Noname manuscript No.
(will be inserted by the editor)

Heuristics for p-class Towers of Imaginary Quadratic Fields

Nigel Boston · Michael R. Bush · Farshid Hajir

With an Appendix by Jonathan Blackhurst

Dedicated to Helmut Koch

Received:... Accepted: ... etc. to be filled in by editors

Abstract Cohen and Lenstra have given a heuristic which, for a fixed odd prime p, leads to many interesting predictions about the distribution of p-class groups of imaginary quadratic fields. We extend the Cohen-Lenstra heuristic to a non-abelian setting by considering, for each imaginary quadratic field K, the Galois group of the p-class tower of K, i.e. $G_K := \mathrm{Gal}(K_\infty/K)$ where K_∞ is the maximal unramified p-extension of K. By class field theory, the maximal abelian quotient of G_K is isomorphic to the p-class group of K. For integers $c \geq 1$, we give a heuristic of Cohen-Lenstra type for the maximal p-class c quotient of G_K and thereby give a conjectural formula for how frequently a given p-group of p-class c occurs in this manner. In particular, we predict that every finite Schur σ-group occurs as G_K for infinitely many fields K. We present numerical data in support of these conjectures.

Mathematics Subject Classification (2000) 11R29, 11R11

Keywords Cohen-Lenstra heuristics, class field tower, ideal class group, Schur σ-group

The research of the first author was supported by National Security Agency Grant MSN115460. The second author received support from several Lenfest Summer Research Grants (an internal college grant).

Nigel Boston
Department of Mathematics, University of Wisconsin - Madison, 480 Lincoln Drive, Madison, WI 53706, USA
E-mail: boston@math.wisc.edu

Michael R. Bush
Department of Mathematics, Washington and Lee University, Lexington, VA 24450, USA
E-mail: bushm@wlu.edu

Farshid Hajir
Department of Mathematics & Statistics, University of Massachusetts - Amherst, 710 N. Pleasant Street, Amherst MA 01003, USA
E-mail: hajir@math.umass.edu

1 Introduction

1.1 Cohen-Lenstra Philosophy

About 30 years ago, Cohen and Lenstra [12,13] launched a heuristic study of the distribution of class groups of number fields. To focus the discussion, we restrict to a specialized setting. Let p be an odd prime. Among the numerous insights contained in the work of Cohen and Lenstra, let us single out two and draw a distinction between them: (1) There is a natural probability distribution on the category of finite abelian p-groups for which the measure of each G is proportional to the reciprocal of the size of $\mathrm{Aut}(G)$; and (2) the distribution of the p-part of class groups of imaginary quadratic fields is the same as the Cohen-Lenstra distribution of finite abelian p-groups. The first statement, a purely group-theoretical one, is quite accessible and Cohen and Lenstra prove many beautiful facts about such distributions (not just for abelian groups viewed as $\mathbb{Z}$-modules but also more generally for modules over rings of integers of number fields) in the first part of [13]. The second, and bolder, insight is much less accessible at present but leads to striking number-theoretical predictions, only a small number of which have been proven, but all of which agree with extensive numerical data. Note that (2) quantifies the notion that the (rather elementary) necessary conditions for a group to occur as the p-part of the class group of an imaginary quadratic field - namely that it be a finite abelian p-group - should also be sufficient.

In the decades since the publication of [12,13], the application of (1) has been broadened to a number of other situations. It should be noted, however, that there are many circumstances where the weighting factor should also involve some power of the order of G. This includes recent investigations into variation of Tate-Shafarevich groups, variation of p-class tower groups (p odd) for real quadratic fields (to be described in a subsequent paper [6] by the authors) and variation in presentations of p-groups as described in [5]. The case under consideration in the current paper, however, does not involve these extra factors.

As regards the combination of (1) and (2), one can speak of a "Cohen-Lenstra strategy," perhaps, as follows. Suppose we have a sequence $G_1, G_2, \ldots$ of groups (arising as invariants attached to some kind of arithmetic objects, say). One can hope to identify a category $\mathcal{C}$ of groups in which the sequence lies and to assign to each G in $\mathcal{C}$ a positive real number $w(G)$ called its weight; we would expect the size of $\mathrm{Aut}_{\mathcal{C}}(G)$ (the set of automorphisms of G in the category $\mathcal{C}$) to appear in the denominator of $w(G)$. We set $w_{\mathcal{C}} = \sum_{G \in \mathcal{C}} w(G)$ for the total weight of $\mathcal{C}$, assumed to be a finite quantity. Suppose we also define the frequency with which any object G of $\mathcal{C}$ occurs in the sequence $(G_n)_{n \geq 1}$ to be the limit

$$\mathrm{Freq}(G) = \lim_{n \to \infty} \frac{\sum_{\nu=1}^{n} \mathrm{ch}_G(G_\nu)}{n}$$

assuming this exists. Here, $\mathrm{ch}_G(H)$ is the characteristic function of G, taking the value 1 if H is isomorphic to G (in the category $\mathcal{C}$) and 0 otherwise. The Cohen-Lenstra philosophy would then say that, assuming the sequence $(G_n)_{n \geq 1}$ is sufficiently general and the category $\mathcal{C}$ is correctly chosen, for each $G \in \mathcal{C}$ we would expect $\mathrm{Freq}(G)$ to equal the Cohen-Lenstra measure of G in the category $\mathcal{C}$, namely $w(G)/w_{\mathcal{C}}$. In such a situation, we can speak of the sequence (G_n) "obeying

a Cohen-Lenstra distribution for the category $\mathcal{C}$ equipped with the weight function w."

As just some of the examples of applications of this philosophy, we cite Cohen-Martinet [14], Wittman [31], and Boston-Ellenberg [7]. In the first two of these, the class groups are in fact studied as modules over the group ring of the Galois group. In [7], the groups under study are non-abelian, and in fact the situation is slightly different because the base field is fixed (to be $\mathbb{Q}$) and the ramifiying set varies; however the essential Cohen-Lenstra idea appears to apply in that situation also.

1.2 The Cohen-Lenstra heuristics for p-class groups

For an algebraic number field K, we let A_K be the p-Sylow subgroup of its ideal class group. If we allow K to vary over all imaginary quadratic fields, ordered according to increasing absolute value of the discriminant d_K, the groups A_K fluctuate with no immediately apparent rhyme or reason. When Cohen and Lenstra investigated their *cumulative* behavior, however, they found a surprising pattern. Namely, they asked what can be said about the frequency with which a given group would occur as A_K when the fields K are ordered by the magnitude of their discriminants. Their heuristic, described above, led them to many predictions, one of which is the following conjecture.

Conjecture 1.1 (Cohen-Lenstra) Fix a finite abelian group $G = \mathbb{Z}/p^{r_1} \times \cdots \times \mathbb{Z}/p^{r_g}$ of rank $g \geq 1$. Among the imaginary quadratic fields K such that A_K has rank g, ordered by discriminant, the probability that A_K is isomorphic to G is

$$\frac{1}{|\mathrm{Aut}(G)|}\, p^{g^2} \prod_{k=1}^{g} (1 - p^{-k})^2.$$

Remark 1.2 We provide more detail on how the above conjecture is related to the heuristic that groups should be weighted according to the inverse of the size of an appropriate automorphism group.

For a finite abelian group G, if we define the Cohen-Lenstra weight of G to be simply $w(G) = 1/|\mathrm{Aut}(G)|$, then it is a theorem of Hall [22] and, in a more general context, of Cohen-Lenstra, that the total weight w_p of all finite abelian p-groups is given by

$$w_p = \sum_H w(H) = \prod_{n \geq 1} (1 - p^{-n})^{-1},$$

where $\sum_H$ means the sum over the isomorphism classes of finite abelian p-groups. By [13, p. 56], the probability that an abelian p-group has generator rank g is given by

$$\frac{\sum_{\{H : d(H) = g\}} w(H)}{w_p} = p^{-g^2} \prod_{n \geq 1} (1 - p^{-n}) \prod_{k=1}^{g} (1 - p^{-k})^{-2}.$$

Thus, under the Cohen-Lenstra distribution, the probability that a randomly chosen abelian p-group of generator rank g is isomorphic to G is given by

$$\frac{w(G)}{\sum_{\{H : d(H) = g\}} w(H)} = \frac{1}{|\mathrm{Aut}(G)|}\, p^{g^2} \prod_{k=1}^{g} (1 - p^{-k})^2.$$

Cohen-Lenstra's fundamental heuristic assumption (2) then yields Conjecture 1.1.

1.3 Heuristics for the distribution of p-class tower groups

In this article, we continue to assume that p is odd and consider a non-abelian extension of the number-theoretical objects studied by Cohen and Lenstra, passing from the p-part of the class group of a number field K to the pro-p fundamental group of the ring of integers of K, namely the Galois group of its maximal everywhere unramified p-extension. For brevity, henceforth we will refer to these groups as "p-class tower groups." The key fact, as pointed out in Koch-Venkov [24], is that p-class tower groups of imaginary quadratic fields (and certain of their quotients) must satisfy a "Schur σ" condition; the precise definitions are given below.

To each finite Schur σ-group, or more generally to each maximal p-class c quotient of any Schur σ-group, we attach a rational number we call its measure; in the spirit of the Friedman-Washington [17] approach to the Cohen-Lenstra heuristics, this measure is given by a count of how likely it is for a randomly chosen set of relations of a certain type to define the given group. Our main heuristic assumption then, is that for the sequence of p-class tower groups of imaginary quadratic fields, ordered by discriminant, or more generally for the sequence of maximal p-class c quotients of these p-class tower groups (where c is any fixed whole number), the frequency of any given group equals the measure of the group.

To describe our situation in more detail, we specify some notation to be used throughout the paper. For a pro-p group G, we write

$$d(G) = \dim_{\mathbb{Z}/p\mathbb{Z}} H^1(G, \mathbb{Z}/p\mathbb{Z}), \qquad r(G) = \dim_{\mathbb{Z}/p\mathbb{Z}} H^2(G, \mathbb{Z}/p\mathbb{Z}),$$

where the action of G on $\mathbb{Z}/p\mathbb{Z}$ is trivial. These invariants give, respectively, the generator rank and relation rank of G as a pro-p group. The Frattini subgroup of G, denoted $\Phi(G)$, is defined to be the closure of $[G,G]G^p$. The groups $G^{\mathrm{ab}} = G/\overline{[G,G]}$ and $G/\Phi(G)$ are, respectively, the maximal abelian quotient and maximal exponent-p abelian quotient of G.

To describe how we pass to a non-abelian generalization, recall that if K_1 is the p-Hilbert class field of K, defined to be its maximal abelian unramified p-extension, then there is a canonical isomorphism $A_K \to \mathrm{Gal}(K_1/K)$ given by the Artin reciprocity map. Now, let us consider the field K_∞ obtained by taking the compositum of *all* finite unramified p-extensions of K, not just the abelian ones. We put $G_K = \mathrm{Gal}(K_\infty/K)$. It is clear that the maximal abelian quotient of G_K is isomorphic to A_K and by Burnside $d(G_K) = d(A_K)$.

The central question we consider in this work is: For a fixed odd prime p, as K varies over all imaginary quadratic fields of ascending absolute value of discriminant, what can one say about the variation of the groups G_K?

Naturally, this is a more difficult question than the variation of class groups, even for venturing a guess. Already, the group G_K is not always finite. Indeed, in [24], Koch and Venkov proved that G_K is infinite if $d(G_K) \geq 3$; they did so by taking into account all the facts they had at their disposal about the group G_K. Namely, G_K is a finitely generated pro-p group with finite abelianization and deficiency 0 (meaning that $r(G_K) - d(G_K) = 0$) and admits an automorphism of order 2 which acts as inversion on its abelianization (complex conjugation is such an automorphism, for example). Since having zero deficiency is equivalent to

having trivial Schur multiplier in this context, Koch and Venkov dubbed groups having this particular set of properties "Schur σ-groups." In Section 2, we review some of the work of Koch and Venkov on Schur σ-groups, and develop a method via counting relations, of measuring how frequently a given group occurs as the maximal p-class c quotient of Schur σ-groups.

Positing our main heuristic assumption that a finite p-group G arises as a p-class tower group over an imaginary quadratic field with the same frequency as G occurs as a randomly chosen group among Schur σ-groups, in Section 3 we arrive at the following conjecture.

Conjecture 1.3 Suppose G is a finite p-group which is a Schur σ-group of generator rank $g \geq 1$ or, more generally, suppose c is a positive integer and G is the maximal p-class c quotient of a Schur σ-group. Then, among the imaginary quadratic fields K such that A_K has rank g, ordered by discriminant, the probability that G_K (or in the fixed p-class case, the maximal p-class c quotient of G_K) is isomorphic to G is equal to

$$\frac{1}{|\mathrm{Aut}_\sigma(G)|}\, p^{g^2} \prod_{k=1}^{g} (1-p^{-k}) \prod_{k=1+g-h}^{g} (1-p^{-k}),$$

where h is the difference between the p-multiplicator rank and nuclear rank of G and $\mathrm{Aut}_\sigma(G)$ is the centralizer in $\mathrm{Aut}(G)$ of an automorphism σ of order 2, acting as inversion on the abelianization of G. We note that $0 \leq h \leq g$ and that $h = g$ for Schur σ-groups, in which case the above formula should be compared with that of Conjecture 1.1.

Remark 1.4 We should point out that the formula above has been proved assuming an additional group-theoretical conjecture, namely that all the Schur σ-groups and their maximal p-class c quotients satisfy a kernel invariance property (KIP), see Definition 2.22. This condition is discussed in more detail in Section 2.5. [1]

Remark 1.5 As in the abelian situation, Conjecture 1.3 can also be viewed as arising from an appropriate choice of weight function for the finite Schur σ-groups. If we define $w'(G) = 1/|\mathrm{Aut}_\sigma(G)|$ for a finite Schur σ-group G, then the conjecture asserts that the density of K for which G_K is isomorphic to G is equal to $w'(G)/w_p$.

Remark 1.6 We do not make a direct prediction about how frequently a given infinite Schur σ-group G occurs as a p-class tower group, but for every $c \geq 1$, the maximal p-class c quotient of G is finite and the conjecture above applies to predict the density of *all* imaginary quadratic K (including the ones where G_K is infinite) for which the maximal p-class c quotients of G_K and G coincide.

If G is a finite Schur σ-group, then its generator rank g is at most 2. The first case of Conjecture 1.3 predicts how frequently such a group occurs as a p-class tower group for imaginary quadratic fields.

[1] In recent work by Boston and Wood [10], completed during the submission process for this paper, it has been shown that this formula can be derived without the KIP assumption. This will be discussed in more detail in a subsequent paper [6] that deals with the case of real quadratic fields.

1.4 Numerical Evidence

As theoretical evidence for their conjecture, Cohen and Lenstra were able to show that a relatively cheap consequence of their heuristic assumption, namely the prediction that the average value of $3^{d_3(A_K)}$ (as K ranges over all imaginary quadratic fields) is 2, is in fact a highly non-trivial theorem of Davenport and Heilbronn [15]. In more recent work, for example see [3], Bhargava and his students have obtained deep refinements and extensions of the Davenport-Heilbronn result, in particular verifying further consequences of the Cohen-Lenstra and Cohen-Martinet conjectures.

As regards numerical evidence, class groups of imaginary quadratic fields can be computed via an efficient algorithm, and so the class group computations available to Cohen and Lenstra were quite extensive. In [13], they derived many consequences of their heuristic, every one of which matched and in some cases even "explained" the *observed* variation of the p-part of the class group of imaginary quadratic fields.

In our non-abelian situation, we do not even know an algorithm for determining whether G_K is finite, much less for computing it, so the numerical investigation of our heuristic is bound to be more tricky. One of the first examples of a computation of G_K in the literature appears in a 1934 article of Scholz and Taussky [29]: for the field $\mathbb{Q}(\sqrt{-4027})$, with $p = 3$, A_K is elementary abelian of rank 2 and the group G_K has size 243 and is isomorphic to the group denoted `SmallGroup(243,5)` in the terminology of the computer algebra software package `Magma` (see [2,4]).

The method of Boston and Leedham-Green [8] can be used for certain K to produce a short list of candidates for the isomorphism class of G_K. In general, it is difficult to identify the isomorphism type of G_K exactly except in special situations. See section 4 for more details, especially the proof of Theorem 4.3. See also [9,11] for additional discussion and applications of this method.

In order to test our heuristic hypothesis, we considered what kind of number-theoretical data (meaning about the groups G_K) was within reach, and settled on the following: we computed the class groups of unramified extensions of K of degree 1 or p. In terms of group theory, this "index $\leq p$ abelianization data" or "IPAD," describes the abelianization of G_K as well as those of its index p subgroups. Though it is impractical at present to attempt the complete computation of G_K for all fields K within a given large discriminant range, it was possible for us to compute the IPADs for over $460,000$ fields with discriminant in the range $-10^8 < d_K < 0$ and to compare the distribution of IPADs to the group-theoretical prediction.

As a summary of the numerical evidence, the second to last column of Table 5.2 in Section 5 lists the observed frequencies of the most common IPADs with $p = 3$ and $g = 2$ over all imaginary quadratic fields K with $|d_K| < 10^8$. The last column then gives the theoretical predictions based on our heuristic. Given the variability of the data and the general convergence trend toward the predicted value, we believe that the data support our conjecture.

1.5 Organization of the paper

As in [13], we have separated the group theory, where we have theorems, from the number theory, where we mostly make conjectures and collect data. We develop

some basic facts about Schur σ-groups in Section 2 and introduce various measures in both the abelian and non-abelian setting. We then relate these measures and derive formulas for them. In Section 3, we give a precise formulation of our conjecture describing the variation of Galois groups of p-towers of imaginary quadratic fields. The distribution of IPADs of Schur σ-groups is investigated in Section 4. This investigation yields a number of results which we prove using a mixture of theory and computation, thanks to the powerful technique of organizing p-groups via O'Brien's p-group generation algorithm [27]. The number-theoretical data we have collected is summarized in Section 5; see in particular, Tables 1 and 2. The computations were carried out using the symbolic algebra packages `Magma` [4] and `PARI/GP` [28]. Finally, the appendix contains a proof, by Blackhurst, of a group-theoretical fact needed in Section 2.

Acknowledgements We acknowledge useful correspondence and conversations with Bettina Eick, Jordan Ellenberg, John Labute, Daniel Mayer, Cam McLeman, Eamonn O'Brien, and Melanie Matchett Wood. We are grateful to Jonathan Blackhurst for providing the Appendix. We would also like to thank Joann Boston for drawing the figure in Section 2.

2 Schur σ-groups

2.1 Preliminaries

Let p be an odd prime.

Definition 2.1 An automorphism of a finitely generated pro-p group G is called a GI-automorphism (meaning "generator-inverting") if it has order 2 and acts as inversion on G^{ab}.

Definition 2.2 A finitely generated pro-p group G is called a Schur σ-group of rank g if it satisfies the following properties: 1) $d(G) = r(G) = g$; 2) G^{ab} is finite; 3) It has a GI-automorphism σ.

We now fix $g \geq 1$, and let F denote the free pro-p group on g generators $x_1, \ldots, x_g$. Let σ be the automorphism of F induced by the assignment $\sigma(x_i) = x_i^{-1}$ for $i = 1, \ldots, g$. Koch and Venkov [24] showed that, given a GI-automorphism σ on G, one can choose an epimorphism from F to G so that this automorphism is induced by the GI-automorphism σ on F. In particular, this means that we can find generators for G which lie in

$$X(G, \sigma) = \{s \in G \mid \sigma(s) = s^{-1}\}.$$

In addition, the relations of a Schur σ-group can always be chosen to lie in

$$X = X(\Phi(F), \sigma) = \{s \in \Phi(F) \mid \sigma(s) = s^{-1}\}.$$

Using refinements of the theorem of Golod and Shafarevich, Koch and Venkov proved that Schur σ-groups of rank $g \geq 3$ are always infinite.

In general, we will use the symbol σ to denote both the specific automorphism of F defined above and a general GI-automorphism on a group G except when there is the potential for confusion.

Suppose G is a pro-p group and σ is a GI-automorphism of G. As shown by Hall (section 1.3 of [21], although sometimes attributed to Burnside), the kernel from $\mathrm{Aut}(G) \to \mathrm{Aut}(G/\Phi(G))$ is a pro-p group and so by Schur-Zassenhaus (e.g. Prop. 1.1 of [19]), all lifts of order 2 of the inversion automorphism on $G/\Phi(G)$ are conjugate to each other. It follows that the sets $X(G,\sigma)$ and $Y(G,\sigma)$ where

$$Y(G,\sigma) = \{x \in G \mid \sigma(x) = x\}$$

are well-defined up to conjugacy and that their orders are independent of the choice of GI-automorphism σ and hence depend only on G. We will denote the order of $Y(G,\sigma)$ by $y(G)$. Also observe that $Y(G,\sigma) = Y(\Phi(G),\sigma)$. This follows since p is odd and the automorphism induced by σ on the elementary p-abelian quotient $G/\Phi(G)$ is inversion.

We now consider certain special finite quotients of a finitely generated pro-p group, namely the maximal quotients of a fixed p-class. To define this, let $P_0(G) = G$ and, for $n \geq 0$, let $P_{n+1}(G)$ denote the (closed) subgroup generated by $[G, P_n(G)]$ and $P_n(G)^p$. The groups $P_0(G) \geq P_1(G) \geq P_2(G) \geq \ldots$ form a descending chain of characteristic subgroups of G called the lower p-central series. Note that $P_1(G)$ is the Frattini subgroup $\Phi(G)$. The p-class c of a finite p-group G is defined to be the smallest $n \geq 0$ for which $P_n(G) = \{1\}$. If N is a normal subgroup of G, and G/N has p-class n, then $P_n(G) \leq N$. Thus, if G has p-class c, then for $n = 0, \ldots, c$, the maximal p-class n quotient of G is $G/P_n(G)$.

Suppose G has p-class c. A pro-p group H satisfying $H/P_c(H) \cong G$ is called a *descendant* of G. If H has p-class $c+1$, then H is called a *child* or *immediate descendant* of G. O'Brien [27] produced an algorithm that computes all children (and so ultimately all descendants of any finite p-class) of a given p-group. It will be important for us to give much consideration to the maximal p-class n quotients of Schur σ-groups so we make the following definition.

Definition 2.3 Let G be a finite p-group of p-class c. We say that G is a *Schur σ-ancestor* if it is the maximal p-class c quotient of a Schur σ-group. Note that this terminology has the slightly unorthodox meaning that every finite Schur σ-group is regarded as a Schur σ-ancestor of itself.

For the O'Brien p-group generation algorithm, two invariants of a p-group G play important roles, namely its p-multiplicator rank and its nuclear rank. We now recall their definitions and some of their important properties. Suppose G is a p-group with $d(G) = g$ and presentation $1 \to R \to F \to G \to 1$; recall that F is the free pro-p group on g generators $x_1, \ldots, x_g$. The isomorphism class of the objects we are about to define does not depend on the choice of presentation. The *p-covering group* G^* of G is F/R^* where R^* is the topological closure of $R^p[F,R]$. The *p-multiplicator* of G is defined to be the subgroup R/R^* of G^*, and the *nucleus* of G is $P_c(G^*)$ where c is the p-class of G. The nucleus is a subgroup of the p-multiplicator. We call the dimension of R/R^* the *p-multiplicator rank*; the dimension of the subgroup $P_c(G^*)$ is called the *nuclear rank*. If a group has nuclear rank 0, then it has no children and is called *terminal*.

Remark 2.4 In [27], the quantities introduced above are shown to be well-defined and independent of the choice of abstract presentation for a finite p-group G rather than for pro-p presentations. This switch does not cause any problems since if E

is an abstract free group on the same finite generating set as F then one can show that $E/P_c(E) \cong F/P_c(F)$ for all $c \geq 1$. If G has p-class c then this isomorphism can be used to show that there is a one-to-one correspondence between the normal subgroups M of E with $E/M \cong G$ and the (open) normal subgroups N of F with $F/N \cong G$. Furthermore, one can see that for each such corresponding pair of subgroups, we have isomorphisms $E/M^* \cong F/N^*$ and $M/M^* \cong N/N^*$ induced by $E/P_{c+1}(E) \cong F/P_{c+1}(F)$. It follows that the definitions of the nucleus and p-multiplicator rank are independent of whether one uses an abstract or pro-p presentation for G.

A Schur σ-ancestor group of p-class c which is terminal has no proper descendants but must be $H/P_c(H)$ for some Schur σ-group H; hence it must be isomorphic to H and so is a Schur σ-group. Thus, terminal Schur σ-ancestor groups are always Schur σ-groups. In the other direction, in the appendix, Blackhurst proves that a non-cyclic p-group with trivial Schur multiplier must be terminal; this is a result to which several authors have referred, but there appears to be no proof in the literature. Since Schur σ-groups satisfy $r(G) = d(G)$, they have trivial Schur multiplier, hence finite non-cyclic Schur σ-groups are terminal. In summary, terminal Schur σ-ancestor groups are precisely finite non-cyclic Schur σ-groups.

2.2 Measures of p-groups

For each positive integer c, let $F_c = F/P_c(F)$ with GI-automorphism σ induced by the GI-automorphism σ on F defined previously. As an analogue of $X \subset \Phi(F)$, we introduce $X_c \subset \Phi(F_c)$ by defining

$$X_c = X(\Phi(F_c), \sigma) = \{s \in \Phi(F_c) \mid \sigma(s) = s^{-1}\}.$$

Let G be a finite p-group of p-class c with generator and relation ranks both equal to g. One can see that G is a quotient of $F_{c'}$ for all $c' \geq c$. We will say that the tuple of elements $v = (t_1, \ldots, t_g) \in \Phi(F_{c'})^g$ *presents* G if $F_{c'}/\langle v \rangle \cong G$ where $\langle v \rangle$ denotes the closed normal subgroup of $F_{c'}$ generated by $t_1, \ldots, t_g$. We let $S_{c'} = S_{c'}(G)$ denote the set of all such tuples in $\Phi(F_{c'})^g$.

When G is a Schur σ-ancestor group we also wish to consider those relations satisfying the additional restriction that they belong to $X_{c'}$ for some fixed $c' \geq c$. To explain this further, we need two lemmas.

Lemma 2.5 *For all $d \geq 1$, we have $X_d = X'_d$ where*

$$X'_d = \{t^{-1}\sigma(t) \mid t \in \Phi(F_d)\}.$$

Hence, for all $g \geq 1$, the map $\phi_d : \Phi(F_d)^g \to X_d^g$ defined by

$$(t_1, \ldots, t_g) \mapsto (t_1^{-1}\sigma(t_1), \ldots, t_g^{-1}\sigma(t_g))$$

is surjective. Furthermore, for each $w \in X_d^g$, the fiber $\phi_d^{-1}(w)$ is a coset of Y_d^g in $\Phi(F_d)^g$ where $Y_d = Y(F_d, \sigma)$.

Proof It is easy to verify that $X'_d \subseteq X_d$. Now consider the map $X_d \to X'_d$ defined by $t \mapsto t^{-1}\sigma(t) = t^{-2}$. This map is injective since p is odd and F_d is a finite p-group. It follows that $|X_d| \leq |X'_d|$ and hence we must have equality $X_d = X'_d$.

The statement that the fibers of ϕ_d are cosets is straightforward and makes use of the fact that $Y_d \subseteq \Phi(F_d)$.

Remark 2.6 Using the fact $X_d = X'_d$ for all d, one can now show that $X = X'$ where $X' = \{t^{-1}\sigma(t) \mid t \in \Phi(F)\}$. Since both sets are closed in F, it suffices to prove that $\psi_d(X) = \psi_d(X')$ for all d where $\psi_d : F \to F_d$ is the natural projection. It is easy to see that $X' \subseteq X$ and hence $\psi_d(X') \subseteq \psi_d(X)$. It follows that

$$X'_d = \psi_d(X') \subseteq \psi_d(X) \subseteq X_d = X'_d$$

and hence the two the middle containments are also equalities.

Lemma 2.7 *If H is a Schur σ-ancestor group of p-class c then it can be presented with a tuple of relations in $X_c^g \subseteq \Phi(F_c)^g$. Conversely, any group H presented by a tuple of relations in X_c^g is a Schur σ-ancestor group of p-class at most c.*

Proof If H is a Schur σ-ancestor group of p-class c then $H = G_c$ for some Schur σ-group G. If one selects a tuple of relations for G in X^g then it is easy to check that their images in X_c^g, under the natural projection from F to F_c, present H as a quotient of F_c.

Now suppose that $H = F_c/\langle v \rangle$ where $v \in X_c^g$. Then by construction H has p-class at most c. We wish to show that $H = G_c$ for some Schur σ-group G. Lift v to a tuple $w \in \Phi(F)^g$. We can always choose the lift in such a way that the abelianization of the quotient $G = F/\langle w \rangle$ is finite. One can see this by taking the images of the components of w in $F^{ab} \cong \mathbb{Z}_p^g$ and forming them into the rows of a $g \times g$ matrix over $\mathbb{Z}_p$. The group G has finite abelianization if and only if this matrix has full rank g. Different lifts give rise to translations of the rows of this matrix by arbitrary row vectors in $p^c\mathbb{Z}_p^g$ since the image of $P_c(F)$ in $F^{ab} \cong \mathbb{Z}_p^g$ is $P_c(F^{ab}) \cong p^c\mathbb{Z}_p^g$. The statement then follows since the matrices of full rank are dense.

At this stage, we've constructed a g-generated group G so that it has finite abelianization and using only g relations. It follows that the relation rank of G is exactly g. However, it is not clear that the group G must inherit a GI-automorphism from F. This can also be arranged using Lemma 2.5 and Remark 2.6. Replace the initial tuple v in the argument above with an inverse image $v' \in \phi_c^{-1}(v) \subseteq \Phi(F_c)^g$. One can lift v' to a tuple w' so that $F/\langle w' \rangle$ has finite abelianization and then take $w = \phi(w')$ where $\phi : \Phi(F)^g \to X^g$ is the map defined by $t \mapsto t^{-1}\sigma(t)$ in each component. The group $G = F/\langle w \rangle$ has the same finite abelianization and relation rank g, but must now also possess a GI-automorphism and hence be a Schur σ-group since $\langle w \rangle$ is invariant under σ.

To finish, one observes that $G_c \cong H$ thus showing that H is a Schur σ-ancestor group as desired.

Definition 2.8 Let G be a Schur σ-ancestor of p-class c and generator rank g. For $c' \geq c$, let $T_{c'} = T_{c'}(G)$ denote the set of all tuples in $X_{c'}^g$ which present G. We then define the *c'-measure of G* by

$$\mathrm{Meas}_{c'}(G) = \frac{|T_{c'}|}{|X_{c'}|^g}.$$

For a finite p-group G which is not a Schur σ-ancestor group, we set $\mathrm{Meas}_{c'}(G) = 0$. We view the c'-measure of a Schur σ-ancestor group G as the probability with which that group arises as a quotient of $F_{c'}$ when one selects a tuple of relations at random from $X_{c'}^g$. Shortly, we will examine the sequence $(\mathrm{Meas}_{c'}(G))_{c' \geq c}$.

Example 2.9 As an example, let $p = 3$ and consider the case where $g = 2$ and $c' = c = 2$. O'Brien's algorithm yields seven finite 2-generated 3-groups of 3-class 2, of which three are Schur σ-ancestor groups. In this case, $F_2 = F/P_2(F)$ has order 3^5 and we calculate that the set X_2 is an elementary abelian subgroup of order 9. Of these three Schur σ-ancestor groups, the one of order 27 - call it G_1 - arises when the ordered 2-tuple taken from X_2 generates X_2. This happens for 48 of the 81 ordered 2-tuples. Thus $\mathrm{Meas}_2(G_1) = 16/27$. The second group, of order 81 - call it G_2 - arises when the ordered 2-tuple generates one of the four subgroups of X_2 of order 3. Each of these four subgroups is generated by 8 of the 81 ordered 2-tuples in $X_2 \times X_2$; hence, $\mathrm{Meas}_2(G_2) = 32/81$. The third group, of order 243 - call it G_3 - is F_2 itself and arises when both entries in the 2-tuple are trivial. Therefore, $\mathrm{Meas}_2(G_3) = 1/81$. Note that $\mathrm{Meas}_{c'}$ of each of these groups is 0 if $c' > 2$. An explanation for this will be given shortly.

Remark 2.10 In the above example, $X_{c'}$ happened to be a subgroup; in general, X and $X_{c'}$ are not subgroups.

We have $\mathrm{Meas}_1(G) = 1$ where G is the elementary abelian p-group of generator rank g. More generally, for $c > 1$, one can use Lemma 2.7 to see that $\mathrm{Meas}_c(G)$ defines a discrete probability measure on the set of isomorphism classes of maximal p-class c quotients of all Schur σ-groups of generator rank g. This set is finite and consists of the Schur σ-ancestor groups of p-class exactly c, together with all Schur σ-groups of p-class less than c. The next theorem shows how these different probability measures are related.

Theorem 2.11 *Let G be a Schur σ-ancestor group of p-class c.*

(i) We have

$$\mathrm{Meas}_c(G) = \mathrm{Meas}_{c+1}(G) + \sum_Q \mathrm{Meas}_{c+1}(Q)$$

where the summation is over all immediate descendants Q of G which are Schur σ-ancestor groups.

(ii) $\mathrm{Meas}_{c'}(G) = \mathrm{Meas}_{c+1}(G)$ *for all* $c' \geq c+1$.

Proof It follows from Lemma 2.5 that fibers over individual elements for the maps $\phi_c : \Phi(F_c)^g \to X_c^g$ and $\phi_{c+1} : \Phi(F_{c+1})^g \to X_{c+1}^g$ are uniform in size. The same statement holds for the natural projection $\psi : \Phi(F_{c+1})^g \to \Phi(F_c)^g$. We have an induced map $\Psi : X_{c+1}^g \to X_c^g$ obtained by restricting ψ to the subset $X_{c+1}^g \subseteq \Phi(F_{c+1})^g$. It is also surjective and must have fibers that are uniform in size since $\Psi \circ \phi_{c+1} = \phi_c \circ \psi$. Thus, we have

$$\mathrm{Meas}_c(G) = \frac{|T_c|}{|X_c|^g} = \frac{|\Psi^{-1}(T_c)|}{|\Psi^{-1}(X_c^g)|} = \frac{|\Psi^{-1}(T_c)|}{|X_{c+1}|^g}.$$

The statement in part (i) will follow once we show $\Psi^{-1}(T_c) = \Psi^{-1}(T_c(G)) = T_{c+1}(G) \cup \bigcup_Q (T_{c+1}(Q))$ where Q runs through the immediate descendants of G.

Note that the union is disjoint by definition of T_{c+1}, and $T_{c+1}(Q) = \emptyset$ if Q is not a Schur σ-ancestor group by Lemma 2.7.

We now check containment in both directions. If $v \in T_{c+1}(G) \cup \bigcup_Q (T_{c+1}(Q))$ and $\langle v \rangle$ is the normal subgroup of F_{c+1} generated by v, then $H = F_{c+1}/\langle v \rangle$ is isomorphic to G or an immediate descendant Q. In either case, $H/P_c(H) \cong G$. It follows that $\Psi(v) \in T_c(G)$ since the normal subgroup $\langle \Psi(v) \rangle$ in F_c is equal to the image of the normal subgroup $\langle v \rangle$ under the natural epimorphism $F_{c+1} \to F_c$, and thus

$$F_c/\langle \Psi(v) \rangle \cong F_{c+1}/\langle v \rangle P_c(F_{c+1}) \cong H/P_c(H) \cong G.$$

For the other direction, suppose that one has a tuple $v \in X_{c+1}^g$ with $\Psi(v) \in T_c(G)$. Then $H = F_{c+1}/\langle v \rangle$ is a Schur σ-ancestor group by Lemma 2.7 with p-class at most $c+1$. We also have $H/P_c(H) \cong G$ which follows again since $\langle \Psi(v) \rangle$ is equal to the image of $\langle v \rangle$ under the natural epimorphism $F_{c+1} \to F_c$. We deduce that H is either G or an immediate descendant and so by definition $v \in T_{c+1}(G)$ or $v \in T_{c+1}(Q)$ for some immediate descendant Q.

The proof of part (ii) reduces to verifying that $\Psi^{-1}(T_{c+1}(G)) = T_{c'}(G)$ where $\Psi : X_{c'}^g \to X_{c+1}^g$ is now the restriction of the natural epimorphism $\psi : \Phi(F_{c'})^g \to \Phi(F_{c+1})^g$. Verifying the containment $T_{c'}(G) \subseteq \Psi^{-1}(T_{c+1}(G))$ is straightforward. For the reverse containment, we must make use of the assumption that G has p-class c. Suppose that $v \in X_{c'}^g$ and $\Psi(v) \in T_{c+1}(G)$. We wish to show that $v \in T_{c'}(G)$. Let $w \in X^g \subseteq F^g$ be a lift of v under the natural epimorphism $F^g \to F_{c'}^g$ such that $\hat{G} = F/\langle w \rangle$ is a Schur σ-group (see the proof of Lemma 2.7). Let $\hat{G}_d$ denote the quotient $\hat{G}/P_d(\hat{G})$. Then we have $\hat{G}_{c+1} \cong G \cong \hat{G}_c$ since $\Psi(v) \in T_{c+1}(G)$ and G has p-class c. Equivalently, $P_{c+1}(\hat{G}) = P_c(\hat{G})$. An inductive argument now shows that $P_d(\hat{G}) \cong P_c(\hat{G})$ and hence $\hat{G}_d \cong G$ for all $d \geq c$. In particular, $F_{c'}/\langle v \rangle \cong \hat{G}_{c'} \cong G$ which shows that $v \in T_{c'}(G)$ as desired.

Definition 2.12 Let G be a Schur σ-ancestor group of p-class c. We define the *measure of* G (denoted $\mathrm{Meas}(G)$) to be the common value of $\mathrm{Meas}_{c'}(G)$ for $c' \geq c+1$. For a finite p-group G which is not a Schur σ-ancestor group, we set $\mathrm{Meas}(G) = 0$.

Theorem 2.13 *Let G be a Schur σ-ancestor group of p-class c.*

1. *If G is a non-cyclic Schur σ-group, then* $\mathrm{Meas}(G) = \mathrm{Meas}_c(G) > 0$.
2. *If G is not a Schur σ-group, then* $\mathrm{Meas}(G) = \mathrm{Meas}_{c+1}(G) = 0$ *and* $\mathrm{Meas}_c(G)$ *is the sum of the $c+1$-measures of its immediate descendants.*

Proof If G is a non-cyclic Schur σ-group of p-class c then, as discussed in Section 2.1, it has no descendants and so by part (i) of Theorem 2.11 we see that $\mathrm{Meas}_c(G) = \mathrm{Meas}_{c+1}(G)$. It follows that $\mathrm{Meas}(G) = \mathrm{Meas}_c(G)$ and $\mathrm{Meas}_c(G) > 0$ since $T_c(G) \neq \emptyset$.

On the other hand, if $\mathrm{Meas}(G) = \mathrm{Meas}_{c+1}(G) > 0$ then $T_{c+1}(G) \neq \emptyset$. Let $v \in T_{c+1}(G)$ and consider a lift $w \in X^g$ such that $\hat{G} = F/\langle w \rangle$ is a Schur σ-group. If G has p-class c then the arguments in the proof of part (ii) of Theorem 2.11 imply that $G \cong \hat{G}_c \cong \varprojlim \hat{G}_{c'} \cong \hat{G}$. Hence G itself is a Schur σ-group of p-class c. Thus if G is a Schur σ-ancestor of p-class c which is not a Schur σ-group then one must have $\mathrm{Meas}_{c+1}(G) = 0$ and so $\mathrm{Meas}_c(G)$ is the sum of the $c+1$-measures of its immediate descendants by part (i) of Theorem 2.11,

2.3 Measures of abelian p-groups

We are now going to define analogous measures on the class of finite abelian p-groups and relate these to the measures introduced above. This can be used to justify the assertion that our conjectures in the non-abelian setting generalize the Cohen-Lenstra heuristics for p-class groups. See Remark 3.3.

In what follows, the role of F and F_c will be played by the abelianizations F^{ab} and F_c^{ab}. Note that $(F_c)^{ab} \cong (F^{ab})_c$. Every abelian pro-$p$ group G comes equipped with a unique σ-automorphism, namely the inversion mapping $x \mapsto x^{-1}$. We define sets X^{ab} and X_c^{ab} in an analogous way to X and X_c but things are now simpler and it is easy to verify that $X^{ab} = \Phi(F^{ab})$ and $X_c^{ab} = \Phi(F_c^{ab})$.

Let G be a finite abelian p-group of p-class c with generator rank g and let $c' \geq c$. We will say that the tuple of elements $v = (t_1, \ldots, t_g) \in \Phi(F_{c'}^{ab})^g$ *presents* G if $F_{c'}^{ab}/\langle v \rangle \cong G$ where $\langle v \rangle$ denotes the (normal) subgroup of $F_{c'}^{ab}$ generated by $t_1, \ldots, t_g$. We let $S_{c'}^{ab} = S_{c'}^{ab}(G)$ denote the set of all such tuples in $\Phi(F_{c'}^{ab})^g$. In the non-abelian setting, we introduced a second set of tuples $T_{c'} \subseteq S_{c'}$. We can do the same in the abelian setting, but the situation now is simpler and we have $T_{c'}^{ab} = S_{c'}^{ab}$ since $X_{c'}^{ab} = \Phi(F_{c'}^{ab})$.

Definition 2.14 Let G be an abelian p-group of p-class c and generator rank g. For $c' \geq c$, we define the *abelian c'-measure of G* by

$$\mathrm{Meas}_{c'}^{ab}(G) = \frac{|T_{c'}^{ab}|}{|X_{c'}^{ab}|^g} \left(= \frac{|S_{c'}^{ab}|}{|\Phi(F_{c'}^{ab})|^g} \right).$$

We view the abelian c'-measure of a finite p-group G as the probability with which that group arises as a quotient of $F_{c'}^{ab}$ when one selects a tuple of relations at random from $(X_{c'}^{ab})^g = \Phi(F_{c'}^{ab})^g$.

Theorem 2.15 *Let G be an abelian p-group of p-class c.*

(i) We have

$$\mathrm{Meas}_c^{ab}(G) = \mathrm{Meas}_{c+1}^{ab}(G) + \sum_Q \mathrm{Meas}_{c+1}^{ab}(Q)$$

where the summation is over all immediate abelian descendants Q of G.

(ii) $\mathrm{Meas}_{c'}^{ab}(G) = \mathrm{Meas}_{c+1}^{ab}(G)$ *for all* $c' \geq c+1$.

Proof The proof is carried out in exactly the same fashion as the proof of Theorem 2.11. We omit the details.

Definition 2.16 Let G be an abelian p-group of p-class c. We define the *abelian measure of G* (denoted $\mathrm{Meas}^{ab}(G)$) to be the common value of $\mathrm{Meas}_{c'}^{ab}(G)$ for $c' \geq c+1$.

Remark 2.17 It follows from part (i) of Theorem 2.15 that if G is an abelian p-group of p-class c then

$$\mathrm{Meas}^{ab}(G) = \mathrm{Meas}_c^{ab}(G) - \sum_Q \mathrm{Meas}_{c+1}^{ab}(Q)$$

where the summation is over all abelian groups Q of p-class $c+1$ with $Q/Q^{p^c} \cong G$; here $Q^{p^c} = P_c(Q)$ is the subgroup of Q generated by all p^c-th powers.

The following theorem and its corollary provide the link between Meas_c^{ab} and Meas_c.

Theorem 2.18 *Let G be an abelian p-group of p-class c. For all $c' \geq c$ we have*

$$\mathrm{Meas}_{c'}^{ab}(G) = \sum_{Q} \mathrm{Meas}_{c'}(Q)$$

where the summation is over all Schur σ-ancestor groups Q with p-class at most c' and $Q^{ab} \cong G$.

Proof It follows from Lemma 2.5 that the fibers of the map $\phi_{c'} : \Phi(F_{c'})^g \to X_{c'}^g$ are uniform in size. The same statement holds for the analogous map on the abelian side, namely $\phi_{c'}^{ab} : \Phi(F_{c'}^{ab})^g \to (X_{c'}^{ab})^g = \Phi(F_{c'}^{ab})^g$ given by $(t_1, \ldots, t_g) \mapsto (t_1^{-1}\sigma(t_1), \ldots, t_g^{-1}\sigma(t_g)) = (t_1^{-2}, \ldots, t_g^{-2})$. Indeed, the latter map is a bijection since p is odd. We also have a projection map $\psi : \Phi(F_{c'})^g \to \Phi(F_{c'}^{ab})^g$ and its restriction $\Psi : X_{c'}^g \to (X_{c'}^{ab})^g = \Phi(F_{c'}^{ab})^g$. Since the projection ψ has uniform fibers and $\Psi \circ \phi_{c'} = \phi_{c'}^{ab} \circ \psi$, we see that Ψ is also onto and has uniform fibers. Thus

$$\mathrm{Meas}_{c'}^{ab}(G) = \frac{|T_{c'}^{ab}|}{|X_{c'}^{ab}|^g} = \frac{|\Psi^{-1}(T_{c'}^{ab})|}{|\Psi^{-1}((X_{c'}^{ab})^g)|} = \frac{|\Psi^{-1}(T_{c'}^{ab})|}{|X_{c'}|^g}.$$

If $v \in \Psi^{-1}(T_{c'}^{ab}) = \Psi^{-1}(T_{c'}^{ab}(G))$ then by definition $F_{c'}^{ab}/\langle \Psi(v) \rangle \cong G$ and so $(F_{c'}/\langle v \rangle)^{ab} \cong F_{c'}^{ab}/\langle \Psi(v) \rangle \cong G$. This means $v \in T_{c'}(Q)$ where $Q = F_{c'}/\langle v \rangle$ is a Schur σ-ancestor group by Lemma 2.7 and we have $Q^{ab} \cong G$. Conversely, the same isomorphisms show that if $v \in T_{c'}(Q)$ for a Schur σ-ancestor group Q with $Q^{ab} \cong G$, then $v \in \Psi^{-1}(T_{c'}^{ab}(G))$. Thus we have $\Psi^{-1}(T_{c'}^{ab}(G)) = \bigcup_Q T_{c'}(Q)$ where the union is taken over all the Schur σ-ancestor groups Q with p-class at most c' and $Q^{ab} \cong G$. The union is disjoint so the statement about the measures now follows.

If $Q^{ab} \cong G$ and G has p-class c then Q must have p-class at least c. We thus have the following corollary.

Corollary 2.19 *Let G be an abelian p-group of p-class c. Then*

$$\mathrm{Meas}_{c}^{ab}(G) = \sum_{Q} \mathrm{Meas}_{c}(Q)$$

where the summation is over all Schur σ-ancestor groups Q with p-class exactly c that satisfy $Q^{ab} \cong G$.

2.4 Formulas for Meas_c^{ab} and Meas_c

We will now derive formulas for the various measures introduced so far starting with the abelian case. The derivation in this case lays the groundwork for the proof of Theorem 2.25 which is more complicated but has a similar structure and begins with the same counting argument.

Theorem 2.20 *Let G be an abelian p-group of p-class c and generator rank g. We have*

$$\mathrm{Meas}_c^{ab}(G) = \frac{1}{|\mathrm{Aut}(G)|}\, p^{g^2} \prod_{k=1}^{g} (1-p^{-k}) \prod_{k=1+g-u}^{g} (1-p^{-k})$$

where u counts the number of cyclic groups of order strictly less than p^c in the direct product decomposition of G.
For $c' > c$, we have

$$\mathrm{Meas}_{c'}^{ab}(G) = \mathrm{Meas}^{ab}(G) = \frac{1}{|\mathrm{Aut}(G)|}\, p^{g^2} \prod_{k=1}^{g} (1-p^{-k})^2.$$

Proof To compute $\mathrm{Meas}_c^{ab}(G)$ we need to count tuples of relations in $\Phi(F_c^{ab})^g$ which present G. We will do this in two stages by following the same strategy as in [5]. First, we will count the number of normal subgroups $\overline{R}$ in F_c^{ab} with $F_c^{ab}/\overline{R} \cong G$ by counting certain collections of epimorphisms. Then we will count the number of generating tuples that generate each such subgroup as a normal subgroup although the normality condition imposes no restriction here since F_c^{ab} is abelian.

Let $\mathrm{Epi}(F, G)$ be the set of epimorphisms from F to G where F is the free pro-p group on g generators. Such epimorphisms are in one-to-one correspondence with ordered g-tuples of elements in G that generate G. By Burnside's basis theorem, a tuple of elements generates G if and only if it generates $G/\Phi(G)$. It follows that

$$|\mathrm{Epi}(F,G)| = |\Phi(G)|^g\,(p^g - p^{g-1})(p^g - p^{g-2})\dots(p^g-1) = |\Phi(G)|^g \prod_{k=1}^{g} (p^g - p^{g-k})$$

since $G/\Phi(G)$ is an $\mathbb{F}_p$-space of dimension g.

Two epimorphisms have the same kernel if and only if they differ by an automorphism of G, so dividing by $|\mathrm{Aut}(G)|$ gives the number of (closed) normal subgroups R of F with quotient isomorphic to G. Since G is abelian and has p-class c we have $P_c(F)[F,F] \subseteq R$ for each such subgroup R and there is a one-to-one correspondence between these subgroups of F and the subgroups $\overline{R}$ of F_c^{ab} such that $F_c^{ab}/\overline{R} \cong G$. Thus the number of such subgroups $\overline{R}$ is

$$\frac{|\mathrm{Epi}(F,G)|}{|\mathrm{Aut}(G)|} = \frac{|\Phi(G)|^g}{|\mathrm{Aut}(G)|} \prod_{k=1}^{g} (p^g - p^{g-k}).$$

Now we need to count how many g-tuples of elements generate each $\overline{R}$ as a (normal) subgroup of F_c^{ab}. A g-tuple of elements generates $\overline{R}$ as a subgroup of F_c^{ab} if and only if their images generate the $\mathbb{F}_p$-space $V = \overline{R}/\Phi(\overline{R})$. Since $F_c^{ab} \cong F^{ab}/(F^{ab})^{p^c}$ is a product of g copies of $\mathbb{Z}/p^c\mathbb{Z}$, the dimension of V is equal to the number of cyclic factors in the decomposition of the abelian group G which are strictly smaller than $\mathbb{Z}/p^c\mathbb{Z}$. This is the quantity u in the statement of the theorem. There are $\prod_{k=1}^{u}(p^g - p^{u-k})$ g-tuples of elements in V which span this space and hence

$$|\Phi(\overline{R})|^g \prod_{k=1}^{u} (p^g - p^{u-k})$$

g-tuples that generate each subgroup $\overline{R}$. Note that $|\Phi(\overline{R})| = |F_c^{ab}|/[F_c^{ab} : \Phi(\overline{R})] = |F_c^{ab}|/(|G|p^u)$ so this quantity is independent of the particular subgroup $\overline{R}$ being considered.

Combining the statements above, we have

$$\begin{aligned}
\mathrm{Meas}_c^{ab}(G) = \frac{|S_c^{ab}(G)|}{|\Phi(F_c^{ab})|^g} &= \frac{1}{|\Phi(F_c^{ab})|^g}\frac{|\Phi(G)|^g}{|\mathrm{Aut}(G)|}\prod_{k=1}^{g}(p^g - p^{g-k})|\Phi(\overline{R})|^g\prod_{k=1}^{u}(p^g - p^{u-k}) \\
&= \frac{1}{|\Phi(F_c^{ab})|^g}\frac{(|\Phi(F_c^{ab})|/|\overline{R}|)^g}{|\mathrm{Aut}(G)|}\prod_{k=1}^{g}(p^g - p^{g-k})\frac{|\overline{R}|^g}{p^{gu}}\prod_{k=1}^{u}(p^g - p^{u-k}) \\
&= \frac{1}{|\mathrm{Aut}(G)|}\frac{1}{p^{gu}}\prod_{k=1}^{g}(p^g - p^{g-k})\prod_{k=1}^{u}(p^g - p^{u-k}) \\
&= \frac{1}{|\mathrm{Aut}(G)|}p^{g^2}\prod_{k=1}^{g}(1-p^{-k})\prod_{k=1+g-u}^{g}(1-p^{-k})
\end{aligned}$$

The second statement about $\mathrm{Meas}_{c'}^{ab}(G)$ for $c' > c$ is verified in exactly the same way. The only difference occurs in the second step. One sees that the space $V = \overline{R}/\Phi(\overline{R})$ has dimension g since G has p-class c which means that all g of its cyclic components are strictly smaller that $\mathbb{Z}/p^{c'}\mathbb{Z}$. Thus the formula one obtains is the one above with $u = g$.

Remark 2.21 If we define $\eta_j(p) = \prod_{k=1}^{j}(1-p^{-k})$ as in [13], then the formulas in Theorem 2.20 can be written

$$\begin{aligned}
\mathrm{Meas}_c^{ab}(G) &= \frac{1}{|\mathrm{Aut}(G)|}p^{g^2}\left(\frac{\eta_g(p)^2}{\eta_{g-u}(p)}\right) \\
\mathrm{Meas}^{ab}(G) &= \frac{1}{|\mathrm{Aut}(G)|}p^{g^2}\eta_g(p)^2.
\end{aligned}$$

To derive similar formulas for the measures in the non-abelian context, we need an additional technical assumption on the groups involved.[2] Recall that F is the free pro-p group of generator rank g. Let G be a Schur σ-ancestor group of p-class c with generator rank g. Given $w \in T_c(G)$, the normal subgroup $\langle w \rangle$ is the kernel of an epimorphism from F_c to G and satisfies $\sigma(\langle w \rangle) = \langle w \rangle$. In the lemma and theorems which follow, we will need to make the much stronger assumption that the kernel of *every* epimorphism from F_c to G is invariant under σ. Or, equivalently, that the kernel of every epimorphism from F to G is invariant under σ.

Definition 2.22 If G is a finite p-group with the same generator rank as the free group F and $\sigma(\ker\psi) = \ker\psi$ for every epimorphism $\psi : F \to G$ then we will say that G satisfies the *kernel invariance property* (KIP).

Some additional remarks about this property and its range of applicability will be made later in Section 2.5.

Lemma 2.23 *Let G be a Schur σ-ancestor group of p-class c satisfying KIP. Let $c' \geq c$ and define $y(G)$, $S_{c'} = S_{c'}(G)$, $T_{c'} = T_{c'}(G)$, $\phi_{c'}$ and $Y_{c'} = Y(F_{c'}, \sigma)$ as discussed in Sections 2.1 and 2.2 prior to Lemma 2.7. The following statements hold.*

[2] See the earlier footnote to Remark 1.4.

(i) *If* $v \in S_{c'}$ *then* $\langle \phi_{c'}(v) \rangle = \langle v \rangle$ *and* $\phi_{c'}(v) \in T_{c'}$.
(ii) *If* $w \in T_{c'}$ *then there exists* $v \in S_{c'}$ *such that* $\phi_{c'}(v) = w$.
(iii) *Let* $v \in S_{c'}$ *and* $\overline{R} = \langle v \rangle$. *Then* $u \in S_{c'}$ *and* $\phi_{c'}(u) = \phi_{c'}(v)$ *if and only if* $uv^{-1} \in (Y_{c'} \cap \overline{R})^g$.
(iv) $[Y_{c'} : Y_{c'} \cap \overline{R}] = y(G)$ *where* $\overline{R}$ *is the kernel of any epimorphism from* $F_{c'}$ *to* G.

Proof We make two preliminary observations. First, since G has p-class c and satisfies KIP, the kernel of every epimorphism from $F_{c'}$ to G where $c' \geq c$ must be invariant under σ. This follows since each such epimorphism is induced by an epimorphism from F to G for which the kernel is invariant by the KIP assumption. This form of KIP is used below and in some of the later proofs in this section.

Second, if we restrict $\phi_{c'}$ to $X_{c'}^g$ then we obtain a map $\phi_{c'} : X_{c'}^g \to X_{c'}^g$ which is the powering map $t \mapsto t^{-2}$ in each component. Since p is odd, there exists $n \geq 1$ such that $(-2)^n \equiv 1$ modulo the exponent of the group $F_{c'}$. The iterate $\phi_{c'}^n$ is then the identity on $X_{c'}^g$.

For part (i), suppose $v \in S_{c'}$. Then the normal subgroup $\overline{R} = \langle v \rangle$ is the kernel of an epimorphism $F_{c'} \to G$ and so is invariant under σ by the KIP assumption. It follows that $t_i^{-1}\sigma(t_i) \in \langle v \rangle$ for all i and so $\langle v \rangle \supseteq \langle \phi_{c'}(v) \rangle$. Further iterates of the tuple continue to lie in $X_{c'}^g$ and so generate normal subgroups invariant under σ. By induction, we then have

$$\overline{R} = \langle v \rangle \supseteq \langle \phi_{c'}(v) \rangle \ldots \supseteq \langle \phi_{c'}^n(v) \rangle \supseteq \ldots$$

Since $\phi_{c'}(\phi_{c'}^n(v)) = \phi_{c'}^n(\phi_{c'}(v)) = \phi_{c'}(v)$, we see that $u = \phi_{c'}^n(v)$ and v both lie in the same fiber of the map $\phi_{c'}$. This fiber is a (right) coset of $Y_{c'}^g$ by Lemma 2.5. The components of u and v also lie in $\overline{R}$, so $uv^{-1} \in (Y_{c'} \cap \overline{R})^g$. The argument in the first part of the proof of (iii) below now shows that $\langle u \rangle = \langle v \rangle$ and hence that the chain of containments above are all equalities. In particular, we have $\langle \phi_{c'}(v) \rangle = \langle v \rangle = \overline{R}$ and so $\phi_{c'}(v) \in T_{c'}$.

For part (ii), let $w \in T_{c'} \subseteq X_{c'}^g$. We then have $w = \phi_{c'}^n(w) = \phi_{c'}(v)$ where $v = \phi_{c'}^{n-1}(w)$. Note that $v \in T_{c'} \subseteq S_{c'}$ by repeated application of part (i).

For part (iii), let $v = (t_1, \ldots, t_g) \in S_{c'}$ and $\overline{R} = \langle v \rangle$. If $uv^{-1} \in (Y_{c'} \cap \overline{R})^g$ then $u = (y_1 t_1, \ldots, y_g t_g)$ for some $(y_1, \ldots, y_g) \in (Y_{c'} \cap \overline{R})^g$ and it follows immediately that $\phi_{c'}(u) = \phi_{c'}(v)$. The assumption that $t_1, ..., t_g$ generate $\overline{R}$ as a normal subgroup of $F_{c'}$ is equivalent to their images spanning the $\mathbb{F}_p$-space $\overline{R}/\overline{R}^* \cong R/P_{c'}(F)R^*$. We will now show that this also holds for $y_1 t_1, \ldots, y_g t_g$ from which it follows that $\langle u \rangle = \langle v \rangle = \overline{R}$ and hence $u \in S_{c'}$.

The key observation is that the induced action of σ on $\overline{R}/\overline{R}^*$ is entirely by inversion. This follows by first using [20] p.100, Prop. 4, to identify the vector space with $H_2(G, \mathbb{F}_p)$. Next, consider the homology long exact sequence associated to the short exact sequence

$$0 \to \mathbb{Z} \to \mathbb{Z} \to \mathbb{Z}/p \to 0.$$

This is

$$\ldots \to H_2(G, \mathbb{Z}) \to H_2(G, \mathbb{Z}) \to H_2(G, \mathbb{Z}/p) \to H_1(G, \mathbb{Z}) \to H_1(G, \mathbb{Z}) \to \ldots$$

which yields the exact sequence

$$0 \to H_2(G, \mathbb{Z})/pH_2(G, \mathbb{Z}) \to H_2(G, \mathbb{Z}/p) \to H_1(G, \mathbb{Z})[p] \to 0.$$

These maps are σ-equivariant and the 3rd and 4th terms have the same dimension over $\mathbb{F}_p$, implying that $H_2(G,\mathbb{Z}/p)$ is σ-isomorphic to $H_1(G,\mathbb{Z})[p]$. Since G is finite, this in turn is σ-isomorphic to $H_1(G,\mathbb{Z}/p)$, which by [20], p.99, Prop. 3, is σ-isomorphic to $G/\Phi(G)$. By assumption, σ acts on the latter by inversion. Thus if $y_1,\dots,y_g$ lie in $Y_{c'}\cap\overline{R}$ then their images in $\overline{R}/\overline{R}^*$ must be trivial (since they are both fixed and inverted), and so the images of y_it_i for $i=1,\dots,g$ will also span this space. It follows that $\langle u\rangle=\langle v\rangle=\overline{R}$ and so $u\in S_{c'}$.

Conversely, suppose $u\in S_{c'}$ and $\phi_{c'}(u)=\phi_{c'}(v)$. Then we can apply part (i) to see that $\langle u\rangle=\langle\phi_{c'}(u)\rangle=\langle\phi_{c'}(v)\rangle=\langle v\rangle=\overline{R}$. By Lemma 2.5, we can write $u=(y_1t_1,\dots,y_gt_g)$ with $y_i\in Y_{c'}$ for all i. Combining the previous two statements, we deduce that $y_i\in Y_{c'}\cap\overline{R}$ for all i and so $uv^{-1}\in(Y_{c'}\cap\overline{R})^g$.

Finally, part (iv) follows since if $G\cong F_{c'}/\overline{R}$ then $y(G)=y(F_{c'}/\overline{R})=|Y(F_{c'}/\overline{R},\sigma)|$ where σ also denotes the automorphism on the quotient induced by $\sigma:F_{c'}\to F_{c'}$. Now observe that $Y(F_{c'}/\overline{R},\sigma)=Y_{c'}\overline{R}/\overline{R}\cong Y_{c'}/Y_{c'}\cap\overline{R}$. The first equality can be verified by checking containment in both directions. First, since $Y_{c'}=Y(F_{c'},\sigma)$ and $\sigma(\overline{R})=\overline{R}$, it is easy to see that $Y_{c'}\overline{R}/\overline{R}\subseteq Y(F_{c'}/\overline{R},\sigma)$. For the reverse containment, suppose that $g=x\overline{R}\in Y(F_{c'}/\overline{R},\sigma)$, then $\sigma(x)=xr$ for some $r\in\overline{R}$. Since $x=\sigma^2(x)=xr\sigma(r)$ one sees that $\sigma(r)=r^{-1}$. Using the fact that the map $s\mapsto s^2$ is a bijection from $\overline{R}$ to $\overline{R}$, we can select $s\in\overline{R}$ such that $s^2=r$. One can then verify that $x'=xs\in Y_{c'}$ and hence $g=x\overline{R}=x'\overline{R}\in Y_{c'}\overline{R}/\overline{R}$.

Before stating the next theorem we need to define one additional quantity.

Definition 2.24 Let G be a finite p-group. Define $h(G)$ to be p-multiplicator rank of G minus the nuclear rank of G. Equivalently, $h(G)$ is the dimension of the $\mathbb{F}_p$-space $R/P_c(F)R^*$ where $G\cong F/R$ and has p-class c.

We note that for any finite p-group G we have $h(G)\geq 0$. It is a fact that $r(Q)\geq h(G)$ for any descendant Q of G (Prop. 2 of [9]). In particular, if G is g-generated and $h(G)>g$ then G and its descendants cannot be Schur σ-ancestor groups.

Theorem 2.25 *Let G be a Schur σ-ancestor group of p-class c and rank g satisfying KIP. Let $h=h(G)$ and $r=r(G)$. Then*

$$\mathrm{Meas}_c(G)=\frac{y(G)^g}{|\mathrm{Aut}(G)|}\,p^{g^2}\prod_{k=1}^{g}(1-p^{-k})\prod_{k=1+g-h}^{g}(1-p^{-k})$$

and for $c'>c$

$$\mathrm{Meas}_{c'}(G)=\frac{y(G)^g}{|\mathrm{Aut}(G)|}\,p^{g^2}\prod_{k=1}^{g}(1-p^{-k})\prod_{k=1+g-r}^{g}(1-p^{-k})$$

Proof To compute $\mathrm{Meas}_c(G)$, we will first find the proportion of g-tuples of relators in $\Phi(F_c)$ that present G. We will then modify this to obtain $\mathrm{Meas}_c(G)$. A similar argument yields the second formula.

For the first step, we use similar arguments as in Theorem 2.20. If $G=F/R$ has p-class c then $P_c(F)\subseteq R$ and we have a one-to-one correspondence between the normal subgroups R of F such that $F/R\cong G$ and the normal subgroups $\overline{R}$ of F_c such that $F_c/\overline{R}\cong G$. The number of such normal subgroups is

$$\frac{|\mathrm{Epi}(F,G)|}{|\mathrm{Aut}(G)|}=\frac{|\Phi(G)|^g}{|\mathrm{Aut}(G)|}\prod_{k=1}^{g}(p^g-p^{g-k}).$$

A g-tuple of elements generates $\overline{R} = R/P_c(F)$ as a normal subgroup of F_c if and only if its image generates the $\mathbb{F}_p$-space $V = R/P_c(F)R^*$. This has dimension h by definition. If we let $\overline{R}^* = P_c(F)R^*/P_c(F) \subseteq F_c$ then the number of g-tuples that generate $\overline{R}$ is

$$|\overline{R}^*|^g \prod_{k=1}^{h}(p^g - p^{h-k}).$$

A similar calculation to the one in Theorem 2.20 now shows that

$$\begin{aligned}\frac{|S_c(G)|}{|\Phi(F_c)|^g} &= \frac{1}{|\Phi(F_c)|^g}\frac{|\Phi(G)|^g}{|\mathrm{Aut}(G)|}\prod_{k=1}^{g}(p^g - p^{g-k})|\overline{R}^*|^g \prod_{k=1}^{h}(p^g - p^{h-k}) \\ &= \frac{1}{|\mathrm{Aut}(G)|}\, p^{g^2}\prod_{k=1}^{g}(1-p^{-k}) \prod_{k=1+g-h}^{g}(1-p^{-k})\end{aligned}$$

where we've made use of the fact that $|\overline{R}^*| = |\overline{R}|/p^h$ and $|\Phi(G)| = |\Phi(F_c)|/|\overline{R}|$.

We now relate this quantity to $\mathrm{Meas}_c(G)$. Using parts (i), (ii) and (iii) of Lemma 2.23, we have

$$|S_c(G)| = |S_c| = |T_c| \cdot |Y_c \cap \overline{R}|^g.$$

It follows that

$$\begin{aligned}\mathrm{Meas}_c(G) = \frac{|T_c|}{|X_c|^g} &= \frac{|S_c|}{|X_c|^g|Y_c \cap \overline{R}|^g} = \frac{1}{|Y_c \cap \overline{R}|^g}\frac{|\Phi(F_c)|^g}{|X_c|^g}\frac{|S_c|}{|\Phi(F_c)|^g} \\ &= \frac{|Y_c|^g}{|Y_c \cap \overline{R}|^g}\frac{|S_c|}{|\Phi(F_c)|^g} = y(G)^g\frac{|S_c|}{|\Phi(F_c)|^g}\end{aligned}$$

where we have also made use of Lemma 2.5 and part (iv) of Lemma 2.23 in the last two steps. Substituting our earlier expression for $|S_c|/|\Phi(F_c)|^g$, we arrive at the formula for $\mathrm{Meas}_c(G)$ in the statement of the theorem.

This completes the verification of the formula for $\mathrm{Meas}_c(G)$. The verification of the formula for $\mathrm{Meas}_{c'}(G)$ where $c' > c$ is almost identical. The only part that changes is the second step where one now counts the number of g-tuples generating a normal subgroup $\overline{R} = R/P_{c'}(F)$ with $F_{c'}/\overline{R} \cong G$. Since G has p-class c we have $P_c(F) \subseteq R$ and so $P_{c'}(F) \subseteq R^*$ for all $c' > c$. It follows that in this case $V = R/P_{c'}(F)R^* = R/R^*$. This is the p-multiplicator whose dimension as an $\mathbb{F}_p$-space is equal to the relation rank r. Thus the formula for the number of g-tuples can be obtained by taking the formula in the first argument and replacing the quantity h with r.

Corollary 2.26 *Let G be a non-cyclic Schur σ-group of p-class c and rank g satisfying KIP. Then*

$$\mathrm{Meas}(G) = \mathrm{Meas}_c(G) = \frac{y(G)^g}{|\mathrm{Aut}(G)|}\, p^{g^2}\prod_{k=1}^{g}(1-p^{-k})^2.$$

Example 2.27 Let's compute $\mathrm{Meas}_2(G)$ for the Schur σ-ancestor groups of 3-class 2 in Example 2.9 using Theorem 2.25. The fact that these three groups satisfy KIP can be verified computationally or by using Theorem 2.32 since the three groups are all immediate descendants of $F_1 = F/P_1(F)$. We have $p = 3$, $g = 2$

and $c = 2$ so the formula reduces to $\mathrm{Meas}_2(G_i) = 48ky(G_i)^2/|\mathrm{Aut}(G_i)|$, where $k = \frac{1}{3^{2h}}\prod_{k=1}^{h}(3^2 - 3^{h-k})$ is the proportion of ordered pairs of vectors that span an h-dimensional vector space over $\mathbb{F}_3$ and $h = h(G_i)$. For G_1, G_2, G_3 we have $y(G_i) = 3$ for all i, $h(G_i) = 2, 1, 0$ respectively, and $|\mathrm{Aut}(G_i)| = 432, 972, 34992$ respectively. Thus, $\mathrm{Meas}_2(G_1) = 48 \times 16/27 \times 3^2 \times 1/432 = 16/27$; $\mathrm{Meas}_2(G_2) = 48 \times 8/9 \times 3^2 \times 1/972 = 32/81$; $\mathrm{Meas}_2(G_3) = 48 \times 1 \times 3^2 \times 1/34992 = 1/81$. These values agree with those obtained earlier by direct enumeration of the tuples of relations.

Definition 2.28 Suppose G is a finite p-group equipped with a GI-automorphism τ. We denote by $\mathrm{Aut}_\tau(G)$ the set of all automorphisms of G which commute with τ.

Theorem 2.29 *Suppose G is a finite p-group equipped with a GI-automorphism τ and which satisfies KIP. Then $|\mathrm{Aut}_\tau(G)| = |\mathrm{Aut}(G)|/y(G)^g$.*

Proof Let $\Sigma(G)$ be the set of all GI-automorphisms of G. The automorphism group $\mathrm{Aut}(G)$ acts on $\Sigma(G)$ by conjugation and this action is transitive by Hall's theorem and Schur-Zassenhaus. The stabilizer of $\tau \in \Sigma(G)$ is $\mathrm{Aut}_\tau(G)$ so we have $|\mathrm{Aut}(G)| = |\mathrm{Aut}_\tau(G)||\Sigma(G)|$. We will now show that $|\Sigma(G)| = y(G)^g$ which implies the statement of the theorem.

Consider the set $\mathcal{E}(F, G)$ of epimorphisms from F to G. We are going to count the number of elements in $\mathcal{E}(F, G)$ in two different ways. Let $\phi \in \mathcal{E}(F, G)$. The kernel of ϕ is invariant under σ since G satisfies KIP. It follows that σ induces a GI-automorphism on G, which we denote by α, satisfying

$$\alpha(\phi(x)) = \phi(\sigma(x)) \quad (*)$$

for all x in F. We thus have a map $\mathcal{E}(F, G) \to \Sigma(G)$ defined by $\phi \mapsto \alpha$. This map is surjective due to work of Koch and Venkov discussed in Section 2.1. To understand the fibers of this map, we fix α and ask which ϕ satisfy $(*)$. First note that ϕ is determined by $(\phi(x_1), ..., \phi(x_g)) \in G^g$ and that any ordered g-tuple is possible so long as they generate G and satisfy $(*)$. The property $(*)$ says that $\alpha(\phi(x_i)) = \phi(x_i^{-1}) = \phi(x_i)^{-1}$, in other words, that $\phi(x_i) \in X(G, \alpha)$ for all i. Thus every ϕ yields an element of $X(G, \alpha)^g$ generating G and vice versa every element of $X(G, \alpha)^g$ generating G specifies a legitimate ϕ. The size of this set of tuples is independent of α, so we see that the fibers are uniform in size and hence $|\mathcal{E}(F, G)|$ is the product of $|\Sigma(G)|$ and the number of elements of $X(G, \alpha)^g$ generating G.

On the other hand, if we fix $\alpha \in \Sigma(G)$ then it is easily seen that $G = Y(G, \alpha)X(G, \alpha)$ and that $X(G, \alpha) \cap Y(G, \alpha) = \{1\}$. Associate to $\phi \in \mathcal{E}(F, G)$, the tuple $(\phi(x_1), ..., \phi(x_g)) \in G^g$. Write this uniquely as $(a_1b_1, ..., a_gb_g)$ where $a_i \in Y(G, \alpha)$ and $b_i \in X(G, \alpha)$. Since ϕ is surjective if and only if $b_1, ..., b_g$ generate G (as $Y(G, \alpha) \subseteq \Phi(G)$), we see that $|\mathcal{E}(F, G)|$ is $|Y(G, \alpha)|^g$ times the number of elements of $X(G, \alpha)^g$ generating G.

Equating the two expressions for $|\mathcal{E}(F, G)|$, we deduce that $|\Sigma(G)| = |Y(G, \alpha)|^g = y(G)^g$ as desired.

Combining Theorem 2.25 and Theorem 2.29, and using the function $\eta_j(p)$ in Remark 2.21, we obtain the following Corollary, which is the basis for Conjecture 1.3 stated in Section 1.

Corollary 2.30 *Let G be a Schur σ-ancestor group of p-class c and rank g satisfying KIP. Let $h = h(G)$ and $r = r(G)$. Then*

$$\mathrm{Meas}_c(G) = \frac{1}{|\mathrm{Aut}_\sigma(G)|}\, p^{g^2} \left(\frac{\eta_g(p)^2}{\eta_{g-h}(p)} \right)$$

and for $c' > c$

$$\mathrm{Meas}_{c'}(G) = \frac{1}{|\mathrm{Aut}_\sigma(G)|}\, p^{g^2} \left(\frac{\eta_g(p)^2}{\eta_{g-r}(p)} \right).$$

Remark 2.31 We have largely set aside the case of cyclic p-groups in this section because, being abelian, they are already covered by the original Cohen-Lenstra heuristics. Hence, one can compute $\mathrm{Meas}^{ab}(G)$ as a predictor for the value of the frequency $\mathrm{Freq}(G)$. However, one can also view a cyclic p-group G as a Schur σ-group. It easy to see that such a group satisfies KIP and we can therefore compute $\mathrm{Meas}(G)$ via our formula; when we do so, we obtain the same value for $\mathrm{Freq}(G)$ since the GI-automorphism σ is just inversion and so $\mathrm{Aut}_\sigma(G) = \mathrm{Aut}(G)$.

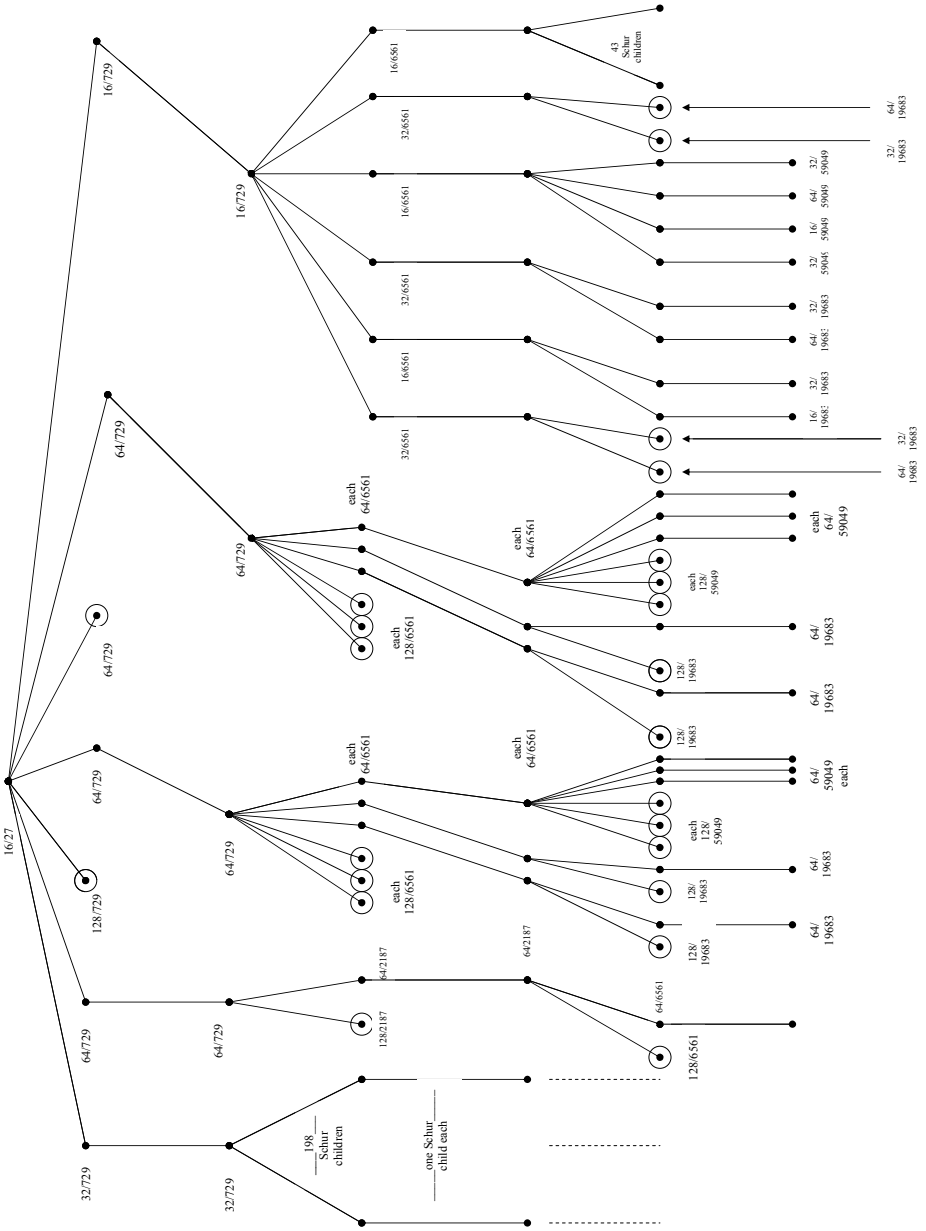
For use later and to illustrate the ideas so far, we display a tree in Figure 2.1 showing the first few levels of Schur σ-ancestor groups G which are descendants of G_1. Each vertex corresponds to a group G and is labeled with the quantity $\mathrm{Meas}_c(G)$ where c is the p-class of G. Certain relationships exist between the labels as explained in Theorem 2.13. Vertices with no descendants (circled in the figure) correspond to Schur σ-groups and in this case the label is also the value of $\mathrm{Meas}(G)$. For vertices that are not terminal, the label is always equal to the sum of the labels of the immediate descendants. These labels were calculated using the formulas from this section and assuming KIP. We confirmed KIP computationally for each group in the figure with p-class at most 6. Unfortunately it would appear to be prohibitively time-consuming to test KIP for all of the groups at the next level.

Identifying whether or not a given p-group is a Schur σ-ancestor group from the definition can be computationally difficult as the order and p-class of the group increases. The p-groups of generator rank g and fixed p-class can be generated using O'Brien's p-group generation algorithm. If one is only interested in those descendants which are Schur σ-ancestor groups, then one can immediately eliminate descendants G which do not possess a GI-automorphism or for which $h(G) > g$. Occasionally, one encounters groups which pass both of these tests and which are not Schur σ-ancestor groups. These groups are hard to distinguish from Schur σ-ancestor groups and so we refer to them as *pseudo-Schur σ-ancestor groups*. For example, the computations in Example 2.9 show that G_1, G_2 and G_3 are the only 2-generated 3-groups which are Schur σ-ancestor groups of 3-class 2. However, the groups $\mathbb{Z}/3\times\mathbb{Z}/9$ and $\mathbb{Z}/9\times\mathbb{Z}/9$ are also 2-generated of 3-class 2 and both possess a GI-automorphism and have the difference between their p-multiplicator rank and nuclear rank equal to 2 making them pseudo-Schur σ-ancestor groups.

In practice, we are often able to exploit Theorem 2.13 to eliminate such groups using an iterative process as follows. Given a known Schur σ-ancestor group G of p-class c, one computes the immediate descendants and eliminates those which do not satisfy the two criteria in the previous paragraph. One then computes the sum of the $c+1$-measures of the groups that remain using the formula in this section. If any pseudo-Schur σ-ancestor groups are present, then the formula will return a

non-zero value even though the actual measure is 0. In this case, the computed sum will exceed $\mathrm{Meas}_c(G)$. On the other hand, if the sum equals and does not exceed the c-measure of G then one can conclude that no such groups are present and the newly generated groups form a complete list of the Schur σ-ancestor groups descended from G.

Fig. 2.1 Descendants of G_1



2.5 Groups satisfying KIP

The KIP condition would seem to be quite restrictive, yet it applies in all the cases where we have needed to compute $\mathrm{Meas}_c(G)$ and we have yet to find a Schur

σ-ancestor group where it does not apply. Finite abelian p-groups certainly satisfy KIP; we also have the following result.

Theorem 2.32 *For all $c \geq 1$, if $G = F_c$ or G is an immediate descendant of F_c possessing a GI-automorphism, then G satisfies KIP.*

Proof If $G = F_c$ then every epimorphism $\alpha : F \to G$ factors through the natural epimorphism $F \to F_c$ since G has p-class c. This gives rise to an epimorphism from F_c to G which must be an isomorphism since $G = F_c$ is finite. It follows that $\ker \alpha = P_c(F)$ which is a characteristic subgroup of F and hence invariant under σ.

If G is an immediate descendant of F_c then every epimorphism $\alpha : F \to G$ factors through the natural epimorphism $F \to F_{c+1}$ and $P_{c+1}(F) \subseteq \ker \alpha \subseteq P_c(F)$. Hence, all the kernels of epimorphisms from F to G will be invariant under σ if and only if all the kernels of the epimorphisms from F_{c+1} to G are invariant. As noted earlier, if G possesses a σ-automorphism then there must exist at least one epimorphism with kernel that is invariant under σ. We now show that this implies all such kernels are invariant.

Observe that if R_1 and R_2 are two such kernels then the isomorphism $F/R_1 \cong F/R_2$ lifts to an automorphism of F which maps R_1 to R_2. It follows that $\mathrm{Aut}(F)$ acts transitively on the set of kernels. If we now consider the images of the kernels in $P_c(F)/P_{c+1}(F) \subseteq F_{c+1}$, then the same statement holds where $\mathrm{Aut}(F)$ acts via the restriction homomorphism $\rho_c : \mathrm{Aut}(F) \to \mathrm{Aut}(P_c(F)/P_{c+1}(F))$. The map ρ_c factors through $\rho_0 : \mathrm{Aut}(F) \to \mathrm{Aut}(F/P_1(F))$. This can be seen by using an inductive argument to verify that $\ker \rho_0 \subseteq \ker \rho_i$ for $i \geq 0$. The induction is straightforward and uses the recursive definition of the central series $\{P_i(F)\}_{i \geq 0}$. More details can be found in the proof of a slightly more general statement appearing in [23, Chapter VIII, Theorem 1.7].

Since $\rho_0(\sigma)$ is the inversion automorphism in $\mathrm{Aut}(F/P_1(F))$, it is clearly central in the image of $\mathrm{Aut}(F)$ under ρ_0. It follows that the image of σ is central in the image of $\mathrm{Aut}(F)$ under ρ_i for all i. Combining this statement for $i = c$ with the transitivity of the action of $\mathrm{Aut}(F)$ on the images of the kernels in $P_c(F)/P_{c+1}(F)$, we see that if one image is invariant under σ then they all must be invariant. Pulling this back to F, we see that if one kernel is invariant under σ then they all must be invariant.

There are examples of finite p-groups with GI-automorphisms that do not satisfy KIP. The five groups SmallGroup(243,i) for $i = 51, ..., 55$ are the smallest ones. Indeed it appears that for any odd prime p there are exactly five groups of order p^5 failing to satisfy KIP, all of which have $g = 3$ generators but $h = 6$ and so are not Schur σ-ancestor groups. Among the groups of order 729, there are exactly 58 such examples, of which 53 are 3-generated (and have $h = 5, 6$, or 8) and five are 4-generated (and have $h = 10$). Therefore none of the examples of order 729 are Schur σ-ancestor groups.

3 Conjectures

In this section, we formulate our main heuristic assumption, then use the group-theoretical results from the previous section to make precise conjectures about the distribution of p-class tower groups of imaginary quadratic fields as well as

the distribution of their maximal p-class c quotients. Recall that A_K denotes the p-Sylow subgroup of the class group of K.

The arithmetic input, as already noted by Koch and Venkov, is three-fold. First, we observe that for an imaginary quadratic field K, complex conjugation has a natural action on arithmetic objects attached to K. In particular, since $\mathbb{Q}$ has trivial class group, $\mathfrak{a}\bar{\mathfrak{a}}$ is principal for every fractional ideal $\mathfrak{a}$ of K, so complex conjugation acts by inversion on A_K. More generally, complex conjugation acts as an involution on G_K, and as inversion on $G_K^{\mathrm{ab}} \cong A_K$ thanks to the functorial properties of the Artin reciprocity map. The last two ingredients are the finiteness of the class group, and the vanishing of the p-rank of the unit group of $\mathcal{O}_K$. The former ensures that G_K has finite abelianization (as does every one of its open subgroups), and the latter that $r(G_K) = d(G_K)$, by a theorem of Shafarevich [30]. Thus, G_K is always a Schur σ-group.

For $x > 0$, let $\mathcal{F}_x$ denote the set of imaginary quadratic fields with absolute value of discriminant not exceeding x, and for each natural number g, let $\mathcal{F}_{x,g}$ be the subset of $\mathcal{F}_x$ consisting of those fields K having $d(A_K) = g$. For pro-p groups G and H, define $\mathrm{ch}_G(H)$ to be 1 if $G \cong H$ and 0 otherwise.

Definition 3.1 Let G be a finitely generated pro-p group with generator rank g. We define

$$\mathrm{Freq}(G) = \lim_{x\to\infty} \frac{\sum_{K\in\mathcal{F}_{x,g}} \mathrm{ch}_G(G_K)}{\sum_{K\in\mathcal{F}_{x,g}} 1},$$

assuming the limit exists. If G is also finite then, for $c \geq 1$, we define

$$\mathrm{Freq}_c(G) = \lim_{x\to\infty} \frac{\sum_{K\in\mathcal{F}_{x,g}} \mathrm{ch}_G(G_K/P_c(G_K))}{\sum_{K\in\mathcal{F}_{x,g}} 1},$$

assuming the limit exists.

Our main heuristic assumption is that the frequencies defined above exist and are given by the group-theoretical measures introduced in Section 2 when G is a finite p-group. More specifically, we make the following conjecture.

Conjecture 3.2 For every finite p-group G and $c \geq 1$, we have

$$\mathrm{Freq}(G) = \mathrm{Meas}(G)$$
$$\mathrm{Freq}_c(G) = \mathrm{Meas}_c(G).$$

In particular, $\mathrm{Freq}(G) \neq 0$ if and only if G is a Schur σ-group, and $\mathrm{Freq}_c(G) \neq 0$ if and only if G is a Schur σ-ancestor group with p-class c or G is a Schur σ-group with p-class at most c. When G satisfies KIP, the measures can be computed using the formulas provided in Section 2.4.

Remark 3.3 The Cohen-Lenstra heuristics for p-class groups (see Conjecture 1.1) follows from Conjecture 3.2 using Theorems 2.18 and 2.20.

As a consequence of Conjecture 3.2, we expect every *finite* Schur σ-group (respectively Schur σ-ancestor group of p-class c) to occur as G_K (respectively $G_K/P_c(G_K)$) for a positive proportion of imaginary quadratic fields K.

We do not have a conjecture about the value of $\mathrm{Freq}(G)$ when G is an infinite pro-p group. It is worth noting that there are infinite Schur σ-groups that we do not

expect to arise as G_K for any K. For example, the Sylow 3-subgroup of $SL_2(\mathbb{Z}_3)$ considered in [1] is a 2-generator 2-relator pro-3 group with finite abelianization and a GI-automorphism, but the tame case of the Fontaine-Mazur conjecture [16, Conjecture 5a] implies that it does not arise as G_K for any K. It is, however, arbitrarily closely approximated by the finite Schur σ-groups in [1].

4 Index-p-Abelianization-Data (IPAD)

As discussed in Section 1, a complete calculation of G_K is prohibitive for most fields K. In order to make comparisons with data coming from number theory, it will be useful to consider abelianizations of low index subgroups. To that end we introduce the notion of IPAD and an associated measure. Thanks to the p-group generation algorithm, and the theory developed in Section 2, we are able to prove precise values for the measures of the most frequent IPADs when $p = 3$ and $g = 2$. We will compare these values with the observed number-theoretical frequencies in Section 5 (see Table 2).

Definition 4.1 The abelian group $\mathbb{Z}/q_1 \times \cdots \times \mathbb{Z}/q_d$ will be denoted $[q_1, \ldots, q_d]$. Given a g-generated pro-p group G, its Index-p Abelianization Data (or IPAD for short) will be the unordered $(p^g - 1)/(p - 1)$-tuple of abelianizations of the index p subgroups of G augmented by the abelianization of G itself; we always list the latter group first. It will be denoted IPAD(G).

For example, the IPAD of the Schur σ-group `SmallGroup`(243,5) is

$$[[3,3]; [3,3,3][3,9]^3],$$

indicating that its abelianization is $[3,3]$ and those of its four index 3 subgroups are $[3,3,3]$, $[3,9]$, $[3,9]$, and $[3,9]$.

Some other terminology that we will use in this section: for brevity, a descendant of a Schur σ-ancestor group G is called a *Schur descendant* of G if it is also a Schur σ-ancestor group. If it is an immediate descendant then we will call it a *Schur child*. We will sometimes simply say that a group is *Schur* to indicate that it is a Schur σ-ancestor group.

There are some things to note in working with IPADs (see also [9]). First, considering g-generated pro-p groups for a fixed p and g, if H is a quotient of G, then each entry of IPAD(H) is a quotient of a corresponding entry of IPAD(G). This gives a partial order on IPADs and we say that IPAD(H) $\leq$ IPAD(G). Second, if IPAD($G/P_n(G)$) = IPAD($G/P_{n-1}(G)$) (we call the IPAD *settled*), then IPAD(G) = IPAD($G/P_n(G)$). Finally, given a fixed IPAD I, one has IPAD(G) = I if and only if IPAD(G_n) = I once n is sufficiently large (where the bound on n depends only on I). This follows from the fact that if IPAD(G) = I and H is a subgroup of index at most p then the quotient $G/[H,H]$ is a finite p-group with p-class c bounded above by $\log_p |G/[H,H]|$. One has $P_c(G) \subseteq [H,H]$ and so it suffices to choose n to be the largest value of c that can occur over all such H.

Definition 4.2 We define the *measure of an IPAD* I to be sum of $\text{Meas}_n(G)$ over all Schur σ-ancestor groups G of p-class at most n with IPAD(G) = I where n is sufficiently large (see preceding paragraph).

The measure of an IPAD is well defined since one can use Theorem 2.11 to show that the defining sum does not change if n is made larger. In practice, the same theorem often enables one to compute the measure of an IPAD by summing the values of $\mathrm{Meas}_c(G)$ for groups G of p-class c where c is smaller than n depending on how quickly the IPAD stabilizes in various parts of the tree of descendants. This will be illustrated in the proof of the next theorem where we compute, in the case $g = 2$ and $p = 3$, the measures of the most common IPADs.

Theorem 4.3 *(1)* IPAD $[[3,3];[3,3,3][3,9]^3]$ *has measure* $128/729 \approx 0.1756$*;*
(2) IPAD $[[3,9];[3,3,9]^2[3,27]^2]$ *has measure* $256/2187 \approx 0.1171$*;*
(3) IPAD $[[3,3];[3,3,3]^3,[3,9]]$ *has measure* $64/729 \approx 0.0878$*;*
(4) IPAD $[[3,3];[3,3,3]^2[3,9]^2]$ *has measure* $64/729 \approx 0.0878$*;*
(5) IPAD $[[3,3];[3,9]^3[9,27]]$ *has measure* $512/6561 \approx 0.0780$*;*
(6) IPAD $[[3,3];[3,3,3][3,9]^2[9,27]]$ *has measure* $512/6561 \approx 0.0780$*;*
(7) IPAD $[[3,27];[3,3,27]^2[3,81]^2]$ *has measure* $256/6561 \approx 0.0390$*;*
(8) IPAD $[[3,3];[3,3,3]^2[9,27]^2]$ *has measure* $2048/59049 \approx 0.0347$*;*
(9) IPAD $[[3,9];[3,3,9][3,9,27][3,27]^2]$ *has measure* $640/19683 \approx 0.0325$*;*
(10) IPAD $[[3,3];[3,9]^4]$ *has measure* $16/729 \approx 0.0219$*;*
(11) IPAD $[[3,9];[3,3,9][3,27]^3]$ *has measure* $128/6561 \approx 0.0195$*;*
(12) IPAD $[[3,9];[3,3,9][3,27]^2[9,9,9]]$ *has measure* $128/6561 \approx 0.0195$*;*
(13) IPAD $[[3,9];[3,3,3,3][3,27]^3]$ *has measure* $128/6561 \approx 0.0195$*;*
(14) IPAD $[[3,9];[3,9,27][3,27]^3]$ *has measure* $1024/59049 \approx 0.0173$.

Proof We first note that all the groups below whose measures are needed satisfy KIP allowing us to make use of the formulas derived in Section 2. Next, note that the abelianizations of G_1, G_2, G_3 (see Example 2.9) are $[3,3],[3,9],[9,9]$ respectively. It follows that any IPAD with first entry $[3,3]$ has to come from descendants of G_1, and moreover that the first entry is settled, and so every descendant of G_1 has abelianization $[3,3]$. Thus, for all the cases above starting with $[3,3]$, we focus on descendants of G_1. The reader may find it helpful to refer to the tree in Figure 2.1 and discussion at the end of Section 2.4 before reading through the computations which follow.

Using O'Brien's p-group generation algorithm we compute that G_1 has 11 children of p-class 3. Of these, 7 have difference between p-multiplicator rank and nuclear rank at most 2 (in fact exactly 2) and all of these turn out to have a GI-automorphism. Call them $H_1, \ldots, H_7$ in the order produced by O'Brien's algorithm as implemented in `Magma` [4], version 2.16. Of these, H_3 and H_5 are terminal and so are Schur σ-groups. In the standard database they are SmallGroup(243,5) and SmallGroup(243,7) respectively. Their IPADs are those on lines (1) and (4) above. We compute that $\mathrm{Meas}(H_3) = \mathrm{Meas}_3(H_3) = 128/729$ and $\mathrm{Meas}(H_5) = \mathrm{Meas}_3(H_5) = 64/729$.

(1) and (4) follow by establishing that none of the Schur descendants of the other H_i have these IPADs. This also shows that these groups are determined by their IPADs. Note that the latter fact for SmallGroup(243,5) is already observed in [1][Prop. 3.1 and Cor. 3.3]. In relation to (1), only $\mathrm{IPAD}(H_4) \leq \mathrm{IPAD}(H_3)$ (in fact equal). The Schur child of H_4 has IPAD including $[9,9]$ and so does not contribute to (1). In relation to (4), we need to consider H_1, which has the same IPAD as H_5. Only one child of H_1, however, is Schur and its IPAD includes a $[9,9]$ and so cannot contribute to (4). Thus, (1) and (4) are complete.

The Schur child of H_1 has 1602 children of p-class 4, of which 198 are Schur. All of these have IPAD $[[3,3];[3,3,3]^2[9,27]^2]$ and nuclear rank between 2 and 4. All the Schur children of 155 of these have the same IPAD, so are settled and they contribute 2048/59049 to line (8) above. The Schur children of the other 43 include $[27,27]$, so do not count towards (8). The IPADs of the remaining H_i are not less than or equal to this IPAD and so (8) is also complete.

The IPAD of H_2 is that on line (3) and all its children have the same IPAD. It therefore contributes $\mathrm{Meas}_3(H_2) = 64/729$ to (3). None of the other H_i has small enough IPAD that their descendants could have IPAD as in (3), and so (3) is proven.

The IPADs of H_6 and H_7 are both that given in line (10). All the children of H_6 have IPADs involving $[9,9]$, whereas the IPADs of all the children of H_7 are settled as (10). It follows that this IPAD has measure $\mathrm{Meas}_3(H_7) = 16/729$, proving (10).

As for cases (5) and (6), these come from further investigation of descendants of H_6 and H_4 respectively. In each case, the group has a unique Schur child, which then has 6 Schur children. These all have the respective IPADs. In each case, 3 of the 6 are terminal, and the other 3 each have one Schur child. Two of those are settled, whereas the remaining group has larger IPAD. Thus 5 of the 6 Schur grandchildren of each H_i, whose measures are each 64/729, contribute to (5) and (6) respectively and the remaining grandchild, whose measure is 64/6561, does not. Thus the IPADs in (5) and (6) each have measure $64/729 - 64/6561 = 512/6561$, and (5) and (6) are proven.

IPADs (2), (7), (9), (11), (12), (13), and (14) above must come from descendants of G_2. This has 22 Schur children of p-class 3. We call these $J_1, \ldots, J_{22}$ in accordance with O'Brien's ordering. Only J_{10}, J_{11}, and J_{12} have IPADs less than or equal to (in fact equal to) that of (2). The last two are terminal and the unique Schur child of J_{10} has larger IPAD. Thus, the IPAD of (2) has measure $\mathrm{Meas}_3(J_{11}) + \mathrm{Meas}_3(J_{12}) = 256/2187$, and (2) is proven.

The unique Schur child mentioned towards the end of the previous paragraph has IPAD $[[3,9];[3,3,9][3,9,9][3,27]^2]$. A Schur descendant of G_2 with IPAD in line (9) or (12) has to descend from this child (by comparing the IPADs of the other J_i). It has 9 Schur children, of which 6 have the IPAD of (9). The others have IPAD $[[3,9];[3,3,9][9,9,9][3,27]^2]$, which is incomparable with (9) but matches (12). Two of these are terminal, the other settled, and so this allows us to obtain the measure in (12). Of the remaining 6, there are 4 terminal groups, 1 settled, and 1 with a unique Schur child with larger IPAD. Summing the measures of the first 5 groups yields 640/19683 and establishes (9).

Case (7) can only arise from descendants of J_5. It has 3 Schur children, with the 2 terminal ones having the desired IPAD and the other having larger IPAD. This establishes (7).

Case (11) arises from descendants of J_{14} and J_{17}, all of which are settled, and so its measure is the sum of their measures. Case (13) similarly arises from J_{13} and J_{16}, which are settled.

As for (14), this has to come from descendants of J_{15} and J_{18}. Each has measure 64/6561 and their trees of descendants are identical. Each has a unique Schur child and 4 Schur grandchildren. Of these, 1 is terminal and 2 others settled with the desired IPAD. The children of the remaining group have larger IPAD, so we

subtract its measure, 64/59049. Since $2(64/6561 - 64/59049) = 1024/59049$, (14) is proven.

Note that none of the 14 given IPADs have first entry greater than or equal to $[9, 9]$ and so no descendants of G_3 will have one of these IPADs. Since $\text{Meas}_2(G_3) = 1/81 = 0.0123$, the IPADs produced by its descendants will all have measure smaller than that of any of the 14 given IPADs.

Remark 4.4 1. As demonstrated in the proof of the theorem, the measure of an IPAD is usually computed as the sum of the measures of terminal and settled groups. If the sum only involves terminal groups, then it determines a finite list of groups having that IPAD. Sometimes, such as for lines (1) and (4) above, it determines a unique group. Now consider the IPAD in line (7), which corresponds to the two terminal Schur children of J_5. An imaginary quadratic number field with that IPAD (such as $\mathbb{Q}(\sqrt{-17399})$) therefore has one of these two groups as the Galois group of its 3-class tower, the first cases of a non-abelian 3-class tower of a quadratic field having 3-class length 4. This group has derived length 2. We have not found an IPAD consisting only of terminal groups of finite derived length exceeding 2; however, see [11] where additional arithmetic constraints are used to achieve this.

2. In [24], Koch and Venkov proved that if a 2-generated Schur σ-group is finite, then it has relations at depth 3 and k where $k \in \{3, 5, 7\}$ in the p-Zassenhaus filtration. McLeman [26] conjectures that the group is finite if and only if both relations have depth 3. Computing dimensions of the first three factors of the Jennings series, we observe that every Schur descendant of G_1 has its relations at this depth. The apparent combinatorial explosion in descendants of H_1 then casts doubt on the "if" part of McLeman's conjecture.

As for Schur descendants of G_2, those not having both relations at depth 3 are precisely those descended from $J_6, \ldots, J_9, J_{19}, \ldots, J_{22}$. The combinatorial explosion in descendants of these groups lends support to the "only if" part of McLeman's conjecture.

3. One might ask for the probability that a 2-generated Schur σ-group is finite. Searching through the tree in the case $p = 3$, we find 90 descendants of G_1 that are Schur σ-groups of 3-class at most 11, 144 descendants of G_2 that are Schur σ-groups of 3-class at most 8, and 222 descendants of G_3 that are Schur σ-groups of 3-class at most 7. Their combined measure is slightly over 0.8533 and so, in this sense, there is at least an 85.33% probability that a 2-generated Schur σ-group is finite when $p = 3$.

As for an upper bound, it is natural, in the spirit of Golod and Shafarevich, to conjecture that "large" IPADs will correspond only to infinite groups, but one must be careful. Extending the above census slightly, we find that J_1 has Schur σ-group descendants of 3-class 9 and order 3^{18} with IPAD $[[3, 243]; [3, 3, 3, 81], [3, 729]^3]$. Thus, having a rank 4 subgroup of index 3 (the highest rank possible by comparison with the free group) is not sufficient to imply that the Schur σ-group is infinite.

5 Computations

As evidence for our conjectures we have collected numerical data in the case of the smallest odd prime $p = 3$. In particular, we have obtained IPADs for all imaginary

quadratic fields K with 3-class group A_K of rank 2 and discriminant d_K satisfying $|d_K| < 10^8$ assuming GRH.

For an imaginary quadratic field K of discriminant $-d$, with 3-class group of rank 2, the four unramified cubic extensions F of K can be computed using results of Fung and Williams [18]. In each case, the maximal real subfield F^+ of F is a cubic field of discriminant $-d$. Fung and William's work allows one to compute a defining polynomial for such fields efficiently. Indeed, one finds many such polynomials, and we can then distinguish the four isomorphism classes of fields by using `PARI/GP`'s program `nfisisom`. It is then straightforward to compute a defining polynomial for $F = F^+(\sqrt{-d})$ and its class group. Taken together with the 3-class group of K, these give IPAD(G_K).

Originally, we had attempted to compute the unramified extensions of K and their class groups directly but there was a large amount of variation in the running times and this approach proved to be very slow. The current approach was suggested in [25]. There, the author refers to an object called the Transfer Target Type (TTT) of K. The notion of TTT is almost the same as our IPAD except that the 3-class group of the base field is not included.

Computations were carried out using the symbolic algebra package `PARI/GP` [28], version 2.5.4 running on 2×2.66 GHz 6-Core Intel Xeon processors running OS X 10.8.5. The computations were run in parallel across multiple cores by dividing up the discriminants into subintervals and searching through a space of potential defining polynomials for the cubic extensions using the coefficient bounds in [18]. Although this created some redundancy, the parallelization limited the real world running time to the maximum length across all of the intervals. Roughly 890 core hours were used in total.

The class group computations were also double-checked using `Magma` [4], version 2.19-5. A tiny number of discrepancies in the results of the two packages were detected (for less than 30 out of around 1.85 million fields F). These have disappeared since updating to `PARI/GP` [28], version 2.7.3.

Table 5.1 Census of the most common IPADs.

	I_1	$I_{3.2}$	I_{10}	I_{32}	I_{100}
$[3,3]$; $[3,3,3]$ $[3,9]^3$	667	2270	7622	25737	83353
$[3,9]$; $[3,3,9]^2$ $[3,27]^2$	406	1497	4974	16821	55310
$[3,3]$; $[3,3,3]^2$ $[3,9]^2$	269	1069	3625	12314	41398
$[3,3]$; $[3,3,3]^3$ $[3,9]$	297	1056	3619	12324	40968
$[3,3]$; $[3,9]^3$ $[9,27]$	276	973	3190	11042	36458
$[3,3]$; $[3,3,3]$ $[3,9]^2$ $[9,27]$	249	889	3113	10739	35923
$[3,27]$; $[3,3,27]^2$ $[3,81]^2$	103	463	1615	5620	18422
$[3,3]$; $[3,3,3]^2$ $[9,27]^2$	112	384	1293	4593	15541
$[3,9]$; $[3,3,9]$ $[3,9,27]$ $[3,27]^2$	101	367	1317	4559	15037
$[3,3]$; $[3,9]^4$	94	323	1019	3284	10426
$[3,9]$; $[3,3,9]$ $[3,27]^3$	75	254	844	2914	9335
$[3,9]$; $[3,3,3,3]$ $[3,27]^3$	64	233	799	2734	9000
$[3,9]$; $[3,3,9]$ $[3,27]^2$ $[9,9,9]$	66	229	786	2740	8955
$[3,9]$; $[3,9,27]$ $[3,27]^3$	61	232	728	2447	8165
Other IPADs (331 types)	350	1505	5741	21222	73634
Total	3190	11744	40285	139090	461925

Table 5.2 Relative proportions of the most common IPADs.

	I_1	$I_{3.2}$	I_{10}	I_{32}	I_{100}	Predicted
$[3,3]$; $[3,3,3]$ $[3,9]^3$	0.2091	0.1933	0.1892	0.1850	0.1804	0.1756
$[3,9]$; $[3,3,9]^2$ $[3,27]^2$	0.1273	0.1275	0.1235	0.1209	0.1197	0.1171
$[3,3]$; $[3,3,3]^2$ $[3,9]^2$	0.0843	0.0910	0.0900	0.0885	0.0896	0.0878
$[3,3]$; $[3,3,3]^3$ $[3,9]$	0.0931	0.0899	0.0898	0.0886	0.0887	0.0878
$[3,3]$; $[3,9]^3$ $[9,27]$	0.0865	0.0829	0.0792	0.0794	0.0789	0.0780
$[3,3]$; $[3,3,3]$ $[3,9]^2$ $[9,27]$	0.0781	0.0757	0.0773	0.0772	0.0778	0.0780
$[3,27]$; $[3,3,27]^2$ $[3,81]^2$	0.0323	0.0394	0.0401	0.0404	0.0399	0.0390
$[3,3]$; $[3,3,3]^2$ $[9,27]^2$	0.0351	0.0327	0.0321	0.0330	0.0336	0.0347
$[3,9]$; $[3,3,9]$ $[3,9,27]$ $[3,27]^2$	0.0317	0.0313	0.0327	0.0328	0.0326	0.0325
$[3,3]$; $[3,9]^4$	0.0295	0.0275	0.0253	0.0236	0.0226	0.0219
$[3,9]$; $[3,3,9]$ $[3,27]^3$	0.0235	0.0216	0.0210	0.0210	0.0202	0.0195
$[3,9]$; $[3,3,3,3]$ $[3,27]^3$	0.0201	0.0198	0.0198	0.0197	0.0195	0.0195
$[3,9]$; $[3,3,9]$ $[3,27]^2$ $[9,9,9]$	0.0207	0.0195	0.0195	0.0197	0.0194	0.0195
$[3,9]$; $[3,9,27]$ $[3,27]^3$	0.0191	0.0198	0.0181	0.0176	0.0177	0.0173
Other IPADs (331 types)	0.1097	0.1282	0.1425	0.1526	0.1594	0.1717

We now present a summary of the data collected. We have broken down the interval of discriminants d_K with $1 < |d_K| < 10^8$ into 5 nested subintervals I_j where $I_j = \mathcal{F}_{j\cdot 10^6,2} = \{d_K \mid 1 < -d_K < j \cdot 10^6 \text{ and } d(A_K) = 2\}$ and we have selected values of j so that the upper bound for each successive subinterval is scaled by a factor of $\sqrt{10} \approx 3.2$.

The first table is a census of the most common IPADs. The second lists their relative proportions obtained by dividing each entry in the first by the corresponding column total. In addition, the last column of the second table lists the values predicted by our conjectures as computed in Theorem 4.3. Note that in lines 1 and 3 of Table 2, the IPAD determines the isomorphism type of the group, namely SmallGroup(243,5) and SmallGroup(243,7) respectively. Thus, on these two lines, the predicted and computed frequencies for individual non-abelian groups are being compared, providing a direct test of our heuristics.

6 Appendix: On the nucleus of certain p-groups – by Jonathan Blackhurst

In this appendix we[3] prove the proposition that if the Schur multiplier of a finite non-cyclic p-group G is trivial, then the nucleus of G is trivial. Our proof of the proposition will use the facts that a p-group has trivial nucleus if and only if it has no immediate descendants and that a finite group has trivial Schur multiplier if and only if it has no non-trivial stem extensions, so we will begin by recalling a few definitions. For the definition of the lower p-central series and p-class of a group, we refer to section 2 of the article.

Definition 6.1 Let G be a finite p-group with minimal number of generators $d = d(G)$ and presentation F/R where F is the free pro-p group on d generators. The *p-covering group* G^* of G is F/R^* where R^* is the topological closure of $R^p[F,R]$,

[3] Jonathan Blackhurst
1037 E Millbrook Way, Bountiful, UT 84010 USA
E-mail: jblackhurst@gmail.com

and the *nucleus* of G is $P_c(G^*)$ where c is the p-class of G. The *p-multiplicator* of G is defined to be the subgroup R/R^* of G^*. The *Schur multiplier* $\mathcal{M}(G)$ of G is defined to be $(R \cap [F,F])/[F,R]$. A group C is a *stem extension* of G if there is an exact sequence

$$1 \to K \to C \to G \to 1$$

where K is contained in the intersection of the center and derived subgroups of C.

We will need to recall some basic properties of Schur multipliers and p-covering groups. First, for a finite group G, the largest stem extension of G has size $|G||\mathcal{M}(G)|$. Hence, the Schur multiplier of a finite group G is trivial if and only if G admits no non-trivial stem extensions. Second, every elementary abelian central extension of G is a quotient of G^*. By this we mean that if H is a d-generated p-group with elementary abelian subgroup Z contained in the center of H such that H/Z is isomorphic to G, then H is a quotient of G^*. Every immediate descendant of G is an elementary abelian central extension of G, hence is a quotient of G^*. A subgroup M of the p-multiplicator of G is said to supplement the nucleus if M and the nucleus together generate the p-multiplicator, that is $MP_c(G^*) = R/R^*$. The immediate descendants of G can be put in one-to-one correspondence with equivalence classes of proper subgroups M of the p-multiplicator of G that supplement the nucleus. The equivalence relation comes from the action of the outer automorphism group of G^*, so M and N are equivalent if there is an outer automorphism σ of G^* such that $\sigma(M) = N$. The reader is referred to O'Brien [27] for more details.

With these preliminaries in place, we can show that the non-cyclic hypothesis in our proposition is necessary by considering the finite cyclic p-group $G = \mathbb{Z}/p^c\mathbb{Z}$. The Schur multiplier is trivial since in this case $F = \mathbb{Z}$ so $[F,F]$ is trivial. On the other hand, the nucleus is non-trivial since in this case $F = \mathbb{Z}_p$ and $R = p^c\mathbb{Z}_p$ so $R^* = p^{c+1}\mathbb{Z}_p$ and $G^* = F/R^* = \mathbb{Z}/p^{c+1}\mathbb{Z}$ which implies that $P_c(G^*) = p^cG^*$ is non-trivial.

Proposition 6.2 *Let G be a finite non-cyclic p-group. If the Schur multiplier of G is trivial, then the nucleus of G is trivial.*

Proof We will prove the following equivalent assertion: if the nucleus of G is non-trivial, then G has a non-trivial stem extension. We divide the problem into two cases depending on whether the abelianization of G has stabilized; that is, whether the abelianization of an immediate descendant of G can have larger order than the abelianization G^{ab} of G. We will see that this is equivalent to whether or not $G^{ab} \simeq (G/P_{c-1}(G))^{ab}$ where G has p-class c.
CASE 1: Suppose that $G^{ab} \simeq (G/P_{c-1}(G))^{ab}$ and that the nucleus of G is non-trivial. Since the nucleus is non-trivial, G has an immediate descendant C and we have the following diagram

$$1 \to K \to C \to G \to 1$$

where $K = P_c(C)$. Note that since $C/P_k(C) \simeq G/P_k(G)$ for $k \leq c$, we have that $(C/P_{c-1}(C))^{ab} \simeq (C/K)^{ab}$. If $P_{c-1}(C)$ were not contained within the derived subgroup C' of C, then its image $\overline{P_{c-1}}(C)$ in C/C' would be non-trivial. Since $K = P_{c-1}(C)^p[C, P_{c-1}(C)]$, the image $\overline{K}$ of K would be $\overline{P_{c-1}}(C)^p$ and thus would

be stricly smaller than $\overline{P_{c-1}}(C)$. Now $(C/H)^{ab} \simeq (C/C')/\overline{H}$ for any $H \triangleleft C$, so, replacing H with K and $P_{c-1}(C)$, we see that $(C/P_{c-1}(C))^{ab}$ would be smaller than $(C/K)^{ab}$, contradicting that they are isomorphic. Thus $P_{c-1}(C) < C'$, hence $K < C'$, so C is a stem extension of G. Since G has a non-trivial stem extension, its Schur multiplier is non-trivial.
CASE 2: Suppose that $G^{ab} \not\simeq (G/P_{c-1}(G))^{ab}$. Let

$$1 \to R \to F \to G \to 1$$

be a presentation of G where F is free pro-p group on d generators and d is the minimal number of generators of G. Induction and the argument in the preceding case shows that $(G/P_k(G))^{ab}$ is strictly smaller than $(G/P_{k+1}(G))^{ab}$ for any $k < c$. Furthermore, since the image $\overline{P_{k+1}}(G)$ of $P_{k+1}(G)$ in G/G' is $\overline{P_k}(G)^p$, there must be a generator b of F such that the image of $b^{p^{c-1}}$ in G lies outside G'. Now consider $R^* = R^p[F,R]$ and let $G^* = F/R^*$ be the p-covering group of G. We have the following diagrams:

$$1 \to R^* \to F \to G^* \to 1$$

and

$$1 \to R/R^* \to G^* \to G \to 1$$

We now show that the image of b^{p^c} in $P_c(G^*)$ is non-trivial so G has non-trivial nucleus. Let G have abelianization isomorphic to $\mathbb{Z}/p^{n_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p^{n_d}\mathbb{Z}$. Consider the topological closure S of $R \cup [F,F]$. Then F/S is isomorphic to G^{ab}. The group $\mathbb{Z}/p^{n_1+1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p^{n_d+1}\mathbb{Z}$ is an elementary abelian central extension of F/S. This implies that b^{p^c} lies outside $S^* = S^p[F,S]$. Since $R \subset S$, we have that $R^* \subset S^*$. Hence b^{p^c} lies outside R^* so it has non-trivial image in G^*. Since its image lies inside $P_c(G^*)$, this group is non-trivial.

We have shown that G has non-trivial nucleus. Now let a be a generator of F independent of b—i.e., one that doesn't map to the same element as b in the elementary abelianization of F—and let $\overline{M}$ be a proper subgroup of R/R^* that contains the image of $b^{p^c}[a, b^{p^{c-1}}]$ and that supplements the subgroup of R/R^* generated by the image of b^{p^c} (so $\overline{M}$ and the image of b^{p^c} generate R/R^*). Now consider $C = G^*/\overline{M}$. Letting $K = (R/R^*)/\overline{M}$, we have the following diagram

$$1 \to K \to C \to G \to 1$$

Since G^* is a central extension of G and C is a quotient of G^*, C is also a central extension of G. Furthermore, $|K| = p$. Now let M be the subgroup of F corresponding to $\overline{M}$ under the lattice isomorphism theorem. Then we have the following diagram:

$$1 \to M \to F \to C \to 1$$

Since M does not contain b^{p^c}, its image in C is non-trivial. Since G has p-class c, the image of b^{p^c} is trivial in G. Also since $|K| = p$, the image of the powers of b^{p^c} constitute K. Since M does contain $b^{p^c}[a, b^{p^{c-1}}]$, the image of b^{p^c} in C equals the image of $[b^{p^{c-1}}, a]$, hence K lies in the derived subgroup of C, so C is a non-trivial stem extension of G. Consequently, the Schur multiplier of G is non-trivial.

References

1. L. Bartholdi and M.R. Bush, *Maximal unramified 3-extensions of imaginary quadratic fields and $SL_2(\mathbb{Z}_3)$*, J. Number Theory **124** (2007), no. 1, 159–166.
2. H. U. Besche, B. Eick, and E. A. O'Brien, *A millennium project: constructing small groups*, Internat. J. Algebra Comp., **12** (2002), no. 5, 623–644.
3. M. Bhargava, *The density of discriminants of quartic rings and fields* Ann. of Math. (2) **162** (2005), no. 2, 10311063.
4. W. Bosma, J. J. Cannon, C. Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput., **24** (1997), 235–265.
5. N. Boston, *Random pro-p groups and random Galois groups*, Annales des Sciences Mathématiques du Québec. **32** (2008), no. 2, 125–138.
6. N. Boston, M.R. Bush, F. Hajir, *Heuristics for p-class towers of real quadratic fields*, in preparation.
7. N. Boston and J. Ellenberg, *Random pro-p groups, braid groups, and random tame Galois groups*, Groups Geom. Dyn. **5** (2011), no. 2, 265-280.
8. N. Boston and C.R. Leedham-Green, *Explicit computation of Galois p-groups unramified at p*, Journal of Algebra **256** (2002), no. 2, 402–413.
9. N. Boston and H. Nover, *Computing pro-p Galois groups*, Lecture Notes in Computer Science **4076**, ANTS VII, 1-10.
10. N. Boston and M. M. Wood, *Non-abelian Cohen-Lenstra heuristics over function fields*, submitted, arXiv: 1604.03433, 2016.
11. M. R. Bush, D. C. Mayer, *3-class field towers of exact length* 3, J. Number Theory **147** (2015), 766–777.
12. H. Cohen and H.W. Lenstra, Jr., *Heuristics on class groups*, in: Number Theory, 26–36, LNM **1052**, Springer, Berlin, 1984.
13. H. Cohen and H. W. Lenstra, Jr., *Heuristics on class groups of number fields*, pp. 33–62 in: Number theory, Noordwijkerhout 1983, LNM **1068**, Springer, Berlin, 1984.
14. H. Cohen and J. Martinet, *Étude heuristique des groupes des classes des corps de nombres*, J. Reine Angew. Math. **404** (1990), 39–76.
15. H. Davenport and H. Heilbronn, *On the density of discriminants of cubic fields (ii)*, Proc. Roy. Soc. Lond. A **322** (1971), 405–420.
16. J.-M. Fontaine and B. Mazur, *Geometric Galois representations.* Elliptic Curves, Modular Forms & Fermats Last Theorem (J. H. Coates and S. T. Yau, eds.), Internat. Press, Massachusetts, 1995, Proceedings of the conference on elliptic curves and modular forms held at the Chinese University of Hong Kong, December 18–21,1993, pp. 41–78
17. E. Friedman and L. C. Washington, *On the distribution of divisor class groups of curves over a finite field*, Théorie des nombres (Quebec, PQ, 1987), 227–239, de Gruyter, Berlin, 1989.
18. G.W. Fung and H.C. Williams, *On the computation of a table of complex cubic fields of discriminant $D > -10^6$*, Proc. AMS **55** (1990), 313–325.
19. D. Gildenhuys, W. Herfort, and L. Ribes, *Profinite Frobenius groups*, Arch. Math. (Basel) **33** (1979/80), no. 6, 518–528.
20. K.W. Gruenberg, *Cohomological topics in group theory*, LNM **143**, Springer, Berlin, 1970.
21. P. Hall, *A contribution to the theory of groups of prime-power order*, Proc. London Math. Soc. **36** (1934), 29–95.
22. P. Hall, *A partition formula connected with Abelian groups*, Comment. Math. Helv. **11** (1938-39), 126-129.
23. B. Huppert, N. Blackburn, *Finite Groups II*, Springer-Verlag, Berlin, 1982.
24. H. Koch and B.B. Venkov, *Über den p-Klassenkörperturm eines imaginär-quadratischen Zahlkörpers*, Soc. Math. France, Astérisque **24-25** (1975), 57–67.
25. D. C. Mayer, *Principalization algorithm via class group structure*, accepted by J. Théor. Nombres Bordeaux, 2014.
26. C. McLeman, *Class field towers over quadratic imaginary number fields*, Annales des Sciences Mathématiques du Québec **32** (2008), no 2, 199–209.
27. E.A. O'Brien, *The p-group generation algorithm*, J. Symbolic Comput. **9** (1990), 677–698.
28. The PARI Group, PARI/GP `2.7.0`, Bordeaux, 2014, `http://pari.math.u-bordeaux.fr/`.
29. A. Scholz and O. Taussky, *Die Hauptideale der kubischen Klassenkörper imaginär-quadratischer Zahlkörper*, J. Reine Angew. Math. **171** (1934), 19–41.
30. I. Shafarevich, *Extensions with prescribed ramification points*, Inst. Hautes Études Sci. Publ. Math. 18 (1964), 7195 [In Russian]; English translation "Amer. Math. Soc. Transl.," Vol. 59, pp. 128149, Amer. Math. Soc., Providence, RI, 1966.

31. C. Wittmann, *p-class groups of certain extensions of degree p*, Math. Comp. **74** (2005), no. 250, 937-947.