1 Primality Testing

First, we will treat the various methods of determining whether or not a number is prime, or, as is usually easier, determining whether a number is composite. At first glance, it seems strange to have methods for such a thing that are distinct from factorization methods - after all, if you can obtain the prime factorization of a number, it is trivial to determine primality from that factorization. However, factorization methods are in general very computationally expensive compared to some of the methods we will examine in this section.

1.1 Primality Tests and Composite Tests

Any viable computational method for testing primality explicitly is composed of a condition on a number n that, when met, necessitates that n is prime. Thus if the condition is not met, n is composite. Such tests are certainly extremely convenient to directly determine the primality of n, but unfortunately methods of this form are usually a combination of very complex and restricted to n of a particular form or within some bounded range.

In addition to these primality tests, which never fail on determining primality, we also have many tests that are comparatively simple and computationally efficent, which occasionally fail to identify primes, but never indicate that a composite number is prime. We will call these tests, which never fail on determining compositeness, compositeness tests.

It is of the utmost importance to note, before moving on, that while the conditions of a primality or compositeness test being met guarantees that the number is either prime or composite respectively, and a failed primality test proves compositeness, a failed compositeness test does not necessitate that the number is prime.

1.2 The Sieve of Eratosthenes

Sieving is a process where a series of operations are applied to every number in a large, regularly spaced set of integers in order to find numbers with certain characteristics.

The Sieve of Eratosthenes is among the first algorithmic methods for factorization and primality testing. Admittedly, it is very crude, but the theoretical basis of the method gives us some of the fundamental ideas for developing further primality tests. Additionally, the Sieve of Eratosthenes also can be used for algorithmic factorization of a number - as such, we will treat it rather lightly here and revisit it in more detail when we examine factorization methods.

The core use of the Eratosthenes' sieve in primality testing stems from the idea of *trial division*, where given an integer n, we attempt to divide n by every integer that could possibly be a factor. If the number divides n, then it is a factor of n, and because possible factors must be bounded above by n itself, we are guaranteed a finite number of computations. In fact, we need only test factors less than $\sqrt{n}$, recovering greater factors as the quotient of a successful

division by a lesser factor. Additionally, after each successful division, only the quotient need be tested further, resulting in easier computation as more factors are found. None of these facts are particularly reassuring computationally, but they provide some background for the use of our sieve.

The sieve can be thought of as a list or array of $N-1$ consecutive integers, beginning with 2. We recognize that 2 is prime, and thus every multiple of 2 is composite. We then remove 4,6, etc. from our list, and are left with 2 and the odd integers. We then recognize 3 as prime, and remove all multiples of 3 in the same way. 5 has not been removed from the list, and so it is prime. We then repeat our method. Continuing this way until the end of the set has been reached, we have constructed a set of all prime numbers up to N.

Note that the sieve has other modifications that can be made to find things like – the least prime factor of each composite up to N, or even the complete factorizations of the numbers in the sieve.

It is a fair observation that this is not at all an individualized test, but rather construction of a set that will prove a number n prime if n is an element of the set, and composite if $n < N$ and n is not in the set. This is not ideal, but combined with our method of trial division, and given that 76% of odd integers have a prime factor less than 100, precompiling a list of primes up to a certain bound and performing trial divisions of only those prime numbers on a given n can prove many numbers composite without having to resort to more rigorous primality tests.

1.3 Fermat's Theorem and Resulting Methods

Among the many limitations of Eratosthenes' sieve is its lack of precision - it is impossible to simply prove a given number n prime or composite with the sieve, you must construct the entire set of primes up to n in the worst case to say anything about n at all. We will next move to our very first formal test, a compositeness test based on a theorem of Fermat.

In passing through the list, we are adding p to find the next multiple. In addition to that, we noted earlier that we are guaranteed a finite number of computations. Therefore, this sieve has the benefit of not being particularly computationally intensive. In practice, the problem is the amount of memory it can consume. This is dependent upon the size of the list of numbers. There are work-arounds for this, like array segmentation, but this may affect the efficiency of the sieve.

1.3.1 Fermat's Theorem

Fermat's Theorem. *If a positive integer p is a prime and $(a,p)=1$, then*

$$a^{p-1} \equiv 1 \pmod{p}$$

Fermat's theorem cannot be used as a primality test, as we will examine further when we define pseudoprimes, but we can use the converse of the theorem to develop a test for compositeness.

Converse of Fermat's Theorem. *If n and a are positive integers such that $(a, n) = 1$ and*

$$a^{n-1} \not\equiv 1 \pmod{n},$$

then n is not a prime. Hence n is composite.

1.3.2 Pseudoprimes and Carmichael Numbers

There are composite numbers that will behave similarly to primes when Fermat's theorem is applied to them for certain a. We will call these composites *Fermat pseudoprimes*.

Definition. *A Fermat pseudoprime base a is an odd composite number m for which*

$$a^{m-1} \equiv 1 \pmod{m}$$

holds.

In general, a *pseudoprime* is a composite number that behaves like a prime in the context of a primality or compositeness test. We will see other pseudoprimes in the future. If it were somehow possible to characterize all Fermat pseudoprimes for a base a, we could then use Fermat's Theorem with base a as a primality test - this is approachable for fixed length numbers.

The question arises - can all Fermat pseudoprimes be detected through a change of base? The answer, unfortunately, is no. There exist numbers that will be Fermat pseudoprimes base k for all valid k, and these numbers motivate the following definition.

Definition. *A composite number n such that $a^{n-1} \equiv 1 \pmod{n}$ for all a relatively prime to n is called a Carmichael Number.*

A Carmichael number, the smallest of which is 561, will never be revealed as composite by a Fermat's Theorem test. The existence of Carmichael numbers are a serious hindrance of the Fermat's theorem method, although fortunately they can be characterized.

Theorem. *n is a Carmichael number if and only if $p-1$ divides $n-1$ for every prime factor of n, n is composite, and n is squarefree.*

Proof. First, suppose that n is a composite number that square free, and $p-1 \mid N-1$ for every prime factor p of n, and that a is an arbitrary integer such that $(a, n) = 1$.

n is square free, so $n = p_1p_2p_3 \dots p_k$. Because a is relatively prime to n, a is also relatively prime to p_i for all $i \in 1, \dots, k$. Thus, $a^{p-1} \equiv 1 \pmod{p_i}$ for all i by Fermat's Theorem. However, because $p-1 \mid n-1$, $a^{n-1} = (a^{p-1})^t \equiv (1)^t = 1 \pmod{p_i}$ for some $t \in \mathbb{Z}$. Because $p_i \mid a^{n-1} - 1$ for all i, and all p_i are distinct primes, $n = p_1p_2p_3 \dots p_k \mid a^{n-1} - 1$. Thus $a^{n-1} \equiv 1 \pmod{N}$ for any arbitrary a relatively prime to n, and n is a Carmichael number.

We now suppose that a composite n is a Carmichael number, so $a^{n-1} \equiv 1 \pmod{n}$ for all a relatively prime to n. $n = p_1^{\alpha_1} \cdot \dots p_k^{\alpha_k}$, and so $p_i^{\alpha_i} \mid a^{n-1} - 1$

and $(a, p_i) = 1$ for all $i \in \{1, \ldots, k\}$. Let r be an integer such that $N = p_i^{\alpha_i} r$. Obviously, $p_i^{\alpha_i}$ and r are relatively prime, so by the chinese remainder theorem we have a unique solution of the system

$$x \equiv a_1 \pmod{p_i^{\alpha_i}}$$

$$x \equiv 1 \pmod{r}$$

where a_1 is a primitive root of $p_i^{\alpha_i}$.

Suppose that $\alpha_i = 1$. x is coprime to both p_i and r, so we have $x \equiv a \pmod{p_i \cdot r = n}$, and so $x^{n-1} \equiv 1 \pmod{n}$, and so $x^{n-1} \equiv 1 \pmod{p}_i$ as well. However, $x^{p_i - 1} \equiv 1 \pmod{p_i}$, and $p_i - 1$ is the least power of x such that x is equivalent to 1 modulo p_i, and so we have $p_i - 1 \mid N - 1$ for all i.

Now suppose that $\alpha_i > 1$ for some i, or that n is not squarefree. $x^{n-1} \equiv 1 \pmod{p_i^{\alpha_i}}$, and by Euler's theorem, $x^{\phi(p_i^{\alpha_i})} = x^{p_i^{\alpha_i - 1}(p_i - 1))} \equiv 1 \pmod{p_i^{\alpha_i}}$. Thus because $p_i^{\alpha_i - 1}(p_i - 1)$ is the least power of a_1 equivalent to 1 modulo pi^{α_i}, $p_i^{\alpha_i - 1}(p_i - 1) \mid n - 1$. However, $p_i^{\alpha_i - 1} \mid n - 1$ is a contradiction, because $(n, n - 1) = 1$ and $p_i^{\alpha_i - 1} \mid n$. Thus n must also be squarefree, and we are done. □

1.3.3 Improvements due to Eulers Criterion

There is clear motivation to improve our method, if possible, to avoid the problems presented by Carmichael numbers. One tool that we have at our disposal is the *Euler Criterion for Quadratic Residues*, which is is stated below and proven in the Background section.

Euler's Criterion for Quadratic Residues. *If p is an odd prime and $(a, p) = 1$,*

$$a^{\frac{n-1}{2}} \equiv \pm 1 \pmod{p}$$

This criterion's logical converse, much like Fermat's theorem, gives us a test for compositeness.

Euler's Criterion as a Compositeness Test. *If n is odd, $(a, n) = 1$, and*

$$a^{\frac{n-1}{2}} \not\equiv \pm 1 \pmod{n}$$

then n is a composite number.

To strengthen this compositeness test, we use *Jacobi's symbol*, $(\frac{a}{n})$, over Legendre's symbol.

Euler's Criterion as a Compositeness Test. *If n is odd, $(a, n) = 1$, and*

$$a^{\frac{n-1}{2}} \not\equiv \pm 1 \pmod{n}$$

then n is a composite number.

If $a^{\frac{n-1}{2}} \equiv \pm 1 \pmod{n}$ *and*

$$a^{\frac{n-1}{2}} \not\equiv (\frac{a}{n}) \pmod{n}$$

then n *is composite.*

However, if $a^{\frac{n-1}{2}} \equiv (\frac{a}{n}) \pmod{n}$ *holds, the test is inconclusive.*

As the last line of our definition implies, this test also has pseudoprimes associated with it. We will call these *Euler pseudoprimes.*

1.3.4 Euler Pseudoprimes and Strong Pseudoprimes

We define an Euler pseudoprime in an analogous way to our definition of a Fermat pseudoprime.

Definition. *Let* n *be an odd composite number such that*

$$a^{\frac{n-1}{2}} \not\equiv (\frac{a}{n}) \pmod{n}$$

for some a *relatively prime to* n*. Then* n *is an Euler pseudoprime base* a*.*

Unfortunately, there are some numbers that are both Carmichael numbers and Euler pseudoprimes. An example of one such number is 1729 -

Despite this disappointing fact, many Carmichael numbers are revealed as composite by Euler's criterion, and additionally, there is no analogue for Carmichael numbers in Euler pseudoprimes. If enough bases are tested, we will eventually prove an Euler pseudoprime composite.

To close this section, we present a final type of pseudoprime.

Definition. *An odd composite number* n *with* $n - 1 = d \cdot 2^s$*,* d *odd, is called a strong pseudoprime for base* a *if either* $a^d \equiv 1 \pmod{n}$ *or* $a^{d \cdot 2r} \equiv -1 \pmod{n}$*, for some* $r \in \{0, 1, \ldots, s-1\}$*.*

This test for pseudoprimes is intended, much like Euler pseudoprimes, to eliminate the problem of Carmichael numbers in the Fermat Compositeness test. Indeed, any Fermat pseudoprime will be eventually proven composite by the strong pseudoprime test if enough bases a are tested.

The motivation for the definition is as follows - Any Fermat pseudoprime n to base a will satisfy the equivalence

$$a^{n-1} - 1 \equiv 0 \pmod{n}$$

Because we assume that n is odd if it is being tested for primality, $n = 2m + 1$ for some integer m and we have

$$a^{2m} - 1 = (a^m - 1)(a^m + 1) \equiv 0 \pmod{()n}$$

If n is a prime, it must divide one of these factors, but it cannot divide both because then it would need to divide all linear combinations of them, including $(a^m - 1) - (a^m + 1) = -2$. Therefore we have

$$a^m \equiv \pm 1 \pmod{n}$$

We can write n as $2^\alpha k + 1$, where k is odd, and have

$$a^{n-1} - 1 = (a^k - 1)(a^k + 1)(a^{2k} + 1) \dots (a^{2^{\alpha-1}k} + 1)$$

If n divides exactly one of these factors but is composite, then it is a strong pseudoprime. Interestingly, if a number is a strong pseudoprime to the base a, it will also be a Euler pseudoprime to A.

1.4 Proving Primality

In the previous section, all the methods we examined could only possibly prove compositeness - because we are unable to test infinitely many bases in an Euler or Strong Pseudoprime test, there is no way for us to computationally prove that a number is prime using these methods, only to say that it is very likely to be prime, or prove that it is composite.

In this section, we examine some methods of proving primality. These primality tests are much more complicated computationally and theoretically than compositeness tests, and unfortunately usually depend on factorization of a number, which is a slow and laborious process. However, they are theoretically motivating, despite not being as generally popular as more modern primality tests.

1.4.1 Lehmer's Theorem

The first method of actually proving primality that we will look at is based on a theorem by Lucas, and proven by Lehmer.

Lehmer's Theorem. *Suppose* $n - 1 = \prod_{j=1}^{n} q_j^{\beta_j}$, *with all* q_j *distinct primes. If an integer* a *exists such that*

$$a^{\frac{n-1}{q_j}} \neq 1 \pmod{n} \text{ for all } j = 1, \dots, n$$

and such that

$$a^{n-1} \equiv 1 \pmod{n},$$

then n *is a prime number.*

We note that this theorem is an extension of Fermat's theorem.

Proof. Consider $a^k \equiv 1 \pmod{n}$. The smallest such k divides all possible k, including $n - 1$. However, every divisor k not equal to $n - 1$ of $n - 1$ divides at least one $\frac{n-1}{q_j}$. Thus if k is less than $n - 1$, $a^{\frac{n-1}{q_j}} \equiv a^{tk} \equiv 1 \pmod{n}$ will hold for some q_j. This contradicts the assumptions of the theorem, and so $k = n - 1$. However, by Euler's theorem, we know that $a^{\phi(n)} \equiv 1 \pmod{n}$ for all a relatively prime to n, and that $\phi n < n - 1$ for all composite numbers. If $\phi n < k$, it contradicts the assumption that k is the smallest power of a equivalent to 1, and thus $\phi n = k = n - 1$, and n must be prime. □

From our knowledge of primitive roots, we know that any primitive root will be a suitable a for Lehmer's Theorem. The problem inherent in this is that there is not an efficient, deterministic method to find primitive roots, or even quadratic non-residues. If n is prime, then we will have $\phi(\phi(n)) = \phi(n-1) = (n-1)\prod(1-\frac{1}{q_j}$ primitive roots, but as the equation shows, if $n-1$ has many factors, especially many small factors, primitive roots become a smaller proportion of possible a values.

Fortunately, Selfridge provides us with a relaxed version of Lehmer's theorem such that we do not need a primitive root to prove n prime.

Theorem. *Suppose* $n-1 = \prod_{j=1}^{n} q_j^{\beta_j}$*, with* q_j *all distinct primes. If for every* q_j*, there is an* a_j *such that*

$$a_j^{\frac{n-1}{q_j}} \not\equiv 1 \pmod{n}$$

while

$$a_j^{n-1} \equiv 1 \pmod{n},$$

then n *is a prime.*

This allows us to perform tests with multiple bases, and prove primality usually much more quickly than Lehmer's theorem applied directly, especially for primes where the least primitive root is large.

However, there is still a major stumbling block in practical applications of Lehmer's Theorem - the need to factorize $n-1$. Luckily, it is possible to relax the conditions of Lehmer's theorem in a way that allows us to only partially factorize $n-1$.

Theorem. *Let* $n-1 = r \cdot f$*, where* f *is the factorized part of* $n-1$ *and* r *is the unfactorized or remaining part,* $(r, f) = 1$,$r < f$*.* $F = \prod_{j=1}^{n} q_j^{\beta_j}$*, with all* q_j *distinct primes.*

If an integer a *exists such that*

$$gcd(a^{\frac{n-1}{q_j}} - 1, N) = 1 \text{ for all } j,$$

and

$$a^{n-1} \equiv 1 \pmod{n},$$

then n *is a prime number.*

1.4.2 Pépin's and Proth's Theorems

Another possible sidestep of the need to factor $n-1$ arises when n happens to be a Fermat number, or a number of the form $2^{2^r}+1$ (These happen to be the only sums of a power of 2 and 1 that can be prime). In this case, $n-1$ has the minimum possible number of distinct prime factors, and thus is very suited to

a Lehmer's Theorem test. In fact, for a Fermat number n, the requirements of Lehmer's Theorem become

$$a^{\frac{n-1}{2}} = a^{2^{2^r-1}} \not\equiv 1 \pmod{n}$$

and

$$a^{n-1} = a^{2^{2^r}} \equiv 1 \pmod{n},$$

which essentially reduces to

$$x^2 \equiv 1 \pmod{n}$$

and

$$x \not\equiv 1 \pmod{n}$$

where $x = a^{2^{2^r-1}}$.

If n is prime, then $x^2 \equiv 1 \pmod{n}$ will have exactly two solutions - $x \equiv \pm 1 \pmod{n}$. $x \equiv 1 \pmod{n}$ violates the conditions of Lehmer's theorem, and so $x \equiv -1 \pmod{n}$ when n is prime. This motivates a need for an a such that

$$x \equiv a^{\frac{n-1}{2}} \equiv -1 \pmod{n}$$

Noting that Euler's Criterion gives us that a must be a quadratic non-residue of n. The fact that 3 is a quadratic non-residue of all primes of the form $12n \pm 5$ and the observation that $2^{2^k} \equiv 4 \pmod{1}2$ for all k motivates the idea of (but does not prove) Pépin's Theorem, which is stated below.

Pépin's Theorem. *A Fermat number $n = 2^{2^r} + 1$, $n \geq 1$ is prime if and only if*

$$a^{\frac{n-1}{2}} \equiv -1 \pmod{n}$$

It is worth mentioning that the binary computer makes the division by n necessary to reduce powers of a modulo n very computationally simple, due to n being a power of 2 plus 1 and therefore very simply expressed in binary. This makes Pépin's test a very attractive test for Fermat number primes.

With both Pépin's Theorem and relaxed version of Lehmer's theorem, we obtain Proth's Theorem, stated below.

Proth's Theorem. *Suppose n is of the form $n = m \cdot 2^k + 1$, with $2^k > m$ and m odd. If there exists an integer a such that*

$$a^{\frac{n-1}{2}} \equiv -1 \pmod{n},$$

then n is prime.

There is another way to possibly ease the difficulty of factorizing $n - 1$, namely by using a different compositeness test than Fermat's theorem, that allows us to attempt to factorize a number other than $N - 1$, which may prove easier to factor. For example, primality tests based on Lucas Sequences give an analogue to Lehmer's theorem in quadratic fields, where $N + 1$ may be factored instead of $N - 1$. The theory of quadratic fields is beyond the scope of this paper, and so we do not address Lucasian or similar methods, but their existence is worth noting.

1.5 Modern Primality Tests

Of the primality tests covered in the previous section, none are of practical use for all numbers. Computationally viable methods such as Pépin's Theorem only apply to primes of a certain form, and the need for factorization and primitive roots cripples Lehmer's Theorem computationally.

1.5.1 The Jacobi Sum Primality Test

1.5.2 Lenstra's Theorem

1.5.3 Elliptic Curve Primality Testing

1.5.4 The Goldwasser-Kilian Test

1.5.5 Atkin's Test

1.6 Closing Notes on Primality Testing