On the subgroup permutability degree of some finite simple groups.

Aivazidis, Stefanos

On the subgroup permutability degree of some finite simple groups

Stefanos Aivazidis

Thesis submitted in partial fulfilment of the requirements of the degree of

Doctor of Philosophy

School of Mathematical Sciences

Queen Mary University of London

February 2015

To the memory of my dear friend Giorgos…

Contents

List of Figures

Declaration of Authorship

I, Stefanos Aivazidis, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below. I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material. I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis. I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university. The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature: Stefanos Aivazidis
Date: 18th July, 2015

Details of collaboration and publications:

- [Aiv13] S. Aivazidis, *The subgroup permutability degree of projective special linear groups over fields of even characteristic*, J. Group Theory **16** (2013), no. 3, 383–396.
- [Aiv15] S. Aivazidis, *On the subgroup permutability degree of the simple Suzuki groups*, Monatsh. Math. **176** (2015), no. 3, 335–358.
- [AS14] S. Aivazidis and E. Sofos, *On the distribution of the density of maximal order elements in general linear groups*, Ramanujan J. (2014), 1-25.

My contribution to the paper written jointly with E. Sofos which appears (almost) verbatim in Appendix A is all material up to and including section A.2 on Singer cycles.

Acknowledgements

I thank my supervisor, Robert A. Wilson, for his guidance during the past 4 years. I would never have had the opportunity to pursue this work without financial support from the Chrysovergis Endowment, under the auspices of the National Scholarships Foundation of Greece. I am grateful to my friends for their help and companionship, both in mathematics and in life. Finally, I thank my parents for their unwavering love and support.

Stefanos Aivazidis
London, February 2015

Abstract

Consider a finite group G and subgroups H,K of G. We say that H and K permute if $HK = KH$ and call H a permutable subgroup if H permutes with every subgroup of G. A group G is called quasi-Dedekind if all subgroups of G are permutable. We can define, for every finite group G, an arithmetic quantity that measures the probability that two subgroups (chosen uniformly at random with replacement) permute and we call this measure the subgroup permutability degree of G. This measure quantifies, among others, how close a finite group is to being quasi-Dedekind, or, equivalently, nilpotent with modular subgroup lattice. The main body of this thesis is concerned with the behaviour of the subgroup permutability degree of the two families of finite simple groups $\mathrm{PSL}_2(2^n)$, and $\mathrm{Sz}(q)$. In both cases the subgroups of the two families of simple groups are completely known and we shall use this fact to establish that the subgroup permutability degree in each case vanishes asymptotically as n or q respectively tends to infinity. The final chapter of the thesis deviates from the main line to examine groups, called $\mathfrak{F}$-groups, which behave like nilpotent groups with respect to the Frattini subgroup of quotients. Finally, we present in the Appendix joint research on the distribution of the density of maximal order elements in general linear groups and offer code for computations in GAP related to permutability.

Caminante, son tus huellas
el camino, y nada más;
caminante, no hay camino,
se hace camino al andar.

Antonio Machado
Proverbios y cantares XXIX, Campos de Castilla

1 Introduction

The application of probabilistic methods to a mathematical field other than probability theory itself (e.g. number theory, graph theory, group theory) has often led to interesting results, which themselves could not be obtained by standard techniques within that field. These results are, generally speaking, of the following nature: it is shown that, in a certain sense, "most" elements of a set of mathematical objects possess a certain property, thereby characterising the underlying structure in terms of that property.

A rigorous definition of the nature of probabilistic group theory is beyond the scope of this thesis or the interests of its author. We elect, instead, to loosely describe the main areas of focus within the field. These are:

1. the characterisation of a group or families of groups by analysing proportions of elements (subgroups, conjugacy classes of elements or subgroups etc.) that satisfy a certain property,

2. the application of probability to the construction of algorithms in computational group theory, and

3. proving deterministic statements about groups via probabilistic methods.

We shall have very little to say about the latter two areas; the interested reader is referred to the expository article of Dixon [Dix02] and the thorough list of references it provides[1].

Despite the fact that the present thesis contains material (Chapter 5 and Appendix A) seemingly unrelated to the main theme, which is subgroup permutability and the subgroup permutability degree, it should be stressed that all, or nearly all, material is about finite groups and probabilities in the sense of proportions. An exception is Chapter 5 where, among others, we attempt to describe structural aspects of a certain class of groups with specific behaviour relative to the Frattini subgroup and offer an improvement[2] of an already existing algorithm of Hulpke for the computation of representatives of the conjugacy classes of subgroups of nilpotent groups. However, Chapter 5 came about from an (unsuccessful) attempt to generalise the argument given in Lemma 4.7 to all finite p-groups, the ultimate goal being to deduce good lower and upper bounds for the size of the table of marks of a candidate group in terms of other invariants of the group.

All in all, this thesis is a very modest attempt at a contribution to the first area of probabilistic group theory mentioned above.

A short description of the contents of each chapter is in order.

- The rest of the present chapter is devoted to a brief overview of some milestones in probabilistic group theory; more specifically in the areas of probabilistic generation of finite simple groups and the commuting probability, an arithmetic measure which quantifies the “abelianness” of a finite group.

- Chapter 2 provides an overview of subgroup permutability and related notions and introduces the concept of the subgroup permutability degree of a finite group. This is an arithmetic measure analogous to the commuting probability, but it is defined in terms of “commuting” pairs of subgroups instead of elements. It measures the extent to which a group is Iwasawa, i.e. has all subgroups permutable. The class of Iwasawa groups certainly contains the class of abelian groups, but it is larger (and is strictly contained within the class of nilpotent groups). The subgroup permutability degree

[1] In fact, the writing of this chapter has benefitted from Dixon’s exposition.

[2] It would probably be more accurate to say that the modified algorithm is *likely* to run faster than the one currently in use, but no guarantee is made, since we haven’t attempted an actual implementation.

can thus be viewed as a slight generalisation of the commuting probability. Note, however, that we will not attempt to make this precise, nor will we explore connections between the two arithmetic measures. We close the chapter with a criterion (Lemma 2.16) for the vanishing of the subgroup permutability degree of a family of finite groups.

- The content of the following two chapters is reproduced from publications of the author ([Aiv13] and [Aiv15] respectively). They constitute the main body of this thesis and examine the behaviour of the subgroup permutability degree of the two families of simple groups $\mathrm{PSL}_2(2^n)$ and $\mathrm{Sz}(q)$. In both cases the subgroups of the two families of simple groups are completely known and we shall use this fact to establish that the subgroup permutability degree vanishes asymptotically as n or q respectively tends to infinity, thus taking a first step towards a proof of the following conjecture:

  Let $\mathcal{S}$ be the set of all nonabelian finite simple groups. Then the probability that two subgroups of G permute tends to 0 as $|G| \to \infty$, $G \in \mathcal{S}$.

- Chapter 5, as we have already mentioned, is independent from the rest of the thesis. In general, if G is a finite group and N is a normal subgroup of G then $\Phi(G/N) \geqslant \Phi(G)N/N$ and there exist pairs (G, N) for which equality fails to hold (consider the Frobenius group of order 20 and its Sylow 5–subgroup). Our goal in this chapter is to extract as much information as possible about finite groups all of whose normal subgroups satisfy $\Phi(G/N) = \Phi(G)N/N$.

- Appendix A presents joint research with Efthymios Sofos and its content is reproduced (verbatim) from our joint publication [AS14]. We examine elements of maximal order in $\mathrm{GL}_n(q)$, also known as Singer cycles, and study the distribution of their density in $\mathrm{GL}_n(q)$.

- Appendix B presents GAP code which can be used to compute or construct objects related to permutability.

1.1 Probabilistic generation

Probabilistic generation is that subfield within probabilistic group theory which is concerned with questions of type:

Given a family of groups G_n, what is the probability that a d-tuple of elements with prescribed properties, chosen uniformly at random, generates G_n as $n \to \infty$?

The earliest recording of a problem of this sort is due to E. Netto who, in 1892, writes[3]:

> If we arbitrarily select two or more substitutions of n elements, it is to be regarded as extremely probable that the group of lowest order which contains these is the symmetric group, or at least the alternating group. In the case of two substitutions the probability in favour of the symmetric group may be taken as about $\frac{3}{4}$, and in favour of the alternating, but not symmetric, group as about $\frac{1}{4}$. In order that any given substitutions may generate a group which is only a part of the $n!$ possible substitutions, very special relations are necessary, and it is highly improbable that arbitrarily chosen substitutions [...] should satisfy these conditions. The exception most likely to occur would be that all the given substitutions were severally equivalent to an even number of transpositions and would consequently generate the alternating group.

The problem stated in the passage can be recast more succinctly in modern language as follows.

Conjecture 1.1 (Netto, 1892) *The probability P that a pair of elements of Σ_n, chosen uniformly at random, generate either A_n or Σ_n tends to 1 as $n \to \infty$.*

Netto's conjecture remained open, perhaps due to obscurity, until Dixon [Dix69] picked up the problem and settled it in the affirmative.

Theorem 1.2 (Dixon, 1969) *Netto's conjecture holds, i.e., $P(A_n) \to 1$ as $n \to \infty$.*

In fact, Dixon does more than merely settling the conjecture, by providing explicit estimates for the rate of convergence of the probability to 1. Dixon's proof relies on a classic theorem of Jordan (see [Isa08, Theorem 8.23]).

Theorem 1.3 (Jordan) *Suppose that H is a subgroup of Σ_n which acts primitively on the set $\{1, \dots, n\}$ and contains a p-cycle for some prime $p \leqslant n-3$. Then $H = A_n$ or Σ_n.*

[3] Reproduced from the second edition, translated in english by Cole [Net64, p. 90].

The proof goes along the following lines: if $x, y \in G$ do not generate G, where $G = A_n$, then they both lie in a maximal subgroup M of G. Given M, the probability that this happens is $|G:M|^{-2}$. Hence

$$1 - P(G) \leqslant \sum_{M \lessdot G} |G:M|^{-2} = \sum_{M \in \mathcal{M}} |G:M|^{-1} \tag{1.1}$$

where $M \lessdot G$ means that M is a maximal subgroup of G and $\mathcal{M}$ is a set of representatives of the conjugacy classes of maximal subgroups of G.

A maximal subgroup of $G = A_n$ is either intransitive, imprimitive, or primitive. Denote the contributions to the sum in (1.1) from each category by S_1, S_2, S_3 respectively, so that $1 - P(G) \leqslant S_1 + S_2 + S_3$. One can prove with elementary methods that the contribution afforded by subgroups from intransitive or imprimitive subgroups is negligible, hence $S_1 + S_2 \to 0$ as $n \to \infty$. The case of primitive subgroups is less straightforward. This is dealt with in Lemma 3 of Dixon's paper (borrowing ideas from Erdős and Turán's second paper [ET67] on the statistics of the symmetric group–more on this in Appendix A), thus concluding the proof.

Lemma 1.4 *Let p_n be the probability that a permutation in Σ_n, chosen uniformly at random, has one of its powers equal to a p-cycle for some prime $p \leqslant n-3$. Then $p_n \to 1$ as $n \to \infty$.*

In most cases, probabilistic results, once established, are subsequently sharpened by others as is often the case in mathematics. Probabilistic group theory is particularly amenable to this practice because probabilistic statements are ultimately phrased in terms of inequalities, which, in turn, seem to admit inexorable improvement[4]. Dixon's theorem (in particular the bounds that Dixon obtains) is no exception and the reader is referred to [Dix02] for a brief account of those improvements.

Dixon went on to conjecture that A_n is not special with regard to 2-generation.

Conjecture 1.5 (Dixon, 1969) *Let G be a finite simple group and let P be the probability that two elements of G, chosen uniformly at random, generate G. Then $P(G) \to 1$ as $|G| \to \infty$.*

[4] That is, until they can provably be no longer improved.

Dixon's conjecture, much like Netto's conjecture, would have to wait quite a bit to be resolved. In 1990 Kantor and Lubotzky proved the assertion for classical groups and certain exceptional groups. The remaining cases were proved by Liebeck and Shalev in 1995.

Theorem 1.6 (KLLS, 1990–1995) *Dixon's conjecture holds, i.e.,* $P(G) \to 1$ *as* $|G| \to \infty$, *G a finite simple group.*

We mention that variants of Dixon's Conjecture have been investigated extensively in recent years, the more famous of these being the $(2,3)$-generation problem and its connections with the modular group $\mathrm{PSL}_2(\mathbb{Z})$. Liebeck's recent exposition [DFO13, Chapter 1] offers a detailed account of the field of probabilistic generation.

1.2 The commuting probability

The commuting probability of a finite group G, which we denote by $\mathrm{cp}(G)$ (but it has known other names), is the probability that two elements of G, chosen uniformly at random, commute:

$$\mathrm{cp}(G) := \frac{\left|\{(x,y) \in G \times G : xy = yx\}\right|}{|G|^2}.$$

It was popularised by Gustafson [Gus73] who, in turn, traces its origins to Erdős and Turán's series of papers on the statistics of the symmetric group. Using the class equation of G one can easily prove that $\mathrm{cp}(G) = k(G)\big/|G|$, where $k(G)$ is the number of conjugacy classes (of elements) of G. Gustafson also established the gap result that if G is nonabelian then $\mathrm{cp}(G) \leqslant 5/8$, with equality if and only $G\big/Z(G) \cong C_2 \times C_2$. Of course, G is abelian if and only if $\mathrm{cp}(G) = 1$ and, despite the $(5/8, 1)$ gap, the commuting probability can be viewed as an arithmetic quantification of the "abelianness" of a group.

Most research on the commuting probability has focussed on obtaining inequalities for $\mathrm{cp}(G)$ in terms of other invariants of the group and proving that if $\mathrm{cp}(G) > \varepsilon$ then G is in some sense close to abelian. For example, Rusin [Rus79] proved that if $\mathrm{cp}(G) > 11/32$ then either $|G'| \in \{2,3\}$, or $\mathrm{cp}(G) \leqslant 7/16$ and $|G : Z(G)| \leqslant 16$. Lescot et al. [LNY14] recently established that if $\mathrm{cp}(G) > 5/16$, then either G is supersoluble, or G or its abelianisation is isoclinic to A_4. We recall that two groups

G_1 and G_2 with centres Z_1, Z_2 respectively are said to be *isoclinic* if there are isomorphisms $\phi : G_1/Z_1 \to G_2/Z_2$, $\psi : G_1' \to G_2'$ such that for all $x, y \in G_1$

$$\psi([xZ_1, yZ_1]) = [\phi(xZ_1), \phi(yZ_1)].$$

Isoclinism is an equivalence relation on the set of all finite groups and each equivalence class contains a so-called *stem group*, i.e., a group G such that $Z(G) \leqslant G'$. In fact, cp is invariant under isoclinism, that is, if G_1 is isoclinic to G_2 then $\mathrm{cp}(G_1) = \mathrm{cp}(G_2)$. A few other pleasing properties that cp enjoys are:

(i) If G_1, G_2 are groups then $\mathrm{cp}(G_1 \times G_2) = \mathrm{cp}(G_1) \cdot \mathrm{cp}(G_2)$.

(ii) If N is a normal subgroup of G then $\mathrm{cp}(G) \leqslant \mathrm{cp}(N) \cdot \mathrm{cp}(G/N)$.

(iii) If H is a subgroup of G then $\mathrm{cp}(H) \geqslant \mathrm{cp}(G)$.

The interested reader is referred to a paper of Guralnick and Robinson [GR06], which contains a wealth of information on the commuting probability and a multitude of novel results. The more striking of those is that $\mathrm{cp}(G) \to 0$ as either the index or the derived length of the Fitting subgroup of G tends to infinity. One can therefore deduce at once that $\mathrm{cp}(G) \to 0$ as $|G| \to \infty$, G a finite simple group.

Let us close this chapter and move on to subgroup permutability by mentioning recent advancements on Joseph's conjectures. In 1977 Keith Joseph, who had already completed his dissertation on the commuting probability 8 years earlier, made the following three conjectures about the set $\mathcal{P} = \{\mathrm{cp}(G) : G \text{ a finite group}\}$.

Conjecture 1.7 (Joseph [Jos77])

J_1. *All limit points of $\mathcal{P}$ are rational.*

J_2. *$\mathcal{P}$ is well ordered by $>$.*

J_3. *$\{0\} \cup \mathcal{P}$ is closed.*

Note that J_1 implies that $\mathcal{P}$ is nowhere dense in $(0, 1]$, and J_2 implies that for every $x \in \mathcal{P}$ there is an $\varepsilon > 0$ such that $(x - \varepsilon, x) \cap \mathcal{P} = \emptyset$. Until recently, the best partial result was due to Hegarty [Heg13], who proved that the first two of Joseph's Conjectures hold for the set $\mathcal{P} \cap (2/9, 1]$. Eberhard [Ebe14], a few months prior to the writing of this thesis, managed to prove both J_1 and J_2 in their generality.

2

Subgroup permutability

In this preliminary chapter we present definitions and results concerning the notion of permutability in finite groups. Permutability can be thought of as a weak form of normality and we shall attempt to provide justification for this claim. We should, however, mention early on that permutability can be a rather awkward property to study, perhaps because two subgroups can permute for a variety of reasons. Nevertheless, this notion has been investigated extensively over the years and a number of strong results have been obtained. Since the number of items in the literature that deal with subgroup permutability is rather large, we shall attempt to give only a brief overview of the more interesting of those results here, and refer the reader to the textbook of Ballester-Bolinches et al. [BBERA10] which explores subgroup permutability (among others) in depth. Our exposition has also benefited from a survey paper of Robinson [Rob99], as well as Chapter 2 of Schmidt's book on subgroup lattices [Sch94], especially as far as modular subgroups and Iwasawa groups are concerned.

2.1 Permutable subgroups and Iwasawa groups

Consider a finite group G and subgroups H,K of G. The *Frobenius product* of H,K is defined to be the set $HK := \{hk : h \in H, k \in K\}$. In general HK is not a subgroup and it is an elementary exercise to show that HK is a subgroup if and only if $HK = KH$, in which case we say that H and K permute and write $H \operatorname{per} K$. This prompts the following definition.

Definition 2.1 *A subgroup H of the finite group G is called a permutable (or quasi-normal) subgroup if H permutes with every subgroup of G.*

Øystein Ore, in an influential early paper [Ore39], was the first to study permutable subgroups. He established that permutable subgroups of finite groups are subnormal, thus justifying the term 'quasi-normal'.

Theorem 2.2 (Ore [Ore39]) *If H is a permutable subgroup of the finite group G then H is subnormal in G.*

One sees effortlessly that either $K \leqslant N_G(H)$, or $H \leqslant N_G(K)$ implies that the subgroups H, K of G permute and thus normal subgroups are always permutable. On the other hand, if H permutes with K it need not follow that either normalises the other. Consider a Sylow 5–subgroup, say P, of A_5, the alternating group on 5 letters. Then $A_5 = PA_4 = A_4P$, but clearly neither P nor A_4 is normal in A_5. Recall that a *Dedekind group* is one all of whose subgroups are normal. Richard Dedekind, as early as 1897, studied the finite groups with this property in an attempt to find the algebraic number fields all of whose subfields are normal. Abelian groups are clearly Dedekind groups, so only the nonabelian ones require description. He named the nonabelian ones *Hamiltonian* after Sir William Rowan Hamilton who discovered the quaternions.

Theorem 2.3 (Dedekind [Ded97]) *Every subgroup of the finite nonabelian group G is normal in G if and only if G is a direct product of the quaternion group Q_8 of order 8, a group of exponent at most 2, and an odd-order abelian group.*

Thus Dedekind groups are primary examples of groups which have lots of permutability built in. We mention in passing that the Hamiltonian 2-groups admit the following equivalent characterisation.

Theorem 2.4 ([BBCLER13, Theorem 2.3]) *A finite 2-group G is Hamiltonian if and only if the Frattini subgroup $\Phi(G)$ of G has order 2, the centre $Z(G)$ of G has exponent 2 and index 4 in G, and the preimages of the generators of $G/Z(G)$ under the natural epimorphism from G onto $G/Z(G)$ have order 4.*

Iwasawa [Iwa41] obtained the full description of finite groups with every subgroup permutable.

Definition 2.5 *A group G is called Iwasawa, or quasi-Dedekind, if all subgroups of G are permutable.*

From Ore's theorem, Iwasawa groups are nilpotent. We remind the reader that a group G is called *modular* if its subgroup lattice is modular, that is, if $\langle H, K \cap L\rangle = \langle H, K\rangle \cap L$ for all subgroups H, K, L of G such that $H \leqslant L$.

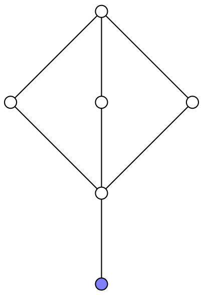


Figure 2.1: Hasse diagram of Q_8.

Theorem 2.6 (Iwasawa [Iwa41]) *Every subgroup of a finite group G is permutable if and only if G is a nilpotent modular group. The finite modular p-groups which are not Dedekind groups are the groups of type $G = \langle t, N\rangle$, where $N \lhd G$, N is abelian, and $n^t = n^{1+p^s}$, with $s > 1$ if $p = 2$.*

Thus one has the containments

$$\text{abelian} \subsetneq \text{Dedekind} \subsetneq \text{Iwasawa} \leftrightarrow \text{nilpotent modular} \subsetneq \text{nilpotent.}$$

The containments are proper: Q_8 is Dedekind but not abelian, the extraspecial group of order 27 and exponent 9 is Iwasawa but not Dedekind, and D_8 is nilpotent but not Iwasawa. In fact, one need only look at sections of order p^3 to decide whether a p-group is Iwasawa.

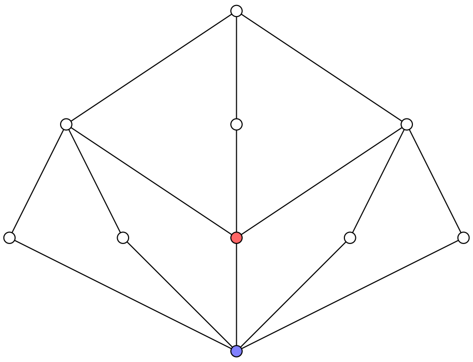


Figure 2.2: Hasse diagram of D_8.

Theorem 2.7 (Napolitani [Nap70, Teorema A]) *Let G be a finite p-group. Then G is Iwasawa if and only if each of its sections of order p^3 is Iwasawa.*

There are five groups of order 8: three of those are abelian and the other two are D_8 and Q_8. Among those, clearly, only D_8 is not Iwasawa. When $p > 2$, there are, again, only five groups of order p^3. The three abelian groups are Iwasawa. But the extraspecial p-group of 'minus' type is Iwasawa because

$$p_-^{1+2} = \left\langle x, y : x^{p^2} = y^p = 1, x^y = x^{1+p} \right\rangle \cong C_{p^2} \rtimes C_p;$$

the action in the semidirect product can be read off from the presentation and agrees with the condition in Iwasawa's theorem. The extraspecial group of 'plus' type

$$p_+^{1+2} = \langle x, y, z : x^p = y^p = z^p = 1, [x,z] = 1, [y,z] = 1, [x,y] = z \rangle,$$

however, is not Iwasawa. For instance, the subgroups $\langle x \rangle$, $\langle y \rangle$ do not permute. For if $\langle x \rangle\langle y \rangle = \langle y \rangle\langle x \rangle$, there would exist integers a, b such that $x^{-1}y^{-1} = y^a x^b$, thus $x^{-1} = y^{a+1}(y^{-1}xy)^b = y^{a+1}(xz)^b = y^{a+1}z^b x^b$, since z commutes with x. Hence $x^{-1-b} = y^{a+1}z^b$. Suppose that $b \neq -1$. Then there exists an $r \in \mathbb{Z}$ such that $r(-1-b) \equiv 1 (\mathrm{mod}\ p)$, and so $x = y^{r(a+1)}z^{rb}$. Substituting this expression for x in $[x,y] = z$ yields $z = 1$, a contradiction. We conclude that $b = -1$, i.e., $y^{a+1} = z$. For the same reason we must have $a \neq -1$, or z would be the trivial element. So there exists an $s \in \mathbb{Z}$ such that $y = z^{s(a+1)}$, forcing y to be central, thus $z = [x,y] = 1$, again a contradiction. We may therefore state the following corollary to Theorem 2.7.

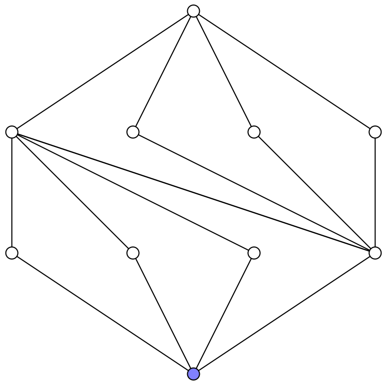


Figure 2.3: Hasse diagram of 3_-^{1+2}.

Corollary 2.8 *If the finite group G is not Iwasawa then there exist subgroups H, K of G, $K \lhd H$, such that either $H/K \cong D_8$, the dihedral group of order 8, or $H/K \cong p_+^{1+2}$, the extraspecial group of order p^3 and exponent p, $p > 2$.*

The following theorem is an interesting result of Longobardi.

Theorem 2.9 (Longobardi [Lon82, Proposizione 1.6]) *Let p a prime and let G be a p-group which is not Iwasawa, all of whose proper factors are Iwasawa. Then G has a unique minimal normal subgroup.*

Remark 2.10 *The dihedral group D_8 is not Iwasawa, but all its proper sections are abelian, hence Iwasawa. The unique minimal normal subgroup of D_8 in Figure 2.2 is displayed in red.*

The definition of a permutable subgroup $H \leqslant G$ requires that H permutes with every subgroup of G. We may, however, require that H permutes only with members of some interesting family of subgroups.

Definition 2.11 *Let G be a finite group and let Θ be a subset of $\mathfrak{s}(G)$. Then H is called a Θ-permutable subgroup, if H permutes with every $K \in \Theta$.*

The most interesting choice is $\Theta = \mathrm{Syl}(G)$, the set of all Sylow subgroups of G, and in that case we call H a Sylow-permutable, or S-permutable, subgroup of G.

The proof of the S-permutable cases in the first half of the following theorem is due to Schmid [Sch98]; for permutability (but not S-permutability) see [Sch94, p. 202].

Theorem 2.12 *Let G be a group.*

(i) *If N is a normal subgroup of G and H/N is an* (S-)*permutable subgroup of G/N then H is an* (S-)*permutable subgroup of G.*

(ii) *If X is a subgroup of G and H is an* (S-)*permutable subgroup of G then $H \cap X$ is an* (S-)*permutable subgroup of X.*

(iii) (Deskins [Des63], Kegel [Keg62]) *If H is an* S-*permutable subgroup of G then the quotient $\left\langle H^G \right\rangle \big/ H_G$ is nilpotent. In particular, H/H_G is contained in the Fitting subgroup $\mathrm{F}(G/H_G)$ of G/H_G.*

(iv) (Kegel [Keg62]) *If H is an* S-*permutable subgroup of G then H is subnormal in G.*

2.2 The subgroup permutability graph

A simple, yet efficient, method to study groups is via graphs. We can associate vertices with objects of interest (e.g. generators, prime divisors of the order of the group, conjugacy classes of noncentral elements, etc.) and require that an edge joins two vertices if and only if a certain property holds between the objects corresponding to the two vertices. The Cayley graph, for example, and the prime graph of a group are well-known instances of this construction. While it is not clear to what extent group theory benefits from purely graph-theoretical methods once a translation of this sort is made, it should be clear that such a translation provides, at the very least, a more transparent framework and often a way to visualise results obtained algebraically.

Let $\mathfrak{X}(G)$ be the set of nonnormal subgroups of a finite group G. Bianchi et al. [BGBMV95] introduced the subgroup permutability graph $\Gamma(G)$ of the finite group G as the graph whose vertex set is $\mathfrak{X}(G)$ and two vertices are joined by an edge if and only if the corresponding subgroups permute. As is often the case with graphs constructed from groups, typical questions that arise address issues of connectivity of the graph, the number of its connected components, and bounds for its diameter. Recall that the diameter of a graph is the maximum of the distances between any two vertices. The authors call a (finite) group *irreducible* if $\Gamma(G)$ is connected, i.e., if there is a path between any two vertices, and *reducible* otherwise. They denote by $\Delta(G)$ the maximum of the diameters of the connected components. The authors' key findings (Theorems A–D) are summarised in the following theorem.

Theorem 2.13 ([BGBMV95]) *Let G be a finite group.*

(i) *If G is soluble then $\Delta(G) \leqslant 4$.*

(ii) *If G is nonabelian simple then it is irreducible and $\Delta(G) \leqslant 16$.*

(iii) *If G is reducible then it is soluble.*

(iv) *If G is insoluble then $\Delta(G) \leqslant \max\{\Delta(S)\} + 4$, where S ranges over all nonabelian finite simple groups.*

The group $\mathcal{G} = \left\langle x, y, z : x^3 = y^3 = z^9 = [y,z] = 1, [z,x] = y, [y,x] = z^6 \right\rangle$ has order 3^4, its subgroup permutability graph is connected, and $\Delta(\mathcal{G}) = 4$, thus $\mathcal{G}$ attains the upper bound of Theorem 2.13 (i).

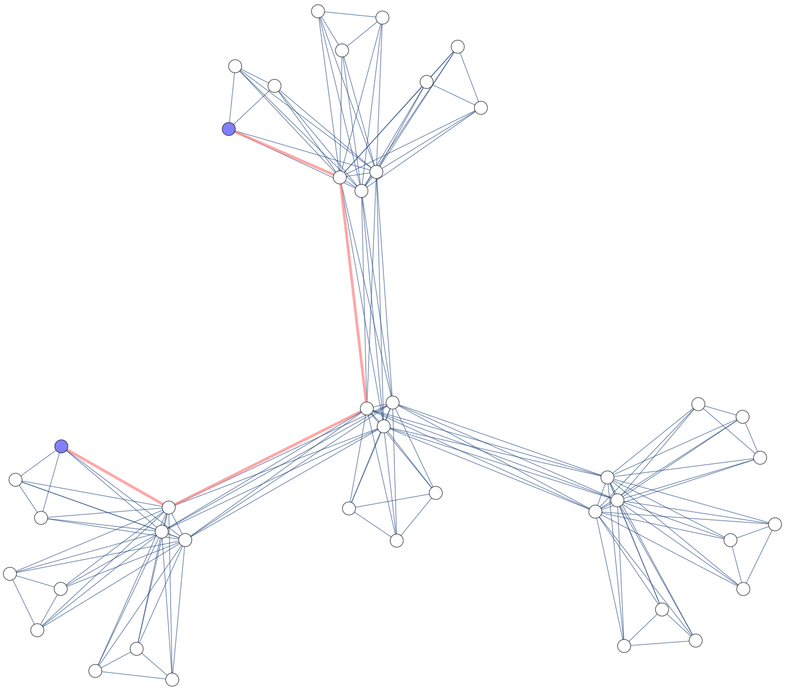


Figure 2.4: The subgroup permutability graph of $\mathcal{G}$ (drawn without loops). The vertices in blue have distance 4.

2.3 An arithmetic measure of subgroup permutability

Recently Tărnăuceanu [Tăr09] introduced the concept of subgroup permutability degree as the probability that two subgroups of G permute:

$$\mathfrak{p}(G) := \frac{|\{(H,K) \in \mathfrak{s}(G) \times \mathfrak{s}(G) : HK = KH\}|}{|\mathfrak{s}(G)|^2} = \frac{1}{|\mathfrak{s}(G)|^2} \sum_{H \leqslant G} |\mathrm{Per}(H)|,$$

where $\mathrm{Per}(H) := \{K \leqslant G : HK = KH\}$ and $\mathfrak{s}(G)$ is the set of subgroups of G. Thus $\mathfrak{p}$ provides us with an arithmetic measure of how close G is to being Iwasawa and we ask what structural information for G can be deduced from knowledge of $\mathfrak{p}(G)$. As explained earlier, a finite group G satisfies $\mathfrak{p}(G) = 1$ if and only if G is Iwasawa; equivalently, if and only if G is nilpotent modular. We can ask what happens if either of the two conditions is dropped.

Nilpotency of a finite group alone cannot be related to its subgroup permutability degree in any meaningful way. Consider the families of groups $\{C_{2^{n-3}} \times Q_8\}_{n \geqslant 5}$

and $\{D_{2^n}\}_{n\geqslant 5}$. In both cases the groups are nilpotent non-modular of the same order, but

$$\lim_{n\to\infty} \mathfrak{p}\left(C_{2^{n-3}} \times Q_8\right) = 1 \neq 0 = \lim_{n\to\infty} \mathfrak{p}\left(D_{2^n}\right).$$

Indeed, in this case the groups lie at the opposite extremes of the range of values of $\mathfrak{p}$, asymptotically speaking.

The modular non-nilpotent case admits a similar answer. Denote by $r_n := p_1 p_2 \dots p_n$ the product of the first n primes and consider the families $\left\{C_{r_n/r_2} \times \Sigma_3\right\}_{n\geqslant 2}$, and $\left\{C_{r_n/2p_n} \times D_{2p_n}\right\}_{n\geqslant 2}$, where Σ_3 denotes the symmetric group on 3 letters. Both families consist of groups that are modular non-nilpotent of the same order, but

$$\lim_{n\to\infty} \mathfrak{p}\left(C_{r_n/r_2} \times \Sigma_3\right) = 5/6 \neq 0 = \lim_{n\to\infty} \mathfrak{p}\left(C_{r_n/2p_n} \times D_{2p_n}\right).$$

We offer a proof of both claims at the end of this section.

One might wonder if $\mathfrak{p}$ shares any of the nice ‘structural’ properties of cp, i.e., whether any of the following holds:

(i) If G_1, G_2 are groups then $\mathfrak{p}(G_1 \times G_2) = \mathfrak{p}(G_1) \cdot \mathfrak{p}(G_2)$.

(ii) If N is a normal subgroup of G then $\mathfrak{p}(G) \leqslant \mathfrak{p}(N) \cdot \mathfrak{p}(G/N)$.

(iii) If H is a subgroup of G then $\mathfrak{p}(H) \geqslant \mathfrak{p}(G)$.

However, counterexamples exist even among groups of small order,[5] which lends evidence to the claim that $\mathfrak{p}$ is much more difficult to handle in general than cp. For instance, the dicyclic group of order 12 with presentation

$$\mathrm{Dic}_3 = \left\langle a, x : a^6 = 1, x^2 = a^3, x^{-1}ax = a^{-1}\right\rangle,$$

has $\mathfrak{p}\left(\mathrm{Dic}_3\right) = 29/32$ and a normal subgroup of order 2 with quotient isomorphic to Σ_3, but $\mathfrak{p}\left(\Sigma_3\right) = 5/6$, thus (ii) is not true. Then the group of order 16 which is a central product of D_8 and C_4 over a common cyclic central subgroup of order 2 has subgroup permutability degree equal to 505/529, but $\mathfrak{p}\left(D_8\right) = 23/25$, thereby disproving (iii). Finally, $\mathfrak{p}\left(C_2 \times \Sigma_3\right) = \mathfrak{p}\left(D_{12}\right) = 101/128$, but $\mathfrak{p}\left(\Sigma_3\right) = 5/6$, hence (i) does not hold either.

[5] But note that (i) holds in general if G_1, G_2 have coprime orders. See the proof of Theorem 2.15 for an explanation of this exception to the general rule.

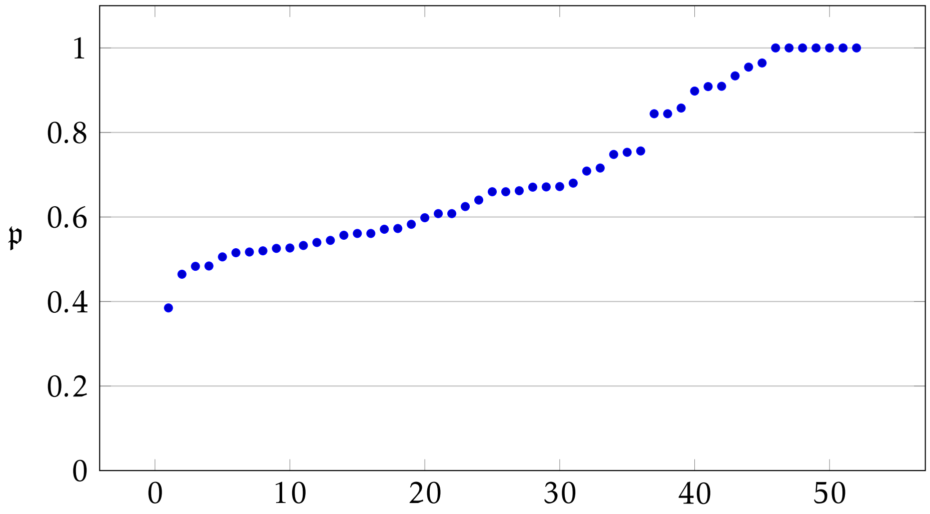


Figure 2.5: A plot of the subgroup permutability degree of all groups of order 80. The x-axis lists the groups in order of increasing $\mathfrak{p}$.

We digress briefly to discuss the case of equilibrated groups, where the computation of the subgroup permutability degree is more straightforward. Blackburn et al. examined the finite groups G with the property that if $H, K \leqslant G$ permute then either $K \leqslant N_G(H)$, or $H \leqslant N_G(K)$. They call these groups *equilibrated* or E-groups. Their findings are summarised in the following theorem.

Theorem 2.14 ([BDM96]) *Let G be a finite E-group.*

(i) *If G is nonabelian simple then* $G \cong \mathrm{PSL}_2(p)$ *where p is a prime such that* $p \equiv 5 (\mathrm{mod}\ 8)$, $p^2 \equiv -1 (\mathrm{mod}\ 5)$, $p+1$ *is twice a prime power, and* $p-1$ *is four times a prime power.*

(ii) *If G is soluble then G is an extension of a p-group by a Dedekind group.*

(iii) *If G is a p-group of odd order then one of the following holds:*

  (a) *G is Iwasawa and* $[x,y,y] = 1$ *for all* $x, y \in G$.

  (b) $G = EN$, *where* $E = \Omega_1(G)$, $|E| = p^3$, $G' = E'$, *N is cyclic, and* $E \cap N = G'$.

  (c) *G is a group of order* p^4 *and class* 3.

  (d) *G is a* 3*-group of maximal class.*

The authors show that every 2-group of maximal class is an E-group, but obtain little further information about 2-groups.

The case of equilibrated 2-groups has attracted further research from Silberberg [Sil06]. The author considers the case of finite 2-generated 2-groups and proves that if such a group is nonabelian and nonmetacyclic then it is not equilibrated (Corollary 3.1). He then determines exactly which nonabelian metacyclic 2-groups are equilibrated (Theorem 4.1). It is shown that if a finite 2-generated 2-group is equilibrated then it is either of class at most two, or of co-class at most two. The latter condition is sufficient for nonabelian metacyclic 2-groups (Proposition 4.4).

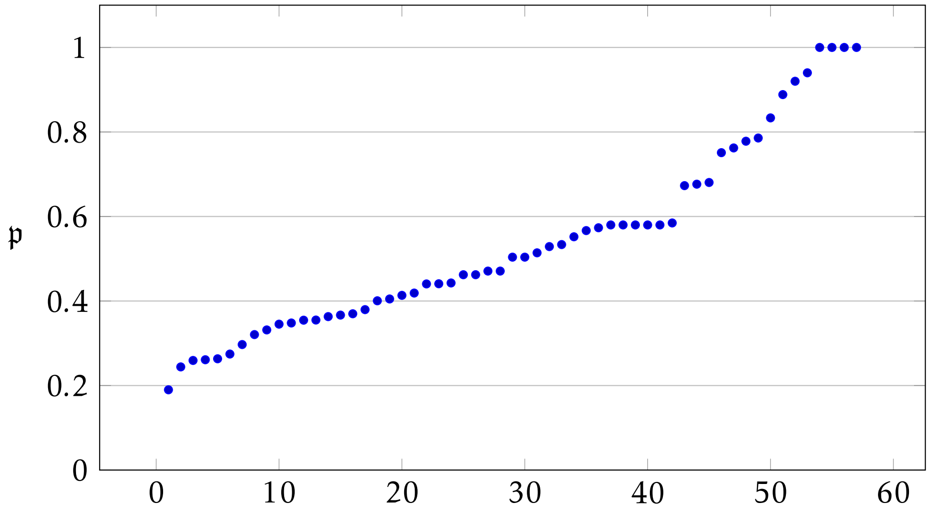


Figure 2.6: A plot of the subgroup permutability degree of all groups of order 168. The x-axis lists the groups in order of increasing $\mathfrak{p}$.

We close this section with a proof of the two claims made earlier.

Theorem 2.15 *Let $r_n := p_1 p_2 \dots p_n$ be the product of the first n primes.*

(i) *Consider the families of groups $\{C_{2^{n-3}} \times Q_8\}_{n \geqslant 5}$ and $\{D_{2^n}\}_{n \geqslant 5}$. In both cases the groups are nilpotent non-modular of the same order, but*

$$\lim_{n\to\infty} \mathfrak{p}(C_{2^{n-3}} \times Q_8) = 1 \neq 0 = \lim_{n\to\infty} \mathfrak{p}(D_{2^n}).$$

(ii) *The families of groups $\left\{C_{r_n/r_2} \times \Sigma_3\right\}_{n \geqslant 2}$, and $\left\{C_{r_n/2p_n} \times D_{2p_n}\right\}_{n \geqslant 2}$, where Σ_3 denotes the symmetric group on 3 letters, both consist of groups that are modular non-nilpotent of the same order, but*

$$\lim_{n\to\infty} \mathfrak{p}(C_{r_n/r_2} \times \Sigma_3) = 5/6 \neq 0 = \lim_{n\to\infty} \mathfrak{p}(C_{r_n/2p_n} \times D_{2p_n}).$$

Proof. (i) First, both families consist of 2-groups, which are nilpotent. Recall that a p-group G is modular if and only if it is Iwasawa, i.e., $\mathfrak{p}(G) = 1$ [Sch94, Lemma 2.3.2], and we shall see in due course that in both cases $\mathfrak{p}(G) < 1$, thereby establishing the desired property of non-modularity. The vanishing of $\mathfrak{p}(D_{2^n})$ follows from [Tăr09, Corollary 3.1.4.], and the explicit formula for the subgroup permutability degree of D_{2^n} given earlier [Tăr09, Corollary 3.1.3.] proves that, indeed, $\mathfrak{p}(D_{2^n}) < 1$ for all $n \geqslant 3$.

We shall make use of Goursat's lemma [Gou89] to find the subgroups of $G_n = C_{2^n} \times Q_8$. An excellent account (and some generalisations) of this useful result is given in [AC09]. Briefly, given subgroups $Q \lhd R \leqslant G_1$, $S \lhd T \leqslant G_2$ and an isomorphism $f : R/Q \to T/S$, $H = \{(a,b) \in R \times T : f(aQ) = bS\}$ is a subgroup of $G_1 \times G_2$ and each subgroup of $G_1 \times G_2$ is of this form. Then sections of Q_8 must necessarily be cyclic: there are 6 sections of order 1, 7 sections of order 2, and 3 (cyclic) sections of order 4. In C_{2^n} there are $n+1$ sections of order 1, n sections of order 2 and $n-1$ sections of order 4. There is a unique isomorphism between cyclic groups of order 1 or 2, but there are two isomorphisms between cyclic groups of order 4 (the identity and the inverting one). So in total we have

$$|\mathfrak{s}(G_n)| = 6(n+1) + 7n + 2 * 3(n-1) = 19n.$$

A necessary and sufficient condition for H to be normal in $G_1 \times G_2$ is that both Q, R and S, T are normal in G_1, G_2 respectively, and that $Q/R \leqslant Z(G_1/R)$, $T/S \leqslant Z(G_2/S)$. The first condition is automatically satisfied because both C_{2^n} and Q_8 are Dedekind groups. The second condition is satisfied by all sections of C_{2^n} and all sections of Q_8 of order 1 or 2 (in the latter case because the unique minimal subgroup of order 2 of Q_8 is its centre and all nontrivial quotients of Q_8 are abelian), but not the sections of order 4 since $\langle i\rangle/\langle 1\rangle$, $\langle j\rangle/\langle 1\rangle$, $\langle k\rangle/\langle 1\rangle$ are not subgroups of $Z(Q_8/\langle 1\rangle) = Z(Q_8) = \langle -1\rangle$. All symbols i, j, k, -1 come from the standard presentation of Q_8.

Therefore, the $6(n+1)+7n = 13n+6$ subgroups corresponding to sections of order 1 or 2 are all normal in G_n, hence permutable. It thus suffices to examine which pairs of subgroups corresponding to sections of order 4 do not permute with each other. Let $\langle c\rangle = C_{2^n}$ and denote by $C(\alpha,\beta)$ the coset $c^{\alpha 2^{n-\beta-2}} C_{2^\beta} \in C_{2^{\beta+2}}/C_{2^\beta} \leqslant C_{2^n}/C_{2^\beta}$.

Then the $6(n-1)$ subgroups under consideration fall into 6 'classes', which we call $\mathcal{H}_1$, $\mathcal{H}_2$, $\mathcal{H}_3$, $\mathcal{J}_1$, $\mathcal{J}_2$, $\mathcal{J}_3$ respectively:

$$\begin{aligned}H_{1,l} &= \{(a,i^x) : 0 \leqslant x \leqslant 3, a \in C(x,l)\},\\ H_{2,m} &= \{(a,j^y) : 0 \leqslant y \leqslant 3, a \in C(y,m)\},\\ H_{3,r} &= \{(a,k^z) : 0 \leqslant z \leqslant 3, a \in C(z,r)\},\end{aligned}$$

for $l,m,r \in \{0,1,...,n-2\}$, and

$$\begin{aligned}J_{1,s} &= \{(a,i^{-u}) : 0 \leqslant u \leqslant 3, a \in C(u,s)\},\\ J_{2,t} &= \{(a,j^{-v}) : 0 \leqslant v \leqslant 3, a \in C(v,t)\},\\ J_{3,d} &= \{(a,k^{-w}) : 0 \leqslant w \leqslant 3, a \in C(w,d)\},\end{aligned}$$

for $s,t,d \in \{0,1,...,n-2\}$.

First, a few observations. Subgroups which belong to the same class clearly permute, since they centralise each other, and the same is true for pairs of subgroups in the classes $\mathcal{H}_i, \mathcal{J}_i$, $0 \leqslant i \leqslant 3$. It should also be clear that the symmetry in the defining relations among i,j,k allows us to consider only pairs of subgroups in $\mathcal{H}_1, \mathcal{H}_2$, since the other cases are identical. Lastly, note that elements in any subgroup of the $\mathcal{H}, \mathcal{J}$ classes with second coordinate either 1 or $i^2 = j^2 = k^2 = -1$ commute with every other element of $C_{2^n} \times Q_8$.

Let $l,m \in \{0,1,\ldots,n-2\}$. We shall show that, unless $l = m$, the subgroup $H_{1,l}$ permutes with $H_{2,m}$. Let $h_1 = \left(c^{\alpha 2^{n-l-2}+\lambda 2^{n-l}}, i^\alpha\right) \in H_{1,l}$, $h_2 = \left(c^{\beta 2^{n-m-2}+\mu 2^{n-m}}, j^\beta\right) \in H_{2,m}$, where $\alpha,\beta \in \{1,3\}$, $\lambda \in \{0,1,\ldots,2^{n-l}-1\}$, $\mu \in \{0,1,\ldots,2^{n-m}-1\}$. Then

$$h_1 h_2 = \left(c^{\alpha 2^{n-l-2}+\lambda 2^{n-l}+\beta 2^{n-m-2}+\mu 2^{n-m}}, i^\alpha j^\beta\right),$$

and we ask if there exist elements $h_2' \in H_{2,m}$ and $h_1' \in H_{1,l}$ such that $h_1 h_2 = h_2' h_1'$. Write $h_1' = \left(c^{\gamma 2^{n-l-2}+\nu 2^{n-l}}, i^\gamma\right)$ and $h_2' = \left(c^{\delta 2^{n-m-2}+\xi 2^{n-m}}, j^\delta\right)$. The question then is whether, given α,β,λ,μ, there exist γ,δ,ν,ξ such that

$$(\alpha-\gamma)2^{n-l-2} + (\lambda-\nu)2^{n-l} + (\beta-\delta)2^{n-m-2} + (\mu-\xi)2^{n-m} \equiv 0 (\text{mod } 2^n), \qquad (2.1)$$

and $i^\alpha j^\beta = j^\delta i^\gamma$ in Q_8. For simplicity, assume that $l > m$.

1. If $(\alpha,\beta) = (1,1)$ then $(\gamma,\delta) = (1,3)$ and $(\nu,\xi) \underset{2^n}{\equiv} (\lambda - 2^{l-m-1}, \mu - 2^m)$,

2. if $(\alpha,\beta) = (1,3)$ then $(\gamma,\delta) = (1,1)$ and $(\nu,\xi) \underset{2^n}{\equiv} (\lambda - 2^{l-m-1}, \mu + 1)$,

3. if $(\alpha,\beta) = (3,1)$ then $(\gamma,\delta) = (3,3)$ and $(\nu,\xi) \underset{2^n}{\equiv} (\lambda - 2^{l-m-1}, \mu)$,

4. if $(\alpha,\beta) = (3,3)$ then $(\gamma,\delta) = (3,1)$ and $(\nu,\xi) \underset{2^n}{\equiv} (\lambda - 2^{l-m-1}, \mu + 1)$

are all solutions to (2.1). The case $l < m$ is similar and solutions always exist. On the other hand, if $l = m$ there is no solution since 2^{m+1} (or 2^{l+1}) is required to divide an odd integer.

Taking into account that all other cases between the classes $\mathcal{H}, \mathcal{J}$ are symmetric, i.e. identical, we deduce at once that the number of pairs of subgroups of G_n which do not permute is $6 * 4(n-1) = 24(n-1)$. Therefore

$$\mathfrak{p}(G_n) = 1 - \frac{24(n-1)}{361n^2},$$

so $\lim_{n\to\infty} \mathfrak{p}(G_n) = 1$; in particular, if $n \geqslant 2$ then $\mathfrak{p}(G_n) < 1$.

(ii) Clearly Σ_3 and D_{2p_n} are not nilpotent, so the two families consist of non-nilpotent groups. Since the orders of the two factors in each case are coprime, it follows (either as a direct consequence of Goursat's lemma itself, or simply by [AC09, Theorem 2]) that the subgroups of both $C_{r_n/r_2} \times \Sigma_3$ and $C_{r_n/2p_n} \times D_{2p_n}$ are subproducts, that is, direct products of subgroups of each factor. So $\mathfrak{p}$ is 'multiplicative' in that case[6], i.e.,

$$\mathfrak{p}(C_{r_n/r_2} \times \Sigma_3) = \mathfrak{p}(C_{r_n/r_2}) \cdot \mathfrak{p}(\Sigma_3) = \mathfrak{p}(\Sigma_3) = 5/6,$$

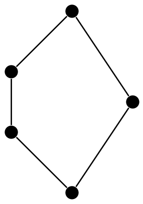
and

$$\mathfrak{p}(C_{r_n/2p_n} \times D_{2p_n}) = \mathfrak{p}(C_{r_n/2p_n}) \cdot \mathfrak{p}(D_{2p_n}) = \mathfrak{p}(D_{2p_n}).$$

Now, if p is an odd prime then $\mathfrak{p}(D_{2p}) = 7p + 9\big/(p+3)^2$. This is easy to see from first principles: D_{2p} has $\tau(p) + \sigma(p) = p + 3$ subgroups, of which only $1, C_p, D_{2p}$ are normal. Each of the p subgroups generated by involutions permutes only with itself and the three normal subgroups, while the three normal subgroups are permutable, bringing the count to $4p + 3(p+3) = 7p + 9$. One may also use Corollary 3.1.2. of [Tăr09] and do the algebra. We conclude that $\mathfrak{p}(C_{r_n/2p_n} \times D_{2p_n})$ vanishes as $n \to \infty$.

[6] This property of $\mathfrak{p}$ appears explicitly as Proposition 2.2. in [Tăr09], but no proof is given there.

So we need only justify why $C_{r_n/r_2} \times \Sigma_3$ and $C_{r_n/2p_n} \times D_{2p_n}$ are modular groups. We shall prove first that D_{2n} is modular if and only if n is a prime.



N_5

Figure 2.7: Hasse diagram of N_5, the smallest non-modular lattice.

Recall that a subgroup lattice is modular if and only if it does not contain the forbidden pentagon N_5 as a sublattice [Sch94, Theorem 2.1.2]. Viewing the subgroup lattice of D_{2n} as an undirected graph, there are always two independent paths from 1 to D_{2n}; one through the cyclic C_n and one through a reflection. Clearly each has length at least 2 and moreover the length is 2 in both cases if and only if n is prime. So, unless n is a prime number, D_{2n} contains a cycle of length at least 5, hence N_5 as a minor, i.e., sublattice. On the other hand, if n is prime then the circumference of the subgroup lattice as a graph is (at most) 4, hence D_{2n} does not contain N_5 as a sublattice. [7]

Now proving that $C \times D$ is modular, given that $\gcd(|C|, |D|) = 1$ and both C, D are modular, is an easy exercise. The proof is complete. □

2.4 A criterion for the vanishing of the subgroup permutability degree

Let us now focus on the criterion for the vanishing of the subgroup permutability degree that we mentioned earlier. In general, working with the definition of $\mathfrak{p}$ seems difficult-there is usually little or no insight when two randomly chosen subgroups of a group permute. Even if one were only to consider groups for which subgroup permutability is reduced to a more manageable property (e.g. E-groups), one should still be able to say something useful about the behaviour of the various sums that would ultimately appear in the resulting expression for $\mathfrak{p}$.

[7] One may also prove that if D_{2n} is modular then n must be prime by observing that dihedral groups are generated by two (distinct) involutions and appealing to [Sch94, Lemma 2.2.4].

One should therefore ask if perhaps 'most' subgroups of the group in question are of a particular type. The simplest case arises when p-subgroups dominate the subgroup lattice for some prime p dividing the order of the group and when, in addition, the Sylow p-subgroups intersect trivially. In this case it suffices to only check permutability between subgroups of the same Sylow p-subgroup. The following lemma makes this precise.

Lemma 2.16 *Let $\{G_n\}_{n\geqslant 1}$ be a family of finite groups such that $p \mid |G_n|$ for some fixed prime p and for all $n \in \mathbb{N}$, satisfying the conditions*

(i) *the Sylow p-subgroups of G_n intersect trivially for all $n \in \mathbb{N}$,*

(ii) $\lim\limits_{n\to\infty} \left|\mathrm{Syl}_p(G_n)\right| = \infty$, *and*

(iii) $\lim\limits_{n\to\infty} \dfrac{|\mathcal{E}_n|}{|\mathfrak{s}(G_n)|} = 1$,

where

$$\mathcal{E}_n := \left\{H \leqslant G_n : |H| = p^k \text{ for some } k \in \mathbb{N}\right\} = \bigcup_{P\in \mathrm{Syl}_p(G_n)} \mathfrak{s}(P).$$

Then $\lim\limits_{n\to\infty} \mathfrak{p}(G_n) = 0.$

Proof. Define the map $f : \mathfrak{s}(G_n) \times \mathfrak{s}(G_n) \to \{0,1\}$ via the rule

$$\left(H_i, H_j\right) \mapsto \begin{cases} 1, & \text{if } H_iH_j = H_jH_i, \\ 0, & \text{otherwise,} \end{cases}$$

and observe that f is symmetric in its arguments. Thus

$$\begin{aligned} \sum_{H\leqslant G_n} |\mathrm{Per}(H)| &= \sum_{X_i,X_j\in\mathcal{E}_n} f\left(X_i,X_j\right) + 2\sum_{\substack{X_i\in\mathcal{E}_n \\ Y_j\in\mathcal{E}_n^c}} f\left(X_i,Y_j\right) + \sum_{Y_i,Y_j\in\mathcal{E}_n^c} f\left(Y_i,Y_j\right) \\ &\leqslant \sum_{X_i,X_j\in\mathcal{E}_n} f\left(X_i,X_j\right) + 2\sum_{\substack{X_i\in\mathcal{E}_n \\ Y_j\in\mathcal{E}_n^c}} 1 + \sum_{Y_i,Y_j\in\mathcal{E}_n^c} 1 \\ &= \sum_{X_i,X_j\in\mathcal{E}_n} f\left(X_i,X_j\right) + 2\,|\mathcal{E}_n|\,|\mathcal{E}_n^c| + |\mathcal{E}_n^c|^2 \\ &= \sum_{X_i,X_j\in\mathcal{E}_n} f\left(X_i,X_j\right) + |\mathfrak{s}(G_n)|^2 - |\mathcal{E}_n|^2. \end{aligned}$$

Divide by $|\mathfrak{s}(G_n)|^2$ both sides to deduce that

$$\mathfrak{p}(G_n) \leqslant 1 - \frac{|\mathcal{E}_n|^2}{|\mathfrak{s}(G_n)|^2} + \frac{\sum_{X_i,X_j \in \mathcal{E}_n} f\left(X_i, X_j\right)}{|\mathfrak{s}(G_n)|^2}. \tag{2.2}$$

Now let $X_i, X_j \in \mathcal{E}_n$. We claim that if X_iX_j is a subgroup of G_n then both X_i, X_j belong to the same Sylow p-subgroup. To see this, let $P \in \mathrm{Syl}_p(G_n)$. Then there exist elements g_i, g_j of G_n such that $X_i \leqslant P^{g_i}$ and $X_j \leqslant P^{g_j}$. Since X_i, X_j are p-groups, so is X_iX_j. Hence there exists an element $g_k \in G_n$ such that $X_iX_j \leqslant P^{g_k}$. Notice that $X_i \leqslant P^{g_i}$ and $X_i \leqslant X_iX_j \leqslant P^{g_k}$. Thus $P^{g_i} \cap P^{g_k} \geqslant X_i > 1$. Since distinct Sylow p-subgroups of G_n intersect trivially, we deduce that $P^{g_i} = P^{g_k}$. Similarly $P^{g_j} \cap P^{g_k} \geqslant X_j > 1$ and this forces $P^{g_j} = P^{g_k}$ for the same reason. We conclude that $P^{g_i} = P^{g_j}$, thus both X_i and X_j are subgroups of the same Sylow p-subgroup, as required. Now let $\mathrm{Syl}_p(G_n) = \left\{P^{g_i} : 0 \leqslant i \leqslant \left|\mathrm{Syl}_p(G_n)\right|\right\}$. By dint of the above observation we may thus write

$$\begin{aligned}
\sum_{X_i,X_j \in \mathcal{E}_n} f\left(X_i, X_j\right) &= \sum_{k=1}^{\left|\mathrm{Syl}_p(G_n)\right|} \sum_{X_i,X_j \in P^{g_k}} f\left(X_i, X_j\right) \\
&\leqslant \sum_{k=1}^{\left|\mathrm{Syl}_p(G_n)\right|} \sum_{X_i,X_j \in P^{g_k}} 1 \\
&= \sum_{k=1}^{\left|\mathrm{Syl}_p(G_n)\right|} (|\mathfrak{s}\,(P^{g_k})| - 1)^2 \\
&= \left|\mathrm{Syl}_p(G_n)\right| (|\mathfrak{s}(P)| - 1)^2 .
\end{aligned}$$

On the other hand we have $|\mathcal{E}_n|^2 = \left|\mathrm{Syl}_p(G_n)\right|^2 (|\mathfrak{s}(P)| - 1)^2$. Hence

$$\begin{aligned}
0 \leqslant \frac{\sum_{X_i,X_j \in \mathcal{E}_n} f\left(X_i, X_j\right)}{|\mathfrak{s}(G_n)|^2} &\leqslant \frac{\sum_{X_i,X_j \in \mathcal{E}_n} f\left(X_i, X_j\right)}{|\mathcal{E}_n|^2} \\
&\leqslant \frac{\left|\mathrm{Syl}_p(G_n)\right| (|\mathfrak{s}(P)| - 1)^2}{\left|\mathrm{Syl}_p(G_n)\right|^2 (|\mathfrak{s}(P)| - 1)^2} \\
&= \frac{1}{\left|\mathrm{Syl}_p(G_n)\right|},
\end{aligned}$$

from which we see that

$$\lim_{n\to\infty} \frac{\sum\limits_{X_i, X_j \in \mathcal{E}_n} f\left(X_i, X_j\right)}{|\mathfrak{s}(G_n)|^2} = 0,$$

since $\lim\limits_{n\to\infty} \left|\mathrm{Syl}_p(G_n)\right| = \infty$, thus $\lim\limits_{n\to\infty} \left|\mathrm{Syl}_p(G_n)\right|^{-1} = 0$. Also

$$\lim_{n\to\infty} \frac{|\mathcal{E}_n|^2}{|\mathfrak{s}(G_n)|^2} = 1,$$

since $\lim\limits_{n\to\infty} \dfrac{|\mathcal{E}_n|}{|\mathfrak{s}(G_n)|} = 1$, by hypothesis. Taking limits in (2.2) yields

$$0 \leqslant \lim_{n\to\infty} \mathfrak{p}(G_n) \leqslant \lim_{n\to\infty} \left(1 - \frac{|\mathcal{E}_n|^2}{|\mathfrak{s}(G_n)|^2} + \frac{\sum\limits_{X_i, X_j \in \mathcal{E}_n} f\left(X_i, X_j\right)}{|\mathfrak{s}(G_n)|^2} \right) = 0,$$

thus concluding the proof. □

3

The subgroup permutability degree of $\mathrm{PSL}_2(2^n)$

The main result of this chapter is the following.

Theorem 3.1 *The subgroup permutability degree of* $\mathrm{PSL}_2(2^n)$ *vanishes asymptotically, i.e.,*

$$\lim_{n\to\infty} \mathfrak{p}(\mathrm{PSL}_2(2^n)) = 0.$$

The following is standard notation from number theory and asymptotic analysis. Let $n \in \mathbb{N}$. Then

(i) $\tau(n)$ is the number of divisors of n.

(ii) $\sigma(n)$ is the sum of divisors of n.

(iii) $\omega(n)$ is the number of distinct prime divisors of n.

(iv) For the sequences $\{f_n\}$, $\{g_n\}$, $g_n \neq 0$, write $f_n \sim g_n$ if $\lim\limits_{n\to\infty} \frac{f_n}{g_n} = 1$.

3.1 The subgroup structure of $\mathrm{PSL}_2(2^n)$

In order to compute the subgroup permutability degree of a group, knowledge of its subgroup structure is essential. It is usually hard to find all the subgroups of a nonabelian finite group G, but in the case of projective special linear groups Dickson gave in [Dic03] a complete list of the subgroups of $\mathrm{PSL}_2(q)$, here adapted to the case $q = 2^n$.

(i) a single conjugacy class of $q+1$ elementary abelian subgroups of order q.

(ii) a single conjugacy class of $q(q+1)/2$ cyclic subgroups of order d for each divisor $d \neq 1$ of $q-1$.

(iii) a single conjugacy class of $q(q-1)/2$ cyclic subgroups of order d for each divisor $d \neq 1$ of $q+1$.

(iv) a single conjugacy class of $q(q^2-1)/2d$ dihedral subgroups of order $2d$ for each divisor $d \neq 1$ of $q-1$.

(v) a single conjugacy class of $q(q^2-1)/2d$ dihedral subgroups of order $2d$ for each divisor $d \neq 1$ of $q+1$.

(vi) a number of conjugacy classes of abelian subgroups of order r for each divisor $r \neq 1$ of q.

(vii) a number of conjugacy classes of subgroups of order rd for each divisor r of q and for d depending on r, all lying inside a group of order $q(q-1)$ isomorphic to $\mathrm{AGL}_1(2^n)$.

(viii) a single conjugacy class of $q(q^2-1)/r(r^2-1)$ subgroups $\mathrm{PSL}_2(r)$, where q is a power of r.

By looking at the list of subgroups of $\mathrm{PSL}_2(2^n)$ we are able to deduce the following.

Corollary 3.2 *Let $n = p_1^{a_1} \cdots p_m^{a_m}$ be the prime factorisation of n, $G := \mathrm{PSL}_2(2^n)$ and $q = 2^n$. Then the maximal subgroups of G of the same order form single conjugacy classes. A set of representatives is*

$$\mathrm{Max}_{\mathrm{c}}(G) = \left\{ D_{2(q-1)}, D_{2(q+1)}, \mathrm{AGL}_1(q), \mathrm{PSL}_2(2^{n/p_i}) : 1 \leqslant i \leqslant m \right\}.$$

In particular, G possesses $\omega(n)+3$ conjugacy classes of maximal subgroups.

At this point we take some time to discuss certain properties and related results concerning a sequence that will play an important role in the proof of Theorem 3.1.

3.2 The magnitude of $G_{n,q}$ and auxiliary results

Let q be a prime power. Then the total number of linear subspaces of a finite-dimensional vector space $\mathbb{F}_q^n$ of degree n over the finite field with q elements $\mathbb{F}_q$ is

$$G_{n,q} := \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q ,$$

where $\left[{n \atop k}\right]_q$ denotes the number of k-dimensional subspaces of $\mathbb{F}_q^n$. Goldman and Rota introduced $G_{n,q}$ as the Galois numbers in [GR69], where they also proved the recurrence

$$G_{0,q} = 1, G_{1,q} = 2 \text{ and } G_{n+1,q} = 2G_{n,q} + (q^n - 1)G_{n-1,q} \tag{3.1}$$

using the symbolic method. The purpose of this section is to establish that $G_{n,q}$ is sub-exponential in q^n and obtain auxiliary results that will come in handy later on. For simplicity we will write G_n instead of $G_{n,2}$ and likewise for other doubly indexed sequences when needed.

Theorem 3.3 *Let $\{G_{n,q}\}$ be the sequence defined in (3.1). Then*

$$q^{\frac{n^2}{4}} < G_{n,q} < q^{\frac{n^2}{4}+4},$$

for all $n \geqslant 2$.

Proof. We use induction to prove both inequalities. For the first one, the claim certainly holds for $n = 2$, since $G_{2,q} = q + 3 > q$. Notice that $G_{n+1,q} - G_{n,q} = G_{n,q} + (q^n - 1)G_{n-1,q} > 0$, for all $n \in \mathbb{N}$. Hence

$$G_{n+1,q} > q^n G_{n-1,q} > q^n q^{(n-1)^2/4} = q^{(n+1)^2/4}.$$

Now define the following auxiliary sequence: $c_{0,q} = 1$, $c_{1,q} = q^{\frac{3}{4}}$ and

$$c_{n+1,q} = c_{n-1,q} + \lambda_{n,q} c_{n,q},$$

where $\lambda_{n,q} := q^{1-\frac{n}{2}}$. We work towards a proof that for fixed q, $\{c_{n,q}\}$ is bounded. By induction

$$c_{n+1,q} = c_{n-1,q} + \lambda_{n,q} c_{n,q} \leqslant c_{n-1} + \lambda_n c_n = c_{n+1}.$$

Thus it suffices to establish boundedness for $q = 2$. First we show that $c_n < (5/4)^n$ for all $n \geqslant 12$. For $n \in \{12, 13\}$ the claim holds, since we can compute $c_n < 14 <$

$(5/4)^n$. Assume $n \geqslant 14$ and that the claim is valid for all values less than or equal to n. Then

$$\begin{aligned} c_{n+1} &= c_{n-1} + 2^{1-\frac{n}{2}} c_n \\ &\leqslant (5/4)^{n-1} + 2^{1-\frac{n}{2}} (5/4)^n \\ &= (5/4)^{n-1} \left(1 + 5 \cdot 2^{-1-\frac{n}{2}}\right). \end{aligned}$$

However, $1 + 2^{-1-\frac{n}{2}} \cdot 5 < 1 + \frac{9}{16} = (5/4)^2$. Hence $c_{n+1} < (5/4)^{n+1}$ and the induction is complete. We deduce that

$$c_{n+1} < c_{n-1} + 2\gamma^n$$

for all $n \geqslant 12$, where $\gamma = \frac{5}{4\sqrt{2}} < 1$. Hence, for n even, we get

$$\begin{aligned} c_n = c_{12} + \sum_{j=6}^{\frac{n-2}{2}} (c_{2j+2} - c_{2j}) &< c_{12} + 2 \sum_{j=6}^{\frac{n-2}{2}} \gamma^{2j+1} \\ &< c_{12} + 2 \sum_{j=6}^{\infty} \gamma^{2j+1} \\ &= c_{12} + \frac{2\gamma^{13}}{1-\gamma^2}, \end{aligned}$$

while for n odd we have

$$\begin{aligned} c_n = c_{11} + \sum_{j=6}^{\frac{n-1}{2}} (c_{2j+1} - c_{2j-1}) &< c_{11} + 2 \sum_{j=6}^{\frac{n-1}{2}} \gamma^{2j} \\ &< c_{11} + 2 \sum_{j=6}^{\infty} \gamma^{2j} \\ &= c_{11} + \frac{2\gamma^{12}}{1-\gamma^2}. \end{aligned}$$

Thus for all $n \geqslant 12$ we have

$$c_n < \max\left\{ c_{12} + \frac{2\gamma^{13}}{1-\gamma^2}, c_{11} + \frac{2\gamma^{12}}{1-\gamma^2} \right\} < 2^4$$

and, in fact, this inequality also holds for $n \leqslant 11$. Now let us show that $G_n \leqslant c_n 2^{\frac{n^2}{4}}$. For $n \in \{0,1\}$ we have equality, hence we may assume that $n \geqslant 2$. We have

$$G_{n+1} = 2G_n + (2^n - 1)G_{n-1} < c_n 2^{1+\frac{n^2}{4}} + c_{n-1} 2^{\frac{(n-1)^2}{4}+n},$$

from the induction hypothesis. We write the right-hand-side as

$$2^{\frac{(n+1)^2}{4}} \left(c_{n-1} + 2^{-\frac{2n-3}{4}} c_n\right) < 2^{\frac{(n+1)^2}{4}} \left(c_{n-1} + 2^{1-\frac{n}{2}} c_n\right) = c_{n+1} 2^{\frac{(n+1)^2}{4}},$$

and this completes the induction. □

Corollary 3.4 *Let $n \in \mathbb{N}$ and let m be a nontrivial divisor of n. Then*

$$G_{\frac{n}{m}+1,2^m} \leqslant G_{\lceil \frac{n}{2} \rceil+1,4}$$

Proof. For $m = 2$ we have equality, hence we may assume that $m \geqslant 3$. If $m = 3$ then we need to show that $G_{\frac{n}{3}+1,8} \leqslant G_{\lceil \frac{n}{2} \rceil+1,4}$ for all n a multiple of 3. But $G_{\frac{n}{3}+1,8} < 8^{\frac{(\frac{n}{3}+1)^2}{4}+4} = 2^{\frac{(n+3)^2}{12}+12}$ and $G_{\lceil \frac{n}{2} \rceil+1,4} > 4^{\frac{(\lceil \frac{n}{2} \rceil+1)^2}{4}} \geqslant 4^{\frac{(\frac{n}{2}+1)^2}{4}} = 2^{\frac{(n+2)^2}{8}}$, by Theorem 3.3. Notice that $\frac{(n+2)^2}{8} \geqslant \frac{(n+3)^2}{12} + 12$ for all $n \geqslant 18$ and in fact $G_{\frac{n}{3}+1,8} \leqslant G_{\lceil \frac{n}{2} \rceil+1,4}$ holds for all $n \in \mathbb{N}$ by a direct computation. Now assume $m \geqslant 4$. By Theorem 3.3 again, we conclude that $G_{\frac{n}{m}+1,2^m} < (2^m)^{\frac{(\frac{n}{m}+1)^2}{4}+4} = 2^{\frac{(n+m)^2}{4m}+4m}$ and $G_{\lceil \frac{n}{2} \rceil+1,4} > 2^{\frac{(n+2)^2}{8}}$. Now $\frac{(n+2)^2}{8} \geqslant \frac{(n+m)^2}{4m} + 4m$ is equivalent to $n^2 \geqslant \frac{2m(17m-2)}{m-2}$ and the right-hand side is strictly increasing for $m \geqslant 4$. But $n^2 \geqslant \frac{2n(17n-2)}{n-2}$ for all $n \geqslant 36$, hence the claim holds for all n in that range. A direct computation yields the validity of the claim for all $n \in \mathbb{N}$. □

From this point on we restrict our attention to the case $q = 2$, although the same arguments work for arbitrary q. The next corollary is a direct consequence of the fact that both G_n and $\frac{G_n}{G_{\lfloor \frac{n}{2} \rfloor}}$ are sub-exponential in $q = 2^n$ as was shown in Theorem 3.3, while the right-hand-side in both cases is polynomial in q. Hence we feel justified in omitting the details of the proof.

Corollary 3.5 *For all $n \geqslant 16$ the following inequalities hold*

(i) $G_n - 1 > (n+2)q(q^2-1) + 1$.

(ii) $\frac{G_n}{G_{\lfloor \frac{n}{2} \rfloor}} > \frac{1}{8} \cdot \frac{q(q^2-1)(q-3)}{q^{\lfloor \frac{n}{2} \rfloor}(q^{\lfloor \frac{n}{2} \rfloor}-3)} \omega(n)$.

Corollary 3.6 *For all $n \geqslant 5$ the sequence $b_n := \frac{G_n-1}{2^n(2^n-3)}$ is strictly increasing.*

Proof. We have

$$b_{n+1} - b_n = \frac{(2^n-3)(G_{n+1}-1) - (2^{n+2}-6)(G_n-1)}{2^n(2^{n+2}-6)(2^n-3)}.$$

Hence $b_{n+1} - b_n > 0$ if and only if $\frac{G_{n+1}-1}{G_n-1} > \frac{2^{n+2}-6}{2^n-3}$. But

$$\frac{G_{n+1}-1}{G_n-1} > \frac{G_{n+1}}{G_n} > 2^{\frac{(n+1)^2}{4}-\frac{n^2}{4}-4} = 2^{\frac{2n-15}{4}} > 5 > \frac{2^{n+2}-6}{2^n-3},$$

for all $n \geqslant 13$, where the second inequality follows from Theorem 3.3. For $n \in \{5, \ldots, 12\}$ the inequality holds by inspection. □

3.3 Bounding the number of subgroups of $\mathrm{PSL}_2(2^n)$

The purpose of this section is to obtain an upper bound for the number of subgroups of $\mathrm{PSL}_2(2^n)$. We know from Corollary 3.2 that the maximal subgroups of $\mathrm{PSL}_2(2^n)$ are either dihedral, $\mathrm{AGL}_1(2^n)$ or type PSL defined over the maximal subfields of $\mathbb{F}_{2^n}$. Cavior obtained a formula for the number of subgroups of dihedral groups in [Cav75]. Hence we need only examine the number of subgroups of $\mathrm{AGL}_1(2^n)$ and then use it to obtain a (necessarily recursive) bound in terms of $\mathfrak{h}$, where $\mathfrak{h}(G)$ is the function defined as the quotient of the total number of nontrivial subgroups of G to the order of the group. The effort culminates in Theorem 3.10, which will then be used in the proof of Theorem 3.1.

Theorem 3.7 *The number of subgroups of $\mathrm{AGL}_1(2^n)$ is given by*

$$|\mathfrak{s}(\mathrm{AGL}_1(2^n))| = G_n + \sum_{\substack{m|n \\ m>1}} a_m \left(G_{\frac{n}{m}+1,2^m} - G_{\frac{n}{m},2^m} \right),$$

where

$$a_m := \#\left\{ d : d \mid 2^m - 1, d \nmid 2^k - 1, k \mid m, k < m \right\}.$$

Proof. Let $C = \mathbb{F}(2^n, \times)$ and $E = \mathbb{F}(2^n, +)$ be the multiplicative and additive groups of the field $\mathbb{F}_{2^n}$ respectively. Note that $\mathrm{AGL}_1(2^n)$ is by definition the semi-direct product $\mathbb{F}(2^n,+) \rtimes \mathbb{F}(2^n,\times) = E \rtimes C$, where C acts fixed-point-freely on E by right multiplication. As such it is a Frobenius group with Frobenius kernel E and Frobenius complement C. The number of subgroups that are wholly contained in the kernel is obviously G_n. On the other hand, any subgroup H not contained in the kernel must intersect a complement or one of its conjugates nontrivially. Hence it can be written as $H = V \rtimes B$, where $V \leqslant E$ and $B \leqslant C^g$ for some $g \in \mathrm{AGL}_1(2^n)$ by Proposition 4.1.8., p. 81 of [KS04]. Note that the normaliser of H in $\mathrm{AGL}_1(2^n)$ is just $V \rtimes C^g$, so that the number of conjugates of H is $\frac{|E|}{|V|}$. Now fix a nontrivial subgroup $B = \langle b \rangle$ of C and notice that b^2 is also a generator of B, since $d := |B|$ is odd. Then b^2 generates a certain subfield of $\mathbb{F}_{2^n}$. Specifically, it is the field of order 2^{m_d}, where m_d is minimal with the property that d divides $2^{m_d} - 1$. Denote by V_d the additive group of this subfield and notice that $V_d \rtimes B$ is a subgroup of $\mathrm{AGL}_1(2^n)$. In fact, because $\mathbb{F}_{2^n}$ is a vector space of dimension $\frac{n}{m_d}$ over $\mathbb{F}_{2^{m_d}}$ (see §70 of [Dic03]), the invariant subgroups of E on which B acts are in $1-1$ correspondence with the subspaces of $\mathbb{F}_{2^{m_d}}^{\frac{n}{m_d}}$. Taking into account the different conjugates, we deduce that there are

$$\sum_{i=0}^{\frac{n}{m_d}} \begin{bmatrix} \frac{n}{m_d} \\ i \end{bmatrix}_{2^{m_d}} 2^{n-im_d}$$

subgroups in $\mathrm{AGL}_1(2^n)$ for each nontrivial divisor d of $2^n - 1$. Hence

$$|\mathfrak{s}(\mathrm{AGL}_1(2^n))| = G_n + \sum_{\substack{d|2^n-1 \\ d>1}} \sum_{i=0}^{\frac{n}{m_d}} \begin{bmatrix} \frac{n}{m_d} \\ i \end{bmatrix}_{2^{m_d}} 2^{n-im_d}. \tag{3.2}$$

However,

$$\begin{bmatrix} \frac{n}{m_d} \\ i \end{bmatrix}_{2^{m_d}} 2^{n-im_d} = \begin{bmatrix} \frac{n}{m_d}+1 \\ i+1 \end{bmatrix}_{2^{m_d}} - \begin{bmatrix} \frac{n}{m_d}+1 \\ i \end{bmatrix}_{2^{m_d}}.$$

This equality is just the q-analogue of Pascal's rule for binomial coefficients (see for example [KC02, Proposition 6.1]). Therefore

$$\begin{aligned} \sum_{i=0}^{\frac{n}{m_d}} \begin{bmatrix} \frac{n}{m_d} \\ i \end{bmatrix}_{2^{m_d}} 2^{n-im_d} &= \sum_{i=0}^{\frac{n}{m_d}} \begin{bmatrix} \frac{n}{m_d}+1 \\ i+1 \end{bmatrix}_{2^{m_d}} - \sum_{i=0}^{\frac{n}{m_d}} \begin{bmatrix} \frac{n}{m_d}+1 \\ i \end{bmatrix}_{2^{m_d}} \\ &= G_{\frac{n}{m_d}+1,2^{m_d}} - G_{\frac{n}{m_d},2^{m_d}}. \end{aligned}$$

and (3.2) can be rewritten as

$$|\mathfrak{s}(\mathrm{AGL}_1(2^n))| = G_n + \sum_{\substack{d|2^n-1 \\ d>1}} \left(G_{\frac{n}{m_d}+1,2^{m_d}} - G_{\frac{n}{m_d},2^{m_d}} \right).$$

Finally set a_m as in the statement of the Theorem. Then

$$G_n + \sum_{\substack{d|2^n-1 \\ d>1}} \left(G_{\frac{n}{m_d}+1,2^{m_d}} - G_{\frac{n}{m_d},2^{m_d}} \right) = G_n + \sum_{\substack{m|n \\ m>1}} a_m \left(G_{\frac{n}{m}+1,2^m} - G_{\frac{n}{m},2^m} \right)$$

and this completes the proof. □

Corollary 3.8 *For all $n \geqslant 13$ we have*

$$G_n < \left|\mathfrak{s}(\mathrm{AGL}_1(2^n))\right| < \left(1 + \frac{1}{2^n+1}\right) G_n.$$

Proof. By Theorem 3.7 we have

$$|\mathfrak{s}(\mathrm{AGL}_1(2^n))| = G_n + \sum_{\substack{m|n \\ m>1}} a_m \left(G_{\frac{n}{m}+1,2^m} - G_{\frac{n}{m},2^m} \right).$$

Since $G_{\frac{n}{m}+1,2^m} - G_{\frac{n}{m},2^m} > 0$, the left-hand side inequality follows immediately. On the other hand $G_{\frac{n}{m}+1,2^m} - G_{\frac{n}{m},2^m} < G_{\frac{n}{m}+1,2^m}$ and $a_m \leqslant \tau(2^n-1)$ for all divisors m of n. Hence

$$|\mathfrak{s}(\mathrm{AGL}_1(2^n))| < G_n + (\tau(2^n-1)-1) \sum_{\substack{m|n \\ m>1}} G_{\frac{n}{m}+1,2^m}.$$

By Corollary 3.4 this reduces to

$$|\mathfrak{s}(\mathrm{AGL}_1(2^n))| < G_n + (\tau(2^n-1)-1)(\tau(n)-1)\, G_{\lceil\frac{n}{2}\rceil+1,4}.$$

Therefore

$$|\mathfrak{s}(\mathrm{AGL}_1(2^n))| < G_n + (\tau(2^n-1)-1)^2\, G_{\lceil\frac{n}{2}\rceil+1,4},$$

which in turn yields

$$|\mathfrak{s}(\mathrm{AGL}_1(2^n))| < G_n + 2^n G_{\lceil\frac{n}{2}\rceil+1,4}. \tag{3.3}$$

However, due to Theorem 3.3 we obtain

$$2^{2n+1}G_{\lceil\frac{n}{2}\rceil+1,4} < 2^{2n+1}4^{\frac{\left(\lceil\frac{n}{2}\rceil+1\right)^2}{4}+4} = 2^{\frac{\left(\lceil\frac{n}{2}\rceil+1\right)^2}{2}+2n+9} \leqslant 2^{\frac{\left(\frac{n}{2}+2\right)^2}{2}+2n+9},$$

hence $2^{2n+1}G_{\lceil\frac{n}{2}\rceil+1,4} < 2^{\frac{n^2+24n+88}{8}}$. But $\frac{n^2+24n+88}{8} < \frac{n^2}{4}$ if and only if $n^2-24n-88>0$, which holds for all $n \geqslant 28$. We deduce that for $n \geqslant 28$

$$2^n(2^n+1)G_{\lceil\frac{n}{2}\rceil+1,4} < 2^{2n+1}G_{\lceil\frac{n}{2}\rceil+1,4} < 2^{\frac{n^2}{4}} < G_n.$$

In fact, $(2^n+1)(\tau(2^n-1)-1)^2\,G_{\lceil\frac{n}{2}\rceil+1,4} < G_n$ holds for all $13 \leqslant n \leqslant 27$ by inspection. Hence (3.3) becomes

$$G_n < \left|\mathfrak{s}(\mathrm{AGL}_1(2^n))\right| < \left(1+\frac{1}{2^n+1}\right)G_n, \quad \text{for all } n \geqslant 13,$$

and this completes the proof. □

In particular, Corollary 3.8 shows that $|\mathfrak{s}(\mathrm{AGL}_1(2^n))| \sim G_n$. We have obtained appropriate bounds for the number of subgroups of $\mathrm{AGL}_1(2^n)$, hence we will now try to give a recursive bound for the number of subgroups of $\mathrm{PSL}_2(2^n)$. We begin with the following easy lemma.

Lemma 3.9 *Let G be a finite simple group with the property that its maximal subgroups of the same order lie in single conjugacy classes. Then*

$$\mathfrak{h}(G) \leqslant \sum_{i=1}^{k} \mathfrak{h}(M_i),$$

where M_i are representatives for each conjugacy class of maximal subgroups.

Proof. Observe that maximal subgroups of simple groups are self-normalising and denote by $|\mathrm{Max}(G)|$ their totality in G. Then $\mathfrak{s}(G) = \{G\}\bigcup\limits_{i=1}^{r}\mathfrak{s}(M_i)$, where $r =$

$\left|\mathrm{Max}(G)\right|$. Write $\mathrm{Max}(G)$ as a union of conjugacy classes with representatives $\{M_1,\ldots,M_k\}$ and sizes $\{r_1,\ldots,r_k\}$ respectively, so that $r = r_1 + \cdots + r_k$. Then

$$\begin{aligned}|\mathfrak{s}(G)| - 1 &\leqslant \sum_{i=1}^{k} |G : N_G(M_i)| \left(|\mathfrak{s}(M_i)| - 1\right) \\ &= \sum_{i=1}^{k} |G : M_i| \left(|\mathfrak{s}(M_i)| - 1\right).\end{aligned}$$

Hence

$$\mathfrak{h}(G) = \frac{\left|\mathfrak{s}(G)\right| - 1}{|G|} \leqslant \sum_{i=1}^{k} \frac{|\mathfrak{s}(M_i)| - 1}{|M_i|} = \sum_{i=1}^{k} \mathfrak{h}(M_i).$$

The proof is now complete. □

From [Cav75] we have

$$\mathfrak{h}(D_{2\lambda}) = \frac{\tau(\lambda) + \sigma(\lambda) - 1}{2\lambda}.$$

But $\tau(\lambda) + \sigma(\lambda) - 1 < \tau(\lambda) + \sigma(\lambda) < \lambda(2 + \log\lambda)$, since

$$\frac{\sigma(\lambda)}{\lambda} = \sum_{d|\lambda} \frac{1}{d} \leqslant \sum_{d \leqslant \lambda} \frac{1}{d} < 1 + \int_1^{\lambda} \frac{dt}{t} = 1 + \log\lambda, \text{ and } \tau(\lambda) \leqslant \lambda.$$

Thus $\mathfrak{h}(D_{2\lambda}) < 1 + \frac{\log\lambda}{2}$, which implies that

$$\mathfrak{h}\left(D_{2(q-1)}\right) + \mathfrak{h}\left(D_{2(q+1)}\right) < 2 + \frac{\log(q-1) + \log(q+1)}{2}.$$

However, from Jensen's inequality we have

$$\frac{\log(q-1) + \log(q+1)}{2} \leqslant \log\left(\frac{q - 1 + q + 1}{2}\right) = \log q < \log_2 q = n.$$

Therefore

$$\mathfrak{h}\left(D_{2(q-1)}\right) + \mathfrak{h}\left(D_{2(q+1)}\right) < n + 2.$$

On the other hand, we have from Corollary 3.8 that

$$\mathfrak{h}(\mathrm{AGL}_1(q)) < \frac{\left(1 + \frac{1}{q+1}\right)(G_n - 1) + \frac{1}{q+1}}{q(q-1)} = \frac{(q+2)(G_n - 1) + 1}{q(q^2 - 1)}.$$

According to Lemma 3.9 and Corollary 3.2 we have

$$\mathfrak{h}(\mathrm{PSL}_2(q)) = \mathfrak{h}\left(D_{2(q-1)}\right) + \mathfrak{h}\left(D_{2(q+1)}\right) + \mathfrak{h}(\mathrm{AGL}_1(q)) + \sum_{i=1}^{m} \mathfrak{h}\left(\mathrm{PSL}_2(2^{n/p_i})\right)$$

But

$$\mathfrak{h}\left(D_{2(q-1)}\right) + \mathfrak{h}\left(D_{2(q+1)}\right) + \mathfrak{h}(\mathrm{AGL}_1(q)) < n + 2 + \frac{(q+2)(G_n - 1) + 1}{q(q^2-1)},$$

and $(n+2)q(q^2-1)+1 < G_n - 1$ from Corollary 3.6. Thus

$$\mathfrak{h}(\mathrm{PSL}_2(q)) < \frac{(q+3)}{q(q^2-1)}(G_n - 1) + \sum_{i=1}^{m} \mathfrak{h}\left(\mathrm{PSL}_2(2^{n/p_i})\right). \tag{3.4}$$

We now arrive at the main result of this section.

Theorem 3.10 *For every positive integer* $n \geqslant 15$ *we have*

$$\mathfrak{h}(\mathrm{PSL}_2(2^n)) < \frac{G_n - 1}{q(q-3)}.$$

Proof. For $n = 15$ the inequality holds by inspection. Hence we may suppose that $n \geqslant 16$ and that the claim has already been established for all values less than n. By Corollary 3.5 we have

$$\frac{G_n - 1}{G_{\lfloor \frac{n}{2} \rfloor} - 1} > \frac{G_n}{G_{\lfloor \frac{n}{2} \rfloor}} > \frac{1}{8} \cdot \frac{q(q^2-1)(q-3)}{2^{\lfloor \frac{n}{2} \rfloor}(2^{\lfloor \frac{n}{2} \rfloor} - 3)} \omega(n),$$

hence

$$\omega(n) \frac{G_{\lfloor \frac{n}{2} \rfloor} - 1}{2^{\lfloor \frac{n}{2} \rfloor}(2^{\lfloor \frac{n}{2} \rfloor} - 3)} < \frac{8}{q(q^2-1)(q-3)}(G_n - 1). \tag{3.5}$$

If p is the smallest prime divisor of n then

$$\begin{aligned} \mathfrak{h}(\mathrm{PSL}_2(2^n)) &< \sum_{i=1}^{m} \mathfrak{h}\left(\mathrm{PSL}_2\left(2^{n/p_i}\right)\right) + \frac{q+3}{q(q^2-1)}(G_n - 1) \\ &< \omega(n) \frac{G_{\frac{n}{p}} - 1}{2^{\frac{n}{p}}(2^{\frac{n}{p}} - 3)} + \frac{(q+3)}{q(q^2-1)}(G_n - 1), \end{aligned}$$

where the first inequality is just (3.4) and the second inequality follows from the inductive hypothesis and Corollary 3.6 . Hence, from Corollary 3.6 again, we have

$$\mathfrak{h}\,(\mathrm{PSL}_2\,(2^n)) \leqslant \omega(n)\frac{G_{\lfloor \frac{n}{2}\rfloor}-1}{2^{\lfloor \frac{n}{2}\rfloor}(2^{\lfloor \frac{n}{2}\rfloor}-3)}+\frac{(q+3)}{q(q^2-1)}(G_n-1).$$

Therefore

$$\mathfrak{h}\,(\mathrm{PSL}_2\,(2^n)) \leqslant \frac{8}{q(q^2-1)(q-3)}(G_n-1)+\frac{(q+3)}{q(q^2-1)}(G_n-1),$$

by (3.5). Thus

$$\mathfrak{h}\,(\mathrm{PSL}_2\,(2^n)) < (G_n-1)\left[\frac{8}{q(q^2-1)(q-3)}+\frac{q+3}{q(q^2-1)}\right]=\frac{G_n-1}{q(q-3)},$$

completing the induction. □

We take a moment to appreciate the strength of the inequality just established. The totality of elementary abelian subgroups in $\mathrm{PSL}_2(2^n)$ is $(q+1)(G_n-1)$. Hence the proportion of elementary abelian subgroups to the total number of subgroups of $\mathrm{PSL}_2(2^n)$ is

$$\frac{(q+1)(G_n-1)}{|\mathfrak{s}(G)|}=\frac{(q+1)(G_n-1)}{\mathfrak{h}\,(\mathrm{PSL}_2\,(2^n))\,q(q^2-1)+1}.$$

This quantity is certainly less than 1. What Theorem 3.10 allows us to do is bound it from below

$$1>\frac{(q+1)(G_n-1)}{\mathfrak{h}\,(\mathrm{PSL}_2\,(2^n))\,q(q^2-1)+1}>\frac{(q+1)(q-3)(G_n-1)}{(q^2-1)(G_n-1)+q-3}>1-\frac{3}{q}.$$

Since

$$\lim_{n\to\infty}\frac{(q+1)(G_n-1)}{|\mathfrak{s}(G)|}=1,$$

Theorem 3.10 essentially says that almost all subgroups of $\mathrm{PSL}_2(2^n)$ are elementary abelian. The proof of Theorem 3.1 now follows from a straightforward application of Lemma 2.16.

4
The subgroup permutability degree of $\mathrm{Sz}(q)$

The main result of this chapter is the following.

Theorem 4.1 *The subgroup permutability degree of* $\mathrm{Sz}\left(2^{2n+1}\right)$ *vanishes asymptotically, i.e.,*

$$\lim_{n\to\infty} \mathfrak{p}\left(\mathrm{Sz}\left(2^{2n+1}\right)\right) = 0.$$

4.1 The subgroup structure of $\mathrm{Sz}(q)$

The discussion in this section follows closely that of Nouacer [Nou82] and Berkovich and Janko [BJ11, §105]. Let $\mathbb{F}_q$ be the finite field with $q := 2^{2n+1}$ elements and set $\theta := 2^{n+1}$. The map $\overline{\theta} : x \mapsto x^\theta$ is an automorphism of the field and, in fact, generates the cyclic group $\mathrm{Gal}\left(\mathbb{F}_q/\mathbb{F}_2\right)$. This is because $\left|\mathrm{Gal}\left(\mathbb{F}_q/\mathbb{F}_2\right)\right| = 2n+1$ and $\overline{\theta}$ acts as a “square root” of the Frobenius automorphism ϕ, that is, $x^{\theta^2} = x^2$ for all $x \in \mathbb{F}_q$, hence both $\overline{\theta}$ and ϕ have the same order in $\mathrm{Gal}\left(\mathbb{F}_q/\mathbb{F}_2\right)$.

Definition 4.2 *Suppose that* $a, b, \in \mathbb{F}_q$ *and* $\lambda \in \mathbb{F}_q^\times$. *Define* 4×4 *matrices over* $\mathbb{F}_q$ *by*

$$S(a,b) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ b & a^\theta & 1 & 0 \\ a^{2+\theta} + ab + b^\theta & a^{1+\theta} + b & a & 1 \end{pmatrix}, C(\lambda) := \begin{pmatrix} \lambda^{1+\frac{\theta}{2}} & 0 & 0 & 0 \\ 0 & \lambda^{\frac{\theta}{2}} & 0 & 0 \\ 0 & 0 & \lambda^{-\frac{\theta}{2}} & 0 \\ 0 & 0 & 0 & \lambda^{-1-\frac{\theta}{2}} \end{pmatrix},$$

and

$$T := \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

The Suzuki group $\mathrm{Sz}(q)$ *is defined to be the following subgroup of* $\mathrm{GL}_4(q)$

$$\mathrm{Sz}(q) := \left\langle S(a,b), C(\lambda), T : a,b \in \mathbb{F}_q, \lambda \in \mathbb{F}_q^\times \right\rangle.$$

In this notation, the set $P := \left\{S(a,b) : (a,b) \in \mathbb{F}_q^2\right\}$ is a Sylow 2-subgroup of $\mathrm{Sz}(q)$. In fact, $P \cong \left(\mathbb{F}_q^2, *\right)$, where $*$ is defined via the rule

$$(a_1,b_1) * (a_2,b_2) = (a_1 + a_2, b_1 + b_2 + a_1 a_2^\theta),$$

the implicit isomorphism being $S(a,b) \mapsto (a,b)$. This writing of P as a Cartesian product endowed with a “twisted” multiplication is particularly convenient, as it captures the essential information contained within each matrix while avoiding the cumbersome matrix notation. Now notice that $(0,0)$ is the identity element and $(a,b)^{-1} = (a, b + a^{1+\theta})$, hence

$$[(a_1,b_1),(a_2,b_2)] = \left(0, a_1 a_2^\theta + a_2 a_1^\theta\right). \tag{4.1}$$

If either $a_1 = 0$ or $a_2 = 0$ then $[(a_1,b_1),(a_2,b_2)] = (0,0)$. Moreover $(0,b_1)*(0,b_2) = (0,b_1+b_2)$ and $(0,b)^2 = (0,0)$, thus $\left\{(0,b) : b \in \mathbb{F}_q\right\} \leqslant Z$. In fact, equality occurs here. For suppose that $(a_1,b_1) \in Z$. Then $\left(0, a_1 a_2^\theta + a_2 a_1^\theta\right) = (0,0)$ for all $a_2 \in \mathbb{F}_q$, thus $a_1 a_2^\theta = a_2 a_1^\theta$, since $\mathrm{char}\mathbb{F}_q = 2$. Because $n \geqslant 1$, we may choose $a_2 \in \mathbb{F}_q \setminus \{0, a_1\}$. Therefore $\left(a_1 a_2^{-1}\right)^\theta = a_1 a_2^{-1}$, i.e., the element $a_1 a_2^{-1}$ is a fixed point of the automorphism $\overline{\theta}$. Since $\left\langle\overline{\theta}\right\rangle = \mathrm{Gal}\left(\mathbb{F}_q/\mathbb{F}_2\right)$, the fixed points of $\overline{\theta}$ are precisely the elements of the prime subfield $\mathbb{F}_2 = \{0,1\}$. Hence $a_1 a_2^{-1} = 0$, that is $a_1 = 0$. Thus $Z \leqslant \left\{(0,b) : b \in \mathbb{F}_q\right\}$, which establishes the claim. We deduce that the centre of P is an elementary abelian group, isomorphic to the additive group of the field. From (4.1) it is clear that $P' \leqslant Z$, since all commutators are central, hence P/Z is abelian. Moreover $(a,b)^2 = (0, a^{1+\theta}) \in Z$, thus all squares are central as well. In view of $|P/Z| = |Z|$, we infer that $P/Z \cong Z$. As all squares lie in the centre, clearly $\mho(P) \leqslant Z$ holds. Recall that for a p-group P the agemo subgroups of P are the series of subgroups: $\mho^i(G) = \langle g^{p^i} : g \in G\rangle$. When $i = 1$ and p is odd, then i is usually omitted from the definition. When p is even, an omitted i may mean either $i = 1$ or $i = 2$, depending on local convention. In this thesis, we use the convention that an omitted i always indicates $i = 1$. Now consider an arbitrary element $(0,b) \in Z$, and notice that the map $x \mapsto x^{1+\theta}$ is a bijection of the field $\mathbb{F}_q$, since

$$\gcd(q-1, 1+\theta) = \gcd\left(2^{2n+1}-1, 1+2^{n+1}\right) = 1. \tag{4.2}$$

Thus there exists a unique element $a_b \in \mathbb{F}_q$ such that $a_b^{1+\theta} = b$. Therefore $(0,b) = (a_b, b)^2 \in \mho(P)$, which proves that $\mho(P) = Z$. Also $\Phi(P) = \mho(P)P'$ when P is a p-group;[8] since $P' \leqslant Z$ so $\Phi(P) = Z$. Proving that P' and Z actually coincide is not difficult. The multiplicative group of the field is a subgroup of $\mathrm{Aut}(P)$ and acts (sharply) transitively on the nonidentity elements of Z, as we shall shortly see. Since P' is a characteristic subgroup of P, it is invariant under the action via automorphisms of $\mathbb{F}_q^\times$. The claim now follows from $P' \leqslant Z$, which we already know. In spite of the simple argument above, we offer an alternative proof that is essentially due to Isaacs. It is more direct and, if modified appropriately, works equally well in a more general setting.

Claim 4.3 *Let $P \in \mathrm{Syl}_2\left(\mathrm{Sz}\left(2^{2n+1}\right)\right)$, $n \geqslant 1$. Then $P' = Z$.*

Proof (Isaacs). It is sufficient to show that the subgroup of the additive group of $\mathbb{F}_q$ generated by the elements of the form $xy^\theta + x^\theta y$ is the whole group. Taking $x = 1$ and letting y vary over $\mathbb{F}_q$ gives all elements of the form $y^\theta + y$. This set is actually a subgroup since the map $y \mapsto y^\theta + y$ is an additive homomorphism. Furthermore, the kernel of this homomorphism is the prime subfield $\mathbb{F}_2$, and thus by taking $x = 1$, we get a subgroup of $\mathbb{F}_q$ of index 2. In fact, every member of this subgroup has trace zero, where the trace of an element $t \in \mathbb{F}_q$ is understood to be $\mathrm{Tr}(t) = \sum_{\sigma \in \langle \overline{\theta} \rangle} t^\sigma$. It is known that the trace map maps $\mathbb{F}_q$ onto the prime subfield, so the kernel of the trace is a subgroup of index 2. Thus taking $x = 1$ yields exactly the elements with trace zero. It suffices now to find x and y such that $xy^\theta + x^\theta y$ does not have trace zero. It will follow that the group generated by the elements of the form $xy^\theta + x^\theta y$ is the whole of $\mathbb{F}_q$. Now in general, $\mathrm{Tr}(t) = \mathrm{Tr}(t^\theta)$, so

$$\begin{aligned}
\mathrm{Tr}\left(xy^\theta + x^\theta y\right) &= \mathrm{Tr}\left(xy^\theta\right) + \mathrm{Tr}\left(x^\theta y\right) \\
&= \mathrm{Tr}\left(x^\theta y^{\theta^2}\right) + \mathrm{Tr}\left(x^\theta y\right) \\
&= \mathrm{Tr}\left(x^\theta\left(y^{\theta^2} + y\right)\right).
\end{aligned}$$

[8] See Rotman [Rot95, Theorem 5.48].

Since $q \geqslant 8$, θ^2 is not the identity automorphism, so choose y so that $y^{\theta^2} + y \neq 0$ and write c to denote this nonzero element. It suffices now to find x such that $\mathrm{Tr}\left(cx^{\theta}\right) \neq 0$. As x varies over $\mathbb{F}_q$, the element cx^{θ} runs over all of $\mathbb{F}_q$, so for some value of x we get an element with nonzero trace. This completes the proof. □

Notice that a subgroup $H \leqslant P$ either contained in Z, or containing Z is normal in P. The first assertion is clear, while the second assertion follows from $h^g = [g,h]h$ being an element of H for all $g \in P$, as $[g,h] \in P' = Z$. We collect what we have established so far.

The group P is a special 2-group of exponent 4 and class 2, with the property that $P/Z \cong Z$.

Remark 4.4 The Sylow 2-subgroups of Sz(q) arise as special cases in Higman's more general theory of so-called Suzuki 2-groups,[9] i.e., nonabelian 2-groups with more than one involution, admitting a cyclic group of automorphisms which permutes their involutions transitively. The purpose of the first joint condition is to avoid considering known (and well understood) families of groups, such as elementary abelian, cyclic or generalised quaternion, which also have cyclic groups of automorphisms acting transitively on their involutions (in the elementary abelian case these are known as Singer cycles.)

Let us now consider the group $C := \left\{C(\lambda) : \lambda \in \mathbb{F}_q\right\}$. This is a cyclic group, generated by $C(\lambda^*)$, where λ^* is any primitive element of $\mathbb{F}_q$. It is clearly isomorphic to the multiplicative group of the field, where $\lambda \mapsto C(\lambda)$ establishes the said isomorphism, and acts via conjugation on the Sylow 2-subgroup P. Since

$$\lambda \cdot (a,b) = (a,b)^{\lambda} = \left(\lambda a, \lambda^{1+\theta} b\right),$$

and in view of (4.2), the action of C on the nonidentity elements of both Z and P/Z is regular. In fact, the action on P is via automorphisms since

$$\begin{aligned} \lambda \cdot (a_1,b_1)(a_2,b_2) &= \lambda \cdot \left(a_1 + a_2, b_1 + b_2 + a_1 a_2^{\theta}\right) \\ &= \left(\lambda a_1 + \lambda a_2, \lambda^{1+\theta} b_1 + \lambda^{1+\theta} b_2 + \lambda a_1 \left(\lambda a_2\right)^{\theta}\right) \\ &= \left(\lambda a_1, \lambda^{1+\theta} b_1\right)\left(\lambda a_2, \lambda^{1+\theta} b_2\right) \\ &= (\lambda \cdot (a_1,b_1))(\lambda \cdot (a_2,b_2)). \end{aligned}$$

[9] See Higman [Hig63] for the original paper that introduces them (the groups P appear as $A_2(n,\theta)$ therein), or Huppert and Blackburn [HB82, Chapter VIII, §7] for a definitive account.

The group $P \rtimes C$ is a Frobenius group with Frobenius kernel P and Frobenius complement C. It is the normaliser of P and is maximal in $\mathrm{Sz}(q)$. The maximal subgroups of $\mathrm{Sz}(q)$ are up to conjugacy[10]

(i) the normaliser $\Gamma = P \rtimes C$ of a Sylow 2-subgroup P,

(ii) $\mathrm{Sz}(q_0)$, where $q = q_0^r$, r is prime, and $q_0 > 2$,

(iii) $D_{2(q-1)}$,

(iv) $C_{q+\theta+1} \rtimes C_4$,

(v) $C_{q-\theta+1} \rtimes C_4$.

The metacyclic groups in the last two cases are Frobenius groups. Of course, $D_{2(q-1)}$ is also Frobenius. This is a consequence of the slightly more general fact that if N is an abelian group of odd order and H is the cyclic group of order 2 acting on N by inversion then the semidirect product $N \rtimes H$ is Frobenius.

4.2 Conjugacy classes of complements and 1-cohomology

In this section we shall discuss an application of Hulpke's method for finding the conjugacy classes of subgroups of a soluble group to a Sylow 2-subgroup P of $\mathrm{Sz}(q)$. The reader is referred to Hulpke [Hul99] for a detailed exposition of said method; in particular section 3, Lemma 3.1. Consider a subgroup H of P and observe that $H \cap Z$ is central in P, thus normal in all subgroups of P that contain it. Since $Z \lhd P$, the group HZ is defined and is normal in P from the discussion preceding Remark 4.4, thus both quotient groups $Z/H \cap Z$, $HZ/H \cap Z$ are defined as well. In fact $Z/H \cap Z$ is a subgroup of $HZ/H \cap Z$ and

$$HZ/H \cap Z/Z/H \cap Z \cong HZ/Z \cong H/H \cap Z.$$

Since $Z/H\cap Z$ and $H/H\cap Z$ intersect trivially, we see that $H/H\cap Z$ is a complement to $Z/H \cap Z$ in $HZ/H \cap Z$. Now let H_1, H_2 be a pair of subgroups of P. We observe the following.

[10] See Wilson [Wil09, §4.2.3] or the original source [Suz62, §15].

Lemma 4.5 *The subgroup H_1 is conjugate to H_2 if and only if $H_1\big/H_1 \cap Z$ is conjugate to $H_2\big/H_2 \cap Z$.*

Proof. Suppose first that $H_2 = H_1^g$ for some $g \in P$. Then

$$H_2 \cap Z = H_1^g \cap Z = H_1^g \cap Z^g = (H_1 \cap Z)^g = H_1 \cap Z,$$

where the last equality holds because $H_1 \cap Z$ is a central subgroup of P. Thus

$$H_2\big/H_2 \cap Z = H_1^g\big/H_1 \cap Z = (H_1/H_1 \cap Z)^{\overline{g}}.$$

Conversely, assume that $H_1\big/H_1 \cap Z$ is conjugate to $H_2\big/H_2 \cap Z$. Then $H_2 \cap Z = H_1 \cap Z$ holds and $H_2\big/H_2 \cap Z = \left(H_1\big/H_2 \cap Z\right)^{\overline{g}} = H_1^g\big/H_2 \cap Z$ for some $\overline{g} \in \overline{P}$. Thus $H_1^g = H_2$, proving the assertion. □

Let us now consider a set of representatives of the conjugacy classes of subgroups of P that contain Z, say $\mathcal{K}$, and a set of representatives of the conjugacy classes of subgroups of Z, say $\mathcal{H}$. Evidently $\mathcal{H}$ is just the set of subgroups of Z, while the members of $\mathcal{K}$ are the full preimages of $\mathfrak{s}(P/Z)$.

Lemma 4.6 *Let $\mathcal{K}, \mathcal{H}$ be as above. For each $K \in \mathcal{K}$, $H \in \mathcal{H}$ denote by $\mathcal{U}_{K,H}$ the full preimages of a set of representatives for the P-classes of complements to Z/H in K/H. Then*

$$\mathcal{C} = \bigcup_{K \in \mathcal{K}} \bigcup_{H \in \mathcal{H}} \mathcal{U}_{K,H} \tag{4.3}$$

is a set of representatives for the P-classes of subgroups of P.

Proof. Consider a subgroup L of P and let $K = \langle L, Z \rangle = LZ$, $H = L \cap Z$. Then L/H is a complement to Z/H in K/H, thus L is conjugate to a member of $\mathcal{U}_{K,H}$. Conversely, the proof of Lemma 4.5 shows that L can be conjugate to at most one group from $\mathcal{C}$. □

We note that the above lemma does not tell us for which pairs of subgroups (K, H) the set $\mathcal{U}_{K,H}$ is nonempty; only that, by considering all such pairs, we will end up with a complete list for the conjugacy classes of subgroups of P. We address this issue in the following lemma, but we hasten to inform the reader that a method which treats the general case has been obtained by Celler et. al. [CNW90].

However, their method assumes knowledge of a polycyclic presentation for the group at hand.

Lemma 4.7 *Suppose that $Z \leqslant K \leqslant P$ and let H be a central subgroup of P. Then Z/H has a complement in K/H if and only if K/H is elementary abelian. If such a complement does exist then $|\Phi(K)| \geqslant |K/Z|$.*

Proof. Recall that K/H is elementary abelian if and only if $\Phi(K) \leqslant H$, since the Frattini subgroup of a finite p-group is the unique normal subgroup of the said group minimal with the property that the quotient is elementary abelian. Now notice that one direction of the first claim follows immediately. In an elementary abelian group all subgroups are direct summands, so if K/H is elementary abelian then Z/H is complemented. Conversely, suppose that C/H is a complement to Z/H in K/H. Let us first note that since C/H is a complement,

$$C/H \cong K/H \Big/ Z/H \cong K/Z.$$

However, since K/Z is elementary abelian C/H is elementary abelian as well, thus $\Phi(C) \leqslant H$. Moreover, since $(Z/H)(C/H) = K/H$, we see that $ZC = K$. Therefore $K' = (ZC)' = Z'C' = C'$, and $\mho(K) = \mho(ZC) = \mho(C)$, since Z is central and elementary abelian. Hence

$$\Phi(K) = K'\mho(K) = C'\mho(C) = \Phi(C) \leqslant H.$$

We deduce that K/H is elementary abelian and this settles the first claim.

In proof of the second claim, let K be a subgroup of P that contains the centre and recall that $\mho(K) = \langle g^2 : g \in K \rangle$ is a subgroup of $\Phi(K) = K'\mho(K)$. We may then define a map

$$f : K/Z \to \mho(K), \;\; gZ \mapsto g^2.$$

This map is well-defined as $g_1 Z = g_2 Z$ implies that $g_1 = g_2 z$ for some $z \in Z$, thus $g_1^2 = (g_2 z)^2 = g_2^2$. Moreover $\ker f = \left\{ gZ \in K/Z : g^2 = 1 \right\}$. But $g^2 = 1$ if and only if $g \in Z$, thus $\ker f$ is trivial. The proof is now complete. □

Remark 4.8 *The inequality that appears in the preceding lemma is no idle observation. The upper bounds for the number of subgroups of a Sylow 2-subgroup P of* Sz(q) *as well as that of the normaliser of P depend on this inequality in a crucial way. This point will soon become clear.*

In view of the above lemma, equation (4.3) assumes the form

$$\mathcal{C} = \bigcup_{Z \leqslant K \leqslant P} \bigcup_{\Phi(K) \leqslant H \leqslant Z} \mathcal{U}_{K,H}. \tag{4.4}$$

We note in passing that the inequality of Lemma 4.7 becomes an equality precisely when K/Z is a subfield of $P/Z \cong \mathbb{F}_q$, that is, if and only if $\log_2|K:Z|$ is a divisor of $\log_2|P:Z|$. We shall now briefly recall some basic concepts from the theory of group extensions. We say that the group G is an extension of N by F if G has a normal subgroup N such that $G/N \cong F$. If G is such an extension, with $\phi : F \to G/N$ realising the isomorphism then a section of G through F is any set $\{\tau(f) : f \in F\}$ such that $\tau(1) = 1$ and $\tau(f)$ is a representative for the coset $\phi(f)$. Assuming that N is abelian, the map $F \to \mathrm{Aut}(N)$, $f \mapsto \left(n \mapsto n^{\tau(f)}\right)$ is well defined and independent of τ. The following

$$Z^1(F,N) := \left\{\gamma : F \to N : \gamma(f_1 f_2) = \gamma(f_1)^{\tau(f_2)}\gamma(f_2),\ \text{for all}\ f_1, f_2 \in F\right\}$$

is known as the group of **1-cocycles**, while

$$B^1(F,N) := \left\{\gamma_n = \left(f \mapsto n n^{-f}\right) : F \to N : n \in N\right\}$$

is the group of **1-coboundaries**. It is easy to see that B^1 is a subgroup of Z^1. Provided the extension G splits over N and $K \leqslant G$ is a fixed complement, every complement of N in G can be written as $\left\{k\gamma(\bar{k}) : k \in K\right\}$ for some $\gamma \in Z^1$, and two complements corresponding to cocycles $\gamma, \delta \in Z^1$ are conjugate in G if and only if $\gamma\delta^{-1}$ lies in B^1. *Thus the factor group $H^1 = Z^1/B^1$ is in one-to-one correspondence to the conjugacy classes of complements of N in G.* Note that if $N \leqslant Z(G)$ then $\gamma_n = \gamma_1$ for all $n \in N$, thus B^1 is the trivial group. Moreover the group of 1-cocycles reduces to

$$Z^1(F,N) = \{\gamma : F \to N : \gamma(f_1 f_2) = \gamma(f_1)\gamma(f_2),\ \text{for all}\ f_1, f_2 \in F\},$$

which is, by definition, the group of homomorphisms $\mathrm{Hom}(F,N)$. Thus, in the case of a central subgroup N, one has

$$H^1(F,N) \cong \mathrm{Hom}(F,N).$$

Taking $G = K/H$ and $N = Z/H$ in the above relation and noting that $F = K/H\Big/Z/H \cong K/Z$ and that Z/H is central in K/H yields

$$\begin{aligned} H^1(K/Z, Z/H) &\cong \mathrm{Hom}(K/Z, Z/H) \\ &\cong \mathrm{Hom}\Big(K/Z, Z\big/\Phi(K)\Big/H\big/\Phi(K)\Big). \end{aligned}$$

Let us rewrite (4.4) as

$$\mathcal{C} = \bigcup_{K/Z \leqslant P/Z} \bigcup_{H/\Phi(K) \leqslant Z/\Phi(K)} \mathcal{U}_{K,H}.$$

We notice that the factor groups K/Z and Z/H are elementary abelian, thus both K/Z and Z/H are vector spaces over $\mathbb{F}_2$. Set $V := V(2,n) \cong P/Z$, $X := K/Z$, $V(X) := Z/\Phi(K)$, and $Y := H/\Phi(K)$ to obtain yet another expression

$$\mathcal{C} = \bigcup_{X \subseteq V} \bigcup_{Y \subseteq V(X)} \mathcal{U}_{X,Y}, \tag{4.5}$$

where $\mathcal{U}_{X,Y}$ is defined naturally in correspondence to $\mathcal{U}_{K,H}$. In this notation

$$\mathrm{Hom}\Big(K/Z, Z\big/\Phi(K)\Big/H\big/\Phi(K)\Big) = \mathrm{Hom}\Big(X, V(X)\big/Y\Big) \cong \mathrm{Hom}(X, Y'),$$

where Y' is such that $Y \oplus Y' = V(X)$. Each element of $\mathrm{Hom}(X, Y')$ is a linear transformation of vector spaces, thus $\mathrm{Hom}(X, Y') \cong \mathcal{L}(X, Y')$. Since $\mathcal{U}_{X,Y}$ and $\mathcal{L}(X, Y')$ are in bijection, equation (4.5) yields

$$|\mathcal{C}| = \sum_{X \subseteq V} \sum_{\substack{Y \subseteq V(X) \\ V(X) = Y \oplus Y'}} \big|\mathcal{L}(X, Y')\big|. \tag{4.6}$$

Of course,

$$\dim \mathcal{L}(X, Y') = \dim X \dim Y', \tag{4.7}$$

but it is important to note that the dimension of the $V(X)$-space (which specifies the range of values for the dimension of the Y-space, thus also for the dimension of the Y'-space), does not depend solely on $\dim X$, but rather on the X-space itself.[11] Now consider an element U of $\mathcal{U}_{X,Y}$. Clearly $K = UZ$ normalises U, thus $P \geqslant N_P(U) \geqslant UZ$. Since $|X| = |K : Z| = |UZ : Z| = |U : U \cap Z|$, one has

$$1 \leqslant |P : N_P(U)| \leqslant \frac{|Z|^2}{|UZ|} = \frac{|Z|}{|U : U \cap Z|} = \frac{|Z|}{|X|} = \left|X'\right|, \tag{4.8}$$

where X' is such that $X \oplus X' = V$. Put informally, the size of each conjugacy class of subgroups with given "X-part" is at most the size of the "X'-part". Assembling equation (4.6) and inequality (4.8) yields

$$|\mathfrak{s}(P)| \leqslant \sum_{\substack{X \subseteq V \\ V = X \oplus X'}} \sum_{\substack{Y \subseteq V(X) \\ V(X) = Y \oplus Y'}} \left|\mathcal{L}(X, Y')\right|\left|X'\right|. \tag{4.9}$$

The proof of the following lemma is now straightforward.

Lemma 4.9 *Let $P \in \mathrm{Syl}_2(\mathrm{Sz}(q))$. The number of subgroups of P satisfies the following inequality*

$$|\mathfrak{s}(P)| \leqslant \sum_{i=0}^{n} \begin{bmatrix} n \\ i \end{bmatrix}_2 \sum_{j=0}^{n-i} \begin{bmatrix} n-i \\ j \end{bmatrix}_2 2^{n+i(n-(i+j+1))}.$$

Proof. In view of the inequality shown in Lemma 4.7, one has $|V(X)| \leqslant |Z||X|^{-1} = |X'|$. Now let $V^*(X)$ be the subspace of the X'-space isomorphic to $V(X)$ under the isomorphism carrying P/Z to Z. The right-hand-side of inequality (4.9) may thus be rewritten as

$$\begin{aligned}\sum_{X \subseteq V} \sum_{Y \subseteq V(X)} \left|\mathcal{L}(X, Y')\right|\left|X'\right| &= \sum_{X \subseteq V} \sum_{W \subseteq V^*(X)} \left|\mathcal{L}(X, W')\right|\left|X'\right| \\ &\leqslant \sum_{X \subseteq V} \sum_{W \subseteq X'} \left|\mathcal{L}(X, W')\right|\left|X'\right|,\end{aligned}$$

with the understanding that the dash symbol refers to a complementary subspace. In turn, the right-hand-side of the above inequality is

$$\sum_{i=0}^{n} \sum_{\substack{X \subseteq V \\ \dim X = i}} \sum_{j=0}^{n-i} \sum_{\substack{W \subseteq X' \\ \dim W = j}} \left|\mathcal{L}(X, W')\right|\left|X'\right|,$$

[11] In general, there exist distinct subgroups $Z \leqslant K_1, K_2$ of P such that $|K_1/Z| = |K_2/Z|$, but $|\Phi(K_1)| \neq |\Phi(K_2)|$.

which, by equation (4.7), is equal to

$$\sum_{i=0}^{n} \begin{bmatrix} n \\ i \end{bmatrix}_2 \sum_{j=0}^{n-i} \begin{bmatrix} n-i \\ j \end{bmatrix}_2 2^{i(n-i-j)} 2^{n-i} = \sum_{i=0}^{n} \begin{bmatrix} n \\ i \end{bmatrix}_2 \sum_{j=0}^{n-i} \begin{bmatrix} n-i \\ j \end{bmatrix}_2 2^{n+i(n-(i+j+1))}.$$

The proof is complete. □

4.3 The subgroups of the normaliser $\Gamma = P \rtimes C$

Recall that the multiplicative group $C = \mathbb{F}_q^\times$ of the field acts via automorphisms on P; in fact, the action of C on the nonidentity elements of both Z and P/Z is regular thus, a fortiori, a Frobenius action.

Lemma 4.10 *Let $B \leqslant C$ and suppose that both U and U^g are B-invariant subgroups of P, where $g \in P$. Then $g \in N_P(U)$.*

Proof. First note that the B-invariance of U implies the B-invariance of $N_P(U)$. To see why, let $b \in B$, $n \in N_P(U)$. Then $U^{b(n)} = b(U^n) = b(U) = U$, where the second equality holds because n normalises U and the last equality holds because U is B-invariant. Therefore $b(n) \in N_P(U)$, as claimed. We infer from this that the induced action of B on $P\big/N_P(U) = \overline{P}$ is Frobenius.[12] Now, suppose that b is a nontrivial element of B. Then $U^g = b(U^g) = U^{b(g)}$, thus $b^{-1}(g)g \in N_P(U)$. Hence $b(\overline{g}) = \overline{g}$, i.e., $\overline{g} \in C_{\overline{P}}(b) = \overline{1} = N_P(U)$, where the first equality holds because b is nontrivial and the action Frobenius. The claim follows. □

We deduce that at most one element from each conjugacy class is B-invariant, thus we may as well consider representatives for the conjugacy classes of subgroups of P and ask which of those representatives are B-invariant. We shall then be able to determine all subgroups of Γ by observing that $U^{g^{-1}}$ is B-invariant if and only if U is B^g-invariant, i.e., the conjugates of U are acted upon by the different inverse-conjugates of B, where U ranges in the set of B-invariant subgroups of P. As mentioned previously, the action of C on the nonidentity elements of both Z and P/Z is regular, thus Dickson's argument, as outlined in the proof of Theorem 3.7, is in effect. In particular, both Z and P/Z are vector spaces over the subfield $\mathbb{F}_b$ that b generates, where $\langle b \rangle = B$ is any subgroup of C, and both are isomorphic

[12] See Isaacs [Isa08, Corollary 6.2].

to $V_b := V\left(2^{m_b}, \frac{n}{m_b}\right)$, where $|\mathbb{F}_b| = 2^{m_b}$, $m_b := \min\{r \in \mathbb{N} : o(b) \mid 2^r - 1\}$. With this in mind, let us retain the notation V_b for the space P/Z and write $\overline{V_b}$ for the Z-space, so that we may distinguish between them. Further, for each $X \subseteq V_b$ define $V_b(X)$ to be the $\mathbb{F}_b$-space $Z\big/\Phi(K)$, where K is the full preimage of X. Let $\mathcal{U}_{X,Y}(\mathbb{F}_b)$ be the full preimages of a set of representatives for the P-classes of complements to Z/H in K/H, where H is the full preimage of the subspace $Y \subseteq V_b(X)$. Similar considerations to the ones established in the first part of this section furnish a proof for the following lemma.

Lemma 4.11 *Let Γ be the normaliser of a Sylow 2-subgroup P of* Sz(q). *Then*

$$|\mathfrak{s}(\Gamma)| \leqslant \sum_{b|q-1} \sum_{i=0}^{\frac{n}{m_b}} \begin{bmatrix} \frac{n}{m_b} \\ i \end{bmatrix}_{2^{m_b}} \sum_{j=0}^{\frac{n}{m_b}-i} \begin{bmatrix} \frac{n}{m_b} - i \\ j \end{bmatrix}_{2^{m_b}} 2^{n+i(n-m_b(i+j+1))}.$$

Proof. The proof is identical to that of Lemma 4.9; the only difference is that instead of $\mathbb{F}_2$, the underlying field now is $\mathbb{F}_b$. The details are thus omitted. □

Setting $\mathrm{I}(P) := |\mathfrak{s}(\Gamma)| - |\mathfrak{s}(P)|$, one has

$$\begin{aligned} \mathrm{I}(P) &\leqslant \sum_{\substack{b|q-1 \\ b>1}} \sum_{i=0}^{\frac{n}{m_b}} \begin{bmatrix} \frac{n}{m_b} \\ i \end{bmatrix}_{2^{m_b}} \sum_{j=0}^{\frac{n}{m_b}-i} \begin{bmatrix} \frac{n}{m_b} - i \\ j \end{bmatrix}_{2^{m_b}} 2^{n+i(n-m_b(i+j+1))} \qquad (4.10) \\ &= \sum_{\substack{b|q-1 \\ b>1}} \sum_{i=0}^{\frac{n}{m_b}} \sum_{j=0}^{\frac{n}{m_b}-i} \begin{bmatrix} \frac{n}{m_b} \\ i \end{bmatrix}_{2^{m_b}} \begin{bmatrix} \frac{n}{m_b} - i \\ j \end{bmatrix}_{2^{m_b}} 2^{n+i(n-m_b(i+j+1))}. \end{aligned}$$

Note that the q-binomial coefficient $\left[\begin{smallmatrix} m \\ k \end{smallmatrix}\right]_q$ satisfies the elementary double inequality

$$q^{k(m-k)} \leqslant \begin{bmatrix} m \\ k \end{bmatrix}_q \leqslant q^{k(m-k+1)}. \qquad (4.11)$$

To see why that must be, recall that

$$\begin{bmatrix} m \\ k \end{bmatrix}_q = \frac{(q^m-1)(q^{m-1}-1)\ldots(q^{m-k+1}-1)}{(q^k-1)(q^{k-1}-1)\ldots(q-1)} = \prod_{i=0}^{k-1} \frac{q^{m-i}-1}{q^{k-i}-1},$$

and notice that for each factor in the product we have

$$q^{m-k} \leqslant \frac{q^{m-i}-1}{q^{k-i}-1} \leqslant q^{m-k+1}.$$

Thus

$$q^{k(m-k)} = \prod_{i=0}^{k-1} q^{m-k} \leqslant \begin{bmatrix} m \\ k \end{bmatrix}_q \leqslant \prod_{i=0}^{k-1} q^{m-k+1} = q^{k(m-k+1)},$$

as claimed. In view of the above upper bound, we may thus write inequality (4.10) as

$$\begin{aligned} \mathrm{I}(P) &\leqslant \sum_{\substack{b \mid q-1 \\ b>1}} \sum_{i=0}^{\frac{n}{m_b}} \sum_{j=0}^{\frac{n}{m_b}-i} 2^{m_b i\left(\frac{n}{m_b}-i+1\right)} 2^{m_b j\left(\frac{n}{m_b}-i-j+1\right)} 2^{n+i(n-m_b(i+j+1))} \\ &= \sum_{\substack{b \mid q-1 \\ b>1}} \sum_{i=0}^{\frac{n}{m_b}} \sum_{j=0}^{\frac{n}{m_b}-i} 2^{i(n-im_b+m_b)} 2^{j(n-im_b-jm_b+m_b)} 2^{n+i(n-m_b(i+j+1))} \\ &= \sum_{\substack{b \mid q-1 \\ b>1}} \sum_{i=0}^{\frac{n}{m_b}} \sum_{j=0}^{\frac{n}{m_b}-i} 2^{f(i,j,m_b,n)}, \end{aligned} \tag{4.12}$$

where

$$f(i,j,m_b,n) := n(2i+j+1) - m_b\left(2i^2+2ij+j^2-j\right).$$

The summation limits of the innermost double sum as well as the nature of the summand make it clear that the quantity

$$\sum_{i=0}^{\frac{n}{m_b}} \sum_{j=0}^{\frac{n}{m_b}-i} 2^{f(i,j,m_b,n)},$$

when viewed as a function of m_b only, attains its maximum at

$$m_0 := \min\{m_b : o(b) \mid q-1, b \neq 1\} = \min\{p \in \mathbb{P} : p \mid n\}.$$

Since n is odd, we see that $m_0 \geqslant 3$. Writing $n' := \lfloor \frac{n}{3} \rfloor$ we obtain

$$\sum_{i=0}^{\frac{n}{m_b}} \sum_{j=0}^{\frac{n}{m_b}-i} 2^{f(i,j,m_b,n)} \leqslant \sum_{i=0}^{\frac{n}{m_0}} \sum_{j=0}^{\frac{n}{m_0}-i} 2^{f(i,j,m_0,n)} \leqslant \sum_{i=0}^{n'} \sum_{j=0}^{n'-i} 2^{f(i,j,3,n)}.$$

Therefore inequality (4.12) becomes

$$\mathrm{I}(P) \leqslant \sum_{\substack{b|q-1 \\ b>1}} \sum_{i=0}^{n'} \sum_{j=0}^{n'-i} 2^{f(i,j,3,n)}. \tag{4.13}$$

In the following section we shall obtain an upper bound for the right-hand-side of the above inequality and use this to establish that Γ and P have the same number of subgroups, asymptotically speaking.

4.4 Proof of $\left|\mathfrak{s}(\Gamma)\right| \sim \left|\mathfrak{s}(P)\right|$

Let us fix n temporarily (thus also n') and define

$$\mathcal{R} := \left\{(x,y) \in \mathbb{R}^2 : 0 \leqslant x \leqslant n', 0 \leqslant y \leqslant n' - x\right\}$$

to be the triangular region of the Cartesian plane lying in the first quadrant and below the line $x + y = n'$. Moreover, let

$$\overline{f} : \mathcal{R} \to \mathbb{R}, \ \ (x,y) \mapsto n(2x+y+1) - 3(2x^2 + 2xy + y^2 - y)$$

be the extension of f over the reals. We shall apply standard techniques from calculus in order to find the (absolute) maximum of $\overline{f}$ in $\mathcal{R}$. We begin by finding the interior local extrema of $\overline{f}(x,y)$. Now,

$$\begin{aligned} \frac{\partial \overline{f}}{\partial x} &= 2n - 12x - 6y, \ \text{ and} \\ \frac{\partial \overline{f}}{\partial y} &= n - 6x - 6y + 3. \end{aligned}$$

At an interior critical point the partial derivatives vanish. This, in our case, is equivalent to $(x_0, y_0) = \left(\frac{n-3}{6}, 1\right)$. The Hessian matrix of $\overline{f}$ equals

$$H_{\overline{f}}(x,y) = \begin{pmatrix} -12 & -6 \\ -6 & -6 \end{pmatrix},$$

hence $\det\left(H_{\overline{f}}(x,y)\right) = 36$ for all x, y. As $\frac{\partial^2 \overline{f}}{\partial x^2} = -12$ the second partial derivative test applies and shows that the critical point (x_0, y_0) is a local maximum. In fact

$$\overline{f}(x_0, y_0) = \frac{(n+3)^2}{6}.$$

We now check the maximum value of $\overline{f}(x,y)$ on the boundary of $\mathcal{R}$. The three cases to consider here correspond to the sides of our triangle and are

$$\begin{aligned} \overline{f}(0,y) &= -3y^2 + (n+3)y + n, \\ \overline{f}(x,0) &= -6x^2 + 2nx + n, \\ \overline{f}(x, n'-x) &= -3x^2 + (n-3)x + 3n' + nn' + n - 3n'^2, \end{aligned}$$

where x, y range in $[0, n']$. In each case the function $\overline{f}$ is a quadratic polynomial $\alpha z^2 + \beta z + \gamma$. Since $\alpha < 0$ in all cases and because $z_0 := -\frac{\beta}{2\alpha}$ is an interior point of the corresponding line segment, we see that $\overline{f}$ peaks at z_0. Thus the desired maximum of $\overline{f}$ is the maximum among

$$\begin{aligned} \overline{f}\left(\frac{n-3}{6}, 1\right) &= \frac{1}{6}n^2 + n + \frac{3}{2}, \\ \overline{f}\left(0, \frac{n+3}{6}\right) &= \frac{1}{12}n^2 + \frac{3}{2}n + \frac{3}{4}, \\ \overline{f}\left(\frac{n}{6}, 0\right) &= \frac{1}{6}n^2 + n, \\ \overline{f}\left(\frac{n-3}{6}, \frac{6n'-n+3}{6}\right) &= \frac{1}{12}n^2 + \left(n' + \frac{1}{2}\right)n - \left(3n'^2 - 3n' - \frac{3}{4}\right). \end{aligned}$$

Using $\frac{n}{3} - 1 \leqslant n' \leqslant \frac{n}{3}$, one easily sees that

$$\frac{1}{12}n^2 + \frac{7}{2}n - \frac{9}{4} \geqslant \overline{f}\left(\frac{n-3}{6}, \frac{6n'-n+3}{6}\right) \geqslant \frac{1}{12}n^2 + \frac{1}{2}n - \frac{9}{4}.$$

Therefore

$$\begin{aligned}\max_{n\geqslant 1}\left\{f(i,j,3,n):(i,j)\in\mathcal{R}\cap\mathbb{N}^2\right\} &\leqslant \max_{n\geqslant 1}\left\{\overline{f}(x,y):(x,y)\in\mathcal{R}\right\}\\ &\leqslant \frac{n^2}{6}+4n.\end{aligned}$$

We may thus write

$$\begin{aligned}\sum_{i=0}^{n'}\sum_{j=0}^{n'-i}2^{f(i,j,3,n)} &\leqslant \sum_{i=0}^{n'}\sum_{j=0}^{n'-i}2^{\frac{n^2}{6}+4n}\\ &< \left(\frac{n}{3}+1\right)^2 2^{\frac{n^2}{6}+4n}.\end{aligned}$$

Substituting this in (4.13), we obtain

$$\begin{aligned}\mathrm{I}(P) &\leqslant (d(q-1)-1)\left(\frac{n}{3}+1\right)^2 2^{\frac{n^2}{6}+4n} \\ &< 2^n n^2 2^{\frac{n^2}{6}+4n}\\ &< 2^{\frac{n^2}{6}+6n}.\end{aligned} \tag{4.14}$$

This bound is sufficient for our purposes. In order to see why that is, we look back at (4.11). Take $m=n$, $k=\frac{n-1}{2}$ and $q=2$ there. Then

$$2^{\frac{n^2-1}{4}}\leqslant\begin{bmatrix} n\\ \frac{n-1}{2}\end{bmatrix}_2.$$

Since Z is an elementary abelian 2-group, the quantity $\left[\begin{smallmatrix} n\\ \frac{n-1}{2}\end{smallmatrix}\right]_2$ counts the number of central subgroups of order $2^{\frac{n-1}{2}}$ in P. Hence

$$2^{\frac{n^2-1}{4}}\leqslant\begin{bmatrix} n\\ \frac{n-1}{2}\end{bmatrix}_2<|\mathfrak{s}(P)|, \tag{4.15}$$

which in turn implies that

$$0<\frac{|\mathfrak{s}(\Gamma)|-|\mathfrak{s}(P)|}{|\mathfrak{s}(P)|}<2^{\frac{n^2}{6}+6n-\frac{n^2}{4}+\frac{1}{4}}.$$

Thus

$$\lim_{n\to\infty}\frac{|\mathfrak{s}(\Gamma)|-|\mathfrak{s}(P)|}{|\mathfrak{s}(P)|}=0;$$

equivalently

$$\lim_{n\to\infty} \frac{|\mathfrak{s}(\Gamma)|}{|\mathfrak{s}(P)|} = 1. \tag{4.16}$$

A similar analysis to the one outlined above will reveal that

$$|\mathfrak{s}(P)| < 2^{\frac{(n+1)^2}{2}} \tag{4.17}$$

for all $n \in \mathbb{N}$, where the maximum of the implied $\overline{f}$ now occurs at an interior point.

4.5 Most subgroups are 2-groups

We begin this section with the following lemma, which is a straightforward application of the Schur-Zassenhaus theorem.

Lemma 4.12 *Let $G = A \rtimes B$ be a finite group, where* $\gcd(|A|, |B|) = 1$*. If $H \leqslant G$ then $H = (H \cap A)(H \cap B^g)$ for some $g \in A$.*

Proof. Observe that $H \cap A$ is a normal subgroup of H and that

$$\gcd\left(|H \cap A|, \left|H\big/H \cap A\right|\right) = 1,$$

since $H\big/H \cap A$ is isomorphic to a subgroup of B. By the Schur-Zassenhaus theorem, $H \cap A$ has a complement in H, say C, thus $H = (H \cap A)C$. Quoting the same theorem there exists a $g \in G$ such that $C^g \leqslant B$ (choose $g = 1$ if C is trivial.) Now write $g = ba$ for some $b \in B$, $a \in A$. Then $C^a \leqslant B$, hence $H^a = (H \cap A)^a C^a \leqslant (H^a \cap A)(H^a \cap B) \leqslant H^a$. We conclude that $H = (H \cap A)(H \cap B^g)$ for $g = a^{-1}$. □

Corollary 4.13 *Suppose that G is a finite Frobenius group satisfying the conditions of Lemma 4.12. Then* $|\mathfrak{s}(G)| \leqslant |A||\mathfrak{s}(A)||\mathfrak{s}(B)|$.

Proof. Note that when G is Frobenius the element g of Lemma 4.12 is unique, since $B \cap B^g = 1$ for all $g \in G \setminus B$. Now consider the (well-defined) map

$$f : \mathfrak{s}(G) \to A \times \mathfrak{s}(A) \times \mathfrak{s}(B), \quad H \mapsto (g, H \cap A, H \cap B^g),$$

where g is such that $H = (H \cap A)(H \cap B^g)$, and observe that f is injective. □

We apply the above corollary, along with the elementary inequality $d(k) \leqslant 2\sqrt{k}$, to the groups $D_{2(q-1)}$, $C_{q-\theta+1} \rtimes C_4$, and $C_{q+\theta+1} \rtimes C_4$:

(i) $\left|\mathfrak{s}\left(D_{2(q-1)}\right)\right| \leqslant 2(q-1)d(q-1) \leqslant 4q^{\frac{3}{2}}$,

(ii) $\left|\mathfrak{s}\left(C_{q-\theta+1} \rtimes C_4\right)\right| \leqslant 3(q-\theta+1)d(q-\theta+1) \leqslant 6q^{\frac{3}{2}}$,

(iii) $\left|\mathfrak{s}\left(C_{q+\theta+1} \rtimes C_4\right)\right| \leqslant 3(q+\theta+1)d(q+\theta+1) \leqslant 6 \cdot 2^{\frac{3}{2}} q^{\frac{3}{2}} < 17q^{\frac{3}{2}}$.

Assuming that $n \geqslant 9$, we see that $|\mathfrak{s}(H)| < q^2$ when H is any of the groups in the above list. In fact this inequality holds for all $n \in \mathbb{N}$ by a direct calculation. We shall also require the following lemma.

Lemma 4.14 *The number of subgroups of* $\mathrm{Sz}(q)$ *satisfies the following inequality*

$$|\mathfrak{s}(\mathrm{Sz}(q))| < 2^{\frac{11}{5}(\log_2 q)^2},$$

for all q an odd power of 2.

Proof. The proof is by induction on the exponent of q. To establish the base case, we use a computer algebra programme to compute the size of the subgroup lattice of $\mathrm{Sz}(8)$ and $\mathrm{Sz}(32)$ and find that $|\mathfrak{s}(\mathrm{Sz}(8))| = 17295 < 2^{15} < 2^{\frac{99}{5}}$, $|\mathfrak{s}(\mathrm{Sz}(32))| = 21170191 < 2^{25} < 2^{55}$, Now set $m := \log_2 q$ and let $\{p_1, \ldots, p_k\}$ be the set of distinct prime divisors of m. Since each subgroup of $\mathrm{Sz}(q)$ is contained in one of its maximal subgroups, we see that

$$\begin{aligned} |\mathfrak{s}(\mathrm{Sz}(q))| < (q^2+1)\,|\mathfrak{s}(\Gamma)| &+ \frac{1}{2}q^2(q^2+1)\left|\mathfrak{s}\left(D_{2(q-1)}\right)\right| \\ &+ \frac{1}{4}q^2(q-1)(q+\theta+1)\left|\mathfrak{s}\left(C_{q-\theta+1} \rtimes C_4\right)\right| \\ &+ \frac{1}{4}q^2(q-1)(q-\theta+1)\left|\mathfrak{s}\left(C_{q+\theta+1} \rtimes C_4\right)\right| \\ &+ \sum_{i=1}^{k} \left|\mathrm{Sz}(q) : \mathrm{Sz}\left(q^{1/p_i}\right)\right| \left|\mathfrak{s}\left(\mathrm{Sz}\left(q^{1/p_i}\right)\right)\right|. \end{aligned}$$

Observe that $\left|\mathrm{Sz}(q):\mathrm{Sz}\left(q^{1/p_i}\right)\right| < q^5$ as $|\mathrm{Sz}(q)| = q^2(q^2+1)(q-1) < q^5$ and recall that $|\mathfrak{s}(H)| < q^2$ when H is either the dihedral group, or one of the two metacyclic Frobenius groups. Hence

$$\begin{aligned} |\mathfrak{s}(\mathrm{Sz}(q))| &< (q^2+1)\,|\mathfrak{s}(\Gamma)| + q^2\left[\frac{1}{2}q^2(q^2+1) + \frac{1}{4}q^2(q-1)(q\pm\theta+1)\right] \qquad (4.18)\\ &\quad + q^5\sum_{i=1}^{k}\left|\mathfrak{s}\left(\mathrm{Sz}\left(q^{1/p_i}\right)\right)\right| \\ &= (q^2+1)\,|\mathfrak{s}(\Gamma)| + q^6 + q^5\sum_{i=1}^{k}\left|\mathfrak{s}\left(\mathrm{Sz}\left(q^{1/p_i}\right)\right)\right|. \end{aligned}$$

The induction hypothesis yields

$$\begin{aligned} |\mathfrak{s}(\mathrm{Sz}(q))| &< (q^2+1)\,|\mathfrak{s}(\Gamma)| + q^6 + q^5\sum_{i=1}^{k} 2^{\frac{11}{5}(m/3)^2} \\ &= (q^2+1)\,|\mathfrak{s}(\Gamma)| + q^6 + q^5\omega(m)2^{\frac{11}{45}m^2}. \end{aligned}$$

Recall that $|\mathfrak{s}(\Gamma)| = |\mathfrak{s}(P)| + \mathrm{I}(P) < 2^{\frac{(m+1)^2}{2}} + 2^{\frac{m^2}{6}+6m}$ by (4.17) and (4.14) respectively, hence

$$\begin{aligned} |\mathfrak{s}(\mathrm{Sz}(q))| &< (2^{2m}+1)\left(2^{\frac{(m+1)^2}{2}} + 2^{\frac{m^2}{6}+6m}\right) + 2^{6m} + 2^{\frac{11}{45}m^2+5m+\log_2\omega(m)} \\ &< 2^{2m+\frac{1}{2}}2^{\frac{(m+1)^2}{2}+\frac{m^2}{6}+6m} + 2^{6m} + 2^{\frac{11}{45}m^2+5m+\log_2\omega(m)} \\ &= 2^{\frac{2}{3}m^2+8m+\frac{3}{4}} + 2^{6m} + 2^{\frac{11}{45}m^2+5m+\log_2\omega(m)}. \end{aligned}$$

But $\max\left\{2^{\frac{2}{3}m^2+8m+\frac{3}{4}}, 2^{6m}, 2^{\frac{11}{45}m^2+5m+\log_2\omega(m)}\right\} = 2^{\frac{2}{3}m^2+8m+\frac{3}{4}}$ for all $m \in \mathbb{N}$, thus

$$|\mathfrak{s}(\mathrm{Sz}(q))| < 2^{\frac{2}{3}m^2+8m+\frac{3}{4}+\log_2 3} < 2^{\frac{11}{5}m^2},$$

since $\frac{11}{5}m^2 > \frac{2}{3}m^2+8m+\frac{3}{4}+\log_2 3$ for all $m \geqslant 7$. The induction is now complete. □

Remark 4.15 *The constant 11/5 which appears at the exponent of the upper bound for $|\mathfrak{s}(\mathrm{Sz}(q))|$ in Lemma 4.14 is by no means the best possible, but it is sufficient for our purposes. The reason becomes clear in the next section.*

4.6 Proof of Theorem 4.1

As explained in section 4.1, the Sylow 2-subgroups of Sz(q) intersect trivially. Moreover

$$\lim_{n\to\infty}\left|\mathrm{Syl}_2\left(\mathrm{Sz}(q)\right)\right| = \lim_{n\to\infty}\left(q^2+1\right) = \infty,$$

thus conditions (i) and (ii) of our Main Lemma 2.16 are satisfied with $p = 2$. Therefore, in order to complete the proof of Theorem 4.1, it suffices to show that condition (iii) is satisfied as well. To that end we look back at (4.18) which, in view of Lemma 4.14, yields

$$\begin{aligned}
\frac{|\mathfrak{s}\left(\mathrm{Sz}(q)\right)|}{(q^2+1)\,|\mathfrak{s}(\Gamma)|} &< 1 + \frac{q^6 + q^5 \sum\limits_{i=1}^{k}\left|\mathfrak{s}\left(\mathrm{Sz}\left(q^{1/p_i}\right)\right)\right|}{(q^2+1)\,|\mathfrak{s}(\Gamma)|}\\
&< 1 + \frac{2^{6\log_2 q} + 2^{\frac{11}{5}(\log_2 q/3)^2+5\log_2 q+\log_2 \omega(\log_2 q)}}{(q^2+1)\,|\mathfrak{s}(\Gamma)|}\\
&< 1 + \frac{2^{\frac{11}{45}(\log_2 q)^2+5\log_2 q+\log_2 \omega(\log_2 q)+1}}{(q^2+1)\,|\mathfrak{s}(\Gamma)|}.
\end{aligned}$$

We recall that $|\mathfrak{s}(\Gamma)| > |\mathfrak{s}(P)| > 2^{\frac{(\log_2 q)^2}{4}-\frac{1}{4}}$ by inequality (4.15), thus

$$(q^2+1)\,|\mathfrak{s}(\Gamma)| > q^2\,|\mathfrak{s}(P)| > 2^{\frac{(\log_2 q)^2}{4}+2\log_2 q-\frac{1}{4}}.$$

In conclusion

$$\begin{aligned}
\frac{|\mathfrak{s}\left(\mathrm{Sz}(q)\right)|}{(q^2+1)\,|\mathfrak{s}(\Gamma)|} &< 1 + 2^{\frac{11}{45}(\log_2 q)^2+5\log_2 q+\log_2 \omega(\log_2 q)+1-\left(\frac{(\log_2 q)^2}{4}+2\log_2 q-\frac{1}{4}\right)}\\
&= 1 + 2^{-\frac{1}{180}(\log_2 q)^2+3\log_2 q+\log_2 \omega(\log_2 q)+\frac{5}{4}},
\end{aligned}$$

hence

$$\lim_{n\to\infty}\frac{|\mathfrak{s}\left(\mathrm{Sz}(q)\right)|}{(q^2+1)\,|\mathfrak{s}(\Gamma)|} = 1.$$

Since $\lim\limits_{n\to\infty}\frac{|\mathfrak{s}(\Gamma)|}{|\mathfrak{s}(P)|} = 1$ by (4.16), so $\lim\limits_{n\to\infty}\frac{(q^2+1)|\mathfrak{s}(\Gamma)|}{|\mathcal{E}_n|} = 1$. Therefore

$$\lim_{n\to\infty}\frac{|\mathfrak{s}\left(\mathrm{Sz}(q)\right)|}{|\mathcal{E}_n|} = \lim_{n\to\infty}\frac{|\mathfrak{s}\left(\mathrm{Sz}(q)\right)|}{(q^2+1)\,|\mathfrak{s}(\Gamma)|}\cdot\lim_{n\to\infty}\frac{(q^2+1)\,|\mathfrak{s}(\Gamma)|}{|\mathcal{E}_n|} = 1\cdot 1 = 1.$$

The universe is no narrow thing and the order within it is not constrained by any latitude in its conception to repeat what exists in one part in any other part. Even in this world more things exist without our knowledge than with it and the order in creation which you see is that which you have put there, like a string in a maze, so that you shall not lose your way. For existence has its own order and that no man's mind can compass, that mind itself being but a fact among others.

Cormac McCarthy
Blood Meridian or the Evening Redness in the West

5

On the Frattini subgroup of homomorphic images

5.1 Introduction

In one of his last papers published in the Journal of Algebra, Doerk [Doe94] examined the finite soluble groups G that behave like nilpotent groups with respect to the Frattini subgroup, i.e., $\Phi(U) \leqslant \Phi(V)$ for all subgroups $U \leqslant V$ of G and $\Phi(G/N) = \Phi(G)N/N$ for all $N \lhd G$. The class $\mathfrak{X}$ of finite soluble groups G with Frattini subgroups $\Phi(U) \leqslant \Phi(V)$ for all subgroups $U \leqslant V$ of G is a subgroup closed local formation. Recall that a class of (soluble) groups is a formation if it closed under homomorphic images and has the following property: if G/M, G/N belong to the class, $M, N \lhd G$, so does $G/(M \cap N)$. The coinage of the term is due to Gaschütz [Gas63]. The groups in $\mathfrak{X}$ have elementary abelian Sylow subgroups and p-length $\leqslant 1$ for all primes p. The class $\mathfrak{F}_{\mathrm{sol}}$ of finite soluble groups G where $\Phi(G/N) = \Phi(G)N/N$ for any normal subgroup N of G is again a local formation and $\mathfrak{X}$ is the largest subgroup closed class within $\mathfrak{F}_{\mathrm{sol}}$. A group G in $\mathfrak{F}_{\mathrm{sol}}$ belongs to $\mathfrak{X}$ if and only if it has p-length $\leqslant 1$ for all primes p.

While Doerk treats the two properties of the Frattini subgroup both separately and in conjunction, we will address only the second one here. Recall that if K/L is a chief factor of G and $K/L \leqslant \Phi(G/L)$, then K/L is said to be a Frattini chief factor of G. Doerk's theorem on the class $\mathfrak{F}_{\mathrm{sol}}$ reads as follows.

Theorem 5.1 ([Doe94, Satz 2']) *Let G be a finite soluble group. Then the following are equivalent:*

(i) *If $N \lhd G$ then $\Phi(G/N) = \Phi(G)N/N$.*

(ii) *$G/\Phi(G)$ has no Frattini chief factors.*

(iii) *$G/\mathrm{F}(G)$ has no Frattini chief factors.*

(iv) *If H/K is a chief factor of G then $G/C_G(H/K)$ has no Frattini chief factors.*

We provide the following definition in order to facilitate reference to the core property of the Frattini subgroup that we shall be concerned with.

Definition 5.2 *A normal subgroup N of the finite group G is called a solution (in G) if N satisfies the equation $\Phi(G/N) = \Phi(G)N/N$, and G is called an $\mathfrak{F}$-group if G is finite and every subgroup $N \lhd G$ is a solution.*

In this chapter we focus on the class $\mathfrak{F}$ and provide (partial) answers to the following dual problem:

Which characteristic subgroups of G control the property of being an $\mathfrak{F}$-group and what are the solutions of an arbitrary finite group?

An answer to the first question (Theorem 5.5) also serves to characterise the finite soluble $\mathfrak{F}$-groups as iterated split extensions of direct products of elementary abelian groups (Corollary 5.6) and a partial answer to the second question establishes that subgroups lying in intervals defined by the terms of a certain "socle-central" characteristic series are always solutions (Corollary 5.9). We close our investigation on $\mathfrak{F}$-groups with an application to the computation of classes of subgroups of finite nilpotent groups.

Henceforth, and in keeping with Gaschütz's original terminology, we shall call a group Φ-free if its Frattini subgroup is trivial. If G is a finite group then we denote by $S_1(G)$ the product of its minimal abelian normal subgroups, i.e., the abelian part of its socle. We shall also assume that the reader is familiar with basic properties of the Frattini subgroup of a finite group. Namely that $\Phi(G_1 \times G_2) = \Phi(G_1) \times \Phi(G_2)$, that Φ is normal-subgroup-monotone, and that

$\Phi(G/N) \geqslant \Phi(G)N/N$, with equality when $N \leqslant \Phi(G)$, where $N \lhd G$. W. Gaschütz's early paper [Gas53] contains a wealth of information on the Frattini subgroup and we shall refer to it when needed (but see also [DH92, p. 30]).

5.2 The Frattini subgroup of homomorphic images

We should, perhaps, mention first that in the realm of finite soluble groups the behaviour of the Frattini subgroup of a homomorphic image is connected to what Gaschütz [Gas62] introduced and called *prefrattini subgroups*. A subgroup $W \leqslant G$ is called a prefrattini subgroup of G if W covers each non-complementable chief factor of G and if each maximal subgroup of G is conjugate to one that contains W. These subgroups are all conjugate and they are related to the (ordinary) Frattini subgroup; if W is a prefrattini subgroup of the finite soluble group G then for every normal subgroup N of G we have $\Phi(G/N) = \left(\cap_{g \in G} W^g N\right)/N$.

In the following lemma we collect a number of observations and easy-to-prove facts concerning the Frattini subgroup of quotients. Some of those will be needed later in the text.

Lemma 5.3 *Let G be a finite group.*

(i) *Suppose that N is a solution, $N \lhd G$. Then M/N is a solution (in G/N), $N \leqslant M \lhd G$, if and only if M is a solution.*

(ii) *If G is an $\mathfrak{F}$-group and $N \lhd G$ then G/N is an $\mathfrak{F}$-group.*

(iii) *G is an $\mathfrak{F}$-group if and only $G/\Phi(G)$ is an $\mathfrak{F}$-group.*

(iv) *Let $N \lhd G$. Then $\Phi(G/N) = \Phi(L)N\big/N$, where L is a minimal supplement to N in G.*[13]

(v) *If G is a Φ-free $\mathfrak{F}$-group then $S_1(G) = \mathrm{F}(G)$ and G splits over $\mathrm{F}(G)$.*

Proof. (i) As $G/M \cong G/N\big/M/N$, so $\Phi\left(G/M\right) \cong \Phi\left(G/N\big/M/N\right)$ and

$$\Phi(G/N) \cdot M/N\big/M/N = \Phi(G)N/N \cdot M/N\big/M/N.$$

[13] This is (also) a result of Bechtell [Bec65].

The right-hand-side is equal to $\Phi(G)M/N\big/M/N \cong \Phi(G)M/M$. The claim now follows.

(ii) Follows easily from (i).

(iii) If G is an $\mathfrak{F}$-group then $G/\Phi(G)$ is an $\mathfrak{F}$-group. Suppose that $G/\Phi(G)$ is an $\mathfrak{F}$-group and let $N \lhd G$. Then $N\Phi(G)/\Phi(G) \lhd G/\Phi(G)$ is a solution as is $\Phi(G)$, always. From (i) $N\Phi(G)$ is a solution and so $\Phi\left(G\big/N\Phi(G)\right) = N\Phi(G)\big/N\Phi(G)$. However, the preimages of $\Phi\left(G\big/N\Phi(G)\right)$ and $\Phi\left(G/N\right)$ are equal since the intersection of those maximal subgroups of G that contain N is precisely the intersection of the maximal subgroups that contain $N\Phi(G)$.

(iv) If $N \leqslant \Phi(G)$ then $L = G$ and the claim follows. So assume that $N \nleqslant \Phi(G)$. Then a minimal supplement L of N in G is proper and

$$\Phi(G/N) = \Phi\left(NL/N\right) \cong \Phi\left(L\big/N \cap L\right).$$

We claim that $N \cap L \leqslant \Phi(L)$. If not then $N \cap L$ has a proper supplement in L, say K. Then $G = NK$, which contradicts the minimality of L. Therefore

$$\Phi(G/N) \cong \Phi\left(L\big/N \cap L\right) = \Phi(L)\big/N \cap L = \Phi(L)\big/N \cap \Phi(L) \cong \Phi(L)N\big/N,$$

proving the claim.

(v) A finite group G is Φ-free if and only if G splits over $S_1(G)$ [Gas53, Satz 9]. But $S_1\left(G/\Phi(G)\right) = \mathrm{F}(G)/\Phi(G) = \mathrm{F}\left(G/\Phi(G)\right)$, where the first equality follows from [Gas53, Satz 13]. So if G is a Φ-free $\mathfrak{F}$-group then $S_1(G) = \mathrm{F}(G)$ and G splits over $\mathrm{F}(G)$. □

5.3 A criterion for a finite group to be an $\mathfrak{F}$-group

The findings presented in this section (the next auxiliary lemma and the theorem that follows it) are due to Isaacs [Isa14].

Lemma 5.4 *Suppose that $\Phi(G) = 1$ and $\Phi\left(G/\mathrm{F}(G)\right) = 1$. Then $\Phi(G/E) = 1$ for every normal nilpotent subgroup E of G.*

Proof. Write $\Phi(G/E) = U/E$. Then every maximal subgroup of G containing $\mathrm{F}(G)$ contains E and thus contains U, and so contains $U\mathrm{F}(G)$. Then $U\mathrm{F}(G)/\mathrm{F}(G) \leqslant$

$\Phi(G/\mathrm{F}(G)) = 1$ and thus $U \leqslant \mathrm{F}(G)$. But U is abelian, since $\Phi(G) = 1$ forces $\mathrm{F}(G)$ to abelian, and so U has a complement in G by [Gas53, Satz 7], say X. Then XE/E is a complement to U/E in G/E, and since $U/E = \Phi(G/E)$, it follows that $U/E = 1$. □

Theorem 5.5 *Suppose that $\Phi(G) = 1$ and assume for all $M \lhd G$ with $\mathrm{F}(G) \leqslant M$ that $\Phi(G/M) = 1$. Then $\Phi(G/K) = 1$ for all $K \lhd G$.*

Proof. We may assume that $K > 1$, and we proceed by double induction, first on $|G|$ and then on $|K|$. If $1 < N < K$ with $N \lhd G$, we argue that G/N satisfies the hypotheses. First, $\Phi(G/N) = 1$ by the inductive hypothesis applied with N in place of K. Next, if $M/N \geqslant \mathrm{F}(G/N)$ with $M \lhd G$ then $M \geqslant \mathrm{F}(G)$, so by hypothesis $\Phi(G/M) = 1$, thus $\Phi\left(G/N \middle/ M/N\right) = 1$. It follows by the inductive hypothesis applied in the group G/N that $\Phi\left(G/N \middle/ K/N\right) = 1$, so $\Phi(G/K) = 1$, as wanted. We may thus assume that K is minimal normal in G. Note that $\mathrm{F}(G)$ is abelian and so if $K \leqslant \mathrm{F}(G)$, we are done by Lemma 5.4. We can therefore assume that $\mathrm{F}(G) \cap K = 1$, and thus K is nonabelian and $K = K'$. Let L be a complement for $\mathrm{F}(G)$ in G. Now $\mathrm{F}(G)K = \mathrm{F}(G) \times K$ so $\mathrm{F}(G) \leqslant Z(\mathrm{F}(G)K)$. Let $C = \mathrm{F}(G)K \cap L$, so $C > 1$ and $C \lhd L$. Also, $\mathrm{F}(G)$ centralises C so $C \lhd G$ and $\mathrm{F}(G)K = \mathrm{F}(G) \times C$. Then $K = K' \leqslant (\mathrm{F}(G)K)' \leqslant C \leqslant L$. Let $U/K = \Phi(G/K)$. Every maximal subgroup of G containing $\mathrm{F}(G)K$ contains U and since $\Phi\left(G \middle/ \mathrm{F}(G)K\right)$ is trivial by hypothesis, it follows that $U \leqslant \mathrm{F}(G)K = \mathrm{F}(G) \times K$. Write $V = U \cap \mathrm{F}(G)$, so $U = V \times K$ and $V \lhd G$. Now V is complemented in G, so V has a complement X in $\mathrm{F}(G)$, and $X \lhd G$. Now $UXL \geqslant VXL = \mathrm{F}(G)L = G$, so $(U/K)(XL/K) = G/K$. Since $U/K = \Phi(G/K)$, we have $XL/K = G/K$, so $XL = G$, and hence by Dedekind's Lemma, $\mathrm{F}(G) = X(\mathrm{F}(G) \cap L) = X$. Since X is a complement for V in $\mathrm{F}(G)$, we have $V = 1$. But $U = VK$, so $U = K$. □

Note that Theorem 5.5 together with Lemma 5.3 (iii) extend the equivalency of (i), (ii), and (iii) of Theorem 5.1 to all finite groups. Theorem 5.5 can also be employed in the proof of the following equivalent characterisation of soluble $\mathfrak{F}$-groups.

Corollary 5.6 *Suppose that the group G is soluble. Then G is an $\mathfrak{F}$-group if and only if the Fitting series of $G/\Phi(G)$ splits and has factors which are direct products of elementary abelian groups.*

Proof. By Lemma 5.3 (iii) we see that G is an $\mathfrak{F}$-group if and only if $G/\Phi(G)$ is an $\mathfrak{F}$-group. Now one direction is clear. The opposite direction follows easily by induction on the Fitting length of $G/\Phi(G)$ and Theorem 5.5, with the observation that if F_i are the terms of the Fitting series of some group H then $F_i/\mathrm{F}(H)$ are the terms of the Fitting series of $H/\mathrm{F}(H)$. □

5.4 Computing subgroups of nilpotent groups

Our goal in this section is to improve upon an algorithm of Hulpke for computing representatives for the classes of subgroups of a finite soluble group, when a candidate group moreover possesses a central series. For the convenience of the reader, let us recall the key lemma underlying Hulpke's algorithm "SubgroupsSolvableGroup" which is implemented in GAP.

Lemma 5.7 ([Hul99]) *Let $N \lhd G$ be an elementary abelian normal subgroup, $\mathcal{A}$ a set of representatives of the conjugacy classes of those subgroups of G that contain N properly and $\mathcal{B}$ (containing N) a set of representatives of the G-classes of subgroups in N. For each $A \in \mathcal{A}$ let $\mathcal{B}_A$ be a set of representatives of the $N_G(A)$-classes of proper subgroups of N, that are normal in A. Finally, for $B \in \mathcal{B}_A$ set $C_{A,B} := N_G(A) \cap N_G(B)$ and let $\mathcal{U}_{A,B}$ be the full preimages of a set of representatives of the $C_{A,B}$-classes of complements to N/B in A/B. Then*

$$\mathcal{R} := \mathcal{A} \cup \mathcal{B} \cup \bigcup_{A \in \mathcal{A}} \bigcup_{B \in \mathcal{B}_A} \mathcal{U}_{A,B}$$

is a set of representatives for the G-classes of subgroups of G.

The main idea is that for any subgroup $U \leqslant G$ and for any normal subgroup $N \lhd G$, $U\big/U \cap N$ is a complement to $N\big/U \cap N$ in $UN\big/U \cap N$. Assuming that we know all subgroups containing or contained in N, we may use the well-known property of the 1-cohomology group being in bijection to conjugacy classes of complements to find all classes of subgroups. The reader should note, however, that for finite soluble groups in general there exists no method that provides a priori knowledge of the existence or nonexistence of a complement to N/H in K/H, where $N \lhd G$ and $1 \leqslant H \leqslant N \leqslant K \leqslant G$; thus the algorithm proceeds "greedily" and examines all pairs of subgroups $(A, B) \in (\mathcal{A}, \mathcal{B}_A)$ for a complement.

Our aim is to find the largest possible subclass in the class of finite soluble groups for which a definite criterion that decides whether a complement exists or not is available. The following lemma aids our purpose and we present it in greatest possible generality because the findings might prove to be of some independent interest. However, let it be noted that a proof specifically for nilpotent groups should be much shorter (in particular, every normal subgroup of a finite nilpotent group is a solution).

Lemma 5.8 *Let N be a central Φ-free subgroup of G, i.e., $N \leqslant \mathrm{Soc}(Z(G))$, H a subgroup of N, and K a subgroup of G containing N. Then*

(i) *N has a complement in G if and only if $N \cap \Phi(G) = 1$. Moreover, N is a solution.*

(ii) *N/H has a complement in K/H if and only if $N \cap \Phi(K) \leqslant H$.*

Proof. (i) Suppose first that N has a complement in G, say C. Then $G = NC$, N is normal in G, and $N \cap C = 1$. As both N and C normalise C, so $C \triangleleft G$. Hence $G = N \times C$, from which $\Phi(G) = \Phi(N) \times \Phi(C)$ follows. However, $\Phi(N) = 1$ by assumption hence $N \cap \Phi(G) = N \cap \Phi(C) = 1$. The reverse direction is a consequence of a theorem of Gaschütz [Gas53, Satz 7].

Now let C be a (not necessarily proper or nontrivial) complement to $N \cap \Phi(G)$ in N so that $N = (N \cap \Phi(G)) \times C$. The existence of C is guaranteed since N, being a subgroup of $\mathrm{Soc}(Z(G))$, is a direct product of elementary abelian subgroups and such a group is complemented [Hal37]. Then $C \cap \Phi(G) = 1$, hence C has a complement in G, say K. So $G = C \times K$ and $\Phi(G) = \Phi(K)$ just as in the proof contained in the above paragraph. Since

$$G/N = K \times C \Big/ (N \cap \Phi(G)) \times C \cong K \Big/ N \cap \Phi(G),$$

we see that

$$\Phi(G/N) \cong \Phi\left(K \Big/ N \cap \Phi(G)\right) = \Phi\left(K \Big/ N \cap \Phi(K)\right).$$

Obviously $N \cap \Phi(K) \leqslant \Phi(K)$, so $\Phi\left(K \Big/ N \cap \Phi(K)\right) = \Phi(K) \Big/ N \cap \Phi(K)$, which in turn is isomorphic to $\Phi(K)N \Big/ N$. Using $\Phi(G) = \Phi(K)$ again yields $\Phi(G/N) = \Phi(G)N \Big/ N$ and the claim follows.

(ii) It is not difficult to see that the property of being central and Φ-free is inherited by homomorphic images. So, by dint of (i), it suffices to establish that $N \cap \Phi(K) \leqslant H$ is equivalent to $N/H \cap \Phi(K/H) = 1$ instead. But H, being a subgroup of N, is central (in every subgroup that contains it) and Φ-free, thus (i) applies to H in K. Therefore

$$N/H \cap \Phi(K/H) = (N \cap \Phi(K)H)\big/H = (N \cap \Phi(K))H\big/H,$$

where the first equality follows from Dedekind's lemma. Clearly $(N \cap \Phi(K))H = H$ if and only if $N \cap \Phi(K) \leqslant H$. The proof is now complete. □

The first part of the above lemma tells us that certain subgroups of a finite group are always solutions.

Corollary 5.9 *Let $R_0 := 1$, $R_{i+1}/R_i := \mathrm{Soc}(Z(G/R_i))$. Then every normal subgroup of the finite group G lying in an interval $[\![R_j, R_{j+1}]\!]$ is a solution. In particular, the ultimate term R^∞ of the R-series is a solution.*

Proof. Induction. The base case is the content of Lemma 5.8 (i) and the induction step follows from Lemma 5.3 (i). □

Note that, in general, expressing "solution spaces" in terms of intervals is all one could ask for because the product of two solutions is not necessarily a solution and thus there is no largest subgroup containing all solutions. We construct a counterexample as follows.

Example 5.10 *Let E be elementary abelian of order 5^2 and assume that $C = \langle c \rangle$ is cyclic of order 4, acting on E by $x^c = x^2$, $x \in E$. Let $G = E \rtimes C$, the semidirect product, and write $E = M \times N$, where each of M and N has order 5. Note that M and N are normal in G. Also $\Phi(G) = 1$. (To see this, note that MC and NC are maximal in G, so $\Phi(G) \leqslant MC \cap NC = C$. Then $\Phi(G) \cap E = 1$ so $\Phi(G)$ centralises E. Since E is selfcentralising in G, this forces $\Phi(G) = 1$). Now $K \lhd G$ is a solution if and only if $\Phi(G/K)$ is trivial. Then M and N are solutions because G/M and G/N are each isomorphic to the Frobenius group of order 20 and so have trivial Frattini subgroups. But $MN = E$ is not a solution.*

When G is a finite nilpotent group, we may choose a central series with elementary abelian factors for G in a canonical way by refining its lower central series according to the primes involved being in strictly ascending (or descending) order, so that quotients are elementary abelian.

Proposition 5.11 *Let G be a finite nilpotent group and let $G = N_0 > N_1 > \ldots > N_r > N_{r+1} = 1$ be a central series of G with elementary abelian factors. Define by induction*

$$\mathcal{S}_0 := \bigcup_{H\in[\![N_0,N_1]\!]} H$$

and

$$\mathcal{S}_{i+1} := \bigcup_{A\in\mathcal{S}_i} \; \bigcup_{B\in[\![N_{i+1},(\Phi(A)\cap N_{i+1})N_{i+2}]\!]} \mathcal{U}_{A,B}$$

for $i \geqslant 0$, where $\mathcal{U}_{A,B}$ is the full preimages of a set of representatives for the $N_G(A)$-classes of complements to N_{i+1}/B in A/B. Then $\mathcal{S}_r$ is a set of representatives for the classes of subgroups of G.

Proof. We use induction on the length r of the central series of G with elementary abelian factors.

When $r = 0$, G is (elementary) abelian and $\mathcal{S}_0$ is simply the set of all subgroups of G, which is, of course, the unique set of representatives for the classes of subgroups of G. Thus the claim holds trivially in this case.

For the induction step we may assume that, since G/N_r has length one less than that of G, a set of representatives for its classes of subgroups is given by $\overline{\mathcal{S}}_{r-1}$, where

$$\overline{\mathcal{S}}_0 := \bigcup_{\overline{H}\in[\![\overline{N_0},\overline{N_1}]\!]} \overline{H}$$

and

$$\overline{\mathcal{S}}_{i+1} := \bigcup_{\overline{A}\in\overline{\mathcal{S}}_i} \; \bigcup_{\overline{B}\in[\![\overline{N}_{i+1},(\Phi(\overline{A})\cap\overline{N}_{i+1})\overline{N}_{i+2}]\!]} \mathcal{U}_{\overline{A},\overline{B}} \;\; ;$$

the overbar denotes quotients mod N_r. We take full preimages by removing the overbar and deduce that $\mathcal{S}_{r-1}$, as originally defined, is a set of representatives for the classes of those subgroups of G that contain N_r. In view of Lemma 5.7 and Lemma 5.8 (ii) we conclude that

$$\mathcal{S}_r = \bigcup_{A\in\mathcal{S}_{r-1}} \; \bigcup_{B\in[\![N_r,\Phi(A)\cap N_r]\!]} \mathcal{U}_{A,B}$$

is a set of representatives for the classes of subgroups of G. □

6

Conclusion and further research

6.1 Subgroup permutability

We have seen that $\mathfrak{p}(G)$ vanishes asymptotically when G is either $\mathrm{PSL}_2(2^n)$, or $\mathrm{Sz}(q)$. At the same time our intuition guides us to believe that all simple groups should have low subgroup permutability degree.

Problem 6.1 *Let $\mathcal{S}$ be the set of all nonabelian finite simple groups. Then the probability that two subgroups of G permute tends to 0 as $|G| \to \infty$, $G \in \mathcal{S}$.*

This problem strengthens Problem 4.3. of Tărnăuceanu [Tăr11], while the content of Chapters 3 and 4 provides a partial solution.

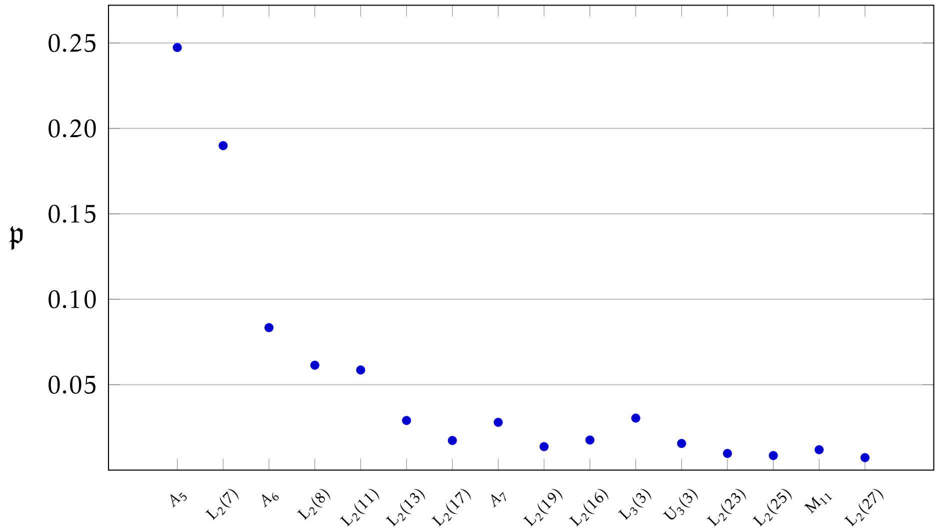


Figure 6.1: Subgroup permutability degrees of all simple groups of order $\leqslant 10^4$.

Guralnick and Robinson [GR06] have obtained an answer to a similar question upon replacing $\mathfrak{p}(G)$ with $\mathrm{cp}(G)$, where $\mathrm{cp}(G)$ is the commuting probability of G.

In particular, Theorem 4 in the aforementioned paper establishes that $\mathrm{cp}(G) \leqslant |G:\mathrm{F}(G)|^{-1/2}$, thereby proving $\mathrm{cp}(G) \to 0$ as $|G| \to \infty$, G a finite simple classical (or alternating) group. However, it seems unlikely to us that a compact bound of this sort is within reach as far as the subgroup permutability degree is concerned.

It seems that any strategy to tackle Problem 6.1 would require some version of Lemma 2.16, where the first condition is weakened. We ask the following.

Question 6.2 *Let $\{G_n\}_{n\geqslant 1}$ be a family of finite groups such that $p \mid |G_n|$ for some fixed prime p and for all $n \in \mathbb{N}$, satisfying the conditions*

(i) $O_p(G) = 1$,

(ii) $\lim\limits_{n\to\infty} \left|\mathrm{Syl}_p(G_n)\right| = \infty$, *and*

(iii) $\lim\limits_{n\to\infty} \dfrac{|\mathcal{E}_n|}{|\mathfrak{s}(G_n)|} = 1$,

where

$$\mathcal{E}_n := \left\{H \leqslant G_n : |H| = p^k \text{ for some } k \in \mathbb{N}\right\} = \bigcup_{P\in\mathrm{Syl}_p(G_n)} \mathfrak{s}(P).$$

Does it follow that $\lim\limits_{n\to\infty} \mathfrak{p}(G_n) = 0$?

A weaker version of the above question, if true, provides an interesting nonsimplicity criterion and stems from the empirical observation that high subgroup permutability degree forces normality.

Problem 6.3 *Let G be a finite group. If $\mathfrak{p}(G) > \mathfrak{p}(A_5)$ then G is not nonabelian simple.*

Finally, it would be interesting to have a clearer picture of the range of values that $\mathfrak{p}$ assumes.

Question 6.4 *Which rational numbers are limit points for $\mathfrak{p}$? Do irrational limit points exist?*

6.2 Insoluble $\mathfrak{F}$-groups

It is natural to speculate that there might be an analogue to Corollary 5.6 for arbitrary finite groups that are not necessarily soluble, in terms of some characteristic series whose ultimate term (if e.g. the series is ascending) is the whole group. For instance, one might be inclined to conjecture the following.

Conjecture 6.5 *Let G be a finite group. Then G is an $\mathfrak{F}$-group if and only if the generalised Fitting series of $G/\Phi(G)$ splits and has factors which are direct products of simple groups.*

Moreover, a result of this type would be useful in obtaining nontrivial information about $\mathfrak{F}$-groups for which Theorem 5.5 yields none at all, i.e., when already $\mathrm{F}(G)$ is trivial. Recall that the generalised Fitting subgroup $\mathrm{F}^*(G)$ of a (finite) group G is defined as $\mathrm{F}^*(G) = \mathrm{F}(G)\mathrm{L}(G)$, where $\mathrm{L}(G)$ is the layer of G, i.e., the subgroup generated by all quasisimple subnormal subgroups of G.

The above statement, unfortunately, is not true. An obvious counter-example is the Mathieu group M_{10} (indeed, this is the counter-example of least possible order).

Note, however, that if G is an $\mathfrak{F}$-group then the quotients defined by the generalised Fitting series of G *are* direct products of simple groups. First an observation: if $\Phi(G) = 1$ then $\Phi(K) = 1$ for all K subnormal in G; this is clear. Now consider a component N of G. Then N is subnormal in G and quasisimple, by definition. So $\Phi(N) = 1$ and we argue that in a quasisimple group the centre and its Frattini subgroup coincide. First, $\Phi(N)$ contains the intersection of N' with $Z(N)$, for all groups N [Gas53, Satz 4]. Since N is perfect by definition, $\Phi(N)$ contains $Z(N)$. But every proper normal subgroup of a quasisimple group is contained in its centre [Isa08, Lemma 9.2] and so $\Phi(N) = Z(N)$. So, in our case, $Z(N)$ is trivial and N is nonabelian simple. Thus $\mathrm{L}(G)$ is the direct product of some nonabelian simple groups and so the intersection of $\mathrm{F}(G)$ with $\mathrm{L}(G)$ is trivial, meaning that $\mathrm{F}^*(G) = \mathrm{L}(G) \times \mathrm{F}(G)$. But $\mathrm{F}(G)$ is normal in G and nilpotent and, moreover, $\Phi(\mathrm{F}(G)) = 1$. So $\mathrm{F}(G)$ is a direct product of abelian simple groups. The claim now follows from $\mathrm{F}^*(G) = \mathrm{L}(G) \times \mathrm{F}(G)$ and induction on the generalised Fitting length of G.

In the interest of dispelling any ambiguities, we mention that this direct-decomposition property of the F^*-quotients of G does not, conversely, characterise

$\mathfrak{F}$-groups. Consider the (permutational) wreath product $G = A_5 \wr B$, where B is a nonsplit extension of the elementary abelian 2-group of rank 3 by $\mathrm{PSL}_3(2)$. Then every quotient of the generalised Fitting series of G is a direct product of simple groups, but the Frattini subgroup of $G/\mathrm{Soc}(G)$ is equal to $\mathrm{Soc}(B) = \mathrm{C}_2^3$.

The main reason, it seems, that $\mathrm{F}(G)$ cannot be replaced by a larger (insoluble in general) characteristic subgroup in Theorem 5.5, can be traced to Gaschütz's Satz 7 which guarantees that normal abelian subgroups which avoid the Frattini subgroup are complemented. In the proof of Gaschütz's theorem we need to assume that the subgroup in question is abelian so as to ensure that every one of its subgroups is normal in that subgroup. Although there is, in principal, no need to consider groups all of whose subgroups are normal (we should only ask that the intersection of that subgroup with any one of its minimal supplements is normal in that subgroup), guaranteeing that only those particular subgroups are normal is an impossible task (also note that there is nothing to be gained by considering Dedekind groups instead of abelian groups).

Who are the inventors of Tlön? The plural is inevitable, because the hypothesis of a lone inventor—an infinite Leibniz labouring away darkly and modestly—has been unanimously discounted. It is conjectured that this brave new world is the work of a secret society of astronomers, biologists, engineers, metaphysicians, poets, chemists, algebraists, moralists, painters, geometers directed by an obscure man of genius. Individuals mastering these diverse disciplines are abundant, but not so those capable of inventiveness and less so those capable of subordinating that inventiveness to a rigorous and systematic plan.

Jorge Luis Borges
Tlön, Uqbar, Orbis Tertius

A

On the density of Singer cycles in $\mathrm{GL}_n(q)$

A.1 Introduction

In a series of papers stretching from 1965 to 1972, and beginning with [ET65], Erdős and Turán examined in detail various questions of a statistical nature regarding the symmetric group. One might ask whether similar work can be done for other classes of groups; indeed, the most natural next candidate is the general linear group $\mathrm{GL}_n(q)$ of $n \times n$ matrices over the finite field $\mathbb{F}_q$. A notable difference between the two classes of groups is the dependence of the latter on more than one parameter. While Σ_Ω is completely determined (up to isomorphism) by the cardinality of the set Ω, the groups $\mathrm{GL}_n(q)$ require 3 variables for their definition: the rank n and the size q of the underlying field, which, in turn, depends both on the characteristic of the field and on the degree of the extension. Stong [Sto93] considered the average order of a matrix in $\mathrm{GL}_n(q)$ for fixed q and varying n and proved that

$$\log \nu_n = n \log q - \log n + o(\log n),$$

where

$$\nu_n = \frac{1}{|\mathrm{GL}_n(q)|} \sum_{A \in \mathrm{GL}_n(q)} \mathrm{ord}(A).$$

Our purpose here is to address a question which is more limited in scope, but rather different in nature from Stong's investigations. In particular, we shall consider elements of maximal order in $\mathrm{GL}_n(q)$, also known as Singer cycles, and examine in detail the mean density of those elements in $\mathrm{GL}_n(q)$. Fixing any two of the parameters n, p, r and letting the remaining one vary accordingly, we show that the density of Singer cycles follows a distribution law and provide its expected value. It is straightforward to show that the maximal order of an element in $\mathrm{GL}_n(q)$ is q^n-1. In section A.2 we shall do this after establishing existence of the said elements and obtain the formula

$$\frac{|\mathrm{GL}_n(q)|}{q^n-1}\cdot\frac{\phi\,(q^n-1)}{n}$$

for the number of Singer cycles, where ϕ is the usual Euler function. The core of the work presented in this chapter is concerned with their density function

$$\mathfrak{t}_n(q) := \frac{1}{n}\frac{\phi(q^n-1)}{q^n-1}, \tag{A.1}$$

i.e., the probability that an element has maximal order in $\mathrm{GL}_n(q)$. We are interested in the distribution of $\mathfrak{t}_n(q)$ in the interval $(0,\frac{1}{n})$. If we fix n and let $q\to\infty$ through prime powers, we see that $\lim_q \mathfrak{t}_n(q)$ does not exist. It thus makes sense to examine its average value instead. In Theorem A.1 (i) we provide the answer to this question. However, the average is not greatly affected by the values that $\mathfrak{t}_n(q)$ assumes when q varies through genuine powers of primes. We therefore investigate the average of $\mathfrak{t}_n(p^r)$ for fixed n, p and for varying r. This is the content of Theorem A.1 (ii). Lastly, by similar methods we provide the average value for the case that Stong deals with, that is when the field is fixed and the rank n varies, and that is the content of Theorem A.1 (iii). We notice here that similar questions have been considered previously in the case where one has the multiplicative group $\mathbb{Z}/m\mathbb{Z}$ instead of the general linear group. See for example [Li98] and the survey [LP02]. Before we state our first theorem, let us introduce some relevant terminology. For a prime p and an integer a coprime to p, we let $\ell_p(a)$ denote its multiplicative order (mod p), that is the least positive integer k for which $p^k\equiv 1(\mathrm{mod}\ a)$. Define for a prime p and integer r the following series

$$\mathfrak{t}(p,r) := \sum_{m\in\mathbb{N}}{}^{*}\ \frac{\mu(m)}{m}\frac{\gcd\left(\ell_p(m),r\right)}{\ell_p(m)}, \tag{A.2}$$

Here and throughout this chapter $\sum^*$ denotes a summation over those positive integers m that are coprime to p. We will show in Lemma A.12 (iii) that this series is convergent. We also define for an integer n the quantity

$$\mathfrak{t}_n := \frac{1}{n} \prod_p \left(1 - \frac{\gcd(p-1,n)}{p(p-1)}\right). \tag{A.3}$$

Notice that $\mathfrak{t}_1$ is the so-called Artin constant, arising in Artin's primitive root conjecture. The infinite product in (A.3) converges since for fixed n one has $\gcd(p-1,n) \leqslant n$, hence

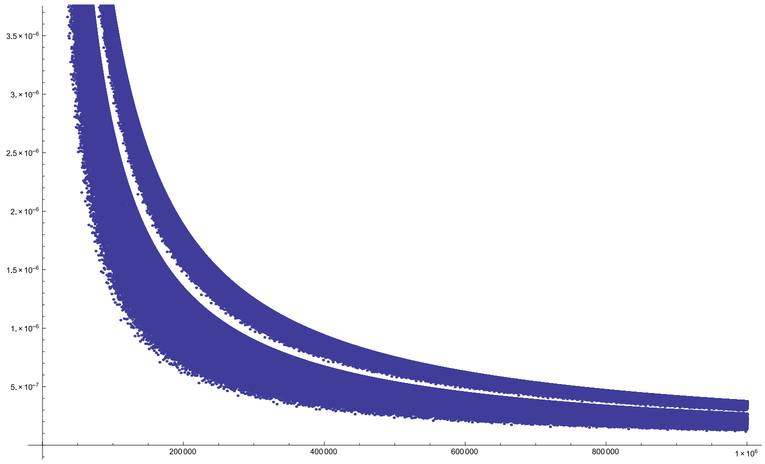


Figure A.1: The first one million values of $\mathfrak{t}_n$.

$$\sum_p \frac{\gcd(p-1,n)}{p(p-1)} \leqslant n \sum_p \frac{1}{p(p-1)} < \infty.$$

Our first result is the following.

Theorem A.1 *Let $x \in \mathbb{R}_{\geqslant 1}$ and denote the cardinality of the powers of primes below x by $Q(x)$.*

(i) *For any fixed $n \in \mathbb{N}$ and any $A \in \mathbb{R}_{>1}$ one has*

$$\frac{1}{Q(x)} \sum_{q \leqslant x} \mathfrak{t}_n(q) = \mathfrak{t}_n + O_{n,A}\left(\frac{1}{(\log x)^A}\right),$$

where the summation is taken over powers of primes.

(ii) *For any fixed prime p and $n \in \mathbb{N}$ one has*

$$\frac{1}{x}\sum_{r \leqslant x} \mathfrak{t}_n(p^r) = \frac{1}{n}\mathfrak{t}(p,n) + O\left(\frac{\log(xn\log p)}{xn}\right),$$

where the implied constant is absolute.

(iii) *For any fixed $q = p^r$ one has*

$$\frac{1}{\log x}\sum_{n \leqslant x} \mathfrak{t}_n(q) = \mathfrak{t}(p,r) + O\left(\frac{\tau(r)\log(r\log p)}{\log x}\right),$$

where the implied constant is absolute and τ denotes the divisor function.

Remark A.2 *The above theorem reveals the noteworthy fact that the density of Singer cycles in $\mathrm{GL}_n(q)$ is approximated on average by a constant multiple of $\frac{1}{n}$ when one allows any of the underlying parameters to vary.*

The special case corresponding to $n = 1$ in Theorem A.1 (i) has been dealt with previously by Stephens [Ste69, Lemma 1] and in an equivalent form by Pillai [Pil41, Theorem 1]. Following the proof of Theorem A.1 (i), we show that the error of approximation can be substantially improved to

$$O_n\left(\frac{(\log x)^{\tau(n)+2}}{\sqrt{x}}\right)$$

under the assumption of the Generalised Riemann Hypothesis. One should notice that the sequence $\mathfrak{t}_n(q)$ oscillates wildly around its mean value in all cases of Theorem A.1 and it would therefore be interesting to obtain information regarding the nature of its distribution. It thus makes sense to examine how $\mathfrak{t}_n(q)$ distributes over subintervals of $\left(0, \frac{1}{n}\right)$. To that end let us recall some standard definitions from Probabilistic Number Theory (see [Ten95, Chapter III] for a detailed discussion). Let b_n be a sequence of real numbers and define for $x, z \in \mathbb{R}$, the frequencies ν_x as follows:

$$\nu_x(n; b_n \leqslant z) := \frac{|\{n \leqslant x : b_n \leqslant z\}|}{x}.$$

Similarly denote

$$\nu_x\left(q; b_q \leqslant z\right) := \frac{\left|\left\{q \leqslant x : q \text{ is a prime power}, b_q \leqslant z\right\}\right|}{Q(x)}.$$

We say that *the frequencies ν_x converge to a limiting distribution as $x \to \infty$*, if for any z in a certain dense subset $E \subset \mathbb{R}$, the following limit exists

$$\lim_{x\to\infty} \nu_x(n; b_n \leqslant z),$$

and furthermore, denoting its value by $F(z)$, that one has

$$\lim_{\substack{z\to\alpha\\ z\in E}} F(z) = \begin{cases} 1, & \text{if } \alpha = +\infty \\ 0, & \text{if } \alpha = -\infty \end{cases}.$$

Thus the existence of a limiting distribution should be interpreted as an equidistribution of b_n with respect to some measure. We are now ready to state our second theorem.

Theorem A.3 *The frequencies ν_x, with respect to any of the involved parametres, converge to a limiting distribution. More precisely:*

(i) *For fixed $n \in \mathbb{N}$ the frequencies $\nu_x(q; \mathfrak{t}_n(q) \leqslant z)$ converge to a continuous limiting distribution.*

(ii) *For fixed prime p and integer n the frequencies $\nu_x(r; \mathfrak{t}_n(p^r) \leqslant z)$ converge to a limiting distribution.*

(iii) *For fixed prime power $q = p^r$ the frequencies $\nu_x(n; \mathfrak{t}_n(q) \leqslant \frac{z}{n})$ converge to a limiting distribution.*

Let us put Theorem A.3 in context. The easier problem of determining the frequencies $\nu_x(n; \phi(n)/n \leqslant z)$, where one ranges general integers n, has received much attention. Shoenfeld [Sch28] proved that these frequencies converge to a limiting distribution, say $F(z)$, and that $F(z)$ is continuous. In a subsequent paper [Sch36] he showed that $F(z)$ is strictly increasing. Erdős [Erd39] discovered the following asymptotic expression:

$$F(1-\varepsilon) = 1 - \frac{e^{-\gamma}}{\log\frac{1}{\varepsilon}} + O\left(\frac{1}{\left(\log\frac{1}{\varepsilon}\right)^2}\right) \tag{A.4}$$

uniformly for all $\varepsilon \in (0,1)$, where γ is the Euler constant. Toulmonde [Tou09] proved that we can in fact get a more precise expression in the right-hand-side of (A.4), involving an asymptotic expansion in negative powers of $\frac{1}{\log(1/\varepsilon)}$ in place of the error term. The problem however becomes less easy when we range over powers of primes rather than general integers. The special case of part (i) of Theorem A.3 corresponding to $n = 1$ has been handled by Kátai [Kat68], who showed that the limiting distribution exists and is continuous. Deshouillers and Hassani [DH12] proved that this limiting distribution possesses infinitely many points of nondifferentiability. It would be desirable to have analogues of (A.4) for part (i) of Theorem A.3, even in the case $n = 1$, a problem which is equivalent to obtaining an analogue of (A.4) regarding the limiting distribution of $\frac{\phi(p-1)}{p-1}$, as p ranges in primes. A few remarks about the proofs of the two theorems are in order. Part (i) of Theorem A.1 is proved via splitting a certain sum over moduli into two ranges according to the size of the moduli. The contribution coming from the range corresponding to small moduli will give the main term after an application of the Bombieri-Vinogradov Theorem A.7 and the range corresponding to large moduli will be shown to give a negligible contribution compared to the main term. The proofs of Theorem A.1 (ii) and Theorem A.1 (iii) are rather similar; both are based on Lemma A.16, which resembles [Shp90, Theorem 3]. There, however, the dependence of the error term on the underlying parameters is not best possible and we will attempt to remedy this. The proof of Lemma A.16 is based on Lemma A.12, where all sums appearing in its statement resemble Romanoff's series

$$\sum_{m\in\mathbb{N}}^{*} \frac{1}{m\ell_p(m)},$$

where p is a fixed prime. We will bound such sums by introducing the auxiliary quantity $E_p(x,d)$, defined later in the proof of Lemma A.12, which is a trick introduced by Erdős [Erd51]. The proof of Theorem A.3 (i) is conducted via restricting attention to prime numbers below x and then applying a theorem in [TF69] regarding distribution laws of $f(|g(p)|)$, for an additive function f and a polynomial $g \in \mathbb{Z}[x]$. The proofs of parts (ii) and (iii) of Theorem A.3 are again quite similar. Here, however, the sequence involved is not additive. Therefore we have to use a result for distribution laws of general arithmetic functions. The rest of the chapter is organised as follows. In section A.2 some background on Singer cycles is given, resulting in the explicit estimation of $\mathfrak{t}_n(q)$. In section A.3 we provide some analytic tools that will be used later in the proof of Theorem A.1. In section A.4 some auxiliary lemmata regarding upper bounds of sums of certain arithmetic functions are provided. Finally, in sections A.5 and A.6 we provide the proof of Theorems A.1 and A.3 respectively.

Notation

Throughout this Chapter p will denote a prime and q a (not necessarily proper) power of a prime.

(i) The notation $\sum_p$ is understood to be a sum taken over primes; similarly, $\sum_q$ should be read as a sum taken over powers of primes, and the same principle applies to products. As already mentioned, $\sum^*$ denotes a summation over positive integers that are coprime to p.

(ii) We shall write $p^\lambda \| n$ for a prime p and positive integers λ, n if $p^\lambda \mid n$ and $p^{\lambda+1} \nmid n$.

(iii) For the real functions $f(x), g(x)$, defined for $x > 0$, the notation $f(x) = O(g(x))$ (or, equivalently, $f(x) \ll g(x)$) means that there exists an absolute constant $M > 0$ independent of x, such that $|f(x)| \leqslant M|g(x)|$ for $x > 0$. We shall write $f(x) = g(x) + O(h(x))$ if $(f - g)(x) = O(h(x))$. When the implied constant depends on a set of parameters $\mathcal{S}$, we shall write $f = O_{\mathcal{S}}(g)$, or $f \ll_{\mathcal{S}} g$. If no such subscript appears then the implied constant is absolute.

(iv) As usual, we let $\mu(n)$ the Möbius function, $\sigma(n)$ the sum of divisors of n, and $\Lambda(n)$ the von Mangoldt function. This, we recall, is defined as

$$\Lambda(n) = \begin{cases} \log p, & \text{if } n \text{ is a power of a prime } p, \\ 0, & \text{otherwise.} \end{cases}$$

A.2 Singer cycles

An element of order $q^n - 1$ in $\mathrm{GL}_n(q)$ is called a *Singer cycle*. That Singer cycles always exist can be seen as follows. Let $\mathbb{F}_{q^n}$ be the n-degree field extension of $\mathbb{F}_q$, and let α be a primitive element of $\mathbb{F}_{q^n}^*$. The map

$$s : \mathbb{F}_{q^n} \to \mathbb{F}_{q^n}, \ x \mapsto \alpha x$$

is $\mathbb{F}_q$-linear and invertible. Further, the order of s is equal to that of α in $\mathbb{F}_{q^n}^*$, that is, $q^n - 1$. In fact the integer $q^n - 1$ is maximal among possible element orders in

$\mathrm{GL}_n(q)$. To see why that must be the case, consider the algebra $\mathrm{Mat}_n(q)$ of all $n \times n$ matrices over $\mathbb{F}_q$ and note that each element $A \in \mathrm{GL}_n(q)$ generates a subalgebra $\mathbb{F}_q[A] \subseteq \mathrm{Mat}_n(q)$. The Cayley-Hamilton theorem then ensures that $\dim \mathbb{F}_q[A] \leqslant n$, thus $o(A) \leqslant q^n - 1$ for all $A \in \mathrm{GL}_n(q)$. Our claim now is that $o(A) = q^n - 1$ if and only if the minimal polynomial $m_A(x)$ of A is primitive of degree n (recall that f is a primitive polynomial if and only if any root of f in the splitting field of f generates the multiplicative group of that field). For this we shall need the following [LN94, Lemma 3.1].

Lemma A.4 *Let $f \in \mathbb{F}_q[x]$ be a polynomial of degree $m \geqslant 1$ with $f(0) \neq 0$. Then there exists a positive integer $e \leqslant q^m - 1$ such that $f(x) \mid x^e - 1$.*

The least such e is called the *order* of f and is denoted by $\mathrm{ord}(f)$. Clearly $A^k = I_n$ if and only if $m_A(x) \mid x^k - 1$, $k \in \mathbb{N}$. Thus[14]

$$o(A) = \mathrm{ord}\,(m_A).$$

From Lidl and Niederreiter [LN94, Theorem 3.3] we know that the order of an irreducible polynomial of degree n over $\mathbb{F}_q[x]$ is equal to the order of any of its roots in $\mathbb{F}_{q^n}^*$. Thus $\mathrm{ord}\,(m_A) = q^n - 1$ if and only if m_A is primitive and $\deg m_A = n$, as wanted. Note that the above argument shows that the minimal and the characteristic polynomial of a Singer cycle coincide. We shall now give a proof for the number of Singer cycles as an application of a theorem of Reiner [Rei61, Theorem 2], but it ought to mentioned that the same formula can be obtained via the familiar Orbit-Stabiliser Theorem (a Singer cyclic subgroup has index exactly n in its normaliser). Reiner computes the number of (not necessarily invertible) $n \times n$ matrices with entries in the finite field $\mathbb{F}_q$ having given characteristic polynomial. Let R_n denote the ring of all $n \times n$ matrices with entries in $\mathbb{F}_q$, and define $F(u,r) = \prod_{i=1}^{r} \left(1 - u^{-i}\right)$, where $F(u,0) = 1$.

Theorem A.5 ([Rei61]) *Let $g(x) \in \mathbb{F}_q[x]$ be a polynomial of degree n, and let*

$$g(x) = f_1^{n_1}(x) \cdots f_k^{n_k}(x)$$

[14] This yields yet another proof that the order of a matrix is at most $q^n - 1$.

be its factorisation in $\mathbb{F}_q[x]$ into powers of distinct irreducible polynomials $f_1(x), \dots f_k(x)$. Set $d_i := \deg(f_i(x))$, $1 \leqslant i \leqslant k$. Then the number of matrices $X \in R_n$ with characteristic polynomial $g(x)$ is

$$q^{n^2-n} \frac{F(q,n)}{\prod_{i=1}^{k} F\left(q^{d_i}, n_i\right)}.$$

Now take $k = 1$, $n_1 = 1$, and $d_1 = n$ in the above formula and notice that a matrix whose characteristic polynomial is thus parameterised is necessarily invertible. In particular there are

$$q^{n^2-n} \frac{F(q,n)}{F(q^n,1)} = \frac{|\mathrm{GL}_n(q)|}{q^n - 1}$$

such matrices. Since there are precisely

$$\frac{\phi(q^n - 1)}{n}$$

primitive (thus also irreducible) polynomials of degree n in $\mathbb{F}_q[x]$, the following has been proved.

Lemma A.6 *The number of Singer cycles in $\mathrm{GL}_n(q)$ is given by the formula*

$$\frac{|\mathrm{GL}_n(q)|}{q^n - 1} \frac{\phi(q^n - 1)}{n}.$$

A.3 Preliminaries

In this section we recall standard results regarding the distribution of primes in arithmetic progressions. Denote by $\pi(x) = \sum_{p \leqslant x} 1, x \geqslant 3$ the number of primes less than or equal to x and by $\mathrm{li}(x)$ the logarithmic integral

$$\mathrm{li}(x) := \int_2^x \frac{\mathrm{d}t}{\log t}.$$

The Prime Number Theorem states that for each $A > 0$ and for all $x \geqslant 3$ one has

$$\pi(x) = \mathrm{li}(x) + O_A\left(\frac{x}{(\log x)^A}\right).$$

For coprime integers m, a and all $x \geqslant 3$, define

$$\psi(x; m, a) := \sum_{\substack{n \leqslant x \\ n \equiv a (\mathrm{mod}\, m)}} \Lambda(n). \tag{A.5}$$

The Bombieri-Vinogradov theorem [Dav00, §28] then states that:

Theorem A.7 *For any fixed $A > 0$ and for all $x \geqslant 3$ one has*

$$\sum_{m \leqslant \sqrt{x}/(\log x)^A} \max_{\substack{a (\mathrm{mod}\ m) \\ (a,m)=1}} \left| \psi(x; m, a) - \frac{x}{\phi(m)} \right| \ll_A \frac{x}{(\log x)^{A-5}}.$$

We shall also make use of the following lemma.

Lemma A.8 *Let a, m be coprime integers. Then one has the following estimate*

$$\sum_{\substack{p \leqslant x \\ p \equiv a (\mathrm{mod}\ m)}} \frac{1}{p} = \frac{1}{\phi(m)} \log\log x + O_m(1),$$

for $x \geqslant 3$.

Proof. Corollary 4.12(c) in Montgomery-Vaughan [MV07] yields

$$\sum_{\substack{p \leqslant x \\ p \equiv a (\mathrm{mod}\ m)}} \frac{1}{p} = \frac{1}{\phi(m)} \log\log x + O_m \left(1 + \sum_{\chi \neq \chi_0 (\mathrm{mod}\ m)} \log |L(1, \chi)| \right),$$

where the summation $\sum_{\chi \neq \chi_0 (\mathrm{mod}\ m)}$ is taken over nontrivial characters (mod m). Using the fact that there are finitely many such characters and that $L(1, \chi) \neq 0$ for each such character proves the lemma. □

A.4 Lemmata

In this section we introduce a certain arithmetical function $\rho_n(m)$ and provide an explicit expression when m is square-free. We then obtain upper bounds for sums that involve this function. At the end of the section we give some auxiliary lemmata regarding sums that involve the multiplicative order function $\ell_p(n)$. Let $n \in \mathbb{N}$ be a fixed positive integer, and define the function $\rho_n : \mathbb{N} \to \mathbb{N}$ via the rule

$$m \mapsto |\{a \ (\mathrm{mod}\ m) : a^n \equiv 1 \ (\mathrm{mod}\ m)\}|. \tag{A.6}$$

It is a direct consequence of the Chinese Remainder Theorem that ρ_n is multiplicative. The verification of the following lemma is left to the reader.

Lemma A.9 *For all primes p and for all positive integers n one has*

$$\rho_n(p) = \gcd(p-1, n).$$

Lemma A.10 *One has the following bounds, valid for all $n \in \mathbb{N}$ and for all $x \geqslant 3$*

(i) $\sum_{m \leqslant x} \mu^2(m) \frac{\rho_n(m)}{m} \ll_n (\log x)^{\tau(n)}$,

(ii) $\sum_{m > x} \mu^2(m) \frac{\rho_n(m)}{m^2} \log m \ll_n \dfrac{1}{x} (\log x)^{\tau(n)+1}$.

Proof. (i) We begin by noticing that if $f : \mathbb{N} \to \mathbb{R}_{\geqslant 0}$ is a multiplicative function then

$$\sum_{m \leqslant x} \mu^2(m) f(m) \leqslant \prod_{p \leqslant x} (1 + f(p)).$$

Setting $f(m) = \dfrac{\rho_n(m)}{m}$ in the above relation yields

$$\sum_{m \leqslant x} \mu^2(m) \frac{\rho_n(m)}{m} \leqslant \prod_{p \leqslant x} \left(1 + \frac{\gcd(p-1, n)}{p}\right)$$

for all $x \geqslant 3$. The inequality $1 + t \leqslant e^t$, valid for any $t > 0$, shows that

$$\prod_{p \leqslant x} \left(1 + \frac{\gcd(p-1, n)}{p}\right) \leqslant \exp\left(\sum_{p \leqslant x} \frac{\gcd(p-1, n)}{p}\right).$$

Using the identity

$$\gcd(a, b) = \sum_{\substack{d \mid a \\ d \mid b}} \phi(d), \tag{A.7}$$

valid for all $a, b \in \mathbb{N}$, yields

$$\sum_{p \leqslant x} \frac{\gcd(p-1, n)}{p} = \sum_{d|n} \phi(d) \left(\sum_{\substack{p \leqslant x \\ p \equiv 1 (\mathrm{mod}\ d)}} \frac{1}{p} \right).$$

Using Lemma A.8 for each inner sum gives

$$\sum_{p \leqslant x} \frac{\gcd(p-1, n)}{p} = \tau(n) \log\log x + O_n(1),$$

which proves that

$$\sum_{m \leqslant x} \mu^2(m) \frac{\rho_n(m)}{m} \ll_n (\log x)^{\tau(n)}.$$

(ii) Splitting the range of summation in disjoint intervals gives

$$\sum_{m > x} \mu^2(m) \frac{\rho_n(m)}{m^2} \log m = \sum_{i=0}^{\infty} \sum_{xe^i < m \leqslant xe^{i+1}} \mu^2(m) \frac{\rho_n(m)}{m^2} \log m.$$

Noticing that

$$\sum_{xe^i < m \leqslant xe^{i+1}} \mu^2(m) \frac{\rho_n(m)}{m^2} \log m \leqslant \frac{\log(xe^{i+1})}{xe^i} \sum_{m \leqslant xe^{i+1}} \mu^2(m) \frac{\rho_n(m)}{m}$$

and applying part (i) yields

$$\sum_{m > x} \mu^2(m) \frac{\rho_n(m)}{m^2} \log m \ll_n \frac{1}{x} \sum_{i=0}^{\infty} \frac{(\log(xe^{i+1}))^{\tau(n)+1}}{e^i}.$$

Using the inequality $\log(ab) \leqslant (\log a)(\log b)$, valid for all $a, b > e$ gives

$$\sum_{m > x} \mu^2(m) \frac{\rho_n(m)}{m^2} \log m \ll_n \frac{(\log x)^{\tau(n)+1}}{x} \sum_{i=0}^{\infty} \frac{(i+1)^{\tau(n)+1}}{e^i}.$$

This proves the assertion of the lemma since the series

$$\sum_{i=0}^{\infty} \frac{(i+1)^{\tau(n)+1}}{e^i}$$

is convergent. □

We record here for future reference a familiar lemma which allows us to translate information about the asymptotic behaviour of weighted sums by $\Lambda(k)$ into one regarding unweighted sums.

Lemma A.11 *Let b_k be a sequence of real numbers and define for any $x \in \mathbb{R}_{>1}$,*

$$\beta(x) := \max\{|b_k| : 1 \leqslant k \leqslant x\}.$$

Then one has the following estimates.

(i) *For any $x \in \mathbb{R}_{\geqslant 2}$, $\sum_{q \leqslant x} b_q = \sum_{p \leqslant x} b_p + O\left(\sqrt{x}\frac{\beta(x)}{\log x}\right)$.*

(ii) *Let c, A be positive constants such that $\sum_{p \leqslant x} b_p \log p = cx + O\left(\frac{x}{(\log x)^A}\right)$, for all $x \geqslant 3$. Then $\sum_{p \leqslant x} b_p = c\,\mathrm{li}(x) + O\left(\frac{x}{(\log x)^A}\right)$, for all $x \geqslant 2$.*

Proof. (i) Define $k_0 := \left[\frac{\log x}{\log 2}\right]$. We partition the sum $\sum_{q \leqslant x} b_q$ according to the values of prime powers that q assumes and this yields

$$\sum_{q \leqslant x} b_q = \sum_{k=1}^{k_0} \sum_{p \leqslant x^{\frac{1}{k}}} b_{p^k}.$$

Now notice that

$$\sum_{k=2}^{k_0} \sum_{p \leqslant x^{\frac{1}{k}}} b_{p^k} \ll \beta(x)\left[\pi\left(\sqrt{x}\right) + (k_0 - 2)\,\pi\left(x^{\frac{1}{3}}\right)\right],$$

and use the bound $\pi(t) \ll \frac{t}{\log t}$, valid for all $t \in \mathbb{R}_{\geqslant 2}$, to get

$$\sum_{k=2}^{k_0} \sum_{p \leqslant x^{\frac{1}{k}}} b_{p^k} \ll \sqrt{x}\frac{\beta(x)}{\log x}.$$

The claim follows. (ii) Applying partial summation yields

$$\sum_{p \leqslant x} b_p = c\,\mathrm{li}(x) + O\left(\frac{x}{(\log x)^{A+1}} + \int_2^x \frac{\mathrm{d}t}{(\log t)^{A+2}}\right).$$

The use of the following inequalities

$$\int_2^{\sqrt{x}} \frac{\mathrm{d}t}{(\log t)^{A+2}} \leqslant \frac{\sqrt{x}}{(\log 2)^{A+2}}, \quad \int_{\sqrt{x}}^{x} \frac{\mathrm{d}t}{(\log t)^{A+2}} \leqslant \frac{x-\sqrt{x}}{\left(\log\left(\sqrt{x}\right)\right)^{A+2}},$$

concludes the proof of the lemma. □

In the following lemma we record auxiliary bounds that will be needed when we deal with cases (ii) and (iii) of Theorem A.1.

Lemma A.12 *Let p be a fixed prime and $d \in \mathbb{N}$. One has the following bounds, uniformly for all $x \geqslant 1$.*

(i) $\displaystyle\sum_{k\leqslant x} \sideset{}{^*}\sum_{\substack{m\in\mathbb{N}\\ \ell_p(m)=kd}} \frac{1}{m} \ll \log(xd\log p),$

(ii) $\displaystyle\sum_{k\geqslant x} \frac{1}{k} \sideset{}{^*}\sum_{\substack{m\in\mathbb{N}\\ \ell_p(m)=kd}} \frac{1}{m} \ll \frac{\log(xd\log p)}{x},$

(iii) *the series $\mathfrak{t}(p,r)$ defined in (A.2) converges for each prime p and each $r \in \mathbb{N}$. Further, one has*

$$\mathfrak{t}(p,r) \ll \tau(r)\log(r\log p),$$

with an absolute implied constant.

Proof. (i) The integers m taken into account in the inner sum satisfy $p^{kd} \equiv 1 (\mathrm{mod}\ m)$ for some $k \leqslant x$. Therefore each such m is a divisor of

$$E_p(x,d) := \prod_{k\leqslant x} (p^{kd}-1).$$

Hence the double sum is at most

$$\sum_{m|E_p(x)} \frac{1}{m} = \frac{\sigma\left(E_p(x,d)\right)}{E_p(x,d)}.$$

We now use the well-known bound

$$\frac{\sigma(n)}{n} \ll \log\log n$$

to deduce that the double sum is $\ll \log\log E_p(x,d)$. One easily sees that

$$E_p(x,d) \leqslant \prod_{k\leqslant x} p^{kd} \leqslant p^{dx^2},$$

which shows that the double sum is $\ll \log\log(p^{dx^2}) \ll \log(xd\log p)$, as asserted. (ii) The term corresponding to $k = x$ makes a contribution only when x is an integer, in which case we get a contribution which is $\ll \frac{1}{x}\log(xd\log p)$, as shown by the first part of this lemma. It remains to examine the contribution made by the terms corresponding to $k > x$. Using partial summation along with part (i) we deduce that

$$\sum_{k>x}\frac{1}{k}\sideset{}{^*}\sum_{\substack{m\in\mathbb{N}\\ \ell_p(m)=kd}}\frac{1}{m} \ll \frac{\log(xd\log p)}{x} + \int_x^\infty \frac{\log(ud\log p)}{u^2}\mathrm{d}u.$$

Alluding to the estimate $\int_x^\infty (\log u)u^{-2}\mathrm{d}u \ll \log(2x)x^{-1}$, valid for all $x \geqslant 1$, proves our claim. (iii) The identity (A.7) and

$$\mathfrak{t}(p,r) = \sum_{k=1}^\infty \frac{\gcd(k,r)}{k}\sideset{}{^*}\sum_{\substack{m\in\mathbb{N}\\ \ell_p(m)=k}}\frac{\mu(m)}{m}$$

show that one has

$$\mathfrak{t}(p,r) = \sum_{d|r}\frac{\phi(d)}{d}\sum_{k=1}^\infty\frac{1}{k}\sideset{}{^*}\sum_{\substack{m\in\mathbb{N}\\ \ell_p(m)=kd}}\frac{\mu(m)}{m},$$

which is bounded in absolute value by

$$\sum_{d|r}\sum_{k=1}^\infty\frac{1}{k}\sideset{}{^*}\sum_{\substack{m\in\mathbb{N}\\ \ell_p(m)=kd}}\frac{1}{m}.$$

Using $x = 1$ in part (ii) concludes the proof of part (iii). □

A.5 Proof of Theorem A.1

In this section we prove Theorem A.1. We begin by proving the auxiliary Lemmata A.13, A.14, and A.15 and use them in succession to provide the proof of Theorem A.1 (i). We then prove Lemma A.16 from which we deduce the validity of Theorem A.1 (ii) and Theorem A.1 (iii). For fixed $n, m \in \mathbb{N}$ and $x \in \mathbb{R}_{>1}$, define the following functions

$$\Psi_n(x) := \sum_{k \leqslant x} \Lambda(k) \frac{\phi(k^n - 1)}{k^n - 1},$$

and

$$\Psi_n(x; m) := \sum_{\substack{a (\mathrm{mod}\, m) \\ a^n \equiv 1 (\mathrm{mod}\, m)}} \psi(x; m, a),$$

where ψ was defined in (A.5).

Lemma A.13 *For all naturals $n \geqslant 1$, and for all $x \in \mathbb{R}_{>1}$ one has*

$$\Psi_n(x) = \sum_{m \leqslant x^n} \frac{\mu(m)}{m} \Psi_n(x; m).$$

Proof. The proof follows readily by noticing that

$$\frac{\phi(k)}{k} = \sum_{m | k} \frac{\mu(m)}{m}, \tag{A.8}$$

and inverting the order of summation. □

Lemma A.14 *For any fixed constant $A > 0$ and any fixed $n \in \mathbb{N}$, one has uniformly for all $x \geqslant 3$*

$$\sum_{\sqrt{x}/(\log x)^A < m \leqslant x^n} \frac{\mu(m)}{m} \Psi_n(x; m) \ll_{A,n} \sqrt{x} (\log x)^{2 + \tau(n) + A}.$$

Proof. We break the summation over m into the disjoint intervals $[\sqrt{x}/(\log x)^A, x]$ and $(x, x^n]$. We first deal with the contribution afforded by the latter interval. We claim that for $m > x$ one has $\psi(x; m, a) \leqslant \log x$. To see why, note that there exists at most one k in $[1, x]$ such that $k \equiv a (\mathrm{mod}\, m)$, and notice that $\Lambda(k) \leqslant \log k$, for all

$k \in \mathbb{N}$, with equality if and only if k is prime. Thus $\psi(x;m,a) \leqslant \log x$, as wanted. Recalling the definition of $\rho_n(m)$, given in equation (A.6), we get for $m > x$,

$$\Psi_n(x;m) = \sum_{\substack{a(\mathrm{mod}\ m)\\ a^n\equiv 1(\mathrm{mod}\ m)}} \psi(x;m,a) \leqslant \rho_n(m)\log x.$$

Therefore

$$\left|\sum_{x<m\leqslant x^n} \frac{\mu(m)}{m}\Psi_n(x;m)\right| \leqslant \log x \sum_{m\leqslant x^n} \mu^2(m)\frac{\rho_n(m)}{m}.$$

We use the first part of Lemma A.10 to conclude that

$$\left|\sum_{x<m\leqslant x^n} \frac{\mu(m)}{m}\Psi_n(x;m)\right| \ll_n (\log x)^{1+\tau(n)}.$$

We proceed by estimating the contribution inherited from m in the range $(\sqrt{x}/(\log x)^A, x]$. Since

$$\sum_{\substack{k\leqslant x\\ k\equiv a(\mathrm{mod}\ m)}} 1 \leqslant \left[\frac{x}{m}\right] + 1 \leqslant 2\frac{x}{m}$$

for $m \leqslant x$, we see that

$$\psi(x;m,a) = \sum_{\substack{k\leqslant x\\ k\equiv a(\mathrm{mod}\ m)}} \Lambda(k) \leqslant \log x \sum_{\substack{k\leqslant x\\ k\equiv a(\mathrm{mod}\ m)}} 1 \leqslant 2\frac{x}{m}\log x.$$

Therefore

$$\begin{aligned}\left|\sum_{\sqrt{x}/(\log x)^A<m\leqslant x} \frac{\mu(m)}{m}\Psi_n(x;m)\right| &\leqslant \sum_{\sqrt{x}/(\log x)^A<m\leqslant x} \frac{\mu^2(m)}{m} \sum_{\substack{a(\mathrm{mod}\ m)\\ a^n\equiv 1(\mathrm{mod}\ m)}} \psi(x;m,a)\\ &\leqslant 2x\log x \sum_{\sqrt{x}/(\log x)^A<m\leqslant x} \mu^2(m)\frac{\rho_n(m)}{m^2}.\end{aligned}$$

Using the second part of Lemma (A.10) we get

$$x\log x \sum_{m>\sqrt{x}/(\log x)^A} \mu^2(m)\frac{\rho_n(m)}{m^2} \ll_{n,A} \sqrt{x}(\log x)^{\tau(n)+A+2},$$

thus completing the proof. □

Lemma A.15 *For any fixed constant $A>0$ one has uniformly for all $x \geqslant 3$*

$$\sum_{m \leqslant \sqrt{x}/(\log x)^A} \frac{\mu(m)}{m} \Psi_n(x;m) = n\mathfrak{t}_n x + O_{n,A}\left(\frac{x}{(\log x)^{A-5}}\right).$$

Proof. By the definition of $\Psi_n(x;m)$ we have

$$\sum_{m \leqslant \sqrt{x}/(\log x)^A} \frac{\mu(m)}{m} \Psi_n(x;m) = \sum_{m \leqslant \sqrt{x}/(\log x)^A} \frac{\mu(m)}{m} \sum_{\substack{a (\mathrm{mod}\ m)\\ a^n \equiv 1 (\mathrm{mod}\ m)}} \psi(x;m,a),$$

which equals

$$\begin{aligned}
&\sum_{m \leqslant \sqrt{x}/(\log x)^A} \frac{\mu(m)}{m} \sum_{\substack{a (\mathrm{mod}\ m)\\ a^n \equiv 1 (\mathrm{mod}\ m)}} \left(\psi(x;m,a) - \frac{x}{\phi(m)}\right)\\
+x &\sum_{m \leqslant \sqrt{x}/(\log x)^A} \frac{\mu(m)}{m} \sum_{\substack{a (\mathrm{mod}\ m)\\ a^n \equiv 1 (\mathrm{mod}\ m)}} \frac{1}{\phi(m)}\\
&= \mathcal{E} + x\mathcal{M}, \text{say.}
\end{aligned}$$

Recalling the definition of $\rho_n(m)$ (equation (A.6)), one has

$$|\mathcal{E}| \leqslant \sum_{m \leqslant \sqrt{x}/(\log x)^A} \frac{\rho_n(m)}{m} \max_{\substack{a (\mathrm{mod}\ m)\\ (a,m)=1}} \left|\psi(x;m,a) - \frac{x}{\phi(m)}\right|. \tag{A.9}$$

Now notice that by the definition of $\rho_n(m)$ one trivially has $\rho_n(m) \leqslant m$. Thus inequality (A.9) becomes

$$|\mathcal{E}| \leqslant \sum_{m \leqslant \sqrt{x}/(\log x)^A} \max_{\substack{a (\mathrm{mod}\ m)\\ (a,m)=1}} \left|\psi(x;m,a) - \frac{x}{\phi(m)}\right| \ll_A \frac{x}{(\log x)^{A-5}},$$

where a use of Theorem A.7 has been made. For the other term we get

$$\mathcal{M} = \sum_{m \leqslant \sqrt{x}/(\log x)^A} \frac{\mu(m)}{m} \frac{\rho_n(m)}{\phi(m)}.$$

We will show that this series converges. To that end, let us bound the tail of the series as follows. Using the inequality

$$\phi(m) \gg \frac{m}{\log m},$$

valid for all $m \geqslant 2$, we deduce that

$$\left| \sum_{m>\sqrt{x}/(\log x)^A} \frac{\mu(m)}{m} \frac{\rho_n(m)}{\phi(m)} \right| \ll \sum_{m>\sqrt{x}/(\log x)^A} \mu^2(m) \frac{\rho_n(m)}{m^2} \log m.$$

By part (ii) of Lemma A.10

$$\sum_{m>\sqrt{x}/(\log x)^A} \mu^2(m) \frac{\rho_n(m)}{m^2} \log m \ll_n \frac{\left(\log\left(\frac{\sqrt{x}}{(\log x)^A}\right)\right)^{\tau(n)+1}}{\frac{\sqrt{x}}{(\log x)^A}}.$$

This in turn is bounded by

$$\frac{(\log x)^{\tau(n)+A+1}}{\sqrt{x}},$$

which tends to 0 as $x \to \infty$. We may therefore write

$$\mathcal{M} = \sum_{m=1}^{\infty} \frac{\mu(m)}{m} \frac{\rho_n(m)}{\phi(m)} - \sum_{m>\sqrt{x}/(\log x)^A} \frac{\mu(m)}{m} \frac{\rho_n(m)}{\phi(m)},$$

which, by the preceding bound, equals

$$\sum_{m=1}^{\infty} \frac{\mu(m)}{m} \frac{\rho_n(m)}{\phi(m)} + O_n\left(\frac{(\log x)^{\tau(n)+A+1}}{\sqrt{x}}\right).$$

Notice that the function $\frac{\mu(m)}{m}\frac{\rho_n(m)}{\phi(m)}$ is multiplicative, being the product of multiplicative functions. We may thus use Euler products to deduce that

$$\sum_{m=1}^{\infty} \frac{\mu(m)}{m} \frac{\rho_n(m)}{\phi(m)} = \prod_p \left(\sum_{k=0}^{\infty} \frac{\mu\left(p^k\right)}{p^k} \frac{\rho_n\left(p^k\right)}{\phi\left(p^k\right)} \right).$$

Recall the definition of $\mathfrak{t}_n$ (equation (A.3)), as well as the fact that $\mu\left(p^k\right) = 0$ for $k \geqslant 2$. Hence

$$\begin{aligned}\sum_{m=1}^{\infty} \frac{\mu(m)}{m} \frac{\rho_n(m)}{\phi(m)} =& \prod_p \left(1 + \frac{\mu(p)}{p} \frac{\rho_n(p)}{\phi(p)}\right) \\ =& \prod_p \left(1 - \frac{\gcd(p-1, n)}{p(p-1)}\right) \\ =& n\mathfrak{t}_n,\end{aligned}$$

where the second equality follows from Lemma A.9. □

We may now combine Lemma A.13, Lemma A.14, and Lemma A.15 to get that for any fixed $n \in \mathbb{N}$ and $A > 0$ one has

$$\Psi_n(x) = n\mathfrak{t}_n x + O_{n,A}\left(\frac{x}{(\log x)^{A-5}}\right), \tag{A.10}$$

for all $x \geqslant 2$.

Proof of Theorem A.1 (i)

Using (A.10) and the first part of Lemma A.11 with

$$b_k = \Lambda(k) \frac{\phi(k^n - 1)}{k^n - 1}$$

we get

$$\sum_{p \leqslant x} \frac{\phi(p^n - 1)}{p^n - 1} \log p = n\mathfrak{t}_n x + O_{n,A}\left(\frac{x}{(\log x)^{A-5}}\right). \tag{A.11}$$

Inserting (A.11) into the second part of Lemma A.11 with $b_k = \frac{\phi(k^n-1)}{k^n-1}$ gives

$$\sum_{p \leqslant x} \frac{\phi(p^n - 1)}{p^n - 1} = n\mathfrak{t}_n \operatorname{li}(x) + O_{n,A}\left(\frac{x}{(\log x)^{A-5}}\right). \tag{A.12}$$

Using the first part of Lemma A.11 with $b_k = n\mathfrak{t}_n(k)$ gives

$$\sum_{q \leqslant x} \mathfrak{t}_n(q) = \frac{1}{n} \sum_{p \leqslant x} \frac{\phi(p^n - 1)}{p^n - 1} + O_n\left(\frac{\sqrt{x}}{\log x}\right),$$

which, when combined with (A.12), yields

$$\sum_{q\leqslant x} \mathfrak{t}_n(q) = \mathfrak{t}_n \operatorname{li}(x) + O_{n,A}\left(\frac{x}{(\log x)^{A-5}}\right). \tag{A.13}$$

Recall that $Q(x)$ denotes the number of prime powers that are at most x. To finish the proof it suffices to notice that by the Prime Number Theorem in the form

$$\pi(x) = \operatorname{li}(x) + O_A\left(\frac{x}{(\log x)^A}\right)$$

and the first part of Lemma A.11 with $b_k = 1$, we have

$$Q(x) = \pi(x) + O\left(\frac{\sqrt{x}}{\log x}\right), \tag{A.14}$$

which implies that

$$Q(x) = \operatorname{li}(x) + O_A\left(\frac{x}{(\log x)^A}\right).$$

Inserting this in (A.13) and using $\operatorname{li}(x) \gg \frac{x}{\log x}$, valid for all $x \geqslant 3$, yields

$$\frac{1}{Q(x)} \sum_{q\leqslant x} \mathfrak{t}_n(q) = \mathfrak{t}_n + O_{n,A}\left(\frac{1}{(\log x)^{A-6}}\right), \tag{A.15}$$

for all $A > 0$ and $x \geqslant 3$. Since once is allowed to use any positive value for A, we can use $A + 6$ instead. This may increase the dependence of the implied constant on A but doesn't affect the validity of Theorem A.1. We therefore conclude that for any positive A, the error term is $O_{n,A}\left(\frac{1}{(\log x)^A}\right)$, thus concluding the proof. □

The error term in (A.15) can be improved conditionally on the Generalised Riemann Hypothesis. Indeed, if one is to assume GRH for all L-functions of any modulus then one can obtain

$$\psi(x; m, a) = \frac{x}{\phi(m)} + O\left(\sqrt{x}(\log x)^2\right),$$

with an absolute implied constant, for all $m \geqslant 1$, $x \geqslant 3$, as shown in [MV07, Corollary 13.8]. Using this in place of Theorem A.7 one can follow the steps in the proof of Theorem A.1 to prove

$$\frac{1}{Q(x)}\sum_{q\leqslant x}\mathfrak{t}_n(q)=\mathfrak{t}_n+O_n\left(\frac{(\log x)^{\tau(n)+2}}{\sqrt{x}}\right),$$

for all $x\geqslant 3$. The following lemma essentially contains the proof of both part (ii) and (iii) of Theorem A.1. It is proved via introducing multiplicative indices in the sum and then applying Lemma A.12.

Lemma A.16 *For all naturals $r\geqslant 1$ and for all $x\in\mathbb{R}_{>1}$ one has*

$$\sum_{n\leqslant x}\frac{\phi(p^{rn}-1)}{p^{rn}-1}=\mathfrak{t}(p,r)x+O\left(\log\left(xr\log p\right)\right).$$

Proof. In view of (A.8) we can write

$$\sum_{n\leqslant x}\frac{\phi(p^{rn}-1)}{p^{rn}-1}=\sum_{m\leqslant p^{rx}}{}^{*}\frac{\mu(m)}{m}\sum_{\substack{n\leqslant x\\ p^{rn}\equiv 1(\mathrm{mod}\ m)}}1.$$

The condition $p^{rn}\equiv 1(\mathrm{mod}\ m)$ is equivalent to

$$\frac{\ell_p(m)}{\gcd\left(\ell_p(m),r\right)}\bigg| n.$$

Grouping terms according to the value of the order, the double sum in the right-hand-side of the above equation is seen to equal

$$\sum_{k\leqslant rx}\left[\frac{x}{k}\gcd(k,r)\right]\left(\sum_{\substack{m\in\mathbb{N}\\ \ell_p(m)=k}}{}^{*}\frac{\mu(m)}{m}\right).$$

Using $[t]=t+O(1)$, valid for all $t\in\mathbb{R}_{\geqslant 0}$, we see that this is

$$x\sum_{k\leqslant rx}\frac{\gcd(k,r)}{k}\left(\sum_{\substack{m\in\mathbb{N}\\ \ell_p(m)=k}}{}^{*}\frac{\mu(m)}{m}\right)+O\left(\sum_{k\leqslant rx}\sum_{\substack{m\in\mathbb{N}\\ \ell_p(m)=k}}{}^{*}\frac{1}{m}\right).$$

Part (i) of Lemma A.12 implies that the error term is $\ll \log(rx\log p)$. Recall the definition of $\mathfrak{t}(p,r)$, stated in (A.2). The inequality $\gcd(k,r)\leqslant r$ implies that the main term above equals

$$x\mathfrak{t}(p,r)+O\left(xr\sum_{k>rx}\frac{1}{k}\sideset{}{^*}\sum_{\substack{m\in\mathbb{N}\\ \ell_p(m)=k}}\frac{1}{m}\right).$$

Using part (ii) of Lemma A.12 to handle the above error term concludes our proof. □

Proof of Theorem A.1 (ii)

Recall the definition of $\mathfrak{t}_n(p^r)$ in equation (A.1). Using Lemma A.16 yields

$$\begin{aligned}\sum_{r\leqslant x}\mathfrak{t}_n(p^r)&=\frac{1}{n}\sum_{r\leqslant x}\frac{\phi(p^{rn}-1)}{p^{rn}-1}\\ &=\frac{\mathfrak{t}(p,n)}{n}x+O\left(\frac{\log(xn\log p)}{n}\right),\end{aligned}$$

which proves the assertion of Theorem A.1 (ii). □

Proof of Theorem A.1 (iii)

Define for $x\geqslant 1$, $r\in\mathbb{N}$, and p a prime

$$E(x,p,r):=\sum_{n\leqslant x}\frac{\phi(p^{rn}-1)}{p^{rn}-1}-x\mathfrak{t}(p,r),$$

so that Lemma A.16 is equivalent to

$$E(x,p,r)\ll\log(xr\log p). \tag{A.16}$$

Using partial summation one sees that for any $x\geqslant 1$,

$$\sum_{n\leqslant x}\mathfrak{t}_n(p^r)=\frac{x\,\mathfrak{t}(p,r)+E(x,p,r)}{x}+\int_1^x\frac{u\,\mathfrak{t}(p,r)+E(u,p,r)}{u^2}\mathrm{d}u.$$

Part (iii) of Lemma A.12 combined with (A.16) shows that this equals

$$\mathfrak{t}(p,r)\log x + O(\tau(r)\log(r\log p)).$$

The proof is complete. □

A.6 Proof of Theorem A.3

We begin by stating two definitions that we will adhere to during the ensuing proofs.

Definition A.17 *A function $f:\mathbb{N}\to\mathbb{R}$ is called strongly-additive if it satisfies*

$$f\left(\prod_{p^\lambda \| n} p^\lambda\right) = \sum_{p^\lambda \| n} f(p)$$

for all $n\in\mathbb{N}$.

Definition A.18 *Let b_n be a sequence of real numbers and $x,z\in\mathbb{R}$. We define the frequencies*

$$\nu_x\left(p; b_p \leqslant z\right) := \frac{|\{p\leqslant x : p \text{ is prime}, b_p\leqslant z\}|}{\pi(x)},$$

where $\pi(x)$ is the number of primes below x.

Let f be a strongly-additive function and let $g\in\mathbb{Z}[x]$. Define for each $m\in\mathbb{N}$

$$\rho_g(m) := |\{a\in[0,m) : g(a)\equiv 0(\mathrm{mod}\ m),\ \gcd(a,m)=1\}|$$

and notice that this is a generalisation of the function ρ_n defined at the beginning of section A.4. The next theorem can be found in [TF69].

Theorem A.19 *Let f and g be as above. Assume that $\rho_g(p)f(p)\to 0$ as $p\to\infty$ and that each of the following three series converges*

$$\sum_{|f(p)|>1}\frac{\rho_g(p)}{p-1},\ \sum_{|f(p)|\leqslant 1}\frac{\rho_g(p)f(p)}{p-1},\ \sum_{|f(p)|\leqslant 1}\frac{\rho_g(p)f^2(p)}{p-1}.$$

Then the frequencies $\nu_x(p; f(|g(p)|) \leqslant z)$ converge to a limiting distribution as $x \to \infty$. Furthermore, the said limiting distribution is continuous if and only if the series

$$\sum_{f(p)\neq 0} \frac{\rho_g(p)}{p-1}$$

diverges.

Proof of Theorem A.3 (i)

We shall use Theorem A.19 to prove that for fixed $n \in \mathbb{N}$ the frequencies

$$\nu_x\left(p; \frac{\phi(p^n-1)}{p^n-1} \leqslant z\right)$$

converge to a limiting distribution as $x \to \infty$. Define the strongly-additive function $f(k) := \log \frac{\phi(k)}{k}$ and notice that the Taylor expansion of the logarithm implies that for any prime p,

$$f(p) = \log\left(1 - \frac{1}{p}\right) = -\frac{1}{p} + O\left(\frac{1}{p^2}\right).$$

Define the polynomial $g_n(x) := x^n - 1$ and notice that by Lemma A.9

$$\rho_{g_n}(p) = \gcd(p-1, n) \leqslant n$$

for all primes p. Therefore

$$\rho_{g_n}(p) f(p) \ll \frac{n}{p} \to 0$$

as $p \to \infty$. We proceed to show that each of the three series in Theorem A.19 converge. First, notice that $f(p) \in [-\log 2, 0)$ for all primes p, hence the first series contains no terms. Regarding the second series one has

$$\begin{aligned}\sum_{|f(p)|\leqslant 1}\left|\frac{\rho_{g_n}(p)}{p-1}f(p)\right| &= \sum_p \frac{\gcd(n,p-1)}{p-1}\left(\frac{1}{p}+O\left(\frac{1}{p^2}\right)\right)\\ &\ll n\sum_p \frac{1}{p^2}\\ &< \infty,\end{aligned}$$

thus the series is convergent. Similarly

$$\begin{aligned}\sum_{|f(p)|\leqslant 1}\left|\frac{\rho_{g_n}(p)}{p-1}\,f^2(p)\right| &\ll n\sum_p \frac{1}{p^3}\\ &< \infty,\end{aligned}$$

and the third series converges as well. To conclude the proof of the theorem we observe that the limiting distribution is continuous due to

$$\begin{aligned}\sum_{|f(p)|\neq 0}\left|\frac{\rho_{g_n}(p)}{p-1}\right| &= \sum_p \frac{\gcd(n,p-1)}{p-1}\\ &\geqslant \sum_p \frac{1}{p-1}\\ &= \infty.\end{aligned}$$

Now notice that (A.14) implies that for each $x \geqslant 1$, $z \in \mathbb{R}$,

$$\nu_x\left(p;\frac{\phi(p^n-1)}{p^n-1}\leqslant z\right) = \nu_x\left(q;\frac{\phi(q^n-1)}{q^n-1}\leqslant z\right)+O\left(\frac{1}{\sqrt{x}}\right).$$

Thus letting $x \to \infty$ shows that the limiting distribution of

$$\nu_x\left(p;\frac{\phi(p^n-1)}{p^n-1}\leqslant z\right)$$

is equal to the limiting distribution of

$$\nu_x\left(q;\frac{\phi(q^n-1)}{q^n-1}\leqslant z\right).$$

The proof is now complete. □

The following theorem, stated in [Ten95, Chapter III.2, Theorem 2], provides a general criterion that ensures the existence of a limiting distribution. Before presenting it we need the following definition.

Definition A.20 *Let $A \subseteq \mathbb{N}$. The density of A is defined as*

$$\mathbf{d}(A) := \lim_{x\to\infty} \frac{|\{n \leqslant x : n \in A\}|}{x},$$

provided that the limit exists, and the upper density of A as

$$\overline{\mathbf{d}}(A) := \limsup_{x\to\infty} \frac{|\{n \leqslant x : n \in A\}|}{x}.$$

Theorem A.21 *Let $f : \mathbb{N} \to \mathbb{R}$ be a function and suppose that for any $\varepsilon > 0$ there exists a function $\alpha_\varepsilon(n) : \mathbb{N} \to \mathbb{N}$ having the following properties:*

(i) $\lim_{\varepsilon\to 0} \limsup_{T\to\infty} \overline{\mathbf{d}}\{n : \alpha_\varepsilon(n) > T\} = 0,$

(ii) $\lim_{\varepsilon\to 0} \overline{\mathbf{d}}\{n : |f(n) - f(\alpha_\varepsilon(n))| > \varepsilon\} = 0,$ *and*

(iii) *for each $\alpha \geqslant 1$ the density $\mathbf{d}\{n : \alpha_\varepsilon(n) = \alpha\}$ exists.*

Then the frequencies $\nu_x(n; f(n) \leqslant z)$ converge to a limiting distribution as $x \to \infty$.

We shall use this theorem to prove the following lemma.

Lemma A.22 *Fix $c \in \mathbb{N}$ and a prime η. Then the frequencies*

$$\nu_x\left(k; \frac{\phi(\eta^{ck}-1)}{\eta^{ck}-1} \leqslant \omega\right)$$

converge to a limiting distribution as $x \to \infty$.

Proof. Define for $k \in \mathbb{N}$

$$f(k) := \log \frac{\phi(\eta^{ck}-1)}{\eta^{ck}-1} = \sum_{p \mid \eta^{ck}-1} \log\left(1 - \frac{1}{p}\right)$$

and for each fixed $\varepsilon > 0$

$$\alpha_\varepsilon(k) := \prod_{\substack{p^\lambda \| k \\ p \leqslant y}} p^\lambda$$

where

$$y := y(\varepsilon) = \max\left\{\eta, \exp\left(\frac{c \log \eta}{\varepsilon^2}\right)\right\}.$$

It is not difficult to verify that with this choice of y the validity of $y > \varepsilon^{-2}$ is guaranteed, so that properties (i) and (iii) of Theorem A.21 hold as in [Ten95, Ex. 1, p. 295]. In order to verify the validity of property (ii), let us begin by noticing that since $\alpha_\varepsilon(k)$ is a divisor of k, we get by the inequality $\log\left(1 - \frac{1}{p}\right) \ll \frac{1}{p}$, valid for each prime p, that

$$|f(k) - f(\alpha_\varepsilon(k))| \ll \sum_{\substack{p | \eta^{ck} - 1 \\ p > y}} \frac{1}{p}. \tag{A.17}$$

Using (A.17) and the fact that $\ell_\eta(p) \mid ck$ implies

$$\frac{\ell_\eta(p)}{\gcd(\ell_\eta(p), c)} \Bigg| k,$$

we deduce that

$$\begin{aligned}\sum_{k \leqslant x} |f(k) - f(\alpha_\varepsilon(k))| &\ll \sum_{p \in (y, \eta^{cx})} \frac{1}{p} \left|\left\{k \leqslant x : \ell_\eta(p) \mid ck\right\}\right| \\ &\leqslant x \sum_{p > y} \frac{1}{p} \frac{\gcd\left(\ell_\eta(p), c\right)}{\ell_\eta(p)}.\end{aligned}$$

In light of the inequalities $\gcd\left(\ell_\eta(p), c\right) \leqslant c$ and $\log p < l_\eta(p) \log \eta$, the last expression is seen to be at most

$$c x \log \eta \sum_{p > y} \frac{1}{p \log p}.$$

Now Lemma A.8 for $a = m = 1$ combined with partial summation implies that

$$\begin{aligned}\sum_{p > y} \frac{1}{p \log p} &= -\frac{\log \log y + O(1)}{\log y} + \int_y^\infty \frac{\log \log t + O(1)}{t \log^2 t} \mathrm{d}t \\ &\ll \frac{1}{\log y},\end{aligned}$$

which shows that

$$\sum_{k\leqslant x}|f(k)-f(\alpha_\varepsilon(k))| \ll \frac{cx\log\eta}{\log y}.$$

We may now use this inequality to deduce that

$$\begin{aligned}\frac{1}{x}|\{k\leqslant x : |f(k)-f(\alpha_\varepsilon(k))|>\varepsilon\}| &\leqslant \sum_{k\leqslant x}\frac{|f(k)-f(\alpha_\varepsilon(k))|}{\varepsilon x}\\ &\ll \frac{c\log\eta}{\varepsilon\log y}\leqslant\varepsilon,\end{aligned}$$

by the definition of y. This establishes the validity of property (ii) and therefore Theorem A.21 applies and shows that the frequencies

$$\nu_x\left(k;\frac{\phi(\eta^{ck}-1)}{\eta^{ck}-1}\leqslant\omega\right)$$

converge to a limiting distribution as $x\to\infty$. □

Now the proof of part (ii) of Theorem A.3 follows by setting $\eta = p, c = n, k = r$ and $\omega = zn$ in Lemma A.22. The proof of part (iii) of Theorem A.3 follows by setting $\eta = p, c = r, k = n$ and $\omega = z$ in the same lemma.

B

GAP code

Here we collect pieces of GAP code which can be used to compute various quantities related to permutability. We start with the subgroup permutability degree. First we need to load the Permut package of Ballester-Bolinches et al. [BBCLER13].

```
gap> LoadPackage("Permut");
Loading  FORMAT 1.3 (Formations of Finite Soluble Groups)
by Bettina Eick (http://www.icm.tu-bs.de/~beick) and
   Charles R.B. Wright (http://www.uoregon.edu/~wright).
Homepage: http://www.uoregon.edu/~wright/RESEARCH/format/
---------------------------------------------------------------------
# Loading the GAP package ``permut'' in version 1.01
# (a package to deal with permutability in finite groups)
# by Adolfo Ballester-Bolinches <Adolfo.Ballester@uv.es>,
#    Enric Cosme-Ll\'opez <Enric.Cosme@uv.es>,
#    and Ramon Esteban-Romero <Ramon.Esteban@uv.es> /
#                             <resteban@mat.upv.es>.
#
#    Use ``?permut:'' for help.
---------------------------------------------------------------------
true
gap>
```

This package contains, among others, the command ArePermutableSubgroups which tests for permutability between two subgroups of a group, and so we can use this command to calculate the subgroup permutability degree of a finite group G. Note, however, that testing for subgroup permutability between two subgroups is not difficult. For example, the command permute defined in the following manner

```
permute := function(H,K)
   if Size(ClosureGroup(H,K))*Size(Intersection(H,K))
   =Size(H)*Size(K) then
      return true;
   fi;
end;
```

accepts two subgroups H, K of a group G as arguments and, when called, returns true if $H \operatorname{per} K$. This is basically what the command ArePermutableSubgroups does, but it also provides for special cases where the full test by order considerations is not necessary. The following snippet of code accepts a finite group G as argument and, when called, returns $\mathfrak{p}(G)$ as a float.

```
spd := function(G)
   local allsubs, count, H, K;
   allsubs := Flat(List(ConjugacyClassesSubgroups(G),Elements));
   count:=0;
      for H in allsubs do;
         for K in allsubs do;
            if ArePermutableSubgroups(H,K) then
               count := count+1;
            fi;
         od;
      od;
   return Float(count/(Size(allsubs)^2));
end;
```

We may then ask for the subgroup permutability degree of, say, A_6.

```
gap> spd(AlternatingGroup(6));
0.0833901
gap> time; # in milliseconds
277900
```

In fact, we can improve the above algorithm by iterating over representatives of the conjugacy classes of subgroups in one of the two for loops. Since $H \operatorname{per} K$ if and only if $H^g \operatorname{per} K^g$, it follows that $|\operatorname{Per}(H)| = |\operatorname{Per}(H^g)|$ for all $g \in G$. We should point out, however, that we cannot iterate over representatives of the conjugacy classes of subgroups in both `for` loops, because $H \operatorname{per} K$ does not imply that $H \operatorname{per} K^g$, nor vice-versa.

```
spd := function(G)
local allsubs, subreps, count, H, K;
subreps := Flat(List(ConjugacyClassesSubgroups(G),Representative));
allsubs := Flat(List(ConjugacyClassesSubgroups(G),Elements));
count:=0;
   for H in subreps do;
      for K in allsubs do;
         if ArePermutableSubgroups(H,K) then
            count := count+Index(G,Normalizer(G,H));
         fi;
      od;
   od;
return Float(count/(Size(allsubs)^2));
end;
```

We ask GAP to compute the subgroup permutability degree of A_6 again

```
gap> spd(AlternatingGroup(6));
0.0833901
gap> time; # in milliseconds
21416
```

and we notice a significant time gain by a factor of almost 13, which is explained by the fact that in A_6 (more generally in every finite simple group) the size of each conjugacy class of subgroups is never too small. If, however, the group in question is close in some sense to Hamiltonian then the difference between the two versions of the algorithm should be negligible.

Next we address the construction of the subgroup permutability graph of a finite group as defined by Bianchi et al. We can use the full strength of the GRAPE package of Leonard Soicher to define our graph. First we load GRAPE

```
gap> LoadPackage("GRAPE");
--------------------------------------------------------------------
```

```
Loading  GRAPE 4.6.1 (GRaph Algorithms using PErmutation groups)
by Leonard H. Soicher (http://www.maths.qmul.ac.uk/~leonard/).
Homepage: http://www.maths.qmul.ac.uk/~leonard/grape/
-----------------------------------------------------------------
true
```

and we construct a list of all nonnormal subgroups of the group G

```
nns := Concatenation(List(Filtered(ConjugacyClassesSubgroups(G),
       c -> Size(c)>1), AsList));;
```

We then define our graph Γ via

```
gamma := Graph(G,nns,OnPoints,ArePermutableSubgroups,true);
```

Note that this construction assumes that we have already loaded the Permut package.

Let us also present a method to obtain the adjacency matrix of the subgroup permutability graph of a group G by recording instances of permutability between two members of nns in a symmetric (0,1) permutability matrix A.

```
PermutabilityMatrix := function(nns)
   local A, i, j;
   A := IdentityMat(Length(nns));
      for i in [1..Length(nns)-1] do;
         for j in [i+1..Length(nns)] do;
            if ArePermutableSubgroups(nns[i],nns[j]) then
               A[i][j]:=1; A[j][i]:=1;
            fi;
         od;
      od;
   return A;
end;
```

Bibliography

[AC09] D.D. Anderson and V. Camillo, *Subgroups of direct products of groups, ideals and subrings of direct products of rings, and Goursat's lemma*, Rings, modules and representations. International conference on rings and things in honor of Carl Faith and Barbara Osofsky, Zanesville, OH, USA, June 15–17, 2007, American Mathematical Society, 2009, pp. 1–12.

[Aiv13] S. Aivazidis, *The subgroup permutability degree of projective special linear groups over fields of even characteristic*, J. Group Theory **16** (2013), no. 3, 383–396.

[Aiv15] ______, *On the subgroup permutability degree of the simple Suzuki groups*, Monatsh. Math. **176** (2015), no. 3, 335–358.

[AS14] S. Aivazidis and E. Sofos, *On the distribution of the density of maximal order elements in general linear groups*, Ramanujan J. (2014), 1–25.

[BBCLER13] A. Ballester-Bolinches, E. Cosme-Llópez, and R. Esteban-Romero, *Algorithms for permutability in finite groups*, Cent. Eur. J. Math. **11** (2013), no. 11, 1914–1922.

[BBERA10] A. Ballester-Bolinches, R. Esteban-Romero, and M. Asaad, *Products of finite groups*, de Gruyter Expositions in Mathematics, vol. 53, Walter de Gruyter GmbH & Co. KG, Berlin, 2010.

[BDM96] N. Blackburn, M. Deaconescu, and A. Mann, *Finite equilibrated groups*, Math. Proc. Cambridge Philos. Soc. **120** (1996), no. 4, 579–588.

[Bec65] H. Bechtell, *Reduced partial products*, Amer. Math. Monthly **72** (1965), 881–882.

[BGBMV95] M. Bianchi, A. Gillio Berta Mauri, and L. Verardi, *Finite groups and subgroup-permutability*, Ann. Mat. Pura Appl. (4) **169** (1995), 251–268.

[BJ11] Y. Berkovich and Z. Janko, *Groups of Prime Power Order. Vol. 3*, de Gruyter Expositions in Mathematics, vol. 56, Walter de Gruyter GmbH & Co. KG, Berlin, 2011.

[Cav75] S.R. Cavior, *The subgroups of the dihedral group*, Math. Mag. **48** (1975), 107.

[CNW90] F. Celler, J. Neubüser, and C. Wright, *Some remarks on the computation of complements and normalizers in soluble groups*, Acta Appl. Math. **21** (1990), no. 1, 57–76.

[Dav00] H. Davenport, *Multiplicative Number Theory*, third ed., Graduate Texts in Mathematics, vol. 74, Springer-Verlag, New York, 2000.

[Ded97] R. Dedekind, *Über Gruppen, deren sämmtliche Theiler Normaltheiler sind*, Math. Ann. **48** (1897), no. 4, 548–561.

[Des63] W. E. Deskins, *On quasinormal subgroups of finite groups*, Math. Z. **82** (1963), 125–132.

[DFO13] Alla Detinko, Dane Flannery, and Eamonn O'Brien (eds.), *Probabilistic group theory, combinatorics, and computing*, Lecture Notes in Mathematics, vol. 2070, Springer, London, 2013, Lectures from the 5th de Brún Workshop "Groups, Combinatorics, Computing" held at the National University of Ireland, Galway, April 11–16, 2011.

[DH92] K. Doerk and T. Hawkes, *Finite soluble groups*, de Gruyter Expositions in Mathematics, vol. 4, Walter de Gruyter & Co., Berlin, 1992.

[DH12] J.-M. Deshouillers and M. Hassani, *A note on the distribution function of $\phi(p-1)/(p-1)$*, J. Aust. Math. Soc. **93** (2012), no. 1-2, 77–83.

[Dic03] L. E. Dickson, *Linear groups with an exposition of galois field theory*, Dover Publications, 2003.

[Dix69] J. D. Dixon, *The probability of generating the symmetric group*, Math. Z. **110** (1969), 199–205.

[Dix02] ———, *Probabilistic group theory*, C. R. Math. Acad. Sci. Soc. R. Can. **24** (2002), no. 1, 1–15.

[Doe94] K. Doerk, *Über endliche auflösbare Gruppen, die sich gegenüber der Frattinigruppe wie nilpotente Gruppen verhalten*, J. Algebra **167** (1994), no. 2, 533–537.

[Ebe14] S. Eberhard, *Commuting probabilities of finite groups*, ArXiv e-prints (2014).

[Erd39] P. Erdős, *On the smoothness of the asymptotic distribution of additive arithmetical functions*, Am. J. Math. **61** (1939), 722–725.

[Erd51] ______, *On some problems of Bellman and a theorem of Romanoff*, J. Chinese Math. Soc. (N.S.) **1** (1951), 409–421.

[ET65] P. Erdős and P. Turán, *On some problems of a statistical group-theory. I*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **4** (1965), 175–186 (1965).

[ET67] P. Erdős and P. Turán, *On some problems of a statistical group-theory. II*, Acta math. Acad. Sci. Hungar. **18** (1967), 151–163.

[Gas53] W. Gaschütz, *Über die Φ-Untergruppe endlicher Gruppen*, Math. Z. **58** (1953), 160–170.

[Gas62] ______, *Praefrattinigruppen*, Arch. Math. (Basel) **13** (1962), 418–426.

[Gas63] ______, *Zur Theorie der endlichen auflösbaren Gruppen*, Math. Z. **80** (1962/1963), 300–305.

[Gou89] E. Goursat, *Sur les substitutions orthogonales et les divisions regulières de l'espace*, Ann. Sci. Éc. Norm. Supér. (3) **6** (1889), 9–102.

[GR69] J. Goldman and G.-C. Rota, *The Number of subspaces of a vector space*, Recent Prog. Comb., Proc. 3rd Waterloo Conf. 1968, 75-83 (1969), 1969.

[GR06] R. M. Guralnick and G. R. Robinson, *On the commuting probability in finite groups*, J. Algebra **300** (2006), no. 2, 509–528.

[Gus73] W. H. Gustafson, *What is the probability that two group elements commute?*, Amer. Math. Monthly **80** (1973), 1031–1034.

[Hal37] P. Hall, *Complemented groups*, J. Lond. Math. Soc. **12** (1937), 201–204.

[HB82] B. Huppert and N. Blackburn, *Finite groups. II*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 242, Springer-Verlag, Berlin, 1982.

[Heg13] P. Hegarty, *Limit points in the range of the commuting probability function on finite groups*, J. Group Theory **16** (2013), no. 2, 235–247.

[Hig63] G. Higman, *Suzuki* 2*-groups*, Illinois J. Math. **7** (1963), 79–96.

[Hul99] A. Hulpke, *Computing subgroups invariant under a set of automorphisms*, J. Symbolic Comput. **27** (1999), no. 4, 415–427.

[Isa08] I. M. Isaacs, *Finite Group Theory*, Graduate Studies in Mathematics, vol. 92, American Mathematical Society, Providence, RI, 2008.

[Isa14] ———, private communication, 2014.

[Iwa41] K. Iwasawa, *Über die endlichen Gruppen und die Verbände ihrer Untergruppen*, J. Fac. Sci. Imp. Univ. Tokyo. Sect. I. **4** (1941), 171–199.

[Jos77] K. S. Joseph, *Research Problems: Several Conjectures on Commutativity in Algebraic Structures*, Amer. Math. Monthly **84** (1977), no. 7, 550–551.

[Kat68] I. Katai, *On distribution of arithmetical functions on the set prime plus one*, Compos. Math. **19** (1968), 278–289.

[KC02] V. G. Kac and P. Cheung, *Quantum Calculus*, Springer Verlag, 2002.

[Keg62] O. H. Kegel, *Sylow-Gruppen und Subnormalteiler endlicher Gruppen*, Math. Z. **78** (1962), 205–221.

[KS04] H. Kurzweil and B. Stellmacher, *The theory of finite groups*, Universitext, Springer-Verlag, New York, 2004, An introduction, Translated from the 1998 German original.

[Li98] S. Li, *On the number of elements with maximal order in the multiplicative group modulo n*, Acta Arith. **86** (1998), no. 2, 113–132.

[LN94] R. Lidl and H. Niederreiter, *Introduction to Finite Fields and their Applications*, first ed., Cambridge University Press, Cambridge, 1994.

[LNY14] P. Lescot, H. N. Nguyen, and Y. Yang, *On the commuting probability and supersolvability of finite groups*, Monatsh. Math. **174** (2014), no. 4, 567–576.

[Lon82] P. Longobardi, *Gruppi finite a fattoriali modulari*, Note Mat. **2** (1982), no. 1, 73–100.

[LP02] S. Li and C. Pomerance, *Primitive roots: a survey*, 219–231.

[MV07] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007.

[Nap70] F. Napolitani, *Modularità e distributività nell'insieme dei sottogruppi subnormali*, Rend. Sem. Mat. Univ. Padova **43** (1970), 215–220.

[Net64] E. Netto, *The theory of substitutions and its applications to algebra*, Second edition. Revised by the Author and translated with his permission by F. N. Cole, Chelsea Publishing Co., New York, 1964.

[Nou82] Z. Nouacer, *Caractères et sous-groupes des groupes de Suzuki*, Diagrammes **8** (1982), ZN1–ZN29.

[Ore39] O. Ore, *Contributions to the theory of groups of finite order*, Duke Math. J. **5** (1939), no. 2, 431–460.

[Pil41] S.S. Pillai, *On the sum function connected with primitive roots*, Proc. Indian Acad. Sci., Sect. A **13** (1941), 526–529.

[Rei61] I. Reiner, *On the number of matrices with given characteristic polynomial.*, Ill. J. Math. **5** (1961), 324–329.

[Rob99] D. J. S. Robinson, *Permutability properties of subgroups*, London Math. Soc. Lecture Note Ser., vol. 261, pp. 633–638, Cambridge Univ. Press, Cambridge, 1999.

[Rot95] J. J. Rotman, *An Introduction to the Theory of Groups*, fourth ed., Graduate Texts in Mathematics, vol. 148, Springer-Verlag, New York, 1995.

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

[Sch28] I. Schoenberg, *Über die asymptotische Verteilung reeller Zahlen mod 1*, Math. Z. **28** (1928), 171–199.

[Sch36] I.J. Schoenberg, *On asymptotic distributions of arithmetical functions*, Trans. Am. Math. Soc. **39** (1936), 315–330.

[Sch94] R. Schmidt, *Subgroup Lattices of Groups*, vol. 14, de Gruyter, 1994.

[Sch98] P. Schmid, *Subgroups permutable with all Sylow subgroups*, J. Algebra **207** (1998), no. 1, 285–293.

[Shp90] I.E. Shparlinskii, *Some arithmetic properties of recurrence sequences*, Mathematical Notes of the Academy of Sciences of the USSR **47** (1990), no. 6, 612–617.

[Sil06] G. Silberberg, *Finite equilibrated 2-generated 2-groups*, Acta Math. Hungar. **110** (2006), no. 1-2, 23–35.

[Ste69] P.J. Stephens, *An average result for Artin's conjecture*, Mathematika **16** (1969), 178–188.

[Sto93] R. Stong, *The average order of a matrix*, J. Combin. Theory Ser. A **64** (1993), no. 2, 337–343.

[Suz62] M. Suzuki, *On a class of doubly transitive groups*, Ann. Math. **75** (1962), no. 2, 105–145.

[Tăr09] M. Tărnăuceanu, *Subgroup commutativity degrees of finite groups*, J. Algebra **321** (2009), no. 9, 2508–2520.

[Tăr11] ______, *Addendum to "Subgroup commutativity degrees of finite groups" [J. Algebra 321 (9) (2009) 2508–2520]*, J. Algebra **337** (2011), 363–368.

[Ten95] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge Studies in Advanced Mathematics, vol. 46, Cambridge University Press, Cambridge, 1995.

[TF69] Ž. Toleuov and A. S. Faĭnleĭb, *The distribution of values of arithmetic functions on certain subsets of a natural series*, Izv. Akad. Nauk UzSSR Ser. Fiz.-Mat. Nauk **13** (1969), no. 5, 23–27.

[Tou09] V. Toulmonde, *Comportement au voisinage de 1 de la fonction de répartition de $\varphi(n)/n$*, Int. J. Number Theory **5** (2009), no. 8, 1347–1384.

[Wil09] R. Wilson, *The Finite Simple Groups*, vol. 251, Springer, 2009.