MPRI – Cours 2.12.2

F. Morain

Groups and applications

2017/09/13-20 & 2017/12/{05,19}

The slides are available on `http://www.lix.polytechnique.fr/Labo/Francois.Morain/MPRI/2017`

Contents

I. Introduction

Why groups? finite groups are used everywhere in crypto (and elsewhere).

Which tasks?

- representing elements;
- drawing elements at random;
- efficient group laws;
- computation of cardinality;
- structure (with generators);
- etc.

Some groups

- $(\mathbb{Z}/N\mathbb{Z})^*$;
- finite fields $\mathbb{F}_{q^n}$ and subfields;
- algebraic curves (elliptic, hyperelliptic, any genus) over finite fields;
- class groups;
- etc.

II. Algorithms for generic groups

$(G, \circ, 1_G)$, Abelian, finite, of order N; computable $\circ$.

Def. $\mathrm{ord}_G(a) = \min\{k > 0, a^k = 1_G\}$.

Thm. (Lagrange) $\mathrm{ord}_G(a) \mid N$.

Coro. $a^{-1} = a^{N-1}$.

Def. $\mathrm{Exp}(G) = \min\{k > 0, \forall\, a \in G, a^k = 1_G\}$.

Prop.
1. $\mathrm{Exp}(G) \mid N$;
2. $\mathrm{Exp}(G) = \mathrm{lcm}(\mathrm{ord}_G(a), a \in G)$.

It can happen that $\mathrm{Exp}(G) < N$, see later.

Finding the order of an element

Pb. $G = \langle g \rangle$, $N = \mathrm{ord}(g)$; what is the order ω of a in G?

Thm. (Lagrange) $\omega \mid N$.

Rem. If N is small, we can enumerate in $O(N)$ or its divisors.

Prop. a is of order ω if and only if
i) $a^\omega = 1_G$;
ii) for all prime $p \mid \omega$, $a^{\omega/p} \neq 1_G$.

Proof:

In practice, if N *and* its factorization are known, easy.

What if we don't know N (completely)? E.g., (hyper)elliptic curves.

Baby-steps giant-steps

Fundamental algorithm in ANT/crypto; due to Shanks.

Write:

$$\omega = cu + d, \quad 0 \le d < u, \quad 0 \le c < N/u.$$
$$a^\omega = 1_G \Leftrightarrow (a^{-u})^c = a^d.$$

Number of group operations: $C_o = u + N/u$ minimized for $u = \sqrt{N}$, hence $2\sqrt{N}$ group operations.

Set operations: u insertions in $\mathcal{B}$ and N/u membership tests in the worst case.

$\Rightarrow \mathcal{B}$ must be a hash table, where both operations take $O(1)$.

Complexity: $O(\sqrt{N})$ in time and space.

Function *BSGS(G, g, N, a)*
Input : $G \supset \langle g \rangle$, g of order N
Output: $\omega = \mathrm{ord}_g(a)$
$u \leftarrow \lceil \sqrt{N} \rceil$;
`// Step 1 (baby steps)`
initialize a table $\mathcal{B}$ for storing u pairs (elt of G, int $< N$);
store($\mathcal{B}$, $(1_G, 0)$);
$H \leftarrow a$; store($\mathcal{B}$, $(H, 1)$);
for $d := 2$ **to** $u - 1$ **do**
  $H \leftarrow H \circ a$; store($\mathcal{B}$, (H, d));
`// Step 2 (giant steps)`
$H \leftarrow H \circ a$; $f \leftarrow 1/H = a^{-u}$;
$H \leftarrow 1_G$;
for $c := 0$ **to** N/u **do**
  // $H = f^c$
  if $\exists (H', d) \in \mathcal{B}$ *such that* $H = H'$ **then**
    // $H = f^c = a^d$ hence $\omega = cu + d$
    return $cu + d$;
  $H \leftarrow H \circ f$;

Exercises

Exo1-0. Fill in the missing details in function BSGS (numerical trials and/or think).

Exo1-1. Decrease the average time by remarking that $c \approx N/(2u)$ on average.

Exo1-2. What if computing $1/x$ is free?

Exo1-3. Design a variant which takes $O(\max(c,d))$ operations. What is its average running time?

III. Divisibility; elementary algorithms

Thm. $\mathbb{Z}$ is an euclidean domain: $a = bq + r$, $0 \leq r < |b|$.

Divisibility: $b \mid a$ iff $r = 0$.

Prime: p with two positive divisors: $\{1, p\}$.

Thm. All integers $N > 1$ can be written as

$$N = \prod_{i=1}^{k} p_i^{\alpha_i}$$

where p_i are distinct primes, $\alpha_i > 0$, in a unique way (up to permutation of factors).

Properties and problems

Properties:

- number of prime factors: $\omega(N) = k$; $1 \leq \omega(N) \leq \log_2 N$, $\log\log N$ on average;
- number of divisors: $d(N) = \prod_{i=1}^{k}(1+\alpha_i)$; $d(N) \geq 2$, $\overline{\lim}\, d(N)$ exponential in $\log N$ (highly composite numbers).

Pb1: compute $\omega(N)$ (resp. $d(N)$) without factoring N.

Pb2: (squarefreeness) given N, is it true that all α_i's are 1?

Elementay factoring (1/2)

Thm. An integer N has a prime factor $\leq \sqrt{N}$.

Easy: enumerate all primes $\leq \sqrt{N}$ and factor N. $O(\sqrt{N})$ operations.

Fermat: find X and Y s.t. $N = X^2 - Y^2 = (X-Y)(X+Y)$.

Function *Fermat(N)*
  Input : N an integer > 1
  Output: a pair of factors of N
  $Y \leftarrow -1$;
  repeat
    $Y \leftarrow Y + 1$;
    $X_2 \leftarrow N + Y^2$;
  until X_2 *is a perfect square*;
  $X \leftarrow \sqrt{X_2}$;
  return $(X - Y, X + Y)$.

Elementay factoring (2/2)

Exo2-1. How to test if X_2 is a square rapidly?

Prop. If $N = ab$ with $a \leq b$, need $1 + (b-a)/2$ loops.
Proof: write $Y = (b-a)/2$, for which $X_2 = N + Y^2 = ((a+b)/2)^2$. $\square$

Generalization: if $N = pq$ with $p/q \approx \ell/m$, we have

$$4\ell mN = (mp + \ell q)^2 - (mp - \ell q)^2.$$

Prop. (Sherman-Lehman) find k s.t. $4kN = X^2 - Y^2$ with $k = \ell m$ and use preceding remark.

$\Rightarrow$ deterministic $O(N^{1/3})$.

Rem. Heuristic $O(N^{1/4+\epsilon})$ by McKee.

IV. The ring $(\mathbb{Z}/N\mathbb{Z})^*$

A) General results

Def. $\mathbb{Z}/N\mathbb{Z} = \{0, 1, \ldots, N-1\}$ set of equivalence classes of $x\mathcal{R}y \iff x - y \in N\mathbb{Z}$ or $x \equiv y \bmod N$; add ring operations.

Prop. $(\mathbb{Z}/N\mathbb{Z})^* = \{x \in \mathbb{Z}/N\mathbb{Z}, \exists y, xy \equiv 1 \bmod N\}$
$= \{x \in \mathbb{Z}/N\mathbb{Z}, \gcd(x, N) = 1\}$.

Thm. (Euler totient function)
$\varphi(N) := \mathrm{Card}((\mathbb{Z}/N\mathbb{Z})^*) = \prod_{i=1}^k \varphi(p_i^{\alpha_i}) = \prod_{i=1}^k p_i^{\alpha_i - 1}(p_i - 1)$ where $N = \prod_{i=1}^k p_i^{\alpha_i}$.

Thm. (Carmichael function) $\mathrm{Exp}((\mathbb{Z}/N\mathbb{Z})^*) = \lambda(N) = \mathrm{lcm}_{i=1}^k \lambda(p_i^{\alpha_i})$ where

$$\lambda(p_i^{\alpha_i}) = \begin{cases} \varphi(p_i^{\alpha_i}) = p_i^{\alpha_i - 1}(p_i - 1) & \text{if } p_i \text{ odd or } \alpha_i \leq 2, \\ 2^{e-2} & \text{if } e \geq 3. \end{cases}$$

More properties

Thm. $\mathbb{Z}/N\mathbb{Z}$ is a field iff N is prime.

Thm. $\mathbb{Z}/N\mathbb{Z} \simeq \prod_{i=1}^k \mathbb{Z}/p_i^{\alpha_i}\mathbb{Z}$.

Rem. Chinese Remaindering Theorem (CRT) Given $(x_i)_{1 \leq i \leq k}$ with $x_i \in \mathbb{Z}/p_i^{\alpha_i}\mathbb{Z}$, $\exists$ unique $x \in \mathbb{Z}/N\mathbb{Z}$, $x \equiv x_i \bmod p_i^{\alpha_i}$ for all i.

Thm. $(\mathbb{Z}/N\mathbb{Z})^*$ is cyclic iff $N = p^\alpha$ or $2p^\alpha$ for odd p, or $N = 2, 4$.

Justification of RSA

Prop. If N is squarefree, then for all $a \in \mathbb{Z}$, $a^{\lambda(N)+1} \equiv a \bmod N$.
Proof:

Coro. RSA is valid: for all x, $x^{ed} \equiv x \bmod N$.
Proof:

B) Application to primality proving

Thm.(Fermat) N is prime if and only if $(\mathbb{Z}/N\mathbb{Z})^*$ is cyclic of ordre $N-1$:

$$\left.\begin{array}{l} a^{N-1} \equiv 1 \bmod N \\ \forall\, p \mid N-1,\ a^{\frac{N-1}{p}} \not\equiv 1 \bmod N \end{array}\right\} \Rightarrow N \text{ is prime}$$

Certificate: $(N, \{p \mid N-1\}, a) \Rightarrow$**isPrime?**$\in$**NP**.

Thm. (**Pocklington**, 1914) Let s s.t. $s \mid N-1$

$$\left.\begin{array}{l} a^{N-1} \equiv 1 \bmod N \\ \forall\, q \text{ prime } \mid s, \gcd(a^{\frac{N-1}{q}} - 1, N) = 1 \end{array}\right\} \Rightarrow \forall\, p \mid N,\ p \equiv 1 \bmod s$$

Coro. $s > \sqrt{N} \Rightarrow N$ is prime.

Example of use

Hyp. We know how to find all prime factors < 20.

$$\begin{array}{rcrl} N_0 & = & 100003, & N_0 - 1 = 2 \times 3 \times 7 \times N_1, \\ N_1 & = & 2381, & N_1 - 1 = 2^2 \times 5 \times 7 \times 17 \end{array}$$

p	2	5	7	17
$3^{(N_1-1)/p} \bmod N_1$	2380	1347	1944	949

$\Rightarrow N_1$ is prime

$$s = N_1 > \sqrt{N_0}$$

$$2^{N_0-1} \equiv 1 \bmod N_0, \gcd(2^{(N_0-1)/N_1} - 1, N_0) = 1$$

$\Rightarrow N_0$ is prime

Rem. We have got a (recursive) primality proof of depth $O(\log N)$.

Compositeness

Fermat: if $\gcd(a, N) = 1$, then $a^{N-1} \equiv 1 \bmod N$.

But: $2^{340} \equiv 1 \bmod 341$: pseudoprime to base 2 (psp-2).

Thm. There exists an infinite number of psp-2 numbers.

Def. $P(N) = \#\{a \in (\mathbb{Z}/N\mathbb{Z})^*, a^{N-1} \equiv 1 \bmod N\}$.

Thm. If $N = \prod_i p_i^{\alpha_i}$, $P(N) = \prod_i \gcd(p_i - 1, N-1)$.

Proof: ■

Thm. There exists an infinite number of psp-a numbers for all possible a's (Carmichael numbers: 561, etc.), i.e., $P(N) = \varphi(N)$.

The test

function isComposite(N)

1. Choose a at random in $\mathbb{Z}/N\mathbb{Z} - \{0\}$.
2. Compute $g = \gcd(a, N)$; **if** $g > 1$, **then** return $(\texttt{yes}, g \mid N)$.
3. **if** $a^{N-1} \not\equiv 1 \bmod N$, **then** return $(\texttt{yes}, a)$
   otherwise return I don't know.

Cost. $O((\log N)\mathsf{M}(\log N))$; typically $O((\log N)^3)$, asymptotically $\tilde{O}((\log N)^2)$.

Prop. Proba("I don't know") $= P(N)/(N-1)$.

Proof. Probability of yes is:

$$\left(1 - \frac{\varphi(N)}{N-1}\right) + \frac{\varphi(N)}{N-1}\left(1 - \frac{P(N)}{\varphi(N)}\right). \square$$

Rem. if N is prime, proba is 1...!

Improvement: Miller-Rabin

Idea: N being odd, write $N-1=2^s t$ with $s \geq 1$ and odd t.

$$a^{N-1}-1=(a^t-1)(a^t+1)(a^{2t}+1)\cdots(a^{2^{s-1}t}+1)$$

$$(MR_a): a^t \equiv 1 \bmod N \text{ or } \exists j, 0 \leq j < s, a^{2^j t} \equiv -1 \bmod N.$$

Pb: $N=2047=23\times 89$ is s.t. $N-1=2\times 1023$ and $2^{(N-1)/2}\equiv 1 \bmod N$: strong-pseudoprime to base 2 (spsp-2).

Def. $F(N)=\#\{a \in (\mathbb{Z}/N\mathbb{Z})^*, (MR_a) \text{ is satisfied}\}$.

Thm. [Monier] $F(N)/(N-1) \leq 1/4$.

Building primes?

function randomProbablePrime(b)

repeat
  choose odd N at random in $[2^{b-1}, 2^b[$
until N passes k tests.

$p_{b,k}=$ Proba$(X=N \text{ is composite} | Y_k=N \text{ passes } k \text{ tests})=$?

Rem. What we know is
Proba$(Y_k=N \text{ passes } k \text{ tests} | X=N \text{ is composite}) \leq (1/4)^k$.

Thm. (Burthe, 1996) $\forall b \geq 2, \forall k \geq 1, p_{b,k} \leq 4^{-k}$.

C) Application to factorization: Pollard's $p-1$

- Invented by Pollard in 1974.
- Williams: $p+1$.
- Bach and Shallit: Φ_k factoring methods.
- Shanks, Schnorr, Lenstra, etc.: quadratic forms.
- Lenstra (1985): ECM.

Rem. Almost all the ideas invented for the classical $p-1$ can be transposed to the other methods.

First phase

Idea: assume $p \mid N$ and a is prime to p. Then

$$(p \mid a^{p-1}-1 \text{ and } p \mid N) \Rightarrow p \mid \gcd(a^{p-1}-1, N).$$

Generalization: if R is known s.t. $p-1 \mid R$,

$$p \mid \gcd((a^R \bmod N)-1, N).$$

How do we find R? Only reasonable hope is that $p-1 \mid B!$ for some (small) B. In other words, $p-1$ is B-smooth.

Algorithm: $R=\prod_{p^\alpha \leq B_1} p^\alpha = \text{lcm}(2,\ldots,B_1)$.
Ex.

k	$b=2^{k!} \bmod 143$	$\gcd(b-1,143)$
2	4	1
3	64	1
4	27	13

since $13-1=2^2\times 3$.

Second phase: the classical one

Let $b = a^R \bmod N$ and $\gcd(b, N) = 1$.
Hyp. $p - 1 = Qs$ with $Q \mid R$ and s prime, $B_1 < s \leq B_2$.

Test: is $\gcd(b^s - 1, N) > 1$ for some s.

$s_j = j$-th prime. In practice all $s_{j+1} - s_j$ are small (Cramer's conjecture implies $s_{j+1} - s_j \leq (\log B_2)^2$).

- Precompute $c_\delta \equiv b^\delta \bmod N$ for all possible δ (small);
- Compute next value with one multiplication $b^{s_{j+1}} = b^{s_j} c_{s_{j+1}-s_j} \bmod N$.

Cost: $O((\log B_2)^2) + O(\log s_1) + (\pi(B_2) - \pi(B_1))$ multiplications $+(\pi(B_2) - \pi(B_1))$ gcd's. When $B_2 \gg B_1$, $\pi(B_2)$ dominates.

Rem. We need a table of all primes $< B_2$; memory is $O(B_2)$.

Record. Nohara (66dd of $960^{119} - 1$, 2006; see `http://www.loria.fr/~zimmerma/records/Pminus1.html`).

Second phase: BSGS

Select $w \approx \sqrt{B_2}$, $v_1 = \lceil B_1/w \rceil$, $v_2 = \lceil B_2/w \rceil$.

Write our prime s as $s = vw - u$, with $0 \leq u < w$, $v_1 \leq v \leq v_2$.

Lem. $\gcd(b^s - 1, N) > 1$ iff $\gcd(b^{wv} - b^u, N) > 1$.

Algorithm:

1. Precompute $b^u \bmod N$ for all $0 \leq u < w$.
2. Precompute all$(b^w)^v$ for all $v_1 \leq v \leq v_2$.
3. For all u and all v evaluate $\gcd(b^{vw} - b^u, N)$.

Number of multiplications: $w + (v_2 - v_1) + O(\log_2 w) = O(\sqrt{B_2})$

Memory: $O(\sqrt{B_2})$.

Number of *gcd*: $\pi(B_2) - \pi(B_1)$.

Second phase: using fast polynomial arithmetic

Algorithm:
1. Compute $h(X) = \prod_{0 \leq u < w} (X - b^u) \in \mathbb{Z}/N\mathbb{Z}[X]$
2. Evaluate all $h((b^w)^v)$ for all $v_1 \leq v \leq v_2$.
3. Evaluate all $\gcd(h(b^{wv}), N)$.

Analysis:

Step 1: $O((\log w)\mathsf{M}_{\mathrm{pol}}(w))$ operations (using a product tree).

Step 2: $O((\log w)\mathsf{M}_{\mathrm{int}}(\log N))$ for b^w; $v_2 - v_1$ for $(b^w)^v$; multi-point evaluation on w points takes $O((\log w)\mathsf{M}_{\mathrm{pol}}(w))$.

Rem. Evaluating $h(X)$ along a geometric progression of length w takes $O(w \log w)$ operations (see Montgomery-Silverman).

Total cost: $O((\log w)\mathsf{M}_{\mathrm{pol}}(w)) = O\left(B_2^{0.5+o(1)}\right)$.

Trick: use $\gcd(u, w) = 1$ and $w = 2 \times 3 \times 5 \ldots$.

Exercises

Program all these versions and try to factor some numbers, e.g., those of the web page.

V. Finite fields

A) General results

Thm. (characteristic) Let $\mathbb{F}$ be a finite field.
a) There exists a smallest $p > 1$ s.t. $p.1_{\mathbb{F}} = 0$; p is prime.
b) The set $\{k.1_{\mathbb{F}}, 0 \leq k < p\}$ is the smallest subfield of $\mathbb{F}$; it is isomorphic to $\mathbb{F}_p$ (prime subfield of $\mathbb{F}$).

Thm.

$$\begin{array}{ccc} \mathbb{F} \times \mathbb{F} & \to & \mathbb{F} \\ (x,y) & \mapsto & x+y \end{array} \quad \text{and} \quad \begin{array}{ccc} \mathbb{F}_p \times \mathbb{F} & \to & \mathbb{F} \\ (a,x) & \mapsto & ax \end{array}$$

turn $\mathbb{F}$ into a $\mathbb{F}_p$-vector space. If n is the dimension of this space, $\mathbb{F}$ has p^n elements.

Thm. $\mathbb{F}^\times$ is cyclic.

Frobenius

Thm.

$$\begin{array}{rccc} \sigma_F : & \mathbb{F} & \to & \mathbb{F} \\ & x & \mapsto & x^p. \end{array}$$

It is a field automorphism, i.e.

$$\sigma_F(1) = 1, \quad \sigma_F(x+y) = \sigma_F(x) + \sigma_F(y), \quad \sigma_F(xy) = \sigma_F(x)\sigma_F(y).$$

Fixed points are the elements of $\mathbb{F}_p$.
Proof: (hint: Lucas)

Building finite fields

Thm. (the canonical way)
Let $f(X)$ be an irreducible polynomial of degree n over $\mathbb{F}_p$.
Then $\mathbb{F}_p[X]/(f(X))$ is a finite field of degree n and cardinality p^n, noted $\mathbb{F}_{p^n}$.

You are kindly invited/strongly encouraged to attend A. Canteaut's lab on finite fields on Thursday October, 5th, 10h15ff. See also 2-13-2 homepage for lecture notes on finite fields.

Quadratic reciprocity (1/2)

Legendre symbol: for prime odd p and $a \in \mathbb{Z}$

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p \mid a \\ 1 & \text{if } \exists\, x \text{ s.t. } a \equiv x^2 \bmod p \\ -1 & \text{otherwise.} \end{cases}$$

Easy properties:
(i) $\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \bmod p$;
(ii) $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$;
(iii) $\left(\frac{a}{p}\right) = \left(\frac{a \bmod p}{p}\right)$;
(iv) $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$;

Quadratic reciprocity (2/2)

Not so easy properties: (Gauss)
(v) $\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}$;
(vi) (Quadratic reciprocity law) p and q odd primes:

$$\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\times\frac{q-1}{2}}\left(\frac{p}{q}\right).$$

Jacobi symbol: $n \in \mathbb{Z}$, $m = \prod_{i=1}^{k} p_i \in \mathbb{Z}$ odd,

$$\left(\frac{n}{m}\right) = \prod_{i=1}^{k}\left(\frac{n}{p_i}\right).$$

Properties: same as for the Legendre symbol.

Ex. Show that $\left(\frac{n}{m}\right) = 0$ iff $\gcd(n,m) > 1$.

Example

Build $\mathbb{F}_{41^2}$, using a quadratic non-residue modulo 41.

$$\begin{aligned}\left(\tfrac{7}{41}\right) &= (-1)^{(41-1)/2\times(7-1)/2}\left(\tfrac{41}{7}\right)\\ &= \left(\tfrac{41}{7}\right) = \left(\tfrac{41 \bmod 7}{7}\right)\\ &= \left(\tfrac{6}{7}\right) = \left(\tfrac{2}{7}\right)\left(\tfrac{3}{7}\right) = \left(\tfrac{3}{7}\right) = (-1)\left(\tfrac{7}{3}\right)\\ &= -\left(\tfrac{1}{3}\right) = -1\end{aligned}$$

$\Rightarrow \mathbf{K}_1 = \mathbb{F}_{41^2} \sim \mathbb{F}_{41}[X]/(X^2-7)$.
This is a vector space of dimension 2 over $\mathbb{F}_{41}$.
Let $\theta = \overline{X}$. All elements can be written $a + b\theta$ where a, b are in $\mathbb{F}_{41}$.
$\theta^2 - 7 = \overline{X^2-7} = 0$. We get

$$\theta^2 = 7, \theta^3 = 7\theta, \theta^4 = 8, \ldots, \theta^{80} = 1,$$

so that θ does not generate $\mathbf{K}^*$, but $\theta + 10$ does.

Application (1/2)

Pb. Given $\left(\frac{a}{p}\right) = 1$, compute $\sqrt{a} \bmod p$.

Case $p \equiv 3 \bmod 4$**:** $r = a^{(p+1)/4} \bmod p$.

Case $p \equiv 1 \bmod 4$**:** find b s.t. $\Delta = b^2 - 4a$ is not a square.

$$\alpha = (-b + \sqrt{\Delta})/2 \quad \Rightarrow \alpha^p = (-b - \sqrt{\Delta})/2 \quad \Rightarrow \alpha\alpha^p = a$$

since $\sqrt{\Delta}^p = \left(\frac{\Delta}{p}\right)\sqrt{\Delta}$.
Let $\beta = \alpha^{(p+1)/2} \bmod (p, X^2 + bX + a)$. Then

$$\beta^2 = \alpha^{p+1} = a;$$

$$\beta^p = \beta(\beta^2)^{(p-1)/2} = \beta a^{(p-1)/2} = \beta$$

$\Rightarrow \beta \in \mathbb{F}_p$.

Application (2/2)

Let $a = 2 \bmod 41$, which is a square;
$b = 1$ is s.t. $\Delta = 1 - 4 \times 2 = -7$ which is not a square; hence
$\mathbb{F}_{41^2} \sim \mathbb{F}_{41}[X]/(X^2 + X + 2)$.

$$\alpha = X, \quad \alpha^p = 40X + 40, \quad \alpha\alpha^p = 2.$$

$$\beta = X^{(p+1)/2} = 17, \quad 17^2 \equiv 2 \bmod 41.$$

Back to compositeness: Solovay-Strassen

Euler: if N is prime and $\gcd(a,N)=1$, then $a^{(N-1)/2} \equiv \left(\frac{a}{N}\right) \bmod N$.

Pb: $2^{(1105-1)/2} \equiv \left(\frac{2}{1105}\right) \bmod 1105$; this is an Euler pseudoprime to base 2 (epsp-2). There are an infinite number of them.

Def. $\mathcal{E}(N) = \{a \in (\mathbb{Z}/N\mathbb{Z})^*, a^{(N-1)/2} \equiv \left(\frac{a}{N}\right) \bmod N\}$; $E(N) = \#\mathcal{E}(N)$.

Them. $E(N)/\varphi(N) \leq 1/2$.

Prop. Proba("I don't know")$= E(N)/(N-1) \leq 1/2$.

Coro. isComposite? $\in$ **RP** (hence **isPrime?**$\in$ **co** $-$ **RP**).

B) En route for **P**

- Gauss and Jacobi sums: L. Adleman, C. Pomerance, S. Rumely (1980, 1983); H. Cohen, H. W. Lenstra, Jr (1981 – 1984) ; H. Cohen, A. K. Lenstra (1982, 1987). W. Bosma & M.-P. van der Hulst (1990) ; P. Mihăilescu (1998). deterministic $O((\log N)^{c_1 \log\log\log N})$.
- almost **RP**: Goldwasser and Kilian using elliptic curves (1986); practical algorithm by Atkin (1986; later FM). See Smith's part.
- **RP**: Adleman and Huang using hyperelliptic curves (1986ff). See Smith's part.

Agrawal, Kayal, Saxena (AKS)

First idea: (Agrawal, Biswas – 1999)

Prop. N is prime iff $P(X) = (X+1)^N - X^N - 1 \equiv 0 \bmod N$.

In pratice: choose $Q(X) \in \mathbb{Z}/N\mathbb{Z}[X]$ at random of degree $O(\log N)$. If

$$(X+1)^N \not\equiv X^N + 1 \bmod (Q(X), N)$$

then N is composite.

The probability of failure is bounded by $1 - 1/(4\log N)$.

Conjecture: If N is composite, there exists $1 \leq r \leq \log N$ s.t. $P(X)$ is not divisible by $X^r - 1$ modulo N.

Agrawal, Kayal, Saxena

Thm. Let N, s be integers, r a prime number and $q = P(r-1)$. If:

(0)
$$\binom{q-1+s}{s} > N^{2\lfloor\sqrt{r}\rfloor};$$

(i) $N \neq M^k$, $k > 1$;

(ii) N has no prime factor $\leq s$;

(iii) $N^{(r-1)/q} \bmod r \notin \{0, 1\}$;

(iv) $\forall a, 1 \leq a \leq s$, $(X-a)^N \equiv X^N - a \bmod (X^r - 1, N)$;

then N is prime.

For a proof, see FM's Bourbaki article.

What next?

- cf. D. Bernstein homepage for more on the history of improvements to the basic test; iIncluding: H. W. Lenstra, Jr. ($\tilde{O}_{\mathit{eff}}((\log N)^{12})$ or $\tilde{O}((\log N)^8)$), S. David.
- Cleaner version of AKS: $\tilde{O}_{\mathit{eff}}((\log N)^{10.5})$ or $\tilde{O}((\log N)^{7.5})$.
- H. W. Lenstra, C. Pomerance : $\tilde{O}_{\mathit{eff}}((\log N)^6)$.
- P. Berrizbeitia / Q. Cheng :
  Let r prime s.t. $r^\alpha \mid\mid N-1$, $r \geq \log^2 N$; $1 < a < N$ s.t. $a^{r^\alpha} \equiv 1 \bmod N$, $\gcd(a^{r^{\alpha-1}} - 1, N) = 1$, $(X+1)^N = X^N + 1 \bmod (X^r - a, N)$, then N is prime. Heuristic complexity would be $\tilde{O}((\log N)^4)$ for these numbers.
- D. Bernstein, P. Mihăilescu: use $e \mid N^d - 1$; inject cyclotomic ideas, $\tilde{O}((\log N)^4)$.

Conclusions for primality

Which algorithm?

- easy to understand / implement, fast: compositeness tests;
- fast, proven: Jacobi;
- fast, heuristic: ECPP;
- certificate: ECPP;
- deterministic polynomial: AKS.

D. Bernstein has an AKS example for $2^{1024} + 643$ (13 hours on 800 MHz PC, 200 Mb memory).

To be compared to FASTECPP:
14/07/03: FM, 7000dd with mpifastECPP.
19/08/03: J. Franke, T. Kleinjung, T. Wirth, 10000dd.
15/10/10: FM, 25,050 dd with mpifastECPP (2000 CPU days).
Rem. 2013: Franke et al., 30,008 dd with combination CIDE.

VI. Generic DLP

DLP: given $h \in G = \langle g \rangle$ of order N, find an integer n, $0 \leq n < N$ such that $h = g^n$.

Z) The Pohlig-Hellman reduction.

A) Enumeration; baby-steps, giant steps (adaptation as exercises).

B) RHO.

C) The theoretical side.

Z) The Pohlig-Hellman reduction

Idea: reduce the problem to the case N prime.

$$N = \prod_i p_i^{\alpha_i}$$

Solving $g^n = h$ is equivalent to knowing $n \bmod N$, i.e. $n \bmod p_i^{\alpha_i}$ for all i (chinese remainder theorem).

Idea: let $p^\alpha \mid\mid N$ and $m = N/p^\alpha$. Then $b = h^m$ is in the cyclic group of ordre p^α generated by g^m. We can find the log of b in this group, which yields $n \bmod p^\alpha$.

Cost: $O(\max(DL(p^\alpha))) = O(\max(DL(p)))$.

Consequence: for DH, N must have at least one large prime factor.

B) The RHO method

Basic model: birthday paradox
Let E be a finite set of cardinality m.

Thm. Suppose we draw uniformly n elements from E with replacement. The probability that all n elements are distinct is $\text{Proba} = \frac{1}{m}\prod_{k=1}^{n-1}\left(1-\frac{k}{m}\right)$.

Taking logarithms, and assuming $n \ll m$, we get

$$\log \text{Proba} \approx \log(n/m) - \frac{n(n-1)}{2m}.$$

$\Rightarrow$ taking $n = O(\sqrt{m})$ will give a somewhat large value for this probability.

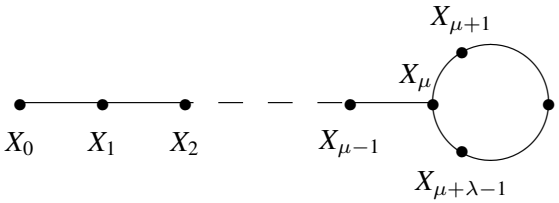
A very simple algorithm

Function *NaiveDL*(G, g, N, h)
  Input : $G \supset \langle g\rangle$, g of order N
  Output: $0 \le n < N$, $g^n = h$
  initialize a table $\mathcal{L}$ for storing u triplets (elt of G, two ints $< N$);
  repeat
    draw u and v at random modulo N;
    $H \leftarrow g^u \circ h^v$;
    if $\exists (H', u', v') \in \mathcal{L}$ *such that* $H = H'$ **then**
      // $H = g^u \circ h^v = g^{u'} \circ h^{v'}$
      // hence $n(v - v') = u' - u \bmod N$
      if $v - v'$ *is invertible modulo* N **then**
        return $(u' - u)/(v - v') \bmod N$;
    else
      store($\mathcal{L}$, (H, u, v));
  until *a collision is found*;

Complexity: $O(\sqrt{n}\log n)$ on average, together with a space $O(\sqrt{n})$, which is no better than BSGS.

Functional digraphs

Let $f : E \to E$ be a function on E.
Consider $X_{n+1} = f(X_n)$ for some starting point $X_0 \in E$.
The functional digraph of X is built with vertices X_i's; an edge is put between X_i and X_j if $f(X_i) = X_j$.

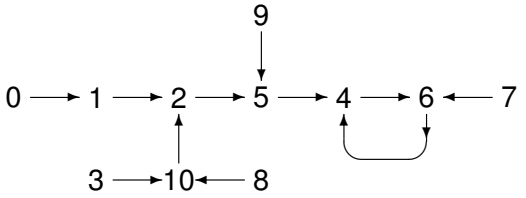


The first part of the sequence is the set of X_i's that are reached only once and there are μ of them.
The second part forms a loop containing λ distinct elements.
Rem. λ and ν cannot be too large on average (use $n = \lambda + \mu$ in the Theorem).

Examples

1) $E = G$ finite group, $f(x) = ax$ and $x_0 = a \Rightarrow (x_n)$ purely is periodic, i.e., $\mu = 0$, and $\lambda = \text{ord}_G(a)$.

2) Take $E = \mathbb{Z}/11\mathbb{Z}$ and $f : x \mapsto x^2 + 1 \bmod 11$



Typical shape: a cycle and trees plugged on the structure.

Epact

Goal: find λ and μ.

Prop. There exists a unique $e > 0$ (epact) s.t. $\mu \leq e < \lambda + \mu$ and $X_{2e} = X_e$.
It is the smallest non-zero multiple of λ that is $\geq \mu$: if $\mu = 0$, $e = \lambda$ and if $\mu > 0$, $e = \lceil \frac{\mu}{\lambda} \rceil \lambda$.

Proof:

Floyd's algorithm

Function *epact(*f, x_0*)*
  Input : A function f, a starting point x_0
  Output: The epact of (x_n) defined by $x_{n+1} = f(x_n)$
  $x \leftarrow x_0$; $y \leftarrow x_0$; $e \leftarrow 0$;
  repeat
    $e \leftarrow e + 1$;
    $x \leftarrow f(x)$;
    $y \leftarrow f(f(y))$;
  until $x = y$;
  return e.

Cost: $3e$ evaluations of f and e comparisons. For decreasing the number of evaluations, see Brent (and Montgomery).

Asymptotic statistics

Convenient source: Flajolet & Odlyzko (EUROCRYPT 1989).

Thm. When $m \to \infty$

$$\overline{\lambda} \sim \overline{\mu} \sim \sqrt{\frac{\pi m}{8}} \approx 0.627\sqrt{m}.$$

Thm. $\overline{e} \sim \sqrt{\frac{\pi^5 m}{288}} \approx 1.03\sqrt{m}$.

Fundamental coro. A collision is expected to be found after $O(\sqrt{m})$ computations.

Application to DL

Pollard: build a function f from G to G appearing to be random, i. e., the epact of f is $c\sqrt{N}$ for some small c.

. . . **Teske**:

- precompute r random elements $z_i = g^{\gamma_i} \circ h^{\delta_i}$ for $1 \leq i \leq r$ for some random exponents (say $r = 20$);
- use some hash function $\mathcal{H} : G \to \{1, \ldots, r\}$;
- define $f(y) = y \circ z_{\mathcal{H}(y)}$ so that

$$x_i = g^{c_i} \circ h^{d_i},$$

  where (c_i) and (d_i) are two integer sequences.

Ex. if G contains integers, we may simply use $\mathcal{H}(x) = 1 + (x \bmod r)$.

Application to DL (cont'd)

When e is found:

$$g^{c_{2e}} \circ h^{d_{2e}} = g^{c_e} \circ h^{d_e}$$

or

$$g^{c_{2e}-c_e} = h^{d_e-d_{2e}}$$

i.e.,

$$n(c_{2e} - c_e) \equiv (d_e - d_{2e}) \bmod N.$$

Function *Iterate(G, N, $\mathcal{H}$, $(z_i, \gamma_i, \delta_i)$, x, u_x, v_x)*
- **Input** : $\mathcal{H} : G \to \{1, \ldots, r\}$; $(z_i)_{1 \le i \le r}$ random powers $z_i = g^{\gamma_i} \circ h^{\delta_i}$ of G; $x = g^{u_x} h^{v_x}$
- **Output:** $f(x, u_x, v_x) = (w, u_w, v_w)$ s.t. $w = g^{u_w} \circ h^{v_w}$
- $i \leftarrow \mathcal{H}(x)$;
- **return** $(x \circ z_i, u_x + \gamma_i \pmod N, v_x + \delta_i \pmod N)$.

The algorithm

Function *RHO(G, g, N, h, $\mathcal{H}$, $(z_i, \gamma_i, \delta_i)$)*
- **Input** : $\mathcal{H} : G \to \{1, \ldots, r\}$; $(z_i)_{1 \le i \le r}$ random powers $z_i = g^{\gamma_i} \circ h^{\delta_i}$ of G
- **Output:** $0 \le n < N$, $g^n = h$
- **if** $h = 1_G$ **then**
  - **return** 0
- $x \leftarrow h$; $u_x \leftarrow 0$; $v_x \leftarrow 1$;
- $y \leftarrow x$; $u_y \leftarrow u_x$; $v_y \leftarrow v_x$;
- **repeat**
  - // invariant: $x = g^{u_x} \circ h^{v_x}$, $y = g^{u_y} \circ h^{v_y}$
  - $(x, u_x, v_x) \leftarrow$ Iterate(G, N, $\mathcal{H}$, $(z_i, \gamma_i, \delta_i)$, x, u_x, v_x);
  - $(y, u_y, v_y) \leftarrow$ Iterate(G, N, $\mathcal{H}$, $(z_i, \gamma_i, \delta_i)$, y, u_y, v_y);
  - $(y, u_y, v_y) \leftarrow$ Iterate(G, N, $\mathcal{H}$, $(z_i, \gamma_i, \delta_i)$, y, u_y, v_y);
- **until** $x = y$;
- // $g^{u_x} \circ h^{v_x} = g^{u_y} \circ h^{v_y}$
- **if** $v_x - v_y$ *is invertible modulo* N **then**
  - **return** $(u_y - u_x)/(v_x - v_y) \pmod N$;
- **else**
  - **return** *Failure*.

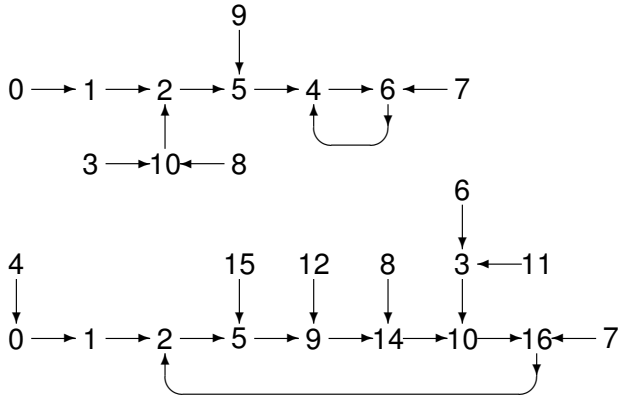
A variant for factoring integers

Pollard's idea: suppose $p \mid N$ and we have a random $f \bmod N$ s.t. $f \bmod p$ is "random".

Function *RHO(N)*
- **Input** : N an integer
- **Output:** a factor of N
- $x \leftarrow 1$; $y \leftarrow 1$; $g \leftarrow 1$;
- **while** $g = 1$ **do**
  - $x \leftarrow f(x, N)$;
  - $y \leftarrow f(f(y, N), N)$;
  - $g \leftarrow \gcd(x - y, N)$;
- **return** g.

Pollard's RHO: factoring $N = 187$

i	x_i	y_i	$\gcd(x_i - y_i, N)$
1	1	2	1
2	2	26	1
3	5	180	1
4	26	70	11



C) The theoretical side

Thm. (Self-reducibility) Let $G = \langle g \rangle$ be a cyclic group of prime order p. Suppose that there exists $A \subset G$ such that DLP is solvable in time T in A. Then DLP is solvable in G with time $O((|A|/|G|)T)$.

Proof: Given $h \in G$, draw elements g^r uniformly at random until $h \circ g^r$ belongs to A, from which the discrete logarithm of h is computed in time T. The time it takes to reach A is proportional to $|A|/|G|$. $\square$

Thm. (Nechaev/Shoup) Any algorithm operating on a generic group of cardinality n takes times $\Omega(\sqrt{n})$.

Thm. DHP: $(g^x, g^y) \leftarrow g^{xy}$.

- DLP $\Rightarrow$ DHP;
- Maurer-Wolf: DHP $\Rightarrow$ DLP for a large number of groups.

Take home messages

To get a feeling, program everything for the simple case of $(\mathbb{Z}/p\mathbb{Z})^*$.

To have a better than square-root algorithm for DL, you need specific ideas for specific groups. $\Rightarrow$ Barbulescu's part

Many crypto problems of size n may have solution algorithms in $O(\sqrt{n})$ (time and/or space; deterministic or probabilistic).