n-1} T^k \zeta \right\| \xrightarrow[n\to\infty]{} 0,$$

which proves the first assertion.

For the second assertion, note that

$$|\langle b(x), \xi\rangle|^2 = \langle b(x)\otimes \overline{b(x)},\, \xi\otimes\overline{\xi}\rangle,$$

hence

$$\frac{1}{n}\sum_x \mu^{*n}(x)\,|\langle b(x),\xi\rangle|^2 = \frac{1}{n}\langle S_n,\, \xi\otimes\overline{\xi}\rangle \le \frac{1}{n}\|S_n\|\,\|\xi\otimes\overline{\xi}\|.$$

Since $\|\xi\otimes\overline{\xi}\| = \|\xi\|^2$ and $\frac{1}{n}\|S_n\| \to 0$, the claim follows. □

7.4.2 Proof of Ozawa's theorem

Definition 7.4.3 (Shalom)**.** A group G has *property* H_{FD} if for every weakly mixing unitary representation π, one has $H^1_{\mathrm{red}}(G,\pi) = 0$.

Theorem 7.4.4 (Ozawa, [40])**.** *If G has slow entropy growth, then G has property H_{FD}.*

Proof. Fix μ witnessing slow entropy growth (Definition 6.3.9). Let $\pi : G \to \mathcal{U}(\mathcal{H})$ be weakly mixing and let $b \in Z^1(G,\pi)$ be a μ-harmonic cocycle. We claim that $b = 0$. By Proposition 7.3.6, this implies that $H^1_{\mathrm{red}}(G,\pi) = 0$ for every weakly mixing π, which is exactly property H_{FD}.

Set $H_n := H(\mu^{*n})$. By slow entropy growth there exist a constant $C_\mu > 0$ and a subsequence $n_k \to \infty$ such that

$$n_k\big(H_{n_k+1} - H_{n_k}\big) \le C_\mu \qquad \text{for all } k. \tag{7.2}$$

Fix $g \in \mathrm{supp}(\mu)$ and $\xi \in \mathcal{H}$. Since b is μ-harmonic, it is μ^{*n}-harmonic for every n, i.e.

$$\sum_{x\in G}\mu^{*n}(x)\,b(x) = b(e) = 0 \qquad \text{and} \qquad \sum_{x\in G}\mu^{*n}(x)\,b(g^{-1}x) = b(g^{-1}).$$

Reindexing the second identity (substitute $x \mapsto gx$) gives

$$\sum_{x\in G}\mu^{*n}(gx)\,b(x) = b(g^{-1}).$$

Subtracting the first identity and taking inner products with ξ yields

$$|\langle b(g^{-1}),\xi\rangle| \le \sum_{x\in G} |\langle b(x),\xi\rangle|\,\big|\mu^{*n}(gx) - \mu^{*n}(x)\big|. \tag{7.3}$$

Apply Lemma 6.2.11 with $p := \mu^{*n}$, $q := g^{-1}\mu^{*n}$ (so $q(x) = \mu^{*n}(gx)$), and $f(x) := |\langle b(x),\xi\rangle|$. We obtain

$$|\langle b(g^{-1}),\xi\rangle| \le \big(8\,\delta(\mu^{*n}, g^{-1}\mu^{*n})\big)^{1/2}\cdot\left(\sum_{x\in G}|\langle b(x),\xi\rangle|^2\big(\mu^{*n}(x)+\mu^{*n}(gx)\big)\right)^{1/2}. \tag{7.4}$$

The concavity defect is controlled by the entropy increment. By Lemma 6.2.12 applied to $\nu = \mu^{*n}$, we have

$$H_{n+1} - H_n = H(\mu * \mu^{*n}) - H(\mu^{*n}) \geq 2 \min\{\mu(e), \mu(g^{-1})\}\, \delta(\mu^{*n}, g^{-1}\mu^{*n}).$$

Hence

$$\delta(\mu^{*n}, g^{-1}\mu^{*n}) \leq \frac{H_{n+1} - H_n}{2 \min\{\mu(e), \mu(g^{-1})\}}. \tag{7.5}$$

For the second factor in (7.4) we use a cocycle estimate. Reindexing and using $b(g^{-1}x) = b(g^{-1}) + \pi(g^{-1})b(x)$ gives

$$\sum_{x \in G} \mu^{*n}(gx)\, |\langle b(x), \xi\rangle|^2 = \sum_{x \in G} \mu^{*n}(x)\, |\langle b(g^{-1}x), \xi\rangle|^2 = \sum_{x \in G} \mu^{*n}(x)\, |\langle b(g^{-1}) + \pi(g^{-1})b(x), \xi\rangle|^2.$$

Using $|a+b|^2 \leq 2|a|^2 + 2|b|^2$ and unitarity of π, we obtain

$$\sum_{x \in G} \mu^{*n}(gx)\, |\langle b(x), \xi\rangle|^2 \leq 2|\langle b(g^{-1}), \xi\rangle|^2 \;+\; 2\sum_{x \in G} \mu^{*n}(x)\, |\langle b(x), \pi(g)\xi\rangle|^2.$$

Therefore, if we set

$$A_n(\eta) := \sum_{x \in G} \mu^{*n}(x)\, |\langle b(x), \eta\rangle|^2 \qquad (\eta \in \mathcal{H}),$$

then

$$\sum_{x \in G} |\langle b(x), \xi\rangle|^2 \big(\mu^{*n}(x) + \mu^{*n}(gx)\big) \leq 2A_n(\xi) + 2A_n(\pi(g)\xi) + 2|\langle b(g^{-1}), \xi\rangle|^2. \tag{7.6}$$

By Lemma 7.4.2 we have $A_n(\xi) = o(n)$ and $A_n(\pi(g)\xi) = o(n)$. Now combine (7.4), (7.5), and (7.6), and apply the resulting estimate with $n = n_k$ from (7.2). After squaring and using $H_{n_k+1} - H_{n_k} \leq C_\mu / n_k$ we obtain

$$|\langle b(g^{-1}), \xi\rangle|^2 \leq \frac{C(g,\mu)^2}{n_k}\Big(o(n_k) + |\langle b(g^{-1}), \xi\rangle|^2\Big),$$

where $C(g,\mu)$ depends only on $\mu(e)$ and $\mu(g)$ (since μ is symmetric). Letting $k \to \infty$ gives $\langle b(g^{-1}), \xi\rangle = 0$. Since $\xi \in \mathcal{H}$ was arbitrary, we conclude $b(g^{-1}) = 0$, hence $b(g) = 0$. Because $\mathrm{supp}(\mu)$ generates G as a semigroup, the cocycle identity forces $b \equiv 0$ on all of G. □

7.4.3 Gromov's theorem and Shalom's strategy

This section develops a direct route to Gromov's theorem using analytic methods.

Theorem 7.4.5 (Gromov)**.** *Every finitely generated group of polynomial growth is virtually nilpotent.*

Our presentation follows Ozawa [40] and Shalom [49]. For background on harmonic-map methods in metric settings that motivate some of these techniques, see Korevaar–Schoen [31] and Mok [39]. See [22] for the original proof and [30] for a different approach.

The key inductive step is to decrease the growth degree. For this it suffices to produce a finite-index subgroup that surjects onto $\mathbb{Z}$: the kernel then has smaller growth degree, so induction and Lemma 4.2.12 imply virtual nilpotence. While the details of this reduction are spelled out in Section 7.4.4, most of this Chapter provides the tools and techniques to produce a virtual surjection onto $\mathbb{Z}$. The emphasis is on Ozawa's idea: slow entropy growth implies Shalom's property H_{FD}, which in the amenable setting forces a virtual surjection onto $\mathbb{Z}$.

Lemma 7.4.6. *If G is amenable and has property H_{FD}, then admits a non-zero finite-dimensional unitary representation π with $H^1_{\mathrm{red}}(G,\pi) \neq 0$.*

Proof. Since G is infinite amenable, it does not have property (T) (Lemma 7.3.2). By Theorem 7.3.12, there exists a unitary representation σ of G such that $H^1_{\mathrm{red}}(G,\sigma) \neq 0$. By property H_{FD}, we know that σ has a non-zero finite-dimensional subrepresentation.

Let $\mathcal{H}$ be the Hilbert space of σ, and let $\mathcal{H}_{\mathrm{fd}} \subseteq \mathcal{H}$ be the (closed) linear span of all finite-dimensional G-invariant subspaces of $\mathcal{H}$. Then $\mathcal{H}_{\mathrm{fd}}$ is G-invariant and so is its orthogonal complement $\mathcal{H}_{\mathrm{wm}} := \mathcal{H}_{\mathrm{fd}}^{\perp}$. Accordingly, σ splits as an orthogonal direct sum $\sigma \cong \sigma_{\mathrm{fd}} \oplus \sigma_{\mathrm{wm}}$, where σ_{fd} is a possibly infinite Hilbert direct sum of finite-dimensional representations and σ_{wm} is weakly mixing. Since reduced cohomology is compatible with orthogonal direct sums, one has $H^1_{\mathrm{red}}(G,\sigma_{\mathrm{fd}}) \neq 0$.

Finally, write $\sigma_{\mathrm{fd}} \cong \bigoplus_{j\in J} \pi_j$ as a Hilbert direct sum of finite-dimensional unitary representations (e.g. a sum of irreducibles). Using again compatibility of H^1_{red} with Hilbert direct sums, we obtain

$$H^1_{\mathrm{red}}(G,\sigma_{\mathrm{fd}}) \;\cong\; \bigoplus_{j\in J} H^1_{\mathrm{red}}(G,\pi_j).$$

Since the left-hand side is non-zero, there exists some $j \in J$ with $H^1_{\mathrm{red}}(G,\pi_j) \neq 0$. Taking $\pi := \pi_j$ yields the desired finite-dimensional unitary representation. □

Theorem 7.4.7 (Gromov)**.** *Let G be finitely generated, infinite and of polynomial growth. Then G has a finite index subgroup with infinite abelianization.*

Proof. We follow Shalom's strategy [49] combined with Ozawa's result (Theorem 7.4.4). Combine:

$$\begin{aligned} \text{polynomial growth} &\Rightarrow \text{slow entropy growth (Lemma 6.3.10)} \\ &\Rightarrow H_{\mathrm{FD}} \text{ (Theorem 7.4.4)} \end{aligned}$$

By Lemma 7.4.6, there exists a finite-dimensional unitary $\pi : G \to U(n)$ with $H^1_{\mathrm{red}}(G,\pi)$ non-zero. If the image $\pi(G)$ were finite, then $\ker(\pi)$ would have finite index in G. But then every non-trivial cocycle on G would restrict to a homomorphism on the finite-index subgroup $\ker(\pi)$, which must be non-trival. Hence, $\ker(\pi)$ has infinite abelianization. Thus, we are left with the case that $\pi(G)$ is infinite.

Since G has polynomial growth, the same holds for $\pi(G)$. By Lemma 3.1.11, $\pi(G)$ is infinite and virtually abelian. This finishes the proof. □

7.4.4 Finishing the proof of Gromov's theorem

This section isolates the remaining steps once one knows that a finite index subgroup surjects onto $\mathbb{Z}$.

Lemma 7.4.8. *Let G be a finitely generated group and let $\varphi : G \twoheadrightarrow \mathbb{Z}$ be a surjective homomorphism. If G has polynomial growth of degree at most d, then $H := \ker(\varphi)$ is finitely generated and has polynomial growth of degree at most $d-1$.*

Proof. By the Lemma 4.2.11, H is finitely generated. Fix a finite symmetric generating set S of H and pick $t \in G$ with $\varphi(t) = 1$. Set $\Sigma := S \cup \{t, t^{-1}\}$. There exist constants $C \geq 1$ and $d \geq 1$ such that with respect to the word metric on G induced by Σ one has

$$|B_G(R)| \leq C\,R^d \qquad \text{for all } R \geq 1,$$

where $B_G(R) := \{g \in G : |g| \leq R\}$.

For any integer n and any $h \in H$ we have

$$|t^n h| \leq |t^n| + |h| \leq |n| + |h|.$$

In particular, if $|n| \leq \frac{R}{2}$ and $h \in H \cap B_G\big(\frac{R}{2}\big)$, then $|t^n h| \leq R$.

Now note that for distinct integers $m \neq n$ the left cosets $t^m H$ and $t^n H$ are disjoint, since $t^{m-n} \in H$ would imply $0 = \varphi(t^{m-n}) = m - n$. Hence the sets

$$t^n\big(H \cap B_G(R/2)\big) \subseteq B_G(R) \qquad \big(|n| \leq \lfloor \tfrac{R}{2} \rfloor\big)$$

are pairwise disjoint. Therefore

$$|B_G(R)| \;\geq \sum_{|n| \leq \lfloor R/2 \rfloor} \Big| t^n\big(H \cap B_G(R/2)\big)\Big| = \big(2\lfloor R/2 \rfloor + 1\big)\,\big|H \cap B_G(R/2)\big|.$$

For $R \geq 2$ we have $2\lfloor R/2 \rfloor + 1 \geq R/2$, so the growth bound on G gives

$$\big|H \cap B_G(R/2)\big| \leq \frac{|B_G(R)|}{R/2} \leq (2C)\,R^{d-1}.$$

Finally, for $r \geq 1$ apply this with $R = 2r$ to obtain

$$|H \cap B_G(r)| \leq (2C)(2r)^{d-1}.$$

If $B_H(r) := \{h \in H : |h|_S \leq r\}$, then $B_H(r) \subseteq H \cap B_G(r)$ since $|h| \leq |h|_S$ for $h \in H$. Thus $|B_H(r)| \leq (2C)(2r)^{d-1}$ for all $r \geq 1$, proving that H has polynomial growth of degree at most $d - 1$. □

Proof of Theorem 7.4.5. By Theorem 7.4.7, G has a finite index subgroup G_1 admitting a surjection onto $\mathbb{Z}$. Let $\varphi : G_1 \to \mathbb{Z}$ be such a surjection and write $H := \ker(\varphi)$. Assume by induction that all groups of polynomial growth of strictly smaller degree are virtually nilpotent. By Lemma 7.4.8, H has polynomial growth of strictly smaller degree, hence H is virtually nilpotent. Therefore G_1 is an extension of $\mathbb{Z}$ by a virtually nilpotent group. By Theorem 4.2.8 or more directly Lemma 4.2.12, polynomial growth then forces G_1 to be virtually nilpotent. Since G_1 has finite index in G, this finishes the proof. □

Chapter 8

Sofic and hyperlinear groups

8.1 Definition and Examples

The purpose of this chapter is to single out two approximation paradigms for countable groups. Very roughly, a group is called *sofic* if its finite pieces can be modelled inside finite symmetric groups, and *hyperlinear* if they can be modelled inside finite-dimensional unitary groups. We phrase this using the standard bi-invariant metrics on these families. For background and original sources, see [23, 53, 13].

For $n \geq 1$ write $[n] := \{1, \dots, n\}$. On the symmetric group $\mathrm{Sym}(n)$ we use the normalized Hamming distance

$$d_{\mathrm{H}}(\sigma, \tau) := \frac{1}{n} \#\{i \in [n] : \sigma(i) \neq \tau(i)\} \qquad (\sigma, \tau \in \mathrm{Sym}(n)).$$

On the unitary group

$$\mathcal{U}(n) := \{u \in M_n(\mathbb{C}) : u^* u = u u^* = \mathbb{1}\}$$

we use the normalized Hilbert–Schmidt norm

$$\|A\|_2 := \Big(\frac{1}{n} \sum_{i,j=1}^{n} |a_{ij}|^2 \Big)^{1/2} \qquad \big(A = (a_{ij}) \in M_n(\mathbb{C})\big)$$

and the associated bounded metric

$$D_2(u, v) := \tfrac{1}{2} \|u - v\|_2 \in [0, 1].$$

Definition 8.1.1. Let G be a group, let $F \subseteq G$ be finite and let $\varepsilon > 0$.

(i) A map $\phi : G \to \mathrm{Sym}(n)$ is an (F, ε)-*sofic approximation* if $\phi(e) = \mathrm{id}$ and

$$d_{\mathrm{H}}\big(\phi(gh), \phi(g)\phi(h)\big) < \varepsilon \qquad (g, h \in F)$$

and

$$d_{\mathrm{H}}\big(\phi(g), \mathrm{id}\big) > 1 - \varepsilon \qquad (g \in F \setminus \{e\}).$$

(ii) A map $\psi : G \to \mathcal{U}(n)$ is an (F, ε)-*hyperlinear approximation* if $\psi(e) = \mathbb{1}$ and

$$\|\psi(gh) - \psi(g)\psi(h)\|_2 < \varepsilon \qquad (g, h \in F)$$

and

$$\|\psi(g) - \mathbb{1}\|_2 > 1 - \varepsilon \qquad (g \in F \setminus \{e\}).$$

Definition 8.1.2. Let G be a countable group.

(i) G is *sofic* if for every finite $F \subseteq G$ and every $\varepsilon > 0$ there exist $n \geq 1$ and an (F, ε)-sofic approximation $\phi : G \to \mathrm{Sym}(n)$.

(ii) G is *hyperlinear* if for every finite $F \subseteq G$ and every $\varepsilon > 0$ there exist $n \geq 1$ and an (F, ε)-hyperlinear approximation $\psi : G \to \mathcal{U}(n)$.

Remark 8.1.3. The separation conditions above are only meant to prevent the trivial collapse of non-identity elements to the identity. Equivalent formulations are common in the literature, for example with a uniform lower bound on the distance to the identity, or via a normalized trace condition.

See [13].

Lemma 8.1.4. *Every sofic group is hyperlinear.*

Proof. For each n let $\iota_n : \mathrm{Sym}(n) \hookrightarrow \mathcal{U}(n)$ be the embedding sending a permutation to its permutation matrix. For $\sigma, \tau \in \mathrm{Sym}(n)$ one checks

$$\|\iota_n(\sigma) - \iota_n(\tau)\|_2^2 = 2\, d_{\mathrm{H}}(\sigma, \tau), \qquad \|\iota_n(\sigma) - \mathbb{1}\|_2^2 = 2\, d_{\mathrm{H}}(\sigma, \mathrm{id}).$$

Let G be sofic and fix finite $F \subseteq G$ and $\varepsilon > 0$. Set $\delta := \min\{\varepsilon^2/2,\ 1/2\}$ and choose an (F, δ)-sofic approximation $\phi : G \to \mathrm{Sym}(n)$. Then $\psi := \iota_n \circ \phi : G \to \mathcal{U}(n)$ satisfies $\psi(e) = \mathbb{1}$ and for $g, h \in F$,

$$\|\psi(gh) - \psi(g)\psi(h)\|_2 = \|\iota_n(\phi(gh)) - \iota_n(\phi(g)\phi(h))\|_2 = \sqrt{2\, d_{\mathrm{H}}(\phi(gh), \phi(g)\phi(h))} < \varepsilon.$$

For $g \in F \setminus \{e\}$,

$$\|\psi(g) - \mathbb{1}\|_2 = \sqrt{2\, d_{\mathrm{H}}(\phi(g), \mathrm{id})} \geq \sqrt{2(1-\delta)} \geq 1 > 1 - \varepsilon.$$

So ψ is an (F, ε)-hyperlinear approximation. □

Lemma 8.1.5. *(i) Finite groups are sofic and hyperlinear.*

(ii) Residually finite groups are sofic (hence hyperlinear).

(iii) Amenable groups are sofic (hence hyperlinear); in particular, abelian and solvable groups are sofic.

Proof. We prove soficity in each case; hyperlinearity then follows from Lemma 8.1.4.

(1) If G is finite, the left-regular action gives an injective homomorphism $\lambda : G \to \mathrm{Sym}(G)$. For any finite $F \subseteq G$ and $\varepsilon > 0$, set $n = |G|$ and let $\phi = \lambda$. Then $\phi(gh) = \phi(g)\phi(h)$ for all $g, h \in F$, and for $g \neq e$ the permutation $\phi(g)$ has no fixed points, so $d_{\mathrm{H}}(\phi(g), \mathrm{id}) = 1 > 1 - \varepsilon$.

(2) Let G be residually finite. Given finite $F \subseteq G$ and $\varepsilon > 0$, choose a finite-index normal subgroup $N \lhd G$ with $N \cap (F \setminus \{e\}) = \emptyset$. Let $q : G \to Q := G/N$ be the quotient map and let $\rho : Q \to \mathrm{Sym}(Q)$ be the left-regular action. Then $\phi := \rho \circ q$ is a homomorphism, and for $g \in F \setminus \{e\}$ we have $q(g) \neq e$, hence $\phi(g)$ has no fixed points and $d_{\mathrm{H}}(\phi(g), \mathrm{id}) = 1$.

(3) Let G be amenable. Fix finite $F \subseteq G$ and $\varepsilon > 0$. Choose a finite set $A \subseteq G$ such that $|gA \triangle A| < \frac{\varepsilon}{3}|A|$ for all $g \in F \cup F^{-1} \cup F^2$ (a Følner set). For each $g \in F$ define a permutation $\phi(g) \in \mathrm{Sym}(A)$ by setting $\phi(g)(a) = ga$ whenever $a \in A$ and $ga \in A$,

and extending arbitrarily to a bijection of A. The Følner condition ensures that for each $g \in F$ the set of points where $a \mapsto ga$ exits A has size $< \frac{\varepsilon}{3}|A|$, so $\phi(g)$ moves at least $(1 - \frac{\varepsilon}{3})|A|$ points and hence $d_{\mathrm{H}}(\phi(g), \mathrm{id}) > 1 - \varepsilon$. Similarly, for $g, h \in F$ the equality $\phi(gh) = \phi(g)\phi(h)$ can fail only if at least one of a, ha, gha lies outside A. Therefore it fails on at most

$$|A \triangle h^{-1}A| + |A \triangle g^{-1}A| + |A \triangle (gh)^{-1}A| \; < \; \varepsilon|A|,$$

so $d_{\mathrm{H}}(\phi(gh), \phi(g)\phi(h)) < \varepsilon$. Thus G is sofic. □

Remark 8.1.6. Two central open problems ask whether every countable group is sofic (a conjecture due to Gromov [23, 53]) and whether every countable group is hyperlinear (Connes' embedding conjecture for groups [7]). See Pestov's survey [44, Open Questions 3.8 and 3.9].

8.1.1 Metric ultraproducts

Metric ultraproducts provide a convenient way to package asymptotic approximation data.

Let (K_n, d_n) be a sequence of groups equipped with uniformly bounded bi-invariant metrics d_n taking values in $[0, 1]$. Fix a non-principal ultrafilter $\mathcal{U}$ on $\mathbb{N}$. For a bounded sequence of real numbers (a_n) we write

$$\lim_{\mathcal{U}} a_n = L$$

if for every $\varepsilon > 0$ the set $\{n : |a_n - L| < \varepsilon\}$ belongs to $\mathcal{U}$.

Lemma. Let (K_n, d_n) be groups equipped with uniformly bounded bi-invariant metrics $d_n \leq 1$, and define

$$\mathcal{N}_{\mathcal{U}} := \Big\{ (k_n) \in \prod_n K_n : \lim_{\mathcal{U}} d_n(k_n, e) = 0 \Big\}.$$

Then $\mathcal{N}_{\mathcal{U}}$ is a normal subgroup of $\prod_n K_n$.

Proof. It is immediate from the triangle inequality and left-invariance that $\mathcal{N}_{\mathcal{U}}$ is a subgroup. To see normality, let $k = (k_n) \in \mathcal{N}_{\mathcal{U}}$ and $h = (h_n) \in \prod_n K_n$. Since each d_n is bi-invariant,

$$d_n(h_n k_n h_n^{-1}, e) = d_n(k_n, e),$$

hence $\lim_{\mathcal{U}} d_n(h_n k_n h_n^{-1}, e) = 0$ and so $hkh^{-1} \in \mathcal{N}_{\mathcal{U}}$. □

Definition 8.1.7. The *metric ultraproduct* of the (K_n, d_n) with respect to $\mathcal{U}$ is the quotient group

$$\prod_{\mathcal{U}} (K_n, d_n) \; := \; \Big(\prod_{n \in \mathbb{N}} K_n \Big) \Big/ \mathcal{N}_{\mathcal{U}}.$$

It carries a natural bi-invariant metric defined by

$$D\big([(k_n)], [(\ell_n)]\big) := \lim_{\mathcal{U}} d_n(k_n, \ell_n).$$

Note that the formula defining D is indeed well-defined (independent of representatives) and turns $\prod_{\mathcal{U}} (K_n, d_n)$ into a metric group.

8.1.2 Ultraproduct characterizations

Lemma 8.1.8. *Let G be a group and let $\phi_n : G \to K_n$ be maps into groups with bi-invariant metrics $d_n \leq 1$. Define $\Phi : G \to \prod_{\mathcal{U}}(K_n, d_n)$ by $\Phi(g) := [(\phi_n(g))]_{\mathcal{U}}$. If for every $g, h \in G$ one has*

$$\lim_{\mathcal{U}} d_n\big(\phi_n(gh), \phi_n(g)\phi_n(h)\big) = 0,$$

then Φ is a homomorphism. Moreover, if for every $g \neq e$ one has $\lim_{\mathcal{U}} d_n\big(\phi_n(g), e\big) > 0$, then Φ is injective.

Proof. For fixed $g, h \in G$ we compute

$$D\big(\Phi(gh), \Phi(g)\Phi(h)\big) = \lim_{\mathcal{U}} d_n\big(\phi_n(gh), \phi_n(g)\phi_n(h)\big) = 0,$$

so $\Phi(gh) = \Phi(g)\Phi(h)$. If $g \neq e$ and $\Phi(g) = e$, then by definition of the quotient one has $\lim_{\mathcal{U}} d_n(\phi_n(g), e) = 0$, contradicting the assumption. □

Theorem 8.1.9. *Let G be a countable group. Then G is sofic if and only if there exist integers $n_k \geq 1$ and an injective homomorphism*

$$G \hookrightarrow \prod_{\mathcal{U}}\big(\mathrm{Sym}(n_k), d_{\mathrm{H}}\big)$$

for some non-principal ultrafilter $\mathcal{U}$.

Proof. (⇒) Fix an increasing sequence of finite subsets $F_1 \subseteq F_2 \subseteq \cdots \subseteq G$ with $\bigcup_k F_k = G$ and $e \in F_k$. Let $\varepsilon_k := 1/k$. By soficity, choose n_k and (F_k, ε_k)-sofic approximations $\phi_k : G \to \mathrm{Sym}(n_k)$. Define $\Phi(g) := [(\phi_k(g))]_{\mathcal{U}}$.

Fix $g, h \in G$. For all sufficiently large k we have $\{g, h, gh\} \subseteq F_k$, hence $d_{\mathrm{H}}(\phi_k(gh), \phi_k(g)\phi_k(h)) < \varepsilon_k$. Since $\varepsilon_k \to 0$, this implies $\lim_{\mathcal{U}} d_{\mathrm{H}}(\phi_k(gh), \phi_k(g)\phi_k(h)) = 0$. Lemma 8.1.8 shows that Φ is a homomorphism.

If $g \neq e$, then for all sufficiently large k we have $g \in F_k$, so $d_{\mathrm{H}}(\phi_k(g), \mathrm{id}) > 1 - \varepsilon_k$. Thus $\lim_{\mathcal{U}} d_{\mathrm{H}}(\phi_k(g), \mathrm{id}) = 1$, and injectivity follows from Lemma 8.1.8.

(⇐) Let $\iota : G \hookrightarrow \prod_{\mathcal{U}}(\mathrm{Sym}(n_k), d_{\mathrm{H}})$ be an injective homomorphism. Choose for each $g \in G$ a representing sequence $(\sigma_k(g)) \in \prod_k \mathrm{Sym}(n_k)$. For $g \neq e$ set $\delta(g) := D(\iota(g), e) > 0$.

Let $F \subseteq G$ be finite and let $\varepsilon > 0$. Set $\delta_F := \min\{\delta(g) : g \in F \setminus \{e\}\} > 0$. Choose $r \geq 1$ such that $(1 - \delta_F/2)^r < \varepsilon$.

For $m \geq 1$ define the homomorphism $\alpha_r : \mathrm{Sym}(m) \to \mathrm{Sym}(m^r)$ by letting $\alpha_r(\sigma)$ act diagonally on $[m]^r$. Then

$$d_{\mathrm{H}}\big(\alpha_r(\sigma), \alpha_r(\tau)\big) = 1 - \big(1 - d_{\mathrm{H}}(\sigma, \tau)\big)^r.$$

In particular, if $d_{\mathrm{H}}(\sigma, \mathrm{id}) \geq \delta_F/2$ then $d_{\mathrm{H}}(\alpha_r(\sigma), \mathrm{id}) \geq 1 - (1 - \delta_F/2)^r > 1 - \varepsilon$.

Since ι is a homomorphism, for each $g, h \in G$ one has $\lim_{\mathcal{U}} d_{\mathrm{H}}(\sigma_k(gh), \sigma_k(g)\sigma_k(h)) = 0$. Also $\lim_{\mathcal{U}} d_{\mathrm{H}}(\sigma_k(g), \mathrm{id}) = \delta(g)$. Hence for some index k we have simultaneously

$$d_{\mathrm{H}}(\sigma_k(gh), \sigma_k(g)\sigma_k(h)) < \varepsilon/r \qquad (g, h \in F),$$
$$d_{\mathrm{H}}(\sigma_k(g), \mathrm{id}) > \delta_F/2 \qquad (g \in F \setminus \{e\}).$$

Fix such a k and define $\phi := \alpha_r \circ \sigma_k : G \to \mathrm{Sym}(n_k^r)$. Then $\phi(e) = \mathrm{id}$ and for $g, h \in F$,

$$\begin{aligned} d_{\mathrm{H}}\big(\phi(gh), \phi(g)\phi(h)\big) &= 1 - \big(1 - d_{\mathrm{H}}(\sigma_k(gh), \sigma_k(g)\sigma_k(h))\big)^r \\ &\leq r\, d_{\mathrm{H}}(\sigma_k(gh), \sigma_k(g)\sigma_k(h)) < \varepsilon, \end{aligned}$$

using $1-(1-x)^r \le rx$ for $x \in [0,1]$. The separation estimate shows $d_{\mathrm{H}}(\phi(g), \mathrm{id}) > 1-\varepsilon$ for $g \in F \setminus \{e\}$. Thus ϕ is an (F,ε)-sofic approximation. □

Theorem 8.1.10. *Let G be a countable group. Then G is hyperlinear if and only if there exist integers $n_k \ge 1$ and an injective homomorphism*

$$G \hookrightarrow \prod_{\mathcal{U}} \big(\mathcal{U}(n_k), D_2\big)$$

for some non-principal ultrafilter $\mathcal{U}$.

Proof. The proof is analogous to Theorem 8.1.9, using the maps given by hyperlinear approximations and Lemma 8.1.8.

For the direction $\Leftarrow$, start from an injective homomorphism ι and representatives $u_k(g) \in \mathcal{U}(n_k)$. For a finite set F define $\delta_F := \min_{g \in F\setminus\{e\}} D(\iota(g), \mathbb{1}) > 0$. Choose r such that

$$\sqrt{2 - 2(1-\delta_F^2/8)^r} > 1-\varepsilon.$$

Consider the tensor-power homomorphism $\beta_r : \mathcal{U}(m) \to \mathcal{U}(m^r)$, $\beta_r(u) = u^{\otimes r}$. If $\|u - \mathbb{1}\|_2 \ge \delta_F/2$, then

$$\|u^{\otimes r} - \mathbb{1}\|_2^2 = 2 - 2\Re\big(\tau_m(u)^r\big) \ge 2 - 2\big(\Re\tau_m(u)\big)^r \ge 2 - 2\big(1 - \tfrac{1}{8}\delta_F^2\big)^r,$$

where $\tau_m(u) := \frac{1}{m}\operatorname{Tr}(u)$ and we used $\Re\tau_m(u) = 1 - \frac{1}{2}\|u-\mathbb{1}\|_2^2$. This yields the required separation. Approximate multiplicativity on F is obtained as in the sofic case, using that β_r is a homomorphism and the estimate $\|u^{\otimes r} - v^{\otimes r}\|_2 \le r\,\|u-v\|_2$. □

8.2 Equations over groups

Let G be a group and let $\langle x\rangle$ be the free group on one generator. An *equation over G in one variable* is an element $w \in G * \langle x\rangle$. We say that w is *solvable over G* if there exists an overgroup $H \ge G$ and an element $h \in H$ such that

$$w(h) = e \quad \text{in } H.$$

If G is finite and one can choose H finite, we say that w is solvable in a *finite extension* of G.

There is a canonical homomorphism

$$\varepsilon : G * \langle x\rangle \longrightarrow \mathbb{Z}$$

defined by sending G to 0 and x to 1. An equation $w \in G * \langle x\rangle$ is called *nonsingular* if $\varepsilon(w) \ne 0$. Equivalently, $\varepsilon(w)$ is the exponent sum of x in w, and nonsingularity means that this sum is non-zero.

Conjecture 8.2.1 (Kervaire–Laudenbach)**.** *Let G be a group and let $w \in G * \langle x\rangle$ be a nonsingular equation in one variable. Then w is solvable over G. Moreover, if G is finite then w is solvable in a finite extension of G.*

In classical work, Gerstenhaber and Rothaus proved that for finite groups G every nonsingular equation in one variable can be solved in a finite extension of G.

8.2.1 Degree of the power map on $U(n)$

For an integer m define the continuous map

$$p_m : U(n) \longrightarrow U(n), \qquad u \longmapsto u^m.$$

Since $U(n)$ is a compact connected Lie group, it is a closed connected smooth manifold of real dimension n^2 and comes with a canonical orientation (e.g. the one induced by any left-invariant volume form). For such an oriented closed manifold M, the degree of a continuous map $f : M \to M$ is the unique integer $\deg(f)$ such that

$$f^*([M]) = \deg(f)\,[M] \qquad \text{in } H^{\dim M}(M;\mathbb{Z}),$$

where $[M] \in H^{\dim M}(M;\mathbb{Z})$ is the cohomological fundamental class. Denote by $H^*(U(n);\mathbb{Z})$ the integral cohomology ring of $U(n)$.

Let $\mu : U(n)\times U(n) \to U(n)$ denote multiplication and $\mu^* : H^*(U(n);\mathbb{Z}) \to H^*(U(n);\mathbb{Z})\otimes H^*(U(n);\mathbb{Z})$ the induced coproduct.

Recall that a cohomology class $x \in H^*(U(n);\mathbb{Z})$ is called *primitive* if

$$\mu^*(x) = x \otimes 1 + 1 \otimes x.$$

Lemma 8.2.2 (Hopf)**.** *The integral cohomology of $U(n)$ is torsion-free and is an exterior algebra on odd-degree generators:*

$$H^*(U(n);\mathbb{Z}) \;\cong\; \Lambda\big(x_1, x_3, \ldots, x_{2n-1}\big), \qquad |x_{2k-1}| = 2k-1 \tag{8.1}$$

where the generators are primitive. In particular, the product $x_1\,x_3\cdots x_{2n-1}$ generates the group $H^{n^2}(U(n);\mathbb{Z}) \cong \mathbb{Z}$.

Idea of proof. There is a fibration $U(n-1) \hookrightarrow U(n) \twoheadrightarrow S^{2n-1}$ obtained by taking the first column of a unitary matrix. Assuming the statement for $U(n-1)$ and using the Serre spectral sequence for this fibration, one sees inductively that $H^*(U(n);\mathbb{Z})$ is torsion-free and that a new generator in degree $2n-1$ is created at each step. This yields (8.1); the statement about the top class follows because $\sum_{k=1}^{n}(2k-1) = n^2 = \dim U(n)$. □

Proposition 8.2.3. *For $m \in \mathbb{Z}$ the power map $p_m : U(n) \to U(n)$ has degree*

$$\deg(p_m) = m^n.$$

In particular, p_m has nonzero degree if and only if $m \neq 0$.

Proof. We define the top-degree class as:

$$\omega := x_1\,x_3\cdots x_{2n-1} \in H^{n^2}(U(n);\mathbb{Z}). \tag{8.2}$$

Write $p_m = \mu^{(m)} \circ \Delta^{(m)}$, where $\Delta^{(m)} : U(n) \to U(n)^m$ is the diagonal map and $\mu^{(m)} : U(n)^m \to U(n)$ is the m-fold multiplication. Iterating the identity for $\mu^*(x)$ shows that for a primitive class x one has

$$(\mu^{(m)})^*(x) = x \otimes 1 \otimes \cdots \otimes 1 + \cdots + 1 \otimes \cdots \otimes 1 \otimes x.$$

Pulling back along the diagonal yields $p_m^*(x) = m\,x$. In particular, $p_m^*(x_{2k-1}) = m\,x_{2k-1}$ for $1 \le k \le n$. Using (8.2) and multiplicativity of p_m^*,

$$p_m^*(\omega) = \prod_{k=1}^{n} p_m^*(x_{2k-1}) = \prod_{k=1}^{n}(m\,x_{2k-1}) = m^n\,\omega.$$

Since ω generates $H^{n^2}(U(n);\mathbb{Z}) \cong \mathbb{Z}$, this shows $\deg(p_m) = m^n$. □

8.2.2 Gerstenhaber–Rothaus for one variable

Let $w \in U(n) * \langle x \rangle$ be a one-variable word with coefficients in $U(n)$. Write it as

$$w(x) = u_0 x^{\varepsilon_1} u_1 x^{\varepsilon_2} \cdots u_{\ell-1} x^{\varepsilon_\ell} u_\ell, \qquad u_j \in U(n),\ \varepsilon_j \in \{\pm 1\}.$$

Its exponent sum is $m := \sum_{j=1}^{\ell} \varepsilon_j \in \mathbb{Z}$. The associated *word map* is

$$f_w : U(n) \to U(n), \qquad f_w(x) = w(x).$$

Lemma 8.2.4. *The word map f_w is homotopic to the power map p_m.*

Proof. Since $U(n)$ is path-connected, for each coefficient u_j choose a continuous path $u_j(t)$ with $u_j(0) = u_j$ and $u_j(1) = \mathbb{1}$. Define a homotopy $H : [0,1] \times U(n) \to U(n)$ by

$$H(t,x) := u_0(t) x^{\varepsilon_1} u_1(t) x^{\varepsilon_2} \cdots u_{\ell-1}(t) x^{\varepsilon_\ell} u_\ell(t).$$

Then $H(0,\cdot) = f_w$ and $H(1,x) = x^{\varepsilon_1 + \cdots + \varepsilon_\ell} = x^m = p_m(x)$. □

Theorem 8.2.5 (Gerstenhaber–Rothaus)**.** *If the exponent sum m is non-zero, then $f_w : U(n) \to U(n)$ is surjective. Equivalently, the equation $w(x) = \mathbb{1}$ has a solution already in $U(n)$.*

Proof. By Lemma 8.2.4, f_w is homotopic to p_m, so $\deg(f_w) = \deg(p_m)$. Proposition 8.2.3 gives $\deg(p_m) = m^n \neq 0$. Since $U(n)$ is a compact connected oriented manifold, any continuous self-map of nonzero degree is surjective. Thus f_w is surjective, and in particular $\mathbb{1} \in U(n)$ has a preimage. □

Corollary 8.2.6. *Let G be a finite group and let $w \in G * \langle x \rangle$ be a nonsingular equation in one variable. Then w is solvable in a finite extension of G.*

Proof. Let $m \neq 0$ be the exponent sum of x in w. Embed G into $U(|G|)$ via the left regular representation (as permutation matrices). Viewing the coefficients of w inside $U(|G|)$ yields a word map $f_w : U(|G|) \to U(|G|)$ with exponent sum m. By Theorem 8.2.5, f_w is surjective, so choose $u \in U(|G|)$ with $f_w(u) = \mathbb{1}$.

Let $H \leq U(|G|)$ be the subgroup generated by G and u. Then H is finitely generated and linear, hence residually finite. Since G is finite, there exists a finite quotient $\pi : H \twoheadrightarrow Q$ that is injective on G. The image $\pi(u)$ satisfies the equation in Q, and $\pi(G) \cong G$. Thus Q is a finite group containing a copy of G in which w has a solution. □

8.2.3 Algebraically closed groups and hyperlinear groups

Definition 8.2.7. A group G is called *algebraically closed* if for every $w \in G * \langle x \rangle$ with $\varepsilon(w) \neq 0$ there exists $g \in G$ such that $w(g) = e$ in G.

Corollary 8.2.8. *For every $n \geq 1$, the unitary group $U(n)$ is algebraically closed in the sense of Definition 8.2.7.*

Proof. Let $w \in U(n) * \langle x \rangle$ be nonsingular. Then its exponent sum m is non-zero and by Theorem 8.2.5, the equation $w(x) = \mathbb{1}$ has a solution in $U(n)$. □

Corollary 8.2.9. *Any metric ultraproduct $\prod_{\mathcal{U}} (U(n_k), D_2)$ is algebraically closed in the sense of Definition 8.2.7.*

Proof. Let $w \in \left(\prod_{\mathcal{U}}(U(n_k), D_2)\right) * \langle x \rangle$ be a nonsingular equation. Write its coefficients as ultraproduct classes represented by sequences in the $U(n_k)$. For each k, form the corresponding word $w_k \in U(n_k) * \langle x \rangle$ by choosing representatives for the coefficients. Then $\varepsilon(w_k) = \varepsilon(w) \neq 0$, hence w_k is nonsingular. By Corollary 8.2.8, choose $u_k \in U(n_k)$ with $w_k(u_k) = \mathbb{1}$. Let $u := [(u_k)]$ be the corresponding element of the metric ultraproduct. Then $w(u) = e$ in $\prod_{\mathcal{U}}(U(n_k), D_2)$. □

Theorem 8.2.10 (Pestov)**.** *Let G be a hyperlinear group. Then every nonsingular one-variable equation $w \in G * \langle x \rangle$ is solvable in a hyperlinear overgroup of G.*

Proof. By hyperlinearity, there exists an injective homomorphism

$$\iota : G \hookrightarrow \prod_{\mathcal{U}}(U(n_k), D_2).$$

View the coefficients of w inside the ultraproduct via ι, obtaining an equation over the ultraproduct. By Corollary 8.2.9, this ultraproduct is algebraically closed, so there exists an element u in the ultraproduct with $w(u) = e$. Thus w is solvable in $\langle G, u \rangle$, which is a subgroup of the ultraproduct and hence hyperlinear. □

8.3 Kaplansky's direct finiteness conjecture

A striking open problem in the theory of group algebras is the following conjecture of Kaplansky.

Conjecture 8.3.1 (Kaplansky)**.** *Let K be a field and let G be a group. Then the group algebra $K[G]$ is* directly finite*, i.e. for every $a, b \in K[G]$ the relation $ab = 1$ implies $ba = 1$.*

We record two cases in which the conjecture is known. The first is due to Kaplansky himself covering the case when the field has characteristic zero, and the second was proved by Elek and Szabó for sofic groups and arbitrary fields.

Lemma 8.3.2. *Let A be a unital C^*-algebra and let $e \in A$ satisfy $e^2 = e$. Then there exists an invertible element $v \in A$ such that vev^{-1} is a projection.*

Proof. Set

$$y := e^*e + (1 - e^*)(1 - e) \in A.$$

Then $y \geq 0$. To see that y is invertible, represent A faithfully on a Hilbert space $\mathcal{H}$. For $\xi \in \mathcal{H}$ we have

$$\langle y\xi, \xi \rangle = \|e\xi\|^2 + \|(1 - e)\xi\|^2.$$

Since $\xi = e\xi + (1 - e)\xi$, the inequality $\|u + v\|^2 \leq 2(\|u\|^2 + \|v\|^2)$ gives

$$\|\xi\|^2 \leq 2\big(\|e\xi\|^2 + \|(1 - e)\xi\|^2\big) = 2\langle y\xi, \xi \rangle,$$

so $y \geq \frac{1}{2}\mathbb{1}$ and hence y is invertible in A. Next observe that $ye = e^*e = e^*y$. Indeed, $(1 - e)e = 0$ shows $ye = e^*ee + (1 - e^*)(1 - e)e = e^*e$. Also $e^*(1 - e^*) = 0$ shows $e^*y = e^*e^*e + e^*(1 - e^*)(1 - e) = e^*e$. Let $v := y^{1/2}$ and set $p := vev^{-1}$. Then

$$p^2 = (vev^{-1})(vev^{-1}) = ve(v^{-1}v)ev^{-1} = ve^2v^{-1} = vev^{-1} = p.$$

Moreover, from $v^2e = e^*v^2$ we get $vev^{-1} = v^{-1}e^*v$, hence

$$p^* = (v^{-1})^*e^*v^* = v^{-1}e^*v = vev^{-1} = p.$$

Thus p is a projection. □

Theorem 8.3.3. *Let K be a field of characteristic 0 and let G be a group. Then $K[G]$ is directly finite.*

Proof. If $a, b \in K[G]$ witness a failure of direct finiteness, then the coefficients of a and b lie in some finitely generated subfield $F \subseteq K$. It is well-known that every finitely generated field F of characteristic 0 admits an embedding $F \hookrightarrow \mathbb{C}$. Thus, it suffices to prove direct finiteness for $\mathbb{C}[G]$.

Now consider the left regular representation $\lambda : G \to \mathcal{U}(\ell^2(G))$ and extend it linearly to a $*$-homomorphism

$$\lambda : \mathbb{C}[G] \longrightarrow C_r^*(G) \subseteq \mathcal{B}(\ell^2(G)).$$

Let τ be the canonical trace on $C_r^*(G)$ from Definition 7.1.3. By Proposition 7.1.5, τ is tracial on $\mathbb{C}[G]$ and hence on $C_r^*(G)$ by continuity.

We also need that τ is positive and faithful (Lemma 7.1.4). Positivity follows immediately since τ is a vector state: if $a \geq 0$ then $a = b^*b$ and

$$\tau(a) = \langle b^*b\delta_e, \delta_e\rangle = \|b\delta_e\|_2^2 \geq 0.$$

For faithfulness, let $x \in C_r^*(G)$ satisfy $x \geq 0$ and $\tau(x) = 0$. Then for every $g \in G$ we have

$$\langle x\delta_g, \delta_g\rangle = \langle \lambda(g)^* x \lambda(g)\, \delta_e, \delta_e\rangle = \tau(\lambda(g)^* x \lambda(g)) = \tau(x) = 0.$$

Since $x \geq 0$, this implies $x^{1/2}\delta_g = 0$ for all g, so $x^{1/2} = 0$ and hence $x = 0$.

Suppose that $a, b \in \mathbb{C}[G]$ satisfy $ab = 1$. In $C_r^*(G)$ we have

$$(ba)^2 = b(ab)a = ba,$$

so $e := ba$ is an idempotent. By traciality,

$$\tau(e) = \tau(ba) = \tau(ab) = \tau(1) = 1.$$

By Lemma 8.3.2 there exists an invertible $v \in C_r^*(G)$ such that $p := vev^{-1}$ is a projection. Using traciality we get $\tau(p) = \tau(e) = 1$, hence $\tau(1-p) = 0$. Since $1 - p \geq 0$ and τ is faithful by Lemma 7.1.4, it follows that $1 - p = 0$, i.e. $p = 1$. Therefore $e = v^{-1}pv = 1$, so $ba = 1$. ◻

The aim of the rest of this section is to prove the following result, which gives a large class of groups for which Kaplansky's conjecture holds for arbitrary fields.

Theorem 8.3.4 (Elek–Szabó)**.** *Let K be a field and let G be a sofic group. Then $K[G]$ is directly finite.*

Proof. Fix a free ultrafilter $\mathcal{U}$ on $\mathbb{N}$. Choose a sofic approximation $\sigma_n : G \to \mathrm{Sym}(V_n)$ and let $d_n := |V_n|$. For each n, let $\pi_n : G \to \mathrm{GL}_{d_n}(K)$ be the corresponding permutation-matrix representation. Extending K-linearly gives ring homomorphisms $\pi_n : K[G] \to M_{d_n}(K)$. For $A \in M_{d_n}(K)$ write

$$\rho_n(A) := \frac{1}{d_n} \operatorname{rank}(A) \in [0, 1].$$

Consider the ideal

$$\mathcal{N}_{\mathcal{U}} := \Big\{(A_n) \in \prod_n M_{d_n}(K) : \lim_{\mathcal{U}} \rho_n(A_n) = 0\Big\}$$

and the corresponding *rank ultraproduct*

$$\prod_{\mathcal{U}}^{\mathrm{rk}} M_{d_n}(K) := \Big(\prod_n M_{d_n}(K)\Big)\Big/\mathcal{N}_{\mathcal{U}}.$$

Passing to the rank ultraproduct yields a homomorphism

$$\Pi : K[G] \longrightarrow \prod_{\mathcal{U}}^{\mathrm{rk}} M_{d_n}(K), \qquad a \longmapsto [(\pi_n(a))].$$

We include a proof that Π is injective. Let $0 \neq a \in K[G]$ and write $a = \sum_{g\in F_0} \alpha_g g$ with $F_0 \subseteq G$ finite and all $\alpha_g \neq 0$. Set $F := F_0 \cup \{e\}$. By the defining property of a sofic approximation, for every $\varepsilon > 0$ there is a set $I_\varepsilon \in \mathcal{U}$ such that for all $n \in I_\varepsilon$ the set

$$W_n := \big\{v \in V_n : \text{the points } \sigma_n(g)v \ (g \in F) \text{ are pairwise distinct}\big\}$$

has size $|W_n| \geq (1-\varepsilon)d_n$.

For such n and $v \in W_n$ define the block

$$\mathcal{O}_v := \{\sigma_n(g)v : g \in F\} \subseteq V_n.$$

Since $e \in F$, we have $v \in \mathcal{O}_v$ and $|\mathcal{O}_v| = |F|$. Choose a subset $W_n' \subseteq W_n$ maximal with the property that the blocks $\{\mathcal{O}_v : v \in W_n'\}$ are pairwise disjoint. Maximality implies $|W_n'| \geq |W_n|/|F| \geq (1-\varepsilon)d_n/|F|$.

For each $v \in V_n$ let $\delta_v \in K^{V_n}$ be the standard basis vector. Then

$$\pi_n(a)\,\delta_v = \sum_{g\in F_0} \alpha_g\,\delta_{\sigma_n(g)v}.$$

If $v \in W_n'$, this vector is supported on $\mathcal{O}_v$. Since the blocks $\mathcal{O}_v$ $(v \in W_n')$ are disjoint, the family

$$\{\pi_n(a)\,\delta_v : v \in W_n'\}$$

is linearly independent in K^{V_n}. Hence $\operatorname{rank}(\pi_n(a)) \geq |W_n'|$ and therefore

$$\rho_n(\pi_n(a)) \geq \frac{|W_n'|}{d_n} \geq \frac{1-\varepsilon}{|F|}.$$

Taking $\mathcal{U}$-limits and then letting $\varepsilon \downarrow 0$ gives $\lim_{\mathcal{U}} \rho_n(\pi_n(a)) > 0$, so $(\pi_n(a)) \notin \mathcal{N}_{\mathcal{U}}$. Thus $\Pi(a) \neq 0$, proving injectivity.

Each matrix algebra $M_{d_n}(K)$ is directly finite, and the same argument works in the rank ultraproduct. Indeed, suppose $[(A_n)]\,[(B_n)] = 1$ in $\prod_{\mathcal{U}}^{\mathrm{rk}} M_{d_n}(K)$. By definition this means that $\lim_{\mathcal{U}} \rho_n(1_{d_n} - A_nB_n) = 0$. For each n set $V_n := \ker(1_{d_n} - A_nB_n)$. Then $\dim V_n = d_n - \operatorname{rank}(1_{d_n} - A_nB_n)$, and A_nB_n is the identity on V_n. In particular, B_n is injective on V_n (if $B_nv = 0$ with $v \in V_n$, then $v = A_nB_nv = 0$). Moreover, for $v \in V_n$ we have $(1_{d_n} - B_nA_n)(B_nv) = 0$, so $B_n(V_n) \subseteq \ker(1_{d_n} - B_nA_n)$. Therefore

$$\dim\ker(1_{d_n} - B_nA_n) \geq \dim B_n(V_n) = \dim V_n$$

and hence

$$\operatorname{rank}(1_{d_n} - B_nA_n) \leq \operatorname{rank}(1_{d_n} - A_nB_n).$$

Dividing by d_n and taking $\mathcal{U}$-limits yields $\lim_{\mathcal{U}} \rho_n(1_{d_n} - B_nA_n) = 0$, i.e. $[(B_n)]\,[(A_n)] = 1$. Since $K[G]$ embeds into a directly finite ring, it is directly finite. □

Chapter 9

The first ℓ^2-Betti number

The first ℓ^2-Betti number $b_1^{(2)}(G)$ is an analytic invariant that measures, in a von Neumann dimension sense, the size of the space of square-summable 1-cocycles on G. Positivity of the first ℓ^2-Betti number is a strong negation of Kazhdan's property (T) and has many remarkable consequences. In order to keep the exposition self-contained, we restrict ourselves to finitely generated groups and give a definition of $b_1^{(2)}(G)$ in terms of the Cayley graph and its cycle space. We refer to Lück's book [35] for a more general treatment of ℓ^2-Betti numbers.

9.1 Definitions and basic properties

There are several equivalent ways to define the first ℓ^2-Betti number. For our purposes it is convenient to work directly with the Cayley graph and its cycle space. Before doing so, we review the notion of Dirichlet functions and harmonicity, and the von Neumann dimension for Hilbert G-modules.

9.1.1 Dirichlet functions and ℓ^2-harmonicity

Let G be a group generated by a finite set S_0. Set $S := S_0 \cup S_0^{-1}$. Write $\ell^2(G)$ for the Hilbert space of square-summable functions $G \to \mathbb{C}$. For a function $f : G \to \mathbb{C}$ and $s \in S$ define the discrete derivative

$$(\partial_s f)(g) := f(gs) - f(g) \qquad (g \in G).$$

The non-normalized Dirichlet energy of f with respect to S is

$$\mathcal{E}_S(f) := \sum_{s \in S} \|\partial_s f\|_2^2 \in [0, \infty].$$

We let $\mathcal{D}_S(G) := \{f : G \to \mathbb{C} : \mathcal{E}_S(f) < \infty\}$ be the space of finite-energy *Dirichlet functions*. Note that $\ell^2(G) \subseteq \mathcal{D}_S(G)$ and that constant functions lie in $\mathcal{D}_S(G)$ with energy 0. Define the non-normalized Laplacian by

$$(\Delta_S f)(g) := \sum_{s \in S} \big(f(g) - f(gs)\big).$$

We say that f is *S-harmonic* if $\Delta_S f = 0$. Equivalently, $f(g)$ is the average of the values of f at the neighbors of g in the Cayley graph $\mathrm{Cay}(G, S)$, i.e., f is μ_S-harmonic, where

μ_S is the simple random walk associated to S. Note however, that we do not require f to be bounded. We write

$$\mathcal{H}_S(G) := \{f \in \mathcal{D}_S(G) : \Delta_S f = 0\}$$

for the space of *harmonic Dirichlet functions*. Any bounded μ_S-harmonic function is a harmonic Dirichlet function, i.e. $\mathcal{H}(G, \mu_S) \subseteq \mathcal{H}_S(G)$, but the converse is not true in general.

9.1.2 von Neumann dimension for Hilbert G-modules

Let ρ be the *right* regular representation of G on $\ell^2(G)$,

$$(\rho(g)\xi)(x) := \xi(xg) \qquad (g, x \in G).$$

We write $\lambda : G \to \mathcal{U}(\ell^2(G))$ for the *left* regular representation,

$$(\lambda(g)\xi)(x) := \xi(g^{-1}x) \qquad (g, x \in G).$$

The *group von Neumann algebra* $L(G)$ is the weak operator closure of $\lambda(\mathbb{C}[G]) \subseteq \mathcal{B}(\ell^2(G))$. It is a finite von Neumann algebra with faithful normal trace

$$\tau(T) := \langle T\delta_e, \delta_e\rangle \qquad (T \in L(G)).$$

For $n \geq 1$ we extend τ to $M_n(L(G))$ by

$$\tau_n\big((T_{ij})\big) := \sum_{i=1}^{n} \tau(T_{ii}).$$

The commutant of $L(G)$ in $\mathcal{B}(\ell^2(G))$ is the right group von Neumann algebra $R(G)$ (generated by the right regular representation), and it carries the trace

$$\tau(S) := \langle S\delta_e, \delta_e\rangle \qquad (S \in R(G)).$$

By abuse of notation we write τ for both traces; the meaning is determined by whether the argument lies in $L(G)$ or in $R(G)$. For $n \geq 1$ we also extend this trace to $M_n(R(G))$ by the same formula

$$\tau_n\big((S_{ij})\big) := \sum_{i=1}^{n} \tau(S_{ii}).$$

Let $\mathcal{H} := \ell^2(G)^n$. A closed subspace $V \leq \mathcal{H}$ is called a *Hilbert G-submodule* if it is G-invariant for the left action (equivalently, $\lambda(g)V \subseteq V$ for all g). Then the orthogonal projection $P_V \in \mathcal{B}(\mathcal{H})$ commutes with $\lambda(G)$, hence $P_V \in M_n(R(G))$. We define its *von Neumann dimension* by

$$\dim_G(V) := \tau_n(P_V) = \sum_{i=1}^{n} \langle P_V(\delta_e \otimes e_i), \delta_e \otimes e_i\rangle.$$

We will use the following basic properties, which follow immediately from the corresponding identities for orthogonal projections.

Lemma 9.1.1. *Let $V, W \leq \ell^2(G)^n$ be Hilbert G-submodules.*

(a) $0 \le \dim_G(V) \le n$ *and* $\dim_G(\ell^2(G)^n) = n$.

(b) If $V \subseteq W$, *then* $\dim_G(V) \le \dim_G(W)$.

(c) If $V \perp W$, *then* $\dim_G(V \oplus W) = \dim_G(V) + \dim_G(W)$.

(d) $\dim_G(V^\perp) = n - \dim_G(V)$.

(e) If V *and* W *are isomorphic as Hilbert* G*-modules, then* $\dim_G(V) = \dim_G(W)$.

(f) One has $\dim_G(V) = 0$ *if and only if* $V = \{0\}$.

Proof. Let $P_V, P_W \in M_n(R(G))$ denote the orthogonal projections onto V and W.

(a) Since $0 \le P_V \le \mathbb{1}$ (operator order) and τ_n is positive, we have $0 \le \tau_n(P_V) \le \tau_n(\mathbb{1}) = n$. Also $\dim_G(\ell^2(G)^n) = \tau_n(\mathbb{1}) = n$.

(b) If $V \subseteq W$, then $P_W - P_V$ is the orthogonal projection onto $W \cap V^\perp$ and in particular is positive. Hence $\dim_G(W) - \dim_G(V) = \tau_n(P_W - P_V) \ge 0$.

(c) If $V \perp W$, then $P_{V\oplus W} = P_V + P_W$, so $\dim_G(V \oplus W) = \tau_n(P_V + P_W) = \tau_n(P_V) + \tau_n(P_W)$.

(d) One has $P_{V^\perp} = \mathbb{1} - P_V$, hence $\dim_G(V^\perp) = \tau_n(\mathbb{1} - P_V) = n - \dim_G(V)$.

(e) Let $U : V \to W$ be a G-equivariant unitary. Extend it to a bounded operator $v \in \mathcal{B}(\mathcal{H})$ by setting $v|_V = U$ and $v|_{V^\perp} = 0$. Then v commutes with $\lambda(G)$, hence $v \in M_n(R(G))$, and it is a partial isometry with $v^*v = P_V$ and $vv^* = P_W$. Since τ_n is tracial on $M_n(R(G))$, we obtain

$$\dim_G(W) = \tau_n(P_W) = \tau_n(vv^*) = \tau_n(v^*v) = \tau_n(P_V) = \dim_G(V).$$

(f) If $V = \{0\}$ then $P_V = 0$ and $\dim_G(V) = 0$. Conversely, if $\dim_G(V) = \tau_n(P_V) = 0$ then for each i we have $\langle P_V(\delta_e \otimes e_i), \delta_e \otimes e_i\rangle = 0$. Since P_V is a projection, this implies $P_V(\delta_e \otimes e_i) = 0$. Using $P_V\lambda(g) = \lambda(g)P_V$ we get $P_V(\delta_g \otimes e_i) = 0$ for all $g \in G$ and all i. Hence $P_V = 0$, so $V = \{0\}$. □

9.1.3 The Cayley graph complex and cycles

Let S be a finite generating set. With $E_{G,S}$ as above we identify

$$\ell^2(E_{G,S}) \cong \ell^2(G)^S, \qquad \omega \longleftrightarrow \big(\omega(g, gs)\big)_{s\in S}.$$

For $s \in S$ we write $\bar{s} := (1_G, s) \in E_{G,S}$. Define the boundary operator on the dense subspace $\mathbb{C}[E_{G,S}]$ by $\partial(g, gs) = gs - g$. It extends to a bounded operator

$$\partial_1 : \ell^2(E_{G,S}) \to \ell^2(G).$$

Its adjoint $\partial_1^* : \ell^2(G) \to \ell^2(E_{G,S})$ is the discrete derivative operator

$$(\partial_1^* f)(g, gs) = f(gs) - f(g).$$

In particular, for $f \in \ell^2(G)$ one has

$$\|\partial_1^* f\|_2^2 = \sum_{s\in S} \|\partial_s f\|_2^2.$$

The space of finite cycles $Z_{G,S} \subseteq \mathbb{C}[E_{G,S}]$ was defined above by $\partial z = 0$. We write

$$\mathcal{Z}_{G,S} := \overline{Z_{G,S}} \le \ell^2(E_{G,S})$$

for its closure. It is a Hilbert G-submodule and satisfies $\mathcal{Z}_{G,S} \subseteq \ker(\partial_1)$.

Definition 9.1.2. We define the *zeroth ℓ^2-Betti number* of a group G by

$$b_0^{(2)}(G) := \dim_G(\ell^2(G))^G \qquad \text{where } (\ell^2(G))^G := \{\xi \in \ell^2(G) : \lambda(g)\xi = \xi \ \forall g \in G\}.$$

Lemma 9.1.3. *Let G be a group. Then $b_0^{(2)}(G) = 0$ if G is infinite, and $b_0^{(2)}(G) = \frac{1}{|G|}$ if G is finite.*

Proof. If $\xi \in \ell^2(G)$ satisfies $\lambda(g)\xi = \xi$ for all $g \in G$, then ξ is constant on G. If G is infinite, the only constant vector in $\ell^2(G)$ is 0, hence $(\ell^2(G))^G = \{0\}$ and $b_0^{(2)}(G) = 0$.

Assume that G is finite. Then $(\ell^2(G))^G$ is the one-dimensional subspace spanned by the constant vector. The orthogonal projection P onto this subspace is given by

$$(P\eta)(x) = \frac{1}{|G|}\sum_{y\in G}\eta(y) \qquad (\eta \in \ell^2(G),\ x \in G).$$

In particular, P commutes with $\lambda(G)$, hence $P \in R(G)$, and

$$b_0^{(2)}(G) = \dim_G(\ell^2(G))^G = \tau(P) = \langle P\delta_e, \delta_e\rangle = \frac{1}{|G|}.$$

□

9.1.4 Harmonic Dirichlet functions and $b_1^{(2)}(G)$

Recall that $\mathcal{D}_S(G)$ is the space of functions $f : G \to \mathbb{C}$ of finite Dirichlet energy $\mathcal{E}_S(f) < \infty$. The discrete derivative map

$$\nabla : \mathcal{D}_S(G) \to \ell^2(E_{G,S}), \qquad (\nabla f)(g, gs) := f(gs) - f(g),$$

has kernel equal to the constant functions. Note that $\nabla = \oplus_s \partial_s$ and $\|\nabla f\|_2^2 = \mathcal{E}_S(f)$. It therefore induces an isometric embedding

$$\mathcal{D}_S(G)/\mathbb{C} \hookrightarrow \ell^2(E_{G,S}), \qquad [f] \mapsto \nabla f,$$

where we equip $\mathcal{D}_S(G)/\mathbb{C}$ with the norm $\|[f]\| := \mathcal{E}_S(f)^{1/2}$.

Lemma 9.1.4. *Let $\omega \in \ell^2(E_{G,S})$. If $\langle \omega, z\rangle = 0$ for every finite cycle $z \in Z_{G,S}$, then there exists $f \in \mathcal{D}_S(G)$ with $\nabla f = \omega$. Moreover f is unique modulo constants and $\mathcal{E}_S(f) = \|\omega\|_2^2$.*

Proof. Assume $\omega \perp Z_{G,S}$. Define $f(e) := 0$. Given $g \in G$, choose a word $g = s_1 \cdots s_k$ with $s_i \in S$ and set

$$f(g) := \sum_{i=1}^{k} \omega(s_1 \cdots s_{i-1},\ s_1 \cdots s_i).$$

If $g = t_1 \cdots t_m$ is another word, then the two edge-paths from e to g determine a finite cycle $z \in Z_{G,S}$. The difference of the two sums above is precisely $\langle \omega, z\rangle$, hence vanishes. Therefore f is well-defined. By construction one has $\nabla f = \omega$.

Since $\omega \in \ell^2(E_{G,S})$ we obtain $\mathcal{E}_S(f) = \|\nabla f\|_2^2 = \|\omega\|_2^2 < \infty$, so $f \in \mathcal{D}_S(G)$. If f' is another Dirichlet function with $\nabla f' = \omega$, then $\nabla(f - f') = 0$, hence $f - f'$ is constant. □

Lemma 9.1.5. *Let $f \in \mathcal{D}_S(G)$. Then f is S-harmonic if and only if $\partial_1(\nabla f) = 0$.*

Proof. With our conventions, the adjoint relation between ∂_1 and $\partial_1^* = \nabla$ yields the concrete formula

$$(\partial_1\omega)(g) = \sum_{s\in S}\big(\omega(g,gs) - \omega(gs,g)\big) = \sum_{s\in S}\big(\omega(g,s) - \omega(gs,s^{-1})\big).$$

Applying this to $\omega = \nabla f$ gives

$$(\partial_1\nabla f)(g) = \sum_{s\in S}\big((f(gs)-f(g))-(f(g)-f(gs))\big) = -2\sum_{s\in S}\big(f(g)-f(gs)\big) = -2(\Delta_S f)(g).$$

In particular, $\partial_1\nabla f = 0$ if and only if $\Delta_S f = 0$. □

Theorem 9.1.6. *Let G be generated by a finite set S.*

(a) The discrete derivative map induces an isometric isomorphism

$$\mathcal{H}_S(G)/\mathbb{C} \;\cong\; \ker(\partial_1)\cap \mathcal{Z}_{G,S}^{\perp} \;\le\; \ell^2(E_{G,S}).$$

(b) For $f \in \mathcal{D}_S(G)$ set

$$b_f : G \to \ell^2(G), \qquad b_f(g) := \rho(g)f - f.$$

Then the assignment $f \mapsto b_f$ induces a linear isometric isomorphism

$$\mathcal{D}_S(G)/\mathbb{C} \;\cong\; Z^1(G,\rho),$$

where $Z^1(G,\rho)$ is equipped with the inner product from the measure $\mu_S = \frac{1}{|S|}\sum_{s\in S}\delta_s$.

(c) Under this identification, the subspace $\mathcal{H}_S(G)/\mathbb{C}$ corresponds to the subspace of μ_S-harmonic cocycles. Consequently, by Proposition 7.3.6, one obtains a canonical isometric isomorphism

$$\mathcal{H}_S(G)/\mathbb{C} \;\cong\; H^1_{\mathrm{red}}(G,\rho).$$

Proof. We first prove (a). Let $f \in \mathcal{H}_S(G)$. For every finite cycle $z \in Z_{G,S}$ we have $\langle \nabla f, z\rangle = 0$ because z is a finite sum of directed edges and the telescoping sum along the cycle is 0. Hence $\nabla f \in \mathcal{Z}_{G,S}^{\perp}$. By Lemma 9.1.5 we also have $\partial_1(\nabla f) = 0$. Thus ∇ maps $\mathcal{H}_S(G)/\mathbb{C}$ into $\ker(\partial_1)\cap\mathcal{Z}_{G,S}^{\perp}$. Injectivity modulo constants is clear.

Conversely, let $\omega \in \ker(\partial_1)\cap\mathcal{Z}_{G,S}^{\perp}$. Since $\omega \perp Z_{G,S}$, Lemma 9.1.4 provides $f \in \mathcal{D}_S(G)$ with $\nabla f = \omega$. Lemma 9.1.5 and $\partial_1\omega = 0$ imply $\Delta_S f = 0$, so $f \in \mathcal{H}_S(G)$. This proves surjectivity and yields the isomorphism.

For the dimension statement, note that $\mathcal{Z}_{G,S} \subseteq \ker(\partial_1)$ and both are closed Hilbert G-submodules. Therefore

$$\ker(\partial_1) = \mathcal{Z}_{G,S} \oplus \big(\ker(\partial_1)\cap\mathcal{Z}_{G,S}^{\perp}\big)$$

orthogonally, and by Lemma 9.1.1(c) we get

$$\dim_G\big(\ker(\partial_1)\cap\mathcal{Z}_{G,S}^{\perp}\big) = \dim_G\ker(\partial_1) - \dim_G\mathcal{Z}_{G,S}.$$

On the other hand, $\ker(\partial_1) = (\overline{\mathrm{im}(\partial_1^*)})^{\perp}$, so Lemma 9.1.1(d) gives

$$\dim_G\ker(\partial_1) = |S| - \dim_G\overline{\mathrm{im}(\partial_1^*)}.$$

The operator $\partial_1^*: \ell^2(G) \to \ell^2(E_{G,S})$ has kernel equal to the ℓ^2-constant functions, hence $\dim_G \ker(\partial_1^*) = b_0^{(2)}(G)$. Therefore

$$\dim_G \overline{\operatorname{im}(\partial_1^*)} = \dim_G \ell^2(G) - \dim_G \ker(\partial_1^*) = 1 - b_0^{(2)}(G),$$

and so

$$\dim_G \ker(\partial_1) = |S| - 1 + b_0^{(2)}(G).$$

Combining the previous identities yields

$$\dim_G\big(\ker(\partial_1) \cap \mathcal{Z}_{G,S}^{\perp}\big) = |S| - 1 + b_0^{(2)}(G) - \dim_G \mathcal{Z}_{G,S}.$$

By Theorem 9.1.9 the right-hand side equals $b_1^{(2)}(G)$, and the result follows from the isomorphism proved above. We now prove (b). We first check that $f \mapsto b_f$ lands in $Z^1(G,\rho)$ and only depends on the class of f modulo constants.

For $s \in S$ we have $b_f(s) = f(\cdot s) - f(\cdot)$, hence $b_f(s) \in \ell^2(G)$ and $\sum_{s\in S} \|b_f(s)\|_2^2 = \mathcal{E}_S(f) < \infty$. If $g = s_1 \cdots s_k$ is a word in S, then the telescoping identity

$$\rho(g)f - f = \sum_{i=1}^{k} \rho(s_1 \cdots s_{i-1})\big(\rho(s_i)f - f\big)$$

shows that $b_f(g)$ is a finite sum of translates of $b_f(s_i)$ and therefore belongs to $\ell^2(G)$. The cocycle identity

$$b_f(gh) = \rho(gh)f - f = b_f(g) + \rho(g)b_f(h)$$

is immediate, so $b_f \in Z^1(G,\rho)$. If f is replaced by $f + c$ with $c \in \mathbb{C}$, then $b_{f+c} = b_f$.

For the isometry, note that with our choice of the uniform measure on S we have

$$\|b_f\|_{Z^1}^2 = \sum_{s\in S} \mu_S(s)\, \|b_f(s)\|_2^2 = \frac{1}{|S|} \sum_{s\in S} \|\partial_s f\|_2^2 = \frac{1}{|S|}\, \mathcal{E}_S(f).$$

Thus $f \mapsto b_f$ descends to an isometric embedding $\mathcal{D}_S(G)/\mathbb{C} \to Z^1(G,\rho)$.

We now construct the inverse map. Given $b \in Z^1(G,\rho)$ define a function $f_b : G \to \mathbb{C}$ by

$$f_b(g) := b(g)(e) \qquad (g \in G),$$

where we view $b(g) \in \ell^2(G)$ as a function on G. Then for $x \in G$ and $s \in S$ we compute using the cocycle identity $b(xs) = b(x) + \rho(x)b(s)$ and evaluating at e:

$$f_b(xs) - f_b(x) = b(xs)(e) - b(x)(e) = (\rho(x)b(s))(e) = b(s)(x).$$

Hence $\partial_s f_b = b(s) \in \ell^2(G)$ for all $s \in S$, and therefore $f_b \in \mathcal{D}_S(G)$ with $\mathcal{E}_S(f_b) = \sum_{s\in S} \|b(s)\|_2^2 < \infty$. Moreover, the same computation shows $b = b_{f_b}$. This proves that $\mathcal{D}_S(G)/\mathbb{C} \to Z^1(G,\rho)$ is onto (and hence an isometric isomorphism).

Finally, b_f is μ_S-harmonic if and only if $\sum_{s\in S} b_f(s) = 0$ in $\ell^2(G)$. Evaluating at $x \in G$ yields

$$\sum_{s\in S} b_f(s)(x) = \sum_{s\in S} (f(xs) - f(x)) = -\sum_{s\in S} (f(x) - f(xs)) = -(\Delta_S f)(x),$$

so b_f is μ_S-harmonic if and only if $\Delta_S f = 0$, i.e. $f \in \mathcal{H}_S(G)$. Thus $\mathcal{H}_S(G)/\mathbb{C}$ corresponds exactly to the μ_S-harmonic cocycles.

By Proposition 7.3.6, every class in $H^1_{\mathrm{red}}(G,\rho)$ has a unique μ_S-harmonic representative, so $H^1_{\mathrm{red}}(G,\rho)$ identifies isometrically with the Hilbert space of μ_S-harmonic cocycles. Combining this with the previous paragraph yields the claimed identification of $\mathcal{H}_S(G)/\mathbb{C}$ with $H^1_{\mathrm{red}}(G,\rho)$. □

In particular, $H^1_{\mathrm{red}}(G,\rho)$ is naturally a Hilbert G-submodule of $\ell^2(E_{G,S})$.

Definition 9.1.7. Let G be a finitely generated group and let ρ be the right regular representation of G on $\ell^2(G)$. The *first ℓ^2-Betti number* of G is

$$b_1^{(2)}(G) := \dim_G H^1_{\mathrm{red}}(G,\rho).$$

The presentation chain complex

Fix a finite generating set S for G and choose a (possibly infinite) set of relators R such that $G \cong \langle S \mid R\rangle$. Let $X = X(G,S,R)$ be the associated Cayley 2-complex. Its cellular chain complex begins with a sequence of free $\mathbb{C}[G]$-modules

$$\mathbb{C}[G]^R \xrightarrow{\partial_2} \mathbb{C}[G]^S \xrightarrow{\partial_1} \mathbb{C}[G] \to \mathbb{C} \to 0.$$

Identifying $\mathbb{C}[G]^S \cong \mathbb{C}[E_{G,S}]$ by sending the basis vector corresponding to (g,s) to the directed edge from g to gs, we view ∂_2 as producing a (finite) edge-cycle from each relator.

Lemma 9.1.8. *With notation as above, one has*

$$\partial_2\big(\mathbb{C}[G]^R\big) = Z_{G,S} \subseteq \mathbb{C}[E_{G,S}], \qquad \textit{and hence} \qquad \overline{\mathrm{im}(\partial_2)} = \mathcal{Z}_{G,S} \subseteq \ell^2(E_{G,S}).$$

Proof. The boundary of each 2-cell of X is a finite cycle in the Cayley graph, so $\partial_2(\mathbb{C}[G]^R) \subseteq Z_{G,S}$. Conversely, every finite cycle is the boundary of a van Kampen diagram over the presentation $\langle S \mid R\rangle$; reading off the 2-cells shows that this cycle is a finite sum of G-translates of relator boundaries, hence lies in $\partial_2(\mathbb{C}[G]^R)$. Taking closures after tensoring with $\ell^2(G)$ gives the second claim. □

Theorem 9.1.9 (Pichot's formula)**.** *Let G be generated by a finite set S. Then*

$$b_1^{(2)}(G) = |S| - 1 + b_0^{(2)}(G) - \dim_G\big(\mathcal{Z}_{G,S}\big).$$

Proof. Choose a presentation $G \cong \langle S \mid R\rangle$ and consider the associated chain complex from the Cayley 2-complex: $\mathbb{C}[G]^R \xrightarrow{\partial_2} \mathbb{C}[G]^S \xrightarrow{\partial_1} \mathbb{C}[G] \to \mathbb{C} \to 0$. After tensoring with $\ell^2(G)$ and taking von Neumann dimensions one obtains

$$b_1^{(2)}(G) = |S| - \dim_G \overline{\mathrm{im}(\partial_1)} - \dim_G \overline{\mathrm{im}(\partial_2)}.$$

Here $\dim_G \overline{\mathrm{im}(\partial_1)} = 1 - b_0^{(2)}(G)$, and by Lemma 9.1.8 we have $\overline{\mathrm{im}(\partial_2)} = \mathcal{Z}_{G,S}$ inside $\ell^2(E_{G,S}) \cong \ell^2(G)^S$. Rearranging yields the claimed identity. □

Corollary 9.1.10. *Let G be finitely generated by S. We have $b_1^{(2)}(G) \leq |S|-1$ for infinite groups with equality only when G is free on S.*

9.2 A Morse inequality for finitely presented groups

The following inequality is often referred to as a Morse inequality in this context.

Theorem 9.2.1 (Morse inequality)**.** *Let G be an infinite finitely presented group with a presentation*

$$G \cong \langle S \mid R\rangle$$

where S is finite and R is a finite set of relations. Then

$$b_1^{(2)}(G) \geq |S| - 1 - |R|.$$

Proof. Let F be the free group on S and let $N \trianglelefteq F$ be the normal closure of R, so that $G \cong F/N$. As in the proof sketch of Theorem 9.1.9, the beginning of the associated free resolution yields a chain complex of free $\mathbb{C}[G]$-modules

$$\mathbb{C}[G]^R \xrightarrow{\partial_2} \mathbb{C}[G]^S \xrightarrow{\partial_1} \mathbb{C}[G] \to \mathbb{C} \to 0.$$

After tensoring with $\ell^2(G)$ we obtain a complex of Hilbert G-modules

$$\ell^2(G)^R \xrightarrow{\partial_2} \ell^2(G)^S \xrightarrow{\partial_1} \ell^2(G) \to 0.$$

By Lemma 9.1.8 we have $\overline{\operatorname{im}(\partial_2)} = \mathcal{Z}_{G,S} \subseteq \ell^2(E_{G,S})$. In particular, Pichot's formula (Theorem 9.1.9) identifies our invariant $b_1^{(2)}(G) = \dim_G H^1_{\text{red}}(G,\rho)$ with

$$b_1^{(2)}(G) = \dim_G \ker(\partial_1) - \dim_G \overline{\operatorname{im}(\partial_2)}.$$

Moreover, $\dim_G \ker(\partial_1) = |S| - 1 + b_0^{(2)}(G)$. Finally,

$$\dim_G \overline{\operatorname{im}(\partial_2)} \leq \dim_G \ell^2(G)^R = |R|.$$

Since G is infinite we have $b_0^{(2)}(G) = 0$ by Lemma 9.1.3. Combining these identities gives the claimed inequality. □

The following refinement of this inequality incorporates also the order of relators.

Theorem 9.2.2. *Let G be an infinite group which admits a presentation*

$$G \cong \langle S \mid r_1^{w_1}, \ldots, r_k^{w_k} \rangle$$

where $r_1, \ldots, r_k$ are words in the free group $\langle g_1, \ldots, g_n \rangle$ and $w_1, \ldots, w_k \in \mathbb{N}$. Assume the presentation is irredundant in the sense that the image of r_i in G has order exactly w_i for all i.

Then,

$$b_1^{(2)}(G) \geq |S| - 1 - \sum_{j=1}^{k} \frac{1}{w_j}.$$

Proof. We recast the argument in terms of reduced cohomology with coefficients in $\ell^2(G)$. Then one has the cohomological identity

$$b_1^{(2)}(G) = \dim_G H^1_{\text{red}}(G,\rho).$$

Let $F_n := \langle g_1, \ldots, g_n \rangle$ be the free group. Via the quotient map $F_n \twoheadrightarrow G$ we view ρ as a unitary representation of F_n on $\ell^2(G)$. Since F_n is free, a 1-cocycle $c \in Z^1(F_n,\rho)$ is determined by the values $c(g_i) \in \ell^2(G)$. Moreover, under the identification $Z^1(F_n,\rho) \cong \ell^2(G)^n$, the subspace of coboundaries corresponds to the (closure of the) image of the coboundary map

$$d : \ell^2(G) \to \ell^2(G)^n, \qquad \xi \mapsto \big(\rho(g_i)\xi - \xi\big)_{i=1}^n.$$

It follows that

$$\dim_G H^1_{\text{red}}(F_n,\rho) = n - 1.$$

The relations impose linear constraints via evaluation. For a cocycle c and a word $r \in F_n$, the cocycle identity gives the telescoping formula

$$c(r^w) = \sum_{\ell=0}^{w-1} \rho(r)^\ell\, c(r).$$

If the image of r in G has order exactly w, then

$$e := \frac{1}{w} \sum_{\ell=0}^{w-1} \rho(r)^\ell \in R(G)$$

is a nonzero projection, and $\tau(e) = 1/w$. In particular, the operator $\sum_{\ell=0}^{w-1} \rho(r)^\ell = we$ has image contained in $e\ell^2(G)$, whose von Neumann dimension equals $1/w$. Thus the constraint $c(r^w) = 0$ cuts out a subspace of codimension at most $1/w$ in von Neumann dimension.

Applying this to the relators $r_j^{w_j}$ and using irredundancy (so that r_j has order w_j in G) gives

$$\dim_G H^1_{\mathrm{red}}(G, \rho) \geq (n-1) - \sum_{j=1}^{k} \frac{1}{w_j},$$

which is the claimed inequality. □

Remark 9.2.3. Theorem 9.1.9 also gives an *upper* bound $b_1^{(2)}(G) \leq |S| - 1 + b_0^{(2)}(G)$ for every finite generating set S, because $\dim_G\big(\mathcal{Z}_{G,S}\big) \geq 0$.

9.3 Some computations

We now record a geometric method for bounding $b_1^{(2)}(G)$ for certain torsion groups, following [14]. The key input is an explicit construction of cycles in the Cayley graph.

9.3.1 A cycle estimate

We also use the following estimate (a special case of a more general character-theoretic formula).

Lemma 9.3.1. *Let G be generated by a finite set S. Then*

$$b_1^{(2)}(G) = |S| - 1 + b_0^{(2)}(G) - \sup_{0 \neq z \in Z_{G,S}} \sum_{s \in S} \frac{|\langle z, \bar{s}\rangle|^2}{\langle z, z\rangle},$$

where the inner product is the standard one on $\ell^2(E_{G,S})$.

Proof. This is a reformulation of Theorem 9.1.9 using the identity

$$\dim_G\big(\mathcal{Z}_{G,S}\big) = \sum_{s \in S} \langle P\bar{s}, \bar{s}\rangle$$

for the orthogonal projection P onto $\mathcal{Z}_{G,S}$, together with the Hilbert-space projection identity

$$\langle Pv, v\rangle = \sup_{0 \neq w \in Z_{G,S}} \frac{|\langle w, v\rangle|^2}{\langle w, w\rangle}.$$

Indeed, $\langle Pv, v\rangle = \|Pv\|^2$. For $w \in \mathcal{Z}_{G,S}$ we have $\langle v, w\rangle = \langle Pv, w\rangle$, so by Cauchy–Schwarz, $|\langle v, w\rangle|^2 \le \|Pv\|^2\,\|w\|^2$. Thus $|\langle v, w\rangle|^2/\|w\|^2 \le \langle Pv, v\rangle$, with equality for $w = Pv$ (if $Pv \neq 0$). Since $Z_{G,S}$ is dense in $\mathcal{Z}_{G,S}$, taking the supremum over $0 \neq w \in Z_{G,S}$ yields the same value. □

9.3.2 A uniform bound for p-torsion groups

Theorem 9.3.2 (Feldkamp–Kionke). *Let p be a prime and let G be a torsion group of exponent p. Then*

$$\bar{b}_1^{(2)}(G) \le 2p - 2.$$

Proof. We may assume G is infinite (otherwise $b_1^{(2)}(G) = 0$). By definition of $\bar{b}_1^{(2)}$ it suffices to treat the case that G is finitely generated.

Let S be a minimal generating set and set $N := |S|$. Since all nontrivial elements have order p, for pairwise distinct $a, b, c \in S$ one has

$$\langle ac\rangle \cap \langle ab\rangle = \{1\}.$$

Indeed, if $\langle ac\rangle \cap \langle ab\rangle \neq \{1\}$, then these cyclic groups of order p coincide, hence $ac = (ab)^k$ for some $k \in \mathbb{N}$, and this implies $c \in \langle a, b\rangle$, contradicting minimality.

Fix $a \in S$. For each $b \in S \setminus \{a\}$ the relation $(ab)^p = 1$ gives a cycle of length $2p$ in $\mathrm{Cay}(G, S)$. By the intersection property above, these $N - 1$ cycles have no common edge except the first edge $\bar{a}$. Summing them yields a cycle $z_a \in Z_{G,S}$ such that

$$\frac{|\langle z_a, \bar{a}\rangle|^2}{\langle z_a, z_a\rangle} = \frac{(N-1)^2}{(N-1)^2 + (N-1)(2p-1)} = \frac{1}{1 + \frac{2p-1}{N-1}}.$$

Applying Lemma 9.3.1 (and using $b_0^{(2)}(G) = 0$ for infinite G) gives

$$b_1^{(2)}(G) \le N - 1 - \sum_{a\in S} \frac{|\langle z_a, \bar{a}\rangle|^2}{\langle z_a, z_a\rangle} = N - 1 - \frac{N}{1 + \frac{2p-1}{N-1}} = \frac{2p-2}{1 + \frac{2p-1}{N-1}} \le 2p - 2.$$

□

Remark 9.3.3 (Feldkamp–Kionke, Theorem 2.2). Let p be a prime and let G be a countable torsion group of exponent p. If G has an infinite normal subgroup N of infinite index, then

$$b_1^{(2)}(G) = 0.$$

Indeed, by Theorem 9.3.2 and the definition of $\bar{b}_1^{(2)}$ we have

$$b_1^{(2)}(N) \le \bar{b}_1^{(2)}(N) \le 2p - 2 < \infty.$$

Gaboriau's theorem implies that if $N \trianglelefteq G$ is infinite and of infinite index and $b_1^{(2)}(N) < \infty$, then $b_1^{(2)}(G) = 0$. We refer to [16, Thm. 6.8]. □

9.3.3 Weakly q-normal subgroups

In this section we show that the presence of a large subgroup forces the first ℓ^2-Betti number to be large as well.

Definition 9.3.4. Let G be a countable group and let $H \leq G$ be an infinite subgroup. We say that H is

(a) *q-normal* in G if the set of elements $g \in G$ with $H \cap gHg^{-1}$ infinite generates G;

(b) *weakly q-normal* (or *wq-normal*) if there exists an ordinal α and an increasing chain of subgroups $H_0 \leq H_1 \leq \cdots \leq H_\alpha$ with $H_0 = H$, $H_\alpha = G$, and such that for every $\beta < \alpha$ the union $\bigcup_{\gamma<\beta} H_\gamma$ is q-normal in H_β.

Remark 9.3.5. The notion of q-normality and wq-normality was introduced by Popa in the context of von Neumann algebras. In some sense, it is easier to say when $H \leq G$ is not wq-normal than to say when it is. Indeed, $H \leq G$ is not wq-normal if and only if there exists a subgroup K with $H \leq K < G$ such that $K \cap gKg^{-1}$ is finite for all $g \in G \setminus K$.

Theorem 9.3.6. *Let G be a countable discrete group and let $H \leq G$ be an infinite wq-normal subgroup. Then*

$$b_1^{(2)}(H) \geq b_1^{(2)}(G).$$

Proof. The proof is by transfinite induction on the ordinal α from the definition of wq-normality. If $\alpha = 0$, then $H = G$ and the inequality is trivial. If α is a successor ordinal, say $\alpha = \beta + 1$, then H_β is q-normal in $H_{\beta+1}$, so we are left to show that the restriction map $H^1_{\text{red}}(H_{\beta+1}, \rho) \to H^1_{\text{red}}(H_\beta, \rho|_{H_\beta})$ is injective, where $\rho\colon H_{\beta+1} \to \mathcal{U}(\ell^2(H_{\beta+1}))$ is the right regular representation and $\rho|_{H_\beta}$ is its restriction to H_β. This follows from the fact that if $c \in Z^1(H_{\beta+1}, \rho)$ is a cocycle which is a coboundary when restricted to H_β, then c is a coboundary on $H_{\beta+1}$. We prove the slightly simpler fact that if $c \in Z^1(H_{\beta+1}, \rho)$ is a cocycle which vanishes on H_β, then c vanishes also on $H_{\beta+1}$.

Let $g \in H_{\beta+1}$ be such that $H_\beta \cap gH_\beta g^{-1}$ is infinite. Then for every $k \in H_\beta \cap gH_\beta g^{-1}$ we have

$$0 = c(g^{-1}kg) = \rho(g^{-1}k)c(g) + \rho(g^{-1})c(k) + c(g^{-1}) = \rho(g^{-1})(\rho(k) - 1)c(g),$$

so $c(g) = 0$ since $H_\beta \cap gH_\beta g^{-1}$ is infinite. Since the set of such g generates $H_{\beta+1}$, we conclude that c vanishes on $H_{\beta+1}$. □

9.4 Lück's approximation theorem

One of the most striking computations of ℓ^2-Betti numbers is Lück's approximation theorem, which allows one to compute ℓ^2-Betti numbers of residually finite groups as limits of normalized Betti numbers of finite quotients.

Theorem 9.4.1 (Lück, see [34])**.** *Let G be a finitely presented residually finite group and let (N_i) be a chain of finite index normal subgroups with trivial intersection. Then,*

$$b_1^{(2)}(G) = \lim_{i\to\infty} \frac{\dim_{\mathbb{Q}}((N_i)_{\text{ab}} \otimes \mathbb{Q})}{[G : N_i]}.$$

Proof. Fix a finite presentation $G = \langle s_1, \dots, s_n \mid r_1, \dots, r_m \rangle$ and let X be the associated Cayley 2-complex. Then the cellular chain complex of the universal cover $\widetilde{X}$ has the form

$$0 \to C_2(\widetilde{X}) \xrightarrow{\partial_2} C_1(\widetilde{X}) \xrightarrow{\partial_1} C_0(\widetilde{X}) \to 0$$

with $C_2(\widetilde{X}) \cong \mathbb{Z}[G]^m$, $C_1(\widetilde{X}) \cong \mathbb{Z}[G]^n$ and $C_0(\widetilde{X}) \cong \mathbb{Z}[G]$.

For each i set $Q_i := G/N_i$ and tensor this complex over $\mathbb{Z}[G]$ with $\mathbb{C}[Q_i]$. We obtain a finite-dimensional chain complex

$$\mathbb{C}[Q_i]^m \xrightarrow{\partial_{2,i}} \mathbb{C}[Q_i]^n \xrightarrow{\partial_{1,i}} \mathbb{C}[Q_i].$$

This is the cellular chain complex of the finite cover $X_i := N_i \backslash \widetilde{X}$. In particular,

$$\dim_{\mathbb{Q}}\big((N_i)_{\mathrm{ab}} \otimes \mathbb{Q}\big) = b_1(X_i; \mathbb{Q}) = \dim_{\mathbb{C}} H^1(X_i; \mathbb{C}).$$

We compute $H^1(X_i;\mathbb{C})$ using the Hodge Laplacian. Endow $\mathbb{C}[Q_i]^n$ and $\mathbb{C}[Q_i]^m$ with the standard Hermitian inner product coming from the basis $\{\delta_q\}_{q\in Q_i}$. Define

$$\Delta_i := \partial_{2,i}\partial_{2,i}^* + \partial_{1,i}^*\partial_{1,i} \qquad \text{on } \mathbb{C}[Q_i]^n.$$

Then $\Delta_i \geq 0$ and the finite-dimensional Hodge decomposition gives

$$\ker(\Delta_i) \;\cong\; H^1(X_i; \mathbb{C}),$$

so

$$\frac{\dim_{\mathbb{Q}}((N_i)_{\mathrm{ab}} \otimes \mathbb{Q})}{[G : N_i]} = \frac{\dim_{\mathbb{C}} \ker(\Delta_i)}{|Q_i|}. \tag{9.1}$$

We now relate the right-hand side to $b_1^{(2)}(G)$. Let $\Delta := \partial_2\partial_2^* + \partial_1^*\partial_1$ be the corresponding Laplacian on $\ell^2(G)^n$. By the standard definition of ℓ^2-Betti numbers in terms of von Neumann dimension of reduced cohomology (cf. Definition 9.1.7 and the discussion around Pichot's formula), one has

$$b_1^{(2)}(G) = \dim_G \ker(\Delta). \tag{9.2}$$

Let μ be the spectral measure of Δ with respect to the canonical trace. Concretely, μ is the unique probability measure on $[0, \|\Delta\|]$ such that

$$\tau\big(f(\Delta)\big) = \int f(t)\, d\mu(t)$$

for every bounded Borel function f (here τ is the normalized trace on the von Neumann algebra acting on $\ell^2(G)^n$). Equivalently, $\mu([0,t]) = \tau\big(\mathbf{1}_{[0,t]}(\Delta)\big)$ for all $t \geq 0$. For each i let μ_i be the normalized spectral measure of Δ_i:

$$\mu_i := \frac{1}{|Q_i|} \sum_{\lambda \in \mathrm{Spec}(\Delta_i)} (\mathrm{mult}_\lambda)\, \delta_\lambda.$$

Then

$$\mu(\{0\}) = \dim_G \ker(\Delta) = b_1^{(2)}(G), \qquad \mu_i(\{0\}) = \frac{\dim \ker(\Delta_i)}{|Q_i|}.$$

Thus, by (9.1), it suffices to show that

$$\mu_i(\{0\}) \longrightarrow \mu(\{0\}). \tag{9.3}$$

Step 1: moment convergence. For every polynomial p one has

$$\int p(t)\,d\mu(t) = \tau\big(p(\Delta)\big), \qquad \int p(t)\,d\mu_i(t) = \frac{1}{|Q_i|}\mathrm{Tr}\big(p(\Delta_i)\big).$$

Since Δ is given by a matrix over $\mathbb{Z}[G]$ with finite support, the same holds for $p(\Delta)$. Writing $p(\Delta) = \sum_{g\in F} a_g g$ (matrix coefficients in $\mathbb{C}$ and a finite set $F \subseteq G$), the quotient map $G \to Q_i$ sends it to $p(\Delta_i)$ and the normalized trace on $\mathbb{C}[Q_i]$ picks the coefficient of the identity in Q_i. Equivalently,

$$\frac{1}{|Q_i|}\mathrm{Tr}\big(p(\Delta_i)\big) = \sum_{g\in F\cap N_i} a_g.$$

Because $\bigcap_i N_i = \{1\}$ and F is finite, for i large enough we have $F \cap N_i = \{1\}$, hence

$$\frac{1}{|Q_i|}\mathrm{Tr}\big(p(\Delta_i)\big) \longrightarrow a_1 = \tau\big(p(\Delta)\big).$$

Therefore $\mu_i \Rightarrow \mu$ weakly.

Step 2: uniform control near 0 *(integrality).* Weak convergence alone does not imply convergence of the atom at 0; we need to prevent spectral mass from accumulating in $(0,\varepsilon)$. Let $M := \|\Delta\|$; then $\|\Delta_i\| \le M$ for all i.

We claim that for each i, the product of the nonzero eigenvalues of Δ_i is a positive integer, hence at least 1. Indeed, Δ_i is represented by a positive semidefinite symmetric matrix with integer entries in the standard basis of $\mathbb{Z}[Q_i]^n \subseteq \mathbb{C}[Q_i]^n$. Consider the polynomial

$$f_i(t) := \det(\Delta_i + t\mathbb{1}) = \prod_{j=1}^{n|Q_i|}(\lambda_{i,j} + t),$$

where $\lambda_{i,1},\dots,\lambda_{i,n|Q_i|} \ge 0$ are the eigenvalues of Δ_i counted with multiplicity. The coefficients of f_i are integers. If $k_i := \dim\ker(\Delta_i)$, then $f_i(t)$ is divisible by t^{k_i} and the coefficient of t^{k_i} equals $\prod_{\lambda_{i,j}>0}\lambda_{i,j}$. Since this coefficient is a nonzero integer, we obtain

$$\prod_{\lambda_{i,j}>0} \lambda_{i,j} \;\ge\; 1. \tag{9.4}$$

Now fix $\varepsilon \in (0,1)$ and let $m_i(\varepsilon)$ be the number of eigenvalues of Δ_i in $(0,\varepsilon]$. All nonzero eigenvalues lie in $(0,M]$, so by (9.4),

$$1 \le \prod_{\lambda_{i,j}>0} \lambda_{i,j} \le \varepsilon^{m_i(\varepsilon)}\, M^{n|Q_i|-k_i-m_i(\varepsilon)} \le \varepsilon^{m_i(\varepsilon)}\, M^{n|Q_i|}.$$

Taking logarithms yields

$$\frac{m_i(\varepsilon)}{|Q_i|} \le \frac{n\log M}{|\log\varepsilon|}.$$

In other words,

$$\sup_i \mu_i\big((0,\varepsilon]\big) = \sup_i \frac{m_i(\varepsilon)}{|Q_i|} \;\le\; \frac{n\log M}{|\log\varepsilon|} \underset{\varepsilon\downarrow 0}{\longrightarrow} 0. \tag{9.5}$$

Step 3: convergence of the atom at 0. Fix $\varepsilon > 0$ which is a continuity point of the distribution function of μ. By weak convergence, $\mu_i([0,\varepsilon]) \to \mu([0,\varepsilon])$. Moreover,

$$\mu_i(\{0\}) = \mu_i([0,\varepsilon]) - \mu_i((0,\varepsilon]).$$

Taking lim inf and using (9.5) gives

$$\liminf_{i\to\infty} \mu_i(\{0\}) \geq \mu([0,\varepsilon]) - \sup_i \mu_i((0,\varepsilon]).$$

Similarly,

$$\limsup_{i\to\infty} \mu_i(\{0\}) \leq \mu([0,\varepsilon]).$$

Now let $\varepsilon \downarrow 0$ through continuity points of μ. Since $\mu([0,\varepsilon]) \downarrow \mu(\{0\})$ and the error term tends to 0 by (9.5), we obtain (9.3).

Combining (9.1), (9.2) and (9.3) proves the theorem. □

Appendix

A.1 Schoenberg's theorem

Definition A.1.1. Let G be a group. A function $\varphi : G \to \mathbb{C}$ is *positive definite* if for every $n \geq 1$, every $g_1, \ldots, g_n \in G$ and every $c_1, \ldots, c_n \in \mathbb{C}$ one has

$$\sum_{i,j=1}^{n} \overline{c_i} c_j \, \varphi(g_i^{-1} g_j) \geq 0.$$

Equivalently, the matrix $\big(\varphi(g_i^{-1} g_j)\big)_{i,j}$ is positive semidefinite.

Theorem A.1.2 (GNS construction)**.** *Let G be a group and let $\varphi : G \to \mathbb{C}$ be positive definite. Then there exist a Hilbert space $\mathcal{H}_\varphi$, a unitary representation $\pi_\varphi : G \to \mathcal{U}(\mathcal{H}_\varphi)$ and a cyclic vector $\xi_\varphi \in \mathcal{H}_\varphi$ such that*

$$\varphi(g) = \langle \pi_\varphi(g)\xi_\varphi, \xi_\varphi \rangle \qquad (g \in G).$$

If $\varphi(e) = 1$, then $\|\xi_\varphi\| = 1$.

Proof. Let $\mathbb{C}[G]$ be the group algebra, i.e. the vector space of finitely supported functions $f : G \to \mathbb{C}$. Define a sesquilinear form on $\mathbb{C}[G]$ by

$$\langle f, h \rangle := \sum_{g,x \in G} \overline{f(g)}\, h(x)\, \varphi(g^{-1}x).$$

Positive definiteness of φ implies $\langle f, f \rangle \geq 0$ for all f. Let $\mathcal{N} := \{f \in \mathbb{C}[G] : \langle f, f \rangle = 0\}$ and set $\mathcal{H}_0 := \mathbb{C}[G]/\mathcal{N}$. The form descends to an inner product on $\mathcal{H}_0$, and we let $\mathcal{H}_\varphi$ be its completion.

For $g \in G$ define $(\pi_\varphi(g)f)(x) := f(g^{-1}x)$ on $\mathbb{C}[G]$. One checks that $\pi_\varphi(g)$ preserves $\langle \cdot, \cdot \rangle$, hence induces a unitary operator on $\mathcal{H}_\varphi$. Let $\delta_e \in \mathbb{C}[G]$ be the delta function at e and let ξ_φ be its class in $\mathcal{H}_\varphi$. Then

$$\langle \pi_\varphi(g)\xi_\varphi, \xi_\varphi \rangle = \langle \delta_g, \delta_e \rangle = \varphi(g).$$

Finally, ξ_φ is cyclic because the linear span of $\{\pi_\varphi(g)\xi_\varphi : g \in G\}$ contains the image of $\mathbb{C}[G]$. If $\varphi(e) = 1$, then $\|\xi_\varphi\|^2 = \langle \delta_e, \delta_e \rangle = \varphi(e) = 1$. □

Definition A.1.3. Let G be a group. A function $\psi : G \to \mathbb{R}$ is *conditionally negative definite* if $\psi(e) = 0$, $\psi(g) = \psi(g^{-1})$ for all $g \in G$, and for every $n \geq 1$, every $g_1, \ldots, g_n \in G$ and every $c_1, \ldots, c_n \in \mathbb{R}$ with $\sum_i c_i = 0$ one has

$$\sum_{i,j=1}^{n} c_i c_j \, \psi(g_i^{-1} g_j) \leq 0.$$

Proposition A.1.4 (Hilbert space model for conditionally negative definite functions). *Let $\psi : G \to \mathbb{R}$ be conditionally negative definite. Then there exist a Hilbert space $\mathcal{K}$ and a map $b : G \to \mathcal{K}$ such that*

$$\psi(g^{-1}h) = \|b(g) - b(h)\|^2 \qquad (g, h \in G).$$

In particular, $\psi(g) = \|b(g)\|^2$ after normalizing $b(e) = 0$.

Proof. Let V be the real vector space of finitely supported functions $f : G \to \mathbb{R}$ with $\sum_{g\in G} f(g) = 0$. Define a symmetric bilinear form on V by

$$\langle f, h\rangle_V := -\frac{1}{2} \sum_{g,x\in G} f(g)h(x)\, \psi(g^{-1}x).$$

Conditional negative definiteness of ψ implies $\langle f, f\rangle_V \geq 0$ for all $f \in V$. Let $\mathcal{N} := \{f \in V : \langle f, f\rangle_V = 0\}$ and let $\mathcal{K}$ be the completion of $V/\mathcal{N}$.

For $g \in G$ let δ_g be the delta function at g. Set $b(g) := [\delta_g - \delta_e] \in \mathcal{K}$. Then for $g, h \in G$ we have

$$\begin{aligned}\|b(g) - b(h)\|^2 &= \langle(\delta_g - \delta_h), (\delta_g - \delta_h)\rangle_V \\ &= -\tfrac{1}{2}\big(\psi(e) + \psi(e) - \psi(g^{-1}h) - \psi(h^{-1}g)\big) + \psi(g^{-1}h),\end{aligned}$$

where we used $\psi(e) = 0$ and $\psi(h^{-1}g) = \psi((g^{-1}h)^{-1}) = \psi(g^{-1}h)$. □

Lemma A.1.5. *Let $\mathcal{K}$ be a (real) Hilbert space and $t > 0$. Then the kernel $k_t(x, y) :=$ $\exp(-t\|x - y\|^2)$ on $\mathcal{K}$ is positive definite.*

Proof. Using $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x, y\rangle$, we can write

$$k_t(x, y) = e^{-t\|x\|^2}\, e^{-t\|y\|^2}\, e^{2t\langle x,y\rangle}.$$

The function $(x, y) \mapsto e^{2t\langle x,y\rangle}$ is positive definite because

$$e^{2t\langle x,y\rangle} = \sum_{n=0}^{\infty} \frac{(2t)^n}{n!}\, \langle x, y\rangle^n,$$

and each kernel $(x, y) \mapsto \langle x, y\rangle^n$ is positive definite (it is the Gram kernel on the n-th symmetric tensor power). Multiplying by the positive diagonal factors $e^{-t\|x\|^2}$ preserves positive definiteness. □

Theorem A.1.6 (Schoenberg). *Let G be a group and let $\psi : G \to \mathbb{R}$ be conditionally negative definite. Then for every $t > 0$ the function $\varphi_t(g) := \exp(-t\psi(g))$ is positive definite.*

Proof. Let $b : G \to \mathcal{K}$ be as in Proposition A.1.4, so that $\psi(g_i^{-1}g_j) = \|b(g_i) - b(g_j)\|^2$. Then for any $g_1, \ldots, g_n \in G$ we have

$$\big(\varphi_t(g_i^{-1}g_j)\big)_{i,j} = \big(e^{-t\|b(g_i)-b(g_j)\|^2}\big)_{i,j},$$

which is positive semidefinite by Lemma A.1.5. □

A.2 Algebraic Numbers

Definition A.2.1. A complex number α is called *algebraic* if it is a root of a non-zero polynomial with coefficients in $\mathbb{Q}$. We write $\overline{\mathbb{Q}}$ for the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$; its elements are precisely the algebraic numbers.

Let $G := \mathrm{Aut}(\overline{\mathbb{Q}}/\mathbb{Q})$ be the absolute Galois group of $\mathbb{Q}$, i.e. all field automorphisms of $\overline{\mathbb{Q}}$. For $\alpha \in \overline{\mathbb{Q}}$ and $\sigma \in G$ we call $\sigma(\alpha)$ a *Galois conjugate* of α. These conjugates are exactly the distinct roots of the minimal polynomial of α over $\mathbb{Q}$.

Proposition A.2.2. *The group G acts by permuting the roots of every polynomial $p(x) \in \mathbb{Q}[x]$. Moreover, the fixed field of G inside $\overline{\mathbb{Q}}$ is $\mathbb{Q}$ itself.*

Proof. If β is a root of p and $\sigma \in G$, then $p\big(\sigma(\beta)\big) = \sigma\big(p(\beta)\big) = 0$, so $\sigma(\beta)$ is another root; thus G permutes the roots.

For the fixed field, let $\alpha \in \overline{\mathbb{Q}}$ be fixed by every $\sigma \in G$ and let m_α be its minimal polynomial. The roots of m_α are the G–orbit of α. If the orbit is a single point, then $\deg m_\alpha = 1$ and $\alpha \in \mathbb{Q}$. Hence the elements fixed by all of G are exactly $\mathbb{Q}$. □

Definition A.2.3. An algebraic number α is an *algebraic integer* if it is a root of a monic polynomial with coefficients in $\mathbb{Z}$. We say that α is integral over $\mathbb{Z}$.

Proposition A.2.4. *If $r \in \mathbb{Q}$ is an algebraic integer, then $r \in \mathbb{Z}$.*

Proof. Write $r = a/b$ in lowest terms with $a \in \mathbb{Z}$, $b \in \mathbb{N}$. If r satisfies a monic polynomial $x^d + c_1x^{d-1} + \cdots + c_d$ with $c_i \in \mathbb{Z}$, then multiplying by b^d shows that b divides a^d. Since $\gcd(a, b) = 1$, this forces $b = 1$, hence $r = a \in \mathbb{Z}$. □

Theorem A.2.5 (Kronecker)**.** *Let $\alpha \neq 0$ be an algebraic integer and let $\{\alpha_1, \ldots, \alpha_d\}$ be its G–conjugates. If $|\alpha_i| \leq 1$ for every i, then $\alpha_1, \ldots, \alpha_d$ are roots of unity.*

Proof. If every $|\alpha_i| < 1$, then the constant term of the monic minimal polynomial m_α, given by $(-1)^d\alpha_1 \cdots \alpha_d$, has absolute value < 1 and is an integer, so it is 0, forcing $\alpha = 0$. Hence we may assume at least one conjugate has modulus 1 and $\alpha \neq 0$.

For $n \geq 1$ consider the algebraic integer α^n and its conjugates $\alpha_1^n, \ldots, \alpha_d^n$. The minimal polynomial of α^n is monic with integer coefficients that are elementary symmetric polynomials in the α_i^n. Each coefficient of degree k is a sum of $\binom{d}{k}$ products of k numbers of modulus ≤ 1, so its absolute value is at most $\binom{d}{k}$. There are only finitely many monic integer polynomials of degree $\leq d$ whose coefficients satisfy these bounds. Consequently, there exist integers $m > n$ such that α^m and α^n have the same minimal polynomial. This means the multisets $\{\alpha_1^m, \ldots, \alpha_d^m\}$ and $\{\alpha_1^n, \ldots, \alpha_d^n\}$ coincide, so there is a permutation π of $\{1, \ldots, d\}$ with $\alpha_i^m = \alpha_{\pi(i)}^n$ for all i.

The permutation π has finite order r, and iterating the equality gives $\alpha_i^{m^r} = \alpha_{\pi^r(i)}^{n^r} = \alpha_i^{n^r}$ for every i. Since none of the α_i vanish, we obtain $\alpha_i^{m^r - n^r} = 1$ for all i. In particular $\alpha^{m^r - n^r} = 1$, so α is a root of unity, and the same is true for each conjugate α_i. □

A.3 Growth of the partition function

Let $p(n)$ be the partition function, with $p(0) = 1$. The function $p(n)$ counts the number of ways to write n as a sum of positive integers, disregarding the order of the summands. The following theorem is a coarse version of a classical result of Hardy and Ramanujan; see [25, 1, 46].

Theorem A.3.1 (Hardy–Ramanujan)**.** *There exist absolute constants $c, C > 0$ such that for all sufficiently large n,*

$$c\sqrt{n} \le \log p(n) \le C\sqrt{n}.$$

In particular, $\log p(n) = \Theta(\sqrt{n})$.

For $|q| < 1$ the partition generating function is

$$\sum_{n\ge 0} p(n)q^n = \prod_{k\ge 1} \frac{1}{1-q^k}.$$

Set $q = e^{-t}$ with $t > 0$ and define

$$F(t) := \sum_{n\ge 0} p(n)\, e^{-nt} = \prod_{k\ge 1} \frac{1}{1-e^{-kt}}, \qquad S(t) := \log F(t).$$

Define a probability distribution μ_t on $\mathbb{N}$ by

$$\mu_t(\{n\}) := \frac{p(n)e^{-nt}}{F(t)}.$$

Then $\sum_{n\ge 0} \mu_t(\{n\}) = 1$. We define its mean by

$$m(t) := \mathbb{E}_{\mu_t}[N] = \sum_{n\ge 0} n\, \mu_t(\{n\}) = \frac{\sum_{n\ge 0} n\, p(n)e^{-nt}}{\sum_{n\ge 0} p(n)e^{-nt}}.$$

Differentiation under the summation sign yields

$$m(t) = -S'(t).$$

Lemma A.3.2. *The sequence $p(n)$ is nondecreasing: $p(n+1) \ge p(n)$ for all $n \ge 0$.*

Proof. Given a partition of n, add a part of size 1 to obtain a partition of $n+1$. This map is injective, hence $p(n+1) \ge p(n)$. □

We use the power series identity (valid for $|x| < 1$)

$$-\log(1-x) = \sum_{r=1}^{\infty} \frac{x^r}{r}.$$

Applying it with $x = e^{-kt}$ and summing over $k \ge 1$ (absolute convergence for $t > 0$) gives

$$\begin{aligned} S(t) &= \sum_{k\ge 1} -\log\bigl(1-e^{-kt}\bigr) = \sum_{k\ge 1}\sum_{r\ge 1} \frac{e^{-krt}}{r} = \sum_{r\ge 1} \frac{1}{r} \sum_{k\ge 1} e^{-krt} \\ &= \sum_{r\ge 1} \frac{1}{r} \cdot \frac{e^{-rt}}{1-e^{-rt}} = \sum_{r\ge 1} \frac{1}{r(e^{rt}-1)}. \end{aligned}$$

Lemma A.3.3. *For $y \ge 0$,*

$$e^y - 1 \ge y.$$

For $0 \le y \le 1$,

$$e^y - 1 \le (e-1)\, y.$$

Proof. The first inequality is $e^y \geq 1 + y$, which follows from convexity of e^y (or its Taylor series). For the second, use convexity of e^y on $[0, 1]$:

$$e^y \leq (1 - y)e^0 + ye^1 = 1 + (e - 1)y,$$

hence $e^y - 1 \leq (e - 1)y$. □

Lemma A.3.4. *For all $t > 0$,*

$$S(t) \leq \frac{2}{t}.$$

Proof. By Lemma A.3.3 we have $e^{rt} - 1 \geq rt$, hence

$$S(t) = \sum_{r\geq 1} \frac{1}{r(e^{rt} - 1)} \leq \sum_{r\geq 1} \frac{1}{r \cdot rt} = \frac{1}{t} \sum_{r\geq 1} \frac{1}{r^2} \leq \frac{2}{t}.$$

□

Lemma A.3.5. *For all $t \in (0, 1/2]$,*

$$S(t) \geq \frac{1}{(e - 1)} \cdot \frac{1}{t}.$$

Proof. Let $R := \lfloor 1/t \rfloor$, so $R \geq 2$ and $rt \leq 1$ for $1 \leq r \leq R$. By Lemma A.3.3 (applied with $y = rt \in [0, 1]$) we have $e^{rt} - 1 \leq (e - 1)rt$, hence

$$S(t) = \sum_{r\geq 1} \frac{1}{r(e^{rt} - 1)} \geq \sum_{r=1}^{R} \frac{1}{r(e^{rt} - 1)} \geq \sum_{r=1}^{R} \frac{1}{r \cdot (e - 1)rt} = \frac{1}{(e - 1)t} \sum_{r=1}^{R} \frac{1}{r^2} \geq \frac{1}{(e - 1)} \cdot \frac{1}{t}.$$

□

Lemma A.3.6. *For all $t \in (0, 1]$,*

$$m(t) \leq \frac{9}{t^2}.$$

Proof. From

$$S'(t) = \frac{F'(t)}{F(t)} = \frac{\sum_{n\geq 0}(-n)p(n)e^{-nt}}{\sum_{n\geq 0} p(n)e^{-nt}},$$

we indeed have $m(t) = -S'(t)$. Also, differentiating the product formula for $F(t)$ yields the standard identity

$$m(t) = \sum_{k\geq 1} \frac{k}{e^{kt} - 1}.$$

Split the sum into $k \leq 1/t$ and $k > 1/t$.

Small k: By Lemma A.3.3 we have $e^{kt} - 1 \geq kt$ for all k, hence for $k \leq 1/t$,

$$\frac{k}{e^{kt} - 1} \leq \frac{k}{kt} = \frac{1}{t}.$$

Therefore

$$\sum_{1\leq k\leq 1/t} \frac{k}{e^{kt} - 1} \leq \sum_{1\leq k\leq 1/t} \frac{1}{t} \leq \frac{1}{t} \cdot \frac{1}{t} = \frac{1}{t^2}.$$

Large k: If $kt \geq 1$, then $e^{kt} - 1 \geq (1 - e^{-1})e^{kt} = \frac{e-1}{e} e^{kt}$, hence

$$\frac{1}{e^{kt}-1} \leq \frac{e}{e-1} e^{-kt} \leq 2\, e^{-kt}.$$

So, for $k > 1/t$,

$$\frac{k}{e^{kt}-1} \leq 2k\, e^{-kt}.$$

Summing and using the geometric-series identity $\sum_{k \geq 1} kq^k = \frac{q}{(1-q)^2}$ for $|q| < 1$ with $q = e^{-t}$,

$$\sum_{k>1/t} \frac{k}{e^{kt}-1} \leq 2 \sum_{k \geq 1} k\, e^{-kt} = 2 \cdot \frac{e^{-t}}{(1-e^{-t})^2} \leq \frac{2}{(1-e^{-t})^2}.$$

For $t \in (0, 1]$ one has $1 - e^{-t} \geq t/2$ (e.g. by convexity of e^{-t} on $[0, 1]$, which gives $e^{-t} \leq 1 - \frac{t}{2}$), thus

$$\frac{2}{(1-e^{-t})^2} \leq \frac{2}{(t/2)^2} = \frac{8}{t^2}.$$

Combining both parts gives $m(t) \leq \frac{1}{t^2} + \frac{8}{t^2} = \frac{9}{t^2}$. □

Lemma A.3.7. *Let ν be a probability measure on $\mathbb{N}$ with finite mean $\mathbb{E}_\nu[X] = m$. Then there exists an integer $k \in \{0, 1, \ldots, \lfloor 2m \rfloor\}$ such that*

$$\nu(\{k\}) \geq \frac{1}{4m+2}.$$

Proof. Markov's inequality gives

$$\nu(X \geq 2m+1) \leq \frac{\mathbb{E}_\nu[X]}{2m+1} = \frac{m}{2m+1} < \frac{1}{2},$$

hence $\nu(X \leq 2m) \geq 1/2$. The set $\{0, 1, \ldots, \lfloor 2m \rfloor\}$ has at most $\lfloor 2m \rfloor + 1 \leq 2m+1$ points, so by pigeonhole there is some k in this set with

$$\nu(\{k\}) \geq \frac{\nu(X \leq 2m)}{2m+1} \geq \frac{1/2}{2m+1} = \frac{1}{4m+2}.$$

□

Proof of Theorem A.3.1. Upper bound. For any $t > 0$,

$$F(t) = \sum_{k \geq 0} p(k) e^{-kt} \;\geq\; p(n) e^{-nt},$$

hence $p(n) \leq e^{nt} F(t)$ and therefore

$$\log p(n) \leq nt + S(t).$$

Using $S(t) \leq 2/t$ (Lemma A.3.4) and choosing $t = n^{-1/2}$ gives

$$\log p(n) \leq n \cdot n^{-1/2} + \frac{2}{n^{-1/2}} = 3\sqrt{n}.$$

Thus we may take $C = 3$.

Lower bound. Fix large n and set

$$t := \frac{6}{\sqrt{n}}.$$

For large n we have $t \le 1/2$ and $t \le 1$, so Lemma A.3.5 and Lemma A.3.6 apply.

Consider the probability measure μ_t on $\mathbb{N}$ defined by

$$\mu_t(\{k\}) = \frac{p(k)e^{-kt}}{F(t)}.$$

Its mean is $m(t)$. By Lemma A.3.6,

$$m(t) \le \frac{9}{t^2} = \frac{9}{36}\,n = \frac{n}{4}.$$

Apply Lemma A.3.7 to $\nu = \mu_t$. There exists $k \le 2m(t) \le n/2 \le n$ such that

$$\mu_t(\{k\}) = \frac{p(k)e^{-kt}}{F(t)} \ge \frac{1}{4m(t)+2}.$$

Rearranging,

$$p(k) \ge e^{kt}\,\frac{F(t)}{4m(t)+2} \ge \frac{F(t)}{4m(t)+2},$$

hence

$$\log p(k) \ge \log F(t) - \log(4m(t)+2) = S(t) - \log(4m(t)+2).$$

Since $k \le n$ and p is nondecreasing (Lemma A.3.2), $p(n) \ge p(k)$ and thus

$$\log p(n) \ge S(t) - \log(4m(t)+2).$$

Now, Lemma A.3.5 gives (since $t \le 1/2$)

$$S(t) \ge \frac{1}{2(e-1)} \cdot \frac{1}{t} = \frac{1}{2(e-1)} \cdot \frac{\sqrt{n}}{6} = \frac{\sqrt{n}}{12(e-1)}.$$

Also $m(t) \le n/4$ implies $4m(t)+2 \le n+2$, hence $\log(4m(t)+2) \le \log(n+2)$. Finally,

$$\frac{\log(n+2)}{\sqrt{n}} \longrightarrow 0 \qquad (n \to \infty),$$

since for real $x \to \infty$, $\frac{\log x}{\sqrt{x}} \to 0$ by l'Hospital:

$$\lim_{x\to\infty} \frac{\log x}{\sqrt{x}} = \lim_{x\to\infty} \frac{1/x}{1/(2\sqrt{x})} = \lim_{x\to\infty} \frac{2}{\sqrt{x}} = 0.$$

Therefore

$$\frac{\log p(n)}{\sqrt{n}} \;\ge\; \frac{1}{12(e-1)} - \frac{\log(n+2)}{\sqrt{n}} \;\longrightarrow\; \frac{1}{12(e-1)} \quad \text{from below.}$$

In particular, for all sufficiently large n one has

$$\log p(n) \ge \frac{1}{24(e-1)}\sqrt{n},$$

so we may take $c = \frac{1}{24(e-1)}$. □

Bibliography

[1] G. E. Andrews, *The theory of partitions*, Addison–Wesley, 1976.

[2] G. Baumslag and S. M. Gersten, *Dehn functions and the length of some shortest proofs of triviality*, in *Algorithms and Classification in Combinatorial Group Theory*, G. Baumslag and C. F. Miller III (eds.), Springer, 1992.

[3] G. Baumslag and D. Solitar, *Some two-generator one-relator non-Hopfian groups*, Bull. Amer. Math. Soc. 68 (1962), 199–201.

[4] B. Bekka and A. Valette, *Groupes, représentations et invariants* L^2, Astérisque 245, Société Mathématique de France, 1997.

[5] I. Benjamini and O. Schramm, *Every graph with a positive Cheeger constant contains a tree with a positive Cheeger constant*, Geom. Funct. Anal. 7 (1997), 403–419.

[6] C. Chou, *Elementary amenable groups*, Illinois J. Math. 24 (1980), 396–407.

[7] A. Connes, *Classification of injective factors. Cases* II_1, II_∞, III_λ, $\lambda \neq 1$, *Ann. of Math.* 104 (1976), 73–115.

[8] Y. de Cornulier and R. Tessera, *Quasi-isometrically embedded free sub-semigroups*, *Geom. Topol.* 12 (2008), no. 1, 461–473.

[9] M. M. Day, *Amenable semigroups*, Illinois J. Math. 1 (1957), 509–544.

[10] P. Delorme, *1-cohomologie des représentations unitaires des groupes de Lie semi-simples et résolubles*, *Bull. Soc. Math. France* 105 (1977), 281–336.

[11] S. Eberhard and E. Maini, *The growth of residually soluble groups*, arXiv:2511.07018 (2025).

[12] G. Elek and E. Szabó, *Sofic groups and direct finiteness*, *J. Algebra* 280 (2004), no. 2, 426–434.

[13] G. Elek and E. Szabó, *Hyperlinearity, essentially free actions and* L^2*-invariants. The sofic property*, *Math. Ann.* 332 (2005), 421–441.

[14] C. Feldkamp and S. Kionke, *On upper bounds for the first* ℓ^2*-Betti number*, *Proc. Amer. Math. Soc.* 151 (2023), no. 6, 2375–2385.

[15] J. Friedman, *Sheaves on Graphs, Their Homological Invariants, and a Proof of the Hanna Neumann Conjecture: with an Appendix by Warren Dicks*, Mem. Amer. Math. Soc. **233** (2014), no. 1100, 106 pp.; preprint: arXiv:1105.0129.

[16] D. Gaboriau, Invariants ℓ^2 de relations d'équivalence et de groupes, *Publications Mathématiques de l'IHÉS* 95 (2002), 93–150.

[17] M. Greendlinger, *Dehn's algorithm for the word problem*, Comm. Pure Appl. Math. 13 (1960), 67–83.

[18] R. I. Grigorchuk, On the Milnor problem of growth, *Sov. Math. Dokl.* 28 (1983), 23–26.

[19] R. I. Grigorchuk, On the Hilbert–Poincaré series of graded algebras associated with groups, *Math. USSR-Sb.* **66** (1990), no. 1, 211–229.

[20] R. I. Grigorchuk, *On the Gap Conjecture concerning group growth*, arXiv:1202.6044 (2012).

[21] R. I. Grigorchuk, *Milnor's Problem on the Growth of Groups and its Consequences*, arXiv:1111.0512 (2011).

[22] M. Gromov, Groups of polynomial growth and expanding maps, *Inst. Hautes Etudes Sci. Publ. Math.* 53 (1981), 53–78.

[23] M. Gromov, *Endomorphisms of symbolic algebraic varieties*, *J. Eur. Math. Soc. (JEMS)* 1 (1999), no. 2, 109–197.

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

[25] G. H. Hardy and S. Ramanujan, *Asymptotic formulae in combinatory analysis*, Proc. London Math. Soc. (2) 17 (1918), 75–115.

[26] P. de la Harpe, *Topics in Geometric Group Theory*, University of Chicago Press, 2000.

[27] G. Higman, *A finitely generated infinite simple group*, J. London Math. Soc. 26 (1951), 61–64.

[28] V. A. Kaimanovich and A. M. Vershik, *Random walks on discrete groups: boundary and entropy*, Ann. Probab. 11 (1983), no. 3, 457–490.

[29] S. Kim and T. Koberda, *Non-freeness of groups generated by two parabolic elements with small rational parameters*, arXiv:1901.06375.

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

[31] N. Korevaar and R. Schoen, *Global existence theorems for harmonic maps to non-locally compact spaces*, Comm. Anal. Geom. 5 (1997), 333–387.

[32] M. Lazard, *Groupes analytiques p-adiques*, Publications Mathématiques de l'IHÉS 26 (1965), 5–219.

[33] R. Lyndon and P. Schupp, *Combinatorial Group Theory*, Springer, 1977.

[34] W. Lück, *Approximation of L^2-invariants by their finite-dimensional analogues*, *Geom. Funct. Anal.* 4 (1994), no. 4, 455–481.

[35] W. Lück, *L^2-Invariants: theory and applications to geometry and K-theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge, vol. 44, Springer, 2002.

[36] R. Lyons, *Random walks and the growth of groups*, C. R. Acad. Sci. Paris Sér. I Math. 320 (1995), 1361–1366.

[37] J. Milnor, Growth of finitely generated solvable groups, *J. Differential Geom.* 2 (1968), 447–449.

[38] I. Mineyev, *Submultiplicativity and the Hanna Neumann Conjecture*, Ann. of Math. **175** (2012), no. 1, 393–414.

[39] N. Mok, *Harmonic forms with values in locally constant Hilbert bundles*, in: *Several Complex Variables in Geometry and Mathematical Physics*, Proc. Sympos. Pure Math. 52, Part 2, AMS, 1991, 307–323.

[40] N. Ozawa, *A functional analysis proof of Gromov's polynomial growth theorem*, *Ann. Sci. Éc. Norm. Supér.* 51 (2018), 549–556.

[41] P. Papasoglu, Strongly geodesically automatic groups are hyperbolic, *Invent. Math.* 121 (1995), 323–334.

[42] A. L. T. Paterson, *Amenability*, American Mathematical Society, 1988.

[43] J. Peterson and A. Thom, *Group cocycles and the ring of affiliated operators*, *Invent. Math.* 185 (2011), 561–592.

[44] V. Pestov, *Hyperlinear and sofic groups: a brief guide*, Bull. Symbolic Logic 14 (2008), no. 4, 449–480; arXiv:0804.3968.

[45] B. Prunaru, *A Poisson boundary for topological semigroups*, *Arch. Math.* 102 (2014), 449–454.

[46] S. Ramanujan, *Collected papers*, Cambridge Univ. Press, 1927; reprint, Chelsea, New York, 1962.

[47] J.-P. Serre, *Trees*, Springer, 1980.

[48] J.-P. Serre, *Lie groups and Lie algebras*, Benjamin, New York, 1965.

[49] Y. Shalom, *Harmonic analysis, cohomology, and the large-scale geometry of amenable groups*, *Acta Math.* 192 (2004), 119–185.

[50] J. R. Stallings, *Topology of finite graphs*, *Invent. Math.* 71 (1983), 551–566.

[51] J. Tits, Free subgroups in linear groups, *J. Algebra* 20 (1972), 250–270.

[52] S. Wagon, *The Banach–Tarski Paradox*, Cambridge University Press, 1985.

[53] B. Weiss, *Sofic groups and dynamical systems*, *Sankhyā Ser. A* 62 (2000), no. 3, 350–359.

[54] K. Whyte, *Amenability, Bilipschitz Equivalence, and the von Neumann Conjecture*, Duke Math. J. 99 (1999), 93–112.

[55] J. A. Wolf, Growth of finitely generated solvable groups and curvature of Riemannian manifolds, *J. Differential Geom.* 2 (1968), 421–446.

Index