ON MINIMAL SETS OF GENERATORS  
FOR PRIMITIVE ROOTS  

FRANCESCO PAPPALARDI  

ABSTRACT. A conjecture of Brown and Zassenhaus (see [2]) states that the first $\log p$ primes generate a primitive root $\pmod{p}$ for almost all primes $p$. As a consequence of a Theorem of Burgess and Elliott (see [3]) it is easy to see that the first $\log^2 p \log \log^k p$ primes generate a primitive root $\pmod{p}$ for almost all primes $p$. We improve this showing that the first $\log^2 p / \log \log p$ primes generate a primitive root $\pmod{p}$ for almost all primes $p$.

For a given odd prime number $p$, we define the function $\kappa$ as

$$\kappa(p) = \min\{r \mid \text{the first } r \text{ primes generate } \mathbb{F}_p^*\}.$$

In 1969, H. Brown and H. Zassenhaus conjectured in [2] that $\kappa(p) \leq \lfloor \log p \rfloor$ with probability almost equal to one.

If we denote by $g(p)$ the least primitive root modulo $p$, then a Theorem of D. A. Burgess and P. D. T. A. Elliott states that

$$\pi(x)^{-1} \sum_{p \leq x} g(p) \ll \log^2 x (\log \log x)^4.$$

If $U$ is the number of primes up to $x$ for which $g(p) \geq T$, then

$$UT \ll \sum_{p \leq x} g(p) \ll \pi(x) \log^2 x (\log \log x)^4.$$

For any $\varepsilon > 0$, we choose $T = \log^2 x (\log \log x)^{4+\varepsilon/2}$ so that $U = o(\pi(x))$ and since $g(p) \leq T$ is product of primes less than $T$, we deduce that for almost all primes $p \leq x$,

$$\kappa(p) \leq \log^2 x (\log \log x)^{4+\varepsilon/2} \leq \log^2 p (\log \log p)^{4+\varepsilon}.$$

We will prove the following:

**Theorem 1.** Let $\pi$ be the prime counting function. For all but

$$O\left( \frac{x}{\exp\left( \frac{x}{\log \log x} \right)} \right)$$

we have

$$\kappa(p) \leq \log^2 x (\log \log x)^{4+\varepsilon/2} \leq \log^2 p (\log \log p)^{4+\varepsilon}.$$
primes $p \leq x$, we have that
\[
\kappa(p) \leq \pi \left( \frac{\log^2 p}{e^2} \exp \left\{ \frac{2 (\log \log \log p)^3}{(\log \log p)^2} \right\} \right).
\]
The proof is based on a uniform estimate for the size of the set
\[
\mathcal{H}_{m,r}(x) = \# \left\{ p \leq x \mid |\Gamma_r| = \frac{p - 1}{m} \right\}
\]
where $m$ and $r$ are given integers strictly greater than one, and
\[
\Gamma_r = \langle p_1, \ldots, p_r \pmod{p} \rangle
\]
is the subgroup of $\mathbb{F}_p^*$ generated by the first $r$ primes.

As a subgroup of the cyclic group $\mathbb{F}_p^*$ with index $m$, $\Gamma_r$ is the subgroup of $m$-th powers (mod $p$). Hence
\[
\mathcal{H}_{m,r}(x) = \{ p \leq x \mid p \equiv 1 \pmod{m} \text{ and } p_i \text{ is an } m\text{-th power } \pmod{p} \forall i = 1, \ldots, r \}.
\]

If $n_m(p)$ is the least prime which is not congruent to an $m$-th power (mod $p$), then we can also write:
\[
\mathcal{H}_{m,r}(x) = \{ p \leq x \mid p \equiv 1 \pmod{m} \text{ and } n_m(p) > p_r \}.
\]

We will need to use the large sieve inequality, the proof of which can be found in [1]. That is:

**Lemma 2 (The Large Sieve).** Let $\mathcal{X}$ be a set of integers contained in the interval $\{1, \ldots, z\}$ and for any prime $p \leq x$, let $\Omega_p = \{ h \pmod{p} \mid \forall n \in \mathcal{X}, n \not\equiv h \pmod{p} \}$ and
\[
L = \sum_{q \leq x} \mu^2(q) \prod_{p|q} \frac{|\Omega_p|}{p - |\Omega_p|},
\]
then
\[
|\mathcal{X}| \leq \frac{z + 3x^2}{L}.
\]

In our case, let $\mathcal{X} = \{ n \leq z \mid \forall q|n, q < p_r \}$ and note that if $p \in \mathcal{H}_{m,r}(x)$, then
\[
\Omega_p \supset \{ h \pmod{p} \mid h \text{ is not an } m\text{-th power } \pmod{p} \}
\]
therefore, for such $p$'s, $|\Omega_p| \geq p - 1 - (p - 1)/m$ and
\[
L \geq \sum_{p \in \mathcal{H}_{m,r}(x)} \frac{|\Omega_p|}{p - |\Omega_p|} \geq \frac{m - 1}{2} |\mathcal{H}_{m,r}(x)|.
\]

If we let $\Psi(s, t)$ denote the number of integers $n \leq s$ free of prime factors exceeding $t$, then
\[
\mathcal{H}_{m,r}(x) \leq \frac{8x^2}{(m - 1)\Psi(x^2, p_r)}.
\]

Estimating the function $\Psi(z, y)$ is a classical problem in Number Theory. In 1983, R. Canfield, P. Erdős and C. Pomerance (see [4]) proved the following:
LEMMA 3. Let \( u = \frac{\log z}{\log y} \). There exists an absolute constant \( c_1 \) such that

\[
\Psi(z, y) \geq z \exp \left\{ -u \left( \log u + \log \log u - 1 + \frac{(\log \log u) - 1}{\log u} + c_1 \frac{(\log \log u)^2}{\log^2 u} \right) \right\},
\]

for all \( z \geq 1 \) and \( u \geq e^c \).

Applying Lemma 3 with \( z = x^2 \) and \( y = p_x \), we get the following:

LEMMA 4. Let \( u = \frac{\log x}{\log p_x} \). There exists an absolute constant \( c_1 \) such that

\[
\mathcal{H}_{m, x}(x) \leq \frac{8}{e^2} \exp \left\{ u \left( \log u + \log \log u - 1 + \frac{(\log \log u) - 1}{\log u} + c_1 \frac{(\log \log u)^2}{\log^2 u} \right) \right\},
\]

for all \( x \geq 1 \) and \( u \geq e^c \).

PROOF OF THEOREM 1. Let us take \( p_x \) as the range

\[
\log^2 x \geq p_x \geq \frac{\log^2 x}{e^2} \exp \left\{ \frac{(\log \log \log x)^3}{(\log \log x)^2} \right\}.
\]

If we set \( \log_2 x = \log \log x \), \( \log_3 x = \log \log \log x \) and \( u = \frac{\log x}{\log p_x} \), then we can write the estimates:

\[
\frac{\log x}{\log_2 x} \leq u \leq \frac{\log x}{\log_2 x - 1 + \log_3 x / 2 \log_2 x};
\]

\[
\log_2 x - \log_3 x \leq \log u \leq \log_2 x - \log_3 x + \frac{1}{\log_2 x};
\]

\[
\log_2 u \leq \log_3 x \frac{\log_3 x}{\log_2 x} + c_2 \frac{\log_3 x}{\log_2 x};
\]

\[
\frac{1}{\log_2 x} - \frac{2}{\log_2^2 x} \leq \frac{1}{\log u} \leq \frac{1}{\log_2 x} + c_3 \frac{\log_3 x}{\log_2 x};
\]

where \( c_2 \) and \( c_3 \) are absolute constants.

Now let us apply Lemma 4 and deduce that

\[
m \mathcal{H}_{m, x}(x) \ll \exp \left\{ \log x \frac{\log_2 x - 1 + c_4 \frac{\log_3 x}{\log_2 x}}{\log_2 x - 1 + \log_3 x / 2 \log_2 x} \right\}
\ll \exp \left\{ \log x \left( 1 - \frac{\log_3 x}{2 \log_2 x} + c_5 \left( \frac{\log_3 x}{\log_2 x} \right) \right) \right\}
\]

where \( c_4 \) and \( c_5 \) are absolute constants.

Now we are ready to estimate

\[
\# \{ p \leq x \mid [\mathbb{F}_p^* : \Gamma_r] > 1 \}.
\]

We note that the index \([\mathbb{F}_p^* : \Gamma_r]\) is at most \( x \) as it is a divisor of \( p - 1 \).
Since for all but \( O(x / \exp \frac{\log x}{\log \log x}) \) primes \( p \), we may assume that 

\[
p > x / \exp(2 \log x / \log \log x),
\]

if we set \( p_r \geq \frac{\log x}{\log_2 x} \exp(2 \log_3 x / \log_2 x) \) then \( p_r \) is in the range of (1) and by (2) the number of such primes \( p \) for which \( [F^*_p : \Gamma_r] > 1 \) is 

\[
\ll \sum_{m=2}^{x} h_{m,r}(x) \leq \left( \sum_{m=2}^{x} \frac{1}{m} \right) \exp \left\{ \log x \left( 1 - \frac{\log^3 x}{2 \log^2 x} + c_5 \left( \frac{\log^3 x}{\log^2 x} \right) \right) \right\} = O \left( \frac{x}{\exp \left( \frac{\log x \log^3 x}{4 \log^2 x} \right)} \right)
\]

and this completes the proof. \( \Box \)

ACKNOWLEDGMENTS. A version of Lemma 4 has been proven recently also by S. Konyagin and C. Pomerance in [5].

I would like to thank Professor Ram Murty for his suggestions and for a number of interesting observations.

REFERENCES


Dipartimento di Matematica
Terza Università degli Studi di Roma
Via Corrado Segre, 4
Roma
00146-Italia

e-mail: pappa@mat.uniroma3.it