THE CHEBOTAREV INVARIANT OF A FINITE GROUP

EMMANUEL KOWALSKI AND DAVID ZYWINA

Abstract. We consider invariants of a finite group G related to the number of random (independent, uniformly distributed) conjugacy classes which are required to generate it. These invariants are intuitively related to problems of Galois theory. We find group-theoretic expressions for them and investigate their values both theoretically and numerically.

1. Introduction

A well-known method to compute the Galois group H of a number field (e.g., of the splitting field of a polynomial $P \in \mathbf{Z}[T]$ with integral coefficients) can be described roughly as follows: (1) find a group G which contains H, e.g., because of symmetry considerations (2) this group G is our "guess" for H, and we try to prove that $H = G$ by reducing modulo successive primes, using the fact that the Frobenius automorphisms give *conjugacy classes* in the Galois group H, and hence conjugacy classes in G.

If the guess in (1) was right, and if one is patient enough in (2) that the conjugacy classes observed are only compatible with the Galois group being our candidate G, then we have succeeded.

This method is particularly simple when G is "guessed" to be the symmetric group acting on the roots of the polynomial: it can lead quickly to examples of equations with this Galois group. In general, however this is not the most efficient algorithm (if only because the first step (1) is hard to formalize!), and thus computer algebra systems use other techniques. Still, this method is well-suited for certain theoretical investigations, for instance for probabilistic Galois theory (see, e.g., [G]), and it can be surprisingly efficient even for fairly complicated groups (see for instance our joint works [JKZ] and [JKZ2] with F. Jouve, for a case where the Weyl group of the exceptional Lie group E_8 is the Galois group involved, and for for further developments along these lines).

In view of this, it is somewhat surprising that no general study of the efficiency of the underlying algorithm seems to have been performed. Among the very few references we know is a paper of Dixon [D1], who considers informally the case of the symmetric groups $\mathfrak{S}_n$ and mentions some earlier work of McKay. On the other hand, there has been a fair amount of interest in the question of determining the probability that a tuple of elements generate a finite group, which is the analogue problem where conjugacy is ignored, see for instance the paper [KL] of Kantor and Lubotzky. The paper [P] of C. Pomerance considers the question for abelian groups, when the conjugacy issue is also irrelevant, and his results do apply to our setting. The current paper will provide the beginning of the theoretical analysis of this type of algorithm for general finite groups. As a specific result, we will prove the following

2000 *Mathematics Subject Classification.* 20P05, 20Dxx, 20K01, 20F69, 60Bxx.

Key words and phrases. Chebotarev density theorem, coupon collector problems, Galois group, conjugacy classes generating a group, probabilistic group theory.

result (informally stated; Theorem 6.1 gives the precise statement using the definitions of Section 2):

Theorem 1.1 (Boundedness of Chebotarev invariants for symmetric groups)**.** *There exists a constant $c > 0$ with the following property: for all integers $n \geqslant 1$, the average number of independently, randomly chosen conjugacy classes*[1] *of the symmetric group $\mathfrak{S}_n$ one must pick before ensuring that any tuple of elements taken from each of these classes generate $\mathfrak{S}_n$, is at most c. In fact, for any $k \geqslant 1$, there exists $c_k \geqslant 0$ such that the average of the k-th power of this number is bounded by c_k for all n.*

Here is the rough outline of this work: we consider probabilistic models in Section 2, and define an invariant, which we call the *Chebotarev invariant* of a finite group, using such a model (the name, based on the Chebotarev density theorem, is justified in Section 8); it makes precise the informal notion in the statement of Theorem 1.1. Computing this invariant is seen to be related to very interesting questions of group theory, independently of any arithmetic motivation. In Section 3, we indicate how to compute this invariant for abelian groups (based on Pomerance's work) and in Section 4 we consider solvable groups of a certain "extremal" type. In Sections 5, 6 and 7, we consider theoretical and numerical examples for non-abelian, often non-solvable, groups – in particular alternating and symmetric groups, proving Theorem 1.1. Finally, Section 8 makes some informal remarks concerning the applicability of our results to arithmetic situations (the original motivation); as we will see, there are non-trivial difficulties involved, and we hope to come back to these questions later.

Notation. As usual, $|X|$ denotes the cardinality of a set, and $\mathbf{F}_q$ is a field with q elements. If G is a finite group, and $H \subset G$, we write

$$\nu_G(H) = \nu(H) = \frac{|H|}{|G|}.$$

We write $G^\sharp$ for the set of conjugacy classes of G, and for $C \subset G^\sharp$, we also write $\nu_G(C)$ or $\nu(C)$ for $\nu(\tilde{C})$, where $\tilde{C} \subset G$ is the union of all conjugacy classes in C.

In fact, as a matter of convenience, we will usually denote in the same way a subset of conjugacy classes and the corresponding set of elements in G, unless it is not clear in context if $c \in C$ means that c is a conjugacy class or an element of G (we will write often $c^\sharp$ for a conjugacy class, avoiding most ambiguity).

We recall that a geometric random variable X with parameter $p \in [0,1]$ on a probability space is a random variable taking almost surely values in the set of positive integers, with

$$\boldsymbol{P}(X = k) = p(1-p)^{k-1} \tag{1.1}$$

for $k \geqslant 1$. We then have

$$\mathbf{E}(X) = p\sum_{k\geqslant 1} k(1-p)^{k-1} = \frac{1}{p}, \qquad \mathbf{E}(X^2) = \frac{2-p}{p^2}, \qquad \mathbf{V}(X) = \frac{1-p}{p^2}. \tag{1.2}$$

By $f \ll g$ for $x \in X$, or $f = O(g)$ for $x \in X$, where X is an arbitrary set on which f is defined, we mean synonymously that there exists a constant $C \geqslant 0$ such that $|f(x)| \leqslant Cg(x)$ for all $x \in X$. The "implied constant" refers to any value of C for which this holds. It may depend on the set X, which is usually specified explicitly, or clearly determined by the

[1] This means distributed in proportion with the size of the conjugacy class.

context. Similarly, $f \asymp g$ means that $f \ll g$ and $g \ll f$ on the same set. On the other hand $f(x) \sim g(x)$ as $x \rightarrow x_0$ means that $f(x)/g(x) \rightarrow 1$ as $x \rightarrow x_0$.

2. The Chebotarev invariant of a finite group

In this section, we describe a natural probabilistic model for the recognition algorithm described previously. Fix a finite group G. We first remark that, whereas it does not make sense to say that a conjugacy class lies in a certain subgroup, unless the latter is a normal subgroup, it does make sense to say that it lies in a conjugacy class of subgroups. With that in mind, we define:

Definition 2.1. Let G be a finite group, and let $C = \{C_1, \ldots, C_m\} \subset G^\sharp$ be a subset of conjugacy classes in G. Then C *generates* G if, for any choice of representatives $g_i \in C_i$ for $1 \leqslant i \leqslant m$, the elements of the tuple $(g_1, \ldots, g_m)$ generate G. Equivalently, C generates G if and only if there is no (proper) maximal subgroup H of G that has non-empty intersection with each of the C_i. [2]

The equivalence of the two definitions is quite clear contrapositively: if there are $g_i \in C_i$ which generate a proper subgroup H_1, then each C_i intersects any maximal proper subgroup H of G that contains H_1, and conversely. Note also that the second condition can be stated by saying that there is a conjugacy class of maximal subgroups containing C.

The following well-known lemma (due to Jordan) is the basic fact underlying the whole technique:

Lemma 2.2. *Let G be a finite group. Then the set $G^\sharp$ of conjugacy classes generates H. In other words, there is no proper subgroup of G which contains a representative from each conjugacy class.*

Simple as this is, one should also recall at this point that the analogue of this lemma is false for infinite groups (even compact groups, such as $SU(n)$, $n \geqslant 2$); for further discussion of various interpretations of this lemma, see [S1].

Now, let $(\Omega, \Sigma, \boldsymbol{P})$ be a fixed probability space with a sequence $X = (X_n)_{n \geqslant 1}$ of G-valued random variables

$$X_n : \Omega \rightarrow G,$$

and let $X_n^\sharp$ be the conjugacy class of X_n in $G^\sharp$: those are $G^\sharp$-valued random variables.

Intuitively, those $(X_n^\sharp)$ are the conjugacy classes that we see coming "one by one"; the Chebotarev invariant(s) looks at when we get enough information to conclude that those conjugacy classes can not all belong to some proper subgroup of G.

We now define a random variable $\tau_{X,G}$ (a *waiting time*) by

$$\tau_{X,G} = \min\{n \geqslant 1 \mid (X_1^\sharp, \ldots, X_n^\sharp) \text{ generate } G\} \in [1, +\infty]. \tag{2.1}$$

This depends on the sequence $X = (X_n)$, and it may be always infinite (e.g., if $X_n = 1$ for all n!). But it is, in an intuitive sense, the "finest" invariant in terms of this probabilistic model. To obtain more compact and purely numerical invariants, it is natural to first take the expectation; this takes values in $[1, +\infty]$.

[2] Alternately, following [D1], one says that elements $(g_1, \ldots, g_m)$ *invariably generate* G if their conjugacy classes generate G in the sense above.

Definition 2.3. Let G be a finite group, $X = (X_n)$ a sequence of G-valued random variables and $\tau_{X,G}$ the waiting time above. The *Chebotarev invariant* of G with respect to X, denoted $c(G; X)$, is the expectation $c(G; X) = \mathbf{E}(\tau_{X,G})$ of this random variable.

Remark 2.4. We focused on conjugacy classes because this is how applications to Galois theory are likely to arise, but of course one can similarly define an invariant using the original random elements (X_n) in G. If G is abelian, the two coincide.

To have an unambiguously defined invariant, we must use a specific choice of sequence (X_n). The natural model is that of independent, uniformly distributed elements in G: if (X_n) are independent and identically uniformly distributed G-valued random variables, so that

$$\boldsymbol{P}(X_n = g) = \frac{1}{|G|} \qquad \text{for all } g \in G, \quad \text{and all } n \geqslant 1,$$

and hence

$$\boldsymbol{P}(X_n^\sharp = g^\sharp) = \frac{|g^\sharp|}{|G|}, \qquad \text{for all } g^\sharp \in G^\sharp, \quad \text{and all } n \geqslant 1,$$

then we call $c(G; X)$ *the* Chebotarev invariant, and we just write $c(G)$.

Remark 2.5. It may be of interest, at least for numerical purposes, to use a sequence (X_n) which is not independent, but is obtained, for instance, by a rapidly mixing random walk on G. Also, the arithmetic analogue for computing Galois groups may be interpreted as involving non-independent and non-uniform choices of conjugacy classes (see Section 8).

Other numerical invariants may of course be derived from $\tau_{X,G}$, starting from the higher moments $\mathbf{E}(\tau_{X,G}^k)$ for $k \geqslant 1$. Since it is probabilistically most important, when the expectation of a random variable is known, to also have a control of its second moment, we define formally the *secondary Chebotarev invariant*:

Definition 2.6. Let G be a finite group, $X = (X_n)$ a sequence of G-valued random variables, and let $\tau_{X,G}$ be the waiting time above. The *secondary Chebotarev invariant* of G with respect to X, is the second moment $c_2(G; X) = \mathbf{E}(\tau_{X,G}^2)$. If (X_n) is a sequence of independent, uniformly distributed random variables, then we write $c_2(G)$ and call it *the* secondary Chebotarev invariant.

We will now give formulas for the two Chebotarev invariants (in the independent case), which are expressed purely in terms of group-theoretic information. This is useful for explicit computations, at least for groups which are very well understood (but often the probabilistic origin of $c(G)$ should also be kept in mind.)

To state the formulas, we must introduce the following data and notation about G. Let $\max(G)$ be the set of conjugacy classes of (proper) maximal subgroups of G (if G is trivial, this is empty); for a conjugacy class of of maximal subgroups $\mathcal{H} \in \max(G)$, let $\mathcal{H}^\sharp$ denote the set of conjugacy classes C of G which "occur in $\mathcal{H}$", i.e., such that $C \cap H_1 \neq \emptyset$ for *some* H_1 in the conjugacy class $\mathcal{H}$.[3] Moreover, if $I \subset \max(G)$ is a set of conjugacy classes of maximal subgroups, we let

$$\mathcal{H}_I^\sharp = \bigcap_{\mathcal{H} \in I} \mathcal{H}^\sharp, \tag{2.2}$$

[3] Note that this depends on the underlying group G.

the set of conjugacy classes of G which appear in all subgroups in I.

Then we have:

Proposition 2.7. *Let G be a* non-trivial *finite group. With notation as above, we have*

$$c(G) = \sum_{\substack{I \subset \max(G) \\ I \neq \emptyset}} \frac{(-1)^{|I|+1}}{1-\nu(\mathcal{H}_I^\sharp)}, \tag{2.3}$$

and

$$c_2(G) = \sum_{\substack{I \subset \max(G) \\ I \neq \emptyset}} \frac{(-1)^{|I|}}{1-\nu(\mathcal{H}_I^\sharp)}\Bigl(1 - \frac{2}{1-\nu(\mathcal{H}_I^\sharp)}\Bigr) = \sum_{\substack{I \subset \max(G) \\ I \neq \emptyset}} (-1)^{|I|+1} \frac{1+\nu(\mathcal{H}_I^\sharp)}{(1-\nu(\mathcal{H}_I^\sharp))^2}. \tag{2.4}$$

These formulas do not apply for the trivial group, since they lead to empty sums which are zero, whereas the definition leads to[4] $c(1) = 1$, $c_2(1) = 1$.

Probabilists will have noticed that the first formula (at least) is very similar to the one for the expectation of the waiting time for a general coupon collector problem. There is indeed a link, which is provided by the next lemma where independence of the random elements X_n is not required.

Lemma 2.8. *Let G be a non-trivial finite group and $X = (X_n)$ a sequence of G-valued random variables. The waiting time $\tau_{X,G}$ is equal to*

$$\tau_{X,G} = \max_{\mathcal{H} \in \max G} \hat{\tau}_{\mathcal{H}},$$

where

$$\hat{\tau}_{\mathcal{H}} = \min\{n \geqslant 1 \mid X_n^\sharp \notin \mathcal{H}^\sharp\} ; \tag{2.5}$$

note that $\hat{\tau}_{\mathcal{H}}$ depends also on X.

In other words, $\tau_{X,G}$ is also the maximal n such that we need to look at X_i for i up to n, before we witness, for every conjugacy class $\mathcal{H}$ of maximal subgroups, some X_n which is incompatible with the groups in this class $\mathcal{H}$. This is very close to a coupon collector problem (see, e.g., [FGT] for a general description of this type of problems), where the "coupons" we need to collect correspond to the conjugacy classes which are not in $\mathcal{H}^\sharp$, as $\mathcal{H}$ ranges over $\max(G)$. But since a single X_n may serve as coupon for more than one $\mathcal{H}^\sharp$, this does not exactly correspond to standard coupon collector problems.[5] Because of this, we state and prove the following general abstract result, which may have other applications.

Proposition 2.9. *Let $(\Omega, \Sigma, \boldsymbol{P})$ be a probability space, D a finite set. Let (Z_n) be a sequence of D-valued random variables. Let $\mathcal{E}$ be a non-empty finite collection of non-empty subsets of D, and let*

$$\tau_{\mathcal{E}} = \min\{n \geqslant 1 \mid \text{ for all } E \in \mathcal{E}, \text{ there exists some } m \leqslant n \text{ with } Z_m \in E\}$$

[4] One may argue that there is no need to look at any elements to be sure of generating the trivial group, but this does not correspond to the definition.

[5] This has been called the "coupon subset collection problem" by Adler and Ross [A].

be the waiting time before all subsets $E \in \mathcal{E}$ have been witnessed in the sequence (Z_n). For $I \subset \mathcal{E}$, non-empty, let

$$T_I = \min\{n \geqslant 1 \mid Z_n \in E \text{ for some subset } E \in I\}. \tag{2.6}$$

(1) *Assume $T_I < +\infty$ almost surely for all non-empty subsets $I \subset \mathcal{E}$. Then we have*

$$\tau_{\mathcal{E}} = \sum_{\emptyset \neq I \subset \mathcal{E}} (-1)^{|I|+1} T_I. \tag{2.7}$$

(2) *Assume the Z_n are independent and identically distributed random variables and let μ be their common law. We have*

$$\mathbf{E}(\tau_{\mathcal{E}}) = \sum_{\substack{I \subset \mathcal{E} \\ I \neq \emptyset}} \frac{(-1)^{|I|+1}}{\boldsymbol{P}(Z_n \in \bigcup_{E \in I} E)} = \sum_{\substack{I \subset \mathcal{E} \\ I \neq \emptyset}} \frac{(-1)^{|I|+1}}{\mu(\bigcup_{E \in I} E)}, \tag{2.8}$$

and if the subsets in $\mathcal{E}$ are disjoint, we have

$$\mathbf{E}(\tau_{\mathcal{E}}) = \int_0^{+\infty} \Bigl(1 - \prod_{E \in \mathcal{E}} (1 - \exp(-\mu(E)t))\Bigr) dt. \tag{2.9}$$

(3) *We have*

$$\mathbf{E}(\tau_{\mathcal{E}}^2) = \sum_{\substack{I \subset \mathcal{E} \\ I \neq \emptyset}} \frac{(-1)^{|I|}}{\mu(\bigcup_{E \in I} E)} \Bigl(1 - \frac{2}{\mu(\bigcup_{E \in I} E)}\Bigr). \tag{2.10}$$

When $\mathcal{E}$ is the set of singletons in D, where we have exactly the coupon collector problem, the formulas for the expectation are well-known (see, e.g., [FGT, Theorem 4.1] for the integral formula); we have not seen general formulas for the second moment in the literature.

Proof of Proposition 2.9. To simplify notation, define

$$E_I = \bigcup_{E \in I} E, \tag{2.11}$$

for each $I \subset \mathcal{E}$. Formula (2.7) – which implies in particular that $\tau_{\mathcal{E}}$ is finite almost surely – can be checked, e.g., by seeing $\tau_{\mathcal{E}}$ as the length of the subset

$$\bigcup_{E \in \mathcal{E}} [0, T_{\{E\}}] \subset \mathbf{R},$$

(which is therefore almost surely finite by assumption on the T_I), and applying the inclusion-exclusion formula for the measure of a union of finitely many sets (for the Lebesgue measure, or the counting measure on $\mathbf{Z}$, equivalently):

$$\Bigl|\bigcup_{E \in \mathcal{E}} [0, T_{\{E\}}]\Bigr| = \sum_{\emptyset \neq I \subset \mathcal{E}} (-1)^{|I|+1} \Bigl|\bigcap_{E \in I} [0, T_{\{E\}}]\Bigr|,$$

at which point it is enough to observe that, for any $I \subset \mathcal{E}$, we have

$$\Bigl|\bigcap_{E \in I} [0, T_{\{E\}}]\Bigr| = \min_{E \in I} T_{\{E\}} = T_I.$$

We can now finish the computation of $\mathbf{E}(\tau_{\mathcal{E}})$ in (2), in the case of independent random variables. Indeed, in that case the random variable T_I is distributed like a geometric random variable with parameter $p = \boldsymbol{P}(Z_n \in E_I)$ (see (1.1)) for any non-empty subset $I \subset \mathcal{E}$, so that taking expectation in (2.7) and applying (1.2), we obtain the result.

The integral expression (2.9) is a consequence of (2.8) and the additivity of measure for disjoint sets: it suffices to expand the product and use

$$\int_0^{+\infty} e^{-at} dt = \frac{1}{a}, \qquad \text{for } a > 0.$$

Finally, to compute the second moment in the independent case, we start with the same formula (2.7) to get

$$\mathbf{E}(\tau_{\mathcal{E}}^2) = \sum_{\substack{\emptyset \neq I \subset \mathcal{E} \\ \emptyset \neq J \subset \mathcal{E}}} \sum (-1)^{|I|+|J|} \mathbf{E}(T_I T_J).$$

We first transform this by applying (2.16) in Lemma 2.14 below to compute $\mathbf{E}(T_I T_J)$. This gives

$$\mathbf{E}(\tau_{\mathcal{E}}^2) = \sum_{\substack{\emptyset \neq I \subset \mathcal{E} \\ \emptyset \neq J \subset \mathcal{E}}} \sum \frac{(-1)^{|I|+|J|}}{\mu(E_{I \cup J})} \Big\{ \frac{1}{\mu(E_I)} + \frac{1}{\mu(E_J)} - 1 \Big\}$$

$$(2.12) \qquad = \sum_{\substack{\emptyset \neq I \subset \mathcal{E} \\ \emptyset \neq J \subset \mathcal{E}}} \sum \frac{(-1)^{|I|+|J|}}{\mu(E_{I \cup J})} \Big\{ \frac{2}{\mu(E_I)} - 1 \Big\} \qquad \text{(by symmetry)}.$$

To continue, consider more generally arbitrary complex coefficients $\beta(I)$ defined for $I \subset \mathcal{E}$, and the expression

$$W(\beta) = \sum_{\substack{\emptyset \neq I \subset \mathcal{E} \\ \emptyset \neq J \subset \mathcal{E}}} \sum \frac{(-1)^{|I|+|J|}}{\mu(E_{I \cup J})} \beta(I);$$

note that $\mathbf{E}(\tau_{\mathcal{E}}^2)$ is a simple combination of two such expressions.

We proceed to reduce $W(\beta)$ to a single sum over $I \subset \mathcal{E}$ by rearranging the sum according to the value of $I \cup J$:

$$W(\beta) = \sum_{\emptyset \neq K \subset \mathcal{E}} \frac{1}{\mu(E_K)} \sum_{\substack{\emptyset \neq I, J \subset \mathcal{E} \\ I \cup J = K}} \sum (-1)^{|I|+|J|} \beta(I).$$

The inner sum is rearranged in turn as

$$\sum_{\substack{\emptyset \neq I, J \subset \mathcal{E} \\ I \cup J = K}} \sum (-1)^{|I|+|J|} \beta(I) = \sum_{\emptyset \neq I \subset K} (-1)^{|I|} \beta(I) \sum_{\substack{\emptyset \neq J \subset K \\ I \cup J = K}} (-1)^{|J|}$$
$$= \sum_{\emptyset \neq I \subset K} (-1)^{|I|+|K-I|} \beta(I) \sum_{\substack{I' \subset I \\ I' \cup (K-I) \neq \emptyset}} (-1)^{|I'|}$$

since the subsets J with $I \cup J = K$ are parametrized by $I' \subset I$ using the correspondence $I' \mapsto (K - I) \cup I'$ with inverse $J \mapsto J \cap I$.

For fixed I, the last summation condition $I' \cup (K - I) \neq \emptyset$ is always valid, *unless* $I = K$. In that last case, it only excludes the set $I' = \emptyset$ from all $I' \subset I$. Since we have, for any finite set X, the binomial relation

$$\sum_{Y\subset X}(-1)^{|Y|}=0,$$

it follows that the double sum is simply given by

$$\sum\sum_{\substack{\emptyset\neq I,J\subset\mathcal{E}\\ I\cup J=K}}(-1)^{|I|+|J|}\beta(I)=(-1)^{|K|+1}\beta(K),$$

and hence

$$W(\beta)=\sum_{\emptyset\neq K\subset\mathcal{E}}\frac{(-1)^{|K|+1}\beta(K)}{\mu(E_K)}.$$

Applied to the expression (2.12), this leads precisely to (2.10). □

To deduce Proposition 2.7, we apply this proposition with

$$Z_n=X_n^\sharp,\qquad D=G^\sharp\qquad \mathcal{E}=\{G^\sharp-\mathcal{H}^\sharp\mid \mathcal{H}\in\max(G)\},$$

in the case where the (X_n) are independent uniformly distributed on G, so that the common distribution is $\mu=\nu$. Since, for $I\subset\max G$, we have

$$\nu\Big(\bigcup_{\mathcal{H}\in I}(G^\sharp-\mathcal{H}^\sharp)\Big)=1-\nu\Big(\bigcap_{\mathcal{H}\in I}\mathcal{H}^\sharp\Big),$$

the formulas (2.8) and (2.10) give exactly the claimed formulas (2.3) and (2.4).

Remark 2.10. Note the following strange-looking "linearity" property, which can be checked from our formulas and (2.16): for $G\neq 1$, we have

$$c(G)=\mathbf{E}(\tau_G)=\sum_{\emptyset\neq I\subset\max G}(-1)^{|I|+1}\mathbf{E}(T_I),$$

$$c_2(G)=\mathbf{E}(\tau_G^2)=\sum_{\emptyset\neq I\subset\max G}(-1)^{|I|+1}\mathbf{E}(T_I^2).$$

Remark 2.11. These formulas can only be useful for practical computation if the number of conjugacy classes of maximal subgroups of G is fairly small, or if they are very well understood. As a theoretical instrument, they suffer from the fact that it is very hard to use them to guess or estimate the actual value of $c(G)$. For instance, it is not clear how to recover even the trivial lower bound

$$c(G)\geqslant\delta(G),\tag{2.13}$$

where $\delta(G)$ is the minimal cardinality of a generating set of G. We will give examples later on where this bound is very close to the truth, even with groups of size growing to infinity.

Remark 2.12. Another natural formula is

$$c(G)=1+\sum_{n\geqslant 1}\boldsymbol{P}(X_1^\sharp,\ldots,X_n^\sharp\text{ do not generate }G),\tag{2.14}$$

which is also valid for $G=1$.

Indeed, since τ_G takes positive integer values, we have the familiar formula

$$\mathbf{E}(\tau_G)=\sum_{n\geqslant 1}\boldsymbol{P}(\tau_G\geqslant n),$$

and clearly

$$\{\tau_G \geqslant n\} = \{X_1^\sharp, \ldots, X_{n-1}^\sharp \text{ do not generate } G\},$$

for $n \geqslant 1$. When $n = 1$, this is the certain event, with probability one, thus leading to (2.14).

A moment's thought shows that one can also identify this formula with the one coming from (2.3) by expanding the geometric series:

$$c(G) = \sum_{\substack{I \subset \max(G) \\ I \neq \emptyset}} \frac{(-1)^{|I|+1}}{1 - \nu(\bigcap_{\mathcal{H} \in I} \mathcal{H}^\sharp)} = \sum_{\substack{I \subset \max(G) \\ I \neq \emptyset}} (-1)^{|I|+1} \sum_{n \geqslant 0} \nu(\bigcap_{\mathcal{H} \in I} \mathcal{H}^\sharp)^n$$

$$= 1 + \sum_{n \geqslant 1} \Bigl(\sum_{\substack{I \subset \max(G) \\ I \neq \emptyset}} (-1)^{|I|+1} \nu(\bigcap_{\mathcal{H} \in I} \mathcal{H}^\sharp)^n \Bigr),$$

(where the term with $n = 0$ is only equal to 1 for $\max(G) \neq \emptyset$, i.e., $G \neq 1$).

This is not really a different proof of Proposition 2.7 since the relation

$$\boldsymbol{P}(X_1^\sharp, \ldots, X_n^\sharp \text{ do not generate } G) = \sum_{\substack{I \subset \max(G) \\ I \neq \emptyset}} (-1)^{|I|+1} \nu(\bigcap_{\mathcal{H} \in I} \mathcal{H}^\sharp)^n$$

is proved by inclusion-exclusion, exactly as in the proof of Proposition 2.9.

The point of (2.14) is rather that it leads to another lower bound (for $G \neq 1$):

$$c(G) \geqslant 1 + \sum_{n \leqslant k} \boldsymbol{P}(X_1^\sharp, \ldots, X_n^\sharp \text{ do not generate } G)$$

for any fixed $k \geqslant 1$, and this may be quite useful because there has been a large amount of work on the estimation of those probabilities when k is small, e.g., $k = 2$ if G is not cyclic. For instance, Dixon (for $G = A_n$ and $n \to +\infty$) and Kantor and Lubotzky (for $G = \mathbf{G}(\mathbf{F}_q)$ a simple classical group and $q \to +\infty$) have shown that in those cases we have

$$\boldsymbol{P}(X_1, X_2 \text{ do not generate } G) \to 0$$

as n (resp. q) goes to infinity, indeed with quantitative estimates (see [KL]) – but note the probabilities with elements and with conjugacy classes may behave rather differently (e.g., for $G = PSL(2, \mathbf{F}_p)$, the probability that two random elements generate G is very close to 1 for large p, but there is a probability converging to $1/2$ that two random conjugacy classes do not generate G, see the proof of Theorem 5.1). In a sense, the Chebotarev invariant is thus a refinement of these type of probabilities. We refer to [D2] for a brief survey of probabilistic Galois theory, and to [D1] for the analysis of the case of symmetric groups.

Remark 2.13. As explained in [S1, Th. 5], we have

$$\nu(\mathcal{H}^\sharp) \leqslant 1 - \frac{1}{|G/H|} \tag{2.15}$$

for any conjugacy class of maximal subgroup of G (this is due to Cameron and Cohen).

We now prove the lemma which supplies the formula (2.16) used in the proof of the proposition.

Lemma 2.14. *With notation as in Proposition 2.9, in particular with independent identically distributed random variables* (Z_n)*, for any two non-empty subsets* I*,* J *in* $\mathcal{E}$*, we have*

$$\mathbf{E}(T_I T_J) = \frac{1}{\mu(E_{I\cup J})}\Big(\frac{1}{\mu(E_I)} + \frac{1}{\mu(E_J)} - 1\Big). \tag{2.16}$$

Proof. This is a fairly direct computation, but not very enlightening (at least in our presentation; there might be other approaches that makes this more transparent).

We first compute the joint distribution of T_I and T_J, and for this, we use the shorthand notation

$$p = \mu(E_I), \qquad q = \mu(E_J), \qquad r = \mu(E_{I\cap J}), \qquad s = \mu(E_{I\cup J}),$$
$$p' = \mu(E_I - E_J), \qquad q' = \mu(E_J - E_I).$$

Then we have (generalizing the geometric distribution of a single T_I):

$$\boldsymbol{P}(T_I = n \text{ and } T_J = m) = \begin{cases} (1-s)^{n-1} r, & \text{if } n = m \geqslant 1, \\ (1-s)^{m-1} q'(1-p)^{n-m-1} p, & \text{if } n > m \geqslant 1, \\ (1-s)^{n-1} p'(1-q)^{m-n-1} q, & \text{if } m > n \geqslant 1. \end{cases} \tag{2.17}$$

To justify, e.g., the second of these, note that $T_I = n > m = T_J$ means that Z_k must not be in $E_{I\cup J}$ for $k \leqslant m-1$, Z_m must be in E_J but not in E_I, Z_k must not be in E_I for $m < k < n$, and finally Z_n must be in E_I; then the independence of the (Z_n) gives the formula.

Now we write

$$\mathbf{E}(T_I T_J) = \sum\sum_{n,m\geqslant 1} nm\boldsymbol{P}(T_I = n \text{ and } T_J = m),$$

and we split the sum according to the three cases, say

$$\mathbf{E}(T_I T_J) = Q_1 + Q_2 + Q_3.$$

Introducing further the functions

$$\varphi_i(x) = \sum_{n\geqslant 1} n^i (1-x)^{n-1},$$

we have the expressions

$$Q_1 = \sum_{n\geqslant 1} n^2 \boldsymbol{P}(T_I = T_J = n) = r\sum_{n\geqslant 1} n^2 (1-s)^{n-1} = r\varphi_2(s),$$

and

$$\begin{aligned} Q_2 &= \sum_{1\leqslant m<n} nm\boldsymbol{P}(T_I = n \text{ and } T_J = m) \\ &= q'p \sum_{1\leqslant m<n} nm(1-s)^{m-1}(1-p)^{n-m-1} \\ &= q'p\sum_{m\geqslant 1} m(1-s)^{m-1}\sum_{k\geqslant 1}(m+k)(1-p)^{k-1} \\ &= q'p(\varphi_0(p)\varphi_2(s) + \varphi_1(p)\varphi_1(s)), \end{aligned}$$

while Q_3 is given by the same expression after exchanging p and q, p' and q'.

Since, by Taylor expansion, we have

$$\varphi_0(x) = \frac{1}{x}, \quad \varphi_1(x) = \frac{1}{x^2}, \quad \varphi_2(x) = \frac{2-x}{x^3} = \frac{2}{x^3} - \frac{1}{x^2},$$

we obtain

$$\mathbf{E}(T_I T_J) = \frac{1}{s}\Big(\frac{2}{s} + \frac{q'}{ps} + \frac{p'}{qs} - 1\Big),$$

by adding the three terms. Finally, the relations

$$q' + p = s, \qquad p' + q = s,$$

lead to the simplified expression

$$\begin{aligned}\mathbf{E}(T_I T_J) &= \frac{1}{s}\Big(\frac{2}{s} + \frac{q'}{ps} + \frac{p'}{qs} - 1\Big) = \frac{1}{s}\Big(\frac{2}{s} + \frac{s-p}{ps} + \frac{s-q}{qs} - 1\Big)\\ &= \frac{1}{s}\Big(\frac{2}{s} + \frac{1}{p} - \frac{1}{s} + \frac{1}{q} - \frac{1}{s} - 1\Big),\end{aligned}$$

which gives (2.16). □

We now present some easy formal properties of the Chebotarev invariants (attached, unless stated otherwise, with a sequence of independent uniformly distributed random variables).

The first lemma may be used to simplify and expand the range of groups covered by certain computations (this is also observed by Pomerance [P], for the case of numbers of generators instead of conjugacy classes):

Lemma 2.15. *Let G be a finite group, and let $\Phi(G)$ be the* Frattini subgroup *of G, i.e., the intersection of all maximal subgroups of G. Then, for any normal subgroup $N \triangleleft G$ such that $N \subset \Phi(G)$, in particular for $N = \Phi(G)$, we have*

$$c(G) = c(G/N), \qquad c_2(G) = c_2(G/N).$$

Proof. Let $H = G/N$. Moreover, we have $\Phi(H) = \Phi(G)/N$ and hence $H/\Phi(H) \simeq G/\Phi(G)$. This means that we need only prove the result when $N = \Phi(G)$, the general case following by applying this to H.

Let $\pi : G \to G/\Phi(G)$ be the quotient map. If (X_n) is a sequence of independent random variables uniformly distributed on G, the $Y_n = \pi(X_n)$ are independent and uniformly distributed on $G/\Phi(G)$. Moreover, for any $n \geqslant 1$, the elements $(X_1^\sharp, \ldots, X_n^\sharp)$ generate G if and only if the elements $(Y_1^\sharp, \ldots, Y_n^\sharp)$ generate $G/\Phi(G)$: indeed, this follows from the basic fact that a subset $S \subset G$ generates G if and only if $\pi(S)$ generates $G/\Phi(G)$ (this is applied to all sets $S = \{x_1, \ldots, x_n\}$ where x_i conjugate to X_i). This gives the result immediately from the definition of the waiting times. □

We next consider products:

Proposition 2.16. *Let G_1, G_2 be finite groups such that the only subgroup $H \subset G_1 \times G_2$ which surjects by projection to both factors is $H = G$. Then we have*

$$c(G_1 \times G_2) \leqslant c(G_1) + c(G_2) - 1.$$

Examples of groups G_1, G_2 satisfying the hypothesis are any pair of non-isomorphic simple groups; note that this proposition suggests that sometimes $c'(G) = c(G)-1$ would be a more natural invariant to consider, since we then have the simpler inequality

$$c'(G_1 \times G_2) \leqslant c'(G_1) + c'(G_2).$$

Proof. With $G = G_1 \times G_2$ and denoting $X_n = (Y_n, Z_n) \in G_1 \times G_2$ a sequence of independent uniformly distributed random variables, it is clear that (Y_n), (Z_n) are similarly independent uniformly distributed on G_1 and G_2 respectively. We then have the inequality

$$\tau_G \leqslant \max(\tau_1, \tau_2) \leqslant \tau_1 + \tau_2 - 1$$

(since $\tau_i \geqslant 1$ and $\max(m, n) \leqslant n + m - 1$ for integers n, $m \geqslant 1$), with

$$\tau_1 = \min\{n \geqslant 1 : (Y_1^\sharp, \ldots, Y_n^\sharp) \text{ generate } G_1\}, \quad \tau_2 = \min\{n \geqslant 1 : (Z_1^\sharp, \ldots, Z_n^\sharp) \text{ generate } G_2\}$$

which are distributed like τ_{G_1}, τ_{G_2} (indeed, if $n \geqslant \max(\tau_1, \tau_2)$, then the group generated by any elements in $X_n^\sharp = (Y_n^\sharp, Z_n^\sharp)$ surjects to G_1 and G_2, hence it must be equal to G by assumption). Taking expectation, we get the inequality stated. □

The next result gives upper and lower estimates for the Chebotarev invariant using smaller sets of maximal subgroups than $\max(G)$. This can be very useful in particular for the asymptotic study of $c(G_n)$ for a sequence of finite groups (G_n), as we will see later on.

Proposition 2.17. *Let G be a finite group, and let $M \subset \max(G)$ be an arbitrary non-empty finite subset of maximal subgroups. Let*

$$\tilde{\tau}_M = \max_{\mathcal{H} \in M} \hat{\tau}_{\mathcal{H}}.$$

with notation as in (2.5) *and*

$$p_M = \nu\Big(G^\sharp - \bigcup_{\mathcal{H} \in \max(G) - M} \mathcal{H}^\sharp\Big). \tag{2.18}$$

We then have

$$\mathbf{E}(\tilde{\tau}_M) = \sum_{\emptyset \neq I \subset M} \frac{(-1)^{|I|+1}}{1 - \nu(\bigcap_{\mathcal{H} \in I} \mathcal{H}^\sharp)} \leqslant c(G) \leqslant \mathbf{E}(\tilde{\tau}_M) - 1 + p_M^{-1}$$

and

$$\mathbf{E}(\tilde{\tau}_M^2) \leqslant c_2(G) \leqslant \mathbf{E}(\tilde{\tau}_M^2) + \frac{2 - p_M}{p_M^2} - 1.$$

Proof. Define the additional waiting time

$$\tau^* = \min\{n \geqslant 1 \mid X_n \notin \bigcup_{\mathcal{H} \notin M} \mathcal{H}^\sharp\}.$$

We then note the inequalities

$$\tilde{\tau}_M \leqslant \tau_G \leqslant \max(\tilde{\tau}_M, \tau^*) \leqslant \tilde{\tau}_M + \tau^* - 1,$$

where the first inequality is obvious, while the second follows because, for $n = \max(\tilde{\tau}_M, \tau^*)$, we know that the group generated by $(X_1^\sharp, \ldots, X_n^\sharp)$ is not contained in any subgroup in a conjugacy class of maximal subgroups $\mathcal{H} \in M$, and that this group also contains one element which is not conjugate to any element in a subgroup not in M.

Now we take expectation on both sides; observing that, by independence, τ^* is distributed like a geometric random variable with parameter p_m given by (2.18), we obtain the first inequalities, using Proposition 2.9 and (1.2).

Similarly, for the secondary invariant, we use the inequalities

$$\tilde{\tau}_M^2 \leqslant \tau_G^2 \leqslant \max(\tilde{\tau}_M, \tau^*)^2 \leqslant \tilde{\tau}_M^2 + (\tau^*)^2 - 1,$$

and get

$$\mathbf{E}(\hat{\tau}_M^2) \leqslant c_2(G) \leqslant \mathbf{E}(\hat{\tau}_M^2) + \mathbf{E}((\tau^*)^2) - 1 = \mathbf{E}(\hat{\tau}_M^2) + \frac{2 - p_M}{p_M^2} - 1.$$

□

The following immediately follows:

Corollary 2.18. *Let (G_n) be a sequence of non-trivial finite groups, and let ν_n denote the corresponding density. For each $n \geqslant 1$, let M_n be a non-empty subset of $\max(G_n)$, and assume that*

$$\lim_{n \to +\infty} \nu_n\Bigl(\bigcup_{\mathcal{H} \in \max(G_n) - M_n} \mathcal{H}^\sharp \Bigr) = 0, \tag{2.19}$$

i.e., the proportion of elements represented by a conjugacy class in some subgroup in M_n goes to zero. Then we have

$$c(G_n) = \mathbf{E}(\tilde{\tau}_{M_n}) + o(1), \qquad \text{and} \qquad c_2(G_n) = \mathbf{E}(\tilde{\tau}_{M_n}^2) + o(1),$$

as $n \to +\infty$, with notation as in Proposition 2.17.

The following sections will now take up the problem of computing, or estimating, the Chebotarev invariants for various classes of groups.

3. Abelian and nilpotent groups

In this section, we look at finite *abelian* and nilpotent groups G. In fact, because nilpotent groups have the (characteristic) property that $[G, G] \subset \Phi(G)$ (see, e.g., [Ro, Th. 11.3, (v)]), Lemma 2.15 shows that if G is a nilpotent group, we have

$$c(G) = c(G/[G, G]), \qquad c_2(G) = c_2(G/[G, G])$$

which are Chebotarev and secondary Chebotarev invariants of abelian groups. This applies, in particular, to all p-groups.

We will not use the formula from Proposition 2.7 (although it is possible, as was done in a first draft, to do some computations using it), because in abelian groups there tends to be many maximal subgroups up to conjugacy – since conjugacy is now trivial. Following the work of Pomerance [P], who computed $c(G)$ (with different terminology) for any abelian group G, we will use another description of the Chebotarev waiting time in the case of abelian groups.

Theorem 3.1 (Pomerance)**.** *Let G be a finite abelian group, and for any prime number $p \mid |G|$, let $r_p(G) = \dim_{\mathbf{F}_p}(G/pG)$ be the p-rank of G. Let $\delta(G) = \max r_p(G)$ be the minimal cardinality of a generating set of G. Then we have*

$$c(G) = \delta(G) + \sum_{j \geqslant 1} \Bigl(1 - \prod_{p \mid |G|} \prod_{1 \leqslant i \leqslant r_p(G)} (1 - p^{-(\delta(G)+j-i)})\Bigr).$$

In particular, for $G = \mathbf{Z}/n\mathbf{Z}$ with $n \geqslant 2$, we have

$$c(G) = -\sum_{\substack{d \mid n \\ d \neq 1}} \frac{\mu(d)}{1 - d^{-1}} \tag{3.1}$$

and for $G = \mathbf{F}_p^k$, where $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$, with p prime and $k \geqslant 1$, we have

$$c(G) = k + \sum_{1 \leqslant j \leqslant k} \frac{1}{p^j - 1}. \tag{3.2}$$

This is [P, Theorem] and immediate corollaries of it. The formula for $G = \mathbf{Z}/n\mathbf{Z}$ might be easier to get directly from Proposition 2.7. Indeed, the subgroups of $\mathbf{Z}/n\mathbf{Z}$ are the groups $H_d = d\mathbf{Z}/n\mathbf{Z}$ for $d \mid n$, with $\nu(H_d) = d^{-1}$ and $H_d \cap H_e = H_{[d,e]}$, and the maximal subgroups among these correspond to minimal divisors of n for divisibility, i.e., to the primes p dividing n. Then a non-empty subset I of $\max(G)$ can be parametrized by the corresponding subset of prime divisors of n, or equivalently by the squarefree divisor $d > 1$ of n which is the product of those primes. In this correspondence, we have

$$\bigcap_{H \in I} H = \bigcap_{p \mid d} H_p = H_d, \qquad \text{hence} \qquad \nu\Big(\bigcap_{H \in I} H\Big) = \nu(H_d) = \frac{1}{d},$$

and $(-1)^{|I|} = \mu(d)$, hence (2.3) gives the stated formula for $c(\mathbf{Z}/n\mathbf{Z})$.

We have similar results for the secondary Chebotarev invariant; Pomerance mentions the possibility of computing these, but does not give any results in his paper.

Theorem 3.2. *Let G be a finite abelian group. With notation as in Theorem 3.1, we have*

$$c_2(G) = \delta(G)^2 + \sum_{j \geqslant 1} (2j + 2\delta(G) - 1)\Big(1 - \prod_{p \mid |G|} \prod_{1 \leqslant i \leqslant r_p(G)} (1 - p^{-(\delta(G)+j-i)})\Big).$$

In particular, we have

$$c_2(\mathbf{Z}/n\mathbf{Z}) = -\sum_{2 \leqslant d \mid n} \mu(d) \frac{1 + d^{-1}}{(1 - d^{-1})^2}$$

for $n \geqslant 1$ and

$$c_2(\mathbf{F}_p^k) = c(\mathbf{F}_p^k)^2 + \sum_{1 \leqslant j \leqslant k} \frac{p^j}{(p^j - 1)^2},$$

for p prime and $k \geqslant 1$.

Proof. The first result is obtained by reasoning as in [P, p. 195], with r and $(r + j)$ there replaced by r^2 and $(r + j)^2$. The point is that he shows that

$$\boldsymbol{P}((X_1, \ldots, X_{\delta(G)+j}) \text{ generate } G) = \prod_{p \mid |G|} \prod_{1 \leqslant i \leqslant r_p(G)} (1 - p^{-(\delta(G)-r_p(G)+j+i)}).$$

To deduce the values for $G = \mathbf{Z}/p^k\mathbf{Z}$, it is simpler to use the description[6]

$$\tau_G = \sum_{j=1}^{k} G_j$$

where the G_j are independent geometric random variables with parameters $p_j = 1 - p^{-j}$. Concretely, they can be defined as follows

$$G_k = \min\{n \geqslant 1 \mid X_n \neq 0\},$$
$$G_{k-1} = \min\{n \geqslant 1 \mid \dim_{\mathbf{F}_p}\langle X_{G_k+n}, X_{G_k}\rangle = 2\}, \quad \ldots$$
$$G_1 = \min\{n \geqslant 1 \mid \dim_{\mathbf{F}_p}\langle X_{G_2+n}, X_{G_2}, \ldots, X_{G_k}\rangle = k\},$$

which, by independence of the (X_n), are easily checked to be indeed independent geometric variables with the stated parameters.

This decomposition leads to the formula for $c_2(G)$ immediately, using (1.2) and additivity of the variance of independent random variables. ∎

The formula of Pomerance gives a quick way to understand the limit values of Chebotarev invariants for abelian groups with a given rank $\delta(G)$.

Corollary 3.3 (Pomerance)**.** *For any fixed integer $k \geqslant 1$, and any abelian finite group G with $\delta(G) = k$, we have*

$$k \leqslant c(G) \leqslant \limsup_{\substack{|G|\to+\infty \\ \delta(G)=k}} c(G) = k + 1 + \sum_{j\geqslant 1}\Bigl(1 - \prod_{1\leqslant j\leqslant k}\zeta(j+k)^{-1}\Bigr).$$

In particular, the Chebotarev invariants for cyclic groups are bounded.

Corollary 3.4. *For any fixed k, we have*

$$c(\mathbf{F}_p^k) = k + O(p^{-1}), \qquad c_2(\mathbf{F}_p^k) = k^2 + O(p^{-1}),$$

and

$$\boldsymbol{P}(\tau_{\mathbf{F}_p^k} \neq k) \ll p^{-1},$$

where the implied constants depend only on k.

This last result shows that, for vector spaces over a finite field, the Chebotarev invariant is strongly peaked around the average, which is itself close to the dimension.

Proof. Only the last inequality needs (maybe) a bit of explanation. Since $\tau_{\mathbf{F}_p^k}$ takes positive integer values $\geqslant k$, we have

$$|\tau_{\mathbf{F}_p^k} - k| \geqslant 1$$

if $\tau_{\mathbf{F}_p^k} \neq k$. Hence, if $\tau_{\mathbf{F}_p^k} \neq k$, we have

$$|\tau_{\mathbf{F}_p^k} - c(\mathbf{F}_p^k)| \geqslant |\tau_{\mathbf{F}_p^k} - k| - |c(\mathbf{F}_p^k) - k| \geqslant 1 - |c(\mathbf{F}_p^k) - k|,$$

and if we furthermore we have $p \geqslant p_0$, where p_0 (depending on k) is chosen so that

$$k \leqslant c(\mathbf{F}_p^k) \leqslant k + 1/2$$

[6] Which is the analogue of the decomposition of the waiting time for the standard Coupon Collecting Problem in a sum of geometric random variables.

for all $p \geqslant p_0$, it follows that

$$\{\tau_{\mathbf{F}_p^k} \neq k\} \subset \{|\tau_{\mathbf{F}_p^k} - c(\mathbf{F}_p^k)| \geqslant 1/2\}$$

for such p, and then the Chebychev inequality gives

$$\boldsymbol{P}(\tau_{\mathbf{F}_p^k} \neq k) \leqslant 4\mathbf{V}(\tau_{\mathbf{F}_p^k}) \ll p^{-1}$$

for $p \geqslant p_0$, where the implied constant depends on k. Increasing this constant if needed (e.g., taking it to be at least p_0), we can also claim that this inequality holds for $p \geqslant 2$. □

Remark 3.5. In particular, for cyclic groups, the Chebotarev invariant is at most, and its limsup is, the constant

$$2 + \sum_{k \geqslant 2} \Big(1 - \frac{1}{\zeta(k)}\Big) = 2.705211140105367764\ldots$$

This asymptotic behavior is not without interest (and some surprise): on the one hand, we see that $c(\mathbf{Z}/n\mathbf{Z})$ remains absolutely bounded, despite the existence of cyclic groups with many subgroups, and on the other hand, we see that it is not always close to the minimal number of generators.

Concerning the first point, notice for instance that a "naive" invariant is given by

$$\sum_{H \in \max(G)} \nu(H) = \sum_{p \mid n} \frac{1}{p},$$

and this is unbounded as n grows for $G = \mathbf{Z}/n\mathbf{Z}$ (though it is $\ll \log\log\log n$, and thus bounded in practice...), see the discussion surrounding (9.2) below for occurrences of such quantities instead of the Chebotarev invariant.

For the second, note that (interpreting $1/\zeta(1) = 0$) the limsup we found is also N where

$$N = \sum_{k \geqslant 1} (1 - \zeta(k)^{-1})$$

is sometimes called the *Niven constant*. Niven obtained it as the mean-value of the maximal exponent of a prime dividing a positive integer:

$$N = \lim_{n \to +\infty} \frac{1}{n-1} \sum_{2 \leqslant j \leqslant n} \alpha(j)$$

with

$$\alpha(j) = \max\{\nu \geqslant 0 \mid p^\nu \mid j \text{ for some prime } p\}$$

for $j \geqslant 2$ (see [N]). The explanation for this coincidence is that $\zeta(k)^{-1}$, for $k \geqslant 2$, is both the (asymptotic) density of primitive vectors in $\mathbf{Z}^k$ and that of k-power-free integers.

Remark 3.6. If one uses Proposition 2.7 instead, one can prove (after some computation) the following formulas

$$c(\mathbf{F}_p^k) = \sum_{1 \leqslant j \leqslant k} \frac{(-1)^{j+1}}{1 - p^{-j}} \binom{k}{j}_p p^{j(j-1)/2},$$

$$c_2(\mathbf{F}_p^k) = \sum_{1 \leqslant j \leqslant k} (-1)^j \frac{1 + p^{-j}}{(1 - p^{-j})^2} \binom{k}{j}_p p^{j(j-1)/2},$$

where

$$\binom{k}{j}_p = \frac{(1-p^k)\cdots(1-p^{k-j+1})}{(1-p^j)\cdots(1-p)}$$

are the p-binomial coefficients. Note that those formulas do not immediately reveal the limiting behavior as $p \to +\infty$, since the summands have different degrees as rational functions of p.

4. A solvable example

The results of the previous section, as well as those we will see in the next one, reveal (or suggest) rather small values of the Chebotarev invariants, in comparison with the size of the groups. The following example in the solvable case exhibits very different behavior (possibly the worse possible).

Proposition 4.1. *For q a power of a prime, let*

$$H_q = \left\{ \begin{pmatrix} a & t \\ 0 & 1 \end{pmatrix} \mid a \in \mathbf{F}_q^\times, \quad t \in \mathbf{F}_q \right\}$$

be the group of translations and dilations of the affine plane $\mathbf{F}_q^2$ of order $q(q-1)$, isomorphic to a semi-direct product $\mathbf{F}_q \rtimes \mathbf{F}_q^\times$.

(1) *We have*

$$c(H_q) = q - q^{-1} \sum_{1 \neq d \mid q-1} \frac{\mu(d)}{(1-d^{-1})(1-d^{-1}+q^{-1})} \tag{4.1}$$

$$c_2(H_q) = q(2q-1) + c_2(\mathbf{Z}/(q-1)\mathbf{Z}) + \sum_{1 \neq d \mid q-1} \mu(d) \frac{1+d^{-1}-q^{-1}}{(1-d^{-1}+q^{-1})^2}. \tag{4.2}$$

(2) *For $q \geqslant 2$, we have*

$$c(H_q) = q + O(\tau(q-1)), \qquad c_2(H_q) = q(2q-1) + O(\tau(q-1)), \tag{4.3}$$

where $\tau(q-1)$ is the number of positive divisors of $q-1$. In particular, $c(H_q) \sim q$ as $q \to +\infty$.

Since we have a split exact sequence

$$1 \to \mathbf{F}_q \to H_q \xrightarrow{\det} \mathbf{F}_q^\times \to 1$$

and the two surrounding groups are isomorphic to $\mathbf{F}_p^k$, where $q = p^k$ with p prime, and to a cyclic group $\mathbf{Z}/(q-1)\mathbf{Z}$, with Chebotarev invariants tending to k as p gets large, and bounded, respectively, this shows in particular that the Chebotarev invariant can jump quite uncontrollably under extensions.

The proof will use Proposition 2.7. We start with a lemma that is certainly well-known, but for which we give a proof for completeness and lack of a suitable reference.

Lemma 4.2. (1) *The conjugcay classes of maximal subgroups of H_q have representatives given by*

$$A = \left\{ \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \mid a \in \mathbf{F}_q^\times \right\},$$

and

$$C_\ell = \left\{ \begin{pmatrix} a & t \\ 0 & 1 \end{pmatrix} \in H_q \ \mid\ a \in (\mathbf{F}_q^\times)^\ell \ \textit{and}\ t \in \mathbf{F}_q \right\},$$

where ℓ runs over the prime divisors of $q-1$.

(2) *There are q conjugacy classes in H_q; they are given, with representatives of them, by*

$$g_b = \begin{pmatrix} b & 0 \\ 0 & 1 \end{pmatrix}, \quad g_b^\sharp = \{g \in H_q \mid \det(g) = b\}, \quad |g_b^\sharp| = q,$$

where $b \in \mathbf{F}_q^\times - \{1\}$, and

$$\mathrm{Id} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathrm{Id}^\sharp = \{\mathrm{Id}\},$$

$$u = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad u^\sharp = \{g \in H_q - \{\mathrm{Id}\} \mid \det(g) = 1\}, \quad |u^\sharp| = q-1.$$

Proof. (1) Denote

$$U = H_q \cap SL(2, \mathbf{F}_q) = \left\{ \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \mid t \in \mathbf{F}_q \right\}.$$

Let $H \subset H_q$ be a maximal subgroup. Let $D = \det(H)$ be the image of the determinant restricted to H. Since

$$H \subset \det{}^{-1}(D),$$

we have either $\det^{-1}(D) = H_q$, i.e., $D = \mathbf{F}_q^\times$, or $H = \det^{-1}(D)$. In this second case, the subgroup D must be a maximal subgroup of $\mathbf{F}_q^\times$ for H to be maximal, which implies that H is of the form C_ℓ for some ℓ. Conversely, such a subgroup is maximal because if we add any extra element g and let $H' = \langle C_\ell, g\rangle$, the fact that $U \subset C_\ell$ implies that some

$$\begin{pmatrix} a' & 0 \\ 0 & 1 \end{pmatrix},$$

with $a' \notin (\mathbf{F}_q^\times)^\ell$, is in H', and then by maximality in $\mathbf{F}_q^\times$, we have $H' = H_q$.

Note that the C_ℓ are normal in H_q, hence also pairwise non-conjugate.

In the first case, when $D = \mathbf{F}_q^\times$, i.e., when the determinant restricted to H is surjective, we claim that the determinant is also injective on H: indeed, otherwise, there exists $u \in U \cap H$, $u \neq 1$, say

$$u = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \quad \text{with} \quad t \neq 0.$$

For any $a \in \mathbf{F}_q^*$, by surjectivity there exists $\alpha(a) \in \mathbf{F}_q$ with

$$\begin{pmatrix} a & \alpha(a) \\ 0 & 1 \end{pmatrix}$$

and by applying the relation

$$\begin{pmatrix} a & \alpha(a) \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & \alpha(a) \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & at \\ 0 & 1 \end{pmatrix}$$

for all $a \in \mathbf{F}_q^\times$, we conclude that in fact $U \subset H$. Then $|H|$ is divisible both by q and by $q-1$, hence $H = H_q$, contradicting the assumption that H is a proper subgroup of H_q. So, in this second case, the determinant gives an isomorphism $H \simeq \mathbf{F}_q^\times$. Then a generator of H

must be diagonalizable (it has distinct eigenvalues in $\mathbf{F}_q^\times$), and therefore H is conjugate to A.

(2) We have the general conjugation formula

$$\begin{pmatrix} a & t \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} b & v \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & t \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} b & av + t(1-b) \\ 0 & 1 \end{pmatrix}$$

from which it immediately follows that all elements with $\det(g) = b \neq 1$ are conjugate, and gives therefore the $q-2$ conjugacy classes with representatives g_b as described. For $b = 1$, it is clear that all elements with $b = 1$, $v \neq 0$ form the conjugacy class $u^\sharp$, and only the identity class remains. □

Proof of Proposition 4.1. First of all, in addition to the maximal subgroups C_ℓ given by Lemma 4.2, there are subgroups C_d for all squarefree divisors $d \mid q-1$, the inverse image under the determinant of the subgroup of order $(q-1)/d$ in the cyclic group $\mathbf{F}_q^\times$.

Given a subset $I \subset \max(H_q)$, we now compute the density of conjugacy classes in

$$\mathcal{H}_I^\sharp = \bigcap_{\mathcal{H} \in I} \mathcal{H}^\sharp,$$

as follows:

– If $A \in I$, then with $I' = I - \{A\}$, and d the product of those primes ℓ for which $C_\ell \in I'$ (including $d = 1$ when $I' = \emptyset$), we have

$$\nu(\mathcal{H}_I^\sharp) = \frac{1}{d} - q^{-1} \quad \text{and in particular } \nu(A^\sharp) = 1 - q^{-1}.$$

Indeed, we have to find the density of those elements of H_q which are diagonalizable with eigenvalues 1 and $a \in C_d$. These are exactly the conjugacy classes $g_b^\sharp$ with $b \in C_d - \{1\}$, and the trivial class, so

$$\nu(\mathcal{H}_I^\sharp) = \frac{1 + ((q-1)/d - 1)q}{q(q-1)} = \frac{q(q-1)/d - (q-1)}{q(q-1)} = \frac{1}{d} - \frac{1}{q}.$$

– If $A \notin I$, then I corresponds to a divisor $d \mid q-1$, $d \neq 1$, and we have

$$\nu(\mathcal{H}_I^\sharp) = \frac{1}{d},$$

since we must now compute the density of elements of H_q which have $\det(g) \in C_d$, and this is

$$\frac{q((q-1)/d - 1) + 1 + q - 1}{q(q-1)} = \frac{1}{d}.$$

Applying (2.3), and isolating the contribution of $I = \{A\}$, leads exactly to (4.1) and to (4.2).

To deduce (4.3) for $c(H_q)$, we may assume $q = p^k$ with p an odd prime, since for q even, we have

$$c(H_q) = q + c(\mathbf{Z}/(q-1)\mathbf{Z}) = q + O(1)$$

by Corollary 3.4. So for q odd, we write

$$c(H_q) = q + c(\mathbf{Z}/(q-1)\mathbf{Z}) - \Delta(q) = q - \Delta(q) + O(1)$$

where

$$\Delta(q) = \sum_{1\neq d\mid q-1} \frac{\mu(d)}{1-d^{-1}+q^{-1}}.$$

Since $1-d^{-1}+q^{-1} \geqslant 1-d^{-1} > 0$, we can bound this from above by

$$|\Delta(q)| \leqslant \sideset{}{^\flat}\sum_{1\neq d\mid q-1} \frac{1}{1-d^{-1}},$$

and then proceeding as in the proof of Corollary 3.4, we obtain

$$\begin{aligned}
|\Delta(q)| &\leqslant \sum_{k\geqslant 0}\Bigl(\prod_{p\mid q-1}(1+p^{-k})-1\Bigr)\\
&\leqslant \tau(q-1) + \frac{\psi(q-1)}{q-1} - 2 + \sum_{k\geqslant 2}\Bigl(\frac{\zeta(k)}{\zeta(2k)}-1\Bigr)\\
&= O(\tau(q-1))
\end{aligned}$$

since the series converges absolutely again.

Finally, the asymptotics for $c_2(H_q)$ are obtained by essentially identical arguments. ☐

The proof confirms the intuitive fact that the large size of $c(H_q)$ is due directly to the existence of a fairly small diagonal subgroup A (of index q) that contains elements conjugate to a very large proportion of elements of H_q. So the waiting time is quite close to the waiting time until a non-diagonalizable element is obtained, which is a geometric random variable T with

$$\boldsymbol{P}(T=k) = \frac{1}{q}\Bigl(1-\frac{1}{q}\Bigr)^{k-1}, \quad \text{for } k\geqslant 1$$

(since very often, it will be the case that sufficiently many diagonalizable elements will have appeared by the time an element of U appears to generate the whole group).

This is confirmed by the large second moment $c_2(H_q)$: it corresponds to a standard deviation of the waiting time which is

$$\sqrt{c_2(H_q)-c(H_q)^2} \;\sim\; q, \quad \text{as } q\to+\infty,$$

i.e., very close to the expectation, similar to the fact that

$$\mathbf{V}(T) = q\sqrt{1-\frac{1}{q}}.$$

The groups $G=H_q$ also show that the inequality (2.15) is best possible (with the maximal subgroup $H=A$), as observed also in [S1], so it is not surprising that they lead to high Chebotarev invariants.

5. Some finite groups of Lie type

For specific complicated non-abelian groups, the Chebotarev invariant may be hard to compute exactly, except numerically using the formulas of Proposition 2.7, when feasible (we will give examples from computer calculations in Section 7). However, if we consider infinite families of non-abelian groups, it may be that the subgroup structure is sufficiently well-known, simple and regular, that one can derive asymptotic information. In fact, using results like Proposition 2.17, it is not needed for this purpose to have complete control over *all* maximal subgroups.

We illustrate this first simplest family of simple groups of Lie type.

Theorem 5.1. (1) *We have*

$$c(PSL(2,\mathbf{F}_p)) = 3 + O(p^{-1}), \qquad c_2(PSL(2,\mathbf{F}_p)) = 11 + O(p^{-1}),$$

for primes $p \geqslant 2$.

(2) *For all* $k \geqslant 2$, *we have*

$$\boldsymbol{P}(\tau_{PSL(2,\mathbf{F}_p)} = k) = \frac{1}{2^{k-1}} + O(p^{-1})$$

where the implied constant depends on k.

(3) *The same results hold for* $SL(2,\mathbf{F}_p)$, *and in fact*

$$c(SL(2,\mathbf{F}_p)) = c(PSL(2,\mathbf{F}_p)), \qquad c_2(SL(2,\mathbf{F}_p)) = c_2(PSL(2,\mathbf{F}_p)) \tag{5.1}$$

for all p.

Note that the limit of $c(SL(2,\mathbf{F}_p))$ is not the minimal number of generators of $SL(2,\mathbf{F}_p)$ (which is 2, since $SL(2,\mathbf{F}_p)$ is generated by the two elementary matrices with 1 over and under the main diagonal.)

For the proof, we will not use the formula of Proposition 2.7, although this could be done at least to prove (1) (since the subgroups of $PSL(2,\mathbf{F}_p)$ are well understood since Dickson, see, e.g., [Gi, Th. 2.2]). Instead, we use the following criterion of Serre [S3, Prop. 19] (which is itself based on knowing the subgroup structure).

Lemma 5.2 (Serre)**.** *Let* $p \geqslant 5$ *be a prime number. Assume that* $G \subset SL(2,\mathbf{F}_p)$ *is a subgroup such that*

(1) *The group* G *contains an element* s *such that* $\mathrm{Tr}(s)^2 - 4$ *is a non-zero square in* $\mathbf{F}_p$, *and such that* $\mathrm{Tr}(s) \neq 0$;

(2) *The group* G *contains an element* s *such that* $\mathrm{Tr}(s)^2 - 4$ *is not a square in* $\mathbf{F}_p$, *and such that* $\mathrm{Tr}(s) \neq 0$;

(3) *The group* G *contains an element* s *such that* $\mathrm{Tr}(s)^2 \in \mathbf{F}_p$ *is not in* $\{0,1,2,4\}$, *and is not a root of* $X^2 - 3X + 1$.

Then we have $G = SL(2,\mathbf{F}_p)$.

Proof of Theorem 5.1. We first notice that we need only check (5.1) and then consider the case of $SL(2,\mathbf{F}_p)$. These equalities are consequences of Lemma 2.15, since[7] $\{\pm I\} \subseteq \Phi_p$, where Φ_p is the Frattini subgroup of $SL(2,\mathbf{F}_p)$. Indeed, we may of course assume that $p \neq 2$; then, if p is such that $-I \notin \Phi_p$, there exists a maximal subgroup H of $SL(2,\mathbf{F}_p)$ which surjects to $PSL(2,\mathbf{F}_p)$. We would then have

$$SL(2,\mathbf{F}_p) = \{\pm I\} \times H$$

which is impossible, since $SL(2,\mathbf{F}_p)$ is generated by the elements

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

both of which are of odd order p, hence contained in H (compare with [S2, IV-23]).

[7] In fact, it is known that there is equality, but we do not need this stronger fact.

Now we consider $SL(2,\mathbf{F}_p)$, and we assume $p \geqslant 5$. Let $\tau = \tau_{SL(2,\mathbf{F}_p)}$ denote the corresponding waiting time, and let τ_1, τ_2, τ_3 denote the waiting times for conjugacy classes satisfying the conditions (1), (2) and (3) in Lemma 5.2, i.e., for instance

$$\tau_1 = \min\{n \geqslant 1 \,:\, s = X_n^\sharp \text{ has } \mathrm{Tr}(s) \neq 0 \text{ and } \mathrm{Tr}(s)^2 - 4 \text{ is in } (\mathbf{F}_p^\times)^2\}.$$

Let also τ_1^*, τ_2^* be the waiting times for conditions (1) and (2) *without* the condition $\mathrm{Tr}(s) \neq 0$. Note that (1) and (2) are exclusive conditions. Moreover, each τ_i is a geometric random variable, with parameters, respectively

$$p_1 = \frac{1}{2} + O(p^{-1}), \quad p_2 = \frac{1}{2} + O(p^{-1}), \quad p_3 = 1 + O(p^{-1}), \tag{5.2}$$

and for τ_1^*, τ_2^*, the parameters are also

$$p_1^* = \frac{1}{2} + O(p^{-1}), \quad p_2^* = \frac{1}{2} + O(p^{-1}) \ ;$$

all these facts can be checked easily, e.g., by looking at tables of conjugacy classes in $SL(2,\mathbf{F}_p)$ (for instance, [FH, p. 71]).

We then have

$$\max(\tau_1^*, \tau_2^*) \leqslant \tau_p \leqslant \max(\tau_1, \tau_2, \tau_3),$$

where the right-hand inequality comes from Lemma 5.2 and the left-hand inequality is due to the fact that the Borel subgroup

$$B = \left\{ \begin{pmatrix} x & a \\ 0 & x^{-1} \end{pmatrix} \right\} \subset SL(2,\mathbf{F}_p)$$

intersects every conjugacy class satisfying (1) (so that $\tau_p \geqslant \tau_2^*$) and the non-split Cartan subgroup

$$C_{ns} = \left\{ \begin{pmatrix} a & b \\ \varepsilon b & a \end{pmatrix} \right\} \subset SL(2,\mathbf{F}_p)$$

intersects every conjugacy class satisfying (2), where $\varepsilon \in \mathbf{F}_p^\times$ is a fixed non-square element (so that $\tau_p \geqslant \tau_1^*$).

By applying Proposition 2.7 to compute the expectation and second moment on the two extreme sides, we find

$$3 + O(p^{-1}) \leqslant \mathbf{E}(\tau_p) \leqslant 3 + O(p^{-1}), \qquad 11 + O(p^{-1}) \leqslant \mathbf{E}(\tau_p^2) \leqslant 11 + O(p^{-1}).$$

which proves (1).

To prove (2), fix some $k \geqslant 2$. We denote

$$\tau_p^* = \max(\tau_1^*, \tau_2^*), \qquad \tau_p' = \max(\tau_1, \tau_2, \tau_3),$$

We have the equality of events

$$\{\tau_p = k\} = \{\tau_p = \tau_p' = k\} \cup \{\tau_p = k < \tau_p'\},$$

which is of course a disjoint union. Then we note that

$$\boldsymbol{P}(\tau_p = k < \tau_p') \leqslant \sum_{1 \leqslant j \leqslant k} \boldsymbol{P}(\tau_p^* = j,\ \tau_p' > j).$$

But clearly, if $\tau_p^* = j$ and $\tau_p^* < \tau_p'$, either one of the conjugacy classes $(X_1^\sharp, \ldots, X_j^\sharp)$ has trace zero, or otherwise we must have $\tau_p' = \tau_3 > j \geqslant 2$. In the first case, since all X_n have the same uniform distribution, the probability is at most

$$j\boldsymbol{P}(\mathrm{Tr}(X_1^\sharp) = 0) \ll jp^{-1}$$

for all $p \geqslant 2$ (again by looking at conjugacy classes for example). In the second case, we have

$$\boldsymbol{P}(\tau_3 > j) \leqslant \boldsymbol{P}(\tau_3 \geqslant 2) \ll p^{-2}.$$

Combining this with the equality of events we found, it follows that for k fixed, we have

$$\boldsymbol{P}(\tau_p = k) = \boldsymbol{P}(\tau_p = \tau_p' = k) + O(p^{-1})$$

where the implied constant depends on k.

Next we note that

$$\{\tau_p' = k\} = \{\tau_p = \tau_p' = k\} \cup \{\tau_k' = p,\ \tau_p < k\},$$

again a disjoint union. As above, we find that

$$\boldsymbol{P}(\tau_k' = p,\ \tau_p < k) \leqslant \sum_{j=1}^{k-1} \boldsymbol{P}(\tau_p^* = j < \tau_p') \ll p^{-1}$$

where the implied constant depends on k, and hence we have finally

$$\boldsymbol{P}(\tau_p = k) = \boldsymbol{P}(\tau_p' = k) + O(p^{-1}),$$

and the result now follows easily: first, by arguments already used, we have

$$\boldsymbol{P}(\tau_p' = k) = \boldsymbol{P}(\max(\tau_1, \tau_2) = k) + O(p^{-1})$$

and then we are left with a coupon collector problem with two coupons of roughly equal probability by (5.2). This gives

$$\boldsymbol{P}(\max(\tau_1, \tau_2) = k) = p_1^{k-1}p_2 + p_2^{k-1}p_1 = 2\Big(\frac{1}{2} + O(p^{-1})\Big)^k = \frac{1}{2^{k-1}} + O(p^{-1})$$

for $p \geqslant 2$, the implied constant depending on k. □

Remark 5.3. Part (2) states that the waiting time $\tau_{SL(2,\mathbf{F}_p)}$ converges in law, as $p \to +\infty$, to the waiting time for a coupon collector problem with two coupons of probability 1/2. Intuitively, those represent finding matrices with split or irreducible characteristic polynomial.

Remark 5.4. Recent results of Fulman and Guralnick (announced in [FG]) should lead to a similar good understanding of $c(\mathbf{G}(\mathbf{F}_q))$ when $\mathbf{G}$ is a fixed (almost simple) algebraic group over $\mathbf{Q}$. Indeed, their results should also be applicable to situations with rank going to infinity, which are analogue of the symmetric and alternating groups that we consider now.

6. Symmetric and alternating groups

We now come to the case of the symmetric groups $\mathfrak{S}_n$ and alternating groups A_n. Here we have the following result, which is a precise formulation of a result essentially conjectured by Dixon [D1, Abstract], following McKay:[8]

Theorem 6.1. *For $n \geqslant 1$, we have*

$$c(\mathfrak{S}_n) \asymp 1, \qquad c(A_n) \asymp 1, \qquad c_2(\mathfrak{S}_n) \asymp 1, \qquad c_2(A_n) \asymp 1.$$

In fact, there exists a constant $c > 1$ such that, for all $n \geqslant 1$, we have

$$\mathbf{E}(c^{\tau_{\mathfrak{S}_n}}) \ll 1, \qquad \mathbf{E}(c^{\tau_{A_n}}) \ll 1.$$

The proof is based on the following difficult result of Łuczak and Pyber, the proof of which involves a lot of information on symmetric groups.

Theorem 6.2 (Łuczak and Pyber)**.** *For any $\varepsilon > 0$, there exists a constant C depending only on ε such that*

$$\boldsymbol{P}((X_1^\sharp, \ldots, X_m^\sharp) \text{ generate } \mathfrak{S}_n) > 1 - \varepsilon$$

for all $m \geqslant C$ and all $n \geqslant 1$. The same applies to A_n.

This is proved in [LP], improving earlier work of Dixon [D1].

Proof. We need only prove that the exponential moments $\mathbf{E}(c^{\tau_n})$ are bounded for some $c > 1$, where $\tau_n = \tau_{G_n}$ with $G_n = \mathfrak{S}_n$ (the A_n case is similar).

¿From Theorem 6.2, there exists $m \geqslant 1$ such that

$$\boldsymbol{P}((Y_1^\sharp, \ldots, Y_m^\sharp) \text{ do not generate } \mathfrak{S}_n) \leqslant \frac{1}{2} \tag{6.1}$$

for any family of independent, uniformly distributed random variables Y_i on G_n.

Now let $k \geqslant 1$ be given; we can partition the set $\{1, \ldots, k-1\}$ in $\lfloor (k-1)/m \rfloor \geqslant 0$ subsets of size m and a remainder, and we observe that if $\tau_n = k$, for each of these subsets I, we have

$$\boldsymbol{P}((X_i^\sharp),\ i \in I) \leqslant \frac{1}{2},$$

by independence and (6.1). Since all those sets are disjoints, we get

$$\boldsymbol{P}(\tau_n = k) \leqslant \Bigl(\frac{1}{2}\Bigr)^{\lfloor (k-1)/m \rfloor} \leqslant 2^{1-(k-1)/m}$$

for $k \geqslant 1$, and then, for any $c \geqslant 1$, we have

$$\mathbf{E}(c^{\tau_n}) = \sum_{k \geqslant 1} c^k \boldsymbol{P}(\tau_n = k) \leqslant 2^{1+1/m} \sum_{k \geqslant 1} (c 2^{1/m})^k$$

which converges, and is independent of n, for any c with $1 < c < 2^{1/m}$. ∎

In view of this, the following question seems natural:

Question 1. Is it true or not that for *all* $c > 1$, we have

$$\mathbf{E}(c^{\tau_{\mathfrak{S}_n}}) \ll 1$$

for $n \geqslant 1$ (and similarly for A_n)?

[8] This conjecture is imprecisely formulated in [D1], where the "expected number of elements needed to generate $\mathfrak{S}_n$ invariably" seems to mean any $r(n)$ for which $\boldsymbol{P}(c(\mathfrak{S}_n) > r(n)) \to 0$.

Another natural question, also suggested by Dixon, is:

Question 2. Do the sequences $(c(\mathfrak{S}_n))$ (or $(c(A_n))$) converge as $n \to +\infty$? If they do, can their limits be computed?

Our guess is that the answer is positive. In fact, we now present a heuristic model that suggests this and predicts the value of

$$\lim_{n\to+\infty} c(A_n).$$

Our first step is to apply Corollary 2.18 to a suitable "essential" set of maximal subgroups of symmetric groups to obtain a simpler waiting time that is asymptotically close to $c(A_n)$ (or to $c(\mathfrak{S}_n)$). The required result is again one due to Łuczak and Pyber [LP].

Theorem 6.3 (Łuczak and Pyber)**.** *For $n \geqslant 1$, let S_n be the set of $g \in \mathfrak{S}_n$ such that g is contained in a subgroup H of $\mathfrak{S}_n$, distinct from A_n, and such that G acts transitively on $\{1, \ldots, n\}$. Then we have*

$$\lim_{n\to+\infty} \nu_n(S_n) = 0,$$

where $\nu_n(A) = |A|/|\mathfrak{S}_n|$ is the uniform density on the symmetric group.

Corollary 6.4. *For $n \geqslant 1$ and $1 \leqslant i < n/2$, let*

$$H_{i,n} = \big\{g \in \mathfrak{S}_n \mid g \cdot \{1, \ldots, i\} = \{1, \ldots, i\}\big\}$$

be the subgroup of $\mathfrak{S}_n$ leaving $\{1, \ldots, i\}$ invariant. Let $H'_{i,n} = H_{i,n} \cap A_n$. Then the $H_{i,n}$ – resp. $H'_{i,n}$ – are maximal subgroups of $\mathfrak{S}_n$ – resp. A_n –. Moreover, let

$$M_n = \{A_n\} \cup \{H_{i,n} \mid 1 \leqslant i < n/2\} \subset \max(\mathfrak{S}_n),$$
$$M'_n = \{H'_{i,n} \mid 1 \leqslant i < n/2\} \subset \max(A_n).$$

As in Proposition 2.17, let $\tilde{\tau}_n$, resp. $\tilde{\tau}'_n$, be the waiting time before conjugacy classes in each subgroup of M_n, resp. M'_n, has been observed. Then we have

$$c(\mathfrak{S}_n) = \mathbf{E}(\tilde{\tau}_n) + o(1), \qquad c_2(\mathfrak{S}_n) = \mathbf{E}(\tilde{\tau}_n^2) + o(1),$$

as $n \to +\infty$, and similarly

$$c(A_n) = \mathbf{E}(\tilde{\tau}'_n) + o(1), \qquad c_2(A_n) = \mathbf{E}((\tilde{\tau}'_n)^2) + o(1).$$

Proof. It is known that the $H_{i,n}$ are (representatives of) the conjugacy classes of maximal intransitive subgroups of $\mathfrak{S}_n$. Thus, we find by definition of S_n that

$$\bigcup_{H\in\max(\mathfrak{S}_n)-M_n} H^\sharp = S_n,$$

and hence the result follows immediately from Corollary 2.18 and Theorem 6.3, which provides us with the assumption (2.19) required. □

In particular, in approaching Question 2, it is enough to consider the expectations and second moment of the random variables $\tilde{\tau}_n$ and $\tilde{\tau}'_n$. Those are combinatorially simpler, or at least more explicit.

In particular, note the following: an element $\sigma \in \mathfrak{S}_n$ is conjugate to an element of $H_{i,n} \subset \mathfrak{S}_n$ *if and only if*, when expressed as a product of disjoint cycles of lengths $\ell_j(\sigma) \geqslant 1$,

$1 \leqslant j \leqslant \varpi(\sigma)$, say, has the property that a sum of a subset of the lengths is equal to i: for some $J \subset \{1, \ldots, \varpi(\sigma)\}$, we have

$$\sum_{j \in J} \ell_j(\sigma) = i.$$

Note also that this applies equally to an element σ in A_n: the element is conjugate to $H'_{i,n} \subset A_n$ if and only if the property above is true for its cycle lengths computed in $\mathfrak{S}_n$ (although these cycle lengths do not always characterize the conjugacy class of σ in A_n).

In particular, conjugacy classes $(\sigma_1^\sharp, \ldots, \sigma_k^\sharp)$ in $\mathfrak{S}_n^\sharp$ (or $A_n^\sharp$) generate a transitive subgroup of $\mathfrak{S}_n$ (or A_n) *if and only* if n (which always occurs as the sum of all lengths) is the only common such sum for all σ_j. (Indeed, if $i < n$ occurs as a common subsum, we can assume that $i \leqslant n/2$, and then it is possible to select elements in each conjugacy class which all belong to $H_{i,n}$, so that the conjugacy classes can not generate invariably a transitive subgroup; the converse is also simple.)

Now, we come to the model when $n \to +\infty$. The distribution of the set of lengths of random permutations is a well-studied subject in probabilistic group theory, and this allows us to make a guess as to the existence and value of the limit.

Indeed, for $i \geqslant 1$, consider the map

$$\varpi_i : \mathfrak{S}_n \to \{0, 1, \ldots\}$$

sending σ to the number of cycles of length i in its decomposition as product of disjoint cycles (for $i = 1$, this is the number of fixed points; for $i \geqslant n+1$, of course, this is zero, but it will be convenient for the asymptotic study to allow arbitrary i). Now consider, for each $n \geqslant 1$, any random variables s_n, σ_n uniformly distributed on $\mathfrak{S}_n$ and A_n, respectively. Then the following is a consequence of well-known results dating back to Goncharov [Go]: for fixed i, as $n \to +\infty$, the random variables $\varpi_i(\sigma_n)$ converge in law to a Poisson random variable with parameter $1/i$, i.e., we have

$$\lim_{n \to +\infty} \boldsymbol{P}(\varpi_i(\sigma_n) = k) = e^{-1/i} \frac{1}{k! i^k}, \qquad \text{for fixed } k \geqslant 0. \tag{6.2}$$

Moreover, the limits for distinct values of i are independent, i.e., for any fixed finite set I of positive integers, we have

$$\lim_{n \to +\infty} \boldsymbol{P}(\varpi_i(\sigma_n) = k_i \text{ for all } i \in I) = \prod_{i \in I} e^{-1/i} \frac{1}{i^{k_i} k_i!}.$$

More precisely, this is proved (and much more precise results) for symmetric groups in, e.g., [AT, Th.1] or [ABT, Th. 1.3]. The case of alternating groups can be deduced from this using the fact that the indicator function of A_n in $\mathfrak{S}_n$ is given in terms of cycle-lengths by

$$\frac{1 + (-1)^{\varpi_2 + \varpi_4 + \cdots}}{2}.$$

For instance, for fixed j, the characteristic function of $\varpi_k(\sigma_n)$ is

$$\mathbf{E}(e^{it\varpi_k(\sigma_n)}) = \mathbf{E}(e^{it\varpi_k(s_n)}) + \mathbf{E}((-1)^{\sum_j \varpi_{2j}(s_n)} e^{it\varpi_k(s_n)}).$$

By Goncharov's result, the first term converges for every $t \in \mathbf{R}$ to the desired characteristic function of a Poisson variable with parameter $1/k$; for the second term, we can use the method of Lloyd and Shepp [LS, §2]. Assuming $k = 2k'$ is even (the other case being similar), one finds (see in particular [LS, (3)]) that the expectation over $\mathfrak{S}_n$ is the coefficient of z^n in

$$\frac{1}{1-z}\exp\Bigl(\frac{z^k}{k}(e^{i(t+\pi)}-1)\Bigr)\prod_{\substack{j\geqslant 1\\ j\neq k'}}\exp(-z^{2j}/j) = (1+z)\exp\Bigl(\frac{z^k}{k}(1-e^{it})\Bigr),$$

and since this function (of $z \in \mathbf{C}$) is regular at $z = 1$, those coefficients converge to 0 for every fixed t. This computation proves (6.2).

It seems therefore reasonable to use a model of Poisson variables to predict the limit of Chebotarev invariants of alternating groups. For this purpose, let $\mathcal{A}$ be the set of sequences $(\ell_i)_{i\geqslant 1}$ of non-negative integers; we denote the i-th component of $\ell \in \mathcal{A}$ by $\varpi_i(\ell)$. Let $\nu_{\mathcal{A}}$ be the infinite product (probability) measure on $\mathcal{A}$ such that the i-th component ℓ_i is distributed like a Poisson random variable with parameter $1/i$. This set $\mathcal{A}$ is meant to be like the set of conjugacy classes of an infinite symmetric group, and indeed, from the above, we see that for any finite I of positive integers and any $k_i \geqslant 0$ defined for $i \in I$, we have

$$\lim_{n\to+\infty}\boldsymbol{P}(\varpi_i(\sigma_n) = k_i \text{ for all } i \in I) = \nu_{\mathcal{A}}(\{\ell \in \mathcal{A} \mid \varpi_i(\ell) = k_i,\ i \in I\}).$$

Now consider an infinite sequence $(X_k)_{k\geqslant 1}$ of $\mathcal{A}$-valued, independent random variables, identically distributed according to ν. We look at the following waiting time:

$$\tau_{\mathcal{A}} = \min\{k \geqslant 1 \mid \bigcap_{1\leqslant j\leqslant k} S(X_j) = \{+\infty\}\},$$

where, for $\ell \in \mathcal{A}$, we denote by $S(\ell) \subset \{0, 1, 2, \ldots,\} \cup \{+\infty\}$ the set of all sums

$$\sum_{i\geqslant 1} ib_i, \qquad \text{where } 0 \leqslant b_i \leqslant \varpi_i(\ell)$$

(note the usual shift of notation from our description of the case of fixed n: the sequence of lengths of cycles occuring in a permutation is replaced by the sequence of multiplicities of each possible length). Then our guess for the limit of $c(A_n)$ is that

$$\lim_{n\to+\infty} c(A_n) = \mathbf{E}(\tau_A).$$

We hope to come back to this question in a future work.

7. Non-abelian groups: numerical experiments

In this section, we give some tables of values of the Chebotarev invariant (and the secondary invariant) for some non-abelian finite groups. Although those are clearly rational numbers, we list real approximations only because the "height" of those rationals grows very fast.

The computations are feasible even for fairly large and complicated non-abelian groups, because they may have few conjugacy classes of maximal subgroups, and not too many conjugacy classes. For instance, the Weyl group $W(E_8)$ (one of our motivating examples) has 9 conjugacy classes of maximal subgroups, and 112 conjugacy classes. If these data are available to suitable software packages, Proposition 2.7 provides a way to compute the Chebotarev invariants, though this is at best an exponential-time algorithm (due to the necessity to sum over all subsets of $\max(G)$).

The computations here were done for the most part with Magma (see [M]), using the script included in the Appendix. The correctness of the results was checked partly by independent computations with the open-source package GAP (see [GAP]), and by checking that the results agree, for cyclic groups and groups $\mathbf{F}_p^k$, with the theoretical formulas of the Section 3. They are also in good agreement, in the case of $PSL(2, \mathbf{F}_p)$, with the asymptotic result of Section 5. Hence, altogether, we have very high confidence in these values.

The computations were relatively fast; usually there was a sharp threshold between computing for one group in a family in less than an hour, and the next one proving infeasible due to the exponential growth of the number of subsets of $\max(G)$. As an indication of timing, the computation for $PSL(6, \mathbf{F}_3)$ with Magma (version 2.14.15) took about 42 seconds on a 2.5 GHz Core 2 processor.

Below, we include tables for the alternating groups A_n, for the symmetric groups $\mathfrak{S}_n$, for the groups $PSL(2, \mathbf{F}_p)$ with p prime $\leqslant 150$ (though the computations can be done for p quite a bit larger, we do not include the results which are not particularly enlightening), for $PSL(3, \mathbf{F}_p)$, $PSL(4, \mathbf{F}_p)$, $PSL(n, \mathbf{F}_2)$, $PSL(n, \mathbf{F}_3)$, $PSL(n, \mathbf{F}_4)$, $PSL(2, \mathbf{F}_{2^n})$, $PSL(3, \mathbf{F}_{2^n})$, $Sp(2g, \mathbf{F}_3)$. (Note that, in general, the computations tend to run quite a bit faster for simple groups.) We also include a table of the "partial" invariants $\mathbf{E}(\tilde{\tau}'_n)$ and $\mathbf{E}((\tilde{\tau}'_n)^2)$ of alternating groups defined in Corollary 6.4. Note that although we have shown that these are asymptotically converging to the Chebotarev invariants themselves, the convergence is by no means visible! There is also a table for the Borel subgroup of $SL(3, \mathbf{F}_p)$, namely

$$B_3(\mathbf{F}_p) = \Big\{ \begin{pmatrix} x & r & s \\ 0 & y & t \\ 0 & 0 & z \end{pmatrix} \mid (r, s, t) \in \mathbf{F}_p^3,\ (x, y, z) \in (\mathbf{F}_p^\times)^3,\ xyz = 1 \Big\},$$

as another example of a solvable group.

Another table lists some more "sporadic" groups; the names of those groups in the table should be self-explanatory. For instance, D_{2n} is the dihedral group of order $2n$, $W(R)$ denotes the Weyl group of a root system of type R; $Sz(8)$ is the Suzuki group. The groups M_n, where $n \in \{11, 12, 22, 23, 24\}$, J_k, $k \in \{1, 2, 3, 4\}$, McL and HS are some sporadic simple groups: the Mathieu groups, the Janko groups, the MacLaughlin group and the Higman-Sims group, respectively. The group Rub at the end of the table is the Rubik's group (the subgroup of $\mathfrak{S}_{48}$ that gives the possible moves on the Rubik's Cube; computing $c(Rub)$ takes about two days on a fast Opteron; this group has 20 conjugacy classes of maximal subgroups and 81120 conjugacy classes). In order to ease checking, the url

`http://www.math.ethz.ch/~kowalski/other-groups.mgm`

contains a Magma file where each group in this list is constructed explicitly.

It also possible to exploit the databases of small groups, or of transitive groups, or primitive groups, to compute the Chebotarev invariants for, say, all groups of a given small order (up to isomorphism), or for all transitive permutations groups of small degree. The latter is of course particularly interesting from the point of view of Galois theory, and the groups $\mathbf{F}_q \rtimes \mathbf{F}_q^\times$ which appear as transitive permutation groups of degree q (and in Galois theory as Galois groups of Kummer extensions of prime-power degree, i.e., splitting fields of polynomials of the type $X^q - a$) are very noticeable, having much higher Chebotarev invariants than the other groups despite their rather small order (see the example in the table for transitive groups of degree 17 – noting that the group with Chebotarev invariant roughly 8.88 is the index 2 subgroup of H_{17} denoted C_2 in Section 4). We include a figure of the empirical distribution of values for the Chebotarev waiting time for H_{31} (chosen because $q-1=30$ has "many" divisors).

We also include a figure with an histogram showing the distribution of the Chebotarev invariant for the 840 distinct groups of order 720 (up to isomorphism). Note that this data also indicates that the invariant is far from injective (as can be guessed from its dependency on relatively little data): there are only 188 distinct values of $c(G)$ for $|G|=720$; the value

$$\frac{469589438194474533813031879}{80462083849550829871525080} \simeq 5.836158\ldots.$$

occurs with maximal multiplicity (it arises 39 times).

Note that for simple groups (or groups which are nearly so), the relation between $c(G)$ and $c_2(G)$ seems relatively regular, but there is certainly no strict monotony in terms of the order; see, e.g., the cases of alternating groups A_n, where sorting according to the value of $c(A_n)$ leads to the following rather bizarre ordering of the segment $2 \leqslant n \leqslant 21$:

$$2, 3, 13, 19, 17, 11, 5, 10, 14, 20, 21, 16, 18, 15, 4, 6, 12, 9, 7, 8\ ;$$

the ordering with respect to $c_2(A_n)$ is slightly different, namely:

$$2, 3, 13, 19, 11, 17, 10, 14, 21, 20, 16, 18, 15, 5, 12, 6, 9, 4, 7, 8.$$

And the orderings for $c(\mathfrak{S}_n)$ and $c_2(\mathfrak{S}_n)$ are also different:

$$2, 3, 7, 11, 13, 9, 17, 19, 5, 15, 21, 16, 20, 4, 14, 18, 12, 10, 8, 6,$$

and

$$2, 3, 7, 11, 13, 9, 17, 19, 15, 5, 21, 16, 20, 14, 18, 12, 8, 10, 4, 6,$$

respectively. Note however that in Table 2, if we fix the parity of n, the invariants $\mathbf{E}(\tilde{\tau}'_{2n})$ and $\mathbf{E}(\tilde{\tau}'_{2n+1})$ seem monotonically increasing. This indicates that they are indeed very natural objects to study.

8. Arithmetic considerations

In this short section, we indicate the (expected) number-theoretic connections of our work.

First, let K be a Galois extension of $\mathbf{Q}$ with group G. For each prime p that is unramified in K, we have a well-defined Frobenius conjugacy class $\mathrm{Fr}_{p,K} \in G^\sharp$. For simplicity, we denote $\mathrm{Fr}_{p,K}=1$ when p is ramified in K. The *Chebotarev density theorem* says that

$$\lim_{y\to+\infty} \frac{|\{p \leqslant y : \mathrm{Fr}_{p,K}=C\}|}{\pi(y)} = \frac{|C|}{|G|}$$

where C is a fixed conjugacy class of G and $\pi(y)$ is the number of primes that are at most y.

Now fix a real number y large enough, so that every conjugacy class of G is of the form $\mathrm{Fr}_{p,K}$ for some $p \leqslant y$. For each $i \geqslant 1$, select a random prime p from the set $\{p : p \leqslant y\}$ and define $X^\sharp_{i,y} = \mathrm{Fr}_{p,K}$. We thus have a sequence of independent and identically distributed random variables $X(y) = (X^\sharp_{i,y})$ in $G^\sharp$. As usual, we define the waiting time

$$\tau_{X(y),G} = \min\{n \geqslant 1 \mid (X^\sharp_{1,y}, \ldots, X^\sharp_{n,y}) \text{ generate } G\} \in [1, +\infty].$$

Using the Chebotarev density theorem, one can show that

$$\lim_{y\to+\infty} \mathbf{E}\big(\tau_{X(y),G}\big) = c(G).$$

FIGURE 1. Empirical distribution of the waiting time for H_{31}

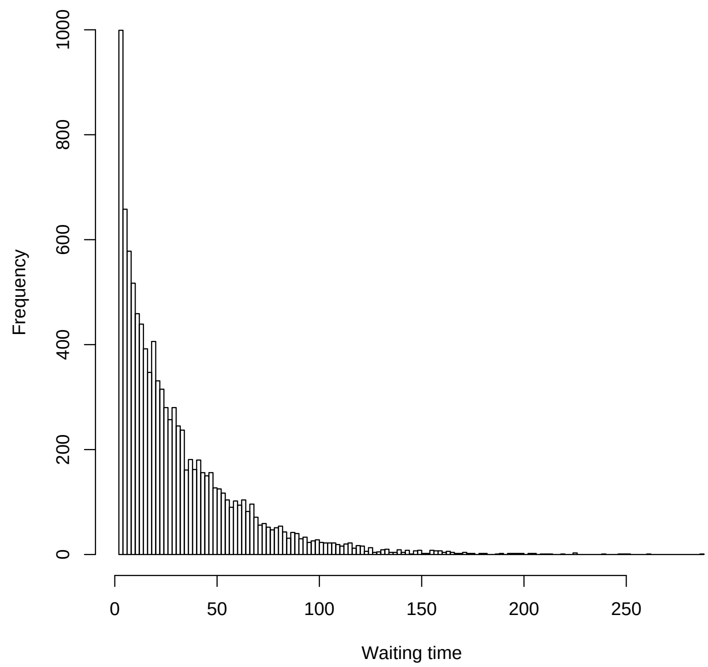
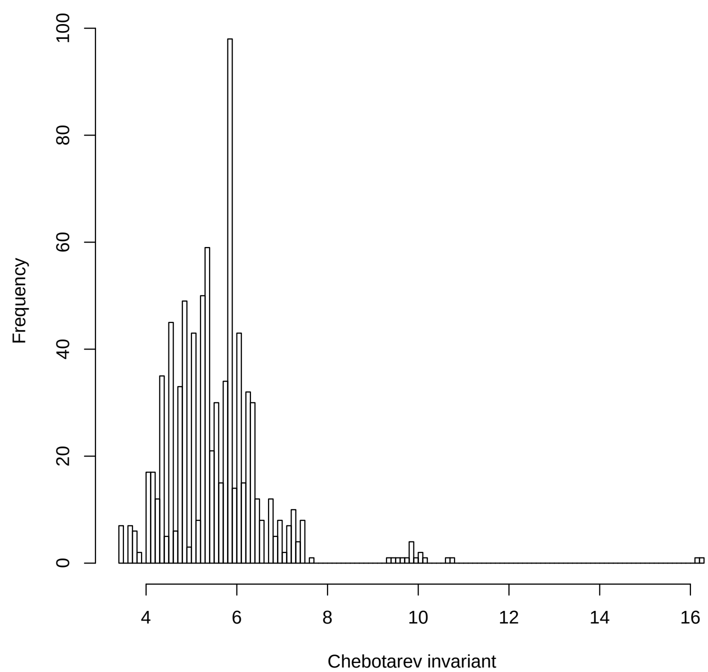


FIGURE 2. Distribution of the Chebotarev invariant for groups of order 720



Therefore, $c(G)$ can be thought of as the expected number of "random" primes p needed for $\mathrm{Fr}_{p,K}$ to generate $G = \mathrm{Gal}(K/\mathbf{Q})$.

Of course in practice, one usually considers the (non-random) sequence $\mathrm{Fr}_{2,K}$, $\mathrm{Fr}_{3,K}$, $\mathrm{Fr}_{5,K}$, $\mathrm{Fr}_{7,K}, \ldots$. We now explain, informally, what can be expected to happen in that situation.

Suppose we have some family $\mathcal{K}$ of finite Galois extensions of $\mathbf{Q}$ (or another base field), all (or almost all) of which have Galois group $\mathrm{Gal}(K/\mathbf{Q}) \simeq G$, a fixed finite group, and that, for all values of some parameter $x \geqslant 1$, we have finite subfamilies $\mathcal{K}_x$ (which exhaust $\mathcal{K}$ as $x \to +\infty$) and some averaging process for invariants of the fields in $\mathcal{K}$, denoted $\mathbf{E}_x$ (for instance, one might take

$$\mathbf{E}_x(\alpha(K)) = \frac{1}{|\mathcal{K}_n|} \sum_{K \in \mathcal{K}_x} \alpha(K)$$

but other weights, involving multiplicities, etc, might be better adapted). For every prime p such that p is unramified in K, we have a well-defined Frobenius conjugacy class $\mathrm{Fr}_{p,K} \in G^\sharp$. For simplicity, we denote $\mathrm{Fr}_{p,K} = 1$ when p is ramified in K. We can now define the following analogue of the Chebotarev waiting time:

$$\tau(K) = \min\{k \geqslant 1 \mid \text{the first } k \text{ conjugacy classes } \mathrm{Fr}_{2,K}, \ldots, \mathrm{Fr}_{p_k,K} \text{ generate } G\},$$

where p_k is the k-th prime number, and the empirical Chebotarev invariants:

$$c(\mathcal{K}_x) = \mathbf{E}_x(\tau(K)), \qquad c_2(\mathcal{K}_x) = \mathbf{E}_x(\tau(K)^2).$$

The basic arithmetic question is then: for a given family, is it true that $c(\mathcal{K}_x)$ is, for x sufficiently large at least, close to $c(G)$ (and similarly for the secondary Chebotarev invariant)?

The basic reason one can expect this is the Chebotarev density theorem, which states that, for a fixed finite Galois extension $K/\mathbf{Q}$ with $\mathrm{Gal}(K/\mathbf{Q}) \simeq G$, we have

$$\lim_{x \to +\infty} \frac{1}{\pi(x)} |\{p \leqslant x \mid \mathrm{Fr}_{p,K} = c^\sharp\}| = \nu_G(c^\sharp),$$

for any conjugacy class $c^\sharp \in G^\sharp$. Thus, $\mathrm{Fr}_{p,K}$ is asymptotically distributed like a random conjugacy class distributed according to the measure ν_G. This strongly suggests that the Chebotarev invariants should be comparable.

We want to point out a few difficulties that definitely arise in trying to make this precise.

First of all, quantifying the Chebotarev density theorem is *hard*: it almost immediately runs into issues related to the Generalized Riemann Hypothesis; even in the seemingly trivial case where $G = \mathbf{Z}/2\mathbf{Z}$ (quadratic extensions of $\mathbf{Q}$), the basic question of estimating the size of the smallest prime p for which the corresponding Frobenius is non-trivial, i.e., the smallest quadratic non-residue modulo p, in terms of the discriminant of the field, is unsolved (see, e.g., [IK, Prop. 5.22, Th. 7.16] for conditional and uncontional results in that case). This is a problem because if we sum with uniform weight, a single "bad" field K_0 can destroy any chance of approaching the Chebotarev invariant. Indeed: note that in that case

$$\mathbf{E}_x(\tau(K)) \geqslant \frac{1}{|\mathcal{K}_x|} k_{min}(K_0) \tag{8.1}$$

where

$$k_{min}(K) = \min\{k \geqslant 1 \mid \mathrm{Fr}_{p,K} \neq 1\}$$

is the index of the first non-trivial Frobenius conjugacy class. In the current state of knowledge, it can be that there exists K_0 with

$$k_{min}(K_0) > \mathrm{disc}(K_0)^A$$

for some constant $A > 0$ (see [LMO]); on the other hand, if the family $\mathcal{K}$ is defined as that of splitting fields of monic polynomials of degree n, and the subfamily $\mathcal{K}_x$ is that of polynomials of height $\leqslant x$, then we know that most $K \in \mathcal{K}$ have Galois group $\mathfrak{S}_n$, that $|\mathcal{K}_x| = (2x+1)^n$ if x is an integer, and the discriminant is obviously often also *at least* a power of x. Thus (8.1) might already be bad enough to preclude any comparison. On the other hand, on the Riemann Hypothesis, we have

$$k_{min}(K) \ll (\log \mathrm{disc}(K))^2,$$

(where the implied constant depends on G), and the problem would then be alleviated.

Another issue is that one can not expect, as stated, to have

$$\lim_{x\to+\infty} c(\mathcal{K}_x) = c(G)$$

for interesting families for the simple reason that the statistic of small primes is typically *not* the uniform one, i.e., if we fix a prime p, we can not expect to have

$$\lim_{x\to+\infty} \mathbf{E}_x(\mathbb{1}_{\{\mathrm{Fr}_{p,K}=c^\sharp\}}) = \nu_G(c^\sharp),$$

even if we assume that all the fields involved are unramified at p.

For instance, consider $\mathcal{K}$ the set of cubic polynomials

$$X^3 + a_2X^2 + a_1X + a_0,$$

with $a_i \in \mathbf{Z}$, with $\mathcal{K}_x$ those where $|a_i| \leqslant x$ for $i = 0, 1, 2$, and count them uniformly. Take $p = 5$ and consider only polynomials with no repeated root modulo 5 and splitting field of degree 6; then, asymptotically, the conjugacy Frobenius at 5 will be distributed in $\mathfrak{S}_3 = G$ as dictated by the factorization of the polynomial modulo 5. One finds easily that there are 100 monic cubic polynomials in $\mathbf{F}_5[X]$ with non-zero discriminant (there are 25 with repeated roots), among which:

- 10 split in linear factors, i.e., a density 1/10;
- 40 are irreducible, i.e., a density 4/10;
- 50 split as a product of one linear factor and one irreducible quadratic factor, i.e., a density 1/2.

This is in sharp constrast with the density of the three corresponding conjugacy classes in $\mathfrak{S}_3$, which are respectively:

- 1/6 for the identity class;
- 1/2 for the 3-cycles;
- 1/3 for the transposition.

In particular, not even the relative frequencies are preserved! On the other hand, it is well-known that if p is increasing, the discrepancy between the distribution of the factorization patterns of squarefree polynomials modulo p and the density of conjugacy classes disappears: we have

$$\frac{1}{p^n}|\{f \in \mathbf{F}_p[X] \mid f \text{ squarefree of degree } n \text{ with } \mathrm{Fr}_f = c^\sharp\}| \sim \nu_G(c^\sharp)$$

uniformly for all conjugacy classes $c^\sharp \in G = \mathfrak{S}_n$.

This suggests that it is likely that one can prove some relevant results: one would consider some increasing starting point $s(x) \geqslant 2$ and a modified waiting time

$$\tau_x(K) = \min\{k \mid \text{the first } k \text{ conjugacy classes } \mathrm{Fr}_{p,K} \text{ with } p \geqslant s(x) \text{ generate } G\}$$

and hope to prove (possibly under the Generalized Riemann Hypothesis, possibly unconditionally after throwing away a few "bad" fields) that

$$\lim_{x \to +\infty} \mathbf{E}_x(\tau_x(K)) = c(G),$$

for suitable $s(x)$. One may guess that for polynomials of height $\leqslant x$ and fixed degree n (and $G = \mathfrak{S}_n$), this would be true with $s(x) \asymp \log x$.

9. Remarks and problems

We finish with a few more remarks and problems.

- (What does the invariant "know"?) As a bare numerical invariant of a finite group, the Chebotarev invariant seems to be fairly subtle. For instance, we see from Section 3 that it "knows" that vector spaces over finite fields are in some sense very similar for varying base field, but that they become also "simpler" as the cardinality of the base field grows. It also seems to know that non-reductive finite matrix groups are worse-behaved than reductive ones (as shown by the results for H_q). What else does the invariant reveal?
- (A method for upper bounds) There are, in the literature, quite a few results about a finite group G of the type: "if a subgroup H contains elements in some set C_1, some other set C_2, ..., some other set C_m, of conjugacy classes, then H is in fact equal to G". For instance, a lemma of Baer quoted by Gallagher [G, Lemma, p. 98] says that there is no proper subgroup H of $\mathfrak{S}_n$ which (1) contains an n-cycle, (2) contains a product of a transposition and cycles of odd lengths, (3) contains an element of order divisible by a *prime* $p > n/2$. Another such result is the Lemma 5.2 of Serre for $SL(2, \mathbf{F}_p)$, and we also mention [JKZ, Lemma 3.2] for another example with the Weyl group $W(E_8)$, and there are many other such results used, e.g., for proving concrete cases of Hilbert's Irreducibility Theorem.

  With this notation, and assuming we work with a sequence of independent and uniformly distributed G-valued random variables (X_n), this means that we have

  $$\tau_G \leqslant \tau_{C_1,\ldots,C_m} = \max(\tau_{C_j},\ 1 \leqslant j \leqslant m),$$

  where

  $$\tau_{C_j} = \min\{n \geqslant 1 \mid X_n^\sharp \in C_j\}.$$

  From Proposition 2.9, we obtain easily an upper bound

  $$c(G) \leqslant \mathbf{E}(\tau_{C_1,\ldots,C_m}) = \sum_{\emptyset \neq I \subset \{1,\ldots,m\}} \frac{(-1)^{|I|+1}}{\nu\Bigl(\bigcup_{j \in I} C_j\Bigr)}, \tag{9.1}$$

  and one may hope to approximate $c(G)$ by choosing wisely the sets (C_j).

  However, it is not clear at all to what extent this can approach the truth. Here are some examples:

  (1) Baer's lemma gives only an upper bound

  $$c(\mathfrak{S}_n) \ll n$$

as $n \to +\infty$, which is quite weak compared with Theorem 6.1 (it is dominated by the time required to obtain an n-cycle). How far is this from the best possible result that can be obtained in this way, and how far is the latter from Theorem 6.1?

(2) Consider $G = \mathbf{F}_p^2$ with p odd. It is possible to take

$$C_1 = \{(x,y) \in \mathbf{F}_p^2 - \{0\} \mid y \neq 0 \text{ and } xy^{-1} \text{ is a square in } \mathbf{F}_p\},$$
$$C_2 = \mathbf{F}_p^2 - \{0\} - C_1.$$

The point is that whenever $(v,w) \in C_1 \times C_2$, w and v are not on the same line through the origin, so (v,w) generate $\mathbf{F}_p^2$. Since

$$|C_1| = |C_2| = (p^2-1)/2,\ |C_1 \cup C_2| = p^2 - 1,$$

this leads to

$$c(\mathbf{F}_p^2) \leqslant \frac{1}{\nu(C_1)} + \frac{1}{\nu(C_2)} - \frac{1}{\nu(C_1 \cup C_2)} = \frac{3p^2}{p^2-1},$$

which asymptotically requires one more step on average than the right Chebotarev invariant (given by (3.2)), namely $c(\mathbf{F}_p^2) = (2p^2+p)/(p^2-1)$. It seems also that this type of sets is essentially best possible for applying this upper bound in this case.

(3) Consider $G = W(E_8)$, the Weyl group of E_8. There is a non-trivial homomorphism

$$\varepsilon : W(E_8) \to \{\pm 1\},$$

and in [JKZ, Lemma 3.2], jointly with F. Jouve, we proved that one could take $C_1 = \ker(\varepsilon)$, C_2 the union of the conjugacy classes of w and w^2, where $w \in W(E_8)$ is a Coxeter element; the density of C_2 is $1/15$ and we then get the upper-bound

$$2 + 15 - 30/17 = 25.23\ldots$$

instead of the correct value $4.194248\ldots$.

(4) For $SL(2,\mathbf{F}_p)$, Theorem 5.1 shows that the sets C_1, C_2, C_3 given by Lemma 5.2 give an asymptotically optimal answer (and this is an essential ingredient in the proof).

Despite this relative inefficiency, it is interesting to notice that in applications of sieve methods to probabilistic Galois theory (as was the case in [G]) and [JKZ],[9] it is this type of distinguishing families which can be used in estimating how rare "small" Galois groups are in certain families, and in fact it is the quantity

$$\sum_{i=1}^{m} \frac{1}{\nu(C_i)} \tag{9.2}$$

which occurs naturally as coefficient in a "saving factor" of the large sieve; see, e.g, [K1, p. 57], where the question of minimizing this by varying the sets was raised explicitly for symmetric groups.

- (General upper bounds?) A first problem is to bound $c(G)$ from above, in a meaningful way. Since we have

$$\tau_G \leqslant \sum_{H \in \max(G)} \hat{\tau}_H.$$

[9] If only implicitly in the latter.

we obtain

$$c(G) \leqslant \sum_{H \in \max(G)} \frac{1}{1-\nu(H^\sharp)},$$

from (2.3). Together with (2.15), this gives an upper bound

$$c(G) \leqslant |G| \sum_{H \in \max(G)} \frac{1}{|H|} \tag{9.3}$$

which is close to being sharp for the groups H_q of Section 4: indeed, if q is odd, then Lemma 4.2 gives

$$\begin{aligned} |H_q| \sum_{H \in \max(H_q)} \frac{1}{|H_q|} &= q(q-1)\Big(\frac{1}{q-1} + \sum_{\ell \mid q-1} \frac{\ell}{q(q-1)}\Big) \\ &= q + \sum_{\ell \mid q-1} \ell = q + 2 + \sum_{2<\ell \mid q-1} \ell \end{aligned}$$

(where ℓ runs over prime divisors of $q-1$). If $q = 2\ell + 1$ (ℓ odd prime) is a Sophie Germain prime, this gives

$$q + 2 + \sum_{2<\ell \mid q-1} \ell = q + 2 + \frac{q-1}{2} = \frac{3(q+1)}{2},$$

which is off, asymptotically, only by a factor $3/2$ from the value

$$c(H_q) = q + O(q^{-1})$$

that follows from (4.1). Of course, it is not known that there are infinitely many Sophie Germain primes, but for $q = 2\ell_1\ell_2 + 1$, with ℓ_i prime, we have

$$\begin{aligned} |H_q| \sum_{H \in \max(H_q)} \frac{1}{|H_q|} &= \begin{cases} q + \ell_1 + \ell_2 + 2 & \text{if } \ell_1 \neq \ell_2 \\ q + \ell_1 + 2 & \text{if } \ell_1 = \ell_2 \end{cases} \\ &\leqslant 2q. \end{aligned}$$

By sieve methods, it is known that there are infinitely many primes q for which either q is a Sophie Germain prime, or is $2\ell_1\ell_2 + 1$, and hence one sees that the "trivial" estimate (9.3) above can not be improved by more than a constant in full generality. On the other hand, it is very far off for many groups: for a random example, it gives

$$4.7820\ldots = c(A_7) \leqslant 93.$$

It would be more interesting to have a decent upper bound in terms of the order of G only. Here, using the set of all conjugacy classes in (9.1), we get as an upper bound from the contribution of singletons that

$$c(G) \leqslant \sum_{g^\sharp \in G^\sharp} \frac{1}{\nu(g^\sharp)} = \sum_{g^\sharp \in G^\sharp} |N_G(g)|,$$

(where $N_G(g)$ is the normalizer of g in G). This gives trivially

$$c(G) \leqslant |G|^2,$$

but this seems unlikely to be close to the truth (for $G \neq 1$). For instance, since

$$c(H_q) = q \sim \sqrt{|H_q|},$$

one may wonder if H_q is also (essentially) extremal in this sense, i.e., one may ask whether an estimate

$$c(G) \ll \sqrt{|G|}$$

holds for all G. (Certainly for $|G| = q(q-1)$ with $q \leqslant 43$ prime, it is experimentally true that H_q maximizes the Chebotarev invariant).

- (Other classes of groups?) There are many classes of groups for which it should be possible to understand the behavior of the Chebotarev invariant, at least asymptotically. For instance, one can consider non-reductive subgroups of finite matrix groups, e.g., the standard Borel subgroup (upper triangular matrices) of $GL(n, \mathbf{F}_q)$. In fact, solvable groups seem particularly interesting.

Appendix: Magma script

The following script can be used to compute the Chebotarev invariant (and the secondary invariant) using Magma, by applying the formulas (2.3) and (2.4). The output is given as real approximations since usually the denominators are unwieldy. Also note that because of the use of the construct `Subsets(M)`, this script only applies to groups with at most 29 conjugacy classes of maximal subgroups;[10] to – try to – compute further, one would have to replace the loop over subsets obtained in this manner with a hand-rolled one.

A similar GAP script is available upon request, as well as a Sage version, which basically calls the GAP group theory routines. However, these versions are much slower.

The last routine in the script is useful for "empirical" study of the probabilistic model.

```
// The following calculates J such that
// J[k][i]=true if and only if the k-th maximal subgroup
// of G intersects the i-th conjugacy class of G

MCIntersectionMatrix:=function(G,C,f,M)
  J := [ [false : i in [1..#C]] : k in [1..#M] ];
  for k in [1..#M] do
    H := M[k]`subgroup;
    CH := ConjugacyClasses(H);
    for j in [1..#CH] do
      J[k][f(CH[j][3])] := true;
    end for;
  end for;
  return J;
end function;

// This returns [c,s] where c is the Chebotarev invariant of G
// and s the secondary invariant.

Chebotarev:= function (G)
  if IsTrivial(G) then
```

[10] For alternating groups A_n, this means $n \leqslant 47$, or $n \in \{49, 51, 53\}$.

```
    return <1.0,1.0>;
  end if;

  C := ConjugacyClasses(G);
  f := ClassMap(G);
  M := MaximalSubgroups(G);
  J := MCIntersectionMatrix(G,C,f,M);
  c:=0.0; s:=0.0;

  for I in Subsets({1..#M}) do
    if #I ne 0 then
      v:=0;
      for i in [1..#C] do
        if forall(t) {k: k in I | J[k][i]} then
          v:= v + C[i][2]/#G;
        end if;
      end for;
      c := c + (-1)^(#I+1)/(1-v);
      s := s+ (-1)^(#I)/(1-v)*(1-2/(1-v));
    end if;
  end for;
  return([c,s]);

end function;

// Compute empirical Chebotarev invariant.
// The optional parameter steps is the number
// of iterations to do. Example:
// > EmpiricalChebotarev(Alt(7):steps:=10000);

EmpiricalChebotarev:=function(G : steps:=1)
  total:=0;
  C := ConjugacyClasses(G);
  f:=ClassMap(G);
  M := MaximalSubgroups(G);
  J := MCIntersectionMatrix(G,C,f,M);
  for count in [1..steps] do
    nb:=0;
    vprint User1: "Iteration, ", count;
    // Start with all subgroups
    possible:=[ 1..#M ];
    while possible ne [] do
      g:=Random(G);
      nb := nb+1;
      index:=f(g);
      // Only those subgroups containing the class of g remain
      possible:=[ k : k in possible | J[k][index] ];
    end while;
    total:=total+nb;
```

```
  end for;
  return total/steps, total/steps*1.0;
end function;
```

References

[A] I. Adler and S. Ross: *The coupon subset collection problem*, J. Applied Prob. 38 (2001), 737–746.

[ABT] R. Arratia, A.D. Barbour and S. Tavaré: *Logarithmic combinatorial structures: a probabilistic approach*, E.M.S. Monographs, 2003.

[AT] R. Arratia and S. Tavaré: *The cycle structure of random permutations*, Annals of Prob. 20 (1992), 1567–1591.

[M] W. Bosma, J. Cannon and C. Playoust: *The Magma algebra system, I. The user language* J. Symbolic Comput., 24 (1997), 235–265; also `http://magma.maths.usyd.edu.au/magma/`

[D1] J.D. Dixon: *Random sets which invariably generate the symmetric group*, Discrete Math. 105 (1992), 25–39.

[D2] J.D. Dixon: *Probabilistic group theory*, C.R. Math. Rep. Acad.. Sci. Canada 24 (2002), 1–15.

[FGT] P. Flajolet, D. Gardy and L. Thimonier: *Birthday paradox, coupon collectors, caching algorithms and self-organizing search*, Discrete Applied Math. 39 (1992), 207–229.

[FG] J. Fulman and R. Guralnick: *Derangements in simple and primitive groups*, in "Groups, combinatorics & geometry" (Durham, 2001), 99–121, World Sci. Publ., River Edge, NJ, 2003.

[FH] W. Fulton and J. Harris: *Representation theory. A first course*, Grad. Texts in Math. 129, Springer 1991.

[G] P.X. Gallagher: *The large sieve and probabilistic Galois theory*, in Proc. Sympos. Pure Math., Vol. XXIV, Amer. Math. Soc. (1973), 91–101.

[GAP] The GAP Group: *GAP – Groups, Algorithms, and Programming, Version 4.4.9*, 2007, `www.gap-system.org`

[Gi] M. Giudici: *Maximal subgroups of almost simple groups with socle $PSL(2, q)$*, `arXiv:math/07033685`.

[Go] V. Goncharov: *Du domaine d'analyse combinatoire*, Bull. Acad. Sci. USSR Ser. Mat. (Izv. Akad. Nauk SSSR) 8 (1944), 3–48; Amer. Math. Soc. Transl. (2) 19 (1962), 1–46.

[IK] H. Iwaniec and E. Kowalski: *Analytic number theory*, A.M.S. Coll. Publ. 53, 2004.

[JKZ] F. Jouve, E. Kowalski and D. Zywina: *An explicit integral polynomial whose splitting field has Galois group $W(E_8)$*, Journal de Théorie des Nombres de Bordeaux 20 (2008), 761–782.

[JKZ2] F. Jouve, E. Kowalski and D. Zywina: *Splitting fields of characteristic polynomials of random elements in arithmetic groups*, preprint (2010).

[KL] W.M. Kantor and A. Lubotzky: *The probability of generating a finite classical group*, Geom. Dedicata 36 (1990), 67–87.

[K1] E. Kowalski: *The large sieve, monodromy and zeta functions of curves*, J. reine angew. Math. 601 (2006), 29–69.

[K2] E. Kowalski: *The large sieve and its applications*, Cambridge Tract in Math. 175, Cambridge Univ. Press 2008.

[LMO] J. C. Lagarias, H. L. Montgomery2 and A. M. Odlyzko: *A bound for the least prime ideal in the Chebotarev Density Theorem*, Inventiones math. 54 (1979), 271–296.

[LS] S.P. Lloyd and L.A. Shepp: *Ordered cycle lengths in a random permutation*, Trans. Amer. Math. Soc. 121 (1966), 340–357.

[LP] T. Łuczak and L. Pyber: *On random generation of the symmetric group*, Combin. Probab. Comput. 2 (1993), 505–512.

[N] I. Niven: *Average of exponents in factoring integers*, Proc. Amer. Math. Soc. 22 (1969), 356–360.

[P] C. Pomerance: *The expected number of random elements to generate a finite abelian group*, Period. Math. Hungar. 43 (2001) 191–198.

[Ro] J.S. Rose: *A course on group theory*, Dover 1994.

[R] J. Rotman: *An introduction to the theory of groups*, 4th edition, GTM 148, Springer 1995.

[S1] J-P. Serre: *On a theorem of Jordan*, Math. Medley 29 (2002), 3–18; also in Bull. A.M.S 40 (2003), 429–440.

[S2] J-P. Serre: *Abelian ℓ-adic representations and elliptic curves*, Res. Notes Math., A.K. Peters, 1998.
[S3] J-P. Serre: *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Invent. math. 15 (1972), 259–331.

ETH Zürich – D-MATH, Rämistrasse 101, 8092 Zürich, Switzerland
E-mail address: kowalski@math.ethz.ch

Department of Mathematics, University of Pennsylvania, Philadelphia, PA 19104-6395, USA
E-mail address: zywina@math.upenn.edu

Table 1. Chebotarev invariants of A_n

n	Order	$c(A_n)$	$c_2(A_n)$
2	1	1.000000...	1.000000...
3	3	1.500000...	3.000000...
4	12	4.409091...	29.71074...
5	60	4.136364...	22.64463...
6	360	4.439574...	25.49003...
7	2520	4.782001...	29.98671...
8	20160	4.939097...	31.98434...
9	181440	4.637463...	26.35009...
10	1814400	4.145282...	21.73709...
11	19958400	4.092974...	21.08692...
12	239500800	4.444074...	24.14188...
13	3113510400	4.016324...	20.51475...
14	43589145600	4.212753...	22.16514...
15	653837184000	4.289698...	22.51291...
16	10461394944000	4.239141...	22.21416...
17	177843714048000	4.089704...	21.12890...
18	3201186852864000	4.248133...	22.38035...
19	60822550204416000	4.072274...	21.08656...
20	1216451004088320000	4.229094...	22.20516...
21	25545471085854720000	4.238026...	22.19523...
22	562000363888803840000	4.240513...	22.33370...
23	12926008369442488320000	4.131077...	21.54514...
24	310224200866619719680000	4.282667...	22.58460...

Table 2. "Partial" Chebotarev invariants of A_n

n	Order	$\mathbf{E}(\tilde{\tau}'_n)$	$\mathbf{E}((\tilde{\tau}'_n)^2)$
3	3	1.500000...	3.000000...
4	12	2.123377...	5.874009...
5	60	2.500000...	10.00000...
6	360	2.649424...	9.187574...
7	2520	3.243247...	16.47701...
8	20160	2.812743...	10.71084...
9	181440	3.133704...	13.97383...
10	1814400	3.115450...	13.08967...
11	19958400	3.399573...	15.88920...
12	239500800	3.225496...	14.16483...
13	3113510400	3.402011...	15.56383...
14	43589145600	3.357361...	15.13742...
15	653837184000	3.504050...	16.37350...
16	10461394944000	3.385358...	15.32752...
17	177843714048000	3.544719...	16.55867...
18	3201186852864000	3.497980...	16.21775...
19	60822550204416000	3.625919...	17.22183...
20	1216451004088320000	3.530703...	16.46076...

Table 3. Chebotarev invariants of $\mathfrak{S}_n$

n	Order	$c(\mathfrak{S}_n)$	$c_2(\mathfrak{S}_n)$
2	2	2.000000...	6.000000...
3	6	3.800000...	19.32000...
4	24	4.498380...	25.91538...
5	120	4.331526...	23.50351...
6	720	5.610738...	37.63260...
7	5040	4.115230...	21.20184...
8	40320	4.626289...	25.71722...
9	362880	4.250355...	22.49197...
10	3628800	4.624666...	25.76898...
11	39916800	4.173683...	21.86294...
12	479001600	4.583705...	25.11338...
13	6227020800	4.213748...	22.21319...
14	87178291200	4.508042...	24.57963...
15	1307674368000	4.365718...	23.39257...
16	20922789888000	4.461633...	24.12713...
17	355687428096000	4.282141...	22.79488...
18	6402373705728000	4.531784...	24.67680...
19	121645100408832000	4.308469...	23.01145...
20	2432902008176640000	4.497047...	24.37207...
21	51090942171709440000	4.391209...	23.61488...
22	1124000727777607680000	4.477492...	24.29632...
23	25852016738884976640000	4.352364...	23.37533...
24	620448401733239439360000	4.523388...	24.57409...

Table 4. Chebotarev invariants of transitive groups of degree 17

Name	Order	$c(G)$	$c_2(G)$
$\mathbf{Z}/17\mathbf{Z}$	17	1.062500...	1.195312...
$C_8 \subset H_{17}$	34	3.094697...	11.81350...
$C_4 \subset H_{17}$	68	4.890000...	35.53580...
$C_2 \subset H_{17}$	136	8.880953...	138.3764...
H_{17}	272	17.21053...	562.3851...
$PSL(2, \mathbf{F}_{16})$	4080	3.200912...	12.73727...
7	8160	4.055261...	20.84364...
8	16320	4.067118...	20.58582...
A_{17}	177843714048000	4.089704...	21.12890...
$\mathfrak{S}_{17}$	355687428096000	4.282141...	22.79488...

TABLE 5. Chebotarev invariants of $PSL(3, \mathbf{F}_p)$

p	Order	$c(PSL(3, \mathbf{F}_p))$	$c_2(PSL(3, \mathbf{F}_p))$
2	168	4.653153...	29.48762...
3	5616	3.845890...	20.67132...
5	372000	3.629464...	18.36114...
7	1876896	3.661481...	18.91957...
11	212427600	3.527819...	17.29354...
13	270178272	3.546344...	17.55063...
17	6950204928	3.511708...	17.12456...
19	5644682640	3.521753...	17.25893...
23	78156525216	3.506462...	17.06878...
29	499631102880	3.504076...	17.04348...
31	283991644800	3.508213...	17.09800...
37	1169948144736	3.505795...	17.06906...
41	7980059337600	3.502051...	17.02191...

TABLE 6. Chebotarev invariants of $PSL(4, \mathbf{F}_p)$

p	Order	$c(PSL(4, \mathbf{F}_p))$	$c_2(PSL(4, \mathbf{F}_p))$
2	20160	4.939097...	31.98434...
3	6065280	4.191257...	23.35082...
5	7254000000	3.768197...	18.89633...
7	2317591180800	3.613602...	17.31973...
11	2069665112592000	3.530797...	16.44109...
13	12714519233969280	3.513963...	16.24990...

TABLE 7. Chebotarev invariants of $PSL(n, \mathbf{F}_2)$

n	Order	$c(PSL(n, \mathbf{F}_2))$	$c_2(PSL(n, \mathbf{F}_2))$
2	6	3.800000...	19.32000...
3	168	4.653153...	29.48762...
4	20160	4.939097...	31.98434...
5	9999360	4.238182...	25.64374...
6	20158709760	4.456089...	27.20052...
7	163849992929280	4.335957...	26.54874...
8	5348063769211699200	4.465723...	27.53266...
9	699612310033197642547200	4.460433...	27.64706...

Table 8. Chebotarev invariants of $PSL(n, \mathbf{F}_3)$

n	Order	$c(PSL(n, \mathbf{F}_3))$	$c_2(PSL(n, \mathbf{F}_3))$
2	12	4.409091...	29.71074...
3	5616	3.845890...	20.67132...
4	6065280	4.191257...	23.35082...
5	237783237120	3.949889...	21.81110...
6	21032402889738240	4.123378...	23.06449...
7	67034222101339041669120	4.066340...	22.81370...

Table 9. Chebotarev invariants of $PSL(n, \mathbf{F}_4)$

n	Order	$c(PSL(n, \mathbf{F}_4))$	$c_2(PSL(n, \mathbf{F}_4))$
2	60	4.136364...	22.64463...
3	20160	4.399979...	26.39681...
4	987033600	3.770618...	19.19928...
5	258492255436800	3.838194...	20.33428...
6	361310134959341568000	4.002927...	21.57223...

Table 10. Chebotarev invariants of $PSL(2, \mathbf{F}_{2^n})$

2^n	Order	$c(PSL(2, \mathbf{F}_{2^n}))$	$c_2(PSL(2, \mathbf{F}_{2^n}))$
2	6	3.800000...	19.32000...
4	60	4.136364...	22.64463...
8	504	3.437879...	14.95188...
16	4080	3.200912...	12.73727...
32	32736	3.096876...	11.82191...
64	262080	3.048732...	11.40623...
128	2097024	3.023623...	11.19773...
256	16776960	3.011765...	11.09826...
512	134217216	3.005965...	11.04945...

Table 11. Chebotarev invariants of $PSL(3, \mathbf{F}_{2^n})$

2^n	Order	$c(PSL(3, \mathbf{F}_{2^n}))$	$c_2(PSL(3, \mathbf{F}_{2^n}))$
2	168	4.653153...	29.48762...
4	20160	4.399979...	26.39681...
8	16482816	3.551417...	17.54363...
16	1425715200	3.549690...	17.47208...
32	1098404364288	3.503357...	17.03581...

Table 12. Chebotarev invariants of the Borel subgroup of $SL(3, \mathbf{F}_p)$

p	Order	$c(B_3(\mathbf{F}_p))$	$c_2(B_3(\mathbf{F}_p))$
2	8	3.333333...	13.55556...
3	108	5.074442...	31.76009...
5	2000	7.686557...	81.14365...
7	12348	10.07528...	150.8724...
11	133100	16.38777...	402.7223...
13	316368	18.85106...	551.0363...
17	1257728	25.31072...	978.0196...
19	2222316	27.79352...	1204.483...
23	5888828	34.28491...	1805.763...
29	19120976	43.27249...	2885.634...
31	26811900	45.75644...	3268.081...
37	65646288	54.75057...	4678.007...
41	110273600	61.26132...	5801.515...
43	140250348	63.74680...	6339.956...

Table 13. Chebotarev invariants of $PSL(2, \mathbf{F}_p)$, $p \leqslant 150$

p	Order	$c(PSL(2, \mathbf{F}_p))$	$c_2(PSL(2, \mathbf{F}_p))$
2	6	3.800000...	19.32000...
3	12	4.409091...	29.71074...
5	60	4.136364...	22.64463...
7	168	4.653153...	29.48762...
11	660	3.981397...	20.76193...
13	1092	3.293965...	13.63659...
17	2448	3.264353...	13.20732...
19	3420	3.259202...	13.08533...
23	6072	3.136600...	12.18536...
29	12180	3.115633...	11.99619...
31	14880	3.111661...	11.92578...
37	25308	3.088522...	11.75723...
41	34440	3.098342...	11.78358...
43	39732	3.071689...	11.61064...
47	51888	3.065454...	11.55651...
53	74412	3.060208...	11.51103...
59	102660	3.051900...	11.43952...
61	113460	3.051897...	11.43943...
67	150348	3.045600...	11.38545...
71	178920	3.046777...	11.38343...
73	194472	3.042989...	11.36306...
79	246480	3.043013...	11.34889...
83	285852	3.036689...	11.30930...
89	352440	3.036100...	11.30056...
97	456288	3.031998...	11.26935...
101	515100	3.032463...	11.26755...
103	546312	3.030308...	11.25228...
107	612468	3.028370...	11.23855...
109	647460	3.029877...	11.24644...
113	721392	3.028016...	11.23330...
127	1024128	3.024393...	11.20309...
131	1123980	3.024148...	11.19945...
137	1285608	3.022889...	11.19063...
139	1342740	3.022686...	11.18747...
149	1653900	3.020586...	11.17269...

Table 14. Chebotarev invariants of some other groups

Name	Order	$c(G)$	$c_2(G)$
$W(G_2) = D_{12}$	12	$4.31\underline{515}\ldots = 717/165$	$23.45407\ldots$
$W(C_4)$	384	$4.864890\ldots$	$29.10488\ldots$
$\mathbf{Z}/6\mathbf{Z} \times PSL(2,\mathbf{F}_7)$	1008	$4.890689\ldots$	$31.42829\ldots$
$W(F_4)$	1152	$5.417656\ldots$	$35.12470\ldots$
$GL(2,\mathbf{F}_7)$	2016	$3.767768\ldots$	$17.29394\ldots$
$A_5 \times A_5$	3600	$5.374156\ldots$	$35.41628\ldots$
$W(C_5)$	3840	$4.863533\ldots$	$28.13517\ldots$
M_{11}	7920	$4.850698\ldots$	$29.72918\ldots$
$GL(3,\mathbf{F}_3)$	11232	$4.110394\ldots$	$22.77077\ldots$
$G_2(\mathbf{F}_2)$	12096	$5.246204\ldots$	$34.24515\ldots$
$SL(4,\mathbf{F}_2) = A_8$	20160	$4.939097\ldots$	$31.98434\ldots$
$Sz(8)$	29120	$3.101639\ldots$	$11.92233\ldots$
$W(C_6)$	46080	$5.792117\ldots$	$39.56093\ldots$
$W(E_6)$	51840	$4.470824\ldots$	$23.93050\ldots$
$Sp(4,\mathbf{F}_3)$	51840	$4.401859\ldots$	$24.03143\ldots$
$PGL(3,\mathbf{F}_4)$	60480	$3.763384\ldots$	$19.49865\ldots$
M_{12}	95040	$4.953188\ldots$	$29.53947\ldots$
J_1	175560	$3.423739\ldots$	$14.76364\ldots$
$GL(2,\mathbf{F}_{23})$	267168	$3.448700\ldots$	$14.55080\ldots$
M_{22}	443520	$4.164445\ldots$	$22.70981\ldots$
J_2	604800	$4.031298\ldots$	$19.07590\ldots$
$W(C_7)$	645120	$4.632612\ldots$	$25.54504\ldots$
$PSp(6,\mathbf{F}_2)$	1451520	$5.270439\ldots$	$34.84139\ldots$
$W(E_7)$	2903040	$5.398250\ldots$	$36.04850\ldots$
$G_2(\mathbf{F}_3)$	4245696	$4.511630\ldots$	$24.06106\ldots$
M_{23}	10200960	$4.030011\ldots$	$20.98580\ldots$
$W(C_8)$	10321920	$4.928996\ldots$	$28.53067\ldots$
$Sz(32)$	32537600	$2.755449\ldots$	$9.107751\ldots$
HS	44352000	$4.484432\ldots$	$25.68549\ldots$
J_3	50232960	$4.304616\ldots$	$23.42082\ldots$
$W(C_9)$	185794560	$4.716359\ldots$	$26.41344\ldots$
M_{24}	244823040	$4.967107\ldots$	$29.84845\ldots$
$Sp(4,\mathbf{F}_7)$	276595200	$3.501127\ldots$	$14.83811\ldots$
$\Omega^+(4,\mathbf{F}_{31})$	442828800	$3.829841\ldots$	$17.60003\ldots$
$\Omega^-(4,\mathbf{F}_{31})$	443751360	$3.003133\ldots$	$11.02613\ldots$
$W(E_8)$	696729600	$4.194248\ldots$	$20.79438\ldots$
McL	898128000	$4.561453\ldots$	$27.45649\ldots$
$Sp(4,\mathbf{F}_9)$	3443212800	$3.409108\ldots$	$14.04475\ldots$
$G_2(\mathbf{F}_5)$	5859000000	$3.855868\ldots$	$18.68766\ldots$
$Sp(6,\mathbf{F}_3)$	9170703360	$3.871692\ldots$	$18.90072\ldots$
$\Omega(5,\mathbf{F}_{31})$	409387254681600	$3.277801\ldots$	$12.90986\ldots$
Rub	43252003274489856000	$5.668645\ldots$	$36.78701\ldots$