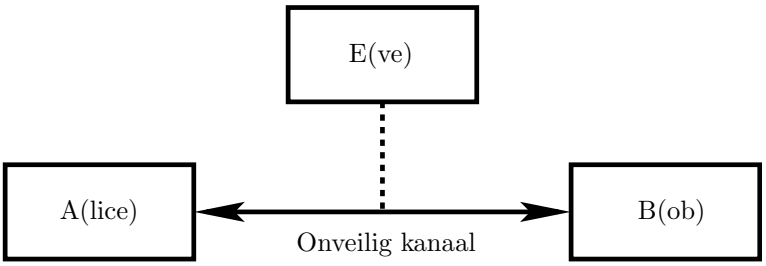
CURSUSNOTA'S GETALTHEORIE EN CRYPTOGRAFIE UHASSELT, 2022-2023

MICHEL VAN DEN BERGH

1. Korte inleiding over het gebruik van cryptografie in de praktijk

1.1. **Symmetrische encryptie.** De standaardsetting in cryptografie is als volgt:



Als Alice een bericht naar Bob wil sturen dan zal ze dit bericht vercijferen zodat Eve het bericht niet kan lezen. Traditioneel gebeurt dit vercijferen (encryptie) door middel van een geheim dat door Alice en Bob gedeeld wordt en dat niet bekend is aan Eve. Zo'n "geheim" noemt men een *sleutel*. We noemen dit een "symmetrisch" encryptiesysteem omdat de rol van Alice en Bob symmetrisch is.

Formeel bestaat een symmetrisch encryptiesysteem uit verzamelingen

$$\begin{aligned}\mathcal{M} &= \{\text{berichten}\}\\ \mathcal{C} &= \{\text{vercijferde berichten}\}\\ \mathcal{K} &= \{\text{sleutels}\}\end{aligned}$$

en functies

$$\begin{aligned}E &: \mathcal{K}\times\mathcal{M}\to\mathcal{C} \qquad \text{(encryptie)}\\ D &: \mathcal{K}\times\mathcal{C}\to\mathcal{M} \qquad \text{(decryptie)}\end{aligned}$$

zodat

$$\forall m\in\mathcal{M}, \forall K\in\mathcal{K} : D(K,E(K,m)) = m$$

Normaal zullen $\mathcal{M},\mathcal{C},\mathcal{K}$ de verzamelingen van woorden zijn bestaande uit een gegeven aantal bits. Voor $\mathcal{M}$ en $\mathcal{C}$ is dit aantal bits normaal hetzelfde en men noemt het de "block size". Het aantal bits voor $\mathcal{K}$ is de "key size". We noteren in het bijzonder dat $\mathcal{M}$, $\mathcal{C}$, $\mathcal{K}$ in het algemeen eindige verzamelingen zijn. Om een lang bericht te vercijferen zal men het dan ook in kleine stukjes kappen die individueel kunnen voorgesteld worden door een element van $\mathcal{M}$. Het is absoluut niet triviaal om dit te doen op een manier die geen informatie lekt na vercijfering. We gaan hier echter niet verder op in.

Een van de grondbeginselen van de cryptografie is het principe van "Kerckhoffs" dat zegt dat in een goed cryptografisch systeem D en E publiek zijn en enkel de

sleutels geheim. Er zijn een hele hoop redenen om dit principe aan te nemen. Hier zijn er enkelen:

- Cryptografische algoritmen worden ingebouwd in hardware. Ze zijn dus moeilijk aan te passen indien ze uitlekken. Nieuwe sleutels zijn daarentegen "gemakkelijk" te verspreiden.
- Cryptografie pretendeert tot op zekere hoogte een exacte wetenschap te zijn. De informatieinhoud van een sleutel wordt uitgedrukt in bits en is welgedefinieerd. De informatieinhoud van een algoritme is een veel vager concept.
- Bij het bedenken van cryptografische algoritmen is het gemakkelijk om denkfouten te maken. Het is dus goed dat cryptografische algoritmen aan een zo groot mogelijke "public scrutiny" worden blootgesteld zodat eventuele zwakheden snel ontdekt worden.

Opmerking 1.1. Een der meest bekende toepassingen van cryptografie is de copiëerbeveiliging van digitale data (bijvoorbeeld films of muziek). Dit noemt men "Digital Rights Management" of kortweg DRM. Deze toepassing valt niet onder de "standaard setting" die we hierboven hebben aangegeven omdat Bob en Eve nu eigenlijk dezelfde persoon zijn. Met andere woorden de ontvanger van de data moet deze kunnen decrypten (bijvoorbeeld om een film af te spelen op een monitor) maar hij/zij mag niet in staat zijn om de gedecrypte data op te slaan om deze bijvoorbeeld verder te verspreiden. Dit is natuurlijk een logische contradictie en DRM systemen zijn in de praktijk enkel geschikt om "casual piracy" tegen te gaan. Gedetermineerde hackers laten zich er niet door afschrikken.

Een der meest geavanceerde DRM systemen is de copiëerbeveiliging van BluRay DVD's. Maar zelfs hiervoor is het niet moeilijk om commerciële pakketten te vinden die deze bescherming omzeilen.

De hoofdvoorwaarde voor de veiligheid van een cryptosysteem is natuurlijk dat decryptie zonder de kennis van de sleutel "heel moeilijk" moet zijn. Er is echter geen enkel cryptosysteem waarvan formeel is aangetoond dat decryptie *echt* heel moeilijk is. De praktijkdefinitie van veiligheid is dat het cryptosysteem bestand moet zijn tegen standaardaanvallen (zoals "differentiële cryptoanalyse"). Dit garandeert natuurlijk niet dat het ook bestand zal zijn tegen eventuele nieuwe soorten aanvallen.

Een andere manier om te "bewijzen" dat een cryptosysteem veilig is (die voornamelijk gebruikt wordt voor publieke sleutel systemen, waarvan hieronder sprake), is aan te tonen dat het kraken van het systeem equivalent is met het oplossen van een wiskundig probleem dat "heel moeilijk" is, bijvoorbeeld ontbinden in factoren. Helaas is het zelfs bij ontbinden in factoren niet bewezen dat dit niet efficiënt kan.

Moeilijkheid van decryptie zonder sleutel is slechts een van de vele voorwaarden die opgelegd worden aan een modern cryptosysteem. Zo moet ook gelden met $c = E(K, m)$

$$(m, c) \xrightarrow{\text{moeilijk}} K$$

Dit wil zeggen: indien een bericht lekt tezamen met zijn vercijferde versie dan moet het "heel moeilijk" zijn om hieruit de sleutel te halen. Dit is zelfs zo indien een

onbeperkt aantal berichtenparen lekt. Schematisch

$$\underbrace{(m_1, c_1), \ldots, (m_N, c_N)}_{\text{onbeperkt}} \xrightarrow{\text{moeilijk}} K$$

De vroegere standaard voor symmetrische encrypties was DES (Data Encryption Standard) met een block size van 64 bits en een key size van 56 bits. Dit laatste is echter veel te weinig voor de huidige technologie. In 1999 werd in een distributed project de volledige key space doorlopen in 22 uur (cfr Wikipedia).

De huidige standaard is AES (Advanced Encryption Standard) die ontwikkeld werd aan de KULeuven en die de winnaar was van een competitie die werd uitgeschreven door het Amerikaanse "National Institute of Standards" met de bedoeling een opvolger te vinden voor DES. AES heeft een block size van 128 bits en key sizes van 128, 192 en 256 bits. Het leuke van AES is dat de niet lineaire component van het algoritme (de zogenaamde S-box) geconstrueerd werd met behulp van een eindig lichaam. De sterk algebraïsche structuur van AES heeft enige nervositeit veroorzaakt vanwege het vermoeden dat het eventueel gekraakt zou kunnen worden met methoden uit de computationele algebraïsche meetkunde.

1.2. **Asymmetrische encryptie.** Een probleem met symmetrische encryptie is dat het enkel werkt indien Alice en Bob reeds een sleutel delen. Ze moeten dus reeds over een veilig kanaal beschikken om die sleutel te kunnen delen. Als Bob bijvoorbeeld Amazon is en Alice een klant dan is het helemaal niet evident hoe zulk een veilig kanaal zou kunnen gecreëerd worden. De oplossing voor dit probleem wordt aangeleverd door de "publieke sleutel cryptografie" of "asymmetrische encryptie".

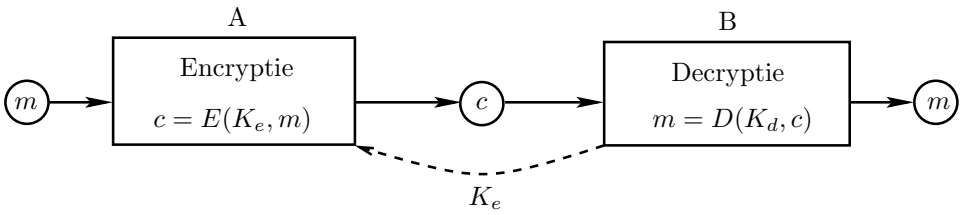
Het principe van publieke sleutel cryptografie is dat van een hangslot. Iedereen kan dit sluiten maar enkel de persoon met de sleutel kan het openen. Bob stuurt zo'n hangslot naar Alice, Alice steekt haar bericht in een doosje, sluit het doosje met het hangslot van Bob en stuurt het vervolgens terug naar Bob. Indien Eve het doosje zou onderscheppen dan zou ze het niet kunnen openen omdat ze de sleutel van het hangslot niet niet heeft.

In publieke sleutel cryptografie heeft Alice geen sleutel maar Bob heeft daarentegen een *sleutelpaar* (K_d, K_e).

- (K_e) is de publieke sleutel die dient voor encryptie. Het is het digitale analogon van het hangslot. Bob's publieke sleutel wordt vrij verspreid en is in het bijzonder bekend aan Eve.
- (K_d) is de private sleutel die dient voor decryptie. Het is het digitale analogon van de sleutel van het hangslot. Het is belangrijk dat Bob deze sleutel geheim houdt.

Zoals voorheen zijn er encryptie- en decryptiefuncties die moeten voldoen aan

$$D(K_d, E(K_e, m)) = m$$

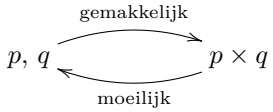


Het principe is dus

- *Iedereen* kan een vercijferd bericht naar Bob sturen.
- Enkel Bob kan het lezen want Bob is de enige die K_d kent.

De voorwaarde is weer: decryptie zonder K_d moet "heel moeilijk" zijn. Dit betekent dat de functie $m \mapsto E(K_e, m)$ heel moeilijk te inverteren moet zijn, tenzij je K_d kent. Het is helemaal niet duidelijk dat zo'n moeilijk te inverteren functies bestaan. Bij de constructie van dergelijke functies komt dan ook heel veel wiskunde kijken. Hier zijn enige voorbeelden.

(RSA) Een standaard publiek sleutel systeem is RSA, genaamd naar de uitvinders Rivest-Shamir-Adleman. RSA steunt op het feit dat ontbinden in factoren wordt verondersteld moeilijk te zijn. Schematisch voor p, q priem



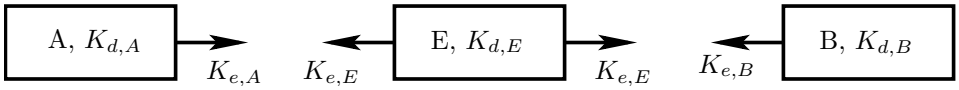
(ECC) Tegenwoordig schakelt men meer en meer over naar de zogenaamde "Elliptic Curve Cryptography". Een elliptische kromme is een kromme van graad 3. Bijvoorbeeld

$$y^2 = x^3 + ax + b \tag{1.1}$$

waarbij a, b elementen van het grondlichaam zijn. In cryptografie is dit grondlichaam typisch $\mathbb{Z}/p\mathbb{Z}$. Het is niet evident maar de oplossingen van (1.1) hebben een groepstructuur en deze is de bron van vele cryptografische algoritmen.

Krommen van graad 2 zijn zogenaamde "kegelsneden". Deze worden geacht wiskundig volledig begrepen te zijn. Over elliptische krommen zijn er echter vandaag de dag nog diepe onopgeloste problemen.

1.3. **"Man In the Middle" (MITM) attack.** Publieke sleutelcryptografie lost de creatie van een veilig kanaal op indien Eve enkel het kanaal afluistert. Men andere woorden: indien haar rol *passief* is. In de praktijk kan ze echter veel meer kwaad aanrichten. Ze kan namelijk *actief* ingrijpen op het kanaal zoals in de volgende figuur



Dus Eve genereert haar eigen sleutelpaar en:

- ze doet zich ten opzichte van Alice voor als Bob;
- ze doet zich ten opzichte van Bob voor als Alice.

Op deze manier beheerst Eve de volledige communicatie.

1.4. **Digitale handtekeningen.** MITM attacks worden voorkomen door zogenaamde *certificaten*. Het is niet echt een bevredigende oplossing maar het is de beste oplossing die er op dit moment bestaat. Certificaten maken gebruik van het concept van een "digitale handtekening". Een *digitale handtekening* is een stukje data verbonden met een boodschap dat

- de integriteit van de boodschap;
- de identiteit van de afzender

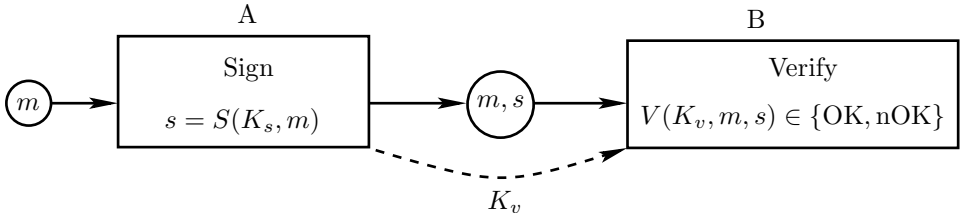
garandeert.

Om een digitale handtekening te genereren steunt men weer op een sleutelpaar (K_s, K_v) waarbij

(K_s) de "signing" key is. Deze is enkel gekend door de persoon die de digitale handtekening plaatst.

(K_v) de "verification" key is. Deze is publiek.

We hebben nu weer twee functies S(ign), V(erify) zoals in de volgende figuur



Met andere woorden. Enkel Alice kan het document tekenen want enkel zij kent K_s. Iedereen kan echter nagaan dat het document van Alice komt omdat K_v publiek is. Merk op dat dit schema op zichzelf weer kwetsbaar is voor een MITM attack. Hoe weet Bob namelijk of de K_v die Alice hem zogenaamd toestuurt echt van haar afkomstig is?

1.5. **Certificaten.** Een certificaat is een electronisch document dat bevat:

(1) een webadres en/of persoonlijke informatie;
(2) een publieke sleutel;
(3) een digitale handtekening door *een derde partij* die door iedereen vertrouwd wordt.

Zo'n derde partij noemt men een *Certificate Authority* of kortweg CA. De informatie om digitale handtekeningen te controleren van bekende CA's zit ingebakken in de browser (of andere software voor online transacties). CA's zijn meestal commerciële ondernemingen (bijvoorbeeld Verizon) maar het kunnen ook nationale overheden zijn (zoals bijvoorbeeld in het geval van een electronisch paspoort).

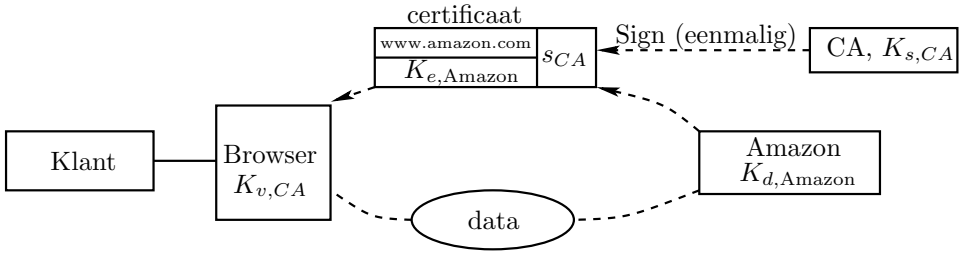
Een certificaat is een document dat een beperkte geldigheidsduur heeft. Het is een publiek document dat met gepaste software kan gedownload worden. Hieronder volgt als voorbeeld het certificaat van Amazon.

```
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number:
            6b:66:ae:56:5f:d0:3f:7d:1e:2b:c0:bd:4a:f3:3c:66
    Signature Algorithm: sha1WithRSAEncryption
        Issuer: C=US, O=VeriSign, Inc., OU=VeriSign Trust Network, OU=Terms of use at https://www.verisign.com/rpa (c)10, CN=VeriSign Class 3 Secure Serv
        Validity
            Not Before: May 17 00:00:00 2013 GMT
            Not After : May 18 23:59:59 2014 GMT
        Subject: C=US, ST=Washington, L=Seattle, O=Amazon.com Inc., CN=www.amazon.com
        Subject Public Key Info:
```

```
            Public Key Algorithm: rsaEncryption
                Public-Key: (2048 bit)
                Modulus:
                    00:b7:5c:95:8f:c9:d9:68:5c:2b:64:13:30:b0:8a:
                    82:49:ff:68:ab:07:b7:50:de:fd:33:4d:a8:cb:a0:
                    78:a8:41:bb:83:55:6b:e5:41:cc:f9:36:41:33:8e:
                    71:7e:22:01:cc:ab:07:3c:d5:34:15:5f:66:88:66:
                    fe:e7:e4:dc:4e:00:37:32:79:a5:11:11:14:b3:3f:
                    1f:ec:65:ea:f9:c1:3c:cb:94:d3:ee:27:a4:46:13:
                    4e:40:a4:f5:a2:35:87:04:ea:e8:35:11:38:81:b8:
                    5a:e7:5c:95:ec:d1:e8:a2:c1:c0:12:b6:68:89:27:
                    07:3a:d2:61:d0:9f:71:0d:c1:b5:8e:e2:b5:18:0c:
                    66:ef:22:fb:d7:2f:2a:b0:46:0d:13:12:4a:15:f0:
                    8f:65:f3:9f:32:48:3c:a9:ed:2c:d0:82:a8:11:4a:
                    a1:04:81:0d:2c:8b:a1:ea:65:e5:88:b1:5f:e1:6f:
                    7c:28:a3:a2:52:97:2c:19:45:d7:b6:75:3f:c0:26:
                    b8:4a:83:03:10:c8:8c:23:cc:42:75:28:66:57:05:
                    b9:af:8b:34:60:15:20:5e:eb:f4:2c:8e:59:ec:18:
                    dc:44:dd:55:ae:5c:d7:be:01:73:71:66:ff:92:75:
                    29:9a:1f:69:f1:02:be:ed:b9:f7:04:de:e3:fd:cb:
                    e6:8f
                Exponent: 65537 (0x10001)
        X509v3 extensions:
            X509v3 Subject Alternative Name:
                DNS:uedata.amazon.com, DNS:amazon.com, DNS:amzn.com, DNS:www.amzn.com, DNS:www.amazon.com
            X509v3 Basic Constraints:
                CA:FALSE
            X509v3 Key Usage: critical
                Digital Signature, Key Encipherment
            X509v3 Extended Key Usage:
                TLS Web Server Authentication, TLS Web Client Authentication
            X509v3 Certificate Policies:
                Policy: 2.16.840.1.113733.1.7.54
                  CPS: https://www.verisign.com/cps

            X509v3 Authority Key Identifier:
                keyid:0D:44:5C:16:53:44:C1:82:7E:1D:20:AB:25:F4:01:63:D8:BE:79:A5

            X509v3 CRL Distribution Points:

                Full Name:
                  URI:http://SVRSecure-G3-crl.verisign.com/SVRSecureG3.crl

            Authority Information Access:
                OCSP - URI:http://ocsp.verisign.com
                CA Issuers - URI:http://SVRSecure-G3-aia.verisign.com/SVRSecureG3.cer

    Signature Algorithm: sha1WithRSAEncryption
         91:0c:aa:e5:3f:82:85:76:85:e2:b1:e8:51:d8:f9:d5:1a:c7:
         0e:fe:75:17:36:e2:4a:25:59:d1:01:21:72:e4:ab:65:7b:07:
         75:08:63:86:f0:cf:ec:60:15:3c:c0:27:ad:ba:a0:54:93:1b:
         9b:1d:e0:93:e1:f8:b3:84:3a:ce:04:8f:47:33:c4:35:7d:e0:
         a2:8b:ce:11:f8:4c:45:78:31:ec:85:40:a2:2e:37:af:5f:a3:
         77:d3:ad:b9:65:71:ec:f6:a9:c4:eb:06:6f:22:3d:49:d8:69:
         73:60:37:1c:fb:58:28:72:98:a2:6b:dd:22:63:47:2a:9f:a1:
         86:8a:27:8b:8f:0b:79:0b:d3:4c:be:78:91:2b:3a:cd:86:50:
         5b:a8:90:f7:2b:ad:28:8d:34:81:51:58:90:ac:ae:f7:56:e5:
         b4:d9:3a:a7:1f:f9:54:3b:cc:57:e9:8f:c6:3a:28:ed:e7:f1:
         d2:eb:66:80:4b:f3:9a:44:69:1b:08:3e:f7:76:66:2e:9c:b0:
         85:12:66:6e:d2:94:d4:47:80:c1:d7:e2:73:c2:fd:43:f6:34:
         fe:a3:2d:2a:e8:8f:e2:4b:4c:33:3d:64:76:b5:a8:d6:ac:bd:
         20:f8:b7:1a:28:4e:50:56:7f:7e:b7:07:e6:06:d2:d9:4b:84:
         ea:33:e1:77
```

Het schema van een moderne online transactie is zoals in Figuur 1. In feite beschrijft Figuur 1 enkel het begin van een transactie. Met name de creatie van een veilig kanaal van de klant naar Amazon. De klant zou zo'n veilig kanaal dan bijvoorbeeld kunnen gebruiken om zelf een publieke sleutel naar Amazon te sturen om aldus een veilig tweerichtingskanaal te creëren. Dergelijke gedragsregels vormen het onderwerp van het *cryptografisch protocol* dat de communicatie beschrijft.

Een cryptografisch protocol is een combinatie van *cryptografische primitieven* zoals symmetrische/asymmetrische encryptie, digitale handtekeningen, certificaten etc...met de bedoeling een veilige communicatie tussen 2 of meerdere partijen tot stand te brengen.



FIGUUR 1. Een moderne online transactie.

Cryptografische protocollen zijn bijzonder moeilijk te ontwerpen. Standaarden zijn:

- SSL (Secure Socket Layer) (tegenwoordig TLS (Transport Layer Security)) voor webbrowsers, email,....
- IPSEC. Dit is een veilige versie van het IP protocol.

2. Complexiteit

2.1. **De big-O-notatie.** De volgende definitie is gelijkaardig aan [Kob94, Def. I.7] maar niet helemaal identiek.

Definitie 2.1.1. Zij f, g functies van de form

$$f : S \subset \mathbb{N}^r \to \mathbb{R}$$
$$g : T \subset \mathbb{N}^r \to \mathbb{R}$$

Dan schrijven we $f = O(g)$ indien er $B, C \in \mathbb{R}^+$ bestaan zodat voor alle $(n_1, \ldots, n_r) \in \mathbb{N}^r$ met $\forall i : n_i \geq B$ er geldt $(n_1, \ldots, n_r) \in S \cap T$ en

$$|f(n_1, \ldots, n_r)| \leq C|g(n_1, \ldots, n_r)|$$

In plaats van $f = O(g)$ schrijven we soms ook

$$f(n_1, \ldots, n_r) = O(g(n_1, \ldots, n_r))$$

De verzamelingen S, T zijn meestal impliciet gegeven als de plaatsen waar f, g gedefinieerd zijn. In het vervolg zullen f, g meestal functies in één variabele zijn.

Het volgende is eenvoudig te bewijzen.

Lemma 2.1.2. *Indien f, g functies in één variabele zijn en $\lim_{n\to\infty} f(n)/g(n)$ bestaat dan $f = O(g)$.*

We geven nu enige voorbeelden van de big-O-notatie.

(1) Indien $f(n) = an^d + bn^{d-1} + \cdots \in \mathbb{R}[n]$ dan

$$f(n) = O(n^d)$$

Dit volgt uit lemma 2.1.2.

(2) Indien $f(n) = \log n$ dan[1]

$$f(n) = O(n^\epsilon)$$

[1]log zonder index betekent altijd de natuurlijke logaritme.

voor alle $\epsilon > 0$. Dit is weer een gevolg van lemma 2.1.2. De benodigde limiet $\lim_{n\to\infty} \log n/n^\epsilon$ kan bijvoorbeeld berekend worden met behulp van de regel van de l'Hopital (de limiet is nul).

(3) Zij $f(n) =$"aantal cijfers van n uitgedrukt in basis b". Dan

$$f(n) = O(\log n)$$

Dit volgt uit de meer preciese formule $f(n) = \lfloor \log_b n \rfloor + 1$.[2]

(4) Zij $T(k) = \text{Tijd}(k\text{-bit}+k\text{-bit})$[3] $\stackrel{\text{def}}{=}$"aantal operaties nodig om 2 k-bit getallen op te tellen". De exacte waarde van $T(k)$ zal natuurlijk sterk afhankelijk zijn van de gebruikte machine maar indien we het standaard algoritme bekijken

```
 1001001
 1011011
--------
10100100
```

dan is het redelijk om aan te nemen dat op gelijk welke computerarchitectuur we zullen hebben $T(k) = O(k)$.

(5) $T(k) =$"aantal operaties nodig om 2 k-bit getallen te vermenigvuldigen". De bepaling van de grootteorde van $T(k)$ is een veel subtieler probleem. Indien we weer naar het standaard algoritme kijken

```
 1001001
 1011011
 -------
 1001001
1001001
0000000
1001001
1001001
0000000
1001001
-------------
1100111110011
```

dan zien we $T(k) = O(k^2)$. Men is er echter in geslaagd veel efficientere vermenigvuldigingsalgoritmen te vinden zodat de volgende sterkere bewering ook correct is[4]

$$T(k) = O(k \cdot \log k \cdot \log\log k)$$

Hieronder zullen we enige eenvoudige voorbeelden geven van efficientere algoritmes. De zogenaamde Karatsuba en Toom-Cook vermenigvuldigingsalgoritmen.

(6) Het wordt bewezen in [Knu80, §4.4.3] dat een vermenigvuldigingsalgoritme kan omgezet worden in een delignsalgoritme dat op een constante factor

[2] $\lfloor x \rfloor$ en $\lceil x \rceil$ zijn respectievelijk het grootste geheel getal kleiner dan of gelijk aan x en het kleinste geheel getal groter dan of gelijk aan x.

[3] Bij het gebruik van de notatie "Tijd" onderstellen we dat er geen parallelisme plaats vindt in de berekening.

[4] Deze sterke begrenzing wordt bereikt door de Schönhage-Strassen vermenigvuldiging. Deze is erg technisch om te implementeren en wordt in de praktijk maar sneller dan bijvoorbeeld de Karatsuba vermenigvuldiging bij honderden decimale cijfers.

na even snel is. Dus met andere woorden: onafhankelijk van het gekozen vermenigvuldingingsalgoritme hebben we

$$\text{Tijd}(k\text{-bit}/k\text{-bit}) = O(\text{Tijd}(k\text{-bit} \times k\text{-bit}))$$

Er zijn veel delingsalgoritmen. Hetwelk in het bewijs van [Knu80, §4.4.3] gebruikt wordt is gebaseerd op de methode van Newton-Raphson.[5]

Hier is een ruwe schets. We beschouwen een deling van reele getallen. Hieruit kan eenvoudig een algoritme voor het delen van gehele getallen met rest afgeleid worden.

We herschrijven een deling u/v eerst als $u \times 1/v$. Dus het is voldoende om $1/v$ te kunnen berekenen. Door v met een factor 2^t, $t \in \mathbb{Z}$ te vermenigvuldigen[6] mogen we onderstellen $1/2 \leq v < 1$. Indien we nu Newton-Raphson toepassen op de vergelijking $1/x - v = 0$ dan bekomen we een rij $(x_k)_k$

$$x_{k+1} = 2x_k - x_k^2 v$$

die zeer snel convergeert (zoals eigen aan Newton-Raphson) naar het nulpunt $x = 1/v$.

Definitie 2.1.3. Een algoritme met k-bit getallen als input is *polynomiaal* indien de looptijd $O(k^\alpha)$ is voor $\alpha \geq 0$.

Zoals we boven gezien hebben kunnen de hoofdbewerkingen in polynomiale tijd uitgevoerd worden. Hetzelfde geldt voor het berekenen (met k-bit precisie) van klassieke wiskundinge functies zoals sin, cos, log, exp,….

Het RSA cryptosysteem is gebaseerd op het feit dat het moeilijk is om grote getallen in priemfactoren te ontbinden. Er is inderdaad geen polynomiaal algoritme bekend voor dit probleem.

Er bestaat daarentegen wel een polynomiaal algoritme om te verifiëren of een getal al dan niet priem is.

Opmerking 2.1.4. De looptijd van een algoritme wordt ook dikwijls uitgedrukt in het aantal k-bit vermenigvuldigingen/delingen dat vereist is. Op deze manier is het resultaat onafhankelijk van het gebruikte vermenigvuldigings/delingsalgoritme.

2.2. **De Karatsuba vermenigvuldiging.** De Karatsuba vermenigvuldiging is een relatief eenvoudig te programmeren algoritme dat reeds een spectaculaire efficientieverhoging geeft.

Zij u, v getallen die uit $2k$ bits bestaan. We kunnen ze schrijven als

$$u = 2^k u_1 + u_0$$
$$v = 2^k v_1 + v_0$$

waarbij u_0, u_1, v_0, v_1 getallen zijn die uit k-bits bestaan. We vinden

$$uv = 2^{2k} u_1 v_1 + 2^k (u_1 v_0 + u_0 v_1) + u_0 v_0$$

Het lijkt er dus op dat het vermenigvuldigen van 2 getallen van $2k$ bits neerkomt op het uitvoeren van 4 vermenigvuldigingen van k-bit getallen. Karatsuba's geniale

[5] De methode van Newton-Raphson berekent benaderingen voor de nulpunten van een vergelijking $f(x) = 0$ via de recursie $x_{k+1} = x_k - f(x_k)/f'(x_k)$. In goede situaties zal $(x_k)_k$ zeer snel convergeren naar een nulpunt van f.

[6] Vermenigvuldigen met een macht van 2 is heel goedkoop omdat computers met binaire getallen werken.

observatie was nu dat we dit aantal kunnen terugbrengen tot 3 vermenigvuldigingen door de volgende identiteit te gebruiken

$$u_1v_0 + u_0v_1 = (u_0 + u_1)(v_0 + v_1) - u_0v_0 - u_1v_1$$

en gebruik te maken van het feit dat we u_0v_0 en u_1v_1 toch reeds moeten berekenen voor de andere termen. Het Karatsuba algoritme bestaat er nu uit dat we k-bit getallen aanvullen tot 2^l-bit getallen met $l = \lceil \log_2 k \rceil$ en dan vervolgens de Karatsuba observatie recursief toepassen.

Als we nu schrijven

$$T(k) = \text{Tijd}(k\text{-bit} \times_{\text{Karatsuba}} k\text{-bit})$$

dan hebben we dus voor k een macht van 2

$$T(2k) \leq 3T(k) + ck \qquad k \geq 1 \tag{2.1}$$

De term ck representeert de "overhead". Deze komt voort uit het feit dat we naast vermenigvuldigingen ook nog optellingen moeten uitvoeren. En verder kunnen $u_0 + u_1$ en $v_0 + v_1$ in feite $k+1$-bit getallen zijn. Dus deze vermenigvuldiging duurt in het algemeen ook iets langer.

Lemma 2.2.1. *We nemen aan dat c groot genoeg gekozen is zodat $T(2) \leq c$. Dan geldt:*

$$T(2^l) \leq c(3^l - 2^l) \qquad \textit{voor } l \geq 1 \tag{2.2}$$

Bewijs. We gebruiken inductie. Het geval $l = 1$ is duidelijk. Onderstel dat de bewering waar is voor een zekere l. Dan geldt

$$\begin{aligned} T(2^{l+1}) &\leq 3T(2^l) + c2^l \qquad \text{(vanwege (2.1))} \\ &\leq 3c(3^l - 2^l) + c2^l \qquad \text{(inductie)} \\ &= 3c3^l - 2c2^l \\ &= c(3^{l+1} - 2^{l+1}) \quad \square \end{aligned}$$

Zij $k \geq 2$ willekeurig. We hebben nu

$$\begin{aligned} T(k) &= T(2^{\lceil \log_2 k \rceil}) \\ &\leq c3^{\lceil \log_2 k \rceil} \qquad \text{(vanwege (2.2))} \\ &\leq 3c3^{\log_2 k} \\ &= 3c2^{\log_2 3 \log_2 k} \\ &= 3ck^{\log_2 3} \end{aligned}$$

Uiteindelijk hebben we aangetoond

$$T(k) = O(k^{\log_2 3}) \tag{2.3}$$

waarbij $\log_2 3 = 1,585$.

Opmerking 2.2.2. Onze versie van het Karatsuba algoritme is uiteraard niet de meest efficiente. In de praktijk zal men Karatsuba's observatie natuurlijk niet toepassen tot op het niveau van individuele bits maar zal men op een zeker moment overschakelen naar het klassiek vermenigvuldigingsalgoritme. Dergelijke "fine tuning" is heel belangrijk voor implementaties maar ze verandert niet de exponent $\log_2 3$ in (2.3).

2.3. **De Toom-Cook vermenigvuldiging.** De Toom-Cook vermenigvuldiging is een veralgemening van de Karatsuba vermenigvuldiging waarbij we de getallen in meer dan 2 delen opsplitsen. Als voorbeeld beschrijven we het veelgebruikte Toom-3 algoritme (Karatsuba is equivalent met Toom-2).

Zij u, v getallen die uit $3k$ bits bestaan. We schrijven ze als

$$u = 2^{2k}u_2 + 2^k u_1 + u_0$$
$$v = 2^{2k}v_2 + 2^k v_1 + v_0$$

waarbij $(u_i)_i$, $(v_i)_i$ bestaan uit k bits. Definieer

$$p(x) = u_2x^2 + u_1x + u_0$$
$$q(x) = v_2x^2 + v_1x + v_0$$

$p(x)$ en $q(x)$ zijn polynomen die in het punt $x = 2^k$ de waarden u en v aannemen. Dus

$$uv = p(x)q(x) \mid_{x=2^k}$$

We moeten dus het polynoom $r(x) \stackrel{\text{def}}{=} p(x)q(x)$ efficient kunnen berekenen. Immers de substitutie $x = 2^k$ vraagt weinig rekenwerk omdat dit neerkomt op het optellen van binaire shifts van de coefficienten van $r(x)$.

Het polynoom $r(x)$ heeft 5 coefficienten. Hieruit volgt dat het voldoende is de waarde van $r(x)$ te kennen in 5 punten $\{\eta_1, \ldots, \eta_5\}$. Typisch neemt men hiervoor (bijvoorbeeld in GMP, de "GNU Multiprecision Library")

$$\{\eta_1, \ldots, \eta_5\} = \{-1, 0, 1, 2, \infty\}$$

waarbij voor $s(x) \in \mathbb{R}[x]$ we $s(\infty)$ (adhoc) definieren als de hoogstegraadscoefficient van $s(x)$.

Het berekenen van $r(\eta_i) = p(\eta_i)q(\eta_i)$ voor $i = 1, \ldots, 5$ vraagt nu 5 vermenigvuldigingen van k-bit getallen plus een zekere overhead die bestaat uit optellingen en vermenigvuldigingen van k-bit getallen met kleine getallen.

Het reconstrueren van $r(x)$ uit $r(\eta_i)_i$ komt neer op het oplossen van een 5×5 lineair stelsel met kleine coefficienten wat we ook als overhead beschouwen. We bekomen dan uiteindelijk de volgende stelling

Stelling 2.3.1. *Definieer*

$$T(k) = \text{Tijd}(k\text{-}bit \times_{Toom\text{-}3} k\text{-}bit)$$

Dan geldt

$$T(k) = O(k^{\log_3 5})$$

Bewijs. Dit wordt op een analooge manier bewezen als in het geval van de Karatsuba vermenigvuldiging. □

Merk op dat

$$\log_3 5 = 1{,}465$$

Dit is beter dan Karatsuba maar de overhead van Toom-3 is veel hoger. Dus Toom-3 zal maar effectief beter worden vanaf een zeker aantal cijfers.

3. Deelbaarheid

3.1. **De grootste gemene deler en het algoritme van Euclides.** Zij $a, b \in \mathbb{Z}$, $(a, b) \neq (0, 0)$ en zij k de *grootste gemene deler* van a en b. Notatie: $k = \text{ggd}(a, b)$. k is het kleinste strikt positief getal in het ideaal $\mathbb{Z}a + \mathbb{Z}b \subset \mathbb{Z}$. Of nog $\mathbb{Z}k = \mathbb{Z}a + \mathbb{Z}b$.

Om de expositie te vereenvoudigen onderstellen we in deze sectie dat $a, b \in \mathbb{N}$. Dit is is geen restrictie aangezien $\text{ggd}(\pm a, b) = \text{ggd}(a, \pm b) = \text{ggd}(a, b)$. De klassieke manier om k te berekenen is via de ontbinding in priemfactoren van a en b. Dus

$$a = p_1^{\alpha_1} \cdots p_u^{\alpha_u}$$
$$b = p_1^{\beta_1} \cdots p_u^{\beta_u}$$

met p_i priem en $\alpha_i \geq 0$, $\beta_i \geq 0$. Dan

$$\text{ggd}(a, b) = p_1^{\min(\alpha_1, \beta_1)} \cdots p_u^{\min(\alpha_u, \beta_u)}$$

Omdat ze berust op ontbinding in priemfactoren is deze methode echter onefficient. Een beter methode is te vinden in de Elementen van Euclides.[7] Zonder verlies aan algemeenheid mogen we onderstellen dat $0 \neq a \geq b \geq 0$. (De moderne versie van) het algoritme van Euclides maakt gebruik van de volgende twee observaties

- Indien $b = 0$ dan $\text{ggd}(a, b) = a$.
- Indien $b \mid a$ dan $\text{ggd}(a, b) = b$.
- Onderstel $b \neq 0$, $a \neq 0$, $b \nmid a$ (dus $b < a$). Schrijf $a = qb + r$ met $0 < r < b$ (dit is een deling met rest). Dan $\text{ggd}(a, b) = \text{ggd}(r, b) = \text{ggd}(b, r)$.

Omdat $b < a$ kunnen we in het laaste geval de redenering herhalen nadat we (a, b) door (b, r) vervangen hebben. Uiteindelijk zullen we in het eerste geval uitkomen.

Voorbeeld 3.1.1. $\text{ggd}(21, 15) = \text{ggd}(15, 6) = \text{ggd}(6, 3) = 3$.

Feit. *Indien a, b k-bit getallen zijn dan vereist het algoritme van Euclides $O(k)$ delingen* [Knu80, §4.5.3 Cor L].

3.2. **Het uitgebreide algoritme van Euclides.** Zijn $k = \text{ggd}(a, b)$ zoals boven. Omdat $\mathbb{Z}k = \mathbb{Z}a + \mathbb{Z}b$ bestaan er u, v zodat $k = au + bv$. De vraag is: hoe berekenen we u, v op een efficiente manier. We onderstellen weer $a \geq b$.

(1) Indien $b = 0$ dan kunnen we nemen $(u, v) = (1, 0)$.
(2) Indien $b \mid a$ dan kunnen we nemen $(u, v) = (0, 1)$.
(3) Onderstel $b \nmid a$. Stel $a = qb + r$, $0 < r < b$. We vinden dus

$$\begin{aligned} k &= au + bv \\ &= (qb + r)u + bv \\ &= b(qu + v) + ru \end{aligned}$$

We hebben $k = \text{ggd}(b, r) = \text{ggd}(a, b)$. Dus indien we u', v' kunnen bepalen zodat

$$k = bu' + rv'$$

dan volstaat het (u, v) zo te nemen dat $(qu + v, u) = (u', v')$. Oftewel $(u, v) = (v', u' - qv')$.

Met andere woorden om (u, v) te bepalen moeten we het algoritme van Euclides in omgekeerde volgorde doorlopen. Dit is gemakkelijker uitgelegd op een voorbeeld.

[7] In de versie van Euclides gaat het over het vinden van een gemeenschappelijke maat voor twee lijnstukken.

Voorbeeld 3.2.1. Beschouw

$$3 = 21u + 15v$$

We gebruiken onze berekening $3 = \text{ggd}(21, 15)$ (zie Voorbeeld 3.1.1) en werken achterwaards

$$\begin{aligned} 3 &= 6 \cdot 0 + 3 \cdot 1 \\ &= 6 \cdot 0 + (15 - 2 \cdot 6) \cdot 1 \\ &= 6 \cdot (-2) + 15 \cdot 1 \\ &= (21 - 1 \cdot 15) \cdot (-2) + 15 \cdot 1 \\ &= 21 \cdot (-2) + 15 \cdot 3 \end{aligned}$$

4. Modulaire aritmetiek

4.1. **Notaties.** We herhalen enige notaties. We schrijven

$$a \equiv b \bmod m$$

indien $m \mid a - b$. We noteren het beeld van a in $\mathbb{Z}/m\mathbb{Z}$ door $a \bmod m$ of $\bar{a}$ indien m duidelijk is uit de context. Dus

$$a \equiv b \bmod m \iff \bar{a} = \bar{b} \text{ in } \mathbb{Z}/m\mathbb{Z}$$

Soms, indien het duidelijk is uit de context, bedoelen we met $a \bmod m$ ook de rest van a na deling door m.

$\mathbb{Z}/m\mathbb{Z}$ is natuurlijk een (commutatieve) ring. We noteren met R^* de *eenhedengroep* van een ring R.

$$R^* = \{a \in R \mid \exists b \in R : ab = ba = 1\}$$

Het volgende is welbekend

Stelling 4.1.1. *(bvb* [Kob94, Prop. I.3.1]*)*

$$(\mathbb{Z}/m\mathbb{Z})^* = \{\bar{a} \mid a \in \mathbb{Z} \text{ en } \text{ggd}(a, m) = 1\}$$

4.2. **Berekening van de inverse.** Voor toepassingen is het belangrijk de inverse van een omkeerbaar element in $\mathbb{Z}/m\mathbb{Z}$ expliciet te kunnen berekenen. Dus zij $\text{ggd}(a, m) = 1$. Hoe vinden we b zodat $ab \equiv 1 \bmod m$? Als we dit expliciet uitschrijven zien we dat we b, u moeten vinden zodat

$$ab = 1 + mu$$

Omdat $\text{ggd}(a, m) = 1$ kunnen we het uitgebreide algoritme van Euclides toepassen. Met andere woorden: de inverse van $\bar{a}$ kan efficient berekend worden.

4.3. **De Chinese reststelling.** We herhalen de formulering van de Chinese reststelling (zie bvb [Kob94, Prop I.3.3]). Onderstel dat $m_1, \ldots, m_p \in \mathbb{Z}$ paarsgewijze relatief priem zijn ($\text{ggd}(m_i, m_j) = 1$ voor $i \neq j$). Definieer $M = m_1 \cdots m_p$. Dan zegt de Chinese reststelling dat de natuurlijke afbeelding

$$\mathbb{Z}/M\mathbb{Z} \to \mathbb{Z}/m_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/m_p\mathbb{Z} : a \mapsto (a \bmod m_1, \ldots, a \bmod m_p)$$

een isomorfisme van ringen is. In gewone taal: voor alle $a_1, \ldots, a_p \in \mathbb{Z}$ bestaat er een $x \in \mathbb{Z}$, uniek modulo M, zodat

$$\forall i : x \equiv a_i \bmod m_i \tag{4.1}$$

Voor toepassingen is het weer belangrijk om x expliciet te kunnen berekenen. We beweren dat het voldoende is het geval $p = 2$ te behandelen. Inderdaad onderstel dat we x_q gevonden hebben zodat $x_q \equiv a_i \bmod m_i$ voor $1 \leq i \leq q$ met $2 \leq q < p$. We hebben $\gcd(m_{q+1}, m_1 \cdots m_q) = 1$. Door gebruik te maken van het $p = 2$ geval kunnen we dan x_{q+1} berekenen zodat $x_{q+1} \equiv a_{q+1} \bmod m_{q+1}$ en $x_{q+1} \equiv x_q \bmod m_1 \cdots m_q$. De laatste congruentie impliceert $x_{q+1} \equiv x_q \equiv a_i \bmod m_i$ voor $1 \leq i \leq q$.

Dus door gebruik te maken van het $p = 2$ geval kunnen we achtereenvolgens $x_2, x_3, \ldots, x_p$ berekenen en $x = x_p$ voldoet dan aan (4.1).

Onderstel $p = 2$. Zij $m, n \in \mathbb{Z}$ zodat $\gcd(m, n) = 1$ en zij gegeven $a, b \in \mathbb{Z}$. We moeten x bepalen zodat

$$\begin{aligned} x &\equiv a \bmod m \\ x &\equiv b \bmod n \end{aligned} \tag{4.2}$$

Expliciet uitgeschreven geeft dit

$$\begin{aligned} x &= a + mu \\ x &= b + nv \end{aligned}$$

Dus met andere woorden we moeten (u, v) vinden zodat $a + mu = b + nv$, oftewel

$$b - a = mu - nv$$

We passen weer het uitgebreide Euclidische algoritme toe om u', v' te vinden zodat

$$1 = mu' - nv'$$

Nadien stellen we $u = (b - a)u'$, $v = (b - a)v'$.

4.4. **De formule van Garner.** Soms moeten we x kunnen berekenen voor verschillende waarden van (a, b). In de literatuur vinden we daarvoor een nuttige formule (oefening).

Stelling 4.4.1. *Onderstel* $0 \leq a \leq m - 1$, $0 \leq b \leq n - 1$. *Definieer*

$$x = (((a - b) \cdot (n^{-1} \bmod m)) \bmod m) \cdot n + b$$

Dan voldoet x *aan* (4.2) *en* $0 \leq x \leq mn - 1$. *De "*mod*" symbolen staan hier voor de rest na deling door* m.

Merk op dat we $n^{-1} \bmod m$ slechts 1 keer moeten berekenen.

4.5. **De Euler ϕ-functie.** We definiëren voor $m \in \mathbb{Z}$

$$\phi(m) \stackrel{\text{def}}{=} |(\mathbb{Z}/m\mathbb{Z})^*|$$

Stelling 4.5.1. *Zij* $m = p_1^{\alpha_1} \cdots p_n^{\alpha_n}$ *met* p_i *priem en* $\alpha_i > 0$. *Dan*

$$\phi(m) = \prod_{i=1}^{n} (p_i^{\alpha_i} - p_i^{\alpha_i - 1})$$

Bewijs. Vanwege de Chinese reststelling hebben we

$$\mathbb{Z}/m\mathbb{Z} \cong \mathbb{Z}/p_1^{\alpha_1}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p_n^{\alpha_n}\mathbb{Z}$$

en dus

$$(\mathbb{Z}/m\mathbb{Z})^* \cong (\mathbb{Z}/p_1^{\alpha_1}\mathbb{Z})^* \oplus \cdots \oplus (\mathbb{Z}/p_n^{\alpha_n}\mathbb{Z})^*$$

Er volgt

$$\phi(m) = \phi(p_1^{\alpha_1}) \cdots \phi(p_n^{\alpha_n})$$

We moeten dus enkel $\phi(p^\alpha)$ bepalen voor p priem en $\alpha \geq 1$.

$$\begin{aligned}(\mathbb{Z}/p^\alpha\mathbb{Z})^* &= \{\bar{a} \mid (a, p^\alpha) = 1\} \\ &= \{\bar{a} \mid p \nmid a\} \\ &= \mathbb{Z}/p^\alpha\mathbb{Z} - p\mathbb{Z}/p^\alpha\mathbb{Z}\end{aligned}$$

en dus $\phi(p^\alpha) = p^\alpha - p^{\alpha-1}$. □

Opmerking 4.5.2. Het is een leuk (en niet volledig triviaal) probleem om de volledige structuur van $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$ te bepalen (in plaats van enkel de orde). We beperken ons enkel tot het antwoord.

(1) Er bestaat een isomorfisme

$$(\mathbb{Z}/p^\alpha\mathbb{Z})^* \cong (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p^\alpha\mathbb{Z})_{1 \bmod p}$$

waarbij

$$(\mathbb{Z}/p^\alpha\mathbb{Z})_{1 \bmod p} = \{\bar{a} \mid a \equiv 1 \bmod p\}$$

(2) $(\mathbb{Z}/p\mathbb{Z})^*$ is een cyclische groep.
(3) $(\mathbb{Z}/p^\alpha\mathbb{Z})_{1 \bmod p}$ is een cyclische groep indien p oneven is.
(4) Voor $p = 2$ en $\alpha = 1$ is $(\mathbb{Z}/2^\alpha\mathbb{Z})^* = (\mathbb{Z}/2\mathbb{Z})^*$ triviaal.
(5) Onderstel $p = 2$ en $\alpha > 1$. Dan hebben we

$$(\mathbb{Z}/2^\alpha\mathbb{Z})_{1 \bmod 2} = (\mathbb{Z}/2^\alpha\mathbb{Z})^* = \{\bar{1}, -\bar{1}\} \times (\mathbb{Z}/2^\alpha\mathbb{Z})_{1 \bmod 4}$$

met

$$(\mathbb{Z}/2^\alpha\mathbb{Z})_{1 \bmod 4} = \{\bar{a} \mid a \equiv 1 \bmod 4\}$$

(6) $(\mathbb{Z}/2^\alpha\mathbb{Z})_{1 \bmod 4}$ is cyclisch.

4.6. **De kleine stelling van Fermat.**

Stelling 4.6.1. *Zij* $\gcd(a, m) = 1$. *Dan geldt*

$$a^{\phi(m)} \equiv 1 \bmod m \tag{4.3}$$

Bewijs. $(\mathbb{Z}/m\mathbb{Z})^*$ is een eindige groep en in een eindige groep G hebben we voor alle $g \in G$: $g^{|G|} = 1$. Dit bewijst het gestelde. □

Opmerking 4.6.2. $\phi(m)$ is niet de beste keuze in (4.3). Zij A een abelse groep. Dan definiëren we

$$\exp A = \min\{n > 0 \mid \forall a \in A : a^n = 1\}$$

Er geldt $\exp A \mid |A|$ met gelijkheid als en slechts als A cyclisch is.

We hebben (kgv =kleinste gemeen veelvoud)

$$\exp(A \times B) = \text{kgv}(\exp A, \exp B)$$

De beste exponent in (4.3) is $\exp(\mathbb{Z}/m\mathbb{Z})^*$. Met behulp van Opmerking 4.5.2 kunnen we $\exp(\mathbb{Z}/m\mathbb{Z})^*$ precies berekenen.

In het geval dat $m = pq$ met p, q priem en $p \neq q$ vinden we

$$\exp(\mathbb{Z}/m\mathbb{Z})^* = \text{kgv}(p-1, q-1)$$

4.7. **Modulaire exponentiatie.** In de cryptografie moeten we dikwijls

$$a^n \bmod m$$

efficient kunnen berekenen. De naieve berekeningsmethode

$$a^n = \underbrace{a \cdots a}_{n \text{ factoren}}$$

vereist $n-1$ vermenigvuldigingen en is dus niet polynomiaal in de input.

Gelukkig is er een veel betere methode: *herhaald kwadrateren*. Dit is gebaseerd op de volgende observaties

(1) Indien n even is dan $a^n = (a^2)^{n/2}$.
(2) Indien n oneven is dan $a^n = a(a^2)^{(n-1)/2}$.

In plaats van de exponent met 1 te verminderen per stap kunnen we hem dus ongeveer halveren. Meer precies kost ons dat 1 vermenigvuldiging indien n even is (het berekenen van a^2) en 2 vermenigvuldigingen indien n oneven is (het berekenen van a^2 en achteraf het vermenigvuldigen van a met $(a^2)^{(n-1)/2}$). Hieruit leiden we de volgende formule af voor het aantal benodigde vermenigvuldigingen (modulo m): schrijf n in binaire notatie $n = n_{k-1} \cdots n_0$ met $n_{k-1} = 1$. Dan is het aantal benodigde vermenigvuldigingen gelijk aan

$$\sum_{i=0}^{k-2} \epsilon(n_i)$$

waarbij

$$\epsilon(n_i) = \begin{cases} 2 & \text{indien } n_i = 1 \\ 1 & \text{indien } n_i = 0 \end{cases}$$

Merk op dat de som maar tot n_{k-2} loopt. Het berekenen van a^1 kost niets.

Opmerking 4.7.1. In de praktijk zal men het herhaald kwadrateren op een iteratieve manier uitvoeren als volgt:

(1) Stel $a' := a$, $y := 1$, $n' := n$.
(2) Indien n' even is vervang a' door $a'^2 \bmod m$ en n' door $n'/2$.
(3) Indien n' oneven is vervang a' door $a'^2 \bmod m$, y door $ya' \bmod m$ en n' door $(n'-1)/2$.
(4) Herhaal (2)(3) tot $n' = 1$. In dat geval $a^n \bmod m = ya'$.

Het algoritme is correct omdat steeds de volgende identiteit geldt

$$a^n \equiv ya'^{n'} \bmod m$$

Als het algoritme stopt bij $n' = 1$ dan vinden we inderdaad $a^n \equiv ya' \bmod m$.

Een vermenigvuldiging $\bmod m$ is hetzelfde als een gewone vermenigvuldiging en dan een deling met rest door m. Onderstel dat n een k-bit getal is en a, m l-bit getallen zijn. Dan vinden we

$$\text{Tijd}((l\text{-bit})^{k\text{-bit}} \bmod l\text{-bit}) = O(kl^2)$$

Dit kan verbeterd worden door het gebruikt van een efficienter vermenigvuldigingsalgoritme.

4.8. **De Montgomery vermenigvuldiging.** Als we het algoritme uit voorgaande sectie bekijken dan zien we dat er veel modulaire vermenigvuldigingen moeten uitgevoerd worden. We kunnen zo'n modulaire vermendigvuldiging implementeren als een vermenigvuldiging gevolgd door een deling met rest. Deze laatste operatie is in het algemeen duur en we zouden ze liefst vermijden. Er bestaat een ingenieuze truuk om dit te doen.

Onderstel dat m een oneven natuurlijk getal is. Kies $r = 2^l$ zodat $m < r$. De *Montgomery vermenigvuldiging* $\bmod m$ is

$$a * b \bmod m \stackrel{\text{def}}{=} abr^{-1} \bmod m$$

(waarbij we hier met "mod" in het rechterlid de rest na deling door m bedoelen). We zien hieronder dat indien we veel modulaire vermenigvuldigingen moeten uitvoeren met een vaste modulus, de Montgomery vermenigvuldiging efficienter is dan de gewone modulaire vermenigvuldiging.

Hoe gebruiken we de Montgomery vermenigvuldiging? Er is een ring isomorfisme

$$\mu : (\mathbb{Z}/m\mathbb{Z}, \cdot, +) \to (\mathbb{Z}/m\mathbb{Z}, *, +) : a \mapsto ra$$

Dus in plaats van in $(\mathbb{Z}/m\mathbb{Z}, \cdot, +)$ te rekenen kunnen we net zo goed in $(\mathbb{Z}/m\mathbb{Z}, *, +)$ rekenen. Hiervoor moeten we de kost inrekening brengen van het toepassen van μ op de inputs van de berekening en μ^{-1} op de outputs van de berekening.

Voorbeeld 4.8.1. Het algoritme uit de vorige sectie is een ideale kandidaat voor het gebruik van de Montgomery vermenigvuldiging. Om $a^n \bmod m$ te berekenen, berekenen we eerst $c = ra \bmod m$. Daarna berekenen we $c^{*n} \bmod m$ (de n'de macht voor de Montgomery vermenigvuldiging) met herhaald kwadrateren en op het einde vinden we $a^n \bmod m$ als $r^{-1}c^{*n} \bmod m$.

We bespreken nu de berekening van de Montgomery vermenigvuldiging. Gebruik het uigebreide algoritme van Euclides om u te bepalen zodat $u = (-m^{-1}) \bmod r$. Indien $0 \leq a, b < m$ dan $0 \leq t = ab < m^2$.

Lemma 4.8.2. *Voor $0 \leq t < m^2$ definieer*

$$w = t + (tu \bmod r) \cdot m$$

Dan is w deelbaar door r. Definieer

$$\text{res}(t) = w/r$$

Dan geldt $0 \leq \text{res}(t) < 2m$ en $\text{res}(t) \equiv tr^{-1} \bmod m$.

Bewijs. We berekenen eerst $w \bmod r$. Er geldt

$$\begin{aligned}(t + (tu \bmod r) \cdot m) \bmod r &= (t + tum) \bmod r \\ &= (t + t \cdot (-1)) \bmod r \\ &= 0 \bmod r\end{aligned}$$

Het is duidelijk dat $\text{res}(t) \geq 0$. We hebben ook

$$\begin{aligned}\text{res}(t) &< (m^2 + (r-1)m)/r \\ &= m(m + r - 1)/r \\ &< 2m\end{aligned}$$

Tenslotte berekenen we $\mathrm{res}(t) \bmod m$. Omdat $\bar{r}$ omkeerbaar is in $\mathbb{Z}/m\mathbb{Z}$ mogen we net zo goed $r\,\mathrm{res}(t) \bmod m = w \bmod m$ berekenen. We vinden $r\,\mathrm{res}(t) \equiv t \bmod m$ wat impliceert $\mathrm{res}(t) \equiv tr^{-1} \bmod m$. □

Uit het vorige lemma volgt nu

$$a * b \bmod m = \begin{cases} \mathrm{res}(ab) & \text{indien } \mathrm{res}(ab) < m \\ \mathrm{res}(ab) - m & \text{anders} \end{cases}$$

We zijn er dus in in geslaagd de delingen door m te vervangen door delingen door r. Dit laatste is veel efficienter omdat r een macht van twee is. We moeten dan gewoon de bits ofschuiven.

4.9. **Polynoomringen.** Indien k een lichaam dan hebben bovenstaande resultaten en definities voor $\mathbb{Z}$ en $\mathbb{Z}/m\mathbb{Z}$ een analogon voor $k[x]$ en $k[x]/(f)$. De reden daarvoor is dat $k[x]$ net zoals $\mathbb{Z}$ een notie van deling met rest heeft. Indien $g, h \in k[x]$ met $h \notin k$ dan kunnen we schrijven in $k[x]$

$$g = qh + r$$

met $\mathrm{graad}\, r < \mathrm{graad}\, h$.[8]

De grootste gemene deler van twee polynomen is in feite maar op een constante na bepaald. Als we toch een keuze moeten maken dan kunnen we eisen dat de ggd monisch is.

5. Eindige lichamen

5.1. **Inleiding.** Eindige lichamen[9] zijn natuurlijke veralgemeningen van de lichamen $\mathbb{F}_p \stackrel{\mathrm{def}}{=} \mathbb{Z}/p\mathbb{Z}$, p priem. Ze worden erg veel gebruikt in praktische toepassingen van wiskunde zoals codetheorie en cryptografie. De eenhedengroep van een eindig lichaam kan bijvoorbeeld gebruikt worden voor Diffie-Hellman sleuteluitwisseling (zie hieronder). Een andere toepassing is de constructie van de S-box voor AES (zie hieronder).

In deze sectie geven we overzicht van de theorie van eindige lichamen. We zullen achteraf eindige lichamen gebruiken om een bewijs van de kwadratische wederkerigheidswet te geven.

5.2. **Hoofdstellingen en voorbeelden.** In deze sectie geven we enkele belangrijke stellingen over eindige lichamen. Deze zouden voldoende moeten zijn om in de praktijk met eindige lichamen te werken.

Zij $\mathbb{F}$ een lichaam met een eindig aantal elementen. We hebben dan inclusies

$$\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} \cong \{a \cdot 1_{\mathbb{F}} \mid a \in \mathbb{Z}\} \subset \mathbb{F} \tag{5.1}$$

waarbij $p > 0$ de karakteristiek is van $\mathbb{F}$. Op deze manier verschijnt $\mathbb{F}_p$ als het *priemlichaam* van $\mathbb{F}$.

---

[8] $k[x]$ is net als $\mathbb{Z}$ een voorbeeld van een *Euclidische ring*. Met andere woorden: het is een ring met een delingsalgoritme. Zie `http://en.wikipedia.org/wiki/Euclidean_domain` voor de preciese definitie.

[9] In deze cursus is een lichaam per definitie commutatief. Voor eindige lichamen maakt dit echter geen verschil omdat een bekende stelling uit de algebra zegt dat er geen niet commutatieve eindige lichamen bestaan.

Indien L/K een lichaamsuitbreiding is dan noteren we $\dim_K L$ met $[L : K]$. Stel $r = [\mathbb{F} : \mathbb{F}_p]$. Dan is $\mathbb{F}$ een r-dimensionale vectorruimte over $\mathbb{F}_p$ en dus

$$|\mathbb{F}| = p^r$$

(dus in het bijzonder $r < \infty$). De orde van een eindig lichaam is dus een priemmacht.

De volgende stelling zegt dat het omgekeerde ook geldt

Stelling 5.2.1. *Voor elke priemmacht $q = p^r$ bestaat er op isomorfisme na een uniek lichaam $\mathbb{F}_q$ zodat $|\mathbb{F}_q| = q$.*

Bewijs. Zonder bewijs. □

Opmerking 5.2.2. Verwar $\mathbb{F}_q$ niet met $\mathbb{Z}/q\mathbb{Z}$ (indien $q \neq p$). Dit zijn beide ringen met q elementen. De ring $\mathbb{Z}/q\mathbb{Z}$ is echter geen lichaam want er zijn elementen die geen inverse hebben (bvb $\overline{p}$).

Lemma 5.2.3. *In een L lichaam van karakteristiek p geldt*

$$(a + b)^p = a^p + b^p$$

Bewijs. Dit is een direct gevolg van het binomiaaltheorema.

$$(a + b)^p = \sum_{i=0}^{p} \frac{p!}{i!(p - i)!} a^i b^{p-i}$$

We zien dat $p!/i!(p - i)!$ deelbaar is door p, behalve wanneer $i = 0, p$. Dit bewijst wat we willen. □

Dus $F_p(a) = a^p$ definieert een ringendomorfisme van L. We noemen dit het *Frobenius endomorfisme* van L. De machten van F_p^n zijn uiteraard ook endomorfismen van L. Omdat

$$F_p^n(a) = a^{p^n}$$

noteert men $F_{p^n} = F_p^n$.

Het Frobenius endomorfisme is injectief want de kern van F_p, beschouwd als een groepshomomorfisme $(L, +) \to (L, +)$, is $\{a \in L \mid a^p = 0\} = 0$, waarbij we gebruiken dat L geen nuldelers heeft. Dus indien L eindig is dan is F_p een automorfisme. We spreken in dat geval van het *Frobenius automorfisme* van L.

Met behulp van de theorie van eindige lichaamsuitbreidingen kan men aantonen

$$\mathbb{F}_q = \mathbb{F}_p[x]/(f(x))$$

waarbij $f(x)$ een irreduciebel polynoom is modulo p van graad r. Dus rekenen in $\mathbb{F}_q$ komt neer op rekenen in $\mathbb{F}_p[x]$ en dan reduceren modulo $f(x)$. De elementen van $\mathbb{F}_q$ hebben een unieke representatie als polynoom over $\mathbb{F}_p$ van graad $< r$. Deze representatie hangt uiteraard wel af van de keuze van $f(x)$.

Voorbeeld 5.2.4. Onderstel $q = 4$. Dus $p = 2$ en $r = 2$. Neem

$$f(x) = x^2 + x + 1$$

Het polynoom $f(x)$ irreduciebel want anders zou er een $a \in \mathbb{F}_2$ bestaan zodat $f(a) = 0$. Dit is niet het geval (aangezien $\mathbb{F}_2 = \{0, 1\}$ moeten we slechts 2 elementen testen).

De elementen van $\mathbb{F}_4$ worden gegeven door de polynomen van graad < 2. De polynomen van hogere graad kunnen immers gereduceerd worden door te delen door $f(x)$. We vinden dus

$$\mathbb{F}_4 = \{0, 1, x, x+1\}$$

waarbij we voor het gemak de overlijningen weggelaten hebben (we zouden bvb $\overline{1+x}$ moeten schrijven om aan te geven dat we in $\mathbb{F}_2[x]/(f(x))$ werken).

Laat ons de vermenigvuldiging in $\mathbb{F}_4$ uitschrijven. We beschouwen enkel de niet volledig triviale gevallen

$$\begin{aligned} x \cdot x &= x^2 \\ &\equiv 1 + x \bmod f(x) \end{aligned}$$

$$\begin{aligned} x(1+x) &= x + x^2 \\ &\equiv 1 \bmod f(x) \end{aligned}$$

$$\begin{aligned} (1+x)(1+x) &= 1 + x^2 \\ &\equiv x \bmod f(x) \end{aligned}$$

Dus de vermenigvuldigingstabel is

$\times$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	$1+x$	1
$1+x$	0	$1+x$	1	x

In de literatuur schrijft men dikwijls een polynoom over $\mathbb{F}_2$: $\cdots + b_1 x + b_0$ als het binair getal $\cdots b_1 b_0$. Met deze notatie hebben

$$\mathbb{F}_4 = \{0, 1, 2, 3\}$$

en de optelling en vermenigvuldiging worden gegeven door

$+$	0	1	2	3
0	0	1	2	3
1	1	0	3	2
2	2	3	0	1
3	3	2	1	0

$\times$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	3	1
3	0	3	1	2

Voorbeeld 5.2.5. Onderstel $q = 8$. Dus $p = 2$ en $r = 3$. Nu kunnen we bijvoorbeeld $f(x) = x^3 + x + 1$ nemen. Indien f reduciebel is dan heeft het een factor van graad 1 en dus een wortel. We verifiëren weer dat dit niet het geval is.

Voorbeeld 5.2.6. Onderstel $q = 16$. Dus $p = 2$ en $r = 4$. We kunnen $f(x) = x^4 + x + 1$ nemen. Het is niet meer zo duidelijk dat dit een irreduciebel polynoom is want het zou een produkt van twee factoren van graad twee kunnen zijn. Met behulp van het Maple commando

```
Factor(x^4+x+1) mod p
```

vinden we dat $f(x)$ inderdaad irreduciebel is.

Voorbeeld 5.2.7. Onderstel $q = 9$. Aangezien -1 geen kwadraat is mod 3 kunnen we nu nemen $f(x) = x^2 + 1$. Dus in $\mathbb{F}_9 = \mathbb{F}_3[\sqrt{-1}]$. Het werken in $\mathbb{F}_9$ vertoont veel gelijkenissen met het werken in het lichaam der *Gaussische rationale getallen* $\mathbb{Q}[\sqrt{-1}]$.

Voorbeeld 5.2.8. AES gebruikt het zogenaamde *AES-lichaam* $F_{256} = \mathbb{F}_2[x]/f(x)$ met $f(x) = x^8 + x^4 + x^3 + x + 1$. De *AES S-box* is de bijectie

$$S : \mathbb{F}_{256} \to \mathbb{F}_{256} : \mathrm{Aff} \circ \mathrm{Inv}$$

waarbij

$$\mathrm{Inv}(u) = \begin{cases} u^{-1} & \text{indien } u \neq 0 \\ 0 & \text{indien } u = 0 \end{cases}$$

en Aff is een affiene transformatie gedefinieerd als volgt: we identificeren $\mathbb{F}_{256}$ met de 8-dimensionale $\mathbb{F}_2$-vectorruimte $\mathbb{F}_2^8$ via

$$b_7x^7 + b_6x^6 + b_5x^4 + b_4x^4 + b_3x^3 + b_2x^2 + b_1x + b_0 \mapsto (b_7, b_6, b_5, b_4, b_3, b_2, b_1, b_0)$$

Dan

$$\mathrm{Aff}((b_i)_i) = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

(zie [AES01])

Als we Aff beschouwen als een afbeelding $\mathbb{F}_{256} \to \mathbb{F}_{256}$ dan heeft Aff de volgende algebraïsche beschrijving (zie [MR02]).

$$\mathrm{Aff}(u) = \mu + \lambda_0 u + \lambda_1 u^2 + \lambda_2 u^4 + \lambda_3 u^8 + \lambda_4 u^{16} + \lambda_5 u^{32} + \lambda_6 u^{64} + \lambda_7 u^{128}$$

waarbij

$$(\mu, \lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7) = (63{,}05{,}09{,}F9{,}25{,}F4{,}01{,}B5{,}8F) \in (\mathbb{F}_{256})^9$$

We hebben hier $\mu, (\lambda_i)_i$ gerepresenteerd als (hexadecimale) getallen zoals aangegeven in Voorbeeld 5.2.4.

Voorbeeld 5.2.9. (Optioneel) Een wel heel erg vreemde constructie van de lichamen $\mathbb{F}_{2^{2^n}}$ wordt gegeven door de "nimbers" uit de combinatorische speltheorie [Con76]. In deze constructie worden de elementen van $\mathbb{F}_{2^{2^n}}$ gegeven door de elementen $0, \ldots, 2^{2^n} - 1$ van $\mathbb{N}$. De optelling is de bitsgewijze XOR van de binaire schrijfwijze van de getallen.

Om de vermenigvuldiging te beschrijven definiëren we eerst voor $S \subset \mathbb{N}$

$$\mathrm{mex}(S) = \min(\mathbb{N} - S)$$

(mex ="minimal excluded"). Dus bvb $\mathrm{mex}(\emptyset) = 0$. De formule voor de vermenigvuldiging van nimbers is recursief:

$$ab = \mathrm{mex}\{a'b + ab' + a'b' \mid 0 \leq a' < a, 0 \leq b' < b\}$$

waarbij "+" staat voor de optelling die we in de eerste paragraaf gedefinieerd hebben.

In feite definiëren nimberoptelling/nimberprodukt gewoon een lichaamstruktuur op $\mathbb{N}$!! De lichamen $\mathbb{F}_{2^{2^n}}$ verschijnen op een natuurlijke manier als deellichamen van $\mathbb{N}$.

Je kan eventueel een programmaatje schrijven in bijvoorbeeld Maple dat het produkt van twee nimbers uitrekent. Te verifiëring: $16 \cdot 16 = 24$.

Een niet recursieve manier om het produkt te beschrijven is als volgt:

(1) Het nimberprodukt van twee of meerdere verschillende Fermat machten[10] is het gewone produkt in $\mathbb{N}$.
(2) Het nimberkwadraat van een Fermat macht x is $3x/2$ (uitgewerkt voor de gewone vermenigvuldiging/deling op $\mathbb{N}$).

Waarom geeft dit voldoende informatie om het produkt van twee willekeurige nimbers uit te rekenen?

Om te rekenen in het lichaam $\mathbb{F}_q$ moeten we een irreduciebel polynoom vinden van graad r. Hoe kunnen we dit efficiënt doen? Stel

$$a_{np} = |\{\text{monische irreduciebele polynomen met graad } n \text{ in } \mathbb{F}_p[x]\}|$$

Dan volgens [Knu80, Ex. 4.6.2.4] (en de bijbehorende oplossing) hebben we

$$\lim_{p\to\infty} \frac{a_{np}}{p^n} = \frac{1}{n}$$

Dus indien we gewoon random monische polynomen proberen dan zullen we (voor grote p) gemiddeld n keer moeten proberen voor we een irreduciebel polynoom te pakken hebben.

Blijft nog de vraag hoe we kunnen nagaan of een polynoom irreduciebel is? Hiervoor bestaat een eenvoudig criterium. Zie Stelling 5.3.2 hieronder.

Er is echter meer. Er bestaat een polynomiaal algoritme om een polynoom over $\mathbb{F}_p[x]$ te ontbinden in factoren: het algoritme van Berlekamp [Knu80, §4.6.2]. Dit is dus een fundamenteel verschil met de situatie over $\mathbb{Z}$.

Hier is een andere fundamentele stelling over eindige lichamen.

Stelling 5.2.10. *De eenhedengroep in $\mathbb{F}_q^*$ is cyclisch.*

Bewijs. We moeten bewijzen

$$\exp \mathbb{F}_q^* = |\mathbb{F}_q^*| = q - 1$$

Onderstel dat $e = \exp \mathbb{F}_q^* < q - 1$. Dan geldt voor alle $a \in \mathbb{F}_q^* : a^e = 1$. Dus het polynoom $x^e - 1$ heeft $q-1 > e$ wortels. Dit is onmogelijk aangezien $\mathbb{F}_q$ een lichaam is. □

[10] Een Fermat macht is een getal van de vorm 2^{2^n}.

Voorbeeld 5.2.11. Beschouw het lichaam $\mathbb{F}_8 = \mathbb{F}_2[x]/(x^3+x+1)$. Dan hebben we

$$\begin{aligned} x \cdot 1 &= x \\ x \cdot x &= x^2 \\ x \cdot x^2 &= x^3 \\ &\equiv x+1 \\ x \cdot (x+1) &= x^2 + x \\ x \cdot (x^2+x) &= x^3 + x^2 \\ &\equiv x^2 + x + 1 \\ x \cdot (x^2+x+1) &= x^3 + x^2 + x \\ &\equiv x^2 + 1 \\ x \cdot (x^2+1) &= x^3 + x \\ &\equiv 1 \end{aligned}$$

Dus $\bar{x}$ is een generator van $\mathbb{F}_8^*$.

We brengen nog een paar feiten uit de Galois theorie in herinnering.

Herinnering 5.2.12. *Onderstel dat L/K een eindige lichaamsuitbreiding is (dwz $K \subset L$ en $\dim_K L < \infty$).*

1. *Elk K-lineair ringhomomorfisme $L \to L$ is een automorfisme.*
2. *Zij $\mathrm{Aut}(L/K)$ de verzameling an K-lineaire automorfismen van L. Dan geldt $|\mathrm{Aut}(L/K)| \leq [L:K]$.*
3. *Indien $|\mathrm{Aut}(L/K)| = [L:K]$ dan noemen we L/K* Galois. *In dat geval definiëren we de* Galoisgroep *van L/K als*

$$\mathrm{Gal}(L/K) = \mathrm{Aut}(L/K)$$

Indien σ een automorfisme is van $\mathbb{F}_q$ dan is σ de identiteit op het priemlichaam $\mathbb{F}_p$. Dit ziet men als volgt (cfr (5.1))

$$\sigma(a \cdot 1_{\mathbb{F}}) = a \cdot \sigma(1_{\mathbb{F}}) = a \cdot 1_{\mathbb{F}}$$

Bijgevolg

$$\mathrm{Aut}(\mathbb{F}_q/\mathbb{F}_p) = \mathrm{Aut}(\mathbb{F}_q)$$

Het Frobenius endomorfisme F_p van $\mathbb{F}_q$ is dus een element van $\mathrm{Aut}(\mathbb{F}_q/\mathbb{F}_p)$. We spreken nu uiteraard over het *Frobenius automorfisme* van $\mathbb{F}_q$.

Stelling 5.2.13. *$\mathbb{F}_q/\mathbb{F}_p$ is een Galois extensie en $\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$ is een cyclische groep van orde r voortgebracht door F_p.*

Bewijs. Zonder bewijs. □

Nu we weten dat $\mathbb{F}_q/\mathbb{F}_p$ Galois is kunnen we de *Galois correspondentie* gebruiken.

Herinnering 5.2.14. *Zij L/K een (eindige) Galois uitbreiding van lichamen met $G = \mathrm{Gal}(L/K)$. Dan bestaat er een orde omkerende bijectie*

$$\{\textit{deelgroepen } H \subset G\} \to \{\textit{tussenlichamen } K \subset L' \subset L\} : H \mapsto L^H$$

Verder geldt $|H| = [L : L^H]$.

We passen de Galois correspondentie toe op de uitbreiding $\mathbb{F}_q/\mathbb{F}_p$. Hierbij merken we op dat *elk* deellichaam $\mathbb{F} \subset \mathbb{F}_q$ het priemlichaam $\mathbb{F}_p$ bevat.

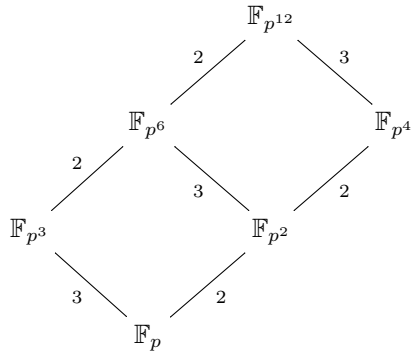
Voorbeeld 5.2.15. Zij $q = p^4$. Dan $G = \mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p) = \{\mathrm{id}, F_p, F_p^2, F_p^3\}$. Naast $\{\mathrm{id}\}$ en G zelf heeft G één enkele niet triviale deelgroep: $\{\mathrm{id}, F_p^2\}$. Dus $\mathbb{F}_q$ heeft één enkel niet triviaal deellichaam: $\mathbb{F} = \mathbb{F}^{F_p^2}$ en $[\mathbb{F}_q : \mathbb{F}] = 2$. We besluiten dat $|\mathbb{F}| = p^2$.

Het algemene geval is net zoals in dit voorbeeld.

Stelling 5.2.16. *De verzameling deellichamen van $\mathbb{F}_q$ wordt gegeven door $\{\mathbb{F}_q^{F_{p^{r'}}} \mid r' \text{ deelt } r\}$, waarbij geldt $|\mathbb{F}_q^{F_{q'}}| = q'$. Dus alle deellichamen van $\mathbb{F}_q$ hebben een verschillend aantal elementen.*

Bewijs. Zij $C_r = \mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$ de cyclische groep met r elementen. We gebruiken lemma 5.2.18 hieronder met $\sigma = F_p$. C_r bevat voor elke $r' \mid r$ een unieke deelgroep $C_{r/r'}$ van de orde r/r' en dit is de cyclische groep voortgebracht door $F_p^{r'} = F_{p^{r'}}$. Zij $\mathbb{F} = \mathbb{F}_q^{C_{r/r'}} = \mathbb{F}_q^{F_{p^{r'}}}$. Dan geldt $[\mathbb{F}_q : \mathbb{F}] = |C_{r/r'}| = r/r'$ en dus $|\mathbb{F}| = p^{r'}$. □

Voorbeeld 5.2.17. Zij $r = 12$. De deellichamen van $\mathbb{F}_{p^{12}}$ worden gegeven door de delers van 12. We vinden de volgende struktuur



Lemma 5.2.18. *Zij C_r de cyclische groep met r elementen en zij σ een generator voor C_r. Dan geldt*

1. *Elke deelgroep G van C_r is cyclisch met generator $\sigma^{r'}$ voor zekere $r' \mid r$.*
2. *We hebben $|G| = r/r'$. Dus r' (en dus ook G) is uniek bepaald door $|G|$.*

Bewijs. Zonder verlies aan algemeenheid mogen we onderstellen dat $C_r = \mathbb{Z}/r\mathbb{Z}$ en $\sigma = \bar{1}$.

(1) Zij $G \subset \mathbb{Z}/r\mathbb{Z}$ een deelgroep. Definieer

$$r' = \min\{\delta \mid 0 < \delta \leq r-1, \bar{\delta} \in G\}$$

We beweren dat $r' \mid r$. Indien niet definieer dan $0 < s < r'$ via $r = qr' + s$. Dan hebben we $\bar{s} = -\bar{r}'q \in G$, in contradictie met het minimaal zijn van r'.

Zij $\bar{\tau}$ een willekeurig element van G met $0 \leq \tau \leq r-1$. We beweren $r' \mid \tau$. Indien niet definieer $0 < \tau' < s$ via $\tau = q'r' + \tau'$. Omdat $\bar{\tau}' \in G$ hebben we weer een contradictie met het minimaal zijn van r'.

(2) Dit is duidelijk. □

5.3. **Irreducibiliteit van polynomen in $\mathbb{F}_p[x]$.**

Lemma 5.3.1. *Onderstel dat $f(x) \in \mathbb{F}_p[x]$ een irreduciebel polynoom van graad r is. Dan geldt*

$$f(x) \mid x^{p^s} - x \iff r \mid s$$

Bewijs. Definieer $\mathbb{F}_q = \mathbb{F}_p[x]/(f(x))$. Dan geldt

$$\begin{aligned} f \mid x^{p^s} - x &\iff \bar{x}^{p^s} = \bar{x} \text{ in } \mathbb{F}_q \\ &\iff \bar{x} \in \mathbb{F}_q^{F_p^s} \\ &\iff \mathbb{F}_q = \mathbb{F}_q^{F_p^s} \\ &\iff F_p^s = \mathrm{id} \\ &\iff r \mid s \end{aligned}$$

□

Stelling 5.3.2. *Zij $f(x) \in \mathbb{F}_p[x]$ een polynoom van graad r. Dan is f irreduciebel als en slechts als de volgende condities gelden*

(1) $f \mid x^{p^r} - x$.
(2) $\gcd(f, x^{p^{r/q}} - x) = 1$ *voor alle priemdelers $q \mid r$.*

Bewijs. Onderstel eerst dat f irreduciebel is. Dan volgt (1) onmiddelijk uit lemma 5.3.1. Onderstel dat (2) vals is voor zekere q. Dan moet noodzakelijk gelden $f \mid x^{p^{r/q}} - x$. Dan volgt weer uit het lemma dat geldt $r \mid r/q$, wat natuurlijk absurd is.

Onderstel nu omgekeerd dat (1) en (2) gelden en dat f niet irreduciebel is. Dus $f = gh$ met g een irreduciebel polynoom van graad $s < r$. Uit (1) halen we dan

$$g \mid f \mid x^{p^r} - x$$

en dus vanwege lemma 5.3.1: $s \mid r$. Er bestaat dus een priemgetal q zodat $s \mid r/q$ en nogmaals vanwege lemma 5.3.1 vinden we $g \mid x^{p^{r/q}} - x$. Hieruit halen we tenslotte $g \mid \gcd(f, x^{p^{r/q}} - x)$ wat een contradictie met (2) is. We zijn dus klaar. □

Voorbeeld 5.3.3. We beschouwen nogmaals het AES polynoom $f(x) \in \mathbb{F}_2[x]$:

$$f(x) = x^8 + x^4 + x^3 + x + 1$$

Met herhaald kwadrateren berekenen we x^{2^r} modulo f. We vinden

r	$x^{2^r} \bmod f$
0	x
1	x^2
2	x^4
3	$x^4 + x^3 + x + 1$
4	$x^6 + x^4 + x^3 + x^2 + x$
5	$x^7 + x^6 + x^5 + x^2$
6	$x^6 + x^3 + x^2 + 1$
7	$x^7 + x^6 + x^5 + x^4 + x^3 + x$
8	x

Er geldt dus $f \mid x^{2^8} - x$. Conditie (1) van Stelling 5.3.2 is dus voldaan. Om conditie (2) te verifieren berekenen we (de enige mogelijkheid is $q = 2$)

$$\mathrm{ggd}(f, x^{2^{8/2}} - x) = \mathrm{ggd}(x^8 + x^4 + x^3 + x + 1, x^6 + x^4 + x^3 + x^2 + x - x) = 1$$

De gegeven f is dus inderdaad irreduciebel.

5.4. **Ontbinden in factoren in $\mathbb{F}_q[x]$.** Zij $f(x) \in \mathbb{F}_q[x]$, $r = \mathrm{graad}\, f > 0$. We weten dat $f(x)$ een ontbinding in irreduciebele factoren heeft

$$f(x) = f_1(x) \cdots f_n(x)$$

Hoe vinden we $f_i(x)$? Het is uiteraard voldoende een algoritme te hebben dat voor reduciebele $f(x)$ een "niet triviale factor" geeft van $f(x)$, m.a.w.

$$h(x) \mid f(x)$$

met $0 < \mathrm{graad}\, h < \mathrm{graad}\, f$. We hebben dat

$$f(x) = h(x) \cdot f(x)/h(x)$$

en we herhalen het algoritme met $h(x)$, $f(x)/h(x)$.

We beschrijven nu het zogenaamde algoritme van Berlekamp. Als eerste stap berekenen we $f'(x)$. Indien $f'(x) = 0$ dan

$$f(x) = \sum_i a_i x^{pi}$$

en dus

$$f(x) = \left(\sum_i a_i^{1/p} x^i \right)^p$$

waarbij $(-)^{1/p}$ de inverse van het Frobenius automorfisme is. We vinden dus een niet triviale factor van $f(x)$.

Onderstel $f'(x) \neq 0$. Als volgende stap berekenen we

$$h(x) = \mathrm{ggd}(f(x), f'(x))$$

Omdat

$$\mathrm{graad}\, f'(x) < \mathrm{graad}\, f(x)$$

geldt

$$\mathrm{graad}\, h(x) < \mathrm{graad}\, f(x)$$

Dus indien $h(x)$ niet constant is dan hebben we weer een niet triviale factor. Onderstel dus dat $h(x)$ constant is. Dit kan enkel indien $f(x)$ niet deelbaar is door een niet constant kwadraat (we zeggen $f(x)$ is "kwadraatvrij"). Inderdaad indien

$$u(x)^2 \mid f(x)$$

dan

$$u(x) \mid \mathrm{ggd}(f(x), f'(x))$$

Dus we mogen onderstellen dat $f(x)$ kwadraatvrij is. Zij

$$f(x) = f_1(x) \cdots f_n(x)$$

de ontbinding van $f(x)$ in irreduciebele factoren. Vanwege de Chinese reststelling is er een ringisomorfisme

$$\mathbb{F}_q[x]/(f(x)) \to \mathbb{F}_q[x]/(f_1(x)) \oplus \cdots \oplus \mathbb{F}_q[x]/(f_n(x)) : \overline{g(x)} \mapsto (\overline{g(x)}, \ldots, \overline{g(x)})$$

Lemma 5.4.1. *Zij* $g(x) \in \mathbb{F}_q[x]$ *zodat*

$$g(x)^q \cong g(x) \bmod f(x)$$

Dan bestaan er unieke $s_i \in \mathbb{F}_q$ *voor alle* i

$$g(x) \cong s_i \bmod f_i(x)$$

Bewijs. Indien

$$g(x)^q \cong g(x) \bmod f(x) \tag{5.2}$$

dan hebben we voor alle i

$$g(x)^q \cong g(x) \bmod f_i(x) \tag{5.3}$$

Echter $\mathbb{F} = \mathbb{F}_q[x]/(f_i(x))$ is een lichaam en $\mathbb{F}_q$ is een deellichaam van $\mathbb{F}$. Vanwege de eigenschappen van eindige lichamen geldt dus

$$\mathbb{F}_q = (\mathbb{F}_q[x]/(f_i(x)))^{F_q}$$

Uit (5.3) halen we dus dat $g(x)$ een unieke representant $s_i \in \mathbb{F}_q$ heeft modulo $f_i(x)$. □

Zij

$$B = \{g(x) \in \mathbb{F}_q[x] \mid g(x)^q \cong g(x) \bmod f(x), \operatorname{graad} g < r\}$$

Merk op dat B een eindigdimensionale $\mathbb{F}_q$-deelvectorruimte van $\mathbb{F}_q[x]$ is zodat $\mathbb{F}_q \subset B$. Dus B is minstens 1-dimensionaal.

We hebben de afbeelding

$$B \mapsto \mathbb{F}_q^n : g(x) \mapsto (s_i)_i$$

Deze afbeelding is $\mathbb{F}_q$-linear en met de Chinese reststelling ziet men dat ze injectief en surjectief is.

We bekomen dan reeds het volgende resultaat.

Stelling 5.4.2. *Het aantal irreduciebele factoren van* $f(x)$ *is gelijk aan* $\dim B$*. In het bijzonder is* $f(x)$ *reduciebel als en slechts als* $\dim B > 1$*.*

We willen echter een expliciete factor van $f(x)$ vinden. Dit doen we met de volgende stelling

Stelling 5.4.3. *Zij* $g(x) \in B$ *niet constant (dit kan enkel indien* $\dim B > 1$*). Dan bestaat er een* $s \in \mathbb{F}_q$ *zodat*

$$h_s(x) = \operatorname{ggd}(f(x), g(x) - s)$$

een niet triviale factor van $f(x)$ *is.*

Bewijs. Aangezien

$$\operatorname{graad}(g(x) - s) = \operatorname{graad} g(x) < r$$

geldt

$$\operatorname{graad} h_s(x) < \operatorname{graad} f(x)$$

Het is dus voldoende aan te tonen dat er een $h_s(x)$ bestaat die niet constant is.

Vanwege Lemma 5.4.1 bestaat er een $s_1 \in \mathbb{F}_q$ zodat

$$g(x) \cong s_1 \bmod f_1(x)$$

Dus

$$f_1(x) \mid h_{s_1}(x)$$

Dus $h_{s_1}(x)$ is niet constant. □

Om deze stelling te kunnen gebruiken moeten we de elementen van B kunnen vinden. Dit is echter een gewoon lineair algebra probleem.

Zij

$$C = \{g(x) \in \mathbb{F}_q[x] \mid \text{graad}(g) < r\}$$

Dit is een $\mathbb{F}_q$-vectorruimte met basis $1, \ldots, x^{r-1}$.

De vectorruimte B is de kern van de $\mathbb{F}_q$-lineaire afbeelding

$$\Phi : C \to C : g(x) \mapsto (g(x)^q - g(x)) \text{ “rest” } f(x)$$

We kunnen Φ in matrix vorm omzetten door de beelden $\Phi(x^i)$ te berekenen. Dit geeft ons een $r \times r$ matrix[11] waarvan we de nulruimte kunnen berekenen met de gebruikelijke Gaussische eliminatie.

[11] De i-de kolom bestaat uit de coefficienten van $\Phi(x^i)$.

6. De kwadratische wederkerigheidswet

6.1. **Kwadratische residuen.** Zij p een oneven priemgetal en $a \in \mathbb{Z}$ zodat $p \nmid a$. We zeggen dat a (of $\bar{a}$) een *kwadratisch residue*[12] is modulo p indien a een kwadraat is modulo p. I.e. indien er een $b \in \mathbb{Z}$ bestaat zodat $b^2 \equiv a \bmod p$. Indien dit niet het geval is dan noemen we a een *kwadratisch non-residue*.

Voorbeeld 6.1.1. Onderstel $p = 11$. Gebruikmakende van het feit dat $\bar{a}^2 = (-\bar{a})^2$ zien we dat de kwadraten modulo p zijn: $1^2 = 1$, $2^2 = 4$, $3^2 = 9$, $4^2 = 16 \equiv 5$, $5^2 = 25 \equiv 3$, i.e. de verzameling $\{\bar{1}, \bar{3}, \bar{4}, \bar{5}, \bar{9}\}$. Dit zijn de kwadratische residuen modulo 11. De kwadratische non-residuen zijn de overblijvende elementen: $\{\bar{2}, \bar{6}, \bar{7}, \bar{8}, \overline{10}\}$.

Na het uitrekenen van meerdere voorbeelden overtuigen we ons dat er precies evenveel kwadratische residuen als non-residuen zijn. Dit kan als volgt verklaard worden. Beschouw de map

$$(-)^2 : (\mathbb{Z}/p\mathbb{Z})^* \to (\mathbb{Z}/p\mathbb{Z})^* : \bar{b} \mapsto \bar{b}^2$$

Omdat $\mathbb{Z}/p\mathbb{Z}$ een lichaam is van karakteristiek $p \neq 2$ hebben de vezels van deze afbeelding kardinaliteit 2 (i.e. ze zijn van de vorm $\{\pm\bar{b}\}$). De kwadratische residuen vormen het beeld van deze afbeelding. Dit beeld heeft dus kardinaliteit $(p-1)/2$.

6.2. **Het Legendre symbool.** Het *Legendre symbool* $\left(\frac{a}{p}\right)$ wordt als volgt gedefinieerd

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{indien } p \mid a \\ 1 & \text{indien } a \text{ een kwadratisch residue is modulo } p \text{ (ihb } p \nmid a) \\ -1 & \text{indien } a \text{ geen kwadratisch residue is modulo } p \text{ (ihb } p \nmid a) \end{cases}$$

Soms schrijven we ook $\left(\frac{\bar{a}}{p}\right)$ om te benadrukken dat het Legendre symbool enkel afhangt van $a \bmod p$.

Stelling 6.2.1. *Er geldt de volgende formule*

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \bmod p \tag{6.1}$$

Bewijs. We gebruiken dat $(\mathbb{Z}/p\mathbb{Z})^*$ cyclisch is (zie Stelling 5.2.10).

Indien $p \mid a$ dan is de formule (6.1) triviaal. Onderstel dus $p \nmid a$. Zij $\bar{g}$ een generator van $(\mathbb{Z}/p\mathbb{Z})^*$. Dan $\bar{a} = \bar{g}^x$.

Wanneer is $\bar{g}^x$ een kwadraat? Dit zal het geval zijn indien er een $\bar{g}^y$ bestaat zodat $\bar{g}^{2y} = \bar{g}^x$, oftewel $2y \equiv x \bmod p-1$. Aangezien p oneven is kan dit enkel indien x even is. Samenvattend

$$\left(\frac{\bar{g}^x}{p}\right) = \begin{cases} 1 & x \text{ is even} \\ -1 & x \text{ is oneven} \end{cases}$$

Zij b de rechterzijde van (6.1). We vinden $\bar{b} = \bar{g}^{x(p-1)/2}$. Indien $x(p-1)/2$ deelbaar is door $p-1$ dan $\bar{b} = \bar{1}$. Indien $x(p-1)/2$ niet deelbaar is door $p-1$ dan $\bar{b} \neq \bar{1}$ maar $\bar{b}^2 = \bar{1}$. Omdat $\mathbb{Z}/p\mathbb{Z}$ een lichaam is volgt $\bar{b} = -\bar{1}$.

$x(p-1)/2$ is deelbaar door $p-1$ enkel en alleen als x even is. Dit bewijst het gestelde. □

[12] De nederlandse term is "kwadraatrest".

Opmerking 6.2.2. Deze stelling toont aan $\left(\frac{a}{p}\right)$ in $O(\log p)$ vermenigvuldigingen $\bmod p$ kan berekend worden. Dit is polynomiaal en dus goed. Hieronder tonen we echter aan dat met behulp van de kwadratische reciprociteitswet de berekening nog veel efficiënter kan.

Gevolg 6.2.3. *Het Legendre symbool is multiplicatief in a. I.e.*

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$$

Bewijs. Dit volgt uit het feit dat het rechterlid van (6.1) triviaal multiplicatief is. □

Gevolg 6.2.4. *We hebben*

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$$

Bewijs. Dit volgt onmiddelijk uit (6.1). □

6.3. **De kwadratische wederkerigheidswetten.** De kwadratische wederkerigheidswetten zijn in zekere zin de eerste "moeilijke" stellingen uit de getaltheorie. Geen van de bekende bewijzen is echt transparant. We zullen later een bewijs geven gebaseerd op de theorie van eindige lichamen. In deze sectie beperken we ons tot de formulering.

Stelling 6.3.1. *(De kwadratische wederkerigheidswet) Zij p, q oneven priemgetallen. Dan geldt*

$$\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}\left(\frac{p}{q}\right)$$

Hierbij merken we op dat

$$(-1)^{(p-1)(q-1)/4} = \begin{cases} -1 & \text{indien } p \equiv q \equiv 3 \bmod 4 \\ 1 & \text{anders} \end{cases}$$

De kwadratische wederkerigheidswet heeft twee zogenaamde *supplementen* die de waarde van het Legendre symbool in speciale gevallen geven. Het eerste supplement is Gevolg 6.2.4. Het tweede supplement is de volgende stelling.

Stelling 6.3.2. *Zij p een oneven priemgetal. Dan geldt*

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}$$

waarbij we hebben

$$(-1)^{(p^2-1)/8} = \begin{cases} 1 & \text{indien } p \equiv 1, 7 \bmod 8 \\ -1 & \text{indien } p \equiv 3, 5 \bmod 8 \end{cases}$$

Voorbeeld 6.3.3. We willen beslissen of 7411 een kwadratisch residue is modulo 9283 (dit zijn priemgetallen). Met andere woorden we moeten $\left(\frac{7411}{9283}\right)$ berekenen.

Zowel 7411 als 9283 zijn $\equiv 3$ modulo 4. Dus we hebben

$$\begin{aligned}\left(\frac{7411}{9283}\right) &= -\left(\frac{9283}{7411}\right)\\ &= -\left(\frac{1872}{7411}\right)\\ &= -\left(\frac{2}{7411}\right)^4\left(\frac{3}{7411}\right)^2\left(\frac{13}{7411}\right)\\ &= -\left(\frac{13}{7411}\right)\end{aligned}$$

We hebben gebruikt dat $9283 - 7411 = 1872 = 2^4 3^2 13$. In de laatste lijn hebben we gebruikt dat we even machten mogen weglaten (want $(-1)^2 = 1$).

We gebruiken opnieuw de kwadratische reciprociteitswet ($13 \equiv 1 \bmod 4$). Verder gebruiken we ook $7411 = 570 \cdot 13 + 1$. Dus

$$\begin{aligned}-\left(\frac{13}{7411}\right) &= -\left(\frac{7411}{13}\right)\\ &= -\left(\frac{1}{13}\right)\\ &= -1\end{aligned}$$

Dus 7411 is geen kwadratisch residue modulo 9283.

6.4. **Het Jacobi symbool.** De kwadratische wederkerigheidswet zoals we die in §6.3 geformuleerd hebben laat ons toe om Legendre symbolen uit te rekenen (zoals we gezien hebben in Voorbeeld 6.3.3). Deze methode is echter nog niet bevredigend omdat ze gebaseerd is op ontbinding in factoren.

We hebben een nieuw concept nodig: het *Jacobi symbool*. Zij n een positief oneven getal en zij $a \in \mathbb{Z}$. Indien $n = 1$ dan definiëren we:

$$\left(\frac{a}{1}\right) = 1$$

Indien $n \neq 1$ dan ontbinden we n in priemfactoren: $n = p_1 \cdots p_r$. Dan definiëren we

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)\cdots\left(\frac{a}{p_r}\right)$$

Waarschuwing 6.4.1. *Indien $\left(\frac{a}{n}\right) = 1$ dan betekent dit niet dat a een kwadraat is modulo n!*

Onderstel bijvoorbeeld $n = pq$, $p \neq q$ en a is een kwadratisch non-residue modulo p en q. Dan geldt

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p}\right)\left(\frac{a}{q}\right) = (-1)(-1) = 1$$

Aangezien a echter geen kwadraat is modulo p zal het zeker geen kwadraat zijn modulo $n = pq$.

Een concreet voorbeeld wordt gegeven door $a = 2$ en $n = 15$.

We geven enkele rekenregels voor het Jacobi symbool.

Stelling 6.4.2. (1) *$\left(\frac{a}{n}\right)$ hangt enkel af van de waarde van $a \bmod n$.*

(2) *Het Jacobi symbool is multiplicatief in de zin dat*

$$\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right)\left(\frac{b}{n}\right)$$

Bewijs. (1) $\left(\frac{a}{n}\right)$ is een produkt van termen van de vorm $\left(\frac{a}{p}\right)$ met p een priemdeler van n. Indien we bij a een veelvoud van n tellen dan verandert zo'n term niet (want $p \mid n$). Het produkt verandert dus ook niet.

(2) Dit volgt uit het feit dat de individuele termen $\left(\frac{a}{p}\right)$ multiplicatief zijn in a. □

Stelling 6.4.3. *Er geldt*

$$\left(\frac{-1}{n}\right) = (-1)^{(n-1)/2}$$

Bewijs. Dit is een beetje vervelend om direct te bewijzen. We zullen volledige inductie gebruiken. Als $n = 1$ dan zijn beide zijden van de gelijkheid 1. De stelling is dus waar. Als n een priemgetal is dan is de stelling waar vanwege (6.1). Onderstel dat n geen priemgetal is en $\neq 1$. Dan geldt $n = n_1 n_2$ met $n_1 < n$, $n_2 < n$. We mogen aannemen dat de stelling waar is voor n_1, n_2. Dus

$$\left(\frac{-1}{n}\right) = \left(\frac{-1}{n_1}\right)\left(\frac{-1}{n_2}\right) = (-1)^{(n_1-1)/2+(n_2-1)/2}$$

We moeten aantonen dat dit gelijk is aan $(-1)^{(n_1 n_2-1)/2}$. Dus we moeten aantonen

$$((n_1 - 1) + (n_2 - 1))/2 \equiv (n_1 n_2 - 1)/2 \bmod 2$$

of te wel

$$n_1 + n_2 - 2 \equiv n_1 n_2 - 1 \bmod 4$$

wat op zijn beurt weer equivalent is met

$$(n_1 - 1)(n_2 - 1) \equiv 0 \bmod 4$$

Dit laatste is waar omdat bij definitie n_1, n_2 oneven zijn en dus zijn $n_1 - 1$, $n_2 - 1$ deelbaar door 2. □

Stelling 6.4.4. *Er geldt*

$$\left(\frac{2}{n}\right) = (-1)^{(n^2-1)/8}$$

Bewijs. We gebruiken dezelfde inductiemethode als in de vorige stelling (steunende op Stelling 6.3.2 in het priemgeval). We moeten dan aantonen

$$((n_1^2 - 1) + (n_2^2 - 1))/8 \equiv (n_1^2 n_2^2 - 1)/8 \bmod 2$$

oftewel

$$(n_1^2 - 1) + (n_2^2 - 1) \equiv (n_1^2 n_2^2 - 1) \bmod 16$$

wat equivalent is met

$$(n_1^2 - 1)(n_2^2 - 1) \equiv 0 \bmod 16$$

Dit is waar omdat n_1, n_2 oneven zijn en dus zijn $n_1^2 - 1$, $n_2^2 - 1$ deelbaar door 4. □

Stelling 6.4.5. *Onderstel dat a, n allebei oneven en positief zijn. Dan geldt*

$$\left(\frac{a}{n}\right) = (-1)^{(a-1)(n-1)/4}\left(\frac{n}{a}\right)$$

Bewijs. We gebruiken inductie op n en a simultaan. Indien $n = 1$ of $a = 1$ dan zijn beide leden van de gelijkheid 1. Indien n en a priemen zijn dan gebruiken we Stelling 6.3.1.

Onderstel nu bijvoorbeeld dat $n \neq 1$ en $n = n_1 n_2$ met $n_1 < n$, $n_2 < n$. We mogen onderstellen dat de stelling geldt voor (a, n_1) en (a, n_2).

Dit leidt tot de verificatie van de volgende identiteit.

$$(a-1)(n_1-1)/4 + (a-1)(n_2-1)/4 \equiv (a-1)(n_1 n_2 - 1)/4 \bmod 2$$

of te wel

$$(a-1)(n_1-1) + (a-1)(n_2-1) \equiv (a-1)(n_1 n_2 - 1) \bmod 8$$

wat equivalent is met

$$(a-1)(n_1-1)(n_2-1) \equiv 0 \bmod 8$$

Dit laatste is waar omdat $a-1$, n_1-1, n_2-1 alle even zijn.

Het geval waarbij a niet priem is, is volledig analoog. □

Voorbeeld 6.4.6. We geven nu een alternatieve berekening van $\left(\frac{7411}{9283}\right)$ die niet ontbinding in factoren gebruikt (we splitsen soms factoren 2 af).

$$\begin{aligned}
\left(\frac{7411}{9283}\right) &= -\left(\frac{9283}{7411}\right) \\
&= -\left(\frac{1872}{7411}\right) \\
&= -\left(\frac{2}{7411}\right)^4 \left(\frac{117}{7411}\right) \\
&= -\left(\frac{7411}{117}\right) \\
&= -\left(\frac{40}{117}\right) \\
&= -\left(\frac{2}{117}\right)^3 \left(\frac{5}{117}\right) \\
&= \left(\frac{117}{5}\right) \\
&= \left(\frac{2}{5}\right) \\
&= -1
\end{aligned}$$

7. Een bewijs van de kwadratische wederkerigheidswet

7.1. **De eigenlijke kwadratische wederkerigheidswet.** We zullen eerst Stelling 6.3.1 bewijzen. Dwz we onderstellen dat p, q oneven priemgetallen en we bewijzen

$$\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4} \left(\frac{p}{q}\right)$$

We onderstellen $p \neq q$ want indien $p = q$ dan zijn beide leden van de gelijkheid 0 en dus is de stelling triviaal.

We maken gebruik van de zogenaamde Gaußsommen.

Herinnering 7.1.1. *Zij $n \in \mathbb{N}_0$. Indien K een lichaam is en $\xi \in K$ dan noemen we ξ een n-*de eenheidswortel *indien geldt $\xi^n = 1$. Indien bovendien geldt $\xi^{n'} \neq 1$ voor $0 < n' < n$ dan noemen we ξ een* primitieve n'de eenheidswortel.

Voorbeeld 7.1.2. Indien $K = \mathbb{C}$ dan is $e^{2\pi i/n}$ een primitieve n-de eenheidswortel.

Zij $\xi \in K$ een primitieve q-de eenheidswortel. Dan wordt de corresponderende Gaußsom gedefinieerd als

$$G = \sum_{j=1}^{q-1} \left(\frac{j}{q}\right) \xi^j$$

Lemma 7.1.3. *Er geldt*

$$G^2 = \left(\frac{-1}{q}\right) q$$

en dus in het bijzonder $G \neq 0$.

Bewijs. Het bewijs is gebaseerd op het uitschrijven van G^2 als een dubbele som.

$$G = \sum_{j=1}^{q-1} \left(\frac{j}{q}\right) \xi^j$$

$$G = \sum_{k=1}^{q-1} \left(\frac{k^{-1}}{q}\right) \xi^k \tag{7.1}$$

In de laatste definitie is k^{-1} een getal zodanig dat $kk^{-1} \equiv 1 \bmod q$. Aangezien het Legendre symbool multiplicatief is hebben we

$$\left(\frac{k^{-1}}{q}\right) = \left(\frac{k}{q}\right)^{-1} = \left(\frac{k}{q}\right)$$

waarbij de laatste gelijkheid volgt uit $\left(\frac{k}{q}\right) \in \{\pm 1\}$. Dus de uitdrukking voor G in (7.1) is inderdaad correct.

We berekenen nu

$$G^2 = \sum_{j,k=1}^{q-1} \left(\frac{jk^{-1}}{q}\right) \xi^{j+k}$$

We stellen nu $j' \equiv jk^{-1} \bmod q$ en dus $j \equiv j'k \bmod q$. Dan geldt

$$G^2 = \sum_{j',k=1}^{q-1} \left(\frac{j'}{q}\right) \xi^{k(j'+1)}$$

$$= \sum_{j'=1}^{q-1} \left(\frac{j'}{q}\right) \sum_{k=1}^{q-1} \xi^{k(j'+1)} \tag{7.2}$$

Vanwege lemma 7.1.4 hieronder geldt

$$\sum_{k=0}^{q-1} \xi^{k(j'+1)} = \begin{cases} q & \text{indien } \xi^{j'+1} = 1 \\ 0 & \text{anders} \end{cases}$$

en dus, aangezien ξ een primitieve q-de eenheidswortel is

$$\sum_{k=0}^{q-1} \xi^{k(j'+1)} = \begin{cases} q & \text{indien } j' \equiv -1 \bmod q \\ 0 & \text{anders} \end{cases} \tag{7.3}$$

We kunnen dit resultaat niet onmiddelijk toepassen omdat de binnenste som in (7.2) vanaf 1 start en niet vanaf 0. We hebben echter

$$\sum_{j'=1}^{q-1}\left(\frac{j'}{q}\right)\sum_{k=1}^{q-1}\xi^{k(j'+1)}=\sum_{j'=1}^{q-1}\left(\frac{j'}{q}\right)\sum_{k=0}^{q-1}\xi^{k(j'+1)}-\sum_{j'=1}^{q-1}\left(\frac{j'}{q}\right)=\sum_{j'=1}^{q-1}\left(\frac{j'}{q}\right)\sum_{k=0}^{q-1}\xi^{k(j'+1)}$$

waarbij de laatste gelijkheid volgt uit het feit dat er evenveel kwadratische residuen als non-residuen zijn. Als we dit combineren met (7.3) vinden we

$$\sum_{j'=1}^{q-1}\left(\frac{j'}{q}\right)\sum_{k=1}^{q-1}\xi^{k(j'+1)}=\left(\frac{-1}{q}\right)q$$

wat het bewijs beëindigt. □

We hebben het volgende lemma gebruikt

Lemma 7.1.4. *Onderstel dat η een element van een lichaam is zodat $\eta^q = 1$. Dan geldt*

$$\sum_{k=0}^{q-1}\eta^k=\begin{cases}q & \textit{indien } \eta=1\\ 0 & \textit{indien } \eta\neq 1\end{cases}$$

Bewijs. Het geval $\eta = 1$ is triviaal. Onderstel dus $\eta \neq 1$. Stel

$$S=\sum_{k=0}^{q-1}\eta^k$$

Dan

$$\eta S=\sum_{k=0}^{q-1}\eta^{k+1}=S$$

en dus $(1-\eta)S = 0$. Indien $\eta \neq 1$ hebben we dus inderdaad $S = 0$ (we werken in een lichaam en we kunnen dus delen door $(1-\eta)$). □

We gaan nu onderstellen dat ons lichaam een eindig lichaam is. Meer precies zij $\mathbb{F} = \mathbb{F}_{p^{q-1}}$. Volgens Stelling 5.2.10 is $\mathbb{F}^*$ cyclisch van orde $p^{q-1} - 1$. Vanwege de kleine stelling van Fermat $q \mid |\mathbb{F}^*|$ en dus vanwege lemma 5.2.18 is er een unieke cyclische deelgroep $C_q \subset \mathbb{F}^*$ met q elementen.

Zij ξ een generator van C_q. Dan geldt $\xi^q = 1$ maar $\xi^{q'} \neq 1$ voor $1 \leq q' < q$. Dus ξ is een primitieve q'de eenheidswortel in $\mathbb{F}$.

Zij nu G de bijbehorend Gaußsom.

Lemma 7.1.5. *Er geldt*

$$G^p=\left(\frac{p}{q}\right)G$$

Bewijs. Dit is de volgende berekening.

$$\begin{aligned} G^p &= \sum_{j=1}^{q-1} \left(\frac{j}{q}\right)^p \xi^{jp} \\ &= \sum_{j=1}^{q-1} \left(\frac{j}{q}\right) \xi^{jp} \\ &= \sum_{j=1}^{q-1} \left(\frac{jp^{-1}}{q}\right) \xi^{j} \\ &= \left(\frac{p^{-1}}{q}\right) G \\ &= \left(\frac{p}{q}\right) G \quad \square \end{aligned}$$

Het bewijs van de kwadratische reciprociteitswet volgt nu uit de volgende berekening.

$$\begin{aligned} G^p &= (G^2)^{(p-1)/2} G \\ &= \left(\frac{-1}{q}\right)^{\frac{p-1}{2}} q^{(p-1)/2} G \end{aligned}$$

Als we dit combineren met lemma 7.1.5 en gebruikmaken van het feit that $G \neq 0$ (lemma 7.1.3) vinden we

$$\left(\frac{p}{q}\right) = \left(\frac{-1}{q}\right)^{\frac{p-1}{2}} q^{(p-1)/2} \qquad \text{in } \mathbb{F}$$

Beide zijden van deze gelijkheid zijn elementen van het priemlichaam. We kunnen ze dus interpreteren als een congruentie modulo p.

$$\left(\frac{p}{q}\right) \equiv \left(\frac{-1}{q}\right)^{\frac{p-1}{2}} q^{(p-1)/2} \qquad \bmod p$$

Als we nu substitueren

$$\left(\frac{-1}{q}\right) = (-1)^{(q-1)/2}$$

$$q^{(p-1)/2} \equiv \left(\frac{q}{p}\right) \qquad \bmod p$$

(Stelling 6.2.1 en Gevolg 6.2.4) dan vinden we

$$\left(\frac{p}{q}\right) \equiv (-1)^{(p-1)(q-1)/4} \left(\frac{q}{p}\right) \qquad \bmod p$$

Beide zijden van deze congruentie zijn echter ± 1 en dus is het een gelijkheid (immers p is oneven en dus ≥ 3). We zijn dus klaar.

7.2. **Het tweede supplement.** We moeten nu bewijzen

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}$$

voor p een oneven priemgetal.

Definieer voor n een oneven getal: $\chi(n) = (-1)^{(n^2-1)/8}$. In het bewijs van Stelling 6.4.4 hebben we aangetoond dat geldt $\chi(n_1 n_2) = \chi(n_1)\chi(n_2)$.

Onderstel dat K een lichaam is dat een primitieve 8'ste eenheidswortel ξ bevat. We definiëren dan weer de bijbehorende Gaußsom

$$G = \sum_{j=1,3,5,7} \chi(j)\xi^j$$

Lemma 7.2.1. *We hebben*

$$G^2 = 8$$

en dus in het bijzonder $G \neq 0$.

Bewijs. Het bewijs is een eenvoudige berekening

$$\begin{aligned} G^2 &= (\xi - \xi^3 - \xi^5 + \xi^7)^2 \\ &= \xi^2 + \xi^6 + \xi^{10} + \xi^{14} - 2\xi^4 - 2\xi^6 + 2\xi^8 + 2\xi^8 - 2\xi^{10} - 2\xi^{12} \\ &= \xi^2 - \xi^2 + \xi^2 - \xi^2 + 2 + 2\xi^2 + 2 + 2 - 2\xi^2 + 2 \\ &= 8 \end{aligned}$$

In de derde lijn hebben we gebruikt dat $\xi^4 = -1$. Immers uit $(\xi^4)^2 = 1$ halen we $\xi^4 = \pm 1$. $\xi^4 = 1$ is onmogelijk aangezien ξ een primitieve achtste eenheidswortel is. □

Omdat p oneven is geldt dat $p^2 - 1$ deelbaar door 8 is. Dus zoals in §7.1 tonen we aan dat $\mathbb{F} = \mathbb{F}_{p^2}$ een primitieve achtste eenheidswortel ξ bevat.

Lemma 7.2.2. *Er geldt*

$$G^p = \chi(p)G$$

Bewijs. Dit bewijs is volledig analoog aan het bewijs van lemma 7.1.5. □

We berekenen nu weer

$$\begin{aligned} G^p &= (G^2)^{(p-1)/2}G \\ &= 8^{(p-1)/2}G \\ &= (2^{(p-1)/2})^3 G \end{aligned}$$

We vinden dus

$$(2^{(p-1)/2})^3 \equiv \chi(p) \qquad \bmod p$$

Als we nu invullen

$$2^{(p-1)/2} \equiv \left(\frac{2}{p}\right) \qquad \bmod p$$

verkrijgen we het gestelde.

8. Modulaire viekantswortels

8.1. **Inleiding.** Zij p een oneven priemgetal en zij $a \in \mathbb{Z}$ zodanig dat $\left(\frac{a}{p}\right) = 1$. Hoe vinden we nu b zodat $b^2 \equiv a \bmod p$? Als een dergelijke b bestaat dan noemen we dit een *vierkantswortel modulo* p en soms noteren we zo'n vierkantswortel met $\sqrt{a} \bmod p$. Merk op dat de notatie $\sqrt{a} \bmod p$ dubbelzinnig is want indien b een vierkantswortel is dan geldt dit ook voor $-b$. Als we *echt* een keuze moeten maken kunnen we bijvoorbeeld eisen $\sqrt{a} \in [1, p/2[$

8.2. **Het eenvoudige geval.** Onderstel $p \equiv 3 \bmod 4$. Dan levert de volgende berekening ons een vierkantswortel van $a \bmod p$.

$$\begin{aligned}(a^{(p+1)/4})^2 &= a^{(p+1)/2} \\ &= a^{(p-1)/2} \cdot a \\ &\equiv \left(\frac{a}{p}\right) \cdot a \bmod p \qquad (6.1) \\ &= a\end{aligned}$$

Op deze manier vinden we dus een vierkantswortel van $a \bmod p$ in $O(\log p)$ modulaire vermenigvuldigingen $\bmod p$.

De reden waarom deze truuk werkt is de volgende. Zij a een element van een abelse groep G zodat $a^n = 1$. We kunnen nu zoeken naar een vierkantswortel van a *binnen de cyclische groep voortgebracht door* a. Als we stellen $b = a^k$ dan moeten we hebben $a^{2k} = a$. Het is dus voldoende dat $n \mid 2k-1$. Dit is uiteraard onmogelijk als n even is. Indien n oneven is dan kunnen we kiezen[13] $k = (n+1)/2$.

Omdat

$$a^{(p-1)/2} \equiv \left(\frac{a}{p}\right) = 1 \qquad \bmod p \tag{8.1}$$

konden we in ons geval $n = (p-1)/2$ kiezen wat een oneven getal is gezien de onderstelling $p \equiv 3 \bmod 4$.

Merk op dat om af te leiden dat $a^{(p-1)/2} \equiv 1 \bmod p$ we niet noodzakelijk beroep moeten doen op het Legendre symbool. Er was immers gegeven dat a een kwadraat $\bmod p$ is. Dus $a \equiv c^2 \bmod p$ en dus

$$a^{(p-1)/2} \equiv c^{p-1} \equiv 1 \qquad \bmod p$$

Dit is een nuttige observering indien we deze truuk willen toepassen op hogere machtswortels.

8.3. **Het moeilijke geval.** Het geval $p \equiv 1 \bmod 4$ is subtieler. $(p-1)/2$ is nu een even getal en we kunnen dus geen wortel van a vinden in de cyclische groep voortgebracht door a. We hebben nu $p - 1 = 2^s n$ met $s \geq 2$ en n oneven. Geinspireerd door het $p \equiv 3 \bmod 4$ geval proberen we

$$r \equiv a^{(n+1)/2} \qquad \bmod p$$

We beweren dat r redelijk dicht tegen $\sqrt{a}$ komt. Immers

$$ar^{-2} \equiv a \cdot a^{-(n+1)} \equiv a^{-n} \qquad \bmod p$$

[13]Dit is wat we vinden indien we de congruentie $2k \equiv 1 \bmod n$ oplossen met behulp van het veralgemeende algoritme van Euclides. In dit geval zien we echter op het zicht wat de oplossing is.

en dus $(ar^{-2})^{2^{s-1}} \equiv a^{-2^{s-1}n} \equiv \left(\frac{a}{p}\right) = 1 \bmod p$.

Als we nu a vervangen door ar^{-2} dan mogen we dus onderstellen dat $a^{2^{s-1}} \equiv 1 \bmod p$. Definieer

$$C_{2^s} = \{c \in (\mathbb{Z}/p\mathbb{Z})^* \mid c^{2^s} \equiv 1 \bmod p\}$$

Omdat $\mathbb{Z}/p\mathbb{Z}$ een lichaam is heeft C_{2^s} tenhoogste 2^s elementen. In dit geval heeft C_{2^s} precies 2^s elementen gegeven door

$$g^{j(p-1)/2^s} \qquad j = 0, \ldots, 2^s - 1$$

waarbij g een generator is van de cyclische groep $(\mathbb{Z}/p\mathbb{Z})^*$.

Indien $c^2 \equiv a \bmod p$ dan geldt $c^{2^s} \equiv 1 \bmod p$. Zowel a als de vierkantswortels van a liggen dus in de cyclische groep C_{2^s}.

Om een vierkantswortel van a te kunnen bepalen hebben we een generator van C_{2^s} nodig. We zouden hiervoor $g^{(p-1)/2^s}$ kunnen nemen maar dan moeten we een generator g van $(\mathbb{Z}/p\mathbb{Z})^*$ vinden.

In dit geval kunnen we dit omzeilen. Zij k' een kwadratisch non-residue modulo p en definieer $k \equiv (k')^n \bmod p$.

We hebben $k^{2^s} \equiv 1 \bmod p$. Aan de ander kant hebben we ook

$$-1 \equiv \left(\frac{k'}{p}\right) \equiv (k')^{(p-1)/2} \equiv k^{2^{s-1}} \qquad \bmod p$$

Met andere woorden k is een generator voor C_{2^s}.

We gaan nu een vierkantswortel van a zoeken van de vorm k^x. Als we schrijven

$$x = x_0 + x_1 2 + \cdots + x_{s-2}2^{s-2} + x_{s-1}2^{s-1} \qquad \text{met } x_i \in \{0, 1\}$$

dan moeten dus oplossen

$$a \equiv k^{x_0 2 + x_1 2^2 + \cdots + x_{s-2}2^{s-1}} \qquad \bmod p$$

We vinden dan

$$a^{2^{s-2}} \equiv k^{x_0 2^{s-1}} \qquad \bmod p$$

en dus

$$x_0 = \begin{cases} 0 & \text{indien } a^{2^{s-2}} \equiv 1 \bmod p \\ 1 & \text{indien } a^{2^{s-2}} \equiv -1 \bmod p \end{cases}$$

Definieer nu $a_1 = ak^{-2x_0} = k^{x_1 2^2 + \cdots}$. Er geldt nu $a_1^{2^{s-2}} \equiv 1 \bmod p$ en verder

$$a_1^{2^{s-3}} \equiv k^{x_1 2^{s-1}} \bmod p$$

en dus

$$x_1 = \begin{cases} 0 & \text{indien } a_1^{2^{s-3}} \equiv 1 \bmod p \\ 1 & \text{indien } a_1^{2^{s-3}} \equiv -1 \bmod p \end{cases}$$

Als we deze procedure voortzetten dan vinden we alle x_i. Merk op dat er geen conditie is op x_{s-1}. Dus a heeft twee vierkantswortels.

Blijft de vraag: hoe vinden we een kwadratisch non-residue k'? In de praktijk is dit triviaal. Kies k' gewoon random. We hebben 50% percent kans op succes. Na gemiddeld twee pogingen vinden we op deze manier een kwadratisch non-residue. Dit is een zogenaamd polynomiaal *probabilistisch algoritme*.

Merkwaardig genoeg is er geen polynomiaal, niet probabilistisch, *niet conditioneel* algoritme bekend om een kwadratisch non-residue te vinden. Met "niet conditioneel" bedoelen we dat de complexiteit niet afhangt van het waar zijn van

een onbewezen vermoeden (in dit geval de veralgemeende Riemann hypothese, zie [Wed01, Thm 6.35]).

8.4. **Vierkantswortels modulo samengestelde n.** Zij $n = pq$ het produkt van twee verschillende oneven priemgetallen. Onderstel dat $a \equiv b^2$ een kwadraat is modulo n. Dan is a ook een kwadraat modulo p en modulo q. Dus heeft a twee vierkantswortels $\pm b_p$ modulo p en twee vierkantswortels $\pm b_q$ modulo q. Uit de Chinese reststelling volgt dat a dus 4 vierkantswortels modulo n heeft. Twee hiervan worden gegeven door $\pm b$. Zij b' een van de ander vierkantswortels. We hebben $n \mid (b - b')(b + b')$ maar n deelt niet $b - b'$ of $b + b'$. Dus $\gcd(n, b - b')$ moet gelijk zijn aan p of q.

De volgende stelling is de basis van het Rabin cryptosysteem (zie verder).

Stelling 8.4.1. *Zij $\mathcal{S}$ de deelverzameling van $\mathbb{N}$ bestaande uit produkten van twee verschillende priemgetallen. Voor $n \in \mathbb{N}$ zij* $\mathrm{sqr}(n) = \{b^2 \bmod n \mid \gcd(b, n) = 1\}$. *Onderstel dat er een algoritme bestaat dat voor $a \in \mathrm{sqr}(n)$ een wortel modulo n berekent. Dan bestaat er een (probabilistisch) algoritme met dezelfde complexiteit dat $n \in \mathcal{S}$ ontbindt in factoren.*

Bewijs. Zij gegeven $n = pq \in \mathcal{S}$. Kies een random $b \in (\mathbb{Z}/n\mathbb{Z})^*$ en bereken $a \equiv b^2 \bmod n$. Dus $a \in \mathrm{sqr}(n)$. Als we nu ons algoritme toepassen op (a, n) dan vinden we een vierkantswortel b' van a modulo n. Omdat b random was en ons algoritme enkel a kende (en niet b) is er 50% kans dat $b \not\equiv \pm b'$. Dan vormt $\gcd(b - b', n)$ een niet triviale factor van n vanwege de voorgaande discusie. □

8.5. **Het Rabin cryptosysteem.** Het Rabin crypto systeem is een publieke sleutel systeem waarvan de parameters twee verschillende priemgetallen p, q zijn. De publieke sleutel is $n = pq$ en de private sleutel is (p, q). De verzamelingen plaintexts $\mathcal{P}$ en ciphertexts $\mathcal{C}$ zijn gelijk en worden gegeven door $(\mathbb{Z}/n\mathbb{Z})^*$. De encryptiefunctie is $E(m) = m^2 \bmod n$ en de decryptiefunctie is $D(c) = \sqrt{c} \bmod n$. Een praktisch nadeel is duidelijk dat decryptie niet uniek is (er zijn 4 vierkantswortels modulo n). Daarom zal Alice voldoende redundantie in haar boodschappen moeten inbouwen zodat Bob een geldige plaintext kan herkennen. In [MvOV97, p293] stelt men voor om de laatste 64 bit van de plaintext te herhalen.[14].

Zoals we gezien hebben in §8.4 kunnen we $D(c)$ efficient berekenen indien we de private sleutel (p, q) kennen. Stelling 8.4.1 toont aan dat het omgekeerde ook geldt. Vandaar dat men zegt dat het Rabin cryptosysteem een "bewijsbaar veilig" cryptosysteem is. Men bedoelt hier mee dat efficiente decryptie equivalent is met efficiente ontbinding in factoren. Dit is niet geweten voor het (veel meer gebruikte) RSA cryptosysteem.

Opmerking 8.5.1. Om de decryptie te vereenvoudigen is het best om $p, q \equiv 3 \bmod 4$ te kiezen. Zie §8.2.

9. RSA

9.1. **Basisprincipe.** De parameters van het RSA cryptosysteem bestaan uit een 5-tuple (n, p, q, d, e) waarbij p, q priemgetallen zijn, $n = pq$ en $de \equiv 1 \bmod t$ met $t = \mathrm{kgv}(p - 1, q - 1)$. Voor de eenvoud onderstellen we hieronder dat p, q oneven zijn.

[14]En dus is de verzameling legale cyphertexts in feite kleiner dan $(\mathbb{Z}/n\mathbb{Z})^*$!

De publieke sleutel is het koppel (n, e) en de private sleutel is het tripel (p, q, d). De verzameling plaintexts en cyphertexts zijn beide gelijk aan $\mathbb{Z}/n\mathbb{Z}$. De encryptie/decryptie functies zijn respectievelijk van de vorm

$$E(m) = m^e \mod n$$
$$D(c) = c^d \mod n$$

We zien dus dat we voor decryptie slechts (n, d) nodig hebben en niet (p, q, d). Het is echter gebruikelijk aan te nemen dat de publieke sleutel uit de private sleutel kan afgeleid worden. Vandaar deze keuze.

Lemma 9.1.1. *Er geldt* $D \circ E = \mathrm{id}$, $E \circ D = \mathrm{id}$.

Bewijs. Beide beweringen zijn equivalent met $m^{de} \equiv m \bmod n$ voor alle $m \in \mathbb{Z}/n\mathbb{Z}$. Vanwege de Chinese reststelling is het voldoende om aan te tonen dat geldt: $m^{de} \equiv m \bmod p$ en $m^{de} \equiv m \bmod q$. We bewijzen de eerste bewering. Het bewijs van de tweede bewering is uiteraard identiek.

Onderstel eerst dat $p \mid m$. In dat geval is zowel m als m^{de} deelbaar door p. Dus

$$0 \equiv m \equiv m^{de} \mod p$$

Onderstel nu $p \nmid m$. Omdat $p-1 \mid t$ hebben we $p-1 \mid de-1$. Schrijf $de-1 = a(p-1)$. Vanwege de kleine stelling van Fermat geldt $m^{p-1} \equiv 1 \bmod p$ en dus

$$\begin{aligned} m^{de} &= m \cdot (m^{p-1})^a \\ &\equiv m \mod p \quad \square \end{aligned} \tag{9.1}$$

9.2. **Toepassing van het Jacobi symbool op RSA.** De encryptiefunctie voor RSA is

$$c \equiv m^e \bmod n$$

met $n = pq$. We vinden dan

$$\left(\frac{c}{n}\right) = \left(\frac{m^e}{n}\right) = \left(\frac{m}{n}\right)^e = \left(\frac{m}{n}\right)$$

waarbij we in de laatste ongelijkheid gebruiken dat e oneven is.

Met andere woorden RSA "lekt" informatie over de waarde van $\left(\frac{m}{n}\right)$. Dit geeft echter zo weinig informatie dat het in de praktijk geen probleem is.

9.3. **RSA en ontbinden in factoren.** In deze sectie bewijzen we dat indien we d efficient kunnen berekenen uit (n, e) dat we dan ook n efficient kunnen ontbinden in factoren. Dit betekent nog niet dat het kraken van RSA equivalent is met efficient ontbinden in factoren. Dit laatste is niet geweten.

Procedure. Schrijf

$$de - 1 = 2^w x \qquad x \text{ oneven}$$

Kies $a \in (\mathbb{Z}/n\mathbb{Z})^*$ random.

Bewering 1. Er is tenminste 25% kans[15] dat er een $1 \leq i \leq w$ bestaat zodat

$$\begin{aligned} a^{2^i x} &\equiv 1 \mod n \\ a^{2^{i-1} x} &\not\equiv \pm 1 \mod n \end{aligned} \tag{9.2}$$

[15]Dit is wat we zullen bewijzen. Een meer preciese analyse toont aan dat de kans $\geq 50\%$ is.

Als we een goede i gevonden hebben dan geldt $n \mid a^{2^i x} - 1$. Er geldt echter ook $a^{2^i x} - 1 = (a^{2^{i-1} x} - 1)(a^{2^{i-1} x} + 1)$ en $n \nmid (a^{2^{i-1} x} \pm 1)$. Dus

$$\mathrm{ggd}(a^{2^{i-1} x} - 1, n)$$

is een niet triviale factor van n.

Als onze a niet voldoet aan (9.2) dan kiezen we gewoon een nieuwe. Op deze manier vinden we snel een factor van n.

Bewijs van Bewering 1. Schrijf

$$\begin{aligned} p - 1 &= 2^u p' \quad p' \text{ oneven} \\ q - 1 &= 2^v q' \quad q' \text{ oneven} \end{aligned}$$

en onderstel dat $u \geq v$. De analyse in het geval $u \leq v$ is analoog.

Uit de Chinese reststelling halen we

$$(\mathbb{Z}/n\mathbb{Z})^* \cong (\mathbb{Z}/p\mathbb{Z})^* \oplus (\mathbb{Z}/q\mathbb{Z})^*$$

en het kiezen van een random element a in $(\mathbb{Z}/n\mathbb{Z})^*$ is natuurlijk hetzelfde als het kiezen van random elementen (a_1, a_2) in $(\mathbb{Z}/p\mathbb{Z})^* \oplus (\mathbb{Z}/q\mathbb{Z})^*$. In het bijzonder zijn de waarden van $\left(\frac{a}{p}\right) = \left(\frac{a_1}{p}\right)$ en $\left(\frac{a}{q}\right) = \left(\frac{a_2}{q}\right)$ onafhankelijk.

De kans is 50% dat geldt $\left(\frac{a_1}{p}\right) = -1$ en eveneens 50% dat geldt $\left(\frac{a_2}{q}\right) = 1$. Dus we vinden dat de kans 25% is dat geldt

$$\left(\frac{a}{p}\right) = -1 \qquad \left(\frac{a}{q}\right) = +1 \tag{9.3}$$

Om het bewijs van Bewering 1 te voltooien bewijzen we het volgende.

Bewering 2. Er geldt $u \leq w$. Indien (9.3) geldt dan

$$\begin{aligned} a^{2^u x} &\equiv 1 \qquad \bmod n \\ a^{2^{u-1} x} &\not\equiv \pm 1 \qquad \bmod n \end{aligned} \tag{9.4}$$

Bewijs Bewering 2. We hebben

$$t = \mathrm{kgv}(p-1, q-1) = 2^{\max(u,v)} \,\mathrm{kgv}(p', q') \mid (de - 1) = 2^w x$$

en dus $u \leq w$ en $p', q' \mid x$.

We berekenen nu

$$\begin{aligned} a^{2^{u-1} x} &= (a^{2^{u-1} p'})^{x/p'} \\ &= (a^{2^{u-1} p'})^{\text{oneven}} \\ &= (a^{(p-1)/2})^{\text{oneven}} \\ &\equiv \left(\frac{a}{p}\right)^{\text{oneven}} \\ &= -1 \qquad \bmod p \end{aligned}$$

$$\begin{aligned} a^{2^{u-1}x} &= (a^{2^{v-1}q'})^{2^{u-v}x/q'} \\ &= (a^{2^{v-1}q'})^{\text{iets}} \\ &= (a^{(q-1)/2})^{\text{iets}} \\ &\equiv \left(\frac{a}{q}\right)^{\text{iets}} \\ &= 1 \qquad \bmod q \end{aligned}$$

(9.4) volgt nu uit de Chinese reststelling. □

9.4. **Aandachtspunt: side channel attacks.** De krachtigste attacks op cryptografische systemen zijn side channel attacks, dwz attacks die gebruik maken van informatie extern aan de mathematische definitie van het systeem. RSA is hier bijzonder kwetsbaar voor.

We hebben bijvoorbeeld gezien dat de tijd nodig om

$$D(c) = c^d \bmod n$$

te berekenen sterkt afhangt van het aantal enen in de binaire schrijfwijze van d. Nu is het kennen van het aantal enen van d niet voldoende om d eenvoudig te kunnen bepalen. Men heeft echter aangetoond dat door het variëren van c men informatie kan krijgen over de individuele bits in d. Zie [Koc96].

Een mogelijke oplossing voor dit probleem is "blinding". In plaats van $c^d \bmod n$ uit te rekenen kiest Bob eerste een random element $u \in (\mathbb{Z}/n\mathbb{Z})^*$ en berekent

$$D(c) = (u^e c)^d u^{-1}$$

De input $u^e c$ van de modulaire exponentiatie is nu niet meer onder controle van de attacker.

9.5. **Aandachtspunt: bit security.** Onderstel dat we in de volgende situatie zijn:

(1) De attacker kan de ciphertext c vrij kiezen.
(2) Het systeem lekt een beetje informatie over de gedecrypte boodschap (bvb door timing informatie).

Een dergelijke setting is dikwijls voldoende om de attacker toe te laten een willekeurige boodschap te decrypten. We hebben hiervan een voorbeeld gezien in de inleidende lessen: de zogenaamde Bleichenbacher attack op SSL.

Hier geven we een eenvoudiger voorbeeld. Onderstel dat het systeem de volgende informatie lekt voor $0 \leq m \leq n-1$

$$\text{half}(m) = \begin{cases} 0 & 0 \leq m < n/2 \\ 1 & n/2 \leq m < n \end{cases}$$

We geven nu aan hoe de attacker dit kan gebruiken om een willekeurige boodschap te decrypten.

Procedure. De te decrypten boodschap is c. De gedecrypte boodschap is $m = D(c)$. Onderstel dat s een getal is zodat $2^{s+1} \geq n$. De attacker voedt het systeem met cyphertexts $2^{ie}c \bmod n$ voor $i = 0, \ldots, s$. Op deze manier verkrijgt hij

de volgende informatie:

$$\begin{aligned} l_i &= \mathrm{half}(D(2^{ie}c \bmod n)) \\ &= \mathrm{half}(2^i D(c) \bmod n) \\ &= \mathrm{half}(2^i m \bmod n) \end{aligned}$$

Dit is voldoende om m te kunnen reconstrueren. Om dit in te zien observeren we

$$l_0 = \mathrm{half}(m) = 0 \iff m \in]0, n/2[$$

$$\begin{aligned} l_1 = \mathrm{half}(2m \bmod n) &= 0 \\ \iff 2m &\in \big[0, \frac{n}{2}\big[\ \cup \ \big[n, \frac{3n}{2}\big[\\ \iff m &\in \big[0, \frac{n}{4}\big[\ \cup \ \big[\frac{n}{2}, \frac{3n}{4}\big[\end{aligned}$$

Dus als we zowel l_0 als l_1 kennen dan weten tot welk interval van de vorm

$$\big[\frac{pn}{4}, \frac{(p+1)n}{4}\big[, \qquad p = 0, 1, 2, 3$$

m behoort.

Als we deze redenering voortzetten dan vinden we dat als we $l_0, \ldots, l_s$ kennen dat we dan weten tot welk interval van de vorm

$$\big[\frac{pn}{2^{s+1}}, \frac{(p+1)n}{2^{s+1}}\big[, \qquad p = 0, \ldots, 2^{s+1} - 1$$

m behoort. Aangezien $2^{s+1} \geq n$ heeft zo'n (halfopen) interval lengte ≤ 1. Het zal dus ten hoogste 1 geheel getal bevatten. We kennen dus m.

Een meer realistische versie van deze setting wordt gegeven indien het systeem de minst significante bit lekt van m. I.e. de attacker kent nu voor $0 \leq m \leq n-1$:

$$\mathrm{LSB}(m) = \begin{cases} 0 & m \text{ is even} \\ 1 & m \text{ is oneven} \end{cases}$$

Lemma 9.5.1. *Er geldt*

$$\mathrm{half}(m) = \mathrm{LSB}(2m \bmod n)$$

Bewijs. Indien $\mathrm{half}(m) = 0$ dan $m < n/2$ en dus $2m < n$ en dus de rest van $2m$ na deling door n is $2m$. Aangezien $2m$ even is vinden we in dit geval $\mathrm{LSB}(2m \bmod n) = 0$.

Onderstel nu dat geldt $\mathrm{half}(m) = 1$. Dan $n/2 \leq m < n$ en dus $n \leq 2m < 2n$. De rest na deling door n van $2m$ is dan $2m - n$ wat oneven is. Bijgevolg $\mathrm{LSB}(2m \bmod n) = \mathrm{LSB}(2m - n) = 1$. □

We vinden nu

$$\begin{aligned} l_i &= \mathrm{half}(D(2^{ie}c \bmod n)) \\ &= \mathrm{LSB}(2D(2^{ie}c \bmod n) \bmod n) \\ &= \mathrm{LSB}(D(2^{(i+1)e}c \bmod n)) \end{aligned}$$

De attacker kan dus weer l_i bepalen en dus m.

En gelijkaardig resultaat geldt voor de andere bits van m. Dit resultaat is echter veel moeilijker. Zie [HN04].

Opmerking 9.5.2. Met noemt dit onderwerp "bit security" omdat men in feite aantoont dat de inviduele bits van m even veilig zijn als de volledige m.

9.6. **Aandachtspunt: verschillende encryptie exponenten.** Onderstel dat er twee stellen RSA parameters met dezelfde modulus in gebruik zijn

$$(n, p, q, d_1, e_1)$$
$$(n, p, q, d_2, e_2)$$

zodanig dat $\text{ggd}(e_1, e_2) = 1$ (bvb $e_1 = 3$ en $e_2 = 2^{16} + 1$) en dat deze gebruikt worden om dezelfde boodschap te encrypten:

$$c_1 \cong m^{e_1} \bmod n$$
$$c_2 \cong m^{e_2} \bmod n$$

Met behulp van het veralgemeende algoritme van Euclides kunnen we f_1, f_2 bepalen zodanig dat $f_1 e_1 + f_2 e_2 = 1$. We hebben dan

$$c_1^{f_1} c_2^{f_2} \cong m^{f_1 e_1 + f_2 e_2} = m$$

Dus we kunnen m berekenen zonder d_1 of d_2 te kennen.

9.7. **Aandachtspunt: lage encryptie exponent.** Onderstel nu dat er drie stellen RSA parameters in gebruik zijn, alle met encryptie exponent $e = 3$:

$$(n_1, p_1, q_1, d_1, 3)$$
$$(n_2, p_2, q_2, d_2, 3)$$
$$(n_3, p_3, q_3, d_3, 3)$$

We onderstellen ook nog dat alle p_i, q_i verschillend zijn. Dit is niet onredelijk aangezien RSA parameters random genereerd worden en de priemgetallen heel groot zijn. We hebben dus $\text{ggd}(n_i, n_j) = 1$ voor $i \neq j$. Tenslotte mogen we zonder verlies aan algemeenheid onderstellen dat $n_1 \leq n_2, n_3$.

We onderstellen weer dat dezelfde boodschap $0 \leq m < n_1$ door de verschillende RSA parameters geëncrypt wordt. Dus

$$c_1 \equiv m^3 \qquad \bmod n_1$$
$$c_2 \equiv m^3 \qquad \bmod n_2$$
$$c_3 \equiv m^3 \qquad \bmod n_3$$

Vanwege de Chinese reststelling bestaat er een $0 \leq c < n_1 n_2 n_3$ zodat $c \equiv c_i \bmod n_i$. We hebben dan $c \equiv m^3 \bmod n_i$ en dus weer vanwege de Chinese reststelling

$$c \equiv m^3 \quad \bmod n_1 n_2 n_3$$

We hebben echter ook $0 \leq m^3 < n_1^3 \leq n_1 n_2 n_3$. Dus deze congruentie is in feite een gelijkheid:

$$c = m^3$$

Maar dan hebben we $m = c^{1/3}$. We hebben dus weer m gevonden zonder een der d_i's te kennen.

Tegenwoordig gebruikt men meestal de encryptie exponent $e = 2^{16} + 1$.

10. Het discrete logaritme probleem

10.1. **Inleiding.** Het RSA systeem is elegant en eenvoudig te implementeren. Het is echter gebaseerd om een eigenaardigheid van de gehele getallen met weinig mogelijkheid tot variatie.

Moderne cryptosystemen zijn dikwijls gebaseerd op een ander principe: het zogenaamde *discrete logaritme probleem*. Het discrete logaritme probleem wordt geformuleerd binnen een gekozen groep en deze keuzemogelijkheid leidt tot extra flexibiliteit.

Zij gegeven een groep G en een element $g \in G$. Het discrete logaritme probleem (DLP) voor (G, g) is het volgende probleem.

> Gegeven $y \in G$ in de cyclische groep voortgebracht door g bepaal $x \in \mathbb{N}$ zodat $y = g^x$.

Soms schrijven we $x = \log_g y$ maar we moeten ons daarbij herinneren dat x slechts bepaald is modulo de orde van g.

De meest voorkomende situatie is wanneer G cyclisch is en voortgebracht door g. In feite kunnen we altijd tot dit geval reduceren door G te vervangen door de deelgroep $\langle g \rangle$ voortgebracht door g. Meestal kennen we $\langle g \rangle$ echter niet expliciet en kunnen we dus enkel in de volledig groep G rekenen.

Voor cryptografie willen we uiteraard een groep gebruiken waarvoor het DLP moeilijk op te lossen is. Een slechte keuze is $G = \mathbb{Z}/n\mathbb{Z}$ want dan wordt het DLP herleid tot een deling in $\mathbb{Z}/n\mathbb{Z}$. Dit is triviaal uit te voeren met behulp van het veralgemeende algoritme van Euclides. Goede keuzes voor G zijn:

- $G = (\mathbb{Z}/p\mathbb{Z})^*$ voor p priem.
- Meer algemeen $G = \mathbb{F}_q^*$ met q een priemmacht.
- Groepen geassocieerd aan elliptische krommen (zie later).

In de literatuur vindt men ook veel artikels over meer exotische keuzes van groepen[16]. Cryptografische systemen zijn echter best gebaseerd op veel bestudeerde, en dus hopelijk goed begrepen, problemen.

We bespreken nu eerst enige cryptografische systemen gebaseerd op het DLP.

10.2. **Diffie-Hellman.** Diffie-Hellman is een key exchange protocol. Het laat toe aan Alice en Bob om een sleutel uit te wisselen zonder dat Eve die kan kennen. De parameters van dit protocol zijn een groep G en een element $g \in G$. Deze parameters zijn publiek. Als Alice en Bob een sleutel willen willen afspreken voor verdere communicatie via symmetrische encryptie dan kiezen ze allebei een geheim randomgetal a en b en sturen mekaar g^a en g^b door. De sleutel is g^{ab} en kan door beide partijen berekend worden. Eve ziet enkel g^a en g^b. De veiligheid van het Diffie-Hellman protocol is gebaseerd op de *Diffie-Hellman onderstelling* voor het koppel (G, g):

> Uit g, g^a en g^b is het moeilijk om g^{ab} te berekenen.

We hebben

$$\text{DH-onderstelling voor } (G, g) \Rightarrow \text{DLP is moeilijk voor } (G, g)$$

Het is echter niet geweten of het omgekeerde ook geldt.

[16] Bijvoorbeeld jacobianen van hyperelliptische krommen en klasgroepen van kwadratisch imaginaire getallenringen.

10.3. **ElGamal.** Dit is is een public key systeem. De parameters van dit systeem zijn een groep G, een element $g \in G$ en een random getal $0 < a < |G|$. De publieke sleutel is

$$(G, g, g^a)$$

en de private sleutel is

$$(G, g, a)$$

De verzamelingen plaintexts is

$$\mathcal{P} = G$$

en de verzameling ciphertexts is

$$\mathcal{C} = G \times G$$

Voor encrytie kiest Alice een geheim random getal $0 < k < |G|$ en berekent

$$E_k(m) = (g^k, mg^{ak})$$

Merk op dat aangezien $g^{ak} = (g^a)^k$ Alice dit kan berekenen door enkel gebruik te maken van de publieke sleutel van Bob.

De decryptie functie wordt gegeven door

$$D(c_1, c_2) = c_2 c_1^{-a}$$

Bob kan dit berekenen aangezien hij a kent.

Het ElGamal cryptosysteem is ook gebaseerd op de Diffie-Hellman onderstelling. Immers wanneer Eve een geëncrypte boodschap (c_1, c_2) ziet dan kent ze g^k, g en g^a (de laatste twee gegevens uit de publieke sleutel). Als ze hieruit g^{ak} kan bepalen dan kan ze de de boodschap decrypten.

Opmerking 10.3.1.

(1) Het is belangrijk dat k geheim is. Anders kan iedereen de boodschap (c_1, c_2) decrypten door $c_2(g^a)^{-k}$ te berekenen.
(2) Het belangrijk dat k steeds een andere waarde krijgt. Indien Alice altijd dezelfde k kiest dan geldt voor encrypties (c_1, c_2), (c_1', c_2') van boodschappen m, m':

$$m(m')^{-1} = c_2(c_2')^{-1}$$

Dus als Eve de decryptie van 1 boodschap kent dan kan ze alle andere boodschappen decrypten.

10.4. **Het Digital Signature Algorithm (DSA).** Het Digital Signature Algorithm is een van de huidige standaarden voor digitale handtekeningen.

De parameters van het DSA algoritme (zie [DSA00, p8]) zijn als volgt:

(1) Een priemgetal p met tussen 512 en 1024 bits. Het aantal bits moet deelbaar door 64 zijn. De eerste bit moet 1 zijn.
(2) Een priemdeler q van $p - 1$ met 160 bits. De eerste bit moet 1 zijn.
(3) $(\mathbb{Z}/p\mathbb{Z})^*$ bevat een unieke deelgroep C_q van orde q. Deze is cyclisch. We kiezen een generator $g \bmod p$ van deze deelgroep.
(4) Een random getal $0 < x < q$.

De parameters (p, q, g) zijn publiek en mogen gedeeld worden door een groep gebruikers.

Opmerking 10.4.1. Om g te bepalen gaan we als volgt te werk. We kiezen een random getal $0 < h < p-1$. Indien $h^{(p-1)/q} \bmod p \neq 1$ dan kiezen we $g = h^{(p-1)/q} \bmod p$. Omdat $h^{(p-1)/q} \bmod p$ een niet triviaal element is van de cyclische groep C_q met priem orde is $h^{(p-1)/q} \bmod p$ een generator.

Het aantal oplossingen van $h^{(p-1)/q} \equiv 1 \bmod p$ is $(p-1)/q$. Dus de kans dat de eerste h *niet* werkt is slechts $1/q$ wat totaal verwaarloosbaar is.

(1) De private sleutel van DSA is (p, q, g, x).
(2) De publieke sleutel van DSA is (p, q, g, g^x).

Om de handtekening op een boodschap M te genereren berekenen we:

(1) Een hash $m = \text{SHA-1}(M)$ van M. Het resultaat m heeft 160 bits en is dus van de grootteorde van q.
(2) Een random getal $0 < k < q$.
(3) $r = (g^k \bmod p) \bmod q$ (De "mod" symbolen staan hier voor "rest").
(4) $s = k^{-1}(m + xr) \bmod q$ ("mod" staat hier weer voor "rest").
(5) Indien $r = 0$ of $s = 0$ dan kiezen we een nieuwe k. De kans hierop is echter verwaarloosbaar klein.
(6) De handtekening op M is (r, s) en wordt samen met de boodschap opgestuurd.

Om een handtekening op een boodschap M te verifiëren gaan we als volgt te werk. We berekenen

$$\begin{aligned} m &= \text{SHA-1}(M) \\ u_1 &\equiv s^{-1} m \bmod q \\ u_2 &\equiv s^{-1} r \bmod q \end{aligned}$$

We aanvaarden de handtekening wanneer geldt

$$(g^{u_1}(g^x)^{u_2} \bmod p) \bmod q = r$$

Dit werkt omdat

$$\begin{aligned} g^{u_1}(g^x)^{u_2} &\equiv g^{s^{-1}(m+xr)} \\ &= g^k \bmod p \end{aligned}$$

Het idee achter DSA is dat we gebruik willen maken van de veiligheid van het DLP voor het grote priem p maar toch de grootte van de digitale handtekeningen willen beperken. Dit wordt bereikt door te reduceren modulo het kleinere priem q. Een DSA handtekening bestaat dus uit 320 bits, onafhankelijk van de grootte van p.

Als we het DLP in C_q of $(\mathbb{Z}/p\mathbb{Z})^*$ kunnen oplossen dan kunnen we uit de publieke sleutel de private sleutel reconstrueren. Voor het DLP in $(\mathbb{Z}/p\mathbb{Z})^*$ bestaat een krachtige algoritme (index calculus, zie verder) maar $(\mathbb{Z}/p\mathbb{Z})^*$ is een heel grote groep. C_q is een veel kleinere groep maar hier lijken we alleen minder krachtige algoritmen op toe te kunnen passen (Baby step, Giant step of Pollard rho, zie verder).

Opmerking 10.4.2. Het DSA algoritme kan aangepast worden aan andere groepen. Zo bestaat er ECDSA wat slaat op DSA over elliptische krommen.

Nu bespreken we enkele standaard algoritmen om het DLP probleem aan te pakken.

10.5. **Baby step, Giant step (Shanks).** Zij G een cyclische groep van orde n en zij g een generator van G. We geven nu een heel eenvoudig algoritme om het DLP in G op te lossen. Dit algoritme vereist $O(\sqrt{n})$ vermenigvuldigingen in de groep G. Daar bovenop is er ook nog opslagruimte voor $O(\sqrt{n})$ groepselementen vereist.

Procedure. Zij gegeven $y \in G$.

- (*Baby steps*) Bereken en slaag op

$$1, g, g^2, \ldots, g^{\lfloor\sqrt{n}\rfloor-1} \tag{10.1}$$

- (*Giant steps*) Bereken achtereenvolgens

$$yg^{-i\lfloor\sqrt{n}\rfloor} \qquad i = 0, 1, 2, \ldots$$

tot we een groepselement vinden dat voorkomt in (10.1). Indien dit het geval is dan

$$yg^{-i\lfloor\sqrt{n}\rfloor} = g^j$$

en dus

$$y = g^{j+i\lfloor\sqrt{n}\rfloor}$$

We vinden dus

$$\log_g y = j + i\lfloor\sqrt{n}\rfloor$$

Bewering De tweede stap van het algoritme zal success hebben voor een waarde $0 \leq i \leq \lfloor\sqrt{n}\rfloor + 1$.

Bewijs. Schrijf de gezochte $0 \leq x < n$ als

$$x = j + i\lfloor\sqrt{n}\rfloor$$

met $0 \leq j < \lfloor\sqrt{n}\rfloor$ (dit is gewoon een deling met rest).

We vinden dan

$$i = \frac{x-j}{\lfloor\sqrt{n}\rfloor} \leq \frac{x}{\lfloor\sqrt{n}\rfloor} \leq \frac{n-1}{\lfloor\sqrt{n}\rfloor} \leq \frac{n-1}{\sqrt{n}-1} = \sqrt{n}+1 \quad \square$$

10.6. **Cycli van random afbeeldingen.** Zij S een eindige verzameling met $|S| = n$. Sommige algoritmen zijn gebaseerd op het vinden van een cycle voor een random afbeelding $f : S \to S$.

Kies $y_0 \in S$ en bereken $(y_i)_i$ via $y_{i+1} = f(y_i)$. Zij j de index waarvoor de eerste herhaling plaats vindt. I.e. $y_0, \ldots, y_{j-1}$ zijn alle verschillend en $y_j \in \{y_0, \ldots, y_{j-1}\}$. Dan definiëren we $\rho(f, y_0) \stackrel{\text{def}}{=} j$. We beschouwen $\rho(f, y_0)$ als een random variabele (afhangende van de random functie f). Het is niet moeilijk om de verwachtingswaarde μ_n van $\rho(f, y_0)$ af te schatten[17]. Er geldt [MvOV97, Fact 2.37][18]

$$\lim_{n\to\infty} \frac{\mu_n}{\sqrt{\pi n/2}} = 1$$

Dus in gemiddelde mogen we na $O(\sqrt{n})$ stappen een herhaling verwachten.

De vraag is nu: hoe vinden we zo'n herhaling? I.e. hoe bepalen we $i < j$ met $y_i = y_j$? Er is een voor de hand liggen naieve aanpak voor dit probleem: sla alle

[17] De distributie van $\rho(f, y_0)$ wordt voor n groot gegeven door $P(\rho \geq j) \cong e^{-j^2/(2n)}$ [MvOV97, Fact 2.27]. Dit is een voorbeeld van een Weibull distributie.

[18] Dus numeriek $\mu_n \simeq 1.252\sqrt{n}$. De standaard afwijking σ_n van $\rho(f, y_0)$ voor $n \to \infty$ is gelijk aan $\sqrt{2\left(1 - \frac{\pi}{4}\right)n} \simeq 0.655\sqrt{n}$.

$(y_i)_i$ op tot we een herhaling vaststellen. Het nadeel van deze methode is dat veel opslagruimte vereist is.

Er is echter een algoritme dat bijna geen opslagruimte vereist. Dit is het *Floyd cycle finding*. Dit algoritme zal in het algemeen niet de eerste herhaling detecteren. De complexiteit blijft echter $O(\sqrt{n})$. De basis is het volgende lemma

Lemma 10.6.1. *Indien $y_i = y_j$ met $i < j$ dan bestaat er een $i \leq m < j$ zodat $y_m = y_{2m}$.*

Bewijs. Stel $k = j - i$. Uit $y_i = y_{i+k}$ leiden we af $y_{i+p} = y_{i+p+qk}$ voor alle $p, q \geq 0$. We beweren dat we p, q zo kunnen kiezen dat $i+p+qk = 2(i+p)$. Dit is equivalent met $qk = i + p$. We kunnen dus i aanvullen tot het eerstvolgende k-voud.

Dus we vinden $y_m = y_{2m}$ waarbij m het eerste k-voud $\geq i$ is. Dus $i \leq m < i + k = j$ en het lemma is bewezen. □

Het volgende algoritme detecteert een herhaling na het evalueren van f voor ten hoogste driemaal zoveel waarden en als bij het naieve algoritme. Er is nu echter bijna geen opslagruimte meer vereist.

Procedure. Start met (y_1, y_2) en bereken $(y_{i+1}, y_{2i+2}) = (f(y_i), f^2(y_{2i}))$. Als we (y_m, y_{2m}) vinden zodat $y_m = y_{2m}$ dan stoppen we.

10.7. **Pollard rho.** Dit is een probabilistisch algoritme met dezelfde complexiteit als Baby step, Giant step: $O(\sqrt{n})$. De vereiste opslagruimte is echter verwaarloosbaar.

Pollard rho lost in feite een iets ander probleem op. Zij G een cyclische groep met priemorde p en zij $\alpha, \beta \in G$. De output van het Pollard rho algoritme bestaat uit getallen $u, v \in \mathbb{Z}/p\mathbb{Z}$ zodat $\alpha^u \beta^v = 1$.

Onderstel dat $\alpha = g$ een generator van G is en zij $\beta = y = g^x$ met x onbekend. Dan geldt $g^{u+vx} = 1$ en dus vinden we $x \equiv -v^{-1}u \bmod p$. Merk op dat het algoritme faalt indien $u \equiv v \equiv 0 \bmod p$. Dit is erg onwaarschijnlijk.

Procedure. Verdeel G in drie verzamelingen G_1, G_2, G_3 van min of meer dezelfde grootte waarvoor het gemakkelijk is om te beslissen of een element $z \in G$ in G_1, G_2 of G_3 zit. Kies G_2 zodanig dat $1 \notin G_2$.

Zij $f : G \to G$ de afbeelding

$$f(y) = \begin{cases} \beta y & \text{indien } y \in G_1 \\ y^2 & \text{indien } y \in G_2 \\ \alpha y & \text{indien } y \in G_3 \end{cases}$$

Kies $y_0 = 1$. Dan is het duidelijk dat $y_i = f^i(y_0)$ van de vorm zal zijn $\alpha^{u_i}\beta^{v_i}$ met $(u_i, v_i) \in (\mathbb{Z}/p\mathbb{Z})^2$. Met behulp van het Floyd cycle finding algoritme vinden we $y_i = y_j$. Er geldt dan $\alpha^{u_i}\beta^{v_i} = \alpha^{u_j}\beta^{v_j}$ en dus $\alpha^{u_j-u_i}\beta^{v_j-v_i} = 1$.

Het is moeilijk om de complexiteit van dit algoritme af te schatten. Als we echter aannemen dat f zich gedraagt als een random afbeelding dan volgt uit §10.6 dat de complexiteit $O(\sqrt{p})$ is.

Voorbeeld 10.7.1. Zij $q = 630803$. De ontbinding in factoren van $q - 1$ is $2 \cdot 17 \cdot 18553$. Dus $(\mathbb{Z}/q\mathbb{Z})^*$ heeft een unieke cyclische deelgroep G van orde 18553. Aangezien $2^{2\cdot 17} \not\equiv 1 \bmod q$ vinden we dat $g = 2^{2\cdot 17} \bmod q = 580282 \bmod q$ een generator is van G.

Zij $y = 591325 \bmod q$. Omdat $y^{18553} \equiv 1 \bmod q$ geldt $y \in G$. We vragen ons af wat $\log_g y$ is.

We kiezen G_1, G_2, G_3 als volgt.

$$\begin{cases} y \in G_1 & \text{indien } y \equiv 0 \bmod 3 \\ y \in G_2 & \text{indien } y \equiv 2 \bmod 3 \\ y \in G_3 & \text{indien } y \equiv 1 \bmod 3 \end{cases}$$

Als we het Pollard rho algoritme toepassen dan vinden we na 872 vermenigvuldigingen in $G \subset (\mathbb{Z}/q\mathbb{Z})^*$ dat $\log_g y = 17681$.

Opmerking 10.7.2. Het is misschien nuttig het volgende op te merken: Baby step, Giant step is een toepassing van het "meet in the middle" principe en Pollard rho is een toepassing van de verjaardagsparadox.

10.8. **Silver-Pohlig-Hellman.** We bespreken nu een algoritme dat het discrete logaritme probleem efficiënt oplost voor een cyclische groep G met een generator g op voorwaarde $|G|$ slechts door kleine priemgetallen deelbaar is.

Als we dit algoritme dus willen toepassen op $G = \mathbb{F}_q^*$ dan moeten we eisen dat $q-1$ slechts door kleine priemgetallen deelbaar is.

Het algoritme bestaat uit twee gedeelten: een "pre-computatie" fase die onafhankelijk is van y en daarna de eigenlijke berekening van $\log_g y$.

Pre-computatie. Zij $n = |G|$. Voor alle priemdelers $p \mid n$ berekenen we (en slaan op)

$$r_{p,j} = g^{jn/p}, \qquad j = 0, \ldots, p-1$$

We hebben dan $r_{p,j}^p = 1$ en voorts $r_{p,i} \neq r_{p,j}$ voor $i \neq j$.

Eigenlijke berekening. Zij $n = \prod_p p^{\alpha_p}$ de ontbinding in priemfactoren van n. Zij $y \in G$. We zoeken x zodanig dat $y = g^x$. Het is voldoende om x te berekenen modulo p^{α_p} voor alle p. Schrijf

$$x = x_0 + x_1 p + \cdots + x_{\alpha_p - 1} p^{\alpha_p - 1} + a_p p^{\alpha_p}$$

$0 \leq x_i < p$, $a_p \in \mathbb{Z}$.

Bereken

$$\begin{aligned} y^{n/p} &= g^{(x_0 + x_1 p + \cdots) n/p} \\ &= g^{x_0 n/p} \\ &= r_{p,x_0} \end{aligned}$$

Hieruit kunnen we dus x_0 bepalen.

Bereken nu $y_1 = y g^{-x_0}$. Dan geldt

$$\begin{aligned} y_1^{n/p^2} &= g^{(x_1 p + x_2 p^2 + \cdots) n/p^2} \\ &= g^{x_1 n/p} \\ &= r_{p,x_1} \end{aligned}$$

Hieruit kunnen we nu x_1 bepalen. Het is duidelijk dat we deze procedure kunnen voortzetten tot we uiteindelijk alle $(x_i)_i$ gevonden hebben.

Opmerking 10.8.1. In plaats van alle $r_{p,j}$ op te slaan kunnen we Silver-Pohlig-Hellman combineren met Baby step, Giant step of Pollard rho. Het idee is dat indien we bijvoorbeeld x_0 moeten bepalen dan moeten we de vergelijking

$$y^{n/p} = g^{x_0 n/p}$$

oplossen. Dit is een DLP probleem in de cyclische groep van orde p voortgebracht door $g^{n/p}$. Hierop is zowel Baby step, Giant step als Pollard rho van toepassing.

10.9. **Index calculus.** De krachtigste algoritmen voor de berekening van discrete logaritmen (en ook voor ontbinden in factoren, zie verder) zijn gebaseerd op het "index calculus principe". Helaas is dit principe enkel op zeer speciale groepen van toepassing en bijvoorbeeld niet op de groepen die men kan construeren met behulp van elliptische krommen[19]. Dit is de reden waarom men denkt dat cryptografie gebaseerd op elliptische krommen fundamenteel veiliger is dan dewelke gebaseerd op eindige lichamen.

We zullen nu onderstellen dat $G = (\mathbb{Z}/p\mathbb{Z})^*$ waarbij p een priemgetal is. We onderstellen dat g een generator van G is en we kiezen $y \in G$. We willen x vinden zodat $g^x = y$.

Algoritmen gebaseerd op index calculus bestaan weer uit twee fases. Een pre-computatie fase, onafhankelijk van y en een eigenlijke berekeningsfase.

Pre-computatie. We kiezen een *factorbasis*

$$B = \{g_1, \ldots, g_r\}$$

van kleine priemgetallen. Meestal de verzameling

$$B = \{\text{priemen} \leq w\}$$

waarbij w een nog te bepalen getal is. De pre-computatie fase bestaat er uit om $x_i = \log_g g_i$ te vinden. Hiervoor gaan we als volgt te werk.

(1) Kies een random getal $0 < k < p$ en bereken $g^k \bmod p$ waarbij we het resultaat beschouwen als een element van $[1, p-1]$.
(2) Ga na of er $a_1, \ldots, a_r \in \mathbb{N}$ bestaan zodat

$$g^k \bmod p = g_1^{a_1} \cdots g_r^{a_r} \qquad (\text{in } \mathbb{Z}) \tag{10.2}$$

Dit kunnen we doen door $g^k \bmod p$ door alle priemen in B te delen (er zijn snellere methoden). Indien we zo'n $a_1, \ldots, a_r$ niet vinden dan gaan we terug naar (1).

Indien er $a_1, \ldots, a_r$ bestaan die voldoen aan (10.2) dan hebben we

$$k \equiv a_1 x_1 + \cdots + a_r x_r \qquad \bmod p - 1$$

We hebben dus een lineare relatie (modulo $p-1$) gevonden tussen de x_i. We slaan deze op en gaan terug naar (1). We blijven deze stappen herhalen tot we genoeg relaties gevonden hebben om x_i uniek te bepalen. We verwachten dat we hiervoor een beetje meer dan r relaties zullen nodig hebben (sommige relaties zullen lineair afhankelijk zijn).

Eigenlijke berekening.

(1) Kies een random getal $0 < l < p$.

[19]Dit is tenminste de huidige opvatting. Zie bijvoorbeeld [Sil98].

(2) Ga na of er $b_1, \ldots, b_r \in \mathbb{N}$ bestaan zodat

$$yg^{-l} \bmod p = g_1^{b_1} \cdots g_r^{b_r} \qquad \text{in } \mathbb{Z}$$

Indien niet ga dan terug naar (1). Indien wel dan hebben we de identiteit

$$\log_g y \equiv l + b_1 x_1 + \cdots + b_r x_r \qquad \bmod p - 1$$

Het DLP is dus opgelost.

De optimale keuze van w is een delicate afweging. Bijvoorbeeld als we kiezen

$$w = e^{\sqrt{\log p}}$$

dan kan men schatten dat de looptijd van het algoritme $O(e^{(\log p)^\alpha})$ is met een zekere $\alpha < 1$. Dit is dus een *subexponentieel algoritme*. Met andere woorden het algoritme is $O(p^\epsilon)$ voor elke $\epsilon > 0$. Dit is beter dan zowel Baby step, Giant step en Pollard rho die beide $O(\sqrt{p})$ zijn.

Opmerking 10.9.1. Dit slechts een naieve implementatie van het index calculus idee. Er zijn vele verfijningen mogelijk. Zie bijvoorbeeld [LO91]. Dit valt echter buiten het bestek van deze cursus.

Opmerking 10.9.2. Een van de goede eigenschappen van de index calculus is dat het zoeken van de lineaire relaties tijdens de pre-computatie parallel kan uitgevoerd worden. Dit geldt ook voor het berekenen van de discrete logaritmen in de tweede fase.

Als we de index calculus willen veralgemenen naar andere groepen G dan hebben we het volgende nodig:

(1) Een factor basis $B = \{g_1, \ldots, g_r\} \subset G$.
(2) Een algoritme dat voor een random $h \in G$ met een kans $\epsilon > 0$ op succes $a_1, \ldots, a_r \in \mathbb{Z}$ vindt zodat $h = g_1^{a_1} \cdots g_r^{a_r}$.

Dit werkt voor groepen van de vorm $\mathbb{F}_{p^n}^*$ met p klein en n groot. We schrijven $\mathbb{F}_{p^n} = \mathbb{F}_p[x]/(f(x))$ waarbij $f(x)$ een irreduciebel polynoom van graad n is. De rol van $\mathbb{Z}$ in bovenstaande bespreking wordt nu overgenomen door $\mathbb{F}_p[x]$. We nemen

$$B = \{\text{monische polynomen van graad} \leq w\} \cup \{g_0\}$$

waarbij g_0 een generator van $\mathbb{F}_p^*$ is.

11. Priemtesten en het genereren van grote priemgetallen

11.1. **Fermat pseudo-priemen.** Om grote priemgetallen te kunnen genereren moeten we eerst een snelle manier hebben om te verifiëren of een getal al dan niet priem is. Dit is ons eerste doel.

De kleine stelling van Fermat zegt dat indien n priem is dan geldt

$$(F_b) \qquad b^{n-1} \equiv 1 \qquad \bmod n$$

voor all b zodat $\gcd(b, n) = 1$. Helaas kunnen we, omgekeerd, uit het gelden van (F_b) niet concluderen dat n priem is.

Voorbeeld 11.1.1. We hebben $341 = 11 \cdot 31$. Dus 341 is niet priem. Anderzijds vertelt Maple ons

$$2^{340} \equiv 1 \bmod 341$$

Een samengesteld getal dat aan (F_b) voldoet zullen we een *Fermat pseudo-priem* met basis b noemen. Hier is een lijstje met de kleinste Fermat pseudo-priemen met basis b voor $b = 1, \dots, 7$ (dit is Sloane sequence A007535, zie [Slo])

b	1	2	3	4	5	6	7
n	4	341	91	15	4	35	6

Het wordt echter nog erger! Er zijn samengestelde getallen n die pseudo-priem zijn ten opzichte van elke basis b (waarvoor geldt $\gcd(b, n) = 1$). Dit noemt men *Carmichael getallen*. De kleinste Carmichael getallen zijn

$$561, 1105, 1729, 2465, 2821, 6601, 8911, 10585, 15841, 29341$$

(dit is Sloane sequence A002997).

Stelling 11.1.2. [AGP94] *Er bestaan oneindig veel Carmichael getallen.*

Stelling 11.1.3. *Een getal n is een Carmichael getal als en slechts als*

(1) *n is niet priem.*
(2) *Voor elk priem $p \mid n$ geldt $p - 1 \mid n - 1$.*
(3) *n is niet deelbaar door een kwadraat > 1 (we zeggen: n is* kwadraatvrij*).*

Bewijs. We bewijzen eerst de $\Rightarrow$ richting. Onderstel dus dat n een Carmichael getal is.

(1) Dit is deel van de definitie.
(2) Kies een generator $\bar{b}$ voor $(\mathbb{Z}/p\mathbb{Z})^*$. We willen b gebruiken als basis voor een Fermat test. Helaas is het mogelijk dat $\gcd(b, n) \neq 1$. We gaan daarom b vervangen door een b' zodat $b' \equiv b \bmod p$ en $\gcd(b', n) = 1$. Zij p^a de hoogste macht van p die n deelt en $n' = n/p^a$. Vanwege de Chinese reststelling bestaat er een b' zodat $b' \equiv b \bmod p^a$ en $b' \equiv 1 \bmod n'$. Dan is b' niet deelbaar door een der priemen die n delen en dus $\gcd(b', n) = 1$. Hieronder onderstellen we dus $\gcd(b, n) = 1$. Dan is n per hypothese Fermat pseudo-priem met basis b en dus
$$b^{n-1} \equiv 1 \qquad \bmod n$$
en dus
$$b^{n-1} \equiv 1 \qquad \bmod p$$
Omdat $\bar{b}$ een generator is van de cyclische groep $(\mathbb{Z}/p\mathbb{Z})^*$ volgt hieruit $p - 1 \mid n - 1$.
(3) Onderstel dat $p^2 \mid n$ voor een priemgetal p. De orde van $(\mathbb{Z}/p^2\mathbb{Z})^*$ is $\phi(p) = p(p-1)$. Zij $\bar{b}$ een generator van de Sylow p-groep van $(\mathbb{Z}/p^2\mathbb{Z})^*$. Dus
$$b \not\equiv 1 \bmod p^2 \tag{11.1}$$
en
$$b^p \equiv 1 \qquad \bmod p^2 \tag{11.2}$$
Zoals boven mogen we onderstellen $\gcd(b, n) = 1$ en dus per hypothese:
$$b^{n-1} \equiv 1 \qquad \bmod n$$
en dus
$$b^{n-1} \equiv 1 \qquad \bmod p^2 \tag{11.3}$$

Aangezien $p \mid n$ hebben we $\gcd(p, n-1) = 1$. Dus de congruenties (11.2)(11.3) impliceren

$$b \equiv 1 \qquad \bmod p^2$$

wat in strijd met (11.1) is.

Nu bewijzen we de omgekeerde implicatie. Onderstel dat $n = p_1 \cdots p_n$ met p_i priem zodat $p_i - 1 \mid n-1$. Alle p_i zijn verschillend. Zij $\gcd(b, n) = 1$. Met andere woorden $p_i \nmid b$. Er geldt voor elke p_i:

$$b^{p_i - 1} \equiv 1 \qquad \bmod p_i$$

Omdat $p_i - 1 \mid n - 1$ geldt ook

$$b^{n-1} \equiv 1 \qquad \bmod p_i$$

Uit de Chinese reststelling halen we dan

$$b^{n-1} \equiv 1 \mod n \quad \square$$

Gevolg 11.1.4. (1) *Een Carmichael getal is oneven.*
(2) *Een Carmichael getal is niet van de vorm pr met p priem en $p > r > 1$.*
(3) *Een Carmichael getal is deelbaar door tenminste drie priemen.*

Bewijs. (1) Onderstel dat n een even Carmichael getal is. Aangezien n samengesteld is en niet deelbaar door 4 bestaat er een oneven priemgetal $p \mid n$. Dan geldt $p - 1 \mid n - 1$. Dit is onmogelijk aangezien $2 \mid p - 1$ en $2 \nmid n - 1$.

(2) Zij $n = pr$ een Carmichael getal met $p > r$. Dan geldt $p - 1 \mid pr - 1$. Oftewel

$$pr \equiv 1 \qquad \bmod p - 1$$

Aangezien $p \equiv 1 \bmod p - 1$ impliceert dit

$$r \equiv 1 \qquad \bmod p - 1$$

en dus

$$p - 1 \mid r - 1$$

Dit is onmogelijk aangezien $2 \leq r \leq p - 1$ en dus $1 \leq r - 1 < p - 1$.

(3) Zij $n = pq$ met $p > q$ priem. Dan volgt uit (2) dat n geen Carmichael getal kan zijn. $\square$

Opmerking 11.1.5. Het eerste Carmichael getal deelbaar door 4 priemen is $41041 = 7 \cdot 11 \cdot 13 \cdot 41$. Het eerste Carmichael getal deelbaar door 5 priemen is $825265 = 5 \cdot 7 \cdot 17 \cdot 19 \cdot 73$. Dit is Sloane sequence A006931.

11.2. **Euler pseudo-priemen.** Onderstel dat $n > 0$ oneven is. Vanwege Stelling 6.2.1 weten we dat indien n priem is dan

$$(E_b) \qquad b^{(n-1)/2} \equiv \left(\frac{b}{n}\right) \qquad \bmod n$$

voor alle b zodat $\gcd(b, n) = 1$. Een samengesteld getal dat aan (E_b) voldoet noemen we een *Euler pseudo-priem* met basis b.

Lemma 11.2.1. *Een Euler pseudo-priem met basis b is een Fermat pseudo-priem met basis b.*

Bewijs. Dit volgt door (E_b) te kwadrateren. $\square$

Voorbeeld 11.2.2. We hadden gezien dat $341 = 11 \cdot 31$ een Fermat pseudo-priem met basis 2 was. Is het ook een Euler pseudo-priem? We berekenen

$$2^{(341-1)/2} \equiv 1 \bmod 341$$

Aan de andere kant geldt

$$\left(\frac{2}{n}\right) = -1$$

want $n \equiv 5 \bmod 8$. Het is dus geen Euler pseudo-priem. Het omgekeerde van lemma 11.2.1 is vals.

Hier is een lijstje met de kleinste Euler pseudo-priemen met gegeven basis

b	1	2	3	4	5	6	7
n	9	561	121	341	781	217	25

Er is een belangrijk onderscheid tussen Euler pseudo-priemen en Fermat pseudo-priemen.

Stelling 11.2.3. *Als n oneven en samengesteld is dan faalt (E_b) voor tenminste de helft van de $b \bmod n$ zodat* $\gcd(b, n) = 1$.

Bewijs.

Stap 1. Zij

$$H = \{\bar{b} \in (\mathbb{Z}/n\mathbb{Z})^* \mid (E_b) \text{ geldt voor } b\}$$

De verzameling H is gesloten onder de vermenigvuldiging en is dus een deelgroep van $(\mathbb{Z}/n\mathbb{Z})^*$. Indien $H \neq (\mathbb{Z}/n\mathbb{Z})^*$ dan geldt $|H| \leq |(\mathbb{Z}/n\mathbb{Z})^*|/2$. Het is dus voldoende aan te tonen $H \neq (\mathbb{Z}/n\mathbb{Z})^*$.

Stap 2. Onderstel $H = (\mathbb{Z}/n\mathbb{Z})^*$. Dus n is een Euler pseudo-priem voor elke basis. Uit lemma 11.2.1 volgt dat n een Fermat pseudo-priem is voor elke basis. Met andere woorden n is een Carmichael getal. Dus vanwege Stelling 11.1.3 weten we dat n kwadraatvrij is.

Stap 3. Schrijf nu $n = p_1 \cdots p_r$ met p_i priem en kies b (met behulp van de Chinese reststelling) zodanig dat

$$\left(\frac{b}{p_1}\right) = -1$$
$$b \equiv 1 \bmod p_i \qquad i = 2, 3, \ldots, r$$

Dan geldt

$$\left(\frac{b}{n}\right) = \left(\frac{b}{p_1}\right)\left(\frac{b}{p_2}\right)\cdots\left(\frac{b}{p_r}\right) = -1$$

en

$$b^{(n-1)/2} \equiv 1 \bmod p_2$$

Dus $b^{(n-1)/2} \not\equiv \left(\frac{b}{n}\right) \bmod p_2$. Met andere woorden $b \notin H$. □

11.3. **De Solovay-Strassen test.** De Solovay-Strassen test gaat als volgt.

(1) Kies $0 < b < n$ random met $\gcd(b, n) = 1$.
(2) Indien (E_b) niet geldt dan is n samengesteld. Indien (E_b) geldt ga terug naar (1).

Indien n priem is dan stopt dit algoritme niet. Daarom voeren we de stappen (1)(2) een vooraf gekozen aantal keer k uit. Indien er geen besluit is dan weten we dat de kans dat n samengesteld is kleiner dan 2^{-k} is. In de praktijk zal men zo'n n als zijnde priem aanvaarden. Men gebruikt bijvoorbeeld $k = 50$.

11.4. **Sterke pseudo-priemen.** We hebben de volgende stelling.

Stelling 11.4.1. *Onderstel dan n een oneven priem is. Schrijf $n-1 = 2^s t$ met t oneven. Dan geldt voor elke b met $\gcd(b,n) = 1$ een van de volgende alternatieven*

$$(S_b) \qquad \begin{array}{c} b^t \equiv 1 \bmod n \qquad \textit{of} \\ \exists r : 0 \leq r < s : b^{2^r t} \equiv -1 \bmod n \end{array}$$

Bewijs. Onderstel n priem. Vanwege de kleine stelling van Fermat hebben we

$$b^{2^s t} \equiv 1 \bmod n$$

Dus $(b^{2^{s-1}t})^2 \equiv 1 \bmod n$. Omdat $\mathbb{Z}/n\mathbb{Z}$ een lichaam is halen we hier uit $b^{2^{s-1}t} \equiv \pm 1 \bmod n$. Als $b^{2^{s-1}t} \equiv -1 \bmod n$ dan zijn we klaar. Indien $b^{2^{s-1}t} \equiv 1 \bmod n$ en $s-1 = 0$ dan zijn we ook klaar.

Indien $b^{2^{s-1}t} \equiv 1 \bmod n$ en $s-1 > 0$ dan herhalen we de redenering. Uiteindelijk komen we uit op een van de vermelde alternatieven. □

Indien (S_b) waar is voor een samengestelde n dan noemen we n een *sterk pseudo-priem* met basis b.

Door de alternatieven in (S_b) tot de gepaste tweemacht te verheffen vinden we dat een sterk pseudo-priem met basis b een Fermat pseudo-priem met basis b is. Het volgende is veel moeilijker te bewijzen (zie [Kob94, Proposition V.1.6]).

Stelling 11.4.2. *Een sterk pseudo priem met basis b is een Euler pseudo-priem met basis b.*

Voorbeeld 11.4.3. We hebben gezien dat 561 een Euler pseudo-priem is met basis 2. Is het ook een sterk pseudo-priem met basis 2? We berekenen $560 = 2^4 \cdot 5 \cdot 7$. Dus $s = 4$, $t = 35$. Er geldt

$$2^{35} \equiv 263 \not\equiv \pm 1 \mod 561$$

We moeten nu $2^{2^1 \cdot 35}$, $2^{2^2 \cdot 35}$ en $2^{2^3 \cdot 35}$ uitrekenen om te kijken of we $-1 \bmod 561$ uit komen. Dit doen we uiteraard door 263 herhaald te kwadrateren modulo 561. We vinden.

$$166, 67, 1$$

Dus 561 is niet sterk pseudo-priem met basis 2.

Hier is een lijstje met de kleinste sterke pseudo-priemen met gegeven basis

b	1	2	3	4	5	6	7
n	9	2047	121	341	781	217	25

Sterke pseudo-priemen hebben echter nog betere eigenschappen dan Euler pseudo-priemen.

Stelling 11.4.4. *Als n oneven en samengesteld is dan faalt (S_b) voor tenminste drie vierde van de $b \bmod n$ zodat $\gcd(b,n) = 1$.*

Dit bewijs is weer wat technisch. Zie [Kob94, Proposition V.1.7].

11.5. **De Miller-Rabin test.** De Miller-Rabin test is heden ten dage de meest gebruikte (pseudo-)priemtest omdat hij heel gemakkelijk te implementeren is en bovendien bijzonder snel is. Hij gaat als volgt.

(1) Kies $0 < b < n$ random met $\mathrm{ggd}(b, n) = 1$.
(2) Indien (S_b) niet geldt dan is n samengesteld. Indien (S_b) geldt ga terug naar (1).

Indien n priem is dan stopt dit algoritme niet. Daarom voeren we de stappen (1)(2) een vooraf gekozen aantal keer k uit (bijvoorbeeld 25 keer). Indien er geen besluit is dan weten we dat de kans dat n samengesteld is kleiner dan 2^{-2k} is. In de praktijk zal men zo'n n als zijnde priem aanvaarden.

Wat is de complexiteit van de Miller-Rabin test? Om een sterke priemtest met basis b uit te voeren moeten we in feite gewoon $b^{(n-1)/2} \bmod n$ uitrekenen met herhaald kwadrateren.[20]

Hiervoor zijn $O(\log n)$ vermenigvuldigingen modulo n nodig. De complexiteit van een enkele test is dus $O(\log(n)^{2+\epsilon})$ waarbij $O(\log(n)^{1+\epsilon})$ de complexiteit van een vermenigvuldiging mod n is. Voor de naieve vermenigvuldiging hebben we $\epsilon = 1$ maar met moderne methoden kunnen we $\epsilon > 0$ willekeurig klein maken.

Omdat we een vast aantal testen uitvoeren is de complexiteit van de volledige Miller-Rabin test ook gelijk aan $O(\log(n)^{2+\epsilon})$.

11.6. **Echte priemtesten.** Pseudo-priemtesten zijn voldoende voor de praktijk maar niet voor een wiskundige. Men zoekt daarom naar testen die ondubbelzinnig vaststellen of een getal al dan niet priem is. Een vereiste is natuurlijk dat zo'n test een accepteerbare snelheid heeft.

Een grote doorbraak was het AKS algoritme [AKS04] (ontdekt door drie Indische informatici). Dit is een niet probabilistisch algoritme dat in polynomiale tijd bepaalt of een getal al dan niet priem is.

Het AKS algoritme is gebaseerd op de volgende simpele observatie: indien n priem is dan hebben we voor alle $a \in \mathbb{Z}/n\mathbb{Z}$ en alle $r \in \mathbb{N}$

$$(x + a)^n = x^n + a \tag{11.4}$$

in $(\mathbb{Z}/n\mathbb{Z})[x]/(x^r - 1)$.

In [AKS04] toont men aan dat indien (11.4) geldt voor een goedgekozen r en een goedgekozen verzameling a's dat dan n noodzakelijk priem is.

Voor een goed overzicht (met bewijzen!) zie [Sch04].

11.7. **De priemgetalstelling.** We weten nu hoe we efficient kunnen verifiëren of een getal priem is. Om nu random priemgetallen te kunnen genereren moeten we een idee hebben hoeveel priemgetallen er zijn. Deze vraag leidt onmiddelijk naar een mooi stuk getaltheorie. Een goede referentie is [Sil98, §13].

We definieren

$$\pi(x) = |\{\text{priemen} \leq x\}|$$

Het bewijs van de volgende stelling is heel moeilijk.

[20] De aandachtige lezer merkt op dat dit een lichtjes andere versie is van herhaald kwadrateren. Immers indien $n - 1 = 2^s t$ en dan rekenen we bij een sterke priemtest met basis b eerst $b^t \bmod n$ uit en daarna gaan we kwadrateren. Bij het algoritme dat we gezien hebben om $b^{2^{s-1}t} \bmod n$ uit te rekenen voeren we het kwadrateren eerst uit. Het aantal benodigde operaties is uiteraard precies hetzelfde.

Stelling 11.7.1. *Er geldt*

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

Dat dit waar was werd als vermoeden gesteld door Gauß en Legendre rond 1800. Het nam bijna honderd jaar in beslag om een bewijs te vinden. Dit gebeurde uiteindelijk in 1896 door (onafhankelijk) Hadamard en de Belgische wiskundige de la Vallée Poussin.

Voorbeeld 11.7.2. Onderstel $x = 10^9$. Dan geldt

$$\pi(x) = 50847534$$
$$x/\log(x) = 48254942{,}43$$

en dus

$$\frac{\pi(x)}{x/\log(x)} = 1,054$$

In de praktijk zijn we geinteresseerd in de dichtheid van de priemgetallen. I.e.

$$\frac{\pi(x+\Delta) - \pi(x)}{\Delta}$$

waarbij Δ "niet te groot" is vergeleken met x. We kunnen vermoeden dat dit ongeveer gelijk is aan

$$\cong \frac{d}{dx}\frac{x}{\log(x)} = \frac{1}{\log(x)} - \frac{1}{\log(x)^2} \approx \frac{1}{\log(x)}$$

Dit blijkt ook waar te zijn.[21]

Voorbeeld 11.7.3. Er geldt

$$|\{\text{priemen} \in [10^7, 10^7 + 999]\}| = 61$$

$$\frac{1}{\log(10^7 + 500)} = 0{,}062$$

Hier is nu een procedure om een randompriemgetal met l-bits te genereren.

(1) Genereer een randomgetal n in $[2^{l-1}, 2^l[$.
(2) Test of n priem is. Indien niet ga terug naar (1). Indien wel dan zijn we klaar.

Vanwege bovenstaande discussie zullen we ongeveer

$$\log(2^l) = \log(2)l = 0{,}69l$$

keer moeten proberen voor we succes hebben. De Miller-Rabin test heeft complexiteit $O(l^{2+\epsilon})$ waarbij $O(l^{1+\epsilon})$ de complexiteit van een vermenigvuldiging mod n is (zie §11.5). We besluiten dat de complexiteit van het genereren van een priemgetal met l bits gelijk is aan $O(l^{3+\epsilon})$.

[21] Een meer nauwkeurige versie van de priemgetalstelling zegt: $\pi(x) \approx \int_{t=2}^{x} \frac{dt}{\log(t)}$. In deze formulering zien we duidelijk dat de dichtheid van de priemgetallen $\approx 1/\log(x)$ is.

11.8. **Priemcertificaten.** Een priemcertificaat is een stukje data dat het eenvoudig maakt om te controleren of een gegeven getal inderdaad priem is. Merk op dat het niet noodzakelijk eenvoudig is om zo'n certificaat te genereren.

Lemma 11.8.1. *Een getal p is priem als en slechts als er een $a \in \mathbb{Z}$ bestaat zodat de volgende condities gelden:*

(1) $a^{p-1} \equiv 1 \bmod p$.
(2) $\forall q \mid p-1$ *priem geldt* $a^{(p-1)/q} \not\equiv 1 \bmod p$.

Bewijs. $\Rightarrow$ Kies a zodat $\bar{a}$ een generator van de cyclische groep $(\mathbb{Z}/p\mathbb{Z})^*$ is.

$\Leftarrow$ Onderstel dat (1)(2) geldt en dat p niet priem is. De orde van $\bar{a}$ in de groep $(\mathbb{Z}/p\mathbb{Z})^*$ is gelijk aan $p-1$. Het aantal elementen van $(\mathbb{Z}/p\mathbb{Z})^*$ is $p-1$ min het aantal getallen in $]1, p[$ dat niet relatief priem is met p. Aangezien p niet priem is bestaan er zo'n getallen. Dus $|(\mathbb{Z}/p\mathbb{Z})^*| < p-1$. Aangezien de orde van een element van een eindige groep altijd kleiner of gelijk is aan de orde van de groep komen we een contradictie uit. □

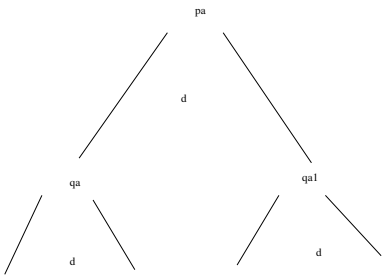
We beschrijven nu een zogenaamd *Pratt certificaat* voor een priem p. Zo'n Pratt certificaat bestaat uit de volgende data

$$(p, q_1, \alpha_1, \cdots, q_r, \alpha_r, a, \ldots)$$

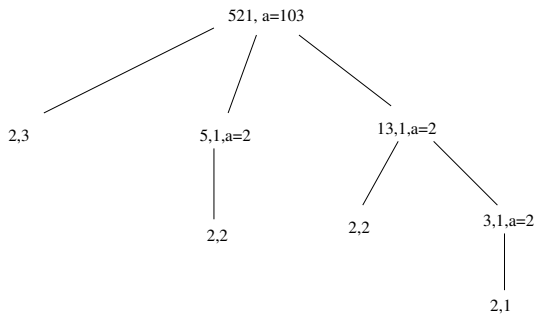
waarbij

(1) $p - 1 = q_1^{\alpha_1} \cdots q_r^{\alpha_r}$ is de ontbinding in priemfactoren van $p-1$.
(2) a is zoals in het lemma.
(3) ... zijn de Pratt certificaten voor q_i indien deze niet behoren tot een lijst "bekende priemen" (bvb alle priemen kleiner dan 1000).

In feite kunnen we zo'n Pratt certificaat dus beter voorstellen als een boom.



Voorbeeld 11.8.2. Hier is het Pratt certificaat van 521 voorgesteld als een boom. We hebben enkel 2 als bekend priem aangenomen.



12. Ontbinding in factoren

12.1. **Inleiding.** Indien een priemtest aantoont dat n niet priem is dan kunnen we ons afvragen wat de ontbinding in factoren van n is. In de eerste plaats kunnen we proberen een niet triviale ontbinding

$$n = n_1 n_2 \tag{12.1}$$

te vinden. Indien we daar in slagen dan kunnen we proberen om n_1 en n_2 verder te ontbinden. In de volgende secties bespreken we dus (12.1).

De (on)efficientie van ontbinden in factoren is uiteraard zeer belangrijk voor de veiligheid van het RSA crypto-systeem. Daarom creëerde RSA Laboratories indertijd een aantal "challenges" waarmee zelfs geld te verdienen was. Zie

http://en.wikipedia.org/wiki/RSA_numbers

Alhoewel de challenges niet langer actief zijn probeert men nog steeds de RSA-getallen te ontbinden in factoren.

Hier is het kleinste getal dat open is (RSA-704, 212 cijfers)

```
74037563479561712828046796097429573142593188889231
28908493623263897276503402826627689199641962511784
39958943305021275853701189680982867331732731089309
00552505116877063299072396380786710086096962537934
650563796359
```

Hiermee kon je indertijd 30.000$ verdienen.

12.2. **Trial division.** De meest naieve methode om een niet triviale factor van n te vinden is te delen door $2, \ldots, \lfloor\sqrt{n}\rfloor$. Dit vereist in het slechtste geval $O(\sqrt{n})$ delingen. Deze methode is zo traag dat ze enkel van toepassing is op heel kleine n.

Een nuttige variant is te delen door priemen kleiner dan een gegeven getal w (bvb 1000). Op deze manier kunnen we snel kleine factoren van n detecteren en wegdelen.

12.3. **Pollard-rho.** Onderstel dat p de kleinste priemfactor is die n deelt. Beschouw de afbeelding

$$f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} : x \mapsto x^2 + \bar{1}$$

Kies $x_0 = \bar{1}$ (of eventueel random) en bereken $(x_i)_i$ via $x_{i+1} = f(x_i)$. Bereken voor $j \geq 0$

$$\text{ggd}(x_{2j} - x_j, n)$$

We hebben een commutatief diagram

$$\begin{array}{ccc} \mathbb{Z}/n\mathbb{Z} & \xrightarrow{f} & \mathbb{Z}/n\mathbb{Z} \\ {\scriptstyle \bmod p}\downarrow & & \downarrow{\scriptstyle \bmod p} \\ \mathbb{Z}/p\mathbb{Z} & \xrightarrow[\bar{f}]{} & \mathbb{Z}/p\mathbb{Z} \end{array}$$

waarbij $\bar{f}$ door dezelfde formule als f gegeven wordt: $\bar{f}(x) = x^2 + \bar{1}$.

Indien $\bar{f}$ zich gedraagt als een randomafbeelding (wat in de praktijk het geval blijkt te zijn) dan vinden zoals uitgelegd in §10.7 na $O(\sqrt{p})$ stappen een congruentie

$$x_{2j} \equiv x_j \bmod p$$

Het is veel minder waarschijnlijk dat ook zal gelden $x_{2j} \equiv x_j \bmod n$ omdat hiervoor in het algemeen $O(\sqrt{n})$ stappen nodig zijn. Met andere woorden p deelt $x_{2j} - x_j$ maar niet n. Dus $\gcd(x_{2j} - x_j, n)$ zal een niet triviale factor van n zijn.

Indien we pech hebben en dit niet zo is dan beginnen we opnieuw met een andere x_0.

Voorbeeld 12.3.1. Zij $n = 2^{32} + 1$, $x_0 = 1$. Dan vinden we $\gcd(x_{44} - x_{22}, n) = 641$.[22]

12.4. **Pollard $p-1$.** Dit is een algoritme dat werkt indien n een priemfactor p heeft zodat $p-1$ slechts "kleine" priemfactoren heeft. Onderstel dat de ontbinding in priemfactoren van $p-1$ wordt gegeven door

$$p - 1 = q_1^{\alpha_1} \cdots q_n^{\alpha_n}$$

waarbij $q_i^{\alpha_i} \leq w$. Hierbij is w een vooraf gekozen parameter. Bereken

$$M = \prod_{q^\alpha \leq w < q^{\alpha+1}} q^\alpha$$

Dan geldt $p - 1 \mid M$. Dus indien $\gcd(a, n) = 1$

$$a^M = (a^{p-1})^{\text{iets}} \equiv 1 \bmod p$$

en dus

$$p \mid a^M - 1$$

Indien we geluk hebben dan is $\gcd(n, a^M - 1)$ een niet triviale factor van n. Merk op dat we $a^M - 1$ uiteraard enkel modulo n moeten uitrekenen.

Opmerking 12.4.1. Indien

$$M = 2^x 3^y 5^z \cdots w^t$$

dan kan men overwegen om

$$a^2, a^4, \cdots, a^{2^x}, a^{3 \cdot 2^x}, a^{3^2 \cdot 2^x}, \ldots, a^M \bmod n$$

uit te rekenen. Dan moet men niet eerst M volledig uitrekenen.

12.5. **De Fermat methode.** Als n een oneven getal van de vorm $n_1 n_2$ is dan kunnen we n schrijven als een verschil van twee kwadraten

$$n = \left(\frac{n_1 + n_2}{2}\right)^2 - \left(\frac{n_1 - n_2}{2}\right)^2$$

Omgekeerd indien n een verschil van twee kwadraten is

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

dan geeft dit een ontbinding in factoren $n = (t - s)(t + s)$. Hierop is de Fermat methode gebaseerd. Met name voor t in de rij

$$\lceil \sqrt{n} \rceil, \lceil \sqrt{n} \rceil + 1, \lceil \sqrt{n} \rceil + 2, \ldots$$

controleren we of $t^2 - n$ een kwadraat s^2 is. Indien dit zo is dan geldt $n = t^2 - s^2$ en we hebben de bijbehorende ontbinding in factoren.

---

[22] In 1640 uitte Fermat in een brief aan Mersenne het vermoeden dat $F_n = 2^{2^n} + 1$ altijd priem is. Dit is immers zo voor F_0, F_1, F_2, F_3, F_4. Fermat en Mersenne hebben nooit kunnen beslissen of $F_5 = 4294967297$ al dan niet priem is. Indien zij de Pollard-rho methode hadden gekend dan hadden zij zelfs met de hand de factor 641 van F_5 kunnen vinden.

Dit algoritme werk indien s klein is. I.e. indien n een produkt is van twee dicht bij elkaar liggende factoren.

Er bestaat een nuttige variant op dit algoritme. Merk op dat indien geldt

$$\begin{aligned} t^2 &\equiv s^2 \bmod n \\ t &\not\equiv \pm s \bmod n \end{aligned} \tag{12.2}$$

dan $n \mid t^2 - s^2 = (t-s)(t+s)$ en $n \nmid (t+s)$, $n \nmid (t-s)$. Dus is $\gcd(n, t+s)$ een niet triviale factor van n.

12.6. **Factorbasis methoden.** Bij de bespreking van het discrete logaritme probleem hebben we reeds kennis gemaakt met een factorbasis methode (de zogenaamde index calculus). Dergelijke methoden bestaan ook voor het ontbinden in factoren, we gaan ze nu beschrijven.

Het basisprincipe is (12.2). We gaan een t zoeken zodat t^2 op een niet triviale manier een kwadraat is modulo n. We kiezen allereerst een zogenaamde *factorbasis* B. Typisch:

$$B = \{-1\} \cup \{\text{priemen} \leq w\}$$

en we zoeken naar een t zodat

$$t^2 \equiv \prod_{p \in B} p^{\epsilon_p} \bmod n \tag{12.3}$$

waarbij alle ϵ_p even zijn. Indien dit het geval is dan is t^2 inderdaad een kwadraat modulo n.

Hier is de methode. Voor $x \in \mathbb{Z}$ definieer $\tilde{x}$ als het unieke getal in $[-n/2, n/2[$ zodat $\tilde{x} \equiv x \bmod n$. Onderstel dat we een procedure hebben die ons getallen t_i aanlevert zodat

$$\widetilde{t_i^2} \equiv \prod_{p \in B} p^{\epsilon_{p,i}} \tag{12.4}$$

(met $\epsilon_{p,i}$ niet noodzakelijk even). We kunnen dan een t zoeken die aan (12.3) voldoet door deze t_i te combineren. Met andere woorden we zoeken

$$t = \prod_i t_i^{\eta_i}$$

met $\eta_i = 0, 1$ zodat (12.3) voldaan is met ϵ_p even. Uitwerken levert

$$t^2 \equiv \prod_{p \in B} p^{\sum_i \epsilon_{p,i}\eta_i} \bmod n$$

Dus we moeten oplossen

$$\sum_i \epsilon_{p,i}\eta_i \equiv 0 \bmod 2$$

Met andere woorden we moeten een stelsel van vergelijkingen modulo 2 oplossen.

Voorbeeld 12.6.1. Beschouw $n = 4633$. We gebruiken als factorbasis $B = \{-1, 2, 3\}$. We hebben $\sqrt{n} = 68{,}07$. We berekenen

$$\begin{aligned} 67^2 = 4489 &\equiv -144 = -2^4 3^2 \\ 68^2 = 4624 &\equiv -9 = -3^2 \\ 69^2 = 4761 &\equiv 128 = 2^7 \end{aligned}$$

Als we nu zoeken naar (η_1, η_2, η_3) zodanig dat $t = 67^{\eta_1}68^{\eta_2}69^{\eta_3}$ een kwadraat is dan moeten we het volgende stelsel over $\mathbb{F}_2$ oplossen.

$$\eta_1 + \eta_2 = 0$$
$$\eta_3 = 0$$

Een oplossing wordt gegeven door $(\eta_1, \eta_2, \eta_3) = (1, 1, 0)$. Dus we vinden $t = 67^1 68^1 69^0 = 4556$ en dus $t^2 \equiv (-2^4 3^2)^1(-3^2)^1(2^7)^0 = 2^4 3^4 = (2^2 3^2)^2$. We berekenen

$$\text{ggd}(n, t + 2^2 3^2) = 41$$

en we hebben dus inderdaad een niet triviale factor van n gevonden.

Opmerking 12.6.2. Het kan voordelig zijn t_i in een verzameling van de vorm

$$T: \qquad -k' + \lfloor\sqrt{jn}\rfloor, \ldots, k' + \lfloor\sqrt{jn}\rfloor,$$

te zoeken met $j = 1, 2, \ldots, l$.

Om echt een oordeel te vellen wat de ideale waarden van k, l etc... zijn moeten we een schatting kunnen maken wat de kans is dat een positief getal $\leq x$ een product is van priemen in B.[23] Na een heuristische berekening bekomt Koblitz [Kob94, §V.3] hiervoor de volgende afschatting:

$$u^{-u}$$

waarbij $u = \log(x)/\log(w)$. Dwz u is ruwweg de verhouding tussen het aantal bits van x en w.

12.7. **Kettingbreuken.** Iedereen kent de benadering

$$\pi \approx \frac{22}{7}$$

en sommigen misschien ook

$$\pi \approx \frac{355}{113}$$

uit de lagere school. Deze benaderingen zijn merkwaardig accuraat.

$$\pi = 3{,}1415926536\cdots$$
$$\frac{22}{7} = 3{,}1428571429\cdots$$

Dus de benadering van π door $22/7$ is accuraat op bijna drie decimalen terwijl we normaal gezien met een breuk van de vorm $y/7$ slechts een nauwkeurigheid van $1/7 \approx 1$ decimaal zouden verwachten.

$$\frac{355}{113} = 3{,}1415929204\cdots$$

Nu is de nauwkeurigheid bijna 7 decimalen terwijl we in feite slechts 3 decimalen zouden verwachten gezien de noemer.

Deze benaderingen komen uit de theorie van de kettingbreuken. Laat ons dit illustreren in het voorbeeld van π. Als eerste benadering hebben we

$$\pi \approx 3$$

Dit is de benadering die in de bijbel gebruikt wordt.[24]

[23] Zo'n getal noemt men B-glad.

[24] 2 Kronieken 4:2: Hij liet ook de Zee maken, een bekken van gegoten brons, vijf el hoog, met een middellijn van tien el en een omtrek van dertig el [NBV].

$$\pi = 3 + \pi_0$$

met

$$\pi_0 = 0{,}1415926536\cdots$$

Omdat π_0 kleiner dan 1 is kunnen we meer inzicht verkrijgen door de inverse van π_0 te beschouwen.

$$\frac{1}{\pi_0} = 7{,}0625133059\cdots$$

Als we $1/\pi_0$ door 7 benaderen dan verkrijgen we

$$\pi = 3 + \cfrac{1}{\cfrac{1}{\pi_0}} \approx 3 + \frac{1}{7} = \frac{22}{7}$$

Gesterkt door dit succes gaan we door.

$$\pi_1 = \frac{1}{\pi_0} - 7 = 0{,}0625133059$$

en

$$\frac{1}{\pi_1} = 15{,}9965944067$$

Als we $1/\pi_1$ door 16 benaderen dan vinden we

$$\pi = 3 + \cfrac{1}{7 + \cfrac{1}{\cfrac{1}{\pi_1}}} \approx 3 + \cfrac{1}{7 + \cfrac{1}{16}} = \frac{355}{113}$$

Helaas is de benadering $1/\pi_1 \cong 16$ niet zo handig voor de theorie (we zouden dan soms naar beneden en soms naar boven afronden). Daarom gebruiken we $1/\pi_1 \cong 15$. De bijbehorende benadering van π is dan

$$\pi = 3 + \cfrac{1}{7 + \cfrac{1}{\cfrac{1}{\pi_1}}} \approx 3 + \cfrac{1}{7 + \cfrac{1}{15}} = \frac{333}{106}$$

wat natuurlijk een beetje minder goed is.

We zullen nog een stap verder gaan. Definieer

$$\pi_2 = \frac{1}{\pi_1} - 15 = 0.996594406684103$$

en

$$\frac{1}{\pi_2} = 1.00341723101500$$

Als we $1/\pi_2$ door 1 benaderen dan vinden we

$$\pi = 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{\cfrac{1}{\pi_2}}}} \approx 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1}}} = \frac{355}{113} \tag{12.5}$$

We vinden dus dezelfde mooie benadering als voorheen maar dan een stapje later. Dit is een algemeen verschijnsel.

Zo'n schuin aflopende breuk als in (12.5) is niet erg mooi.[25] Daarom wordt er meestal een alternatieve schrijfwijze gebruikt

$$\pi = 3 + \frac{1}{7+}\frac{1}{15+}\frac{1}{1+}\frac{1}{292+}\frac{1}{1+}\cdots$$

De bijbehorende benaderingen (officieel *convergenten* genaamd) zijn:

$$\frac{3}{1}, \frac{22}{7}, \frac{333}{106}, \frac{355}{113}, \frac{103993}{33102}, \frac{104348}{33215}, \ldots$$

Laat ons nu een meer systematische bespreking van kettingbreuken geven. We onderstellen $x \in \mathbb{R}_{>0}$. Dit is om wat vervelende speciale gevallen uit te sluiten.

Definieer

$$a_0 = \lfloor x \rfloor$$
$$x_0 = x - a_0$$

Indien we x_i, a_i bepaald hebben dan bepalen we x_{i+1}, a_{i+1} als volgt.

$$a_{i+1} = \lfloor 1/x_i \rfloor$$
$$x_{i+1} = 1/x_i - a_{i+1}$$

Onderstel eerst dat x een rationaal getal is

$$x = \frac{b}{c}$$

met $\mathrm{ggd}(b, c) = 1$. Voer de deling met rest uit $b = rc + q$. Dan hebben we

$$r = \left\lfloor \frac{b}{c} \right\rfloor = a_0$$

en

$$x_0 = \frac{b}{c} - r = \frac{q}{c}$$

en tenslotte

$$\frac{1}{x_0} = \frac{c}{q}$$

Met andere woorden de overgang $(b, c) \mapsto (c, q)$ correspondeert precies met een stap in het algoritme van Euclides. We weten dat dit algoritme stopt. Dus na een zekere tijd zullen we hebben $x_i = 0$. We hebben dan

$$x = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_i}$$

De rechterzijde (of meer formeel de getallen $a_0, \ldots, a_i$) noemen we de kettingbreuk van x. We schrijven die soms verkort als

$$[a_0, \ldots, a_i]$$

Onderstel dat x geen rationaal getal is. Dan schrijven we

$$x = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_i+}\cdots \tag{12.6}$$

en we noemen de rechterzijde de kettingbreuk van x. Het is weer niet ongebruikelijk om deze verkort te schrijven als

$$[a_0, a_1, \cdots, a_i, \ldots]$$

[25] En werd vroeger ook niet op prijs gesteld door drukkers!

De *convergenten* van x worden gegeven door de breuken

$$\frac{b_i}{c_i} = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_i}$$

Voor een rationaal getal zijn er dus een eindig aantal convergenten waarbij de laatste het rationaal getal zelf is. Voor een irrationaal getal zijn er echter een oneindig aantal convergenten.

Hieronder tonen we aan dat geldt

$$x = \lim_{i\to\infty}\frac{b_i}{c_i}$$

Op deze manier is (12.6) meer dan een formele uitdrukking.

Merk op dat we ook de volgende identiteiten hebben

$$x = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_{i-1}+}\ \frac{1}{a_i + x_i} \tag{12.7}$$

$$x_i = \frac{1}{a_{i+1}+}\frac{1}{a_{i+2}+}\cdots \tag{12.8}$$

We geven nu nog enige voorbeelden. Beschouw $x = \sqrt{3}$. Dan hebben we

i	a_i	x_i
0	1	$\sqrt{3}-1$
1	1	$\frac{\sqrt{3}-1}{2}$
2	2	$\sqrt{3}-1$

We hebben hier gerekend in het lichaam $\mathbb{Q}(\sqrt{3})$.

Omdat $x_2 = x_0$ zien we dat de kettinbreuk $\sqrt{3}$ periodiek is. Maw

$$\sqrt{3} = 1 + \frac{1}{1+}\frac{1}{2+}\frac{1}{1+}\frac{1}{2}+\cdots$$

Dit is een algemeen verschijnsel. We hebben namelijk de volgende stelling

Stelling 12.7.1. *Een niet rationaal reeel getal heeft een periodieke kettingbreuk als en slechts als het van de vorm*

$$a + b\sqrt{n}$$

is met $a, b \in \mathbb{Q}$ en $n \in \mathbb{Z}$ een niet kwadraat.

Hier is nog een ander voorbeeld.

Voorbeeld 12.7.2. De kettingbreuk van e heeft een hele leuke vorm

$$e = 2 + \frac{1}{1+}\frac{1}{2+}\frac{1}{1+}\frac{1}{1+}\frac{1}{4+}\frac{1}{1+}\frac{1}{1+}\frac{1}{6+}\frac{1}{1+}\frac{1}{1+}\frac{1}{8+}\cdots$$

Dit is klassiek maar niet triviaal om te bewijzen. Zie [Hen06].

We zullen nu de convergenten van een kettingbreuk van wat meer van nabij bekijken. We hebben de volgende stelling

Stelling 12.7.3. *Zij*

$$x = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots$$

de kettingbreuk van een reel getal $x > 0$ (eventueel stoppend indien x rationaal is) en zij b_i/c_i de bijbehorende convergenten (met $\mathrm{ggd}(b_i, c_i) = 1$). Er geldt

$$\begin{pmatrix} b_i \\ c_i \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{12.9}$$

Bewijs. Definieer voor $0 \leq j \leq i$

$$\frac{b_i^{(j)}}{c_i^{(j)}} = a_j + \frac{1}{a_{j+1}+} \cdots \frac{1}{a_i}$$

met $\mathrm{ggd}(b_i^{(j)}, c_i^{(j)}) = 1$. Dus

$$\frac{b_i^{(i)}}{c_i^{(i)}} = a_i = \frac{a_i}{1}$$

$$\frac{b_i^{(0)}}{c_i^{(0)}} = \frac{b_i}{c_i}$$

We tonen hieronder met dalende inductie op j aan dat geldt

$$\begin{pmatrix} b_i^{(j)} \\ c_i^{(j)} \end{pmatrix} = \begin{pmatrix} a_j & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{j+1} & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{12.10}$$

Het volstaat nu om $j = 0$ te stellen om (12.9) te verkrijgen.

Nu de inductie. Het geval $j = i$ is OK. Onderstel $j \leq i - 1$.

$$\begin{aligned} \frac{b_i^{(j)}}{c_i^{(j)}} &= a_j + \frac{1}{\frac{b_i^{(j+1)}}{c_i^{(j+1)}}} \\ &= \frac{a_j b_i^{(j+1)} + c_i^{(j+1)}}{b_i^{(j+1)}} \end{aligned}$$

Er geldt

$$\mathrm{ggd}(a_j b_i^{(j+1)} + c_i^{(j+1)}, b_i^{(j+1)}) = \mathrm{ggd}(c_i^{(j+1)}, b_i^{(j+1)}) = 1$$

Dus

$$\begin{aligned} b_i^{(j)} &= a_j b_i^{(j+1)} + c_i^{(j+1)} \\ c_i^{(j)} &= b_i^{(j+1)} \end{aligned}$$

Of in matrix vorm

$$\begin{pmatrix} b_i^{(j)} \\ c_i^{(j)} \end{pmatrix} = \begin{pmatrix} a_j & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b_i^{(j+1)} \\ c_i^{(j+1)} \end{pmatrix}$$

Per inductie bekomen we dus inderdaad (12.10). □

Gevolg 12.7.4. *Er geldt:*

$$\begin{pmatrix} b_i & b_{i-1} \\ c_i & c_{i-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} \tag{12.11}$$

Bewijs. We berekenen

$$\begin{aligned}\begin{pmatrix} b_i \\ c_i \end{pmatrix} &= \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\cdots\begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\cdots\begin{pmatrix} a_{i-1} & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_i \\ 1 \end{pmatrix}\end{aligned}$$

$$\begin{pmatrix} b_{i-1} \\ c_{i-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\cdots\begin{pmatrix} a_{i-1} & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Combineren geeft de gezochte identiteit. □

Gevolg 12.7.5.

$$\det\begin{pmatrix} b_i & b_{i-1} \\ c_i & c_{i-1} \end{pmatrix} = (-1)^{i+1}$$

Bewijs. Dit volgt onmiddelijk uit (12.11). □

Gevolg 12.7.6. *Er geldt de volgende recursiebetrekking*

$$\begin{pmatrix} b_{i+1} \\ c_{i+1} \end{pmatrix} = a_{i+1}\begin{pmatrix} b_i \\ c_i \end{pmatrix} + \begin{pmatrix} b_{i-1} \\ c_{i-1} \end{pmatrix} \tag{12.12}$$

Bewijs. Uit Gevolg 12.7.4 halen we

$$\begin{pmatrix} b_{i+1} & b_i \\ c_{i+1} & c_i \end{pmatrix} = \begin{pmatrix} b_i & b_{i-1} \\ c_i & c_{i-1} \end{pmatrix}\begin{pmatrix} a_{i+1} & 1 \\ 1 & 0 \end{pmatrix}$$

De eerste kolom hiervan is

$$\begin{pmatrix} b_{i+1} \\ c_{i+1} \end{pmatrix} = \begin{pmatrix} b_i & b_{i-1} \\ c_i & c_{i-1} \end{pmatrix}\begin{pmatrix} a_{i+1} \\ 1 \end{pmatrix}$$

wat precies (12.12) is. □

Stelling 12.7.7. *We hebben*

$$b_i/c_i \leq x \leq b_{i+1}/c_{i+1}$$

indien i even is en

$$b_{i+1}/c_{i+1} \leq x \leq b_i/c_i$$

indien i oneven is. (In het rationaal geval: indien de breuken b_i/c_i en b_{i+1}/c_{i+1} gedefinieerd zijn).

Bewijs. Uiteraard is

$$\mathbb{R}_{>0} \to \mathbb{R}_{>0} : y \to a_i + y$$

een stijgende functie. Dus is

$$\mathbb{R}_{>0} \to \mathbb{R}_{>0} : y \to \frac{1}{a_i + y}$$

een dalende functie. Dit geldt dan ook voor

$$\mathbb{R}_{>0} \to \mathbb{R}_{>0} : y \to a_{i-1} + \frac{1}{a_i + y}$$

Dan is

$$\mathbb{R}_{>0} \to \mathbb{R}_{>0} : y \to \frac{1}{a_{i-1} + \dfrac{1}{a_i + y}}$$

weer een stijgende functie. We kunnen deze redering voortzetten tot we uiteindelijk uitkomen dat

$$\phi : \mathbb{R}_{>0} \to \mathbb{R}_{>0} : y \to a_0 + \frac{1}{a_1+} \cdots \frac{1}{a_{i-1}+} \frac{1}{a_i + y}$$

een stijgende functie is indien i even is en anders een dalende functie.

Onderstel nu dat i even is (het geval i oneven is analoog). Vanwege (12.7) en de voorgaande discussie hebben we

$$x = a_0 + \frac{1}{a_1+} \cdots \frac{1}{a_{i-1}+} \frac{1}{a_i + x_i} \geq a_0 + \frac{1}{a_1+} \cdots \frac{1}{a_{i-1}+} \frac{1}{a_i} = \frac{b_i}{c_i}$$

maar ook

$$x = a_0 + \frac{1}{a_1+} \cdots \frac{1}{a_i+} \frac{1}{a_{i+1} + x_{i+1}} \leq a_0 + \frac{1}{a_1+} \cdots \frac{1}{a_i+} \frac{1}{a_{i+1}} = \frac{b_{i+1}}{c_{i+1}} \tag{12.13}$$

wat het gestelde bewijst. □

Opmerking 12.7.8. Indien x irrationaal is dan zijn de ongelijkheden in Stelling 12.7.7 uiteraard strikt (indien niet dan zou x rationaal zijn).

Als we kijken naar het bewijs dan zien we dat de functies die we gedefinieerd hebben ofwel *strikt* stijgend ofwel *strikt* dalend zijn. De bijbehorende ongelijkheden zijn dus ook meestal strikt. Het *enige* geval waarin we geen strikte ongelijkheid hebben is wanneer x rationaal is en $x = b_{i+1}/c_{i+1}$ de laatste convergent is. In het bewijs zien we dat in dat geval $x_{i+1} = 0$ en dus is (12.13) een gelijkheid.

Laat ons de volgende adhoc definitie invoeren

$$\{\!\{a, b\}\!\} = \begin{cases} [a, b] & \text{als } a \leq b \\ [b, a] & \text{als } a > b \end{cases}$$

We kunnen de vorige stelling dan compact formuleren als

$$x \in \left[\!\!\left[\frac{b_i}{c_i}, \frac{b_{i+1}}{c_{i+1}} \right]\!\!\right] \tag{12.14}$$

De lengte van het interval $\{\!\{b_i/c_i, b_{i+1}/c_{i+1}\}\!\}$ wordt gegeven door

$$\left| \frac{b_i}{c_i} - \frac{b_{i+1}}{c_{i+1}} \right| = \left| \frac{b_i c_{i+1} - b_{i+1} c_i}{c_i c_{i+1}} \right| = \frac{1}{c_i c_{i+1}}$$

vanwege Gevolg 12.7.5. Dus

$$\left| x - \frac{b_i}{c_i} \right| \leq \frac{1}{c_i c_{i+1}} < \frac{1}{c_i^2} \tag{12.15}$$

aangezien de c_i strikt stijgend zijn (vanwege de recursiebetrekking)[26].

Gevolg 12.7.9. *Onderstel dat x irrationaal is. Er geldt*

$$\lim_{i \to \infty} \frac{b_i}{c_i} = x$$

Bewijs. Omdat c_i strikt stijgt volgt uit (12.15) dat $|x - b_i/c_i|$ naar nul gaat. Dit is precies wat we nodig hebben. □

We hebben ook het volgende bewezen.

[26]Een speciaal geval is wanneer b_i/c_i de laatste convergent van x is en c_{i+1} dus niet bestaat. Maar dan geldt $|x - b_i/c_i| = 0 < 1/c_i^2$. De bewering is dus nog steeds waar.

Stelling 12.7.10. *Onderstel dat x irrationaal is. Dan bestaan er oneindig veel (b, c) zodat*

$$\left| x - \frac{b}{c} \right| < \frac{1}{c^2}$$

Het gebied van de getaltheorie dat zich bezig houdt met het benaderen van reële getallen door middel van rationale getallen noemt men de *Diophantische approximatietheorie*. Merk op dat indien we c willekeurig kiezen de beste benadering b/c van x slechts aan de ongelijkheid

$$\left| x - \frac{b}{c} \right| \leq \frac{1}{2c}$$

zal voldoen.

Een mooie stelling is de volgende:

Stelling 12.7.11. *(Thue-Siegel-Roth) Onderstel dat x een irrationale wortel is van een polynoom over $\mathbb{Q}$. Dan zijn er voor elke $\epsilon > 0$ slechts eindig veel (b, c) zodat*

$$\left| x - \frac{b}{c} \right| \leq \frac{1}{c^{2+\epsilon}}$$

Deze stelling zegt dat we (12.7.10) niet kunnen verbeteren. Een minder evidente toepassing is de transcendentie van getallen zoals Liouville's getal

$$\sum_{j=1}^{\infty} 10^{-j!}$$

We bewijzen nu een variant op (12.15)

Stelling 12.7.12. *Zij $n, b, c \in \mathbb{N}$, $c \neq 0$ zodat*

$$\left| \sqrt{n} - \frac{b}{c} \right| \leq \frac{1}{c^2} \tag{12.16}$$

Dan geldt

$$|\widetilde{b^2}| \leq 2\sqrt{n} + 1$$

Bewijs. Uit (12.16) halen we

$$\sqrt{n} - \frac{1}{c^2} \leq \frac{b}{c} \leq \sqrt{n} + \frac{1}{c^2}$$

waaruit we achtereenvolgends de volgende begrenzingen halen

$$\begin{aligned} c\sqrt{n} - \frac{1}{c} \leq & b \leq c\sqrt{n} + \frac{1}{c} \\ c^2 n - 2\sqrt{n} + \frac{1}{c^2} \leq & b^2 \leq c^2 n + 2\sqrt{n} + \frac{1}{c^2} \\ -2\sqrt{n} + \frac{1}{c^2} \leq & b^2 - c^2 n \leq 2\sqrt{n} + \frac{1}{c^2} \\ & |b^2 - c^2 n| \leq 2\sqrt{n} + 1 \end{aligned}$$

$b^2 - c^2 n$ is congruent met b^2 modulo n. Er moet dus gelden $|\widetilde{b^2}| \leq |b^2 - c^2 n| \leq 2\sqrt{n} + 1$. □

12.8. **Ontbinden in factoren met behulp van kettingbreuken.** Om t_i's te vinden die aan (12.4) voldoen kunnen we in twee stappen te werk gaan.

(1) Eerst zoeken we naar t_i zodanig dat $\widetilde{t_i^2}$ klein is.
(2) Daarna zoeken we of $\widetilde{t_i^2}$ kan geschreven worden als een produkt van elementen in B (door "trial division").

Kettingbreuken kunnen ons helpen bij de eerste stap. Immers volgens Stelling 12.7.12 geldt $\widetilde{b_i^2} \leq 2\sqrt{n}+1$. Dus het volstaat om $t_i = b_i$ te nemen. Merk op dat we enkel $b_i \bmod n$ moeten berekenen. Dit kunnen we doen met de recursiebetrekking (12.12).

Voorbeeld 12.8.1. Zij $n = 197209$ en $B = \{-1, 2, 3, 5\}$.

i	a_i	x_i	$b_i \bmod n$	$\widetilde{b_i^2}$	ontbinding
0	444	$\sqrt{n} - 444$	444	-73	
1	12	$1/73\sqrt{n} - 432/73$	5329	145	
2	6	$1/145\sqrt{n} - 438/145$	32418	-37	
3	23	$1/37\sqrt{n} - 413/37$	159316	720	$2^4 3^2 5^1$
4	1	$1/720\sqrt{n} - 307/720$	191734	-143	
5	5	$1/143\sqrt{n} - 408/143$	131941	215	
6	3	$1/215\sqrt{n} - 237/215$	193139	-656	
7	1	$1/656\sqrt{n} - 419/656$	127871	33	
8	26	$1/33\sqrt{n} - 439/33$	165232	-136	
9	6	$1/136\sqrt{n} - 377/136$	133218	405	$3^4 5^1$
10	2	$1/405\sqrt{n} - 433/405$	37250	-24	$2^1 3^3 5^1$
11	36	$1/24\sqrt{n} - 431/24$	93755	477	
12	1	$1/477\sqrt{n} - 46/477$	131005	-409	
13	1	$1/409\sqrt{n} - 363/409$	27551	160	$2^5 5$

Visuele inspectie toont aan dat de eerste twee "hits" kunnen gecombineerd worden tot een kwadraat. We hebben dus

$$(159316 \cdot 133218)^2 \equiv 2^4 3^2 5^1 \cdot 3^4 5^1 = (2^2 3^3 5^1)^2 \bmod n$$

We berekenen

$$\gcd(n, 159316 \cdot 133218 + 2^2 3^3 5^1) = 991$$

en we hebben dus inderdaad een niet triviale factor gevonden.

12.9. **De kwadratische zeef.** Het idee achter de kwadratische zeef komt van de zeef van Eratosthenes die we kennen uit de lagere school. Dit is een zeer efficient algoritme om *alle* priemgetallen kleiner dan een gekozen getal N te genereren. De tijdcomplexiteit hiervan is $O(N \log\log N)$.[27] Dus per priemgetal $O(\log\log N \log N)$. Vergelijk dit met het feit (§11.5) dat de complexiteit van de Miller-Rabin test voor een getal van de grootteorde van N gelijk is aan $O((\log N)^{2+\epsilon})$.

Terug naar ons onderwerp. We nemen nu

$$B = \{\text{priemen} \leq w\}$$

We gaan zeven om te zoeken naar een produkt van elementen van B in de rij

$$S = \{t^2 - n \mid t = \lceil\sqrt{n}\rceil, \cdots, \lceil\sqrt{n}\rceil + A\}$$

[27] De $\log\log N$ factor is zo klein dat hij in de praktijk ondetecteerbaar is.

waarbij A een voorafgekozen constante is (naast w).

Het idee is dat we de lijst S aflopen en eerst alle getallen die deelbaar zijn door 2 door 2 delen. Dit herhalen we tot alle getallen oneven zijn. Vervolgens herhalen we dit met 3, 5, etc....

Het probleem is dat detecteren of een getal deelbaar is door p eigenlijk neerkomt op het delen door p en dus een dure operatie is. Het is veel beter dat we op voorhand weten welke elementen van S die door p^α deelbaar zijn voor $\alpha \geq 1$. Dan moeten we enkel deze elementen in beschouwing nemen.

We hebben

$$\begin{aligned} p^\alpha \mid t^2 - n &\iff t^2 \equiv n \bmod p^\alpha \\ &\iff t \text{ is een vierkantswortel van } n \text{ in } \mathbb{Z}/p^\alpha\mathbb{Z} \end{aligned}$$

en

$$\mathbb{Z}/p^\alpha\mathbb{Z} = \{\overline{a_0 + pa_1 + \cdots + p^{\alpha-1}a_{\alpha-1}} \mid a_i \in \{0, \ldots, p-1\}\}$$

Gegeven n, p dan kunnen we $\sqrt{n} \in \mathbb{Z}/p^\alpha\mathbb{Z}$ berekenen (zie hieronder). Dus we kennen dan de $t \bmod p^\alpha$ waarvoor $t^2 - n$ deelbaar is door p^α.

We leggen de berekening van $\sqrt{n} \bmod p^\alpha$ uit met een voorbeeld. Zij $n = 1042387$, $p = 3$. We moeten oplossen

$$(a_0 + 3a_1 + 9a_2 + \cdots)^2 \equiv n \bmod 3^\alpha \tag{12.17}$$

waarbij

$$n = 1+3\cdot 2+3^2\cdot 2+3^3\cdot 2+3^4+3^5\cdot 2+3^6+3^7\cdot 2+3^8\cdot 2+3^9+3^{10}\cdot 2+3^{11}\cdot 2+3^{12} \tag{12.18}$$

Eerst rekenen we modulo 3. Dan reduceert (12.17) tot

$$a_0^2 \equiv 1 \bmod 3$$

en dus $a_0 = 1$ or 2. We onderstellen $a_0 = 1$. Het geval $a_0 = 2$ is analoog.

Nu werken we modulo 9. We moeten dan oplossen

$$(1 + 3a_1)^2 \equiv 1 + 3 \cdot 2 \bmod 9$$

Oftewel

$$1 + 6a_1 \equiv 7 \bmod 9$$

en dus $a_1 = 1$.

Nu werken we modulo 27. We moeten dan oplossen

$$(1 + 3 \cdot 1 + 9a_2)^2 \equiv 1 + 3 \cdot 2 + 9 \cdot 2 \bmod 27$$

oftewel

$$16 + 72a_2 \equiv 25 \bmod 27$$

Dus

$$18a_2 \equiv 9 \bmod 27$$

Na delen door 9 vinden we

$$2a_2 \equiv 1 \bmod 3$$

end dus $a_2 = 2$. Het blijkt dat na de keuze van a_0 er geen verdere keuzes mogelijk zijn. Dus ongeacht α heeft n twee wortels in $\mathbb{Z}/p^\alpha\mathbb{Z}$.

We kunnen nu heel snel de t bepalen zodat $3 \mid t^2 - n$, $9 \mid t^2 - n$, $27 \mid t^2 - n$ etc…. Namelijk

$$\begin{aligned} 3 \mid t^2 - n &\iff t \equiv 1, 2 \bmod 3 \\ 9 \mid t^2 - n &\iff t \equiv 4, 5 \bmod 9 \\ 27 \mid t^2 - n &\iff t \equiv 5, 22 \bmod 27 \end{aligned}$$

etc….

Opmerking 12.9.1. Als factorbasis kunnen we in feite nemen

$$B' = \left\{ \text{priemen } p \leq w, \left(\frac{n}{p}\right) = 1 \right\}$$

want indien

$$t^2 \equiv n \bmod p$$

dan geldt

$$\left(\frac{n}{p}\right) = 1$$

Opmerking 12.9.2. (Lezen) Zoals al gezegd is bovenstaande berekening in feite onafhankelijk van α. Na nog wat verder rekenen[28] vinden we dat de twee wortels van n gegeven worden door

$$1 + 3 + 3^2 \cdot 2 + 3^4 + 3^5 \cdot 2 + 3^8 \cdot 2 + 3^{10} + 3^{11} \cdot 2 + 3^{12} \cdot 2 + \cdots \tag{12.19}$$

en

$$2 + 3 + 3^3 \cdot 2 + 3^4 + 3^6 \cdot 2 + 3^7 \cdot 2 + 3^9 \cdot 2 + 3^{10} + \cdots \tag{12.20}$$

We kunnen dus eigenlijk doen alsof $\alpha = \infty$ en (12.19-12.20) als oneindige reeksen beschouwen in p (in dit geval $p = 3$). Dergelijke reeksen vormen een ring: de ring der "p-adische getallen". Notatie: $\mathbb{Z}_p$. Elementen van deze ring zijn oneindige reeksen

$$a_0 + a_1 p + a_2 p^2 + \cdots$$

In feite kunnen we p-adische getallen beschouwen als getallen met basis p:

$$(\cdots a_2 a_1 a_0)_p \tag{12.21}$$

die (eventueel) "oneindig naar links doorlopen". Het is gemakkelijk in te zien dat de traditionele lagere school algoritmen voor optellen en vermenigvuldigen ("onder elkaar schrijven") goed gedefinieerd zijn voor dergelijke oneindig naar links doorlopende getallen.

Indien de reeks eindig is (i.e. $a_i = 0$ voor $i \gg 0$) dan is (12.21) gewoon de voorstelling van een geheel getal met basis p. Dus $\mathbb{Z} \subset \mathbb{Z}_p$. In $\mathbb{Z}/p^\alpha\mathbb{Z}$ maken we p^α gelijk aan nul. We hebben $\mathbb{Z}/p^\alpha\mathbb{Z} = \mathbb{Z}_p/p^\alpha\mathbb{Z}_p$.

De quotientenring van $\mathbb{Z}_p$ is $\mathbb{Q}_p$: het *lichaam* der p-adische getallen. Dit zijn de getallen van de vorm

$$a_h p^{-h} + \cdots + a_{-2} p^{-2} + a_{-1} p^{-1} + a_0 + a_1 p + a_2 p^2 + \cdots$$

Of geschreven als getal met basis p:

$$(\cdots a_2 a_1 a_0, a_{-1} a_{-2} \cdots a_{-h})_p$$

[28] Of het gebruik van een computerprogramma zoals het open source "sage": `http://www.sagemath.org`.

Merk op dat in tegenstelling met de reele getallen zo'n getal niet oneindig naar rechts doorloopt. Inderdaad voor getallen die oneindig in de twee richtingen zouden doorlopen is de vermenigvuldiging niet gedefinieerd (oefening).

De p-adische rationale getallen zijn uitgerust met een norm:

$$|a_h p^{-h} + \cdots + a_{-2}p^{-2} + a_{-1}p^{-1} + a_0 + a_1 p + a_2 p^2 + \cdots| \stackrel{\text{def}}{=} p^h$$

waarvoor de lichaamsoperaties continue zijn. Bovendien is $\mathbb{Q}_p$ compleet voor deze norm (alle Cauchy rijen convergeren).

We zien dus dat er een grote analogie is tussen $\mathbb{R}$ en $\mathbb{Q}_p$. Er bestaat dan ook een analogon over $\mathbb{Q}_p$ van de analyse over $\mathbb{R}$: de zogenaamde *p-adische analyse*. Sommige aspecten van de p-adische analyse zij echter eenvoudiger. Dit komt omdat in $\mathbb{Q}_p$ de volgende sterke ongelijkheid geldt

$$|x + y| \leq \max(|x|, |y|) \tag{12.22}$$

terwijl in $\mathbb{R}$ slechts geldt

$$|x + y| \leq |x| + |y|$$

Uit (12.22) kan men bijvoorbeeld het volgende resultaat afleiden: een reeks over $\mathbb{Q}_p$

$$\sum_{n=0}^{\infty} x_n$$

is convergent als en slechts als $|x_n| \to 0$. Het corresponderende resultaat over $\mathbb{R}$ is natuurlijk vals.

13. Kettingbreuken extra's

13.1. **De benaderingseigenschap.** We gaan nog wat dieper in op de goede benaderingseigenschappen van convergenten.

We hebben reeds gezien dat indien b/c een convergent is van $x > 0$ dat b/c dan een goede rationale benadering is van x. Hoe goed? Om dit te kwantificeren kunnen we kijken naar het verschil $|x - b/c|$ maar dat is niet helemaal eerlijk omdat dit beïnvloed wordt door de grootte van c. Een betere maatstaf is

$$|cx - b| \tag{13.1}$$

De minimale waarde van (13.1) wordt gegeven door cx af te ronden op een geheel getal. Dus

$$0 \leq |cx - b| < 0.5$$

Laat ons de gulden snede snede beschouwen

$$\phi = \frac{1 + \sqrt{5}}{2} = 1.618\ldots$$

c	$\boxed{1}$	$\boxed{2}$	$\boxed{3}$	4	$\boxed{5}$	6	7
$\min_b \lvert cx - b\rvert$	0.382	0.236	0.146	0.472	0.090	0.292	0.326
b	2	3	5	6	8	10	11

c	$\boxed{8}$	9	10	11	12	$\boxed{13}$	14
$\min_b \lvert cx - b\rvert$	0.056	0.438	0.180	0.202	0.416	0.034	0.347
b	13	15	16	18	19	21	23

We hebben met een $\boxed{}$ aangegeven welke benaderingen beter waren dan alle voorgaande. We merken dat dit zich precies voordoet wanneer c een Fibonacci getal is.

De bijbehorende b is dan ook een Fibonacci getal. De theorie der kettingbreuken geeft een verklaring voor dit fenomeen.

Men toont gemakkelijk aan dat de kettingbreuk van ϕ gelijk is aan

$$1+\frac{1}{1+}\frac{1}{1+}\frac{1}{1+}\frac{1}{1+}+\cdots$$

en de convergenten worden gegeven door de verhoudingen van de opeenvolgende Fibonacci getallen[29]

$$\frac{2}{1},\frac{3}{2},\frac{5}{3},\frac{8}{5},\ldots$$

Het lijkt er dus op dat een benadering door een convergent beter is dan de beste benaderingen met strikt kleinere noemer. Dit wordt geformalizeerd in de volgende stelling.

Stelling 13.1.1. *Zij $x>0$, $c>2$. Dan is b/c met $\gcd(b,c)=1$ een convergent van x als en slechts als voor alle $0<c'<c$ en alle b' we hebben*

$$|c'x-b'|>|cx-b|$$

We hebben het volgende voorbereidende lemma nodig.

Lemma 13.1.2. *Zij $x>0$ en zij x_i zoals in de §12.7 van de cursus. Dan geldt*

$$x_i=-\frac{c_ix-b_i}{c_{i-1}x-b_{i-1}} \tag{13.2}$$

waarbij zoals gebruikelijk b_i/c_i de convergenten van x zijn.

Bewijs. We hebben

$$x=a_0+\frac{1}{a_1+}\cdots\frac{1}{a_i+x_i}$$

Als we het bewijs van Stelling 12.7.3 nalezen dan zien we dat $a_0,\ldots,a_i$ in feite geen gehele getallen hoeven te zijn. Maar (12.9) geldt dan slechts op een factor na (we hebben nu immers geen ggd=1 conditie meer).

We krijgen dan dat er een u bestaat zodat

$$\begin{aligned}u\begin{pmatrix}x\\1\end{pmatrix}&=\begin{pmatrix}a_0&1\\1&0\end{pmatrix}\cdots\begin{pmatrix}a_{i-1}&1\\1&0\end{pmatrix}\begin{pmatrix}a_i+x_i&1\\1&0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\\&=\begin{pmatrix}a_0&1\\1&0\end{pmatrix}\cdots\begin{pmatrix}a_{i-1}&1\\1&0\end{pmatrix}\begin{pmatrix}a_i&1\\1&0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}+x_i\begin{pmatrix}a_0&1\\1&0\end{pmatrix}\cdots\begin{pmatrix}a_{i-1}&1\\1&0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\\&=\begin{pmatrix}b_i\\c_i\end{pmatrix}+x_i\begin{pmatrix}b_{i-1}\\c_{i-1}\end{pmatrix}\end{aligned}$$

We krijgen dan

$$x=\frac{b_i+x_ib_{i-1}}{c_i+x_ic_{i-1}}$$

waaruit we eenvoudig (13.2) halen. □

Omdat $0\leq x_i<1$ krijgen we dus ondermeer

$$|c_ix-b_i|<|c_{i-1}x-b_{i-1}| \tag{13.3}$$

M.a.w. de rij $|c_ix-b_i|$ is strikt dalend.

[29]Normaal zou 1/1 ook een convergent zijn maar die heeft dezelfde noemer als de volgende convergent 2/1. Om de discussie te versimpelen laten we 1/1 buiten beschouwing.

Lemma 13.1.3. *Zij $x > 0$, $b, c \in \mathbb{N}$, $\gcd(b,c) = 1$. Onderstel dat er een i is zodanig dat $c_{i-1} \leq c < c_i$ Dan geldt $\forall j \geq i-1$*

$$|c_{i-1}x - b_{i-1}| \leq |cx - b|$$
$$|c_j x - b_j| < |cx - b| \qquad \forall j \geq i$$

De eerste ongelijkheid is strikt, tenzij $(b,c) = (b_{i-1}, c_{i-1})$.

Bewijs. Gebruikmakende van Gevolg 12.7.5 vinden we dat er $u, v \in \mathbb{Z}$ bestaan zodat

$$\begin{pmatrix} b_i & b_{i-1} \\ c_i & c_{i-1} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix}$$

Het is duidelijk dat $v \neq 0$. Verder volgt uit de ongelijkheid $c_i > c$ dat u, v niet beide hetzelfde teken kunnen hebben.

Links vermenigvuldigen met de rijvector $(-1x)$ geeft

$$(c_i x - b_i)u + (c_{i-1}x - b_{i-1})v = cx - b$$

Uit (13.2) volgt dat $c_i x - b_i$ en $c_{i-1}x - b_{i-1}$ ook verschillende tekens hebben. M.a.w. $(c_i x - b_i)u$ en $(c_{i-1}x - b_{i-1})v$ hebben dezelfde tekens. Er geldt dus

$$|c_i x - b_i||u| + |c_{i-1}x - b_{i-1}||v| = |cx - b|$$

waaruit volgt

$$|c_{i-1}x - b_{i-1}| \leq |cx - b|$$

(immers $|v| \geq 1$). Gelijkheid is enkel mogelijk indien $|u| = 0$, $|v| = 1$. M.a.w. indien $(b,c) = (b_{i-1}, c_{i-1})$.

De conclusie $|c_j x - b_j| < |cx - b|$, $j \geq i$ volgt uit het feit dat $|c_j x - b_j|$ een strikt dalende rij is (zie (13.3)). □

Gevolg 13.1.4. *Zij $x > 0$, $b, c \in \mathbb{N}$, $\gcd(b,c) = 1$. Onderstel dat er een j is zodanig dat $c < c_j$ Dan geldt*

$$|c_j x - b_j| < |cx - b|$$

Bewijs. Zij i zodanig dat $c_{i-1} \leq c < c_i$. Dan geldt $i \leq j$ en de conclusie volgt uit Lemma 13.1.3. □

Bewijs van Stelling 13.1.1. Onderstel eerst dat b/c een convergent is en zij $0 < c' < c$, b' willekeurig. Vanwege Gevolg 13.1.4 volgt onmiddelijk

$$|cx - b| < |c'x - b'|$$

(de rol van (b', c') wordt hier gespeeld door (b,c) in Gevolg 13.1.4).

Omgekeerd onderstel dat b/c geen convergent is. We tonen aan dat er $0 < c' < c$, b' bestaan zodat $|c'x - b'| < |cx - b|$.

(1) Onderstel nu dat c wel de noemer is van een convergent maar b niet de teller. Zij b''/c de eigenlijke convergent. We hebben dan $|cx - b''| < 1/c$ en dus $|cx - b| \geq \big||cx - b''| - |b - b''|\big| > 1 - 1/c > 1/2$.
    We nemen nu $c' = 1$ en b' de afgeronde waarde van x.

(2) We mogen nu onderstellen dat c niet de noemer is van een convergent. Onderstel eerst dat er een i bestaat zodanig dat $c_{i-1} \leq c < c_i$. Dan is de eerste ongelijkheid strikt. Uit Lemma 13.1.3 volgt dat we $(c', b') = (c_{i-1}, b_{i-1})$ kunnen nemen.

(3) Als i niet bestaat dan is x rationaal en c is strikt groter dan de noemer van de laatse convergent b_i/c_i. Dan geldt $|c_i x - b_i| = 0$. Indien $|cx - b| = 0$ dan is b/c een convergent wat we uitgesloten hadden. Dus ook in dit geval $|c'x - b'| < |cx - b|$. □

13.2. **Herkennen van convergenten.** We bewijzen de volgende stelling

Stelling 13.2.1. *Zij $x > 0$. Indien $b, c \in \mathbb{N}$, $\gcd(b, c) = 1$ en*

$$|cx - b| < \frac{1}{2c}$$

dan is b/c een convergent van x.

Bewijs. We gaan Stelling 13.1.1 toepassen. Helaas is $c = 1$ een speciaal geval. Onderstel $|x - b| < 1/2$. Indien $b \leq x$ dan is $(a_0, 1) = (b_0, c_0) = (b, 1)$. Indien $b > x$ dan hebben we $(a_0, 1) = (b_0, c_0) = (b - 1, 1)$ en $x_0 = x - (b - 1) > 1/2$. Dus $a_1 = 1$. De tweede convergent is dan

$$a_0 + \frac{1}{a_1} = b - 1 + \frac{1}{1} = \frac{b}{1}$$

Onderstel nu $c > 1$. Zij $0 < c' < c$, b' willekeurig. We moeten aantonen $|cx - b| < |c'x - b'|$. Dit steunt op de volgende berekening

$$\begin{aligned} 1 &\leq |bc' - cb'| \\ &= |(c(c'x - b') - (cx - b)c'| \\ &\leq c|c'x - b'| + |cx - b|c' \\ &< c|c'x - b'| + c'/2c \end{aligned}$$

Dus

$$|c'x - b'| > \frac{1}{c}(1 - c'/2c) > \frac{1}{c}(1 - c/2c) = \frac{1}{2c} > |cx - b|$$

□

13.3. **De Wiener attack op RSA in geval van kleine decryptieexponent.** Zij (n, p, q, d, e) parameters van een RSA communicatie. Dus p, q zijn priem en $de \cong 1 \bmod t$ met $t = \mathrm{kgv}(p - 1, q - 1)$.

Zij $k = \gcd(p - 1, q - 1)$. Dan geldt $kt = \phi(n)$. k zal in het algemeen klein zijn. p, q worden immers random gekozen en de kans dat twee random gekozen getallen deelbaar zijn door een priemgetal h is $1/h^2$.

Er geldt

$$de - st = 1 \tag{13.4}$$

voor zekere s. Dus ook $kde - s\phi(n) = k$. We kennen $\phi(n)$ niet maar we kunnen het benaderen door n. Dus

$$kde \cong sn$$

en ook

$$\frac{ke}{n} \cong \frac{s}{d}$$

De observatie van Wiener is dat indien d klein is dat dan s/d zo dicht bij ke/n ligt dat het een convergent is. We kunnen dan s/d en dus hopelijk d vinden door de kettingbreuken te beschouwen van e/n, $2e/n$, etc….

We gaan dit nu preciseren. We maken de volgende additionele hypothesen

$$q < p < 2q$$

M.a.w. het aantal bits van p en q verschilt ten hoogste 1.

$$d < n^{1/4}/3$$

M.a.w. d is "klein" t.o.v. n.

$$0 \leq d, e < t$$

d, e worden meestal zo gekozen.

We hebben

$$k = \phi(n)/t \leq \phi(n)/(p-1) = q - 1$$

Uit (13.4) halen we ook

$$st < st + 1 = de$$

waaruit we halen $s < d(e/t) < d < t$.

Aangezien $pq > q^2$ geldt $q < \sqrt{n}$. Dus

$$0 < n - \phi(n) = p + q - 1 < 3q - 1 < 3q < 3\sqrt{n}$$

We maken nu de volgende berekening

$$\begin{aligned}
\left|\frac{ke}{n} - \frac{s}{d}\right| &= \left|\frac{ked - ns}{dn}\right| \\
&= \left|\frac{k + kst - ns}{dn}\right| \\
&= \left|\frac{k + s(\phi(n) - n)}{dn}\right| \\
&\leq \left|\frac{s(n - \phi(n))}{dn}\right| \qquad \text{(aangezien } k \leq q - 1 \leq p + q - 1 = n - \phi(n)\text{)} \\
&< \frac{3s}{d\sqrt{n}} \\
&< \frac{3d}{d \cdot 9d^2} \qquad \text{(aangezien } \sqrt{n} > 9d^2,\ s < d\text{)} \\
&= \frac{1}{3d^2}
\end{aligned}$$

Uit Stelling 13.2.1 volgt dus dat s/d inderdaad een convergent is van ke/n.

14. Elliptische krommen

14.1. **Basis definities.** Hieronder is K een lichaam. We zullen de volgende voorbeelden beschouwen.

- $K = \mathbb{F}_q$ (voor cryptografie).
- $K = \mathbb{R}, \mathbb{C}$ (voor intuitie).
- $K = \mathbb{Q}$ (voor getaltheorie).

Elliptische krommen kan je maar goed bestuderen mits een flinke brok algebraische meetkunde. Niettegenstaande dit zullen we toch hieronder proberen om een aantal aspecten te belichten.

Een *elliptische kromme* E over K is een equivalentieklasse (nog te definiëren) van vergelijkingen van de vorm

$$F = \alpha_1 y^2 + \alpha_2 xy + \alpha_3 y - \alpha_4 x^3 - \alpha_5 x^2 - \alpha_6 x - \alpha_7 \tag{14.1}$$

met $(\alpha_i)_i \in K$ en $\alpha_1 \neq 0$, $\alpha_4 \neq 0$ zodat $\forall (u,v) \in \bar{K}^2$ die voldoet aan $F(u,v) = 0$ we hebben

$$\left(\frac{\partial F}{\partial x}(u,v), \frac{\partial F}{\partial y}(u,v)\right) \neq (0,0)$$

Of anders geschreven: $\forall (u,v) \in \bar{K}^2$

$$\left(F(u,v), \frac{\partial F}{\partial x}(u,v), \frac{\partial F}{\partial y}(u,v)\right) \neq (0,0,0)$$

We noemen F de *vergelijking* van E. Ze wordt meestal geschreven in de vorm

$$\alpha_1 y^2 + \alpha_2 xy + \alpha_3 y = \alpha_4 x^3 + \alpha_5 x^2 + \alpha_6 x + \alpha_7$$

Opmerking 14.1.1. De monomen in x en y die in (14.1) voorkomen worden als volgt bepaald: geef x graad 2 en y graad 3. Dan bevat (14.1) precies de monomen van graad ≤ 6.

We zeggen dat F, G zoals in (14.1) equivalent zijn indien je G uit F kan verkrijgen door met een van nul verschillend element van K te vermenigvuldigen en een lineaire coordinatentransformatie van de vorm

$$x = px' + q$$
$$y = ry' + sx' + t$$

met $p, r \neq 0$ uit te voeren. Het is eenvoudig na te gaan dat zo'n coordinatentransformatie de algemene vorm (14.1) behoudt.

Als we het eenvoudigste geval $x = px'$, $y = ry'$ beschouwen dan wordt de vergelijking van de vorm

$$\alpha_1 r^2 (y')^2 + \cdots = \alpha_4 p^3 (x')^3 + \cdots$$

Dus als we nemen $r = p$ en $p = \alpha_1/\alpha_4$ dan krijgen we

$$\alpha_1^3/\alpha_4^2 (y')^2 + \cdots = \alpha_1^3/\alpha_4^2 (x')^3 + \cdots$$

en dus op equivalentie na

$$(y')^2 + \cdots = (x')^3 + \cdots$$

Over elk lichaam kunnen we dus de vergelijking van een elliptische kromme in de volgende vorm gieten

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6 \tag{14.2}$$

Het idee achter de indexering van de a's is dat indien we a_i graad i geven (met x, y zoals in Opmerking 14.1.1) dan is de vergelijking (14.2) homogeen van graad 6.

Afhankelijk van de karakteristiek van K kunnen we de vergelijking van een elliptische kromme verder vereenvoudigen.

Stelling 14.1.2. *De vergelijking van een elliptische kromme E kan in de volgende vorm gegoten worden*

(1) *Indien* $\operatorname{char} K \neq 2, 3$*:*

$$y^2 = x^3 + ax + b \tag{14.3}$$

zodat $f(x) = x^3 + ax + b$ geen meervoudige wortels heeft. Dit laatste is equivalent met $4a^3 + 27b^2 \neq 0$.[30]

[30]Voor elk polynoom $f(x) = x^n + ax^{n-1} + bx^{n-2} + \cdots$ is er een polynoom $D(a, b, \cdots)$ in de coefficienten van f zodat $D = 0$ equivalent is met het hebben van een meervoudige wortel

(2) *(lezen) Indien* $\operatorname{char} K = 2$*. Nu zijn er twee mogelijkheden*
   (a) $y^2 + cy = x^3 + ax + b$ *met* $c \neq 0$*.*
   (b) $y^2 + xy = x^3 + ax^2 + b$ *met* $b \neq 0$*.*
(3) *(lezen) Indien* $\operatorname{char} K = 3$*:*

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

zodanig dat $f(x) = x^3 + ax^2 + bx + c$ *geen meervoudige wortels heeft over* $\bar{K}$*.*

Bewijs. Dit nemen we aan. □

Opmerking 14.1.3. Als we hieronder een vergelijking schrijven van een elliptische krommen dan zullen we meestal de eenvoudige vorm (14.3) nemen. Dit is gewoon om schrijfwerk te besparen. Meestal gaan de argumenten ook door in karakteristieek 2 en 3.

Definitie 14.1.4. Zij $K \subset L$ een lichaamsuitbreiding en zij E een elliptische kromme over K. Dan definieren we

$$E_{\text{aff}}(L) = \{(u, v) \in L^2 \mid F(u, v) = 0\}$$

en

$$E(L) = E_{\text{aff}}(L) \cup \{P_\infty\}$$

Hierbij noemen we P_∞ het *punt op oneindig*.

Opmerking 14.1.5. Het punt op oneindig kan zichtbaar gemaakt worden door over te gaan naar homogene coordinaten $x = X/Z$, $y = Y/Z$. De vergelijking wordt dan

$$Y^2 Z = X^3 + aXZ^2 + bZ^3$$

Deze vergelijking heeft $(0, 1, 0)$ als oplossing. Dit is P_∞. Merk op dat dit het enige punt is met $Z = 0$. Er is dus slechts 1 punt op oneindig!

Definitie 14.1.6. Zij $P = (u, v) \in E_{\text{aff}}(L)$. Dan wordt de *raaklijn* aan E in P gedefinieerd als de rechte met vergelijking

$$\frac{\partial F}{\partial x}(P)(x - u) + \frac{\partial F}{\partial y}(P)(y - v) = 0 \tag{14.4}$$

Indien $P = P_\infty$ dan kunnen we de raaklijn aan P als de rechte op oneindig definieren. Dit is de rechte in het projectieve vlak met vergelijking $Z = 0$.

Opmerking 14.1.7. Indien $K = L = \mathbb{R}$ dan is het eenvoudig in de zien dat de rechte met vergelijking (14.4) precies de raaklijn is aan $\{F = 0\}$ in de differentiaalmeetkundige zin.

14.2. **Elliptische krommen over** $\mathbb{R}$**.** In de vergelijking

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

heeft het polynoom $f(x) = x^3 + ax + b$ een of drie reele wortels. Bovendien gaat $f(x)$ naar $+\infty$ als $x \to +\infty$.

---

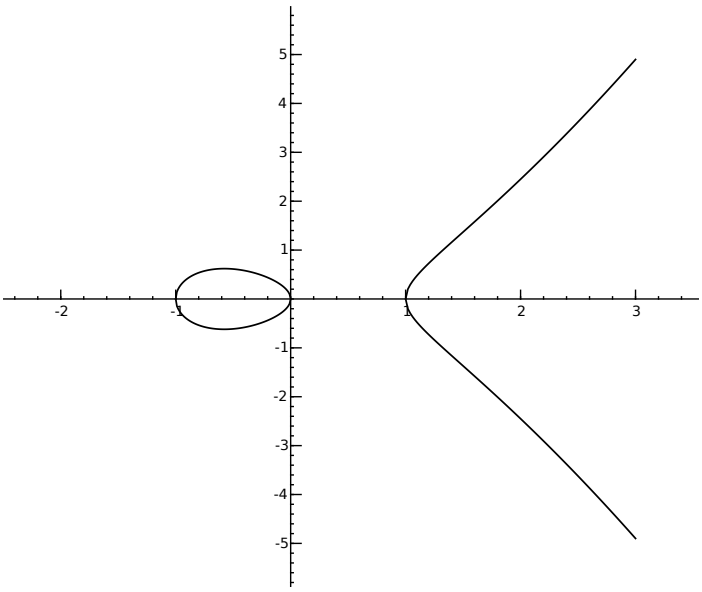
van f. Dit polynoom D is de zogenaamde discriminant van f. De discriminant kan bijvoorbeeld gedefinieerd worden als de resultante van f en f'. In ons geval is deze gelijk aan $4a^3 + 27b^2$.

(1) Beschouw eerst het geval waarbij er drie reele wortels x_1, x_2, x_3 zijn. Dan geldt

$$f(x) \geq 0 \iff x \in [x_1, x_2] \cup [x_3, +\infty[$$

Dus indien (x, y) in $E_{\text{aff}}(\mathbb{R})$ dan $x \in [x_1, x_2] \cup [x_3, +\infty[$ en omgekeerd indien $x \in [x_1, x_2] \cup [x_3, +\infty[$ dan zijn er een of twee (meestal twee) corresponderende punten $(x, y) \in E(\mathbb{R})$.

Hier is de tekening van $y^2 = x^3 - x$.

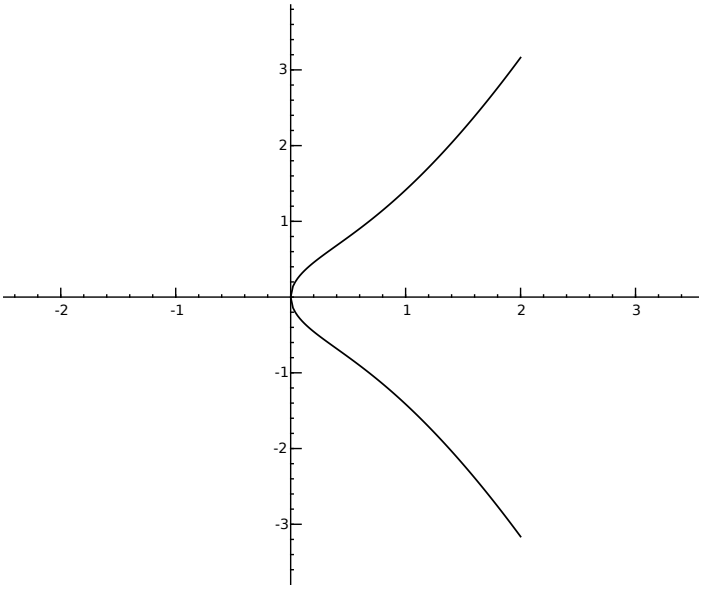


(2) We beschouwen nu het geval waarbij er precies een reele wortel x_1 is. Dan geldt

$$f(x) \geq 0 \iff x \in [x_1, +\infty[$$

Dus indien (x, y) in $E_{\text{aff}}(\mathbb{R})$ dan $x \in [x_1, +\infty[$ en omgekeerd indien $x \in [x_1, +\infty[$ dan zijn er een of twee (meestal twee) corresponderende punten $(x, y) \in E(\mathbb{R})$.

Hier is de tekening van $y^2 = x^3 + x$.



14.3. **Elliptische krommen over $\mathbb{C}$.** Om een echt inzicht te krijgen in elliptische krommen moeten we een complex grondlichaam beschouwen. Beschouw weer de vergelijking

$$y^2 = x^3 + ax + b \tag{14.5}$$

waarbij zowel x, y als de scalairen a, b complexe getallen zijn. Schrijf

$$\begin{aligned} x &= x_0 + ix_1 \\ y &= y_0 + iy_1 \\ a &= a_0 + ia_1 \\ b &= b_0 + ib_1 \end{aligned}$$

Dan komen de oplossingen van (14.5) in $\mathbb{C}$ overeen met de oplossingen in $\mathbb{R}$ van het volgende stelsel.

$$\begin{aligned} -y_1^2 + y_0^2 &= b_0 - a_1x_1 + a_0x_0 - 3x_0x_1^2 + x_0^3 \\ 2y_0y_1 &= b_1 + a_0x_1 + a_1x_0 - x_1^3 + 3x_0^2x_1 \end{aligned}$$

Aangezien we 2 vergelijkingen met 4 onbekenden hebben is het redelijk te onderstellen dat we op een of andere manier te maken hebben met een oppervlak.

14.4. **Dubbelperiodieke functies.** De trigoniometrische functies zijn periodiek met period 2π.

$$\begin{aligned} \sin(z + 2\pi) &= \sin(z) \\ \cos(z + 2\pi) &= \cos(z) \end{aligned}$$

en deze periodiciteit geldt zelf voor z complex. In dit geval is 2π de enige periode. Met andere woorden indien bijvoorbeeld geldt

$$\sin(z+L)=\sin(z)$$

dan is L een geheel veelvoud van 2π. De trigoniometrische functies zijn *enkelvoudig periodieke functies*.

Om diverse redenen is het interessant om *dubbel* periodieke functies te bestuderen. Dit wil zeggen meromorfe functies $f:\mathbb{C}\to\mathbb{C}$ (meromorf=zonder essentiele singulariteiten)[31] zodat er twee lineair onafhankelijke (over $\mathbb{R}$) elementen ω_1, ω_2 van $\mathbb{C}$ zijn zodat

$$f(z+\omega_1)=f(z)$$
$$f(z+\omega_2)=f(z)$$

We kunnen dit meer intrinsiek formuleren als volgt Een *rooster* $L\subset\mathbb{C}$ is een abelse deelgroep van $(\mathbb{C},+)$ voortgebracht door twee generatoren ω_1, ω_2, linear onafhankelijk over $\mathbb{R}$. Een dubbel periodieke functie is een meromorfe functie zodanig dat er een rooster L bestaat zodat

$$\forall l: f(z+l)=f(z)$$

Opmerking 14.4.1. Waarom meromorf? Onderstel dat $f:\mathbb{C}\to\mathbb{C}$ dubbelperiodiek is en *analytisch*. Zij P het gesloten parallellogram opgespannen door ω_1 en ω_2. Aangezien elk element van $\mathbb{C}$ kan geschreven worden als $l+z$ met $l\in L$ en $z\in P$ hebben we

$$\max_{z\in\mathbb{C}}|f(z)|=\max_{z\in P}|f(z)|<\infty$$

(aangezien P compact is). Dus f is analytisch en begrensd op het volledig complex vlak. Dan moet f constant zijn.

Het natuurlijk niet duidelijk dat dubbelperiodieke functies bestaan. Hier is een beroemd voorbeeld. De zogenaamde Weierstraß $\mathcal{P}$-functie. We onderstellen steeds dat $L\subset\mathbb{C}$ een rooster is.

$$\mathcal{P}(z)=\frac{1}{z^2}+\sum_{l\in L'}\left(\frac{1}{(z-l)^2}-\frac{1}{l^2}\right) \tag{14.6}$$

met $L'=L-\{0\}$.

Opmerking 14.4.2. In feite zouden we graag schrijven

$$\mathcal{P}(z)=\sum_{l\in L}\frac{1}{(z-l)^2} \tag{14.7}$$

Dit zou duidelijk dubbel periodiek zijn. Helaas is de rechterzijde van (14.7) divergent. We corrigeren dit door er een andere divergente som van af te trekken (renormalizatie!)

$$\sum_{l\in L'}\frac{1}{l^2}$$

Stelling 14.4.3. (1) *De rechterzijde van* (14.6) *is absoluut convergent op* $\mathbb{C}-L$ *en definieert een analytische functie op* $\mathbb{C}-L$ *die meromorf is op* $\mathbb{C}$.
(2) $\mathcal{P}(z+l)=\mathcal{P}(z)$ *voor* $l\in L$.

[31] Strikt genomen is een meromorfe functie maar een functie op $\mathbb{C}-P$ waarbij P de verzameling polen is.

(3) $\mathcal{P}'(z) = \sum_{l \in L} \frac{-2}{(z-l)^3}$. *Dit is absoluut convergent op* $\mathbb{C} - L$.

(4) $\mathcal{P}$ *voldoet aan de volgende differentiaalvergelijking.*

$$\mathcal{P}'(z)^2 = 4\mathcal{P}(z)^3 - g_2\mathcal{P}(z) - g_3$$

waarbij

$$g_2 = g_2(L) = 60 \sum_{l \in L'} \frac{1}{l^4}$$
$$g_3 = g_3(L) = 140 \sum_{l \in L'} \frac{1}{l^6}$$

Het bewijs van deze stelling berust op het volgende resultaat uit de complexe analyse.

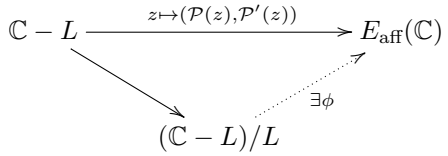
Stelling 14.4.4. *Zij* D *een open deel van* $\mathbb{C}$ *en zij* $(f_n)_n$ *analytische functies op* D *die uniform convergeren op compacte delen naar een functie* f. *Dan is* f *analytisch en verder convergeren de afgeleiden* $(f_n^{(k)})_n$ *uniform op compacte delen naar* $f^{(k)}$.

Hiermee kunnen we dus het gedrag van sommen als (14.6) onder controle krijgen. Alhoewel het bewijs van Stelling 14.4.3 niet echt moeilijk is vereist het toch redelijk wat schrijfwerk. We slaan het dus over. Zie bijvoorbeeld [Sil92, §VI.3].

We definieren nu de elliptische kromme $E/\mathbb{C}$ met vergelijking.

$$y^2 = 4x^3 - g_2x - g_3 \tag{14.8}$$

Dan hebben we het volgende commutatieve diagram



Men kan nu aantonen dat de afbeelding ϕ een bijectie is

$$(\mathbb{C} - L)/L = \mathbb{C}/L - L/L \cong E_{\text{aff}}(\mathbb{C})$$

Als we dan definieren $\phi(L/L) = P_\infty$ dan bekomen we uiteindelijk een bijectie

$$\mathbb{C}/L \cong E(\mathbb{C})$$

Nu is $\mathbb{C}/L$ op natuurlijke manier een torus. Een manier om dat in te zien is $\mathbb{C}/L$ te beschouwen als het parallellogram P uit Opmerking 14.4 met overstaande zijden geidentificeerd.

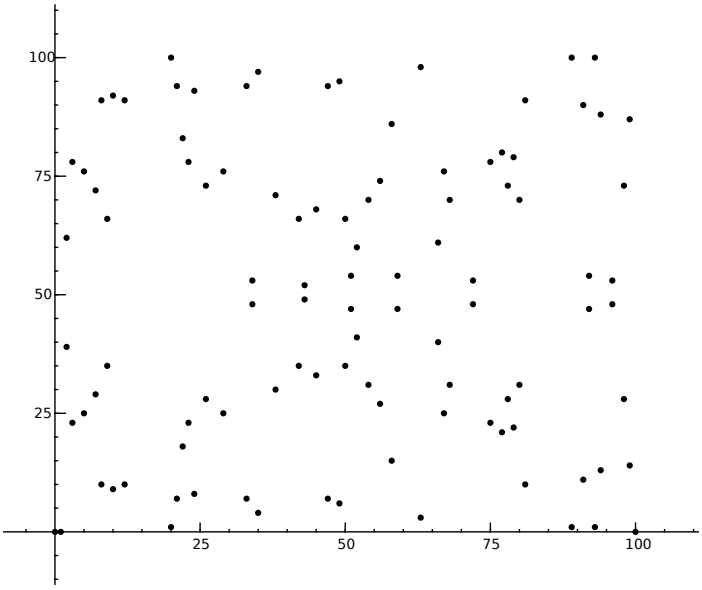
Omgekeerd hebben we het volgende

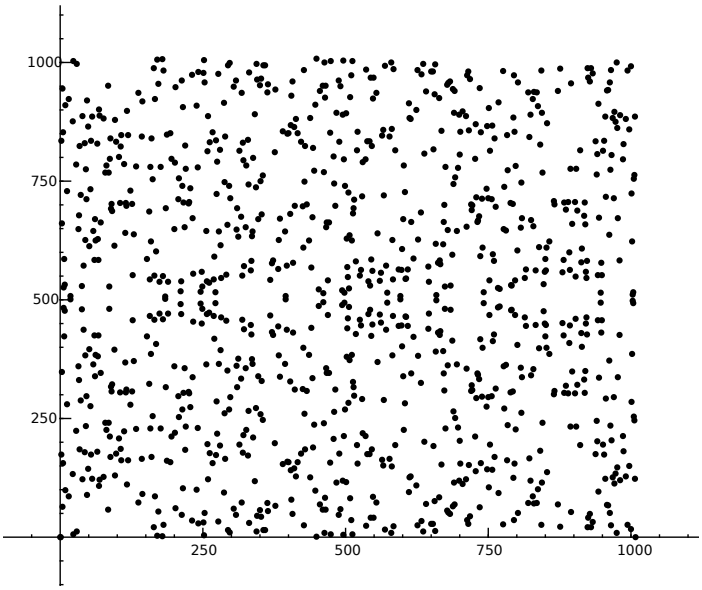
Stelling 14.4.5. *Elke elliptische kromme over* $\mathbb{C}$ *heeft een vergelijking van de vorm* (14.8) *voor een zeker rooster* L.

Als we deze resultaten combineren dan hebben we

Elliptische krommen over $\mathbb{C}$ vormen op natuurlijke wijze een torus

14.5. **Elliptische krommen over $\mathbb{F}_q$.** In deze sectie onderstellen we $K = \mathbb{F}_q$. Dan geldt natuurlijk $|E(\mathbb{F}_q)| < \infty$. Hier zijn twee tekeningen van $y^2 = x^3 - x$, eerst over het lichaam $\mathbb{F}_{101}$ en dan over het lichaam $\mathbb{F}_{1009}$.





Het is duidelijk dat deze tekeningen niet veel zeggen over de struktuur van $E(\mathbb{F}_q)$. Toch kunnen we onze intuitie van $\mathbb{R}$ en $\mathbb{C}$ gebruiken om na te denken over $E(\mathbb{F}_q)$. De taal om dit te doen is de algebraische meetkunde (die we dus niet gebruiken in deze nota's).

Een natuurlijk vraag is wat $|E(\mathbb{F}_q)|$ precies is. Dit is het onderwerp van veel huidig onderzoek. We hebben de volgende begrenzing (Hasse).

Stelling 14.5.1. *Er geldt*[32]

$$|q+1-|E(\mathbb{F}_q)|| \leq 2\sqrt{q} \tag{14.9}$$

Dit resultaat is erg moeilijk te bewijzen. Men kan het echter wel aannemelijk maken op de volgende manier. Voor de gemakkelijkheid onderstellen we dat $q=p$ priem is. We hebben

$$\begin{aligned} |E_{\text{aff}}(\mathbb{F}_p)| &= \sum_{x\in\mathbb{F}_p} \left[\left(\frac{x^3+ax+b}{p}\right)+1\right] \\ &= p + \sum_{x\in\mathbb{F}_p} \left(\frac{x^3+ax+b}{p}\right) \end{aligned}$$

We moeten dus het stuk

$$\sum_{x\in\mathbb{F}_p} \left(\frac{x^3+ax+b}{p}\right) \tag{14.10}$$

kunnen afschatten.

Om een heuristische afschatting te maken beschouwen we

$$x \mapsto \left(\frac{x^3+ax+b}{p}\right) \tag{14.11}$$

als een randomvariabele met waarden $\{-1,0,1\}$. De kans dat de waarde 0 wordt aangenomen is erg klein dus die verwaarlozen we. Onderstel dus dat de waarden $-1,1$ worden aangenomen met kans $1/2$. Het gemiddelde van deze randomvariabele is dan $(1/2)(1+(-1))=0$ en de variantie $(1/2)(1-0)^2+(1/2)(-1-0)^2=1$. Dus de standaardafwijking is $\sqrt{1}=1$. Het volgt dat de verwachtingswaarde van (14.10) gelijk is aan $p\cdot 0=0$ en de standaardafwijking is gelijk aan $\sqrt{p}\cdot 1=\sqrt{p}$. De wet van de grote getallen zegt dat (14.10) bij benadering normaal verdeeld is en de symmetrische overschrijdingskans van tweemaal de standaard afwijking is 4,56%. Dus om statistische redenen vermoeden we dat (14.9) "meestal" zal waar zijn. Het feit dat (14.9) *altijd* waar is zegt dus in feite dat (14.11) zich niet volledig als een randomvariabele gedraagt!

Er bestaat een meer preciese versie van (14.9). Definieer

$$a_r \stackrel{\text{def}}{=} q^r + 1 - |E(\mathbb{F}_{q^r})|$$

Stelling 14.5.2. *Er bestaat een $\alpha\in\mathbb{C}$ met $|\alpha|=\sqrt{q}$ zodat*

$$a_r = 2\,\mathrm{Re}(\alpha^r)$$

Opmerking 14.5.3. Stelling 14.5.2 impliceert Stelling 14.5.1. Inderdaad schrijf

$$\alpha = \alpha_0 + i\alpha_1$$

[32]Indien q een oneven macht is van een priemgetal dan is deze ongelijkheid noodzakelijkerwijze strikt. Indien q een even macht is dan kan er gelijkheid optreden.

Dan geldt $|\alpha_0| \leq |\alpha| = \sqrt{q}$ en dus

$$\begin{aligned} |q+1-|E(\mathbb{F}_q)|| &= |a_1| \\ &= 2|\alpha_0| \\ &\leq 2\sqrt{q} \end{aligned}$$

Gevolg 14.5.4. *a_r is bepaald door a_1.*

Bewijs. Schrijf zoals boven $\alpha = \alpha_0 + i\alpha_1$. Dan hebben we de volgende beperkingen

$$2\alpha_0 = a_1$$
$$\alpha_0^2 + \alpha_1^2 = q$$

We kennen dus α_0 en α_1 kennen we op teken na. Dat betekent dat we α op conjugatie na kennen. Aangezien

$$\mathrm{Re}(\alpha^r) = \mathrm{Re}(\bar{\alpha}^r)$$

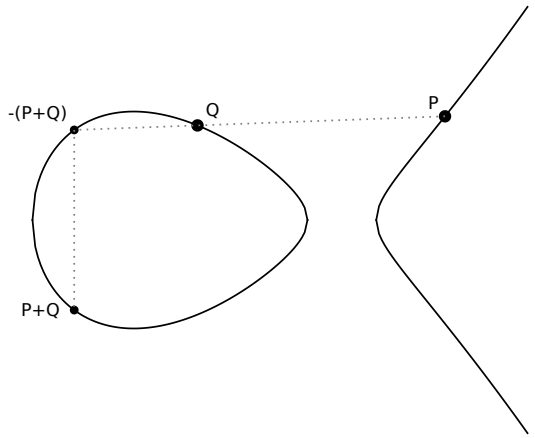
betekent dit dat we a_r kennen. □

15. De groepswet

15.1. **Definitie van de groepswet.** De punten op een elliptische kromme vormen een abelse groep. We zullen deze struktuur beschrijven indien char$K \neq 2, 3$ en de vergelijking de standaardvorm

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

heeft. De constructie van de groepswet is aangegeven in de volgende figuur



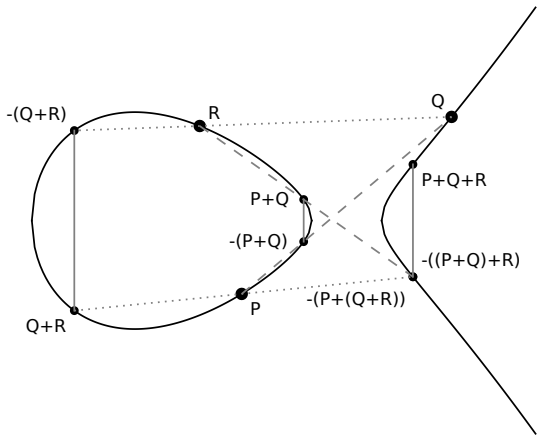
Dus gegeven punten P, Q nemen we de lijn L door P, Q. Deze snijdt de elliptische kromme E in een derde punt, zeg R. De som van $P + Q$ is het punt R, gespiegeld ten opzichte van de x-as.

Helaas zijn er redelijk wat speciale gevallen waarbij deze constructie op de juiste manier moet geinterpreteerd worden. We proberen die hier allemaal aan te geven.[33] We onderstellen $P = (x_1, y_1)$, $Q = (x_2, y_2)$.

(1) $P + P_\infty = P_\infty + P = P$.
(2) $P_\infty + P_\infty = P_\infty$.
(3) Indien $P \neq Q$ en $x_1 = x_2$ dan geldt $P + Q = P_\infty$.
(4) Indien $P = Q$ en $y_1 = y_2 = 0$ dan geldt $P + Q = P_\infty$. (1-4) kunnen we samenvatten door te zeggen dat P_∞ het neutraal element is en de inverse van een punt P is de gespiegelde van P ten opzichte van de x-as.
(5) Indien $P \neq Q$ en L raakt aan E in P dan is $P + Q = -P$.
(6) Indien $P \neq Q$ en L raakt aan E in Q dan is $P + Q = -Q$.
(7) Indien $P = Q$, $y_1 = y_2 \neq 0$ dan nemen we voor L de raaklijn aan E in P. Indien L de elliptische kromme in een tweede punt $R \neq P$ snijdt dan $P + P = -R$.
(8) Indien $P = Q$, $y_1 = y_2 \neq 0$ en indien L de elliptische kromme niet in een tweede punt snijdt dan $P + P = -P$.

Stelling 15.1.1. *Met bovenstaande definitie is $E(K)$ een abelse groep met P_∞ als neutraal element.*

We slaan het bewijs van deze stelling over. Het enige moeilijke punt is de associativiteit. Die wordt geillustreerd op de volgende figuur.



Opmerking 15.1.2. (Optioneel) Gegeven alle speciale gevallen in de definitie kan men zich afvragen of er geen betere manier is om de groepswet te beschrijven. Die is er.

[33]Eventuele speciale gevallen die ik gemist zou hebben worden met graagte aanvaard.

Eerst en vooral moeten we projectief werken. Zij $F \in K[X,Y,Z]$ een homogene vergelijking van graad drie zodaning dat

$$\partial F/\partial X, \partial F/\partial Y, \partial F/\partial Z$$

geen gemeenschappelijke wortels hebben.

Hieronder is $C(K)$ de verzameling oplossingen van $F(P) = 0$ met $P \in \mathbb{P}^2_K$ (dus de coordinaten van P zijn maar op een van nul verschillende scalair na bepaald). We fixeren ook een willeurig element $O \in C(K)$.[34]

Zij $P, Q \in C(K)$ en beschouw

$$f(\lambda, \mu) = F(\lambda P + \mu Q)$$

$f(\lambda,\mu)$ is een homogeen polynoom van graad 3 in (λ,μ). Aangezien $f(1,0) = f(0,1) = 0$ moet f de volgende vorm hebben

$$f(\lambda, \mu) = \lambda\mu(p\lambda + q\mu)$$

met $(p,q) \neq (0,0)$.[35] Defineer $R = qP - pQ$. Dan geldt $R \in C(K)$ en verder liggen P, Q, R op een rechte. We definieren $P \wedge Q = R$ en

$$P + Q \stackrel{\text{def}}{=} O \wedge (P \wedge Q)$$

In het geval van elliptische krommen (indien $\text{char}(K) \neq 2, 3$) nemen we

$$F = Y^2Z - (X^3 + aXZ^2 + bZ^3)$$

en $O = P_\infty$. Dus indien we projectief werken is het mogelijk om de groepswet te beschrijven zonder speciale gevallen aan te moeten geven.

We zullen nu de groepswet met formules beschrijven. Voor de gedetailleerde bewijzen verwijzen we naar [Kob94]. Onderstel $P = (x_1, y_1)$, $Q = (x_2, y_2)$ en zij $P + Q = (x_3, y_3)$ (deze zijn dus te bepalen).

(1) Onderstel $P \neq Q$, $x_1 \neq x_2$. De rechte L door P en Q heeft parameter voorstelling

$$(x_1, y_1) + \lambda(x_2 - x_1, y_2 - y_1)$$

Door op te lossen naar λ kunnen we het derde snijpunt van L met E bepalen. Uiteindelijk vinden we de volgende formule

$$\begin{aligned} x_3 &= \left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2 - x_1 - x_2 \\ y_3 &= -y_1 + \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x_1 - x_3) \end{aligned} \tag{15.1}$$

(2) Onderstel nu $P = Q$ en $y_1 \neq 0$. De rechte L is nu de raaklijn. Uit de vergelijking voor de raaklijn

$$(x - x_1)\frac{\partial F}{\partial x}(x_1, y_1) + (y - y_1)\frac{\partial F}{\partial y}(x_1, y_1) = 0$$

halen we dat de richtingsvector van L gelijk is aan

$$\left(\frac{\partial F}{\partial y}(x_1, y_1), -\frac{\partial F}{\partial x}(x_1, y_1)\right) = (2y_1, 3x_1^2 + a)$$

[34] In deze algemeenheid is het mogelijk dat $C(K) = \emptyset$. In dat geval zitten strop. Dit geldt bijvoorbeeld voor $F = 5X^3 + 11Y^3 + 13Z^3$, wat men kan zien door modulo 13 te werken.

[35] Dit vergt een argument dat we overslaan.

Dus de parameter voorstelling van L is

$$(x_1, y_1) + \lambda(2y_1, 3x_1^2 + a)$$

We krijgen dan de volgende formule

$$x_3 = \left(\frac{3x_1^2 + a}{2y_1}\right)^2 - 2x_1$$
$$y_3 = -y_1 + \left(\frac{3x_1^2 + a}{2y_1}\right)(x_1 - x_3)$$

15.2. **Speciale lichamen.** We geven nu een beschrijving van de groep $E(K)$ in enige speciale gevallen.

Onderstel $K = \mathbb{C}$. In dat geval geldt

Stelling 15.2.1. *De bijectie $E(\mathbb{C}) \cong \mathbb{C}/L$ die we boven geconstrueerd hebben (zie §14.3) is een groepsisomorfisme.*

Onderstel nu $K = \mathbb{R}$. Zij $S^1 = \mathbb{R}/\mathbb{Z}$ (de "cirkelgroep"). Het is niet moeilijk in te zien dat de rechtersamenhangingscomponent van $E(\mathbb{R})$ (zie de tekeningen in §14.2) een deelgroep van $E(\mathbb{R})$ is. Men kan aantonen dat deze deelgroep isomorf is met S^1. De andere samenhangingscomponent (indien hij bestaat) is een nevenklasse. Dus

$$E(\mathbb{R}) \cong S^1$$

of

$$E(\mathbb{R}) \cong \mathbb{Z}/2\mathbb{Z} \times S^1$$

Onderstel nu $K = \mathbb{Q}$. Dit geval is ontzettend belangrijk voor de getaltheorie. Bijvoorbeeld het bewijs van het Fermat vermoeden door Taylor-Wiles maakt essentieel gebruik van elliptische krommen over $\mathbb{Q}$.

We hebben de volgende fundamentele stelling.

Stelling 15.2.2. *(Mordell-Weil) $E(\mathbb{Q})$ is een eindig voortgebrachte abelse groep.*

Dus

$$E(\mathbb{Q}) = E(\mathbb{Q})_{\text{tors}} + \mathbb{Z}^r$$

waarbij $E(Q)_{\text{tors}}$ een eindige abelse groep is. Het getal r noemen we de *rang* van E. Notatie $r = \operatorname{rk} E$.

Stelling 15.2.3. *(Mazur) $E(\mathbb{Q})_{tors}$ is een van de volgende groepen*

$$\mathbb{Z}/n\mathbb{Z} \qquad n=1,2,3,\ldots,10,12$$

of

$$\mathbb{Z}/2n\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z} \qquad n=1,2,3,4$$

In het bijzonder geldt $|E(\mathbb{Q})_{tors}| \leq 16$.

Er is een vermoeden dat $\operatorname{rk} E$ onbegrensd is. De elliptische kromme met hoogst bekende rank werd gevonden in 2006 door Elkies. Hiervoor geldt $r \geq 28$. Zie [Duj].

15.3. **l-torsie.** Hieronder onderstellen we dat $K = \bar{K}$ algebraisch gesloten is. We definieren de *l-torsie groep* van $E(K)$ als

$$E(K)_l = \{P \in E(K) \mid lP = P_\infty\}$$

We vragen ons af: wat is $E(K)_l$?

Laat ons eerst het geval $K = \mathbb{C}$ beschouwen. Dan hebben we $E(\mathbb{C}) = \mathbb{C}/L$. Dus

$$\begin{aligned}(\mathbb{C}/L)_l &= \{\bar{z} \mid z \in \mathbb{C}, lz \in L\}\\ &= ((1/l)L)/L\\ &\overset{\times l}{\cong} L/lL\\ &\cong \mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/l\mathbb{Z}\end{aligned}$$

Dit resultaat blijkt ook voor andere lichamen te gelden.

Stelling 15.3.1. *Zij E een elliptische kromme over K met K algebraisch gesloten. Onderstel* $\operatorname{char} K \nmid l$. *Dan geldt*

$$E(K)_l \cong \mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/l\mathbb{Z}$$

Voor een lichaam van karakteristiek nul kan men deze stelling bewijzen met het zogenaamde *Lefschetz principe* dat ruwweg zegt dat elke algebraisch resultaat dat waar is voor $\mathbb{C}$ ook waar is voor alle algebraisch gesloten lichamen van karakteristiek nul. We zeggen dat een resultaat algebraisch is als het kan geformuleerd worden met behulp van de predikatenlogica en de lichaamsstructuur (0,1,+,×). Het zou duidelijk moeten zijn dat men Stelling 15.3.1 inderdaad op deze manier kan formuleren.[36]

Uit Stelling 15.3.1 halen we in het bijzonder

$$|E(\mathbb{C})_l| = l^2$$

15.4. **Het Frobenius automorfisme van $\bar{\mathbb{F}}_q$.** Zij K een lichaam. We weten dat K een algebraische sluiting $\bar{K}$ heeft. De lichaamsextensie $\bar{K}/K$ wordt op K-isomorfisme na gekarakterizeerd door twee eigenschappen.

(1) $\bar{K}/K$ is algebraisch. Met andere woorden voor elke $a \in \bar{K}$ is er een monisch polynoom $f(x) \in K[x]$ zodat $f(a) = 0$.
(2) Elk niet constant polynoom $f(x) \in \bar{K}[x]$ heeft een wortel in $\bar{K}$.

Indien $a \in \bar{K}$ dan is a algebraisch over K en dus is $K(a) : K$ eindig. Door dit te itereren vinden we dat indien $a_1, \ldots, a_n \in \bar{K}$ dan is $K(a_1, \ldots, a_n) : K$ eindig. Daarom is het werken in $\bar{K}$ bijna hetzelfde als het werken in eindige lichaamsuitbreidingen van K. Immers elk eindig aantal elementen van $\bar{K}$ zit in zo'n eindige lichaamsuitbreiding.

We gaan ons nu beperken tot $K = \mathbb{F}_q$. Beschouw de volgende afbeelding

$$F : \bar{\mathbb{F}}_q \to \bar{\mathbb{F}}_q : x \mapsto x^q$$

Het volgende bewijst men precies zoals voor eindige lichamen.

(1) $F(x + y) = F(x) + F(y)$, $F(xy) = F(x)F(y)$.
(2) $F \mid \mathbb{F}_q = \mathrm{id}$.

Lemma 15.4.1. (1) *F is een automorfisme van $\bar{\mathbb{F}}_q$.*

[36] Alhoewel dat natuurlijk uiterst omslachtig is.

(2) $(\bar{\mathbb{F}}_q)^F \stackrel{def}{=} \{x \in \bar{\mathbb{F}}_q \mid F(x) = x\} = \mathbb{F}_q$.

Bewijs. (1) F is een ringhomomorfisme tussen lichamen. Het is dus zeker injectief. We moeten dus enkel bewijzen dat F surjectief is. Onderstel dus $x \in \bar{\mathbb{F}}_q$. We moeten aantonen dat x in het beeld van F zit. Uit bovenstaande discussie volgt dat $x \in \mathbb{F}_{q^r} \subset \bar{\mathbb{F}}_q$ voor zekere r. Aangezien de Frobenius voor eindige lichamen een automorfisme is volgt dat er een $y \in \mathbb{F}_{q^r}$ is zodat $F(y) = x$. Dit is voldoende.

(2) Zij $x \in \bar{\mathbb{F}}_q$ zodat $F(x) = x$. x zit weer in een eindig lichaam $x \in \mathbb{F}_{q^r} \subset \bar{\mathbb{F}}_q$ en dus kunnen we weer de theorie van eindige lichamen gebruiken om te besluiten dat $x \in \mathbb{F}_q$. □

In overeenstemming met het gelijknamige vroeger ingevoerde begrip zullen we F het *Frobenius automorfisme* van $\bar{\mathbb{F}}_q$ noemen.

15.5. **De Frobenius afbeelding voor elliptische krommen.** Hieronder is E een elliptische kromme over $\mathbb{F}_q$. We onderstellen $2, 3 \nmid q$. Zoals gebruikelijk nemen we als vergelijking van E:

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

met $a, b \in \mathbb{F}_q$.

Zij $P = (u, v) \in E_{\text{aff}}(\bar{\mathbb{F}}_q)$. Dan definieren we $FP = (u^q, v^q)$. We definieren ook $FP_\infty = P_\infty$.[37]

Omdat $(u, v) \in E_{\text{aff}}(\bar{\mathbb{F}}_q)$ geldt

$$v^2 = u^3 + au + b$$

Als we dit tot de q-de macht verheffen en gebruiken dat $a^q = a$, $b^q = b$

$$(v^q)^2 = (u^q)^3 + au^q + b$$

dan vinden we $FP \in E_{\text{aff}}(\bar{\mathbb{F}}_q)$. Met andere woorden F definieert een afbeelding

$$F : E(\bar{\mathbb{F}}_q) \to E(\bar{\mathbb{F}}_q)$$

die we de Frobenius afbeelding noemen. Het is duidelijk dat F een bijectie is aangezien de Frobenius op $\bar{\mathbb{F}}_q$ een bijectie is. Verder hebben we ook

$$E(\mathbb{F}_q) = E(\bar{\mathbb{F}}_q)^F \tag{15.2}$$

Dus het begrijpen van $E(\mathbb{F}_q)$ komt neer op het begrijpen van $E(\bar{\mathbb{F}}_q)$ tezamen met het begrijpen van de actie van de Frobenius.

Stelling 15.5.1. *De Frobenius afbeelding is een groepshomomorfisme.*

Bewijs. We moeten aantonen $F(P + Q) = FP + FQ$. We kunnen dit geval per geval nakijken. We zullen een voorbeeld geven. Zij $P = (x_1, y_1)$, $Q = (x_2, y_2)$ en $P + Q = (x_3, y_3)$. Onderstel $P \neq Q$ en $x_1 \neq x_2$. Dan $F(P + Q) = (x_3^q, y_3^q)$. Door de formules (15.1) tot de q-de macht te verheffen vinden we

$$x_3^q = \left(\frac{y_2^q - y_1^q}{x_2^q - x_1^q}\right)^2 - x_1^q - x_2^q$$

$$y_3^q = -y_1^q + \left(\frac{y_2^q - y_1^q}{x_2^q - x_1^q}\right)(x_1^q - x_3^q)$$

De rechterzijde van deze formules is precies $FP + FQ$. □

[37] Deze definitie volgt weer logisch als we met homogene coordinaten werken.

Met behulp van deze stelling kunnen we de formule (15.2) herschrijven als

$$E(\mathbb{F}_q) = \ker(\mathrm{id} - F)$$

Immers $x \in \ker(\mathrm{id} - F) \iff (\mathrm{id} - F)(x) = 0 \iff \mathrm{id}(x) - F(x) = 0 \iff x = F(x)$. Analoog

$$E(\mathbb{F}_{q^r}) = \ker(\mathrm{id} - F^r)$$

De Frobenius heeft nog veel meer mooie eigenschappen dan de vermelde. Het volgende speciale geval van de beroemde "spoorformule" leidt tot een algoritme voor het berekenen van $|E(\mathbb{F}_q)|$.

Zij l zodanig dat $\mathrm{ggd}(q, l) = 1$ en beschouw het volgende commutatieve diagram

$$\begin{array}{ccccc} E(\bar{\mathbb{F}}_q) & \xleftarrow{\supset} & E(\bar{\mathbb{F}}_q)_l & \xrightarrow{\cong} & (\mathbb{Z}/l\mathbb{Z})^2 \\ {\scriptstyle F}\downarrow & & {\scriptstyle F_l}\downarrow & & \downarrow{\scriptstyle F_l} \\ E(\bar{\mathbb{F}}_q) & \xleftarrow[\supset]{} & E(\bar{\mathbb{F}}_q)_l & \xrightarrow[\cong]{} & (\mathbb{Z}/l\mathbb{Z})^2 \end{array}$$

Hierbij is F_l de restrictie van F tot de deelgroep $E(\bar{\mathbb{F}}_q)_l$.[38]

We beschouwen F_l als een 2×2-matrix met coefficienten in $\mathbb{Z}/l\mathbb{Z}$ via het isomorfisme $E(\bar{\mathbb{F}}_q)_l \cong (\mathbb{Z}/l\mathbb{Z})^2$ gegeven door Stelling 15.3.1.

Stelling 15.5.2. *Definieer $a = q + 1 - |E(\mathbb{F}_q)|$. Dan geldt*

$$a \equiv \mathrm{Tr}(F_l) \qquad \mathrm{mod}\, l \tag{15.3}$$

We hebben hier het spoor van een matrix A met $\mathrm{Tr}\, A$ aangeduid ("Tr" is "trace").

Opmerking 15.5.3. (ter info) Voor diegenen die ooit in contact komen met algebraische topologie: de spoorformule is een algebraische versie van de Lefschetz fixpuntstelling die de kardinaliteit van de fixpuntverzameling van een afbeelding $F : M \to M$ van een varieteit (manifold) M naar zichzelf uitdrukt in functie van de (co)homologie van M. Zie [Lef]. Het idee is dat $E(\bar{\mathbb{F}}_q)_l$ een algebraische versie is van de eerste cohomologie groep van E (opgetensored met $\mathbb{Z}/l\mathbb{Z}$). Dit laatste is een speciale eigenschap van elliptische krommen. Er bestaan echter diverse algebraische cohomologietheorieen die men voor meer algemene algebraische varieteiten kan gebruiken. De meest bekende is de zogenaamde l-adische cohomologie. Hiervoor kreeg Alexander Grothendieck in 1966 de Fields Medaille.[39]

Deze l-adische cohomologie werd gebruikt voor het bewijs van de zogenaamde Weil vermoedens door de Belgische wiskundige Pierre Deligne. Deze laatste kreeg hiervoor op zijn beurt in 1978 de Fields Medaille. De Weil vermoedens vormen een gigantische veralgemening van Stelling 14.5.1.

---

[38] Omdat F een groepshomomorfisme is wordt een l-torsie punt natuurlijk op een l-torsie punt afgebeeld: indien $lP = P_\infty$ dan $F(lP) = lF(P) = lP_\infty = P_\infty$.

[39] De Fields Medaille is in zekere zin de wiskundige versie van de Nobelprijs. Er is echter een belangrijk verschil. Terwijl de Nobelprijs meestal wordt toegekend voor werk dat vele jaren eerder gebeurde wordt de Fields Medaille enkel toegekend aan wiskundigen onder de 40 jaar.

Enkel jaren geleden werd de zogenaamde Abelprijs gecreeerd. Deze wordt toegekend in Noorwegen (zoals de Nobelprijs voor de vrede) en is meer gelijkaardig aan de Nobelprijs (ook in geldwaarde). Het prestige van de Fields Medaille blijft echter zeer groot.

De spoorformule leidt tot een polynomiaal algoritme voor het berekenen van $|E(\mathbb{F}_q)|$ (Rene Schoof). Kies

$$l = 2, 3, 5, 7, \ldots, L$$

Dan bepalen (15.3) plus de Chinese reststelling a_1 modulo $M = \prod_{p \leq L} p$. Aangezien volgens Stelling 14.5.1 geldt

$$|a_1| \leq 2\sqrt{q}$$

is het voldoende om $M > 4\sqrt{q}$ te kiezen opdat $|E(\mathbb{F}_q)|$ volledig bekend zou zijn.

16. Gebruik van elliptische krommen in cryptografie

16.1. **Inleiding.** Het idee achter elliptische krommen cryptografie (ECC) is het gebruik van het discrete logaritme probleem in $E(\mathbb{F}_q)$. Belangrijk hiervoor is te weten dat er geen subexponentieel algoritme bekend is om het DLP op te lossen in $E(\mathbb{F}_q)$. Zo'n subexponentieel algoritme bestaat wel voor $\mathbb{F}_q^*$: de zogenaamde index calculus.

We moeten echter wel een belangrijke proviso maken. Opdat het DLP in $E(\mathbb{F}_q)$ moeilijk zou zijn is het belangrijk dat $|E(\mathbb{F}_q)|$ minstens 1 grote priemfactor heeft. Anders kunnen we immers het Silver-Pohlig-Hellman algoritme gebruiken dat werkt in willekeurige groepen. Gelukkig hebben we gezien dat er efficiente algoritmen bestaan om $|E(\mathbb{F}_q)|$ uit te rekenen.

16.2. **Voorstellen van een boodschap door een punt op een elliptische kromme.** Bij het gebruik van ECC moeten allerlei kleine praktische probleempjes opgelost worden. Een daarvan is het representeren van een boodschap door een punt op een elliptische kromme.

Zij p een priemgetal. Het voorstellen van een "boodschap" $0 \leq m < M$ door een element van $\mathbb{F}_p^*$ is triviaal. Kies bijvoorbeeld $p > M + 1$ en stel m voor als $m + 1 \in \mathbb{F}_p^*$.[40]

In het geval van elliptische krommen is het idee dat we m voorstellen als x-coordinaat van een elliptische kromme E over $\mathbb{F}_p$. Helaas komt niet met elke x-coordinaat een punt overeen.

Hier is een mogelijk schema om dit probleem op te lossen. We kiezen een $\kappa \in \mathbb{N}$ (bijvoorbeeld $\kappa = 50$, zie hieronder wat de keuze van κ bepaalt), een priemgetal $p > M\kappa$ en een elliptische kromme E

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

over $\mathbb{F}_p$. We gaan m representeren als een punt op E met x-coordinaat van de vorm $x = m\kappa + j$ met $0 \leq j < \kappa$. De ontvanger kan dan m reconstrueren als $m = \lfloor x/\kappa \rfloor$.

We zoeken zo'n representatie door $j = 0, 1, 2, \ldots$ te proberen, Indien we een $x = m\kappa + j$ vinden zodat $x^3 + ax + b$ een kwadraat is dan representeren we onze boodschap als het punt

$$(x, \sqrt{x^3 + ax + b})$$

Wat is de kans dat we geen goede j vinden? Dit gebeurt indien

$$\left(\frac{x^3 + ax + b}{p}\right) = -1$$

[40]We brengen in herinnering dat dit een simplistische voorstelling van zaken is. In de praktijk zal men een vorm van padding gebruiken om m ruwweg hetzelfde hetzelfde aantal bits als p te geven.

voor $x = m\kappa, \ldots m\kappa + \kappa - 1$. Als we aannemen dat de waarde van $\left(\frac{x^3+ax+b}{p}\right)$ random is dan is de kans hierop $2^{-\kappa}$. Vandaar onze keuze $\kappa = 50$. Een kans 2^{-50} wordt geacht verwaarloosbaar klein te zijn. In elk geval is ze veel veel kleiner dan de kans dat onze computer een hardware probleem krijgt tijdens de berekening.

16.3. **Genereren van elliptische krommen over eindige lichamen.** In de vorige sectie moesten we reeds een elliptische kromme over een lichaam $\mathbb{F}_p$ kiezen. We hebben toen niet gespecifieerd hoe we dat doen. Voor het implementeren van cryptografische algoritmen moeten we dikwijls een paar kiezen (E, B) waarbij E een elliptische kromme is over $\mathbb{F}_p$ en $B \in E_{\text{aff}}(\mathbb{F}_p)$. Dit kunnen we als volgt doen.

(1) Kies $(u, v) \in \mathbb{F}_p^2$ random.
(2) Kies $a \in \mathbb{F}_p$ random.
(3) Bereken
$$b = v^2 - (u^3 + au)$$
(4) Indien $4a^3 + 27b^2 \neq 0$ in $\mathbb{F}_p$ dan nemen we
$$(E : y^2 = x^3 + ax + b, B = (u, v))$$
Anders gaan we terug naar (1).

Opmerking 16.3.1. Indien (E, B) geconstrueerd werden als data voor een DLP dan moeten we uiteraard nog verifieren dat $|E(\mathbb{F}_p)|$ een grote priemfactor heeft.

16.4. **ElGamal.** Als voorbeeld geven we aan hoe het ElGamal cryptosysteem werkt over elliptische krommen. In feite is dit overbodig want we hebben het ElGamal systeem ingevoerd over een willekeurige groep. Zie §10.3. We doen het toch maar om een duidelijker idee te geven hoe een ECC systeem werkt.

De parameters van het ElGamal ECC systeem zijn een priemgetal p, een elliptische kromme E over $\mathbb{F}_p$, een element $B \in E_{\text{aff}}(\mathbb{F}_p)$ en een random getal $0 < a < |E(\mathbb{F}_p)|$.

Een boodschap wordt gerepresenteerd door een element P van $E(\mathbb{F}_p)$.

De publieke sleutel is
$$(E, p, B, aB)$$
De private sleutel is
$$(E, p, B, a)$$
Om een boodschap $P \in E(\mathbb{F}_p)$ te encrypten kiezen we een randomgetal $0 < k < |E(\mathbb{F}_p)|$. De encryptie is
$$(kB, P + kaB)$$
De decryptie van
$$(C_1, C_2) \in E(\mathbb{F}_p) \times E(\mathbb{F}_p)$$
wordt gegeven door $C_2 - aC_1$.

16.5. **Diffie-Hellman.** Ook Diffie-Hellman hebben we voor een willekeurige groep ingevoerd. Zie §10.2. De ECC versie ziet er als volgt uit. De publiek parameters van het systeem zijn (E, p, B) waarbij p een priemgetal is, E een elliptisch kromme over $\mathbb{F}_p$ en $B \in E_{\text{aff}}(\mathbb{F}_p)$.

Als Alice en Bob een sleutel willen kiezen voor symmetrische cryptografie dan kiezen ze random getallen u en v. Ze sturen elkaar uB en vB. De sleutel kan dan bijvoorbeeld de x-coordinaat van uvB zijn.

17. Onbinden in factoren met behulp van elliptische krommen

17.1. **Inleiding.** Het ontbinden in factoren met elliptische krommen is in essentie een variant op de Pollard $p-1$-methode. Zie §12.4. In deze methode maken we de onderstelling dat het te ontbinden getal n een priemfactor p heeft zodat $p-1$ enkel deelbaar is door kleine priemfactoren.

We herinneren ons dat de Pollard $p-1$ methode essentieel neerkomt op het rekenen in $\mathbb{F}_p^*$, waarbij we gebruik maken van

$$|\mathbb{F}_p^*| = p-1$$

Er bestaat ook een $p+1$-methode (uitgevonden door Hugh C. Williams in 1982). Deze komt neer op het rekenen in $\mathbb{F}_{p^2}$ gebruikmakende van

$$|\mathbb{F}_{p^2}^*| = (p-1)(p+1)$$

Hiermee lijken we echter de limiet bereikt te hebben. We zouden kunnen proberen van bijvoorbeeld $\mathbb{F}_{p^3}$ te gebruiken. Maar

$$|\mathbb{F}_{p^3}^*| = (p-1)(p^2+p+1)$$

en de factor p^2+p+1 is niet meer $O(p)$ maar $O(p^2)$. Intuïtief verwachten we dat de kans dat er een priemfactor p is zodat p^2+p+1 enkel door kleine priemgetallen deelbaar is een stuk kleiner is dan dat dit zou gelden voor $p-1$ of $p+1$ (en dat kan men ook heuristisch verantwoorden).

Het idee is nu dat we $\mathbb{F}_p^*$ vervangen door $E(\mathbb{F}_p)$. Omdat

$$p+1-2\sqrt{p} \leq |E(\mathbb{F}_p)| \leq p+1+2\sqrt{p}$$

is de kans groot dat we na een aantal keer proberen uitkomen op een elliptische kromme zodat $|E(\mathbb{F}_p)|$ enkel deelbaar is door kleine priemfactoren. Deze methode zal hieronder uitgelegd worden.

In de $p-1$-methode hadden we het probleem dat we niet echt in $\mathbb{F}_p^*$ konden rekenen aangezien we p niet kennen. Dus we rekenen in feite in $(\mathbb{Z}/n\mathbb{Z})^*$ en gebruiken impliciet het groepshomorfisme $(\mathbb{Z}/n\mathbb{Z})^* \to \mathbb{F}_p^* : (a \bmod n) \mapsto (a \bmod p)$.

In de elliptische krommen context zouden we iets analoog willen doen. We zouden graag een elliptische kromme E over $\mathbb{Z}/n\mathbb{Z}$ definieren en dan gebruikmaken van een eventueel groepshomomorfism $E(\mathbb{Z}/n\mathbb{Z}) \to E(\mathbb{F}_p)$. Dit is in principe mogelijk maar het vraagt algebraische meetkunde om het correct te doen. Om dit te vermijden gebruiken we een veel meer adhoc aanpak.

17.2. **Elliptische krommen over $\mathbb{Z}/n\mathbb{Z}$.** Hieronder is n het te ontbinden getal. We onderstellen $2, 3 \nmid n$. We definieren een elliptische kromme E over $\mathbb{Z}/n\mathbb{Z}$ als een vergelijking

$$y^2 = x^3 + \bar{a}x + \bar{b}$$

met $\text{ggd}(4a^3+27b^2, n) = 1$. We definieren ook

$$E_{\text{aff}}(\mathbb{Z}/n\mathbb{Z}) = \{(u,v) \in (\mathbb{Z}/n\mathbb{Z})^2 \mid v^2 = u^3 + \bar{a}u + \bar{b}\}$$

Indien $p \mid n$ een priemdeler is and hebben we een elliptische kromme E_p over $\mathbb{F}_p$ met vergelijking

$$y^2 = x^3 + (a \bmod p)x + (b \bmod p)$$

en een afbeelding van verzamelingen

$$- \bmod p : E_{\text{aff}}(\mathbb{Z}/n\mathbb{Z}) \to E_{p,\text{aff}}(\mathbb{F}_p) : (u,v) \mapsto (u \bmod p, v \bmod p)$$

We gaan een partiele optelling "+" definieren op $E_{\text{aff}}(\mathbb{Z}/n\mathbb{Z})$. We zeggen dat $P = (\bar{x}_1, \bar{y}_1)$, $Q = (\bar{x}_2, \bar{y}_2) \in E_{\text{aff}}(\mathbb{Z}/n\mathbb{Z})$ *goed* zijn indien een der volgende voorwaarden geldt.

(1) $\bar{x}_1 - \bar{x}_2 \in (\mathbb{Z}/n\mathbb{Z})^*$ (en dus $P \neq Q$).
(2) $P = Q$ en $\bar{y}_1 \in (\mathbb{Z}/n\mathbb{Z})^*$.

Indien P, Q goed zijn dan gebruiken we gewoon de formules die we in §15.1 gezien hebben. In andere gevallen is de optelling niet gedefinieerd.

Lemma 17.2.1. *De afbeelding*

$$-\bmod p : E_{aff}(\mathbb{Z}/n\mathbb{Z}) \to E_{p,aff}(\mathbb{F}_p) \subset E_p(\mathbb{F}_p)$$

die we boven gedefinieerd hebben is compatiebel met de optelling.

Bewijs. Dit is gewoon omdat $\bar{x} \mapsto x \bmod p$ een ringhomomorfisme $\mathbb{Z}/n\mathbb{Z} \to \mathbb{F}_p$ is en ringhomomorfismen behouden eenheden en hun inversen. □

Definitie 17.2.2. Een paar van punten $P = (\bar{x}_1, \bar{y}_1)$, $Q = (\bar{x}_2, \bar{y}_2) \in E_{\text{aff}}(\mathbb{Z}/n\mathbb{Z})$ is *tegengesteld* indien $(\bar{x}_2, \bar{y}_2) = (\bar{x}_1, -\bar{y}_1)$.

Het is eenvoudig in te zien dat goed en tegengesteld incompatiebel zijn. Dus we hebben de volgende trichotomie.

(1) P, Q zijn goed.
(2) P, Q zijn tegengesteld.
(3) P, Q zijn noch goed noch tegengesteld. We beweren dat we in het laatste geval onmiddelijk een niet triviale factor van n vinden.

    Aangezien P, Q niet goed zijn hebben we $\bar{x}_1 - \bar{x}_2 \notin (\mathbb{Z}/n\mathbb{Z})^*$. Indien $\bar{x}_1 - \bar{x}_2 \neq \bar{0}$ dan is $\gcd(x_1 - x_2, n)$ een niet triviale factor van n. Onderstel dus $\bar{x}_1 - \bar{x}_2 = 0$. Uit de vergelijking van E volgt dan $\bar{y}_1^2 = \bar{y}_2^2$. Oftewel $(\bar{y}_1 - \bar{y}_2)(\bar{y}_1 + \bar{y}_2) = \bar{0}$. Omdat P, Q niet tegengesteld zijn en $\bar{x}_1 = \bar{x}_2$ volgt ook $\bar{y}_1 + \bar{y}_2 \neq \bar{0}$. Dus indien $\bar{y}_1 \neq \bar{y}_2$ dan is $\gcd(y_1 - y_2, n)$ een niet triviale factor van n. We mogen dus onderstellen $\bar{y}_1 = \bar{y}_2$. Opniew gebruiken dat P, Q niet tegengesteld zijn levert $\bar{y}_1 \neq 0$. Op dit punt geldt $P = Q$ en aangezien P, Q niet goed zijn moet gelden $\bar{y}_1 \notin (\mathbb{Z}/n\mathbb{Z})^*$. Dus $\gcd(y_1, n)$ is een niet triviale factor van n.

17.3. **Beschrijving van het algoritme.** We kiezen random een elliptische kromme E over $\mathbb{Z}/n\mathbb{Z}$

$$y^2 = x^3 + \bar{a}x + \bar{b}$$

en een punt $P = (u, v) \in E(\mathbb{Z}/n\mathbb{Z})$. We gebruiken hier uiteraard de techniek uit §16.3. Dus we kiezen $\bar{a}, u, v$ eerst en dan berekenen we $\bar{b}$. Tenslotte controleren we de conditie $\gcd(4a^3 + 27b^2, n) = 1$.

Kies een getal M dat een produkt is van kleine priemmachten (zie de beschrijving van de Pollard $p-1$ methode). We maken de volgende gok.[41]

Gok Er bestaat een priemgetal $p \mid n$ zodanig dat de exponent van $E_p(\mathbb{F}_p)$ een deler is van M. In het bijzonder geldt $M(P \bmod p) = P_\infty$ in $E_p(\mathbb{F}_p)$.

Hier en verder noteren we voor de compactheid het punt op oneindig in alle $E_q(\mathbb{F}_q)$ (voor q een priemdeler van n) met P_∞ (beter zou misschien zijn $P_\infty \bmod q$, alhoewel in onze setup P_∞ zelf niet in het domein van de "$\bmod q$" afbeelding zit).

[41] Meestal beschrijft men zo'n gok met het Duitse woord "Ansatz".

Gaan nu *proberen* om MP in $E_{\text{aff}}(\mathbb{Z}/n\mathbb{Z})$ te berekenen met behulp van "herhaald verdubbelen" (het additieve analogon van herhaald kwadrateren). Dus indien we bijvoorbeeld $6P$ zouden moeten berekenen dan berekenen we

$$6P = 2(2P + P)$$

We schrijven "proberen" omdat we niet zeker zijn dat de stappen $2P$, $2P + P$, $2(2P + P)$ wel allemaal gedefinieerd zijn.

Om MP te berekenen proberen we een rij punten in $E_{\text{aff}}(\mathbb{Z}/n\mathbb{Z})$ te construeren

$$P = P_1, P_2, P_3, \ldots, P_r \tag{17.1}$$

waarbij $P_i = P_{i'} + P_{i''}$ voor zekere $i', i'' < i$. Deze rij is zodanig dat indien alle optellingen gedefinieerd zijn en de optelling associatief is dan $P_r = MP$. Die associativiteit is een beetje een subtiel punt. Alhoewel die geldt in $E_{\text{aff}}(\mathbb{Z}/n\mathbb{Z})$ zullen we er niet op te steunen.

We beweren nu dat we een rij als in (17.1) *niet kunnen construeren*. Inderdaad als zo'n rij zou bestaan dan zouden we hebben $P_r \bmod p = M(P \bmod p) = P_\infty$ (de afbeelding $\bmod p$ is immers compatiebel met de optelling en we nemen uiteraard wel de associativiteit van de optelling in $E_p(\mathbb{F}_p)$ aan). Langs de andere kant $P_r \in E_{\text{aff}}(\mathbb{Z}/n\mathbb{Z})$ en dus $P_r \bmod p \in E_{p,\text{aff}}(\mathbb{F}_p)$. Dit is een contradictie.

Dus er moet iets mis gaan in de berekening van MP en het enige dat kan mis gaan is dat we op een gegeven moment een paar $P_{i'}, P_{i''}$hebben dat niet goed is. Volgens onze trichotomie hebben we dan ofwel dat $P_{i'}, P_{i''}$ tegengesteld zijn, ofwel dat we onmiddelijk een factor van n kunnen vinden.

Het slechte geval is dus wanneer $P_{i'}, P_{i''}$ tegengesteld zijn. Indien dit gebeurt dan geldt voor alle priemdelers $q|n$: $P_{i'} \bmod q = L'(P \bmod q)$, $P_{i''} \bmod q = L''(P \bmod q)$ voor zekere $L', L'' \in \mathbb{N}$ (onafhankelijk van q!) en $(P_{i'} \bmod q) + (P_{i''} \bmod q) = P_\infty$. Dus $(L' + L'')(P \bmod q) = P_\infty$ en dus $M'(P \bmod q) = P_\infty$ voor $M' = L' + L'' \leq M$.

De kans dat $M'(P \bmod q) = P_\infty$ in alle groepen $E_q(\mathbb{F}_q)$ voor een M' die niet van q afhangt is erg klein (deze groepen hebben bijvoorbeeld hoogstwaarschijnlijk allemaal verschillende orde). We kunnen nu besluiten dat de kans dat $P_{i'}, P_{i''}$ tegengesteld zouden zijn erg klein is, en vanwege onze trichotomie vinden we dus inderdaad een factor van n.

Indien er niets mis loopt dan was onze gok hoogstwaarschijnlijk fout. In ieder geval nemen gewoon een nieuwe elliptische kromme E over $\mathbb{Z}/n\mathbb{Z}$ en we beginnen opnieuw.

18. Het congruente getallenprobleem

18.1. **Inleiding.** Elliptische krommen komen te voorschijn bij allerlei wiskundinge problemen. Een voorbeeld hiervan is het *congruente getallen probleem* dat reeds bij de oude Grieken opduikt. De volgende nota's zijn gebaseerd op [Kob94]

We noemen een strikt positief rationaal getal *congruent* indien het de oppervlakte is van een rechthoekige driehoek is met rationale zijden. De vraag wat precies de congruente getallen zijn is heden ten dage nog altijd open. Er is wel een precies criterium bekend dat er echter vanuit gaat dat het Birch & Swinnerton-Dyer vermoeden geldt.[42]

[42]Het Birch-Swinnerton-Dyer vermoeden is een van de Millenium Problemen. Dit is een lijst van 7 beroemde problemen waarop reeds vele wiskundigen hun tanden stuk gebeten hebben. Voor het oplossen van (een van:-) die problemen krijg je dan ook een miljoen dollar. Zie

Zij n een strikt positief rationaal getal. Opdat n congruent zou zijn moet het volgende stelsel een oplossing in rationale getallen hebben.

$$Z^2 = X^2 + Y^2$$
$$n = \frac{1}{2}XY$$

Indien we X, Y, Z met een constant k vermenigvuldigen dan vermenigvuldigen we n met k^2. Het is dus voldoende het geval te beschouwen waarbij n een kwadraatvrij strikt positief natuurlijk getal is. *Dit zullen we hieronder steeds doen ook als we het niet expliciet vermelden.*

We zullen hieronder het congruente getallenprobleem vertalen naar een probleem over elliptische krommen. Dit laatste probleem kunnen we in veel gevallen oplossen maar een algemene oplossing is niet gekend, *behalve* indien het BSD vermoeden waar zou zijn.[43]

Als voorsmaakje geven we hier enige voorbeelden van al dan niet congruente getallen.

(1) $1, 2, 3$ zijn niet congruent.
(2) 5 is congruent. Een mogelijke driehoek is $(3/2, 20/3, 41/6)$.
(3) 6 is congruent. Een mogelijke driehoek is $(3, 4, 5)$.
(4) 7 is congruent. Een mogelijke driehoek is $(24/5, 35/12, 337/60)$.

18.2. **Vertaling naar elliptische krommen.** We onderstellen dat n een kwadraatvrij strikt positief natuurlijk getal is. We moeten nagaan of het volgende stelsel een oplossing in rationale getallen heeft

$$Z^2 = X^2 + Y^2 \tag{18.1}$$
$$\frac{1}{2}XY = n \tag{18.2}$$

Onderstel dat we zo'n oplossing hebben. Door (18.1)$\pm$4(18.2) te bekijken vinden we

$$(X + Y)^2 = Z^2 + 4n \tag{18.3}$$
$$(X - Y)^2 = Z^2 - 4n \tag{18.4}$$

Vermenigvuldigen en delen door 4 levert dit

$$\left(\frac{X^2 - Y^2}{4}\right)^2 = \left(\frac{Z}{2}\right)^4 - n^2$$

Definieer

$$u = \frac{Z}{2} \tag{18.5}$$
$$v = \frac{X^2 - Y^2}{4} \tag{18.6}$$

Dan verkrijgen we

$$v^2 = u^4 - n^2$$

`http://www.claymath.org/millennium/`. Een Millenium Probleem dat intussen opgelost is, is het Poincare vermoeden. Dit vermoeden werd bewezen door de Russische wiskundige Perelman, die hiervoor de Fields Medaille werd toegekend. Hij weigerde deze echter om niet volledig duidelijke redenen.

[43]Wat zeer waarschijnlijk is, gezien de grote hoeveelheid evidentie.

We vermenigvuldigen met u^2 en we stellen

$$x = u^2 \tag{18.7}$$

$$y = uv \tag{18.8}$$

Dit levert ons

$$y^2 = x^3 - n^2 x$$

Dit is de vergelijking van een elliptische kromme die we met E_n zullen noteren. Een oplossing van (18.1-18.2) heeft ons dus een element van $E_n(\mathbb{Q})$ opgeleverd.

Wanneer kunnen we nu terug? Hiervoor kunnen we alvast een paar noodzakelijke voorwaarden formuleren.

(1) x is een kwadraat.

Indien we (18.3-18.4) combineren met (18.5)(18.7) vinden we

$$\left(\frac{X+Y}{2}\right)^2 = x + n$$

$$\left(\frac{X-Y}{2}\right)^2 = x - n$$

Dus we vinden als extra nodige voorwaarden.

(2) $x - n$ is een kwadraat.
(3) $x + n$ is een kwadraat.

Indien ze condities gelden dan kunnen we X, Y, Z uit x, y berekenen via

$$X = \sqrt{x+n} + \sqrt{x-n}$$

$$Y = \sqrt{x+n} - \sqrt{x-n}$$

$$Z = 2\sqrt{x}$$

Het blijkt dat deze voorwaarden in direct verband staan met de groepswet.

Stelling 18.2.1. *Zij E/K een elliptische kromme met* $\operatorname{char} K \neq 2, 3$ *en vergelijking*

$$y^2 = (x - e_1)(x - e_2)(x - e_3)$$

Onderstel $(u, v) \in E_{\mathit{aff}}(K)$.

$$(u, v) \in 2E(K) \iff u - e_1,\ u - e_2,\ u - e_3 \text{ zijn kwadraten in } K$$

Het bewijs hiervan is nogal lastig, dus dat slaan we over.

Stelling 18.2.2. *Een punt $(x, y) \in E_{n,\mathit{aff}}(\mathbb{Q})$ correspondeert met een oplossing van het stelsel (18.1-18.2) als en slechts als $(x, y) \in 2E(\mathbb{Q})$.*

Bewijs. De vergelijking van E_n is

$$y^2 = (x - 0)(x - n)(x + n)$$

Dus een punt $(x, y) \in E_{\mathrm{aff}}(\mathbb{Q})$ is in $2E(\mathbb{Q})$ als en slechts als $x - 0$, $x - n$, $x + n$ kwadraten zijn. Dit zijn precies de condities die we boven gevonden hadden. □

Vanwege de Mordel-Weil stelling (zie Stelling 15.2.2) weten we nu

$$E_n(\mathbb{Q}) = E_n(\mathbb{Q})_{\mathrm{tors}} \oplus \mathbb{Z}^r$$

met $E_n(\mathbb{Q})_{\mathrm{tors}}$ eindig. Men kan het volgende aantonen.

Stelling 18.2.3. *We hebben* $E(\mathbb{Q})_{tors} = E(\mathbb{Q})_2$. *Met andere woorden*

$$E(\mathbb{Q})_{tors} = \{P_\infty, (0,0), (-n,0), (n,0)\}$$

Gebruikmakende van deze stelling vinden we

$$2E_n(\mathbb{Q}) = 2\mathbb{Z}^r$$

en dus

Gevolg 18.2.4. *n is een congruent getal als en slechts als* $\operatorname{rk} E_n(\mathbb{Q}) \geq 1$.

Het probleem met dit gevolg is dat er geen volledig algoritme bestaat om de rang van een elliptische kromme uit te rekenen. Er bestaan echter wel partiele algoritmen die in vele gevallen het antwoord geven. Deze zijn geintegreerd in het computer algebra pakket sage (www.sagemath.org).

In het volgende stukje code bepalen we de rang van de kromme E_5 alsook een generator van de groep $E_5(\mathbb{Q})$. Tweemaal die generator geeft ons een element van $2E_5(\mathbb{Q})$.

```
sage: n=5
sage: E=EllipticCurve([-n^2,0]); E
Elliptic Curve defined by y^2  = x^3 - 25*x over Rational Field
sage: E.rank()
1
sage: g=E.gen(0); g
(-4 : 6 : 1)
sage: g2=2*g; g2
(1681/144 : -62279/1728 : 1)
```

We gebruiken nu dit gevonden punt in $2E_5(\mathbb{Q})$ om een corresponderende driehoek te construeren.

```
sage: x,y=g2[0:2]
sage: X,Y,Z=sqrt(x+n)+sqrt(x-n),sqrt(x+n)-sqrt(x-n),2*sqrt(x)
sage: X,Y,Z
(20/3, 3/2, 41/6)
```

We kunnen de groepswet gebruiken om een oneindige reeks rechthoekige driehoeken te construeren met rationale zijden. Hier construeren we de driehoek die met vier maal de generator overeenkomt.

```
sage: g4=2*g2; g4
(11183412793921/2234116132416 : 1791076534232245919/3339324446657665536 : 1)
sage: x,y=g4[0:2]
sage: X,Y,Z=sqrt(x+n)+sqrt(x-n),sqrt(x+n)-sqrt(x-n),2*sqrt(x)
sage: X,Y,Z
(4920/1519, 1519/492, 3344161/747348)
```

Koblitz bespreekt in [Kob94] het geval $n = 157$. Met behulp van sage vinden we in een fractie van een seconde de volgende rationale rechthoekige driehoek met

oppervlakte 157.

$$X = \frac{411340519227716149383203}{21666555693714761309610}$$
$$Y = \frac{6803298487826435051217540}{411340519227716149383203}$$
$$Z = \frac{22440351770433696992455751309067486316094847204 1}{8912332268928859588025535178967163570016480830}$$

18.3. **Het Birch-Swinnerton-Dyer vermoeden.** De volledige formulering van het Birch-Swinnerton-Dyer vermoeden valt ver buiten het bestek van deze cursus. We zullen een ruw idee geven waarover het gaat. Onderstel dat we een elliptische kromme E

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

over $\mathbb{Q}$ hebben met $a, b \in \mathbb{Z}$. Dan is er voor elk priem p zodat $\operatorname{ggd}(p, 4a^3+27b^2) = 1$ een corresponderende elliptische kromme E_p over $\mathbb{F}_p$ met vergelijking

$$y^2 = x^3 + (a \bmod p)x + (b \bmod p)$$

Het Birch-Swinnerton Dyer vermoeden zegt nu dat er een link is tussen de getallen $|E(\mathbb{F}_p)|$ en $r = \operatorname{rk} E$. Een variant van het vermoeden (niet voldoende voor onze toepassingen) is het volgende (zie [Kna92, Conj 1.9]).

Vermoeden 18.3.1. *De limiet*

$$\lim_{R\to\infty} \frac{1}{(\log R)^{\operatorname{rk} E}} \prod_{p\leq R} \frac{|E(\mathbb{F}_p)|}{p}$$

bestaat en is een strikt positief reel getal.

Het idee achter dit vermoeden is is the volgende. Indien $a/b \in \mathbb{Q}$ zodanig is dat $\operatorname{ggd}(p, b) = 1$ dan kunnen we de "reductie modulo p" van a/b definieren als $ab^{-1} \bmod p$. Het idee is dat de elementen van $E(\mathbb{Q})$ bijdragen aan $E_p(\mathbb{F}_p)$ door reductie modulo p. Jammer genoeg heeft nooit iemand dit naieve idee kunnen omzetten in een bewijs.

De toepasbaarheid van het BSD vermoeden op het congruente getallen probleem volgt uit het feit dat dat voor de elliptische krommen E_n die we boven hebben ingevoerd de getallen $|E(\mathbb{F}_p)|$ bekend zijn. Zie [Kob93, §II.2]. Onderstel $p \nmid 2n$ dan geldt

$$|E(\mathbb{F}_p)| = \begin{cases} p+1 & \text{als } p \equiv 3 \bmod 4 \\ p+1-2a & \text{als } p \equiv 1 \bmod 4, \text{ waarbij } \exists b\colon\ a^2+b^2 = p,\ 2+2i \mid a+bi-\left(\frac{n}{p}\right) \end{cases}$$

Indien p een priemgetal is $\equiv 1 \bmod 4$ dan bestaan er gehele getallen zodat $p = a^2+b^2$. In feite kan dit op 8 manieren want we kunnen a, b vervangen door $\pm a, \pm b$ en we kunnen ze ook omwisselen. De conditie $2+2i \mid a+bi-\left(\frac{n}{p}\right)$ legt vast welk van de 8 oplossingen we moeten nemen.

18.4. **Tunnell's criterium.** Definieer

$$f(n) = \#\{x,y,z \in \mathbb{Z} \mid 2x^2 + y^2 + 8z^2 = n\}$$
$$g(n) = \#\{x,y,z \in \mathbb{Z} \mid 2x^2 + y^2 + 32z^2 = n\}$$
$$h(n) = \#\{x,y,z \in \mathbb{Z} \mid 4x^2 + y^2 + 8z^2 = \frac{n}{2}\}$$
$$k(n) = \#\{x,y,z \in \mathbb{Z} \mid 4x^2 + y^2 + 32z^2 = \frac{n}{2}\}$$

Als we het Birch & Swinnerton-Dyer vermoeden aannemen dan is n congruent als en slechts als $h(n) = 2k(n)$, indien n even is, ofwel $f(n) = 2g(n)$ indien n oneven is.

Men kan uit dit criterium het volgende gevolg halen.

Gevolg 18.4.1. *(modulo B&S-D). Indien $n \cong 5,6,7 \bmod 8$ dan is n congruent.*

19. De j-invariant

We hebben het in feite nog niet over de classificatie van elliptische krommen gehad. Het blijkt dat over een algebraisch afgesloten lichaam elliptische krommen afhangen van een enkele parameter. De zogenaamde j-invariant.

Hieronder is K een algebraisch gesloten lichaam met karakteristiek verschillend van twee en drie.

Herinner je dat we een elliptische kromme adhoc gedefinieerd hebben als een bepaald soort vergelijking modulo coordinatentransformaties. Door onze hypothese op de karakteristiek mogen we onderstellen dat de vergelijking van de vorm $y^2 = x^3 + ax + b$ is met $4a^3 + 27b^2 \neq 0$. Met andere woorden

$$\{\text{ell. krommen}/K\} \cong \{y^2 = x^3 + ax + b; 4a^3 + 27b^2 \neq 0\}/\{\text{coord. transformaties}\}$$

De coordinatentransformaties waarvan hier spraken is zijn deze van §14.1 met de bijkomende voorwaarde dat de meer specifieke vorm $y^2 = x^3 + ax + b$ voor de vergelijking bewaard blijft. Men verifieert dat zulke coordinatentransformaties noodzakelijk van de vorm moeten zijn

$$\begin{aligned} x &= \alpha x' \\ y &= \gamma y' \end{aligned} \tag{19.1}$$

met $\alpha \neq 0$, $\gamma \neq 0$.

De nieuwe vergelijking wordt dan

$$\gamma^2 (y')^2 = \alpha^3 (x')^3 + a\alpha(x') + b$$

Om de kopcoefficienten weg te kunnen delen moeten we dus de bijkomende voorwaarde $\gamma^2 = \alpha^3$ opleggen in (19.1). Indien α gegeven is dan kunnen we altijd een γ vinden die aan deze bijkomende voorwaarde voldoet. We moeten ons dus enkel op α concentreren.

Aannemende dat $\gamma^2 = \alpha^3$ vinden we als nieuwe vergelijking

$$(y')^2 = (x')^3 + \frac{a}{\alpha^2} x' + \frac{b}{\alpha^3}$$

We kunnen dus besluiten

$$\{\text{ell. krommen/K}\} \cong \underbrace{\{(a,b) \in K^2 \mid 4a^3 + 27b^2 \neq 0\}}_{\stackrel{\text{def}}{=} U} / \sim$$

waarbij

$$(a,b) \sim (a',b') \iff \exists \alpha \in K - \{0\} : a' = \frac{a}{\alpha^2}, b' = \frac{b}{\alpha^3}$$

Definieer de volgende functie

$$j : U \to K : (a,b) \mapsto 1728 \cdot \frac{4a^3}{4a^3 + 27b^2}$$

De rare voorfactor 1728 is puur conventie.[44]

Stelling 19.1. *Hieronder zijn $(a,b), (a',b')$ elementen van U.*

(1) $(a,b) \sim (a',b') \Rightarrow j(a,b) = j(a',b')$.
(2) $j(a,b) = j(a',b') \Rightarrow (a,b) \sim (a',b')$.
(3) $\forall j \in K : \exists (a,b) \in U : j(a,b) = j$ *(maw: de afbeelding j is surjectief).*

Bewijs. (1)

$$\begin{aligned} j(a',b') &= 1728 \cdot \frac{4\left(\frac{a}{\alpha^2}\right)^3}{4\left(\frac{a}{\alpha^2}\right)^3 + 27\left(\frac{b}{\alpha^3}\right)^2} \\ &= 1728 \cdot \frac{4a^3}{4a^3 + 27b^2} \\ &= j(a,b) \end{aligned}$$

(2) Onderstel $j(a,b) = j(a',b')$. Er geldt

$$1728 \cdot \frac{4a^3}{4a^3 + 27b^2} = 1728 \cdot \frac{4a'^3}{4a'^3 + 27b'^2}$$

Hieruit halen we snel

$$a^3 b'^2 = a'^3 b^2 \tag{19.2}$$

Onderstel bijvoorbeeld $a = 0$. Dan vinden we achtereenvolgens $b \neq 0$ (want $4a^3 + 27b^2 \neq 0$), $a' = 0$, $b' \neq 0$. Het volstaat nu van

$$\alpha = \sqrt[3]{\frac{b}{b'}}$$

te stellen.

De gevallen waarbij een van de volgende condities $b = 0$, $a' = 0$, $b' = 0$ gelden worden op analoge manier behandeld.

Onderstel tenslotte $a \neq 0$, $b \neq 0$, $a' \neq 0$, $b' \neq 0$. We kunnen dan (19.2) herschrijven als

$$\left(\frac{a}{a'}\right)^3 = \left(\frac{b}{b'}\right)^2$$

Stel

$$\alpha = \frac{a'}{a} \cdot \frac{b}{b'}$$

We vinden

$$\alpha^2 = \left(\frac{a'}{a}\right)^2 \left(\frac{b}{b'}\right)^2 = \left(\frac{a'}{a}\right)^2 \left(\frac{a}{a'}\right)^3 = \frac{a}{a'}$$

[44] De eigenlijke reden van de voorfactor is dat wanneer je de j-invariant uitrekent van de elliptische kromme $\mathbb{C}/L$ (zie §14.4) je op deze manier een elegantere uitdrukking bekomt.

en op een gelijkaardige manier

$$\alpha^3 = \frac{b}{b'}$$

Dus $(a,b) \sim (a',b')$.

(3) We moeten de volgende vergelijking oplossen in U:

$$1728 \cdot \frac{4a^3}{4a^3 + 27b^2} = j$$

Oftewel

$$(1728 - j)4a^3 = j(27b^2)$$

We schrijven dit als

$$ua^3 = vb^2$$

met $u = (1728 - j)4$, $v = 27j$. Merk op dat noodzakelijk $(u,v) \neq (0,0)$.

Indien bijvoorbeeld geldt $v \neq 0$ (en dus $j \neq 0$) dan stellen we

$$a = 1, \qquad b = \sqrt{\frac{u}{v}}$$

We moeten dan nog nagaan dat (a,b) in U zit. Met andere woorden $4a^3 + 27b^2 \neq 0$. We hebben

$$4a^3 + 27b^2 = 4 + 27\frac{u}{v} = \frac{4 \cdot 1728}{j} \neq 0$$

In het geval $v = 0$ kunnen we $(a,b) = (0,1)$ nemen. □

Zij $E_{a,b}$ de elliptische kromme met vergelijking $y^2 = x^3 + ax + b$. Stel

$$j(E_{a,b}) = j(a,b)$$

Uit de voorgaande stelling halen we dat j goed gedefinieerd is en een 1-1 verband geeft tussen de elliptische krommen/K en de elementen van K.

Opmerking 19.2. Voor de theorie van de j-invariant is het erg belangrijk dat K algebraisch gesloten is. Inderdaad: in het congruente getallen probleem zijn we de elliptische krommen tegengekomen

$$E_n : y^2 = x^3 - n^2 x$$

met n een kwadraatvrij strikt positief natuurlijk getal. Alle E_n's hebben dezelfde j-invariant. Met name 1728. Over $\mathbb{Q}$ zijn ze echter alle verschillend.

Inderdaad indien $E_n = E_m$ dan moeten er $(\alpha,\gamma) \in \mathbb{Q}_0^2$ bestaan zodat

$$-n^2 = (-m^2)/\alpha^2, \qquad \gamma^2 = \alpha^3$$

Uit de tweede vergelijking volgt dat α een kwadraat is en uit de eerste volgt $\alpha = \frac{m}{n}$. Het quotient van twee kwadraatvrije getallen is een kwadraat als en slechts als ze aan elkaar gelijk zijn.

Referenties

[AES01] *Announcing the ADVANCED ENCRYPTION STANDARD (AES)*, Federal Information Processing Standards Publication 197, November 2001, http://www.csrc.nist.gov/publications/fips/fips197/fips-197.pdf.

[AGP94] W. R. Alford, A. Granville, and C. Pomerance, *There are infinitely many Carmichael numbers*, Ann. of Math. (2) **139** (1994), no. 3, 703–722.

[AKS04] Manindra Agrawal, Neeraj Kayal, and Nitin Saxena, *PRIMES is in P*, Ann. of Math. (2) **160** (2004), no. 2, 781–793.

[Con76] J. H. Conway, *On numbers and games*, vol. 6, London Mathematical Society, 1976.

[DSA00] *DIGITAL SIGNATURE STANDARD (DSS)*, FEDERAL INFORMATION PROCESSING STANDARDS PUBLICATION 186-2, January 2000, http://csrc.nist.gov/publications/fips/fips186-2/fips186-2-change1.pdf.

[Duj] Andrej Dujella, *History of elliptic curves rank records*, http://web.math.hr/~duje/tors/rankhist.html.

[Hen06] C. Henry, *A short proof of the simple continued fraction expansion of e*, Amer. Math. Monthly **113** (2006), no. 1, 57–62.

[HN04] J. Hastad and M. Naslund, *The security of all RSA and discrete log bits*, J. ACM **51** (2004), no. 2, 187–230.

[Kna92] Anthony W. Knapp, *Elliptic curves*, Mathematical Notes, vol. 40, Princeton University Press, Princeton, NJ, 1992.

[Knu80] D. E. Knuth, *The art of computer programming*, second ed., vol. 2, Addison-Wesley, 1980.

[Kob93] Neal Koblitz, *Introduction to elliptic curves and modular forms*, second ed., Graduate Texts in Mathematics, vol. 97, Springer-Verlag, New York, 1993.

[Kob94] N. Koblitz, *A course in number theory and cryptography*, second ed., Graduate Texts in Mathematics, vol. 114, Springer Verlag, 1994.

[Koc96] P. Kocher, *Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems*, CRYPTO 1996, Springer Verlag, 1996, pp. 104–113.

[Lef] *Lefschetz fixed-point theorem*, http://en.wikipedia.org/wiki/Lefschetz_fixed-point_theorem.

[LO91] B. LaMacchia and A. Odlyzko, *Computation of discrete logarithms in prime fields*, CRYPTO '90, Lecture Notes in Comput. Sci., vol. 537, Springer-Verlag, 1991, pp. 109–133.

[MR02] S. Murphy and M. Robshaw, *Essential algebraic structure within the AES*, Advances in cryptology—CRYPTO 2002, Lecture Notes in Comput. Sci., vol. 2442, Springer, Berlin, 2002, pp. 1–16.

[MvOV97] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of applied cryptography*, CRC Press, 1997.

[NBV] *De nieuwe bijbelvertaling*, http://www.biblija.net/biblija.cgi?m=2+Kronieken+4%3A2 &id18=1&pos=0&set=10&lang=nl.

[Sch04] R. Schoof, *Four primality testing algorithms.*, http://www.mat.uniroma2.it/~schoof/millerrabinpom.pdf, 2004.

[Sil92] Joseph H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 106, Springer-Verlag, New York, 1992, Corrected reprint of the 1986 original.

[Sil98] J. Silverman, *Elliptic curves, discrete logarithms and the index calculus*, Talk at the University of Waterloo, September 1998, http://www.cacr.math.uwaterloo.ca/conferences/1998/ecc98/silverman.ps.

[Slo] N. J. A. Sloane, *On-line encyclopedia of integer sequences*, http://www.research.att.com/~njas/sequences/index.html.

[Wed01] S. Wedeniwski, *Primality tests on commutator curves*, Ph.D. thesis, Eberhard-Karls-Universität Tübingen, 2001.