The Pollard $p-1$ factoring algorithm.

The Pollard $p-1$ algorithm, is a factoring algorithm introduced by John Pollard in 1974. It takes its name from the fact that it detects the prime factors p of an integer n, with the property that $p-1$ is a *smooth number*, i.e. it factors into relatively small primes. Equivalently, the group $\mathbf{Z}_p^*$ has *smooth* order.

Despite its limited probability of success, the Pollard $p-1$ algorithm is important because in Lenstra's elliptic curve method for factoring (ECM), the same principles are applied to the group of points of an elliptic curve over $\mathbf{Z}_p$.

Integers all of whose prime factors are $\leq B$, are called *B-smooth*.

Let $n \in \mathbf{N}$ be the integer to be factored. Fix a smoothness bound $B \in \mathbf{N}$ and define

$$\mathbf{k} := \prod_{\substack{p \leq B \\ p^\alpha \leq B}} p^\alpha.$$

We may think of $\mathbf{k}$ as "the mother of all B-smooth numbers", in the sense that "any" B-smooth number divides $\mathbf{k}$. The Pollard $p-1$ algorithm consists of taking a random class $a \in \mathbf{Z}_n^*$ and computing

$$\gcd(a^{\mathbf{k}} - 1 \bmod n, n).$$

When is the above *gcd* going to be > 1?

Heuristically, we can say that $\gcd(a^{\mathbf{k}} - 1 \bmod n, n) > 1$ *if and only if n admits some prime factor p of n with the property that $p - 1 = \#\mathbf{Z}_p^*$ is B-smooth.*

Proof. If p is a prime factor of n with the property that $p-1$ is B-smooth, then we may assume that $p-1$ divides $\mathbf{k}$. From the Fermat Little Theorem applied to $\mathbf{Z}_p^*$, one deduces that $a^{\mathbf{k}} = 1 \bmod p$ and that p divides $\gcd(a^{\mathbf{k}} - 1 \bmod n, n)$.
Conversely, if a prime p divides $\gcd(a^{\mathbf{k}} - 1 \bmod n, n)$, then p is a factor of n and $a^{\mathbf{k}} = 1 \bmod p$. This implies that the order of a in $\mathbf{Z}_p^*$ divides $\mathbf{k}$ and therefore it is B-smooth. On the other hand, in the cyclic group $\mathbf{Z}_p^*$, the order of an element is typically given by $\frac{p-1}{r}$, for some small integer r. Hence $p-1$ is B-smooth.

Remark. If a smoothness bound B is fixed, then the $\gcd(a^{\mathbf{k}} - 1 \bmod n, n)$ coincides with *the product of all prime factors p* of n such that $p-1$ is B-smooth.

The complexity of the algorithm is dominated by the complexity of the calculation

$$a^{\mathbf{k}} \bmod n,$$

which is

$$\mathcal{O}(\log \mathbf{k} (\log n)^2) = \mathcal{O}(B(\log n)^2),$$

since $\mathbf{k}$ is roughly $\exp B$.

Once a smoothness bound B is fixed, the B-smooth integers become more and more sparse as their size grows. This means that factors p such that $p-1$ is B-smooth are rare. If the algorithm fails to find a factor for a given smoothness bound B, the only option is to increase B. On the other hand increasing B is very expensive from a computational point of view (the complexity of the algorithm is exponential in $\log B$). This is why this algorithm has limited probability of success.

Example. Consider $n = 11951438413903$.
For $B = 50$ and $a = 57$, we get $\gcd(a^{\mathbf{k}} - 1, n) = 1$;
for $B = 100$ and $a = 57$, we get $\gcd(a^{\mathbf{k}} - 1, n) = 1$;
for $B = 150$ and $a = 57$, we get $\gcd(a^{\mathbf{k}} - 1, n) = 108769$.

Indeed $n = p * q$, with $p = 108769$ and $q = 109879087$, and

$$p - 1 = 2^5 \cdot 3 \cdot 11 \cdot 103, \qquad q - 1 = 2 \cdot 3 \cdot 29 \cdot 373 \cdot 1693.$$

Exercise. Consider $n = p \cdot q = 687442130387521$, where $p = 686989$ and $q = 1000659589$.
As $p - 1 = 2^2 \cdot 3^3 \cdot 6361$ and $q - 1 = 2^2 \cdot 3 \cdot 59 \cdot 1413361$.
How big should one choose B in order for the algorithm to detect the non-trivial factor p?
What happens if one takes $B > 1413361$?

The Miller-Rabin theorem and the Miller-Rabin primality test.

Miller-Rabin primality test is a very efficient algorithm to detect whether a positive integer $n \in \mathbf{N}$ is composite or it is possibly prime. It is based on the Miller-Rabin Theorem (MRT in short), which is a refinement of the Fermat Little Theorem (FLT in short).

Theorem. (Fermat Little Theorem). *Let $n \in \mathbf{N}$ be a prime number and let a be an integer satisfying* $\gcd(a, n) = 1$. *Then*

$$a^{n-1} = 1 \bmod n.$$

The FLT provides a necessary condition for an integer n to be prime: if $a^{n-1} \neq 1 \bmod n$, then we are sure that n is composite. On the other hand, if $a^{n-1} = 1 \bmod n$, then we cannot conclude anything about the primality of n. Even worse, there infinitely many composite integers, the *Carmicheal numbers*, for which $a^{n-1} = 1 \bmod n$, for every a with $\gcd(a, n) = 1$. The MRT combines the FLT with the following fact:

Let p be a prime number. If p divides $x^2 - 1$, then either $x = 1 \bmod p$ or $p = -1 \bmod p$.

Theorem. (Miller-Rabin Theorem). *Let $n \in \mathbf{N}$ be a prime number and let a be an integer satisfying* $\gcd(a, n) = 1$. *Write $n - 1 = m2^k$, for some m odd, and set $b := a^m$. Then either $b = 1 \bmod n$ or there is $1 \leq s \leq k$ such that $b^{2^{s-1}} = -1 \bmod n$.*

Proof. By FLT one has

$$a^{n-1} = (a^m)^{2^k} = 1 \bmod n.$$

Now we examine how such a power of a can become to 1, modulo n.
One way is that already $b := a^m = 1 \bmod n$.
Otherwise, if $b := a^m \neq 1 \bmod n$, then the power a^{n-1} is obtained from b by successive squarings modulo n

$$b^2 = b \cdot b, \qquad b^{2^2} = b^2 \cdot b^2, \quad \ldots \quad b^{2^k} = b^{2^{k-1}} \cdot b^{2^{k-1}}.$$

Let s be the smallest positive in integer, $1 \leq s \leq k$, for which

$$b^{2^s} = b^{2^{s-1}} \cdot b^{2^{s-1}} = 1.$$

Then the element $b^{2^{s-1}}$ is a zero of the polynomial $X^2 - 1$. Since n is prime, one has that either $b^{2^{s-1}} = 1 \bmod n$ or $b^{2^{s-1}} = -1 \bmod n$. Since s is the *smallest* integer for which $b^{2^s} = 1 \bmod n$, then necessarily

$$b^{2^{s-1}} = -1 \bmod n,$$

as claimed.

An integer n is called a-pseudoprime, provided that it satisfies the Miller-Rabin theorem for some class $a \in \mathbf{Z}_n^*$. Like the FLT, also the MRT only provides a necessary condition for an integer n to be prime: if $a^{n-1} \neq 1 \bmod n$, then n is certainly composite. On the other hand, given an *odd composite integer* $n > 9$, then

$$\#\{a \in \mathbf{Z}_n^* \mid n \text{ is } a\text{-pseudoprime}\} \leq \varphi(n)/4,$$

where φ denotes the Euler φ function (see Schoof R., Four primality tests).

This means in particular that if n is a-pseudoprime for k different bases a, the probability that it is composite is $\leq 1/4^k$. This fact combined with the Prime Number Theorem enables one to determine integers n, with probability of being primes arbitrarily close to 1, in polynomial time.