INTRODUCTION

by

H.W. LENSTRA, JR.

This introductory lecture is devoted to a specific problem from computational number theory. The discussion will provide us with an opportunity to indicate which type of questions will be considered in the other lectures.

A classical theorem due to Fermat asserts that for every prime number p with $p \equiv 1 \bmod 4$ there exist integers x and y, unique up to order and sign, such that

$$p = x^2 + y^2.$$

For example, the prime factor $p = 1238926361552897$ of $2^{2^8} + 1$ discovered by BRENT and POLLARD [2] can be written as

$$p = 25515304^2 + 24246559^2.$$

How were these values determined? More generally, given p, how does one determine x and y in the most efficient way? That is the problem to be discussed in this lecture. Throughout p denotes a prime number that is $1 \bmod 4$.

DAVENPORT, in [5, Chapter V, Section 3], gives four methods of constructing x and y. Before we analyze their efficiency let us set ourselves a standard by first considering the trivial method. If we assume $x > y$ then $\sqrt{p/2} < x < \sqrt{p}$, so it suffices to test, for each x in this range, whether $p - x^2$ is a square. This takes time $O(p^{(1/2)+\varepsilon})$ for any $\varepsilon > 0$, the p^{ε} accounting for the arithmetic that must be done for each x; see Turk's lecture for a more precise analysis of the cost of arithmetic operations.

One of Davenport's constructions, due to Jacobsthal, is as follows. Let $(\frac{m}{p})$ denote the Legendre symbol, and choose $a,b \in \mathbb{Z}$ with $(\frac{a}{p}) = 1$, $(\frac{b}{p}) = -1$. Then the integers

$$x = \frac{1}{2}\sum_{n=1}^{p-1}\left(\frac{n(n^2-a)}{p}\right), \qquad y = \frac{1}{2}\sum_{n=1}^{p-1}\left(\frac{n(n^2-b)}{p}\right)$$

satisfy $x^2 + y^2 = p$. This can be proved by relating x and y to the number of solutions of each of the congruences

$$u^2 \equiv v^3 - av \bmod p,$$

$$u^2 \equiv v^3 - bv \bmod p,$$

see [6, Chapter 18, Theorem 5]. Using this construction for x and y in a straightforward way leads to an $O(p^{1+\varepsilon})$-algorithm, much slower than our standard.

Davenport's second construction is obtained by putting $a = -1$ in the above formula. Writing $p = 4k + 1$ and using that

$$\left(\frac{n(n^2+1)}{p}\right) \equiv (n^3 + n)^{2k} \bmod p,$$

$$\sum_{n=1}^{p-1} n^i \equiv \begin{cases} 0 \bmod p & \text{if } i \not\equiv 0 \bmod p-1 \\ -1 \bmod p & \text{if } i \equiv 0 \bmod p-1 \end{cases}$$

one readily finds that

$$x \equiv -\frac{1}{2}\binom{2k}{k} \bmod p,$$

as was first proved by Gauss. Together with $|x| < \frac{1}{2}p$ this suffices to determine x and hence y. Calculating $\binom{2k}{k} \bmod p$ in the trivial way we arrive again at an $O(p^{1+\varepsilon})$-algorithm. Using the technique described in Section 4 of Pomerance's paper we can reduce this to $O(p^{(1/2)+\varepsilon})$, exactly our standard but much slower in practice. At the end of the author's lecture on primality testing it will be seen that there is a much faster way to calculate $\binom{2k}{k} \bmod p$ if arithmetic operations on ordinary integers are assumed to be doable in unit time. But the size of the numbers that appear is such that this is a very unrealistic assumption; in the terminology of the lecture by Van Emde Boas we are using the wrong *machine model*.

The third method that we discuss is basically due to Legendre. Davenport formulates it in terms of continued fractions, but here we shall use quadratic forms. Define two sequences of integers $a_0, a_1, \ldots, b_0, b_1, \ldots$

as follows:

$$a_0 = 1,$$

$$b_0 = \text{greatest odd integer} < \sqrt{p},$$

$$a_{n+1} = (b_n^2 - p)/(4a_n),$$

$$b_{n+1} \equiv -b_n \bmod 2a_{n+1}, \quad \sqrt{p} - |2a_{n+1}| < b_{n+1} < \sqrt{p}.$$

For some n it will happen that $a_{n+1} = -a_n$, and then we have $(2a_n)^2 + b_n^2 = p$. For example, for p = 73 we have

n :	0	1	2	3	4	5
a_n:	1	-6	2	-3	4	-4
b_n:	7	5	7	5	3	

$$73 = (2a_4)^2 + b_4^2 = 8^2 + 3^2.$$

From Schoof's lecture it will be clear that the forms $F_n = a_nX^2 + b_nXY + a_{n+1}Y^2$ are precisely the binary quadratic forms of discriminant p in the *principal cycle*. It can be shown that the length ℓ of this cycle is 2 mod 4, and that $a_{n+1} = -a_n$ occurs first for $n = (\ell-2)/4$. The known estimate $\ell = O(p^{(1/2)+\varepsilon})$ thus implies that this is again an $O(p^{(1/2)+\varepsilon})$-algorithm. But Shanks' technique of jumping through the principal cycle, explained by Schoof, improves this significantly: the desired form $F_{(\ell-2)/4}$ can be found in time $O(p^{(1/4)+\varepsilon})$, and if the generalized Riemann hypothesis is assumed even in time $O(p^{(1/5)+\varepsilon})$. In several other contributions we shall encounter algorithms in which the Riemann hypothesis plays a role. In the paper of Brent *et al.* attention is paid to numerical techniques related to the Riemann hypothesis.

The sequences (a_n), (b_n) defined above can also be used to solve the Pell equation

$$x^2 - py^2 = -4.$$

More general equations such as

$$ax^n + by^m = c$$

are considered, from different angles, in the contribution of Stroeker and Tijdeman.

In Schoof's lecture it is explained how binary quadratic forms can be used to determine the class number and the units of a quadratic field. In the lectures by Brentjes and Zantema the same questions are considered for number fields of higher degree.

The fourth method discussed by Davenport is due to Serret, and again we give a slightly different formulation, as in [3]. If $p = x^2 + y^2$ then $u = xy^{-1}$ (division mod p) satisfies $u^2 \equiv -1 \bmod p$, and up to sign it is the only such integer modulo p. Suppose now that, conversely, an integer u is given such that $u^2 \equiv -1 \bmod p$. We claim that it is easy to recover x and y. One method to do this is by calculating the greatest common divisor of p and u+i in the ring $\mathbb{Z}[i]$ of Gaussian integers. This can be done by means of the Euclidean algorithm, which is valid in this ring, and the result is

$$\gcd(p, u+i) = x + yi$$

where $x, y \in \mathbb{Z}$ are such that $x^2 + y^2 = p$.

The second method to recover x and y from u employs the Euclidean algorithm only for ordinary integers. It proceeds as follows. Calculate the gcd of p and u by means of the ordinary Euclidean algorithm, until two consecutive remainders are less than $\sqrt{p}$; then these can be taken as x and y. Example: for $p = 73$ we have $u^2 \equiv -1 \bmod p$ for $u = 27$, and the sequence of successive remainders is

$$73, 27, 19, 8, 3, \ldots$$

so that we can take $x = 8$, $y = 3$. The proof of the correctness of this algorithm depends on the symmetry appearing in the sequence of congruences

$$\begin{aligned}
0.27 &\equiv 73 \bmod 73 \\
1.27 &\equiv 27 \bmod 73 \\
-2.27 &\equiv 19 \bmod 73 \\
3.27 &\equiv 8 \bmod 73 \\
-8.27 &\equiv 3 \bmod 73 \\
19.27 &\equiv 2 \bmod 73 \\
-27.27 &\equiv 1 \bmod 73 \\
73.27 &\equiv 0 \bmod 73.
\end{aligned}$$

This symmetry is caused by the next-to-last congruence $-u.u \equiv 1 \bmod p$.

This construction of x and y has a geometric interpretation: the pair (x,y) is a "short" vector in the two-dimensional lattice $\{(v,w) \in \mathbb{Z} \times \mathbb{Z}: v \equiv uw \bmod p\}$. For a method to find short vectors in higher dimensional lattices and an application to computational number theory we refer to [7]. The subject is closely related to diophantine approximation, as discussed in Brentjes' lecture.

How fast is the above method to construct x and y? The Euclidean algorithm takes time $O((\log p)^2)$, or in a faster version [8] only $O(\log p\ (\log\log p)^2\ \log\log\log p)$. But to this the time needed to find u should be added.

This leads to the question how the equation $u^2 \equiv -1 \bmod p$ can be solved. For the prime divisor $p = 1238926361552897$ of $2^{28}+1$ we can clearly take $u = 2^{27}$, and from this the values for x and y stated at the beginning can be easily computed. For general p we can take $u = (\frac{1}{2}(p-1))!$, but this formula is useless for computational purposes.

A.K. Lenstra discusses in his lecture a method to find zeros of polynomials over finite fields. Applying this to the polynomial X^2+1 over the field $\mathbb{Z}/p\mathbb{Z}$ we obtain a solution for our problem that is quite efficient in practice, but for which it is difficult to estimate the time needed in a satisfactory way.

The following method has a similar problem. Let b be the least positive integer with $(\frac{b}{p}) = -1$; then $b^{(p-1)/2} \equiv -1 \bmod p$, so we can take $u \equiv b^{(p-1)/4} \bmod p$. Using the reciprocity law for the Jacobi symbol one can calculate $(\frac{n}{p})$ in time $O((\log p)^2)$, for $0 < n < p$; perhaps this can be improved with the techniques of [8]. Further, $b^{(p-1)/4} \bmod p$ can be calculated in time $O((\log p)^{2+\varepsilon})$. Hence u can be determined in time $O(b(\log p)^2 + (\log p)^{2+\varepsilon})$; here we have $b = O(p^{1/(4\sqrt{e})+\varepsilon})$ (see [4]), and if the truth of the generalized Riemann hypothesis is assumed then $b = O((\log p)^2)$ (see [1]).

We conclude that Serret's method to solve $p = x^2+y^2$ takes time $O(p^{1/(4\sqrt{e})+\varepsilon})$, where $1/(4\sqrt{e}) = 0.15163...$, and $O((\log p)^4)$ if the generalized Riemann hypothesis is true.

An improvement of theoretical value was recently obtained by SCHOOF [9], who showed without any unproved assumption that $p = x^2+y^2$ can be solved in time $O((\log p)^6)$. His algorithm makes use of the elliptic curve $u^2 = v^3 - v$ (over $\mathbb{Z}/p\mathbb{Z}$) that we mentioned in connection with Jacobsthal's construction. It proceeds by investigating the action of the "Frobenius automorphism" on the ℓ-torsion points of the curve, for several small primes ℓ.

It may be expected that Schoof's algorithm is only the first of many applications of arithmetical algebraic geometry to computational number theory.

REFERENCES

[1] ANKENY, N.C., *The least quadratic non residue*, Ann. of Math. 55 (1952), 65-72.

[2] BRENT, R.P. & J.M. POLLARD, *Factorization of the eighth Fermat number*, Math. Comp. 36 (1981), 627-630.

[3] BRILLHART, J., *Note on representing a prime as a sum of two squares*, Math. Comp. 26 (1972), 1011-1013.

[4] BURGESS, D.A., *The distribution of quadratic residues and non-residues*, Mathematika 4 (1957), 106-112.

[5] DAVENPORT, H., *The higher arithmetic*, Hutchinson, London, 1968.

[6] IRELAND, K. & M. ROSEN, *A classical introduction to modern number theory*, Springer, New York, 1982.

[7] LENSTRA, A.K., H.W. LENSTRA, JR., & L. LOVÁSZ, *Factoring polynomials with rational coefficients*, Math. Annalen 261 (1982), 515-534.

[8] SCHÖNHAGE, A., *Schnelle Berechnung von Kettenbruchentwicklungen*, Acta Inform. 1 (1971), 139-144.

[9] SCHOOF, R.J., in preparation.