VINOGRADOV'S ESTIMATES FOR THE LEAST QUADRATIC NON-RESIDUES

STEVE FAN

ABSTRACT. For an odd prime p, denote by n_p the least (positive) quadratic non-residue modulo p. Vinogradov [15] proved that $n_p = O(p^\alpha (\log p)^2)$, where $\alpha = 1/(2\sqrt{e})$. Here we present an elementary proof of this result due to Davenport and Erdős [4]. We shall also discuss upper bounds for the least (positive) primitive root g_p modulo p that are related to Vinogradov's work [16], and in particular, Hua's result [11] that $g_p < 2^{m+1}\sqrt{p}$, where m denotes the number of distinct prime factors of $p-1$.

1. Introduction

Let p be an odd prime and let n_p denote the least (positive) quadratic non-residue modulo p. By definition, we know that n_p must be prime. It is also easy to show that $n_p \le (p-1)/2$ for all $p \ge 5$. Indeed, this is clear if $p \equiv 1 \pmod 4$, since $(-1/p) = 1$, where $(\cdot/p)$ is the Legendre symbol (mod p). Suppose now that $p \equiv 3 \pmod 4$. If $(p-1)/2$ is a quadratic non-residue (mod p), then $n_p \le (p-1)/2$. If $(p-1)/2$ is a quadratic residue (mod p), say $x^2 \equiv (p-1)/2 \pmod p$ for some $x \in \mathbb{Z}$, then $2x^2 \equiv -1 \pmod p$. Since $(-1/p) = -1$, this implies that 2 is a quadratic non-residue (mod p) and hence $n_p = 2 \le (p-1)/2$. In the case $p \equiv 3 \pmod 4$, this argument actually shows that $n_p \le \max(d, (p-1)/d)$, where d is any positive divisor of $p-1$. By choosing d to be the largest divisor of $p-1$ with $d \le \sqrt{p-1}$, we may expect that n_p is at most $O(\sqrt{p})$. Such a non-trivial upper bound for n_p (with an extra $\log p$ factor) can be obtained from the Pólya-Vinogradov inequality:

$$\sum_{n=M+1}^{M+N} \chi(n) \ll \sqrt{q}\log q,$$

where M, N are any integers, $q \ge 1$ is a positive integer, and χ is any non-principle Dirichlet character (mod q). Indeed, taking $q = p$, $M = 1$, $N = n_p - 1$ and $\chi(n) = (n/p)$ we obtain $n_p = O(\sqrt{p}\log p)$. For an elementary proof of the Pólya-Vinogradov inequality, see [5, §23]. See also [8] for a short proof using Fourier analysis and for results on various generalized character sums. Vinogradov [15] proved that $n_p = O(p^\alpha(\log p)^2)$, where $\alpha = 1/(2\sqrt{e})$. This was further improved by Burgess [2] who showed that $n_p = O(p^\alpha)$ for any given $\alpha > 1/(4\sqrt{e})$. Burgess derived this result based on Weil's estimate for the complete sum of the Legendre symbols of polynomial values:

$$\left|\sum_{x=1}^{p} \left(\frac{f(x)}{p}\right)\right| \le (n-1)\sqrt{p},$$

where $n \ge 1$ is an odd integer, p is an odd prime, and $f \in \mathbb{F}_p[x]$ is a polynomial of degree n. The case $n = 1$ is trivial, for the sum on the left side is always 0. Weil's estimate is a consequence of the proof of the Riemann hypothesis for curves over finite fields due to

Weil himself, though improvements have been obtained by Korobov [12] and Grechnikov [9] using elementary methods. It was conjectured by Vinogradov that $n_p = O(p^\epsilon)$ for any given $\epsilon > 0$. Vinogradov's conjecture is important in that it is intimately related to deep questions about smooth numbers and the zeros of quadratic Dirichlet L-functions. Linnik [13] proved this conjecture under the generalized Riemann hypothesis. He also showed by means of the large sieve that for any $\epsilon > 0$, the number of primes $p \leq N$ with $n_p > N^\epsilon$ is $O_\epsilon(1)$. Thus Vinogradov's conjecture holds for most primes. Later Ankeny [1] showed that the generalized Riemann hypothesis implies $n_p = O((\log p)^2)$.

In the next section of this note, we shall present an elementary proof of Vinogradov's bound due to Davenport and Erdős [4]. In fact, we shall prove the following slight improvement.

Theorem 1. *$n_p = O((\sqrt{p}\log p)^\alpha)$ for all odd primes p, where $\alpha = 1/\sqrt{e}$.*

Among all the quadratic non-residues modulo a prime p, the primitive roots, namely the generators of $\mathbb{F}_p^\times := \mathbb{F}_p \setminus \{0\}$, are of special interest. For a fixed prime $p \geq 3$, denote by g_p the least (positive) primitive root modulo p. It is clear that g_p is a quadratic non-residue (mod p) and $g_p \geq n_p$. Let m denote the number of distinct prime factors of $p-1$. Vinogradov [16] proved that $g_p < 2^m\sqrt{p}(p-1)/\varphi(p-1)$ for sufficiently large p, improving his earlier result that $g_p < 2^m\sqrt{p}\log p$. Here φ is Euler's totient function. Hua [11] showed that $g_p < 2^{m+1}\sqrt{p}$. Since $2^{m+1} = O(p^\epsilon)$ for every fixed $\epsilon > 0$, Hua's result implies that $g_p = O(p^\alpha)$ for every fixed $\alpha > 1/2$. Using Brun's sieve, Erdős [6] proved that $g_p < \sqrt{p}(\log p)^{17}$ for sufficiently large p, which is better than Hua's estimate when m is large compared to $\log\log p$. Later Erdős and Shapiro [7] improved Hua's result slightly to $g_p = O(m^c\sqrt{p})$, where $c > 0$ is a constant. Using his estimates for character sums, Burgess [3] obtained $g_p = O(p^\alpha)$ for any given $\alpha > 1/4$. However, these results are substantially weaker than expected, since Shoup [14] proved under the assumption of the generalized Riemann hypothesis that $g_p = O((m\log(m+1))^4(\log p)^2)$. We shall present a short proof of Hua's result due to Erdős and Shapiro [7] in the last section.

Theorem 2. *$g_p < 2^{m+1}\sqrt{p}$ for all sufficiently large p, where m is the number of distinct prime factors of $p-1$.*

2. Proof of Theorem 1

The proof of Theorem 1 depends on the following simple identity [4, Lemma 1]:

$$\sum_{x=1}^{p}\left|\sum_{n=1}^{h}\chi(x+n)\right|^2 = h(p-h), \tag{1}$$

where $1 \leq h \leq p$ and χ is any non-principle Dirichlet character (mod p). To prove (1), we expand the square of the inner sum and observe that the contribution from the diagonal terms is

$$\sum_{n=1}^{h}\sum_{x=1}^{p}|\chi(x+n)|^2 = h(p-1).$$

Thus, to prove (1) it suffices to show that the contribution from the non-diagonal terms is

$$\sum_{\substack{n_1,n_2=1\\ n_1\neq n_2}}^{h}\sum_{x=1}^{p}\chi(x+n_1)\overline{\chi}(x+n_2) = -h(h-1).$$

This would follow if we can show

$$\sum_{x=1}^{p} \chi(x+n_1)\overline{\chi}(x+n_2) = -1 \tag{2}$$

for all $n_1, n_2 \in \mathbb{Z}$ with $n_1 \not\equiv n_2 \pmod{p}$. There are a few ways to prove (2). The proof that Davenport and Erdős gave in their paper makes use of the substitution $x + n_1 \equiv y(x + n_2) \pmod{p}$, which gives a bijection between $x \not\equiv -n_1 \pmod{p}$ and $y \not\equiv 1 \pmod{p}$. It then follows from the orthogonality relation that

$$\sum_{x=1}^{p} \chi(x+n_1)\overline{\chi}(x+n_2) = \sum_{y=2}^{p} \chi(y) = -\chi(1) = -1.$$

The argument that the author came up with by himself goes as follows. It is easily seen that (2) is equivalent to the statement that

$$\sum_{x=1}^{p} \chi(x)\overline{\chi}(x+a) = -1 \tag{3}$$

holds for all $a \in (\mathbb{Z}/p\mathbb{Z})^{\times}$, where $(\mathbb{Z}/p\mathbb{Z})^{\times}$ is the multiplicative group of $\mathbb{Z}/p\mathbb{Z}$. Denote by $f(a)$ the expression on the left side of (3). Then

$$f(a) = \sum_{x=1}^{p} \chi(ax)\overline{\chi}(ax+a) = \sum_{x=1}^{p} \chi(x)\overline{\chi}(x+1) = f(1).$$

Thus f is constant on $(\mathbb{Z}/p\mathbb{Z})^{\times}$. By the orthogonality relation we have

$$f(a) = \frac{1}{p-1}\sum_{x=1}^{p} \chi(x) \sum_{b=1}^{p-1} \overline{\chi}(x+b) = \frac{1}{p-1}\left|\sum_{x=1}^{p}\chi(x)\right|^2 - \frac{1}{p-1}\sum_{x=1}^{p}|\chi(x)|^2 = -1$$

for all $a \in (\mathbb{Z}/p\mathbb{Z})^{\times}$. This completes the proof of (3), and hence the proof of (2).

It may be worth noting that Burgess obtained his estimate for the least quadratic non-residue $\pmod{p}$ by treating the more general $2r$-th moment

$$\sum_{x=1}^{p}\left|\sum_{n=1}^{h}\chi(x+n)\right|^{2r}$$

with $\chi(n) = (n/p)$. Based on Weil's estimate mentioned earlier, he showed that the above sum is less than $(2r)^r p h^r + r(2\sqrt{p}+1)h^{2r}$. The reader is referred to [2] for further details.

We are now in a position to prove Theorem 1. Suppose $p \geq 5$. Take $h = \lfloor \sqrt{p}\log p \rfloor \geq 3$ and $\chi(n) = (n/p)$, where $\lfloor \sqrt{p}\log p \rfloor$ is the integer part of $\sqrt{p}\log p$. For every positive integer $1 \leq x \leq h$, denote by $N(x, x+h)$ the number of quadratic non-residues $\pmod{p}$ in the interval $(x, x+h]$. Observe that

$$\sum_{n=1}^{h}\chi(x+n) = h - 2N(x, x+h).$$

Since every positive quadratic non-residue $\pmod{p}$ must have a prime divisor q which satisfies $(q/p) = -1$ and hence satisfies $q \geq n_p$, it follows that

$$N(x, x+h) \leq \#\{m \in (x, x+h] \colon m \text{ has a prime divisor } q \geq n_p\}.$$

If $n_p > 2h$, then $N(x, x+h) = 0$ for all $1 \leq x \leq h$. Thus we have

$$\sum_{n=1}^{h} \chi(x+n) = h$$

for all $1 \leq x \leq h$. By (1) we have $h^3 \leq h(p-h)$, i.e., $h^2 + h - p \leq 0$. But this is false, since

$$h^2 + h > \frac{(h+1)^2}{2} > \frac{p(\log p)^2}{2} > p.$$

Hence we must have $n_p \leq 2h$. This yields the bound that we previously derived from the Pólya-Vinogradov inequality. By Chebyshev's estimate [10, Theorem 7] and Mertens' theorem [10, Theorem 427] we have

$$\begin{aligned} N(x, x+h) \leq \sum_{n_p \leq q \leq 2h} \left(\left\lfloor \frac{x+h}{q} \right\rfloor - \left\lfloor \frac{x}{q} \right\rfloor \right) &= h \sum_{n_p \leq q \leq 2h} \frac{1}{q} + O\left(\frac{h}{\log h} \right) \\ &= h(\log\log 2h - \log\log n_p) + O\left(\frac{h}{\log h} \right). \end{aligned}$$

Hence

$$\sum_{n=1}^{h} \chi(x+n) \geq h \left(1 - 2\log\log 2h + 2\log\log n_p + O\left(\frac{1}{\log h} \right) \right). \tag{4}$$

If the right side of (4) is negative, then we have

$$\frac{\log n_p}{\log 2h} < e^{-1/2 + O(1/\log h)} = e^{-1/2 + \log(1 + O(1/\log h))} = e^{-1/2} \left(1 + O\left(\frac{1}{\log h} \right) \right),$$

which implies that $\log n_p < e^{-1/2} \log 2h + O(1)$. This gives $n_p = O((\sqrt{p} \log p)^{\alpha})$, where $\alpha = 1/\sqrt{e}$. Suppose now that the right side of (4) is non-negative. By (3) we obtain

$$h^3 \left(1 - 2\log\log 2h + 2\log\log n_p + O\left(\frac{1}{\log h} \right) \right)^2 \leq h(p-h) < hp.$$

It follows that

$$1 - 2\log\log 2h + 2\log\log n_p + O\left(\frac{1}{\log h} \right) < \frac{\sqrt{p}}{h} < \frac{2\sqrt{p}}{h+1} < \frac{2}{\log p} < \frac{2}{\log h}.$$

Thus we have

$$1 - 2\log\log 2h + 2\log\log n_p + O\left(\frac{1}{\log h} \right) < 0.$$

We can conclude as before that $n_p = O((\sqrt{p} \log p)^{\alpha})$. This finishes the proof of Theorem 1.

3. Proof of Theorem 2

The proof of Theorem 2 depends on a simple inequality for character sums [7, Lemma]. It states that if $A, B \subseteq \mathbb{F}_p$ with cardinality $|A|$ and $|B|$, respectively, then

$$\left| \sum_{a \in A} \sum_{b \in B} \chi(a+b) \right| \leq \sqrt{p|A||B|} \tag{5}$$

for any non-principle Dirichlet character (mod p). To prove this, we consider the Gauss sum

$$\tau(\chi) := \sum_{h\in\mathbb{F}_p} \chi(h)e_p(h),$$

where $e_p(h) := e^{2\pi ih/p}$. It can be shown easily that

$$\chi(h')\tau(\bar{\chi}) = \sum_{h\in\mathbb{F}_p} \chi(h)e_p(hh').$$

and that $|\tau(\chi)| = \sqrt{p}$ (see [5, §2]). Thus we have

$$\tau(\bar{\chi})\sum_{a\in A}\sum_{b\in B}\chi(a+b) = \sum_{h\in\mathbb{F}_p}\chi(h)\left(\sum_{a\in A}e_p(ha)\right)\left(\sum_{b\in B}e_p(hb)\right).$$

It follows that

$$\sqrt{p}\left|\sum_{a\in A}\sum_{b\in B}\chi(a+b)\right| \le \sum_{h\in\mathbb{F}_p}\left|\sum_{a\in A}e_p(ha)\right|\left|\sum_{b\in B}e_p(hb)\right|.$$

By Cauchy-Schwarz inequality, the right side is

$$\le \left(\sum_{h\in\mathbb{F}_p}\left|\sum_{a\in A}e_p(ha)\right|^2\right)^{\frac{1}{2}}\left(\sum_{h\in\mathbb{F}_p}\left|\sum_{b\in B}e_p(hb)\right|^2\right)^{\frac{1}{2}} \le p\sqrt{|A||B|},$$

since

$$\sum_{h\in\mathbb{F}_p}\left|\sum_{a\in A}e_p(ha)\right|^2 = \sum_{a,a'\in A}\sum_{h\in\mathbb{F}_p}e_p((a-a')h) = \sum_{a\in A}p = p|A|$$

and similarly

$$\sum_{h\in\mathbb{F}_p}\left|\sum_{b\in B}e_p(hb)\right|^2 = p|B|.$$

Hence

$$\sqrt{p}\left|\sum_{a\in A}\sum_{b\in B}\chi(a+b)\right| \le p\sqrt{|A||B|},$$

which gives (5).

Another ingredient needed for the proof of Theorem 2 concerns the values of the sum $S(h)$ defined for every $h\in\mathbb{Z}$ with $\gcd(h,p)=1$ by

$$S(h) := \sum_{d|p-1}\frac{\mu(d)}{\varphi(d)}\sum_{\operatorname{ord}(\chi)=d}\chi(h),$$

where μ is the Möbius function and the inner sum is over all characters χ of order d in the character group (mod p). Let g be any primitive root (mod p), so that $h\equiv g^v \pmod p$ for some $0\le v<p$. For every $d\mid(p-1)$, put $u_d := \gcd(v,d)$. Then

$$\sum_{\operatorname{ord}(\chi)=d}\chi(h) = \sum_{\substack{k=1\\ \gcd(k,d)=1}}^{d} e_d(kv) = c_d(v),$$

where $c_d(v)$ is Ramanujan's sum which is multiplicative as a function of d. Hence

$$S(h)=\sum_{d|p-1}\frac{\mu(d)c_d(v)}{\varphi(d)}.$$

Note that

$$\sum_{d|n}\frac{\mu(d)c_d(v)}{\varphi(d)}$$

is multiplicative as a function of n. By [10, Theorem 272] we have

$$c_d(v)=\frac{\mu(d/u_d)\varphi(d)}{\varphi(d/u_d)}.$$

Let q be a prime and $r\geq 1$ a positive integer. Then

$$\sum_{d|q^r}\frac{\mu(d)c_d(v)}{\varphi(d)}=1-\frac{\mu(q/u_q)}{\varphi(q/u_q)}.$$

It follows that

$$\sum_{d|n}\frac{\mu(d)c_d(v)}{\varphi(d)}=\prod_{q|n}\left(1-\frac{\mu(q/u_q)}{\varphi(q/u_q)}\right),$$

If h is a primitive root (mod p), then $u_q=1$ for all $q\mid(p-1)$. Thus we have

$$S(h)=\prod_{q|(p-1)}\left(1+\frac{1}{q-1}\right)=\frac{p-1}{\varphi(p-1)}.$$

On the other hand, if h is not a primitive root (mod p), then $u_{p-1}>1$. This implies that there exists a prime divisor q of $p-1$ for which $u_q=q$, so that $1-\mu(q/u_q)/\varphi(q/u_q)=0$. Therefore, we have $S(h)=0$.

We are now ready to prove Theorem 2. We may assume that $g_p\geq 3$. Note that $S(h)=0$ for all $1\leq h<g_p$. Taking $A=B=\{1,2,...,\lfloor(g_p-1)/2\rfloor\}$, where $\lfloor x\rfloor$ is the integer part of $x\in\mathbb{R}$, we obtain

$$\begin{aligned}0=\sum_{a\in A}\sum_{b\in B}S(a+b)&=\sum_{d|p-1}\frac{\mu(d)}{\varphi(d)}\sum_{\operatorname{ord}(\chi)=d}\sum_{a\in A}\sum_{b\in B}\chi(a+b)\\&=\lfloor(g_p-1)/2\rfloor^2+\sum_{\substack{d|p-1\\d>1}}\frac{\mu(d)}{\varphi(d)}\sum_{\operatorname{ord}(\chi)=d}\sum_{a\in A}\sum_{b\in B}\chi(a+b).\end{aligned}$$

It follows that

$$\lfloor(g_p-1)/2\rfloor^2\leq\sum_{\substack{d|p-1\\d>1}}\frac{|\mu(d)|}{\varphi(d)}\sum_{\operatorname{ord}(\chi)=d}\left|\sum_{a\in A}\sum_{b\in B}\chi(a+b)\right|.$$

By (5) we have

$$\lfloor(g_p-1)/2\rfloor^2\leq\sqrt{p}\lfloor(g_p-1)/2\rfloor\sum_{\substack{d|p-1\\d>1}}|\mu(d)|,$$

where we have used the fact that the number of elements of $\mathbb{F}_p^\times$ of order d equals $\varphi(d)$ (see [10, Theorem 110]). Note that the sum on the right side represents the number of square-free positive divisors $d > 1$ of $p-1$. It follows that

$$\lfloor (g_p - 1)/2 \rfloor \leq (2^m - 1)\sqrt{p}.$$

But

$$\left\lfloor \frac{g_p - 1}{2} \right\rfloor + 1 \geq \frac{g_p - 2}{2} + 1 = \frac{g_p}{2}.$$

Therefore, we have

$$g_p \leq 2(2^m - 1)\sqrt{p} + 2 < 2^{m+1}\sqrt{p}.$$

This completes the proof of Theorem 2.

References

[1] N. C. Ankeny, *The least quadratic non-residue*, Ann. of Math. **55** (2) (1952), 65–72.
[2] D. A. Burgess, *The distribution of quadratic residues and non-residues*, Mathematika **4** (1957), 106–112.
[3] D. A. Burgess, *On character sums and primitive roots*, Proc. Lond. Math. Soc. **12** (3) (1962), 179–192.
[4] H. Davenport and P. Erdős, *The distribution of quadratic and higher residues*. Publ. Math. Debrecen **2** (1952), 252–265.
[5] H. Davenport, *Multiplicative Number Theory*, 3rd. ed., Grad. Texts in Math., vol. 74, Springer-Verlag, New York, 2000. Revised and with a preface by H. L. Montgomery.
[6] P. Erdős, *On the least primitive root of a prime*, Bull. Lond. Math. Soc. **55** (1945), 131–132.
[7] P. Erdős and H. N. Shapiro, *On the least primitive root of a prime*, Pacific J. Math. **7** (1957), 861–865.
[8] J. Friedlander and H. Iwaniec, *Estimates for character sums*, Proc. Amer. Math. Soc. **119** (1993), 365–372.
[9] E. A. Grechnikov, *An estimate for the sum of Legendre symbols*, Math. Notes **88** (2010), 819–826.
[10] G. H. Hardy and E. M. Wright, *An introduction to the Theory of Numbers*, 6th. ed., Oxford Univ. Press, Oxford, 2008. Revised by D. R. Heath-Brown and J. H. Silverman; With a forward by A. J. Wiles.
[11] L.-K. Hua, *On the least primitive root of a prime*, Bull. Amer. Math. Soc. **48** (1942), 726–730.
[12] P. M. Korobov, *An estimate of the sum of the Legendre symbols*, Dokl. Akad. Nauk SSSR **196** (4) (1971), 764–767.
[13] U. V. Linnik, *A remark on the least quadratic non-residue*, C. R. (Doklady) Acad. Sci. URSS (N.S.) **36** (1942), 119–120.
[14] V. Shoup, *Searching for primitive roots in finite fields*, Math. Comp. **58** (1992), 369–380.
[15] I. M. Vinogradov, *Sur la distribution des résidus et des non-résidus des puissances*, Journal Physico-Math. Soc. Univ. Perm, no. 1 (1918), 94–96.
[16] I. M. Vinogradov, *On the least primitive root of a prime (in Russian)*, Dokl. Akad. Nauk SSSR **1** (1930), 7–11.

Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA
Email address: steve.fan.gr@dartmouth.edu