1. Artin’s primitive root conjecture

Given a prime $p$, a primitive root modulo $p$ is a generator of the cyclic group $(\mathbb{Z}/p\mathbb{Z})^\times$ of invertible residues modulo $p$, that is its order in the multiplicative group is $p - 1$, the maximal possible value. Gauss seemed to have observed that 10 occurs often as a primitive root, for instance in 39 of the first 100 primes. Likewise, 2 is a primitive root for 41 of the first 100 primes.

**Exercise 1.**

i) If $p \nmid 10$ then $\frac{1}{p}$ has a periodic decimal expansion, e.g. $\frac{1}{7} = 0.142857142857\ldots$ has period 6, $\frac{1}{11} = 0.0909\ldots$ has period 2.

ii) The order of 10 mod $p$ is the length of the minimal period.

**Exercise 2.** If $p$ is a prime of the form $p = 4p' + 1$ where $p'$ is also prime, then 2 is a primitive root modulo $p$.

The problem with this approach is that we do not know that there are infinitely many primes of this form.

It is clear that a perfect square cannot be a primitive root if $p > 2$. In 1927, Artin conjectured that for any integer $g \neq -1, \square$, there are infinitely many prime $p$ for which $g$ is a primitive root modulo $p$. A quantitative version is that

**Conjecture.** If $g \neq -1$ or a perfect square, then there is $C(g) > 0$ such that

$$
\#\{p \leq x : g \text{ is a primitive root modulo } p\} \sim C(g) \frac{x}{\log x}. \quad x \to \infty
$$

The constant $C(g)$ is known; for the simple case $g = 2$, we have

$$
C(2) = \prod_{q \text{ prime}} \left(1 - \frac{1}{q(q - 1)}\right) = 0.3739\ldots
$$

In 1967, Hooley [1] proved Artin’s conjecture, assuming the Generalized Riemann Hypothesis (GRH) for the Dedekind zeta function of a certain infinite family of number fields (Kummer extensions). Below we will explain his argument. For further reading, see the surveys of Murty [3] and Moree [2].

---

*Date: June 17, 2015.*
2. Hooley’s approach

From now on, we will take \( g = 2 \), so we want primes \( p \) for which 2 is a primitive root modulo \( p \). Set

\[ N(x) := \{ p \leq x \text{ prime, } p \nmid 2, \ 2 \text{ is a primitive root modulo } p \} \]

and we want to show that \( \#N(x) \sim C(2)x/\log x \).

We observe that for \( p \nmid 2 \), the condition 2 is a primitive root modulo \( p \) is equivalent to the condition

\[ \forall \text{ prime } q \text{ s.t. } q | p - 1, \ 2^{(p-1)/q} \neq 1 \mod p \]

that is we have \( \text{not}(R(p; q)) \) for all primes \( q \), where \( R(p; q) \) is the condition

\[ R(p; q) : \quad p = 1 \mod q \quad \text{and} \quad 2^{(p-1)/q} = 1 \mod p \]

For \( z < x \), set

\[ N'(x, z) := \{ 2 < p \leq x : \forall \text{ prime } q \leq z, \ \text{not } R(p; q) \} \]

so that

\[ N(x) = N'(x, x - 1) \]

and

\[ N(x) \subseteq N'(x, z) \]

for all \( z < x \).

We also set, for \( w < z \),

\[ N''(x; w, z) = \{ 2 < p \leq x : \exists \text{ prime } w < q \leq z, \ \text{s.t. } R(p, q) \text{ holds} \} \]

Then clearly

\[ N'(x; z) \subseteq N(x) \cup N''(x; z, x) \]

and hence

\[ \#N(x) = \#N'(x; z) + O \left( \#N''(x; z, x) \right) \]

We will take \( z = \log x/6 \) and show

\[ \#N'(x; \frac{1}{6} \log x) = C(2) \frac{x}{\log x} + O \left( \frac{x}{(\log x)^2} \right) \]

and

\[ \#N''(x; \frac{1}{6} \log x, x) \ll \frac{x}{(\log x)^2} \log \log x \]

which will give our Theorem.
3. Evaluating $\#N'(x; \frac{1}{6} \log x)$

Let

$$P(z) := \prod_{2 < p \leq z} p \approx x^{1/3}$$

if $z \approx \frac{\log x}{6}$. For $d \mid P(z)$ (necessarily squarefree), set

(5) \hspace{1cm} P(x; d) := \# \{ p \leq x : R(p; q) \text{ holds } \forall \text{ prime } q \mid d \}

(for $d = 1$ there is no condition).

**Theorem 3.1.** Assume the Generalized Riemann Hypothesis. Then for squarefree $d$,

$$P(x; d) = \frac{1}{n(d)} \Li(x) + O \left( x^{1/2} \log(dx) \right)$$

where $n(d) = d \varphi(d)$.

To explain Theorem 3.1, we will need a major bit of input from algebraic number theory, the explanation of which is deferred to later on.

By the sieve of Eratosthenes,

$$\#N'(x; z) = \sum_{d \mid P(z)} \mu(d) P(x; d)$$

and inputing Theorem 3.1 gives

$$\#N'(x; z) = \sum_{d \mid P(z)} \mu(d) \left( \frac{\Li(x)}{d \varphi(d)} + O \left( x^{1/2} \log(dx) \right) \right)$$

$$= C(2) \left( 1 + O \left( \frac{1}{z} \right) \right) \Li(x) + O \left( x^{1/2} \log x \sum_{d \mid P(z)} 1 \right)$$

$$= C(2) \left( 1 + O \left( \frac{1}{z} \right) \right) \Li(x) + O \left( x^{1/2} \log x \cdot 2^z \right)$$

because

$$\sum_{d \mid P(z)} \frac{1}{d \varphi(d)} = \prod_{q \mid P(z) \text{ prime}} \left( 1 - \frac{1}{q(q-1)} \right) = C(2) \left( 1 + O \left( \frac{1}{z} \right) \right)$$

Taking into account $z \approx (\log x)/6$, so that $2^z \ll x^{1/3}$, we get

$$\#N'(x; \frac{\log x}{6}) = C(2) \frac{x}{\log x} + O \left( \frac{x}{(\log x)^2} \right)$$

giving (3).
4. Estimating $\#N''(x; \frac{1}{6} \log x, x)$

To bound $\#N''(x; \frac{1}{6} \log x, x)$, which is the number of primes $2 < p \leq x$ for which there is some prime $z < q < x$ such that $R(p; q)$ holds, that is such that $p = 1 \mod q$ and $2^{(p-1)/q} = 1 \mod p$, we use a union bound

$$\#N''(x; \frac{1}{6} \log x, x) \leq \#N''(x; \frac{1}{6} \log x, \frac{x}{\log x}) + \#N''(x; \frac{x}{\log x}, \frac{x}{\log x}) + \#N''(x; x, \sqrt{x} \log x)$$

where the summands put conditions on the existence of a prime $q$ which is “small” (that is $\frac{(\log x)}{6} < q < \frac{x}{\log x}$), “medium”, meaning $\frac{x}{\log x} < q < \sqrt{x} \log x$, and “large”, meaning $\sqrt{x} \log x < q < x$. We will apply separate considerations for each summand.

4.1. Small primes. For the small primes, we use a union bound together with Theorem 3.1 (so we use GRH here)

$$\#N''(x; \frac{1}{6} \log x, \frac{x}{\log x}) \leq \sum_{\frac{1}{6} \log x < q \leq \frac{x}{\log x}} P(x; q) \lesssim \sum_{\frac{1}{6} \log x < q \leq \frac{x}{\log x}} \left( \frac{x}{q(q-1) \log x} + \sqrt{x} \log x \right) \leq \frac{x}{\log x} \sum_{\frac{1}{6} \log x < q \leq \frac{x}{\log x}} \frac{1}{q^2} + \sqrt{x} \log x \cdot \pi(\frac{x}{\log x^2}) \ll \frac{x}{(\log x)^2}$$

which is an admissible bound.

4.2. Medium primes. To handle the contribution of “medium” primes $q$, we replace the condition $p = 1 \mod q$ and $2^{(p-1)/q} - 1 \mod p$ with just the first condition, so that

$$P(x; q) \leq \#\{p \leq x : p = 1 \mod q\} = \pi(x; q, 1)$$

Now we use the Brun Titchmarsh theorem, which gave a good upper bound for the number of primes in an arithmetic progression with large modulus:

$$\pi(x; q, 1) \leq \frac{x}{\varphi(q) \log(x/q)}$$

Taking into account that we are in the range that $q$ is close to $\sqrt{x}$ gives

$$P(x; q) \leq \pi(x; q, 1) \ll \frac{x}{q \log x}$$
and hence we find
\[
\# \mathcal{N}''(x; \frac{\sqrt{x}}{(\log x)^2}, \sqrt{x} \log x) \leq \sum_{\frac{\sqrt{x}}{(\log x)^2} < q \leq \sqrt{x} \log x} P(x; q) \\
\ll \sum_{\frac{\sqrt{x}}{(\log x)^2} < q \leq \sqrt{x} \log x} \frac{x}{q \log x} \\
= \frac{x}{\log x} \sum_{\frac{\sqrt{x}}{(\log x)^2} < q \leq \sqrt{x} \log x} \frac{1}{q}
\]

To estimate the sum over \( q \) (which are prime), we use Merten’s theorem
\[
\sum_{q < y \text{ prime}} \frac{1}{q} = \log \log y + C + O\left(\frac{1}{\log y}\right)
\]
which gives
\[
\sum_{\frac{\sqrt{x}}{(\log x)^2} < q \leq \sqrt{x} \log x} \frac{1}{q} \ll \frac{\log \log x}{\log x}
\]
and therefore
\[
\# \mathcal{N}''(x; \frac{\sqrt{x}}{(\log x)^2}, \sqrt{x} \log x) \ll \frac{x \log \log x}{(\log x)^2}
\]
which is an admissible bound.

4.3. Large primes. Finally, we need to bound the contribution of “large” primes, that is \( \sqrt{x} \log x < q < x \).

We note that the primes \( p \) counted by \( \mathcal{N}''(x; \sqrt{x} \log x, x) \) satisfy \( q \mid p - 1 \) and \( 2^{(p-1)/q} = 1 \mod p \) and that in our range of \( q \)'s, the fraction \( m := (p - 1)/q \leq \sqrt{x} / \log x \). Thus these \( p \)'s must all divide some \( 2^m - 1 \) for some \( m \leq \sqrt{x} / \log x \), so that they are at most the number of prime divisors of the product of these factors \( 2^m - 1 \):
\[
\# \mathcal{N}''(x; \sqrt{x} \log x, x) \leq \omega\left( \prod_{m \leq \sqrt{x} / \log x} (2^m - 1) \right)
\]
Using the crude bound \( \omega(n) \leq \log_2 n \) gives
\[
\omega\left( \prod_{m \leq \sqrt{x} / \log x} (2^m - 1) \right) \ll \sum_{m \leq \sqrt{x} / \log x} m \ll \frac{x}{(\log x)^2}
\]
giving
\[
\# \mathcal{N}''(x; \sqrt{x} \log x, x) \ll \frac{x}{(\log x)^2}
\]
which is an admissible bound.
5. Algebraic number theory

We now give some background in algebraic number theory needed for understanding Theorem 3.1.

5.1. Splitting of primes. Given a number field $K$, that is a finite extension of the rationals, a principal goal of algebraic number theory is to understand the splitting of rational primes in the ring of integers of $K$. Here the ring of integers of $K$ is the set of all algebraic integers contained in $K$, namely $\alpha \in \overline{\mathbb{Q}}$ which are roots of a monic polynomial with integer coefficients.

Example: The Gaussian integers $K = \mathbb{Q}(\sqrt{-1})$. Here the ring of integers is $O_K = \mathbb{Z}[\sqrt{-1}]$, the Gaussian integers, which is a Euclidean ring, hence has unique factorization into irreducibles. To find what are the irreducibles of $\mathbb{Z}[\sqrt{-1}]$, we check the factorization of rational primes. The result is that there are three possibilities:

- The split case $p = 1 \text{ mod } 4$, in which case $p = \pi \overline{\pi}$ splits as a product of two nonassociate primes of $K$, so that if $\pi = a + ib$ then $p = a^2 + b^2$.
- The inert case $p = 3 \text{ mod } 4$, in which case $p$ remains irreducible in $K$.
- The ramified case $p = 2$ which factors as $2 = -i(1 + i)^2$.

For other number fields, even quadratic, there is no longer unique factorization into irreducibles and what replaces it is the unique factorization of ideals in the ring of integers $O_K$ into prime ideals. Recall an ideal $P \subset O_K$ is prime if $a \cdot b \in P$ iff $a \in P$ or $b \in P$.

Given a rational prime, we can uniquely factor the principal ideal $pO_K$ as

$$pO_K = P_1^{e_1} \cdots P_g^{e_g}$$

where $P_j$ are distinct prime ideals. Defining the norm of a nonzero ideal $(0) \neq I \subset O_K$ as $N(I) = \#O_K/I$ (which is finite if $I \neq (0)$), one has

$$N(P_j) = p^{f_j}$$

for some $f_j \geq 1$, called the degree of the prime ideal $P_j$, and there is a conservation law involved in the numbers here:

$$\sum_{j=1}^{g} e_j f_j = [K : \mathbb{Q}]$$

We say that a rational prime $p$ splits completely in $K$ if all $e_j = 1 = f_j$, so that

$$pO_K = P_1 \cdots P_n, \quad n = [K : \mathbb{Q}]$$

is a product of degree one primes.

5.2. Examples. i) In the case of the Gaussian integers, the split primes are precisely $p = 1 \text{ mod } 4$.

ii) Another important example are the cyclotomic fields $Z_q = \mathbb{Q}(\zeta_q)$, where $\zeta_q$ is a primitive $q$-th root of unity. These have degree $[Z_q : \mathbb{Q}] = \varphi(q)$, and the split primes are precisely those such that $p = 1 \text{ mod } q$. 
iii) The example we shall need is that of a Kummer extension, specifically for prime $q > 2$, let

$$K_q = \mathbb{Q}(\sqrt[1/q]{2}, \zeta_q)$$

be the splitting field of the polynomial $x^q - 2$ over the rationals, where $\zeta_q$ is a primitive $q$-th root of unity. For $q$ prime (odd),

$$[K_q : \mathbb{Q}] = q(q - 1)$$

since $K_q$ is obtained from the rationals by the sequence $\mathbb{Q} \subset \mathbb{Q}(\sqrt[2]{2}) \subset \mathbb{Q}(\sqrt[2]{2})((\zeta_q))$ and assuming the extension $\mathbb{Q}(\sqrt[2]{2})$, whose degree is $q$, is disjoint from the cyclotomic extension $\mathbb{Q}(\zeta_q)$, whose degree is $\varphi(q) = q - 1$, we obtain $[K_q : \mathbb{Q}] = q(q - 1)$. It is then a fact that for $p \nmid 2$,

$$p \text{ splits completely in } K_q \iff p = 1 \mod q \text{ and } 2^{(p-1)/q} = 1 \mod p.$$  

iv) For (odd) squarefree $d$, define $K_d$ to be the compositum of all the fields $K_q$ for prime $q \mid d$, whose degree we denote by $n(d) := [K_d : \mathbb{Q}]$. Then $p \nmid 2d$ splits completely in $K_d$ iff $p \nmid 2$ and for all primes $q \mid d$,

$$p = 1 \mod q \text{ and } 2^{(p-1)/q} = 1 \mod p$$

Thus the number of primes $p \leq x$, $p \nmid 2d$, which split completely in $K_d$ is (maybe up to $O(\omega(d))$) the quantity $P(x; d)$ defined in (5).

5.3. Using GRH. For any normal extension $K/\mathbb{Q}$ (equivalently, Galois here because we are in characteristic zero), Landau showed that there are always infinitely many split primes, in fact that

$$\#\{p \leq x : p \text{ splits completely in } K\} \sim \frac{1}{[K : \mathbb{Q}]} \text{Li}(x), \quad x \to \infty.$$  

This is valid for $K/\mathbb{Q}$ fixed, and $x \to \infty$. We need a version where $K$ varies with $x$, much as we needed to study the prime number theorem in arithmetic progressions with growing modulus; the case of the progressions $p = 1 \mod q$ being precisely that of the cyclotomic fields.

For a number field $K/\mathbb{Q}$, the Dedekind zeta function is defined as

$$\zeta_K(s) := \sum_{(0) \neq I \subset O_K} \frac{1}{N(I)^s}$$

the sum over all nonzero ideals of $O_K$, which is shown to converge absolutely for $\text{Re}(s) > 1$, and in that region by the unique factorization into prime ideals one has an Euler product

$$\zeta_K(s) = \prod_{P \subset O_K \text{ prime}} (1 - \frac{1}{N(P)^s})^{-1}$$

It is known that $\zeta_K(s)$ has an analytic continuation to the entire complex plane, save for a simple pole at $s = 1$, and satisfies a functional equation $s \mapsto 1 - s$. The Generalized Riemann Hypothesis for $\zeta_K(s)$ is that all (nontrivial) zeros of $\zeta_K(s)$ lie on the critical line $\text{Re}(s) = 1/2$. 
Hooley showed that the assumption of the Generalized Riemann Hypothesis for the Dedekind zeta function of $K_d$ implies that the number of primes $p \leq x$ which split completely in $K_d$, satisfies

$$\#\{p \leq x : p \text{ splits completely in } K_d\} = \frac{\text{Li}(x)}{[K_d : \mathbb{Q}]} + O\left(x^{1/2} \log(xd)\right)$$

Since this number is essentially our $P(x;d)$, we obtain Theorem 3.1.

REFERENCES