A simple improvement of a theorem of Pólya

Jin-Yi Cai*

The distribution of quadratic residues modulo a prime number p is a fascinating problem. In this note I present a simple improvement to a classical theorem due to Pólya [1].

Theorem: Let p be an odd prime, then for all m and n, the difference of the numbers of quadratic residues and quadratic non-residues between m and n satisfy

$$\left|\sum_{t=m}^{n-1}\left(\frac{t}{p}\right)\right| \leq \frac{2}{\pi}\sqrt{p}\log_e p + O(\sqrt{p}),$$

where $\left(\frac{t}{p}\right)$ is the Legendre symbol.

Intuitively, such a theorem says that the quadratic residues and non-residues modulo a prime number is more or less uniformly distributed in the range from 1 to $p-1$. It serves as a basis for many interesting applications in diverse areas of mathematics, not only in number theory proper. For example, in combinatorics and in applications to computer science, the properties of the *Paley graphs*, which are defined using quadratic residuosity, are frequently used as a test case for properties of random graphs or used in deterministic pseudorandom constructions.

Compared with the classical theorem, our improvement gets a better constant of $2/\pi$. Better asymptotic estimates for interval length $n-m$ between $p^{1/4+\epsilon}$ and $p^{1/2}$ are known, due to Burgess [2]; however, the proof replies on Weil's proof, in an essential way, of the Riemann hypothesis for finite fields. Our proof is more elementary; it uses a change of order in summation involving Gauss sums, and a simple estimation of sums by integrals.

Let p be an odd prime. Let $\xi = e^{\frac{2\pi i}{p}}$ be a primitive pth root of unity. Then the Gauss sum G is defined as $\sum_{s=0}^{p-1}\left(\frac{s}{p}\right)\xi^s$. More generally

$$G_t = \sum_{s=0}^{p-1}\left(\frac{s}{p}\right)\xi^{st}.$$

One of the basic properties of Gauss sums is the relation $G_t = \left(\frac{t}{p}\right)G$.

Consider

$$\sum_{t=m}^{n-1}\left(\frac{t}{p}\right)G = \sum_{t=m}^{n-1} G_t = \sum_{t=m}^{n-1}\sum_{s=0}^{p-1}\left(\frac{s}{p}\right)\xi^{st} = \sum_{s\neq 0}\left(\frac{s}{p}\right)\sum_{t=m}^{n-1}\xi^{st}.$$

*Department of Computer Science, State University of New York at Buffalo, Buffalo, NY 14260. Email: cai@cs.buffalo.edu. Supported in part by NSF grants CCR 9057486 and CCR 9319093, and an Alfred P. Sloan Fellowship.

Note that the geometric sum

$$\sum_{t=m}^{n-1} \xi^{st} = \xi^{sm} \cdot (1 + \xi^s + \cdots + \xi^{s(n-m-1)}) = \frac{\xi^{sn} - \xi^{sm}}{\xi^s - 1} = \xi^{\frac{n+m-1}{2}s} \cdot \frac{\sin \frac{\pi s(n-m)}{p}}{\sin \frac{\pi s}{p}},$$

we have

$$\left| \sum_{t=m}^{n-1} \left(\frac{t}{p}\right) G \right| \leq \sum_{s \neq 0} \frac{1}{\sin \frac{\pi s}{p}}.$$

Our theorem will follow from a (slightly) better estimate of this last sum.

There is a one-to-one correspondence $s \mapsto p-s$ between $1 \leq s \leq \frac{p-1}{2}$ and $\frac{p+1}{2} \leq s \leq p-1$, and note that $\sin \frac{\pi s}{p} = \sin \left(\pi - \frac{\pi s}{p}\right)$, we have

$$\sum_{s \neq 0} \frac{1}{\sin \frac{\pi s}{p}} = 2 \sum_{s=1}^{\frac{p-1}{2}} \frac{1}{\sin \frac{\pi s}{p}}.$$

Now we will estimate the sum using an integral for the monotone function $\frac{1}{\sin \frac{\pi x}{p}}$, for $1 \leq x < \frac{p}{2}$:

$$\sum_{s=1}^{\frac{p-1}{2}} \frac{1}{\sin \frac{\pi s}{p}} \leq \frac{1}{\sin \frac{\pi}{p}} + \int_{\frac{\pi}{p}}^{\frac{p-1}{p} \cdot \frac{\pi}{2}} \frac{p}{\pi} \cdot \frac{dy}{\sin y}.$$

Now the indefinite integral

$$\int \frac{dy}{\sin y} = \frac{1}{2} \log_e \left(\frac{1 - \cos y}{1 + \cos y} \right),$$

with the help of a substitution $t = \cos y < 1$, we have

$$\int_{\frac{\pi}{p}}^{\frac{p-1}{p} \cdot \frac{\pi}{2}} \frac{dy}{\sin y} < \frac{1}{2} \log_e \frac{1 + \cos \frac{\pi}{p}}{1 - \cos \frac{\pi}{p}} = \log p + \log \left(\frac{2}{\pi}\right) + O\left(\frac{1}{p^2}\right).$$

Thus,

$$\left| \sum_{t=m}^{n-1} \left(\frac{t}{p}\right) G \right| \leq \frac{2}{\sin \frac{\pi}{p}} + \frac{2p}{\pi} \cdot \left[\log p + \log \left(\frac{2}{\pi}\right) + O\left(\frac{1}{p^2}\right) \right] = \frac{2}{\pi} p \log p + O(p).$$

As $|G| = \sqrt{p}$, we get

$$\left| \sum_{t=m}^{n-1} \left(\frac{t}{p}\right) \right| \leq \frac{2}{\pi} \sqrt{p} \log p + O(\sqrt{p}).$$

Acknowledgment

I thank Andrew Odlyzko for correspodence on the problem.

References

[1] Hua Loo Keng, *Introduction to Number Theory*. Springer-Verlag 1982.

[2] Burgess, D. A., The distribution of quadratic residues and non-residues. *Mathematika* **4** (1957) 106–112.