COSET CORRECT MEANS ON GROUPS AND THE PROBABILITY THAT TWO ELEMENTS COMMUTE

ARMANDO MARTINO AND MOTIEJUS VALIUNAS

ABSTRACT. Amenable groups are those admitting an *invariant mean*—a finitely additive probability mean that assigns equal "weight" to any two translates of the same set.

We introduce *coset correct means* (*CCMs*), a class of finitely additive means that, for any subgroup, assigns equal weight to all its cosets, weakening and therefore generalising the notion of an invariant mean. We show that, unlike the case for invariant means, *every* group admits a CCM and give two constructions—one via the Ultrafilter Lemma and one via the Hahn–Banach Theorem—both relying on a Theorem of B. H. Neumann. Using CCMs, we define a degree of commutativity for arbitrary groups, measuring the "probability" that two random elements of a group commute. We prove that this degree of commutativity is independent of the choice of CCM and is positive precisely for finite-by-abelian-by-finite groups, recovering and unifying previous characterisations.

We also introduce a defect function that quantifies the failure of left invariance for finitely additive means, and define the defect of a group as the infimum of these. We then prove a dichotomy: the defect for a group is either 0 or 1, with 0 characterising amenable groups.

Contents

1. INTRODUCTION

The goal of this paper is to define finitely additive probability means for any group which give the "correct" answer for subgroups and their cosets, namely the reciprocal of the index of the subgroup. We call such a mean a coset correct mean, a CCM. We then use these means to define a probability that two elements of a group commute—the degree of commutativity—and thence deduce structural results about groups where this probability is positive. Such means are already seen to exist for amenable groups but this is well known to exclude many groups.

As a starting point for thinking about such means it is perhaps instructive to consider the following well-known Theorem:

$$\mathbb{R}^n \text{ is not the union of finitely many affine hyperplanes.}$$

There are various proofs of this fact but one of the more straightforward ways to verify this is to consider the Lebesgue measure; the union of finitely many affine hyperplanes has measure 0 and so cannot equal Euclidean space. This is straightforward up to constructing the Lebesgue measure, of course.

After some work one can verify the analogous result holds for any infinite field but the most general version of this type of result is due to B. H. Neumann.

Theorem 1.1 (Neumann [Neu54, Lemma 4.5])**.** *Suppose that* $G = \bigcup_{i=1}^{k} g_i H_i$ *for a group* G, *elements* $g_1, \ldots, g_k \in G$ *and subgroups* $H_1, \ldots, H_k \leq G$. *Then we have*

$$\sum_{i=1}^{k} \frac{1}{[G : H_i]} \geq 1,$$

where the reciprocal $\frac{1}{[G:H_i]}$ *is understood to be zero whenever* H_i *has infinite index in* G.

In particular, no coset of a finite index subgroup of G *can be covered by finitely many cosets of infinite index subgroups.*

Our contribution is to show that this result actually allows one to construct a CCM, a mean on G whereby all subsets are given some non-negative mean and for which all cosets of a given subgroup are given the same mean. Thus, just as in the case of affine hyperplanes in Euclidean space, no group can be realised as a finite union of cosets of infinite index subgroups because the mean of the latter is always zero. We give two constructions of our CCMs, one using the Ultrafilter Lemma and one using the Hahn–Banach Theorem, but both of these constructions rely on the result of B. H. Neumann above to verify consistency before appealing to some version of the axiom of choice to extend the mean to all subsets of the group.

In order to motivate the definition of a coset correct mean we recall the definition of an amenable group as a group admitting a finitely additive probability mean. Equivalently, an amenable group is one which does not admit a paradoxical decomposition in the sense of the Banach–Tarski paradox. We will largely take the former point of view, although there are many equivalent definitions. Precisely,

Definition. A group G is said to be *amenable* if there exists a function, $\mu \colon \mathcal{P}(G) \to [0,1]$, satisfying the following:

    (i) $\mu(G) = 1$;
    (ii) $\mu(A \sqcup B) = \mu(A) + \mu(B)$, for all disjoint subsets $A, B \subseteq G$;
    (iii) $\mu(gA) = \mu(A)$, for all $A \subseteq G$ and all $g \in G$.

The function μ is called an *invariant mean* on G. It is then clear that:

Corollary. If G is amenable and μ is an invariant mean on G then for any subgroup H of G and any $g \in G$ we have

$$\mu(gH) = \frac{1}{[G : H]},$$

where the right hand side is understood to be zero when H has infinite index.

We recall that not all groups are amenable. In particular, any group containing a non-abelian free group is not amenable.

The definition of a coset correct mean (CCM) weakens condition (iii) in the Definition above and replaces it with the statement of the Corollary. Equivalently, the definition of a CCM simply takes the Definition above of amenability and only asks that left invariance is enjoyed by subgroups (and their cosets). As a minor note, left invariance on subgroups is equivalent to right invariance on subgroups, see Remark 3.2.

Definition 1.2 (see Definition 3.1)**.** Let G be a group. A *coset correct mean* (*CCM*) on G is a function $\mu \colon \mathcal{P}(G) \to [0,1]$ satisfying the following properties:

    (i) $\mu(G) = 1$ (*probability mean*);
    (ii) if $A, B \subseteq G$ are disjoint, then $\mu(A \sqcup B) = \mu(A) + \mu(B)$ (*finitely additive*);
    (iii) $\mu(gH) = \frac{1}{[G:H]}$ for all subgroups $H \leq G$ and all $g \in G$ (*left invariant on subgroups*).

As always, $\frac{1}{[G:H]}$ is understood to be zero when H has infinite index.

Our first main result—for which we give two proofs, using two weaker versions of the Axiom of Choice—is then as follows:

Theorem 1.3 (see Theorems 3.4 and A.3). *Every group admits a CCM.*

As an application of this construction, we use this to measure the *degree of commutativity* of a group G. That is, we consider the set, $\mathrm{Comm}(G) := \{(g, h) \in G^2 : gh = hg\}$, and measure the size of this set in G^2.

For example, for a finite group G, we would take the uniform measure—count elements, in other words—and define $\mathrm{dc}(G) = \frac{|\mathrm{Comm}(G)|}{|G^2|}$. For a residually finite group G, we could measure the degree of commutativity by looking at increasingly large finite quotients and define $\mathrm{dc}_{\mathrm{RF}}(G)$ as the infimum of $\mathrm{dc}(G/N)$ over all finite quotients G/N—see Definition 5.1.

In the case of a group G equipped with a CCM μ, we first define the product mean $\mu^{(2)}$ on G^2, see Definition 4.1. From there we define the degree of commutativity of G with respect to μ as $\mathrm{dc}_\mu(G) = \mu^{(2)}(\mathrm{Comm}(G))$: see Definition 4.5.

Our results show that defining the degree of commutativity in various senses, this number is positive if and only if the group in question is finite-by-abelian-by-finite, or FAF. We recall that, given families of groups $\mathcal{A}$ and $\mathcal{B}$, a group G is said to be $\mathcal{A}$-by-$\mathcal{B}$ if G admits a normal subgroup N so that N is in $\mathcal{A}$ and G/N is in $\mathcal{B}$. For instance, using the core construction one can show that a group is virtually abelian (admits a finite index abelian subgroup) if and only if it is abelian-by-finite. We note that the apparent ambiguity in calling a group FAF is genuinely only apparent as the class of (finite-by-abelian)-by-finite groups equals the class of finite-by-(abelian-by-finite) groups. More pithily, $(\mathrm{FA})\mathrm{F} = \mathrm{FAF} = \mathrm{F}(\mathrm{AF})$: see Lemma 2.15.

We then characterise the groups having positive degree of commutativity in the following Theorems.

Theorem 1.4 (see Theorem 6.10). *Let G be a group equipped with a CCM μ. Then $\mathrm{dc}_\mu(G) > 0$ if and only if G is finite-by-abelian-by-finite. Moreover, $\mathrm{dc}_{\mathrm{CCM}}(G) := \mathrm{dc}_\mu(G)$ is a rational number that does not depend on the choice of μ.*

In the case of a residually finite group we note that one can use the Haar measure on the profinite completion and we show that this also leads to the same result, on noting that a residually finite FAF group is AF, i.e. virtually abelian.

Theorem 1.5 (see Theorem 7.2). *Let G be a residually finite group with profinite completion $\widehat{G}$, and let $\widehat{\mu}$ be the Haar measure on $\widehat{G}^2$. Then we have*

$$\widehat{\mu}(\mathrm{Comm}(\widehat{G})) = \widehat{\mu}(\overline{\mathrm{Comm}(G)}) = \mathrm{dc}_{\mathrm{RF}}(G) = \mathrm{dc}_{\mathrm{CCM}}(G).$$

Moreover, this number is positive if and only if G is virtually abelian.

We also look at this degree of commutativity when G is an amenable group and taking an invariant mean (not necessarily the product mean) on the square. In this setting we get:

Theorem 1.6 (see Theorem 8.4). *Let G be an amenable group, and let μ be a finitely additive left invariant probability mean on G^2. Then $\mu(\mathrm{Comm}(G)) > 0$ if and only if G is finite-by-abelian-by-finite.*

Finally, since we are considering finitely additive probability means on groups which are invariant on some family of subsets it makes sense to define the *left defect* of such a probability mean μ on a group G as

$$\delta_\ell(\mu) = \sup\{|\mu(gA) - \mu(A)| : g \in G, A \subseteq G\} \in [0, 1].$$

We then define the (*left*) *defect* of G as

$$\delta_G := \inf_\mu \delta_\ell(\mu),$$

as in Definition 9.1. We then get:

Theorem 1.7 (see Theorem 9.5). *Let G be a group. Then the defect δ_G of G is equal to $\delta_\ell(\mu)$ for some finitely additive probability mean μ on G. Moreover, δ_G equals either 0 or 1, and is 0 precisely when G is amenable.*

Therefore, while the defect provides a natural way to quantify the failure of (left) invariance in finitely additive means, its behaviour turns out to be unexpectedly rigid as it is always either 0 or 1. This dichotomy mirrors the amenability of the group (defect zero if and only if the group is amenable) and thus, while conceptually clean, the defect offers no finer gradation among non-amenable groups. In this sense, it is both striking and somewhat limited.

Results in the literature. Amenable groups were introduced by John von Neumann in 1920s [Neu29], in order to understand the group-theoretic nature of the Banach–Tarski Paradox. Since then, amenability has been widely studied: see [Bar18] and references therein. Throughout the years, many equivalent definitions of amenability have been introduced; the most relevant to us are the existence of invariant means—closely related to the CCMs studied here—and the Følner condition (Proposition 3.10), related to our construction of CCMs in Section 3.

Even though we deal with discrete groups in this paper, it is worth mentioning that such concepts make sense, more generally, for locally compact Hausdorff topological groups. Such groups G admit Borel measures, called *Haar measures*, that are left-invariant and countably additive, see [DS14] and references therein; however, the total Haar measure for the whole G is infinite unless G is compact. A notion of amenability for (locally compact Hausdorff) topological groups has also been extensively studied [Wil03].

The degree of commutativity for finite groups was first introduced by Erdős and Turán in 1960s [ET68], and has been widely studied since then [Gal70, Gus73, Rus79, Neu89]. This has been later generalised to profinite groups by Lévai and Pyber [LP00] and, more generally, to compact groups by Hofmann and Russo [HR12]. In there, the degree of commutativity $\mathrm{dc}_{\mathrm{Cpt}}(G)$ of a compact group μ was defined analogously to our definition of $\mathrm{dc}_\mu(G)$, by taking μ to be the Haar measure on G. It was shown for any compact G, the number $\mathrm{dc}_{\mathrm{Cpt}}(G)$ is rational, and positive if and only if G has an open abelian subgroup; this can be therefore viewed as a topological version of Theorem 1.4 (or Theorem 1.6) for compact groups.

In [AMV17], Antolín, Ventura and the first author introduced the study of the degree of commutativity $\mathrm{dc}_S(G)$ of a finitely generated group $G = \langle S \rangle$, by computing the proportion of commuting elements in a ball $B_{G,S}(n)$ of radius n in the Cayley graph $\mathrm{Cay}(G, S)$, and taking the limit (superior) as $n \to \infty$. They showed that $\mathrm{dc}_S(G)$ is zero for non-elementary hyperbolic G, and similar results were later obtained by the second author for non-(virtually abelian) groups G that are graph products [Val19b] and higher rank solvable Baumslag–Solitar groups [Val19a]. The main obstacle in these cases, however, was that for many groups, the ball counting measure does not measure the index of subgroups correctly.

In order to avoid such problems, Tointon [Toi20] studied sequences of measures that measure indices of subgroups correctly (in the limit), as well as the degree of commutativity for such sequences. This is the case, for instance, for the measures obtained from random walks (in a finitely generated group), as well as Følner sequences (in a countable amenable group); in particular, a version of Theorem 1.3 for finitely generated groups follows from the results in [Toi20]. A version of Theorem 1.4 was also proved there, and this has been later extended by Tointon, Ventura and the authors [MTVV21] to the *degree of k-step nilpotence* of a group G— the probability that the iterated commutator $[x_0, x_1, \ldots, x_k]$ is trivial for "random" elements $x_0, \ldots, x_k \in G$.

Structure of the paper. In Section 2 we give some preliminaries on Boolean rings, filters, ultrafilters and ultralimits, and basic structural results for finite-by-abelian-by-finite groups. We prove Theorem 1.3 in Section 3 using the Ultrafilter Lemma, and in the Appendix using the Hahn–Banach Theorem. We prove Theorem 1.4 in Sections 5 and 6, Theorem 1.5 in Section 7, Theorem 1.6 in Section 8, and Theorem 1.7 in Section 9. Finally, in Section 10 we discuss an alternative generalisation of the degree of commutativity to infinite groups, by measuring sets of conjugacy class representatives.

2. Preliminaries

Boolean rings and filters. Recall that a *Boolean ring* is a ring in which $x^2 = x$ for every element x of the ring. Note that any Boolean ring is commutative and has characteristic 2. Our Boolean rings will always be unital.

The main examples of Boolean rings arise as powersets. For any set X, the powerset of X, $\mathcal{P}(X)$, is a Boolean ring where the addition is symmetric difference and the multiplication is intersection: for $A, B \subseteq X$, the operations are

$$A \oplus B \coloneqq (A \cup B) - (A \cap B),$$
$$AB \coloneqq A \cap B.$$

It is straightforward to see that this defines a ring structure on noting the bijection between $\mathcal{P}(X)$ and $(\mathbb{Z}/2\mathbb{Z})^X$; the operations in $\mathcal{P}(X)$ correspond to pointwise addition and multiplication in $(\mathbb{Z}/2\mathbb{Z})^X$. The set X is the multiplicative identity of the ring and $\varnothing$ is the additive identity.

It is also worth noting that a Boolean ring is equivalent to a Boolean algebra, meaning that union, complementation and set difference can be expressed in terms of the ring operations:

$$A \cup B = A \oplus B \oplus AB,$$
$$A^c = 1 \oplus A,$$
$$A - B = A \oplus AB.$$

Hence any (unital) subring of $\mathcal{P}(X)$ is also closed under finite unions, complementation and set difference.

Definition 2.1. Let X be a nonempty set. A *filter* on X is a collection $\mathcal{F} \subseteq \mathcal{P}(X)$ of subsets of X such that:

  (i) $\varnothing \notin \mathcal{F}$, $X \in \mathcal{F}$;
  (ii) if $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$;
  (iii) if $A \in \mathcal{F}$ and $A \subseteq B \subseteq X$, then $B \in \mathcal{F}$.

A filter $\mathcal{F}$ on X is called *principal* if there exists a subset $A \subseteq X$ such that

$$\mathcal{F} = \{B \subseteq X : A \subseteq B\}.$$

In this case, $\mathcal{F}$ is said to be *generated* by A, and we may write $\mathcal{F} = \langle A \rangle$.

A filter $\mathcal{U}$ on X is called an *ultrafilter* if it is a maximal proper filter on X, that is, there is no filter $\mathcal{F}'$ on X satisfying $\mathcal{F}' \supsetneq \mathcal{U}$. Equivalently, a filter $\mathcal{U}$ is an ultrafilter if for every subset $A \subseteq X$, either $A \in \mathcal{U}$ or $X - A \in \mathcal{U}$, but not both.

Remark 2.2. Note that an ultrafilter is non-principal if and only if it contains no finite subsets of X. Hence all ultrafilters on finite sets are principal. A principal ultrafilter is always generated by a singleton.

Given any infinite set X, the Fréchet filter on X is the collection of all co-finite sets (sets with finite complement). An ultrafilter on X is non-principal if and only if it contains the Fréchet filter.

The following is standard.

Lemma 2.3. *Let X be a set and $\mathcal{F} \subseteq \mathcal{P}(X)$. Define $\mathcal{F}^c \coloneqq \{A \subseteq X : A^c \in \mathcal{F}\}$. Then,*

  (i) *$\mathcal{F}$ is a filter if and only if $\mathcal{F}^c$ is a proper ideal.*
  (ii) *$\mathcal{F}$ is a principal filter if and only if $\mathcal{F}^c$ is a principal ideal (generated by a single element).*
  (iii) *$\mathcal{F}$ is an ultrafilter if and only if $\mathcal{F}^c$ is a maximal ideal.*

Remark 2.4. Since $\mathcal{P}(X)$ is a Boolean ring, any quotient of $\mathcal{P}(X)$ is also a Boolean ring. In particular, any quotient which is a field satisfies $x^2 = x$ for every element and so must be the field with two elements. Therefore, any maximal ideal has index 2.

The existence of (non-principal) ultrafilters require a weak version of the Axiom of Choice. For those not worried about relative strengths, the following is an easy consequence of Zorn's Lemma.

Theorem 2.5 (Ultrafilter Lemma). *Let X be a set and $\mathcal{F}$ a filter on X. Then $\mathcal{F}$ is contained in an ultrafilter.*

Remark 2.6. Note that the Ultrafilter Lemma gives us the existence of non-principal ultrafilters, since one can apply it to the Fréchet filter.

We make the following definitions.

Definition 2.7. Let X be a set and $S \subseteq \mathcal{P}(X)$. Then S is said to have the *finite intersection property* if the intersection of finitely many elements of S is never the empty set.

We will also say that a subset $S \subseteq \mathcal{P}(X)$ generates a filter if the following set is a filter

$$\mathcal{F}_S := \left\{ A \subseteq X : \bigcap_{i=1}^{k} S_i \subseteq A \text{ for some } S_1, \ldots, S_k \in S \right\}.$$

Thus $\mathcal{F}_S$ is the smallest "upwards closed" subset of $\mathcal{P}(X)$ containing all the finite intersections of elements of S. Saying that $\mathcal{F}_S$ is a filter is equivalent to saying that S has the finite intersection property. Hence we obtain the following, which is the main tool we will use to construct filters, and thereby ultrafilters.

Proposition 2.8. *Let X be a set and $S \subseteq \mathcal{P}(X)$. Then S generates a filter on X if and only if it satisfies the finite intersection property.*

In this case, by the Ultrafilter Lemma, S will be contained in some ultrafilter on X.

Ultralimits. The notions of a limit via filters and ultrafilters are fairly general but we shall be concerned with the following more restrictive notion. Throughout this one should regard X as an indexing set.

Definition 2.9. Let X be a set, $f : X \to \mathbb{R}$ a function and $\mathcal{F}$ a filter on X. Then for $a \in \mathbb{R}$, we say that the *limit* of f with respect to $\mathcal{F}$ is a, written $\lim_{\mathcal{F}} f = a$, if for all $0 < \epsilon \in \mathbb{R}$,

$$\{x \in X : |f(x) - a| < \epsilon\} \in \mathcal{F}.$$

In the case where $\mathcal{F}$ is an ultrafilter, this is called the *ultralimit* of f with respect to $\mathcal{F}$.

Remark 2.10. We are not guaranteed that this limit exists, but when it does it is unique since the topology on $\mathbb{R}$ is Hausdorff.

Remark 2.11. We note that Definition 2.9 can also be obtained in the following way; let Y be a topological space and $\mathcal{F}_Y$ a filter on Y. Then we say that $\mathcal{F}_Y \to y \in Y$ if every (open) neighbourhood of y is an element of $\mathcal{F}_Y$.

If $f : X \to Y$ is a function and $\mathcal{F}_X$ is a filter on X then the push-forward of $\mathcal{F}_X$ is $f^*\mathcal{F}_X := \{A \subseteq Y : f^{-1}(A) \in \mathcal{F}_X\}$. This is a filter on Y (and an ultrafilter when $\mathcal{F}_X$ is an ultrafilter).

The notion of limit in Definition 2.9 is the same as saying that $f^*\mathcal{F}_X \to y$. However, the formulation in Definition 2.9 suits us better since we will keep the set X constant and consider various functions f whose limits are all defined via the same ultrafilter on X. In particular, we will be able to utilise an algebra of limits—Lemma 2.12.

The following is straightforward to prove:

Lemma 2.12. *Let X be a set and $\mathcal{U}$ an ultrafilter on X.*

   (i) *If $f : X \to \mathbb{R}$ is bounded, then $\lim_{\mathcal{U}} f$ exists and is unique.*

  (ii) *If $f, g : X \to \mathbb{R}$ and $a, b \in \mathbb{R}$, then $\lim_{\mathcal{U}}(af + bg) = a \lim_{\mathcal{U}} f + b \lim_{\mathcal{U}} g$, whenever these limits exist.*

 (iii) *If $\mathcal{U}$ is non-principal then $\lim_{\mathcal{U}} f$ is a limit point of f, where $y \in \mathbb{R}$ is called a* limit point *of f if $f^{-1}(O)$ is infinite for any open neighbourhood O of y.*

 (iv) *Conversely, if y is a limit point of f, then there is some non-principal ultrafilter $\mathcal{U}$ such that $\lim_{\mathcal{U}} f = y$.*

  (v) *If $\mathcal{U}$ is principal and hence generated by some $\{x_0\}$, then $\lim_{\mathcal{U}} f = f(x_0)$.*

Remark 2.13. More generally, limits via filters always obey the algebra of limits when these exist and make sense. That is, these limits preserve sums and products in the usual way.

Finite-by-abelian-by-finite (FAF) groups. We end these preliminaries by a short discussion of finite-by-abelian-by-finite groups, which we shall be analysing in significant detail. For now, we simply justify the lack of bracketing in the terminology.

As in the introduction, given classes $\mathcal{A}$ and $\mathcal{B}$ of groups, a group G is said to be $\mathcal{A}$-by-$\mathcal{B}$ if G admits a normal subgroup N so that N is in $\mathcal{A}$ and G/N is in $\mathcal{B}$. We also recall that a group is called virtually $\mathcal{A}$ if it admits a finite index subgroup in $\mathcal{A}$.

Remark 2.14. If $\mathcal{A}$ is a class of groups closed under taking (finite index) subgroups, then a virtually $\mathcal{A}$ group is the same as an $\mathcal{A}$-by-finite group, as the following construction shows. For a group G and a subgroup $H \leq G$, we define the core of H (in G) as $\mathrm{Core}_G(H) := \bigcap_{g \in G} H^g$—the intersection of all the conjugates of H in G; thus $\mathrm{Core}_G(H)$ is the largest normal subgroup of G contained in H. Crucially, if H is a finite index subgroup of G then $\mathrm{Core}_G(H)$ is also a finite index subgroup of G which is now normal. Hence a virtually $\mathcal{A}$ group is the same as an $\mathcal{A}$-by-finite group, as claimed. For instance, this will be the case when the class $\mathcal{A}$ consists of abelian groups.

By a *FAF group* we mean a finite-by-abelian-by-finite group. The following result shows that this definition does not depend on the chosen bracketing of the phrase "finite-by-abelian-by-finite". Recall that a group is *residually finite* if the intersection of all its (normal) finite index subgroups is trivial.

Lemma 2.15 (Equivalence of bracketings for FAF groups: $(FA)F = FAF = F(AF)$). *Let G be a group. Then G is (finite-by-abelian)-by-finite if and only it is finite-by-(abelian-by-finite). Moreover, if a FAF group G is either finitely generated or residually finite, then G is abelian-by-finite.*

Proof. Suppose first that G is finite-by-(abelian-by-finite). Then we have short exact sequences of groups

$$1 \to N \to G \xrightarrow{\pi_H} H \to 1 \qquad \text{and} \qquad 1 \to A \to H \xrightarrow{\pi_Q} Q \to 1$$

with N, Q finite and A abelian. Viewing A as a normal subgroup of H and restricting the first sequence to $G_0 := \pi_H^{-1}(A)$, we get a short exact sequence $1 \to N \to G_0 \to A \to 1$ and so G_0 is finite-by-abelian, while $G/G_0 \cong Q$ is finite. This shows that G is (finite-by-abelian)-by-finite.

Conversely, let G be a (finite-by-abelian)-by-finite group. Then we have a short exact sequence

$$1 \to H \to G \to Q \to 1$$

with H finite-by-abelian and Q finite. Note that the fact that H is finite-by-abelian means precisely that the derived subgroup $H' = [H, H]$ is finite. Note that H' is characteristic in H, so since H is normal in G it follows that H' is normal in G as well. Since H/H' is abelian and G/H is finite, it follows that G/H' is abelian-by-finite. Thus G is finite-by-(abelian-by-finite), as required.

Suppose now that G is a finitely generated FAF group. Then G has a finite index finite-by-abelian normal subgroup H, and since G is finitely generated, so is H; let $\{s_1, \ldots, s_k\}$ be a finite generating set of H. Since H is finite-by-abelian it follows that $|H'| < \infty$, and in particular the conjugacy class of each s_i, $\{g^{-1} s_i g : g \in G\} = \{s_i[s_i, g] : g \in G\}$, is contained in $s_i H'$ and so is finite. It follows that $C_H(s_i)$ has finite index in H for $1 \leq i \leq k$, and therefore the centre $Z(H) = \bigcap_{i=1}^k C_H(s_i)$ also has finite index in H. Since $Z(H)$ is characteristic in H and H is normal in G, it follows that $Z(H)$ is normal in G, and since both $[H : Z(H)]$ and $[G : H]$ are finite, it follows that so is $[G : Z(H)]$. As $Z(H)$ is clearly abelian, it follows that G is abelian-by-finite, as required.

Finally, suppose that G is a residually finite FAF group. Then G has a finite normal subgroup N such that G/N is abelian-by-finite. Since N is finite and G is residually finite, G contains a finite index subgroup H such that $H \cap N = \{1\}$. It follows that the composite $H \hookrightarrow G \twoheadrightarrow G/N$ is injective, implying that H is isomorphic to a subgroup of G/N. Since G/N is abelian-by-finite, it follows that H has a finite index abelian subgroup A_0, which is thus a finite index abelian subgroup of G. In particular, G has a normal finite index abelian subgroup $A = \mathrm{Core}_G(A_0)$, and so is abelian-by-finite, as required. $\qquad\square$

Remark 2.16. It can be seen from the proof that a FAF group G contains a finite index normal subgroup H that has a finite derived subgroup H'. The fact that H' is finite then implies that so is $\mathrm{Aut}(H')$, and in particular the kernel of the map $H \to \mathrm{Aut}(H')$ sending $h \in H$ to conjugation by h has finite index in H. But this kernel is precisely $C_H(H')$; moreover, since H' is characteristic in H, so is $C_H(H')$. In particular, $C_H(H')$ is characteristic and of finite index in H, and H is normal and of finite index in G, implying that $C_H(H')$ is a normal finite index subgroup of G. Moreover, $C_H(H')/Z(H')$ is isomorphic to the subgroup $C_H(H')H'/H'$ of H/H', and in particular is abelian. It follows, by considering $N_0 = Z(H')$ and $H_0 = C_H(H')$, that any FAF group G has a finite normal subgroup $N_0 \trianglelefteq G$ and a finite index normal subgroup $H_0 \trianglelefteq G$ containing N_0 such that N_0 is central in H_0 and H_0/N_0 is abelian.

That is, any FAF group is virtually a 2-step nilpotent group whose derived subgroup (which is central) is finite.

3. Construction of coset correct means

The goal of this section is to prove that every group admits a CCM, defined below. Our argument here is via ultrafilters, using the Ultrafilter Lemma (Theorem 2.5), but we also give an argument using the Hahn–Banach Theorem (Theorem A.1) in the Appendix.

Definition 3.1. Let G be a group. A *coset correct mean* (*CCM*) on G is a function $\mu \colon \mathcal{P}(G) \to [0, 1]$ satisfying the following properties:

   (i) $\mu(G) = 1$ (*probability mean*);
   (ii) if $A, B \subseteq G$ are disjoint, then $\mu(A \sqcup B) = \mu(A) + \mu(B)$ (*finitely additive*);
   (iii) $\mu(gH) = \frac{1}{[G:H]}$ for all subgroups $H \leq G$ and all $g \in G$ (*left invariant on subgroups*).

As always, $\frac{1}{[G:H]}$ is understood to be zero when H has infinite index.

More generally, let $\mathcal{R}$ be a subring of the Boolean ring $\mathcal{P}(G)$. We similarly define a *coset correct mean* (*CCM*) on $\mathcal{R}$ to be a function $\mu \colon \mathcal{R} \to [0, 1]$ satisfying the conditions (i)–(iii) above whenever μ is defined on the corresponding subsets of G.

Remark 3.2. We also note that a *left* CCM must also be a *right* CCM since

$$\mu(Hg) = \mu(g(g^{-1}Hg)) = \frac{1}{[G:H^g]} = \frac{1}{[G:H]} = \mu(H).$$

That is, left invariance on cosets is equivalent to right invariance on cosets.

The strategy in our construction is the following: given a group G, we denote by $\mathcal{P}_{\mathrm{fin}}(G)$ the set of finite subsets of G, which will serve as an "indexing set". We then define functions for each $A \subseteq G$, $f_A \colon \mathcal{P}_{\mathrm{fin}}(G) \to [0, 1]$, via

$$f_A(F) = \frac{|A \cap F|}{|F|}.$$

We then take an ultrafilter $\mathcal{U}$ on $\mathcal{P}_{\mathrm{fin}}(G)$ and define $\mu(A) = \lim_{\mathcal{U}} f_A$. It is immediate that μ has total weight one, since f_G is the constant function and it is also clear that μ is finitely additive since $f_{A \sqcup B} = f_A + f_B$ (using the algebra of ultralimits, Lemma 2.12).

All the remains to show is that μ is left invariant on cosets. This is not true for all ultrafilters so the goal is to find some (ultra)-filter where it is true. However, the property that μ is left invariant on cosets determines certain subsets of $\mathcal{P}_{\mathrm{fin}}(G)$ that would need to belong to the (ultra)-filter and hence our strategy becomes simply showing that these sets generate a filter, as in Proposition 2.8. We actually show this in some strong sense in the following:

Lemma 3.3. *Given a subgroup $H \leq G$ and a real number $\varepsilon > 0$, we define*

$$\mathcal{F}_\varepsilon(H) \coloneqq \left\{ F \in \mathcal{P}_{\mathrm{fin}}(G) : \left| \frac{|gH \cap F|}{|F|} - \frac{1}{[G:H]} \right| < \varepsilon \text{ for all } g \in G \right\}.$$

Then the collection $\mathcal{G} = \mathcal{G}_G \coloneqq \{\mathcal{F}_\varepsilon(H) : H \leq G, \varepsilon > 0\}$ generates a filter on $\mathcal{P}_{\mathrm{fin}}(G)$, the set of finite subsets of G.

Proof. Note that, by Proposition 2.8, it is enough to show that $\mathcal{G}$ satisfies the finite intersection property, Definition 2.7. Thus, let $\varepsilon_1, \ldots, \varepsilon_k > 0$ and $H_1, \ldots, H_k \leq G$. We then need to show that $\bigcap_{i=1}^{k} \mathcal{F}_{\varepsilon_i}(H_i) \neq \varnothing$.

After possibly permuting the indices, we may assume that, for some m with $0 \leq m \leq k$, we have $[G : H_i] = \infty$ if $i \leq m$ and $[G : H_i] < \infty$ if $i > m$. Then $H := \bigcap_{i=m+1}^{k} H_i$ is a finite index subgroup of G. Let $N = r[G : H]$ for some integer $r \geq 1$ large enough so that $\frac{1}{N} < \varepsilon_i$ for $1 \leq i \leq m$. We aim to construct a subset $F \in \mathcal{F}$ of cardinality N that contains at most one element in each left coset of H_i for $i \leq m$, and exactly r elements in each left coset of H. This will imply that for each $g \in G$, we have $\frac{|gH_i \cap F|}{|F|} \leq \frac{1}{N} < \varepsilon_i$ for $i \leq m$, and $\frac{|gH_i \cap F|}{|F|} = \frac{r[H_i:H]}{r[G:H]} = \frac{1}{[G:H_i]}$ for $i > m$, and so $F \in \bigcap_{i=1}^{k} \mathcal{F}_{\varepsilon_i}(H_i)$, as required.

Let $H = g_1 H, \ldots, g_n H$ be all the left cosets of H, where $n = [G : H]$. We construct the subset $F = \{f_1, \ldots, f_N\}$, where $f_{ir+j} \in g_{i+1} H$ for $0 \leq i \leq n - 1$ and $1 \leq j \leq r$, inductively as follows. Having constructed $\{f_1, \ldots, f_{s-1}\}$ for some $s \geq 1$, let $f_s \in g_i H$, where $i = \lceil \frac{s}{r} \rceil$, be such that $f_s H_j \neq f_t H_j$ for any $j \leq m$ and $t < s$. Such a choice is possible: indeed, if it wasn't then we would have

$$g_i H \subseteq \bigcup_{j=1}^{m} \bigcup_{t=1}^{s-1} f_t H_j,$$

i.e. $g_i H$ could be covered by a finite union of cosets of its infinite index subgroups of G, which is impossible by Theorem 1.1 since H has finite index in G.

Thus the subset $F = \{f_1, \ldots, f_N\} \in \bigcap_{i=1}^{k} \mathcal{F}_{\varepsilon_i}(H_i)$ exists, as required. $\qquad\square$

Theorem 3.4. *The Ultrafilter Lemma implies that every group G admits a CCM.*

Proof. Let $\mathcal{U}$ be any ultrafilter containing $\mathcal{G}_G$ from Lemma 3.3, the existence of which follows from the Ultrafilter Lemma (Theorem 2.5). As above, we have functions for each $A \subseteq G$, $f_A : \mathcal{P}_{\text{fin}}(G) \to [0, 1]$ defined by

$$f_A(F) = \frac{|A \cap F|}{|F|}.$$

We then set,

$$\mu(A) := \lim_{\mathcal{U}} f_A.$$

From the remarks above, it is clear that this is a finitely additive probability mean. Moreover, by construction, $\mu(gH) = \mu(H)$ for all subgroups H and all $g \in G$. Hence μ is a CCM. $\qquad\square$

Remark 3.5. Note that as long as G is infinite, for any $\varepsilon > 0$ the subset $\mathcal{F}_\varepsilon(\{1\}) \in \mathcal{G}$ does not contain any $F \in \mathcal{P}_{\text{fin}}(G)$ of cardinality $\leq \frac{1}{\varepsilon}$. This implies that the filter constructed above is non-principal.

Alternatively, if $\mathcal{U}$ is a principal ultrafilter on $\mathcal{P}_{\text{fin}}(G)$ then it is generated by some fixed $F_0 \in \mathcal{P}_{\text{fin}}(G)$; that is, $\mathcal{U} = \langle F_0 \rangle$. Then,

$$\lim_{\langle F_0 \rangle} \frac{|F_0 \cap F|}{|F|} = \frac{|F_0 \cap F_0|}{|F_0|} = 1.$$

But F_0 is a disjoint union of cosets of the trivial subgroup and so a CCM will assign it density $\frac{|F_0|}{[G:\{1\}]} = \frac{|F_0|}{|G|}$. Hence $\mathcal{U} = \langle F_0 \rangle$ can only define a CCM on G if $F_0 = G$ is finite.

Thus we are observing that, when G is infinite, our construction produces a non-principal ultrafilter and that *any* ultrafilter used to construct a CCM must also be non-principal.

Remark 3.6. Instead of the subsets $\mathcal{F}_\varepsilon(H)$, given any $H \leq G$, $g \in G$ and $\varepsilon > 0$ one may consider the subset $\mathcal{F}'_\varepsilon(gH) := \left\{ F \in \mathcal{P}_{\text{fin}}(G) : \left| \frac{|gH \cap F|}{|F|} - \frac{1}{[G:H]} \right| < \varepsilon \right\}$. Since clearly $\mathcal{F}'_\varepsilon(gH) \supseteq \mathcal{F}_\varepsilon(H)$ for all H, g and ε, it follows that the collection $\mathcal{G}' := \{\mathcal{F}'_\varepsilon(gH) : H \leq G, g \in G, \varepsilon > 0\}$ generates a filter (that is a sub-filter of that generated by $\mathcal{G}$).

In a sense, the filter we construct in Lemma 3.3 is the uniform version of what we need, as it deals with all cosets of a given subgroup at the same time rather than one-by-one.

While the following results are known, we present elementary proofs that are particularly suited to our context. This is not merely for completeness: the combinatorial nature of our setting, especially the role played by finite sets and the Boolean algebra they generate, makes it natural to revisit these results from a more foundational perspective. The reader may wish to compare with [Hop21], particularly Theorems 2.6 and 5.9, where related arguments are given in a broader setting.

Proposition 3.7. *Let X be a set and let μ be a finitely additive probability mean on X. That is, $\mu\colon \mathcal{P}(X) \to [0,1]$ is a function such that (i) $\mu(X) = 1$ and (ii) $\mu(A \sqcup B) = \mu(A) + \mu(B)$ for all disjoint subsets A, B of X. Moreover, suppose that $\mu(F) = 0$ for any finite subset F of X.*

Then there exists an ultrafilter $\mathcal{U}$ on $\mathcal{P}_{\mathrm{fin}}(X)$ such that for all $A \subseteq X$,

$$\mu(A) = \lim_{\mathcal{U}} \frac{|A \cap F|}{|F|}.$$

Proof. For any $\epsilon > 0$ and any $A \subseteq X$, consider the sets

$$\mathcal{F}_\varepsilon(A) = \left\{ F \in \mathcal{P}_{\mathrm{fin}}(X) : \left| \frac{|F \cap A|}{|F|} - \mu(A) \right| < \varepsilon \right\} \subseteq \mathcal{P}_{\mathrm{fin}}(X).$$

Note that if these sets satisfy the finite intersection property (Definition 2.7) then they will be contained in some ultrafilter $\mathcal{U}$ on $\mathcal{P}_{\mathrm{fin}}(X)$. It then follows, by construction, that

$$\mu(A) = \lim_{\mathcal{U}} \frac{|A \cap F|}{|F|}.$$

Hence it is enough to show that these sets satisfy the finite intersection property.

To that end, let $A_1, \ldots, A_k \subseteq X$ and $\varepsilon_1, \ldots, \varepsilon_k > 0$. We then need to show that we have $\bigcap_{i=1}^k \mathcal{F}_{\varepsilon_i}(A_i) \neq \varnothing$.

Given any function $f\colon \{1, \ldots, k\} \to \{0,1\}$, consider $B_f = \bigcap_{i=1}^k A_i^{(f(i))} \subseteq X$, where $A_i^{(0)} = X - A_i$ and $A_i^{(1)} = A_i$. Then the sets B_f are pairwise disjoint and cover X. In other language, we are taking the Boolean subalgebra of $\mathcal{P}(X)$ generated by the A_i and considering the atoms, which are exactly the B_j.

Since A_i is the union of all those B_f for which $f(i) = 1$, it is clear by finite additivity of μ and the triangle inequality that we have

$$\mathcal{F}_{\varepsilon_i}(A_i) \supseteq \bigcap_{\substack{f\colon \{1,\ldots,k\} \to \{0,1\} \\ f(i) = 1}} \mathcal{F}_{\varepsilon_i/2^{k-1}}(B_f).$$

In particular, if the intersection $\bigcap_{f\colon \{1,\ldots,k\} \to \{0,1\}} \mathcal{F}_\varepsilon(B_f)$ is non-empty, where $\varepsilon = \frac{\min\{\varepsilon_1, \ldots, \varepsilon_k\}}{2^{k-1}}$, then so is $\bigcap_{i=1}^k \mathcal{F}_{\varepsilon_i}(A_i)$. We can therefore assume, without loss of generality, that the A_i are pairwise disjoint and cover X, and that all the ε_i are equal to a fixed $\varepsilon > 0$.

Now suppose that A_i are ordered so that we have $\mu(A_i) > 0$ for $i \leq m$ and $\mu(A_i) = 0$ for $i > m$. It then follows that A_i is infinite for all $i \leq m$. Pick an integer $N > \frac{1}{\varepsilon}$, and for $1 \leq i \leq m$ let $F_i' \subset A_i$ be any subset of cardinality $\lfloor \mu(A_i)N \rfloor$. Note that we have $\sum_{i=1}^m \mu(A_i) = \sum_{i=1}^k \mu(A_i) = \mu(X) = 1$ since the A_i are pairwise disjoint and cover X and since μ is finitely additive. It follows that

$$N - m = \sum_{i=1}^m (\mu(A_i)N - 1) \leq \sum_{i=1}^m \lfloor \mu(A_i)N \rfloor = \sum_{i=1}^m |F_i'| \leq \sum_{i=1}^m \mu(A_i)N = N.$$

This implies that we have subsets $F_i \subset A_i$, each obtained by adding at most one element to F_i', such that we have $\sum_{i=1}^m |F_i| = N$. Note that $\lfloor \mu(A_i)N \rfloor \leq |F_i| \leq \lfloor \mu(A_i)N \rfloor + 1$, and consequently $\big| |F_i| - \mu(A_i)N \big| \leq 1$.

We now claim that $F = \bigcup_{i=1}^m F_i$ belongs to $\mathcal{F}_\varepsilon(A_i)$ for each i. Indeed, for $i > m$ we have $F \cap A_i = \varnothing$ and $\mu(A_i) = 0$, implying that $\left| \frac{|F \cap A_i|}{|F|} - \mu(A_i) \right| = 0 < \varepsilon$. On the other hand, for

$i \leq m$ we have $F \cap A_i = F_i$ and therefore

$$\left| \frac{|F \cap A_i|}{|F|} - \mu(A_i) \right| = \left| \frac{|F_i|}{N} - \mu(A_i) \right| = \frac{||F_i| - \mu(A_i)N|}{N} \leq \frac{1}{N} < \varepsilon,$$

as claimed. Thus we indeed have $\bigcap_{i=1}^{k} \mathcal{F}_\varepsilon(A_i) \neq \varnothing$.

As we observed at the start, this is sufficient to prove the result. $\qquad\square$

Remark 3.8. It is easy to verify that if $X = G$ is a group and the ultrafilter constructed above satisfies the following Følner condition, namely: $\lim_{\mathcal{U}} \frac{|F \oplus gF|}{|F|} = 0$ for all $g \in G$, then the mean constructed will be a (left) invariant mean and G will be amenable. However, notice that in the construction we can always insist that our finite sets satisfy $F \cap gF = \varnothing$ and so $\lim_{\mathcal{U}} \frac{|F \oplus gF|}{|F|} = 2$, even if the mean itself is (left) invariant. We now do this explicitly.

Corollary 3.9. *Let G be an infinite group and μ a finitely additive probability mean on G with $\mu(F) = 0$ for all finite $F \subset G$. Then, using the construction of the finite approximating sets in the proof of Proposition 3.7, we may choose each finite set F so that, for any finite $S \subset G$ and any finite family of subsets $\mathcal{A} \subset \mathcal{P}(G)$,*

$$F \cap sF = \varnothing \quad \text{for all } 1 \neq s \in S,$$

while simultaneously

$$\left| \frac{|F \cap A|}{|F|} - \mu(A) \right| < \varepsilon \quad \text{for all } A \in \mathcal{A}.$$

In particular, there is an ultrafilter $\mathcal{U}$ on $\mathcal{P}_{\text{fin}}(G)$ such that

$$\mu(A) = \lim_{\mathcal{U}} \frac{|F \cap A|}{|F|} \quad \text{and} \quad \lim_{\mathcal{U}} \frac{|F \oplus gF|}{|F|} = 2 \ \forall g \in G.$$

Proof. The corollary follows directly from the construction in the proof of Proposition 3.7, with a minor refinement to ensure the disjointness condition for finite subsets, S of G. The existence of the ultrafilter then follows from Proposition 2.8 and Theorem 2.5.

As in the proof of the Proposition, for a finite family of sets $\mathcal{A} = \{A_1, \ldots, A_k\}$, we consider the Boolean algebra they generate and its atoms $B_1, \ldots, B_m$. For each atom B_j, we choose a finite number p_j of elements to include in F, where $p_j > 0$ only if B_j is infinite, and p_j is determined by the required approximation of $\mu(A_i)$.

The key observation is that each B_j from which we choose is infinite, so we are free to avoid finitely many "forbidden" elements at each stage. In particular, as we select the p_j elements from B_j, we can always ensure that $F \cap sF = \varnothing$ for all $1 \neq s \in S$: explicitly, we may pick elements one-by-one from the complement of $\bigcup_{s \in S}(F_0 \cup sF_0 \cup s^{-1}F_0)$, where $F_0 \subseteq F$ is the subset already constructed at a certain stage. Since the number of forbidden elements at each stage is finite, and each "operative" B_j is infinite, this process can be carried out without affecting the required counts p_j. Thus the resulting finite set F simultaneously approximates μ on all A_i and satisfies the disjointness conditions $F \cap sF = \varnothing$. $\qquad\square$

We then show that for an amenable group, any invariant mean may be obtained via an ultralimit that does satisfy the Følner condition. To do that, we use the classical Følner condtion.

Proposition 3.10 ([Føl55])**.** *Let G be a group. Then G is amenable if and only if for every finite subset $K \subseteq G$ and every $\varepsilon > 0$, there exists a non-empty finite subset $S \subseteq G$ such that*

$$\frac{|S \oplus gS|}{|S|} < \varepsilon \quad \text{for all } g \in K.$$

Proposition 3.11. *Let G be an infinite amenable discrete group and let $\mu \colon \mathcal{P}(G) \to [0,1]$ be a left invariant mean on G. Then there exists an ultrafilter $\mathcal{U}$ on $\mathcal{P}_{\text{fin}}(G)$ which both realises μ and satisfies the Følner criterion as follows,*

$$\mu(A) = \lim_{\mathcal{U}} \frac{|A \cap F|}{|F|} \quad \text{and} \quad \lim_{\mathcal{U}} \frac{|F \oplus gF|}{|F|} = 0 \ \forall g \in G.$$

Proof. Note that as G is an infinite amenable group, μ returns a density of 0 for any finite subset of G and hence we may use Corollary 3.9, which we shall do below.

To construct the ultrafilter we invoke Proposition 2.8 as before. It is clearly sufficient to verify the finite intersection property for the following family of subsets,

$$\mathcal{F}_\varepsilon(\mathcal{A}) \coloneqq \left\{ F \in \mathcal{P}_{\mathrm{fin}}(G) : \left| \frac{|F \cap A_i|}{|F|} - \mu(A_i) \right| < \varepsilon \,\forall A_i \in \mathcal{A} \right\}, \ \mathcal{A} = \{A_1, \ldots, A_r\} \subseteq \mathcal{P}(G), \ \varepsilon > 0,$$

$$\mathcal{E}_\varepsilon(K) \coloneqq \left\{ F \in \mathcal{P}_{\mathrm{fin}}(G) : \frac{|F \oplus gF|}{|F|} < \varepsilon \,\forall g \in K \right\}, \ K \in \mathcal{P}_{\mathrm{fin}}(G), \ \varepsilon > 0.$$

Notice that Corollary 3.9 already gives us that $\mathcal{F}_\varepsilon(\mathcal{A})$ is non-empty for any $\mathcal{A} = \{A_1, \ldots, A_r\} \subset \mathcal{P}(G)$ and any $\varepsilon > 0$. One easily observes that to verify the finite intersection property for this family, it is enough to show that

$$\mathcal{F}_\varepsilon(\mathcal{A}) \cap \mathcal{E}_\varepsilon(K) \neq \varnothing \quad \text{for any } \mathcal{A}, K \text{ and } \varepsilon > 0.$$

So let's assume that $\mathcal{A}$, K and ε are given and proceed to show that the intersection above is non-empty.

We first invoke Proposition 3.10 to find a finite subset S of G such that $|S \oplus gS| < \varepsilon|S|$ for all $g \in S$. Now we invoke Corollary 3.9 to find an $F_0 \in \mathcal{F}_\varepsilon(K\mathcal{A})$ such that $F_0 \cap gF_0 = \varnothing$ for all $1 \neq g \in S^{-1}S$, where $K\mathcal{A} \coloneqq \{kA_i : k \in K, A_i \in \mathcal{A}\}$. We claim that $F \coloneqq SF_0 = \bigcup_{g \in S} gF_0$ is in $\mathcal{F}_\varepsilon(\mathcal{A}) \cap \mathcal{E}_\varepsilon(K)$.

To see this, note that $F_0 \cap gF_0 = \varnothing$ for all $1 \neq g \in S^{-1}S$ implies that the sets gF_0, for $g \in S$, are pairwise disjoint. Hence $|F| = |S||F_0|$. In particular, $|F \oplus gF| \leq |S \oplus gS||F_0| < \varepsilon|S||F_0| = \varepsilon|F|$ for all $g \in K$. Therefore, $F \in \mathcal{E}_\varepsilon(K)$. To see that $F \in \mathcal{F}_\varepsilon(\mathcal{A})$, we calculate

$$\left| |F \cap A_i| - \mu(A_i)|F| \right| \leq \sum_{g \in S} \left| |gF_0 \cap A_i| - \mu(A_i)|F_0| \right| \qquad \text{as } |F| = |S||F_0|$$

$$= \sum_{g \in S} \left| |F_0 \cap g^{-1}A_i| - \mu(g^{-1}A_i)|F_0| \right| \qquad \text{as } \mu(A_i) = \mu(g^{-1}A_i)$$

$$< \sum_{g \in S} \varepsilon|F_0| \qquad \text{since } F_0 \in \mathcal{F}_\varepsilon(K\mathcal{A})$$

$$= \varepsilon|F|.$$

Hence F is also in $\mathcal{F}_\varepsilon(\mathcal{A})$ and we are done. $\qquad\square$

Remark 3.12. The construction above uses a two-scale idea: first, we choose a "core" set F_0 that approximates the mean on all relevant translates of the given subsets and satisfies strong disjointness conditions. Then we amplify F_0 by multiplying with a large Følner set S to form $F = SF_0$. This ensures that the relative boundary $\frac{|F \oplus gF|}{|F|}$ becomes arbitrarily small while preserving the density approximations, thanks to the invariance of μ. In essence, the invariance allows us to control averages over translates, and the use of S maintains the Følner property.

4. Product means and a degree of commutativity

As an application of the existence of a CCM, we use it to measure the "probability" that two elements of a group commute and eventually derive structural results about groups from this quantity. Our point of view is to understand the probability of $\mathrm{Comm}(G)$ in G^2 and so we extend our CCM from G to G^2 via a product mean.

Definition 4.1. Suppose that μ is a CCM on a group G. We may then define a mean $\mu^{(2)}$ on G^2, which we refer to as a *product mean*, by integrating: given $S \subseteq G^2$ we set

$$\mu^{(2)}(S) \coloneqq \int \mu(S_g) \,\mathrm{d}\mu(g),$$

where $S_g = \{h \in G : (g, h) \in S\}$; see [DS88, Chapter III].

Remark 4.2. There is certain asymmetry in this definition. Indeed, suppose that μ is a CCM on a group G. We could have also defined a mean $\widehat{\mu}^{(2)}$ on G^2 by setting

$$\widehat{\mu}^{(2)}(S) := \int \mu(\widehat{S}_g) \, d\mu(g),$$

where $\widehat{S}_g = \{h \in G : (h, g) \in S\}$.

These two definitions agree when G is finite (in which case $\mu^{(2)}(S) = \frac{|S|}{|G|^2} = \widehat{\mu}^{(2)}(S)$ for any $S \subseteq G^2$), but not when G is infinite. For instance, if $G = \{g_0, g_1, \ldots\}$ is countable and infinite, then we have $\mu^{(2)}(S) = 1$ but $\widehat{\mu}^{(2)}(S) = 0$, where $S = \{(g_i, g_j) \in G^2 : i < j\}$. A slightly more involved argument shows the same result when $G = \{g_i : i < \kappa\}$ is any infinite group, where κ is the initial ordinal of the cardinal $|G|$.

We note that if μ is a CCM, then so is $\mu^{(2)}$.

Lemma 4.3. *Let μ be a CCM on a group G. Then $\mu^{(2)}$ is a CCM on G^2.*

Proof. It is easy to see that $\mu^{(2)}$ is a finitely additive probability mean. It remains to show that it is left invariant on cosets. In particular, given $H \leq G^2$ and $(x, y) \in G^2$, we aim to show that $\mu((x, y)H) = \frac{1}{[G^2 : H]}$.

First we let $H_2 = \{h \in G : (1, h) \in H\}$ and $S = (x, y)H$. Note that H_2 is a subgroup of G, and that each S_g (for $g \in G$) is either empty, or a single coset of H_2 in G. If H_2 has infinite index, therefore, we conclude that $\mu^{(2)}((x, y)H) = 0$.

Otherwise, we have that H_2 is a finite index subgroup of G. In this case, set $\widehat{H}_1 = \pi_1(H)$, where $\pi_1 : G^2 \to G$ is the projection to the first coordinate. Note that $\widehat{H}_1 \times G$ is a finite union of left cosets of H (we can take these cosets to be $(1, g)H$, where g ranges over a left transversal of H_2 in G). Therefore, if H has infinite index in G^2 then $\widehat{H}_1$ has infinite index in G, and consequently

$$\mu^{(2)}((x, y)H) \leq \mu^{(2)}(x\widehat{H}_1 \times G) = \mu(x\widehat{H}_1) \cdot \mu(G) = 0 \cdot 1 = 0,$$

and hence $\mu^{(2)}((x, y)H) = 0$ whenever H has infinite index in G.

In the case where H has finite index in G, both $H_1 = \{g \in G \mid (g, 1) \in H\}$ and H_2 have finite index in G, and we have $H_1 \times H_2 \leq H$. Then every coset of $H_1 \times H_2$ in G has the same mean, $\mu(H_1)\mu(H_2) = \frac{1}{[G:H_1][G:H_2]}$, since μ is left invariant on cosets. In particular, since each coset of H will be a union of $[H : H_1 \times H_2] = \frac{[G^2:H_1 \times H_2]}{[G^2:H]} = \frac{[G:H_1][G:H_2]}{[G^2:H]}$ cosets of $H_1 \times H_2$, we have

$$\mu^{(2)}((x, y)H) = \frac{[G : H_1][G : H_2]}{[G^2 : H]}\mu^{(2)}(H_1 \times H_2) = \frac{1}{[G^2 : H]},$$

as required. $\qquad\square$

Remark 4.4. Note that not every CCM on G^2 arises as a product mean—for instance, one could take the average of the CCMs $\mu^{(2)}$ and $\widehat{\mu}^{(2)}$, as in Definition 4.1 and Remark 4.2. We return to this fact and its relationship to the degree of commutativity in Example 8.6.

Nevertheless, armed with the product mean, we can define a degree of commutativity with respect to it.

Definition 4.5. Let μ be a CCM on a group G. We define the *degree of commutativity* of G with respect to μ as

$$dc_\mu(G) := \mu^{(2)}(\text{Comm}(G)),$$

where $\text{Comm}(G) = \{(g, h) \in G^2 : [g, h] = 1\}$.

We will subsequently prove that this number is the same for *any* CCM μ in Proposition 6.9.

In the case where G is finite, there is clearly a unique CCM on G and it is already (left) invariant: namely, the uniform measure. Hence the following definition is consistent with the previous one and has been widely studied.

Definition 4.6. Let G be a finite group. We define the *degree of commutativity* of G as

$$\mathrm{dc}(G) \coloneqq \frac{|\mathrm{Comm}(G)|}{|G|^2}.$$

5. GROUPS WITH POSITIVE DEGREE OF COMMUTATIVITY ARE FAF

Finite quotients and residual finiteness. Since CCMs are, by definition, well behaved for subgroups it is natural to ask about the relationship between $\mathrm{dc}_\mu(G)$ and $\mathrm{dc}(G/N)$, whenever N is a finite index normal subgroup. This leads us naturally to a degree of commutativity which is seen by the finite quotients of a group. This will actually be important for showing that a positive degree of commutativity with respect to a CCM implies the group is FAF, but we need some preliminary results relating this to finite quotients first.

Definition 5.1. For a group G, we define

$$\mathrm{dc}_{\mathrm{RF}}(G) = \inf\{\mathrm{dc}(G/N) : N \trianglelefteq G, [G : N] < \infty\}.$$

We call this the *residually finite degree of commutativity* for G.

There is an easy relationship between the residually finite degree of commutativity and that arising from a CCM.

Lemma 5.2. *Let G be a group and μ a CCM on G. Then $\mathrm{dc}_\mu(G) \le \mathrm{dc}_{\mathrm{RF}}(G)$.*

Proof. Let $N \trianglelefteq G$ be a normal subgroup of finite index, and let $\pi_N \colon G^2 \to (G/N)^2$ be the natural projection. Then any pair of commuting elements also commute modulo N, implying that $\mathrm{Comm}(G) \subseteq \pi_N^{-1}(\mathrm{Comm}(G/N))$ and therefore

$$\mu^{(2)}(\mathrm{Comm}(G)) \le \mu^{(2)}(\pi_N^{-1}(\mathrm{Comm}(G/N))).$$

The right hand side equals $\mathrm{dc}(G/N)$ because the set $\pi_N^{-1}(\mathrm{Comm}(G/N))$ is a union of exactly $|\mathrm{Comm}(G/N)|$ cosets of N^2 in G^2, each of which has measure $1/[G : N]^2 = 1/|G/N|^2$ under $\mu^{(2)}$. Hence,

$$\mu^{(2)}\left(\pi_N^{-1}(\mathrm{Comm}(G/N))\right) = \frac{|\mathrm{Comm}(G/N)|}{|G/N|^2} = \mathrm{dc}(G/N),$$

and taking the infimum over all such N gives $\mathrm{dc}_\mu(G) \le \mathrm{dc}_{\mathrm{RF}}(G)$. $\qquad\square$

Note that $\mathrm{dc}_{\mathrm{RF}}(G)$ only really gives information if G admits sufficiently many finite quotients. Residually finite groups are defined as precisely the class of groups where the finite quotients can distinguish elements. More concretely, G is called *residually finite* if $\bigcap_{N \trianglelefteq_{\mathrm{fin}} G} N = \{1\}$, where we use the symbol $N \trianglelefteq_{\mathrm{fin}} G$ to denote that N is a normal subgroup of finite index in G. (Via the core construction—see Remark 2.14—one can drop the normality assumption on the finite index subgroups.)

It is then fairly straightforward to see the following. We note that this result fits the pattern of saying that the degree of commutativity is positive exactly when the group is FAF, since residually finite FAF groups are virtually abelian by Lemma 2.15.

Lemma 5.3. *Let G be a residually finite group. Then $\mathrm{dc}_{\mathrm{RF}}(G) > 0$ if and only if G is virtually abelian.*

Proof. If G is virtually abelian, then it has an abelian subgroup $H \le G$ of index $k < \infty$. Given any finite index normal subgroup $N \trianglelefteq G$, the subgroup HN/N of G/N is an abelian subgroup of index $\le k$, implying that $\mathrm{dc}(G/N) \ge 1/k^2$. Thus $\mathrm{dc}_{\mathrm{RF}}(G) \ge 1/k^2 > 0$.

Conversely, suppose that G is not virtually abelian. We will inductively construct a sequence $G = N_0, N_1, \ldots$ of finite index normal subgroups of G such that $\mathrm{dc}(G/N_i) \le (5/8)^i$, from which the result will follow.

Indeed, suppose that for some $i \ge 0$ the group N_i has been constructed. Since G is not virtually abelian, it follows that N_i is not abelian and so contains elements $g_i, h_i \in N_i$ such that $[g_i, h_i] \ne 1$. Since G is residually finite, there exists a finite index normal subgroup $M_{i+1} \trianglelefteq G$ such that $[g_i, h_i] \notin M_{i+1}$. We set $N_{i+1} = N_i \cap M_{i+1}$.

Then N_{i+1} is a normal finite index subgroup of G since so are N_i and M_{i+1}, and $N_{i+1} \trianglelefteq N_i$. Moreover, by construction N_i/N_{i+1} is non-abelian, and therefore $\mathrm{dc}(N_i/N_{i+1}) \leq 5/8$ by [Gus73, §1]. It then follows by the result of [Gal70] and the inductive hypothesis that

$$\mathrm{dc}(G/N_{i+1}) \leq \mathrm{dc}(G/N_i) \cdot \mathrm{dc}(N_i/N_{i+1}) \leq (5/8)^i \cdot (5/8) = (5/8)^{i+1},$$

as required. □

Remark 5.4. For a group G the residual of G, $\mathrm{Res}(G)$, is defined to be the intersection of all the normal finite index subgroups of G. The residual of a group G defines its largest residually finite quotient. The Lemma above then clearly shows that even if G is not residually finite, $\mathrm{dc}_{\mathrm{RF}}(G) > 0$ if and only if $G/\mathrm{Res}(G)$ is virtually abelian.

Centralisers and degree of commutativity. We are now ready to show that having $\mathrm{dc}_\mu(G) > 0$ implies that G is FAF. We will do this by looking at centralisers.

Throughout the remainder of the paper, we will refer to subsets of G consisting of elements whose centralisers have index $\leq n$, using the following notation.

Definition 5.5. Let G be a group and let $n \in \mathbb{N}$. We denote

$$X_n(G) = \{g \in G : [G : C_G(g)] \leq n\},$$

and we set $X_{\mathrm{fin}}(G) = \bigcup_{i=1}^\infty X_n(G)$.

Note that for any $g \in X_n(G)$ and $h \in X_m(G)$, we have

$$[G : C_G(gh)] \leq [G : C_G(g) \cap C_G(h)] \leq [G : C_G(g)] \cdot [G : C_G(h)] \leq mn,$$

implying that $X_n(G)X_m(G) \subseteq X_{nm}(G)$. In particular, $X_{\mathrm{fin}}(G)$ is a subgroup of G (and it is normal, since each $X_n(G)$ is invariant under conjugation).

Next we observe that having positive degree of commutativity implies that "many" elements have centraliser of small index.

Lemma 5.6. *Let μ be a CCM on a group G, and suppose that $\mathrm{dc}_\mu(G) = \alpha > 0$. Then $\mu(X_n(G)) > 0$ for any $n > \alpha^{-1}$.*

Proof. We have

$$\begin{aligned}
\mathrm{dc}_\mu(G) &= \mu^{(2)}(\{(x,y) \in G^2 : xy = yx \text{ and } x \in X_n(G)\}) \\
&\quad + \mu^{(2)}(\{(x,y) \in G^2 : xy = yx \text{ and } x \notin X_n(G)\}) \\
&\leq \int \mathbb{1}_{X_n(G)} d\mu + \frac{1}{n} \int \mathbb{1}_G d\mu \\
&= \mu(X_n(G)) + \frac{1}{n}.
\end{aligned}$$

The third line is obtained by estimating $\mu(C_G(x)) \leq 1$ for all $x \in X_n(G)$ and $\mu(C_G(x)) < \frac{1}{n}$ for all $x \notin X_n(G)$, as well as $\mathbb{1}_{X_n(G)^c} \leq \mathbb{1}_G$.

Hence $\mu(X_n(G)) > 0$. □

We now prove that having a positive degree of commutativity implies that a group is FAF. We do this in two stages; we first prove it for 2-step nilpotent groups (i.e. nilpotent groups of class at most 2) and then, for the general case, reduce the problem to precisely the class of 2-step nilpotent groups.

Remark 5.7. We recall that a group is nilpotent of class c if its lower central series terminates at the identity in c steps. In particular, 2-step nilpotent groups have central derived subgroups.

Lemma 5.8. *Let G be a 2-step nilpotent group equipped with a CCM μ. Suppose that for some $m \in \mathbb{N}$, there exists subset $S \subseteq X_m(G)$ such that $\mu(S) > 0$. Then G is FAF.*

Proof. Let $k = \lceil \mu(S)^{-1} \rceil$. We prove the result by induction on (m, k), in the lexicographic order.

For the base case, suppose that $m = 1$. Then we have $S \subseteq Z(G)$, and thus $\langle S \rangle \leq Z(G)$. On the other hand, we have $\mu(\langle S \rangle) \geq \mu(S) > 0$, implying that $\langle S \rangle$ has finite index in G. Thus G is virtually abelian and hence FAF, as required.

Suppose now that $m \geq 2$, and pick any non-central element $s_0 \in S$. The argument splits into two cases, depending on the relationship between $\mu(S \cap C_G(s_0))$ and $\mu(S)$.

Suppose first that $\mu(S \cap C_G(s_0)) = \mu(S)$. Let $G_1 = C_G(s_0)$ and $S_1 = S \cap G_1$, and define a mean μ_1 on G_1 by setting $\mu_1(A) = [G : G_1]\mu(A)$. It is easy to check that μ_1 is a CCM on G_1; moreover, for any $s \in S_1$ we have $[G_1 : C_{G_1}(s)] = [C_G(s)G_1 : C_G(s)] \leq m$ (note that $C_G(s)$ is normal in G since G is 2-step nilpotent). Since s_0 is non-central, we have $\mu(S_1) \leq \mu(C_G(s_0)) \leq \frac{1}{2}$ and thus

$$\mu_1(S_1)^{-1} = \frac{1}{[G : G_1]\mu(S_1)} \leq \frac{1}{2}\mu(S_1)^{-1} \leq \mu(S_1)^{-1} - 1 = \mu(S)^{-1} - 1,$$

and so $\lceil \mu_1(S_1)^{-1} \rceil < \lceil \mu(S)^{-1} \rceil$. It then follows by the inductive hypothesis that G_1 is FAF. But $[G : G_1] \leq m$, implying that G is virtually FAF and so FAF as well, as required.

Suppose now that $\mu(S \cap C_G(s_0)) < \mu(S)$. Since $C_G(s_0)$ has finite index in G and since μ is finitely additive, it follows that $\mu(S \cap gC_G(s_0)) > 0$ for some $g \in G - C_G(s_0)$. Let $z = [g, s_0]$, and note that, since G is 2-step nilpotent, the subgroup $\langle z \rangle \leq G$ is central and finite, as the fact that $m_0 := [G : C_G(s_0)] \leq m < \infty$ implies that $z^{m_0} = [g^{m_0}, s_0] = 1$. Let $G_2 = G/\langle z \rangle$ and $S_2 = (S\langle z \rangle \cap C_G(s_0))/\langle z \rangle \subseteq G_2$, and define a mean μ_2 on G_2 by setting $\mu_2(A\langle z \rangle/\langle z \rangle) = \mu(A\langle z \rangle)$. It is easy to check that μ_2 is a CCM on G_2.

We claim that $[G_2 : C_{G_2}(s)] \leq \frac{m}{2}$ for all $s \in S_2$. Indeed, given any $g \in G$, we have $|[g, G]| = |g^{-1} \cdot g^G| = |g^G| = [G : C_G(g)]$; similarly, $|[g\langle z \rangle, G_2]| = [G_2 : C_{G_2}(g\langle z \rangle)]$. Now for any element $s = s'\langle z \rangle \in S_2$ (where $s' \in S \cap gC_G(s_0)$), we have $1 \neq z = [g, s_0] = [s', s_0] \in [s', G]$, implying that

$$[G_2 : C_{G_2}(s)] = |[s, G_2]| = \left| \frac{[s', G]\langle z \rangle}{\langle z \rangle} \right| = \frac{|[s', G]|}{|\langle z \rangle|} = \frac{[G : C_G(s')]}{|\langle z \rangle|} \leq \frac{m}{|\langle z \rangle|} \leq \frac{m}{2},$$

as claimed.

Now it follows from the claim that $S_2 \subseteq G_2$ satisfies $\mu_2(S_2) > 0$ and $[G_2 : C_{G_2}(s)] \leq \frac{m}{2} \leq m-1$ for all $s \in S_2$. Thus, by the inductive hypothesis, $G_2 = G/\langle z \rangle$ is FAF. Since $\langle z \rangle$ is finite, it follows that G is FAF as well, as required. $\qquad\square$

We now reduce the general case to the nilpotent case in the following Proposition.

Proposition 5.9. *Let G be a group equipped with a CCM μ such that $\mathrm{dc}_\mu(G) > 0$. Then G is FAF.*

Proof. Choose some $m \in \mathbb{N}$ such that $\frac{1}{\mathrm{dc}_\mu(G)} < m$ and let $S = X_m(G)$. By Lemma 5.6, $\mu(S) > 0$. We have $\mu(\langle S \rangle) \geq \mu(S) > 0$, implying that the subgroup $G_0 := \langle S \rangle$ has finite index in G. Note also that S is invariant under conjugacy and so $G_0 \trianglelefteq_{\mathrm{fin}} G$.

Now the centre $Z(G_0)$ of G_0 is equal to $\bigcap_{s \in S} C_{G_0}(s)$, and in particular it is the intersection of finite index subgroups of G_0. Thus $G_0/Z(G_0)$ is residually finite, and hence so is $G_1 := G/Z(G_0)$.

By Lemma 5.2, $\mathrm{dc}_{\mathrm{RF}}(G) \geq \mathrm{dc}_\mu(G)$ and it is clear that $\mathrm{dc}_{\mathrm{RF}}(G_1) \geq \mathrm{dc}_{\mathrm{RF}}(G)$ and thus $\mathrm{dc}_{\mathrm{RF}}(G_1) \geq \mathrm{dc}_\mu(G) > 0$. It follows by Lemma 5.3 that G_1 is virtually abelian, and hence G is virtually 2-step nilpotent.

Let $G_2 \leq G$ be a finite index 2-step nilpotent subgroup. Since $\mu(S) > 0$, it then follows that $\mu(S_0) > 0$, where $S_0 = S \cap gG_2$, for some $g \in G$. Let $s_0 \in S_0$. Since $S^{-1}S = X_m(G)^2 \subseteq X_{m^2}(G)$, we then get $S_2 := s_0^{-1}S_0 \subseteq G_2 \cap X_{m^2}(G) \subseteq X_{m^2}(G_2)$ and moreover $\mu_2(S_2) > 0$, where μ_2 is a mean on G_2 defined by $\mu_2(A) = [G : G_2]\mu(s_0 A)$. It is easy to see that μ_2 is a CCM on G_2 and thus, by Lemma 5.8, G_2 is FAF and hence so is G, as required. $\qquad\square$

6. Finite-by-abelian-by-finite (FAF) groups

Since we have shown that positivity of the degree of commutativity corresponds to a group being FAF, the invariance of the actual quantity reduces to the study of it in FAF groups. This section is therefore devoted to showing that in a FAF group, the degree of commutativity is the same for every CCM. From this it immediately follows that the degree of commutativity is independent of the CCM used.

Indices of centralisers. Recall from Definition 5.5, that for any group G the set $X_n(G)$ denotes the set of elements whose centraliser has index at most n. Our next aim is to describe the sets $X_n(G)$ in a FAF group and relate these to the degree of commutativity.

We start by looking at the Boolean subring of $\mathcal{P}(G)$ generated by the cosets of subgroups.

Definition 6.1. Let G be a group and $\mathcal{P}(G)$ the power set of G considered as a Boolean ring. We denote by $\mathcal{C}$ the subring of $\mathcal{P}(G)$ generated by the cosets of subgroups of G.

Remark 6.2. Note that the identity $gH = (gHg^{-1})g$ means that we do not need to specify whether we intend to use left cosets or right cosets or both, as these all define the same object.

Lemma 6.3. *Let G be a group, $\mathcal{P}(G)$ the power set of G and $\mathcal{C}$ be the subring of generated by the cosets of subgroups of G. Then,*
  (i) *$\mathcal{C}$ is equal to the additive subgroup generated by the cosets of subgroups of G.*
  (ii) *In particular, every element of $\mathcal{C}$ can be written as a finite sum $H_1g_1 \oplus \cdots \oplus H_kg_k$ for some subgroups H_i of G and elements g_i.*

Proof. Note that (i) follows from (ii), since the ring has characteristic 2 and hence the finite sums form a subgroup. To observe (ii), simply note that the intersection of two cosets is either empty or is a coset of the intersection, and hence the finite sums are closed under multiplication. $\square$

We next analyse some sub-objects of $\mathcal{C}$ in order to show that $\mathcal{C}$ always admits a unique CCM.

Definition 6.4. Given the Boolean subring $\mathcal{C}$ of $\mathcal{P}(G)$, we let
  (i) $\mathcal{C}_{\mathrm{fin}}$ denote the additive subgroup generated by the cosets of finite index subgroups of G; and
  (ii) $\mathcal{C}_{\mathrm{inf}}$ denote the additive subgroup generated by the cosets of infinite index subgroups of G.

We then have:

Proposition 6.5. *Let G be a group and $\mathcal{C}$ be the Boolean subring generated by the cosets of subgroups of G. Then,*
  (i) *$\mathcal{C}_{\mathrm{fin}}$ is a subring of $\mathcal{C}$;*
  (ii) *$\mathcal{C}_{\mathrm{inf}}$ is an ideal of $\mathcal{C}$;*
  (iii) *$\mathcal{C}_{\mathrm{fin}} \cap \mathcal{C}_{\mathrm{inf}} = \{\varnothing\}$ and $\mathcal{C} = \mathcal{C}_{\mathrm{fin}} \oplus \mathcal{C}_{\mathrm{inf}}$ as abelian groups;*
  (iv) *there is a (ring) projection, $\pi\colon \mathcal{C} \to \mathcal{C}_{\mathrm{fin}}$;*
  (v) *there exists a unique CCM ν on $\mathcal{C}_{\mathrm{fin}}$, and it takes only rational values;*
  (vi) *there exists a unique CCM μ on $\mathcal{C}$, given by the composition $\nu \circ \pi$ and only taking rational values.*

Proof. We deal with these in turn.
  (i) This follows since the intersection of two cosets of finite index is either empty or a coset of the intersection, which is again of finite index.
  (ii) This follows similarly, since the non-empty intersection of a coset of infinite index with an arbitrary coset is a coset of infinite index.
  (iii) Every non-empty element of $\mathcal{C}_{\mathrm{fin}}$ is equal to a disjoint union of cosets of a single finite index subgroup. In contrast, elements of $\mathcal{C}_{\mathrm{inf}}$ are finite symmetric differences of cosets of infinite index subgroups, and are not generally disjoint unions of such cosets.

   Suppose $X \in \mathcal{C}_{\mathrm{fin}} \cap \mathcal{C}_{\mathrm{inf}}$ is non-empty. Then X contains a coset of a finite index subgroup and, being in $\mathcal{C}_{\mathrm{inf}}$, is contained in a finite union of cosets of infinite index subgroups. This contradicts Theorem 1.1.

   Hence, $\mathcal{C}_{\mathrm{fin}} \cap \mathcal{C}_{\mathrm{inf}} = \{\varnothing\}$. The decomposition $\mathcal{C} = \mathcal{C}_{\mathrm{fin}} \oplus \mathcal{C}_{\mathrm{inf}}$ as abelian groups now follows from Lemma 6.3.
  (iv) This is immediate from (ii) and (iii).
  (v) Since every element X of $\mathcal{C}_{\mathrm{fin}}$ is a disjoint union of cosets of a single finite index subgroup $H = H(X)$, the values of a CCM ν on elements of this subring are uniquely determined and rational. The only thing to check is that the values of ν do not depend on the

choices of the $H(X)$, and that ν is finitely additive; this comes down to saying that if $K \leq H \leq G$ are finite index subgroups of G, then every coset gH is a disjoint union of $[H : K]$ cosets of K, and hence the CCM gives the same value whether we think of gH as an H-coset or a union of K-cosets.

(vi) Since π is a ring homomorphism, it is readily seen that μ is a CCM on $\mathcal{C}$, and it takes only rational values since so does ν. Moreover, if μ' were another CCM on $\mathcal{C}$ then we can use (iii) to write any $X \in \mathcal{C}$ as $X = X_{\mathrm{fin}} \oplus X_{\mathrm{inf}}$ uniquely, where $X_{\mathrm{fin}} \in \mathcal{C}_{\mathrm{fin}}$ and $X_{\mathrm{inf}} \in \mathcal{C}_{\mathrm{inf}}$. From finite additivity we get that

$$\mu'(X_{\mathrm{fin}}) = \mu'(X_{\mathrm{fin}}) - \mu'(X_{\mathrm{inf}}) \leq \mu'(X_{\mathrm{fin}} \oplus X_{\mathrm{inf}}) \leq \mu'(X_{\mathrm{fin}}) + \mu'(X_{\mathrm{inf}}) = \mu'(X_{\mathrm{fin}}),$$

from which it follows that $\mu' = \mu$ by (v). $\square$

We now aim to show that the sets $U_m := X_m(G) - X_{m-1}(G)$ are elements of $\mathcal{C}$ in an arbitrary FAF group G. The idea is to take a FAF group G with its normal finite index subgroup $H \trianglelefteq_{\mathrm{fin}} G$ such that $N := [H, H]$ is central in H and finite (the existence of such an H is guaranteed by Remark 2.16), and consider the inclusions $C_G(g) \leq \overline{C}(gN) \leq G$, where $\overline{C}(gN)$ is the preimage of $C_{G/N}(gN)$ under the quotient map $G \to G/N$. The index $[\overline{C}(gN) : C_G(g)]$ is studied in Lemma 6.7, and the index $[G : \overline{C}(gN)]$ in Proposition 6.6 (which applies here since G/N is virtually abelian). These results are then combined in Corollary 6.8 to show that $U_m \in \mathcal{C}$.

Proposition 6.6. *Let G be a virtually abelian group, and let $\mathcal{K}$ be the set of all centralisers of elements $g \in X_{\mathrm{fin}}(G)$. Then $\mathcal{K}$ is finite, and we have*

$$S_K := \{g \in G : C_G(g) = K\} \in \mathcal{C} \quad \text{for all } K \in \mathcal{K}.$$

Proof. Let $A \leq G$ be a finite index abelian subgroup. Then for all $g \in G$ and $a \in A$, we have $C_A(g) = C_A(ga)$: indeed, an arbitrary element $b \in A$ commutes with a since A is abelian, and therefore b commutes with g if and only if it commutes with ga. In particular, since A has finite index in G, there are only finitely many subgroups of the form $C_A(g)$ when g ranges over the elements of G. Note that if $g \in X_{\mathrm{fin}}(G)$ then $[G : C_A(g)] = [G : A][A : C_A(g)] < \infty$, and therefore $N_0 = \bigcap_{g \in X_{\mathrm{fin}}(G)} C_A(g)$ has finite index in G.

Let $N = \mathrm{Core}_G(N_0) \trianglelefteq G$, and note that N has finite index in G. Then for every $g \in X_{\mathrm{fin}}(G)$, the centraliser $C_G(g)$ contains N, and therefore is the preimage in G of a subgroup of G/N. Since G/N is finite it only has finitely many subgroups, implying that the set $\mathcal{K}$ is finite.

Now note that given an element $g \in X_{\mathrm{fin}}(G)$ and a subgroup $K \in \mathcal{K}$, we have $C_G(g) = K$ if and only if g centralises K but does not centralise K_0 for any $K \lneq K_0 \in \mathcal{K}$. This shows that

$$S_K = C_G(K) - \bigcup_{K \lneq K_0 \in \mathcal{K}} C_G(K_0).$$

In particular, since $\mathcal{C}$ is closed under finite unions and set differences, we have $S_K \in \mathcal{C}$, as required. $\square$

The following Lemma is technical. Given a finite index subgroup $K \leq_{\mathrm{fin}} G/N$, we consider its preimage $\overline{K} \leq G$ as well as the subgroup $\overline{C}_K \leq G$ of elements that centralise $\overline{K}$ modulo N. We show that given any $g \in \overline{C}_K$, the set of commutators $M_g := [g, H \cap \overline{K}] \subseteq N$ is a subgroup, and for any subgroup $M \leq N$, the set $L_M := \{g \in \overline{C}_K : M_g \leq M\}$ is a subgroup as well. This allows us to control the set of elements for which M_g has a certain cardinality. Moreover, we show that given $g \in L_M$, the coset $[g, h]M$ for $h \in \overline{K}$ depends only on the coset $h(H \cap \overline{K})$, which allows us to have control over the number $|[g, \overline{K}]|/|M_g|$. Combining these results, we get a description of the set of elements $g \in \overline{C}_K$ for which $[g, \overline{K}]$ has some fixed cardinality $m \in \mathbb{N}$. But this cardinality is the same as the cardinality of the $\overline{K}$-conjugacy class of g, i.e. equal to the index $[\overline{K} : C_{\overline{K}}(g)]$. The result then follows by noting that if $C_{G/N}(gN) = K$, then we have $C_G(g) \subseteq \overline{K}$ and therefore $C_{\overline{K}}(g) = C_G(g)$.

Lemma 6.7. *Let G be a group that has a finite normal subgroup $N \trianglelefteq G$ and a finite index normal subgroup $H \trianglelefteq_{\mathrm{fin}} G$ containing N such that N is central in H and H/N is abelian. Let*

$K \leq_{\mathrm{fin}} G/N$, let $m \in \mathbb{N} \cup \{\infty\}$, and write $\pi_N \colon G \to G/N$ for the quotient map. Then we have

$$S_{K,m} := \{g \in \overline{S}_K : [\pi_N^{-1}(K) : C_G(g)] = m\} \in \mathcal{C},$$

where $\overline{S}_K = \{g \in G \colon C_{G/N}(gN) = K\}$, and $S_{K,m} = \varnothing$ for all $m > |N|$.

Proof. We write $\overline{K} := \pi_N^{-1}(K)$ and $\overline{C}_K := \pi_N^{-1}(C_{G/N}(K))$; note that $\overline{S}_K \subseteq \overline{C}_K$ and that since G/N is virtually abelian, we have $\overline{S}_K \in \mathcal{C}$ by Proposition 6.6. Given a subgroup $M \leq N$, we let

$$L_M := \{g \in \overline{C}_K : [g, H \cap \overline{K}] \subseteq M\},$$

where $[g, H \cap \overline{K}] = \{[g, h] : h \in H \cap \overline{K}\}$. The proof then proceeds in the following steps:

  (i) showing that $[g, H \cap \overline{K}]$ is a subgroup of N for every $g \in \overline{C}_K$;

  (ii) showing that $[g, \overline{K}]$ is a union of cosets of $[g, H \cap \overline{K}]$ in N for every $g \in \overline{C}_K$;

  (iii) showing that L_M is a subgroup of G;

  (iv) showing that for any $g \in L_M$, the function (which is not necessarily a group homomorphism) $\phi_g \colon \overline{K} \to N/M$, defined by $\phi_g(h) = [g, h]M$, factors through the quotient map $\overline{K} \xrightarrow{q} \overline{K}/(H \cap \overline{K})$, i.e. $\phi_g = \Phi_g \circ q$ for some function $\Phi_g \colon \overline{K}/(H \cap \overline{K}) \to N/M$;

  (v) showing that the function $\Psi \colon L_M \to (N/M)^{\overline{K}/(H \cap \overline{K})}$ defined by $g \mapsto \Phi_g$ is a group homomorphism.

After proving all these points, the proof can be finalised as follows. Given a subgroup $M \leq N$, we set $T_M = \{g \in \overline{C}_K : [g, H \cap \overline{K}] = M\}$. It then follows from point (i) that we have

$$T_M = L_M - \bigcup_{M_0 \lneq M} L_{M_0}.$$

By point (iii) we have $L_{M_0} \in \mathcal{C}$ for all $M_0 \leq M$; in particular, since M is finite (and so has finitely many subgroups) and since $\mathcal{C}$ is closed under finite unions and set differences, it follows that $T_M \in \mathcal{C}$.

Now given any $g \in \overline{S}_K$, note that $\overline{K} = \pi_N^{-1}(K)$ is precisely the set of elements $h \in G$ such that $[g, h] \in N$, and in particular $C_G(g) \leq \overline{K}$. On the other hand, $\overline{K}$ acts transitively on the set of $\overline{K}$-conjugacy classes of g, $\{h^{-1}gh : h \in \overline{K}\} = g[g, \overline{K}]$, and $C_G(g)$ is precisely the stabiliser of g under this action. In particular, we have $[\overline{K} : C_G(g)] = |[g, \overline{K}]|$. Since $[g, \overline{K}] \subseteq N$, it also follows that $S_{K,m} = \varnothing$ for all $m > |N|$.

Note that if $g \in T_M$ for some $M \leq N$, then it follows from point (ii) and the definition of ϕ_g that we have $|[g, \overline{K}]| = |M| \cdot |\phi_g(\overline{K})|$. In particular, if given any $r \in \mathbb{N}$ we set $\mathcal{F}_r := \{g \in L_M : |\phi_g(\overline{K})| = r\}$, we then have

$$S_{K,m} = \overline{S}_K \cap \bigsqcup_{\substack{M \leq N \\ m/|M| \in \mathbb{N}}} T_M \cap \mathcal{F}_{m/|M|}.$$

Now by point (v), it follows that for each r, the set $\mathcal{F}_r$ is a union of cosets of $\ker(\Psi)$. Since H has finite index in G and since N is finite, it follows that both $\overline{K}/(H \cap \overline{K})$ and N/M are finite (for all $M \leq N$), and therefore $(N/M)^{\overline{K}/(H \cap \overline{K})}$ is also finite. We therefore have that $\ker(\Psi)$ is a finite index subgroup of L_M and so $\mathcal{F}_r$ is a finite union of cosets of $\ker(\Psi)$ for each $r \in \mathbb{N}$; thus $\mathcal{F}_r \in \mathcal{C}$. Since $\mathcal{C}$ is closed under finite unions and intersections, and since $T_M \in \mathcal{C}$ for all $M \leq N$ (by the argument above) and $\overline{S}_K \in \mathcal{C}$ (by Proposition 6.6), it then follows that $S_{K,m} \in \mathcal{C}$, as required.

It thus remains to prove the points (i)–(v).

  (i) Given $h_1, h_2 \in H \cap \overline{K}$, note that $[g, h_1] \in N$ since $g \in \overline{C}_K$ and $h_1 \in \overline{K}$, and therefore $[g, h_1]^{h_2} = [g, h_1]$ since N is central in H. We then have, using commutator calculus and the fact that N is abelian,

$$[g, h_1 h_2] = [g, h_2][g, h_1]^{h_2} = [g, h_2][g, h_1] = [g, h_1][g, h_2],$$

and so the map $H \cap \overline{K} \to N$ sending $h \mapsto [g, h]$ is a group homomorphism. It follows that the image of this map, $[g, H \cap \overline{K}]$, is a subgroup of N, as required.

(ii) This is similar to the previous point: it is enough to show that given $k \in \overline{K}$ we have $[g, k][g, H \cap \overline{K}] \subseteq [g, \overline{K}]$—that is, $[g, k][g, h] \in [g, \overline{K}]$ for every $h \in H \cap \overline{K}$. But we have $[g, k] \in N$ since $g \in \overline{C}_k$ and therefore $[g, k]^h = [g, k]$ since N is central in H. We thus have
$$[g, k][g, h] = [g, k]^h[g, h] = [g, h][g, k]^h = [g, kh] \in [g, \overline{K}],$$
as required.

(iii) Clearly $1 \in L_M$. We first claim that $g^{-1} \in L_M$ for a given $g \in L_M$. Indeed, for any $h \in H \cap \overline{K}$ we have
$$[g^{-1}, h] = [h, g]^{g^{-1}} = [h^{g^{-1}}, g] = [g, h^{g^{-1}}]^{-1}.$$
On the other hand, note that $h^{g^{-1}} \in H$ since H is normal in G, and $h^{g^{-1}} \in \overline{K}$ since the fact that $g \in \overline{C}_K$ implies that $\overline{K}$ is invariant under conjugation by g. Thus we have $[g^{-1}, h] \in [g, H \cap \overline{K}]^{-1} \subseteq M$ for all $h \in H \cap \overline{K}$, showing that $g^{-1} \in L_M$, as claimed.

We now claim that $g_1g_2 \in L_M$ for given $g_1, g_2 \in L_M$. Indeed, for any $h \in H \cap \overline{K}$ we have, by commutator calculus,
$$[g_1g_2, h] = [g_1, h]^{g_2}[g_2, h] = [g_1, h][[g_1, h], g_2][g_2, h] = [g_1, h][g_2, [g_1, h]]^{-1}[g_2, h].$$
Since $[g_1, h] \in N \subseteq H \cap \overline{K}$, it follows that all three terms in the equation above are elements of M for any $h \in H \cap \overline{K}$ and so $[g_1g_2, h] \in M$, implying that $g_1g_2 \in L_M$, as claimed.

(iv) It is enough to show that we have $\phi_g(k) = \phi_g(kh)$ for all $k \in \overline{K}$ and $h \in H \cap \overline{K}$. But this follows from the computation in point (ii) above: indeed, we have
$$\phi_g(kh) = [g, kh]M = [g, k][g, h]M = [g, k]M = \phi_g(k),$$
since the fact that $g \in L_M$ implies that $[g, h] \in M$.

(v) Given $g_1, g_2 \in L_M$ and $k \in \overline{K}$, note that we have
$$[g_2, [g_1, k]] \in [g_2, N] \subseteq [g_2, H \cap \overline{K}] \subseteq M.$$
We thus have, by commutator calculus,
$$\phi_{g_1g_2}(k) = [g_1g_2, k]M = [g_1, k]^{g_2}[g_2, k]M = [g_1, k][g_2, [g_1, k]]^{-1}[g_2, k]M$$
$$= [g_1, k][g_2, k]M = \phi_{g_1}(k)\phi_{g_2}(k) = (\phi_{g_1}\phi_{g_2})(k),$$
and thus Ψ is a group homomorphism, as required. $\square$

Corollary 6.8. *Let G be a FAF group, and let $m \in \mathbb{N} \cup \{\infty\}$. Then we have*
$$U_m := \{g \in G : [G : C_G(g)] = m\} \in \mathcal{C},$$
and only finitely many of the U_m are non-empty.

Proof. By Remark 2.16, we may assume that G has subgroups N and H as in the statement of Lemma 6.7. Note that the group G/N is virtually abelian.

We first observe that only finitely many of the U_m are non-empty. Indeed, note that by Proposition 6.6, only finitely many finite index subgroups of G/N can appear as centralisers of elements of G/N. On the other hand, it follows from Lemma 6.7 that for any $g \in X_{\mathrm{fin}}(G)$ we have $[\pi_N^{-1}(C_{G/N}(gN)) : C_G(g)] \leq |N|$. Combining these two facts, we get an upper bound on the finite indices of centralisers of elements in G, as required.

In particular, it is enough to show that $U_m \in \mathcal{C}$ for all $m \in \mathbb{N}$, as then U_∞ will be the complement of a finite union of elements of $\mathcal{C}$ and so will also belong to $\mathcal{C}$.

Now by Proposition 6.6, there is a finite set $\mathcal{K}$ of finite index subgroups of G/N that occur as centralisers of elements. Moreover, by Lemma 6.7, for any $K \in \mathcal{K}$ and any $r \in \mathbb{N}$ we have
$$S_{K,r} := \{g \in G : C_{G/N}(gN) = K, [\pi_N^{-1}(K) : C_G(g)] = r\} \in \mathcal{C}.$$
In particular, if we denote by i_K the index $[G : \pi_N^{-1}(K)]$ for any $K \in \mathcal{K}$, it then follows that
$$U_m = \bigsqcup_{\substack{K \in \mathcal{K} \\ m/i_K \in \mathbb{N}}} S_{K, m/i_K},$$

and so $U_m \in \mathcal{C}$ for any $m \in \mathbb{N}$, as required. $\qquad \square$

Degree of commutativity. We are now in a position to prove the degree of commutativity is independent of the CCM used for a FAF group.

Proposition 6.9. *Let μ and ν be CCMs on a FAF group G. Then $\mathrm{dc}_\mu(G) = \mathrm{dc}_\nu(G) \in \mathbb{Q}$.*

Proof. By Corollary 6.8, we have $G = \left(\bigsqcup_{m=1}^M U_m\right) \sqcup U_\infty$ for some $M \in \mathbb{N}$, where

$$U_m = \{g \in G : [G : C_G(g)] = m\} \in \mathcal{C}.$$

In particular, we have

$$\mathrm{dc}_\mu(G) = \int \mu(C_G(g)) \, \mathrm{d}\mu(g) = \int \frac{1}{[G : C_G(g)]} \, \mathrm{d}\mu(g)$$

$$= \left(\sum_{m=1}^M \int \frac{1}{m} \mathbb{1}_{U_m} \, \mathrm{d}\mu\right) + \int 0 \cdot \mathbb{1}_{U_\infty} \, \mathrm{d}\mu = \sum_{m=1}^M \frac{\mu(U_m)}{m},$$

and similarly $\mathrm{dc}_\nu(G) = \sum_{m=1}^M \frac{\nu(U_m)}{m}$. But for $1 \leq m \leq M$ we have $U_m \in \mathcal{C}$ and therefore $\mu(U_m) = \nu(U_m) \in \mathbb{Q}$ by Lemma 6.5(vi), implying that $\mathrm{dc}_\mu(G) = \mathrm{dc}_\nu(G) \in \mathbb{Q}$, as required. $\qquad \square$

Theorem 6.10. *Let G be a group equipped with a CCM μ. Then $\mathrm{dc}_\mu(G) > 0$ if and only if G is finite-by-abelian-by-finite. Moreover, $\mathrm{dc}_{\mathrm{CCM}}(G) := \mathrm{dc}_\mu(G)$ is a rational number that does not depend on the choice of μ.*

Proof. Suppose that G is FAF. In particular, G has a finite index normal subgroup $H \trianglelefteq G$ such that $N := [H, H]$ is finite and central in H (see Remark 2.16). Since H is 2-step nilpotent, it follows from commutator calculus that the map $\psi_g \colon H \to N$ defined by $\psi_g(x) = [g, x]$ is a group homomorphism for any $g \in H$. In particular, we have $[H : C_H(g)] = [H : \ker(\psi_g)] = |\psi_g(H)| \leq |N|$ for every $g \in H$, implying that

$$\mathrm{dc}_\mu(G) = \int \mu(C_G(g)) \, \mathrm{d}\mu = \int [G : C_G(g)]^{-1} \, \mathrm{d}\mu \geq \int [G : C_G(g)]^{-1} \mathbb{1}_H \, \mathrm{d}\mu$$

$$\geq \int [G : C_H(g)]^{-1} \mathbb{1}_H \, \mathrm{d}\mu \geq (|N| \cdot [G : H])^{-1} \int \mathbb{1}_H \, \mathrm{d}\mu = \frac{1}{|N| \cdot [G : H]^2}$$

and thus $\mathrm{dc}_\mu(G) > 0$, as required.

Suppose now instead that $\mathrm{dc}_\mu(G) > 0$. By Lemma 5.6, there exists $m \in \mathbb{N}$ such that $\mu(X_m(G)) > 0$. It then follows by Proposition 5.9 that G is FAF.

Finally, the rationality and independence from the choice of μ follows from Proposition 6.9 if G is FAF, and from the fact that $\mathrm{dc}_\mu(G) = 0$ otherwise. $\qquad \square$

7. RESIDUALLY FINITE GROUPS

Having defined the residually finite degree of commutativity in Definition 5.1 it is natural to see whether this quantity agrees with the analogous number obtained using the Haar measure.

Definition 7.1. For a residually finite group G, the *profinite completion* $\widehat{G}$ is the inverse limit of the directed system of finite quotients of G:

$$\widehat{G} := \varprojlim \{G/N : N \trianglelefteq_{\mathrm{fin}} G\}.$$

The diagonal embedding $g \mapsto (gN)$ embeds G as a dense subgroup into $\widehat{G}$, and we will therefore identify G with a subgroup of $\widehat{G}$. The group $\widehat{G}$, with the inverse limit topology (induced by the discrete topology on the finite quotients of G), is a compact Hausdorff totally disconnected topological group and has a unique left invariant inner-regular and outer-regular Borel probability measure $\widehat{\mu}$, called the *Haar measure*.

Theorem 7.2. *Let G be a residually finite group with profinite completion $\widehat{G}$, and let $\widehat{\mu}$ be the Haar measure on $\widehat{G}^2$. Then we have*

$$\widehat{\mu}(\mathrm{Comm}(\widehat{G})) = \widehat{\mu}(\overline{\mathrm{Comm}(G)}) = \mathrm{dc}_{\mathrm{RF}}(G) = \mathrm{dc}_{\mathrm{CCM}}(G).$$

Moreover, this number is positive if and only if G is virtually abelian.

Proof. The last statement of this Theorem is simply Lemma 5.3, so it suffices to prove the displayed equalities.

Note that $\mathrm{Comm}(\widehat{G})$ is closed in $\widehat{G}^2$, since the commutator map $(g, h) \mapsto [g, h]$ is continuous and since $\{1\}$ is closed in $\widehat{G}$. As $\mathrm{Comm}(G) = G^2 \cap \mathrm{Comm}(\widehat{G}) \subseteq \mathrm{Comm}(\widehat{G})$, it follows that $\overline{\mathrm{Comm}(G)} \subseteq \mathrm{Comm}(\widehat{G})$. This implies that $\widehat{\mu}(\overline{\mathrm{Comm}(G)}) \leq \widehat{\mu}(\mathrm{Comm}(\widehat{G}))$.

On the other hand, if $N \trianglelefteq G$ is a normal finite index subgroup of G, then the preimage of $\mathrm{Comm}(G/N)$ in $\widehat{G}^2$ is equal to $\{(g, h) \in \widehat{G}^2 : [g, h] \in \overline{N}\}$ and therefore contains $\mathrm{Comm}(\widehat{G})$, and the Haar measure of this preimage is exactly equal to $\mathrm{dc}(G/N)$. It follows that $\widehat{\mu}(\mathrm{Comm}(\widehat{G})) \leq \mathrm{dc}(G/N)$ for all $N \trianglelefteq_{\mathrm{fin}} G$, and thus $\widehat{\mu}(\mathrm{Comm}(\widehat{G})) \leq \mathrm{dc}_{\mathrm{RF}}(G)$. In particular, we have

$$\widehat{\mu}(\overline{\mathrm{Comm}(G)}) \leq \widehat{\mu}(\mathrm{Comm}(\widehat{G})) \leq \mathrm{dc}_{\mathrm{RF}}(G).$$

Similarly, $\mathrm{dc}_\nu(G) \leq \mathrm{dc}_{\mathrm{RF}}(G)$ by Lemma 5.2. It is hence enough to show that $\mathrm{dc}_{\mathrm{RF}}(G) \leq \widehat{\mu}(\overline{\mathrm{Comm}(G)})$ and $\mathrm{dc}_{\mathrm{RF}}(G) \leq \mathrm{dc}_\nu(G)$.

Now if $\mathrm{dc}_{\mathrm{RF}}(G) = 0$, then the result follows immediately, so we may assume that $\mathrm{dc}_{\mathrm{RF}}(G) > 0$. Then, by Lemma 5.3, G is virtually abelian, and so contains a normal finite index abelian subgroup $A \trianglelefteq G$. Recall that the set $X_{\mathrm{fin}}(G)$ of elements of G whose centraliser has finite index forms a normal subgroup, and note that $X_{\mathrm{fin}}(G) = \{g \in G : [G : C_A(g)] < \infty\}$ as A has finite index in G. We aim to construct:

(i) a finite index normal subgroup $M \trianglelefteq_{\mathrm{fin}} G$ that is contained in $X_{\mathrm{fin}}(G)$ such that we have $\mathrm{Comm}(G) \cap X_{\mathrm{fin}}(G)^2 = \{(g, h) \in X_{\mathrm{fin}}(G)^2 : [g, h] \in M\}$; and

(ii) for any $\varepsilon > 0$, a finite index normal subgroup $N_\varepsilon \trianglelefteq_{\mathrm{fin}} G$ contained in M such that $(G/N_\varepsilon)^2$ contains at most $\varepsilon |G/N_\varepsilon|^2$ commuting pairs outside of $(X_{\mathrm{fin}}(G)/N_\varepsilon)^2$.

Now having constructed N_ε, note that since $N_\varepsilon \subseteq M \subseteq X_{\mathrm{fin}}(G)$, we have that $\mathrm{Comm}(G) \cap X_{\mathrm{fin}}(G)^2$ is a union of cosets of N_ε^2 in G^2, and that the image of $\mathrm{Comm}(G) \cap X_{\mathrm{fin}}(G)^2$ in $(G/N_\varepsilon)^2$ is precisely $\mathrm{Comm}(G/N_\varepsilon) \cap (X_{\mathrm{fin}}(G)/N_\varepsilon)^2$. It follows that $\overline{\mathrm{Comm}(G)} \subseteq \widehat{G}^2$ contains the preimage of $\mathrm{Comm}(G/N_\varepsilon) \cap (X_{\mathrm{fin}}(G)/N_\varepsilon)^2$ under the surjection $\widehat{G}^2 \to (G/N_\varepsilon)^2$, and therefore

$$\widehat{\mu}(\overline{\mathrm{Comm}(G)}) \geq \frac{|\mathrm{Comm}(G/N_\varepsilon) \cap (X_{\mathrm{fin}}(G)/N_\varepsilon)^2|}{|G/N_\varepsilon|^2}.$$

Similarly, $\mathrm{Comm}(G)$ contains the preimage of $\mathrm{Comm}(G/N_\varepsilon) \cap (X_{\mathrm{fin}}(G)/N_\varepsilon)^2$ under the surjection $G^2 \to (G/N_\varepsilon)^2$, and therefore

$$\mathrm{dc}_\nu(G) \geq \frac{|\mathrm{Comm}(G/N_\varepsilon) \cap (X_{\mathrm{fin}}(G)/N_\varepsilon)^2|}{|G/N_\varepsilon|^2}.$$

On the other hand, we have $\frac{|\mathrm{Comm}(G/N_\varepsilon) - (X_{\mathrm{fin}}(G)/N_\varepsilon)^2|}{|G/N_\varepsilon|^2} \leq \varepsilon$ by the choice of N_ε, and therefore

$$\mathrm{dc}(G/N_\varepsilon) - \frac{|\mathrm{Comm}(G/N_\varepsilon) \cap (X_{\mathrm{fin}}(G)/N_\varepsilon)^2|}{|G/N_\varepsilon|^2} \leq \varepsilon.$$

It follows that $\widehat{\mu}(\overline{\mathrm{Comm}(G)}) \geq \mathrm{dc}(G/N_\varepsilon) - \varepsilon \geq \mathrm{dc}_{\mathrm{RF}}(G) - \varepsilon$, and similarly $\mathrm{dc}_\nu(G) \geq \mathrm{dc}_{\mathrm{RF}}(G) - \varepsilon$. But $\varepsilon > 0$ was arbitrary, implying that we have $\widehat{\mu}(\overline{\mathrm{Comm}(G)}) \geq \mathrm{dc}_{\mathrm{RF}}(G)$ and $\mathrm{dc}_\nu(G) \geq \mathrm{dc}_{\mathrm{RF}}(G)$, as required.

It remains to construct the subgroups M and N_ε as above.

(i) Note that $[G : X_{\mathrm{fin}}(G)] < \infty$ (since $A \subseteq X_{\mathrm{fin}}(G)$) and set

$$M_0 := \bigcap_{g \in X_{\mathrm{fin}}(G)} C_A(g).$$

Note that $M_0 \trianglelefteq G$ (since $X_{\mathrm{fin}}(G) \trianglelefteq G$ and $A \trianglelefteq G$). Moreover, given any $g \in G$ and $a \in C_A(g)$, a given element $b \in A$ commutes with a, and so it commutes with g if and only if it commutes with ga, showing that $C_A(g) = C_A(ga)$. Thus $C_A(g)$ is constant as g ranges over a coset of A in G; as $[G : A] < \infty$, it follows that M_0 is the intersection of *finitely many* finite index subgroups of G and therefore $[G : M_0] < \infty$.

Now let $\mathcal{T}$ be a transversal of M_0 in $X_{\mathrm{fin}}(G)$ containing 1 (note that $M_0 \subseteq A \subseteq X_{\mathrm{fin}}(G)$), and let $C = \{[g, h] : (g, h) \in \mathcal{T}^2\}$. Then C is a finite subset of G containing 1. Since G is residually finite, there exists a finite index normal subgroup $M_1 \trianglelefteq G$ such that the quotient map $G \to G/M_1$ is injective on C, and in particular $C \cap M_1 = \{1\}$.

We set $M = M_0 \cap M_1$; by construction, M is a normal finite index subgroup of G contained in $X_{\mathrm{fin}}(G)$. Given any $g, h \in \mathcal{T}$ and $a, b \in M_0$, note that $[ga, b] = 1 = [a, h]$ by the definition of M_0, implying that

$$[ga, hb] = [ga, b][ga, h]^b = [ga, h]^b = [g, h]^{ab}[a, h]^b = [g, h]^{ab}.$$

In particular, since $M \trianglelefteq G$, we have $[ga, hb] \in M$ if and only if $[g, h] \in M$, which happens if and only if $[g, h] = 1$ by the choice of M_1. It follows that for any $g, h \in X_{\mathrm{fin}}(G)$, we have either $[g, h] = 1$ or $[g, h] \notin M$, and therefore $\mathrm{Comm}(G) \cap X_{\mathrm{fin}}(G)^2 = \{(g, h) \in X_{\mathrm{fin}}(G)^2 : [g, h] \in M\}$, as required.

(ii) Note that given $g \in G$ and $a, b \in A$, we have

$$[ga, b] = [g, b]^a[a, b] = [g, b]$$

since A is normal and abelian, and in particular $[g, A] = [ga, A]$. Thus $[g, A]$ is constant as g ranges over elements in a coset of A in G, and so, since $[G : A] < \infty$, there are only finitely many subsets of G of the form $[g, A]$, for $g \in G$. Note also that we have $|[g, A]| = |g[g, A]| = |g^A| = [G : C_A(g)]$ by the orbit-stabiliser theorem, and in particular $|[g, A]| < \infty$ if and only if $g \in X_{\mathrm{fin}}(G)$.

Let $n \in \mathbb{N}$ be such that $2/n \leq \varepsilon$. By the paragraph above, there exists a finite subset $S \subseteq G$ such that for each $g \in G - X_{\mathrm{fin}}(G)$, the subset S contains at least n commutators of the form $[g, a]$ for $a \in A$. Since G is residually finite, there exists a finite index normal subgroup $N_\varepsilon \trianglelefteq_{\mathrm{fin}} G$ such that the quotient map $G \to G/N_\varepsilon$ is injective on S; by replacing N_ε with $N_\varepsilon \cap A \cap M$ if necessary, we may further assume that $N_\varepsilon \subseteq A$ and $N_\varepsilon \subseteq M$.

It then follows that for any $g \in G - X_{\mathrm{fin}}(G)$, the subset $[gN_\varepsilon, A/N_\varepsilon] = [g, A]N_\varepsilon/N_\varepsilon \subseteq G/N_\varepsilon$ contains at least n elements, and therefore

$$[G/N_\varepsilon : C_{G/N_\varepsilon}(gN_\varepsilon)] = |[gN_\varepsilon, G/N_\varepsilon]| \geq |[gN_\varepsilon, A/N_\varepsilon]| \geq n$$

by the orbit-stabiliser theorem. Now note that we have

$$\mathrm{Comm}(G/N_\varepsilon) - (X_{\mathrm{fin}}(G)/N_\varepsilon)^2 = \left(\bigcup_{g \in G - X_{\mathrm{fin}}(G)} \{gN_\varepsilon\} \times C_{G/N_\varepsilon}(gN_\varepsilon) \right)$$
$$\cup \left(\bigcup_{h \in G - X_{\mathrm{fin}}(G)} C_{G/N_\varepsilon}(hN_\varepsilon) \times \{hN_\varepsilon\} \right),$$

and therefore, by symmetry,

$$\left| \mathrm{Comm}(G/N_\varepsilon) - (X_{\mathrm{fin}}(G)/N_\varepsilon)^2 \right| \leq 2 \sum_{gN_\varepsilon \in (G - X_{\mathrm{fin}}(G))/N_\varepsilon} |C_{G/N_\varepsilon}(gN_\varepsilon)|$$
$$\leq 2 \cdot |(G - X_{\mathrm{fin}}(G))/N_\varepsilon| \cdot \frac{|G/N_\varepsilon|}{n}$$
$$\leq \frac{2|G/N_\varepsilon|^2}{n} \leq \varepsilon |G/N_\varepsilon|^2,$$

as required. $\qquad\square$

8. AMENABLE GROUPS

We note that our construction to date has been to look at a CCM on a group G and extend to a product mean on G^2 from which we define a degree of commutativity and deduce structural results on G. However, one could instead look at an arbitrary CCM on G^2 and, as we shall see—Example 8.6—this can produce a different answer.

However we can also ask the same questions for invariant means and amenable groups. Namely, what can we say about a degree of commutativity for an amenable group which is defined using a non-product mean on G^2? In particular, one should bear in mind that FAF groups are all amenable.

Our result here is that the positivity of these quantities corresponds to the group being FAF in the amenable case, even if one doesn't use the product mean.

Non-product CCMs on products. We recall, Definition 5.5, that $X_n(G)$ consists of all the elements of G whose centraliser has index at most n, and that $X_{\mathrm{fin}}(G)$ is a normal subgroup of G that is the union of all the $X_n(G)$.

Lemma 8.1. *Suppose that either $X_n(G) \neq X_{\mathrm{fin}}(G)$ for any $n \in \mathbb{N}$, or $[G : X_{\mathrm{fin}}(G)] = \infty$. Then for every $n \in \mathbb{N}$ and every $\varepsilon > 0$, there exist $h_1, \ldots, h_n \in G$ such that*

$$\sum_{1 \leq i < j \leq n} [G : C_G(h_i^{-1} h_j)]^{-1} < \varepsilon.$$

Proof. Suppose first that $[G : X_{\mathrm{fin}}(G)] = \infty$. We then choose any elements $h_1, \ldots, h_n$ that represent n distinct cosets of $X_{\mathrm{fin}}(G)$ in G. This implies that $h_i^{-1} h_j \notin X_{\mathrm{fin}}(G)$ and therefore $[G : C_G(h_i^{-1} h_j)] = \infty$ for all $1 \leq i < j \leq n$, implying that $\sum_{1 \leq i < j \leq n} [G : C_G(h_i^{-1} h_j)]^{-1} = 0 < \varepsilon$, as required.

Suppose now that $X_n(G) \neq X_{\mathrm{fin}}(G)$ for any $n \in \mathbb{N}$. We will then construct $h_1, \ldots, h_n \in X_{\mathrm{fin}}(G)$ such that

$$c_k := \sum_{1 \leq i < j \leq k} [G : C_G(h_i^{-1} h_j)]^{-1} \leq \varepsilon(1 - 2^{1-k})$$

for $1 \leq k \leq n$, inductively as follows. We pick an arbitrary $h_1 \in X_{\mathrm{fin}}(G)$. Having constructed $h_1, \ldots, h_k$ for some $k \geq 1$, let $N_k \in \mathbb{N}$ be such that $N_k \geq \frac{2^k k M_k}{\varepsilon}$, where $M_k = \max\{[G : C_G(h_i)] : 1 \leq i \leq k\}$. Since $X_{N_k}(G) \neq X_{\mathrm{fin}}(G)$, there exists $h_{k+1} \in X_{\mathrm{fin}}(G)$ such that $[G : C_G(h_{k+1})] > N_k$. We then have

$$N_k < [G : C_G(h_{k+1})] \leq [G : C_G(h_i) \cap C_G(h_i^{-1} h_{k+1})] \leq M_k \cdot [G : C_G(h_i^{-1} h_{k+1})]$$

for $1 \leq i \leq k$, implying that

$$[G : C_G(h_i^{-1} h_{k+1})] \geq \frac{N_k}{M_k} \geq \frac{2^k k}{\varepsilon}.$$

In particular, it follows that

$$c_{k+1} = c_k + \sum_{i=1}^{k} [G : C_G(h_i^{-1} h_{k+1})]^{-1} \leq \varepsilon(1 - 2^{1-k}) + k \cdot \frac{\varepsilon}{2^k k} = \varepsilon(1 - 2^{-k}),$$

as required. $\qquad\square$

Corollary 8.2. *Suppose that G is amenable, and let μ be any invariant mean on G^2. If $\mu(\mathrm{Comm}(G)) > 0$, then there exists $n \in \mathbb{N}$ such that $\mu(X_n(G) \times G) > 0$.*

Proof. Suppose, for contradiction, that $\mu(X_n(G) \times G) = 0$ for all $n \in \mathbb{N}$. If we have $X_n(G) = X_{\mathrm{fin}}(G)$ for some $n \in \mathbb{N}$, then we have $\mu(X_{\mathrm{fin}}(G) \times G) = 0$ and therefore, as $X_{\mathrm{fin}}(G) \times G$ is a subgroup of G^2 and μ is an invariant mean, we have $[G : X_{\mathrm{fin}}(G)] = \infty$. Thus either way, the assumption of Lemma 8.1 is satisfied.

Taking $\varepsilon = \frac{\mu(\mathrm{Comm}(G))}{2}$ and taking $n \in \mathbb{N}$ such that $n > \frac{1}{\varepsilon}$, we then get elements $h_1, \ldots, h_n \in G$ such that

$$\sum_{1 \leq i < j \leq n} \mu(C_G(h_i^{-1} h_j) \times G) < \varepsilon.$$

In particular, taking $Y = \bigcup_{1 \leq i < j \leq n} C_G(h_i^{-1} h_j) \subseteq G$, we then have $\mu(Y \times G) < \varepsilon$ and therefore

$$\mu(\mathrm{Comm}(G) - Y \times G) \geq \mu(\mathrm{Comm}(G)) - \mu(Y \times G) > 2\varepsilon - \varepsilon = \varepsilon.$$

On the other hand, for $1 \leq i \leq n$ define $Y_i = (1, h_i)(\mathrm{Comm}(G) - Y \times G)$. Then $\mu(Y_i) \geq \varepsilon$ for all i since μ is left-invariant. However, given $(g, h) \in G^2 - Y \times G$ we have

$$(g, h) \in Y_i \iff [g, h_i^{-1} h] = 1 \iff h \in h_i C_G(g),$$

and therefore if $(g, h) \in Y_i \cap Y_j$ for some $i < j$, then $h_i C_G(g) = h_j C_G(g)$, or equivalently $g \in C_G(h_i^{-1} h_j)$, which is impossible since $g \notin Y$. Thus the subsets $Y_1, \ldots, Y_n$ are pairwise disjoint, implying that

$$1 = \mu(G^2) \geq \mu(Y_1 \cup \cdots \cup Y_n) = \sum_{i=1}^{n} \mu(Y_i) \geq n\varepsilon,$$

contradicting the fact that $n > \frac{1}{\varepsilon}$.

Hence we must have $\mu(X_n(G) \times G) > 0$ for some $n \in \mathbb{N}$, as required. $\qquad\square$

Proposition 8.3. *Let G be a FAF group. Then there exists a constant $\alpha > 0$ such that $\mu(\mathrm{Comm}(G)) \geq \alpha$ for any invariant mean μ on G^2.*

Proof. By Remark 2.16, G has a finite index subgroup $G_0 \leq G$ such that G_0' is finite and central in G_0. Then the function $\mu_0 \colon \mathcal{P}(G_0^2) \to [0, 1]$ defined by $\mu_0(A) = [G : G_0]^2 \mu(A)$ is an invariant mean on G_0^2, and if we have $\mu_0(\mathrm{Comm}(G_0)) > 0$ then it also follows that

$$\mu(\mathrm{Comm}(G)) \geq \mu(\mathrm{Comm}(G_0)) = [G : G_0]^{-2} \mu_0(\mathrm{Comm}(G_0)) > 0.$$

We may thus assume, by replacing G with G_0 and multiplying α by $[G : G_0]^{-2}$ if necessary, that G' is finite and central in G. In particular, G is 2-step nilpotent.

Now let $g_1 = [x_1, y_1], \ldots, g_n = [x_n, y_n]$ be all the possible commutators in G, and let $H = \bigcap_{i=1}^{n} (C_G(x_i) \cap C_G(y_i))$. Note that we have $[G : C_G(x_i)], [G : C_G(y_i)] \leq |G'| < \infty$ for all i, and therefore $[G : H] < \infty$. On the other hand, since μ is an invariant mean on G^2, we obtain

$$\mu(\mathrm{Comm}(G)) = \frac{1}{n} \sum_{i=1}^{n} \mu\big((x_i, y_i^{-1}) \mathrm{Comm}(G)\big) \geq \frac{1}{n} \mu\left(\bigcup_{i=1}^{n} (x_i, y_i^{-1}) \mathrm{Comm}(G) \right)$$

$$\geq \frac{1}{n} \mu\left(H^2 \cap \bigcup_{i=1}^{n} (x_i, y_i^{-1}) \mathrm{Comm}(G) \right)$$

$$= \frac{1}{n} \mu\big(\{(x, y) \in H^2 : [x_i^{-1} x, y_i y] = 1 \text{ for some } i \in \{1, \ldots, n\}\}\big)$$

$$= \frac{1}{n} \mu\big(\{(x, y) \in H^2 : [x, y] = [x_i, y_i] \text{ for some } i \in \{1, \ldots, n\}\}\big),$$

where the last equality follows from commutator calculus (in the 2-step nilpotent group G) and the fact that $[x_i, y] = [x, y_i] = 1$ for all $(x, y) \in H^2$. But since the $g_i = [x_i, y_i]$ are all the possible commutators in G, it follows that $\mu(\mathrm{Comm}(G)) \geq \frac{1}{n} \mu(H^2) = \frac{1}{n[G:H]^2} > 0$, so we obtain the result by setting $\alpha = \frac{1}{n[G:H]^2}$. $\qquad\square$

Theorem 8.4. *Let G be an amenable group, and let μ be a finitely additive left invariant probability mean on G^2. Then $\mu(\mathrm{Comm}(G)) > 0$ if and only if G is finite-by-abelian-by-finite.*

Proof. If G is FAF, then $\mu(\mathrm{Comm}(G)) > 0$ by Proposition 8.3. Conversely, if $\mu(\mathrm{Comm}(G)) > 0$, then we have $\mu(X_n(G) \times G) > 0$ for some $n \in \mathbb{N}$ by Corollary 8.2. Define $\mu' \colon \mathcal{P}(G) \to [0, 1]$ by setting $\mu'(A) = \mu(A \times G)$ for all $A \subseteq G$, and note that μ' is clearly a left invariant finitely additive probability mean, and therefore a CCM, on G. Since $\mu'(X_n(G)) > 0$, it then follows from Proposition 5.9 that G is FAF. $\qquad\square$

Remark 8.5. A natural question is whether $\mu(\mathrm{Comm}(G))$ depends on the choice of an invariant mean μ on a FAF group G. We have not been able to prove or to find a counterexample for this. However, if a FAF group G is residually finite (for instance, when it is finitely generated), then the independence of the choice of μ follows from the proof of Theorem 7.2.

Residually amenable groups. We may similarly deduce that if G is a residually amenable group, then it has non-zero degree of commutativity if and only if it is is FAF. In particular, for a residually amenable G, we can set $\mathrm{dc}_{\mathrm{RAm}}(G)$ to be the infimum of all $\mu(\mathrm{Comm}(H))$, where H ranges over amenable quotients of G and where μ ranges over invariant means on H^2. If G is FAF, then it is amenable (since finite and abelian groups are amenable and since amenability is closed under extensions), and so $\mathrm{dc}_{\mathrm{RAm}}(G)$ is equal to the infimum of all $\mu(\mathrm{Comm}(G))$, where μ ranges over invariant means on G^2, which is positive by Proposition 8.3.

Conversely, if $\mathrm{dc}_{\mathrm{RAm}}(G) > 0$, then any amenable quotient of G is FAF by Theorem 8.4, and therefore virtually metabelian by Remark 2.16. Since G is residually amenable, it follows that the residual $\mathrm{Res}(G)$ of G is residually metabelian, and therefore metabelian since being metabelian is defined by a group law. On the other hand, since finite groups are amenable we have $\mathrm{dc}_{\mathrm{RF}}(G/\mathrm{Res}(G)) \geq \mathrm{dc}_{\mathrm{RAm}}(G) > 0$, and so $G/\mathrm{Res}(G)$ is virtually abelian: see Remark 5.4. It follows that G is metabelian-by-(virtually abelian), and so virtually solvable of derived length 3; in particular, G is amenable. We therefore have $\mu(\mathrm{Comm}(G)) \geq \mathrm{dc}_{\mathrm{RAm}}(G) > 0$ for any invariant mean μ on G^2, and so G is FAF by Theorem 8.4.

Counterexamples. Here we show that Theorem 8.4 does not hold for CCMs. That is, there exists a group G and CCMs on G^2 such that the degrees of commutativity with respect to these CCMs on G^2 are 0 and 1.

Example 8.6. In this example we shall find a group H and two CCMs μ_0, μ_1 on H^2 such that $\mu_0(\mathrm{Comm}(H)) = 0$ whereas $\mu_1(\mathrm{Comm}(H)) = 1$. Here H will be a nilpotent group, and therefore an amenable group. In fact, H will be finite-by-abelian, FA (and hence also FAF). But H is not abelian, and so one has that product means on H^2 give a degree of commutativity strictly between 0 and 1. The point being that neither μ_0 nor μ_1 are product means on the square, nor invariant means.

Let $\mathbb{F}_2$ be the field of order 2, and consider the group

$$H = \left\{ h(\mathbf{x}, \mathbf{y}, z) := \begin{pmatrix} 1 & \mathbf{x} & z \\ 0 & 1 & \mathbf{y} \\ 0 & 0 & 1 \end{pmatrix} : \mathbf{x}, \mathbf{y} \in V, z \in \mathbb{F}_2 \right\},$$

where $V = \bigoplus_{i=1}^{\infty} \mathbb{F}_2$ and the group operation is "matrix multiplication":

$$h(\mathbf{x}, \mathbf{y}, z) \cdot h(\mathbf{x}', \mathbf{y}', z') = h(\mathbf{x} + \mathbf{x}', \mathbf{y} + \mathbf{y}', z + z' + \langle \mathbf{x}, \mathbf{y}' \rangle);$$

here $\langle \mathbf{x}, \mathbf{y}' \rangle = \sum_{i=1}^{\infty} x_i y_i'$ for $\mathbf{x} = (x_1, x_2, \ldots)$ and $\mathbf{y}' = (y_1', y_2', \ldots)$. One may check that $Z(H) = \{h(\mathbf{0}, \mathbf{0}, z) : z \in \mathbb{F}_2\}$ and that $[H : C_H(h)] = 2$ for all $h \in H - Z(H)$. We should thus expect to have "$\mathrm{dc}(H) = \frac{1}{2}$" for any "reasonable" definition of dc. In particular, $\mathrm{dc}_{\mathrm{CCM}}(H) = \frac{1}{2}$.

However, we will show that there exists a CCM μ_1 on H^2 for which $\mu_1(\mathrm{Comm}(H)) = 1$, by using the construction of CCMs in §3. In order to do this, it is enough to show that the set $\mathcal{P}_{\mathrm{fin}}(\mathrm{Comm}(H))$ generates a filter together with the sets of the form

$$\mathcal{F}_{\varepsilon}(K) := \left\{ F \in \mathcal{P}_{\mathrm{fin}}(H^2) : \left| \frac{|gK \cap F|}{|F|} - \frac{1}{[H^2 : K]} \right| < \varepsilon \text{ for all } g \in H^2 \right\}$$

where $K \leq H^2$ and $\varepsilon > 0$. In particular, let $K_1, \ldots, K_n \leq H^2$ and $\varepsilon > 0$. It is then enough to show that $\bigcap_{i=1}^{n} \mathcal{F}_{\varepsilon}(K_i)$ contains an element $F \subset H^2$ such that $[h_1, h_2] = 1$ for all $(h_1, h_2) \in F$.

Suppose, without loss of generality, that for some m we have $[H^2 : K_i] < \infty$ for $1 \leq i \leq m$ and $[H^2 : K_j] = \infty$ for $m + 1 \leq j \leq n$. For $m + 1 \leq j \leq n$, note that K_j is a finite index subgroup of $K_j Z(H)^2$, and thus $\widehat{K}_j := K_j Z(H)^2 / Z(H)^2$ is an infinite index subgroup of $(H/Z(H))^2 \cong V^4 \cong (\bigoplus_{i=1}^{\infty} \mathbb{F}_2, +)$. It follows that $\widehat{K}_j$ is an $\mathbb{F}_2$-subspace of $(H/Z(H))^2$ of infinite codimension, and hence is contained in an $\mathbb{F}_2$-subspace $\widehat{L}_j$ of finite codimension $N > -\log_2 \varepsilon$. Letting L_j be the preimage of $\widehat{L}_j$ under the quotient map $H^2 \to (H/Z(H))^2$, it then follows that we have $K_j < L_j \leq H^2$ and $\frac{1}{\varepsilon} < [H^2 : L_j] < \infty$.

Now let $K = \bigcap_{i=1}^{m} K_i \cap \bigcap_{j=m+1}^{n} L_j$. Then K is a finite index subgroup of H^2. By Lemma 8.7 below, there exists a (left) transversal F of K in H^2 such that $F \subset \mathrm{Comm}(H)$. By construction,

for any $g \in H^2$ we have

$$\frac{|gK_i \cap F|}{|F|} = \frac{1}{[H^2 : K_i]}$$

for $1 \leq i \leq m$ and

$$\left| \frac{|gK_j \cap F|}{|F|} - \frac{1}{[H^2 : K_j]} \right| = \frac{|gK_j \cap F|}{|F|} \leq \frac{|gL_j \cap F|}{|F|} = \frac{1}{[H^2 : L_j]} < \varepsilon$$

for $m + 1 \leq j \leq n$, implying that $F \in \bigcap_{i=1}^n \mathcal{F}_\varepsilon(K_i)$, as required.

Using a similar argument, we may show that $\mathcal{P}_{\mathrm{fin}}(\mathrm{Comm}(H)^c)$ together with the $\mathcal{F}_\varepsilon(K)$ generates a filter, and therefore there exists another CCM μ_0 on H^2 such that $\mu_0(\mathrm{Comm}(H)) = 0$. $\square$

This is the technical Lemma used in the example above.

Lemma 8.7. *Let H be as in Example 8.6, let $K \leq H^2$ be a finite index subgroup, and let $g \in H^2$. Then there exists an element $(h_1, h_2) \in gK$ such that $[h_1, h_2] = 1$.*

Proof. Since K has finite index in H^2 we have $K \cap (H \times \{1\}) = L_1 \times \{1\}$ and $K \cap (\{1\} \times H) = \{1\} \times L_2$ for finite index subgroups $L_1, L_2 \leq_{\mathrm{fin}} H$. Let $L = L_1 \cap L_2$. Then $L^2 \leq_{\mathrm{fin}} K \leq_{\mathrm{fin}} H^2$ and so we may assume, without loss of generality, that $K = L^2$.

Let $g = (g_1, g_2) \in H^2$, and suppose for contradiction that $[g_1 h_1, g_2 h_2] \neq 1$ for all $h_1, h_2 \in L$. Note that, as H is 2-step nilpotent, we have

$$[g_1 h_1, h_2] = [g_1 h_1, g_2]^{-1}[g_1 h_1, g_2 h_2],$$

and thus, since $[H, H]$ is cyclic of order 2, we have $[g_1 h_1, h_2] = 1$. Since $h_2 \in L$ was arbitrary, it follows that $L \subseteq C_H(g_1 h_1)$ for all $h_1 \in L$. Taking $h_1 = 1$ we get $L \subseteq C_H(g_1)$, implying that $L \subseteq C_H(g_1 h) \cap C_H(g_1) \subseteq C_H(h)$ for all $h \in L$. In particular, L is abelian.

On the other hand, L has finite index in H, so this contradicts the fact that H is not virtually abelian (see [MTVV21, §7.3]). $\square$

Remark 8.8. The CCM μ in Example 8.6 is not an invariant mean, as for any invariant mean μ on G^2, where G is a non-abelian group, we have $\mu(\mathrm{Comm}(G)) < 1$. Indeed, if $g \in G$ is any non-central element, then we have

$$\mathrm{Comm}(G)^c \cup (1, g)\,\mathrm{Comm}(G)^c = \{(x, y) \in G^2 : [x, y] \neq 1 \text{ or } [x, g^{-1}y] \neq 1\}$$

$$= \{(x, y) \in G^2 : [x, y] = [x, g^{-1}y] = 1\}^c$$

$$\supseteq \{(x, y) \in G^2 : [x, g] = 1\}^c = (G - C_G(g)) \times G,$$

so if μ is an invariant mean on G^2 then we get

$$\mu(\mathrm{Comm}(G)^c) \geq \frac{1}{2}\mu((G - C_G(g)) \times G) = \frac{1}{2}\left(1 - \frac{1}{[G : C_G(g)]}\right) \geq \frac{1}{4}$$

and therefore $\mu(\mathrm{Comm}(G)) \leq \frac{3}{4}$.

9. MEANS ON NON-AMENABLE GROUPS

Since the paper has discussed various means on groups that exhibit partial invariance, it is natural to ask whether this can be quantified. We do this via a defect function which measures the worst case in which a mean on a group deviates from invariance. We show that while this is defined as a real number in the interval $[0, 1]$, it turns out to be equal to either 0 or 1 with the former case corresponding to the case of amenability. Therefore, even though we construct means which are invariant on subgroups, the defect function sees these as generally no closer to an invariant mean than an arbitrary one.

Definition 9.1. Given a finitely additive probability mean μ on a group G, we define the *left defect* of μ as

$$\delta_\ell(\mu) = \sup\{|\mu(gA) - \mu(A)| : g \in G, A \subseteq G\} \in [0, 1],$$

and the *right defect* of μ as

$$\delta_r(\mu) = \sup\{|\mu(Ag) - \mu(A)| : g \in G, A \subseteq G\} \in [0, 1].$$

Moreover, we define the (*left*) *defect* of G to be

$$\delta_G = \inf_\mu \delta_\ell(\mu),$$

where the infimum ranges over all finitely additive probability means μ on G.

Remark 9.2. We are omitting the right defect for a group but it has an analogous definition and is easily seen to be equal to the left defect. That is, if μ is a finitely additive probability mean, we can define $\nu(A) := \mu(A^{-1})$, and then $\delta_\ell(\nu) = \delta_r(\mu)$.

Lemma 9.3. *Let μ be a finitely additive probability mean on a group G.*
  (i) *Suppose that $\delta_\ell(\mu) > 0$ and $\delta_r(\mu) < 1$. Then G admits a finitely additive probability mean $\overline{\mu}$ such that $\delta_\ell(\overline{\mu}) < \delta_\ell(\mu)$ and $\delta_r(\overline{\mu}) \leq \delta_r(\mu)$.*
  (ii) *Suppose that $\delta_r(\mu) > 0$ and $\delta_\ell(\mu) < 1$. Then G admits a finitely additive probability mean $\overline{\mu}$ such that $\delta_r(\overline{\mu}) < \delta_r(\mu)$ and $\delta_\ell(\overline{\mu}) \leq \delta_\ell(\mu)$.*

Proof. Suppose first that $\delta_\ell(\mu) > 0$ and $\delta_r(\mu) < 1$. Define $\overline{\mu}\colon \mathcal{P}(G) \to [0,1]$ by setting

$$\overline{\mu}(A) = \int \mu(hA)\, \mathrm{d}\mu(h)$$

for all $A \subseteq G$. It is then easy to see that $\overline{\mu}$ is a finitely additive probability mean on G satisfying $\delta_r(\overline{\mu}) \leq \delta_r(\mu)$. Now let $A \subseteq G$, let $m_- = \inf_{g \in G} \mu(gA)$, let $m_+ = \sup_{g \in G} \mu(gA)$, let $m = \frac{m_- + m_+}{2}$, let $D = \frac{m_+ - m_-}{2}$, and let $\delta = \delta_\ell(\mu)$. Note that $D = m_+ - m = m - m_-$ and that $m_+ - m_- \leq \delta$; in particular, $D \leq \frac{\delta}{2}$. Setting $B = \{h \in G : \mu(hA) \geq m\}$ we get $Bg^{-1} = \{h \in G : \mu(hgA) \geq m\}$ for any $g \in G$, which allows us to estimate

$$m_- + D\mu(Bg^{-1}) \leq \overline{\mu}(gA) \leq m + D\mu(Bg^{-1}).$$

In particular, we have

$$\overline{\mu}(gA) - \overline{\mu}(A) \leq \big(m + D\mu(Bg^{-1})\big) - (m_- + D\mu(B)) = D(1 + \mu(Bg^{-1}) - \mu(B))$$
$$\leq \frac{\delta}{2}\left(1 + \delta_r(\mu)\right)$$

for all $g \in G$ and $A \subseteq G$. Since $\delta_r(\mu) < 1$ and $\delta > 0$, this implies that $\delta_\ell(\overline{\mu}) \leq \frac{1+\delta_r(\mu)}{2}\delta < \delta$, proving (i).

Part (ii) follows in a similar way to part (i), or by applying part (i) to the finitely additive probability mean μ' on G defined by $\mu'(A) = \mu(A^{-1})$. $\qquad\square$

Proposition 9.4. *Let ν be a finitely additive probability mean on a group G, and suppose that $\delta_\ell(\nu) < 1$. Then G is amenable.*

Proof. Throughout this proof, all means on G will be assumed to be finitely additive probability means.

Let $\delta_G = \inf_{\nu'} \delta_\ell(\nu')$, where the infimum ranges over all means ν' on G. Then there is a sequence $(\mu_n)_{n \in \mathbb{N}}$ of means on G such that $\delta_G = \lim_{n \to \infty} \delta_\ell(\mu_n)$. Let ω be a non-principal ultrafilter on $\mathbb{N}$, and let $\mu\colon \mathcal{P}(G) \to [0,1]$ be defined by setting $\mu(A) = \lim_\omega \mu_n(A)$. It is then easy to verify that μ is a finitely additive probability mean on G and that $|\mu(gA) - \mu(A)| \leq \delta_G$ for all $g \in G$ and $A \subseteq G$. In particular, we have $\delta_\ell(\mu) = \delta_G$. Note that if $\delta_G = 0$ (respectively $\delta_r(\mu) = 0$), then μ is a left invariant (respectively right invariant) mean on G and so G is amenable.

Thus, suppose for contradiction that $\delta_G > 0$ and $\delta_r(\mu) > 0$; note that we also have $\delta_G \leq \delta_\ell(\mu) < 1$. By Lemma 9.3(ii), there exists a mean $\overline{\mu}$ on G satisfying $\delta_\ell(\overline{\mu}) \leq \delta_G$ (and so $\delta_\ell(\overline{\mu}) = \delta_G > 0$ by the definition of δ_G) and $\delta_r(\overline{\mu}) < \delta_r(\mu)$ (and so $\delta_r(\overline{\mu}) < 1$). But then, by Lemma 9.3(i) (applied to $\overline{\mu}$), it follows that there exists a mean $\overline{\overline{\mu}}$ on G satisfying $\delta_\ell(\overline{\overline{\mu}}) < \delta_\ell(\overline{\mu}) = \delta_G$, contradicting the definition of δ_G. $\qquad\square$

Theorem 9.5. *Let G be a group. Then the defect δ_G of G is equal to $\delta_\ell(\mu)$ for some finitely additive probability mean μ on G. Moreover, δ_G equals either 0 or 1, and is 0 precisely when G is amenable.*

Proof. Note that the proof of Proposition 9.4 already shows that the defect of G is realised on an actual mean, via ultralimits. However, we can also argue *a posteriori*.

If G is amenable then it admits a left invariant mean μ whose defect is 0, so $\delta_G = \delta_\ell(\mu) = 0$. If G is not amenable then Proposition 9.4 implies that $\delta_\ell(\mu) = 1$ for all means μ on G and hence $\delta_G = 1$. $\square$

10. Conjugacy classes

For a finite group, we have $dc(G) = cc(G)/|G| =: k(G)$, where $cc(G)$ is the number of conjugacy classes of G. For a residually finite group G, we similarly get $k_{RF}(G) := \inf_{N \trianglelefteq_{fin} G} k(G/N) = dc_{RF}(G)$. This motivates the following definition.

Definition 10.1. Let μ be a CCM on a group G. Define
$$k_\mu(G) := \inf\{\mu(S) : S \text{ is a set of representatives of conjugacy classes of } G\}.$$

For the following result, note that any FAF group is amenable and therefore admits a finitely additive bi-invariant probability mean.

Lemma 10.2. *Let G be a FAF group equipped with a finitely additive bi-invariant probability mean μ. Then $k_\mu(G) > 0$.*

Proof. Let $H \trianglelefteq_{fin} G$ be such that $N := [H, H]$ is finite. Suppose that S is a set of representatives of conjugacy classes of G, so that we have
$$H \cap S = \{h^{\varphi(h)} : h \in H\},$$
for some function $\varphi \colon H \to G$. Given any coset $A = gH \in G/H$, let $T_A = \{h \in H : \varphi(h) \in A\}$. Since μ is finitely additive, we have $\frac{1}{[G:H]} = \mu(H) = \sum_{A \in G/H} \mu(T_A)$; in particular, we have $\mu(T_{A_0}) \geq \frac{1}{[G:H]^2}$ for some $A_0 = g_0 H \in G/H$.

Given any $h \in T_{A_0}$, note that we have $g_0 = \varphi(h)\psi(h)$ for some $\psi(h) \in H$, and therefore
$$h^{g_0} = h^{\varphi(h)\psi(h)} = h^{\varphi(h)}[h^{\varphi(h)}, \psi(h)] \in h^{\varphi(h)}N,$$
implying that $g_0^{-1} T_{A_0} g_0 \subseteq SN = \bigcup_{z \in N} Sz$. Since μ is bi-invariant, we get $\mu(g_0^{-1} T_{A_0} g_0) = \mu(T_{A_0})$, and thus
$$\mu(S) = \frac{1}{|N|} \sum_{z \in N} \mu(Sz) \geq \frac{1}{|N|} \mu\left(\bigcup_{z \in N} Sz\right) \geq \frac{1}{|N|} \mu(T_{A_0}) \geq \frac{1}{|N| \cdot [G:H]^2},$$
implying that $k_\mu(G) \geq \frac{1}{|N| \cdot [G:H]^2} > 0$, as required. $\square$

The following example shows that the assumption that μ is bi-invariant, and not merely left invariant, is necessary in Lemma 10.2. By inspecting the argument, however, one may deduce that an assumption that μ is conjugacy invariant and left invariant on cosets would be enough.

Example 10.3. Let $G = \mathbb{Z}^2 \rtimes (\mathbb{Z}/4\mathbb{Z})$, where the generator t of $\mathbb{Z}/4\mathbb{Z}$ acts on $H = \mathbb{Z}^2$ as the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Then G has a set S of representatives of conjugacy classes of G such that $S \cap H \subseteq \{(x, y) \in \mathbb{Z}^2 : |y| \leq x\}$. Moreover, one can verify that G contains only finitely many conjugacy classes outside of H, and therefore $\mu(S) = \mu(S \cap H)$ for any CCM μ on G.

Now let μ_1 be a finitely additive left invariant probability mean on $\mathbb{Z}$, and define $\mu_2 \colon \mathcal{P}(\mathbb{Z}^2) \to [0, 1]$ by setting
$$\mu_2(A) = \int \mu_1(A_x) \, d\mu_1(x),$$
where $A_x = \{y \in \mathbb{Z} : (x, y) \in A\}$. It is easy to see that μ_2 is a finitely additive left invariant probability mean on $\mathbb{Z}^2$. Finally, define $\mu \colon \mathcal{P}(G) \to [0, 1]$ by setting
$$\mu(A) = \frac{1}{4} \sum_{i=0}^{3} \mu_2(t^i A \cap H),$$

which is again a finitely additive left invariant probability mean. We then get

$$\mu(S) = \mu(S \cap H) = \frac{1}{4}\mu_2(S \cap H) \leq \frac{1}{4}\int_{x \in \mathbb{N}} \mu_1(\{-x, -x+1, \ldots, x\})\,\mathrm{d}\mu_1(x) = \frac{1}{4}\int 0\,\mathrm{d}\mu_1 = 0,$$

even though G is virtually abelian.

Proposition 10.4. *Let G be a group equipped with a CCM μ, and suppose that $k_\mu(G) > 0$. Then G is FAF.*

Proof. Let $n \in \mathbb{N}$ be such that $k_\mu(G) > \frac{1}{n+1}$. We claim that then we have $\mu(X_n(G)) > 0$, where $X_n(G) = \{g \in G : [G : C_G(g)] \leq n\}$. The result will then follow from Proposition 5.9.

Indeed, we can choose representatives $S_0, \ldots, S_n$ of conjugacy classes of G as follows. For $g \in X_n(G)$, choose the representative of $[g]$ in each S_i arbitrarily. For $g \in G - X_n(G)$, note that the conjugacy class of g contains $[G : C_G(g)] \geq n+1$ elements, and so we can choose $n+1$ different representatives of $[g]$ in the sets $S_0, \ldots, S_n$. Then each element of $X_n(G)$ appears in the sets $S_0, \ldots, S_n$ at most $n+1$ times, whereas each element of $G - X_n(G)$ appears there at most once. Since μ is finitely additive, we then have

$$1 < (n+1)k_\mu(G) \leq \sum_{i=0}^{n} \mu(S_i) \leq \mu(G - X_n(G)) + (n+1)\mu(X_n(G)) = 1 + n \cdot \mu(X_n(G)),$$

and thus $\mu(X_n(G)) > 0$, as claimed. $\qquad\square$

APPENDIX: USING THE HAHN-BANACH THEOREM TO CONSTRUCT CCMs

In this appendix we give another proof of the existence of CCMs for any group. The reason to do so is that in Theorem 3.4 we invoke the Ultrafilter Lemma which is a weak version of the Axiom of Choice. However, one can also prove the existence of CCMs by using the Hahn–Banach Theorem which is strictly weaker than the Ultrafilter Lemma.

Theorem A.1 (Hahn–Banach Theorem for real normed vector spaces). *Let $(V, \|\cdot\|)$ be a real normed vector space. Suppose that S is a subspace of V and $f_S \colon S \to \mathbb{R}$ a bounded linear functional on S. Then there exists an extension f of f_S to V such that $f|_S = f_S$ and $\|f\|_{V^*} = \|f_S\|_{S^*}$.*

Remark A.2. Recall that if $(V, \|\cdot\|)$ is a normed vector space with $V \neq \{0\}$ and $f \colon V \to \mathbb{R}$ a linear functional then $\|f\|_{V^*} := \sup_{0 \neq x \in V} \frac{|f(x)|}{\|x\|}$.

Theorem A.3. *The Hahn–Banach Theorem for real normed vector spaces implies the existence of a CCM for any group G.*

Proof. It is sufficient to prove the existence of a linear functional, $m \colon \ell^\infty(G) \to \mathbb{R}$, of norm 1 such that $m(\mathbb{1}_{gH}) = \frac{1}{[G:H]}$ for all cosets gH in G. Indeed, if this is the case then setting $\mu(A) = m(\mathbb{1}_A)$ for any subset $A \subseteq G$ defines a CCM μ on G. To see that this defines a CCM observe that $\mu(G) = m(\mathbb{1}_G) = 1$ by definition, the finite additivity of μ is a consequence of the linearity of m, and positivity of μ follows from the fact that for all $A \subsetneq G$,

$$m(\mathbb{1}_A) = m(\mathbb{1}_G) - m(\mathbb{1}_{A^c}) = 1 - m(\mathbb{1}_{A^c}) \geq 1 - \|m\|_{\ell^\infty(G)^*}\|\mathbb{1}_{A^c}\|_\infty = 0.$$

Therefore, $\mu(A) \geq 0$ for all $A \subseteq G$.

Recall that $\mathcal{C} \subseteq \mathcal{P}(G)$ denotes the Boolean subring generated by cosets of subgroups of G, Definition 6.1. By Proposition 6.5, there exists a unique CCM μ on $\mathcal{C}$.

Now let $\mathfrak{C} \leq \ell^\infty(G)$ be the subspace spanned by $\{\mathbb{1}_S : S \in \mathcal{C}\}$. Note that, since $\mathcal{C}$ is a subring of $\mathcal{P}(G)$, any element of $\mathfrak{C}$ can be expressed as $\sum_{i=1}^{n} a_i \cdot \mathbb{1}_{S_i}$, where $a_1, \ldots, a_n \in \mathbb{R}$ and where $S_1, \ldots, S_n \in \mathcal{C}$ form a partition of G. We then define a functional $m_{\mathfrak{C}} \colon \mathfrak{C} \to \mathbb{R}$ by setting

$$m_{\mathfrak{C}}\left(\sum_{i=1}^{n} a_i \cdot \mathbb{1}_{S_i}\right) := \sum_{i=1}^{n} a_i \cdot \mu(S_i).$$

To see that $m_{\mathfrak{C}}$ is well-defined, note that if $\sum_{i=1}^{n} a_i \cdot \mathbb{1}_{S_i} = \sum_{j=1}^{m} b_j \cdot \mathbb{1}_{T_j}$ are two different expressions of an element $f \in \mathfrak{C}$ as above, then we have

$$\sum_{i=1}^{n} a_i \cdot \mathbb{1}_{S_i} = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} \cdot \mathbb{1}_{U_{ij}} = \sum_{j=1}^{m} b_j \cdot \mathbb{1}_{T_j},$$

where $U_{ij} = S_i \cap T_j$ and where $a_i = c_{ij} = b_j$ whenever $U_{ij} \neq \varnothing$. It then follows by the finite additivity of μ that

$$\sum_{i=1}^{n} a_i \cdot \mu(S_i) = \sum_{i=1}^{n} \left(a_i \sum_{j=1}^{m} \mu(U_{ij}) \right) = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} \cdot \mu(U_{ij}) = \sum_{j=1}^{m} \left(b_j \sum_{i=1}^{n} \mu(U_{ij}) \right) = \sum_{j=1}^{m} b_j \cdot \mu(T_j),$$

as required. A similar argument shows that $m_{\mathfrak{C}}$ is linear.

Finally, $m_{\mathfrak{C}}$ is bounded since for any $f = \sum_{i=1}^{n} a_i \cdot \mathbb{1}_{S_i} \in \mathfrak{C}$ as above we have,

$$|m_{\mathfrak{C}}(f)| \leq \sum_{i=1}^{n} |a_i| \cdot \mu(S_i) \leq \max_{1 \leq i \leq n} |a_i| \cdot \sum_{i=1}^{n} \mu(S_i) = \max_{1 \leq i \leq n} |a_i| \cdot \mu(G) = \max_{1 \leq i \leq n} |a_i| = \|f\|_{\infty}.$$

Together with the fact that $m_{\mathfrak{C}}(\mathbb{1}_G) = \|\mathbb{1}_G\|_{\infty} = 1$, this forces the norm of $m_{\mathfrak{C}}$ to be 1.

By Theorem A.1, the functional $m_{\mathfrak{C}}$ then extends to a linear functional $m \colon \ell^{\infty}(G) \to \mathbb{R}$ which satisfies the required properties. $\qquad\square$

Remark A.4. We note that Theorem 1.1 from [Neu54] is used in Proposition 6.5 and so is an ingredient in both the construction using ultrafilters as well as this one using the Hahn–Banach Theorem.

REFERENCES

[AMV17] Yago Antolín, Armando Martino, and Enric Ventura, *Degree of commutativity of infinite groups*, Proc. Amer. Math. Soc. **145** (2017), no. 2, 479–485.

[Bar18] Laurent Bartholdi, *Amenability of groups and G-sets*, Sequences, groups, and number theory (Valérie Berthé and Michel Rigo, eds.), Springer, 2018, pp. 433–544.

[DS88] Nelson Dunford and Jacob T. Schwartz, *Linear operators, Part 1: General theory*, Pure and Applied Mathematics, no. VII, Interscience Publishers, 1988.

[DS14] Joe Diestel and Angela Spalsbury, *The joys of Haar measure*, Graduate Studies in Mathematics, no. 150, American Mathematical Society, 2014.

[ET68] Paul Erdős and Paul Turán, *On some problems of a statistical group-theory. IV*, Acta Math. Hung. **19** (1968), no. 3-4, 413–435.

[Føl55] Erling Følner, *On groups with full Banach mean value*, Math. Scand. **3** (1955), 243–254.

[Gal70] Patrick X. Gallagher, *The number of conjugacy classes in a finite group*, Math. Z. **118** (1970), no. 3, 175–179.

[Gus73] William H. Gustafson, *What is the probability that two group elements commute?*, Am. Math. Mon. **80** (1973), no. 9, 1031–1034.

[Hop21] John Hopfensperger, *When is an invariant mean the limit of a Følner net?*, Topol. Appl. **300** (2021), 107755.

[HR12] Karl H. Hofmann and Francesco G. Russo, *The probability that x and y commute in a compact group*, Math. Proc. Camb. Philos. Soc. **153** (2012), no. 3, 557–571.

[LP00] L. Lévai and L. Pyber, *Profinite groups with many commuting pairs or involutions*, Arch. Math. (Basel) **75** (2000), no. 1, 1–7.

[MTVV21] Armando Martino, Matthew C. H. Tointon, Motiejus Valiunas, and Enric Ventura, *Probabilistic nilpotence in infinite groups*, Isr. J. Math. **244** (2021), no. 2, 539–588.

[Neu29] John von Neumann, *Zur allgemeinen Theorie des Masses*, Fundam. Math. **13** (1929), 73–116.

[Neu54] B. H. Neumann, *Groups covered by permutable subsets*, J. London Math. Soc. **s1-29** (1954), no. 2, 236–248.

[Neu89] Peter M. Neumann, *Two combinatorial problems in group theory*, Bull. London Math. Soc. **21** (1989), no. 5, 456–458.

[Rus79] David Rusin, *What is the probability that two elements of a finite group commute?*, Pac. J. Math. **82** (1979), no. 1, 237–247.

[Toi20] Matthew C. H. Tointon, *Commuting probabilities of infinite groups*, J. London Math. Soc. **101** (2020), no. 3, 1280–1297.

[Val19a] Motiejus Valiunas, *Negligibility of elliptic elements in ascending HNN-extensions of $\mathbb{Z}^m$*, Comm. Algebra **47** (2019), no. 12, 5101–5120.

[Val19b]   _____, *Rational growth and degree of commutativity of graph products*, J. Algebra **522** (2019), 309–331.

[Wil03]   George A. Willis, *Harmonic analysis and amenability*, Introduction to Banach Algebras, Operators, and Harmonic Analysis, London Mathematical Society Student Texts, no. 57, Cambridge University Press, 2003, pp. 73–134.

(A. Martino) Mathematical Sciences, University of Southampton, University Road, Southampton SO17 1BJ, United Kingdom
Email address: `A.Martino@soton.ac.uk`

(M. Valiunas) Instytut Matematyczny, Uniwersytet Wrocławski, plac Grunwaldzki 2, 50-384 Wrocław, Poland
Email address: `motiejus.valiunas@math.uni.wroc.pl`