NIEUW ARCHIEF VOOR WISKUNDE (4), Vol.1 , (1983), 133-144

FAST PRIME NUMBER TESTS

Text of a lecture held at the Wintersymposium "Elementen van de Wiskunde" of the Wiskundig Genootschap, Zeist, Januari 8, 1983(*)

H.W. Lenstra, Jr.

Throughout this lecture, n denotes an integer larger than 1. We call n *composite* if there exist integers a, b for which

$$n = a \cdot b, \quad a > 1, \quad b > 1.$$

Otherwise n is called *prime*.

Every positive integer can be decomposed into prime factors, and this decomposition is unique up to the order of the factors. It is a basic and difficult problem from elementary number theory, how to find the *prime factor decomposition* of a given integer n in an efficient manner. In this lecture we shall only consider a subproblem, which turns out to be much more tractable: the problem how to *recognize prime numbers*, i.e. how to determine whether a given integer n is prime or not.

Recent advances on this problem have been made by the American mathematicians L.M. Adleman and R.S. Rumely. A substantial simplification of their method was developed by me, in cooperation with H. Cohen. The simplified algorithm was programmed for the CDC Cyber 170-750 computer system of the SARA computing centre in Amsterdam, with the help of D.T. Winter and A.K. Lenstra.

Before I discuss the theory on which the program is based I treat a few numerical examples. Mathematically speaking they are mere curiosities, but they illustrate at least what can be done at the moment.

The number

$$10^{100} + 267 = \underbrace{100 \ldots 00}_{97\times}267$$

(*) An earlier version of this text appeared (in Dutch) in *Wiskunde en Onderwijs*, journal of the Vlaamse Vereniging Wiskunde Leraars.

is the least prime number of 101 digits. Our program proved the primality of this number within 42.8 seconds. This running time is typical for prime numbers of this order of magnitude. Older methods were not able to deal with the number $10^{100} + 267$. In fact, one of the few good reasons we had to program our algorithm was to obtain convincing evidence that is better than earlier algorithms.

If our program is adapted to 200 digit numbers, the typical prime number of that size will require 6 minutes approximately.

The number

$$\frac{10^{1031} - 1}{9} = \underbrace{111 \ldots 111}_{1031 \times}$$

has a *high probability* of being prime; the exact meaning of this will be seen later. We estimate that our method would need roughly one week for this number, on the same machine. It may well be that the number is within reach of future improvements of our algorithm.

For numbers of certain special types one can go much further. For example, D. Slowinski proved with the help of a CRAY-1 computer that the 25962-digit number

$$2^{86243} - 1 = 536 \ldots 207$$

is prime. This is presently the largest prime number known. It must have required several hours of computing time.

The problem to decompose a number into prime factors is much harder. Existing methods can usually deal with numbers of up to 40 or 50 digits in a few hours. No prime factor is known of the 89-digit number

$$2^{293} - 1 = 159 \ldots 791.$$

It is known that this number is composite.

It may seem surprising that a number can be known to be composite without a factor being known. This state of affairs is usually due to the following theorem or one of its variants:

FERMAT'S THEOREM (Pierre de Fermat, 1601-1665).

$$n \text{ is prime} \Rightarrow \forall a \in \mathbb{Z} : a^n \equiv a \bmod n.$$

We remark that, for given a and n, it is easy to check whether the congruence in Fermat's theorem holds, at least if one uses an electronic computer. This is even true if a and n are very large (e.g., $\approx 10^{100}$). To do this, one should *not* start by calculating a^n *itself*: even for $a = 3$, $n \approx 10^{100}$ this number is much too large to be computed on any computer. Instead, one calculates the remainder of a^n upon division by n, which can be done by a series of successive squarings and multiplications modulo n.

A single a not satisfying $a^n \equiv a \bmod n$ suffices to show that n is composite, without yielding a non-trivial divisor of n however.

In order to prove that a number n is *prime* rather than composite we need a *converse* to Fermat's theorem. Two problems present themselves in this connection.

I. The first problem is that the direct converse, in which "$\Rightarrow$" is replaced by "$\Leftarrow$", is false. Ramanujan's number $1729 = 7 \cdot 13 \cdot 19$ is composite, but we do have

$$a^{1729} \equiv a \bmod 1729 \quad \text{for all } a \in \mathbb{Z}.$$

Composite numbers with this property are called *Carmichael numbers*, and there are probably infinitely many of them.

II. The second problem is that even if the direct converse of Fermat's theorem were true it would not help us much, since checking all integers a (mod n) is completely unfeasible, even for moderately sized n.

How can these problems be solved?

The first problem is solved by the use of sharper versions of Fermat's theorem, which do admit a converse. We give two examples.

The first is an *algebraic generalization* of Fermat's theorem:

THEOREM. *If* n *is prime, then for any commutative ring* R *we have*

$$(a+b)^n \equiv a^n + b^n \bmod nR \quad \textit{for all } a, b \in R.$$

Here nR denotes the ideal $\{x + x + \ldots + x$ (n terms): $x \in R\}$ of R. The proof of the theorem makes use of the binomial theorem of Newton. It depends on the observation that the binomial coefficients $\binom{n}{i}$, for $0 < i < n$, are divisible by n if n is prime.

Taking $R = \mathbb{Z}$, $b = 1$, and using induction on a, we regain Fermat's theorem.

It can be proved that the *converse* of the above theorem is also valid: if the congruence in the theorem holds for all R and all $a, b \in R$ then n is prime. It suffices, in fact, to take R equal to the polynomial ring $\mathbb{Z}[X]$, and $a = X$, $b = 1$.

Before we go into a *number theoretic generalization* of Fermat's theorem we treat a few properties of the *Jacobi symbol* $\left(\frac{a}{n}\right)$; for more details we refer to the textbooks, such as [3].

For the rest of this lecture we assume that n is *odd*. If n is prime, then Fermat's theorem implies that

$$a^{n-1} \equiv 1 \bmod n \quad \text{for } a \in \mathbb{Z}, \ \gcd(a,n) = 1$$

and therefore

$$a^{(n-1)/2} \equiv 1 \quad \text{or} \quad -1 \bmod n.$$

The *Jacobi symbol* $\left(\frac{a}{n}\right) \in \{1, -1\}$ is defined, for $a \in \mathbb{Z}$, $\gcd(a,n) = 1$, by

$$\left(\frac{a}{n}\right) \equiv a^{(n-1)/2} \bmod n \quad \text{if } n \text{ is prime,}$$

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right) \cdots \left(\frac{a}{p_t}\right) \quad \text{if } n = p_1 p_2 \cdots p_t, \ p_i \text{ prime.}$$

Further we put $\left(\frac{a}{n}\right) = 0$ if $\gcd(a,n) \neq 1$, and $\left(\frac{a}{1}\right) = 1$ for all a.

The Jacobi symbol is studied in a theory developed by Gauss (1801), the main theorem of which is the *quadratic reciprocity law*, see [3]. I do not state this law here, but I just note one consequence of it that is important to know for the sequel: *using the reciprocity law one can efficiently calculate* $\left(\frac{a}{n}\right)$, even if the prime factorization of n is not known.

The definition of $(\frac{a}{n})$ immediately implies the following strengthening of Fermat's theorem.

THEOREM.

$$n \text{ prime} \Rightarrow (\forall a \in \mathbb{Z} : \gcd(a,n) = 1 \Rightarrow a^{(n-1)/2} \equiv (\tfrac{a}{n}) \bmod n).$$

Again it is, even for large a and n, easy to check whether the congruence $a^{(n-1)/2} \equiv (\frac{a}{n}) \bmod n$ holds, at least with the help of a computer.

D.H. Lehmer proved in 1976 that the converse to the above theorem is also valid. More precisely:

THEOREM. *If* n *is an odd composite number, then*

$$a^{(n-1)/2} \not\equiv (\tfrac{a}{n}) \bmod n$$

for at least half of all $a \in \{1, 2, \ldots, n-1\}$ *with* $\gcd(a,n) = 1$.

With these results problem I is solved in a satisfactory way. In the actual prime number test we use in fact a combination of both approaches: we consider congruences in certain *extension rings* of $\mathbb{Z}$, and these congruences contain a symbol that generalizes the *Jacobi symbol*.

We are still faced with problem II: it is not computationally feasible to test all a (mod n), and even less to try all rings R.

The first method to solve this problem that we discuss, is a *probabilistic* method. It is as follows. Choose 100 random values of a from $\{1, 2, \ldots, n-1\}$, and test for each a that is chosen, whether

$$a^{(n-1)/2} \equiv (\tfrac{a}{n}) = \pm 1 \bmod n.$$

If this is false for at least one a, then n is certainly composite. On the other hand, if it is true for *all* 100 values of a, then n has a *high probability* of being prime.

To see this, suppose that n is not prime. By Lehmer's theorem, every single a has a probability $\leq \frac{1}{2}$ to satisfy the above congruence,

so the probability that *all* 100 values of a satisfy the congruence is

$$\leq \left(\frac{1}{2}\right)^{100} < 0.00000000000000000000000000001.$$

It is therefore very difficult to doubt that n is prime.

This method, proposed by R. Solovay and V. Strassen, is basically incapable of yielding *rigorous primality proofs*. On the other hand it is very fast, and it may be applied before the more time consuming method discussed below is used.

G.L. Miller and M.O. Rabin proposed an improvement of the above test, in which the same error probability is achieved with less than half the same amount of computation.

The second method to solve problem II is a *futuristic* one, since it depends on an unproved conjecture, the *generalized Riemann hypothesis*. This is an assertion about the location of the zeros of certain complex functions that arise in analytic number theory. The hypothesis is likely to be correct, but it remains unproved to this day.

Assuming this hypothesis, G.L. Miller proved in 1976 that it suffices, in the above test, to try only the numbers a with

$$a < 70\ (\log n)^2, \quad a \text{ prime}.$$

If all these a satisfy $a^{(n-1)/2} \equiv \left(\frac{a}{n}\right) \bmod n$ then with the help of the generalized Riemann hypothesis it can be shown that n is prime.

This method has two shortcomings. First, a rough estimation shows that for numbers of about 100 digits it is approximately 500 times slower than the method discussed below, although it is faster for much larger values of n. Secondly, the generalized Riemann hypothesis is still unproved.

We now turn to the methods by which the primality of a number can be proved with *mathematical certainty*. The number of such methods is greater than I have time to enumerate in this lecture. Ultimately, they all depend on the following theorem, which is so obvious that nobody else is likely to claim it for himself:

LENSTRA'S THEOREM.

$$n \text{ is prime} \iff \text{every divisor } r \text{ of } n \text{ is a power of } n.$$

To prove $\Rightarrow$, remark that $1 = n^0$ and $n = n^1$. The proof of $\Leftarrow$ is left to the audience.

Which role does this theorem play in prime number tests? Inaccurately speaking it comes down to the following. Suppose that n satisfies many conditions of the sort

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

Then it can be proved that every divisor r of n behaves *in a certain way* as a power of n. In certain circumstances this information can be used to deduce that 1 and n are the only divisors of n, so that n is prime.

The following theorem may serve to illustrate this procedure. It is now also assumed that n is not divisible by 3.

THEOREM. *Suppose that*

$$a^{(n-1)/2} \equiv \left(\frac{a}{n}\right) \bmod n \quad \text{for} \quad a = -1, 2, 3$$

and that there exists $a \in \mathbb{Z}$ *for which*

$$a^{(n-1)/2} \equiv -1 \bmod n.$$

Then there exists, for every divisor r *of* n, *an integer* $i \geq 0$ *such that*

$$r \equiv n^i \bmod 24.$$

The condition that $a^{(n-1)/2} \equiv -1 \bmod n$ for some $a \in \mathbb{Z}$ cannot be removed, as can be seen from the example $n = 1729$. If n is indeed a prime number then it is usually easy to find an integer a for which $a^{(n-1)/2} \equiv -1 \bmod n$.

The conclusion of the theorem is, that every divisor r of n

behaves "in a certain way" (viz.: modulo 24) as a power n^i of n. In fact we can take i = 0 or i = 1, since $n^2 \equiv 1 \bmod 24$.

The proof of the theorem depends on the following assertion, which is valid for positive integers n_1, n_2 that are not divisible by 2 or 3:

$$n_1 \equiv n_2 \bmod 24 \iff \left(\frac{a}{n_1}\right) = \left(\frac{a}{n_2}\right) \quad \text{for } a = -1,\ 2,\ 3.$$

This assertion is an easy consequence of the quadratic reciprocity law mentioned above.

To prove the theorem it clearly suffices to consider a *prime* divisor r of n. Write

$$n-1 = u \cdot 2^k, \qquad r-1 = v \cdot 2^{\ell}$$

with u, v odd and k, $\ell \geq 1$. By hypothesis, there exists $a \in \mathbb{Z}$ with $a^{u \cdot 2^{k-1}} = a^{(n-1)/2} \equiv -1 \bmod n$. Raising this to the power v we see that

$$a^{uv2^{k-1}} \equiv -1 \bmod r.$$

Since $v2^{\ell} = r-1$ we have by Fermat's theorem

$$a^{uv2^{\ell}} \equiv 1 \bmod r$$

for the same a. Comparing these two congruences we see that we must have $k-1 < \ell$, so $\ell \geq k$.

Now let a = -1, 2 or 3. As before, the hypotheses of the theorem imply that

$$a^{uv2^{k-1}} \equiv \left(\frac{a}{n}\right) \bmod r$$

whereas we see from $a^{(r-1)/2} \equiv \left(\frac{a}{r}\right) \bmod r$ that

$$a^{uv2^{\ell-1}} \equiv \left(\frac{a}{r}\right) \bmod r.$$

It follows that

$$\left(\frac{a}{r}\right) = \left(\frac{a}{n}\right)^{2^{\ell-k}} \quad \text{for } a = -1,\ 2,\ 3.$$

If now $\ell = k$, then this reduces to $(\frac{a}{r}) = (\frac{a}{n})$, so the consequence of the quadratic reciprocity law stated above implies that $r \equiv n = n^1 \bmod 24$. If $\ell > k$, then $(\frac{a}{r}) = 1 = (\frac{a}{1})$ for $a = -1, 2, 3$, so for the same reason $r \equiv 1 = n^0 \bmod 24$. This proves the theorem.

The theorem just proved is not a very useful one for the purpose of primality testing, since the conclusion of the theorem is only very weak. It would be far more useful to have a similar theorem in which 24 were replaced by a considerably larger number. The crucial property of the number 24 on which the theorem depends is the following:

$$m^2 \equiv 1 \bmod 24 \text{ for every } m \in \mathbb{Z} \text{ with } \gcd(m,24) = 1.$$

It can be proved that 24 is the largest number with this property. If one wishes to replace 24 by a larger number, then squares have to be replaced by higher powers, and this leads to the consideration of symbols that generalize the Jacobi symbol.

Replacing squares by twelfth powers we find that 24 can be replaced by a much larger number:

$$m^{12} \equiv 1 \bmod 65520 \text{ for every } m \in \mathbb{Z} \text{ with } \gcd(m, 65520) = 1.$$

Here we have $65520 = 2^4 \cdot 3^2 \cdot 5 \cdot 7 \cdot 13$.

The computer program mentioned at the beginning of this lecture makes use of considerably larger numbers:

$$m^{5040} \equiv 1 \bmod s \text{ for every } m \in \mathbb{Z} \text{ with } \gcd(m, s) = 1.$$

Here $5040 = 2^4 \cdot 3^2 \cdot 5 \cdot 7$, and s is a number of 53 digits:

$$s = 153219867888544432846626127356636113800104312257712 00$$

$$= 2^6 \cdot 3^3 \cdot 5^2 \cdot 7^2 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 29 \cdot 31 \cdot 37 \cdot 41 \cdot 43 \cdot 61 \cdot 71 \cdot 73 \cdot 113 \cdot 127 \cdot 181 \cdot$$

$$211 \cdot 241 \cdot 281 \cdot 337 \cdot 421 \cdot 631 \cdot 1009 \cdot 2521.$$

Notice that $s > \sqrt{n}$ if n has at most 100 digits.

I now give a sketchy and not very accurate description of our prime number test. For more details I refer to [1].

Prime number test for $n < 10^{100}$.

First step. Test whether $\gcd(n, s) = 1$, with s as above. (This gcd can be determined with Euclid's algorithm.) If $\gcd(n, s) > 1$ then n is divisible by one of the prime factors of s, and we stop.

Second step. Test 67 congruences that are analogous to congruences of the form

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

but with a replaced by suitably chosen elements of the rings

$$\mathbb{Z}[e^{2\pi i/p^k}], \text{ with } p \text{ prime, } k \geq 1, \text{ and } p^k$$

$$\text{dividing } 5040 = 2^4 \cdot 3^2 \cdot 5 \cdot 7,$$

and with $\left(\frac{a}{n}\right)$ replaced by a generalized symbol, whose values are powers of $e^{2\pi i/p^k}$ (for $p^k = 2$ we have $e^{2\pi i/p^k} = -1$, and then we regain the Jacobi symbol).

If at least one of the 67 congruences is not satisfied, then n is composite, and one stops.

Suppose now that all 67 congruences are valid. The congruences are selected in such a way, that in this case it can be proved that for every divisor r of n there exists $i \in \mathbb{Z}$ such that

$$r \equiv n^i \bmod s, \quad 0 \leq i < 5040.$$

Third step. Starting from this information we determine all divisors r of n with $r \leq \sqrt{n}$. This clearly suffices to factor n into primes and hence to see whether n is prime.

From $r \leq \sqrt{n}$ it follows that $r < s$, so r is completely determined if we know r modulo s. Because of the information from the second step we can therefore proceed as follows: calculate for each $i = 0, 1, \ldots, 5039$, the number r_i for which

$$r_i \equiv n^i \bmod s, \quad 0 \le r_i < s.$$

Then all divisors $\le \sqrt{n}$ of n are found among the r_i, so we can finish the test with 5040 trial divisions.

From the above description of the third step one should not get the impression that the algorithm is helpful in factoring n if n is composite. In practice all composite numbers will be eliminated in one of the first two steps.

The algorithm can be adapted to deal with larger values of n. It was proved by C. Pomerance and A.M. Odlyzko that the running time is

$$O((\log n)^{c \log\log\log n})$$

for some constant c.

I finish with a prime number of 100 digits that was specially made for the *Wintersymposium;* our program proved it to be prime in 31.581 seconds:

$$2309142005181925131615190921130023091911211404090 7$$

$$0007051415152019030801160026050919200019830108096 7.$$

REFERENCES

In [1] one finds a detailed description of the new prime number test. Two surveys on the subject of primality testing are [2] and [4], the latter concentrating on the older tests that are not discussed in this lecture. One should also consult the bibliographies to [1], [2] and [4]. The theory of the Jacobi symbol and the quadratic reciprocity law are treated in [3].

[1] COHEN, H. & H.W. LENSTRA, JR., *Primality testing and Jacobi sums*, to appear. Preliminary version: Report 82-18, Mathematisch Instituut, Universiteit van Amsterdam.

[2] LENSTRA, H.W., JR., *Primality testing*, in: H.W. Lenstra, Jr. & R. Tijdeman (eds), Computational methods in number theory, Mathematisch Centrum, Amsterdam 1982.

[3] SCHOLZ, A. & B. SCHOENEBERG, *Einführung in die Zahlentheorie*, Sammlung Göschen 1131, Walter de Gruyter.

[4] WILLIAMS, H.C., *Primality testing on a computer*, Ars Combin. 5 (1978), 127-185.