Open priemproblemen

Jan van de Craats

Misschien denk je dat over priemgetallen, de bouwstenen van het rekenen, wel zo ongeveer alles bekend is. Dat er op dat terrein geen onopgeloste vraagstukken meer zijn. Vergeet het maar: in de getallentheorie wemelt het van de open priemproblemen: uitdagende vragen over priemgetallen waar wiskundigen nog steeds geen antwoord op weten.

Oneindig veel

Natuurlijk is er ook veel dat wél bekend is. Zo weten we al sinds Euclides (ongeveer 300 voor Christus) dat er oneindig veel priemgetallen bestaan. Het bewijs is simpel, maar toch ook erg mooi. Hier is het.

Een geheel getal n heet een *priemgetal* als het groter dan 1 is en geen delers heeft die groter dan 1 en kleiner dan n zijn. De rij van alle priemgetallen begint als volgt:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, \ldots$$

Noem het k-de priemgetal p_k, dus $p_1 = 2$, $p_2 = 3$, $p_3 = 5$ enzovoort. Maak nu een nieuwe rij $\{q_k\}$ als volgt:

$$q_k = p_1 \times p_2 \times \cdots \times p_k + 1$$

dus $q_1 = 2 + 1 = 3$, $q_2 = 2 \times 3 + 1 = 7$, $q_3 = 2 \times 3 \times 5 + 1 = 31$, ...

Je ziet dat q_k altijd groter dan p_k is en dat q_k rest 1 geeft bij deling door elk priemgetal p_i dat kleiner dan of gelijk aan p_k is. Het getal q_k is dus niet deelbaar door zo'n priemgetal p_i. Daarom is q_k òf zelf een priemgetal, òf het is een getal dat ontbonden kan worden worden in priemgetallen *die allemaal groter zijn dan* p_k. Hieruit volgt dat er een priemgetal moet zijn dat groter is dan p_k. Maar omdat deze redenering voor elke p_k geldig is, kan de rij van alle priemgetallen nooit ophouden. Er zijn dus inderdaad oneindig veel priemgetallen.

De onderstaande tabel geeft een overzicht van het begin van de rijen p_k en q_k. Je ziet hoe snel die tweede rij groeit, en ook dat sommige q_k priem zijn, terwijl de andere zijn samengesteld uit priemfactoren die inderdaad allemaal groter zijn dan de bijbehorende p_k.

k	p_k	q_k	
1	2	3	(priem)
2	3	7	(priem)
3	5	31	(priem)
4	7	211	(priem)
5	11	2311	(priem)
6	13	30031	$= 59 \times 509$
7	17	510511	$= 19 \times 97 \times 277$
8	19	9699691	$= 347 \times 27953$
9	23	223092871	$= 317 \times 703763$
10	29	6469693231	$= 331 \times 571 \times 34231$
11	31	200560490131	(priem)
...	...	...	

Over de rij $\{q_k\}$ kun je al gelijk twee voor de hand liggende vragen stellen die nog geen van beide beantwoord zijn:

Open probleem 1: Bevat de rij $\{q_k\}$ oneindig veel priemgetallen?

Open probleem 2: Bevat de rij $\{q_k\}$ oneindig veel samengestelde getallen (dat wil zeggen getallen die geen priemgetal zijn)?

Priemtweelingen

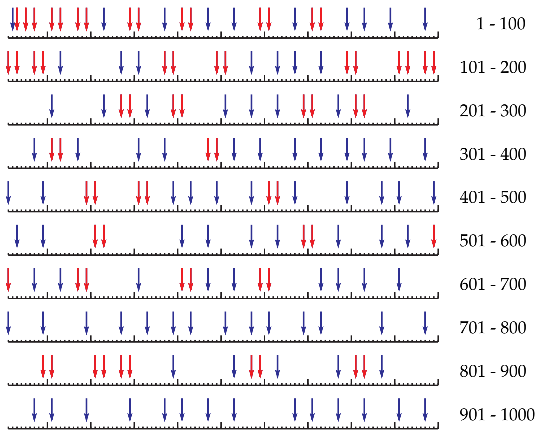
Maar ook over de rij $\{p_k\}$ van alle priemgetallen zelf is nog veel onbekend. In Figuur 1 zijn de priemgetallen onder de duizend met pijlen in beeld gebracht. Met rode pijlen zijn de *priemtweelingen* aangegeven, dat wil zeggen de paren priemgetallen die een verschil van precies 2 hebben. Omdat alle priemgetallen groter dan 2 oneven zijn, is er behalve het paar $(2, 3)$ geen tweetal met verschil 1. De priemtweelingen zijn dus de 'naaste buren' in de rij van alle priemgetallen. In Figuur 1 zie je heel wat priemtweelingen: $(3, 5)$, $(5, 7)$, $(11, 13)$ ..., maar ook bijvoorbeeld $(599, 601)$ en $(881, 883)$. Tussen 700 en 800 zit geen enkele tweeling en evenmin tussen 900 en 1000. Worden de priemtweelingen op den duur steeds zeldzamer? Houdt de rij misschien zelfs op een zeker moment op? We weten het niet. Een heel beroemd open probleem is namelijk:

Open probleem 3: Zijn er oneindig veel priemtweelingen?

En, daaraan verwant:

Open probleem 4: Als er oneindig veel priemtweelingen zijn, hoe liggen die dan verdeeld?

Naar deze problemen is al veel onderzoek gedaan, maar de definitieve antwoorden kennen we nog steeds niet. Bijna alle experts geloven dat er inderdaad



Figuur 1: De priemgetallen kleiner dan duizend. Met rode pijlen zijn de priemtweelingen aangegeven.

oneindig veel priemtweelingen zijn en er zijn ook plausibele vermoedens over hun verdeling, maar bewijzen ontbreken. Wel zijn er veel priemtweelingen bekend. De recordhouder op het moment dat we dit schrijven, is het paar $33218925 \times 2^{169690} \pm 1$, bestaande uit twee priemgetallen van elk 51090 cijfers. Dat tweetal is in 2002 ontdekt. Op het internet kun je nakijken of dit record inmiddels gebroken is (googlen met *twin primes*).

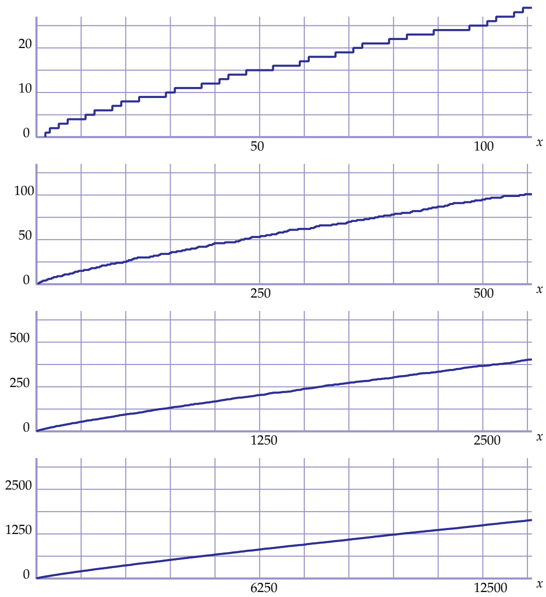
De priemgetallen-telfunctie $\pi(x)$.

Figuur 1 geeft de indruk dat de verdeling van de priemgetallen erg chaotisch is. Er lijkt weinig systeem in te zitten. In zekere zin is dat ook zo: als je bijvoorbeeld het duizendste priemgetal p_{1000} wilt berekenen, zit er weinig anders op dan de hele rij $\{p_k\}$ af te lopen totdat je p_{1000} hebt. Er is geen eenvoudige formule voor of een snelle sluiproute. (Voor wie het weten wil: $p_{1000} = 7919$.)

Maar toch, een totale chaos kan de verdeling van de priemgetallen ook niet zijn. Dat zie je wanneer je een grafiek maakt van de *priemgetallen-telfunctie* $\pi(x)$ die gedefinieerd is als het aantal priemgetallen kleiner dan of gelijk aan x. Wiskundigen gebruiken traditioneel de letter π om deze functie aan te geven en wij sluiten ons daarbij aan. Maar let op: dit functiesymbool π heeft dus niets te maken met het beroemde getal $\pi = 3,1415926\ldots$!

In Figuur 2 zie je vier maal de grafiek van $\pi(x)$, telkens uitgezoomd met een factor vijf. In het bovenste plaatje, dat hoort bij het traject $0 \leq x \leq 110$, zijn de trapjes van de grafiek duidelijk te zien: bij elk nieuw priemgetal neemt

$\pi(x)$ met één toe, en daartussen is $\pi(x)$ constant. Het volgende plaatje laat het traject $0 \leq x \leq 550$ zien, en de grafiek begint er al gladder uit te zien. Op $0 \leq x \leq 2750$ zijn er nog nauwelijks trapjes zichtbaar, en op $0 \leq x \leq 13750$ is er geen enkele oneffenheid meer zichtbaar. De gladheid van deze grafiek is ongetwijfeld een van de verbluffendste verschijnselen in de wiskunde: op kleine schaal lijken de priemgetallen chaotisch verdeeld te zijn, maar globaal gezien is er sprake van een prachtige glad verlopende grafiek.



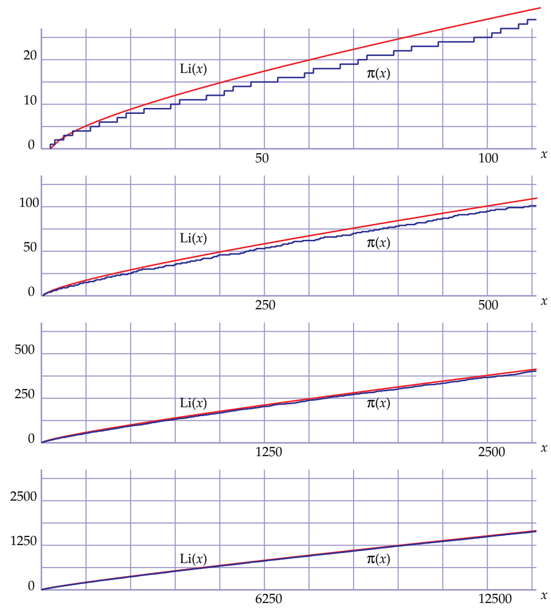
Figuur 2: De grafiek van de functie $\pi(x)$ op vier schalen getekend, telkens uitgezoomd met een factor 5.

De logaritmische integraal

Globaal gezien lijkt de grafiek van $\pi(x)$ een gladde kromme, op het eerste gezicht zelfs een rechte lijn. Maar als je beter kijkt, zie je dat de helling ervan geleidelijk afneemt naarmate x toeneemt. De grote wiskundige Carl Friedrich Gauss (1777-1855) was die regelmaat ook al op het spoor gekomen. Hij probeerde een eenvoudig formule voor een functie te vinden die bij benadering zo'n zelfde grafiek heeft, en ontdekte daarbij op 15-jarige leeftijd dat de globale helling van de grafiek van $\pi(x)$ in het punt x goed benaderd kan worden door $1/\ln x$. Dat wil zeggen dat

$$\frac{\pi(x+a) - \pi(x)}{a} \approx \frac{1}{\ln x}$$

wanneer a veel kleiner is dan x, maar toch nog groot genoeg om geen last te hebben van de kleine oneffenheden in de grafiek van $\pi(x)$. In Figuur 2 kun je in de onderste grafiek bijvoorbeeld $x = 12500$ en $a = 1250$ nemen. Omdat $1/\ln 12500 \approx 0,1060$ zou dan moeten gelden dat $\pi(13750)-\pi(12500) \approx 0,1060\times 1250 = 132,5$ en dat klopt uitstekend want $\pi(13750) = 1625$ en $\pi(12500) = 1492 = 1625 - 133$.



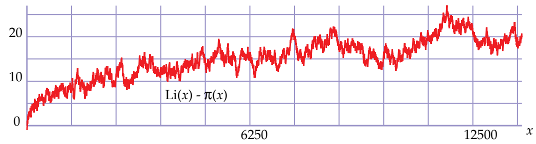
Figuur 3: De grafieken van de functies $\pi(x)$ en $\mathrm{Li}(x)$.

Als $F(x)$ een echt gladde, differentieerbare functie is die de priemgetallentelfunctie $\pi(x)$ goed benadert, zou volgens het vermoeden van Gauss dus voor alle x moeten gelden dat $F'(x) = 1/\ln x$. Dat wil zeggen dat

$$F(x) = \int_2^x \frac{1}{\ln t}\, dt$$

waarbij de ondergrens 2 zo gekozen is dat $F(2) = 0$. Deze functie wordt de *logaritmische integraal* genoemd, afgekort $\mathrm{Li}(x)$. In Figuur 3 zie je de grafieken van de functies $\pi(x)$ en $\mathrm{Li}(x)$ samen getekend. Vooral op grote schaal is de overeenkomst verbazingwekkend goed.

Hoe goed de overeenkomst tussen $\pi(x)$ en $\mathrm{Li}(x)$ ook lijkt, helemaal gelijk zijn ze niet. De grafiek van de verschilfunctie $\mathrm{Li}(x) - \pi(x)$ die in Figuur 4 getekend is, laat dat duidelijk zien Je ziet dat het verschil $\mathrm{Li}(x) - \pi(x)$ geweldig fluctueert, maar toch relatief klein is. Het lijkt ook (behalve in het allereerste begin) altijd positief te blijven, maar dat is niet zo: de Britse wiskundige Littlewood



Figuur 4: De grafiek van de verschilfunctie $\mathrm{Li}(x) - \pi(x)$.

toonde in 1914 aan dat $\mathrm{Li}(x) - \pi(x)$ oneindig vaak van teken wisselt. Toch heeft niemand ooit een grote x gevonden waarvoor $\mathrm{Li}(x) - \pi(x) < 0$ is. Veel wiskundigen hebben wel bovengrenzen aan kunnen geven waaronder er zeker zo'n x moet zijn. De beste bovengrens op dit moment is $1,39822 \times 10^{316}$.

Een verband met Riemanns vermoeden

Als we precieze, hanteerbare formules zouden hebben voor het verschil $\mathrm{Li}(x) - \pi(x)$, zouden we de verdeling van de priemgetallen helemaal kennen. In feite was dat ook het doel van het onderzoek van Riemann dat uitmondde in zijn artikel uit 1859 waarin ook zijn beroemde, nog steeds onopgeloste vermoeden te vinden is ("alle niet-triviale nulpunten van de zètafunctie liggen op de kritische lijn", zie het artikel over de milleniumproblemen in het vorige nummer van Pythagoras). Als je dat vermoeden bewijst of weerlegt, verdien je een miljoen dollar. Dat dit vermoeden nauw verbonden is met de verschilfunctie $\mathrm{Li}(x) - \pi(x)$ is in 1901 door de Zweedse wiskundige Helge von Koch nog duidelijker gemaakt. Hij liet namelijk zien dat Riemanns vermoeden equivalent is met het bestaan van een positieve constante C waarvoor geldt dat

$$|\mathrm{Li}(x) - \pi(x)| < C\sqrt{x}\ln x$$

voor alle x. Die ongelijkheid houdt in dat de verschilfunctie niet al te snel kan groeien, want zowel de wortelfunctie als de logaritme stijgen maar langzaam.

De priemgetallenstelling

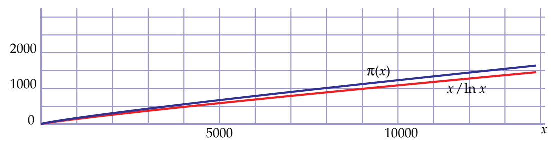
Tot nu toe hebben we over de priemgetallen-telfunctie $\pi(x)$ eigenlijk nog niets bewezen behalve dat $\lim_{x\to\infty} \pi(x) = \infty$ (want er zijn oneindig veel priemgetallen). Gauss vermoedde dat $\mathrm{Li}(x)$ een goede benadering van $\pi(x)$ is, maar bewijzen kon hij het niet. Trouwens, wat is een 'goede benadering'? Het Riemann-vermoeden houdt in dat het verschil $\mathrm{Li}(x) - \pi(x)$ 'niet te snel stijgt'. Maar het Riemann-vermoeden is nog steeds niet bewezen. Het wordt nu dus tijd om iets positiefs te vermelden. Dat vinden we in de *relatieve fout*

$$\frac{\mathrm{Li}(x) - \pi(x)}{\mathrm{Li}(x)} = 1 - \frac{\pi(x)}{\mathrm{Li}(x)}$$

Veel wiskundigen, ook Gauss en Riemann, hadden het idee dat die relatieve fout naar nul gaat voor $x \to \infty$, dus dat

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

Dit staat bekend als de *priemgetallenstelling*. Dat is ook heel lang een beroemd open probleem geweest. Maar in 1896 werd door twee wiskundigen onafhankelijk van elkaar een bewijs gevonden. Het waren de Fransman Jacques Hadamard en de Belg Charles de la Vallée Poussin, die daarbij voortbouwden op het werk van Riemann.



Figuur 5: De grafieken van $\pi(x)$ en $x/\ln x$.

Overigens, het is niet zo heel moeilijk om te bewijzen dat

$$\lim_{x \to \infty} \frac{\mathrm{Li}(x)}{x/\ln x} = 1$$

zodat ook de eenvoudige functie $f(x) = x/\ln x$ een benadering van $\pi(x)$ geeft waarvoor de relatieve fout naar nul gaat. De priemgetallenstelling wordt daarom ook wel geformuleerd als

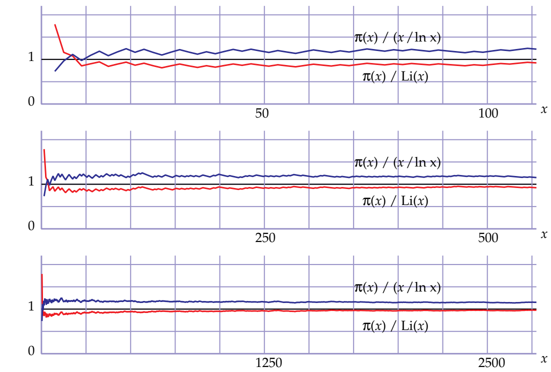
$$\lim_{x \to \infty} \frac{\pi(x)}{x/\ln x} = 1$$

Deze benaderingsfunctie is echter slechter dan $\mathrm{Li}(x)$ in die zin dat het quotient veel langzamer naar 1 gaat, zoals je in Figuur 6 kunt zien. Voor $x \approx 2500$ is $\pi(x)/\mathrm{Li}(x)$ (rode grafiek) al nauwelijks van 1 te onderscheiden terwijl $\pi(x)/(x/\ln x)$ (blauwe grafiek) er nog een flink eind boven zit.

Ook in absolute zin is de benadering $x/\ln x$ veel slechter dan $\mathrm{Li}(x)$. Dat is goed te zien in Figuur 5, waarin de grafieken van $\pi(x)$ (blauw) en $x/\ln x$ (rood) getekend zijn. Die van $\mathrm{Li}(x)$ valt op deze schaal volledig met $\pi(x)$ samen, terwijl de grafiek van $x/\ln x$ er duidelijk een flink eind onder ligt.

Nogmaals: priemtweelingen

Wat we met de priemgetallen-telfunctie $\pi(x)$ gedaan hebben, kunnen we natuurlijk ook voor priemtweelingen doen. Definieer de *priemtweelingen-telfunctie* $\pi_2(x)$ als het aantal priemtweelingen $(p, p+2)$ waarvoor p kleiner dan of gelijk

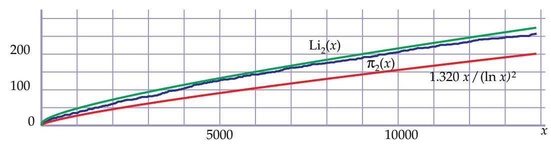


Figuur 6: De priemgetallenstelling in beeld gebracht.

aan x is. Als we van die functie een grafiek maken, zien we weer een stijgende trapfunctie die er op grote schaal bekeken tamelijk glad uitziet. Met de truc van Gauss kun je ook hiervan de helling proberen te schatten als functie van x. Gecombineerd met nog wat theoretische overwegingen die we hier onvermeld moeten laten, is men gekomen tot een schatting voor die helling van de vorm $C/(\ln x)^2$ waarbij C gegeven wordt door een formule die hoort bij een numerieke waarde $C \approx 1,320$. Hiermee kan men als volgt een *logaritmische-tweelingenintegraal* $\mathrm{Li}_2(x)$ definiëren:

$$\mathrm{Li}_2(x) = \int_2^x \frac{C}{(\ln t)^2}\, dt$$

die, naar men hoopt, de functie $\pi_2(x)$ goed benadert in de zin dat de relatieve fout naar nul gaat. Een eenvoudiger, maar naar verwachting minder goede benadering is $Cx/(\ln x)^2$. In Figuur 7 zie je ter illustratie een grafiek van de functies $\pi_2(x)$, $\mathrm{Li}_2(x)$ en $1,320x/(\ln x)^2$.



Figuur 7: De priemtweelingen-telfunctie $\pi_2(x)$ en twee benaderingsfuncties.

Helaas, dit zijn allemaal alleen nog maar vermoedens, die overigens wel door veel indrukwekkende numerieke gegevens worden ondersteund (zie het internet), maar die nog geen van alle bewezen zijn!