6.2. The Integer Factorization Problem

Given composite n. Assume that n has a prime factor p such that $p-1$ has all prime factors $\leq B$.

Algorithm Pollard $-(p-1)$

- Choose $a > 1$ (often $a=2$)
- Compute $b = a^{B!} \bmod n$
- Compute $d = \gcd(b-1, n)$
- If $1 < d < n$ then d is a nontrivial factor n.

Proof that Pollard $-(p-1)$ is correct:

Assume p is a prime factor of n s.t. $p-1$ has all prime factors $\leq B$.

Then $p-1 \mid B!$, i.e., $B! = k(p-1)$ for some $k \in \mathbb{N}$.

By Fermat's theorem

$$a^{B!} \equiv (a^{p-1})^k \equiv 1 \bmod p$$

Hence, $a^{B!} - 1 \equiv 0 \bmod p$, such that $\gcd(a^{B!}-1, n)$ is a factor of n. $\blacksquare$

Remarks:

a) Compute $a^{B!} \bmod n$ as follows

$b_1 = a \bmod n$, $b_j = b_{j-1}^{j} \bmod n$, $j = 2, \dots, B$.

b) $B!$ can be substituted by

$$\prod_{\substack{q \le B \\ q \text{ prime}}} q^{\lfloor \ln n / \ln q \rfloor}$$

Note $q^{\lfloor \ln n / \ln q \rfloor} \le n$

since: $l \le \frac{\ln n}{\ln q} \Leftrightarrow l \ln q \le \ln n$

$\Leftrightarrow q^{l} \le n$

still $(p-1) \mid \prod_{\substack{q \le B \\ q \text{ prime}}} q^{\lfloor \ln n / \ln q \rfloor}$ holds.

To protect against Pollard-$(p-1)$ select $n = p \cdot q$ s.t. $p-1$ and $q-1$ have at least one large prime factor. How $\rightarrow$ exercise (Trappe&Wash.)

Improvement of Pollard-$(p-1)$ is "elliptic curve factoring".

Another principle of factoring alg. is:

Example: Factor 8051

$$8051 = 8100 - 49 = 90^2 - 7^2$$
$$= (90+7)(90-7) = 97 \cdot 83$$

<u>Prop. 6.8.</u> $x \not\equiv \pm y \pmod n$, $x^2 \equiv y^2 \pmod n$
$\Rightarrow \gcd(x-y, n)$ is a nontrivial divisor of n. ⌟

<u>Proof.</u> $x^2 \equiv y^2 \pmod n$, i.e., $n \mid (x^2-y^2)$.
Hence $n \mid (x+y)(x-y)$. By assumption
$n \nmid (x+y)$ and $n \nmid (x-y)$, which shows the assertion. ∎

- Prop. 6.8. forms the basis of "quadratic sieve factoring". Problem: determine x, y as above. [Stinson (2002), p. 182–194]
- $\gcd(a,b)$ can be efficiently computed by the Euclidean algorithm, see section 6.3.

<u>Factoring in practice</u>

Presently 3 methods are most successful
- quadratic sieve
- elliptic curve factorization
- number field sieve (most powerful)

They
Are "subexponential", not polynomial

History of factoring

1994: RSA-129 quadr. sieve, 600 workstations

1996: RSA-130

1999: RSA-155 8400 MIPS-years, 300 PCs

2003: RSA-174

2005: RSA-193 5 months, 80 2.2 GHz Opteron CPU

2010: RSA-232 almost 2 years (12 authors)

Factoring is considered as a <u>one-way function</u>.

- "Easy": Given 2 prime p, q - Compute $n = p \cdot q$
- "Computationally infeasible", "hopeless"
  Given $n = p \cdot q$, p, q prime, unknown
  Determine p and q.

6.3. The Extended Euclidean Algorithm

Known: $\gcd(a, n) = 1 \iff \exists$ inverse s s.t. $a \cdot s \equiv 1 \pmod{n}$

Aim: efficient alg. for computing $s = a^{-1}$.

Euclidean algorithm: Let $r_0 > r_1 \in \mathbb{N}$ (*)

$r_0 = q_1 r_1 + r_2$, $0 < r_2 < r_1$	$r_2 = r_0 - q_1 r_1 = s_2 r_0 + t_2 r_1$
$r_1 = q_2 r_2 + r_3$, $0 < r_3 < r_2$	$r_3 = r_1 - q_2 r_2 = s_3 r_0 + t_3 r_1$
$\vdots$	$\vdots$
$r_{k-2} = q_{k-1} r_{k-1} + r_k$, $0 < r_k < r_{k-1}$	$r_k = s_k r_0 + t_k r_1$
$r_{k-1} = q r_k$	

It holds that

$$\gcd(r_0, r_1) = \gcd(r_1, r_2) = \cdots = \gcd(r_{k-1}, r_k) = r_k$$

[since $\gcd(a,b) = \gcd(b, a-qb)$]

If $\gcd(r_0, r_1) = 1$, then $s_k r_0 + t_k r_1 = 1$

$\Rightarrow t_k r_1 \equiv 1 - s_k r_0 \equiv 1 \pmod{r_0}$, i.e., $t_k \equiv r_1^{-1} \pmod{r_0}$

Define recursively (according to the r.h.s. of (*))

$$t_0 = 0, \quad t_1 = 1, \quad t_j = (t_{j-2} - q_{j-1} t_{j-1}) \bmod r_0, \; j \geq 2$$

<u>Th. 6.8.</u> $r_j \equiv t_j \cdot r_1 \pmod{r_0}$, $j = 0, \dots, k$ ⌟

<u>Proof.</u> (by induction)

$j=0$: $r_0 \equiv t_0 r_1 \pmod{r_0}$

$j=1$: $r_1 \equiv t_1 r_1 \pmod{r_0}$

Thinking the Future
Zukunft denken

$(j-2, j-1) \to j$:

$$r_j \overset{\text{Eucl. alg.}}{\equiv} r_{j-2} - q_{j-1} r_{j-1} \overset{\text{I.V.}}{\equiv} t_{j-2} r_1 - q_{j-1} t_{j-1} r_1$$

$$\equiv (t_{j-2} - q_{j-1} t_{j-1}) r_1 \equiv t_j \cdot r_1 \pmod{r_0}. \quad \boxed{\equiv}$$

Corollary 6.9. $\gcd(r_0, r_1) = 1 \Rightarrow t_k = r_1^{-1} \pmod{r_0}$.

Number of divisions in the Eucl. alg.

$$\leq \log_\phi(\sqrt{5}\, r_0) - 2 \quad , \quad \phi = \tfrac{1}{2}(1 + \sqrt{5})$$

"golden ratio"

(cf. Knuth II, Chap. 4.5.3, p. 320)

Least favorable if r_0, r_1 are succive Fibonacci numbers.

1 1 2 3 5 8 13 21 .....

6.4. The Chinese Remainder Theorem

Theorem 6.10.

Suppose $m_1, \dots, m_r$ are pairwise relatively prime, $a_1, \dots, a_r \in \mathbb{N}$. The system of r congruences

$$x \equiv a_i \pmod{m_i} \quad , \quad i = 1, \dots, r$$

has a unique solution modulo $M = m_1 \cdots m_r$, given by

$$x = \sum_{i=1}^{r} a_i M_i y_i \mod M$$

where $M_i = M/m_i$, $y_i = M_i^{-1} \mod m_i$, $i = 1, \dots, r$. $\perp$

Example:

$r=3, \quad m_1=7, \quad m_2=11, \quad m_3=13$

$\Rightarrow M=1001, \quad M_1=143, \quad M_2=91, \quad M_3=77$

$y_1 = 143^{-1} \bmod 7 = 3^{-1} \bmod 7 = 5$

$y_2 = 4, \quad y_3 = 12$

Solution of
$$x \equiv 5 \pmod 7 \qquad (a_1=5)$$
$$x \equiv 3 \pmod{11} \qquad (a_2=3)$$
$$x \equiv 10 \pmod{13} \qquad (a_3=10)$$

is
$$x = 5\cdot 143\cdot 5 + 3\cdot 91\cdot 4 + 10\cdot 77\cdot 12 \bmod 1001$$
$$= 894.$$