# PRIMALITY TESTING

CHARLES DENIS XAVIER

## 1. INTRODUCTION

Primality Testing is a fundamental problem of Number Theory, for which despite centuries of study no provably efficient algorithms have been devised. Further it has several applications especially in Cryptography. In this treatise we shall survey this beautiful and interesting area.

## 2. PRIME NUMBERS

In this section we shall enumerate the important properties of prime numbers.

**Definition 2.1.** Let $R$ be a commutative ring with identity. Then

(1) $p \in R, p \neq 0$ is said to be *irreducible* if $(p = ab) \Rightarrow (a \text{ is a unit}) \vee (b \text{ is a unit})$.
(2) $p \in R$ is said to be *prime* if $p \backslash ab \Rightarrow (p \backslash a) \vee (p \backslash b)$.

The definition of a prime element as above, does not immediately suggest an algorithm for detecting whether an element is a prime element in the ring. In fact for arbitrary rings, the question whether an element is prime is difficult to decide. However the following theorem gives us some hope in the case of the ring of integers.

**Theorem 2.2.** *If $R$ is a* Euclidean domain *then $p \in R$ is prime iff it is irreducible.*

We shall for the most part be interested in the prime elements in the ring $(\mathbb{Z}, +, \times)$, which is a euclidean domain with respect to the function $\varphi(n) = |n|$. Further the ordering of the numbers by the same function suggests we can exhaustively divide $p$ by all numbers below it and check that the only divisor is 1 (which is a unit). This will prove that $p$ is irreducible which by the above theorem also proves that it is a prime element. In fact *Fibonacci* observed that if $n$ is composite then one of its divisors is less than $\lfloor \sqrt{n} \rfloor$. The problem with this naïve algorithm is the fact that though the input requires only $\lfloor \lg n \rfloor + 1$ bits to represent the above algorithm requires $O(2^{\lg n})$ divisions and is hence an exponential time algorithm. Before we search for criteria for primality which can be efficiently verified, we digress first into the properties of prime numbers.

### 2.1. Distribution of Primes.

Euclid proved that the set of primes is infinite. A natural question to ask then is "How dense are the primes?" Or more precisely, if $\pi(n) = |\{p \mid p \leq n, p \text{ prime } \}|$, how does $\pi(n)$ vary with $n$? Estimates to this function were given independently by *Gauss* and *Legendre*. Both the estimates were asymptotically the same and they were :

$$\pi(n) \sim \int_2^n \frac{1}{\ln x} dx \sim \frac{n}{\ln n}.$$

The integral in the above expression occurs very often and is called the *Log-integral function* $\text{Li}\,(n) = \int_2^n \frac{1}{\ln x} dx$. The first proof that the function behaved asymptotically as above was by *Hadamard* and *de la valée Poussin* nearly a century after the original conjecture. It is interesting to give a heuristic justification for this bound. The argument is from [Sch97]. Let $W(n)$ denote the probability that the number $n$ is prime[1]. Now assuming that divisibility of a number by two other numbers is independent, we have :

$$W(n) \sim \prod_{2 \leq p < n, p \text{ prime}} \left(1 - \frac{1}{p}\right)$$

Since 1 out of every $p$ numbers is divisible by a prime $p$.

[1]Note that a number is either prime or not so this heuristic argument makes no sense if interpreted formally, it is given so that we have an intuitive understanding of the distribution of primes.

Taking logarithms on both sides we have

$$\ln W(n) \sim \sum_{2\leq p<n} \ln(1-\frac{1}{p})$$

But $\ln(1+x) \approx x$ for small values of $x$. So we have

$$\ln W(n) \sim \sum_{2\leq p<n} -\frac{1}{p} = -\sum_{2\leq p<n} \frac{1}{p}.$$

We can extend this sum over all integers as follows :

$$\ln W(n) \sim -\sum_{2\leq i<n} \frac{1}{i} W(i)$$

Now sums can be approximated very well using integrals so we have :

$$\ln W(n) \sim -\int_2^n \frac{1}{i} W(i)$$

Let $A(x) = \frac{1}{W(x)}$ This yields a differential equation :

$$\frac{1}{A(x)} A'(x) = \frac{1}{xA(x)}$$

or

$$A'(x) = \frac{dx}{x}$$
$$A(x) = \ln x.$$

So $W(n) \sim \frac{1}{\ln n}$. Thus we have

$$\pi(n) \sim \sum_{2\leq i<n} 1 \times W(i) \sim \sum_{2\leq i<n} \frac{1}{\ln i} \sim \frac{n}{\ln n}.$$

The sharpest result for the behaviour of $\pi(n)$ is the following.

**Theorem 2.3.** *There is a $c > 0$ such that*

$$\pi(n) = \mathrm{Li}\,(n) + O(ne^{-c\lambda(n)}),$$

*where*

$$\lambda(n) = (\ln n)^{\frac{3}{5}} (\ln \ln n)^{-\frac{1}{5}}.$$

It is instructive to look at the proof of the infinitude of primes by Euler, since it resulted in the development of new machinery in number theory.
Euler considered the following function :

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s}.$$

This is the famous *Zeta* function. We can rewrite the above function in the following interesting form,

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{\left(1-\frac{1}{p^s}\right)}.$$

The expressions are correct using the *Fundamental Theorem of Arithmetic* and from the fact that $\sum_{i\geq 0} r^i = \frac{1}{(1-r)}$ for $|r| < 1$. The above can be interpreted as the quantitative equivalent of the fundamental theorem.
Now taking logarithms on both sides on obtains :

$$\ln\zeta(s) = \sum_{p \text{ a prime}} -\ln\left(1-\frac{1}{p^s}\right) \approx \sum_{p \text{ a prime}} \frac{1}{p^s}.$$

From the other expression $\lim_{s\to 1}\ln\zeta(s)$ diverges so we have $\sum_p \frac{1}{p}$ diverges. Thus the sum over the reciprocals of the primes diverges so there must be an infinite number of them. In fact we can use the prime number theorem to prove that the $n$-th prime number is $\approx n\log n$. This allows us to obtain an estimate of the growth rate of this sum as follows : For the moment we shall treat the $\log n$ as though it is $\lg n$, this will simplify our argument and allow us to ignore the constant $\frac{1}{\lg e}$.

Let $S_k = \sum_{p_i, i\le k} \frac{1}{p_i}$, where $p_i$ is the $i$-th prime

$$\begin{aligned}
S_{2k} &= S_k + \sum_{p_i, k<i\le 2k} \frac{1}{p_i} \\
&\approx S_k + \sum_{k<i\le 2k} \frac{1}{i\log i} \\
&\ge S_k + \sum_{k<i\le 2k} \frac{1}{i\log 2k} \\
&= S_k + \frac{1}{\log 2k}\left(\sum_{k<i\le 2k}\frac{1}{i}\right) \\
&= S_k + \frac{1}{\log 2k}(H_{2k} - H_k) \\
&\approx S_k + \frac{1}{\log 2k}(\log 2k + \gamma - \log k - \gamma) \\
&= S_k + \frac{1}{\log 2k} \\
&= S_k + \frac{1}{\log k + 1}.
\end{aligned}$$

Thus by induction we have $S_{2^k} \ge H_k \approx \log k$ or $S_n \ge \log\log n$. In fact it has been shown that $S_n$ is asymptotically this quantity.

*Lejeune Dirichlet* used essentially the same idea as Euler, with a new "selecting" function to drop off terms which are $p \not\equiv a \mod n, a \perp n$ to prove the following theorem[2].

**Theorem 2.4** (Dirichlet,1826)**.** *If $a \perp n$ there are infinitely many primes $p \equiv a \mod n$.*

Here are some interesting facts about the most important number theoretic function, the Riemann Zeta Function (so called because of a deep paper by *Bernhard Riemann* concerning this function).
The series $\zeta(s) = \sum_{1\le n} \frac{1}{n^s}$ is absolutely and uniformly convergent in the domain $\Re(s) \ge 1+\delta$, for every $\delta > 0$. It represents an analytic function in the half-plane $\Re(s) > 0$.
Recall that the $\Gamma$ function is defined by :

$$\Gamma(s) = \int_0^\infty \frac{e^{-y}y^s}{y} dy.$$

The *Completed Zeta function* is defined as

$$Z(s) = \pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s).$$

[2]Clearly this is an understatement of the proof of this theorem, which is an exceedingly non-trivial argument.

The nice property of the completed zeta function is that $Z(s) = Z(1-s)$. So we have

$$\pi^{-\frac{s}{2}}\Gamma(\frac{s}{2})\zeta(s) = \pi^{\frac{(s-1)}{2}}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s)$$

$$\text{so } \zeta(s) = \pi^{\frac{(s-1)}{2}}\frac{\Gamma(\frac{1-s}{2})}{\Gamma(\frac{s}{2})}\zeta(1-s).$$

It is clear from the definition of the zeta function that it does not vanish for $\Re(s) > 1$. Now at $s = -2, -4, ..., \Gamma(\frac{s}{2})$ has a simple pole, since $1-s$ is positive $\zeta(1-s)$ is not zero and finite. So $\zeta(s)$ vanishes at every negative even integer. These zeros of $\zeta(s)$ are called the trivial zeros of the zeta function $(s = -2, -4, -6, ...)$. All other zeros lie on the critical strip

$$0 \leq \Re(s) \leq 1.$$

This brings us to the most famous unsolved problem of mathematics.

**Riemann Hypothesis :** The nontrivial zeros of $\zeta(s)$ lie on the line $\Re(s) = \frac{1}{2}$.

The importance of the Riemann hypothesis to prime numbers comes from the following connection.

$$\pi(x) = R(x) - \sum_{\rho} R(x^{\rho})$$

where $\rho$ varies over all zeros of $\zeta(s)$ and

$$R(x) = 1 + \sum_{1 \leq n} \frac{1}{n\zeta(n+1)}\frac{(\ln x)^n}{n!}.$$

The surprising consequences are the following :

**Theorem 2.5.** *If the Riemann Hypothesis is true then*

$$\pi(x) = \text{Li }(x) + O(\sqrt{x}\ln x).$$

$$\psi(x) = \sum_{p^i \leq x} \ln p = x + O(\sqrt{x}(\ln x)^2)$$

$$p_n = \text{Li }^{-1}(n) + O(\sqrt{n}(\ln n)^{\frac{5}{2}}).$$

$$\prod_{p \leq x}\left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\ln x} + O(x^{-\frac{1}{2}}).$$

($p_n$ is the $n$ prime number).

The following is known without the assumption of the Riemann Hypothesis.

**Theorem 2.6.**

$$\pi(x) = \text{Li }(x) + O(xe^{-c\lambda(x)}) \sim \frac{x}{\ln x}.$$

$$\psi(x) = x + O(xe^{c\lambda(x)}) \sim x.$$

$$p_n = \text{Li }^{-1}(n) + O(ne^{-c\lambda(n)}) \sim n\ln n.$$

$$\prod_{p \leq x}\left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\ln x} + O(e^{-c\lambda(x)}) \sim \frac{e^{-\gamma}}{\ln x}.$$

Where $\gamma \approx 0.57721566490153286060651 2\cdots$ is the *Euler-Mascheroni* constant.

## 3. Important Developments in Primality Testing

Here is an approximate timeline of some of the important developments in the field of primality testing[3].

1935 *Lucas-Lehmer* Mersenne Primes $\in$ P.

1975 *Pratt V. R.* **PRIMES** $\in$ NP $\cap$ coNP.

[3]No attempt at completeness has been made.

1976 *Gary L. Miller* Under ERH **PRIMES** $\in$ P.
1977 *Solovay-Strassen* **PRIMES** $\in$ coRP.
1980 *Rabin M. O.* Improved coRP algorithm for Primes.
1983 *Adleman, Pomerance, Rumely* **PRIMES** $\in$ DTIME$[n^{O(\lg\lg n)}]$.
1986 *Goldwasser, Kilian* There is an infinite set $S \subset$ **PRIMES** such that $S \in$ ZPP. Also the set $S$ is a dense subset of the primes.
1992 *Fellows-Koblitz* If the factorization of $n-1$ is available then primality of $n$ can be checked in deterministic polynomial time.
1992 *Adleman, Huang* **PRIMES** $\in$ ZPP.

The problem of whether there is an infinite subset of primes which can be checked in polynomial time was open for a long time, until Pintz, Steiger and Szemeredi settled the problem in 1989 [PSS89]. However the subset they constructed is not very dense. It is curious that the existing algorithms for primality sacrifice exactly one of efficiency, generality, or certainity. We shall look at one algorithm from each category in what follows.

## 4. Fermat's Theorem

Many of the primality algorithms are based on partial converses to the following theorem.

**Theorem 4.1.** *If $a \perp p$ where $p$ is prime then $a^{p-1} \equiv 1 \mod p$.*

**Proof :** Consider the set $\Phi(p) = \{1, \cdots, p-1\}$, each element $s \in \Phi(p)$ is relatively prime to $p$. Now if $a \perp p$ $(\gcd(a,p) = 1)$, then we have

$$\forall b, c \in \Phi(p) \ : \ ba \equiv ca \Rightarrow b \equiv c \Rightarrow b = c.$$

So $a.\Phi(p) = \{a.1, a.2, \cdots, a.(p-1)\} = \Phi(p)$. So we have

$$\begin{aligned} a.1.a.2\cdots a.(p-1) &\equiv 1.2.\cdots(p-1) \mod p \\ a^{p-1}(p-1)! &\equiv (p-1)! \mod p \\ a^{p-1} &\equiv 1 \mod p \text{ since } k \perp p \text{ for } 1 \le k < p. \end{aligned}$$

□

This yields a natural algorithm for testing whether $n$ is a prime, pick $m \in \{1, \cdots, (n-1)\}$, check whether $m^{n-1} \equiv 1 \mod n$, and if so declare $n$ prime and otherwise declare $n$ to be composite. The problem is that the direct converse of the above theorem is false.

**Example 4.2.** $4 \perp 15$, $4^{14} \equiv 1$, but 15 is not prime.

There are numbers $n$ which are composite for which $\forall a \in (\frac{\mathbb{Z}}{n\mathbb{Z}})^{\times} : a^{n-1} \equiv 1 \mod n$. Such pathological numbers are called *Carmichael Numbers*. The first example of which is $561 = 3 \times 11 \times 17$ and the next being $1729 = 7 \times 13 \times 19$. Such cases occur infinitely often on which our poor algorithm has no chance of succeeding, more precisely.
Let $C(x) = |\{n \ : \ n \le x, n \text{ is a Carmichael number}\}|$.

**Theorem 4.3** (Alford, Pomerance, Granville)**.** *For all large enough $x$, $C(x) > x^{\frac{2}{7}}$.*

We will require the following properties of Carmichael numbers later.

**Theorem 4.4.** *If $n$ is a Carmichael number, then $n$ is odd, squarefree and has at least 3 distinct prime factors.*

## 5. Lucas-Lehmer Test

Numbers of special forms can be proved to be prime or composite in deterministic polynomial time. We give two examples of this phenomenon, however we do not know that there are infinitely many primes of this form.

**Definition 5.1.** If $q$ is a prime of the form $2^p - 1$, where $p$ is prime, then $q$ is called a *Mersenne Prime*.

An easy observation is that if $n = 2^k - 1$ is prime then $k$ is prime. Since if $n = 2^{ab} - 1$ then $n = (2^a - 1)(1 + 2^a + \cdots + 2^{a(b-1)})$. So this immediately reduces our search space for Mersenne Primes. However Mersenne primes are very rare and only 38 such primes are known (as of now - 1999). Interestingly each Mersenne prime yields an even perfect number and every even perfect number has to be a Mersenne prime times a power of two. The belief is that there are infinitely many such primes.

**Theorem 5.2.** *Let $p > 2$ be a prime, let $n = 2^p - 1$ and*

$$S_1 = 4,$$
$$S_{k+1} = S_k^2 - 2 \mod n, k \geq 1.$$

*Then n is prime iff $S_{p-1} \equiv 0 \mod n$.*

**Proof :** Consider the polynomial

$$f(x) = x^2 - 2^{\frac{(p+1)}{2}}x - 1.$$

Let $q$ be any prime dividing $n$, and let $\alpha, \beta \in \mathbb{F}_{q^2}$ be the roots of $f(x) = 0$ over $\mathbb{F}_q$. Note that if $f$ had no roots over $\mathbb{F}_q$, then $f$ is irreducible in this smaller field, so that we have $\mathbb{F}_{q^2} \cong \frac{\mathbb{F}_q}{\langle f \rangle}$, and if $f$ had a root over $\mathbb{F}_q$, then since $\mathbb{F}_{q^2}$ is an extension field of the other and $f$ has a root in the bigger field.

$$(x-\alpha)(x-\beta) = x^2 - (\alpha+\beta)x + \alpha\beta.$$

So we have $\alpha + \beta = 2^{\frac{p+1}{2}}, \alpha\beta = -1$.
Let $V(k) = \alpha^k + \beta^k$, and let $\overline{S_k} = S_k \mod q$.
*Claim :* $\overline{S_k} = V(2^k)$. Proof of this is by induction on $k$.
[*Basis*] For $k = 1$.

$$\begin{aligned} V(2) = \alpha^2 + \beta^2 &= (\alpha+\beta)^2 - 2\alpha\beta \\ &= (2^{\frac{(p+1)}{2}})^2 - 2(-1) \\ &= 2^{p+1} + 2 = 2(2^p - 1) + 4 \mod q = 4. \end{aligned}$$

and $S_1 = 4$, by definition.

[*Inductive Hypothesis*] $\overline{S_k} = V(2^k) = \alpha^{2^k} + \beta^{2^k}$.
[*Inductive Step*]

$$\begin{aligned} \overline{S_{k+1}} = (\alpha^{2^k} + \beta^{2^k})^2 - 2 &= \alpha^{2^{k+1}} + \beta^{2^{k+1}} + 2(\alpha\beta)^{2^k} - 2 \\ &= \alpha^{2^{k+1}} + \beta^{2^{k+1}} = V(2^{k+1}). \end{aligned}$$

This proves the claim.
Let $n = 2^p - 1$ be a prime.
We have to show that $S_{p-1} \equiv 0 \mod n$.

*Claim :* $f(x)$ is irreducible over $\mathbb{F}_n$.
We know that $f(x) = ax^2 + bx + c$ is irreducible iff $\Delta = b^2 - 4ac$ is not a quadratic residue. We have

$$\Delta = \left(2^{\frac{(p+1)}{2}}\right)^2 - 4(-1) = 2^{p+1} + 4 = 2(2^p - 1) + 6 \equiv 6 \mod n.$$

Now $\left(\frac{\Delta}{n}\right) = \left(\frac{2}{n}\right)\left(\frac{3}{n}\right)$.
Since

$$\begin{aligned} (2^{\frac{p+1}{2}})^2 &\equiv 2^{p+1} \\ &= 2n + 2 \\ &\equiv 2 \mod n. \end{aligned}$$

So $\left(\frac{2}{n}\right) = +1$, as 2 is a quadratic residue. Now $\left(\frac{3}{n}\right) = -\left(\frac{n}{3}\right)$, Since $2^p - 1 \equiv -1 \mod 4, 3 \equiv -1 \mod 4$.
Further $2 \equiv -1 \mod 3$, so $2^k \equiv (-1)^k$. Since $p$ is odd we have $2^p \equiv -1 \mod 3$ So $2^p - 1 \equiv 1 \mod 3$. And finally we have $\left(\frac{n}{3}\right) = \left(\frac{1}{3}\right) = 1$. Thus $\left(\frac{\Delta}{n}\right) = -1$. So $f(x)$ is irreducible.
Thus we have that $\alpha, \beta$ are algebraic over $\mathbb{F}_n$ and that $f(x)$ is their minimal polynomial. Now the norm of an element in $\mathbb{F}_{q^m}$ is $N(x) = x^{1+q+q^2+\cdots+q^{m-1}}$. Since $\alpha \in \mathbb{F}_{n^2}$, we have $m = 2$ so we have $N(\alpha) = \alpha^{n+1}$ and $N(\beta) = \beta^{n+1}$. But $N(\alpha)$ is the product of the conjugates of $\alpha$ so we have $N(\alpha) = N(\beta) = \alpha\beta = -1$.

Hence $\alpha^{n+1} = \beta^{n+1} = -1$, or $\alpha^{n+1} + \beta^{n+1} = \overline{S_p} = -2$. And $S_p = S_{p-1}^2 - 2$ or $S_{p-1}^2 \equiv 0 \mod n$, as $n$ is prime we have $S_{p-1} \equiv 0 \mod n$.
Now let us assume that $S_{p-1} \equiv 0 \mod n$ and $n$ is not prime. Since $n$ is composite, we have $\exists q : (q \setminus n) \wedge (q^2 \leq n)$.
Now $S_{p-1} \equiv 0 \mod n \Rightarrow S_{p-1} \equiv 0 \mod q$.
Let $\alpha, \beta$ be as before.
$\alpha^{2^{p-1}} + \beta^{2^{p-1}} = 0$ as $\overline{S_{p-1}} = V(2^{p-1})$.
Now

$$\alpha^{2^p} + (\alpha\beta)^{2^{p-1}} = 0$$

(by multiplying by $\alpha^{2^{p-1}}$). So we have $\alpha^{2^p} = -1$, and $\alpha^{2^{p+1}} = 1$.
Since the power is even we have that $\text{ord}(\alpha) = 2^{p+1}$ in $\mathbb{F}_{q^2}$. Also $\alpha \in \mathbb{F}_{q^2}$. So we have $\text{ord}(\alpha) \setminus q^2 - 1$ or $2^{p+1} \setminus q^2 - 1$. But $2^{p+1} > n$ and $q^2 \leq n$. This is a contradiction so $n$ is prime. □

Though the proof was rather non-trivial, the test that it yields can be efficiently implemented, and is essentially the test used by the *GIMPS* (Great Internet Mersenne Primes Search). The search has been very successful in isolating new "titanic" primes. The largest known prime has been isolated by this project and is $2^{6972593} - 1$.

## 6. Primes of the form $4p^2 + 1$

There are many algorithms which can test primality of specific types of primes. Here is an example of a test, which makes use of the fact that $n - 1$ is a simple number whose factorization is known. There is a more general algorithm by *Fellows* and *Koblitz* which given the factorization of $n - 1$ will prove in polynomial time the primality status of $n$.

**Theorem 6.1.** *If $n = 4p^2 + 1$ with $p$ a prime, then $n$ is prime iff $p \setminus \text{Ord}(p)$ in $\left(\frac{\mathbb{Z}}{n\mathbb{Z}}\right)^\times$.*

**Proof :** It is clear the $p \perp n$ (relatively prime) so $p \in \left(\frac{\mathbb{Z}}{n\mathbb{Z}}\right)^\times$.

**[$n$ is prime]** Then the multiplicative group has order $n - 1 = 4p^2$ and since $p$ belongs to this group we have that $\text{Ord}(p) \setminus 4p^2$, so if we show that $\text{Ord}(p)$ is not one of 1,2, or 4 then from the above we have $p \setminus \text{Ord}(p)$. Clearly $p \neq 1$ and $p^2 < n$ so its order is not 1 or 2. Now if $p^4 \equiv 1 \mod n$ then $n \setminus (p^4 - 1)$ so $n$ divides $4p^4 - 4$ also $n$ divides $4p^4 + p^2$ by the definition of $n$, so $n$ should divide $p^2 + 4$ which is impossible since this is smaller than $n$. So $p \setminus \text{Ord}(p)$.

**[$p \setminus \text{Ord}(p)$]** Suppose $n$ is not a prime then we have the following facts :

$$p \setminus \varphi(n) \text{ by Euler-Fermat and the fact that } p \setminus \text{Ord}(p)$$
$$\text{also } p \setminus (n-1).$$

The idea is that the order of $p$ is now big enough to make this impossible if $n$ were composite.
Let $n = p_1^{e_1} p_2^{e_2} \cdots p_j^{e_j}$ be the prime factorization of $n$ then we have $\varphi(n) = p_1^{e_1 - 1} \cdots p_j^{e_j - 1}(p_1 - 1) \cdots (p_j - 1)$. Now since $p \setminus \varphi(n)$ and $p \perp n$ so $p$ must divide one of the $p_i - 1$ factors. Let us assume that it actually divides $p_j - 1$ so that $p_j = kp + 1$ where $k$ is a positive constant.

Now $n - 1 = mp_j - 1 = m(kp + 1) - 1 = mkp + (m - 1)$ but $p$ divides $n - 1$ (by definition) so $p$ divides $m - 1$, so $m = lp + 1$ for some positive integer $l$ since $n$ is composite. Thus $4p^2 + 1 = n = mp_j = (lp + 1)(kp + 1) + 1 = klp^2 + (k + l)p + 1$, this gives us that $p = \frac{(k+l)}{(4-kl)}$. Since the denominator needs to be positive the only acceptable value of $k$ is 2 since $p_j$ is a prime bigger than $p$ and so it should be odd. But then there are no values of $l$ for which the equality holds. Thus $n$ is prime. □

## 7. Solovay-Strassen Algorithm

The Solovay-Strassen Algorithm is a simple and direct algorithm for checking whether a number is prime. It uses $O(\lg n)$ random bits at exactly one step. The *Jacobi symbol* can be evalutated in a style similar to the Euclidean GCD algorithm, and can be implemented to run in $O((\lg n)^3)$ steps.

**Algorithm 7.1** Solovay-Strassen Algorithm for Primality

**Input :** $n$ an odd integer.
**Steps :**
Pick $a \in_R \{1, \cdots, n-1\}$.
if $(a \perp n)$
{
$\varepsilon \leftarrow a^{\frac{(n-1)}{2}} \mod n$ with $-1 \leq \varepsilon \leq n-2$.
$\delta \leftarrow \left(\frac{a}{n}\right)$
if $(\varepsilon \neq \delta)$
Declare Not Prime
else
Declare Prime
}
else
Declare Not Prime

**Theorem 7.1.** *If n is prime, the algorithm outputs "Prime". If n is composite, then the algorithm outputs "Prime" with probability* $\leq \frac{1}{2}$*, where the probability is taken over the random bits of the algorithm.*

**Proof :** [$n$ is Prime] In this case $\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \mod n$. So the decision is correct.

[$n$ is Composite] Consider the set $G = \{a + \langle n \rangle \mid a \in \mathbb{Z}, a \perp n, a^{\frac{(n-1)}{2}} \equiv \left(\frac{a}{n}\right) \mod n\}$.
If $b = a^{-1} \mod n$ then $\left(\frac{a}{n}\right)\left(\frac{b}{n}\right) \equiv \left(\frac{ab}{n}\right) \equiv \left(\frac{1}{n}\right) \equiv 1 \mod n$. Thus $\left(\frac{b}{n}\right) \equiv \left(\frac{a}{n}\right)^{-1} \mod n$. And

$$(a^{-1})^{\frac{n-1}{2}} = (a^{\frac{n-1}{2}})^{-1} = \left(\frac{a}{n}\right)^{-1} = \left(\frac{a^{-1}}{n}\right).$$

It is also clear that the property is closed under modular multiplication. So $G$ is a subgroup of $\left(\frac{\mathbb{Z}}{n\mathbb{Z}}\right)^{\times}$. If we show that this is a proper subgroup in case $n$ is composite then we are done, since then by Lagrange's theorem

$$|G| \leq \frac{\left|\left(\frac{\mathbb{Z}}{n\mathbb{Z}}\right)^{\times}\right|}{2}$$

.
Let $a^{\frac{n-1}{2}} \equiv \left(\frac{a}{n}\right) \mod n$ for all $a \perp n$. Then clearly $n$ is a Carmichael number and so is squarefree and has more than two factors.
So let $n = rs$ with $r \perp s$.

$$\forall a \ : \ a \perp n \Rightarrow a^{\frac{(n-1)}{2}} \equiv \pm 1 \mod n.$$

*Claim :* In this case we actually have

$$\forall a \ : \ a \perp n \Rightarrow a^{\frac{(n-1)}{2}} \equiv 1 \mod n.$$

If

$$\exists a \ : \ (a \perp n) \ \wedge \ (a^{\frac{(n-1)}{2}} \equiv -1 \mod n),$$

then using the CRT we can find $b$ such that

$$b \equiv 1 \mod r, \text{ and } b \equiv a \mod s.$$

So we have $b^{\frac{n-1}{2}} \equiv 1 \mod r$ and $b^{\frac{n-1}{2}} \equiv a^{\frac{n-1}{2}} \equiv -1 \mod s$. Now $b^{\frac{n-1}{2}} \not\equiv \pm 1 \mod n$.Which is a contradiction.
Now let $n = p.r$ where $p$ is a prime and $p \perp r$ ($n$ is squarefree). Let $g$ be a quadratic non-residue $\mod p$. Using CRT find an $a$ such that $a \equiv g \mod p, a \equiv 1 \mod r$.
So

$$\left(\frac{a}{n}\right) = \left(\frac{a}{pr}\right) = \left(\frac{a}{p}\right)\left(\frac{a}{r}\right) = -1\left(\frac{1}{r}\right) \equiv -1 \mod n.$$

This is a contradiction. So $G$ is a proper subgroup. And so

$$|G| \leq \frac{\left|\left(\frac{\mathbb{Z}}{n\mathbb{Z}}\right)^{\times}\right|}{2} = \frac{\varphi(n)}{2} \leq \frac{n-1}{2}.$$

So the number of bad choices for $a$ in the algorithm is bounded above by $\frac{1}{2}$. □

## 8. The APR Algorithm

In 1983, *Adleman, Pomerance and Rumely* [APR83] developed a powerful test for primality using some pretty advanced mathematics, in specific from the theory of Cylotomic Fields. This test is currently the fastest *deterministic* test for primality which does not rely on any unproven hypothesis. In this section we shall describe a simplified version of the algorithm, which covers the main ideas used in the actual algorithm, and briefly touch upon the actual algorithm. Lenstra and Cohen later simplified the procedure by using a variation of the techniques used here, resulting in an algorithm which can be implemented relatively easily.
Here is the basic idea behind the approach.

**Definition 8.1.** Let $I$ be a finite set of primes. Define the set of *Euclidean Primes* with respect to I, as

$$\text{Euclid}(I) = \{q \mid q-1 \text{ is squarefree and } p \backslash (q-1) \Rightarrow p \in I\}.$$

Let $n$ be the number to be tested for primality.

(1) Find a set of primes $I_n$, such that

$$\prod_{q \in \text{Euclid}(I_n)} q > \sqrt{n}.$$

We will find that the running time of the algorithm will be proportional to the total length of primes in the set $I_n$ for the number $n$.

(2) If $n$ is composite then it has a divisor $r$ such that $0 < r \leq \sqrt{n}$.
We try to find $r$ as follows.
Find $r \mod q$ for each $q \in I_n$. Since $r \leq \sqrt{n}$, CRT can then be used to find $r$.
Clearly if we have $I_n$ we can verify that $q \neq r$, so we can assume $r \not\equiv 0 \mod q$.

(3) Now how do we find $r \mod q$. Since $r$ belongs to the multiplicative group of each of the primes in $I_n$, if we can find the order of this element in each of the groups then we can find $r \mod q$. So how do we find the order of this $r$ in $\mathbb{Z}/q\mathbb{Z}$?
Fix a primitive root for $q$ say $t_q$.
For $x \perp q$, let $\text{Ind}_q(x)$ be the least integer such that $x \equiv t_q^{\text{Ind}_q(x)} \mod q$ (just the discrete-log modulo $q$ for the generator $t_q$).
If $\text{Ind}_q(r)$ is known, then so is $r \mod q$.

(4) Since $q-1$ is squarefree (which is the order of the multiplicative subgroup) if we know $\text{Ind}_q(r) \mod p$ for each prime $p \backslash (q-1)$ , then once again using CRT we can determine $\text{Ind}_q(r)$.

(5) Clearly the difficult part is to determine $\text{Ind}_q(r) \mod p$ for each $p \backslash (q-1)$.
The winning idea is the realization that the Legendre symbol $\left(\frac{x}{q}\right)$ is a special case of this problem and is that of finding $\text{Ind}_q(x) \mod 2$ and so our approach should generalize the reciprocity laws observed in the ordinary quotient fields of integers. A nice generalization to higher power reciprocity laws can be found in Cyclotomic fields. Which is what we shall study first, following which we shall give a more detailed description of the algorithm.

### 8.1. Results from Cyclotomic Field Theory.

Since very little can be done in the improvement of the presentation of the following, we include this section almost at verbatim from the paper.

**Definition 8.2.** Let $\zeta_p = e^{\frac{2\pi i}{p}}$, be the $p$th root of unity, whose irreducible polynomial over $\mathbb{Q}$ is the $p$th cyclotomic polynomial

$$\Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + 1.$$

We shall be working in the extension field $\mathbb{Q}(\zeta_p)$.

By definition the extension is a finite dimensional extension over the rationals so is a number field. It can be shown that the ring of the algebraic integers in this field is $\mathbb{Z}[\zeta_p]$.

A rational prime $q$ factorizes into prime ideals in $\mathbb{Z}[\zeta_p]$, and the factorization is determined by its congruence class mod $p$. In particular $\langle q\rangle$ is ramified iff $q = p$ in fact

$$\langle p\rangle = \langle \lambda\rangle^{p-1}$$

where $\lambda = 1-\zeta_p$, otherwise $\langle q\rangle$ is the product of $g = \frac{p-1}{f}$ distinct prime ideals $\langle q\rangle = I_1 \cdots I_g$, where $f$ is $\mathrm{Ord}(q)$ in $(\mathbb{Z}/p\mathbb{Z})^\times$.

The Galois group of $\mathbb{Q}(\zeta_p)$ is also known, we have $\mathrm{Gal}_{\mathbb{Q}}(\mathbb{Q}(\zeta_p)) \cong (\mathbb{Z}/p\mathbb{Z})^\times$.

The *norm* of ideals is defined as $\mathbf{N}(I) = k$, where $k$ is the number of equivalence classes in $\mathbb{Z}[\zeta_p]/I$, this is finite for ideals $I$ in number fields. In the case of the primes occuring in the prime factorization above we have :

$$\mathbf{N} I_i = q^f \equiv 1 \mod p.$$

The polynomial $\Phi_p(x)$ factors $\mod q$ into $g$ factors

$$\Phi_p(x) \equiv \prod_{q \le i \le g} h_i(x) \mod q.$$

The $h_i$ are monic of degree $f$ and are irreducible. A theorem of *Kummer* (who incidentally coined the idea of *ideal numbers* which was later re-introduced in the form of ideals by *Dedekind*), allows us to find the $I_i$.

Let $h_i$ be as above.

**Proposition 8.3** (Kummer)**.** *The prime ideals in $\mathbb{Z}[\zeta_p]$ lying over $\langle q\rangle$ are*

$$I_i = \langle q, h_i(\zeta_p)\rangle, \text{ for } 1 \le i \le g.$$

We need a modification of the above result in case the number we are considering happens to be composite, but not "wildly" so.

**Proposition 8.4.** *Let $n \ge 2$ be an integer, $n$ has order $f$ in $(\mathbb{Z}/p\mathbb{Z})^\times$ and $g = \frac{p-1}{f}$,*

$$\Phi_p(x) \equiv h_i(x) \mod n$$

*where $h_i(x)$ are monic polynomials in $\mathbb{Z}[x]$ each of degree $f$, and let*

$$\mathfrak{I}_i = \langle n, h_i(\zeta_p)\rangle \text{ in } \mathbb{Z}[\zeta_p].$$

*If $r$ is a prime factor of $n$, then in $\mathbb{Z}[\zeta_p]$ we have*

$$\langle r\rangle = \prod_{1 \le i \le g} \langle r, \mathfrak{I}_i\rangle$$

*and each $\langle r, \mathfrak{I}_i\rangle$ is divisible by the same number of prime ideals of $\mathbb{Z}[\zeta_p]$ associated with $r$.*

**Definition 8.5.** Let $I$ be a non-zero prime ideal of $\mathbb{Q}(\zeta_p)$ not dividing $p$, and let $\mathbf{v}_I(\cdot)$ denote the corresponding exponential valuation. For any $\alpha \in \mathbb{Q}(\zeta_p)$ with $\mathbf{v}_I(\alpha) = 0$ (a generalization of the notion that $\alpha \perp I$), we define the *pth power residue symbol* $\left(\frac{\alpha}{I}\right)_p$ to be the unique $p$th root of unity satisfying the congruence,

$$\left(\tfrac{\alpha}{I}\right)_p = \zeta_p^j \equiv \alpha^{\frac{\mathbf{N}I-1}{p}} \mod I.$$

Similar to the generalization of the Legendre symbol to the Jacobi symbol we define

$$\left(\tfrac{\alpha}{\gamma}\right)_p = \prod_{I \not| p,\alpha} \left(\tfrac{\alpha}{I}\right)^{\mathbf{v}_I(\gamma)}.$$

**Theorem 8.6** ($p$th Power Reciprocity Law)**.** *Let $p > 2$ and let $\alpha, \gamma$ be elements of $\mathbb{Q}(\zeta_p)$ relatively prime to $\lambda = 1-\zeta_p$ and to each other. Then there is an independently defined $p$th root of unity $(\alpha,\gamma)_\lambda$ called the* Norm Residue Symbol *such that*

$$\left(\tfrac{\alpha}{\gamma}\right)_p = \left(\tfrac{\gamma}{\alpha}\right)_p (\alpha,\gamma)_\lambda.$$

The symbol $(\alpha,\gamma)_\lambda$ is multiplicative in both arguments and is not changed when $\alpha$ or $\gamma$ is multiplied by a $p$th power. In the cases we will be handling the norm residue symbol becomes trivial, for which we require the following theorem.

**Proposition 8.7.** *If* $\alpha \equiv 1 \mod \lambda^i, \gamma \equiv 1 \mod \lambda^j$ *and* $i+j \geq p+1$ *then* $(\alpha,\gamma)_\lambda = 1$.

Observe that if $\sigma \in \mathrm{Gal}_{\mathbb{Q}}(\mathbb{Q}(\zeta_p))$ then from the defining congruence that

$$\left(\frac{\sigma\alpha}{\sigma I}\right)_p = \sigma\left(\frac{\alpha}{I}\right)_p.$$

We consider power residue symbols in the special case where $q$ is a rational prime and $p \backslash (q-1)$. Let $t = t_q$ be a primitive root for $q$.

The $\Phi_p(x) = \prod_{1\leq i\leq (p-1)}\left(x - t^{\frac{q-1}{p}i}\right) \mod q$. By the Proposition, there is a "canonical" prime $I$ associated to $q$ in $\mathbb{Z}[\zeta_p]$ given by

$$I = \langle q, \zeta_p - t^{\frac{q-1}{p}}\rangle.$$

If $x$ is a rational integer, then similar to the case $p = 2$ we have

$$\left(\frac{x}{I}\right)_p = \zeta_p^{\mathrm{Ind}_q(x)}.$$

Where the indices are computed with the primitive root $t$. Since

$$\begin{aligned}\left(\frac{x}{I}\right)_p &\equiv x^{\frac{q-1}{p}} \\ &\equiv t^{\mathrm{Ind}_q(x)\frac{q-1}{p}} \\ &\equiv \zeta_p^{\mathrm{Ind}_q(x)} \mod I\end{aligned}$$

as

$$t^{\mathrm{Ind}_q(x)\frac{q-1}{p}} - \zeta_p^{\mathrm{Ind}_q(x)} \in I.$$

Thus knowing $\left(\frac{x}{I}\right)_p$ for the canonical prime $I$ is equivalent to knowing $\mathrm{Ind}_q(x) \mod p$.

The algorithm to be designed depends on our ability to produce "good" elements in the ring with known prime factorization.

**Definition 8.8.** If $I$ is a prime of $\mathbb{Q}(\zeta_p)$ not dividing $p$ and if $a,b \in \mathbb{Z}$, we define the *Jacobi sum*

$$J_{a,b}(I) = \sum \left(\frac{x}{I}\right)_p^{-a}\left(\frac{1-x}{I}\right)_p^{-b}$$

where the sum is over a set of coset representatives $x$ of $\mathbb{Z}[\zeta_p]/I$ other than $0, 1 \mod I$.

The prime factorization of $J_{a,b}(I)$ is given by the following result.

**Proposition 8.9.** *(Stickelberger) Suppose* $a,b \in \mathbb{Z}$ *with* $ab(a+b) \not\equiv 0 \mod p$. *For* $u \in \mathbb{Z}$, *let*

$$\theta_{a,b}(u) = \left\lfloor \frac{(a+b)}{p}u \right\rfloor - \left\lfloor \frac{a}{p}u \right\rfloor - \left\lfloor \frac{b}{p}u \right\rfloor.$$

*Then*

$$\langle J_{a,b}(I)\rangle = \prod_{1\leq u\leq p-1} \sigma_u^{-1}(I)^{\theta_{a,b}(u)}.$$

*Where* $\sigma_u$ *is the automorphism sending* $\zeta_p \mapsto \zeta_p^u$.

Note: Since

$$\left\lfloor \frac{x+y}{z} \right\rfloor = \left\lfloor \lfloor \frac{x}{z} \rfloor + \lfloor \frac{y}{z} \rfloor + \left\{\frac{x}{z}\right\} + \left\{\frac{y}{z}\right\} \right\rfloor$$

the quantity $\theta_{a,b}(u)$ is either 0 or 1.

Let $I$ be as before.

**Proposition 8.10** (Iwasawa)**.** *For all* $a,b \in \mathbb{Z}$

$$-J_{a,b}(I) \equiv 1 \mod \lambda^2.$$

**Proposition 8.11.** *If $p > 2$, there exist $a, b \in \mathbb{Z}$ such that $ab(a+b) \neq 0 \mod p$ and*

$$\hat{\theta}_{a,b} \equiv_{def} \sum_{1 \leq u \leq p-1} \theta_{a,b}(u) u^{-1} \not\equiv 0 \mod p.$$

The above result suggests that Jacobi sums are useful substitutes for primes, in that some useful information can still be gleaned from them.

8.2. **Probabilistic Version of the Algorithm.** Here we give an easy to digest version of the algorithm which resembles the actual algorithm except that we simplify certain steps using random strings.
Let $n \in \mathbb{N}$ be the number to be tested.

**Preparation Step** (a) Compute $f(n)$, the least square-free number such that

$$\prod_{(q-1) \setminus f(n), q \text{ prime}} q > \sqrt{n}$$

$$\text{Initial-Primes}(n) = \{p \mid p \setminus f(n)\}$$
$$\text{Euclidean-Primes}(n) = \{q \mid (q-1) \setminus f(n)\}$$

How do we know that this can always be done? We have the following theorem.
**Theorem 8.12** (Mirsky)**.** *There are infinitely many primes $q$ such that $q-1$ is square-free.*
So $f(n)$ always exists. We will be able to show that $f(n) \leq (\ln n)^{c \ln \ln \ln n}$ for all large enough $n$.

(b) Compute and fix a primitive root $t_q$ for each Euclidean prime $q$. Also check that $n \perp p, n \perp q$. If any of the tests fail declare $n$ composite.

(c) For each initial prime $p > 2$, find $a, b \in \mathbb{Z}$ such that $0 < a, b < p, a + b \equiv 0 \mod p$, and

$$\hat{\theta}_{a,b} = \sum_{1 \leq u \leq p-1} \theta_{a,b}(u) u^{-1} \not\equiv 0 \mod p.$$

The proposition above guarantees this. For $p = 2$, let $a = b = \hat{\theta}_{a,b} = 1$.

(d) Compute a "Jacobi Sum" $J_p(q)$ for each initial prime $p$ and euclidean prime $q$ with $p \setminus (q-1)$ as follows.

If $p = 2$, set $J_p(q) = -q$.
If $p > 2$, let

$$\begin{aligned} J_p(q) &= -J_{a,b}(I) \\ &= -\sum_{2 \leq x \leq q-1} \left(\tfrac{x}{I}\right)_p^{-a} \left(\tfrac{1-x}{I}\right)_p^{-b} \in \mathbb{Q}(\zeta_p), \end{aligned}$$

where $a, b$ are the integers defined in the previous step, and $I$ is defined by $\langle q, \zeta_p - t_q^{\frac{q-1}{p}} \rangle$ is the canonical prime associated to $q$, with respect to $t_q$. Since $\mathbb{Z}[\zeta_p]/I \cong \mathbb{Z}/q$, rational integers can be used as the coset representatives in the sum.
To compute $\left(\frac{x}{I}\right) = \zeta_p^j$ we can make a table of $t^{\frac{j(q-1)}{p}}$ for $0 \leq j \leq p-1$ and look up $x^{\frac{q-1}{p}}$ to determine the $j$.
If $p > 2$ and if $r$ is any number $r \perp p$, then the norm residue symbol $(J_p(q), r)_\lambda$ is trivial. This can be seen as follows :

$$(J_p(q), r)_\lambda = (J_p(q), r^{1-p})_\lambda = (J_p(q), (r^{-1})^{p-1})_\lambda$$

since the norm residue symbol is unaffected by $p$th powers in both arguments. We also have

$$(r^{-1})^{p-1} \equiv 1 \mod \lambda^{p-1}$$

by Fermat, and $\langle p \rangle = \langle \lambda^{p-1} \rangle$. Also $J_p(q) \equiv 1 \mod \lambda^2$, by the proposition 8.10. Thus $\langle J_p(q), r \rangle_\lambda = 1$ follows from proposition 8.7. We have used the multiplicative nature of the norm residue symbol in both arguments here.

(e) **[Probabilistic Step]** For each $p$, factor $n$ into ideals in $\mathbb{Z}[\zeta_p]$ ($p$ is an initial prime). We attempt this as follows.
Let $f = \mathrm{Ord}(n)$ in $(\mathbb{Z}/p\mathbb{Z})^\times$. Let $g = \frac{p-1}{f}$.
Attempt to factor $\Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + x + 1 \equiv \prod_{1 \le i \le g} h_i(x) \mod n$, where each $h_i(x) \in \mathbb{Z}[x]$ is monic and has degree $f$.
If $n$ is in fact prime, then with high probability $\Phi_p(x)$ can be factored over $\mathbb{Z}/n\mathbb{Z}$, by using the Berlekamp or Rabin algorithm in time polynomial in $p$ and $\lg n$.
If factorization can be carried out, then we set

$$\mathfrak{I}_i = \langle n, h_i(\zeta_p) \rangle, \text{ for } 1 \le i \le g.$$

**Extraction Step** Suppose $\mathfrak{I}$ is an ideal of $\mathbb{Z}[\zeta_p]$ not dividing $\langle \lambda \rangle$ and $\alpha \in \mathbb{Z}[\zeta_p]$. We have from the previous section, that if $\mathfrak{I}$ is not a proper prime ideal and if $\alpha \in \mathfrak{I}$, then

$$\alpha^{\frac{\mathbf{N}\mathfrak{I}-1}{p}} \equiv \zeta_p^j \mod \mathfrak{I}.$$

However if $\mathfrak{I} \not| \langle \lambda \rangle$, this can hold for at most one such root of unity, if $\zeta_p^i \equiv \zeta_p^j \mod \mathfrak{I}$, then

$$\begin{aligned} \zeta_p^i(\zeta_p^{j-i} - 1) &\in \mathfrak{I}, \\ \text{as } \zeta_p^i \notin \mathfrak{I}, \text{ and } \mathfrak{I} &\text{ is prime} \\ \zeta_p^{j-i} - 1 &\in \mathfrak{I} \\ \zeta_p^{j-i} &\equiv 1 \mod \mathfrak{I} \\ \text{or } j - i \equiv_p 0 &\Rightarrow i = j. \end{aligned}$$

We define the *mock residue symbol*

$$\left\langle \frac{\alpha}{\mathfrak{I}} \right\rangle_p = \begin{cases} \zeta_p^j \equiv \alpha^{\frac{\mathbf{N}\alpha-1}{p}} \mod \mathfrak{I}, & \text{if it holds.} \\ 0, & \text{otherwise.} \end{cases}$$

(a) For each initial prime $p$, and each euclidean prime $q$ with $p \setminus (q-1)$, compute the mock residue symbols

$$\left\langle \frac{J_p(q)}{\mathfrak{I}_i} \right\rangle_p \text{ for } 1 \le i \le g,$$

for each of the ideals found in the last step of the previous stage. If any of the mock residue symbols are 0, declare $n$ as composite. It can be shown that the computation of the mock residue symbol can be accomplised in time polynomial in $p$ and $\lg n$. It is clear that the number of euclidean primes $< f(n)$, so this step requires time polynomial in $f(n)$.

(b) **[Probabilistic Step]** For each initial prime $p$ do the following. If the mock residue symbols for $p$ are not all equal to 1, select some non-trivial one, $\left\langle \frac{\gamma}{\mathfrak{I}} \right\rangle_p$ and call it the distinguished symbol corresponding to $p$. If they are all 1, compute mock residue symbols $\left\langle \frac{\gamma}{\mathfrak{I}_i} \right\rangle_p$ for other $\gamma$ selected at random. If $n$ is prime the probability of an arbitrarility chosen $\gamma$ working is roughly $\frac{p-1}{p}$.

(c) For each pair $p, q$ with $p \setminus (q-1)$ compute the exponents $m_{i,q}$ such that

$$\left\langle \frac{\gamma}{\mathfrak{I}} \right\rangle_p^{m_{i,q}} = \left\langle \frac{J_p(q)}{\mathfrak{I}_i} \right\rangle_p, 0 \le m_{i,q} < p.$$

These relations are preserved for the power residue symbols at prime ideals dividing the $\mathfrak{I}_i$.

Finally we have an "extraction" lemma which allows relations among the mock residue symbols to translate into useful relations among the actual residue symbols.

**Theorem 8.13** (Extraction Lemma)**.** *Let $\mathfrak{I}$ and $\mathfrak{I}_1$ be ideals of $\mathbb{Z}[\zeta_p]$ such that $p \not| \mathbf{N}\mathfrak{I} = \mathbf{N}\mathfrak{I}_1$ and let $\mathcal{R}, \mathcal{R}_1$ be associated primes dividing $\mathfrak{I}$ and $\mathfrak{I}_1$ respectively.*
*Suppose there is some $\gamma \in \mathbb{Z}[\zeta_p]$ such that $\left\langle \frac{\gamma}{\mathfrak{I}} \right\rangle_p$ is not 0 or 1. Then for any $\alpha \in \mathbb{Z}[\zeta_p]$, the relation*

$$\left\langle \frac{\gamma}{\mathfrak{I}} \right\rangle_p^m = \left\langle \frac{\alpha}{\mathfrak{I}_1} \right\rangle_p$$

*implies the following relation*

$$\left(\frac{\gamma}{\mathcal{K}}\right)_p^m = \left(\frac{\gamma}{\mathcal{K}_1}\right)_p.$$

Let $k$ be the largest power of $p$ which divides $\mathbf{N}\mathfrak{D}-1$.

$$\left(\gamma^{\frac{\mathbf{N}\mathfrak{D}-1}{p^k}}\right)^{p^{k-1}} \equiv \zeta_p^j \mod \mathfrak{D}$$

The relations among mock residue symbols relating it to the mock residue symbol of the jacobi sum, allow us to evaluate the power residue symbols in terms of one unknown, by the following type of calculation.
Suppose $r$ is a prime divisor of $n$. Let $p>2$ be an initial prime.
Let $\left\langle\frac{\gamma}{\mathfrak{D}}\right\rangle_p$ be the distinguished symbol corresponding to $p$. Knowing this mock residue does *not* allows us to compute $\left(\frac{\gamma}{\langle r,\mathfrak{D}\rangle}\right)_p$, but there are only $p$ possibilities for it.
The value of $\left(\frac{\gamma}{\langle r,\mathfrak{D}\rangle}\right)_p$ completely determines each $\mathrm{Ind}_q(r) \mod p$, for every euclidean prime $q$ with $p\backslash q-1$.
So how do we find it? The trick is to evaluate $\left(\frac{J_p(q)}{r}\right)_p$ in two ways. First

$$\left(\frac{J_p(q)}{r}\right)_p = \left(\frac{r}{J_p(q)}\right)_p (J_p(q),r)_\lambda = \left(\frac{r}{J_p(q)}\right)_p$$

by the power reciprocity law. By the step $(d)$ we have

$$\begin{aligned}\left(\frac{r}{J_p(q)}\right)_p &= \prod_{1\le u\le p-1}\left(\frac{r}{\sigma_u^{-1}I}\right)_p^{\theta_{a,b}(u)}\\ &= \prod_{1\le u\le p-1}\sigma_u^{-1}\left(\frac{r}{I}\right)_p^{\theta_{a,b}(u)}\\ &= \prod_{1\le u\le p-1}\left(\frac{r}{I}\right)_p^{u^{-1}\theta_{a,b}(u)}\\ &= \left(\frac{r}{I}\right)_p^{\hat{\theta}_{a,b}},\end{aligned}$$

where $I$ is the canonical prime associated to $q$. We have used the factorization of Jacobi sums and functoriality of the residue symbol. On the other hand,

$$\begin{aligned}\left(\frac{J_p(q)}{r}\right)_p &= \prod_{1\le i\le g}\left(\frac{J_p(q)}{\langle r,\mathfrak{D}_i\rangle}\right)_p\\ &= \prod_{1\le i\le g}\left(\frac{\gamma}{\langle r,\mathfrak{D}\rangle}\right)_p^{m_{i,q}} \text{ where } \mathfrak{D} \text{ is the chosen ideal,}\\ &= \left(\frac{\gamma}{\langle r,\mathfrak{D}\rangle}\right)_p^{\sum_{1\le i\le g} m_{i,q}}.\end{aligned}$$

By Proposition 8.4 there is a bijection between primes dividing $\langle r,\mathfrak{D}_i\rangle$ and $\langle r,\mathfrak{D}\rangle$.
Now $a,b$ were selected such that $\hat{\theta}_{a,b}$ is invertible $\mod p$.
So

$$\left(\frac{r}{I}\right)_p^{\hat{\theta}_{a,b}} = \left(\frac{\gamma}{\langle r,\mathfrak{D}\rangle}\right)_p^{\sum m_{i,q}}$$

$$\left(\frac{r}{I}\right)_p = \left(\frac{\gamma}{\langle r,\dashv\rangle}\right)^{\hat{theta}_{a,b}^{-1}\sum m_{i,q}}$$

So if $\left(\frac{\gamma}{\langle r,\mathfrak{D}\rangle}\right)_p = \zeta_p^k$ then

$$\mathrm{Ind}_q(r) \equiv r\hat{\theta_{a,b}}^{-1}\Big(\sum_{1\le i\le g} m_{i,q}\Big) \mod p.$$

**Consolidation Step** If $p_1,\cdots,p_d$ are the initial primes and $\left\langle\frac{\gamma_i}{\mathfrak{D}_i}\right\rangle_{p_i}$ are the corresponding distinguished symbols found in the extraction step, then for each prime factor $r$ or $n$ there are integers $k_1,\cdots,k_d$ such that

$$\left(\frac{\gamma_i}{\langle r,\mathfrak{D}_i\rangle}\right)_{p_i} = \zeta_{p_i}^{k_i}, 1\le i\le d.$$

By the CRT we have a single $k \mod \prod p_i$ such that

$$\left(\frac{\gamma_i}{\langle r, \mathcal{O}_i\rangle}\right)_{p_i} = \zeta_{p_i}^k, 1 \le i \le d.$$

For each $k, 1 \le k \le f(n) = \prod p_i$ , we assemble and test a possible divisor $r = r(k)$ of $n$.

(a) Use CRT to compute for each $q > 2$ integers $I(k,q)$ such that

$$I(k,q) \equiv k\hat{\theta}_{a,b}^{-1} \sum_{1 \le i \le g} m_{i,q} \mod p$$

for each $p$ with $p \backslash q-1$. Let $I(k,2) = 1$. If $k$ is an actual prime factor $r$ of $n$, then $I(k,q)$ are the $\mathrm{Ind}_q(r)$ mod $p$.

(b) For each $q$ compute the least positive integer such that

$$r(k,q) \equiv t_q^{I(k,q)} \mod q.$$

If $k$ is a factor then $r(k,q)$ are the $r \mod q$.

(c) Use CRT to compute the least positive integer such that for each $q$, $r(k) \equiv r(k,q) \mod q$ .
If $k$ corresponds to an actual factor $r$ or $n$, then $r(k) \equiv r \mod Q = \prod q$. Since $r \le \sqrt{n} < Q$ we have $r(k) = r$.

(d) Check if $r(k) \backslash n$, if it does and $r(k) \neq 1$, declare $n$ composite and halt. Otherwise check the next value of $k$.

(e) Declare $n$ prime.

**Theorem 8.14.** *The above algorithm correctly determines whether n is prime or composite, if it terminates. There is an absolute, calculable constant $c > 0$ such that for every $k \ge 1$, if n is prime, the algorithm terminates within $T_k(n)$ steps with probability greater than $1 - 2^{-k}$, where $f(n) \le T_k(n) \le k f(n)^c$, and $f(n) = (\lg n)^{d \lg\lg\lg n}$.*

The two probabilistic steps can be replaced using a stronger extraction lemma and a gcd like computation which eliminates the need for factoring, yielding an algorithm which runs in time $f(n)$.

## 9. Goldwasser-Kilian Test

In §6 we saw how helpful an element of large order can be in proving that a number is prime. The problem is that there is no way of telling that a number has such and such order without knowing the prime factorization of $n-1$, where $n$ is the number that we are testing for primality. However it turns out that if we can pick elements of high order in other groups then we can still do the trick. Specifically we will look at the group of points on *elliptic curves*, over the ring $\mathbb{Z}/n\mathbb{Z}$, which becomes a field if $n$ is prime. Elliptic curves have the advantage that a large number of them can be generated at random easily, and also that there is a deterministic polytime algorithm for computing the order of the elliptic group over finite fields. In this section we shall sketch the Goldwasser-Kilian test, which works for an infinite subset of primes.

### 9.1. Elliptic Curves.

Here we shall give a sketchy account of the properties of Elliptic Curves that we will use in the algorithm.

Let $\mathbb{F}$ be an arbitrary field, we define an elliptic curve (in Weierstrass Normal form), to be set $\{(x,y) \in \mathbb{F}^2 \mid y^2 = x^3 + Ax + B\}$, where $A, B \in \mathbb{F}$. We shall require that the elliptic curve be non-singular and this happens iff the discriminant of the curve is not zero. The discriminant for an elliptic curve as given above is $\Delta = 4A^3 + 27B^2$, and we assume char $\mathbb{F} \neq 2$ or 3 (this is required to derive the Weierstrass form).

Let us consider the set $E(A,B) = \{(x,y) \in \mathbb{F}^2 \mid y^2 = x^3 + Ax + B\} \cup I$, which is the point at infinity (this is where a projective version of the definition will seem to make more sense). Now we define an addition operation on the points of $E(A,B)$ as follows (this is the well known tangent and chord construction for generating other solutions to the equation given two solutions).

Let $L$ and $M$ be two points not equal to $I$. If $L = M$, then consider the tangent of the curve at $L$, otherwise look at the line joining $L$ and $M$. If this line is vertical define $L + M = I$, otherwise define $L + M$ to be the reflection on the $x$-axis of the intersection of the line with the curve. This yields an algorithm for adding two points which is given below.

One can apply this addition algorithm to elliptic curves over arbitrary rings provided the required inverses exist. In the situation in which we use the above, $n$ is ostensibly a prime number so if some inverses do not exist we can immediately conclude that the number is not prime.

**Algorithm 9.1** Procedure for adding two points on an Elliptic Curve

Add$(P_1, P_2)$
Let $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$
if $(x_1 = x_2) \wedge (y_1 = -y_2)$ (Line is vertical)
  return$(I)$
if $(P_1 = P_2)$  Let $\lambda = \frac{3x_1^2 + A}{2y_1}$ (The derivative of the curve at this point)
else
  $\lambda = \frac{y_2 - y_1}{x_2 - x_1}$.
Let $\beta = y_1 - \lambda x_1$ (The $x$-intercept of the line joining the two points).
$x_s = \lambda^2 - x_1 - x_2$ (The intersection with the curve).
$y_s = -(\lambda x_s + \beta)$ (Reflect the point).
return$P_s = (x_s, y_s)$

It turns out that with the above operation, the set $E(A,B)$ becomes an abelian group with respect to the addition operation with $I$ as the identity element. We can also define multiplication of the points by integers (i.e., raising to a power in the additive group) as follows :

$$\begin{aligned} 0M &= I, \\ qM &= (q-1)M + M, q \text{ odd}, \\ qM &= \frac{q}{2}M + \frac{q}{2}M, q \text{ even} \end{aligned}$$

So we can evaluate $qM$ using $O(\lg q)$ additions.
Let $E_n(A,B)$ be the set of points on the elliptic curve over $\mathbb{Z}/n\mathbb{Z}$. We know that $E_p(A,B)$ is a group and in this case let $N_p(A,B)$ be the order of this group. There is a nice algorithm by Schoof, which computes $N_p(A,B)$ in time $O(\lg^9 p)$. By the Weil's famous theorem (actually Hasse proved this, but Weil generalized his theorem to curves of arbitrary genus), we have $N_p(A,B) = p + 1 - t$, where $|t| \leq 2\sqrt{p}$.

**Definition 9.1.** Let $M \in E_n(A,B)$, define $(M)_p \in E_p(A,B)$ as $(x \mod p, y \mod p)$ where $p > 3$ is a prime divisor of $n$.

Now $4A^3 + 27B^2 \neq 0 \mod p$, since $n \perp (4A^3 + 27B^2)$. So this projection is well defined.
The projection is useful in the following sense:

**Lemma 9.2.** *Let $L, M \in E_n$, and $p > 3$ be a prime such that $p \backslash n$, if $L + M$ is well defined (the addition algorithm was able to complete), then $(L+M)_p = (L)_p + (M)_p$.*

The above lemma can be proved from the definition of the addition operation. It plays the role of the extraction lemma which we showed for the APR test.

**Corollary 9.3.** *If $M \in E_n$ and $qM = I$ and if $p > 3$ is a prime such that $p \backslash n$, then $(qM)_p = I$.*

9.2. **The Algorithm.** A simple version of the algorithm works as follows :

Step 1 Let $p$ be the number we wish to show is prime. Select a random elliptic group $\mod p$, this is done by picking at random $A, B \in \mathbb{Z}/p\mathbb{Z}$, making sure that the discriminant is non-zero. Let $E_p(A,B)$ be the group. Now compute the order of this order using Schoof's algorithm $O(\lg^9 p)$.
Using a standard probabilistic test, verify that $\frac{N_p(A,B)}{2} = q$ is a prime. If the order is not of this form, then repeat the test.

Step 2 Once we have such a group, randomly select points on the group. This can be done by randomly picking $x \in \mathbb{Z}/p\mathbb{Z}$ and then checking whether $x^3 + Ax + B$ is a quadratic residue, and if so compute the square root of this $\mod p$, all this can be done in expected polynomial time. Check whether the point is of order $q$, if not repeat till one is found.

Step 3 Now recursively prove that $q$ is prime, and if $q$ is sufficiently small then a deterministic algorithm like the APR algorithm can then be used to prove $q$ prime.

Since we are generating a certificate for primality which can be verified, the use of a coRP algorithm inside the RP type algorithm is harmless. The following theorem proves the correctness of the reduction.

**Theorem 9.4.** *For all n not divisible by* 2 *or* 3*, if* $\exists M,q,A,B$ *such that* $q > n^{\frac{1}{2}} + 1 + 2n^{\frac{1}{4}}$*, q prime,* $n \perp (4A^3 + 27B^2)$*,* $M \neq I, M \in E_n(A,B)$*, and* $qM = I$*, then n is prime.*

**Proof :** Suppose $n$ was composite. Then $\exists p : p < \sqrt{n}$, $p$ prime such that $p \backslash n$. If $qM = I$, then $qM_p = I$ by the corollary above. Thus $\text{Ord}(M_p) \backslash q$. However $\text{Ord}(M_p) \leq N_p \leq p + 1 + 2\sqrt{p}$ (By Weil), which is less than $q$. Since $q$ is prime, we have $\text{Ord}(M_p) = 1$. This implies $M_p = I$ which is a contradiction since $M \neq I$. □

## 10. Conclusion

Primality testing is an important primitive in Cryptography. Though the complexity of this problem has not been completely ascertained, we know for example that there are "efficient" probabilistic algorithms for it. The related celebrated problem of finding the prime factorization of numbers seems to be far more difficult, with no polynomial algorithms randomized or otherwise known.

The primality tests which are being used in cryptographic systems, have a small probability of declaring a number prime when a series of tests did not find a witness to its compositeness, to eliminate this uncertainity we need to use a ZPP type algorithm but the Adleman-Huang algorithm seems to be impractical for implementation. It would be a very useful pursuit to try and find a simplified version of the test which can be implemented easily.

## REFERENCES


[AH92] Leonard M. Adleman, Ming-Deh A. Huan, *Primality Testing and Abelian Varieties Over Finite Fields*, Vol. 1512, Lecture notes in Mathematics, Springer-Verlag, 1992.

[APR83] Leonard M. Adleman, Carl Pomerance, Robert S. Rumely, *On Distinguishing Prime Numbers from Composite Numbers*, Annals of Mathematics, **117**, 173-206, 1983.

[AGP94] Alford W. R., Andrew Granville, Carl Pomerance, *There are infinitely many Carmichael numbers*, Annals of Mathematics, **140**, 703-722, 1994.

[Ank52] Ankeny N. C., *The Least Quadratic Non-Residue,* Annals of Mathematics **55**, 65-72, 1952.

[BS96] Eric Bach, Jeffrey Shallit, *Algorithmic Number Theory : Efficient Algorithms*, MIT-Press, 1996.

[Car12] Carmichael R. D., *On Composite numbers $p$ which satisfy the fermat congruence $a^{p-1} \equiv p \mod p$,* American Mathematical Monthly, **19**,22-27, 1912.

[GK86] Shafi Goldwasser, Joe Kilian, *Almost all primes can be quickly certified,* ACM STOC, 316-329, 1986.

[PSS89] Janos Pintz, William L. Steiger, and Endre Szemerédi. *Infinite sets of primes with fast primality tests and quick generation of large primes.* Mathematics of Computation, **53**, 399-406, 1989.

[Rab80] Michael O. Rabin, *Probabilistic Algorithm for testing Primality*, Journal of Number Theory **12**, 128-138, 1980.

[Sch97] Manfred Schroeder, *Number Theory in Science and Communication: With Applications in Cryptography, Physics, Digital Information, Computing, and Self-Similarity* , Springer-Verlag, 1997.

[SS77] Solovay R., Strassen V., *A fast Monte-Carlo test for Primality,* SIAM Journal of Computing, **6**, 84-85, 1977.