Average Number of Quadratic Frobenius Pseudoprimes

Andrew Fiori

PIMS Postdoctoral Fellow - University of Calgary

Spring 2016

This work is joint with Andrew Shallue.

Background - Fermat Pseudoprimes

Theorem

If p is prime and $\gcd(a, p) = 1$*, then* $a^{p-1} \equiv 1 \mod p$*.*

One application of Fermat's theorem gives a very efficient way to confirm that a number is not prime, unfortunately it cannot show that a number is prime.

Definition

Suppose n is composite and $a^{n-1} \equiv 1 \mod n$. Then n is called a base-a Fermat pseudoprime, or a-psp.

A very natural question to ask is, how likely is it that a number is pseudoprime? or rather, how many false positives are we likely to get?

Average Number of Fermat Pseudoprimes

We have a formula for $F(n)$, exact number of $a \pmod{n}$ for which n is a base a-psp:

$$F(n) = \prod_{p|n} \gcd(p-1, n-1)$$

Erdős and Pomerance showed the average value of $F(n)$ satisfies:

$$x^{15/23-o(1)} < \frac{1}{x}\sum_{n \le x} F(n) \le x \cdot \exp\left\{ \frac{-(1+o(1)) \log x \log\log\log x}{\log\log x} \right\}.$$

It is conjectured that $\frac{15}{23}$ can be replaced by 1.

However, if one wants concrete bounds for a fixed a, we have $\mathcal{P}_a(x)$, the number of a-psp up to some bound x, satisfies:

$$x^{1/3} \le \mathcal{P}_a(x) \le x/\sqrt{\exp(\log x \cdot \log\log\log x / \log\log x)}.$$

The lower bound based on the count of Carmichael numbers, the upper bound by Pomerance.

Other Fast Primality Tests

There are many natural generalizations of this test, such as Euler pseudoprimes or strong pseudoprimes, which will reduce the number of false positives. We will quickly introduce one test which moves us in the direction of quadratic Frobenius pseudoprimes.

Definition (Lucas Pseudoprime)

Consider the Fibonacci sequence f_n :

$$f_0 = 0 \quad f_1 = 1 \quad f_n = f_{n-1} + f_{n-2}.$$

Set $e_n = \left(\frac{5}{n}\right)$. If p is a prime other than 2 or 5 then it is known that:

$$f_{p-e_p} = 0 \pmod{p}$$

A composite number n which satisfies the above condition is called a Lucas pseudoprime.

There are no known Lucas pseudoprimes n for which $e_n = -1$. There is $620 waiting for the first person to find one (or prove they don't exist).

How does Lucas test work

Let α be one root of $x^2 - x - 1$.
If $e_p = -1$ then, $x^2 - x - 1$ does not factor over $\mathbb{F}_p$ and hence α generates $\mathbb{F}_{p^2}$. It follows that

$$\alpha^{p+1} = \alpha^p \cdot \alpha = N_{\mathbb{F}_{p^2}/\mathbb{F}_p}(\alpha) = -1 \pmod{p}.$$

One can readily observe that:

$$\alpha^{n+1} = f_{n+1}\alpha + f_n$$

from which we derive the conditions of the Lucas test (the $e_n = 1$ case has a different, but vaguely similar argument).

This test admits many generalizations, such as using other recurances, and refinements, including strong Lucas pseudoprimes and Lehmer pseudoprimes. We shall bypass these and pass straight to a rather broad generalization, Frobenius pseudoprimes.

Frobenius Pseudoprimes

Fix a polynomial $f(x)$. If p is a prime number, and $(\Delta_f f(0), p) = 1$ then we know that:

- For all d we have that:
$$d \mid \deg\left(\gcd_{\text{mod-p}}\left(f(x), \frac{x^{p^d}-1}{x^{p^{d-1}}-1}\right)\right)$$
- The map $x \mapsto x^p$ permutes the roots of each:
$$f_d(x) = \gcd_{\text{mod-p}}\left(f(x), \frac{x^{p^d}-1}{x^{p^{d-1}}-1}\right).$$
- We have the factorization $f(x) = \prod f_d(x) \pmod{p}$.
- The factorization structure of $f(x)$, that is the collection of degrees above, determines $\left(\frac{\Delta_f}{p}\right)$.

Without getting into the technical details, we will call a number n a Frobenius pseudoprime if it satisfies these four conditions.

Frobenius pseudoprimes for $x^2 - x - 1$ will be Lucas pseudoprimes (the converse does not hold).

Some natural questions

1) Fix n, are there quadratic $f(x)$ with (f, n) a Frobenius liar pair?
   Yes.
2) Do there exist such $f(x)$ with $\left(\frac{\Delta_f}{n}\right) = \pm 1$?
   Yes for $+1$, sometimes no for -1 (for example 21).
3) Given n, how many f are there modulo n?
   The number with $\left(\frac{\Delta_f}{n}\right) = 1$ is:

$$\frac{1}{2}\prod_{p|n}\frac{1}{2}\left((n-1,p^2-1)+(n-1,p-1)^2-2(n-1,p-1))\right)+\frac{1}{2}\prod_{p|n}\frac{1}{2}\left((n-1,p-1)^2-(n-1,p^2-1))\right)$$

The number with $\left(\frac{\Delta_f}{n}\right) = -1$ is:

$$\frac{1}{2}\prod_{p|n}\frac{1}{2}\left((\frac{n}{p}-1,p^2-1)+(n^2-1,p-1)-2(n-1,p-1)\right)-\frac{1}{2}\prod_{p|n}\frac{1}{2}\left(-(\frac{n}{p}-1,p^2-1)+(n^2-1,p-1)\right)$$

One can ask the following questions in the other direction:

4) Given f, are there n for which (f, n) gives a Frobenius liar pair?

Yes.

We just need a Carmichael number n for which f splits at all $p|n$.

5) Given f, are there n with $\left(\frac{\Delta_f}{n}\right) = \pm 1$ and the pair is a liar?

Unclear.

In particular the -1 case is unclear, the above argument handles the $+1$ case.

6) How many f are there on average? that is, what are:

$$\frac{1}{x}\sum_{n<x} F_+(n) \qquad \text{and} \qquad \frac{1}{x}\sum_{n<x} F_-(n)?$$

7) How many n are there on average? that is, what are:

$$\frac{1}{x^2}\sum_{n<x} F_+(n) \qquad \text{and} \qquad \frac{1}{x^2}\sum_{n<x} F_-(n)?$$

These last two questions are what me and Andrew Shallue are working on.

How to count liars for a fixed n

Fix n, and $p|n$, suppose we want to construct a polynomial $f(x)$ which is a liar for n. What can we say about the roots α and β of $f(x)$ modulo p?

The values of $\left(\frac{\Delta_f}{n}\right)$ and $\left(\frac{\Delta_f}{p}\right)$ restrict the possible multiplicative orders of α and β modulo p. These orders must always divide $\gcd(n^2-1, p^2-1)$. However, in the various cases the order must specifically divide (or not divide) some of $n-1$, $p-1$ or $n-p$.

For $\left(\frac{\Delta_f}{n}\right) = +1$ we obtain terms like:

$$\gcd(n-1, p-1)^2 \quad \text{and} \quad \gcd(n-1, p^2-1)$$

Whereas for $\left(\frac{\Delta_f}{n}\right) = -1$ we obtain terms like:

$$\gcd(n^2-1, p-1) \quad \text{and} \quad \gcd(n^2-1, p^2-1, n-p)$$

By the Chinese remainder theorem, one can combine the options from each p.

Optimizing this bound - Erdős-Pomerance

We now want to use these formulas to obtain lower bounds on:

$$\sum_{n\leq x} F(n).$$

We will use the strategy of Erdős-Pomerance, namely to get a large collection of n for which the $F(n)$ are large. We do this by optimizing the contribution from almost all $p|n$.

We will sketch the argument for the discriminant -1 case, where by taking $p+1$ smooth we can ensure:

$$\gcd(n^2-1, p^2-1, n-p) > \gcd(n+1, p+1) = p+1.$$

The discriminant $+1$ case is similar, except one takes $p-1$ smooth to ensure:

$$\gcd(n-1, p-1)^2 = (p-1)^2.$$

The Setup

- Let $M(x) = \mathrm{LCM}\left(1, \ldots, \frac{\log(x)}{\log\log(x)}\right)$, and fix α such that:
$$|\{p < \log^{\alpha}(x) \mid p+1|M(x)\}| = (\log(x))^{\alpha - o(1)}.$$
Note that $M(x) = x^{o(1)}$, and it is known that such $\alpha > 1$ exist.
- Let $0 < \epsilon < \alpha - 1$ be arbitrarily small.
- Set:
$$P(x) = \{\log^{\alpha-\epsilon}(x) < p < \log^{\alpha}(x) \mid p+1|M(x)\}.$$
By construction $|P(x)| = (\log(x))^{\alpha - o(1)}$.
- Set k to be the largest odd number less than $\frac{\log(x) - 5\log(M(x))}{\log\log(x) + \log(\alpha)}$.
- Let $S(x)$ be the collection of integers which are the product of exactly k distinct elements of $P(x)$.
By construction $|S(x)| = x^{1-\alpha^{-1}-o(1)}$ and the elements $s \in S(x)$ all satisfy $x^{1-o(x)} < s < \frac{x}{M^5}$.

For each $s \in S(x)$, set $q(s)$ to be the smallest prime such that $sq(s) = -1 \pmod{M}$ (note $q(s) < M^5$).
We then have $x^{1-o(x)} < sq(s) < x$ and $n = sq(s)$ has at least:

$$\begin{aligned} F_-(n) &> \prod_{p' \neq p \mid s} \tfrac{1}{2}(\gcd(n^2-1, p^2-1, n-p) - \gcd(n-1, p-1)) \\ &> 2^{-k} \frac{1}{\log^\alpha(x)} \prod_{p \mid s} (p-1) \\ &= x^{1-o(1)} \end{aligned}$$

many liars for which the discriminant is -1.

Note, we may need to exclude one prime factor p' of s, this explains the $\log^\alpha(x)$ factor.

The Average Number of Liars

Denoting by $F_{\pm}(n)$ the number of quadratic Frobenius liar pairs (f, n) for which $\left(\frac{\Delta_f}{n}\right) = -1$ the argument above gave us the following:

$$\sum_{n \leq x} F_{-}(n) > x^{2-\alpha^{-1}-o(1)}.$$

A similar argument for the $+1$ discriminant case would yield:

$$\sum_{n \leq x} F_{+}(n) > x^{3-\alpha^{-1}-o(1)}.$$

In the formulas above, α is at least $\frac{23}{8}$ by the work of Balog. It is conjectured that α can be taken to be arbitrarily large.

Note, improving the 2 to a 3 in the -1 case is plausible from the formula and can be done if one assumes conjectures about either smooth values of polynomials evaluated at primes in arithmetic progressions, or the average number of prime factors of smooth values of polynomials evaluated at primes.

Conclusions

The results we have suggest:

- For a randomly chosen n of size x you can expect at least $x^{2-\alpha^{-1}-o(1)}$ many liars with $\left(\frac{\Delta_f}{n}\right) = 1$.
- For a randomly chosen n of size x you can expect at least $x^{1-\alpha^{-1}-o(1)}$ many liars with $\left(\frac{\Delta_f}{n}\right) = -1$.
- For a randomly chosen f with coefficients less than x you can expect at least $x^{1-\alpha^{-1}-o(1)}$ many liars with $\left(\frac{\Delta_f}{n}\right) = 1$.
- For a randomly chosen f with coefficients less than x you can expect at least $x^{-\alpha^{-1}-o(1)} > 0$ many liars with $\left(\frac{\Delta_f}{n}\right) = -1$.

Note these are only lower bounds, and conjecturally the -1 bounds can be improved to match the $+1$ bounds. But the sharp difference between what is proven in these cases really does hint that for fixed f it is harder to find liars in the -1 in the $+1$ case.

The End.

Thank you.