LECTURE 9, MONDAY 08.03.04

FRANZ LEMMERMEYER

We continue where we left off.

Pollard's p-1 Method Revisited. Let us now recast Pollard's algorithm in a different language. We were working in the groups $(\mathbb{Z}/N\mathbb{Z})^\times$ and $(\mathbb{Z}/p\mathbb{Z})^\times$. Letting $\mathcal{H}: XY = 1$ denote the standard hyperbola with neutral element $O = (1,1)$, observe that $\mathcal{H}(\mathbb{Z}/N\mathbb{Z}) = (\mathbb{Z}/N\mathbb{Z})^\times$ and $\mathcal{H}(\mathbb{Z}/N\mathbb{Z}) = (\mathbb{Z}/N\mathbb{Z})^\times$. We have chosen a point $P = (a, \frac{1}{a})$ with $b = \frac{1}{a}$ on $\mathcal{H}(\mathbb{Z}/N\mathbb{Z})$ and computed $nP = (x_n, y_n)$ with $x_n = a^n$, as well as the factor $\gcd(N, x_n - 1)$ of N.

Factoring using Conics. The idea is the same: take a Pell conic $\mathcal{C}: X^2 - \Delta Y^2 = 1$ and some point $P \in \mathcal{C}(\mathbb{Z}/N\mathbb{Z})$ (actually, in real life it's the other way round: pick a point $X \in \mathbb{Z}/n\mathbb{Z}$ at random, check that $\gcd(X, N) = 1$ (if not, we have found a factor and are done), and compute $\Delta \equiv X^2 + 1 \bmod N$; then $(X, 1)$ is a point on $\mathcal{C}$). Now compute the point $kP = (x_k, y_k)$ using the group law on $\mathcal{C}(\mathbb{Z}/N\mathbb{Z})$ with neutral element $O = (1, 0)$, and where k is a product of powers of small primes (as in Pollard's $p-1$-method). Now assume that $p \mid N$ has the property that $\#\mathcal{C}(\mathbb{Z}/p\mathbb{Z}) \mid k$: then $kP = (1, 0)$ in $\mathcal{C}(\mathbb{Z}/p\mathbb{Z})$, and therefore we will have $x_k \equiv 1 \bmod p$ and $y_k \equiv 0 \bmod p$. Thus $\gcd(x_k - 1, N)$ and $\gcd(y_k, N)$ are both divisors of N (possibly trivial if $kP = (1, 0)$ in $\mathcal{C}(\mathbb{Z}/N\mathbb{Z})$).

Note that $\#\mathcal{C}(\mathbb{Z}/p\mathbb{Z}) = p - (\frac{\Delta}{p})$, so if Δ happens to be a quadratic nonresidue modulo p, then p will be found if the factorization of $p + 1$ involves only small primes.

1. Factoring using elliptic curves

As above, pick random numbers x, y modulo N and then pick a, b such that $y^2 \equiv x^3 + ax + b \bmod N$. Now the points on this curve do not form a group with respect to the geometric group law: the problem is that when you compute $P + Q$, the denominator in the formula for x_{P+Q} might be divisible by a prime dividing N. But if the goal is to find such factors, such a failure of the addition formulas is exactly what you want. Thus take $P = (x, y)$ and compute $kP = (x_k, y_k)$ on $E(\mathbb{Z}/N\mathbb{Z})$ for a highly composite number k. For quite a while nothing exciting will happen. But if k becomes so large that the order of the group $E(\mathbb{Z}/p\mathbb{Z})$ divides k, then kP will be the point at infinity on $E(\mathbb{Z}/p\mathbb{Z})$, which means that the denominator of both x_k and y_k must be divisible by p. Thus if we compute the gcd of N and the denominators of x_k during the calculations, eventually we will find that this gcd becomes divisible by p. Except in the rare case when this gcd becomes N, we will have found a nontrivial factor.

In practice one does not work with a single curve, but with several curves simultaneously; if no factor is found, more elliptic curves are used, and the bound for k is increased.

Example. Let $M_n = 2^n - 1$ denote the n-th Mersenne number. On April 25, 1998, the complete factorization of the Mersenne number M_{589} was found. Known factors at the time were $M_{19} = 524287$, $M_{31} = 2147483647$, as well as 18083479 and 36064471. The factorization

$$\begin{aligned} p_{46} &= 2023706519999643990585239115064336980154410119 \\ p_{103} &= 13635133929781911357360183447731257848357221022 \\ &\quad 11913963639355051056896705852735103386975412732016027769 \end{aligned}$$

of the remaining factor was found using the elliptic curve

$$E : y^2 \equiv x^3 + Ax^2 + x \bmod p_{46},$$

where $A \equiv 78012041994340464943289779036551782426830367 6 \bmod p_{46}$. Its group order is

$$\begin{aligned} \#E(\mathbb{Z}/p_{46}\mathbb{Z}) &= 2023706519999643990585250126089270445504437408 \\ &= 2^5 \cdot 3 \cdot 7^2 \cdot 223 \cdot 661 \cdot 2141 \cdot 2621 \cdot 847031 \cdot 5965699 \cdot 6047191 \cdot 17020639711 \end{aligned}$$

The factorizations of $p_{46} \pm 1$ are

$$\begin{aligned} p_{46} - 1 &= 2 \cdot 7 \cdot 19 \cdot 31 \cdot 53 \cdot 181 \cdot 641 \cdot 399109188494863188876561949283238 41, \\ p_{46} + 1 &= 2^3 \cdot 3^3 \cdot 5 \cdot 17 \cdot 97 \cdot 113632646048089975438831565430470598351 1, \end{aligned}$$

hence p_{46} could not have been discovered by the $p - 1$ or $p + 1$ method.

2. Elliptic Curves over Finite Fields

Now in order to find the best strategy for factoring numbers using the elliptic curve method ECM, one needs to know a lot about the possible orders of elliptic curves over finite fields. It is clear that $E(\mathbb{F}_p)$ is finite since there are only p^2+p+1 points in the projective plane $\mathbb{P}^2\mathbb{F}_p$. Heuristically, we would expect a group order in the vicinity of $p + 1$: for approximately half the elements of $\mathbb{F}_p$ the values of $f(x) = x^3 + ax + b$ should be squares, and those squares correspond to two points $(x, \pm y)$ on the elliptic curve.

The simplest method to improve the trivial bound $\#E(\mathbb{F}_p) \le p^2 + p + 1$ is due to Postnikov.

Let $\mathbb{F}_q$ be a finite field of characteristic $p > 2$, and $f \in \mathbb{F}_q[X]$ a nonzero polynomial. In order to determine how many values of f are squares, we consider the equation

$$f(X)^{(q-1)/2} - 1 = 0 \tag{1}$$

and the polynomial

$$R(X) = 2f(X)(1 - f(X)^{(q-1)/2}) + f'(X)(X^q - X). \tag{2}$$

Clearly any root of (1) is a root of (2). Since

$$\begin{aligned} R'(X) &= 2f'(X)(1 - f(X)^{(q-1)/2}) \\ &\quad + f'(X)(f(X)^{(q-1)/2} - 1) + f''(X)(X^q - X), \end{aligned}$$

any root of (1) is a double root of (2).

Now $\deg R = \frac{q+1}{2} \deg f$, hence the number N_f of solutions of (1) satisfies $2N_f \le \frac{q+1}{2} \deg f$.

Now R is a polynomial of degree $\deg R = \frac{p+1}{2} \deg f$ with at least N_f roots of multiplicity ≥ 2 and at least $\delta \leq \deg f$ roots of multiplicity ≥ 1, where δ is the number of roots of f in $\mathbb{F}_p$. Thus the number N_f of solutions of (1) satisfies $2N_f + \delta \leq \frac{p+1}{2} \deg f$.

Similarly, any root of

$$f(X)^{(p-1)/2} + 1 = 0 \tag{3}$$

is at least a double root of

$$R(X) = 2f(X)(1 + f(X)^{(p-1)/2}) + f'(X)(X^p - X). \tag{4}$$

If we denote the number of roots of (3) by M_f, then $2M_f + \delta \leq \frac{p+1}{2} \deg f$. Moreover, $N_f + M_f + \delta = p$.

In the special case of a cubic polynomial f, we get $2N_f + \delta - p \leq \frac{p+3}{2}$ and $2M_f + \delta - p \leq \frac{p+3}{2}$. This implies $|2N_f + \delta - p| \leq \frac{p+3}{2}$.

Since the number of solutions of $y^2 = f(x)$ over $\mathbb{F}_p$ is given by $2N_f + \delta$, we find

Proposition 1. *The number N of $\mathbb{F}_p$-rational points (including the point $\mathcal{O}$ at infinity) on a Weierstrass elliptic curve satisfies $|N - (p+1)| \leq \frac{p+3}{2}$.*

This takes care of the existence of $\mathbb{F}_p$-rational points as well as of a bound for the number of $\mathbb{F}_p$-rational points. Already in the 1930s, a much sharper bound had been given by Hasse:

Theorem 2. *The number N of $\mathbb{F}_p$-rational points on an elliptic curve $Y^2 = X^3 + aX + b$ over $\mathbb{F}_p$ (including the point $\mathcal{O}$ at infinity) satisfies $|N - (p+1)| \leq 2\sqrt{p}$.*