Average Number of Quadratic Frobenius Pseudoprimes

Andrew Fiori

University of Lethbridge

Fall 2017

This work is joint with Andrew Shallue.

Context and Motivation

It is well known that if p is a prime number then $a^p = a \pmod{p}$.

A consequence is that if $a^n \neq a \pmod{n}$ then n is not a prime.
A useful feature of this is that it gives a computationally efficient way to rule out primality.

However, it is not an if and only if... so we define:
A number n is a **base a (Fermat)-pseudoprime** if $a^n = a \pmod{n}$ (and a and n are coprime).

How often does this fail?

Let $F(n) = |\{a \pmod n \mid n \text{ is a base } a \text{ pseudoprime}\}|$.

- Monir showed that the number of failures for a fixed n is:

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

- Erdős and Pomerance showed that (with $\alpha = \frac{23}{15}$) we have:

$$x^{2-\alpha^{-1}-o(1)} < \sum_{n<x} F(n) < x^2 e^{-\log(x)\frac{\log\log\log(x)}{\log\log(x)}}.$$

(and conjectures one can take $\alpha \to \infty$).

But neither result actually tells us that if I fix a, if there are any base a Fermat pseudoprimes.
However, a result of Alford, Granville, Pomerance gives that $\mathcal{P}_a(x)$, the number of a-psp up to some bound x, satisfies:

$$x^{1/3} \leq \mathcal{P}_a(x) \leq xe^{-\log x \cdot \frac{\log\log\log x}{2\log\log x}}.$$

Are there better tests?

Depends what you mean by better...

- Efficiently computable.
- Fewer false positives.

What other tests are there?

- Run the test again with another choice of a. But heuristics + Bayes rule, and experimentation suggests this is ineffective.
- There are a bunch of refinements to this tests (consider $a^{(p-1)/2}$). But these tend to have the similar asymptotic failure rates.
- You can look at other divisibility relations like those that come from recurrence relations. (For example the Lucas test) These are slightly slower, but we have no idea how often they fail.

The Lucas Test.

If p is prime then it is "well known" that p divides the $(p - (5 \mid p))$th Fibonacci number.
This gives another primality test.
It turns out there are no known composite numbers n such that:

- n is a strong Fermat pseudoprime for base 2.
- n "is innert" in $\mathbb{Q}(\sqrt{5})$ (ie check the Jacobi symbol $(5 \mid n) = -1$).
- n does divides the $(n+1)$st Fibonacci number.

Just something to think about:
Recurance relations have closed forms... so what does that tell us about this test?

These sorts of observations lead to a more general class of test.

The Frobenius Test.

This was first proposed by Grantham, and tries to use the maximum amount of information you can extract about how Frobenius should behave on the roots of a polynomial... (Given that we don't really know how that polynomial splits modulo p).

Definition (Grantham):

Let $f(x) \in \mathbb{Z}[x]$ be a monic polynomial of degree d and discriminant Δ. Then composite n is a Frobenius pseudoprime with respect to $f(x)$ if the following conditions all hold.

1. (Integer Divisibility) We have $\gcd(n, f(0)\Delta) = 1$.
2. (Factorization) Let $f_0(x) = f(x) \pmod{n}$. Define $F_i(x) = \gcd(x^{n^i} - x, f_{i-1}(x))$ and $f_i(x) = f_{i-1}(x)/F_i(x)$ for $1 \leq i \leq d$. All of the gcds exist and $f_d(x) = 1$.
3. (Frobenius) For $2 \leq i \leq d$, $F_i(x) \mid F_i(x^n)$.
4. (Jacobi) Let $S = \sum_{2|i} \deg(F_i(x))/i$. Have $(-1)^S = (\Delta \mid n)$, where $(\Delta \mid n)$ is the Jacobi symbol.

The Frobenius Test for degree 2 polynomials.

For $d = 1, 2$ we define:

$$f_d(x) = \gcd_{\text{mod-n}}\left(f(x), \frac{x^{n^d-1}-1}{x^{n^{d-1}-1}-1}\right)$$

Then the conditions are:

- Either $f_1 = f$ or $f_2 = f$ (and the other is 1) depending on $(\Delta_f \mid n)$ being $+1$ or -1.
- If $f_2 = f$ the map $x \mapsto x^n$ permutes the roots.

As we can see the conditions greatly simplify in this case and we get different natural assertions about the multiplicative order of the roots of f based on two natural cases.

These conditions can also be mostly interpretted as giving independent conditions for each $p|n$.

Monir type formula

We consider seperately the number of f modulo n which "split" vs are "innert". For square free n we will ultimately obtain:

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

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

Contribution as it depends on $(\Delta_f \mid n)$ and $(\Delta_f \mid p)$

In each case the modulo p contribution comes from conditions on the multiplicative order of the roots (or as an assertion about one root determining the other).

	+	-
+	Two distinct elements of $\mathbb{F}_p$ whose order divides $n-1$	One element of $\mathbb{F}_{p^2} \setminus \mathbb{F}_p$ whose order divides $n-1$, choice symmetric in picking α^p
-	One element of $\mathbb{F}_p$ whose order divides n^2-1 but not $n-1$, choice is symmetric in picking α^n	One element of $\mathbb{F}_{p^2} \setminus \mathbb{F}_p$ whose order divides n^2-1 and $n-p$, choice symmetric in picking α^p.

	+	-
+	$\frac{1}{2}((n-1,p-1)^2-(n-1,p-1))$	$\frac{1}{2}((n-1,p^2-1)-(n-1,p-1))$
-	$\frac{1}{2}((n^2-1,p-1)-(n-1,p-1))$	$\frac{1}{2}((n^2-1,p^2-1,n-p)-(n-1,p-1))$

Note that $p-1$ is the order of $\mathbb{F}_p^*$ and p^2-1 is the order of $\mathbb{F}_{p^2}^*$.

The only other ingredient you need for the formula is 'anti-symmetrization' to account for the relationship between $(\Delta_f \mid n)$ and $\prod_{p|n}(\Delta_f \mid p)$.

One surprise is that these numbers can be zero. ($F_-(n) = 0$ infinitely often.)

Erdos-Pomerance type formula

We have the following bounds on the average:

$$x^{3-\alpha^{-1}-o(1)} < \sum_{n<x} F_{\pm}(n) < x^3 e^{-\log(x)\frac{\log\log\log(x)}{\log\log(x)}}$$

* best known α depends on $\pm$

The strategy of proof is very similar to that of Erdos-Pomerance.

Sketch:

1. There are lots (in a sense related to α) of p such that $p+1$ and $p-1$ are smooth (divisible by only "small" primes)
2. There are lots (roughly $x^{1-\alpha^{-1}}$) of ways of making $n_0 \sim x^{1-\epsilon}$ which are only products of these p
3. I can find auxilliary primes q_i so that $n = n_0 q_1 ... q_n$ has $n+1$ and $n-1$ also smooth.
4. These numbers n has will have aproximately $n^{2-\epsilon}$ many liars.
5. This gives me the lower bounds of the theorem.

What we need about smoothness

For classical pseudoprimes you set $M(x) = \prod_{p<\log(x)} p$ and you need there to exist a value $\alpha > 1$ so that the set

$$|\{\text{prime } p < (\log(x))^{\alpha} \text{ such that } (p-1) \mid M(x)\}| > \log(x)^{\alpha - o(1)}$$

(Erdős proved the first such result, Balog proved the current best).

In our case we need to factor $M(x) = M_1(x)M_2(x)$ so that $(M_1(x), M_2(x)) = 2$ and

$$|\{\text{prime } p < (\log(x))^{\alpha} \text{ such that } (p-1) \mid M_1(x) \text{ and } (p+1) \mid M_2(x)\}|$$

has the same type of growth rate (note that the best known α will be lower).
Using results of Dartyge-Martin-Tenenbaum which consider $p^2 - 1$ smooth, we can prove the above with $\alpha = 4/3$.

The effect is that there are $x^{1-\alpha^{-1}-o(1)}$ many ways to construct the desired numbers n_0 which have size at least $x^{1-\epsilon}$.

How we use this

In the classical case, by picking q prime so that $n_0 q - 1 = 0 \pmod{M(x)}$ I immediately force all the gcds of $(n_0 q - 1, p - 1)$ to be large so that Monir's formula is maximized for all $p | n_0$.

In this case we may need two primes to account for the fact that the Monir type formula can give zero. These primes are easily chosen from a congruence modulo $M(x)$ (based on ones modulo $M_1(x)$ and $M_2(x)$.
The effect is that we have $x^{1-\alpha^{-1}-o(1)}$ many numbers $n = n_0 q_1 q_2$ with $x^{2-\epsilon-o(1)}$ many liars each. This gives us the lower bound by taking $\epsilon \to 0$.

Frobenius-Carmichael Numbers

Grantham already proved that there cannot be any in the most naive form of the natural definition.
However, if we allow a slightly weaker definition:

Fix a field K and consider only those polynomials f which define K

Then these could still exist.

Our attempts to prove any bounds on this (using the ideas of Alford-Granville-Pomerance) have failed.

However, one can get partial results in this direction.

Frobenius-Carmichael Numbers

n can't be both a classical Carmichael number, and a Frobenius-Carmichael number for a quadratic field in which it is "innert" unless every $p|n$ is actually innert. In fact, there would be 0 many liars which generate such a field!
Just observe the tension between the different options in the boxes:

	+	-
+	$\frac{1}{2}((n-1,p-1)^2-(n-1,p-1))$	$\frac{1}{2}((n-1,p^2-1)-(n-1,p-1))$
-	$\frac{1}{2}((n^2-1,p-1)-(n-1,p-1))$	$\frac{1}{2}((n^2-1,p^2-1,n-p)-(n-1,p-1))$

Any one entry being large tends to give *less space* for other entries to be large.
This tension gives at least some heuristic support to the following hybrid test.

Baillie PSW Test

The Baillie PSW (Pomerance-Selfridge-Wagstaff) test is a hybrid test which basically functions as follows:

1. Do a (strong) Fermat pseudoprime test base 2.
2. Do the Lucas test for the recurrence with smallest descriminant such that n is "innert".

The connection to the Frobenius test is that the Quadratic Frobenius test is strictly stronger than the Lucas test (or the same if you use a certain modified version of the Lucas test).
However, this test is known to deterministically identify primes up to at least 2^{50} (the bound is probably higher now).
Our bounds tend to suggest this test should fail eventually, but the heuristic I just mentioned suggests it might be a while.

The End.

Thank you.