Appendix 10A. Pseudoprime tests using square roots of 1

In section 7.6 we noted that the converse to Fermat's Little Theorem may be used to give a quick proof that a given integer n is composite: One simply finds an integer a, not divisible by n, for which $a^{n-1} \not\equiv 1 \pmod{n}$ (if this fails, that is if $a^{n-1} \equiv 1 \pmod{n}$ and n is composite, then n is called a *base-a pseudoprime*). Such a search often works quickly, especially for randomly chosen values of n, but can fail if the tested n have some special structure. For example, it always fails for Carmichael numbers, which have the property that n is a base-a pseudoprime for every a with $(a, n) = 1$. What can we do in these cases? Can we construct a test, based on similar ideas, that is guaranteed to recognize even these composite numbers?

10.8. The difficulty of finding all square roots of 1

Lemma 10.1.1 implies that there are *at least* four distinct square roots of $1 \pmod{n}$, for any odd n which is divisible by at least two distinct primes. This suggests that we might try to prove that a given base-a pseudoprime n is composite by finding a square root of $1 \pmod{n}$ which is neither 1 nor -1. (If we can find such a square root of $1 \pmod{n}$ then we can partially factor n, as discussed in section 10.1) The issue then becomes: How do we efficiently search for a square root of 1?

This is not difficult: Since n is a base-a pseudoprime, we have

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

and so $a^{\frac{n-1}{2}} \pmod{n}$ is a square root of $1 \pmod{n}$. By Euler's criterion we know that if p is prime then $a^{\frac{p-1}{2}} \equiv (a/p) \pmod{p}$, so that $a^{\frac{p-1}{2}} \equiv 1$ or $-1 \pmod{p}$. If n is a base-a pseudoprime (and therefore composite), it is feasible that $a^{\frac{n-1}{2}} \not\equiv (a/n) \pmod{n}$, which would imply that n is composite. If $a^{\frac{n-1}{2}} \pmod{n}$ is neither 1 nor

-1 this allows us to factor n into two parts, since

$$n \;=\; \gcd(a^{\frac{n-1}{2}} - 1, n) \;\cdot\; \gcd(a^{\frac{n-1}{2}} + 1, n).$$

If n is composite and $a^{\frac{n-1}{2}} \equiv (a/n) \pmod n$ then we call n a base-a *Euler pseudoprime*.

For example, 1105 is a Carmichael number, and so $2^{1104} \equiv 1 \pmod{1105}$. We take the square root, and determine that $2^{552} \equiv 1 \pmod{1105}$. So this method fails to prove that 1105 is composite, since 1105 is a base-2 Euler pseudoprime. But, wait a minute, 552 is even, so we can take the square root again, and a calculation reveals that $2^{276} \equiv 781 \pmod{1105}$. That is, 781 is a square root of 1 mod 1105, which proves that 1105 is composite. Moreover, since $\gcd(781-1, 1105) = 65$ and $\gcd(781+1, 1105) = 17$, we can even factor 1105 as 65×17.[13]

This property is even more striking mod 1729. In this case $1728 = 2^6 \cdot 27$ so we can take square roots many times. Indeed, taking successive square roots of 2^{1728} we determine that

$$1 \equiv 2^{1728} \equiv 2^{864} \equiv 2^{432} \equiv 2^{216} \pmod{1729}, \text{ but then } 2^{108} \equiv 1065 \pmod{1729}.$$

This proves that 1729 is composite, and even that

$$1729 = \gcd(1064, 1729) \times \gcd(1066, 1729) = 133 \times 13.$$

This protocol of taking successive square roots can fail to identify that our given pseudoprime is indeed composite; for example, we cannot use 103 to prove that either 561 or 1729 is composite, since

$$103^{35} \equiv 1 \pmod{561}, \text{ and so } 103^{70} \equiv \ldots \equiv 103^{560} \equiv 1 \pmod{561},$$
$$103^{27} \equiv -1 \pmod{1729}, \text{ and so } 103^{54} \equiv \ldots \equiv 103^{1728} \equiv 1 \pmod{1729},$$

but such failures are rare (see exercise 10.8.7).

Suppose that n is a composite integer with $n-1 = 2^k m$ for some integer $k \ge 1$ with m odd. We call n a base-a *strong pseudoprime* if the sequence of residues

$$a^{n-1} \pmod n,\ a^{(n-1)/2} \pmod n, \ldots,\ a^{(n-1)/2^k} \pmod n \tag{10.8.1}$$

is equal to either

$$1, 1, \ldots, 1 \quad \text{or} \quad 1, 1, \ldots, 1, -1, *, \ldots, *$$

where the $*$'s stand for any residue mod n. These are the only two possibilities if n is prime, and so if the sequence of residues in (10.8.1) looks like one of these two possibilities then this information does not allow us to deduce that n is composite.

On the other hand, if n is a not a base-a strong pseudoprime then we say that a is a *witness* (to n being composite). To be more precise:

Definition. *Suppose that n is a composite odd integer and $n-1 = 2^k m$ for some integer $k \ge 1$ with m odd. Assume that n is a base-a pseudoprime, that is $a^{n-1} \equiv 1 \pmod n$. If $a^m \equiv 1 \pmod n$ or $a^{m \cdot 2^j} \equiv -1 \pmod n$ for some integer $j \ge 0$, then n is* a base-a strong pseudoprime. *Otherwise a is* a witness *(to the compositeness of n) and if ℓ is the largest integer for which $a^{m \cdot 2^\ell} \not\equiv -1$ or $1 \pmod n$ then $\gcd(a^{m \cdot 2^\ell} - 1, n)$ is a non-trivial factor of n.*

[13] We have not factored 1105 into prime factors (since 65 factors further as $65 = 5 \times 13$), but rather into two non-trivial factors.

One can compute high powers modulo n very rapidly using "fast exponentiation" (a technique we discussed in section 7.13 of appendix 7A), so this strong pseudoprime test can be done quickly and easily.

In exercise 10.8.7 we will show that at least three-quarters of the integers a, $1 \le a \le n$ with $(a, n) = 1$ are witnesses for n, for each odd composite $n > 9$. So can we find a witness quickly if n is composite?

• The most obvious idea is to try $a = 2, 3, 4, \dots$ consecutively until we find a witness. It is believed that there is a witness $\le 2(\log n)^2$, but we cannot prove this (though we can deduce this from a famous conjecture, the Generalized Riemann Hypothesis[14]).

• Pick integers $a_1, a_2, \dots, a_\ell, \dots$ from $\{1, 2, 3, \dots, n-1\}$ at random until we find a witness. By what we wrote above, if n is composite then the probability that none of $a_1, a_2, \dots, a_\ell$ are witnesses for n is $\le 1/4^\ell$. Thus with a hundred or so such tests we get a probability that is so small that it is inconceivable that it could occur in practice; so we believe that any integer n for which none of a hundred randomly chosen a's is a witness, is prime. We call such n "*industrial strength primes*" since they have not been proven to be prime, but there is an enormous weight of evidence that they are not composite.

This test is a *random polynomial time* test for compositeness (like our test for finding a quadratic non-residue given at the end of appendix 8B). If n is composite then the randomized witness test is almost certain to provide a short proof of n's compositeness in 100 runs of the test. On the other hand, if 100 runs of the test do not produce a witness then we can be almost certain that n is prime, but we cannot be *absolutely* certain since no proof is provided, and therefore we have an industrial strength prime.

In practice the witness test accomplishes Gauss's dream of quickly distinguishing between primes and composites, for either we will quickly get a witness to n being composite or, if not, we can be almost certain that our industrial strength prime is indeed prime. Although this solves the problem in practice, we cannot be absolutely certain that we have distinguished correctly when we claim that n is prime since we have no proof, and mathematicians like proof. Indeed if you claim that industrial strength primes are prime, without proof, then a cynic might not believe that your randomly chosen a are so random, or that you are unlucky, or ... No, what we need is a proof that a number is prime when we think that it is.

Exercise 10.8.1. Find all bases b for which 15 is a base-b Euler pseudoprime.

Exercise 10.8.2.† We wish to show that every odd composite n is not a base-b Euler pseudoprime for some integer b, coprime to n. Suppose not, i.e., that n is a base-b Euler pseudoprime for every integer b with $(b, n) = 1$.

(a) Show that n is a Carmichael number.
(b) Show that if prime p divides n then $p - 1$ cannot divide $\frac{n-1}{2}$.
(c) Deduce that $(b/n) \equiv (b/p) \pmod p$ for each prime p dividing n.
(d) Explain why (c) cannot hold for every integer b coprime to n.

[14] We discussed the Riemann Hypothesis, and its generalizations, in section 5.15 of appendix 5D. Suffice to say that this is one of the most famous and difficult open problems of mathematics, so much so that the Clay Mathematics Institute has now offered one million dollars for its resolution (see `http://www.claymath.org/millennium/`).

Exercise 10.8.3. Prove that $F_n = 2^{2^n} + 1$ is either a prime, or a base-2 strong pseudoprime.

Exercise 10.8.4. Prove that if n is a base-2 pseudoprime then $2^n - 1$ is a base-2 strong pseudoprime and a base-2 Euler pseudoprime. Deduce that there are infinitely many base-2 strong pseudoprimes.

Exercise 10.8.5. Pépin showed that one can test Fermat numbers F_m for primality by using just one strong pseudoprime test; i.e., F_m is prime if and only if $3^{(F_m-1)/2} \equiv -1 \pmod{F_m}$.

(a) Use exercise 8.9.22 to show if F_m is prime then $3^{(F_m-1)/2} \equiv -1 \pmod{F_m}$.

(b) In the other direction show that if $3^{(F_m-1)/2} \equiv -1 \pmod{F_m}$ then $\mathrm{ord}_p(3) = 2^{2^m}$ whenever prime $p|F_m$.

(c) Deduce that $F_m - 1 \leq p - 1$ in (b) and so F_m is prime.

Exercise 10.8.6.† (a) Prove that $A := (4^p + 1)/5$ is composite for all primes $p > 3$.

(b) Deduce that A is a base-2 strong pseudoprime.

Exercise 10.8.7.‡ How many witnesses are there mod n? Suppose that $n - 1 = 2^k m$ with m odd and $k \geq 1$, and that n has ω distinct prime factors. Let g_p be the largest odd integer dividing $(p - 1, n - 1)$, and 2^{R+1} be the largest power of 2 dividing $\gcd(p - 1 : p|n)$.

(a) Prove that $R \leq k - 1$.

(b) Show that (10.8.1) is $1, 1, \ldots, 1$ if and only if $a^{g_p} \equiv 1 \pmod{p^e}$ for every prime power $p^e \| n$.

(c) Show that there are $\prod_{p|n} g_p$ such integers $a \pmod{n}$.

(d) Show that if (10.8.1) is $1, 1, \ldots, 1, -1, *, \ldots, *$, with r *'s at the end then $0 \leq r \leq R$, and that this holds if and only if $a^{2^r g_p} \equiv -1 \pmod{p^e}$ for every prime power $p^e \| n$.

(e) Show that there are $\leq \prod_{p|n} 2^r g_p$ such integers $a \pmod{n}$.

(f) Show the number of strong pseudoprimes mod n is

$$\prod_{p|n} (2^R g_p) \cdot \left(1 + \frac{1}{2^\omega} + \frac{1}{2^{2\omega}} + \ldots + \frac{1}{2^{(R-1)\omega}} + \frac{2}{2^{R\omega}}\right).$$

(g) Prove that $2^R g_p \leq \frac{p-1}{2}$ and so deduce that the quantity in (f) is $\leq \frac{\phi(n)}{2^{\omega-1}}$, and so is $< \frac{1}{4}\phi(n)$ if $\omega \geq 3$.

(h) Show that there are $\leq \frac{1}{4}\phi(n)$ reduced residues mod n which are not witnesses, whenever $n \geq 10$ with equality holding if and only if either

- $n = pq$ where $p = 2m + 1, q = 4m + 1$ are primes with m odd, or
- $n = pqr$ is a Carmichael number with p, q, r primes each $\equiv 3 \pmod 4$ (e.g. $7 \cdot 19 \cdot 67$).