Proof of lemma of AKS algorithm

lemma: suppose $a, n \in \mathbb{Z}$ and $\gcd(a, n) = 1$ then n is prime iff $(X+a)^n \equiv (X^n + a) \bmod n$

Proof:

The coefficient of X^i in $((X+a)^n - (X^n + a))$ is ${}^nC_i \cdot a^{n-i}$

where,

$0 < i < n$

Let suppose n is prime, then ${}^nC_i \equiv 0 \bmod n$

Hence all coefficients are zero and we are done.

Now, let's prove in ~~another~~ backward direction,

let us suppose n is composite, let us consider a prime q that is a factor of n and $q^k | n$, where $k \geq 1$

Let $n = q^k \cdot t$

Here q^k does not divide $\binom{n}{q}$ and coprime to q^q { since $\binom{n}{q} = \frac{n(n-1)\cdots(n-q+1)}{q!}$, when numerator is divisible by q^k and not by q^{k+1} and denominator is divisible by q.

Therefore coefficient of x^{n-q} is not ~~0 mod~~ 0 mod n. thus $(x+a)^n - (x^n + a)$ is not identically zero over $\mathbb{Z}/n\mathbb{Z}$.

Hence our lemma proved.

Fermat Factorization and Factor bases

Fermat factorization

Way to factor a composite number n is efficient if n is a product of two integers which are close to one another. This method called "Fermat Factorization" is based on the fact that n is equal to product of 2 squares (one of which is very small).

Proposition:

Let n be a positive odd integer. There is a 1-to-1 correspondence between factorizations of n in form of ab where $a \geq b > 0$ and representation of n in form $t^2 - s^2$ where s & t are non-negative integers. The correspondence is given by the equations:

$$n = ab$$

$$n = t^2 - s^2$$

$$t = \frac{a+b}{2}, \quad s = \frac{a-b}{2}$$

$$a = t + s$$
$$b = t - s$$

Proof:

$$n = ab$$
$$= \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2$$
$$= t \qquad = s$$
$$= t^2 - s^2$$

we can factor $t^2 - s^2$ as $(t+s)(t-s)$

Case 1:

NOTE: If $n = ab$ with a & b close together then $s = \frac{a-b}{2}$ is small, so t is only slightly larger than $\sqrt{n}$

$$n = t^2 - s^2$$
$$n + s^2 = t^2$$
$$t = \sqrt{n+s^2}$$ slightly greater than $\sqrt{n}$

In this case, we can find a and b by trying all values for t starting $[\sqrt{n}]+1$, until we find one for which $t^2 - n = s^2$ is a perfect square.

Greatest integer function: denoted by [ ], it gives the integer less than or equal to number in brackets.

$[448.12] = 448$

Ex 1) Factor 200819 ?

Soln: we have to starting from $[\sqrt{n}]+1$

$n = 200819$

$[\sqrt{n}]+1 = [\sqrt{200819}]+1 = 448+1 = 449$

$t = 449$: $t^2 - n = 449^2 - 200819 = 782$

not perfect square

increment t by 1 :

$t = 450$: $t^2 - n = 450^2 - 200819 = 1681$

perfect square

Therefore,

$200819 = 450^2 - 1681$

$= 450^2 - 41^2$

$= (450+41)(450-41)$

$= 491 \cdot 409$

Therefore 491 and 409 are the factors of 200819.

Case 2: If a and b are not close together, for any factorization of $n = ab$, then the Fermat factorization method will find a and b, but after trying many values of $t = [\sqrt{n}]+1$, $[\sqrt{n}]+2$ - - - continued.

There is a generalization of Fermat factorization which works better in such a situation:

→ Choose a small k, successively set $t = [\sqrt{kn}]+1$, $[\sqrt{kn}]+2$, etc. until we get a t such that $t^2 - kn = s^2$ is perfect square.

→ Now $(t+s)(t-s) = kn$, so $t+s$ has a non-trivial common factor with n which can be found by $\boxed{gcd(t+s, n)}$

Ex 2) Factor 141467?

Soln: If we go by fermat factorization algo, we have to set t for many values.

$[\sqrt{n}]+1 = [\sqrt{141467}]+1 = 377$

$t = 377, 378$ ---- till we get tire of finding t.

Now, we apply the algorithm of generalization of Fermat Factorization as follows:

(i) let $k = 3$

$t = [\sqrt{kn}]+1$

$= [\sqrt{3 \times 141467}]+1 = 652$

$652^2 - 141467 \times 3 = 703$ → not perfect square

$t = t+1$

$t = 653$: $653^2 - 3 \times 141467 = 2008$ → not perfect square

$t = t+1$: $654 = 654^2 - 3 \times 141467 = 3315$

↙ not perfect square

$t = t+1$: $t = 655 \Rightarrow 655^2 - 3 \times 141467$

$= 4624$ = Perfect square

$s^2 = 4624 \Rightarrow s = 68$

Now compute $\gcd(t+s, n)$

$= \gcd(655 + 68, 141467)$

$= \gcd(723, 141467)$

$= \gcd(723, 141467) = 241$

Therefore $= 141467 = 241 \times 587$

141467 | 241
=587

Complexity Analysis:

with $k=3$, generalized Fermat factorization, we need to try only four t's whereas with simple factorization, it would take 38 t's. ($k=1$)

Another efficient factoring method

Find a congruence of the form $t^2 \equiv s^2 \bmod n$ with $t \not\equiv \pm s \bmod n$.

We immediately find a factor of n by computing $\gcd(t+s, n)$ (OR $\gcd(t-s, n)$).

Ex3) Factor 4633?

Soln: By Hit & Try, we get:

118^2 leaves a remainder 25 when divided by 4633.

Therefore, the congruence equation is:

$$118^2 \equiv 25 \bmod 4633$$
$$118^2 \equiv 5^2 \bmod 4633$$
$$t = 118,\ s = 5$$

Factors are: $\gcd(t+s, n)$, $\gcd(t-s, n)$

$= \gcd(118+5, 4633)$, $\gcd(118-5, 4633)$

$= \gcd(123, 4633)$, $\gcd(113, 4633)$

$= 41, 113$

$$4633 = 41 \cdot 113$$

Definitions:

least absolute residue = The least absolute residue of a number a mod n is the integer in the interval $-n/2$ to $n/2$ to which a is congruent.

Factor Base = A factor base is a set $B = \{p_1, p_2 - - p_n\}$ of distinct primes except that p_1 may be the integer -1.

B-number = we say that square of an integer b is a B-number (for a given n) if the least absolute residue of $b^2 \bmod n$ can be written as a product of numbers from B.

Ex 4) $n = 4633$, $B = \{-1, 2, 3\}$, Find the B-numbers?

Soln: we can see that:

$$67^2 = -144 \bmod 4633$$
$$68^2 = -9 \bmod 4633$$
$$69^2 = 128 \bmod 4633$$

67, 68, 69 are B-numbers since:
- the least absolute residues $(-144, -9, 128)$ are product of numbers from B.

$$-144 = (-1)^1 \times 2^4 \times 3^2$$
$$-9 = (-1)^1 \times 2^0 \times 3^2$$
$$128 = (-1)^0 \times 2^7 \times 3^0$$

Vector representation of B-numbers

Given n and a factor base B containing h numbers, To correspond a vector $\in$ to every B-number follow the steps:

① write $b^2 \bmod n$ in form of $\prod_{j=1}^{h} p_j^{\alpha_j}$

② Set the jth component $\in_j$ equal to $\alpha_j \bmod 2$

$\in_j = 0$ if α_j = even

$\in_j = 1$, if α_j = odd

Ex5) Find the vector form of the B-numbers for $n = 4633$ & $B = \{-1, 2, 3\}$?

Soln: B-numbers are:

$67^2 = -144 \bmod 4633$

$68^2 = -9 \bmod 4633$

$69^2 = 128 \bmod 4633$

writing $b^2 \bmod n$ in form of $\prod_{j=1}^{h} P_j^{a_j}$

$-144 \bmod 4633 = (-1)^1 \times 2^4 \times 3^2$

$\vec{\epsilon}_1 = (1, 0, 0)$

$-9 \bmod 4633 = (-1)^1 \times 2^0 \times 3^2$

$\vec{\epsilon}_2 = (1, 0, 0)$

$128 \bmod 4633 = (-1)^0 \times 2^7 \times 3^0$

$\vec{\epsilon}_3 = (0, 1, 0)$

Algorithm

Factor Base factoring:

(1) find the B-numbers by the method discussed above.

(2) find 2 numbers such that their $(\text{product})^2$ gives t and s

$(t \neq s)$

if $t = s$, we cannot find factor since the gcd $(t+s, n)$ will come out to be 1.

(3) Calculate gcd $(t+s, n)$ for the factors.

Ex6) Using Factor Base Algorithm, factorize $n = 2043221$ using factor base $B = \{2, 3, 5, 7, 11\}$?

Soln:

$n = 2043221$

$B = \{2, 3, 5, 7, 11\}$

finding the B-numbers:

$1439^2 \bmod 2043221 = 27500 = 2^2 5^4 11$

$2878^2 \bmod 2043221 = 110000 = 2^4 5^4 11$

$3197^2 \bmod 2043221 = 4704 = 2^5 . 3 . 7^2$

$3199^2 \bmod 2043221 = 17496 = 2^3 3^7$

$3253^2 \bmod 2043221 = 365904 = 2^4 . 3^3 . 7 . 11^2$

let us consider the 3rd & 4th number:

$[(3197 \times 3199)]^2 \equiv 2^8 3^6 7^2 \bmod 2043221$

comparing $t^2 \equiv s^2 \bmod n$

$t = 3197 \times 3199 \bmod 2043221 = 11098$

$s = 2^4 \times 3^4 \times 7 \bmod 2043221 = 9072$

$\gcd(t+s, n) = \gcd(11098 + 9072, 2043221)$

$= 2017$

So, we have $2043221 = (2017) \times (1013)$

NOTE:
Be careful in choosing the numbers, if we chose first 2 numbers,

$$[(1439 \times 2878)]^2 = 2^6 5^8 11^2 \bmod n$$

$\Rightarrow t = 1439 \times 2878 \bmod n = 55000$
$s = 2^3 \times 5^4 \times 11 \bmod n = 55000$

$\Rightarrow \boxed{t = s}$

$\Rightarrow \boxed{gcd(t+s, n) = 1}$ which does not lead to factorization