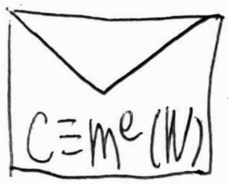
Pollard's and Lenstra's factorization algorithms.

Let's start by recalling the RSA cryptosystem.

Step 1. Bob chooses two distinct odd primes p and q and a number e with $\gcd(e,(p-1)(q-1))=1$. Then he publishes $N=pq$ and e.

Step 2. Alice — 

→ Bob

Step 3. Bob recovers m via finding $d \equiv e^{-1} \pmod{\varphi(N)}$ and taking $c^d \equiv m^{ed} \equiv m^{k\varphi(N)+1} \equiv m \pmod{N}$.

The key takeaway: to break RSA, one needs to find $d \equiv e^{-1} \pmod{\varphi(N)}$, in turn, to find d it suffices to factorize N.

Pollard's $p-1$ method.

Idea: suppose that we managed to find a number L, s.t. $p-1$ divides L,
$q-1$ does not divide L.

Then $a^L \equiv 1 \pmod{p}$,
$a^L \equiv a^r \pmod{q}$, where r is the residue of the division of L by $q-1$

we recover p as

$$p = \gcd(a^L - 1, N).$$

But how can we find such L?

Suppose that all factors of $p-1$ are relatively small:

$$p-1 = \prod_{i=1}^{s} a_i, \quad |a_i| \leq B \text{ for all } i \text{ and } B \text{ is 'reasonable'}.$$

Then $L = n!$ will work for sufficiently large value of n.

So we can simply check the values of $\gcd(a^{n!}-1, N)$ for $n = 2, 3, \ldots$ and some chosen base a, say, $a=2$.

1. If $\gcd(a^{n!}-1, N) = 1$, take the next value of n.
2. If $\gcd(a^{n!}-1, N) = N$, choose a different base a.
3. If $1 < \gcd(a^{n!}-1, N) < N$, then $\gcd(a^{n!}-1, N)$ is one of the two prime divisors of N (p or q).

<u>Rmk.</u> $a^{n!}$ is a very large number, but we only care about $a^k \pmod N$.

<u>Rmk.</u> We know that the fast powering algorithm computes $a^{n!}$ in $\Theta(2\log_2 n!)$ steps. On the other hand, Stirling's formula states that $\ln(n!) \sim n\ln(n) - n + \Theta(\ln(n))$

for large n. Hence, $n! \sim \left(\frac{n}{e}\right)^n$ and we can compute $a^{n!} \pmod N$ in $\sim 2\log_2\left(\frac{n}{e}\right)^n \approx 2n\log_2 n$ steps, which is feasible.

Rmk. The algorithm works well in case of the numbers $p-1$ or $q-1$ factorizes into the product of small primes.

Lenstra's elliptic curve factorization algorithm.

Let E be an elliptic curve with defining equation $y^2 = x^3 + ax + b$ over $\mathbb{Z}_N$ ($N = pq$ as before, in particular $\mathbb{Z}_N$ is not a field).

Take a point $P = (p_x, p_y)$ on E, i.e.

$$P_y^2 \equiv P_x^3 + aP_x + b \pmod N.$$

Key observation. Since N is not prime, the group law formulas will not always work. For instance, let $Q_k = kP$ and $Q_s = sP$.

We would like to find $Q = Q_k \oplus Q_s$. Using the formulas, we get

$$m \equiv \frac{(Q_k)_y - (Q_s)_y}{(Q_k)_x - (Q_s)_x} \quad \text{and} \quad Q_x \equiv m^2 - (Q_k)_x - (Q_s)_x.$$

In order for the formula to make sense, the solution to the congruence $m \equiv \frac{(Q_k)_y - (Q_s)_y}{(Q_k)_x - (Q_s)_x} \pmod N$ must exist, equivalently, $\gcd((Q_k)_x - (Q_s)_x, N) = 1$.

Thus, in case we fail to compute $Q_k \oplus Q_s$, we get a divisor of N on the way ☺

The algorithm works as follows. We compute (attempt to) the points $2!\cdot P$, $3!\cdot P$, $4!\cdot P$, ... There are 3 possibilities:

1. The computation is successful: we find $n!\cdot P$.
2. Somewhere in the computation, we had to find the inverse of a number d, divisible by N, and failed.
3. While the computation, we had to find the inverse of a number d with $1 < \gcd(d, N) < N$.

In this case we find a prime factor of N.

<u>Rmk.</u> We need to find a point P on E to start with.

Easy trick: start with $P = (\alpha, \beta)$, pick any a and set $b \equiv \beta^2 - \alpha^3 - a\cdot\alpha \pmod N$.

Rmk. Let $N=pq$, then we fail to add two points P and Q on $E(\mathbb{Z}_N)$, when $P \oplus Q = \mathcal{O}$ on $E(\mathbb{F}_p)$ or $P \oplus Q = \mathcal{O}$ on $E(\mathbb{F}_q)$ (or both). In case we are computing multiples of P, the failure occurs when m is divisible by the order of P on $E(\mathbb{F}_p)$ or the order of P on $E(\mathbb{F}_q)$.

Example (Pollard's algorithm). Let $N=437$, choose $a=2$.

① $n=1$: $\gcd(2^{1}-1, 437)=1$

② $n=2$: $\gcd(2^{2!}-1, 437) = \gcd(3, 437) = 1$.

③ $n=3$: $\gcd(2^{3!}-1, 437) = \gcd(2^{6}-1, 437) = \gcd(63, 437) = 1$.

④ $n=4$: $\gcd(2^{4!}-1, 437) = \gcd(348, 437) = 1$.
$(2^{4!} \equiv 349 \pmod{437})$

⑤ $n=5$: $\gcd(2^{5!}-1, 437) = \gcd(333, 437) = 1$.
$(2^{5!} \equiv 334 \pmod{437})$

⑥ $n=6$: $\gcd(2^{6!}-1, 437) = \gcd(399, 437) = 19$. $\left(\begin{array}{l} 437 = 1\cdot 399 + 38 \\ 38 = 19\cdot 2 + 0 \end{array}\right.$
$(2^{6!} \equiv 400 \pmod{437})$

$437 = 19 \cdot 23$.

Rmk. $19-1 = 18 = 2\cdot 3^2$, $6!$ is the first factorial divisible by 18.
$23-1 = 2\cdot 11$

__Example__. Consider the elliptic curve E given by equation $y^2 = x^3+4x+4$ over $\mathbb{Z}/_{21\mathbb{Z}}$ (the discriminant $D_f = 4\cdot 4^3+27\cdot 4^2 \equiv$
$\equiv 4+6\cdot 16 \equiv 100 \equiv 16 \not\equiv 0 \pmod{21}$, so E is smooth).
Let $P=(1,3)$ and $Q=(15,4)$ be two points on E.
The line through points P and Q has slope $m = \frac{4-3}{15-1} =$
$= \frac{1}{14}$, but $\gcd(14,21)=7$ is a divisor of 21, so 14 isn't invertible.
Next we take a look at E over $\mathbb{F}_3$ and $\mathbb{F}_7$ (here 3 and 7 are two prime factors of 21).
__Over $\mathbb{F}_3$__: the defining equation of E simplifies to $y^2 = x^3+x+1$. We notice that P has coordinates $(1,0)$, so the order of P is 2 (it is on the x-axis).
__Over $\mathbb{F}_7$__: E has 10 points, and the order of P is equal to 5 (check it!)
Let's compute $P \oplus P$:

- over $\mathbb{F}_3$: $P\oplus P = \theta$;
- over $\mathbb{F}_7$: the slope of the tangent line is $\frac{3\cdot 1^2+4}{2\cdot 3} \equiv 0$; x-coordinate of $P\oplus P$ is $0^2-2\cdot 1 \equiv -2 \equiv 5$ and y-coordinate is $-3\equiv 4$, hence $P\oplus P = (5,4)$;
- over $\mathbb{Z}_{21}$: the slope of the tangent line is $\frac{3\cdot 1^2+4}{2\cdot 3} \equiv \frac{7}{6}$, but $\gcd(6,21)=3$, hence, 6 is not invertible modulo 21 and 3 is a divisor of it.