Universiteit Utrecht

Faculteit Bètawetenschappen

Factorisatie met de algoritmes van Pollard en Lenstra

Bachelorscriptie

Timotheus Weijers

Wiskunde

Scriptiebegeleider:

Dr. P.H. Koymans Universiteit Utrecht

januari 2025

Samenvatting

In deze scriptie analyseren we het $p-1$ factorisatiealgoritme van Pollard en het elliptische kromme factorisatiealgoritme (ECM) van Lenstra. De focus ligt op de getaltheoretische fundamenten van Pollards algoritme, waarbij we de complexiteitsanalyse als secundair beschouwen.

We onderzoeken in detail de keuze van de exponent L in Pollards algoritme. In de literatuur worden hiervoor doorgaans twee opties gebruikt: de faculteit $L!$ of de functie $M(L)$, gedefiniëerd als het product van priemmachten kleiner gelijk L. Door middel van analytische getaltheorie bestuderen we de eigenschappen van $L!$, terwijl we $M(L)$ analyseren met behulp van de tweede functie van Chebyshev.

Als alternatieve exponent in Pollards algoritme introduceren we een exponent gebaseerd op het product van priemgetallen kleiner dan L. We schatten vanaf de basis de slagingskans van Pollards methode in met deze nieuwe exponent, waarbij we enkele noodzakelijke resultaten over de Möbiusfunctie μ en zijn relatie met de Riemann-zètafunctie behandelen.

Voor Lenstra's ECM algoritme presenteren we eerst een theoretisch fundament van algebraïsche vlakkrommen, gevolgd door een grondige behandeling van elliptische krommen over Q en over eindige lichamen en hun toepassing in het factorisatiealgoritme.

Inhoudsopgave

1 Inleiding

Een priemgetal p is een natuurlijk getal met precies twee delers, 1 en p. Priemgetallen vormen de fundamentele bouwstenen van de natuurlijke getallen. Al in de Elementen bewees Euclides dat elk natuurlijk getal uniek te schrijven is als product van priemgetallen. Twee centrale vragen in de getaltheorie zijn:

1. **Primaliteit:** Gegeven een natuurlijk getal n, is het priem of samengesteld?
2. **Factorisatie:** Gegeven een samengesteld getal, wat zijn zijn priemfactoren?

De eerste vraag is praktisch opgelost: moderne algoritmen kunnen getallen met tienduizenden cijfers testen. In 2002 vonden Agrawal, Kayal en Saxena het AKS-algoritme om in polynomiale tijd primaliteit te testen, hoewel dit in de praktijk niet het snelste algoritme is. In paragraaf 5 geven we een kort overzicht van priemtesten en leggen we de Miller-Rabin-priemtest uit, een eenvoudig te begrijpen maar krachtige primaliteitstest die geschikt is voor cryptografie.

1.1 Geschiedenis van factorisatie

De tweede vraag is lastiger. Tot de 19e eeuw werden priemfactoren voornamelijk gevonden door met de hand delers te zoeken. Deze methode heet proefdeling en wordt beschreven in paragraaf 6. Hoewel Euler al wel manieren kende om hierbij minder berekeneningen te hoeven doen, was deze methode tijdrovend. Ook op moderne computers kunnen we alleen getallen van minder dan ≈ 15 tot 20 cijfers factoriseren met proefdeling.

De geschiedenis van factorisatiealgoritmen kent drie belangrijke perioden:

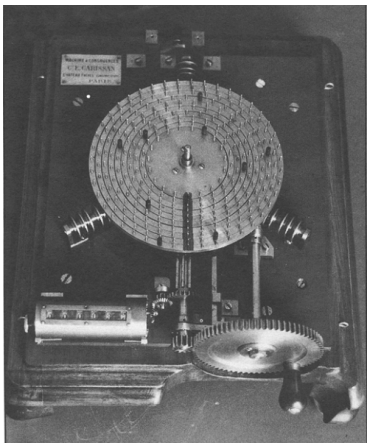
- **Klassieke periode** (tot 19e eeuw): Handmatige methoden zoals proefdeling en Fermats methode
- **Mechanische periode** (begin 20e eeuw): Factoriseermachines van Carissan en Lehmer
- **Moderne periode** (vanaf 1970): Computationele algoritmen zoals:
  - Lehmans $O(n^{1/3})$-algoritme (1974)
  - Pollards rho- en $p-1$-methoden (1974)
  - De kwadratische zeef (Pomerance, 1984)
  - Lenstra's elliptische-krommemethode (1987)
  - De getallenlichaamzeef (Pollard, 1988)

In de 17e eeuw ontwikkelde Fermat een factoriseermethode die berust op het schrijven van een samengesteld getal als het verschil van twee kwadraten. Deze methode blijkt echter in veel gevallen minder efficiënt te zijn dan de traditionele proefdeling.

Aan het begin van de 20e eeuw bouwden Carissan (Figuur 1) en Lehmer (Figuur 2) enkele mechanische factoriseermachines.

Deze machines bleven tot de jaren 70 de snelste manier om te factoriseren.

In de jaren 70 werden vele snelle factoriseermethoden ontdekt. In 1974 ontdekte Lehman een $O(n^{\frac{1}{3}})$-algoritme [R S74]. In datzelfde jaar vond Pollard de Pollard-rho-methode en het Pollard-$p-1$-algoritme, die significant sneller zijn. In 1987 ontwikkelde Lenstra het elliptische-krommefactorisatiealgoritme, dat voortbouwde op Pollards $p-1$-algoritme maar significant beter werkt. Het is nog steeds het beste algoritme voor wanneer n een kleine deler (tot 50–60 cijfers) heeft.



Figuur 1: De factoriseermachine van Carissan uit 1920, foto uit [SWM95]

Figuur 2: Factoriseerstencils gebaseerd op kwadratische residuën, gemaakt door Derrick Norman Lehmer in 1929, foto van [Smi29]. Later bouwde Lehmer een elektronische versie met fotocellen.

In de jaren 80 werden ook andere methoden gevonden. De methode van Fermat vormde de basis voor de in 1984 door Carl Pomerance ontwikkelde kwadratische zeef en de in 1988 door John Pollard ontwikkelde getallenlichaamszeef, die momenteel het snelste algoritme is voor getallen met enkel grote priemfactoren.

De publicatie van het RSA-cryptografiesysteem in 1977 stimuleerde de zoektocht naar snelle factorisatiealgoritmen. Het systeem is gebaseerd op de moeilijkheid van factorisatie. Indien iemand een snel factorisatiealgoritme vindt, wordt RSA gebroken.

In deze scriptie behandelen we het algoritme van Pollard in paragraaf 7 en de methode van Lenstra in paragraaf 12.

1.2 Het idee achter Pollards algoritme

Zij n het te factoriseren getal en p een priemfactor van n. Als $p-1$ alleen kleine priemfactoren heeft is, dan vindt het algoritme mogelijk p via:

$$p = \text{ggd}(2^{L(B)} - 1, n)$$

waarbij L één van de volgende functies is:

- $L(B) = B!$ (faculteit)
- $L(B) = M(B) = \prod_{p \text{ priem}} p^{\lfloor \log_p(B) \rfloor}$.
- $L(B) = \prod_{p \text{ priem}, p \leq B} p$

In paragraaf 7 bekijken we kort de eigenschappen van de priemfactorisatie van de faculteit.

In paragraaf 8 bestuderen we de eigenschappen van de functie $M(B)$ die, verrassend genoeg, in de limiet naar e^B gaat.

In hoofdstuk paragraaf 9 beschouwen we de derde mogelijkheid voor L: het product van alle priemgetallen kleiner dan B. Het blijkt dat Pollards algoritme voor deze keuze van L slaagt als de priemfactorisatie van $p-1$ geen kwadraten (behalve 1^2) als deler heeft. Met stellingen uit de analytische getaltheorie geven we een heuristische schatting van de waarschijnlijkheid dat dit het geval is.

De methode van Lenstra vereist voorkennis van elliptische krommen. Deze behandelen we in paragraaf 10 en paragraaf 11.

Een uitgebreide behandeling van de complexiteitsgraad van de algoritmen valt helaas buiten het bereik van deze scriptie. Er is wel een heuristische complexiteit van de methoden van Pollard en Lenstra bekend, maar deze afleiden zou een gehele scriptie op zich zijn en vereist complexe voorkennis. (De theoretische complexiteit van de vermenigvuldiging van zeer grote getallen is pas sinds 2007 bewezen en biedt voldoende materiaal voor een tweede scriptie.)

Het is verrassend hoeveel deelgebieden van de wiskunde samenkomen bij het bestuderen van factorisatiemethoden: niet alleen algoritmiek, maar ook analytische getaltheorie, combinatoriek, abstracte algebra en algebraïsche meetkunde.

2 Eindige groepen en ringen

De theorie van groepen en ringen is belangrijk voor de algoritmes van Pollard en Lenstra. Pollards $p-1$ algoritme werkt over de multiplicatieve groep $(\mathbb{Z}/p\mathbb{Z})^\times$ voor een priemfactor p van n. Om te begrijpen hoe het algoritme werkt is de chinese reststelling nuttig.

Voor Lenstra's algoritme gebruiken we dat een elliptische kromme modulo een samengesteld getal n zich gedraagt als het product van elliptische krommen modulo de priemfactoren van n.

2.1 Groepentheorie

Definitie 2.1 (Groep van eenheden)**.** Zij $n \in \mathbb{N}_{\geq 1}$. We noemen een element $a \in \mathbb{Z}/n\mathbb{Z}$ een *eenheid* als er een $b \in \mathbb{Z}/n\mathbb{Z}$ bestaat zodat $ab \equiv 1 \pmod{n}$. De verzameling van alle eenheden in $\mathbb{Z}/n\mathbb{Z}$ vormt een groep onder vermenigvuldiging modulo n, genoteerd als $(\mathbb{Z}/n\mathbb{Z})^\times$. Een element $a \in \mathbb{Z}/n\mathbb{Z}$ is een eenheid dan en slechts dan als $\text{ggd}(a,n) = 1$. De eenheid e van deze groep is het getal 1.

Stelling 2.2 (Orde van de groep van eenheden)**.** *Voor $n \in \mathbb{N}_{\geq 1}$ wordt de orde van de groep van eenheden gegeven door de Euler ϕ-functie: $|(\mathbb{Z}/n\mathbb{Z})^\times| = \phi(n)$, waarin $\phi(n)$ het aantal getallen k tussen 1 en n is waarvoor* $\text{ggd}(k,n) = 1$.

Definitie 2.3 (Orde van een element)**.** Zij G een groep en $g \in G$. De *orde* van g, genoteerd als $\text{ord}(g)$, is het kleinste positieve gehele getal n waarvoor geldt dat $g^n = e$. Indien zo'n n niet bestaat zeggen we $\text{ord}(g) = \infty$. In een eindige groep heeft elk element een eindige orde.

Definitie 2.4 (Ondergroep van een eindige groep)**.** Een niet-lege deelverzameling H van een eindige groep $(G, \cdot)$ heet een *ondergroep* als $H \neq \emptyset$ en voor alle $a, b \in H$ geldt: $a \cdot b \in H$ (geslotenheid). Notatie: $H \leq G$

Stelling 2.5 (Stelling van Lagrange)**.** *Zij G een eindige groep en $H \leq G$ een ondergroep. Dan geldt dat de orde van H de orde van G deelt: $|H| \mid |G|$.*

Gevolg 2.6. *Voor elk element $g \in G$ geldt* $\text{ord}(g) \mid |G|$.

2.2 Ringtheorie

Definitie 2.7 (Lichaam)**.** Een lichaam is een verzameling $\mathbb{F}$ samen met twee bewerkingen $+$ en $\cdot$, geschreven als $(\mathbb{F}, +, \cdot)$, zodanig dat:

1. $(\mathbb{F}, +)$ is een abelse groep met neutraal element 0:
   - $\forall a, b \in \mathbb{F} : a + b = b + a$ (commutativiteit)
   - $\forall a, b, c \in \mathbb{F} : (a + b) + c = a + (b + c)$ (associativiteit)
   - $\forall a \in \mathbb{F} : a + 0 = a$ (neutraal element)
   - $\forall a \in \mathbb{F}, \exists (-a) \in \mathbb{F} : a + (-a) = 0$ (inverse elementen)
2. $(\mathbb{F} \setminus \{0\}, \cdot)$ is een abelse groep met neutraal element 1:
   - $\forall a, b \in \mathbb{F} \setminus \{0\} : a \cdot b = b \cdot a$ (commutativiteit)
   - $\forall a, b, c \in \mathbb{F} \setminus \{0\} : (a \cdot b) \cdot c = a \cdot (b \cdot c)$ (associativiteit)
   - $\forall a \in \mathbb{F} \setminus \{0\} : a \cdot 1 = a$ (neutraal element)
   - $\forall a \in \mathbb{F} \setminus \{0\}, \exists a^{-1} \in \mathbb{F} \setminus \{0\} : a \cdot a^{-1} = 1$ (inverse elementen)

3. De vermenigvuldiging distribueert over de optelling:

$$\forall a, b, c \in \mathbb{F} : a \cdot (b + c) = (a \cdot b) + (a \cdot c)$$

Het aantal elementen van een lichaam heet de *orde* en wordt genoteerd als ord($\mathbb{F}$). Een eindig lichaam is een lichaam waarvoor ord($\mathbb{F}$) eindig is.

De karakteristiek van een lichaam $\mathbb{F}$, genoteerd char($\mathbb{F}$) is het minimale aantal keer dat men 1 moet optellen om 0 te krijgen. Indien zo'n n niet bestaat zeggen we char($\mathbb{F}$) $= 0$. Voor een eindig lichaam is char($\mathbb{F}$) een priemgetal.

Stelling 2.8 (Eigenschappen van eindige lichamen)**.** *Voor elk eindig lichaam $\mathbb{F}$ geldt:*

1. *De orde van $\mathbb{F}$ is altijd een macht van een priemgetal: $|\mathbb{F}| = p^n$ waarin p priem is en $n \geq 1$*
2. *Voor elke priemmacht $q = p^n$ bestaat er een uniek (op isomorfisme na) eindig lichaam van orde q, genoteerd als $\mathbb{F}_q$ of* GF(q)
3. *Een lichaam van priem orde p is isomorf aan het lichaam $\mathbb{Z}/p\mathbb{Z}$. De elementen zijn de restklassen $0, \ldots, p-1$. De operaties $+$, $-$, $\cdot$ worden op de representanten uitgevoerd, en daarna wordt de rest bij deling door p genomen. Deling door a gebeurt door vermenigvuldiging met de inverse a^{-1}, welke met het uitgebreide algoritme van Euclides toegepast op (a, p) wordt gevonden.*

De structuur van een lichaam $\mathbb{F}$ met $|\mathbb{F}| = p$ is volgens het bovenstaande lemma eenvoudig. Voor $|\mathbb{F}| = p^n, n \geq 2$ is de structuur lastiger. Het lichaam met 4 elementen is niet isomorf aan de ring $\mathbb{Z}/4\mathbb{Z}$, aangezien daarin het element 2 geen multiplicatie inverse heeft, omdat $2 \cdot 2 \equiv 0 \pmod 4$, dus 2 is een nuldeler.

De constructie van $\mathbb{F}_4$: Om een lichaam $\mathbb{F}_{p^n}$ te construeren kiezen we een irreducibel polynoom r in $\mathbb{F}_p[X]$ van graad n. Dan

$$\mathbb{F}_{p^n} = \mathbb{F}_p[X]/(r) \quad ((r) \text{ is het ideaal gegenereerd door } r)$$

Zij nu α een wortel van r, dan

$$\mathbb{F}_p[X]/(r) \cong \mathbb{F}_p(\alpha) \qquad \text{waarin } \mathbb{F}_p(\alpha) \text{ een lichaamsuitbreiding van } \mathbb{F}_p \text{ is.}$$

Nu $\mathbb{F}_4$: er is maar één irreducibel polynoom van graad 2, namelijk $X^2 + X + 1$. Indien α een wortel is, dan $\alpha^2 + \alpha + 1 = 0$, dus $\alpha^2 = -\alpha - 1 = \alpha + 1$.

$$\mathbb{F}_2[X]/(X^2 + X + 1) \cong \mathbb{F}_2(\alpha)$$

Hiermee vinden we de vier elementen. $\mathbb{F}_4 = \{0, 1, \alpha, 1 + \alpha\}$. Met de regel $\alpha^2 = \alpha + 1$, kunnen nu eenvoudig tabellen voor optelling en vermenigvuldiging gemaakt worden.

Hoewel het factorisatiealgoritme van Lenstra gebruikt maakt van eindige lichamen $\mathbb{F}_q$ is het voldoende om deze te begrijpen voor het geval dat q priem is.

Definitie 2.9 (Ringisomorfisme)**.** Een ringhomomorfisme $\phi : R \to S$ is een functie die de ringstructuur behoudt, dwz voor alle $a, b \in R$:

$$\begin{aligned} \phi(1) &= 1 \\ \phi(a + b) &= \phi(a) + \phi(b) \\ \phi(a \cdot b) &= \phi(a) \cdot \phi(b) \end{aligned}$$

Indien ϕ bijectief is heet ϕ een ringisomorfisme en schrijven we $R \cong S$ en zeggen we dat R en S isomorf zijn.

Stelling 2.10 (Chinese reststelling [DF04, § 7.6])**.** *Zij n een positief getal met factorisatie $p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$ in priemmachten. Dan bestaat het volgende isormorphisme tussen ringen*

$$\mathbb{Z}/n\mathbb{Z} \cong (\mathbb{Z}/{p_1}^{\alpha_1}\mathbb{Z}) \times (\mathbb{Z}/{p_2}^{\alpha_2}\mathbb{Z}) \times \cdots \times (\mathbb{Z}/{p_k}^{\alpha_k}\mathbb{Z})$$

Gevolg 2.11. *In het bijzonder bestaat het volgende groepsisomorfisme tussen de multiplicatieve groepen van eenheden*

$$(\mathbb{Z}/n\mathbb{Z})^\times \cong (\mathbb{Z}/{p_1}^{\alpha_1}\mathbb{Z})^\times \times (\mathbb{Z}/{p_2}^{\alpha_2}\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/{p_k}^{\alpha_k}\mathbb{Z})^\times$$

Dit wordt gebruikt in Pollards $p-1$ algoritme.

We kunnen de chinese reststelling ook op de volgende manier opschrijven (equivalent). Zij $m_1, m_2, \dots, m_k$ paarsgewijs copriem. Dan heeft het stelsel congruenties:

$$\begin{aligned} x &\equiv a_1 \pmod{m_1} \\ x &\equiv a_2 \pmod{m_2} \\ &\vdots \\ x &\equiv a_k \pmod{m_k} \end{aligned}$$

een unieke oplossing modulo $m = \prod_{i=1}^k m_i$. Deze wordt gegeven door

$$x \equiv \sum_{i=1}^{k} a_i M_i y_i \qquad \text{waarin } M_i = \frac{m}{m_i} \qquad \text{en } y_i \text{ is de multiplicatieve inverse van } M_i \bmod m_i$$

Voorbeeld: We lossen het volgende stelsel congruenties op:

$$\begin{aligned} x &\equiv 2 \pmod 3 \\ x &\equiv 3 \pmod 5 \\ x &\equiv 2 \pmod 7 \end{aligned}$$

Hier geldt $m = 3 \cdot 5 \cdot 7 = 105$ en

$$\begin{aligned} M_1 &= 5 \cdot 7 = 35 & y_1 &\equiv 2 \pmod 3 \\ M_2 &= 3 \cdot 7 = 21 & y_2 &\equiv 1 \pmod 5 \\ M_3 &= 3 \cdot 5 = 15 & y_3 &\equiv 1 \pmod 7 \end{aligned}$$

Dus:

$$x \equiv 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 \equiv 23 \pmod{105}$$

De chinese reststelling is voldoende voor Pollards $p-1$ algoritme. Voor Lenstra's algoritme hebben we echter een meer algemene stelling nodig:

Stelling 2.12 (Hoofdstelling van eindige abelse groepen [DF04, § 5.2])**.** *Zij G een eindige abelse groep, dan*

$$G \cong \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \times \mathbb{Z}_{n_s}$$

boor r, n_1, n_2, …, n_s die voldoen aan

(a) $n_j \geq 2$ voor alle j

(b) $n_{i+1} \mid n_i$ *voor* $1 \leq i \leq s-1$.

De uitdrukking Stelling 2.12 is uniek: als $G \cong \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \times \cdots \times \mathbb{Z}_{m_u}$ *en* m_j *voldoen aan (a), (b) dan* $u = s$ *en* $\forall\ i :\ m_i = n_i$.

Zie [DF04, § 6.1] voor een bewijs.

Lemma 2.13. *De multiplicatieve groep* $(\mathbb{Z}/p\mathbb{Z})^\times$ *is een cyclische groep van orde* $p-1$.

3 Getaltheorie

We behandelen in dit hoofdstuk de noodzakelijke voorkennis van getaltheorie waaronder enkele definities van gladheid (alleen kleine priemfactoren hebben), en enkele stellingen over de verdeling van priemgetallen.

3.1 Deelbaarheid

Definitie 3.1 (Deelbaarheid)**.** Zij $a, b \in \mathbb{Z}$. We zeggen dat a deelbaar is door b, geschreven als $b \mid a$, als er een $k \in \mathbb{Z}$ bestaat zodat $a = bk$, b heet een *deler* of *factor* van a.

Definitie 3.2 (Grootste gemene deler)**.** Zij $a, b \in \mathbb{Z}$. De *grootste gemene deler* van a en b, is het grootste getal $d \in \mathbb{Z}$ waarvoor $d \mid a$ en $d \mid b$. Het wordt genoteerd als $\gcd(a, b)$ of soms als (a, b).

Lemma 3.3 (Elke deler deelt grootste gemene deler)**.** *Als $c|a$ en $c|b$ dan $c|\gcd(a, b)$.*

Lemma 3.4 (Bézouts identiteit)**.** *Zij a, b getallen met $\gcd(a, b) = d$. Dan bestaan er getallen x, y zodanig dat $ax + by = d$. Bovendien zijn de getallen d van de vorm $ax + by = d$ precies de veelvouden van $\gcd(a, b)$.*

Definitie 3.5 (Copriem)**.** Twee gehele getallen a en b zijn *copriem* als $\gcd(a, b) = 1$.

Stelling 3.6 (Delingsalgorithme)**.** *Voor elk geheel getal a en elk positief geheel getal n bestaan er unieke gehele getallen q en r zodanig dat $0 \leq r < n$ en $a = qn + r$.*

Definitie 3.7 (Quotiënt en rest)**.** Bij de deling van a door n noemen we $q = \lfloor a/n \rfloor$ het *quotiënt* van de deling. De waarde $r = a \bmod n$ is de *rest* (of *residu*) van de deling, zodat $n \mid a$ dan en slechts dan als $a \bmod n = 0$.

Stelling 3.8 (Uitgebreid algoritme van Euclides)**.** *Het uitgebreide algoritme van Euclides toegepast op $a, b \in \mathbb{N}_{\geq 1}$ eindigt met $d = \gcd(a, b)$ en de coëfficienten $u, v \in \mathbb{Z}$ waarvoor $d = ua + vb$.*

Definitie 3.9 (Congruentie en restklassen)**.** De gehele getallen worden verdeeld in n restklassen modulo n. De *restklasse modulo n* die een geheel getal a bevat is:

$$\overline{a}_n = \{a + kn : k \in \mathbb{Z}\}$$

Bijvoorbeeld, $\overline{3}_7 = \{\ldots, -11, -4, 3, 10, 17, \ldots\}$, en $\overline{-4}_7$ en $\overline{10}_7$ geven dezelfde verzameling aan. Met de notatie $a \equiv b \pmod{n}$ bedoelen we dat $a \in \overline{b}_n$. De verzameling van alle restklassen is:

$$\mathbb{Z}/n\mathbb{Z} = \{\overline{a}_n : 0 \leq a \leq n - 1\} = \{0, 1, \ldots, n - 1\}.$$

3.2 Priemgetallen

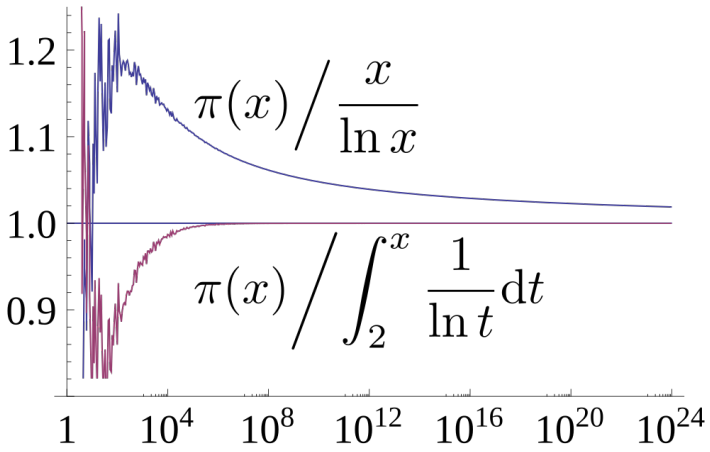
Definitie 3.10 (Priemgetal)**.** Zij a een positief geheel getal. Een *echte deler* van a is een positieve deler van a die niet gelijk is aan 1 of a. Een geheel getal $p \geq 2$ heet een *priemgetal* als het geen echte delers heeft. Een geheel getal dat wel echte delers heeft heet *samengesteld*

Definitie 3.11 (Priemfactor)**.** Een getal $a \in \mathbb{N}$ heet priemfactor van $b \in \mathbb{N}$ als het b deelt en priem is

Stelling 3.12 (Hoofdstelling van de rekenkunde)**.** *Elk natuurlijk getal $n > 1$ kan op precies één manier worden geschreven als een product van priemgetallen (afgezien van de volgorde van de factoren). Dit heet de priemfactorisatie van een getal.*

$$\forall n \in \mathbb{N}_{\geq 1} \; \exists \text{ unieke } p_1^{a_1} p_2^{a_2} \ldots p_k^{a_k} = n$$

Waarbij $p_1, p_2, \ldots, p_k$ priemgetallen zijn met $p_1 < p_2 < \ldots < p_k$ en $a_i \in \mathbb{N}_{\geq 0}$.



Figuur 3: De twee schattingen $\frac{x}{\ln(x)}$ en $\mathrm{Li}(x)$ convergeren naar $\pi(x)$. Merk op dat $\mathrm{Li}(x)$ veel sneller convergeert. [Dco13]

Lemma 3.13. *Indien p oneven priem en $x \in \mathbb{Z}_p$, dan*

$$x^2 \equiv 1 \pmod{p} \iff x \equiv \pm 1 \pmod{p}$$

Stelling 3.14 (De priemgetalstelling van Hadamard en De La Vallée-Poussin)**.** *Zij $\pi(x)$ de functie de priemgetal-telfunctie, die het aantal priemgetallen $\leq x$ telt. Dan*

$$\lim_{x\to\infty} \frac{\pi(x)\ln(x)}{x} = 1$$

3.2.1 De logaritmische integraal

De (verschoven) logaritmische integraal voor $x \geq 2$ is gedefiniëerd als

$$\mathrm{Li}(x) = \int_2^x \frac{1}{\ln(t)}\,\mathrm{d}t$$

3.2.2 Intuïtie achter de benadering

De functie $\mathrm{Li}(x)$ geeft een veel betere benadering voor $\pi(x)$ dan $\frac{x}{\ln(x)}$. De intuïtie erachter is als volgt.

Stel dat priemgetallen echt willekeurig verdeeld zijn, met een kans van $\frac{1}{\ln(n)}$ dat een positief geheel getal n priem is (voor $n \geq 2$). Dan zou het verwachte aantal priemgetallen tussen 2 en n gelijk zijn aan:

$$\sum_{k=2}^{n} \frac{1}{\ln(k)} \approx \int_2^n \frac{1}{\ln(x)}\,\mathrm{d}x = \mathrm{Li}(n)$$

waarbij de LHS een soort Riemann-som is die de RHS benadert. De benadering $\frac{n}{\ln(n)}$ is te klein omdat deze voor elk getal x in $\{2, \ldots n\}$ uitgaat van een kans van $\frac{1}{\ln(n)}$ dat x priem is, terwijl de dichtheid van de priemgetallen juist veel groter is voor kleine x.

3.2.3 Priemgetalverschillen (prime gaps)

Zij $g_n = p_{n+1} - p_n$ het verschil tussen het $(n+1)$-de en n-de priemgetal. De eerste waarden zijn $g_1 = 3 - 2 = 1$, $g_2 = 5 - 3 = 2$, $g_3 = 2$. Aangezien alle priemgetallen $p > 2$ oneven zijn, is g_n even voor $n > 1$.

Lemma 3.15. *We kunnen twee priemgetallen vinden waarvoor g_n arbitrair groot is.*

Bewijs. Zij m een willekeurig getal. Bekijk de rij $(m! + i)$ voor $2 \leq i \leq m$. Elke getal in deze rij is samengesteld, immers $2 \mid m! + 2$, $3 \mid m! + 3$, enzovoorts. Hiermee hebben we een priemgetalverschil van minstens $m - 1$ gevonden. □

Toch kunnen we iets zeggen over het gemiddelde priemgetalverschil rondom p_n. In elk geval geldt $g_n < p_n$, omdat er voor elke m altijd een priemgetal tussen m en $2m$ zit (postulaat van Bertrand), dus $g_n = p_{n+1} - p_n < p_n$. Ook volgt uit de priemgetalstelling van Hadamard dat het gemiddelde verschil tussen priemgetallen van grootte N ongeveer $\ln(N)$ is.

Het aantal verschillende priemfactoren van een getal n is ongeveer $\ln\ln(n)$, de stelling van Hardy-Ramanujan, en de sterkere stelling van Erdös-Kac zeggen hier iets over.

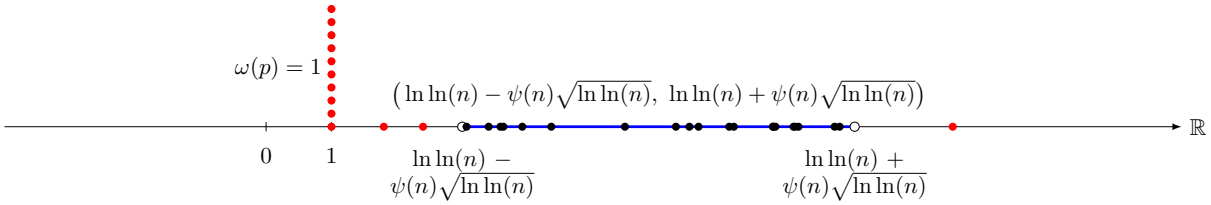
Stelling 3.16 (Hardy-Ramanujan)**.** *Zij $\omega(n)$ het aantal verschillende priemfactoren van n. Zij ψ een willekeurige functie $\psi : \mathbb{R} \to \mathbb{R}$ met $\psi(x) \to \infty$ voor $x \to \infty$ en zij*

$$U = \{1 \leq n \leq B : |\omega(n) - \ln\ln(n)| < \psi(n)\sqrt{\ln\ln(n)}\}$$
$$V = \{1 \leq n \leq B\}$$

Dan geldt

$$\lim_{B\to\infty} \frac{\#U}{\#V} = 1$$

Zie Figuur 4 voor een voorstelling.



Figuur 4: De zwarte punten zijn $n \in \mathbb{N}_{\geq 1}$ waarvoor $\omega(n)$ binnen het interval uit de stelling valt. De rode punten vallen er buiten. Merk op dat voor p priem geldt $\omega(p) = 1$. De stelling zegt dat voor B voldoende groot $\omega(n)$ meestal binnen het interval zal vallen.

De convergentiesnelheid hangt natuurlijk wel af van de keuze voor ψ.

3.3 Gladde getallen

Definitie 3.17 (Gladde getallen [Wag13, §3.1])**.** Een positief geheel getal is y-glad als al zijn priemfactoren kleiner of gelijk aan y zijn. De de Bruijn-functie, $\psi(x, y)$ geeft het aantal y-gladde gehele getallen tussen 1 en x (inclusief x).

Definitie 3.18 (Machtgladde getallen)**.** Een positief geheel getal is n-machtglad als voor al zijn priemmachten geldt $p_i^{\alpha_i} \leq n$.

Definitie 3.19 (Kwadraatvrij getal)**.** Een kwadraatvrij getal is een getal dat niet door kwadraatgetal, behalve 1^2, kan worden gedeeld. Dit zijn precies de getallen waarvan alle priemfactoren de exponent 1 hebben.

Voorbeeld: $35 = 5 \cdot 7$ is kwadraatvrij, maar $12 = 2^2 \cdot 3$ is niet kwadraatvrij.

Stelling 3.20 (Dickman [HT93]). *Voor elk vast reëel getal $u > 0$ bestaat er een reëel getal $\rho(u) > 0$ zodanig dat*

$$\psi(x, x^{1/u}) \sim \rho(u)x,$$

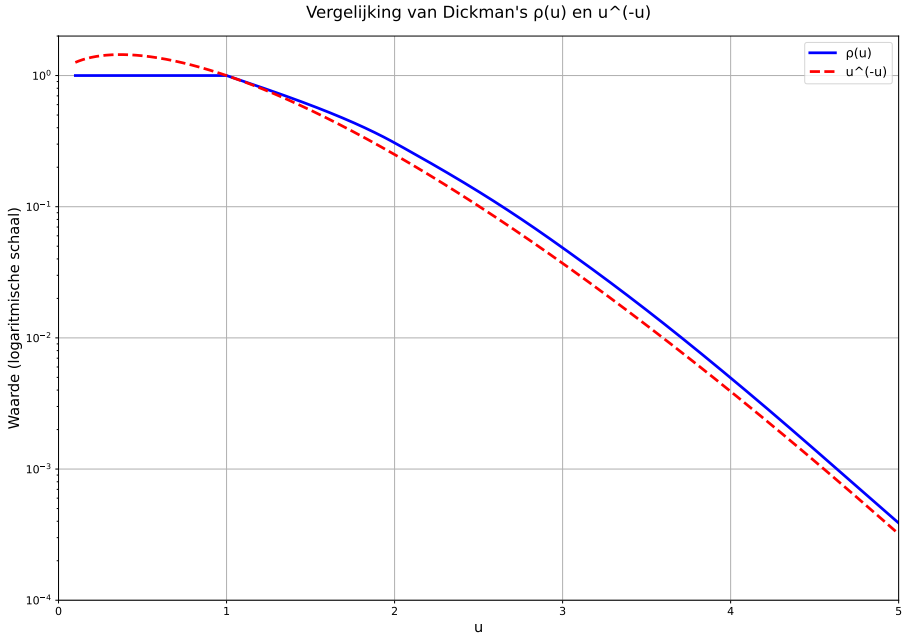
waarbij ρ de oplossing van de volgende differentiaalvergelijking is

$$u\rho'(u) = -\rho(u-1) \qquad \text{als } u > 1$$

met beginwaarde

$$\rho(u) = 1 \qquad \text{als } 0 \leq u \leq 1.$$

Er bestaat geen oplossing in gesloten vorm van de differentiaalvergelijking $\rho'(u) = \frac{-\rho(u-1)}{u}$, maar de functie $u \mapsto u^{-u}$ is een goede benadering.



Figuur 5: Het verschil tussen de Dickman-functie $\rho(u)$ en de benadering u^{-u}.

4 Computationele tijdscomplexiteit

Definitie 4.1 (Grote-O notatie [Cor+22, §3.2])**.** Voor een gegeven functie $g(n)$ schrijven we $\mathcal{O}(g(n))$ voor de verzameling functies:

$$\mathcal{O}(g(n)) = \{f(n) : \text{ er bestaan positieve constanten } c \text{ en } n_0 \text{ zodat } 0 \leq f(n) \leq cg(n) \text{ voor alle } n \geq n_0\}$$

Een functie $f(n)$ behoort tot de verzameling $\mathcal{O}(g(n))$ als er een positieve constante c bestaat zodat $f(n) \leq c \cdot g(n)$ voor voldoende grote n.

4.1 Complexiteit in aantal cijfers

Bij het analyseren van algoritmen in de getaltheorie is het essentieel onderscheid te maken tussen complexiteit in termen van de grootte van het getal zelf en de complexiteit in termen van het aantal bits of cijfers dat nodig is om het getal te representeren.

Voor een natuurlijk getal n met m decimale cijfers geldt:

- Het aantal bits k benodigd voor representatie is: $k = \lfloor \log_2(n) \rfloor + 1$
- De relatie tussen decimale cijfers en bits is: $k \approx m \log_2(10) \approx 3.32 \cdot m$

Definitie 4.2 (Complexiteitsnotatie)**.** We gebruiken de volgende notatie voor complexiteitsanalyse:

- $\mathcal{O}_{\text{op}}(n)$: complexiteit uitgedrukt in het aantal operaties als functie van n
- $\mathcal{O}_{\text{bit}}(k)$: complexiteit uitgedrukt in het aantal bits k van de invoer

4.2 Complexiteitsklassen

We onderscheiden de volgende belangrijke complexiteitsklassen: Zij k de grootte van de invoer in bits en c een constante.

4.2.1 Polynomiale tijd

Een algoritme heeft polynomiale tijdscomplexiteit als zijn looptijd wordt begrensd door een polynoom in het aantal bits van de invoer. Formeel: $\mathcal{O}_{\text{bit}}(k^c)$.

4.2.2 Exponentiële tijd

Een algoritme heeft exponentiële tijdscomplexiteit als zijn looptijd exponentieel groeit met de grootte van de invoer. Formeel: $\mathcal{O}_{\text{bit}}\left(c^k\right)$, voor $c > 1$.

4.2.3 Subexponentiële tijd

Een algoritme heeft subexponentiële tijdscomplexiteit als zijn looptijd asymptotisch kleiner is dan elke exponentiële functie maar groter dan elke polynomiale functie. Een voorbeeld is $2^{\sqrt{k}}$.

Het simpele factoriseeralgoritme proefdeling loopt in exponentiële tijd. De beste factoriseeralgoritmes lopen in subexponentiële tijd.

5 Priemtesten

Priemtesten zijn een belangrijk onderdeel van factorisatiealgoritmen. Voordat we een getal proberen te factoriseren is het belangrijk om te checken of het wel samengesteld is. Ook geven de meeste factorisatiealgoritmen een niet-triviale factor van een getal, maar dat hoeft geen priemgetal te zijn. Indien het geen priemgetal is moeten we de factor verder factoriseren.

5.1 Classificatie van priemtests

Priemtests kunnen worden ingedeeld in twee categorieën:

1. **Deterministische tests**
2. **Probabilistische tests**

Deterministische tests geven met absolute zekerheid aan of een getal priem is. De probabilistische tests zeggen dat een getal ofwel zeker samengesteld is ofwel waarschijnlijk priem is. De probabilistische tests kunnen veel grotere getallen testen op samengesteldheid dan we kunnen factoriseren.

Er bestaan zeer efficiënte polynomiale priemtests voor Mersenne en- Fermatgetallen. Priemgetallen van deze types zijn echter zo zeldzaam dat ze in de cryptografie niet relevant zijn.

5.2 Deterministische priemtests

5.2.1 Proefdelingen

Het proefdelingsalgoritme kan worden gebruikt om te bepalen of n priem is. Het checkt voor alle gehele getallen $k \leq \sqrt{n}$ of $k \mid n$. Het is voldoende om de getallen tot $\sqrt{n}$ te controleren, stel immers dat $n = ab$ en $a > \sqrt{n}$ en $b > \sqrt{n}$, dan $n = ab > \sqrt{n}\sqrt{n}$, contradictie. Dus indien n samengesteld is heeft n een factor k met $k \leq \sqrt{n}$.

De complexiteit is $\mathcal{O}_{\text{op}}(\sqrt{n})$ of $\mathcal{O}_{\text{bit}}\left(2^{\frac{n}{2}}\right)$.

5.3 Probabilistische priemtests

Definitie 5.1 (Priemtest van Fermat [Weg22, §3.1])**.** Zij n het te testen getal en a copriem met n. Als $a^{n-1} \not\equiv 1 \pmod{n}$, dan is n zeker samengesteld. Als $a^{n-1} \equiv 1 \pmod{n}$, dan is n mogelijk priem. In dit laatste geval:

- Als n samengesteld is, heet n een pseudopriem voor grondtal a
- Het getal a heet een Fermat-leugenaar voor n

Bijvoorbeeld voor $a = 2, n = 341 = 11 \cdot 31$ geldt $2^{340} \equiv 1 \pmod{341}$. In dit geval is 2 een Fermat-leugenaar voor 341, en 341 heet pseudopriem voor 2.

Definitie 5.2 (Carmichael-getal)**.** Een samengesteld getal n heet een Carmichael-getal als voor alle a met $\gcd(a, n) = 1$ geldt:

$$a^{n-1} \equiv 1 \pmod{n}$$

Het kleinste Carmichael-getal is 561. Er bestaan oneindig veel Carmichael-getallen.

We kunnen met het criterium van Korselt de Carmichael-getallen eenvoudig herkennen als we hun priemfactorisatie kennen.

Stelling 5.3 (Criterium van Korselt)**.** *Een samengesteld getal n is een Carmichael-getal dan en slechts dan als:*

1. *n is kwadraatvrij*
2. *n heeft ten minste drie priemfactoren*
3. *n is oneven*
4. *Voor elke priemfactor p van n geldt: $p-1 \mid n-1$.*

5.4 Miller-Rabin

Vanwege het bestaan van de Carmichael getallen is de priemtest van Fermat niet betrouwbaar. De Miller-Rabin test gebruikt daarom een gewijzigde versie van de kleine stelling van Fermat.

Lemma 5.4 ([CP05, § 3.5])**.** *Zij p een oneven priemgetal en schrijf $p-1 = 2^s t$ met t oneven. Dan geldt voor alle $a \in \mathbb{Z}$ met $\gcd(a,p) = 1$ precies één van de volgende beweringen:*

1. *$a^t \equiv 1 \pmod{p}$, of*
2. *$\exists i \in \{0,1,\ldots,s-1\} : a^{2^i t} \equiv -1 \pmod{p}$*

Bewijs. We gebruiken de kleine stelling van Fermat en Lemma 3.13. Zij p een oneven priemgetal en schrijf $p-1 = 2^s t$ met t oneven. Veronderstel dat $p \nmid a$.

Volgens de kleine stelling van Fermat geldt:

$$a^{p-1} = a^{2^s t} \equiv 1 \pmod{p}$$

Definieer de verzameling

$$U = \{x \in \mathbb{N}_{\geq 0} : 0 \leq x \leq s \text{ en } a^{2^x t} \equiv 1 \pmod{p}\}$$

Merk op dat $U \neq \emptyset$ aangezien $s \in U$. Volgens het welgeordendheidsprincipe heeft U een kleinste element; noem dit y. We onderscheiden twee gevallen:

Geval 1: Als $y = 0$, dan geldt $a^t \equiv 1 \pmod{p}$ en zijn we klaar.

Geval 2: Stel $y > 0$. We weten dat $a^{2^y t} \equiv 1 \pmod{p}$. Beschouw nu $y-1$. Volgens Lemma 3.13 geldt:

$$a^{2^{y-1} t} \equiv \pm 1 \pmod{p}.$$

Vanwege de minimaliteit van y kan niet gelden dat $a^{2^{y-1} t} \equiv 1 \pmod{p}$, dus moet wel gelden dat $a^{2^{y-1} t} \equiv -1 \pmod{p}$. □

Definitie 5.5. Een getal $n > 3$ heet een *strong probable prime* voor grondtal a, met $1 < a < n-1$, als een van de gevallen van Lemma 5.4 voor n geldt.

Definitie 5.6. Een oneven samengesteld getal $n = 2^s \cdot t + 1$, met t oneven heet een *sterk pseudopriem* voor grondtal a als voor n een van de beweringen van Lemma 5.4 geldt. Merk op dat elk getal n een sterk pseudopriem is voor grondtal $a \equiv 1 \pmod{n}$.

Voorbeeld Zij grondtal $n = 341 = 11 \cdot 31$, $n - 1 = 2^2 \cdot 85$, $t = 85$, $s = 2$. We laten zien dat n voor grondtal $a = 2$ niet aan *Lemma* 5.4 voldoet, en dus niet een sterk pseudopriem

is voor grondtal 2.

$$
\begin{aligned}
a^t = 2^{85} &\equiv 32 \pmod n \not\equiv 1 \pmod n && \text{voldoet niet aan voorwaarde (1)}\\
a^{2^0 t} = 2^{1\cdot 85} &\equiv 1 \pmod n \not\equiv -1 \pmod n && \text{voldoet niet aan voorwaarde (2) voor } i = 0\\
a^{2^1 t} = 2^{2\cdot 85} &\equiv 1 \pmod n \not\equiv -1 \pmod n && \text{voldoet niet aan voorwaarde (2) voor } i = 1\\
a^{2^2 t} = 2^{4\cdot 85} &\equiv 1 \pmod n \not\equiv -1 \pmod n && \text{voldoet niet aan voorwaarde (2) voor } i = 2
\end{aligned}
$$

Zij nu $n = 91 = 7 \cdot 13$, $n - 1 = 2 \cdot 3^2 \cdot 5$, $t = 45$, $s = 1$ en $a = 10$. Nu geldt $10^{2^0} \equiv -1 \pmod n$, dus 91 is een sterk pseudopriem voor grondtal $a = 10$.

De kans is klein dat een samengesteld getal n pseudopriem is voor een grondtal a_1, en nog kleiner dat n pseudopriem is voor verschillende grondtallen $a_1, a_2, \ldots a_j$. Een bovengrens voor de kans wordt gegeven door Rabin in [Rab80].

We definiëren de verzameling

$$\mathcal{S}(n) = \{a \pmod n : n \text{ is sterk pseudopriem voor grondtal } a\ \}$$

En zij $S(n) = |\mathcal{S}(n)|$ (kardinaliteit van $\mathcal{S}(n)$.

Lemma 5.7. *Voor een samengestelde $n > 9$ geldt $S(n) \leq \frac{\phi(n)}{4}$*

Definitie 5.8. Een getuige $1 \leq a \leq n - 1$ voor een oneven samengestelde n is een grondtal waarvoor Lemma 5.4 onwaar is. Het is genoeg om 1 getuige te vinden als bewijs dat n samengesteld is, aangezien

$$n \text{ niet sterk pseudopriem} \implies n \text{ niet priem.}$$

De priemtest van Miller-Rabin voor n probeert net zolang grondtallen a totdat ofwel een getuige is gevonden, ofwel de kans voldoende groot is dat n niet samengesteld is.

Algorithm 1 Miller-Rabin priemtest iteratie

Require: Een oneven getal $n > 3$
Ensure: Een getuige a voor de samengesteldheid van n, of een grondtal a waarvoor n een strong probable prime is
  Bereken s, t zodanig dat $n - 1 = 2^s \cdot t$ met t oneven
  $a \leftarrow$ willekeurig getal in $\{2, 3, \ldots, n - 1\}$
  $b \leftarrow a^t \pmod n$
  if $b = 1 \vee b = n - 1$ **then** ▷ Geval (1) en Geval (2) voor $i = 0$ van Lemma 5.4
    Print "n is een *strong probable prime* voor grondtal a" **return**
  end if
  for $r \in \{1, 2, \ldots, s - 1\}$ **do**
    $b \leftarrow b^2 \pmod n$ ▷ dus $b = a^{2^r \cdot t}$
    if $b = n - 1$ **then** ▷ Geval (2) voor $i \geq 1$ van Lemma 5.4
      Print "n is een *strong probable prime* voor grondtal a" **return**
    end if
  end for
  Print "n is samengesteld met getuige a" **return**

Door de Miller-Rabin test te herhalen voor enkele verschillende a is de kans dat er een getuige wordt gevonden voor een samengestelde n zeer groot. Indien er geen getuige wordt gevonden mag men er van uit gaan dat n priem is.

6 Proefdelingen

De meest eenvoudige manier om een getal te factoriseren is door voor alle getallen $k \leq \sqrt{N}$ te controleren of $k \mid N$. De complexiteit is $\mathcal{O}_{\text{op}}\left(\sqrt{N}\right)$, en $\mathcal{O}_{\text{bit}}\left(2^{\frac{N}{2}}\right)$ in het aantal bits van de invoer. Dit is exponentiëel in het aantal bits. In praktijk werkt proefdeling maar tot getallen van ≈ 50 bits.

Algorithm 2 Proefdeling Factorisatie - geeft alle factoren

Require: $N > 1$
  $m \leftarrow N$
  $d \leftarrow 2$
  while $d \leq \sqrt{m}$ **do**
    if $m \pmod d = 0$ **then**
      print 'd deelt N'
      $m \leftarrow m/d$
    else
      $d \leftarrow d + 1$
    end if
  end while
  if $m = N$ **then**
    print 'N is priem'
  else if $m > 1$ **then**
    print 'm deelt N'
  end if

6.1 Verbeteringen

6.1.1 Priemlijst

Het is niet nodig om voor alle $d \leq \sqrt{N}$ te testen of $d \mid N$, maar alleen voor priemgetallen $p \leq \sqrt{N}$. Wanneer we een lijst met alle priemgetallen p_i hebben van 2 tot $\sqrt{N}$ hoeven we slechts $\pi(\sqrt{N})$ getallen te checken, dus $\approx \frac{\sqrt{N}}{\ln \sqrt{N}}$.

Zeef van Eratosthenes De lijst met priemgetallen kan snel gemaakt worden met de zeef van Eratosthenes. Het is in praktijk sneller om priemgetallen te genereren dan om ze uit het computergeheugen te lezen. Een goede parallele implementatie van de zeef kan de eerste $\approx 10^9$ priemgetallen in een seconde vinden op een moderne computer.

6.2 Wielfactorisatie

Wielfactorisatie is een optimalisatie van proefdeling waarbij we systematisch getallen overslaan die deelbaar zijn door een set van kleine priemgetallen. Het wiel voor een product W van eerste priemgetallen wordt geconstrueerd door alle getallen modulo W te nemen die copriem zijn met W.

Bijvoorbeeld, voor $W = 2 \cdot 3 = 6$ zijn de getallen die we moeten controleren: $\{n : n \equiv 1, 5 \pmod 6\}$

Voor $W = 2 \cdot 3 \cdot 5 = 30$ zijn dit: $\{n : n \equiv 1, 7, 11, 13, 17, 19, 23, 29 \pmod{30}\}$

Deze methode is efficiënter dan alleen even getallen overslaan, omdat we direct grotere sets van samengestelde getallen kunnen uitsluiten.

Lemma 6.1. *Elk priemgetal $p > 3$ is van de vorm $6m \pm 1$.*

Bewijs. Alle $n \geq 3$ zijn van de vorm $6k, 6k \pm 1$, $6k \pm 2$, $6k \pm 3$, en $2|6k$, $2|6k \pm 2$, $3|6k \pm 3$. Dus alle priemgetallen zijn van de vorm $6k \pm 1$. □

We kunnen dit uitbreiden tot elk priemgetal is van de vorm $30m \pm n$ met $n \in \{1, 7, 11, 13\}$, en verder. Een zeer snelle multithreaded implementatie van de zeef van Eratosthenes wordt gegeven door [Wal25]. Deze vindt op een MacBook M2 (2022) alle priemgetallen tot 10^{14} in een uur, zie Figuur 6.

Figuur 6: Een snelle implementatie van de priemzeef vindt alle priemgetallen tussen 1 tot 10^{14} in iets meer dan een uur op een MacBook M2

7 Pollards $p-1$-algoritme

Het $p-1$-factorisatiealgoritme van Pollard [Pol74] uit 1974 was samen met Pollards Rho algoritme het eerste snelle factorisatiealgoritme. In 1987 zijn enkele verbeteringen gevonden door Montgomery, zie [Mon87].

Het algoritme is gebaseerd op de kleine stelling van Fermat.

Lemma 7.1 (Kleine stelling van Fermat)**.** *Voor priemgetal p en geheel getal a met $p \nmid a$ geldt:*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Opmerking 7.2. Merk op dat de orde van een element in $(\mathbb{Z}/p\mathbb{Z})^\times$ een deler van $p-1$ is, maar niet gelijk aan $p-1$ hoeft te zijn. Bijvoorbeeld $2^3 \equiv 1 \pmod{7}$.

De looptijd van Pollards algoritme hangt af van de grootte van de kleinste priemdeler. Het is een probabilistisch algoritme, dat wil zeggen dat er geen garantie is dat het een priemfactor van n zal vinden. Na het vinden van een priemfactor p wordt het algoritme opnieuw toegepast op $\frac{n}{p}$ net zolang tot alle priemfactoren zijn gevonden.

7.1 Werking van het algoritme

Zij $n \geq 2$ een samengesteld getal, dus $n = pq$ met p priem. Kies een willekeurige c zodanig dat $\gcd(c, p) = 1$. Zij nu L zodanig dat $p-1$ een factor is van L, dus $L = (p-1)k$, dan

$$\begin{aligned} c^L &\equiv c^{(p-1)k} \pmod{p} \\ &\equiv (1)^k \pmod{p} \\ &\equiv 1 \pmod{p} \end{aligned}$$

Hieruit volgt dat $c^L - 1 \equiv 0 \pmod{p}$ en dus $p \mid c^L - 1$. Omdat $p \mid n$, geldt ook dat $p \mid \gcd(c^L - 1, n)$.

Nu is het zo dat we p niet weten en juist zoeken. Maar we kunnen wel een slimme keuze maken voor L en dan $\gcd(c^L - 1, n)$ uitrekenen. Er is een goede kans dat L een deler $p-1$ voor één of meerdere priemfactoren p van n heeft, maar niet voor alle priemfactoren. In dat geval is

$$c^L - 1 \equiv 0 \pmod{p} \qquad \text{én} \qquad c^L - 1 \not\equiv 0 \pmod{n}$$

De reden dat Pollards algoritme aanzienlijk sneller is dan proefdeling komt doordat mogelijke priemfactoren van n in één stap van het algoritme kunnen proberen. Door L zo te kiezen dat L veel delers van de vorm $p-1$ heeft kunnen we veel priemgetallen p als mogelijke delers van n in één keer proberen.

7.2 Versimpelde versie van het algoritme

We geven nu eerst de allersimpelste versie van het algoritme, daarna een verbeterde versie die in de praktijk toepasbaar is.

Eerst het geval dat het algoritme (meestal) slaagt. Zij n samengesteld met priemfactor p. Zij $p-1$ een 5-glad getal. Nu $p - 1 = 2^{\alpha_2} \cdot 3^{\alpha_3} \cdot 5^{\alpha_5}$. Voor de α_i moet gelden dat $2^{\alpha_2} \leq p-1$, $3^{\alpha_3} \leq p-1$ en $5^{\alpha_5} \leq p-1$.

Door $\alpha_2 = \lfloor \log_2(n) \rfloor$, $\alpha_3 = \lfloor \log_3(n) \rfloor$, $\alpha_5 = \lfloor \log_5(n) \rfloor$ te kiezen wordt hieraan voldaan.

In regel 4 is nu $L = 2^{\lfloor \log_2 n \rfloor} \cdot 3^{\lfloor \log_3 n \rfloor} \cdot 5^{\lfloor \log_5 n \rfloor}$. Uit de 5-gladheid van $p-1$ en de keuzes voor α_i volgt $p-1 \mid L$. De stelling van Fermat zegt nu dat $c^L \equiv 1 \pmod{p}$. Zij $g = \gcd(c^L - 1, n)$.

Algorithm 3 Simpelste versie van Pollards $p-1$ algoritme voor $p-1$ 5-glad

Require: Een samengestelde $n > 2$.
Ensure: Het algoritme print een niet-triviale factor van n of 'mislukt'

```
e_2 ← 2^⌊log_2 n⌋
e_3 ← 3^⌊log_3 n⌋
e_5 ← 5^⌊log_5 n⌋
L ← e_2 · e_3 · e_5
c ← 2^L (mod n)
g ← ggd(c − 1, n)
if 1 < g < n then
    print 'g is een factor van n'
else
    stop en print 'mislukt'
end if
```

Indien $1 < g < n$ hebben we een niet-triviale factor van n gevonden. Merk op dat hoewel $p \mid g$, niet hoeft te gelden dat g priem is. In dat geval passen we opnieuw het algoritme toe met $n = g$.

Nu het geval dat $p-1$ niet 5-glad is, dan $p-1 \nmid L$. Nu is waarschijnlijk (maar niet zeker, Opmerking 7.2) $c^L \not\equiv 1 \pmod{p}$, dus $\text{ggd}(c^L - 1, n) = 1$. We kunnen nu L vermenigvuldigen met $7^{\lfloor \log_7 n \rfloor}$ en dan weer $g = \text{ggd}(2^L - 1, n)$ uitrekenen. Indien $p-1$ wel 7-glad is, zal dit mogelijk een factor van n opleveren.

7.3 De functie L

Het bovenstaande Algoritme 3 werkt alleen als $p-1$ 5-glad is. Hoe kunnen we het algoritme aanpassen zodat het ook werkt als $p-1$ niet 5-glad is? We definiëren hiervoor een functie $L : \mathbb{N}_{\geq 2} \to \mathbb{N}_{\geq 2}, B \mapsto L(B)$, waarbij B de zoekgrens is. Deze B bepaalt hoeveel iteraties van het algoritme we maximaal willen uitvoeren voordat we het zoeken opgeven. Wat is een goede keuze voor de functie L? We geven eerst twee opties die niet goed blijken te werken en daarna twee betere keuzes.

7.3.1 Eerste poging

We kiezen $L(B) = B$. We willen dat $p-1 \mid L$, dus als we in elke iteratie $B \leftarrow B+1$ doen, zijn er $p-1$ iteraties nodig voordat $L(B) = p-1$. Dit is niet efficiënt. Er zijn immers $p-1$ operaties nodig, waardoor het ongeveer even efficiënt is als proefdeling.

Merk op dat niet hoeft te gelden dat $L = p-1$, het is voldoende dat $p-1 \mid L$ voor p een priemfactor van n, want dan $p \mid \text{ggd}(c^L - 1, n)$. Dit geeft het inzicht dat we door L zeer groot en met veel kleine priemdelers te kiezen meer kans hebben dat $p-1 \mid L$.

7.3.2 Tweede poging: product van priemgetallen

We definiëren $L(B)$ als het product van alle priemgetallen kleiner dan B:

$$L(B) = \prod_{p \text{ priem}, p<B} p$$

Met p_j het j-de priemgetal zodanig dat $p_j < B$. L groeit nu sneller, en er is veel kans dat L alle priemfactoren van $p-1$ bevat.

Wat zijn de priemfactoren van $p-1$? We bekijken eerst twee extreme gevallen. Indien $p = 2q+1$, met q priem, dan zijn 2 en q de enige priemfactoren en we zouden pas bij $B = q$ de factor q vinden. Indien p een Fermatpriem is dan $p-1 = 2^{2^n}$ dus heeft $p-1$ alleen maar de factor 2.

Ervan uitgaande dat de priemfactoren p van n willekeurig gekozen zijn, mogen we aannemen dat $p-1$ dezelfde deelbaarheidseigenschappen heeft als een willekeurig gekozen even (want p is oneven) getal (zie deelparagraaf 7.8). Het nadeel van deze keuze voor L is dat alle priemmachten van L van de vorm p_i^1 zijn. Als $p-1$ niet kwadraatvrij dan $p-1 \nmid L$.

Lemma 7.3. *Zij $n \geq 1$, zij s_n het aantal kwadraatvrije getallen in $\{1, 2, \ldots, n\}$ dan*

$$\lim_{n\to\infty} \frac{s_n}{n} = \frac{6}{\pi^2} \approx 0,61.$$

De kans is dus vrij groot dat $p-1$ niet kwadraatvrij is.

7.3.3 Derde poging: faculteit

$L(B) = B!$. In iedere iteratie kan $c^{L(B+1)}$ kan snel berekend worden als $c = c^{L(B)}$ door $c \leftarrow c^{B+1}$. Voorbeeld

$$a^{3!} \pmod n \equiv \left(a^{2!}\right)^3 \pmod n$$

Merk op dat L nu het product is van veel kleine priemfactoren, dus de kans dat $p-1 \mid L$ is groot. We bekijken de priemfactorisatie van de faculteit van een getal $n!$ in meer detail.

Lemma 7.4. *Voor elke priemfactor p van $n!$ geldt $p \leq n$. En voor elke priem $2 \leq q \leq n$ geldt dat q een priemfactor is van $n!$*

Bewijs. Zij p een priemfactor van $n!$. Merk op dat

$$p \mid a_1 \cdot a_2 \cdot \ldots \cdot a_k \implies p \mid a_1 \vee p \mid a_2 \vee \ldots \vee p \mid a_k$$

Dit impliceert dat p ten minste één getal j deelt voor $1 \leq j \leq n$, dus $p \leq n$. Het tweede deel van het lemma is triviaal. □

We weten nu dat $n!$ precies alle priemfactoren $2 \leq q \leq n$ bevat. Maar wat kunnen we zeggen over de exponenten?

Lemma 7.5 (Formule van Legendre)**.** *Zij p priem en $n \in \mathbb{N}_{\geq 1}$. Zij $\nu_p(n)$ de exponent van de grootste macht van p die n deelt. Dan*

$$\nu_p(n!) = \sum_{i=1}^{\lfloor \log_p(n) \rfloor} \left\lfloor \frac{n}{p^i} \right\rfloor$$

Bewijs. Het aantal factoren van de vorm p in $\{1, 2, \ldots, n\}$ is $\left\lfloor \frac{n}{p} \right\rfloor$. Het aantal factoren van de vorm p^2 is $\left\lfloor \frac{n}{p^2} \right\rfloor$ enzovoort. Merk op dat er voor $p^a > n$ geen factoren zijn, vandaar de bovengrens van de sommatie. □

Gevolg 7.6.

$$\nu_p(n!) \geq \nu_q(n!) \qquad \textit{voor } p, q \textit{ priem met } p < q$$

Bewijs. Zij p, q priem en $p < q$, dan $\left\lfloor \frac{n}{p^i} \right\rfloor \geq \left\lfloor \frac{n}{q^i} \right\rfloor$ voor alle $i \geq 1$.

Bovendien is de bovengrens van de sommatie $\lfloor \log_p(n) \rfloor \geq \lfloor \log_q(n) \rfloor$. Dus $\nu_p(n!)$ heeft meer termen. □

We zien nu dat de exponenten van de priemfactorisatie van $n!$ een niet-stijgende rij vormen. Bijvoorbeeld

$$37! = 2^{32} \cdot 3^{15} \cdot 5^7 \cdot 7^4 \cdot 11^3 \cdot 13^2 \cdot 17^2 \cdot 19^1 \cdot 23 \cdot 29 \cdot 31$$

7.3.4 Vierde poging: product van priemmachten

Definitie 7.7. Zij $M(B)$ het product van priemmachten $p_i^{\alpha_i} \leq B$.

$$M(B) = \prod_{j=1}^{n} p_j^{\lfloor \log_{p_j}(B) \rfloor}.$$

Met p_j het j-de priemgetal en n het grootste getal zodanig dat $p_n \leq B$.

Kies $L(B) = M(B)$. Dan geldt $p-1 \mid M(B)$ als $p-1$ B-machtglad is (Lemma 7.8), dus wanneer alle priemmachten van $p-1$ kleiner of gelijk B zijn.

Lemma 7.8. *$M(B)$ is deelbaar door alle B-machtgladde getallen*

Bewijs. Zij k een B-machtglad getal. Zij $p_i^{\alpha_i}$ een priemmacht van k, dan $p_i^{\alpha_i} \leq B$. Dus $p_i^{\alpha_i} \mid M(B)$. Aangezien $p_i^{\alpha_i}$ een willekeurige priemmacht van k is geldt nu voor alle priemmachten van k dat ze $M(B)$ delen. Omdat de priemmachten copriem zijn volgt $k \mid M(B)$. □

De keuze voor $L = M(B)$ werkt dus als $p-1$ is B-machtglad. De keuze voor $L = B!$ of $L = M(B)$ maakt in praktijk niet veel verschil. De functie $B!$ is makkelijker te berekenen. Voor $M(B)$ is het nodig om eerst de priemgetallen tot B te berekenen, maar dat kan heel snel.

7.4 Relatie met de chinese reststelling

Zij $n = pq$ met p, q priem. Neem B zodanig dat $p-1$ is B-glad en $q-1$ is niet B-glad. Beschouw het systeem

$$\begin{aligned} x &\equiv a^{M(B)} - 1 \pmod{p} \\ x &\equiv a^{M(B)} - 1 \pmod{q} \end{aligned}$$

Nu volgt uit het feit dat $p-1$ B-glad is dat $a^{M(B)} \equiv 1 \pmod{p}$, dus $a^{M(B)} - 1 \equiv 0 \pmod{p}$. Omdat $q-1$ niet B-glad is geldt waarschijnlijk (maar niet zeker) dat $a^{M(B)} - 1 \not\equiv 0 \pmod{q}$. Als dit wel zo is mislukt het algoritme dus neem voor de eenvoud aan dat $a^{M(B)} - 1 \equiv m \pmod{q}$, met $m \neq 0$.

We hebben nu dus het systeem

$$\begin{aligned} a^{M(B)} - 1 &\equiv 0 \pmod{p} \\ a^{M(B)} - 1 &\equiv m \pmod{q} \end{aligned}$$

Omdat p, q copriem zijn kunnen we de chinese reststelling toepassen, die zegt dat dit systeem een unieke oplossing heeft. Noem deze oplossing x, dan

$$x \equiv a^{M(B)} - 1 \pmod{n}$$

Wanneer we nu $\gcd(a^{M(B)-1}, n)$ toepassen zullen we de priemfactor p vinden, omdat $a^{M(B)} - 1 \equiv 0 \pmod{p}$, dus $p \mid a^{M(B)} - 1$, en $a^{M(B)} - 1 \not\equiv 0 \pmod{q}$, dus $q \nmid a^{M(B)} - 1$, dus $\gcd(a^{M(B)-1}, n) \neq n$.

7.5 Verbeterde versie van Pollards algoritme

We passen nu de keuze $L = M(k)$ toe voor $1 \leq k \leq B$. Ook controleren we niet meer alleen aan het einde de $\gcd(c^L - 1, n)$, maar elke G stappen.

7.5.1 Backtracking

Het verbeterde algoritme maakt gebruik van backtracking voor het geval dat $\text{ggd}(c^L - 1, n) = n$. Stel dat $n = pq$ (argument werkt ook voor 3 of meer priemfactoren) met p, q priem en $p-1$ is 13-machtglad en $q-1$ is 17-machtglad. Stel dat $G = 4$, dus we controleren de ggd elke 4 stappen.

Neem bijvoorbeeld $n = 5459$, $p = 53$, $q = 103$, $p - 1 = 2^2 \cdot 13$, $q - 1 = 2 \cdot 3^2 \cdot 17$. Dus $p-1$ is 13-machtglad en $q-1$ is 17-machtglad, kies $c = 2$.

We doen de eerste ggd-test bij $k = 4$ (dus product van eerste 4 priemgetallen $\leq B$) en vinden $\text{ggd}(c^{2^3 \cdot 3^2 \cdot 5 \cdot 7} - 1, n) = 1$. We gaan door en doen we de tweede ggd-test bij $k = 8$ en vinden dat $\text{ggd}(c^{2^4 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19} - 1, n) = n$.

Nu moeten we terug (backtracken) naar $k = 4$ en een ggd-test doen voor $k = 5$ tot $k = 6$ waar we vinden $\text{ggd}(c^{2^4 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 13} - 1, n) = 53$. Zie Algoritme 4 voor het verbeterde algoritme.

Tabel 1: Voorbeeld van backtracking bij $G = 4$

Iteratie	**k**	**Berekening**
1	4	$\text{ggd}(2^{2^3 \cdot 3^2 \cdot 5 \cdot 7} - 1, 5459) = 1$
2	8	$\text{ggd}(2^{2^4 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19} - 1, 5459) = 5459$
3	5	$\text{ggd}(2^{2^4 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11} - 1, 5459) = 1$
4	6	$\text{ggd}(2^{2^4 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 13} - 1, 5459) = 53$

7.6 Wanneer vindt het algoritme geen factoren?

Zij $g = \text{ggd}(a^L - 1, n)$. Het algoritme mislukt als aan het einde $g = 1$ of $g = n$.

1. $\text{ggd}(a^L - 1, n) = 1$. In dit geval zijn er geen priemfactoren p van n waarvoor $p-1$ B-machtglad is. In dit geval heeft $p-1$ een priemfactor $p_i^{\alpha_i} > B$. (We gaan er hier vanuit dat L de functie M uit Definitie 7.7 is). Oplossingen zijn om B te verhogen of naar de tweede fase van het algoritme te gaan (zie deelparagraaf 7.9).

   Voorbeeld Zij $n = 1829 = 31 \cdot 59$, zij $B = 4$ dus de priemmachten zijn $p_1 = 2^2$, $p_2 = 3$. Kies $c = 2$ en bereken $c^L - 1 = 2^{2^2 \cdot 3} - 1 \equiv 437 \pmod{n}$, nu is $\text{ggd}(437, 1829) = 1$.

   We verhogen B naar 5, nu is $c^L - 1 = 2^{2^2 \cdot 3 \cdot 5} - 1 \equiv 62 \pmod{n}$, dus $\text{ggd}(62, 1829) = 31$. We hebben de factor 31 van 1829 gevonden.

2. $\text{ggd}(a^L - 1, n) = n$. Er kunnen meerdere oorzaken zijn

   - De ggd-test wordt niet vaak genoeg uitgevoerd, en alle priemfactoren worden tegelijk gevonden. Backtracken is de oplossing.
   - Alle priemfactoren van n zijn B-machtglad. Ze worden dus in dezelfde iteratie gevonden. Hier is geen oplossing voor.

Algorithm 4 Verbeterde versie van Pollards $p-1$ algoritme

Require: Een samengestelde $n > 2$, een zoekgrens B, een staplengte G hoe vaak de ggd wordt gestest.

Ensure: Een niet-triviale factor g van n, of stop en print 'mislukt' anders

```
Maak een lijst p_1 = 2, p_2 = 3, p_3, ..., p_k : ∀ 1 ≤ i ≤ k : p_i ≤ B
Maak een lijst a_i van exponenten bij elke p_i zodat ∀ 1 ≤ i ≤ k : p_i^{a_i} ≤ B. Dus a_i = ⌊log_{p_i} B⌋
c ← 2                                                    ▷ Mag willekeurig getal zijn
for i ← 1 to k do
    f ← p_i^{a_i}                                        ▷ Gebruik snelle exponentiatie
    c ← c^f (mod n)                                      ▷ Gebruik snelle exponentiatie
    if G | k then          ▷ Elke G iteraties een ggd check of de priemfactor al gevonden is
        g ← ggd(c − 1, n)
        if 1 < g < n then
            print 'g is een factor van n'
        else
            c' ← c                          ▷ c en k opslaan voor als backtracken nodig blijkt
            k' ← k
        end if
    end if
end for
g ← ggd(c − 1, n)
if 1 < g < n then
    print 'g is een factor van n'
else
    if g = n then                                        ▷ Backtracken
        c ← c'
        for i = k' to k do
            f ← p_i^{a_i}
            c ← c^f (mod n)
            g ← ggd(c − 1, n)
            if 1 < g < n then
                print 'g is een factor van n'
            end if
        end for
    end if
    if g = 1 then
        stop en print 'mislukt'
    end if
end if
```

Voorbeeld $n = 1247 = 29 \cdot 43$, en $29 - 1 = 2^2 \cdot 7$ en $43 - 1 = 2 \cdot 3 \cdot 7$. Nu hebben 42 en 28 dezelfde grootste priemmacht 7. Dus indien $B \geq 7$ zullen ze in dezelfde ggd-test gevonden worden.

- $g = 1$ en door het verhogen van B zodat er 1 priemfactor bij komt wordt meteen $g = n$. We illustreren dit met een voorbeeld.

Voorbeeld Zij $n = 2047 = 2^{11} - 1 = 39 \cdot 89$. We gebruiken $c = 2$. Voor n geldt dat $g = \text{ggd}(2^M - 1, n) = n$ als $11 \mid M$ en $g = 1$ als $11 \nmid M$. Hierdoor zullen we nooit een factor vinden als $c = 2$.
Lemma 7.9. *Zij M samengesteld met $d \mid M$ en $d > 1$, dan heeft $2^M - 1$ een niet-triviale factor $2^d - 1$.*

Bewijs.

$$2^M - 1 = 2^{dk} - 1 = (2^d)^k - 1 = (2^d - 1)((2^d)^{k-1} + (2^d)^{k-2} + \ldots + (2^d)^1 + (2^d)^0).$$

□

We nemen nu $d = 11$ en dan volgt uit het lemma inderdaad dat voor $11 \mid M$ geldt dat $n = 2^{11} - 1$ een factor is van $2^M - 1$.

In het geval dat $11 \nmid M$ geldt $\text{ggd}(2^M - 1, n) = 1$. Dit is eenvoudig te controleren doordat M in dit geval van de vorm is: $M = 2$, $M = 2 \cdot 3$, $M = 2 \cdot 3 \cdot 5$ of $M = 2 \cdot 3 \cdot 5 \cdot 7$.
De oplossing is om nu voor c een andere waarde te kiezen. Door proberen vinden we dat $c = 12$ de oplossing $\text{ggd}(12^{2^3 \cdot 3 \cdot 5 \cdot 7} - 1, 2047) = 89$.

- c heeft kleine orde in $(\mathbb{Z}/n\mathbb{Z})^\times$, dus $c^M \equiv 1 \pmod{n}$ voor een kleine M. Dan geldt voor alle positieve veelvouden van M dat $c^M - 1 = 0$ dus $g = n$. Aangezien we de orde van een element in $(\mathbb{Z}/n\mathbb{Z})^\times$ niet makkelijk kunnen bepalen zonder de factorisatie van n te kennen is de enige oplossing een willekeurige andere c te kiezen. De meeste elementen hebben echter een grote orde.

7.7 Cryptografische bescherming tegen Pollard's algoritme

In de cryptografie is het belangrijk dat een gekozen samengesteld getal n lastig of niet praktisch te factoriseren is. Na het verschijnen van Pollards algoritme werd aangeraden om voor RSA sterke priemgetallen te kiezen. Sinds het algoritme van Lenstra en de getallenlichaamzeef zijn deze richtlijnen echter niet meer relevant. We noemen ze omdat ze inzicht geven in het verschil in effectiviteit van Pollard en Lenstra's algoritmes.

7.7.1 Sterke priemgetallen

Definitie 7.10 (Sterk priemgetal)**.** Let op dit is de cryptografiedefinitie, die is anders dan de getaltheoriedefinitie van sterk priemtgetal. Een sterk priemgetal p voldoet aan de volgende eigenschappen [RS01]

1. Als p, q sterk zijn dan is pq voldoende groot om niet (met proefdeling) te kunnen factoriseren.
2. $p - 1$ heeft één of meerdere grote priemfactoren, dus $p = a_1 q_1 + 1$ voor een a_1 en groot priemgetal q_1. Dit beschermt tegen Pollards $p - 1$ factorisatiealgoritme.
3. $q_1 - 1$ heeft één of meerdere grote priemfactoren, dus $q_1 = a_2 q_2 + 1$ voor een a_2 en groot priemgetal q_2. Dit beschermt tegen een zekere 'cycling' aanval op RSA, beschreven in

[RS01, § 9].

4. $p+1$ heeft één of meerdere grote priemfactoren, dus $p+1 = a_3 q_3 - 1$ voor een a_3 en groot priemgetal q_3. Dit beschermt tegen Williams $p+1$ factorisatiealgoritme. [Wil82]

7.7.2 Sterke priemgetallen vinden

Het algoritme van Gordon [Gor85] kan efficiënt sterke priemgetallen genereren. Het genereren van een k-bits sterk priemgetal kost ongeveer 19% meer werk dan het vinden van een willekeurig k-bits priemgetal.

7.7.3 Zijn sterke priemgetallen nuttig in 2025?

Sinds de ontdekking van elliptische kromme factorisatie door Lenstra in 1985 hebben sterke priemgetallen geen nut meer. In Pollards algoritme is de belangrijkste parameter $|(\mathbb{Z}_p)^\times| = p-1$. Wanneer $p-1$ niet B-machtglad is, dus zeker bij p sterk, mislukt het algoritme.

Door de groep $(\mathbb{Z}_p)^\times$ te vervangen door $E_{a,b}$ kan de orde $|E_{a,b}|$ van de groep (afhankelijk van a en b) variëren tussen

$$p+1-2\sqrt{p} \leq |E_{a,b}| \leq p+1+2\sqrt{p}$$

Men kan nu net zo lang waarden voor a, b blijven proberen totdat $|E_{a,b}|$ B-machtglad is voor een kleine B. Het heeft dus geen zin meer om een sterke p te maken waarvoor $p \pm 1$ een grote priemfactor heeft, aangezien $|E_{a,b}|$ praktisch een willekeurig getal in de buurt van p is. Merk op dat p zodanig construeren dat voor alle x:

$$p+1-2\sqrt{p} \leq x \leq p+1+2\sqrt{p}$$

geldt dat x niet glad is niet praktisch haalbaar is.

7.8 Eigenschappen van $p-1$

Om de kans op succes van Pollards algoritme te bepalen willen we meer weten van $p-1$ voor p een priemgetal. Voor beide keuzes van exponent $L(B) = B!$ als $L(B) = M(B)$ geldt dat $p-1 \mid L$ als $p-1$ B-machtglad is. We vatten de belangrijkste resultaten samen

Lemma 7.11 (Fundamentele eigenschappen van $p-1$)**.** *Voor elk priemgetal $p > 2$ gelden de volgende eigenschappen:*

1. *$p-1$ is even (want p is oneven).*
2. *$p-1$ is ruwweg even waarschijnlijk B-glad als andere getallen van vergelijkbare grootte.*
3. *Het aantal priemfactoren van $p-1$ is ongeveer $\ln\ln(p)$ [EP85, p. 344].*
4. *Voor elk priemgetal $q < p$ geldt: $p-1 \not\equiv q \pmod{q+1}$ Lemma 7.12*
5. *Het Lemma 7.8 geeft ruwweg de waarschijnlijkheid dat $p-1$ B-glad is. Dit is ongeveer $\frac{1}{\log_B(p-1)^{\log_B(p-1)}}$*

Lemma 7.12. *Voor elk priemgetal $q < p$ geldt: $p-1 \not\equiv q \pmod{q+1}$*

Bewijs. Veronderstel, voor een tegenspraak, dat er priemgetallen p, q bestaan met $q < p$ zodanig dat

$$p-1 \equiv q \pmod{q+1}$$

Dan geldt $p = k(q+1)$ voor zekere $k \geq 0$. Maar dit betekent dat p deelbaar is door $(q+1)$, wat in tegenspraak is met de primaliteit van p. □

We weten dat de waarschijnlijkheid dat $p-1$ B-glad is ruwweg hetzelfde is als voor algemene getallen van dezelfde grootte. Zij m een willekeurig getal ongeveer even groot als $p-1$. We kunnen de de Bruijn functie gebruiken en de benadering van Dickman, zie Stelling 3.20

$$\begin{aligned}\frac{\psi(m,B)}{m} &= \frac{\psi(m,m^{\frac{1}{u}})}{m} \quad \text{voor } u = \log_B(m) \\ &= \frac{m\rho(u)}{m} \approx u^{-u} = \frac{1}{\log_B(m)^{\log_B(m)}}.\end{aligned}$$

Rekenvoorbeeld: Zij p, q priemgetallen van 64-bits en $n = pq$, zoals in 128-bits RSA. Zij B een 16-bits getal. De kans dat $p-1$ nu B-glad is, is ongeveer

$$\frac{1}{\log_B(p-1)^{\log_B(p-1)}} \approx \frac{1}{4^4} = \frac{1}{256}$$

7.9 Tweede fase van het algoritme

Wanneer aan het einde van het algoritme nog steeds $\text{ggd}(a^L(B)-1, n) = 1$, kan ofwel de zoekgrens B verhoogd worden, ofwel de tweede fase van het algoritme uitgevoerd worden. We beschrijven de versie van Pollard [Pol74].

Men kiest een tweede zoekgrens $B_2 > B$, bijvoorbeeld $B_2 = B\ln(B)$. Zij $q_1 < q_2 < \cdots < q_t$ de priemgetallen in $[B, B_2]$.

In de eerste fase hebben we de exponenten $M(1), M(2), \ldots, M(B)$ bekeken (of $1!, 2!, \ldots B!$, afhankelijk van de implementatie).

In de tweede fase bekijken we de exponenten $q_i M(B)$ voor $1 \leq i \leq t$. Indien n een priemfactor p heeft waarvoor $p-1 = q_i u$, voor $u \mid M(B)$, dan wordt deze in de tweede fase gevonden. Immers geldt dan $p-1 \mid q_i M(B)$, dus $p \mid \text{ggd}(c^{M(B)q_i}, n)$.

In elke stap van de tweede fase bepalen we eerst op een efficiënte manier $c^{Lq_i} \pmod n$ en daarna $\text{ggd}(c^{Lq_i} - 1, n)$.

Vanuit de eerste fase weten we $c^L \pmod n$. Bereken nu eerst $c^{L\cdot q_1} \pmod n \equiv (c^L)^{q_1} \pmod n$.

We kunnen de volgende waarden op een efficiënte manier berekenen. Het verschil $q_{i+1} - q_i$ is even en veel kleiner dan q_i, ongeveer $\ln q_i$, zie deel-deelparagraaf 3.2.3.

Maak nu een tabel van $c^{Ld} \pmod n$ voor $d = 2, 4, \ldots 1000$. Indien $d := q_{i+1} - q_i \leq 1000$ kunnen we nu in $O(1)$ de waarde van c^{Ld} uit de tabel halen. Er geldt

$$c^{Lq_{i+1}} \pmod n \equiv c^{L(q_{i+1}-q_i+q_i)} \pmod n \equiv c^{Lq_i} \cdot c^{L(q_{i+1}-q_i)} \pmod n \equiv c^{Lq_i} \cdot c^{Ld} \pmod n$$

Indien $d > 1000$ moeten we c^{Ld} eerst berekenen en dan aan de tabel toevoegen. Aangezien de meeste priemverschillen < 1000 zullen zijn is de geamortiseerde complexiteit van het berekenen van c^{Lq_i} een enkele vermenigvuldiging $\pmod n$.

8 De functie M

Dit hoofdstuk analyseert de functie M die een centrale rol speelt in Pollards algoritme. We bestuderen twee aspecten:

1. Een recursieve methode om M te berekenen.
2. De asymptotische relatie tussen M en de exponentiële functie.

Lemma 8.1. *De functie M uit Definitie 7.7 is het* kgv *van de getallen* 1 *tot en met k.*

$$M(k) = \text{kgv}\{1, 2, \ldots, k\}$$

De eerste waarden van M zijn $M(1) = 1, M(2) = 2, M(3) = 6, M(4) = 12, M(5) = 60$.

Lemma 8.2 (Recursieve berekening)**.** *De functie $M(k)$ kan recursief worden berekend via:*

$$M(k+1) = \begin{cases} M(k) & \text{als } k+1 \text{ geen priemmacht is} \\ p \cdot M(k) & \text{als } k+1 = p^\alpha \text{ een priemmacht is} \end{cases}$$

Bewijs. Als $k+1$ geen priemmacht is, dan $k+1 = ab$ met $1 < a, b < k$ en a, b copriem. Hieruit volgt dat $\text{kgv}(1, \ldots, k, k+1) = \text{kgv}(1, \ldots, a, \ldots, b, \ldots k) = M(k)$. Stel dat $k+1$ wel een priemmacht p^α is. Voor $\alpha = 1$ geldt duidelijk

$$\begin{aligned} M(k+1) &= p \prod_{j=1}^{n} p_j^{\lfloor \ln k / \ln p_j \rfloor} \qquad \text{Met } p_j \text{ het } j\text{-de priem en } n \text{ het grootste getal zdd } p_n \leq k \\ &= pM(k) \end{aligned}$$

Voor $\alpha > 1$ geldt

$$\left\lfloor \frac{\ln(k)}{\ln(p)} \right\rfloor = \alpha - 1 \qquad \left\lfloor \frac{\ln(k+1)}{\ln(p)} \right\rfloor = \alpha$$

dus ook $M(k+1) = pM(k)$. □

8.1 De functie M benadert e^n

Een opmerkelijk resultaat is dat $M(n)$ asymptotisch benaderd kan worden door de exponentiële functie, zie Stelling 8.9. Deze benadering is wiskundig interessant en geeft ons inzicht in de orde van grootte van $M(n)$, hoewel dit niet essentieel voor het begrip van Pollards algoritme.

Lemma 8.3 ([Eve24b])**.** *Zij $\pi(x)$ het aantal priemgetallen $\leq x$, en zij $\theta(x) = \sum_{p \leq x} \ln(p)$, waarbij p priem is, dan*

$$\frac{\pi(x) \cdot \ln(x)}{x} = \frac{\theta(x)}{x} + \mathcal{O}\left(\frac{1}{\ln(x)}\right)$$

Voor het bewijs hebben we eerst een aantal lemma's nodig

Lemma 8.4. *Voor elke reële $x \leq 2$ geldt, (p staat voor priemgetal)*

$$\prod_{p \leq x} p \leq 4^x.$$

Zie voor een bewijsschets [Eve24a, opgave 1.4].

Lemma 8.5.

$$\theta(x) = \sum_{p \leq x} \ln(p) \leq x \ln(4) = \mathcal{O}(x) \qquad \text{voor } x \to \infty$$

Bewijs. Neem de logaritme van de vergelijking in Lemma 8.4. □

Lemma 8.6. *Zij*

$$\psi(x) = \sum_{k,p:p^k \leq x} \ln(p).$$

Dit is de Chebyshev ψ functie. Bijvoorbeeld $\psi(10) = 3\ln(2) + 2\ln(3) + \ln(5) + \ln(7)$.

Er geldt dat

$$\psi(x) = \theta(x) + \mathcal{O}(\sqrt{x}) \text{ voor } x \to \infty.$$

Bewijs.

$$\begin{aligned}\psi(x) &= \sum_{k,p:p^k \leq x} \ln(p) = \sum_{p \leq x} \ln(p) + \sum_{p^2 \leq x} \ln(p) + \sum_{p^3 \leq x} \ln(p) + \dots \\ &= \theta(x) + \theta(x^{\frac{1}{2}}) + \theta(x^{\frac{1}{3}}) + \dots\end{aligned}$$

Merk op dat $\theta(t) = 0$ voor $t < 2$ dus

$$\theta(x^{\frac{1}{k}}) = 0 \iff x^{\frac{1}{k}} < 2.$$

De logaritme nemen geeft in dit geval

$$\theta(x^{\frac{1}{k}}) = 0 \iff k > \frac{\ln(x)}{\ln(2)}.$$

Aangezien voor $k < 2$ en een geheel getal $k > \left\lfloor \frac{\ln(x)}{\ln(2)} \right\rfloor$ geldt dat $\theta(x^{1/k}) = 0$ volgt

$$\psi(x) - \theta(x) = \sum_{k=2}^{\left\lfloor \frac{\ln(x)}{\ln(2)} \right\rfloor} \theta(x^{1/k}).$$

Nu de som opsplitsen, en merk op dat $\theta(x^{\frac{1}{3}}) \geq \theta(x^{\frac{1}{m}})$ voor $m \geq 3$

$$\psi(x) - \theta(x) \leq \theta(x^{\frac{1}{2}}) + \sum_{k=3}^{\left\lfloor \frac{\ln(x)}{\ln(2)} \right\rfloor} \theta(x^{\frac{1}{3}}) = \theta(x^{\frac{1}{2}}) + \sum_{k=0}^{\left\lfloor \frac{\ln(x)}{\ln(2)} \right\rfloor} \theta(x^{\frac{1}{3}}) - \sum_{k=0}^{2} \theta(x^{\frac{1}{3}})$$

Omdat $\left\lfloor \frac{\ln(x)}{\ln(2)} \right\rfloor \leq \frac{\ln(x)}{\ln(2)}$ volgt

$$\psi(x) - \theta(x) \leq \theta(x^{\frac{1}{2}}) + \left(\frac{\ln(x)}{\ln(2)} - 3 \right) \theta(x^{\frac{1}{3}})$$

Merk op dat $\theta(x) = \mathcal{O}(x)$, dus

$$\begin{aligned}\psi(x) - \theta(x) &= \mathcal{O}(\sqrt{x} + \sqrt[3]{x} \cdot \ln x) \\ &= \mathcal{O}(\sqrt{x}) \text{ voor } x \to \infty.\end{aligned}$$

□

Lemma 8.7. *Voor elke reële $A > 0$ geldt*

$$\int_2^x \frac{1}{(\ln t)^A} \, \mathrm{d}t = \mathcal{O}\left(\frac{x}{(\ln x)^A} \right) \text{ voor } x \to \infty,$$

waarbij de constante die geïmpliceerd wordt door de $\mathcal{O}$-notatie afhankelijk is van A. Dus er bestaan $C_A > 0$, $x_{0A} > 0$, zodanig dat

$$\left| \int_2^x \frac{1}{(\ln t)^A} \, \mathrm{d}t \right| \leq C \cdot \frac{x}{(\ln x)^A} \text{ voor } x \geq x_0.$$

Zie voor een bewijsschets [Eve24a, opgave 1.3].

Lemma 8.8 (Partiële sommatie (Abels somformule))**.** *Zij $(a_n)_{n=0}^{\infty}$ een rij reële getallen en zij $A(t) = \sum_{n=0}^{t} a_n$. Zij $x \in \mathbb{R}$, $x > 0$, en ϕ een C^1 functie op $[0, x]$, dan*

$$\sum_{0 \leq n \leq x} a_n \phi(n) = A(x)\phi(x) - \int_0^x A(u)\phi'(u) \, \mathrm{d}u$$

Bewijs van Lemma 8.3. We nemen aan dat de priemgetalstelling reeds bewezen is.

$$\pi(x) = \sum_{p \leq x} 1 = \sum_{p \leq x} \ln(p) \cdot \frac{1}{\ln(p)}.$$

Nu gebruiken we partiële sommatie met $a_n = \mathbb{1}_P(n) \cdot \ln(n)$ waarbij $\mathbb{1}_P(n)$ de priemindicatorfunctie is, daaruit volgt $A(t) = \theta(t)$. Ook kiezen we $\phi(t) = \frac{1}{\ln(t)}$

$$\begin{aligned} &= \theta(x)\frac{1}{\ln(x)} - \int_0^x \theta(t) \cdot \left(\frac{1}{\ln(t)}\right)' \mathrm{d}t = \theta(x)\frac{1}{\ln(x)} - \int_2^x \theta(t) \cdot \left(\frac{1}{\ln(t)}\right)' \mathrm{d}t \\ &= \frac{\theta(x)}{\ln(x)} + \int_2^x \frac{\theta(t)}{t \ln^2(t)} \, \mathrm{d}t \end{aligned}$$

We gaan nu de integraal afschatten. De integraal is ≥ 0. Uit Lemma 8.5 volgt dat er een constante $C > 0$ bestaat waarvoor $\theta(t) \leq Ct$ voor alle $t \geq 2$, dus $\frac{\theta(t)}{t} \leq C$

$$0 \leq \int_2^x \frac{\theta(t)}{t \ln^2(t)} \, \mathrm{d}t \leq C \int_2^x \frac{1}{\ln^2(t)} \, \mathrm{d}t$$

Met Lemma 8.7 volgt

$$C \int_2^x \frac{1}{\ln^2(t)} \, \mathrm{d}t = \mathcal{O}\left(\frac{x}{\ln^2(x)}\right) \text{ voor } x \to \infty.$$

Nu kunnen we het bewijs afronden

$$\begin{aligned} \pi(x) \cdot \frac{\ln(x)}{x} &= \frac{\theta(x)}{\ln(x)} \cdot \frac{\ln(x)}{x} + \mathcal{O}\left(\frac{\ln(x)}{x} \cdot \frac{x}{\ln^2(x)}\right) \\ &= \frac{\theta(x)}{x} + \mathcal{O}\left(\frac{1}{\ln(x)}\right) \quad \text{voor } x \to \infty. \end{aligned}$$

□

Stelling 8.9. *De functie $M(n)$ benadert e^n voor $n \to \infty$. Om precies te zijn*

$$\lim_{n \to \infty} \frac{\ln M(n)}{n} = 1.$$

Bewijs. De definitie van $M(n)$ is (merk op dat p de priemgetallen $\leq n$ zijn)

$$M(n) = \prod_{p \leq n} p^{\left\lfloor \frac{\ln n}{\ln p} \right\rfloor}.$$

Door de logaritme van beide kanten te nemen volgt

$$\ln(M(n)) = \sum_{p \le n} \left\lfloor \frac{\ln n}{\ln p} \right\rfloor \cdot \ln p = \sum_{p \le \sqrt{n}} \left\lfloor \frac{\ln n}{\ln p} \right\rfloor \cdot \ln p + \sum_{\sqrt{n} < p \le n} \left\lfloor \frac{\ln n}{\ln p} \right\rfloor \cdot \ln p.$$

Voor de rechtersommatie: $\sqrt{n} < p \le n \implies \frac{1}{2}\ln(n) < \ln(p) \le \ln(n) \implies \frac{\ln(n)}{\ln(p)} \in [1,2) \implies \left\lfloor \frac{\ln n}{\ln p} \right\rfloor = 1$, dus

$$\ln(M(n)) = \sum_{p \le \sqrt{n}} \left\lfloor \frac{\ln n}{\ln p} \right\rfloor \cdot \ln p + \sum_{\sqrt{n} < p \le n} \ln(p)$$

Voor de linkersommatie geldt $0 < \left\lfloor \frac{\ln n}{\ln p} \right\rfloor \cdot \ln p \le \ln n$, dus

$$\begin{aligned} \ln(M(n)) &= \mathcal{O}\left(\sqrt{n}\ln(n)\right) + \sum_{\sqrt{n} < p \le n} \ln(p) \\ &= \mathcal{O}\left(\sqrt{n}\ln(n)\right) + \sum_{p \le n} \ln(p) - \sum_{p < \sqrt{n}} \ln(p) \\ &= \mathcal{O}\left(\sqrt{n}\ln(n)\right) + \sum_{p \le n} \ln(p) = \theta(n) + \mathcal{O}\left(\sqrt{n}\ln(n)\right) \quad \theta \text{ is de eerste functie van Chebyshev.} \end{aligned}$$

Nu delen we beide kanten door n

$$\frac{\ln(M(n))}{n} = \frac{\theta(n)}{n} + \frac{\mathcal{O}\left(\sqrt{n}\ln(n)\right)}{n}.$$

De limiet nemen voor $n \to \infty$ geeft, met Lemma 8.3

$$\lim_{n\to\infty} \frac{\ln(M(n))}{n} = \lim_{n\to\infty} \frac{\theta(n)}{n} + \lim_{n\to\infty} \frac{\mathcal{O}\left(\sqrt{n}\ln(n)\right)}{n} = 1.$$

□

9 De verdeling van kwadraatvrije getallen

In dit hoofdstuk bepalen we de verdeling van kwadraatvrije getallen. Dit geeft inzicht in de kans op succes van Pollards algoritme indien we voor de exponent $L(B)$ het product van priemgetallen kleiner of gelijk aan B kiezen. Hoewel dit in de praktijk geen goede keuze voor L is, is het voor deze keuze aanzienlijk makkelijker om de kans op succesvol factoriseren te vinden.

Definitie 9.1 (Möbiusfunctie [AS65, p. 826])**.** De Möbiusfunctie $\mu : \mathbb{N} \to \{-1, 0, 1\}$ is gedefinieerd als:

$$\mu(n) = \begin{cases} 1 & \text{als } n = 1 \\ (-1)^k & \text{als } n = p_1 \cdots p_k \text{ met } p_i \text{ verschillende priemgetallen} \\ 0 & \text{als } n \text{ deelbaar is door een kwadraat } > 1 \end{cases}$$

Opmerking 9.2. De Möbiusfunctie is multiplicatief: voor coprieme getallen a en b geldt $\mu(ab) = \mu(a)\mu(b)$.

Het volgende lemma komt van pas om het aantal kwadraatvrije getallen tussen 1 en x uit te drukken met de Möbiusfunctie.

Lemma 9.3.

$$\sum_{d^2 \mid n} \mu(d) = \begin{cases} 0 & n \textit{ is niet kwadraatvrij} \\ 1 & n \textit{ is kwadraatvrij.} \end{cases}$$

Bewijs. Indien n kwadraatvrij, dan geldt $\sum_{d^2|n} \mu(d) = \mu(1) = 1$.

Indien n niet kwadraatvrij, dan bevat n tenminste één kwadratische factor. Zij k het aantal verschillende priemdelers p van n waarvoor $p^2 \mid n$.

Voorbeeld voor enkele waarden van k:

$$k = 1 \quad p^2 \mid n: \quad \sum_{d^2 \mid n} \mu(d) = \mu(1) + \mu(p) = 1 - 1 = 0$$

$$k = 2 \quad p^2, q^2 \mid n: \quad \sum_{d^2 \mid n} \mu(d) = \mu(1) + \underbrace{\mu(p) + \mu(q)}_{-2} + \mu(pq) = 1 - 2 + 1 = 0$$

$$k = 3 \quad p^2, q^2, r^2 \mid n: \quad \sum_{d^2 \mid n} \mu(d) = \mu(1) + \underbrace{\mu(p) + \mu(q) + \mu(r)}_{-3} + \underbrace{\mu(pq) + \mu(pr) + \mu(qr)}_{3} + \mu(pqr) = 1 - 3 + 3 - 1 = 0$$

Inderdaad lijkt voor alle k te gelden dat $\sum_{d^2|n} \mu(d) = 0$. Het bewijs volgt uit de volgende combinatorische identiteit:

Lemma 9.4 (Alternerende som van binomiaalcoëfficiënten)**.** *Voor alle $n \in \mathbb{N}_{\geq 1}$ geldt:*

$$\sum_{j=0}^{n} (-1)^j \binom{n}{j} = 0.$$

Toepassing van deze identiteit maakt het bewijs. Merk op dat we geen rekening hoeven te houden met hogere exponenten: als $p^4 \mid n$, dan volgt $\mu(p^2) = 0$ uit de definitie van μ. □

De Möbiusfunctie heeft een directe relatie met de Riemann-zèta-functie.

We geven eerst een hulplemma waarin we het Eulerproduct van de zèta-functie bepalen.

Lemma 9.5 (Eulerproduct van ζ)**.** *Zij $\mathbb{P}$ de verzameling priemgetallen, dan geldt voor* $\mathrm{Re}(s) > 1$

$$\zeta(s) = \prod_{p\in\mathbb{P}} \frac{1}{1-p^{-s}}.$$

Bewijs. Het bewijs is slechts een schets, men dient convergentie van de reeksen nog aan te tonen.

$$\begin{aligned}
\zeta(s) &= 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} + \frac{1}{6^s} + \frac{1}{7^s} + \frac{1}{8^s} + \cdots \\
\frac{1}{2^s}\zeta(s) &= \quad\frac{1}{2^s} + \quad\frac{1}{4^s} + \quad\frac{1}{6^s} + \quad\frac{1}{8^s} + \cdots
\end{aligned}$$

De bovenste van de onderste vergelijking aftrekken geeft

$$\begin{aligned}
\left(1-\frac{1}{2^s}\right)\zeta(s) &= 1 + \frac{1}{3^s} + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{9^s} + \frac{1}{11^s} + \frac{1}{13^s} + \frac{1}{15^s} + \cdots \\
\frac{1}{3^s}\left(1-\frac{1}{2^s}\right)\zeta(s) &= \quad\frac{1}{3^s} + \quad\frac{1}{9^s} + \quad\frac{1}{15^s} + \frac{1}{21^s} + \cdots \qquad \text{delen door } 3^s
\end{aligned}$$

Weer de bovenste van de onderste vergelijking aftrekken geeft

$$\left(1-\frac{1}{3^s}\right)\left(1-\frac{1}{2^s}\right)\zeta(s) = 1 + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{11^s} + \frac{1}{13^s} + \frac{1}{17^s} + \frac{1}{19^s} + \frac{1}{23^s} + \cdots$$

We kunnen deze techniek voortzetten voor meer priemgetallen

$$(\cdots)\left(1-\frac{1}{11^s}\right)\left(1-\frac{1}{7^s}\right)\left(1-\frac{1}{5^s}\right)\left(1-\frac{1}{3^s}\right)\left(1-\frac{1}{2^s}\right)\zeta(s) = 1.$$

Delen door het oneindige product geeft nu het Eulerproduct

$$\zeta(s) = \frac{1}{\left(1-\frac{1}{2^s}\right)\left(1-\frac{1}{3^s}\right)\left(1-\frac{1}{5^s}\right)\left(1-\frac{1}{7^s}\right)\left(1-\frac{1}{11^s}\right)(\cdots)}.$$

□

Lemma 9.6. *Er geldt voor* $\mathrm{Re}(s) > 1$

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}.$$

Bewijs. We geven een bewijsschets. Merk op dat μ een multiplicatieve functie is, en

$$\mu(p^k) = \begin{cases} 1 & \text{voor } k = 1 \\ 0 & \text{voor } k \geq 2. \end{cases}$$

We kunnen de vergelijking dus als Dirichletreeks schrijven

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \prod_{p\in\mathbb{P}} \left(1 + \mu(p)p^{-s} + \mu(p^2)p^{-2s} + \cdots\right) = \prod_{p\in\mathbb{P}} \left(1-p^{-s}\right)$$

Uit Lemma 9.5 volgt ook

$$\frac{1}{\zeta(s)} = \prod_{p\in\mathbb{P}} \left(1 - p^{-s}\right).$$

Dus de LHS is gelijk aan de RHS. □

Stelling 9.7 (Dichtheid van kwadraatvrije getallen)**.** *Zij $Q(x)$ het aantal kwadraatvrije getallen $\leq x$. Dan geldt:*

$$Q(x) = \frac{6x}{\pi^2} + O(\sqrt{x}).$$

Bewijs. We herschrijven $Q(x)$ eerst met Lemma 9.3.

$$\begin{aligned}
Q(x) &= \sum_{n\leq x} \sum_{d^2|n} \mu(d) && \text{sommatie van Lemma 9.3 over } 1\ldots x \\
&= \sum_{d\leq x} \mu(d) \sum_{\substack{n\leq x \\ d^2|n}} 1 && \text{verwissel sommatievolgorde} \\
&= \sum_{d\leq x} \mu(d) \left\lfloor \frac{x}{d^2} \right\rfloor && \text{aantal delers tellen} \\
&= \sum_{d\leq \sqrt{x}} \mu(d) \left\lfloor \frac{x}{d^2} \right\rfloor && \text{want } \left\lfloor \frac{x}{d^2} \right\rfloor = 0 \text{ voor } d > \sqrt{x}
\end{aligned}$$

Nu willen we de floorfunctie vervangen door een afschatting met grote-$\mathcal{O}$ notatie:

$$\begin{aligned}
&= \sum_{d\leq \sqrt{x}} \mu(d) \frac{x}{d^2} + \mathcal{O}\left(\sum_{d\leq\sqrt{x}} 1\right) && \text{want } b-1 < \lfloor b \rfloor \leq b \quad \text{dus } \sum \lfloor b \rfloor = \sum b + \mathcal{O}\left(\sum 1\right) \\
&= x \sum_{d\leq\sqrt{x}} \frac{\mu(d)}{d^2} + \mathcal{O}\left(\sqrt{x}\right) && x \text{ buiten som halen en } \mathcal{O}\text{-term versimpelen} \\
&= x \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} - x \sum_{d>\sqrt{x}} \frac{\mu(d)}{d^2} + \mathcal{O}\left(\sqrt{x}\right) && \text{beide sommen zijn convergent want } |\mu(d)| \leq 1 \\
&= x \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} + \mathcal{O}\left(x \sum_{d>\sqrt{x}} \frac{1}{d^2} + \sqrt{x}\right) && \mu(d) \text{ afschatten op } 1
\end{aligned}$$

We kunnen $\sum_{d>\sqrt{x}} \frac{1}{d^2}$ afschatten met een integraal: $\sum_{d>\sqrt{x}} \frac{1}{d^2} < \int_{\lfloor\sqrt{x}\rfloor-1}^{\infty} \frac{1}{t^2}\,\mathrm{d}t = \left[-\frac{1}{t}\right]_{\lfloor\sqrt{x}\rfloor-1}^{\infty} \leq \frac{2}{\sqrt{x}}$ voor $x \geq 4$

$$= x \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} + \mathcal{O}\left(\sqrt{x} + \sqrt{x}\right) \qquad \text{integraalafschatting gebruikt}$$

Nu kan Lemma 9.6 gebruikt worden

$$= \frac{x}{\zeta(2)} + \mathcal{O}\left(\sqrt{x}\right) = \frac{6x}{\pi^2} + \mathcal{O}(\sqrt{x}) \quad \text{Zie [Cha99] voor 14 manieren om } \zeta(2) \text{ te evalueren.}$$

□

Gevolg 9.8. *De dichtheid van kwadraatvrije getallen is $\frac{6}{\pi^2} \approx 0.608$.*

10 Algebraïsche vlakkrommen

In 1987 ontdekte Lenstra [Len87] een grote verbetering van het $p-1$ algoritme van Pollard door elliptische krommen over eindige lichamen te gebruiken.

De verzameling van elliptische krommen is een deelverzameling van de verzameling van kubische krommen, die weer een deelverzameling is van de vlakkrommen. We beschouwen in dit hoofdstuk eerst enkele algemene eigenschappen van vlakkrommen, en gaan in het volgende hoofdstuk verder in op elliptische krommen.

Definitie 10.1. Een algebraïsche (affiene) vlakkromme van graad d over een lichaam k is een kromme γ van de vorm [Tao11]

$$\gamma = \{(x, y) \in k^2 : P(x, y) = 0\}$$

Waarin P een polynoomfunctie is met totale graad d. We kunnen de kromme ook beschouwen als de nulverzameling van de polynoomfunctie

$$g : k^2 \to k, (x, y) \mapsto \sum_{i+j \leq d} a_{ij} x^i y^j \qquad \text{waarbij } a_{ij} \in k.$$

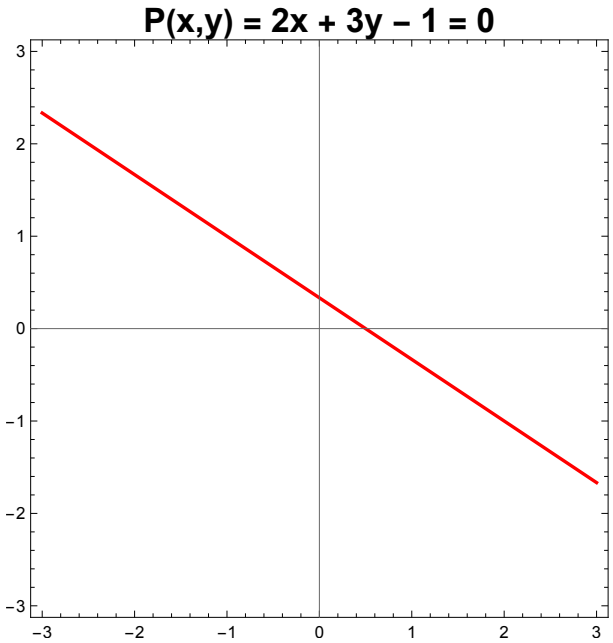
We bekijken vlakkrommen van graad $d \in \{1, 2, 3\}$.

10.1 Graad 1 - lineaire krommen

Voor $d = 1$ is de kromme een lijn

$$\{(x, y) \in k^2 : ax + by = c\}$$

Zie Figuur 7



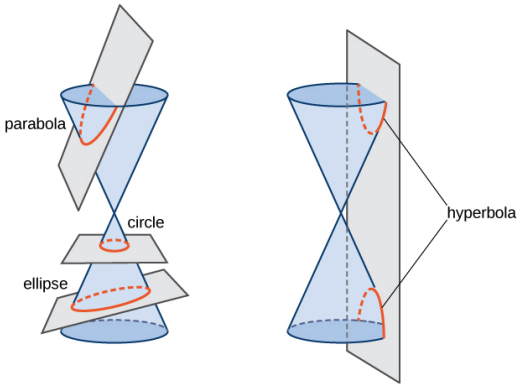
Figuur 7: Vlakkromme van graad 1 over $\mathbb{R}$

10.2 Graad 2 - kwadratische krommen

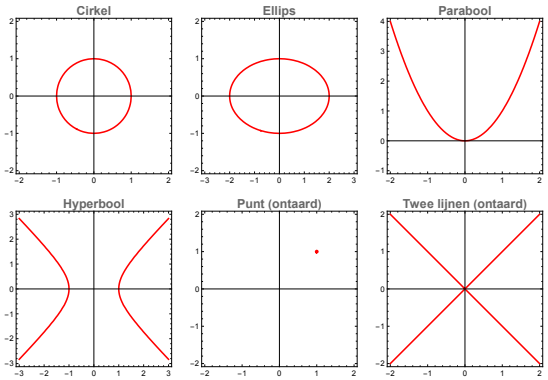
Voor graad $d = 2$ hebben de we krommen van de vorm

$$\{(x, y) \in k^2 : ax^2 + bxy + cy^2 + dx + ey + f = 0\}.$$

Dit zijn de kegelsneden (en de ontaarde gevallen: twee parallele lijnen, twee snijdende lijnen, en een punt). Zie Figuur 8



(a) Kegelsneden, bijgesneden afbeelding van [SH16]



(b) Kegelsneden in het vlak

Figuur 8: Links: kegelsneden als snijding van een kegel met een vlak in k^3. Rechts: Kegelsneden als vlakkrommen

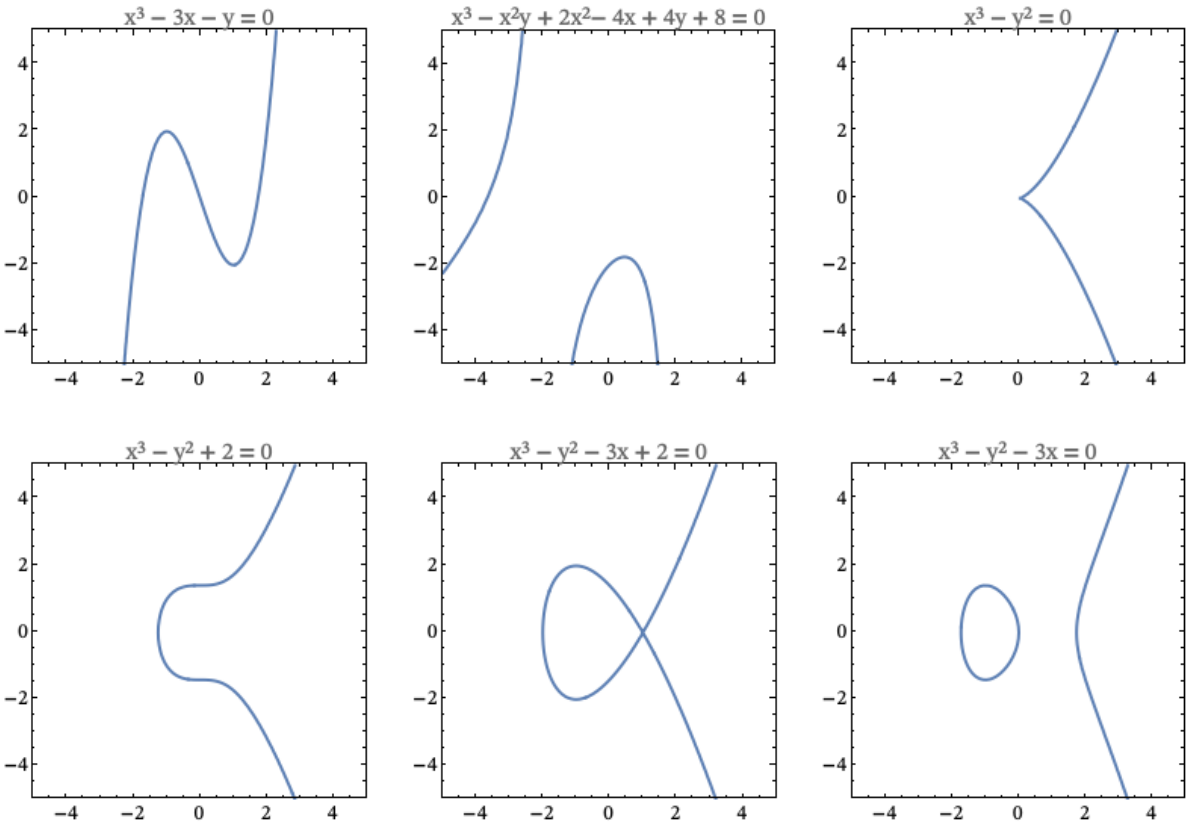
10.3 Graad 3 - kubische vlakkrommen

Voor graad $d = 3$ hebben de kubische vlakkrommen, deze zijn van de vorm

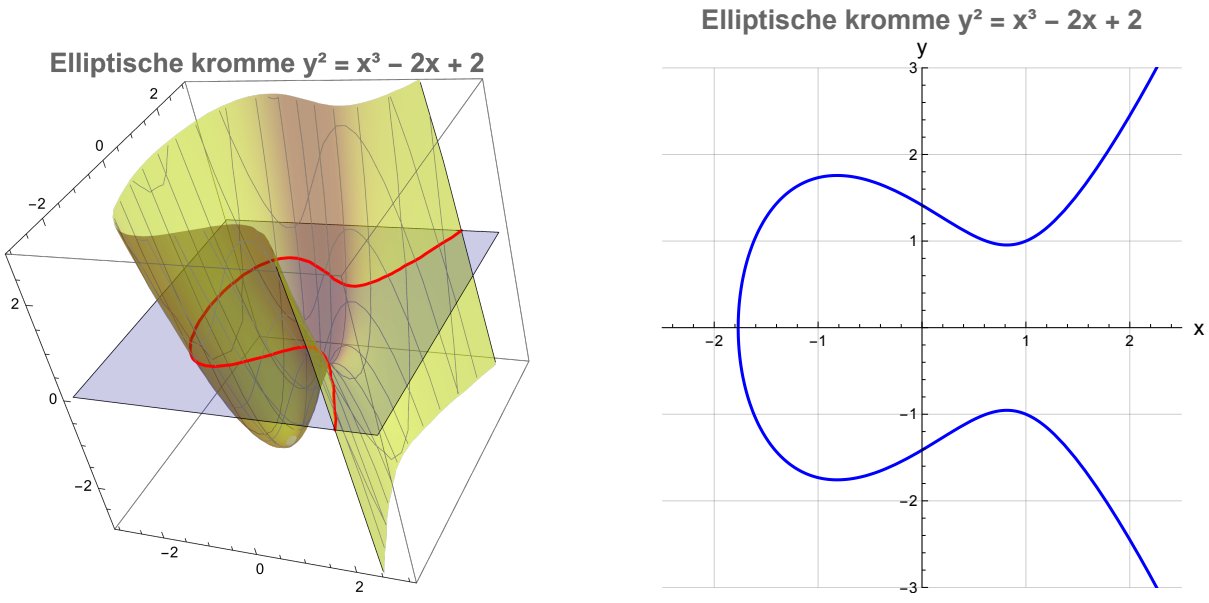
$$\{(x, y) \in k^2 : ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + iy + j = 0\}.$$

De kromme heet rationaal indien de coëfficienten $a, b, \ldots, j$ rationaal zijn.

Deze verzameling bevat onder andere de elliptische krommen, zie volgende hoofdstuk, maar ook de vereniging van een kegelsnede en een lijn, en de vereniging van drie lijnen. Zie bijvoorbeeld [Wil] voor een overzicht van 45 soorten kubische krommen. We tonen er hier een paar in Figuur 9, en een elliptische kromme in Figuur 10.



Figuur 9: Grafieken van enkele kubische krommen op $[-5,5] \times [-5,5]$



Figuur 10: Links: Een elliptische kromme als nulverzameling van het oppervlak $z = y^2 - x^3 + 2x - 2$. Rechts: dezelfde elliptische kromme als vlakkromme $y^2 = x^3 - 2x + 2$.

10.4 Homogene coördinaten

De algebraïsche vlakkrommen kunnen worden uitgebreid naar het projectieve vlak.

Definitie 10.2. [Ras11] Een functie $f : R^n \to \mathbb{R}$ heet homogeen van graad r als

$$f(\lambda x_1, \lambda x_2, \ldots, \lambda x_n) = \lambda^r f(x_1, x_2, \ldots, x_n).$$

Het doel van homogeniteit is invariantie onder schaling. Indien een homogene functie een kromme $f(x_1, x_2, \ldots, x_n) = 0$ definiëert en een punt $p = (x_1, x_2, \ldots x_n)$ voldoet aan $f(p) = 0$ (dus p ligt op de kromme), dan voldoen ook alle scalaire veelvouden van p (dus alle punten op de lijn door de oorsprong en p) aan f.

Om een kubische kromme te homogeniseren, voegen we een derde variabele z toe, en vermenigvuldigen we elke term net zo vaak met z totdat deze graad 3 heeft.

De eerste vergelijking is niet homogeen, en de tweede is homogeen gemaakt.

$$\begin{aligned} ax^3 + bx^2y + cxy^2 + dy^3 + \ ex^2 + \ fxy + \ gy^2 + \ \ hx + \ \ iy + \ \ j = 0 \\ ax^3 + bx^2y + cxy^2 + dy^3 + ex^2z + fxyz + gy^2z + hxz^2 + iyz^2 + jz^3 = 0. \end{aligned}$$

Merk op dat de nulpolynoom ook aan deze definitie voldoet en dus homogeen is.

Voorbeeld: [ST15, Appendix A] Beschouw de vergelijking

$$x^3 + y^3 = 1 \qquad \text{voor } x, y \in \mathbb{Q}. \tag{10.1}$$

Zij nu $x = \frac{a}{c}, y = \frac{b}{d}$ met a, c en b, d copriem, en c, d positief, een oplossing. Dan geldt

$$\left(\frac{a}{c}\right)^3 + \left(\frac{b}{d}\right)^3 = 1 \implies \frac{a^3d^3 + b^3c^3}{c^3d^3} = 1 \implies a^3d^3 + b^3c^3 = c^3d^3.$$

Uit $a^3d^3 = c^3(d^3 - b^3)$ volgt $c^3 \mid a^3d^3$, en samen met $\text{ggd}(a, c) = 1$ geldt $c^3 \mid d^3$. Evenzo geldt $d \mid c$, dus $c = d$ De oplossing is dus van de vorm $x = \frac{a}{c}, y = \frac{b}{c}$. Vermenigvuldigen met c geeft nu een oplossing in gehele getallen.

$$a^3 + b^3 = c^3.$$

Dit kunnen we zien als een oplossing van de homogene vergelijking

$$X^3 + Y^3 = Z^3. \tag{10.2}$$

Andersom kan elke homogene oplossing $(X = a, Y = b, Z = c)$, waarvoor $c \neq 0$ van (10.2) omgezet worden in een rationale oplossing van (10.1) door $f : (a, b, c) \mapsto (\frac{a}{c}, \frac{b}{c}, 1)$. Dit heet een projectieve correspondentie tussen oplossingen.
Er zijn drie dingen om op te letten

1. Merk op dat een homogene oplossing (ta, tb, tc) dezelfde rationale oplossing geeft als (a, b, c). We zien beide als dezelfde homogene oplossing.
2. De oplossing $(0, 0, 0)$ is altijd een oplossing van (10.2). We laten deze triviale oplossing buiten beschouwing.
3. De (10.2) heeft de oplossingen $(1, -1, 0)$ en $(-1, 1, 0)$, waarvoor geen afbeelding naar (10.1) bestaat. We noemen dit de *punten op oneindig*.

Door de projectieve correspondentie te gebruiken, kunnen we om punten te vinden zowel naar rationale oplossingen van $x^3 + y^3 = 1$ kijken, als naar geheeltallige oplossingen van $X^3 + Y^3 = Z^3$.

Definitie 10.3 (Het projectieve vlak)**.** Het projectieve vlak $\mathbb{P}^2(K)$ over een lichaam K bestaat uit de drietallen $[a, b, c]$ met a, b, c niet alledrie 0, modulo de equivalentierelatie $\sim$, waarbij

$$[a, b, c] \sim [a', b', c'] \quad \iff \quad a = ta', b = tb', c = tc' \quad \text{voor een } t \neq 0.$$

We noteren een punt in het projectieve vlak als $(a : b : c)$.

10.5 Algebraïsche krommen door punten

Stelling 10.4 (Stelling van Bézout)**.** *Een kromme γ van graad d en een kromme γ' van graad d' snijden elkaar in hoogstens dd' punten.*

Wanneer we een verzameling punten hebben kunnen we een algebraïsche kromme construeren die door deze punten gaat. Wanneer we twee verschillende punten P_1, P_2 hebben kunnen we altijd een lijn vinden die door deze punten gaat.

Definitie 10.5. Voor een algebraïsche vlakkrome ontstaat een lineaire beperking door het invullen van een punt $P = (x_0, y_0)$ in de vergelijking van het polynoom van de kromme.

Bijvoorbeeld $P = (2, 7)$ invullen in de kromme $F(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0$ geeft

$$4a + 14b + 49c + 2d + 7e + f = 0$$

Een punt waar de kromme door heen gaat correspondeert dus met een lineaire beperking.

De relatie tussen het aantal lineaire beperkingen en het aantal krommen wat door deze punten gaat (specifieker: de dimensie van de oplossingsruimte) wordt gegeven door de dimensiestelling uit de lineaire algebra. Indien er m lineaire beperkingen zijn, en n onbekenden, dan heeft het system minstens één oplossing als $m < n$.

Voorbeeld: Zij $P_1 = (1, 2)$ en $P_2 = (3, 4)$, dit correspondeert met het stelsel (rechts in matrixvorm)

$$\begin{cases} a + 2b + c = 0 \\ 3a + 4b + c = 0 \end{cases} \quad \left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 3 & 4 & 1 & 0 \end{array}\right)$$

Oplossen geeft

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} t \\ -t \\ t \end{pmatrix} = t \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad c \in \mathbb{R} \setminus \{0\}$$

Dit is de lijn $x - y + 1 = 0$.

Voor 5 punten in algemene positie (d.w.z. er zijn geen drie collineaire punten) gaat dit als volgt. Zij de volgende 5 punten gegeven: $P_1 = (2, 0)$, $P_2 = (4, -3)$, $P_3 = (0, -3)$, $P_4 = (\frac{18}{5}, -\frac{6}{5})$, $P_5 = (\frac{2}{5}, -\frac{6}{5})$.
Deze punten invullen in de formule van de kwadratische vlakkromme

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

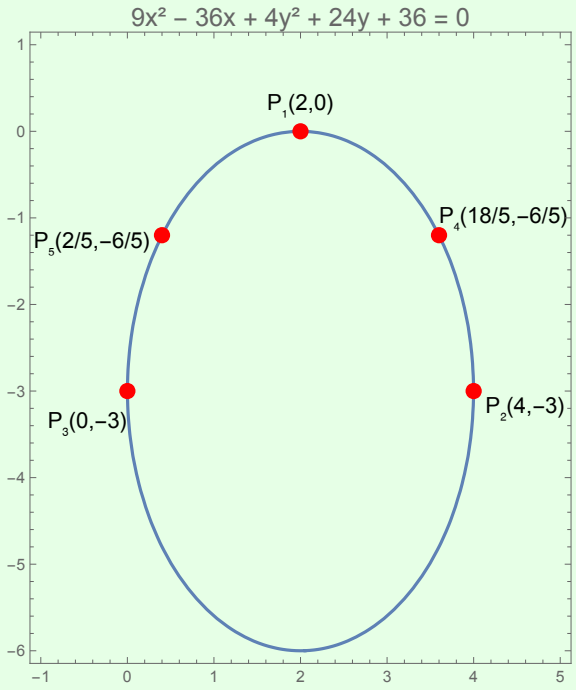
geeft de matrixvergelijking

$$\begin{pmatrix} 4 & 0 & 0 & 2 & 0 & 1 \\ 16 & -12 & 9 & 4 & -3 & 1 \\ 0 & 0 & 9 & 0 & -3 & 1 \\ 324 & -108 & 36 & 90 & -30 & 25 \\ 4 & -12 & 36 & 10 & -30 & 25 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Oplossen geeft de coëfficienten van de vlakkromme

$$9x^2 + 0xy + 4y^2 - 36x + 24y + 36 = 0$$

Dit is de vergelijking van een ellips, zie Figuur 11.



Figuur 11: De 5 gegeven punten definiëren een ellips, omschrijven naar standaardvorm geeft $\frac{(x-2)^2}{4} + \frac{(y+3)^2}{9} = 1$

We bekijken nu kubische krommen door acht of negen punten.

Stelling 10.6 (Cayley-Bacharach)**.** *Zij* $\gamma_0 = \{P_0(x,y) = 0\}$ *en* $\gamma_1 = \{P_1(x,y) = 0\}$ *twee kubische krommen die elkaar (over een algebraïsch gesloten lichaam* k*) snijden in exact* 9 *verschillende punten* $A_1, \ldots, A_9 \in k^2$*. Zij* P *een kubische polynoom waarvoor* $P(A_i) = 0$ *op* 8 *van deze punten, z.v.v.a op* $1 \leq i \leq 8$*. Dan is* P *een lineaire combinatie van* P_0 *en* P_1 *en* $P(A_9) = 0$*. Zie [Tao11] voor een bewijs.*

11 Elliptische krommen

We bekijken nu de elliptische krommen. Dit is een deelverzameling van niet-singuliere, gladde kubische krommen die in Weierstrassvorm geschreven kunnen worden. Deze worden gebruikt in het algoritme van Lenstra.

11.1 De Weierstrassvormen

Er zijn meerdere definities van de Weierstrassvorm. De gereduceerde Weierstrassvorm:

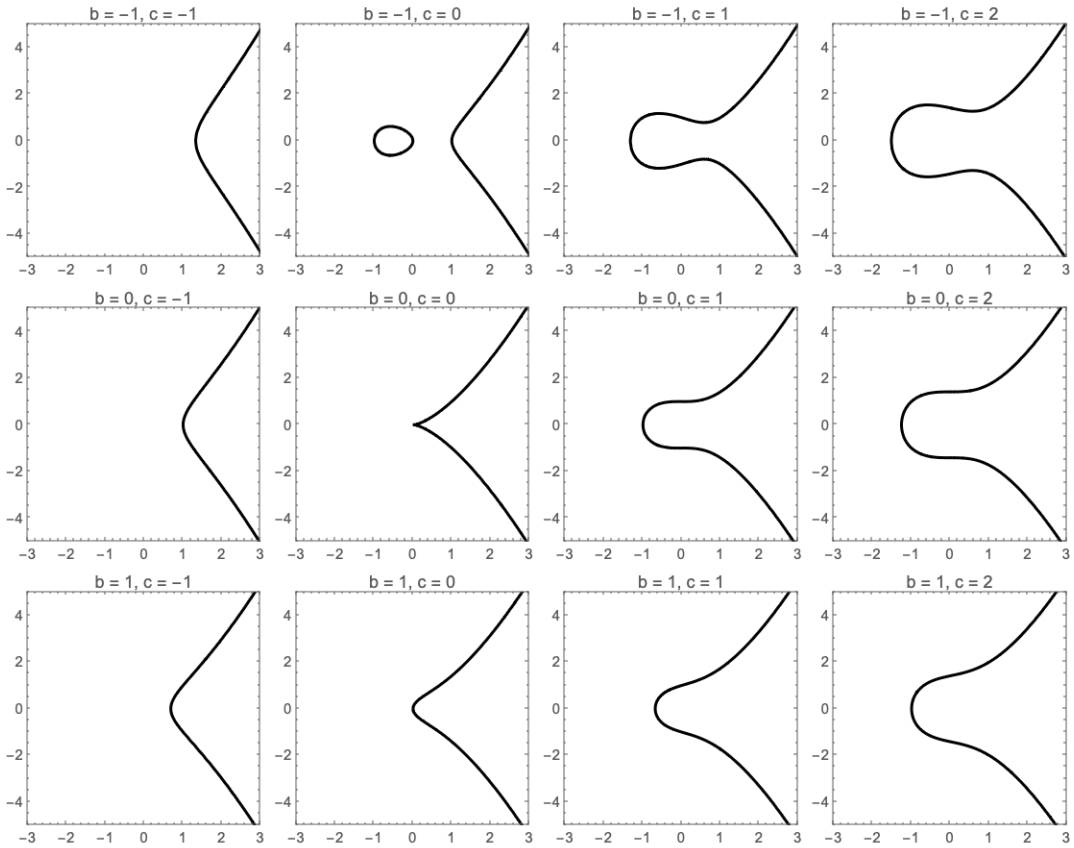
$$y^2 = f(x) = x^3 + bx + c.$$

We noteren deze kromme als $E_{a,b}$. De Weierstrass-normaalvorm is

$$y^2 = f(x) = x^3 + ax^2 + bx + c.$$

Indien de (complexe) wortels van f verschillend zijn, heet deze kubische kromme in Weierstrass-normaalvorm een *elliptische kromme* [ST15, p. 20]. In cryptografie wordt meestal de gereduceerde Weierstrassvorm gebruikt.

Zie Figuur 12 voor verschillende elliptische krommen.



Figuur 12: De kromme $y^2 = x^3 + bx + c$ voor verschillende waarden van b en c. Merk op dat de kromme voor $b = c = 0$ singulier is, dus geen elliptische kromme is. Voor de andere waarden is het wel een elliptische kromme.

Lemma 11.1. *Een kromme in gereduceerde Weierstrassvorm*

$$F(x, y) = y^2 - x^3 - bx - c$$

is singulier dan en slechts dan als $4b^3 + 27c = 0$

Bewijs. F heeft een singulier punt P als beide partiële afgeleiden in P nul zijn. De partiële afgeleiden zijn:

$$\frac{\partial F}{\partial x} = -3x^2 - b \qquad \frac{\partial F}{\partial y} = 2y$$

In een singulier punt geldt dus

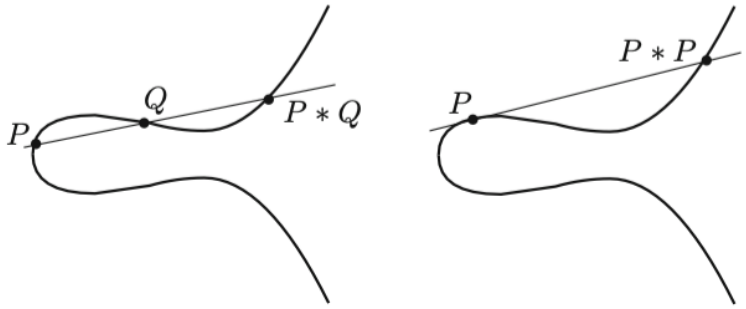
$$\begin{aligned} 3x^2 + b &= 0 \\ y &= 0 \end{aligned}$$

Deze invullen in de vergelijking van F geeft

$$\begin{aligned} 0 &= -x^3 - bx - c \\ &= -x^3 - (-3x^3) - c \quad \text{(vanwege } b = -3x^2\text{)} \\ &= 2x^3 - c. \end{aligned}$$

Dus $x^3 = \frac{c}{2}$ en $x^2 = -\frac{b}{3}$. Dit geeft $x^6 = \frac{c^2}{4} = -\frac{b^3}{27}$, dus $4b^3 + 27c^2 = 0$. □

Wanneer we twee rationale punten P en Q, $P \neq Q$ op de kromme gevonden hebben, kunnen we als volgt een derde punt vinden. Teken de lijn l door P en Q, l snijdt de kromme in een rationaal derde punt $P * Q$. Dit snijpunt is weer rationaal, zie Lemma 11.7 voor het bewijs. Het tweede snijpunt van de raaklijn in een rationaal punt P is ook een rationaal punt, zie Figuur 13.



Figuur 13: De operatie $*$ [ST15, p. 9]

Lemma 11.2. *Een kromme in Weierstrass-normaalvorm met rationale coëfficienten a, b, c van $f(x)$ heeft één of drie wortels in $\mathbb{R}$.*

Bewijs. We gebruiken het feit dat f continu is en de tussenwaardestelling. Er bestaat een M waarvoor $f(-M) < 0$ en $f(M) > 0$, want $\lim_{x\to\infty} f(x) = +\infty$ en $\lim_{x\to-\infty} f(x) = -\infty$. Nu passen we de tussenwaardestelling toe en zien dat er een $c \in \mathbb{R}$ bestaat waarvoor $f(c) = 0$.

Uit de hoofdstelling van de algebra volgt dat als f een complexe wortel z heeft, $\bar{z}$ ook een wortel is, dus twee reële wortels is niet mogelijk. □

Lemma 11.3. *Het polynoom $f(x) = x^3 + ax^2 + bx + c$ met rationale coëfficienten heeft drie verschillende wortels dan en slechts dan als $F(x, y) = y^2 - x^3 - ax^2 - bx - c$ niet-singulier is.*

Bewijs. Stel dat F een singulier punt (x_0, y_0) heeft. Dan is $\frac{\partial F}{\partial y}(y_0) = 2y_0 = 0$, dus $y_0 = 0$. Ook $f(x_0) = y_0^2 = 0$, en $f'(x_0) = \frac{\partial F}{\partial x} = 0$.

Aangezien nu $f(x) = f'(x_0) = 0$, en f is een polynoom, kunnen we deze schrijven als

$$\begin{aligned} f(x) &= (x - x_0)g(x) \\ f'(x) &= (x - x_0)h(x) = g(x) + (x - x_0)g'(x) \qquad \text{RHS volgt uit productregel op } f(x) \end{aligned}$$

waarbij g en h polynomen zijn. Invullen van $x = x_0$ geeft nu $g(x_0) = 0$, dus $g(x) = (x - x_0)k(x)$, met k een polynoom. Hieruit volgt dat $f(x) = (x - x_0)^2 k(x)$, dus f heeft een dubbele wortel.

Neem voor de implicatie in de andere richting aan dat f een dubbele wortel x_0 heeft. We kunnen f schrijven als

$$f(x) = (x - x_0)^2(x - \beta)$$

Merk op dat

$$\begin{aligned} \frac{\partial F}{\partial y}(0) &= 2y \\ \frac{\partial F}{\partial x} &= 2(x - x_0)(x - \beta) + (x - x_0)^2 \end{aligned}$$

Het punt $(x_0, 0)$ ligt op de kromme want $F(x_0, 0) = 0^2 - f(x) = 0$, ook zijn de partiële afgeleiden 0 in $(x_0, 0)$, dus dit is een singulier punt. □

Lemma 11.4. *Elke kubische kromme met coëfficienten a_i in een lichaam K, waarvoor* $\mathrm{char}(K) \neq 2, 3$ *met een rationaal punt op de kromme kan in de Weierstrass-normaalvorm geschreven worden, en elke kromme in de Weierstrass-normaalvorm kan met een transformatie in de gereduceerde Weierstrassvorm geschreven worden.*

Bewijs. Zij E een kubische kromme die aan de bovenstaande eisen voldoet, met coëfficienten in bijvoorbeeld $\mathbb{Q}$. We schrijven de vergelijking van E in homogene vorm

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3.$$

Het feit dat een elliptische kromme in deze vorm geschreven kan worden volgt uit de Riemann-Roch stelling. De lezer kan hier meer over vinden in [Sil09, p. 42].

Nu schrijven we de vergelijking in niet-homogene coördinaten door $x = \frac{X}{Z}$, $y = \frac{Y}{Z}$

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

Merk op dat er een extra punt O op oneindig op de kromme ligt dat alleen in homogene coördinaten bestaat. Substitutie van $y \mapsto \frac{1}{2}(y - a_1x - a_3)$ geeft

$$\begin{aligned} &\frac{1}{4}(y^2 - 2a_1xy - 2a_3y + a_1^2x^2 + 2a_1a_3x + a_3^2) \\ &+ \frac{1}{2}(a_1xy - a_1^2x^2 - a_1a_3x) \\ &+ \frac{1}{2}(a_3y - a_1a_3x - a_3^2) \\ &= x^3 + a_2x^2 + a_4x + a_6. \end{aligned}$$

Haakjes uitwerken geeft

$$\frac{y^2}{4} - \frac{a_1^2x^2}{4} - \frac{a_1a_3x}{2} - \frac{a_3^2}{4} = x^3 + a_2x^2 + a_4x + a_6.$$

Beide zijden met 4 vermenigvuldigen

$$y^2 - a_1^2x^2 - 2a_1a_3x - a_3^2 = 4x^3 + 4a_2x^2 + 4a_4x + 4a_6.$$

Coëfficienten groeperen

$$y^2 = 4x^3 + (4a_2 + a_1^2)x^2 + 2(a_1a_3 + 2a_4)x + a_3^2 + 4a_6.$$

Definieer nu $b_2 = a_1^2 + 4a_2$, $b_4 = 2a_4 + a_1a_3$, $b_6 = a_3^2 + 4a_6$, dan wordt de vergelijking

$$E : y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6.$$

Om de x^2 term weg te werken doen we de substituties

$$x \mapsto \frac{x - 3b_2}{36} \qquad y \mapsto \frac{y}{108}.$$

Dit geeft

$$y^2 = x^3 + (648b_4 - 27b_2^2)x + 54b_2^3 - 1944b_2b_4 + 11664b_6.$$

Nu doen we de substituties $c_4 = b_2^2 - 24b_4$, $c_6 = -b_2^3 + 36b_2b_4 - 216b_6$, dit geeft

$$y^2 = x^3 - 27c_4x - 54c_6.$$

Ten slotte, met de substituties $b = -27c_4$, en $c = -54c_6$ krijgen we de gewenste gereduceerde Weierstrassvorm

$$y^2 = x^3 + bx + c.$$

□

11.2 De groep van punten op een elliptische kromme

De punten in een elliptische kromme vormen een groep met optelling van punten.

11.2.1 De eenheid

De eenheid is het punt O op oneindig. Dit punt bestaat alleen in homogene coördinaten. Beschouw een kromme in Weierstrass-normaalvorm

$$C : y^2 = x^3 + ax^2 + bx + c$$

Omzetten naar homogene coördinaten geeft

$$Y^2Z = X^3 + aX^2Z + bXZ^2 + cZ^3$$

Het punt O wordt nu gevonden door $Z = 0$ in te vullen, dat geeft immers $X^3 = 0$, dus $X = 0$. We kunnen Y vrij kiezen en nemen per conventie $O = [0 : 1 : 0]$. Dit geeft het unieke punt op oneindig O. Door dit punt toe te voegen geldt voor alle lijnen dat zij de kromme drie keer snijden (of raken), (waarbij x, y ook complex mogen zijn).

11.2.2 Optelling van punten

Om een formule voor optelling van punten te vinden hebben we de hoofdstelling van de algebra en één van Vieta's formules over polynomen nodig.

Stelling 11.5 (Hoofdstelling van de algebra)**.** *Een niet-constant polynoom in één variabele van graad n met rationale coëfficienten heeft n complexe wortels $r_1, r_2, \dots r_n$ waarbij meervoudige wortels meerdere keren worden meegeteld. Bovendien kan een polynoom*

$$P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \qquad a_n \neq 0$$

geschreven worden als

$$P(x) = a_n(x - r_1)(x - r_2)\dots(x - r_n).$$

Indien P een complexe wortel z heeft, dan is $\bar{z}$ ook een wortel.

Lemma 11.6 (Vieta)**.** *Zij P(x) een polynoom van graad n met coefficiënten in $\mathbb{Q}, \mathbb{R}$ of $\mathbb{C}$,*

$$P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \qquad a_n \neq 0.$$

Vanwege de hoofdstelling van de algebra heeft P(x) nu n wortels $r_1, r_2, \dots r_n$. Er geldt

$$r_1 + r_2 + \cdots + r_{n-1} + r_n = -\frac{a_{n-1}}{a_n}.$$

Bewijs. Schrijf $P(x)$ als

$$P(x) = a_n(x - r_1)(x - r_2)\dots(x - r_n).$$

Beschouw de coëfficient van x^{n-1}. Deze wordt verkregen uit het product door één term r_i te kiezen en de rest x.

$$\begin{aligned}
&a_n(x - \underline{r_1})(\underline{x} - r_2)(\underline{x} - r_3)\cdots(\underline{x} - r_n)\\
&a_n(\underline{x} - r_1)(x - \underline{r_2})(\underline{x} - r_3)\cdots(\underline{x} - r_n)\\
&a_n(\underline{x} - r_1)(\underline{x} - r_2)(x - \underline{r_3})\cdots(\underline{x} - r_n)\\
&\vdots\\
&a_n(\underline{x} - r_1)(\underline{x} - r_2)(\underline{x} - r_3)\cdots(x - \underline{r_n})
\end{aligned}$$

Optellen geeft $a_n(-r_1 - r_2 - \cdots - r_n)$ voor de coëfficient van a_{n-1}, dus volgt

$$-a_n(r_1 + r_2 + \cdots + r_n) = a_{n-1}.$$

Deling door $-a_n$ concludeert het bewijs. □

Lemma 11.7 (Optelling)**.** *Zij $P = (x_1, y_1)$, $Q = (x_2, y_2)$, $x_1 \neq x_2$, $y_1 \neq y_2$ rationale punten op een kromme $y^2 = x^3 + ax^2 + bx + c$, met $a, b, c \in \mathbb{Q}$. Dan is $P * Q$, het derde snijpunt van de lijn door P en Q met de kromme ook een rationaal punt. De optelling wordt nu gedefiniëerd door $P + Q = (P * Q) * O$, wat hetzelfde is als de reflectie van $P * Q$ in de x-as.*

Bewijs. We bepalen eerst $P * Q$. De lijn door P en Q wordt gegeven door

$$y = \lambda x + \nu \qquad \text{waar } \lambda = \frac{y_2 - y_1}{x_2 - x_1} \qquad \nu = y_1 - \lambda x_1.$$

Invullen in de vergelijking van de kromme geeft

$$\begin{aligned}(\lambda x+\nu)^2 &= x^3+ax^2+bx+c\\ \lambda^2x^2+2\lambda\nu x+\nu^2 &= x^3+ax^2+bx+c\\ x^3+(a-\lambda^2)x^2+(b-2\lambda\nu)x+(c-\nu^2) &= 0.\end{aligned}$$

Dit is een polynoom in één variabele, waarvan we de twee wortels x_1, x_2 weten. Met Vieta's formule voor een derdegraads polynoom hebben we:

$$x_1+x_2+x_3=\lambda^2-a$$

Daaruit volgt:

$$x_3=\lambda^2-a-x_1-x_2$$

Aangezien λ, a, x_1, x_2 rationaal zijn, is x_3 ook rationaal. De bijbehorende y_3 kan worden gevonden door x_3 terug te substitueren in de lijnvergelijking, wat een rationale waarde oplevert.

$$y_3=\lambda x_3+\nu \quad \text{Merk op dit is voor } P*Q\text{, neem } -y_3 \text{ voor } P+Q$$

Nu bepalen we $P+Q$. De lijn door $P*Q$ en O is de verticale lijn door $P*Q$. Aangezien een elliptische kromme in Weierstrassvorm symmetrisch is in de x-as geldt dat $P+Q$ de reflectie van $P*Q$ in de x-as is. □

Lemma 11.8 ([ST15, opgave 1.10])**.** *Zij C een kubische kromme. De operatie $*$ is commutatief.*

Bewijs. $P*Q$ is het unieke derde snijpunt van de lijn door P en Q met de kromme. De lijn door P en Q is hetzelfde als de lijn door Q en P. □

Lemma 11.9 (Duplicatie van een punt)**.** *Zij C eem kubische kromme en een rationaal punt P. We definiëren $P*P$ als het snijpunt van de raaklijn door P met de kromme, dan is $P*P$ ook weer rationaal. Nu is $P+P$ de spiegeling van $P*P$ in de x-as.*

Bewijs. De helling λ van de raaklijn in $P_0=(x_0,y_0)$ wordt gegeven door impliciet differentiëren van de kromme $y^2=f(x)$

$$\lambda=\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_P=\frac{f'(x_0)}{2y_0}$$

Deze invullen in de formule voor optelling Lemma 11.7 geeft $P+P$. □

Lemma 11.10 ([ST15, opgave 1.19])**.** *Zij $P=(x_0,y_0)$ een punt op de kromme $C: y^2=f(x)=x^3+ax^2+bx+c$, en $2P:=P+P$. De x en y coördinaten van $2P$ worden ook gegeven door*

$$\begin{aligned}x(2P) &= \frac{x_0^4-2bx_0^2-8cx_0+b^2-4ac}{4x_0^3+4ax_0^2+4bx_0+4c}\\ y(2P) &= y_0+\frac{3x_0^2+2ax_0+b}{2y_0}\cdot(x(2P)-x_0).\end{aligned}$$

Bewijs. Eerst $x(2P)$: uit Lemma 11.9 volgt

$$\lambda=\frac{3x_0^2+2ax_0+b}{2y_0}$$

Invullen in de optelformule geeft

$$x(2P) = \lambda^2 - a - 2x_0 = \frac{(3x_0^2 + 2ax_0 + b)^2}{4y_0^2} - a - 2x_0$$

We weten uit de definitie van C dat $y_0^2 = f(x_0) = x_0^3 + ax_0^2 + bx_0 + c$, dat geeft

$$\begin{aligned} x(2P) &= \frac{(3x_0^2 + 2ax_0 + b)^2}{4(x_0^3 + ax_0^2 + bx_0 + c)} - a - 2x_0 \\ &= \frac{b^2 + 4abx_0 + 4a^2x_0^2 + 6bx_0^2 + 12ax_0^3 + 9x_0^4}{4(x_0^3 + ax_0^2 + bx_0 + c)} \\ &- \frac{4a(c + bx_0 + ax_0^2 + x_0^3)}{4(x_0^3 + ax_0^2 + bx_0 + c)} \\ &- \frac{8x_0(c + bx_0 + ax_0^2 + x_0^3)}{4(x_0^3 + ax_0^2 + bx_0 + c)}. \end{aligned}$$

Haakjes uitwerken en vereenvoudigen geeft de gewenste formule voor $x(2P)$.
De formule van $y(2P)$ volgt direct uit Lemma 11.7.

$$y(2P) = \lambda \cdot x(2P) + y_0 - \lambda x_0.$$

Invullen van λ geeft de gewenste formule. □

Een snelle manier om kP te berekenen voor een grote k is door herhaald de verdubbelingsformule te gebruiken (double and add). We demonstreren dit aan de hand van een voorbeeld.

Zij $k = 19$ en P een punt op een elliptische kromme. We schrijven eerst k in binaire representatie als som van machten van 2:

$$19 = (10011)_2 = 2^4 + 2^1 + 2^0.$$

We maken een tabel van 2^jP voor $0 \leq j \leq \lfloor \log_2(k) \rfloor$

$$\begin{aligned} a_0 &= 2^0P = P \\ a_1 &= 2^1P = 2P = 2(a_0) \\ a_2 &= 2^2P = 4P = 2(a_1) \\ a_3 &= 2^3P = 8P = 2(a_2) \\ a_4 &= 2^4P = 16P = 2(a_3). \end{aligned}$$

Dan volgt:

$$19P = 2^4P + 2^1P + 2^0P = a_4 + a_1 + a_0.$$

Dit algoritme vereist slechts:

- 4 verdubbelingen (voor a_1 t/m a_4)
- 2 optellingen (voor de finale som)

Dit is significant sneller dan de naïeve methode die 18 optellingen zou vereisen ($\underbrace{P + P + \cdots + P}_{18 \text{ keer}}$).

11.2.3 De inverse van een element

De inverse van een punt $P = (x, y)$ wordt gegeven door $-P = (x, -y)$.

Bewijs. Met gevalsonderscheiding. Zij P een punt op de kromme $C : y^2 = f(x) = x^3+ax^2+bx+c$.

1. $P = (x_0, y_0) \neq O$, P voldoet aan de vergelijking van de kromme dus $-P = (x_0, -y_0)$ ook:
$$(-y_0)^2 = y_0^2 = x_0^3 + ax_0^2 + bx_0 + c$$
De lijn door P en $-P$ is de verticale lijn $x = x_0$, het derde snijpunt is O.
2. $P = O$. Aangezien O de eenheid is, volgt $-P = O$.

□

Lemma 11.11. *$P + Q + R = O$ dan en slechts dan als P, Q, R colineair zijn.*

Bewijs.
$$P + Q + R = O \iff P + Q = -R \iff P * Q = R \iff P, Q, R \text{ collineair.}$$

□

11.2.4 Associativiteit

Het algebraïsch bewijs van associativiteit van de groep is erg lang en geeft geen inzicht. Zie [Sut21] voor een algebraïsch bewijs in Sage. We geven hier een meetkundig bewijs naar [Tao11].

Lemma 11.12. *De optelling van punten op de elliptische kromme is associatief. Zij*
$$\gamma = \{(x, y) \in k^2 : y^2 = x^3 + ax + b\} \cup \{O\}$$
een niet-singuliere elliptische kromme in homogene coördinaten met $O = [0, 1, 0]$ het punt op oneindig, en zij A, B, C punten op γ. Dan $(A + B) + C = A + (B + C)$.

Associativiteit [Tao11]. Neem aan dat $O, A, B, C, A+B, B+C, -(A+B), -(B+C)$ verschillend van elkaar zijn en van $-((A + B) + C)$ en $-(A + (B + C))$.

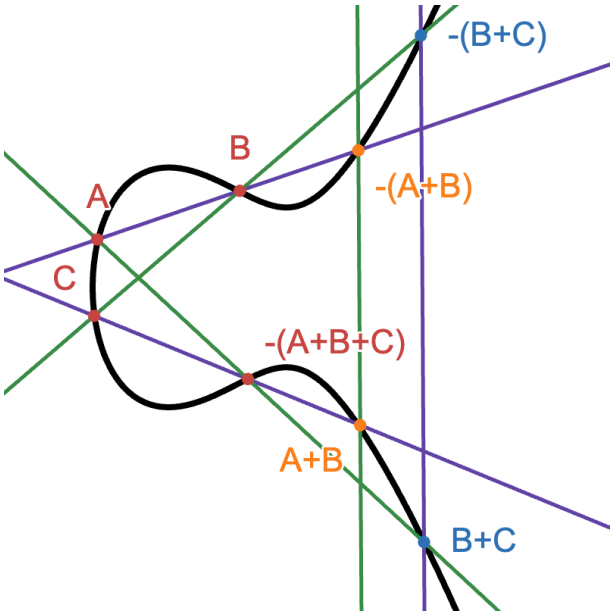
Zij nu γ_1 de vereniging van de drie lijnen $AB, C(A + B), O(B + C)$ (paars), en γ_2 de vereniging van de drie lijnen $O(A + B), BC, A(B + C)$ (groen). Merk op dat een vereniging van drie lijnen een kubische kromme is, dus γ_1, γ_2 zijn kubische krommen.

De kubische krommen γ en γ_1 hebben geen gemeenschappelijke component (γ bevat immers geen lijn). Ze treffen elkaar in de 9 verschillende punten (zie Figuur 14).
$$O, A, B, C, A + B, -(A + B), B + C, -(B + C), -((A + B) + C).$$
De kubische kromme γ_2 gaat door al de punten behalve $-((A + B) + C)$. De stelling van Cayley-Bacharach (Stelling 10.6), zegt dat γ_2 ook door het negende punt $-((A + B) + C)$ gaat.

Hieruit volgt dat de lijn door A en $B+C$ de kromme γ in zowel $-(A+(B+C))$ als $-((A+B)+C)$ treft, dus deze punten zijn gelijk.

De inverse nemen geeft nu $(A + B) + C = A + (B + C)$.

Zie Figuur 14



Figuur 14: Aangepast van [Tao11]. γ is de elliptische kromme (zwart), γ_1 is de vereniging van de paarse lijnen, γ_2 is de vereniging van de groene lijnen.

□

11.3 De orde van een punt

Een punt P op een elliptische kromme heeft orde m als

$$mP = P + P + \ldots + P = O$$

Terwijl $m'P \neq O$ voor alle $1 \leq m' < m$. Indien een dergelijke m bestaat dan heeft P een eindige orde, indien niet dan heeft P een oneindige orde [ST15][§ 2.1].

Dit is hetzelfde als de groepentheoretische orde van het element P in de groep van de elliptische kromme.

Zij $C : y^2 = f(x)$ een elliptische kromme. De punten van orde 2 zijn de punten waarvoor $2P = O \iff P = -P \iff (x, y) = (x, -y) \iff y = 0$. Dit zijn de punten $(\alpha, 0)$ waarvoor α een wortel is van $f(x)$.

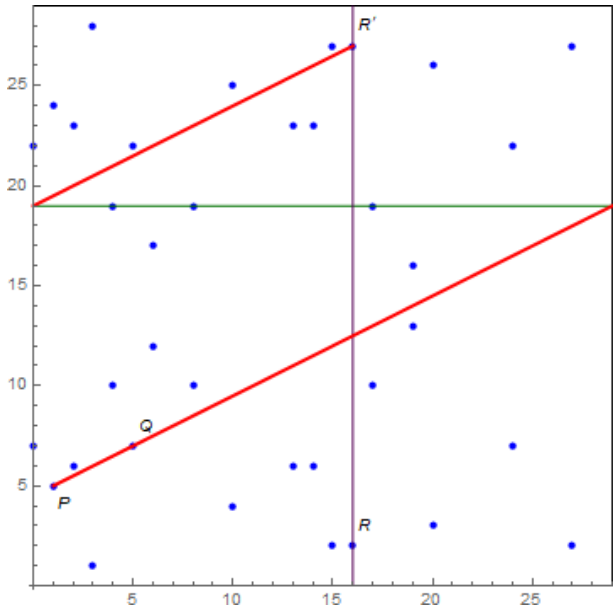
11.4 Elliptische krommen over eindige lichamen

Definitie 11.13 (Rationale punten)**.** Zij $\mathbb{F}_p$ het eindige lichaam met p elementen en zij

$$C : F(x, y) = 0$$

een polynoomvergelijking met coëfficienten in $\mathbb{F}_p$. Een oplossing (x, y) met x, y in F_p heet een rationaal punt op de kromme C.

We kunnen een elliptische kromme over een eindig lichaam als volgt visualiseren, zie Figuur 15.



Figuur 15: $y^2 \equiv x^3 + 4x + 20$ over $\mathbb{F}_{29}$. Merk op dat de kromme horizontaal symmetrisch is. De lijn door P en Q zal altijd een derde punt R treffen. Figuur van [kel21]

De formules voor optelling Lemma 11.7 en duplicatie Lemma 11.9 zijn ook geldig over eindige lichamen.

De groep van rationale punten $C(\mathbb{F}_7)$ van de kromme [ST15, Opgave 4.2]

$$C : y^2 = f(x) = x^3 + x + 1$$

over $\mathbb{F}_7$ wordt gegeven door

$$\{(0, \pm 1), (2, \pm 2).\}$$

Merk op dat we alleen x hoeven te controleren waarvoor $f(x) \in \{0, 1, 2, 4\}$, de verzameling kwadratische residuën modulo 7.

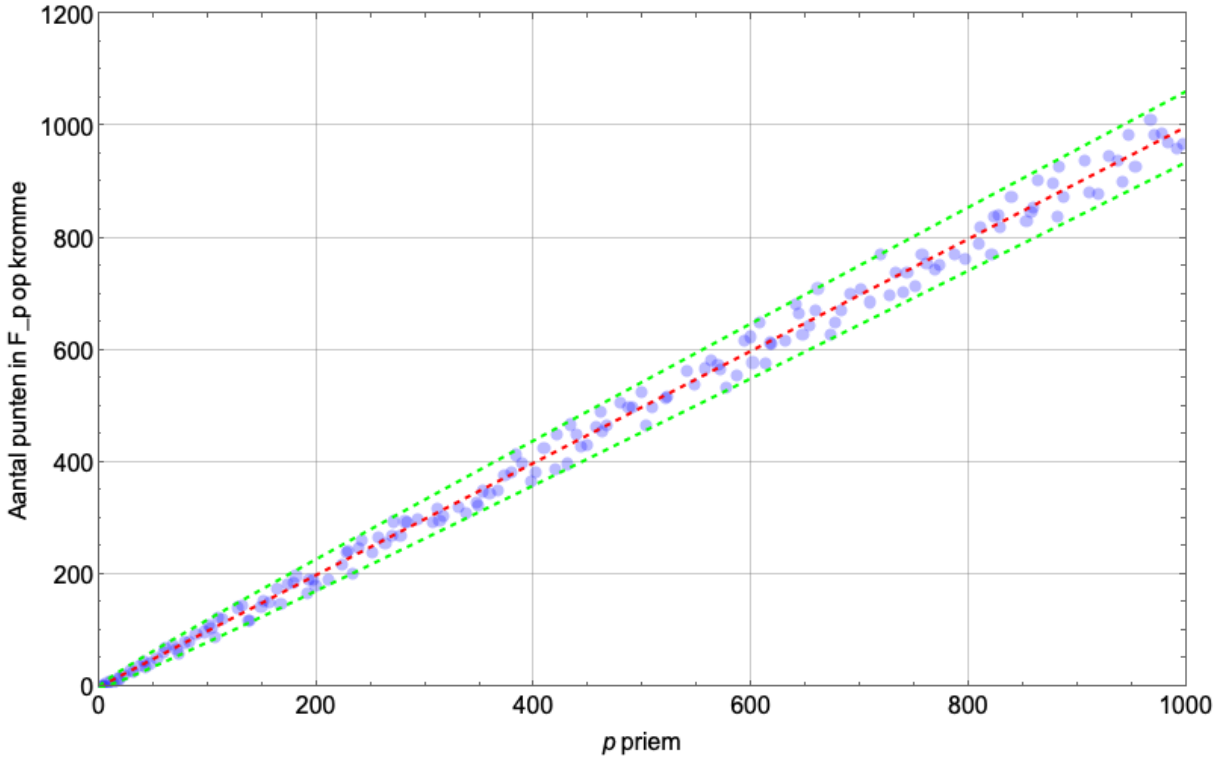
11.4.1 Het aantal punten op een elliptische kromme

Als C een niet-singuliere elliptische kromme is over een eindig lichaam $\mathbb{F}_p$, dan voldoet het aantal punten op C met coördinaten in $\mathbb{F}_p$ aan

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

Een elliptische kromme $E_{a,b}$ over $\mathbb{F}_{59}$ heeft dus tussen de 45 en 75 punten, afhankelijk van de keuze voor a, b. Er is geen eenvoudige formule om het aantal punten op $E_{a,b}$ over $\mathbb{F}_p$ te berekenen, wel zijn er algoritmes, waaronder Schoofs algoritme.

Een voorbeeld van Hasse-Weil voor één kromme is te zien in Figuur 16.



Figuur 16: Demonstratie van Hasse-Weil voor de kromme $y^2 = x^3 + x + 5$ over $\mathbb{F}_p$ voor p priem in $\{2, 3, 5, 7, \ldots, 997\}$

Kunnen we voor elk getal binnen de grenzen van Hasse-Weil een kromme vinden met precies zoveel punten?

Lemma 11.14 (Deuring)**.** *[Len87] Voor alle gehele getallen m in het interval $]p + 1 - 2\sqrt{p}, p + 1 + 2\sqrt{p}[$ bestaat een paar (a, b) in de verzameling*

$$\{(a, b) : a, b \in \mathbb{F}_p : 4a^3 + 27b^2 \neq 0\}$$

waarvoor $|E_{a,b}(\mathbb{F}_p)| = m$. Dit betekent er een niet-singuliere elliptische kromme over $\mathbb{F}_p$ bestaat

$$E : y^2 = x^3 + ax + b$$

met exact m punten.

Hieruit volgt dat de grenzen van Hasse-Weil optimaal zijn.

12 Het algoritme van Lenstra

Het factorisatiealgoritme van Lenstra, ook bekend als Lenstra's Elliptic Curve Method (ECM), maakt gebruik van de groepsstructuur van elliptische krommen om getallen te factoriseren. Het algoritme werd gepresenteerd door Lenstra in zijn paper "Factoring integers with elliptic curves" [Len87] uit 1987. Het is een verbetering van Pollards $p-1$. Lenstra's algoritme werkt ook als $p-1$ niet (macht)glad is.

Het idee van het algoritme is hetzelfde als van Pollards $p-1$ algoritme. Waar Pollard gebruik maakt van het feit dat $\left|\mathbb{F}_p^*\right| = p-1$, gebruikt Lenstra's algoritme de groep $E_{a,b}(\mathbb{F}_p)$ waarvan we weten dat de orde variëert voor verschillende a, b. Door verschillende krommen te proberen is de kans dat het algoritme succesvol factoriseert veel groter.

12.1 Pseudokromme

Zij n een samengesteld getal. Merk op dat $\mathbb{Z}/n\mathbb{Z}$ geen lichaam is, want niet alle elementen hebben een inverse.

We kunnen een *pseudokromme* [Len87, p. 664] definiëren over de ring $\mathbb{Z}/n\mathbb{Z}$

$$E_{a,b}(\mathbb{Z}/n\mathbb{Z}) = \{(x : y : z) \in \mathbb{P}^2(\mathbb{Z}/n\mathbb{Z}) : y^2 = x^3 + ax^2 + bz^3\}$$

De groepsstructuur van $E_{a,b}(\mathbb{Z}/n\mathbb{Z})$ is ingewikkelder dan de structuur van elliptische krommen over eindige lichamen. We werken daarom met een pseudo-optelling op een deelverzameling van $E_{a,b}(\mathbb{Z}/n\mathbb{Z})$.

Zij $O := (0 : 1 : 0) \in \mathbb{P}^2(\mathbb{Z}/n\mathbb{Z})$ en zij $V_n \subset \mathbb{P}^2(\mathbb{Z}/n\mathbb{Z})$ gegeven door

$$V_n = \{(x : y : 1) : x, y \in \mathbb{Z}/n\mathbb{Z}\} \cup O$$

Wanneer we een punt $P \in V_n$ hebben en een priemgetal $p \mid n$, dan is P_p het punt van $\mathbb{P}^2(\mathbb{Z}/n\mathbb{Z})$ dat wordt verkregen door iedere coördinaat van P modulo p te reduceren. Daarbij geldt $P_p = O_p \iff P = O$.

12.2 Pseudo-optelling op de pseudokromme

De optelling op de pseudokromme is niet voor alle elementen van $E_{a,b}(\mathbb{Z}/n\mathbb{Z})$ gedefiniëerd. Het briljante inzicht van Lenstra is dat het mislukken van de optelling ons een priemfactor p van n kan geven.

Zij gegeven een samengestelde $n > 1$, $a \in \mathbb{Z}/n\mathbb{Z}$ en twee punten $P, Q \in V_n$. Het algoritme geeft ofwel een niet-triviale deler g van n (mislukte optelling), ofwel een nieuw punt $R \in V_n$ met de volgende eigenschappen:

Zij p priem en een deler van n. Zij $\bar{a} = a \pmod p$. Indien er een $b \in \mathbb{F}_p$ bestaat waarvoor $E_{\bar{a},b}$ niet singulier is en $P_p, Q_p \in E_{\bar{a},b}(\mathbb{F}_p)$ dan is $R_p = P_p + Q_p$ in de groep $E_{\bar{a},b}$ (optelling op elliptische kromme over eindig lichaam).

De mogelijke uitkomsten van de pseudo-optelling van P en Q zijn

1. Als $P = O$ return $R = Q$.
2. Als $P \neq O$ en $Q = O$ return $R = P$.
3. Als $P \neq O$ en $Q \neq O$, zij dan $P(x_1 : y_1 : 1)$ en $Q = (x_2 : y_2 : 1)$. Bereken $g = \text{ggd}(x_1 - x_2, n)$.
   (a) Als $1 < g < n$ return g, een niet-triviale deler van n.

(b) Als $g = 1$, bereken

$$\lambda = (y_1 - y_2)(x_2 - x_1)^{-1} \qquad x_3 = \lambda^2 - x_1 - x_2 \qquad y_3 = \lambda(x_1 - x_3) - y_1.$$

Return $R = (x_3 : y_3 : 1)$.

(c) Als $\text{ggd}(x_1 - x_2, n) = n$, dan $x_1 = x_2$. Bereken $g = \text{ggd}(y_1 + y_2, n)$.

i. Als $1 < g < n$ return g, een niet-triviale deler van n.

ii. Als $g = n$ return $R = O$.

iii. Als $g = 1$ dan $P = Q$. Bereken

$$\lambda = (3x_1^2 + a)(y_1 + y_2)^{-1} \qquad x_3 = \lambda^2 - x_1 - x_2 \qquad y_3 = \lambda(x_1 - x_3) - y_1.$$

Return $R = (x_3 : y_3 : 1)$.

12.3 Vermenigvuldiging

Vermenigvuldiging van een punt met een getal k, genoteerd als kP, is gedefiniëerd als herhaald een punt bij zichzelf optellen. Deze operatie is daarom ook niet voor alle elementen van $E_{a,b}(\mathbb{Z}/n\mathbb{Z})$ gedefiniëerd. We kunnen een punt in verschillende volgorde bij zichzelf optellen, bijvoorbeeld

$$5P = P + (P + (P + (P + P))) \qquad \text{of } 5P = ((P + P) + (P + P)) + P$$

Of de vermenigvuldiging slaagt kan afhangen van de volgorde van optelling.

12.4 Factoriseren met het algoritme

Zij n samengesteld en kies een pseudokromme $E_{a,b}(\mathbb{Z}/n\mathbb{Z})$ met een punt P op de kromme. Laat B de zoekgrens zijn. We proberen met de hierboven gedefiniëerde vermenigvuldiging $B!P$ te berekenen.

Een noodzakelijke voorwaarde voor het slagen van het algoritme is dat voor p een priemdeler van n geldt dat $\#E(\mathbb{F}_p) \mid L!$.

Voorbeeld [Sil09, p. 371]: Zij $N = 6887$, we kiezen willekeurig een elliptische kromme en een punt P op de kromme

$$E : y^2 = x^3 + 14x + 19 \qquad P = (1512, 3166)$$

Ook kiezen we een zoekgrens B. Nu berekenen we in iedere stap van het algoritme $j!P$ voor $2 \leq j \leq B$. Dit mislukt in de 7-de stap

$$6!P = 6 \cdot 5!P \equiv (6141, 5581)$$
$$7!P = 7 \cdot 6!P \equiv (6!P + 2 \cdot 6!P) + 4 \cdot 6!P = 3 \cdot 6!P + 4 \cdot 6!P$$

Hier mislukt de optelling in de stap $3 \cdot 6!P + 4 \cdot 6!P$ want

$$3 \cdot 6!P = (984, 589) := (x_1, y_1) \qquad 4 \cdot 6!P = (203, 2038) := (x_2, y_2)$$

Voor de optelling van deze punten hebben we

$$\lambda = (y_1 - y_2)(x_2 - x_1)^{-1}$$

nodig. Maar $\text{ggd}((x_2 - x_1), N) = \text{ggd}(203 - 984, 6887) = 71$, dus $(x_2 - x_1)$ heeft geen inverse modulo N: de optelling mislukt. Het gevolg is dat we een factor $p = 71$ met $p \mid N$ gevonden hebben. De factorisatie is $N = 71 \cdot 97$.

Dit werkt omdat in $E(\mathbb{F}_{71})$ geldt

$$63P \equiv O \pmod{71}$$

Terwijl in $E(\mathbb{F}_{97})$ geldt

$$107P \equiv O \pmod{97}$$

En er geldt dat $63 \mid 7!$, maar $107 \nmid 7!$

In Algoritme 5 geven we pseudocode van een (versimpelde) versie van het algoritme.

12.5 Algoritme

Algorithm 5 Algoritme van Lenstra

Require: Een samengestelde $n > 2$. Een zoekgrens B
Ensure: Een niet-triviale factor g van n, of stop en print 'mislukt' anders

```
if ggd(n, 6, =)1 then
    stop en print 6 is een factor
end if
label 'Kies kromme'
Kies random b, x_1, y_1 ∈ Z/nZ
P_0 ← (x_1, y_1)
c ← y_1^2 − x_1^3 − bx_1
E ← y^2 = x^3 + bx + c                ▷ dit is een elliptische kromme E (mod n)
P ← P_0
for L ← 2 to B do
    Q ← LP (mod n)                    ▷ Gebruik double and add, optelling in E(Z/nZ)
                                      ▷ Merk op dat nu Q = L!P_0
                                      ▷ Voor de optelling is het nodig een element a te inverteren in Z/nZ

    if Inverteren van a mislukt then
        g ← ggd(a, n)
        if g < n then
            stop en print g is een factor van n
        end if
        if g = n then
            GOTO 'Kies Kromme'
        end if
    else
        P ← Q
    end if
end for
GOTO 'Kies Kromme'
```

12.6 Complexiteit

Zij n samengesteld, en p de kleinste priemfactor van n, dan is de (heuristische) looptijd [Sil09, p. 367]

$$g(p)_c = \exp\left(c\sqrt{\ln(p)(\ln\ln(p))}\right) \qquad \text{stappen}$$

De lezer kan hier meer over vinden in Lenstra's originele artikel [Len87].

Indien n geen kleine priemfactoren heeft, dus voor $p \approx \sqrt{n}$, is het algoritme niet efficient. Bijvoorbeeld voor $c = 1$ en alle $n \in [10^3, 10^{100}]$ geldt

$$n^{\frac{1}{7}} \leq g(\sqrt{n})_1 \leq n^{\frac{1}{2}},$$

dus exponentiële tijd en nauwelijks beter dan proefdeling. Het veelgebruikte crypotografiealgoritme RSA berust op de onfactoriseerbaarheid van een $n = pq$ waarbij p, q priem van dezelfde orde van grootte. Lenstra's algoritme is dus niet effectief om RSA te kraken.

Referenties

[AS65] Milton Abramowitz en Irene A. Stegun, red. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Revised edition. New York, NY: Dover Publications, 1 jun 1965. 1046 p. ISBN: 978-0-486-61272-0.

[Cha99] Robin Chapman. *Evaluating zeta(2)*. Originele publicatie 1999-04-30, herzien op 2003-07-07. 30 apr 1999. URL: https://empslocal.ex.ac.uk/people/staff/rjchapma/etc/zeta2.pdf (bezocht op 30-12-2024).

[Cor+22] Thomas H. Cormen e.a. *Introduction to algorithms*. Fourth edition. Cambridge, Massachusetts London: The MIT Press, 2022. 1 p. ISBN: 978-0-262-04630-5 978-0-262-36750-9.

[CP05] Richard E. Crandall en Carl Pomerance. *Prime numbers: a computational perspective*. 2nd ed. New York: Springer, 2005. ISBN: 978-0-387-25282-7.

[Dco13] Dcoetzee. *Prime number theorem ratio convergence*. 21 mrt 2013. URL: https://commons.wikimedia.org/wiki/File:Prime_number_theorem_ratio_convergence.svg (bezocht op 02-01-2025).

[DF04] David Steven Dummit en Richard M. Foote. *Abstract algebra*. 3rd ed. Hoboken, NJ: Wiley, 2004. 932 p. ISBN: 978-0-471-43334-7.

[EP85] Paul Erdos en Carl Pomerance. „On the normal number of prime factors of $\phi(n)$". In: *Rocky Mountain Journal of Mathematics* 15.2 (1 jun 1985). ISSN: 0035-7596. DOI: 10.1216/RMJ-1985-15-2-343. URL: https://projecteuclid.org/journals/rocky-mountain-journal-of-mathematics/volume-15/issue-2/On-the-normal-number-of-prime-factors-of-phin/10.1216/RMJ-1985-15-2-343.full (bezocht op 08-12-2024).

[Eve24a] Jan-Hendrik Evertse. „Introduction to prime number theory". In: *Analytic Number Theory (Mastermath)*. 2024. URL: https://pub.math.leidenuniv.nl/%E2%88%BCevertsejh.

[Eve24b] Jan-Hendrik Evertse. „The Prime number theorem for arithmetic progressions". In: *Analytic Number Theory (Mastermath)*. 2024. URL: https://pub.math.leidenuniv.nl/%E2%88%BCevertsejh.

[Gor85] John Gordon. „Strong Primes are Easy to Find". In: *Advances in Cryptology*. Red. door Thomas Beth, Norbert Cot en Ingemar Ingemarsson. Berlin, Heidelberg: Springer, 1985, p. 216–223. ISBN: 978-3-540-39757-1. DOI: 10.1007/3-540-39757-4_19.

[HT93] Adolf Hildebrand en Gerald Tenenbaum. „Integers without large prime factors". In: *Journal de théorie des nombres de Bordeaux* 5.2 (1993), p. 411–484. ISSN: 2118-8572. URL: http://www.numdam.org/item/JTNB_1993__5_2_411_0/ (bezocht op 31-12-2024).

[kel21] kelalaka. *Graphically representing points on Elliptic Curve over finite field*. Cryptography Stack Exchange. 15 jan 2021. URL: https://crypto.stackexchange.com/questions/12093/graphically-representing-points-on-elliptic-curve-over-finite-field (bezocht op 11-01-2025).

[Len87] H. W. Lenstra. „Factoring Integers with Elliptic Curves". In: *Annals of Mathematics* 126.3 (1987). Publisher: [Annals of Mathematics, Trustees of Princeton University on Behalf of the Annals of Mathematics, Mathematics Department, Princeton University], p. 649–673. ISSN: 0003-486X. DOI: 10.2307/1971363. URL: https://www.jstor.org/stable/1971363 (bezocht op 04-12-2024).

[Mon87] Peter L. Montgomery. „Speeding the Pollard and Elliptic Curve Methods of Factorization". In: *Mathematics of Computation* 48.177 (1987). Publisher: American Mathematical Society, p. 243–264. ISSN: 0025-5718. DOI: 10.2307/2007888. URL: https://www.jstor.org/stable/2007888 (bezocht op 10-12-2024).

[Pol74] J. M. Pollard. „Theorems on factorization and primality testing". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 76.3 (nov 1974), p. 521–

528. ISSN: 1469-8064, 0305-0041. DOI: 10.1017/S0305004100049252. URL: https://www.cambridge.org/core/journals/mathematical-proceedings-of-the-cambridge-philosophical-society/article/abs/theorems-on-factorization-and-primality-testing/6762E84DBD34AEF13E6B1D1A8334A989 (bezocht op 04-12-2024).

[R S74] R. Sherman Lehman. „Factoring Large Integers". In: *Mathematics of Computation* 28.126 (apr 1974), p. 637–646. URL: https://www.ams.org/journals/mcom/1974-28-126/S0025-5718-1974-0340163-2/S0025-5718-1974-0340163-2.pdf.

[Rab80] Michael O Rabin. „Probabilistic algorithm for testing primality". In: *Journal of Number Theory* 12.1 (1 feb 1980), p. 128–138. ISSN: 0022-314X. DOI: 10.1016/0022-314X(80)90084-0. URL: https://www.sciencedirect.com/science/article/pii/0022314X80900840 (bezocht op 14-12-2024).

[Ras11] Raskolnikov. *Answer to "What Does Homogenisation Of An Equation Actually Mean?"* Mathematics Stack Exchange. 10 jan 2011. URL: https://math.stackexchange.com/a/17027 (bezocht op 15-12-2024).

[RS01] Ron Rivest en Robert Silverman. *Are 'Strong' Primes Needed for RSA*. Publication info: Published elsewhere. Unknown where it was published. 2001. URL: https://eprint.iacr.org/2001/007 (bezocht op 11-12-2024).

[SWM95] Jeffrey Shallit, Hugh C. Williams en François Morain. „Discovery of a lost factoring machine". In: *The Mathematical Intelligencer* 17.3 (1 jan 1995), p. 41–47. ISSN: 0343-6993. DOI: 10.1007/BF03024369. URL: https://doi.org/10.1007/BF03024369 (bezocht op 12-01-2025).

[Sil09] Joseph H. Silverman. *The Arithmetic of Elliptic Curves*. Deel 106. Graduate Texts in Mathematics. New York, NY: Springer New York, 2009. ISBN: 978-0-387-09493-9 978-0-387-09494-6. DOI: 10.1007/978-0-387-09494-6. URL: http://link.springer.com/10.1007/978-0-387-09494-6 (bezocht op 20-10-2024).

[ST15] Joseph H. Silverman en John T. Tate. *Rational Points on Elliptic Curves*. Undergraduate Texts in Mathematics. Cham: Springer International Publishing, 2015. ISBN: 978-3-319-18587-3 978-3-319-18588-0. DOI: 10.1007/978-3-319-18588-0. URL: https://link.springer.com/10.1007/978-3-319-18588-0 (bezocht op 20-10-2024).

[Smi29] Smithsonian Institution. *Factor Stencils by Derrick N. Lehmer*. Smithsonian Institution. 1929. URL: https://www.si.edu/object/factor-stencils-derrick-n-lehmer%3Anmah_1214102 (bezocht op 12-01-2025).

[SH16] Gilbert Strang en Edwin Herman. *11.5: Conic Sections*. Mathematics LibreTexts. 11 jul 2016. URL: https://math.libretexts.org/Bookshelves/Calculus/Calculus_(OpenStax)/11%3A_Parametric_Equations_and_Polar_Coordinates/11.05%3A_Conic_Sections (bezocht op 07-01-2025).

[Sut21] Andrew Sutherland. *Lecture 2 Proof of associativity*. Algebraic proof of the associativity of the elliptic curve group law on curves defined by a short Weierstrass equation, as presented in Lecture 2. 22 feb 2021. URL: https://cocalc.com/share/public_paths/a6a1c2b188bd61d94c3dd3bfd5aa73722e8bd38b (bezocht op 04-01-2025).

[Tao11] Terry Tao. *Pappus theorem and elliptic curves*. What's new. 15 jul 2011. URL: https://terrytao.wordpress.com/tag/pappus-theorem/ (bezocht op 07-01-2025).

[Wag13] Samuel S. Wagstaff. *The joy of factoring*. Student mathematical library volume 68. Providence: American mathematical society, 2013. ISBN: 978-1-4704-1048-3.

[Wal25] Kim Walisch. *kimwalisch/primesieve*. original-date: 2013-09-24T17:58:41Z. 10 jan 2025. URL: https://github.com/kimwalisch/primesieve (bezocht op 10-01-2025).

[Weg22] Benne de Weger. *Elementaire Getaltheorie en Asymmetrische Cryptografie*. 4de ed. Epsilon uitgaven, 2022. 192 p. ISBN: 978-90-5041-108-0. URL: https://www.epsilon-

uitgaven.nl/wetenschappelijke-reeks/elementaire-getaltheorie-en-asymmetrische-cryptografie/11026 (bezocht op 20-10-2024).

[Wil82] H. C. Williams. „A $p + 1$ Method of Factoring". In: *Mathematics of Computation* 39.159 (1982). Publisher: American Mathematical Society, p. 225–234. ISSN: 0025-5718. DOI: 10.2307/2007633. URL: https://www.jstor.org/stable/2007633 (bezocht op 11-12-2024).

[Wil] Steven Wilson. *A Gallery of Cubic Plane Curves*. URL: http://www.milefoot.com/math/planecurves/cubics.htm (bezocht op 02-01-2025).