Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

Commuting Probability in Compact Groups

by

João Pedro Papalardo Azevedo

Advisor: Prof. Dr. Pavel Shumyatsky

Brasília

2023

Contents

Resumo

Seja G um grupo topológico compacto com um subgrupo fechado K e medidas de Haar normalizadas μ e ν, respectivamente. Considere o subconjunto fechado $C = \{(x, y) \in K \times G \,|\, xy = yx\}$ de $K \times G$ e defina a probabilidade de comutação relativa de K em G por $Pr(K, G) = (\mu \times \nu)(C)$. Esse valor representa a probabilidade de escolher aleatoriamente um elemento de K e um de G que comutam. Se $K = G$, obtemos a probabilidade de comutação de G, uma medida de quão abeliano o grupo é. Ao longo do tempo, estudou-se o impacto dos valores $Pr(G)$ e $Pr(K, G)$ na estrutura de G. Por exemplo, um teorema de P.M. Neumann [40] assegura que, se G é finito e ϵ é um número positivo, $Pr(G) \geq \epsilon$ implica que G possui subgrupo H tal que $[G : H]$ e $|H'|$ são ϵ-limitados. Nosso objetivo é o de estudar propriedades similares relacionadas à probabilidade de comutação relativa.

Em [9], Detomi e Shumyatsky obtêm resultados estruturais sobre um grupo finito G admitindo subgrupo K tal que $Pr(K, G) \geq \epsilon$. Eles provam que existem subgrupos T de G e B de K tais que os índices $[G : T]$ e $[K : B]$ e a ordem de $[T, B]$ são ϵ-limitados. Nós estendemos esse resultado para grupos compactos e demonstramos alguns corolários.

Se G é um grupo topológico compacto e $x \in G$, denote por $\langle x \rangle$ o subgrupo fechado gerado por x. Será demonstrado que, se $Pr(\langle x \rangle, G) \geq \epsilon$ para todo x em um subgrupo fechado K de G, então existem um subgrupo aberto T de G e um inteiro e tais que o índice $[G : T]$ e o número e são ϵ-limitados e $[T, K^e] = 1$. Este resultado representa uma interpretação probabilística da noção de expoente num grupo. Diversos corolários serão demonstrados, todos relacionados à noção de expoente de um grupo. Por fim, consideramos a situação mais geral em que $Pr(\langle x \rangle, G)$ é positivo para todo x em $K \leq G$. Provaremos que G possui subgrupo aberto T de forma que todo $x \in K$ possui uma potência x^l, onde l não necessariamente é fixo, que centraliza T.

Palavras-chave: Probabilidade de comutação, medida de Haar, grupos compactos, subgrupos monotéticos

Abstract

Let G be a compact topological group with a closed subgroup K and normalized Haar measures μ and ν, respectively. Consider the closed subset $C = \{(x, y) \in K \times G \,|\, xy = yx\}$ of $K \times G$ and define the relative commuting probability of K in G by $Pr(K, G) = (\mu \times \nu)(C)$. This value represents the probability of choosing at random an element of K and one of G that commute. If $K = G$, we get the commuting probability of G, a measure of how close to be abelian the group is. For years, the influence of $Pr(G)$ and $Pr(K, G)$ on the structure of G has been studied. For example, a theorem of P.M. Neumann [40] ensures that, if G is finite and ϵ is a positive number, $Pr(G) \geq \epsilon$ implies that G has a subgroup H such that $[G : H]$ and $|H'|$ are ϵ-bounded. Our goal is to study similar properties concerning relative commuting probability.

In [9], Detomi and Shumyatsky prove structural resuts about a finite group G having a subgroup K such that $Pr(K, G) \geq \epsilon$. They prove that there exist subgroups T of G and B of K such that the indices $[G : T]$ and $[K : B]$ and the order of $[T, B]$ are ϵ-bounded. We extend this result to compact groups and prove corollaries of it.

If G is a topological group and $x \in G$, denote by $\langle x \rangle$ the closed subgroup generated by x. We prove that, if $Pr(\langle x \rangle, G) \geq \epsilon$ for every x in a closed subgroup K of G, then there are an open subgroup T of G and an integer e such that the index $[G : T]$ and the number e are ϵ-bounded and $[T, K^e] = 1$. This result represents a probabilistic interpretation of the notion of exponent in a group. Several corollaries are proved, all related to the notion of exponent. Finally, we consider the more general situation where $Pr(\langle x \rangle, G)$ is positive for all x in $K \leq G$. We prove that G has an open subgroup T in such a way that every $x \in K$ has a power x^l, where l is not necessarily fixed, centralizing T.

Keywords: Commuting probability, Haar measure, compact groups, monothetic subgroups.

Chapter 1

Introduction

Let G be a finite group. Erdös and Turán introduced the commuting probability of G in [13] as the ratio

$$Pr(G) = \frac{|\{(x,y) \in G \times G \,|\, xy = yx\}|}{|G|^2},$$

which measures the probability that two randomly chosen elements in G commute. For an arbitrary $x \in G$, the number of ordered pairs (x,y) such that x and y commute is $|C_G(x)|$, so the numerator above can be rewritten as

$$Pr(G) = \frac{\sum_{x \in G} |C_G(x)|}{|G|^2} = \frac{1}{|G|} \sum_{x \in G} \frac{1}{[G : C_G(x)]}. \qquad (*)$$

In an analogous way, if K is a subgroup of G, the relative commuting probability of K in G was defined in [14] as the number

$$Pr(K,G) = \frac{|\{(x,y) \in K \times G \,|\, xy = yx\}|}{|K||G|}.$$

If we let $K = G$, then what we have is the commuting probability of G.

The concept of commuting probability can be extended to infinite groups following ideas of Gustafson [21]. Let G be a compact Hausdorff topological group. The Borel σ-algebra $\mathcal{M}$ of G is the one generated by all closed subsets of G. There is a unique measure μ that can be defined in the measurable space $(G, \mathcal{M})$ such that $\mu(G) = 1$ and $\mu(xS) = \mu(S)$ for every measurable subset S of G and $x \in G$, the so-called normalized left Haar measure of G. If K is a closed subgroup of G, consider the set $C = \{(x,y) \in K \times G \,|\, xy = yx\}$, which is closed in $K \times G$, since it is the preimage of 1 under the continuous map $f(x,y) = [x,y]$.

Denoting the normalized Haar measures of K and G by ν and μ, respectively, and equipping $K \times G$ with the product measure $\nu \times \mu$, the probability that a random element of G commutes with a random element of K is $(\nu \times \mu)(C)$. Since finite groups are compact with the discrete topology, this definition is indeed a generalization of the previous one.

The notion of commuting probability has been used to derive structural results about G, and conversely the structure of the group can also be used to derive lower and upper bounds for its commuting probability. For example, if G is a compact group, Gustafson [21] proves that either G is abelian or $Pr(G) \leq \frac{5}{8}$. If G is a finite group, representation theory was used by Lescot in [34] to prove that $Pr(G) > \frac{1}{2}$ implies that G is nilpotent and G' has order at most 2, and in [35] he proved that $Pr(G) > \frac{1}{12}$ implies that G is soluble without recurring to the Classification of Finite Simple Groups. On the other hand, R.M. Guralnick and G. Robinson prove in [18], among several other results, that a soluble group of derived length n has its commuting probability bounded from above by a function depending on n only.

An arbitrary group is said to be a BFC-group if its conjugacy classes are finite and have bounded size. In particular, if G is finite and all conjugacy classes of G have at most n elements, then $[G : C_G(x)] \leq n$ and, by $(*)$, $Pr(G) \geq \frac{1}{n}$, proving that BFC-groups have high commuting probability. This result also holds in the compact case and suggests that the property of having high commuting probability is related to having finite conjugacy classes. Moreover, in the context of BFC-groups, B. H. Neumann [39] proved the following result in 1954.

Theorem. *If G is a BFC-group, then G' is finite.*

We say that a quantity is "$(a, b, \dots)$"-bounded whenever it is possible to bound it from above by a function of the parameters $a, b \dots$ only. As we remarked before, compact BFC-groups have positive commuting probability. It turns out that compact groups with positive commuting probability have an open subgroup which is a BFC-group, as we will prove as a corollary of Theorem A. In the context of finite groups, this is the content of P.M. Neumann's theorem [40].

Theorem. *Let G be a finite group such that $Pr(G) \geq \epsilon$. Then G has a subgroup H, of ϵ-bounded index, such that $|H'|$ is also ϵ-bounded.*

In recent years, some generalizations of B.H. Neumann's theorem were obtained. For example, in [2], Acciarri and Shumyatsky prove that if G is an

arbitrary group having a subgroup K such that every conjugacy class in G of an element of K has size at most n, then the normal closure, that is, the smallest normal subgroup of G containing K, has finite commutator subgroup of n-bounded order. Using this result in particular, Detomi and Shumyatsky [9] obtained a generalization of P.M. Neumann's theorem for finite groups, in the direction of relative commuting probability, as stated below.

Theorem. *Let $\epsilon > 0$ and let G be a finite group having a subgroup K such that $Pr(K, G) \geq \epsilon$. Then there is a normal subgroup $T \leq G$ and a subgroup $B \leq K$ such that the indices $[G : T]$ and $[K : B]$ and the order of $[T, B]$ are ϵ-bounded.*

Here we denote by $[T, B]$ the subgroup generated by all commutators $[t, b]$ with $t \in T$ and $b \in B$. Our first main result, Theorem A, is the extension of the above theorem to the class of compact groups. It was proved in [5].

Theorem A. *Let G be a compact group having a closed subgroup K such that $Pr(K, G) \geq \epsilon$ for $\epsilon > 0$. Then there is a normal closed subgroup $T \leq G$ and a closed subgroup $B \leq K$ such that the indices $[G : T]$ and $[K : B]$ and the order of $[T, B]$ are ϵ-bounded.*

If we let $K = \gamma_n(G)$, the nth term of the lower central series of G, then it can be proved following [9] that G has an open normal subgroup of ϵ-bounded index which is nilpotent of class at most $n + 1$. In the same direction, if we let $K = G^{(n)}$, we prove in Theorem 3.3.5 that G has an open normal subgroup which is soluble of derived length at most $n + 1$. If G is a profinite group and K is a Sylow pro-p subgroup of G, we show in Theorem 3.3.6 that K has an open class-2-nilpotent subgroup L such that the index $[K : L]$ and the order of L' are ϵ-bounded. In Corollary 4.2.5 we prove that if $K = G_0$, the connected component of 1 in G, then there is a normal profinite subgroup Δ of G such that $G = G_0\Delta$ and the centralizer of G_0 in G is open.

The study of different structural properties in groups requires defining different probability measures. As an example, we mention [12], where the authors study probabilistic nilpotency in finite groups by investigating the proportion of ordered triples (x, y, z) of $G \times G \times G$ such that $[x, y, z] = 1$. The commuting probability has been used to conclude, mostly, properties related to commutativity in groups, such as nilpotency, solubility and finiteness of commutator subgroups. Our next main result, however, concerns a probabilistic interpretation of the exponent of a compact group.

If S is a subset of a topological group G, then we denote by $\langle S \rangle$ the subgroup topologically generated by S, and a monothetic subgroup of G is a topologically 1-generated subgroup. For a given $\epsilon > 0$, we study a compact group G such that $Pr(\langle x \rangle, G) \geq \epsilon$ for every monothetic subgroup $\langle x \rangle \leq K$, where K is a closed subgroup of G. For example, if G has exponent e, then $Pr(\langle x \rangle, G) \geq \frac{1}{e}$ for all $x \in G$. More generally, if $G^e \leq Z(G)$, where $Z(G)$ denotes the center of G, then $Pr(\langle x \rangle, G) \geq \frac{1}{e}$ for all $x \in G$. Theorem B states a partial converse of this idea. It appeared in [5].

Theorem B. *Let $\epsilon > 0$ and let G be a compact group such that $Pr(\langle x \rangle, G) \geq \epsilon$ for every x in a closed subgroup K of G. Then there is an ϵ-bounded number e and an open normal subgroup T of G such that $[G : T]$ is ϵ-bounded and $[K^e, T] = 1$.*

Theorem B will allow us to give probabilistic interpretations of some well-known results in group theory. For example, we give a variation of Zelmanov's solution of the Restricted Burnside Problem [52, 53] stating that an r-generated finite group of exponent e has order bounded in terms of r and e solely. We prove in Theorem 4.3.2 that if G is an r-generated compact group and $Pr(\langle x \rangle, G) \geq \epsilon$ for all $x \in G$, then there is an open abelian subgroup N such that $[G : N]$ is ϵ-bounded. Moreover, if G is a profinite group such that $Pr(\langle x \rangle, G) \geq \epsilon$ for every x in a Sylow pro-p subgroup, we prove in Theorem 4.3.9 that G has a series of ϵ-bounded lenght where all factors are pro-p, pro-p' or Cartesian products of finite simple groups of orders not divisible by p. This is analogous to Hall and Higman's famous result [23]. Automorphisms of profinite groups are also considered. Let G be a profinite group and A be a noncyclic finite group of coprime automorphisms of G of order p^2, where p is a prime. Assume that $Pr(\langle x \rangle, G) \geq \epsilon$ for every $x \in C_G(\phi)$ and a nontrivial $\phi \in A$. Then there is an ϵ-bounded number e and a normal open subgroup of ϵ-bounded index T of G such that $[G^e, T] = 1$, as we will prove in Theorem 4.3.15. This extends the main result of [31].

In Chapter 5 we deal with a probabilistic variation of the concept of torsion in compact groups. We consider compact groups G having a subgroup K such that $Pr(\langle x \rangle, G) > 0$ for any $x \in K$. Note that this condition is satisfied whenever the subgroup K is torsion. More generally, the condition is satisfied whenever the image of K in $G/Z(G)$ is torsion. Theorem C was proved in [4].

Theorem C. *Let K be a subgroup of a compact group G. Then $Pr(\langle x \rangle, G) > 0$*

for any $x \in K$ if and only if G has an open normal subgroup T such that $K/C_K(T)$ is torsion.

In the past, compact torsion groups attracted considerable attention. Wilson showed that any such group possesses a characteristic series of finite length each of whose factors either is a pro-p group for some prime p or is isomorphic to a Cartesian product of isomorphic finite simple groups [50]. Again, if we let G be profinite and K be a Sylow pro-p subgroup, we prove in Theorem 5.2.5 that G admits a finite-length series where factors are pro-p, pro-p' or Cartesian products of finite simple groups. We prove other corollaries of Theorem C, as well as investigate, in Section 5.3, the relation between property $Pr(\langle x\rangle, G) > 0$ for every $x \in K$ and $Pr(\langle x\rangle, G) \geq \epsilon$ for every $x \in K$ and a positive ϵ. This is related to the question whether compact torsion groups have finite exponent.

Finally, some words on the organization of this text. Chapter 2 contains an overview of the results that are necessary to understand our work, as well as a brief introduction to measure theory and Haar measure on compact groups self-contained enough to serve as a first reading on the topic. Chapter 3 is where we define properly the commuting and relative commuting probability in compact groups, examples of groups with well-behaved commuting probabilities are considered, and Theorem A is proved in Section 3.3. In Chapter 4 we prove Theorem B and its corollaries. Theorem C is proved in Chapter 5, as well as its consequences. An Appendix, which can be read independently, is included. There we prove from scratch the existence and unicity of Haar measure up to a scaling factor. In the Appendix we avoid the language and methods of functional analysis everywhere it was possible for the benefit of the reader.

Chapter 2

Preliminaries

In this chapter we collect some general facts from group theory, topological groups and measure theory that are needed in the course of the text. References for the results not proved here are provided.

2.1 Groups

Given a positive integer k, a *group-word* $w = w(x_1, x_2, \ldots, x_k)$ is a nontrivial element of the free group F on countably many free variables $x_1, x_2, \cdots$. Only reduced words are considered in the text. Given a group G, a group word can also be regarded as a map from the direct product of k copies of G on G itself:

$$\begin{aligned} w\colon \overbrace{G \times \cdots \times G}^{k \text{ times}} &\to G \\ (g_1, \ldots, g_k) &\mapsto w(g_1, \ldots, g_k). \end{aligned}$$

If $g_1, \ldots, g_k \in G$, an element x of G of the form $x = w(g_1, \ldots, g_k)$ is called a w-value of G. The subgroup generated by all w-values is denoted by $w(G)$. It is called a *verbal* subgroup and is characteristic in G. Let us define the words, verbal subgroups and related concepts that we need along this work.

- Given a positive integer n, consider the power word x^n. The corresponding verbal subgroup is $G^n = \langle x^n \mid x \in G \rangle$. We say that G has *finite exponent* if there is an n such that $G^n = 1$.
- The *commutator word* $[x_1, x_2]$ is defined as $x_1^{-1}x_2^{-1}x_1x_2$. The corresponding verbal subgroup is the *commutator subgroup* of G. Moreover, if A and B

are arbitrary subsets of G, the subgroup $[A, B]$ is defined as

$$[A, B] = \langle [a, b] \,|\, a \in A, b \in B \rangle.$$

- A *multilinear commutator word* is a word obtained by nesting commutators, but always in different variables. As an example we have the word $[[x_1, x_2], [x_3, x_4]]$, whereas the Engel word $[x, y, y]$ is not.
- As a particular case of the above, the *lower central words* are defined recursively: if $x_1, x_2, \ldots, x_n \in G$ define $[x_1] = x_1$ and if $n \geq 2$ define $\gamma_n(x_1, x_2, \ldots, x_n) = [x_1, \ldots, x_n]$ as $[[x_1, x_2, \ldots, x_{n-1}], x_n]$. The verbal subgroup corresponding to γ_n is denoted by $\gamma_n(G)$. If $\gamma_{n+1}(G) = 1$, the group G is said to be *nilpotent* of class at most n.
- The *derived words* δ_n are defined recursively: $\delta_0(x_1)$ equals x_1 simply and for $n \geq 1$ we have $\delta_n = [\delta_{n-1}(x_1, \ldots, x_{2^{n-1}}), \delta_{n-1}(x_{2^{n-1}+1}, \ldots, x_{2^n})]$. The verbal subgroup corresponding to δ_n is called the *nth derived group* of G and is denoted by $G^{(n)}$. If $G^{(n)} = 1$ we say that G is *soluble* of derived length at most n.

The weight of a commutator is defined recursively: x has weight 1 and if the commutators u and v have weights n_1 and n_2, respectively, then $[u, v]$ has weight $n_1 + n_2$. The commutator $[x, y, [x, y, z]]$, for example, has weight 5.

If $x, g \in G$, the *conjugate* of x by g is $x^g = g^{-1}xg$. The set of all conjugates of x by elements of G is denoted x^G and called the *conjugacy class* of x. This notion can be extended to a subset X of G, and we denote by X^G the set

$$X^G = \bigcup_{x \in X} x^G$$

of conjugates of all elements of X. The *normal closure* of X in G is the subgroup $\langle X^G \rangle$, which is the smallest normal subgroup containing X, in the sense of inclusion. The *centralizer* of x in G is the subgroup $C_G(x) = \{g \in G \,|\, x^g = x\}$, and it is a well-known result [44, Theorem 1.6.1] that $|x^G| = [G : C_G(x)]$. If K is a subgroup of G, the centralizer of $x \in K$ is $C_K(x) = K \cap C_G(x)$. The *center* of G can be defined as $Z(G) = \bigcap_{x \in G} C_G(x)$.

Remark 2.1.1. Let $x \in G$ be any element and consider x^G. There is a natural bijection between x^G and the set $\{[x, g] \,|\, g \in G\}$ given by $x^g \mapsto x^{-1}x^g$. If we

need to count how many conjugates an element $x \in G$ has, it is often useful to count how many commutators of weight two are there which have x as an entry.

The following result will be used along the text without explicit mention.

Proposition 2.1.2. *Let G be a group and $K_1, K_2, \ldots, K_n$ be finite-index subgroups of G. Then*

$$\left[G : \bigcap_{i=1}^{n} K_i\right] \leq \prod_{i=1}^{n} [G : K_i].$$

Proof. Let $H = \bigcap_{i=1}^{n} K_i$. Let X denote the set of cosets of G modulo H and let X_i denote the set of cosets of G modulo H_i, $i = 1, \ldots, n$. The map from X to the cartesian product $X_1 \times \cdots \times X_n$ sending xH to $(xK_1, \ldots, xK_n)$ is a well-defined injection. □

Recall that we use the expression "$(a, b, \ldots)-$bounded" to mean that a quantity is bounded from above by a function depending on the parameters $a, b, \cdots$ only. In the previous proposition, we say that the index $[G : \bigcap_{i=1}^{n} K_i]$ is $([G : K_1], \ldots, [G : K_n])$-bounded, for example.

Let $\mathcal{X}$ be any group-theoretic property. We say that a group G is virtually-$\mathcal{X}$ if G has a finite-index subgroup satisfying property $\mathcal{X}$. For example, a virtually abelian group G has a finite-index subgroup K such that $[K, K] = 1$.

If ϕ is an automorphism of G, the centralizer of ϕ in G is the subgroup defined as $C_G(\phi) = \{x \in G \,|\, \phi(x) = x\}$, and if A is an automorphism group of G define $C_G(A) = \bigcap_{\phi \in A} C_G(\phi)$. The symbol $A^{\#}$ will be used to denote the set of nontrivial elements of the group A. Moreover, if N is a normal A-invariant subgroup of G, then every element of A induces an automorphism of G/N and A can be regarded as a group of automorphisms of G/N too. If G is a finite group, an automorphism group A of G is said to be *coprime* if $(|G|, |A|) = 1$. The following fact about coprime automorphisms will be needed. It is [17, Theorem 6.2.2 (iv)].

Lemma 2.1.3. *Let G be a finite group admitting a coprime group of automorphisms A and let N be a normal A-invariant subgroup of G. Then*

$$C_{G/N}(A) = C_G(A)N/N.$$

2.2 Topological Groups

For a more complete exposition, the reader can check [51].

Definition 2.2.1. A topological group is a triple $(G, \cdot, \tau)$ where τ is a topology on G such that the following maps are continuous with respect to the product topology on $G \times G$ and τ on G

$$\begin{array}{rccc} m\colon & G\times G & \to & G \\ & (x,y) & \mapsto & x\cdot y \end{array} \quad \text{and} \quad \begin{array}{rccc} \iota\colon & G & \to & G \\ & x & \mapsto & x^{-1} \end{array} .$$

It follows from Definition 2.2.1 that all group-words are continuous maps. In particular, the most important for us is the commutator map $(x, y) \mapsto [x, y]$. If $g \in G$ is a fixed element, the maps $x \mapsto xg$ and $x \mapsto gx$ are both homeomorphisms. As examples of topological groups, there are finite groups with discrete topology and $GL_n(\mathbb{R})$ with the subspace topology inherited from $\mathbb{R}^{n^2}$. We record some useful properties of topological groups.

Definition 2.2.2. A topological space X is compact if, for every open cover $X \subseteq \bigcup_{\lambda\in\Lambda} O_\lambda$ of X, there is a finite subcover $X \subseteq \bigcup_{i=1}^{n} O_i$.

Definition 2.2.3. A topological space X is *locally compact* if, for every element $x \in X$, there exists an open set O and a compact set K such that $x \in O \subseteq K$.

The real line $\mathbb{R}$ is an example of a locally compact space that is not compact, since every point is contained in a closed and bounded interval, but the collection of open sets $\{(n, n+2) \mid n \in \mathbb{Z}\}$ covers $\mathbb{R}$ and does not have a finite subcover.

The next result states some basic facts about the structure of a topological group. It is [51, Lemma 0.3.1].

Lemma 2.2.4. *Let G be a topological group.*

(a) Every open subgroup of G is closed, and every closed subgroup of finite index is open. If G is compact then every open subgroup of G has finite index.

(b) If H is a subgroup of G and K is a normal subgroup of G then H is a topological group with respect to the subgroup topology, G/K is a topological group with respect to the quotient topology and the quotient map q from G to G/K takes open sets to open sets.

We will also need the following fact.

Lemma 2.2.5. *Let X be a compact topological space and Y be a closed subset of X. Then Y is also compact.*

Proof. First we observe that $X \setminus Y$ is open on X. If $Y \subseteq \bigcup_{\lambda\in\Lambda} O_\lambda$ is an open cover of Y, then $(X \setminus Y) \cup \left(\bigcup_{\lambda\in\Lambda} O_\lambda\right)$ is an open cover of X. Extract a finite subcover $\bigcup_{i=1}^{n} O_i$ of it. Removing $X \setminus Y$ from this cover, if needed, provides a finite subcover of Y. □

If G is a topological group and $C, D \subseteq G$, consider the sets

$$CD = \{cd \,|\, c \in C, d \in D\} \text{ and } C^{-1} = \{c^{-1} \,|\, c \in C\}$$

and let bars denote topological closures. If $f : X \to Y$ is a continuous function between topological spaces, then the inclusion $f(\overline{A}) \subseteq \overline{f(A)}$ holds for any subset A of X. Indeed, if O is an open set containing $f(a)$ for $a \in \overline{A}$, then $f^{-1}(O)$ is open in X and contains a, hence there is $b \in f^{-1}(O) \cap A$. Furthermore $f(b) \in O \cap f(A)$, so a belongs to the closure of $f(A)$. If K is a subgroup of the compact group G, the previous argument and continuity of the group operation and inversion ensure that

$$\overline{K} \cdot \overline{K}^{-1} \subseteq \overline{K} \cdot \overline{K^{-1}} \subseteq \overline{K \cdot K^{-1}} = \overline{K}.$$

The closure $\overline{K}$ is then a subgroup of G, which will be compact by Lemma 2.2.5. As in the upcoming discussion we will be interested only in compact groups and their compact subgroups, a subgroup of G will always be considered closed from now on, unless explicitly stated otherwise.

The well-known Baire Category Theorem [30, p. 200] is going to be needed in the following form.

Theorem 2.2.6. *Let X be a compact topological space and let $\{F_n \,|\, n \in \mathbb{N}\}$ be a countable family of closed subsets of X. If $X = \bigcup_n F_n$, there exists some n such that F_n has nonempty interior.*

Concerning separation axioms, topological groups are not Hausdorff spaces in general, as we can always endow G with the trivial topology $\{\emptyset, G\}$. However, there is the following equivalence, which is [51, Lemma 0.3.1].

Proposition 2.2.7. *Let G be a topological group. Then G is Hausdorff if and only if all singletons $\{x\}$ with $x \in G$ are closed.*

Consider the commutator map $f : G \times G \to G$ given by $f(x, y) = [x, y]$. The subset $\{(x, y) \in G \times G \,|\, xy = yx\}$ is the preimage of 1 under the continuous map

f, and will be closed in $G \times G$ if 1 is a closed subgroup of G. In what comes next we need this set to be closed, so this fact and Proposition 2.2.7 are the reason why all topological groups considered in this work are Hausdorff. In short, the groups in this work satisfy the following convention.

Convention 2.2.8. A topological group G will always be compact and Hausdorff and a subgroup K of G will always be considered closed, unless stated otherwise.

We also need the following fact, which is Lemma 0.1.1(a) in [51]. Topological spaces satisfying the conclusion of Proposition 2.2.9 are called *normal spaces*.

Proposition 2.2.9. *Let X be a compact Hausdorff space. If U and V are disjoint closed subsets of X, there are disjoint open subsets A and B such that $U \subseteq A$ and $V \subseteq B$.*

A group G is *topologically generated* by a subset S if the *abstract* subgroup generated by S is a dense subgroup of G. We will refer to this notion simply as being *generated* by S.

Definition 2.2.10. A group G is *monothetic* if it is generated by a single element.

In profinite groups (see Section 2.3), monothetic groups can be isomorphic to $\mathbb{Z}/n\mathbb{Z}$ for some integer n or to the p-adic integers $\mathbb{Z}_p$, for example. In the non-profinite case we have the circle group $\mathbb{R}/\mathbb{Z}$, which is generated by the image of any irrational number. To see this, we prove that $H = \mathbb{Z} + \alpha\mathbb{Z}$ is dense in $\mathbb{R}$ for any irrational α. Define $frac(x) = x - \lfloor x \rfloor$, the factional part of x and see that $frac(x) \in H$ for every $x \in H$. Moreover, as α is irrational, if k and m are distinct integers, then $frac(m\alpha) \neq frac(k\alpha)$. Dividing $[0, 1]$ in n parts of equal length, the pigeonhole principle ensures that two of the numbers $0, frac(\alpha), \ldots, frac(n\alpha)$ belong to the same interval, thus there are two integers r and s such that $0 \leq y = frac(r\alpha) - frac(s\alpha) \leq 1/n$. In particular, $y \in H$. Now let $\epsilon > 0$ and $z \in \mathbb{R}$ be arbitrary. There is $y \in H$ with $0 \leq y < \epsilon$, by the previous argument, and an integer n such that $ny \leq z < (n+1)y$. This implies that $0 \leq z - ny < y < \epsilon$ and we conclude that H is dense in $\mathbb{R}$. The image of H in $\mathbb{R}/\mathbb{Z}$ generates this quotient topologically, and considering suitable products of such generators we can prove that the Cartesian products $\underbrace{(\mathbb{R}/\mathbb{Z}) \times \cdots \times (\mathbb{R}/\mathbb{Z})}_{k}$ are monothetic, with $k \in \mathbb{N}$, or even $k = \aleph_0$.

Definition 2.2.11. Let G be any group. The abstract subgroup consisting of all $x \in G$ such that $|x^G|$ is finite is the FC-center of G.

If G is a topological group, $FC(G)$ is not necessarily a closed subgroup of G. If H is any nonabelian finite group and H_i is an isomorphic copy of H for each $i \in \mathbb{N}$, then the Cartesian product $G = \prod_{i\in\mathbb{N}} H_i$ is a profinite group (see Section 2.3), $FC(G)$ consists of elements with finite support and any element with infinite support is in the closure of $FC(G)$. To see this, let $x = (x_i)_{i\in\mathbb{N}}$ be an element where infinitely many x_i are nontrivial and let $y_n = (y_{ni})_{i\in\mathbb{N}}$, where $y_{ni} = x_i$ for $i \leq n$ and 1 for $i > n$. If O is an open set containing x, it can be proved that some open set of the form xU is contained in O, where U is an open normal subgroup of G of the form $\prod_{i>m} H_i$ with $m \geq n$. Notice that xU consists entirely of elements whose m first entries coincide with the first m entries of x and the i-th entry, with $i > m$, is arbitrary. Then xU contains y_n for every $n \geq m$ and we conclude that the closure of $FC(G)$ is the whole G.

Definition 2.2.12. If G is a topological group, G_0 denotes the connected component of the identity subgroup, i.e., the largest connected subgroup of G.

The Euclidean space $\mathbb{R}^n$ with topology induced by the Euclidean norm is connected and a profinite group is always totally disconnected, i.e., the connected subsets of it are only the singletons. Let P be a profinite goup, $Q = \mathbb{R}$ and consider $G = P \times Q$. If C is a connected subgroup of G, its projection on P is also connected, so it must be trivial and $C \leq 1 \times Q$. Since $1 \times Q$ itself is connected, it must coincide with G_0.

The class of compact groups such that the connected component of identity is central admits an useful decomposition, as stated in [26, Lemma 3.5].

Proposition 2.2.13. *Let G be a compact group such that G_0 is central. Then there is a profinite normal subgroup Δ such that $G = G_0\Delta$*

We are going to need the following technical result from [5].

Lemma 2.2.14. *Let G be a compact group and n a positive integer. The set $X = \{x \in G \,|\, |x^G| \leq n\}$ is a closed subset of G.*

Proof. It is sufficient to show that if $a \in G \setminus X$, then a is contained in an open subset U which has empty intersection with X. Since $a \notin X$, we can choose $n+1$ elements $x_1, ..., x_{n+1}$ in such a way that the conjugates a^{x_i} are distinct for $i = 1, \ldots, n+1$. Set

$$U = \{u \in G;\ [u, x_i x_j^{-1}] \neq 1 \text{ for } 1 \leq i \neq j \leq n+1\}.$$

We observe that $a \in U$ and every element in U has at least $n+1$ conjugates. Indeed, suppose that $u \in U$. Then $[u, x_i x_j^{-1}] \neq 1$ for every choice of $i \neq j$. Standard commutator identities show that

$$[u, x_i x_j^{-1}] = [u, x_j^{-1}][u, x_i]^{x_j^{-1}}.$$

Thus $[u, x_j^{-1}] \neq [x_i, u]^{x_j^{-1}}$, which implies that $[u^{x_j}, x_j^{-1}] \neq [x_i, u]$. Now, since $[u^{x_j}, x_j^{-1}] = [x_j, u]$, we conclude that $[x_j, u] \neq [x_i, u]$ for every choice of different i and j. This implies that u has at least $n+1$ conjugates and $U \cap X = \emptyset$. Further, since the commutator map is continuous, U is open. The proof is complete. □

Remark 2.2.15. If S is any subset of G, we denote by x^S the set of S-conjugates of x in G. The previous argument also proves that the set $\{x \in G \,|\, |x^S| \leq n\}$ is closed in G for any (not necessarily closed) subset S of G.

2.3 Profinite Groups

Profinite groups are a special class of compact groups, as we define now. A topological space is *totally disconnected* if its only connected subsets are singletons.

Definition 2.3.1. Let G be a topological group. If G is compact, Hausdorff and totally disconnected, then G is called *profinite*.

There is an equivalent way of defining a profinite group, in terms of inverse limits. If $\mathcal{N} = \{N \,|\, N \text{ open and normal in } G\}$ is the set of open normal subgroups of G, then

$$\mathcal{C} = \{G/N \,|\, N \in \mathcal{N}\}$$

is an inverse system of finite groups. The group G can be defined as the inverse limit of the inverse system defined above: $G = \varprojlim_{N \in \mathcal{N}} G/N$. For more details see [43, Section 1.1].

If the members of the inverse system are p-groups (resp. soluble), their inverse limit is called a pro-p-group (resp. prosoluble). Moreover, as profinite groups are generalizations of finite groups, many theorems about finite groups have profinite variations. For example, if $\mathcal{C}$ is an inverse system of finite groups and p is a prime, the Sylow p-subgroups of the groups in $\mathcal{C}$ form an inverse system, and its inverse limit is a Sylow pro-p subgroup of G. Sylow pro-p subgroups exist and are conjugate in G [43, Theorem 2.3.5].

The order of a profinite group is either finite or uncountable [43, Proposition 2.3.1 (b)], but it is possible to define the set of primes dividing the order of a profinite group even in the infinite case. If $\{G_i \mid i \in I\}$ is the inverse system giving rise to G, define $\pi(G)$ as

$$\pi(G) = \{p \text{ a prime} \mid p \text{ divides } |G_i| \text{ for some } i\}.$$

If A is a group of continuous automorphisms of G, it can be proved that the topology of G induces a topology on A under which A becomes a topological group [43, Section 4.4]. A continuous automorphism ϕ of G is said to be *coprime* if it has finite order coprime to every number in $\pi(G)$. The following lemma is an extension to profinite groups of a well-known fact on coprime automorphisms of finite groups (see for instance Lemma 2.1.3).

Lemma 2.3.2. *Let G be a profinite group admitting a coprime automorphism ϕ and let N be a ϕ-invariant normal subgroup of G. Then $C_{G/N}(\phi) = C_G(\phi)N/N$.*

2.4 The Length of a Finite Group

In this section we give some definitions needed in the text. For a prime number p, a finite group G is called a p-group if its order is a power of p, and is called a p'-group if p does not divide the order of G. We say that G is *p-soluble* if G has a normal series where all factors are p- or p'-groups. The *p-length* of G, denoted by $l_p(G)$, is the smallest number of p-factors in such a series.

Every finite group has a normal series where all factors are soluble or a direct product of nonabelian finite simple groups. To construct such a series, one might consider the soluble radical S of G, i.e., the product of all soluble normal subgroups of G, and pass to G/S. Consider the product N/S of all minimal normal subgroups of G/S and see that N/S is a direct product of nonabelian finite simple groups, then pass to G/N and repeat the two steps until we get back to G. The smallest number of non-soluble factors in such a series is the *nonsoluble length* of G and is denoted by $\lambda(G)$. Similarly, if p is a prime, every finite group has a normal series where all factors are either p-soluble or a direct product of finite simple groups of orders divisible by p. The *non-p-soluble length* of G, denoted $\lambda_p(G)$, is defined as the smallest number of non-p-soluble factors in such a series.

Of course $\lambda_2(G) = \lambda(G)$ as all odd order groups are soluble, by the Feit-Thompson Theorem [15]. Hall and Higman bounded the p-length of a p-soluble

group G in terms of some identities satisfied by the Sylow p-subgroups of G in [23], reducing the Restricted Burnside Problem to the class of p-groups. Nonsoluble and non-p-soluble lengths were defined in [33] by Shumyatsky and Khukhro, and the authors also bound both lengths in terms of certain identities satisfied by Sylow subgroups of G. Recently, Fumagalli, Leinen and Puglisi [16] bounded $\lambda(G)$ in terms of an arbitrary identity satisfied by the 2-Sylow subgroup of a finite group, confirming a conjecture posed in [33].

If G is a profinite group, analogues of the previous notions can be defined. A profinite group G has finite p-length, also denoted by $l_p(G)$, if it has a finite length series of closed normal subgroups where all factors are pro-p or pro-p' groups. The minimal number of pro-p factors is $l_p(G)$. Say that G has finite nonprosoluble length if it has a finite length series of closed normal subgroups where all factors are either prosoluble or a Cartesian product of finite simple groups. The smallest number of nonprosoluble factors in this series is $\lambda(G)$. Replacing "prosoluble" by "pro-p-soluble" and "simple" by "simple of order divisible by p" we get the notion of non-pro-p-soluble length, also denoted by $\lambda_p(G)$. In a similar way to what Hall and Higman did in [23], Wilson proved in [50] that a profinite torsion group has finite nonprosoluble length and a prosoluble torsion group G has finite p-length for every $p \in \pi(G)$. This has enabled Zelmanov [54] to prove that compact torsion groups are locally finite, that is, any finite subset in such a group generates a finite subgroup.

2.5 Measure theory

In this section basic results related to measure theory in general measure spaces are established. The exposition follows mainly [3].

Definition 2.5.1. Let X be a set. A collection $\mathcal{A}$ of subsets of X is called a σ-algebra if the following hold.

(i) $X \in \mathcal{A}$,

(ii) If $A \in \mathcal{A}$, then the complement $A^c \in \mathcal{A}$, and

(iii) If $\{A_i \mid i \in \mathbb{N}\}$ is a countable collection of sets in $\mathcal{A}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

In this case, we say that the ordered pair $(X, \mathcal{A})$ is a *measurable space*. If $M \in \mathcal{A}$, we say that M is a *measurable subset of* X. Moreover, see that a σ-algebra $\mathcal{A}$ is closed under finite unions and intersections. For finite unions let

$A_1, \ldots, A_n \in \mathcal{A}$, let $A_s = \emptyset$ for $s \geq n+1$ and use property (iii), and for finite intersections use (ii) and de Morgan laws.

Example 2.5.2.

- Let X be a set and denote by $\mathcal{P}(X)$ the set of all subsets of X. Then $\mathcal{P}(X)$ is a σ-algebra on X.
- Let $X = \{a, b, c\}$. Both $\mathcal{A}_1 = \{\emptyset, X\}$ and $\mathcal{A}_2 = \{\emptyset, \{a\}, \{b, c\}, X\}$ are σ-algebras on X.

If X is a set and $\{\mathcal{A}_i \,|\, i \in I\}$ is a collection of σ-algebras on X, then the intersection $\bigcap_{i\in I} \mathcal{A}_i$ is also a σ-algebra on X. This idea can be used to construct the smallest σ-algebra on X that contains a particular collection of subsets.

Definition 2.5.3. Let X be a set and $\mathcal{C}$ be a collection of subsets of X. The smallest σ-algebra that contains $\mathcal{C}$ is called the σ-algebra *generated* by $\mathcal{C}$. We denote it by $\sigma(\mathcal{C})$.

Definition 2.5.4. Let X be a Hausdorff topological space. The *Borel σ-algebra* on X is the one generated by the open subsets of X and is denoted by $\mathcal{B}(X)$. A subset belonging to $\mathcal{B}(X)$ is called Borel-measurable.

We observe that different collections of subsets of X can generate the same σ-algebra. In particular, $\mathcal{B}(X)$ can also be generated by the closed subsets of X. Furthermore, subsets that are not open nor closed are contained in $\mathcal{B}(X)$, as the next example shows.

Example 2.5.5.

- Half-open intervals belong to $\mathcal{B}(\mathbb{R})$:

$$(a, b] = \bigcup_{r=1}^{\infty} \left(a, b - \frac{1}{r}\right).$$

- If G is a compact topological group and k is a positive integer, we know that $G_k = \{x \in G \,|\, |x^G| \leq k\}$ is a closed subset of G for every k, by Lemma 2.2.14. The union $FC(G) = \bigcup_{k=1}^{\infty} G_k$ belongs to $\mathcal{B}(G)$ and is not necessarily closed, as we remarked before.

Definition 2.5.6. Let $(X, \mathcal{A})$ be a measurable space. A function $\mu : \mathcal{A} \to [0, \infty]$ is called a *measure* on $\mathcal{A}$ if the following hold.

(i) $\mu(\emptyset) = 0$, and

(ii) The function μ is *countably additive*: if $A_1, A_2, \ldots$ is a countable collection of pairwise disjoint measurable subsets of X, then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

If $(X, \mathcal{A})$ is a measurable space and μ is a measure on $\mathcal{A}$, we call the ordered triple $(X, \mathcal{A}, \mu)$ a *measure space*. If $\mu(X) = 1$, we call μ a *probability measure* and $(X, \mathcal{A}, \mu)$ is a *probability space*. If X is a topological space, then a measure defined on $(X, \mathcal{B}(X))$ is called a *Borel measure* on X. We will often say that a measure μ is defined on X, rather than on $(X, \mathcal{B}(X))$.

Remark 2.5.7.

- Property (ii) also holds for finite unions of pairwise disjoint subsets, in a similar fashion to the remark made after Definition 2.5.1.
- Let μ be a measure on X and A, B be measurable subsets, with $A \subseteq B$. Then $\mu(A) \leq \mu(B)$. Indeed, since $B = A \cup (A^c \cap B)$ and $A^c \cap B$ is measurable, it follows that $\mu(A) \leq \mu(A) + \mu(A^c \cap B) = \mu(B)$. This property is called *monotonicity*.

Example 2.5.8 (The Lebesgue measure on $\mathbb{R}$)**.** Consider the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and let $a, b \in \mathbb{R}$. It is possible to define a measure λ on $\mathbb{R}$ such that, for every half-closed interval $(a, b]$, the measure $\lambda((a, b])$ equals the length $b - a$. In particular, $\lambda((-\infty, b])$ and $\lambda((a, \infty))$ are both ∞. This is the so-called *Lebesgue measure* on $\mathbb{R}$, and it generalizes the notion of length to subsets of $\mathbb{R}$ which are not intervals. For example, both the Cantor ternary set and the rational numbers $\mathbb{Q}$ are Borel-measurable and have zero measure under λ.

Example 2.5.9.

- Let X be a countable set and $\mathcal{A} = \mathcal{P}(X)$. If μ is the function that associates to each subset A of X its cardinality, then μ is a measure. It is called the counting measure.

- Consider $(X, \mathcal{P}(X), \mu)$ where X is finite and $\mu(A) = |A|/|X|$. Then μ is called the discrete uniform probability on X. See that $\mu(X) = 1$.

We are also going to need the following. It is [3, Theorem 1.2.7(b)].

Lemma 2.5.10. *Let $(X, \mathcal{A}, \mu)$ be a probability space. If $A_1, A_2, \ldots$ is a countable family of measurable subsets of X such that $A_n \supseteq A_{n+1}$ for every n, then*

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

One of the applications of measures is the possibility to develop a stronger notion of integral than the usual Riemann integral. We begin with some more definitions.

Definition 2.5.11. Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. A function $f : X \to Y$ is *measurable with respect to $\mathcal{A}$ and $\mathcal{B}$* if $f^{-1}(B) \in \mathcal{A}$ for every $B \in \mathcal{B}$. If X and Y are Hausdorff topological spaces, a function $f : X \to Y$ is said to be *Borel-measurable* if f is measurable with respect to $\mathcal{B}(X)$ and $\mathcal{B}(Y)$.

Example 2.5.12.

- If X and Y are topological spaces, then every continuous map $f : X \to Y$ is Borel measurable.
- Let $(X, \mathcal{A})$ be a measurable space and let A be any subset of X. Define the *indicator of A* as the map $\chi_A : X \to \mathbb{R}$, assuming the value 1 in A and 0 in A^c. Then χ_A is Borel-measurable if and only if A is Borel-measurable.
- If $f, g : X \to \mathbb{R}$ are measurable, then $f + g$, $f - g$, $f \cdot g$ and f/g (provided $g \neq 0$) are also measurable.

Definition 2.5.13. Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f : X \to \mathbb{R}$ be a Borel-measurable function. We say that a f is *simple* if it takes only finitely many values. In particular, there exist measurable subsets $A_1, \ldots, A_n$ of X and real numbers $c_1, \ldots, c_n$ such that $f = \sum_{i=1}^{n} c_i \chi_{A_i}$, where χ_{A_i} is the characteristic function of A_i.

Now let $(X, \mathcal{A}, \mu)$ be a measure space and f be a Borel-measurable simple function. If $f = \sum_{i=1}^{n} c_i \chi_{A_i}$, we define the integral of f with respect to μ as

$$\int_X f(x) d\mu(x) = \sum_{i=1}^{n} c_i \mu(A_i).$$

A general non-negative function can be approximated by simple functions, as the next result states. It is [3, Theorem 1.5.5 (a)].

Proposition 2.5.14. *Let $(X, \mathcal{A})$ be a measurable space and $f : X \to \mathbb{R}$ be a non-negative Borel-measurable function. Then there exists an increasing sequence of Borel-measurable simple functions (f_i) converging pointwise from below to f.*

If $(X, \mathcal{A}, \mu)$ is a measure space and f is a non-negative Borel-measurable function, we define the integral of f with respect to μ as

$$\int_X f(x)d\mu(x) = \sup\left\{\int_X \phi(x)d\mu(x) \,|\, 0 \leq \phi \leq f \text{ and } \phi \text{ is simple}\right\}.$$

Moreover, if f is an arbitrary Borel-measurable function, the positive part of f is $f^+ = max\{f, 0\}$ and the negative part of f is $f^- = max\{-f, 0\}$, both non-negative. Then $f = f^+ - f^-$ and the integral $\int_X f(x)d\mu(x)$ is

$$\int_X f(x)d\mu(x) = \int_X f^+(x)d\mu(x) - \int_X f^-(x)d\mu(x),$$

provided this difference is well defined. We say that f is *μ-integrable* if the integral $\int_X f(X)d\mu(x)$ is finite.

Remark 2.5.15.

- If μ is a measure on X and A is a measurable subset of X, the measure of A can be calculated as the integral

$$\int_X \chi_A(x)d\mu(x) = 1 \cdot \mu(A) + 0 \cdot \mu(A^c) = \mu(A).$$

- Consider the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, where λ is the Lebesgue measure on $\mathbb{R}$. The indicator of rational numbers, $\chi_{\mathbb{Q}}$, is an example of function which is not Riemann-integral, as $\chi_{\mathbb{Q}}$ is discontinuous in every point. However, since $\mathbb{Q}$ is Borel-measurable, this function is Lebesgue-integrable and

$$\int_{\mathbb{R}} \chi_{\mathbb{Q}}(x)d\lambda(x) = \lambda(\mathbb{Q}) = 0.$$

If A is a measurable subset of X and $f : X \to \mathbb{R}$ is a measurable function,

define the integral of f over A as

$$\int_A f(x)d\mu(x) = \int_X f(x) \cdot \chi_A(x)d\mu(x).$$

See that the integrand is measurable because A is measurable. The Lebesgue integral has the following properties.

Proposition 2.5.16. *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $f : X \to \mathbb{R}$ be a measurable function. Then*

(i) If $f(x) \leq g(x)$ for every $x \in X$, then $\int_X f(x)d\mu(x) \leq \int_X g(x)d\mu(x)$.

(ii) If $A \subseteq B$ are measurable sets and f is non-negative, then

$$\int_A f(x)d\mu(x) \leq \int_B f(x)d\mu(x).$$

Proof. For (i), if f, g are non-negative it suffices to see that every simple function ϕ such that $\phi \leq f$ also satisfies $\phi \leq g$. In general, $f \leq g$ implies that $f^+ \leq g^+$ and $f^- \geq g^-$ so the result follows from the definition of integral. For (ii), see that $f(x)\chi_A(x) \leq f(x)\chi_B(x)$ and apply (i). □

We consider one last construction related to Lebesgue integral. Let $(X, \mathcal{A}_X, \mu_X)$ and $(Y, \mathcal{A}_Y, \mu_Y)$ be measure spaces. We can form a measure space $(X \times Y, \mathcal{A}, \mu)$, called *product space*, such that the σ-algebra $\mathcal{A}$ contains all sets of the form $A \times B$ and $\mu(A \times B) = \mu_X(A)\mu_Y(B)$ whenever $A \in \mathcal{A}_X$ and $B \in \mathcal{A}_Y$. The details can be verified in [3, Section 2.6].

Now let $f : X \times Y \to \mathbb{R}$ be a Borel-measurable function. For each $x \in X$, $y \in Y$ and $C \in \mathcal{A}$, define $C_x = \{y \in Y \,|\, (x, y) \in C\}$, the "slice" of Y having x as first coordinate, and analogously $C_y = \{x \in X \,|\, (x, y) \in C\}$. It is possible to prove that C_x is $\mathcal{A}_Y$-measurable and C_y is $\mathcal{A}_X$-measurable for every $x \in X$ and $y \in Y$. Moreover, the map $x \mapsto \mu_Y(C_x)$ is $\mathcal{A}_X$-measurable, and the map $y \mapsto \mu_X(C_y)$ is $\mathcal{A}_Y$-measurable. Then we have

Theorem 2.5.17. *With the notation introduced before, let $C \in \mathcal{A}$. Then*

$$\mu(C) = \int_{X \times Y} \chi_C(x, y)d\mu_X(x)d\mu_Y(y) = \int_X \mu_Y(C_x)d\mu_X(x) = \int_Y \mu_X(C_y)d\mu_Y(y).$$

This is a weaker version of the classical Fubini Theorem [3, Theorem 2.6.6]. See that the integrands are indeed measurable functions on the appropriate measure spaces.

2.6 The Haar measure

In this section we state existence and unicity of a Haar measure on locally compact groups, together with some useful results. We also stress out that a proof of both existence and unicity of Haar measure in locally compact groups is presented in the Appendix for the benefit of the reader who might not be acquainted with the subject. We avoided the language and machinery of functional analysis everywhere where it was possible in the Appendix to make it rely only on group theory and topology.

Definition 2.6.1. Let G be a Hausdorff locally compact topological group and $\mathcal{B}(G)$ be its Borel σ-algebra. A *left Haar measure* μ on $(G, \mathcal{B}(G))$ is a nonzero measure with the following properties:

(H1) If $x \in G$ and $S \in \mathcal{B}(G)$, then $\mu(xS) = \mu(S)$.

(H2) The measure μ is finite on every compact subset $K \subseteq G$.

(H3) The measure μ is outer regular on Borel sets: if $S \subseteq G$ is measurable, then

$$\mu(S) = \inf\{\mu(U) \,|\, S \subseteq U,\ U \text{ is open}\}.$$

(H4) The measure μ is inner regular on open sets: if $U \subseteq G$ is open in G, then

$$\mu(U) = \sup\{\mu(K) \,|\, K \subseteq U,\ K \text{ is compact}\}.$$

If we replace xS by Sx on (H1), we get a *right Haar measure*. The existence of a measure satisfying (H1)-(H4) was first proved for locally compact groups by A. Haar [22] in 1933, although his proof only worked for separable groups. The existence was proved in full generality first by A. Weil [48] and later by H. Cartan [7] and J. von Neumann [46], and unicity, up to a scaling factor, was proved by several mathematicians including J. von Neumann [47] and S. Kakutani [28]. We can resume their results in the following theorem.

Theorem 2.6.2. *Let G be a Hausdorff locally compact topological group. Then there exists a nonzero left Haar measure μ on $(G, \mathcal{B}(G))$. If ν is another nonzero left Haar measure on $(G, \mathcal{B}(G))$, there is a positive real number k such that $\mu(S) = k\nu(S)$ for every Borel set S.*

The Haar measure on the real line consists on Lebesgue measure and its scalar multiples. It is translation-invariant with respect to measure of sets and, more generally, with respect to integration: if f is any Lebesgue-integrable function, then

$$\int_{\mathbb{R}} f(x+a)dx = \int_{\mathbb{R}} f(x)dx,$$

for every $a \in \mathbb{R}$, provided this integral is well-defined.

Property (H2) in Definition 2.6.1 ensures that, if G is a compact group with Haar measure μ, then $\mu(G)$ is finite. It turns out that the converse is also true.

Proposition 2.6.3. *Let G be a locally compact group and μ be a left Haar measure on G. Then $\mu(G)$ is finite if and only if G is compact.*

Proof. Suppose that G is compact. As we remarked before, property (H2) ensures that every compact subset of G has finite measure, so $\mu(G)$ must be finite. Assume now that $\mu(G) < \infty$ and let V be a compact neighbourhood of the identity of positive measure. Such a compact set must exist, combining (H3) and (H4) in Definition 2.6.1 and the fact that μ is nonzero. Let X be the collection of all finite subsets $\{x_1, \ldots, x_n\}$ of G such that $x_iV \cap x_jV = \emptyset$. Then

$$n\mu(V) = \mu\left(x_1V \cup \cdots \cup x_nV\right) \leq \mu(G).$$

and it follows that $n \leq \frac{\mu(G)}{\mu(V)}$. Therefore, we can choose a maximal element $\{y_1, \ldots, y_n\}$ in X, in the sense of inclusion. For each $y \in G$, it follows that $yV \cap y_iV \neq \emptyset$ for some i, thus $y \in y_iVV^{-1}$. Continuity of the group operation implies that VV^{-1} is compact, and as $G = \bigcup_{i=1}^{n} y_iVV^{-1}$, we conclude that G is compact. □

We are concerned with probability measures, so $\mu(G)$ will always be 1 in the text. This is the reason why we consider compact rather than locally compact groups. Moreover, we say that a left Haar measure μ on a compact group G is *normalized* provided $\mu(G) = 1$.

Let G be a compact Hausdorff group and $N \leq G$ be a closed normal subgroup of G. The topology of G naturally induces one on G/N by declaring as open on

G/N the images of open subsets of G, and analogously the σ-algebra of Borel subsets of G/N will coincide with the σ-algebra induced by $\mathcal{B}(G)$. The quotient G/N is again compact and Hausdorff, since N is closed. Finally, if $\pi : G \to G/N$ is the natural projection and S is a measurable subset of G/N, we can define a Haar measure $\bar{\mu}$ on G/N by

$$\bar{\mu}(S) = \mu(\pi^{-1}(S)).$$

It follows directly that $\mu(S) \leq \bar{\mu}(\pi(S))$.

The following three Lemmas are going to be useful in the sequence. In all of them we consider a compact group G equipped with the normalized measure μ.

Lemma 2.6.4. *Let G be a compact group and let K be a subgroup of G. Then either $\mu(K) = 0$ or $\mu(K) > 0$ and K is open on G. Furthermore, in the latter case, $\mu(K) = [G : K]^{-1}$.*

Proof. Assume $\mu(K) > 0$, let n be the smallest integer such that $\mu(K) \geq \frac{1}{n}$ and assume that $[G : K]$ is infinite. Then there exist $x_1, \ldots, x_{n+1}$ such that the cosets $x_i K$ are all different, $i = 1, \ldots, n+1$. It follows that

$$\mu(G) \geq \mu\left(\bigcup_{i=1}^{n+1} x_i K\right) = (n+1)\frac{1}{n} > 1,$$

a contradiction. Thus K has finite index on G and is open. Moreover, we have $\mu(K) = [G : K]^{-1}$ by property (H1). □

It follows that $\mu(H) = [G : H]^{-1}$ even if H has infinite index, interpreting $''\frac{1}{\infty}''$ as zero. If G is a compact group, then compactness and Lemma 2.6.4 can be used to argue that left and right Haar measures coincide. Next we use the Haar measure to estimate the index of a subgroup.

Lemma 2.6.5. *Let G be a compact group, and let K and H be subgroups of G with $K \leq H$. Assume further that $\mu(K) \geq \epsilon\mu(H) > 0$ for some positive ϵ. Then $[H : K] \leq \epsilon^{-1}$.*

Proof. Since $\mu(K), \mu(H) > 0$, the previous lemma implies that both subgroups are of finite index and $\mu(K) = [G : K]^{-1}$ and $\mu(H) = [G : H]^{-1}$. Hence the result. □

We say that a subset X of a group G is *symmetric* provided $X = X^{-1}$. Moreover, for a positive integer s, let X^s be the set of products of at most s elements from X. The next lemma is essentially Lemma 2.1 in [11].

Lemma 2.6.6. *Let G be a compact group with normalized Haar measure μ and let $r \geq 1$. Suppose that X is a closed symmetric subset of G containing the identity. If $\mu(X) > \frac{1}{r+1}$, then $\langle X \rangle = X^{3r}$.*

Proof. Suppose $x_i \in X^{3i+1} \setminus X^{3i}$ for $i = 0, \ldots, r$. Then for each i, as long as $X^{3i+1} \setminus X^{3i}$ is nonempty, we have

$$x_i X \subseteq X^{3i+2} \setminus X^{3i-1}.$$

So, assume that the sets $X^{3i+1} \setminus X^{3i}$ are nonempty for $i = 0, \ldots, r$. Then $x_0 X, \ldots, x_r X$ are disjoint subsets of G, each of measure $\mu(X)$, and

$$\mu\left(\bigcup_{i=0}^{r} x_i X\right) = (r+1)\mu(X) > 1.$$

Therefore $X^{3i+1} = X^{3i}$ for some $i \leq r$. In particular, $X^{3r} = \langle X \rangle$. □

Example 2.6.7. Let $G = GL_n(\mathbb{R})$ and consider $(G, \mathcal{B}(G), \mu)$ where μ is the normalized Haar measure on G. If $f : G \to \mathbb{R}$ is a continuous function, there is a relation between Riemann and Lebesgue integrals of f:

$$\int_G f(x) d\mu(x) = \int_G f(x) \cdot \frac{1}{|det(x)|^n} dx.$$

As another example of a well-behaved integral consider $G = (0, \infty)$, the multiplicative group of positive real numbers. Then, for every continuous function $f : G \to \mathbb{R}$ we have

$$\int_G f(x) d\mu(x) = \int_G f(x) \cdot \frac{1}{x} dx.$$

For example, the measure of (a, b) is $log(b/a)$.

2.6.1 Why Haar measure?

There are some reasons why normalized Haar measure is a suitable idea to define probability in a compact group. If G is an infinite compact group, the probability of randomly choosing a single element must be zero, but if H is an open subgroup of G it makes sense that the probability of choosing an element from H randomly is $[G : H]^{-1}$, as stated in Lemma 2.6.4. Moreover, if G is a profinite group with normalized Haar measure μ, then μ induces the uniform

discrete probability on the finite quotients of G: if N is a normal open subgroup of G, the induced measure μ_N on G/N satisfies $\mu_N(N/N) = \mu(N) = [G : N]^{-1}$.

On the other hand, let $\mathcal{C} = \{G_i \mid i \in I\}$ be the inverse system of finite groups whose inverse limit is G and let μ_i be the discrete measure on G_i. It is possible to define a measure on G which is an inverse limit for the discrete probability measures of each G_i, and this inverse limit coincides with the normalized Haar measure on G [6, Chapter 7, Paragraph 6]. As another illustration of the connection of such measures, we have the following result. We provide a proof of it in the special case where G admits a countable family of subgroups with trivial intersection.

Proposition 2.6.8. *Let G, $\mathcal{C}$, μ and μ_N be as defined before. If S is any measurable subset of G, then $\mu(S) \leq \inf_N \mu_N(SN/N)$, with equality when S is closed.*

Proof. The definition of μ_N ensures that $\mu(S) \leq \mu_N(SN/N)$, thus it follows that $\mu(S) \leq \inf_N \mu_N(SN/N)$. Now assume S is closed on G and let $\mathcal{M}$ be any collection of closed subsets of G such that, for every $S_1, S_2 \in \mathcal{M}$ there is $S_3 \in \mathcal{M}$ with $S_3 \subseteq S_1 \cap S_2$. Proposition 2.1.4(a) in [43] states that

$$\bigcap_{M \in \mathcal{M}} MC = \left(\bigcap_{M \in \mathcal{M}} M \right) C$$

for every closed subset C of G. In particular, if $\{N_i \mid i = 1, 2, \dots\}$ is a countable family of nested neighbourhoods of identity with trivial intersection, we conclude that $\bigcap_{i=1}^{\infty} SN_i = S$ and, by Lemma 2.5.10, $\mu(S) = \lim_{i \to \infty} \mu(SN_i)$. Denoting by π_i the projection of G onto G/N_i, then

$$\mu(SN_i) = \mu(\pi_i^{-1}(SN_i/N_i)) = \mu_{N_i}(SN_i/N_i),$$

so we conclude that $\mu(S) = \inf_i \mu_{N_i}(SN_i/N_i)$. The measure $\mu(S)$ is a lower bound for $\{\mu_N(SN/N) \mid N \in \mathcal{C}\}$, by the first part of the proof. However, for any $\epsilon > 0$, there exists an i such that $\mu(S) \leq \mu_{N_i}(SN_i/N_i) < \mu(S) + \epsilon$, so it follows that $\mu(S) = \inf_N \mu_N(SN/N)$, as we wanted. □

This inequality can be strict. Let H be a countable residually finite group, let G be its profinite completion and μ the normalized Haar measure of G. As singletons are measurable subsets of G, then H can be regarded as a measurable subgroup of G. Moreover, the measure of a singleton is zero unless G is finite, so $\mu(H) = 0$ whereas $\mu(G) = \inf_N \mu_N(G/N) = \mu_N(HN/N) = 1$, since H projects surjectively on all finite images of G.

Chapter 3

The commuting probability in groups

3.1 Definitions and Historical Remarks

The definition of commuting probability in groups dates back to Erdös and Turán [13]. If G is a finite group, we can define the probability that two elements of G commute as

$$Pr(G) = \frac{|\{(x,y) \in G \times G \,|\, xy = yx\}|}{|G|^2},$$

which we call the *commuting probability* of G. Fixing $x \in G$, the number of ordered pairs (x,y) where y commutes with x is $|C_G(x)|$, so we can rewrite the above probability as

$$Pr(G) = \frac{\sum_{x \in G} |C_G(x)|}{|G|^2} = \frac{1}{|G|} \sum_{x \in G} \frac{1}{[G : C_G(x)]}.$$

Now, $[G : C_G(x)] = [G : C_G(y)]$ for every $y \in x^G$. If G has $k(G)$ conjugacy classes and $x_1, \ldots, x_{k(G)}$ are conjugacy class representatives of G, the summation can be written as

$$\frac{1}{|G|} \sum_{x \in G} \frac{1}{[G : C_G(x)]} = \frac{1}{|G|} \sum_{i=1}^{k(G)} |x_i^G| \frac{1}{[G : C_G(x_i)]} = \frac{k(G)}{|G|}.$$

This direct method can be used to compute commuting probabilty for finite groups and to prove, for example, that $Pr(G \times H) = Pr(G) \times Pr(H)$. We also conclude that nonabelian groups can have arbitrarily small commuting probabili-

ties: if G is a finite group such that $Pr(G) < 1$, then $Pr(\underbrace{G \times \cdots \times G}_{m\,\text{times}}) = Pr(G)^m$ is arbitrarily small for large m.

Following Gustafson [21], it is also possible to define commuting probability in infinite compact groups. Let G be a compact Hausdorff group and consider the map $f : G \times G \to G$ taking (x, y) to the commutator $[x, y]$. Then f is continuous and, since 1 is a closed subgroup of G, the preimage $f^{-1}(1)$ is a closed subset of $G \times G$. This preimage is the set

$$C = \{(x, y) \in G \times G \,|\, xy = yx\},$$

which is closed in $G \times G$, thus measurable. If μ is the normalized Haar measure on G, we equip $G \times G$ with the product measure $\mu \times \mu$ and define the commuting probability of G as the value

$$(\mu \times \mu)(C) = \int\limits_{G \times G} \chi_C(x, y) d(\mu(x) \times \mu(y)),$$

where χ_C is the characteristic function of C. As it was defined in the end of section 2.5, the set C_x is

$$C_x = \{y \in G \,|\, (x, y) \in C\} = \{y \in G \,|\, xy = yx\} = C_G(x).$$

Now, using Theorem 2.5.17, we can rewrite the above integral as

$$\int\limits_{G \times G} \chi_C(x, y) d(\mu(x) \times \mu(y)) = \int\limits_G \mu(C_x) d\mu(x) = \int\limits_G \mu(C_G(x)) d\mu(x).$$

Recall that if $C_G(x)$ has infinite index on G, we interpret $[G : C_G(x)]^{-1}$ as zero. Then, by Lemma 2.6.4, we have

$$Pr(G) = \int\limits_G \mu(C_G(x)) d\mu(x) = \int\limits_G \frac{1}{[G : C_G(x)]} d\mu(x). \qquad (**)$$

Commuting probability has been used, among other applications, to derive structural results about groups. The intuition behind some problems is: if $Pr(G)$ is sufficiently high, is it true that G will be "close" to being abelian in some sense? For example, Gustafson proved in [21] that, if $Pr(G) \approx 1$, then $Pr(G) = 1$ already: every nonabeilan compact group G satisfies $Pr(G) \leq \frac{5}{8}$. Representation theory tools have been used by Lescott [34] to prove that, if G is

finite and $Pr(G) > \frac{1}{2}$, then G is nilpotent. Not relying on the Classification of Finite Simple Groups, Lescott [35] and Dixon (see Can. Math. Bul., 16 (1973) p. 302) independently proved that if $Pr(G) > \frac{1}{12}$, then G is soluble. Both bounds are sharp, since the symmetric group S_3 of degree 3 satisfies $Pr(S_3) = \frac{1}{2}$ and the alternating group of degree 5, A_5, satisfies $Pr(A_5) = \frac{1}{12}$.

We say that G is a BFC-group when there is a positive integer n such that for every $x \in G$ we have $|x^G| \leq n$. This is equivalent to saying that $\frac{1}{[G:C_G(x)]} \geq \frac{1}{n}$. In particular, if G is a compact group such that $|x^G| \leq n$ for all $x \in G$, then $Pr(G) \geq \frac{1}{n}$, by equation $(**)$. A well-known theorem due to B. H. Neumann states that in a BFC-group the derived group G' is finite [39]. In particular, if $|x^G| \leq n$ for all $x \in G$, the order of G' is n-bounded. J. Wiegold [49] found the first explicit bound for the order of G', and the best known bound was found in [19]. To prove some of our results we are going to use tools developed in recent generalizations of B. H. Neumann's result [2, 8, 10].

P.M. Neumann [40] proved one of the first structure theorems on commuting probability for finite groups, which in particular follows from Theorem A. The theorem of P.M. Neumann is as follows.

Theorem 3.1.1. *Let ϵ be positive and G be a finite group such that $Pr(G) \geq \epsilon$. Then G has a subgroup H such that $[G : H]$ and $|H'|$ are ϵ-bounded.*

Considering a similar problem on relative commuting probability, Detomi and Shumyatsky [9] proved the following theorem.

Theorem 3.1.2. *Let $\epsilon > 0$ and let G be a finite group having a subgroup K such that $Pr(K, G) \geq \epsilon$. Then there is a normal subgroup $T \leq G$ and a subgroup $B \leq K$ such that the indices $[G : T]$ and $[K : B]$ and the order of $[T, B]$ are ϵ-bounded.*

In particular, if one considers $K = G$, what we get is Theorem 3.1.1. Our first main result, Theorem A, is an extension of Theorem 3.1.2 to the class of compact groups. In particular, our proof follows closely ideas from [9]. All the corollaries present in [9] can also be extended suitably to compact groups using Theorem A, which we restate here.

Theorem A. *Let G be a compact group having a closed subgroup K such that $Pr(K, G) \geq \epsilon$ for $\epsilon > 0$. Then there is a normal closed subgroup $T \leq G$ and a closed subgroup $B \leq K$ such that the indices $[G : T]$ and $[K : B]$ and the order of $[T, B]$ are ϵ-bounded.*

Theorem A will be proved in Section 3.3. We finish this introduction shedding light in the intuition behind some theorems we prove. Suppose that G is a compact group such that $Pr(G) \geq \epsilon$. Theorem A implies that G has a subgroup H of ϵ-bounded index such that H' is finite of ϵ-bounded order, so we conclude that G has a large abelian section, namely H/H'. We can also prove that G is virtually abelian. Indeed, since H' is finite, there exists an open subgroup L of H avoiding all nontrivial elements of H', so L is open in G and must be virtually abelian.

3.2 Relative Commuting Probability

The notion of commuting probability can be generalized in the following way: let G be a compact group and K be a subgroup of G. Consider the set of commuting pairs $C = \{(x, y) \in K \times G \,|\, xy = yx\}$. This is closed in $K \times G$ since it is the preimage of 1 under the continuous map $f\colon K \times G \to G$ given by $f(x, y) = [x, y]$. Denoting the normalized Haar measures of K and G by ν and μ, respectively, the probability that a random element from K commutes with a random element from G is defined as $Pr(K, G) = (\nu \times \mu)(C)$ and is called the *relative commuting probability* of K in G. This notion has received increasing attention in recent years and basic properties about it can be found in [14].

We dedicate this section to record the results on relative commuting probability that are needed to prove Theorem A. The first result states that, roughly saying, it is easier to commute elements chosen in smaller subgroups.

Lemma 3.2.1. *Let H and K be subgroups of a compact group G, with $H \leq K$. Then*

$$Pr(K, G) \leq Pr(H, G) \leq Pr(H, K).$$

In particular, $Pr(G, G) \leq Pr(K, G) \leq Pr(K, K)$.

Proof. Let μ, ν and λ be the normalized Haar measures of G, K and H, respectively. Given $x \in G$, the map $\alpha : \{hC_H(x) \,|\, h \in H\} \to \{kC_K(x) \,|\, k \in K\}$ taking $hC_H(x)$ to $hC_K(x)$ is injective. We deduce that $[H : C_H(x)] \leq [K : C_K(x)]$ and $\nu(C_K(x)) \leq \lambda(C_H(x))$ by Lemma 2.6.4. We have

$$Pr(H, G) = \int_G \lambda(C_H(x)) d\mu(x) \geq \int_G \nu(C_K(x)) d\mu(x) = Pr(K, G).$$

The other inequality is proved in an analogous way. □

In the case of finite groups Lemma 3.2.5 was established in [9] and once again there is some intuition behind it: if N is a normal subgroup of G, Lemma 3.2.5 states that commutation is easier in the quotient group G/N than on G itself. To prove it we need some preliminary results, the first one being the extended Weil formula ([41, p.88]).

Lemma 3.2.2. *Let G be a compact group and H be a (closed) normal subgroup of G. If f is a measurable function and μ, ν and λ are the Haar measures on G, H and G/H respectively, then*

$$\int_{G/H} \left(\int_H f(xh) d\nu(h) \right) d\lambda(xH) = \int_G f(x) d\mu(x).$$

If $\phi : G \to H$ is a continuous isomorphism between compact groups, there is a natural way to relate the integral of a function f over G with an integral over H. This is Corollary 2.5 in [42].

Lemma 3.2.3. *Let G and H be compact groups with normalized Haar measures μ_G and μ_H. If $\phi : G \to H$ is a topological isomorphism and $f : H \to \mathbb{R}$ is a measurable function, then*

$$\int_G f(\phi(x)) d\mu_G(x) = \int_H f(x) d\mu_H(x).$$

The following is an identity used in the proof of Lemma 3.2.5.

Lemma 3.2.4. *Let G be a compact group with normal subgroup N and let μ_G and μ_N be the normalized Haar measures of G and N, respectively. If $x \in G$ is an FC-element, then*

$$\mu_G(C_G(x)N)\mu_N(C_N(x)) = \mu_G(C_G(x)).$$

Proof. We have that

$$\mu_G(C_G(x)N) = [G : C_G(x)N]^{-1} = \frac{[C_G(x)N : C_G(x)]}{[G : C_G(x)N][C_G(x)N : C_G(x)]},$$

which equals $[C_G(x)N : C_G(x)]\mu_G(C_G(x))$. Since

$$[C_G(x)N : C_G(x)] = [N : C_N(x)],$$

it follows that $\mu_G(C_G(x)N)\mu_N(C_N(x)) = \mu_G(C_G(x))$. □

Now we proceed to the proof of Lemma 3.2.5. The proof is technical and requires integrals of several functions. It is possible to prove that the integrands are indeed measurable with respect to the suitable σ-algebra considered on each step.

Lemma 3.2.5. *Let G be a compact group and let N be a normal subgroup of G. For any subgroup K of G, we have*

$$Pr(K,G) \leq Pr(KN/N, G/N)Pr(K \cap N, N).$$

Proof. If X is a compact group, we denote by μ_X the normalized Haar measure of X. Let $FC(K)$ be the abstract subgroup of K consisting of elements having finite conjugacy class in G and recall $FC(K)$ is a measurable subgroup of G. Applying Lemma 3.2.4 we have

$$\begin{aligned}\int_K \mu_G(C_G(x))d\mu_K(x) &= \int_{FC(K)} \mu_G(C_G(x))d\mu_K(x) \\ &= \int_{FC(K)} \mu_G(C_G(x)N)\mu_N(C_N(x))d\mu_K(x) \\ &\leq \int_K \mu_G(C_G(x)N)\mu_N(C_N(x))d\mu_K(x),\end{aligned}$$

where the last inequality follows from Proposition 2.5.16(ii). We now apply the extended Weil formula 3.2.2 to the last integral and obtain

$$\begin{aligned}&Pr(K,G) \leq \\ &\int_{\frac{K}{K\cap N}} \left(\int_{K\cap N} \mu_G(C_G(xk)N)\mu_N(C_N(xk))d\mu_{K\cap N}(k) \right) d\mu_{\frac{K}{K\cap N}}(x(K\cap N)) \\ \leq &\int_{\frac{K}{K\cap N}} \left(\int_{K\cap N} \mu_{\frac{G}{N}}(C_{\frac{G}{N}}(xN))\mu_N(C_N(xk))d\mu_{K\cap N}(k) \right) d\mu_{\frac{K}{K\cap N}}(x(K\cap N)) \\ = &\int_{\frac{K}{K\cap N}} \mu_{\frac{G}{N}}(C_{\frac{G}{N}}(xN)) \left(\int_{K\cap N} \mu_N(C_N(xk))d\mu_{K\cap N}(k) \right) d\mu_{\frac{K}{K\cap N}}(x(K\cap N)). \quad (1)\end{aligned}$$

If x is any element of K, define the set

$$\begin{aligned} A_x &= \{(k,n) \in (K \cap N) \times N \mid [xk,n] = 1\} \\ &= \{(k,n) \in (K \cap N) \times N \mid xk \in C_G(n) \cap x(K \cap N)\}. \end{aligned}$$

If $C_G(n) \cap x(K \cap N)$ is nonempty, then it equals $tC_{K\cap N}(n)$ for some $t \in x(K \cap N)$. Thus,

$$\begin{aligned} A_x &= \{(k,n) \in (K \cap N) \times N \mid xk \in tC_{K\cap N}(n)\} \\ &= \{(k,n) \in (K \cap N) \times N \mid k \in x^{-1}tC_{K\cap N}(n)\}. \end{aligned}$$

We use the Fubini Theorem 2.5.17 to give an estimate for the expression in parenthesis in (1):

$$\begin{aligned} \int_{K\cap N} \mu_N(C_N(xk))d\mu_{K\cap N}(k) &= \int_{(K\cap N)\times N} \chi_{A_x}(k,n)d(\mu_{K\cap N} \times \mu_N)(k,n) \\ &\leq \int_N \mu_{K\cap N}(x^{-1}tC_{K\cap N}(n))d\mu_N(n) \\ &= \int_N \mu_{K\cap N}(C_{K\cap N}(n))d\mu_N(n) \\ &= Pr(K \cap N, N). \end{aligned}$$

Replacing this back in (1) we have

$$\begin{aligned} Pr(K,G) \leq & \int_{\frac{K}{K\cap N}} \mu_{\frac{G}{N}}(C_{\frac{G}{N}}(xN)) \left(\int_{K\cap N} \mu_N(C_N(xk))d\mu_{K\cap N}(k) \right) d\mu_{\frac{K}{K\cap N}}(x(K \cap N)) \\ \leq & Pr(K \cap N, N) \int_{\frac{K}{K\cap N}} \mu_{\frac{G}{N}}(C_{\frac{G}{N}}(xN))d\mu_{\frac{K}{K\cap N}}(x(K \cap N)). \end{aligned}$$

Finally, since $K/K \cap N$ and KN/N are isomorphic, we can apply Lemma 3.2.3 with respect to the last integral above to conclude that

$$\begin{aligned} \int_{\frac{K}{K\cap N}} \mu_{\frac{G}{N}}(C_{\frac{G}{N}}(xN))d\mu_{\frac{K}{K\cap N}}(x(K \cap N)) &= \int_{\frac{KN}{N}} \mu_{\frac{G}{N}}(C_{\frac{G}{N}}(xN))d\mu_{\frac{KN}{N}}(xN) \\ &= Pr(KN/N, G/N). \end{aligned}$$

The lemma follows. □

Using Lemma 3.2.5, if G is compact and $Pr(K,G) \geq \epsilon$, then this property is inherited by all homomorphic images of G, since $Pr(K,G) \leq Pr(KN/N, G/N)$ for any normal subgroup N of G.

3.2.1 Examples

Here we collect some examples of groups having well-behaved commuting probability, to illustrate the methods that can be used to calculate it.

Example 3.2.6. [27, Lemma 5.1] Let $G = \prod_{i=1}^{\infty} H_i$ be the Cartesian product of the finite groups H_i. Then

$$Pr(G) = \lim_{n\to\infty} Pr(H_1)Pr(H_2)\cdots Pr(H_n),$$

in case it exists.

See that the sequence whose general term is $Pr(H_1)Pr(H_2)\ldots Pr(H_n)$ is non-increasing and bounded, so it must converge to its infimum.

Proof. Let μ be the product measure on $G\times G$. If H is a compact group, denote by $C(H)$ the set $\{(x,y) \in H \times H \mid [x,y] = 1\}$. Define the sets

$$C_n = C(H_1) \times \cdots \times C(H_n) \times H_{n+1} \times H_{n+1} \times \cdots \subseteq G \times G.$$

Then $\mu(C_n) = Pr(H_1)\cdots Pr(H_n)$. Identifying $\prod_{i=1}^{\infty} H_i \times H_i$ with $G \times G$, then $\bigcap_{i=1}^{\infty} C_n = C(G)$. We have that $\mu(\bigcap_{i=1}^{\infty} C_n) = \lim_{n\to\infty} \mu(C_n)$ by Lemma 2.5.10, as we wanted to prove. Moreover, it is clear that $Pr(G) > 0$ if and only if all but finitely many of the H_n are abelian, as nonabelian finite groups have commutativity degree $\leq \frac{5}{8}$ and this product converges if and only if $Pr(H_n) \to 1$. □

Still in the context of profinite groups we consider the problem to calculate $Pr(G)$ when G is the profinite completion of a given residually finite group.

Example 3.2.7. [6, Chapter 7, Paragraph 6] Let G be a profinite group equipped with normalized Haar measure μ. Consider $C = \{(x,y) \in G\times G \mid [x,y] = 1\}$ and let $\mu \times \mu$ be the product measure in $G \times G$. If N is an open normal subgroup of G, let μ_N be the discrete probability measure in G/N, which is induced by μ. The subset C is closed in G, so $\mu(C) = \inf\{\mu_N(CN/N)\}$ by Proposition 2.6.8, and as CN/N equals the set of commuting pairs of G/N, it follows that $Pr(G) = \inf\{Pr(G/N) \mid N \text{ normal open in } G\}$.

Now consider $D = \langle a, b \,|\, a^b = a^{-1}, b^2 = 1\rangle$, the infinite dihedral group, and let G be its profinite completion. The nonabelian finite images of D are all the finite dihedral groups $D_n = \langle a, b \,|\, a^b = a^{-1}, a^n = b^2 = 1\rangle$, where $|D_n| = 2n$. Let $k(H)$ denote the class number of a finite group H and recall that $Pr(H) = \frac{k(H)}{|H|}$. In particular, $k(D_n)$ equals $\frac{n+3}{2}$ if n is odd and $\frac{n+6}{2}$ if n is even, so we deduce that $Pr(D_n)$ equals $\frac{n+3}{4n}$ or $\frac{n+6}{4n}$ as n is odd or even, respectively. By the previous paragraph, $Pr(G)$ can be calculated taking the infimum of $Pr(G/N)$ only with respect to nonabelian quotients. We conclude that

$$Pr(G) = \inf_n Pr(D_n) = \frac{1}{4}.$$

The group G, as defined before, is also suitable to calculate an explicit example of relative commuting probability.

Example 3.2.8. We consider G as the profinite completion of the infinite dihedral group again. Now, recall that $G = \hat{\mathbb{Z}} \rtimes C_2$, where C_2 is cyclic of order 2 and acts by inversion on $\hat{\mathbb{Z}}$, the profinite completion of the integers. Let H be the subgroup of G isomorphic to $\hat{\mathbb{Z}}$, and let μ and ν be the normalized Haar measures on G and H, respectively. Then $Pr(H, G) = \frac{1}{2}$: if $x \in H$, then $C_G(x) = H$, so

$$Pr(H, G) = \int_H \mu(C_G(x)) d\nu(x) = \int_H [G : H]^{-1} d\nu H(x) = \frac{1}{2}.$$

In the next example we consider the problem whether nonmeasurable subsets of a compact group exist.

Example 3.2.9. Infinite profinite groups have nonmeasurable subsets. Let G be an infinite profinite group. Then G has a countable abstract subgroup H. If we let T be a left transversal for H in G, then $G = \bigcup_{h \in H} hT$. Assuming T is measurable, we get $1 = \sum_{h \in H} \mu(hT)$ and, since μ is translation-invariant, $1 = \sum_{h \in H} \mu(T)$, which is impossible. Thus T is nonmeasurable. See that the existence of T is guaranteed, but one needs Axiom of Choice to actually construct T.

3.3 Proof of Theorem A

In this section we prove our first main theorem, which we restate here one more time for the benefit of the reader.

Theorem A. *Let $\epsilon > 0$ and let G be a compact group having a subgroup K such that $Pr(K,G) \geq \epsilon$. Then there is a normal subgroup $T \leq G$ and a subgroup $B \leq K$ such that the indices $[G:T]$ and $[K:B]$ and the order of $[T,B]$ are ϵ-bounded.*

If A and B are normal subgroups of an arbitrary group G such that the indexes $[A : C_A(B)] \leq m$ and $[B : C_B(A)] \leq m$, then $[A,B]$ has m-bounded order. This result is due to Baer, cf. [44, 14.5.2]. We need a variation of it, which is Lemma 2.1 in [9].

Lemma 3.3.1. *Let $m \geq 1$ and let G be a group containing a normal subgroup A and a subgroup B such that $[A : C_A(y)] \leq m$ and $[B : C_B(x)] \leq m$ for all $x \in A, y \in B$. Assume further that $\langle B^G \rangle$ is abelian. Then $[A,B]$ has finite m-bounded order.*

The next theorem holds in any group and plays a key role in the proof of Theorem A. It is Theorem 1.1 in [2] and generalizes B.M. Neumann's result [39] on finiteness of the commutator subgroup of a BFC-group.

Theorem 3.3.2. *Let m be a positive integer, let G be a group having a subgroup K such that $|x^G| \leq m$ for each $x \in K$, and let $H = \langle K^G \rangle$. Then the order of the commutator subgroup $[H,H]$ is finite and m-bounded.*

Proof of Theorem A. Let μ and ν be the normalized Haar measures of G and K, respectively. Set

$$X = \{x \in K \,|\, |x^G| \leq 2/\epsilon\}.$$

Note that X is measurable, by Lemma 2.2.14: it is the intersection of K and the closed set $\{x \in G \,|\, |x^G| \leq 2/\epsilon\}$. We have

$$K \setminus X = \{x \in K \,|\, |x^G| > 2/\epsilon\}.$$

Since $\mu(C_G(x)) < \epsilon/2$ for all $x \in K \setminus X$, it follows that

$$\begin{aligned}
\epsilon \leq Pr(K,G) &= \int_K \mu(C_G(x))d\nu(x) \\
&= \int_X \mu(C_G(x))d\nu(x) + \int_{K\setminus X} \mu(C_G(x))d\nu(x) \\
&\leq \int_X d\nu(x) + \int_{K\setminus X} \frac{\epsilon}{2} d\nu(x) \\
&= \nu(X) + \frac{\epsilon}{2}(1-\nu(X)) \leq \nu(X) + \frac{\epsilon}{2}.
\end{aligned}$$

This implies that $\epsilon/2 \leq \nu(X)$. Let B be the subgroup generated by X. Then, by Lemma 2.6.6, every element of B is a product of at most $6/\epsilon$ elements of X. Clearly, $\nu(B) \geq \nu(X) \geq \epsilon/2$, so the index of B in K is at most $2/\epsilon$, by Lemma 2.6.5. Furthermore, $|b^G| \leq (2/\epsilon)^{6/\epsilon}$ for every $b \in B$.

Let $L = \langle B^G \rangle$. Theorem 3.3.2 tells us that the commutator subgroup $[L, L]$ has finite ϵ-bounded order. Let us use the bar notation for images of subgroups of G in $G/[L, L]$. By Lemma 3.2.5, $Pr(\overline{K}, \overline{G}) \geq Pr(K, G) \geq \epsilon$. Moreover, $[\overline{K} : \overline{B}] \leq [K : B] \leq \epsilon/2$ and $|\overline{b}^{\overline{G}}| \leq |b^G| \leq (2/\epsilon)^{6/\epsilon}$ for any $b \in B$. Thus we can pass to the quotient over $[L, L]$ and assume that L is abelian.

Now we set

$$Y = \{y \in G \,|\, |y^K| \leq 2/\epsilon\}.$$

Note that Y is closed, by Remark 2.2.15. Arguing as before, as $\nu(C_K(y)) < \epsilon/2$ for all $y \in G \setminus Y$, we have

$$\begin{aligned}
\epsilon \leq Pr(K, G) &= \int_G \nu(C_K(y)) d\mu(y) \\
&= \int_Y \nu(C_K(y)) d\mu(y) + \int_{G \setminus Y} \nu(C_K(y)) d\mu(y) \\
&\leq \int_Y d\mu(y) + \int_{G \setminus Y} \epsilon/2 d\mu(y) \\
&= \mu(Y) + \epsilon/2(1 - \mu(Y)) \leq \mu(Y) + \epsilon/2.
\end{aligned}$$

Therefore, $\mu(Y) \geq \epsilon/2$. Let E be the subgroup generated by Y. Lemma 2.6.6 ensures that every element of E is a product of at most $6/\epsilon$ elements of Y. Also, we have $\mu(E) \geq \mu(Y) \geq \epsilon/2$, so the index of E in G is at most $2/\epsilon$, by Lemma 2.6.5. Since $|y^K| \leq 2/\epsilon$ for every $y \in Y$, it follows that $|g^K| \leq (2/\epsilon)^{6/\epsilon}$ for every $g \in E$. Let T be the maximal normal subgroup of G contained in E. Then the index $[G : T]$ is ϵ-bounded. Moreover, $|b^G| \leq (2/\epsilon)^{6/\epsilon}$ for every $b \in B$ and $|g^K| \leq (2/\epsilon)^{6/\epsilon}$ for every $g \in T$. As L is abelian, we can apply Lemma 3.3.1 and deduce that $[T, B]$ has finite ϵ-bounded order. The theorem follows. □

In particular, Theorem A implies that if G is a compact group then either $Pr(G) = 0$ or G is virtually a BFC-group. Indeed, if we suppose that $Pr(G) > 0$, applying Theorem A we conclude that there is an open subgroup H of G such that H' is finite, and this condition is equivalent to H being a BFC-group, by B. H. Neumann's result [39].

Before we proceed to the applications of Theorem A, we need to make a few remarks. First we state a criterion for an element $x \in G$ to have a finite conjugacy class.

Remark 3.3.3. Let G be a group and $K \leq G$ with $[G : K] = m$ finite. If $|x^K| = n$ is finite, then $|x^G| \leq mn$.

Proof. To see this, just note that the index $[K : C_K(x)]$ equals n, so

$$[G : C_G(x)] \leq [G : C_K(x)] = [G : K][K : C_K(x)] = mn.$$

□

Following notation of Theorem A, the second remark concerns finiteness of the normal closure of $[T, B]$.

Remark 3.3.4. Under the Hypothesis of Theorem A, the subgroup $\langle [T, B]^G \rangle$ has finite ϵ-bounded order.

Proof. Observe that $[t_1, b]^{t_2} = [t_1 t_2, b][t_2, b]^{-1}$ for any $t_1, t_2 \in T$ and $b \in B$. So $[T, B]$ is normal in T and the normalizer of $[T, B]$ is open of ϵ-bounded index, proving that $[T, B]$ has boundedly many conjugates. Normality of T implies that all conjugates of $[T, B]$ normalize each other. Since $\langle [T, B]^G \rangle$ equals the product of those subgroups, the result follows. □

Theorem A allows us to derive different structural properties depending on the subgroup K considered. In [9], for example, the authors study the case where $K = \gamma_i(G)$, that is, $Pr(\gamma_i(G), G) \geq \epsilon$. They prove that G has a nilpotent normal subgroup R of of class at most $i + 1$ such that the index $[G : R]$ and the order of $\gamma_{i+1}(R)$ are ϵ-bounded. We consider a variation of this result.

Theorem 3.3.5. *Let G be a compact group, let $K = G^{(i)}$ and assume that $Pr(K, G) \geq \epsilon$. Then G has an open normal subgroup R, soluble of derived length at most $i + 2$, such that the index $[G : R]$ and the order of $R^{(i+1)}$ are ϵ-bounded. Moreover, G has a soluble open subgroup M of derived length at most $i + 1$.*

Proof. By Theorem A, there is a normal subgroup T of G and a subgroup B of K such that $[G : T]$, $[K : B]$ and the order of $[T, B]$ are ϵ-bounded. Since $T^{(i)} \leq K$ we know that $[T^{(i)}, B]$ has ϵ-bounded order, so $|x^B|$ is bounded for every $x \in T^{(i)}$ and $|x^K|$ has bounded size by Remark 3.3.3. In particular, $T^{(i)}$ is

a BFC-group with ϵ-bounded BFC-number so $T^{(i+1)} = [T^{(i)}, T^{(i)}]$ must be finite of ϵ-bounded order. Let $R = C_T(T^{(i+1)})$. Then R is open in G, the index $[G : R]$ and the order of $R^{(i+1)}$ are ϵ-bounded and R is soluble of derived length at most $i + 2$.

Now let M be an open subgroup of R avoiding all nontrivial elements of $R^{(i+1)}$ and see that $M^{(i+1)} = 1$. As M is open in G and soluble of derived length at most $i + 1$, the proof is complete. □

In [9] the authors consider the case where G is a finite group such that $Pr(P, G) \geq \epsilon$ for a Sylow p-subgroup P of G. They conclude that G has a class-2-nilpotent normal p-subgroup L such that both the index $[P : L]$ and the order of $[L, L]$ are ϵ-bounded. We consider a profinite variant of this result. Here, if G is any group, we denote by $Z_i(G)$ the i-th term of the upper central series of G, that is,

$$Z_i(G) = \{x \in G \,|\, [x, g_1, \ldots, g_i] = 1\},$$

which is a characteristic subgroup of G.

Theorem 3.3.6. *Let p be a prime and let G be a profinite group with a Sylow pro-p subgroup P such that $Pr(P, G) \geq \epsilon$. Then G has a normal p-subgroup L, nilpotent of class at most 2, such that the index $[P : L]$ and the order of $[L, L]$ are ϵ-bounded. Moreover, P is virtually abelian.*

Proof. Applying Theorem A, we have a normal subgroup T of G and a subgroup B of P such that the indexes $[G : T]$ and $[P : B]$ and the order of $[T, B]$ are ϵ-bounded. By Remark 3.3.4 the subgroup $N = \langle [T, B]^G \rangle$ has finite ϵ-bounded order, and $C = C_T(N)$ has ϵ-bounded index in G. Set $B_0 = B \cap C$ and see that $[C, B_0] \leq Z(C)$. It follows that $B_0 \leq Z_2(C)$ and all conjugates of B_0 lie within $Z_2(C)$. Moreover, since B_0 is a pro-p subgroup of $Z_2(C)$, it is contained in some Sylow pro-p subgroup B_1 of $Z_2(C)$. The subgroup $Z_2(C)$ is nilpotent, thus B_1 is normal in G and we conclude that B_0 is contained in all Sylow pro-p subgroups of G. Let $L = \langle B_0^G \rangle$. It is clear that L is contained in P as a subgroup of ϵ-bounded index. Moreover, $[L, L] \leq N$, so the order of $[L, L]$ is finite. The first part of the theorem is proved.

Now, since $[L, L]$ is finite and L is profinite, there exists an open normal subgroup M of L avoiding all nontrivial elements from $[L, L]$. It follows that M is abelian and open in P. The proof is complete. □

The following is an upper bound for $Pr(H)$ whenever H is a nonabelian pro-p group. It is Lemma 3.4 in [42].

Lemma 3.3.7. *Let p be a prime and let H be a nonabelian pro-p group. Then*

$$Pr(H) \leq \frac{p^2+p-1}{p^3}.$$

In Theorem 3.3.6 we assumed that $Pr(P,G) \geq \epsilon$ for a Sylow p-subgroup P of G for a fixed prime p. However, if we assume that for all prime numbers a Sylow pro-p subgroup of G satisfies $Pr(P,G) \geq \epsilon$, then a stronger result holds.

Theorem 3.3.8. *Let G be a profinite group such that $Pr(P,G) \geq \epsilon$ whenever P is a Sylow subgroup. Then G has a nilpotent normal subgroup R, of class at most 2, such that the index $[G:R]$ and the order of $[R,R]$ are ϵ-bounded. Moreover, G is virtually abelian.*

Proof. For each prime number p we fix a Sylow pro-p subgroup S_p of G. Applying Theorem 3.3.6, for each p there is a normal class-2-nilpotent subgroup L_p of G such that the index $[S_p : L_p]$ and the order of $[L_p, L_p]$ are ϵ-bounded. Moreover, using Lemma 3.2.1, it follows that $Pr(S_p) \geq Pr(S_p, G) \geq \epsilon$. Using Lemma 3.3.7, whenever S_p is nonabelian we have

$$Pr(S_p) \leq \frac{p^2+p-1}{p^3}.$$

Since $Pr(S_p) \geq \epsilon$ and taking the limit of the above fraction in p, we see that there exists a constant C depending only on ϵ such that $p \geq C$ implies that S_p is abelian. In particular, if $p \geq C$, then $L_p = S_p$. Set

$$R = \prod_{p \text{ a prime}} L_p.$$

See that G/R has boundedly many nontrivial Sylow subgroups, all of which have ϵ-bounded order, thus we conclude that R has finite ϵ-bounded index in G. Furthermore, as

$$[R,R] = \prod_{p \text{ a prime}} [L_p, L_p],$$

it follows that the order of $[R,R]$ is finite and ϵ-bounded. Note that R is also nilpotent of class 2.

Now, let M be an open normal subgroup of R avoiding all nontrivial elements of $[R,R]$. It follows that M is abelian and open in G. □

We finish this section proving a converse for Theorem A.

Proposition 3.3.9. *Let G be a compact group with a subgroup K, let T be an open subgroup of G, let B be an open subgroup of K and suppose that $[T,B]$ has finite order. Then $Pr(K,G)$ is bounded away from zero in terms of $[G:T]$, $[K:B]$ and the order of $[T,B]$.*

Proof. Let $[G:T]=l$, $[K:B]=m$ and $|[T,B]|=n$. As we did before, it suffices to prove that for every $x\in K$ the size of the conjugacy class x^G is (l,m,n)-bounded, and we do so by counting how may weight 2 commutators there are with x as an entry, by Remark 2.1.1. Fix a right transversal for T in G and one for B in K and let $x\in K$ and $g\in G$ be arbitrary. Write $x=ba$ and $g=ts$, where $b\in B$, $t\in T$, a is a coset representative of K modulo B and s is a coset representative of G modulo T. Using commutator identities, we have

$$[x,g]=[ba,ts]=[b,s]^a[b,t]^{sa}[a,s][a,t]^s.$$

Remark 2.1.1 ensures that $|t^B|\leq n$, so it follows that $|t^K|\leq mn$ by Remark 3.3.3. In the same way we prove that $|b^G|\leq ln$. We conclude that there are at most lmn possibilities for the first and last commutators in the product above. There are also at most lmn possibilities for the second and lm for the third. Thus $|x^G|\leq l^4m^4n^3$. □

Chapter 4

A probabilistic notion of exponent

Results in this chapter concern a probabilistic generalization of the notion of exponent in a group. Let G be a compact group and recall that we denote by $\langle x \rangle$ the subgroup topologically generated by $x \in G$. If $\epsilon > 0$ is given and K is a subgroup of G, we impose the condition that $Pr(\langle x \rangle, G) \geq \epsilon$ for every $x \in K$ and study the consequences of this condition on the structure of G. In this chapter we prove Theorem B, which we restate here for the reader's convenience:

Theorem B. *Let $\epsilon > 0$ and let G be a compact group such that $Pr(\langle x \rangle, G) \geq \epsilon$ for every x in a closed subgroup K of G. Then there is an ϵ-bounded number e and an open normal subgroup T of G such that $[G : T]$ is ϵ-bounded and $[K^e, T] = 1$.*

We also give some applications of the above theorem on the context of finite groups. The corollaries are related to structural results on the exponent of a group and, unless stated otherwise, all results presented here come from [5].

4.1 Preliminaries

In this and the next sections, G will be an arbitrary compact group. Let $\epsilon > 0$ be given. As it was mentioned, we consider a group G with a subgroup K such that $Pr(\langle x \rangle, G) \geq \epsilon$ for all monothetic subgroups of K. When this is the case, we say that K has high commuting probability on monothetic subgroups.

Observe that if G has finite exponent e then $|\langle x \rangle| \leq e$ and $Pr(\langle x \rangle, G) \geq \frac{1}{e}$ for all $x \in G$. Indeed, if μ and ν denote normalized Haar measures in G and

$\langle x\rangle$, respectively, then

$$Pr(\langle x\rangle, G) = \int_G \nu(C_{\langle x\rangle}(y))d\mu(y) = \int_G [\langle x\rangle : C_{\langle x\rangle}(y)]^{-1}d\mu(y) \geq \int_G \frac{1}{e}d\mu(y) = \frac{1}{e}.$$

More generally, if $G^e \leq Z(G)$, then $Pr(\langle x\rangle, G) \geq \frac{1}{e}$ for all $x \in G$. This explains that high commuting probability on monothetic subgroups is, in some sense, a generalization of having finite exponent. The concept was developed to treat probabilistically some finiteness conditions that are related to having finite exponent or related properties.

The condition $Pr(\langle x\rangle, G) \geq \epsilon$ for all $x \in K$ is strictly weaker than the condition in Theorem A. If $Pr(K, G) \geq \epsilon$, then $Pr(\langle x\rangle, G) \geq \epsilon$ for all $x \in K$, by Lemma 3.2.1, but the converse is not true. To see this, let $G = \prod_{i\in I} S_i$ be the Cartesian product of infinitely many copies of the same nonabelian finite simple group S, of exponent e. Then $Pr(\langle x\rangle, G) \geq \frac{1}{e}$, but $Pr(G) = 0$. If $Pr(G)$ was positive, G would admit a normal open subgroup T such that T' is finite. However, we know that T must contain a subgroup of the form $\prod_{j\in J} S_i$, where $I\backslash J$ is finite, and T coincides with T' and is infinite. In Corollary 4.2.5 we explore a situation where both conditions are equivalent.

Let us recall the result on Lemma 3.2.4. Let G be a compact group with a normal subgroup N, and let μ_G and μ_N be the normalized Haar measures of G and N, respectively. It is true that, for every $g \in G$, the following equality holds

$$\mu_G(C_G(g)N)\mu_N(C_N(g)) = \mu_G(C_G(g)).$$

Using Lemma 3.2.4 we can derive an easy criterion to verify whether a subgroup K of G has high commuting probability on monothetic subgroups.

Lemma 4.1.1. *Let l, n be positive integers. Let K be a subgroup of G such that any conjugacy class containing an lth power of an element $x \in K$ is of size at most n. Then $Pr(\langle x\rangle, G) \geq \frac{1}{ln}$ for every $x \in K$.*

Proof. Note that $\langle x^l\rangle$ is normal in $\langle x\rangle$. Using Lemma 3.2.4, we have

$$\begin{aligned} Pr(\langle x\rangle, G) = \int_G \mu_{\langle x\rangle}(C_{\langle x\rangle}(g))d\mu_G(g) &= \int_G \mu_{\langle x\rangle}(C_{\langle x\rangle}(g)\langle x^l\rangle)\mu_{\langle x^l\rangle}(C_{\langle x^l\rangle}(g))d\mu_G(g) \\ &\geq \int_G \frac{1}{l}\mu_{\langle x^l\rangle}(C_{\langle x^l\rangle}(g))d\mu_G(g) \\ &= \frac{1}{l}\cdot Pr(\langle x^l\rangle, G) \geq \frac{1}{ln}. \end{aligned}$$

□

We end this introduction with the remark that Theorem B admits a converse. The hypothesis are a little more general here.

Proposition 4.1.2. *Let s, e, m be positive integers. Assume that G is a compact group with a normal open subgroup T of index at most s and a subgroup K such that $[K^e, T]$ is finite of order at most m. Then there is $0 < \epsilon \leq 1$ depending only on s, e, m such that $Pr(\langle x \rangle, G) \geq \epsilon$ for every $x \in K$.*

To see this simply note that if K is as above, then $[G : C_G(x^e)] \leq ms$ for every $g \in K$ and use Lemma 4.1.1.

4.2 Proof of Theorem B

First of all, observe that Lemma 4.1.1 admits a converse.

Lemma 4.2.1. *Let $\epsilon > 0$ and let G be a compact group such that $Pr(\langle x \rangle, G) \geq \epsilon$ for every x in a closed subgroup K of G. Then there exist ϵ-bounded integers l and n with the property that $[G : C_G(x^l)] \leq n$ for all $x \in K$.*

Proof. Let $x \in K$. Since $Pr(\langle x \rangle, G) \geq \epsilon$, in view of Theorem A there is a normal subgroup T of G and a subgroup B of $\langle x \rangle$ such that the indices $[G : T]$ and $[\langle x \rangle : B]$ and the order of $[T, B]$ are ϵ-bounded. Hence, as required, there are ϵ-bounded numbers l and n such that $[G : C_G(x^l)] \leq n$ for all $x \in K$. □

Taking this into consideration, Theorem B will follow from the next proposition.

Proposition B′. *Let G be a compact group and let l, n be positive integers. Suppose that there is a subgroup K of G such that $[G : C_G(x^l)] \leq n$ for every $x \in K$. Then there exist a positive integer e and a normal subgroup T of G such that $[K^e, T] = 1$ and both e and the index $[G : T]$ are (l, n)-bounded.*

We need to make some considerations before we proceed to the proof of Proposition B′. Let G be a topological group generated by a symmetric set X, that is, $X = X^{-1} = \{x^{-1} \mid x \in X\}$. If it is possible to write $g \in G$ as a product of finitely many elements from X, we denote by $w(g)$ the shortest length of such an expression. If g cannot be written as a product of finitely many elements of X, we simply say that $w(g)$ is infinite. The next result is Lemma 2.1 in [10].

Lemma 4.2.2. *Let G be a group generated by a symmetric set X and let D be a subgroup of index m in G. Then every coset Db contains an element such that $w(g) \leq m - 1$.*

We remark that Lemma 4.2.2 holds for topological groups and their closed (open) subgroups of finite index. Indeed, for an integer $r \geq 0$ let D_r be the union of the cosets of D containing some element g with $w(g) \leq r$. Then $D_r \subseteq D_{r+1}$ and $D_r X \subseteq D_{r+1}$ for all r. Let R be the minimal integer such that $D_{R+1} = D_R$. Then D_R is a closed set containing the group generated by X, so $D_R = G$. Since D has m cosets and $D_0 = D$, it follows that $R < m$.

Now we are able to embark on the proof of Proposition B′, which follows ideas from [2] and [10]. Also, we observe that some auxiliary lemmas will be proved within the main demonstration.

Proof of Proposition B′. Let G be a compact group and let n, l be positive integers. Assume that $[G : C_G(x^l)] \leq n$ for every x in a closed subgroup K of G. Let X be the union of the conjugacy classes of G containing an lth power of an element of K and let H be the subgroup generated by X. Define m as the maximum of the indices of $C_H(x)$ in H, where $x \in X$. Obviously, $m \leq n$.

Lemma 4.2.3. *For any $x \in X$ the order of the subgroup $[H, x]$ is m-bounded.*

Proof. Since the index of $C_H(x)$ in H is at most m, Lemma 4.2.2 guarantees that there are elements $y_1, \ldots, y_m$ in H such that each y_i is a product of at most $m-1$ elements of X and the subgroup $[H, x]$ is generated by the commutators $[y_i, x]$, for $i = 1, \ldots, m$. For any such i write $y_i = y_{i1} \ldots y_{i(m-1)}$, where y_{ij} belongs to X. Using the standard commutator identities, we can rewrite $[y_i, x]$ as a product of conjugates in H of the commutators $[y_{ij}, x]$. Let $\{h_1, \ldots, h_s\}$ be the set of conjugates in H of all elements from the set $\{x, y_{ij} \mid 1 \leq i, j \leq m-1\}$. Note that the number s here is m-bounded. This follows from the fact that $C_H(x)$ has index at most m in H for every $x \in X$. Let D be the subgroup of H generated by $h_1, \ldots, h_s$. Since $[H, x]$ is contained in the commutator subgroup D', it suffices to show that D' has finite m-bounded order. Observe that the center $Z(D)$ has index at most m^s in D, since the index of $C_H(h_i)$ is at most m for every h_i. Thus, by Schur's theorem [44, 10.1.4], we conclude that D' has finite m-bounded order. □

We now argue by induction on m. If $[H : C_H(g^l)] \leq m-1$ for all $g \in K$, then by induction the result follows. We therefore assume that there is $d \in K$ such that $[H : C_H(d^l)] = m$. Set $a = d^l$ and choose $b_1, \ldots, b_m$ in H such that $a^H = \{a^{b_i} \mid i = 1, \ldots, m\}$ and $w(b_i) \leq m-1$ (the existence of the elements b_i is guaranteed by Lemma 4.2.2). Moreover, observe that $[H : C_H(a)] \leq m$. If

$c \in H$, we can rewrite the conjugate a^c as

$$a^c = (d^l)^c = (d^c)^l = (d^{b_i})^l = (d^l)^{b_i} = a^{b_i}$$

for some i. We conclude that the set $\{(a^l)^{b_i} \mid i = 1, \dots, m\}$ contains the conjugacy class of d^l, possibly with repetitions. Set $U = C_G(\langle b_1, \dots, b_m \rangle)$. Note that the index of U in G is n-bounded. Indeed, since $w(b_i) \leq m-1$ we can write $b_i = b_{i1} \dots b_{i(m-1)}$, where $b_{ij} \in X$ and $i = 1, \dots, m$. By the hypothesis, the index of $C_G(b_{ij})$ in G is at most n for any such element b_{ij}. Thus, $[G : U] \leq n^{(m-1)m}$.

Lemma 4.2.4. *Suppose that $u \in U$ and $ua \in X$. Then $[H, u] \leq [H, a]$.*

Proof. For each $i = 1, \dots, m$ we have $(ua)^{b_i} = ua^{b_i}$, since u belongs to U. By hypothesis, $ua \in X$. Hence, taking into account the assumption on the size of the conjugacy class of ua in H, we deduce that $(ua)^H$ consists exactly of the elements $(ua)^{b_i}$, for $i = 1, \dots, m$. Therefore, given an arbitrary element $h \in H$, there exists $b \in \{b_1, \dots, b_m\}$ such that $(ua)^h = (ua)^b$ and so $u^h a^h = ua^b$. It follows that $[u, h] = a^b a^{-h} \in [H, a]$, and the lemma holds. □

Let R be the normal closure in G of the subgroup $[H, a]$, that is, $R = [H, a^{b_1}] \dots [H, a^{b_n}]$, where a^{b_i} are all the conjugates of a in G (if $|a^G| \leq n-1$, then not all the $a^{b_1}, \dots, a^{b_n}$ are pairwise distinct). By Lemma 4.2.3, each of the subgroups $[H, a^{b_i}]$ has n-bounded order. Thus, the order of R is n-bounded as well.

Let $Y_1 = Xa^{-l} \cap U$ and $Y_2 = Xa^{-1} \cap U$. Note that for any $y \in Y_1$, the product ya^l belongs to X. So, by Lemma 4.2.4 applied with a^l in place of a and y in place of u, the subgroup $[H, y]$ is contained in $[H, a^l]$, which is contained in R. Similarly, for any $y \in Y_2$, we have $[H, y] \leq R$. Set $Y = Y_1 \cup Y_2$. Thus, $[H, Y] \leq R$.

Let U_0 be the maximal normal subgroup of G contained in U. Observe that the index of U_0 in G is n-bounded. Observe further that for any $u \in U_0$ the commutators $[u, a^{-l}]$ and $[u, a^{-1}]$ lie in Y. Since $[U_0, a^{-1}] = [U_0, a]$, we deduce that

$$[H, [U_0, a]] \leq R.$$

Let $K_0 = K \cap U_0$. We remark that $(ua)^l (a^l)^{-1} \in Y$ whenever $u \in K_0$. We pass to the quotient $\overline{G} = G/R$ and use the bar notation to denote images in $\overline{G}$. We know that $\overline{Y}$ is central in $\overline{H}$. We also deduce that $[\overline{U}_0, \overline{a}] \leq Z(\overline{H})$.

Since $(ua)^l(a^l)^{-1} \in Y$ whenever $u \in K_0$ and since $\overline{Y} \leq Z(\overline{H})$, it follows that in the quotient $\overline{G}/Z(\overline{H})$ the element $\overline{a}$ commutes with $\overline{U}_0$ and $(\overline{ua})^l(\overline{a})^{-l} = 1$ for every $\overline{u} \in \overline{K}_0$. This implies that $\overline{K}_0$ has exponent dividing l modulo $Z(\overline{H})$. It follows that $\overline{K}_0^l$ is abelian and every element of $\overline{K_0}^{l^2}$ is again an lth power of an element in $\overline{K}_0$. We therefore deduce that

$$Pr(\overline{K}_0^{l^2}, \overline{G}) \geq \frac{1}{n}.$$

By Theorem A there is a normal subgroup $\overline{T}$ in $\overline{G}$ and a subgroup $\overline{V}$ in $\overline{K}_0^{l^2}$ such that the indices $[\overline{G} : \overline{T}]$ and $[\overline{K}_0^{l^2} : \overline{V}]$ and the order of $[\overline{T}, \overline{V}]$ are (n,l)-bounded. Let T be the inverse image of $\overline{T}$ in G and V the inverse image of $\overline{V}$ in $K_0^{l^2}$. Bearing in mind that the order of R is n-bounded, we conclude that the indices $[G : T]$ and $[K_0^{l^2} : V]$ are (n,l)-bounded, as also is the order of $[T, V]$. As the index of V in $K_0^{l^2}$ is bounded, there is a positive (n,l)-bounded integer e such that $K^e \leq V$. So far we have proved that $[K^e, T]$ has finite (n,l)-bounded order, where the integer e and the index $[G : T]$ are (n,l)-bounded.

Assume that $|[K^e, T]| = d$, let $k \in K$, $t \in T$ and consider the commutators $[k^{ei}, t]$ for $i = 0, 1, \ldots, d$. The pigeonhole principle ensures that there are two integers $u \geq v$ such that $[x^{eu}, t] = [x^{ev}, t]$, which in turn implies that $x^{e(u-v)}$ centralizes t. Now, as $1 \leq u - v \leq d$, it follows that $[x^{d!e}, t] = 1$ for arbitrary $x \in K$ and $t \in T$. Defining $f = d!e$, it follows that $[K^f, T] = 1$. This completes the proof of the proposition. □

Theorem B has different consequences according to the choice of K. For example, if $K = G_0$, then $Pr(G_0, G) \geq \epsilon$ is equivalent to $Pr(\langle x \rangle, G) \geq \epsilon$ for every $x \in G_0$.

Corollary 4.2.5. *Let G be a compact group and let G_0 be the connected component of the identity in G. Then, the following are equivalent*

(i) The probability $Pr(G_0, G)$ is positive;

(ii) The centralizer $C_G(G_0)$ is open in G;

(iii) There exists $\epsilon > 0$ such that $Pr(\langle x \rangle, G) \geq \epsilon$ for any $x \in G_0$.

Proof. Suppose first that $Pr(G_0, G) > 0$. Then there are subgroups of finite index T of G and B of G_0 such that $[T, B]$ is finite, by Theorem A. Since G_0 is divisible [37], the only finite index subgroup of G_0 is G_0 itself. Therefore $[T, G_0]$

is a finite connected subgroup of G, what implies that $[T, G_0] = 1$. We conclude that $T \leq C_G(G_0)$, so (i) implies (ii).

Now, assume that $C_G(G_0)$ is open in G, let $x \in G_0$ and let μ and μ_0 be the normalized Haar measures of G and $\langle x \rangle$, respectively. For any $y \in G_0$, the inclusion $C_G(G_0) \leq C_G(y)$ holds, thus $\mu(C_G(G_0)) \leq \mu(C_G(y))$. We have

$$Pr(\langle x \rangle, G) = \int\limits_{\langle x \rangle} \mu(C_G(y)) d\mu_0(y) \geq \mu(C_G(G_0)) > 0.$$

We conclude that $Pr(\langle x \rangle, G) \geq \epsilon$ where $\epsilon = [G : C_G(G_0)]^{-1}$, and so (ii) implies (iii).

Now we assume the validity of (iii) and prove that (i) holds. By Theorem B, there are a finite index subgroup T of G and a natural number e such that $[G_0^e, T] = 1$. Since G_0 is divisible, $G_0 = G_0^e$ and so T centralizes G_0 and $T \leq C_G(x)$ for any $x \in G_0$. Writing μ and μ_0 for the normalized Haar measures of G and G_0, respectively, we have

$$Pr(G_0, G) = \int\limits_{G_0} \mu(C_G(x)) d\mu_0(x) \geq \mu(T) > 0.$$

□

4.3 Applications of Theorem B

We say that a group G is *torsion* if every element of G has finite order. It was proved in [25, p.69] that a compact torsion group must be profinite. In particular, concerning groups with finite exponent, the Restricted Burnside Problem was whether the order of an r-generated finite group of exponent e is bounded in terms of e and r alone, and it was famously solved in the affirmative by Zelmanov [52, 53]. Thus we have the following result.

Theorem 4.3.1. *Let G be a compact group of exponent e, (topologically) generated by r elements. Then G is finite of (e, r)-bounded order.*

The first application of Theorem B is a probabilistic generalization of the theorem above.

Theorem 4.3.2. *Let G be an r-generated compact group where $Pr(\langle x \rangle, G) \geq \epsilon$ for all $x \in G$. Then G has a normal abelian subgroup N such that the index $[G : N]$ is (r, ϵ)-bounded.*

Proof. By Theorem B there is an ϵ-bounded number e and a normal subgroup $T \leq G$ such that the index $[G : T]$ is ϵ-bounded and $[G^e, T] = 1$. Theorem 4.3.1 implies that the index $[G : G^e]$ is (e, r)-bounded. Therefore the subgroup $N = G^e \cap T$ has the required properties. □

In [36], Mann proves the following result. Here, G does not need to be a compact group.

Theorem 4.3.3. *Let G be a group such that $G/Z(G)$ is locally finite of exponent e. Then the commutator subgroup $[G, G]$ has finite e-bounded exponent.*

Using Theorem 4.3.3 we can derive further structural properties of a compact group G where all monothetic subgroups have high commuting probability.

Theorem 4.3.4. *Let G be a compact group such that $Pr(\langle x \rangle, G) \geq \epsilon$ for all $x \in G$. Then there is an open normal subgroup T of G such that the index $[G : T]$ and the exponent of $[T, T]$ are ϵ-bounded.*

Proof. Theorem B implies there is an ϵ-bounded number e and a normal subgroup T such that the index $[G : T]$ is ϵ-bounded and $[G^e, T] = 1$. It is clear that all eth powers of elements of G centralize T. In particular, $T/Z(T)$ has exponent e, and Theorem 4.3.1 ensures that $T/Z(T)$ is locally finite. Theorem 4.3.3 can then be used to deduce that $[T, T]$ has finite e-bounded exponent. □

We now turn to structural results on profinite groups. J. S. Wilson, in [50, Lemma 2], considered the following result.

Lemma 4.3.5. *Let $\mathfrak{X}_1, \ldots, \mathfrak{X}_n$ be classes of finite groups closed with respect to taking normal subgroups and subdirect products and let $\mathfrak{X}$ be the class of groups H having a series*

$$1 = H_0 \leq H_1 \leq \cdots \leq H_n = H$$

with $H_i/H_{i-1} \in \mathfrak{X}_i$ for each i. Then every pro-$\mathfrak{X}$ group has a series of closed characteristic subgroups

$$1 = G_0 \leq G_1 \leq \cdots \leq G_n$$

such that G_i/G_{i-1} is a pro-$\mathfrak{X}_i$ group for each i.

Furthermore, Lemma 3.2.5 allows us to make the following consideration.

Remark 4.3.6. Let G be a profinite group with a subgroup K and let N be an open normal subgroup of G. If we suppose that K has high commuting probability on monothetic subgroups, then Lemma 3.2.5 ensures that this property is inherited by the image KN/N of K in G/N.

Using Lemma 4.3.5 and Remark 4.3.6 we are able to apply the usual inverse limit argument to reduce the next two applications of Theorem B to the class of finite groups. Both of them concern the existence of a finite length series in a profinite group G with some prescribed properties. This is the reason why we turn our attention to finite groups now.

An important part of the eventual solution of the restricted Burnside problem was developed by Hall and Higman in their paper [23]. They proved the following theorem.

Theorem 4.3.7. *Let p be a prime and G a finite group with Sylow p-subgroups of exponent p^s. Then G has a normal series all of whose factors are p-groups, p'-groups or direct products of nonabelian simple groups of order divisible by p and whose length is s-bounded.*

We remark that every finite group G has a series as stated in Theorem 4.3.7. The important conclusion of the theorem is that the length of this series is s-bounded. We establish the following result.

Theorem 4.3.8. *Let p be a prime and G a finite group such that $Pr(\langle x\rangle, G) \geq \epsilon$ for x in a Sylow p-subgroup. Then G has a normal series of ϵ-bounded length all of whose factors are p-groups, p'-groups or direct products of nonabelian simple groups of order divisible by p.*

Proof. Let P be a Sylow p-subgroup of G. By Theorem B there is an ϵ-bounded positive integer e and a normal subgroup $T \leq G$ such that the index $[G : T]$ is ϵ-bounded and $[P^e, T] = 1$. Since T has ϵ-bounded index on G, it is sufficient to show that T has a normal series with the required properties. Let Z be the center of T and consider the normal series of T given by $1 \leq Z \leq T$. The Sylow p-subgroup $Z \cap P$ of Z is normal in G and the quotient $Z/(Z \cap P)$ is a p'-group. Furthermore, since $[P^e, T] = 1$, we deduce that T/Z has Sylow p-subgroups of exponent dividing e. According to the Hall-Higman Theorem 4.3.7, T/Z has a normal series of ϵ-bounded length with the required properties. Thus, the result follows. □

Theorem 4.3.8 can be extended to profinite groups, as we remarked before. The previous discussion yields the following result.

Theorem 4.3.9. *Let G be a profinite group such that $Pr(\langle x\rangle, G) \geq \epsilon$ for x in a Sylow pro-p subgroup P of G. Then G has a series of ϵ-bounded length of closed normal subgroups whose factors are pro-p, pro-p' or Cartesian products of nonabelian finite simple groups of order divisible by p.*

Let G be a finite group. A group-word w is said to be a law in G if $w(G) = 1$. In view of Theorem B, if $Pr(\langle x\rangle, G) \geq \epsilon$ for every x in $w(G)$, then a law of (ϵ, w)-bounded length holds in the group G. Indeed, there is an ϵ-bounded number e and an normal subgroup T of G such that the index $[G : T]$ is ϵ-bounded and $[w(G)^e, T] = 1$. There exists an ϵ-bounded integer f such that $G^f \leq T$, so the word $[w^e, x^f]$ is a law on G. This simple observation provides a tool for obtaining extensions of results about finite groups satisfying certain laws. We will illustrate this with a theorem bounding the nonsoluble length of a finite group. The following is the main result of [16].

Theorem 4.3.10. *Let G be a finite group, P be a Sylow 2-subgroup of G and let w be a group-word. If w is a law on P, then the nonsoluble length $\lambda(G)$ is bounded in terms of w only.*

This result can be extended as follows.

Theorem 4.3.11. *Let w be a group-word and P a Sylow 2-subgroup of a finite group G such that $Pr(\langle x\rangle, G) \geq \epsilon$ for every $x \in w(P)$. Then the nonsoluble length $\lambda(G)$ is (ϵ, w)-bounded.*

Proof. By Theorem B, there is a normal subgroup T of G and a positive integer e such that the index $[G : T]$ and e are ϵ-bounded and $[w(P)^e, T] = 1$. Let $[G : T] = f$. Then $[w(P)^e, G^f] = 1$ and, in particular, $[w(P)^e, P^f] = 1$. We conclude that P satisfies an identity of (ϵ, w)-bounded length and, by Theorem 4.3.10, $\lambda(G)$ is (ϵ, w)-bounded. □

We have thus proved

Theorem 4.3.12. *Let G be a profinite group and let P be a Sylow pro-2 subgroup of G. If $Pr(\langle x\rangle, G) \geq \epsilon$ for every $x \in w(P)$, then the nonsoluble length $\lambda(G)$ is finite and (ϵ, w)-bounded.*

The next two applications of Theorem B concern automorphisms of profinite groups. Recall that $A^{\#}$ represents the set of nontrivial elements in the group A. The main result of [31] states the following.

Theorem 4.3.13. *Let G be a finite group admitting an elementary abelian coprime group of automorphisms A of order p^2. Suppose further that $C_G(\phi)$ has exponent dividing d for each $\phi \in A^{\#}$. Then the exponent of G is (d,p)-bounded.*

Using Lemma 2.3.2, Theorem 4.3.13 can be extended in the following way.

Theorem 4.3.14. *Let G be a profinite group admitting a noncyclic group of coprime automorphisms of order p^2. If $C_G(\phi)$ has exponent dividing d for each $\phi \in A^{\#}$, then G has (d,p)-bounded exponent.*

We can now consider a probabilistic variant of the above result.

Theorem 4.3.15. *Let $\epsilon > 0$ and let G be a profinite group admitting an elementary abelian coprime group of automorphisms A of order p^2. Suppose that $Pr(\langle x\rangle, G) \geq \epsilon$ for every $x \in C_G(\phi)$ and for each $\phi \in A^{\#}$. Then there is a (p,ϵ)-bounded number e and an A-invariant normal subgroup T such that the index $[G:T]$ is (p,ϵ)-bounded and $[G^e, T] = 1$.*

Proof. Let $A_1, \ldots, A_{p+1}$ be the subgroups of order p in A and set $G_i = C_G(A_i)$ for $i = 1, \ldots, p+1$. According to Theorem B there is an ϵ-bounded number d and, for $i = 1, \ldots, p+1$, there are subgroups $H_i \leq G$ such that the index $[G : H_i]$ is ϵ-bounded and $[G_i^d, H_i] = 1$. If $\phi \in A$, let H_i^ϕ be the image of H_i under ϕ. Replace H_i with $T_i = \bigcap_{\phi \in A} H_i^\phi$ and note that T_i is A-invariant, the index $[G : T_i]$ is (p,ϵ)-bounded and $[G_i^d, T_i] = 1$.

Now, let $T = \bigcap T_i$ and observe that T also is A-invariant, the index $[G:T]$ is (p,ϵ)-bounded and $[G_i^d, T] = 1$ for every $i = 1, \ldots, p+1$. If $C = C_G(T)$, then the subgroup C is A-invariant and $G_i^d \leq C$ for $i = 1, \ldots, p+1$. Hence, $C_{G/C}(A_i)$ has exponent dividing d for each $i = 1, \ldots, p+1$, by Lemma 2.3.2. Theorem 4.3.14 says that the exponent of G/C is (d,p)-bounded. Therefore there exists a (p,ϵ)-bounded number e such that $G^e \leq C$, that is, $[G^e, T] = 1$. This completes the proof. □

The situation considered in Theorem 4.3.15 has the following stronger implication.

Corollary 4.3.16. *Let $\epsilon > 0$ and let G be a profinite group admitting an elementary abelian coprime group of automorphisms A of order p^2. Suppose that $Pr(\langle x\rangle, G) \geq \epsilon$ for every $x \in C_G(\phi)$ and for each $\phi \in A^{\#}$. Then there exists a number $\epsilon_0 > 0$ depending only on ϵ and p such that $Pr(\langle x\rangle, G) \geq \epsilon_0$ for every $x \in G$.*

Proof. By Theorem 4.3.15, there is a (p,ϵ)-bounded number e and a normal subgroup T such that the index $[G:T]$ is (p,ϵ)-bounded and $[G^e,T]=1$. Then we can apply Proposition 4.1.2 to obtain the desired conclusion. □

In the spirit of the work [20] we record the following theorem.

Theorem 4.3.17. *Let G be a finite group admitting an elementary abelian coprime group of automorphisms A of order p^3 such that the exponent commutator subgroup of $C_G(\phi)$ divides d for every $\phi \in A^{\#}$. Then there is a (d,p)-bounded number e such that $[G,G]^e=1$.*

Once again, Lemma 2.3.2 allows us to derive the next theorem.

Theorem 4.3.18. *Let G be a profinite group admitting an elementary abelian coprime group of automorphisms A of rank three such that the commutator subgroup of $C_G(\phi)$ has exponent dividing d for every $\phi \in A^{\#}$. Then there is a (d,p)-bounded number e such that $[G,G]^e=1$.*

This result also admits a probabilistic version.

Theorem 4.3.19. *Let $\epsilon > 0$ and let G be a finite group admitting an elementary abelian coprime group of automorphisms A of order p^3. Suppose that $Pr(\langle x\rangle, G) \geq \epsilon$ for every $x \in C_G(\phi)'$ and for each $\phi \in A^{\#}$. Then there is a (p,ϵ)-bounded number e and an A-invariant normal subgroup T such that the index $[G:T]$ is (p,ϵ)-bounded and $[[G,G]^e,T]=1$.*

Proof. Let $A_1,\dots,A_s$ be the subgroups of order p of A and let D_i denote the commutator subgroup of $C_G(A_i)$ for $i=1,\dots,s$. According to Theorem B there is an ϵ-bounded number d and, for $i=1,\dots,s$, normal open subgroups $H_i \leq G$ such that the index $[G:H_i]$ is ϵ-bounded and $[D_i^d,T_i]=1$. Set $T_i=\bigcap_{\phi\in A} H_i^{\phi}$ and observe that T_i is A-invariant and $[G:T_i]$ is (p,ϵ)-bounded.

Now let $T=\bigcap T_i$ and see that $[G:T]$ is (p,ϵ)-bounded and $[G_i^d,T]=1$ for $i=1,\dots,p+1$. Denote by C the centralizer $C_G(T)$, which is A-invariant and contains D_i^d for $i=1,\dots,s$. Hence, $C_{G/C}(A_i)$ has commutator subgroup of exponent dividing d for each $i=1,\dots,s$, by Lemma 2.3.2. Theorem 4.3.18 says that the exponent of the commutator subgroup of G/C is (d,p)-bounded. Therefore there exists a (p,ϵ)-bounded number e such that $[G,G]^e \leq C$, that is, $[[G,G]^e,T]=1$. This completes the proof. □

Chapter 5

Positive commuting probability of monothetic subgroups

This chapter deals with a non-quantitative version of Theorem B, so this discussion would not be interesting for finite groups, for example, and G always denotes a compact group here. We treat groups G having a subgroup K such that $Pr(\langle x\rangle, G) > 0$ for any $x \in K$. Note that this condition is satisfied whenever the subgroup K is torsion, and, more generally, whenever the image of K in $G/Z(G)$ is torsion. Theorem C shows that an "almost converse" to the latter statement also holds:

Theorem C. *Let K be a subgroup of a compact group G. Then $Pr(\langle x\rangle, G) > 0$ for any $x \in K$ if and only if G has an open normal subgroup T such that $K/C_K(T)$ is torsion.*

The techniques employed in the proof of Theorem C are different from the ones in the previous chapter. This was published in [4].

5.1 Proof of Theorem C

Recall that $FC(G)$ denotes the set of all elements x of a group G such that $C_G(x)$ has finite index. Note that in general $FC(G)$ is an abstract subgroup which is not necessarily closed in G. If an element x belongs to $FC(G)$ we call x an *FC-element*, and a subgroup of $FC(G)$ is an FC-subgroup of G. We need the following theorem, which is [44, Theorem 14.5.9]. Here, G does not need to be compact.

Theorem 5.1.1. *If G is an FC-group, then G' is torsion.*

Let G be a group and let $x \in G$. If $Pr(\langle x\rangle, G) > 0$, then, by Theorem A, some power x^e of x is an FC-element of G. Indeed, Theorem A tells us that there exist open subgroups T of G and B of $\langle x\rangle$ such that $[T, B]$ is finite. In particular, B contains x^e for some positive integer e, and x^e is an FC-element of G. It turns out that, under the same hypothesis, a stronger result holds.

Lemma 5.1.2. *Let G be a compact group and $x \in G$ an element such that $Pr(\langle x\rangle, G) > 0$. Then there is a positive integer e such that x^e is contained in an abstract torsion-free abelian normal subgroup N of G such that $N \subseteq FC(G)$.*

Proof. Since $Pr(\langle x\rangle, G) > 0$, an application of Theorem A gives that x^{e_1} has finitely many conjugates, for some positive integer e_1. Let $\{y_1, y_2, \dots, y_k\}$ be the conjugacy class of x^{e_1} and let L be the abstract subgroup generated by $y_1, y_2, \dots, y_k$. Since $[G : C_G(y_i)] = k$ for $i = 1, 2, \dots, k$, it follows that the intersection $\bigcap_{i=1}^{k} C_L(y_i) = Z(L)$ has index at most k^k in L. Set $e_2 = k^k$ and note that the power $x^{e_1 e_2}$ belongs to $Z(L)$. Let M be the abstract subgroup generated by $y_1^{e_2}, y_2^{e_2}, \dots, y_k^{e_2}$. Obviously, M is a finitely generated abelian group. Thus, we conclude that the torsion part of M is finite and denote by e_3 the exponent of the torsion subgroup of M. Write $e = e_1 e_2 e_3$ and observe that the minimal abstract normal subgroup N containing x^e is torsion-free, abelian, and normal. Since N is generated by finitely many FC-elements, it follows that $N \subseteq FC(G)$. The proof is complete. □

We will now prove Theorem C, which we restate here for the reader's convenience.

Theorem C. *Let K be a subgroup of a compact group G. Then $Pr(\langle x\rangle, G) > 0$ for any $x \in K$ if and only if G has an open normal subgroup T such that $K/C_K(T)$ is torsion.*

Proof of Theorem C. Suppose first that G has an open normal subgroup T, say of index i, such that $K/C_K(T)$ is torsion. Choose $x \in K$ and note that the probability that a random element of $\langle x\rangle$ centralizes T is at most $\frac{1}{j}$, where j is the order of the image of $\langle x\rangle$ in $K/C_K(T)$. On the other hand, the probability that a random element of G belongs to T is $\frac{1}{i}$ and so we deduce that $Pr(\langle x\rangle, G) \geq \frac{1}{ij}$. In particular, $Pr(\langle x\rangle, G) > 0$ for every $x \in K$.

We now need to prove the other part of the theorem, that is, assuming that G is a compact group having a subgroup K such that $Pr(\langle x\rangle, G) > 0$ for every $x \in K$ we need to show that G has an open normal subgroup T such that $K/C_K(T)$ is torsion.

For an element $x \in K$ write $e(x)$ to denote the least positive integer e satisfying the conclusion of Lemma 5.1.2. Also, define A_x as the minimal abstract normal subgroup containing $x^{e(x)}$. Thus, A_x is torsion-free and abelian.

Let x and y be arbitrary elements of K. The product A_xA_y is an abstract FC-subgroup of G, and by Theorem 5.1.1 the commutator subgroup of A_xA_y is torsion. On the other hand, the commutator subgroup of A_xA_y is contained in the intersection $A_x \cap A_y$, which is torsion-free. It follows that A_xA_y is abelian. Since this happens for any $x, y \in K$, we conclude that the product $\prod_{x\in K} A_x$ is abelian.

Let N be the topological closure of the abstract subgroup $\prod_{x\in K} A_x$. We see that N is an abelian normal subgroup of G and KN/N is torsion. Set $M = N \cap K$ and, for positive integers k, s, define

$$M_{k,s} = \{x \in M \mid x^k \text{ has at most } s \text{ conjugates in } G\}.$$

Lemma 5.1.2 shows that the sets $M_{k,s}$ cover M. We also note that the sets $M_{k,s}$ are closed. The Baire Category Theorem, Theorem 2.2.6, ensures that at least one of the above sets has non-empty interior. Therefore, for some positive integers k, s, there is an open subset V of M such that x^k has at most s conjugates for every $x \in V$. As M is compact, there is a finite subcover $M = \bigcup_{i=1}^{n} x_iV$ of the cover $M = \bigcup_{x\in M} xV$. Moreover, by Theorem A, for each $i = 1, 2, \ldots, n$ there exist positive integers l_i and m_i such that $[G : C_G(x_i^{l_i})] = m_i$. Let $l = kl_1l_2 \cdots l_n$ and $m = max\{m_1, m_2 \ldots, m_n\}$. Taking into account that M is abelian we deduce that any element of M^l has at most ms conjugates in G. Therefore $Pr(M^l, G) \geq \frac{1}{ms}$ and, by Theorem A, there is an open normal subgroup $T \leq G$ and an open subgroup $B \leq M^l$ such that $[T, B]$ is finite.

Suppose that the order of $[T, B]$ is f, let $t \in T$, $b \in B$ and consider the commutators $[t, b^r]$ for $r = 0, 1, \ldots, f$. The pigeonhole principle ensures that for two exponents $u > v$ the commutators must coincide, and this implies that b^{u-v} centralizes t. Since $1 \leq u - v \leq f$, we deduce that $B^{f!}$ centralizes T. In the series

$$K \geq M \geq B \geq B^{f!} \geq 1$$

the quotients, K/M, M/B and $B/B^{f!}$ are torsion. We conclude that the quotient $K/B^{f!}$ is torsion so the proof is complete. □

5.2 Applications of Theorem C

Here we collect corollaries of Theorem C. Some of them can be regarded as non-quantitative versions of the corollaries of Theorem B, roughly saying.

5.2.1 General Corollaries

An immediate corollary of Theorem C is the following.

Corollary 5.2.1. *If G is a compact group such that $Pr(\langle x\rangle, G) > 0$ for any $x \in G$, then G is virtually central-by-torsion, that is, there is an open normal subgroup T of G such that $T/Z(T)$ is torsion.*

Recall that G_0 stands for the connected component of identity of G, i.e., the largest connected subgroup of G.

Corollary 5.2.2. *Let G be a compact group such that $Pr(\langle x\rangle, G) > 0$ for every $x \in G_0$. Then G has a normal profinite subgroup Δ such that $G_0\Delta$ is open in G and $G_0 \leq Z(G_0\Delta)$.*

Proof. By Theorem C, the group G has an open normal subgroup T such that $G_0/C_{G_0}(T)$ is torsion. Since G_0 is connected and divisible, we deduce that $G_0 \leq Z(T)$. The structure of compact groups in which the identity component is central is determined in Proposition 2.2.13. Therefore $T = G_0D$, where D is a profinite subgroup that is normal in T. Set $\Delta = \prod_{x\in G} D^x$ and observe that $G_0\Delta = T$ is open in G and $G_0 \leq Z(G_0\Delta)$. The proof is complete. □

5.2.2 On the p-structure of a profinite group

In this subsection G denotes a profinite group. Let w be a multilinear commutator word. Recall that we denote by G_w the set of all values of w on elements of G and by $w(G) = \langle G_w\rangle$ the corresponding (closed) verbal subgroup. Let P be a subgroup of G. Following [32] we denote by $W_G(P)$ the subgroup of P generated by all elements of P that are conjugate in G to w-values on elements of P, that is, $W_G(P) = \langle P_w^G \cap P\rangle$. When it causes no confusion, we write $W(P)$ in place of $W_G(P)$. We need the following results from [32].

Lemma 5.2.3. *Let G be a profinite group and let P be a Sylow p-subgroup of G such that $W_G(P)$ is torsion. It follows that*

1. *The non-p-soluble length $\lambda_p(G)$ is finite;*
2. *If G is pro-p-soluble, then the p-length $l_p(G)$ is finite.*

In the case where $w = x$ and p is odd the above results were obtained in Wilson [50]. Furthermore, we also need the following lemma from [32].

Lemma 5.2.4. *Let w be a multilinear commutator, p a prime, and P a Sylow p-subgroup of a profinite group G.*

1. *If $G_1 \leq G$ and $P_1 \leq P$, then $W_{G_1}(P_1) \leq W_G(P)$;*
2. *If N is a normal subgroup of G, then $W_{G/N}(PN/N) = W_G(P)N/N$.*

We can now establish the following theorem.

Theorem 5.2.5. *Let w be a multilinear commutator, p a prime, and P a Sylow p-subgroup of a profinite group G such that $Pr(\langle x\rangle, G) > 0$ for every $x \in W_G(P)$. Then $\lambda_p(G) < \infty$. If G is pro-p-soluble, then $l_p(G) < \infty$.*

Proof. Set $K = W_G(P)$. By Theorem C the group G contains an open normal subgroup T such that $K/C_K(T)$ is torsion. Let $Z = Z(T)$. Obviously, the image of K in G/Z is torsion. In view of Lemma 5.2.4 the aforementioned results from [32] imply that

1. $\lambda_p(G/Z)$ is finite;
2. If G/Z is pro-p-soluble, then $l_p(G/Z)$ is finite.

Since the subgroup Z is abelian, the theorem follows. □

5.2.3 Centralizers of coprime automorphisms of profinite groups

Recall that $A^{\#}$ stands for the set of nontrivial elements of a group A. We first restate Theorems 4.3.14 and 4.3.18.

Remark 5.2.6. Let G be a profinite group admitting an elementary abelian coprime group of automorphisms A.

- **Theorem 4.3.14:** Suppose that A has order p^2. If $C_G(\phi)$ has exponent dividing d for each $\phi \in A^{\#}$, then the exponent of G is (d,p)-bounded.
- **Theorem 4.3.18:** Suppose that A has order p^3. If the commutator subgroup of $C_G(\phi)$ has exponent dividing d for each $\phi \in A^{\#}$, then the exponent of $[G,G]$ is (d,p)-bounded.

Theorems 4.3.14 and 4.3.18 have non-quantitative profinite variations, that were obtained in [45] and [1], respectively.

Theorem 5.2.7. *Let p be a prime and G a pro-p' group admitting an elementary abelian p-group of automorphisms A.*

1. *If A is noncyclic and $C_G(\phi)$ is torsion for all $\phi \in A^{\#}$, then G is torsion.*
2. *If A is of rank at least three and the commutator subgroup $C_G(\phi)'$ is torsion for all $\phi \in A^{\#}$, then G' is torsion.*

Now we are able to supply probabilistic variants:

Theorem 5.2.8. *Let p be a prime and G a pro-p' group admitting an elementary abelian p-group of automorphisms A.*

1. *If A is noncyclic and $Pr(\langle x\rangle, G) > 0$ for every $x \in C_G(\phi)$ and $\phi \in A^{\#}$, then G is virtually central-by-torsion.*
2. *If A is of rank at least three and $Pr(\langle x\rangle, G) > 0$ for every $x \in C_G(\phi)'$ and $\phi \in A^{\#}$, then G' is virtually central-by-torsion.*

Proof. We prove part (1). Assume that A is of rank 2 and $Pr(\langle x\rangle, G) > 0$ for every $x \in C_G(\phi)$ and $\phi \in A^{\#}$. Let $A_1, \ldots, A_{p+1}$ be the maximal subgroups of A and write K_i for $C_G(A_i)$. Applying Theorem C conclude that there are open normal subgroups $T_1, \ldots, T_{p+1}$, such that $K_i/C_{K_i}(T_i)$ is torsion for every $i = 1, \ldots, p+1$. Let $T = \bigcap_{i=1}^{p+1} T_i$ and $Z = Z(T)$. Then, by Lemma 2.3.2, the centralizers $C_{G/Z}(A_i)$ are torsion. In view of Theorem 5.2.7 (1) we conclude that G/Z is torsion, as required.

The proof of Part (2) is similar to the above. Assume that the rank of A is at least 3 and write K_i for for the commutator subgroup $C_G(A_i)'$, where $A_1, \ldots, A_s$ are the maximal subgroups of A. Then proceed as above and use Theorem 5.2.7 (2) in place of Theorem 5.2.7 (1). □

5.3 Connection with high commuting probability of monothetics

It is natural to ask what is the connection between the properties $Pr(\langle x\rangle, G) > 0$ for all $x \in G$ and $Pr(\langle x\rangle, G) \geq \epsilon$ for a fixed $\epsilon > 0$ and all $x \in G$, since clearly the latter implies the former. However, this connection is related to a nonsolved

problem in Group Theory: the problem whether compact torsion groups have finite exponent, which remains open for many years (cf [25, p. 70]). We remark that if indeed compact torsion groups have finite exponent, then under the hypotheses of Theorem C there is $\epsilon > 0$ such that $Pr(\langle x \rangle, G) \geq \epsilon$ for any $x \in K$ and so there is a number e such that $[K^e, T] = 1$.

Proposition 5.3.1. *The following statements are equivalent.*

(i) Any compact torsion group has finite exponent.

(ii) If G is a compact group such that $Pr(\langle x \rangle, G) > 0$ for every $x \in G$, then $Pr(\langle x \rangle, G) \geq \epsilon$ for some $\epsilon > 0$.

Proof. Assume that (i) is true, let G be a compact group and suppose that $Pr(\langle x \rangle, G) > 0$ for every $x \in G$. There exists an open normal subgroup T of G such that $G/C_G(T)$ is torsion. Since (i) is true, then $G/C_G(T)$ has finite exponent, say e, what in turn implies that $[G^e, T] = 1$. Applying the converse of Theorem B, Proposition 4.1.2, we get that $Pr(\langle x \rangle, G) \geq \epsilon$ for $\epsilon = ([G : T]e)^{-1}$.

Conversely, assume (ii) to be true, let G be a compact torsion group and see that $Pr(\langle x \rangle, G) > 0$ for every $x \in G$. Then, since (ii) is true, there exists a positive ϵ such that $Pr(\langle x \rangle, G) \geq \epsilon$. By Theorem B, there is some positive integer e and an open subgroup T of G such that $[G^e, T] = 1$. Let f be a positive integer such that $G^f \leq G^e \cap T$ and observe that G^f is abelian. Considering the series

$$1 \leq G^f \leq G,$$

the factor G/G^f has finite exponent and G^f has finite exponent. Indeed, if $T_n = \{x \in G^f \mid x^n = 1\}$, then T_n is closed and $\bigcup_n T_n$ covers G^f, which is compact. Baire Category Theorem 2.2.6 ensures that there is some open coset xU contained in T_k for some k, and in particular $x^k = 1$. Suppose that the index $[G^f : U]$ is s. Then $g^{sf} \in U$ for every $g \in G$, thus $(xg^{sf})^k = 1$ and $g^{sfk} = 1$ as G^f is abelian. This amounts to G having finite exponent, as we wanted to prove. □

In some very particular cases, the equivalence between $Pr(\langle x \rangle, G) \geq \epsilon$ and property $Pr(\langle x \rangle, G) > 0$ for all $x \in G$ holds, as we prove next.

Corollary 5.3.2. *Let G be a compact group such that $Pr(\langle x\rangle, G) > 0$ for every $x \in G$. Assume further that G is finitely generated. Then there exists a positive ϵ such that $Pr(\langle x\rangle, G) \geq \epsilon$ for every $x \in G$.*

Proof. By Theorem C, there is an open normal subgroup T of G such that $G/C_G(T)$ is torsion. Being G finitely generated, we can apply Theorem 4.3.1 to deduce that $C_G(T)$ must be open in G. Observe that $[C_G(T), T] = 1$, where both $C_G(T)$ and T are open in G. The converse of Theorem B, Proposition 4.1.2, implies that there is some $\epsilon > 0$ such that $Pr(\langle x\rangle, G) \geq \epsilon$ for every $x \in G$. □

Moreover, if we assume that $Pr(\langle x\rangle, G) > 0$ for every x in the connected component of identity G_0 of G, then it is true that G_0 has high commuting probability on monothetic subgroups.

Corollary 5.3.3. *Whenever G is a compact group such that $Pr(\langle x\rangle, G) > 0$ for every $x \in G_0$, there is $\epsilon > 0$ with the property that $Pr(\langle x\rangle, G) \geq \epsilon$ for every $x \in G_0$.*

Proof. Theorem C implies that there exists an open normal subgroup T such that $G_0/C_{G_0}(T)$ is torsion. However, G_0 is divisible [37] and $G_0/C_{G_0}(T)$ must be trivial. It follows that the open subgroup T centralizes G_0 and the result follows from Corollary 4.2.5 (ii). □

Chapter 6

Appendix: Existence of Haar measure

In this chapter we give a proof of the existence of a translation-invariant regular measure in locally compact groups. The one we give here differs from the average proof found in textbooks like [24, 38] since it avoids formally functional analysis and its terminology.

6.1 Preliminaries

First we need to record some topological and measure-theoretic preliminaries, the first one being Tychonoff's Theorem [30, p. 143]. It is important to remark that Tychonoff's theorem is equivalent to the Axiom of Choice, as it is stated in [29]. A constructive proof of existence of Haar measure that does not rely on the Axiom of Choice can be found in [7].

Theorem 6.1.1. *Let I be any set and K_i be a compact topological space for every $i \in I$. Then the product $\prod_{i \in I} K_i$ is compact under the product topology.*

Next we define an important property in the context of compact topological spaces.

Definition 6.1.2 (Finite Intersection Property)**.** Let X be any set and let $\mathcal{C}$ be a collection of subsets of X. We say that $\mathcal{C}$ has the *finite intersection property* if every finite subcolletion of $\mathcal{C}$ has nonempty intersection.

Compact topological spaces also have a characterization in terms of the finite intersection property, as it is stated in the next result.

Proposition 6.1.3. *Let X be a topological space. Then X is compact if and only if every collection of closed subsets of X having the finite intersection property has nonempty intersection.*

We need a further topological result. Here we do not need G to be locally compact, but we need the group to be Hausdorff.

Lemma 6.1.4. *Let G be a Hausdorff topological group and let K_1 and K_2 be disjoint compact subsets of G. Then there exists an open neighbourhood of identity O such that $K_1O^{-1} \cap K_2O^{-1} = \emptyset$.*

Proof. Continuity of the group operation and inversion imply that the set $K_1^{-1}K_2$ is compact, and since G is Hausdorff, it follows that $K_1^{-1}K_2$ is closed in G. The subset $U = G\backslash K_1^{-1}K_2$ is open in G and contains 1, since K_1 and K_2 are disjoint. Again, continuity of the group operation and inversion imply that there is an open neighbourhood O of 1 such that $O = O^{-1}$ and $OO \subseteq U$. We claim that $K_1O^{-1} \cap K_2O^{-1} = \emptyset$. To see this, assume that $K_1O^{-1} \cap K_2O^{-1} \neq \emptyset$ or, equivalently, that $K_1 \cap K_2O^{-1}O \neq \emptyset$. There exist $k_1 \in K_1$, $k_2 \in K_2$ and $v_1, v_2 \in O$ such that $k_1 = k_2v_1^{-1}v_2$, and it follows that $k_1^{-1}k_2 = v_2^{-1}v_1$. This element lies simultaneously in $K_1^{-1}K_2$ and OO, which is contained in U, a contradiction. □

Recall that $\mathcal{P}(X)$ denotes the collection of all subsets of a set X. Next we define a set function that generalizes the notion of a measure.

Definition 6.1.5. Let X be a set. An *outer measure* on X is a set function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ such that the following hold.

(i) $\mu^*(\emptyset) = 0$ and $\mu^*(A) \geq 0$ for every $A \subseteq X$,

(ii) If $A \subseteq B \subseteq X$, then $\mu^*(A) \leq \mu^*(B)$, and

(iii) If $\{A_i\}_{i=1}^{\infty}$ is a countable collection of subsets of X, then

$$\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu^*(A_i).$$

Property (iii) is called *countable subadditivity*. If one chooses on X a σ-algebra that is different from the set of all subsets of X, a measure will fail to be an outer measure, and an outer measure can fail to be countably additive and hence fail to be a measure. Furthermore, it is natural to pass from an outer measure to a measure. To do so, it suffices to restrict the outer measure to a specific σ-algebra consisting of sets having the following property.

Definition 6.1.6. Let X be a set and μ^* be an outer measure on X. A subset B of X is said to be μ^*-measurable if, for every $A \subseteq X$, we have

$$\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c)$$

Since the inequality $\mu^*(A) \leq \mu^*(A \cap B) + \mu^*(A \cap B^c)$ always holds, as μ^* is a subadditive set function, to prove that a set B is μ^*-measurable, it suffices to prove the reverse inequality. As it was said before, there is a natural way of passing from an outer measure to a measure, as it is stated in the next result. It is [3, Theorem 1.3.5].

Proposition 6.1.7. *Let X be a set, let μ^* be an outer measure on it and let M be the collection of all μ^*-measurable subsets of X. Then*

(i) M is a σ-algebra on X, and

(ii) The restriction of μ^* to M is a measure on M.

6.2 Existence of Haar measure

Let G be a locally compact topological group, fixed for the rest of the section, with Borel σ-algebra $\mathcal{B}(G)$. We want to define a nonzero measure μ on $\mathcal{B}(G)$ satisfying the following properties:

(H1) $\mu(S) = \mu(xS)$ for every $x \in G$ and $S \in \mathcal{B}(G)$;

(H2) $\mu(K)$ is finite if K is compact;

(H3) $\mu(S) = \inf\{\mu(U) \,|\, S \subseteq U, U \text{ is open}\}$ for evey $S \in \mathcal{B}(G)$;

(H4) $\mu(U) = \sup\{\mu(K) \,|\, K \subseteq U, K \text{ compact}\}$ for every open subset U of G.

We fix more notation. Let $\mathcal{C}$ be the set of all compact subsets of G and $\mathcal{U}$ be the collection of all open neighbourhoods of identity. Also fix a compact subset K_0 of G having nonempty interior. Such a subset exists since G is locally compact and every point x is contained in an open set which, in turn, is contained in a compact set.

If K is a compact subset of G and U is a set with nonempty interior, let $(K : U)$ denote the smallest n such that there exist $x_1, x_2, \ldots, x_n \in G$ with $K \subseteq \bigcup_{i=1}^n x_i U$. We call $(K : U)$ the *index* of U in K. Note that $(K : U)$ is always finite since K is compact and U has nonempty interior, and $(K : U) = 0$

if and only if $K = \emptyset$. The index $(K : U)$ has the following properties.

Proposition 6.2.1. *Let K_0 be as before, $K, K_1, K_2 \in \mathcal{C}$ and let U be a fixed set with nonempty interior. Then*

(i) $0 \leq (K : U) \leq (K : K_0) \cdot (K_0 : U)$;

(ii) $(xK : U) = (K : U)$ *for every* $x \in G$;

(iii) If $K_1 \subseteq K_2$, *then* $(K_1 : U) \leq (K_2 : U)$;

(iv) $(K_1 \cup K_2 : U) \leq (K_1 : U) + (K_2 : U)$;

(v) If $K_1 U^{-1} \cap K_2 U^{-1} = \emptyset$, *then* $(K_1 \cup K_2 : U) = (K_1 : U) + (K_2 : U)$.

Proof.

(i) Let $n = (K : K_0)$, $m = (K_0 : U)$ and suppose that $K \subseteq \bigcup_{i=1}^{n} x_i K_0$ and $K_0 \subseteq \bigcup_{j=1}^{m} y_j U$. Conclude that $K \subseteq \bigcup_{i=1}^{n} \bigcup_{j=1}^{m} x_i y_j U$ and $(K : U) \leq mn$.

(ii) See that $K \subseteq \bigcup_{i=1}^{n} x_i U$ if and only if $xK \subseteq \bigcup_{i=1}^{n} x x_i U$.

(iii) Observe that every cover of K_2 by left-translates of U also covers K_1, hence the result.

(iv) If $K_1 \subseteq \bigcup_{i=1}^{m} x_i U$ and $K_2 \subseteq \bigcup_{j=1}^{n} y_j U$, then $K_1 \cup K_2$ is contained in $(\bigcup_{i=1}^{m} x_i U) \cup \left(\bigcup_{j=1}^{n} y_i U\right)$.

(v) Let $x \in G$ be arbitrary. Observe that if $xU \cap K_1 \neq \emptyset$ and $xU \cap K_2 \neq \emptyset$, then there exist $u_1, u_2 \in U$, $k_1 \in K_1$ and $k_2 \in K_2$ such that $xu_1 = k_1$ and $xu_2 = k_2$. It follows that $x = k_1 u_1^{-1} = k_2 u_2^{-1}$ belongs to $K_1 U^{-1} \cap K_2 U^{-1}$, a contradiction. Then every left translate xU of U must intercept at most one of the subsets K_1 or K_2, and the result follows.

□

Now, for every $U \in \mathcal{U}$, we define $h_U : \mathcal{C} \to [0, \infty)$ as

$$h_U(K) = \frac{(K : U)}{(K_0 : U)}.$$

Proposition 6.2.1 implies that h_U has the following properties.

Proposition 6.2.2. *Let $K, K_1, K_2 \in \mathcal{C}$ and K_0 as before. Then*

(i) $0 \leq h_U(K) \leq (K : K_0)$ *and* $h_U(K_0) = 1$*;*

(ii) $h_U(xK) = h_U(K)$ *for every* $x \in G$*;*

(iii) If $K_1 \subseteq K_2$*, then* $h_U(K_1) \leq h_U(K_2)$*;*

(iv) $h_U(K_1 \cup K_2) \leq h_U(K_1) + h_U(K_2)$*;*

(v) If $K_1 U^{-1} \cap K_2 U^{-1} = \emptyset$*, then* $h_U(K_1 \cup K_2) = h_U(K_1) + h_U(K_2)$*.*

Next, for every $K \in \mathcal{C}$, consider the compact interval $[0, (K : K_0)] \subseteq \mathbb{R}$ and let

$$X = \prod_{K \in \mathcal{C}} [0, (K : K_0)].$$

Here, X can be seen as a space of functions from $\mathcal{C}$ to $\mathbb{R}$, where for each function $f \in X$ and $K \in \mathcal{C}$ we have $f(K) \leq (K : K_0)$. Applying Theorem 6.1.1 we deduce that X is compact under the product topology. Furthermore, for every $U \in \mathcal{U}$, we have that h_U belongs to X.

For each open neighbourhood U of unit, define the set $T_U = \{h_V \,|\, V \subseteq U\}$ and let $V(U)$ be the closure of T_U on X. If $U_1, U_2, \ldots, U_n \in \mathcal{U}$, let $V = \bigcap_{i=1}^{n} U_i$ and see that $h_V \in T_{U_i}$ for $i = 1, 2, \ldots, n$. We deduce that the collection of sets $\{V(U) \,|\, U \in \mathcal{U}\}$ has the finite intersection property and thus has nonempty intersection by Proposition 6.1.3. Let h be an element in this intersection. Before we pass to the properties of h we state two results.

Lemma 6.2.3. *Let K be a fixed compact subset of G and let $\phi : X \to \mathbb{R}$ be $\phi(f) = f(K)$. Then ϕ is continuous.*

Proof. See that $f(K) \leq (K : K_0)$. For a given $\epsilon > 0$, the interval

$$S_K = (f(K) - \epsilon, f(K) + \epsilon) \cap [0, (K : K_0)]$$

is an open subset of $[0, (K : K_0)]$. Consider $Y = \prod_{K \in \mathcal{C}} S_K$ where

$$S_{K'} = [0, (K' : K_0)]$$

for $K' \neq K$ and S_K is as above. Then Y is open on X and every function g of Y satisfies $|\phi(g) - \phi(f)| < \epsilon$. □

Now we can derive a criterion for a subset of X to contain h.

Lemma 6.2.4. *Let C be a closed subset of X containing T_U for some $U \in \mathcal{U}$. Then C contains h.*

Proof. Since C is closed and contains T_U, it also contains the closure $V(U)$. The function h belongs to the intersection of all $V(U)$, so in particular $h \in C$. □

The function $h\colon \mathcal{C} \to [0, \infty)$ has the following properties.

Proposition 6.2.5. *Let $K, K_1, K_2 \in \mathcal{C}$ and K_0 as before. Then*

(i) $h(K)$ is finite and non-negative, $h(\emptyset) = 0$ and $h(K_0) = 1$;

(ii) $h(xK) = h(K)$ for every $x \in G$;

(iii) If $K_1 \subseteq K_2$, then $h(K_1) \leq h(K_2)$;

(iv) $h(K_1 \cup K_2) \leq h(K_1) + h_U(K_2)$;

(v) If K_1 and K_2 are disjoint compact sets, then $h(K_1 \cup K_2) = h(K_1) + h(K_2)$.

Proof. We prove (i) first. To prove that $h(K) \leq (K : K_0)$, consider the subset $A = \{f \in X \mid f(K) \leq (K : K_0)\}$ of X and see that A is closed, by Lemma 6.2.3. Note that, for every $U \in \mathcal{U}$, the function h_U belongs to A. It folows that T_U is contained in A for every U so Lemma 6.2.4 implies that $h \in A$. To verify that h is nonnegative, mimic the argument with respect to the subset $\{f \in X \mid f(K) \geq 0\}$ of X. Furthermore, the properties that $h(\emptyset) = 0$ and $h(K_0) = 1$ follow from the same argument, applied with respect to the closed subsets $\{f \in X \mid f(\emptyset) = 0\}$ and $\{f \in X \mid f(K_0) = 1\}$ of X.

The same arguments used to prove (i) can be used to prove (ii), (iii) and (iv). It suffices to mimic the proof of (i) with respect to the following subsets of X, respectively:

(ii) $\{f \in X \mid f(K) - f(xK) = 0\}$;

(iii) $\{f \in X \mid f(K_2) - f(K_1) \geq 0\}$, and;

(iv) $\{f \in X \mid f(K_1) + f(K_2) - f(K_1 \cup K_2) \geq 0\}$.

Now we proceed to the proof of (v). Let K_1 and K_2 be disjoint compact subsets of G. We need to prove that $h(K_1 \cup K_2) = h(K_1) + h(K_2)$. Let O be a neighbourhood of 1 such that $K_1 O^{-1} \cap K_2 O^{-1} = \emptyset$, by Proposition 6.1.4, and observe that, if $U \subseteq O$ is another open neighbourhood of 1, it follows that

$K_1U^{-1} \cap K_2U^{-1} = \emptyset$. Using Lemma 6.2.5(v), we deduce that T_O is contained in the closed subset of X given by

$$C = \{f \in X \mid f(K_1 \cup K_2) - f(K_1) - f(K_2) = 0\}.$$

Thus $h \in C$, by Lemma 6.2.4. The proof is complete.

□

We now are able to define an outer measure μ^* on G in terms of h. Let U be any open subset of G and define

$$\mu^*(U) = \sup\{h(K) \mid K \subseteq U \text{ and } K \in \mathcal{C}\}.$$

We extend this map to all subsets of A of G by

$$\mu^*(A) = \inf\left\{\sum_{n=1}^{\infty} \mu * (A_n) \mid A \subseteq \bigcup_{i=1}^{\infty} A_n \text{ and } A_n \text{ is open, } n = 1, 2, \dots\right\}$$

We verify that μ^* indeed satisfies Definition 6.1.5.

Proposition 6.2.6. *The set function $\mu^* : \mathcal{P}(G) \to [0, \infty]$ is a translation-invariant nonzero outer measure on G.*

Proof. First of all, since h is non-negative and $h(K_0) = 1$ we conclude that μ^* is non-negative and nonzero. Also, as $h(\emptyset) = 0$, the definition of μ^* implies that $\mu^*(\emptyset) = 0$, and if $A \subseteq B$ are arbitrary subsets of G, the inequality $\mu^*(A) \leq \mu^*(B)$ and the fact that μ^* is translation-invariant both follow from the definition of μ^*. It remains to prove countable subadditivity.

For this reason, let $A_1, A_2, \dots$ be subsets of G. We need to prove that

$$\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu^*(A_i).$$

Define $A = \bigcup_{i=1}^{\infty} A_i$ and see that, if $\mu^*(A_i) = \infty$ for any i, monotonicity of μ^* implies that $\mu^*(A) = \infty$ also. Hence we may assume that $\mu^*(A_i)$ is finite for every i. Let $\epsilon > 0$. The definition of $\mu^*(A_i)$ implies that, for each i there exists an open cover $\bigcup_{n=1}^{\infty} B_{in} \supseteq A_i$ such that

$$\mu^*(A_i) \leq \sum_{n=1}^{\infty} \mu^*(B_{in}) < \mu^*(A_i) + \frac{\epsilon}{2^i}.$$

Now, see that the union $\bigcup_{i=1}^{\infty}\bigcup_{n=1}^{\infty} B_{in}$ is a countable cover of A. The definition of μ^* implies that

$$\mu*(A) \leq \sum_{i=1}^{\infty}\left(\sum_{n=1}^{\infty}\mu^*(B_{in})\right) \leq \sum_{i=1}^{\infty}\mu^*(A_i) + \frac{\epsilon}{2^i} \leq \sum_{i=1}^{\infty}\mu^*(A_i) + \epsilon.$$

As ϵ is arbitrary, we proved that $\mu^*(A) \leq \sum_{i=1}^{\infty}\mu^*(A_i)$. □

Let $A \subseteq G$ be any subset. We defined $\mu^*(A)$ in terms of open covers of A, but there is an equivalent definition in terms of open sets containing A. This is the content of the next result.

Proposition 6.2.7. *For any subset A of G, we have*

$$\mu^*(A) = \inf\{\mu^*(U) \,|\, U \supseteq A \text{ and } U \text{ is open}\}.$$

Proof. Let $\mu_{op}(A) = \inf\{\mu^*(U) \,|\, U \supseteq A \text{ and } U \text{ is open}\}$. Monotonicity of μ^* implies that $\mu^*(A) \leq \mu^*(U)$ for every open subset U of G containing A, so $\mu^*(A) \leq \mu_{op}(A)$. For the reverse inequality, let $\epsilon > 0$. By definition of $\mu^*(A)$, there is an open cover $\bigcup_{n=1}^{\infty} A_n$ of A such that

$$\mu^*(A) \leq \sum_{n=1}^{\infty}\mu^*(A_n) < \mu^*(A) + \epsilon.$$

Observe that $\bigcup_{n=1}^{\infty}$ is an open set containing A. Then

$$\mu_{op}(A) \leq \mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty}\mu^*(A_n),$$

by subadditivity of μ^*. It follows that $\mu_{op}(A) \leq \mu^*(A) + \epsilon$ and, as ϵ is arbitrary, we conclude that $\mu_{op}(A) \leq \mu^*(A)$. □

The collection of μ^*-measurable subsets of G forms a σ-algebra and $\mathcal{B}(G)$ is the smallest σ-algebra containing the open subsets of G. Then, to prove that $\mathcal{B}(G)$ consists on μ^*-measurable sets, it suffices to verify this property for the topology of G. To do so, we first need the following lemma.

Lemma 6.2.8. *Let U and V be open subsets of G. Then*

$$\mu^*(V) = \mu^*(V \cap U) + \mu^*(V \cap U^c).$$

.

Proof. Let $\epsilon > 0$ and choose a compact subset K of $V \cap U$ such that

$$h(K) > \mu^*(V \cap U) - \epsilon.$$

Also, choose a compact subset L of $V \cap K^c$ such that

$$h(L) > \mu^*(V \cap K^c) - \epsilon.$$

Then K and L are disjoint. Moreover, since $V \cap U^c \subseteq V \cap K^c$ and h is monotonic, it follows that $h(L) > \mu^*(V \cap U^c) - \epsilon$. We conclude that

$$h(K \cup L) = h(K) + h(L) \geq \mu^*(V \cap U) + \mu^*(V \cap U^c) - 2\epsilon.$$

Now, as $h(K \cup L) \leq \mu^*(V)$ and taking the limit with $\epsilon \to 0$, we conclude that $\mu^*(V) \geq \mu^*(V \cap U) + \mu^*(V \cap U^c)$. The reverse inequality follows from monotonicity of μ^*. □

With the aid of Lemma 6.2.8 we can prove that all open subset of G are μ^*-measurable.

Proposition 6.2.9. *Let U be an open subset of G and let $A \subseteq G$ be arbitrary. Then $\mu^*(A) \geq \mu^*(A \cap U) + \mu^*(A \cap U^c)$.*

Proof. Let V be an open subset if G containing A. Then

$$\mu^*(V) \geq \mu^*(V \cap U) + \mu^*(V \cap U^c) \geq \mu^*(A \cap U) + \mu^*(A \cap U^c).$$

We conclude that $\mu^*(A \cap U) + \mu^*(A \cap U^c)$ is a lower bound for the set

$$\{\mu^*(V) \,|\, V \supseteq A, V \in \mathcal{U}\}.$$

Definition of $\mu^*(A)$ implies that $\mu^*(A) \geq \mu * (A \cap U) + \mu^*(A \cap U^c)$, as we wanted. □

Proposition 6.2.9 implies that the restriction μ of μ^* to $\mathcal{B}(G)$ is a measure. We claim that μ is the desired Haar measure. Left-invariance of μ^* implies that μ itself is left-invariant, and properties (H3) and (H4) follow from the definition of μ^*. It remains to prove that $\mu(K)$ is finite for every compact subset K of G. To see this, let U be an open subset containing K whose closure $\overline{U}$ is compact. The existence of U is guaranteed because K is compact. Then $\mu(K) \leq \mu(U)$ by monotonicity, and if L is any compact subset of U, then $h(L) \leq h(\overline{U})$. By property (H4) we have $\mu(U) \leq h(\overline{U})$, which is finite. We conclude that $\mu(K)$ is finite, as we wanted to prove.

6.3 Unicity of Haar measure

In this section we reproduce Kakutani's proof [28] of the following result.

Theorem 6.3.1. *Let G be a locally compact topological group and let μ and ν be Haar measures defined on G. Then there is a positive constant c such that $\mu(S) = c\nu(S)$ for every Borel set S of G.*

To prove Theorem 6.3.1 it suffices to prove that

$$\mu(K)\nu(K') = \mu(K')\nu(K) \qquad (\star)$$

for compacts K and K'. Indeed, if $\nu(K')$ is nonzero, then $\mu(K) = \frac{\mu(K')}{\nu(K')}\nu(K)$ for every pair of compact sets. Making $c = \frac{\mu(K')}{\nu(K')}$ for a fixed subset K' of G, we deduce that $\mu(K) = c\nu(K)$ for every compact set $K \subseteq G$. By property (H4), the measure of open subsets of G can be deduced from the measure of compact ones, and it follows from property (H3) that the measure of any Borel set can be calculated in terms of measures of open sets. We therefore conclude that $\mu(S) = c\nu(S)$ for every measurable subset S of G.

Let A and B be measurable subsets of G, equip $G \times G$ with the product measure $\mu \times \nu$ and recall that $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$. Also, recall that χ_A denotes the characteristic function of A. Moreover, note $\chi_{A\times B}(x, y) = \chi_A(x)\chi_B(y)$. We deduce equation $(\star)$ as a consequence of an inequality, which is proved in the next lemma.

Lemma 6.3.2. *Let K and K' be compact subsets of G and let O be an open subset of G. Then $\mu(K)\nu(K') \leq \mu(K'O^{-1})\nu(KO)$.*

Proof. First of all, see that

$$\begin{aligned}\mu(K)\nu(O) &= \iint_{G\times G} \chi_K(x)\chi_O(y)d\mu(x)d\nu(y) \\ &= \iint_{G\times G} \chi_K(x)\chi_O(x^{-1}y)d\mu(x)d\nu(y),\end{aligned}$$

since ν is left-invariant. Now observe that $\chi_K(x)\chi_O(x^{-1}y) = 1$ implies that $x^{-1}y \in O$ and $y \in KO$. Moreover, we remark that $\chi_O(x^{-1}y) = \chi_{O^{-1}}(y^{-1}x)$.

Following the calculations above, we have

$$\begin{aligned}\iint\limits_{G\times G} \chi_K(x)\chi_O(x^{-1}y)d\mu(x)d\nu(y) &\leq \iint\limits_{G\times G} \chi_{KO}(y)\chi_O(x^{-1}y)d\mu(x)d\nu(y)\\ &= \iint\limits_{G\times G} \chi_{KO}(y)\chi_{O^{-1}}(y^{-1}x)d\mu(x)d\nu(y)\\ &= \iint\limits_{G\times G} \chi_{KO}(y)\chi_{O^{-1}}(x)d\mu(x)d\nu(y) \quad \text{(left-invariance)}\\ &= \mu(O^{-1})\nu(KO).\end{aligned}$$

So far, we have proved the following inequality

$$\mu(K)\nu(O) \leq \mu(O^{-1})\nu(KO). \tag{1}$$

Replacing μ, ν, K and O above by ν, μ, K' and O^{-1}, respectively, we obtain

$$\nu(K')\mu(O^{-1}) \leq \mu(K'O^{-1})\nu(O), \tag{2}$$

where K' is an arbitrary compact set. If μ is nonzero, the measure $\mu(O)$ of an open set can never be zero, by (H3) and (H4). Hence, multiplying inequalities (1) and (2) we obtain

$$\mu(K)\nu(K') \leq \mu(K'O^{-1})\nu(KO) \tag{3}$$

for any compact sets K and K' and open subset O. □

To finish the proof we need one further lemma.

Lemma 6.3.3. *Let K and K' be compact subsets of G and let U and U' be open sets such that $U \supseteq K$ and $U' \supseteq K'$. Then there is an open neighbourhood of identity O such that $KO \subseteq U$ and $K'O^{-1} \subseteq U'$.*

Proof. Let $x \in K$ be arbitrary. Since $x = x\cdot 1$, continuity of the group operation ensures that there exist neighbourhoods U_x of x and V_x of 1 such that $U_xV_x \subseteq U$. Then $\bigcup_{x\in K} U_x$ is an open cover of K, from which we extract the finite subcover $\bigcup_{i=1}^n U_{x_i}$. Let $V = \bigcap_{i=1}^n V_{x_i}$ and see that V is nonempty. Then

$$KV \subseteq \left(\bigcup_{i=1}^n U_{x_i}\right)\cdot V \subseteq \bigcap_{i=1}^n U_{x_i}\cdot V_{x_i} \subseteq U.$$

Mimic the above argument with K' in place of K and conclude that there is some open neighbourhood V' of 1 such that $K'V' \subseteq U'$. Let $O = V \cap V^{-1} \cap V' \cap (V')^{-1}$ and observe that $O = O^{-1}$. Then O is a symmetric neighbourhood of identity satisfying the conditions of the statement. □

We can now conclude the proof of equation $(\star)$.

Proof of $(\star)$. Let K and K' be compact subsets of G and let $\epsilon > 0$. By (H4), there are open sets $U \supseteq K$ and $U' \supseteq K'$ such that $\nu(U) < \nu(K) + \epsilon$ and $\mu(U') < \mu(K') + \epsilon$. Let further O be an open neighbourhood of 1 such that $KO \subseteq U$ and $K'O^{-1} \subseteq U'$, whose existence is ensured by Lemma 6.3.3. Then Lemma 6.3.2 implies

$$\mu(K)\nu(K') < (\mu(K') + \epsilon)(\nu(K) + \epsilon).$$

Taking the limit with $\epsilon \to 0$, we get $\mu(K)\nu(K') \leq \mu(K')\nu(K)$. The reverse inequality follows from the same argument, swapping K and K'. □

Bibliography

[1] Acciarri, C., Lima, A., and Shumyatsky, P. Derived subgroups of fixed points in profinite groups. *Glasg. Math. J. 54* (2012), 97–105.

[2] Acciarri, C., and Shumyatsky, P. A stronger form of Neumann's BFC-theorem. *Israel J. Math. 242* (2021), 269 – 278.

[3] Ash, R. B., and Doleans-Dade, C. A. *Probability and Measure Theory*, second ed. Harcourt/Academic Press, 1999.

[4] Azevedo, J., and Shumyatsky, P. Compact groups with high commuting probability of monothetic subgroups. *arXiv e-prints* (Jul. 2022), arXiv:2207.07966.

[5] Azevedo, J., and Shumyatsky, P. Compact groups with probabilistically central monothetic subgroups. *Israel J. Math.* (to appear).

[6] Bourbaki, N. *Integration II, chapters 7-9*. Springer, 2004.

[7] Cartan, H. Sur la mesure de Haar. *Comptes Rendus de l'Académie des Sciences de Paris 211* (1940), 759 – 762.

[8] Detomi, E., Morigi, M., and Shumyatsky, P. BFC-theorems for higher commutator subgroups. *Quarterly J. Math. 70* (2019), 849 – 858.

[9] Detomi, E., and Shumyatsky, P. On the commuting probability for subgroups of a finite group. *Proc. Roy. Soc. Edinburgh Sect. A 152(6)* (2022), 1551 – 1564.

[10] Dierings, G., and Shumyatsky, P. Groups with boundedly finite conjugacy classes of commutators. *Quarterly J. Math. 69 no. 3* (2018), 1047 – 1051.

[11] Eberhard, S. Commuting probabilities of finite groups. *Bull. London Math. Soc. 57* (2015), 796 – 808.

[12] Eberhard, S., and Shumyatsky, P. Probabilistically nilpotent groups of class two. *Math. Ann.* (2023), to appear.

[13] Erdös, P., and Turán, P. On some problems of statistical group theory. *Acta Math. Acad. Sci. Hung. 19* (1968), 413 – 435.

[14] Erfanian, A., Rezaei, R., and Lescot, P. On the relative commutativity degree of a subgroup of a finite group. *Comm. Algebra 35* (2007), 4183 – 4197.

[15] Feit, W., and Thompson, J. G. Solvability of groups of odd order. *Pacific J. Math. 13* (1963), 773 – 1029.

[16] Fumagalli, F., Leinen, F., and Puglisi, O. An upper bound for the nonsolvable length of a finite group in terms of its shortest law. *Proc. London Math. Soc.*, to appear.

[17] Gorenstein, D. *Finite Groups.* Chelsea Publishing Company, New York, 1980.

[18] Guralnick, R., and Robinson, G. On the commuting probability in finite groups. *J. Algebra 300* (2006), 509 – 528.

[19] Guralnick, R. M., and Maroti, A. Average dimension of fixed point spaces with applications. *J. Algebra 226* (2011), 298 – 308.

[20] Guralnick, R. M., and Shumyatsky, P. Derived subgroups of fixed points. *Israel Math. J. 126* (2001), 345 – 362.

[21] Gustafson, W. H. What is the probability that two group elements commute? *Amer. Math. Monthly 80* (1973), 1031 – 1034.

[22] Haar, A. Der Massbegriff in der Theorie der kontinuierlichen Gruppen. *Ann. Math. 34* (1933), 147 – 169.

[23] Hall, P., and Higman, G. The p-length of a p-soluble group and reduction theorems for Burnside's problem. *Proc. London Math. Soc. (3) 6* (1956), 1–42.

[24] Hewitt, E., and Ross, K. *Abstract Harmonic Analysis. Vol. I.* Springer, Berlin, 1963.

[25] Hewitt, E., and Ross, K. *Abstract Harmonic Analysis. Vol. II.* Springer, Berlin, 1963.

[26] Hofmann, K., and Russo, F. The probability that x and y commute in a compact group. *Math. Proc. Cambridge Phil. Soc. 153* (2012), 557 – 571.

[27] Hofmann, K. H., and Russo, F. G. The probability that x^m and y^n commute in a compact group. *Bull. Aust. Math. Soc. 87* (2013), 503 – 513.

[28] Kakutani, S. A proof of the uniqueness of Haar's measure. *Ann. Math. 49* (1948), 225 – 226.

[29] Kelley, J. L. The Tychonoff product theorem implies the axiom of choice. *Fund. Math. 35* (1950), 75 – 76.

[30] Kelley, J. L. *General Topology.* Van Nostrand, Toronto - New York - London, 1955.

[31] Khukhro, E. I., and Shumyatsky, P. Bounding the exponent of a finite group with automorphisms. *J. Algebra 212* (1999), 363 – 374.

[32] Khukhro, E. I., and Shumyatsky, P. Words and pronilpotent subgroups in profinite groups. *J. Aust. Math. Soc. 97* (2014), 343 – 364.

[33] Khukhro, E. I., and Shumyatsky, P. Nonsoluble and non-p-soluble length of finite groups. *Israel J. Math. 207* (2015), 507 – 525.

[34] Lescot, P. Sur certains groupes finis. *Rev. Math. Spéciales* (Avril 1987), 276–277.

[35] Lescot, P. Degré de commutativité et structure d'un groupe fini (1). *Rev. Math. Spéciales* (Avril 1988), 276–279.

[36] Mann, A. The exponent of central factors and commutator groups. *J. Group Theory 10* (2007), 435–436.

[37] Mycielski, J. Some properties of connected compact groups. *Colloq. Math. 5* (1958), 162 – 166.

[38] Nachbin, L. *The Haar Integral*. Van Nostrand, 1965.

[39] Neumann, B. H. Groups covered by permutable subsets. *J. London Math. Soc. (3) 29* (1954), 236–248.

[40] Neumann, P. M. Two combinatorial problems in group theory. *Bull. Lond. Math. Soc. 21* (1989), 456 – 458.

[41] Reiter, H. *Classical Harmonic Analysis and Locally Compact Groups*. Oxford, 2000.

[42] Rezaei, R., and Erfanian, A. On the commutativity degree of compact groups. *Arch. Math. 43* (2009), 345 – 356.

[43] Ribes, L., and Zalesskii, P. *Profinite Groups*, second edition ed. Springer Verlag, Berlin - New York, 2010.

[44] Robinson, D. J. S. *A course in the theory of groups*, second edition ed., vol. 80. Springer-Verlag, New York, 1986.

[45] Shumyatsky, P. Coprime automorphisms of profinite groups. *Quart. J. Math. 53* (2002), 371 – 386.

[46] von Neumann, J. Zum Haarschen Maß in Topologischen Gruppen. *Compos. Math. 1* (1934), 106 – 114.

[47] von Neumann, J. The uniqueness of Haar's measure. *Mat. Sb. 1* (1936), 721 – 734.

[48] Weil, A. *L'Intégration dans les Groupes Topologiques et ses Applications*, second edition ed. Herman & C^{ie}, Paris, 1940.

[49] Wiegold, J. Groups with boundedly finite classes of conjugate elements. *Proc. Roy. Soc. London Ser. A 238* (1957), 389 – 401.

[50] Wilson, J. On the structure of compact torsion groups. *Monatsh. Math. 96* (1983), 57 – 66.

[51] Wilson, J. S. *Profinite Groups*, vol. 19. Clarendon Press, Oxford, 1998.

[52] Zelmanov, E. I. Solution of the restricted Burnside problem for groups of odd exponent. *Math. USSR-Izv. 36* (1991), 41–60.

[53] Zelmanov, E. I. Solution of the restricted Burnside problem for 2-groups. *Math. USSR-Sb. 72* (1992), 543 – 565.

[54] Zelmanov, . E. I. On periodic compact groups. *Israel J. Math. 77(1-2)* (1992), 83 – 95.

Index of Notations

Symbol	Meaning
$\|\ \|$	order of a finite set or group
$[G:H]$	the index of H in G
$C_G(x)$	the centralizer of x in G
G^n	the closed subgroup of G generated by nth powers
$[x,y]$	the commutator of x and y; $x^{-1}y^{-1}xy$
$[A,B]$	the closed subgroup generated by $[a,b], a \in A, b \in B$
G'	the derived group of G
x^g	the conjugate of x by g; $g^{-1}xg$
X^G	the set of conjugates of all $x \in X$
$\langle X^G \rangle$	the normal closure of X in G
$\gamma_n(G)$	nth term of lower central series of G
$Z_n(G)$	nth term of upper central series of G
$G^{(n)}$	nth term of derived series of G
$\langle S \rangle$	the closed subgroup generated by $S \subseteq G$
$FC(G)$	the FC-center of G
G_0	the connected component of G
ϕ	a continuous automorphism of a profinite group
$A^\#$	the set of nontrivial elements in the group A
$C_G(\phi)$	the centralizer of ϕ in G
$l_p(G)$	the p-length of a profinite group G
$\lambda(G)$	the nonsoluble length of a profinite group G
$\lambda_p(G)$	the non-p-soluble length of a profinite group G
$\mathcal{B}(X)$	the Borel σ-algebra on a set X
μ, ν	measures
$\mu \times \nu$	the product measure on a product space
χ_A	the characteristic function of a set A
X^r	set of products of at most r elements from the set X
$Pr(G)$	the commuting probability of G
$Pr(K,G)$	the relative commuting probability of K in G

$\hat{Z}$	the profinite completion of the integers
$\prod_{i\in I} G_i$	the Cartesian product of the groups G_i, $i \in I$
$w(x_1, \ldots, x_k)$	a group word
$w(G)$	the verbal subgroup determined by w
G_w	the set of w-values of G
$W_G(P)$	the subgroup of P defined by $\langle P_w^G \cap P \rangle$
$\mu*$	an outer measure
$(K : K_0)$	the index of K_0 in K