An introduction to amenable groups

Alejandra Garrido

These notes are based on a series of four talks which I gave at the Oxford Advanced Class in Algebra in Michaelmas Term 2013. As the title suggests, they are intended to be an introduction to the theory of amenable groups, starting from their origins in measure theory and the Banach–Tarski paradox running through some of the important results that have led to the solution of the von Neumann–Day problem. The material is largely based on a variety of excellent sources surveying the topic, which I heartily recommend, namely the books by Lubotzky ([7], Chapter 2), Paterson ([10]) and Wagon ([14]), and Appendix G of [2] by Valette. Terry Tao's blog entries ([11] and [12]) on the subject also provide great introductory material.

1 The origins of amenability: the Banach–Tarski paradox

1.1 The Banach–Tarski paradox

In 1924 Banach and Tarski ([1]) proved a remarkable theorem which is nowadays stated as "given a ball in 3-dimensional space, there is a way of decomposing it into finitely many disjoint pieces that can be rearranged to form two balls of the same size as the original one". This counterintuitive result is essentially a statement about measure theory. It says that there is no finitely additive measure on $\mathbb{R}^3$ which is defined on every subset of $\mathbb{R}^3$, is isometry-invariant and gives the unit ball non-zero measure. Consequently, one cannot extend Lebesgue measure to all subsets of $\mathbb{R}^3$; that is, there are non-Lebesgue-measurable sets. On the other hand, there is no analogue of the Banach–Tarski paradox for smaller dimensions; indeed, there do exist such measures on $\mathbb{R}$ and $\mathbb{R}^2$. The reason behind this dichotomy, as we shall see, is that the isometry groups of $\mathbb{R}$ and $\mathbb{R}^2$ are amenable while that of $\mathbb{R}^3$ is not.

In this section, we will prove the Banach–Tarski paradox and see how it leads to the definition of an amenable group.

Definition 1.1. Let a group G act on a set X and $A, B \subseteq X$ be two subsets of X. We say that A and B are *(finitely) G-equidecomposable* if each can be partitioned into finitely many subsets $A_1, \dots, A_n$ and $B_1, \dots, B_n$ and there exist elements $g_1, \dots, g_n \in G$ such that $B_i = g_i A_i$ for each i.

We denote this by $A \sim B$ and write $A \precsim B$ if $A \sim C$ for some subset $C \subseteq B$.

A *realization* h of $A \sim B$ is a bijection $h : A \to B$ such that for each i in a decomposition as above we have $h(a_i) = g_i a_i$ for all $a_i \in A_i$.

For fixed G, being G-equidecomposable is a transitive relation: Suppose $A \sim B$ and $B \sim D$ with decompositions $A = \sqcup_{i_1}^n A_i$, $B = \sqcup_{i=1}^n B_i = \sqcup_{j=1}^m C_j$, $D = \sqcup_{j=1}^m D_j$ and group elements $g_1, \dots, g_n$, $h_1, \dots, h_m$. Then we can form new partitions by defining $A_{ij} := g_i^{-1}(B_i \cap C_j)$, $g_{ij} := g_i|_{A_{ij}}$, $h_{ij} := h_j|_{B_i \cap C_j}$ and $D_{ij} := h_{ij} g_{ij}(A_{ij})$ for $1 \le i \le n, 1 \le j \le m$. This gives a realization of $A \sim D$.

Note that if $h : A \to B$ is a realization of $A \sim B$ and $S \subseteq A$ then $S \sim h(S)$, almost by definition.

We will use both of the above facts in what follows.

Theorem 1.2 (Banach–Schröder–Bernstein). *Let G act on X and $A, B \subseteq X$. If $A \lesssim B$ and $B \lesssim A$ then $A \sim B$.*

Proof. Since $A \lesssim B$ and $B \lesssim A$, there are bijections $f : A \to B_1$ and $g : A_1 \to B_1$ for $A_1 \subseteq A$ and $B_1 \subseteq B$. Put $C_0 := A \setminus A_1$ and $C_{n+1} := g^{-1}f(C_n)$ for $n \geq 0$. Let $C := \bigcup_{n=1}^{\infty} C_n$ be the union of these sets. Then for $a \in A \setminus C$ we have $a \notin C_n$ for each n, so $g(a) \notin f(C_n)$ for each n. Hence, $g(A \setminus C) \subseteq B \setminus f(C)$. Similarly, $B \setminus f(C) \subseteq g(A \setminus C)$. Thus $A \setminus C \sim B \setminus f(C)$. Since $C \sim f(C)$, we conclude that $A \sim B$. □

Corollary 1.3. *The following are equivalent:*

1. *There exist proper disjoint subsets A, B of X such that $A \sim X \sim B$;*
2. *There exist proper disjoint subsets A, B of X such that $A \cup B = X$ and $A \sim X \sim B$.*

Proof. Since $X \sim B \subseteq (X \setminus A) \subset X$, we have $X \lesssim X \setminus A$. Containment $X \setminus A \subset X$ gives $X \setminus A \lesssim X$, so the theorem gives $A \sim X \sim X \setminus A$. □

Definition 1.4. Let a group G act on a set X. We say that X is *(finitely) G-paradoxical* if any and hence both of the conditions in the above corollary hold.

Proposition 1.5. 1. *The free group F_2 of rank 2 is F_2-paradoxical (where F_2 acts on itself by left multiplication).*

2. *If F_2 acts freely on a set X then X is F_2-paradoxical.*

Proof. 1. Let F_2 be generated by $\{a, b\}$ and let $W(x)$ denote the set of reduced words starting with x. Then writing

$$\begin{aligned} F_2 &= \{1\} \sqcup W(a) \sqcup W(a^{-1}) \sqcup W(b) \sqcup W(b^{-1}) \\ &= W(a) \sqcup aW(a^{-1}) \\ &= W(b) \sqcup bW(b^{-1}) \end{aligned}$$

we see that $A := W(a) \sqcup W(a^{-1})$ and $B := W(b) \sqcup W(b^{-1})$ satisfy the conditions of the first part of the corollary.

2. Let M be a set of representatives for the F_2-orbits of X. For $c \in F_2$, define $X_c := \{zm \mid z \in W(c), m \in M\}$. Then the sets $X_a, X_{a^{-1}}, X_b, X_{b^{-1}}$ are disjoint and, as in the previous part, $X = X_a \sqcup aX_{a^{-1}} = X_b \sqcup bX_{b^{-1}}$ gives the desired decomposition.

□

Proposition 1.6. *The special orthogonal group $SO(3, \mathbb{R})$ contains a copy of F_2.*

Proof. The rotations ρ and σ given by the matrices below generate a copy of F_2 (for more details, see Theorem 2.1 of [14]):

$$\rho = \begin{pmatrix} 1/3 & -2\sqrt{2}/3 & 0 \\ 2\sqrt{2}/3 & 1/3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ \sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/3 & -2\sqrt{2}/3 \\ 0 & 2\sqrt{2}/3 & 1/3 \end{pmatrix}.$$

□

This proposition will enable us to prove the Banach–Tarski paradox, which is sometimes called the Hausdorff–Banach–Tarski paradox because it uses the following result by Hausdorff.

Theorem 1.7 (Hausdorff Paradox)**.** *There is a countable subset D of the 2-sphere $\mathbb{S}^2$ such that $\mathbb{S}^2 \setminus D$ is finitely $SO(3,\mathbb{R})$-paradoxical.*

Proof. Every nontrivial element of $SO(3,\mathbb{R})$ fixes exactly two points in $\mathbb{S}^2$ (a pair of antipodal points where the axis of rotation meets the sphere). Let D be the union of all these fixed points. Since F_2 embeds in $SO(3,\mathbb{R})$, it acts freely on $\mathbb{S}^2 \setminus D$. The result follows from Proposition 1.5. $\square$

Proposition 1.8. *If D is a countable subset of the 2-sphere $\mathbb{S}^2$ then $\mathbb{S}^2$ and $\mathbb{S}^2 \setminus D$ are $SO(3,\mathbb{R})$-equidecomposable.*

Proof. Let l be a line through the origin that misses the given subset D (this exists because D is countable). Since D is countable, there is an angle θ such that for every $n \in \mathbb{Z}^+$ the image $\rho^n(D)$ of D under the rotation ρ^n by $n\theta$ around l does not intersect D. Let $\bar{D} := \bigcup_{n=0}^{\infty} \rho^n(D)$ be the union of the rotations of D. Then $\mathbb{S}^2 = \bar{D} \cup (\mathbb{S}^2 \setminus \bar{D}) \sim \rho(\bar{D}) \cup \mathbb{S}^2 \setminus \bar{D}$. $\square$

Since being $SO(3,\mathbb{R})$-equidecomposable is a transitive relation, these results immediately yield the following form of the Banach–Tarski paradox.

Corollary 1.9 (Banach–Tarski)**.** *The 2-sphere $\mathbb{S}^2$ is $SO(3,\mathbb{R})$-paradoxical. In fact, the n-sphere $\mathbb{S}^n$ is $SO(n+1,\mathbb{R})$-paradoxical for $n \geq 2$.*

Proof. It suffices to show the inductive step. Suppose $\mathbb{S}^{n-1} = A \cup B$ is a paradoxical decomposition. For $X = A, B$, define $X_1 := \{(x_1, \ldots, x_n, x_{n+1} \in \mathbb{S}^n \mid (x_1, \ldots, x_n)/\sqrt{x_1^2 + \cdots + x_n^2} \in X\}$. Then it is not hard to see that A_1 and B_1 are still disjoint and their union is the whole n-sphere $\mathbb{S}^n$ minus the two poles $(0, \ldots, 0, \pm 1)$. Further, $\mathbb{S}^n \setminus (0, \ldots, 0, \pm 1) = A_1 \sqcup B_1$ is a paradoxical decomposition and, by the same arguments as in Proposition 1.8, $\mathbb{S}^n \setminus (0, \ldots, 0, \pm 1)$ is $SO(n+1,\mathbb{R})$-equidecomposable with $\mathbb{S}^n$. $\square$

This implies that one cannot put a finitely additive rotation-invariant probability measure on all subsets of $\mathbb{S}^n$ for $n \geq 2$.

Using this we can establish the more familiar form of the paradox.

Corollary 1.10 (Banach–Tarski paradox)**.** *Let $E(3)$ denote the group of isometries of $\mathbb{R}^3$. Any solid ball in $\mathbb{R}^3$ is $E(3)$-paradoxical. Furthermore, $\mathbb{R}^3$ is $E(3)$-paradoxical.*

Proof. We will only consider balls centered at the origin as $E(3)$ contains all translations. In fact, the same proof works for balls of any size so we may take the unit ball $\mathbb{B}$. Because every subset A of $\mathbb{S}^2$ corresponds one-to-one with a subset $\{\lambda A \mid 0 < \lambda \leq 1\}$ of $\mathbb{B}$, the $E(3)$-paradoxical decomposition of $\mathbb{S}^2$ yields such a decomposition of $\mathbb{B} \setminus \{0\}$. Thus, we must show that $\mathbb{B} \sim \mathbb{B} \setminus \{0\}$. Let ρ be a rotation of infinite order about an axis that crosses $\mathbb{B}$ but misses the origin. Then, as in the proof of Proposition 1.8, $\mathbb{B} = \bar{D} \cup (\mathbb{B} \setminus \bar{D}) \sim \rho(\bar{D}) \cup \mathbb{B} \setminus \bar{D}$, where $D := \{0\}$ and $\bar{D} = \{\rho(D)^n \mid n \geq 0\}$.

The proof for $\mathbb{R}^3$ uses the same argument, taking instead the radial correspondence of $\mathbb{S}^2$ with $\mathbb{R}^3 \setminus \{0\}$. $\square$

As a converse to the Banach–Tarski paradox, we have the following theorem, whose proof we omit (see [14], Corollary 9.2)

Theorem 1.11 (Tarski)**.** *Let G be a group acting on a set X and let $E \subseteq X$. There is a finitely additive G-invariant measure $\mu : \mathcal{P}(X) \to [0, \infty]$ with $\mu(E) = 1$ if and only if E is not G-paradoxical.*

1.2 Amenability

We have just seen that a way of obtaining a paradoxical decomposition of a set X is to find a paradoxical decomposition of a group G acting on the set and then 'transfer' it. Conversely, if there is a finitely additive left-invariant measure on $\mathcal{P}(G)$ then we can use it to find a finitely additive G-invariant measure on $\mathcal{P}(X)$. Such a measure shows that X cannot be G-paradoxical. It was von Neumann ([13]) who realized in the 1920s that this transference was possible and began to classify which groups had measures of this sort.

Definition 1.12. Let G be a discrete (resp. locally compact) group. A *measure on* G is a finitely additive measure μ on $\mathcal{P}(G)$ (respectively, $\mathcal{B}(G)$, the Borel sets of G), with $\mu(G) = 1$ and which is left-invariant; that is, $\mu(gA) = \mu(A)$ for every $g \in G$ and $A \subseteq G$. We say that G is *amenable* (or *'mittelbar'* in the original German) if it has such a measure.

We can immediately notice that if a group is paradoxical with respect to left multiplication then it cannot be amenable. Thus any group containing F_2 is not amenable. This led von Neumann to conjecture that if a group is non-amenable then it must contain F_2. We shall see a counterexample to this conjecture later on.

Definition 1.13. Let G be a locally compact group equipped with the Haar measure μ. Recall that $L^\infty(G)$ is the set of (equivalence classes of) essentially bounded measurable functions $f : G \to \mathbb{R}$ where a function is essentially bounded if it is bounded outside a set of zero measure. If G is discrete, μ is just the counting measure and $L^\infty(G)$ becomes $\ell^\infty(G)$.

Construct an integral on $(G, \mathcal{B}(G), \mu)$, so $\int f \,\mathrm{d}\mu$ defines a linear functional on $L^\infty(G)$ such that

1. $\int f \,\mathrm{d}\mu \geq 0$ if $f(g) \geq 0$ for all $g \in G$;
2. $\int \mathbf{1}_G \,\mathrm{d}\mu = 1$ where $\mathbf{1}_G$ denotes the indicator function on G;
3. $\int {}_g f \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu$ for every $g \in G$ and $f \in L^\infty(G)$ where ${}_g f(h) := f(g^{-1}h)$.

Such a linear functional is a *left-invariant mean on* G.

Thus an amenable group always has a left-invariant mean. Conversely, if $m : L^\infty(G) \to \mathbb{R}$ is a left-invariant mean, then defining $\mu(A) := m(\mathbf{1}_A)$ for $A \in \mathcal{B}(G)$ gives a measure on G.

Remark. There is nothing special about left-multiplication in the definitions of a measure and a mean on G. If μ is a left-invariant measure then $\mu_{-1}(A) := \mu(A^{-1})$ is a right-invariant measure on G as $\mu_{-1}(Ax) = \mu(x^{-1}A^{-1}) = \mu(A^{-1}) = \mu_{-1}(A)$. Furthermore, the integral with respect to μ_{-1} is invariant under the right action of G on $L^\infty(G)$: $\int f_g \,\mathrm{d}\mu_{-1} = \int f \,\mathrm{d}\mu_{-1}$ where $f_g(h) = f(gh^{-1})$.

Note that a left-invariant measure on an amenable group is not necessarily right-invariant. Nevertheless, a group G is amenable if and only if there is some left- and right-invariant measure on G.

Proof. Suppose μ is a left-invariant measure witnessing the amenability of G and let μ_{-1} be the corresponding right-invariant measure, as above. For any Borel subset A of G, define $f_A \in L^\infty(G)$ by $f_A(g) = \mu(Ag^{-1})$ so we can define $\nu : \mathcal{B}(G) \to [0, \infty]$ by $\nu(A) := \int f_A \,\mathrm{d}\mu_{-1}$. Then ν fulfills the requirements:

1. $\nu(G) = 1$ as $\mu(G) = \mu_{-1}(G) = 1$;

2. ν is finitely additive since if $A, B \in \mathcal{B}(G)$ are disjoint we have $f_{A\cup B} = f_A + f_B$;

3. ν is left-invariant as $f_{gA}(x) = \mu(gAx^{-1}) = \mu(Ax^{-1}) = f_A(x)$, and right-invariant since $f_{Ag}(x) = \mu(Agx^{-1}) = f_A(xg^{-1}) = (f_A)_g(x)$ and $\int (f_A)_g \,\mathrm{d}\mu_{-1} = \int f_A \,\mathrm{d}\mu_{-1}$.

□

Proposition 1.14. *Let G act on X and suppose that G is amenable. Then there is a finitely additive G-invariant probability measure on $\mathcal{P}(X)$. Therefore X is not G-paradoxical (by Tarski's theorem).*

Proof. Let μ be a measure on G that witnesses its amenability. Choose a point $x \in X$. Define $\nu : \mathcal{P}(X) \to [0,1]$ by $\nu(A) := \mu(\{g \in G \mid gx \in A\})$. This is easily seen to satisfy the requirements of the statement. □

So now we have the following list of equivalent definitions of amenability. We will add a few more later on.

Theorem 1.15. *For a group G, the following are equivalent:*

1. *G is amenable, that is, there is a finitely additive left-invariant probability measure on $\mathcal{B}(G)$;*
2. *there is a left-invariant mean on G;*
3. *G is not paradoxical.*

Historical note. It was Mahlon Day ([4]) who proved in the 1950s the equivalence between von Neumann's definition of amenability and the existence of an invariant mean. As we have seen, this was not a particularly difficult result, but it is an important change of perspective as it made the extensive tools of functional and harmonic analysis.

Day also coined the term 'amenable', apparently as a pun on 'mean'.

2 Properties and examples of amenable groups

With our current running list of equivalent definitions of amenability we can already prove quite a few properties. For ease of exposition, we will mostly focus on discrete groups, but these properties also hold in the locally compact case. The interested reader is directed to [2], Appendix G, or [10] for these proofs and further topics in amenability of a more analytic flavour, alluded to at the end of the last section.

We start with an easy example.

Example 2.1. All finite groups are amenable: for any subset $A \subseteq G$ of G define $\mu(A) := |A|/|G|$.

Further, all compact groups are amenable: take the unique left Haar measure[1] and normalize to make it a probability measure.

[1] Recall Haar's theorem: for a locally compact Hausdorff topological group G, there is a unique (up to positive multiplicative constant) countably additive nontrivial measure μ on $\mathcal{B}(G)$ which is

- left-invariant,
- finite on compact subsets,
- outer regular on $E \in \mathcal{B}(G)$, that is, $\mu(E) = \inf\{\mu(U) \mid E \subseteq U \in \mathcal{B}(G),\, U \text{ open}\}$,
- inner regular on open Borel sets $E \in \mathcal{B}(G)$, that is, $\mu(E) = \sup\{\mu(K) \mid K \subseteq E,\, K \text{ compact}\}$.

Proposition 2.2. *Let G be a (discrete) group. Then*

1. *if G is amenable, any subgroup H of G is amenable and any quotient G/N of G is amenable;*
2. *if $N \trianglelefteq G$ and G/N are amenable, so is G;*
3. *if all groups in the direct system $\{G_i\}_{i\in I}$ are amenable, so is their direct union $G := \bigcup_{i\in I} G_i$.*

Proof. 1. Let $\mu : \mathcal{P}(G) \to [0, 1]$ be a measure on G witnessing its amenability. Any subgroup H of G (by the axiom of choice) has a right transversal M in G. So we can define $\nu : \mathcal{P}(G) \to [0, \infty]$ by $\nu(A) := \mu(AM)$ which is easily seen to be a measure on H. For a quotient G/N of G, define $\lambda : \mathcal{P}(G/N) \to [0, \infty]$ by $\lambda(A) := \lambda(AN)$, which is a measure on G/N.

2. Suppose that ν_1 and ν_2 are measures on (respectively) $N \trianglelefteq G$ and G/N. Then, for $A \subseteq G$, define $f_A : G \to \mathbb{R}$ by $f_A(g) := \nu_1(N \cap g^{-1}A)$. Since ν_1 is left-invariant, f_A is well-defined on (left) cosets of N so we can view it as a function on G/N. Define $\mu : \mathcal{P}(G) \to [0, \infty]$ by $\mu(A) := \int f_A \,\mathrm{d}\nu_2$, which is easily seen to be a finitely-additive probability measure on G. To see that it is left-invariant, note that $\int f_{xA} \,\mathrm{d}\nu_2 = \int x(f_A) \,\mathrm{d}\nu_2 = \int f_A \,\mathrm{d}\nu_2$ since ν_2 is left-invariant.

3. We have $G = \bigcup_{i\in I} G_i$ where for each $i, j \in I$ there exists $k \in I$ such that $G_i, G_j \le G_k$. For each $i \in I$, denote by μ_i the measure on G_i that makes it amenable and define the set

$$M_i := \{\mu : \mathcal{P}(G) \to [0, 1] \mid \mu \text{ is a finitely additive measure and } \mu(gA) = \mu(A) \text{ for all } g \in G_i\}.$$

Then each M_i is nonempty since we can define μ by $\mu(A) := \mu_i(A \cap G_i)$. Note that $[0, 1]^{\mathcal{P}(G)}$ is compact by Tychonoff's theorem and it is easy to check that each M_i is closed in $[0, 1]^{\mathcal{P}(G)}$. Now, if $G_i, G_j \le G_k$ then $M_k \subseteq M_i \cap M_j$ as $\mu(gA) = \mu(A)$ for all $g \in G_k$ implies that the same holds for all $g \in G_i, G_j$. Thus the family $\{M_i\}_{i\in I}$ of closed subsets of $[0, 1]^{\mathcal{P}(G)}$ has the finite intersection property and so there exists a measure $\mu \in \bigcap_{i\in I} M_i$ which makes G amenable.

$\square$

Proposition 2.3. *All abelian (discrete) groups are amenable.*

Proof. We start by noting that every group is the direct union of its finitely generated subgroups (every element is contained in the cyclic group it generates, which is part of the direct union). By Proposition 2.2, it suffices to consider finitely generated abelian groups. Since every such group is the direct sum of $\mathbb{Z}^n$ and a finite group T (which is amenable) we reduce, again by Proposition 2.2, to showing that $\mathbb{Z}$ is amenable.

It is enough to show that for each $\varepsilon > 0$ there is a finitely-additive probability measure μ_ε on $\mathcal{P}(\mathbb{Z})$ which is almost invariant with respect to the generator a of $\mathbb{Z}$, that is,

$$|\mu_\varepsilon(A) - \mu_\varepsilon(aA)| \le \varepsilon \text{ for every } A \subseteq \mathbb{Z}.$$

This would show that each set M_ε of ε-invariant finitely additive probability measures on $\mathbb{Z}$ is nonempty. Each M_ε is also closed. The family $\{M_\varepsilon\}_{\varepsilon>0}$ of all these sets satisfies the finite intersection property since $\bigcap_{i=1}^n M_{\varepsilon_i} = M_{\min\{\varepsilon_i\}}$. Hence, compactness of $[0, 1]^{\mathcal{P}(\mathbb{Z})}$ once more gives the existence of a finitely additive probability measure $\mu \in \bigcap_{\varepsilon>0} M_\varepsilon$ which is left-invariant with respect to a and hence with respect to $\mathbb{Z}$. In order to find μ_ε, for given $\varepsilon > 0$ choose $N \in \mathbb{N}$ such that $2/N \leq \varepsilon$ and define

$$\mu_\varepsilon(A) := \frac{|\{i \mid 1 \leq i \leq N,\, a^i \in A\}|}{N}.$$

Then $|\mu_\varepsilon(A) - \mu_\varepsilon(aA)| \leq 2/N \leq \varepsilon$. □

Corollary 2.4. *All virtually solvable groups are amenable.*

This is immediate from the previous propositions.

We can now introduce the class EG of *elementary amenable groups*. This is the smallest class of groups containing all abelian groups and all finite groups and closed under taking subgroups, quotients, extensions and direct unions. It is in some sense the 'smallest' class of amenable groups and so it is reasonable to ask whether EG is properly contained in AG, the class of all amenable groups. This is part of the von Neumann–Day conjecture, which we will explore in the last section in more detail.

Corollary 2.5. *The Banach–Tarski paradox has no analogue for dimensions 1 and 2.*

For this we need the Invariant Extension Theorem.

Theorem 2.6 (Invariant Extension Theorem)**.** *Recall Carathéodory's Extension Theorem: If $\mathcal{R}$ is a subring of the boolean algebra $\mathcal{A}$ and μ is a measure on $\mathcal{R}$, then μ can be extended to a measure $\bar{\mu}$ on $\mathcal{A}$.*

If G is an amenable group of automorphisms of $\mathcal{A}$ and $\mathcal{R}$, μ are G-invariant, then $\bar{\mu}$ can be chosen to be G-invariant.

Proof. Use Carathéodory's Extension Theorem to find a measure ν on $\mathcal{A}$ extending μ and let θ be a measure on G. If $b \in \mathcal{A}$ then define $f_b : G \to \mathbb{R}$ by $f_b(g) := \nu(g^{-1}b)$ and

$$\bar{\mu}(b) := \begin{cases} \int f_b \,\mathrm{d}\theta, & f_b \in \ell^\infty(G) \\ \infty, & \text{otherwise.} \end{cases}$$

Then $\bar{\mu}$ is a G-invariant extension of μ. □

In fact, this is only true for amenable groups. Suppose it is true for G, then apply it to $\mathcal{A} = \mathcal{P}(G)$, $\mathcal{R} = \{\varnothing, G\}$ with $\mu(\varnothing) = 0$, $\mu(G) = 1$ and G acting on $\mathcal{A}$ by left-translation. Then the $\bar{\mu}$ given by the Invariant Extension Theorem is a measure on G, making it amenable.

This gives us another characterization of amenability. Thus our running list of equivalent definitions now looks like.

Theorem 2.7. *For a group G, the following are equivalent:*

1. *G is amenable, that is, there is a finitely additive left-invariant probability measure on $\mathcal{B}(G)$;*
2. *there is a left-invariant mean on G;*
3. *G is not paradoxical;*
4. *G satisfies the Invariant Extension Theorem.*

To prove the corollary, we apply this theorem to extend Lebesgue measure to an invariant finitely additive measure defined on all subsets of $\mathbb{R}$ or $\mathbb{R}^2$. We can do this as the groups of isometries $E(1)$ and $E(2)$ of $\mathbb{R}$ and $\mathbb{R}^2$ (respectively) are solvable and hence amenable. Since the Banach–Tarski paradox implies the non-existence of such an extension of Lebesgue measure (i.e., it implies that there are non-Lebesgue-measurable sets in $\mathbb{R}^n$, $n \geq 3$), the paradox cannot hold in dimensions 1 and 2.

3 Amenable groups and growth

In the previous section we proved that all abelian groups are amenable, finding almost-invariant measures, where 'almost' is arbitrarily close to 0, and then 'making the measures converge' to an actual invariant measure. In this section, we will see a generalization of this idea, the *Følner condition*, which we will show to be equivalent to amenability; in fact, it is typically given as the definition. We will also see how this condition can be used to show that all groups of subexponential growth are (supra)amenable.

3.1 The Følner condition

Definition 3.1. A discrete group G satisfies the *Følner condition* if for every finite subset $A \subseteq G$ and every $\varepsilon > 0$ there exists a finite nonempty subset $F \subseteq G$ such that for each $a \in A$ we have

$$\frac{|aF \bigtriangleup F|}{|F|} \leq \varepsilon.$$

If G is locally compact we use the same definition but A is a compact subgroup, F is a Borel set with positive finite Haar measure and we use Haar measure instead of cardinality.

We can view this in terms of the Cayley graph of G. For any finite subset A of G, the Cayley graph $\Gamma(G, A)$ of G with respect to A is the graph with elements of G as vertices and edges (g, ga) for $a \in A$. The Følner condition then states that for every finite subset F of the vertices, the boundary of F (that is, the number of edges leaving F) is at most $\varepsilon|F|$. The reader familiar with expander graphs will notice that this is a sort of 'opposite' condition to expansion.

Example 3.2. As an easy example, we see that all finite (or compact in the topological case) groups satisfy the Følner condition, by simply taking $F = G$.

Frequently in the literature one finds the following definition.

Definition 3.3. For a discrete and countable (resp. locally compact) group G, a *Følner sequence* is a sequence $\{F_n\}$ of nonempty finite (resp. compact) subsets of G such that

$$\frac{|gF_n \bigtriangleup F_n|}{|F_n|} \to 0 \quad \left(\text{resp. } \frac{\mu(gF_n \bigtriangleup F_n)}{\mu(F_n)} \to 0\right)$$

for every $g \in G$.

If G is discrete and uncountable, we define a *Følner net* in the obvious way.

This definition is conveniently equivalent to the previous one.

Lemma 3.4. *A group G satisfies the Følner condition if and only if it has a Følner sequence.*

Proof. Suppose G satisfies the condition and write it as $G = \bigcup_n A_n$, the ascending union of finite subsets $A_1 \subset A_2 \subset \dots$. Let $\varepsilon_n = 1/n$ for each n. The Følner condition then implies that for each n, there is a finite subset F_n such that for every $a \in A_n$ we have $|aF_n \triangle F_n|/|F_n| \leq 1/n$. Then for any $g \in G$ there is some A_n containing it (and therefore all larger A_m contain it too) so that $|gF_n \triangle F_n|/|F_n| \leq 1/n \to 0$.

The converse proof is easy, picking an appropriate F_n for given ε.

□

Example 3.5. The group $\mathbb{Z}$ has a Følner sequence, namely $F_n = \{-n, \dots, n\}$.

Exercise: extend this to finitely generated abelian groups.

If every finitely generated subgroup of a discrete group G satisfies the Følner condition, so does G, since every finite subset of G generates a subgroup which satisfies the condition. Hence all abelian groups satisfy the Følner condition.

Theorem 3.6. *A group satisfies the Følner condition if and only if it is amenable.*

Proof. For any given finite $A \subseteq G$ and $\varepsilon > 0$ define $M_{A,\varepsilon}$ to be the set of finitely additive probability measures μ on G such that $|\mu(B) - \mu(aB)| \leq \varepsilon$ for every $B \subseteq G$ and every $a \in A$. It is easy to check that each of these sets is closed in the compact $[0,1]^{\mathcal{P}(G)}$, so it suffices to show that they are nonempty. For this note that we can define $\mu(B) := |B \cap F|/|F|$ where F is given by the Følner condition (since $B \cap F, aB \cup F \subseteq F \cup aF$, we have $|\mu(B) - \mu(aB)| \leq |F \triangle aF|/|F| \leq \varepsilon$.)

For the converse, we give an argument of Namioka ([8]). Let

$$\Phi := \{f \in \ell^1(G) \mid f \geq 0 \text{ is finitely supported and } \|f\|_{\ell^1(G)} = \sum_{g \in G} |f(g)| = 1\}.$$

We first show that if G is amenable, for every finite $A \subseteq G$, $\varepsilon > 0$ there exists $f \in \Phi$ such that $\|f - {}_af\|_{\ell^1(G)} \leq \varepsilon$ for all $a \in A$. Suppose this is not the case, so there exist $A \subseteq G$, $\varepsilon > 0$ such that for every $f \in \Phi$ there is some $a \in A$ with $\|f - {}_af\|_{\ell^1(G)} > \varepsilon$. Then $\{f - {}_af \mid f \in \Phi\}$ is a convex subset of $\ell^1(G)$ bounded away from 0 so, by the Hahn–Banach Separation Theorem, there exist $\beta \in \ell^1(G)^*$ and $t \in \mathbb{R}$ such that $\beta(f - {}_af) \geq t > 0$ for all $f \in \Phi$. Since $\ell^(G)^* \cong \ell^\infty(G)$ there is some $b \in \ell^\infty(G)$ such that $\langle f - \delta_a * f, b\rangle = \sum_{x \in G}(f - \delta_a * f)(x)m(x) \geq t$ for all $f \in \Phi$. Taking $f = \delta_y$ for $y \in G$, we obtain

$$\langle \delta_y - \delta_a * \delta_y, m\rangle = \sum_{x \in G} \delta_y(x)m(x) - \delta_y(a^{-1}x)m(x) = m(y) - {}_{a^{-1}}m(y) \geq t.$$

Thus $m(y) - {}_{a^{-1}}m(y) \geq t$ for every $y \in G$. Since G is amenable, there is a left-invariant $M : \ell^\infty(G) \to \mathbb{R}$. Applying M to the above we obtain $M(m - {}_{a^{-1}}m) \geq t > 0$, contradicting the left-invariance of M.

In particular, for fixed $A \subseteq G$ and $\varepsilon > 0$ there is some $f \in \Phi$ such that $\|f - {}_af\|_{\ell^1(G)} \leq \varepsilon/|A|$ for every $a \in A$. Since f is finitely supported, we can find a 'layer cake' representation for it: $f = \sum_{i=1}^n c_i \mathbf{1}_{F_i}$ for nonempty finite $F_1 \supseteq \dots \supseteq F_n$ and $c_i > 0$. We also have $\sum_{i=1}^n c_i|F_i| = 1$ as $f \in \Phi$. Now, $|f(g) - {}_af(g)| \geq c_i$ for $g \in (aF_i \triangle F_i)$, so

$$\sum_{i=1}^n c_i |aF_i \triangle F_i| \leq \|f - {}_af\|_{\ell^1(G)} \leq \varepsilon/|A| \sum_{i=1}^n c_i|F_i|$$

for each $a \in A$. Hence

$$\sum_{i=1}^{n} \sum_{a \in A} c_i |aF_i \triangle F_i| \leq \varepsilon \sum_{i=1}^{n} c_i |F_i|.$$

By the pigeonhole principle there is some i such that $\sum_{a \in A} |aF_i \triangle F_i| \leq \varepsilon |F_i|$ whence $|aF_i \triangle F_i|/|F_i| \leq \varepsilon$ for all $a \in A$. $\square$

Remark. Alternatively, to prove that Følner's condition implies amenability, one can use the equivalence between the condition and the existence of a Følner sequence. Given a Følner sequence $\{F_n\}$ one would like to define $\mu(B) := \lim_n |B \cap F_n|/|F_n|$. However, this limit may not exist. To get around this, one takes an ultralimit $\lim_\omega |B \cap F_n|/|F_n|$ along some non-principal ultrafilter ω on $\mathbb{N}$.

3.2 Growth

We start by recalling the definition of word growth of a finitely generated group.

Definition 3.7. Let G be generated by a finite symmetric subset S. Define the *length function* $l_S : G \to \mathbb{N}$ (with respect to S) by taking $l_S(g)$ to be the length of a shortest representative of g as a word in S.
This defines a metric on G, where $B_S(n)$ (the *ball of radius* n centred at the identity) is the set of all elements $g \in G$ with $l_S(g) \leq n$.
Define the *growth function* of G with respect to S by $\gamma_G^S(n) := |B_S(n)|$.
For two functions γ_1, γ_2 write $\gamma_1 \precsim \gamma_2$ if there exist $C, \alpha > 0$ such that $\gamma_1(n) \leq C\gamma_2(\alpha n)$ for all n. Write $\gamma_1 \sim \gamma_2$ if $\gamma_1 \precsim \gamma_2$ and $\gamma_2 \precsim \gamma_1$. This is easily seen to be an equivalence relation and it is left as an exercise to check that all growth functions (with respect to finite generating sets) of a given finitely generated group are equivalent. We will therefore usually omit the superscript S.

There are three types of growth:

1. polynomial growth, where $\gamma(n) \sim n^\alpha$ for some $\alpha > 0$;
2. exponential growth, where $\gamma(n) \sim \mathrm{e}^{\mathrm{n}}$;
3. intermediate growth, where $\gamma(n)$ is equivalent to neither of the above.

The limit $\lim_{n\to\infty} \gamma_G(n)^{1/n}$ exists for all finitely generated G (because $\gamma_G(n)$ is submultiplicative). If this limit is strictly greater than 1, the group has exponential growth; if it is at most 1, we say it has *subexponential growth*. We make use of this to prove the following.

Theorem 3.8. *All subgroups of subexponential growth are amenable.*

Proof. The idea is to use the balls $B(n)$ as a Følner sequence. Let G have subexponential growth, so $\gamma(n)^{1/n} = |B(n)^{1/n}| \to x \leq 1$. This means that for every $\varepsilon > 0$ there is some k_ε such that $|B(k_\varepsilon + 1)|/|B(k_\varepsilon)| < 1 + \varepsilon$. Put $n_i := k_{1/i}$ so for every s in a fixed generating set S, we have

$$\frac{|sB(n_i) \triangle B(n_i)|}{|B(n_i)|} \leq \frac{2|B(n_i + 1)| - |B(n_i)|}{|B(n_i)|} < 2(1 + 1/i) - 2 \to 0.$$

Hence for every $g \in G$ (which is a word in S) we obtain

$$\frac{|gB(n_i) \triangle B(n_i)|}{|B(n_i)|} \to 0,$$

as required. $\square$

In fact, groups of subexponential growth are more than amenable, they are supramenable, a concept introduced by Rosenblatt.

Definition 3.9. A group G is *supramenable* if for every $\varnothing \neq A \subseteq G$ there is a finitely additive left-invariant measure $\mu : \mathcal{P}(G) \to [0, 1]$ such that $\mu(A) = 1$.

This implies in particular that no nonempty subset of G is paradoxical. Conversely, if no nonempty subset of G is paradoxical, Tarski's theorem (Theorem 1.11) yields that G is supramenable. We will use this to show that all groups of subexponential growth are supramenable.

Theorem 3.10. *Let G be finitely generated.*

1. *If G has subexponential growth and acts on a set X then no nonempty $A \subseteq X$ is G-paradoxical.*
2. *If G has subexponential growth then G is supramenable.*

Proof. 1. Suppose that A is G-paradoxical. Then there are two piecewise G-transformations $h_1, h_2 : A \to A$ such that $h_1(A) \cap h_2(A) = \varnothing$. Let $S := \{g_1, \dots, g_n\}$ be the elements of G occurring as multipliers in h_1, h_2. Since G has subexponential growth, there exists n such that $\gamma_G^S(n) < 2^n$. Now, consider the functions (2^n of them) which are obtainable as 'words' in h_1, h_2 of length n (i.e., compositions of h_1, h_2 of length n). Denote each of them by f_i for $1 \leq i \leq 2^n$. By the properties of h_1, h_2, each f_i is still a function $A \to A$ and $f_i(A) \cap f_j(A) = \varnothing$ for distinct i, j. For any $x \in A$, the set $\{f_i(x) \mid 1 \leq i \leq 2^n\}$ must contain 2^n elements. However, by construction of the f_i, each $f_i(x)$ is of the form wx where w is a word in S of length n and there are strictly fewer than 2^n of them by assumption.

2. This follows from the previous part, letting G act on itself by left multiplication. Let $\varnothing \neq A \subseteq G$. By the previous part, A is not G-paradoxical, so Tarski's theorem yields a finitely additive left-invariant measure on $\mathcal{P}(G)$ which assigns measure 1 to A.

$\square$

We have included the proof of this theorem because it illustrates how a G-paradoxical set in a G-action implies that G has exponential growth.

4 The von Neumann–Day problem

In previous sections, we saw that containing a nonabelian free group is an obstacle to being amenable. We also saw that the class of groups EG generated by finite and amenable groups and closed under taking subgroups, quotients, extensions and direct unions consists of amenable groups. We thus have the following inclusions:

$$EG \subseteq AG \subseteq NF$$

where AG denotes the class of amenable groups and NF is the class of groups which do not contain nonabelian free subgroups.

The *von Neumann–Day*[2] *problem* asks whether these inclusions are strict.

This problem was completely solved in the 1980s. In 1980, Ol'shanskii [9] proved that $AG \neq NF$ by showing that the Tarski monster group (which is an infinite group in which every nontrivial proper subgroup is cyclic of order a fixed prime p) is not amenable. The proof of this, which involves intricate combinatorial arguments, is beyond the scope of these talks, so we avoid it. The other part of the problem was solved in 1985 by Grigorchuk ([6]) with the next theorem which we will prove in this section, following the exposition in [5].

Theorem 4.1. *The (first) Grigorchuk group* Γ *is amenable but not elementary amenable.*

4.1 The class of elementary amenable groups

We start with a crucial and elegant result of Chou ([3]).

Theorem 4.2. *Every torsion group in* EG *is locally finite.*

The proof of this requires an auxiliary result.

Theorem 4.3. *The class* EG *is the class generated by finite groups and abelian groups and the operations of taking extensions and direct unions.*

To prove this we make the following definition.

Definition 4.4. Denote by EG_0 the class of all finite and all abelian groups. Let $\alpha > 0$ be an ordinal and suppose EG_β has been defined for all ordinals $\beta < \alpha$. If α is a limit ordinal, define $EG_\alpha := \bigcup\{EG_\beta \mid \beta < \alpha\}$; otherwise, define EG_α as the class of groups which have been obtained from a group in $EG_{\alpha-1}$ by one, and only one, extension or direct union.

Lemma 4.5. *Each* EG_α *is closed under taking subgroups and quotients.*

Proof. This is clear for EG_0. Suppose $\alpha > 0$ and that EG_β is closed under the stated operations if $\beta < \alpha$. Let $G \in EG_\alpha$, B a subgroup of G and $C = G/N$ a quotient of G. If α is a limit ordinal then $G \in EG_\beta$ for some $\beta < \alpha$ and so $B, C \in EG_\beta$. If α is not a limit ordinal we have two cases:

a) There is an exact sequence $1 \to K \to G \to Q$ with $K, Q \in EG_{\alpha-1}$. Then $1 \to K \cap B \to B \to P \to 1$ is an exact sequence with $P \leq Q$, and so is $1 \to K_1 \to C \to Q_1 \to 1$ with K_1, Q_1 homomorphic images of K, Q.

b) G is a direct union of groups $\{G_i\}_{i \in I}$ in $EG_{\alpha-1}$. Then B is a direct union of $\{B \cap G_i\}_{i \in I}$ and C is a direct union of $\{\varphi(G_i)\}_{i \in I}$ where $\varphi : G \twoheadrightarrow C$ is the quotient map.

The lemma follows by transfinite induction. □

Proof of Theorem 4.3. It suffices to show that $EG = \bigcup\{EG_\alpha \mid \alpha \text{ an ordinal}\}$. The right hand side is clearly contained in EG. It is, almost by construction, closed under taking extensions and, by the lemma, under taking subgroups and quotients. To show that it is closed under taking direct

[2]The problem was suggested by a number of people in the 1950s, 20 years after von Neumann's work on amenable groups first appeared. It was first stated in print, with von Neumann's name attached to it, in a 1957 paper by Day.

unions, let $G = \bigcup_{i\in I} G_i$ be the direct union of groups G_i contained in the right hand side. Each G_i is contained in some EG_{α_i}, so let α be the supremum of the α_i. Then $G_i \in EG_\alpha$ for each $i \in I$, whence $G \in EG_{\alpha+1}$ which is contained in the right hand side.

Since EG was defined to be the smallest class containing EG_0 closed under the above operations, equality follows and the theorem is proved. □

Proof of Theorem 4.2 Recall that we wish to show that all torsion groups in EG are locally finite. The class EG_0 clearly satisfies this. Suppose then that $\alpha > 0$ is an ordinal and that EG_β satisfies the theorem for every ordinal $\beta < \alpha$. Let $G \in EG_\alpha$ be torsion. If α is a limit ordinal then G is in some EG_β with $\beta < \alpha$ and so is locally finite by inductive hypothesis. If α is not a limit ordinal then G is either an extension $1 \to K \to G \to Q \to 1$ or a direct union $\bigcup_{i\in I} G_i$ of groups in $EG_{\alpha-1}$. In the former case, any finitely generated subgroup H of G will have finite image in the quotient Q, so the kernel $H \cap K$ will have finite index in H and therefore will be a finitely generated subgroup of K, making H finite. In the latter case, the generators of any finitely generated subgroup H will be contained in some G_i and so H will be finite too. Hence, in all cases G is locally finite and the theorem follows by transfinite induction and Theorem 4.3. □

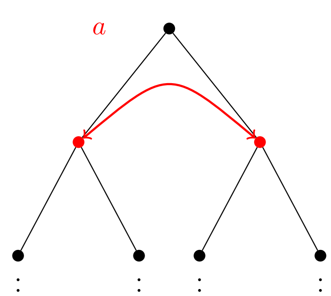
4.2 The (first) Grigorchuk group

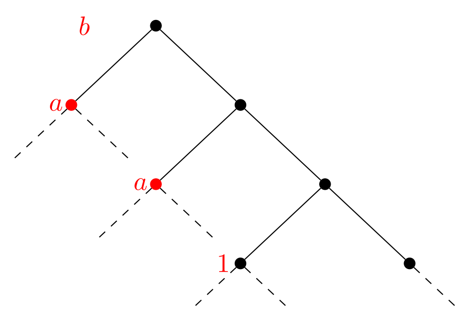
The result of Chou that we have just proved traces a clear route for solving the Day problem. Namely, it suffices to find an amenable torsion group which is not locally finite. We will show that the first Grigorchuk group is one such example.

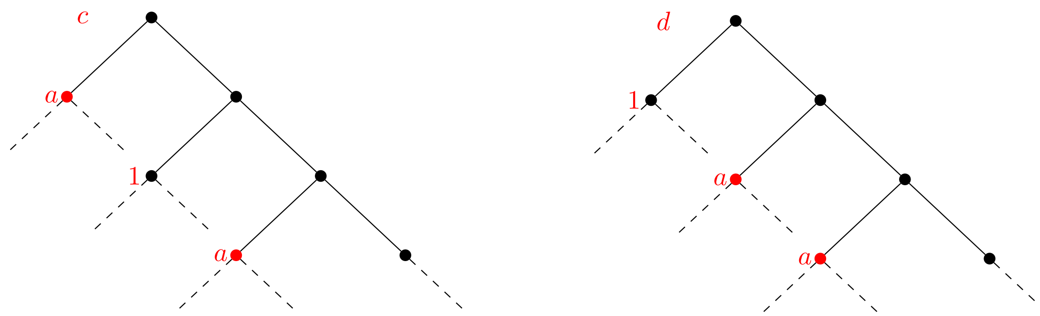
Definition 4.6. Denote by T the infinite rooted binary tree. We will generally identify the vertices of T with finite words in the alphabet $\{0, 1\}$. The *(first) Grigorchuk group* Γ is a group of automorphisms of T generated by four automorphisms denoted a, b, c, d. The automorphism a rigidly swaps the two subtrees T_0, T_1 rooted at the first level of T, while b, c and d leave the first level intact and are defined recursively by

$$b = (a, c), c = (a, d), d = (1, b).$$

This notation means that, for instance, b acts on T_0 (the subtree rooted at the leftmost vertex of the first level) like a acts on T, while acting on T_1 like c does on T. This is best illustrated with pictures:







For any vertex v of T, there is a unique path to the root. The *level* of v is the number of edges in this unique path to the root. The stabilizer of v is denoted by $St(v)$ and the *nth level stabilizer* $St(n)$ is the intersection of the stabilizers of all vertices of level n.

Writing any element $u \in St(1)$ as $u = (u_0, u_1)$, it is easy to see that conjugation by a, induces a 'swapped' action: $aua = (u_1, u_0)$. It is also not hard to see that $\{1, b, c, d\}$ is a Klein 4 group and that $St(1)$ is generated by $\{b, c, b^a, c^a\}$.

Remark. The group Γ is infinite. In fact, $St(1) \twoheadrightarrow \Gamma$. To see this, define $\varphi_0, \varphi_1 : St(1) \to \Gamma$ by $\varphi_0(g_0, g_1) = g_0$, $\varphi_1(g_0, g_1) = g_1$. Then $\varphi_0(b) = a$, $\varphi_0(b^a) = c$, $\varphi_0(c^a) = d$ and $\varphi_1(b^a) = a$, $\varphi_1(b) = c$, $\varphi_1(c) = d$ prove our claim.

This also shows that the map $\psi : St(1) \to \Gamma \times \Gamma$, $g \mapsto (g_0, g_1)$ is a monomorphism. Similarly, so is the map $\psi_n : St(n) \to \Gamma^{2^n}$ from the nth level stabilizer to the 2^n-fold direct power of Γ.

Proposition 4.7. *Every element of Γ has order a power of 2.*

Proof. This is not very hard but we omit it for conciseness. See [5], Chapter VIII for a proof. $\square$

Hence Γ is not in EG. This last proposition also means that Γ provides a solution to the general Burnside problem (recall that this asks whether all finitely generated torsion groups are finite). It was first answered by Golod and Shafarevich in 1964 using very different methods.

4.2.1 Growth of the Grigorchuk group

In order to show that Γ is amenable, we will show that it is of subexponential growth.

Throughout, we will abuse notation and write $l(g)$ to mean both the length of a shortest word representing g and for the length of g as a word.

Lemma 4.8. *For each $g \in St(3)$ we have $\sum_{i,j,k=0,1} l(g_{ijk}) \leq \frac{3}{4} l(g) + 8$, where g_{ijk} is the element obtained by reducing $\varphi_k(\varphi_j(\varphi_i(g)))$.*

Proof. If $g = 1$ the statement is trivial. Suppose then that $g \neq 1$ and let w be a word representing g such that $l(w) = l(g)$. For $x \in \{a, b, c, d\}$ and any word v, we denote by $|v|_x$ the number of instances of x in v. Then $|w|_a$ must be even as g stabilizes the first level of T and we have

$$\frac{l(g) - 1}{2} \leq |w|_{b,c,d} \leq \frac{l(g) + 1}{2}. \tag{1}$$

Now, $l(g) \geq 4$ since $axa \notin St(3)$ for any $x \in \{b, c, d\}$, and w is of the form $[x_0]ax_1a \ldots x_{2m-1}a[x_{2m}]$, of length between $4m-1$ and $4m+1$. Then

$$l(\varphi_0(w)) + l(\varphi_1(w)) \leq l(g) + 1 - |w|_d \text{ and } |\varphi_0(w)|_{c,d} + |\varphi_1(w)|_{c,d} = |w|_{b,c}. \tag{2}$$

To see this, take for example $w = dabac$, so $\varphi_0(w) = 1ca$, $\varphi_1(w) = bad$.

For any word v, the reduced form $\varphi_i(v)$ will be denoted by v_i. Define $\rho(v)$ to be the weighted number of reductions required to put v in reduced form. A reduction of the form $x_1x_2 \to x_3$ where $x_1, x_2, x_3 \in \{b, c, d\}$ has weight 1 and one of the form $xx \to 1$ for $x \in \{a, b, c, d\}$ has weight 2. Writing $\rho_1 := \rho(\varphi_0(w)) + \rho(\varphi_1(w))$, we obtain from 4.2.1

$$\begin{aligned} &l(w_0) + l(w_1) \leq l(g) + 1 - |w|_d - \rho_1, \\ &|w_0|_{c,d} + |w_1|_{c,d} \geq |w|_{b,c} - 2\rho_1. \end{aligned} \tag{3}$$

For example, take $w = cabadadabaca = (ac1bad, dab1ca) = (adad, dada)$, so $|w_0|_{c,d} = |w_1|_{c,d} = 2$, $\rho_1 = 2$ and $|w|_{b,c} = 4$.

Since $|v|_{b,c,d} = |v|_{b,c} + |v|_d$, equations 3 and 1 give

$$|w_0|_{c,d} + |w_1|_{c,d} \geq \frac{l(g)-1}{2} - |w|_d - 2\rho_1. \tag{4}$$

Repeating this process we obtain, using 4,

$$\begin{aligned} \sum_{i,j=0,1} l(w_{ij}) &\leq l(g) + 3 - |w|_d - \rho_1 - |w_0|_d - |w_1|_d - \rho_2, \\ \sum_{i,j=0,1} |w_{ij}|_d &\geq \frac{l(g)-1}{2} - |w|_d - 2\rho_1 - |w_0|_d - |w_1|_d - 2\rho_2, \end{aligned} \tag{5}$$

where ρ_2 is the sum of the weighted number of reductions $\varphi_j(w_i) \to w_{ij}$.

Repeating the procedure and using 5 yields

$$\sum_{i,j,k=0,1} l(w_{ijk}) \leq \frac{l(g)}{2} + 8 + \rho_1 + \rho_2.$$

There are now two cases, either *a)* $\rho_1 + \rho_2 \leq l(g)/4$ or *b)* $\rho_1 + \rho_2 > l(g)/4$.

a) $\sum_{i,j,k=0,1} l(w_{ijk}) \leq \frac{3l(g)}{4} + 8$, as required.

b) Equation 5 yields $\sum_{i,j=0,1} l(w_{ij}) \leq 3l(g)/4 + 3$. Now, for $i, j = 0, 1$ we have $l(w_{ij0}) + l(w_{ij1}) \leq l(w_{ij}) + 1$, so we conclude

$$\sum_{i,j,k=0,1} l(w_{ijk}) \leq \sum_{i,j=0,1} (l(w_{ij}) + 1) \leq \frac{3l(g)}{4} + 7,$$

as required.

To illustrate this, take as an example $w = (babadaba)^2$. Then

$$\begin{aligned} w &= (\varphi_0(w), \varphi_1(w)) = (ac1cac1c, cabacaba) \\ &= (w_0, w_1) =_{(\rho_1=6)} (1, cabacaba) \\ &=_{(\rho_2=0)} ((1,1), (acac, dada)) = (((1,1),(1,1)),((da,ad),(1b,b1))). \end{aligned}$$

Hence $\sum_{i,j,k=0,1} l(w_{ijk}) = 6 \leq 17 = \frac{3\cdot 12}{4} + 8 = \frac{3l(g)}{4} + 8$.

□

Theorem 4.9. *The group Γ has subexponential growth.*

Proof. For this proof we will use the following fact: if G is generated by a finite set S and H is a subgroup of index n in G, writing $\gamma(k) := |B_s(k)|$ and $\gamma_0 := |B_S(k) \cap H|$, we have

$$\gamma(k) \leq n\gamma_0(k + n - 1) \text{ for all } k \geq 0.$$

This follows from the Schreier rewriting system for finding generators of H.

Let Γ be generated by $S = \{a, b, c, d\}$. We will take $St(3)$ as our subgroup of finite index, so $\gamma(k) = |B_S(k)|$ and $\gamma_0(k) := |B_S(k) \cap St(3)|$. Notice that $|\Gamma : St(3)| = 2^7$ since $\Gamma/St(3)$ acts on the third level of T as $(\mathbb{Z}/2\mathbb{Z} \wr \mathbb{Z}/2\mathbb{Z}) \wr \mathbb{Z}/2\mathbb{Z}$ which has size 2^7.

Let $\omega := \lim_{k\to\infty} \gamma(k)^{1/k}$. Then for each $\varepsilon > 0$ there is some k_0 such that $(\omega - \varepsilon)^k \leq \gamma(k) \leq (\omega + \varepsilon)^k$ for any $k \geq k_0$. In particular, for every $k \geq 0$ we have

$$\gamma(k) \leq \gamma(k_0)(\omega + \varepsilon)^k. \tag{6}$$

Now, since $\psi_3 : St(3) \to \Gamma^8$ is injective, the previous lemma gives

$$\gamma_0(k) \leq \sum_X \gamma(k_1) \cdots \gamma(k_8) \tag{7}$$

where X is the set of 8-tuples $(k_1, \ldots, k_8) \in \mathbb{N}^8$ such that $k_1 + \ldots + k_8 \leq 3k/4 + 8$. Thus, by the fact quoted at the start of the proof and (7) we get

$$\gamma(k) \leq 2^7\gamma_0(k + 2^7 - 1) \leq 2^7 \sum_{Y_k} \gamma(k_1) \cdots \gamma(k_8)$$

where Y_k is the set of tuples $(k_1, \ldots, k_8) \in \mathbb{N}^8$ such that $k_1 + \ldots + k_8 \leq 3(k + 2^7 - 1)/4 + 8 =: N$. By a combinatorial 'stars and bars' argument, the size of Y_k is $\binom{N-1}{8}$ which is a polynomial in N and thus a polynomial $P(k)$ in k.

Using (6) we now have

$$\gamma(k) \leq 2^7\gamma(k_0)^8 P(k)(\omega + \varepsilon)^N$$

for all $k \geq 0$. Thus, since $(2^7\gamma(k_0)^8 P(k)(\omega+\varepsilon)^{3/4(2^7-1)+8})^{1/k}$ tends to 1 as k tends to ∞, we obtain

$$\omega = \lim_{k\to\infty} (\gamma(k))^{1/k} \leq \lim_{k\to\infty} ((\omega + \varepsilon)^{3/4k})^{1/k} = (\omega + \varepsilon)^{3/4}.$$

This holds for every $\varepsilon > 0$, so we must have $\omega \leq \omega^{3/4}$. Now notice that ω must be at least 1 as Γ is infinite. Thus $\omega = 1$ and so Γ has subexponential growth.

□

By Theorem 3.8, Γ is amenable, showing that there is a finitely generated group which is amenable but not elementary amenable.

References

[1] St. Banach and A. Tarski. Sur la décomposition des ensembles de points en parties respectivement congruentes. Fund. math. 6, 244-277 (1924)., 1924.

[2] B. Bekka, P. de la Harpe, and A. Valette. *Kazhdan's Property (T).* New Mathematical Monographs. Cambridge University Press, 2008.

[3] Ching Chou. Elementary amenable groups. *Ill. J. Math.*, 24:396–407, 1980.

[4] Mahlon M. Day. Amenable semigroups. *Ill. J. Math.*, 1:509–544, 1957.

[5] P. de la Harpe. *Topics in Geometric Group Theory.* Chicago Lectures in Mathematics. University of Chicago Press, 2000.

[6] R. I. Grigorchuk. Degrees of growth of finitely generated groups, and the theory of invariant means. *Mathematics of the USSR-Izvestiya*, 25(2):259, 1985.

[7] Alexander Lubotzky. *Discrete groups, expanding graphs and invariant measures. With an appendix by Jonathan D. Rogawski. Reprint of the 1994 original.* Basel: Birkhäuser, reprint of the 1994 original edition, 2010.

[8] I. Namioka. Folner's conditions for amenable semi-groups. *Math. Scand.*, 15:18–28, 1964.

[9] A.Yu. Ol'shanskii. An infinite group with subgroups of prime orders. *Math. USSR, Izv.*, 16:279–289, 1981.

[10] A.L.T. Paterson. *Amenability.* Mathematical surveys and monographs. American Mathematical Society, 1988.

[11] Terry Tao. Some notes on amenability. Accessed: 08/10/2013.

[12] Tery Tao. Amenability, the ping-pong lemma, and the Banach–Tarski paradox. Accessed: 08/10/2013.

[13] J. von Neumann. Zur allgemeinen Theorie des Maßes. *Fundam. Math.*, 13:73–116, 1929.

[14] Stan Wagon. *The Banach-Tarski paradox.* 1985.