ARTIN'S PRIMITIVE ROOT CONJECTURE
- a survey -

PIETER MOREE

ABSTRACT. This is an expanded version of a talk given 19-10-2000 at the DMV-seminar 'The Riemann zeta function and Random Matrix Theory' in Oberwolfach. A large part of it was intended to be understandable by non-number theorists with a mathematical background. The talk covers some of the history, results and ideas connected to Artin's celebrated primitive root conjecture dating from 1927.

Gauss considers in articles 315-317 of his Disquisitiones Arithmeticae (1801) the decimal expansion of numbers of the form $\frac{1}{p}$ with p prime. For example, $\frac{1}{7} = \cdot 142857142857\ldots$, $\frac{1}{11} = \cdot 090909\ldots$.

These decimal expansions for $p \neq 2$ and $p \neq 5$ are purely periodic. It is easy to see that the period is equal to the smallest integer k such that $10^k \equiv 1 (\bmod p)$. This integer k is the *multiplicative order* of 10 modulo p and is denoted by $\mathrm{ord}_p(10)$. The integer k equals the order of the subgroup generated by 10 in $(\mathbb{Z}/p\mathbb{Z})^*$, the multiplicative subgroup of residue classes modulo p. By Lagrange's theorem the order of a subgroup of a finite group is a divisor of the number of elements in the group. Since $(\mathbb{Z}/p\mathbb{Z})^*$ has $p-1$ elements, we conclude that $\mathrm{ord}_p(10) \mid p-1$. If $\mathrm{ord}_p(10) = p-1$, we say that 10 is a *primitive root* mod p. The decimal expansion we have given for $\frac{1}{7}$ shows that 7 is a primitive root mod 10. Gauss' table gives several other such examples and he must have asked himself whether there are infinitely many such primes, that is, primes for which the decimal period is $p-1$.

Some light on this question can be shed on using the simple observation (due to Chebyshev) that if $q = 4p+1$ is a prime with p a prime satisfying $p \equiv 2 (\bmod 5)$, then 10 is a primitive root modulo q. To this end notice that, a priori, $\mathrm{ord}_q(10) = \{1, 2, 4, p, 2p, 4p\}$. Now, using the law of quadratic reciprocity, see e.g. [11], one deduces that

$$10^{2p} \equiv \left(\frac{10}{q}\right) = \left(\frac{2}{q}\right)\left(\frac{5}{q}\right) = (-1)^{\frac{q^2-1}{8}}\left(\frac{q}{5}\right) = -\left(\frac{4}{5}\right) \equiv -1 (\bmod q).$$

Since no prime divisor of $10^4 - 1$ satisfies the requirements (and hence $\mathrm{ord}_q(10) \nmid 4$), one sees that $\mathrm{ord}_q(10) = 4p = q-1$.

[1]1991 *Mathematics Subject Classification.* Primary: 11R45; Secondary: 11A07, 11N69, 11R42
Key words and phrases. Artin primitive root, natural density, number field

The *logarithmic integral* $\mathrm{Li}(x)$ is defined as $\int_2^x dt/\log t$. The Prime Number Theorem in the form

$$\pi(x) := \#\{p \le x\} \sim \mathrm{Li}(x), \quad x \to \infty,$$

suggests that the 'probability that a number n is prime' is $1/\log n$. Thus for both n and $4n+1$ to be prime and $n \equiv 2 (\mathrm{mod}\ 5)$ we expect a probability of $1/(5\log^2 n)$, assuming independence of these three events. Since they are not, we have to correct by some positive constant c and hence expect that up to x there are at least $cx/\log^2 x$ (note that $\sum_{n\le x}\log^{-2} n \sim x\log^{-2} x$) primes p such that 10 is a primitive root mod p. Hence we expect that there are infinitely many primes p having 10 as a primitive root mod p.

Emil Artin in 1927, led by probability arguments, made the following conjecture, where for a given $g \in \mathbb{Q}^*$ we define $\mathcal{P}(g) = \{p : \mathrm{ord}_p(g) = p-1\}$ and $\mathcal{P}(g)(x) = \#\{p \in \mathcal{P}(g) : p \le x\}$. (In general, if S is a set, we define $S(x) = \#\{s \in S : s \le x\}$.)

Artin's primitive root conjecture (1927) Let $g \in \mathbb{Q}\backslash\{-1, 0, 1\}$.
(Qualitative form) The set $\mathcal{P}(g)$ is infinite if g is not a square of a rational number.
(Quantitative form) Let h be the largest integer such that $g = g_0^h$ with $g_0 \in \mathbb{Q}$. We have, as x tends to infinity,

$$\mathcal{P}(g)(x) = \prod_{p\nmid h}\left(1 - \frac{1}{p(p-1)}\right)\prod_{p|h}\left(1 - \frac{1}{p-1}\right)\frac{x}{\log x} + o\left(\frac{x}{\log x}\right). \quad (1)$$

Denote the product appearing in the latter equation by $A(h)$. Note that $A(h) = 0$ when h is even. Then $\mathcal{P}(g)$ is finite and (1) is trivial. Thus the condition that g is not a square is necessary (and according to Artin sufficient) for $\mathcal{P}(g)$ to be infinite. For h odd we have that $A(h)$ equals a positive rational multiple of

$$A(1) = A = \prod_p \left(1 - \frac{1}{p(p-1)}\right) \approx \cdot 3739558136192\ldots,$$

the *Artin constant*. Thus the quantitative form implies the qualitative form.

The product defining the Artin constant does not converge too quickly. It turns out that for numerical approximations it is possible and indeed much more efficient to write $A = \prod_{k=2}^{\infty}\zeta(k)^{-e_k}$, with $e_k \in \mathbb{N}$. This then allows one to determine A without too much effort with 1000 decimal precision, cf [19].

Usually one talks about the Artin primitive root conjecture, rather than Artin's conjecture since there are various unresolved conjectures due to Artin (most notably the Artin holomorphy conjecture).

The remainder of the talk consists of the following 7 parts :

I Naive heuristic approach

II Algebraic number theory

III Artin's heuristic approach

IV Modified heuristic approach (à la Artin).

V Hooley's work

VI Probabilistic model

VII Some variations of Artin's problem.

Our starting point in our analysis of Artin's primitive root conjecture is the following observation :

$$p \in \mathcal{P}(g) \iff g^{\frac{p-1}{q}} \not\equiv 1 (\text{mod } p) \text{ for every prime } q \text{ dividing } p-1. \quad (2)$$

"$\Longrightarrow$" Obvious.
"$\Longleftarrow$" Suppose $p \notin \mathcal{P}(g)$. Then $g^{\frac{p-1}{k}} \equiv 1(\text{mod } p)$ for some $k \mid (p-1)$, $k > 1$. But this implies that $g^{\frac{p-1}{q_1}} \equiv 1(\text{mod } p)$ for some prime divisor q_1 of k. This is a contradiction.

Thus, associated with a prime p we have conditions for various q. We are going to interchange the role of p and q, that is for a fixed q we consider the set of primes p such that $p \equiv 1(\text{mod } q)$ and $g^{\frac{p-1}{q}} \not\equiv 1(\text{mod } p)$.

1 Naive heuristic approach

Fix any prime q. We try to compute the density of primes p such that both $p \equiv 1(\text{mod } q)$ and $g^{\frac{p-1}{q}} \equiv 1(\text{mod } p)$. By Dirichlet's theorem $p \equiv 1(\text{mod } q)$ is true for primes p with frequency $1/\varphi(q) = 1/(q-1)$. Note that $g^{\frac{p-1}{q}}$ is a solution of $x^q \equiv 1(\text{mod } p)$. We expect there to be q solutions and we want a solution to be 1 modulo q. Thus we expect to be successful with probability $\frac{1}{q}$, except when $q \mid h$. Then $g^{\frac{p-1}{q}} = g_0^{h\frac{p-1}{q}} \equiv 1(\text{mod } p)$, trivially. If we assume that these events are independent then the probability that both events occur is $\frac{1}{q(q-1)}$ if $q \nmid h$ and $\frac{1}{q-1}$ otherwise.

By (2) the above events must not occur for any q in order to ensure that $p \in \mathcal{P}(g)$. This suggests a natural density of

$$\prod_{q \nmid h} \left(1 - \frac{1}{q(q-1)}\right) \prod_{q \mid h} \left(1 - \frac{1}{q-1}\right) = A(h)$$

for such primes. That is we expect (1) to hold true.

2 Algebraic number theory

In order to understand Artin's approach to his conjecture we need some facts about algebraic number theory. We say α is *algebraic* over $\mathbb{Q}$ if it satisfies an equation of the form $f(\alpha) = 0$ with $f(x) \in \mathbb{Z}[x]$. We say α is an *algebraic* integer over $\mathbb{Q}$ if it satisfies an equation of the form $f(\alpha) = 0$, where $f(x) \in \mathbb{Z}[x]$ is monic. A *number field* over $\mathbb{Q}$ is a field obtained by adjoining finitely many algebraic numbers $\alpha_1, \ldots, \alpha_s$ to $\mathbb{Q}$. That is, $K = \mathbb{Q}(\alpha_1, \ldots, \alpha_s)$. By $\mathbb{Q}(\alpha_1, \ldots, \alpha_s)$ we denote the smallest field containing $\alpha_1, \ldots, \alpha_s$ and $\mathbb{Q}$. The so called *theorem of the primitive element* asserts that for a given number field K there exists an algebraic integer α such that $K = \mathbb{Q}(\alpha)$. The element α is said to be a *primitive element*. For an algebraic integer α let $f_\alpha(x)$ be the unique (as it turns out) monic polynomial $f_\alpha(x)$ of smallest degree such that $f_\alpha(\alpha) = 0$. Then $K = \mathbb{Q}(\alpha)$ is isomorphic to $\mathbb{Q}[x]/(f_\alpha(x))$. The *degree* $[K : \mathbb{Q}]$ of K is defined as $\deg f_\alpha(x)$. The *compositum* of two fields K and L is defined as the smallest field containing both K and L.

Example. Let $\alpha = i$. Then $f_i(x) = x^2 + 1$. Note that $\mathbb{Q}[x]$ contains elements of the form $\sum_{j=0}^{m} a_j x^j$. These elements correspond with an element in $\mathbb{Q}[x]/(x^2 + 1)$ obtained by replacing every x^2 by -1. Thus, as a set, $\mathbb{Q}[x]/(x^2 + 1) \cong \{a + bx \mid a, b \in \mathbb{Q}\}$. Similarly, as a set, $\mathbb{Q}(i) = \{a + bi \mid a, b \in \mathbb{Q}\}$. It is not difficult to see that a field isomorphism between $\mathbb{Q}(i)$ and $\mathbb{Q}[x]/(x^2 + 1)$ is given by sending X to i.

For a general number field K, we have the following picture :

$$\begin{array}{ccc} \mathbb{Q}(\alpha) = K & \supset & \mathcal{O}_K \\ \cup & & \cup \\ \mathbb{Q} & \supset & \mathbb{Z} \end{array}$$

Here $\mathcal{O}_K$ is the *ring of integers* of the field K. It plays a role analogous to that of $\mathbb{Z}$ in $\mathbb{Q}$ (in our example $\mathcal{O}_K$ is $\mathbb{Z}[i] = \{a + bi | a, b \in \mathbb{Z}\}$ the *ring of Gaussian integers*). The *main theorem of arithmetic* states that, up to the order of factors, factorisation into primes in $\mathbb{Z}$ is unique. One might hope for a ring inside K with similar properties. Let us consider the set of algebraic integers over $\mathbb{Q}$ that are inside K. This set is actually a ring, the ring of integers, $\mathcal{O}_K$, of K. It usually does not have unique factorization in elements. It, however, has unique factorization into terms of *prime ideals* (up to order of factors). The prime ideal (p) in $\mathbb{Z}$ turns out to factor as $\mathfrak{P}_1^{e_1} \ldots \mathfrak{P}_g^{e_g}$ inside $\mathcal{O}_K$. The equivalent statement of $p = \#\mathbb{Z}/p\mathbb{Z}$ in $\mathbb{Z}$, reads in $\mathcal{O}_K$: $p^{f_i} = \#\mathcal{O}_K/\mathfrak{P}_i\mathcal{O}_K$. Here f_i is called the *degree* of the prime ideal $\mathfrak{P}_i$. We have the relationship $\sum_{i=1}^{g} e_i f_i = n$, with n the degree of K. We say that a rational prime p splits completely in $\mathcal{O}_K$ if $e_i = f_i = 1$ for $i = 1, \ldots, g$. Note that in this case $g = n$. If a prime p splits completely in the ring of integers $\mathcal{O}_K$, with $K \cong \mathbb{Q}[x]/(f(x))$, then over $\mathbb{Z}/p\mathbb{Z}$, $f(x)$ splits into n distinct linear factors.

2.1 Analytic algebraic number theory

Let $\mathfrak{O}$ run through the non-zero integral ideals of $\mathcal{O}_K$, with K a number field. For $\Re(s)>1$, we define $\zeta_K(s)=\sum_{\mathfrak{O}} N\mathfrak{O}^{-s}$ and elsewhere by analytic continuation. Note that if $K=\mathbb{Q}$, then $\zeta_K(s)=\zeta(s)$, the usual zeta-function. For $\Re(s)>1$ we have an Euler product, $\zeta_K(s)=\prod_{\mathfrak{P}}(1-N\mathfrak{P}^{-s})^{-1}$, where the product runs over all prime ideals $\mathfrak{P}$ of $\mathcal{O}_K$. Hecke, in the 1920's showed in a complicated work that $\zeta_K(s)$ satisfies a functional equation, analogous to the celebrated functional equation for the Riemann zeta-function. Tate, in his famous thesis (written around 1950), using adeles and harmonic analysis, greatly simplified Hecke's work. In 1903, Landau proved the *Prime Ideal Theorem* to the effect that

$$\pi_K(x):=\sum_{N\mathfrak{P}\leqq x}1\sim\frac{x}{\log x},\quad x\to\infty.$$

Assuming the *Riemann Hypothesis* for the field K, that is, assuming that all the non-trivial zeros of $\zeta_K(s)$ are on $\Re(s)=\frac{1}{2}$, one obtains the much sharper estimate

$$\pi_K(x)=\mathrm{Li}(x)+O_K(\sqrt{x}\log x),\quad x\to\infty,\tag{3}$$

where the implied constant depends at most on K.

In the analysis of Artin's primitive root conjecture an infinite family of fields will play a role and we need an error term in (3) that is explicit in its dependency on K. Such a result was folklore and finally written down by Lang [9]:

$$\pi_K(x)=\mathrm{Li}(x)+O(\sqrt{x}\log(x^{[K:\mathbb{Q}]}|D_K|)),\tag{4}$$

where D_K denotes the discriminant of the field K.

The function $\pi_K(x)$ is heavily biased towards prime ideals of degree one:

$$\begin{aligned}\pi_k(x) &= \sum_{\substack{N\mathfrak{P}\leqq x\\ N\mathfrak{P}=p}}1+\sum_{\substack{N\mathfrak{P}\leqq x\\ N\mathfrak{P}=p^2}}1+\ldots\\ &= \sum_{\substack{N\mathfrak{P}\leqq x\\ N\mathfrak{P}=p}}1+O(\sum_{\substack{p^j\leqq x\\ j\geqq 2}}1)=\sum_{\substack{N\mathfrak{P}\leqq x\\ N\mathfrak{P}=p}}1+O(\sqrt{x}\log x).\end{aligned}$$

In Artin's problem one is specifically interested in so called *normal fields*. For such a field all prime ideals of degree one with at most finitely many exceptions, come from rational primes p that split completely. Throwing out these exceptions we have the following picture :

$$\begin{array}{ccc}\mathfrak{P}_1\mathfrak{P}_2\ldots\mathfrak{P}_n & \qquad & \mathfrak{P}_1'\mathfrak{P}_2'\ldots\mathfrak{P}_n'\\ \ddots\vdots\iddots & & \ddots\vdots\iddots\\ p_1 & & p_1'\end{array}$$

(A $[K:\mathbb{Q}]$-fold covering of the rational primes.) Thus, for a normal field we have

$$\pi_K(x) = [K:\mathbb{Q}] \sum_{\substack{p \leqq x \\ p \text{ splits completely in } K}} 1 + O(\sqrt{x}\log x),$$

that is, by the Prime Ideal Theorem,

$$\sum_{\substack{p \leqq x \\ p \text{ splits completely in } K}} 1 \sim \frac{1}{[K:\mathbb{Q}]}\frac{x}{\log x}, \quad x \to \infty. \tag{5}$$

This is a particular case of a very important result called the *Chebotarev density theorem*. For a nice introductory account of this result see Lenstra and Stevenhagen [26].

Example. (Cyclotomic fields). A *cyclotomic field* K is a field of the form $K = \mathbb{Q}(\zeta_n)$ with $\zeta_n = e^{2\pi i/n}$ and n a natural number. One can show that

$$f_{\zeta_n}(x) = \prod_{\substack{a=1 \\ (a,n)=1}}^{n} (x - \zeta_n^a) = \Phi_n(x) \in \mathbb{Z}[x].$$

The polynomial $\Phi_n(x)$ is the nth *cyclotomic polynomial*. From this we deduce that $[\mathbb{Q}(\zeta_n):\mathbb{Q}] = \deg \Phi_n(x) = \varphi(n)$. The cyclotomic fields are normal. It can be shown that p splits completely in $K \iff p \equiv 1 (\text{mod } n) \iff x^n - 1$ factors completely over $\mathbb{Z}/p\mathbb{Z}$.

On applying (5) we deduce that

$$\pi(x,n,1) := \sum_{\substack{p \leqq x \\ p \equiv 1 (\text{mod } n)}} 1 \sim \frac{1}{\varphi(n)}\frac{x}{\log x}, \quad x \to \infty,$$

a particular case of Dirichlet's theorem. The same result can be deduced for the progressions $a(\text{mod } n)$ with $(a,n) = 1$ on invoking Chebotarev's density theorem.

3 Artin's heuristic approach

We have now the preliminaries available to follow the rest of the story on the progress made on Artin's primitive root conjecture. Remember that we were interested in the set of all primes p satisfying

$$\begin{cases} p \equiv & 1(\text{mod } q) \\ g^{\frac{p-1}{q}} \equiv & 1(\text{mod } p), \end{cases}$$

where q is any fixed prime. By what we have learned about cyclotomic fields the primes $p \equiv 1(\text{mod } q)$ are precisely those for which the equation $x^q \equiv$

$1 \pmod p$) has q distinct solutions mod p. I claim that if $g^{\frac{p-1}{q}} \equiv 1 \pmod p$, then $x^q \equiv g \pmod p$ has a solution mod p. To see this, write $g = \gamma^a$ with γ a primitive root mod p. Then $\gamma^{a\frac{p-1}{q}} \equiv 1 \pmod p$. A power of a primitive root can only be congruent to $1 \pmod p$ if the exponent is a multiple of $p-1$, hence $q \mid a$. This says $a = bq$. Then $(\gamma^b)^q \equiv g \pmod p$ and this proves the claim. Let $\alpha_1, \ldots, \alpha_q$ be the distinct mod p solutions of $x^q \equiv 1 \pmod p$. Then we see that $\alpha_1\gamma^b, \ldots, \alpha_q\gamma^b$ are all distinct solutions of $x^q \equiv g \pmod p$. We conclude that a prime that satisfies $p \equiv 1 \pmod q$ and $g^{\frac{p-1}{q}} \equiv 1 \pmod p$ splits completely in the number field

$$k_q := \mathbb{Q}(\zeta_q, g^{\frac{1}{q}}). \tag{6}$$

(Note that it does not make a difference which qth root of g we take, in that we always end up with the same field.) We leave it to the reader to show that if p splits completely in k_q, then $p \equiv 1 \pmod q$ and $g^{\frac{p-1}{q}} \equiv 1 \pmod p$. Thus we have

$$\left\{ \begin{array}{rl} p \equiv & 1 \pmod q \\ g^{\frac{p-1}{q}} \equiv & 1 \pmod p \end{array} \right. \iff p \text{ splits completely in } k_q.$$

Let us denote the degree of k_q by $n(q)$. It is not difficult to show that k_q is normal. Hence we may apply (5) to deduce that the number of primes $p \leq x$ that do not split completely in k_q equals $(1-1/n(q))x/\log x$, asymptotically. So one expects $\prod_q(1-1/n(q))$ as a density of primes p for which g is a primitive root mod p.

The latter heuristic argument was believed until about 1960, when the Lehmers [10] made some numerical calculations that did not seem to always match with Artin's heuristic. Then, in 1968, Heilbronn realized that the events 'p does not split completely in k_q' are not necessarily independent, as p and q range through all primes and published a corrected quantitative conjecture (Artin when he received a letter from Lehmer on his numerical work, came to the same correction much earlier, but did not publish it).

Take for example $g = 5$, thus $h = 1$. Then $k_2 = \mathbb{Q}(\sqrt{5}) \subsetneq k_5 = \mathbb{Q}(\zeta_5, 5^{1/5})$, i.e. k_2 is a subfield of k_5. Now if $K \subsetneq L$ and p splits completely in L, then p must split completely in K. This means that the condition 'p does not split completely in k_2' implies the condition 'p does not split completely in k_5'. So, assuming that there are no further dependencies between the various conditions on q, we expect :

$$\mathcal{P}(5)(x) \sim \prod_{p \neq 5} \left(1 - \frac{1}{p(p-1)}\right) \frac{x}{\log x}, \quad x \to \infty.$$

Note that the constant involved equals $20A/19$. This turns out to be more consistent with the numerical data than Artin's prediction A.

4 Modified heuristic approach (à la Artin)

Recall that k_q is defined in (6) for prime q. Let m be squarefree. We define k_m to be the compositum of the fields k_q with $q|m$ and prime. It can be shown that $k_m = \mathbb{Q}(\zeta_m, g^{1/m})$. We are interested in the density of primes p that do not split completely in any k_q (if this exists). We can try to compute it by inclusion-exclusion. So we start writing down $1-\frac{1}{n(2)}-\frac{1}{n(3)}\cdots$. However, we have counted double here the primes p that split completely in both k_2 and k_3. It can be shown that the primes that split completely in both k_n and k_m are precisely the primes that split completely in $k_{lcm(n,m)}$. Thus we have to add $\frac{1}{n(6)}$. Continuing in this fashion we arrive at the heuristic

$$\mathcal{P}(g)(x) \sim \sum_{k=1}^{\infty} \frac{\mu(k)}{[\mathbb{Q}(\zeta_k, g^{1/k}) : \mathbb{Q}]} \frac{x}{\log x}, \quad x \to \infty. \tag{7}$$

Here μ is the *Möbius function*. It is defined by $\mu(1) = 1$. If $n > 1$, then $\mu(n) := (-1)^s$, where s denotes the number of distinct prime factors of n if n is squarefree, and $\mu(n) := 0$ otherwise. If we are only interested in the primes p that do not split completely in any of $k_{q_1}, \ldots, k_{q_s}$, with s finite then there would be little to prove. It is the infinitely many q that makes the problem hard.

Artin's heuristic amounts to assuming that fields k_{q_1} and k_{q_2} are linearly independent over $\mathbb{Q}$, for distinct primes q_1 and q_2. This has as a consequence that $[k_{q_1 q_2} : \mathbb{Q}] = [k_{q_1} : \mathbb{Q}][k_{q_2} : \mathbb{Q}]$ and in general, for squarefree k that $n(k) = \frac{k\varphi(k)}{(k,h)}$. Thus, according to Artin we should have

$$\sum_{k=1}^{\infty} \frac{\mu(k)}{n(k)} = \sum_{k=1}^{\infty} \frac{\mu(k)(k,h)}{k\varphi(k)} = \prod_{p\nmid h}\left(1 - \frac{1}{p(p-1)}\right) \prod_{p|h}\left(1 - \frac{1}{p-1}\right) = A(h),$$

where in the derivation of the last equality we used the fact that if $f(k)$ is a multiplicative function, then

$$\sum_{k=1}^{\infty} |f(k)| < \infty \Rightarrow \sum_{k=1}^{\infty} f(k) = \prod_p (1 + f(p) + f(p^2) + \cdots).$$

It can be shown that if $g \neq -1$ and g not equal to a square, then

$$\sum_{k=1}^{\infty} \frac{\mu(k)}{n(k)} = c_g A,$$

where $c_g > 0$ is a rational number depending on g. It turns out that numerically the heuristic (7) does not seem to be inconsistent with the numerical data. Let $g \neq -1$ or a square. Hooley [6] proved in 1967 that if the Riemann Hypothesis holds for the number fields $\mathbb{Q}(\zeta_k, g^{1/k})$ with all squarefree k, then we have

$$\mathcal{P}(g)(x) = \sum_{k=1}^{\infty} \frac{\mu(k)}{[\mathbb{Q}(\zeta_k, g^{1/k}) : \mathbb{Q}]} \frac{x}{\log x} + O\left(\frac{x \log\log x}{\log^2 x}\right) \tag{8}$$

and he explicitely evaluated the latter sum. For convenience we denote the assumption on the fields $\mathbb{Q}(\zeta_k, g^{1/k})$ made by Hooley as Hooley's Riemann Hypothesis (HRH) and the error term $x \log\log x/\log^2 x$ as $R(x)$.

5 Hooley's work

For simplicity we assume $g = 2$ and sketch Hooley's proof [6].

A large part of his paper is devoted to proving an estimate for

$$P(x, l) = \#\{p \leqq x : p \text{ splits completely in } k_l\},$$

for which the implied constant in the error term does not depend on l. His result on this is a special case of the result (4) of Lang (proven a few years later) :

$$P(x, l) = \mathrm{Li}(x) + O(\sqrt{x} \log(lx)),$$

where we assume the Riemann Hypothesis for k_l. Another quantity needed is

$$N(x, \eta) = \#\{p \leqq x : p \text{ does not split completely in } k_q,\ q \leqq \eta\}.$$

By inclusion-exclusion we have $N(x, \eta) = \sum_{l'} \mu(l')P(x, l')$, where l' ranges over all divisors of $\prod_{q \leqq \eta} q$. Note that

$$\prod_{q \leqq \eta} q = e^{\sum_{q \leqq \eta} \log q} \leqq e^{2\eta},$$

where we used that $\sum_{q \leqq \eta} \log q \sim \eta$ (which is equivalent with the Prime Number Theorem). Note that $\mathcal{P}(2)(x) = N(x, x-1)$. The problem in using this equation is that our estimates for $P(x, l)$ become meaningless if $l > x^{\frac{1}{2}+\epsilon}$ say. So we are forced to work with $N(x, \eta)$ for η rather smaller than x. We will actually choose $\eta = \frac{\log x}{6}$. Then $l' < e^{2\eta} = x^{1/3}$ and in the inclusion-exclusion we only work with $P(x, l)'s$ for which our estimate is meaningful. Let us introduce a third quantity $M(x, \eta_1, \eta_2) = \#\{p \leqq x : p \text{ splits completely in some } k_q,\ \eta_1 < q < \eta_2\}$. It is easy to see that $N(x, \zeta_1) - M(x, \zeta_1, \zeta_2) - M(x, \zeta_2, \zeta_3) \leqq \mathcal{P}(2)(x) \leqq N(x, \zeta_1)$, that is,

$$\mathcal{P}(2)(x) = N(x, \zeta_1) + O(M(x, \zeta_1, \zeta_2)) + O(M(x, \zeta_2, \zeta_3)) + O(M(x, \zeta_3, x-1)).$$

We will choose $\zeta_1 = \frac{\log x}{6}$, $\zeta_2 = \frac{\sqrt{x}}{\log^2 x}$ and $\zeta_3 = \sqrt{x}\log x$. Note the small gap between ζ_2 and ζ_3. This is the 'hard region' and unfortunately it is out of reach to close it and work say with

$$\mathcal{P}(2)(x) = N(x,\zeta_1) + O(M(x,\zeta_1,\zeta_2)) + O(M(x,\zeta_2,x-1)).$$

Now using the estimte for $P(x,l)$ one easily arrives at

$$\begin{aligned} N(x,\zeta_1) &= \mathrm{Li}(x)\sum_{l'\leqq x^{\frac{1}{3}}}\frac{\mu(l')}{l'\varphi(l')} + O\left(\frac{x}{\log^2 x}\right) \\ &= A\mathrm{Li}(x) + O\left(\frac{x}{\log^2 x}\right) \end{aligned}$$

Again using the estimte for $P(x,q)$ we arrive at

$$M(x,\zeta_1,\zeta_2) \leqq \sum_{\zeta_1\leqq q\leqq\zeta_2} P(x,q) = O\left(\frac{x}{\log^2 x}\right).$$

In the region (ζ_2,ζ_3) we use the fact that the primes p counted by $P(x,q)$ certainly satisfy $p\equiv 1(\mathrm{mod}\ q)$. Thus $P(x,q)\leqq\pi(x,q,1)$. Then using the Brun-Titchmarsh estimate

$$\pi(x,q,1) < \frac{cx}{\log(x/q)}, \quad 1\leqq q < x,$$

we obtain

$$M(x,\zeta_2,\zeta_3) \leqq \sum_{\zeta_2\leqq q\leqq\zeta_3} P(x,q) \leqq \sum_{\zeta_2\leqq q\leqq\zeta_3}\pi(x,q,1) = O(R(x)).$$

Finally, if a prime p is counted by $M(x,\zeta_3,x-1)$, then its order is less than $\frac{x}{\zeta_3} = \frac{\sqrt{x}}{\log x}$ and thus it divides 2^m-1 for some $m < \sqrt{x}/\log x$. Now

$$2^{M(x,\zeta_3,x-1)} < \prod_{p \text{ counted by } M(x,\zeta_3,x-1)} p \leqq \prod_{m<\frac{\sqrt{x}}{\log x}}(2^m-1).$$

Thus

$$M(x,\zeta_3,x-1) < \sum_{m<\frac{\sqrt{x}}{\log x}} m = O\left(\frac{x}{\log^2 x}\right).$$

On gathering all terms we obtain $\mathcal{P}(2)(x) = A\mathrm{Li}(x) + O(R(x))$.

Repeating this argument for arbitrary $g\neq -1$ or a square, we arrive at (8) under HRH, which implies the truth of the quantitative form of Artin's primivite root conjecture and a modified form of the qualitative version.

5.1 Unconditional results

If one asks for unconditional results, the state of affairs is quite appalling. There is no known number g for which $\mathcal{P}(g)$ is known to be infinite ! In 1986, however, Heath-Brown [5] proved (improving on earlier work by Gupta and Ram Murty) a result which implies that there are at most two primes q_1 and q_2 for which $\mathcal{P}(q_1)$ and $\mathcal{P}(q_2)$ are finite and at most three squarefree numbers. (Sarnak called this a result with a theological flavour).

Remember the observation that if $q = 4p + 1$ and $p \equiv 2(\text{mod } 5)$, then $q \in \mathcal{P}(10)$. Sieve theory cannot prove presently that there are infinitely many such primes. However, the so called lower bound sieve method combined with the Chen-Iwaniec switching method gave rise to $\gg \frac{x}{\log^2 x}$ primes $p \leqq x$ such that either $p - 1 = 2^{e_q}$ for some prime q or $p - 1 = 2^{e_{q_1 q_2}}$ with $p^{\alpha} < q_1 < p^{\delta}$ for some $\alpha > \frac{1}{4}$ and $\delta < \frac{1}{2}$. This together with a rather elementary argument, is then enough to establish Heath-Brown's result.

6 Probabilistic model

Let me consider the problem of finding a probabilistic model for a prime p to be such that g is a primitive root mod p. Herewith I mean finding a simple function $f_g(p)$ (if it exists) such that

$$\sum_{\substack{p \in S \\ p \leqq x}} f_g(p) \sim \sum_{\substack{p \in S,\ p \leqq x \\ g \text{ is a primitive root mod } p}} 1, \tag{9}$$

where S has to be a set of primes having a natural density.

There are $\varphi(p-1)$ primitive roots mod p and thus one could try to take $f_g(p) = \varphi(p-1)/(p-1)$ assuming the primitive roots are randomly distributed over $1, \ldots, p-1$. Results of Elliott, Cobeli & Zaharescu, Rudnick & Zaharescu indeed all suggest that the distribution of the primitive roots over $1, \ldots, p-1$ is to a large extent governed by Poisson processes. Thus we would hope to find something like

$$\sum_{p \leqq x} \frac{\varphi(p-1)}{p-1} \sim \mathcal{P}(g)(x), \quad x \to \infty. \tag{10}$$

However, by Hooley's theorem there are, under HRH, g_1 and g_2 such that $\mathcal{P}(g_1)(x) \nsim \mathcal{P}(g_2)(x)$. Thus (10) cannot be true. Let us not be scared by this and try to evaluate $\sum_{p \leqq x} \frac{\varphi(p-1)}{p-1}$. Note that $\frac{\varphi(n)}{n} = \prod_{p|n}(1 - \frac{1}{p}) = \sum_{d|n} \frac{\mu(d)}{d}$. Thus

$$\sum_{p \leqq x} \frac{\varphi(p-1)}{p-1} = \sum_{p \leqq x} \sum_{d|(p-1)} \frac{\mu(d)}{d} = \sum_{d \leqq x} \frac{\mu(d)}{d} \sum_{\substack{p \leqq x \\ p \equiv 1(d)}} 1 = \sum_{d \leqq x} \frac{\mu(d)}{d} \pi(x, d, 1).$$

Let c_1 be a fixed real number. Then the estimate

$$\pi(x,d,a) := \sum_{\substack{p\le x\\ p\equiv a(\bmod\ d)}} 1 = \frac{\mathrm{Li}(x)}{\varphi(d)} + O(xe^{-c_2\sqrt{\log x}})$$

holds uniformly for all integers a and d such that $(a,d)=1$ and $1\le d\le \log^{c_1} x$, with c_2 some positive constant. This is a well-known result due to Siegel and Walfisz. On applying it we obtain

$$\sum_{p\leqq x}\frac{\varphi(p-1)}{p-1} = \mathrm{Li}(x)\sum_{d\leqq \log^{c_1} x}\frac{\mu(d)}{d\varphi(d)} + O\left(\frac{x}{\log^{c_1} x}\right) = A\mathrm{Li}(x) + O\left(\frac{x}{\log^{c_1} x}\right).$$

Recall that Artin's heuristic answer was not always correct because he failed to take into account quadratic interactions. So let us try to incorporate this into our model. In order that g is a primitive root mod p it is necessary that $(g/p)=-1$ and this is a quadratic condition. In particular if g has to be a primitive root it must be a non-square mod p. There are $\frac{p-1}{2}$ of them and then $2\varphi(p-1)/(p-1)$ is the density. Thus it makes sense to try to evaluate the sum in the identity below, which yields, assuming HRH, by a computation a little more complicated than for the sum with the condition $(g/p)=-1$ omitted:

$$2\sum_{\substack{p\leqq x\\ \left(\frac{g}{p}\right)=-1}}\frac{\varphi(p-1)}{p-1} = \mathcal{P}(g)(x) + O(R(x)).$$

In case $h>1$ equality does not always hold. The effect we have to take into account there is that if $(p-1,h)>1$, then

$$g^{\frac{p-1}{(p-1,h)}} \equiv \left(g_0^{\frac{h}{(p-1,h)}}\right)^{p-1} \equiv 1(\bmod\ p)$$

and thus g cannot be a primitive root mod p.

Here is my proposal for $f_g(p)$.

$$f_g(p) = \begin{cases} 2\frac{\varphi(p-1)}{p-1} & \text{if } \left(\frac{g}{p}\right)=-1 \text{ and } (p-1,h)=1;\\ 0 & \text{otherwise}\end{cases} \tag{11}$$

Then, it can be shown assuming HRH that

$$\sum_{p\leqq x} f_g(p) = \mathcal{P}(g)(x) + O(R(x)). \tag{12}$$

To my amazement, this result appeared to be new.

The traditional approach after the failure of the naive heuristic, is to show that is holds true on average :

$$\frac{1}{N}\sum_{g\leqq N}\mathcal{P}(g)(x) = \frac{1}{N}\sum_{p\leqq x} M_p(N),$$

where $M_p(N)$ is the number of integers $g \leqq N$ which are primitive roots modulo p. Clearly $M_g(N) = \varphi(p-1)\{\frac{N}{p} + O(1)\}$. Thus

$$\frac{1}{N}\sum_{g\leqq N}\mathcal{P}(g)(x) = \sum_{p\leqq x}\frac{\varphi(p-1)}{p-1} + O\left(\frac{x^2}{N\log x}\right) + O\left(\frac{x}{\log^D x}\right),$$

provided that $N > x^2(\log^{D-1} x)$. Less trivial unconditional results in this direction were obtained by Goldfeld [4] and Stephens [24].

It turns out to be relatively easy to compute the first sum in (9) with $S = \{p : p \equiv a(\mathrm{mod}\ f)\}$ and one obtains a relatively complicated, but completely explicit, Euler product. The question is whether the true behaviour of $\mathcal{P}(g,f,a)(x) := \#\{p \leqq x : p \in \mathcal{P}(g),\ \ p \equiv a(\mathrm{mod}\ f)\}$ matches up with this.

Lenstra's work [12], which introduced Galois theory into the subject, implies the following result on $\mathcal{P}(g,f,a)(x)$.

Theorem 1 *Let $1 \leq a \leq f$, $(a,f) = 1$. Let σ_a be the automorphism of $\mathbb{Q}(\zeta_f)$ determined by $\sigma_a(\zeta_f) = \zeta_f^a$. Let $c_a(n)$ be 1 if the restriction of σ_a to the field $\mathbb{Q}(\zeta_f) \cap \mathbb{Q}(\zeta_n, g^{1/n})$ is the identity and $c_a(n) = 0$ otherwise. Put*

$$\delta(a,f,g) = \sum_{n=1}^{\infty}\frac{\mu(n)c_a(n)}{[\mathbb{Q}(\zeta_f,\zeta_n,g^{1/n}) : \mathbb{Q}]}.$$

Then, assuming RH for all number fields $\mathbb{Q}(\zeta_f,\zeta_n,g^{1/n})$ with n squarefree,

$$\mathcal{P}(g,f,a)(x) = \delta(a,f,g)\frac{x}{\log x} + O(R(x)).$$

Inspired by the relatively easy Euler product coming from quadratic heuristics, I set out to explicitely evaluate $\delta(a,f,g)$ and managed to find an Euler product for it. This is the same as the one found with quadratic heuristics ! As a consequence we have the following result [17], which generalizes (12):

Theorem 2 *Assume RH for all number fields $\mathbb{Q}(\zeta_f,\zeta_n,g^{1/n})$ with n squarefree. Then*

$$\sum_{\substack{p\leqq x \\ p\equiv a(\mathrm{mod}\ f)}} f_g(p) = \mathcal{P}(g,f,a)(x) + O(R(x)).$$

A rather more insightful derivation of the Euler product for $\delta(a,f,g)$ will appear in [13], as a consequence of a more explicit approach involving quadratic characters in the line of Lenstra's earlier work. This approach is the most powerful sofar in finding explicit Euler products for primitive root densities.

The explicit formula for $\delta(a,f,g)$ allows one to determine the f such that the primes in $\mathcal{P}(g)$ are equidistributed in the various primitive residue classes mod f. This was done earlier by me [16] using an Euler product for $\delta(1,f,g)$ (which is easier to obtain) and some Galois theoretic arguments. Basically f must be a power of two, in case $h=1$. Also $2^\alpha 3^\beta$ can occur, the smallest integer instance being $g=21^7$.

If S is a union of arithmetic progressions, then assuming a sufficiently strong generalisation of RH, it follows by Theorem 2 that equality holds with $f_g(p)$ chosen as in (11). This choice of $f_g(p)$ does not always work for arbitrary sets S having a natural density, for example when one chooses S to be the set of primes that split completely in $\mathbb{Q}(\zeta_3, g^{1/3})$, in case g is not a cube.

7 Indicator function

Let G be a finite cyclic group and put $f(g)=1$ if G is generated by g and $f(g)=0$ otherwise. It was already observed by Landau that

$$f(g)=\frac{\varphi(n)}{n}\sum_{d|n}\frac{\mu(d)}{\varphi(d)}\sum_{\text{ord}\chi=d}\chi(g), \tag{13}$$

where the sum is over the characters χ of G having order d. Using this, one deduces that

$$\mathcal{P}(g)(x)=\sum_{p\le x}\frac{\varphi(p-1)}{p-1}\sum_{d|p-1}\frac{\mu(d)}{\varphi(d)}\sum_{\text{ord}\chi=d}\chi(g).$$

Now write $\mathcal{P}(g)(x)=\sum_{d\le x}S_d(x)$. Note that

$$S_1(x)=\sum_{p\le x}\frac{\varphi(p-1)}{p-1}$$

and that

$$S_2(x)=-(\frac{g}{p})\sum_{p\le x}\frac{\varphi(p-1)}{p-1}.$$

The term $S_d(x)$ can be written as a linear combination of terms

$$\frac{1}{\pi(x)}\sum_{\substack{p\le x,\ p\equiv 1(\text{mod } d)\\ p\equiv 1(\text{mod } \delta)}}\left(\frac{g}{p}\right)_d,$$

where $(\frac{g}{p})_d$ denotes the d-th power residue symbol, cf. [11, p. 111]. In case g is squarefree these terms disappear for every $d>2$ en δ. Hence we expect that

$$\mathcal{P}(g)(x)\sim(S_1(x)+S_2(x))\sim 2\sum_{\substack{p\le x\\ (\frac{g}{p})=-1}}\frac{\varphi(p-1)}{p-1},\quad x\to\infty.$$

A slightly more complicated argument leads to the conjecture

$$\mathcal{P}(g)(x) \sim 2 \sum_{\substack{p \le x,\ (p-1,h)=1 \\ (\frac{g}{p})=-1}} \frac{\varphi(p-1)}{p-1}, \quad x \to \infty,$$

where g is not required to be squarefree. The problem in proving this (and hence a quantitative Artin's primitive root conjecture) is that we have 'too many error terms' and can only achieve this if one can establish that there is sufficient cancellation, which is out of reach presently.

The analog of the latter conjecture for the general index t case (cf. section 8.6) is far from easy to infer by ad hoc methods (as we managed to do for $t = 1$ in the previous section), but can be merely computed using the approach sketched here. Furthermore, under the usual RH proviso this conjecture can then be shown to be true [18, Theorem 1].

The unconditional results on Artin's conjecture on average in [4, 24] where obtained on using the indicator function and estimates for character sums.

We now give another application of the identity (13). Let $\mathbb{F}_q$ be the finite field with $q = p^n$ elements. Recall that its multiplicative group is cyclic. Let $a(x)$ be a polynomial in the polynomial ring $\mathbb{F}_p[x]$. We are interested in the number of irreducible polynomials $p(x) \in \mathbb{F}_p[x]$ such that $a(x)$ generates $(\mathbb{F}_p[x]/(p(x))^*$. Recall that if $\deg p(x) = n$ then $\mathbb{F}_p[x]/(p(x))$ is isomorphic with $\mathbb{F}_{p^n}$, where the isomorphism is given explicitly as follows: for $h(x) \in \mathbb{F}_p[x]$, we write, using the Euclidean algorithm,

$$h(x) = p(x)q(x) + r(x), \text{ with either } r(x) = 0 \text{ or } 0 \le \deg r < \deg p = n.$$

Let $\rho \in \mathbb{F}_{p^n}$ be a root of $p(x)$. Then

$$h(\rho) = r(\rho) = a_0 + a_1\rho + \cdots + a_{n-1}\rho^{n-1},\ a_i \in \mathbb{F}_p$$

describes all the elements of $\mathbb{F}_{p^n}$. Thus, $a(x)$ generating $(\mathbb{F}_p[x]/(p(x)))^*$ is equivalent to $a(\rho)$ generating $\mathbb{F}_{p^n}^*$.

Hence, to count the number of irreducible $p(x)$ of degree n for which $a(x)$ is a generator is tantamount to counting the number of ρ of degree n for which $a(\rho)$ generates $\mathbb{F}_{p^n}^*$. Indeed, since each $p(x)$ has n roots, we find that the number of irreducible polynomials of degree n in $\mathbb{F}_p[x]$ such that $a(x)$ generates $(\mathbb{F}_p[x]/(p(x)))^*$ is equal to the number of elements $\rho \in \mathbb{F}_{p^n}$ of degree n such that $a(\rho)$ generates $\mathbb{F}_{p^n}^*$ divided by n. The latter number, is on using (13), equal to

$$\frac{1}{n} \sum_{\substack{\rho \in \mathbb{F}_{p^n} \\ \deg\rho = n}} \frac{\varphi(p^n-1)}{p^n-1} \sum_{d \mid p^n-1} \frac{\mu(d)}{\varphi(d)} \sum_{\mathrm{ord}\chi = d} \chi(a(\rho)).$$

By inclusion-exclusion we find that the number of irreducible polynomials of degree n over $\mathbb{F}_q[x]$ equals

$$\frac{1}{n}\sum_{\substack{\rho\in\mathbb{F}_{p^n}\\ \deg\rho=n}} 1 = \frac{1}{n}\sum_{d|n}\mu(d)p^{\frac{n}{d}} = \frac{p^n}{n} + O\left(\frac{p^{\frac{n}{2}}}{n}\right).$$

Thus the contribution from the main term (the term with $d=1$) is

$$\frac{p^n + O(p^{\frac{n}{2}})}{n}\frac{\varphi(p^n-1)}{p^n-1},$$

and the error term is

$$\frac{\varphi(p^n-1)}{n(p^n-1)}\sum_{\substack{d|p^n-1\\ d>1}}\frac{\mu(d)}{\varphi(d)}\sum_{\mathrm{ord}\chi=d}\sum_{\substack{\rho\in\mathbb{F}_{p^n}\\ \deg\rho=n}}\chi(a(\rho)).$$

On applying Weil's estimate for character sums and imposing the appropriate conditions on $a(x)$, the truth of Artin's conjecture for this situation is then established, cf. [23].

In [8] the authors show that in case $a(x) = x^m + c$ there is no need to invoke such a powerful result as Weil's estimates and the required estimate for the error term can be obtained on using Gauss sums over finite fields.

8 Some variations of Artin's problem

8.1 Function fields

It is a well-known principle in number theory that the analogue of problems for number fields in the function field setting is usually easier to establish and indeed sometimes a source of inspiration for making progress in the number field setting. Thus for example the ABC conjecture only arose in the course of considerations in function fields and its fundamental importance has been realized only quite recently. In the function field setting the ABC conjecture is easily proved.

In the face of our failure to prove the Artin primitive root conjecture it thus seems natural to try resolve it in the function field setting. This was accomplished by a PhD student of Hasse, Bilharz [2], under the proviso that the RH for the zeta function corresponding to the function field would hold true. Indeed, much later this was proved by Weil and hence Artin's primitive root conjecture is true for function fields ! Bilharz result has been subsequently generalized by Hsu [7] to Carlitz modules (a particular case of Drinfeld modules of rank one).

8.2 Elliptic Artin

An excellent introduction to elliptic Artin can be found in [22] and we quote it here almost verbatim. For references we refer to that paper.

In 1976 Lang and Trotter formulated elliptic analogues of the Artin primitive root conjecture. Suppose E is an elliptic curve over $\mathbb{Q}$ with a rational point of infinite order. A natural question is : How often does the prescribed point generate $E(\mathbb{F}_p)$, the group of points (mod p) ?

More precisely, let a be a rational point of infinite order. The problem is to determine the density of primes p for which $E(\mathbb{F}_p)$ is generated by $\bar{a}$, the reduction of $a(\mathrm{mod}\ p)$. (Here in addition to primes of bad reduction, we may need to exclude primes dividing the denominators of the coordintes of a.) Such a point will be called a primitive point for these primes. Lang and Trotter conjectured that the density of primes for which a is a primitive point always exists. This is the elliptic analogue of Artin's primitive roots conjecture.

Of course, situations may (and do) arise when the density is zero. In such a case the set of primes for which a is a primitive point is finite (and often empty). If this is the case, then it is so for obvious reasons. So the statement is that if it is not the case for obvious reasons that there are only infinitely many, then there are infinitely many and they have a density.

In considering the elliptic analogue, we see that two assertions are being made about a prime p for which a is a primitive point :
1) $E(\mathbb{F}_p)$ is cyclic.
2) $\langle\bar{a}\rangle \cong E(\mathbb{F}_p)$,
where with $\langle\bar{a}\rangle$ we denote the subgroup in $E(\mathbb{F}_p)$ generated by the reduction mod p of the point a. Is it even true that $E(\mathbb{F}_p)$ is cyclic infintely often ? It was Serre who pointed out the relevance of this question. In a course at Harvard in the fall of 1976, he proved that Hooley's method can be adapted to show that the set of primes p with $E(\mathbb{F}_p)$ cyclic is

$$\delta_E \frac{x}{\log x} + O\left(\frac{x}{\log x}\right),$$

where

$$\delta_E = \sum_{k=1}^{\infty} \frac{\mu(k)}{[\mathbb{Q}(E[k]) : \mathbb{Q}]}.$$

Here $\mathbb{Q}(E[k])$ is the number field obtained by adjunction of the coordinates of the k-division points of the elliptic curve E. It is the analogue of the Kummerian extension $\mathbb{Q}(\zeta_k, g^{1/k})$. As noted by Serre, $\delta_E > 0$ if and only if E has an irrational 2-division point. That is, if E has the Weierstrauss model, $y^2 = x^3+cx+d$, we require that x^3+cx+d has an irrational root. In the non-CM case, it is still unknown if $E(\mathbb{F}_p)$ is cyclic for a positive density of primes whenever E has an irrational 2-division point.

8.3 Even order

Let $g \geq 2$ be an integer. As we saw in the introduction the question whether $\text{ord}_p(g) = p - 1$ for a prime p, is equivalent with the period of the decimal expansion of $1/p$ in base g being maximal. Krishnamurty wondered how often the period in these decimal expansions are even. This can be attacked by algebraic methods similar to those sketched above.

We say that a prime p *divides a sequence* of integers, if it divides at least one term of the sequence. Notice that

$$\text{ord}_g(p) \text{ is even} \iff g^x \equiv -1(\text{mod } p) \text{ has a solution} \iff p | \{g^k + 1\}_{k=0}^{\infty}.$$

There is an unique $j \geq 1$ such that $p \equiv 1 + 2^j (\text{mod } 2^{j+1})$. Now $\text{ord}_g(p)$ is even if and only if

$$g^{\frac{p-1}{2^j}} \not\equiv 1(\text{mod } p).$$

It is more natural to consider the g such that $\text{ord}_g(p)$ is odd, that is the set $\{p : 2 \nmid \text{ord}_g(p)\}$. Let

$$P_j = \{p : p \equiv 1 + 2^j(\text{mod } 2^{j+1}),\ 2 \nmid \text{ord}_g(p)\}.$$

Note that

$$P_j = \{p : p \equiv 1 + 2^j(\text{mod } 2^{j+1}),\ g^{\frac{p-1}{2^j}} \equiv 1(\text{mod } p)\}.$$

Now P_j consists of the primes p that split completely in $\mathbb{Q}(g^{1/2^j}, \zeta_{2^j})$, but do not split completely in $\mathbb{Q}(g^{1/2^j}, \zeta_{2^{j+1}})$. Since $\{p : \ 2 \nmid \text{ord}_g(p)\} = \cup_{j=1}^{\infty} P_j$, and the P_j are disjoint, one expects that the natural density of the set $\{p : \ 2 \nmid \text{ord}_g(p)\})$ equals $\sum_{j=1}^{\infty} \delta(P_j)$, where $\delta(P_j) = \lim_{x \to \infty} P_j(x)/\pi(x)$. Now

$$\sum_{j=1}^{\infty} \delta(P_j) = \sum_{j=1}^{\infty} \left(\frac{1}{[\mathbb{Q}(g^{1/2^j}, \zeta_{2^j}) : \mathbb{Q}]} - \frac{1}{[\mathbb{Q}(a^{1/2^j}, \zeta_{2^{j+1}}) : \mathbb{Q}]} \right).$$

On taking $g = 2$ one expects, on computing the field degrees involved, that the density of prime divisors of the sequence $\{2^k + 1\}_{k=0}^{\infty}$ equals

$$\delta = (\frac{1}{2} - \frac{1}{4}) + (\frac{1}{8} - \frac{1}{8}) + \sum_{j \geq 3} (\frac{4}{4^j} - \frac{2}{4^j}) = \frac{7}{24}.$$

With natural density replaced by Dirichlet density, this result was proved by Hasse in 1966. The natural density of sequences of the form $\{a^k + b^k\}_{k=0}^{\infty}$ with a and b integers was computed independently along similar lines by Ballot, Odoni and Wiertelak, see e.g. [1]. For the 'canonical' sequence $\{a^k + b^k\}$ the density is $2/3$.

Note that this result, unlike Hooley's, is unconditional. Both involve infinite towers of field extensions, but the one appearing in this problem is so sparse (in the sense that the degrees of the fields quickly increase) that the profusion of error terms that occurs in the Artin primitive root conjecture, does not arise and one can afford working with the relatively large error terms in $\pi_K(x)$ that are conditionally known.

Instead for asking when $2|\mathrm{ord}_p(g)$, one can ask for $m|\mathrm{ord}_p(g)$ or for $(m, \mathrm{ord}_p(g)) = 1$ etc.. These questions, with sharp error terms, are dealt with in various papers of Wiertelak.

8.4 Divisors of second order recurrences

The sequences $\{a^k+b^k\}_{k=0}^{\infty}$ are special cases of sequences satisfying a second order linear recurrence and one can wonder about prime divisors of such recurrences in general. A general linear second order recurrence has the form $u_{n+2} = cu_{n+1} + du_n$. Once the values of u_0 and u_1 are given, it is uniquely determined. The *characteristic polynomial* associated to the recurrence is defined as $X^2 - cX - d$. If this polynomial is reducible over $\mathbb{Q}$, then the recurrence has the form $R = \{ca^k - db^k\}_{k=0}^{\infty}$, finding the prime divisors amounts to finding the prime divisors of the sequence with general term $(a/b)^k - d/c = \alpha^k - \beta$. Now p is a prime divisor of R, with at most finitely many exceptions, if and only if $\mathrm{ord}_p(\beta)|\mathrm{ord}_p(\alpha)$. This can be phrased differently by saying that, with at most finitely many exceptions, p divides R if and only if the subgroup mod p generated by β, $\langle\beta\rangle$, is contained in the subgroup $\langle\alpha\rangle$. One can then base a two variable Artin conjecture on this, of which the qualitative form says that the sequence R should have infinitely many prime divisors. This form is not too difficult to prove, as was done by Polya around 1920, cf. [20]. To investigate the quantitative form, we put $S(x) = \sum_{p\le x,\ p|R} 1$. Then

$$S(x) = \sum_{p\le x} \sum_{\substack{w|p-1 \\ \mathrm{ord}_p(\beta)|\mathrm{ord}_p(\alpha)=(p-1)/w}} 1 = \sum_{w\le x-1} \sum_{\substack{p\le x \\ p\equiv 1(\mathrm{mod}\ w) \\ \mathrm{ord}_p(\beta)|\mathrm{ord}_p(\alpha)=(p-1)/w}} 1.$$

Now $p \equiv 1(\mathrm{mod}\ w)$ and $\mathrm{ord}_p(\beta)|\mathrm{ord}_p(\alpha) = (p-1)/w$ if and only if p splits completely in $\mathbb{Q}(\alpha^{1/w}, \beta^{1/w}, \zeta_w)$ and p does not split completely in any of the fields $\mathbb{Q}(\alpha^{1/wq}, \beta^{1/w}, \zeta_{wq})$, with q a prime. Denote the degree of the latter field by $n(qw)$. Then, as $x \to \infty$, the limit $S(x)/\pi(x)$ tends to

$$\sum_{w=1}^{\infty}\sum_{k=1}^{\infty} \frac{\mu(k)}{n(kw)}, \tag{14}$$

provided a sufficiently strong generalisation of RH is true. The sum (14) equals a rational number times the *Stephens constant*

$$S = \prod_p \left(1 - \frac{p}{p^3-1}\right) \approx 0.5759599688929..$$

This result is due to Moree and Stevenhagen [20] and improves on earlier work by Stephens [25]. Stephens evaluation of (14) (in the case where both α and β are integers) is not always correct and corrected in [20].

Note that the problem considered here is a variant of Artin's primitive root conjecture that is more complicated than the original one, whereas the variant discussed in section 8.3 is easier, although both deal with second order recurrences.

8.5 Lenstra's work

The most far reaching generalisation of the Artin primitive root conjecture was considered by Lenstra [12], in the context of his research on Euclidean number fields : Let K be a global field and F a finite normal extension of K. Let C be a subset of $\mathrm{Gal}(F/K)$ which is stable under conjugation and let d be a positive integer (coprime to the characteristic of K in the case of a function field). Consider a finitely generated subgroup W of K^* which has, modulo torsion, rank $r \geqq 1$, and let M be the set of prime ideals $\mathfrak{P}$ of K satisfying

1. the Artin symbol $(\mathfrak{P}, F/K) \subseteq C$
2. the normalized exponential valuation attached to $\mathfrak{P}$ satisfies $\mathrm{ord}_{\mathfrak{P}}(w) = 0$ for all $w \in W$.
3. if $\psi : W \to \bar{K}_{\mathfrak{P}}^*$ is the natural map, then the index of $\psi(W)$ in $\bar{K}_{\mathfrak{P}}^*$ divides d.

Here $\bar{K}_{\mathfrak{P}}^*$ denotes the multiplicative group of $\bar{K}_{\mathfrak{P}}$, the residue class field at $\mathfrak{P}$. Lenstra conjectured that M has a density. He also obtained necessary and sufficient conditions for this density to be nonzero, under a sufficiently strong generalisation of RH. His main result implies Theorem 1 for example. Lenstra applied it to show, under the usual RH proviso, that certain principal ideal rings are euclidean.

8.6 Varia

We mention some other generalisations without going in to them. Note that g is a primitive root mod p if the subgroup generated by $\bar{g}$, the reduction of g mod p, equals $(\mathbb{Z}/p\mathbb{Z})^*$. If $g_1, \cdots, g_r$ are rational numbers, one can wonder about the density of primes $p \le x$ such that $\langle \bar{g}_1, \cdots, \bar{g}_r \rangle$ is the full multiplicative group $(\mathbb{Z}/p\mathbb{Z})^*$, see e.g. [3], or about the density such that each g_i, $1 \le i \le r$, is a primitive root [15].

Fix an integer t. For the primes $p \equiv 1 (\mathrm{mod}\ t)$ we can ask for the subset of them such that $\langle \bar{g} \rangle$ generates a subgroup of index t in $(\mathbb{Z}/p\mathbb{Z})^*$, see [18] and the references there.

The definition of primitive roots was extended by Carmichael to that of *primitive* λ*-roots* for composite moduli n, which are integers with the maximal order modulo n. Let $N_g(x)$ be the number of natural numbers up to x for which g is a primitive λ-root. One might wonder then, in analogy with Artin's conjecture, whether $N_g(x) \sim B(g)x$ for some positive constant $B(g)$ depending on g, perhaps with some exceptional values of g. Li [14] provides results suggesting that $N_g(x)/x$ does not typically tend to a limit.

Recall that g is a primitive root mod p if p does not split completely in any of the fields k_q. Goldstein considered the more general problem where the fields k_q are replaced by normal number fields L_q and made a conjecture which was subsequently disproved by Weinberger. Goldstein then proved a restrictive version of his conjecture, which result was improved upon by Murty [21].

Acknowledgement. I'd like to thank Zeev Rudnick for inviting me to give the talk this manuscript is based on and Jilyana Cazaran (Louisiana State University) for kindly typing these lecture notes from my handwritten ones. My trip to Oberwolfach was paid for from Prof. Opdam's NWO-PIONEER fund.

References

[1] C. Ballot, Density of prime divisors of linear recurrences, *Mem. Amer. Math. Soc.* **115** (1995), no. 551, viii+102 pp..

[2] H. Bilharz, Primdivisoren mit vorgegebener Primitivwurzel, *Math. Ann.* **114** (1937), 476-492.

[3] L. Cangelmi and F. Pappalardi, On the r-rank Artin conjecture. II, *J. Number Theory* **75** (1999), 120-132.

[4] M. Goldfeld, Artin's conjecture on the average, *Mathematika* **15** (1968), 223-226.

[5] D.R. Heath-Brown, Artin's conjecture for primitive roots, Quart. J. Math. Oxford **37** (1986), 27-38.

[6] C. Hooley, Artin's conjecture for primitive roots, *J. Reine Angew. Math.* **225** (1967), 209-220.

[7] C.-N. Hsu, On Artin's conjecture for the Carlitz module, *Compositio Math.* **106** (1997), 247-266.

[8] E. Jensen and M.R. Murty, Artin's conjecture for polynomials over finite fields, Number theory, 167-181, Trends Math., Birkhuser, Basel, 2000 (also available at `http://www.mast.queensu.ca/~murty/`).

[9] S. Lang, On the zeta function of number fields, Invent. Math. **12** (1971), 337–345.

[10] D.H. Lehmer and Emma Lehmer, Heuristics, anyone ?, in "Studies in Mathematical Analysis and Related Topics," pp. 202-210, Stanford Univ. Press, Stanford, CA, 1962.

[11] F. Lemmermeyer, *Reciprocity laws. From Euler to Eisenstein.*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000. xx+487 pp..

[12] H.W. Lenstra, jr., On Artin's conjecture and Euclid's algorithm in global fields, *Invent. Math.* **42** (1977), 202-224.

[13] H.W. Lenstra, jr., P. Moree and P. Stevenhagen, Character sums for primitive root densities, in preparation.

[14] Shuguang Li, Artin's conjecture on average for composite moduli, *J. Number Theory* **84** (2000), 93-118.

[15] K.R. Matthews, A generalisation of Artin's conjecture for primitive roots, *Acta Arith.* **29** (1976), 113-146.

[16] P. Moree, Uniform distribution of primes having a prescribed primitive root, *Acta Arith.* **89** (1999), 9-21.

[17] P. Moree, Primes in arithmetic progression having a prescribed primitive root, *J. Number Theory* **78** (1999), 85-98.

[18] P. Moree, Asymptotically exact heuristics for (near) primitive roots, *J. Number Theory* **83** (2000), 155-181.

[19] P. Moree, Approximation of singular series and automata, *Manuscripta Math.* **101** (2000), 385-399.

[20] P. Moree and P. Stevenhagen, A two variable Artin conjecture, to appear in Journal of Number Theory.

[21] M.R. Murty, An analogue of Artin's conjecture for abelian extensions, *J. Number Theory* **18** (1984), 241-248.

[22] M.R. Murty, Artin's conjecture and elliptic analogues, Sieve methods, exponential sums, and their applications in number theory (Cardiff, 1995), 325–344, London Math. Soc. Lecture Note Ser., 237, Cambridge Univ. Press, Cambridge, 1997.

[23] F. Pappalardi and I. Shparlinski, On Artin's conjecture over function fields, *Finite Fields Appl.* **1** (1995), 399-404.

[24] P.J. Stephens, An average result for Artin's conjecture, *Mathematika* **16** (1969), 178-188.

[25] P.J. Stephens, Prime divisors of second-order linear recurrences. I, II, *J. Number Theory* **8** (1976), 313-332; 333-345.

[26] P. Stevenhagen and H.W. Lenstra, jr., Chebotarev and his density theorem, *Math. Intelligencer* **18** (1996), 26-37.

Pieter Moree
Faculty of Mathematics,
University of Amsterdam
e-mail address: moree@wins.uva.nl
homepage: http://web.inter.nl.net/hcc/J.Moree/