R. van Bommel

Using the Chebotarev density theorem to calculate the size of Galois groups

Bachelor's thesis, 20 July 2012

Supervisor: dr. L. Taelman

Mathematisch Instituut, Universiteit Leiden

Contents

Introduction

Let $f \in \mathbb{Z}[X]$ be monic separable of degree n. Let L be a splitting field of f over $\mathbb{Q}$ and let G be the Galois group of $L/\mathbb{Q}$. For a prime number p consider the factorization of $f \bmod p$. Consider the list of the degrees of the irreducible factors as a partition of n and call this partition the factorization type of $f \bmod p$. Furthermore, for each element g of G consider the cycle type of g, induced by the action of G on the set of roots of f, also as a partition of n.

Let C be a partition of n. The Chebotarev density theorem states that the fraction of elements of G having cycle type C equals the density of prime numbers p for which $f \bmod p$ has factorization type C. In particular, the latter density exists and is rational. In the first chapter of this thesis, a precise statement of the Chebotarev density theorem will be made and some background will be given.

In particular, the theorem implies that the fraction of primes for which f mod p totally splits into linear factors is equal to the fraction of elements that have cycle type $(1, 1, \ldots, 1)$. As only the identity has that cycle type, this fraction equals $\frac{1}{|G|}$. This suggests an algorithm to find the size of G. Namely, count the primes $p < x$ for which $f \bmod p$ splits into linear factors. As $x \to \infty$ the fraction of primes having this property will tend to $\frac{1}{|G|}$. In principle we can use an effective version of the Chebotarev density theorem to turn this into a correct but very slow algorithm.

Note, in the typical case the group G is isomorphic to S_n (see [17]). In this case there are very few primes for which $f \bmod p$ splits into linear factors. We would expect x to need to be at least $n!$ to hope to be able to distinguish between the size of S_n and the size of A_n. This makes this algorithm very inefficient to use for very many of the polynomials.

F. Rodriguez Villegas (personal communication, 27 March 2012) came up with the idea of using representation theory to improve upon this algorithm. In this thesis we will explain this idea, and discuss two algorithms based on it.

In the second chapter all necessary representation theory will be treated. The third chapter will contain a description of this improved algorithm together with a correctness proof and runtime analysis. In the fourth and final chapter a probabilistic model will be used to quantify heuristically how much better the improved algorithm is in comparison to the original algorithm.

1 Chebotarev density theorem

1.1 Setting

First we will describe the setting in which the Chebotarev density theorem will be stated. The following definitions and notations will be used throughout the whole chapter.

For a group H denote by $C(H)$ its set of conjugacy classes. If $h \in H$ is an element, then $C(h)$ is the conjugacy class of h.

Let $f \in \mathbb{Z}[X]$ be a monic polynomial of degree n and let $L/\mathbb{Q}$ be a splitting field of f. Let G be the Galois group of $L/\mathbb{Q}$. Let $\overline{\mathbb{Q}}$ be an algebraic closure of $\mathbb{Q}$. Furthermore, assume that f has no multiple roots in $\overline{\mathbb{Q}}$, i.e. assume that the discriminant $\Delta(f)$ is non-zero.

We will define a map $\iota\colon C(G) \to C(S_n)$ as follows. Fix a bijection between the set of roots of f in $\overline{\mathbb{Q}}$ and $\{1, \dots, n\}$. Consider G as subgroup of S_n via this bijection and let $\iota\colon C(G) \to C(S_n)$ be the map induced by the inclusion $G \subset S_n$. This map does not depend on the chosen bijection. Furthermore, this map generally is not injective or surjective.

We recall some algebraic number theory.

Definition 1.1.1 (Ring of integers)**.** The *ring of integers* of L is

$$\mathcal{O}_L = \{x \in L : \text{there is a } g \in \mathbb{Z}[X] \text{ such that } g \text{ is monic and } g(x) = 0\} \subset L.$$

Remark 1.1.2. Since f is monic, the roots $\alpha_1, \dots, \alpha_n \in L$ of f are elements of $\mathcal{O}_L$.

The following proposition states that the ring of integers is indeed a ring and it also states some useful properties of $\mathcal{O}_L$.

Proposition 1.1.3. *The ring of integers $\mathcal{O}_L$ is a subring of L. It is a Dedekind domain. In particular, every non-zero ideal in $\mathcal{O}_L$ factors uniquely into prime ideals.*

Proof. We give references for the assertions of the proposition. The fact that $\mathcal{O}_L$ is a ring follows from Proposition 5 of [9, I.§2] applied on the ring $\mathbb{Z} \subset L$. By Theorem 1 of [9, I.§2] $\mathcal{O}_K$ is finitely generated as $\mathbb{Z}$-module and hence it is a Noetherian ring. By Corollary 5.5 of [1, ch. 4] $\mathcal{O}_K$ is integrally closed. By Proposition 10 of [9, I.§3] every non-zero prime ideal of $\mathcal{O}_K$ is maximal.

Hence $\mathcal{O}_L$ is a Dedekind domain and Theorem 2 of [9, I.§6] implies that every non-zero ideal in $\mathcal{O}_L$ factors uniquely into prime ideals. $\square$

1.2 Frobenius substitution

Let $p \in \mathbb{Z}$ be a prime number. Let $\overline{\mathbb{F}}_p$ be an algebraic closure of $\mathbb{F}_p$. The following definition of a place of L over p is equivalent to the definition given in [15] and it is not equivalent to the standard definition of a place of a number field.

Definition 1.2.1 (Place of L over p)**.** A *place ψ of L over p* is a morphism $\psi\colon \mathcal{O}_L \to \overline{\mathbb{F}}_p$ of rings.

Proposition 1.2.2. *A place of L over p exists.*

Proof. Let $\mathfrak{B} \subset \mathcal{O}_L$ be some maximal ideal containing p. Let $q\colon \mathcal{O}_L \to \mathcal{O}_L/\mathfrak{B}$ be the natural quotient map. Then $\mathcal{O}_L/\mathfrak{B}$ is a field of characteristic p. Furthermore $\mathcal{O}_L/\mathfrak{B}$ is an algebraic extension of $\mathbb{F}_p$, since L is an algebraic extension of $\mathbb{Q}$. Hence there exists an injection $i\colon \mathcal{O}_L/\mathfrak{B} \to \overline{\mathbb{F}}_p$. Then $i \circ q$ is a place of L over p. $\square$

Proposition 1.2.3. *Let ψ be a place of L over p and let $\theta \in \mathrm{Aut}(\overline{\mathbb{F}}_p)$ and $\tau \in G$ be automorphisms of $\overline{\mathbb{F}}_p$ respectively L. Then $\theta \circ \psi \circ \tau$ is a place of L over p.*

Proof. This follows immediately from the fact that compositions of ring morphisms are ring morphisms. $\square$

Lemma 1.2.4. *Suppose that ψ and ψ' are places of L over p. Then there exists a $\tau \in G$ such that $\psi' = \psi \circ \tau$. Furthermore, if $p \nmid \Delta(f)$ then τ is unique.*

Proof. The existence of τ follows from Corollary 1 of [9, I.§5]. Suppose that $p \nmid \Delta(f)$. Since $p \nmid \Delta(f)$ and f is monic, $\overline{f} \in \mathbb{F}_p[X]$ has n distinct roots in $\overline{\mathbb{F}}_p$. In particular if $\alpha_1, \ldots, \alpha_n \in \mathcal{O}_L$ are the roots of f then $\psi(\alpha_1), \ldots, \psi(\alpha_n) \in \overline{\mathbb{F}}_p$ are distinct. If $\tau, \tau' \in G$ satisfy $\psi' = \psi \circ \tau = \psi \circ \tau'$, then $\psi = \psi \circ \tau(\tau')^{-1}$ and hence $\tau(\tau')^{-1}$ fixes $\alpha_1, \ldots, \alpha_n$. Therefore, $\tau(\tau')^{-1} = \mathrm{id}$ and hence $\tau = \tau'$. This proves the uniqueness of τ. $\square$

Suppose that $p \nmid \Delta(f)$. Let ψ be a place of L over p, which exists because of Proposition 1.2.2. Let $F\colon \overline{\mathbb{F}}_p \to \overline{\mathbb{F}}_p\colon x \mapsto x^p$ be the Frobenius automorphism. By Proposition 1.2.3 the map $F \circ \psi$ is also a place of L over p and by Lemma 1.2.4 there exists a unique element $\tau_\psi \in G$ such that $F \circ \psi = \psi \circ \tau_\psi$. If we chose the place ψ' of L over p instead of ψ, then $\psi' = \psi \circ \sigma$ for some unique $\sigma \in G$. Hence $F \circ \psi' = F \circ \psi \circ \sigma = \psi \circ \tau_\psi \sigma = \psi' \circ \sigma^{-1} \tau_\psi \sigma$, i.e. $\tau_{\psi'} = \sigma^{-1}\tau_\psi\sigma$. Therefore, the following is well-defined.

Definition 1.2.5 (Frobenius substitution)**.** Let $p \in \mathbb{Z}$ be a prime such that $p \nmid \Delta(f)$. Then the *Frobenius substitution of* p is $F_p := C(\tau_\psi) \in C(G)$, where ψ is some place of L over p.

Example 1.2.6. Take $f = X^3 - 2$. Then $L = \mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ has Galois group $G = S_3$ over $\mathbb{Q}$. Furthermore, $\mathcal{B} = (5, \sqrt[3]{2} - 3)\mathcal{O}_L$ is prime and $\mathcal{O}_L/\mathcal{B} \cong \mathbb{F}_{25}$. The roots of f in $\mathcal{O}_L/\mathcal{B}$ are $\overline{3}, \overline{3\zeta_3}, \overline{3\zeta_3^2}$. Then the Frobenius automorphism maps $\overline{3\zeta_3^i}$ to $\overline{3\zeta_3^{2i}}$ for $i = 0, 1, 2$. Let $\sigma \in G$ be the element of the Galois group for which $\sigma(\zeta_3^i\sqrt[3]{2}) = \zeta_3^{2i}\sqrt[3]{2}$ for $i = 0, 1, 2$, then $F_5 = C(\sigma) = C((12))$.

Definition 1.2.7 (Factorization type)**.** Let p be a prime. Then the *factorization type of f modulo p* is the unordered partition $(n_1, \ldots, n_t)$ of n consisting of the degrees of the irreducible factors of $\overline{f} \in \mathbb{F}_p[X]$. Denote by $C(f, p) \in C(S_n)$ the class consisting of the permutations that have cycle type $(n_1, \ldots, n_t)$.

The following useful lemma links the Frobenius substitution with the factorization of $\overline{f} \in \mathbb{F}_p[X]$.

Lemma 1.2.8. *For all primes p such that $p \nmid \Delta(f)$ we have $\iota(F_p) = C(f, p)$.*

Proof. Notice that by definition F_p permutes the roots of f in L in the same way as F permutes the roots of $f \bmod p$ in $\overline{\mathbb{F}}_p$. It is a known fact (see for example [14, §22]) that F permutes the roots of each irreducible factor cyclically. The statement of the lemma then follows immediately. □

1.3 Densities

Let $\mathcal{P} \subset \mathbb{Z}$ be the set of prime numbers. There are different notions of the density of subsets of $\mathcal{P}$.

Definition 1.3.1 (Natural density)**.** Let $A \subset \mathcal{P}$ be a subset and suppose that the limit

$$d(A) := \lim_{x \to \infty} \frac{|\{p \in A : p \leqslant x\}|}{|\{p \in \mathcal{P} : p \leqslant x\}|}$$

exists. Then $d(A)$ is called the *natural density* of A.

The natural density is perhaps the most natural notion of density. The following notion of density, the Dirichlet density, is much harder to come up with and it might feel unnatural. However, the Chebotarev density theorem and many other density theorems in number theory were originally proven for the Dirichlet density.

Definition 1.3.2 (Dirichlet density)**.** Let $A \subset \mathcal{P}$ be a subset and suppose that the limit

$$\delta(A) := \lim_{s \downarrow 1} \frac{\sum_{p \in A} \frac{1}{p^s}}{\sum_{p \in \mathcal{P}} \frac{1}{p^s}}$$

exists. Then $\delta(A)$ is called the analytic or *Dirichlet density* of A.

The natural density and the Dirichlet density are related in the following way.

Lemma 1.3.3. *Let $A \subset \mathcal{P}$ be a subset and suppose that the natural density $d(A)$ of A exists. Then the Dirichlet density of A exists and $\delta(A) = d(A)$.*

Proof. This follows from Theorem 2 and Theorem 3 of [16, p.272–274]. □

However, the converse is not true. There are subsets of $\mathcal{P}$ which have a Dirichlet density and do not have a natural density. One of them is the following subset.

Example 1.3.4. The subset $\{p \in \mathcal{P} : \text{the first digit of } p \text{ is a } 1\}$ has Dirichlet density $\frac{\log 2}{\log 10}$, but it does not have a natural density, see [3].

We derive some useful results for the Dirichlet density.

Proposition 1.3.5. *Let $A, B \subset \mathcal{P}$ be such that $A \cap B = \emptyset$. Suppose that two of the densities $\delta(A), \delta(B), \delta(A \cup B)$ exist, then the third one exists and they satisfy:*

$$\delta(A) + \delta(B) = \delta(A \cup B).$$

In particular, if $C \subset D \subset \mathcal{P}$ are subsets and $\delta(C)$ and $\delta(D)$ exist, then $\delta(C) \leqslant \delta(D)$.

Proof. For every $s > 1$ we have

$$\frac{\sum_{p\in A}\frac{1}{p^s}}{\sum_{p\in\mathcal{P}}\frac{1}{p^s}} + \frac{\sum_{p\in B}\frac{1}{p^s}}{\sum_{p\in\mathcal{P}}\frac{1}{p^s}} = \frac{\sum_{p\in A\cup B}\frac{1}{p^s}}{\sum_{p\in\mathcal{P}}\frac{1}{p^s}}.$$

By using the fact that addition and subtraction are continuous the result follows for the limit $s \downarrow 1$. In particular, $\delta(D) = \delta(C) + \delta(D \setminus C) \geqslant \delta(C)$, because densities are clearly non-negative. □

Proposition 1.3.6. *Let $A \subset \mathcal{P}$ be finite. Then $\delta(A) = 0$.*

Proof. Notice that $\lim_{s\downarrow 1}\sum_{p\in A}\frac{1}{p^s} < \infty$ and $\lim_{s\downarrow 1}\sum_{p\in\mathcal{P}}\frac{1}{p^s} = \infty$. The result now follows immediately. □

Corollary 1.3.7. *Let $A, B \subset \mathcal{P}$ such that $A\backslash B$ and $B\backslash A$ are finite. Suppose that $\delta(A)$ exists. Then $\delta(B)$ exists and $\delta(A) = \delta(B)$.*

Proof. By applying Propositions 1.3.5 and 1.3.6 we find

$$\delta(A) = \delta(A) + \delta(B \setminus A) = \delta(A \cup B) = \delta(B) + \delta(A \setminus B) = \delta(B).$$

□

Remark 1.3.8. Propositions 1.3.5, 1.3.6 and Corollary 1.3.7 are also true if the Dirichlet density is replaced with the natural density. The proofs are analogous to the proofs for the Dirichlet density and will not be given in detail.

1.4 Chebotarev density theorem

Theorem 1.4.1 (Chebotarev density theorem)**.** *The following holds for every conjugacy class $C \in C(G)$:*

$$\delta\left(\{p \in \mathcal{P} : p \nmid \Delta(f) \text{ and } F_p = C\}\right) = \frac{|C|}{|G|}.$$

Proof. See [15] or [10, p.545]. □

The theorem is also true if the Dirichlet density is replaced by the natural density. However, the result was first proven by Chebotarev for the Dirichlet density in [4]. The following famous theorems are special cases of the Chebotarev density theorem.

Corollary 1.4.2 (Dirichlet's theorem)**.** *Let $n \in \mathbb{Z}$ be a positive integer. Then for each $a \in \mathbb{Z}$ with $\gcd(a, n) = 1$ the following holds:*

$$\delta\left(\{p \in \mathcal{P} : p \equiv a \bmod n\}\right) = \frac{1}{\varphi(n)}.$$

Proof. Take $f = X^n - 1$. Then we get $L = \mathbb{Q}(\zeta_n)$ and $\rho\colon (\mathbb{Z}/n\mathbb{Z})^* \to G\colon (a \bmod n) \mapsto (\zeta_n \mapsto \zeta_n^a)$ is an isomorphism. Notice that we have $F_p = C(\rho(p \bmod n))$. Furthermore, note that as G is abelian, conjugacy classes consist of 1 element. Also, note that due to Corollary 1.3.7 it does not matter if we consider or exclude the finitely many primes p such that $p \mid \Delta(f)$. Therefore, the Chebotarev density theorem (1.4.1) immediately yields the desired result. □

Corollary 1.4.3 (Frobenius' theorem)**.** *The following holds for all $C \in \mathcal{C}(S_n)$:*

$$\delta(\{p \in \mathcal{P} : C(f, p) = C\}) = \frac{|\{g \in G : \iota(g) \in C\}|}{|G|}.$$

Proof. Notice that $\{g \in G : \iota(g) \in C\} = \iota^{-1}(C) \subset G$ is a union of conjugacy classes. Then use the Chebotarev density theorem (1.4.1) for these conjugacy classes. Also, note that due to Corollary 1.3.7 it does not matter if we consider or exclude the finitely many primes p such that $p \mid \Delta(f)$. Lemma 1.2.8 then finishes the proof of the statement. □

2 Representation theory

2.1 Definitions

Let G be a finite group. Denote by $C(G)$ its set of conjugacy classes. If $s \in G$ is an element, then $C(s) \in C(G)$ is the conjugacy class of s.

Definition 2.1.1 (Group algebra)**.** The *group algebra* $\mathbb{C}[G]$ is the $\mathbb{C}$-algebra whose elements are formal sums $\sum_{s\in G} c_s s$ where $c_s \in \mathbb{C}$ for all $s \in G$. If $a = \sum_{s\in G} a_s s$ and $b = \sum_{s\in G} b_s s$ are two elements, then their sum is $a + b := \sum_{s\in G}(a_s + b_s)s$ and their product is

$$a \cdot b := \sum_{s,t\in G} a_s b_t (st).$$

Definition 2.1.2 (Representation)**.** A *representation V of G* is a (left) $\mathbb{C}[G]$-module that is finite-dimensional as $\mathbb{C}$-vector space. A morphism of representations is a morphism of $\mathbb{C}[G]$-modules.

Remark 2.1.3. A representation V gives rise to the morphism $\rho_V : G \to \mathrm{Aut}_{\mathbb{C}}(V) : s \mapsto (v \mapsto s \cdot v)$. Conversely, if V is a finite dimensional $\mathbb{C}$-vector space and $\rho : G \to \mathrm{Aut}_{\mathbb{C}}(V)$ a morphism, then V together with ρ defines a representation by $s \cdot v = \rho(s)(v)$ for all $s \in G$ and $v \in V$.

Examples 2.1.4.

1. The trivial representation is the $\mathbb{C}[G]$-module $T = \mathbb{C}$ where G acts trivially, i.e. $s \cdot v = v$ for all $v \in T$ and $s \in G$. This representation is sometimes also denoted by $\mathbb{C}$.
2. Let $G = S_n$. The sign representation is the $\mathbb{C}[G]$-module $S = \mathbb{C}$ where G acts via the sign morphism, i.e. $s \cdot v = \mathrm{sgn}(s) \cdot v$ for all $v \in T$ and $s \in G$.
3. Let $G = S_n$. Let W be the $(n-1)$-dimensional subspace of $\mathbb{C}^n$ of vectors whose sum of coordinates is zero. Let G act on W by permuting the coordinates of the vectors, i.e. $s \cdot (v_1, \ldots, v_n) = (v_{s^{-1}(1)}, \ldots, v_{s^{-1}(n)})$. This representation W is called the standard representation of S_n.

Representations can be restricted to a subgroup or induced to a larger group.

Definition 2.1.5 (Restricted representation)**.** Let $H \subset G$ be a subgroup and let V be a representation of G. Then the *restricted representation* $V|_H$ or $\mathrm{Res}^G_H V$ is V where the action of $\mathbb{C}[G]$ is restricted to $\mathbb{C}[H]$.

Definition 2.1.6 (Induced representation)**.** Let $H \subset G$ be a subgroup and let V be a representation of H. Consider $\mathbb{C}[G]$ as right $\mathbb{C}[H]$-module by the multiplication in $\mathbb{C}[G]$. Then the *induced representation* $\mathrm{Ind}_H^G V$ is $\mathbb{C}[G] \otimes_{\mathbb{C}[H]} V$ where $\mathbb{C}[G]$ acts on the left factor, i.e. $s\cdot(t\otimes v) = (s\cdot t)\otimes v$ for all $s, t \in \mathbb{C}[G]$ and $v \in V$.

Example 2.1.7. Take $H = 1$ and $V = \mathbb{C}$. Then $\mathrm{Ind}_1^G V$ is $\mathbb{C}[G]$ where $\mathbb{C}[G]$ acts on the induced representation by left multiplication.

As in many other categories there are some useful ways to construct representations of G out of other representations of G. We discuss some of them.

Definition 2.1.8 (Direct sum of representations)**.** Let V and W be representation of G. Their *direct sum* $V \oplus W$ is their direct sum as $\mathbb{C}[G]$-modules, i.e. G acts as follows: $s \cdot (v, w) = (s \cdot v, s \cdot w)$ for all $s \in G, v \in V$ and $w \in W$.

Definition 2.1.9 (Tensor product of representations)**.** Let V and W be representations of G. The *tensor product* $V \otimes_{\mathbb{C}} W$ is a representation of G with G acting on both factors, i.e. $s \cdot (v \otimes w) = (s \cdot v) \otimes (s \cdot w)$ for all $s \in G, v \in V$ and $w \in W$.

Definition 2.1.10 (Dual representation)**.** Let V be a representation of G. The *dual representation* is $V^\dagger = \mathrm{Hom}_{\mathbb{C}}(V, \mathbb{C})$ where G acts as follows: $s \cdot f \colon v \mapsto f(s^{-1} \cdot v)$ for all $s \in G, f \in V^\dagger$ and $v \in V$.

Some character theory now follows.

Definition 2.1.11 (Character)**.** Let V be a representation of G. Then the *character of* V is the function

$$\chi_V \colon C(G) \to \mathbb{C} \colon C(s) \mapsto \mathrm{Tr}(\rho_V(s)).$$

Example 2.1.12. Let $G = S_3$. Then the characters of the representations defined in Examples 2.1.4 are as follows.

	$C(\mathrm{id})$	$C((12))$	$C((123))$
χ_T	1	1	1
χ_S	1	-1	1
χ_W	2	0	-1

Definition 2.1.13 (Irreducible representation)**.** A representation V is called *irreducible* if V has exactly two submodules: the zero module and V itself.

Definition 2.1.14 (Class function)**.** A *class function* is a function $C(G) \to \mathbb{C}$. The *space of class functions* is the inner product space $\mathbb{C}^{C(G)}$ equipped with the usual addition and scalar multiplication and the following inner product:

$$\langle \cdot, \cdot \rangle_G \colon \mathbb{C}^{C(G)} \times \mathbb{C}^{C(G)} \to \mathbb{C} \colon (\alpha, \beta) \mapsto \langle \alpha, \beta \rangle_G := \frac{1}{|G|} \sum_{s \in G} \alpha(C(s)) \overline{\beta(C(s))}.$$

Definition 2.1.15 (Irreducible character)**.** A class function χ is called an *irreducible character* if there exists an irreducible representation V such that $\chi = \chi_V$.

Definition 2.1.16 (Virtual character)**.** A class function $\chi \in \mathbb{C}^{C(G)}$ is called a *virtual character* if there exist representations V and W of G such that $\chi = \chi_V - \chi_W$.

2.2 Results

A lot is known about group representations. In this section some of the results of the representation theory of finite groups will be presented.

Characters behave well with respect to the direct sum, tensor product and dual of representations.

Proposition 2.2.1. *Suppose that V and W are representations of G. Then the characters of the representations $V \oplus W, V \otimes_{\mathbb{C}} W$ and $V^\dagger$ are as follows.*

$$\chi_{V \oplus W} = \chi_V + \chi_W \tag{1}$$
$$\chi_{V \otimes_{\mathbb{C}} W} = \chi_V \cdot \chi_W \tag{2}$$
$$\chi_{V^\dagger} = \overline{\chi_V} \tag{3}$$

Proof. Let $s \in G$ be arbitrary and suppose that $n = \dim_{\mathbb{C}}(V)$ and $m = \dim_{\mathbb{C}}(W)$. Furthermore, suppose that $\lambda_1, \ldots, \lambda_n$ and $\kappa_1, \ldots, \kappa_m$ are the eigenvalues of $\rho_V(s)$ respectively $\rho_W(s)$ (see Remark 2.1.3).
Then the eigenvalues of $\rho_{V \oplus W}(s)$ are equal to $\lambda_1, \ldots, \lambda_n, \kappa_1, \ldots, \kappa_m$ yielding $\chi_{V \oplus W}(C(s)) = (\chi_V + \chi_W)(C(s))$. Furthermore, the eigenvalues of $\rho_{V \otimes_{\mathbb{C}} W}(s)$ are equal to $\lambda_i \kappa_j$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$, yielding $\chi_{V \otimes_{\mathbb{C}} W}(C(s)) = \sum_{i=1}^n \sum_{j=1}^m \lambda_i \kappa_j = (\chi_V \cdot \chi_W)(C(s))$. Finally, the matrix $\rho_{V^\dagger}(s)$ is the conjugate transpose of $\rho_V(s)$. Hence its eigenvalues are $\overline{\lambda_1}, \ldots, \overline{\lambda_n}$ yielding $\chi_{V^\dagger}(C(s)) = \overline{\chi_V(C(s))}$. $\square$

The following lemma makes use of the fact that $\mathbb{C}$ has characteristic 0.

Lemma 2.2.2. *Let V and W be representations of G. Then*

$$\chi_V = \chi_W \quad \Longleftrightarrow \quad V \cong_{\mathbb{C}[G]} W.$$

Proof. This follows from Theorem 9.2, 9.6 and 10.7 of [5]. $\square$

Remark 2.2.3. The previous lemma shows that the irreducible characters are exactly the characters corresponding to irreducible representations.

Lemma 2.2.4. *The irreducible characters form an orthonormal basis of $\mathbb{C}^{C(G)}$.*

Proof. This follows from Theorem 10.17 of [5]. $\square$

Example 2.2.5. The characters in Examples 2.1.12 are in fact the irreducible characters of S_3 and one can check that these form an orthonormal basis of $\mathbb{C}^{C(S_3)}$.

Let $1_G = \chi_T$ be the character of the trivial representation; it is given by $1_G(C) = 1$ for all $C \in C(G)$. Then the following corollary of Lemma 2.2.4 will turn out to be very useful.

Corollary 2.2.6. *If χ is a virtual character, then $\langle \chi, 1_G \rangle_G \in \mathbb{Z}$.* $\square$

Frobenius reciprocity gives the relation between the characters of induced and restricted representations.

Lemma 2.2.7 (Frobenius reciprocity)**.** *Let $H \subset G$ be a subgroup, let V be a representation of G and let W be a representation of H. Then the following holds:*

$$\langle \chi_V, \chi_{\mathrm{Ind}_H^G W} \rangle_G = \langle \chi_{\mathrm{Res}_H^G V}, \chi_W \rangle_H.$$

Proof. See Theorem 8.1.3 of [13]. $\square$

3 The Rodriguez Villegas algorithm

In this chapter two algorithms will be given. Both algorithms are based on an idea of F. Rodriguez Villegas (personal communication, 27 March 2012) of using character theory and the Chebotarev density theorem to find the order of Galois groups.

3.1 Goal and notations

Let $f \in \mathbb{Z}[X]$ be a monic irreducible polynomial of degree n and let $L/\mathbb{Q}$ be a splitting field of f. Let G be the Galois group of $L/\mathbb{Q}$. Let $\overline{\mathbb{Q}}$ be an algebraic closure of $\mathbb{Q}$. Note that the assumption that f is irreducible automatically implies that f has no multiple roots in $\overline{\mathbb{Q}}$, i.e. that the discriminant $\Delta(f)$ is non-zero. Furthermore, it implies that G acts transitively on the set of n roots of f in $\overline{\mathbb{Q}}$.

Our goal is to compute $|G|$. We will make use of the following reformulation of the Chebotarev density theorem.

Theorem 3.1.1. *Let S be the set of primes not dividing $\Delta(f)$. Then, for all functions $\phi\colon C(G) \to \mathbb{C}$ we have*

$$\lim_{x\to\infty} \frac{\sum_{p\leqslant x} \phi(F_p)}{|\{p \in S : p \leqslant x\}|} = \sum_{C\in C(G)} \frac{|C|}{|G|}\phi(C) = \langle \phi, 1_G\rangle_G.$$

Proof. Apply the Chebotarev density theorem for the natural density as follows.

$$\begin{aligned}\lim_{x\to\infty} \frac{\sum_{p\leqslant x} \phi(F_p)}{|\{p \in S : p \leqslant x\}|} &= \sum_{C\in C(G)} \phi(C)\cdot \lim_{x\to\infty} \frac{|\{p \in S : F_p = C \text{ and } p \leqslant x\}|}{|\{p \in S : p \leqslant x\}|} \\ &= \sum_{C\in C(G)} \frac{|C|}{|G|}\phi(C) = \sum_{g\in G} \frac{1}{|G|}\phi(C(g)) = \langle \phi, 1_G\rangle_G. \quad \square\end{aligned}$$

For each subgroup $H \subset S_n$ define the class function

$$\delta_1^H : C(H) \to \mathbb{C} \colon C(h) \mapsto \begin{cases} 1 & \text{if } h = 1; \\ 0 & \text{otherwise.}\end{cases}$$

The next lemma explains why theorem 3.1.1 is useful to us.

Lemma 3.1.2. *Suppose that $H \subset S_n$ is a subgroup, then $\langle \delta_1^H, 1_H \rangle_H = \frac{1}{|H|}$.*

Proof. This is just a trivial calculation. $\square$

For a subgroup $H \subset S_n$ let $\iota_H \colon C(H) \to C(S_n)$ be the map induced by the inclusion (see page 5). Note that G acts transitively on the set of n roots of f. Hence, by fixing a bijection between $\{1, \dots, n\}$ and the set of roots G can be seen as subgroup of S_n. As already seen on page 5 the map ι_G does not depend on the choice of this bijection.

3.2 Precalculation

To compute the order of the Galois group of a polynomial f of degree n, the algorithm will make use of a list of transitive subgroups of S_n, up to conjugacy in S_n. This list of transitive subgroups is known for $n \leqslant 31$ and can be found, for example, by using Magma (see [2]).

Let k be the number of conjugacy classes of S_n and let $\mathbb{C}^{C(G)}$ be the class function space with its inner product as defined in Definition 2.1.14. The algorithm needs an orthonormal basis $\psi_1, \dots, \psi_k \colon C(S_n) \to \mathbb{C}$ of $\mathbb{C}^{C(G)}$. For example, take the standard basis of $\mathbb{C}^{C(G)}$ and normalize its vectors.

For each transitive subgroup H in our list define

$$p_H := (\langle \psi_i \circ \iota_H, 1_H \rangle_H)_{i=1}^k \in \mathbb{R}^k.$$

We will also precalculate these p_H and store them in a table.

We will assume that these data are already available and we will not consider the construction of the list of transitive groups, the orthonormal basis and the table containing the p_H's, as a part of the actual algorithm. Furthermore, notice that these data do not depend on f but only on its degree.

3.3 Algorithm

The input of the algorithm is a monic irreducible separable polynomial f and an integer x. Let n be the degree of f. As stated in section 3.2 we will assume that the list of transitive subgroups of S_n, the orthonormal basis of $\mathbb{C}^{C(S_n)}$ and the table containing the p_H's are known. The output of the algorithm will be a natural number, which will equal the order of the Galois group G if x is chosen large enough.

The algorithm proceeds as follows. First of all, calculate $\Delta(f)$. Let S be the set of primes $p \leqslant x$ such that $p \nmid \Delta(f)$. For all primes p in S calculate the factorization type $C(f,p)$ of f mod p, and calculate

$$E_i = \frac{1}{|S|} \sum_{p \in S} \psi_i(C(f,p)).$$

Consider $E := (E_i)_{i=1}^k$ as a point of $\mathbb{R}^k$. Choose a transitive subgroup $H \subset S_n$ from our list such that the Euclidean distance between p_H and E is minimal (note that there might be more than one closest point p_H and more than one group H representing the point) and output its order.

3.4 Correctness

In this section we will argue why the algorithm will output $|G|$ for x large enough. First we start with a lemma.

Lemma 3.4.1. *Suppose that $H, H' \subset S_n$ are transitive subgroups. Suppose that $p_H = p_{H'}$. Then $|H| = |H'|$.*

Proof. By Lemma 2.2.4 there are coefficients $c_1, \ldots, c_k \in \mathbb{C}$ such that we have $\sum_{i=1}^k c_i \psi_i = \delta_1^{S_n}$. Then it is just a matter of calculation to verify that

$$(c_i)_{i=1}^k \cdot p_H = \sum_{i=1}^k c_i \langle \psi_i \circ \iota_H, 1_H \rangle_H = \langle \delta_1^H, 1_H \rangle_H = \frac{1}{|H|}.$$

Analogously $(c_i)_{i=1}^k \cdot p_H = (c_i)_{i=1}^k \cdot p_{H'} = \frac{1}{|H'|}$. Hence, $|H| = |H'|$. □

This lemma proves that the output of the algorithm only depends on the choice of p_H and not on the choice of a particular H.

Remark 3.4.2. Note that the output of the algorithm also does not depend on the choice of the orthonormal basis as the Euclidean distance is preserved under orthogonal transformations.

Furthermore, define $\phi_i = \psi_i \circ \iota_G$ for $i = 1, \ldots, k$. By Lemma 1.2.8 we have $\phi_i(F_p) = \psi_i(C(f,p))$ for all $p \in S$. Hence, by Theorem 3.1.1, E_i will be an estimate of $\langle \phi_i, 1_G \rangle_G$ for all $i = 1, \ldots, k$ in the sense that there exists an $x_0 \in \mathbb{N}$ such that p_G is the closest p_H for all $x \geqslant x_0$.

As we already know that the output only depends on p_H, we get the following corollary that proves that our algorithm is correct if x is large enough.

Corollary 3.4.3. *There exists an $x_0 \in \mathbb{N}$ such that for all $x \geqslant x_0$ the output of the Rodriguez Villegas algorithm is equal to $|G|$.* □

Remark 3.4.4. There are effective versions of the Chebotarev density theorem that give rise to effectively computable x_0. However, these x_0 are too large to be of practical use to us.

3.5 Runtime analysis

The size of n is practically bounded by the requirement of the precalculations. For the runtime analysis we will also assume that the coefficients of the polynomial f are all bounded. Hence, we will only consider runtime in terms of x.

The computation of the discriminant does not depend on x in any way and can be done in $\mathcal{O}(1)$. In the algorithm we consider the primes $p \leqslant x$. By the prime number theorem there are $\mathcal{O}(\frac{x}{\log x})$ such primes. To find them we may use a prime number sieve requiring $\mathcal{O}(x)$ operations (see [11]). For each prime p we test divisibility of $\Delta(f)$ by p, which takes $\mathcal{O}(1)$ time. For the primes that do not divide $\Delta(f)$ we need to factor $f \bmod p$. This factoring can be done quite efficiently, namely in average run time $\mathcal{O}((\log p)n^{2+\varepsilon}) = \mathcal{O}(\log p)$ for all $\varepsilon > 0$, by using the probabilistic Cantor-Zassenhaus algorithm (see [12]). Hence, the second part of the algorithm takes at most $\mathcal{O}(x + \frac{x}{\log x} \cdot \log x) = \mathcal{O}(x)$ time. To find the closest p_H we will look up all p_H and calculate all distances, this takes $\mathcal{O}(T(n) \cdot P(n)) = \mathcal{O}(1)$ time, where $T(n)$ is the number of transitive subgroups on our precalculated list and $P(n)$ the number of conjugacy classes of S_n. Notice that the last part can be made more efficient by using space partitioning methods. However, as this does not impose a practical problem on the runtime, we will not do so.

3.6 Alternative algorithm

In this alternative version of the algorithm we will not need the list of transitive groups. Now $\psi_1, \ldots, \psi_k : C(S_n) \to \mathbb{C}$ are the irreducible characters of

S_n and we assume that these irreducible characters are precalculated. For example, by using Magma (see [2]) one can calculate the character table of S_n for $n \leqslant 25$.

The input consists of the polynomial f and an integer x and the output will again be a natural number, which will be the group order $|G|$ if x is chosen large enough.

The following lemma has a corollary that will be useful for the alternative algorithm.

Lemma 3.6.1. *Let ψ be a virtual character of S_n. Then, for all subgroups $H \subset S_n$ we have $\langle \psi \circ \iota_H, 1_H \rangle_H \in \mathbb{Z}$.*

Proof. By definition there are representations V and W of S_n such that $\psi = \chi_V - \chi_W$. We get that $\psi \circ \iota_H = \chi_{\mathrm{Res}_H^{S_n} V} - \chi_{\mathrm{Res}_H^{S_n} W}$ and $1_H = \chi_{\mathbb{C}}$, where $\mathbb{C}$ is viewed as representation of H. Lemma 2.2.7 and Corollary 2.2.6 give that $\langle \psi \circ \iota_H, 1_H \rangle_H = \langle \psi, \chi_{\mathrm{Ind}_H^{S_n} \mathbb{C}} \rangle_{S_n} \in \mathbb{Z}$. □

Corollary 3.6.2. *We have $\langle \phi_i, 1_G \rangle_G \in \mathbb{Z}$ for all $i = 1, \ldots, k$, and hence $p_G \in \mathbb{Z}^k$.* □

The calculation of E is exactly the same as in the original algorithm, however we will not look for the closest p_H, instead, we will look for the closest point in $q \in \mathbb{Z}^k$. Then we will calculate $P := (c_i)_{i=1}^k \cdot q$ where the c_i are as in the proof of Lemma 3.4.1. The output is a divisor d of $n!$ such that $|d - P|$ is minimal.

Again notice that E converges to p_G if $x \to \infty$. By Corollary 3.6.2 we have that $p_G \in \mathbb{Z}^k$. Hence, the point q will eventually become p_G. Hence, also this alternative version of the algorithm outputs $|G|$ if x is large enough. The runtime, as function of x, is similar to the runtime we found for the original algorithm.

3.7 Examples

Consider the following polynomials.

$$\begin{aligned} f_1 :=\ & x^{12} - x^{11} + \ldots - x + 1; \\ f_2 :=\ & x^{12} + 4x^{11} + 8x^{10} - 160x^9 + 144x^8 + 612x^7 - 276x^6 \\ & - 1164x^5 + 1209x^4 - 380x^3 + 22x^2 + 8x - 1; \\ f_3 :=\ & x^{12} - x^9 - x^4 + x + 1. \end{aligned}$$

All these polynomials are irreducible. For $j = 1, 2, 3$, let L_j be a splitting field of f_j over $\mathbb{Q}$, then the Galois group of $L_j/\mathbb{Q}$ is the cyclic group C_{12} of order 12 if $j = 1$, the Mathieu group M_{12} of order 95040 if $j = 2$ and the symmetric group S_{12} of order 479001600 if $j = 3$ (see [8]).

In the following tables, for a number of values of x the output Q of the Rodriguez Villegas algorithm, the output Q' of the alternative version and the distance between p_H and E are depicted. Recall the naive algorithm in which we count the totally split primes and then round the inverted fraction to the nearest divisor of $n!$ (or ∞ if no totally split primes were found). For $j = 1$ the output W of the naive algorithm is also presented. As for $j = 2, 3$ no primes smaller than x were found for which $f \bmod p$ totally splits into linear factors, the output of the naive algorithm is not included in these cases.

x	Q	$\|p_H - E\|$	Q'	W
10^1	15552	117.50	46200	∞
10^2	12	324.35	12	12
10^3	12	14438	12	14
10^4	12	3529.5	12	12
10^5	12	8.1572	12	12
10^6	12	0.4541	12	12

Table for $j = 1$.

x	Q	$\|p_H - E\|$	Q'
10^1	239500800	9.0000	1
10^2	95040	62.698	56320
10^3	95040	1.0484	95040
10^4	95040	0.1951	95040
10^5	95040	0.0806	95040
10^6	95040	0.0563	95040

Table for $j = 2$.

x	Q	$\lvert p_H - E\rvert$	Q'
10^1	479001600	8.6875	1
10^2	479001600	1.7184	479001600
10^3	479001600	0.2900	479001600
10^4	479001600	0.0593	479001600
10^5	479001600	0.0075	479001600
10^6	479001600	0.0008	479001600

Table for $j = 3$.

Note that the naive algorithm only works for very small groups. The Rodriguez Villegas algorithm appears to be the best algorithm. It outputs the correct group order already for $x = 10^2$, though the large distance $|p_H - E|$ suggests that this might be coincidental.

The distance $|p_H - E|$ converges faster to 0 for large groups G. In the case $j = 3$, for example, we only need to consider the primes up to 10^4 to find an E within distance 0.1 of p_G. By comparison, the naive algorithm would need at least $12! \approx 4{,}8 \cdot 10^8$ primes to hope to be able to distinguish between the order of S_{12} and the order of A_{12}.

4 A probabilistic model

Let $f \in \mathbb{Z}[X]$ be a monic irreducible polynomial of degree n, let $L/\mathbb{Q}$ be a splitting field of f and let G be the Galois group of $L/\mathbb{Q}$. Furthermore, let p be a prime number and let $C \in C(S_n)$ be a conjugacy class. The Chebotarev density theorem suggests that the 'probability' that f mod p has factorization type C equals the probability that a random element of G has cycle type C.

To analyse the Rodriguez Villegas algorithm of the previous chapter, we will consider a probabilistic model in which factorization types will be drawn randomly according to the probability distribution implied by the Chebotarev density theorem. This analysis will give us an idea why the Rodriguez Villegas algorithm is better than the naive algorithm (see page 20), at least for large Galois groups.

In principle, we could also use an effective version of the Chebotarev density theorem to further analyze the algorithm. However, this amounts to a lot of calculations and the bounds will become quite weak.

4.1 The model

Now let $G \subset S_n$ be a transitive subgroup (not necessarily a Galois group). For a subgroup $H \subset S_n$ let $\iota_H \colon C(H) \to C(S_n)$ be the map induced by the inclusion (see page 5). Let k be a natural number and let $\psi_1, \ldots, \psi_k \colon C(S_n) \to \mathbb{R}$ be real-valued class functions of S_n such that $\langle \psi_i, 1_H \rangle_H \in \mathbb{Z}$ for all transitive subgroups $H \subset S_n$ and all $i = 1, \ldots, k$. Furthermore, define $\phi_i = \psi_i \circ \iota_G$. We will also assume that there are coefficients $c_1, \ldots, c_k \in \mathbb{C}$ such that we have $\sum_{i=1}^{k} c_i \psi_i = \delta_1^{S_n}$ (see page 15). This will assure us that the conclusion of Lemma 3.4.1 holds.

Let X be a random variable with state space G such that for all $g \in G$ we have that $\Pr(X = g) = \frac{1}{|G|}$. Define $Y = \iota_G(C(X))$, i.e. the cycle type of the element X. Furthermore, define the random variable $Z^{(i)}$ on the state space $\mathbb{C}$ by $Z^{(i)} = \psi_i(Y) = \phi_i(C(X))$. Let σ_Z^i be the standard deviation of $Z^{(i)}$.

Just like in the Rodriguez Villegas algorithm our goal is to find the value of

$$\langle \phi_i, 1 \rangle_G = \sum_{g \in G} \Pr(X = g) \cdot \phi_i(C(g)) = \sum_{z \in \phi_i(C(G))} \Pr(Z^{(i)} = z) \cdot z = \mathbb{E}(Z^{(i)}).$$

We will consider a Monte Carlo experiment with an oracle that outputs elements of $C(S_n)$ according to the probability distribution of Y. In the experiment we estimate $\mu_i := \mathbb{E}(Z^{(i)})$ by calculating the sample average of $\psi_i(Y)$. More formally, let $N > 0$ be an integer, let $Z_j^{(i)}$ for $j = 1, \ldots, N$ be independent and identically distributed copies of $Z^{(i)}$ and let $A^{(i)} = \frac{1}{N}\sum_{j=1}^{N} Z_j^{(i)}$ be the sample average.

4.2 Analysis

The following proposition tells us what $\mathrm{Var}(A^{(i)})$ is, which measures the expected error in $A^{(i)}$.

Proposition 4.2.1. *The variance* $\mathrm{Var}(A^{(i)})$ *is equal to* $\frac{1}{N}\langle \phi_i', \phi_i' \rangle_G$ *where* $\phi_i' = \phi_i - \langle \phi_i, 1_G \rangle_G \cdot 1_G$.

Proof. By basic probability theory we derive

$$\mathrm{Var}(A^{(i)}) = \frac{1}{N^2}\sum_{j=1}^{N} \mathrm{Var}(Z_j^{(i)}) = \frac{1}{N^2} \cdot N(\sigma_Z^i)^2 = \frac{1}{N}(\sigma_Z^i)^2.$$

Furthermore, it is just a matter of calculation to find

$$\begin{aligned}
(\sigma_Z^i)^2 &= \mathbb{E}((Z^{(i)} - \mu_i)(\overline{Z^{(i)}} - \mu_i)) = \mathbb{E}(Z^{(i)}\overline{Z^{(i)}}) - \mu_i^2 - \mu_i^2 + \mu_i^2 \\
&= \mathbb{E}(Z^{(i)}\overline{Z^{(i)}}) - \mu_i^2 = \sum_{g \in G} \frac{1}{|G|}\phi_i(g)\overline{\phi_i(g)} - \left(\sum_{g \in G} \frac{1}{|G|}\phi_i(g)\right)^2 \\
&= \langle \phi_i, \phi_i \rangle_G - (\langle \phi_i, 1_G \rangle_G)^2 = \langle \phi_i', \phi_i' \rangle_G.
\end{aligned}$$

This proves the assertion. □

Remark 4.2.2. Suppose that $\psi_1, \ldots, \psi_k$ are the irreducible characters of G. Then the variance is bounded by $\frac{1}{N}\langle \phi_i, \phi_i \rangle_G \leqslant \frac{\langle \psi_i, \psi_i \rangle_{S_n} \cdot |S_n|}{N \cdot |G|} = \frac{1}{N}[S_n : G]$. Therefore, if $[S_n : G]$ is small we would expect to have faster convergence. This is completely in line with our observations in section 3.7.

Let $r\colon \mathbb{R}^k \to \mathbb{Z}$ be any function that has the property that it maps a point $q \in \mathbb{R}^k$ to the order of a transitive subgroup $H \subset S_n$ such that $|q - p_H|$ is minimal, where p_H is defined as on page 16. As Lemma 3.4.1 holds by our assumptions, it is obvious that the probability that $\eta := r((A^{(i)})_{i=1}^k)$ equals $|G|$ tends to 1 as $N \to \infty$. In the following theorem a more precise statement is made.

Theorem 4.2.3. *Let M_i be the maximum value that $|Z^{(i)} - \mathbb{E}(Z^{(i)})|$ may attain. Then*

$$\Pr(\eta = |G|) \geqslant 1 - \sum_{i=1}^{k} 2e^{-\frac{\frac{1}{4}N}{2(\sigma_Z^i)^2 + M_i/3}}.$$

Proof. For $i = 1, \ldots, k$, let Q_i be the event that $|A^{(i)} - \mu_i| < \frac{1}{2}$. Applying Theorem 2.6 of [6] gives that

$$\Pr\left(A^{(i)} \geqslant \mathbb{E}(Z^{(i)}) + \tfrac{1}{2}\right) \leqslant e^{-\frac{\frac{1}{4}N^2}{2(N(\sigma_Z^i)^2 + M_i N/6)}}.$$

By applying the same inequality to the case where $A^{(i)} \leqslant \mathbb{E}(Z^{(i)}) - \frac{1}{2}$ we find that

$$\Pr(\text{not } Q_i) \leqslant 2e^{-\frac{\frac{1}{4}N}{2(\sigma_Z^i)^2 + M_i/3}}.$$

Note that $\eta = |G|$ certainly holds if Q_i happens for all $i = 1, \ldots, k$. By using the laws of probability we find

$$\Pr(\eta = |G|) \geqslant 1 - \sum_{i=1}^{k} \Pr(\text{not } Q_i) \geqslant 1 - \sum_{i=1}^{k} 2e^{-\frac{\frac{1}{4}N}{2(\sigma_Z^i)^2 + M_i/3}}.$$

□

In the next section, this upper bound will be calculated for a few example cases to give an idea about the size of N that is sufficient to have a small error probability.

4.3 Examples

For the example groups occurring in section 3.7 we have calculated the bound

$$B := \sum_{i=1}^{k} e^{-\frac{\frac{1}{4}N}{2(\sigma_Z^i)^2 + M_i/3}}$$

given in Theorem 4.2.3. The following tables contain the bounds for different values of N.

N	10^7	10^8	10^9
B	15.6	0.181	$2.042 \cdot 10^{-12}$

Bounds for $j = 1$ (order 12)

N	10^4	10^5	10^6
B	11.2	$3.42 \cdot 10^{-3}$	$7.91 \cdot 10^{-29}$

Bounds for $j = 2$ (order 95040)

N	10^4	10^5	10^6
B	7.62	$1.35 \cdot 10^{-4}$	$1.08 \cdot 10^{-42}$

Bounds for $j = 3$ (order 479001600)

For $j = 2$ and $j = 3$ the bound is not useful for $N = 10^4$ as $B \geqslant 1$ in these cases, but it is already quite strong for $N = 10^5$. This is due to the quite large value of the M_i (the largest M_i is 7700).

For $j = 1$ it takes very long for the bound to become useful. This is due to the large variance that occurs (the largest $(\sigma_Z^i)^2$ is 4526132). In section 3.7 we have already seen that for $j = 1$ the convergence of E to p_G was very slow, which is completely in line with the above result.

Acknowledgements

I would first like to thank my thesis supervisor Lenny Taelman. His enthusiasm and guidance were of great value to me. Without his help I could never have written this thesis.

Furthermore, I would like to thank Fernando Rodriguez Villegas for sharing his ideas that formed the basis of this thesis.

Moreover, I would also like to thank Peter Stevenhagen, Hendrik Lenstra, Michiel Kosters and Nick Towner for the corrections and suggestions they made.

Finally, I would like to thank all other people that are not mentioned here, but that have contributed in some way to this thesis.

Raymond van Bommel

References

[1] M.F. Atiyah & I.G. MacDonald. *Introduction to Commutative Algebra.* Westview Press, Colorado Oxford, 1969.

[2] Wieb Bosma, John Cannon & Catherine Playoust. The Magma algebra system. I. The user language. *Journal of Symbolic Computation* **24** (1997): 235–265.

[3] Daniel I.A. Cohen & Talbot M. Katz. Prime Numbers and the First Digit Phenomenon. *Journal of Number Theory* **18** (1984): 261–268.

[4] N. Tschebotareff (Chebotarev). Die Bestimmung der Dichtigkeit einer Menge von Primzahlen, welche zu einer gegebenen Substitutionsklasse gehören. *Mathematische Annalen* **95** (1925): 191–228.

[5] G. Dalla Torre. *Representation theory.* Accessed 9 May 2012, <http://www.win.tue.nl/mm-representation-theory/representation_theory.pdf>.

[6] Fan Chung. *Old and New Concentration Inequalities* Accessed 19 July 2012, <http://www.math.ucsd.edu/~fan/complex/ch2.pdf>.

[7] Kenneth Ireland & Michael Rosen. *A Classical Introduction to Modern Number Theory.* Second Edition. Springer-Verlag, Berlin Heidelberg New York, 1990.

[8] Jürgen Klüners & Gunter Malle. *A Database for Number Fields.* Accessed 17 July 2012, <http://www.math.uni-duesseldorf.de/~klueners/minimum/>.

[9] Serge Lang. *Algebraic Number Theory.* Addison Wesley, Massachusetts, 1970.

[10] Jürgen Neukirch. *Algebraic Number Theory.* Translated by Norbert Schappacher. Springer-Verlag, Berlin Heidelberg New York, 1999.

[11] Paul Pritchard. Fast compact prime number sieves (among others). *Journal of Algorithms* **4**.4 (1983): 332–344.

[12] Victor Shoup. On the deterministic complexity of factoring polynomials over finite fields. *Information Processing Letters* **33** (1990): 261–267.

[13] Benjamin Steinberg. *Representation Theory of Finite Groups: An Introductory Approach.* Springer-Verlag, Berlin Heidenberg New York, 2012.

[14] P. Stevenhagen. *Algebra 3*. Accessed 9 May 2012, <`http://websites.math.leidenuniv.nl/algebra/algebra3.pdf`>.

[15] P. Stevenhagen & H.W. Lenstra, Jr. Chebotarëv and his Density Theorem. *The Mathematical Intelligencer* **18**.2 (1996): 26–37.

[16] Gérald Tenenbaum. *Introduction to analytic probabilistic number theory*. Translated by C.B. Thomas. Cambridge University Press, Cambridge, 1995.

[17] B.L. van der Waerden. Die Seltenheit der Gleichungen mit Affekt. *Mathematische Annalen* **109** (1934): 13–16.