The smallest quadratic nonresidue modulo a prime

Paul Pollack

University of Georgia

August 29, 2012

Introduction

Definition

For each odd prime p, let $n_2(p)$ denote the least quadratic nonresidue modulo p. For example, $n_2(5) = 2$ and $n_2(7) = 3$. For completeness, put $n_2(2) = 0$.

Introduction

Definition

For each odd prime p, let $n_2(p)$ denote the least quadratic nonresidue modulo p. For example, $n_2(5) = 2$ and $n_2(7) = 3$. For completeness, put $n_2(2) = 0$.

Problem (**Normal order**)

How large is $n_2(p)$ typically?

Problem (**Maximal order**)

What is the largest $n_2(p)$ can be as a function of p?

Problem (**Average order**)

What is the mean value of $n_2(p)$? In other words, what do the finite averages $\frac{1}{\pi(x)} \sum_{p \le x} n_2(p)$ converge to as $x \to \infty$?

The normal order

Since the Legendre symbol is multiplicative,

$$\left(\frac{n}{p}\right) = -1 \Longrightarrow \left(\frac{q}{p}\right) = -1 \text{ for some prime } q \text{ dividing } n.$$

Hence, $n_2(p)$ is always a prime number.

The normal order

Since the Legendre symbol is multiplicative,

$$\left(\frac{n}{p}\right) = -1 \implies \left(\frac{q}{p}\right) = -1 \text{ for some prime } q \text{ dividing } n.$$

Hence, $n_2(p)$ is always a prime number.

Let p_k be the kth prime. Let's ask how often $n_2(p) = p_k$. For example,

$$\begin{aligned} n_2(p) = 2 &\iff \left(\frac{2}{p}\right) = -1 \\ &\iff p \equiv \pm 3 \pmod 8. \end{aligned}$$

By the prime number theorem for arithmetic progressions, it follows that $n_2(p) = 2$ for asymptotically half of all primes p.

A random variable perspective

In general, if BLAH is a property primes p might have, let me write

$$\mathbb{P}(\text{BLAH}) = \lim_{x\to\infty} \frac{1}{\pi(x)} \#\{p \le x : p \text{ satisfies BLAH}\}.$$

[**WARNING**: The word "probability" is a bit misplaced, since natural density is not a probability measure.]

Then for any fixed prime q, quadratic reciprocity and the prime number theorem for progressions combine to show that

$$\mathbb{P}(\left(\frac{q}{p}\right) = 1) = \frac{1}{2}.$$

A random variable perspective

For distinct primes q, the events "q is a square mod p" are independent. This follows (for example) from the Chebotarev density theorem, using that $[\mathbb{Q}(\sqrt{q_1}, \dots, \sqrt{q_k}) : \mathbb{Q}] = 2^k$.

Hence,

$$\mathbb{P}(n_2(p) = p_k) = \frac{1}{2^k}.$$

A random variable perspective

For distinct primes q, the events "q is a square mod p" are independent. This follows (for example) from the Chebotarev density theorem, using that $[\mathbb{Q}(\sqrt{q_1}, \dots, \sqrt{q_k}) : \mathbb{Q}] = 2^k$.

Hence,

$$\mathbb{P}(n_2(p) = p_k) = \frac{1}{2^k}.$$

As a corollary, we find that the "random variable" $n_2(\cdot)$ is "bounded in probability":

Theorem

If ξ is any function that tends to infinity (however slowly), then

$$\mathbb{P}(n_2(p) > \xi(p)) = 0.$$

The maximal order

Let's pretend that for each prime p, the number $n_2(p)$ is determined by flipping coins until one gets a 'heads'; if this occurs on the kth flip, set $n_2(p) = p_k$. And let's pretend that for distinct primes p, these experiments are independent.

The Borel-Cantelli theorem suggests the following conjecture:

Conjecture

Let $\epsilon > 0$. Then for all large primes p,

$$n_2(p) < \left(\frac{1}{\log 2} + \epsilon\right) \cdot (\log p)(\log\log p).$$

On the other hand, the reverse inequality holds for infinitely many primes p if $1+\epsilon$ is replaced by $1-\epsilon$.

The maximal order, on GRH

Let p be a prime. Let's assume that the Riemann Hypothesis holds for the Dirichlet L-function

$$L(s, \left(\frac{\cdot}{p}\right)) := \sum_{n=1}^{\infty} \left(\frac{n}{p}\right) n^{-s}.$$

In this case, the proof of the prime number theorem for arithmetic progressions (see, for example, Davenport) shows that

$$\left| \sum_{\substack{q \leq x \\ q \text{ prime}}} \left(\frac{q}{p}\right) \right| < C x^{1/2} \log(px),$$

for all $x \geq 2$.

The maximal order, on GRH

If $x < p$ and all primes $q \leq x$ are quadratic residues modulo p, then

$$\sum_{q \leq x} \left(\frac{q}{p}\right) = \pi(x),$$

which is asymptotically $x/\log x$. Once x is at all large (in terms of p), this exceeds the upper bound on the previous slide.

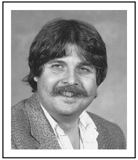
More precisely, we have:

Theorem

Suppose p is sufficiently large. If the Riemann Hypothesis holds for $L(s, (\frac{\cdot}{p}))$, then

$$n_2(p) < (C' \log p \log\log p)^2.$$

The maximal order, on GRH



Theorem (Bach, improving on Ankeny)

If RH holds for $L(s, (\frac{\cdot}{p}))$, then

$$n_2(p) < 2(\log p)^2.$$

On GRH, Montgomery has proved that

$$n_2(p) > c(\log p)(\log\log p)$$

infinitely often, where $c > 0$ is a small positive constant. (Without needing to assume GRH, the double-log can be replaced with a triple log, as shown by Graham and Ringrose.)

For most of the rest of this talk, we will focus attention on unconditional upper bounds on $n_2(p)$.

The main results

Theorem (Folklore)

For all large primes p, we have

$$n_2(p) < p^{1/2}.$$

The main results

Theorem (Folklore)

For all large primes p, we have

$$n_2(p) < p^{1/2}.$$

Theorem (Pólya–Vinogradov)

For all primes p, we have

$$n_2(p) \leq 1 + \sqrt{p}\log p.$$

The main results, ctd.

Theorem (Pólya–Vinogradov + Vinogradov's trick)

Let $\epsilon > 0$. Then for all large primes p, we have

$$n_2(p) \leq p^{\frac{1}{2\sqrt{e}}+\epsilon}.$$

The main results, ctd.

Theorem (Pólya–Vinogradov + Vinogradov's trick)

Let $\epsilon > 0$. Then for all large primes p, we have

$$n_2(p) \le p^{\frac{1}{2\sqrt{e}}+\epsilon}.$$

Theorem (Burgess + Vinogradov's trick)

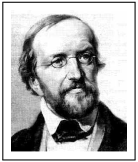
Let $\epsilon > 0$. Then for all large primes p, we have

$$n_2(p) \le p^{\frac{1}{4\sqrt{e}}+\epsilon}.$$

Remark

We have $1/2\sqrt{e} = 0.303265\dots$ and $1/4\sqrt{e} = 0.151632\dots$.

A digression: The probability two integers are relatively prime

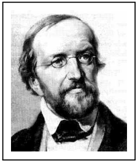


Theorem (Dirichlet)

The probability that two integers are relatively prime is $1/\zeta(2) = 6/\pi^2$. *More precisely:*

$$\lim_{N\to\infty} \frac{\#\{(a,b) : 1 \le a, b \le N, \gcd(a,b) = 1\}}{N^2} = \frac{6}{\pi^2}.$$

A digression: The probability two integers are relatively prime



Theorem (Dirichlet)

The probability that two integers are relatively prime is $1/\zeta(2) = 6/\pi^2$*. More precisely:*

$$\lim_{N\to\infty} \frac{\#\{(a,b) : 1 \le a, b \le N, \gcd(a,b) = 1\}}{N^2} = \frac{6}{\pi^2}.$$

Proof: Let us say (a,b) is *visible from the origin* if $\gcd(a,b) = 1$. The visible lattice points are symmetric about $y = x$. Moreover, the only visible lattice point of the form (a,a) is $(1,1)$.

A digression: The probability two integers are relatively prime

Hence,

$$\sum_{\substack{1 \le a,b \le N \\ \gcd(a,b)=1}} 1 = \left(2 \sum_{a=1}^{N} \sum_{\substack{1 \le b \le a \\ \gcd(a,b)=1}} 1 \right) - 1$$

$$= 2 \sum_{a=1}^{N} \phi(a) - 1.$$

A digression: The probability two integers are relatively prime

Hence,

$$\sum_{\substack{1\le a,b\le N\\ \gcd(a,b)=1}} 1 = \left(2\sum_{a=1}^{N}\sum_{\substack{1\le b\le a\\ \gcd(a,b)=1}} 1\right) - 1$$
$$= 2\sum_{a=1}^{N}\phi(a) - 1.$$

To evaluate the remaining sum, notice that

$$\phi(a) = a\prod_{p\mid a}(1-1/p)$$
$$= a\sum_{d\mid a}\frac{\mu(d)}{d}.$$

A digression: The probability two integers are relatively prime

Therefore,

$$\begin{aligned}
\sum_{a=1}^{N} \phi(a) &= \sum_{a=1}^{N} a \sum_{d \mid a} \frac{\mu(d)}{d} \\
&= \sum_{d \le N} \frac{\mu(d)}{d} \sum_{\substack{1 \le a \le N \\ d \mid a}} a \\
&= \sum_{d \le N} \frac{\mu(d)}{d} \sum_{1 \le e \le N/d} (de) \\
&= \sum_{d \le N} \mu(d) \sum_{e \le N/d} e.
\end{aligned}$$

The inner sum is $\frac{1}{2}(N/d)^2 + O(N/d)$.

A digression: The probability two integers are relatively prime

We find that ignoring (easily-handled) error terms,

$$\#\{1 \le a, b \le N : \gcd(a,b) = 1\} \approx N^2 \sum_{d=1}^{N} \frac{\mu(d)}{d^2}.$$

A digression: The probability two integers are relatively prime

We find that ignoring (easily-handled) error terms,

$$\#\{1 \le a, b \le N : \gcd(a,b) = 1\} \approx N^2 \sum_{d=1}^{N} \frac{\mu(d)}{d^2}.$$

For large N, we can extend the sum to infinity making only a small error (of size $O(1/N)$). Moreover,

$$\sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} = \prod_p \left(1 - \frac{1}{p^2}\right) = \prod_p \left(1 + \frac{1}{p^2} + \frac{1}{p^4} + \dots\right)^{-1} = \zeta(2)^{-1}.$$

Since $\zeta(2) = 6/\pi^2$, the theorem of Dirichlet follows.

An elementary proof that $n_2(p) < p^{1/2}$

Given a fraction $\frac{a}{b}$ with $p \nmid b$, we identify $\frac{a}{b}$ with $ab^{-1} \pmod{p}$. Notice that

$$\frac{a}{b} \equiv \frac{c}{d} \pmod{p} \Longleftrightarrow p \mid ad - bc.$$

An elementary proof that $n_2(p) < p^{1/2}$

Given a fraction $\frac{a}{b}$ with $p \nmid b$, we identify $\frac{a}{b}$ with $ab^{-1} \pmod p$. Notice that

$$\frac{a}{b} \equiv \frac{c}{d} \pmod p \iff p \mid ad - bc.$$

Now consider the following set of fractions:

$$\mathfrak{F} = \left\{ \frac{a}{b} : 1 \le a, b \le \sqrt{p} \text{ and } \gcd(a,b) = 1 \right\}.$$

An elementary proof that $n_2(p) < p^{1/2}$

Given a fraction $\frac{a}{b}$ with $p \nmid b$, we identify $\frac{a}{b}$ with $ab^{-1} \pmod p$. Notice that

$$\frac{a}{b} \equiv \frac{c}{d} \pmod p \Longleftrightarrow p \mid ad - bc.$$

Now consider the following set of fractions:

$$\mathfrak{F} = \left\{ \frac{a}{b} : 1 \le a, b \le \sqrt{p} \text{ and } \gcd(a,b) = 1 \right\}.$$

By Dirichlet's result on visible lattice points,

$$\#\mathfrak{F} \sim \frac{6}{\pi^2} p; \quad \text{this gives} \quad \#\mathfrak{F} > \frac{p}{2}$$

for large p. (Since $6/\pi^2 = 0.607927\ldots$.)

An elementary proof that $n_2(p) < p^{1/2}$

Lemma

No two elements of $\mathfrak{F}$ are congruent modulo p.

Proof.

If $\frac{a_1}{b_1}, \frac{a_2}{b_2} \in \mathfrak{F}$ (and not the same), then $0 < |a_1 b_2 - a_2 b_1| < p$.

An elementary proof that $n_2(p) < p^{1/2}$

Lemma

No two elements of $\mathfrak{F}$ are congruent modulo p.

Proof.

If $\frac{a_1}{b_1}, \frac{a_2}{b_2} \in \mathfrak{F}$ (and not the same), then $0 < |a_1 b_2 - a_2 b_1| < p$.

Since $\#\mathfrak{F} > p/2$ and there are only $\frac{p-1}{2}$ (nonzero) squares mod p, some $\frac{a}{b} \in \mathfrak{F}$ reduces to a nonsquare mod p. So either a is a nonsquare or b is a nonsquare. Hence,

$$n_2(p) \leq \sqrt{p}.$$

(Of course, equality is impossible here.)

The Pólya-Vinogradov inequality

We now turn to our next upper bound on $n_2(p)$. Since there are the same number of squares as nonsquares modulo p, and since the Legendre symbol is periodic modulo p, it is trivial that

$$\left|\sum_{n=1}^{N}\left(\frac{n}{p}\right)\right| < p$$

for all N.

Theorem (Pólya-Vinogradov)

For every natural number N, we have

$$\left|\sum_{n=1}^{N}\left(\frac{n}{p}\right)\right| < \sqrt{p}\log p.$$

A quick corollary

Corollary

For all primes $p \geq 3$, we have

$$n_2(p) < 1 + \sqrt{p}\log p.$$

Proof.

Obvious if $1 + \sqrt{p}\log p \geq p$. So suppose otherwise. If $n_2(p) \geq 1 + \sqrt{p}\log p$, then

$$\sqrt{p}\log p \leq \sum_{n < 1+\sqrt{p}\log p} 1 = \sum_{n < 1+\sqrt{p}\log p} \left(\frac{n}{p}\right) < \sqrt{p}\log p.$$

Not cool.

Quadratic Gauss sums

For each integer a, define

$$g_a = \sum_{r \bmod p} \left(\frac{r}{p}\right) \exp\left(2\pi i \frac{ar}{p}\right).$$

Note that g_a depends only on the residue class of a mod p. [For Fourier transform fans, g_a is $\hat{\chi}(-a)$, where $\hat{\cdot}$ is the Fouier transform on $\mathbb{Z}/p$.]

Quadratic Gauss sums

For each integer a, define

$$g_a = \sum_{r \bmod p} \left(\frac{r}{p}\right) \exp\left(2\pi i \frac{ar}{p}\right).$$

Note that g_a depends only on the residue class of a mod p. [For Fourier transform fans, g_a is $\hat{\chi}(-a)$, where $\hat{\cdot}$ is the Fouier transform on $\mathbb{Z}/p$.]

If $a \not\equiv 0 \pmod p$, the change of variables $r \mapsto a^{-1}r$ shows that

$$g_a = \left(\frac{a^{-1}}{p}\right) g_1 = \left(\frac{a}{p}\right) g_1.$$

This also holds if $a \equiv 0 \pmod p$, since g_a and $\left(\frac{a}{p}\right)$ both vanish.

Evaluation of the Gauss sum

Theorem

Each sum g_a *with* $a \not\equiv 0 \pmod{p}$ *has* $|g_a| = \sqrt{p}$*. In fact,* $g_1 = \pm\sqrt{p}$ *if* $p \equiv 1 \pmod{4}$*, and* $g_1 = \pm i\sqrt{p}$ *if* $p \equiv 3 \pmod{4}$*.*

The determination of the sign . . . has vexed me for many years. This deficiency overshadowed everything that I found over the last four years. . . . Finally, a few days ago, I succeeded – but not as a result of my search but rather, I should say, through the mercy of God. As lightning strikes, the riddle has solved itself.

Evaluation of the Gauss sum

Theorem (Gauss)

If $p \equiv 1 \pmod 4$, *then*

$$g_1 = \sqrt{p},$$

and if $p \equiv 3 \pmod 4$, *then*

$$g_1 = \mathrm{i}\sqrt{p}.$$

In what follows, we only need the easy result that

$$|g_a| = \sqrt{p} \quad \text{for all } a \not\equiv 0 \pmod p.$$

Proof of the Pólya–Vinogradov inequality

If $a \not\equiv 0 \pmod{p}$, then $g_a = \left(\frac{a}{p}\right) g_1$. Solving for $\left(\frac{a}{p}\right)$, we find that

$$\left(\frac{a}{p}\right) = \frac{g_a}{g_1} = \frac{1}{g_1} \sum_{r \text{ mod } p} \left(\frac{r}{p}\right) \exp\left(2\pi i \frac{ar}{p}\right).$$

Hence: $$\sum_{a=1}^{N} \left(\frac{a}{p}\right) = \frac{1}{g_1} \sum_{r \text{ mod } p} \left(\frac{r}{p}\right) \sum_{a=1}^{N} \exp\left(2\pi i \frac{ar}{p}\right).$$

Proof of the Pólya–Vinogradov inequality

If $a \not\equiv 0 \pmod p$, then $g_a = \left(\frac{a}{p}\right) g_1$. Solving for $\left(\frac{a}{p}\right)$, we find that

$$\left(\frac{a}{p}\right) = \frac{g_a}{g_1} = \frac{1}{g_1} \sum_{r \bmod p} \left(\frac{r}{p}\right) \exp\left(2\pi \mathrm{i} \frac{ar}{p}\right).$$

Hence:
$$\sum_{a=1}^{N} \left(\frac{a}{p}\right) = \frac{1}{g_1} \sum_{r \bmod p} \left(\frac{r}{p}\right) \sum_{a=1}^{N} \exp\left(2\pi \mathrm{i} \frac{ar}{p}\right).$$

For $r \not\equiv 0$, the inner sum is a geometric series with value

$$\exp\left(2\pi \mathrm{i} \frac{r(N+1)}{2p}\right) \frac{\exp(\pi \mathrm{i} \frac{rN}{p}) - \exp(-\pi \mathrm{i} \frac{rN}{p})}{\exp(\pi \mathrm{i} \frac{r}{p}) - \exp(-\pi \mathrm{i} \frac{r}{p})}.$$

This has absolute value $|\sin(\pi r N/p)|/|\sin(\pi r/p)|$.

Proof of the Pólya–Vinogradov inequality

Thus,

$$\left|\sum_{a=1}^{N}\left(\frac{a}{p}\right)\right| \le \frac{1}{|g_1|}\sum_{r=1}^{p-1}\frac{|\sin(\pi r N/p)|}{|\sin(\pi r/p)|} = \frac{1}{\sqrt{p}}\sum_{r=1}^{p-1}\frac{1}{|\sin(\pi r/p)|}.$$

Lemma

For any real number θ, we have

$$|\sin(\pi\theta)| \ge 2\|\theta\|,$$

where $\|\theta\|$ denotes the distance from θ to the nearest integer.

Proof of the Pólya–Vinogradov inequality

Lemma

For any real number θ, we have

$$|\sin(\pi\theta)| \geq 2\|\theta\|,$$

where $\|\theta\|$ denotes the distance from θ to the nearest integer.

Proof: Using periodicity mod 1 and the even-ness of both sides, it's enough to verify this for $0 \leq \theta \leq 1/2$. This amounts to proving

$$\sin(\pi\theta) \geq 2\theta,$$

which is an exercise in calculus.

Proof of the Pólya–Vinogradov inequality

It follows that

$$\sum_{r=1}^{p-1} \frac{1}{|\sin(\pi r/p)|} \leq \frac{1}{2}\sum_{r=1}^{p-1}\frac{1}{\|r/p\|}$$

$$= \sum_{r=1}^{(p-1)/2} \frac{1}{r/p} = p \sum_{r=1}^{(p-1)/2} \frac{1}{r}.$$

Putting this in above,

$$\left|\sum_{a=1}^{N}\left(\frac{a}{p}\right)\right| \leq \frac{1}{\sqrt{p}}\sum_{r=1}^{p-1}\frac{1}{|\sin(\pi r/p)|} \leq \sqrt{p}\sum_{r=1}^{(p-1)/2}\frac{1}{r}.$$

Proof of the Pólya–Vinogradov inequality

Putting this in above,

$$\left|\sum_{a=1}^{N}\left(\frac{a}{p}\right)\right| \le \frac{1}{\sqrt{p}}\sum_{r=1}^{p-1}\frac{1}{|\sin(\pi r/p)|} \le \sqrt{p}\sum_{r=1}^{(p-1)/2}\frac{1}{r}.$$

Proof of the Pólya–Vinogradov inequality

Putting this in above,

$$\left|\sum_{a=1}^{N}\left(\frac{a}{p}\right)\right| \leq \frac{1}{\sqrt{p}}\sum_{r=1}^{p-1}\frac{1}{|\sin(\pi r/p)|} \leq \sqrt{p}\sum_{r=1}^{(p-1)/2}\frac{1}{r}.$$

Using our knowledge of the partial sums of the harmonic series, the final sum is

$$\log\left(e^{\gamma+o(1)}\frac{p-1}{2}\right) < \log p,$$

for large p. (In fact, with some cleverness, one sees that this holds for all $p \geq 3$.)

This completes the proof of Pólya-Vinogradov.

Vinogradov's trick

I ask your indulgence for another digression.

Problem

For all integers $2 \leq n \leq N$, write down the largest prime factor of n. What is the median element of this list?

Vinogradov's trick

I ask your indulgence for another digression.

Problem

For all integers $2 \leq n \leq N$, write down the largest prime factor of n. What is the median element of this list?

Theorem

For any constant $A \geq 1/2$, the limiting proportion of $n \leq N$ with largest prime factor $> N^A$ is

$$\log \frac{1}{A}.$$

As a consequence, if $A > \frac{1}{\sqrt{e}}$, then the limit is strictly less than $1/2$, and if $A < \frac{1}{\sqrt{e}}$, the limit is strictly larger than $1/2$.

Proof of the theorem

Proof: Suppose $\frac{1}{2} \leq A \leq 1$. If an integer $n \leq N$ has a prime factor $p \geq N^A$, then p is the only prime factor which is that large. For each prime $p \in (N^A, N]$, the number of $n \leq N$ which are divisible by p is $\lfloor N/p \rfloor$. So with $P^+(\cdot)$ the largest prime factor function,

$$\begin{aligned}
\#\{n \leq N : P^+(n) > N^A\} &= \sum_{N^A < p \leq N} \lfloor N/p \rfloor \\
&\approx N \sum_{N^A < p \leq N} \frac{1}{p}.
\end{aligned}$$

Proof of the theorem

Proof: Suppose $\frac{1}{2} \leq A \leq 1$. If an integer $n \leq N$ has a prime factor $p \geq N^A$, then p is the only prime factor which is that large. For each prime $p \in (N^A, N]$, the number of $n \leq N$ which are divisible by p is $\lfloor N/p \rfloor$. So with $P^+(\cdot)$ the largest prime factor function,

$$\begin{aligned} \#\{n \leq N : P^+(n) > N^A\} &= \sum_{N^A < p \leq N} \lfloor N/p \rfloor \\ &\approx N \sum_{N^A < p \leq N} \frac{1}{p}. \end{aligned}$$

According to Mertens,

$$\sum_{p \leq x} \frac{1}{p} = \log\log x + C + O(1/\log x).$$

Proof of the theorem

So for the coefficient of N in the last estimate, we have

$$\sum_{N^A < p \leq N} \frac{1}{p} = \log\log N - \log\log N^A + o(1)$$
$$= \log\frac{1}{A} + o(1),$$

as $N \to \infty$.

This gives our claim that

$$\frac{1}{N}\#\{n \leq N : P^+(n) > N^A\} \to \log\frac{1}{A}.$$

More on the median

Remark

Eric Naslund, a UBC undergradute, has proved the following very nice result: Among the integers $2 \leq n \leq N$, the median largest prime factor is asymptotic to

$$e^{(\gamma-1)/\sqrt{e}} N^{1/\sqrt{e}}, \quad \text{as } N \to \infty.$$

In particular, the median is strictly less than $N^{1/\sqrt{e}}$ for large N.

Getting more from Pólya–Vinogradov

Our previous argument was very simple: If $N > \sqrt{p}\log p$, then

$$\sum_{n=1}^{N}\left(\frac{n}{p}\right) < \sqrt{p}\log p < N,$$

and so it cannot be that $\left(\frac{n}{p}\right) = 1$ for all $n \leq N$.

Getting more from Pólya–Vinogradov

Our previous argument was very simple: If $N > \sqrt{p}\log p$, then

$$\sum_{n=1}^{N}\left(\frac{n}{p}\right) < \sqrt{p}\log p < N,$$

and so it cannot be that $\left(\frac{n}{p}\right) = 1$ for all $n \leq N$.

Now we note a different consequence of P–V. Whenever $\frac{N}{\sqrt{p}\log p} \to \infty$ (for example, if $N = p^{1/2+\epsilon}$) we have

$$\sum_{n=1}^{N}\left(\frac{n}{p}\right) = o(N), \quad \text{as } p \to \infty.$$

That is, our Legendre symbol sums **display cancelation**.

Getting more from Pólya–Vinogradov

For concreteness, let $\epsilon > 0$, and take

$$N = p^{1/2+\epsilon}.$$

Since

$$\sum_{n \le N} \left(\frac{n}{p}\right) = o(N),$$

and since $\left(\frac{n}{p}\right) = \pm 1$ for $1 \le n < p$, it follows that (as $p \to \infty$), asymptotically 50% of the values $n \le N$ are squares mod p and 50% are non-squares.

The $p^{1/2\sqrt{e}}$ bound

We are now ready to prove the following result:

Theorem

For large p, we have $n_2(p) \le p^{\frac{1}{2\sqrt{e}}+\epsilon}$.

Proof: Let $N = p^{\frac{1}{2}+\epsilon}$ and let $M = p^{\frac{1}{2\sqrt{e}}+\epsilon}$. Notice that $M > N^{1/\sqrt{e}}$. In fact,

$$M > N^{\frac{1}{\sqrt{e}}+\frac{1}{100}\epsilon}.$$

This means that the proportion of $n \le N$ which have a prime factor $> M$ is below 50%; in fact, at most $(50-\eta)\%$ for some $\eta > 0$.

The $p^{1/2\sqrt{e}}$ bound

Suppose for the sake of contradiction that $n_2(p) > M$. Then every prime $q \leq M$ satisfies $\left(\frac{q}{p}\right) = 1$. So every integer $n \leq N$ composed only of primes $q \leq M$ also satisfies $\left(\frac{n}{p}\right) = 1$. But this accounts for at least $(50 + \eta)\%$ of the $n \leq N$.

But in the limit, only 50% of the $n \leq N$ should be squares mod p. So this is a contradiction once p is large.

When does cancelation "kick in" for character sums?

By Pólya–Vinogradov, as soon as N grows a bit faster than $p^{1/2}$,

$$\sum_{n=1}^{N} \left(\frac{n}{p}\right) = o(N).$$

The following is a consequence of some work of D.A. Burgess in the 1960s:

Theorem (Burgess)

As soon as N grows a bit faster than $p^{1/4}$,

$$\sum_{n=1}^{N} \left(\frac{n}{p}\right) = o(N).$$

Burgess's proof is intricate and not "elementary" in the usual sense of the word: A key innovation is the use of Weil's Riemann Hypothesis for curves (for certain hyperelliptic curves) to bound certain auxiliary sums.

Burgess's proof is intricate and not "elementary" in the usual sense of the word: A key innovation is the use of Weil's Riemann Hypothesis for curves (for certain hyperelliptic curves) to bound certain auxiliary sums. Applying Vinogradov's trick in the same manner as before, we halve the exponent:

Corollary

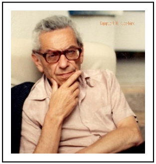
Let $\epsilon > 0$. For large primes p, we have

$$n_2(p) \le p^{\frac{1}{4\sqrt{e}}+\epsilon}.$$

Open problem

Remove the $+\epsilon$.

Average order of the least quadratic nonresidue



Theorem (Erdős)

We have

$$\lim_{x\to\infty}\left(\frac{1}{\pi(x)}\sum_{p\le x} n_2(p)\right) = A,$$

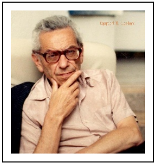
where

$$A := \sum_{k=1}^{\infty}\frac{p_k}{2^k},$$

and p_k *denotes the* k*th prime.*

This is what one would expect from the random-variables model.

Average order of the least quadratic nonresidue



Theorem (Erdős)

We have

$$\lim_{x\to\infty}\left(\frac{1}{\pi(x)}\sum_{p\le x} n_2(p)\right) = A,$$

where

$$A := \sum_{k=1}^{\infty}\frac{p_k}{2^k},$$

and p_k denotes the kth prime.

This is what one would expect from the random-variables model. **Proof**: Another time!

Thank you!