Fibres of Words in Finite Groups, a Probabilistic Approach

Ashurst, Carolyn

Award date:
2012

Awarding institution:
University of Bath

Fibres of Words in Finite Groups, a Probabilistic Approach

submitted by

Carolyn Ashurst

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Mathematical Sciences

June 2012

Summary

We investigate the relationship between a finite group and the set of probabilities associated with evaluating words over the group.

Given a finite group, a group element and a word, one may consider the probability that a uniformly random evaluation of the associated verbal mapping yields that particular element of the group. For a fixed group, we consider the set of such probabilities obtained by varying over all words and all group elements. It is known that properties of the group are reflected in this associated set of probabilities. For example, if a group is nilpotent then the set of non-zero probabilities associated with that group has a positive lower bound.

We seek to further establish the link between a finite group and its set of probabilities. We show how properties of the group, such as nilpotency and verbal subgroup structure are manifested in the properties of its set of probabilities, such as cardinality, the infimum and the corresponding set of accumulation points. We calculate the set of probabilities explicitly for several groups.

Acknowledgements

I would like to express my gratitude to everyone who gave me the help and support I needed throughout my PhD. Thank you to the EPSRC whose sponsorship allowed me the opportunity to undertake this research, and to the University of Bath Mathematics Department for providing such an excellent working environment and supportive community. Thank you in particular to my supervisor Geoff Smith for his ideas, encouragement, humour, and his expert supervision. Thanks to the algebraists Gunnar, Simon and Peter and to many other postgraduates at Bath for numerous useful discussions and suggestions. Thanks also to those who did time in 1W4.17 for many unproductive but hugely entertaining afternoons, in particular to James, Phil, Adam and Jane. Thanks to everyone who made my time as a postgraduate such an enjoyable one, whether through maths, music, adventures away from Bath or surprise visits to my office. A big thank you of course goes to my parents and sister for their continuous love and support in everything I do. And finally a special thanks to Chris Guiver for proof reading numerous drafts and providing many helpful suggestions, for his encouragement and support, and for his huge amount of help in matters technical, mathematical and in countless other ways.

Contents

Chapter 1

Introduction

In [9] Gustafson answered the following question.

> Given a finite group G, what is the probability that two randomly selected group elements commute?

Gustafson employed techniques used by Erdös and Turán in [7] to show that the answer is kn^{-1}, where k is the number of conjugacy classes of G, and n is the order of the group. Gustafson's question might be rephrased as a question about commutators. The expression $x^{-1}y^{-1}xy$ is known as the **commutator word** and is often denoted by $[x, y]$. Given a group G, we may use the commutator word to define a map γ_2 from $G \times G$ to G in the obvious way, namely

$$\gamma_2 : (g, h) \mapsto g^{-1}h^{-1}gh.$$

A pair of group elements g, h commute if, and only if, $\gamma_2(g, h) = 1$, that is, (g, h) is in the fibre of 1 under γ_2. Hence the original question may be rephrased as follows.

> Given a finite group G, what is the probability that two randomly selected group elements g, h satisfy $\gamma_2(g, h) = 1$?

Or, equivalently,

> Given a finite group G, what is the probability that two randomly selected group elements are in the fibre of 1 under γ_2?

We denote this probability by $P(G, \gamma_2 = 1)$. We might ask similar questions about the fibres of other group elements under maps defined using expressions other than $[x, y]$.

A **word** is an expression of the form

$$w(x_1, \ldots, x_k) := \prod_{j=1}^{l} x_{i_j}^{\varepsilon_{i_j}},$$

where $k \in \mathbb{N}$, $i_j \in \{1, \ldots, k\}$ and $\varepsilon_{i_j} \in \{\pm 1\}$ for all j. Given a word w and a group element g, we denote the probability that a random k-tuple of group elements lies in the fibre of g under w by $P(G, w = g)$. We denote the image of w over G by G_w^+.

It is known that properties of the group G give restrictions on the possible values of $P(G, w = g)$. The example most relevant to our work is the following theorem, proved by Nikolov and Segal in [23].

Theorem 1.0.1 (Nikolov and Segal)**.** *Let G be a finite group, and put $\varepsilon(G) = p^{-|G|}$ where p is the largest prime divisor of $|G|$.*

(i) G is soluble if and only if

$$\inf_w P(G, w = 1) > 0,$$

where w ranges over all words, and this holds if and only if

$$\inf_w P(G, w = 1) > \varepsilon(G).$$

(ii) G is nilpotent if and only if

$$\inf_{w,g} P(G, w = g) > 0,$$

where w ranges over all words and g ranges over G_w^+, and this holds if and only if

$$\inf_{w,g} P(G, w = g) > \varepsilon(G).$$

This result raises several questions which have motivated the work in this thesis. Firstly, as asked by Segal in [27], "Is the bound $\varepsilon(G)$ best possible?" In Section 2.4 we shall see that for abelian groups the infimum is easily shown to be $|G|^{-1}$. This is not very close to the bound given in Theorem 1.0.1. In order to investigate this bound, we may ask, "For a given finite group G, how do we calculate this infimum explicitly?" These questions are explored in Chapters 3 and 4 where we calculate the infimum associated with several small nilpotent groups of class two, as well as two infinite families of nilpotent groups.

The second part of Theorem 1.0.1 may be restated as follows.

Theorem. *A finite group G is non-nilpotent if, and only if, there exists a sequence of words (w_i) and a sequence of elements $(g_i) \in G$, such that*

$$\lim_{i \to \infty} P(G, w_i = g_i) = 0.$$

This leads one to ask, "For a fixed finite group G, what values in the interval $[0,1]$ may be written as the limits of a sequence of probabilities associated with G?" Let $S(G)$ denote the set of all probabilities associated with G, that is

$$S(G) := \{P(G, w = g) \mid w \text{ a word}, g \in G\}.$$

Let T be a set of real numbers. An element $r \in \mathbb{R}$ is called an **accumulation point** of T if r is the limit of a sequence of elements of T that is not eventually constant. Using this terminology, Theorem 1.0.1 says that 0 is an accumulation point of $S(G)$ if and only if G is non-nilpotent. We thus know exactly which finite groups G are such that 0 is an accumulation point of $S(G)$, but we are interested to know what other accumulation points are associated with a given finite group. For example, although nilpotent groups cannot have 0 as accumulation point of their set of probabilities, can they have other accumulation points? What do these accumulation points tell us about the structure of G?

It is known that the set of probabilities associated with some finite groups are dense in the interval $[0,1]$. In [1], Miklós Abért proves that if G is a finite just non-solvable group then $S(G)$ is dense in $[0,1]$. We are interested to know exactly which finite groups are such that $S(G)$ is dense in $[0,1]$. If G is nilpotent, then by Theorem 1.0.1, we know that $S(G)$ cannot be dense in $[0,1]$, since it does not have probabilities arbitrarily close to 0. In this thesis we shall prove that the associated set of probabilities is dense for certain non-nilpotent (but soluble) groups.

In short, the overall aim of this thesis is to explore the relationship between properties of G and properties of $S(G)$.

1.1 Probabilistic group theory

In 1708, Montmort considered the following game for one player. The player has n cards, labeled 1 to n, which he shuffles face down. He then reveals the cards, one by one, without replacement. Whilst doing this he counts from 1 to 10 out loud as each card is shown. If at any point he reveals a card that matches the number spoken, he wins. If he reaches the last card without having had any matches, he loses. Montmort asked himself, "What is the probability of losing at this game?" We could view this as a probabilistic question about the symmetric group of degree n, which we denote by $\mathrm{Sym}(n)$. A **derangement** is a permutation with no fixed points. The problem is equivalent to asking, "What is the probability $P(n)$ that a randomly chosen element of $\mathrm{Sym}(n)$ is a derangement?" For example, in $\mathrm{Sym}(3)$, the derangements are the 3-cycles, $(1,2,3)$ and $(1,3,2)$. Thus $P(3) = 2/6 = 1/3$. In "Essay d'analyse sur les jeux

de hazard", Montmort proved that

$$P(n) = 1 - \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{(n-k)!}{n!} = \frac{1}{2!} - \frac{1}{3!} \ldots + \frac{(-1)^n}{n!}$$

and

$$P(n) \to e^{-1} \text{ as } n \to \infty.$$

Perhaps this is the first example of probabilistic group theory, albeit unknown to Montmort.

Two and a half centuries later, Erdös and Turán wrote a series of papers on the statistics of symmetric groups [4] – [7]. In the fourth paper [7], they included the following.

Theorem 1.1.1 (Erdös and Turán)**.** *The number of commutable (a, b) pairs from an arbitrary group G of order n is nk where k stands for the number of conjugacy classes in G.*

In our notation,

$$P(G, [x_1, x_2] = 1) = \frac{k}{n}.$$

In [9], Gustafson went on to show that for any finite non-abelian group

$$P(G, [x_1, x_2] = 1) \leq \frac{5}{8}.$$

These are probably the first results that consider the probability of obtaining a certain result when a word is randomly evaluated in a finite group.

Group theorists have used the language and sometimes methods of probability theory in a wide range of problems. In the survey [3], Dixon gives an overview of three such areas, some of which we shall briefly touch on here.

Perhaps the majority of questions regarding probability have been to do with the probability that some random collection of group elements form a generating set. In 1882, Netto [22] conjectured the following which was proved by Dixon in 1969 in [2].

Theorem 1.1.2 (Dixon)**.** *Let $Alt(n)$ denote the alternating group of degree n. Let p_n denote the probability that 2 random elements of $Alt(n)$ generate $Alt(n)$. Then p_n tends to 1 as n tends to infinity.*

Dixon then suggested that the following might hold.

Theorem 1.1.3. *Let G be a finite simple group of order n. Then the probability that two randomly chosen elements of G generate G tends to 1 as n tends to infinity.*

This was proved to be the case by Kantor and Lubotsky in 1990 [14] and Liebeck and Shalev in 1995 [18]. Liebeck and Shalev extend these results in [19]. Here they use group zeta functions to prove several results of the following form.

> Given two random elements with a certain property (e.g. a certain order) within a finite simple group, what happens to the probability that these elements generate the group, as the order of G tends to infinity?

In their paper, Kantor and Lubotsky also prove the following.

Theorem 1.1.4. *Let S be a finite simple group and let S_0 be a group with $S \leq S_0 \leq Aut(S)$. If $P(S_0)$ is the probability that two randomly chosen elements of S_0 generate a subgroup containing S, then $P(S_0)$ tends to 1 as $|S_0|$ tends to infinity.*

This result played a key role in the proof of the following theorem, which was the main result presented in [13], a paper by Dixon, Pyber, Seress and Shalev.

Theorem 1.1.5 (Dixon, Pyber, Shalev and Seress)**.** *Let S be a finite simple group and let $w(x, y)$ be a non-trivial word on two variables. Then the probability that two randomly chosen elements x, y in S satisfy $w(x, y) \neq 1$ tends to 1 as $|S|$ tends to infinity.*

Here we see the language of probability theory and words reunited. In contrast to the work in this thesis, the authors of [13] explore what happens when one fixes a word and varies the group over which the evaluation takes place. In our work, we are interested instead in what happens when we fix a group and vary the words.

In [1], Abért considers exactly this for non-soluble groups. The main result of the paper shows the existence of words that test whether a subgroup generated by elements in the group will be soluble.

Theorem 1.1.6 (Abért)**.** *Let G be a finite group. Then for all n there exists a word $w \in F_n$ such that for all $g_1, g_2, \ldots, g_n \in G$, the tuple $(g_1, g_2, \ldots, g_n)$ satisfies w if and only if the subgroup $\langle g_1, \ldots, g_n \rangle \leq G$ is soluble.*

As a corollary, Abért explains that the probabilities associated with satisfying words in a non-soluble group can be made arbitrarily small. That is,

Theorem 1.1.7 (Abért)**.** *Let G be a finite non-soluble group, then*

$$\inf_w P(G, w = 1) = 0.$$

In our terminology, this theorem says that if G is non-soluble, 0 is an accumulation point of $S(G, 1) := \{\ P(G, w = 1) \mid w \text{ a word } \}$. In the same paper Abért also proves the following.

Theorem 1.1.8 (Abért)**.** *Let G be a finite just non-soluble group. Then the set*

$$\{P(G, w = 1) \mid w \in F_\infty\}$$

is dense in $[0, 1]$.

In [23], Nikolov and Segal show the converse of Theorem 1.1.7. Thus they give the following result.

Theorem 1.1.9 (Nikolov and Segal)**.** *Let G be a finite group. Then G is soluble if and only if the numbers $P(G, w = 1)$ are bounded away from zero as w ranges over all group words.*

It is in this paper that they also prove the following characterisation of nilpotent groups.

Theorem 1.1.10 (Nikolov and Segal)**.** *Let G be a finite group. Then G is nilpotent if and only if the positive values of $P(G, w = c)$ are bounded away from zero as c ranges over G and w ranges over all group words.*

It is this last result that has been of the greatest importance to the work in this thesis. For example, we have sought to find out what this bound is for specific groups, and asked ourselves whether groups may have accumulation points other than 0. A discussion of Theorem 1.1.8 and Theorem 1.1.9 can be found in the 2009 Groups St Andrews conference proceedings [27], along with results concerning *ellipticity* of groups. Although we do not consider questions regarding ellipticity in this thesis, since they have had far reaching consequences, we shall discuss them now briefly.

We say that a word w has **width** $m_G(w)$ in a group G if every element of its verbal subgroup (that is, the subgroup generated by the image of w) can by written as the product of m elements, each of which is in the image of w, or is the inverse of such an element. If $m_G(w)$ is finite the group G is said to be w**-elliptic**. G is **verbally elliptic** if it is w-elliptic for every word w. The study of verbal subgroups in infinite groups was begun by P. Hall in [11] and [12]. He and his students sought to find exactly which groups are verbally elliptic. Stroud [30] proved that all finitely generated abelian-by-nilpotent groups G are verbally elliptic. Rhemtulla [25] proved that every not infinite dihedral free product G = A . B of non-trivial groups A and B is **verbally parabolic**, i.e., every proper verbal subgroup of G has infinite width. Romankov [26] and George [8] also made contributions to answering this question. A discussion of such results can be found in [28]. In 2007 Nikolov and Segal used such ideas to prove a generalisation of the following theorem by Serre.

Theorem 1.1.11 (Serre)**.** *In a finitely generated pro-p group every subgroup of finite index is open.*

Serre asked whether this holds for all finitely generated profinite groups. In [24], Nikolov and Segal proved that it was, and hence showed the following.

Theorem 1.1.12 (Nikolov and Segal)**.** *In a finitely generated profinite group the topology is uniquely determined by the group structure.*

Other important results that have come from the study of verbal width within recent years include Shalev's proof [29] that for any nontrivial word w every element of a sufficiently large finite simple group is a product of three values of w, and a proof of Ore's conjecture. This states that every element of every non-abelian finite simple group is a commutator, and was proved by Liebeck, O'Brien, Shalev and Tiep in [20].

1.2 Thesis overview

We begin Chapter 2 by introducing the basic concepts used in this thesis and establishing the notation we shall use. We then present some preliminary results. These include those concerned with finding accumulation points, and those concerning various group constructions, (direct products for example), which shall be used throughout this thesis. This chapter also includes what is known on the subject regarding abelian groups, simple groups and verbally simple groups.

The next two chapters contain all results regarding non-abelian nilpotent groups. In Chapter 3 we demonstrate how the set of probabilities may be explicitly calculated for small nilpotent groups. We include calculations for several small groups of nilpotency class 2. In Chapter 4 we discuss what is known regarding the infimum problem. We state and prove the value of the infimum of the set of non-zero probabilities associated with each nilpotent dihedral group and generalised quaternion group.

In Chapter 5 we discuss our findings concerning non-nilpotent groups. We show that the set of probabilities associated with G is dense in the interval $[0, 1]$ whenever G is a non-nilpotent generalised dihedral group, non-nilpotent alternating group or non-nilpotent symmetric group. We conclude this chapter with a summary of results concerning the alternating and symmetric groups.

In the final chapter we present a summary of our work, along with a discussion of related conjectures and open problems.

Chapter 2

Definitions and preliminary results

In this chapter we include several definitions and results that will be used throughout the main body of this thesis. We begin by defining the basic concepts and introducing the notation we shall use. In Section 2.2 we shall introduce two methods used to show the existence of limit points in the set of probabilities associated with a group. In Section 2.3 we collect results regarding direct products, quotient groups and verbal subgroups. In the last three sections we include results regarding abelian groups, simple groups and verbally simple groups.

2.1 Words, verbal mappings and associated sets of probabilities

Definition 2.1.1. *A **word on** k **variables** is an expression of the form*

$$w(x_1, \ldots, x_k) := \prod_{j=1}^{l} x_{i_j}^{\varepsilon_{i_j}}, \tag{2.1}$$

*where $i_j \in \{1, \ldots, k\}$ and $\varepsilon_{i_j} \in \{\pm 1\}$ for all j. We call l the **length** of the word.*

A word on k or fewer variables may be viewed as an element of F_k, the free group on k letters. We shall also use the notation F_∞ to denote the set of all words of finite length.

Given a group G, we may use w to define a map on k copies of G in the obvious way. Let $G^{(k)}$ denote the direct product of k copies of G. The associated map $w : G^{(k)} \to G$ is given by

$$w(\mathbf{g}) = w(g_1, \ldots, g_k) = \prod_{j=1}^{l} g_{i_j}^{\varepsilon_{i_j}},$$

where $\mathbf{g} = (g_1, \ldots, g_k) \in G^{(k)}$. When viewed in this way, we call $w : G^{(k)} \to G$ a **word over** G. We refer to this map as the **verbal mapping** associated with w. We denote the image of w over G by G_w^+.

Given a finite group G, and two words w_1 and w_2 on k variables, it is possible that w_1 and w_2 may be different elements of F_k, but give rise to the same verbal mapping over G. For example, if G is the cyclic group of order 3, then the words x^2 and x^5 give rise to the same verbal mapping over G. Since we are interested in words only when viewed as maps, in such a situation we do not distinguish between the two words, and say that over G, w_1 may be written as w_2, or that w_2 is G-**equivalent** to w_1.

In general, the ε_{i_j} in (2.1) may have a value of 1 or -1. If w is such that $\varepsilon_{i_j} = 1$ for all j then w is called a **positive word**. If G is a finite group, then every word is G-equivalent to a positive word – one may replace x_i^{-1} by x_i^{m-1} where m is the exponent of G. Since we are only concerned with finite groups in this thesis, we therefore often assume all words are positive, without loss of generality.

Definition 2.1.2. *Let G be a finite group and w a word on k variables. The image of $G^{(k)}$ under w is denoted by G_w^+, and the subgroup generated by the image of w is denoted by $w(G)$. Thus $w(G) = \langle G_w^+ \rangle$. Similarly, if T is a subset of G, then we denote the image of $T^{(k)}$ under w by T_w^+, and the subgroup generated by the image of T by $w(T)$. If g is an element of G we denote the fibre or preimage of g under w by $w^{-1}(g)$. If $g_1, \ldots, g_k \in G$ are such that $w(g_1, \ldots, g_k) = 1$, that is $(g_1, \ldots, g_k)$ is in the fibre of 1, we say that $g_1, \ldots, g_k$ **satisfy** w.*

Given a word it is natural to ask about the fibres of its associated verbal mapping. When G is a finite group, we may ask about the size of these fibres. In order to make meaningful comparisons that are independent of k, the number of variables, we shall talk about the *relative size* of a fibre. That is, the size of a fibre as a proportion of $G^{(k)}$. To do this, we use the following definition.

Definition 2.1.3. *Let G be a finite group, g an element of G and w a word on k variables. The **probability that** w **evaluates to** g is*

$$P(G, w = g) := \frac{|w^{-1}(g)|}{|G|^k}.$$

*The **set of probabilities associated with** G, denoted $S(G)$ is*

$$S(G) := \{P(G, w = g) \mid w \text{ a word}, g \in G\}.$$

*The **set of probabilities associated with** g **in** G, denoted $S(G, g)$ is*

$$S(G, g) := \{P(G, w = g) \mid w \text{ a word}\}.$$

$S(G)$ is of course the union of the $S(G, g)$, as g ranges over G.

Throughout this thesis we are interested in how $S(G)$ and $S(G, g)$ reflect properties of the group G. In Theorem 1.0.1, Segal and Nikolov show that 0 is an accumulation of $S(G)$ if and only if G is non-nilpotent. We might therefore wonder whether non-nilpotent groups have any other accumulation points, and whether nilpotent groups may have any at all. In order to address these questions, we use two methods to demonstrate the existence of accumulation points. We describe these methods in the next section.

2.2 Finding accumulation points

Definition 2.2.1. *Let T be a set of real numbers. An element $r \in \mathbb{R}$ is called an **accumulation point** of T if every open interval containing r contains infinitely many elements of T. Equivalently, r is the limit of a sequence of elements of T that is not eventually constant.*

Note that some authors refer to such an element as an ω-accumulation point.

Definition 2.2.2. *Let T be a set of real numbers. Then $r \in \mathbb{R}$ is a **limit point** of T if r is an accumulation point of T, or $r \in T$. Equivalently, r is the limit of a sequence whose elements are in T.*

There are two methods we have used to construct sequences of words which yield limit points of $S(G)$. These are called the *concatenation method* and the *substitution method*.

2.2.1 The concatenation method

Let $w_1(x_1, \ldots, x_{k_1})$ and $w_2(y_1, \ldots, y_{k_2})$ be words on k_1 and k_2 variables respectively. We define the word $w_1 * w_2$ to be the word on $k_1 + k_2$ variables given by

$$(w_1 * w_2)(x_1, \ldots, x_{k_1+k_2}) = w_1(x_1, \ldots, x_{k_1})w_2(x_{k_1+1}, \ldots, x_{k_1+k_2}).$$

We call this **concatenation**. Note that in the concatenation, w_1 and w_2 are given independent variables. We are at liberty to change the names of variables for clarity.

We may use this construction to create an infinite sequence of words in the following way. Let w be a word on k variables. We define $(w_i)_{i\in\mathbb{N}}$ as follows.

$$\begin{aligned} w_1(x_1, \ldots, x_k) &:= w(x_1, \ldots, x_k) \\ w_{r+1}(x_1, \ldots, x_{(r+1)k}) &:= w_r * w_1 = w_r(x_1, \ldots, x_{rk}) * w(x_{rk+1}, \ldots, x_{(r+1)k}) \qquad \forall r \geq 1. \end{aligned}$$

Example Let $w(x_1, x_2) = [x_1, x_2]$. The sequence constructed from this word by the concatenation method is then

$$\begin{aligned} w_1(x_1, x_2) &= [x_1, x_2], \\ w_2(x_1, \ldots, x_4) &= [x_1, x_2][x_3, x_4], \\ w_3(x_1, \ldots, x_6) &= [x_1, x_2][x_3, x_4][x_5, x_6], \quad \text{etc.} \end{aligned}$$

Proposition 2.2.3. *Let G be a finite group, w a word and $g \in w(G)$. If $(w_i)_{i\in\mathbb{N}}$ is the sequence of words constructed by concatenation as above, then*

$$P(G, w_i = g) \to |w(G)|^{-1} \text{ as } i \to \infty.$$

Thus $|w(G)|^{-1}$ is a limit point of $S(G)$.

Proof. The probabilities associated with such a sequence may be thought of within the context of a **Markov chain**, that is, a sequences of random variables in which the state of the next variable depends only on the current state, and no others before it. We shall use a well known result from probability theory that gives sufficient conditions for a Markov chain to have equally likely limiting probabilities.

Let $g_1, g_2, g_3, \ldots$ be uniformly randomly selected elements from G. Define the random variables (in the sense of probability theory) $Y_1, Y_2, \ldots$ to be

$$Y_i := w_i(g_1, \ldots, g_{ik}).$$

Then $Pr(Y_i = g) = P(G, w_i = g)$. Since the variables in each concatenated copy of w in w_i are independent, we have

$$P(G, w_{i+1} = g) = \sum_{h\in G} P(G, w_i = h)P(G, w = h^{-1}g)$$

and so Y_{i+1} depends only on Y_i and not on $Y_1, \ldots, Y_{i-1}$. Thus we have a Markov chain. We define our **state space** (i.e. the set of possible values of the random variables Y_i) to be $w(G)$.

The following four definitions may be found in any standard text on probability theory, see [21] for example.

Definition 2.2.4. *A Markov chain is **irreducible** if it is possible to reach any state from any other state.*

Since G is finite and $w(G) = \langle G_w^+ \rangle$, there exists some $m \in \mathbb{N}$ such that any element g of the state space $w(G)$ may be written as a product of m elements of the image G_w^+. One could take $m = |w(G)|$ for example. (In the language of Dan Segal, finite groups

are verbally elliptic. See [28] for details). Let $g = g_1 \dots g_m$ where g_i is in G_w^+ for all i. Then g is in the image of w_m. Suppose h_i and h_j are in $w(G)$ and we wish to reach the state h_j from h_i. Since $g := h_j h_i^{-1} \in w(G)$ is in the image of w_m for some $m \geq 1$,

$$Pr(Y_{m+1} = h_j \mid Y_1 = h_i) = P(G, w_m = h_i^{-1} h_j) > 0.$$

Thus we have an irreducible Markov chain.

Definition 2.2.5. *A Markov chain is **aperiodic** if there exists some $n' \in \mathbb{N}$ such that for all $n > n'$*

$$Pr(Y_n = i \mid Y_1 = i) > 0.$$

Suppose $Y_1 = g$. That is, $w(g_1, \dots, g_k) = g$. If $g_{k+1} = \dots = g_{2k} = 1$ (which happens with non-zero probability) then $Y_2 = g$. Similarly, if $g_{k+1}, \dots, g_{nk} = 1$ (which again happens with non-zero probability) then $Y_n = g$. Thus $Pr(Y_n = g \mid Y_1 = g) > 0$ for all $n > 1$. Thus we have an aperiodic Markov chain.

Definition 2.2.6. *The **transition matrix** of a Markov chain with state space*

$$y_1, \dots, y_m$$

is the matrix $(a_{ij}) \in \mathbb{R}^{m \times m}$ with entries

$$a_{ij} = Pr(Y_{n+1} = y_j \mid Y_n = y_i).$$

Definition 2.2.7. *A Markov chain is **doubly stochastic** if every row and column of its transition matrix has entries that sum to 1.*

Let $A = (a_{ij})$ denote the transition matrix associated with our Markov chain. The sum of the i^{th} row of A is

$$\sum_j a_{ij} = \sum_j P(G, w_n = g_j \mid w_{n-1} = g_i) = P(G, w_n \in G \mid w_{n-1} = g_i) = 1.$$

The sum of the j^{th} column of A is

$$\sum_i a_{ij} = \sum_i P(G, w_n = g_j \mid w_{n-1} = g_i) = \sum_i P(G, w_1 = g_j g_i^{-1}) = P(G, w_1 \in G) = 1.$$

Thus A is doubly stochastic.

It is known (see for example p151 of [21]) that an irreducible and aperiodic Markov chain with a finite state space has states that are equally likely in the limit if and only if it has a transition matrix that is doubly stochastic. Thus the limiting probabilities

of our Markov chain are equal. Since the state space is $w(G)$ we have

$$\lim_{i\to\infty} P(G, w_i = g) = |w(G)|^{-1}$$

for any $g \in w(G)$. □

We shall use the above result to show that if G is a finite group such that the left normed commutator γ_i is not trivial, then $|\gamma_i(G)|^{-1}$ is an accumulation point of $S(G)$. To prove this we will use the following lemma.

Lemma 2.2.8. *Let $\gamma_i(x_1, \ldots, x_i)$ be the left normed commutator of length i, and let G be a finite group such that $\gamma_i(G)$ is not trivial (i.e. G is either non-nilpotent, or has nilpotency class at least i). Then for any $1 \neq g \in G$*

$$P(G, \gamma_i = 1) > P(G, \gamma_i = g).$$

Proof. Let G and γ_n satisfy the conditions of the lemma, for some $n \geq 2$. For simplicity of notation define $\gamma_1(x_1) := x_1$. Let $g_1, \ldots, g_n \in G$. Then

$$\begin{aligned} \gamma_n(g_1, \ldots, g_n) = 1 &\iff [\gamma_{n-1}(g_1, \ldots, g_{n-1}), g_n] = 1 \\ &\iff g_n \in C_G(\gamma_{n-1}(g_1, \ldots, g_{n-1})). \end{aligned}$$

So there are

$$\sum_{g_1,\ldots,g_{n-1}\in G} |C_G(\gamma_{i-1}(g_1, \ldots, g_{n-1}))|$$

$n-$tuples that satisfy γ_n.

Now suppose $\gamma_n(g_1, \ldots, g_n) = g \neq 1$ for some $g, g_i \in G$. Then

$$\begin{aligned} & \gamma_n(g_1, \ldots, g_n) &&= \gamma_n(g_1, \ldots, g_{n-1}, h_n) \\ \iff \quad & [\gamma_{n-1}(g_1, \ldots, g_{n-1}), g_n] &&= [\gamma_{n-1}(g_1, \ldots, g_{n-1}), h_n] \\ \iff \quad & \gamma_{n-1}(g_1, \ldots, g_{n-1})^{g_n} &&= \gamma_{n-1}(g_1, \ldots, g_{n-1})^{h_n} \\ \iff \quad & h_n &&\in g_n.C_G(\gamma_{n-1}(g_1, \ldots, g_{n-1})). \end{aligned}$$

There are $|C_G(\gamma_{n-1}(g_1, \ldots, g_{n-1}))|$ such h_n. Thus if T is the set of $(n-1)-$tuples in $G^{(n-1)}$ such that there is at least one solution to $\gamma(g_1, \ldots, g_{n-1}, x_n) = g$ then there are

$$\sum_{(g_1,\ldots,g_{n-1})\in T} |C_G(\gamma_{i-1}(g_1, \ldots, g_{n-1}))|$$

solutions to $\gamma_n = g$. Since $T \subsetneq G^{(n-1)}$ (for example $(1, \ldots, 1) \notin T$) it follows that

$|\gamma_n^{-1}(g)| < |\gamma_n^{-1}(1)|$. Thus

$$P(G, \gamma_n = g) < P(G, \gamma_n = 1). \qquad \square$$

Lemma 2.2.9. *Let G be a finite group and $d \in \mathbb{N}$ be such that the left normed commutator γ_d is not trivial in G, i.e. G is non-nilpotent, or is nilpotent of class at least d. Let $m = |\gamma_d(G)|$ be the order of the verbal subgroup associated with γ_d. Then m^{-1} is an accumulation point of $S(G)$. Thus if G is not abelian then $S(G)$ is infinite.*

Before we prove this, we shall prove the following.

Lemma 2.2.10. *Let G be a finite group such that $\gamma_d(G)$ is not trivial for some $d \geq 2$. Then for all $g \in G$,*

$$|\gamma_d^{-1}(g)| = |\gamma_d^{-1}(g^{-1})|.$$

Proof. We shall use the following well known commutator identities that may be easily verified

(i) $[x, y]^{-1} = [y, x]$,

(ii) $[x, y]^{-1} = [x^{-1}, y^x]$.

By the first identity, if g is a commutator, then

$$\phi_2 : \gamma_2^{-1}(g) \to \gamma_2^{-1}(g^{-1}), \qquad (a, b) \mapsto (b, a)$$

is well defined. Since ϕ_2 is a bijection (it is self inverse) it follows that

$$|\gamma_2^{-1}(g)| = |\gamma_2^{-1}(g^{-1})| \qquad \forall g \in G_{\gamma_2}^+. \tag{2.2}$$

Note too that by identity (i), $g \in G_{\gamma_2}^+$ if and only if $g^{-1} \in G_{\gamma_2}^+$, and thus (2.2) holds for all $g \in G$.

Now consider γ_3. Suppose $\gamma_3(a, b, c) = g$. Then

$$\begin{aligned}
\gamma_3(b, a, c^{[a,b]}) &= \left[[b, a], c^{[a,b]} \right] \\
&= \left[[a, b]^{-1}, c^{[a,b]} \right] \quad \text{(by identity (i))} \\
&= [[a, b], c]^{-1} \quad \text{(by identity (ii))} \\
&= \gamma_3(a, b, c)^{-1}.
\end{aligned}$$

So for all $g \in G_{\gamma_3}^+$,

$$\phi_3 : \gamma_3^{-1}(g) \to \gamma_3^{-1}(g^{-1}), \qquad (a, b, c) \mapsto (b, a, c^{[a,b]})$$

is well defined and since it is self inverse (this is easily checked) ϕ_3 is a bijection and $|\gamma_3^{-1}(g)| = |\gamma_3^{-1}(g^{-1})|$.

We might write ϕ_3 as

$$\phi_3(a,b,c) = (\phi_2(a,b), c^{\gamma_2(a,b)}).$$

Now define ϕ_i inductively as follows.

$$\begin{aligned}\phi_2(a,b) &:= (b,a)\\ \phi_i(g_1,\ldots,g_i) &:= \left(\phi_{i-1}(g_1,\ldots,g_{i-1}), g_i^{\gamma_{i-1}(g_1,\ldots,g_{i-1})}\right) \quad \forall i > 2.\end{aligned}$$

Suppose for some i, ϕ_{i-1} restricted to $\gamma_{i-1}^{-1}(g)$ maps onto $\gamma_{i-1}^{-1}(g^{-1})$ for all g in G. Then

$$\begin{aligned}\gamma_i(\phi_i(g_1,\ldots,g_i)) &= \left[\gamma_{i-1}(g_1,\ldots,g_{i-1})^{-1}, g_i^{\gamma_{i-1}(g_1,\ldots,g_{i-1})}\right]\\ &= [\gamma_{i-1}(g_1,\ldots,g_{i-1}), g_i]^{-1}\\ &= \gamma_i(g_1,\ldots,g_i)^{-1}.\end{aligned}$$

Therefore ϕ_i restricted to $\gamma_i^{-1}(g)$ maps onto $\gamma_i^{-1}(g^{-1})$ for all g in $G_{\gamma_i}^+$. Note that ϕ_i is self-inverse and therefore a bijection. Thus

$$|\gamma_i^{-1}(g)| = |\gamma_i^{-1}(g^{-1})| \qquad \forall g \in G_{\gamma_i}^+.$$

Note also that since $\gamma_i(g_1,\ldots,g_i) = g$ if and only if $\gamma_i(\phi_i(g_1,\ldots,g_i)) = g^{-1}$, it follows that $g \in G_{\gamma_i}^+$ if and only if $g^{-1} \in G_{\gamma_i}^+$.

Thus by induction,

$$|\gamma_i^{-1}(g)| = |\gamma_i^{-1}(g^{-1})|$$

for any $g \in G$ and any $i > 2$ for which $\gamma_i(G)$ is non-trivial. □

In the proof of Lemma 2.2.9 we shall also use the following well known result.

Lemma 2.2.11 (Rearrangement inequality)**.** *Let $p_1 \leq \ldots \leq p_n$ and $q_1 \leq \ldots \leq q_n$ be real values. If $\sigma \in Sym(n)$ then*

$$\sum_{i=1}^{n} p_i q_{\sigma(i)} \leq \sum_{i=1}^{n} p_i q_i.$$

Equivalently, the sum

$$\sum_{i=1}^{n} p_i q_{\sigma(i)}$$

is maximised when σ is the identity.

Proof of Lemma 2.2.9. Let G be a finite group where $\gamma_d(G) \neq 1$, for some fixed $d > 2$.

For all $i \in \mathbb{N}$, let w_i be the word constructed inductively as follows.

$$\begin{aligned} w_1 &:= \gamma_d \\ w_i &= w_{i-1} * w_{i-1} \qquad \forall i > 1 \end{aligned}$$

so that w_i is constructed from 2^{i-1} concatenated copies of γ^d.

By Proposition 2.2.3 we have that

$$P(G, w_i = 1) \to |\gamma_d(G)|^{-1} = \tfrac{1}{m}$$

as i tends to infinity. To see this, note that if $(\widetilde{w}_i)$ is the sequence of words given by $\widetilde{w}_1 := \gamma_d, \widetilde{w}_{i+1} := \widetilde{w}_i * \gamma_d$ then (w_i) is a subsequence of $(\widetilde{w}_i)$, which has the above limit by Proposition 2.2.3.

Thus m^{-1} is a limit point of $S(G)$. To show that m^{-1} is also an accumulation point (i.e. the limit of a *non-constant* sequence) of $S(G)$, we shall show that

$$P(G, w_i = g) < P(G, w_i = 1) \tag{2.3}$$

for all $g \in G$ and $i \in \mathbb{N}$. We shall prove this by induction.

By Lemma 2.2.8, (2.3) holds for $i = 1$. Now suppose that $i > 1$.

Claim 2.2.12. *For all $g \in G, i \in \mathbb{N}$*

$$P(G, w_i = g) = P(G, w_i = g^{-1}).$$

Proof of Claim 2.2.12. Define

$$\begin{aligned} \phi : w_i^{-1}(g) &\to w_i^{-1}(g^{-1}) \\ (g_1, \dots, g_{id}) &\mapsto (\phi_i(g_{(i+1)d+1}, \dots, g_{id}), \dots, \phi_i(g_1, \dots, g_d)) \end{aligned}$$

where ϕ_i is as defined in Lemma 2.2.10. This is well defined since if $w_i(g_1, \dots, g_{id}) = g$ then

$$\begin{aligned} &w_i(\phi_i(g_{(i+1)d+1}, \dots, g_{id}), \dots, \phi_i(g_1, \dots, g_d)) \\ &= \gamma_d(\phi_i(g_{(i+1)d+1}, \dots, g_{id})) \dots \gamma_d(\phi_i(g_1, \dots, g_d)) \\ &= \gamma_d(g_{(i-1)d+1}, \dots, g_{id})^{-1} \dots \gamma_d(g_1, \dots, g_d)^{-1} \\ &= \big(\gamma_d(g_1, \dots, g_d) \dots \gamma_d(g_{(i-1)d+1}, \dots, g_{id})\big)^{-1} \\ &= w_i(g_1, \dots, g_{id})^{-1} = g^{-1} \end{aligned}$$

and the claim is proved. □

We now continue with the proof of Lemma 2.2.9. Since $w_i = w_{i-1} * w_{i-1}$ it follows that for all $g \in G$,

$$P(G, w_i = g) = \sum_{h \in G} P(G, w_{i-1} = h) P(G, w_{i-1} = h^{-1} g) \tag{2.4}$$

and in particular, using Claim 2.2.12,

$$\begin{aligned} P(G, w_i = 1) &= \sum_{h \in G} P(G, w_{i-1} = h) P(G, w_{i-1} = h^{-1}) \\ &= \sum_{h \in G} P(G, w_{i-1} = h)^2 \end{aligned}$$

Let $g_1, \ldots, g_n$ denote the n elements of G and let $p_j := P(G, w_{i-1} = g_j)$. Without loss of generality let $p_1 \leq \ldots \leq p_n$. By the induction hypothesis

$$P(G, w_{i-1} = 1) > P(G, w_{i-1} = g_j) \qquad \forall g_j \neq 1$$

therefore $p_n = P(G, w_{i-1} = 1)$ and

$$p_n > p_j \qquad \forall j \neq n. \tag{2.5}$$

Using (2.4), for any $g \in G$ we may write

$$P(G, w_i = g) = \sum_{j=1}^{n} p_j p_{\sigma(j)}$$

for some $\sigma \in \mathrm{Sym}(n)$.

We wish to show that for any $1 \neq g \in G$

$$P(G, w_i = g) < P(G, w_i = 1).$$

Let $g \neq 1$. By the rearrangement inequality (see Lemma 2.2.11)

$$P(G, w_i = g) = \sum_{j=1}^{n} p_j p_{\sigma(j)} \leq \sum_{j=1}^{n} p_j^2 = P(G, w_i = 1)$$

for some permutation σ. We wish to show that this inequality is strict.

Suppose for a contradiction that this is not the case, and $1 \neq g \in G$ is such that

$$P(G, w_i = g) = P(G, w_i = 1). \tag{2.6}$$

Let $\sigma \in \mathrm{Sym}(n)$ be such that $P(G, w_i = g) = \sum_j p_j p_{\sigma(j)}$.
Since $p_j := P(G, w_{i-1} = g_j)$, from (2.4) we see that $p_{\sigma(j)} = P(G, w_{i-1} = g_j^{-1} g)$. Since $g_n = 1$ it follows that $\sigma(n) \neq n$. Hence $\sigma(n) = r$ and $\sigma(k) = n$ for some $k, r < n$. By (2.5),

$$p_{\sigma(n)} = p_r < p_n = p_{\sigma(k)} \tag{2.7}$$

$$\text{and } p_k < p_n \tag{2.8}$$

therefore

$$(p_{\sigma(k)} - p_{\sigma(n)})(p_n - p_k) > 0 \tag{2.9}$$

$$p_{\sigma(k)} p_n + p_{\sigma(n)} p_k > p_{\sigma(n)} p_n + p_{\sigma(k)} p_k. \tag{2.10}$$

Define $\tau \in \mathrm{Sym}(n)$ to be

$$\tau(i) = \begin{cases} \sigma(n) = r & i = k \\ \sigma(k) = n & i = n \\ \sigma(i) & i \neq k, n \end{cases}$$

(that is, τ is obtained from σ by swapping the values of $\sigma(n)$ and $\sigma(k)$ and keeping the remaining $\sigma(i)$ as before). Then

$$\begin{aligned} \sum_{i=1}^{n} p_i p_{\tau(i)} &= \sum_{i \neq k,n} p_i p_{\sigma(i)} + p_k p_{\tau(k)} + p_n p_{\tau(n)} \\ &= \sum_{i \neq k,n} p_i p_{\sigma(i)} + p_k p_{\sigma(n)} + p_n p_{\sigma(k)} \\ &> \sum_{i \neq k,n} p_i p_{\sigma(i)} + p_n p_{\sigma(n)} + p_k p_{\sigma(k)} \qquad \text{(by (2.10))}. \\ &= \sum_i p_i p_{\sigma(i)} \\ &= \sum_i p_i^2 \quad \text{(by (2.6))}. \end{aligned}$$

This contradicts the rearrangement inequality. Thus for all $g \neq 1$

$$P(G, w_i = g) < P(G, w_i = 1).$$

Since $P(G, w_i = g) = 0$ for all $g \notin \gamma_d(G)$ and $|\gamma_d(G)| = m$ it follows that

$$P(G, w_i = 1) > \tfrac{1}{m}$$

for all $i \in \mathbb{N}$. Since by Proposition 2.2.3 we also have that

$$P(G, w_i = 1) \to \tfrac{1}{m}$$

as i tends to infinity, it follows that m^{-1} is an accumulation point of $S(G, 1)$ (and $S(G, g)$ for some $g \neq 1$) and hence of $S(G)$. □

The following is an open question.

Question Is it true in general that for a finite group G, $w \in F_\infty$ and $g \in G$

$$P(G, w = g) = P(G, w = g^{-1})? \tag{2.11}$$

Of course if g has order 2, or is conjugate to g^{-1}, or is auto-equivalent to g^{-1} (see Definition 4.2.3 in §4.2.3) then (2.11) holds for any word. We suspect that it is true in general.

Suppose that equation (2.11) holds for some word w and finite group G. If there exists some $g \in G_w^+$ such that

$$P(G, w = g) < P(G, w = 1) \tag{2.12}$$

it will follow by a proof analogous to Lemma 2.2.9 that $|w(G)|^{-1}$ is an accumulation point of $S(G, 1)$ and $S(G)$.

Proposition 2.2.3 shows that if we use the concatenation method, the only accumulation points we will find for $S(G)$ are those of the form $|H|^{-1}$, where H is a verbal subgroup of G. In order to prove the existence of other accumulation points, we may use the substitution method.

2.2.2 The substitution method

This is a generalisation of left normed commutators. Given w, a word on k variables, we construct a sequence of words $(w_i)_{i \in \mathbb{N}}$ by repeatedly replacing the first variable with a copy of w. Explicitly,

$$\begin{aligned} w_1(x_1, \ldots, x_k) &:= w(x_1, \ldots, x_k) \\ w_2(x_1, \ldots, x_{2k-1}) &:= w(w(x_1, \ldots, x_k), x_{k+1}, \ldots, x_{2k-1}) \\ w_{i+1}\left(x_1, \ldots, x_{((i+1)(k-1)+k)}\right) &:= w_i\left(w(x_1, \ldots, x_k), x_{k+1}, \ldots, x_{((i+1)(k-1)+k)}\right) \quad \forall i \geq 1. \end{aligned}$$

Example If $w_1(x_1, x_2) = [x_1, x_2]$ then

$$\begin{aligned} w_2(x_1, x_2, x_3) &= w(w(x_1, x_2), x_3) \\ &= [\ [\ x_1, x_2\ ]\ ,\ x_3\ ], \qquad \text{etc,} \end{aligned}$$

so that $w_i = \gamma_{i+1}$, the left normed commutator of length $i+1$.

We will use the substitution method in many of the proofs in Chapter 5.

2.3 Constructions

Here we shall briefly discuss some results regarding direct products, quotient groups and verbal subgroups. These will be of use throughout this thesis, in particular in the calculations in Chapter 3 and for the results in Sections 5.5.1 and 5.5.2.

2.3.1 Direct products

Suppose a finite group G is the direct product of N and M. Since

$$w(n_1m_1, \ldots, n_km_k) = w(n_1, \ldots, n_k)w(m_1, \ldots, m_k)$$

for any $n_i \in N$ and $m_i \in M$, it follows that

$$|w_G^{-1}(nm)| = |w_N^{-1}(n)||w_M^{-1}(m)|.$$

Thus

$$P(G, w = nm) = \frac{|w_G^{-1}(nm)|}{|G|^k} = \frac{|w_N^{-1}(n)|}{|N|^k}\frac{|w_M^{-1}(m)|}{|M|^k}.$$

Hence (as has been noted by authors before, for example in Theorem 1.2 in [28]),

$$P(N \times M, w = nm) = P(N, w = n)P(M, w = m) \tag{2.13}$$

and so

$$S(N \times M) \subseteq S(N)S(M).$$

2.3.2 Quotient groups

Let G be a finite group with normal subgroup N. Since

$$w(g_1N, \ldots, g_kN) = gN \iff w(g_1, \ldots, g_k) \in gN$$

we have

$$|w_{G/N}^{-1}(gN)| = \sum_{n \in N} |w_G^{-1}(gn)| \frac{1}{|N|^k}$$

(we divide by $|N|^k$ since each element $(g_1N, \ldots, g_kN)$ corresponds to $|N|^k$ elements in $G^{(k)}$). Thus

$$\frac{|w_{G/N}^{-1}(gN)|}{|G/N|^k} = \sum_{n \in N} \frac{|w_G^{-1}(gn)|}{|N|^k}\frac{|N|^k}{|G|^k}$$

hence

$$P(G/N, w = gN) = \sum_{n \in N} P(G, w = gn) \tag{2.14}$$

and so

$$S(G/N, gN) \subseteq \sum_{n \in N} S(G, gn).$$

2.3.3 Verbal subgroups

Lemma 2.3.1. *Let G be a finite group with verbal subgroup H. Then $S(H)$ is contained in the closure of $S(G)$. In particular, any limit point of $S(H)$ is also a limit point of $S(G)$, and if $S(H)$ is dense in $[0, 1]$ then so is $S(G)$.*

Proof. Let $a \in S(H)$. Then there exists $h \in H$ and $w \in F_\infty$ such that $P(H, w = h) = a$. Since H is verbal in G, there is a word w^* such that $w^*(G) = H$. Let w_i be the word constructed by concatenating w^* with itself i times (see §2.2.1). We have seen that

$$P(G, w_i = h) \to |H|^{-1} \text{ as } i \to \infty.$$

Now consider the word $\tilde{w}_i$ constructed from w with each variable replaced by w_i. Then $P(G, \tilde{w}_i = h) \to P(H, w = h) = a$ as $i \to \infty$. Since $P(G, \tilde{w}_i = h) \in S(G)$ for all i, a is a limit point of $S(G)$ and thus is in the closure of $S(G)$, and so the first part of the lemma is proved. The latter parts follow immediately. □

2.4 Abelian groups

If G is an abelian group, then for any word w, the associated verbal mapping is necessarily a homomorphism since

$$w(\mathbf{gh}) = w(g_1h_1, \ldots, g_kh_k) = \prod_{j=1}^{l} (g_{i_j}h_{i_j})^{\varepsilon_{ij}} = \prod_{j=1}^{l} {g_{i_j}}^{\varepsilon_{ij}} \prod_{j=1}^{l} {h_{i_j}}^{\varepsilon_{ij}} = w(\mathbf{g})w(\mathbf{h})$$

where $\mathbf{g}$ denotes the k-tuple $(g_1, \ldots, g_k) \in G^{(k)}$ and $\mathbf{h} = (h_1, \ldots, h_k) \in H^{(k)}$.

This means that the image of any word in a finite abelian group G must be a subgroup of G (rather than merely a subset), i.e. $G_w^+ = w(G)$. The following well known result is useful.

Lemma 2.4.1. *Let f be a group homomorphism. Then for any element h in the image of f, there is a bijection between the preimage of h and the kernel of f.*

Proof. Let $h = f(g)$ be in the image of f. Define $\phi : f^{-1}(1) \to f^{-1}(h), a \mapsto ag$. Note that if $a \in f^{-1}(1)$ then $f(ag) = f(a)f(g) = 1h$, so ϕ is well defined, and has inverse $a \mapsto ag^{-1}$. □

Since every verbal mapping over an abelian group is a homomorphism, by Lemma 2.4.1, if G is a finite abelian group, w is a word and g is in the image of w, then

$$|w^{-1}(g)| = |w^{-1}(1)|,$$

and so $P(G, w = g) = P(G, w = 1)$. Thus $P(G, w = g) = |w(G)|^{-1}$. Hence the following result.

Corollary 2.4.2. *Let G be a finite abelian group, $w \in F_\infty$ and $g \in G_w^+$. Then*

$$P(G, w = g) = |w(G)|^{-1} = |G_w^+|^{-1}.$$

Consequently

$$P(G, w = g) \geq |G|^{-1}.$$

Subsequently, the set $S(G)$ can be easily discovered.

Lemma 2.4.3. *Let G be a finite abelian group. Then*

$$S(G) = \left\{|H|^{-1} \mid H \text{ is a verbal subgroup of } G\right\} \cup \{0\}.$$

Of course this means that $S(G)$ must be finite when G is abelian. We shall now see that the reverse implication also holds.

Lemma 2.4.4. *Let G be a finite group. Then $S(G)$ is finite if, and only if, G is abelian.*

Proof. By Lemma 2.4.3 we see that $S(G)$ must be finite for any abelian group G. Suppose now that G is not abelian. By Lemma 2.2.9 we see that $|G'|^{-1}$ is an accumulation point of $S(G)$, and thus $S(G)$ cannot be finite. □

We shall now see that if G is an abelian group, the verbal subgroups of G are exactly the power subgroups of G, i.e. subgroups of the form

$$G^a := \langle g^a \mid g \in G \rangle$$

for some $a \in \mathbb{N}$. Note that if G is abelian, $G^a = \{g^a \mid g \in G\}$.

Lemma 2.4.5. *Let G be a finite abelian group, and w a word. Then w is G-equivalent to a word of the form*

$$w(x_1, \ldots, x_k) = x_1^{r_1} \ldots x_k^{r_k}.$$

Proof. By Hall's collecting process (see [10]), any word $w \in F_k$ is F_k-equivalent to a word of the form

$$x_1^{r_1} \ldots x_k^{r_k} K(x_1, \ldots, , x_k)$$

where $K \in F_k'$. If G is abelian, this is G-equivalent to $x_1^{r_1} \ldots x_k^{r_k}$. □

Once written in this form, finding the verbal subgroup generated by a word w over an abelian group G is easy.

Lemma 2.4.6. *Let G be a finite abelian group and w a word written in the form $w(x_1, \ldots, x_k) = x_1^{r_1} \ldots x_k^{r_k}$. Then $w(G) = G^a := \langle g^a \mid g \in G \rangle$ where a is the greatest common divisor of $\{r_1, \ldots, r_k\}$.*

Proof. By Euclid's algorithm there exist integers $b_1, \ldots, b_k$ such that

$$a = b_1 r_1 + \ldots + b_k r_k.$$

Let $g \in G$. Then

$$g^a = g^{(b_1 r_1 + \ldots + b_k r_k)} = (g^{b_1})^{r_1} \ldots (g^{b_k})^{r_k}.$$

Thus

$$w(g^{b_1}, \ldots, g^{b_k}) = g^a$$

and so $G^a \leq G_w^+$. Now let

$$g = {g_1}^{r_1} \ldots {g_k}^{r_k} \in G_w^+.$$

Since a divides r_i for all i we have

$${g_i}^{r_i} = ({g_i}^{b_i})^a \in G^a$$

for some integer b_i. Thus g is the product of elements of G^a, so $g \in G^a$. Hence $G_w^+ \leq G^a$ and equality follows. □

We shall now show that if G is abelian, $S(G)$ characterises G up to isomorphism. Suppose we are given $S(G)$ for some finite abelian group G, and wish to determine G. By Lemma 2.4.3

$$S(G) = \left\{ |K|^{-1} \mid K \text{ is a verbal subgroup of } G \right\} \cup \{0\}. \tag{2.15}$$

The smallest non-zero element of $S(G)$ must be $|G|^{-1}$, and so the order of G is determined.

Let p be a prime that divides $|G|$. Write $|G| = bp^n$ for some positive integer n and some integer b that is coprime to p. G may be written as the direct product of cyclic p groups and some group of order b. Let

$$G \cong C_{p^{a_1}} \times \ldots \times C_{p^{a_t}} \times G_2$$

where G_2 has order b and $a_i \leq a_{i+1}$ for all i. By Lemma 2.4.4 the verbal subgroups of an abelian groups are exactly the power subgroups. So from (2.15) we can obtain a list of the orders of the power subgroups. Consider those that are divisible by b.

Let n_i be the integer such that the verbal subgroup associated with x^{p^i} has order bp^{n_i}. Since

$$p^{n-n_1} = |G/G^p| = p^t$$

we have that $t = n - n_1$ and so t is determined by $S(G)$.

Let t_i be the number of a_j such that $a_j \geq i$. Then $t_1 = t$, which we have seen is determined by $S(G)$. Now consider x^{p^2}. We have

$$p^{n-n_2} = \left|G/G^{p^2}\right| = p^{t_1+t_2}$$

and so $t_2 = n - (t_1 + n_2)$ is uniquely determined by $S(G)$. Continuing in this manner it is possible to determine t_i and hence a_i for all i. By repeating this process for all prime divisors of G it is possible to determine the summands of G and hence the group is uniquely determined. Hence the following lemma.

Lemma 2.4.7. *Let G and H be finite groups, and let G be abelian. Then $S(G) = S(H)$ if and only if G is isomorphic to H.*

Proof. Suppose $S(G) = S(H)$ for G and H as described in the lemma. By Lemma 2.4.4, $S(G)$ and hence $S(H)$ are finite. Thus H is abelian, and by the process described before this lemma, G and H are uniquely determined by $S(G)$, up to isomorphism. Thus $G \cong H$. The other implication is trivial. □

If we expand our horizons to the world of non-abelian groups, the result of Lemma 2.4.7 fails to hold. We shall see later that if D_8 denotes the dihedral group of order 8 and Q_8 the quaternion group, then $S(G) = S(H)$. Perhaps this should not be surprising. The classification of abelian groups shows that there is only one way to build an abelian group out of its constituent parts. The fact that this is no longer the case for non-abelian groups perhaps indicates that we should expect to find groups that are similar enough to yield the same set of probabilities, whilst not being isomorphic.

2.5 Simple groups

Definition 2.5.1. *A group G is called **just non-solvable** if every proper quotient of G is solvable, but G is not.*

The following theorem is by Miklós Abért, see [1].

Theorem 2.5.2. *Let G be a finite just non-solvable group. Then the set*

$$\{P(G, w = 1) \mid w \in F_\infty\}$$

is dense in [0,1].

Let G be a finite simple group. If G is abelian, G must be cyclic of order p, where p is prime. Then by Lemma 2.4.3

$$S(G) = \left\{0, \tfrac{1}{p}, 1\right\}.$$

If G is non-abelian, it must be just non-solvable, since it has no proper quotient except the trivial group. Thus by Theorem 2.5.2 we have that $S(G, 1)$ and thus $S(G)$ are both dense in $[0, 1]$.

2.6 Verbally simple groups

Definition 2.6.1. *A group G is called **verbally simple** if the only verbal subgroups of G are G itself and the trivial subgroup.*

The following result is by Kovacs and Newman, see [15].

Theorem 2.6.2. *A finite verbally simple group is a direct product of isomorphic simple groups.*

Let G be a finite verbally simple group. If G is abelian, then $S(G)$ is finite (see Lemma 2.4.4). If G is non-abelian, then by Theorem 2.6.2, G is the direct product of isomorphic (non-abelian) simple groups, H. Suppose G is the direct product of n copies of H. By equation (2.13) in Section 2.3.1, we see that for any word w,

$$P(G, w = 1) = P(H, w = 1)^n.$$

Since by Theorem 2.5.2 $S(H, 1)$ is dense in $[0, 1]$, it follows that $S(G, 1)$ is also dense in $[0, 1]$. Hence

Corollary 2.6.3. *Let G be a finite non-abelian verbally simple group. Then $S(G, 1)$ (and hence $S(G)$) is dense in $[0, 1]$.*

Throughout this thesis we shall discuss the following conjecture.

Conjecture 2.6.4. *Let G be a finite group. Then $S(G)$ is dense in $[0, 1]$ if and only if G is non-nilpotent.*

Let G be a finite group, and suppose we wanted to know whether $S(G)$ is dense in $[0, 1]$ or not. If G is verbally simple, then by Lemma 2.4.4 and Corollary 2.6.3, $S(G)$ is dense in $[0, 1]$ if and only if G is non-abelian, in agreement with the conjecture. Suppose now that G is not verbally simple, i.e. G has some proper verbal subgroup H. If $S(H)$ is dense in $[0, 1]$, then by Lemma 2.3.1 $S(G)$ is also dense in $[0, 1]$. Since the verbal subgroup of a verbal subgroup is verbal, G must have a minimal verbal subgroup H (i.e. a verbal subgroup that does not contain another non-trivial verbal subgroup), which must be verbally simple. If H is non-abelian then by Corollary 2.6.3 and Lemma

2.3.1 it follows that $S(H)$ and thus $S(G)$ are dense in $[0,1]$ and again, the conjecture holds in this case. However, if H is an abelian subgroup, then $S(H)$ is finite, and this tells us nothing of whether or not $S(G)$ is dense in $[0,1]$. For example, let Sym(3) be the symmetric group on 3 letters. Sym(3) is not verbally simple – it has one proper verbal subgroup H, namely the subgroup of order 3 containing the two 3-cycles and the identity. Since H is a abelian, $S(H)$ is finite, so we cannot tell from this whether $S(\mathrm{Sym}(3))$ is dense in $[0,1]$ or not. We shall see in §5.1 that $S(\mathrm{Sym}(3))$ is in fact dense in $[0,1]$, though we could not conclude this by looking at the verbal subgroups. Thus this line of attack can only get us so far, though it does mean that we may concentrate our efforts on non-nilpotent groups whose verbal subgroups are abelian.

Chapter 3

Calculating the set of probabilities for nilpotent groups

We have seen in Lemma 2.4.3 that the set of probabilities associated with a finite abelian group G is easily calculated – one simply finds the orders of the verbal subgroups of G. Explicitly, if G is abelian then

$$S(G) = \left\{ |H|^{-1} \mid H \text{ is a verbal subgroup of } G \right\} \cup \{0\}.$$

Suppose now that G is nilpotent but not abelian. We are interested to see what implications this has for $S(G)$, and what similarities and differences there are between the description of $S(G)$ for non-abelian nilpotent groups and abelian groups. We first note that by Lemma 2.4.4, $S(G)$ is necessarily infinite for non-abelian G, so our focus shifts to describing the accumulation points in the set. Lemma 2.2.9 states that for any finite group and any integer $d > 1$, if the left normed commutator γ_d is not trivial over G, then $|\gamma_d(G)|^{-1}$ is an accumulation point of $S(G)$. We may ask ourselves whether there may be other accumulation points however. For nilpotent groups with a straightforward structure, i.e. those of low nilpotency class and order, $S(G)$ may be calculated exactly without too much difficulty. Since nilpotent groups are the direct product of p-groups, and the probabilities associated with direct products are easily dealt with (see section 2.3.1), we focus here on p-groups. In this chapter we shall calculate $S(G)$ for all nilpotent groups of class 2 and order 8, 16 or 27, as well as a few examples of order 32. As far as we are aware, this has not been done before. We begin with a brief outline of the method used.

3.1 The method

For each finite group G of nilpotency class 2 that we consider, the overarching method is the same. We begin by taking an arbitrary word over G, and writing it in standard

form

$$w(x_1, \ldots, x_k) = x_1^{r_1} \ldots x_k^{r_k} \prod_{i<j} [x_i, x_j]^{r_{ij}}\,, \tag{3.1}$$

with $r_i \in \{0, \ldots, m_1\}$ and $r_{ij} \in \{0, \ldots, m_2\}$, where m_1 is the exponent of G and m_2 is the exponent of the derived subgroup G'.

In Section 3.2 we shall prove that if r_i is coprime to $|G|$ for some i, then $P(G, w = g) = |G|^{-1}$ for any g in G (see Lemma 3.2.1). We eliminate this case, and assume henceforth that the r_i are not coprime to $|G|$.

Next we convert the problem into a question about fibres of polynomials over finite fields. We do this by fixing a generating set $h_1, \ldots, h_m$ for G, and replace each x_i in the word by $h_1^{\alpha_{1,i}} \ldots h_m^{\alpha_{m,i}}$. Evaluating a word at some $g_1, \ldots, g_k$ now becomes equivalent to substituting in the appropriate indices $\alpha_{j,i}$. The new expression is simplified by collecting together occurrences of the same element. We are left with a product of the generators with polynomial indices. Thus counting solutions to $w = g$ becomes a question of counting solutions to these polynomial equations over some finite field, depending on the exponent of the group elements concerned.

We count solutions to these polynomials in similar but slightly varied ways. In general, we split it into a few cases, depending on the powers r_i in the original word (see (3.1)). We write the polynomials as matrix equations, then use various results from Section 3.2 to count the solutions. We then use these findings to calculate the probabilities associated with the original word.

In the next section we provide some preliminary results. We shall first include a result that eliminates some words for which the associated probabilities are easily found (i.e. when r_i is coprime to $|G|$ for some i). We then establish some results about counting solutions of matrix equations over finite fields, in order to make the calculations for specific groups more concise.

3.2 Preliminary results

Let G be a finite nilpotent group and w a word on k variables. By Hall's collecting process (see [10]) we may write w in the form

$$w(x_1, \ldots, x_k) = \prod_{i=1}^{k} x_i^{r_i} K(x_1, \ldots, x_k) \tag{3.2}$$

for some integers $r_1, \ldots, r_k$, where $K \in F_k'$ is a commutator word on $x_1, \ldots, x_k$. We say that w is written in **standard form**.

Lemma 3.2.1. *Let G be a finite, nilpotent group. Let w be a word on k variables written in standard form over G as in (3.2). Using the notation of (3.2), if r_i is coprime to $|G|$ for some i, then*

(i) $G_w^+ = G$,

(ii) $P(G, w = g) = |G|^{-1}, \quad \forall g \in G.$

Proof. Let $(1, \ldots, G, \ldots, 1)$ be the subgroup of $G^{(k)}$ consisting of elements whose j^{th} component is 1 for all $j \neq i$. Since r_i is coprime to $|G|$ it follows that $\phi : G \to G,\ g \mapsto g^{r_i}$. Now

$$w(1, \ldots, 1, G, 1, \ldots, 1) = \phi(G) = G,$$

so that $G_w^+ = G$ and thus (i) is proved.

We shall now prove (ii) by induction on c, the nilpotency class of G. Let G be abelian. By part (i), $w(G) = G$ and by Lemma 2.4.2, $P(G, w = g) = |w(G)|^{-1}$. Thus (ii) holds in the case $c = 1$.

Let $c > 1$. Suppose r_i is coprime to $|G|$ for some i. Then r_i is also coprime to $|G/Z|$. Since G/Z has nilpotency class $c - 1$, by the induction hypothesis

$$(G/Z)_w^+ = G/Z \text{ and} \tag{3.3}$$

$$P(G/Z, w = gZ) = |G/Z|^{-1}, \quad \forall g \in G. \tag{3.4}$$

Since r_i must also be coprime to $|Z|$, and Z is abelian, it follows that

$$Z_w^+ = Z. \tag{3.5}$$

Let $z \in Z$. By (3.5) there exists some $\mathbf{z} \in Z^{(k)}$ such that $w(\mathbf{z}) = z$. Define

$$\psi : w^{-1}(g) \to w^{-1}(gz), \qquad \mathbf{g} \mapsto \mathbf{g} \cdot \mathbf{z}.$$

Note that since $\mathbf{z} \in Z^{(k)}$,

$$w(\mathbf{g} \cdot \mathbf{z}) = w(\mathbf{g})w(\mathbf{z}) = gz.$$

Then since ψ is a bijection (it has inverse $\mathbf{g} \mapsto \mathbf{g} \cdot \mathbf{z}^{-1}$), $|w^{-1}(g)| = |w^{-1}(gz)|$ for all z in Z. By (2.14) we have that

$$P(G/Z, w = gZ) = \sum_{z \in Z} P(G, w = gz)$$

and so, using (3.4),

$$P(G, w = g) = \tfrac{1}{|Z|} P(G/Z, w = gZ) \geq \tfrac{1}{|Z|} \tfrac{1}{|G/Z|} = \tfrac{1}{|G|}. \tag{3.6}$$

Since by part (i) the image of w is the whole of G, and $P(G, w = g) \geq |G|^{-1}$ by (3.6), it follows that

$$P(G, w = g) = |G|^{-1}$$

for all g in G. Therefore part (ii) also holds. □

Lemma 3.2.2. *Let $\mathbb{F}_2 = \{0, 1\}$ denote the field of two elements. Let $M := (m_{ij})$ be a symmetric matrix in $\mathbb{F}_2^{k\times k}$. Then there exists a row vector $a \in \mathbb{F}_2^k$ such that*

$$aM = (m_{11}, m_{22}, \ldots, m_{kk}).$$

That is, the vector consisting of the diagonal entries of M is in the row span of M.

Proof. Let σ be the rank of M. If $\sigma = k$ then M has full rank and thus

$$(m_{11}, \ldots, m_{kk})$$

must be in the image of M, and there is nothing to prove.

Suppose then that $\sigma < k$. Let $C := \{m_{i_1}, \ldots, m_{i_\sigma}\}$ be a set of σ linearly independent columns of M. Extend this to a basis for $\mathbb{F}_2^k$

$$B := \{m_{i_1}, \ldots, m_{i_\sigma}, n_{\sigma+1}, \ldots, n_k\}. \tag{3.7}$$

Let N be a matrix whose columns consist of the vectors in B where each m_{i_j} is in the j^{th} column. Since N has full rank there exists some $a \in \mathbb{F}_2^k$ such that

$$aN = (m_{11}, \ldots, m_{kk}),$$

so that

$$am_{i_j} = m_{i_j,i_j} \tag{3.8}$$

for all vectors in C.

Let $1 \leq r \leq k$ be such that m_r is not in C. If m_r is the zero vector then $m_{rr} = 0$ and so

$$am_r = m_{rr}$$

is automatically satisfied. Suppose $m_r \neq 0$. Since m_r is in the span of C and all entries are in $\mathbb{F}_2$ we have

$$m_r = \sum_{l=1}^{s} m_{i_{j_l}} \tag{3.9}$$

for some non-zero subset

$$\{m_{i_{j_1}}, \ldots, m_{i_{j_s}}\} \subseteq C.$$

That is, for each component $m_{b,r}$ of m_r,

$$m_{b,r} = \sum_{l=1}^{s} m_{b,i_{j_l}}. \tag{3.10}$$

Then

$$\begin{aligned} m_{rr} &= \sum_{l=1}^{s} m_{r,i_{j_l}}, && \text{(by equation (3.10))}, \\ &= \sum_{l=1}^{s} m_{i_{j_l},r}, && \text{(since } M \text{ symmetric)}, \\ &= \sum_{l=1}^{s}\sum_{t=1}^{s} m_{i_{j_l},i_{j_t}}, && \text{(by equation (3.10))}. \end{aligned}$$

Note that if $1 \leq a, b \leq s$ and $a \neq b$ then both m_{i_a,i_b} and m_{i_b,i_a} occur exactly once in the above expression. But as M is symmetric, $m_{i_a,i_b} + m_{i_b,i_a} = 2m_{i_a,i_b} = 0$, so that only the diagonal elements contribute to the sum. Thus

$$\begin{aligned} m_{rr} &= \sum_{l=1}^{s} m_{i_{j_l},i_{j_l}} && \text{(by above)}, \\ &= \sum_{l=1}^{s} a m_{i_{j_l}} && \text{(by (3.8))}, \\ &= a\left(\sum_{l=1}^{s} m_{i_{j_l}}\right), \\ &= a m_r, && \text{(by (3.9))}. \end{aligned}$$

Hence for all $i \in \{1, \ldots, k\}$,

$$a m_i = m_{ii},$$

and so

$$aM = (m_{11}, \ldots, m_{kk})$$

as claimed. □

For a vector v we denote the transpose of v by v^t. We denote the transpose of a matrix M by M^T.

Lemma 3.2.3. *Let p be prime and let $a \in \mathbb{F}_p$, where $\mathbb{F}_p$ denotes the field of order p. Let $v \in \mathbb{F}_p^k$ be a non-zero row vector. Define $f : \mathbb{F}_p^k \to \mathbb{F}_p, x \mapsto vx^t$. The number of solutions $x \in \mathbb{F}_p^k$ to the equation*

$$f(x) = vx^t = a$$

is p^{k-1}.

Proof. Note that $f(x+y) = v(x+y)^t = vx^t + vy^t = f(x) + f(y)$. Hence f is a homomorphism from $\mathbb{F}_p^k$ to $\mathbb{F}_p$, and thus has fibres of equal size. Since v is non-zero, $\mathrm{Im}f = \mathbb{F}_p$ and so $|f^{-1}(a)| = |\mathbb{F}_p^k|/p = p^k/p = p^{k-1}$. □

Lemma 3.2.4. *Let* $M = (m_{ij}) \in \mathbb{F}_2^{k\times k}$ *be a symmetric matrix, and*

$$r = (m_{11}, \ldots, m_{kk}) \in \mathbb{F}_2^k$$

be the row vector whose entries are the diagonal entries of M*. Then for any* $x \in \mathbb{F}_2^k$,

$$xMx^t = rx^t.$$

Proof. Let m_i be the i^{th} row of M. Then for all $x = (x_1, \ldots, x_k)$ in $\mathbb{F}_2^k$,

$$xMx^t = \sum_i \sum_j m_{ij} x_i x_j.$$

Now for $i \neq j$, both $m_{ij}x_ix_j$ and $m_{ji}x_jx_i$ appear exactly once, and since M is symmetric $m_{ij}x_ix_j + m_{ji}x_jx_i = 2m_{ij}x_ix_j = 0$. Thus

$$xMx^t = \sum_i m_{ii}x_i^2 = \sum_i m_{ii}x_i = rx^t.$$

So for all x in $\mathbb{F}_2^k$,

$$xMx^t = rx^t.$$

□

Lemma 3.2.5. *Let* $M \in \mathbb{F}_2^{k\times k}$ *be such that*

(i) M *is symmetric.*

(ii) M *has zero diagonal.*

Then M *has even rank.*

Proof. We induct on k, the size of the matrix M. The result is trivially true for $k = 1$, since M must be the zero matrix.

Let $k > 1$. We may write

$$M = \begin{pmatrix} M_{k-1} & b^t \\ b & 0 \end{pmatrix}$$

where b is some row vector in $\mathbb{F}_2^{k-1}$.

Suppose $b^t = M_{k-1}x^t$ for some $x \in \mathbb{F}_2^{k-1}$. By Lemma 3.2.4 $xMx^t = 0$. Thus since

$$xb^t = xM_{k-1}x^t = xr^t = 0,$$

we have

$$\begin{pmatrix} I \\ x \end{pmatrix} M_{k-1}(I \;\; x^t) = \begin{pmatrix} I \\ x \end{pmatrix} (M_{k-1} \;\; b^t) = \begin{pmatrix} M_{k-1} & b^t \\ b & 0 \end{pmatrix} = M.$$

Thus since

$$\text{rank } (AB) \leq \min\{ \text{ rank}(A), \text{rank}(B) \},$$

we have that

$$\text{rank}(M) \leq \text{rank}(M_{k-1}).$$

Hence, since clearly the reverse inequality holds (this is clear when rank is considered to be the number of linearly independent columns of a matrix), we have that $\text{rank}(M) = \text{rank}(M_{k-1})$ is even by induction.

If instead b is linearly independent of the columns of M_{k-1} then

$$\text{rank}(M) = \text{rank}\begin{pmatrix} M_{k-1} \\ b \end{pmatrix} + 1 = \text{rank}(M_{k-1}) + 2$$

must be even by induction. □

We use the above lemma to eliminate a special case in the following result.

Lemma 3.2.6. *Let $M \in \mathbb{F}_2^{k \times k}$ be such that*

(i) M is symmetric.

(ii) Excluding diagonal entries, M has an even number of non-zero entries in each row.

Then M has even rank if, and only if, M has an even number of non-zero entries on its diagonal.

Proof. We induct on k, the size of M, noting that the result is trivially true when $k = 1$.

Now suppose $k > 1$. For any matrix A, let $\text{diag}(A)$ denote the number of diagonal entries in A equal to 1. If $\text{diag}(M) = 0$, then by Lemma 3.2.5, M has even rank and the result holds.

Assume then that $\text{diag}(M) \geq 1$. We wish to show that the parity of the rank is the same as the parity of the number of non-zero diagonal entries of M. To do this we will perform various row and column operations on M that yield a matrix that shares

the same parity of rank and number of non-zero entries, and equivalently show that the result holds for this final matrix. We may assume that $m_{kk} = 1$, else swap the appropriate row and column to ensure this is the case.

Consider the k^{th} row of M. For each $i < k$ such that $M_{ki} = 1$, add row k to row i. Note that by (ii), there are an even number of such elements. Each row must still have an even number of ones off the diagonal, since row k has an odd number of ones (since $M_{kk} = 1$) but since one of these is necessarily M_{ki} which will be added to M_{ii}, only an even number are added to non-diagonal positions. The i^{th} column now has an odd number of off-diagonal ones, since although an even number of ones have been added, one of these is necessarily on the diagonal. Note finally that the k^{th} column of our matrix is now $(0, \ldots, 0, 1)^t$.

Now do the same for the columns, i.e. add the new column k to each of the columns i for which M_{ik} was one in the original matrix. Note that since this merely adds 1 to some off diagonal entry an even number of times, the matrix still has an even number of non-zero off diagonal entries. Thus (ii) holds for our new matrix. Our final matrix $\widetilde{M}$ will by symmetric, since if R_{ij} is the elementary matrix corresponding to adding row j to i then

$$\begin{aligned}
\widetilde{M}^T &= \left(R_{k,i_1} \ldots R_{k,i_s} M R_{k,i_1}^T \ldots R_{k,i_s}^T\right)^T, \\
&= \left(R_{k,i_s} \ldots R_{k,i_1} M R_{k,i_1}^T \ldots R_{k,i_s}^T\right)^T, \\
&= R_{k,i_s} \ldots R_{k,i_1} M R_{k,i_1}^T \ldots R_{k,i_s}^T, \\
&= \widetilde{M}.
\end{aligned}$$

Note that the parity of $\text{diag}(\widetilde{M})$ is the same as $\text{diag}(M)$, since 1 was added to some diagonal entry an even number of times.

Now write

$$\widetilde{M} = \begin{pmatrix} \widetilde{M}_{k-1} & \tilde{b}^t \\ \tilde{b} & 1 \end{pmatrix}, \tag{3.11}$$

where $\tilde{b}$ is now the zero vector. Then

$$\begin{aligned}
\text{rank}(M) &= \text{rank}(\widetilde{M}), \\
&= \text{rank}(\widetilde{M}_{k-1}) + 1, \\
&\equiv \text{diag}(\widetilde{M}_{k-1}) + 1 \qquad \text{mod } 2, \qquad\qquad \text{(by induction)}, \\
&= \text{diag}(\widetilde{M}), \\
&= \text{diag}(M),
\end{aligned}$$

and thus the result is shown. □

Lemma 3.2.7. *Let $M = (m_{ij})$ be a symmetric matrix in $\mathbb{F}_2^{k\times k}$, and let r be the vector whose components are the diagonal entries of M. Define*

$$W_1 := \{\, x \in \mathbb{F}_2^k \mid xM + r = 0 \,\},$$
$$W_2 := \{\, x \in \mathbb{F}_2^k \mid rx^t = 0 \,\}.$$

Then either $W_1 \cap W_2 = \emptyset$ or $W_1 \subseteq W_2$.

Proof. Suppose $W_1 \cap W_2 \neq \emptyset$. Let $v_0 \in W_1 \cap W_2$. Then

$$W_1 = v_0 + V_1,$$

where $V_1 = \{x \in \mathbb{F}_2^k \mid xM = 0\}$.

Let $w_1 \in W_1$, so $w_1 = v_0 + v_1$ for some $v_1 \in V_1$. Then

$$\begin{aligned}
rw_1^t &= r\,(v_0 + v_1)^t, \\
&= rv_0^t + rv_1^t, \\
&= rv_1^t, && \text{(since } v_0 \in W_2\text{)}, \\
&= v_1 M v_1^t, && \text{(by Lemma 3.2.4)}, \\
&= 0v_1^t, && \text{(since } v_1 \in V_1\text{)}, \\
&= 0.
\end{aligned}$$

Thus $w_1 \in W_2$, and we have shown that $W_1 \subseteq W_2$. □

Lemma 3.2.8. *Define M, r, W_1, W_2 as in Lemma 3.2.7. Then $W_1 \subseteq W_2$ if, and only if, the rank of M is even.*

Proof. We proceed by induction on k, where $M \in \mathbb{F}_2^{k\times k}$. Suppose $k = 1$. Then $M = 0$ or $M = 1$.

If $M = r = 0$ then the rank of M is even and

$$W_1 = \{x \in \mathbb{F}_2 \mid x.0 + 0 = 0\} = \mathbb{F}_2$$
$$W_2 = \{x \in \mathbb{F}_2 \mid 0.x = 0\} = \mathbb{F}_2.$$

Therefore $W_1 \subseteq W_2$ which agrees with the lemma.

Suppose $r = M = 1$. Then

$$W_1 = \{x \in \mathbb{F}_2 \mid x.1 + 1 = 0\} = \{1\} \quad \text{and}$$
$$W_2 = \{x \in \mathbb{F}_2 \mid 1.x = 0\} = \{0\}.$$

Thus $W_1 \cap W_2 = \emptyset$, and the lemma holds for $k = 1$.

Suppose now that $k > 1$. We consider two cases.

Case i. Suppose that W_1 contains an element w which has a zero component. Without loss of generality suppose that this is the k^{th} component of w. Otherwise, swap the appropriate row and column of M with the k^{th} row and column and proceed as follows.

Let $\widetilde{M} \in \mathbb{F}_2^{(k-1)\times(k-1)}$ be the matrix obtained by removing the last row and column of the matrix M. So there must be some row vector $b \in \mathbb{F}_2^{k-1}$ such that we may write

$$M = \begin{pmatrix} \widetilde{M} & b^t \\ b & m_{kk} \end{pmatrix}. \tag{3.12}$$

For any vector $v \in \mathbb{F}_2^k$ let $\tilde{v} \in \mathbb{F}_2^{k-1}$ be the vector obtained from v by removing the k^{th} component.

Let $\widetilde{W}_1, \widetilde{W}_2$ be analogous to W_1, W_2 with M and r replaced by $\widetilde{M}$ and $\tilde{r}$. Explicitly,

$$\widetilde{W}_1 = \{x \in \mathbb{F}_2^{k-1} \mid x\widetilde{M} + \tilde{r} = 0\},$$
$$\widetilde{W}_2 = \{x \in \mathbb{F}_2^{k-1} \mid \tilde{r}x^t = 0\}.$$

Claim 3.2.9. $\tilde{w} \in \widetilde{W}_1$.

Proof of claim. Since by definition, w is in W_1, we have $wM = r$. Therefore,

$$\begin{aligned}(\tilde{r}, m_{kk}) &= r = wM, \\ &= (\tilde{w}, 0)M, \\ &= (\tilde{w}, 0)\begin{pmatrix} \widetilde{M} & b^t \\ b & m_{kk} \end{pmatrix}, \qquad \text{where } b \text{ is as given in (3.12)}, \\ &= (\tilde{w}\widetilde{M}, \tilde{w}b^t).\end{aligned}$$

Thus $\tilde{r} = \tilde{w}M$ i.e. $\tilde{w} \in \widetilde{W}_1$ and the claim holds. □

Note also that

$$rw^t = \tilde{r}\tilde{w}^t + 0 = \tilde{r}\tilde{w}^t.$$

So $rw^t = 0$ if, and only if, $\tilde{r}\tilde{w}^t = 0$. Thus $w \in W_2$ if, and only if $\tilde{w} \in \widetilde{W}_2$.

Hence if $W_1 \subseteq W_2$ then $w \in W_1 \cap W_2$ and by the above $\tilde{w} \in \widetilde{W}_1 \cap \widetilde{W}_2$. Thus since $\widetilde{W}_1 \cap \widetilde{W}_2$ is non-empty, by Lemma 3.2.7, $\widetilde{W}_1 \subseteq \widetilde{W}_2$.

On the other hand if $W_1 \cap W_2 = \emptyset$, we must have $w \in W_1$ but $w \notin W_2$. Then $\tilde{w} \in \widetilde{W}_1$ and $\tilde{w} \notin \widetilde{W}_2$. Thus $\widetilde{W}_1 \not\subseteq \widetilde{W}_2$ and so $\widetilde{W}_1 \cap \widetilde{W}_2 = \emptyset$.

Hence

$$W_1 \cap W_2 = \emptyset \qquad \Longleftrightarrow \qquad \widetilde{W_1} \cap \widetilde{W_2} = \emptyset.$$

Since $\widetilde{M} \in \mathbb{F}_2^{(k-1)\times(k-1)}$, by induction $\widetilde{M}$ satisfies the lemma, i.e. $\widetilde{W_1}$ is contained in $\widetilde{W_2}$ if, and only if, the rank of $\widetilde{M}$ is even.

Since $W_1 \subseteq W_2$ if, and only if, $\widetilde{W_1} \subseteq \widetilde{W_2}$, it remains to show that rank(M) and rank$(\widetilde{M})$ have the same parity.

As above, write

$$M = \begin{pmatrix} \widetilde{M} & b^t \\ b & m_{kk} \end{pmatrix}.$$

If b is not in the span of the columns of $\widetilde{M}$ then

$$\text{rank}(M) = \text{rank} \begin{pmatrix} \widetilde{M} \\ b \end{pmatrix} + 1 = \text{rank}(\widetilde{M}) + 2.$$

Thus rank(M) and rank $(\widetilde{M})$ have the same parity as required.

Suppose finally that b *is* in the span of the columns of $\widetilde{M}$. We wish to show that $\binom{b^t}{m_{kk}}$ is in the span of the columns of $\binom{\widetilde{M}}{b}$.

Suppose for a contradiction that this is not the case. Since b is in the span of $\widetilde{M}$ we would have

$$b = \sum_{i=1}^{k-1} \lambda_i \widetilde{m_i}, \tag{3.13}$$

for some $\lambda_i \in \mathbb{F}_2$. Componentwise this is

$$b_j = \sum_{i=1}^{k-1} \lambda_i m_{ij}. \tag{3.14}$$

But if $\binom{b}{m_{kk}}$ is not in the span of the columns of $\binom{\widetilde{M}}{b^t}$ then

$$m_{kk} = \sum_{i=1}^{k-1} \lambda_i b_i + 1. \tag{3.15}$$

Then

$$\begin{aligned}
m_{kk} - 1 &= \sum_{i=1}^{k-1} \lambda_i b_i && \text{(by (3.15))}, \\
&= \sum_{i=1}^{k-1} \lambda_i \left(\sum_{j=1}^{k-1} \lambda_j m_{ji} \right) && \text{(by (3.14))}, \\
&= \sum_{i=1}^{k-1} \lambda_i^2 m_{ii} && (\lambda_i \lambda_j m_{ij} + \lambda_j \lambda_i m_{ji} = 0 \quad \forall i \neq j), \\
&= \sum_{i=1}^{k-1} \lambda_i m_{ii} \\
&= \sum_{i=1}^{k-1} \lambda_i (\tilde{w} \tilde{m}_i) && \text{(since } \tilde{w} \in \widetilde{W}_1) \\
&= \tilde{w} \left(\sum_{i=1}^{k-1} \lambda_i \tilde{m}_i \right) \\
&= \tilde{w} \tilde{b} = wM_k = m_{kk}.
\end{aligned}$$

This contradiction shows that $\binom{b}{m_{kk}}$ must be in the span of the columns of $\binom{\widetilde{M}}{b^t}$. Therefore

$$\operatorname{rank}(M) = \operatorname{rank}\begin{pmatrix} \widetilde{M} \\ b^t \end{pmatrix} = \operatorname{rank}(\widetilde{M}).$$

Thus in either case, the parity of $\operatorname{rank}(M)$ and $\operatorname{rank}(\widetilde{M})$ is the same.

Case ii. Suppose no element of W_1 has a zero entry. By Lemma 3.2.2, r is in the row span of M, so W_1 must be non-empty. Thus we must have that $W_1 = \{(1, \ldots, 1)\}$. Since $(1, \ldots, 1)M = (m_{11}, \ldots, m_{kk})$ we have

$$\begin{aligned}
& (1, \ldots, 1) m_i = m_{ii}, \\
\iff \quad & \sum_{j=1}^{k} m_{ji} = m_{ii}, \\
\iff \quad & \sum_{j \neq i} m_{ji} = 0.
\end{aligned}$$

Thus excluding diagonal entries, M has an even number of non-zero entries in each column (or row). Now W_1 is contained in W_2 if, and only if, $(1, \ldots, 1) r^t = 0$. This happens exactly when M has an even number of non-zero entries on its diagonal.

By Lemma 3.2.6, if M has an even number of non-zero entries off the diagonal in every column, then $\operatorname{rank}(M)$ is even if, and only if, r has an even number of non-zero entries. Thus the result holds in this case.

Thus $W_1 \subseteq W_2$ if, and only if, the rank of M is even. $\square$

Lemma 3.2.10. *Let $M = (m_{ij}) \in \mathbb{F}_2^{k\times k}$ be a symmetric matrix and let r be the vector whose components are the diagonal entries of M. Let $p : \mathbb{F}_2^k \times \mathbb{F}_2^k \to \mathbb{F}_2$ be given by*

$$p(\beta, \gamma) = \beta M \gamma^t + r\beta^t + r\gamma^t.$$

Then $|p^{-1}(0)| = 2^{2k-1} + 2^{2k-\sigma-1}$ when σ, the rank of M is even and $|p^{-1}(0)| = 2^{2k-1} - 2^{2k-\sigma-1}$ when σ is odd.

Proof. Let

$$W_1 := \{\ x \in \mathbb{F}_2^k \mid xM + r = 0\ \},$$
$$W_2 := \{\ x \in \mathbb{F}_2^k \mid rx^t = 0\ \}.$$

We know from Lemma 3.2.2 that r is in the rowspan of M. Let $f : \mathbb{F}_2^k \to \mathbb{F}_2^k$ be given by $f(x) := xM$. Then f is a homomorphism whose image has size 2^σ, where σ is the rank of M. Hence

$$|W_1| = |\{\ x \mid xM = r\ \}| = 2^k/2^\sigma = 2^{k-\sigma}.$$

Let γ_0 be in W_1. Then for all β in $\mathbb{F}_2^k$,

$$p(\beta, \gamma_0) = (\gamma_0 M + r)\beta^t + r\gamma_0^t = r\gamma_0^t.$$

Then (β, γ_0) is a solution if, and only if, $r\gamma_0^t = 0$, i.e. if γ_0 is also in W_2.

We therefore wish to know the size of the set $W := W_1 \cap W_2$. By Lemma 3.2.7, W is either empty or is equal to W_1, and thus contains 0 or $2^{k-\sigma}$ elements. In the latter situation, if γ_0 is in W, then (β_0, γ_0) is a solution for any $\beta_0 \in \mathbb{F}_2^k$. In this case there are then $2^{k-\sigma}2^k = 2^{2k-\sigma}$ solutions of this form. So there are either 0 or $2^{2k-\sigma}$ solutions of this sort.

Now fix γ_0 to be any of the $(2^k - 2^{k-\sigma})$ vectors such that $\gamma_0 M + r \neq 0$, i.e. $\gamma_0 \notin W_1$. Let $v_0 := \gamma_0 M + r \neq 0$. Then by Lemma 3.2.3 $f(\beta) := v_0\beta^t = r\gamma_0^t$ has 2^{k-1} solutions. For each of these vectors β_0 we have

$$f(\beta_0, \gamma_0) = (\gamma_0 M + r)\beta_0^t + r\gamma_0^t = v_0\beta_0^t + r\gamma_0^t = 2r\gamma_0^t = 0.$$

So we have $(2^k - 2^{k-\sigma})2^{k-1} = 2^{2k-1} - 2^{2k-\sigma-1}$ solutions here.

Thus if $W = \emptyset$, i.e. when M has odd rank (see Lemma 3.2.8) the number of solutions in total is

$$0 + 2^{2k-1} - 2^{2k-\sigma-1} = 2^{2k-1} - 2^{2k-\sigma-1}.$$

If $W \neq \emptyset$, i.e. when M has even rank (see Lemma 3.2.8) then the number of solutions is

$$2^{2k-\sigma} + 2^{2k-1} - 2^{2k-\sigma-1} = 2^{2k-1} + 2^{2k-\sigma-1}(2-1) = 2^{2k-1} + 2^{2k-\sigma-1}$$

and the claim is proved. □

We shall now explicitly determine $S(G)$ for some small finite nilpotent groups, beginning with D_8. We will use the method set out in Section 3.1.

3.3 The dihedral group of order 8, D_8

Let D_8 be the dihedral group of order 8, given by the presentation

$$D_8 := \langle a, b \mid a^4 = b^2 = 1, bab = a^{-1} \rangle . \tag{3.16}$$

D_8 has exponent 4. The subgroup $\langle a^2 \rangle$ which has order 2 is both the derived subgroup and the centre of D_8. This subgroup may also be described as $\langle g^2 \mid g \in D_8 \rangle$ and is the only proper verbal subgroup of D_8.

Theorem 3.3.1. *Let D_8 denote the dihedral group of order 8. Then*

$$S(D_8) = \left\{\tfrac{1}{8}\right\} \cup \left\{\tfrac{1}{2} \pm (\tfrac{1}{2})^n \mid n \in \mathbb{N}\right\} .$$

Furthermore,

$$\begin{aligned} S(D_8, 1) &= \left\{\tfrac{1}{8}\right\} \cup \left\{\tfrac{1}{2} + (\tfrac{1}{2})^n \mid n \in \mathbb{N}\right\} , \\ S(D_8, a^2) &= \left\{\tfrac{1}{8}\right\} \cup \left\{\tfrac{1}{2} - (\tfrac{1}{2})^n \mid n \in \mathbb{N}\right\} , \\ S(D_8, g) &= \left\{0, \tfrac{1}{8}\right\} , \quad \forall g \notin \langle a^2 \rangle . \end{aligned}$$

Proof. Let w be a word on k variables. Since D_8 is nilpotent of class 2, has exponent 4 and has derived subgroup of exponent 2, we may write $w : D_8^k \to D_8$ in standard form as

$$w(x_1, \ldots, x_k) = x_1^{r_1} \ldots x_k^{r_k} \prod_{i<j} [x_i, x_j]^{r_{ij}} ,$$

where $r_i \in \{0, 1, 2, 3\}$ and $r_{ij} \in \{0, 1\}$ for all i, j.

Case i. Suppose that r_i is odd for some $i \in \{1, \ldots, k\}$. By Lemma 3.2.1 we have that for all g in D_8,

$$P(D_8, w = g) = |D_8|^{-1} = \tfrac{1}{8}.$$

Case ii. The only other possibility is that $r_i \in \{0, 2\}$ for all i. Since

$$\langle g^2 \mid g \in D_8 \rangle = Z(D_8) = (D_8)' = \langle a^2 \rangle ,$$

it follows that the image of w lies in $\langle a^2 \rangle$.

Using the presentation for D_8 given in (3.16) we may write any $g \in G$ uniquely in the form $g = b^\beta a^\alpha$ where $\beta \in \{0,1\}, \alpha \in \{0,1,2,3\}$. In the expression for w, we replace each x_i with $b^{\beta_i} a^{\alpha_i}$.

Now

$$\begin{aligned}(b^{\beta_i} a^{\alpha_i})^2 &= b^{\beta_i} a^{\alpha_i} b^{\beta_i} a^{\alpha_i}, \\ &= b^{2\beta_i} a^{2\alpha_i} [b,a]^{\alpha_i \beta_i}, \\ &= a^{2\alpha_i + 2\beta_i \alpha_i}, \qquad (\text{as } b^2 = 1, [b,a] = a^2).\end{aligned}$$

So if $r_i \in \{0,2\}$ we have

$$(b^{\beta_i} a^{\alpha_i})^{r_i} = a^{r_i \alpha_i (1+\beta_i)}.$$

Also

$$\left[b^{\beta_i} a^{\alpha_i}, b^{\beta_j} a^{\alpha_j} \right] = [b,a]^{\beta_i \alpha_j + \beta_j \alpha_i} = a^{2(\beta_i \alpha_j + \alpha_j \beta_i)}.$$

Thus

$$w(b^{\beta_1} a^{\alpha_1}, \ldots, b^{\beta_k} a^{\alpha_k}) = a^{\sum_{i=1}^k r_i \alpha_i (1+\beta_i)} (a^2)^{\sum_{i<j} r_{ij} (\beta_i \alpha_j + \beta_j \alpha_i)}. \tag{3.17}$$

For all i, define $r_{ii} := \frac{r_i}{2} \in \{0,1\}$. Then equation (3.17) becomes

$$w(b^{\beta_1} a^{\alpha_1}, \ldots, b^{\beta_k} a^{\alpha_k}) = (a^2)^{\sum_{i=1}^k r_{ii} \alpha_i (1+\beta_i) + \sum_{i<j} r_{ij} (\beta_i \alpha_j + \beta_j \alpha_i)}.$$

Now in order to count $\{ (g_1, \ldots, g_k) \in G^{(k)} \mid w(g_1, \ldots, g_k) = 1 \}$ it suffices to count

$$\{ (\alpha, \beta) \in A \times B \mid \sum_{i=1}^k r_{ii} \alpha_i (1+\beta_i) + \sum_{i<j} r_{ij} (\beta_i \alpha_j + \beta_j \alpha_i) \text{ is even } \},$$

where

$$A = \{ \alpha := (\alpha_1, \ldots, \alpha_k) \mid \alpha_i \in \{0,1,2,3\} \},$$

$$B = \{ \beta := (\beta_1, \ldots, \beta_k) \mid \beta_i \in \{0,1\} \}.$$

Since a^2 has order 2, only the parity of the expression is important. This means we need only consider values of α_i in $\{0,1\}$, so long as we remember to account for this later. Hence we calculate the size of the fibres of the map $p : \mathbb{F}_2^k \times \mathbb{F}_2^k \longrightarrow \mathbb{F}_2$ given by

$$p(\alpha, \beta) := \sum_{i=1}^k r_{ii} \alpha_i (1+\beta_i) + \sum_{i<j} r_{ij} (\beta_i \alpha_j + \beta_j \alpha_i).$$

Let $r := (r_{11}, \ldots, r_{kk})$ and

$$M := \begin{pmatrix} r_{11} & r_{12} & \ldots & r_{1k} \\ r_{12} & r_{22} & \ldots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{1k} & r_{2k} & \ldots & r_{kk} \end{pmatrix}.$$

Then we may write $p(\alpha, \beta) = (\beta M + r)\alpha^t$ and count solutions to

$$p(\alpha, \beta) = (\beta M + r)\alpha^t = 0. \tag{3.18}$$

By Lemma 3.2.2, r is in the row span of M, i.e. $\beta_0 M = r$ for some $\beta_0 \in \mathbb{F}_2$. Since

$$W_1 := \{\beta \in \mathbb{F}_2^k \mid \beta M = r\} = \beta_0 + \{\beta \in \mathbb{F}_2^k \mid \beta M = 0\},$$

W_1 must contain $2^{k-\sigma}$ elements, where σ denotes the rank of M. For each of these $2^{k-\sigma}$ vectors β such that $(\beta M + r) = 0$, along with any $\alpha \in \mathbb{F}_2^k$, (α, β) satisfies equation (3.18). Thus there are $2^{k-\sigma}2^k = 2^{2k-\sigma}$ such pairs (α, β).

For the remaining $2^k - 2^{k-\sigma}$ vectors β such that $(\beta M + r) \neq 0$, by Lemma 3.2.3 there are 2^{k-1} vectors α such that $(\beta M + r)\alpha^t = 0$. Thus there are $(2^k - 2^{k-\sigma})2^{k-1} = 2^{2k-1} - 2^{2k-\sigma-1}$ such pairs.

So in total the number of pairs $(\alpha, \beta) \in \mathbb{F}_2^k \times \mathbb{F}_2^k$ such that $p(\alpha, \beta) = 0$ is

$$2^{2k-\sigma} + 2^{2k-1} - 2^{2k-\sigma-1} = 2^{2k-1} + 2^{2k-\sigma-1}.$$

Thus the number of pairs $(\alpha, \beta) \in \{0, 1\} \times \{0, 1, 2, 3\}$ such that $p(\alpha, \beta)$ is even is

$$2^k(2^{2k-1} + 2^{2k-\sigma-1}).$$

Each of these corresponds to some k-tuple $\mathbf{g} \in G^{(k)}$ such that $w(\mathbf{g}) = 1$.

Therefore

$$P(D_8, w = 1) = \frac{2^k(2^{2k-1} + 2^{2k-\sigma-1})}{8^k} = \tfrac{1}{2} + \left(\tfrac{1}{2}\right)^{\sigma+1}.$$

Since in this case $G_w^+ = \left\{1, a^2\right\}$, it follows that

$$P(D_8, w = a^2) = 1 - \left(\tfrac{1}{2} + (\tfrac{1}{2})^{\sigma+1}\right) = \tfrac{1}{2} - (\tfrac{1}{2})^{\sigma+1}.$$

Since cases i and ii constitute all possibilities we have demonstrated that

$$S(D_8, 1) \subseteq \{\tfrac{1}{8}\} \cup \{\tfrac{1}{2} + (\tfrac{1}{2})^n \mid n \in \mathbb{N}\}, \tag{3.19}$$

$$S(D_8, a^2) \subseteq \{\tfrac{1}{8}\} \cup \{\tfrac{1}{2} - (\tfrac{1}{2})^n \mid n \in \mathbb{N}\}, \tag{3.20}$$

$$S(D_8, g) \subseteq \{0, \tfrac{1}{8}\}, \quad \forall g \notin \langle a^2 \rangle . \tag{3.21}$$

To complete the proof, it suffices to show the reverse inclusions in the above expressions.

First, let $w(x) := x$. Then clearly $P(G, w = g) = \frac{1}{8}$ for all g in D_8. If $w(x) \equiv 1$ is the trivial word, then $P(G, w = 1) = 1$, $P(G, w = g) = 0$ for all $g \neq 1$. Now consider the sequence of words defined by

$$\begin{aligned} w_1(x_1) &:= x_1^2, \\ w_2(x_1, x_2) &:= x_1^2 x_2^2, \\ w_i(x_1, \ldots, x_2) &:= w_{i-1} * x_i^2, \qquad \forall i > 1, \end{aligned}$$

where " $*$ " denotes concatenation.

Claim 3.3.2. *For all $i \in \mathbb{N}$,*

$$P(G, w_i = 1) = \tfrac{1}{2} + (\tfrac{1}{2})^{i+1}.$$

Proof of claim. Since $w_1(a) = w_1(a^3) = a^2$ and all other elements have exponent dividing 2, we have that

$$P(G, w_1 = 1) = \tfrac{6}{8} = \tfrac{3}{4} = \tfrac{1}{2} + (\tfrac{1}{2})^2,$$

and hence the claim holds for $i = 1$.

Suppose now $i > 1$. Since $w_i = w_{i-1} * x_i^2$ we have that

$$\begin{aligned} P(G, w_i = 1) &= P(G, w_{i-1} = 1)P(G, x_i^2 = 1) + P(G, w_{i-1} = a^2)P(G, x_i^2 = a^2), \\ &= (\tfrac{1}{2} + (\tfrac{1}{2})^{i-1+1})\tfrac{3}{4} + (\tfrac{1}{2} - (\tfrac{1}{2})^{i-1+1})\tfrac{1}{4}, \\ &= \tfrac{1}{2} + (\tfrac{1}{2})^i(\tfrac{3}{4} - \tfrac{1}{4}), \\ &= \tfrac{1}{2} + (\tfrac{1}{2})^{i+1}, \end{aligned}$$

as claimed. □

For these words w_i we have $P(D_8, w_i = a^2) = 1 - \left(\frac{1}{2} + (\frac{1}{2})^{i+1}\right) = \frac{1}{2} - (\frac{1}{2})^{i+1}$. Thus we have shown the reverse inclusions, and thus the proof is complete. □

3.3.1 Remarks

In the proof of Theorem 3.3.1 we have actually proved more.

Corollary 3.3.3. *For any word w and any g in $(D_8)_w^+$,*

$$P(D_8, w = 1) \geq P(D_8, w = g).$$

This result (which holds for abelian groups) is not true for all nilpotent groups of class 2. It is seen to fail for the quaternion group Q_8 if one considers the word $w(x) = x^2$.

Corollary 3.3.4.

$$\inf_{w,g} P(D_8, w = g) \;=\; \inf_{w} P(D_8, w = 1) \;=\; |D_8|^{-1}$$

where w varies over F_∞ and g varies over G_w^+.

Corollary 3.3.5. *The only accumulation point of $S(D_8)$ is $\frac{1}{2}$.*

Note that $\frac{1}{2}$ is the reciprocal of the order of the only proper verbal subgroup of D_8.

Consider an abelian group of order 8 whose only proper verbal subgroup has order 2. Then we would have $S(G) = \{0, \frac{1}{8}, \frac{1}{2}, 1\}$. Compare this to $S(D_8) = \left\{0, \frac{1}{8}, 1\right\} \cup \left\{\frac{1}{2} \pm (\frac{1}{2})^n \mid n \in \mathbb{N}\right\}$. One might argue that the similarity here testifies to the fact that D_8 is very close to being abelian. Of course there is exactly one such abelian group, namely $\mathbb{Z}_4 \oplus \mathbb{Z}_2$. Since D_8 is the semidirect product of $\mathbb{Z}_4$ and $\mathbb{Z}_2$, these two groups have very similar structure, so perhaps this should not be surprising in this case.

The above theorem motivates the following.

Conjecture 3.3.6. *Suppose G is a nilpotent group (of class 2). Then every accumulation point of $S(G)$ is of the form $|H|^{-1}$ for some verbal subgroup H of G.*

One might ask whether the accumulation points of $S(G)$ is exactly the set

$$\left\{|H|^{-1} \mid H \text{ is a proper verbal subgroup of } G\right\} \tag{3.22}$$

for every finite group of nilpotency class 2, or indeed every nilpotent group, as is the case for D_8. However, we shall see in §3.7 that there is a finite nilpotent group G of class 2 for which (3.22) is strictly contained in the set of accumulation points of $S(G)$. That is, G has a proper verbal subgroup the order of which does not appear as the reciprocal of an accumulation point.

We shall now apply the same method to the quaternion group – the only other group of order 8 and nilpotency class 2.

3.4 The quaternion group

Let

$$Q_8 := \langle -1, s, t, u \mid (-1)^2 = 1, s^2 = t^2 = u^2 = stu = -1 \rangle \tag{3.23}$$

denote the quaternion group. Then Q_8 is nilpotent of class 2, has order 8 and exponent 4. It has only one proper verbal subgroup, $\langle -1 \rangle = \langle g^2 \mid g \in Q_8 \rangle$, which has order 2 and is also the centre and the derived subgroup.

Theorem 3.4.1.

$$S(Q_8) = \left\{\tfrac{1}{8}\right\} \cup \left\{\tfrac{1}{2} \pm (\tfrac{1}{2})^n \mid n \in \mathbb{N}\right\}.$$

Furthermore,

$$\begin{aligned} S(Q_8, 1) &= \left\{\tfrac{1}{8}\right\} \cup \left\{\tfrac{1}{2} + (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\} \cup \left\{\tfrac{1}{2} - (\tfrac{1}{2})^{2n} \mid n \in \mathbb{N}\right\}, \\ S(Q_8, -1) &= \left\{\tfrac{1}{8}\right\} \cup \left\{\tfrac{1}{2} - (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\} \cup \left\{\tfrac{1}{2} + (\tfrac{1}{2})^{2n} \mid n \in \mathbb{N}\right\}, \\ S(Q_8, g) &= \left\{0, \tfrac{1}{8}\right\}, \quad \forall g \notin \{-1, 1\}. \end{aligned}$$

3.4.1 Remarks

1. $S(Q_8) = S(D_8)$. Perhaps this should not be surprising given their similar structure – both groups are extensions of $\mathbb{Z}_4$ by $\mathbb{Z}_2$. This is in contrast to abelian groups, where $S(G) = S(H)$ if, and only if, G is isomorphic to H, i.e. if G is abelian, $S(G)$ characterises G (see Lemma 2.4.7). A word of caution however – the groups $[16, 3]$ and $[16, 13]$ are both semidirect products of $\mathbb{Z}_4 \times \mathbb{Z}_2$ and $\mathbb{Z}_2$ and yet

   $$S([16, 3]) \neq S([16, 13]).$$

2. The proof of Theorem 3.4.1 tells us that there are infinitely many words that satisfy

   $$P(Q_8, w = g) > P(Q_8, w = 1),$$

   for some group element $g \in Q_8$. This is in contrast to Corollary 3.3.3 which states that for any word w and any element g of D_8, $P(D_8, w = 1) \geq P(D_8, w = g)$. It is unknown to the author which groups satisfy this property.

3.4.2 Proof of Theorem 3.4.1

Proof of Theorem 3.4.1. Let w be a word on k variables. As in the D_8 case, since Q_8 is nilpotent of class 2, has exponent 4 and Q_8' (the derived subgroup) has exponent 2, we may write $w : Q_8^k \to Q_8$ in the standard form

$$w(x_1, \ldots, x_k) = x_1^{r_1} \ldots x_k^{r_k} \prod_{i<j} [x_i, x_j]^{r_{ij}},$$

where $r_i \in \{0,1,2,3\}$ and $r_{ij} \in \{0,1\}$.

Case i. Suppose $r_i \in \{1,3\}$ for some i, then by Lemma 3.2.1, for all g in Q_8,

$$P(G, w = g) = \tfrac{1}{8}.$$

Case ii. Suppose instead that w is such that $r_i \in \{0,2\}$ for all i. Recall the presentation for Q_8,

$$Q_8 := \left\langle -1, s, t, u \mid (-1)^2 = 1, s^2 = t^2 = u^2 = stu = -1 \right\rangle.$$

We may denote any element $g_i \in Q_8$ by

$$g_i = (-1)^{\alpha_i} s^{\beta_i} t^{\gamma_i}$$

for unique $\alpha_i, \beta_i, \gamma_i \in \{0,1\}$.

Let $g_i = (-1)^{\alpha_i} s^{\beta_i} t^{\gamma_i}$ and $g_j = (-1)^{\alpha_j} s^{\beta_j} t^{\gamma_j}$. Then

$$\begin{aligned}
g_i^2 &= (-1)^{\alpha_i} s^{\beta_i} t^{\gamma_i} . (-1)^{\alpha_i} s^{\beta_i} t^{\gamma_i}, \\
&= (-1)^{2\alpha_i} s^{\beta_i} (t^{\gamma_i} s^{\beta_i}) t^{\gamma_i}, && \text{(as } (-1) \text{ is central)}, \\
&= (-1)^{2\alpha_i + \beta_i \gamma_i} s^{2\beta_i} t^{2\gamma_i}, && \text{(as } [s,t] = -1), \\
&= (-1)^{2\alpha_i + \beta_i \gamma_i + \beta_i + \gamma_i}, && \text{(as } s^2 = t^2 = -1), \\
&= (-1)^{\beta_i \gamma_i + \beta_i + \gamma_i}, && \text{(as } (-1)^2 = 1),
\end{aligned}$$

and

$$\begin{aligned}
[g_i, g_j] &= [\ (-1)^{\alpha_i} s^{\beta_i} t^{\gamma_i}, (-1)^{\alpha_j} s^{\beta_j} t^{\gamma_j}\ ], \\
&= [\ s, t\ ]^{(\beta_i \gamma_j + \beta_j \gamma_i)}, \\
&= (-1)^{(\beta_i \gamma_j + \beta_j \gamma_i)}.
\end{aligned}$$

Thus if

$$w(x_1, \ldots, x_k) = x_1^{r_1} \ldots x_k^{r_k} \prod_{i<j} [x_i, x_j]^{r_{ij}}$$

is such that $r_i \in \{0,2\}, r_{ij} \in \{0,1\}$ for all i, j, then writing $x_i = (-1)^{\alpha_i} s^{\beta_i} t^{\gamma_i}$, $r_{ii} := \frac{r_i}{2}$ for all i we have

$$w(x_1, \ldots, x_k) = (-1)^{\sum_i r_{ii}(\beta_i \gamma_i + \beta_i + \gamma_i) + \sum_{i<j} r_{ij}(\beta_i \gamma_j + \beta_j \gamma_i)}.$$

Let $\alpha := (\alpha_1, \ldots, \alpha_k)$, $\beta := (\beta_1, \ldots, \beta_k)$, $\gamma := (\gamma_1, \ldots, \gamma_k)$. Since α has no influence on

the value of w, we define $p : \mathbb{F}_2^{2k} \to \mathbb{F}_2$ by

$$p(\beta, \gamma) = \sum_{i=1}^{k} r_{ii}(\beta_i\gamma_i + \beta_i + \gamma_i) + \sum_{i<j} r_{ij}(\beta_i\gamma_j + \beta_j\gamma_i) \tag{3.24}$$

and wish to calculate $|p^{-1}(0)|$.

Let $r := (r_{11}, \ldots, r_{kk})$ and

$$M := \begin{pmatrix} r_{11} & \cdots & r_{1k} \\ \vdots & \ddots & \vdots \\ r_{1k} & \cdots & r_{kk} \end{pmatrix}.$$

Then we may rewrite (3.24) as

$$p(\beta, \gamma) = \beta M\gamma^t + r\beta^t + r\gamma^t.$$

Now by Lemma 3.2.10 we have that $|p^{-1}(0)| = 2^{2k-1} + 2^{2k-\sigma-1}$, when σ is even, and $|p^{-1}(0)| = 2^{2k-1} - 2^{2k-\sigma-1}$ when σ is odd. Since $\alpha \in \{0,1\}^k$ may be chosen arbitrarily, we have

$$P(Q_8, w = 1) = \tfrac{1}{8^k}2^k\left(2^{2k-1} \pm 2^{2k-\sigma-1}\right) = \tfrac{1}{2} \pm (\tfrac{1}{2})^{\sigma+1}. \tag{3.25}$$

Since cases i and ii constitute all possibilities, we have that

$$S(Q_8, 1) \subseteq \{\tfrac{1}{8}\} \cup \{\tfrac{1}{2} + (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{2} - (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N}\} \tag{3.26}$$

$$S(Q_8, -1) \subseteq \{\tfrac{1}{8}\} \cup \{\tfrac{1}{2} - (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{2} + (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N}\} \tag{3.27}$$

$$S(Q_8, g) \subseteq \{0, \tfrac{1}{8}\} \qquad \forall g \neq \pm 1. \tag{3.28}$$

To complete the proof it remains to prove the reverse inclusions to (3.26) – (3.28).

Define the sequence of words w_i by

$$\begin{aligned} w_1(x_1) &:= x_1^2, \\ w_2(x_1, x_2) &:= x_1^2x_2^2, \\ w_i(x_1, \ldots, x_2) &:= w_{i-1} * x_i^2 \quad \forall i > 1, \end{aligned}$$

where " $*$ " denotes concatenation (see §2.2.1).

Claim 3.4.2. *For all i in $\mathbb{N}$,*

$$P(Q_8, w_i = 1) = \begin{cases} \frac{1}{2} + (\frac{1}{2})^{i+1}, & \textit{for } i \textit{ even,} \\ \frac{1}{2} - (\frac{1}{2})^{i+1}, & \textit{for } i \textit{ odd.} \end{cases} \tag{3.29}$$

Proof of claim. We proceed by induction. Since $1^2 = (-1)^2 = 1$ and all other elements in Q_8 square to -1, we have that

$$P(Q_8, w_1 = 1) = \tfrac{1}{4} = \tfrac{1}{2} - (\tfrac{1}{2})^2,$$

and the claim holds for $i = 1$.

Suppose the claim is true for some $k \geq 1$. Since $w_{k+1} = w_k * w_1$ we have

$$\begin{aligned}
P(Q_8, w_{k+1} = 1) &= P(Q_8, w_k = 1)P(Q_8, w_1 = 1) \\
&\quad + P(Q_8, w_k = -1)P(Q_8, w_1 = -1), \\
&= \begin{cases} (\frac{1}{2} + (\frac{1}{2})^{k+1})\frac{1}{4} + (\frac{1}{2} - (\frac{1}{2})^{k+1})\frac{3}{4}, & \text{for } k \text{ even}, \\ (\frac{1}{2} - (\frac{1}{2})^{k+1})\frac{1}{4} + (\frac{1}{2} + (\frac{1}{2})^{k+1})\frac{3}{4}, & \text{for } k \text{ odd}, \end{cases} \\
&= \begin{cases} \frac{1}{2} + (\frac{1}{2})^{k+1}(\frac{1}{4} - \frac{3}{4}), & \text{for } k \text{ even}, \\ \frac{1}{2} + (\frac{1}{2})^{k+1}(-\frac{1}{4} + \frac{3}{4}), & \text{for } k \text{ odd}, \end{cases} \\
&= \begin{cases} \frac{1}{2} - \frac{1}{2}(\frac{1}{2})^{k+1}, & \text{for } k \text{ even}, \\ \frac{1}{2} + \frac{1}{2}(\frac{1}{2})^{k+1}, & \text{for } k \text{ odd}, \end{cases} \\
&= \begin{cases} \frac{1}{2} - (\frac{1}{2})^{k+2}, & \text{for } k \text{ even}, \\ \frac{1}{2} + (\frac{1}{2})^{k+2}, & \text{for } k \text{ odd}, \end{cases}
\end{aligned}$$

and the claim holds for $k + 1$. Thus the claim holds for all i in $\mathbb{N}$. □

Since $w_i(Q_8) = \langle -1 \rangle$, it follows that for all i in $\mathbb{N}$,

$$P(Q_8, w_i = -1) = 1 - P(Q_8, w_i = 1) = \begin{cases} \frac{1}{2} - (\frac{1}{2})^{i+1} & \text{for } i \text{ even}, \\ \frac{1}{2} + (\frac{1}{2})^{i+1} & \text{for } i \text{ odd}. \end{cases} \tag{3.30}$$

Let $\tilde{w}(x) := x$ and $w_0 \equiv 1$. Then

$$P(Q_8, \tilde{w} = g) = \tfrac{1}{8} \quad \forall g \in Q_8, \tag{3.31}$$

$$P(Q_8, w_0 = 1) = 1, \quad P(Q_8, w_0 = -1) = 0. \tag{3.32}$$

Combining (3.29), (3.30), (3.31) and (3.32) we have shown that the reverse inclusions hold and the theorem is proved. □

Remark In (3.25) we accounted for the fact that the variable γ may be arbitrarily chosen by multiplying the number of solutions by 2^k. However, we could instead not do this, but divide by 8^k instead of 16^k. We shall do this in future without comment.

There are only two groups of order 8 that have nilpotency class 2, namely D_8 and Q_8. We shall now calculate $S(G)$ for all groups of nilpotency class 2 and order 16.

3.5 [16,3]

Let G be the group with GAP ID $[16, 3]$, i.e. the group given by the presentation

$$G := \langle a, b, c \mid a^4 = b^2 = c^2 = 1, ab = ba, bc = cb, cac^{-1} = ab \rangle .$$

Then G has 16 elements, is nilpotent of class 2 and has exponent 4. $\langle a, b\rangle$ is isomorphic to $\mathbb{Z}_4 \times \mathbb{Z}_2$, and c acts on $\langle a, b\rangle$ by conjugation by fixing b and sending a to ab. The centre of G is $Z(G) = \langle a^2, b\rangle$, which is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. $G/Z(G)$ is also isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. The commutator subgroup is $G' = \langle b\rangle$ and is isomorphic to $\mathbb{Z}_2$. G has two proper verbal subgroups, G' and $\langle g^2 \mid g \in G\rangle = \langle a^2, b\rangle$ of orders 2 and 4 respectively.

Theorem 3.5.1. *Let $S(G)$ denote the set of probabilities associated with G. Then*

$$S(G) = \{\tfrac{1}{16}, \tfrac{1}{4}\} \cup \{\tfrac{1}{4} \pm (\tfrac{1}{2})^{n+1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{2} \pm (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\}.$$

Furthermore,

$$\begin{aligned}
S(G, 1) &= \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} + (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{4} + (\tfrac{1}{2})^{n+1} \mid n \in \mathbb{N}\},\\
S(G, b) &= \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} - (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{4} - (\tfrac{1}{2})^{n+1} \mid n \in \mathbb{N}\},\\
S(G, a^2) &= S(G, a^2b) = \{0, \tfrac{1}{16}, \tfrac{1}{4}\},\\
S(G, g) &= \{0, \tfrac{1}{16}\} \qquad \forall g \notin \langle a^2, b\rangle .
\end{aligned}$$

Proof. Let w be a word on k variables. Since G is nilpotent of class 2, has exponent 4 and G' has exponent 2, we may write $w : G^k \to G$ in standard form as

$$w(x_1, \ldots, x_k) = x_1^{r_1} \ldots x_k^{r_k} \prod_{i<j} [x_i, x_j]^{r_{ij}} ,$$

where $r_i \in \{0, 1, 2, 3\}$ and $r_{ij} \in \{0, 1\}$ for all i, j.

Suppose $r_i \in \{1, 3\}$ for some i, then by Lemma 3.2.1, if $g \in G$,

$$P(G, w = g) = \tfrac{1}{16}. \tag{3.33}$$

Suppose instead that w is such that r_i is even for every i. Then $G_w^+ \leq \langle a^2, b\rangle$. Note then that

$$P(G, w = g) = 0$$

for every $g \notin \langle a^2, b\rangle$. Every element of G may be uniquely written in the form $a^{\alpha_i} b^{\beta_i} c^{\gamma_i}$

where $\alpha_i \in \{0,1,2,3\}, \beta_i, \gamma_i \in \{0,1\}$.

Claim 3.5.2. *Let $g = a^{\alpha_i} b^{\beta_i} c^{\gamma_i} \in G$ for some $\alpha_i \in \{0,1,2,3\}, \beta_i, \gamma_i \in \{0,1\}$. Then $g^2 = a^{2\alpha_i} b^{\alpha_i \gamma_i}$.*

Proof of Claim 3.5.2. Now

$$(a^{\alpha_i} b^{\beta_i} c^{\gamma_i})^2 = a^{\alpha_i} b^{\beta_i} (c^{\gamma_i} a^{\alpha_i}) b^{\beta_i} c^{\gamma_i} = a^{\alpha_i} b^{\beta_i} a^{\alpha_i} c^{\gamma_i} [c^{\gamma_i}, a^{\alpha_i}] b^{\beta_i} c^{\gamma_i}.$$

Now $[c^{\gamma_i}, a^{\alpha_i}] = [c,a]^{\gamma_i \alpha_i}$, since $G' \leq Z(G)$. From the presentation of G we know that

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

so that $[c,a] = [a,c]^{-1} = b^{-1} = b$. Thus $[c,a]^{\gamma_i \alpha_i} = b^{\alpha_i \gamma_i}$. Thus

$$\begin{aligned}
(a^{\alpha_i} b^{\beta_i} c^{\gamma_i})^2 &= a^{\alpha_i} b^{\beta_i} a^{\alpha_i} c^{\gamma_i} b^{\alpha_i \gamma_i} b^{\beta_i} c^{\gamma_i}, \\
&= a^{2\alpha_i} b^{(2\beta_i + \alpha_i \gamma_i)} c^{2\gamma_i} && (b \in Z(G)), \\
&= a^{2\alpha_i} b^{\alpha_i \gamma_i} && (b^2 = c^2 = 1). \qquad \square
\end{aligned}$$

Claim 3.5.3. *Let $g, h \in G$ be given by*

$$\begin{aligned}
g &= a^{\alpha_i} b^{\beta_i} c^{\gamma_i}, \\
h &= a^{\alpha_j} b^{\beta_j} c^{\gamma_j},
\end{aligned}$$

for some $\alpha_i, \alpha_j \in \{0,1,2,3\}, \beta_i, \beta_j, \gamma_i, \gamma_j \in \{0,1\}$. Then $[g,h] = b^{(\alpha_i \gamma_j + \gamma_i \alpha j)}$.

Proof of Claim 3.5.3.

$$\begin{aligned}
[g,h] &= [a^{\alpha_i} b^{\beta_i} c^{\gamma_i}, a^{\alpha_j} b^{\beta_j} c^{\gamma_j}], \\
&= [a^{\alpha_i} c^{\gamma_i}, a^{\alpha_j} c^{\gamma_j}], && (b \in Z(G)), \\
&= [a,c]^{(\alpha_i \gamma_j + \alpha_j \gamma_i)}, \\
&= b^{(\alpha_i \gamma_j + \alpha_j \gamma_i)}. && \square
\end{aligned}$$

Since $r_i \in \{0,2\}$ for all i, let $r_{ii} := \frac{r_i}{2} \in \{0,1\}$. Let $g_i = a^{\alpha_i} b^{\beta_i} c^{\gamma_i}$ for all i. Then

$$\begin{aligned}
w(g_1, \ldots, g_k) &= \prod_i a^{2 r_{ii} \alpha_i} \prod_{i<j} b^{r_{ij} (\alpha_i \gamma_j + \alpha_j \gamma_i)} \\
&= a^{2 \sum_i r_{ii} \alpha_i} b^{\sum_i r_{ii} \alpha_i \gamma_i + \sum_{i<j} r_{ij} (\alpha_i \gamma_j + \alpha_j + \gamma_i)}.
\end{aligned}$$

Since only the parity of the sums is important, we may restrict $\alpha_i \in \{0,1\}$ for all i, so

long as we account for this later. Hence define $p, q : \mathbb{F}_2^k \times \mathbb{F}_2^k \to \mathbb{F}_2$ by

$$\begin{aligned} p(\alpha, \gamma) &:= \sum_i r_{ii}\alpha_i, \\ q(\alpha, \gamma) &:= \sum_i r_{ii}\alpha_i\gamma_i + \sum_{i<j} r_{ij}(\alpha_i\gamma_j + \alpha_j\gamma_i). \end{aligned} \tag{3.34}$$

To calculate $P(G, w = g)$ where $g = a^{2s}b^t$, we need to count the number of pairs α, γ that satisfy $p(\alpha, \gamma) = s$ and $q(\alpha, \gamma) = t$.

As in the proofs for D_8 and Q_8, let $M \in \mathbb{F}_2^{k\times k}$ be given by

$$m_{ij} = \begin{cases} r_{ij} & i \leq j, \\ r_{ji} & i > j. \end{cases}$$

Let $r = (r_{11}, \ldots, r_{kk})$, $\alpha = (\alpha_1, \ldots, \alpha_k)$ and $\beta = (\beta_1, \ldots, \beta_k)$. Then we may rewrite (3.34) as

$$\begin{aligned} p(\alpha, \gamma) &= r\alpha^t, \\ q(\alpha, \gamma) &= \alpha M \beta^t. \end{aligned} \tag{3.35}$$

Let us first count the solutions to $p = q = 0$ and $p = 0, q = 1$, which correspond to elements in $w^{-1}(1)$ and $w^{-1}(b)$ respectively.

Case i. Suppose w is a commutator word, i.e. r is the zero vector.

Then $p = 0$ is automatically satisfied, and we need to only consider the fibres of q. Let σ be the rank of M. Then there are $2^{k-\sigma}$ vectors α such that $\alpha M = 0$. For any of the 2^k vectors $\beta \in \mathbb{F}_2^k$ we have $q(\alpha, \beta) = 0$. Thus there are $2^{2k-\sigma}$ pairs of this form.

If instead α is one of the $(2^k - 2^{k-\sigma})$ vectors not in the row kernel of M, then $\alpha M \neq 0$, and by Lemma 3.2.3, 2^{k-1} elements β in $\mathbb{F}_2^k$ yield $q(\alpha, \beta) = 0$, and 2^{k-1} yield $q(\alpha, \beta) = 1$.

Hence the number of solutions to $p = q = 0$ is

$$\begin{aligned} 2^{2k-r} + (2^k - 2^{k-r})2^{k-1} &= 2^{2k-1} + 2^{2k-r-1}(2-1) \\ &= 2^{2k-1} + 2^{2k-r-1}. \end{aligned}$$

Since the remaining pairs must be solutions to $p = 0, q = 1$ it follows that the number of such pairs is

$$2^{2k} - (2^{2k-1} + 2^{2k-r-1}) = 2^{2k-1} - 2^{2k-r-1}.$$

Thus in this case,

$$P(G, w = 1) = (2^{2k-1} + 2^{2k-\sigma-1}).2^{-2k} = \tfrac{1}{2} + (\tfrac{1}{2})^{\sigma+1},$$
$$P(G, w = b) = (2^{2k-1} - 2^{2k-\sigma-1}).2^{-2k} = \tfrac{1}{2} - (\tfrac{1}{2})^{\sigma+1}.$$

Note that by Lemma 3.2.5, $\sigma :=$rank(M) is even when $r = 0$.

Case ii. Suppose w is not a commutator word, i.e. r is not the zero vector.

First note that by Lemma 3.2.3 there are 2^{k-1} vectors α for which $r\alpha^t = 0$.

Suppose α is in the row kernel of M, i.e. $\alpha M = 0$. Then by Lemma 3.2.4 we have that $\alpha r^t = 0$. Then $p(\alpha, \beta) = q(\alpha, \beta) = 0$ for any $\beta \in \mathbb{F}_2^k$. Since the kernel of M has $2^{k-\sigma}$, this gives $2^{2k-\sigma}$ solutions to $p = q = 0$.

Suppose now that α is one of the remaining $2^k - 2^{k-\sigma}$ vectors such that $r\alpha = 0$ but $\alpha M \neq 0$. Then again by Lemma 3.2.3, 2^{k-1} vectors β satisfy $\alpha M\beta^t = 0$ and 2^{k-1} vectors β satisfy $\alpha M\beta^t = 1$. So here we have found $2^{k-1}(2^{k-1} - 2^{k-\sigma})$ solutions to $p = q = 0$ and the same number for $p = 0, q = 1$.

Thus in total, the number of solutions to $p = q = 0$ is $2^{2k-\sigma} + 2^{2k-2} - 2^{2k-1-\sigma} = 2^{2k-2} + 2^{2k-\sigma-1}$, and the number of solutions to $p = 0, q = 1$ is $2^{2k-2} - 2^{2k-\sigma-1}$.

Thus in this case

$$\begin{aligned} P(G, w = 1) &= (2^{2k-2} + 2^{2k-\sigma-1})2^{-2k} = \tfrac{1}{4} + (\tfrac{1}{2})^{\sigma+1}, \\ P(G, w = b) &= (2^{2k-2} - 2^{2k-\sigma-1})2^{-2k} = \tfrac{1}{4} - (\tfrac{1}{2})^{\sigma+1}. \end{aligned} \tag{3.36}$$

Combining (3.33) and (3.36) we have that

$$\begin{aligned} S(G, 1) &\subseteq \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} + (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{4} + (\tfrac{1}{2})^{n+1} \mid n \in \mathbb{N}\} \\ S(G, b) &\subseteq \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} - (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{4} - (\tfrac{1}{2})^{n+1} \mid n \in \mathbb{N}\}. \end{aligned} \tag{3.37}$$

To see the reverse inclusions, consider the following:

- Let $w(x) = x$. Then $P(G, w = g) = \frac{1}{16}$ for all $g \in G$.
- Let w be the trivial word. Then $P(G, w = 1) = 1, P(G, w = b) = 0$.
- For all $n \in \mathbb{N}$, let w_n be the word constructed by concatenating n commutators.

  Claim 3.5.4. *Let $(w_n)_{n\in\mathbb{N}}$ denote the sequence of words given by*

$$\begin{aligned} w_1(x_1, x_2) &:= [x_1, x_2], \\ w_{i+1}(x_1, \ldots, x_{2(i+1)}) &:= [x_1, x_2] * w_i(x_3, \ldots, x_{2i+2}) \quad \forall i \geq 1. \end{aligned}$$

where "$*$" denotes concatenation of words. Then for all $i \in \mathbb{N}$,

$$P(G, w_i = 1) = \tfrac{1}{2} + (\tfrac{1}{2})^{2i+1}, \tag{3.38}$$
$$P(G, w_i = b) = \tfrac{1}{2} - (\tfrac{1}{2})^{2i+1}. \tag{3.39}$$

Proof of claim 3.5.4. Since the commutator subgroup of G is $\langle b \rangle$,

$$P(G, w_i = b) = 1 - P(G, w_i = 1).$$

Thus it suffices to prove (3.38), and (3.39) will follow. We prove (3.38) by induction. Since

$$P(G, w_1 = 1) = \tfrac{5}{8} = \tfrac{1}{2} + (\tfrac{1}{2})^3,$$

the lemma holds for $i = 1$. Let $i \geq 1$. Then

$$\begin{aligned} P(G, w_{i+1}) &= P(G, w_1 = 1)P(G, w_i = b) + P(G, w_1 = b)P(G, w_i = 1) \\ &= \tfrac{5}{8}(\tfrac{1}{2} + (\tfrac{1}{2})^{1+2i}) + \tfrac{3}{8}(\tfrac{1}{2} - (\tfrac{1}{2})^{1+2i}) \\ &= \tfrac{1}{2} + (\tfrac{1}{2})^{2(i+1)+1} \end{aligned}$$

and we are done. □

- For all $n \in \mathbb{N}$, let w_n be the word constructed by concatenating n copies of x^2.

  Claim 3.5.5. *Let $(w_n)_{n\in\mathbb{N}}$ denote the sequence of words given by*

$$w_1(x_1) := x_1^2,$$
$$w_{i+1}(x_1, \ldots, x_{i+1}) := x_1^2 * w_i(x_2, \ldots, x_i).$$

  Then for all $i \geq 1$,

$$P(G, w_i = g) = \begin{cases} \frac{1}{4} + (\frac{1}{2})^{i+1}, & g = 1, \\ \frac{1}{4} - (\frac{1}{2})^{i+1}, & g = b, \\ \frac{1}{4}, & g \in \{a^2, a^2b\}. \end{cases} . \tag{3.40}$$

  Proof of claim 3.5.5. We proceed by induction. Since

$$P(G, w_1 = g) = \begin{cases} \frac{1}{2}, & g = 1, \\ 0, & g = b, \\ \frac{1}{4}, & g \in \{a^2, a^2b\}, \end{cases} \tag{3.41}$$

  the claim holds for $i = 1$.

Let $i \geq 2$. Then since $P(G, w_i = z) = \sum_{g \in \mathbb{Z}} P(G, w_1 = g) P(G, w_{i-1} = g^{-1}z)$,

$$\begin{aligned}
P(G, w_i = 1) = &\quad \tfrac{1}{2}\left(\tfrac{1}{4} + \left(\tfrac{1}{2}\right)^i\right) + 2\left(\tfrac{1}{4}\right)^2 + 0 &&= \tfrac{1}{4} + \left(\tfrac{1}{2}\right)^{i+1}, \\
P(G, w_i = b) = &\quad \tfrac{1}{2}\left(\tfrac{1}{4} - \left(\tfrac{1}{2}\right)^i\right) + 2\left(\tfrac{1}{4}\right)^2 + 0 &&= \tfrac{1}{4} - \left(\tfrac{1}{2}\right)^{i+1}, \\
P(G, w_i = a^2) = &\quad \tfrac{1}{2}\tfrac{1}{4} + \tfrac{1}{4}(\tfrac{1}{4} + (\tfrac{1}{2})^i) - \tfrac{1}{4}(\tfrac{1}{4} + (\tfrac{1}{2})^i) + 0 &&= \tfrac{1}{4}, \\
P(G, w_i = a^2 b) = &\quad \tfrac{1}{2}\tfrac{1}{4} + \tfrac{1}{4}\left(\tfrac{1}{4} + \left(\tfrac{1}{2}\right)^i\right) - \tfrac{1}{4}\left(\tfrac{1}{4} + \left(\tfrac{1}{2}\right)^i\right) + 0 &&= \tfrac{1}{4}.
\end{aligned} \tag{3.42}$$

Thus the claim is proved. □

Thus we have that

$$\begin{aligned}
S(G, 1) &= \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} + (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{4} + (\tfrac{1}{2})^{n+1} \mid n \in \mathbb{N}\}, \\
S(G, b) &= \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} - (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{4} - (\tfrac{1}{2})^{n+1} \mid n \in \mathbb{N}\}.
\end{aligned} \tag{3.43}$$

Let us now calculate $S(G, g)$ for $g = a^2$ and $g = a^2 b$. This corresponds to looking at solutions to $p = q = 1$ and $p = 1, q = 0$.

Case i. Suppose w is a commutator word, i.e. r is the zero vector. Then there are no solutions to $p = 1$. So in this case $P(G, w = a^2) = P(G, w = a^2 b) = 0$.

Case ii. Suppose w is not a commutator word, i.e. r is not the zero vector. From Lemma 3.2.4 we know that $\alpha r^t = 1$ implies $\alpha M \neq 0$. There are 2^{k-1} vectors α such that $\alpha r^t = 1$. For each of these, since $\alpha M \neq 0$ there are 2^{k-1} vectors β such that $q(\alpha, \beta) = \alpha M \beta^t = 1$ and 2^{k-1} vectors β such that $q(\alpha, \beta) = \alpha M \beta^t = 0$. Thus there are 2^{k-2} solutions to $p = q = 1$ and the same number to $p = 1, q = 0$. So for a word of this form,

$$P(G, w = a^2) = P(G, w = a^2 b) = \tfrac{1}{4}.$$

So by amalgamating the last two cases along with (3.33) we see that

$$S(G, a^2) = S(G, a^2 b) \subseteq \{0, \tfrac{1}{16}, \tfrac{1}{4}\}.$$

By considering the words $w(x) = x, w(x) = x^2$ and the trivial word, we see that

$$S(G, a^2) = S(G, a^2 b) = \{0, \tfrac{1}{16}, \tfrac{1}{4}\}.$$

Thus our proof is complete.

□

3.6 [16,4]

Let G be the group with presentation

$$G := \langle a, b \mid a^4 = b^4 = 1, bab^{-1} = a^3 \rangle .$$

Then G is a semi-direct product of C_4 and C_4. The centre of G is $\langle a^2, b^2 \rangle \cong V_4$ and the derived subgroup is $\langle a^2 \rangle \cong C_2$. G has two proper verbal subgroups, $\langle a^2, b^2 \rangle = \langle g^2 \mid g \in G \rangle$ and $G' = \langle a^2 \rangle$. G has exponent 4.

Theorem 3.6.1.

$$S(G) = \left\{ \tfrac{1}{16}, \tfrac{1}{4} \right\} \cup \left\{ \tfrac{1}{4} \pm (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N} \right\} \cup \left\{ \tfrac{1}{2} \pm (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N} \right\}.$$

More specifically

$$\begin{aligned}
S(G,1) &= \left\{ \tfrac{1}{16}, \tfrac{1}{4} \right\} \cup \left\{ \tfrac{1}{2} + (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N} \right\} \cup \left\{ \tfrac{1}{4} + (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N} \right\}, \\
S(G,a^2) &= \left\{ \tfrac{1}{16}, \tfrac{1}{4} \right\} \cup \left\{ \tfrac{1}{2} - (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N} \right\} \cup \left\{ \tfrac{1}{4} - (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N} \right\}, \\
S(G,b^2) &= \left\{ 0, \tfrac{1}{16}, \tfrac{1}{4} \right\} \cup \left\{ \tfrac{1}{4} + (\tfrac{1}{2})^{2n} \mid n \in \mathbb{N} \right\}, \\
S(G,a^2b^2) &= \left\{ \tfrac{1}{16}, \tfrac{1}{4} \right\} \cup \left\{ \tfrac{1}{4} - (\tfrac{1}{2})^{2n} \mid n \in \mathbb{N} \right\}, \\
S(G,g) &= \left\{ 0, \tfrac{1}{16} \right\} \ \forall g \notin \langle a^2, b^2 \rangle .
\end{aligned}$$

Proof. We may write any word over G in standard form as

$$w(x_1, \ldots, x_k) = x_1^{r_1} \ldots x_k^{r_k} \prod_{i<j} [x_i, x_j]^{r_{ij}}$$

for some $r_i \in \{0,1,2,3\}, r_{ij} \in \{0,1\}$.

Case i. Suppose r_i is odd for some i.

Then by Lemma 3.2.1, $G_w^+ = G$ and $P(G, w = g) = \frac{1}{16}$ for all g in G.

Case ii. Suppose r_i is even for all i.

By the presentation, we may write any element $g_i \in G$ as $g_i = a^{\alpha_i} b^{\beta_i}$, for some $\alpha_i \in \{0,1,2,3\}$ and $\beta_i \in \{0,1,2,3\}$. The following routine calculations will be needed.

- $[g_i, g_j] = [a^{\alpha_i} b^{\beta_i}, a^{\alpha_j} b^{\beta_j}] = [a,b]^{\alpha_i\beta_j + \alpha_j\beta_i} = a^{2(\alpha_i\beta_j + \alpha_j\beta_i)}$.
- $(g_i)^2 = (a^{\alpha_i} b^{\beta_i})^2 = a^{2\alpha_i} b^{2\beta_i} a^{2\alpha_i\beta_i} = a^{2\alpha_i(1+\beta_i)} b^{2\beta_i}$.
  Thus if $r_i \in \{0,1\}$, $(g_i)^{r_i} = a^{r_i\alpha_i(1+\beta_i)} b^{r_i\beta_i}$.

If we define $\alpha := (\alpha_1, \ldots, \alpha_k)$, $\beta := (\beta_i, \ldots, \beta_k)$, $r_{ii} := \frac{r_i}{2}$ then we may write

$$\begin{aligned} w(\alpha, \beta) &= \prod_i a^{r_i \alpha_i (1+\beta_i)} b^{r_i \beta_i} \prod_{i<j} a^{2 r_{ij} (\alpha_i \beta_j + \alpha_j \beta_i)} \\ &= a^{2(\sum_i r_{ii} \alpha_i (1+\beta_i) + \sum_{i<j} r_{ij} (\alpha_i \beta_j + \alpha_j \beta_i))} b^{2 r_{ii} \beta_i}. \end{aligned}$$

Since we are only interested in the parity of the exponents of a^2 and b^2, we thus consider $p, q : \mathbb{F}_2^k \times \mathbb{F}_2^k \to \mathbb{F}_2$,

$$\begin{aligned} p(\alpha, \beta) &= \sum_i r_{ii} \alpha_i (1 + \beta_i) + \sum_{i<j} r_{ij} (\alpha_i \beta_j + \alpha_j \beta_i), \\ q(\alpha, \beta) &= \sum_i r_{ii} b_i. \end{aligned}$$

If, as per usual we define $r := (r_{11}, \ldots, r_{kk})$, and $M = (m_{ij})_{ij}$ where

$$m_{ij} := \begin{cases} r_{ij}, & i \leq j, \\ r_{ji}, & i > j, \end{cases}$$

then we may write p, q as

$$\begin{aligned} p(\alpha, \beta) &= r \alpha^t + \beta M \alpha^t, \\ q(\alpha, \beta) &= r \beta^t. \end{aligned}$$

Case iia. w is a commutator word, i.e. $r = 0$.

Then we have

$$\begin{aligned} p(\alpha, \beta) &= \beta M \alpha^t, \\ q(\alpha, \beta) &= 0. \end{aligned}$$

We shall first count solutions to $p = q = 0$.

- If $\alpha \in \mathrm{Ker}(M)$, (there are $2^{k-\sigma}$ such α), then any $\beta \in \mathbb{F}_2^k$ yields a solution. There are therefore $2^{2k-\sigma}$ such pairs.
- If $\alpha \notin \mathrm{Ker}(M)$, (there are $2^k - 2^{k-\sigma}$ such α), since $M\alpha^t \neq 0$, by Lemma 3.2.3, there are $2^{k-1}\beta$ such that $\beta M \alpha^t = 0$.

Therefore in total there are $2^{2k-\sigma} + (2^k - 2^{k-\sigma}) 2^{k-1} = 2^{2k-1} - 2^{2k-\sigma-1}$ solutions.

Thus if w is a commutator word,

$$P(G, w = g) = \begin{cases} \frac{1}{2} + (\frac{1}{2})^{\sigma+1}, & g = 1, \\ \frac{1}{2} - (\frac{1}{2})^{\sigma+1}, & g = a^2, \\ 0, & g \notin \langle a^2 \rangle, \end{cases}$$

where $\sigma = \text{rank}(M)$, which must be even by Lemma 3.2.5.

Case iib. Suppose w is not a commutator word, i.e. $r \neq 0$.

Recall we are counting solutions to

$$p(\alpha, \beta) = \alpha(r^t + M\beta^t),$$
$$q(\alpha, \beta) = \beta r^t.$$

We shall consider the cases where σ is even and odd separately. We will make reference to the following subsets of $\mathbb{F}_2^k$,

$$W_1 := \{x \in \mathbb{F}_2^k \mid Mx^t = r\},$$
$$W_2 := \{x \in \mathbb{F}_2^k \mid rx^t = 0\}.$$

Recall from Lemma 3.2.7 and Lemma 3.2.8 that $W_1 \subseteq W_2$ if $\sigma := \text{rank}(M)$ is even, and $W_1 \cap W_2 = \emptyset$ if $\text{rank}(M)$ is odd.

- Suppose σ is even. Let us first count solutions to $q = p = 0$ and $q = 0, p = 1$. By Lemma 3.2.3, since $r \neq 0$, there are 2^{k-1} solutions to $r\beta^t = 0$. We wish to know how many of these β also satisfy $M\beta^t + r = 0$. By Lemma 3.2.7 $W_1 \subseteq W_2$, so all of the $2^{k-\sigma}$ vectors β such that $M\beta^t = r$ also satisfy $r\beta^t = 0$. Thus for any such β and any $\alpha \in \mathbb{F}_2^k$ we have $p(\alpha, \beta) = 0$. There are $2^{2k-\sigma}$ such pairs.

  For the remaining $2^{k-1} - 2^{k-\sigma}$ vectors β such that $r\beta^t = 0$ but $M\beta^t \neq r$, by Lemma 3.2.3, there are 2^{k-1} vectors α such that $\alpha(r^t + M\beta^t) = 0$ and 2^{k-1} such that $\alpha(r^t + M\beta^t) = 1$.

  Thus there are $(2^{k-1} - 2^{k-\sigma})2^{k-1} = 2^{2k-2} - 2^{2k-\sigma-1}$ solutions to $p = q = 0$ and the same for $q = 0, p = 1$.

  In total, $p = q = 0$ has $2^{2k-\sigma} + 2^{2k-2} - 2^{2k-\sigma-1} = 2^{2k-2} + 2^{2k-\sigma-1}$ solutions.

  Thus if w is a non-commutator words, $\sigma =$rank(M) is even and,

$$P(G, w = g) = \begin{cases} \frac{1}{4} + (\frac{1}{2})^{\sigma+1}, & g = 1, \\ \frac{1}{4} - (\frac{1}{2})^{\sigma+1}, & g = a^2. \end{cases}$$

Let us now count solutions to $q = 1$ and $p = 0$ and $q = p = 1$. By Lemma 3.2.3 there are 2^{k-1} elements β such that $r\beta^t = 1$, giving $q = 1$. Since $W_1 \subseteq W_2$, it follows that all such β satisfy $(M\beta^t + r) \neq 0$. Thus for each such β, there are 2^{k-1} values of α that yield $p(\alpha, \beta) = 0$ and the same number yield $p(\alpha, \beta) = 1$. Thus if w is a non-commutator word, M has even rank and

$$P(G, w = a^2b^2) = P(G, w = b^2) = \tfrac{1}{4}.$$

- Suppose σ is odd. Let us first count solutions to $q = p = 0$ and $q = 0, p = 1$. By Lemma 3.2.8 $W_1 \cap W_2 = \emptyset$. Thus all of the 2^{k-1} vectors β that satisfy $r\beta^t = 0$, we also have $M\beta^t + r \neq 0$. Thus for each such β, there are 2^{k-1} values for α that yield $p(\alpha, \beta) = 0$, and the same number give $p(\alpha, \beta) = 1$. Therefore if w is a commutator word

$$P(G, w = 1) = P(G, w = a^2) = \tfrac{1}{4}.$$

Let us now count solutions to $q = 1, p = 0$ and $q = p = 1$. There are 2^{k-1} vectors β such that $r\beta^t = 1$. This includes 2^{k-r} vectors that also satisfy $(M\beta + r) = 0$. For these vectors, $p(\alpha, \beta) = 0$ for any value of α. Therefore there are 2^{2k-r} solutions to $p = 0, q = 1$. For the remaining $2^{k-1} - 2^{k-\sigma}$ vectors β such that $r\beta^t = 1$ but $M\beta^t + r \neq 0$, 2^{k-1} values of α yield $p = 0, q = 1$ and the rest yield $p = q = 1$. Therefore, if w is a non-commutator word, σ is odd and

$$P(G, w = g) = \begin{cases} \frac{1}{4} + (\frac{1}{2})^{\sigma+1}, & g = b^2, \\ \frac{1}{4} - (\frac{1}{2})^{\sigma+1}, & g = a^2b^2. \end{cases}$$

Thus amalgamating cases i, iia and iib we have

$$\begin{aligned} S(G, 1) &\subseteq \left\{\tfrac{1}{16}, \tfrac{1}{4}\right\} \cup \left\{\tfrac{1}{2} + (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\} \cup \left\{\tfrac{1}{4} + (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N}\right\}, \\ S(G, a^2) &\subseteq \left\{\tfrac{1}{16}, \tfrac{1}{4}\right\} \cup \left\{\tfrac{1}{2} - (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\} \cup \left\{\tfrac{1}{4} - (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N}\right\}, \\ S(G, b^2) &\subseteq \left\{0, \tfrac{1}{16}, \tfrac{1}{4}\right\} \cup \left\{\tfrac{1}{4} + (\tfrac{1}{2})^{2n} \mid n \in \mathbb{N}\right\}, \\ S(G, a^2b^2) &\subseteq \left\{\tfrac{1}{16}, \tfrac{1}{4}\right\} \cup \left\{\tfrac{1}{4} - (\tfrac{1}{2})^{2n} \mid n \in \mathbb{N}\right\}, \\ S(G, g) &\subseteq \left\{0, \tfrac{1}{16}\right\} \ \forall g \notin \left\langle a^2, b^2 \right\rangle. \end{aligned}$$

To see the reverse inequalities consider the following.

- $w(x) = 1, w(x) = x$.
- Let w_i be the word constructed from i concatenations of commutators. Since

$$P(G, [x_1, x_2] = g) = \begin{cases} \frac{5}{8}, & g = 1, \\ \frac{3}{8}, & g = a^2, \end{cases}$$

it follows from (3.38) and (3.39) in §3.5 that

$$P(G, w_i = g) = \begin{cases} \frac{1}{2} + (\frac{1}{2})^{2i+1}, & g = 1, \\ \frac{1}{2} - (\frac{1}{2})^{2i+1}, & g = a^2. \end{cases}$$

- Now let $w_i(x_1, \ldots, x_i) := x_1^2 \ldots x_i^2$ for $i \geq 1$.

  Claim 3.6.2. *If i is odd*

$$P(G, w_i = g) = \begin{cases} \frac{1}{4}, & g = 1, a^2, \\ \frac{1}{4} + (\frac{1}{2})^{i+1}, & g = b^2, \\ \frac{1}{4} - (\frac{1}{2})^{i+1}, & g = a^2b^2, \\ 0, & g \notin \langle a^2, b^2 \rangle. \end{cases}$$

  If i is even then

$$P(G, w_i = g) = \begin{cases} \frac{1}{4} + (\frac{1}{2})^{i+1}, & g = 1, \\ \frac{1}{4} - (\frac{1}{2})^{i+1}, & g = a^2, \\ \frac{1}{4}, & g = b^2, a^2b^2, \\ 0, & g \notin \langle a^2, b^2 \rangle. \end{cases}$$

Proof of Claim 3.6.2. By considering the orders of the group elements, or by consulting GAP we find that

$$P(G, w_1 = g) = \begin{cases} \frac{1}{4}, & g = 1, a^2 \\ \frac{1}{2} & g = b^2, \\ 0, & \text{otherwise.} \end{cases} \qquad P(G, w_2 = g) = \begin{cases} \frac{3}{8}, & g = 1, \\ \frac{1}{8} & g = a^2, \\ \frac{1}{4} & g = b^2, a^2b^2, \\ 0, & g \notin \langle a^2, b^2 \rangle. \end{cases}$$

The claim therefore holds for $i \in \{1, 2\}$. We now prove the claim holds for even i, by induction. Suppose the result holds for some $i \geq 2$, where i is even. Then

$$\begin{aligned} P(G, w_{i+2} = 1) &= \sum_g P(G, w_i = g) P(G, w_2 = g^{-1}), \\ &= \left(\tfrac{1}{4} + \tfrac{1}{2^{i+1}}\right) \tfrac{3}{8} + \left(\tfrac{1}{4} - \tfrac{1}{2^{i+1}}\right) \tfrac{1}{8} + \tfrac{1}{4^2} + \tfrac{1}{4^2}, \\ &= \tfrac{1}{4} + \left(\tfrac{1}{2}\right)^{(i+2)+1}. \end{aligned}$$

$$\begin{aligned}P(G, w_{i+2} = 1) &= \sum_g P(G, w_i = g)P(G, w_2 = g^{-1}a^2),\\ &= \left(\tfrac{1}{4} + \tfrac{1}{2^{i+1}}\right)\tfrac{1}{8} + \left(\tfrac{1}{4} - \tfrac{1}{2^{i+1}}\right)\tfrac{3}{8} + \tfrac{1}{4^2} + \tfrac{1}{4^2},\\ &= \tfrac{1}{4} - \left(\tfrac{1}{2}\right)^{(i+2)+1}.\end{aligned}$$

$$\begin{aligned}P(G, w_{i+2} = b^2) &= \sum_g P(G, w_i = g)P(G, w_2 = g^{-1}b^2),\\ &= \left(\tfrac{1}{4} + \tfrac{1}{2^{i+1}}\right)\tfrac{1}{4} + \left(\tfrac{1}{4} - \tfrac{1}{2^{i+1}}\right)\tfrac{1}{4} + \tfrac{1}{4}\tfrac{3}{8} + \tfrac{1}{4}\tfrac{1}{8},\\ &= \tfrac{1}{4}.\end{aligned}$$

$$\begin{aligned}P(G, w_{i+2} = a^2b^2) &= \sum_g P(G, w_i = g)P(G, w_2 = g^{-1}a^2b^2),\\ &= \left(\tfrac{1}{4} + \tfrac{1}{2^{i+1}}\right)\tfrac{1}{4} + \left(\tfrac{1}{4} - \tfrac{1}{2^{i+1}}\right)\tfrac{1}{4} + \tfrac{1}{4}\tfrac{1}{8} + \tfrac{1}{4}\tfrac{3}{8},\\ &= \tfrac{1}{4}.\end{aligned}$$

By induction the result holds for even i. Let $i > 1$ be odd. Write $i = 2j + 1$. We shall use

$$P(G, w_i = h) = \sum_g P(G, w_{2j} = g)P(G, w_1 = g^{-1}h).$$

Then

$$\begin{aligned}P(G, w_i = 1) = &\quad \left(\tfrac{1}{4} + \tfrac{1}{2^{2j+1}}\right)\tfrac{1}{4} + \left(\tfrac{1}{4} - \tfrac{1}{2^{2j+1}}\right)\tfrac{1}{4} + \tfrac{1}{2}\tfrac{1}{4} + 0 &&= \tfrac{1}{4},\\ P(G, w_i = a^2) = &\quad \left(\tfrac{1}{4} + \tfrac{1}{2^{2j+1}}\right)\tfrac{1}{4} + \left(\tfrac{1}{4} - \tfrac{1}{2^{2j+1}}\right)\tfrac{1}{4} + 0 + \tfrac{1}{4}\tfrac{1}{2} &&= \tfrac{1}{4},\\ P(G, w_i = b^2) = &\quad \left(\tfrac{1}{4} + \tfrac{1}{2^{2j+1}}\right)\tfrac{1}{2} + 0 + \tfrac{1}{4}\tfrac{1}{4} + \tfrac{1}{4}\tfrac{1}{4} &&= \tfrac{1}{4} + \left(\tfrac{1}{2}\right)^{i+1},\\ P(G, w_i = a^2b^2) = &\quad 0 + \left(\tfrac{1}{4} - \tfrac{1}{2^{2j+1}}\right)\tfrac{1}{2} + \tfrac{1}{4}\tfrac{1}{4} + \tfrac{1}{4}\tfrac{1}{4} &&= \tfrac{1}{4} - \left(\tfrac{1}{2}\right)^{i+1}.\end{aligned}$$

Thus the claim holds. □

The above show the reverse inclusions and so the proof of the theorem is complete. □

Remark We have seen that there are infinitely many words that satisfy

$$P(G, w = 1) < P(G, w = b^2),$$

namely $w_i(x_1, \ldots, x_i) := x_1^2 \ldots x_i^2$ where i is odd. This property also holds for Q_8, but not for D_8 where we have that

$$P(D_8, w = 1) \geq P(D_8, w = g)$$

for all words w and all g in D_8.

3.7 [16,6]

Let M_{16} be the group given by the presentation

$$M_{16} = \langle a, b, \mid a^8 = b^2 = 1, bab = a^5 \rangle .$$

Then M_{16} has exponent 8 and has two proper verbal subgroups. These are the centre,

$$Z(M_{16}) = \langle g^2 \mid g \in M_{16} \rangle = \langle a^2 \rangle \cong \mathbb{Z}_4,$$

and the derived subgroup

$$M_{16}' = \langle a^4 \rangle \cong \mathbb{Z}_2.$$

Theorem 3.7.1.

$$S(M_{16}) = \left\{ \tfrac{1}{16}, \tfrac{1}{4}, \tfrac{1}{2} \right\} \cup \left\{ \tfrac{1}{2} \pm (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N} \right\}.$$

In particular

$$\begin{aligned} S(M_{16}, 1) &= \left\{ \tfrac{1}{16}, \tfrac{1}{4} \right\} \cup \left\{ \tfrac{1}{2} + (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N} \right\}, \\ S(M_{16}, a^4) &= \left\{ \tfrac{1}{16}, \tfrac{1}{4} \right\} \cup \left\{ \tfrac{1}{2} - (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N} \right\}, \\ S(M_{16}, g) &= \left\{ 0, \tfrac{1}{16}, \tfrac{1}{4} \right\}, \ \forall g \in \left\{ a^2, a^6 \right\}, \\ S(M_{16}, g) &= \left\{ 0, \tfrac{1}{16} \right\} \ \textit{otherwise.} \end{aligned}$$

Note that M_{16} has a verbal subgroup of order 4, but $\frac{1}{4}$ is not an accumulation point of $S(M_{16})$.

Proof. First we perform some routine calculations.

- $[a^{\alpha_i} b^{\beta_i}, a^{\alpha_j} b^{\beta_j}] = [a, b]^{\alpha_j \beta_i + \alpha_i \beta_j} = a^{4(\alpha_j \beta_i + \alpha_i \beta_j)}$
- $(a^{\alpha_i} b^{\beta_i})^{2r_{ii}} = (a^{2\alpha_i} b^{\beta_i} [b, a]^{\beta_i \alpha_i})^{r_{ii}} = (a^{2\alpha_i + 4\alpha_i \beta_i})^{r_{ii}} = a^{2r_{ii}\alpha_i(1+\beta_i)}$.

As per usual, write

$$w(x_1, \ldots, x_k) = \prod_i x_i^{r_i} \prod_{i<j} [x_i, x_j]^{r_{ij}}$$

with $0 \leq r_i \leq 7$, $0 \leq r_{ij} \leq 1$.

Case i. r_i odd for some i.

By Lemma 3.2.3, $(M_{16})_w^+ = M_{16}$ and $P(M_{16}, w = g) = \frac{1}{16}$ for all g in M_{16}.

Case ii. $r_i \in \{0, 2, 4, 6\}$ for all i.

Write $\alpha = (\alpha_i, \ldots, \alpha_k)$, $\beta = (\beta_1, \ldots, \beta_k)$ then

$$w(a^{\alpha_i} b^{\beta_1}, \ldots, a^{\alpha_k} b^{\beta_k}) = a^{2\left(\sum_i r_{ii}\alpha_i(1+\beta_i) + \sum_{i<j} 2(\alpha_j\beta_i + \alpha_i\beta_j)\right)}.$$

Thus we consider $p : \mathbb{F}_4^k \times \mathbb{F}_2^k \longrightarrow \mathbb{F}_4$ given by

$$p(\alpha, \beta) = r\alpha^t + 2\beta M\alpha^t = (r + 2\beta M)\beta^t.$$

Let $v \in \mathbb{F}_4^k$. Consider the map $f : \mathbb{F}_4^k \longrightarrow \mathbb{F}_4$, $f(x) := vx^t$. Since f is a homomorphism of groups, it has equally sized fibres. We therefore have the following cases.

- If $v = 0$, then $\mathrm{Im}f = \{0\}$. Then $P(\mathbb{F}_4^k, f = 0) = 1$.
- If $v_i \in \{1, 3\}$ for some i then $\mathrm{Im}f = \mathbb{F}_4$. Then $P(\mathbb{F}_4^k, f = 0) = \frac{1}{2}$.
- If $v_i \in \{0, 2\}$ for all i (but not all 0) then $\mathrm{Im}f = \{0, 2\}$ and $P(\mathbb{F}_4^k, f = 0) = \frac{1}{4}$.

Thus we need to calculate how many times $v := \phi(\beta) := (r + 2\beta M)$ falls into each of the above categories.

- If r has a component equal to 1 or 3, then ϕ evaluated at any β has a component equal to 1 or 3. Then $\mathrm{Im}\phi = \mathbb{F}_4$ and all outcomes have probability $\frac{1}{4}$.
- Suppose r_{ii} is even for all i, i.e. $r_i \in \{0, 4\}$. Then let $\widetilde{r_{ii}} = \frac{r_{ii}}{2} = \frac{r_i}{4} \in \{0, 1\}$. Then p becomes
$$p(\alpha, \beta) = \sum_i 2\widetilde{r_{ii}}\alpha_i + \sum_{i<j} 2r_{ij}(\alpha_j\beta_i + \alpha_i\beta_j).$$
Since p is always divisible by 2, define $p' : \mathbb{F}_2^k \times \mathbb{F}_2^k \longrightarrow \mathbb{F}_2$ by
$$p'(\alpha, \beta) = \sum_i \widetilde{r_{ii}}\alpha_i + \sum_{i<j} r_{ij}(\alpha_j\beta_i + \alpha_i\beta_j). \tag{3.44}$$

Let $r := (\widetilde{r_{11}}, \ldots, \widetilde{r_{kk}})$ and $M = (m_{ij})$ where

$$m_{ij} = \begin{cases} r_{ij} & i < j, \\ r_{ji} & i > j. \\ r_{ij} = 0 & i = j. \end{cases}$$

Then we have

$$p'(\alpha, \beta) = (r + \beta M)\alpha^t. \tag{3.45}$$

Note that by Lemma 3.2.5, M has even rank.

Case iia. r is in the rows span of M.

There are $2^{k-\rho}$ vectors β such that $(\beta M + r) = 0$. For any of these $p(\alpha, \beta) = 0$ for any α. If β_0 is one of the $2^k - 2^{k-\rho}$ vectors such that $(\beta_0 M + r) \neq 0$, then by Lemma 3.2.3 there are 2^{k-1} vectors α such that $p(\beta_0, \alpha) = 0$ and the same for $p(\beta_0, \alpha) = 1$. Therefore

$$P(M_{16}, w = 1) = \tfrac{1}{2^{2k}}(2^{2k-\rho} + (2^k - 2^{k-\rho})2^{k-1}) = \tfrac{1}{2} + \left(\tfrac{1}{2}\right)^{\rho+1}, \tag{3.46}$$

and

$$P(M_{16}, w = a^4) = 1 - P(M_{16}, w = 1) = \tfrac{1}{2} - \left(\tfrac{1}{2}\right)^{\rho+1}, \tag{3.47}$$

and ρ is necessarily even.

Case iib. r is not in the row span of M.

Then $(\beta M + r) \neq 0$ for any vector β. Then by Lemma 3.2.3 for each β_0 we have equal numbers of solutions to $(\beta_0 M + r)\alpha = 0$ and $(\beta_0 M + r)\alpha = 1$. Thus

$$P(M_{16}, w = 1) = P(M_{16}, w = a^4) = \tfrac{1}{2}.$$

Combining all the various cases, we see that

$$\begin{aligned}
S(M_{16}, 1) &\subseteq \left\{\tfrac{1}{16}, \tfrac{1}{4}\right\} \cup \left\{\tfrac{1}{2} + (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N}\right\}, \\
S(M_{16}, a^4) &\subseteq \left\{\tfrac{1}{16}, \tfrac{1}{4}\right\} \cup \left\{\tfrac{1}{2} - (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\}, \\
S(M_{16}, g) &\subseteq \left\{0, \tfrac{1}{16}, \tfrac{1}{4}\right\}, \ \forall g \in \left\{a^2, a^6\right\}, \\
S(M_{16}, g) &\subseteq \left\{0, \tfrac{1}{16}\right\}.
\end{aligned}$$

To see the reverse inclusions, consider

- $w(x) = 1, w(x) = x, w(x) = x^2$.
- Let $w_i(x_1, \ldots, x_{2i}) := [x_1, x_2] \ldots [x_{2i-1}, x_{2i}]$. Then

$$P(M_{16}, w_1 = 1) = \tfrac{5}{8}, \qquad P(M_{16}, w_1 = a^4) = \tfrac{3}{8}.$$

  Thus by (3.38) and (3.39) in §3.5 we have that

$$P(M_{16}, w_i = 1) = \tfrac{1}{2} + \left(\tfrac{1}{2}\right)^{2i+1}, \qquad P(M_{16}, w_i = a^4) = \tfrac{1}{2} - \left(\tfrac{1}{2}\right)^{2i+1}.$$

Thus we have the reverse inclusions and the theorem is proved. □

Remark $\frac{1}{2}$ is the only accumulation point associated with this group. Yet M_{16} has verbal subgroups of order 2 and order 4. Every other group we have looked at so far has been such that the reciprocal of the order of every proper subgroup has been an accumulation point.

3.8 [16,11]

Proposition 3.8.1. *Let $G := D_8 \times C_2$. Then*

$$S(G) = \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} \pm (\tfrac{1}{2})^n \mid n \in \mathbb{N}\}.$$

In particular, if z is the non-trivial element of the centre of the subgroup D_8,

$$\begin{aligned} S(G,1) &= \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} + (\tfrac{1}{2})^n \mid n \in \mathbb{N}\}, \\ S(G,z) &= \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} - (\tfrac{1}{2})^n \mid n \in \mathbb{N}\}, \\ S(G,g) &= \{0, \tfrac{1}{16}\}, \qquad \forall g \in \langle z \rangle \,. \end{aligned}$$

Proof. Any word over G may be written in standard form as

$$w(x_1, \ldots, x_k) = \prod_{i=1}^{k} x_i^{r_i} K(x_1, \ldots, x_k),$$

for some integers $r_1, \ldots, r_k$, where $K \in F_k'$ is a commutator word on $x_1, \ldots, x_k$. As in the proof of Theorem 3.3.1, we consider two cases.

Suppose r_i is odd for some i. By Lemma 3.2.1, since r_i is coprime to $|G|$, we have that for all g in G,

$$P(G, w = g) = \tfrac{1}{16}.$$

Now suppose that r_i is even for every i. By (2.13), we have that if $g \in D_8$ and $h \in C_2$,

$$P(G, w = gh) = P(D_8, w = g)P(C_2, w = h).$$

For each of these words, $P(C_2, w = 1) = 1$. Thus in this case, for $g \in D_8$,

$$P(G, w = g) = P(D_8, w = g).$$

By the proof of Theorem 3.3.1, if z is the non-trivial element of the centre,

$$\begin{aligned} \{P(D_8, w = 1) \mid w \text{ is such that } r_i \text{ is even for all } i\} &= \{\tfrac{1}{2} + (\tfrac{1}{2})^n \mid n \in \mathbb{N}\}, \\ \{P(D_8, w = z) \mid w \text{ is such that } r_i \text{ is even for all } i\} &= \{\tfrac{1}{2} - (\tfrac{1}{2})^n \mid n \in \mathbb{N}\}, \\ \{P(D_8, w = g) \mid w \text{ is such that } r_i \text{ is even for all } i\} &= \{0\}, \qquad \forall g \notin \langle z \rangle \,. \end{aligned}$$

Thus the proposition holds. □

3.9 [16,12]

Proposition 3.9.1. *Let $G := Q_8 \times C_2$. Then*

$$S(G) = \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} \pm (\tfrac{1}{2})^n \mid n \in \mathbb{N}\}.$$

In particular, if -1 is the non-trivial element of the centre of the subgroup Q_8, then

$$\begin{aligned} S(G,1) &= \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} + (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{2} - (\tfrac{1}{2})^{2n} \mid n \in \mathbb{N}\}, \\ S(G,-1) &= \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} - (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{2} + (\tfrac{1}{2})^{2n} \mid n \in \mathbb{N}\}, \\ S(G,g) &= \{0, \tfrac{1}{16}\}, \qquad \forall g \in \langle -1 \rangle . \end{aligned}$$

Proof. This follows in exactly the same way as that for $D_8 \times C_2$ (Proposition 3.8.1), using the proof of Theorem 3.4.1 in place of Theorem 3.3.1. □

3.10 [16,13]

Let G be the group with presentation

$$G = \left\langle a, b, c \mid a^4 = b^2 = 1, a^2 = c^2, bab = a^{-1}, ab = ba, bc = cb \right\rangle .$$

G is a non-abelian extension of $C_4 \times C_2$ by C_2. It is the central product of D_8 (or Q_8) over a common cyclic central subgroup of order 2. $\langle a, b \rangle \cong D_8$ and $\langle y \rangle \cong \mathbb{Z}_4$.

The derived subgroup of G is $\left\langle c^2 \right\rangle = \left\langle a^2 \right\rangle \cong C_2$, which is strictly contained in the centre, $Z(G) = \langle c \rangle \cong C_4$. G has one proper verbal subgroup, $\left\langle c^2 \right\rangle$. G has exponent 4.

Theorem 3.10.1.

$$S(G) = \left\{\tfrac{1}{16}\right\} \cup \left\{\tfrac{1}{2} \pm (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\}.$$

More specifically

$$\begin{aligned} S(G,1) &= \left\{\tfrac{1}{16}\right\} \cup \left\{\tfrac{1}{2} + (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\}, \\ S(G,c^2) &= \left\{\tfrac{1}{16}\right\} \cup \left\{\tfrac{1}{2} - (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\}, \\ S(G,g) &= \left\{0, \tfrac{1}{16}\right\}, \qquad \forall g \notin \left\langle c^2 \right\rangle . \end{aligned}$$

Proof. According to the presentation, we might write any element g_i of G as

$$g_i = a^{\alpha_i} b^{\beta_i} c^{\gamma_i},$$

for some $\alpha_i \in \{0,1,2,3\}$, $\beta_i, \gamma_i \in \{0,1\}$.

Then

$$\begin{aligned}
[g_i, g_j] &= [a^{\alpha_i} b^{\beta_i} c^{\gamma_i}, a^{\alpha_j} b^{\beta_j} c^{\gamma_j}], \\
&= [a^{\alpha_i} b^{\beta_i} c^{\gamma_i}, a^{\alpha_j} b^{\beta_j} c^{\gamma_j}], && \text{(since } c \in Z(G)\text{)}, \\
&= [a, b]^{\alpha_i \beta_j + \alpha_j \beta_i}, \\
&= c^{2(\alpha_i \beta_j + \alpha_j \beta_i)}, && \text{(since } [a, b] = a^2 = c^2\text{)},
\end{aligned}$$

and

$$\begin{aligned}
g_i^2 &= (a^{\alpha_i} b^{\beta_i} c^{\gamma_i})^2, \\
&= a^{\alpha_i} b^{\beta_i} a^{\alpha_i} b^{\beta_i} c^{2\gamma_i}, && \text{(since } c \in Z(G)\text{)}, \\
&= a^{2\alpha_i} b^{\beta_i} [b^{\beta_i}, a^{\alpha_i}] b^{\beta_i} c^{2\gamma_i}, \\
&= a^{2\alpha_i} b^{2\beta_i} c^{2\gamma_i} c^{2\alpha_i \beta_i}, && \text{(by the calculation above)}, \\
&= c^{2\alpha_i} c^{2\gamma_i} c^{2\alpha_i \beta_i}, && \text{(since } a^2 = c^2, b = 1\text{)}, \\
&= c^{2(\alpha_i + \gamma_i + \alpha_i \beta_i)}.
\end{aligned}$$

Let w be any word over G. We may write w in standard form as

$$w(x_1, \ldots, x_k) = x_1^{r_1} \ldots x_k^{r_k} \prod_{i<j} [x_i, x_j]^{r_{ij}},$$

where $r_i \in \{0, 1, 2, 3\}$ and $r_{ij} \in \{0, 1\}$.

Case i. Suppose r_i is odd for some i.

Then by Lemma 3.2.1, since r_i is coprime to $|G|$, we have $G_w^+ = G$ and for any $g \in G$,

$$P(G, w = g) = \tfrac{1}{16}.$$

Case ii. Suppose instead that $r_i \in \{0, 2\}$ for all i.

Define $r_{ii} := \frac{r_i}{2} \in \{0, 1\}$. Then if we write $g_i = a^{\alpha_i} b^{\beta_i} c^{\gamma_i}$ for all i, we have

$$\begin{aligned}
w(g_1, \ldots g_k) &= \prod_i c^{r_i(\alpha_i + \gamma_i + \alpha_i \beta_i)} \prod_{i<j} c^{2r_{ij}(\alpha_i \beta_j + \alpha_j \beta_i)}, \\
&= (c^2)^{\left(\sum_i r_{ii}(\alpha_i + \gamma_i + \alpha_i \beta_i) + \sum_{i<j}(\alpha_i \beta_j + \alpha_j \beta_i)\right)}.
\end{aligned}$$

Thus we are concerned with the parity of

$$p(\alpha, \beta, \gamma) = \sum_i r_{ii}(\alpha_i + \gamma_i + \alpha_i \beta_i) + \sum_{i<j}(\alpha_i \beta_j + \alpha_j \beta_i)$$

for $\alpha = (\alpha_1, \ldots, \alpha_k)$, $\beta = (\beta_1, \ldots, \beta_k)$ and $\gamma = (\gamma_i, \ldots, \gamma_k)$.

Since only the parity of p is important, we may consider p as a map from $\mathbb{F}_2^k \times \mathbb{F}_2^k \times \mathbb{F}_2^k$ to $\mathbb{F}_2^k$, so long as we later account for the fact that α_i may actually be any value in $\{0,1,2,3\}$, not just $\{0,1\}$.

As in the previous proofs, let $r := (r_{11}, \ldots, r_{kk})$ and $M := (m_{ij})$ where

$$m_{ij} = \begin{cases} r_{ij} & i \leq j, \\ r_{ji} & i > j. \end{cases}$$

Then we may write $p : \mathbb{F}_2^k \times \mathbb{F}_2^k \times \mathbb{F}_2^k \to \mathbb{F}_2^k$, as

$$p(\alpha, \beta, \gamma) = r\alpha^t + r\gamma^t + \alpha M \beta^t.$$

Case iia. Suppose w is a commutator word, i.e. $r = 0$.

Then

$$p(\alpha, \beta, \gamma) = p(\alpha, \beta) = \alpha M \beta^t.$$

- If $\beta_0 \in \mathrm{Ker}(M)$, then (α, β_0) is a solution for any $\alpha \in \mathbb{F}_2^k$. There are $2^{k-\sigma}2^k$ such pairs, where $\sigma = \mathrm{rank}(M)$.
- If $\beta_0 \notin \mathrm{Ker}(M)$, then by Lemma 3.2.3, $f(\alpha) := \alpha(M\beta_0) = 0$ has 2^{k-1} solutions. Since there are $2^k - 2^{k-\sigma}$ such β_0, we have $(2^k - 2^{k-\sigma})2^{k-1}$ pairs of this form.

In total, since γ may be arbitrarily chosen, there are $2^k(2^{2k-\sigma} + 2^{2k-1} - 2^{2k-\sigma-1})$ solutions to $p(\alpha, \beta, \gamma) = 0$.

Thus

$$\begin{aligned} P(G, w = 1) &= 2^k 2^k (2^{2k-\sigma} + 2^{2k-1} - 2^{2k-\sigma-1}) 2^{-3k} \\ &= \tfrac{1}{2} + \left(\tfrac{1}{2}\right)^{\sigma} - \left(\tfrac{1}{2}\right)^{\sigma+1} = \tfrac{1}{2} + \left(\tfrac{1}{2}\right)^{\sigma+1}. \end{aligned}$$

where $\sigma :=$rank(M). By Lemma 3.2.5 σ must be even.

Case iib. Suppose w is not a commutator word, i.e. $r \neq 0$.

Then

$$p = 0 \qquad \Longleftrightarrow \qquad r\gamma^t = r\alpha^t + \alpha M \beta^t.$$

Choose α_0, β_0 arbitrarily. Then, as explained in Lemma 3.2.3,

$$f(\gamma) := r\gamma^t = r\alpha_0^t + \alpha_0 M \beta_0^t$$

has 2^{k-1} solutions, out of the 2^k possible values for γ. Therefore,

$$P(G, w = 1) = P(G, w = c^2) = \tfrac{1}{2}.$$

Combining all three cases we have

$$\begin{aligned} S(G,1) &\subseteq \left\{\tfrac{1}{16}\right\} \cup \left\{\tfrac{1}{2} + (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\}, \\ S(G,c^2) &\subseteq \left\{\tfrac{1}{16}\right\} \cup \left\{\tfrac{1}{2} - (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\}, \\ S(G,g) &\subseteq \left\{0, \tfrac{1}{16}\right\}, \qquad \forall g \notin \left\langle c^2 \right\rangle. \end{aligned}$$

We conclude by demonstrating the reverse inclusions.

Consider the words w_i given by

$$\begin{aligned} w_1 &:= [x_1, x_2], \\ w_i &= w_{i-1} * [x_{2i-1}, x_{2i}], \quad \forall i > 1. \end{aligned}$$

Then since $P(G, [x_1, x_2] = 1) = \frac{5}{8}$, and $P(G, [x_{1,2}] = c^2) = \frac{3}{8}$, we see by Claim 3.5.4 in the proof of Theorem 3.8.1 that

$$P(G, w_i = g) = \begin{cases} \frac{1}{2} + (\frac{1}{2})^{2i+1}, & g = 1, \\ \frac{1}{2} - (\frac{1}{2})^{2i+1}, & g = c^2. \end{cases}$$

These words along with $w(x) = x$, $w(x) = x^2$ and $w(x) = 1$ show the reverse inclusion. □

Remarks

1. From the proof we see that if w is a non-commutator word, it has equally sized fibres. This does not hold for nilpotent groups in general, for example consider $w(x) = x^2$ in D_8.

2. Although the groups G and H given respectively by the GAP IDs $[16, 13]$ and $[16, 3]$ are both non-abelian extensions of $C_4 \times C_2$ by C_2, we have that $S(G) \neq S(H)$.

3. G satisfies the property that for any word w and any element $g \in G$,

$$P(G, w = 1) \geq P(G, w = g).$$

4. The accumulation points of $S(G)$ are exactly the reciprocals of the set of orders of proper verbal subgroups of G.

We have now considered all groups of order 16 and nilpotency class 2. We now consider some of those groups of order 32 with nilpotency class 2.

Since the method only various slightly with each calculation, and the results are always very similar, we shall just look at 3 examples of groups of order 32 (there are 26 in total). We shall perform the calculation for the groups with GAP ID $[32, 2]$, $[32, 4]$ (the first two in the catalogue) and $[32, 27]$, since this is an example of a group with non-cyclic centre.

3.11 [32,2]

Let G be the group given by the presentation

$$\left\langle a, b, c \mid a^4 = b^4 = c^2 = 1, [a, b] = c, [a, c] = [b, c] = 1 \right\rangle.$$

Then G has exponent 8. It has derived subgroup $\langle c \rangle \cong \mathbb{Z}_2$ and centre $Z(G) = \left\langle a^2, b^2, c \right\rangle = \left\langle g^2 \mid g \in G \right\rangle$. G has two proper verbal subgroups, G' and $Z(G)$, which have orders 2 and 8 respectively.

Theorem 3.11.1.

$$S(G) = \{0, \tfrac{1}{32}, \tfrac{1}{8}\} \cup \{\tfrac{1}{2} \pm \left(\tfrac{1}{2}\right)^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{8} \pm \left(\tfrac{1}{2}\right)^{2n+1} \mid n \in \mathbb{N}\}.$$

More specifically

$$\begin{aligned}
S(G, 1) &= \{\tfrac{1}{32}\} \cup \{\tfrac{1}{2} + \left(\tfrac{1}{2}\right)^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{8} + \left(\tfrac{1}{2}\right)^{2n+1} \mid n \in \mathbb{N}\}, \\
S(G, c) &= \{\tfrac{1}{32}\} \cup \{\tfrac{1}{2} - \left(\tfrac{1}{2}\right)^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{8} - \left(\tfrac{1}{2}\right)^{2n+1} \mid n \in \mathbb{N}\}, \\
S(G, a^2) = S(G, b^2) = S(G, a^2b^2c) &= \{0, \tfrac{1}{32}\} \cup \{\tfrac{1}{8} + \left(\tfrac{1}{2}\right)^{2n+1} \mid n \in \mathbb{N}\}, \\
S(G, a^2c) = S(G, b^2c) = S(G, a^2b^2) &= \{0, \tfrac{1}{32}\} \cup \{\tfrac{1}{8} - \left(\tfrac{1}{2}\right)^{2n+1} \mid n \in \mathbb{N}\}.
\end{aligned}$$

Proof. We may write any word over G in standard form as

$$w(x_1, \ldots, x_k) = \prod_{i=1}^{k} x_i^{r_i} \prod_{i<j} [x_i, x_j]^{r_{ij}},$$

where $r_i \in \{0, 1, 2, 3\}, r_{ij} \in \{0, 1\}$. We shall write

$$g_i = a^{\alpha_i} b^{\beta_i} c^{\gamma_i},$$

where $\alpha_i, \beta_i \in \{0,1,2,3\}, \gamma_i \in \{0,1\}$. Note that

$$\begin{aligned}[a^{\alpha_i}b^{\beta_i}c^{\gamma_i}, a^{\alpha_j}b^{\beta_j}c^{\gamma_j}] &= [a^{\alpha_i}b^{\beta_i}, a^{\alpha_j}b^{\beta_j}], \qquad (c \in Z(G)),\\ &= [a,b]^{\alpha_i\beta_j+\alpha_j\beta_i},\\ &= c^{\alpha_i\beta_j+\alpha_j\beta_i}.\end{aligned}$$

$$\begin{aligned}(a^{\alpha_i}b^{\beta_i}c^{\gamma_i})^2 &= (a^{\alpha_i}b^{\beta_i})^2, \qquad (c \in Z(G), c^2 = 1),\\ &= a^{2\alpha_i}b^{2\beta_i}c^{\alpha_i\beta_i}.\end{aligned}$$

Case i. Suppose r_i is odd for some i.

Then by Lemma 3.2.1, $G_w^+ = G$ and for all g in G,

$$P(G, w = g) = \tfrac{1}{32}.$$

Case ii. Suppose r_i even for all i.

Then if $r_{ii} := \frac{r_i}{2} \in \{0,1\}$,

$$\begin{aligned}w(\alpha,\beta,\gamma) &= \prod_{i=1}^{k} a^{2r_i\alpha_i}b^{2r_i\beta_i}c^{r_{ii}\alpha_i\beta_i}\prod_{i<j} r_{ij}(\alpha_i\beta_j+\alpha_j\beta_i),\\ &= a^{\sum_i r_i\alpha_i}b^{\sum r_i\beta_i}c^{\sum_i r_{ii}\alpha_i\beta_i}\sum_{i<j} r_{ij}(\alpha_i\beta_j+\alpha_j\beta_i).\end{aligned}$$

Let $r := (r_{11}, \ldots, r_{kk})$, and $M = (m_{ij})_{ij}$ where

$$m_{ij} := \begin{cases} r_{ij}, & i \le j,\\ r_{ji}, & i > j.\end{cases}$$

Then we have

$$w(\alpha,\beta,\gamma) = a^{2r\alpha^t}b^{2r\beta^t}c^{\alpha M\beta^t}.$$

Thus we consider $p_1, p_2 : \mathbb{F}_2^k \longrightarrow \mathbb{F}_2$, $p_3 : \mathbb{F}_2^k \times \mathbb{F}_2^k \longrightarrow \mathbb{F}_2$ given by

$$\begin{aligned}p_1(\alpha) &= r\alpha^t,\\ p_2(\beta) &= r\beta^t,\\ p_3(\alpha,\beta) &= \alpha M\beta^t.\end{aligned}$$

Case iia. Suppose $r = 0$, i.e. w is a commutator word.

Then we have $p_1 = p_2 \equiv 0$ and $p_3 = \alpha M\beta^t$. By Lemma 3.2.5, M has even rank.

- If $\alpha_0 \in \mathrm{Ker}(M)$ (there are $2^{k-\rho}$ of these) then $p(\alpha_0, \beta) = 0$ for any β. Thus there are $2^{2k-\rho}$ such solutions.

- If $\alpha_0 \notin \mathrm{Ker}(M)$ (there are $2^k - 2^{k-\rho}$ of these) then $\alpha M \neq 0$ and by Lemma 3.2.3 there are 2^{k-1} β such that $\alpha M \beta^t = 0$. Therefore there are $2^{k-1}(2^k - 2^{k-\rho})$ such solutions.

Therefore

$$\begin{array}{rcll} P(G, w = 1) = & \frac{1}{2^{2k}}(2^{2k-\rho} + 2^{2k-1} - 2^{2k-\rho-1}) & = \frac{1}{2} + \left(\frac{1}{2}\right)^{\rho+1}, \\ P(G, w = c) = & 1 - P(G, w = 1) & = \frac{1}{2} - \left(\frac{1}{2}\right)^{\rho+1}, \end{array}$$

and note that ρ is even by Lemma 3.2.5.

Case iib. Suppose $r \neq 0$, but r_i even for all i.

Note first that $\mathrm{rank}(M)$ is at least 1. We shall use the following results from §3.4.

(i) Since r is always in the row span of M (see Lemma 3.2.4) then

$$Mx^t = 0 \qquad \Longrightarrow \qquad rx^t = 0.$$

(ii) Define

$$\begin{aligned} W_1 &:= \{x \in \mathbb{F}_2^k \mid xM = r\}, \\ W_2 &:= \{x \in \mathbb{F}_2^k \mid rx^t = 0\}, \\ W_3 &:= \{x \in \mathbb{F}_2^k \mid rx^t = 1\} = \mathbb{F}_k^2 \backslash W_2. \end{aligned}$$

Then by Lemma 3.2.7 and Lemma 3.2.8,

- if $\mathrm{rank}(M)$ is even then $W_1 \subseteq W_2$,
- if $\mathrm{rank}(M)$ is odd then $W_1 \cap W_2 = \emptyset$, i.e. $W_1 \subseteq W_3$.

1. **Solving** $p_1 = p_2 = p_3 = 0$.

   - Suppose $\alpha M = 0$ (there are $2^{k-\rho}$ such α). Then by (i) $r\alpha^t = 0$. Thus $p_1 = p_3 = 0$. We may choose any β such that $r\beta^t = 0$ (there are 2^{k-1} such β). We therefore have $2^{2k-\rho}$ solutions of this kind.
   - Suppose $\alpha M = r$, (there are $2^{k-\rho}$ such α). By (ii), if $\mathrm{rank}(M)$ is odd, then $r\alpha^t = 1$ and there are no solutions of this sort. If $\mathrm{rank}(M)$ is even then

   $$\begin{aligned} p_1(\alpha) = r\alpha^t = 0, \qquad & \text{and} \\ p_3(\alpha, \beta) = \alpha M \beta^t = r\beta^t = p_2(\beta). & \end{aligned}$$

   There are 2^{k-1} solutions to $p_2 = 0$. Therefore we have $2^{k-\rho-1}$ solutions if $\mathrm{rank}(M)$ is even, and none if $\mathrm{rank}(M)$ is odd.

- Suppose α is such that $r\alpha^t = 0$ but $\alpha M \notin \{0, r\}$. If $\text{rank}(M)$ is even there are $2^{k-1} - 2^{k-\rho} - 2^{k-\rho}$ of these. If $\text{rank}(M)$ is odd there are $2^{k-1} - 2^{k-\rho}$ of these. We now need to solve

$$p_2(\beta) = r\beta^t = 0,$$
$$p_3(\beta) = (\alpha M)\beta^t = 0.$$

  There are 2^{k-2} such β. Therefore we have

$$2^{k-2}(2^{k-1} - 2^{k-\rho+1}) = 2^{2k-3} - 2^{2k-\rho-1}$$

  solutions of this type if $\text{rank}(M)$ is even, and

$$2^{k-2}(2^{k-1} - 2^{k-\rho}) = 2^{2k-3} - 2^{2k-\rho-2}$$

  solutions of this type if $\text{rank}(M)$ is odd.

Thus combining all three cases we see that if ρ is even then

$$P(G, w = 1) = \tfrac{1}{2^{2k}}(2^{2k-\rho-1} + 2^{2k-\rho-1}) + 2^{2k-3} - 2^{2k-\rho-1} = \tfrac{1}{8} + \left(\tfrac{1}{2}\right)^{\rho+1},$$

and if ρ is odd then

$$P(G, w = 1) = \tfrac{1}{2^{2k}}(2^{2k-\rho-1} + 0 + 2^{2k-3} - 2^{2k-\rho-2}) = \tfrac{1}{8} + \left(\tfrac{1}{2}\right)^{\rho+2}.$$

2. **Solving** $p_1 = p_2 = 0, p_3 = 1$.

   - Suppose $\alpha M = 0$. Then $p_3 = 0$ and there are no solutions.
   - Suppose $\alpha M = r$. Then $p_2 = r\beta^t = \alpha M \beta^t = p_3$ and there are no solutions.
   - Suppose $r\alpha^t = 0$ but $\alpha M \notin \{0, r\}$. As we saw in the last part of (1), if ρ is even, there are $2^{k-1} - 2^{k-\rho+1}$ of these. If $\text{rank}(M)$ is odd, there are $2^{k-1} - 2^{k-\rho}$ of these. We now need to solve

$$p_2(\beta) = r\beta^t = 0,$$
$$p_3(\beta) = (\alpha M)\beta^t = 1.$$

     Again, there are 2^{k-2} such β. Thus if ρ is even

$$P(G, w = c) = \tfrac{1}{2^{2k}}(2^{k-1} - 2^{k-\rho+1})2^{k-2} = \tfrac{1}{8} - \left(\tfrac{1}{2}\right)^{\rho+1},$$

and if ρ is odd

$$P(G, w = c) = \frac{1}{2^{2k}}(2^{k-1} - 2^{k-\rho})2^{k-2} = \frac{1}{8} - \left(\frac{1}{2}\right)^{\rho+2}.$$

3. **Solving** $p_1 = 0, p_2 = 1, p_3 = 0$.
   By the symmetry of M, this has the same number of solutions as $p_1 = 1, p_2 = p_3 = 0$.

   - If $\alpha M = 0$, (there are $2^{k-\rho}$ of these), then by (i) $r\alpha^t = 0$ and we have $p_1 = p_3 = 0$. There are 2^{k-1} choices for β such that $r\beta^t = 1$. Therefore there are $2^{2k-\rho-1}$ solutions of this sort.
   - If $\alpha M = r$ then $p_2 = p_3$ so there are no solutions.
   - If $\alpha r^t = 0$ but $\alpha M \notin \{0, r\}$ then just as in the last case, if ρ is even then there are $2^{2k-3} - 2^{2k-\rho-1}$ solutions, and if ρ is odd there are $2^{2k-3} - 2^{2k-\rho-2}$ solutions.

   Therefore if ρ is even then

   $$P(G, w = b^2) = P(G, w = a^2) = \frac{1}{2^{2k}}(2^{2k-\rho-1} + 2^{2k-3} - 2^{2k-\rho-1}) = \frac{1}{8}.$$

   If ρ is odd then

   $$P(G, w = b^2) = P(G, w = a^2) = \frac{1}{2^{2k}}(2^{2k-\rho-1} + 2^{2k-3} - 2^{2k-\rho-2}) = \frac{1}{8} + \left(\frac{1}{2}\right)^{\rho+2}.$$

4. **Solving** $p_1 = 1, p_2 = 0, p_3 = 1$.
   Or equivalently $p_1 = 0, p_2 = 1, p_3 = 1$.

   - If $\alpha M = 0$ there are no solutions.
   - If $\alpha M = r$ then $p_2 = p_3$ so there are no solutions.
   - Suppose $r\alpha^t = 1$ but $\alpha M \notin \{0, r\}$. If ρ is even then $W_1 \cap W_3 = \emptyset$, so there are 2^{k-1} such α. If ρ is odd, $W_1 \subseteq W_3$, so there are $2^{k-1} - 2^{k-\rho}$ such α. For each of these α_0 there are 2^{k-2} β such that $r\beta^t = 0$ and $(\alpha M)\beta^t = 1$.

   Therefore if ρ is even

   $$P(G, w = a^2c) = P(G, w = b^2c) = \frac{1}{2^{2k}}(2^{k-2}2^{k-1}) = \frac{1}{8}.$$

   If ρ is odd

   $$P(G, w = a^2c) = P(G, w = b^2c) = \frac{1}{2^{2k}}(2^{k-1} - 2^{k-\rho})2^{k-2} = \frac{1}{8} - \left(\frac{1}{2}\right)^{\rho+2}.$$

5. **Solving** $p_1 = p_2 = 1, p_3 = 0$.

- If $\alpha M = 0$ then there are no solutions.
- If $\alpha M = r$ then $p_2 = p_3$ and there are no solutions.
- Suppose $\alpha M \notin \{0, r\}$ and $r\alpha^t$. If $\text{rank}(M)$ is even then $W_1 \subseteq W_2$ so there are 2^{k-1} such α. There are then 2^{k-2} elements β such that $r\beta^t = 1$ and $(\alpha M)\beta^t = 0$. Thus there are 2^{2k-3} solutions. If $\text{rank}(M)$ is odd then $W_1 \subseteq W_3$, so there are $2^{k-1} - 2^{k-\rho}$ such α. Thus there are $2^{2k-2} - 2^{2k-\rho-2}$ solutions.

Thus if ρ is even then

$$P(G, w = a^2b^2) = \tfrac{1}{8},$$

and if ρ is odd then

$$P(G, w = a^2b^2) = \tfrac{1}{8} - \left(\tfrac{1}{2}\right)^{\rho+2}.$$

6. **Solving** $p_1 = p_2 = p_3 = 1$.

- If $\alpha M = 0$ there are no solutions.
- Suppose $\alpha M = r$. If $\text{rank}(M)$ is even, then $W_1 \subseteq W_2$ so that $r\alpha^t = 0$ and there are no solutions. If $\text{rank}(M)$ is odd, then $W_1 \subseteq W_2$, so all $2^{k-\rho}$ such α also satisfy $r\alpha^t = 1$. We then need only solve $2\beta^t = 1$, which has 2^{k-1} solutions.
- Suppose that $\alpha M \notin \{0, r\}$ but $r\alpha^t = 1$. If ρ is even, there are 2^{k-1} such α (since $W_1 \subseteq W_2$) and if ρ is odd there are $2^{k-1} - 2^{k-\rho}$ such α (since $W_1 \cap W_2 = \emptyset$). As in the last case, there are then 2^{k-2} appropriate β.

If ρ is even then

$$P(G, w = a^2b^2c) = \tfrac{1}{2^{2k}} 2^{k-2} 2^{k-1} = \tfrac{1}{8}.$$

If ρ is odd then

$$P(G, w = a^2b^2c) = \tfrac{1}{2^{2k}}(2^{2k-\rho-1} + 2^{2k-3} + 2^{2k-\rho-2}) = \tfrac{1}{8} + \left(\tfrac{1}{2}\right)^{\rho+2}.$$

Thus

$$\begin{aligned}
S(G, 1) &\subseteq \{\tfrac{1}{32}\} \cup \{\tfrac{1}{2} + \left(\tfrac{1}{2}\right)^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{8} + \left(\tfrac{1}{2}\right)^{2n+1} \mid n \in \mathbb{N}\}, \\
S(G, c) &\subseteq \{\tfrac{1}{32}\} \cup \{\tfrac{1}{2} - \left(\tfrac{1}{2}\right)^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{8} - \left(\tfrac{1}{2}\right)^{2n+1} \mid n \in \mathbb{N}\}, \\
S(G, a^2), S(G, b^2), S(G, a^2b^2c) &\subseteq \{0, \tfrac{1}{32}\} \cup \{\tfrac{1}{8} + \left(\tfrac{1}{2}\right)^{2n+1} \mid n \in \mathbb{N}\}, \\
S(G, a^2c), S(G, b^2c), S(G, a^2b^2) &\subseteq \{0, \tfrac{1}{32}\} \cup \{\tfrac{1}{8} - \left(\tfrac{1}{2}\right)^{2n+1} \mid n \in \mathbb{N}\}.
\end{aligned}$$

It remains to show the reverse inclusions.

- Let $w(x) = 1$. Then $P(G, w = 1) = 1$, and $P(G, w = g) = 0$ if $g \neq 1$.
- Let $w(x) = x$. Then $P(G, w = g) = \frac{1}{32}$.

- Let w_i be constructed from i concatenated commutators. Then

$$P(G, w_1 = 1) = \tfrac{5}{8} \qquad P(G, w_1 = c) = \tfrac{3}{8}.$$

  By (3.38) and (3.39) in §3.5 it follows that

$$P(G, w_i = 1) = \tfrac{1}{2} + \left(\tfrac{1}{2}\right)^{2i+1},$$
$$P(G, w_i = c) = \tfrac{1}{2} - \left(\tfrac{1}{2}\right)^{2i+1}.$$

- Let w_i be constructed from i concatenated commutators. Then the matrix M associated with i as defined above is the identity matrix I_i, which has rank i. We thus attain all the remaining values.

Thus the reverse inclusions hold, and the proof is complete. □

3.12 [32,4]

Let G be the group given by the presentation

$$G := \left\langle a, b \mid a^8 = b^4 = 1, bab^{-1} = a^5 \right\rangle.$$

Then G is the semi-direct product of $\mathbb{Z}_8$ and $\mathbb{Z}_4$ of M-type. G has exponent 8 and derived subgroup $\left\langle a^4 \right\rangle \cong \mathbb{Z}_2$. It has centre $\left\langle a^2, b^2 \right\rangle \cong \mathbb{Z}_4 \times \mathbb{Z}_2$ which is also the verbal subgroup $\left\langle g^2 \mid g \in G \right\rangle$. These are the only two proper verbal subgroups of G.

Theorem 3.12.1.

$$S(G) = \{\tfrac{1}{32}, \tfrac{1}{8}, \tfrac{1}{2}\} \cup \{\tfrac{1}{2} \pm \left(\tfrac{1}{2}\right)^{2n-1} \mid n \in \mathbb{N}\},$$
$$S(G, 1) = \{\tfrac{1}{32}, \tfrac{1}{8}, \tfrac{1}{2}\} \cup \{\tfrac{1}{2} + \left(\tfrac{1}{2}\right)^{2n-1} \mid n \in \mathbb{N}\},$$
$$S(G, a^4) = \{\tfrac{1}{32}, \tfrac{1}{8}, \tfrac{1}{2}\} \cup \{\tfrac{1}{2} - \left(\tfrac{1}{2}\right)^{2n-1} \mid n \in \mathbb{N}\},$$
$$S(G, g) = \{0, \tfrac{1}{32}, \tfrac{1}{8}\} \qquad g \in \left\langle a^2, b^2 \right\rangle \setminus \left\langle a^4 \right\rangle,$$
$$S(G, g) = \{0, \tfrac{1}{32}\} \qquad g \notin \left\langle a^2, b^2 \right\rangle.$$

Proof. Let w be a word over G. We write w in standard form as

$$w(x_1, \ldots, x_k) = \prod_i x_i^{r_i} \prod_{i<j} [x_i, x_j]^{r_{ij}},$$

where $0 \leq r_i \leq 7$, and $0 \leq r_{ij} \leq 1$. We shall write elements of G in the form $g_i = a_i^{\alpha_i} b_i^{\beta_i}$ where $0 \leq \alpha_i \leq 7$, $0 \leq \beta_i \leq 3$.

Case i. Suppose r_i is odd for some i.

Then by Lemma 3.2.1, $G_w^+ = G$ and for any g in G,

$$P(G, w = g) = \tfrac{1}{32}.$$

Case ii. Suppose $r_i \in \{0, 4\}$ for all i i.e. $w(G) = \langle a^4 \rangle$.

We shall use the following calculations.

$$[a^{\alpha_i} b^{\beta_i}, a^{\alpha_j} b^{\beta_j}]^{r_{ij}} = [a, b]^{r_{ij}(\alpha_i\beta_j - \alpha_j\beta_i)} = a^{4r_{ij}(\alpha_i\beta_j - \alpha_j\beta_i)}.$$

$$\begin{aligned}
(a^{\alpha_i} b^{\beta_i})^2 &= a^{2\alpha_i + 4\alpha_i\beta_i} b^{2\beta_i}, \\
(a^{\alpha_i} b^{\beta_i})^4 &= a^{4\alpha_i}, \\
(a^{\alpha_i} b^{\beta_i})^6 &= a^{6\alpha_i + 4\alpha_i\beta_i} b^{2\beta_i}, \\
(a^{\alpha_i} b^{\beta_i})^{2m} &= a^{2m\alpha_i + 4m\alpha_i\beta_i} b^{2m\beta_i}.
\end{aligned}$$

Thus if r_i is even for all i we may write

$$w(a^{\alpha_1} b^{\beta_1}, \ldots, a^{\alpha_k} b^{\beta_k}) = \prod_i (a^{\alpha_i} b^{\beta_i})^{r_i} \prod_{i<j} [a^{\alpha_i} b^{\beta_i}, a^{\alpha_j} b^{\beta_j}]^{r_{ij}}, \tag{3.48}$$

$$= a^{\sum_i r_i\alpha_i(1+2\beta_i) + \sum_{i<j} 4r_{ij}(\alpha_i\beta_j - \alpha_j\beta_i)} b^{\sum_i r_i\beta_i}. \tag{3.49}$$

Since in this case we have the stronger condition that $r_i \in \{0, 4\}$ for all i, we let $4\tilde{r_i} = r_i$ and have

$$\begin{aligned}
w(a^{\alpha_1} b^{\beta_1}, \ldots, a^{\alpha_k} b^{\beta_k}) &= a^{\sum_i 4\tilde{r_i}\alpha_i(1+2\beta_i) + \sum_{i<j} 4r_{ij}(\alpha_i\beta_j - \alpha_j\beta_i)} b^{\sum_i 4\tilde{r_i}\beta_i}, \\
&= a^{4\left(\sum_i \tilde{r_i}\alpha_i + \sum_{i<j} r_{ij}(\alpha_i\beta_j - \alpha_j\beta_i)\right)},
\end{aligned}$$

since a has order 8 and b has order 4.

Thus we consider $p : \mathbb{F}_2^k \times \mathbb{F}_2^k \longrightarrow \mathbb{F}_k^2$ given by

$$p(\alpha, \beta) = (r + \beta M)\alpha^t,$$

where $r := (\tilde{r_i}, \ldots, \tilde{r_k})$ and $M := (m_{ij})$ is given by

$$m_{ij} = \begin{cases} r_{ij}, & i \leq j, \\ 0, & i = j, \\ -r_{ji}, & i > j. \end{cases}$$

By Lemma 3.2.5, M has even rank. p is the same map as (3.45) used to calculate the probabilities associated with M_{16} in §3.7. From (3.46) and (3.47) in that proof we see

that if r is in the row span of M, then

$$P(G, w = 1) = \tfrac{1}{2} + \left(\tfrac{1}{2}\right)^{\rho+1},$$
$$P(G, w = a^4) = \tfrac{1}{2} - \left(\tfrac{1}{2}\right)^{\rho+1},$$

where ρ is the rank of M. We also see that if r is not in the row span of M then

$$P(G, w = 1) = P(G, w = a^4) = \tfrac{1}{2}.$$

Case iii. Suppose that r_i is even for i, but that $r_i = 2$ for some i, i.e. $w(G) = \left\langle a^2, b^2 \right\rangle$. Then letting $2\tilde{r_i} = r_i$ in (3.49) we have

$$w(a^{\alpha_1} b^{\beta_1}, \ldots, a^{\alpha_k} b^{\beta_k}) = a^{\sum_i 2\tilde{r_i}\alpha_i(1+2\beta_i) + \sum_{i<j} 4r_{ij}(\alpha_i\beta_j - \alpha_j\beta_i)} b^{\sum_i 2\tilde{r_i}\beta_i}.$$

Thus we consider $p_1 : \mathbb{F}_4^k \times \mathbb{F}_4^k \longrightarrow \mathbb{F}_4^k$ and $p_2 : \mathbb{F}_4^k \longrightarrow \mathbb{F}_2^k$ given by

$$p_1(\alpha, \beta) = (r + 2\beta M)\alpha^t, \tag{3.50}$$
$$p_2(\beta) = r\beta^t, \tag{3.51}$$

where $r = (\tilde{r_1}, \ldots, \tilde{r_k})$ and $M = (m_{ij})_{ij}$ is given by

$$m_{ij} = \begin{cases} r_{ij}, & i \leq j, \\ \tilde{r_i}, & i = j, \\ -r_{ji}, & i > j. \end{cases}$$

Consider first $p_2(\beta) = r\beta^t$. This has equally sized fibres. For each β_0, since $(r + 2\beta_0 M)$ must contain an odd component (r_i contains an odd component, $2\beta M$ does not) it follows that $\alpha \mapsto (r + 2\beta_0 M)\alpha^t$ has image $\mathbb{F}_4$ and equally sized fibres. Thus for all $g \in G_w^+$,

$$P(G, w = g) = \tfrac{1}{8}.$$

Combining all of the above cases we see that

$$S(G, 1) \subseteq \{\tfrac{1}{32}, \tfrac{1}{8}, \tfrac{1}{2}\} \cup \{\tfrac{1}{2} + \left(\tfrac{1}{2}\right)^{2n-1} \mid n \in \mathbb{N}\},$$
$$S(G, a^4) \subseteq \{\tfrac{1}{32}, \tfrac{1}{8}, \tfrac{1}{2}\} \cup \{\tfrac{1}{2} - \left(\tfrac{1}{2}\right)^{2n-1} \mid n \in \mathbb{N}\},$$
$$S(G, g) \subseteq \{0, \tfrac{1}{32}, \tfrac{1}{8}\} \qquad g \in \left\langle a^2, b^2 \right\rangle \setminus \left\langle a^4 \right\rangle,$$
$$S(G, g) \subseteq \{0, \tfrac{1}{32}\} \qquad g \notin \left\langle a^2, b^2 \right\rangle.$$

It remains to see the reverse inclusions. Consider

- $w(x) = 1, w(x) = x$.

- w_i given by concatenating i commutators.
- $w(x) = x^2$.

It is routine to show that these give all the values in the above, and so the theorem holds. □

Remark G contains a verbal subgroup of order 8, but $\frac{1}{8}$ is not an accumulation point of $S(G)$.

3.13 [32,27]

Let G be the group with presentation

$$\begin{aligned}\langle a,b,c,d,f \mid\ & a^2=b^2=c^2=d^2=f^2=1, ab=ba,\\ & bc=bc, dcd^{-1}=ac, fcf^{-1}=bc,\\ & da=ad, db=bd, fa=af, fb=bf\rangle\,.\end{aligned}$$

Then G may be thought of as the subgroup of the upper triangular unipotent matrix group $U(4,2)$ that consists of matrices whose entry in position $(1,2)$ is zero, i.e. G is the group of matrices of the form

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

where the starred entries lie in the field with two elements. G has exponent 4. The subgroup $\langle a,b\rangle$ is both the centre and derived subgroup, and is equal to the group $\left\langle g^2 \mid g \in G\right\rangle \cong V_4$. This is the only proper subgroup of G.

Theorem 3.13.1.

$$\begin{aligned} S(G) &= \{\tfrac{1}{32}\} \cup \{\tfrac{1}{4} + 3\left(\tfrac{1}{2}\right)^{n+1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{4} - \left(\tfrac{1}{2}\right)^{n+1} \mid n \in \mathbb{N}\},\\ S(G,1) &= \{\tfrac{1}{32}\} \cup \{\tfrac{1}{4} + 3\left(\tfrac{1}{2}\right)^{n+1} \mid n \in \mathbb{N}\},\\ S(G,a), S(G,b), S(G,ab) &= \{\tfrac{1}{32}\} \cup \{\tfrac{1}{4} - \left(\tfrac{1}{2}\right)^{n+1} \mid n \in \mathbb{N}\},\\ S(G,g) &= \{0, \tfrac{1}{32}\}, \qquad g \notin Z(G). \end{aligned}$$

Proof. Let w be a word over G. As usual write

$$w(x_1,\ldots,x_k) = \prod_{i=1}^{k} x_i^k \prod_{i<j} [x_i,x_j]^{r_{ij}},$$

with $r_i \in \{0,1,2,3\}$ and $r_{ij} \in \{0,1\}$.

Case i. Suppose r_i is odd for some i.

By Lemma 3.2.1, $G_w^+ = G$ and for any g in G,

$$P(G, w = g) = \tfrac{1}{32}.$$

Case ii. Suppose r_i is even for all i.

We may write any element g of G in the form

$$g = a^\alpha b^\beta c^\gamma d^\delta f^\epsilon$$

where $\alpha, \beta, \gamma, \delta, \epsilon \in \{0,1\}$. Note that

$$\begin{aligned}
[a^{\alpha_i} b^{\beta_i} c^{\gamma_i} d^{\delta_i} f^{\epsilon_i}, a^{\alpha_j} b^{\beta_j} c^{\gamma_j} d^{\delta_j} f^{\epsilon_j}] &= [c^{\gamma_i} d^{\delta_i} f^{\epsilon_i}, c^{\gamma_j} d^{\delta_j} f^{\epsilon_j}], \qquad (a, b \in Z(G)), \\
&= [c,d]^{\gamma_i\delta_j - \gamma_j\delta_i} [c,f]^{\gamma_i\epsilon_j - \gamma_j\epsilon_i} [d,f]^{\delta_i\epsilon_j - \epsilon_i\delta_j}, \\
&= a^{\gamma_i\delta_j - \gamma_j\delta_i} b^{\gamma_i\epsilon_j - \gamma_j\epsilon_i}
\end{aligned}$$

and

$$\begin{aligned}
(a^{\alpha_i} b^{\beta_i} c^{\gamma_i} d^{\delta_i} f^{\epsilon_i})^2, &= (c^{\gamma_i} d^{\delta_i} f^{\epsilon_i})^2 \qquad (a, b \in Z(G), Z \text{ has exponent } 2), \\
&= c^{2\gamma_i} (d^{\delta_i} e^{\epsilon_i})^2 [d^{\delta_i} e^{\epsilon_i}, c^{\gamma_i}], \\
&= a^{\gamma_i\delta_i} b^{\gamma_i\epsilon_i}.
\end{aligned}$$

So if r_i is even, then since a, b are central,

$$(a^{\alpha_i} b^{\beta_i} c^{\gamma_i} d^{\delta_i} f^{\epsilon_i})^{r_i} = a^{r_i\gamma_i\delta_i} b^{r_i\gamma_i\epsilon_i}.$$

Thus

$$\begin{aligned}
w(g_1, \ldots, g_k) &= \prod_i a^{r_i\gamma_i\delta_i} b^{r_i\gamma_i\epsilon_i} \prod_{i<j} a^{r_{ij}(\gamma_i\delta_j - \gamma_j\delta_i)} b^{r_{ij}(\gamma_i\epsilon_j - \gamma_j\epsilon_i)}, \\
&= a^{\sum_i r_i\gamma_i\delta_i + \sum_{i<j} r_{ij}(\gamma_i\delta_j - \gamma_j\delta_i)} b^{\sum_i r_i\gamma_i\epsilon_i + \sum_{i<j} r_{ij}(\gamma_i\epsilon_j - \gamma_j\epsilon_i)}.
\end{aligned}$$

We therefore define $p_1, p_2 : \mathbb{F}_2^k \times \mathbb{F}_2^k \longrightarrow \mathbb{F}_2$ by

$$p_1(\gamma, \delta) := \gamma M \delta, \tag{3.52}$$

$$p_2(\gamma, \epsilon) := \gamma M \epsilon, \tag{3.53}$$

where $M = (m_{ij})_{ij}$ is given by

$$m_{ij} = \begin{cases} r_{ij}, & i \leq j, \\ \tilde{r_i}, & i = j, \\ r_{ji}, & i > j. \end{cases}$$

1. **Solving** $p_1 = p_2 = 0$.

   - Suppose $\gamma M = 0$ (there are $2^{k-\rho}$ such γ, where ρ is the rank of M). Then any δ, ϵ yield a solution. Therefore there are $2^{3k-\rho}$ solutions of this form.
   - Suppose $\gamma M \neq 0$ (there are $2^k - 2^{k-\rho}$ such γ). Then by Lemma 3.2.3 there are 2^{k-1} δ such that $p_1 = 0$ and 2^{k-1} ϵ such that $p_2 = 0$. Thus there are $(2^k - 2^{k-\rho})2^{k-1}2^{k-1}$ solutions of this form.

   Thus in total we have $2^{3k-2} + 3.2^{3k-\rho}$ solutions. Thus

   $$P(G, w = 1) = \tfrac{1}{4} + 3\left(\tfrac{1}{2}\right)^{\rho+2}.$$

2. **Solving** $p_1 = 0, p_2 = 1$.
   This will have the same number of solutions as $p_1 = 0, p_2 = 0$.

   - Suppose $\gamma_0 M = 0$. Then there are no solutions containing γ_0.
   - Suppose $\gamma_0 M \neq 0$ there are $(2^{k-\rho})$ such γ_0. Then by Lemma 3.2.3 there are 2^{2k-2} pairs (δ, ϵ) that yield a solution along with γ_0. Hence there are $(2^{k-\rho})2^{2k-2}$ solutions.

   Thus

   $$P(G, w = a) = P(G, w = b) = \tfrac{1}{4} - \left(\tfrac{1}{2}\right)^{\rho+2}.$$

3. **Solving** $p_1 = p_2 = 1$.
   It follows from the last two cases that

   $$P(w = ab) = 1 - \tfrac{1}{4} + 3\left(\tfrac{1}{2}\right)^{\rho+2} - 2\left(\tfrac{1}{4} - \left(\tfrac{1}{2}\right)^{\rho+2}\right) = \tfrac{1}{4} - \left(\tfrac{1}{2}\right)^{\rho+2}.$$

Thus we see that

$$\begin{aligned} S(G, 1) &\subseteq \{\tfrac{1}{32}\} \cup \{\tfrac{1}{4} + 3\left(\tfrac{1}{2}\right)^{n+1} \mid n \in \mathbb{N}\}, \\ S(G, a), S(G, b), S(G, ab) &\subseteq \{\tfrac{1}{32}\} \cup \{\tfrac{1}{4} - \left(\tfrac{1}{2}\right)^{n+1} \mid n \in \mathbb{N}\}, \\ S(G, g) &\subseteq \{0, \tfrac{1}{32}\} \qquad g \notin Z(G). \end{aligned}$$

It remains to see the reverse inclusions. If w_i is the word constructed by concatenating i copies of x_2, where i is a non-negative integer, then the associated matrix M as defined

above has rank i. This along with the word $w(x) = x$ gives us the reverse inclusion. Thus the theorem holds. □

There are many more groups of order 32 and nilpotency class 2. We leave these calculations as an open problem. We turn our attention instead to a couple of groups of order 27, and again class 2.

3.14 [27,3]

Let G be the group given by the presentation

$$G = \langle a, b, c \mid a^3 = b^3 = c^3 = 1, [a, b] = c^{-1}, [a, c] = [b, c] = 1 \rangle .$$

This is the upper triangular unipotent matrix group $U(3, 3)$. It has exponent 3 (in fact it is the only non-abelian group of order 27 and exponent 3) and is equal to the Burnside group $B(2, 3)$, i.e. the quotient of the free group of rank 2 by the subgroup generated by all cubes in the group. The derived subgroup is equal to the centre of the group, which is the subgroup $\langle c \rangle$.

Theorem 3.14.1.

$$\begin{aligned}
S(G) &= \{\tfrac{1}{27}\} \cup \{\tfrac{1}{3} + 2\left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{3} + 2\left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\},\\
S(G, 1) &= \{\tfrac{1}{27}\} \cup \{\tfrac{1}{3} + 2\left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\},\\
S(G, c) &= S(G, c^2) = \{\tfrac{1}{27}\} \cup \{\tfrac{1}{3} - \left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\},\\
S(G, g) &= \{0, \tfrac{1}{27}\} \qquad g \notin \langle c \rangle .
\end{aligned}$$

Proof. Let w be a word over G. Write w in the form

$$w = \prod_{i=1}^{k} x_i^{r_i} \prod_{i<j} [x_i, x_j]^{r_{ij}},$$

where $r_i \in \{0, 1, 2\}$ and $r_{ij} \in \{0, 1, 2\}$ for all i, j.

Case i. Suppose $r_i \in \{1, 2\}$ for some i.

Then by Lemma 3.2.1 we have that $G_w^+ = G$ and for all $g \in G$,

$$P(G, w = g) = \tfrac{1}{27}.$$

Case ii. Suppose $r_i = 0$ for all i, i.e. w is a commutator word.

Write $g_i = a^{\alpha_i} b^{\beta_i} c^{\gamma_i}$. Then we have

$$[g_i, g_j] = [a^{\alpha_i} b^{\beta_i} c^{\gamma_i}, a^{\alpha_j} b^{\beta_j} c^{\gamma_j}] = c^{2\alpha_i\beta_j - \beta_i\alpha_j}.$$

Thus we may write

$$w(\alpha, \beta) = \prod_{i<j} c^{2\alpha_i\beta_j - \beta_i\alpha_j} = c^{\left(\sum_{i<j} 2\alpha_i\beta_j + \beta_i\alpha_j\right)}.$$

So we define $p : \mathbb{F}_3^k \times \mathbb{F}_3^k \longrightarrow \mathbb{F}_3^k$ by

$$p(\alpha, \beta) := \beta M \alpha^t,$$

where $M = (m_{ij})_{ij}$ is given by

$$m_{ij} = \begin{cases} r_{ij}, & i \leq j, \\ 0, & i = j, \\ -r_{ji}, & i > j. \end{cases}$$

Note, since M is skew symmetric, it has even rank.

- Suppose β_0 is in the row kernel of M (there are $3^{k-\rho}$ such vectors) then $p(\alpha, \beta_0) = 0$ for any α. Thus there are $3^{k-\rho}3^k$ such solutions to $p = 0$.
- Suppose β_0 is not in the row kernel of M (there are $3^k - 3^{k-\rho}$ such vectors) then since $x \mapsto (\beta_0 M)x^t$ is a homomorphism with image $\mathbb{F}_3$, it follows that it has equally sized fibres, each of size 3^{k-1}.

Thus

$$\begin{aligned} P(G, w = 1) &= \tfrac{1}{3^{2k}} \left(3^{k-\rho}3^k + (3^k - 3^{k-\rho})3^{k-1}\right) = \tfrac{1}{3} + 2\left(\tfrac{1}{3}\right)^{\rho+1}, \\ P(G, w = c) = P(G, w = c^2) &= \tfrac{1}{3^{2k}} \left((3^k - 3^{k-\rho})3^{k-1}\right) = \tfrac{1}{3} - \left(\tfrac{1}{3}\right)^{\rho+1}. \end{aligned}$$

Since ρ must be even, combining the two cases we have

$$\begin{aligned} S(G, 1) &\subseteq \{\tfrac{1}{27}\} \cup \{\tfrac{1}{3} + 2\left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\}, \\ S(G, c), S(G, c^2) &\subseteq \{\tfrac{1}{27}\} \cup \{\tfrac{1}{3} - \left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\}, \\ S(G, g) &\subseteq \{0, \tfrac{1}{27}\} \qquad g \notin \langle c \rangle. \end{aligned}$$

To see the reverse inclusion, consider the word $w(x) = x$, and the words w_i constructed from concatenating i commutators, for non-negative integers i. The associated matrices have rank $2i$ (since each row of the $2i$ by $2i$ matrix will have exactly one non-zero entry) and thus all of the above values may be attained. Thus the reverse inclusions hold,

and the theorem is proved. □

Remark Let $B(n,m)$ denote the free Burnside group of rank n and exponent m. Suppose $G = B(2,p)/\gamma_3$ where p is prime, and G has nilpotency class 2. Then the above generalises exactly to give

$$\begin{aligned} S(G,1) &= \{|G|^{-1}\} \cup \{\tfrac{1}{p} + (1-p)\left(\tfrac{1}{p}\right)^{2n-1} \mid n \in \mathbb{N}\}, \\ S(G,g) &= \{|G|^{-1}\} \cup \{\tfrac{1}{p} - \left(\tfrac{1}{p}\right)^{2n-1} \mid n \in \mathbb{N}\}, \qquad \text{for } g \in G', \\ S(G,g) &= \{0, |G|^{-1}\}, \qquad \text{for } g \notin \langle c \rangle . \end{aligned}$$

3.15 [27,4]

Let G be the group given by the presentation

$$G = \left\langle a, b, c \mid a^9 = b^3 = 1, [a,b] = a^3 \right\rangle .$$

Then G is a semidirect product of $\mathbb{Z}_9$ and $\mathbb{Z}_3$. G has exponent 9. The derived subgroup and centre coincide, and are equal to $\left\langle g^3 \mid g \in G \right\rangle = \left\langle a^3 \right\rangle$.

Theorem 3.15.1.

$$\begin{aligned} S(G,1) &= \{\tfrac{1}{27}, \tfrac{1}{3}\} \cup \{\tfrac{1}{3} + 2\left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\}, \\ S(G,a^3) = S(G,a^6) &= \{\tfrac{1}{27}, \tfrac{1}{3}\} \cup \{\tfrac{1}{3} - \left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\}, \\ S(G,g) &= \{0, \tfrac{1}{27}\}, \qquad g \notin \left\langle a^3 \right\rangle . \end{aligned}$$

Proof. Let w be a word over G. Write w in the form

$$w = \prod_{i=1}^{k} x_i^{r_i} \prod_{i<j} [x_i, x_j]^{r_{ij}},$$

where $0 \le r_i \le 8$ and $0 \le r_{ij} \le 2$ for all i, j.

Case i. Suppose r_i is not a multiple of 3 for some i.

Then by Lemma 3.2.1 we have that $G_w^+ = G$ and for all $g \in G$,

$$P(G, w = g) = \tfrac{1}{27}.$$

Case ii. Suppose $r_i = \{0, 3, 6\}$ for all i.

Write $g_i = a^{\alpha_i} b^{\beta_i}$. Then we have

$$[g_i, g_j] = [a^{\alpha_i} b^{\beta_i}, a^{\alpha_j} b^{\beta_j}] = [a,b]^{\alpha_i\beta_j - \beta_i\alpha_j} = a^{3(\alpha_i\beta_j - \beta_i\alpha_j)},$$

and

$$\begin{aligned} g_i^3 &= (a^{\alpha_i} b^{\beta_i})^3, \\ &= a^{\alpha_i} b^{\beta_i} a^{2\alpha_i} b^{2\beta_i} [b^{\beta_i}, a^{\alpha_i}], \\ &= a^{3\alpha_i} b^{3\beta_i} [b^{\beta_i}, a^{\alpha_i}][b^{\beta_i}, a^{2\alpha_i}], \\ &= a^{3\alpha_i} [b, a]^{3\alpha_i\beta_i} = a^{3\alpha_i}. \end{aligned}$$

Thus we may write

$$w(\alpha, \beta) = \prod_i a^{3\alpha_i} \prod_{i<j} a^{3(2\alpha_i\beta_j - \beta_i\alpha_j)} = a^{3\left(\sum_i \alpha_i + \sum_{i<j} \alpha_i\beta_j + \beta_i\alpha_j\right)}.$$

So we define $p : \mathbb{F}_3^k \times \mathbb{F}_3^k \longrightarrow \mathbb{F}_3^k$ by

$$p(\alpha, \beta) := (r + \beta M)\alpha^t,$$

where $r = (r_1, \ldots, r_k)$ and $M = (m_{ij})_{ij}$ is given by

$$m_{ij} := \begin{cases} r_{ij}, & i \leq j, \\ 0, & i = j, \\ -r_{ji}, & i > j. \end{cases}$$

Note, since M is skew symmetric, it has even rank.

Case iia. Suppose $-r$ is in the row span of M.

- Then there are $3^{k-\rho}$ vectors β_0 such that $r + \beta_0 M = 0$. Then $p(\alpha, \beta_0) = 0$ for any α_0 and so there are $3^{2k-\rho}$ solutions to $p(\alpha, \beta) = 0$ of this form.
- There are $3^k - 3^{k-\rho}$ vectors β_0 such that $r + \beta_0 M \neq 0$. Then there are 3^{k-1} vectors α_0 such that $p(\alpha_0, \beta_0) = 0$, and the same number such that $p(\alpha_0, \beta_0) = 1$ and similarly for $p(\alpha_0, \beta_0) = 2$.

Thus

$$\begin{aligned} P(G, w = 1) &= \tfrac{1}{3} + 2(\tfrac{1}{3})^{\rho+1}, \\ P(G, w = s^3) = P(G, w = s^6) &= \tfrac{1}{3} - (\tfrac{1}{3})^{\rho+1}. \end{aligned}$$

Case iib. Suppose $-r$ is not in the row span of M.

Then for any β_0, $r + \beta_0 M \neq 0$. Thus since $\alpha \longmapsto (r + \beta_0 M)\alpha_0^t$ has equally sized fibres, we see that

$$P(G, w = c^m) = \tfrac{1}{3},$$

for $m \in \{0, 1, 2\}$.

Combining the above three cases we find that

$$\begin{aligned} S(G,1) &\subseteq \{\tfrac{1}{27},\tfrac{1}{3}\} \cup \{\tfrac{1}{3} + 2\left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\}, \\ S(G,a^3), S(G,a^6) &\subseteq \{\tfrac{1}{27},\tfrac{1}{3}\} \cup \{\tfrac{1}{3} - \left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\}, \\ S(G,g) &\subseteq \{0,\tfrac{1}{27}\}, \qquad g \notin \langle a^3 \rangle . \end{aligned}$$

To see the reverse inclusions consider the words $w(x) = x$, the sequence of words constructed from concatenating commutators, and the sequence of words constructed from concatenating x^3. □

These are the only groups of order 27 and nilpotency class 2.

3.16 Summary of results and conjectures

For convenience, we shall list $S(G)$ for the above groups, and include some remarks on those that satisfy unusual properties.

1. D_8 has one proper verbal subgroup, the derived subgroup, of order 2.

$$S(D_8) = \left\{\tfrac{1}{8}\right\} \cup \left\{\tfrac{1}{2} \pm (\tfrac{1}{2})^n \mid n \in \mathbb{N}\right\}.$$

2. Q_8 has one proper verbal subgroup, the derived subgroup, of order 2.

$$S(Q_8) = \left\{\tfrac{1}{8}\right\} \cup \left\{\tfrac{1}{2} \pm (\tfrac{1}{2})^n \mid n \in \mathbb{N}\right\}.$$

   There are infinitely many words such that

$$P(Q_8, w = g) > P(Q_8, w = 1)$$

   for some group element g. D_8 and Q_8 have the same set of associated probabilities, though the sets $S(D_8, g)$ and $S(Q_8, g)$ do differ.

3. $G = [16, 3]$ has two proper subgroups, of orders 2 (the derived subgroup) and 4.

$$S(G) = \{\tfrac{1}{16},\tfrac{1}{4}\} \cup \{\tfrac{1}{4} \pm (\tfrac{1}{2})^{n+1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{2} \pm (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N}\}.$$

4. $G = [16, 4]$ has two proper subgroups, of orders 2 (the derived subgroup) and 4.

$$S(G) = S(G) = \left\{\tfrac{1}{16},\tfrac{1}{4}\right\} \cup \left\{\tfrac{1}{4} \pm (\tfrac{1}{2})^{2n+1} \mid n \in \mathbb{N}\right\} \cup \left\{\tfrac{1}{2} \pm (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\}.$$

   There are infinitely many words such that

$$P(G, w = g) > P(G, w = 1)$$

for some group element g.

5. M_{16} has two proper verbal subgroups, of orders 2 (the derived subgroup) and 4 (the centre).
$$S(M_{16}) = \left\{\tfrac{1}{16}, \tfrac{1}{4}, \tfrac{1}{2}\right\} \cup \left\{\tfrac{1}{2} \pm (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\}.$$
Although M_{16} has a verbal subgroup of order 4, $\frac{1}{4}$ is not an accumulation point of $S(M_{16})$.

6. $G = [16, 11]$ is the direct product of D_8 and C_2.
$$S(G) = \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} \pm (\tfrac{1}{2})^{n} \mid n \in \mathbb{N}\}.$$

7. $G = [16, 12]$ is the direct product of Q_8 and C_2.
$$S(G) = \{\tfrac{1}{16}\} \cup \{\tfrac{1}{2} \pm (\tfrac{1}{2})^{n} \mid n \in \mathbb{N}\}.$$

8. $G = [16, 13]$ has only one proper verbal subgroup, the derived subgroup, which has order 2.
$$S(G) = \left\{\tfrac{1}{16}\right\} \cup \left\{\tfrac{1}{2} \pm (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\}.$$
Any non-commutator word has equally sized fibres over G. Although $[16, 3]$ and $[16, 13]$ are both non-abelian extensions of $C_4 \times C_2$ by C_2, they do not have the same set of associated probabilities.

9. $G = [32, 2]$ has two proper verbal subgroups, of order 2 (the derived subgroup) and 8 (the centre).
$$S(G) = \{0, \tfrac{1}{32}, \tfrac{1}{8}\} \cup \{\tfrac{1}{2} \pm \left(\tfrac{1}{2}\right)^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{8} \pm \left(\tfrac{1}{2}\right)^{2n+1} \mid n \in \mathbb{N}\}.$$

10. $G = [32, 4]$ has two proper verbal subgroups, of order 2 (the derived subgroup) and 8 (the centre).
$$S(G) = \{0, \tfrac{1}{32}, \tfrac{1}{8}, \tfrac{1}{2}\} \cup \{\tfrac{1}{2} \pm \left(\tfrac{1}{2}\right)^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{8} \pm \left(\tfrac{1}{2}\right)^{2n+1} \mid n \in \mathbb{N}\}.$$
Although G has a verbal subgroup of order 8, $\frac{1}{8}$ is not an accumulation point.

11. $G = [32, 27]$ has only one verbal subgroup, the derived subgroup, which has order 4.
$$S(G) = \{\tfrac{1}{32}\} \cup \{\tfrac{1}{4} + 3\left(\tfrac{1}{2}\right)^{n+1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{4} - \left(\tfrac{1}{2}\right)^{n+1} \mid n \in \mathbb{N}\}.$$

12. $G = [27, 3] = U(3, 3)$ has only one proper verbal subgroup, the derived subgroup,

which has order 3.

$$S(G) = \{\tfrac{1}{27}\} \cup \{\tfrac{1}{3} + 2\left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{3} - \left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\}.$$

If $G = B(2, p)$ where $p \geq 3$ is any prime, then the above generalises exactly so that

$$S(G) = \{|G|^{-1}\} \cup \{\tfrac{1}{p} + (p-1)\left(\tfrac{1}{p}\right)^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{p} - \left(\tfrac{1}{p}\right)^{2n-1} \mid n \in \mathbb{N}\}.$$

13. $G = [27, 4]$ has only one verbal subgroup, the derived subgroup, of order 3.

$$S(G) = \{\tfrac{1}{27}, \tfrac{1}{3}\} \cup \{\tfrac{1}{3} + 2\left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\} \cup \{\tfrac{1}{3} - \left(\tfrac{1}{3}\right)^{2n-1} \mid n \in \mathbb{N}\}.$$

In light of the above results for several small nilpotent groups, we state some conjectures regarding the set of probabilities associated with finite nilpotent groups.

Conjecture 3.16.1. *Let G be a finite nilpotent group (of class 2), and n be an accumulation point of $S(G)$. Then G has a proper verbal subgroup of order $\frac{1}{n}$.*

Conjecture 3.16.2. *Let G be a finite nilpotent group (of class 2). Then*

$$P(G, w = g) \geq |G|^{-1}$$

for any g in the image of w.

This last conjecture is concerned with what we call the *infimum problem*. We shall discuss this further in the next chapter.

Chapter 4

The infimum problem

4.1 The problem

In [23], Nikolov and Segal present a proof of the following theorem.

Theorem 4.1.1 (Nikolov and Segal)**.** *Let G be a finite group, and put $\varepsilon(G) = p^{-|G|}$ where p is the largest prime divisor of $|G|$. G is nilpotent if and only if*

$$\inf_{w,g} P(G, w = g) > 0,$$

where w ranges over all words and g ranges over G_w^+, and this holds if and only if

$$\inf_{w,g} P(G, w = g) > \varepsilon(G).$$

There is no reason to believe that the bound $\varepsilon(G)$ is best possible. In an attempt to gain some insight into whether this bound may be improved, we have investigated what we call the 'infimum problem' for individual groups. By this we simply mean, "Given a finite group G, what is $\inf_{w,g} P(G, w = g)$?" Since we are taking the infimum over the probabilities associated with all words, and all elements in the image of those words (i.e. those that have a non-zero probability), this infimum can be more concisely described as

$$\inf S(G) \setminus \{0\}.$$

We shall use both types of notation in this chapter.

The infimum problem is of course easily answered for abelian groups. Since any word over an abelian group is a homomorphism, we have that if G is a finite abelian group and g is in G_w^+, then $P(G, w = g) \geq |G|^{-1}$ (see Corollary 2.4.2). Thus for abelian groups

$$\inf S(G) \setminus \{0\} = |G|^{-1}.$$

Nikolov and Segal have provided an answer to the infimum problem for all finite non-nilpotent groups through Theorem 4.1.1. If G is a finite non-nilpotent group,

$$\inf S(G)\setminus\{0\} = 0,$$

i.e. we may find arbitrarily small probabilities associated with G.

In the previous chapter, we calculated the set of probabilities associated with several small nilpotent groups of class 2. Looking at the results one sees that for these groups

$$\inf S(G)\setminus\{0\} = |G|^{-1}$$

just as for abelian groups. In [16], Levy has has provided a part solution to the infimum problem for nilpotent groups of class 2. He proves that if G is a finite group of nilpotency class 2, then $P(G, w = 1) \geq |G|^{-1}$ for any word w. The question of whether there might exist some finite group of nilpotency class 2 such that $P(G, w = g) < |G|^{-1}$ for some word w and some non-zero group element g remains open.

In this chapter, we look beyond groups of nilpotency class 2, and answer the infimum question for two infinite families of nilpotent groups that contain groups of every nilpotency class larger than 2. These families are the nilpotent dihedral groups and the generalised quaternion groups. In each of these cases, the infimum is again $|G|^{-1}$. The content of this chapter is all original work.

4.2 Nilpotent dihedral groups

First note that for any finite group G, the word $w(x) = x$ satisfies

$$P(G, w = g) = |G|^{-1}$$

for all g in G. Thus for any finite group,

$$\inf S(G)\setminus\{0\} \leq |G|^{-1}.$$

Thus in order to prove that the above infimum is equal to $|G|^{-1}$ for some group G it suffices to show that $P(G, w = g) \geq |G|^{-1}$ for any word w and any element g in the image of w.

We shall first explore the infimum question for the nilpotent dihedral groups, namely those whose order is a power of 2. Let D_{2n} be the dihedral group of order $2n$ given by the presentation

$$D_{2n} = \left\langle a, b \mid a^n = b^2 = 1, bab = a^{-1} \right\rangle.$$

In this section we shall prove the following.

Theorem 4.2.1. *Let G be a nilpotent dihedral group. Then*

$$\inf_{w,g} P(G, w = g) = |G|^{-1}.$$

where w ranges over all words and g ranges over G_w^+.

For completeness and comparison we shall first explicitly state the set of probabilities associated with D_4 and D_8 before proving Theorem 4.2.1 by induction.

4.2.1 Probabilities associated with D_4

Let us quickly consider D_4. This group is more commonly known as V_4, the Klein 4-group, and is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. Since V_4 is abelian and verbally simple, the set of associated probabilities is given by

$$S(D_4) = \left\{0, \tfrac{1}{4}, 1\right\}.$$

(See §2.4, particularly Lemma 2.4.3, for a discussion about the probabilities associated with abelian groups). Thus Theorem 4.2.1 holds for D_4.

4.2.2 Probabilities associated with D_8

We have seen in §3.3 that $S(D_8) = \{\frac{1}{8}\} \cup \{\frac{1}{2} \pm (\frac{1}{2})^n \mid n \in \mathbb{N}\}$ and

$$S(D_8, 1) = \{\tfrac{1}{8}\} \cup \{\tfrac{1}{2} + (\tfrac{1}{2})^n \mid n \in \mathbb{N}\}, \tag{4.1}$$

$$S(D_8, a^2) = \{\tfrac{1}{8}\} \cup \{\tfrac{1}{2} - (\tfrac{1}{2})^n \mid n \in \mathbb{N}\}, \tag{4.2}$$

$$S(D_8, g) = \{0, \tfrac{1}{8}\}, \quad \forall g \notin Z(G). \tag{4.3}$$

By inspection of (4.1) - (4.3) we conclude that $P(D_8, w = g) \geq \frac{1}{8}$ for all words w and $g \in G_w^+$. Therefore D_8 satisfies the result of Theorem 4.2.1.

4.2.3 Probabilities associated with D_{2^n}

Let $G = D_{2^n}$ where $n \geq 2$. We shall show by induction that $P(G, w = g) \geq |G|^{-1}$ for any word w and any g in the image of w. Before we begin the proof, we shall first prove the following lemma.

Lemma 4.2.2. *Let $1 \leq i \leq 2^{n-1} - 1$ be such that $i \neq 2^{n-2}$. Then there exists some $s \in \mathbb{Z}_{2^{n-1}}^\times$ such that*

$$i(s-1) \equiv 2^{n-2} \mod 2^{n-1}. \tag{4.4}$$

By $\mathbb{Z}_{2^{n-1}}^{\times}$ we mean the group of units of the ring of integers modulo 2^{n-1}. Since the units, i.e. those elements with multiplicative inverses, are the integers modulo 2^{n-1} that are coprime to 2^{n-1}, $\mathbb{Z}_{2^{n-1}}^{\times}$ consists of the odd integers mod 2^{n-1}.

Proof of Lemma 4.2.2. Suppose i is odd. Then $i \in \mathbb{Z}_{2^{n-1}}^{\times}$ so has some inverse i^{-1}. Let

$$s := i^{-1}2^{n-2} + 1 \qquad \mod 2^{n-1}.$$

Then s satisfies (4.4) and since s is odd, it must be in $\mathbb{Z}_{2^{n-1}}^{\times}$.

Suppose i is even. Then we may write $i = 2^j u$ where $u \in \mathbb{Z}_{2^{n-1}}^{\times}$ is odd and $j \in \{1, \ldots, n-2\}$. Suppose $j = n-2$. Then if $u = 1$ we would have $i = 2^{n-2}$ which we have supposed is not the case, and if $u \geq 2$ then $i = 2^{n-2}u \geq 2^{n-1}$ which again cannot happen. So in fact $j \in \{1, \ldots, n-3\}$. Now let $s := u^{-1}2^{n-2-j} + 1$ (note that this makes sense since $j \leq n-2$). Since s is odd, s must be in $\mathbb{Z}_{2^{n-1}}^{\times}$ and

$$i(s-1) = 2^j u(u^{-1}2^{n-2-j} + 1 - 1) = 2^{n-2}.$$

Thus s is a solution of (4.4). □

Proof of Theorem 4.2.1. We have seen in §4.2.1 and §4.2.2 that the result holds for D_4 and D_8, so assume $G = D_{2^n}$ where $n \geq 4$. Let

$$G = \left\langle a, b \mid a^{2^{n-1}} = b^2 = 1, bab = a^{-1} \right\rangle.$$

We shall proceed by induction.

Definition 4.2.3. *For a group G call $g, h \in G$ **auto-equivalent** and write $g \sim h$ if there is an automorphism γ of G such that $\gamma(g) = h$.*

It is easily shown that this is an equivalence relation. We denote the equivalence class of $g \in G$ by $[g]$. Note that for any word w we have that $g \sim h$ in G implies $P(G, w = g) = P(G, w = h)$, since if $\gamma(g) = h$ for some automorphism γ then

$$w(g_1, \ldots, g_n) = g \qquad \iff \qquad w(\gamma(g_1), \ldots, \gamma(g_n)) = \gamma(g) = h.$$

It is known that for $m \in \mathbb{N}$, the automorphism group of D_{2m} is given by

$$Aut(D_{2m}) = \{\gamma_{s,t} \mid s \in \mathbb{Z}_m^{\times}, t \in \mathbb{Z}_m\}$$

where $\gamma_{s,t}$ is given by

$$\gamma_{s,t}(a^i) = a^{is},$$
$$\gamma_{s,t}(a^i b) = a^{is+t} b.$$

Claim 4.2.4. *Let g be a non-central element of D_{2^n}, i.e. $g \notin \left\{1, a^{2^{n-2}}\right\}$. Then g is auto-equivalent to $ga^{2^{n-2}}$.*

Proof of Claim 4.2.4. If $g = a^i b$ for some $0 \leq i \leq 2^n - 1$ then $g \sim ga^{2^{n-2}}$ because

$$\gamma_{1,2^{n-2}}(g) = \gamma_{1,2^{n-2}}(a^i b) = a^{i+2^{n-2}} b = ga^{2^{n-2}}.$$

If $g = a^i$ for some $0 \leq i \leq 2^n - 1, i \neq 2^{n-1}$, then by Lemma 4.2.2 there exists some $s \in \mathbb{Z}^\times_{2^{n-1}}$ such that $is \equiv 2^{n-2} + i$. Then

$$a^{is} = a^{2^{n-2}+i}.$$

Thus the automorphism $\gamma_{s,1}$ satisfies

$$\gamma_{s,1}(g) = \gamma_{s,1}(a^{is}) = a^{2^{n-2+i}} = ga^{2^{n-2}}$$

so that $g \sim ga^{2^{n-2}}$. Combining both cases we see that the claim holds. □

Let z denote the non-trivial central element $a^{2^{n-2}}$ and let g be a non-central element. By the claim, $g \sim gz$, so for any word w,

$$P(D_{2^n}, w = g) = P\left(D_{2^n}, w = gz\right).$$

Since $Z := Z(G) = \{1, z\}$,

$$P(D_{2^n}, w = g) = \tfrac{1}{2} P(D_{2^n}, w \in gZ) = \tfrac{1}{2} P(D_{2^{n-1}}, w = g^*)$$

where g^* is the image of g under the canonical homomorphism $D_{2^n} \to D_{2^n}/Z \cong D_{2^{n-1}}$. Thus for any non-central element g,

$$S(D_{2^n}, w = g) = \tfrac{1}{2} S(D_{2^{n-1}}, g^*).$$

Thus by induction, if $w \in F_\infty$ and $g \in G_w^+$,

$$P(D_{2^n}, w = g) = \tfrac{1}{2} P(D_{2^{n-1}}, w = g^*) \geq \tfrac{1}{2} \tfrac{1}{2^{n-1}} = |G|^{-1}.$$

It remains to consider $g \in Z(G)$. If $w \equiv 1$ is the trivial word, then clearly $P(G, w = g) = 1 \geq |G|$ for all $g \in G_w^+$. So suppose w is not the trivial word.

We first make the following observation.

Claim 4.2.5. *If w is not the trivial word $w \equiv 1$ then $Z(G) \leq w(G)$.*

Proof of Claim. Since $[b^{\beta_1}a^{\alpha_1}, b^{\beta_2}a^{\alpha_2}] = a^{2(\beta_1\alpha_2 - \alpha_1\beta_2)}$ it follows that the derived subgroup G' is $\langle a^2 \rangle$. So every commutator word must satisfy $w(G) \leq \langle a^2 \rangle$. Since every non-trivial subgroup of $\langle a^2 \rangle$ contains $a^{2^{n-2}}$ it follows that $Z(G) \leq w(G)$.

If m is odd then $G^m := \langle g^m \mid g \in G \rangle = G$ and so $Z(G) \leq G^m$. If m is even then $G^m \leq \langle a^2 \rangle$, (since elements of the form ba^i have order 2).

By Hall's collecting process (see [10]), every word w may be written in the form

$$w = x_1^{r_1} \dots x_k^{r_k} K(x_1, \dots, x_k)$$

where $K \in F_k'$. Thus it follows that $w(G)$ must either be G or a subgroup of $\langle a^2 \rangle$. Either way, it must contain $Z(G)$. □

We consider two cases.

Case i. $w(G)$ contains a non-central element.

Let $H \leq G$ be the subgroup generated by a^2 and b. Then

$$H = \left\langle a^2, b \mid (a^2)^{2^{n-1}} = b^2 = 1, ba^2b = (a^2)^{-1} \right\rangle \cong D_{2^{n-1}}.$$

Consider $w(G/Z) \cong w(D_{2^{n-1}})$. This cannot be trivial, since then we would have $w(G) \leq Z(G)$ which is not the case. Thus by claim 4.2.5, $\bar{z} \in w(D_{2^{n-1}})$ where $\bar{z}$ is the non-trivial central element of $D_{2^{n-1}}$.

Now we use the fact that $G = D_{2^n}$ is the semi-direct product of $A := \langle a \rangle \cong \mathbb{Z}_{2^{n-1}}$ and $B := \langle b \rangle \cong \mathbb{Z}_2$ where b acts on a by inversion. Fix some k-tuple $\mathbf{b} = (b_1, \dots, b_k) \in B^{(k)}$ and define the map $w_{\mathbf{b}} : A^{(k)} \to G$ by

$$w_{\mathbf{b}}(\mathbf{a}) = w_{\mathbf{b}}(a_1, \dots, a_k) = w(a_1b_1, \dots, a_kb_k).$$

Since b acts on a by inversion, $w_{\mathbf{b}}$ is either a word over A or may be written as $b\tilde{w}(x_1, \dots, x_k)$ for some word $\tilde{w}$ over A. Either way, since $A \cong \mathbb{Z}_{2^{n-1}}$ is abelian, $w_{\mathbf{b}}$ has equally sized fibres.

Call $\mathbf{b} \in B^{(k)}$ **successful** if z, the non-trivial central element of D_{2^n}, is in $H_{\mathbf{b}}^+$, i.e. the image of $w_{\mathbf{b}}$ restricted to H. So $\mathbf{b}$ is successful if $w_{\mathbf{b}}$ is a word over $\langle a^2 \rangle$ that is non-trivial. Of course if $\mathbf{b}$ is successful then 1 is also in the image of $w_{\mathbf{b}}$ restricted to H.

Since A is abelian, we may write any word w over A in the form

$$w(x_1, \ldots, x_k) = x_1^{r_1} \ldots x_k^{r_k}.$$

Let $A^2 := \langle a^2 \rangle$. Then by Lemma 2.4.6, if $\gcd(r_1, \ldots, r_k) = 2^i j$ where j is odd we have

$$|w_{\mathbf{b}}(A)| = |w_{\mathbf{b}}(\mathbb{Z}_{2^{n-1}})| = |(\mathbb{Z}_{2^{n-1}})^{2^i j}| = 2^{n-1-i}, \tag{4.5}$$

$$|w_{\mathbf{b}}(A^2)| = |w_{\mathbf{b}}(\mathbb{Z}_{2^{n-2}})| = |(\mathbb{Z}_{2^{n-2}})^{2^i j}| = 2^{n-2-i}. \tag{4.6}$$

Thus if $\mathbf{b}$ is successful, the number of solutions in $H^{(k)}$ yielded by $w_{\mathbf{b}}$, denoted by f_H is

$$f_H(\mathbf{b}) = \tfrac{|A^2|^k}{|w_{\tilde{\mathbf{b}}}(A^2)|} \tag{4.7}$$

and the number of solutions in $G^{(k)}$ yielded by $w_{\mathbf{b}}$ is

$$f_G(\mathbf{b}) = \tfrac{|A|^k}{|w_{\mathbf{b}}(A)|} = \tfrac{2^k |A^2|^k}{2|w_{\mathbf{b}}(A^2)|} = 2^{k-1} f_H(\mathbf{b}).$$

Thus

$$|w_{D_{2^n}}^{-1}(z)| \geq \sum_{\text{successful } \mathbf{b}} f_G(\mathbf{b}) = \sum_{\text{successful } \mathbf{b}} 2^{k-1} f_H(\mathbf{b}) = 2^{k-1} |w_{D_{2^{n-1}}}(\bar{z})|.$$

Hence

$$P(D_{2^n}, w = z) \geq \tfrac{1}{2^{nk}} 2^{k-1} |w_{D_{2^{n-1}}}(\tilde{z})| = \tfrac{1}{2} P(D_{2^{n-1}}, w = \tilde{z}).$$

Thus by induction, $P(D_{2^n}, w = z) \geq |G|^{-1}$. If $\mathbf{b}$ is successful in D_{2^n}, then certainly $1 \in w_{\mathbf{b}}(A)$. Thus

$$P(D_{2^n}, w = 1) \geq P(D_{2^n}, w = z) \geq |G|^{-1}.$$

This concludes case i.

Case ii. $w(G) = Z(G)$.

Here $w(D_{2^{n-1}}) = \{1\}$, so we cannot proceed as above. Instead we adopt the method used in Chapter 3 of converting the problem of counting solutions of a word over a group into the problem of counting solutions of a polynomial over a finite field.

The following calculations will be of use.

1. $(a^{\alpha_1} b^{\beta_1})^2 = 0$ if $\beta_1 = 1$ and $(a^{\alpha_1} b^{\beta_1})^2 = a^{2\alpha_1}$ if $\beta_1 = 0$. Thus $(a^{\alpha_1} b^{\beta_1})^2 = a^{2\alpha_1(1-\beta_1)}$. Therefore $(a^{\alpha_1} b^{\beta_1})^{2^{n-2}} = a^{2^{n-2}\alpha_1(1-\beta_1)}$.

2. Let γ_i denote the left normed commutator of length i. Then it is routine to check

that

$$\gamma_2(b^{\beta_i}a^{\alpha_i}, b^{\beta_j}a^{\alpha_j}) := [b^{\beta_i}a^{\alpha_i}, b^{\beta_j}a^{\alpha_j}] = a^{2(\beta_j\alpha_i-\beta_j\alpha_i)},$$
$$\gamma_3(b^{\beta_i}a^{\alpha_i}, b^{\beta_j}a^{\alpha_j}, b^{\beta_l}a^{\alpha_l}) = [a^{2(\beta_j\alpha_i-\beta_j\alpha_i)}, b^{\beta_l}a^{\alpha_l}] = a^{4\alpha_l(\alpha_i\beta_j-\alpha_j\beta_i)},$$
$$\gamma_m(\alpha,\beta) = a^{(-1)^m 2^{m-1}\alpha_3\ldots\alpha_m(\beta_1\alpha_2-\beta_2\alpha_1)}.$$

Let the variables of w be written in the form $x_i = a^{\alpha_i}b^{\beta_i}$. Let $\alpha := (\alpha_1, \ldots, \alpha_k)$ and $\beta := (\beta_1, \ldots, \beta_k)$. Then if

$$\tilde{w}(\alpha,\beta) := w(a^{\alpha_1}b^{\beta_1}, \ldots, a^{\alpha_k}b^{\beta_k}),$$

since $w(G) = Z(G) = \left\langle a^{2^{n-2}} \right\rangle$, by the above calculations we may write w as

$$w(\alpha,\beta) = \left(a^{2^{n-2}}\right)^{p(\alpha,\beta)} = z^{p(\alpha,\beta)}$$

where $p : \mathbb{F}_2^k \times \mathbb{F}_2^k \to \mathbb{F}_2$ is a polynomial of degree at most $n-1$. We now use the following Chevalley-Warning Theorem (see [17], Theorem 6.11).

Theorem 4.2.6. *Let $f_1, \ldots, f_m \in \mathbb{F}_p[x_1, \ldots, x_k]$ with $d = deg(f_1) + \ldots + deg(f_m) < k$. If the number N of $(c_1, \ldots, c_k) \in \mathbb{F}_p^n$ with $f_i(c_1, \ldots, c_k) = 0$ for $1 \leq i \leq m$ satisfies $N \geq 1$, then $N \geq p^{k-d}$.*

By this theorem, $p = 0$ and $p = 1$ have at least $2^{k-(n-1)}$ solutions each. So if $g \in Z(G)$,

$$P(G, w = g) \geq \tfrac{2^{k-(n-1)}}{2^k} = 2^{n-1} \geq 2^{-n} = |G|^{-1}.$$

We have shown that for any word and any element in its image, $P(G, w = g) \geq |G|^{-1}$. Since this bound is attained by the word $w(x) = x$, it follows that for any finite nilpotent dihedral group G,

$$\inf_{w,g} P(G, w = g) = |G|^{-1}.$$

□

In the next section we shall show that this result also holds for the generalized quaternion groups.

4.3 Generalized quaternion groups

For $n \geq 3$ let Q_{2^n} denote the group given by the presentation

$$Q_{2^n} = \left\langle a, b \mid a^{2^{n-1}} = 1, b^2 = a^{2^{n-2}}, b^{-1}ab = a^{-1} \right\rangle. \tag{4.8}$$

Q_{2^n} is called the **generalized quaternion group** of order 2^n.

The structure of Q_{2^n} is similar to that of the dihedral group of order 2^n in many ways, because of the similarity of their presentations. Q_{2^n} is nilpotent of class $n-1$ and the centre of the group is $\left\langle a^{2^{n-2}} \right\rangle \cong \mathbb{Z}_2$. By inspecting the presentation one sees that $Q_{2^n}/Z(Q_{2^n}) \cong D_{2^{n-1}}$ and $\left\langle a^2, b \right\rangle \cong Q_{2^{n-1}}$.

Let us discuss the verbal subgroups of $G = Q_{2^n}$. Since G is a 2-group, if $w(x) = x^i$ then

$$w(G) = \begin{cases} G & \text{if } i \text{ is odd,} \\ \left\langle a^j \right\rangle & \text{if } i = 2^j m \text{ where } m \text{ odd.} \end{cases}$$

Since $[b^{\beta_i} a^{\alpha_i}, b^{\beta_j} a^{\alpha_j}] = a^{2(\beta_j \alpha_i - \beta_i \alpha_i)}$, it follows that $G' = \left\langle a^2 \right\rangle$. Thus since any word w can be written in the form

$$w(x_1, \ldots, x_k) = \prod_i x_i^{r_i} K(x_1, \ldots, x_k)$$

where $K \in F_k'$, the proper verbal subgroups of G are those of the form

$$\left\langle a^{2^i} \right\rangle \tag{4.9}$$

for some $1 \leq i \leq n-2$.

Theorem 4.3.1. *Let G be a generalized quaternion group. Then*

$$\inf_{w,g} P(G, w = g) = |G|^{-1}$$

where w ranges over all words, and g over G_w^+.

Proof. This follows in much the same way as Theorem 4.2.1, the analogous theorem for nilpotent dihedral groups.

By Theorem 3.4.1,

$$\begin{aligned} S(Q_8) &= \left\{\tfrac{1}{8}\right\} \cup \left\{\tfrac{1}{2} \pm (\tfrac{1}{2})^n \mid n \in \mathbb{N}\right\}, \\ S(Q_8, 1) &= \left\{\tfrac{1}{8}\right\} \cup \left\{\tfrac{1}{2} + (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\} \cup \left\{\tfrac{1}{2} - (\tfrac{1}{2})^{2n} \mid n \in \mathbb{N}\right\}, \\ S(Q_8, -1) &= \left\{\tfrac{1}{8}\right\} \cup \left\{\tfrac{1}{2} - (\tfrac{1}{2})^{2n-1} \mid n \in \mathbb{N}\right\} \cup \left\{\tfrac{1}{2} + (\tfrac{1}{2})^{2n} \mid n \in \mathbb{N}\right\}, \\ S(Q_8, g) &= \left\{0, \tfrac{1}{8}\right\}, \quad \forall g \notin \{-1, 1\}. \end{aligned}$$

Thus $\inf S(Q_8) \setminus \{0\} = |Q_8|^{-1}$.

Let $G = Q_{2^n}$ for $n \geq 4$. Let $w \in F_\infty$ and let $g \in G$ be a non-central element. Let $z := a^{2^{n-2}} = b^2$ denote the non-trivial central element of G.

Claim 4.3.2. *g is auto-equivalent to gz.*

Proof. Recall that by auto-equivalent we mean that there exists some automorphism ϕ such that $\phi(g) = gz$.

- Suppose $g = ba^i$ for some $0 \leq i \leq 2^{n-1} - 1$. Let $\phi_1 : G \to G$ be given by

$$\begin{aligned}\phi_1(a^j) &= a^j \\ \phi_1(ba^j) &= ba^{j+2^{n-2}}.\end{aligned}$$

  It is routine to check that ϕ_1 is an automorphism. Since

$$\phi_1(g) = \phi_1(ba^i) = ba^{2^{n-2}} = gz,$$

  we have $g \sim gz$.

- Suppose $g = a^i$ for some $1 \leq i \leq 2^{n-1}$, $i \neq 2^{n-2}$. By Lemma 4.2.2 there exists some $s \in \mathbb{Z}_{2^{n-1}}^{\times}$ such that

$$i(s-1) \equiv 2^{n-2} \mod 2^{n-1}.$$

  Let $\phi_2 : G \to G$ be given by

$$\begin{aligned}\phi_2(a^j) &= a^{js} \\ \phi_2(ba^j) &= ba^{js+1}.\end{aligned}$$

  It is routine to check that ϕ_2 is an automorphism. Now

$$\phi_2(g) = \phi_2(a^i) = a^{is} = a^{i+2^{n-2}} = gz.$$

  Thus $g \sim gz$.

In either case $g \sim gz$ and the claim holds. □

Now back to the proof of the theorem. For non-central g, since $g \sim gz$ and $gZ = \{g, gz\}$ we have

$$\begin{aligned}P(G, w = g) &= \tfrac{1}{2}P(G, w \in gZ) \\ &= \tfrac{1}{2}P(G/Z, w = gZ) \\ &= \tfrac{1}{2}P(D_{2^{n-1}}, w = g^*)\end{aligned}$$

where g^* is the image of g under the canonical homomorphism $G \to G/Z \to D_{2^{n-1}}$. Thus by Theorem 4.2.1 we have

$$P(G, w = g) \geq \tfrac{1}{2}\tfrac{1}{2^{n-1}} = |G|^{-1}.$$

It remains to consider solutions to $w = g$ where g is central. If $w \equiv 1$ is the trivial word, the result holds, so assume that w is not trivial. By inspection of (4.9) we see that every non-trivial verbal subgroup contains the centre. As in the proof for nilpotent dihedral groups, we consider two cases.

Case i. $w(G) \neq Z(G)$.

Here $w(G)$ strictly contains $Z(G)$. Any element of G may be (uniquely) written as $b^{\beta_i}a^{\alpha_i}$ for $\beta_i \in \{0,1\}$ and $\alpha_i \in \{0,\dots,2^n-1\}$. Let $A = \langle a \rangle \cong \mathbb{Z}_{2^{n-1}}$. For $\mathbf{b} = (b^{\beta_1},\dots,b^{\beta_k})$, $\beta_i \in \{0,1\}$ let $w_{\mathbf{b}} : A^{(k)} \to G$ be given by

$$w_{\mathbf{b}}(x_1,\dots,x_k) := w(b^{\beta_1}x_1,\dots,b^{\beta_k}x_k).$$

Since $b^2 = a^{2^{n-2}}$ and $b^{-1}ab = a^{-1}$, $w_{\mathbf{b}}$ is either a word over A or $w_{\mathbf{b}} = b^i\tilde{w}$ where $\tilde{w}$ is a word over A and $i \in \{1,2,3\}$.

Let $H := \langle a^2, b \rangle \leq G$. Then $H \cong Q_{2^{n-1}}$ (to see this, consider the presentation of H given the presentation for Q_{2^n} as stated in (4.8)).

Call $\mathbf{b} \in B^{(k)}$ **successful** if $z \in w_{\mathbf{b}}(H)$. Since $z = a^{2^{n-2}} = b^2$, this can only occur if $w_{\mathbf{b}}$ is a word over A, or $w_{\mathbf{b}} = \tilde{w}b^2$, where $\tilde{w}$ is a word over A. In either case, $w_{\mathbf{b}}$ must have equally sized fibres. Let $f_H(\mathbf{b})$ denote the number of solutions to $w_{\mathbf{b}} = b^2$ in H, and $f_G(\mathbf{b})$ the number of solutions in G. Then analogously to (4.7), successful $\mathbf{b}$ satisfy

$$f_G(\mathbf{b}) = 2^{k-1}f_H(\mathbf{b}).$$

Therefore for $g \in Z(G)$

$$|w^{-1}_{Q_{2^n}}(g)| \geq \sum_{\mathbf{b}\text{ successful}} f_G(\mathbf{b}) = \sum_{\mathbf{b}\text{ successful}} 2^{k-1}f_H(\mathbf{b}) = 2^{k-1}|w^{-1}_{Q_{2^{n-1}}}(g)|$$

and so by induction

$$P(Q_{2^n}, w = g) \geq \tfrac{1}{2}P(Q_{2^{n-1}}, w = g) \geq \tfrac{1}{2}(\tfrac{1}{2})^{n-1} = |G|^{-1}$$

and the result holds.

Case ii. $w(G) = Z(G)$.

Since

- $(b^\beta a^\alpha)^{2r} = a^{r(j2^{n-2}+2^i(1-j))}$
- $[\dots[b^{\beta_1}a^{\alpha_1}, b^{\beta_2}a^{\alpha_2}],\dots,b^\beta_m a^\alpha_m] = a^{(-1)^m 2^{m-1}(\beta_1\alpha_2-\beta_2\alpha_1)\beta_3\dots\beta_m}$

it follows that any such word over G may be written as

$$w(\alpha, \beta) = z^{p(\alpha,\beta)}$$

where the polynomial $p(\alpha, \beta)$ has degree $\leq n-1$. Consider p as a map from $\mathbb{F}_2^k \times \mathbb{F}_2^k \to \mathbb{F}_2$. Then by Theorem 4.2.6 we find that the number of solutions N to $p = 0$ or $p = 1$ satisfies $N \geq 2^{2k-(n-1)}$. Thus for $g \in Z(G)$

$$P(G, w = g) = 2^{-2k}N \geq 2^{n-1} \geq 2^n = |G|^{-1}$$

and the result holds. □

4.4 Summary

In the previous chapter, we calculated $S(G)$ for several small nilpotent groups of class 2, and found that in these examples, the infimum of $S(G)\setminus\{0\}$ was $|G|^{-1}$, just as for all abelian groups. In this chapter we have proved that if G is a nilpotent dihedral group or a generalized quaternion group (of any nilpotency class) then the infimum of the positive probabilities associated with G is again $|G|^{-1}$. This tentatively leads to the following conjecture.

Conjecture 4.4.1. *Let G be a finite nilpotent group. Then*

$$\inf S(G)\setminus\{0\} = |G|^{-1}.$$

Of course this is based on very little evidence, so may well not be the case. If not, it would be interesting to see which nilpotent groups satisfy this property. This remains an open problem.

Chapter 5

Non-nilpotent groups

Throughout this thesis we are interested in the relationship between a finite group G and its associated set of probabilities $S(G)$. We are interested in how the group properties of G are reflected in properties of $S(G)$ and vice versa. We suspect that the following conjecture may be true.

Conjecture 5.0.2. *Let G be a finite group. The set of probabilities associated with G is dense in the interval $[0, 1]$ if, and only if, G is non-nilpotent.*

In this chapter we explain what we have found regarding non-nilpotent groups. We show that the conjecture holds for some significant infinite families of non-nilpotent groups.

We begin by giving a constructive proof that $S(\mathrm{Sym}(3))$ is dense in the interval $[0, 1]$, where $\mathrm{Sym}(3)$ is the symmetric group on three letters. The proof is constructive in the sense that not only does it explicitly define an infinite set of words whose associated probabilities are shown to be a dense subset of $[0, 1]$, but given a desired probability within that dense set, it provides a recipe for finding a word with which that probability is associated. Theoretically, given a target probability p in $[0, 1]$ and a margin of error ε, one could use the proof to construct a word w and a find a group element g such that $\mid P(G, w = g) - p \mid < \varepsilon$.

In the second section we generalise this result to include all dihedral groups of order $2m$ where m is odd. The proof is analogous to that given for $\mathrm{Sym}(3)$ in the previous section.

In §5.3 we prove that $S(G)$ is dense in $[0, 1]$ for any non-nilpotent dihedral group. We will use the proof of the result from §5.2 to do this.

In §5.4 we extend the result to include all generalised dihedral groups that are not 2-groups. By a generalised dihedral group we mean any group that is the semi-direct

product of an abelian group H by $\mathbb{Z}_2$, where the non-trivial element of $\mathbb{Z}_2$ acts on H by inversion.

In the final section we give a proof very similar to that for Sym(3) to show that the set of probabilities associated with Alt(4), the alternating group on four elements, is also dense in $[0, 1]$. We end this section with two corollaries that state that the above conjecture holds for all alternating and symmetric groups.

5.1 The symmetric group on three elements

Theorem 5.1.1. *Let Sym(3) denote the symmetric group on three elements. Then $S(Sym(3))$, the set of probabilities associated with Sym(3), is dense in the interval $[0, 1]$.*

Before we begin the proof, we first establish some notation and methodology.

5.1.1 Using conjugacy classes

Firstly, let C_1, C_2, C_3 denote the three conjugacy classes of Sym(3), where C_1 contains only the identity, C_2 contains the two 3-cycles, and C_3 contains the three 2-cycles.

Lemma 5.1.2. *Let g, h be conjugate elements of a finite group G. Then for any word w,*

$$P(G, w = g) = P(G, w = h).$$

Proof. Let g and h be conjugate elements of a finite group G, so that $h = g^a$ for some element a. Thus

$$w(g_1, \ldots, g_k) = g \iff w(g_1, \ldots, g_k)^a = g^a \iff w({g_1}^a, \ldots, {g_k}^a) = g^a = h.$$

Since $G \to G$, $x \mapsto x^a$ is a bijection it follows that $|w^{-1}(g)| = |w^{-1}(h)|$ and thus $P(G, w = g) = P(G, w = h)$. This is a special case of the result proved in Section 4.2.3 which says that if two elements are auto-equivalent then they are equally likely to be the result of a random evaluation. □

Since $|C_2| = 2$ and $|C_3| = 3$ we have

$$P(\mathrm{Sym}(3), w = (1, 2, 3)) = \tfrac{1}{2} P(\mathrm{Sym}(3), w \in C_2), \tag{5.1}$$

$$P(\mathrm{Sym}(3), w = (1, 2)) = \tfrac{1}{3} P(\mathrm{Sym}(3), w \in C_3). \tag{5.2}$$

For simplicity of calculation, we first consider $P(\mathrm{Sym}(3), w \in C_i)$ for all i and then use the above identities to establish the probabilities associated with individual elements of the group.

5.1.2 Constructing the set of words

In order to prove the theorem, we will construct an infinite set of words, whose associated probabilities form a dense subset of $[0, 1]$. We construct each of these words from two short words w_A and w_B given by

$$\begin{aligned} w_A(x_1, x_2) &:= x_1^2 x_2^2, \\ w_B(x_1, x_2) &:= [x_1, x_2]\,. \end{aligned}$$

We construct new words by repeated substitution into the first variable. For example, we may construct a new word w_{AB} by replacing the first variable of w_B with w_A as follows.

$$\begin{aligned} w_{AB}(x_1, x_2, x_2) :&= w_B(w_A(x_1, x_2), x_3) \\ &= \left[x_1^2 x_2^2, x_3\right]. \end{aligned}$$

Let M be some product of matrices A and B, i.e. $M = \prod_{i=1}^{l} M_i$, where $M_i \in \{A, B\}$ for all i. We denote the set of all such matrices by $\langle A, B\rangle$, and say that M has *length* l.

We define the word $w_M = w_{M_1 \ldots M_l}$ by

$$\begin{aligned} w_{M_1 \ldots M_l}(x_1, \ldots, x_{l+1}) :&= w_{M_l}(w_{M_1 \ldots M_{l-1}}(x_1, \ldots, x_l), x_{l+1}) \\ &= w_{M_l}(\ldots w_{M_2}(w_{M_1}(x_1, x_2), x_3), \ldots, x_l). \end{aligned}$$

Suppose one took the infinite set of words that can be built from w_A and w_B in this way and calculated the associated probabilities for every word in that set. In the proof of Theorem 5.1.1 we shall show that this yields a dense subset of $[0, 1]$.

5.1.3 Using transition matrices

We define A, B to be the transition matrices with entries

$$\begin{aligned} a_{ij} &:= P(\mathrm{Sym}(3), w_A(g_1, g_2) \in C_j \mid g_1 \in C_i), \\ b_{ij} &:= P(\mathrm{Sym}(3), w_B(g_1, g_2) \in C_j \mid g_1 \in C_i). \end{aligned}$$

We may use these matrices to calculate the probabilities associated with the words constructed in the last section in the following way. Let v_0 be the row vector whose i^{th} component is the probability that a uniformly random element of $\mathrm{Sym}(3)$ lies in C_i. Then v_0 is given by

$$v_0 = (\tfrac{1}{6}, \tfrac{1}{3}, \tfrac{1}{2}).$$

Let v_A be the vector whose j^{th} component is $P(\text{Sym}(3), w_A \in C_j)$. Then since

$$P(\text{Sym}(3), w_A \in C_j) = \sum_{i=1}^{3} P(\text{Sym}(3), w_A(x_1, x_2) \in C_j \mid x_1 \in C_i) P(\text{Sym}(3), x_1 \in C_i),$$

it follows that $v_A = v_0 A$.

Given any word w_M constructed from w_A and w_B by the method given in Section 5.1.2, we may use the matrices A and B to calculate the probabilities associated with w_M in a similar fashion. Consider the example word w_{AB} described in that section. Then

$$\begin{aligned}
&P(\text{Sym}(3), w_{AB} \in C_j) \\
&= \sum_{i=1}^{3} P(\text{Sym}(3), w_B(w_A, x_3) \in C_j \mid w_A \in C_i) P(\text{Sym}(3), w_A \in C_i) \\
&= \sum_{i=1}^{3} P(\text{Sym}(3), w_B(x_1, x_2) \in C_j \mid x_1 \in C_i) P(\text{Sym}(3), w_A \in C_i) \\
&= v_A B_j
\end{aligned}$$

where B_j is the j^{th} column of B. Thus if v_{AB} is defined to be the vector whose i^{th} component is

$$P(\text{Sym}(3), w_{AB} \in C_i),$$

then

$$v_{AB} = v_A B = v_0 AB.$$

In general, if $I \neq M = M_1 \dots M_l \in \langle A, B \rangle$, with $w_{M_1 \dots M_l}$ as defined as in §5.1.2, then we define the row vector $v_{M_1 \dots M_l}$ to be that which has i^{th} component $P(\text{Sym}(3), w_{M_1 \dots M_l} \in C_i)$. Then

$$v_{M_1 \dots M_l} = v_0 M_1 \dots M_l.$$

Thus we may easily calculate the probabilities of any of the words constructed in §5.1.2 once A and B are known.

5.1.4 Properties of A and B

Recall that we define matrices A and B to be the matrices whose entries are

$$\begin{aligned}
a_{ij} &:= P(\text{Sym}(3), w_A(g_1, g_2) \in C_j \mid g_1 \in C_i), \text{ and} \\
b_{ij} &:= P(\text{Sym}(3), w_B(g_1, g_2) \in C_j \mid g_1 \in C_i).
\end{aligned}$$

We begin by calculating A and B.

Calculating A

Recall that $w_A(x_1, x_2) := x_1^2 x_2^2$, $C_1 := \{1\}, C_2 := \{(1,2,3),(1,3,2)\}$ and $C_3 := \{(1,2),(1,3),(2,3)\}$.

- The map $C_2 \to C_2, g \mapsto g^2$ is a bijection. For any g in C_1 or C_2, $g^2 = 1$.
  Thus

$$\begin{aligned} P(\mathrm{Sym}(3), w_A(x_1,x_2) \in C_i \mid x_1 \in C_1) &= P(\mathrm{Sym}(3), w_A(x_1,x_2) \in C_i \mid x_1 \in C_3) \\ &= P(\mathrm{Sym}(3),\ x^2 \in C_i), \end{aligned}$$

  and

$$P(\mathrm{Sym}(3),\ x^2 \in C_i) = \begin{cases} \frac{4}{6} = \frac{2}{3} & i = 1, \\ \frac{2}{6} = \frac{1}{3} & i = 2, \\ 0 & i = 3. \end{cases}$$

  Hence $(a_{11}, a_{12}, a_{13}) = (a_{31}, a_{32}, a_{33}) = (\frac{2}{3}, \frac{1}{3}, 0)$.

- Now if g_1 is a 3-cycle, then $g_1^2 g_2^2 = 1$ if, and only if, $g_2 = g_1^{-1}$. Otherwise $g_1^2 g_2^2$ is a 3-cycle, since $\left\langle g^2 \mid g \in \mathrm{Sym}(3) \right\rangle = \langle (1,2,3) \rangle$. Therefore

$$P(\mathrm{Sym}(3), w_A(x_1,x_2) \in C_j \mid x_1 \in C_2) = \begin{cases} \frac{1}{6} & i = 1, \\ \frac{5}{6} & i = 2, \\ 0 & i = 3. \end{cases}$$

  Hence $(a_{21}, a_{22}, a_{23}) = (\frac{1}{6}, \frac{5}{6}, 0)$.

A is therefore given by the following.

$$A = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{5}{6} & 0 \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix}.$$

Calculating B

Recall $w_B(x_1, x_2) := [x_1, x_2]$.

- If $g_1 = 1$, then $[g_1, g_2]$ is trivial for any g_2. Therefore

$$P(\mathrm{Sym}(3), w_B \in C_i \mid g_1 \in C_1) = \begin{cases} 1 & i = 1, \\ 0 & i = 2, \\ 0 & i = 3. \end{cases}$$

  So $(b_{11}, b_{12}, b_{13}) = (1, 0, 0)$.

- If g_1 is a 3-cycle, then g_1 commutes with g_2 exactly when $g_2 \in \langle g_1 \rangle$, i.e. when g_2 is a member of C_1 or C_2. Otherwise $[g_1, g_2]$ is a member of C_2, since the derived subgroup of Sym(3) is $\langle (1,2,3) \rangle$. Therefore

$$P(\mathrm{Sym}(3), w_B \in C_i \mid g_1 \in C_2) = \begin{cases} \frac{3}{6} = \frac{1}{2} & i = 1, \\ \frac{3}{6} = \frac{1}{2} & i = 2, \\ 0 & i = 3. \end{cases}$$

  Thus $(b_{21}, b_{22}, b_{23}) = (\frac{1}{2}, \frac{1}{2}, 0)$.

- If g_1 is a transposition, then it commutes only with itself and the trivial element. Otherwise $[g_1, g_2]$ is a member of C_2. Therefore

$$P(\mathrm{Sym}(3), w_B \in C_i \mid g_1 \in C_3) = \begin{cases} \frac{2}{6} = \frac{1}{3} & i = 1, \\ \frac{4}{6} = \frac{2}{3} & i = 2, \\ 0 & i = 3. \end{cases}$$

  Hence $(b_{31}, b_{32}, b_{33}) = (\frac{1}{3}, \frac{2}{3}, 0)$.

B is therefore given by

$$B = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix}.$$

Eigenvalues and eigenvectors

These will be useful later. It is easy to verify that A and B have the same eigenvalues, $0, \frac{1}{2}$ and 1, and that the associated eigenvectors are as follows.

- The left eigenvectors of A corresponding to $0, \frac{1}{2}, 1$ respectively are $(1, 0, -1), (1, -1, 0), (1, 2, 0)$.
- The right eigenvectors of A corresponding to $0, \frac{1}{2}, 1$ respectively are $(0, 0, 1)^t, (-2, 1, -2)^t, (1, 1, 1)^t$.
- The left eigenvectors of B corresponding to $0, \frac{1}{2}, 1$ respectively are $(1, -4, 3), (1, -1, 0), (1, 0, 0)$.
- The right eigenvectors of B corresponding to $0, \frac{1}{2}, 1$ respectively are $(0, 0, 1)^t, (0, 3, 4)^t, (1, 1, 1)^t$.

5.1.5 Products of A and B

Let $\langle A, B\rangle := \left\{\prod_{i=1}^{l} M_i \mid M_i \in \{A, B\}, l \in \mathbb{N}_0\right\}$ and let $M(\neq I) \in \langle A, B\rangle$ be of length l (i.e. M may be written as the product of l matrices, each of which is either A or B). We begin by investigating the eigenvalues and eigenvectors of M.

Eigenvalues of M

Referring to §5.1.4 we see that A and B both have right eigenvectors $v_1 := (0, 0, 1)^t$ and $v_3 := (1, 1, 1)^t$ corresponding to eigenvalues $0, 1$ respectively. Thus it follows that $Mv_1 = 0 = 0v_1$ and $Mv_3 = v_3$, i.e. 0 and 1 are also eigenvalues of M.

In the same section we also find that A and B both have left eigenvector $w_2 := (1, -1, 0)$ corresponding to the eigenvalue $\frac{1}{2}$. Thus we see that

$$w_2 M \;=\; w_2 \prod_{i=1}^{l} M_i \;=\; (w_2 M_1) \prod_{i=2}^{l} M_i \;=\; \tfrac{1}{2} w_2 \prod_{i=2}^{l} M_i \;= \ldots = \; \left(\tfrac{1}{2}\right)^l w_2.$$

Hence we conclude that M has three distinct eigenvalues, $0, (\frac{1}{2})^l$ and 1.

Remark Note that since M has three distinct eigenvalues, M is diagonalizable. What's more, since the limit $\lim_{i\to\infty} \lambda^i$ exists for all eigenvalues λ, it follows that the limit $\lim_{i\to\infty} M^i$ exists for any product M of matrices A and B. This will be useful later.

Eigenvectors of M

We are concerned with finding the *left* eigenvectors of M, which we shall denote by w_1, w_2, w_3.

- As we saw in the last section, the eigenvalue $(\frac{1}{2})^l$ has corresponding left eigenvector $w_2 = (1, -1, 0)$.
- Suppose that $M_1 = A$. Then since $(1, 0, -1)A = 0$, it follows that $(1, 0, -1)M = 0$, and $(1, 0, -1)$ is an eigenvector corresponding to the eigenvalue 0. Similarly, if $M_1 = B$, then $(1, -4, 3)$, the left eigenvector of B corresponding to eigenvalue 0, is an eigenvector of M corresponding to 0. Thus w_1 is $(1, 0, -1)$ in the first case and $(1, -4, 3)$ in the second.

Sadly finding w_3 is not quite so simple. We wish to find the unique vector w_3 that satisfies $w_3 M = w_3$ whose components sum to one. We refer to a vector whose components sum to 1 as being **normalised**. This will simplify calculations later. Since the last column of M will necessarily be a column of zeros, (since this is true of A and B), it follows that the last component of w_3 will be zero. Thus w_3 is in the span of $V := \{(1, 0, 0), (0, 1, 0)\}$, and we restrict our search to this subspace. We choose

$\{u_1 := (\frac{1}{3}, \frac{2}{3}, 0), u_2 := (1, 0, 0)\}$ as a convenient basis for V. Note that these are the normalised eigenvectors of A and B corresponding to eigenvalue 1. It is readily seen that V is both A-invariant and B-invariant, and so A and B restrict to linear operators on V. With respect to this basis, the matrix representations of the restricted operators A and B are given by

$$P = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$$

respectively. This follows since $u_1 B = u_2 A = \frac{1}{2}u_1 + \frac{1}{2}u_2$. Note that for consistency we deal here with *row* vectors, where matrices act on the right.

Suppose $w_3 = \alpha u_1 + \beta u_2$. First note that insisting that w_3 is normalised means that $\beta = 1 - \alpha$. Let $N = \prod_{i=1}^{l} N_i \in \langle P, Q \rangle$ be the matrix corresponding to M with respect to the new basis (i.e. the matrix obtained from $\prod M_i$ by replacing every A with P and B with Q). Then we wish to solve

$$(\alpha, 1-\alpha)N = (\alpha, 1-\alpha). \tag{5.3}$$

Lemma 5.1.3. *Let $N = \prod_{i=1}^{l} N_i$, where $N_i \in \{P, Q\}$ for all i. Then*

$$(\alpha, 1-\alpha)\, N = \left(\tfrac{\alpha+a}{2^l}, 1 - \tfrac{\alpha+a}{2^l}\right)$$

where $a = a(N) = a_0 2^0 + a_1 2^1 + \ldots a_{l-1} 2^{l-1}$ and $a_i = \begin{cases} 1 & \text{if } N_{i+1} = P, \\ 0 & \text{if } N_{i+1} = Q. \end{cases}$

Proof. We prove this by induction on l, the length of the product $\prod_{i=1}^{l} N_i$.

Suppose $l = 1$. Then either (i) $N = P$, or (ii) $N = Q$.

(i) $(\alpha, 1-\alpha)P = \left(\frac{\alpha+1}{2}, 1 - \frac{\alpha+1}{2}\right)$ so $a = 1$ and the claim holds in this case.

(ii) $(\alpha, 1-\alpha)Q = \left(\frac{\alpha+0}{2}, 1 - \frac{\alpha+0}{2}\right)$ so $a = 0$ and the claim holds in this case.

Suppose the claim holds for all products of length l for some $l \geq 1$. Suppose N has length $l+1$. Then we consider the cases (i) $N_{l+1} = P$ and (ii) $N_{l+1} = Q$.

(i) $N_{l+1} = P$.

$$\begin{aligned}
(\alpha, 1-\alpha)N &= (\alpha, 1-\alpha)\left(\prod_{i=0}^{l-1} N_i\right) N_{l+1} \\
&= \left(\frac{\alpha+a}{2^l}, 1-\frac{\alpha+a}{2^l}\right) P, \\
&\quad \text{where } a = \sum_{i=0}^{l-1} a_i 2^i, \quad a_i = \begin{cases} 0 & \text{if } N_{i+1} = Q, \\ 1 & \text{if } N_{i+1} = P, \end{cases} \\
&= \left(\frac{\alpha+a}{2^{l+1}} + \frac{1}{2}, 1-\left(\frac{\alpha+a}{2^{l+1}} + \frac{1}{2}\right)\right) \\
&= \left(\frac{\alpha+(a+2^l)}{2^{l+1}}, 1-\frac{\alpha+(a+2^l)}{2^{l+1}}\right).
\end{aligned}$$

$$\text{Thus } a_i = \begin{cases} 0 & \text{if } N_{i+1} = Q, \\ 1 & \text{if } N_{i+1} = P. \end{cases} \quad \forall \quad i \leq l+1.$$

(ii) $N_{l+1} = Q$.

$$\begin{aligned}
(\alpha, 1-\alpha)N &= (\alpha, 1-\alpha)\left(\prod_{i=0}^{l-1} N_i\right) N_{l+1} \\
&= \left(\frac{\alpha+a}{2^l}, 1-\frac{\alpha+a}{2^l}\right) Q, \\
&\quad \text{where } a = \sum_{i=0}^{l-1} a_i 2^i, \quad a_i = \begin{cases} 0 & \text{if } N_{i+1} = P, \\ 1 & \text{if } N_{i+1} = Q, \end{cases} \\
&= \left(\frac{\alpha+a}{2^{l+1}}, 1-\frac{\alpha+a}{2^{l+1}}\right).
\end{aligned}$$

$$\text{Thus } a_i = \begin{cases} 0 & \text{if } N_{i+1} = Q, \\ 1 & \text{if } N_{i+1} = P. \end{cases} \quad \forall i \leq l+1.$$

Thus, by induction, the claim holds. □

Remark Note that by the above lemma, for any $l \in \mathbb{N}$, and $a \in \{0, \ldots, 2^l - 1\}$, there exists $I \neq N \in \langle P, Q \rangle$ such that $(\alpha, 1-\alpha)\, N = \left(\frac{\alpha+a}{2^l}, 1-\frac{\alpha+a}{2^l}\right)$.

Now to solve equation (5.3) we must solve

$$\left(\tfrac{\alpha+a}{2^l}, 1-\tfrac{\alpha+a}{2^l}\right) = (\alpha, 1-\alpha)$$

for a as described in the above lemma. This has solution

$$\alpha = \frac{a}{2^l - 1}. \tag{5.4}$$

Thus

$$w_3 = \alpha u_1 + (1-\alpha)u_2 = \left(1 - \tfrac{2a}{3(2^l-1)}, \tfrac{2a}{3(2^l-1)}, 0\right) \tag{5.5}$$

is the normalised left eigenvector of M corresponding to eigenvalue 1.

5.1.6 The proof of Theorem 5.1.1

We may now begin the proof of Theorem 5.1.1, using w_A, w_B, A and B as above.

Proof. Let $I \neq M \in \langle A, B\rangle$ have length $l \geq 1$, and let w_M be the word constructed from w_A and w_B in the order dictated by M, as described in §5.1.2. For any positive integer i, let w_{M^i} denote the word obtained by substituting w_M into the first variable of itself i times, as in §5.1.2. Let v_{M^i} be the row vector whose j^{th} component is $P(\mathrm{Sym}(3), w_{M^i} \in C_j)$. We shall consider the limit of the sequences

$$(P(\mathrm{Sym}(3), w_{M^i} \in C_j))_{i\in\mathbb{N}}$$

and shall show that as M varies over $\langle A, B\rangle$ this gives a dense subset of probabilities within $[0, 1]$. Let $v_0 := \left(\frac{1}{6}, \frac{1}{3}, \frac{1}{2}\right)$. As we saw in §5.1.3

$$v_{M^i} := (P(\mathrm{Sym}(3), w_{M^i} \in C_j))_{j=1}^3 = v_0 M^i$$

and

$$v_{M^\infty} := \lim_{i\to\infty} v_{M^i} = \lim_{i\to\infty} (P(\mathrm{Sym}(3), w_{M^i} \in C_j))_{j=1}^3 = \lim_{i\to\infty} v_0 M^i.$$

We know that this limit exists, since we saw in §5.1.5 that M is diagonalisable, with eigenvalues $0, (\frac{2}{3})^l$ and 1. Note that

$$v_{M^\infty} M = \left(\lim_{i\to\infty} v_0 M^i\right) M = \lim_{i\to\infty} v_0 M^i = v_{M^\infty}.$$

Thus v_{M^∞} is a normalised eigenvector of M corresponding to eigenvalue 1. Since there is only one such vector (we have seen above that the eigenvalue 1 has geometric multiplicity 1), we have from (5.5) that

$$v_{M^\infty} = w_3 = \left(1 - \tfrac{2a}{3(2^l-1)}, \tfrac{2a}{3(2^l-1)}, 0\right)$$

where $a = a(M)$ is as defined in Lemma 5.1.3. Thus for any $g_2 \in C_2$,

$$\lim_{i\to\infty} P(\mathrm{Sym}(3), w_{M^i} = 1) = 1 - \frac{2a}{3(2^l-1)}, \tag{5.6}$$

$$\lim_{i\to\infty} P(\mathrm{Sym}(3), w_{M^i} = g_2) = \frac{a}{3(2^l-1)}. \tag{5.7}$$

Note, the numerator in equation (5.7) is 'a' rather than '$2a$' since by equation (5.1),

we have

$$P(\mathrm{Sym}(3), w_{M^i} = g_2) = \tfrac{1}{2}P(\mathrm{Sym}(3), w_{M^i} \in C_2) = \tfrac{1}{2}\tfrac{2a}{3(2^l-1)} = \tfrac{a}{3(2^l-1)}.$$

Now (5.6) and (5.7) correspond to limit points of $S(\mathrm{Sym}(3))$, so it remains to show that such points are dense in the interval $[0,1]$. The following lemma does the hard work.

Lemma 5.1.4. *For $I \neq M \in \{A, B\}$ define $a(M)$ as in Lemma 5.1.3 and let $l(M)$ be the length of M. Then the set*

$$A := \left\{ \frac{a(M)}{2^{l(M)} - 1} \;\middle|\; I \neq M \in \langle A, B\rangle \right\}$$

is dense in $[0,1]$.

Proof. Recall from Lemma 5.1.3 that if $M = \prod_{i=1}^{l} M_i$ then $a(M) := \sum_{i=0}^{l-1} a_i 2^i$ where $a_i = 1$ if $M_{i+1} = P$ and $a_i = 0$ if $M_{i+1} = Q$. For any $l \in \mathbb{N}$, the set

$$\{a(M) \mid M \text{ of length } l\}$$

is $\{0, 1, \ldots, 2^l - 1\}$. Hence

$$A = \bigcup_{l \in \mathbb{N}} \left\{ \frac{i}{2^l - 1} \mid 0 \leq i \leq 2^l - 1 \right\}.$$

We need to show that this set is dense in $[0,1]$. Let $p \in [0,1]$. We shall define a sequence of elements $(q_l)_{l\in\mathbb{N}}$ that lie in A and that tend to p. Consider $p(2^l - 1) \in [0, 2^l - 1]$. Choose $y_l \in \{0, 1, \ldots, 2^l - 1\}$ such that $|y_l - p(2^l - 1)| < 1$. Then $q_l := \frac{y_l}{2^l - 1} \in A$ and

$$|q_l - p| = \left| \frac{y_l - p(2^l - 1)}{2^l - 1} \right| < \frac{1}{2^{l-1}}.$$

Since $\frac{1}{2^{l-1}} \to 0$ as $l \to \infty$, we have defined a sequence of elements in A such that $q_l \to p$ as $l \to \infty$. Thus A is dense in the interval $[0,1]$ as claimed. □

In light of Lemma 5.1.4, since $\left\{ \frac{a(M)}{2^l - 1} \;\middle|\; I \neq M \in \langle A, B\rangle \right\}$ is dense in $[0,1]$ it follows that

- $\left\{ \frac{a(M)}{3(2^l-1)} \mid M \text{ is a word on } \{A, B\} \right\}$ is dense in $[0, \frac{1}{3}]$, and
- $\left\{ 1 - \frac{2a(M)}{3(2^l-1)} \mid M \text{ is a word on } \{A, B\} \right\}$ is dense in $[\frac{1}{3}, 1]$.

Since the above are subsets of the closure of $S(\mathrm{Sym}(3), g_2)$ for any $g_2 \in C_2$ and $S(\mathrm{Sym}(3), 1)$ respectively, we have

- $S(\text{Sym}(3), g_2)$ is dense in $[0, \frac{1}{3}]$ for all $g_2 \in C_2$ and
- $S(\text{Sym}(3), 1)$ is dense in $[\frac{1}{3}, 1]$.

Thus we conclude that $S(\text{Sym}(3))$ is dense in $[0, 1]$ and the theorem is proved. □

In fact, the proof has shown more. It has shown that $S(G, 1)$ is dense in the interval $[\frac{1}{3}, 1]$ and if g is a 3-cycle, $S(G, g)$ is dense in $[0, \frac{1}{3}]$. One might ask whether these intervals are maximal, i.e. what is the largest interval T of $[0, 1]$ such that $S(G, 1)$ is dense in T? This remains an open problem.

It is worth noting that the above proof is constructive – it supplies a recipe for creating words whose associated probabilities are as close as desired to any target probability in the interval $[0, 1]$.

It is natural to ask whether the result holds for other symmetric groups. Since Sym(1) and Sym(2) are abelian, their associated sets of probabilities must be finite, so of course the result does not hold for these groups. We shall see at the end of this chapter however that the theorem holds for any symmetric group Sym(n) where $n \geq 3$.

It is also natural to ask about other non-nilpotent dihedral groups. Sym(3) is isomorphic to the dihedral group D_6, the group consisting of the symmetries of a triangle. One thus might wonder whether the theorem holds for other dihedral groups of this form, i.e. is $S(D_{2n})$ dense for any dihedral group of order $2n$ where n is odd? We shall see in the next section that the answer is yes, and that the proof is a straightforward generalisation of that for Sym(3).

5.2 Dihedral groups of order $2m$, where m is odd

In the §5.1 we proved that the set of probabilities associated with Sym(3) is a dense subset of the interval $[0, 1]$. Since Sym(3) is isomorphic to D_6, the dihedral group of order 6, it is natural to ask whether the result may be generalised to other dihedral groups. In this section we prove that the result holds for all dihedral groups of order $2n$, where n is odd. The proof of this is an exact generalisation of that for Sym(3). As such, we omit much of the explanation, which is analogous to that given in Section 5.1.

Theorem 5.2.1. *Let D_{2n} denote the dihedral group of order $2n$. If n is odd, then $S(D_{2n})$ is dense in the interval $[0, 1]$.*

As before, we begin by defining two words w_A and w_B from which we shall build an infinite set of words, by repeated substitution into the first variable. By calculating the transition matrices for A and B, we can then calculate the limiting probabilities associated with repeated substitution of any such word, which we shall prove is a dense

subset of the interval $[0,1]$. We begin by defining a partition of the group as we did for Sym(3).

Partitioning the group

Let n be an odd integer greater than 2. We denote the **dihedral group** of order $2n$ by D_{2n}. This is the group defined by the presentation

$$D_{2n} = \left\langle a, s \mid a^n = 1, s^2 = 1, sas = a^{-1} \right\rangle.$$

As in the corresponding proof for Sym(3), instead of calculating $P(D_{2n}, w = g)$ for every single *element* g of D_{2n}, we first calculate the probability that a uniformly random evaluation lies with certain *subsets* of D_{2n}. For Sym(3) we used the three conjugacy classes. In general our subsets are not conjugacy classes, but are the sets of elements of a certain order, i.e. we partition D_{2n} into the identity element, the non-trivial rotations and the reflections.

$$\begin{aligned} T_1 &:= \{1\}, & |T_1| &= 1, \\ T_2 &:= \left\{a^i, \ \ 1 \leq i \leq n-1\right\}, & |T_2| &= n-1, \\ T_3 &:= \left\{sa^i, \ \ 0 \leq i \leq n-1\right\}, & |T_3| &= n. \end{aligned}$$

Constructing an infinite set of words

Again our proof relies on constructing words out of the two short words w_A and w_B. Let $w_A, w_B : D_{2n} \times D_{2n} \to D_{2n}$ be given by

$$\begin{aligned} w_A(x_1, x_2) &= x_1^2 x_2^2, \\ w_B(x_1, x_2) &= [x_1, x_2]. \end{aligned}$$

Let $I \neq M := M_1 \dots M_l \in \langle A, B \rangle$. Then $w_M = w_{M_1 \dots M_l}$ is given by

$$\begin{aligned} w_{M_1 \dots M_l}(x_1, \dots, x_{l+1}) : &= w_{M_l}(w_{M_1 \dots M_{l-1}}(x_1, \dots, x_l), x_{l+1}) \\ &= w_{M_l}(\dots w_{M_2}(w_{M_1}(x_1, x_2), x_3), \dots, x_{l+1}). \end{aligned}$$

As before we consider the associated probabilities of each of the infinite number of words that can be built from w_A and w_B in this way. We show that this yields a dense subset of $[0,1]$.

5.2.1 Using transition matrices

We define A, B to be the transition matrices with entries

$$a_{ij} := P(D_{2n}, w_A(g_1, g_2) \in T_j \mid g_1 \in T_i),$$
$$b_{ij} := P(D_{2n}, w_B(g_1, g_2) \in T_j \mid g_1 \in T_i).$$

Let v_0 be the row vector whose i^{th} component is the probability that a uniformly random element of D_{2n} lies in T_i. Then v_0 is given by

$$v_0 = \left(\tfrac{|T_1|}{2n}, \tfrac{|T_2|}{2n}, \tfrac{|T_3|}{2n}\right) = \left(\tfrac{1}{2n}, \tfrac{n-1}{2n}, \tfrac{1}{2}\right).$$

If $I \neq M := M_1 \dots M_l \in \langle A, B\rangle$ with $w_{M_1 \dots M_l}$ as defined above, then we define the row vector $v_{M_1 \dots M_l}$ to be that which has i^{th} component $P(D_{2n}, w_{M_1 \dots M_l} \in T_i)$. As in the previous proof we have

$$v_{M_1 \dots M_l} = v_0 M_1 \dots M_l.$$

Thus we may easily calculate the probabilities associated with any word constructed from w_A and w_B in the above manner once A and B are known.

5.2.2 Properties of A and B

Recall that we define A and B to be the matrices whose entries are

$$a_{ij} := P(D_{2n}, w_A(g_1, g_2) \in T_j \mid g_1 \in T_i),$$
$$b_{ij} := P(D_{2n}, w_B(g_1, g_2) \in T_j \mid g_1 \in T_i).$$

We begin by calculating A and B explicitly, in terms of n.

Calculating A

Recall that $w_1(x_1, x_2) := x_1^2 x_2^2$ and that $T_1 := \{1\}, T_2 := \langle a\rangle \setminus \{1\}$ is the set of non-trivial rotations, and T_3 is the set of transpositions.

- Since n is odd, the map $T_2 \to T_2, g \mapsto g^2$ is bijective and $g^2 = 1$ for any g in T_1 or T_3. Thus

$$\begin{aligned} P(D_{2n}, w_A(x_1, x_2) \in T_i \mid x_1 \in T_1) &= P(D_{2n}, w_A(x_1, x_2) \in T_i \mid x_1 \in T_3) \\ &= P(D_{2n},\ x^2 \in T_i), \end{aligned}$$

and

$$P(D_{2n},\ x^2 \in T_i) = \begin{cases} \frac{n+1}{2n} & i = 1, \\ \frac{n-1}{2n} & i = 2, \\ 0 & i = 3. \end{cases}$$

Thus $(a_{11}, a_{12}, a_{13}) = (a_{31}, a_{32}, a_{31}) = \left(\frac{n+1}{2n}, \frac{n-1}{2n}, 0\right)$.

- Now if g_1 is a non-trivial rotation, then $g_1^2 g_2^2 = 1$ if and only if $g_2 = g_1^{-1}$. Otherwise $g_1^2 g_2^2$ is a rotation. Therefore

$$P(D_{2n}, w_A(x_1, x_2) \in T_j \mid x_1 \in T_1) = \begin{cases} \frac{1}{2n} & i = 1, \\ \frac{2n-1}{2n} & i = 2, \\ 0 & i = 3. \end{cases}$$

Hence $(a_{21}, a_{22}, a_{23}) = \left(\frac{1}{2n}, \frac{2n-1}{2n}, 0\right)$.

Thus

$$A = \begin{pmatrix} \frac{n+1}{2n} & \frac{n-1}{2n} & 0 \\ \frac{1}{2n} & \frac{2n-1}{2n} & 0 \\ \frac{n+1}{2n} & \frac{n-1}{2n} & 0 \end{pmatrix}.$$

Calculating B

Recall that $w_B(x_1, x_2) := [x_1, x_2]$. We use that for $0 \leq r, s \leq 1, \quad 0 \leq i, j \leq n-1$, $\left[x^r a^i, x^s a^j\right] = a^{2(rj-si)}$.

- If $g_1 = 1$, then $[g_1, g_2]$ is trivial for any g_2. Therefore

$$P(D_{2n}, w_B \in T_i \mid g_1 \in T_1) = \begin{cases} 1 & i = 1, \\ 0 & i = 2, \\ 0 & i = 3. \end{cases}$$

Hence $(b_{11}, b_{12}, b_{13}) = (1, 0, 0)$.

- If g_1 is a non-trivial rotation, then g_1 commutes with g_2 exactly when g_2 is in $\langle g_1 \rangle$, i.e. when g_2 is a member of T_1 or T_2. Otherwise $[g_1, g_2]$ is a member of T_2, (since ${D_{2n}}' = \langle a \rangle$). Thus

$$P(D_{2n}, w_B \in T_i \mid g_1 \in T_2) = \begin{cases} \frac{1}{2} & i = 1, \\ \frac{1}{2} & i = 2, \\ 0 & i = 3. \end{cases}$$

Therefore $(b_{12}, b_{22}, b_{23}) = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$.

- If g_1 is a reflection, then it commutes only with itself and the trivial element. Otherwise $[g_1, g_2]$ is in T_2. Therefore

$$P(w_B \in T_i \mid g_1 \in T_3) = \begin{cases} \frac{1}{n} & i = 1, \\ \frac{n-1}{n} & i = 2, \\ 0 & i = 3. \end{cases}$$

  Therefore $(b_{31}, b_{32}, b_{33}) = (\frac{1}{n}, \frac{n-1}{n}, 0)$.

Thus

$$B = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{n} & \frac{n-1}{n} & 0 \end{pmatrix}.$$

Eigenvalues and eigenvectors

It is easily verified that A and B have the same eigenvalues, $0, \frac{1}{2}$ and 1, and that the corresponding eigenvectors are as follows.

- The left eigenvectors of A corresponding to $0, \frac{1}{2}, 1$ respectively are $(1, 0, -1), (1, -1, 0), (1, n-1, 0)$.
- The right eigenvectors of A corresponding to $0, \frac{1}{2}, 1$ respectively are $(0, 0, 1)^t, (1-n, 1, 1-n)^t, (1, 1, 1)^t$.
- The left eigenvectors of B corresponding to $0, \frac{1}{2}, 1$ respectively are $(n-2, 2(1-n), n), (1, -1, 0), (1, 0, 0)$.
- The right eigenvectors of B corresponding to $0, \frac{1}{2}, 1$ respectively are $(0, 0, 1)^t, (0, n, 2(n-1))^t, (1, 1, 1)^t$.

5.2.3 Products of A and B

Let $I \neq M \in \langle A, B \rangle$ have length $l \geq 1$. We begin by investigating the eigenvalues and eigenvectors of M.

Eigenvalues of M

Referring to §5.2.2 we see that A and B both have right eigenvectors $v_1 := (0, 0, 1)^t$ and $v_3 := (1, 1, 1)^t$ corresponding to eigenvalues $0, 1$ respectively. Thus it follows that $Mv_1 = 0 = 0v_1$ and $Mv_3 = v_3$, i.e. 0 and 1 are also eigenvalues of M.

In the same section we also find that A and B both have left eigenvector $w_2 := (1, -1, 0)$ corresponding to the eigenvalue $\frac{1}{2}$. Thus we see that

$$w_2 M \;=\; w_2 \prod_{i=1}^{l} M_i \;=\; (w_2 M_1) \prod_{i=2}^{l} M_i \;=\; \tfrac{1}{2} w_2 \prod_{i=2}^{l} M_i \;= \ldots = \; \left(\tfrac{1}{2}\right)^l w_2.$$

Hence we conclude that M has three distinct eigenvalues, $0, (\frac{1}{2})^l$ and 1. Again, since M has three distinct eigenvalues, M is diagonalizable. Write $M = PDP^{-1}$. Since $\lim_{i\to\infty} \lambda^i$ exists for all eigenvalues λ, it follows that the limit $D^\infty := \lim_{i\to\infty} D^i$ exists and hence $\lim_{i\to\infty} M^i = PD^\infty P^{-1}$ exists.

Eigenvectors of M

We are concerned with finding the *left* normalised eigenvector of M corresponding to eigenvalue 1, i.e. the unique vector w that satisfies $wM = w$ and whose components sum to one. Since the last column of M will necessarily be a column of zeros, (since this is true of A and B) it follows that the last component of w will be zero. Thus w is in the span of $V := \{(1, 0, 0), (0, 1, 0)\}$, and we restrict our search to this subspace. We choose the normalised eigenvectors of A and B that correspond to 1, $\{u_1 := (\frac{1}{n}, \frac{n-1}{n}, 0), u_2 := (1, 0, 0)\}$ as a convenient basis for V.

With respect to this basis, the linear maps represented by A and B are represented by the matrices

$$P = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}.$$

This follows since $u_1 B = u_2 A = \frac{1}{2} u_1 + \frac{1}{2} u_2$.

Suppose $w = \alpha u_1 + \beta u_2$. Again, note that since w is normalised, we have $\beta = 1 - \alpha$. Let $N = \prod_{i=1}^{l} N_i \in \langle P, Q \rangle$ be the matrix corresponding to M with respect to the new basis (i.e. the matrix obtained from $\prod M_i$ by replacing every A with P and B with Q). As in the proof for Sym(3), finding this eigenvector boils down to solving

$$(\alpha, 1 - \alpha) N = (\alpha, 1 - \alpha). \tag{5.8}$$

By Lemma 5.1.3 we have that

$$(\alpha, 1 - \alpha)\, N = \left(\tfrac{\alpha + a}{2^l}, 1 - \tfrac{\alpha + a}{2^l}\right)$$

where $a = a_0 2^0 + a_1 2^1 + \ldots a_{l-1} 2^{l-1}$. Thus (5.8) is equivalent to

$$\left(\tfrac{\alpha + a}{2^l}, 1 - \tfrac{\alpha + a}{2^l}\right) = (\alpha, 1 - \alpha)$$

which has solution $\alpha = \frac{a}{2^l-1}$.

Remark Note that this again implies that for any $l \in \mathbb{N}$, and $a \in \{0, \ldots, 2^l - 1\}$, there exists a word N on P, Q such that $(\alpha, 1-\alpha)\, N = \left(\frac{\alpha+a}{2^l}, 1 - \frac{\alpha+a}{2^l}\right)$.

Substituting $\alpha = \frac{a}{2^l-1}$ and $\beta = 1 - \alpha$ into $w = \alpha u_1 + \beta u_2$ we find that our eigenvector associated with eigenvalue 1 is

$$w = \left(1 - \tfrac{a}{2^l-1}\left(\tfrac{n-1}{n}\right), \tfrac{a}{2^l-1}\left(\tfrac{n-1}{n}\right), 0\right).$$

5.2.4 The proof of Theorem 5.2.1

Having done much of the work, we now begin the proof of Theorem 5.2.1, using w_A, w_B, A and B as above.

Proof. Let $I \neq M \in \langle A, B\rangle$ have length $l \geq 1$. Let w_{M^i} be the word constructed from w_A and w_B in the order dictated by M^i and let v_{M^i} be the row vector whose j^{th} component is $P(D_{2n}, w_{M^i} \in T_j)$. If $v_0 := \left(\frac{1}{2n}, \frac{n-1}{n}, \frac{1}{2}\right)$, then as we saw in §5.1.3

$$v_{M^i} = v_0 M^i$$

and

$$v_{M^\infty} := \lim_{i\to\infty} v_{M^i} = \lim_{i\to\infty} \left(P(D_{2n}, w_{M^i} \in T_j)\right)_{j=1}^{3} = \lim_{i\to\infty} v_0 M^i.$$

We know that this limit exists, since we saw in §5.2.3 that M is diagonalisable, with eigenvalues $0, (\frac{1}{2})^l$ and 1. As before

$$v_{M^\infty} M = \left(\lim_{i\to\infty} v_0 M^i\right) M = \lim_{i\to\infty} v_0 M^i = v_{M^\infty}.$$

Thus v_{M^∞} is a normalised eigenvector of M corresponding to eigenvalue 1. Since there is only one such vector, we have that

$$v_{M^\infty} = \left(1 - \tfrac{a}{2^l-1}(\tfrac{n-1}{n}), \tfrac{a}{2^l-1}(\tfrac{n-1}{n}), 0\right)$$

where $a = a(M)$ is as defined above.

In order to say something about $P(D_{2n}, w = g)$ for individual $g \in T_2$ we require the following lemma.

Lemma 5.2.2. *Let w_M be a word constructed from w_A and w_B via the method described in Section 5.2. Then for any $g, h \in T_2$*

$$P(D_{2n}, w_M = g) = P(D_{2n}, w_M = h). \tag{5.9}$$

Proof. We proceed by induction. Suppose that w_M is a word that satisfies (5.9). We shall show that for all g, h in T_2,

(i) $P(D_{2n}, w_A(w_M, x) = g) = P(D_{2n}, w_A(w_M, x) = h)$,

(ii) $P(D_{2n}, w_B(w_M, x) = g) = P(D_{2n}, w_B(w_M, x) = h)$.

Here are the proofs.

(i) First note the following.

(a) Since $\phi : x \mapsto x^2$ is a bijection on T_2, we must have $P(D_{2n}, x^2 = g) = P(D_{2n}, x^2 = h)$ for all g, h in T_2.

(b) For all g, h in T_2, $(w_M)^2 = g \Leftrightarrow w_M = \phi^{-1}(g)$. Note that $\phi^{-1}(g) \in T_2$. Property (5.9) applied to w_M tells us that

$$P(D_{2n}, w_M = \phi^{-1}(g)) = P(D_{2n}, w_M = \phi^{-1}(h))$$

$$\therefore \quad P(D_{2n}, (w_M)^2 = g) = P(D_{2n}, (w_M)^2 = h) \quad \forall g, h \in T_2.$$

Now let $g_1, g_2 \in T_2$. Then

$$\begin{aligned}
P(D_{2n}, w_A(w_M, x) = g_1) &= P(D_{2n}, (w_M)^2 x^2 = g_1) \\
&= P(D_{2n}, (w_M)^2 = 1) P(D_{2n}, x^2 = g_1) \\
&\quad + P(D_{2n}, (w_M)^2 = g_1) P(D_{2n}, x^2 = 1) \\
&\quad + \sum_{h \in T_2 \setminus \{g_1\}} P(D_{2n}, (w_M)^2 = h) P(D_{2n}, x^2 = h^{-1} g_1) \\
&= P(D_{2n}, (w_M)^2 = 1) P(D_{2n}, x^2 = g_2) \quad (\text{ by } (a)) \\
&\quad + P(D_{2n}, (w_M)^2 = g_2) P(D_{2n}, x^2 = 1) \quad (\text{ by } (b)) \\
&\quad + \sum_{h \in T_2 \setminus \{g_2\}} P(D_{2n}, (w_M)^2 = h) P(D_{2n}, x^2 = h^{-1} g_2) \\
&\qquad (\text{ by } (a), (b) \text{ and since } \forall h, h^{-1} g_1 \in T_2) \\
&= P(D_{2n}, w_A(w_M, x) = g_2).
\end{aligned}$$

(ii) First note that

$$\left[s^r a^i, s^t a^j\right] = a^{2(rj - ti)}.$$

Write $g_1 = a^{2b_1}, g_2 = a^{2b_2}$ where $b_1, b_2 \neq 0$. Note that $(rj - ti) = b$ if and only if

$$(r, i, s, j) \in \{(1, i, 1, b + i), (1, i, 0, b), (0, b, 1, j) \mid i, j \in \{0, \ldots, n - 1\}\}.$$

Note also that since any two elements g, h in T_3 are autoequivalent, (i.e. there

exists an automorphism $\phi : D_{2n} \to D_{2n}$ such that $\phi(g) = h$) it follows that $P(D_{2n}, w = g) = P(D_{2n}, w = h)$ for any word w and any $g, h \in T_3$.

Thus

$$\begin{aligned} P(D_{2n}, w_B(w_M, x) = g_1) &= P(D_{2n}, [w_M, s^t a^j] = a^{2b_1}) \\ &= \sum_{i=0}^{n} P(D_{2n}, w_M = sa^i) P(D_{2n}, x \in \{sa^{b_1+i}, a^{b_1}\}) \\ &\quad + P(D_{2n}, w_M = a^{b_1}) P(D_{2n}, x \in T_3) \\ &= \sum_{i=0}^{n} P(D_{2n}, w_M = sa^i) P(D_{2n}, x \in \{sa^{b_2+i}, a^{b_2}\}) \\ &\quad + P(D_{2n}, w_M = a^{b_2}) P(D_{2n}, x \in T_3) \\ &= P(D_{2n}, w_B(w_M, x) = g_2). \end{aligned}$$

Thus (i) and (ii) are proved.

It remains to note that the word $w(x) = x$ satisfies

$$P(D_{2n}, x = g) = P(D_{2n}, x = h)$$

for all g, h in T_i. Thus by induction any word constructed by the method of Section 5.2 satisfies property (5.9) and the lemma is proved. □

By the above lemma, we have that for any $g_2 \in T_2$

$$P(D_{2n}, w_{M^i} = g_2) = \tfrac{1}{n-1} P(D_{2n}, w_{M^i} \in T_2).$$

Thus

$$\lim_{i \to \infty} P(D_{2n}, w_{M^i} = g_2) = \tfrac{1}{n-1} \tfrac{a}{2^l - 1} \tfrac{n-1}{n} = \tfrac{a}{n(2^l-1)}, \tag{5.10}$$

$$\lim_{i \to \infty} P(D_{2n}, w_{M^i} = 1) = 1 - \tfrac{a(n-1)}{n(2^l-1)}. \tag{5.11}$$

Now (5.10) and (5.11) correspond to limit points of $S(D_{2n})$, so it remains to show that such points are dense in the interval $[0, 1]$. We have seen in §5.1.4 that

$$\left\{ \frac{a(M)}{2^{l(M)} - 1} \mid I \neq M \in \langle A, B \rangle \right\}$$

is dense in $[0, 1]$. It therefore follows that

- $\left\{ \frac{a(M)}{n\left(2^{l(M)}-1\right)} \mid M \text{ is a word on } \{A, B\} \right\}$ is dense in $[0, \frac{1}{n}]$, and

- $\left\{1 - \frac{a(M)(n-1)}{n\left(2^{l(M)}-1\right)} \mid M \text{ is a word on } \{A, B\}\right\}$ is dense in $[\frac{1}{n}, 1]$.

Since the above are subsets of the closure of $S(D_{2n}, g_2)$ and $S(D_{2n}, 1)$ respectively, we have

- $S(D_{2n}, g_2)$ is dense in $[0, \frac{1}{n}]$ for any $g_2 \in T_2$ and
- $S(D_{2n}, 1)$ is dense in $[\frac{1}{n}, 1]$.

Thus we conclude that $S(D_{2n})$ is dense in $[0, 1]$ and the theorem is proved. □

We get the following result for free.

Corollary 5.2.3. *Let $n \in \mathbb{N}$ be odd and larger than 2. Then $S(D_{4n})$ is dense in $[0, 1]$.*

Proof. It is known that if n is odd and larger than 2, then

$$D_{4n} \cong D_{2n} \times \mathbb{Z}_2.$$

Since w_A and w_B (and thus any w_M constructed from them as above) are trivial over $\mathbb{Z}_2$, and by (2.13)

$$P(G \times H, w = (g, h)) = P(G, w = g)P(G, w = h).$$

It therefore follows that for any $g \in D_{2n}$,

$$P(D_{2n} \times \mathbb{Z}_2, w_M = (g, 1)) = P(D_{2n}, w_M = g).$$

Thus the result follows from the proof Theorem 5.3. □

Once again the question of whether the intervals $[0, \frac{1}{n}]$ and $[\frac{1}{n}, 1]$ are maximal is an open question.

5.3 Non-nilpotent dihedral groups

In the last section we proved the following theorem.

Theorem. *Let D_{2n} denote the dihedral group of order $2n$. If $n > 1$ is odd, then $S(D_{2n})$ is dense in the interval $[0, 1]$.*

We proved this by constructing an infinite set of words from $w_A(x_1, x_2) = x_1^2 x_2^2$ and $w_B(x_1, x_2) = [x_1, x_2]$ by repeated substitution, and proved that the set of probabilities associated with such words is dense in $[0, 1]$.

Here we extend this result to all non-nilpotent finite dihedral groups, i.e. those whose order is not a power of two.

Theorem 5.3.1. *Let G be a non-nilpotent dihedral group. Then the set $S(G)$ is dense in $[0,1]$.*

Proof. Let $i \in \mathbb{N}$ and m be odd. Let $D_{2^i m}$ be given by the following presentation, with generators α and β.

$$D_{2^i m} = \left\langle \alpha, \beta \mid \alpha^{2^{i-1}m} = 1, \beta^2 = 1, \beta\alpha\beta = \alpha^{-1} \right\rangle.$$

Let D_{2m} be given by the following presentation, with generators a and b.

$$D_{2m} = \left\langle a, b \mid a^m = 1, b^2 = 1, bab = a^{-1} \right\rangle.$$

In place of w_A and w_B in the proof for the usual dihedral groups, we use the following words.

$$w_{A,i}(x_1, x_2) := x_1^{2^i} x_2^{2^i},$$
$$w_{B,i}(x_1, x_2) := [x_1, x_2]^{2^{i-1}}.$$

In the last section, we considered words built from $w_A = w_{A,1}$ and $w_B = w_{B,1}$. Let $I \neq M \in \langle A, B\rangle$ and let w_M be the associated word, as described in §5.1.2. For any positive integer i, we define $w_{M,i}$ to be the word built in the same way as w_M, but with each w_A replaced with $w_{A,i}$ and each w_B replaced by $w_{B,i}$. We have seen that if

$$W := \{w_M \mid I \neq M \in \langle A, B\rangle\},$$

then the set of probabilities associated with words in W with respect to D_{2m} form a dense subset of the interval $[0,1]$. We proceed by showing that

$$\{P(D_{2m}, w_M = g) \mid I \neq M \in \langle A, B\rangle\}$$

is also contained in $S(D_{2^i m})$, and thus the result follows from the proof of Theorem 5.2.1.

Recall from Lemma 5.2.2 that for any $j_1, j_2 \in \{1, \ldots, m-1\}$

$$P(D_{2m}, w_M = a^{j_1}) = P(D_{2m}, w_M = a^{j_2}).$$

Let $\phi : \langle a\rangle \to \langle a\rangle$ be given by $x \mapsto x^{2^{i-1}}$. Since $\langle a\rangle \cong \mathbb{Z}_m$ and m is odd, ϕ is a bijection and $\phi(1) = 1$. Thus since $w_{A,i} = \phi \circ w_A$ and $w_{B,i} = \phi \circ w_B$ we have

$$P(D_{2m}, w_A = a^j) = P(D_{2m}, \phi \circ w_A = a^j) = P(D_{2m}, w_{A,i} = a^j), \tag{5.12}$$
$$P(D_{2m}, w_B = a^j) = P(D_{2m}, \phi \circ w_B = a^j) = P(D_{2m}, w_{B,i} = a^j). \tag{5.13}$$

Indeed if $I \neq M \in \langle A, B \rangle$ then it follows that the word $w_{M,i}$ as defined above satisfies

$$P(D_{2m}, w_M = a^j) = P(D_{2m}, w_{M,i} = a^j).$$

Consider now $w_{M,i}(D_{2^i m})$. Since $2^i m$ is even, it is known that the derived subgroup of $D_{2^i m}$ is $\langle \alpha^2 \rangle$. Thus $w_{B,i}(D_{2^i m}) = \left\langle \alpha^{2^i} \right\rangle \cong \mathbb{Z}_m$. Since the image of $x \mapsto x^{2^i}$ over $D_{2^i m}$ is $\left\langle \alpha^{2^i} \right\rangle$ it follows that $w_{A,i}(D_{2^i m}) = \left\langle \alpha^{2^i} \right\rangle$. It follows that $w_{M,i}(D_{2^i m}) \leq \left\langle \alpha^{2^i} \right\rangle$ for any $I \neq M \in \langle A, B \rangle$.

From (2.14) we know that for an arbitrary group G with normal subgroup N,

$$P(G/N, w = gN) = \sum_{n \in N} P(G, w = gn).$$

We let $G = D_{2^i m}$ and $N := \langle \alpha^m \rangle$ in the above. Let a^j be in $w_M(D_{2m})$, the image of w_M over D_{2m}. Then for any $0 \leq j \leq m-1$,

$$\begin{aligned} P(D_{2m}, w_M = a^j) &= P(D_{2m}, w_{M,i} = a^j) && (5.14)\\ &= P(D_{2^i m}/N, w_{M,i} = \alpha^j N) && (5.15)\\ &= \sum_{a^{lm} \in \langle \alpha^m \rangle} P(D_{2^i m}, w_{M,i} = \alpha^j \alpha^{lm}) && (5.16)\\ &= \sum_{l=0}^{2^{(i-1)}-1} P(D_{2^i m}, w_{M,i} = \alpha^{(j+lm)}). && (5.17) \end{aligned}$$

Since $w_{M,i}(D_{2^i m}) = \left\langle \alpha^{2^i} \right\rangle$, $P(D_{2^i m}, w_{M,i} = \alpha^c) = 0$ if c is not a multiple of 2^i.

Claim 5.3.2. *There is at most one value of l in $\{0, \ldots, 2^i - 1\}$ such that $j + lm$ is a multiple of 2^i.*

Proof. Suppose l_1 and l_2 satisfy the conditions stated in the claim, i.e. there exist $z_1, z_2 \in \mathbb{Z}$ such that

$$j + l_1 = 2^i z_1, \qquad j + l_2 = 2^i z_2.$$

Then $m(l_1 - l_2) = 2^i(z_1 - z_2)$. Since m is odd, 2^i must divide $l_1 - l_2$. Since $0 \leq |l_1 - l_2| < 2^i$ it follows that $l_1 - l_2 = 0$, and so $l_1 = l_2$ and the claim is proved. □

We know there must be at least one value of l that provides a non-zero probability, since the left hand side of (5.15) is non-zero. Denote this by l^*. Thus $j + l^* m = 2^i z$ for some $z \in \mathbb{Z}$. Hence

$$P(D_{2m}, w_M = a^j) = P(D_{2^{i+1} m}, w_{M,i} = \alpha^{j + l^* m}).$$

Thus the set of probabilities associated with words in W and the group D_{2m} is also

contained in $S(D_{2^i m})$. Since such probabilities form a dense subset of the interval $[0,1]$ it follows that $S(D_{2^{i+1}m})$ is dense in $[0,1]$, and the theorem is proved. □

5.4 Generalised dihedral groups

We now go one step further and show that the result holds for generalised dihedral groups.

Let G be a generalised dihedral group. By this we mean that $G = H \rtimes K$ is the semi-direct product of some abelian subgroup H by some subgroup $K \cong \mathbb{Z}_2$, where the non-trivial element of K acts on H by inversion.

If $|H| = 2^n$ for some positive integer n, then G is a 2-group, and must therefore be nilpotent. If H is not a 2-group we claim that $S(G)$ is dense in $[0,1]$. Let $|H| = 2^i m$ where $m > 1$ is odd. Write H as a direct product $H = H_1 \times H_2$ where $|H_1| = 2^i$ and $|H_2| = m$. We prove that this group has a dense set of associated probabilities by proving three claims.

Claim 5.4.1. *Let $I \neq M \in \langle A, B\rangle$. Then $w_{M,i}(G) = H_2$.*

Claim 5.4.2. *If $1 \neq h_1, h_2 \in H_2$ then $P(G, w_{M,i} = h_1) = P(G, w_{M,i} = h_2)$.*

Claim 5.4.3. *Let $D_{2^{i+1}m} = \widetilde{H} \rtimes \mathbb{Z}_2$, so $\widetilde{H}$ is cyclic of order $2^i m$. Write $\widetilde{H} = \widetilde{H}_1 \times \widetilde{H}_2$ where $\widetilde{H}_1 \cong \mathbb{Z}_{2^i}$ and $\widetilde{H}_2 \cong \mathbb{Z}_m$. Let $I \neq M \in \langle A, B\rangle$. Then for all $1 \neq h \in H_2, 1 \neq \tilde{h} \in \widetilde{H}_2$,*

$$P(G, w_{M,i} = 1) = P(D_{2^{i+1}m}, w_{M,i} = 1), \tag{5.18}$$

$$P(G, w_{M,i} = h) = P(D_{2^{i+1}m}, w_{M,i} = \tilde{h}). \tag{5.19}$$

Since we have already seen that the set of words

$$W := \{w_{M,i} \mid I \neq M \in \langle A, B\rangle\}$$

yields a set of probabilities that are dense in $[0,1]$, it will follow from the claims that $S(G)$ is dense in $[0,1]$ and thus we shall have the following result.

Theorem 5.4.4. *Let G be a non-nilpotent generalised dihedral group. Then $S(G)$ is dense in $[0,1]$.*

We now prove the claims.

Proof of Claim 5.4.1. Recall that

$$w_{A,i}(x_1, x_2) := {x_1}^{2^i} {x_2}^{2^i},$$
$$w_{B,i}(x_1, x_2) := [x_1, x_2]^{2^{i-1}}.$$

Let b denote the non-trivial element of the subgroup $K \cong \mathbb{Z}_2$. Then $G = H \cup bH$, and this is a disjoint union. Note that for any h in H, $(bh)^2 = bhbh = h^b h = h^{-1}h = 1$. Since $|H_1| = 2^i$, $x \mapsto x^{2^i}$ is trivial over H_1. Since $|H_2|$ is odd, $x \mapsto x^{2^i}$ is a bijection over H_2. Thus

$$\begin{aligned} w_{A,i}(G,G) &= H_2H_2 = H_2 \qquad \text{and} \\ w_{A,i}(H_2,G) &= H_2H_2 = H_2. \end{aligned}$$

Note that for any $s_1, s_2 \in \{0,1\}, h_1, h_2 \in H$ we have

$$[b^{s_1}h_1, b^{s_2}h_2] = [b^{s_1}, b^{s_2}h_2]^{h_1}[h_1, b^{s_2}h_2] = [b^{s_1}, h_2]^{h_1}[h_1, b^{s_2}]^{h_2} \tag{5.20}$$

$$= \left((h_2^{-1})^{a^{s_1}}h_2\right)^{h_1}\left({h_1}^{-1}(h_1)^{a^{s_2}}\right)^{h_2} = {h_2}^{2^{s_1}}{h_1}^{-2s_2}. \tag{5.21}$$

Thus if G^i denotes $\left\langle g^i \mid g \in G \right\rangle$ where i is a positive integer,

$$\begin{aligned} w_{B,i}(G,G) &= [G,G]^{2^i} = G^{2^i}G^{2^i} = H_2 \qquad \text{and} \\ w_{B,i}(H_2,G) &= [H_2,G]^{2^i} = H^{2^{i+1}} = H_2. \end{aligned}$$

The latter equation follows since, taking $s_1 = 0$, in (5.20) we obtain

$$[h_1, a^{s_2}h_2]^{2^i} = ({h_2}^0{h_1}^{-2s_2})^{2^i} = ({h_1}^{-2s_2})^{2^i} \in H^{2^{i+1}} = H_2.$$

Thus inductively $w_{M,i}(G) = H_2$ and claim 5.4.1 is proved. □

Proof of Claim 5.4.2. Since $x \mapsto x^{2^i}$ is trivial over aH and H_1, and is a bijection over H_2, it follows that

$$P(G, w_{A,i} = h_1) = P(G, w_{A,i} = h_2) \quad \forall\, 1 \neq h_1, h_2 \in H$$

and similarly if w_M satisfies the claim, then if $h_1, h_2 \in H_2 \backslash \{1\}$

$$\begin{aligned} P(G, w_{A,i}(w_M, \cdot) = h_1) &= \sum_{h \in H_1} P(G, w_M = h)P(G, x^{2^i} = h^{-2^i}h_1) \\ &= \sum_{h \in H_1} P(G, w_M = h)P(G, x^{2^i} = h^{-2^i}). \end{aligned}$$

As above, it can easily be shown that over G, $[b^{s_1}h_1, b^{s_2}h_2]^{2^{i-1}} = {h_2}^{2^i s_1}{h_1}^{-2^i s_2}$, for any h_1, h_2 in H_2 and s_i in $\{0,1\}$. Again, since $x \mapsto x^{2^i}$ is a bijection over H_2 and trivial elsewhere it follows that

$$P(G, w_{B,i} = k_1) = P(G, w_{B,i} = k_2)$$

for all $1 \neq k_1, k_2 \in H_2$ (to see this, note that for any fixed s_i, $({h_2}^{s_1})^{2^i}({h_1}^{s_2})^{2^i}$ will have equally sized fibres for any non-trivial element of H_2).

Now consider $w_{B,i}(w_{M,i}, \cdot)$ for some M where $w_{M,i}$ satisfies the claim. By claim 5.4.1, $w_{M,i}(G) = H_2$. Note that for a fixed $h \in H_2$,

$$w_{B,i}(h, b^{s_2}h_2) = (h^{-s_2})^{2^i}.$$

If $s_2 = 0$ this is trivial. If $s_2 = 1$ then since $x \mapsto x^{-2^i}$ is a bijection over H_2 and we have assumed that occurrences of $w_{M,i} = h$ are equally likely for all h in H_2 it follows that

$$P(G, w_{B,i}(w_{M,i}, x_k) = h_1) = P(G, w_{B,i}(w_{M,i}, x_k) = h_2) \quad \forall h_1, h_2 \in H$$

and the claim holds. □

Proof of Claim 5.4.3. By (5.3) it follows that

$$P(D_{2^i m}, w_{M,i} = h_1) = P(D_{2^i m}, w_{M,i} = h_2) \quad \forall\, h_1, h_2 \in H_2 \backslash \{1\}.$$

Thus for any $h \in H_2$

$$P(D_{2^i m}, w_{M,i} = h) = \frac{1}{m-1}\left(1 - P(D_{2^i m}, w_{M,i} = 1)\right).$$

By claims 5.4.1 and 5.4.2, we also have

$$P(G, w_{M,i} = h) = \frac{1}{m-1}\left(1 - P(G, w_{M,i} = 1)\right).$$

Thus to prove this claim, we need only prove (5.18) and (5.19) will follow.

Since $x \mapsto x^{2^i}$ is trivial over aH and H_1 and a bijection over H_2, (5.18) and hence (5.19) holds for $w_{A,i}$.

To prove the claim we show that if G is a generalised dihedral group (that is not a 2-group), then $P(G, w_{M,i} = 1)$ is dependent only on the order of G, i.e. given generalised dihedral groups $G, \widetilde{G}$ both of order $2^{i+1}m$, then

$$P(G, w_{M,i} = 1) = P(\widetilde{G}, w_{M,i} = 1) \tag{5.22}$$

and from this the claim immediately follows from §5.3. As above let $G = (H_1 \times H_2) \rtimes \mathbb{Z}_2$ and $\widetilde{G} = (\widetilde{H_1} \times \widetilde{H_2}) \rtimes \mathbb{Z}_2$.

Consider $w_{A,i}(x_1, x_2) = {x_1}^{2^i}{x_2}^{2^i}$. Since $H_2 \to H_2, x \mapsto x^{2^i}$ is a bijection and g^{2^i} is the

identity element for any g in $H_1, \widetilde{H_1}$ or $\mathbb{Z}_2$, it follows that

$$P(G, w_{A,i} = 1) = P(\widetilde{G}, w_{A,i} = 1).$$

Suppose that $w_{M,i}$ satisfies (5.22) (and thus (5.18) and (5.19)). Then $w_{A,i}(w_{M,i}, x_k) = (w_{M,i})^{2^i} {x_k}^{2^i}$ must satisfy (5.22).

Consider $w_{B,i}(x_1, x_2) = [x_1, x_2]^{2^{i-1}}$. As above, for $s_1, s_2 \in \{0, 1\}$, $h_1, h_2 \in H$, $1 \neq b \in \mathbb{Z}_2$

$$[b^{s_1} h_1, b^{s_2} h_2]^{2^{i-1}} = ({h_2}^{2s_1} {h_1}^{-2s_1})^{2^{i-1}} = ({h_2}^{s_1})^{2^i} ({h_1}^{-s_2})^{2^i}.$$

Fix $s_1, s_2 \in \{0, 1\}$. Then $w_{s_1,s_2}(h_1, h_2) = {h_2}^{2^i s_1} {h_1}^{2^i s_2}$ is a word over H with image H_2 or the trivial subgroup. Since H is abelian, it follows that $P(H, w_{s_1 s_2} = 1) = |H_2|^{-1}$ in the first case, or $P(H, w_{s_1 s_2} = 1) = 1$ in the latter. In either case, the probability depends only on $|H_2|$ and (5.22) follows.

Let $w_{M,i}$ satisfy (5.22) and consider $w_{B,i}(w_{M,i}, x_k) = [w_{M,i}, x_k]^{2^{i-1}}$. First note that $w_{M,i}(G) = H_2$ and $[h, b^{s_2} h_2]^{2^{i-1}} = ({h_2}^0)^{2^i} (h^{-s_2})^{2^i} = h^{-2^i s_1}$. If $s_1 = 0$ this is trivial; if $s_1 = 1$ this is h^{-2^i}. Since we have assumed that $w_{M,i}$ satisfies the claim and $x \mapsto x^{-2^i}$ is a bijection over H_2, it follows that the same follows for $w_{B,i}(w_{M,i}, x_k)$.

Thus by induction (5.22) follows for all $w_{M,i}$ and claim 5.4.3 is proved. □

We saw in the proof of Lemma 5.3.1 that if $m > 1$ is odd then

$$\{P(D_{2^{i+1}m}, w_{M,i} = g) \mid g \in D_{2^{i+1}m}, I \neq M \in \langle A, B \rangle\}$$

is dense in $[0, 1]$.

We have just seen that if G is a generalised dihedral group of order $2^{i+1}m$ then for all $I \neq M \in \langle A, B \rangle$

$$\{P(G, w_{M,i} = g) \mid g \in G\} = \{P(D_{2^{i+1}m}, w_{M,i} = g) \mid g \in D_{2^{i+1}m}\}.$$

Thus it follows that if G is a generalised dihedral group that is not a 2-group then

$$\{P(G, w_{M,i} = g) \mid I \neq M \in \langle A, B \rangle, g \in G\}$$

is dense in $[0, 1]$. We conclude that $S(G)$ is dense in $[0, 1]$ and the theorem is proved.

A consequence of Theorem 4.1.1 is that if G is nilpotent, then $S(G)$ cannot be dense in $[0, 1]$, since the associated probabilities are bounded below by some positive value. Since we have shown that any generalised dihedral group that is not a 2-group has

probabilities that are dense in $[0,1]$, we have provided an extremely long winded proof that a generalised dihedral group is nilpotent if, and only if, it is a 2-group!

Having settled the case for generalised dihedral groups, we now consider Alt(4), the alternating group on four letters.

5.5 The alternating group on four elements

Theorem 5.5.1. *Let Alt*(4) *denote the alternating group on four letters. Then the set of probabilities associated with Alt*(4)*, is dense in the interval* $[0,1]$.

Proof. The proof is analogous to that for Sym(3), as seen in 5.1.

Alt(4) has one proper verbal subgroup, which we shall denote by H, which consists of the three double transpositions along with the identity element. H is the derived subgroup of Alt(4), and is isomorphic to the Klein Four Group.

We define the words w_A and w_B by $w_A(x_1,x_2) := x_1^3x_2^3$ and $w_B(x_1,x_2) := [x_1,x_2]$. It may be shown that $w_A(\mathrm{Alt}(4)) = w_B(\mathrm{Alt}(4)) = H$. As in the proof for Sym(3) we proceed by considering those words constructed from w_A and w_B by repeated substitution into the first variable. We then show that the set of probabilities associated with these words forms a dense subset of $[0,1]$. As before, we calculate the probabilities via transition matrices, starting with those for w_A and w_B.

Let $T_1 := \{1\}$ and $T_2 := H \setminus \{1\}$. Define the transition matrices $A = (a_{ij})$, and $B = (b_{ij})$ by

$$a_{ij} := P(\mathrm{Alt}(4), w_A \in T_j \mid x_1 \in T_i),$$
$$b_{ij} := P(\mathrm{Alt}(4), w_B \in T_j \mid x_1 \in T_i).$$

Then explicitly, A and B are given by

$$A = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{12} & \frac{11}{12} \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}.$$

It is easily checked that the normalised eigenvectors corresponding to eigenvalue 1 for A and B are $u_1 := \left(\frac{1}{4}, \frac{3}{4}\right)$ and $u_2 := (1,0)$ respectively.

Let M be some product of the matrices A and B, i.e. $M = \prod_{i=1}^{l} N_i$ where $N_i \in \{A, B\}$ for all i. We denote the set of such matrices $\langle A, B \rangle$. We wish to find the normalised eigenvector of M corresponding to eigenvalue 1. Take $\{u_1, u_2\}$ as given above as a

basis. Then the actions of A, B correspond to the matrices P, Q given by

$$P = \begin{pmatrix} 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}, \quad Q = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ 0 & 1 \end{pmatrix}.$$

This follows since $u_1.A = u_1$, $u_2A = (1,0)A = (\frac{3}{4}, \frac{1}{4}) = \frac{1}{3}u_1 + \frac{2}{3}u_2$, $u_1B = (\frac{1}{4}, \frac{3}{4})B = (\frac{1}{2}, \frac{1}{2}) = \frac{1}{2}u_1 + \frac{1}{2}u_2$, $u_2B = u_2$.

It is easily verified that P and Q both have right eigenvector $(1,1)^t$ associated with eigenvalue 1 and left eigenvector $(1,-1)$ associated with left eigenvector $\frac{2}{3}$. Suppose N is some product of the matrices P and Q of length $l \geq 1$. It will necessarily have right eigenvector $(1,1)^t$ associated with eigenvalue 1, and left eigenvector $(1,-1)$ associated with eigenvalue $(\frac{2}{3})^l$. The important thing to note here is that N is necessarily diagonalisable, as it has two distinct eigenvalues. Thus N has a unique normalised eigenvector associated with eigenvalue 1. It is also worth noting that N (and hence the similar matrix M) is diagonalisable, and since it has eigenvalues 1 and $\frac{2}{3}$, the limit of N^i (and M^i) as i tends to infinity must exist.

Now suppose we have a normalised eigenvector of M corresponding to eigenvalue 1 (with the usual basis) written as $(\alpha u_1 + (1-\alpha)u_2)$, for some $\alpha \in \mathbb{R}$, i.e. $(\alpha u_1 + (1-\alpha)u_2)M = \alpha u_1 + (1-\alpha)u_2$. Then written in terms of the new basis $\{u_1, u_2\}$ we have

$$(\alpha, 1-\alpha)N = (\alpha, 1-\alpha)$$

where N is the matrix representing the action of M in the new basis, i.e. the matrix constructed by multiplying matrices P and Q in the same order that M was built by multiplying A and B. Thus in order to find eigenvectors for M, we now turn our attention to finding the eigenvectors of N, i.e. solving the equation

$$(\alpha, 1-\alpha)N = (\alpha, 1-\alpha)$$

for any matrix $N \in \langle P, G \rangle$.

Claim 5.5.2. *Let $N = \prod_{i=1}^{l} N_i$ where $N_i \in \{P, Q\}$ $\forall i$. Then*

$$(\alpha, 1-\alpha)N = \left(\frac{2^l\alpha + a}{3^l}, 1 - \frac{2^l\alpha + a}{3^l} \right)$$

where $a = \sum_{i=1}^{l} a_i 2^{l-i} 3^{i-1}$ and $a_i = 1$ if $N_i = P$, and $a_i = 0$ if $N_i = Q$.

Proof. We induct on l. If $l = 1$ then $N = P$ or $N = Q$.

- $(\alpha, 1-\alpha)P = (\alpha, 1-\alpha) \begin{pmatrix} 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} = \left(\frac{2\alpha+1}{3}, 1 - \frac{2\alpha+1}{3}\right)$

and we have $a_1 = 1$ and $N_1 = P$.

- $(\alpha, 1-\alpha)Q = (\alpha, 1-\alpha)\begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ 0 & 1 \end{pmatrix} = \left(\frac{2\alpha}{3}, 1 - \frac{2\alpha}{3}\right)$
  and we have $a_1 = 0$ and $N_1 = Q$.

Hence the claim holds for $l = 1$.

Now suppose the claim holds for all matrices $N \in \langle P, Q\rangle$ of length l, for some $l \geq 1$. Let N have length $l+1$. We consider the two cases (i) $N_{l+1} = P$ and (ii) $N_{l+1} = Q$.

- Suppose $N_{l+1} = P$. Then if $a = \sum_{i=1}^{l} a_i 2^{l-i} 3^{i-1}$ and the a_i are as described above, then

$$\begin{aligned}(\alpha, 1-\alpha)\prod_{i=1}^{l} N_i.P &= \left(\frac{2^l\alpha + a}{3^l}, 1 - \frac{2^l\alpha + a}{3^l}\right)\begin{pmatrix} 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \\ &= \left(\frac{2^{l+1}\alpha + (2a + 3^l)}{3^{l+1}}, 1 - \frac{2^{l+1}\alpha + (2a+3^l)}{3^{l+1}}\right).\end{aligned}$$

  Now

$$\begin{aligned}2a + 3^l &= 2\sum_{i=1}^{l} a_i 2^{l-i} 3^{i-1} + 3^l \\ &\quad \text{(where } a_i = 1 \text{ when } N_i = P \text{ and } a_i = 0 \text{ when } N_i = Q\text{)} \\ &= \sum_{i=1}^{l+1} a_i 2^{(l+1)-i} 3^{i-1} \\ &\quad \text{(where } a_i = 1 \text{ when } N_i = P \text{ and } a_i = 0 \text{ when } N_i = Q\text{).}\end{aligned}$$

- Suppose $N_{l+1} = Q$. Then if $a = \sum_{i=1}^{l} a_i 2^{l-i} 3^{i-1}$ and the a_i are as described above, then

$$\begin{aligned}(\alpha, 1-\alpha)\prod_{i=1}^{l} N_i.Q &= \left(\frac{2^l\alpha + a}{3^l}, 1 - \frac{2^l\alpha + a}{3^l}\right)\begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ 0 & 1 \end{pmatrix} \\ &= \left(\frac{2^{l+1}\alpha + 2a}{3^{l+1}}, 1 - \frac{2^{l+1}\alpha + 2a}{3^{l+1}}\right).\end{aligned}$$

  Now

$$\begin{aligned}2a &= 2\sum_{i=1}^{l} a_i 2^{l-i} 3^{i-1} \quad \text{(where } a_i = 1 \text{ when } N_i = P \text{ and } a_i = 0 \text{ when } N_i = Q\text{)} \\ &= \sum_{i=1}^{l+1} a_i 2^{(l+1)-i} 3^{i-1} \quad \text{(where } a_i = 1 \text{ when } N_i = P \text{ and } a_i = 0 \text{ when } N_i = Q\text{).}\end{aligned}$$

Thus by induction the claim holds for all $N \in \langle P, Q \rangle$ of length at least 1. $\square$

Having established the claim, we continue with the proof Theorem 5.5.1.

Let N be a word on $\{P, Q\}$. By the claim above,

$$(\alpha, 1-\alpha)N = \left(\frac{2^l\alpha + a}{3^l}, 1 - \frac{2^l\alpha + a}{3^l} \right)$$

where $a = \sum_{i=1}^{l} a_i 2^{l-i} 3^{i-1}$ and $a_i = 1$ if $N_i = P$ and $a_i = 0$ if $N_i = Q$. Thus

$$\begin{aligned} (\alpha, 1-\alpha)N = (\alpha, 1-\alpha) &\iff \alpha = \frac{2^l\alpha + a}{3^l}, \quad (a \text{ as above}) \\ &\iff \alpha = \frac{a}{3^l - 2^l}. \end{aligned}$$

Thus the eigenvector associated with $\lambda = 1$ for N is

$$\left(\frac{a}{3^l - 2^l}, 1 - \frac{a}{3^l - 2^l} \right).$$

We now use this to find the eigenvector for M corresponding to $\lambda = 1$. Since we chose u_1 and u_2 as a basis, it follows that the normalised eigenvector (with respect to the standard basis) for M corresponding to eigenvalue 1 is

$$\frac{a}{3^l - 2^l} u_1 + \left(1 - \frac{a}{3^l - 2^l} \right) u_2 = \left(1 - \frac{3a}{4(3^l - 2^l)}, \frac{3a}{4(3^l - 2^l)} \right).$$

We now see how we use this eigenvector.

Let $I \neq M \in \langle A, B \rangle$ and let w_M be the word constructed from w_A and w_B in the order dictated by M. If $v_0 := \left(\frac{1}{|\text{Alt}(4)|}, \frac{|T_2|}{|\text{Alt}(4)|} \right)$, then

$$v_{M^i} := (P(\text{Alt}(4), w_{M^i} = 1), P(\text{Alt}(4), w_{M^i} \in T_2)) = v_0.M^i$$

and

$$v_{M^\infty} := \lim_{i \to \infty} v_{M^i} = \lim_{i \to \infty} \left(\, P(\text{Alt}(4), w_{M^i} = 1), P(\text{Alt}(4), w_{M^i} \in T_2) \, \right).$$

Note that

$$v_{M^\infty} M = \left(\lim_{i \to \infty} v_0 M^i \right) M = \lim_{i \to \infty} v_0 M^i = v_{M^\infty}.$$

Thus v_{M^∞} is a normalised eigenvector of M corresponding to eigenvalue 1. Thus since we have seen there is only one such vector we have that

$$v_{M^\infty} = \left(1 - \frac{3a}{4(3^l - 2^l)}, \frac{3a}{4(3^l - 2^l)} \right)$$

where $a = a(M)$ is as defined above. Thus for any $g_2 \in T_2$,

$$\lim_{i\to\infty} P(\mathrm{Alt}(4), w_{M^i} = 1) = 1 - \frac{3a}{4(3^l - 2^l)}, \tag{5.23}$$

$$\lim_{i\to\infty} P(\mathrm{Alt}(4), w_{M^i} = g_2) = \frac{a}{4(3^l - 2^l)}. \tag{5.24}$$

Note, since w_M is constructed from substitutions of w_A and w_B we have that

$$P(\mathrm{Alt}(4), w_M = g_2) = P(\mathrm{Alt}(4), w_M = h_2) \qquad \forall g_2, h_2 \in T_2,$$

and thus $P(\mathrm{Alt}(4), w_M = g_2) = \frac{1}{3}P(G, w_M \in T_2)$. Now (5.23) and (5.24) correspond to limit points of $S(\mathrm{Alt}(4))$, so it remains to show that such points are dense in the interval $[0, 1]$.

Lemma 5.5.3. *Define* $a(l, a_1, \ldots, a_l) := \sum_{i=1}^{l} a_i 2^{l-i} 3^{i-1}$. *Then the set*

$$A := \left\{ \frac{a(l, a_1, \ldots, a_l)}{3^l - 2^l} \mid l \in \mathbb{N}, a_i \in \{0, 1\} \ \forall i \right\}$$

is dense in $[0, 1]$.

Proof. Let $p \in [0, 1]$. We show that there is a sequence of elements $(q_l)_{l\in\mathbb{N}}$ in A such that $q_l \to p$ as $l \to \infty$. Let $x_l := p(3^l - 2^l)$. Note that $x_l \in [0, 3^l - 2^l]$. Let $A(l)$ be the ordered set $\left\{3^{l-1}, 2.3^{l-1}, \ldots, 2^{l-1}\right\}$. We define $y_l := \sum_{i=1}^{l} a_i 2^{l-i} 3^{i-1}$ to be the sum of elements of $A(l)$ given by the following algorithm.

1. If $3^{l-1} \leq x_l$ then set $a_l := 1$. Otherwise set $a_1 := 0$.
2. Now if $2.3^{l-2} + a_l 3^{l-1} \leq x_l$ then set $a_{l-1} = 1$. Otherwise let $a_{l-1} := 0$.
3. Continue likewise, namely if

$$2^i 3^{l-i-1} + \sum_{j=l-i+1}^{l} a_j 2^{l-j} 3^{j-1} \quad \leq \quad x_l$$

then set $a_{l-i} = 1$. Otherwise let $a_{l-1} := 0$.

Thus $y_l \leq x_l$, and $\frac{y_l}{3^l - 2^l} \in A$. Note that the elements of $A(l)$ may be written

$$2^{l-1}\left(\tfrac{3}{2}\right)^{l-1}, 2^{l-2}\left(\tfrac{3}{2}\right)^{l-2}, \ldots, 2^{l-2}\left(\tfrac{3}{2}\right)^{0},$$

and that

$$1 + \tfrac{3}{2} < \left(\tfrac{3}{2}\right)^2 \leq \left(\tfrac{3}{2}\right)^i \quad \forall i \geq 2.$$

Suppose $2^{l-1}(\frac{3}{2})^i$ was rejected for some $i \geq 2$ (i.e. $a_{i+1} = 0$). Then since

$$\left(\tfrac{3}{2}\right)^i < 1 + \tfrac{3}{2}$$
$$\therefore \quad 2^{l-1}\left(\tfrac{3}{2}\right)^i < 2^{l-1} + 2^{l-1}\left(\tfrac{3}{2}\right).$$

Since $2^{l-1}\left(\frac{3}{2}\right)^i$ was too large to be accepted, 2^{l-1} and $2^{l-1}\left(\frac{3}{2}\right)$ could not both have been accepted, as their sum is larger. So if this were the case, at least one of $2^{l-1}, 2^{l-2}$ must have been rejected. Thus either

1. $2^{l-1}\left(\frac{3}{2}\right)^i$ was accepted for all i, i.e. $a_i = 1\ \forall i$ and $y_l = \sum_{i=1}^{l} 2^{l-i}3^{i-1} = 3^l - 2^l$. Then $x_l = y_l$ and so $p = \frac{y_l}{3^l - 2^l} \in S$. So let $q_l = p \quad \forall l$. Or,

2. At least one of 2^{l-1} or 3.2^{l-2} was rejected for being too large. This means

$$|x_l - y_l| \leq 3.2^{l-2}.$$

   Define $q_l := \frac{y_l}{3^l - 2^l}$. This is an element of A by definition of y_l. Then

$$|p - q_l| \quad = \quad \tfrac{x_l}{3^l - 2^l} - \tfrac{y_l}{3^l - 2^l} \quad \leq \quad \tfrac{3.2^{l-2}}{3^l - 2^l}.$$

   Note that $\frac{3.2^{l-2}}{3^l - 2^l} \to 0$ as $l \to \infty$ (to see this note that $\frac{3^l - 2^l}{3.2^{l-2}} = 3\left(\frac{3}{2}\right)^{l-2} - \frac{2^2}{3} \to \infty$ as $l \to \infty$). Thus $q_l \to p$ as $l \to \infty$.

Hence in either case, p is the limit point of elements in A, and so A is dense in $[0, 1]$, and the lemma is proved. □

In light of Lemma 5.5.3 it follows that

- $\left\{ \frac{a(M)}{4\left(3^l - 2^l\right)} \mid M \text{ is a word on } \{A, B\} \right\}$ is dense in $[0, \frac{1}{4}]$ and
- $\left\{ 1 - \frac{3a(M)}{4\left(3^l - 2^l\right)} \mid M \text{ is a word on } \{A, B\} \right\}$ is dense in $[\frac{1}{4}, 1]$.

Thus it follows that

- $S(\mathrm{Alt}(4), g_2)$ is dense in $[0, \frac{1}{4}]$ for any $g_2 \in T_2$ and
- $S(\mathrm{Alt}(4), 1)$ is dense in $[\frac{1}{4}, 1]$.

Thus we conclude that $S(\mathrm{Alt}(4))$ is dense in $[0, 1]$ and the theorem is proved. □

Corollary 5.5.4. *Let Sym*(4) *denote the symmetric group on four elements. Then*

$$S(Sym(4)) \text{ is dense in } [0, 1].$$

Proof. The verbal subgroup of Sym(4) associated with the word x^2 is isomorphic to the alternating group Alt(4). Thus by Lemma 2.3.1 it follows that $S(\mathrm{Sym}(4))$ is dense in $[0,1]$. □

We now have a complete picture of the sets of probabilities associated with the alternating and symmetric groups, which we outline below.

5.5.1 The alternating groups

- If $n \leq 3$ then $\mathrm{Alt}(n)$ is abelian, and thus $S(\mathrm{Alt}(n))$ is finite, as described in Section 2.4.
- If $n = 4$ then $S(\mathrm{Alt}(n))$ is dense in $[0,1]$ as stated in Theorem 5.5.1 above.
- If $n \geq 5$ then $\mathrm{Alt}(n)$ is simple, and thus by Theorem 2.5.2, $S(\mathrm{Alt}(n))$ is dense in $[0,1]$.

Corollary 5.5.5. *If G is an alternating group, then $S(G)$ is dense in $[0,1]$ if, and only if, G is non-nilpotent.*

5.5.2 The symmetric groups

- If $n \leq 2$ then $\mathrm{Sym}(n)$ is abelian, and $S(\mathrm{Sym}(n))$ is finite, as described in §2.4.
- If $n = 3$ then $S(\mathrm{Sym}(n))$ is dense in $[0,1]$ as seen in Theorem 5.1.1.
- If $n \geq 4$ then since $S(\mathrm{Alt}(n))$ is dense in $[0,1]$ and $\mathrm{Alt}(n)$ is a verbal subgroup of $\mathrm{Sym}(n)$, it follows from Lemma 2.3.1 that $S(\mathrm{Sym}(n))$ is also dense in $[0,1]$.

Corollary 5.5.6. *If G is a symmetric group, then $S(G)$ is dense in $[0,1]$ if, and only if, G is non-nilpotent.*

Chapter 6

Conclusion

6.1 Summary

Within this thesis we have sought to further establish the relationship between a finite group G and its associated set of probabilities $S(G)$. We shall briefly recap the content here.

In Section 2.2 we established two original methods of constructing sequences of words that can be used to prove the existence of accumulation points of $S(G)$. The concatenation method (§2.2.1) may be used to prove the existence of accumulation points of the form $|H|^{-1}$ where H is a verbal subgroup. This method applied to the commutator word was used to prove that $S(G)$ is infinite for any non-abelian finite group (see Lemma 2.2.9). The second method, called the substitution method (§2.2.2) is a generalisation of left normed commutators. We have used this construction elsewhere in this thesis to prove the existence of accumulation points that are not the reciprocals of the order of some verbal subgroup. It is this method that is used in Chapter 5 to show that $S(G)$ is dense in $[0, 1]$ for certain non-nilpotent groups.

In Section 2.3 we collected results regarding direct products, quotient groups and verbal subgroups, most of which have been noted by other authors before.

Section 2.4 includes what we know of abelian groups. As has been noted by other authors, we show that words over abelian groups are homomorphisms (Lemma 2.4.1), and use this to describe $S(G)$ explicitly, as the reciprocals of the orders of the groups verbal subgroups, along with zero (Lemma 2.4.3). This, along with Lemma 2.2.9 shows that $S(G)$ is finite if and only if G is abelian (Lemma 2.4.4). We also go on to show that for finite abelian groups, $S(G)$ characterises G up to isomorphism (Lemma 2.4.7).

In Sections 2.5 and 2.6 we consider simple and verbally simple groups and conjecture that for a finite group G, $S(G)$ is dense in $[0, 1]$ if and only if G is non-nilpotent. We

return to this conjecture throughout the thesis.

In Chapter 3 we provide a method for calculating $S(G)$ in the case that G has low order and nilpotency class. As far as we are aware this has not been attempted before. Our method involves converting a word written in standard form into a product of elements from a generating set with polynomial exponents. We then count solutions to these polynomial equations using results we have established throughout the chapter, many of which are collected in Section 3.2. We then go on to calculate $S(G)$ for all groups of nilpotency class two and order 8, 16 and 27. We also perform these calculations for three groups of order 32. A summary of these results can be found in Section 3.16. In response to our findings, we conjecture that every accumulation point associated with a nilpotent group is the reciprocal of the order of a verbal subgroup of that group, since this holds for the groups investigated in this chapter. We also note that for these groups, the infimum of the positive probabilities is $|G|^{-1}$ in every case. In Section 3.3 and 3.4 we show that the dihedral group of order 8 and the quaternion group of order 8 have the same set of probabilities. This is in contrast to abelian groups for which $S(G)$ is unique to the group G (Lemma 2.4.7). We also show the existence of a nilpotent group that has a verbal subgroup whose order does not appear as the reciprocal of an accumulation point of $S(G)$ (see §3.7).

The infimum problem was explored in Chapter 4. We proved that if G is a nilpotent dihedral group (§4.2) or a nilpotent generalised quaternion group (§4.3) then the infimum of the positive probabilities associated with G is $|G|^{-1}$, just as for abelian groups. This gives us examples of nilpotent groups of every nilpotency class that satisfy this property, along with those of class 2 from the previous chapter. This led us tentatively to the conjecture that this is the case for every nilpotent group (Conjecture 4.4.1).

In Chapter 5 we turned our attention to non-nilpotent groups. We proved that if G is a non-nilpotent dihedral group (Theorem 5.3.1) or a non-nilpotent generalised dihedral group (Theorem 5.4.4) then $S(G)$ is dense in $[0, 1]$. We did this by using the substitution method to generate a set of words whose associated probabilities could be shown to be dense in the interval. We concluded this chapter by stating how our results show that if G is an alternating group (§5.5.1) or a symmetric group (§5.5.2), then $S(G)$ is dense in $[0, 1]$ if and only if G is non-nilpotent, just as for generalised dihedral groups.

6.2 Open problems, conjectures and further work

The purpose of this thesis has been to further establish the link between certain properties of a finite group and properties of the set of probabilities associated with that group. Here we list some remaining open problems and potential further work, and give reference to the material in the thesis to which they relate.

6.2.1 Further calculations

In Chapter 3 we used generating sets and results about polynomials over finite fields to explicitly calculate $S(G)$ for a handful of small groups of nilpotency class 2. This work could be continued either to calculate $S(G)$ for groups of higher order or higher nilpotency class. However it is worth noting that the complexity of the problem will increase with nilpotency class. Further data here may shed light on several of the questions and conjectures regarding nilpotent groups set out below, in particular those concerning accumulation points and the infima of sets of associated probabilities.

In order to explore the remaining open problems, one might consider using GAP or another software package to gather data. Indeed when we began investigating the probabilities associated with D_8, we first used GAP to calculate the probabilities associated with words on 2, 3, 4, 5 and 6 variables. From this data we were able to conjecture what $S(D_8)$ might be. We were then able to prove that we were correct. If one were doubtful that Conjecture 6.2.1 is true in general for example, one might consider which groups and which words might be likely to provide a counter example. Provided one does not cast the net too wide, a computer program such as GAP can then be used to quickly calculate these probabilities. Of course there are many questions and many approaches that could be explored along these lines.

6.2.2 Conjectures

Here we collect the conjectures given throughout this thesis.

Conjecture 6.2.1. *Let G be a finite nilpotent group. Then*

$$\inf_{w,g} P(G, w = g) = |G|^{-1}.$$

where w ranges over F_∞ and g ranges over G_w^+.

If this is not true in general, then it would be interesting to know which groups satisfy this property. In Chapter 3 we show this to be the case for the handful of groups with nilpotency class 2 for which we calculated $S(G)$. In Chapter 4 we show that the nilpotent generalised dihedral groups and generalised quaternion groups satisfy this property.

Conjecture 6.2.2. *Let G be a finite nilpotent group. Every accumulation point of $S(G)$ is of the form m^{-1}, where m is the order of some verbal subgroup of G.*

The groups for which we calculated the set of probabilities in Chapter 3 all satisfied this property, though of course this is very limited evidence. The groups we investigated were all of class 2, so it is possible that this is not the case for groups of higher nilpotency class.

Conjecture 6.2.3. *Let G be a finite group. Then $S(G)$ is dense in the interval $[0,1]$ if and only if G is non-nilpotent.*

Segal and Nikolov have shown that if G is nilpotent, 0 is not an accumulation point of $S(G)$ (see Theorem 1.1.9 and [23]). Thus the forwards implication is true. In Chapter 5 we find that this is true for the generalised nilpotent groups, generalised quaternion groups, the alternating groups and the symmetric groups. However since our methods here are constructive and use specific properties of the groups involved, it is probable a general proof would need to be very different. In light of Conjecture 6.2.2, one might also consider the following.

Conjecture 6.2.4. *Let G be a finite group. Then $S(G)$ is nowhere dense if and only if G is nilpotent.*

6.2.3 Other open problems

Accumulation points

In Chapter 3 we found that all the accumulation points of $S(G)$ for these nilpotent groups were the reciprocals of the orders of verbal subgroups. For the majority of groups investigated, if the group had a proper verbal subgroup of order m, m^{-1} would appear as an accumulation point of $S(G)$. However, this was not always the case. The groups $[16,6]$ and $[32,4]$ both have proper verbal subgroups whose orders are not reflected in $S(G)$ in this way. Hence the following open question.

Question 1. *Which nilpotent groups and verbal subgroups satisfy this phenomenon?*

Maximal intervals

In Theorem 5.1.1 we established that the set of probabilities associated with the symmetric group on three letters is dense in $[0,1]$. In the proof we showed that $S(G,1)$ is dense in $[\frac{1}{3},1]$ and that if g is a 3-cycle, $S(G,g)$ is dense in $[0,\frac{1}{3}]$. One might ask whether these intervals are maximal. Hence

Question 2. *What is the largest interval T such that $S(G,1)$ is dense in T? What about for $S(G,g)$ where g is a 3-cycle?*

Of course this question could be explored for any number of groups, and may shed light on whether the bound given in Theorem 1.0.1 as an infimum for $S(G,1)$ for soluble groups is best possible.

Calculating the set of probabilities for dihedral and generalised quaternion groups

In Section 4.2 we prove that if G is a dihedral group of order 2^n for some integer n, then

$$P(G, w = g) \geq |G|^{-1}$$

for any $w \in F_\infty$ and $g \in G_w^+$. In the first half of the proof, we show that if g is a non-central element

$$P(D_{2^n}, w = g) = \tfrac{1}{2} P(D_{2^{n-1}}, w = g^*)$$

where g^* is the image of g under the canonical homomorphism

$$D_{2^n} \to D_{2^n}/Z(D_{2^n}) \cong D_{2^{n-1}}.$$

This means that if we know $S(D_{2^{n-1}}, g^*)$, then we know $S(D_{2^n}, g)$. However if g is central, we cannot use this method. For example, since we have calculated $S(D_8, g)$ for all group elements g, we can calculate $S(D_{16}, g)$ for all g except the two central elements. Hence the following is still unknown.

Question 3. *What is $S(D_{16})$? What is $S(D_{2^n})$ for any n strictly larger than 3?*

The same problem occurs in the proof of Theorem 4.3.1 for generalised quaternion groups. Hence we may also ask the following.

Question 4. *What is $S(Q_{16})$? What is $S(Q_{2^n})$ for any n strictly larger than 3?*

Inverse elements

If two elements g, h in a finite group are conjugate or auto-equivalent (see Definition 4.2.3 in §4.2.3) then for any word w

$$P(G, w = g) = P(G, w = h).$$

But the following is still unknown.

Question 5. *If G is a finite group, $g \in G$ and $w \in F_\infty$ is it necessarily true that*

$$P(G, w = g) = P(G, w = g^{-1})?$$

Groups for which the largest fibre is always associated with 1

In Chapter 3 we found that for some groups, given any word, no group element ever has a larger fibre than 1 (D_8 for example). We also found examples of groups in which there were infinitely many words for which this is not the case (Q_8 for example). Hence the following question.

Question 6. *Which finite groups satisfy*

$$P(G, w = 1) \geq P(G, w = g)$$

for all words w and all group elements g?

There are of course many other directions that could be taken and countless other questions that could be asked. It is the hope of the author that others may be interested enough to develop this work further and shed light on some of these unknowns. If so, please feel free to get in touch with the author to ask questions or to share any new developments.

Bibliography

[1] M. Abért. On the probability of satisfying a word in a group. *J. Group Theory*, 9:685–694, 2006.

[2] J. D. Dixon. The probability of generating the symmetric group. *Math.Z*, 110:199 – 205, 1969.

[3] J. D. Dixon. Probabilistic group theory. *C.R. Math. Rep. Acad. Sci. Canada*, 24:1 –15, 2002.

[4] P. Erdös and P. Turán. On some problems of statistical group-theory. I. *Z.Wahrschein. Verw. Gebeite*, 4:175–186, 1965.

[5] P. Erdös and P. Turán. On some problems of a statistical group-theory. II. *Acta Math. Acad. Sci. Hungar.*, 18:151–163, 1967.

[6] P. Erdös and P. Turán. On some problems of a statistical group-theory. III. *Acta Math. Acad. Sci. Hungar.*, 18:309–320, 1967.

[7] P. Erdös and P. Turán. On some problems of a statistical group-theory. IV. *Acta Math. Acad. Sci. Hungar.*, 19:413–435, 1967.

[8] K. George. *Verbal properties of certain groups.* PhD thesis, University of Cambridge, 1976.

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

[10] M. Hall. *The Theory of Groups, Second Edition.* AMS Chelsea Publishing, 1976.

[11] P. Hall. Verbal and marginal subgroups. *Journal fur die reine und angewandte Mathematik (Crelles Journal)*, 182:156 – 157, 1940.

[12] P. Hall. Nilpotent groups, (Lectures given at the Canadian Math. Congress, University of Alberta, 1957) Queen Mary College. *Journal fur die reine und angewandte Mathematik (Crelles Journal)*, 1969.

[13] Á. Seress J. D. Dixon, L. Pyber and A. Shalev. Residual properties of free groups and probabilistic methods. *J. reine angew. Math. (Crelles)*, 556:159172, 2003.

[14] W. M. Kantor and A. Lubotzky. The probability of generating a finite classical group. *Geometriae Dedicata*, 36:67 –87, 1990.

[15] L. G. Kovács and M. F. Newman. Minimal verbal subgroups. *Proc. Camb. Phil. Soc.*, 62:347–350, 1966.

[16] M. Levy. On the probability of satisfying a word in nilpotent groups of class 2, preprint. `http://arxiv.org/abs/1101.4286`, 2011.

[17] R. Lidl and H. Niederreiter. *Finite fields*. Number 20 in Encyclopedia of mathematics and its applications. Cambridge University Press, 1997.

[18] M. W. Liebeck and A. Shalev. The probability of generating a finite simple group. *Geometriae Dedicata*, 56:103–113, 1995.

[19] M. W. Liebeck and A. Shalev. Classical groups, probabilistic methods, and the (2,3) generation problem. *Ann. of Math*, 144:77–125, 1996.

[20] A. Shalev M. W. Liebeck, E. A. O'Brien and P. H. Tiep. The ore conjecture. *J. Eur. Math. Soc. (JEMS)*, 12:939 – 1008, 2010.

[21] J. Medhi. *Stochastic processes, Second edition*. New International Publishers, 1994.

[22] E. Netto. *Substitutionentheorie and ihre Anwendungen auf die Algebra*. Teuber, Leipzig, 1882.

[23] N. Nikolov and D. Segal. A characterization of finite soluble groups. *Bull. London Math. Soc.*, 39:209–213, 2007.

[24] N. Nikolov and D. Segal. On finitely generated profinite groups, I: strong completeness and uniform bounds. *Annals of Math.*, 165:171–238, 2007.

[25] A. H. Rhemtulla. A problem of bounded expressibility in free products. *Proc. Cambridge Philos. Soc.*, 64:573 – 584, 1968.

[26] V. A. Romankov. Width of verbal subgroups in solvable groups. *Algebra and Logic*, 21:41 – 49, 1982.

[27] D. Segal. Words and groups. In *Groups St Andrews 2009 in Bath*, LMS Lecture Note Series 388, pages 344–374. Cambridge University Press, Cambridge, 2009.

[28] D. Segal. *Words: notes on verbal width in groups*. LMS Lecture Note Series 361. Cambridge University Press, Cambridge, 2009.

[29] A. Shalev. Word maps, conjugacy classes, and a noncommutative Waring-type theorem. *Annals of Math.*, 170:1383 – 1416, 2009.

[30] P. W. Stroud. *Topics in the theory of verbal subgroups.* PhD thesis, University of Cambridge, 1966.