Industrial Grade Primes
with a Money-Back Guarantee

Michael P. Abramson

Abstract

A subset of the integers is exhibited for which the converse of Fermat's Little Theorem holds. Strong evidence is given that this set contains infinitely many primes, though a proof of this is known to be very hard.

Many methods for generating large prime numbers begin with Fermat's Little Theorem:

**Theorem 1 (Fermat).** If \( n \) is prime and \( a \) is relatively prime to \( n \), then

\[
a^{n-1} \equiv 1 \pmod{n}.
\]

(1)

The converse statement is:

*If \( a^{n-1} \equiv 1 \pmod{n} \) for some \( a \) relatively prime to \( n \), then \( n \) is prime.*

The converse is false. However, it is false so rarely that Henri Cohen jokingly coined the term *industrial grade prime* to denote any number \( n \) for which \( 2^{n-1} \equiv 1 \pmod{n} \) (see [5, p. 5]). The choice of \( a = 2 \) is what is usually done in practice, and even though we write \( a \) everywhere for generality, we are thinking of \( a = 2 \). Many prime generation methods start with this converse, and then add a secondary test to eliminate the composite numbers for which (1) holds. A composite number \( n \) which satisfies (1) for some \( a \) relatively prime to \( n \) is called a base \( a \) pseudo-prime. If \( n \) is a base \( a \) pseudo-prime for all \( a \) relatively prime to \( n \), then \( n \) is called a Carmichael number. Pseudo-primes and Carmichael numbers have been studied extensively [4]; in fact, it was only recently proved that there are an infinite number of Carmichael numbers [1]. Instead of studying when composite numbers fail (1), we investigate the following dual question:

**Question.** When does satisfying (1) guarantee primality for \( n \)?

Or, in the colloquial language of [3], when does an industrial grade prime come with a money-back guarantee that it is prime? In this note, we will describe a classic method of prime number generation, and show when its secondary test can be eliminated, thus exhibiting a set of integers for which passing (1) guarantees primality. We then give strong evidence for the conjecture that our set contains infinitely many primes.

1 A Theorem from Hardy and Wright.

We present here a slight generalization of part of Theorem 101 in [2]. Before doing so, we recall the definition of the Euler \( \phi \)-function: \( \phi(n) \) is the
number of positive integers less than or equal to \( n \) which are relatively prime to \( n \). Euler showed that \( a^{\phi(n)} \equiv 1 \pmod{n} \) for all \( a \) relatively prime to \( n \). This is a generalization of Fermat's Little Theorem, because \( \phi(n) = n - 1 \) for prime \( n \).

**Theorem 2.** Let \( n = hp + 1 \), where \( p \) is prime and \( h \) is an even positive integer such that \( h < 4p + 4 \). If \( n \) satisfies (1) for some \( a \) relatively prime to \( n \) and if
\[
a^h \not\equiv 1 \pmod{n},
\]
then \( n \) is prime.

**Proof:** Assume \( n \) is not prime. Let \( x \) be the order of \( a \) modulo \( n \); that is, \( x \) is the smallest positive integer such that \( a^x \equiv 1 \pmod{n} \). Then \( x \) divides every integer \( k \) for which \( a^k \equiv 1 \pmod{n} \). In particular, (1) and (2) imply that \( x \) divides \( hp \) but not \( h \). This means that \( x \) is a multiple of \( p \) or, in other words, \( p \) divides \( x \). Similarly, \( a^{\phi(n)} \equiv 1 \pmod{n} \) implies that \( x \) divides \( \phi(n) \) and, by transitivity, \( p \) divides \( \phi(n) \). Let \( n = q_1^{e_1}q_2^{e_2} \cdots q_k^{e_k} \) be the unique prime factorization of \( n \). Using properties of Euler's \( \phi \)-function, one can show that
\[
\phi(n) = \prod_{i=1}^{k} q_i^{e_i-1}(q_i - 1).
\]
Since \( p \) does not divide \( n \), \( p \) cannot divide any factor \( q_i \) of \( n \). But \( p \) does divide \( \phi(n) \); hence, \( p \) divides \( q_i - 1 \) for some \( i \). Let \( P \) denote this prime \( q_i \). Then \( P \) is a factor of \( n \) such that \( P \equiv 1 \pmod{p} \). Since \( n \) is not prime, we must have \( n = PM \) for some \( M > 1 \). Now \( n \equiv 1 \equiv P \pmod{p} \), and since \( n \) and \( p \) are odd, we see that \( P \equiv 1 \equiv M \pmod{2p} \). Thus, \( hp + 1 = n = PM = (2pu + 1)(2pv + 1) \) for some \( u, v \geq 1 \), so \( h = 4puv + 2u + 2v \geq 4p + 4 \), contradicting our hypothesis. Therefore, \( n \) is prime.

We can generate large primes by iterating Theorem 2; that is, at each step, the newly found prime \( n \) plays the role of \( p \) in the next step, and \( h \) is randomly chosen until we find a new prime \( n \).

2 Eliminating the Secondary Test.

The next theorem gives one way of eliminating the secondary test (2).

**Theorem 3.** Let \( n = 2pq + 1 \), where \( p \) and \( q \) are odd primes satisfying
\[
\frac{1}{2}(p - 2) < q < 2(p + 1).
\]
Then \( n \) is prime if and only if (1) holds for some \( a \) with \( a^2 \not\equiv 1 \pmod{n} \).

(Note that the condition \( a^2 \not\equiv 1 \pmod{n} \) is really not restrictive, especially since \( a = 2 \) is usually used in practice.)

**Proof:** If \( n \) is prime, then (1) holds by Fermat's Little Theorem. Conversely, suppose (1) holds for some \( a \) with \( a^2 \not\equiv 1 \pmod{n} \). If \( p \equiv \pm 1 \pmod{3} \),
then \(2p^2 + 1 = 2(3k + 1)^2 + 1 = 3(6k^2 + 4k + 1)\) for some \(k > 1\). Then since \(p\) is prime, \(2p^2 + 1\) is prime if and only if \(p = 3\). Hence, we may assume \(p \neq q\). Writing \(h = 2q\) and \(k = 2p\), condition (3) implies \(n = hp + 1 = kq + 1\) with \(h < 4p + 4\) and \(k < 4q + 4\). Suppose \(a^h \equiv a^k \equiv 1 \pmod{n}\). Then \(a^{2p} \equiv a^{2q} \equiv 1 \pmod{n}\), so \(a^2 = a^{\gcd(2p, 2q)} \equiv 1 \pmod{n}\), a contradiction. Therefore, \(a^h \not\equiv 1 \pmod{n}\) or \(a^k \not\equiv 1 \pmod{n}\). Either way, \(n\) is prime by Theorem 2.

Unfortunately, the method of generating primes suggested by iterating Theorem 3 may be impractical because two primes are needed to construct one new prime at each step. Practicality aside, if we let

\[
\begin{align*}
\overline{M} &= \{2pq + 1: p \text{ and } q \text{ are odd primes}\} \\
M &= \{2pq + 1 \in \overline{M}: p \text{ and } q \text{ satisfy (3)}\},
\end{align*}
\]

then \(M\) has the property that a given integer \(n\) in \(M\) is prime if and only if (1) holds. In other words, \(M\) contains no base pseudo-primes. In order for the set \(M\) to be of any real interest, \(M\) must contain infinitely many primes.

**Conjecture.** The sets \(M\) and \(\overline{M}\) contain infinitely many primes.

### 3 Evidence for the Conjecture.

Let \(p_k\) represent the \(k^{th}\) odd prime. Figure 1 depicts \(M\) and \(\overline{M}\), where the axes represent the indices of the odd primes \(p_i\) and \(p_j\). \(M\) is then represented by the shaded region and \(\overline{M}\) is represented by the entire (square) region inside the axes. Theorem 3 asserts that the primality of integers of the form \(n = 2p_i p_j + 1\), where \((i, j)\) lies in the shaded region of Figure 1, can be determined using only Fermat's Little Theorem.

![Figure 1: The region (3)](image)

Since primes tend to be uniformly distributed in intervals, we expect that the ratio of the number of primes in \(M\) to the number of primes in \(\overline{M}\) should be close to the ratio of the areas of \(M\) (the shaded region) to \(\overline{M}\) (the whole square). Let \(R(t)\) be this ratio of areas of \(M\) and \(\overline{M}\), where \(t\) is the
size of the square ($t = 1000$ in Figure 1). We approximate the boundaries of the shaded area in Figure 1 by (least-squares) lines through the origin $j = a_1(t)i$ and $j = a_2(t)i$, where $a_1$ and $a_2$ are dependent on $t$. Then we can approximate the desired area of the shaded region by subtracting the areas of the two triangles, $\frac{a_1(t)t^2}{2}$ and $\frac{t^2}{2a_2(t)}$, from the area $t^2$ of the square to obtain $t^2 \left(1 - \frac{a_1(t)}{2} - \frac{1}{2a_2(t)}\right)$. Thus,

$$R(t) \approx 1 - \frac{a_1(t)}{2} - \frac{1}{2a_2(t)}.$$  

We note here that although $R(t)$ is dependent on $t$ the dependence is very small: as $t$ increases, the slopes $a_1$ and $a_2$ will change only slightly. For $t = 1000$ as depicted in Figure 1, $a_1 = .5645$ and $a_2 = 1.788$. Thus, $R(1000) \approx .438$. Let $R_p(t)$ be the true ratio of the number of primes in $M$ to the number of primes in $\overline{M}$. $R_p(t)$ can be easily computed for specific values of $t$. For example, $R_p(1000) \approx .422$. Figure 2 below shows both $R$ and $R_p$ as $t$ increases up to $t = 1000$. Note that $R(t)$ looks like a horizontal line because of the very slight dependence on $t$.

Figure 2: The functions $R$ and $R_p$

Although these approximations are somewhat crude, the point of this discussion is not the actual values of the ratios, but rather that the ratios appear to be positive numbers, well away from zero. If we could prove rigorously that $\lim_{t \to \infty} R_p(t) > 0$, then this would imply that $M$ and $\overline{M}$ both contain finitely many primes or both contain infinitely many primes.

We now introduce two functions to study the growth of these primes. Let $Q(k)$ be the number of odd primes $q$ such that (3) holds (with $p = p_k$) and $2p_kq + 1$ is prime. We also need a function that counts all such primes in $M$ up to $p_k$. To prevent over-counting, we define $\overline{Q}(k)$ to be the number of odd primes $q$ such that $\frac{1}{2}(p_k^2 - 2) < q < p_k$ and $2p_kq + 1$ is prime and let $\sigma(k) = \sum_{i=1}^{k} \overline{Q}(i)$. Figure 3 on the next page shows the functions $Q(k)$
and \( \sigma(k) \) for \( k \leq 2000 \). The apparent growth of these functions leads to the stronger conjecture that \( \sigma(k) \geq k \) for all \( k > 0 \); we have verified this statement computationally for the first 100,000 odd primes \( p_k \).

![Figure 3: The functions \( Q(k) \) and \( \sigma(k) \)](image)

In this paper, we have exhibited a set of integers whose primality can be determined using only Fermat’s Little Theorem. Although we have not proved that there are infinitely many primes in this set, we have given strong experimental evidence for it. Proving there are infinitely many primes of a form which is not linear is generally a very hard problem that requires deep results from analytic number theory.

Acknowledgments. The author thanks Robert L. Ward, who generalized Theorem 2 as stated, and Ezra Brown for kind comments on expository style.

References


Michael P. Abramson  
National Security Agency  
9800 Savage Road, Suite 6704  
Ft. Meade, MD 20755, USA  
mabramso@restarea.ncsc.mil