4. Primality testing

4.1. **Introduction.** Factorisation is concerned with the problem of developing efficient algorithms to express a given positive integer $n > 1$ as a product of powers of distinct primes. With primality testing, however, the goal is more modest: given n, decide whether or not it is prime. If n does turn out to be prime, then of course you've (trivially) factorised it, but if you show that it is not prime (i.e., *composite*), then in general you have learnt nothing about its factorisation (apart from the fact that it's not a prime!).

One way of testing a number n for primality is the following: suppose a certain theorem, Theorem X say, whose statement depends on a number n, is true when n is prime. Then if Theorem X is false for a particular n, then n cannot be prime. For instance, we know (Fermat) that $a^{n-1} \equiv 1 \pmod{n}$ when n is prime and $n \nmid a$. So if for such an a we have $a^{n-1} \not\equiv 1 \pmod{n}$, then n is not prime. This test is called the *Pseudoprime Test to base* a. Moreover, a composite number n that passes this test is called a *Pseudoprime to base* a.

(It would be good if we could find a Theorem Y that was true *iff* n was prime, and was moreover easy to test. Then we would know that if the theorem was true for n then n was prime. A result of this type is the following (also on a problem sheet): n is prime iff $a^{n-1} \equiv 1 \pmod{n}$ for $a = 1, 2, \ldots, n-1$. This is, however, not easy to test; it is certainly no easier than testing whether n is divisible by a for $a = 1, \ldots, n$.)

4.2. **Proving primality of n when $n-1$ can be factored.** In general, primality tests can only tell you that a number n either 'is composite', or 'can't tell'. They cannot confirm that n is prime. However, under the special circumstance that we can factor $n-1$, primality can be proved:

Theorem 4.1 (Lucas Test, as strengthened by Kraitchik and Lehmer)**.** *Let $n > 1$ have the property that for every prime factor q of $n-1$ there is an integer a such that $a^{n-1} \equiv 1 \pmod{n}$ but $a^{(n-1)/q} \not\equiv 1 \pmod{n}$. Then n is prime.*

Proof. Define the subgroup G of $(\mathbb{Z}/n\mathbb{Z})^\times$ to be the subgroup generated by all such a's. Clearly the exponent of G is a divisor of $n-1$. But it can't be a proper divisor of $n-1$, for then it would divide some $(n-1)/q$ say, which is impossible as $a^{(n-1)/q} \not\equiv 1 \pmod{n}$ for the a corresponding to that q. Hence G has exponent $n-1$. But then $n - 1 \le \#G \le \#(\mathbb{Z}/n\mathbb{Z})^\times = \varphi(n)$. Hence $\varphi(n) = n-1$, which immediately implies that n is prime. $\square$

Corollary 4.2 (Pepin's Test, 1877)**.** *Let $F_k = 2^{2^k} + 1$, the kth Fermat number, where $k \ge 1$. Then F_k is prime iff $3^{\frac{F_k-1}{2}} \equiv -1 \pmod{F_k}$.*

Proof. First suppose that $3^{\frac{F_k-1}{2}} \equiv -1 \pmod{F_k}$. We apply the theorem with $n = F_k$. So $n - 1 = 2^{2^k}$ and $q = 2$ only, with $a = 3$. Then $3^{\frac{F_k-1}{2}} \not\equiv 1 \pmod{F_k}$ and (on squaring) $3^{F_k-1} \equiv 1 \pmod{F_k}$, so all the conditions of the Theorem are satisfied.

Conversely, suppose that F_k is prime. Then, by Euler's criterion and quadratic reciprocity (see Chapter 5) we have

$$3^{\frac{F_k-1}{2}} \equiv \left(\frac{3}{F_k}\right) = \left(\frac{F_k}{3}\right) = \left(\frac{2}{3}\right) = -1,$$

as 2 is not a square $\pmod 3$.

□

We can use this to show that $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$ and $F_4 = 65537$ are all prime. It is known that F_k is composite for $5 \leq k \leq 32$, although complete factorisations of F_k are known only for $0 \leq k \leq 11$, and there are no known factors of F_k for $k = 20$ or 24. Heuristics suggest that there may be no more k's for which F_k is prime.

4.3. **Carmichael numbers.** A *Carmichael number* is a (composite) number n that is a pseudoprime to every base a with $1 \leq a \leq n$ and $\gcd(a, n) = 1$. Since it it immediate that $a^{n-1} \not\equiv 1 \pmod n$ when $\gcd(a, n) > 1$, we see that Carmichael numbers are pseudoprimes to as many possible bases as any composite number could be. They are named after the US mathematician Robert Carmichael (1879 – 1967).

[But even *finding* an a with $\gcd(a, n) > 1$ gives you a factor of n. (Imagine that n is around 10^{300} and is a product of three 100-digit primes – such a's are going to be few and far between!)]

For examples of Carmichael numbers, see problem sheet 3.

4.4. **Strong pseudoprimes.** Given $n > 1$ odd and an a such that $a^{n-1} \equiv 1 \pmod n$, factorise $n - 1$ as $n - 1 = 2^f q$, where q is odd, $f \geq 1$ and consider the sequence

$$\mathcal{S} = [a^q, a^{2q}, a^{4q}, \ldots, a^{2^f q} \equiv 1],$$

taken $\pmod n$. If n is prime then, working left to right, either $a^q \equiv 1 \pmod n$, in which case $\mathcal{S}$ consists entirely of 1's, or the number before the first 1 must be -1. This is because the number following any x in the sequence is x^2, so if $x^2 \equiv 1 \pmod n$ for n prime, then $x \equiv \pm 1 \pmod n$. (Why?) A composite number n that has this property, (i.e., is a pseudoprime to base a and for which either $\mathcal{S}$ consists entirely of 1's or the number before the first 1 in $\mathcal{S}$ is -1) is called a *strong pseudoprime to base a*.

Clearly, if n is a prime or pseudoprime but not a strong pseudoprime, then this stronger test proves that n isn't prime. This is called the *Miller-Rabin Strong Pseudoprime Test.*

Perhaps surprisingly:

Theorem 4.3. *If n is a pseudoprime to base a but not a strong pseudoprime to base a, with say $a^{2^t q} \equiv 1 \pmod n$ but $a^{2^{t-1} q} \not\equiv \pm 1 \pmod n$, then n factors nontrivially as $n = g_1 g_2$, where $g_1 = \gcd(a^{2^{t-1} q} - 1, n)$ and $g_2 = \gcd(a^{2^{t-1} q} + 1, n)$.*

Proof. For then we have, for $n - 1 = 2^f q$ and some $t \leq f$, that $a^{2^t q} \equiv 1 \pmod n$ but $a^{2^{t-1} q} \not\equiv \pm 1 \pmod n$. Now $a^{2^t q} - 1 = AB \equiv 0 \pmod n$, where $A = (a^{2^{t-1} q} - 1)$ and

$B = (a^{2^{t-1}q} + 1)$, and neither A nor B is divisible by n. Hence g_1 is a nontrivial ($\neq 1$ or n) factor of n. Since $g_1 \mid n$, we have

$$\gcd(g_1, g_2) = \gcd(n, g_1, g_2) = \gcd(n, g_1, g_2 - g_1) = \gcd(n, g_1, 2) = 1,$$

the last step because n is odd. Hence any prime dividing n can divide at most one of g_1 and g_2. So from $n = \prod_p p^{e_p}$, say, and $n \mid AB$, we see that each prime power p^{e_p} dividing n divides precisely one of A or B, and so divides precisely one of g_1 or g_2. Hence $g_1 g_2 = n$. □

Example. Take $n = 31621$, a pseudoprime to base $a = 2$. We have $n - 1 = 2^2 \cdot 7905$, $2^{7905} \equiv 31313 \pmod{n}$ and $2^{15810} \equiv 2^{31620} \equiv 1 \pmod{n}$, so n is not a strong pseudoprime to base 2. Then $g_1 = \gcd(n, 31312) = 103$ and $g_2 = \gcd(n, 31314) = 307$, giving $n = 103 \cdot 307$.

Note that if $n = n_1 n_2$ where n_1 and n_2 are coprime integers, then by the Chinese Remainder Theorem we can solve each of the four sets of equations

$$x \equiv \pm 1 \pmod{n_1} \qquad\qquad x \equiv \pm 1 \pmod{n_2}$$

to get four distinct solutions of $x^2 \equiv 1 \pmod{n}$. For instance, for $n = 35$ get $x = \pm 1$ or ± 6. For the example $n = 31621$ above, we have $31313 \equiv 1 \pmod{103}$ and $31313 \equiv -1 \pmod{307}$, so that four distinct solutions of $x^2 \equiv 1 \pmod{31621}$ are ± 1 and ± 31313.

So what is happening when the strong pseudoprime test detects n as being composite is that some $x \in \mathcal{S}$ is a solution to $x^2 \equiv 1 \pmod{n}$ with $x \not\equiv \pm 1 \pmod{n}$ because $x \equiv 1 \pmod{n_1}$ and $x \equiv -1 \pmod{n_2}$ for some coprime n_1, n_2 with $n_1 n_2 = n$. And then both $\gcd(x - 1, n)$ (divisible by n_1) and $\gcd(x + 1, n)$ (divisible by n_2) are nontrivial factors of n.

4.5. **Strong pseudoprimes to the smallest prime bases.** It is known that

- 2047 is the smallest strong pseudoprime to base 2;
- 1373653 is the smallest strong pseudoprime to both bases 2, 3;
- 25326001 is the smallest strong pseudoprime to all bases 2, 3, 5;
- 3215031751 is the smallest strong pseudoprime to all bases 2, 3, 5, 7;
- 2152302898747 is the smallest strong pseudoprime to all bases 2, 3, 5, 7, 11;
- 3474749660383 is the smallest strong pseudoprime to all bases 2, 3, 5, 7, 11, 13;
- 341550071728321 is the smallest strong pseudoprime to all bases 2, 3, 5, 7, 11, 13, 17.

(In fact 341550071728321 is also a strong pseudoprime to base 19.)

Hence any odd $n < 341550071728321$ that passes the strong pseudoprime test for all bases 2, 3, 5, 7, 11, 13, 17 must be prime. So this provides a cast-iron primality test for all such n.

4.6. **Primality testing in 'polynomial time'.** In 2002 the Indian mathematicians Agrawal, Kayal and Saxena invented an algorithm, based on the study of the polynomial ring $(\mathbb{Z}/n\mathbb{Z})[x]$, that was able to decide whether a given n was prime in time $O((\log n)^{6+\varepsilon})$.

(Here the constant implied by the 'O' depends on ε and so could go to infinity as $\varepsilon \to 0$.) (Search for 'AKS algorithm' on web.)

4.7. **The Lucas-Lehmer primality test for Mersenne numbers.** Given an odd prime p, let $M_p = 2^p - 1$, a *Mersenne number* (and a Mersenne prime iff it is prime). [It is an easy exercise to prove that if p is composite, then so is M_p.]

Define a sequence $S_1, S_2, \ldots, S_n, \ldots$ by $S_1 = 4$ and $S_{n+1} = S_n^2 - 2$ for $n = 1, 2, \ldots$. so we have

$$S_1 = 4, S_2 = 14, S_3 = 194, S_4 = 37634, S_5 = 1416317954, \ldots.$$

There is a very fast test for determining whether or not M_p is prime.

Theorem 4.4 (Lucas-Lehmer Test)**.** *For an odd prime p, the Mersenne number M_p is prime iff M_p divides S_{p-1}.*

So $M_3 = 7$ is prime as $7 \mid S_2$, $M_5 = 31$ is prime as $31 \mid S_4, \ldots$. In this way get M_p prime for $p = 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, 1279, 2203, 2281, \ldots$ (47th) 43112609. There may be others between the 41st and 47th. [as at October 2012.]

For the proof, we need two lemmas.

Lemma 4.5. *Put $\omega = 2 + \sqrt{3}$ and $\omega_1 = 2 - \sqrt{3}$. Then $\omega\omega_1 = 1$ (immediate) and*

$$S_n = \omega^{2^{n-1}} + \omega_1^{2^{n-1}}$$

for $n = 1, 2, \ldots$.

The proof is a very easy induction exercise.

Lemma 4.6. *Let r be a prime $\equiv 1 \pmod 3$ and $\equiv -1 \pmod 8$ (i.e., $\equiv 7 \pmod{24}$). Then*

$$\omega^{\frac{r+1}{2}} \equiv -1 \pmod r.$$

(So it's equal to $a + b\sqrt{3}$ where $a \equiv -1 \pmod r$ and $b \equiv 0 \pmod r$.)

Proof. Put

$$\tau = \frac{1+\sqrt{3}}{\sqrt{2}} \qquad \text{and} \qquad \tau_1 = \frac{1-\sqrt{3}}{\sqrt{2}}.$$

Then we immediately get $\tau\tau_1 = -1$, $\tau^2 = \omega$ and $\tau_1^2 = \omega_1$. Next, from $\tau\sqrt{2} = 1 + \sqrt{3}$ we have $(\tau\sqrt{2})^r = (1+\sqrt{3})^r$, so that

$$\begin{aligned} \tau^r 2^{\frac{r-1}{2}}\sqrt{2} &= 1 + \sum_{j=1}^{r-1} \binom{r}{j} (\sqrt{3})^j + 3^{\frac{r-1}{2}}\sqrt{3} \\ &\equiv 1 + 3^{\frac{r-1}{2}}\sqrt{3} \pmod r, \end{aligned} \tag{1}$$

as $r \mid \binom{r}{j}$. Since $r \equiv -1 \pmod 8$ we have

$$2^{\frac{r-1}{2}} \equiv \left(\frac{2}{r}\right) = (-1)^{\frac{r^2-1}{8}} \equiv 1 \pmod r,$$

using Euler's Criterion, and Prop. 5.3. Further, since $r \equiv 1 \pmod 3$ and $r \equiv -1 \pmod 4$ we have

$$3^{\frac{r-1}{2}} \equiv \left(\frac{3}{r}\right) = \left(\frac{r}{3}\right)(-1)^{\frac{r-1}{2}\cdot\frac{3-1}{2}} = \left(\frac{1}{3}\right)\cdot(-1) \equiv -1 \pmod r,$$

using Euler's Criterion again, and also Quadratic Reciprocity (Th. 5.1). So, from (1), we have successively

$$\begin{aligned}\tau^r\sqrt{2} &\equiv 1-\sqrt{3} \pmod r\\ \tau^r &\equiv \tau_1 \pmod r\\ \tau^{r+1} &\equiv \tau\tau_1 = -1 \pmod r\\ \omega^{\frac{r+1}{2}} &\equiv -1 \pmod r,\end{aligned}$$

the last step using $\tau^2 = \omega$. □

Proof of Theorem 4.4. $\mathbf{M_p}$ **prime** $\Rightarrow \mathbf{M_p \mid S_{p-1}}$. Assume M_p prime. Apply Lemma 4.6 with $r = M_p$, which is allowed as $M_p \equiv -1 \pmod 8$ and $M_p \equiv (-1)^p - 1 \equiv 1 \pmod 3$. So

$$\omega^{\frac{M_p+1}{2}} = \omega^{2^{p-1}} \equiv -1 \pmod{M_p} \tag{2}$$

and, using Lemma 4.5, including $\omega_1^{-1} = \omega$, we have

$$S_{p-1} = \omega^{2^{p-2}} + \omega_1^{2^{p-2}} = \omega_1^{2^{p-2}}\left(\left(\omega_1^{-1}\right)^{2^{p-2}}\omega^{2^{p-2}}+1\right) = \omega_1^{2^{p-2}}\left(\omega^{2^{p-1}}+1\right) \equiv 0 \pmod{M_p}, \tag{3}$$

the last step using (2).

$\mathbf{M_p \mid S_{p-1}} \Rightarrow \mathbf{M_p}$ **prime**. Assume $M_p \mid S_{p-1}$ but M_p composite. We aim for a contradiction. Then M_p will have a prime divisor q (say) with $q^2 \le M_p$.

Now consider the multiplicative group $G = \left(\frac{\mathbb{Z}[\sqrt{3}]}{(q)}\right)^{\times}$ of units of the ring $\frac{\mathbb{Z}[\sqrt{3}]}{(q)}$. Then G has coset representatives consisting of numbers $a + b\sqrt{3}$ with $a, b \in \{0, 1, 2, \ldots, q-1\}$ that are also invertible $\pmod q$. So G is a group of size (order) at most $q^2 - 1$, with multiplication defined modulo q. From $\omega(\omega_1 + q\sqrt{3}) \equiv 1 \pmod q$ we see that $\omega = 2 + \sqrt{3}$ is invertible, and so $\omega \in G$. [Strictly speaking, the coset $\omega \pmod q \in G$.]

Now, using $M_p \mid S_{p-1}$ we see that (3) holds even when M_p is composite, so we have successively that $\omega^{2^{p-1}} + 1 \equiv 0 \pmod{M_p}$, $\omega^{2^{p-1}} \equiv -1 \pmod q$ and $\omega^{2^p} \equiv 1 \pmod q$. Hence the order of ω in G is 2^p. Then $2^p \mid \#G \le q^2 - 1 \le M_p - 1 = 2^p - 2$, a contradiction. Hence M_p must be prime.

□

In practice, to test M_p for primality using Theorem 4.4, one doesn't need to compute $S_j (j = 1, 2, \ldots, p-1)$, but only the much smaller (though still large!) numbers $S_j \pmod{M_p} (j = 1, 2, \ldots, p-1)$.

A good source of information on Mersenne numbers is

http://primes.utm.edu/mersenne/index.html