Illustratie: Ryu Tajiri

Frits Beukers
Mathematisch Instituut
Universiteit Utrecht
f.beukers@uu.nl

Vakantiecursus

Hoe bewijs je het priemtweelingvermoeden?

In het voorjaar van 2013 overrompelde de Chinees-Amerikaanse wiskundige Yitang Zhang de wiskundige wereld met de aankondiging van het bewijs dat er een geheel getal A bestaat zó dat er oneindig veel paren gehele getallen $n, n+A$ bestaan die beide priem zijn. Als A gelijk aan 2 zou zijn geweest dan zou dit een bewijs voor het beruchte priemtweelingvermoeden zijn geweest. Dit vermoeden is waarschijnlijk even oud als de getaltheorie zelf. Uit het bewijs van Zhang volgt iets zwakkers, namelijk dat A kleiner is dan zeventig miljoen. Niettemin was het bewijs van Zhang uniek in zijn soort en tot dusver voor onmogelijk gehouden door andere experts in de analytische getaltheorie. In dit artikel geeft Frits Beukers een overzicht van een aantal feiten over priemtweelingen en de ontwikkelingen die leidden tot Zhangs ontdekking en het vervolg erop. Door inspanningen van veel wiskundigen is in de loop van 2013–15 de grens voor A teruggebracht tot 246. In dit artikel zal ook geprobeerd worden het idee van Zhangs bewijs over te brengen. Of dit lukt is nog onzeker, het bewijs is niet helemaal triviaal.

Een geheel getal >1 dat alleen zichzelf en 1 als deler heeft noemen we een priemgetal. De rij priemgetallen begint als volgt:

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

Bij het opschrijven van deze rij komt meteen al een stel vragen naar voren. Bijvoorbeeld, gaat deze rij oneindig lang door? Waarschijnlijk wel, maar we moeten er wel een bewijs voor geven. Dat zullen we in dit artikel op twee manieren doen, namelijk een bewijs afkomstig van Euclides en een ander van Euler. Nog een vraag die rijst. Gegeven X, hoeveel priemgetallen $\le X$ zijn er? Kun je daar een eenvoudige formule voor geven? Dat is wat lastiger te beantwoorden, maar het is in elk geval mogelijk formules te geven die deze aantallen bij grove benadering geven. Dit wordt mogelijk gemaakt door de zogenaamde Priemgetalstelling uit 1899. Nog een vraag: zijn er oneindig veel priemgetallen die op een 1 eindigen? Of een 7? Het antwoord daarop is 'ja' en wordt gegeven door een stelling van Dirichlet uit 1825, die zegt dat elke rij van de vorm

$$a, a+q, a+2q, a+3q, a+4q, \ldots$$

met $a, q \in \mathbb{N}$ en $\mathrm{ggd}(a,q)=1$, inderdaad oneindig veel priemgetallen bevat. In het bijzonder geldt dit voor de rij $10n+1$ met $n=1,2,3,\ldots$.

Maar zoals iedereen weet zijn er ook onopgeloste problemen. Bijvoorbeeld het *Goldbachvermoeden* dat zegt dat elk geheel even getal ≥ 4 te schrijven is als som van twee priemgetallen. In de novelle van Dioxiades, *Uncle Petros and the Goldbach Conjecture*, kunnen we lezen hoe iemand tot wanhoop kan worden gedreven door de mislukte pogingen een dergelijk vermoeden aan te tonen. Empirisch bewijsmateriaal is er in overvloed. Het getal 100 is bijvoorbeeld op 6 manieren som van twee priemgetallen en het getal 10000 op 165 manieren. De kunst is echter aan te tonen dat elk even getal op minstens één manier een som van twee priemgetallen is. Dat is tot nog toe niemand gelukt.

Een ander notoir probleem is het *priemtweelingvermoeden*. Merk namelijk op dat

$$(5,7),(11,13),(17,19),(29,31),(41,43),\\ (59,61),(71,73),(101,103),\ldots$$

paren opeenvolgende oneven getallen zijn waarvan beide elementen priem zijn. De vraag is of er oneindig veel van bestaan. Waarschijnlijk wel, maar dat is tot dusver nog niet aangetoond. Men heeft wel heel grote priemtweelingen gevonden, bijvoorbeeld

$$3756801695685 \cdot 2^{666669} \pm 1.$$

De enig bekende technieken om dit soort problemen aan te pakken is via de analytische getaltheorie, in het bijzonder de *cirkelmethode* en *zeefmethoden*. Dit zijn lastige technieken die heel veel gebruik maken van slimme afschattingen en resultaten uit de Fouriertheorie en complexe functietheorie. De vooruitgang in concrete resultaten is over het algemeen heel langzaam. Maar soms zijn er verrassingen. Zo lieten bijvoorbeeld Terence Tao en Ben Green in 2008 zien dat er in de rij priemgetallen eindige rekenkundige rijen kunnen voorkomen van elke gekozen lengte. De tot nu toe (februari 2015) langst bekende rekenkundige rij heeft lengte 26 en wordt gegeven door

$$161004359399459161 + 47715109 \cdot 223092870 \cdot n$$

voor $n = 0, 1 \ldots, 25.$

Een andere spectaculaire ontwikkeling is nog recenter. In het voorjaar van 2013 kondigde de Chinees-Amerikaanse wiskunde Yitang Zhang aan dat er een getal A bestaat zó dat er oneindig veel paren priemgetallen van de vorm $n, n+A$ zijn. Als A gelijk zou zijn aan 2, zou dit een bewijs van het priemtweelingvermoeden hebben betekend. Zhang kon alleen maar aantonen dat $A < 70.000.000$. Maar zelfs dit resultaat is uniek in zijn soort en eigenlijk door niemand voorzien. Het artikel, dat Zhang bij het prestigieuze *Annals of Mathematics* indiende, werd al vrij snel door de referees goedgekeurd en gepubliceerd.

Het verhaal rond deze doorbraak is een bijzondere. Wiskundige onderzoekers worden geacht hun meest creatieve werk rond hun 30ste te produceren. Zhang was 58 jaar oud en bovendien had hij tot dan toe weinig gepubliceerd op wiskundegebied. Op Wikipedia vind je een uitgebreide beschrijving van de levensloop van Zhang. Een interview kun je vinden in een artikel uit de *New Yorker* [4].

De grens van 70 miljoen in Zhangs stelling is een gevolg van de methode die hij gebruikte. Het was toen al duidelijk dat met een wat voorzichtiger benadering deze waarde wel kleiner gemaakt zou kunnen worden. En dat is precies wat gebeurde. Een paar weken na Zhangs publicatie stelde Terence Tao een zogenaamd Polymathproject voor. Een publiek project via internet waar in principe iedereen aan kon bijdragen. Het eerste deel van het project werd afgesloten met een grens van 4680. Tijdens dit project kwam de jonge postdoc James Maynard met een iets andere aanpak die eenvoudiger bleek te zijn en ook nog eens betere grenzen opleverde. Dit werd deel twee van het project dat uiteindelijk afgesloten werd met een grens van 246. Voor een overzicht van deze historie en de huidige stand van zaken, zie [5].

Zoals het er nu naar uitziet zullen er nieuwe ideeën nodig zijn om de grens 2 van het priemtweelingvermoeden te bereiken.

Het nadeel van veel recente opzienbarende ontwikkelingen in de wiskunde is dat de details vaak te ingewikkeld en specialistisch zijn voor een breder publiek van niet-ingewijden. Het aantal wiskundigen dat bijvoorbeeld de details van het bewijs van de laatste stelling van Fermat kent is zeer beperkt. De reden is dat je eerst enkele jaren moet wijden aan het bestuderen van onderwerpen als modulaire vormen, galoisrepresentaties, et cetera, die de basis vormen van het bewijs. Het prettige van Maynards bewijs is dat, als je bereid bent een 'black box' uit 1965 te accepteren, de afleiding relatief weinig voorkennis vereist, maar wel een flinke portie technische vaardigheid en uithoudingsvermogen. In principe zou een goede derdejaars wiskundestudent het bewijs moeten kunnen volgen, met eventueel wat begeleiding. En dit is inderdaad wat er gebeurd is, Ilja Nelen, vorig jaar nog bachelorstudent, heeft Maynards bewijs als onderwerp voor zijn bachelorscriptie gehad.

Gesterkt door deze ervaring zal ik aan het eind van deze tekst heel kort een indruk geven van de ingrediënten van het bewijs van Maynards resultaat. Maar voor het zover is geven we een inleiding in wat we weten van priemtweelingen, hetgeen hopelijk wat lichtere kost vormt.

Priemgetallen tellen

In deze paragraaf kijken we naar de oneindigheid van de verzameling priemgetallen en hoe vaak ze voorkomen onder de natuurlijke getallen (dat zijn gehele getallen ≥ 1).

Stelling 1 (Euclides). *Bij elk natuurlijk getal N bestaat er een priemgetal p dat groter is dan N.*

Hieruit volgt meteen dat er oneindig veel priemgetallen zijn. Je kunt bijvoorbeeld voor N de getallen $10, 100, 1000, \ldots$ kiezen.

Bewijs. Kijk naar het getal $N! + 1$, waarin $N!$ staat voor het product van alle getallen van 1 tot en met N. Deze heeft een priemdeler. Noem deze p. Stel dat $p \leq N$. Dan is p ook een deler van $N!$ Maar een priemgetal kan niet twee opeenvolgende getallen delen. Dus $p \leq N$ kan niet gelden. Conclusie: $p > N$. □

Hier is een wat sterker resultaat.

Stelling 2 (Euler). *De oneindige reeks*

$$\sum_{p \text{ priem}} \frac{1}{p}$$

is divergent. Dat wil zeggen, als je de inverse priemgetallen $\frac{1}{2}, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \frac{1}{11}, \ldots$ achtereenvolgens bij elkaar optelt, dan gaat de rij gevonden deelsommen naar oneindig toe.

Uiteraard volgt hieruit dat er oneindig veel priemgetallen bestaan. Het bewijs hiervan is wat lastiger. Het is misschien wel bekend dat de som van de inverse natuurlijke getallen divergent is. Dus

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = \infty.$$

We noemen dit de harmonische reeks. Om de divergentie in te zien breken we de reeks op in stukjes waarbij de som van elk stukje $> \frac{1}{2}$ is, als volgt:

$$\begin{aligned} \frac{1}{3} + \frac{1}{4} \qquad &> 2 \times \frac{1}{4} = \frac{1}{2}, \\ \frac{1}{5} + \cdots + \frac{1}{8} &> 4 \times \frac{1}{8} = \frac{1}{2}, \\ \frac{1}{9} + \cdots + \frac{1}{16} &> 8 \times \frac{1}{16} = \frac{1}{2}, \end{aligned}$$

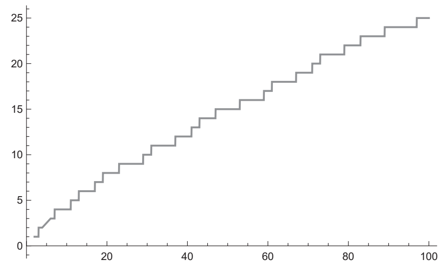
et cetera. De som van deze waarden $\frac{1}{2}$ gaat naar oneindig. Dit bewijs werd al in de veertiende eeuw gevonden door Nicole d'Oresme.

We bewijzen nu Eulers stelling.

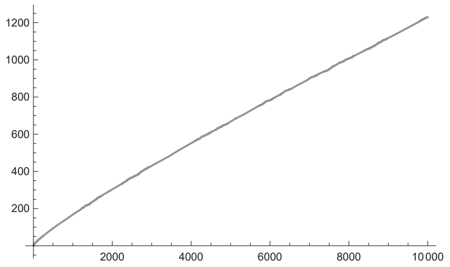
Bewijs. We kiezen N en merken allereerst op dat

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

waarin het product over de priemgetallen $p \leq N$ loopt. Bij het wegwerken van de haakjes uit het product moet je uit elke factor een term nemen en vervolgens die termen vermenigvuldigen. Als je bijvoorbeeld de termen $\frac{1}{2}, \frac{1}{3}$ en 1 voor de andere factoren neemt, krijg je $\frac{1}{6}$. Dit wordt precies de term $\frac{1}{6}$ rechts. Op deze manier zie je dat het product links alle inversen van gehele getallen geeft, die samengesteld zijn met priemfactoren $\leq N$. Dat zijn in ieder geval 1 tot en met N. Vandaar het ongelijkheidsteken.



Figuur 1



Figuur 2

Beschouw nu de volgende ongelijkheid

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

Deze volgt door in elke factor de ongelijkheid $(1+x)^2 > 1 + x + x^2 + x^3 + \cdots$ voor $x \le \frac{1}{2}$ te gebruiken. Dit is een kleine opgave, die laat ik aan de lezer over. Nu nog een derde ongelijkheid,

$$\sum_{p \le N} \frac{1}{p} > \sum_{p \le N} \log\left(1 + \frac{1}{p}\right) = \log \prod_{p \le N} \left(1 + \frac{1}{p}\right). \quad (3)$$

Deze volgt uit het feit dat we in elke term $x > \log(1+x)$ kunnen gebruiken voor $x > 0$ (NB: de $\log$ hier is de $\ln$ op school).

In de ongelijkheid (1) staat rechts de harmonische reeks. Die gaat naar oneindig als $N \to \infty$. Dus, de linkerzijde van ongelijkheid (1),

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

gaat ook naar oneindig als $N \to \infty$. Uit de ongelijkheid (2) concluderen we vervolgens dat $\prod_{p \le N}(1 + \frac{1}{p})$ naar oneindig gaat als $N \to \infty$. Tenslotte, met ongelijkheid (3) concluderen we dat $\sum_{p \le N} \frac{1}{p}$ naar oneindig gaat als $N \to \infty$. □

We gaan nu priemgetallen tellen en definiëren de zogenaamde priemtelfunctie

$$\pi(x) = \#\{p \le x \mid p \text{ priem}\},$$

waarin we 'aantal' met het symbool # aangeven. Figuur 1 toont een grafiek van $\pi(x)$ voor $x \le 100$. Hierin zien we duidelijk de stapjes die gemaakt worden als we met x een priemgetal passeren. Deze onregelmatigheden verdwijnen (op het oog) als we een wat groter domein voor x kiezen, zie Figuur 2. De grafiek is nu een mooie vloeiende figuur geworden en men kan zich afvragen of er een eenvoudige formule bestaat om deze grafiek (bij benadering) te beschrijven. Hier is veel over gespeculeerd, maar het was Gauss die een duidelijk heuristisch argument gaf. Kies X heel groot en Δ klein ten opzichte daarvan. Het principe van Gauss zegt dat het aantal priemgetallen in het interval $[X, X+\Delta]$ ongeveer gelijk is aan $\Delta/\log X$. Met andere woorden, de gemiddelde dichtheid van de priemgetallen in de buurt van X is $1/\log X$. Ter ondersteuning hiervan berekende Gauss een aantal tabellen zoals Tabel 1. Daarin is $\Delta = 10^5$ genomen en met $s(x)$ geven we het aantal priemgetallen in $[x, x+10^5]$ aan. Merk op dat de getallen in de tweede en derde kolom meestal in hun eerste twee cijfers overeenkomen. Als we aannemen dat de dichtheid van de priemgetallen in de buurt van X gelijk is aan $1/\log X$, dan is het aantal priemgetallen $\le x$ gelijk aan de som van alle dichtheden van, zeg, 2 tot en met X en dus

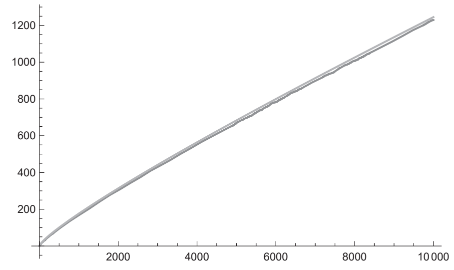
$$\mathrm{li}(x) = \int_2^x \frac{dt}{\log t}.$$

In Figuur 3 wordt een plot getoond van de functies $\pi(x)$ en $\mathrm{li}(x)$ samen. In deze plot geldt dat $\pi(x) < \mathrm{li}(x)$, maar dat is niet altijd zo. Het blijkt namelijk dat $\pi(x) - \mathrm{li}(x)$ oneindig vaak van teken wisselt. Voor welke x dat voor het eerst gebeurt weet men echter niet.

Dat $\pi(x)$ en $\mathrm{li}(x)$ inderdaad dicht bij elkaar liggen wordt bevestigd door de *priemgetalstelling*:

x	$s(x)$	$10^5/\log(x)$
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



Figuur 3

Stelling 3 (Hadamard, De la Vallée-Poussin, 1899). *De verhouding van $\pi(x)$ en $\mathrm{li}(x)$ gaat naar 1 als x naar oneindig gaat.*

Het blijkt dat $\mathrm{li}(x)$ redelijk in de buurt ligt van $x/\log x$. Het is een kleine opgave om aan te tonen dat hun verhouding naar 1 gaat als $x \to \infty$.

Priemtweelingen tellen

We voeren nu een telfunctie voor priemtweelingen in:

$$\pi_2(x) := \#\{n \le x \mid n \text{ en } n+2 \text{ priem}\}.$$

Het priemtweelingvermoeden komt dus neer op $\pi_2(x) \to \infty$ als $x \to \infty$.

Om een grof idee van $\pi_2(x)$ te krijgen herhalen we het idee van Gauss. Kies X heel groot en Δ iets kleiner. We vragen ons af hoeveel priemtweelingen er in het interval $[X, X+\Delta]$ voorkomen. We laten daartoe n van X tot $X+\Delta-1$ lopen en bepalen de kans dat zowel n als $n+2$ priem zijn. Als deze kansen onafhankelijk zouden zijn is dat ongeveer $1/(\log X)^2$ waarmee het aantal priemparen op ongeveer $\Delta/(\log X)^2$ komt. Helaas is de kans dat $n+2$ priem is enigszins afhankelijk van de waarde van n. Als bijvoorbeeld toevallig n even is dan is $n+2$ zeker geen priem.

We kunnen dit omzeilen door alleen naar de oneven getallen in $[X, X+\Delta]$ te kijken. Het aantal priemgetallen in de verzameling is nog steeds ongeveer $\Delta/\log X$. Het aantal oneven getallen $\Delta/2$. De 'kans' dat een oneven getal in het interval priem is, is dus $2/\log X$. De 'kans' dat zowel n als $n+2$ voor oneven n priem zijn wordt $4/(\log X)^2$. Maal het aantal oneven getallen $\Delta/2$ geeft dit de verwachting van $2\Delta/(\log X)^2$ priemtweelingen in $[X, X+\Delta]$. Dit is tweemaal zo groot als wat we eerst hadden als verwachting. Een dergelijke correctie kunnen we voor alle kleine priemgetallen p uitvoeren en we vinden dat het verwachte aantal priemtweelingen tussen X en $X+\Delta$ gelijk is aan

$$C_2 \frac{\Delta}{(\log X)^2}$$

met

$$C_2 = 2 \times \prod_{\substack{p \text{ oneven} \\ \text{priem}}} \frac{(1-2/p)}{(1-1/p)^2} \approx 1.32.$$

Dit geeft aanleiding tot het volgende vermoeden.

Vermoeden 1 (Hardy, Littlewood, 1923), *Definieer*

$$\mathrm{li}_2(x) := C_2 \int_2^x \frac{dt}{(\log t)^2}.$$

Dan gaat de verhouding van $\mathrm{li}_2(x)$ *en* $\pi_2(x)$ *naar* 1 *als* $x \to \infty$.

Om in te zien hoe goed deze benadering is geven we een tabel (zie Tabel 2). Hierin is $s_2(x)$ gelijk aan het aantal priemtweelingen in $x, x+10^5$. Dit lijkt goed te gaan. Helaas is er echter weinig bewezen. Met behulp van *zeefmethoden* heeft men bewezen dat er een constante $C > 0$ bestaat zó dat $\pi_2(x) < C \cdot \mathrm{li}_2(x)$ geldt als x groot genoeg is. Een gevolg van deze bovengrens is de volgende stelling.

Stelling 4 (V. Brun, 1919). *De oneindige som*

$$\sum_{\substack{p \text{ en } p+2 \\ \text{priem}}} \frac{1}{p}$$

convergeert.

Dit is het eerste niet-triviale resultaat op het gebied van priemtweelingen. Voor we verder gaan noem ik nog een ander resultaat. Een getal n heet *bijna priem* als het ofwel priem is, ofwel een product van twee priemgetallen.

x	$s_2(x)$	$C_2 \times 10^5 / (\log x)^2$
10^8	377	389
10^9	341	307
10^{10}	262	249
10^{11}	182	205
10^{12}	171	172
10^{13}	142	147
10^{14}	123	127
10^{15}	117	110

Tabel 2

Links Viggo Brun (1885–1978) en rechts een postzegel ter ere van Jing-Run Chen (1933–1996)

Stelling 5 (Jing-Run Chen, 1973). *Er bestaan oneindig veel priemgetallen* p *zó dat* $p+2$ *een bijna priem is.*

Dit was destijds een opzienbarend resultaat en de Chinese posterijen hebben een postzegel aan het werk van Chen gewijd (zie Figuur 4).

Generalisaties

In plaats van priemtweelingen kunnen we ook naar andere paren kijken. Paren priemgetallen van de vorm

- $n, n+2$ heten priemtweelingen
- $n, n+4$ heten priemneven/nichten
- $n, n+6$ heten sexy paren, et cetera

Ook kunnen we drietallen nemen. Bijvoorbeeld $n, n+2, n+4$. Het is een eenvoudige oefening om te zien dat deze drie getallen alleen priem kunnen zijn als $n=3$. Voor $n>3$ is minstens één van de drie getallen deelbaar door 3.

Kies nu een willekeurige rij van k gehele getallen $0 = a_1 < a_2 < \cdots < a_k$. We beschouwen k-tupels $n+a_1, n+a_2, \ldots, n+a_k$ en vragen ons af of er oneindig veel n bestaan zó dat elk van deze getallen priem is.

We noemen een priemgetal p een *obstructie* voor rijen van het type $n+a_1, \ldots, n+a_k$ als voor elke gehele n minstens één van de getallen $n+a_1, n+a_2, \ldots, n+a_k$ deelbaar is door p. We zagen net dat $p=3$ een obstructie is voor $n, n+2, n+4$. In het algemeen, als er een obstructie p is, dan moet gelden dat $p \le a_k + 1 - a_1$.

We noemen de rij $a_1 < a_2 < \cdots < a_k$ *toelaatbaar* als rijen van het type $n+a_1, \ldots, n+a_k$ geen enkele obstructie hebben. Als bijvoorbeeld $p_1 = 2, p_2 = 3, \ldots, p_k$ de eerste k priemgetallen zijn, dan is het niet lastig te zien dat $p_1, \ldots, p_k$ een toelaatbare rij vormt. Er komen dus toelaatbare rijen van willekeurige lengte voor. Hier is een generalisatie van het priemtweelingvermoeden:

Vermoeden 2. *Stel dat* $a_1 < a_2 < \cdots < a_k$ *toelaatbaar is. Dan zijn er oneindig veel* n *zó dat* $n+a_1, n+a_2, \ldots, n+a_k$ *allen priem zijn.*

Net als bij de priemtweelingen is er ook een vermoeden over de aantallen van dergelijke k-tupels onder een gegeven x, het zogenaamde k-tupel vermoeden van Hardy en Littlewood. En net als bij de priemtweelingen zijn we nog heel ver verwijderd van enige vorm van bewijs. Wel speelt het gebruik van toelaatbare rijen een belangrijke rol in het werk van Zhang en Maynard.

Gaten vullen

We gaan langzamerhand naar de vraag toe hoe klein de afstand tussen opeenvolgende priemgetallen kan zijn. Daartoe noteren we de rij priemgetallen met $p_1 < p_2 < p_3 < \cdots < p_n < \cdots$ In het bijzonder, $p_1 = 2, p_2 = 3$, et cetera.

Vraag: Kies $c > 0$. Kunnen we aantonen dat er oneindig veel n zijn zó dat

$$p_{n+1} - p_n < c \cdot \log p_n?$$

Om te illustreren hoe lastig deze vraag is, geven we een chronologisch overzicht van de bewezen resultaten op dit gebied.

1. Als $c > 1$ dan is het antwoord 'ja'. Dit is een gevolg van de priemgetalstelling. Dit is een opgave voor de lezer.
2. Ja, een dergelijke c bestaat met $c \le 1$ (Erdős, 1940). Dit is geen opgave voor de lezer meer, maar een diep resultaat.
3. Ja, voor elke $c \ge \frac{1}{2}$ (Bombieri–Davenport, 1966).
4. Ja, voor elke $c \ge \frac{1}{4}$ (Maier, 1988).
5. Ja, voor elke $c > 0$ (Goldston–Yilderim–Pintz, 2005).

We zien dat de ontwikkeling heel langzaam gaat. Het laatste resultaat van Goldston, Yilderim en Pintz uit 2005 was zelfs een onverwachte doorbraak. Er geldt nog iets meer.

Stelling 6 (GYP, 2005). *Voor elke* $\epsilon > 0$ *heeft de ongelijkheid*

$$p_{n+1} - p_n < (\log p_n)^{1/2+\epsilon}$$

oneindig veel oplossingen.

Na het bekend worden van deze stelling had men enige tijd hoop dat er een mogelijk bewijs richting priemtweelingen in zat. Helaas, na een aanvankelijk opti-


Foto: John D. and Catherine T. MacArthur Foundaation

Yitang Zhang

misme werd het weer stil rond deze kwestie. Totdat in 2013 een volgende doorbraak plaatsvond.

Stelling 7 (Yitang Zhang, 2013). *Er zijn oneindig veel* n zó dat

$$p_{n+1} - p_n < 70.000.000.$$

Zoals in de inleiding gezegd, werd dit resultaat al snel vereenvoudigd en verbeterd door James Maynard. In de laatste paragraaf geven we een overzicht van de stappen van dit alternatieve bewijs.

Het bewijs van Maynard

De techniek is gebaseerd op telresultaten van priemgetallen in een rekenkundige rij. In de inleiding hebben we de stelling van Dirichlet genoemd, die zegt dat als $\mathrm{ggd}(a,q)=1$, dan bevat de rij $a, a+q, a+2q, a+3q, \ldots$ oneindig veel priemgetallen. Het aantal priemgetallen $\leq x$ in deze rij geven we aan met $\pi(x,a,q)$. De verwachting is dat alle restklassen a modulo q die in aanmerking komen ongeveer evenveel priemgetallen zullen bevatten. Gegeven q, dan geven we het aantal getallen $0 < a \leq q$ met $\mathrm{ggd}(a,q)=1$ aan met $\phi(q)$, de Euler ϕ-functie. Bijvoorbeeld $\phi(4)=2, \phi(5)=4, \phi(6)=2, \ldots$ We verwachten dat voor gegeven q de $\phi(q)$ functies $\pi(x,a,q)$ allen ongeveer even snel met x zullen groeien. Dit werd voor het eerst kwantitatief gemaakt in de volgende stelling.

Stelling 8 (Siegel–Walfisz, 1936). *Bij elke* $A > 1$ *bestaat een* $C > 0$ *zó dat voor alle* x,

$$\left|\pi(x,a,q) - \frac{\pi(x)}{\phi(q)}\right| < C\frac{x}{(\log x)^A}.$$

Omdat $\pi(x)$ ongeveer gelijk is aan $x/\log x$ zien we in het bijzonder dat de verhouding tussen $\pi(x,a,q)$ en $\pi(x)/\phi(q)$ naar 1 gaat als $x \to \infty$. Met andere woorden, de priemgetallen zijn ongeveer gelijk verdeeld over de restklassen $a \bmod q$.

Het blijkt dat voor het merendeel van de x-waarden het verschil tussen $\pi(x,a,q)$ en $\pi(x)/\phi(q)$ veel kleiner is dan dat gegeven door de stelling van Siegel–Walfisz. Dit wordt geaccentueerd door de volgende stelling die simpelweg een heleboel verschillen by elkaar optelt als we q over een groot bereik laten lopen.

Stelling 9 (Bombieri–Vinogradov, 1965). *Bij elke* $A > 1$ *en elke* $\theta < \frac{1}{2}$ *bestaat er een* $C > 0$ *zó dat voor elke* x,

$$\sum_{q \leq x^{\theta}} \max_{\mathrm{ggd}(a,q)=1}\left|\pi(x,a,q) - \frac{\pi(x)}{\phi(q)}\right| < C\frac{x}{(\log x)^A}.$$

Dit is een stelling die met de zogenaamde *grote zeef* bewezen kan worden, zie bijvoorbeeld het boek van Friedlander en Iwaniec [1]. Het is het uitgangspunt voor het werk van Goldston, Yilderim en Pintz. Om hun werk te verbeteren richting priemtweelingen zou je eigenlijk een sterkere versie moeten hebben van Bombieri–Vinogradov met een waarde van $\theta > \frac{1}{2}$. Ondanks alle pogingen in die richting is het nog niemand gelukt een dergelijk resultaat te bewijzen. De bijdrage van Zhang bestaat eruit dat hij een iets andere, verfijnde versie van Bombieri–Vinogradov ontwierp waarbij $\theta > \frac{1}{2}$ is. Dit was een onverwachte wending die niemand had voorzien. Het bewijs van Zhang is behoorlijk technisch en niet makkelijk uit te leggen. Het is daarom prettig dat Maynard met een heel ander idee kwam waarbij het gebruik van de stelling van Bombieri–Vinogradov volstaat. We kunnen deze stelling nu als een soort 'black box' gebruiken in de volgende schets van Maynards aanpak. Tussen haakjes, het bewijs van de stelling van Bombieri–Vinogradov is van dezelfde moeilijkheidsgraad als het vervolg van Maynard.

Met $\chi_{\mathbb{P}}(n)$ geven we de karakteristieke functie aan van de priemgetallen. Dat wil zeggen, $\chi_{\mathbb{P}}(n) = 1$ als n priem is en 0 indien niet. Zij $a_1, \ldots, a_k$ een toelaatbare rij en kies $\rho > 1$. Stel dat we kunnen bewijzen dat

$$\sum_{N \leq n \leq 2N} \sum_{i=1}^{k} \chi_{\mathbb{P}}(n+a_i) > \rho N$$

voor grote N. Dat betekent dat er een $n \in [N, 2N]$ bestaat zó dat $\sum_{i=1}^{k} \chi_{\mathbb{P}}(n+a_i) > \rho$. En dus bevinden zich onder de getallen $n+a_1, \ldots, n+a_k$ minstens $\lfloor \rho + 1 \rfloor$ priemgetallen. In het bijzonder, als $\rho > 1$ hebben we minstens twee priemgetallen in het interval $[n+a_1, \ldots, n+a_k]$ te pakken. Door N steeds tweemaal zo groot te kiezen vinden we een oneindige rij van dergelijke n.

Helaas laat bovenstaande ongelijkheid zich lastig bewijzen. Gelukkig brengt de zeeftheorie hier uitkomst door gewogen sommen te nemen. Voor een geschikt gekozen gewichtsfunctie $w(n) \geq 0$ (later te kiezen) bekijken we de sommen,

$$S_1(N) = \sum_{N \leq n \leq 2N} w(n),$$

$$S_2(N) = \sum_{N \leq n \leq 2N} w(n) \sum_{i=1}^{k} \chi_{\mathbb{P}}(n+a_i).$$

Hierin is $S_1(N)$ de (positieve) som van de gewichten. Als we erin slagen om aan te tonen dat voor voldoend grote N geldt $S_2(N) > \rho S_1(N)$, dan zijn we om dezelfde reden als boven klaar.

We gaan nu geschikte gewichten $w(n)$ kiezen. Goldston, Yilderim en Pintz kozen $\delta > 0, R = N^{1/2-\delta}$ en vervolgens

$$w(n) = \left(\sum_{d \mid (n+a_1)\cdots(n+a_k), d \le R} \lambda(d) \right)^2$$

waarin

$$\lambda(d) := \mu(d) \left(\frac{\log d}{\log R} \right)^{k+l}$$

voor een geschikte l. Hier is $\mu(d)$ de zogenaamde Möbius μ-functie die gedefinieerd wordt door $\mu(d) = (-1)^t$ als $d = p_1 \dots p_t$ met alle p_i verschillend en 0 als d deelbaar is door een kwadraat > 1.

Zoals we reeds gemeld hebben, leek deze keuze van $w(n)$ zeer effectief in 2005, maar niet effectief genoeg richting priemtweelingen. James Maynard bedacht een andere gewichtsfunctie, namelijk

$$w(n) = \left(\sum_{d_1 \mid (n+a_1), \dots, d_k \mid (n+a_k), d_1 \cdots d_k \le R} \lambda(d_1, \dots, d_k) \right)^2$$

met $R = N^{1/2-\delta}$ en

$$\lambda(d_1, \dots, d_k) := \prod_{i=1}^{k} \mu(d_i) d_i \times \sum_{\substack{r_1, \dots, r_k \\ d_i \mid r_i}} \prod_{i=1}^{k} \frac{1}{\phi(r_i)} F\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R} \right)$$

voor een slim gekozen functie $F(x_1, \dots, x_k)$ op het domein in $\mathbb{R}^k$ gegeven door de ongelijkheden $x_i \ge 0$ voor $i = 1, \dots, k$ en $x_1 + \cdots + x_k \le 1$.

We willen nu graag een $\rho > 1$ vinden zó dat $S_1(N)/S_2(N) > \rho$ als N groot genoeg is. Door gebruik te maken van de gewichtsfunctie van Maynard en de stelling van Bombieri–Vinogradov kan men aantonen dat

$$\frac{S_1(N)}{S_2(N)} = \left(\frac{1}{2} - \delta + o(1) \right) \frac{\sum_{m=1}^{k} J_k^{(m)}(F)}{I_k(F)}.$$

waarin

$$I_k(F) = \int_0^1 \cdots \int_0^1 (F(t_1, \dots, t_k))^2 \, dt_1 \cdots dt_k$$

en

$$J_k^{(m)}(F) = \int_0^1 \cdots \int_0^1 \left(\int_0^1 F(t_1, \dots, t_k) \, dt_m \right)^2 dt_1 \cdots dt_{m-1} \, dt_{m+1} \cdots dt_k.$$

Met $o(1)$ bedoelen we een functie die naar 0 gaat als $N \to \infty$. Om het bewijs af te maken is het zaak een functie F te vinden zó dat

$$\frac{\sum_{m=1}^{k} J_k^{(m)}(F)}{I_k(F)} > 2.$$

Alleen dan kunnen we, door δ klein genoeg te kiezen, ervoor zorgen dat we een $\rho > 1$ vinden zó dat $S_1(N)/S_2(N) > \rho$ als N voldoende groot is. Het vinden van een geschikte F is een subtiel optimalisatieprobleem en het is bijzonder verrassend dat het inderdaad mogelijk is. We moeten daarvoor wel zoeken onder functies in minstens 105 variabelen. Dat wil zeggen, $k \ge 105$. Hier is een toelaatbare rij van lengte 105,

$$\begin{array}{l} 0, 10, 12, 24, 28, 30, 34, 42, 48, 52, 54, 64, 70, \\ 72, 78, 82, 90, 94, 100, 112, 114, 118, 120, 124, \\ 132, 138, 148, 154, 168, 174, 178, 180, 184, \\ 190, 192, 202, 204, 208, 220, 222, 232, 234, \\ 250, 252, 258, 262, 264, 268, 280, 288, 294, \\ 300, 310, 322, 324, 328, 330, 334, 342, 352, \\ 358, 360, 364, 372, 378, 384, 390, 394, 400, \\ 402, 408, 412, 418, 420, 430, 432, 442, 444, \\ 450, 454, 462, 468, 472, 478, 484, 490, 492, \\ 498, 504, 510, 528, 532, 534, 538, 544, 558, \\ 562, 570, 574, 580, 582, 588, 594, 598, 600. \end{array}$$

Merk op dat $a_{105} - a_1 = 600$ en dat verklaart de grens in de volgende stelling.

Stelling 10 (James Maynard, 2014). *Er zijn oneindig veel n zó dat*

$$p_{n+1} - p_n \le 600.$$

Foto: Eleanor Grant

James Maynard

Latere verbeteringen hebben ervoor gezorgd dat de grens op 246 kwam. Door k nog groter te kiezen kunnen we

$$\frac{\sum_{m=1}^{k} J_k^{(m)}(F)}{I_k(F)},$$

en dus ρ, zo groot maken als we willen. Gevolg,

Stelling 11 (Maynard, 2014). *Stel $r \in \mathbb{N}$. Dan bestaat er een C_r zó dat*

$$p_{n+r} - p_n < C_r$$

voor oneindig veel n.

Met andere woorden, bij elke $r \in \mathbb{N}$ bestaat een getal C_r zó dat er oneindig veel intervallen met lengte C_r zijn die minstens $r+1$ priemgetallen bevatten.

Referenties

1. J. Friedlander en H. Iwanies, Opera de Cribro, *AMS Colloquium Publications* 57 (2010), AMS.
2. J. Maynard, Small gaps between primes, *Annals of Math.* 181 (2015), 383–413.
3. Y. Zhang, Bounded gaps between primes, *Annals of Math.* 179 (2014), 1121–1174.
4. http://www.newyorker.com/magazine/2015/02/02/pursuit-beauty.
5. http://michaelnielsen.org/polymath1/index.php?title=Bounded_gaps_between_primes.