Lenstra Elliptic Curve Factorization

Miles Benton and Skip Moses

MATH 317

2021

Background

- Hendrik Lenstra Jr. recieved his doctorate from the University of Amsterdam in 1977.

Background

- Hendrik Lenstra Jr. recieved his doctorate from the University of Amsterdam in 1977.
- Discovered Elliptic Curve Factorization (ECM) in 1987.

Background

- Hendrik Lenstra Jr. recieved his doctorate from the University of Amsterdam in 1977.
- Discovered Elliptic Curve Factorization (ECM) in 1987.
- ECM is third-fastest known factoring algorithm and the best algorithm for finding divisors not exceeding 50-60 digits.

Background

- Hendrik Lenstra Jr. recieved his doctorate from the University of Amsterdam in 1977.
- Discovered Elliptic Curve Factorization (ECM) in 1987.
- ECM is third-fastest known factoring algorithm and the best algorithm for finding divisors not exceeding 50-60 digits.
- The largest factor found using ECM has 83 digits.

Group Theory behind Pollard $p-1$

Definition Let $(G, +)$ be a group. If $H \subset G$ is also a group under $+$, then we call H a subgroup of G.

Group Theory behind Pollard $p - 1$

Definition Let $(G, +)$ be a group. If $H \subset G$ is also a group under $+$, then we call H a subgroup of G.
Lagrange's Theorem If $H \subset G$ with $|G| < \infty$, then $|H|$ divides $|G|$.

Group Theory behind Pollard $p-1$

Definition Let $(G,+)$ be a group. If $H \subset G$ is also a group under $+$, then we call H a subgroup of G.

Lagrange's Theorem If $H \subset G$ with $|G| < \infty$, then $|H|$ divides $|G|$.

- Define the relation $a \sim b$ whenever $a = bh$ for some $h \in H$.

Group Theory behind Pollard $p-1$

Definition Let $(G, +)$ be a group. If $H \subset G$ is also a group under $+$, then we call H a subgroup of G.
Lagrange's Theorem If $H \subset G$ with $|G| < \infty$, then $|H|$ divides $|G|$.

- Define the relation $a \sim b$ whenever $a = bh$ for some $h \in H$.
- Let S be an equivalence class of $\sim$ and pick an arbitrary $a \in S$.

Definition Let $(G, +)$ be a group. If $H \subset G$ is also a group under $+$, then we call H a subgroup of G.
Lagrange's Theorem If $H \subset G$ with $|G| < \infty$, then $|H|$ divides $|G|$.

- Define the relation $a \sim b$ whenever $a = bh$ for some $h \in H$.
- Let S be an equivalence class of $\sim$ and pick an arbitrary $a \in S$.
- Define $f : S \to H$, where $f(b) = b^{-1}a$ and $g : H \to S$ by $g(h) = ah^{-1}$

Definition Let $(G,+)$ be a group. If $H \subset G$ is also a group under $+$, then we call H a subgroup of G.

Lagrange's Theorem If $H \subset G$ with $|G| < \infty$, then $|H|$ divides $|G|$.

- Define the relation $a \sim b$ whenever $a = bh$ for some $h \in H$.
- Let S be an equivalence class of $\sim$ and pick an arbitrary $a \in S$.
- Define $f : S \to H$, where $f(b) = b^{-1}a$ and $g : H \to S$ by $g(h) = ah^{-1}$
- Observe f and g are inverses.

$$f(g(h)) = f(ah^{-1}) = (ah^{-1})^{-1}a = h$$

$$g(f(b)) = g(b^{-1}a) = a(b^{-1}a)^{-1} = b$$

- Thus, $|S| = |H|$ for all equivalence classes S.

Group Theory behind Pollard $p-1$

Definition Let $(G, +)$ be a group. If $H \subset G$ is also a group under $+$, then we call H a subgroup of G.

Lagrange's Theorem If $H \subset G$ with $|G| < \infty$, then $|H|$ divides $|G|$.

- Define the relation $a \sim b$ whenever $a = bh$ for some $h \in H$.
- Let S be an equivalence class of $\sim$ and pick an arbitrary $a \in S$.
- Define $f : S \to H$, where $f(b) = b^{-1}a$ and $g : H \to S$ by $g(h) = ah^{-1}$
- Observe f and g are inverses.

$$f(g(h)) = f(ah^{-1}) = (ah^{-1})^{-1}a = h$$

$$g(f(b)) = g(b^{-1}a) = a(b^{-1}a)^{-1} = b$$

- Thus, $|S| = |H|$ for all equivalence classes S.
- Therefore, $|G| = n|H|$.

A different perspective

A different perspective

Lemma 2.2.5 Suppose that m, $n \in \mathbb{N}$ and $\gcd(m, n) = 1$. Then the map

$$\psi : (\mathbb{Z}/mn\mathbb{Z})^* \to (\mathbb{Z}/m\mathbb{Z})^* \times (\mathbb{Z}/n\mathbb{Z})^*$$

defined by

$$\psi(c) = (c \pmod{m}, c \pmod{n})$$

is a bijection.

Example

Example

- Let $B_i = \text{lcm}(1, \ldots, i)$.

Example

- Let $B_i = \text{lcm}(1, \ldots, i)$.

B_i	$2^{B_i} \pmod{1763}$	$(2^i \pmod{41}, 2^i \pmod{43})$
1	2	(2, 2)

Example

- Let $B_i = \operatorname{lcm}(1, \dots, i)$.

B_i	$2^{B_i} \pmod{1763}$	$(2^i \pmod{41}, 2^i \pmod{43})$
1	2	(2, 2)
2	4	(4, 4)

Example

- Let $B_i = \text{lcm}(1, \ldots, i)$.

B_i	$2^{B_i} \pmod{1763}$	$(2^i \pmod{41}, 2^i \pmod{43})$
1	2	(2, 2)
2	4	(4, 4)
6	570	(37, 11)

Example

- Let $B_i = \text{lcm}(1, \ldots, i)$.

B_i	$2^{B_i} \pmod{1763}$	$(2^i \pmod{41}, 2^i \pmod{43})$
1	2	(2, 2)
2	4	(4, 4)
6	570	(37, 11)
60	575	(1, 16)

- We compute $gcd(574, 1763) = 41$

Set up for ECM

Set up for ECM

- Let E be an elliptic curve over $\mathbb{Z}/N\mathbb{Z}$ of the form

$$y^2 = x^3 + ax + 1$$

such that $4a^3 + 27 \in (\mathbb{Z}/N\mathbb{Z})^*$. This forces non singularity and ensures $P = (0, 1)$ is on the curve.

Set up for ECM

- Let E be an elliptic curve over $\mathbb{Z}/N\mathbb{Z}$ of the form

$$y^2 = x^3 + ax + 1$$

  such that $4a^3 + 27 \in (\mathbb{Z}/N\mathbb{Z})^*$. This forces non singularity and ensures $P = (0, 1)$ is on the curve.

- Definition 6.3.1 (Power Smooth). Let B be a positive integer. If n is a positive integer with prime factorization

$$n = \prod p_i^{e_i},$$

  then n is B-power smooth if $p_i^{e_i} \leq B$ for all i.

Set up for ECM

- Let E be an elliptic curve over $\mathbb{Z}/N\mathbb{Z}$ of the form

$$y^2 = x^3 + ax + 1$$

  such that $4a^3 + 27 \in (\mathbb{Z}/N\mathbb{Z})^*$. This forces non singularity and ensures $P = (0, 1)$ is on the curve.

- Definition 6.3.1 (Power Smooth). Let B be a positive integer. If n is a positive integer with prime factorization

$$n = \prod p_i^{e_i},$$

  then n is B-power smooth if $p_i^{e_i} \leq B$ for all i.

- Example $30 = 2 \cdot 3 \cdot 5$ is B power smooth for $B \geq 5$, but $150 = 2 \cdot 3 \cdot 5^2$ is not 5-power smooth.

Motivation

Motivation

- Fix $B \in \mathbb{N}$. Let $p \in \mathbb{N}$ such that $p-1$ is not B- power smooth.

Motivation

- Fix $B \in \mathbb{N}$. Let $p \in \mathbb{N}$ such that $p-1$ is not B- power smooth.
- Recall, in Pollard $p-1$, this would be equivalent to not having $p-1 \not| m = \text{lcm}(1, 2, \ldots, B)$; i.e. $a^m \not\equiv 1 \pmod{p}$.

Motivation

- Fix $B \in \mathbb{N}$. Let $p \in \mathbb{N}$ such that $p-1$ is not B- power smooth.
- Recall, in Pollard $p-1$, this would be equivalent to not having $p-1 \not| m = \text{lcm}(1,2,\ldots,B)$; i.e. $a^m \not\equiv 1 \pmod{p}$.
- On the interval $[10^{15}, 10^{15}+10000]$ 15 percent of the primes p are such that $p-1$ is not 10^6-power smooth.

Motivation

- Fix $B \in \mathbb{N}$. Let $p \in \mathbb{N}$ such that $p - 1$ is not B- power smooth.
- Recall, in Pollard $p - 1$, this would be equivalent to not having $p - 1 \not| m = \text{lcm}(1, 2, \ldots, B)$; i.e. $a^m \not\equiv 1 \pmod{p}$.
- On the interval $[10^{15}, 10^{15} + 10000]$ 15 percent of the primes p are such that $p - 1$ is not 10^6-power smooth.
- The idea of ECM is to replace modular exponentiation on $(\mathbb{Z}/N\mathbb{Z})^*$ by repeated addition of points on $E((\mathbb{Z}/N\mathbb{Z})^*)$

Motivation

- Fix $B \in \mathbb{N}$. Let $p \in \mathbb{N}$ such that $p - 1$ is not B- power smooth.
- Recall, in Pollard $p - 1$, this would be equivalent to not having $p - 1 \not| m = \text{lcm}(1, 2, \ldots, B)$; i.e. $a^m \not\equiv 1 \pmod{p}$.
- On the interval $[10^{15}, 10^{15} + 10000]$ 15 percent of the primes p are such that $p - 1$ is not 10^6-power smooth.
- The idea of ECM is to replace modular exponentiation on $(\mathbb{Z}/N\mathbb{Z})^*$ by repeated addition of points on $E((\mathbb{Z}/N\mathbb{Z})^*)$
- Recall, by the Hasse-Weil bound we can reduce the size of our group by $2 \cdot \sqrt{p}$.

Algorithm 6.3.10 (Elliptic Curve Factorization Method). Let N and B be positive integers.

Algorithm 6.3.10 (Elliptic Curve Factorization Method). Let N and B be positive integers.

1. Compute $m = \operatorname{lcm}(1, 2, \ldots, B)$.

Elliptic Curve Factorization

Algorithm 6.3.10 (Elliptic Curve Factorization Method). Let N and B be positive integers.

1. Compute $m = \mathrm{lcm}(1, 2, \ldots, B)$.
2. Choose $a \in \mathbb{Z}/N\mathbb{Z}$ such that $4a^3 + 27 \in (\mathbb{Z}/N\mathbb{Z})^*$. This forces $P = (0, 1)$ to be a point on $y^2 = x^3 + ax + 1$ over $\mathbb{Z}/N\mathbb{Z}$.

Algorithm 6.3.10 (Elliptic Curve Factorization Method). Let N and B be positive integers.

1. Compute $m = \text{lcm}(1, 2, \ldots, B)$.
2. Choose $a \in \mathbb{Z}/N\mathbb{Z}$ such that $4a^3 + 27 \in (\mathbb{Z}/N\mathbb{Z})^*$. This forces $P = (0, 1)$ to be a point on $y^2 = x^3 + ax + 1$ over $\mathbb{Z}/N\mathbb{Z}$.
3. Try to compute mP. If at some point we cannot compute a sum of points, then some denominator g is not coprime to N, then $gcd(g, N)$ is a nontrivial divisor of N.

Analogy to Pollard $p - 1$

Analogy to Pollard $p - 1$

Table: Let E be an elliptic curve, and $m = lcm(1, 2, \ldots, B)$ for some B

Analogy to Pollard $p-1$

Table: Let E be an elliptic curve, and $m = lcm(1, 2, \ldots, B)$ for some B

Pollard $p-1$	**ECM**

Analogy to Pollard $p-1$

Table: Let E be an elliptic curve, and $m = lcm(1, 2, \ldots, B)$ for some B

Pollard $p-1$	**ECM**
$\mathbb{Z}/N\mathbb{Z}$	$E(\mathbb{Z}/N\mathbb{Z})$

Analogy to Pollard $p-1$

Table: Let E be an elliptic curve, and $m = lcm(1, 2, \ldots, B)$ for some B

Pollard $p-1$	**ECM**
$\mathbb{Z}/N\mathbb{Z}$	$E(\mathbb{Z}/N\mathbb{Z})$
$g \in (\mathbb{Z}/N\mathbb{Z})^*$	$(0, 1)$

Analogy to Pollard $p-1$

Table: Let E be an elliptic curve, and $m = lcm(1, 2, \ldots, B)$ for some B

Pollard $p-1$	**ECM**
$\mathbb{Z}/N\mathbb{Z}$	$E(\mathbb{Z}/N\mathbb{Z})$
$g \in (\mathbb{Z}/N\mathbb{Z})^*$	$(0, 1)$
$g^m \equiv 1 \pmod{N}$	$mP \notin E(\mathbb{Z}/N\mathbb{Z})$

Analogy to Pollard $p-1$

Table: Let E be an elliptic curve, and $m = lcm(1, 2, \ldots, B)$ for some B

Pollard $p-1$	**ECM**
$\mathbb{Z}/N\mathbb{Z}$	$E(\mathbb{Z}/N\mathbb{Z})$
$g \in (\mathbb{Z}/N\mathbb{Z})^*$	$(0, 1)$
$g^m \equiv 1 \pmod{N}$	$mP \notin E(\mathbb{Z}/N\mathbb{Z})$
$gcd(g^m - 1, N)$	$gcd(m, N)$

- If Pollard $p-1$ fails, we have no choice but to increase B.

Analogy to Pollard $p-1$

Table: Let E be an elliptic curve, and $m = lcm(1, 2, \ldots, B)$ for some B

Pollard $p-1$	**ECM**
$\mathbb{Z}/N\mathbb{Z}$	$E(\mathbb{Z}/N\mathbb{Z})$
$g \in (\mathbb{Z}/N\mathbb{Z})^*$	$(0, 1)$
$g^m \equiv 1 \pmod{N}$	$mP \notin E(\mathbb{Z}/N\mathbb{Z})$
$gcd(g^m - 1, N)$	$gcd(m, N)$

- If Pollard $p-1$ fails, we have no choice but to increase B.
- However, ECM has a second option. We can choose another random elliptic curve.

Why ECM "Works"

Why ECM "Works"

We can consider an analogous mapping

$$"g : E(\mathbb{Z}/N\mathbb{Z}) \to \prod_{p|N} E(\mathbb{Z}/p\mathbb{Z})"$$

where p are prime divisors of N.

Why ECM "Works"

We can consider an analogous mapping

$$"g : E(\mathbb{Z}/N\mathbb{Z}) \to \prod_{p|N} E(\mathbb{Z}/p\mathbb{Z})"$$

where p are prime divisors of N.

- Note the quotations. There is a subtly in the difference between $E(\mathbb{Z}/N\mathbb{Z})$ and $\mathbb{Z}/N\mathbb{Z}$.

Implementation

- Generate a random elliptic curve $E \pmod N$ and let $P = (0, 1)$.
- Compute $m = lcm(1, 2, ..., B)$.
- Compute mP (don't be naive!).
- If the calculation fails, you have found a non-trivial factor of N.
- Otherwise, just generate a new Elliptic curve and try again.

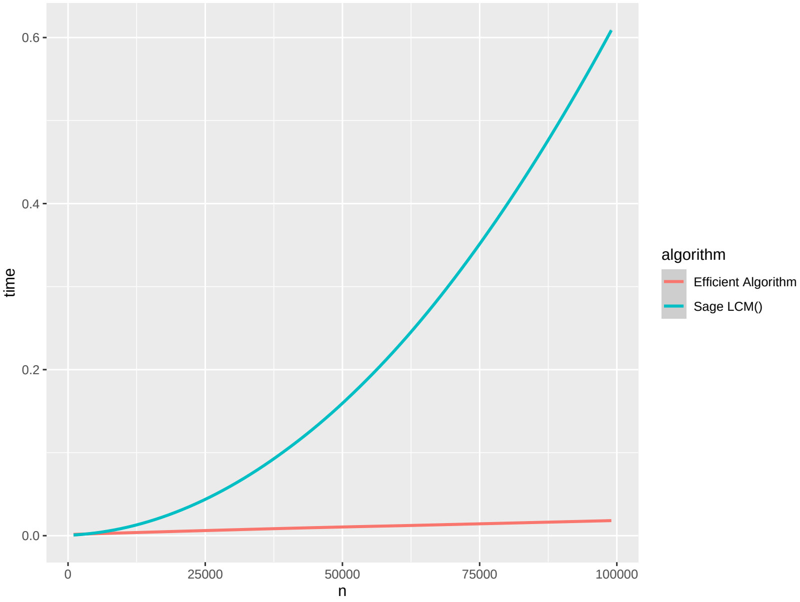
Computing lcm(1,2,...,B)

Recall,

$$lcm(1, 2, \dots B) = \prod_{p \in P} p^r$$

where $r = \max\{r \in \mathbb{Z} \mid p^r \leq B\}$.

$$\begin{aligned} p^r &\leq B \\ r\log(p) &\leq \log(B) \\ r &\leq \log_p(B) \\ r &= \lfloor \log_p(B) \rfloor \end{aligned}$$


0.6
0.4
0.2
0.0
time
0
25000
50000
75000
100000
n
algorithm
Efficient Algorithm
Sage LCM()

Computing mP

$$mP = \overbrace{P + P + P \dots P}^{m \text{ times}}$$

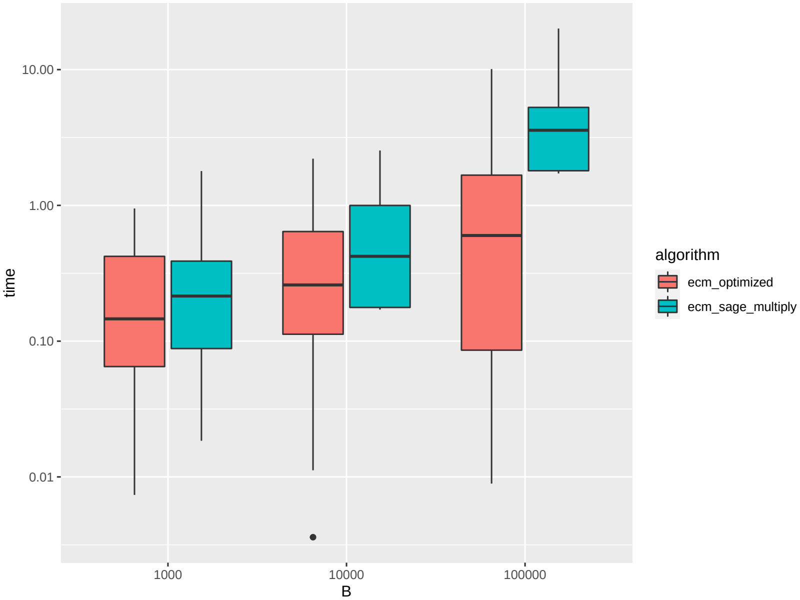
A very bad way to compute mP.

There are many algorithms for computing general elliptic curve point multiplication efficiently, and given the known make-up of m, we can save time by being thoughtful here.
Consider,

$$m_n = q_1^{r_1} \cdot q_2^{r_2} \dots q_n^{r_n}$$

then

$$m_n P = q_n^{r_n} \cdot m_{n-1} P$$


10.00
1.00
0.10
0.01
time
1000
10000
100000
B
algorithm
ecm_optimized
ecm_sage_multiply

Coded Example

```
def ecm(n, B=10^4, trials=100):
    R = Zmod(n)
    primes = list(prime_range(B+1))

    for _ in range(trials):
        while True:
            a = R.random_element()
            if gcd(4 * Integer(a)^3 + 27, n) == 1:
                break

        E = EllipticCurve([a, 1])
        P = E([0,1])

        try:
            for p in primes:
                P = P * p^floor(math.log(B,p))

        except ZeroDivisionError as e:
            return gcd(Integer(str(e).split()[2]), n)

    return -1
```

Complexity analysis

Complexity analysis

- Sieve all primes less than $B \implies O(B \log \log B)$

Complexity analysis

- Sieve all primes less than $B \implies O(B \log \log B)$
- Elliptic curve point multiplication is $O(k)$ where k is the number of bits (double-and-add).

Complexity analysis

- Sieve all primes less than $B \implies O(B \log \log B)$
- Elliptic curve point multiplication is $O(k)$ where k is the number of bits (double-and-add).
- $LCM(1, 2, 3, ..., B) \approx e^B$.

Complexity analysis

- Sieve all primes less than $B \implies O(B \log \log B)$
- Elliptic curve point multiplication is $O(k)$ where k is the number of bits (double-and-add).
- $LCM(1, 2, 3, ..., B) \approx e^B$.
- This implies computing mP given m is $O\left(\frac{B}{\log 2}\right)$.

Complexity analysis

- Sieve all primes less than $B \implies O(B \log \log B)$
- Elliptic curve point multiplication is $O(k)$ where k is the number of bits (double-and-add).
- $LCM(1, 2, 3, ..., B) \approx e^B$.
- This implies computing mP given m is $O\left(\frac{B}{\log 2}\right)$.
- Therefore, computing mP should take roughly $O(B \log \log B)$

Complexity analysis

Ok, but what are the odds we find a curve that works in the first place?

Complexity analysis

Ok, but what are the odds we find a curve that works in the first place?

- Canfield-Erdös-Pomerance theorem tells us about the probability that a random number $x < B$ is B smooth.

Complexity analysis

Ok, but what are the odds we find a curve that works in the first place?

- Canfield-Erdös-Pomerance theorem tells us about the probability that a random number $x < B$ is B smooth.
- Recall Hasse-Weil bound lets us limit the group by $2 \cdot \sqrt{p}$

Complexity analysis

Ok, but what are the odds we find a curve that works in the first place?

- Canfield-Erdös-Pomerance theorem tells us about the probability that a random number $x < B$ is B smooth.
- Recall Hasse-Weil bound lets us limit the group by $2 \cdot \sqrt{p}$
- $O\left(e^{\sqrt{2\log(p)\log(\log(p))}}\right)$