4.2 Lenstras elliptiske kurve algoritme

I [2] Lenstra præsenterede H. W. Lenstra en algoritme til faktorisering af heltal, som anvender elliptiske kurver. Vi skal benytte koncepterne vi udviklede til gruppeloven, men vi skal anvende dem på $E(\mathbb{Z}_n)$, hvor n er et sammensat tal. Når n er et sammensat tal er $\mathbb{Z}_n$ ikke er legeme og derfor er $E(\mathbb{Z}_n)$ heller ikke en gruppe. Derfor vælger vi at definere elliptiske pseudokurver for at kunne give mening til dette:

Definition 6 *Lad* $A, B \in \mathbb{Z}_n$*, hvor* $\gcd(n, 6) = 1$*. En elliptisk pseudokurve er da mængden*

$$E(\mathbb{Z}_n) = \{\infty\} \cup \{(x, y) \in \mathbb{Z}_n \times \mathbb{Z}_n \mid y^2 = x^3 + Ax + B\},$$

hvor $\gcd(4A^3 + 27B^2, n) = 1$.

Med denne definition følger det, at en elliptisk kurve specielt er en elliptisk pseudokurve, da hvis p er et primtal er $\mathbb{Z}_p = \mathbb{F}_p$ som er et endeligt legeme. Vi foretager addition af punkter på en elliptisk pseudokurve på samme måde, som for elliptiske kurver. På grund af denne definition vil vi for to punkter P og Q på en elliptisk pseudokurve kunne komme ud for at $P + Q$ ikke vil være defineret. Et tilfælde hvor $P+Q$ ikke er defineret vil blive fanget i udregningen af hældningen m i definition 2, da $\mathbb{Z}_n$ ikke er et legeme når n er et sammensat, så det der går galt er at $x_2 - x_1$ eller y_1 ikke har en invers i $\mathbb{Z}_n$. At en addition kan "gå galt"er motivationen for at kalde disse kurver for pseudokurver.

Vi ser nu på, hvordan vi kan finde en faktor, når additionen af to punkter ikke er defineret. Algoritmen vi her præsenterer er inspireret af algoritmerne i [3] og [4]. Lad n være et sammensat tal. Vi husker fra gruppeloven, at vi ved additionen af to punkter skal bruge værdien af inverserne til $x_2 - x_1$ og y_1 alt efter, hvilket tilfælde vi er i. Disse inverser eksisterer kun modulo n, hvis (se bevis i appendiks)

$$\gcd(x_1 - x_2, n) = 1 \quad \text{og} \quad \gcd(y_1, n) = 1.$$

Men hvis vi er i stand til at finde punkter $P = (x_1, y_1)$ og $Q = (x_2, y_2)$ sådan, at summen $P + Q$ ikke er defineret, da er $\gcd(x, n) > 1$ hvor $x = x_1 - x_2$ eller $x = y_1$ og dermed har vi muligvis fundet en ikke-triviel faktor i n. Vi kan da se på, hvordan vi kan udnytte dette faktum til at lave en algoritme, som giver os en faktor i n. Vælg først tilfældige heltal x, y, A mellem 1 og n, og sæt $B = y^2 - x^3 - Ax \pmod{n}$. Vi har da en elliptisk kurve (egentlig ikke en elliptisk kurve, da n er et sammensat tal, men vi lader som om)

$$E : y^2 = x^3 + Ax + B,$$

hvor vi ved at punktet $P = (x, y)$ er placeret. Vi tjekker da, at

$$d = \gcd(4A^3 + 27B^2, n) = 1.$$

Hvis det ikke er tilfældet og $1 < d < n$ har vi fundet en faktor i n. Hvis $d = n$ vælger vi et nyt A. For et heltal K lader vi $k = \text{lcm}[1, 2, \ldots, K]$ og vi forsøger da, at bestemme

$$kP = \underbrace{P + P + \ldots + P}_{k \text{ led}}.$$

Det er ikke praktisk at udregne $P + P + \ldots + P$ så vi gør ligesom i Pollards $p - 1$ algoritme og skriver k som den binære udvidelse

$$k = k_0 + k_1 \cdot 2 + k_2 \cdot 2^2 + k_3 \cdot 2^3 + \ldots + k_r \cdot 2^r,$$

hvor alle k_i er 0 eller 1. Vi kan da udregne

$$\begin{aligned} P_0 &= P \\ P_1 &= 2P_0 = 2P \\ P_2 &= 2P_1 = 2^2 P \\ &\vdots \\ P_r &= 2P_{r-1} = 2^r P. \end{aligned}$$

Da kan vi bestemme $kP =$ (summen af P_i'erne hvor $k_i = 1$). I hver udregning regner vi modulo n, da tallene ellers bliver alt for store og meget langsommelige at arbejde med. Vi håber på, at der i løbet af de udregninger er en addition, som ikke kan lade sig gøre og at dette giver os vores faktor.

Med algoritmen på plads kan vi nu se på et eksempler:

Vi opsummerer diskussionen i algoritmen nedenfor:

Algoritme 2 (Lenstras algoritme) *Lad $n \geq 2$ være et sammensat tal.*

1. *Vælg heltal x, y og A mellem 1 og n. Lad da $B = y^2 - x^3 - Ax \pmod{n}$, så vi har den elliptiske kurve*

$$E : y^2 = x^3 + Ax + B,$$

hvor punktet $P = (x, y)$ er placeret.

2. *Tjek at $d = \gcd(4A^3 + 27B^2, n) = 1$. Hvis $d = n$ går vi tilbage til (1) og vælger et nyt B. Hvis $1 < d < n$ har vi fundet en faktor af n og vi er færdige.*

3. *Vælg et positivt heltal k som et produkt af mange små primtal, lad eksempelvis*

$$k = \text{lcm}[1, 2, 3, ..., K],$$

hvor $K \in \mathbb{Z}^+$.

4. *Forsøg at bestemme $kP = P + P + \ldots + P$. Hvis udregningen kan lade sig gøre går vi tilbage til (1) og vælger en ny kurve, eller går til (3) og vælger et større k.*

For at se hvorfor der er en god chance for, at vi støder på et valg af x, y, A sådan at vi finder en faktor, lader vi p være en primfaktor i n. Til den elliptiske kurve E har vi den abelske gruppe $E(\mathbb{F}_p)$ og pr. sætning 3 ved vi, at

$$p + 1 - 2\sqrt{p} < \#E(\mathbb{F}_p) < p + 1 + 2\sqrt{p}.$$

En sætning af Deuring [1] siger, at for ethvert heltal $m \in (p + 1 - 2\sqrt{p}, p + 1 + 2\sqrt{p})$ findes der talpar (A, B) i mængden

$$\{(A, B) \mid A, B \in \mathbb{F}_p, 4A^3 + 27B^2 \neq 0\},$$

sådan at $\#E(\mathbb{F}_p) = m$. For hvert tal i intervallet findes der altså en elliptisk kurve, som har denne orden. Der er altså en positiv sandsynlighed for, at vi kan finde sådan en kurve.

Lenstras sandsynlighedsteoretiske overvejelser.

Eksempel 5 Lad nu

$$n = 753161713$$

være det tal, som vi ønsker at faktorisere. Da $2^{n-1} = 437782651 \pmod{n}$ er n ikke et primtal. Vi vælger da $x = 0$, $y = 1$ og $A = 164$. Vi har dermed, at $B = 1^2 - 0^3 - 164 \cdot 0 = 1$ og den elliptiske kurve vi vil arbejde over bliver

$$E : y^2 = x^3 + 164x + 1,$$

hvorpå punktet $P = (0, 1)$ er placeret. Vi ser, at

$$\begin{aligned} D &= \gcd(4 \cdot 164^3 + 27 \pmod{753161713}, 753161713) \\ &= \gcd(17643803, 753161713) = 1, \end{aligned}$$

så vi fortsætter derfor med algoritmen. Vi lader

$$k = \operatorname{lcm}[1, 2, \ldots, 10] = 2520.$$

Da $2520 = 2^{11} + 2^8 + 2^7 + 2^6 + 2^4 + 2^3$ skal vi beregne $2^i P \pmod{753161713}$ for $0 \leq i \leq 11$. Dette gøres med additionsformlen og vi opsummerer vores resultater i tabellen nedenfor:

i	$2^i P \pmod{753161713}$	i	$2^i P \pmod{753161713}$
0	$(0, 1)$	6	$(743238772, 703386057)$
1	$(6724, 752610344)$	7	$(309161840, 219780637)$
2	$(293427237, 450490340)$	8	$(116974611, 722899047)$
3	$(468952095, 385687511)$	9	$(329743899, 182819134)$
4	$(288125200, 446796094)$	10	$(163952469, 456288424)$
5	$(106753239, 115973502)$	11	$(15710788, 301760412)$

Vi kan nu addere disse punkter igen vha. additionsformlerne, hvor vi stadigvæk regner modulo n:

$$(2^3 + 2^4)P = (606730980, 447512524).$$

Algoritmen giver os en faktor netop når additionen bryder sammen, hvilket kan ske da $\mathbb{Z}/n\mathbb{Z}$ ikke er et legeme. Dette problem viser sig i dette eksempel allerede ved den næste addition, hvor vi forsøger at udregne

$$\begin{aligned}(2^3 + 2^4 + 2^6)P = &(743238772, 703386057)\\ &+(606730980, 447512524) \pmod{n}.\end{aligned}$$

For at denne addition skal kunne lade sig gøre, skal differensen af deres x-koordinater have en invers modulo n. Dette er kun tilfældet, hvis $\gcd(x_2 - x_1, n) = 1$ (se appendiks, sætning k). Men vi ser, at

$$\gcd(606730980 - 743238772, 753161713) = 19259,$$

så der findes altså ikke en invers, men vi har i stedet fundet en faktor i n. Dermed har vi faktoriseringen

$$753161713 = 19259 \cdot 39107.$$

Nu kan det virke til, at det var spild da vi lavede hele tabellen, men i beregningerne af $2^i P \pmod{753161713}$ ville vi også kunne have løbet ind i et element, som ikke havde en invers og som dermed kunne give os en faktor.