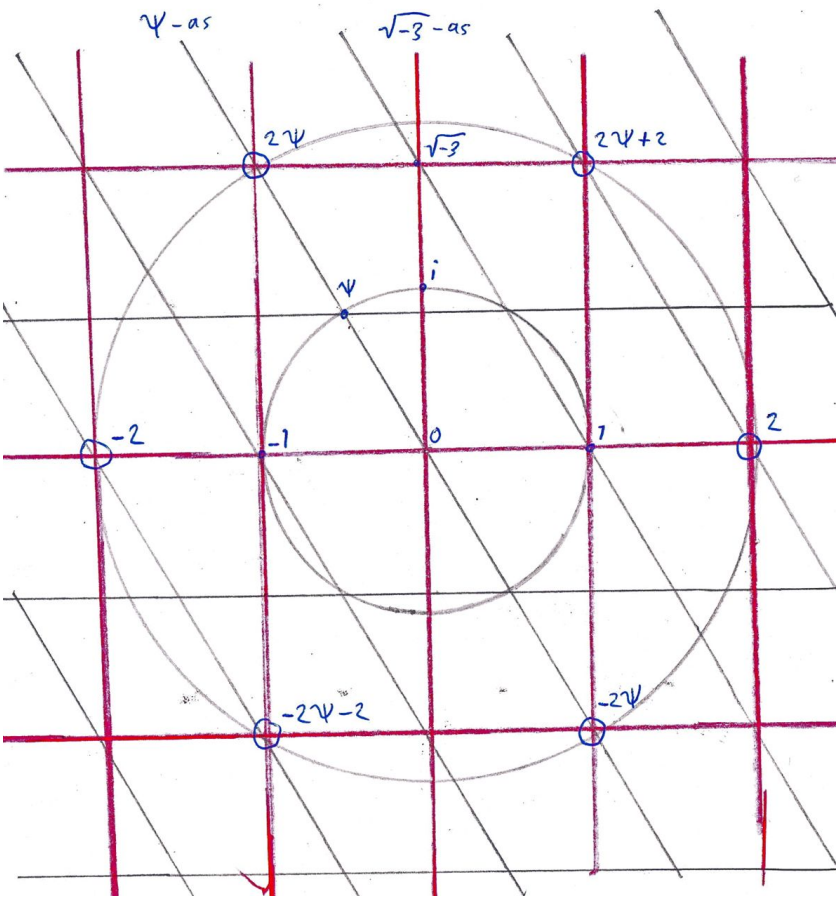
De Laatste Stelling van Fermat

Een introductie tot groepen, ringen, lichamen en idealen

Lars van den Berg, Merlijn Koek, Serop Lazarian
Onder begeleiding van prof. dr. J.P. Hogendijk

20 november 2011



Voorwoord

Over de Laatste stelling van Fermat en alle ontwikkelingen eromheen is in de recente geschiedenis enorm veel geschreven, van bestseller-romans tot mathematische meesterwerken. Wat wij echter misten was een goed leesbare, wiskundige introductie tot dit boeiende onderwerp die voor eerste- en tweedejaars wiskundestudenten toegankelijk is; we hebben geprobeerd dit gat in de literatuur een beetje op te vullen. Vrijwel geen voorkennis wordt verondersteld: alles wat we nodig hebben op onze reis, zoals groepen, ringen en lichamen voeren we ter plekke in, zodat dit boek misschien zelfs voor gemotiveerde VWO-leerlingen waardevol is. Alleen een aantal basisgebrippen uit de wiskunde, zoals priemontbindingen, modulorekenen en complexe getallen veronderstellen we als bekend, de eerste hoofdstukken van Frits Beukers' boekje [2] leveren meer dan genoeg voorkennis. In de voorbeelden is soms wat meer wiskunde nodig, maar daar moet je je niet door laten afschrikken. Achterin is een uitgebreide index opgenomen, waar ook notaties en gebruikte afkortingen zijn terug te vinden. Vooral in het laatste hoofdstuk van dit boek, die als korte inleiding tot de algebraïsche getaltheorie kan worden gezien, loopt de moeilijkheidsgraad aardig op, maar ook als je dat overslaat kun je een aardige indruk krijgen van een stukje mooie wiskunde dat gebruikt is om de Stelling te beteugelen.

Dit 'boek' is begonnen als werkstuk voor het vak Caleidoscoop, de auteurs waren toen het werd ingeleverd (april 2011) eerstejaars wiskundestudenten aan de Universiteit Utrecht. Het was nooit tot stand gekomen zonder de prachtige boeken en dictaten die we met plezier doorgewerkt en gebruikt hebben, zie de literatuurlijst achterin. We willen professor Jan Hogendijk bedanken voor het lezen en verbeteren van het manuscript, en het stimuleren een verbeterde versie openbaar te maken.

Er zullen ongetwijfeld nog fouten en onduidelijkheden in dit werk zitten en we zouden het fijn vinden als je die aan ons doorgeeft, dat kan via e-mail naar lars.vd.berg@kpnmail.nl. Suggesties voor verbetering zijn ook van harte welkom.

Let op. Als je dit boek slechts gedeeltelijk wilt lezen, loont het niet de moeite om alleen de eerste vier hoofdstukken door te spitten. Dat levert een heel verkeerd beeld van de moderne getaltheorie, de hoofdstukken op zichzelf zijn niet echt motiverend om je er verder in te verdiepen. Beter is het dan om bij hoofdstuk 5 te beginnen; het nadeel is dan wel dat je de historische aanleiding en de bewijzen van Fermat's laatste stelling voor $n = 2, 3, 4$ (gedeeltelijk) mist, maar in elk geval zijn deze hoofdstukken meer representatief voor de moderne wiskunde. Als je al bent ingewijd in de theorie van groepen, ringen en/of lichamen kun je hoofdstuk 5, 6 en/of 7 overslaan of snel doorlezen. Hoe je het ook leest, we hopen dat we erin geslaagd zijn een mooi doorkijkje te geven naar een enorm grote en rijke theorie.

Inhoudsopgave

Hoofdstuk 1

Introductie

De Laatste Stelling van Fermat is een van de meest tot de verbeelding sprekende problemen uit de wiskunde. Het begint al met de bijzondere ontstaansgeschiedenis. De geestelijk vader van de stelling, Pierre de Fermat (1601 – 1665), wordt vaak gezien als een van de grootste wiskundigen van zijn tijd, maar eigenlijk was hij geen beroepswiskundige. Fermat was jurist aan het hof van Toulouse, en in zijn vrije tijd deed hij bij wijze van uit de hand gelopen hobby veel aan wiskunde. Er bestonden nog geen wetenschappelijke tijdschriften, en Fermat publiceerde nooit boeken. In plaats daarvan correspondeerde hij uitgebreid met andere wiskundigen via brieven, waarin hij hen vaak uitdaagde een probleem op te lossen. Bovendien had hij een Latijnse vertaling van Diophantus' *Arithmetica*, waarin hij regelmatig notities maakte. Fermat's oudste zoon besefte het belang van zijn vaders werk, en publiceerde na diens dood een nieuwe versie van de *Arithmetica*, nu inclusief de notities.

Figuur 1.1: Pierre de Fermat

Onder de noties was een groot aantal elegante, diepgaande stellingen, maar helaas bijna altijd met slechts een cryptische schets van een bewijs, of zelfs zonder bewijs. Eén van de notities is wereldberoemd geworden. Fermat schreef, vrij vertaald, in de kantlijn van zijn boek:

Het is onmogelijk een getal die een macht is groter dan de tweede te schrijven als een som van twee soortgelijke machten.

Dit staat bekend als de Laatste Stelling van Fermat. Met een getal bedoelde Fermat een positief geheel getal, dus Fermat's opmerking kunnen we schrijven als:

Stelling 1.0.1. De Laatste Stelling van Fermat. *Voor geen enkel geheel getal n groter dan 2 bestaan er positieve gehele getallen x, y, z die voldoen aan*

$$x^n + y^n = z^n.$$

Hij voegde eraan toe: *Ik heb een wonderbaarlijk mooi bewijs voor deze stelling, maar de kantlijn is te klein om het te bevatten.* Niemand weet of Fermat ooit een bewijs had, het had dus eigenlijk het vermoeden van Fermat moeten heten. Wiskundigen na Fermat bewezen (en in een enkel geval ontkrachtten) in de loop van vele jaren één voor één al de stellingen uit de kantlijn, en deze bleef als laatste over. Waarschijnlijk komt daar de naam 'laatste stelling' vandaan.

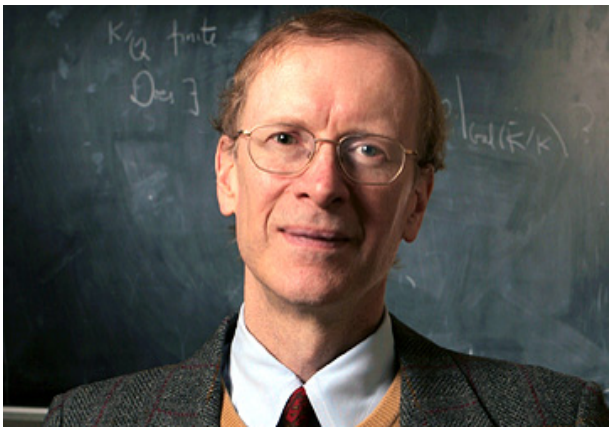
Fermat's laatste stelling is er een van uitersten. De stelling oogt heel eenvoudig en is voor iedereen te begrijpen, maar om het te bewijzen is uiterst moeilijk. En het probleem heeft zowel een grote aantrekkingskracht op beroepswiskundigen als op amateurs. Zeker nadat de wiskundige Wolfskehl in 1908 een prijs van 100.000 mark uitloofde (in die tijd meer waard dan nu een miljoen euro) voor het eerste correcte bewijs, stroomde de 'bewijzen' binnen. Elk bewijs moest officieel nauwkeurig worden nagekeken door het wiskundig instituut van Göttingen, en professor Landau die hiervoor tussen 1909 en 1934 verantwoordelijk was kreeg tientallen inzendingen per maand te verwerken die stuk voor stuk fout waren. Hij loste dit slim op door elke nieuwe inzending bij wijze van huiswerkopgave aan een van zijn studenten mee te geven, samen met een voorgedrukt invulformulier dat naar de afzender werd teruggestuurd. Hierop stond onder meer, vrij vertaald: "Dank u voor uw bewijs. De eerste fout staat op pagina ..., regel .... Dit ontkracht het bewijs."

De belangstelling van beroepswiskundigen voor het probleem daalde geleidelijk, maar er bleven er altijd die erdoor gefascineerd werden. Een van hen is de Britse wiskundige Andrew Wiles (geboren in 1953), die als tienjarig jongetje voor het eerst over het probleem las. In de jaren daarna probeerde hij het als scholier op te lossen, en zodoende werd het een levenslange passie. Na zijn wiskundestudie raadden wiskundigen hem af om er serieus aan te gaan werken, want de Laatste Stelling van Fermat werd niet meer als echt belangrijk voor de ontwikkeling van de wiskunde beschouwd, eerder als een uitdagende puzzel. Maar in 1986 werd een link gelegd tussen Fermat en de hedendaagse getaltheorie: Ken Ribet bewees dat de Laatste stelling van Fermat volgt uit het veel grotere Taniyama-Shimura vermoeden. Dit vermoeden legt een diepe link tussen de algebra en de analyse: heel globaal zegt het dat elke *elliptische kromme*, dat is een meetkundig geval met verbazingwekkende getaltheoretische eigenschappen, te beschouwen is als een *modulaire vorm*, een gevaarte uit de complexe analyse die zich bevindt in de hyperbolische ruimte en onvoorstelbaar veel symmetrie heeft. In de woorden van Barry Mazur: Het Taniyama-Shimura vermoeden is niet alleen een brug tussen twee schijnbaar totaal verschillende werelden, het is een compleet woordenboek dat op heel vruchtbare manier inzichten, stellingen en intuïties van de ene naar de andere wereld vertaalt.

Een oplossing van $x^n + y^n = z^n$ zou aanleiding geven tot een zo eigenaardige elliptische kromme dat het geen modulaire vorm kan zijn: als het Taniyama-Shimura vermoeden waar is, volgt dus ook Fermat's bewering. Toen Wiles dit hoorde, besloot hij al zijn energie aan dit probleem te wijden. Zeven jaar lang heeft hij eraan gewerkt, en om ongewenste publiciteit te voorkomen deed hij dat in het geheim. Vrijwel niemand geloofde in die tijd serieus dat de wiskunde zo ver was dat het vermoeden van Taniyama en Shimura kon worden opgelost.

In juni 1993 organiseerde Wiles een drietal lezingen in Cambridge, en er gingen geruchten rond dat hij het Taniyama-Shimura vermoeden, en daarmee de Laatste stelling van Fermat zou gaan bewijzen. Dat bleek inderdaad zo te zijn, en Wiles werd in één klap wereldberoemd. In kranten over de hele wereld stond op de voorpagina, naast een foto van Wiles, eindelijk

eens een formule. Wiles wist echter dat hij niet te vroeg moest juichen, want het bewijs was behalve erg mooi ook erg moeilijk, en het moest nog worden gecontroleerd door experts. Dit duurde maanden, en dagelijks correspondeerde het team met Wiles over problemen die vaak relatief eenvoudig waren op te lossen. Maar op een gegeven moment kwam een subtiel maar fundamenteel gat in het bewijs aan het licht. Het werd steeds duidelijker dat het niet zomaar was te dichten, het gat bleek een ravijn te zijn waar je niet makkelijk omheen kunt. Wiles zei hierover: 'It was an error in a crucial part of the argument, but it was something so subtle that I'd missed it completely until that point. The error is so abstract that it can't really be described in simple terms. Even explaining it to a mathematician would require the mathematician to spend two or three months studying that part of the manuscript in great detail.'



Figuur 1.2: Andrew Wiles

Wiles besloot helemaal terug te gaan naar een idee dat hij jaren daarvoor opzij had gezet, om zo met een grote omweg toch bij een bewijs van de stelling te komen. Na nog een jaar hard werken kreeg hij een schitterende openbaring die het hele probleem oploste, en nog op elegantere manier dan eerst ook. 'It was so simple and so elegant. I just stared in disbelief for 20 minutes. During the day I walked around the departement, I kept coming to my desk looking to see it was still there. It was still there. The first night I went back and slept on it, I checked through it again the next morning and by eleven o' clock I was satisfied and I went down to my wife; "I've got it, I think I've got it, I've found it." It was so unexpected, I think she thought I was talking about a toy or something and said "what?". I said, "I fixed my proof. I've got it." ' Ruim 350 jaar nadat Fermat zijn notitie maakte, was het raadsel eindelijk opgelost.

Het is onmogelijk dat het bewijs van Andrew Wiles ook maar enigszins lijkt op het bewijs van Fermat, als die er al een had. Het ruim honderd pagina's tellende bewijs van Wiles gebruikt de meest geavanceerde wiskunde van de twintigste eeuw, zoals elliptische krommen, modulaire vormen, Galois representaties, algebraïsche meetkunde en algebraïsche getaltheorie. Uiteindelijk is het meesterwerk van Wiles gebouwd op onmisbare ideeën van duizenden wiskundigen die leefden ná Fermat.

Toen het bewijs verscheen was het slechts voor een handjevol experts te begrijpen, maar inmiddels is het door een aantal wiskundigen bewerkt en vereenvoudigd waardoor het voor masterstudenten wiskunde toegankelijk is gemaakt in het boek [16]. Dat vereist echter een stuk meer ervaring dan dat wij hebben, daarom zullen we ons in het geheel niet met het bewijs van Wiles bezighouden. Wat we wél gaan doen, is de stelling bewijzen voor bepaalde exponenten n. Fermat had in elk geval een bewijs voor $n = 4$, we zullen een bewijs geven dat zelfs voor gevorderde vwo-leerlingen te begrijpen is. Meer dan honderd jaar later, rond 1750, gaf Leonhard Euler een bewijs voor $n = 3$, hoewel die eerst een fundamentele fout bevatte.

Later werd het gat gedicht met lemma's afkomstig van Euler, en daarom wordt het eerste overgeleverde bewijs van het geval $n = 3$ meestal aan Euler toegeschreven. Na Euler werden ook nog $n = 5$, $n = 14$ en $n = 7$ opgelost, maar omdat alle methoden eigenschappen van die specifieke getallen n gebruikten, leek een algemene oplossing nog ver weg. Grote vooruitgang werd pas weer geboekt toen rond 1850 Ernst Kummer met een schitterende methode kwam die het probleem in één keer voor een grote klasse van n oploste. In dit boek proberen we, naast de gevallen $n = 3$, 4 en 5 te behandelen, toe te werken naar Kummer's methode. Met een variant van die methode kunnen de Laatste Stelling van Fermat in elk geval oplossen voor alle n kleiner dan 23.

Het gebeurt in de wiskunde vaak dat men een stelling probeert te bewijzen, maar dat de daarvoor ontwikkelde theorie uiteindelijk mooier of in elk geval belangrijker blijkt dan de stelling zelf. Zo is het met de Laatste Stelling van Fermat ook. Allereerst is met het Taniyama-Shimura vermoeden een van de belangrijkste problemen uit de hedendaagse wiskunde opgelost. Maar ook de voor ons te begrijpen theorie van vroegere wiskundigen die Fermat's laatste stelling probeerden te bewijzen, heeft de wiskunde vaak een enorme groeispurt gegeven. Kummer's ideeën legden bijvoorbeeld de fundamenten van de algebraïsche getaltheorie. Daarom zullen we, behalve Fermat's stelling zelf, ook veel aandacht besteden aan de wiskunde die eromheen werd ontwikkeld.[1]

[1]Bij het schrijven van deze introductie hebben we vooral Singh's boekje [6] gebruikt. De citaten van Wiles komen grotendeels uit een interview met Nova, te vinden op www.pbs.org/wgbh/nova/physics/andrew-wiles-fermat.html.

Hoofdstuk 2

Pythagorese drietallen

De Laatste Stelling van Fermat (of afgekort FLT, van *Fermat's Last Theorem*) heeft, met een beetje fantasie, zijn wortels bij de oude Grieken. De notitie in de marge van de Arithmetica maakte Fermat namelijk bij een probleem dat ging over Pythagorese drietallen, dat zijn geheeltallige oplossingen van de vergelijking $x^2 + y^2 = z^2$. De Grieken wisten al veel over zulke drietallen, onder meer dat er oneindig veel oplossingen zijn. Fermat vroeg zich af wat er zou gebeuren als je 2 door willekeurige n vervangt, en schreef zijn beroemde 'ontdekking' op.

Misschien juist omdat FLT het geval $n = 2$ uitsluit, is het interessant om hier ons onderzoek te beginnen. Wat is er immers zo speciaal aan het getal 2? Bovendien kunnen we misschien meer van de vergelijking $x^n + y^n = z^n$ begrijpen als we inzien waarom de argumenten die we voor $n = 2$ gebruiken, niet opgaan voor andere n. Het blijkt zelfs zo te zijn dat die argumenten na wat aanpassingen heel goed bruikbaar zijn om voor bepaalde n juist te bewijzen dat er *geen* oplossingen bestaan.


DIOPHANTI
ALEXANDRINI
ARITHMETICORVM
LIBRI SEX.
ET DE NVMERIS MVLTANGVLIS
LIBER VNVS.
Nunc primum Græcè & Latinè editi, atque absolutissimis Commentariis illustrati.
AVCTORE CLAVDIO GASPARE BACHETO
MEZIRIACO SEBVSIANO, V.C.
LVTETIAE PARISIORVM,
Sumptibus SEBASTIANI CRAMOISY, via Iacobæa, sub Ciconiis.
M. DC. XXI.
CVM PRIVILEGIO REGIS

Figuur 2.1: Voorkant van Bachet's vertaling van de Arithmetica

In dit en het volgende hoofdstuk bedoelen we met 'getallen' altijd 'natuurlijke getallen', tenzij anders vermeld. We beschouwen in navolging van Fermat 0 niet als natuurlijk getal, dus met $\mathbb{N}$ bedoelen we de verzameling $\{1, 2, 3, \ldots\}$; de verzameling $\{\ldots -2, -1, 0, 1, 2, \ldots\}$ van gehele getallen duiden we aan met $\mathbb{Z}$. In Fermat's tijd stond men nog huiverig tegenover het getal 0, en zeker de negatieve getallen. De Laatste stelling van Fermat is echter waar voor *alle* gehele getallen x, y, z allen ongelijk nul. Voor $n = 4$ en $n = 3$ richten we ons echter alleen op natuurlijke getallen om tot snelle resultaten te komen. Later gaan we ook negatieve getallen erbij betrekken.

2.1 Een paar handigheidjes voor alle exponenten

Het is niet moeilijk om wat geheeltallige oplossingen te vinden van $x^2 + y^2 = z^2$. Na even proberen zien we bijvoorbeeld dat $3^2 + 4^2 = 5^2$ en $5^2 + 12^2 = 13^2$. Onder een Pythagorees drietal verstaan wij een rijtje (x, y, z) van natuurlijke getallen die voldoen aan $x^2 + y^2 = z^2$. Een paar vragen komen al snel op: Hoeveel van die drietallen zijn er? Zijn het er oneindig veel? Kunnen we een lijst met alle Pythagorese drietallen genereren, en zo ja hoe? Op de eerste twee vragen kunnen we meteen antwoord geven: het zijn er oneindig veel. Uit $3^2 + 4^2 = 5^2$ volgen bijvoorbeeld ook de oplossingen $6^2+8^2 = 10^2$, $9^2+12^2 = 15^2$, ... Dit geldt ook in het algemeen, en omdat het toch geen extra moeite kost en straks wel bruikbaar is, generaliseren we het gelijk naar alle n. Eerst voeren we wat notatie in. Zij n een natuurlijk getal groter dan 1. Een rijtje (x, y, z) van natuurlijke getallen die voldoen aan de Fermat-vergelijking $x^n + y^n = z^n$ noemen we kortweg een *n-drietal*. De termen 2-drietal en Pythagorees drietal zullen we door elkaar gebruiken.

Lemma 2.1.1. *Zij (x, y, z) een n-drietal en a een natuurlijk getal. Dan is (ax, ay, az) ook een n-drietal. Als x, y, z een gemene deler d hebben, dan is ook $(x/d, y/d, z/d)$ een n-drietal.*

Bewijs. Als we de vergelijking $x^n + y^n = z^n$ links en rechts met a^n vermenigvuldigen, dan krijgen we $a^n x^n + a^n y^n = a^n z^n$, ofwel $(ax)^n + (ay)^n = (az)^n$. Dit bewijst de eerste bewering. De tweede bewering volgt door in dezelfde redenering $a = 1/d$ in te vullen. □

We hoeven dus alleen nog te zoeken naar Pythagorese drietallen waarvoor 1 de enige gemene deler is van x, y en z. Men kort dit vaak af door te zeggen dat x, y, z relatief priem of copriem zijn, of dat hun *grootste gemene deler* (ggd) 1 is. De volgende stelling laat zien dat x, y, z zelfs alleen *paarsgewijs copriem* hoeven te zijn, dat wil zeggen dat elk van de paren (in dit geval (x, y), (x, z) en (y, z)) copriem is.[1]

Lemma 2.1.2. *Zij (x, y, z) een n-drietal waarvoor x, y, z copriem zijn. Dan zijn ze ook paarsgewijs copriem.*

Bewijs. Stel dat p een gemene priemdeler is van x en y. Dan $p|x^n$ en $p|y^n$, dus ook $p|(x^n + y^n)$, ofwel $p|z^n$. Een fundamentele eigenschap van priemgetallen is dat uit $p|ab$ volgt dat $p|a$ of $p|b$. Dus uit $p|z \cdots z$ volgt dat $p|z$ of ... of $p|z$, ofwel $p|z$. We concluderen dat p een gemene priemdeler is van x, y en z, tegenspraak. We concluderen dat x en y geen gemene priemdelers hebben, dus $\text{ggd}(x, y) = 1$.

Op dezelfde manier volgt dat $\text{ggd}(x, z) = 1$ en $\text{ggd}(y, z) = 1$. Als namelijk p priemdeler zou zijn van x en z, dan zou p deler zijn van $z^n - x^n = y^n$ zodat $p|y$, tegenspraak. Wegens symmetrie in x en y hebben ook y en z geen gemene priemdeler. We concluderen dat $\text{ggd}(x, y) = \text{ggd}(x, z) = \text{ggd}(y, z) = 1$, dus x, y en z zijn paarsgewijs copriem. □

Voor het gemak zullen we n-drietallen (x, y, z) met x, y, z paarsgewijs copriem soms *primitieve* n-drietallen noemen. De vorige twee lemma's laten zien dat we ons alleen op de primitieve n-drietallen hoeven te richten: de n-drietallen zijn precies de veelvouden hiervan.

[1] De notatie $a|b$ betekent 'a is deler van b'.

Nu we toch bezig zijn, doen we nog wat algemene opmerkingen die het zoeken naar n-drietallen gemakkelijker maken. We kunnen ons bijvoorbeeld afvragen welke elementen van zo'n drietal even of oneven zijn.

Lemma 2.1.3. *Stel (x, y, z) is een primitief n-drietal. Dan is precies één van de x, y, z even.*

Bewijs. Uit Lemma 2.1.2 volgt dat x, y, z paarsgewijs copriem zijn, dus in het bijzonder is hoogstens één van hen even. Aan de andere kant kunnen ze niet alledrie oneven zijn. Stel namelijk dat x en y oneven zijn, dan zijn ook x^n en y^n oneven, en hun som z^n is dus even. Maar dan moet z wel even zijn. □

Als de exponent n even is, kunnen we zelfs nog meer zeggen over de pariteit (het even of oneven zijn) van x, y, z.

Gevolg 2.1.4. *Stel (x, y, z) is een n-drietal met n even. Dan is x of y even, de andere twee zijn oneven.*

Bewijs. Na het vorige lemma is het voldoende om te bewijzen dat z niet degene kan zijn die even is. Stel z is even, en dus x en y oneven. Dan is z^n deelbaar door 4, dus $z^n \equiv 0 \pmod 4$. Kwadraten van oneven getallen zijn altijd congruent 1 modulo 4, want $1^2 \equiv 1 \pmod 4$ en ook $3^2 \equiv 9 \equiv 1 \pmod 4$. De getallen x^n en y^n zijn kwadraten, namelijk van $x^{n/2}$ en $y^{n/2}$. Dus $x^n + y^n \equiv 1 + 1 \equiv 2 \not\equiv 0 \equiv z^n \pmod 4$, dus $x^n + y^n \neq z^n$, tegenspraak. □

Tot slot doen we een eenvoudige maar heel belangrijke opmerking. Als we voor een zekere exponent n hebben bewezen dat $x^n + y^n = z^n$ geen oplossingen heeft, dan geldt dat ook voor alle veelvouden van n. Als er namelijk getallen x_0, y_0, z_0 zijn zodat $x_0^{kn} + y_0^{kn} = z_0^{kn}$, dan volgt uit $(x_0^k)^n + (y_0^k)^n = (z_0^k)^n$ meteen een oplossing van $x^n + y^n = z^n$. Omdat elk natuurlijk getal m groter dan twee een viervoud is of een oneven priemdeler heeft, volgt:

Lemma 2.1.5. *Om de Laatste stelling van Fermat aan te tonen, is het voldoende te bewijzen dat $x^n + y^n = z^n$ geen oplossingen heeft voor $n = 4$ en voor de oneven priemgetallen n.*

We kunnen ons dus hoofdzakelijk op de priemexponenten n richten, een feit dat later nog van fundamenteel belang zal blijken te zijn.

2.2 Hoe vinden we Pythagorese drietallen?

We zullen eerst kijken aan wat voor voorwaarden een drietal moet voldoen om een Pythagorees drietal te kunnen zijn, we filteren dus de kandidaten eruit. Daarna zullen we onderzoeken welke van hen inderdaad Pythagorees zijn. Stel (x, y, z) is een primitief Pythagorees drietal. Eén van beide x, y is even en de andere twee zijn oneven, en omdat de vergelijking $x^n + y^n = z^n$ symmetrisch is in x en y kunnen we zonder beperking aannemen dat x even is. We schrijven $x = 2u$ voor een natuurlijk getal u. Het belangrijkste idee is nu dat we de algebraïsche uitdrukking $y^2 - z^2$ ontbinden in factoren, in dit geval als $(y + z)(y - z)$. Dit is een kwadraat, namelijk x^2, en we proberen te bewijzen dat de factoren copriem zijn, en dat daaruit volgt dat beide factoren kwadraten zijn.

Omdat z en y beide oneven zijn, is hun verschil en hun som even, zeg $z+y=2v$ en $z-y=2w$. We hebben dus

$$(2u)^2 = y^2 - z^2 = (z+y)(z-y) = 2v2w,$$

ofwel $u^2 = vw$. We willen laten zien dat de grootste gemene deler van v en w gelijk is aan 1. Elke deler van v en w is deler van hun som en hun verschil, dat zijn

$$v+w = \tfrac{1}{2}(z+y) + \tfrac{1}{2}(z-y) = z \qquad \text{en} \qquad v-w = \tfrac{1}{2}(z+y) - \tfrac{1}{2}(z-y) = y.$$

Het is dus een gemene deler van de getallen y en z die we copriem verondersteld hebben, en is daarom gelijk aan 1. Dus v en w zijn relatief priem. Hun product is een kwadraat, en we willen aantonen dat ook v en w kwadraten zijn. We generaliseren weer naar willekeurige n.

Stelling 2.2.1. *Stel dat $a^n = bc$ voor gehele getallen a, b, c met b, c copriem, met n een natuurlijk getal groter dan 1. Dan zijn b en c allebei n-de macht van een geheel getal.*

Bewijs. Allereerst merken we op dat een geheel getal x een n-de macht is precies dan als alle exponenten in de priemontbinding van x veelvoud zijn van n; met andere woorden, n deelt de *orde* van elke priemfactor van x. Stel namelijk dat x een n-de macht is, dus $x = k^n$ voor een geheel getal k. Zij $k = \pm p_1^{m_1} \cdots p_t^{m_t}$ de priemfactorontbinding van k. Dan is[2]

$$x = k^n = (\pm p_1^{m_1} \cdots p_t^{m_t})^n = \pm (p_1^{m_1})^n \cdots (p_t^{m_t})^n = \pm p_1^{m_1 n} \cdots p_t^{m_t n},$$

dus de exponenten in de priemontbinding van x zijn allen veelvoud van n. Stel omgekeerd dat y een geheel getal is waarvoor n een gemeenschappelijk deler is van de exponenten in de priemontbinding van y. Er zijn dus priemgetallen $p_1, \ldots, p_t$ en natuurlijke getallen $m_1, \ldots, m_t$ zodat

$$y = \pm p_1^{m_1 n} \cdots p_t^{m_t n} = \pm (p_1^{m_1})^n \cdots (p_t^{m_t})^n = (\pm p_1^{m_1} \cdots p_t^{m_t})^n,$$

dus y is een n-de macht.

Om de stelling te bewijzen, beschouwen we de priemontbindingen van a^n, b en c:

$$a^n = \pm p_1^{m_1 n} \cdots p_t^{m_t n}, \qquad b = \pm q_1^{k_1} \cdots q_u^{k_u}, \qquad c = \pm r_1^{l_1} \cdots r_v^{l_v}$$

voor priemgetallen p_i, q_i, r_i en gehele getallen m_i, k_i, l_i. Omdat b en c copriem zijn, zijn alle priemfactoren van b en c verschillend: $q_i \neq r_j$ voor alle i, j. Een gemeenschappelijke priemfactor zou immers een gemeenschappelijke deler groter dan 1 zijn. De priemontbinding van bc is dus gelijk aan het product van de ontbindingen van b en c, en omdat $a^n = bc$ volgt dat

$$\pm p_1^{m_1 n} \cdots p_t^{m_t n} \qquad \text{en} \qquad \pm\, q_1^{k_1} \cdots q_u^{k_u} r_1^{l_1} \cdots r_v^{l_v}$$

twee priemontbindingen zijn van hetzelfde getal. Omdat priemontbindingen op volgorde na eenduidig bepaald zijn, volgt dat er voor elke i een j is waarvoor $q_i^{k_i} = p_j^{m_j n}$, en dat er voor elke i een j is waarvoor $r_i^{l_i} = p_j^{m_j n}$. De exponenten in de priemontbindingen van b en c zijn dus veelvouden van n, dus b en c zijn beide een n-de macht. □

[2]Met het teken $\pm$ bedoelen we in dit bewijs dat we niet weten of het teken plus of min is, maar dat dat ons ook niet uitmaakt.

Inderdaad zijn v en w dus kwadraten van natuurlijke getallen, zeg $v = p^2$ en $w = q^2$. Verder zijn p en q relatief priem, want elke gemene deler deelt ook de relatief prieme v en w. Bovendien hebben we

$$z = v + w = p^2 + q^2 \qquad \text{en} \qquad y = v - w = p^2 - q^2.$$

We zien hieruit dat p groter is dan q, want y is groter dan nul. Ook merken we op dat p en q van tegengestelde pariteit zijn, want z is oneven. We kunnen nu ook x uitdrukken in termen van p en q:

$$x^2 = z^2 - y^2 = (p^2 + q^2)^2 - (p^2 - q^2)^2 = (p^4 + 2p^2q^2 + q^4) - (p^4 - 2p^2q^2 + q^4) = 4p^2q^2 = (2pq)^2.$$

Dus $x = 2pq$ (niet $-2pq$, want x, p, q zijn positief). We hebben nu flink wat noodzakelijke voorwaarden gevonden wil (x, y, z) een primitief Pythagorees drietal zijn: er moeten p, q zijn, relatief priem, van tegengestelde pariteit en met $p > q$, zodat $x = 2pq, y = p^2 - q^2, z = p^2 + q^2$. Omgekeerd, als we p en q willekeurig nemen, dan volgt door haakjes uitwerken dat

$$(2pq)^2 + (p^2 - q^2)^2 = (p^2 + q^2)^2,$$

dus $(2pq, p^2 - q^2, p^2 + q^2)$ is Pythagorees. De primitieve Pythagorese drietallen zijn dus *precies* de drietallen van die vorm, met bovendien $p > q$, en p, q copriem van tegengestelde pariteit.[3] We kunnen dus in principe een tabel met alle primitieve Pythagorese drietallen opstellen door om te beginnen $p = 2$ te stellen, q alle getallen kleiner dan p te laten doorlopen die copriem zijn met p, en $(x, y, z) = (2pq, p^2 - q^2, p^2 + q^2)$ te stellen. Dan hogen we p een op, et cetera. In het bijzonder zijn er dus oneindig veel 'echt verschillende' Pythagorese drietallen. Uit gebrek aan papier hebben we de tabel beperkt tot $p = 8$.

Tabel 2.1: Enkele Pythagorese drietallen

p	q	x	y	z
2	1	4	3	5
3	2	12	5	13
4	1	8	15	17
4	3	24	7	25
5	2	20	21	29
5	4	40	9	41
6	1	12	35	37
6	5	60	11	61
7	2	28	45	53
7	4	56	33	65
7	6	84	13	85
8	1	16	63	65
8	3	48	55	73
8	5	80	39	89
8	7	112	15	113

[3] We beschouwen hierbij bijvoorbeeld $(3, 4, 5)$ en $(4, 3, 5)$ als gelijk.

Hoofdstuk 3

Fermat en het geval $n = 4$

Toen Pierre de Fermat schreef dat hij een opmerkelijk mooi bewijs had van zijn 'stelling', zal hij ongetwijfeld geïnspireerd zijn geweest door de constructie van Pythagorese drietallen. Fermat stuurde, waarschijnlijk nadat hij zijn mysterieuze aantekening had gemaakt, wel eens brieven waarin hij bijvoorbeeld beweerde dat $x^4 + y^4 = z^4$ geen oplossingen heeft, maar het algemene geval $x^n + y^n = z^n$ komt in zijn brieven niet voor. Het is daarom niet ondenkbaar dat hij later ontdekte dat zijn bewijs alleen klopt voor bepaalde exponenten n. In elk geval heeft hij $n = 4$ opgelost, met de door hem bedachte methode van oneindige afdaing (*infinite descent*). Die berust op het feit dat er geen oneindig lange strikt monotoon dalende rij natuurlijke getallen bestaat: begin je met n, dan kom je na hoogstens $n - 1$ stappen uit bij 1. Stel we willen bewijzen dat er geen natuurlijk getal is met bepaalde eigenschappen. Als we uit het bestaan van zo'n getal een *kleiner* natuurlijk getal kunnen construeren met dezelfde eigenschappen, dan zijn we klaar. We kunnen hieruit immers weer een kleiner getal construeren, daaruit een nog kleiner getal, enzovoort; zo krijgen we een oneindig lange dalende rij natuurlijke getallen, en dat kan niet.

We zullen zien dat het bewijs veel weg heeft van de constructie van Pythagorese drietallen. We zoeken op vergelijkbare wijze noodzakelijke voorwaarden voor (x, y, z) om een 4-drietal te zijn, maar dit keer zijn die voorwaarden zo beperkend dat geen enkel drietal voldoet. De methode van oneindige afdaling werkt echter niet voor de vergelijking $x^4 + y^4 = z^4$ zelf: we weten niet hoe we hier kleinere oplossingen van kunnen vinden. Wat we wel kunnen doen, is uit $x^4 + y^4 = z^4$ een oplossing van een andere vergelijking construeren, en daarop kunnen we wél het principe van oneindige afdaling toepassen.

Stel dat (α, β, γ) een primitief 4-drietal is. We weten van Gevolg 2.1.4 dat α of β even is, zonder beperking nemen we aan dat dat α is. Omdat $\alpha^4 + \beta^4 = (\gamma^2)^2$, is $(\alpha, \beta, \gamma^2)$ oplossing van de vergelijking $x^4 + y^4 = z^2$. Elke gemene priemdeler van γ^2 deelt ook γ. Omdat α, β, γ paarsgewijs copriem zijn, zijn α, β, γ^2 dat ook. We hebben dus een oplossing gevonden van de vergelijking[1]

$$x^4 + y^4 = z^2, \qquad x, y, z \text{ zijn copriem}, \qquad x \text{ is even}. \tag{3.1}$$

We gaan met oneindige afdaling laten zien dat déze vergelijking geen oplossing heeft. Stel er

[1] Met 'x, y, z zijn copriem' bedoelen we dat 1 het enige positieve getal is dat hen *alledrie* deelt, copriem is dus minder sterk dan paarsgewijs copriem.

is een oplossing x, y, z. Dan is $(x^2)^2 + (y^2)^2 = z^2$, *dus* (x^2, y^2, z) *is een Pythagorees drietal.* Wegens de aannamen op x, y, z zijn ook x^2, y^2, z copriem, en volgens Lemma 2.1.2 zijn ze paarsgewijs copriem. Verder is x^2 even. Van het vorige hoofdstuk weten we dat er daarom natuurlijke, coprieme getallen p, q zijn van tegengestelde pariteit, $p > q$, zodat

$$x^2 = 2pq, \qquad y^2 = p^2 - q^2, \qquad z = p^2 + q^2. \tag{3.2}$$

Uit de middelste vergelijking volgt dat $y^2 + q^2 = p^2$, dus (y, q, p) *is een Pythagorees drietal.* Bovendien zijn p, q relatief priem, dus zeker y, q, p zijn copriem, en uit Lemma 2.1.2 volgt dat (y, q, p) primitief is. We weten dat dit betekent dat p oneven is, en omdat p, q tegengestelde pariteit hebben is q even. Er zijn dus natuurlijke getallen r, s zodat

$$q = 2rs, \qquad y = r^2 - s^2, \qquad p = r^2 + s^2, \tag{3.3}$$

met r, s copriem, van tegengestelde pariteit en met $r > s$. Met (3.2) en (3.3) kunnen we x^2 in r en s uitdrukken:

$$x^2 = 2pq = 4(r^2 + s^2)rs.$$

Omdat x even is, is $(r^2 + s^2)rs$ een kwadraat:

$$(r^2 + s^2)rs = (\tfrac{1}{2}x)^2. \tag{3.4}$$

Stel t is een gemene priemdeler van $r^2 + s^2$ en rs. Omdat $t|rs$ is t deler van r of van s, laten we aannemen dat $t|s$. Dan deelt t ook s^2, en dus ook het verschil $(r^2 + s^2) - s^2 = r^2$. Dus t deelt r en is dus gemene priemdeler van r en s, tegenspraak want die zijn copriem. Wegens symmetrie in r en s volgt die tegenspraak ook als we $t|r$ hadden aangenomen. We concluderen dat $r^2 + s^2$ en rs geen gemene priemdelers hebben, ze zijn dus copriem. Uit (3.4) en Stelling 2.2.1 volgt dat het kwadraten moeten zijn, zeg

$$r^2 + s^2 = a^2 \qquad \text{en} \qquad rs = b^2. \tag{3.5}$$

Maar r en s zijn relatief priem, en door Stelling 2.2.1 nóg eens toe te passen volgt uit de tweede vergelijking dat r en s kwadraten zijn, zeg

$$r = X^2 \qquad \text{en} \qquad s = Y^2.$$

Uit (3.5) zien we nu een nieuwe oplossing van de vergelijking $x^4 + y^4 = z^2$, namelijk

$$X^4 + Y^4 = a^2.$$

Bovendien zijn r, s ofwel X^2, Y^2 copriem, dus X, Y zijn copriem zodat ook X, Y, a copriem zijn. Verder weten we dat r, s van tegengestelde pariteit zijn, dus X of Y is even. Door zo nodig de rollen van X en Y te verwisselen, hebben we dus weer een oplossing van vergelijking (3.1). Bovendien is de oplossing in zekere zin kleiner dan die waarmee we begonnen:

$$X^4 + Y^4 = a^2 = r^2 + s^2 = p < p^2 + q^2 = z \leq z^2 = x^4 + y^4.$$

We kunnen dus een oneindige rij oplossingen $(x_0, y_0, z_0), (x_1, y_1, z_1), (x_2, y_2, z_2), \ldots$ construeren van de vergelijking $x^4 + y^4 = z^2$, en steeds is het natuurlijke getal $x_{k+1}^4 + y_{k+1}^4$ kleiner dan zijn voorganger $x_k^4 + y_k^4$. Met oneindige afdaling leidt dit tot een tegenspraak. We concluderen dat (3.1) geen oplossing heeft, dus er bestaat ook vier 4-drietal (α, β, γ). Dit bewijst de Laatste stelling van Fermat voor $n = 4$.[2]

[2]Dit bewijs is een bewerking van het analoge bewijs in [4], die weer een bewerking van Fermat's bewijs is.

Hoofdstuk 4

Euler en het geval $n = 3$

Toen de vertaling van de Arithmetica met Fermat's notities eenmaal bekendheid had gekregen onder wiskundigen, probeerden velen van hen de beweringen van Fermat te begrijpen en op te lossen. Onder hen was de Zwitser Leonhard Euler (1707 – 1783), misschien wel de grootste wiskundige van zijn tijd. Hij begon zijn studie met de bedoeling geestelijke te worden, maar hij raakte al snel in de ban van de wiskunde. Het grootste deel van zijn leven werkte hij aan de wetenschappelijke academies van St. Petersburg en Berlijn, en aan bijna alle takken van de wiskunde die toen bestonden heeft hij grote bijdragen geleverd. Hij was wat dat betreft universalist en bovendien heel productief. Toen hij ongeveer 31 jaar oud was, werd hij blind aan één oog, en rond zijn 58ste werd hij geheel blind. Hij had echter een fotografisch geheugen en kon uitzonderlijk goed hoofdrekenen, zijn productiviteit nam eerder toe dan af. De helft van zijn wetenschappelijk werk heeft hij verricht terwijl hij blind was, en na zijn dood heeft men nog 50 jaar gewerkt aan de publicatie van zijn artikelen.[1]

Figuur 4.1: Leonhard Euler, geschilderd rond 1750 door Johann Brucker

Euler heeft een aantal mooie stellingen bewezen die Fermat zonder echt bewijs in de kantlijn had geschreven. Om er een paar te noemen:

Elk priemgetal van de vorm $4n + 1$ *is de som van twee kwadraten.*

Elk priemgetal van de vorm $3n + 1$ *is van de vorm* $a^2 + 3b^2$.

Als a, p *getallen zijn en* p *priem, dan is* a^p *van de vorm* $a + np$.

[1] Voor meer biografische gegevens over Euler en andere wiskundigen zie www-history.mcs.st-andrews.ac.uk.

Deze stellingen zullen we later in dit boek bewijzen, als 'bijproduct' van de theorie die we ontwikkelen om de Laatste stelling van Fermat aan te pakken. Het bevestigt weer eens dat veel ogenschijnlijk verschillende wiskundige problemen met elkaar samenhangen.

Euler probeerde ook Fermat's laatste stelling op te lossen, maar dat lukt niet. Wel gaf hij het eerste overgeleverde bewijs van het speciale geval $n = 3$. Dit bewijs bevatte echter een fundamentele fout, het maakte namelijk gebruik van een bewerking die hij niet juist bewees. Later werd deze bewering door anderen bewezen, in essentie bleek het een gevolg van eerder door Euler bewezen lemma's.

Het bewijs dat we geven is dat van Euler, alleen dan met moderne notatie en in eigen woorden opgeschreven.[2] Het is duidelijk gecompliceerder dan het bewijs van $n = 4$, maar op één punt na is het nog steeds elementair. Dat ene punt is precies het deel waar Euler de mist in ging. Hier ontbindt hij een algebraïsche uitdrukking in factoren die een bepaald soort complexe getallen zijn, en heeft het over relatief priem zijn van deze complexe getallen. Hier zullen we na het hoofdbewijs van de volgende paragraaf verder op ingaan, en de observaties zullen aanleiding geven tot het ontwikkelen van een rijke theorie.

4.1 Euler's bewijs van het geval $n = 3$

Stel (x, y, z) is een primitief 3-drietal. De strategie is om hieruit eerst de volgende bewering af te leiden:

$$\begin{gathered}\textit{Er zijn } a, b \in \mathbb{N} \textit{ met } a, b \textit{ copriem en van tegengestelde pariteit, zodat} \\ 2a(a^2 + 3b^2) \textit{ een derde macht is.}\end{gathered} \tag{4.1}$$

Vervolgens proberen we hieruit af te leiden dat $2a$ en $a^2 + 3b^2$, behalve in een speciaal geval, onderling priem zijn, zodat uit Lemma 2.2.1 volgt dat het beide derde machten zijn. (Het speciale geval behandelen we apart.) Dit leidt tenslotte tot de conclusie dat er een primitief 3-drietal is dat 'strikt kleiner' is dan (x, y, z), zodat we met de methode van oneindige afdaling op een tegenspraak stuiten. Als we zeggen dat een drietal (x_1, y_1, z_1) strikt kleiner is dan (x_0, y_0, z_0), bedoelen we hier dat $z_1 < z_0$.

Uit Lemma 2.1.3 volgt dat precies één van de x, y, z even is. We onderscheiden twee gevallen.

Geval 1: z is even. Zonder beperking van de algemeenheid kunnen we veronderstellen dat $x \geq y$. Als $x = y$, dan volgt uit $\mathrm{ggd}(x, y) = 1$ dat $x = y = 1$ en dus is $x^3 + y^3 = 2$ geen derde macht, tegenspraak. Er geldt dus $x > y$.

Omdat x en y oneven zijn, zijn hun som en verschil even, zeg

$$x + y = 2a \qquad \text{en} \qquad x - y = 2b.$$

Omdat $x, y > 0$ en $x - y > 0$ volgt bovendien dat a en b natuurlijke getallen zijn.

We gaan $z^3 = x^3 + y^3$ uitdrukken in termen van a en b. Er geldt,

$$\begin{aligned} x &= \tfrac{1}{2}((x + y) + (x - y)) = \tfrac{1}{2}(2a + 2b) = a + b, \quad \text{en} \\ y &= \tfrac{1}{2}((x + y) - (x - y)) = \tfrac{1}{2}(2a - 2b) = a - b. \end{aligned}$$

[2] Het is een bewerking van het analoge bewijs dat we in [4] vonden.

Dus

$$\begin{aligned} z^3 = x^3 + y^3 &= (x+y)(x^2 - xy + y^2) \\ &= 2a\big((a^2 + 2ab + b^2) - (a^2 - b^2) + (a^2 - 2ab + b^2)\big) \\ &= 2a(a^2 + 3b^2), \end{aligned} \tag{4.2}$$

dus $2a(a^2 + 3b^2)$ is een derde macht. Omdat $x = a + b$ oneven is, zijn a en b van tegengestelde pariteit. Bovendien, elke gemene deler van a en b deelt ook $a + b = x$ en $a - b = y$ en is dus 1, dus a en b zijn copriem. We hadden al gezien dat $a, b \in \mathbb{N}$, dus bewering (4.1) is waar.

Geval 2: x of y is even; zonder beperking van de algemeenheid veronderstellen we dat x even is. Dus z en y zijn oneven, zodat

$$z - y = 2a \qquad \text{en} \qquad z + y = 2b$$

voor gehele getallen a en b. Omdat $z, y > 0$ is b een natuurlijk getal. En uit $z^3 = x^3 + y^3 > y^3$ volgt dat $z > y$, dus ook a is een natuurlijk getal. We gaan nu de derde macht $x^3 = z^3 - y^3$ schrijven in termen van a en b, dat gaat precies hetzelfde als daarnet: er geldt

$$\begin{aligned} z &= \tfrac{1}{2}((z+y) + (z-y)) = \tfrac{1}{2}(2b + 2a) = b + a \quad \text{en} \\ y &= \tfrac{1}{2}((z+y) - (z-y)) = \tfrac{1}{2}(2a - 2b) = b - a, \end{aligned}$$

zodat

$$\begin{aligned} x^3 = z^3 - y^3 &= (z-y)(z^2 + zy + y^2) \\ &= 2a((b^2 + 2ab + a^2) + (b^2 - a^2) + (b^2 - 2ab + a^2)) \\ &= 2a(a^2 + 3b^2). \end{aligned} \tag{4.3}$$

Omdat $z = b + a$ oneven is, zijn b en a van tegengestelde pariteit; elke gemene deler van b en a deelt ook $b + a = z$ en $b - a = y$, dus b en a zijn copriem. We zagen al dat $a, b \in \mathbb{N}$, dus ook nu is (4.1) waar.

We weten nu dus dat hoe dan ook (4.1) waar is. Als we konden bewijzen dat $2a$ en $a^2 + 3b^2$ onderling priem zijn, dan zou volgen dat het beide derde machten zijn. Dit is echter niet altijd waar: als bijvoorbeeld a veelvoud is van 3, dan is 3 deler van $2a$ en van $a^2 + 3b^2$ zodat ze niet copriem zijn. Omgekeerd, stel dat d een gemene priemfactor is van $2a$ en $a^2 + 3b^2$. Omdat a en b van tegengestelde pariteit zijn, is $a^2 + 3b^2$ oneven,[3] dus $d \neq 2$. Uit $d|2a$ volgt dus $d|a$. Dus d deelt $a^2 + 3b^2$ en a^2, en dus ook hun verschil $3b^2$. Als $d \neq 3$ volgt dat $d|b^2$, en omdat d priem is, is d deler van b. Maar d is ook deler van a, tegenspraak want a en b zijn copriem. Dus $d = 3$, en dit is een deler van a. We concluderen: $2a$ en $a^2 + 3b^2$ zijn copriem precies dan als 3 geen deler is van a. We onderscheiden daarom weer twee gevallen.

Geval 1: a is niet deelbaar door 3. Dan zijn dus $2a$ en $a^2 + 3b^2$ copriem, en omdat hun product een derde macht is, volgt uit Lemma 2.2.1:

$$2a \text{ en } a^2 + 3b^2 \text{ zijn beide derdemachten.} \tag{4.4}$$

[3] Immers, als a oneven en b even is, dan is $a^2 + 3b^2 \equiv 1^2 + 3 \cdot 0^2 \equiv 1 \pmod 2$, en in het andere geval is $a^2 + 3b^2 \equiv 0^2 + 3 \cdot 1^2 \equiv 3 \equiv 1 \pmod 2$.

Net als bij de constructie van Pythagorese drietallen zoeken we uitdrukkingen voor a en b, zodat $a^2 + 3b^2$ een derde macht is precies dan als a en b van deze vorm zijn. Dat is echter lang niet zo makkelijk als daar.

We beweren:

Stel twee gehele getallen a en b zijn copriem en van tegengestelde pariteit, en $a^2 + 3b^2$ is een derde macht. Dan zijn er gehele p, q zodat

$$a = p^3 - 9pq^2, \qquad b = 3p^2q - 3q^3. \tag{4.5}$$

Het bewijs hiervan stellen we uit tot Lemma 4.2.1. Dit is het deel waar Euler zijn fout maakte.

Uitgaande van (4.5) zijn er voor 'onze' a en b gehele getallen p en q zodat

$$a = p^3 - 9pq^2 = p(p^2 - (3q)^2) = p(p - 3q)(p + 3q), \qquad \text{en} \tag{4.6}$$

$$b = 3p^2q - 3q^3 = 3q(p^2 - q^2) = 3q(p - q)(p + q). \tag{4.7}$$

Uit de getallen p en q proberen we een strikt kleiner 3-drietal te construeren dan die waar we mee begonnen. Elke gemeenschappelijke deler d van p en q is volgens bovenstaande vergelijkingen deler van a en b, en omdat die copriem zijn, volgt dat $d = 1$. Dus

$$p \text{ en } q \text{ zijn copriem.} \tag{4.8}$$

Uit (4.4) en (4.6) volgt:

$$2p(p - 3q)(p + 3q) \text{ is een derdemacht.} \tag{4.9}$$

We willen laten zien dat de factoren $2p$, $p - 3q$ en $p + 3q$ paarsgewijs copriem zijn. Als p en q dezelfde pariteit zouden hebben, dan zouden $p + 3q$ en $p + q$ beide even zijn, en dus a en b ook, tegenspraak want zij zijn copriem. Dus p en q, en dus ook p en $3q$, hebben tegengestelde pariteit, zodat $p - 3q$ en $p + 3q$ beide oneven zijn. Elke gemene priemdeler d van $2p$ en $p \pm 3q$ is dus oneven, dus uit $d|2p$ volgt $d|p$. Dus d deelt $\pm 3q$, het verschil van $p \pm 3q$ en p. Als d niet 3 is, dan volgt $d|q$, dus d is gemene priemdeler van p en q, in tegenspraak met (4.8). De enige mogelijkheid is dus $d = 3$, maar uit $3|p$ volgt $3|a$ en we hadden juist het tegendeel aangenomen. We concluderen:

$$\text{ggd}(2p, p + 3q) = \text{ggd}(2p, p - 3q) = 1. \tag{4.10}$$

Laat nu d een gemene priemdeler zijn van $p - 3q$ en $p + 3q$. Dan deelt d hun som $2p$, en uit (4.10) volgt $d = 1$, dus ook

$$\text{ggd}(p + 3q, p - 3q) = 1.$$

We concluderen dat $2p$, $p - 3q$ en $p + 3q$ paarsgewijs copriem zijn, en omdat hun product een derdemacht is, volgt door twee keer toepassen van Lemma 2.2.1 dat het alledrie derdemachten zijn, laten we zeggen

$$p - 3q = X^3, \quad p + 3q = Y^3, \quad 2p = Z^3.$$

Uit deze formules volgt direct dat $X^3 + Y^3 = 2p = Z^3$, dus we hebben een 3-drietal te pakken indien X, Y, Z groter dan 0 zijn. Ze zijn in elk geval niet 0, want de derde macht van hun product is $2p(p - 3q)(p + 3q) = 2a$, dat is een natuurlijk getal en dus groter dan 0. Uit $X^3Y^3Z^3 > 0$ en uit het feit dat het teken van een geheel getal gelijk is aan het teken van zijn

derde macht (immers, $(-z)^3 = -(z^3)$) volgt bovendien dat een even aantal van X, Y, Z, dus nul of twee, negatief is. Zijn het er nul, dan zijn we klaar. Als het er twee zijn, dan zijn het niet X en Y, want dan zou ook $Z^3 = X^3 + Y^3$ negatief zijn, en dus ook Z. Dus als er twee negatief zijn, dan is Z negatief en één van beide X en Y is negatief, en door ze naar de andere kant te brengen krijgen we wél een 3-drietal. Bijvoorbeeld, als X en Z negatief zijn, dan volgt uit $X^3 + Y^3 = Z^3$ dat $(-Z)^3 + Y^3 = (-X)^3$, en $-Z, Y, -X$ zijn allen natuurlijke getallen. In het geval dat $Y < 0$ vinden we het 3-drietal $(X, -Z, -Y)$.

Nu rest ons nog te bewijzen dat het gevonden 3-drietal strikt kleiner is dan degene waar we mee begonnen; we moeten dus in de zojuist genoemde gevallen aantonen dat respectievelijk

$$Z < z; \quad -X < z; \quad -Y < z. \tag{4.11}$$

Het is niet nodig deze gevallen apart te behandelen. Omdat $2a > 0$ volgt namelijk

$$|XYZ|^3 = |X^3Y^3Z^3| = |2p(p - 3q)(2 + 3q)| = |2a| = 2a. \tag{4.12}$$

Als z even is (Geval 1 op pagina 13) volgt uit (4.2) dat $2a$ deler is van z^3, in het andere geval volgt uit (4.3) dat $2a$ deler is van $x^3 = z^3 - y^3$. In elk geval is dus $2a \leq z^3$, ofwel, $|XYZ|^3 \leq z^3$. Dit betekent dat

$$|X| \cdot |Y| \cdot |Z| \leq z. \tag{4.13}$$

Omdat geen van de X, Y, Z gelijk is aan 1 (want in dat geval kan onmogelijk voldaan zijn aan $X^3 + Y^3 = Z^3$, ga maar na) volgt dat $|X|, |Y|, |Z|$ stuk voor stuk strikt kleiner zijn dan z^3. Dus aan de beweringen in (4.11) is altijd voldaan, we hebben daarmee een strikt kleiner 3-drietal gevonden dan die waarmee we begonnen. Met de methode van oneindige afdaling leidt dit tot een tegenspraak: er zijn dus geen oplossingen waarbij a geen drievoud is. Het overgebleven geval $3|a$ is nu makkelijk: we hoeven slechts een paar dingen uit het vorige bewijs te veranderen.

Geval 2: a is deelbaar door 3 (met a, b zoals in (4.1)). Dan is er een $c \in \mathbb{N}$ zodat

$$a = 3c,$$

en dus

$$2a(a^2 + 3b^2) = 6c(9c^2 + 3b^2) = 18c(b^2 + 3c^2). \tag{4.14}$$

Terwijl in dit geval $2a$ en $a^2 + 3b^2$ niet copriem zijn omdat 3 hen beide deelt, zijn $18c$ en $b^2 + 3c^2$ dat wel. Als namelijk d een gemene priemdeler is, volgt uit $3|a$ en $\text{ggd}(a, b) = 1$ dat 3 niet b deelt, en dus ook niet $b^2 + 3c^2$. Dus $d \neq 3$, en omdat $d|18c$ en d priem is, volgt $d|2c$. Omdat $a = 3c$ en b van tegengestelde pariteit zijn, zijn c en b dat ook, dus $b^2 + 3c^2$ is oneven. De priemdeler d hiervan is dus niet 2, dus uit $d|2c$ volgt $d|c$ en dus ook $d|3c^2$. Dat betekent dat d ook b^2 deelt, het verschil van $b^2 + 3c^2$ en $3c^2$; en omdat d priem is, volgt $d|b$. Bovendien volgt uit $d|c$ dat $d|3c = a$, dus d is gemene priemdeler van a en b, tegenspraak. We concluderen dat $18c$ en $b^2 + 3c^2$ onderling priem zijn. Uit (4.14) en (4.1) volgt dat hun product een derde macht is, dus uit Lemma 2.2.1 volgt:

$$18c \text{ en } b^2 + 3c^2 \text{ zijn derde machten.} \tag{4.15}$$

Omdat $b^2 + 3c^2$ een derde macht is en b, c copriem en van tegengestelde pariteit zijn (want $a = 3c, b$ zijn dat ook), volgt uit (4.5) dat er gehele p, q bestaan zodat

$$b = p(p - 3q)(p + 3q), \qquad c = 3q(p - q)(p + q). \tag{4.16}$$

En $18c$, ofwel $18 \cdot 3q(p - q)(p + q) = 3^3 \cdot 2q(p - q)(p + q)$ is een derde macht, dus

$$2q(p - q)(p + q) \text{ is een derde macht.} \tag{4.17}$$

We willen laten zien dat deze drie factoren derde machten zijn, door aan te tonen dat ze relatief priem zijn. De getallen p en q zijn copriem, want volgens (4.16) deelt elke gemene deler ook b en c. Verder zijn p en q van tegengestelde pariteit, want anders zouden b en c beide even zijn. Dus $p \pm q$ zijn beide oneven, zodat elke deler d van $2q$ en $p \pm q$ oneven is en dus ook q deelt. Hieruit volgt dat d ook p deelt, en omdat p en q copriem zijn, volgt dat $d = \pm 1$. Dus $2q$ en $p \pm q$ zijn copriem. Elke priemdeler van de oneven getallen $p + q$ en $p - q$ deelt ook hun som $2p$ en hun verschil $2q$, en omdat zo'n priemdeler oneven is deelt hij ook p en q, die relatief priem zijn. Dus ook $p + q$ en $p - q$ zijn relatief priem. We concluderen dat de drie getallen $2q$, $p - q$ en $p + q$ paarsgewijs copriem zijn, zodat uit (4.17) na twee keer toepassen van Lemma 2.2.1 volgt dat $2q$, $p - q$ en $p + q$ alledrie derde machten zijn. Er zijn dus gehele X, Y, Z zodat

$$p - q = X^3, \quad 2q = Y^3, \quad p + q = Z^3, \tag{4.18}$$

en er volgt direct dat $X^3 + Y^3 = Z^3$. Op vrijwel precies dezelfde manier als in het vorige geval concluderen we dat we een strikt kleiner 3-drietal hebben gevonden dan (x, y, z). Hun product tot de macht 3 maal 9 is namelijk $18q(p - q)(p + q)$, en volgens (4.16) is dit gelijk aan $6c$, ofwel $2a$ en dat is groter dan nul. Dus opnieuw kunnen we door eventueel Z te verwisselen met X of Y een 3-drietal vinden van natuurlijke getallen. Omdat $9X^3Y^3Z^3 = 2a$, volgt dezelfde manier als in (4.12) volgt dat $9|X|^3|Y|^3|Z|^3 = 2a$ deler is van z^3 of x^3 en dus kleiner dan of gelijk aan z^3, en met exact hetzelfde argument als vanaf (4.13) volgt met oneindige afdaling een tegenspraak.

We concluderen dat we, als we het bestaan van een oplossing van $x^3 + y^3 = z^3$ aannemen, in welk geval dan ook op een tegenspraak stuiten, dus zo'n oplossing bestaat niet. Dit bewijst de Laatste stelling van Fermat voor het geval $n = 3$.

4.2 Complexe getallen en gemene delers

In het bewijs van het geval $n = 3$ gebruikten we de nog onbewezen bewering (4.5) van Euler, die we hier nogmaals als lemma formuleren.

Lemma 4.2.1. *Stel twee gehele getallen a en b zijn copriem en van tegengestelde pariteit,*[4] *en $a^2 + 3b^2$ is een derde macht. Dan is*

$$a = p^3 - 9pq^2, \qquad b = 3p^2q - 3q^3$$

voor zekere gehele p, q.

[4] Euler liet zelfs de woorden 'tegengestelde pariteit' weg, maar omdat we dat niet nodig hebben zullen we alleen deze zwakkere variant beschouwen.

Een variant van de omgekeerde bewering geldt ook:

Als we gehele getallen p en q nemen, en $a = p^3 - 9pq^2$, $b = 3p^2q - 3q^3$ stellen, dan is $a^2 + 3b^2$ een derde macht. (4.19)

Door 'domweg' haakjes uitwerken volgt namelijk

$$a^2 + 3b^2 = (p^3 - 9pq^2)^2 + 3(3p^2q - 3q^3)^2 = (p^2 + 3q^2)^3.$$

Hoewel het elementair is om deze formule te controleren, is het natuurlijk de vraag hoe je op zo'n formule komt. Daar komen we zodadelijk op terug.

Het bewijs van bovenstaand Lemma is niet eenvoudig. Euler publiceerde een 'bewijs', maar daar bleek een fundamentele fout in te zitten. Dat betekent niet dat het bewijs waardeloos is, want de concepten die worden gebruikt zijn op zichzelf al heel interessant, en worden ook gebruikt bij het uiteindelijke (correcte) bewijs van het lemma. Bovendien roept de methode vragen op over eenduidige priemontbinding in andere getalsystemen dan de ons zo vertrouwde natuurlijke getallen. Vragen als deze leidden tot de ontdekking van de algebraïsche getaltheorie, en die blijkt de sleutel bij het bewijzen van Fermat's laatste stelling voor meerdere n tegelijk.

Daarom presenteren we nu Euler's 'foute' ideeën. De kern ervan is dat we overgaan op een deelverzameling van de complexe getallen, vervolgens de derde macht $a^2 + 3b^2$ ontbinden als product van zulke getallen die 'relatief priem' zijn, en uit een analogon van Lemma 2.2.1 te concluderen dat elk van deze complexe factoren een derde macht is.

Beschouw de deelverzameling R van complexe getallen van de vorm

$$a + b\sqrt{-3} = a + b\sqrt{3}i$$

waarbij a en b *gehele* getallen zijn;[5] deze verzameling noteren we als $\mathbb{Z}[\sqrt{-3}]$. Deze notatie heeft een wiskundige betekenis (het is de ring van gehelen in het uitbreidingslichaam $\mathbb{Q}[\sqrt{-3}]$ van $\mathbb{Q}$), maar daar zullen we verder niet op ingaan.[6] Deze verzameling is *gesloten onder optelling en vermenigvuldiging* (dat wil zeggen, als $p, q \in R$ dan $p + q \in R$ en $pq \in R$), want

$$(a + b\sqrt{-3}) + (c + d\sqrt{-3}) = (a + c) + (b + d)\sqrt{-3},$$

en

$$\begin{aligned}(a + b\sqrt{-3})(c + d\sqrt{-3}) &= ac + ad\sqrt{-3} + bc\sqrt{-3} + bd(\sqrt{-3})^2 \\ &= (ac - 3bd) + (ad + bc)\sqrt{-3}.\end{aligned}$$

Euler merkte op dat $a^2 + 3b^2$ ontbonden kan worden als product van elementen van R:

$$a^2 + 3b^2 = (a + b\sqrt{-3})(a - b\sqrt{-3}) = \alpha\overline{\alpha} = |\alpha|^2,$$

waarbij α het complexe getal $a + b\sqrt{-3}$ is en $\overline{\alpha}$ zijn geconjugeerde. Deze ontbinding maakt de afleiding van de 'omgekeerde' bewering (4.19) eenvoudiger, of het maakt in elk geval duidelijk

[5] Eigenlijk heeft -3 twee complexe wortels die met evenveel recht de notatie $\sqrt{-3}$ verdienen, namelijk $\sqrt{3}i$ en $-\sqrt{3}i$. Met $\sqrt{-3}$ zullen wij echter altijd $\sqrt{3}i$ bedoelen.

[6] We schrijven $\mathbb{Q}$ voor de verzameling rationale getallen, dat zijn de 'breuken' a/b met $a, b \in \mathbb{Z}, b \neq 0$.

hoe men op het idee komt. De voorwaarde $a = p^3 - 9pq^2$, $b = 3p^2q - 3q^3$ is namelijk equivalent aan

$$a + b\sqrt{-3} = (p^3 - 9pq^2) + (3p^2q - 3q^3)\sqrt{-3},$$

ofwel,

$$a + b\sqrt{-3} = (p + q\sqrt{-3})^3.$$

Als aan deze voorwaarde voldaan is, die we kunnen schrijven als $\alpha = \rho^3$ met $\rho = p + q\sqrt{-3}$, dan volgt uit de multiplicatieve eigenschap $\overline{\gamma}\overline{\delta} = \overline{\gamma\delta}$ van conjugatie dat

$$a^2 + 3b^2 = \alpha\overline{\alpha} = \rho^3\overline{(\rho^3)} = \rho^3(\overline{\rho})^3 = (\rho\overline{\rho})^3 = (p^2 + 3q^2)^3,$$

dus $a^2 + 3b^2$ is een derde macht.

Euler lijkt in zijn 'bewijs' van Lemma 4.2.1 in essentie als volgt te redeneren.[7] Stel a en b zijn copriem, en $a^2 + 3b^2$ is een derde macht. Als een willekeurig product pq van getallen p, q een derde macht is, dan zijn er getallen m, s, t zodat $p = ms^3$ en $q = mt^3$. Uit $a^2 + 3b^2 = (a + b\sqrt{-3})(a - b\sqrt{-3}) = \alpha\overline{\alpha}$ is een derde macht volgt dus dat $\alpha = ms^3$, $\overline{\alpha} = mt^3$ voor getallen $m, s, t \in R$. Maar $a + b\sqrt{-3}$ en $a - b\sqrt{-3}$ zijn relatief priem in de zin dat a en b dat zijn, dus $m = 1$. Er volgt dat α en $\overline{\alpha}$ derde machten zijn, zeg $\alpha = \rho^3$ en bijgevolg $\overline{\alpha} = \overline{\rho^3} = (\overline{\rho})^3$ voor een $\rho = p + q\sqrt{-3} \in R$. Dus

$$a + b\sqrt{-3} = (p + q\sqrt{-3})^3 = (p^3 - 9pq^2) + (3p^2q - 3q^3)\sqrt{-3},$$

dus a en b zijn van de gewenste vorm.

Euler's redenering is vaag op een aantal plekken. Ten eerste is niet duidelijk wat hij bedoelt met '$a + b\sqrt{-3}$ en $a - b\sqrt{-3}$ zijn relatief priem in de zin dat a en b dat zijn', waaruit 'volgt' dat $m = 1$. Ten tweede, en dat is misschien wel het belangrijkste, de bewering dat $p = ms^3$, $q = mt^3$ als pq een derde macht is geldt voor *gehele* getallen p en q, maar dat zegt nog niets over getallen van de vorm $a + b\sqrt{-3}$. Voor gehele p, q volgt het namelijk uit Lemma 2.2.1, maar die berust op de unieke priemfactorisatie van gehele getallen. Het is helemaal niet duidelijk welke getallen van de vorm $a + b\sqrt{-3}$ we als 'priemgetallen' kunnen beschouwen, laat staan dat er sprake is van unieke priemontbinding.

Met dezelfde redenering beweert Euler algemener dat als c geheel is, a, b copriem en $a^2 + cb^2$ een kwadraat, dan zijn diens factoren $a + b\sqrt{-c} =: \alpha$ en $a - b\sqrt{-c} = \overline{\alpha}$ kwadraten, zeg $\alpha = (p + q\sqrt{-c})^2$. Dus

$$a + b\sqrt{-c} = (p + q\sqrt{-c})^2 = (p^2 - cq^2) + 2pq\sqrt{-c},$$

zodat $a = p^2 - cq^2$ en $b = 2pq$. Behalve dat het bewijs onvolledig is, is de conclusie ook onjuist. Neem bijvoorbeeld $a = 2$, $b = 3$ en $c = 5$. Dan zijn a en b copriem en $a^2 + cb^2 = 2^2 + 5 \cdot 3^2 = 49$ is een kwadraat, dus uit de methode zou volgen dat er gehele getallen p, q zijn zodat $2 = p^2 - 5q^2$, $3 = 2pq$. Dit betekent dat de hyperbolen $2 = x^2 - 5y^2$ en $3 = 2xy$ elkaar snijden in een 'roosterpunt' (een punt in $\mathbb{Z}^2$), maar het is niet moeilijk om na te gaan dat de enige twee snijpunten geen roosterpunten zijn.

[7] Bron: zie [4].

We moeten dus oppassen met begrippen als 'priemgetal' en 'relatief priem' in andere getalstelsels. Toch is men er in de loop van de geschiedenis in geslaagd om deze begrippen te definiëren voor bepaalde andere (abstracte) getalstelsels, en er nuttige stellingen over te bewijzen. Hoewel deze theorie op zich niet makkelijk is, worden veel andere bewijzen hierdoor wél veel makkelijker en bovendien inzichtelijker; en stellingen die vroeger niet te temmen waren kunnen ineens worden aangepakt. Het correcte bewijs dat Euler uiteindelijk gaf voor Lemma 4.2.1, waarmee $n = 3$ eindelijk bewezen was, is vrij omslachtig en niet zo transparant. Daarom loont het voor ons de moeite om een duik te nemen in de moderne algebra, waarna het bewijs van het lemma eenvoudiger wordt. Bovendien kunnen we met de zo verkregen inzichten de Laatste stelling van Fermat bewijzen voor nog meer exponenten n. Veel van de begrippen die we zullen tegenkomen, waren in Euler's tijd nog niet bekend. Het begrip 'groep' is bijvoorbeeld ontdekt door de veel te jong gestorven Franse wiskundige Évariste Galois (1811 – 1832).

Het blijkt dat $\mathbb{Z}$ niet het enige getalstelsel is waar we eenduidige priemontbinding hebben. Een voorbeeld is het systeem van *gehele getallen van Gauss*, dat zijn de complexe getallen van de vorm $a+b\sqrt{-1}$ met a, b geheel. Maar voor bijvoorbeeld de verzameling $\mathbb{Z}[\sqrt{-5}]$ van getallen van de vorm $a + b\sqrt{-5}$ met a, b geheel geldt dat niet. We zullen echter zien dat we, behalve met de getallen zelf, ook met bepaalde *deelverzamelingen* van getallen kunnen rekenen, de *idealen*. Deze gedragen zich in veel opzichten prettiger dan de getallen zelf. Zo is er in $\mathbb{Z}[\sqrt{-5}]$ bijvoorbeeld wél eenduidige priemontbinding in 'priemidealen'. De Duitse wiskundige Ernst Kummer (1810 – 1893), de grondlegger van het begrip ideaal, slaagde erin in één keer de Laatste stelling van Fermat bewijzen voor een waarschijnlijk flink groot deel van de priemgetallen, in elk geval voor bijna alle priemgetallen onder de honderd. De studie van idealen is onderdeel van de moderne algebra, waar we in de volgende hoofdstukken een kijkje gaan nemen. De Laatste stelling van Fermat zal hierbij een beetje op de achtergrond raken, maar vanaf hoofdstuk 7 zullen die stelling en de behandelde theorie elkaar ontmoeten.

Hoofdstuk 5

Groepen, ringen en lichamen

De kracht van de moderne algebra is dat het, juist door de grote abstractie, heel breed toepasbaar is. Je hoeft maar één algemene stelling te bewijzen voor een bepaalde 'algebraïsche structuur', zoals een ring, en de stelling geldt meteen voor alle voorbeelden die deze structuur hebben, zoals $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ en $\mathbb{C}$.[1] Bovendien worden bewijzen vaak transparanter als je zoveel mogelijk 'overbodige informatie' weglaat. Dat wil niet zeggen dat alles zo abstract mogelijk moet worden gemaakt: zonder concrete voorbeelden wordt de theorie niet alleen gortdroog, maar ook waardeloos. De observaties aan concrete voorbeelden kunnen met een beetje geluk en veel oefening worden opgelift naar 'hogere', abstracte inzichten, die op hun beurt, na een soms lange, horizontale reis over anders onoverbrugbare landschappen weer afdalen naar andere concrete voorbeelden, die vaak in een andere wereld lijken te leven dan die van het oorspronkelijke voorbeeld.

Bij het opstellen van een wiskundige definitie voor een bepaalde structuur is één van de problemen vaak hoe 'algemeen' die moet zijn. Hoe breder de definitie, hoe meer gevallen eraan voldoen, en hoe groter je 'winst' als je een algemene waarheid over die structuur ontdekt. Aan de andere kant, hoe breder de definitie, hoe minder er te bewijzen valt (er is immers minder informatie die je kunt gebruiken), en je wilt dat bepaalde stellingen die voor de 'standaardvoorbeelden' gelden ook voor de algemene structuur geldt.

De definities van de begrippen *groep*, *ring* en *lichaam* zijn wat dat betreft heel goed gelukt. Zo duikt het begrip 'groep' bijna overal in de wiskunde op, en toch zijn er enorm veel diepe resultaten voor te bewijzen. Bovendien is het 'algemeenheids-dilemma' opgelost door een hiërarchie van specialisatie aan te brengen: een ring is een groep met een extra structuur, een lichaam is een ring met een speciale eigenschap.

De essentie van bovenstaande begrippen is om een bewerking op een bepaalde verzameling te definieëren die aan bepaalde eigenschappen voldoet; bijvoorbeeld $\mathbb{Z}$ heeft de bewerkingen 'optellen' en 'vermenigvuldigen'. Onder een bewerking op een verzameling G verstaan we een functie

$$f : G \times G \to G.$$

Als f bijvoorbeeld voor optelling in $\mathbb{Z}$ staat, dan is $f(5, 7) = 12$. Dit schrijven we natuurlijk als $5 + 7 = 12$, en algemeen neemt men vaak een symbool voor de bewerking, bijvoorbeeld

[1] Met $\mathbb{R}$ en $\mathbb{C}$ bedoelen we de reële en de complexe getallen.

$\circ$, en schrijft $f(a,b)$ als $a \circ b$. Meestal laat men het symbool zelfs weg en schrijft ab. Als er twee bewerkingen op G gedefinieerd zijn, zullen wij deze altijd 'optelling' en 'vermenigvuldiging' noemen en met $+$ en $\cdot$ aanduiden, maar in principe hoeven ze niks met de bekende bewerkingen op getallen te maken te hebben.

Een aantal mogelijke eigenschappen van een bewerking op G zijn:

1). *associativiteit:* $a(bc) = (ab)c$ voor alle $a, b, c \in G$.

2). *eenheidselement:* Er is een $e \in G$ zodat $ae = ea = a$ voor alle $a \in G$.

3). *inverse:* Voor alle $a \in G$ is er een $a^{-1} \in G$ zodat $aa^{-1} = a^{-1}a = e$.

4). *commutativiteit:* $ab = ba$ voor alle $a, b \in G$.

5). *distributiviteit:* $a(b+c) = ab + ac$ en $(a+b)c = ac + bc$ voor alle $a, b, c \in G$.

Inversen kunnen natuurlijk alleen bestaan als er ook een eenheidselement is. Verder kan er alleen sprake zijn van distributiviteit als G twee bewerkingen 'plus' en 'keer' heeft. Distributiviteit is een bijzonder geval omdat het een verband legt tussen de twee bewerkingen. In het geval dat er twee bewerkingen $+$ en $\cdot$ zijn, kunnen de overige eigenschappen voor beide bewerkingen gelden. Bijvoorbeeld commutativiteit voor optelling zegt dat $a+b = b+a$, en voor vermenigvuldiging dat $a \cdot b = b \cdot a$. In plaats van eenheidselement spreken we van 'nulelement' of gewoon 'nul' als het om optelling gaat, en schrijven 0 in plaats van e. Dus 2) luidt in dit geval $a+0 = 0+a = a$. Het eenheidselement (of kortweg één) voor vermenigvuldiging schrijven we als 1, en we lezen $a \cdot 1 = 1 \cdot a = a$. Het zal nu niet als een verassing komen dat we bij optelling spreken van 'tegengestelde' in plaats van inverse, en $-a$ schrijven in plaats van a^{-1}. Dus 3) luidt voor plus en keer dat $a + (-a) = (-a) + a = 0$, en $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

Hoewel niet expliciet genoemd als eigenschap, is het altijd belangrijk om te controleren dat een bewerking *inwendig* is, dat wil zeggen dat $ab \in G$ als $a, b \in G$. Anders hebben we immers niet met een functie van $G \times G$ naar G te maken.

5.1 Groepen

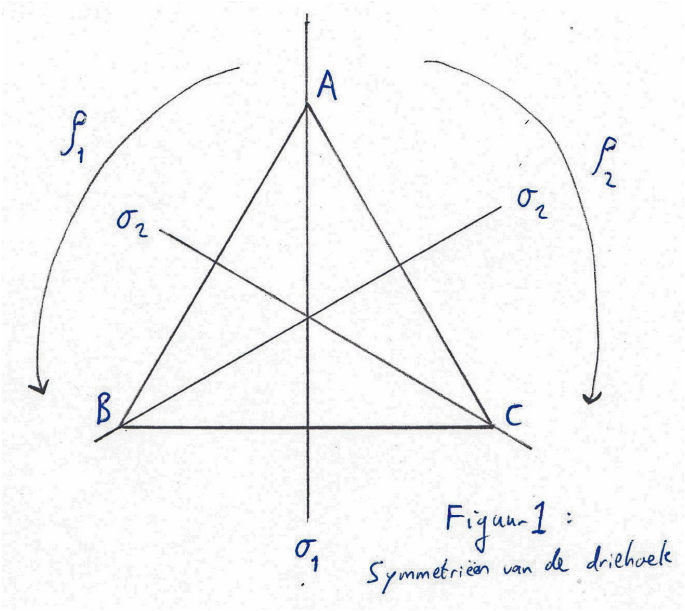
Definitie 5.1.1. Groep. *Een* groep *is een verzameling G voorzien van een bewerking die voldoet aan bovenstaande eigenschappen 1), 2) en 3): er is associativiteit, een eenheidselement en voor elk element een inverse.*

Formeel is een groep een tweetal $(G, \circ)$ met G een verzameling en $\circ$ een bewerking, maar voor het gemak spreken we gewoon van de groep G. Willekeurige groepen schrijven we meestal 'multiplicatief': de bewerking noemen we vermenigvuldiging, en we schrijven $a \circ b$ als ab en noemen dit het product van a en b. In veel voorbeelden schrijven we de groep echter 'additief': de bewerking heet optellen, en we schrijven $a \circ b$ als $a + b$ en noemen dit de som van a en b. Dit is echter alleen een kwestie van notatie. De additieve notatie gebruikt men eigenlijk alleen als de groep commutatief is, dat wil zeggen als $a + b = b + a$ voor alle a, b. In de rest van deze paragraaf bedoelen we met G altijd een groep.

5.1.1 Een voorbeeld: symmetrieën van een driehoek

Voorbeelden van groepen kom je overal tegen, maar het meest tot de verbeelding sprekend zijn de zogenaamde symmetriegroepen. Een symmetrie van een figuur, bijvoorbeeld de gelijkzijdige driehoek van Figuur 1, is een afstand-bewarende bijectieve transformatie van de ruimte (in dit geval van het platte vlak) die de figuur in zichzelf overvoert.[2] Voorbeelden van symmetrieën van de driehoek zijn de rotaties ρ_n over n derde-slagen, en de spiegelingen $\sigma_1, \sigma_2, \sigma_3$ door de aangegeven assen.[3]

De verzamelingen symmetrieën van de driehoek noemen we G, en als bewerking $\circ$ nemen we de samenstelling van functies. Dus als f en g functies zijn, dan betekent $f \circ g$ 'eerst g toepassen, dan f". Deze bewerking is inwendig, want als je twee keer achter elkaar een figuur in zichzelf overvoert, voer je het nog steeds in zichzelf over. Zo is bijvoorbeeld $\rho_4 \circ \rho_1 = \rho_5$: eerst een derde slag draaien en dan vierderde slagen draaien, komt (als functie) op hetzelfde neer als in één keer vijf derde slagen draaien. Bovendien is ρ_5 weer gelijk aan ρ_2, zoals je kunt nagaan. Het eenheidselement is de identieke functie $\mathrm{id} : x \mapsto x$, die duidelijk een symmetrie is, en de inverse van een symmetrie f is gewoon de inverse functie f^{-1} in de zin van de verzamelingenleer, die ook een symmetrie is.[4] Associativiteit voor samenstellen van functies geldt altijd, dus ook voor symmetrieën. Dus G is een groep.



Figuur 1: Symmetrieën van de driehoek

Een symmetrie van de driehoek ligt vast als we de werking op de hoekpunten kennen. We kunnen de hoekpunten bijvoorbeeld labelen met A, B en C zoals in de figuur. Omdat σ_1 de punten C en B verwisselt en A vasthoudt, en omdat ρ_1 het punt A naar B, B naar C en C naar A stuurt, kunnen we $\rho_1 \circ \sigma_1$ omschrijven als

$$A \mapsto A \mapsto B, \qquad B \mapsto C \mapsto A, \qquad C \mapsto B \mapsto C,$$

ofwel A gaat naar B, B naar A, en C wordt vastgehouden. Maar dat is precies wat σ_3 doet, dus

$$\rho_1 \circ \sigma_1 = \sigma_3,$$

[2]Een functie $f : X \to Y$ heeft bijectief als hij surjectief en injectief is. Surjectief betekent dat f heel Y bereikt: voor elke $y \in Y$ is er een $x \in X$ zodat $f(x) = y$. Injectief betekent dat de functie 1-op-1 is: als $x_1 \neq x_2$, dan is $f(x_1) \neq f(x_2)$.

[3]Correctie op de tekening: de linker σ_2 moet σ_3 zijn.

[4]Immers, omdat f afstanden bewaart, volgt uit $\|f^{-1}(x) - f^{-1}(y)\| = \|f(f^{-1}(x)) - f(f^{-1}(y))\| = \|x - y\|$ dat ook f^{-1} afstanden bewaart. Bovendien induceert f een bijectie van de figuur naar zichzelf, dus f^{-1} ook.

want een symmetrie ligt vast door de werking op de hoekpunten. Op dezelfde manier zien we dat $\sigma_1 \circ \rho_1 = \sigma_2$, dus de bewerking is *niet commutatief*. Omdat er maar 3! verschillende 'herrangschikkingen' zijn van A, B, C, is G eindig en heeft hoogstens 6 elementen. Het zijn er precies 6 omdat de σ_i, ρ_j allen verschillen voor $i, j = 1, 2, 3$. (Er geldt $\rho_3 = \mathrm{id}$.) De inverse van een rotatie ρ_k is ρ_{3-k}, en een spiegeling σ_k is zijn eigen inverse.

5.1.2 Ordes

In een groep G kunnen we herhaald een element x met zichzelf vermenigvuldigen. Voor $n \geq 1$ noemen we

$$x^n := x \cdot x \cdot \ldots \cdot x$$

het n-voudige product van x. Het is duidelijk dat $x^{m+n} = x^m x^n$ voor alle $m, n \geq 1$. Als we bovendien voor $n > 1$ definiëren

$$x^{-n} = x^{-1} \cdot x^{-1} \cdot \ldots \cdot x^{-1} \qquad \text{en} \qquad x^0 = 1,$$

dan geldt zelfs dat $x^{m+n} = x^m x^n$ voor alle $m, n \in \mathbb{Z}$. Het bewijs hiervan bestaat simpelweg uit het controleren van de formule voor alle mogelijkheden met m, n danwel positief, negatief of nul. Uit de formule volgt meteen voor alle $m \in \mathbb{Z}$, $n \in \mathbb{N}$ dat $x^{mn} = (x^m)^n$, want $x^{mn} = x^{m+\ldots+m} = x^m \cdots x^m = (x^m)^n$. Bovendien is $x^{-k} = (x^k)^{-1}$, want $x^{-k}x^k = x^{-k+k} = x^0 = e$, en analoog $x^k x^{-k} = e$. Door hier gebruik van te maken zien we dat de formule $x^{mn} = (x^m)^n$ zelfs geldt voor alle $n \in \mathbb{Z}$. We concluderen:

$$x^{m+n} = x^m x^n \qquad \text{en} \qquad x^{mn} = (x^m)^n \qquad \text{voor alle } m, n \in \mathbb{Z}.$$

Als we G additief schrijven, dan hebben we het over de *n-voudige som*, en schrijven deze als $nx = x + x + \ldots + x$. voor $n \geq 1$. Analoog schrijven we $(-n)x = (-x) + (-x) + \ldots + (-x)$ voor $n > 1$ en $0x = 0$, en we hebben de rekenregels

$$(m + n)x = mx + nx \qquad \text{en} \qquad (mn)x = m(nx) \qquad \text{voor alle } m, n \in \mathbb{Z}. \tag{5.1}$$

Dit is niets nieuws, het is gewoon precies dezelfde formule als in de multiplicatieve notatie, maar nu additief genoteerd. Wel kan de notatie nx verwarrend zijn omdat het lijkt alsof we n met x vermenigvuldigen. Zo betekenen de formules in (5.1) niet hetzelfde als de distributiviteit en associativiteit van de groepsbewerking. Als er verwarring kan ontstaan, proberen we duidelijk te zijn in wat we bedoelen.

De *orde* van een groep G, notatie $|G|$, is de kardinaliteit van de verzameling G; voor eindige G is dit gewoon het aantal elementen. De orde van een element $x \in G$ is de kleinste exponent $n > 0$ waarvoor $x^n = e$. Als zo'n n niet bestaat, noemen we de orde van x oneindig. Dus bijvoorbeeld de orde van het eenheidselement e is 1, en in het bovenstaande voorbeeld van de driehoek hebben de rotaties ρ_1 en ρ_2 orde 3, en de spiegelingen $\sigma_1, \sigma_2, \sigma_3$ orde 2. De elementen in de optelgroep $\mathbb{Z}$, behalve 0 die orde 1 heeft, hebben allemaal oneindige orde. Als x eindige orde n heeft, dan geldt voor alle $q, r \in \mathbb{Z}$ dat

$$x^{qn+r} = x^{qn}x^r = (x^n)^q x^r = e^q x^r = x^r.$$

We kunnen de exponenten van x dus modulo n lezen. Bovendien zijn de n elementen x^0, x^1, x^2, $\ldots$, x^{n-1} allemaal verschillend. Stel namelijk dat $x^l = x^k$ voor $k, l \in M = \{0, 1, 2, \ldots, n-1\}$ met $l > k$, dan is $x^{l-k} = x^l x^{-k} = x^l (x^k)^{-1} = e$. Omdat $l - k \in M$ volgt uit de minimaliteit van n dat $l - k = 0$, dus $l = k$. Nu hebben we bijna bewezen:

Lemma 5.1.2. *Zij x een element van een groep G. Als x eindige orde n heeft, dan geldt*

$$x^k = x^l \quad \Longleftrightarrow \quad k \equiv l \pmod{n}.$$

In het bijzonder bevat $\{x^k : k \in \mathbb{Z}\}$ precies n verschillende elementen. Als x oneindige orde heeft, dan zijn alle oneindig veel elementen x^k met $k \in \mathbb{Z}$ verschillend. In het bijzonder heeft elke eindige groep alleen elementen van eindige orde.

Bewijs. Stel eerst x heeft eindige orde n. Als $k \equiv l \pmod{n}$, dan $k = qn + l$ voor een zekere $q \in \mathbb{Z}$, zodat $x^k = x^{qn+l} = x^l$. Stel omgekeerd dat $x^k = x^l$. Schrijf $k = q_1 n + r_1$, $l = q_2 n + r_2$ met $0 \leq r_1, r_2 < n$. We hebben dus $x^k = x^{r_1}$ en $x^l = x^{r_2}$, en dus $x^{r_1} = x^{r_2}$. Omdat $r_1, r_2 \in M$ met M als hiervoor, volgt dat $r_1 = r_2$, en dus $k \equiv l \pmod{n}$.

Stel nu dat x oneindige orde heeft. Uit $x^l = x^k$ volgt dat $x^{l-k} = x^l x^{-k} = e$, dus $l - k = 0$ zodat $l = k$. De elementen $\ldots, x^{-2}, x^{-1}, x^0, x^1, x^2, \ldots$ zijn dus allemaal verschillend. □

Een groep G heet *cyclisch* als er een $x \in G$ is zodat

$$G = \{x^k : k \in \mathbb{Z}\}$$

(in de additieve notatie moeten we kx in plaats van x^k schrijven); G bestaat dus geheel uit de machten van een element x, en we zeggen dat x de groep voortbrengt. Bijvoorbeeld de optelgroep $\mathbb{Z}$ is een oneindige cyclische groep, we kunnen als voortbrenger 1 of -1 kiezen, want elk geheel getal is te schrijven als $k \cdot 1$ voor een gehele k.[5] Voor elk natuurlijk getal n vormen de getallen modulo n een eindige cyclische groep: we kunnen als voortbrenger bijvoorbeeld 1 mod n nemen, deze heeft orde n want $n \cdot 1 \equiv 0 \pmod{n}$.

5.1.3 Ondergroepen en delers van de groepsorde

Een deelverzameling H van een groep G die e bevat, gesloten is onder de bewerking van G en 'onder het nemen van inversen', noemen we een *ondergroep* van G. Om te laten zien dat een deelverzameling H een ondergroep van G is, moeten we dus nagaan dat:

1. $e \in H$;
2. Voor alle $a, b \in H$ is $ab \in H$;
3. Voor alle $a \in H$ is $a^{-1} \in H$.

Elke ondergroep H van G is een groep, want het 'erft' de associativiteit van G. We hebben dus een groep in een groep, en bestudering van bepaalde ondergroepen (die in zekere zin eenvoudiger zijn, want kleiner) levert vaak belangrijke inzichten over de groep zelf. Als de

[5] Hier betekent $k \cdot 1$ de k-voudige som $1 + 1 + \ldots + 1$ en niet k maal 1, maar omdat dat precies hetzelfde is zorgt dat niet voor verwarring.

bewerking van een groep commutatief is, noemen we het een *Abelse groep*, naar de Noor Niels Hendrik Abel (1802 – 1829). Is G Abels, dan is H dat ook: als $ab = ba$ voor alle $a, b \in G$, dan zeker voor alle $a, b \in H$. Elke groep G heeft de triviale ondergroepen $\{e\}$ en G. Een interessanter voorbeeld is de verzameling even getallen van de optelgroep van $\mathbb{Z}$, of in het algemeen de verzameling van n-vouden.

Een natuurlijke vraag is nu: hoe bepalen we wat de ondergroepen van een bepaalde groep zijn? Het is behalve voor heel kleine groepen veel te veel werk om alle combinaties van elementen te controleren! Voor eindige groepen is er gelukkig een handig criterium: de orde van een ondergroep is altijd een deler van de groepsorde. In het bijzonder betekent dit dat als de orde van G een priemgetal is, dan zijn G en $\{e\}$ de enige ondergroepen van G.

Voor het bewijs hiervan beschouwen we de verzamelingen van de vorm

$$aH := \{ah : h \in H\}$$

waarbij a een willekeurig maar vast element is van G. We hebben de groepsbewerking hier multiplicatief genoteerd; als de bewerking $+$ is schrijven we $a + H = \{a + h : h \in H\}$. Als we bijvoorbeeld voor H de ondergroep $\{\ldots, -16, -8, 0, 8, 16, \ldots\} \subset \mathbb{Z}$ van achtvouden nemen, dan is $3 + H = \{\ldots, -13, -5, 3, 11, 19, \ldots\}$. Dit is precies de verzameling getallen met rest 3 bij deling door 8.

Stelling 5.1.3. *Zij G een eindige groep met ondergroep H. Dan is $|H|$ deler van $|G|$.*

Bewijs. We definiëren een relatie $\sim$ op G door

$$a \sim b \iff a^{-1}b \in H.$$

We tonen eerst aan dat dit een equivalentierelatie[6] is, met als equivalentieklassen de verzamelingen aH. Omdat H een groep is, is $a^{-1}a = e \in H$, dus $a \sim a$. Als $a \sim b$, dan $a^{-1}b \in H$ en dus ook $b^{-1}a = (a^{-1}b)^{-1} \in H$, dus $b \sim a$. Als $a \sim b$ en $b \sim c$, dan $a^{-1}b \in H$ en $b^{-1}c \in H$, dus ook $a^{-1}c = (a^{-1}b)(b^{-1}c) \in H$. Dus $\sim$ is een equivalentierelatie. De equivalentieklasse van een $a \in G$ bestaat uit alle $b \in G$ waarvoor er een $h \in H$ is zodat $a^{-1}b = h$ ofwel $b = ah$. Dat zijn precies de elementen $b \in aH$.

De verzameling $L = \{aH : a \in G\}$ van equivalentieklassen vormt, zoals we weten, een *partitie* van G. Voor elk van de equivalentieklassen (laten we zeggen dat het er m zijn) kunnen we een a_i kiezen, zodat $L = \{a_1H, a_2H, \ldots, a_mH\}$. De orde van G is het aantal elementen per equivalentieklasse, gesommeerd over al deze klassen:

$$|G| = \sum_{k=1}^{m} |a_kH|. \tag{5.2}$$

Voor elke $a \in G$ is de functie $f_a : G \to aG : g \mapsto ag$ een bijectie (de inverse is $f_{a^{-1}}$), dus f_a beeldt H bijectief naar aH af. Dus $|a_kH| = |H|$ voor alle k, en uit (5.2) volgt dat $|G| = \sum_{k=1}^{m} |H| = m|H|$. Dus $|H|$ is deler van $|G|$. □

[6]Een relatie $\sim$ op een verzameling X is een equivalentierelatie als $a \sim a$ voor alle a (reflectiviteit), $a \sim b$ impliceert $b \sim a$ (symmetrie), en $a \sim b$, $b \sim c$ impliceert $a \sim b$ (transitiviteit). Is hieraan voldaan, dan wordt X opgedeeld in deelverzamelingen van onderling equivalente elementen: de deelverzameling van elementen equivalent met a wordt de equivalentieklasse van a genoemd, en de equivalentieklassen vormen een partitie (opdeling) van X.

Voor elke $x \in G$ is de verzameling $H = \{x^k : k \in \mathbb{Z}\}$ een ondergroep van G. Van Lemma 5.1.2 weten we dat de orde van H precies de orde van x is. Omdat de orde van H die van G deelt, volgt:

Gevolg 5.1.4. *De orde van elk element van G deelt de orde van G.*

Als de orde van G een *priemgetal* p is, dan bevat G meer dan 1 element, dus ook een x ongelijk aan het eenheidselement e. De orde van x deelt p en is dus 1 of p, maar omdat $x^1 = x \neq e$ kan de orde niet 1 zijn. Dus x heeft orde p, en omdat G maar p elementen bevat concluderen we:

$$G = \{x^0, x^1, x^2, \ldots, x^{p-1}\}.$$

We hebben nu bewezen:

Gevolg 5.1.5. *Elke groep met als orde een priemgetal is cyclisch.*

5.2 Ringen en lichamen

Veel interessante eigenschappen van getallen hebben te maken met een verband tussen optelling en vermenigvuldiging. Ook in abstracte structuren is het vaak de moeite waard om twee in plaats van één bewerking te bestuderen, en vooral hoe die met elkaar combineren. Daarom heeft men het begrip 'ring' ingevoerd.

Definitie 5.2.1. Ring. *Een* ring *is een verzameling R voorzien van twee bewerkingen $+$ (plus) en $\cdot$ (keer), zodat voldaan is aan:*

- *R is een optelgroep; daarmee bedoelen we dat $(R, +)$ een groep is.*
- *De vermenigvuldiging is associatief, en er is een eenheidselement* 1*.*
- *De distributieve wet geldt.*
- $1 \neq 0$.

Het standaardvoorbeeld van een ring is $\mathbb{Z}$, maar er zijn er vele anderen. Neem bijvoorbeeld een vast geheel getal c: de verzameling getallen $a + b\sqrt{-c}$ met a, b geheel vormt een ring, zoals we in essentie al zijn nagegaan aan het eind van het vorige hoofdstuk. Een ander belangrijk voorbeeld is de veeltermring $\mathbb{Z}[X]$ over $\mathbb{Z}$, die bestaat uit de polynomen

$$k_1 + k_2X + k_2X^2 + \ldots + k_nX^n$$

waarbij de k_i gehele getallen zijn, en n een niet-negatief geheel getal. Hier is X een formele variabele die door elk ander symbool kan worden vervangen, de k_i heten de coëffieciënten. Bij het optellen en vermenigvuldigen van polynomen doen we alsof we met X mogen rekenen 'alsof het een getal is', en gebruiken bijvoorbeeld de distributieve wet en de regel $k_mX^m \cdot k_nX^n = k_mk_nX^{m+n}$. Op deze manier is vrij snel duidelijk dat $\mathbb{Z}[X]$ een ring is. Algemeen kunnen we voor een ring R de veeltermring $R[X]$ definiëren met coëffieciënten in R.

Een ander belangrijk voorbeeld zijn de matrixringen. Bijvoorbeeld $\mathrm{Mat}_n(\mathbb{C})$ is de verzameling $n \times n$ matrices met coëfficiënten in $\mathbb{C}$. Zoals we weten kunnen we matrices componentsgewijs optellen en via de 'rij-maal-kolom regel' vermenigvuldigen, en deze bewerkingen voldoen aan de ringeigenschappen. Het nulelement is de nulmatrix, het eenheidselement is I_n, en de tegengestelde van A is $(-1) \cdot A$. Matrixvermenigvudliging is (behalve voor $n = 1$) niet-commutatief, zoals je kunt nagaan.

5.2.1 Enkele basiseigenschappen

Een paar belangrijke eigenschappen van groepen en ringen zijn direct uit de definities af te leiden. We hebben het bijvoorbeeld steeds over 'de' inverse en 'het' eenheidselement, waarmee we suggereren dat ze uniek zijn. Dat is inderdaad het geval. Laat G een groep zijn, en stel dat e, e' beide eenheidselementen zijn. Dan volgt direct dat

$$e' = ee' = e.$$

Stel p en q zijn inversen van a, dan volgt dat

$$p = pe = p(aq) = (pa)q = eq = q.$$

Op dezelfde manier volgt dat in een ring, behalve de nul en de tegengestelden, ook de één en de inversen (indien ze bestaan) uniek zijn.

Een ring waarvoor vermenigvuldiging commutatief is, noemen we een *commutatieve ring*. We zouden ook nog een aparte naam kunnen verzinnen voor ringen waarbij de optelling commutatief is, maar dat is niet nodig. *De optelgroep van een ring R is namelijk altijd Abels.* Er geldt namelijk voor alle $x, y \in R$ dat

$$\begin{aligned} (x+y)(1+1) &= x(1+1) + y(1+1) = x + x + y + y, \qquad \text{en ook} \\ (x+y)(1+1) &= (x+y)1 + (x+y)1 = x + y + x + y, \end{aligned}$$

dus $x + x + y + y = x + y + x + y$. Aan beide kanten van de vergelijking links $-x$ en rechts $-y$ optellen, levert $x + y = y + x$.

Een andere belangrijk feit is dat $0 \cdot a = a \cdot 0 = 0$ voor alle a in een ring R. Er geldt namelijk

$$0a = (0 + 0)a = 0a + 0a,$$

dus door aan beide kanten $-(0a)$ op te tellen volgt $0a = 0$. Analoog volgt $a0 = 0$.

Wij hebben de ietwat vreemd ogende voorwaarde voor een ring gesteld dat $1 \neq 0$, maar veel auteurs doen dat niet. We verliezen echter niet veel door het wel te stellen. Er is namelijk maar één ring mogelijk waarvoor $0 = 1$, de *nulring* $\{0\}$. We kunnen $\{0\}$ voorzien van een optelling $0 + 0 = 0$ en een vermenigvuldiging $0 \cdot 0 = 0$, waarmee het een ring zou worden. Dit is een erg flauwe ring, en door hem uit te sluiten hoeven we in bewijzen de nulring niet steeds apart te behandelen. Het is makkelijk in te zien dat de nulring de enige is met $1 = 0$: als $1 = 0$, dan volgt voor elke $x \in R$ dat $x = 1 \cdot x = 0 \cdot x = 0$, dus $R = \{0\}$.

Net als dat een groep ondergroepen heeft, kunnen we spreken van een *deelring* van een ring R. Dat is een deelverzameling van R waarvoor de optelgroep een ondergroep is van die van

R, en die bovendien 1 bevat en gesloten is onder vermenigvuldiging. Op deze manier is een deelring ook weer een ring. Om na te gaan of een deelverzameling van R een deelring is, moeten we nagaan of het 0 en 1 bevat, gesloten is onder optelling en vermenigvuldiging, en 'gesloten is onder het nemen van tegengestelden'. Elke ring R heeft de triviale deelring R. Voor $R = \mathbb{Z}$ is dit de enige, want omdat elke deelring behalve 0 ook 1 bevat en een optelgroep is, bevat het elk natuurlijk getal $n = 1 + \ldots + 1$ en daarom ook $-n$.

5.2.2 Eenheden en lichamen

In een ring heeft elk element een tegengestelde, maar niet per se een inverse. Er is bijvoorbeeld geen geheel getal a zodat $2a = 1$. Een inverteerbaar element van een ring wordt een *eenheid* genoemd, en de verzameling eenheden van een ring R duiden we aan met R^*. Een fijne eigenschap van R^* is dat het een *groep* is onder vermenigvuldiging. Associativiteit wordt van R geërft, 1 is een duidelijk een eenheid met inverse 1, en als a eenheid is, dan volgt uit $aa^{-1} = a^{-1}a = 1$ dat ook a^{-1} eenheid is met inverse a. We hoeven dus alleen nog na te gaan dat R^* gesloten is onder vermenigvuldiging. Als $a, b \in R^*$, dan bestaan er inversen $a^{-1}, b^{-1} \in R$ van a, b. Uit

$$abb^{-1}a^{-1} = aa^{-1} = 1$$

volgt nu dat ab inverteerbaar is met inverse $b^{-1}a^{-1} \in R$, dus $ab \in R^*$. We concluderen dat R^* een groep is onder vermenigvuldiging.

De eenhedengroep $\mathbb{Z}^*$ van $\mathbb{Z}$ bestaat uit slechts de twee getallen 1 en -1, die inderdaad een groep vormen. In $\mathbb{Z}[i] = \{a + bi \in \mathbb{C} : a, b \in \mathbb{Z}\}$, de ring van gehele getallen van Gauss, zijn $i, -1, -i, 1$ duidelijk eenheden, want de vierde macht van elk van hen is 1 dus hun derde macht is hun inverse. Omgekeerd, als α een eenheid is in $\mathbb{Z}[i]$, dan is er een β zodat $\alpha\beta = 1$, en dus ook $|\alpha|^2|\beta|^2 = |\alpha\beta|^2 = 1$. Het kwadraat van de absolute waarde van een element van $\mathbb{Z}[i]$ is geheel, dus $|\alpha|^2$ deelt 1 en is dus 1. Dus α ligt op de eenheidscirkel, en de enige 'roosterpunten' op de eenheidscirkel zijn $i, -1, -i, 1$. We concluderen dat

$$\mathbb{Z}[i]^* = \{i, -1, -i, 1\} = \{i, i^2, i^3, i^4\}, \tag{5.3}$$

en dit is inderdaad een (cyclische) groep onder vermenigvuldiging. De ring $\mathbb{Z}[i]$ is voor het eerst bestudeerd door de Duitser Carl Friedrich Gauss (1777 – 1855), één van de grootste wiskundigen van zijn tijd.

Per definitie is optelling in een ring R een groepsbewerking: $(R, +)$ is een groep. We kunnen ons afvragen of er ook ringen zijn waarvoor $(R, \cdot)$ een groep is. Zo'n ring blijkt niet te bestaan. In dat geval zou 0 namelijk een inverse a hebben, dus $0a = 1$. Maar we weten ook dat $0a = 0$, dus $1 = 0$, tegenspraak. Er zijn echter wel ringen R waarin behalve nul elk element inverteerbaar is, met andere woorden, $R^* = R - \{0\}$. Zulke ringen komen vaak voor, en als zo'n ring bovendien commutatief is, noemen we het een lichaam.

Definitie 5.2.2. Lichaam. *Een* lichaam *is een commutatieve ring R waarvoor $R^* = R - \{0\}$.*

Standaardvoorbeelden van lichamen zijn $\mathbb{Q}$, $\mathbb{R}$ en $\mathbb{C}$. Een voor ons belangrijk voorbeeld is de ring van gehele getallen modulo n, die een lichaam is precies dan als n een priemgetal is. Deze ringen zijn het onderwerp van het volgende hoofdstuk.

Hoofdstuk 6

Modulorekenen met groepen

Bij het bewijzen van de Laatste stelling van Fermat voor $n = 2, 3, 4$ (met als $n = 2$ de constructie van Pythagorese drietallen) hebben we al gebruik gemaakt van modulorekenen in $\mathbb{Z}$, en gebruikten zonder bewijs eigenschappen die intuïtief duidelijk zijn. Twee gehele getallen k, l noemen we congruent modulo m (met $m \in \mathbb{N}$) als ze een veelvoud van m van elkaar verschillen, notatie

$$k \equiv l \pmod{m}.$$

Deze notatie is bedacht door Gauss. We hebben dus $k \equiv l \pmod{m}$ precies dan als $m|(k-l)$. De notatie $k = l \bmod m$, die we ook wel eens gebruiken, betekent iets anders. Met $l \bmod m$ bedoelen we een *getal*, namelijk de rest van l bij deling door m: als $l = qm + r$ met $0 \leq r < m$, dan is $l \bmod m = r$. We hebben dus bijvoorbeeld $32 \equiv 7 \pmod{5}$ en $2 = 7 \bmod 5$, maar niet $32 = 7 \bmod 5$.

Modulorekenen is een heel krachtig hulpmiddel, maar in de loop van de geschiedenis bleek het ook op zichzelf van groot theoretisch belang. Eén van de redenen hiervoor is dat modulorekenen veel algemener is dan rekenen met getallen: het concept staat centraal in de algebra. Eerst diepen we het abstracte modulorekenen uit aan de hand van een voorbeeld, dit blijkt later goed van pas te komen. Daarna bekijken we getal- en groepentheoretische eigenschappen van modulorekenen met gehele getallen.

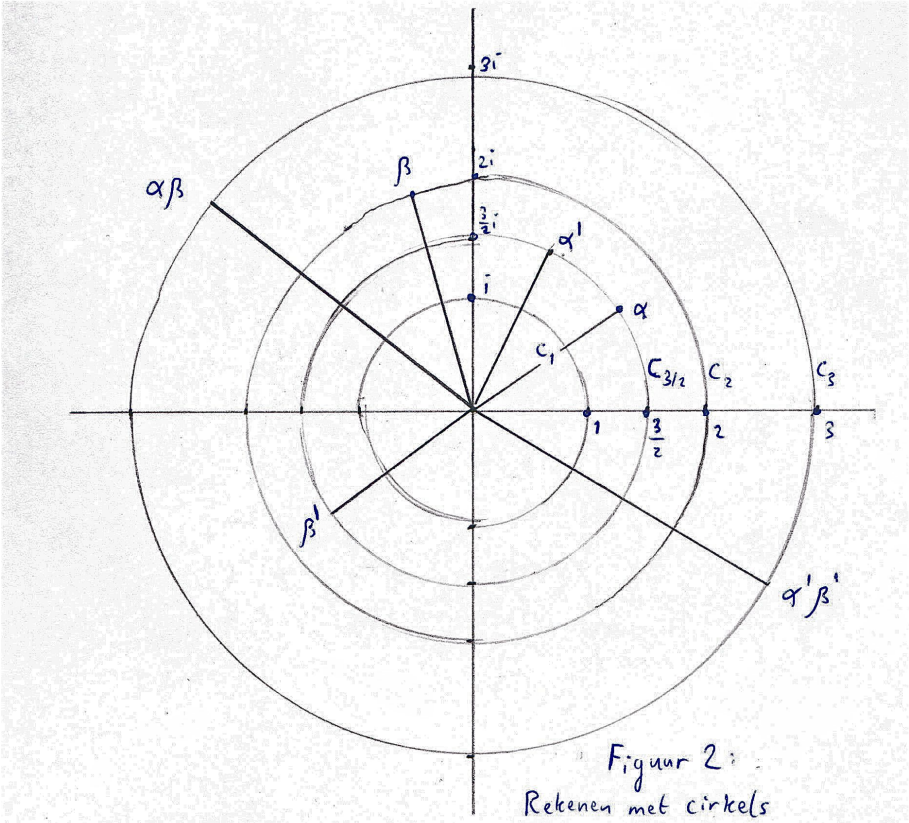
6.1 Rekenen met cirkels en lijnen

Soms is het handig om bepaalde informatie (tijdelijk) te 'vergeten'. In de groep $\mathbb{C}^*$ kunnen we ons bijvoorbeeld alleen richten op de absolute waarde, en het argument vergeten. We kunnen van $\mathbb{C}^*$ overgaan tot de eenvoudigere groep $L := \mathbb{R}_{>0}$ via de absolute-waarde functie, hier berekeningen uitvoeren, en het resultaat terugvertalen naar $\mathbb{C}^*$. Dit heen en weer vertalen is mogelijk dankzij de eigenschap

$$|\alpha\beta| = |\alpha| \cdot |\beta|, \tag{6.1}$$

die zegt dat getallen vermenigvuldigen in $\mathbb{C}^*$ en dan met de absolute waarde 'vertalen' naar L, hetzelfde resultaat geeft als gelijk al de absolute waarde nemen en de vermenigvuldiging uitvoeren in L. We zeggen dat de absolute-waarde functie de bewerkingen van deze groepen

respecteert. Je kunt je dit visualiseren door te zeggen dat je in $\mathbb{C}^*$ 'met cirkels kunt vermenigvuldigen' zoals je in L getallen vermenigvuldigd. Hiermee bedoelen we dat als we twee cirkels rond de oorsprong nemen, zeg C_r en C_s met straal r en s, dan kunnen we twee willekeurige elementen van deze cirkels (de *representanten*) vermenigvuldigen, en het resultaat ligt, onafhankelijk van de keuze van de representanten, altijd op de cirkel C_{rs}. Zie Figuur 2 ter illustratie. We kunnen daarom een eenduidige bewerking op de verzameling cirkels definiëren:



Figuur 2: Rekenen met cirkels

$$C_r \cdot C_s = C_{rs}. \tag{6.2}$$

De informatie die in de cirkels verstopt zit, dat is dus het argument, zijn we vergeten: we beschouwen een cirkel als één element van de 'cirkelgroep' $\{C_r : r \in L\}$. Wegens formule 6.2 is deze groep in essentie gelijk is aan L: de vermenigvuldiging $C_r \cdot C_s = C_{rs}$ kunnen we evengoed schrijven als $r \cdot s = rs$, we veranderen alleen de notatie (en de interpretatie).

We kunnen deze cirkelgroep zelfs geheel omschrijven in termen van de eenheidscirkel C_1, die we voor het gemak C noemen. Hiervoor voeren we het begrip *nevenklasse* in. Als H een ondergroep is van een groep G, en a een element van G, dan definiëren we het 'product' van a

en H door $aH := \{ah : h \in H\}$. We noemen aH een nevenklasse van H.[1] We zijn nevenklassen eigenlijk al tegengekomen in het bewijs van Stelling 5.1.3. Daar zagen we dat de nevenklassen van H een partitie vormen van G, en dat ze allemaal dezelfde kardinaliteit hebben. Uit het bewijs van deze stelling volgt verder de belangrijke eigenschap

$$aH = bH \iff a^{-1}b \in H. \tag{6.3}$$

We zagen daar immers dat de relatie $\sim$ gedefinieerd door $a \sim b \iff a^{-1}b \in H$ een equivalentierelatie is, en de equivalentieklasse van a is aH. Dus (6.3) zegt gewoon dat de equivalentieklassen van a en b overeenkomen precies dan als a en b equivalent zijn.

In ons geval is C de enige cirkel die een ondergroep is van $\mathbb{C}^*$ (ga maar na), daarom richten we ons daarop. Zij $\alpha \in \mathbb{C}^*$ een element met absolute waarde r, ofwel $\alpha \in C_r$. We vragen ons af wat de nevenklasse $\alpha C = \{\alpha c : c \in C\}$ is. Elke $\alpha c \in \alpha C$ heeft absolute waarde $|\alpha| \cdot |c| = |\alpha| = r$, dus $\alpha C \subset C_r$. Omgekeerd, als $x \in C_r$, dan is $|\alpha^{-1}x| = |\alpha^{-1}| \cdot |x| = r^{-1}r = 1$, dus $\alpha^{-1}x \in C$. Dat betekent dat $x = \alpha(\alpha^{-1}x) \in \alpha C$, en dus $C_r \subset \alpha C$. We concluderen dat

$$C_r = \alpha C.$$

Met andere woorden, de cirkel C_r is gelijk aan de nevenklasse αC, met α een willekeurige *representant* van C_r. We kunnen nu vergelijking (6.2) schrijven als

$$\alpha C \cdot \beta C = \alpha\beta C.$$

Met andere woorden: *vermenigvuldiging van twee nevenklassen αC en βC geschiedt door twee elementen van deze nevenklassen, bijvoorbeeld α en β, met elkaar te vermenigvuldigen in $\mathbb{C}^*$, en van het resultaat weer een nevenklasse te maken.* Vermenigvuldiging gebeurt dus in essentie in $\mathbb{C}^*$, maar we vergeten de informatie die in C zit: de elementen van de nevenklassen beschouwen we als één element. Zoals we net zagen is de zo verkregen vermenigvuldiging eenduidig (hangt niet af van de representanten). We zeggen ook wel dat we *modulo de eenheidscirkel* rekenen: het is een abstracte vorm van modulorekenen. De cirkelgroep $\{\alpha C : \alpha \in \mathbb{C}^*\}$ duiden we aan als $\mathbb{C}^*/C$. Deze notatie doet eraan denken dat we $\mathbb{C}^*$ delen door C, en dat is waar in de zin dat we alle nevenklassen van C als één element beschouwen.

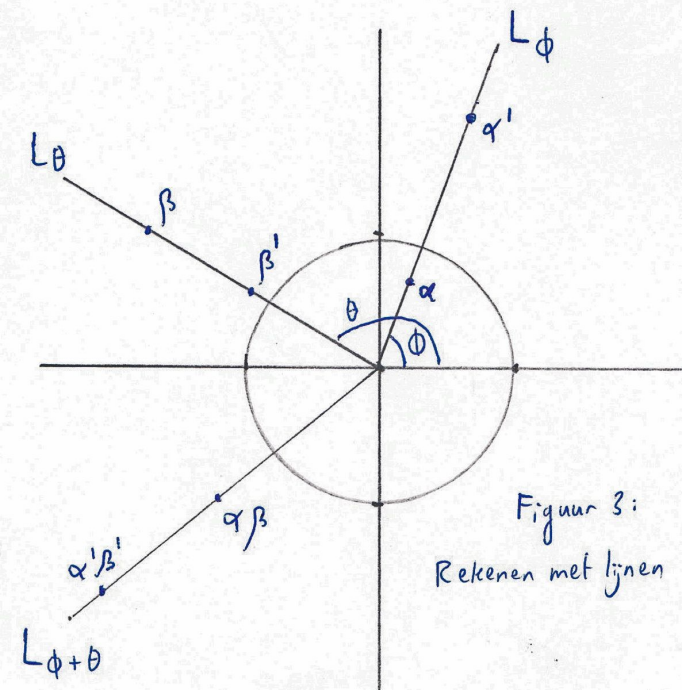
Op dezelfde manier kunnen we alleen naar het argument kijken door $\mathbb{C}^*$ te vertalen naar de eenheidscirkel: we projecteren de elementen van $\mathbb{C}^*$ op de eenheidscirkel, en vergeten zo de absolute waarde. Het vertalen tussen $\mathbb{C}^*$ en C wordt mogelijk gemaakt doordat het argument de bewerkingen respecteert:

$$\arg(\alpha\beta) = \arg(\alpha) + \arg(\beta). \tag{6.4}$$

We kunnen nu twee halve lijnen door de oorsprong[2], zeg L_ρ en L_τ die hoek ρ en τ maken met de positieve x-as, met elkaar te vermenigvuldigen door van elk van hen een representant te nemen, deze met elkaar te vermenigvuldigen, en $L_\rho L_\tau$ te definiëren als de lijn waar dit product op ligt. Het is natuurlijk wel belangrijk dat het resultaat niet afhangt van de keuze van de

[1] Eigenlijk moeten we onderscheid maken tussen linkernevenklassen aH en rechternevenklassen Ha, maar omdat in dit geval de groep commutatief is, en we in het vervolg eigenlijk alleen commutatieve groepen bestuderen, maakt dat voor ons niet uit.

[2] zonder de oorsprong zelf, omdat we in $\mathbb{C}^* = \mathbb{C} - \{0\}$ rekenen



representanten: de vermenigvuldiging moet *welgedefinieerd* zijn. Dat dit zo is, volgt direct uit (6.4), waar we ook uit zien dat

$$L_\rho L_\tau = L_{\rho+\tau}. \tag{6.5}$$

Zie figuur 3 ter illustratie. We hebben dus een lijnengroep $\{L_\rho : \rho \in \mathbb{R}\}$, en deze is wegens formule (6.5) in essentie gelijk aan C. Weer geldt dat de elementen van de lijnengroep de nevenklassen zijn van de lijn $L = L_0$, de enige lijn die een ondergroep is van $\mathbb{C}^*$. Stel namelijk dat $\alpha \in L_\rho$. Een element $\alpha l \in \alpha L$ heeft argument $\arg(\alpha) + \arg(l) = \arg(\alpha)$, en ligt dus op L_ρ. Omgekeerd, als $x \in L_\rho$, dan is

$$\arg(\alpha^{-1}x) = \arg(\alpha^{-1}) + \arg(x) = -\arg(\alpha) + \arg(x) = -\rho + \rho = 0.$$

Dus $\alpha^{-1}x \in L_0$, zodat $x = \alpha(\alpha^{-1}x) \in \alpha L_0$. We concluderen dat $L_\rho = \alpha L$. De lijnengroep is dus gelijk $\{\alpha L : \alpha \in \mathbb{C}^*\}$, en (6.5) legt de vermenigvuldiging van nevenklassen vast:

$$\alpha L \cdot \beta L = \alpha\beta L.$$

Het product van de nevenklassen is dus de nevenklasse van het product. We zeggen weer dat we modulo de lijn L rekenen: de informatie verstopt in L, dat is de absolute waarde, zijn we vergeten. We noteren de lijnengroep als $\mathbb{C}^*/L$.

6.1.1 Homomorfismen en isomorfismen

De ideeën geïntroduceerd in bovenstaande voorbeelden zijn heel belangrijk in de algebra, en komen overal terug. Formeel zeggen we dat de absolute-waarde functie een *homomorfisme* is

van $\mathbb{C}^*$ naar L, en de cirkelgroep $\mathbb{C}^*/C$ is zelfs *isomorf* met L. Analoog is de functie $\alpha \mapsto e^{i \arg \alpha}$ een homomorfisme van $\mathbb{C}^*$ naar C, en de lijnengroep $\mathbb{C}^*/L$ is zelfs isomorf met C.

Definitie 6.1.1. Homomorfisme, Isomorfisme. *Een* homomorfisme *van een groep G naar een groep H is een functie $f : G \to H$ die de bewerkingen respecteert:*

$$f(xy) = f(x)f(y) \qquad \text{voor alle } x, y \in G.$$

Met andere woorden, de afbeelding van het product is het product van de afbeeldingen. Een bijectief homomorfisme heet een isomorfisme. *Als er een isomorfisme tussen G en H bestaat, noemen we G en H isomorf, notatie*

$$G \cong H.$$

Homomorfismen hebben twee fundamentele eigenschappen die heel vaak gebruikt worden. Laat G, H twee groepen zijn met respectievelijk e, e' als eenheidselement, en zij $f : G \to H$ een homomorfisme.

1. Het eenheidselement wordt afgebeeld naar het eenheidselement: $f(e) = e'$. Dit volgt door in $f(e) = f(ee) = f(e)f(e)$ links en rechts te vermenigvuldigen met de inverse van $f(e)$.
2. De afbeelding van de inverse is de inverse van de afbeelding: voor alle $x \in G$ geldt $f(x^{-1}) = f(x)^{-1}$. Immers, $f(x)f(x^{-1}) = f(xx^{-1}) = f(e) = e'$, en op dezelfde manier $f(x^{-1})f(x) = e'$, dus $f(x^{-1})$ is de inverse van $f(x)$.

In de voorbeelden van zojuist merkten we op dat de cirkelgroep en L in essentie gelijk zijn. Dit kunnen we nu hard maken. De functie $f : \mathbb{C}^*/C \to L : \alpha C \mapsto |\alpha|$ een homomorfisme. Dit volgt direct uit

$$f(\alpha C \cdot \beta C) = f(\alpha\beta C) = |\alpha\beta| = |\alpha| \cdot |\beta| = f(\alpha C)f(\beta C).$$

Bovendien is f bijectief. Surjectiviteit is duidelijk; als bijvoorbeeld $r \in L$, dan geldt voor een willekeurige α uit de cirkel rC met straal r dat $f(\alpha C) = r$. Stel dat $f(\alpha C) = f(\beta C)$, ofwel $|\alpha| = |\beta|$. Dan is $|\alpha\beta^{-1}| = |\alpha| \cdot |\beta^{-1}| = |\alpha| \cdot |\beta|^{-1} = 1$, dus $\alpha\beta^{-1} \in C$. Uit (6.3) volgt nu dat $\alpha C = \beta C$, dus f is ook injectief. We concluderen dat f een isomorfisme is, en dus

$$\mathbb{C}^*/C \cong L.$$

We kunnen nu ook laten zien dat de lijnengroep en C inderdaad isomorf zijn. De functie $g : \mathbb{C}^*/L \to C : \alpha L \mapsto \alpha/|\alpha|$ is een homomorfisme, want

$$g(\alpha L \cdot \beta L) = g(\alpha\beta L) = \frac{\alpha\beta}{|\alpha\beta|} = \frac{\alpha}{|\alpha|}\frac{\beta}{|\beta|} = g(\alpha L)g(\beta L).$$

Bovendien is g bijectief. Voor de surjectiviteit merken we op dat als $\alpha \in C$, dan is $g(\beta) = \alpha$ voor alle β op de lijn αL. Als $g(\alpha L) = g(\beta L)$, dan is $\alpha/|\alpha| = \beta/|\beta|$, en dus $\alpha\beta^{-1} = |\alpha|/|\beta| \in L$. Uit (6.3) volgt dat $\alpha L = \beta L$, dus g is ook injectief. Dus g is een isomorfisme, en dus

$$\mathbb{C}^*/L \cong C.$$

In de algebra maken we niet echt onderscheid tussen isomorfe groepen, ze zijn als groep 'in essentie gelijk'. Dit doet denken aan het grapje dat topologen geen verschil zien tussen een koffiekopje en een donut: die zijn 'topologisch hetzelfde' (hom*e*omorf). Isomorfe groepen verschillen alleen in de symbolen of 'labels' die we voor de elementen gebruiken. Bijvoorbeeld de vermenigvuldiggroep $\{1, -1\}$ is isomorf met de optelgroep $\{0, 1\}$ waarin we modulo 2 rekenen. Hier is $f(1) = 0, f(-1) = 1$ een isomorfisme, die we kunnen zien als 'label-verandering'.[3]

Om isomorfe groepen als 'hetzelfde' te kunnen zien, moet de relatie $\cong$ wel aan wat voorwaarden voldoen. Het zou bijvoorbeeld vreemd zijn als G wel hetzelfde zou zijn als H, maar H niet als G! Gelukkig is hieraan voldaan, want $\cong$ is een equivalentierelatie. De equivalentieklassen bestaan dus uit groepen die we als hetzelfde beschouwen, informeel zouden we zo'n equivalentieklasse zelfs als één groep kunnen beschouwen. Zij G, H, J willekeurige groepen. De identieke afbeelding $f : G \to G : x \mapsto x$ is een isomorfisme, dus $G \cong G$: de relatie is reflectief. Als $f : G \to H$ een isomorfisme is, dan is $f^{-1} : H \to G$ een isomorfisme. Voor het bewijs merken we ten eerste op dat de inverse van een bijectieve functie bestaat en bijectief is. Ten tweede, als $u, v \in H$ dan zijn er (omdat f surjectief is) $x, y \in G$ zodat $f(x) = u, f(y) = v$, dus

$$f^{-1}(uv) = f^{-1}(f(x)f(y)) = f^{-1}(f(xy)) = xy = f^{-1}(u)f^{-1}(v),$$

dus f^{-1} is een isomorfisme. Uit $G \cong H$ volgt dus $H \cong G$: de relatie is symmetrisch. Als $f : G \to H$ en $g : H \to J$ homomorfismen zijn, dan is de samenstelling $g \circ f : G \to J$ een homomorfisme. Voor alle $x, y \in G$ geldt namelijk

$$(g \circ f)(xy) = g(f(xy)) = g(f(x)f(y)) = g(f(x))g(f(y)) = (g \circ f)(x) \cdot (g \circ f)(y).$$

Als bovendien f en g isomorfismen zijn, dan is $g \circ f$ als samenstelling van bijectieve functies bijectief en dus een isomorfisme. Dus uit $G \cong H$ en $H \cong J$ volgt $G \cong J$: de relatie is transitief. We hebben dus inderdaad een equivalentierelatie.

Het rekenen modulo een ondergroep van een groep, zoals we deden met L en C, is van grote algemeenheid. We kunnen dit doen met *elke* ondergroep van een willekeurige Abelse groep.[4]

Stelling 6.1.2. *Zij G een Abelse groep, en H een ondergroep van G. We definiëren een bewerking op de verzameling $G/H = \{aH : a \in G\}$ van nevenklassen van H, namelijk $aH \cdot bH = abH$. Met deze bewerking wordt G/H een abelse groep.*

Bewijs. We moeten eerst nagaan dat de bewerking welgedefinieerd is. Stel dat $aH = bH$ en $cH = dH$, we willen dus bewijzen dat $acH = bdH$. Van (6.3) weten we dat $a^{-1}b \in H$ en $c^{-1}d \in H$, er zijn dus $h_1, h_2 \in H$ zodat $a^{-1}b = h_1$ en $c^{-1}d = h_2$. Dat betekent dat $c^{-1}a^{-1}bd = h_1h_2 \in H$ (hier gebruiken we dat G Abels is.) Omdat $c^{-1}a^{-1}$ de inverse van ac is, kunnen we dit schrijven als $(ac)^{-1}bd \in H$, en door weer (6.3) te gebruiken volgt dat $acH = bdH$, zoals gewenst.

[3] Joseph Rotman omschrijft het zo: het enige verschil tussen twee isomorfe groepen is dat de een in het Engels is geschreven en de ander in het Frans, een isomorfisme tussen de twee is een woordenboek die de een in de ander vertaalt. (Joseph J. Rotman, *An Introduction to the Theory of Groups*, Springer Verlag 1995, Fourth Edition pagina 17.)

[4] Het kan ook met een bepaalde klasse van ondergroepen van niet-Abelse groepen, de zogenaamde *normale* ondergroepen. Omdat wij eigenlijk alleen met Abelse groepen zullen werken, gaan we hier niet verder op in.

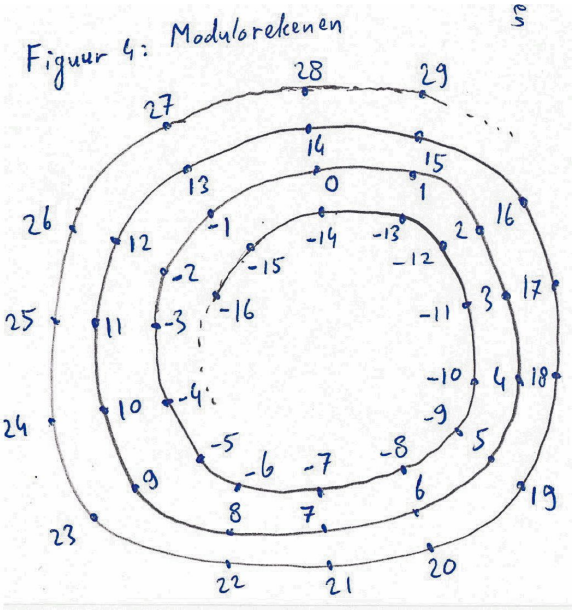
Dat G/H een groep is, volgt bijna direct uit het feit dat G een groep is. Het eenheidselement is eH, want voor alle $aH \in G/H$ is $aH \cdot eH = aeH = aH$. De tegengestelde van aH is $a^{-1}H$, want $aH \cdot a^{-1}H = aa^{-1}H = eH$. Associativiteit volgt uit die van G, want

$$(aH \cdot bH) \cdot cH = abH \cdot cH = (ab)cH = a(bc)H = aH \cdot bcH = aH \cdot (bH \cdot cH).$$

Dus G/H is een groep. □

Als $aH = bH$, dan zegt men vaak dat a en b congruent zijn modulo H, notatie $a \equiv b \pmod{H}$. Dit is dus precies de relatie $\sim$ die we voor het eerst tegenkwamen in Stelling 5.1.3, dus het is een equivalentierelatie. Bovendien is het consistent met de bewerking, dat wil zeggen $a \equiv b \pmod{H}$ en $c \equiv d \pmod{H}$ impliceert dat $ac \equiv bd \pmod{H}$. Dit betekent namelijk precies dat de bewerking op G/H welgedefinieerd is.

6.2 Modulorekenen in $\mathbb{Z}$



In $\mathbb{Z}$ is het soms handig om de veelvouden van een natuurlijk getal n te 'vergeten': we kijken alleen naar de *rest* bij deling door n. We kunnen dit visualiseren door in gedachte de getallenlijn op te rollen tot een cirkel van n punten, en de punten die op elkaar komen te liggen met elkaar te identificeren. Zie Figuur 4 ter illustratie. Twee getallen k en m die we identificeren noemen we congruent modulo n, notatie $k \equiv m \pmod{n}$. Zoals we in het begin van dit hoofdstuk al opmerkten, is $k \equiv m \pmod{n}$ precies dan als $n|(m-k)$.

Formeel gaan we als volgt te werk. De verzameling $\{kn : k \in \mathbb{Z}\}$ van n-vouden schrijven we als[5] $n\mathbb{Z}$, dit is een ondergroep van $\mathbb{Z}$. Volgens Stelling 6.2.2 is $\mathbb{Z}/n\mathbb{Z}$ een groep, met als bewerking $(k + n\mathbb{Z}) + (m + n\mathbb{Z}) = (k + m) + n\mathbb{Z}$. Net als daar schrijven we $k \equiv m \pmod{n\mathbb{Z}}$ als $k + n\mathbb{Z} = m + n\mathbb{Z}$. Een nevenklasse $k + n\mathbb{Z} = \{\ldots, k - 2n, k - n, k, k + n, k + 2n, \ldots\}$ bestaat precies uit de getallen met rest k bij deling door n, dus $k \equiv m \pmod{n\mathbb{Z}}$ betekent niets anders dan dat k en m dezelfde rest hebben bij deling door n. Met andere woorden, $k \equiv m \pmod{n\mathbb{Z}}$ betekent hetzelfde als $k \equiv m \pmod{n}$. Omdat het alleen maar verwarrend is om ingewikkelde notatie te gebruiken, zullen we voortaan gewoon $k \equiv m \pmod{n}$ schrijven.

[5] De notatie $n\mathbb{Z}$ is enigszins verwarrend, het is namelijk geen nevenklasse van de groep $\mathbb{Z}$. We schrijven $\mathbb{Z}$ namelijk additief, dus voor een ondergroep H van $\mathbb{Z}$ zouden we een nevenklasse als $a + H$ noteren.

De nevenklassen $k + n\mathbb{Z}$ korten we af met $\overline{k}$, en we noemen ze de *restklassen modulo n*. Er zijn precies n verschillende restklassen, namelijk[6] $\overline{0}, \overline{1}, \ldots, \overline{n-1}$, en we hebben dus

$$\mathbb{Z}/n\mathbb{Z} = \{\overline{k} : k \in \mathbb{Z}\} = \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{n}.\}$$

De optelling in deze groep kunnen we nu kort noteren als $\overline{k} + \overline{l} = \overline{k+l}$. Maar we kunnen gelukkig ook vermenigvuldigen in $\mathbb{Z}/n\mathbb{Z}$: het is een *ring*. Het zal geen verassing zijn dat we de vermenigvuldiging definiëren als $\overline{k} \cdot \overline{l} := \overline{kl}$. Wel moeten we eerst nagaan dat deze bewerking welgedefinieerd is, en dat het resultaat dus niet van de representanten k en l afhangt. Stel a en b zijn andere representanten: $\overline{a} = \overline{k}$ en $\overline{b} = \overline{l}$. Dan zijn er $p, q \in \mathbb{Z}$ zodat $a = k + pn$ en $b = l + qn$. Dus $ab = (k + pn)(l + qn) = kl + (kq + pl + pqn)n$, en hieruit volgt dat $\overline{ab} = \overline{kl}$. Vermenigvuldiging is dus welgedefinieerd. Nu volgt makkelijk:

Stelling 6.2.1. *Voor alle natuurlijke getallen n is $\mathbb{Z}/n\mathbb{Z}$ een commutatieve ring.*

Bewijs. We weten al dat het een groep is onder optelling. De overige ringeigenschappen volgen eigenlijk meteen uit die van $\mathbb{Z}$. Bijvoorbeeld distributiviteit gaat zo:

$$\overline{k}(\overline{l} + \overline{m}) = \overline{k}(\overline{l+m}) = \overline{k(l+m)} = \overline{kl + km} = \overline{kl} + \overline{km} = \overline{k} \cdot \overline{l} + \overline{k} \cdot \overline{m}.$$

Op dezelfde manier volgen associativiteit en commutativiteit van vermenigvuldiging. Het nulelement is $\overline{0}$ en het eenheidselement $\overline{1}$, want $\overline{k} + \overline{0} = \overline{k+0} = \overline{k}$ en $\overline{k} \cdot \overline{1} = \overline{k \cdot 1} = \overline{k}$. De tegengestelde van $\overline{k}$ is $\overline{-k}$, want $\overline{k} + \overline{-k} = \overline{k + (-k)} = \overline{0}$. Hiermee wordt $\mathbb{Z}/n\mathbb{Z}$ een commutatieve ring. □

6.2.1 Een voorbeeld: Fermat-getallen

De kracht van het modulorekenen kunnen we illustreren door te berekenen wat de rest is van

$$p = 2^{2^7} = 340282366920938463463374607431768211456$$

bij deling door 19, iets wat met de hand een heel karwij lijkt als we gewoon proberen q en s te vinden zodat $p = 19q + s$. Het kan gelukkig veel makkelijker, namelijk door de rij $2, 2^2, 2^{2^2}, 2^{2^3}, \ldots$ 'in stapjes' modulo 19 te bestuderen. Deze rij groeit ongelooflijk hard. Het $n+1$-ste element 2^{2^n} heeft ongeveer ${}^{10}\log(2^{2^n}) = {}^{10}\log(2) \cdot 2^n \approx 0.3 \cdot 2^n$ cijfers: het aantal cijfers groeit exponentieel!

Voor $n \geq 1$ geldt dat $2^{2^n} = (2^{2^{n-1}})^2$, dus voor de bijbehorende restklassen in $\mathbb{Z}/19\mathbb{Z}$ hebben we

$$\overline{2^{2^n}} = \overline{2^{2^{n-1}} \cdot 2^{2^{n-1}}} = \overline{2^{2^{n-1}}} \cdot \overline{2^{2^{n-1}}} = \overline{2^{2^{n-1}}}^2.$$

Dus $2^{2^n} \equiv (2^{2^{n-1}})^2 \pmod{19}$, zodat

$$2^{2^0} \equiv 2, \quad 2^{2^1} \equiv 2^2 \equiv 4, \quad 2^{2^2} \equiv 4^2 \equiv 16 \equiv -3, \quad 2^{2^3} \equiv (-3)^2 \equiv 9, \quad 2^{2^4} \equiv 9^2 \equiv 81 \equiv 5,$$
$$2^{2^5} \equiv 5^2 \equiv 6, \quad 2^{2^6} \equiv 6^2 \equiv -2, \quad 2^{2^7} \equiv (-2)^2 \equiv 4, \quad 2^{2^8} \equiv 4^2 \equiv -3, \ldots \pmod{19}.$$

[6] We hadden ook andere representanten dan 1 t/m $n-1$ kunnen nemen, maar dit rekent wel zo makkelijk.

De gezochte rest is dus 4. We kunnen zelfs van *alle* $n \geq 1$ zeggen wat de rest van 2^{2^n} bij deling door 19 is. Voor $n = 0$ is de rest 2. Vanaf $n = 1$ zien we dat het patroon

$$\overline{4}, \overline{-3}, \overline{9}, \overline{5}, \overline{6}, \overline{-2}, \quad \overline{4}, \overline{-3}, \overline{9}, \overline{5}, \overline{6}, \overline{-2}, \ldots$$

van restklassen $\overline{2^n} \in \mathbb{Z}/19\mathbb{Z}$ zich steeds blijft herhalen op een manier die doet denken aan een repeterende breuk. Voor $n \geq 1$ en voor n congruent $1, 2, 3, 4, 5, 0$ modulo 6 geldt dus respectievelijk

$$2^{2^n} \equiv 4, -3, 9, 5, 6, -2 \pmod{19}.$$

De getallen $F_n = 2^{2^n} + 1$ worden vaak de *Fermat-getallen* genoemd. Modulo 19 zijn deze dus altijd congruent aan $5, -2, 10, 6, 7, -1$, en nooit aan 0. Dat betekent dat geen enkel Fermat-getal priemdeler 19 heeft. Fermat berekende dat $F_1 = 5, F_2 = 17, F_3 = 257$ en $F_4 = 65537$ priem zijn, en vermoedde dat ze wel eens allemaal priem zouden kunnen zijn. Fermat had het niet vaak mis, maar Euler liet zien dat hij hier toch fout zat: F_5 is namelijk niet priem.[7] Het is niet moeilijk te bewijzen $p-1$ veelvoud is van 2^{n+1} als p priemdeler van F_n is, zie bijvoorbeeld §5.3 van [2] voor een bewijs. We hoeven F_5 dus alleen op priemfactoren te controleren die congruent 1 zijn modulo 64. Euler probeerde daarom $193, 257, 449, 577, 641, \ldots$ en bij 641 is het raak. Dit volgt na een kleine berekening van getallen 2^{2^n} modulo 641:

$$2^{2^1} \equiv 4, 2^{2^2} \equiv 16, 2^{2^3} \equiv 256, 2^{2^4} \equiv 65536 \equiv 64100+1436 \equiv 154, 2^{2^5} \equiv 23716 \equiv 37 \cdot 641 - 1 \equiv -1.$$

Dus $2^{2^5} + 1 \equiv 0 \pmod{641}$, ofwel 641 is deler van F_5. Bovenstaande berekening kunnen we in principe voortzetten voor grotere n. Na hoogstens 641 stappen komen we uit op een rest die we al eerder hebben gehad, en vanaf dan krijgen we een 'repeterende staart'. Als de restklasse $\overline{-1}$ in deze staart zou voortkomen, dan zouden er oneindig veel Fermat-getallen deelbaar zijn door 641 en dus niet priem zijn. Maar omdat $2^{2^5} \equiv -1 \pmod{641}$ is $2^{2^6} \equiv 1, 2^{2^7} \equiv 1, 2^{2^8} \equiv 1, \ldots \pmod{641}$, we krijgen dus een oneindige staart van enen. Er is dus behalve F_5 geen enkel Fermat-getal deelbaar door 641. Dit geldt ook in het algemeen: als F_n het kleinste Fermat-getal is dat deelbaar is door een bepaald (priem)getal p, dan is $2^{2^n} \equiv -1 \pmod{p}$ en dus $2^{2^k} \equiv 1 \pmod{p}$ ofwel $F_k \equiv 2 \pmod{p}$ voor alle $k > n$. Elk priemgetal deelt dus hoogstens één Fermat-getal.[8] Bovendien, als p deler is van $F_n = 2^{2^n} + 1$, dan is p deler van het *Mersenne-getal* $F_{n+1} - 2 = 2^{2^{n+1}} - 1$.[9] De Mersenne-getallen zijn de getallen van de vorm $2^k - 1$, dat zijn dus $1, 3, 7, 15, 31, 63, 127, 255, 511, 1023, \ldots$.

6.2.2 Ondergroepen, delers en de formule van Gauss

We richten ons nu op eigenschappen van de ringen $\mathbb{Z}/n\mathbb{Z}$ zelf. Eerst beschouwen we alleen de optelgroep van $\mathbb{Z}/n\mathbb{Z}$, die op zich al een aantal interessante eigenschappen heeft. Als voorbeeld bekijken we $\mathbb{Z}/8\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}, \overline{6}, \overline{7}\}$. Omdat

$$\mathbb{Z}/8\mathbb{Z} = \{\ldots, \overline{3}, \overline{6}, \overline{1}, \overline{4}, \overline{7}, \overline{2}, \overline{5}, \overline{0} \ , \ \overline{3}, \overline{6}, \overline{1}, \overline{4}, \overline{7}, \overline{2}, \overline{5}, \overline{0}, \ldots\} = \{k\overline{3} : k \in \mathbb{Z}\},$$

[7] Bron: zie [2].

[8] Alleen voor $p = 2$ volgt dit niet uit het bovenstaande, maar het is duidelijk dat de Fermat-getallen oneven zijn.

[9] Vanwege de zojuist genoemde staart met enen deelt p zelfs *alle* Mersenne-getallen van de vorm $2^{2^k} - 1$ met $k > n$, maar dat volgt ook direct met inductie uit $2^{2^{k+1}} - 1 = (2^{2^k} - 1)(2^{2^k} + 1)$.

is de groep cyclisch, en wordt voortgebracht door $\overline{3}$. Op dezelfde manier kun je zien dat $\overline{1}$, $\overline{5}$ en $\overline{7}$ de groep voortbrengen, maar $\overline{2}, \overline{4}, \overline{6}$ en $\overline{8} = \overline{0}$ niet. Bijvoorbeeld

$$\{k\overline{6} : k \in \mathbb{Z}\} = \{\overline{6}, \overline{4}, \overline{2}, \overline{0}\} \qquad \text{en} \qquad \{k\overline{4} : k \in \mathbb{Z}\} = \{\overline{4}, \overline{0}\}.$$

Het is duidelijk dat dit komt doordat $\text{ggd}(x, 8) = 1$ voor $x = 1, 3, 5, 7$ en $\text{ggd}(x, 8) \neq 1$ voor $x = 2, 4, 6, 0$. De verzamelingen $\{k\overline{x} : k \in \mathbb{Z}\}$ zijn echter wel allemaal *ondergroepen* van $\mathbb{Z}/8\mathbb{Z}$. Zo vinden we, naast de triviale ondergroepen $\{\overline{0}\}$ (voor $x = 0$) en $\mathbb{Z}/8\mathbb{Z}$ (voor $x = 1, 3, 5, 7$) ook de ondergroepen $\{\overline{0}, \overline{2}, \overline{4}, \overline{6}\}$ (voor $x = 2, 6$) en $\{\overline{0}, \overline{4}\}$ (voor $x = 4$). Merk op dat de orde van de ondergroep $\{k\overline{x} : k \in \mathbb{Z}\}$ steeds gelijk is aan $8/\text{ggd}(x, 8)$. In het bijzonder is er voor elke deler d van 8 een ondergroep van orde d; neem gewoon $x = 8/d$.

De volgende stelling generaliseert wat we hier gezien hebben. Het bewijs wordt intuïtief duidelijker als we er een plaatje bij maken. Daarom hebben we in Figuur 5 de ondergroepen van $\mathbb{Z}/30\mathbb{Z}$ weergegeven. De elementen van orde d in de ondergroepen H_d (zie de stelling voor de betekenis) hebben we omcirkeld. Omdat volgens Stelling 5.1.3 de orde van elke ondergroep en elk element deler is van $|\mathbb{Z}/n\mathbb{Z}| = n$, hoeven we ons alleen te richten op de delers van n.

We gebruiken de *totient-functie* (of ϕ-functie) van Euler, een belangrijke functie in de getaltheorie gedefinieerd door

$$\phi : \mathbb{N} \to \mathbb{N} : \phi(n) \text{ is het aantal gehele getallen } k \text{ met } 1 \leq k \leq n \text{ waarvoor } \text{ggd}(n, k) = 1.$$

Dus bijvoorbeeld $\phi(1) = \phi(2) = 1$, en $\phi(p) = p - 1$ voor alle priemgetallen p. Voor $n = 8$ zijn $1, 3, 5, 7$ de getallen $1 \leq k \leq 8$ die copriem zijn met 8, dus $\phi(8) = 4$. In het algemeen is de verhouding tussen n en $\phi(n)$ kleiner naarmate n meer delers heeft.

Stelling 6.2.2. .

1). De optelgroep $\mathbb{Z}/n\mathbb{Z}$ heeft voor elke positieve deler d van n precies één ondergroep H_d van orde d, voortgebracht door de restklasse van $x_d := n/d$.

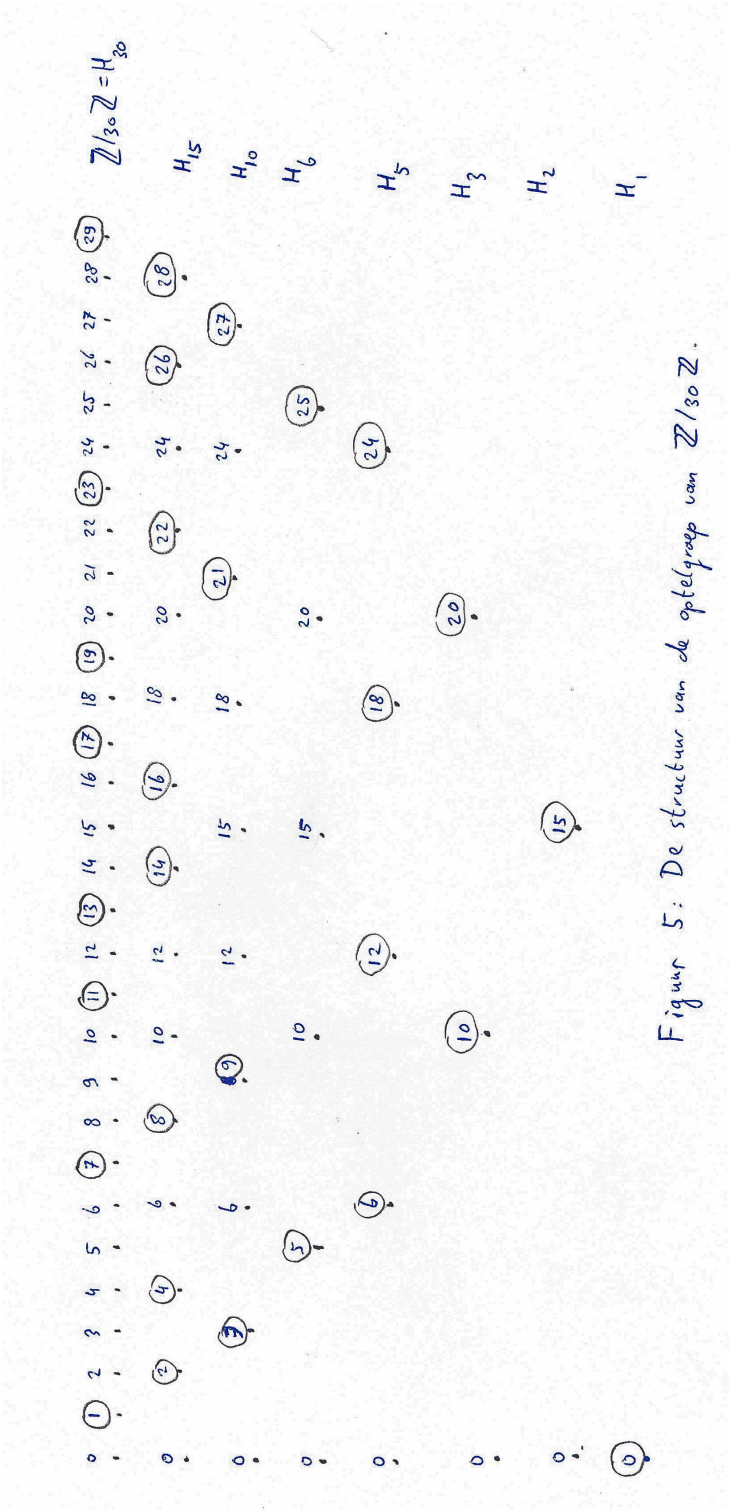
2.) De elementen van orde d in $\mathbb{Z}/n\mathbb{Z}$, met d een deler van n, zijn precies de $\phi(d)$ elementen $k\overline{x_d} \in H_d$ met $k \in \{1, 2, \ldots, d\}$ waarvoor $\text{ggd}(d, k) = 1$.

Bewijs. Zij d een deler van n. Dan is er een $x_d \in \mathbb{Z}$ zodat $dx_d = n$. Definieer H_d als de ondergroep voortgebracht door $\overline{x_d}$, dus $H_d = \{k\overline{x_d} : k \in \mathbb{Z}\}$. Omdat $dx_d = n$, is duidelijk dat de orde van x_d gelijk is aan d, dus ook de orde van H_d is d. Hiermee hebben we het bestaan van een ondergroep van orde d aangetoond, en moeten we alleen nog bewijzen dat hij uniek is. Stel J is een ondergroep van orde d, en zij k het kleinste natuurlijke getal waarvoor $\overline{k} \in J$. Stel $\overline{m} \in J$. Er zijn $r, q \in \mathbb{Z}$ zodat $m = qk + r$ met $0 \leq r < k$. Omdat J een groep is, is ook de q-voudige som $\overline{k} + \ldots + \overline{k} = q\overline{k}$ bevat in J, en dus ook $\overline{r} = \overline{m} - q\overline{k} \in J$. Omdat $0 \leq r < k$ volgt uit de minimaliteit van k dat $r = 0$, dus $\overline{m} = q\overline{k}$. We concluderen dat J cyclisch is en wordt voortgebracht door $\overline{k}$. Omdat J orde d heeft, heeft ook $\overline{k}$ orde d, dus $dk \equiv 0 \pmod{n}$, ofwel $n|dk$. Omdat $dx_d = n$ betekent dit dat $dx_d|dk$, dus $x_d|k$. Uit $dx_d = n$ volgt ook dat $x_d|n$, dus

$$x_d|k \qquad \text{en} \qquad x_d|n. \tag{6.6}$$

J heeft orde d, we schrijven $J = \{y_1, y_2, \ldots, y_d\}$. Voor elke $y_i \in J$ kunnen we een unieke representant $m_i \in \{0, 1, 2, \ldots, n-1\}$ kiezen zodat $\overline{m_i} = y_i$. Zij

$$K = \{m_1, m_2, \ldots, m_d\}$$



Figuur 5: De structuur van de opteldgroep van ℤ/30ℤ.

de verzameling van de d verschillende representanten. Omdat J wordt voortgebracht door $\overline{k}$, is er voor alle m_i een $q_i \in \mathbb{Z}$ zodat $\overline{m_i} = q_i\overline{k}$. Dus er is een $s_i \in \mathbb{Z}$ zodat $m_i = q_ik + s_in$. Uit (6.6) volgt dat

$$x_d | m_i \qquad \text{voor alle } i,$$

dus verzameling K bestaat uit d verschillende veelvouden van x_d die bovendien in $\{0, 1, 2, \ldots, n-1\}$ liggen. Maar er zijn *precies* d van die veelvouden, namelijk 0, x_d, $2x_d$, $\ldots$, $(d-1)x_d$. We concluderen dat al deze veelvouden in K voorkomen, dus ook $x_d \in K$ en bijgevolg $\overline{x_d} \in J$. Dus alle d de verschillende elementen van H_d liggen in J, en omdat J maar d elementen bevat volgt dat $J = H_d$. Hiermee is bewezen dat H_d de unieke ondergroep van orde d is, waarmee 1) is bewezen.

Stel $\overline{a}$ is een element van orde d. Dan is $d\overline{a} = \overline{0}$, dus $da = qn$ voor een $q \in \mathbb{Z}$, ofwel $da = qdx_d$ zodat $a = qx_d$. Dus

$$\overline{a} = \overline{qx_d} = \overline{x_d + \ldots + x_d} = \overline{x_d} + \ldots + \overline{x_d} = q\overline{x_d},$$

dus $\overline{a} \in H_d$. De elementen van orde d in $\mathbb{Z}/n\mathbb{Z}$ zijn dus precies de elementen van orde d in H_d; we gaan nu achterhalen welke dat zijn. Zij $k\overline{x_d} \in H_d$. Stel eerst dat $\text{ggd}(d, k) = p \neq 1$. Dan zijn d/p en k/p geheel, en

$$(d/p) \cdot kx_d = (k/p) \cdot dx_d = (k/p) \cdot n.$$

Dus $\overline{(d/p)kx_d} = 0$, dus kx_d heeft hoogstens orde $d/p < d$. Als $\text{ggd}(d, k) \neq p$, dan is de orde van kx_d dus niet d.

Stel nu dat $\text{ggd}(d, k) = 1$. Zij q de orde van kx_d. Dan is $q\overline{kx_d} = \overline{0}$, en dus $n | qkx_d$, ofwel $dx_d | qkx_d$. Dus $d | qk$, en omdat $\text{ggd}(d, k) = 1$ volgt dat $d | q$ en dus $d \leq q$. Omgekeerd zijn de q elementen $0\overline{kx_d}, 1\overline{kx_d}, 2\overline{kx_d}, \ldots, (q-1)\overline{kx_d} \in H_d$ allen verschillend, zoals we op pagina 25 al opmerkten, dus $d \geq q$. Conclusie: $d = q$. Dus als $\text{ggd}(d, k) = 1$, dan is de orde van kx_d wél d.

We concluderen dat de elementen van orde d in $\mathbb{Z}/n\mathbb{Z}$ precies de $\phi(d)$ verschillende elementen $k\overline{x_d} \in H_d$ zijn met $1 \leq k \leq d$ waarvoor k copriem is met d. Hiermee is ook 2) bewezen. □

Voor elke $d | n$ is er dus precies één ondergroep van orde d, en we weten van Stelling 5.1.3 dat dit de enige ondergroepen zijn. Het aantal delers van n is dus bepalend voor de structuur van $\mathbb{Z}/n\mathbb{Z}$: de groep wordt 'complexer' of 'symmetrischer' naarmate n meer delers krijgt. In Figuur 5 zien we geïllustreerd dat hoe meer delers een getal m gemeen heeft met n, in hoe meer ondergroepen het voorkomt. Dat aantal ondergroepen is zelfs gelijk aan het aantal delers dat het gemeen heeft. Bijvoorbeeld de gemene delers van 24 en 30 zijn de getallen $k \in \{1, 2, 3, 6\}$, en de ondergroepen waar $\overline{24}$ in voorkomt zijn precies de corresponderende $H_{30/k}$. Ook zien we in het plaatje dat de orde van een element precies de orde is van de kleinste ondergroep waar het in voorkomt. Merk ook op dat het plaatje symmetrisch is rond de '15-as'.

De ϕ-functie van Euler heeft een aantal mooie eigenschappen. Zo geldt bijvoorbeeld de volgende formule van Gauss, waarbij we sommeren over alle delers d van n.

Gevolg 6.2.3. (Gauss) *Voor alle $n \in \mathbb{N}$ geldt:*

$$\sum_{d|n} \phi(d) = n. \tag{6.7}$$

Bewijs. Het aantal elementen van een eindige groep is gelijk aan het aantal elementen met orde k, gesommeerd over alle mogelijke ordes k. In de groep $\mathbb{Z}/n\mathbb{Z}$ zijn de mogelijke ordes de delers d van n, en voor elk van deze d zijn er $\phi(d)$ elementen. Het linkerlid van (6.7) is dus het aantal elementen van $\mathbb{Z}/n\mathbb{Z}$, en dat zijn er n. □

Hoewel het bestuderen van cyclische groepen eigenlijk onder groepentheorie valt, kunnen we het net zo goed tot de getaltheorie rekenen. Elke cyclische groep is namelijk isomorf met $\mathbb{Z}$ of met één van de groepen $\mathbb{Z}/n\mathbb{Z}$. Stel namelijk dat G cyclisch is, ofwel

$$G = \{x^k : k \in \mathbb{Z}\}.$$

met x een voortbrenger van G. Stel eerst dat x eindige orde n heeft. De functie $f : \mathbb{Z}/n\mathbb{Z} \to G : \overline{k} \mapsto x^k$ is een homomorfisme, want

$$f(\overline{k+l}) = x^{k+l} = x^k x^l = f(\overline{k})f(\overline{l}).$$

Als $f(\overline{k}) = f(\overline{l})$, ofwel $x^k = x^l$, dan weten van Lemma 5.1.2 dat $k \equiv l \pmod{n}$, ofwel $\overline{k} = \overline{l}$. Dus f is injectief. Uit hetzelfde lemma volgt dat G orde n heeft, dus G en $\mathbb{Z}/n\mathbb{Z}$ bevatten evenveel elementen en we concluderen dat f bijectief is. Dus f is een isomorfisme.

Stel nu dat x oneindige orde heeft. Op dezelfde manier als net zien we dat $f : \mathbb{Z} \to G : k \mapsto x^k$ een homomorfisme is. We weten van het lemma dat alle machten x^k verschillend zijn, en omdat bovendien G geheel opgebouwd is uit deze machten, is f bijectief. Dus f is een isomorfisme. We hebben nu bewezen:

Lemma 6.2.4. *Zij G een cyclische groep. Als G eindige orde n heeft, dan is $G \cong \mathbb{Z}/n\mathbb{Z}$. Als G oneindig is, dan is $G \cong \mathbb{Z}$.*

Dit lemma komt straks goed van pas bij de bestudering van groepen die met de Fermat-vergelijking te maken hebben.

6.2.3 Vermenigvuldigen modulo n en de Kleine stelling van Fermat

Veel diepere eigenschappen van $\mathbb{Z}/n\mathbb{Z}$ komen pas aan het licht als we ook vermenigvuldiging erbij halen. Behalve voor praktische toepassingen zoals het berekenen van $2^{2^7}+1 \pmod{19}$, is de ringstructuur ook theoretisch belangrijk. Voor veel getallen n verschilt de ringstructuur in zekere zin veel van die van $\mathbb{Z}$. Zo zijn bijvoorbeeld -1 en 1 de enige inverteerbare elementen (eenheden) van $\mathbb{Z}$, ofwel $\mathbb{Z}^* = \{\pm 1\}$. In de ring

$$\mathbb{Z}/12\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}, \overline{6}, \overline{7}, \overline{8}, \overline{9}, \overline{10}, \overline{11}\}$$

is dat echter niet zo. Behalve $\overline{1}$ en $\overline{11} = \overline{-1}$ zijn bijvoorbeeld ook $\overline{5}$ en $\overline{7}$ inverteerbaar, want $\overline{5} \cdot \overline{5} = \overline{25} = \overline{1}$ en $\overline{7} \cdot \overline{7} = \overline{49} = \overline{1}$. Merk op dat $1, 11, 5, 7$ precies de getallen tussen 0 en 12 zijn die copriem zijn met 12. De andere elementen kunnen niet inverteerbaar zijn. Voor hen is er namelijk een element ongelijk aan nul waarmee het product nul is, want

$$\overline{2} \cdot \overline{6} = \overline{0}, \qquad \overline{3} \cdot \overline{4} = \overline{0}, \qquad \overline{8} \cdot \overline{9} = \overline{0}, \qquad \overline{10} \cdot \overline{6} = \overline{0}.$$

Dit is ook een belangrijk punt waarin $\mathbb{Z}$ en $\mathbb{Z}/n\mathbb{Z}$ verschillen: in $\mathbb{Z}$ kan $ab = 0$ alleen gelden als $a = 0$ of $b = 0$. Een element $a \neq 0$ van een ring R heet een *nuldeler* als er een $b \in R$ ongelijk

aan 0 is zodat $ab = 0$. In ons geval zijn $\overline{2}, \overline{3}, \overline{4}, \overline{6}, \overline{8}, \overline{9}, \overline{10}$ de nuldelers. Als $x \in R$ inverteerbaar is, dan volgt uit $xb = 0$ dat $b = x^{-1}xb = x^{-1} \cdot 0 = 0$, dus x is geen nuldeler. Nuldelers zijn dus niet inverteerbaar.

In een ring zonder nuldelers, zoals $\mathbb{Z}$, volgt uit $ab = 0$ dat $a = 0$ of $b = 0$. Hierdoor kunnen we in een vergelijking 'aan beide kanten door hetzelfde getal delen': als $px = qx$, dan is $x = 0$ of $p = q$. Uit $px = qx$ volgt namelijk dat $(p - q)x = 0$, dus $x = 0$ of $p - q = 0$. In bijvoorbeeld $\mathbb{Z}/12\mathbb{Z}$ geldt dit niet. De vergelijking $\overline{4x} = \overline{8x}$ heeft bijvoorbeeld de oplossingen $x = \overline{0}, \overline{3}, \overline{6}, \overline{9}$, dus uit $\overline{4x} = \overline{8x}$ volgt echt niet $\overline{4} = \overline{8}$. Ook een lineaire vergelijking als $\overline{9x} = \overline{6}$ kan meerdere oplossingen hebben, in dit geval $x = \overline{2}, \overline{6}, \overline{10}$.

Er is echter een bepaalde klasse van ringen $\mathbb{Z}/n\mathbb{Z}$ zonder nuldelers, namelijk die waarvoor n priem is. Sterker nog, dit zijn *lichamen*. Dit is een direct gevolg van de volgende stelling, die gemotiveerd wordt door wat we bij $n = 12$ hebben gezien.

Stelling 6.2.5. *De eenhedengroep van $\mathbb{Z}/n\mathbb{Z}$ bestaat precies uit de restklassen $\overline{a}$ waarvoor* $\gcd(a, n) = 1$. *In formule:*

$$(\mathbb{Z}/n\mathbb{Z})^* = \{\overline{a} \in \mathbb{Z}/n\mathbb{Z} : \gcd(a, n) = 1\}. \tag{6.8}$$

De orde van $(\mathbb{Z}/n\mathbb{Z})^*$ is dus $\phi(n)$, en $\mathbb{Z}/n\mathbb{Z}$ is een lichaam dan en slechts dan als n priem is.

Bewijs. Allereerst merken we voor de volledigheid op dat $\gcd(a, n)$ niet afhangt van de representant a van $\overline{a}$: elk element van $\overline{a}$ is van de vorm $a + qn$, en de gemeenschappelijke delers van a en n zijn precies die van $a + qn$ en n.

Als een restklasse $\overline{a}$ inverteerbaar is, dan is er een $\overline{x}$ zodat $\overline{ax} = \overline{1}$, en dus $ax \equiv 1 \pmod{n}$. Er is dus een $y \in \mathbb{Z}$ zodat $ax = 1 + (-y)n$, ofwel $ax + ny = 1$. Omgekeerd, als er een $y \in \mathbb{Z}$ is zodat $ax + ny = 1$, dan is $\overline{ax} = \overline{1 - ny} = \overline{1}$, dus $\overline{a}$ is inverteerbaar. We concluderen dat de volgende beweringen equivalent zijn:

1). $\overline{a} \in (\mathbb{Z}/n\mathbb{Z})^*$

2). Er zijn $x, y \in \mathbb{Z}$ zodat $ax + ny = 1$.

We gaan nu het *Euclidisch algoritme* gebruiken om te laten zien dat $\gcd(a, n)$ altijd van de vorm $ax + ny$ is, ofwel, een lineaire combinatie is van a en n. Zij $b_1, b_2 \in \mathbb{Z}$ met $b_2 \neq 0$, en zij $d = \gcd(b_1, b_2)$. We kunnen restdeling van b_1 door b_2 uitvoeren: er zijn $q_1, b_3 \in \mathbb{Z}$ met $0 \leq b_2 < b_3$ zodat $b_1 = q_1 b_2 + b_3$. Als $b_3 \neq 0$ kunnen we ook weer b_2 door b_3 delen zodat we $b_2 = q_2 b_3 + b_4$ krijgen. Dit kunnen we voortzetten tot we een rest 0 tegenkomen: zolang $b_{k+1} \neq 0$ kunnen we q_k, b_{k+2} vinden zodat $b_k = q_k b_{k+1} + b_{k+2}$. We krijgen zo een rij positieve resten $b_3 > b_4 > b_5 > \ldots$, en omdat er geen oneindig afdalende rij natuurlijke getallen bestaat, houdt dit proces een keer op: we stuiten een keer op een rest 0, zeg $b_{t+1} = 0$. We hebben dus een rij vergelijkingen van de vorm

$$b_1 = q_1 b_2 + b_3, \qquad b_2 = q_2 b_3 + b_4, \qquad b_3 = q_3 b_4 + b_5, \qquad \ldots, \qquad b_{t-1} = q_{t-1} b_t.$$

We beweren dat $\gcd(b_1, b_2) = b_t$. Als $d|b_1$ en $d|b_2$, dan volgt uit bovenstaande vergelijkingen dat ook $d|b_1 - q_1 b_2 = b_3$, en zo doorgaande zien we $d|b_4$, $d|b_5$, $\ldots$, $d|b_t$. Omgekeerd is elke deler van b_t deler van $q_{t-1} b_t = b_{t-1}$, en dus ook van $q_{t-2} b_{t-1} + b_t = b_{t-3}$, en dus ook van b_{t-4},

b_{t-5}, ..., b_2, b_1. De gemeenschappelijke delers van b_1 en b_2 zijn dus precies de delers van b_t, dus in het bijzonder is $\text{ggd}(b_1, b_2) = b_t$. Omdat $b_1 = 1b_1 + 0b_2$ en $b_2 = 0b_1 + 1b_2$, zijn b_1 en b_2 lineaire combinaties van b_1 en b_2. Als $b_1, b_2, \ldots, b_{k-1}, b_k$ lineaire combinaties zijn van b_1 en b_2, zeg $b_{k-1} = x_1b_1 + y_1b_2$ en $b_k = x_2b_1 + y_2b_2$ dan is

$$b_{k+1} = b_{k-1} - q_{k-1}b_k = x_1b_1 + y_1b_2 - q_{k-1}(x_2b_1 + y_2b_2) = (x_1 - q_{k-1}x_2)b_1 + (y_1 - q_{k-1}y_2)b_2,$$

dus ook b_{k+1} is een lineaire combinatie van b_1 en b_2. Met inductie concluderen we dat $b_t = \text{ggd}(b_1, b_2)$ een lineaire combinatie is van b_1, b_2.

Terugkerend naar ons geval, er zijn dus $x, y \in \mathbb{Z}$ zodat $\text{ggd}(a, n) = ax+ny$, en als $\text{ggd}(a, n) = 1$ dan is dus aan 2) voldaan. Omgekeerd, als er $x, y \in \mathbb{Z}$ zijn zodat $ax + ny = 1$, dan geldt voor elke gemene deler d van a en n dat $d|ax + ny = 1$, dus a en n zijn copriem. We concluderen dat $\text{ggd}(a, n) = 1$ equivalent is aan 2), en dus aan 1). Hiermee is (6.8) bewezen.

De orde van $(\mathbb{Z}/n\mathbb{Z})^*$ is dus het aantal getallen $a \in M = \{0, 1, 2, \ldots, n-1\}$ waarvoor $\text{ggd}(a, n) = 1$, en dat zijn er precies $\phi(n)$. Per definitie is $\mathbb{Z}/n\mathbb{Z}$ een lichaam dan en slechts dan als $(\mathbb{Z}/n\mathbb{Z})^* = \mathbb{Z}/n\mathbb{Z} - \{0\}$, ofwel, 0 is het enige element van M waarvoor $\text{ggd}(a, n) \neq 1$. Hieraan is voldaan precies dan als n priem is. □

Een belangrijk gevolg is de volgende formule van Euler.

Gevolg 6.2.6. (Euler) *Zij n een natuurlijk getal en a een geheel getal copriem met n. Dan is*

$$a^{\phi(n)} \equiv 1 \pmod{n}.$$

Opmerking: het is illustratief om dit voor een paar specifieke a en n te verifiëren.

Bewijs. Omdat $\text{ggd}(a, n) = 1$, is $\overline{a}$ bevat in de eenhedengroep $(\mathbb{Z}/n\mathbb{Z})^*$. De orde van deze groep is $\phi(n)$, en de orde m van $\overline{a}$ is een deler van de groepsorde, zeg $mk = \phi(n)$. Er geldt dus

$$\overline{a}^{\phi(n)} = (\overline{a}^m)^k = \overline{1}^k = \overline{1},$$

en dit is equivalent met Euler's formule. □

De formule van Euler is een generalisatie van de 'Kleine stelling van Fermat'.

Gevolg 6.2.7. Kleine stelling van Fermat. *Zij p een priemgetal, en a een willekeurig geheel getal. Dan is*

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

Bewijs. Als p deler is van a, ofwel $a \equiv 0 \pmod{p}$, dan is het duidelijk, want dan staat er $0^p \equiv 0 \pmod{p}$. In het andere geval volgt omdat p priem is dat $\text{ggd}(a, p) = 1$. Bovendien is $\phi(p) = p - 1$, dus uit de formule van Euler volgt dat $a^{p-1} \equiv 1 \pmod{p}$. De stelling volgt nu door aan beide kanten met a te vermenigvuldigen. □

Als p priem is, dan is $\mathbb{Z}/p\mathbb{Z}$ een lichaam, en elk element ongelijk $\overline{0}$ heeft dus een inverse. In bijvoorbeeld $\mathbb{Z}/13\mathbb{Z}$ geldt

$$\overline{1} \cdot \overline{1} = \overline{1}, \qquad \overline{2} \cdot \overline{7} = \overline{13+1} = \overline{1}, \qquad \overline{3} \cdot \overline{9} = \overline{26+1} = \overline{1}, \qquad \overline{4} \cdot \overline{10} = \overline{39+1} = \overline{1},$$
$$\overline{5} \cdot \overline{8} = \overline{39+1} = \overline{1}, \qquad \overline{6} \cdot \overline{11} = \overline{65+1} = \overline{1}, \qquad \overline{12} \cdot \overline{12} = \overline{-1} \cdot \overline{-1} = \overline{1}.$$

Voor kleine p zoals 13 kunnen we de inversen nog door bepalen door gewoon te proberen, voor grote p kost dat zelfs met een computer erg veel rekentijd. Het bewijs van Stelling 6.2.5 geeft gelukkig meteen een efficiënte methode om inversen te berekenen. We kunnen namelijk het Euclidisch algoritme toepassen op a en p om getallen x, y te vinden zodat $ax + py = \text{ggd}(a, p) = 1$. Omdat $\overline{1} = \overline{ax + py} = \overline{ax} = \overline{a} \cdot \overline{x}$, is $\overline{x}$ de inverse van $\overline{a}$.

6.2.4 Fermat's laatste stelling voor geval 1 van $n = 5$

Een mooie toepassing van de Kleine stelling van Fermat is dat we heel makkelijk zijn Laatste stelling kunnen bewijzen voor een speciaal geval van $n = 5$, namelijk het geval waarin 5 geen van de getallen x, y, z deelt.[10] Het blijkt dat veel bewijzen van FLT voor specifieke (verzamelingen van) exponenten n op een natuurlijke manier in twee gevallen uiteenvallen, die traditioneel geval 1 en geval 2 worden genoemd. In geval 1 deelt n géén van de x, y, z, het andere geval heet geval 2. We bewijzen dus geval 1 van FLT voor $n = 5$. Als x, y, z oplossing is van $x^5 + y^5 = z^5$, dan is $x, y, -z$ oplossing van $x^5 + y^5 + z^5 = 0$, want 5 is oneven. We kunnen ons dus net zo goed op die tweede, meer symmetrische vergelijking richten: als die geen oplossing heeft, heeft de eerste er ook geen.

Stelling 6.2.8. *Er zijn geen gehele getallen x, y, z die geen van allen deelbaar zijn door 5, en die voldoen aan $x^5 + y^5 + z^5 = 0$.*

Bewijs. We bewijzen dit uit het ongerijmde: stel er zijn wél x, y, z zoals in het bovenstaande. Uit de Kleine stelling van Fermat volgt dat

$$x^5 \equiv x, \qquad y^5 \equiv y, \qquad z^5 \equiv z \pmod 5.$$

Dus $0 \equiv x^5 + y^5 + z^5 \equiv x + y + z \pmod 5$. Omdat 5 geen van de x, y, z deelt, zijn ze niet congruent 0 modulo 5. De enige mogelijkheden voor x, y, z modulo 5 zijn dus $\pm 1, \pm 2$. Als ze allen verschillend zouden zijn modulo 5, dan zouden twee van hen, zonder beperking kunnen we aannemen dat het y en z zijn, tegengesteld zijn modulo 5 (dat wil zeggen ± 1 of ± 2). Dus $0 \equiv x + y + z \equiv x + y - y \equiv x \pmod 5$, tegenspraak want 5 deelt niet x. We concluderen dat twee van hen congruent zijn modulo 5, laten we zeggen dat het x en y zijn (wegens de symmetrie is dat geen beperking). Er geldt dus dat $5|(x - y)$, en we beweren dat $25|(x^5 - y^5)$. Er geldt namelijk dat

$$x^5 - y^5 = (x - y)(x^4 + x^3y + x^2y^2 + xy^3 + y^4), \tag{6.9}$$

en

$$x^4 + x^3y + x^2y^2 + xy^3 + y^4 \equiv x^4 + x^3x + x^2x^2 + xx^3 + x^4 \equiv 5x^4 \equiv 0 \pmod 5.$$

Beide factoren in (6.9) zijn dus deelbaar door 5, dus het product $x^5 - y^5$ is deelbaar door 25, dus inderdaad $x^5 \equiv y^5 \pmod{25}$. Hieruit volgt dat

$$-z^5 \equiv x^5 + y^5 \equiv 2x^5 \pmod{25}.$$

[10] De essentie van het bewijs vonden we op [11].

Maar we weten ook dat $-z \equiv x + y \equiv 2x \pmod 5$, dus 5 deelt $-z - 2x$. Op dezelfde manier als net volgt hieruit dat 25 deler is van $(-z)^5 - (2x)^5$, dus

$$-z^5 \equiv 32x^5 \pmod{25}.$$

We concluderen dat $2x^5 \equiv 32x^5 \pmod{25}$. Stel dat $\gcd(x^5, 25) \neq 1$. De delers van 25 zijn, behalve 1, veelvoud van 5, dus er zou volgen dat 5 deler is van x^5 en dus ook van x, tegenspraak. We concluderen dat x^5 en 25 copriem zijn, dus x^5 is een eenheid in $\mathbb{Z}/25\mathbb{Z}$. Door met zijn inverse te vermenigvuldigen, volgt dus uit $2x^5 \equiv 32x^5 \pmod{25}$ dat $2 \equiv 32 \pmod{25}$, tegenspraak. □

6.2.5 De structuur van de eenhedengroep

We weten van Stelling 6.2.2 aardig wat van de structuur van de optelgroep $\mathbb{Z}/n\mathbb{Z}$, maar nog niet veel van die van de vermenigvuldiggroep $(\mathbb{Z}/n\mathbb{Z})^*$. Wat zijn bijvoorbeeld de ondergroepen? Als voorbeeld nemen we eerst een getal met relatief veel delers, namelijk $n = 30$. De groep $(\mathbb{Z}/30\mathbb{Z})^*$ heeft orde $\phi(30) = 8$:

$$(\mathbb{Z}/30\mathbb{Z})^* = \{\overline{1}, \overline{7}, \overline{11}, \overline{13}, \overline{17}, \overline{19}, \overline{23}, \overline{29}\}.$$

Merk op dat de tegengestelde van elk element ook weer in de groep zit: $\overline{1} = \overline{-29}, \overline{7} = \overline{-23}, \overline{11} = \overline{-19}$ en $\overline{13} = \overline{-17}$. Het is echter geen groep onder optelling, het bevat bijvoorbeeld niet $\overline{3} = \overline{1} + \overline{1} + \overline{1}$.

Laten we kijken wat de orde van elk element $\overline{a}$ is,[11] de kleinste exponent $m \geq 1$ waarvoor $\overline{a}^m = \overline{1}$. We weten in elk geval dat de orde van elk element de groepsorde 8 deelt, en dus gelijk is aan $1, 2, 4$ of 8. De orde van $\overline{1}$ is uiteraard 1, en die van $\overline{29} = \overline{-1}$ is 2. Omdat $\overline{11^2} = \overline{120 + 1} = \overline{1}$, heeft $\overline{11}$ orde 2. Hieruit volgt meteen dat ook $\overline{19} = \overline{-11}$ orde 2 heeft, want $(-11)^2 \equiv 11^2 \equiv 1 \pmod{30}$. De orde van $\overline{7}$ is 4, want

$$7^2 \equiv 49 \equiv 19 \equiv -11, \qquad 7^3 \equiv 7 \cdot -11 \equiv -77 \equiv 13, \qquad 7^4 \equiv (-11)^2 \equiv 1 \pmod{30}.$$

Dus ook $\overline{23} = \overline{-7}$ heeft orde 4: er geldt namelijk

$$(-7)^2 \equiv 7^2 \equiv -11, \qquad (-7)^3 \equiv -7^3 \equiv -13, \qquad (-7)^4 \equiv 7^4 \equiv 1 \pmod{30}.$$

We hebben dus in elk geval al twee ondergroepen van orde 4 gevonden, namelijk die bestaande uit de machten van 7 respectievelijk -7. Als we de ondergroep $\{\overline{1}, \overline{7}, \overline{19}, \overline{13}\} = \{\overline{7^0}, \overline{7^1}, \overline{7^2}, \overline{7^3}\}$ van machten van $\overline{7}$ als uitgangspunt nemen, dan hebben we

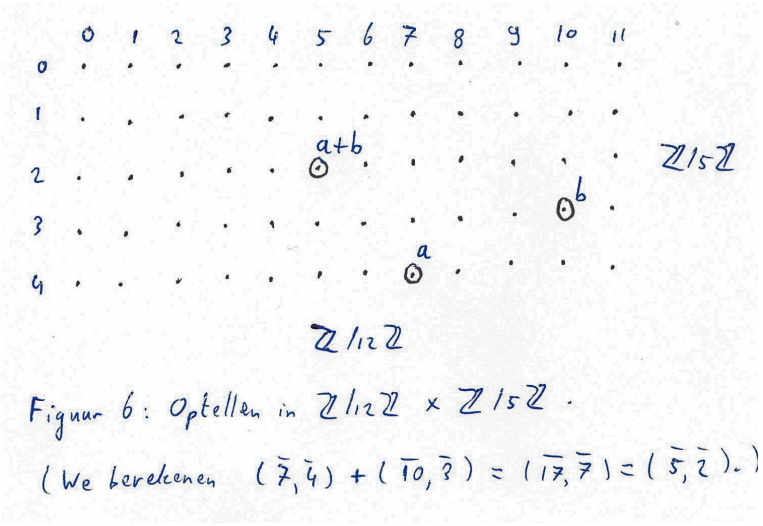
$$\begin{aligned} (\mathbb{Z}/30\mathbb{Z})^* &= \{\overline{1}, \quad \overline{7}, \quad \overline{19}, \quad \overline{13}, \quad \overline{29}, \quad \overline{23}, \quad \overline{11}, \quad \overline{17}\} \\ &= \{\overline{7^0}, \quad \overline{7^1}, \quad \overline{7^2}, \quad \overline{7^3}, \quad -\overline{7^0}, \quad -\overline{7^1}, \quad -\overline{7^2}, \quad -\overline{7^3}\}. \end{aligned} \tag{6.10}$$

Nu is rekenen in deze groep veel makkelijker geworden: we hoeven alleen de regel $7^k \cdot 7^l = 7^{k+l}$ te kennen, samen met de gebruikelijke regels van vermenigvuldiging met mintekens. We kunnen de groep nog verder 'ontbinden', namelijk in een deel waarin we met machten van 7 rekenen, en een deel waarin we met machten van -1 rekenen. Deze delen werken onafhankelijk van elkaar. Dit idee van 'ontbinden' wordt hardgemaakt door het begrip direct product.

[11] Eigenlijk moeten we onderscheid maken tussen de 'additieve orde' en de 'multiplicatieve orde' van elementen in de ring $\mathbb{Z}/n\mathbb{Z}$, maar omdat we ons nu alleen op $(\mathbb{Z}/n\mathbb{Z})^*$ richten is het duidelijk dat we die tweede bedoelen.

Definitie 6.2.9. Direct product. *Zij G, H twee groepen. Het directe product van G en H, notatie $G \times H$, is als verzameling het cartesisch product $\{(g,h) : g \in G, h \in H\}$. We definiëren hierop coördinaatsgewijs een bewerking $(g_1,h_1)(g_2,h_2) = (g_1g_2, h_1h_2)$, waarmee $G \times H$ een groep wordt.*

Het is makkelijk na te gaan dat $G \times H$ inderdaad aan de groepsaxioma's voldoet: associativiteit volgt bijna direct uit die van G en H, het eenheidselement is (e,e') en de inverse van (g,h) is (g^{-1},h^{-1}). In elk geval voor eindige G en H kunnen we ons $G \times H$ meetkundig voorstellen als een rooster in het vlak met een G-as en een H-as, waarin we de vectoren 'coördinaatsgewijs vermenigvuldigen' via de vermenigvuldiging op de G-as en de H-as. Het G-deel en het H-deel zijn dus geheel van elkaar onafhankelijk.[12] Zie Figuur 6 ter illustratie, waar we $G = \mathbb{Z}/12\mathbb{Z}$ en $H = \mathbb{Z}/5\mathbb{Z}$ hebben genomen. Omdat we modulorekenen kunnen

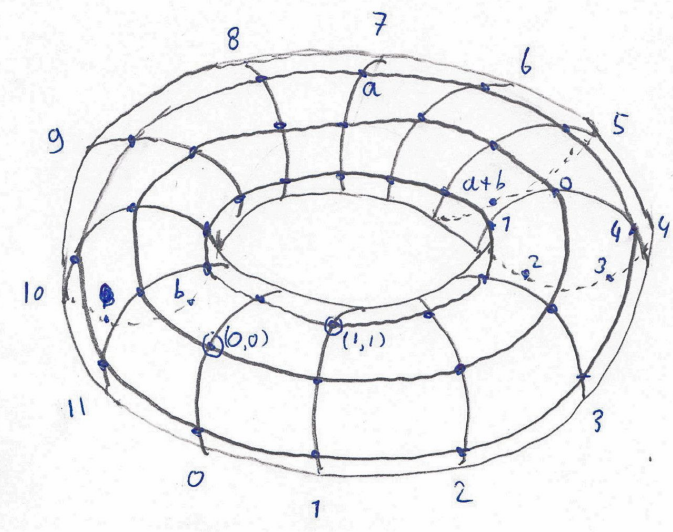


Figuur 6: Optellen in $\mathbb{Z}/12\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$.

(We berekenen $(\bar{7},\bar{4}) + (\bar{10},\bar{3}) = (\bar{17},\bar{7}) = (\bar{5},\bar{2})$.)

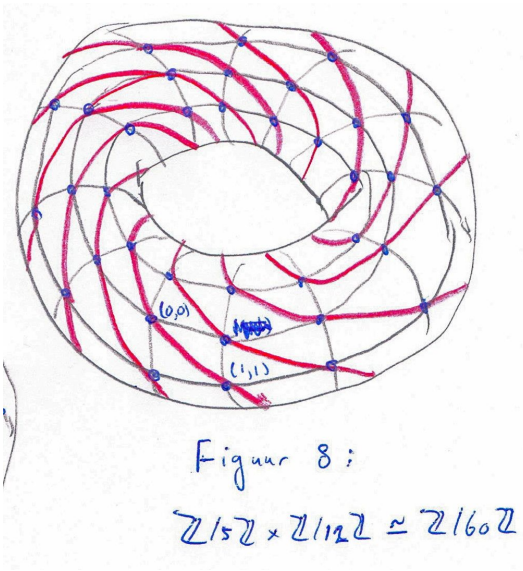
zien als rekenen op cirkels, is het natuurlijk om in het plaatje de G-as en de H-as op te rollen, zodat we een soort torus van roosterpunten krijgen. Rekenen in $\mathbb{Z}/12\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$ kunnen we zien als punten optellen op deze torus, waarbij we bij elk van beide 'cirkelrichtingen' een nulpunt hebben gekozen. Zie figuur 7. Merk op dat de veelvouden van $(1,1)$, dat zijn de punten $(1,1),(2,2),(3,3),(4,4),(5,0),(6,1),\ldots$, heel de torus doorlopen. Het lijkt wel alsof de roosterpunten van de torus allemaal op één lang, aaneengesloten touw (of eigenlijk lus) liggen die rond de torus is gewikkeld, zie Figuur 8 ter illustratie (de lus is roodgekleurd). Voor bijvoorbeeld $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/9\mathbb{Z}$ geldt dit niet: al na 9 stappen komt de rij $(1,1),(2,2)\ldots$, uit bij $(0,0)$. De torus lijkt nu te zijn opgebouwd uit drie in elkaar verstrengelde lussen, zie Figuur 9. Dit verschijnsel heeft te maken met het feit dat 5 en 12, in tegenstelling tot 3 en 9, copriem zijn. Het kleinste gemene veelvoud van 5 en 12 is daarom gelijk aan $5 \cdot 12 = 60$. Dus $(k,k) = (0,0)$ precies dan als k veelvoud is van 5 en van 12, ofwel, veelvoud van 60. De orde van $(1,1)$ is dus 60, dus de groep is cyclisch: $\mathbb{Z}/12\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z} \cong \mathbb{Z}/60\mathbb{Z}$. Op precies dezelfde manier volgt dat $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/mn\mathbb{Z}$ voor alle $m,n \in \mathbb{N}$ die copriem

[12] In sommige opzichten kan het geheel echter groter zijn dan de som der delen. Denk bijvoorbeeld aan de (topologische) rijkheid van $\mathbb{R} \times \mathbb{R}$ ten opzichte van $\mathbb{R}$.



Figuur 7:
$\mathbb{Z}/12\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$
afgebeeld op de torus,
met a, b, $a+b$ zoals
in Figuur 6.

zijn. Ze zijn zelfs als ring hetzelfde; we gaan hier echter niet verder op in omdat we het niet over ringisomorfismen hebben gehad. Dit resultaat staat bekend als de Chinese reststelling.



Figuur 8:
$\mathbb{Z}/5\mathbb{Z} \times \mathbb{Z}/12\mathbb{Z} \simeq \mathbb{Z}/60\mathbb{Z}$

Een heel ander voorbeeld is $\mathbb{C}^*$: deze is isomorf met het directe product van de 'cirkelgroep' en de 'lijnengroep'. In formule: $\mathbb{C}^* \cong \mathbb{C}^*/C \times \mathbb{C}^*/L$ met C en L zoals in het begin van dit hoofdstuk. Dit komt doordat een complex getal (behalve nul) geheel wordt vastgelegd door zijn absolute waarde en zijn argument, en bij het vermenigvuldigen van twee getallen zijn deze twee onafhankelijk van elkaar: de absolute waarden worden vermenigvuldigd, de argumenten worden opgeteld. Een expliciet isomorfisme is $f : \mathbb{C}^* \to \mathbb{C}^*/C \times \mathbb{C}^*/L : f(\alpha) = (\alpha C, \alpha L)$. Stel $(\alpha C, \beta L) \in \mathbb{C}^*/C \times \mathbb{C}^*/L$. Nemen we het complexe getal γ met absolute waarde $|\alpha|$ en argument $\arg(\beta)$, dan is $f(\gamma) = (\gamma C, \gamma L) = (\alpha C, \beta L)$, dus f is surjectief. Als $f(\alpha) = f(\beta)$, dan volgt uit $\alpha C = \beta C$ dat ze dezelfde absolute waarde hebben, en uit $\alpha L = \beta L$ dat ze hetzelfde argument hebben. Dus $\alpha = \beta$, dus f is ook injectief. Dat het een homomorfisme is, volgt bijna direct uit de definities:

$$f(\alpha\beta) = (\alpha\beta C, \alpha\beta L) = (\alpha C \cdot \beta C, \alpha L \cdot \beta L) = (\alpha C, \alpha L)(\beta C, \beta L) = f(\alpha)f(\beta).$$

Door herhaald toepassen kunnen we voor groepen $G_1, \ldots G_n$ het directe product $G_1 \times G_2 \times \cdots \times G_n$ definieëren. Gelukkig hoeven we geen haakjes te zetten: uitschrijven leert dat $(G \times H) \times J \cong G \times (H \times J)$, met $((g,h),j) \mapsto (g,(h,j))$ als isomorfisme. Bovendien maakt de volgorde niet

uit: $G \times H \cong H \times G$, want $(g,h) \mapsto (h,g)$ is een isomorfisme. We kijken hoe we deze theorie op formule (6.10) voor $(\mathbb{Z}/30\mathbb{Z})^*$ kunnen toepassen. We vermoeden dat de groep isomorf is met $G \times H := \{1,-1\} \times \{\overline{7^0}, \overline{7^1}, \overline{7^2}, \overline{7^3}\}$. De functie

$$f : G \times H \to (\mathbb{Z}/30\mathbb{Z})^* = \{\overline{7^0}, \overline{7^1}, \overline{7^2}, \overline{7^3}, -\overline{7^0}, -\overline{7^1}, -\overline{7^2}, -\overline{7^3}\} : (a,b) \mapsto ab$$

is duidelijk een bijectie. Bovendien is het een homomorfisme, want omdat vermenigvuldiging commutatief is, geldt

$$f((a,b)(c,d)) = f(ac,bd) = acbd = abcd = f(a,b)f(c,d). \tag{6.11}$$

Dus f is een isomorfisme, en we zien dat $(\mathbb{Z}/30\mathbb{Z})^* \cong G \times H$. Omdat G priemorde 2 heeft is hij cyclisch en dus isomorf met $\mathbb{Z}/2\mathbb{Z}$. Verder is H duidelijk cyclisch van orde 4 met voortbrenger $\overline{7}$, en dus isomorf met $\mathbb{Z}/4\mathbb{Z}$. We concluderen:

$$(\mathbb{Z}/30\mathbb{Z})^* \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}.$$

Hetzelfde kunnen we proberen voor andere waarden van n. Zo vinden we voor $n = 20$ weer een groep van orde $8 = \phi(20)$, die bovendien weer isomorf is met $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$:

$$\begin{aligned}(\mathbb{Z}/20\mathbb{Z})^* &= \{\overline{1}, \quad \overline{3}, \quad \overline{9}, \quad \overline{7}, \quad \overline{19}, \quad \overline{17}, \quad \overline{11}, \quad \overline{13}\} \\ &= \{\overline{3^0}, \quad \overline{3^1}, \quad \overline{3^2}, \quad \overline{3^3}, \quad -\overline{3^0}, \quad -\overline{3^1}, \quad -\overline{3^2}, \quad -\overline{3^3}\} \\ &\cong \{1,-1\} \times \{\overline{3^0}, \overline{3^1}, \overline{3^2}, \overline{3^3}\} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}.\end{aligned}$$

Maar $(\mathbb{Z}/24\mathbb{Z})^*$, ook een groep van orde 8, is anders. We hebben

$$(\mathbb{Z}/24\mathbb{Z})^* = \{\overline{1}, \overline{5}, \overline{7}, \overline{11}, \overline{13}, \overline{17}, \overline{19}, \overline{23}\}.$$

Behalve $\overline{1}$ heeft elk element orde 2. Namelijk,

$$5^2 \equiv 25 \equiv 1, \qquad 7^2 \equiv 49 \equiv 1, \qquad 11^2 \equiv 121 \equiv 1 \pmod{24},$$

en omdat $\overline{23} = \overline{-1}, \overline{19} = \overline{-5}, \overline{17} = \overline{-7}$ en $\overline{13} = \overline{-11}$ hebben ook de overige 4 elementen orde 2. We herkennen een ondergroep van orde 4, namelijk $H = \{\overline{1}, \overline{5}, \overline{7}, \overline{11}\}$. Deze verzameling bevat $\overline{1}$ en elk element is zijn eigen inverse, dus we hoeven alleen nog te controleren dat hij gesloten is onder vermenigvuldiging. Producten met $\overline{1}$ erin hoeven we niet te controleren, net als de kwadraten, dus we hoeven maar drie gevallen te controleren:

$$5 \cdot 7 \equiv 35 \equiv 11, \qquad 5 \cdot 11 \equiv 55 \equiv 7, \qquad 7 \cdot 11 \equiv 77 \equiv 5 \pmod{24}.$$

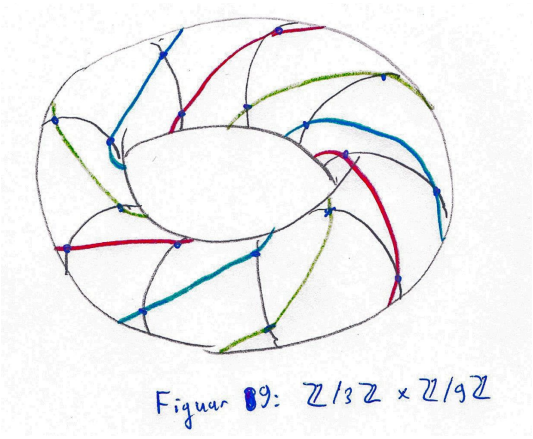
Dus H is een ondergroep. Deze heeft weer onder meer de ondergroepen $\{\overline{1}, \overline{5}\}$ en $\{\overline{1}, \overline{7}\}$, beide van priemorde 2 en dus isomorf met $\mathbb{Z}/2\mathbb{Z}$. Het is makkelijk na te gaan dat $f : \{\overline{1}, \overline{5}\} \times \{\overline{1}, \overline{7}\} : (a,b) \mapsto ab$ een bijectie is, en zoals in (6.11) volgt dat het ook een homomorfisme is. We hebben dus $H \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, en we concluderen dat

$$\begin{aligned}(\mathbb{Z}/24\mathbb{Z})^* &= \{\overline{1}, \quad \overline{5}, \quad \overline{7}, \quad \overline{11}, \quad -\overline{1}, \quad -\overline{5}, \quad -\overline{7}, \quad -\overline{11}\} \\ &\cong \{1,-1\} \times \{\overline{1}, \overline{5}, \overline{7}, \overline{11}\} \cong \mathbb{Z}/2\mathbb{Z} \times H \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}.\end{aligned}$$

Zo kunnen we nog een tijdje doorgaan. We vinden bijvoorbeeld op dezelfde manier als bij $n = 30$ en $n = 24$ dat $(\mathbb{Z}/36\mathbb{Z})^* \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$. En $(\mathbb{Z}/26\mathbb{Z})^*$, ook een groep van orde 12, is isomorf met

$$\begin{aligned} & \{\overline{1}, \quad \overline{5}, \quad \overline{25}, \quad \overline{21}, \quad \overline{3}, \quad \overline{15}, \quad \overline{23}, \quad \overline{11}, \quad \overline{9}, \quad \overline{19}, \quad \overline{17}, \quad \overline{7}\} \\ = \; & \{\overline{3^0 5^0}, \quad \overline{3^0 5^1}, \quad \overline{3^0 5^2}, \quad \overline{3^0 5^3}, \quad \overline{3^1 5^0}, \quad \overline{3^1 5^1}, \quad \overline{3^1 5^2}, \quad \overline{3^1 5^3}, \quad \overline{3^2 5^0}, \quad \overline{3^2 5^1}, \quad \overline{3^2 5^2}, \quad \overline{3^2 5^3}\ \} \\ \cong \; & \{\overline{3^0}, \overline{3^1}, \overline{3^2}\} \times \{\overline{5^0}, \overline{5^1}, \overline{5^2}, \overline{5^3}\} \;\cong\; \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}. \end{aligned}$$

Er lijkt op het eerste gezicht niet veel systeem te zitten in de groepsstructuren, maar dat is schijn. Het blijkt namelijk zo te zijn dat $\mathbb{Z}/mn\mathbb{Z}^* \cong \mathbb{Z}/m\mathbb{Z}^* \times \mathbb{Z}/n\mathbb{Z}^*$ als m, n copriem zijn.[13] We schetsen een bewijs, de lezer mag de details invullen. Voor willekeurige groepen G, H met $G \cong H$ is ook $G^* \cong H^*$: als $f : G \to H$ een isomorfisme is, dan is de restrictie van f tot G^* een isomorfisme $G^* \to H^*$ (in het bijzonder moet je dus nagaan dat $f(u)$ een eenheid is als u dat is). Verder, voor willekeurige G, H is $(G \times H)^* \cong G^* \times H^*$: ze zijn zelfs gelijk aan elkaar, want (x, y) is een eenheid precies dan als x en y dat zijn. Dus in ons geval is


Figuur 89: $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/9\mathbb{Z}$

$$\mathbb{Z}/mn\mathbb{Z}^* \cong (\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z})^* \cong \mathbb{Z}/m\mathbb{Z}^* \times \mathbb{Z}/n\mathbb{Z}^*;$$

het eerste isomorfisme volgt omdat m, n copriem zijn met behulp van de torus-discussie op pagina 47.

We hoeven daarom alleen nog de priemmachten te bestuderen: als n priemfactorontbinding $p_1^{k_1} \cdots p_t^{k_t}$ heeft, dan is

$$\mathbb{Z}/n\mathbb{Z}^* = \mathbb{Z}/(p_1^{k_1} \cdots p_t^{k_t})\mathbb{Z}^* \cong \mathbb{Z}/p_1^{k_1}\mathbb{Z}^* \times \cdots \times \mathbb{Z}/p_t^{k_t}\mathbb{Z}^*,$$

dus we zijn klaar als we de structuur van de $\mathbb{Z}/p^k\mathbb{Z}^*$ kennen voor p priem. Het blijkt dat deze een prachtige structuur hebben. Als voorbeeld beschouwen we $(\mathbb{Z}/13\mathbb{Z})^*$. We kunnen met wat rekenwerk van alle elementen de machten uitrekenen, of zelfs een vermenigvuldigtabel maken met 12^2 producten, maar dat is niet nodig. Na even rekenen zien we namelijk dat het element 2 orde 12 heeft, en dus heel de groep voortbrengt: de groep is *cyclisch*. Expliciet uitgeschreven:

$$\begin{aligned} (\mathbb{Z}/13\mathbb{Z})^* &= \{\overline{1}, \quad \overline{2}, \quad \overline{4}, \quad \overline{8}, \quad \overline{3}, \quad \overline{6}, \quad \overline{12}, \quad \overline{11}, \quad \overline{9}, \quad \overline{5}, \quad \overline{10}, \quad \overline{7}\} \\ &= \{\overline{2^0}, \quad \overline{2^1}, \quad \overline{2^2}, \quad \overline{2^3}, \quad \overline{2^4}, \quad \overline{2^5}, \quad \overline{2^6}, \quad \overline{2^7}, \quad \overline{2^8}, \quad \overline{2^9}, \quad \overline{2^{10}}, \quad \overline{2^{11}}\} \quad \cong \mathbb{Z}/12\mathbb{Z}. \end{aligned}$$

Nu is vermenigvuldigen in $(\mathbb{Z}/13\mathbb{Z})^*$ voor zover het dat nog niet was een fluitje van een cent: we lezen bijvoorbeeld gewoon af dat

$$11 \cdot 9 \equiv 2^7 \cdot 2^8 \equiv 2^{15} \equiv 2^3 \equiv 8 \pmod{13}.$$

[13] Met bijvoorbeeld $\mathbb{Z}/mn\mathbb{Z}^*$ bedoelen we $(\mathbb{Z}/mn\mathbb{Z})^*$, niet $\mathbb{Z}/mn(\mathbb{Z}^*)$.

Bovendien kunnen we meteen de ordes van alle elementen aflezen. Bijvoorbeeld $\overline{5}$ heeft orde 4, want

$$\overline{5} = \overline{2^9}, \qquad \overline{5}^2 = \overline{2^6} = \overline{12}, \qquad \overline{5}^3 = \overline{2^3} = \overline{8}, \qquad \overline{5}^4 = \overline{2^0} = \overline{1}.$$

De orde is dus 4; bovenstaande berekening kan worden samengevat door op te merken dat $\overline{9}$ orde 4 heeft in de *additieve* groep $\mathbb{Z}/12\mathbb{Z}$. Van Stelling 6.2.2 weten we dat er voor elke deler d van 12 precies één ondergroep van orde d is, namelijk die voortgebracht door $\overline{2^{12/d}}$, en er zijn $\phi(d)$ elementen van orde d. De elementen van orde 12 zijn bijvoorbeeld $\overline{2^1}, \overline{2^5}, \overline{2^7}, \overline{2^{11}}$ ofwel $\overline{2}, \overline{6}, \overline{11}, \overline{7}$. Evenzo hebben $\overline{4}$ en $\overline{10}$ orde zes, $\overline{8}$ en $\overline{5}$ orde vier, $\overline{3}$ en $\overline{9}$ orde drie, $\overline{12}$ orde twee en $\overline{1}$ orde één. Als $\mathbb{Z}/n\mathbb{Z}^*$ cyclisch is, en $\overline{q}$ een voortbrenger van deze groep (waarbij we $q \in \{1, 2, 3, \ldots, n-1\}$ kiezen), dan noemen we q een *primitieve wortel modulo* n. Dus de primitieve wortels modulo 13 zijn precies de $\phi(12)$ getallen $2, 6, 11$ en 7. De primitieve wortels x van n hebben de mooie eigenschap dat de getallen $x, x^2, x^3, \ldots, x^{p-1} \mod p$ op volgorde na precies de getallen $1, 2, 3, \ldots, n-1$ zijn, een eigenschap die vaak handig van pas komt. Het blijkt zo te zijn dat $\mathbb{Z}/p^k\mathbb{Z}^*$ cyklisch is voor *alle* priemgetallen $p > 2$ en natuurlijke getallen k, waarmee het eerder geopperde probleem behalve voor $p = 2$ is opgelost.[14] Dit zullen we niet in zijn algemeenheid bewijzen, maar wel een bijzonder geval dat we later nodig hebben, namelijk $k = 1$: we bewijzen dus dat $\mathbb{Z}/p\mathbb{Z}^* \cong \mathbb{Z}/(p-1)\mathbb{Z}$, zoals we net voor $p = 13$ hebben daan. Dit is een gevolg van Stelling 7.4.4, die we nu nog niet kunnen bewijzen. Elk priemgetal heeft dus in het bijzonder $\phi(p-1)$ primitieve wortels. In het geval $p = 17$ vinden we bijvoorbeeld dat 3 een primitieve wortel is. De $\phi(16) = 8$ primitieve wortels van 17 zijn dus $3, 3^3, 3^5, 3^7, 3^9, 3^{11}, 3^{13}, 3^{15} \mod 17$, ofwel $3, 10, 5, 11, 14, 7, 12, 6$.

[14] Voor $p = 2$ blijkt voor alle $k \geq 3$ dat $\mathbb{Z}/2^k\mathbb{Z}^* \cong \mathbb{Z}/2^{k-2}\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Dat $\mathbb{Z}/p^k\mathbb{Z}^*$ voor oneven p cyclisch is, kan worden begrepen doordat $\mathbb{Z}/p^k\mathbb{Z}$ een eindig lichaam is (dit zijn samen met $\mathbb{Z}/2\mathbb{Z}$ zelfs de enige eindige lichamen), en we bewijzen later dat de eindige ondergroepen van F^* cyclisch zijn voor lichamen F.

Hoofdstuk 7

Domeinen en idealen

We zagen in het vorige hoofdstuk dat $\mathbb{Z}/n\mathbb{Z}$ geen nuldelers heeft als n priem is. Commutatieve ringen zonder nuldelers zijn erg belangrijk, en hebben daarom een speciale naam gekregen.

Definitie 7.0.10. Domein. *Een* domein *is een commutatieve ring zonder nuldelers.*

In een domein impliceert $ab = 0$ dus dat $a = 0$ of $b = 0$. Het 'standaardvoorbeeld' van een domein is $\mathbb{Z}$. We merkten al op dat inverteerbare elementen geen nuldelers zijn, dus een lichaam heeft geen nuldelers (want 0, het enige niet-inverteerbare element, is per definitie geen nuldeler). Per definitie is een lichaam commutatief, en dus een domein.

Het rijtje *groep, ring, domein* komt steeds een stapje dichter bij $\mathbb{Z}$ te liggen, in de zin dat er steeds een extra eigenschap is toegevoegd van $\mathbb{Z}$ zodat de voorbeelden die eraan voldoen in het algemeen steeds meer op $\mathbb{Z}$ gaan lijken. De reden dat we deze volgorde kiezen is dat we geïnteresseerd zijn in structuren met eenduidige priemontbinding, en we zoeken de 'eenvoudigste' structuur met deze eigenschap. De volgende in ons rijtje is het *hoofdideaaldomein*, en we zullen laten zien dat hier eenduidige priemontbinding mogelijk is. Eerst voeren we wat begrippen in. Overigens, in dit hele hoofdstuk zullen we met R steeds een ring bedoelen.

Definitie 7.0.11. Ideaal. *Zij R een ring. Een* ideaal *I in R is een deelverzameling van R waarvoor de optelgroep een ondergroep is van die van R, met de eigenschap dat*

$$rx \in I \qquad en \qquad xr \in I$$

voor alle $x \in I$ en $r \in R$.

Voor onze toepassing op de Laatste stelling van Fermat beschouwen we eigenlijk alleen commutatieve ringen (in het bijzonder domeinen), zodat we alleen hoeven te controleren dat $rx \in I$, niet ook dat $xr \in I$.

Een ideaal is gesloten onder vermenigvuldiging, want omdat $rx \in I$ voor alle $r \in R, x \in I$ geldt dit zeker voor alle $r \in I, x \in I$. De enige reden waarom idealen niet altijd deelringen zijn, is dus dat ze niet per se 1 bevatten. Voorbeelden van idealen van $\mathbb{Z}$ zijn de verzamelingen $n\mathbb{Z}$ voor $n \in \mathbb{Z}$.

Als een ideaal I van R een eenheid a bevat, dan is $I = R$, want als $r \in R$ dan volgt uit $ra^{-1} \in R$ en $a \in I$ dat

$$r = (ra^{-1})a \in I.$$

In het bijzonder is R het enige ideaal dat 1 bevat, en dus ook het enige ideaal van R dat een ring is. Elk domein R heeft de triviale idealen $\{0\}$ en R. Als R een lichaam is, zijn dit de enige idealen. Immers, een ideaal $I \neq \{0\}$ in een lichaam bevat een $x \neq 0$, en omdat dit een eenheid is volgt dat $I = R$.

In Stelling 6.1.2 lieten we zien dat als G een Abelse groep is en H een ondergroep van G, dan is de nevenklasseverzameling G/H op een natuurlijke manier een groep. Vervolgens zagen we dat $\mathbb{Z}/n\mathbb{Z}$ zelfs een *ring* wordt. Dit komt doordat $n\mathbb{Z}$ een *ideaal* is van de ring $\mathbb{Z}$, zoals uit de volgende stelling blijkt. Het bewijs is eigenlijk gewoon een generalisatie van wat we in Stelling 6.2.1 deden voor $\mathbb{Z}/n\mathbb{Z}$.

Stelling 7.0.12. *Zij R een ring, en I een ideaal van R. In het bijzonder is I dus een optelgroep van R. Definiëer op de verzameling $R/I = \{a+I : a \in R\}$ naast de optelling $(a+I)+(b+I) = (a+b)+I$ ook een vermenigvuldiging, namelijk $(a+I)(b+I) = (ab+I)$. Met deze bewerkingen wordt R/I een ring.*[1]

Bewijs. We weten al dat R/I een optelgroep is, want I is ondergroep van de Abelse optelgroep van R. We hoeven ons dus alleen op de eigenschappen van vermenigvuldiging te richten. Allereerst gaan we na dat vermenigvuldiging welgedefinieerd is. Stel dat $a + I = b + I$ en $c + I = d + I$, we willen laten zien dat $(a + I)(c + I) = (b + I)(d + I)$. We weten (zie (6.3)) dat $b - a \in I$ en $d - c \in I$. Dus ook $(b - a)c \in I$ en $b(d - c) \in I$, en omdat I een optelgroep is volgt dat

$$bd - ac = b(d - c) + (b - a)c \in I.$$

Opnieuw uit (6.3) volgt dat $ac+I = bd+I$ ofwel $(a+I)(c+I) = (b+I)(d+I)$, zoals gewenst.

Nu volgen de ringeigenschappen vrijwel direct uit die van R, net als in Stelling 6.1.2. Distributiviteit volgt net als daar, maar nu met $\overline{x}$ vervangen door $x + I$. Associativiteit van vermenigvuldiging gaat zo:

$$\Big((a+I)(b+I)\Big)(c+I) = (ab+I)(c+I) = (ab)c+I = a(bc)+I = (a+I)(bc+I) = (a+I)\Big((b+I)(c+I)\Big).$$

Net als in Stelling 6.1.2 volgt dat $0 + I$ het nulelement is, $1 + I$ het eenheidselement, en $-a + I$ de tegengestelde van $a + I$. Hiermee is R/I een ring. □

We zeggen weer dat a en b congruent zijn modulo I als $a + I = b + I$, notatie $a \equiv b \pmod{I}$. Zoals we na het bewijs van Stelling 6.1.2 al opmerkten, is dit een equivalentierelatie die consistent is met optelling. Dat wil zeggen, $a \equiv b \pmod{I}$ en $c \equiv d \pmod{I}$ impliceert $a + c \equiv b + d \pmod{I}$.[2] Nu is de relatie zelfs ook consistent met vermenigvuldiging: $a \equiv b \pmod{I}$ en $c \equiv d \pmod{I}$ impliceert $ac \equiv bd \pmod{I}$.

7.1 Hoofdidealen

We kunnen idealen in een domein R construeren door bijvoorbeeld een willekeurig element $a \in R$ te nemen, en het ideaal (a) te definieëren als

$$(a) := \{ra : r \in R\}.$$

[1] Ons bewijs is een bewerking van [8], Stelling 11.14.

[2] Om verwarring te voorkomen, merken we op dat we de bewerking daar multiplicatief hebben genoteerd, en hier additief.

Het is makkelijk na te gaan dat dit inderdaad een ideaal is. Het heet het ideaal *voortgebracht door* a, en is het kleinste ideaal dat a bevat. Als I namelijk een ideaal is zodat $a \in I$, dan is $ra \in I$ voor alle $r \in R$, dus $(a) \subset I$. Algemeen kunnen we het ideaal $(a_1, \ldots, a_n)$ *voortgebracht door* $a_1, \ldots a_n$ definiëren als de verzameling van alle 'lineaire combinaties' van deze elementen:

$$(a_1, \ldots, a_n) := \{r_1 a_1 + \ldots + r_n a_n : r_i \in R \text{ voor alle } i\}.$$

Dit is het kleinste ideaal dat de elementen $a_1, \ldots, a_n$ bevat.

In $\mathbb{Z}$ bestaat bijvoorbeeld het ideaal $(9, 12)$ uit alle getallen van de vorm $9k + 15l$ met $k, l \in \mathbb{Z}$. In elk geval zit $\text{ggd}(9, 15) = 3$ in dit ideaal, want $9 \cdot 2 + 15 \cdot -1 = 3$. Dus $(3) \subset (9, 12)$. Aan de andere kant, uit $9k + 15l = 3(3k + 5l)$ wordt duidelijk dat elk element van het ideaal een drievoud is, dus $(9, 12) \subset (3)$. We concluderen dat

$$(9, 12) = (3).$$

Een ideaal voortgebracht door één element heet een *hoofdideaal*. Zo'n ideaal is dus te schrijven als (a) voor een $a \in R$. Bijvoorbeeld $(9, 12)$ is dus een hoofdideaal. Dit is geen toeval: we zullen straks laten zien dat *elk* ideaal in $\mathbb{Z}$ is een hoofdideaal is. Een domein met deze fijne eigenschap heet een hoofdideaaldomein.

Definitie 7.1.1. Hoofdideaaldomein. *Een* hoofdideaaldomein *is een domein waarin elk ideaal hoofdideaal is.*

Hoofdideaaldomeinen hebben veel gemeen met $\mathbb{Z}$. De belangrijkste overeenkomst is dat we kunnen spreken van eenduidige priemfactorisatie, zoals we straks zullen bewijzen. Maar daarvoor moeten we eerst begrippen als deelbaarheid en priemgetal generaliseren naar willekeurige domeinen R. Deelbaarheid generaliseren is eenvoudig: we zeggen dat $a \in R$ deler is van $b \in R$, notatie $a|b$, als er een $k \in R$ is zodat $ak = b$. Omdat het hoofdideaal (a) precies bestaat uit de elementen ak met $k \in R$, de 'veelvouden' van a, is dit laatste equivalent met $b \in (a)$. Omdat (b) het kleinste ideaal is dat b bevat, en omdat $b = 1 \cdot b \in (b)$, is op zijn beurt $b \in (a)$ equivalent met $(b) \subset (a)$. We concluderen dat voor a, b in een domein R geldt:

$$a|b \iff (b) \subset (a). \tag{7.1}$$

Dit is een prettige eigenschap die later handig van pas zal komen. Er is nog een manier om tegen deling aan te kijken, namelijk door te rekenen modulo een hoofdideaal. Zij R een domein, en $a, b, c \in R$. We brengen in de herinnering dat als I een ideaal is, we met $a \equiv b \pmod{I}$ bedoelen dat $a + I = b + I$, en dat is weer equivalent met $a - b \in I$. Dus $a \equiv b \pmod{(c)}$ betekent dat $a - b \in (c)$, ofwel c is deler van $a - b$. De congruentierelatie is dus vertrouwder dan die op het eerste gezicht leek: het betekent hetzelfde als in $\mathbb{Z}$. We schrijven $a \equiv b \pmod{(c)}$ liever gewoon als $a \equiv b \pmod{c}$, en we hebben dus

$$a \equiv b \pmod{c} \quad \iff \quad c|(a - b). \tag{7.2}$$

In het bijzonder is c deler van a precies dan als $a \equiv 0 \pmod{c}$.

7.2 Unieke priemfactorisatie

7.2.1 Wanneer kunnen we spreken van een priemontbinding?

In $\mathbb{Z}$ worden de priemgetallen meestal informeel ingevoerd als de verzameling $\mathbb{P}$ van *positieve* getallen p ongelijk aan 1 met als enige positieve delers 1 en p zelf. Dus $\mathbb{P} = \{2, 3, 5, 7, 11, \ldots\}$. In willekeurige domeinen kunnen we meestal niet spreken van positieve getallen (denk bijvoorbeeld aan de gehele getallen van Gauss), dus voor de algemene definitie is het handig om die restrictie weg te laten. We beschouwen daarom $-\mathbb{P} \cup \mathbb{P} = \{2, -2, 3, -3, 5, -5, \ldots\}$ als de verzameling priemgetallen van $\mathbb{Z}$. De priemgetallen p zijn dus precies de getallen ongelijk aan ± 1 met de eigenschap dat als $p = ab$, dan is een van hen, zeg a, gelijk aan ± 1, en bijgevolg $b = \pm p$. De priemgetallen zijn dus de niet-triviale 'irreducibele elementen' in de zin dat ze niet verder ontbonden kunnen worden, behalve op triviale manieren. Dit kunnen we als volgt generaliseren.

Definitie 7.2.1. Irreducibel. *Zij R een domein. Een element $p \in R$ heet* irreducibel *als $p \notin R^*$, en voor alle $a, b \in R$ geldt:*

$$\textit{Als} \quad p = ab, \quad \textit{dan} \quad a \in R^* \quad \textit{of} \quad b \in R^*.$$

Als R een domein is, $u \in R^*$ een eenheid en $p \in R$ irreducibel, dan is up irreducibel. Stel namelijk $up = ab$, dan volgt $p = (u^{-1}a)b$ dus $u^{-1}a \in R^*$ of $b \in \mathbb{R}^*$. Als $b \in R^*$ zijn we klaar, als $u^{-1}a \in R^*$ dan volgt omdat $u \in R^*$ en omdat R^* een groep is onder vermenigvuldiging dat $a = uu^{-1}a \in R^*$. Dus up is irreducibel.

We willen bewijzen dat in een hoofdideaaldomein elk element een in essentie unieke priemontbinding heeft. We zeggen 'in essentie' omdat de factorisatie uniek is op twee dingen na: de volgorde en vermenigvuldiging met eenheden. Bijvoorbeeld

$$(-1) \cdot 2^2 \cdot 3 \cdot 5 = (-1) \cdot 2 \cdot 5 \cdot 3 \cdot 2$$

beschouwen we als 'in essentie gelijke' priemontbindingen van -60. Deze zijn ook in essentie gelijk aan de priemontbinding

$$1 \cdot (-5) \cdot 2 \cdot (-2) \cdot (-3).$$

In een willekeurig domein R verstaan we onder een priemontbinding van $x \in R$ een product

$$x = u \cdot p_1 \cdot p_2 \cdot \ldots \cdot p_n$$

van een eenheid $u \in R^*$ en eindig veel (niet noodzakelijk verschillende) *irreducibele* elementen $p_i \in R$. Net als in $\mathbb{Z}$ kunnen we, omdat R commutatief is, in zo'n ontbinding de volgorde verwisselen of vermenigvuldigen met eenheden. Zo zijn, als a en b eenheden zijn,

$$x = u \cdot p_1 \cdot p_2 \cdot p_3 = ua^{-1}b^{-1} \cdot ap_1 \cdot p_2 \cdot bp_3$$

twee priemontbindingen van hetzelfde getal, want $ua^{-1}b^{-1}$ is een eenheid, en ap_1 en bp_3 zijn, zoals we al opmerkten, irreducibel. Deze discussie motiveert de volgende definitie.

Definitie 7.2.2. Priemontbinding. *Zij R een domein en $x \in R$. Onder een priemontbinding van x verstaan we een ontbinding $x = u \cdot p_1 \cdot p_2 \cdot \ldots \cdot p_m$ van x als product van een eenheid $u \in R^*$ en eindig veel (namelijk $m \geq 0$) irreducibele elementen $p_i \in R$. Twee priemontbindingen*

$$u \cdot p_1 \cdot p_2 \cdot \ldots \cdot p_m \;=\; v \cdot q_1 \cdot q_2 \cdot \ldots \cdot q_n$$

van x noemen we 'in essentie gelijk' als

$$m = n, \quad \textit{en} \quad \textit{voor alle } p_i \textit{ is er een } q_j \textit{ en een eenheid } w \textit{ zodat } p_i = wq_j.$$

We zeggen dat de priemontbinding van x (als die bestaat) in essentie uniek is, als alle priemontbindingen van x in essentie gelijk zijn.

7.2.2 Priemontbinding in hoofdideaaldomeinen

Het bewijs van unieke priemfactorisatie in een hoofdideaaldomein valt in twee delen uiteen: eerst bewijzen we dat (behalve 0) elk element een priemontbinding heeft, en vervolgens dat deze ontbinding in essentie uniek is.

Stelling 7.2.3. *Zij R een hoofdideaaldomein. Elk element van R verschillend van 0 heeft een priemontbinding.*[3]

Bewijs. We bewijzen dit uit het ongerijmde. Stel dat er een $x_0 \neq 0$ in R is zonder priemontbinding. Dan is x_0 geen eenheid, want in dat geval zou $x_0 = x_0$ een priemontbinding zijn (met 0 irreducibele elementen). Echter, x_0 is ook niet irreducibel, want dan zou $x_0 = 1 \cdot x_0$ een priemontbinding zijn. Uit deze twee gegevens volgt dat er $a, b \in R$ zijn die *beide geen eenheid* zijn, waarvoor $x_0 = ab$. Als a een priemontbinding $u \cdot p_1 \cdots p_m$ heeft, en b een priemontbinding $v \cdot q_1 \cdots q_n$, dan zou $uv \cdot p_1 \cdots p_m q_1 \cdots q_n$ een priemontbinding van x_0 zijn, en omdat we hadden aangenomen dat die niet bestaat volgt dat a of b, zeg a, óók geen priemontbinding heeft. We schrijven $a = x_1$.

Als $R = \mathbb{Z}$, dan zouden we nu al met Fermat's methode van oneindige afdaling op een tegenspraak stuiten. Immers, b is geen eenheid ofwel $b \neq \pm 1$, dus uit $x_0 = x_1 b$ en $x_0 \neq 0$ volgt dat $|x_1| < |x_0|$. Bovendien volgt uit $x_0 \neq 0$ dat $x_1 \neq 0$, en x_1 heeft geen priemontbinding, dus x_1 is een strikt kleiner getal dat aan de eisen van x_0 voldoet. We kunnen dus met inductie een oneindige rij $x_0, x_1, x_2, \ldots$ gehele getallen construeren waarvoor $|x_0| > |x_1| > |x_2| > \ldots$, en dat kan niet.

Voor een willekeurig hoofdideaaldomein R geldt dit argument niet, omdat we geen 'orde van grootte' voor de elementen van R hebben gedefinieerd. Toch kunnen we bovenstaand argument met een kleine aanpassing laten werken voor willekeurige R.

Uit $x_0 = x_1 b$ volgt dat $(x_0) \subset (x_1)$, en deze inclusie is geen gelijkheid. Stel namelijk dat $(x_1) \subset (x_0)$. Dan volgt uit $x_1 \in (x_0)$ dat $x_1 = x_0 c$ voor een $c \in R$. Maar ook $x_0 = x_1 b$, dus $x_1 = x_1 bc$. Dus $x_1(1 - bc) = 0$, en omdat een domein geen nuldelers heeft en omdat $x_1 \neq 0$ volgt dat $1 - bc = 0$, zodat $bc = 1$. Dus b is een eenheid, tegenspraak. We concluderen dat de aanname $(x_1) \subset (x_0)$ onjuist is, dus

$$(x_0) \subsetneq (x_1).$$

[3] Ons bewijs is een bewerking van [8], Stelling 12.8.

Omdat x_1 geen priemontbinding heeft en niet nul is (want $x_0 = x_1 b$ en x_0 is niet nul), kunnen we op dezelfde manier als net een x_2 construeren zodat $(x_1) \subsetneq (x_2)$, en zo doorgaande krijgen we een oneindige keten

$$(x_0) \subsetneq (x_1) \subsetneq (x_2) \subsetneq (x_3) \subsetneq (x_4) \subsetneq \dots$$

van steeds groter wordende idealen. We hoeven alleen nog te bewijzen dat zo'n keten niet bestaat. Definieer I als de vereniging van deze oneindig veel idealen:

$$I = \bigcup_{k \geq 0} (x_k).$$

We laten zien dat I een ideaal is. Als $p, q \in I$, dan zijn er m, n, zeg met $n \geq m$, zodat $p \in (x_m)$ en $q \in (x_n)$. Omdat $(x_m) \subset (x_n)$ volgt dat $p, q \in (x_n)$, dus $p + q \in (x_n)$, want het ideaal (x_n) is een optelgroep. Dus $p + q \in I$, dat wil zeggen I is gesloten onder optelling. Bovendien volgt omdat p in de optelgroep (x_m) zit, dat ook $-p \in (x_m)$, dus $-p \in I$ zodat I 'gesloten is onder tegengestelde nemen'. Bovendien zit 0 in I, want 0 zit zelfs in elk van de idealen (x_k), en we concluderen dat I een ondergroep van de optelgroep van R is. Als $r \in R$ en $y \in I$, dan is $y \in (x_m)$ voor een m, dus $ry \in (x_m)$ zodat $ry \in I$. We concluderen: I is een ideaal.

Omdat R een *hoofdideaaldomein* is, is er een $a \in R$ zodat $I = (a)$. Dus $a \in I$, zodat $a \in (x_n)$ voor een zekere n. Maar dat betekent dat $(a) \subset (x_n)$, ofwel $I \subset (x_n)$. En omdat $(x_{n+1}) \subset I$, volgt dat $(x_{n+1}) \subset (x_n)$, in tegenspraak met het feit dat $(x_n) \subsetneq (x_{n+1})$. De aanname is dus onjuist: elke $x_0 \neq 0$ in R heeft een priemontbinding. □

Nu we bewezen hebben dat er in een hoofdideaaldomein priemontbinding mogelijk is, hoeven we alleen nog te bewijzen dat zo'n ontbinding in essentie uniek is. Hiervoor hebben we de 'priemeigenschap' van irreducibele elementen p nodig: $p|ab$ impliceert dat $p|a$ of $p|b$. Voor gehele getallen hebben we dit al gebruikt (maar niet bewezen). Er geldt in $\mathbb{Z}$ zelfs dat de priemgetallen *precies* de getallen zijn met deze eigenschap, zodat het als alternatieve definitie kan worden gezien. In andere ringen hoeft dat niet te gelden; daarom introduceren we een nieuw begrip.

Definitie 7.2.4. Priemelement. *Een element p van een domein R heet een* priemelement *als $p \notin R^*$, $p \neq 0$ en voor alle $a, b \in R$ geldt:*

$$\textit{Als} \quad p|ab, \quad \textit{dan} \quad p|a \quad \textit{of} \quad p|b.$$

Alle priemelementen in een domein zijn irreducibel. Stel namelijk p is priemelement in een domein R, en er zijn $a, b \in R$ zodat dat $p = ab$. Dan geldt zeker dat $p|ab$, dus $p|a$ of $p|b$, zeg $p|a$. Er is dus een $r \in R$ zodat $pr = a$. Dus $p = ab = prb$, zodat $p(1 - rb) = 0$. Omdat we in een domein werken en p niet nul is, volgt dat $1 - rb = 0$, dus $rb = 1$ en we hebben $b \in R^*$. (Analoog volgt in het geval $p|b$ dat $a \in R^*$.) Dus $p = ab$ impliceert $a \in R^*$ of $b \in R^*$, en omdat per definitie $p \notin R^*$ is p irreducibel.

Omgekeerd hoeft niet altijd te gelden dat elk irreducibel element ook priemelement is, maar hoofdideaaldomeinen hebben de fijne eigenschap dat dat wel altijd zo is.

Lemma 7.2.5. *De irreducibele elementen in een hoofdideaaldomein zijn precies de priemelementen.*[4]

[4]Ons bewijs is een bewerking van [8], Lemma 12.10.

Bewijs. Zoals we net zagen hoeven we alleen aan te tonen dat alle irreducibele elementen priem zijn. Zij R een hoofdideaaldomein, en $p \in R$ irreducibel. Stel voor $a, b \in R$ geldt $p|ab$. Als $p|a$ dan zijn we klaar, stel dus p deelt niet a. Dit betekent dat $a \notin (p)$, dus het ideaal (a, p) voortgebracht door a en p is strikt groter dan (p):

$$(p) \subsetneq (a, p).$$

Omdat R een hoofdideaaldomein is, is er een $c \in R$ zodat $(a, p) = (c)$. Uit $(p) \subsetneq (c)$ volgt nu dat er een $d \in R$ bestaat zodat

$$p = dc, \qquad \text{en bovendien} \qquad d \notin R^*. \tag{7.3}$$

Immers, uit $d \in R^*$ zou volgen dat $c = d^{-1}p$ en dus $(c) \subset (p)$, maar dat is niet zo. Omdat p irreducibel is, volgt uit (7.3) dat $c \in R^*$, dus zoals we eerder al opmerkten volgt dat $(c) = R$, ofwel

$$(a, p) = R.$$

In het bijzonder is dus $1 \in (a, p)$, dus er zijn $x, y \in R$ zodat $ax + py = 1$. Omdat $p|ab$ is er een $s \in R$ zodat $ps = ab$, en we zien dat

$$b = b(ax + py) = abx + pby = psx + pby = p(sx + by),$$

dus $p|b$. We concluderen: uit $p|ab$ volgt $p|a$ of $p|b$. Bovendien is $p \neq 0$, want uit $0 = 0 \cdot 0$ volgt dat 0 is niet irreducibel is. Per definitie is p geen eenheid. We concluderen dat p een priemelement is. □

Nu kunnen we ons hoofdresultaat van dit hoofdstuk bewijzen.

Stelling 7.2.6. *Zij R een hoofdideaaldomein. Elk element van R verschillend van 0 heeft een priemontbinding, en deze is (in de zin van Definitie 7.2.2) in essentie uniek.*[5]

Bewijs. Na Stelling 7.2.3 hoeven we alleen nog aan te tonen dat de ontbindingen in essentie uniek zijn. Stel we hebben twee priemontbindingen

$$up_1p_2 \cdots p_m = vq_1q_2 \cdots q_n \tag{7.4}$$

van een element ongelijk aan 0; dat wil zeggen, u en v zijn eenheden en de p_i, q_j irreducibel (en dus priem). Om te bewijzen dat ze in essentie gelijk zijn, moeten we aantonen dat $m = n$, en dat er voor elke p_i een q_j en een eenheid a is zodat $p_i = aq_j$. We bewijzen dit met inductie naar m.

Stel eerst $m = 0$. Dan staat er $u = vq_1 \cdots q_n$, dus $n = 0$. Immers, als $n \neq 0$, dan is q_1 deler van u, dus $zq_1 = u$ voor een $z \in R$. Maar dan volgt $u^{-1}zq_1 = 1$, dus q_1 is een eenheid, tegenspraak want hij is irreducibel. Dus $n = 0$, en dus $m = n$. Aan de bewering 'voor elke p_i is er een q_j en een $a \in R^*$ zodat $p_i = aq_j$' is op triviale manier voldaan, want er zijn geen p_i.

Stel nu dat we het voor alle $m \leq k - 1$ bewezen hebben voor een $k \geq 1$; we willen het voor $m = k$ bewijzen. Uit

$$up_1p_2 \cdots p_k = vq_1q_2 \cdots q_n$$

[5] Ons bewijs is een bewerking van [8], Stelling 12.11.

volgt dat p_k deler is van $vq_1q_2\cdots q_n$, en dus volgens Lemma 7.2.5 deler is van minstens één van de $v, q_1, \ldots, q_n$. Deler van de eenheid v kan niet, want dan zou zoals net volgen dat p_k zelf een eenheid is. Er is dus een i zodat $p_k|q_i$, en dus een $w \in R$ zodat $wp_k = q_i$. Bovendien moet w een eenheid zijn, want q_i is irreducibel en p_k is geen eenheid. De vergelijking waar we mee begonnen, kunnen we nu omschrijven als $up_1p_2\cdots p_k = vq_1q_2\cdots q_{i-1}q_{i+1}\cdots q_nwp_k$, ofwel,

$$(up_1p_2\cdots p_{k-1} - vq_1q_2\cdots q_{i-1}q_{i+1}\cdots q_nw)p_k = 0.$$

Omdat p_k priem is, is $p_k \neq 0$, en omdat een domein geen nuldelers heeft volgt dat de linkerfactor nul is, dus

$$up_1p_2\cdots p_{k-1} = vwq_1q_2\cdots q_{i-1}q_{i+1}\cdots q_n.$$

Omdat vw een eenheid is, zijn dit weer twee priemontbindingen, maar nu staan links $k-1$ priemelementen en rechts $n-1$. Uit de inductiehypothese volgt dat $k-1 = n-1$, zodat $k = n$. Bovendien is er voor elk van de p_r met $1 \leq r \leq k-1$ een q_s en een eenheid a zodat $p_r = aq_s$. We wisten al dat dit ook voor p_k geldt: $p_k = wq_i$ met w een eenheid. De twee ontbindingen zijn dus in essentie gelijk.

Met inductie volgt dat (voor alle $m \geq 0$) de ontbindingen in (7.4) in essentie gelijk zijn, zoals we wilden bewijzen. □

7.3 Deling met rest

Nu we dit belangrijke resultaat over hoofdideaaldomeinen bewezen hebben, willen we natuurlijk voorbeelden vinden van hoofdideaaldomeinen. We merkten al op dat $\mathbb{Z}$ er een is, maar dat hebben we nog niet bewezen. Een ander belangrijk voorbeeld is de ring $\mathbb{Z}[i]$ van gehele getallen van Gauss, en in verband met het geval $n = 3$ van FLT kunnen we ons afvragen of de ring $\mathbb{Z}[\sqrt{-3}]$ van getallen van de vorm $a+b\sqrt{-3}$ met a, b geheel een hoofdideaaldomein is. We zagen in Lemma 4.2.1 dat Euler een deelresultaat voor het geval $n = 3$ wilde bewijzen door over te gaan op dit soort getallen, en daar 'priemfactorisatie' toe te passen. Het lukte hem echter niet om een overtuigend bewijs te geven dat er bij deze getallen sprake was van priemontbindingen, en we zagen dat zijn redenering in elk geval fout is voor de ring $\mathbb{Z}[\sqrt{-5}]$ van getallen van de vorm $a + b\sqrt{-5}$, a, b geheel. Dat het daar verkeerd gaat komt doordat deze ring geen hoofdideaaldomein is.

Een manier om te bewijzen dat een bepaald domein een hoofdideaaldomein is, is om een variant van *deling met rest* deling met rest toe te passen. Het is intuïtief duidelijk (hoewel niet makkelijk formeel te bewijzen) dat $\mathbb{Z}$ een domein is, dit nemen we daarom als uitgangspunt. Stel I is een ideaal in $\mathbb{Z}$. Als $I = \{0\}$ dan is I het hoofdideaal (0) voortgebracht door 0. Als $I \neq \{0\}$, dan bevat I een element $d \neq 0$, en omdat ook $-d \in I$, bevat I een positief getal. Zij a het kleinste positieve getal in I; er geldt dus $(a) \subset I$. Stel $b \in I$. Via deling met rest kunnen we $q, r \in \mathbb{Z}$ vinden, met $0 \leq r < a$, zodat $b = qa + r$. Omdat I een ideaal is, volgt uit $a \in I$ en $b \in I$ dat $qa \in I$ en dus ook

$$r = b - qa \in I.$$

Omdat $0 \leq r < a$ volgt uit de minimaliteit van a dat $r = 0$, dus $b = qa$ zodat $b \in (a)$. Dit geldt voor alle $b \in I$, dus $I \subset (a)$, en we concluderen dat

$$I = (a).$$

Hiermee hebben we bewezen:

Stelling 7.3.1. *$\mathbb{Z}$ is een hoofdideaaldomein.*

Hiermee is dus meteen eenduidige priemfactorisatie in $\mathbb{Z}$ bewezen.

In veel domeinen waarin de elementen op een natuurlijke manier een 'orde van grootte' hebben, kunnen we een variant van restdeling toepassen. Daarom is het handig wat we net deden te generaliseren: het voorkomt dat we dingen dubbel doen.

Stelling 7.3.2. *Zij R een domein. Stel er is een functie $f : R \to \mathbb{R}$ die voldoet aan de volgende eigenschappen.*

1). Het beeld van f ligt in $\mathbb{Z}_{\geq 0}$.

2). $f(x) = 0$ dan en slechts dan als $x = 0$.

3). Voor alle $a, b \in R$ met $b \neq 0$ zijn er elementen $q, r \in R$ zodat

$$a = qb + r \qquad en \qquad f(r) < f(b).$$

Dan is R een hoofdideaaldomein.

Bewijs. Stel I is een ideaal in R. Als $I = \{0\}$, dan is $I = (0)$ een hoofdideaal. Stel dus $I \neq \{0\}$. Dan is er een $d \neq 0$ in I, en omdat $f(d) > 0$ neemt f op I ook positieve waarden aan. Omdat f alleen *gehele* waarden aanneemt, neemt de functie op I een *kleinste positieve waarde* m aan. Zij $b \in I$ een element waarvoor $f(b) = m$. Omdat $b \in I$ is

$$(b) \subset I,$$

en we willen ook de omgekeerde inclusie bewijzen. Stel $a \in I$. Dan zijn er $q, r \in R$ zodat $a = qb + r$ en $f(r) < f(b)$. Omdat I een ideaal is en $a, b \in I$, is ook $r = a - qb \in I$. Maar $f(r) < f(b) = m$, en per definitie van m volgt dat $f(r) = 0$. Dus $r = 0$, en we concluderen dat $a = qb$ en dus $a \in (b)$. Dus

$$I \subset (b),$$

en we concluderen dat $I = (b)$. Dus R is een hoofdideaaldomein. □

Als $R = \mathbb{Z}$ kunnen we bijvoorbeeld voor f de absolute waarde nemen. De identieke functie $f : x \mapsto x$ werkt in het voorgaande bewijs niet, omdat het beeld van f in $\mathbb{Z}_{\geq 0}$ moet liggen.

7.4 Polynomen ontbinden

Een ander voorbeeld waarop we bovenstaande stelling kunnen toepassen, is de veeltermring $R[X]$ van polynomen met coëfficiënten in een ring R, bijvoorbeeld $R = \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ of $\mathbb{C}$. De *graad* $\deg(p)$ van een polynoom $p = c_0 + c_1X + c_2X^2 + \ldots + c_nX^n \in R[X]$ is de grootste coëfficient c_i ongelijk aan nul, dus als $c_n \neq 0$ dan is $\deg(p) = n$. (Als $c_n = 0$ kunnen we de term c_nX^n net zo goed weglaten.) Een 'constant polynoom' $c_0 \neq 0$ heeft graad 0, en de graad van het

nulpolynoom 0 definieëren we als -1. Als we in bovenstaande stelling $f(p) = \deg(p)+1$ nemen, dan voldoet f dus in elk geval aan eigenschap 1) en 2). Stel we hebben twee veeltermen

$$a = c_m X^m + \ldots + c_1 X + c_0 \qquad \text{en} \qquad b = d_n X^n + \ldots + d_1 X + d_0$$

van graad m en n, *en de kopcoëfficiënt d_n van b ligt in R^*.*[6] Als $n > m$, dan kunnen we gewoon $q = 0$, $r = a$ nemen. Als $m \geq n$, dan kunnen we 'een veelvoud van b aftrekken van a zodat de kopterm verdwijnt':

$$c_m X^m + \ldots + c_1 X + c_0 = (c_m d_n^{-1} X^{m-n})(d_n X^n + \ldots + d_1 X + d_0) + r$$

voor een r met graad kleiner dan n; kortom, $a = qb + r$ met $f(r) < f(b)$.

Stel dat R een lichaam is. Dan ligt de kopkoëfficiënt van b altijd in R^*, want omdat b graad n heeft is $d_n \neq 0$. Het bovenstaande werkt dus voor alle $a, b \in R$ met $b \neq 0$. Dus f voldoet aan de voorwaarden van Stelling 7.3.2, en we concluderen:

Gevolg 7.4.1. *Zij R een lichaam. Dan is de veeltermring $R[X]$ een hoofdideaaldomein.*

We kunnen polynomen met coëfficiënten in $\mathbb{Q}$, $\mathbb{R}$ of $\mathbb{C}$ dus eenduidig ontbinden in 'irreducibele veeltermen'. In $\mathbb{R}$ is bijvoorbeeld X^2+1 irreducibel. Immers, als we hem kunnen ontbinden als fg dan hebben f, g graad $0, 2$ of $1, 1$ of $2, 0$, want de som van de graden moet 2 zijn (ga maar na). Als een van de graden nul is zijn we klaar: f of g is dan een veelterm van graad nul, dat wil zeggen een reëel getal ongelijk aan nul, en dit is een eenheid in $\mathbb{R}[X]$ met inverse $1/f$ of $1/g$. Als hij dus een ontbinding in niet-eenheden heeft, dan hebben f en g graad 1 en zijn dus van de vorm $aX - b$, $cX - d$ met $a, b, c, d \in \mathbb{R}, a, c \neq 0$. Maar uit $X^2 + 1 = (aX - b)(cX - d)$ volgt dat de reële getallen b/a en d/c nulpunt zijn van $X^2 + 1$ en dat kan niet: een kwadraat van een reëel getal is nooit -1. We concluderen dat de veelterm irreducibel is in $\mathbb{R}$. In $\mathbb{C}$ is dat niet zo: we hebben bijvoorbeeld de ontbinding $X^2 + 1 = (X + i)(X - i)$. Dit is een algemeenheid: het lichaam $\mathbb{C}$ heeft de fijne eigenschap dat de irreducibele veeltermen precies de lineaire veeltermen $X - a$ zijn, dus *elke* polynoom van graad $n \geq 0$ kan ontbonden worden in lineaire factoren als $(X - a_1) \cdots (X - a_n)$ en heeft in het bijzonder precies n complexe nulpunten. Dat wordt samengevat door te zeggen dat $\mathbb{C}$ een *algebraïsch gesloten* lichaam is. Zie bijvoorbeeld een willekeurig boek over Galois-theorie voor een bewijs.

7.4.1 Nulpunten van veeltermen

Als R geen lichaam is, kunnen we wel andere nuttige dingen over $R[X]$ zeggen. Zij $f \in R[X]$. Neem een willekeurige $a \in R$, en beschouw het lineaire polynoom $g = X - a$. Omdat g graad 1 en kopcoëfficiënt $1 \in R^*$ heeft, kunnen we door restdeling polynomen q, r vinden zodat $f = qg + r$, met r van hoogstens graad 0. Dat betekent dat r een element van R is. We gaan nu laten zien dat het een *bijzonder* element van R is, in elk geval bijzonder ten opzichte van f.

Als f een polynoom in $R[X]$ is, zeg $f = c_n X^n + \ldots + c_2 X^2 + c_1 X + c_0$, dan kunnen we f 'evalueren' in een punt $a \in R$:

$$f(a) := c_n a^n + \ldots + c_2 a^2 + c_1 a + c_0.$$

[6] De kopcoëfficiënt van een veelterm is de coëfficiënt van de grootste macht van X waarvoor de coëfficiënt niet nul is.

Dit is weer een element van R. Voor $R = \mathbb{R}$ of $\mathbb{C}$ kunnen we op deze manier de bekende veeltermfuncties uit de analyse construeren. Er geldt

$$\begin{aligned} f - f(a) &= (c_n X^n + \ldots + c_2 X^2 + c_1 X + c_0) - (c_n a^n + \ldots + c_2 a^2 + c_1 a + c_0) \\ &= c_n(X^n - a^n) + \ldots + c_2(X^2 - a^2) + c_1(X - a). \end{aligned}$$

Uit de formule

$$X^k - a^k = (X - a)(X^{k-1} + aX^{k-2} + a^2 X^{k-3} \ldots + a^{k-2}X + a^{k-1})$$

volgt dat elk van de termen $c_k(X^k - a^k)$ veelvoud is van het lineaire polynoom $X - a$, dus ook $f - f(a)$ is veelvoud van $(X - a)$. We concluderen:

Lemma 7.4.2. *Zij R een ring, $f \in R[X]$ en $a \in R$. Dan is er een veelterm $q \in R[X]$ zodat*

$$f = q(X - a) + f(a).$$

In de volgende stelling gebruiken we de 'homomorfisme-achtige' eigenschap $\deg(fg) = \deg(f) + \deg(g)$ voor polynomen $f, g \in R[X]$ ongelijk aan nul in een *domein* R, wat we al gebruikt hebben voor $fg = X^2 + 1 \in \mathbb{R}[X]$. Als a_m en b_n de kopcoëfficiënten zijn van twee veeltermen f en g van graad m en n, dan is de coëfficiënt van X^{m+n} van fg namelijk gelijk aan $a_m b_n$. Omdat a_m, b_n niet nul zijn en R geen nuldelers heeft, is dit niet nul, en het is duidelijk dat de coëfficiënten van de grotere machten van X wel nul zijn. Dus fg heeft graad $m + n$.

We gebruiken nog een belangrijke eigenschap. Stel dat R commutatief is. Omdat het 'rekenen met X' bijna per definitie gebeurt door X als een 'element' te beschouwen dat met alle andere elementen commuteert, geldt voor alle $f, g \in R[X]$:[7]

$$(f + g)(a) = f(a) + g(a) \qquad \text{en} \qquad (f \cdot g)(a) = f(a) \cdot g(a). \tag{7.5}$$

Een *nulpunt* van een polynoom f is een $a \in R$ waarvoor $f(a) = 0$. We weten uit ervaring dat we polynomen met reële of complexe coëfficiënten met nulpunten $a_1, \ldots, a_k$ altijd kunnen ontbinden door factoren $(x - a_i)$ 'buiten haakjes te halen'. We zijn nu zo ver dat we dat kunnen bewijzen, en bovendien voor algemene domeinen R.

Stelling 7.4.3. *Zij R een domein, en laat $a_1, a_2, \ldots, a_m$* verschillende *nulpunten van een veelterm $f \in R[X]$ van graad n zijn. Dan is er een $q \in R[X]$ zodat*

$$f = q(X - a_1)(X - a_2) \cdots (X - a_m).$$

Als $f \neq 0$, dan heeft f hoogstens n nulpunten.[8]

Bewijs. We bewijzen het met inductie naar m. Voor $m = 0$ is er niets te bewijzen. Voor $m = 1$ is het een speciaal geval Lemma 7.4.2: als a een nulpunt is van f, dan is daar $f(a) = 0$ en dus $f = q(X - a)$. Stel het geldt voor $m = k \geq 1$, en laat f een polynoom zijn met $k + 1$ verschillende nulpunten $a_1, a_2, \ldots, a_{k+1}$. We passen Lemma 7.4.2 toe op het nulpunt $a = a_{k+1}$:

[7] De eerste formule geldt ook als R niet commutatief is.

[8] Ons bewijs is een bewerking van [8], Gevolg 12.3.

er is een polynoom $q_1 \in R[X]$ zodat $f = q_1(X - a_{k+1})$. We schrijven $h = (X - a_{k+1})$, en laten zien dat de nulpunten $a_1, \ldots, a_k$ van f ook nulpunten zijn van q_1. Neem zo'n element a_i. Omdat R commutatief is, volgt uit (7.5) voor alle $i = 1, 2, \ldots, k$:

$$0 = f(a_j) = q_1(a_i)h(a_i) = q_1(a_i)(a_i - a_{k+1}).$$

Omdat de nulpunten $a_1, \ldots, a_{k+1}$ allen verschillend zijn, is $a_i - a_{k+1} \neq 0$. Omdat bovendien R geen nuldelers heeft, volgt dat $q_1(a_i) = 0$. We concluderen dat $a_1, \ldots, a_k$ verschillende nulpunten zijn van q_1, en uit de inductiehypothese volgt dat er een q is zodat

$$q_1 = q(X - a_1)(X - a_2) \cdots (X - a_k).$$

Dus $f = q_1(X - a_{k+1})$ is van de gewenste vorm.

Als f niet 0 is, dan is q dat ook niet, dus de graad n van f is $m + \deg(q) \geq m$. Dus het aantal nulpunten van f is hoogstens n. □

7.4.2 Cyclische eenhedengroepen

Nu zijn we zover dat we de beloofde generalisatie kunnen bewijzen van het vermoeden dat $(\mathbb{Z}/p\mathbb{Z})^*$ cyclisch is voor alle priemgetallen p.

Stelling 7.4.4. *Zij R een domein, en H een eindige ondergroep van R^*. Dan is H cyclisch.*[9]

Bewijs. De orde van H noemen we n. We laten met behulp van polynomen en de formule van Gauss zien dat H een element van orde n moet bevatten. Zij d een deler van n, het aantal elementen van H met orde d noemen we $\mu(d)$. Alle elementen van orde d zijn nulpunt van het polynoom

$$f = X^d - 1 \in R[X]$$

Stel $\mu(d)$ is niet nul. Dan is er een element x van orde d. De d verschillende machten $1, x, x^2, x^3, \ldots, x^{d-1}$ zijn allemaal nulpunt van f, want $(x^k)^d = (x^d)^k = 1^k = 1$. Omdat f graad d heeft, volgt uit Stelling 7.4.3 dat hij hoogstens d nulpunten heeft, dus $1, x, x^2, x^3, \ldots, x^{d-1}$ zijn precies zijn nulpunten. Dus elk element van orde d is een macht van x. De machten van x met orde d zijn precies de $\phi(d)$ elementen x^k met $\gcd(k, d) = 1$ (ga maar na), en we concluderen dat $\mu(d) = \phi(d)$. Dus voor alle delers d van n is $\mu(d) = 0$ of $\mu(d) = \phi(d)$, dus in elk geval $\mu(d) \leq \phi(d)$. Dus ook

$$\sum_{d|n} \mu(d) \leq \sum_{d|n} \phi(d), \tag{7.6}$$

met gelijkheid precies dan als $\mu(d) = \phi(d)$ voor alle $d|n$. Maar de gelijkheid geldt inderdaad, want beide sommen zijn gelijk aan n: voor de rechtersom is dit de formule van Gauss (6.7), de linkersom is het aantal elementen van H met orde d gesommeerd over alle mogelijke ordes d, dat is dus $|H| = n$. Dus inderdaad $\mu(d) = \phi(d)$ voor alle $d|n$. Dus ook $\mu(n) = \phi(n) \geq 1$: er is een element van orde n in H. Dat betekent dat H cyclisch is. □

We zagen al dat $\mathbb{Z}^* = \{-1, (-1)^2\}$ en $\mathbb{Z}[i]^* = \{i, i^2, i^3, i^4\}$, deze zijn inderdaad cyclisch.

[9] Ons bewijs is een bewerking van [8], Gevolg 12.4.

Gevolg 7.4.5. Primitieve wortels. *Zij p een priemgetal. Dan is*

$$(\mathbb{Z}/p\mathbb{Z})^* \cong \mathbb{Z}/(p-1)\mathbb{Z}.$$

Het aantal primitieve wortels van p is $\phi(p-1)$.

Bewijs. Volgens Stelling 6.2.5 is $\mathbb{Z}/p\mathbb{Z}$ een lichaam, en dus zeker een domein. De eenhedengroep $(\mathbb{Z}/p\mathbb{Z})^*$ is ondergroep van zichzelf met orde $p-1$, en is dus cyclisch. Uit Lemma 6.2.4 volgt dat hij isomorf is met $\mathbb{Z}/(p-1)\mathbb{Z}$. Let wel, in $(\mathbb{Z}/p\mathbb{Z})^*$ is de bewerking vermenigvuldiging, en in $\mathbb{Z}/(p-1)\mathbb{Z}$ optelling. We kunnen $\mathbb{Z}/(p-1)\mathbb{Z}$ zien als de 'optelgroep van de exponenten'. De primitieve wortels van p zijn per definitie de elementen van orde $p-1$ in $(\mathbb{Z}/p\mathbb{Z})^* \cong \mathbb{Z}/(p-1)\mathbb{Z}$, en volgens Stelling 6.2.2 zijn dat er $\phi(p-1)$. □

Hoofdstuk 8

Opnieuw het geval $n = 3$

Om het gat in het bewijs van FLT voor het geval $n = 3$ te repareren, zou het handig zijn als er sprake was van unieke priemfactorisatie in $\mathbb{Z}[\sqrt{-3}]$. We willen hiervoor bovenstaande theorie toepassen. Dat is niet eenvoudig, maar dat geeft niet: de zo gevormde theorie is niet alleen handig voor het bewijs van het geval $n = 3$, maar is ook op zichzelf heel mooi, en geeft nieuwe inzichten die ook van pas komen voor latere ontwikkelingen rond de stelling van Fermat waarvan we een klein deel in het volgende hoofdstuk beschrijven. De priemgetallen (behalve 2) vallen uiteen in de viervouden plus één en de viervouden plus drie, dat zijn dus

$$5, 13, 17, 29, 37, 41, 53, 61, 73, 89, 97, \ldots \quad \text{en} \quad 3, 7, 11, 19, 23, 31, 43, 47, 59, 67, 71, 79, 83 \ldots$$

Deze twee klassen van priemgetallen gedragen zich in veel opzichten verschillend; een voorbeeld daarvan werd al ontdekt door Fermat. Hij onderzocht welke priemgetallen te schrijven zijn als de som van twee kwadraten (van gehele getallen). Met de viervouden plus drie zijn we snel klaar: geen van hen is zo'n som. Stel namelijk $p = a^2 + b^2$ voor a, b geheel, p priem congruent 3 modulo 4. Kwadraten zijn altijd 0 of 1 modulo 4, want $0^2 \equiv 0, 1^2 \equiv 1, 2^2 \equiv 0$ en $3^2 \equiv 1 \pmod 4$. Dus modulo 4 zijn a^2, b^2 gelijk aan 0 of 1, en hun som dus aan $0, 1$ of 2, in tegenspraak met het feit dat $p \equiv 3 \pmod 4$.

Voor de priemen $p \equiv 1 \pmod 4$ geldt een heel ander verhaal. Het viel Fermat op dat welke p hij ook nam, telkens was het de som van twee kwadraten. Bijvoorbeeld

$$5 = 2^2 + 1^2, \quad 13 = 3^2 + 2^2, \quad 17 = 4^2 + 1^2, \quad 29 = 5^2 + 2^2, \quad 37 = 6^2 + 1^2, \quad 41 = 5^2 + 4^2,$$
$$53 = 7^2 + 2^2, \quad 61 = 6^2 + 5^2, \quad 73 = 8^2 + 3^2, \quad 89 = 8^2 + 5^2, \quad 97 = 9^2 + 4^2, \ldots.$$

Dit verschijnsel was zo hardnekkig aanwezig dat Fermat het probeerde te bewijzen, en hij beweerde dat hij daarin geslaagd was samen met nog wat soortgelijke resultaten. Hij schreef namelijk in 1658 in een brief aan Digby onder andere dat hij kon bewijzen:[1]

Stelling 8.0.6. (Fermat)

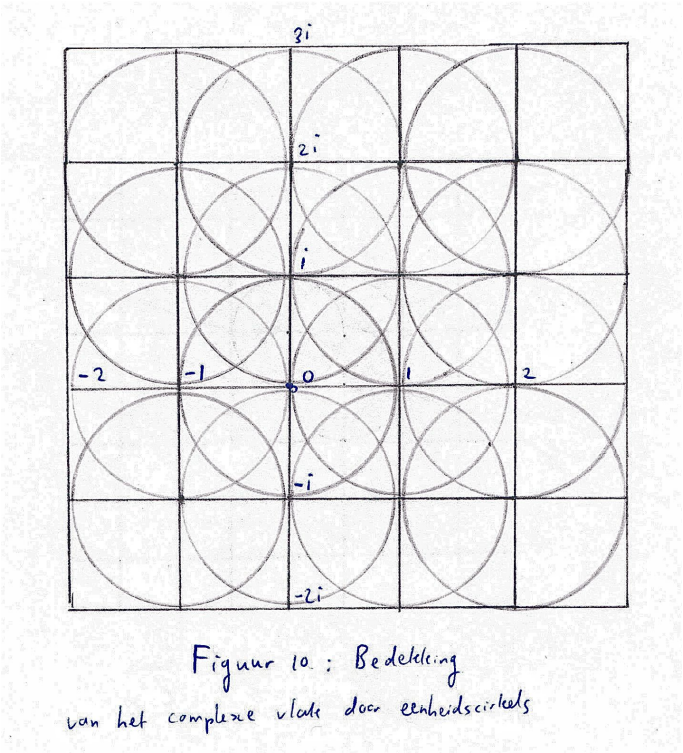
a). Elk priemgetal van de vorm $4n + 1$ *is van de vorm* $a^2 + b^2$*;*

b). Elk priemgetal van de vorm $3n + 1$ *is van de vorm* $a^2 + 3b^2$*.*

[1] Bron: zie [4]

Of Fermat ooit een correct bewijs had weten we niet, maar het eerste kloppende overgeleverde bewijs is van Euler. Zijn bewijs is vrij lang en ondoorzichtig (zie [4]), maar met de hier opgebouwde theorie kunnen we een moderner en eigenlijk eleganter bewijs geven. Het is een bijproduct van ons onderzoek naar het geval $n = 3$ van FLT. We merkten al op dat het gat in Euler's bewijs van het geval $n = 3$, dat te maken heeft met factorisatie van getallen van de vorm $a^2 + 3b^2$, te dichten is door eerder bewezen stellingen van Euler te gebruiken. Eén van die stellingen is deel b van hierboven.

8.1 Gehele getallen van Gauss



Figuur 10 : Bedekking van het complexe vlak door eenheidscirkels

Om te onderzoeken of $\mathbb{Z}[\sqrt{-3}]$ een hoofdideaaldomein is, is het instructief om eerst het eenvoudigere, en op zichzelf interessante voorbeeld

$$\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$$

te bestuderen, de ring van gehele getallen van Gauss. Deze werden door Gauss geïntroduceerd bij zijn studie naar 'kwadratische reciprociteit', een verschijnsel dat met de oplossingen van de congruentievergelijking $y \equiv x^2 \pmod{n}$ te maken heeft. We vragen ons af of we een soort restdeling kunnen bedenken in $\mathbb{Z}[i]$. Allereerst hebben we een functie $f : \mathbb{C} \to \mathbb{Z}_{\geq 0}$ nodig die voldoet aan de voorwaarden van Stelling 7.3.2. De meest voor de hand liggende functie, de absolute waarde, voldoet niet, want $|\alpha|$ is niet per se een geheel getal. Het *kwadraat* van de absolute waarde voldoet echter wel, want voor $a + bi \in \mathbb{Z}[i]$ is $|a + bi|^2 = a^2 + b^2 \in \mathbb{Z}_{\geq 0}$, want $a, b \in \mathbb{Z}$. We definiëren daarom de functie

$$\|\cdot\| : \mathbb{Z}[i] \to \mathbb{Z}_{\geq 0} : \|a + bi\| \mapsto a^2 + b^2,$$

ofwel $\|\alpha\| = \alpha\overline{\alpha}$, en deze voldoet aan voorwaarde 1) en 2). We noemen dit de *norm* op $\mathbb{Z}[i]$. Merk op dat deze norm voor gehele λ niet voldoet aan $\|\lambda\alpha\| = |\lambda| \cdot \|\alpha\|$, maar wel aan $\|\lambda\alpha\| = |\lambda|^2\|\alpha\|$. Als $\alpha, \beta \in \mathbb{Z}[i]$ en $\beta \neq 0$, dan zoeken we $q, r \in \mathbb{Z}[i]$ zodat $\alpha = q\beta + r$ en $\|r\| < \|\beta\|$. Dit betekent dat $\|r\|/\|\beta\| = \|r/\beta\|$ kleiner dan 1 moet zijn, ofwel, $\|\alpha/\beta - q\|$ moet kleiner zijn dan 1. Dit wordt ons uitgangspunt van het bewijs van

Stelling 8.1.1. *$\mathbb{Z}[i]$ is een hoofdideaaldomein.*[2]

[2]Ons bewijs is een bewerking van [8], Stelling 12.19.

Bewijs. Zij $z \in \mathbb{C}$. Dan is er een $q \in \mathbb{Z}[i]$ zodat $\|z - q\| < 1$. Meetkundig betekent dit dat de open eenheidscirkels rond de 'standaardroosterpunten' het hele complexe vlak overdekken, en dat wordt intuïtief duidelijk uit Figuur 10. We kunnen dit hard maken door allereerst op te merken dat er voor alle $x \in \mathbb{R}$ een $k \in \mathbb{Z}$ is zodat $|x-k| \leq \frac{1}{2}$: voldoet entier(x) niet, dan voldoet entier$(x) + 1$ wel. Schrijf $z = c + di$ met $c, d \in \mathbb{R}$, en kies $m, n \in \mathbb{Z}$ zo dat $|c - m|, |d - n| \leq \frac{1}{2}$; we nemen $q = m + ni$. Er geldt nu dat

$$\|z - q\| = (c - m)^2 + (d - n)^2 \leq (\tfrac{1}{2})^2 + (\tfrac{1}{2})^2 = \tfrac{1}{2} < 1,$$

zoals gewenst.

Stel nu dat $\alpha, \beta \in \mathbb{Z}[i]$ en $\beta \neq 0$. Neem een $q \in \mathbb{Z}[i]$ zodat $\|\alpha/\beta - q\| < 1$, en definieer $r = \alpha - q\beta$. Dan is

$$\alpha = q\beta + r \qquad \text{en} \qquad \|r\| = \|\beta(\alpha/\beta - q)\| = \|\beta\| \cdot \|\alpha/\beta - q\| < \|\beta\|,$$

dus de functie $\|\cdot\|$ voldoet aan alle voorwaarden van f in Stelling 7.3.2. We concluderen: $\mathbb{Z}[i]$ is een hoofdideaaldomein. □

8.1.1 Priemelementen

Een interessante vraag is nu wat de priemelementen (of equivalent, de irreducibele elementen) zijn in $\mathbb{Z}[i]$. Door deze vraag te beantwoorden kunnen we gelijk Fermat's bewering 8.0.6 a) bewijzen.

Een groot voordeel van de norm $\|\cdot\|$ boven de gewone absolute waarde is dat het een geheel getal is, zodat we eigenschappen van deelbaarheid in $\mathbb{Z}$ kunnen gebruiken. Die eigenschappen van de gehele getallen $\|\alpha\|$ kunnen we naar $\mathbb{Z}[i]$ vertalen via de formule

$$\|\alpha\beta\| = \|\alpha\| \cdot \|\beta\|,$$

die zegt dat de norm-functie de vermenigvuldiging in de verschillende structuren respecteert. De formule, die we in het bewijs van daarnet al gebruikt hebben, volgt gelijk uit die voor de absolute waarde:

$$\|\alpha\beta\| = |\alpha\beta|^2 = (|\alpha| \cdot |\beta|)^2 = |\alpha|^2|\beta|^2 = \|\alpha\| \cdot \|\beta\|.$$

Hoe bepalen we of een element van $\mathbb{Z}[i]$ irreducibel is? Rechtstreeks gebruik maken van Definitie 7.2.1 of 7.2.4 is vaak een lastige klus, maar het kan gelukkig makkelijker. Als namelijk $\|\alpha\| = p$ voor een priemgetal $p \in \mathbb{Z}$, dan is α een priemelement. Stel namelijk dat $\alpha = \beta\gamma$ voor $\beta, \gamma \in \mathbb{Z}[i]$. Dan is

$$\|\beta\| \cdot \|\gamma\| = \|\beta\gamma\| = \|\alpha\| = p,$$

dus $\|\beta\| = 1, \|\gamma\| = p$ of andersom. We zagen in (5.3) al dat de eenheden van $\mathbb{Z}[i]$ precies de elementen zijn met norm 1, dus β of γ is een eenheid en we concluderen dat α irreducibel is.

Stel δ is een priemelement van $\mathbb{Z}[i]$. Dan is δ deler van $\delta\overline{\delta} = \|\delta\|$. Ontbinden we $\|\delta\|$ *in* $\mathbb{Z}$ als product $p_1 \cdots p_n$ van priemgetallen, dan volgt uit $\delta | p_1 \cdots p_n$ dat δ minstens één van de p_i deelt; immers, de priemgetallen p_i zijn ook elementen van $\mathbb{Z}[i]$. Dus δ treedt op als irreducibele factor in de priemontbinding in $\mathbb{Z}[i]$ van een priemgetal $p \in \mathbb{Z}$. Om alle irreducibele elementen van $\mathbb{Z}[i]$ te bepalen, hoeven we dus alleen de gehele priemgetallen te ontbinden in $\mathbb{Z}[i]$.

Stelling 8.1.2. *De priemgetallen $p \in \mathbb{Z}$ hebben de volgende priemontbinding in $\mathbb{Z}[i]$.*[3]

1. *Het priemgetal $p = 2$ heeft ontbinding $2 = -i \cdot (1+i)^2$.*
2. *Voor $p \equiv 1 \pmod 4$ zijn er priemelementen $\alpha, \overline{\alpha} \in \mathbb{Z}[i]$ zodat $p = \alpha\overline{\alpha}$.*
3. *Voor $p \equiv 3 \pmod 4$ is p zelf priemelement in $\mathbb{Z}[i]$.*

Bewijs. Het element $1+i$ is priem in $\mathbb{Z}[i]$, want zijn norm 2 is priem. Verder is $-i$ een eenheid in $\mathbb{Z}[i]$, dus $-i \cdot (1+i)^2$ is een priemontbinding. En omdat

$$-i \cdot (1+i)^2 = -i(1 + 2i + i^2) = -i(2i) = 2,$$

volgt onderdeel 1 van de stelling. Zij p een priemgetal van $\mathbb{Z}$. Stel p is reducibel (dat wil zeggen niet-irreducibel). Omdat p geen eenheid is, zijn er $\alpha, \beta \in \mathbb{Z}[i]$, beide geen eenheid, zodat $p = \alpha\beta$. Dus

$$p^2 = \|p\| = \|\alpha\beta\| = \|\alpha\| \cdot \|\beta\|,$$

en omdat p priem is en $\|\alpha\| \neq 1$ en $\|\beta\| \neq 1$ (want anders zouden het eenheden zijn), volgt dat $\|\alpha\| = p$ en $\|\beta\| = p$. Maar we zagen net dat dat betekent dat α en β priemelementen zijn. Bovendien is $\|\overline{\alpha}\| = \|\alpha\| = p$, dus ook $\overline{\alpha}$ is irreducibel. En omdat

$$p = \|\alpha\| = \alpha\overline{\alpha}$$

volgt dat $p = \alpha\overline{\alpha}$ de priemontbinding is van p. We hoeven dus alleen nog na te gaan welke priemgetallen reducibel zijn.

Stel $p \equiv 3 \pmod 4$. Als p reducibel is, dan is er dus een $\alpha = a + bi \in \mathbb{Z}[i]$ zodat $p = \alpha\overline{\alpha}$, ofwel $p = a^2 + b^2$. Maar we zagen in de inleiding van dit hoofdstuk al dat dat niet kan, dus p is irreducibel waarmee onderdeel 3 bewezen is.

Stel tenslotte dat $p \equiv 1 \pmod 4$. Uit Gevolg 7.4.5 volgt dat de groep $(\mathbb{Z}/p\mathbb{Z})^*$ cyclisch is van orde $p - 1 \equiv 0 \pmod 4$. Dus 4 deelt de groepsorde, dus er is een element $\overline{x}$ van orde 4. Dus $\overline{x}^2$ heeft orde 2, net als $-\overline{1}$. Maar er is maar $\phi(2) = 1$ element van orde 2, dus $\overline{x}^2 = -\overline{1}$, ofwel $x^2 \equiv -1 \pmod p$. Dus

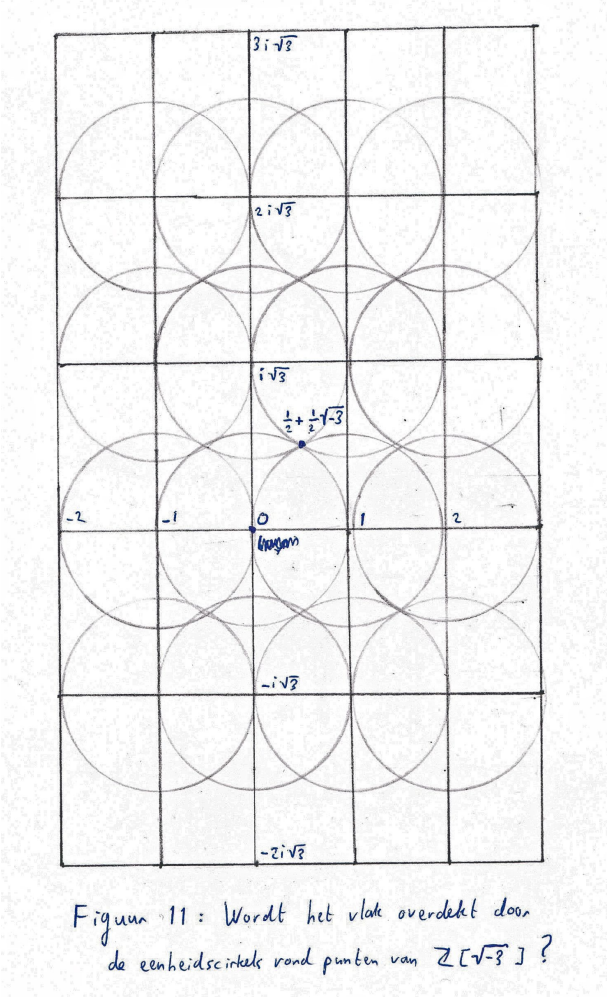
$$p | x^2 + 1, \qquad \text{ofwel} \qquad p | (x - i)(x + i)$$

met $x - i$ en $x + i$ elementen van $\mathbb{Z}[i]$. De absolute waarde van het imaginaire deel van deze twee elementen is 1, en die van $p\alpha$ is $p|b| \neq 1$ voor willekeurige $\alpha = a + bi \in \mathbb{Z}[i]$; dus $p\alpha \neq x \pm 1$ voor elke α, met andere woorden p deelt niet $x - i$ of van $x + i$. Dus p is *niet* priem, ofwel p is reducibel. Zoals we zagen is er dan een $\alpha \in \mathbb{Z}[i]$ zodat $p = \alpha\overline{\alpha}$ de priemontbinding van p is. Dit bewijst 2). □

Het bepalen van de priemontbinding van een willekeurig element is nu niet meer zo moeilijk, we gaan hier verder niet op in. Zie bijvoorbeeld [8] voor een paar voorbeelden. We hebben nu meteen deel a) van Stelling 8.0.6 bewezen: elk priemgetal p die een viervoud plus één is, is te ontbinden als $\alpha\overline{\alpha}$ met $\alpha = a + bi \in \mathbb{Z}[i]$, ofwel $p = a^2 + b^2$ met a, b geheel.

8.2 Gehele getallen van Eisenstein

[3] Ons bewijs is een bewerking van [8], Stelling 12.20.



Figuur 11: Wordt het vlak overdekt door de eenheidscirkels rond punten van $\mathbb{Z}[\sqrt{-3}]$?

Om Fermat's laatste stelling te bewijzen voor $n = 3$, zou het handig zijn als er sprake is van unieke priemfactorisatie in de ring $\mathbb{Z}[\sqrt{-3}]$. We kunnen proberen te bewijzen dat het een hoofdideaaldomein is door op dezelfde manier als net deling met rest proberen toe te passen. Als norm nemen we weer het kwadraat van de absolute waarde, dus $\|a + b\sqrt{-3}\| = a^2 + 3b^2$; dit is weer een geheel getal, en de norm voldoet aan $\|\alpha\beta\| = \|\alpha\| \cdot \|\beta\|$. Zij $\alpha, \beta \in \mathbb{Z}[\sqrt{-3}]$ en $\beta \neq 0$. We zoeken $q, r \in \mathbb{Z}[\sqrt{-3}]$ zodat $\alpha = q\beta + r$ en $\|r\| < \|\beta\|$. Dit impliceert zoals net dat

$$\|\frac{\alpha}{\beta} - q\| = \frac{\|\alpha - q\beta\|}{\|\beta\|} = \frac{\|r\|}{\|\beta\|} < 1. \qquad (8.1)$$

We kunnen dit weer intuïtief proberen in te zien door $\mathbb{Z}[\sqrt{-3}]$ te visualiseren als 'standaardrooster' waarin de y-as uitgerekt is met factor $\sqrt{3}$. Het is nu echter veel minder duidelijk of de eenheidscirkels rond deze roosterpunten het complexe vlak opvullen, zie Figuur 11. In bijvoorbeeld $\mathbb{Z}[\sqrt{-5}]$ is de y-as nog verder uitgerekt, en het is duidelijk dat de eenheidscirkels rond deze roosterpunten het complexe vlak niet meer overdekken: onze methode van restdeling werkt daar niet. Het lijkt alsof $\mathbb{Z}[\sqrt{-3}]$ een soort grensgeval is, maar is er net wel of net niet restdeling mogelijk? De duidelijkste kandidaten voor eventuele punten die niet in het inwendige van een cirkel van Figuur 11 liggen, zijn die midden tussen vier roosterpunten, bijvoorbeeld $\frac{1}{2} + \frac{1}{2}\sqrt{-3}$. Als we in het bovenstaande $\alpha = 1 + \sqrt{-3}$ en $\beta = 2$ nemen, dan is $\alpha/\beta = \frac{1}{2} + \frac{1}{2}\sqrt{-3}$. Schrijven we $q = c + d\sqrt{-3}$ met c, d geheel, dan is $\frac{1}{2} - c, \frac{1}{2} - d \geq \frac{1}{2}$ voor alle mogelijke q, en dus

$$\|\alpha/\beta - q\| = \|(\tfrac{1}{2} - c) + (\tfrac{1}{2} - d)\sqrt{-3}\| = (\tfrac{1}{2} - c)^2 + 3(\tfrac{1}{2} - d)^2 \geq (\tfrac{1}{2})^2 + 3(\tfrac{1}{2})^2 = 1.$$

Er is dus geen $q \in \mathbb{Z}[\sqrt{-3}]$ zodat voldaan is aan (8.1), dus restdeling van α door β is op deze manier niet mogelijk. We kunnen niet zo snel een andere functie dan de norm verzinnen die aan de voorwaarden van f in Stelling 7.3.2 voldoet. Dat betekent natuurlijk niet meteen dat $\mathbb{Z}[\sqrt{-3}]$ geen hoofdideaaldomein is, maar we beginnen er wel aan te twijfelen.

Daarom pakken we het anders aan. We kijken naar een ring waarvan we wél kunnen bewijzen dat het een hoofdideaaldomein is, en waar bovendien $\mathbb{Z}[\sqrt{-3}]$ een deelring van is. Een geschikte ring is die van de *gehele getallen van Eisenstein*. Gotthold Eisenstein (1823 – 1852) was een Duitse wiskundige die (ondanks zijn vroege dood veroorzaakt door tuberculose)

belangrijk werk op het gebied van getaltheorie en analyse heeft verricht. De gehelen van Eisenstein hebben een aantal zo mooie eigenschappen met betrekking tot Fermat's laatste stelling dat het haast geen 'toeval' kan zijn: het lijkt alsof er een dieperliggende structuur achter zit. De getallen gaan niet uit van i, maar van het getal

$$\psi = -\tfrac{1}{2} + \tfrac{1}{2}\sqrt{-3} = e^{2\pi i/3} = \zeta_3.$$

Hierbij is ζ_3 de 'standaard derdemachts eenheidswortel'. In het algemeen definiëren we $\zeta_n = e^{2\pi i/n}$. De n-de machts eenheidswortels, dat zijn de complexe getallen α waarvoor $\alpha^n = 1$, zijn precies de n machten $1, \zeta_n, \zeta_n^2, \zeta_n^3, \ldots, \zeta_n^{n-1}$ van ζ_n. Meetkundig vormen ze een regelmatige n-hoek rond de oorsprong, ingeschreven in de eenheidscirkel.

We merken op dat ψ nulpunt is van de veelterm $X^2 + X + 1$. Er geldt namelijk dat $\psi^2 = \psi^{-1} = \overline{\psi}$, dus

$$\psi^2 + \psi + 1 = \overline{\psi} + \psi + 1 = 2\mathrm{Re}(\psi) + 1 = -1 + 1 = 0.$$

De ring van gehele getallen van Eisenstein is

$$\mathbb{Z}[\psi] = \{a + b\psi : a, b \in \mathbb{Z}\}.$$

Om te bewijzen dat dit een ring is, hoeven we alleen na te gaan dat het een deelring is van $\mathbb{C}$. Dat is triviaal, behalve de inwendigheid van vermenigvuldiging. Zij $a+b\psi, c+d\psi \in \mathbb{Z}[\psi]$. Hun product is $(a + b\psi)(c + d\psi) = ac + (ad + bc)\psi + bd\psi^2$. Omdat

$$\psi^2 = \overline{\psi} = -\tfrac{1}{2} - \tfrac{1}{2}\sqrt{-3} = -1 - (-\tfrac{1}{2} + \tfrac{1}{2}\sqrt{-3}) = -1 - \psi$$

is $bd\psi^2 = -bd - bd\psi$. We hebben dus

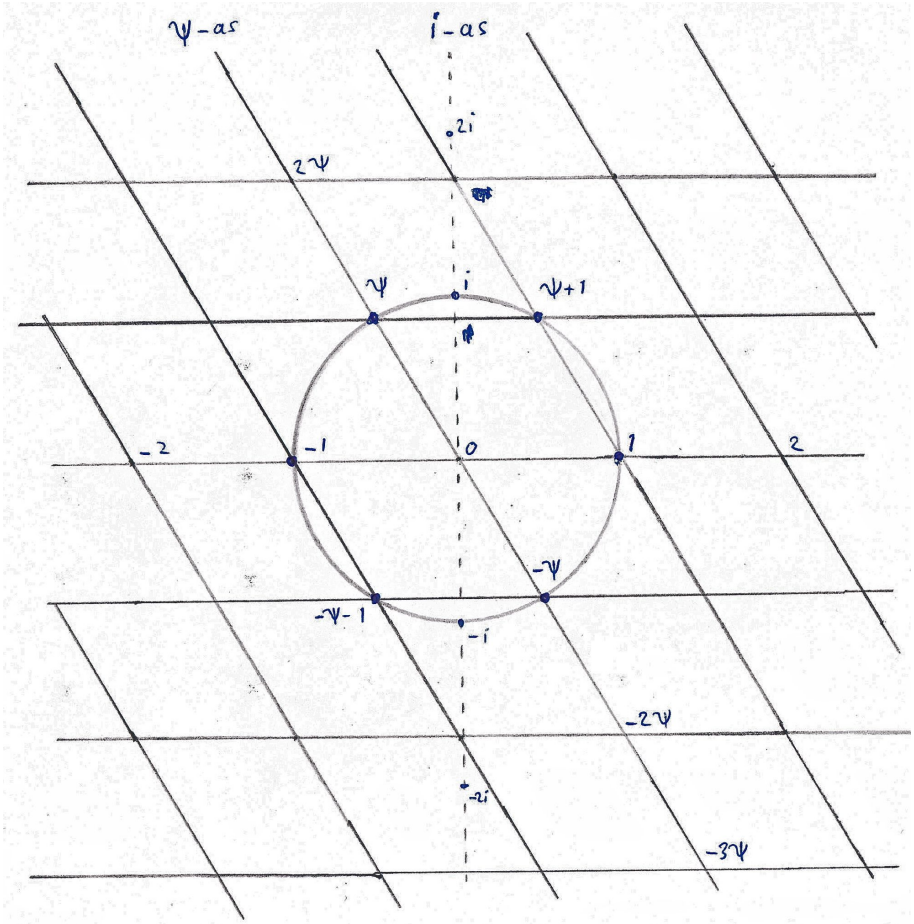
$$(a + b\psi)(c + d\psi) = ac + (ad + bc)\psi + bd\psi^2 = (ac - bd) + (ad + bc - bd)\psi \in \mathbb{Z}[\psi].$$

Vermenigvuldiging is dus inwendig, en we hebben een ring. Meetkundig kunnen we ons $\mathbb{Z}[\psi]$ voorstellen als het 'standaardrooster' in $\mathbb{C}$, maar dan met de y-as (lees: ψ-as) dertig graden naar links gekanteld. Zie Figuur 12 voor een schets. Ons bewijs dat het een hoofdideaaldomein is, gaat ongeveer net zo als bij de gehelen van Gauss. We definiëren weer de norm $\|\cdot\|$ op $\mathbb{Z}[\psi]$ als het kwadraat van de gewone absolute waarde. Nu geldt echter niet $\|a + b\psi\| = a^2 + b^2$, want ψ is niet zuiver imaginair. We hebben wel

$$\begin{aligned} \|a + b\psi\| &= |a + b(-\tfrac{1}{2} + \tfrac{1}{2}\sqrt{3}i)|^2 = |(a - \tfrac{1}{2}b) + (\tfrac{1}{2}\sqrt{3}b)i|^2 \\ &= (a - \tfrac{1}{2}b)^2 + (\tfrac{1}{2}\sqrt{3}b)^2 = a^2 - ab + \tfrac{1}{4}b^2 + \tfrac{3}{4}b^2 = a^2 - ab + b^2, \end{aligned} \tag{8.2}$$

dus de norm van een element in $\mathbb{Z}[\psi]$ is weer een *geheel getal*. Bovendien is het als kwadraat van de absolute waarde positief. Nu kunnen we met restdeling bewijzen dat $\mathbb{Z}[\psi]$ een hoofdideaaldomein is.

Stelling 8.2.1. $\mathbb{Z}[\psi]$ *is een hoofdideaaldomein.*



Figuur 12 : De gehelen van Eisenstein. De eenheden vormen de hoekpunten van een regelmatige zeshoek in de eenheidscirkel.

Bewijs. Zij $z \in \mathbb{C}$. We beweren dat er een $q \in \mathbb{Z}[\psi]$ is zodat $\|z - q\| < 1$. Meetkundig betekent dit dat de open eenheidscirkels rond het vervormde standaardrooster van Figuur 12 het hele complexe vlak overdekken, en Figuur 13 maakt dit intuïtief duidelijk.[4] Om het echt te bewijzen, schrijven we $z = c + di$ met $c, d \in \mathbb{R}$. Neem een geheel getal n zo dat $|\frac{2}{\sqrt{3}}d - n| \leq \frac{1}{2}$, en neem vervolgens $m \in \mathbb{Z}$ zodat $|(c + \frac{1}{2}n) - m| \leq \frac{1}{2}$. Er geldt dus dat

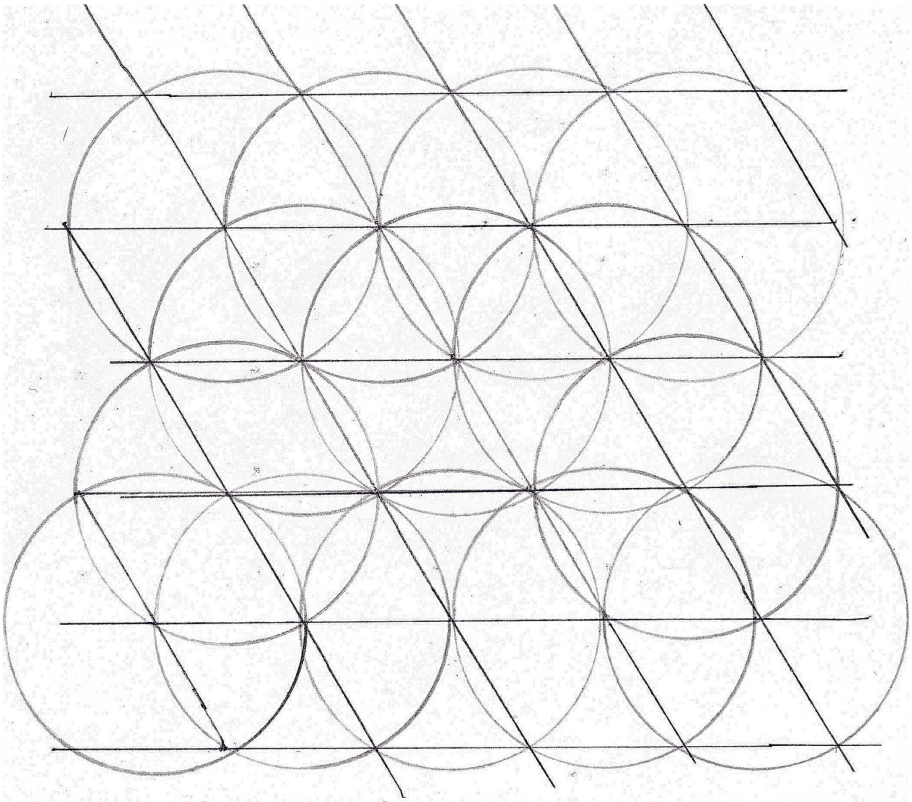
$$|d - \tfrac{1}{2}\sqrt{3}n| = \tfrac{1}{2}\sqrt{3} \cdot |\tfrac{2}{\sqrt{3}}d - n| \leq \tfrac{1}{4}\sqrt{3} \qquad \text{en} \qquad |c - (m - \tfrac{1}{2}n)| \leq \tfrac{1}{2}.$$

We nemen $q = m + n\psi$, ofwel $q = (m - \frac{1}{2}n) + (\frac{1}{2}\sqrt{3}n)i$. Er volgt dat

$$\|z - q\| = (c - (m - \tfrac{1}{2}))^2 + (d - \tfrac{1}{2}\sqrt{3}n)^2 \leq (\tfrac{1}{2})^2 + (\tfrac{1}{4}\sqrt{3})^2 = \tfrac{4}{16} + \tfrac{3}{16} = \tfrac{7}{16} < 1,$$

zoals gewenst.

[4] Correctie op de tekening: er missen twee cirkels in de derde rij.



Stel nu dat $\alpha, \beta \in \mathbb{Z}[\psi]$ en $\beta \neq 0$. Neem een $q \in \mathbb{Z}[\psi]$ zodat $\|\alpha/\beta - q\| < 1$, en definieer $r = \alpha - q\beta$. Dan is

$$\alpha = q\beta + r \qquad \text{en} \qquad \|r\| = \|\beta(\alpha/\beta - q)\| = \|\beta\| \cdot \|\alpha/\beta - q\| < \|\beta\|,$$

dus de functie $\|\cdot\|$ voldoet aan alle voorwaarden van f is Stelling 7.3.2. We concluderen: $\mathbb{Z}[\psi]$ is een hoofdideaaldomein. □

8.2.1 Priemelementen

Door de priemelementen van $\mathbb{Z}[\psi]$ te bepalen, kunnen we ook deel b) van Fermat's bewering bewijzen, namelijk dat elk priemgetal van de vorm $3n+1$ te schrijven is als $a^2 + 3b^2$. We gaan weer net zo te werk als in $\mathbb{Z}[i]$. Eerst bepalen we de eenheden. Als $\|u\| = 1$, dan is $u\overline{u} = 1$, dus u is een eenheid. Omgekeerd, als u een eenheid is, dan is er een $\alpha \in \mathbb{Z}[\psi]$ zodat $u\alpha = 1$. Dus $\|u\| \cdot \|\alpha\| = 1$, en dus $\|u\| = 1$. De eenheden zijn dus de roosterpunten van Figuur 12 die op de eenheidscirkel liggen, en we zien uit het plaatje dat dit er zes zijn:

$$\begin{aligned}\mathbb{Z}[\psi]^* &= \{1, 1+\psi, \psi, -1, -1-\psi, -\psi\} = \{1, \tfrac{1}{2} + \tfrac{1}{2}\sqrt{3}i, -\tfrac{1}{2} + \tfrac{1}{2}\sqrt{3}i, -1, -\tfrac{1}{2} - \tfrac{1}{2}\sqrt{3}i, \tfrac{1}{2} - \tfrac{1}{2}\sqrt{3}i\} \\ &= \{1, e^{\pi i/3}, e^{2\pi i/3}, -1, e^{4\pi i/3}, e^{5\pi i/3}\} = \{\zeta_6^0, \zeta_6^1, \zeta_6^2, \zeta_6^3, \zeta_6^4, \zeta_6^5\}. \qquad (8.3)\end{aligned}$$

De eenhedengroep is dus inderdaad cyclisch. Als $\|\alpha\| = p$ met p een priemgetal in $\mathbb{Z}$, dan is α een priemelement. Dit volgt weer door op te merken dat $\alpha = \beta\gamma$ impliceert dat $\|\beta\| \cdot \|\gamma\| = p$, dus β of γ heeft norm 1 en is dus een eenheid. Stel dat δ een priemelement is. Ontbinden we het natuurlijke getal $\delta\overline{\delta} = \|\delta\| = p_1 p_2 \dots p_n$ in priemgetallen p_i, dan volgt uit $\delta | \delta\overline{\delta} = p_1 p_2 \dots p_n$ dat δ een van de p_i deelt. Om de priemelementen van $\mathbb{Z}[\psi]$ te vinden, hoeven we dus weer alleen de priemgetallen van $\mathbb{Z}$ te ontbinden in $\mathbb{Z}[\psi]$.

Stelling 8.2.2. *Een priemgetal p van $\mathbb{Z}$ is als volgt te ontbinden in priemelementen van $\mathbb{Z}[\psi]$.*

1. *Het priemgetal 3 heeft de ontbinding $-1 \cdot (1 + 2\psi)^2$.*
2. *Als $p \equiv 1 \pmod 3$, dan is $p = \alpha\overline{\alpha}$ met $\alpha, \overline{\alpha}$ priemelementen van norm p.*
3. *Als $p \equiv 2 \pmod 3$, dan is p irreducibel.*

Bewijs. Door gebruik te maken van $\psi^2 = \overline{\psi}$ zien we dat

$$-1 \cdot (1 + 2\psi)^2 = -(1 + 4\psi + 4\overline{\psi}) = -(1 + 4 \cdot 2\mathrm{Re}(\psi)) = -(1 + 8 \cdot -\frac{1}{2}) = -(1 - 4) = 3.$$

Omdat $\|1 + 2\psi\| = 1^2 - 1 \cdot 2 + 2^2 = 3$ een priemgetal is, is $1 + 2\psi$ een priemelement. Omdat bovendien -1 een eenheid is, volgt dat $-1 \cdot (1 + 2\psi)^2$ inderdaad een priemontbinding is van 3. Hiermee is deel 1 bewezen.

Stel dat p reducibel is in $\mathbb{Z}[\psi]$. Dan is $p = \alpha\beta$ met α, β geen eenheden, en we zien opnieuw dat $p^2 = \|\alpha\| \cdot \|\beta\|$ zodat $\|\alpha\| = \|\overline{\alpha}\| = p$, en dus $p = \alpha\overline{\alpha}$ voor irreducibele $\alpha, \overline{\alpha}$. We moeten dus weer nagaan wanneer p reducibel is.

We merken eerst op dat er voor elke $\alpha \in \mathbb{Z}[i]$ gehele getallen a, b zijn zodat $\alpha\overline{\alpha} = a^2 + 3b^2$. Om dit te bewijzen, schrijven we $\alpha = c + d\psi$ met $c, d \in \mathbb{Z}$, zodat $\alpha\overline{\alpha} = c^2 - cd + d^2$. Als d even is, is $\frac{1}{2}d$ geheel, en door haakjes uit te werken[5] zien we dat $\alpha\overline{\alpha} = (c - \frac{1}{2}d)^2 + 3(\frac{1}{2}d)^2$. Omdat $c^2 - cd + d^2$ symmetrisch is in c en d, kunnen we hetzelfde trucje toepassen als c even is: we hebben dan $\alpha\overline{\alpha} = (d - \frac{1}{2}c)^2 + 3(\frac{1}{2}c)^2$ met $\frac{1}{2}c$ geheel. Tenslotte, als c en d beide oneven zijn, dan zijn $c - d$ en $c + d$ even, zodat $\frac{1}{2}(c + d)$ en $\frac{1}{2}(c - d)$ geheel zijn. We hebben nu

$$\alpha\overline{\alpha} = c^2 - cd + d^2 = (c-d)^2 + cd = 3(\tfrac{1}{2}(c-d))^2 + (\tfrac{1}{2}(c-d))^2 + cd = 3(\tfrac{1}{2}(c-d))^2 + (\tfrac{1}{2}(c+d))^2.$$

We concluderen dat er altijd $a, b \in \mathbb{Z}$ zijn zodat $\alpha\overline{\alpha} = a^2 + 3b^2$.

Als een priemgetal p reducibel is, dan is er een $\alpha \in \mathbb{Z}[\psi]$ zodat $p = \alpha\overline{\alpha}$, en dus zijn er gehele a, b zodat $p = a^2 + 3b^2$. Maar kwadraten zijn altijd congruent 0 of 1 modulo 3, want $0^2 \equiv 0, 1^2 \equiv 1, 2^2 \equiv 1$. Door de vier (of twee) mogelijkheden na te gaan, zien we dat $a^2 + 3b^2$ congruent is aan 0 of 1 modulo 3. Als $p \equiv 2 \pmod 3$, dan is er dus niet zo'n α, dus p is irreducibel. Hiermee hebben we deel 3 bewezen.

Stel tenslotte dat $p \equiv 1 \pmod 3$. Dan is de eenhedengroep $(\mathbb{Z}/p\mathbb{Z})^*$ cyclisch van orde $p - 1$, en omdat $3 | p - 1$ bevat deze groep een element $\overline{x}$ van orde 3. Er geldt dus dat

$$(x - 1)(x^2 + x + 1) \equiv x^3 - 1 \equiv 1 - 1 \equiv 0 \pmod p.$$

[5] De formule volgt ook direct uit $\alpha = c + d\psi = (c - \frac{1}{2}d) + (\frac{1}{2}\sqrt{3}d)i$.

Bovendien is $x - 1 \not\equiv 0 \pmod{p}$, want dan zou $\overline{x}$ gelijk zijn aan $\overline{1}$ en orde 1 hebben in plaats van 3. Omdat $\mathbb{Z}/p\mathbb{Z}$ geen nuldelers heeft (want het is een lichaam), volgt dat $x^2 + x + 1 \equiv 0 \pmod{p}$, ofwel $p|x^2 + x + 1$. We weten dat ψ nulpunt is van het polynoom $X^2 + X + 1$, dus we proberen met restdeling een factor $X - \psi$ buiten haakjes te halen. Gebruik makend van $(1+\psi)\psi = e^{\pi i/3}e^{2\pi i/3} = e^{\pi i} = -1$ vinden we de ontbinding $X^2+X+1 = (X-\psi)(X+(1+\psi))$. Ter controle:

$$(X-\psi)(X+(1+\psi)) = X^2-\psi X+(1+\psi)X-\psi(1+\psi) = X^2-\psi X+X+\psi X--1 = X^2+X+1.$$

We hebben dus

$$p|x^2 + x + 1, \qquad \text{ofwel} \qquad p|(x - \psi)((x + 1) + \psi)$$

met $x - \psi$ en $(x + 1) + \psi$ elementen van $\mathbb{Z}[\psi]$ (want x is geheel). Het imaginaire deel van deze twee elementen is $\frac{1}{2}\sqrt{3}$ in absolute waarde, terwijl dat van $(a+b\psi)p$ gelijk is aan $|b|p\frac{1}{2}\sqrt{3} \neq \frac{1}{2}\sqrt{3}$ voor willekeurige $a+b\psi \in \mathbb{Z}[\psi]$. Dat betekent dat p geen deler is van $(x-\psi)$ of van $((x+1)+\psi)$ maar wel van het product, dus p is reducibel. Zoals we zagen is er dus een $\alpha \in \mathbb{Z}[\psi]$ zodat $p = \alpha\overline{\alpha}$ de priemontbinding is van p. Dit bewijst deel 2. □

Zoals het bewijs laat zien, is elk priemgetal van de vorm $3n + 1$ te schrijven als $\alpha\overline{\alpha}$, en dat is weer te schrijven als $a^2 + 3b^2$ met a, b geheel. Hiermee hebben we ook deel b van Fermat's stelling 8.0.6 bewezen.

8.2.2 Is $\mathbb{Z}[\sqrt{-3}]$ een hoofdideaaldomein?

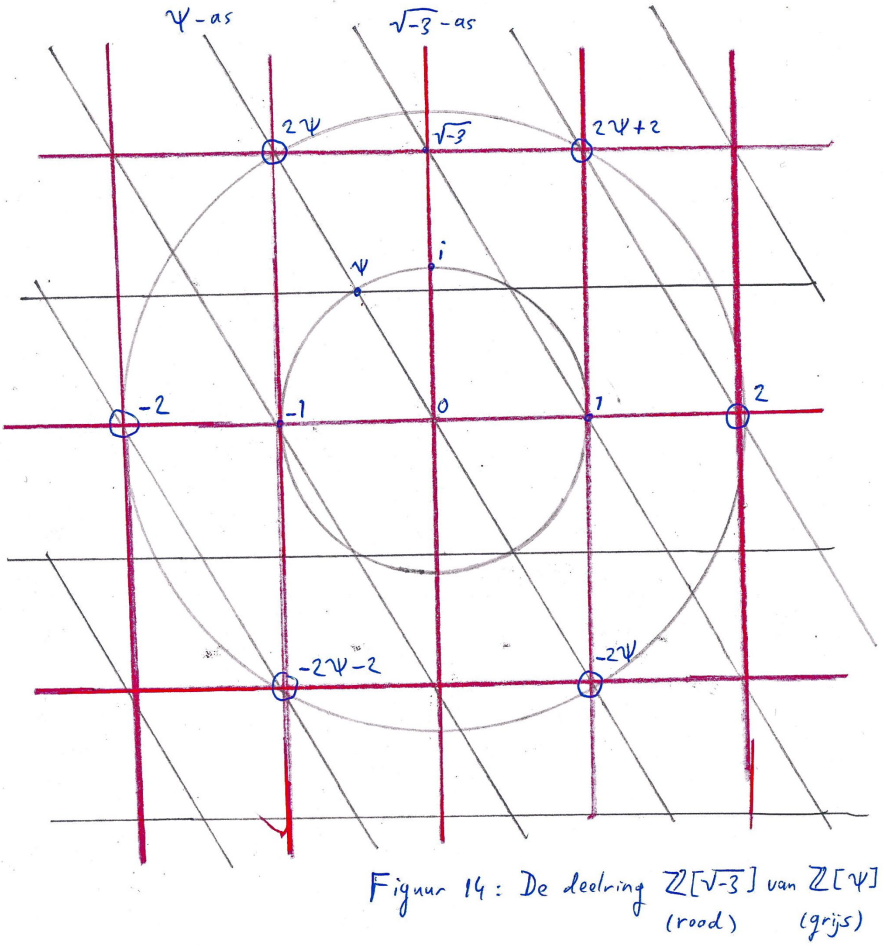
We hebben $\mathbb{Z}[\psi]$ onderzocht omdat we geïnteresseerd zijn in zijn deelring $\mathbb{Z}[\sqrt{-3}]$. We gaan eerst na dat het inderdaad een deelring is. Alle $a + b\sqrt{-3} \in \mathbb{Z}[\sqrt{-3}]$ kunnen we schrijven als $(a + b) + (2b)(-\frac{1}{2} + \frac{1}{2}\sqrt{-3}) = (a + b) + (2b)\psi \in \mathbb{Z}[\psi]$. Een element van $\mathbb{Z}[\sqrt{-3}]$ is dus van de vorm $c + d\psi$ met d even. Omgekeerd, als $c + d\psi \in \mathbb{Z}[\psi]$ met d even, dan is $c + d\psi = (c - \frac{1}{2}d) + (\frac{1}{2}d\sqrt{-3}) \in \mathbb{Z}[\sqrt{-3}]$. We concluderen:

$$\mathbb{Z}[\sqrt{-3}] = \{c + d\psi \in \mathbb{Z}[\psi] : \ d \text{ is even } \} \subset \mathbb{Z}[\psi].$$

Met deze karakterisering volgt vrijwel meteen dat $\mathbb{Z}[\sqrt{-3}]$ een deelring is van $\mathbb{Z}[\psi]$, maar dat hoeven we niet te controleren. We weten namelijk al dat $\mathbb{Z}[\sqrt{-3}]$ een ring is met dezelfde bewerkingen als in $\mathbb{Z}[\psi]$ (namelijk die van $\mathbb{C}$), en omdat het er een deelverzameling van is, is het een deelring. Zie Figuur 14 ter illustratie: we hebben hier $\mathbb{Z}[\psi]$ getekend met daarin de deelring $\mathbb{Z}[\sqrt{-3}]$. Merk op dat de punten van $\mathbb{Z}[\psi]$ die wel of niet tot $\mathbb{Z}[\sqrt{-3}]$ behoren, gelijk verdeeld zijn: de 'even rijen' behoren tot $\mathbb{Z}[\sqrt{-3}]$, de oneven rijen niet. We kunnen nu inzien dat $\mathbb{Z}[\sqrt{-3}]$ *geen* hoofdideaaldomein is. Geïnspireerd door de observatie dat restdeling van $1 + \sqrt{-3}$ door 2 misgaat, bekijken we het ideaal $I = (2, 1 + \sqrt{-3}) = (2, 2\psi)$. Stel dat I een hoofdideaal van $\mathbb{Z}[\sqrt{-3}]$ is, zeg $I = (m)$ met $m \in \mathbb{Z}[\sqrt{-3}]$. Omdat $2 \in I$ is er een $\gamma \in \mathbb{Z}[\sqrt{-3}]$ zodat $2 = \gamma m$. Dus $4 = \|2\| = \|\gamma\| \cdot \|m\|$, dus $\|m\|$ deelt 4. Aan de andere kant, zoals elk element van I is m van de vorm $\alpha \cdot 2 + \beta \cdot 2\psi$ met $\alpha, \beta \in \mathbb{Z}[\sqrt{-3}]$. Dus $\|m\| = \|2(\alpha + \beta\psi)\| = 4\|\alpha + \beta\psi\|$, en omdat $\alpha + \beta\psi \in \mathbb{Z}[\psi]$ is $\|\alpha + \beta\psi\|$ geheel, dus 4 deelt $\|m\|$. We concluderen dat $\|m\| = 4$; in Figuur 14 hebben we de zes mogelijkheden voor m omcirkeld, waaronder 2 en 2ψ.

Uit $\gamma m = 2$ volgt

$$4\|\gamma\| = \|\gamma\| \cdot \|m\| = \|\gamma m\| = \|2\| = 4,$$



Figuur 14: De deelring $\mathbb{Z}[\sqrt{-3}]$ van $\mathbb{Z}[\psi]$ (rood) (grijs)

dus $\|\gamma\| = 1$, en omdat $\gamma \in \mathbb{Z}[\sqrt{-3}]$ lezen we uit Figuur 14 af dat $\gamma = 1$ of $\gamma = -1$. Dus uit $\gamma m = 2$ volgt $m = 2$ of $m = -2$. Maar ook $\delta m = 2\psi$ voor een $\delta \in \mathbb{Z}[\sqrt{-3}]$, want $2\psi \in (m)$. Het argument van boven kunnen we dus toepassen op δ in plaats van γ en 2ψ in plaats van 2, want ook 2ψ heeft norm 4. We vinden dus $m = 2\psi$ of $m = -2\psi$, in tegenspraak met $m = \pm 2$. We concluderen dat I geen hoofdideaal is, en dus $\mathbb{Z}[\sqrt{-3}]$ geen hoofdideaaldomein. Onze stelling over eenduidige priemfactorisatie gaat dus niet op. We kunnen de situatie redden door de elementen van $\mathbb{Z}[\sqrt{-3}]$ te ontbinden in $\mathbb{Z}[\psi]$, die wél een hoofdideaal is.

8.3 Een nieuwe poging voor het geval $n = 3$

Het enige wat we nog moeten doen om FLT te bewijzen voor $n = 3$, is een bewijs geven van Lemma 4.2.1. Die zegt dat als a en b copriem en van tegengestelde pariteit zijn, en $a^2 + 3b^2$ een derde macht, dan zijn er $p, q \in \mathbb{Z}$ zodat $a = p^3 - 9pq^2$ en $b = 3p^2q - 3q^2$. We zagen al dat dit equivalent is met

Stelling 8.3.1. *Stel a en b zijn copriem en van tegengestelde pariteit,*[6] *en $a^2+3b^2 = (a+b\sqrt{-3})(a-b\sqrt{-3})$ is een derde macht van een geheel getal. Dan is $a+b\sqrt{-3}$ een derde macht in $\mathbb{Z}[\sqrt{-3}]$. Met andere woorden, er zijn gehele getallen p, q zodat*

$$a+b\sqrt{-3} = (p+q\sqrt{-3})^3.$$

Bewijs. We schrijven $\alpha = a+b\sqrt{-3}$, en gaan deze ontbinden in $\mathbb{Z}[\psi]$. Om te beginnen schrijven we

$$\alpha = (a+b)+(2b)(-\tfrac{1}{2}+\tfrac{1}{2}\sqrt{-3}) = c+2d\psi$$

met $c = a+b, d = b$. Zij μ een priemelement in de priemfactorontbinding van α in $\mathbb{Z}[\psi]$; we proberen extra beperkingen aan μ op te leggen. We merkten al op dat elk priemelement van $\mathbb{Z}[\psi]$ een priemgetal van $\mathbb{Z}$ deelt, dus μ treedt op als factor in één van de drie soorten ontbindingen uit Stelling 8.2.2. Dat wil zeggen, op vermenigvuldiging met eenheden na is

1. $\mu = 1+2\psi$, of
2. $\mu\overline{\mu} = p$ voor een priemgetal $p \equiv 1 \pmod 3$, of
3. $\mu = p$ voor een priemgetal $p \equiv 2 \pmod 3$.

In geval 3 is p dus deler van α, dus er zijn e, f zodat $\alpha = c+2d\psi = p(e+f\psi) = pe+pf\psi$. Dus $c = pe$ en $2d = pf$, en dus $p|c, p|2d$. Als $p \neq 2$, dan volgt uit $p|2d$ dat $p|d$, dus p is priemdeler van $c = a+b$ en $d = b$. Dus p deelt a en b, tegenspraak want die zijn copriem. Conclusie: in geval 3 kan p alleen 2 zijn.

In geval 1 is $\|\mu\| = \mu\overline{\mu} = 3$, in geval 2 is $\|\mu\| = p$. De norm van μ is in deze gevallen dus een oneven *priemgetal*, en bovendien is $\mu\overline{\mu}$ een priemontbinding van dit priemgetal. We concluderen dat we α kunnen schrijven als

$$\alpha = c+2d\psi = u\cdot 2^m \cdot \pi_1\pi_2\cdots\pi_n$$

met $u \in \mathbb{Z}[\psi]^*$ een eenheid en de π_k niet noodzakelijk verschillende priemelementen met als norm een oneven priemgetal. We weten dat $a^2+3b^2 = \|\alpha\|$ een derde macht is, dus

$$a^2+3b^2 = \|u\cdot 2^m\cdot\pi_1\pi_2\cdots\pi_n\| = \|u\|\cdot\|2^m\|\cdot\|\pi_1\|\cdot\|\pi_2\|\cdots\|\pi_n\| = 2^{2m}\cdot\|\pi_1\|\cdot\|\pi_2\|\cdots\|\pi_n\|$$

is een derdemacht. Maar omdat 2 priem is en de getallen $\|\pi_1\|$ oneven priem, staat hier de priemontbinding van a^2+3b^2 in $\mathbb{Z}$! Uit het bewijs van Lemma 2.2.1 weten we dat de orde[7] van elke priemfactor in een derde macht veelvoud is van 3. De orde van 2 in a^2+3b^2 is $2m$ (want de $\|\pi_1\|$ zijn oneven), dus $3|2m$, en dus ook $3|m$, zeg $m = 3l$. De overige priemfactoren $\|\pi_k\|$ kunnen we in groepjes van 3 gelijke priemgetallen verdelen, we kunnen $\alpha\overline{\alpha}$ dus eventueel na hernummering schrijven als

$$\alpha\overline{\alpha} = 4^{3l}(\|\pi_1\|\cdot\|\pi_2\|\cdot\|\pi_3\|)\cdot(\|\pi_4\|\cdot\|\pi_5\|\cdot\|\pi_6\|)\cdots(\|\pi_7\|\cdot\|\pi_8\|\cdot\|\pi_9\|)$$

[6] De stelling is ook waar als a en b niet van tegengestelde pariteit zijn. Zie hoofdstuk 2 van [4] voor een bewijs.

[7] De orde van een priemgetal in een geheel getal is de multipliciteit waarmee dat priemgetal in de priemfactorontbinding voorkomt.

waarbij de priemgetallen tussen haakjes aan elkaar gelijk zijn, dus bijvoorbeeld $\|\pi_1\| = \|\pi_2\| = \|\pi_3\|$. Maar dat betekent dat bijvoorbeeld $\pi_1\overline{\pi_1} = \pi_2\overline{\pi_2} = \pi_3\overline{\pi_3}$ drie verschillende priemontbindingen zijn van hetzelfde priemgetal, en wegens eenduidigheid van de ontbinding hebben we $\pi_2 = v\pi_1$ of $\pi_2 = v\overline{\pi_1}$ voor een eenheid v. Als $\pi_2 = v\overline{\pi_1}$ zou gelden, dan is $c+2d\psi$ deelbaar door $\pi_1\pi_2 = v\pi_1\overline{\pi_1} = vp$ voor een oneven priemgetal p, en dus ook deelbaar door p. We zagen net dat dit tot de tegenspraak leidt dat a en b niet copriem zijn. Dus $\pi_2 = v_1\pi_1$ voor een eenheid v_1. Met hetzelfde argument volgt dat $\pi_3 = w_1\pi_1$ voor een eenheid w_1. De priemelementen π_1, π_2, π_3 zijn dus op eenheden na gelijk, en met dezelfde redenering volgt dit ook voor de andere groepjes van drie: $\pi_5 = v_4\pi_4$ en $\pi_6 = w_4\pi_4$; $\pi_8 = v_7\pi_7$ en $\pi_9 = w_7\pi_7$; et cetera. Uit $\alpha = u \cdot 2^m \cdot \pi_1\pi_2\cdots\pi_n$ en $m = 3l$ volgt nu

$$\begin{aligned}\alpha &= u \cdot 2^{3l} \cdot \pi_1^3 v_1 w_1 \cdot \pi_4^3 v_4 w_4 \cdot \pi_7^3 v_7 w_7 \cdots \pi_{n-2}^3 v_{n-2} w_{n-2} \\ &= w \cdot 2^{3l} \cdot \pi_1^3\pi_4^3\pi_7^3\cdots\pi_{n-2}^3 = w(2^l\pi_1\pi_4\pi_7\cdots\pi_{n-2})^3 \\ &= w(p+q\psi)^3\end{aligned} \tag{8.4}$$

waarbij $w = u\cdot v_1 w_1 v_4 w_4 \cdots v_{n-2} w_{n-2}$ een eenheid is, en $p+q\psi = 2^l\pi_1\pi_4\pi_7\cdots\pi_{n-2}$ een element van $\mathbb{Z}[\psi]$.

We stuiten nu tegen een probleem aan: α is een eenheid maal een derde macht, maar dat betekent nog niet dat α een derde macht is! We willen dus graag dat w een derde macht is. Een ander probleem is dat we weliswaar weten dat $p + q\psi$ een element is van $\mathbb{Z}[\psi]$, maar we willen dat het ook bevat is in de deelring $\mathbb{Z}[\sqrt{-3}]$. Gelukkig kunnen we na even nadenken deze problemen oplossen. Het tweede probleem kunnen we omzeilen door $p+q\psi$ te vermenigvuldigen met een eenheid, waardoor het een element van $\mathbb{Z}[\psi]$ wordt. Dit kunnen we als volgt inzien. Als q even is zijn we klaar: we weten dat dat betekent dat $p + q\psi \in \mathbb{Z}[\sqrt{-3}]$. Stel dus q is oneven. Als ook p oneven is, dan vermenigvuldigen we met de eenheid ψ:

$$\psi(p+q\psi) = p\psi + q\psi^2 = p\psi + q(-1-\psi) = -q + (p-q)\psi.$$

Nu is $p-q$ even, dus $\psi(p+q\psi) \in \mathbb{Z}[\sqrt{-3}]$. In het andere geval is p even, en we vermenigvuldigen met de eenheid $\psi + 1$: uit $(\psi+1)\psi = -1$ (teken een plaatje) volgt

$$(\psi+1)(p+q\psi) = p - q + p\psi \in \mathbb{Z}[\sqrt{-3}].$$

We concluderen dat er een eenheid e is zodat $e(p+q\psi) = r+s\sqrt{-3}$ voor gehele r, s. We kunnen (8.4) dus schrijven als

$$a + b\sqrt{-3} = v(r+s\sqrt{-3})^3 \tag{8.5}$$

met $v = we^{-3}$ een eenheid. We weten van (8.3) dat ofwel $v = \pm 1$, ofwel v is één van de vier getallen $\frac{1}{2}(\pm 1 \pm \sqrt{-3})$. We willen laten zien dat die laatste niet kunnen. We schrijven $(r+s\sqrt{-3})^3 = x + y\sqrt{-3}$ met x, y geheel. Als $v = \frac{1}{2}(\pm 1 \pm \sqrt{-3})$, dan is

$$a + b\sqrt{-3} = \tfrac{1}{2}(\pm 1 \pm \sqrt{-3})(x + y\sqrt{-3}) = \tfrac{1}{2}(\pm x \pm 3y) + \tfrac{1}{2}(\pm x \pm y)\sqrt{-3},$$

dus $b = \frac{1}{2}(\pm x \pm y)$. Dat betekent dat $\pm x \pm y$ even is, dus x en y hebben dezelfde pariteit. Omdat

$$x + y\sqrt{-3} = (r+s\sqrt{-3})^3 = r^3 - 9rs^2 + (3r^2s - 3s^3)\sqrt{-3},$$

volgt dat $x = r^3 - 9rs^2$ en $y = 3r^2s - 3s^2$ dezelfde pariteit hebben. Dus ook r en s moeten dezelfde pariteit hebben, anders klopt het modulo 2 niet. Als ze beide even zouden zijn, dan zou 8 deler zijn van x en van y, en dus 4 deler van a en van b. Maar a en b zijn copriem, tegenspraak. Dus r en s zijn beide oneven. We weten dat

$$a^2 + 3b^2 = \|a + b\sqrt{-3}\| = \|v\| \cdot \|r + s\sqrt{-3}\|^3 = (r^2 + 3s^2)^3,$$

en kwadraten van oneven getallen zijn altijd 1 modulo 4. Dus $(r^2 + 3s^2)^3 \equiv 1 + 3 \cdot 1 \equiv 0 \pmod 4$, en we concluderen dat ook $a^2 + 3b^2 \equiv 0 \pmod 4$. Maar a en b zijn van tegengestelde pariteit, dus ofwel $a^2 + 3b^2 \equiv 1 + 3 \cdot 0 \equiv 1 \pmod 4$, ofwel $a^2 + 3b^2 \equiv 0 + 3 \cdot 1 \equiv 3 \pmod 4$, tegenspraak. De aanname was dus onjuist: v is niet één van de vier eenheden $\frac{1}{2}(\pm 1 \pm 1\sqrt{-3})$, dus v is 1 of -1. Uit (8.5) volgt

$$a + b\sqrt{-3} = (r + s\sqrt{-3}) \qquad \text{of} \qquad a + b\sqrt{-3} = (-r - s\sqrt{-3})^3, \qquad r, s \text{ geheel},$$

dus $a + b\sqrt{-3}$ is een derde macht in $\mathbb{Z}[\sqrt{-3}]$. □

Hiermee is dus (eindelijk) de Laatste stelling van Fermat bewezen voor $n = 3$.

Hoofdstuk 9

Fermat's laatste stelling voor ontbindingsringen

In 1847 veroorzaakte Gabriel Lamé (1795 – 1870) veel opschudding door tijdens een voordracht enthousiast te verkondigen dat hij de Laatste stelling van Fermat volledig had bewezen. Zijn idee was als volgt. Hij merkte op dat in de bewijzen van de gevallen $n = 3, 4, 5, 7$ een algebraïsche factorisatie nodig was, zoals $x^3 + y^3 = (x + y)(x^2 - xy + y^2)$ in het geval $n = 3$. Voor grotere n worden dit steeds ingewikkeldere factorisaties, dus het lijkt er niet op dat dit een manier is om de stelling te bewijzen voor willekeurige n. Wat we echter wél kunnen proberen, zei hij, is $x^n + y^n$ *geheel* in n lineaire factoren te ontbinden in een deelring van $\mathbb{C}$. Als $z^n = x^n + y^n$, dan is het product van deze factoren een n-de macht, en Lamé beweerde dat elk van de factoren dan ook een n-de macht is. Hij had namelijk 'bewezen' dat in de deelring van $\mathbb{C}$ waar hij het over had, sprake is van unieke priemfactorisatie.[1]

Zijn argumenten waren erg vaag, en niet iedereen was even optimistisch. Achteraf bleken de argumenten inderdaad niet te kloppen, maar het idee is wel centraal in andere, wel succesvolle ontwikkelingen rond FLT. De factorisatie die Lamé voorstelde is

$$x^n + y^n = (x + y)(x + \zeta_n y)(x + \zeta_n^2 y) \cdots (x + \zeta_n^{n-1} y) \tag{9.1}$$

waarin $\zeta_n = e^{2\pi i/n}$ een n-de machts complexe eenheidswortel is, en n oneven. Deze ontbinding kunnen we afleiden door eerst de veelterm $X^n - 1$ te ontbinden in $\mathbb{C}$. De getallen $1, \zeta_n^1, \zeta_n^2, \zeta_n^3 \ldots, \zeta_n^{n-1}$ zijn n verschillende nulpunten van dit polynoom, en uit Stelling 7.4.3 volgt dat we hem kunnen ontbinden als

$$X^n - 1 = q(X - 1)(X - \zeta_n)(X - \zeta_n^2) \cdots (X - \zeta_n^{n-1}). \tag{9.2}$$

De graad van het linkerlid is n, de graad van het rechterlid is $\deg(q) + n$, dus er volgt dat q een constant polynoom is (ofwel $q \in \mathbb{C}$). Bovendien is de coëfficiënt van X^n links 1 en rechts

[1] De historische achtergrond in dit hoofdstuk is grotendeels ontleend aan [4], en veel bewijzen, inzichten en eigenlijk heel de opzet van dit hoofdstuk is geïnspireerd door dat boek.

q, dus $q = 1$. Als we $X = -x/y$ invullen, krijgen we

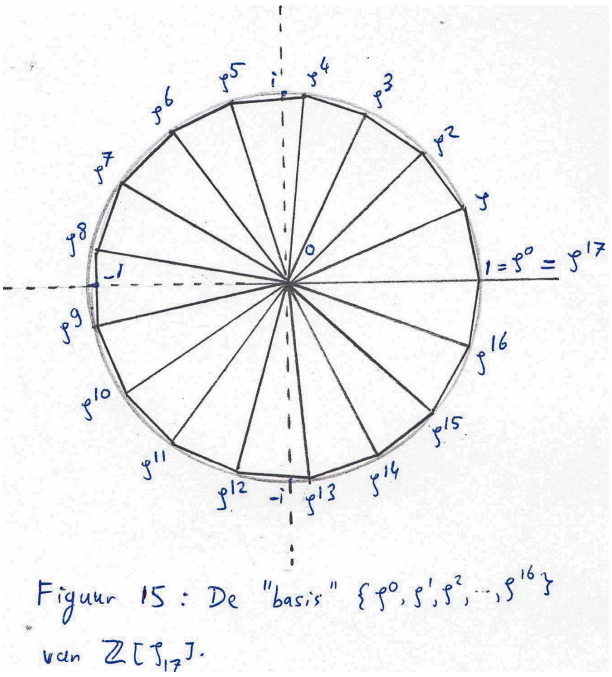
$$\begin{aligned}(-\frac{x}{y})^n - 1 &= (-\frac{x}{y} - 1)(-\frac{x}{y} - \zeta_n)(-\frac{x}{y} - \zeta_n^2)\cdots(-\frac{x}{y} - \zeta_n^{n-1}) \\ &= (-\frac{1}{y})^n(x+y)(x+\zeta_n y)(x+\zeta_n^2 y)\cdots(x+\zeta_n^{n-1}y). \qquad (9.3)\end{aligned}$$

Omdat n oneven is, is

$$(-\frac{x}{y})^n - 1 = (-\frac{1}{y})^n(x^n + y^n),$$

dus door in (9.3) aan beide kanten met $(-y)^n = -y^n$ te vermenigvuldigen, volgt de ontbinding (9.1).

9.1 Cyclotomische getallen



Figuur 15 : De "basis" $\{\zeta^0, \zeta^1, \zeta^2, \ldots, \zeta^{16}\}$ van $\mathbb{Z}[\zeta_{17}]$.

De deelring van $\mathbb{C}$ waar Lamé het over had, is de ring $\mathbb{Z}[\zeta_n]$ van *cyclotomische gehele getallen*. Dit is de ring die bestaat uit alle getallen van de vorm

$$a_0 + a_1\zeta_n + a_2\zeta_n^2 + \ldots + a_{n-1}\zeta_n^{n-1}$$

met alle a_k in $\mathbb{Z}$. Analoog zijn er ook cyclotomische rationale getallen, maar omdat we die niet gebruiken laten we vaak het woord 'gehele' weg en noemen ze gewoon 'cyclotomische getallen'. Bovendien moeten we eigenlijk een n specificeren als we het over cyclotomische getallen hebben, maar om de notatie kort te houden doen we dat niet als uit de context duidelijk wordt om welke n het gaat of als de waarde van n niet belangrijk is. Om dezelfde reden schrijven we vaak ζ_n gewoon als ζ.

De cyclotomische getallen zijn we, hoewel we ze niet zo hebben genoemd, al eerder tegengekomen. Een wat flauw voorbeeld is $\mathbb{Z}[\zeta_2]$ die bestaat uit de getallen van de vorm $a + \zeta_2 b = a - b$: dat is 'gewoon' $\mathbb{Z}$. Ook $\mathbb{Z}[\zeta_1]$ is gelijk aan $\mathbb{Z}$. Maar we kennen ook $\mathbb{Z}[\zeta_4]$ al. Deze bestaat uit de getallen van de vorm

$$a_0 + a_1 i + a_2 i^2 + a_3 i^3 = (a_0 - a_2) + (a_1 - a_3)i,$$

met a_0, a_1, a_2, a_3 geheel, en dit zijn elementen van $\mathbb{Z}[i]$, de ring van gehele getallen van Gauss. Omgekeerd wordt uit $a + bi = a + bi + 0i^2 + 0i^3$ duidelijk dat $\mathbb{Z}[i]$ deelverzameling is van $\mathbb{Z}[\zeta_4]$, dus $\mathbb{Z}[i] = \mathbb{Z}[\zeta_4]$. Overigens, er zou iets mis zijn met onze notatie als dit *niet* zo was geweest, want $i = \zeta_4$. Verder hebben we gewerkt met de ring $\mathbb{Z}[\psi]$ van gehele getallen van Eisenstein, de

getallen van de vorm $a+b(-\frac{1}{2}+\frac{1}{2}\sqrt{-3}) = a+b\zeta_3$ met a, b geheel. Omdat $a+b\zeta_3 = a+b\zeta_3+0\zeta_3^2$ is elk zo'n getal bevat in $\mathbb{Z}[\zeta_3]$. Omgekeerd volgt uit $\zeta_3^2 = \overline{\zeta_3} = -1-\zeta_3$ dat elk zo'n cyclotomisch getal $a_0 + a_1\zeta_3 + a_2\zeta_3^2$ gelijk is aan $(a_0 - a_2) + (a_1 - a_2)\zeta_3$, een gehele van Eisenstein. Dus inderdaad is $\mathbb{Z}[\psi] = \mathbb{Z}[\zeta_3]$.

Het is duidelijk hoe we met cyclotomische getallen kunnen rekenen (zonder ze uit te schrijven in $\mathbb{C}$): optelling gebeurt coördinaatsgewijs, bij vermenigvuldiging maken we (behalve van de standaard rekenregels als associativiteit en distributiviteit) gebruik van $\zeta^k \cdot \zeta^l = \zeta^{k+l}$ en $\zeta^n = 1$. We rekenen dus eigenlijk net als met polynomen $a_0 + a_1X + \ldots + a_{n-1}X^{n-1}$, maar nu lezen we de exponenten van X modulo n. Hiermee is direct duidelijk dat elke $\mathbb{Z}[\zeta_n]$ een deelring is van $\mathbb{C}$. Als voorbeeld voeren we een berekening uit in $\mathbb{Z}[\zeta]$ met $\zeta = \zeta_5$:

$$\begin{aligned}
&(1 + 3\zeta + 5\zeta^2 + 7\zeta^3 + 9\zeta^4)(2 + 4\zeta + 6\zeta^2 + 8\zeta^3 + 10\zeta^4) \\
&= 2 + (4 + 6)\zeta + (6 + 12 + 10)\zeta^2 + (8 + 18 + 20 + 14)\zeta^3 + (10 + 24 + 30 + 28 + 18)\zeta^4 \\
&\quad + (30 + 40 + 42 + 36)\zeta^5 + (50 + 56 + 54)\zeta^6 + (70 + 72)\zeta^7 + 90\zeta^8 \qquad (9.4) \\
&= 2 + 10\zeta + 28\zeta^2 + 60\zeta^3 + 110\zeta^4 + 148\zeta^0 + 160\zeta^1 + 142\zeta^2 + 90\zeta^3 \\
&= 150 + 170\zeta + 170\zeta^2 + 150\zeta^3 + 110\zeta^4.
\end{aligned}$$

Meetkundig vormen de machten ζ_n^k de hoekpunten van een regelmatige n-hoek ingeschreven in de eenheidscirkel, zie Figuur 15 waar we als voorbeeld $n = 17$ hebben genomen. De elementen van $\mathbb{Z}[\zeta]$ zijn precies de punten in het vlak die verkregen kunnen worden door een eindig aantal van deze vectoren 'kop aan staart' te leggen (dezelfde vector mag hierin vaker voorkomen). Bijvoorbeeld $1+\zeta+\zeta^2+\ldots+\zeta^{n-1}$ kunnen we 'berekenen' door $1, \zeta, \zeta^2, \ldots, \zeta^{n-1}$ achter elkaar kop aan staart leggen. We krijgen zo (voor $n \geq 3$) een regelmatige n-hoek, zie Figuur 16 waar we weer $n = 17$ hebben genomen. Zo'n n-hoek is gesloten, het beginpunt is hetzelfde als het eindpunt: de som zou dus nul moeten zijn. Dit kunnen we hard maken door op te merken dat

$$1 + \zeta + \zeta^2 + \ldots + \zeta^{n-1} = \zeta^n + \zeta + \zeta^2 + \ldots + \zeta^{n-1} = \zeta(1 + \zeta + \zeta^2 + \ldots + \zeta^{n-1}),$$

zodat

$$(1 - \zeta)(1 + \zeta + \zeta^2 + \ldots + \zeta^{n-1}) = 0.$$

Omdat $\zeta \neq 1$ en omdat $\mathbb{C}$ geen nuldelers heeft, volgt dat inderdaad $1+\zeta+\zeta^2+\ldots+\zeta^{n-1} = 0$. We kunnen daarom bij elke coördinaat van een cyclotomisch getal dezelfde constante optellen:

$$\begin{aligned}
&a_0 + a_1\zeta + a_2\zeta^2 + \ldots + a_{n-1}\zeta^{n-1} \\
&= a_0 + a_1\zeta + a_2\zeta^2 + \ldots + a_{n-1}\zeta^{n-1} + c(1 + \zeta + \zeta^2 + \ldots + \zeta^{n-1}) \\
&= (a_0 + c) + (a_1 + c)\zeta + (a_2 + c)\zeta^2 + \ldots + (a_{n-1} + c)\zeta^{n-1}.
\end{aligned}$$

Er is dus geen éénduidige relatie tussen de cyclotomische getallen en de rijtjes $(a_0, \ldots, a_{n-1}) \in \mathbb{Z}^n$. We kunnen ons afvragen of er nog meer van die 'lineaire afhankelijkheidsrelaties' zijn tussen de 'basisvectoren' $1, \zeta, \zeta^2, \ldots, \zeta^{n-1}$: zijn er behalve $1 + X + X^2 + \ldots + X^{n-1}$ andere veeltermen met gehele coëfficiënten waarvan ζ een nulpunt is? Als d een echte deler is van n,[2] zeg $n = kd$, dan is bijvoorbeeld

$$0 = 1 - \zeta^n = 1 - (\zeta^d)^k = (1 - \zeta^d)(1 + \zeta^d + \zeta^{2d} + \ldots + \zeta^{(k-1)d}),$$

[2] dat wil zeggen, niet 1 of n

en omdat $1-\zeta^d \neq 0$ moet de tweede factor $1+\zeta^d+\zeta^{2d}+\ldots+\zeta^{n-d}$ nul zijn. Als n *priem* is, dan zijn er geen echte delers. Het blijkt dat er dan *geen* andere relaties tussen de ζ^k zijn dan $1+\zeta+\ldots+\zeta^{n-1}=0$. Met andere woorden, als een cyclotomisch getal $a_0+a_1\zeta+a_2\zeta^2+\ldots+a_{n-1}\zeta^{n-1}$ gelijk is aan nul, dan is $a_0=a_1=a_2=\ldots=a_{n-1}$. Dit werd bewezen door Gauss, en in *exercise 15* van §4.2 van [4] wordt een voor eerstejaars begrijpelijk bewijs geschetst. Om het boek niet te lang te maken, zullen we hier een veel korter bewijs geven, waarvoor je echter wel wat Galois-theorie nodig hebt, en meer ringentheorie dan we hier behandelen (bijvoorbeeld [8] en [9] bieden meer dan genoeg vookennis, de eerste helft van [12] voldoet ook).

Lemma 9.1.1. *Als we een lineaire relatie $a_0+a_1\zeta+a_2\zeta^2+\ldots+a_{n-1}\zeta^{n-1}=0$ hebben in $\mathbb{Z}[\zeta]$ met $\zeta=\zeta_n$, dan zijn de coëfficiënten a_i allemaal aan elkaar gelijk.*

Bewijs. Wat we zojuist lieten zien is in essentie dat het polynoom $1+X+X^2+\ldots+X^{n-1}$ reducibel is als n niet priem is: een mogelijke ontbinding is dan $(1-X^d)(1+X^d+X^{2d}+\ldots+X^{(k-1)d})$ met $n=kd$. Als $n=p$ priem is, blijkt hij echter wél irreducibel. Noemen we de veelterm namelijk f_p (dit wordt vaak het *p-de cyclotomische polynoom* genoemd), dan hebben we

$$f_p(X)=1+X+X^2+\ldots+X^{p-1}=\frac{X^p-1}{X-1}$$

waarbij we de deling in de ring $\mathbb{Z}[X]$ uitvoeren; dit volgt door aan beide kanten met $X-1$ te vermenigvuldigen. De belangrijkste observatie is dat de functie $\phi:\mathbb{Z}[X]\to\mathbb{Z}[X]:g(X)\mapsto g(X+c)$ een homomorfisme is voor alle $c\in\mathbb{Z}$; overtuig jezelf van de waarheid daarvan. Bovendien is ϕ 'graad-bewarend': $\deg(\phi(g))=\deg(g)$ voor alle g. Vervolgens merken we op dat $g(X)$ irreducibel is dan en slechts dan als $g(X+c)$ dat is. Immers, als $g(X)$ reducibel is, laten we zeggen dat $h(X)k(X)$ een ontbinding is in niet-eenheden, dan volgt omdat ϕ een homomorfisme is dat $g(X+c)=h(X+c)k(X+c)$, en deze is reducibel want de factoren hebben dezelfde graad als $h(X)$ en $k(X)$. De omgekeerde bewerking '$g(X+c)$ reducibel impliceert $g(X)$ reducibel' volgt door hetzelfde argument toe te passen met $-c$ in plaats van c.

Nu passen we dit toe op de cyclotomische polynoom f_p met $c=1$. We hebben

$$f_p(X+1)=\frac{(X+1)^p-1}{X}=\frac{-1+\sum_{k=0}^{p}\binom{p}{k}X^k}{X}=X^{p-1}+pX^{p-2}+\binom{p}{2}X^{p-3}+\ldots+\binom{p}{2}X+p.$$

De binominaalcoëfficiënten $\binom{p}{k}$ zijn allen deelbaar door p voor $1\leq k\leq p-1$, de kopcoëfficiënt 1 is geen p-voud en de 'staart-coëfficiënt' p is geen p^2-voud. Eisenstein's criterium is dus van toepassing: $f_p(X+1)$ is irreducibel. Uit het voorgaande volgt dat ook $f_p(X)$ irreducibel is. Dit is dus een monische, irreducibele veelterm in $\mathbb{Q}[X]$ met ζ als wortel, het is dus de *minimaalpolynoom* van ζ. Nu volgt het resultaat vrijwel meteen. Stel $a_0+a_1\zeta+a_2\zeta^2+\ldots+a_{n-1}\zeta^{p-1}=0$ met $a_i\in\mathbb{Z}$, dan is $a_{p-1}\neq 0$ omdat de minimaalpolynoom graad $p-1$ heeft. We mogen daarom door a_{p-1} delen, dus ζ is nulpunt van de monische veelterm

$$g(X)=\frac{a_0}{a_{p-1}}+\frac{a_1}{a_{p-1}}X+\frac{a_2}{a_{p-1}}X^2+\ldots+X^{p-1}\in\mathbb{Q}[X]$$

en omdat $g(X)$ de minimale graad $p-1$ heeft, volgt $g(X)=f_p(X)$, dus $a_0/a_{p-1}=a_1/a_{p-1}=\ldots=a_{p-2}/a_{p-1}=1$ zoals gewenst. □

Stel twee cyclotomische getallen $a_0+a_1\zeta+\ldots+a_{p-1}\zeta^{p-1}$ en $b_0+b_1\zeta+\ldots+b_{p-1}\zeta^{p-1}$ zijn aan elkaar gelijk (met p priem). Dan is hun verschil $(a_0-b_0)+(a_1-b_1)\zeta+\ldots+(a_{p-1}-b_{p-1})\zeta^{p-1}$ gelijk aan nul, en dat betekent dat $a_0-b_0=a_1-b_1=\ldots=a_{p-1}-b_{p-1}$. Omgekeerd, als a_0-b_0, a_1-b_1, $\ldots a_{p-1}-b_{p-1}$ allen gelijk zijn aan dezelfde constante c, dan is

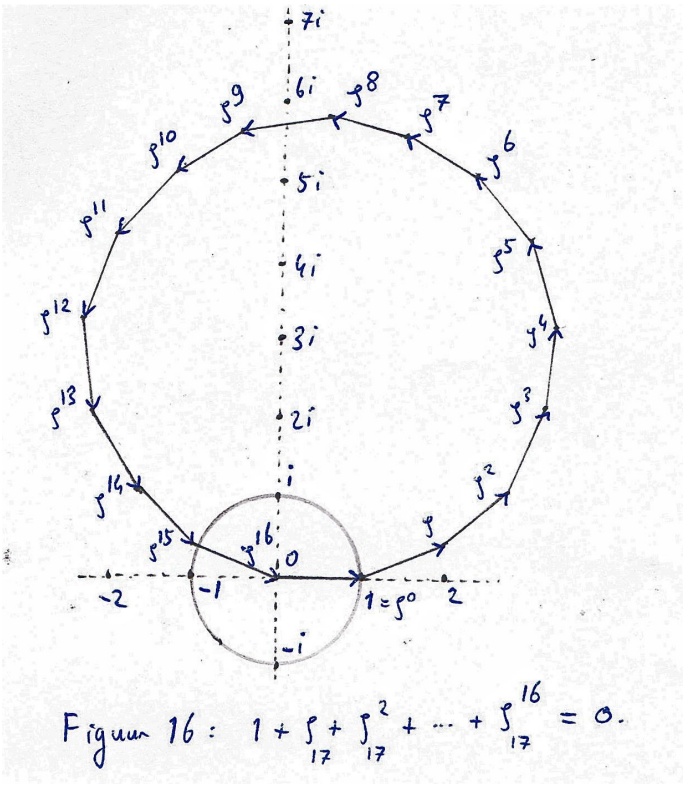
$$a_0+a_1\zeta+\ldots+a_{p-1}\zeta^{p-1}=b_0+b_1\zeta+\ldots+b_{p-1}\zeta^{p-1}+c(1+\zeta+\ldots+\zeta^{p-1})=b_0+b_1\zeta+\ldots+b_{p-1}\zeta^{p-1}.$$

We concluderen:

Gevolg 9.1.2. *Zij p een priemgetal. Twee cyclotomische getallen $a_0+a_1\zeta+\ldots+a_{p-1}\zeta^{p-1}$ en $b_0+b_1\zeta+\ldots+b_{p-1}\zeta^{p-1}$ zijn aan elkaar gelijk precies dan als*

$$a_0-b_0=a_1-b_1=\ldots=a_{p-1}-b_{p-1}.$$

9.1.1 Ontbindingsringen en reguliere priemgetallen



Figuur 16: $1+\zeta_{17}+\zeta_{17}^2+\ldots+\zeta_{17}^{16}=0$.

We hebben al gezien dat $\mathbb{Z}[\zeta_n]$ een hoofdideaaldomein is voor $n=1,2,3,4$, en het zou mooi zijn als dat ook zo zou zijn voor in elk geval alle priemgetallen n. Eigenlijk hoeft het niet eens een hoofdideaaldomein te zijn, het gaat erom dat er sprake is van unieke priemfactorisatie. Een domein waarin elk element een in essentie unieke priemontbinding heeft in de zin van Definitie 7.2.2, *maar dan met 'irreducibel' vervangen door 'priem'*,[3] wordt een ontbindingsring genoemd. Lamé en anderen probeerden een geschikte methode van deling met rest te zoeken die werkt voor alle priemgetallen $n=p$, maar dat lukte niet. In mei 1874, zo'n drie maanden nadat Lamé tijdens zijn voordracht beweerde dat $\mathbb{Z}[\zeta_p]$ een ontbindingsring is voor alle priemgetallen p, maakte Kummer een brief openbaar waarin hij het tegendeel beweerde, met bijgevoegd een artikel dat hij *drie jaar daarvoor* al had gepubliceerd waarin hij dat bewees. Kummer zij echter ook dat unieke factorisatie kon worden 'gered' door een nieuw soort 'complex getal' in te voeren, een zogenoemd 'ideaal complex getal'. Kummer wordt hiermee gezien als (afhankelijk van je visie) de uitvinder of ontdekker van

[3] De reden hiervoor zal straks duidelijk worden. Sommige auteurs houden de zwakkere voorwaarde 'irreducibel' aan, dat doen wij niet.

het begrip ideaal. Hoewel Kummer in principe net zo rekende met 'ideale complexe getallen' in $\mathbb{Z}[\zeta]$ als dat we nu met idealen rekenen, was Kummer's definitie nogal vaag. De moderne definitie die wij gebruiken is in 1876 geopperd door Richard Dedekind.

Kummer bewees dat de Laatste stelling van Fermat waar is voor de zogenaamde *reguliere* priemgetallen. Men weet niet eens zeker of er hier oneindig veel van zijn, maar berekeningen suggereren dat ongeveer 61% van de priemgetallen regulier is.[4] Priemgetallen p waarvoor $\mathbb{Z}[\zeta_p]$ een ontbindingsring is, lijken veel schaarser te zijn, maar in elk geval is het een ontbindingsring voor alle $p < 23$. Het is dus niet zo verwonderlijk dat Lamé en anderen ervan overtuigd waren dat het een ontbindingsring is voor alle p. Omdat het bewijs veel eenvoudiger is als $\mathbb{Z}[\zeta_p]$ een ontbindingsring is dan wanneer p slechts regulier is, zullen we ons hier op ontbindingsringen richten. Fermat's vermoeden bewijzen voor alle priemgetallen onder de 23 is op zich al niet slecht. Als je ook het veel meer omvattende bewijs over reguliere priemgetallen wilt bestuderen, kunnen we [4] aanraden; dit hoofdstuk levert in principe genoeg voorkennis om dat te begrijpen.

9.2 Congjungatie en norm

In het vervolg nemen we aan dat we een zeker priemgetal $p > 2$ hebben gekozen waarvoor we FLT willen bewijzen. We schrijven daarom ζ_p gewoon als ζ. We nemen echter niet aan dat $\mathbb{Z}[\zeta]$ een ontbindingsring is, want het zal blijken dat we dat maar op één plaats nodig hebben.

Om de notatie eenvoudig te houden, schrijven we cyclotomische getallen $a_0 + a_1\zeta + \ldots + a_{p-1}\zeta^{p-1}$ vaak als $f(\zeta)$. We kunnen f als polynoom met gehele coëfficiënten beschouwen, bijvoorbeeld $f = a_0 + a_1X + \ldots + a_{p-1}X^{p-1}$, en $f(\zeta)$ is dan deze veelterm geëvalueerd in het punt ζ. Wegens de regel $\zeta^p = 1$ kunnen we ons beperken tot veeltermen van graad ten hoogste $p-1$. We kunnen Gevolg 9.1.2 nu omschrijven in termen van veeltermen. Als $f = a_0 + a_1X + \ldots + a_{p-1}X^{p-1}$ en $g = b_0 + b_1X + \ldots + b_{p-1}X^{p-1}$ twee veeltermen zijn met gehele coëfficiënten en graad kleiner dan p, dan is $f(\zeta) = g(\zeta)$ precies dan als $a_0 - b_0 = a_1 - b_1 = \ldots = a_{p-1} - b_{p-1}$. En dat is weer equivalent met $f = g + c(1 + X + X^2 + \ldots + X^{p-1})$ voor een geheel getal c (namelijk, $c = a_0 - b_0 = \ldots = a_{p-1} - b_{p-1}$). Als we de p-de cyclotomische polynoom $1 + X + X^2 + \ldots + X^{p-1}$ met Φ noteren (wat we in het vervolg blijven doen), dan krijgen dus:

Lemma 9.2.1. *Zij $f, g \in \mathbb{Z}[X]$ twee veeltermen met gehele coëfficiënten en graad kleiner dan p. Dan is $f(\zeta) = g(\zeta)$ precies dan als er een geheel getal c is zodat $f = g + c\Phi$. Dus in elk geval $f \equiv g \pmod{\Phi}$, met congruentie in de zin van §7.1.*

In het vervolg zullen we met f, g en h, tenzij anders vermeld, steeds veeltermen met gehele coëfficiënten en graad kleiner dan p bedoelen.

We kunnen nu bijvoorbeeld ook van $f(\zeta^k)$ spreken, met $1 \leq k \leq p-1$:

$$f(\zeta^k) = a_0 + a_1\zeta^k + a_2\zeta^{2k} + \ldots + a_{p-1}\zeta^{(p-1)k},$$

waarbij we de machten van ζ modulo p lezen zoals in de berekening (9.4). Met deze notatie moeten we echter wel oppassen. Als $f(\zeta) = g(\zeta)$, dan willen we graag dat ook $f(\zeta^k) = g(\zeta^k)$

[4] Berekeningen als deze kunnen echter verraderlijk zijn: het is al een aantal keer gebeurd dat men een vermoeden opstelde aan de hand van miljarden berekeningen, waarna het toch niet waar bleek te zijn omdat het gedrag van getallen op nog veel grotere schaal anders blijkt te zijn, vaak op schalen waar zelfs computers geen raad mee weten.

voor gehele k, maar het is niet meteen duidelijk of dit wel zo is. Stel $f(\zeta) = g(\zeta)$. Dan is $f = g + c\Phi$ voor een gehele c. We hebben dus

$$f(\zeta^k) = g(\zeta^k) + c(1 + \zeta^k + \zeta^{2k} + \ldots + \zeta^{k(p-1)}). \tag{9.5}$$

Stel k is geen veelvoud van p. Omdat p priem is, doorlopen de getallen $0, k, 2k, 3k, \ldots, (p-1)k$ mod p op volgorde na precies de getallen $0, 1, 2, 3 \ldots, p-1$. Voor elke j is er namelijk precies één l zodat $lk \equiv j \pmod{p}$, namelijk $l = jm$ met m de inverse van k in het lichaam $\mathbb{Z}/p\mathbb{Z}$. De factor $1 + \zeta^k + \zeta^{2k} + \ldots + \zeta^{(p-1)k}$ is dus gelijk aan $1 + \zeta + \zeta^2 + \ldots + \zeta^{p-1} = 0$, dus uit (9.5) volgt dat $f(\zeta^k) = g(\zeta^k)$. Als k veelvoud is van p, ofwel $\zeta^k = 1$, dan hoeft dit niet te gelden. Neem bijvoorbeeld $f = 5 + 2X + 4X^2$ en $g = 7 + 4X + 6X^2$. Dan is $f(\zeta) = g(\zeta)$, maar $f(\zeta^k) = 5 + 2 + 4 = 11$ en $g(\zeta^k) = 7 + 4 + 6 = 17$. We vatten het bovenstaande samen in een lemma. Omdat we de exponenten van ζ modulo p kunnen lezen, richten we ons alleen op de $k \in \{0, 1, 2, \ldots, p-1\}$.

Lemma 9.2.2. *Als $f(\zeta) = g(\zeta)$, dan is $f(\zeta^k) = g(\zeta^k)$ voor alle $1 \leq k \leq p-1$.*

De cyclotomische getallen $f(\zeta), f(\zeta^2), \ldots, f(\zeta^{p-1})$ worden vaak de *geconjugeerden* van $f(\zeta)$ genoemd. Als er verwarring kan ontstaan, zullen we $\overline{a + bi} = a - bi$ de *complex* geconjugeerde van $a + bi$ noemen. Van ζ overgaan op ζ^k kunnen we zien als een verandering in keuze van de complexe eenheidswortel.

Een belangrijke eerste stap in ons onderzoek naar de Laatste stelling van Fermat is het definiëren van een norm op $\mathbb{Z}[\zeta]$. Net als bij de gehelen van Gauss en Eisenstein willen we dat de norm een positief geheel getal is, en dat hij voldoet aan $\|\alpha\beta\| = \|\alpha\| \cdot \|\beta\|$, want dan kunnen we weer delingseigenschappen in $\mathbb{Z}$ gebruiken. Zo'n norm blijkt gelukkig te bestaan, namelijk

$$N(f(\zeta)) = f(\zeta)f(\zeta^2)\cdots f(\zeta^{p-1}).$$

Dit is dus het product van alle $p-1$ geconjugeerden van $f(\zeta)$. Vaak schrijven we gewoon $Nf(\zeta)$ in plaats van $N(f(\zeta))$. De functie N is welgedefinieerd, want uit Lemma 9.2.2 volgt dat $Nf(\zeta) = Ng(\zeta)$ als $f(\zeta) = g(\zeta)$. We gaan na dat inderdaad aan de bovenstaande eisen van een norm is voldaan. Allereerst laten we zien dat $Nf(\zeta)$ een *niet-negatief* reëel getal is. Omdat ζ op de eenheidscirkel ligt, is $\zeta\overline{\zeta} = |\zeta|^2 = 1$, dus $\zeta^{-1} = \overline{\zeta}$ zodat $\zeta^{p-k} = \zeta^{-k} = \overline{\zeta}^k$ voor alle k. Schrijven we $f = a_0 + a_1X + \ldots + a_{p-1}X^{p-1}$, dan zien we dat

$$\begin{aligned} f(\zeta^{p-k}) = f(\overline{\zeta}^k) &= a_0 + a_1\overline{\zeta}^k + a_2\overline{\zeta}^{2k} + \ldots + a_{p-1}\overline{\zeta}^{(p-1)k} = \overline{a_0} + \overline{a_1}\overline{\zeta}^k + \overline{a_2}\overline{\zeta}^{2k} + \ldots + \overline{a_{p-1}}\overline{\zeta}^{(p-1)k} \\ &= \overline{a_0 + a_1\zeta^k + a_2\zeta^{2k} + \ldots a_{p-1}\zeta^{(p-1)k}}, \end{aligned}$$

dus $f(\zeta^{p-k}) = \overline{f(\zeta^k)}$. We kunnen de factoren in de norm daarom in groepjes van twee samen nemen:

$$\begin{aligned} Nf(\zeta) &= f(\zeta)f(\zeta^{p-1})f(\zeta^2)f(\zeta^{p-2})\cdots f(\zeta^{\frac{1}{2}(p-1)})f(\zeta^{\frac{1}{2}(p+1)}) \\ &= f(\zeta)\overline{f(\zeta)}f(\zeta^2)\overline{f(\zeta^2)}\cdots f(\zeta^{\frac{1}{2}(p-1)})\overline{f(\zeta^{\frac{1}{2}(p-1)})} = |f(\zeta)f(\zeta^2)\cdots f(\zeta^{\frac{1}{2}(p-1)})|^2. \end{aligned}$$

De norm van $f(\zeta)$ is dus een niet-negatief reëel getal. Nu bewijzen we dat het bovendien een geheel getal is. De kern van het bewijs is dat $Nf(\zeta)$ invariant is onder een conjugatie $\zeta \mapsto \zeta^k$:

daarmee bedoelen we dat $Nf(\zeta) = Nf(\zeta^k)$ voor alle $1 \leq k \leq p-1$. Dit komt weer doordat de getallen $0, k, 2k, \ldots, (p-1)k \mod p$ op volgorde na precies $0, 1, 2, \ldots, p-1$ zijn. De factoren van $Nf(\zeta^k)$ zijn daarom ook op volgorde na gelijk aan die van $Nf(\zeta)$:

$$Nf(\zeta^k) = f(\zeta)f(\zeta^k)f(\zeta^{2k})\cdots f(\zeta^{(p-1)k}) = f(\zeta)f(\zeta^2)f(\zeta^3)\cdots f(\zeta^{p-1}) = Nf(\zeta). \tag{9.6}$$

We schrijven $Nf(\zeta)$ uit als cyclotomisch getal: $Nf(\zeta) = a_0 + a_1\zeta + \ldots + a_{p-1}\zeta^{p-1}$. Uit bovenstaande formule volgt dat $a_0 + a_1\zeta^k + \ldots + a_{p-1}\zeta^{(p-1)k} = a_0 + a_1\zeta + \ldots + a_{p-1}\zeta^{p-1}$ voor alle $1 \leq k \leq p-1$. De coëfficiënt van ζ^k in het linkerlid is a_1, in het rechterlid is die a_k. Omdat de coëfficiënt van ζ^0 aan beide kanten a_0 is, volgt uit Gevolg 9.1.2 dat $a_1 - a_k = a_0 - a_0 = 0$, dus $a_1 = a_k$. Dit geldt voor alle $1 \leq k \leq p-1$, dus

$$Nf(\zeta) = a_0 + a_1\zeta + a_1\zeta^2 + \ldots + a_1\zeta^{p-1} = a_0 - a_1 + a_1(1 + \zeta + \zeta^2 + \ldots + \zeta^{p-1}) = a_0 - a_1.$$

We concluderen dat $Nf(\zeta)$ een geheel getal is. Bovendien merkten we bij (7.5) al op dat $(fg)(\zeta) = f(\zeta)g(\zeta)$, dus

$$\begin{aligned} N(f(\zeta)g(\zeta)) &= N((fg)(\zeta)) = (fg)(\zeta)\cdots(fg)(\zeta^{p-1}) \\ &= f(\zeta)\cdots f(\zeta^{p-1})g(\zeta)\cdots g(\zeta^{p-1}) = Nf(\zeta)\cdot Ng(\zeta). \end{aligned}$$

We vatten dit alles samen in

Lemma 9.2.3. *De norm van een cyclotomisch getal is een niet-negatief geheel getal, en de normfunctie N respecteert de bewerkingen van $\mathbb{Z}$ en $\mathbb{Z}[\zeta]$. Met andere woorden, als α, β cyclotomische getallen zijn, dan is $N(\alpha) \in \mathbb{Z}_{\geq 0}$ en $N(\alpha\beta) = N(\alpha)N(\beta)$.*[5]

Als m een geheel getal is, dan is $N(m) = m \cdot m \cdot \ldots \cdot m = m^{p-1}$. In het bijzonder is $N(0) = 0$. Men kan ook bewijzen dat $N(\alpha) = 0$ precies dan als $\alpha = 0$, maar omdat we dat niet nodig hebben doen we dat hier niet. Ten slotte merken we op dat de norm die we al op de gehelen van Eisenstein hebben gedefiniëerd, hetzelfde is als deze N. Immers, $Nf(\zeta_3) = f(\zeta_3)f(\zeta_3^{3-1}) = f(\zeta_3)\overline{f(\zeta_3)} = \|f(\zeta_3)\|^2$. De norm op $\mathbb{Z}[\zeta_4]$, de ring van gehelen van Gauss, is echter anders dan N. We weten zelfs niet of N hier voldoet aan de drie normeigenschappen genoemd in bovenstaand lemma, want 4 is niet priem.

De norm is ons belangrijkste gereedschap in het onderzoek naar irreducibele en priemelementen van $\mathbb{Z}[\zeta]$. Een aantal dingen die we bij de gehelen van Gauss en Eisenstein deden, kunnen we direct generaliseren naar willekeurige ζ_p. De eenheden zijn bijvoorbeeld weer precies de elementen van norm 1. Als $Nf(\zeta) = 1$, dan is $f(\zeta)\cdot f(\zeta^2)\cdots f(\zeta^{p-1}) = 1$, dus $f(\zeta)$ is een eenheid met inverse $f(\zeta^2)\cdots f(\zeta^{p-1})$. Stel omgekeerd dat een cyclotomisch getal α een eenheid is, zeg $\alpha\beta = 1$. Dan is $N(\alpha)N(\beta) = N(\alpha\beta) = N(1) = 1^{p-1} = 1$, dus het niet-negatieve gehele getal $N(\alpha)$ deelt 1 en is dus 1. We kunnen ook wat zeggen over deelbaarheid in $\mathbb{Z}[\zeta]$. Als α deelbaar is door β, zeg $\alpha = \beta\gamma$, dan is $N(\alpha) = N(\beta)N(\gamma)$, dus de norm van α is deelbaar door de norm van β. Dit beperkt enorm het aantal kandidaten van delers van α. In het bijzonder, als de norm van α een priemgetal is, dan moet β of γ norm 1 hebben, en dus een eenheid zijn.

[5] Ons bewijs is een bewerking van het analoge bewijs in [4]. (Helaas is het boek terug naar de bieb, we kunnen dus geen pagina- of stellingnummers noemen. De opzet van het boek is echter zo duidelijk dat de lezer eventueel wel snel kan achterhalen om welke stellingen het gaat.)

Omdat α geen eenheid is (want zijn norm is niet 1), betekent dit dat α irreducibel is. Hieruit volgt nog niet dat α ook een priemelement is, want $\mathbb{Z}[\zeta]$ hoeft geen hoofdideaaldomein te zijn. Dat is jammer, want de priemeigenschap $x|yz \implies [\ x|y \text{ of } x|z\ ]$ zou vaak erg goed van pas komen.

Tenslotte noemen we een paar eenvoudige maar handige eigenschappen van deelbaarheid, die eigenlijk voor elk domein gelden. Zij α, β, γ cyclotomische getallen, en u een eenheid. Stel α is deler van $u\beta$, zeg $\alpha\delta = u\beta$. Dan is $\alpha\delta u^{-1} = \beta$, dus α deelt ook β. Stel nu dat α niet nul is en dat $\alpha\beta$ deler is van $\alpha\gamma$, zeg $\alpha\beta\epsilon = \alpha\gamma$. Dan is $\alpha(\beta\epsilon - \gamma) = 0$, en omdat $\alpha \neq 0$ volgt dat $\beta\epsilon = \gamma$, dus β is deler van γ. We vatten de nieuwe observaties weer samen in een lemma.

Lemma 9.2.4. *Zij α, β, γ cyclotomische getallen, en $u \in \mathbb{Z}[\zeta]^*$ een eenheid.*

1. *De eenheden in $\mathbb{Z}[\zeta]$ zijn precies de elementen met norm 1.*
2. *Als α deler is van β, dan is $N(\alpha)$ deler[6] van $N(\beta)$.*
3. *Als $N(\alpha)$ een priemgetal is, dan is α irreducibel (niet noodzakelijk priem).*
4. *Als $\alpha|u\beta$, dan ook $\alpha|\beta$.*
5. *Als α niet nul is en $\alpha\beta|\alpha\gamma$, dan ook $\beta|\gamma$.*

9.3 Modulorekenen en het priemelement $\zeta - 1$

Het blijkt in het algemeen lastig te zijn om direct uit de definitie van 'priemelement' te bepalen of een element priem is. Ter vermaak merken we op dat uit berekening (9.4) volgt dat het getal 5 niet priem is in $\mathbb{Z}[\zeta_5]$. Het product van de factoren daar is duidelijk deelbaar door 5. Met Gevolg 9.1.2 wordt duidelijk dat een cyclotomisch getal $a_0 + a_1\zeta_5 + \ldots + a_4\zeta_5^4$ deelbaar is door vijf precies dan als $a_0 \equiv a_1 \equiv \ldots \equiv a_4 \pmod 5$, dus geen van de factoren is deelbaar door 5. Dat betekent dat 5 hier niet priem is.[7]

Er is gelukkig in elk geval voor *elk* priemgetal p een element in $\mathbb{Z}[\zeta_p]$ waarvan we zeker weten dat het priem is, namelijk $\zeta_p - 1$. Dit zullen we zo dadelijk bewijzen, maar eerst hebben we een paar deelresultaten nodig. Als eerste leiden we een handige ontbinding af van het getal p in $\mathbb{Z}[\zeta]$.

Lemma 9.3.1. *We kunnen p ontbinden als*[8]

$$p = (1-\zeta)(1-\zeta^2)\cdots(1-\zeta^{p-1}).$$

Bewijs. De $p-1$ verschillende getallen ζ^k met $k = 1, 2, \ldots, p-1$ zijn p-de machts eenheidswortels, en dus nulpunt van het polynoom Φ. Uit Stelling 7.4.3 volgt dat deze te ontbinden is als

$$X^{p-1} + X^{p-2} + \ldots + X + 1 = (X-\zeta)(X-\zeta^2)\cdots(X-\zeta^{p-1}),$$

[6] We bedoelen hier deelbaarheid in $\mathbb{Z}$, niet in $\mathbb{Z}[\zeta]$. Dat maakt echter niet uit, want ze komen hier op hetzelfde neer.

[7] Het is zelfs niet irreducibel, zoals blijkt uit het volgende Lemma.

[8] Ons bewijs is een bewerking van het analoge bewijs in [4].

want met hetzelfde argument als na (9.2) volgt dat $q = 1$. Het lemma volgt door aan beide kanten $X = 1$ in te vullen, want $1^{p-1} + 1^{p-2} + \ldots + 1 + 1 = p$. □

Dit lemma zegt dus dat $p = N(1-\zeta)$. Dat betekent dus in elk geval dat $1-\zeta$ irreducibel is. Om te bewijzen dat het ook priem is, hebben we nog wat meer nodig. Stel α is een cyclotomisch getal en deler van twee *gehele* getallen a, b. Als $\text{ggd}(a, b) = 1$, dan kunnen met het Euclidisch algoritme gehele x, y vinden zodat $ax + by = 1$. Omdat α deler is van a en b, deelt het ook $ax + by = 1$. Maar de delers van 1 zijn per definitie precies de eenheden. We hebben nu bewezen:

Lemma 9.3.2. *Als een cyclotomisch getal twee coprieme gehele getallen deelt, dan is het een eenheid. Met andere woorden: copriem zijn in $\mathbb{Z}$ komt op hetzelfde neer als copriem zijn in $\mathbb{Z}[\zeta]$.*

Hieruit volgt in het bijzonder:

Lemma 9.3.3. *De veelvouden van $\zeta - 1$ die in $\mathbb{Z}$ zitten, zijn precies de veelvouden van p. Preciezer geformuleerd: $(\zeta - 1) \cap \mathbb{Z} = p\mathbb{Z}$.*[9]

Bewijs. Stel m is een geheel getal en veelvoud van p, zeg $m = kp$. Dan is m zeker deelbaar door $\zeta - 1$, want[10] $(\zeta - 1)(\zeta^2 - 1) \cdots (\zeta^{p-1} - 1)k = pk = m$. Stel omgekeerd m is geheel en deelbaar door $\zeta - 1$. Als m geen veelvoud is van p, dan zijn m en p relatief priem, en omdat $\zeta - 1$ hen beide deelt, volgt uit Lemma 9.3.2 dat $\zeta - 1$ eenheid is. Tegenspraak, want zijn norm is p en niet 1. We concluderen dat m een p-voud is. □

Eerder in dit boek hebben we modulorekenen gegeneraliseerd naar willekeurige domeinen, en we merkten in Formule (7.2) op dat modulo een hoofdideaal rekenen eigenlijk net zo gaat als in $\mathbb{Z}$. Dat wil zeggen, $\alpha \equiv \beta \pmod{\gamma}$ betekent niets anders dan $\gamma|(\beta - \alpha)$. We zagen al dat $\equiv$ een equivalentierelatie is, en consistent met optelling en vermenigvuldiging. Dit modulorekenen blijkt vaak handig te zijn in $\mathbb{Z}[\zeta]$. We kunnen er bijvoorbeeld makkelijk mee bewijzen dat $\zeta - 1$ priem is. Bovenstaand lemma zegt voor gehele m in deze notatie dat $m \equiv 0 \pmod{\zeta - 1}$ precies dan als p deler is van m. Dit laatste kunnen we schrijven als $m \equiv 0 \pmod{p}$, maar dat kan verwarring veroorzaken: bedoelen we de congruentierelatie gedefinieerd op $\mathbb{Z}$ of op $\mathbb{Z}[\zeta]$? Gelukkig is het makkelijk na te gaan dat deze hier overeenstemmen, de relatie op $\mathbb{Z}[\zeta]$ is een uitbreiding van die op $\mathbb{Z}$. Als a, b geheel zijn en $a\gamma = b$ voor een cyclotomisch getal γ, dan volgt namelijk uit Gevolg 9.1.2 dat γ geheel moet zijn. We kunnen Lemma 9.3.3 daarom als volgt (iets gegeneraliseerd) opschrijven.

Lemma 9.3.4. *Voor gehele getallen is congruentie modulo $\zeta - 1$ equivalent met congruentie modulo p. In formule: voor alle $a, b \in \mathbb{Z}$ geldt*

$$a \equiv b \pmod{\zeta - 1} \qquad \Longleftrightarrow \qquad a \equiv b \pmod{p}.$$

[9] Ons bewijs is een bewerking van het analoge bewijs in [4].

[10] We hebben in Lemma 9.3.1 eigenlijk bewezen dat $(1-\zeta)(1-\zeta^2)\cdots(1-\zeta^{p-1}) = p$, maar omdat het aantal factoren even is (namelijk $p-1$ met p een priemgetal groter dan 2), geeft het niet dat we de tekens van de factoren hebben omgeklapt. Dit omklappen zullen we voortaan steeds doen zonder het te vermelden.

Bewijs. Dit volgt direct uit

$$a \equiv b \pmod{\zeta - 1} \iff (\zeta - 1) | (a - b) \iff p | (a - b) \iff a \equiv b \pmod{p}.$$

□

We bewijzen nu dat $\zeta - 1$ priem is, en bijgevolg ook de $\zeta^k - 1$.

Lemma 9.3.5. *De elementen $\zeta - 1$, $\zeta^2 - 1$, ..., $\zeta^{p-1} - 1$ zijn allen priemelement van $\mathbb{Z}[\zeta_p]$. Bovendien zijn ze op eenheden na gelijk aan elkaar: voor alle $0 \leq k, l \leq p-1$ is er een eenheid u zodat $\zeta^l - 1 = u(\zeta^k - 1)$.*[11]

Bewijs. Stel $f(\zeta), g(\zeta)$ zijn twee cyclotomische getallen zodat $\zeta - 1$ deler is van hun product. Met andere woorden,

$$f(\zeta)g(\zeta) \equiv 0 \pmod{\zeta - 1}.$$

Een wat flauwe maar toch heel belangrijke opmerking is dat $\zeta \equiv 1 \pmod{\zeta - 1}$. Hiermee volgt namelijk uit het bovenstaande dat $f(1)g(1) \equiv 0 \pmod{\zeta - 1}$. Om dit in te zien, schrijven we $f(\zeta)g(\zeta)$ uit als $a_0 + a_1\zeta + a_2\zeta^2 + \ldots + a_{p-1}\zeta^{p-1}$. Omdat de congruentierelatie consistent is met vermenigvuldiging, volgt uit $\zeta \equiv 1 \pmod{\zeta - 1}$ dat $a_k\zeta^k \equiv a_k 1^k = a_k \pmod{\zeta - 1}$ voor alle termen $a_k\zeta^k$. Door herhaald toepassen van de consistentie met optelling, volgt dat

$$\begin{aligned} &a_0 + a_1\zeta \equiv a_0 + a_1, \quad a_0 + a_1\zeta + a_2\zeta^2 \equiv a_0 + a_1 + a_2, \quad \ldots, \\ &a_0 + a_1\zeta + a_2\zeta^2 + \ldots + a_{p-1}\zeta^{p-1} \equiv a_0 + a_1 + \ldots + a_{p-1} \pmod{\zeta - 1}. \end{aligned}$$

Dus $0 \equiv f(\zeta)g(\zeta) \equiv f(1)g(1) \pmod{\zeta - 1}$ zoals gewenst. Omdat $f(1)g(1)$ een geheel getal is, volgt uit Lemma 9.3.3 dat p priemdeler is van het product $f(1)g(1)$ met $f(1), g(1)$ geheel. Dat betekent dat $p | f(1)$ of $p | g(1)$. Weer uit Lemma 9.3.3 volgt dat $f(1) \equiv 0 \pmod{\zeta - 1}$ of $g(1) \equiv 0 \pmod{\zeta - 1}$, en omdat $1 \equiv \zeta \pmod{\zeta - 1}$ volgt zoals net dat $f(\zeta) \equiv 0 \pmod{\zeta - 1}$ of $g(\zeta) \equiv 0 \pmod{\zeta - 1}$. Dus $\zeta - 1$ deelt $f(\zeta)$ of $g(\zeta)$. We concluderen dat $\zeta - 1$ priem is.

Nu volgt makkelijk de rest van de stelling. De belangrijkste observatie is dat

$$\zeta^k - 1 = (\zeta - 1)(\zeta^{k-1} + \zeta^{k-2} + \ldots + \zeta + 1)$$

voor alle $1 \leq k \leq p-1$. Van (9.6) weten we dat $N(\zeta^k - 1) = N(\zeta - 1)$, en met de multiplicatieve eigenschap $N(\alpha)N(\beta) = N(\alpha\beta)$ volgt uit het bovenstaande dat $N(\zeta^{k-1} + \ldots + \zeta + 1) = 1$. Dus $\zeta^{k-1} + \ldots + \zeta + 1$ is een eenheid, laten we hem u_k noemen. We zien dat $\zeta^k - 1 = u_k(\zeta - 1)$, en dus $\zeta^k - 1 = u(\zeta^l - 1)$ voor alle k, l, waarbij $u = u_k u_l^{-1}$ een eenheid is. Als een cyclotomisch getal α priem is, en u een eenheid, dan is ook $u\alpha$ priem. Stel namelijk dat $u\alpha$ deler is van $\beta\gamma$, dan zeker ook $\alpha | \beta\gamma$, dus α deelt β of γ. Er is dus een ϵ zodat $\alpha\epsilon = \beta$ of $\alpha\epsilon = \gamma$, en dus $(u\alpha)\epsilon u^{-1} = \beta$ of $(u\alpha)\epsilon u^{-1} = \gamma$. Dus $u\alpha$ deelt β of γ, en we concluderen dat hij priem is. Omdat $\zeta - 1$ priem is en op een eenheid na $\zeta^k - 1$ is, volgt dat ook deze priem is. □

Nu is er in elk geval één element waar we de volledige priemontbinding van weten, namelijk p. We weten immers dat $p = (\zeta - 1)(\zeta^2 - 1) \cdots (\zeta^{p-1} - 1)$, en voor elke factor $\zeta^k - 1$ is er een eenheid u_k zodat $\zeta^k - 1 = u_k(\zeta - 1)$. Dus $p = u_1 u_2 \cdots u_{p-1}(\zeta - 1)^{p-1}$, met $\zeta - 1$ priem en $u_1 u_2 \cdots u_{p-1}$ een eenheid.

[11] Ons bewijs is een bewerking van het analoge bewijs in [4].

Gevolg 9.3.6. *De priemontbinding van van p in $\mathbb{Z}[\zeta]$ is $p = u(\zeta-1)^{p-1}$ voor een eenheid u.*

Voor elke eenheid u en alle cyclotomische α, β, γ is $\alpha \equiv \beta \pmod{\gamma}$ equivalent met $\alpha \equiv \beta \pmod{u\gamma}$. Immers, het eerste betekent dat γ deler is van $\beta - \alpha$, het tweede dat $u\gamma$ deler is van $\beta - \alpha$. Het is duidelijk dat het tweede het eerste impliceert, en andersom volgt het wegens symmetrie in γ en $u\gamma$ omdat $\gamma = u^{-1}(u\gamma)$. In het bijzonder volgt nu uit Gevolg 9.3.6:

Lemma 9.3.7. *Congruentie modulo $(\zeta-1)^{p-1}$ is equivalent met congruentie modulo p. In formule: voor alle $\alpha, \beta \in \mathbb{Z}[\zeta_p]$ geldt*

$$\alpha \equiv \beta \pmod{(\zeta-1)^{p-1}} \qquad \Longleftrightarrow \qquad \alpha \equiv \beta \pmod{p}.$$

9.4 Gemene delers en eenheden

Om de Laatste stelling van Fermat te bewijzen voor exponenten p waarvoor $\mathbb{Z}[\zeta_p]$ een ontbindingsring is, hebben we nog wat deelresultaten nodig. Het volgende lemma is een generalisatie van wat we van $\mathbb{Z}$ weten, en het bewijs gaat net zo als daar. Voor het eerst hebben we nodig dat $\mathbb{Z}[\zeta_p]$ een ontbindingsring is; dit geldt dus niet meer voor alle oneven priemen p.

Lemma 9.4.1. *Stel $\mathbb{Z}[\zeta_p]$ is een ontbindingsring, en stel dat α relatief priem is met de (eindig veel) elementen $\beta_1, \beta_2, \ldots, \beta_n \in R$. Dan is α relatief priem met hun product $\beta_1\beta_2\cdots\beta_n$.*

Bewijs. Stel α is niet relatief priem met het product, en zij γ een gemene deler die geen eenheid is. Omdat we bij aanname in een ontbindingsring werken, is er een priemelement δ dat γ deelt. Dit is dus een gemene priemdeler van α en van het product $\beta_1\beta_2\cdots\beta_n$. Omdat δ priem is, deelt het minstens één van de β_i, dus δ is een gemene deler van α en β_i, en omdat die relatief priem zijn is δ een eenheid. Tegenspraak, want δ is priem. □

Een prettige eigenschap van modulorekenen is dat we sommige uitdrukkingen modulo een getal enorm kunnen vereenvoudigen. Een voor ons belangrijk voorbeeld is de uitdrukking $(\alpha+\beta)^p$ modulo p voor cyclotomische getallen α, β. We kunnen dit met het binomium van Newton uitschrijven:

$$(\alpha+\beta)^p = \alpha^p + \binom{p}{1}\alpha^{p-1}\beta + \ldots + \binom{p}{p-1}\alpha\beta^{p-1} + \beta^p.$$

De binomiaalcoëfficiënten $\binom{p}{k}$ zijn uitgeschreven $\frac{p!}{k!(p-k)!}$. De teller is deelbaar door p, en omdat p priem is, is de noemer niet deelbaar door p als k niet 0 of p is. De coëfficiënten $\binom{p}{1}, \binom{p}{2}, \ldots, \binom{p}{p-1}$ zijn dus congruent nul modulo p, en we concluderen dat

$$(\alpha+\beta)^p \equiv \alpha^p + \beta^p \pmod{p}.$$

Met inductie kunnen we nu makkelijk inzien dat dit voor alle eindige sommen geldt: $(\alpha_1+\alpha_2+\ldots+a_m)^p \equiv \alpha_1^p + \alpha_2^p + \ldots + \alpha_m^p \pmod{p}$ voor cyclotomische $\alpha_1, \ldots, \alpha_m$. Voor $m=2$ hebben we het net bewezen. Stel het geldt voor $m=k$, dan volgt het vrijwel direct voor $m=k+1$:

$$\begin{aligned}(\alpha_1+\ldots+\alpha_{k+1})^p &\equiv ((\alpha_1+\ldots+\alpha_k)+\alpha_{k+1})^p \\ &\equiv (\alpha_1+\ldots+\alpha_k)^p + \alpha_{k+1}^p \equiv \alpha_1^p+\ldots+\alpha_k^p+\alpha_{k+1}^p \pmod{p}.\end{aligned}$$

Deze eigenschap zelf zullen we niet gebruiken, maar wel een bijzonder geval hiervan. Stel dat γ een cyclotomisch getal is, we schrijven $\gamma = a_0 + a_1\zeta + a_2\zeta^2 + \ldots + a_{p-1}\zeta^{p-1}$. Nemen we $m = p$ en $\alpha_j = a_j\zeta^j$, dan zien we modulo p:

$$\gamma^p \equiv (a_0+a_1\zeta+\ldots+a_{p-1}\zeta^{p-1})^p \equiv a_0^p+(a_1\zeta)^p+\ldots+(a_{p-1}\zeta^{p-1})^p \equiv a_0^p+a_1^p\zeta^p+\ldots+a_{p-1}^p(\zeta^p)^{p-1}.$$

Maar $\zeta^p = 1$, dus dit laatste is gelijk aan het *gehele* getal $a_0^p + a_1^p + \ldots + a_{p-1}^p$. We concluderen:

Lemma 9.4.2. *Zij $\gamma \in \mathbb{Z}[\zeta_p]$. Dan is er een geheel getal c zodat $\gamma^p \equiv c \pmod p$.*[12]

We kunnen nog meer handige 'trucjes' verzinnen met modulorekenen. Het blijkt bijvoorbeeld dat we elk cyclotomisch getal kunnen schrijven als 'polynoom' met gehele 'coëfficiënten' en als 'variabele' het priemelement $\zeta - 1$, tenminste als we modulo een macht van $\zeta - 1$ rekenen. Dit zullen we wat duidelijker maken. Stel $f(\zeta)$ is een cyclotomisch getal. Omdat $\zeta \equiv 1 \pmod{\zeta - 1}$, volgt net als in het bewijs van Lemma 9.3.5 dat $f(\zeta) \equiv f(1) \pmod{\zeta - 1}$, dus elk cyclotomisch getal is modulo $\zeta - 1$ congruent met een *geheel* getal. Met inductie kunnen we nu laten zien er voor elk cyclotomisch getal α en elke $k \in \mathbb{N}$ gehele $a_0, a_1, \ldots, a_{k-1}$ zijn zodat $\alpha \equiv a_0 + a_1(\zeta - 1) + \ldots + a_{k-1}(\zeta - 1)^{k-1} \pmod{(\zeta - 1)^k}$. Voor $k = 1$ staat er $\alpha \equiv a_0 \pmod{\zeta - 1}$ voor een gehele a_0, en dat is precies wat we net opmerkten. Stel het geldt voor $k = j$, we willen het voor $k = j + 1$ bewijzen. Uit de inductiehypothese volgt dat er gehele $a_0, a_1, \ldots, a_j$ zijn, en een cyclotomisch getal β, zodat

$$\alpha = a_0 + a_1(\zeta - 1) + \ldots + a_{j-1}(\zeta - 1)^{j-1} + \beta(\zeta - 1)^j.$$

Bovendien is β modulo $\zeta - 1$ congruent met een geheel getal, laten we hem a_j noemen. Er is dus een cyclotomisch getal γ zodat $\beta = a_j + \gamma(\zeta - 1)$, dus $\beta(\zeta - 1)^j = a_j(\zeta - 1)^j + \gamma(\zeta - 1)^{j+1}$. Hieruit volgt dat

$$\alpha \equiv a_0 + a_1(\zeta - 1) + \ldots + a_j(\zeta - 1)^j \pmod{(\zeta - 1)^{j+1}},$$

het is dus ook waar voor $k = j + 1$. Hiermee hebben we de eerste uitspraak bewezen van

Lemma 9.4.3. *Voor elk cyclotomisch getal α en elk natuurlijk getal k zijn er gehele getallen $a_0, a_1, \ldots, a_{k-1}$ zodat*

$$\alpha \equiv a_0 + a_1(\zeta - 1) + \ldots + a_{k-1}(\zeta - 1)^{k-1} \pmod{(\zeta - 1)^k}.$$

Als $k \leq p - 1$, dan zijn de gehele getallen bovendien modulo p eenduidig bepaald: kort opgeschreven, als

$$b_0 + b_1(\zeta - 1) + b_2(\zeta - 1)^2 + \ldots + b_{k-1}(\zeta - 1)^{k-1} \equiv 0 \pmod{(\zeta - 1)^k} \tag{9.7}$$

voor gehele b_i, dan is $b_0 \equiv b_1 \equiv \ldots \equiv b_{k-1} \equiv 0 \pmod p$.[13]

[12] Ons bewijs is een bewerking van het analoge bewijs in [4].
[13] Ons bewijs is een bewerking van het analoge bewijs in [4].

Bewijs. We hoeven alleen nog de tweede bewering te bewijzen, en die volgt makkelijk uit de al bewezen lemma's. Stel $k \leq p-1$ en er zijn $b_0, \ldots, b_{k-1}$ zoals in (9.7). We bewijzen met inductie naar j dat $b_j \equiv 0 \pmod{p}$ voor alle $0 \leq j \leq k-1$. Het linkerlid van (9.7) is deelbaar door $(\zeta-1)^k$, en dus zeker door $\zeta - 1$. Behalve b_0 zijn de termen zelf ook deelbaar door $\zeta - 1$, dus b_0 is deelbaar door $\zeta - 1$. Omdat b_0 geheel is, volgt uit Lemma 9.3.3 dat b_0 deelbaar is door p. Stel we hebben bewezen dat alle b_l met $l \leq j$ deelbaar zijn door p, voor een zekere $j < k-1$. Dan is ook $b_l(\zeta-1)^l \equiv 0 \pmod{p}$ voor al deze l, en uit Lemma 9.3.7 volgt dat $b_l(\zeta-1)^l \equiv 0 \pmod{(\zeta-1)^{p-1}}$. Omdat $k \leq p-1$ is ook $b_l(\zeta-1)^l \equiv 0 \pmod{(\zeta-1)^k}$. We kunnen (9.7) dus schrijven als

$$b_{j+1}(\zeta-1)^{j+1} + b_{j+2}(\zeta-1)^{j+2} + \ldots + b_{k-1}(\zeta-1)^{k-1} \equiv 0 \pmod{(\zeta-1)^k},$$

dus $(\zeta-1)^k$ is deler van $(\zeta-1)^{j+1}\Big(b_{j+1}+b_{j+2}(\zeta-1)+\ldots+b_{k-1}(\zeta-1)^{k-j-2}\Big)$. Van Lemma 9.2.4 weten we dat dit betekent dat $(\zeta-1)^{k-j-1}$ deler is van $b_{j+1}+b_{j+2}(\zeta-1)+\ldots+b_{k-1}(\zeta-1)^{k-j-2}$, dus zeker ook $\zeta-1$ is hier deler van. Omdat $\zeta-1$ alle termen behalve b_{j+1} deelt, volgt dat b_{j+1} deelbaar is door $\zeta-1$, en dus ook door p want b_{j+1} is geheel. Met inductie naar j zien we dat alle coëfficiënten deelbaar zijn door p. □

9.4.1 Eenheden wegwerken

Een van de belangrijkste obstakels die we tegenkomen in ons onderzoek naar de Laatste stelling van Fermat, is de zo onschuldig lijkende vermenigvuldiging met eenheden. Toen we FLT bewezen voor $n = 3$, kregen we op een gegeven moment een derde macht maal een eenheid, en we wilden dat het een derde macht was. In $\mathbb{Z}$ hebben we geen last van dit soort problemen, want de enige eenheden zijn ± 1 en dat zijn zelf derde machten. In $\mathbb{Z}[\zeta]$ ligt dit veel subtieler, en we hebben twee relatief grote lemma's nodig om 'de eenheden te temmen'. Eén ervan gaat over eenheden die machten zijn van ζ_p. De eenhedengroep $\mathbb{Z}[\zeta_p]^*$ is zoals we weten altijd cyclisch, dus als deze eindig is vormen de elementen de hoekpunten van een regelmatige veelhoek op de eenheidscirkel, net als de machten van ζ_p. Dit betekent nog niet dat de eenhedengroep ook gelijk is aan de verzameling van die machten. Dit zagen we al bij de gehelen van Eisenstein, waar de eenheden $1, \zeta_6, \zeta_6^2, \zeta_6^3, \zeta_6^4, \zeta_6^5$ waren, en niet $1, \zeta_3, \zeta_3^2$. Het blijkt echter wél zo te zijn dat $u/\overline{u}$ macht is van ζ_p voor alle eenheden u. Bijvoorbeeld, $\zeta_6^5/\zeta_6^1 = \zeta_6^4 = \zeta_3^2$.

Lemma 9.4.4. *Als u een eenheid is in $\mathbb{Z}[\zeta]$, dan is $u/\overline{u}$ een macht van ζ.*[14]

Bewijs. We schrijven $u/\overline{u} = f(\zeta) = a_0 + a_1\zeta + \ldots + a_{p-1}\zeta^{p-1}$, met andere woorden, we kiezen een veelterm $f(X) = a_0 + a_1X + \ldots + a_{p-1}X^{p-1}$ met gehele coëfficiënten die $u/\overline{u}$ oplevert als hij geëvalueerd wordt in ζ. We kiezen de a_i bovendien zo dat $-\frac{1}{2}p \leq a_0 + a_1 + \ldots + a_{p-1} < \frac{1}{2}p$. Dat kan omdat een cyclotomisch getal onveranderd blijft als we bij alle coëfficiënten van ζ dezelfde constante optellen, dus als $a_0 + a_1 + \ldots + a_{p-1} = qp + r$ met $-\frac{1}{2}p \leq r < \frac{1}{2}p$, dan kunnen we door van elke coëfficiënt q af te trekken ervoor zorgen dat hun som tussen $-\frac{1}{2}p$ en $\frac{1}{2}p$ komt te liggen.

[14] Ons bewijs is een bewerking van de analoge bewijzen in [4] en [10].

Het is duidelijk wat we met het polynoom $f(X^{p-1})$ bedoelen. We voeren restdeling uit van $f(X^{p-1})f(X)$ door $X^p - 1$:

$$f(X^{p-1})f(X) = q(X)(X^p - 1) + r(X), \qquad \deg(r(X)) < p. \tag{9.8}$$

We schrijven $r(X) = b_0 + b_1X + \ldots + b_{p-1}X^{p-1}$. Als we bovenstaande veeltermen evalueren in $X = 1$, dan volgt omdat $1^{p-1} = 1$ en $1^p - 1 = 0$ dat $f(1)^2 = r(1)$, ofwel

$$(a_0 + a_1 + \ldots + a_{p-1})^2 = b_0 + b_1 + \ldots + b_{p-1}. \tag{9.9}$$

We kunnen ook in ζ evalueren, dan volgt omdat $\zeta^p - 1 = 0$ dat $r(\zeta) = f(\zeta^{p-1})f(\zeta)$, en met eerder afgeleid rekenregels van complexe conjugatie volgt

$$b_0 + b_1\zeta + \ldots + b_{p-1}\zeta^{p-1} = f(\zeta^{-1})f(\zeta) = f(\overline{\zeta})f(\zeta) = \overline{f(\zeta)}f(\zeta) = \overline{(e/\overline{e})}(e/\overline{e}) = \frac{\overline{e}}{e} \cdot \frac{e}{\overline{e}} = 1.$$

Dus $(b_0 - 1) + b_1\zeta + \ldots + b_{p-1}\zeta^{p-1} = 0$, waaruit volgt dat $b_0 - 1 = b_1 = b_2 = \ldots = b_{p-1}$. Laten we deze constante k noemen. Uit (9.9) volgt nu dat $(a_0 + a_1 + \ldots + a_{p-1})^2 = 1 + pk \equiv 1 \pmod{p}$ en dus $a_0 + a_1 + \ldots + a_{p-1} \equiv \pm 1 \pmod{p}$. Omdat we verondersteld hebben dat $\frac{1}{2}p \leq a_0 + a_1 + \ldots + a_{p-1} < \frac{1}{2}p$ en $p \neq 2$, volgt dat $a_0 + a_1 + \ldots + a_{p-1} = \pm 1$. Dus $1 + kp = 1$, zodat $k = 0$. Dit betekent dat

$$b_0 = 1, \quad b_1 = b_2 = \ldots = b_{p-1} = 0. \tag{9.10}$$

We gaan nu de veelterm $f(X^{p-1})f(X)$ op een heel andere manier uitschrijven: het is de som van alle termen van de vorm $a_ia_jX^{pi-i}X^j$ met $0 \leq i, j \leq p-1$. Zo'n term kunnen we schrijven als

$$a_ia_jX^{pi+j-i} = a_ia_jX^{qp+r} = a_ia_jX^r(X^{qp} - 1) + a_ia_jX^r = Q_{i,j}(X)(X^p - 1) + a_ia_jX^r,$$

waarbij $r = j - i \bmod p$, en q een bijbehorend geheel getal, en $Q_{i,j}(X)$ een veelterm. Namelijk,

$$Q_{i,j}(X) = \begin{cases} a_ia_jX^r(X^{q(p-1)} + \ldots + X^p + 1) & \text{als } q > 1, \\ a_ia_jX^r & \text{als } q = 1, \\ 0 & \text{als } q = 0. \end{cases}$$

De som van de veeltermen $Q_{i,j}$ over alle i, j noemen we Q. We hebben dus

$$f(X^{p-1})f(X) = \sum_{i,j} a_ia_jX^{pi+j-i} = Q(X)(X^p - 1) + \sum_{i,j} a_ia_jX^r$$

waarbij $r = r_{i,j} = j - i \bmod p$ nog van i en j afhangt, maar wel geldt altijd $0 \leq r < p$. Het is makkelijk na te gaan dat rest en quotiënt van een veelterm bij deling door een andere veelterm uniek is: als $A = BC + D = EC + F$ met D, F beide met graad kleiner dan die van C, dan is $D = F$ en $B = E$. In ons geval heeft de veelterm $\sum_{i,j} a_ia_jX^r$ hoogstens graad $\max_{i,j}\{r_{i,j}\} < p$, dus uit (9.8) volgt dat $r(X) = \sum_{i,j} a_ia_jX^r$. De termen in dit laatste polynoom die bijdragen aan de coëfficiënt van X^0 zijn precies die met $0 = r = j - i \bmod p$, ofwel, die met $i = j$ (want

$0 \leq i, j \leq p-1$). Met andere woorden, de coëfficiënt van X^0 is $a_0^2 + a_1^2 + \ldots + a_{p-1}^2$. Dit moet gelijk zijn aan de coëfficiënt b_0 van X^0 in $r(X)$, dus

$$b_0 = a_0^2 + \ldots + a_{p-1}^2.$$

Omdat de a_i gehele getallen zijn, en omdat $b_0 = 1$, volgt dat precies één van hen ± 1 is, en de anderen 0. Laten we zeggen $a_m = 1$. We hebben dus

$$u/\overline{u} = f(\zeta) = \pm\zeta^m.$$

We moeten nu alleen nog bewijzen dat het teken plus is en niet min. Stel het is min, dus $u/\overline{u} = -\zeta^m$. Dan is ook $u/\overline{u} = -\zeta^{p+m}$, en omdat m of $p+m$ even is, is er een geheel getal s zodat $u/\overline{u} = -\zeta^{2s}$. Hieruit volgt dat $u\zeta^{-s} = -\overline{u}\zeta^s$. We schrijven deze eenheid $u\zeta^{-s}$ als $F(\zeta) = c_0 + c_1\zeta + \ldots + c_{p-1}\zeta^{p-1}$ met $c_0 = 0$ (dat kan altijd omdat we dezelfde constante van alle coëfficiënten mogen aftrekken). Dit getal voldoet aan

$$F(\zeta^{-1}) = F(\overline{\zeta}) = \overline{F(\zeta)} = \overline{u\zeta^{-s}} = \overline{u} \cdot \overline{\zeta^{-s}} = \overline{u}\zeta^s = -F(\zeta),$$

dus $F(\zeta) = -F(\zeta^{-1})$, zodat $2F(\zeta) = F(\zeta) - F(\zeta^{-1})$. Met andere woorden,

$$2F(\zeta) = c_1(\zeta - \zeta^{-1}) + c_2(\zeta^2 - \zeta^{-2}) + \ldots + c_{p-1}(\zeta^{p-1} - \zeta^{-(p-1)}).$$

Elke factor $\zeta^k - \zeta^{-k}$ kunnen we schrijven als $\zeta^{-k}(\zeta^{2k} - 1)$, en ζ^{-k} is een eenheid. Op zijn beurt weten we dat $\zeta^{2k} - 1 = u(\zeta - 1)$ voor een eenheid u, dus $\zeta^k - \zeta^{-k} = w(\zeta - 1)$ voor een eenheid w. Dat betekent dat bovenstaande som deelbaar is door de priem $\zeta - 1$. De norm moet dus deelbaar zijn door de norm van $\zeta - 1$, en dat is p. Maar $F(\zeta)$ is een eenheid, dus de norm van het getal is $N(2F(\zeta)) = N(2)NF(\zeta) = 2^{p-1}$, en dat is niet deelbaar door p want $p \neq 2$. We stuiten dus op een tegenspraak, en concluderen dat het teken plus moet zijn. Hiermee is de stelling bewezen. □

In zijn onderzoek naar al dan geen eenduidige ontbinding in priemfactoren (of eigenlijk de zogenaamde priemidealen) in $\mathbb{Z}[\zeta_p]$ stelde Kummer op een gegeven moment twee 'voorwaarden' op voor p. De eerste is dat het priemgetal p *regulier* is. We gaan hier niet verder op in wat dat betekent. De tweede voorwaarde is: Als u een eenheid is die congruent een geheel getal is modulo p, dan is u een p-de macht. Kummer bewees dat als p aan deze twee voorwaarden voldoet, de Laatste stelling van Fermat waar is voor exponent p. Hij bewees later ook dat de tweede voorwaarde eigenlijk overbodig is: hij volgt uit de eerste voorwaarde 'p is regulier'. Men kan bewijzen dat als $\mathbb{Z}[\zeta_p]$ een ontbindingsring is, automatisch p regulier is. Dat valt buiten het bestek van dit boek,[15] en daarom spreken we van een voorwaarde in plaats van stelling.

Voorwaarde 9.4.5. *Als u een eenheid is, en $u \equiv c \pmod{p}$ voor een geheel getal c, dan is u een p-de macht: er is een eenheid w zodat $u = w^p$.*

Tenslotte nog een laatste lemma, en dan kunnen we beginnen met het 'echte' bewijs. Hoewel, dit voorbereidende werk is eigenlijk ook deel van het bewijs. Het bewijs van het lemma gaat eigenlijk precies hetzelfde als het analoge lemma dat we eerder voor $\mathbb{Z}$ bewezen.

[15] Zie bijvoorbeeld de hoofdstukken 5 en 6 van [4] voor een bewijs.

Lemma 9.4.6. *Zij R een ontbindingsring, en stel dat $\alpha_1, \alpha_2, \ldots, \alpha_m \in R$ paarsgewijs relatief priem zijn. Stel dat $\alpha_1\alpha_2\cdots\alpha_m$ op een eenheid na een n-de macht is. Dan zijn alle factoren op eenheden na n-de machten.*

We formuleren dit iets preciezer. Stel dat er $\gamma, e \in R$ zijn, met e een eenheid, zodat $e\gamma^n = \alpha_1\cdots\alpha_m$. Dan zijn er $\delta_1, \ldots, \delta_n \in R$ en eenheden $u_1, \ldots, u_n \in R^*$ zodat $\alpha_k = u_k\delta_k^n$ voor alle $k = 1, 2, \ldots, m$.

Bewijs. Eerst bewijzen we het voor $m = 2$, met inductie kunnen we het daarna makkelijk uitbreiden naar alle m. Het bewijs voor $m = 2$ gaat geheel analoog aan dat van Stelling 2.2.1 (waarbij $n = p$). We vervangen daar overal $\pm$ door e, en zeggen zoals in de voetnoot daar: 'In het hele bewijs bedoelen we met e een eenheid waarvan we niet weten welke het is, maar het doet er ons ook niet toe welke het is.' We gebruiken hetzelfde symbool e dus voor verschillende eenheden. Na deze conventie kunnen we het bewijs van die stelling precies herhalen, *behalve dat we e niet binnen de haakjes van een n-de macht mogen halen.* Bij ± 1 konden we dat doen omdat die zelf n-de machten zijn, maar in het algemeen zijn eenheden dat niet. Op deze manier volgt dat de p-de machten *op vermenigvuldiging met eenheden na* precies de cyclotomische getallen zijn waarvan p de orde van elke priemfactor deelt. Verder kunnen we het bewijs van daar zonder problemen herhalen, en er volgt dat $e\gamma^p = \alpha_1\alpha_2$ impliceert dat $\alpha_k = u_k\delta_k^p$ voor een eenheid u_k en cyclotomisch getal δ_k (met $k = 1, 2$).

Stel het geldt voor $m = j \geq 2$, we willen het bewijzen voor $m = j + 1$. Omdat de α_i paarsgewijs copriem zijn, is α_{j+1} relatief priem met het product $\alpha_1\cdots\alpha_j$. Door het zojuist gegeven bewijs bij $m = 2$ los te laten op $\alpha_1\cdots\alpha_j$ en α_{j+1}, zien we dat het beide eenheden maal p-de machten zijn. Uit de inductiehypothese volgt hiermee dat de $\alpha_1, \alpha_2, \ldots, \alpha_j$ ook allemaal eenheden maal p-de machten zijn. De stelling geldt dus ook voor $m = j + 1$, en met inductie naar m voor alle $m \geq 2$. □

9.5 Fermat's Laatste Stelling voor ontbindingsringen

Nu zijn we (eindelijk) zo ver dat we het hoofdresultaat van dit hoofdstuk kunnen bewijzen. Hoewel we dit niet bewezen hebben, voldoen in elk geval alle priemgetallen onder de 23 aan de voorwaarden van de stelling, zie pagina 97 van [4].

Stelling 9.5.1. De Laatste Stelling van Fermat voor ontbindingsringen.[16] *Zij p een priemgetal waarvoor $\mathbb{Z}[\zeta_p]$ een ontbindingsring is, en waarvoor bovendien aan Voorwaarde 9.4.5 is voldaan. Er bestaan geen gehele getallen $x, y, z \in \mathbb{Z} - \{0\}$ die voldoen aan $x^p + y^p = z^p$.*

Om te beginnen maken we het probleem wat eenvoudiger. Als er wél gehele x, y, z ongelijk aan nul zijn zodat $x^p + y^p = z^p$, dan is er zoals we weten ook een oplossing met x_0, y_0, z_0 paarsgewijs copriem. Bovendien is dan $x_0, y_0, -z_0$ oplossing van de meer symmetrische vergelijking $x^p + y^p + z^p = 0$, want p is oneven; en $x_0, y_0, -z_0$ zijn nog steeds ongelijk 0 en paarsgewijs copriem. Het is daarom voldoende om te laten zien dat er geen oplossingen zijn van

$$x^p + y^p + z^p = 0 \qquad \text{met } x, y, z \in \mathbb{Z} - \{0\} \text{ paarsgewijs copriem.} \tag{9.11}$$

[16] Ons bewijs is een bewerking van de analoge bewijzen in [4] en [10].

We bewijzen dit uit het ongerijmde: we gaan er vanuit dat er wél een oplossing x, y, z bestaat. De eerste stap is dat we $x^p + y^p = (-z)^p$ ontbinden zoals in (9.1):

$$(-z)^p = (x+y)(x+\zeta y)(x+\zeta^2 y)\cdots(x+\zeta^{p-1}y). \tag{9.12}$$

Een natuurlijke vraag is of de factoren in deze ontbinding relatief priem zijn. Stel dat twee van hen, zeg $a = (x+\zeta^{k+l}y)$ en $b = (x+\zeta^k y)$, een niet-triviale deler[17] $\mu \in \mathbb{Z}[\zeta]$ gemeen hebben. Dan deelt μ ook het verschil

$$(x+\zeta^{k+l}y) - (x+\zeta^k y) = \zeta^k(\zeta^l - 1)y = u(\zeta-1)y \tag{9.13}$$

met u een eenheid, want ζ^k is een eenheid, en $\zeta^l - 1$ is volgens Lemma 9.3.5 een eenheid maal $\zeta - 1$. Bovendien deelt μ ook $a - \zeta^l b$, dat is

$$(x+\zeta^{k+l}y) - \zeta^l(x+\zeta^k y) = (1-\zeta^l)x = w(\zeta-1)x$$

met w een eenheid. We hebben dus

$$\mu|(\zeta-1)y \qquad \text{en} \qquad \mu|(\zeta-1)x. \tag{9.14}$$

Als a en b niet relatief priem zouden zijn, dan zouden ze ook een priemdeler gemeen hebben. Stel dus μ is priem, dan is μ deler van $(\zeta-1)$ of van y, en μ deelt $(\zeta-1)$ of van x. Als μ geen deler zou zijn van $\zeta - 1$, dan hebben we dus dat μ een gemene priemdeler is van x en y. Maar die zijn copriem in $\mathbb{Z}$, we krijgen dus een tegenspraak met Lemma 9.3.2. Dus μ is deler van $\zeta - 1$, en omdat $\zeta - 1$ priem is volgt dat $\mu = \zeta - 1$. Dit is dus de enige mogelijke gemene priemdeler van elk willekeurig paar verschillende factoren a en b. Elke gemene deler van hen is dus een macht van $\zeta - 1$. Maar we kunnen hoogstens één keer een gemene priemfactor $\zeta - 1$ van a en b wegdelen, want $(\zeta-1)^2$ kan geen gemene deler van hen zijn. Immers, als we $\mu = (\zeta-1)^2$ invullen in (9.14), dan volgt met Lemma 9.2.4 dat $\zeta - 1$ gemene deler is van y en x, en we zagen net dat dat niet kan. We concluderen:

De enige mogelijke niet-triviale gemene deler van een paar factoren van (9.12) is $\zeta - 1$. (9.15)

Stel nu dat $\zeta - 1$ ook maar een van de factoren deelt, zeg $(x+\zeta^k y)$. We kunnen (9.13) schrijven als

$$(x+\zeta^{k+l}y) = u(\zeta-1)y + (x+\zeta^k y)$$

waarbij we l laten variëren[18] van 0 tot en met $p-1$ en waarbij u nog van l afhangt. Omdat $\zeta - 1$ de twee termen in het rechterlid deelt, deelt hij ook $(x+\zeta^{k+l}y)$, en omdat die *elk* van de $p-1$ factoren doorloopt, concluderen we dat $\zeta - 1$ alle factoren van (9.12) deelt. Hiermee zien we dat $(\zeta-1)^{p-1}$ deler is van $(-z)^p = -z^p$, en van Gevolg 9.3.6 weten we dat dit betekent dat p deler is van $-z^p$. Maar p is priem, dus dit impliceert dat p deler is van z. Dus als p niet z deelt, dat deelt $\zeta - 1$ geen van de factoren, en de factoren zijn dus relatief priem. Omgekeerd, als p wel z deelt, dan is $(\zeta-1)^{p-1}$ deler van z, en dus is $\zeta - 1$ deler van $(-z)^p$. Omdat $\zeta - 1$

[17] dat wil zeggen, een deler die geen eenheid is

[18] We hoeven ons geen zorgen te maken over negatieve machten van ζ, want die machten zijn alleen modulo p bepaald.

priem is, moet hij een van de factoren van (9.12) delen, en deelt ze dus gelijk allemaal. We kunnen dus op natuurlijke manier twee gevallen onderscheiden, die traditioneel geval 1 en 2 worden genoemd: p is respectievelijk niet of wel deler van z. Geval 2 is veruit het lastigst. Omdat (9.11) symmetrisch is in x, y, z, kunnen we indien p minstens één van de x, y, z deelt, zonder beperking van de algemeenheid veronderstellen dat het z is. Daarom hebben we:

Geval 1. *Geen van de getallen x, y, z is deelbaar door p.* Als gevolg zijn de factoren in (9.12) paarsgewijs copriem.

Geval 2. *p is deler van z.* Als gevolg deelt $\zeta - 1$ elk van de factoren in (9.12). Uit (9.15) volgt bovendien dat de quotiënten van de factoren bij deling door $\zeta - 1$ paarsgewijs copriem zijn.

Bewijs van geval 1. De factoren $(x + y), (x + \zeta y), \ldots, (x + \zeta^{p-1} y)$ zijn paarsgewijs relatief priem en hun product $(-z)^p$ is een p-de macht, dus volgens Lemma 9.4.6 zijn het allen op eenheden na p-de machten. In het bijzonder is er een eenheid u en een cyclotomisch getal γ zodat

$$x + y\zeta = u\gamma^p.$$

We gaan nu een paar eigenschappen van rekenen modulo p gebruiken. Uit Lemma 9.4.2 volgt dat er een geheel getal c is zodat

$$x + y\zeta \equiv u\gamma^p \equiv uc \pmod{p}. \tag{9.16}$$

Voor complexe getallen $\alpha = a + bi$ duiden we de complex geconjugeerde aan met $\overline{\alpha}$, dus $\overline{\alpha} = a - bi$.[19] De inverse van ζ is $\overline{\zeta}$, want omdat ζ op de eenheidscirkel ligt is $\zeta\overline{\zeta} = |\zeta|^2 = 1$. Bovendien zijn x en y geheel, dus

$$\overline{x + y\zeta} = \overline{x} + \overline{y}\overline{\zeta} = x + y\overline{\zeta} = x + y\zeta^{-1}.$$

Er volgt dat

$$x + y\zeta^{-1} \equiv \overline{x + y\zeta} \equiv \overline{uc} \equiv \overline{u} \cdot \overline{c} \equiv \overline{u}c \pmod{p}. \tag{9.17}$$

Uit Lemma 9.4.4 volgt dat $u/\overline{u} = \zeta^k$ voor een gehele k, en met (9.16) en (9.17) volgt

$$x + y\zeta \equiv u/\overline{u} \cdot \overline{u}c \equiv u/\overline{u}(x + y\zeta^{-1}) \equiv \zeta^k(x + y\zeta^{-1}) \pmod{p}. \tag{9.18}$$

We proberen voorwaarden te vinden waar k, x en y aan moeten voldoen. Allereerst merken we op dat k alleen modulo p is bepaald; we kunnen zonder beperking aannemen dat $0 \leq k \leq p-1$. Stel dat $k = 0$. Dan hebben we dus $x + y\zeta \equiv x + y\zeta^{-1} \pmod{p}$, zodat $y(\zeta - \zeta^{-1}) \equiv 0 \pmod{p}$. Door aan beide kanten met $-\zeta$ te vermenigvuldigen, volgt dat $y(1 - \zeta^2) \equiv 0 \pmod{p}$, ofwel

$$p | y(1 - \zeta^2).$$

Er is dus een cyclotomisch getal α zodat $\alpha p = y(\zeta^2 - 1)$. Uit Lemma 9.3.1 volgt hiermee dat

$$\alpha(1 - \zeta)(1 - \zeta^2) \cdots (1 - \zeta^{p-1}) = y(\zeta^2 - 1),$$

[19] De notatie $\overline{k}$ voor restklassen modulo een geheel getal zullen we in dit bewijs niet gebruiken.

en omdat we in een domein werken en $(1-\zeta^2)$ niet nul is, mogen we aan beide kanten de factor (ζ^2-1) wegstrepen: er volgt dat $1-\zeta$ deler is van y. We weten dat $1-\zeta$ ook deler is van p, en omdat we in Geval 1 zitten, zijn y en p copriem. Bovendien weten we van Lemma 9.3.5 dat $1-\zeta$ irreducibel is. Kortom, we hebben

$$(1-\zeta)|y, \qquad (1-\zeta)|p, \qquad y \text{ en } p \text{ zijn copriem}, \qquad 1-\zeta \text{ is irreducibel.}$$

Dit is in tegenspraak met Lemma 9.3.2. We concluderen dat de aanname onjuist was, dus $k \neq 0$.

We weten dus dat $1 \leq k \leq p-1$. De congruentie $\zeta^k(x+y\zeta^{-1}) \equiv x+y\zeta \pmod p$ van (9.18) kunnen we schrijven als

$$\zeta^{k-1}(x\zeta+y) \equiv x+y\zeta \pmod p, \tag{9.19}$$

ofwel, in termen van $\zeta-1$,

$$\Big(1+(\zeta-1)\Big)^{k-1}\Big((x+y)+x(\zeta-1)\Big) \equiv (x+y)+y(\zeta-1) \pmod{(\zeta-1)^{p-1}}, \tag{9.20}$$

met $k-1 \geq 0$. We hebben hierbij Lemma 9.3.7 gebruikt, die zegt dat congruentie modulo p hetzelfde is als congruentie modulo $(\zeta-1)^{p-1}$. Dit zullen we voortaan gebruiken zonder het te noemen.

Met bijvoorbeeld het binomium van Newton zouden we het verschil van de linker- en rechterterm van (9.20) kunnen uitschrijven als 'polynoom' met 'variabele' $\zeta-1$ en gehele 'coëfficiënten'. We krijgen dan gehele getallen $a_0, \ldots, a_{p-1}$ zodat

$$a_0 + a_1(\zeta-1) + a_2(\zeta-1)^2 + \ldots + a_{p-1}(\zeta-1)^{p-1} \equiv 0 \pmod{(\zeta-1)^{p-1}}.$$

(De term $a_{p-1}(\zeta-1)^{p-1}$ is natuurlijk congruent 0 modulo $(\zeta-1)^{p-1}$ en kunnen we net zo goed weglaten.) Merk op dat het rechterlid van (9.20) alleen bijdraagt aan a_0 en a_1. Uit Lemma 9.4.3 volgt dat $a_0 \equiv a_1 \equiv \ldots \equiv a_{p-2} \equiv 0 \pmod p$.

Stel dat $2 \leq k \leq p-2$. Dan zou de hoogste orde term in het linkerlid van (9.20) gelijk zijn aan $(\zeta-1)^{k-1} \cdot x(\zeta-1) = x(\zeta-1)^k$, en in het rechterlid komt (wegens $k \geq 2$) geen term van die orde voor, dus $a_k = x$. Omdat $k \leq p-2$, volgt nu dat $x \equiv a_k \equiv 0 \pmod p$. Dus $p|x$, tegenspraak want we zitten in geval 1. De enige overgebleven gevallen zijn $k=1$ en $k=p-1$. Stel $k=p-1$. Dan is de term van orde $k-1$ gelijk aan

$$\begin{aligned}(\zeta-1)^{k-1}\cdot(x+y)+(k-1)(\zeta-1)^{k-2}\cdot x(\zeta-1) &= (x+y+x(k-1))(\zeta-1)^{k-1} \\ &= (x(p-1)+y)(\zeta-1)^{p-2},\end{aligned}$$

dus $a_{p-2} = x(p-1)+y = px+y-x$. Er volgt dat $y-x \equiv px+y-x \equiv a_{p-2} \equiv 0 \pmod p$, dus

$$x \equiv y \pmod p.$$

In het andere geval, $k=1$, krijgen we dezelfde conclusie. Dan luidt (9.19) namelijk $x\zeta+y \equiv x+y\zeta \pmod p$, dus

$$(x-y)(\zeta-1) \equiv 0 \pmod{(\zeta-1)^{p-1}}.$$

Uit Lemma 9.4.3 volgt nu dat $x-y \equiv a_1 \equiv 0 \pmod p$, dus ook nu volgt $x \equiv y \pmod p$.

We hebben nu afgeleid dat als x, y, z oplossing is van (9.11) in Geval 1, dan is $x \equiv y \pmod{p}$. Omdat de vergelijking symmetrisch is in x, y, z, volgt dat ook $x \equiv z, y \equiv z \pmod{p}$, ofwel

$$x \equiv y \equiv z \pmod{p}.$$

Omdat p priem is, weten we van de Kleine stelling van Fermat (Gevolg 6.2.7) dat

$$x^p \equiv x, \qquad y^p \equiv y, \qquad z^p \equiv z \pmod{p}.$$

We concluderen dat

$$0 = x^p + y^p + z^p \equiv x + y + z \equiv 3x \pmod{p},$$

dus $p|3x$. Omdat p priem is volgt $p|3$ of $p|x$. We zitten in geval 1, dus p deelt niet x, en er volgt $p = 3$. We hebben FLT echter al bewezen voor $p = 3$. We concluderen dat geval 1 onmogelijk is voor priemgetallen p waarvoor $\mathbb{Z}[\zeta_p]$ een ontbindingsring is.

Bewijs van geval 2. Het idee is weer dat we Fermat's methode van *infinite descent* gebruiken. Net als in het bewijs van $n = 4$ lukt het niet om een kleinere oplossing te vinden van de vergelijking $x^p + y^p + z^p = 0$ zelf, maar wel van een andere vergelijking die er veel op lijkt. Van die andere vergelijking kunnen we met infinite descent bewijzen dat hij geen oplossingen heeft.

In geval 2 is p deler van z, en $\zeta - 1$ deelt alle factoren van (9.12), de qotiënten $\frac{x+y\zeta^k}{\zeta-1}$ zijn paarsgewijs relatief priem. Het product van de quotiënten is $\frac{z^p}{(\zeta-1)^p} = (\frac{z}{\zeta-1})^p$. Hierbij is $\frac{z}{\zeta-1}$ een cyclotomisch getal, want $\zeta - 1$ is deler van p en dus van z. Het product van de quotiënten is dus een p-de macht, en omdat $\mathbb{Z}[\zeta]$ bij aanname een ontbindingsring is, volgt uit Lemma 9.4.6 dat er voor elke k een cyclotomisch getal τ_k en een eenheid u_k is[20] zodat

$$\frac{x + y\zeta^k}{\zeta - 1} = u_k \tau_k^p. \tag{9.21}$$

Elke gemene deler van twee getallen τ_j, τ_l is ook gemene deler van $\frac{x+y\zeta^j}{\zeta-1}, \frac{x+y\zeta^l}{\zeta-1}$, en omdat die relatief priem zijn, zijn ook τ_j, τ_l relatief priem. De τ_k zijn dus paarsgewijs relatief priem. In het bijzonder kan het priemelement $\zeta - 1$ maar één van hen delen. We zullen nu laten zien dat hij er inderdaad een deelt, namelijk τ_0. Omdat $p|z$ is $-z^p \equiv 0 \pmod{p}$. Uit de Kleine stelling van Fermat volgt dat $x^p \equiv x, y^p \equiv y \pmod{p}$, dus

$$0 \equiv -z^p \equiv x^p + y^p \equiv x + y \pmod{p}.$$

Hieruit volgt dat $x + y \equiv 0 \pmod{(\zeta - 1)^{p-1}}$, dus $(\zeta - 1)^{p-2}$ deelt $\frac{x+y}{\zeta-1} = u_0\tau_0^p$. Dus $\zeta - 1$ deelt τ_0^p, en omdat hij priem is deelt hij τ_0.

We halen de grootste factor $\zeta - 1$ uit τ_0: we schrijven $\tau_0 = (\zeta - 1)^m\delta$ met $m \geq 1$ en δ niet deelbaar door $\zeta - 1$. Nu schrijven we (9.21) uit voor $k = -1, 0, 1$, waarbij we gebruik maken van deze formule voor τ_0.

$$\begin{aligned} x + y\zeta^{-1} &= (\zeta - 1)u_{-1}\tau_{-1}^p \\ x + y &= (\zeta - 1)u_0(\zeta - 1)^{mp}\delta^p \\ x + y\zeta &= (\zeta - 1)u_1\tau_1^p \end{aligned} \tag{9.22}$$

[20] We lezen de getallen k modulo p.

Hierbij zijn $\tau_{-1}, \delta, \tau_1$ paarsgewijs copriem, want $\tau_{-1}, \tau_0, \tau_1$ zijn dat ook. Uit deze vergelijkingen kunnen we x en y elimineren. De tweede van de derde aftrekken, en de eerste van de tweede levert achtereenvolgens

$$\begin{aligned}(\zeta-1)y &= (\zeta-1)\Big(u_1\tau_1^p - u_0(\zeta-1)^{mp}\delta^p\Big),\\ \zeta^{-1}(\zeta-1)y = (1-\zeta^{-1})y &= (\zeta-1)\Big(u_0(\zeta-1)^{mp}\delta^p - u_{-1}\tau_{-1}^p\Big).\end{aligned}$$

Van deze twee vergelijkingen trekken we de laatste ζ maal af van de eerste, zodat het linkerlid wegvalt:

$$\begin{aligned}0 &= (\zeta-1)\Big(u_1\tau_1^p - u_0(\zeta-1)^{mp}\delta^p - \zeta u_0(\zeta-1)^{mp}\delta^p + \zeta u_{-1}\tau_{-1}^p\Big)\\ &= (\zeta-1)\Big(u_1\tau_1^p - (1+\zeta)u_0(\zeta-1)^{mp}\delta^p + \zeta u_{-1}\tau_{-1}^p\Big)\\ &= u_1\tau_1^p - (1+\zeta)u_0(\zeta-1)^{mp}\delta^p + \zeta u_{-1}\tau_{-1}^p.\end{aligned}$$

De laatste gelijkheid volgt omdat $\zeta-1 \neq 0$ en $\mathbb{Z}[\zeta]$ geen nuldelers heeft. Omdat $(\zeta-1)(\zeta+1) = \zeta^2-1$, en $\zeta-1, \zeta^2-1$ norm p hebben, volgt dat $\zeta+1$ norm 1 heeft en dus een eenheid is. We kunnen de vergelijking daarom herschrijven als

$$e_0(\zeta-1)^{mp}\delta^p = \tau_1^p + e_{-1}\tau_{-1}^p \tag{9.23}$$

waarbij $e_0 = u_1^{-1}(1+\zeta)u_0$ en $e_{-1} = u_1^{-1}\zeta u_{-1}$ eenheden zijn. Deze vergelijking heeft al veel weg van $z^p = x^p + y^p$, en we kunnen er nog dichterbij komen door de eenheid e_{-1} te elimineren. We weten van Lemma 9.4.2 dat cyclotomische p-de machten modulo p congruent zijn met een geheel getal, dus $\tau_1^p \equiv c_1$ en $\tau_{-1}^p \equiv c_2 \pmod p$ voor zekere gehele c_1, c_2. Bovendien is p een eenheid maal $(\zeta-1)^p$, dus p deelt $e_0(\zeta-1)^{mp}\delta^p$, ofwel $e_0(\zeta-1)^{mp}\delta^p \equiv 0 \pmod p$. Bovenstaande vergelijking wordt modulo p dus

$$0 \equiv c_1 + e_{-1}c_2 \pmod p,$$

en er volgt $e_{-1} \equiv -c_1c_2^{-1} \pmod p$: modulo p is e_{-1} congruent met een *geheel getal*. Uit Voorwaarde 9.4.5 volgt hiermee dat er een eenheid w is zodat $e_{-1} = w^p$. Dus (9.23) kunnen we schrijven als

$$\tau_1^p + (w\tau_{-1})^p = e_0(\zeta-1)^{mp}\delta^p.$$

Dit is een oplossing van de Fermat-achtige vergelijking

$$\alpha^p + \beta^p = e(\zeta-1)^{mp}\gamma^p, \tag{9.24}$$

waarbij e een eenheid is, $m \geq 1$ een geheel getal en α, β, γ paarsgewijs coprieme *cyclotomische* getallen die geen van allen deelbaar zijn door $\zeta-1$. Van deze vergelijking kunnen we met infinite descent bewijzen dat hij geen oplossing heeft. Stel er is wél een oplossing α, β, γ. We factoriseren $\alpha^p + \beta^p$ weer als in (9.12):

$$e(\zeta-1)^{mp}\gamma^p = (\alpha+\beta)(\alpha+\zeta\beta)(\alpha+\zeta^2\beta)\cdots(\alpha+\zeta^{p-1}\beta). \tag{9.25}$$

Dat α en β niet noodzakelijk geheel zijn, doet er voor deze factorisatie niet toe. Omdat $\zeta - 1$ het linkerlid deelt, deelt het minstens één van de factoren in het rechterlid, en zoals hierboven (pagina 96 na (9.15)) volgt dat hij alle factoren deelt. In die redenering hebben we namelijk niet gebruikt dat x, y, z geheel zijn. Bovendien zijn de quotiënten bij deling door $\zeta - 1$ relatief priem, want anders zou zoals na (9.14) volgen dat $\zeta - 1$ gemene deler is van α en β, maar die zijn copriem. In het bijzonder deelt $\zeta - 1$ hoogstens een van de quotiënten. Het argument dat $\alpha - 1$ *precies* een van hen deelt (in het vorige geval was dat $u_0\tau_0^p$), gaat echter niet meer op. Het is echter wel weer het geval, zoals we nu gaan bewijzen.

Volgens Lemma 9.4.3 zijn er gehele getallen a_0, a_1, b_0, b_1 zodat

$$\begin{aligned}\alpha &\equiv a_0 + a_1(\zeta - 1),\\ \beta &\equiv b_0 + b_1(\zeta - 1) \pmod{(\zeta - 1)^2}.\end{aligned}$$

We kunnen elke factor $\alpha + \beta\zeta^k$ daarom modulo $(\zeta - 1)^2$ schrijven als

$$\alpha + \beta\zeta^k \equiv (a_0 + a_1(\zeta - 1)) + (1 + (\zeta - 1))^k(b_0 + b_1(\zeta - 1)) \pmod{(\zeta - 1)^2}.$$

Als we $(1 + (\zeta - 1))^k$ in gedachten met het binomium van Newton uitwerken, zien we dat 1 en $k(\zeta-1)$ de enige termen zijn die geen veelvoud zijn van $(\zeta-1)^2$. Dus $(1+(\zeta-1))^k \equiv 1+k(\zeta-1)$ (mod $(\zeta - 1)^2$), dus we kunnen bovenstaande vergelijking schrijven als

$$\alpha + \beta\zeta^k \equiv a_0 + b_0 + (\zeta - 1)(a_1 + b_1 + kb_0) \pmod{(\zeta - 1)^2}. \tag{9.26}$$

Deze vergelijking bekijken we even 'met een graad van nauwkeurigheid minder', namelijk modulo $\zeta - 1$: omdat $\alpha + \beta\zeta^k \equiv 0 \pmod{\zeta - 1}$ krijgen we $a_0 + b_0 \equiv 0 \pmod{\zeta - 1}$. Omdat a_0, b_0 geheel zijn, volgt uit Lemma 9.3.4 dat $a_0 + b_0 \equiv 0 \pmod{p}$. In het bijzonder is $a_0 + b_0 \equiv 0$ (mod $(\zeta - 1)^2$), want $(\zeta - 1)^2$ deelt p; dus (9.26) kunnen we vereenvoudigen tot

$$\alpha + \beta\zeta^k \equiv (\zeta - 1)(a_1 + b_1 + kb_0) \pmod{(\zeta - 1)^2}. \tag{9.27}$$

Stel dat een van de factoren $\alpha + \beta\zeta^j$ deelbaar is door $(\zeta - 1)^2$. Dan volgt $0 \equiv (\zeta - 1)(a_1 + b_1 + jb_0) \pmod{(\zeta-1)^2}$, dus $(\zeta-1)^2$ deelt $(\zeta-1)(a_1+b_1+jb_0)$ zodat $\zeta-1$ deelt $a_1+b_1+jb_0$. Omdat dit een geheel getal is, volgt zoals net dat $a_1 + b_1 + jb_0 \equiv 0 \pmod{p}$. Omgekeerd, als $a_1 + b_1 + jb_0 \equiv 0 \pmod{p}$, dan is $(\zeta - 1)^2$ deler van $a_1 + b_1 + jb_0$ en volgens (9.27) is $(\zeta - 1)^2$ deler van $\alpha + \beta\zeta^j$. Dus

$$(\zeta - 1)^2 \text{ deelt } \alpha + \beta\zeta^j \quad\Longleftrightarrow\quad jb_0 \equiv -a_1 - b_1 \pmod{p}.$$

Omdat $\mathbb{Z}/p\mathbb{Z}$ een lichaam is, is er een $c \in \mathbb{Z}$ zodat $cb_0 \equiv 1 \pmod{p}$, dus $jb_0 \equiv -a_1-b_1 \pmod{p}$ precies dan als $j \equiv c(-a_1 - b_1) \pmod{p}$. Er is dus precies één $j \in \{0, 1, 2, \ldots, p - 1\}$, laten we die k noemen, waarvoor $(\zeta - 1)^2$ deler is van de factor $\alpha + \beta\zeta^j$. Omdat elk van de andere factoren deelbaar is door $\zeta - 1$, kunnen we in het rechterlid van (9.25) minstens $p + 1$ factoren $\zeta - 1$ buiten haakjes halen. Het linkerlid bevat precies mp factoren $\zeta - 1$, want e en γ^p zijn niet deelbaar door $\zeta - 1$, en we concluderen dat $m > 1$, zeg $m = M + 1$ met $M \geq 1$ geheel.

Door eventueel de volgorde van de factoren te verwisselen, kunnen we (9.25) herschrijven als

$$\begin{aligned} e(\zeta-1)^{mp}\gamma^p &= (\alpha+\zeta^k\beta)(\alpha+\zeta^{k+1}\beta)(\alpha+\zeta^{k+2}\beta)\cdots(\alpha+\zeta^{k-1}\beta) \\ &= (\alpha+\beta_0)(\alpha+\zeta\beta_0)(\alpha+\zeta^2\beta_0)\cdots(\alpha+\zeta^{p-1}\beta_0) \\ &= \alpha^p+\beta_0^p \end{aligned}$$

met $\beta_0=\zeta^k\beta$ niet deelbaar door $\zeta-1$ (want ζ^k is een eenheid). Nu is $\alpha+\beta_0$ dus de factor die deelbaar is door $(\zeta-1)^2$. Elk van de overige $p-1$ factoren $\alpha+\zeta^j\beta_0$ bevat precies één factor $\zeta-1$, en in totaal bevat het rechterlid mp factoren $\zeta-1$. Dat betekent dat $\alpha+\beta_0$ precies $mp-(p-1)=(m-1)p+1=Mp+1$ factoren $\zeta-1$ bevat. Omdat

$$e\Big(\frac{(\zeta-1)^m\gamma}{\zeta-1}\Big)^p=\frac{\alpha+\beta_0}{\zeta-1}\cdot\frac{\alpha+\zeta\beta_0}{\zeta-1}\cdot\ldots\cdot\frac{\alpha+\zeta^{p-1}\beta_0}{\zeta-1}$$

een eenheid maal een p-de macht is, volgt uit Lemma 9.4.6 dat elk van de factoren $\frac{\alpha+\beta_0\zeta^j}{\zeta-1}$ dat ook zijn: we schrijven ze als $u_j\tau_j^p$ voor een eenheid u_j en een cyclotomisch getal τ_j. De τ_i zijn paarsgewijs copriem, dat volgt uit het stukje na (9.25). Bovendien is τ_j voor $j\geq 1$ niet deelbaar door $\zeta-1$, en voor $j=0$ kunnen we nog M factoren $\zeta-1$ uit τ_0 halen: $\tau_0=(\zeta-1)^M\delta^p$ voor een δ niet deelbaar door $\zeta-1$. We hebben dus voor $j=-1,0,1$

$$\begin{aligned} \alpha+\beta_0\zeta^{-1} &= (\zeta-1)u_{-1}\tau_{-1}^p \\ \alpha+\beta_0 &= (\zeta-1)u_0(\zeta-1)^{Mp}\delta^p \\ \alpha+\beta_0\zeta &= (\zeta-1)u_1\tau_1^p. \end{aligned}$$

Dit is precies (9.22), maar dan met α,β_0 in plaats van x,y, en M in plaats van m.[21] Door met deze nieuwe de redenering tussen (9.22) en (9.23) te doorlopen, die alleen gebruik maakt van eigenschappen van x,y,u_i,τ_i waar de nieuwe $\alpha,\beta_0,u_i,\tau_i$ ook aan voldoen, krijgen we weer een oplossing van vergelijking (9.24), *maar nu met $M=m-1$ in plaats van m*. Door dit proces steeds weer te herhalen, krijgen we een oneindig lange afdalende rij $m=m_0,M=m_1,m_2,m_3,\ldots$ van natuurlijke getallen, en dat kan niet. We concluderen dat er geen oplossing is van (9.24). En omdat uit het bestaan van een oplossing van geval 2 het bestaan van een oplossing van (9.24) volgt, heeft geval 2 ook geen oplossing. Dit voltooit het bewijs.

[21] En met in het algemeen andere getallen u_i,τ_i, maar omdat dat alleen hulpvariabelen zijn hebben we ze niet omgenoemd.

Hoofdstuk 10

Epiloog

'I think I'll stop here', met deze woorden eindigde Andrew Wiles zijn beroemde Cambridge-lezing van 1993 waarin hij het Taniyama-Shimura vermoeden en daarmee de Laatste Stelling van Fermat bewees. Hier zijn deze woorden niet zo op z'n plaats: we hebben de stelling slechts in het geheel bewezen voor $n = 3$ en 4, en voor Geval 1 van $n = 5$. Zojuist hebben we het bewezen voor alle $n = p > 2$ waarvoor $\mathbb{Z}[\zeta_p]$ een ontbindingsring is,[1] en we merkten op dat in elk geval alle $p < 23$ hieraan voldoen. Dat hebben we echter niet bewezen, we hebben zelfs niet geschetst hoe je dat aan zou kunnen pakken. Toch hopen we dat we, ondanks deze onvolledigheden, toch in zekere zin de kern van het probleem hebben geraakt, en de lezer geïnspireerd hebben tot verdere studie. En daar was het ons vooral om te doen: in de woorden van Wiles, 'The definition of a good mathematical problem is the mathematics it generates rather than the problem itself.'

Als je je verder wilt verdiepen in de wiskunde die direct over de Laatste Stelling gaat, kunnen we vooral [4] aanraden; dit is een pre-Wiles boek, en het bewijst de stelling voor alle reguliere priemgetallen, die we in de inleiding van hoofdstuk 9 al genoemd hebben. Er zijn veel meer reguliere priemen (naar het schijnt zo'n 61%) dan priemen waarvoor $\mathbb{Z}[\zeta_p]$ een ontbindingsring is, en bovendien wordt in dat boek beschreven hoe je kunt achterhalen of een priem regulier is.

Goede introducties tot de (elementaire) getaltheorie zijn bijvoorbeeld [2] van Frits Beukers en het beroemde boek [14] van Hardy en Wright. Op een hoger niveau raden we het boek [20] van Cohen aan over 'expliciete getaltheorie', en het toegankelijkere boek [15] van Apostol bedoeld voor bachelorstudenten over analytische getaltheorie.

Toen Wiles zijn bewijs publiceerde was men bang dat, hoe mooi het ook was, het slechts toegankelijk was voor een klein groepje experts, maar zeker nu een aantal wiskundigen het bewijs behoorlijk vereenvoudigd hebben is daar geen sprake meer van. Er is zelfs in 1997 een mooi boek [16] verschenen waarin het bewijs in vijfhonderd pagina's inzichtelijk wordt gemaakt voor gevorderde masterstudenten. Hiervoor is echter wel een basispakket aan voorkennis en -kunde nodig. Vele takken van de huidige wiskunde worden in Wiles' stelling naar voren gebracht, en behalve noodzakelijk om het bewijs te begrijpen is het ook zeer de moeite waard om daar vertrouwd mee te raken, en essentieel als je verder wilt in de algebra. Globaal gaat

[1] hoewel ook niet helemaal, want we hebben beweerd maar niet bewezen dat de 'voorwaarde' altijd geldt voor ontbindingsringen

het bewijs over de in de inleiding al genoemde interactie tussen *elliptische krommen* en *modulaire vormen*. Van deze twee zijn de eerste het meest toegankelijk, het beroemde boek [13] van Silverman en Tate bedoeld voor gevorderde bachelorstudenten raden we ten sterkste aan. Modulaire vormen laten zich niet zo makkelijk vangen zonder er een hoop voorkennis op los te laten, met name complexe analyse is belangrijk voor je aan de studie daarvan kan beginnen. Een goede introductie voor masterstudenten is [17].

Een groot deel van dit boek gaat over algebra en algebraïsche getaltheorie. Tegenwoordig het meest invloedrijke boek over algebra is [18] van Serge Lang, dit is een zeer goed boek en bijvoorbeeld de dictaten [7], [8] en [9] van Peter Stevenhagen bieden genoeg voorkennis. Als je je verder wilt verdiepen in de algebraïsche getaltheorie heb je eerst een heel arsenaal aan algebra nodig (Lang's boek biedt meer dan genoeg voorkennis), we kunnen het beste [19] aanbevelen. Het bewijs van Wiles maakt gebruik van resultaten uit nog vele andere vakgebieden, om de belangrijkste te noemen: *representatietheorie* van (galois)groepen (zie bijvoorbeeld [21], geschikt voor iedereen die een basiscursus groepentheorie heeft gevolgd) en *Algebraïsche Meetkunde* (het beroemde boek [22] is in elk geval toegankelijk voor wie het minder veeleisende [23] heeft doorgewerkt).

Al met al kunnen we concluderen dat geduld en doorzettingsvermogen bij de belangrijkste ingrediënten horen om Wiles' bewijs te begrijpen en in het algemeen om verder te komen in de wiskunde, net als het besef dat niet de vermoedens en stellingen op zichzelf het belangrijkst zijn, maar de wiskundige ideeën daaromheen.

Bibliografie

[1] M. A. Armstrong, *Groups and Symmetry*, Springer-Verlag 1988. Dit boek diende als inspiratiebron voor hoofdstuk 5 en 6.

[2] Frits Beukers, *Getaltheorie voor Beginners*, Epsilon Uitgaven 2008. Hier hebben we informatie over Fermat- en Mersennegetallen uit ontleend, evenals wat biografische gegevens van wiskundigen.

[3] Leonard Eugene Dickson, *Introduction tot the theory of numbers*, Dover publications 1957. Dit hebben we als algemene inspiratiebron gebruikt voor hoofdstuk 6.

[4] Harold M. Edwards, *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*, Springer-Verlag 1977 (GTM 50). Dit boek hebben we als leidraad gebruikt in hoofdstuk 2 t/m 4 en hoofdstuk 9. De aanpak van veel bewijzen uit deze hoofdstukken hebben we van dit boek overgenomen.

[5] M. Riemersma, *Algebra*, Epsilon Uitgaven 2003. Dit boek gebruikten we vooral bij de theorie van het ontbinding van veeltermen.

[6] Simon Singh, *Fermat's Last Theorem*, Fourth Estate Limited 1997. Veel informatie uit de introductie hebben we uit dit populair-wetenschappelijke boek gehaald. Het is geschreven naar aanleiding van de prachtige gelijknamige BBC 'horizons' documentaire.

[7] P. Stevenhagen, dictaat *Algebra I*, Universiteit Leiden en TU Delft, 2010. Dit en vooral het volgende dictaat hebben we als leidraad gebruikt voor delen van hoofdstuk 5 t/m 8. Op dit moment zijn de dictaten te downloaden van http://websites.math.leidenuniv.nl/algebra/.

[8] P. Stevenhagen, dictaat *Algebra II*, Universiteit Leiden en TU Delft, 2010.

[9] P. Stevenhagen, dictaat *Algebra III*, Universiteit Leiden en TU Delft, 2011.

[10] Ruden Teuben, kleine scriptie *De laatste stelling van Fermat voor reguliere priemen,* 2005. Hiervan hebben we een aantal ideeën gebruikt bij sommige bewijzen in hoofdstuk 9.

[11] http://en.wikipedia.org/wiki/Proof_of_Fermat%27s_Last_Theorem_for_specific_exponents. Hier stuitten we op een bewijsschets van geval 1 voor $n = 5$.

[12] Joseph Rotman, *Galois Theory*, Springer-Verlag 1990.

De volgende boeken, naast de hierboven genoemden, zijn toegankelijk voor iedereen die dit boek gelezen en begrepen heeft, en zeker voor gevorderde bachelorstudenten wiskunde.

[13] Joseph Silverman, John Tate, *Rational Points on Elliptic Curves, Springer-Verlag 1992*

[14] G.H. Hardy en E.M. Wright, *An introduction to the Theory of Numbers*, Oxford University Press 1938 (zesde editie 2008).

[15] Tom Apostol, *Introduction tot Analytic Number Theory*, Springer-Verlag 1976

[16] Gary Cornell, Joseph Silverman, Glenn Stevens e.a., *Modular Forms and Fermat's Last Theorem*, Springer-Verlag 1997

Deze boeken zijn alleen toegankelijk voor lezers met meer wiskundige ervaring.

[17] Fred Diamond, Jerry Schurman, *A first course in Modular Forms*, Springer-Verlag 2005 (GTM 228)

[18] Serge Lang, *Algebra*, Springer Verlag 2002 (GTM 211)

[19] Jrgen Neukirch, *Algebraic Number Theory*, Springer-Verlag 1999 (SMM 322)

[20] Henry Cohen, *Number Theory* Volume 1, Springer Verlag 2007 (GTM 239)

[21] Gordon James, Martin Liebeck, *Representations and Characters of Groups*, Cambridge University Press 1993

[22] Robin Hartshorne, *Algebraic Geometry*, Springer-Verlag 1977 (GTM 52)

[23] Joe Harris, *Algebraic Geometry: a first course*, Springer-Verlag 1992 (GTM 133)

Index