On the Average of the Least Primitive Root Modulo $p$

Leo Murata

Department of Mathematics, Meijigakuin University
Kamikurata, Totsuka, Yokohama, 244 Japan

Here I discuss about the value distribution of the least primitive root to a prime modulus, as the modulus varies. This is a joint work with P.D.T.A. Elliott.

We describe only a summary of our results in this short paper. As for the details we refer to our full-paper [3].

For each odd prime number $p$, $g(p)$ will denote the least primitive root mod $p$. In order to estimate the magnitude of $g(p)$, we start from a probabilistic argument:

Among the $p-1$ invertible residue classes modulo $p$, $\varphi(p-1)$ classes are primitive, where $\varphi$ is Euler's totient function. So, on the assumption of good distribution of the primitive classes, we can surmise that

for almost all $p$, $g(p)$ is not very far from $\frac{p-1}{\varphi(p-1)}$.

This function fluctuates irregularly, but we can prove:

$$\pi(x)^{-1} \sum_{p \leq x} \frac{p-1}{\varphi(p-1)} = C + O\left(\frac{1}{\log x}\right),$$

where $\pi(x)$ denotes the number of primes not exceeding $x$, and

$$C = \prod_p \left(1 + \frac{1}{(p-1)^2}\right) \approx 2.827\cdots.$$ 

Thus we can surmise that

for almost all $p$, $\frac{p-1}{\varphi(p-1)}$ is not very far from the constant $C$.

Combining these two, we can expect that, for almost all $p$, $g(p)$ is not very far from the constant $C$. Then we arrive at the following conjecture:

**Conjecture.** As $x$ tends to $\infty$,

$$\pi(x)^{-1} \sum_{p \leq x} g(p) \rightarrow C',$$  \hspace{1cm} (1)

where $C'$ is a constant.

In this direction, more than 25 years ago, Burgess-Elliott obtained the following wonderful result:

**Theorem 1 (Burgess-Elliott [2], 1968).**

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

And a few years ago, I proved
Theorem 2 (L. Murata [7], 1991). Under G.R.H., we have

$$\pi(x)^{-1} \sum_{p \leq x} g(p) \ll (\log x)(\log \log x)^7.$$  

Where G.R.H. means the Riemann Hypothesis for the Dedekind $\zeta$-function of certain Kummer fields.

Now, Elliott and I introduce a real parameter $\delta$ and consider the average of $g(p)\delta$. The intention of our joint work is to find out (or identify) a plausible general conjecture which will allow the bound of Theorem 2 to be improved to the asymptotic estimate of the type (1).

Our first result is

**Theorem 1.** We assume G.R.H. Then
1) for any $\delta < \frac{1}{2}$, $\lim_{x \to \infty} \pi(x)^{-1} \sum_{p \leq x} g(p)^{\delta} = E_{\delta}$ exists.  
2) for any $\delta$ with $\frac{1}{2} \leq \delta < 1$, and for any $\epsilon > 0$, $\pi(x)^{-1} \sum_{p \leq x} g(p)^{\delta} \ll (\log x)^{2\delta-1}(\log \log x)^{1+\epsilon}$.  

When we take $\delta = 1$, this gives, for any $\epsilon > 0$,

$$\pi(x)^{-1} \sum_{p \leq x} g(p)^{\delta} \ll (\log x)(\log \log x)^{1+\delta}$$  

which is an improvement of Theorem 2.

Here I refer to another result in this field.

**Theorem C** (Wang [8], 1961). Under G.R.H.,

$$g(p) \ll (\log x)^2 \omega(p-1)^6,$$

where $\omega(n)$ denotes the number of distinct prime which divides $n$.

**Theorem D** (Montgomery [6], 1971). Under G.R.H.,

$$g(p) = \Omega((\log p)(\log \log p)).$$

See also [1] and [4].

Wang proved his result by complex analysis and sieve method, more than 30 years ago. When we replace his old sieve lemma by a modern version, the exponent 6 can be improved into $4 + \epsilon$, for any $\epsilon > 0$. And, taking into account of Hardy-Ramanujan’s theorem, we can regard as, for almost all $p$, $\omega(p-1) \approx \log \log p$. Therefore we notice that

unconditional estimate of the average of $g(p)$ $\approx$ G.R.H.-estimate for individual $g(p)$.

In addition, comparing (3) and Theorem D, we find

G.R.H.-estimate of the average $\approx$ G.R.H. $\Omega$-estimate for individual $g(p)$.

We want to know are these coincides accidental or not?

By Theorem D, Montgomery proved that, for a series of infinite primes, $g(p)$ are actually rather big. As for this type of primes, we have

**Corollary.** We assume G.R.H. Let $B$ be an arbitrary positive constant, then we have, for any $\epsilon > 0$,

$$|\{p \leq x; g(p) \geq B(\log x)(\log \log x)\}| \ll \pi(x)^{1+\epsilon} \sqrt{\log x}.$$
So, the primes of "Montgomery type" are rather exceptional.

Our next result shows that, if we add the following Hypothesis A to G.R.H., then we can extend the validity of (2) to any $\delta < 1$.

For primes $w$ and $q$, we define

$$P_w(x;q) = \{p \leq x; p \equiv 1 \pmod{q}, w \text{ is a } q \text{-th power residue modulo } p\}.$$

**Hypothesis A.** For any prime $q$ with $\sqrt{x}(\log x)^{-6} < q \leq \sqrt{x}(\log x)^3$, and for any $w$ with $w < (\log \log x)^4(\log \log \log x)^3$, we have

$$|P_w(x;q)| \ll \frac{x}{\varphi(q)(\log \frac{2x}{q})^2}$$

where the constant implied by the $\ll$-symbol is absolute.

**Theorem 2.** We assume G.R.H. and Hypothesis A.

1) for any $\delta < 1$, $\lim_{x \to \infty} \pi(x)^{-1} \sum_{p \leq x} g(p)^\delta = E_\delta$ exists.

2) for any $\epsilon > 0$,

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

We can prove Theorems 1 and 2 almost in the same way.

For comparatively small value of $g(p)$, G.R.H. and the use of a linear sieve allow us to accurately calculate the frequencies $\lim_{x \to \infty} \pi(x)^{-1} \sum_{p \leq x, g(p)=n} 1 = c_n$, uniformly for $n < \log \log \log x$. Then we have

$$\sum_{n < \log \log \log x} c_n n^\delta = \sum_{n=1}^\infty c_n n^\delta + \text{(error term)}$$

and the first term of the right hand side gives the constant $E_\delta$ in our Theorems 1 and 2.

For comparatively large $g(p)$, Burgess-Elliott [2] shows that large sieve gives satisfactory control. Over the middle range, particularly, for a fixed $\eta > 0$, $(\log x)^{2-\eta} < g(p) < (\log x)^2(\log \log x)^\eta$, it is very difficult to show that

$$\sum_{p: g(p) \text{ is in the middle range}} g(p) = o(\pi(x)).$$

The Hypothesis A attends this difficulty.

Recently, I received a result of computation by polish mathematician Paszkiewicz. He has a conjecture

$$\pi(x)^{-1} \sum_{p \leq x} g(p) \sim \sqrt{\log x},$$

and he got a numerical example, for $x = 10^9$,

$$\frac{\sum_{p \leq x} g(p)}{\pi(x) \sqrt{\log x}} = 1.0816\ldots$$

But, on our recent result, I am suspicious about his conjecture.

**Remark (about Hypothesis A).** If we cut off the last condition from the definition of $P_w(x;q)$, then $|P_w(x;q)|$ turns into the number of primes in an arithmetic progression, $\pi(x;1,q)$. We can regard as, in some sense, the Hypothesis A is a variation of Brun-Titchmarsh’s Theorem. When $q$ is rather big, the
last condition is very strict. So, at least from the probabilistic point of view, the hypothesis is moderate! C. Hooley [5] introduced the set

$$P_b(x; q, r) = \{ p \leq x; \ p \equiv 1 \pmod{q}, \ b^{2^r} \text{ is a } q-\text{th power residue modulo } p \}$$

and he assumed, for any \( q \) with \( x^{\frac{1}{4}} < q \leq x \),

$$|P_b(x; q, r)| \ll \frac{x}{\varphi(q)(\log \frac{2x}{q})^2}.$$  

Under G.R.H. and this Hypothesis, he succeeded in proving that, for an odd integer \( b \neq \pm 1 \),

$$|\{n \leq x; 2^n + b \text{ is a prime number}\}| = o(x).$$

With respect to the range of \( q \), Hypothesis A is much weaker than his, and we have no need of \( q \), but we need a uniformity concerning \( w \).

References