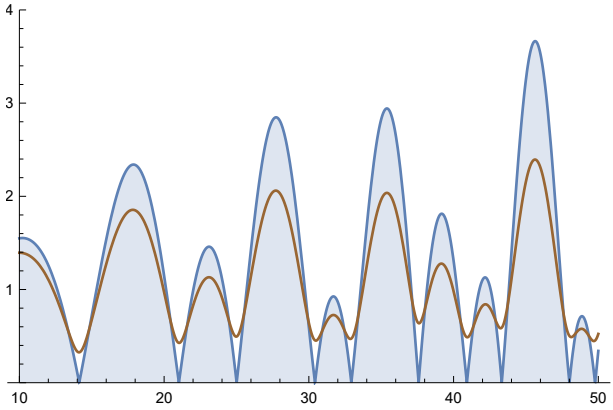
platform wiskunde nederland

PLATFORM WISKUNDE VLAANDEREN

Priemgetallen

Syllabus Vakantiecursus 2023

Antwerpen, 25 en 26 augustus 2023
Amsterdam, 1 en 2 september 2023



platform wiskunde nederland

PLATFORM WISKUNDE VLAANDEREN

Priemgetallen

Syllabus Vakantiecursus 2023

(gecorrigeerde versie)

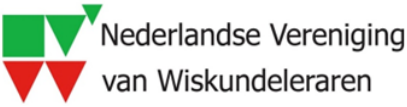
Antwerpen, 25 en 26 augustus 2023
Amsterdam, 1 en 2 september 2023

Programmacommissie

prof. dr. Wil Schilders (PWN, TU/e) (voorzitter)
dr. Jeroen Spandaw (TUD)
Joanne de Jager MSc (NVVW, Metis Montessori Lyceum)
dr. Benne de Weger (TU/e) (eindredactie syllabus)
drs. Peter Ypma (Goudse Waarden, Betapartners)

e-mail: `vakantiecursus@platformwiskunde.nl`

Sponsoren



Platform Wiskunde Nederland, Science Park 123, 1098 XG Amsterdam
Telefoon: 020-592 4006, Website: `http://www.platformwiskunde.nl`

Platform Wiskunde Vlaanderen
Website: `http://www.platformwiskunde.be`

Vakantiecursus 2023

De Vakantiecursus Wiskunde voor leraren in de exacte vakken op middelbare/hogere scholen en andere belangstellenden wordt al sinds 1946 jaarlijks gehouden, in eerste instantie op het in 1946 opgerichte Mathematisch Centrum (MC) in Amsterdam, tegenwoordig het Centrum Wiskunde en Informatica (CWI). Inmiddels is het een initiatief van de Nederlandse Vereniging van Wiskundeleraren (NVvW). Vanaf 2010 wordt de vakantiecursus georganiseerd door Platform Wiskunde Nederland (PWN), en de afgelopen jaren werd de cursus steeds in twee opeenvolgende weekeinden gehouden, met als locaties Amsterdam en Eindhoven. Omdat er ook vaak Vlaamse deelnemers zijn, en om de banden met het recent opgerichte Platform Wiskunde Vlaanderen (PWV) te versterken, ontstond in 2022 het idee om de Vakantiecursus voortaan gezamenlijk te gaan organiseren. Vandaar dat in 2023 de cursus voor het eerst in Vlaanderen plaatsvindt, namelijk aan de Universiteit van Antwerpen. We hopen dat veel Vlaamse deelnemers acte-de-prśence zullen geven!

De Vakantiecursus wordt mede mogelijk gemaakt door een subsidie van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), en bijdragen van 4TU.AMI, het toegepaste wiskunde-instituut van de 4 Nederlandse technische universiteiten, alsmede PWN. Organisatie vindt plaats in nauwe samenwerking met het Centrum voor Wiskunde en Informatica (CWI), de Technische Universiteit Eindhoven (TU/e), en de Universiteit van Antwerpen.

De presentaties van de sprekers zullen zo veel mogelijk beschikbaar komen op de PWN-website: `https://www.platformwiskunde.nl`.

Met dank aan

Ondersteuning UIA: prof. dr. Paul Levrie en Els Vanlommel.

Historie

De eerste vakantiecursus wordt in het jaarverslag 1946 van het Mathematisch Centrum als volgt vermeld:

> *Op 29 en 31 Oct. '46 werd onder auspiciën van het M.C. een druk bezochte en uitstekend geslaagde vacantiecursus gehouden voor wiskundeleeraren in Nederland. Op 29 October stond de wiskunde, op 31 October de didactiek van de wiskunde op de voorgrond. De sprekers waren: Prof.Dr. O. Bottema, "De prismoide", Dr. A. Heyting, "Punten in het oneindige", Mr. J. v. IJzeren, "Abstracte Meetkunde en haar betekenis voor de Schoolmeetkunde.", Dr. H.D. Kloosterman, "Ontbinding in factoren", Dr. G. Wielenga, "Is wiskunde-onderwijs voor alpha's noodzakelijk?", Dr. J. de Groot, "Het scheppend vermogen van den wiskundige" en Dr. N.L.H. Bunt, "Moeilijkheden van leerlingen bij het beginnend onderwijs in de meetkunde".*
>
> *Aan het einde van de vacantiecursus werden diverse zaken besproken die het wiskunde-onderwijs in Nederland betroffen. Een Commissie werd ingesteld, die het M.C. over de verder te organiseren vakantiecursussen van advies zou dienen. Hierin namen zitting een vertegenwoordiger van de Inspecteurs van het V.H. en M.O. benevens vertegenwoordigers van de lerarenverenigingen Wimecos en Liwenagel.*
>
> *Ook werd naar aanleiding van "wenschen" die tijdens de cursus naar voren gekomen waren ingesteld: "een colloquium over moderne Algebra, een dispuut over de didactiek van de wiskunde, beiden hoofdzakelijk bedoeld voor de leeraren uit Amsterdam en omgeving, terwijl tevens vanwege het M.C. een cursus over Getallenleer werd toegezegd te geven door de heeren v.d. Corput en Koksma. (Colloquium, dispuut en cursus zijn in 1947 gestart en verheugen zich in blijvende belangstelling).*

Docenten

prof. dr. Frits Beukers (hoofddocent)
Mathematisch Instituut, Universiteit Utrecht,
web: https://webspace.science.uu.nl/~beuke106/
e-mail: fritsbeukers@gmail.com

prof. dr. Roland van der Veen
Bernoulli Instituut, Universiteit Groningen,
web: http://www.rolandvdv.nl/
e-mail: r.i.van.der.veen@rug.nl

dr. Benne de Weger
Faculteit Wiskunde en Informatica, Technische Universiteit Eindhoven,
web: https://www.win.tue.nl/~bdeweger/
e-mail: b.m.m.d.weger@tue.nl

Programma

Vrijdag 25 augustus 2023 / 1 september 2023

15.00–15.30		*Ontvangst, koffie*
15.30–15.35		*Openingswoord*
15.35–16.20	Frits Beukers	Kennismaken met priemgetallen
16.20–16.45		*Pauze*
16.45–17:30	Frits Beukers	Middelbare school-methodenen
17.30–18.30		*Diner*
18.30–19.15	Frits Beukers	Priemtesten I
19.15–19.45		*Pauze*
19.45–20.30	Roland van der Veen	Priemgetallen volgens de zeta-functie I

Zaterdag 27 augustus 2022 / 3 september 2022

09.00–10.00		*Ontvangst, koffie*
10.00–10.45	Frits Beukers	Priemtesten II
10.45–11.15		*Pauze*
11.15–12.00	Benne de Weger	Priemgetal-heuristiek
12.00–13.00		*Lunch*
13.00–13.45	Frits Beukers	Priemgetal-formules
13.45–14.30	Roland van der Veen	Priemgetallen volgens de zeta-functie II
14.30		*Afsluiting*

1 Priemgetallen
Frits Beukers

1.1 Priemgetallen tellen

We beginnen met de hoofdrolspelers van deze cursus,

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61,$$
$$67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113, 127, 131, 137$$
$$139, 149, 151, 157, 163, 167, 173, 179, 181, 191, 193, 197, 199,$$
$$211, 223, 227, \ldots$$

Dit zijn de eerste 49 priemgetallen en de vraag rijst of deze rij oneindig lang doorgaat. Anders gezegd, neem een willekeurig getal N, bijvoorbeeld 10^{100}, en vervolgens kun je je afvragen of er een priemgetal bestaat dat groter is dan deze N. In de Griekse oudheid heeft Euclides deze vraag al beantwoord.

Stelling 1.1.1. *Euclides: Voor elk getal N bestaat een priemgetal $> N$.*

In het bijzonder impliceert dit dat er oneindig veel priemgetallen bestaan.

Bewijs. We voeren de volgende stappen uit:

- Neem het product van $1, 2, 3, \ldots, N$. Dit noteren we als $N!$ (spreek uit: 'N faculteit').
- Tel daar 1 bij op.
- Kies een priemdeler p van $N! + 1$.
- Stel dat $p \leq N$. Dan is p een deler van $N!$ maar p kan geen deler van zowel $N!$ als $N! + 1$ zijn.
- Dus concluderen we dat $p > N$. □

Het enige dat wij hier gebruikt hebben is dat elk geheel getal > 1 deelbaar is door een priemgetal. Kies hiertoe de kleinste deler > 1 van dit getal. Dat moet wel priem zijn.

Merk op dat we in het bewijs ook $N! - 1$ hadden kunnen nemen (er van uitgaande dat $N > 2$).

Opgave 1.1.2. *Laat op analoge manier zien dat er bij elke $N \geq 4$ een priemgetal groter dan N en van de vorm $4n-1$ bestaat. (Hint: bedenk dat $N!-1$ een 4-voud min 1 is).*

In de literatuur kun je talloze bewijzen vinden van de oneindigheid van de priemgetallen. Eén van de mooiste (maar niet de simpelste) naar mijn mening is dat van Euler. Leonhard Euler (1707-1783) is één van dè bekendste wiskundigen aller tijden.

Stelling 1.1.3. *De oneindige reeks*

$$\sum_{p \ priem} \frac{1}{p} = \frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{11} + \cdots$$

divergeert.

Ter herinnering

Opmerking 1.1.4 (Harmonische reeks)**.** *De oneindige reeks*

$$\sum_{n=1}^{\infty} \frac{1}{n} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \cdots$$

divergeert.

Bewijs van de stelling van Euler. In dit bewijs zijn sommen en producten genomen over p altijd genomen over *priemgetallen* p.
Kies N. Dan geldt door haakjes wegwerken,

$$\prod_{p \leq N} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right) > \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{N}.$$

Ten tweede,

$$\prod_{p \leq N} \left(1 + \frac{1}{p}\right)^2 > \prod_{p \leq N} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right)$$

Ten derde,

$$\sum_{p \leq N} \frac{1}{p} > \sum_{p \leq N} \log\left(1 + \frac{1}{p}\right) = \log \prod_{p \leq N} \left(1 + \frac{1}{p}\right)$$

NB: De logaritme log die we hier gebruiken is de *natuurlijke logaritme*. In het voortgezet onderwijs wordt deze veelal met ln aangegeven.

Zetten we deze drie ongelijkheden achter elkaar dan zien we dat

$$\sum_{p\leq N}\frac{1}{p} > \log\left(\prod_{p\leq N}\left(1+\frac{1}{p}\right)\right) > \frac{1}{2}\log\left(\frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{N}\right).$$

Omdat de harmonische reeks divergeert, divergeert ook de som over de priemgetallen. □

Hier is nog een bewijs.

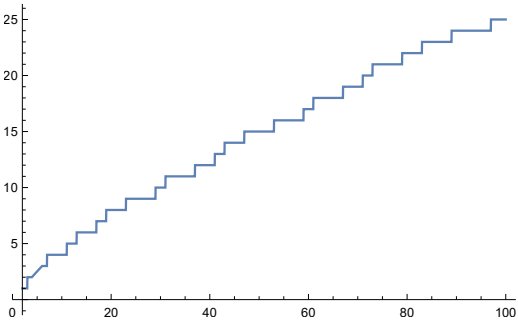
Derde bewijs oneindigheid van de priemgetallen.
Stel er zijn eindig veel priemgetallen. Noem ze $p_1, p_2, \ldots, p_n$.
Dan kan elk getal geschreven worden in de vorm $p_1^{k_1}p_2^{k_2}\cdots p_n^{k_n}$ met $k_1, k_2, \ldots, k_n \geq 0$.
Hoeveel getallen zijn er van deze vorm en $\leq x$?
Dan moet zeker gelden, $p_1^{k_1} \leq x, p_2^{k_2} \leq x, \ldots, p_n^{k_n} \leq x$.
Omdat $p_i \geq 2$ voor alle i volgt hieruit dat $k_i \leq \log x/\log 2$ voor alle i.
Het aantal getallen dat hiermee gevormd kan worden is hooguit $(\log x/\log 2)^n$.
Omdat $x > (\log x/\log 2)^n$ als x groot genoeg is hebben we een tegenspraak. □

We gaan nu tellen hoeveel priemgetallen er beneden een vaste grens zijn.

Definitie 1.1.5. *Definieer de priemtelfunctie*

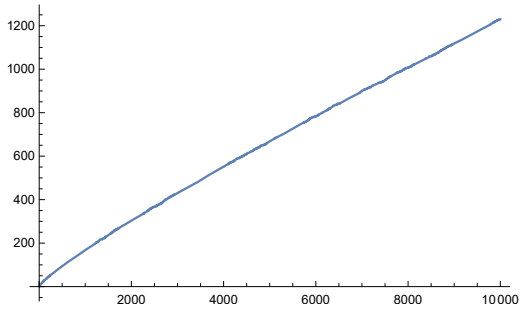
$$\pi(x) = \#\{p \leq x |\ p \text{ priem}\}.$$

Hier is een plot van $\pi(x)$ voor $x \leq 100$,



Figuur 1.1: $\pi(x)$ voor $x \leq 100$.

En een plot voor $x \leq 10000$,



Figuur 1.2: $\pi(x)$ voor $x \leq 10000$.

Een mooie glad uitziende grafiek. Vanaf de 18e eeuw is men op zoek gegaan naar eenvoudige functies die deze grafiek zo goed mogelijk benaderen.

Het dichtstbij kwam Johann Carl Friedrich Gauss (1777-1855), een van dè bekendste wiskundigen aller tijden. Vanaf zijn jeugd stelde hij tabellen op met intervallen waarin hij priemgetallen telde.

Vermoeden 1.1.6 (Gauss)**.** *Neem X groot en Δ minder groot. Dan is het aantal priemgetallen in $[X, X+\Delta]$ ongeveer gelijk aan $\Delta/\log(X)$.*

Anders gezegd: dichtheid van de priemgetallen ter grootte X is ongeveer $1/\log(X)$.

Hier is een tabel met priemtellingen in intervallen $[x, x+10^5]$ voor diverse waarden van x,

x	aantal	$10^5/\log(x)$
10^8	5411	5428
10^9	4832	4825
10^{10}	4306	4342
10^{11}	4019	3948
10^{12}	3614	3619
10^{13}	3382	3340
10^{14}	3045	3102
10^{15}	2804	2895

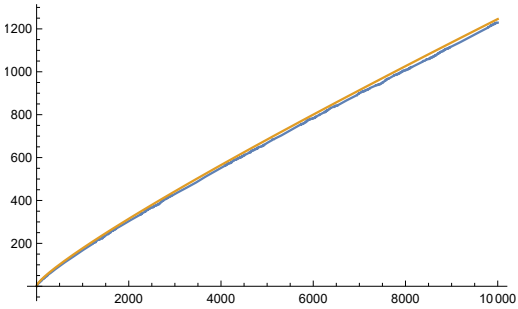
Tabel 1.1: Priemgetallen in intervallen.

Merk op hoe sterk de tweede en derde kolom in de tabel overeenkomen!

Dit soort tellingen leidden bij Gauss tot de volgende benaderingsfunctie,

$$\mathrm{li}(x) := \int_2^x \frac{dt}{\log t}$$

Hier is een plot van $\pi(x)$ en $\mathrm{li}(x)$ voor $x \leq 10000$.



Figuur 1.3: $\pi(x)$ en $\mathrm{li}(x)$ voor $x \leq 10000$.

De bovenste lijn is de grafiek van $\mathrm{li}(x)$. Maar dat geldt niet altijd, men kan aantonen dat $\mathrm{li}(x) - \pi(x)$ oneindig vaak van teken wisselt. Wanneer de eerste tekenwisseling plaatsvindt is niet bekend, wel dat het gebeurt bij een x met $x < 10^{372}$.

Rond 1896 bewezen Hadamard en De la Vallée-Poussin onafhankelijk van elkaar de volgende stelling.

Stelling 1.1.7 (Priemgetalstelling)**.**

$$\pi(x) \sim \mathrm{li}(x).$$

De notatie $f(x) \sim g(x)$ betekent dat de verhouding van $f(x)/g(x)$ naar 1 gaat als $x \to \infty$.

Voor het bewijs gebruikt men eigenschappen van de Riemann ζ-functie die later in deze cursus ter sprake komt.

Omdat $\mathrm{li}(x)$ niet een heel gebruikelijke functie in het onderwijs is, merken we op dat hij nog het meest lijkt op $x/\log x$.

Opgave 1.1.8. *Laat zien dat* $\mathrm{li}(x) \sim x/\log(x)$.

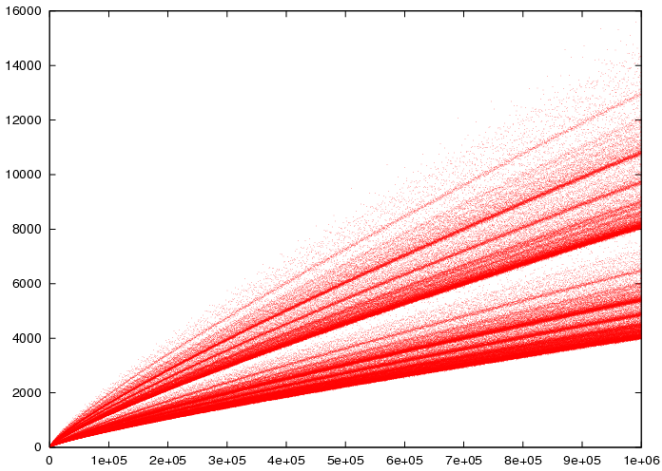
1.2 De grote priemgetalvermoedens

Het onderwerp priemgetallen bevat een groot aantal open vragen, waarvan een aantal zelfs bij een breder, niet-wiskundig publiek bekend zijn. We

noemen er hier een tweetal.

Vermoeden 1.2.1 (Goldbach, 1752)**.** *Elk even getal* ≥ 4 *is te schrijven als som van twee priemgetallen.*

Hier is een plot van het aantal manieren om even getallen N te schrijven als som van twee priemgetallen voor $N \leq 10^6$.



Figuur 1.4: Aantal manieren om N te schrijven als som van twee priemgetallen.

Dit probleem is nog steeds onopgelost. Het houdt zowel amateurs als professioneel wiskundigen nog altijd bezig, getuige de roman van A.Doxiadis: *Oom Petros en het Goldbach vermoeden.*

Als een probleem onoplosbaar lijkt, kan men altijd een probleem formuleren dat iets makkelijker lijkt.

Vermoeden 1.2.2 (Zwak Goldbach vermoeden)**.** *Elk oneven getal* $N \geq 7$ *is te schrijven als som van drie priemgetallen.*

Uit het Goldbach vermoeden volgt dat $N-3$ te schrijven is als $p+q$ met p, q priem. Dus $N = 3 + p + q$. Dit verklaart de naam *zwak* Goldbach vermoeden. Met behulp van de *cirkelmethode*, een gevanceerde analytische methode, kan men wel vat krijgen op het zwakke Goldbach vermoeden. I.M.Vinogradov bewees rond 1937 dat dit vermoeden klopt als N groot genoeg is. Heel recent (2013) rondde H.Helfgott het vraagstuk af door de 'kleine' N ook aan te pakken.

Het tweede vermoeden behelst paren priemgetallen die twee verschillen, zoals $(11, 13)$, $(41, 43)$, $(101, 103)$, etc. We noemen dit *priemtweelingen*.

Vermoeden 1.2.3 (Priemtweelingvermoeden)**.** *Er bestaan oneindig veel priemgetallen p zó dat $p + 2$ ook priem is.*

Ondanks de ontwikkeling van geavanceerde technieken was de vooruitgang uiterst moeizaam. Totdat in 2014 een spectaculaire vooruitgang in deze richting werd geboekt door Yitang Zhang. Hij toonde aan dat er een getal $A < 70.000.000$ bestaat zó dat er oneindig veel priemgetalparen $p, p + A$ bestaan. Hij gebruikte hiertoe een ingenieuze uitbreiding van de zogenaamde *zeefmethoden*, een andere geavanceerde techniek in de analytische getaltheorie. Het jaar daarop gaf James Maynard een aanzienlijke vereenvouding van Zhang's methode en een verbetering van $A < 246$. Het liefst zouden we natuurlijk $A = 2$ hebben, maar het ziet er naar uit dat daarvoor nieuwe ontwikkelingen nodig zijn.

Tenslotte een opgave naar aanleiding van de priemtweelingen.

Opgave 1.2.4. *Laat zien dat er oneindig veel priemgetallen p bestaan zó dat $p + 2$ niet priem is.*

Laat zien dat er maar één priemgetaldrieling $p, p + 2, p + 4$ bestaat.

Tenslotte vermelden we een tweetal diepe en opmerkelijke recente resultaten.

Een rij getallen $k_0 < k_1 < k_2 < \cdots < k_n$ heet een *rekenkundige rij* van lengte n als alle verschillen $k_{i+1} - k_i$ met $i = 0, \ldots, n - 1$ dezelfde waarde hebben.

Stelling 1.2.5 (Green, Tao 2008)**.** *Er bestaan willekeurig lange rekenkundige rijen bestaande uit priemgetallen.*

Voorbeelden: $150n + 7$ voor $n = 0, 1, \ldots, 6$ (Lemaire, 1910) en $26n + 4943$ voor $n = 0, 1, \ldots, 12$ (Seredinskij, 1963).

Een ander opmerkelijk resultaat,

Stelling 1.2.6 (Maynard 2016)**.** *Kies $a_0 \in \{0, 1, \ldots, 9\}$. Dan zijn er oneindig veel priemgetallen waarvoor a_0 niet in de decimale expansie voorkomt.*

1.3 Fermat en de priemgetallen

Pierre de Fermat (1607-1665) geldt algemeen als grondlegger van de moderne getaltheorie. Hoewel hier aan toegevoegd moet worden dat Fermat veel correspondeerde met tijdgenoten die ook in getaltheoretische proble-

men geinteresseerd waren.

Eén van Fermat's bekendste resultaten is de volgende.

Stelling 1.3.1 (Kleine stelling van Fermat)**.** *Zij p een priemgetal en a een getal niet deelbaar door p. Dan geldt dat*

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

Opmerking 1.3.2. *Als we de congruentie aan beide zijden met a vermenigvuldigen, krijgen we*

$$a^p \equiv a \mod p.$$

Als a deelbaar is door p geldt deze congruentie trivialerwijs. Dus is de congruentie waar voor alle gehele getallen a.

De kleine stelling van Fermat was indertijd een totaal nieuw type resultaat. Het geldt zeker niet als p geen priem is en geeft hiermee een interessante typering van de priemgetallen. Het bewijs van de stelling staat in voorbereidingstekst van deze cursus.

Een ander populair onderwerp in Fermat's tijd was dat van de typering van zogenaamde *Pythagoreïsche driehoeken.* Dat zijn rechthoekige driehoeken met gehele zijden. We geven hier een aantal,

$$5^2 = 3^2 + 4^2, \qquad 13^2 = 12^2 + 5^2 \qquad 17^2 = 15^2 + 8^2$$

$$29^2 = 21^2 + 20^2 \qquad 37^2 = 35^2 + 12^2 \qquad 41^2 = 40^2 + 9^2$$

Geordend naar de schuine zijde zijn dit de zes kleinste driehoeken als we ook aannemen dat de grootste gemene deler van de zijden 1 is. Maar er zijn ook grotere, zoals

$$904281937^2 = 281660465^2 + 859298088^2$$

De algemene formule voor deze driehoeken, aannemend dat de ggd van de zijden gelijk aan 1 is,

$$(a^2 + b^2)^2 = (a^2 - b^2)^2 + (2ab)^2 \quad \text{met } a, b \text{ geheel.}$$

De schuine zijdes zijn dus getallen van de vorm $a^2 + b^2$. Het is opvallend dat onze eerste zes voorbeelden allemaal een priemgetal als schuine zijde hebben. Het duurde dan ook niet lang voordat men het volgende vermoedde.

Vermoeden 1.3.3 (Stelling van Girard, 1630)**.** *Elk priemgetal van de vorm $4k + 1$ is te schrijven als som van twee kwadraten.*

Voorbeelden zijn: $13 = 2^2 + 3^2, 61 = 6^2 + 5^2, 113 = 8^2 + 7^2$, etc. Een echt bewijs werd pas geleverd door Euler rond 1740.

De priemgetallen van de vorm $4k - 1$ zijn daarentegen *niet* te schrijven als som van twee kwadraten. Voorbeelden: $3, 7, 11, \ldots$. Sterker nog, geen enkel getal van de vorm $4k - 1$ is som twee kwadraten.

De reden is dat een even kwadraat $(2m)^2 = 4m^2$ altijd deelbaar door 4 is en een oneven kwadraat $(2m + 1)^2 = 4m^2 + 4m + 1$ altijd een 4-voud plus 1. Hun som is dus ook 4-voud plus 1.

Er is zelfs nog een sterkere bewering.

Lemma 1.3.4. *Zij p een oneven priemgetal en a, b geheel, met* $\gcd(a, b) = 1$. *Stel dat $a^2 + b^2$ deelbaar door p. Dan geldt dat $p \equiv 1 \mod 4$.*

Voorbeeld $671^2 + 444^2 = 17 \times 113 \times 337$, met allemaal priemfactoren die 4-voud plus 1 zijn.

Bewijs. Omdat $\gcd(a, b) = 1$ en p deler van $a^2 + b^2$, volgt dat zowel a als b niet deelbaar door p is.

Omdat p een deler is van $a^2 + b^2$ geldt dat $a^2 \equiv -b^2 \mod p$. Verhef aan beide zijden tot de macht $(p - 1)/2$. We krijgen $a^{p-1} \equiv (-1)^{(p-1)/2} b^{p-1} \mod p$. Volgens Fermat's kleine stelling geldt dat $a^{p-1} \equiv b^{p-1} \equiv 1 \mod p$. Dus we houden over dat $1 \equiv (-1)^{(p-1)/2} \mod p$. Dit kan alleen maar als $(p - 1)/2$ even is en dus $p \equiv 1 \mod 4$. □

Dit lemma heeft een leuk gevolg.

Opgave 1.3.5. *Zij $N \geq 2$. Dan is er een priemgetal van de vorm $4k + 1$ dat groter is dan N.*
(Hint: bekijk $(N!)^2 + 1$)

Hier is een tweede gevolg van Fermat's kleine stelling.

Lemma 1.3.6. *Zij p een priemgetal van de vorm $4k + 1$. Dan is er een geheel getal A zó dat $A^2 + 1$ deelbaar is door p.*

Bewijs. Volgens Fermat geldt voor elke $a \in \{1, 2, \ldots, p - 1\}$ dat $a^{p-1} - 1$ deelbaar is door p. Omdat $a^{p-1} - 1 = (a^{(p-1)/2} - 1)(a^{(p-1)/2} + 1)$ moet één van de factoren $a^{(p-1)/2} - 1$ of $a^{(p-1)/2} + 1$ deelbaar zijn door p. Stel dat voor alle a geldt dat $a^{(p-1)/2} - 1$ deelbaar is door p. Maar dan zou de veelterm $x^{(p-1)/2} - 1$ modulo p meer nulpunten dan zijn graad hebben. Dat kan niet, Dus moet er een getal a zijn zó dat $a^{(p-1)/2} + 1$ deelbaar is door p. Kies nu $A = a^{(p-1)/4}$, dan is $A^2 + 1$ deelbaar door p. □

In dit bewijs hebben we gebruik gemaakt van het feit dat een polynoom modulo p niet meer nulpunten dan zijn graad kan hebben. Dat is niet

geheel vanzelfsprekend. Immers, de vergelijking $x^2 - 1 \equiv 0 \mod 24$ heeft de acht oplossingen $1, 5, 7, 11, 13, 17, 19, 23 \mod 24$. Maar dat komt natuurlijk doordat 24 geen priemgetal is. We formuleren daarom een aparte stelling.

Stelling 1.3.7. *Gegeven een polynoom $f(x)$ met gehele coefficienten en van graad d. Zij p een priemgetal. Dan zijn er hooguit d getallen $a \in \{1, 2, \ldots, p-1\}$ zó dat $f(a) \equiv 0 \mod p$.*

We zijn nu in staat een bewijs van de stelling van Girard te geven.

Bewijsschets van de stelling van Girard. Stel p priem met $p \equiv 1 \mod 4$. We hebben zojuist gezien dat er een geheel getal A is, zó dat $A^2 \equiv -1 \mod p$. Bekijk nu de verzameling

$$\Lambda = \{(a, b) \in \mathbb{Z}^2 | a \equiv Ab \mod p\}.$$

Dit is een verzameling punten met gehele coördinaten in het platte vlak met de eigenschappen

1. $(a, b) \in \Lambda \Rightarrow (ka, kb) \in \Lambda$ voor elke gehele k
2. $(a, b), (a', b') \in \Lambda \Rightarrow (a + a', b + b') \in \Lambda$.

We noemen een dergelijke verzameling een *rooster* in $\mathbb{Z}^2$. Kies nu $(a, b) \in \Lambda$. Dan geldt dat $a^2 + b^2 \equiv A^2 b^2 + b^2 \equiv 0 \mod p$. Dus $a^2 + b^2$ is een veelvoud van p. Kies $(r, s) \in \Lambda$ met $(r, s) \neq (0, 0)$ zó dat $r^2 + s^2$ minimaal is. Het blijkt dat $r^2 + s^2$ niet alleen een veelvoud van p is, maar zelfs gelijk er aan.

Hiermee is de stelling van Girard bewezen. □

Fermat was ook geinteresseerd in de zogenaamde *perfecte getallen*. Dit zijn gehele getallen N waarvan de som van de delers $< N$ weer gelijk aan N is. Voorbeelden zijn $N = 6, 28, 496, 8128, \ldots$. Euclides observeerde al dat als $2^n - 1$ een priemgetal is, dan is de som van de delers van $2^{n-1}(2^n - 1)$ gelijk aan

$$1 + 2 + \cdots + 2^{n-1} + (2^n - 1) + 2 \times (2^n - 1) + 2^{n-2} \times (2^n - 1) =$$
$$2^n - 1 + (2^{n-1} - 1)(2^n - 1) = 2^{n-1}(2^n - 1).$$

Met andere woorden, $2^{n-1}(2^n - 1)$ is perfect. Later bewees Euler dat een even perfect getal altijd van de vorm is die Euclides beschreef. Dat brengt ons op de vraag voor welke n het getal $2^n - 1$ priem is. Getallen van de vorm $2^n - 1$ noemt men *Mersenne getallen*, naar een tijdgenoot van Fermat, en als $2^n - 1$ priem is spreken van een *Mersenne priemgetal*.

Opgave 1.3.8. *Bewijs:* $2^n - 1$ *is priem* $\Rightarrow$ n *is priem.*

(Hint: m *deelt* $n \Rightarrow 2^m - 1$ *deelt* $2^n - 1$*)*

NB: n priem hoeft niet te impliceren dat $2^n - 1$ priem is. Bijvoorbeeld $2^{11} - 1$ is niet priem.

Pari opgave 1.3.9. *De volgende code print de exponenten* n *van Mersenne priemgetallen voor* $n \leq 500$*.*

```
for(n=2,500,if(isprime(2^n-1),print(n),))
```

Probeer 500 te verhogen en zie wanneer het mis gaat.

Opgave 1.3.10. *Gegeven gehele getallen* $a, n \geq 2$*. Stel dat* $a^n + 1$ *priem is. Bewijs dat* a *even is en* $n = 2^k$*.*

(Hint: Stel dat m *een deler van* n *is en* n/m *oneven. Dan is* $a^m + 1$ *een deler van* $a^n + 1$*)*

Getallen van de vorm $2^{2^k} + 1$ heten *Fermatgetallen* en *Fermat priemgetal* als het priem is.

De eerste vijf: $5, 17, 257, 65537, 4294967297$, waarvan de eerste vier priem. Fermat meende dat alle Fermatgetallen priem zijn. Tegenwoordig vermoedt men dat alleen de eerste vier priem zijn.

Pari opgave 1.3.11. *Getallen van de vorm* $a^{2^k} + 1$*, met* a *even, noemen we Fermat-achtige priemgetallen. Bepaal met het commando* `isprime(n)` *in PARI en een eventuele for-loop zelf een aantal Fermat-achtige priemgetallen. Pas op, met* $k > 10$ *loopt* `isprime` *al langzamer. Dit omdat de exponent* 2^k *dan wel erg groot wordt. Je zult er achter komen dat ze best wel zeldzaam zijn.*

Tenslotte, ter vermaak en recreativiteit.

Pari opgave 1.3.12. *Een getal zoals* 1251521 *of* 113311 *noemt men een palindroom getal. Bepaal met PARI een aantal grote palindroom priemgetallen. De rep-unit met 19 of 23 enen is een voorbeeld daarvan. Of* 32000000023*. Palindroompriemgetallen hebben altijd een oneven aantal cijfers, behalve* 11*. Om je op gang te helpen een loop:*

```
for(a=1,9,for(b=1,9,if(isprime((a+10*b)*10^19+(b+10*a)),
print(a+10*b))))
```

De exponent 19 kun je door ieder ander oneven getal vervangen. Maar niets houdt je tegen om ook andere mogelijkheden te proberen.

1.4 Middelbare school methoden

Stelling 1.4.1 (Chebyshev, 1850)**.** *Er bestaan $\alpha, \beta > 0$ zó dat voor voldoend grote x geldt*

$$\alpha \frac{x}{\log x} < \pi(x) < \beta \frac{x}{\log x}.$$

Chebyshev liet zien dat je $\alpha = 0.82, \beta = 1.11$ kunt nemen.

Wij zullen laten zien dat we $\alpha = 0.66$ kunnen nemen. Hiertoe merken we eerst een verband tussen $\pi(n)$ en $\mathrm{kgv}(1, 2, \ldots, n)$ op.

Lemma 1.4.2. *Voor elk positief geheel getal n geldt* $\mathrm{kgv}(1, 2, 3, \ldots, n) \leq n^{\pi(n)}$.

Bewijs. Het $\mathrm{kgv}(1, 2, 3, \ldots, n)$ is gelijk aan het product van alle priemmachten p^k met $p^k \leq n$, waarbij k maximaal gekozen is. Dus bijvoorbeeld:

$$\mathrm{kgv}(2, \ldots, 10) = 2^3 \times 3^2 \times 5 \times 7.$$

We hebben dus een product van $\pi(n)$ factoren met maximale waarde n. En dat is kleiner dan $n^{\pi(n)}$. □

Bewijs van Chebyshev's ondergrens. Bekijk de integraal

$$I_n = \int_0^1 x^n (1-x)^n dx.$$

- Omdat $x(1-x) \leq 1/4$ voor $0 \leq x \leq 1$ geldt $I_n < 1/4^n$.
- I_n is de integraal is van een polynoom $a_n x^n + a_{n+1} x^{n+1} + \cdots + a_{2n} x^{2n}$ met gehele coefficienten. Dus

$$I_n = \frac{a_n}{n+1} + \frac{a_{n+1}}{n+2} + \cdots + \frac{a_{2n}}{2n+1} \in \frac{\mathbb{Z}}{\mathrm{kgv}(1, 2, \ldots, 2n+1)}$$

- Omdat $I_n > 0$ moet gelden dat

$$I_n \geq \frac{1}{\mathrm{kgv}(1, 2, \ldots, 2n+1)}.$$

Combinatie van de onder- en bovengrens voor I_n geeft: $\mathrm{kgv}(1, 2, \ldots, 2n+1) > 4^n$. Gebruik vervolgens ons lemma en we vinden dat $(2n+1)^{\pi(2n+1)} > 4^n$.

Neem de logaritme en deel door $\log(2n+1)$,

$$\pi(2n+1) > \frac{n \log 4}{\log(2n+1)} > 0.66 \frac{2n+1}{\log(2n+1)} \quad \text{als } n > 10.$$

□

Chebyshev gebruikte dit soort methoden ook in zijn bewijs van het volgende.

Stelling 1.4.3 (Postulaat van Bertrand)**.** *Voor elk geheel getal $n \geq 2$ bestaat er een priemgetal p met $n \leq p < 2n$.*

Uit de priemgetalstelling volgt een sterkere uitspraak.

Stelling 1.4.4. *Kies $\epsilon > 0$. Dan geldt voor voldoend grote x dat het interval $[x, (1+\epsilon)x]$ een priemgetal bevat.*

Bewijs. We laten zien dat $\pi((1+\epsilon)x)/\pi(x) \to 1+\epsilon$ als $x \to \infty$.

$$\frac{\pi((1+\epsilon)x)}{\pi(x)} \sim \frac{(1+\epsilon)x}{\log((1+\epsilon)x)} \Big/ \frac{x}{\log x} \sim 1+\epsilon.$$

□

Deze resultaten geven aanleiding tot de volgende vraag.

Vraag 1.4.5. *Geef de rij opeenvolgende priemgetallen aan met $p_1 = 2, p_2 = 3, p_3, \ldots, p_n, \ldots$. Hoe groot en hoe klein kunnen de verschillen $g_n = p_{n+1} - p_n$ worden?*

Hier zijn een paar resultaten.

- Uit het postulaat van Bertrand volgt dat $g_n < p_n$ voor alle n.
- Uit de priemgetalstelling volgt dat voor elke $\epsilon > 0$ geldt dat $g_n < \epsilon p_n$ als n groot genoeg is.
- Ingham (1937) Elk interval van de vorm $[n^3, (n+1)^3]$ en n groot genoeg bevat een priemgetal. Dit komt neer op $g_n < 3p_n^{2/3}$.
- Het is niet bekend of elk interval $[n^2, (n+1)^2]$ een priemgetal bevat.
- Het priemtweeling vermoeden komt neer op de vraag of $g_n = 2$ voor oneindig veel n.

In de volgende opgave kun je een indruk krijgen van de gaten tussen opeenvolgende priemgetallen.

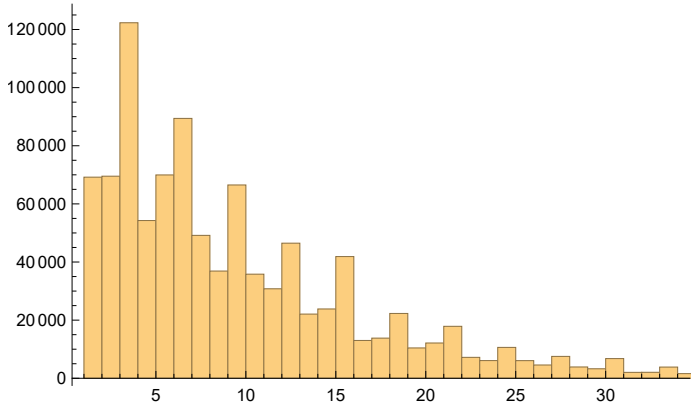
Pari opgave 1.4.6. *De PARI opdracht* `nextprime(n)` *geeft het kleinste priemgetal $\geq n$. Voor het kleinste priemgetal met meer dan 100 cijfers:* `nextprime(10^100)`*. Probeer ook eens wat andere waarden van n. Daarmee krijg je een idee hoe lang het duurt voor je een priemgetal tegenkomt. Vergelijk dit met* `log(n)`*.*

Typ nu de volgende functiedefinitie in:

```
nextprimes(n,k)=q=n;for(m=1,k,q=nextprime(q+1);print(q))
```

De opdracht `nextprimes(10^100,15)` *geeft dan de eerste 15 priemgetallen groter dan 10^{100}. Speel een beetje met deze functie om te een idee te krijgen van de verschillen tussen heel grote opeenvolgende priemgetallen.*

Hier is een histogram van de halve verschillen $g_n/2$ met $n = 10\times10^6, \ldots, 11\times 10^6$.



Figuur 1.5: Histogram van $g_n/2$ voor n van 10 miljoen tot 11 miljoen.

NB: $\log(p_{10^7}) \approx 19$
Hier is een rij van $m-1$ opeenvolgende getallen die niet priem zijn:

$$m!+2, m!+3, \ldots, m!+m.$$

Er geldt dat $m > \dfrac{\log(m!+1)}{\log\log(m!+1)}$. Stel p_n is het grootste priemgetal $\leq m!+1$. Dan geldt $g_n \geq m > \dfrac{\log(p_n)}{\log\log(p_n)}$.
Tegenwoordig weten we dat $g_n > \log(p_n)$ voor oneindig veel n.

Vermoeden 1.4.7 (Cramér, 1936)**.** *Er bestaat een constante C zó dat $g_n < C(\log p_n)^2$ voor alle n*

Tenslotte is hier een programma om priemtweelingen te vinden.

Pari opgave 1.4.8. *Maak zelf de opdracht* `nexttwin(n)` *waarmee je de eerstvolgende priem-tweeling $p, p+2$ met $p \geq n$ kunt vinden. De waarde van p wordt berekend.*

```
nexttwin(n)=q=nextprime(n);while(!isprime(q+2),
q=nextprime(q+1));q
```

Wat is de kleinste priemtweeling boven 10^{100}? Bepaal nog een aantal tweelingen $\geq n$ voor zelfgekozen n. Vergelijk de antwoorden met $(\log n)^2$. We kunnen ook lijstjes opeenvolgende priemtweelingen laten uitprinten. Dit

gaat met de functie `nexttwins(n,k)` *die de waarde van p voor de eerstvolgende k priemtweelingen na n uitprint.*

```
nexttwins(n,k)=q=nexttwin(n);for(m=1,k,print(q);
q=nexttwin(q+1))
```

Probeer eens `nexttwins(10^100,15)` *en bekijk de achtereenvolgende verschillen. Neem ook wat andere waarden van n.*

1.5 Priemtesten

We beginnen met de vraag: is

$$N = 412068292467224099305295867294454137284725652929 61$$

priem? Om dit vast te stellen kunnen we van alle (priem) getallen groter dan 1 en $< \sqrt{N} \approx 6 \cdot 10^{24}$ testen of het een deler van N is. We noemen dit de *naïeve methode*.
Stel dat 1 deling een nanoseconde duurt (10^{-9} seconde). Dan is de geschatte looptijd:

$$6 \times 10^{24} \times 10^{-9} \text{ sec} = 6 \times 10^{15} \text{ sec} \approx 200.000.000 \text{ jaar}.$$

Voor algemene N is de geschatte looptijd van dit naïeve algorithme: $C\sqrt{N} = C\exp(\frac{1}{2}\log N)$. We noemen dit een *exponentieel* algoritme. Het idee is dat er ongeveer $\log N$ toetsaanslagen op de computer nodig zijn om N in te voeren, terwijl het algoritme exponentieel is in dit aantal aanslagen. Een algoritme met verwachte looptijd $C(\log N)^k$ voor vaste k, C noemen we een *polynomiaal algoritme*. We noemen het zelfs *lineair* als $k = 1$.

Vraag 1.5.1. *Zijn er polynomiale algoritmen om een getal op primaliteit te testen of in priemfactoren te ontbinden?*

Hier is een eerste poging die uitgaat van de kleine stelling van Fermat. We weten, gegeven een getal n en a niet deelbaar door n:

$$n \text{ is priem } \Rightarrow a^{n-1} \equiv 1 \mod n.$$

Neem contrapositieve,

$$n \text{ is niet priem } \Leftarrow a^{n-1} \not\equiv 1 \mod n.$$

De laatste uitspraak kan dus dienen als test of het getal n samengesteld (niet priem) is. We noemen dit de *Fermat test*.

Pari opgave 1.5.2. *PARI Opgave Neem* $n = 2^{127} - 1$ *en bereken* a^{n-1} mod n *voor een paar waarden van* a*. Dit kunnen we niet klakkeloos doen want de exponent* $n - 1$ *is veel te groot voor directe berekening van* a^{n-1}*. We moeten expliciet met de restklassen* a *modulo* n *rekenen. Dat gaat als volgt.*

Tik in: `n=2^127-1` *en vervolgens*

- `Mod(3,n)^k` *voor een paar exponenten* k*.*
- `Mod(a,n)^(n-1)` *voor een paar waarden van* a*.*

Neem $n = 2^{67} - 1$ *en bepaal* $a^{n-1} \mod n$ *voor een paar waarden van* a*. Wat is je conclusie?*

In de vorige opgaven hebben we impliciet gebruikt dat machtsverheffing van restklassen heel snel uitgevoerd kan worden. Dat gaat als volgt.

Als voorbeeld berekenen we $a^{1111} \mod n$ voor de een of andere a, n.

1. Schrijf de exponent $1111 = 2^{10} + 2^6 + 2^4 + 2^2 + 2 + 1$ (binaire schrijfwijze)
2. Bepaal achtereenvolgens $a^2 \mod n, a^{2^2} \mod n, \ldots, a^{2^{10}} \mod n$ (herhaald kwadrateren).
3. Bepaal vervolgens het product $a^1 \times a^2 \times a^{2^2} \times a^{2^4} \times a^{2^6} \times a^{2^{10}} \equiv a^{1111} \mod n$.

In het algemeen willen we $a^k \mod n$ bepalen. Het aantal kwadrateringen is $< \log k / \log 2$. Het aantal vermenigvuldigen is $< \log k / \log 2$. De looptijd is dus $C \log k$ voor zekere constante C.

De Fermat test is een *samengesteldheidstest* voor n. Een a, niet deelbaar door n, zó dat $a^{n-1} \not\equiv 1 \mod n$ noemen we een *getuige* van de samengesteldheid van n.

Helaas zijn er samengestelde getallen n die vrijwel geen getuige hebben.

Definitie 1.5.3. *Een samengesteld getal* n *heet Carmichael getal als voor elke* a *met* $\gcd(a, n) = 1$ *geldt* $a^{n-1} \equiv 1 \mod n$*.*

Voorbeelden: $561, 1729, 294409, \ldots$

Stelling 1.5.4 (Alford, Granville, Pomerance, 1992)**.** *Er bestaan oneindig veel Carmichaelgetallen.*

We kunnen de stelling van Fermat iets verfijnen. Stel, a, n geheel en n oneven en geen deler van a. Dan geldt:

$$n \text{ priem} \Rightarrow a^{\frac{n-1}{2}} \equiv \pm 1 \mod n.$$

Dit volgt direct uit het feit dat als n priem is en dus $a^{n-1} - 1 = (a^{(n-1)/2} - 1)(a^{(n-1)/2} + 1)$ deelt, hij één van de twee factoren deelt. De contrapositieve

uitspraak luidt

$$n \text{ samengesteld} \Leftarrow a^{\frac{n-1}{2}} \not\equiv \pm 1 \mod n.$$

Deze samengesteldheidstest noemen we gemakshalve de *Euler-test*. Als n samengesteld is dan kan men aantonen dat minstens 50% van de restklassen $a \mod n$ daar onder de Euler-test getuige van is.

Als voor veel verschillende a een 'positief' uitblijft, dan stijgt de kans dat n priem is. We spreken daarom van een *pseudo priemtest*. Een bekende, en meest gebruikte, verfijning van de Euler-test is de zogenaamde *Rabin-test*. Als het zogenaamde Gegeneraliseerde Riemann-vermoeden waar is, dan is er een getuige a voor de Rabin-test met $0 < a < 2(\log n)^2$ (Miller, 1976). Dat zou een polynomiale priemtest geven. Helaas is het Riemann vermoeden nog verre van bewezen.

Vanaf 1985 zijn er diverse priemtesten (Cohen-Lenstra, Atkin-Morain) ontwikkeld waarmee men getallen tot enkele honderden cijfers deterministisch kon testen. Echter, geen van deze tests is aantoonbaar polynomiaal.

In 2002 bewezen Agrawal, een informaticus uit India en twee van zijn bachelorstudenten de volgende zeer verrassende stelling.

Stelling 1.5.5 (Agrawal, Kayal, Saxena, 2002)**.** *Er bestaat een polynomiale priemtest.*

In contrast met eerdere (niet-polynomiale) priemtesten ziet deze er verrassend eenvoudig. Zonder het algoritme precies uit te leggen reproduceren we het hier.

Input: geheel getal $n > 1$

1. Als $n = a^b$ voor $a \in \mathbb{N}$ en $b > 1$:
   output *Samengesteld*
2. Kies de kleinste $r \in \mathbb{N}$ zó dat $\text{ord}_r(n) > 4(\log n)^2$.
3. Als $1 < \text{ggd}(a, n) < n$ voor zekere $a \leq r$:
   output *Samengesteld*
4. Als $n \leq r$, output *Priem*
5. For $a = 1$ to $[2\sqrt{r}\log n]$ do
   If $(X + a)^n \neq X^n + a \mod X^r - 1, n$:
   output *Samengesteld*
6. output *Priem*

Voor getallen van speciale vorm zijn er aangepaste tests die zeer snel kunnen werken. Bijvoorbeeld de Mersenne priemtest.

Stelling 1.5.6 (Lucas, 1876)**.** *Stel $M_n = 2^n - 1$. Construeer de rij $S_1, S_2, S_3, \ldots$ modulo M_n door $S_1 = 4$ en $S_k = S_{k-1}^2 - 2$ voor $k \geq 2$.*

Dan,

$$M_n \text{ is priem} \Leftrightarrow S_{n-1} \equiv 0 \mod M_n.$$

Hier is de lijst tot nu toe bekende waarden n waarvoor $2^n - 1$ priem is.

$$\begin{aligned} n \;=\; & 2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, 1279, 2203, \\ & 2281, 3217, 4253, 4423, 9689, 9941, 11213, 19937, 21701, \\ & 23209, 44497, 86243, 110503, 132049, 216091, 756839, 859433, \\ & 1257787, 1398269, 2976221, 3021377, 6972593, 13466917, \\ & 20996011, 24036583, 25964951, 30402457, 32582657, \\ & 37156667, 42643801, 43112609, 57885161, 74207281, \\ & 77232917, 82589933 (2018) \end{aligned}$$

De exponent 127 is het laatste met de hand uitgewerkte geval (1876), vanaf de 1940-er jaren nam de computer het over. De laatste 17 exponenten in de lijst zijn ontdekt door het internet project GIMPS, waar een ieder, die nog CPU-tijd heeft, aan kan deelnemen. Voor de stand van zaken zie de pagina `mersenne.org/primes` in de uitgebreide GIMPS-site. Gemiddeld wordt nu om de paar jaar een nieuwe exponent aan de lijst toegevoegd.

Nauw verwant aan de Mersennegetallen zijn de zogenaamde *rep-units*

$$\frac{10^n - 1}{9} = 111 \ldots 11 \quad n \text{ enen.}$$

Het is bekend dat deze priem zijn voor $n = 2, 19, 23, 317, 1031, 49081, 86453, \ldots$

Pari opgave 1.5.7. *Kies een geheel getal $g \geq 2$. Getallen van de vorm $\frac{g^n-1}{g-1}$ noemen we rep-units in basis g. We gaan op zoek naar rep-units die priem zijn. Hiervoor kunnen we de PARI-opdracht* `isprime(n)` *gebruiken. Om type-werk te besparen definieren we eerst een zelfgemaakte functie.*

```
rep(g)=for(n=2,50,if(isprime((g^n-1)/(g-1)),print(n)))
```

Geef vervolgens de opdracht `rep(10)`*. Met de opdracht* `rep(g)` *wordt $\frac{g^n-1}{g-1}$ op primaliteit getest voor $n = 2, 3, ..., 50$. Je kunt 50 natuurlijk ook in een ander getal veranderen. Of de lus over priemwaarden van n te laten lopen door* `n` *door* `prime(n)` *te vervangen.*

1.6 Ontbinding in priemfactoren

Met de Fermat- en Eulertest is vaak eenvoudig vast te stellen of een samengesteld getal inderdaad samengesteld is. Veel lastiger is het om dat getal

vervolgens te ontbinden. Eerst geven we een *niet-serieus* maar hopelijk wel vermakelijk voorbeeld van een ontbinding.

- Beschouw het getal
$$14989086475$$
- Splits de cijfers in twee groepjes
$$14989086475$$
- Neem som en verschil van de twee delen:
$$149890 + 86475 = 236365$$
$$149890 - 86475 = 63415$$
- Vermenigvuldig ze
$$236365 \times 63415 = 14989086475.$$

Nu wat serieuzer, tegenwoordig zijn er diverse niet-naïeve methoden om een groot getal te ontbinden. Eén daarvan is de zogenaamde *kwadratische zeef*. Deze berust op het volgende idee. Stel we willen N in priemfactoren ontbinden. We gaan op zoek naar gehele getallen x, y zó dat $x^2 - y^2 = (x-y)(x+y)$ deelbaar is door N. Vervolgens berekenen we $\text{ggd}(N, x-y)$ en $\text{ggd}(N, x+y)$ en met enig geluk is één van deze getallen een echte deler van N.

De ggd van twee getallen kan in lineaire tijd worden berekend met het Euclidisch algoritme. Bijvoorbeeld $\text{ggd}(2913, 1533)$.

- $2913 - 1533 = 1380$
- $1533 - 1380 = 153$
- $1380 - 9 \times 153 = 3$
- $153 - 51 \times 3 = 0$

De laatste rest $\neq 0$ is de gewenste ggd. Dus $\text{ggd}(2913, 1533) = 3$.

We schetsen nu hoe we x, y berekenen zó dat $x^2 - y^2$ deelbaar is door N. Stel $r = \lfloor \sqrt{N} \rfloor + 1$ en kies een grens B. Bereken $(k+r)^2 - N$ voor kleine k en selecteer de getallen met alle priemdelers $\leq B$. Merk op dat de getallen $(k+r)^2 - N$ van de orde van grootte van $k\sqrt{N}$ zijn. Dat is een stuk kleiner dan N zelf en dus is de "kans" dat ze kleine priemdelers (onder B) hebben des te groter.

Voorbeeld: $N = 123889$. Dan $r = \lfloor \sqrt{N} \rfloor + 1 = 352$. Kies $B = 13$ en probeer $-20 \leq k \leq 20$. We vinden de volgende waarden van k waarvan $(r+k)^2 - N$ uit priemfactoren ≤ 13 bestaat,

k	$(r+k)^2 - N$
-19	$-2^3 \cdot 5^3 \cdot 13$
-17	$-2^4 \cdot 3^6$
-9	$-2^5 \cdot 3 \cdot 5 \cdot 13$
0	$3 \cdot 5$
1	$2^4 \cdot 3^2 \cdot 5$
7	$2^7 \cdot 3 \cdot 13$
15	$2^4 \cdot 3^3 \cdot 13$

Kies nu de verschillen met $k = -19, -9, 15$,

$$\begin{aligned} (r-19)^2 &\equiv -2^3 \cdot 5^3 \cdot 13 \mod N \\ (r-9)^2 &\equiv -2^5 \cdot 3 \cdot 5 \cdot 13 \mod N \\ (r+15)^2 &\equiv 2^4 \cdot 3^3 \cdot 5^2 \mod N \end{aligned}$$

Vermenigvuldiging van deze congruenties geeft

$$(r-19)^2(r-9)^2(r+15)^2 \equiv (2^6 \cdot 3^2 \cdot 5^3 \cdot 13)^2 \mod N$$

We hebben twee verschillende kwadraten, gelijk modulo N. Kijk of dit een factor van N geeft,

$$\begin{aligned} & \gcd(N, (r-19)(r-9)(r+15) - 2^6 \cdot 3^2 \cdot 5^3 \cdot 13) \\ = \; & \gcd(123889, 40982373) = 541 \end{aligned}$$

We hebben geluk en vinden $N = 541 \cdot 229$.

Voor degenen die bekend zijn met lineaire algebra, de stap die de geschikte rijen uit de tabel kiest is niets anders dan lineaire algebra modulo 2. De exponent vectoren van onze ontbindingen zijn modulo 2,

k	-1	2	3	5	7	11	13
-19	1	1	0	1	0	0	1
-17	1	0	0	0	0	0	0
-9	1	1	1	1	0	0	1
0	0	0	1	1	0	0	0
1	0	0	0	1	0	0	0
7	0	1	1	0	0	0	1
15	0	0	1	0	0	0	0

De som van de rijen bij $k = -19, -9, 15$ is nul modulo 2, en dit verklaart onze keuze. In werkelijkheid zijn deze matrices veel groter en moeten we daadwerkelijk lineaire vergelijkingen modulo 2 oplossen.

Uit lineaire algebra volgt dat als het aantal rijen groter is dan het aantal kolommen dan is er zeker een oplossing.

Als we allerlei getaltheoretische vermoedens aannemen dan kan men berekenen dat de geschatte looptijd ongeveer gelijk is aan

$$L(N) = \exp\left(2\sqrt{\log(N)\log\log(N)}\right).$$

Je kunt zelf controleren dat

$$\lim_{N\to\infty} L(N)/N^{\epsilon} = 0$$

voor elke $\epsilon > 0$. We noemen de kwadratische zeef daarom een *subexponentieel algoritme*.

Dat de geschatte looptijd $L(N)$ gebaseerd is op allerlei onbewezen vermoedens is niet erg. Als je eenmaal een deler van N hebt gevonden dan ben je tevreden en klaag je niet over hoe het gevonden is.

In de praktijk is het nog niet mogelijk om willekeurige getallen van ongeveer 600 cijfers te ontbinden binnen een mensenleven. Ook is er geen polynomiaal ontbindingsalgoritme bekend.

1.7 RSA-cryptografie

Kies twee priemgetallen p, q van circa 300 cijfers en bereken $N = pq$.

Lemma 1.7.1. *Stel $M = k(p-1)(q-1)+1$ met k willekeurig geheel. Dan geldt $a^M \equiv a \mod N$ voor alle gehele getallen a.*

Bewijs. We berekenen eerst $a^M \mod p$.

- Stel dat p niet a deelt. Pas Fermat toe: $a^M \equiv a^{k(p-1)(q-1)+1} \equiv 1 \cdot a \equiv a \mod p$.
- Als a deelbaar is door p dan geldt $a^M \equiv 0 \equiv a \mod p$.
- Conclusie $a^M \equiv a \mod p$ voor alle a. Op dezelfde manier $a^M \equiv a \mod q$.
- Dus $a^M - a$ wordt gedeeld door zowel p als q en daarmee door $pq = N$.

□

We beschrijven het RSA-protocol, genoemd naar Rivest-Shamir-Adleman (1977), om berichten te versleutelen en ontsleutelen.

De ingrediënten zijn de eerder genoemde p, q van 300 cijfers en hun product N.

Kies een willekeurig getal e, je publieke sleutel, en bepaal het getal f, je geheime sleutel, zó dat $ef = k(p-1)(q-1)+1$ voor zekere k.

Stel, iemand wil jou een bericht sturen, gepresenteerd door een getal B. Uit voorgaand lemma weten we dat $B^{ef} \equiv B \mod N$.

In plaats van B direct naar je te sturen, stuurt verzender het versleutelde bericht $C = B^e \mod N$. Bij ontvangst bereken je $C^f \equiv (B^e)^f \equiv B^{ef} \equiv B \mod N$, waarmee je het bericht hersteld hebt en gewoon kunt lezen.

Als een malafide persoon je geheime sleutel f uit e wil afleiden heb je p en q nodig en dus de ontbinding van N. Zoals gezegd is dit doorgaans een onmogelijke zaak.

Je kunt het RSA-systeem in PARI simuleren.

Pari opgave 1.7.2. *Voer de volgende opdrachten uit:*

- `p=randomprime(10^50)` *en* `q=randomprime(10^50)`
- `n=p*q`
- `e=137` *(publieke sleutel, kun je zelf kiezen)*
- `f=lift(1/Mod(e,(p-1)*(q-1)))` *(geheime sleutel. Dit kan misgaan, kies dan een andere e)*
- `b=123456789` *(Kies zelf een bericht)*
- `c=Mod(b,n)^e` *(versleutel het)*
- `c^f` *(ontsleuteling)*
- `lift(c^f)` *(ontsleuteling zonder dat n geprint wordt)*

Dit kan levensecht worden gemaakt met twee personen. Eén persoon maakt n en de sleutels e, f aan en stuurt n en e naar persoon twee. Deze verzint een bericht B en stuurt $B^e \mod n$ naar persoon één. Deze ontsleutelt het bericht.

1.8 Priemgetalformules

Euler observeerde dat het polynoom $x^2 - x + 41$ een priemwaarde heeft voor $x = 0, 1, 2, \ldots, 40$, hetgeen een uitzonderlijk lange serie is. Duidelijk is dat de waarde bij $x = 41$ deelbaar is door 41. Het blijkt dat geen enkel polynoom alleen priemwaarden heeft.

Stelling 1.8.1. *Gegeven een niet-constant polynoom $F(x)$. Dan zijn oneindig veel van de waarden $F(0), F(1), F(2), \ldots$ niet priem.*

Bewijs. Neem bijvoorbeeld $F(x) = x^2 + 1$. Merk op $F(2) = 5$. Dan geldt $F(2 + 5k) \equiv F(2) \equiv 0 \mod 5$. Dus alle waarden $F(2 + 5k)$ zijn deelbaar door 5. □

Hoe zit het met de priemwaarden van $F(x)$?

Vermoeden 1.8.2 (Bunyakowsky, 1857)**.** *Stel dat $F(x)$ een irreducibel polynoom is en stel dat* $\text{ggd}(F(1), F(2), \ldots) = 1$*. Dan heeft $F(x)$ oneindig veel priemwaarden.*

NB:

- Het is duidelijk dat bijvoorbeeld $(x+1)(x+2)$ hooguit eindig veel priemwaarden kan aannemen. Vandaar de conditie $F(x)$ irreducibel.
- Het polynoom $3x^2 - x + 2$ is irreducibel, maar neemt alleen even waarden aan. Vandaar de ggd-conditie.
- In de praktijk blijken vaak de eerste paar waarden $F(1), F(2), \ldots$ al ggd $= 1$ te hebben. Deze conditie is dus niet zo zwaar als hij lijkt.

Volgens het Buyakowsky-vermoeden zou n^2+1 voor oneindig veel waarden van n priem moeten zijn. Stel

$$\pi_F(x) = \#\{n^2 + 1 \leq x \text{ en } n^2 + 1 \text{ priem}\}.$$

Uit de volgende tabel zou men kunnen vermoeden dat $\pi_F(x) \sim 1.55\frac{\sqrt{x}}{\log x}$,

x	$\pi_F(x)$	$1.55\sqrt{x}/\log x$
10^3	10	7.09
10^4	19	16.82
10^5	51	45.57
10^6	112	112.19
10^7	316	304.10
10^8	841	841.44
10^9	2378	2365.23
10^{10}	6656	6731.56

Tabel 1.2: $\pi_F(x)$.

Het vermoeden van Bunyakowksy is bewezen voor lineaire polynomen.

Stelling 1.8.3 (Dirichlet, 1837)**.** *Zij a, q een tweetal positieve getallen met* $\text{ggd}(a, q) = 1$*. De rij getallen $qn + a$ met $n = 0, 1, 2, 3, \ldots$ oneindig veel priemwaarden.*

Anders gezegd, er zijn oneindig veel priemgetallen van de vorm $p \equiv a \mod q$.

Het blijkt zelfs dat de priemgetallen min of meer gelijkmatig over de restklassen $a \mod q$ verdeeld zijn.

We kijken nu naar een 'echte' priemproducerende formule.

Stelling 1.8.4 (Mills, 1947)**.** *Er bestaat een reëel getal A zó dat $\lfloor A^{3^n} \rfloor$ priem voor alle gehele $n \geq 0$.*

Opmerking 1.8.5. *Hoewel vermakelijk is deze stelling volkomen nutteloos. Om aan de waarde van A te komen, moet je daar eerst allerlei informatie over de priemgetallen in stoppen. Dit blijkt uit onderstaand bewijs. Daar zal ook blijken dat er oneindig veel mogelijkheden voor A zijn.*

Bewijs. We gebruiken dat elk interval $[N^3, (N+1)^3 - 1]$ met $N \geq N_0$ een priemgetal bevat (Ingham, 1937).
Kies een priem getal $P_0 \geq N_0$ en vervolgens $P_1, P_2, P_3, \ldots$ zó dat

$$P_n^3 < P_{n+1} < (P_n + 1)^3 - 1 \quad n = 0, 1, 2, \ldots$$

Bewering: het gewenste getal A is gelijk aan $\lim_{n\to\infty} P_n^{1/3^n}$ en er geldt $\lfloor A^{3^n} \rfloor = P_n$ voor alle $n \geq 0$.
Om de bewering te bevestigen definieren we $u_n = P_n^{1/3^n}, v_n = (P_n+1)^{1/3^n}$ voor elke $n \geq 0$. Merk op dat $v_n > u_n$ voor alle n.
De rij $u_0, u_1, u_2, \ldots$ is een stijgende rij,

$$\frac{u_{n+1}}{u_n} = \frac{P_{n+1}^{1/3^{n+1}}}{P_n^{1/3^n}} = \left(\frac{P_{n+1}}{P_n^3}\right)^{1/3^{n+1}} > 1.$$

Op analoge manier zien we dat $v_0, v_1, v_2, \ldots$ een dalende rij is. Dus

$$u_0 < u_1 < u_2 < \ldots\ldots < v_2 < v_1 < v_0.$$

Voor elke n geldt nu $u_n < A < v_n$.
Verheffen tot de macht 3^n geeft $P_n < A^{3^n} < P_n + 1$. Dus

$$\lfloor A^{3^n} \rfloor = P_n.$$

□

Een andere beroemde priemproducerende formule volgt uit het gebied van de mathematische logica. Een Diophantische vergelijking is een vergelijking van de vorm $P(x_1, \ldots, x_n) = 0$, met P een polynoom in n variabelen met gehele getallen als onbekenden.
Voorbeelden:

- De vergelijking van Pell voor gegeven $D > 0$: $x^2 - Dy^2 = 1$. Deze vergelijking heeft oneindig veel oplossingen als D geen kwadraat is, en alleen $x = \pm 1, y = 0$ als D een kwadraat is.

  Kleinste oplossing van $x^2 - 313y^2 = 1$ met $y > 0$,

  $$x = 32188120829134849, \quad y = 1819380158564160.$$

- De vergelijking van Mordell voor gegeven $k \neq 0$: $y^2 = x^3 + k$. Deze heeft hooguit eindig veel oplossingen.
  - Oplossingen van $y^2 = x^3 + 7$: geen
  - Oplossingen van $y^2 = x^3 - 2$: $x = 3$
  - Oplossingen van $y^2 = x^3 + 17$: $x = -2, -1, 2, 4, 8, 43, 52$

Vraag 1.8.6 (Hilbert, 1900 (moderne versie))**.** *Bestaat er een computerprogramma dat in staat is om van iedere Diophantische vergelijking te beslissen of er wel of geen oplossing in gehele getallen ≥ 0 bestaat?*

Hier is heel veel werk verzet door met name Hilary Putnam, Martin Davis, Julia Robinson. Met de afronding door Yuri Matijasevich in 1970.

Stelling 1.8.7. *Zo'n computerprogramma bestaat niet.*

In het bewijs van deze stelling speelt het begrip Diophantische verzameling een cruciale rol.

Definitie 1.8.8. *Een verzameling S van positief gehele getallen heet Diophantisch als er een polynoom $M(y, x_1, \ldots, x_n)$ bestaat zó dat $s \in S$ precies dan als $M(s, x_1, \ldots, x_n) = 0$ oplosbaar is in gehele $x_1, \ldots, x_2 \geq 0$.*

Hier zijn een paar voorbeelden van Diophantische verzamelingen en hun bijbehorende polynoom.

- S: kwadraten, $M(y, x) = y - x^2$.
- S: niet-priemgetallen, $M(y, x_1, x_2) = y - (x_1 + 2)(x_2 + 2)$.
- S: niet-kwadraten, $M(y, x_1, x_2) = x_1^2 - y(x_2 + 1)^2 - 1$.
- S: Fibonaccigetallen $F_n = 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \ldots$. Beschouw ook de rij $G_n = 2, 1, 3, 4, 7, 11, 18, 29, \ldots$. Dan geldt

  $$G_n^2 - 5F_n^2 = (-1)^n 4.$$

  We nemen $M(y, x) = (x^2 - 5y^2)^2 - 16$.
- S: De machten van twee, 2^k. M is zeer gecompliceerd.

Als klap op de vuurpijl hebben we het volgende.

Stelling 1.8.9 (Matiyusevich)**.** *De verzameling priemgetallen is Diophantisch.*

Er bestaat dus een polynoom $M(y, x_1, \ldots, x_n)$ zó dat $k+2$ priem is, precies dan als $M(k, x_1, \ldots, x_n) = 0$ oplosbaar is in gehele $x_1, \ldots, x_n \geq 0$.
Het blijkt dat M een som van kwadraten is en dus waarden ≥ 0 heeft.
Hieruit volgt:

Gevolg 1.8.10. *De verzameling positieve waarden van het polynoom* $(k+2)(1-M(k,x_1,\ldots,x_n))$ *met* $k,x_1,\ldots,x_n \geq 0$ *is precies de verzameling priemgetallen.*

Er zijn vele mogelijkheden voor dit priemproducerende polynoom. Eén ervan is

THEOREM 1. *The set of prime numbers is identical with the set of positive values taken on by the polynomial*

$$(1)\quad (k+2)\{1-[wz+h+j-q]^2-[(gk+2g+k+1)\cdot(h+j)+h-z]^2-[2n+p+q+z-e]^2$$
$$-[16(k+1)^3\cdot(k+2)\cdot(n+1)^2+1-f^2]^2-[e^3\cdot(e+2)(a+1)^2+1-o^2]^2-[(a^2-1)y^2+1-x^2]^2$$
$$-[16r^2y^4(a^2-1)+1-u^2]^2-[((a+u^2(u^2-a))^2-1)\cdot(n+4dy)^2+1-(x+cu)^2]^2-[n+l+v-y]^2$$
$$-[(a^2-1)l^2+1-m^2]^2-[ai+k+1-l-i]^2-[p+l(a-n-1)+b(2an+2a-n^2-2n-2)-m]^2$$
$$-[q+y(a-p-1)+s(2ap+2a-p^2-2p-2)-x]^2-[z+pl(a-p)+t(2ap-p^2-1)-pm]^2\}$$

as the variables range over the nonnegative integers.

Opmerking 1.8.11. *Als n een samengesteld getal is dan kan dit bewezen worden met 1 vermenigvuldiging.*

Met behulp van het polynoom M kunnen we de primaliteit van een getal $k+2$ bevestigen in hooguit 87 optellingen en vermenigvuldigingen. Je moet dan wel de juiste waarden van $x_1,\ldots,x_n$ kennen zó dat $M(k,x_1,\ldots,x_n)=0$. En dat is geen sinecure.

Hoe construeren we het polynoom M?

Stelling 1.8.12 (Wilson, 1770)**.** *Voor elk priemgetal p geldt* $(p-1)!\equiv -1 \mod p$.

Bewijs. $x^{p-1}-1$ heeft modulo p de nulpunten $1,2,\ldots,p-1$ (kleine stelling van Fermat).
Dus $x^{p-1}-1\equiv(x-1)(x-2)\cdots(x-(p-1)) \mod p$.
Vul links en rechts $x=0$ in: $-1\equiv 1\cdot 2\cdot 3\cdots(p-1) \mod p$. □

Opgave 1.8.13. *Stel n niet priem en >4. Toon aan dat* $(n-1)!\equiv 0 \mod n$.

Uit bovenstaande twee uitspraken concluderen we dat $k>1$ priem is precies dan als k een deler is van $(k-1)!+1$.

Anders gezegd: k is priem precies dan als er een gehele x bestaat zó dat $(k-1)!+1=kx$.

We moeten nu aantonen dat de de verzameling faculteitswaarden $n!$ een Diophantische verzameling vormt. We citeren

LEMMA 2.11. *For any positive integers k and f, in order that $f = k!$ it is necessary and sufficient that there exist nonnegative integers j, h, n, p, q, w and z such that*

$$\text{(I)}\quad q = wz + h + j, \qquad \text{(IV)}\quad p = (n+1)^k,$$

$$\text{(II)}\quad z = f(h+j) + h, \qquad \text{(V)}\quad q = (p+1)^n,$$

$$\text{(III)}\quad (2k)^3(2k+2)(n+1)^2 + 1 = \square, \qquad \text{(VI)}\quad z = p^{k+1}.$$

Het Diophantische karakter is nu gereduceerd tot het probleem van het Diophantisch karakter van de exponentiele functie a^n. Dit laatste was de cruciale stap in Matyasevich's bewijs.

2 De zeta-functie is een genererende functie voor de priemgetallen
Roland van der Veen

2.1 Inleiding

De bedoeling van deze twee voordrachten is een indruk te geven hoe de techniek van genererende functies kan worden toegepast om de priemgetallen beter te begrijpen. Telkens als we een rij getallen willen bestuderen zoeken we er een geschikte functie bij waar de hele rij op een mooie manier in verstopt zit. Deze zogenaamde genererende functie laat zich vaak gemakkelijker analyseren bijvoorbeeld door hem in factoren te ontbinden. In het geval van de priemgetallen heet de genererende functie zeta en zullen we zien hoe de factorizatie van zeta samenhangt met grote vraagstukken zoals de Riemann hypothese en de priemgetalstelling.
Aan het einde van elk hoofdstukje staan vraagstukken om zelf mee aan de slag te gaan, hier en daar met behulp van het computerprogramma PARI.

2.2 Wat is de zeta-functie en wat heeft die met priemgetallen te maken?

We hebben al een glimp van de zeta-functie gezien in Stelling 1.3 van het college van Frits. Euler liet zien dat de som $\sum_{p \text{ priem}} \frac{1}{p}$ divergeert. Het bewijs van die stelling begon met het product (afgekapt op een willekeurige $N \in \mathbb{N}$)

$$\prod_{p \text{ priem} \leq N} (1 + \frac{1}{p} + \frac{1}{p^2} + \dots) > \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{N}$$

het bewijs van deze ongelijkheid volgde door haakjes wegwerken.
Euler zag in dat precies hetzelfde ook werkt als je alle noemers tot de x-de macht verheft, voor een vastgekozen $x \in \mathbb{R}$. Immers $\frac{1}{p^x}\frac{1}{q^x}$ is nog steeds

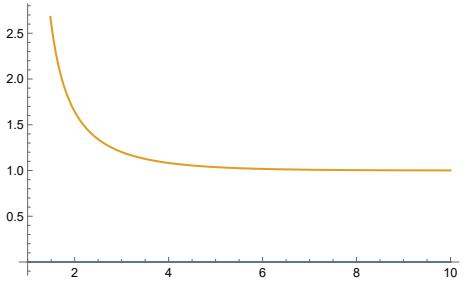
$\frac{1}{(pq)^x}$. Weer haakjes wegwerken geeft ons

$$\prod_{p \text{ priem} \leq N} (1 + \frac{1}{p^x} + \frac{1}{p^{2x}} + \dots) > \frac{1}{1^x} + \frac{1}{2^x} + \frac{1}{3^x} + \dots + \frac{1}{N^x}$$

De reeks aan de rechterkant wordt de (Riemann) zeta-functie (notatie: ζ) wanneer we N naar oneindig sturen.

$$\zeta(x) = \frac{1}{1^x} + \frac{1}{2^x} + \frac{1}{3^x} + \frac{1}{4^x} + \dots$$

We moeten natuurlijk wel voorzichtig zijn met deze limiet maar als $x > 1$ dan is deze oneindige reeks convergent en kunnen we inderdaad over een functie van x spreken $\zeta : (1, \infty) \to \mathbb{R}$. Hieronder is een grafiek van de functie getekend.



Figuur 2.1: Een grafiek van ζ op het domein $(1, 10)$

Het geval $x = 2$ gaf Euler een manier om na te denken over het in die tijd beroemde Basel probleem: de bepaling van de waarde van de reeks $\sum_{n=1}^{\infty} \frac{1}{n^2}$, dat is dus $\zeta(2)$. Euler gebruikte een ingenieuze factorizatie van de sinus om het probleem op te lossen en vond $\zeta(2) = \frac{\pi^2}{6}$. Daar komen we later op terug.

Wanneer we in de bovenstaande ongelijkheid de N naar oneindig sturen wordt de rechterkant gelijk aan de zeta functie maar nog beter is dat het ongelijkteken een gelijkheid wordt:

Stelling 2.2.1. ***(Eulers productstelling)***
Als $x > 1$ dan geldt:

$$\prod_{p \text{ priem}} (1 + \frac{1}{p^x} + \frac{1}{p^{2x}} + \dots) = \zeta(x)$$

Deze stelling vertelt ons dat een relatief eenvoudige functie $\zeta(x)$ alle informatie over de gehele rij van priemgetallen bevat. Dit is een thema in het werk van Euler en nu een vakgebied op zichzelf: probeer een rijen getallen te begrijpen door die te vangen in een functie en pas dan technieken uit de calculus en analyse toe op deze functie om informatie over de oorspronkelijke rij getallen te vinden.

Zo een functie wordt een genererende functie van de rij getallen genoemd. We kunnen de stelling van Euler dus losjes parafraseren door te zeggen: De zeta-functie is een genererende functie van de rij van priemgetallen.

Opgaven

Opgave 2.2.2.

De grafiek van zeta. In de hele opgave nemen we aan dat $x > 1$.

a). Ga door de termen in de reeksen handig te vergelijken na dat

$$\zeta(x) < \frac{1}{1^x} + \frac{2}{2^x} + \frac{4}{4^x} + \frac{8}{8^x} + \dots$$

b). Gebruik de meetkundige reeks om de reeks uit de vorige opgave te herschrijven als $v(x) = \dfrac{1}{1-2^{1-x}}$ en concludeer dat $\zeta(x) < v(x)$.

c). Gebruik de meetkundige reeks opnieuw om met een vergelijkbaar argument aan te tonen dat $u(x) < \zeta(x)$ met $u(x) = \dfrac{1}{1-2^{-x}}$.

d). Gebruik PARI om $\zeta(2)$ en $\pi^2/6$ te berekenen (gebruik `Pi^2/6`). Bereken ook eens $\zeta(3)$[1].

e). Gebruik het PARI commando `ploth` om de grafieken van ζ en u en v te vergelijken. hint: `ploth(x=1.1,3,[zeta(x),u(x),v(x)])`

f). Verklaar waarom de grafiek van ζ een verticale asymptoot heeft bij $x = 1$. Waarom is dat niet in tegenspraak met de gevonden grens $\zeta(x) < v(x)$?

g). Heeft ζ ook een horizontale asymptoot?

2.3 Genererende functies

Voordat we laten zien hoe de analyse van de zeta-functie ons informatie over de rij van priemgetallen geeft, bestuderen we eerst wat eenvoudigere

[1] Frits heeft over deze constante een beroemd artikel geschreven, zie de sectie Meer lezen aan het eind van deze tekst.

voorbeelden van genererende functies. Hoewel er meerdere types genererende functies zijn is het thema altijd hetzelfde. De genererende functie is een soort waslijn waar we een hele rij getallen aan kunnen ophangen. Zo kunnen we de rij collectief bestuderen zonder afgeleid te worden door individuele leden van de rij.

Het eenvoudigste type genererende functie voor een rij getallen $a_0, a_1, \ldots$ is een machtreeks

$$G(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \ldots$$

Aan de hand van zulke genererende functies zullen we de filosofie en technieken illustreren die ook voor priemgetallen en de zeta-functie relevant zijn.

De genererende functie G die hoort bij de rij $1, r, r^2, r^3, r^4, \ldots$ van machten van een getal r is de reeks die we kennen als de meetkundige reeks $1 + rx + (rx)^2 + (rx)^3 + \cdots = \dfrac{1}{1-rx}$. Opzichzelf is de meetkudige reeks geen interessant voorbeeld van een genererende functie maar het is een belangrijk bouwsteentje in heel veel complexere genererende functies.[2]

Een interessanter voorbeeld is de rij van Fibonaccigetallen: 0, 1, 1, 2, 3, 5, 8, 13,... gedefineerd door $f_0 = 0, f_1 = 1$ en $f_n = f_{n-1} + f_{n-2}$. De bijbehorende genererende functie ziet er uit als:

$$G(x) = x + x^2 + 2x^3 + 3x^4 + 5x^5 + \ldots$$

de recursierelatie $f_n = f_{n-1} + f_{n-2}$ maakt dat $G(x)$ een nogal eenvoudige vorm krijgt: Zet de volgende functies maar eens onder elkaar $G(x)$, $xG(x)$, $x^2G(x)$. Door met een macht van x te vermenigvuldigen schuiven de coefficienten op:

$$\begin{aligned}
G(x) &= x+ \quad x^2+ \quad 2x^3+ \quad 3x^4+ \quad 5x^5+ \quad 8x^6+ \quad 13x^7 + \ldots && (2.1)\\
xG(x) &= 0+ \quad x^2+ \quad x^3+ \quad 2x^4+ \quad 3x^5+ \quad 5x^6+ \quad 8x^7 + \ldots && (2.2)\\
x^2G(x) &= 0+ \quad 0+ \quad x^3+ \quad x^4+ \quad 2x^5+ \quad 3x^5+ \quad 5x^7 + \ldots && (2.3)
\end{aligned}$$

Hieruit volgt dat $G(x) - xG(x) - x^2G(x) = x$ en dus $G(x) = \dfrac{x}{1-x-x^2}$.

De eenvoudige genererende functie $G(g) = \dfrac{x}{1-x-x^2}$ weet blijkbaar alles over de Fibonacci getallen.

Hoe krijgen we onze G zover dat hij ons informatie over de Fibonaccigetallen geeft? Het sleutelbegrip is om de genererende functie te factorizeren.

[2] zie bijvoorbeeld ook de Productstelling van Euler, dat is een product van meetkundige reeksen!

In ons geval is de noemer kwadratisch dus als we dit polynoom factorizeren kunnen we G herschrijven in termen van meetkundige reeksen als volgt. In termen van de gulden getallen $\varphi = \dfrac{1+\sqrt{5}}{2}$ en $\phi = \dfrac{1-\sqrt{5}}{2}$ factoriseren we (met de 1 voorop met het oog op de meetkundige reeks):

$$G(x) = \frac{x}{(1-x\varphi)(1-x\phi)} = \frac{1}{\sqrt{5}}\left(\frac{1}{1-x\varphi} - \frac{1}{1-x\phi}\right)$$

We zien dus dat G de som is van twee meetkundige reeksen en daaruit volgt direct de beroemde formule van Binet voor de Fibonaccigetallen:

$$f_n = \frac{\varphi^n - \phi^n}{\sqrt{5}}$$

Opgave 2.3.1. *In deze opgave proberen we de techniek van genererende functies uit op de rij* $0, 1, 3\ldots$ *gegeven door* $a_0 = 0$ *en* $a_n = 2a_{n-1} + 1$. *In dit geval kunnen we de rij ook makkelijk direct bestuderen maar onze bedoeling is om de techniek te illustreren.*

a). Definieer een genererende functie $A(x) = \sum_{n=0}^{\infty} a_n x^n$ *en schrijf de eerste paar termen van* $A(x)/x$ *en* $\dfrac{1}{1-x}$ *en* $2A(x)$ *onder elkaar zoals we in de tekst voor* $G(x), xG(x)$ *en* $x^2G(x)$ *deden.*

b). Vind een verband tussen deze drie reeksen en bewijs hieruit dat $A(x) = \dfrac{x}{(1-x)(1-2x)}$.

c). Laat door breuksplitsing zien dat $A(x) = x\left(\dfrac{2}{1-2x} - \dfrac{1}{1-x}\right)$.

d). Vind uit de vorige formule voor $A(x)$ *een expliciete formule voor* a_n.

Pari opgave 2.3.2. *Gebruik PARI om de gevonden formule van Binet voor* f_n *te controleren. Controleer ook de formule die je in de vorige opgave gevonden hebt.*

```
phi1=(1+sqrt(5))/2;
phi2=(1-sqrt(5))/2;
f(n)=(phi1^n-phi2^n)/sqrt(5);
print(vector(10,n,f(n)))
```

2.4 Complexe getallen

Door het factoriseren van genererende functies komen we vanzelf met complexe getallen in aanraking. Hadden we bijvoorbeeld niet de Fibonaccigetallen genomen maar in plaats daarvan de reeks $0, 1, 1, 0, -1, -1, 0, 1, \ldots$ gegeven door $b_0 = 0, b_1 = 1$ en $b_n = b_{n-1} - b_{n-2}$ dan hadden we de genererende functie $G(x) = \dfrac{x}{1 - x + x^2}$. Volgens de abc-formule zijn de nulpunten in dit geval $\theta = \dfrac{1 + \sqrt{-3}}{2}$ en $\theta^{-1} = \dfrac{1 - \sqrt{-3}}{2}$. Als we bereid zijn verder te rekenen met het getal $\sqrt{-3}$ dan levert de techniek uit de vorige sectie weer een aantrekkelijke formule op:

$$b_n = \frac{\theta^n - \theta^{-n}}{\sqrt{-3}}$$

Iets systematischer werken we met het imagnaire getal $\mathbf{i}$, een nieuw getal dat niet op de getallenlijn te vinden is omdat we afspreken dat het voldoet aan $\mathbf{i}^2 = -1$. Verder proberen we alle bekende rekenregels vast te houden zodat $\sqrt{-3}$ geinterpreteerd wordt als $\sqrt{3}\mathbf{i}$. De getallen die zo ontstaan zijn allemaal van de vorm $u + v\mathbf{i}$ waarbij u, v reeele getallen zijn. Bij de berekeningen wordt $\mathbf{i}$ behandeld als een variabele maar wordt $\mathbf{i}^2$ steeds vervangen door -1. Vermenigvuldigen gaat bijvoorbeeld zo: $(2 + 3\mathbf{i})(4 + \mathbf{i}) = 8 + 3\mathbf{i}^2 + 14\mathbf{i} = 8 - 3 + 14\mathbf{i} = 5 + 14\mathbf{i}$. Net als rekenen met $\sqrt{5}$ dus.

De complexe getallen zijn essentieel om de priemgetallen te begrijpen via de zeta-functie. Kijk nog maar eens naar de grafiek van de zeta-functie, de grafiek laat geen nulpunten zien en vullen we een positieve waarde in voor x dan zal $\zeta(x)$ zeker ook positief zijn. Willen we de zetafunctie toch factorizeren dan zullen we dus complexe waarden moeten toelaten als nulpunten.

Pari opgave 2.4.1. *PARI begrijpt iets als* `sqrt(-3)` *en herschrijft het meteen als* `1.73*I` *waarbij* `I` *onze* $\mathbf{i}$ *is*[3]*. Doe de volgende berekeningen in PARI en met de hand:*

a). `2+3*I+1+2*I`

b). `(2+3*I)*(1+2*I)`

c). `(2+3*I)/(1+2*I)`

Opgave 2.4.2. *Ga na dat de berekening van de genererende functie G voor de rij b_n correct is en gebruik PARI om te checken dat de formule voor b_n klopt.*

[3]Pas op: vergeet in PARI nooit `*` te schrijven!

Opgave 2.4.3. *De norm (ook wel absoluute waarde) van een complex getal $z = x+\mathbf{i}y$ is gedefinieerd als $|z| = \sqrt{x^2+y^2}$. Als we het getal z meetkundig voorstellen als een punt (x, y) in het platte vlak dan is $|z|$ dus de afstand tot de oorsprong.*

a). Bereken $|1+2\mathbf{i}|$ en $|\theta|$ waarbij θ in deze sectie gedefinieerd was.

b). Laat zien dat $|z||w| = |zw|$

c). In PARI is het commando `norm(z)`.
Probeer eens `norm(cos(Pi/5)+I*sin(Pi/5))`

2.5 Nulpunten van de zetafunctie

Om daadwerkelijk nulpunten van ζ te kunnen berekenen is het nodig om het domein van de zeta-functie nog verder uit te breiden met behulp variant op de reeks die bekend staat als de eta-functie (notatie η):

$$\eta(x) = \frac{1}{1^x} - \frac{1}{2^x} + \frac{1}{3^x} - \frac{1}{4^x} \cdots$$

Dankzij de extra mintekens convergeert deze reeks voor alle $x > 0$. De relatie met de zeta-functie volgt door te kijken naar het verschil $\zeta(x) - \eta(x)$. Alle oneven termen vallen daarin weg en we houden juist een reeks over die we kunnen herkennen als $\frac{2}{2^x}\zeta(x)$.

We kunnen onze ζ nu uitbreiden door de stellen dat voor alle $x > 0$ geldt:

$$\zeta(x) = \frac{\eta(x)}{1 - 2^{1-x}}$$

als $x > 1$ dan klopt dit met onze originele definitie van $\zeta(x)$ maar voor $0 < x < 1$ werkt het nu ook!

Zoals gezegd hebben we complexe getallen nodig om de nulpunten van de zeta-functie te zien. Bernhard Riemann was de eerste die systematisch naar de zeta-functie en zijn complexe nulpunten keek. Hij vond twee nulpunten $\rho \approx \frac{1}{2} \pm 14.134725\mathbf{i}$ en nog een paar meer, altijd van de vorm $\frac{1}{2} \pm y\mathbf{i}$ met $y \in \mathbb{R}$.

Zonder een bewijs te geven merkte Riemann in zijn artikel over de zeta-functie op dat alle (niet-reële) nulpunten van de zeta-functie waarschijnlijk van deze vorm waren. Dit is later de *Riemann-hypothese* gaan heten. Een van de beroemdste onopgeloste vraagstukken in de wiskunde.

Opgaven

Opgave 2.5.1. *Eta-functie*

a). Schrijf alle termen uit in $\zeta(x)-\eta(x)$ *en schrijf ook de reeks* $2^{1-x}\zeta(x)$ *uit en controleer zo dat* $(1-2^{1-x})\zeta(x)=\eta(x)$.

b). Laat zien dat $\eta<1$ *voor alle* $x>1$ *met behulp van de functie* $v(x)$ *uit de opgave van sectie 1.*

c). Bereken $\eta(1)$ *met behulp van de reeks voor de logaritme* $\log(1-x)=\sum_{n=1}^{\infty}\frac{x^n}{n}$

d). Gebruik PARI om na te gaan dat $\zeta(1+2\pi i/\log(2))\neq 0$ *en laat zien toch geldt:* $\eta(1+2\pi i/\log(2))=0$.

Pari opgave 2.5.2. *Hieronder staat een PARI programma om het eerste nulpunt van de zeta functie te vinden. Het ligt ergens tussen de* $\frac{1}{2}+14\mathbf{i}$ *en* $\frac{1}{2}+15\mathbf{i}$. *Run het programma om de waarde te vinden die Riemann ook vond (maar dan in meer decimalen!). Pas de code nu aan om meer nulpunten te vinden! Hint: de eerstvolgende ligt ergens bij* $\frac{1}{2}+21\mathbf{i}$ *maar er zijn er nog veel meer*[4]. *Hoeveel kan je er vinden met dit PARI script? De tekst achter de* \\ *is commentaar en kan worden genegeerd.*

```
{
stappen = 0;              \\houdt het aantal stappen bij
gok1 = 1/2 + 14*I;        \\zoek een zeta-nulpunt tussen
gok2 = 1/2 + 15*I;        \\gok1 en gok2.
precisie = 10^(-50);      \\met gewenste precisie
waarde1 = zeta(gok1);     \\bereken de zeta-waarde
until (norm(waarde2) < precisie||stappen>10^3,
                          \\totdat preciesie is bereikt
stappen+=1;               \\of na maximaal 10^3 stappen
waarde2 = zeta(gok2);     \\bereken zeta-waarde bij gok2
if (norm(waarde2) < norm(waarde1),
                          \\als waarde2 kleiner is dan waarde1
wissel = gok1; gok1 = gok2; gok2 = wissel;
                          \\dan wordt gok2 de nieuwe gok1
waarde1 = waarde2;        \\en waarde2 de nieuwe waarde1
);
gok2 = (gok1+gok2) / 2.; \\de nieuwe gok2 is het gemiddelde
```

[4] Misschien ook bij $\frac{1}{2}-14\mathbf{i}$?

```
                              van de vorige gokken.
print("na " stappen " stappen is het");
print("laagste punt: " gok1);
print("de waarde daar is: " waarde1)
}
```

2.6 Oneindige producten

Euler experimenteerde al met factorizatie van functies met oneindig veel nulpunten. Een beroemd voorbeeld is de sinus-functie. We weten dat de nulpunten precies de gehele veelvouden van π zijn. Dus schreef Euler:

$$\sin x = x\left(1-\frac{x}{\pi}\right)\left(1-\frac{x}{-\pi}\right)\left(1-\frac{x}{2\pi}\right)\left(1-\frac{x}{-2\pi}\right)\left(1-\frac{x}{3\pi}\right)\left(1-\frac{x}{-3\pi}\right)\ldots$$

we kunnen de termen $\left(1-\frac{x}{n\pi}\right)\left(1-\frac{x}{-n\pi}\right)$ samen nemen tot $(1-\frac{x^2}{n^2\pi^2})$ en krijgen dan

$$\sin x = x\left(1-\frac{x^2}{\pi^2}\right)\left(1-\frac{x^2}{2^2\pi^2}\right)\left(1-\frac{x^2}{3^2\pi^2}\right)\ldots$$

Doen we weer alsof de sinus een polynoom is waarvan de eerste paar termen gelijk zijn aan $\sin x = x - \frac{x^3}{6} + \ldots$ dan hebben we nu twee formules voor de coefficient van x^3 dus we vinden door haakjes weg te werken in de factorizatie van de sinus:

$$\frac{1}{6} = \frac{1}{\pi^2} + \frac{1}{2^2\pi^2} + \frac{1}{3^2\pi^2} + \ldots$$

Hieruit volgt Eulers oplossing van het Basel probleem: $\zeta(2) = \frac{\pi^2}{6}$.

Riemann vermoedde dat een vergelijkbare factorisatie ook mogelijk moest zijn voor de zeta-functie. Hadamard slaagde er later in om inderdaad een factorisatie te vinden in termen van de nulpunten[5]:

$$\zeta(x) = \frac{1}{x-1}\left(\frac{2\pi}{e}\right)^x \prod_\rho \left(1-\frac{x}{\rho}\right) e^{\frac{x}{\rho}}$$

Door deze productformule te vergelijken met Eulers product formule vond Riemann de zogenaamde expliciete formule voor een priem-telfunctie $\psi(x)$

[5] en nog wat andere termen die te maken hebben met hoe snel ζ groeit als je x groot laat worden of juist nadert tot de asymptoot in $x = 1$.

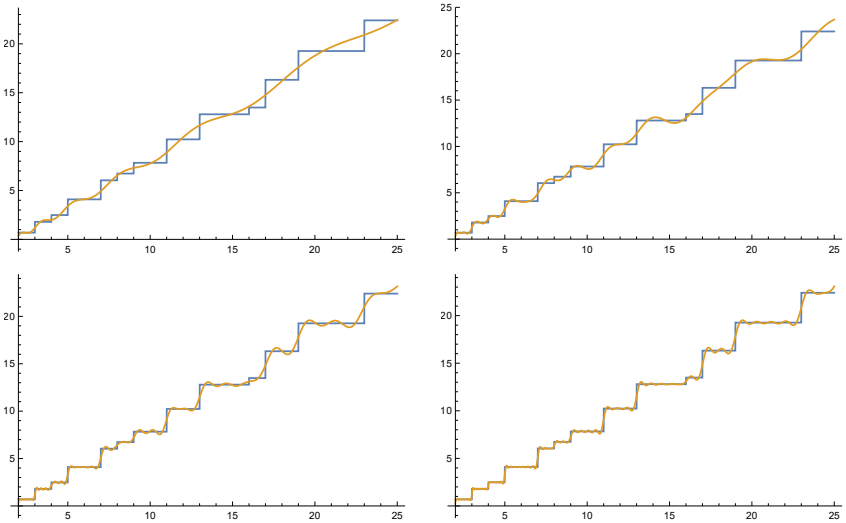
die erg lijkt op de priemtelfunctie $\pi(x)$. In het college van Frits hebben we $\psi(x)$ al ontmoet in termen van het kleinst gemene veelvoud: $\psi(x) = \log \mathrm{kgv}(1, 2, 3, \ldots, [x])$ waarbij $[x]$ het grootste gehele getal kleiner dan x betekent. De expliciete formule zegt nu dat:

$$\psi(x) = x - \sum_{\rho} \frac{x^{\rho}}{\rho} - \log(2\pi)$$

De priemgetalstelling $\pi(x) \sim \dfrac{x}{\log x}$ is overigens equivalent met de benadering

$$\psi(x) = \log \mathrm{kgv}(1, 2, 3, \ldots, [x]) \sim x$$

De expliciete formule vertelt ons dus aan de hand van de zeta-nulpunten hoe ver $\psi(x)$ afwijkt van x. In Figuur 2.2 zien we wat er gebeurt als we de expliciete formule afkappen bij de eerste 1 of 5 of 20 of 50 paren zeta-nulpunten van de vorm $\dfrac{1}{2} \pm y\mathbf{i}$.



Figuur 2.2: Hoe meer zeta-nulpunten we meenemen in de expliciete formule, hoe beter we de priem-telfunctie $\psi(x)$ in blauw benaderen. Hier zijn dat $1, 5, 20$ en 50 paren nulpunten $\dfrac{1}{2} \pm y\mathbf{i}$.

De betekenis van de Riemann hypothese komt duidelijk naar voren in deze formule want de snelheid waarmee x^{ρ} groeit wordt precies bepaald door

het reëele deel van ρ. Volgens Riemann is dat nooit meer dan $\frac{1}{2}$. De hypothese zegt dus dat er geen grote fluctuaties in de expliciete formule zijn en dat de priemgetallen zich zo goed mogelijk gedragen.

Opgave 2.6.1. *De sinus-product formule.*

a). Vul eens $x = \frac{\pi}{2}$ *in in de product-formule voor de sinus. Kun je zo uitkomen op het beroemde Wallisproduct voor* π*?*

$$\frac{\pi}{2} = \frac{2\cdot 2}{1\cdot 3}\cdot\frac{4\cdot 4}{3\cdot 5}\cdot\frac{6\cdot 6}{5\cdot 7}\cdot\frac{8\cdot 8}{7\cdot 9}\cdots$$

b). Neem de coefficient van x^5 *in de Taylor expansie van de sinus:* $\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots$ *en vergelijk die met de coefficient van* x^5 *in de product-formule voor de sinus. Kun je zo de waarde van* $\zeta(4)$ *achterhalen?*

Opgave 2.6.2. *Paren nulpunten. De geconjugeerde van een complex getal* $z = x + y\mathbf{i}$ *is het getal* $\bar{z} = x - \mathbf{i}y$*.*

a). Laat zien dat voor alle complexe z *geldt:* $z\bar{z} = |z|^2$ *en* $\bar{z} + z \in \mathbb{R}$*.*

b). Bewijs dat voor $z, w \in \mathbb{C}$ *geld:* $\overline{z+w} = \bar{z} + \bar{w}$ *en* $\overline{zw} = \bar{z}\bar{w}$*.*

c). Stel we hebben een polynoom $f(z) = a_0 + a_1 z + a_2 z^2 + \ldots$ *met reëele coëfficienten* $a_0, a_1 \ldots$*. Bewijs dat* $f(z) = 0$ *betekent dat ook* $f(\bar{z}) = 0$*.*

d). Concludeer dat nulpunten die niet reëel zijn dus in paren voorkomen. Iets vergelijkbaars werkt ook voor oneindige reeksen (met reëele coëfficienten)en verklaart waarom de nulpunten van de zeta-functie in paren voorkomen.

2.7 Meer lezen?

F. Beukers, (1979), "A Note on the Irrationality of $\zeta(2)$ and $\zeta(3)$", Bull. London Math. Soc., 11 (3): 268–272.

J. van de Craats, R. van der Veen, *De Riemann-hypothese*, Epsilon 2015. Zie ook de uitgebreide tweede editie uitgegeven bij de Mathematical Association of America, 2018.

J. Stopple, *A primer of analytic number theory*, Cambridge 2003.

H. Wilf, *generatingfunctionology*, Academic Press, 1994.

3 Priemgetal-heuristiek
Benne de Weger

3.1 Inleiding

Wiskunde wordt traditioneel nog wel eens gezien als het bewijzen van stellingen. Je wilt toch graag wel zeker weten wat er aan de hand is, en een bewijs geeft je wat wel gezien wordt als de hoogste soort zekerheid, die gebaseerd op logica. Daar is natuurlijk veel voor te zeggen, maar er zijn ook wel beroemde vermoedens in de wiskunde, zoals het Vermoeden van Goldbach waarf Frits het over heeft, en de Hypothese van Riemann die centraal staat in de bijdrage van Roland. Kurt Gödel heeft weliswaar aangetoond dat er meer ware beweringen zijn dan bewijsbare, maar de meeste wiskundigen blijven positief ingesteld en zeggen dat we die bekende vermoedens *nóg* niet bewezen hebben, bijvoorbeeld omdat we daar met z'n allen nog niet aan toe zijn. Er is veel 'echte', d.w.z. bewezen, wiskunde ontwikkeld gestuurd door het verlangen om een bekend vermoeden te ontwikkelen, zoals in het geval van de Laatste Stelling van Fermat: de zoektocht die uiteindelijk leidde tot het bewijs van Wiles en Taylor heeft een enorme hoeveelheid fraaie wiskunde opgeleverd in de algebraische getaltheorie en arithmetische meetkunde.

Vermoedens spelen dus zeker een belangrijke rol in de wiskunde. Er is meer dan alleen stellingen en bewijzen. Er zijn zelfs vele stellingen van de vorm "Als Vermoeden X waar is, dan is ook Vermoeden Y waar", en het is zeker geen schande zo'n stelling te bewijzen. Maar je moet ook weer niet zomaar een of andere bewering verzinnen en die een "Vermoeden" gaan noemen, het is dan wel de bedoeling dat je *aannemelijk* maakt dat jouw vermoeden wel eens waar zou kunnen zijn. Dan kom je in het gebied van de *heuristiek* terecht. Het Griekse 'heuristein' betekent 'vinden' (denk aan 'Eureka'), heuristiek is dan zoiets als 'de kunst van het vinden'. Van Dale zegt ook 'de wetenschap die langs methodische weg tot ontdekkingen leert komen' en 'de methode die de leerling bij het onderwijs gelegenheid geeft zelf waarheden en regels te vinden'. In de praktijk betekent het in de getaltheorie vaak een 'probabilistische' redenering die via het opstellen van een kansmodel tot ideeën komt over wat wel eens waar zou kunnen

zijn. In dit deel van de cursus gaan we dat doen voor diverse telproblemen in de verzameling van de priemgetallen.

3.2 De Priemgetalstelling bekeken vanuit de kansrekening

De priemgetalstelling, al door Frits besproken, is er in een nauwkeuriger maar wat abstractere vorm:

$$\pi(x) \sim \mathrm{li}(x) = \int_2^x \frac{1}{\log t}\,\mathrm{d}t \quad \text{voor} \quad x \to \infty,$$

en in een wat 'slordiger' maar makkelijker te hanteren vorm:

$$\pi(x) \sim \frac{x}{\log x} \quad \text{voor} \quad x \to \infty.$$

Slordiger in de zin dat voor eindige x de benadering over het algemeen minder nauwkeurig is: de fout $|\pi(x) - \mathrm{li}(x)|$ is, als je de Riemann-Hypothese gelooft, kleiner dan $\sqrt{x}\log x$ voor x groot genoeg, terwijl de andere fout $\left|\pi(x) - \dfrac{x}{\log x}\right|$ van de grootteorde $\dfrac{x}{(\log x)^2}$ is, veel groter dus.
Van belang is om op te merken dat de Priemgetalstelling een *asymptotisch* resultaat is, dus iets zegt over de situatie waarbij je x naar ∞ laat gaan. Vaak wil je iets meer weten over hoe snel $\pi(x)$ naar $\mathrm{li}(x)$ of naar $\dfrac{x}{\log x}$ gaat, en wil je een verfijning van de asymptotische formule met een foutschatting. Het is bijvoorbeeld zo dat uit de Riemann-hypothese volgt dat

$$\pi(x) = \frac{x}{\log x} + \frac{x}{(\log x)^2} + \text{ een fout van grootteorde } \frac{x}{(\log x)^3} \text{ voor } x \to \infty.$$

Dit is nog steeds een asymptotische bewering. Maar ook wil je wel enige richtlijn hebben over wat er bij x van een bepaalde grootte ongeveer gebeurt. Als cryptoloog wil je bijvoorbeeld redelijke zekerheid (niet perse een bewijs) hebben over de vraag hoeveel priemgetallen er in een bepaald interval zullen zijn. Daar kan de Priemgetalstelling geen bewezen uitspraken over doen, maar wel een heuristiek geven, bijvoorbeeld door uit een asymptotische uitspraak simpelweg het gedeelte over de fout en het "voor $x \to \infty$" weg te laten. Dan mag je er formeel geen '='-teken meer bij gebruiken, en schrijven we "$\approx$", daarbij in het midden latend hoe groot de fout is in de benadering. Dus:

$$\pi(x) \approx \frac{x}{\log x} + \frac{x}{(\log x)^2}.$$

Dit lijkt aardig te kloppen, ook al voor redelijk kleine x, zie Tabel 3.1.

x	$\pi(x)$	$\frac{x}{\log x}$	$\frac{x}{\log x} + \frac{x}{(\log x)^2}$	$\epsilon(x)$
10^8	5761455	5428681	5723387	2.38
10^9	50847534	48254942	50583482	2.35
10^{10}	455052511	434294482	453155652	2.32
10^{11}	4118054813	3948131654	4104009089	2.28
10^{12}	37607912018	36191206825	37501010277	2.26
10^{13}	346065536839	334072678387	345233133832	2.23
10^{14}	3204941750802	3102103442166	3198333899825	2.21
10^{15}	29844570422669	28952965460217	29791239669157	2.20

Tabel 3.1: Benaderingen voor $\pi(x)$, $\epsilon(x)$ is de relatieve fout t.o.v. $\frac{x}{(\log x)^3}$,

$$\text{dus } \epsilon(x) = \frac{\pi(x) - \frac{x}{\log x} - \frac{x}{(\log x)^2}}{\frac{x}{(\log x)^3}}.$$

Een manier om hier tegenaan te kijken is met behulp van kansrekening. Frits deed dat al een beetje toen hij zei dat de 'dichtheid' van de priemgetallen ter grootte X ongeveer $1/\log X$ is. Dat betekent dat in de buurt van X ongeveer 1 op de $\log X$ getallen een priemgetal is. Je kunt ook zeggen dat de kans dat een willekeurig getal in de buurt van X een priemgetal is, ongeveer $1/\log X$ is. Met behulp van de vuistregel 'kans = aantal gunstige gebeurtenissen gedeeld door aantal mogelijke gebeurtenissen' kunnen we het zo uit de Priemgetalstelling afleiden: neem een interval $(X-\Delta, x+\Delta)$ rondom X, waarbij $\Delta \ll X$, dan zijn er ongeveer 2Δ getallen, waarvan $\pi(X+\Delta) - \pi(X-\Delta)$ priemgetallen zijn. De kans dat een willekeurig getrokken getal uit het interval een priemgetal is, is dan

$$\frac{\pi(X+\Delta) - \pi(X-\Delta)}{2\Delta} \approx \frac{\frac{X+\Delta}{\log(X+\Delta)} - \frac{X-\Delta}{\log(X-\Delta)}}{2\Delta}.$$

Hier kunnen we grip op krijgen door te gebruiken $\log(X \pm \Delta) = \log X + \log\left(1 + \frac{\pm\Delta}{X}\right)$, waarbij $\frac{\pm\Delta}{X}$ nu in de buurt van 0 zit, zodat we maar een heel kleine fout maken als we binnen de $\log(X \pm \Delta)$ de Δ verwaarlozen. Dan zien we meteen

$$\frac{\pi(X+\Delta) - \pi(X-\Delta)}{2\Delta} \approx \frac{\frac{X+\Delta}{\log X} - \frac{X-\Delta}{\log X}}{2\Delta} = \frac{1}{\log X}.$$

Opgave 3.2.1. *Maak een soortgelijke redenering door nu* $\mathrm{li}(x)$ *te gebruiken in plaats van* $\dfrac{x}{\log x}$. *Gebruik dat de functie* $\log x$ *op het interval* $(X - \Delta, X + \Delta)$ *vrijwel constant is.*

Opgave 3.2.2. *Frits heeft het over gaten tussen de priemgetallen:* $g_n = p_{n+1} - p_n$, *waarbij* p_n *het* n*-e priemgetal is, waarbij* g_n *zo klein als* 2 *kan zijn, ook vaak groter dan* $\log p_n$, *maar vermoedelijk (Cramér) niet groter dan* $C(\log p_n)^2$. *Leid uit de Priemgetalstelling een heuristiek af voor de grootte van het gemiddelde gat* g_n, *in termen van* p_n.

Opgave 3.2.3. *Cryptologen willen graag weten dat er voldoende priemgetallen zijn van een bepaalde grootte, bv. die in binaire schrijfwijze een voorgeschreven aantal bits hebben, zeg* b. *Geef een zo eenvoudig mogelijke formule die het aantal priemgetallen telt in het interval* $[2^{b-1}, 2^b - 1]$. *Het aantal elementaire deeltjes in het universum wordt geschat op* 3×10^{80}. *Hoe groot moet je* b *kiezen om op ongeveer evenveel priemgetallen van* b *bits uit te komen? Ter vergelijking: de meeste websites gebruiken voor hun beveiliging tegenwoordig* $b = 1024$.

3.3 Priemgetallen in congruentieklassen: het blijkt niet eerlijk te zijn

We gaan een stapje verder. Frits noemt de stellingen van Dirichlet en Siegel-Walfisz. Laat $q > 1$ een positief geheel getal zijn, en a een geheel getal ≥ 1 en $\leq q - 1$, dat relatief priem is met q. Dan tellen we het aantal $\pi(x, q, a)$ van de priemgetallen $p \leq x$ zodat $p \equiv a \pmod{q}$. De genoemde stellingen zeggen dan in feite dat

$$\pi(x, q, a) \sim \frac{1}{\phi(q)} \pi(x).$$

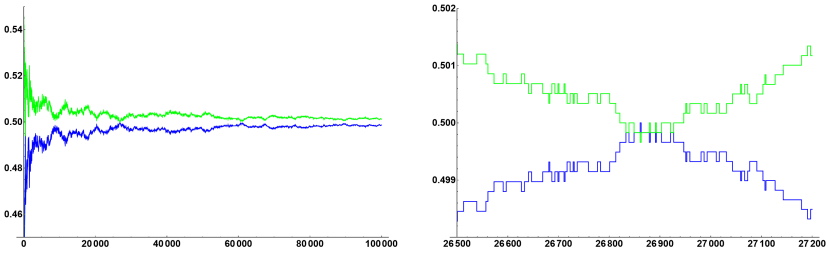
Merk op dat $\phi(q)$ precies het aantal mogelijke a's is.

Dit betekent dus een asymptotisch eerlijke verdeling van de priemgetallen over de congruentieklassen modulo q, de gegeven uitdrukking hangt niet van a af.

In termen van kansen zeggen we dan: voor een gegeven a zijn de kansen op de gebeurtenissen "x is een priemgetal" en "$x \equiv a \pmod{q}$" onafhankelijk, dus de kans dat beide gebeurtenissen tegelijk optreden is het product van de afzonderlijke kansen, respectievelijk $\dfrac{1}{\log x}$ en $\dfrac{1}{\phi(q)}$. Er is kennelijk geen reden om een afhankelijkheid van die gebeurtenissen te veronderstellen.

Pari opgave 3.3.1. *Met Pari kunnen we dit wel testen. De opdracht*
`tel(q,x) = t=vector(q-1); forprime(p=1,x,if(gcd(p,q)==1, t[p%q]++)); t`
definieert een functie `tel(q,s)` *die een lijstje geeft van de aantallen priemgetallen* $\leq x$ *in de congruentieklassen* $1, 2, \ldots, q-1 \pmod q$. *Bijvoorbeeld,* `tel(12,10000)` *geeft* `[300, 0, 0, 0, 309, 0, 311, 0, 0, 0, 307]`. *Speel hier wat mee; begin bijvoorbeeld met* $q = 3$ *of* $q = 4$ *en probeer verschillende waarden van* x*, je kunt rustig tot 1 miljoen gaan. Doe het ook eens met* $q = 10$*, dat geeft tellingen van priemgetallen met een gegeven laatste decimaal.*
Valt je iets bijzonders op?

In Figuur 3.1 staan de grafieken van $\dfrac{\pi(x,4,1)}{\pi(x)-1}$ (blauw) en $\dfrac{\pi(x,4,3)}{\pi(x)-1}$ (groen), op het interval $(2, 10^5)$, en ingezoomd rond 26900.



Figuur 3.1: $\dfrac{\pi(x,4,1)}{\pi(x)-1}$ (blauw) en $\dfrac{\pi(x,4,3)}{\pi(x)-1}$ (groen).

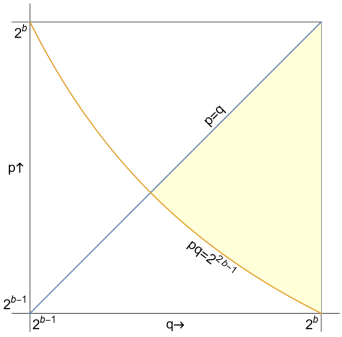
Wat nou eerlijk verdeeld. Het lijkt erop dat 3 (mod 4) systematisch meer van de priemgetallenkoek krijgt dan 1 (mod 4). Dit verschijnsel treedt bij alle moduli in meer of mindere mate op. het is niet in tegenspraak met de stelling van Siegel-Walfisz, want die zegt alleen iets over $x \to \infty$, en die heeft geen verder uitgewerkte foutterm.

Dit verschijnsel staat bekend als *Chebyshev's bias*. Rubinstein en Sarnak wisten in 1994 er een verklaring voor te geven, door te laten zien dat de fouttermen in de asymptotische ontwikkelingen van $\pi(x,q,a)$ niet allemaal dezelfde vorm hebben voor alle a. Het voert voor deze cursus veel te ver daar dieper op in te gaan. maar het verschijnsel is wel aan den lijve te ondervinden, als je gereedschap als Pari hebt.

3.4 RSA-moduli

Voor een RSA-sleutelpaar heb je twee grote priemgetallen p, q nodig, zo laat Frits zien. Cryptologen willen ze graag ongeveer even groot hebben, zeg elk van b bits (dus in het interval $(2^{b-1}, 2^b)$), en ook nog zodanig dat hun product $N = pq$ precies $2b$ bits heeft. En ze moeten natuurlijk verschillend zijn. Dan leveren ze een goede RSA-modulus N op.

Hoeveel goede RSA-moduli van $2b$ bits zijn er? Dat kunnen we met behulp van de Priemgetalstelling en onze heuristische aanpak goed schatten. We kunnen zonder verlies van algemeenheid stellen dat $p < q$. De eisen zijn dat p, q priemgetallen zijn, met $2^{b-1} < p < q < 2^b$, en $2^{2b-1} < pq < 2^{2b}$. Zie Figuur 3.2.



Figuur 3.2: Het toegestane gebied V (geel) voor goede RSA-moduli.

Voor ieder punt $(p, q) \in V$ hebben we nu als eis dat zowel p als q priemgetallen moeten zijn, met kansen respectievelijk $\dfrac{1}{\log p}$ en $\dfrac{1}{\log q}$, die niet overal in het gebied dezelfde zijn. Maar per punt zijn die twee kansen wel onafhankelijk, althans, we zien geen reden om anders te veronderstellen. We moeten dus $\dfrac{1}{(\log p)(\log q)}$ voor alle (p, q) in het gebied optellen. Je kunt proberen van deze dubbele som een tweevoudige integraal te maken en die uit te rekenen; soms gaat dat goed, maar in dit geval is het helaas nogal lastig.

Daarom gebruiken we voor p en q een overschatting van 2^b elk, dat zal een onderschatting geven van de kansen en dan zitten we aan de veilige kant. Dan hebben we opeens een kans die niet meer van de variabelen p en q afhangt, en om de som uit te rekenen hoeven we dan alleen het aantal punten in V met gehele coördinaten te hebben. Een goede benadering

daarvan is de oppervlakte van V, en die is uit te rekenen:

$$\int_{2^{b-\frac{1}{2}}}^{2^b} \left(q - \frac{2^{2b-1}}{q}\right) \mathrm{d}q = \left(\frac{1}{2}q^2 - 2^{2b-1}\log q\right)\Big|_{2^{b-\frac{1}{2}}}^{2^b} = 2^{2b-2}(1-\log 2).$$

Vermenigvuldigd met de schattingen voor de kansen, elk $\dfrac{1}{\log 2^b}$, krijgen we nu als ondergrens voor het aantal goede RSA-moduli

$$\frac{2^{2b}}{b^2}\frac{1-\log 2}{4(\log 2)^2} \approx 0.16\frac{2^{2b}}{b^2}.$$

Om een bovengrens te krijgen gebruiken we onderschattingen, 2^{b-1} voor p en q (beetje grof...). Eenzelfde redenering geeft dan een bovengrens voor het aantal goede RSA-moduli:

$$\frac{2^{2b}}{(b-1)^2}\frac{1-\log 2}{4(\log 2)^2} \approx 0.16\frac{2^{2b}}{(b-1)^2}.$$

Omdat we in de cryptografie tegenwoordig pas bij $b = 1024$ beginnen, is de bovengrens maar ongeveer 0.2% groter dan de ondergrens. Dat valt weg in de afronding naar 0.16 die we toch al maakten. Met $b = 1024$ is het aantal goede RSA-moduli dus $0.16 \times 2^{2028} \approx 5 \times 10^{609}$. Meer dan we als mensheid ooit zullen kunnen opmaken.

Een andere vraag die van groot belang is in de cryptografie is hoe we dit soort grote priemgetallen kunnen maken. Een theoretisch wiskundige kan gewoon zeggen: kies een priemgetal van 1024 bits. Maar een toegepast wiskundige heeft daar niet zoveel aan, die heeft een methode nodig die zowel effectief is als efficiënt. De methode die in de praktijk gebruikt wordt is in essentie verbijsterend simpel: genereer een willekeurige rij bits van 1022 lang, zet er een 1-bit voor en een 1-bit achter zodat je een getal krijgt dat echt 1024 bits heeft en niet minder, en oneven is; laat daar de priemtest van Rabin op los (zie de bijdrage van Frits), en als daar uitkomt dat het getal samengesteld is, dan gooi je het weg en begin je opnieuw.

Deze methode is efficiënt om twee redenen: Rabin is heel goed te doen voor getallen van 1024 bits (alleen een paar machtsverheffingen zijn nodig maar Frits heeft laten zien dat dat snel kan), en het verwachte aantal keren dat je opnieuw moet beginnen is 1 gedeeld door de kans op succes, en volgens de Priemgetalstelling is die kans $\dfrac{2}{\log 2^{1024}} \approx \dfrac{1}{355}$. Dus je verwacht dat je gemiddeld zo'n 300 keer hoeft te proberen. Fluitje van een cent.

De methode is ook effectief. Hoewel Rabin geen zekerheid geeft over een correct antwoord, is de kans op een misser (een getal dat door de test

heenkomt als vermoedelijk priem maar in feite samengesteld is) zo klein te maken dat je je daar echt geen zorgen over hoeft te maken. Dat klein maken doe je door de test een aantal keren te herhalen: de test is gerandomiseerd dus een paar keer toepassen op hetzelfde getal geeft toch iedere keer een andere test. Men kan aantonen dat die testresultaten met hoge kans onafhankelijk zijn, en dat je het maar een stuk of 6 ker hoeft te doen om de kans op een misser astronomisch klein te maken. Naar verwachting gaat het pas voor de eerste keer (wereldwijd) mis als de zon al uitgeblust is.

Pari opgave 3.4.1. *De volgende Pari-code maakt priemgetallen zoals zojuist beschreven.*

```
maakpriem(b) = x=0; t=0; while(!ispseudoprime(x),t++; x=2*
random(2^(b-2))+2^(b-1)+1); printf(``aantal pogingen:  %d'',t);x
```

Experimenteer er maar wat mee. De code laat zien hoeveel verschillende getallen getest zijn. Gebruik de met deze methode gemaakte priemgetallen ook eens bij Opgave 1.7.2 van Frits.

3.5 Priemtweelingen en Sophie Germain-priemgetallen

Een priemtweeling is een paar priemgetallen met een gat van slechts 2, dus $p, p+2$. Het is niet bekend of er oneindig veel van zijn, maar het sterke vermoeden is dat dat wel zo is. Een argument is de heuristiek aan de hand van de Priemgetalstelling. Uitgaande van de veronderstelling dat het priem zijn van p en $p+2$ onafhankelijke gebeurtenissen zijn, zou de kans erop $\dfrac{1}{(\log p)^2}$ zijn (eigenlijk $\dfrac{1}{(\log p)(\log(p+2))}$, maar dat is vrijwel

x	$\pi_2(x)$	$c(x)$
10^3	35	1.06562
10^4	205	1.27805
10^5	1224	1.29672
10^6	8169	1.30806
10^7	58980	1.32546
10^8	440312	1.32016
10^9	3424506	1.32002
10^{10}	27412679	1.32038

x	$\pi_2(x)$	$c(x)$
10^{11}	224376048	1.32037
10^{12}	1870585220	1.32034
10^{13}	15834664872	1.32033
10^{14}	135780321665	1.32032
10^{15}	1177209242304	1.32032
10^{16}	10304195697298	1.32032
10^{17}	90948839353159	1.32032
10^{18}	808675888577436	1.32032

Tabel 3.2: Priemtweelingen tellen, $c(x) = \dfrac{\pi_2(x)}{\mathrm{li}_2(x)}$.

hetzelfde), en zou de priemtweelingtelfunctie $\pi_2(x)$, het aantal priemgetallen $p \leq x$ zodanig dat ook $p+2$ priem is, asymptotisch equivalent zijn met $\mathrm{li}_2(x) = \displaystyle\int_2^x \frac{1}{(\log t)^2}\mathrm{d}t$, oftewel ook met $\dfrac{x}{(\log x)^2}$. Zie Tabel 3.2.

Hier is iets vreemds aan de hand: we verwachtten dat $c(x)$ naar 1 zou gaan convergeren, maar dat lijkt niet te gebeuren, en wel tamelijk overtuigend. Misschien is onze onafhankelijkheidsaanname niet goed?

Inderdaad is dat hier het probleem. Laten we beginnen met een priemgetal p, groter dan 2, dus oneven. Maar dan heeft $p+2$ al geen keus meer tussen even en oneven, en zal dus een grotere kans hebben om een priemgetal te zijn, en wel met een factor 2. En datzelfde kunnen we doen, niet alleen voor modulus 2 (even/oneven), maar voor iedere priem q: als we al weten dat q geen deler is van p, dan is de kans dat q ook geen deler is van $p+2$ niet meer $\dfrac{q-1}{q}$, maar $\dfrac{q-2}{q-1}$. Want we weten al dat er voor p slechts $q-1$ mogelijkheden $\pmod q$ zijn, namelijk $1, 2, \ldots, q-3, q-2, q-1$, maar niet 0. En dan zijn er voor $p+2 \pmod q$ ook $q-1$ mogelijkheden, namelijk $3, 4, \ldots, q-1, 0, 1$, maar niet 2. De kans dat $p+2$ niet deelbaar is door q, gegeven dat p dat niet is, is dus $\dfrac{q-2}{q-1}$. Daar staat tegenover dat de kans dat een volledig willekeurig getal niet deelbaar door q is, gelijk is aan $\dfrac{q-1}{q}$, en we moeten nu de impliciet aanwezige factor $\dfrac{q-1}{q}$ in de kans vervangen door $\dfrac{q-2}{q-1}$. Dat geeft dus een wijzigingsfactor van $\dfrac{q-2}{q-1}/\dfrac{q-2}{q-1} = \dfrac{q(q-2)}{(q-1)^2} = 1 - \dfrac{1}{(q-1)^2}$. Voor de priemgetallen $3, 5, 7, 11, \ldots$ geeft dit extra factoren $\dfrac{3}{4}, \dfrac{15}{16}, \dfrac{35}{36}, \dfrac{99}{100}, \ldots$.

Dit suggereert dat we de hierboven gebruikte heuristiek kunnen corrigeren, met een factor 2 voor het even/oneven-effect, en voor ieder oneven priemgetal met een factor $1 - \dfrac{1}{(q-1)^2}$. Dit leidt tot het invoeren van de *priemtweelingconstante*:

$$C_2 = \prod_{q \text{ oneven priem}} \left(1 - \frac{1}{(q-1)^2}\right).$$

Is dit oneindige product convergent? Ja, want je vermenigvuldigt altijd met een getal > 0 en < 1. Maar dan loop je het risico dat er 0 uitkomt. Neem de logarithme om daar achter te komen:

$$\log C_2 = \sum_{q \text{ oneven priem}} \log\left(1 - \frac{1}{(q-1)^2}\right),$$

en omdat $-\log\left(1 - \frac{1}{x}\right) = \log\left(1 + \frac{1}{x-1}\right) < \frac{1}{x-1}$ voor $x \geq 2$, vinden we

$$\begin{aligned} -\log C_2 &< \sum_{q \text{ oneven priem}} \frac{1}{(q-1)^2 - 1} < \sum_{n=3}^{\infty} \frac{1}{(n-1)^2 - 1} \\ &= \sum_{n=3}^{\infty} \frac{1}{n(n-2)} = \frac{1}{2} \sum_{n=3}^{\infty} \left(\frac{1}{n-2} - \frac{1}{n}\right) = \frac{3}{4}. \end{aligned}$$

Dus is $C_2 > e^{-3/4} = 0.47\ldots$. Een precieze berekening geeft

$$C_2 = 0.66016\ldots,$$

en dus is ons vermoeden nu dat

$$\pi_2(x) \sim 2C_2 \cdot \mathrm{li}_2(x) = 1.32032\ldots \cdot \mathrm{li}_2(x).$$

Tabel 3.2 geeft duidelijk ondersteuning voor dit vermoeden. Kennelijk hebben we nu de afhankelijkheden in de kansen goed te pakken. Maar let wel, bewezen is hier niets. Het blijft pure heuristiek.

Frits keek naar de som van de omgekeerde priemgetallen $\sum_{p \text{ priem}} \frac{1}{p}$, en uit de divergentie ervan concludeerde hij dat er dus oneindig veel priemgetallen bestaan. Dat was een net bewijs, maar vanuit de heuristiek is dit ook aannemelijk te maken: omdat het n-e priemgetal ongeveer $n \log n$ zal zijn, kunnen we de som benaderen met $\sum_{n=1}^{\infty} \frac{1}{n \log n}$, en dus met de integraal $\int_2^{\infty} \frac{1}{x \log x} \mathrm{d}x$. Een primitieve van $\frac{1}{x \log x}$ is $\log \log x$, en voor $x \to \infty$ divergeert dit inderdaad.

Net zo kunnen we nu kijken naar de som van de omgekeerde priemtweelingen: $\sum_{p \text{ priem en } p+2 \text{ ook}} \frac{1}{p}$ (of, zo je wilt, $\sum_{p \text{ priem en } p+2 \text{ ook}} \left(\frac{1}{p} + \frac{1}{p+2}\right)$). Maar de heuristiek zegt ons dat nu de n-e priemtweeling bij $n(\log n)^2$ in de buurt zal liggen, en de oneindige som zal zich gedragen als $\sum_{n-1}^{\infty} \frac{1}{n(\log n)^2}$, en

dus als de integraal $\int_2^\infty \frac{1}{x(\log x)^2}\mathrm{d}x$. Maar een primitieve van $\frac{1}{x(\log x)^2}$ is $-\frac{1}{\log x}$, en deze som zal dus wel convergeren. Dat dit echt zo is is bewezen door Brun. Brun's constante is

$$\sum_{p \text{ priem en } p+2 \text{ ook}} \left(\frac{1}{p} + \frac{1}{p+2}\right) = 1.90216\ldots.$$

We kunnen niet uit dit bewijs van Brun concluderen dat er eindig veel priemtweelingen zijn, want een convergerende som kan eindig maar ook oneindig veel termen hebben. Gelukkig geeft ons vermoeden over $\pi_2(x)$ wel forse steun aan het vermoeden dat er oneindig veel zijn.

Dan noemen we nog even kort de Sophie Germain-priemgetallen. Een priemgetal p is vernoemd naar Sophie Germain (1776-1831, een van de eerste vrouwen die bekend werden in de wiskunde) als ook $2p+1$ een priemgetal is. Deze priemgetallen zijn nuttig in de cryptologie, omdat $\phi(2p+1) = 2p$ weinig delers heeft, en delers van $\phi(q)$ voor een priemgetal q zitten de veiligheid nog wel eens in de weg. Zo'n priem $2p+1$ wordt daarom ook wel een 'veilige priem' genoemd.

Opgave 3.5.1. *Bepaal de eerste* 10 *Sophie Germain-priemgetallen. Als je lui bent mag je Pari gebruiken.*

Dan willen we natuurlijk ook weten hoeveel er van zijn.

Opgave 3.5.2. *We voeren de Sophie Germanpriemtelfunctie $\pi_{SG}(x)$ in als het aantal priemgetallen $p \leq x$ waarvoor ook $2p+1$ priem is. Laat zien dat de volgende heuristiek geldt:*

$$\pi_{SG}(x) \sim 2C_2 \cdot \mathrm{li}_2(x),$$

waar C_2 de priemtweelingconstante is.

3.6 Patronen in opeenvolgende priemgetallen

Je kunt natuurlijk veel verder gaan met varianten bedenken op het thema 'priemtweelingen', zo zijn er de 'priemneefjes' $p, p+4$, de 'sexy priemgetallen' $p, p+6$, of het algemener maken met bijvoorbeeld het Vermoeden van Bunyakovsky (Zie de bijdrage van Frits). Enkele jaren geleden keken Robert Lemke Oliver en Kannan Soundararajan naar patronen van restklassen modulo q die optreden in rijtjes van opeenvolgende priemgetallen. Voor een vector $\mathbf{a} = (a_1, \ldots, a_r)$ van restklassen modulo q (met

alle $\gcd(a_i, q) = 1$) keken ze naar de priemtelfunctie $\pi(x, q, \mathbf{a})$ die het aantal priemgetallen $p \leq x$ telt zodat voor de opeenvolgende priemgetallen $p_1 = p, p_2, \ldots, p_r$ geldt dat $p_i \equiv a_i \pmod q$. Bijvoorbeeld, $\pi(x, 10, (f, g))$ telt het aantal priemgetallen $p \leq x$ zodat p als laatste decimaal f heeft, en het eerstvolgende priemgetal na p als laatste decimaal g heeft. Je zou verwachten dat de boel weer een beetje eerlijk verdeeld is, dat zou dan zijn

$$\pi(x, q, \mathbf{a}) \sim \frac{1}{\phi(q)^r} \operatorname{li}(x),$$

maar tot hun verbazing vonden ze experimenteel sterke afwijkingen, en tot ieders verbazing vonden ze er ook een verklaring voor. Tabel 3.3 geeft een indruk van wat er uit experimenten kwam.

	$g = 1$	$g = 3$	$g = 7$	$g = 9$
$f = 1$	4623042	7429438	7504612	5442345
$f = 3$	6010982	4442562	7043695	7502896
$f = 7$	6373981	6755195	4439355	7431870
$f = 9$	7991431	6372941	6012739	4622916

Tabel 3.3: $\pi(x_0, 10, (f, g))$ met $\pi(x_0) = 10^8 + 2$ voor $f, g \in \{1, 3, 7, 9\}$.

Dat zijn wel erg grote afwijkingen. Hun vermoeden is als volgt:

$$\begin{aligned} \pi(x, q, \mathbf{a}) \quad &\sim \quad \frac{1}{\phi(q)^r} \operatorname{li}(x) \left(1 + c_1(q, \mathbf{a}) \frac{\log\log x}{\log x} + c_2(q, \mathbf{a}) \frac{1}{\log x} \right. \\ &+ \quad \text{een foutterm van ordegrootte } \left. \frac{1}{(\log x)^{7/4}} \right), \end{aligned}$$

waarbij $c_1(q, \mathbf{a})$ en $c_2(q, \mathbf{a})$ alleen van q en $\mathbf{a}$ afhangen. Ze geven expliciete formules voor deze constanten, maar die zijn nogal gecompliceerd. Hoe dan ook, asymptotisch blijft het vermoeden $\pi(x, q, \mathbf{a}) \sim \frac{1}{\phi(q)^r} \operatorname{li}(x)$ fier overeind staan, maar er is wel degelijk meer aan de hand.

3.7 Alda-priemgetallen

Toen ik mijn echtgenote Alda, geen wiskundige, uitlegde wat priemtweelingen zijn, vroeg ze ogenblikkelijk of er ook priemtweelingen bestaan zodanig dat de eerstvolgende twee priemgetallen ook een priemtweeling zijn. Dus definieer ik een *Alda-priemgetal*[1] als een priemgetal p zodanig dat het met

[1] Later kwam ik er achter dat het ook wel een *priemtweelingcluster van orde* 2 heet, maar ik verkies dat te negeren.

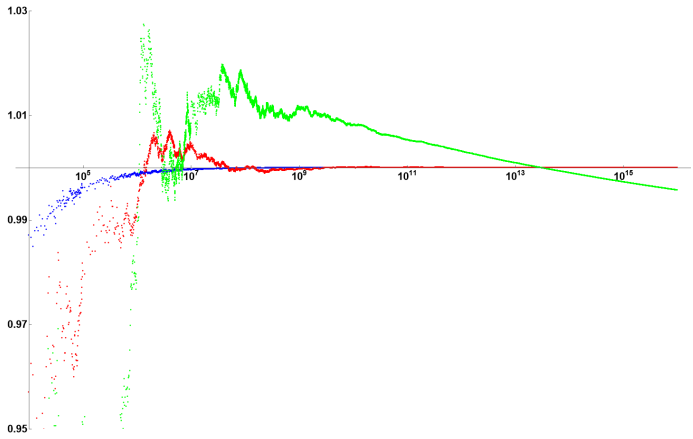
de eerstvolgende drie priemgetallen twee priemtweelingen vormt. De kleinste is 5, en 11 is de enige andere onder de 100, maar 101 is er ook weer eentje.

We definiëren nu uiteraard de Alda-priemgetaltelfunctie $\pi_{\mathrm{A}}(x)$ als het aantal Alda-priemgetallen $\leq x$. Ik legde dit probleem voor aan Pieter Moree, die schakelde zijn collega Efthymios Sofos in, en die kwam met de volgende heuristiek:

$$\pi_{\mathrm{A}}(x) \sim 4C_2^2 \cdot \mathrm{li}_3(x), \text{ waarbij } \mathrm{li}_3(x) = \int_2^x \frac{1}{(\log t)^3}\mathrm{d}t \sim \frac{x}{(\log x)^3}.$$

Opgave 3.7.1. *Kun je deze heuristiek verklaren? Waarom een derdemacht van* $\log x$ *terwijl er toch 4 priemgetallen zijn? Waarom de* C_2 *in het kwadraat?*

Toch is er iets vreemds aan de hand. Ik heb het "high performance cluster" van de TU/e zo'n 2065 uur Alda-priemgetallen laten tellen, en kwam tot $x < 10^{16}$. Deze berekeningen zijn onafhankelijk gecontroleerd dooe Alexander Weisse uit Bonn. Figuur 3.3 laat een vergelijking zien tussen $\dfrac{\pi(x)}{\mathrm{li}(x)}$ (priemgetallen, blauw), $\dfrac{\pi_2(x)}{2C_2 \cdot \mathrm{li}_2(x)}$ (priemtweelingen, rood), en $\dfrac{\pi_{\mathrm{A}}(x)}{4C_2^2 \cdot \mathrm{li}_3(x)}$ (Alda-priemgetallen, groen); alle drie de grafieken zouden naar hoogte 1 moeten convergeren.



Figuur 3.3: Alda-priemgetallen tellen vergeleken met gewone priemgetallen en priemtweelingen.

Hieruit blijkt dat de Alda-priemgetallen zich niet aan de heuristiek lijken

te houden. Of hebben we gewoon nog niet ver genoeg gerekend, en gaat de groene grafiek alsnog omhoog? Later heb ik ook nog een slordige 1000 uur Alda-priemgetallen in intervallen ter grootte 10^{10} en 10^{11} voor x tot aan 10^{100} geteld, en met veel goede wil zie je daar een vage trend de goede kant uit. We hebben nog geen goed idee over wat er hier an de hand is.

Het probleem van Alda-priemgetallen is lastiger dan de priemgetallen die Lemke Oliver en Soundararajan bekeken, omdat nu het gat tussen de twee priemtweelingen niet vastgelegd is.

We kunnen natuurlijk ook *sexy Alda-priemgetallen* bedenken: een viertal priemgetallen $p, p+6, q, q+6$ waarbij er tussen p en $p+6$ wel een priemgetal mag zitten, zo ook tussen q en $q+6$, maar niet tussen $p+6$ en q. Hun telfunctie $\pi_{6\mathrm{A}}(x)$ heeft als heuristiek

$$\pi_{6\mathrm{A}}(x) \sim 16C_2^2 \cdot \mathrm{li}_3(x),$$

en hier vinden we ook de onverklaarde afwijkingen die we bij $\pi_{\mathrm{A}}(x)$ tegenkwamen.

Het is overigens best aardig iets te laten zien van hoe deze tellingen uitgevoerd worden. Allereerst worden priemgetallen in intervallen opgedeeld, van bv. lengte 10^{10}. De priemgetallen worden niet zelf opgeslagen, maar we werken met een positiesysteem: elk getal krijgt een bit op een makkelijk terug te vinden positie toegewezen, en een priemgetal krijgt dan een 1-bit, een samengesteld getal een 0-bit. Omdat computers nu eenmaal met bytes van 8 bits werken, en omdat $n = 30$ het grootste getal is met $\phi(n) = 8$, slaan we alleen informatie op over getallen die relatief priem zijn met 30, dus alleen getallen $1, 7, 11, 13, 17, 19, 23, 29 \pmod{30}$, dan vergeten we alleen de priemgetallen $2, 3, 5$. Dus per 30 opeenvolgende getallen hoeven we slechts 1 byte op te slaan. Voor een interval van 10^{10} getallen kost dat slechts 318 MB.

We beginnen met alle bits op 1 te zetten, en voeren dan gewoon de Zeef van Eratosthenes uit, om alle samengestelde getallen er uit te zeven. Als we zo alle priemgetallen in het interval hebben, gaan we ze allemaal af, met het bijhouden van een toestand (p, g, b), hierbij is p het vorige bezochte priemgetal, g het gat tussen het vorige bezochte priemgetal en het priemgetal waar we nu zijn, en b een bit dat 1 is als we al een priemtweeling hadden gezien, 0 anders. Dan voeren we het volgende uit:

- als $g = 2$ en $b = 0$ dan hebben we een nieuwe priemtweeling gevonden en wordt $b = 1$,
- anders: als $g = 2$ en $b = 1$ dan hebben we een nieuw Alda-priemgetal gevonden, en hogen we de teller met 1 op,
- anders: als $g > 2$ en $b = 0$ dan doen we niets,

- anders: als $g > 2$ en $b = 1$ dan leidt de priemtweeling die we al gezien hadden niet tot een Alda-priemgetal, en wordt $b = 0$,
- als $g = 2$ ga naar het volgende priemgetal, en pas p en g aan,
- ga naar het volgende priemgetal, en pas p en g aan.

Aan het einde van een interval moet je even goed opletten dat je de toestand opslaat en doorgeeft als begintoestand voor het nieuwe interval. Een Alda-priemgetal kan immers net op de rand van een interval zitten.

3.8 Om het af te leren

We sluiten af met een aantal opgaven. Je mag in deze opgaven best een beetje grof redeneren, bv. een makkelijk gemiddelde nemen van een $\log x$ over een interval (bv. op $[1000, 2000]$ stijgt $\log x$ van ongeveer 6.9 naar ongeveer 7.6, doe dan net alsof dat over het hele interval ongeveer 7.3 is).

Opgave 3.8.1. *Een palindroomgetal is een getal dat van achter naar voren hetzelfde getal is, zoals* 123454321*. Geef een heuristiek voor het aantal palindroom-priemgetallen van n cijfers.*

Opgave 3.8.2. *Geef een heuristiek voor het aantal priemgetallen van n cijfers waar het cijfer* 3 *niet in voorkomt.*

Opgave 3.8.3. *Een Mersenne-priemgetal is een priemgetal van de vorm $2^n - 1$. Het is eenvoudig in te zien dat dat alleen kan als n zelf een priemgetal is. Denk je dat er eindig of oneindig veel Mersenne-priemgetallen zijn?*

Opgave 3.8.4. *Een Fermat-priemgetal is een priemgetal van de vorm $2^n + 1$. Het is eenvoudig in te zien dat dat alleen kan als $n = 2^k$. Denk je dat er eindig of oneindig veel Fermat-priemgetallen zijn?*

Opgave 3.8.5. *Geef een heuristiek voor het gemiddelde aantal manieren waarop je een even getal n kunt schrijven als de som van twee priemgetallen. Zie wat Frits schrijft over het Goldbach-vermoeden.*

Opgave 3.8.6. *Geef een heuristiek voor het aantal priemgetallen $\leq x$ die van de vorm $n^2 + 1$ zijn (zie de bijdrage van Frits over het Vermoeden van Bunyakovsky). Maak je geen zorgen over de constante, alleen over de juiste grootteorde.*

Pari opgave 3.8.7. *Maak een priemgetal dat begint met je eigen telefoonnummer. Bedenk van te voren hoeveel cijfers je verwacht er aan toe te moeten voegen.*

Literatuur

Paul Levrie en Rudi Penne – De pracht van priemgetallen, Prometheus / Bert Bakker, 2014.

Richard Crandall and Carl Pomerance – Prime Numbers, A Computational Perspective, Springer, 2nd. Ed., 2005.

Jean-Marie De Koninck and Nicolas Doyon – The Life of Primes in 37 Episodes, AM. Math. Soc., 2020.

Voor wie is PWN interessant?

Beroepswiskundigen
Wiskundeleraren
Bedrijven
Leerlingen en studenten
Breed publiek

Platform Wiskunde Nederland is hét landelijke loket voor alles wat met wiskunde te maken heeft.

PWN behartigt de belangen van, en fungeert als spreekbuis voor, de gehele Nederlandse wiskunde.

Platform Wiskunde Nederland | Science Park 123 | kamer L013 | 1098 XG Amsterdam | 020 592 40 06

Ga voor meer informatie naar:
www.platformwiskunde.nl

platform
wiskunde nederland