As Gauss wrote in 1801 [7],

"*Problema, numeros primos a compositis dignoscendi, hosque in factores suos primos resolvendi, ad gravissima ac utilissima totius arithmeticae pertinere, et geometrarum tum veterum tum recentiorum industriam ac sagacitatem occupavisse, tam notum est, ut de hac re copiose loqui superfluum foret. [P]raetereaque scientiae dignitas requirere videtur, ut omnia subsidia ad solutionem problematis tam elegantis ac celebris sedulo excolantur.*

[The problem of distinguishing prime numbers from composite numbers and of resolving the latter into their prime factors is known to be one of the most important and useful in arithmetic. It has engaged the industry and wisdom of ancient and modern geometers to such an extent that it would be superfluous to discuss the problem at length. Further, the dignity of the science itself seems to require that every possible means be explored for the solution of a problem so elegant and so celebrated.]"

After two centuries of exploring "every possible means" we finally have an essentially satisfactory solution to the first part of the problem posed by Gauss: distinguishing primes from composites. This is a critical first step toward addressing the second part of the problem (resolving composites into their prime factors), which remains open. Without the ability to distinguish primes and composites we have no way of knowing when we have computed the prime factorization of a number. What is the prime factorization of 2017? How about 2021?

Primality testing is a huge topic that we cannot hope to cover in one lecture, we refer the interested reader to [6] for an in depth treatment. Our goal in this lecture is to first give a brief overview of the topic, including how primality testing is done in practice, and then present the main theoretical result: a polynomial-time algorithm for determining whether a given integer is prime or composite.

8.1 Primality testing using modular arithmetic

A key component of all the primality testing algorithms we will consider is modular arithmetic (and generalizations thereof). Let us begin by briefly reviewing some basic facts that are probably familiar to most of you.

The set of integers $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ is an example of a *commutative ring*: it comes equipped with binary operations $+$ and $\times$, both of which are each commutative, associative, have an identity element, and satisfy the distributive law $a(b+c) = ab + ac$. Every element of a ring R has an additive inverse, but not every element necessarily has a multiplicative inverse; the set of elements that do have a multiplicative inverse is called the *unit group* of the ring, denoted $R^\times$; for example, $\mathbb{Z}^\times = \{\pm 1\}$. A *group* is a set with an associative binary operation and an identity element in which every element has an inverse; if the operation is commutative the group is said to be *abelian.* There are two abelian groups associated to a commutative ring R: the additive group $(R, +)$ and the multiplicative group $(R^\times, \times)$. Rings whose multiplicative group consists of all nonzero elements are called *fields.*

Given a ring R one can form additional rings by taking the *quotient* of R modulo an *ideal* I. An ideal of a ring is a nonempty subset $I \subseteq R$ that is closed under addition and by multiplication by elements in R; for example, for any $n \in \mathbb{Z}$ the set $n\mathbb{Z} := \{nz : z \in \mathbb{Z}\}$ is an ideal. Every ideal I is an abelian group under addition, but I is a ring if and only if $I = R$. But associated to each ideal I of R we have a quotient ring R/I whose elements consist of *equivalence classes* under the following equivalence relation: two elements $a, b \in R$ are in the same equivalence class modulo I if and only if $a - b \in I$. We can then use the map

Lecture by Andrew Sutherland

$a \to [a]$ that sends each $a \in R$ to its equivalence class $[a] \in R/I$ to define addition and multiplication in R/I via $[a]+[b] := [a+b]$ and $[a][b] := [ab]$. In order for this to make sense one needs to check that if $[a]=[a']$ and $[b]=[b']$ then $[a'+b']=[a+b]$ and $[a'b']=[ab]$, but this is always the case. A familiar example of a quotient ring is the ring $\mathbb{Z}/n\mathbb{Z}$ of integers modulo n; we can add and multiply in $\mathbb{Z}/n\mathbb{Z}$ using a set of unique representatives $[0,1,\ldots,n-1]$ and reducing modulo n as we go.

The rings $\mathbb{Z}/n\mathbb{Z}$ are of particular interest to us because the structure of this ring, and in particular its unit group $(\mathbb{Z}/n\mathbb{Z})^\times$, depends critically on whether n is prime or not. Indeed, the ring $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime. To see this, note that

$$[a] \in (\mathbb{Z}/n\mathbb{Z})^\times \quad \Longleftrightarrow \quad \gcd(a,n)=1,$$

since if $\gcd(a,n)>1$ then no multiple of a can be congruent to 1 modulo n (so $[a]$ has no multiplicative inverse) and if $\gcd(a,n)=1$ then we can use the extended Euclidean algorithm to compute integers s and t such that $sa+tn=1$, in which case $[s]$ is the multiplicative inverse of $[a]$ because $sa \equiv 1 \bmod n$.

Definition 8.1. The function $\phi(n) := \#(\mathbb{Z}/n\mathbb{Z})^\times$ is known as *Euler's totient function.* It can be computed via the formula

$$\phi(n) = n\prod_{p|n}\left(1-\frac{1}{p}\right)$$

where p ranges over the prime factors of n.

The formula for $\phi(n)$ is easy to prove: the probability that a random $a \in \{0,\ldots,n-1\}$ is not divisible by a particular prime divisor p of n is $1-1/p$, and the probability that a is not divisible by any prime divisor of n, equivalently, $\gcd(a,n)=1$, is the product of these probabilities, which is also equal to $\phi(n)/n$; for distinct primes p and q the probabilities are independent because an integer a is divisible by pq if and only if it is divisible by both p and q. The formula for $\phi(n)$ gives us a simple criterion for primality.

Proposition 8.2. *An integer n is prime if and only if $\phi(n)=n-1$.*

The proposition is obviously true, but at first glance it does not appear to be very useful, since computing $\phi(n)$ with our formula requires us to know the prime factorization of n (and in particular, whether n is prime or not). The key to making Proposition 8.2 useful is to use the fact that $(\mathbb{Z}/n\mathbb{Z})^\times$ is a group.

A basic fact about groups is that the order (cardinality) of any subgroup H of a group G divides the order of G, this follows from the fact that we can partition G into cosets of H, each of which has cardinality $\#H$. For any $a \in G$ the *coset* $aH := \{ah : h \in H\}$ has cardinality $\#H$ because the map $h \mapsto ah$ is a bijection, and two cosets aH and bH are either disjoint or equal: if $ah_1 = bh_2$ with $h_1,h_2 \in H$ then $aH = \{bh_2h_1^{-1}h : h \in H\} = bH$. This applies, in particular, to any cyclic subgroup $\langle a\rangle := \{a,a^2,a^3,\ldots\}$ with $a \in G$. Thus the order of $a \in G$, by which we mean the order of $\langle a\rangle$, is always a divisor of $\#G$.

Corollary 8.3. *If $a^{n-1} \not\equiv 1 \bmod n$ with $1 \le a < n$ then n is composite.*

Proof. If n is prime then $(\mathbb{Z}/n\mathbb{Z})^\times$ has order $\phi(n)=n-1$ and the order of $[a] \in (\mathbb{Z}/n\mathbb{Z})^\times$ is a divisor d of $n-1$. Therefore $a^d = 1 \bmod n$ and $a^{n-1} = (a^d)^{(n-1)/d} = 1^{(n-1)/d} = 1 \bmod n$ □

An integer a that satisfies the condition of the proposition is said to be a *witness* for n; it can be viewed as a *certificate* of the fact that n is composite. One verifies this certificate by checking that $a^{n-1} \not\equiv 1 \bmod n$, which can be done efficiently using binary exponentiation. Every integer n has a unique binary representation, and we can use this representation compute a^{n-1} efficiently using a product of repeated squarings. For example, if $n-1 = 34 = 10010_2 = 2^5 + 2^1$ then $a^{n-1} = a^{2^5}a^2$ and we can compute $a^2, a^4 = (a^2)^2, a^8 = (a^4)^2, a^{16} = (a^8)^2, a^{32} = (a^{16})^2$. With this approach no more than $2\lg n$ multiplications are required to compute $a^{n-1} \bmod n$; here $\lg n$ denotes the base-2 logarithm of n. By reducing modulo n at each step, we can are always working with $(\lg n)$-bit integers; this leads to an algorithm that uses $O(\log^3 n)$ bit operations using standard "schoolbook" methods for multiplication and division with remainder; this can be improved to $O(\log^{2+\epsilon} n)$ using more sophisticated methods based on the fast Fourier transform (FFT).

Remark 8.4. Here and throughout we use the standard asymptotic notation $O(g(t))$ to denote a positive function $f(t)$ for which

$$\lim_{t\to\infty} \frac{f(t)}{g(t)} \le c$$

for some constant c. When measuring the asymptotic complexity of an algorithm we typically express it as a function of the size of the input (measured in bits, the number of 0s and 1s in its binary representation). A *polynomial-time* algorithm is one whose running time is bounded by a polynomial function of the size of its input; in the context of primality testing, this means an algorithm that takes an integer n as input and determines whether or not n is prime using $O(\log^k n)$ bit operations, for some fixed exponent k.

If we knew that every composite integer n has a witness $a \le B = c\log^k n$ for some constants c and k then Corollary 8.3 would give us a polynomial-time primality test: we could simply check each positive integer $a \le B$ to see whether it is a witness for n. But unfortunately this is not the case, in fact there are composite integers n that do not have any witnesses other than those that share a prime factor of n (so finding one amounts to finding a proper divisor of n),

Definition 8.5. A (positive) composite integer n for which $a^{n-1} \equiv 1 \bmod n$ for all integers a relatively prime to n is called a *Carmichael number.*

The first three Carmichael numbers are 561, 1105, and 1729. More can be found in the OEIS sequence A002997 for more), and it is known that there are infinitely many [2].

8.2 A probabilistic primality test

The existence of Carmichael numbers implies that we cannot use Corollary 8.3 to efficiently test primality, but if we strengthen our notion of a witness we can. Let us now assume that the integer n whose primality we wish to test is a positive odd number (even numbers are easy to test for primality). By repeatedly dividing $n-1$ by 2 until we obtain an odd number m we can always write n in the form $n = 2^r m + 1$ with m odd and $r \ge 1$; note that we don't need to factor $n-1$ to do this, we are just computing its 2-adic valuation).

Proposition 8.6. *Let* $n = 2^s t + 1$ *with* $t \ge 1$ *odd and* $s \ge 1$. *If* $a^m \not\equiv 1 \bmod n$ *and* $a^{m2^i} \not\equiv -1$ *for* $0 \le i < s$ *with* $1 \le a < n$ *then* n *is composite.*

Proof. Suppose n is prime. Then $[a] \in (\mathbb{Z}/n\mathbb{Z})^\times$ has order dividing $\phi(n) = n-1 = 2^r t$ and $[a^t]$ has order dividing 2^r, since $(a^t)^{2^s} = a^{n-1} \equiv 1 \bmod n$. For each $0 \le i < s$ the order of $[a^{t2^i}]$ divides 2^{r-i}, and if $a^t \not\equiv 1 \bmod n$ then $[a^{t2^i}]$ has order 2 for some $0 \le i < s$. The residue class $[-1]$ clearly has order 2 (note that $n \ge 3$), and moreover, it is the unique element of $(\mathbb{Z}/n\mathbb{Z})^\times$ of order 2 because for prime n the unit group $(\mathbb{Z}/n\mathbb{Z})^\times$ is cyclic (see the exercises). It follows that $a^{t2^i} \equiv -1 \bmod n$ for some $0 \le i < s$. □

An integer a that satisfies the properties of the Proposition is called a *Miller-Rabin witness*. As with our earlier notion of a witness we can efficiently determine whether or not a given integer a is a witness or not using the Euclidean algorithm and binary exponentiation in $O((\log n)^{2+\epsilon})$ time. The key difference is that while not every composite number has a witness, every odd composite number has a Miller-Rabin witness, in fact it has many.

Theorem 8.7. *Let n be an odd composite number. The probability that a random integer $a \in [1, n-1]$ is Miller-Rabin witness for n is at least $3/4$.*

The theorem implies that if n is composite and we pick, say, 100 random integers $a \in [1, n-1]$, then we are almost certainly going to find a witness for n. On the other hand, if n is prime then Proposition 8.6 guarantees that we will never find a witness. If we don't find a witness among our 100 randomly chosen a we don't know for sure that n is prime, but we can be very confident that this is the case.

Proof.[1] Let n be an odd composite number of the form $n = 2^s t + 1$ with t odd, and let $n = q_1 \cdots q_r$ be the factorization of n into powers of distinct primes. Let us consider a particular $a \in [1, n-1]$, and let $b = a^t$. If a is not a witness then either $b \equiv 1 \bmod n$, in which case $b \equiv 1 \bmod q_j$ for all of the q_j, or $b^{2^i} \equiv -1 \bmod n$ for some $0 \le i < s$, in which case $b^{2^i} \equiv -1 \bmod q_j$ for all of the q_j. In either case, $b \bmod q_j$ is an element of the 2-Sylow subgroup[2] of $(\mathbb{Z}/q_j\mathbb{Z})^\times$ for all q_j, and it has order 2^{i+1} in $(\mathbb{Z}/q_j\mathbb{Z})^\times$ for each q_j (with $i = -1$ in the first case). We will show that the probability of this happening for a random choice of a is at most $1/4$. We consider three cases.

Case 1: n is not square-free. Then some $q_j = p^k$ with $k > 1$. Since p is odd, the group $(\mathbb{Z}/p^k\mathbb{Z})^\times$ is cyclic of order $\phi(p^k) = p^{k-1}(p-1)$. The prime p cannot divide t, since t divides $n-1$ and p divides n. It follows that if a has order divisible by p in $(\mathbb{Z}/p^k\mathbb{Z})^\times$, then so does $b = a^t$, and in this case b cannot be an element of the 2-Sylow subgroup of $(\mathbb{Z}/p^k\mathbb{Z})^\times$. Thus the probability that b lies in the 2-Sylow subgroup of $(\mathbb{Z}/p^k\mathbb{Z})^\times$ is at most $1/p^{k-1}$, which is less than $1/4$ for all $p^k > 9$. For $p^k = 9$ one checks that for $t = \pm \bmod 6$, at most $2/9 < 1/4$ of the possible values of $a \bmod 9$ yield $b = a^t = \pm 1 \bmod 9$ in the 2-Sylow subgroup of $(\mathbb{Z}/9\mathbb{Z})^\times$.

Case 2: $r \ge 3$. For each q_j the 2-Sylow subgroup G_j of $(\mathbb{Z}/q_j\mathbb{Z})^\times$ is a cyclic of order 2^{k_j}, for some $k_j > 1$, and at most half the elements in G_j have any particular order. Assuming $b = a^t \bmod q_j$ lies in G_j for $j = 1, 2, 3$, the probability that it has the same order in each case is at most $1/4$.[3]

Case 3: $n = pq$ for distinct primes p and q. We may write $p = 2^{s_p} t_p + 1$, and $q = 2^{s_q} t_q + 1$, with t_p and t_q odd. Define the random variable X_p to be -1 if $b \bmod p$ does not lie in the 2-Sylow subgroup G_p of $(\mathbb{Z}/p\mathbb{Z})^\times$, and otherwise let $X_p = i$, where $b \bmod p$ has order 2^i in G_p. Similarly define the random variable X_q. We wish to show that $\Pr[X_p = X_q \ge 0] \le 1/4$.

[1] The proof we give here is a bit different (and more elementary) than the proofs of Monier and Rabin.

[2] A p-Sylow subgroup of a group G is a subgroup whose order is the largest power of p dividing $\#G$.

[3] This rules out all Carmichael numbers, since they all have at least 3 distinct prime factors.

We first suppose $s_p > s_q$. Half the elements in the 2-Sylow subgroup of $(\mathbb{Z}/p\mathbb{Z})^\times$) have order $2^{s_p} > 2^{s_q}$, so $\Pr[0 \le X_p \le s_q] \le 1/2$. We also have $\Pr[X_q = X_p | 0 \le X_p \le s_q] \le 1/2$, thus $\Pr[X_p = X_q \ge 0] \le 1/4$.

We now suppose that $s_p = s_q$. We have

$$2^s t = n - 1 = pq - 1 = (p-1)(q-1) + (p-1) + (q-1) = 2^s t_p t_q + 2^{s_p} t_p + 2^{s_p} t_q,$$

so if t_p divides t then it divides t_q and conversely. It follows that t_p and t_q cannot both divide t (if this were true then we would have $t_p = t_q$ and $p = q$ but we assumed $p \ne q$). So we may assume without loss of generality that t_p does not divide t. This means $t_p \ne 1$, so t_p is divisible by an odd prime $\ell \ge 3$ that does not divide t. It follows that $\Pr[X_p \ge 0] \le 1/3$, and we also have $\Pr[X_q = X_p | X_p \ge 0] \le 1/2$, hence $\Pr[X_p = X_q \ge 0] \le 1/6 < 1/4$. □

Theorem 8.7 yields the following probabilistic primality test, due to Gary Miller [11] and Michael Rabin [14].

Algorithm 8.8 (Miller-Rabin)**.** Given an odd integer $n \ge 3$:

```
1. Pick a random integer $a \in [1, n-1]$.
2. Write $n = 2^s t + 1$, with $t$ odd, and compute $x = a^t \bmod n$.
   If $x \equiv \pm 1 \bmod m$, return true ($a$ is not a witness, $n$ could be prime).
3. For $i$ from 1 to $s-1$:
    a. Set $x \leftarrow x^2 \bmod n$.
    b. If $x \equiv -1 \bmod n$, return true ($a$ is not a witness, $n$ could be prime).
4. Return false ($a$ is a witness, $n$ is definitely composite).
```

Example 8.9. For $n = 561$, $a = 2$: $561 = 2^4 \cdot 35 + 1$. We find that

$$2^{35} \equiv 263 \bmod 561$$

is not $\pm 1 \bmod 561$ so we continue:

$$236^2 \equiv 166 \bmod 561$$
$$166^2 \equiv 67 \bmod 561$$
$$67^2 \equiv 1 \bmod 561$$

None of these values is congruent to -1, so $a = 2$ is a witness for $n = 561$ and we return **false**, meaning that 561 is definitely not a prime.

8.3 A deterministic primality test under GRH

We can make Algorithm 8.8 deterministic by having it test every integer a in $[1, B]$ for some bound B, rather than picking a random integer in $[1, n]$. Proposition 8.6 implies that $B = n/4$ would work, but this is far larger than necessary. Indeed, under a generalized form of the Riemann hypothesis (GRH) it is known the set of integers $a \in [1, 2(\log n)^2]$ is guaranteed to contain a complete set of generators for the group $(\mathbb{Z}/n\mathbb{Z})^\times$, for any positive integer n, prime or otherwise [3]. One can show that the elements $[a] \in (\mathbb{Z}/n\mathbb{Z})^\times$ for which the Miller-Rabin algorithm outputs **true** (the non-witnesses) are contained in a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^\times$; it follows that if n is composite then any set of generators for $(\mathbb{Z}/n\mathbb{Z})^\times$ contains at least one witness. This yields a deterministic algorithm that runs in $O(\log^{4+\epsilon} n))$ time, but its correctness depends on a conjecture that has yet to be proved.

8.4 A deterministic primality test

We now come to the primality test of Agrawal, Kayal, and Saxena (AKS), which is a deterministic algorithm that runs in polynomial time. It is based on the following lemma.

Lemma 8.10. *Let $n \geq 2$ be an integer and suppose* $\gcd(a, n) = 1$. *Then n is prime if and only if*

$$(X + a)^n \equiv X^n + a \bmod n$$

Here X is a formal variable and we are working in the quotient of the polynomial ring $\mathbb{Z}[X]$ modulo the ideal $n\mathbb{Z}[X]$, which is equivalent to the polynomial ring $(\mathbb{Z}/n\mathbb{Z})[X]$.

Proof. By the binomial theorem, we have

$$(X + a)^n = X^n + \binom{n}{1} X^{n-1}a + \cdots + \binom{n}{n-1} X a^{n-1} + a^n.$$

If n is prime then $\binom{n}{k} \equiv 0 \bmod n$ for $0 < k < n$ and the desired congruence holds. Otherwise n is divisible by some prime p with $1 < p < n$. The larges power p^e of p dividing n does not divide $\binom{n}{p}$, nor does it divide a^{n-p}, since $\gcd(a, n) = 1$. The coefficient of X^p is therefore nonzero modulo p^e, hence nonzero modulo n. The lemma follows. □

The good thing about Lemma 8.10 is that every a coprime to n works as a witness that can determine if n is composite, but this does not give us a practical algorithm because the polynomial $(X + a)^n$ has far too many terms for us to compute them all. However we can compute $(X + a)^n$ efficiently using binary exponentiation if we work modulo an ideal of the form $(X^r - 1, n)$; this means that not only reduce all coefficients modulo n, we can replace any monomial X^k with $X^{k \bmod r}$, which means we are always working with polynomials of degree less than r.

If n is prime then we have

$$(X + a)^n \equiv X^n + a \bmod (X^r - 1, n)$$

for every $a, r \geq 1$. The key to the AKS primality test is to show that if n is composite then there is a pair (a, r) with both a and r bounded by a polynomial in $\log n$ for which the above identity will fail to hold.

Let us fist give the algorithm, which is very simple. As a matter of notation, for any integer x let $\operatorname{ord}_r(x)$ to denote the order of $[x]$ in the group $(\mathbb{Z}/r\mathbb{Z})^\times$.

Algorithm 8.11 (Agrawal-Kayal-Saxena)**.** Given an integer $n \geq 3$:

```
1. If $n = m^k$ for some $m, k > 1$ then output **false** ($n$ is composite).
2. Find the smallest $r \geq 1$ such that $\operatorname{ord}_r(n) > \lg^2 n$.
3. If $1 < \gcd(a, n) < n$ for some $a \in [1, r]$ then output **false** ($n$ is composite).
4. If $n \leq r$ then output **true** ($n$ is prime).
5. For $a = 1$ to $\lfloor \sqrt{r} \lg n \rfloor$ do
     a. If $(X + a)^n \not\equiv X^n + a \bmod (X^r - 1, n)$ then output **false** ($n$ is composite).
6. Output **true** ($n$ is prime).
```

It is easy to see that if the input n is prime then Algorithm 8.11 will output **true**: it cannot output **false** in steps 1, 3, or 5 so it must output **true**, either in step 4 or in step 6.

8.5 Correctness of the AKS algorithm

To prove the converse, that if Algorithm 8.11 outputs **true** then its input n is prime, requires more work. Let us fix the input n and assume that the output is **true**. If $n \le r$ then the algorithm must have terminated in step 4, in which case n is clearly prime, so we may assume $n > r$, and moreover, that n has no prime factors $p \le r$, since any such factor would have been found in step 3. So let $p > r$ be a prime factor of n, and let us choose p so that $p \not\equiv 1 \bmod r$; such a p exists since $\operatorname{ord}_r(n) > \lg^2 n > 1$. Our goal is to prove that $n = p$, and it suffices to show that n is a power of p, since in this case if $n \ne p$ then the algorithm would have output **false** in step 1.

Let $b := \lfloor \sqrt{r} \lg n \rfloor$. Since we are assuming that $n > r$ and that the algorithm output **true**, it most have reached step 6, and therefore we know that for every integer $a \in [1, b]$ we have

$$(X + a)^n \equiv X^n + a \bmod (X^r - 1, n),$$

and since $p|n$ this implies

$$(X + a)^n \equiv X^n + a \bmod (X^r - 1, p).$$

We also note that $(X + a)^p \equiv X^p + a \bmod p$; indeed for any polynomial $f \in \mathbb{Z}[X]$ we have $f(X)^p \equiv f(X^p) \bmod p$ (this follows from the fact that every multinomial coefficients in the expansion of $f(X)^p$ is divisible by p except those that involve a single term of $f(X)$ raised to the pth power. Corollary 8.3 implies that $a^p \equiv a \bmod p$, thus

$$(X + a)^p \equiv X^p + a \bmod (X^r - 1, p).$$

We also note that

$$((X + a)^{n/p} - (X^{n/p} + a))^p \equiv (X + a)^n - (X^n + a) \equiv 0 \bmod (X^r - 1, p),$$

so

$$(X + a)^{n/p} \equiv X^{n/p} + a \bmod (X^r - 1, p).$$

Definition 8.12. $m \in \mathbb{Z}_{>0}$ is *introspective* for $f \in \mathbb{Z}[X]$ if $f(X)^m \equiv f(X^m) \bmod (X^r - 1, p)$.

The integers $n, p, n/p$ are all introspective for the polynomials $X + a$ for all $a \in [1, b]$.

Lemma 8.13. *Introspectivity is multiplicative in both m and f, that is:*

(i) *if m_1 and m_2 are both introspective for f then so is $m_1 m_2$;*

(ii) *if m is introspective for both f_1 and f_2 then m is introspective for $f_1 f_2$.*

Proof. (i) If m_1 and m_2 are introspective for f then

$$f(X)^{m_1 m_2} \equiv f(X^{m_1})^{m_2} \bmod (X^r - 1, p),$$

since m_1 is introspective for f, and substituting $X^{m_1 - 1}$ for X yields

$$\begin{aligned} f(X^{m_1})^{m_2} &\equiv f(X^{m_1 m_2}) \bmod (X^{m_1 r} - 1, p) \\ &\equiv f(X^{m_1 m_2}) \bmod (X^r - 1, p) \end{aligned}$$

since m_2 is introspective for f and $X^r - 1$ divides $X^{m_1 r} - 1$. So $m_1 m_2$ is introspective for f.

(ii) If m is introspective for f_1 and f_2 then

$$(f_1 f_2)(X)^m \equiv f_1(X)^m f_2(X)^m \equiv f_1(X^m) f_2(X^m) \equiv f_1 f_2(X^m) \bmod (X^r - 1, p),$$

so m is introspective for $f_1 f_2$. □

It follows from Lemma 8.13 that every integer m in the set

$$S := \{(n/p)^i p^j : i, j \geq 0\}$$

is introspective for every polynomial f in the set

$$P := \left\{ \prod_{a=0}^{b} (X+a)^{e_a} : e_a \geq 0 \right\}.$$

We have $\gcd(n,r) = 1$ (checked in step 3), so $\gcd(n/p, r) = 1$, and clearly $\gcd(p,r) = 1$, since $p > r$ is prime. It follows that $\gcd(m,r) = 1$ for all $m \in S$. The set of residues

$$\overline{S} := \{[m] \in \mathbb{Z}/r\mathbb{Z} : m \in S\} \subseteq (\mathbb{Z}/r\mathbb{Z})^\times$$

is a multiplicatively closed subset of the finite group $(\mathbb{Z}/r\mathbb{Z})^\times$, hence a subgroup, and it is generated by the elements $[n]$ and $[p]$. Let $s := \#S$. We have $\operatorname{ord}_r(n) > \lg^2 n$ (by the choice of r in step 2), and $[n]_r \in \overline{S}$, so $s > \lg^2 n$.

We now want to similarly define a finite set of residues $\overline{P}$ for the set P, but to do this we first need to define a suitable finite quotient of the ring $\mathbb{Z}[X]$ in which P lies.

8.5.1 Cyclotomic polynomials and finite field extensions

The roots of the polynomial $X^r - 1$ can be viewed as rth roots of unity (over the complex numbers, its roots $\{e^{2\pi ik/r} : k \in [1,r]\}$ lie on the unit circle). For each divisor d of r the polynomial $X^r - 1$ is divisible by the *cyclotomic polynomial* $\Phi_d(X)$ whose roots ζ are *primitive* dth roots of unity, meaning that $\zeta^d = 1$ but $\zeta^k \neq 1$ for $1 \leq k < d$. In $\mathbb{Z}[X]$ we have the factorization

$$X^r - 1 = \prod_{d|r} \Phi_d(X),$$

in which each cyclotomic polynomial Φ_d is an irreducible polynomial of degree $\phi(d)$. For example,

$$X^4 - 1 = \Phi_1(X)\Phi_2(X)\Phi_4(X) = (X-1)(X+1)(X^2+1).$$

In the ring $(\mathbb{Z}/p\mathbb{Z})[X]$ the polynomial $\Phi_r(X)$ may factor into lower degree polynomials; indeed it is a product of irreducible polynomials of degree $k := \operatorname{ord}_r(p)$ (this follows easily from the Galois theory of finite fields which we won't delve into here). Let $h(X)$ be one of these irreducible factors, and define the quotient ring

$$K := \frac{(\mathbb{Z}/p\mathbb{Z})[X]}{(h(X))}.$$

The fact that $h(X)$ is irreducible in $(\mathbb{Z}/p\mathbb{Z})[X]$ implies that $\gcd(f(X), h(X)) = 1$ for all nonzero polynomials $f(X)$ of degree $\deg f < \deg h$. The set of all such polynomials forms a set of distinct representatives for the nonzero residue classes in K, of which there are $p^k - 1$. The fact that $\gcd(f(X), h(X) = 1$ implies that we can use the extended Euclidean algorithm to compute $s(X)f(X) + t(X)h(X) = 1$ for some $s, t \in (Z/p\mathbb{Z})[X]$, and we then have $[s][f] = 1$ in K; it follows that every nonzero element of K has a multiplicative inverse.

Thus K is a field with p^k elements; up to isomorphism this uniquely determines K. Every $a \in K^\times$ has order dividing $\#K^\times = p^k - 1$ and is therefore a root of the polynomial $Y^{p^k-1} - 1$, which is divisible by $Y^r - 1$, because $p^k \equiv 1 \bmod r$. There are $p^k - 1$ nonzero elements of K, the same as the degree of $Y^{p^k-1}-1$, so every root of $Y^{p^k-1}-1$ lies in K, as does every root of Y^r-1; thus K contains all the rth roots of unity.

In fact everything above remains true if we replace $h(X)$ with any irreducible polynomial in $(\mathbb{Z}/p\mathbb{Z})[X]$ of degree k; the reason to use $h(X)$ is so that the residue class $[X]$ of the polynomial X is a primitive rth root of unity in K, which will be useful in what follows.

8.5.2 Completing the proof

We now return to the task of proving that if Algorithm 8.11 returns **true** then its input n is prime. Let us define the set of residue classes

$$\overline{P} := \{[f] \in K : f \in P\} \subseteq K^\times,$$

Here $[f] \in K$ is obtained by reducing the coefficients of f modulo p and taking the residue class of the result in the quotient ring $K = (\mathbb{Z}/p\mathbb{Z})[X]/(h(X))$ (a finite field). The set $\overline{P}$ is a multiplicatively closed subset of the finite group $K^\times$ (hence a subgroup), generated by the elements $[X], [X+1], \ldots, [X+b]$. Note that $b = \lfloor \sqrt{r} \lg n \rfloor < r < p$, since we can have $\operatorname{ord}_r(n) > \lg^2 n$ only for $r > \lg^2 n$, and $\deg h = \operatorname{ord}_r(p) > 1$ (recall that we chose $p \not\equiv 1 \bmod r$), so these residue classes are all distinct. Recall that $s = \#\overline{S}$, where $\overline{S}$ is the set of reductions modulo r of the set $S = \{(n/p)^i p^j : i, j \geq 0\}$.

Lemma 8.14. $\#\overline{P} \geq \binom{s+b}{s-1}$.

Proof. Let $f, g \in P$ be distinct polynomials of degree less than s. We claim that the residue classes $[f], [g] \in \overline{P}$ are distinct elements of K.

Suppose not. Then $[f] = [g]$ and $[f]^m = [g]^m$ for any integer $m \geq 1$, and for any $m \in S$ we have $[f(X^m)] = [g(X^m)]$, since every $m \in S$ is introspective for all $f, g \in P$. Let $\tilde{f} \in K[Y]$ be the polynomial obtained by reducing the coefficients of f modulo p and viewing them as elements of K, and similarly define $\tilde{g}$. We then have $\tilde{f}([X^m]) = \tilde{g}([X^m])$, and $[X^m] \in K$ is thus a root of the polynomial $Q(Y) := \tilde{f}(Y) - \tilde{g}(Y) \in K[Y]$.

Recall that $h(X)$ divides $X^r - 1$, and it thus also divides $X^{mr} - 1$, so $[X^m]^r = 1$ in K and $[X^m]$ is thus an rth root of unity. Now $[X]$ is a primitive rth root of unity, and $\gcd(m, r) = 1$, so $[X^m]$ is also a primitive rth root unity. For $m_1, m_2 \in S$ we thus have $[X^{m_1}] = [X^{m_2}]$ if and only if $m_1 \equiv m_2 \bmod r$. For each $[m] \in \overline{S} \subseteq (\mathbb{Z}/r\mathbb{Z})^\times$ we get a distinct root $[X^m]$ of $Q(Y)$, so $Q(Y)$ has at least $s = \#\overline{S}$ distinct roots, but $\deg Q < s$, a contradiction. So our supposition that $[f] = [g]$ must be false and $[f]$ and $[g]$ are distinct elements of K as claimed.

By multiplying up to $s-1$ not necessarily distinct elements of $\{X, X+1, \ldots, X+b\} \subseteq P$ we can form a total of $\binom{s-1+b+1}{s-1} = \binom{s+b}{s-1}$ distinct polynomials $f \in P$ of degree less than s, each of which represents a distinct residue class in $\overline{P}$.[4] Therefore $\#\overline{P} \geq \binom{s+b}{s-1}$. □

Lemma 8.14 gives us a lower bound on $\#\overline{P}$. We now prove an upper bound on $\#\overline{P}$ that applies whenever n is not a power of p.

Lemma 8.15. *If n is not a power of p then $\#\overline{P} \leq n^{\sqrt{s}}$.*

[4] Recall that the number of distinct ways of choosing up to n balls of m possible colors is $\binom{n+m}{n}$; here we have up to $n = s-1$ balls (linear factors $X + a$ of f) and $m = b+1$ colors (values of $a \in [0, b]$).

Proof. Assume n is not a power of p. Then the integer n/p is divisible by a prime $q \neq p$ and the set $T := \{(n/p)^i p^j : 0 \le i, j \le \sqrt{s}\} \subseteq S$ has cardinality $\#T = (\lfloor\sqrt{s}\rfloor + 1)^2 > s = \#\overline{S}$. It follows that T must contain distinct integers m_1 and m_2 that lie in the same residue class in $\overline{S}$, meaning that $m_1 \equiv m_2 \bmod r$. It follows that $X^{m_1} \equiv X^{m_2} \bmod X^r - 1$. For every $f \in P$ we have

$$f(X)^{m_1} \equiv f(X^{m_1}) \equiv f(X^{m_2}) \equiv f(X)^{m_2} \bmod (X^{r-1}, p),$$

since $m_1, m_2 \in S$ are introspective for every $f \in P$. The fact that $h(X)$ divides X^{r-1} in $(\mathbb{Z}/p\mathbb{Z})[X]$ implies that we have

$$[f(X)]^{m_1} = [f(X)^{m_1}] = [f(X)^{m_2}] = [f(X)]^{m_2}$$

as an identity in $K = (\mathbb{Z}/p\mathbb{Z})[X]/(h(X))$. It follows that every $[f] \in \overline{P}$ is a root of $Q(Y) := Y^{m_1} - Y^{m_2} \in K[Y]$, so $Q(Y)$ has at least $\#\overline{P}$ distinct roots in K. We also have

$$\deg Q \le \max(m_1, m_2) \le (n/p)^{\sqrt{s}} p^{\sqrt{s}} = n^{\sqrt{s}},$$

since $m_1, m_2 \in T$, and this implies $\#\overline{P} \le n^{\sqrt{s}}$ as claimed. □

We can now prove the correctness of Algorithm 8.11.

Theorem 8.16 (Agrawal-Kayal-Saxena)**.** *Algorithm 8.11 outputs* ***true*** *if and only if its input n is prime.*

Proof. As we observed earlier, if n is prime the algorithm cannot output **false** and must output **true**. For the converse, either the algorithm output **true** in step 4, in which case $n \le r$ is prime and the theorem holds, or n is divisible by a prime $p > r$, and we can assume that $p \not\equiv 1 \bmod r$, since r was chosen in step 2 so that $\operatorname{ord}_r(n) > \lg^2 n > 1$.

Let $\overline{S} \subseteq (\mathbb{Z}/r\mathbb{Z})^\times$ and $\overline{P} \subset K^\times$ be defined as above, and let

$$s := \#\overline{S}, \qquad b = \lfloor\sqrt{r}\lg n\rfloor, \qquad c := \lfloor\sqrt{s}\lg n\rfloor.$$

We have $s < r$, so $b > c$, and $s \ge \operatorname{ord}_r(n) > \lg^2 n$, so $s > c$. We also have $c > 1$ since $\sqrt{s} > \lg n > 1$. Applying Lemma 8.14 and these inequalities yields

$$\#\overline{P} \ge \binom{s+b}{s-1} \ge \binom{2c+1}{c} > 2^{c+1} = 2^{\lfloor\sqrt{s}\lg n\rfloor+1} \ge 2^{\sqrt{s}\lg n} = n^{\sqrt{s}},$$

and Lemma 8.15 then implies that n must be a power of p. But n cannot be a power of p, since this would have caused the algorithm to output **false** in step 1, so $n = p$ is prime. □

8.6 Complexity analysis of the AKS algorithm

In order to analyze the complexity of Algorithm 8.11 we first need to bound the value of r chosen by the algorithm in step 2.

Lemma 8.17. *For every integer $n \ge 3$ there is an $r \le \lg^5 n$ such that $\operatorname{ord}_r(n) > \lg^2 n$.*

Proof. If $n = 3$ then $r = 5$ works, so we assume $n \ge 4$. Let $m := \lfloor\lg^2 n\rfloor \ge 4$ and let q be the least prime that does not divide

$$N := n(n-1)(n^2-1)(n^3-1)\cdots(n^m-1).$$

The order of n in $(\mathbb{Z}/q\mathbb{Z})^\times$ must be greater than m and therefore greater than $\lg^2 n$. It follows that $r \le q$. The integer N is bounded by

$$N < n^{1+1+2+3+\cdots m} = n^{m(m+1)/2+1} \le n^{m^2} \le n^{\lg^4 n} = 2^{\lg^5 n}$$

As you will prove in the exercises below, for all $x \ge 31$ we have

$$\prod_{p\le x} p > 2^x.$$

Applying this to $x = \lg^5 n \ge 32$, we see that the product of the primes less than x is greater than $2^{\lg^5 n} > N$, but the product of the primes dividing N cannot exceed N, so some prime $q < \lg^5 n$ must not divide n. The lemma follows. □

Theorem 8.18 (Agrawal-Kayal-Saxena)**.** *The time complexity of Algorithm 8.11 is bounded by $O(\log^{10.5+\epsilon} n)$, for any $\epsilon > 0$.*

Proof. We assume throughout that fast (FFT-based) arithmetic is used, which implies that all arithmetic operations on operands of size $\log n$ take $O(\log^{1+\epsilon} n)$ time (bit operations). We now analyze the complexity of each step of the algorithm.

1. By using Newton's method to compute an approximate kth root of n for $k = 2, 3, \ldots \lg n$ we can complete step 1 in $O(\log^{2+\epsilon} n)$ time (or even $O(\log^{1+\epsilon} n)$ time if we use [4]).
2. By Lemma 8.17 it suffices to check each $r \in [1, \lg^5 n]$ in step 2. Using the sieve of Eratosthenes we can assume compute the factorization of every value of r in this interval in $O(\log^{5+\epsilon} n)$ time. For each r we can us this to compute $\phi(r) = \#(\mathbb{Z}/r\mathbb{Z})^\times$, and it then takes $O(\log^{1+\epsilon} n)$ time to compute $\operatorname{ord}_r(n)$ (once we compute $n \bmod r$, after that we are working with $O(\log\log n)$-bit integers and the costs are negligible). The total time for step 2 is thus $O(\log^{5+\epsilon} n)$.
3. Using the fast Euclidean algorithm we can compute $\gcd(a, n)$ for each $a \in [1, r]$ in $O(\log^{1+\epsilon} n)$ time; the total time for step 3 is then $O(\log^{5+\epsilon} n)$.
4. Step 4 compares n and r and takes negligible time.
5. In step 5 we perform $\sqrt{r}\lg n$ or $O(\log^{3.5} n)$ iterations. In each iteration we need to compute $(X + a)^n \bmod (X^{r-1}, n)$, which we can do using binary exponentiation using polynomials of degree less than r whose coefficients are reduced modulo n. Using FFT-based techniques, the time to multiply two integer polynomials of degree r with coefficients bounded by n is $O((r \log n)^{1+\epsilon}) = O(\log^{6+\epsilon} n)$ time; each exponentiation in step 5 requires $O(\log n)$ polyn omial multiplications and thus takes $O(\log^{7+\epsilon} n)$ time. Summing over all iterations yields a total time of $O(\log^{10.5+\epsilon} n)$ for step 5.
6. Step 6 simply returns **true**, which takes no time.

Step 5 dominates the complexity and the theorem follows. □

In fact one expects that the value of r chosen in step 2 of the algorithm can be bounded by $O(\log^2 n)$, which would bring the time complexity down to $O(\log^{6+\epsilon} n)$. In [1] the authors note that a result of Fouvry [8] yields an $O(\log^3 n)$ bound on r, yielding a total complexity of $O(\log^{7.5+\epsilon} n)$. More recent work by Lenstra and Pomerance [10] achieves an $O(\log^{6+\epsilon} n)$ complexity bound not by sharpening the bound on r but by generalizing the algorithm to incorporate alternative parameters that they can bound more tightly.

In [5], Bernstein presents a randomized version of the AKS algorithm that yields an $O(\log^{4+\epsilon} n)$ expected running time. The Miller-Rabin primality test is an example of a *Monte Carlo algorithm*: its running time is bounded but its output might not be correct. By contrast, Bernstein's algorithm is a *Las Vegas algorithm*: its output is always correct but its running time depends on random choices and may be arbitrarily large, but has a bounded expectation.

Finally, we should mention that in practice when one wants to unequivocally prove that an integer is prime (as opposed to running the Miller-Rabin repeatedly in order to gain a high degree of confidence), the preferred algorithm is not the AKS algorithm, or even a randomized version of it. Instead one uses a fast version of the elliptic curve primality proving method [9, 13]. This algorithm is heuristically (but not provably) expected to run in $O(\log^{4+\epsilon} n)$ time, with much better constant factors than a randomized version of AKS, and it produces a *certificate* of primality that can be verified in $O(\log^{3+\epsilon} n)$ time. Essentially every integer with more than 1000 digits that is known to be prime and not of a special form (e.g. a Mersenne prime or generalized Fermat prime) has been found using this method (the current record, as of November 2016, is 34,093 digits). You can learn more about elliptic curve primality proving by taking 18.783.

8.7 Exercises

The prime number theorem implies that $\prod_{p<x} p$ approaches e^x as $x \to \infty$ (as usual, the symbol p is implicitly assumed to take only prime values). In the exercises below you will prove bounds on $\prod_{p\le x} p$ that are weaker than this but sufficient to prove Lemma 8.17. These are adapted from Exercises 1.25–1.28 in [6], which you should feel free to consult.

(1) Show that for any integer n we have

$$\prod_{n+1<p<2n+1} p \le \binom{2n+1}{n} \le 4^n,$$

and use this to prove that $\prod_{p\le x} p < 4^x$ for all $x > 0$.

(2) For any positive integer N let

$$C(N) := \frac{(6N)!N!}{(3N)!(2N)!(2N)!}.$$

(a) Show that $C(N)$ is an integer.

(b) Show that if $p > (6N)^{1/k}$ is prime then p^k does not divide $C(N)$.

(c) Using (1), prove that

$$\prod_{p\le 6N} p > C(N)/4^{(6N)^{1/2}+(6N)^{1/3}\lg(3N/2)}.$$

(d) Stirling's formula yields the bounds $\sqrt{2\pi}n^{n+1/2}e^{-n} \le n! \le en^{n+1/2}e^{-n}$, valid for all positive integers n. Use this to prove that $C(N) = 108^N/(4\sqrt{N})$ for all N.

(e) Show that $\prod_{p\le x} p > 2^x$ for all $x > 2^{12}$.

(f) Use a computer to verify that the bound in (e) actually holds for all $x \ge 31$.

References

[1] M. Agrawal, N. Kayal, and N. Saxena, *PRIMES is in P*, Ann. of Math. **160** (2004), 781–793.

[2] W.R. Alford, A. Granville, and C. Pomerance, *There are infinitely many Carmichael numbers*, Annals of Mathematics **140** (1994), 703–722.

[3] E. Bach, *Explicit bounds for primality testing and related problems*, Math. Comp. **55** (191), 355–380.

[4] D.J. Bernstion, *Detecting perfect powers in essentially linear time*, Mathematics of Computation **67** (1998), 1253–1283.

[5] D.J. Bernstion, *Proving primality in essentially quartic random time*, Mathematics of Computation **76** (2007), 389–403.

[6] R. Crandall and C. Pomerance, *Prime numbers: A computational perspective*, 2nd edition, Springer, 2005.

[7] Carl Friedrich Gauss, *Disquisitiones Arithmeticae*, translated by Arthur A. Clarke, Revised by William C. Waterhouse, Cornelius Greither, and A.W. Grootendorst, Springer-Verlag, 1986.

[8] E. Fouvry, *Theoreme de Brun-Titchmarsh; application au theoreme de Fermat*, Invent. Math. **79** (1985), 383–407.

[9] S. Goldwasser and J. Kilian, *Almost all primes can be quickly certified*, Proceedings of the Eighteenth ACM Symposium on the Theory of Computing (1986), 316–329.

[10] H. Lenstra and C. Pomerance, *Primality testing with Gaussian periods*, preprint, 2015.

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

[12] L. Monier, *Evaluation and comparison of two efficient probabilistic primality testing algorithms*, Theoretical Computer Science **12** (1980), 97–108.

[13] F. Morain, *Implementing the asymptotically fast version of the elliptic curve primality proving algorithm*, Mathematics of Computation **76** (2007), 493–505.

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

[15] Lasse Rempe-Gillen and Rebecca Waldecker, *Primality testing for beginners*, American Mathematical Society, 2014.