Miller-Rabin Primality Testing and the Extended Riemann Hypothesis

David Brandfonbrener
Math 354

May 7, 2017

It is an important problem in number theory as well as computer science to determine when an integer is prime rather than composite. This is a problem that people have been working on for hundreds of years, and now is made even more interesting with the reliance of cryptography on large prime numbers. Currently, one of the most popular algorithms to do this is a probabilistic algorithm called Miller-Rabin primality testing that was first published by Rabin in 1980 [8]. This algorithm is based off an earlier, deterministic algorithm by Miller in 1976 [6]. However, the runtime of Miller's algorithm depends on the Extended Riemann Hypothesis (ERH). This connection is especially interesting because it is not obvious why the ERH and primality testing would be connected. This paper will be broken into three sections: first some preliminaries including a statement of the main theorem and motivation and statement of the algorithm, second a proof of the connection to the ERH, and last a proof that the algorithm works as desired assuming the ERH. The first and third sections mainly follow Miller [6], while the second mainly follows Montgomery's proof [7] of a theorem by Ankeny [1].

1 Preliminaries

Definition 1. An algorithm *tests primality in time $O(f(n))$* if there exists a deterministic Turing machine that implements the algorithm in $c \cdot f(n)$ steps for some constant c.

Notation. Let n, the number to be tested for primality be an odd integer. Note that $\lceil \log n \rceil$ is the size of the binary representation for n.

Definition 2. A *Dirichlet character* is a function $\chi : \mathbb{Z} \to \mathbb{C}$ such that

1. There exists a positive integer k such that $\chi(m) = \chi(m + k)$ for all integers m.

2. If $(m, k) > 1$ then $\chi(m) = 0$, and if $(m, k) = 1$ then $\chi(m) \neq 0$.

3. For all integers a, b, $\chi(ab) = \chi(a)\chi(b)$

And a *Dirichlet L-function*, $L(s, \chi)$, is defined for $s \in \mathbb{C}$ with $\text{Re}(s) > 1$ and χ a Dirichlet character by the analytic continuation to a meromorphic function on the complex plane of

$$\sum_{m=1}^{\infty} \frac{\chi(m)}{m^s}.$$

The existence of this continuation is proved in [5].

Extended Riemann Hypothesis (ERH). For $s \in \mathbb{C}$ and χ a Dirichlet character, the zeros of $L(s, \chi)$ in the critical strip, where $0 \leq \text{Re}(s) \leq 1$, all lie on the line $\text{Re}(s) = 1/2$.

Assuming this hypothesis, we will get the following main result.

Theorem 1. (Miller [6]) (ERH) There exists an algorithm which tests the primality of n in $O(\log(n)^4 \log\log\log(n))$ steps.

To get this result, we must present an algorithm that realizes this runtime. First, we will motivate such an algorithm.

Motivation of Algorithm. We have by Fermat's little theorem that for p prime and $(a, p) = 1$ that

$$a^{p-1} \equiv 1 \ (p).$$

By the contrapositive of this statement, if for some $1 < a < n$ we have that $a^{n-1} \not\equiv 1 \ (n)$, then n is composite. So, the idea is to test different values for a, and determine whether $a^{n-1} \equiv 1 \ (n)$. There are two reasons that this strategy may fail

1. If n is composite, but for all a with $(a, n) = 1$, $a^{n-1} \equiv 1 \ (n)$, the strategy fails.

2. If the first a such that $a^{n-1} \not\equiv 1 \ (n)$ is very large, it will be inefficient to find such an a.

The goal of Miller's algorithm is to solve both of these problems. Before we give the algorithm, we need some more notation, and will provide an example of the first problem.

Definition 3. Let $n = p_1^{t_1} \cdots p_m^{t_m}$ be the prime factorization of our *odd* integer n. Then define the following two functions from $\mathbb{Z} \to \mathbb{Z}$

(1) Carmichael's λ-function. $\lambda(n) = \text{lcm}\{p_1^{t_1-1}(p_1-1), \ldots, p_m^{t_m-1}(p_m-1)\}$.

(2) $\lambda'(n) = \text{lcm}\{p_1 - 1, \ldots, p_m - 1\}$.

With this definition, we introduce a theorem for precisely when the first problem with our testing strategy occurs.

Theorem 2. (Carmichael [4]) $a^{n-1} \equiv 1 \ (n)$ for all a with $(a,n) = 1$ if and only if $\lambda(n) \mid n-1$.

Proof. ($\Rightarrow$). Assume that $a^{n-1} \equiv 1 \ (n)$ for all a with $(a,n) = 1$. Let $n = p_1^{t_1} \cdots p_m^{t_m}$ be the prime factorization of n. For each i, let b_i be a generator of the unit group mod $p_i^{t_i}$, which is cyclic of order $\phi(p_i^{t_i}) = (p_i - 1)p_i^{t_i-1} = \lambda(p_i^{t_i})$. Note that the $p_i^{t_i}$ are relatively prime to each other, so by the Chinese Remainder Theorem (CRT) the system of congruences $b \equiv b_i \ (p_i^{t_i})$ has a unique solution b mod n. Since each b_i was a unit mod $p_i^{t_i}$ there exist c_i such that $b_i c_i \equiv 1 \ (p_i^{t_i})$ and by CRT there is a c such that $c \equiv c_i \ (p_i^{t_i})$ for all i. Then, $bc \equiv 1 \ (p_i^{t_i})$ for each i and thus by CRT $bc \equiv 1 \ (n)$, so b is a member of the unit group mod n. Therefore, by the assumption, $b^{n-1} \equiv 1 \ (n)$, and thus for each i, $b_i^{n-1} \equiv 1 \ (p_i^{t_i})$. Since each b_i has order $\lambda(p_i^{t_i})$, we have that $\lambda(p_i^{t_i}) \mid (n-1)$ for each i. Therefore, we get the desired conclusion that $\lambda(n) = \text{lcm}\{\lambda(p_1^{t_1}), \ldots, \lambda(p_m^{t_m})\} \mid (n-1)$.

($\Leftarrow$). Assume that $\lambda(n) \mid n-1$. Let $n = p_1^{t_1} \cdots p_m^{t_m}$ be the prime factorization of n. Now take any a with $(a,n) = 1$. Then, as above, we have that $\lambda(p_i^{t_i}) = \phi(p_i^{t_i})$ for each i. So, by Euler's Theorem, we have that $a^{\lambda(p_i^{t_i})} \equiv 1 \ (p_i^{t_i})$ for each i. Note the $p_i^{t_i}$ are relatively prime and $\lambda(p_i^{t_i}) \mid \lambda(n)$ so that $a^{\lambda(n)} \equiv 1 \ (p_i^{t_i})$ for each i. So, by CRT we have that $a^{\lambda(n)} \equiv 1 \ (n)$. And by the above assumption, $\lambda(n) \mid n-1$ so that $a^{n-1} \equiv a^{\lambda(n)k} \equiv 1^k \equiv 1 \ (n)$, as desired.

□

Example 1. One example of a composite number n that satisfies $a^{n-1} \equiv 1 \ (n)$ for all a with $(a,n) = 1$ is $n = 1105$. Here we have that $1105 = 5 \cdot 13 \cdot 17$. Then $\lambda(1105) = \text{lcm}\{4, 12, 16\} = 16$ and $(1105-1)/16 = 69$, so by the above theorem, 1105 satisfies Fermat's congruence. Now the goal of the algorithm will be to identify these numbers as composite quickly even if they satisfy Fermat's congruence.

Now we will present the algorithm, and then appeal to the Riemann hypothesis to show that the second concern with the strategy is not worrisome. First, we introduce one more piece of notation.

Notation. Let the number of times that 2 divides n be denoted $\#_2(n) = \max\{k : 2^k | n\}$.

Algorithm A_f. Let f be a computable function on the natural numbers. Then define the algorithm A_f as follows (where maximum number of iterations of the loop depends on f):

1. Check if $n = a^b$ for $b \geq 2$ by binary search for a for each $b < \lceil \log(n) \rceil$. If n is a prime power, then return “composite” and halt.
2. Compute $p_1, \ldots, p_m$ where p_i is the ith prime number and m is chosen so that $p_m \leq f(n) < p_{m+1}$. Then compute $S = \#_2(n-1)$ so that $n - 1 = Q2^S$ where Q is odd. Let $i = 1$ and proceed to (b) (letting $a = p_i$ throughout):
   (a) If $i < m$ set i to $i + 1$. If $i = m$, output “prime” and halt.
   (b) If $a|n$ output “composite” and halt.
   (c) Compute $a^Q\ (n), a^{Q2}\ (n), \ldots, a^{Q2^S} = a^{n-1}\ (n)$.
   (d) If $a^{n-1} \not\equiv 1\ (n)$ output “composite” and halt.
   (e) If $a^Q \equiv 1\ (n)$ go to (a).
   (f) Let $J = \max\{j : a^{Q2^j} \not\equiv 1\ (n)\}$. If $a^{Q2^J} \equiv -1\ (n)$ go to (a).
   (g) Output “composite” and halt.

Now we move on to presenting a preliminary lemma that will help to bound the number of a we need to check in our algorithm. First we need more definitions.

Definition 4. For primes p, q, a *qth non-residue mod p* is some value a such that there does not exist b with $b^q \equiv a\ (p)$ where $a \not\equiv 0\ (p)$.

Definition 5. Let p be prime and b be a generator of $(\mathbb{Z}/p\mathbb{Z})^\times$. Then define the index $\text{ind}_{b,p}a = \min\{m : b^m \equiv a\ (p)\}$.

Lemma 1. If $\lambda'(n) \not| \, n-1$ then there exists p, q prime such that:

(1) $p \mid n$, $p - 1 \not| \, n - 1$, $q^m \mid p - 1$, and $q^m \not| \, n - 1$ for some integer $m \geq 1$.

(2) if a is a qth non-residue mod p, then $a^{n-1} \not\equiv 1\ (n)$.

Proof. (1) Let $q_1, \ldots, q_m$ be the distinct prime divisors of n. Now $\lambda'(n) \not| \, n-1$, so $\text{lcm}\{q_1 - 1, \ldots, q_m - 1\} \not| \, n - 1$. Thus, there exists some i such that $q_i - 1 \not| \, n - 1$. Define $p = q_i$. Then, we have that $p \mid n$ and $p - 1 \not| \, n - 1$. Since $p - 1 \not| \, n - 1$, by prime factorization there must exist a prime q and $m \geq 1$ such that $q^m \mid p - 1$ but $q^m \not| \, n - 1$. Choose such a q and we have (1).

(2) Assume that a is a qth non-residue mod p. By means of contradiction, let $a^{n-1} \equiv 1\ (n)$. Then, by (1) we have that $p \mid n$, so that $a^{n-1} \equiv 1\ (p)$. Then, let b be a generator mod p, so we have that $b^{(\text{ind}_{b,p}a)(n-1)} \equiv a^{n-1} \equiv 1\ (p)$. Now, since $b^m \equiv 1\ (p)$ implies that $m \mid p - 1$ since b was a generator mod p, we have that

$$(p-1) \mid (\text{ind}_{b,p}a)(n-1) \tag{i}$$

Since a is a qth non-residue mod p, and $b^{\text{ind}_{b,p}a} \equiv a\ (p)$, we have $q \not| \text{ind}_{b,p}a$. If it did, then for some k, $qk = \text{ind}_{b,p}a$, and then $(b^k)^q \equiv a\ (p)$, which is impossible. So, we have that this fact combined with (1) yields

$$q \not| \text{ind}_{b,p}a \qquad \text{and} \qquad q^m \mid p-1 \tag{ii}$$

Combining (i) and (ii) yields $q^m \mid n-1$, which is a contradiction with (1), giving us (2). □

Definition 6. For p, q prime and $q \mid p-1$, let $N(p,q)$ be the least a such that a is a qth non-residue mod p. Note that a must be prime, since if a were composite such that $a = rs$ with $r, s \neq 1$, then since $r, s < a$ they must be qth residues mod p. So, there exist b, c such that $b^q \equiv r\ (p)$ and $c^q \equiv s\ (p)$, but then $(bc)^q \equiv rs \equiv a\ (p)$, a contradiction, so a must be prime.

Theorem 3. (Ankeny [1]) (Assuming ERH) $N(p,q) = O(\log(p)^2)$.

Note that with Ankeny's Theorem and Lemma 1 we have that if $\lambda'(n) \not| n-1$ that there exists $a \leq O(\log(n)^2)$ such that $a^{n-1} \not\equiv 1\ (n)$.

2 Connection to ERH

First, we will attempt to motivate the connection between least non-residues and the Extended Riemann Hypothesis. The connection stems from the fact that the Legendre symbol, Jacobi symbol, and power residue symbol are Dirichlet characters. Let q be prime with χ one of these characters mod q. So, we want to bound the largest value of N such that for all primes $p \leq N$ we have $\chi(p) = 0$ or 1, this will give us an upper bound on the least pth non-residue mod q. To do this, we will consider the sum

$$\sum_{p \leq N} (1 - \frac{p}{N})(\log p)\chi(p)$$

The proof comes from showing that cN is less than this sum for some constant $c > 0$ on the basis of the prime number theorem. And that the sum is equal to $O(N^{1/2}\log(q))$ on the basis of the ERH, which gives us the result.

Ankeny's proof in [1] is elaborate and relies on some more convoluted facts proved by Selberg. So instead I will present following proof of Montgomery from [7] that is somewhat simpler. Before we begin, it will be useful to define some notation:

Notation. Let a Dirichlet character χ be non-principal if there exists some integer k such that $\chi(k) \neq 0$ and $\chi(k) \neq 1$. Moreover, χ is primitive if it is not induced by a character of smaller modulus dividing the modulus of χ. Define the Von Mangoldt function $\Lambda : \mathbb{N} \to \mathbb{R}$ by

$$\Lambda(n) = \begin{cases} \log(p) & \text{if } n = p^k \text{ with } k \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Let $\rho = \beta + i\gamma$ denote a zero of $L(s, \chi)$ in the critical strip, with $0 \leq \rho \leq 1$. Let $\sum_\rho$ be a sum over all such zeros. Define the function $N(\sigma, t, \chi)$ to be the number of zeros ρ of $L(s, \chi)$ with $\sigma \leq \beta \leq 1$ and $-t \leq \gamma \leq t$.

Now we will need a few lemmas before proceeding to the theorem. This section will follow ideas from Davenport [5]. The proofs of these lemmas are not worked out in their entirety here as they are somewhat involved (and take up most of Davenport's textbook). But, these are standard results in analytic number theory and that textbook provides the proofs in full detail.

Lemma 2. (Montgomery [7], Davenport [5]) For $N > 1$ and χ primitive we have the formula:

$$\sum_{n \leq N} (1 - \frac{n}{N}) \Lambda(n) \chi(n) = -\sum_\rho \frac{N^\rho}{\rho(\rho + 1)} + O(\log qN).$$

Proof. The idea here is to begin with an analog to Perron's formula and then follow a traditional proof of the explicit formula for Chebyshev's function. So, to begin, we have the analog to Perron's formula that for $c > 0$:

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{y^s}{s(s+1)} ds = \begin{cases} 0 & 0 < y \leq 1 \\ 1 - \frac{1}{y} & y > 1. \end{cases}$$

This formula is proved by integrating over the rectangle with vertices $c-iT, c+iT, d+it, d-iT$ as $T \to \infty$ with $d > c$ for the case where $0 < y < 1$ and then the rectangle with vertices $c - iT, c + iT, -d + it, -d - iT$ where $-d < -1$ so that the contour contains both polls of the function $\frac{y^s}{s(s+1)}$. Essentially, it is the contribution of the polls that yields the value of $1 - \frac{1}{y}$, with the 1 from the pole at 0 and the $\frac{-1}{y}$ from the pole at y.

Now we want to prove the above formula. First, we note that $L(s, \chi)$ has an Euler product formula (much in the same manner as the Riemann zeta function) of

$$L(s, \chi) = \prod_{p \ prime} \frac{1}{1 - \chi(p) p^{-s}}.$$

Then, we can take the logarithmic derivative of this formula to get that for $\sigma = Re(s) > 1$ we have that:

$$-\frac{L'(s, \chi)}{L(s, \chi)} = \sum_{n=1}^{\infty} \Lambda(n) \chi(n) n^{-s}.$$

Keeping $\sigma > 1$ and letting $c > 1$, we use the logarithmic derivative formula, the fact that if $n < N$ then $\frac{N}{n} = y > 1$, and the analog to Perron's formula to get that:

$$\sum_{n \leq N} (1 - \frac{n}{N})\Lambda(n)\chi(n) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \sum_{n=1}^{\infty} \frac{\Lambda(n)\chi(n)}{n^s} \frac{N^s}{s(s+1)} ds$$
$$= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \left(\frac{L'(s,\chi)}{L(s,\chi)} \right) \frac{N^s}{s(s+1)} ds.$$

Now we have an expression that is easier to deal with. To evaluate this integral, we consider the rectangle with vertices $c - iT, c + iT, -M + iT, -M - iT$, call it R. The idea is to let $M, T \to \infty$ and show that the above integral is then equal to the residues of $\left(\frac{L'(s,\chi)}{L(s,\chi)} \right) \frac{N^s}{s(s+1)}$ inside of this rectangle. Here we will not be particular about ensuring that the borders of the rectangle do not pass through poles of the function, but know that a careful proof of the lemma would need to do so, and that these details are in Davenport [5]. Another detail we will omit is the case of a double poll at 0, which must be dealt with separately.

Avoiding those complications, the residues associated with each zero ρ of $L(s, \chi)$ will be $\frac{N^\rho}{\rho(\rho+1)}$. The residues at 0, -1 from $\frac{1}{s(s+1)}$ will be $\frac{L'(0,\chi)}{L(0,\chi)}$ and $\frac{L'(-1,\chi)}{L(-1,\chi)}$ respectively. And the trivial zeros correspond to residues of $\frac{x^{a-2m}}{2m-a}$ where a is 0 or 1 depending on whether $\chi(-1) = 1$ or -1. For the full details see section 19 in Davenport [5]. This gives us that the residues will sum to

$$-\sum_{\rho} \frac{N^\rho}{\rho(\rho+1)} - \frac{L'(0,\chi)}{L(0,\chi)} - \frac{L'(-1,\chi)}{L(-1,\chi)} + \sum_{m=1}^{\infty} \frac{x^{a-2m}}{2m-a}.$$

This gives us the desired term we were looking for. All that remains is to show that the integration over the other bounds of the rectangle is not substantial and that the last three terms of the above sum also reduce to $O(\log qN)$. Going through all of this in detail would exceed the scope of this paper, but it important to note that most of the bounds result from using the functional equation for $L(s, \chi)$ that:

$$L(1 - s, \chi) = \varepsilon(\chi) 2^{1-s} \pi^{-s} q^{s-1/2} \cos(\frac{\pi}{2}(s - a)) \Gamma(s) L(s, \overline{\chi}).$$

Where $|\varepsilon(\chi)| = 1$, a is as above, and Γ is the gamma function. This representation of L allows us better access to bounding L'/L since we know more properties of Γ. Again, the details of these excluded steps are important and non-trivial, but there is not space here to explain them fully. What this proof showed was where the main term of our sum comes from to give us the formula that:

$$\sum_{n \leq N} (1 - \frac{n}{N})\Lambda(n)\chi(n) = -\sum_{\rho} \frac{N^\rho}{\rho(\rho+1)} + O(\log qN).$$

□

Lemma 3. (Montgomery [7], Davenport [5]) For $t \geq 2$ we can bound the number of zeros in the critical strip by:

$$N(0, t+2, \chi) - N(0, t, \chi) = O(\log qt).$$

Proof. By equation (1) in section 16 of [5] we have that for $t \geq 2$ that

$$N(0, t, \chi) = \frac{t}{\pi} \log(\frac{qt}{2\pi}) - \frac{t}{\pi} + O(\log t + \log q).$$

So performing the desired subtraction, we clearly get that

$$\begin{aligned} N(0, t+2, \chi) - N(0, t, \chi) &= \frac{2}{\pi}(\log(\frac{q(t+2)}{2\pi}) - \log(\frac{qt}{2\pi})) - \frac{2}{\pi} + O(\log(qt)) \\ &= \frac{2}{\pi}(\log(\frac{(t+2)}{t}) + O(\log(qt)) = O(\log(qt)). \end{aligned}$$

Deriving the above formula is complicated, but the general idea is as follows. First, we must define the function $\xi(s, \chi)$ based on the functional equation for L where $a = 0$ or 1:

$$\xi(s, \chi) = (q/\pi)^{\frac{1}{2}s + \frac{1}{2}a} \Gamma(\frac{1}{2}s + \frac{1}{2}a) L(s, \chi).$$

Now the proof relies on considering the variation in arg $\xi(s, \chi)$ over s along the rectangle R defined by the vertices $\frac{5}{2} - it$, $\frac{5}{2} + it$, $\frac{-3}{2} + it$, $\frac{-3}{2} - it$. This strategy uses the argument principle from complex analysis (which is again relying on the logarithmic derivative) to count the zeros of the function. This region includes 1 trivial zero at either -1 or 0, so the remaining zeros are the ones we want to count, ie we have:

$$2\pi(N(0, t, \chi) + 1) = \Delta_R \arg \xi(s, \chi).$$

Note that the contribution of the left half of the contour is equal to that of the right because for come constant c we have:

$$\arg \xi(\sigma + it, \chi) = \arg \overline{\xi(1 - \sigma + it, \chi)} + c.$$

Now, using the definition of ξ, we can form two parts of the argument relatively easily from complex analysis as:

$$\begin{aligned} \Delta \text{arg } (q/\pi)^{\frac{1}{2}s + \frac{1}{2}a} &= t \log(q/\pi) \\ \Delta \text{arg } \Gamma(\frac{1}{2}s + \frac{1}{2}a) &= t \log(\frac{1}{2}t) - t + O(1). \end{aligned}$$

Now, we have that the total argument will be twice the sum of these two factors plus arg $L(s, \chi)$. So, we already have the dominant terms of the sum, it just remains to show that arg $L(s, \chi) = O(\log t + \log q)$ for $\sigma = 5/2$. This turns out to again be fairly complicated and this paper does not have the space to prove this in sufficient detail. However, the proof relies on a lemma that says that for $\rho = \beta + i\gamma$ the non-trivial zeros of $L(s,\chi)$ and t real we have that

$$\sum_{\rho} \frac{1}{1+(t-\gamma)^2} = O(\log(q|t|)).$$

This fact is proved by considering the real part of the logarithmic derivative of $L(s,\chi)$ with respect to the functional equation for $\xi(s,\chi)$ that

$$-Re\,\frac{L'(s,\chi)}{L(s,\chi)} = \frac{1}{2}\log\frac{q}{\pi} + \frac{1}{2}Re\,\frac{\Gamma'(\frac{1}{2}s+\frac{1}{2}a)}{\Gamma(\frac{1}{2}s+\frac{1}{2}a)} - Re\,\frac{\xi'(0,\chi)}{\xi(0,\chi)} - Re\,\sum_{\rho}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right)$$

But, elaborating this fully is beyond the scope of this paper, full details can be found in sections 12, 14, and 16 of Davenport [5]. This yields the result we wanted, that:

$$N(0,t+2,\chi) - N(0,t,\chi) = O(\log qt).$$

□

With these lemmas, we can now proceed to proving the theorem that connects the extended Riemann hypothesis to the least non-residues that the primality testing algorithm relies upon.

Theorem 4. (Ankeny [1], Montgomery [7]) (Assuming ERH) If χ is a non-principal Dirichlet character mod q, and $L(s,\chi)$ has no zeros $\rho = \beta + i\gamma$ with $\beta > 1/2$, then there exists an integer n with $\chi(n) \neq 0$ and $\chi(n) \neq 1$ such that $1 < n = O(\log(q)^2)$.

Proof. Choose some integer $N > 1$, we will try to bound N as the least non-residue mod q. We begin with the formula from lemma 2:

$$\sum_{n\le N}(1-\frac{n}{N})\Lambda(n)\chi(n) = -\sum_{\rho}\frac{N^{\rho}}{\rho(\rho+1)} + O(\log qN). \tag{iii}$$

Then, restricting the sum on the left to only primes will alter the sum by at most $O(\sqrt{N})$, based on the well known formula that $\sum_{n\le N}\Lambda(n) = \sum_{m=1}^{\log N}\sum_{p\le x^{1/m}}\log(p) = \sum_{p\le N}\log(p) + O(\sqrt{N})$, since $\sum_{p\le N}\log(p) = O(N)$ by the prime number theorem. Moreover, letting χ not be primitive will alter the sum by at most $O(\log q)$, a full explanation of this fact exceeds the scope of this paper, but to see the details of the change to a primitive character see sections 5 and 14 of Davenport [5]. Thus we have that:

$$\sum_{p\le N}(1-\frac{p}{N})\log(p)\chi(p) = -\sum_{\rho}\frac{N^{\rho}}{\rho(\rho+1)} + O(\sqrt{N}) + O(\log q). \tag{iv}$$

Now we note that for $t \geq 2$ we can bound the number of zeros in the critical strip by lemma 3,

$$N(0, t+2, \chi) - N(0, t, \chi) = O(\log qt). \tag{v}$$

Now, under our assumption of the ERH, we have that the real part $\beta \leq 1/2$ for each zero ρ, thus we have that

$$\sum_{\rho} \left| \frac{N^{\rho}}{\rho(\rho+1)} \right| \leq \sum_{t=2}^{\infty} O(\log qt) \left| \frac{\sqrt{N} N^{it}}{(\beta + it)(\beta + 1 + it)} \right| \tag{vi}$$

$$\leq O(\sqrt{N}) \sum_{t=2}^{\infty} \left| \frac{\log q + \log t}{-t^2} \right| = O(\sqrt{N} \log(q)). \tag{vii}$$

Combining this result with (iv) gives us that

$$\sum_{p \leq N} (1 - \frac{p}{N}) \log(p) \chi(p) = O(\sqrt{N} \log(q)). \tag{viii}$$

Now, we move on to proving the desired result. Assume that $\chi(p) = 1$ or 0 for all primes $p \leq N$ (i.e. that N is less than the least non-residue). Then, by a corollary to the prime number theorem (cited as lemma 1C in [1]) we have that

$$\sum_{p \leq N} \log(p) = N + o(N).$$

Then we put in $\chi(p)$ and $1 - \frac{p}{N}$, and as above this will add a factor of $O(\log q)$ for the Dirichlet character and get that

$$\sum_{p \leq N} (1 - \frac{p}{N}) \log(p) \chi(p) \geq \sum_{p \leq N} \log(p) \chi(p) = N + o(N) + O(\log q). \tag{ix}$$

Thus, we combine (viii) and (ix) to get our result that $N = O(\sqrt{N} \log(q))$ so that $N = O(\log(q)^2)$ as desired. □

Note the above Theorem 3 follows as a corollary of this theorem since we can let χ be the q-th power residue symbol mod p, which is a non-principal Dirichlet character. Then, finding such an n with $\chi(n) \neq 0$ and $\chi(n) \neq 1$ is equivalent to finding a qth non-residue mod p. So the Theorem implies that $N(p, q) = O(\log(p)^2)$.

3 Proof Of Main Theorem

Now, we return to Miller to finish off the proof of the main theorem in his paper and this one [6]. This will first require a few lemmas. Above in Section 1, we considered the case where $\lambda'(n) \nmid n-1$, now we must consider $\lambda'(n) \mid n-1$.

Motivation for Lemmas. Suppose that n is composite of the form $n = pq$ for p, q odd primes. Then, if we can find r such that

$$r \equiv 1 \ (q) \quad \text{and} \quad r \equiv -1 \ (p)$$

Then, we have that $q \mid r-1$ and that $r \not\equiv 1 \ (n)$ thus, we would have that $q = (r-1, n)$. Finding such an r would let us know that n is composite. Moreover, this strategy will let us identify composite numbers even when $a^{n-1} \equiv 1 \ (n)$ for all a in the unit group mod n.

Also note that when $\lambda'(n) \mid n-1$, if $q_1, \ldots, q_m$ are the distinct prime divisors of n, then $\#_2(\lambda'(n)) = \max\{\#_2(q_1 - 1), \ldots, \#_2(q_m - 1)\}$. Thus, there must exist some $1 \leq i \leq m$ such that $\#_2(\lambda'(n)) = \#_2(q_i - 1)$. This provokes the following definition.

Definition 7. Let $q_1, \ldots, q_m$ be the distinct prime divisors of n. We say that n is *type A* if for some $1 \leq j \leq m$ we have $\#_2(\lambda'(n)) > \#_2(q_j - 1)$. Alternatively, we say that n is *type B* if for all $1 \leq j \leq m$ we have $\#_2(\lambda'(n)) = \#_2(q_j - 1)$.

Lemma 4 A. Let n be a composite number of type A. Let $p, q \mid n$ and $\#_2(\lambda'(n)) = \#_2(p-1) > \#_2(q-1)$. Assume further that $0 < a < n$ with $(a, n) = 1$ and $(a/p) = -1$ where (a/p) is the Jacobi symbol, then $a^{\lambda'(n)/2} \equiv 1(q)$ and $a^{\lambda'(n)/2} \equiv -1(p)$.

Proof. Since $q - 1 \mid \lambda'(n)$ and $\#_2(q-1) < \#_2(\lambda'(n))$ we have that $q - 1 \mid \frac{\lambda'(n)}{2}$. So, by Fermat:

$$a^{\lambda'(n)/2} \equiv 1 \ (q).$$

Moreover, $(a^{\lambda'(n)/2})^2 \equiv 1 \ (p)$ since $p - 1 \mid \lambda'(n)$ by Fermat. So we have that $a^{\lambda'(n)/2} \equiv \pm 1 \ (p)$. Suppose that $a^{\lambda'(n)/2} \equiv 1 \ (p)$ and b is a primitive root mod p, then $p - 1 \mid (\text{ind}_{b,p} a)\lambda'(n)/2$ which means that $\text{ind}_{b,p} a$ is even since $\#_2(\lambda'(n)) = \#_2(p-1)$. However, $(a/p) = -1$ so that $\text{ind}_{b,p} a$ is odd, otherwise a would be the quadratic residue of $b^{\text{ind}_{b,p} a/2}$. Thus, we must have

$$a^{\lambda'(n)/2} \equiv -1 \ (p).$$

This gives us the desired result.

□

Lemma 4 B. Let n be a composite number of type B. Let $p, q \mid n$ and $\#_2(\lambda'(n)) = \#_2(p-1) = \#_2(q-1)$. Assume further that $0 < a < n$ with $(a, n) = 1$ and $(a/pq) = -1$. Then $a^{\lambda'(n)/2} \equiv 1(q)$ and $a^{\lambda'(n)/2} \equiv -1(p)$.

Proof. Again assume that (a, n) is trivial and thus equal to 1. We know by properties of the Jacobi symbol that $(a/pq) = (a/p)(a/q) = -1$. Without loss of generality, assume that $(a/p) = -1$ and $(a/q) = 1$. Using the same indexing argument as above, we get that $a^{\lambda'(n)/2} \equiv -1\ (p)$ and similarly that $a^{\lambda'(n)/2} \equiv 1\ (q)$. □

Lemma 5. If $p \mid n$ is an odd prime and $\lambda'(n) \mid m$ and $k = \#_2(\frac{m}{\lambda'(n)}) + 1$, then $a^{\frac{\lambda'(n)}{2}} \equiv a^{\frac{m}{2^k}}\ (p)$.

Proof. Since $a^{\lambda'(n)} \equiv 1\ (p)$ is follows that $a^{\lambda'(n)/2} \equiv \pm 1\ (p)$. This gives us two cases:

First, if $a^{\lambda'(n)/2} \equiv 1\ (p)$, then $a^{m/2^k} \equiv 1\ (p)$ since by our choice of k we have $2^k \mid \frac{2m}{\lambda'(n)}$, so that $\frac{\lambda'(n)}{2} \mid \frac{m}{2^k}$.

Second, if $a^{\lambda'(n)/2} \equiv -1\ (p)$, then expanding and substituting in this assumption:

$$a^{\frac{m}{2^k}} \equiv \left(a^{\frac{\lambda'(n)}{2}}\right)^{\frac{m}{\lambda'(n)2^{k-1}}} \equiv (-1)^{\frac{m}{\lambda'(n)2^{k-1}}}\ (p).$$

And, since $\frac{m}{\lambda'(n)2^{k-1}}$ is odd by our choice of k, we must have that $a^{m/2^k} \equiv -1\ (p)$. □

Definition 8. We define the function $N(pq)$ for $p \neq q$ prime to be the minimal a such that $(a/pq) \neq 1$. As above for $N(p, q)$, we have that $N(pq)$ is prime.

Theorem 5. (Assuming ERH) $N(pq) = O(\log(p)^2)$.

This is a slight generalization of Ankeny's result that is not actually proved in his paper. But, it follows from the more general theorem in section 2 with the use of the Jacobi symbol rather than the Legendre symbol as the Dirichlet character.

Choice of f. By the theorems of Ankeny, we can choose $c \geq 1$ such that $N(p, q) \leq c \log(p)^2$ and $N(pq) \leq c \log(pq)^2$. So, we let $f = c \log(n)^2$, and prove that A_f works and runs in the desired amount of time.

Proof of Correctness of A_f. If n is prime, then A_f will return "prime". This is because any prime n will have no divisors a and will have $a^{n-1} \equiv 1\ (n)$ by Fermat's Little Theorem. Moreover, let $A = \#_2(n-1)$ and $n - 1 = Q2^S$, then if $a^Q \not\equiv 1\ (n)$ we must have $a^{Q2^j} \equiv -1\ (n)$ for some $j < S$ and 1 for all greater j since $a^{Q2^S} \equiv 1\ (n)$. Thus, our algorithm always runs until it returns "prime" when n is prime.

So, we need only show that A_f recognizes each composite n as such. Let n be composite, then it falls into one of three cases: (1) n is a prime power, (2) $\lambda'(n) \not| n-1$, (3) $\lambda'(n) \mid n - 1$ and n is not a prime power. We consider each of these below:

(1) If n is a prime power, then step 1 of the algorithm will find this and correctly return that n is composite.

(2) If $\lambda'(n) \not| \, n-1$, then by Lemma 1 we have p, q such that if $a = N(p,q)$ (which is prime by Definition 6) then $a^{n-1} \neq 1$ (n). So, step 2(d) will return "composite" correctly as long as $N(p,q) \leq f(n)$, which it must be by our choice of $f(n)$ above.

(3) A) Suppose that n is of type A so that $p, q \mid n$ and $\#_2(\lambda'(n)) = \#_2(p-1) > \#_2(q-1)$. Then let $a = N(p,2)$ prime by Definition 6. We want to show that either step 2(b), 2(d), or 2(g) outputs composite for a. So, assume that $a \not| \, n$ and that $a^{n-1} \equiv 1$ (n). We will show that the algorithm reaches 2(g) and thus correctly outputs "composite". This means we need to show that the criteria in 2(e) and 2(f) are not fulfilled. First, note that since $a = N(p,2)$ we have that a is not a quadratic residue mod p. Since p is also odd, we have that if $a^k \equiv 1$ (p) then $2|k$ since the order of a is not divisible by 2, but must divide $p-1$ which is even. Then, since $p|n$, for $a^Q \equiv 1$ (n) we must have $a^Q \equiv 1$ (p), but this is impossible since any such Q must be even, but Q is odd. Thus, $a^Q \not\equiv 1$ (p) and we will move on to step 2(f). Now, we have that since $\lambda'(n) \mid n-1$, letting $m = n-1$ in Lemma 3 gives us the existence of a k such that by Lemma 4A, $a^{Q2^k} \equiv 1$ (q) and $a^{Q2^k} \equiv -1$ (p). Now, by means of contradiction, assume that $a^{Q2^J} \equiv -1$ (n). Then we have that $a^{Q2^J} \equiv -1$ (p) and (q). Then $a^{Q2^k} \equiv a^{Q2^J} \equiv -1$ (p) implies that $k = J$ since taking powers of 2 will yield 1. But, $a^{Q2^k} \equiv 1$ (q) while $a^{Q2^J} \equiv -1$ (q) implies that $k > J$, a contradiction. Thus, $a^{Q2^J} \not\equiv -1$ (n) and we must reach step 2(g) as desired.

(B) Suppose n is of type B. The argument proceeds in the same manner except that $a = N(pq)$ and we apply Lemma 4B instead of 4A (which had the same outcome as 4A).

Proof of Run Time of A_f. We will go through each step of the algorithm and calculate the run time. Let $M(n)$ be the run time of multiplication of numbers less than n which is $O(\log(n) \log\log(n) \log\log\log(n))$ using the Schonhage-Strassen algorithm [9].

1. For each b it will take $O(\log(n))$ steps of binary search for the appropriate a. For each of these steps, to calculate a^b will take at most $O(\log(n)M(n))$ steps to perform $b = O(\log(n))$ multiplications of numbers less than n. Then, since there are at most $O(\log(n))$ values of b we have that this whole step takes $O(\log(n)^2 M(n))$ which will be dominated by step 2.

2. To compute the p_i less than $f(n)$, we can use the sieve of Atkin which runs in time linear in $f(n) = O(\log(n)^2)$, which will be dominated [2]. Note that by the prime number theorem there are $O(\frac{f(n)}{\log(f(n))})$ prime numbers less than $f(n) = O(\log(n)^2)$. So, we have that $m = O(\frac{\log(n)^2}{\log\log(n)})$. Computing S, Q only needs to be done once and takes $O(\log(n)M(n))$ which will be dominated.

(a) Takes constant time $O(1)$.

(b) Takes $O(\log(n)^2)$ to do long division naively.

(c) Modular exponentiation by squaring will take $O(\log(n)M(n))$ to compute a^Q (n) since $a, Q, n \leq n$. Then again we have that modular exponentiation by squaring will take $O(\log(n)M(n))$ to calculate all of the a^{Q2^j} (n) since $2^j < n$. Thus, we have that this whole step takes $O(\log(n)M(n))$, which will be the dominant cost.

(d) Takes constant time $O(1)$.

(e) Takes constant time $O(1)$.

(f) Takes $O(S) = O(\log(n))$ to find the max over S elements.

(g) Takes constant time $O(1)$.

Thus, step 2 loops at most $O(\frac{\log(n)^2}{\log\log(n)})$ times, and each time performs at most $O(\log(n)M(n))$ computations, giving us a total run time of:

$$O(\frac{\log(n)^2}{\log\log(n)}\log(n)\log(n)\log\log(n)\log\log\log(n)) = O(\log(n)^4\log\log\log(n)).$$

So, we have shown that A_f for $f = O(\log(n)^2)$ we have an algorithm that tests primality in $O(\log(n)^4\log\log\log(n))$, proving Theorem 1. A paper by Bach in 1990 provides an even stronger result by showing that we need only take $f(n) = 2\log(n)^2$ [3]. The actual necessary $f(n)$ is not known, but the upper bound on the least quadratic non-residue cannot be improved beyond $\Omega(\log(n)\log\log(n))$ by a result from Montgomery that relies on the ERH [7]. But, these polynomial run times are fast enough and with low enough constants that this algorithm does make sense in practice.

References

[1] N. C. Ankeny, *The least quadratic non residue*, Annals of Mathematics **55** (1952), 65–72.

[2] D.J. Bernstein A.O.L. Atkin, *Prime sieves using binary quadratic forms*, Math. Comp. **73** (2004), 1023–1030.

[3] E. Bach, *Explicit bounds for primality testing and related problems*, Math. Comp. **55** (1990), 355–380.

[4] R.D. Carmichael, *On composite numbers p which satisfy the fermat congruence*, The American Mathematical Monthly **19** (1912), 22–27.

[5] H. Davenport, *Multiplicative number theory*, Graduate Texts in Mathematics (1980).

[6] G. L. Miller, *Riemann's hypothesis and tests for primality*, Journal of Computer and System Sciences **13** (1976), 300–317.

[7] H.L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Math. (1971).

[8] M. O. Rabin, *Probabilistic algorithm for testing primality*, Journal of Number Theory **12** (1980), 128–138.

[9] A. Shonhage and V. Strassen, *Schnelle multiplikation grosser zahlen*, Computing **7** (1971), 281–292.