Wieb Bosma
Mathematisch Instituut
Katholieke Universiteit Nijmegen
Postbus 9010, 6500 GL Nijmegen
bosma@sci.kun.nl

Jaap Top
Instituut voor wiskunde en informatica
Rijksuniversiteit Groningen
Postbus 800, 9700 AV Groningen
top@math.rug.nl

De p van primaliteit

Op 6 augustus 2002 publiceerden Manindra Agrawal, Neeraj Kayal en Nitin Saxena van het Indian Institute of Technology een preprint op hun website met als titel 'PRIMES is in P'. Al snel verbreidde zich het nieuws dat nu eindelijk bewezen was dat priemgetallen en samengestelde getallen in polynomiale tijd te onderscheiden zijn. Dezelfde maand berichtten diverse kranten en tijdschriften hierover. Veel lezers zullen zich afgevraagd hebben wat de betekenis van dit resultaat is — was zoiets niet eigenlijk al bekend? Hoe werkt het algoritme, en wat zijn de praktische consequenties? En kunnen we nu sneller getallen in factoren ontbinden? Op dergelijke vragen geven de auteurs een antwoord.

Het is de laatste 25 jaar gebruik een artikel over primaliteitstests te beginnen met een citaat van Gauss, waarin hij 200 jaar geleden aangaf hoe belangrijk het herkennen van priemgetallen en het ontbinden van getallen in factoren is.[1] In het citaat worden twee problemen aangestipt. Voor het eerste probleem, om van een gegeven getal N vast te stellen of het priem is of niet, hebben de Indiase wiskundigen voor het eerst een algoritme gegeven dat deterministisch is met een rekentijd die polynomiaal in de grootte $\log N$ begrensd is zonder dat in het bewijs daarvan onbewezen vermoedens gebruikt worden. Voor het tweede probleem, om van gegeven N alle factoren te bepalen, heeft dit geen directe gevolgen. Een polynomiaal factorisatie-algoritme zou verstrekkende gevolgen hebben voor de complexiteitstheorie, de cryptografie, enzovoorts.

Op het eerste gezicht is het verrassend dat de twee problemen verschillend zijn: is een priemgetal immers niet een getal waarvoor het onmogelijk is een ontbinding te vinden?

Delers

Laten we, bij wijze van voorbeeld, de multiplicatieve 'geaardheid' van de getallen in de verzameling $\{314021, 314721, 314821, 314921, 315021\}$ proberen te bepalen. We doen het op een manier die vergelijkbaar is aan wat een modern computeralgebra-pakket zal proberen. Omdat machten vaak tot complicaties leiden, is het verstandig allereerst te onderzoeken of een getal een echte macht is; zo loopt $m = 314721$ als eerste tegen de lamp, als 561^2. Het uitproberen of $\sqrt[k]{N}$ geheel is, voor $k = 2, 3, \ldots, {}^2\log N$, kan zeer efficiënt gebeuren.

Een *deler* is natuurlijk het meest overtuigende argument voor samengesteldheid – op heterdaad, mogen we wel zeggen. Soms zijn delers overduidelijk aanwezig, zoals de deler 3 in $o = 315021$. Helaas is dat niet altijd zo — wij kennen geen methode om aan de schrijfwijze in ons positionele stelsel eenvoudig een deler van elk samengesteld getal af te lezen. Het eenvoudigste argument is niet altijd het makkelijkst te vinden argument; immers een samengesteld getal dat geen macht is hoeft niet meer dan twee echte delers te hebben. Dat is zoeken naar een speld in een hooiberg. Natuurlijk is het nuttig, bijvoorbeeld door test-deling, uit te proberen of er 'kleine' factoren zijn (zeg onder de 10^6; in onze mini-voorbeelden vinden we onder de 10^1 alleen factoren 3 in m en o).

Getuigen

Gelukkig zijn er ook andere, vaak gemakkelijker te vinden argumenten die samengesteldheid aangeven. We draaien als het ware de bewijslast om: ga er vanuit dat je met een onschuldig priemgetal van doen hebt en verkrijg dan een tegenspraak voor samengestelde N omdat een bepaalde eigenschap van priemgetallen niet geldt. Een veel gebruikte mogelijkheid biedt de kleine stelling van Fermat: als N priem is en a niet deelt, zal $a^{N-1} - 1$ deelbaar zijn door N. Machtsverheffen kan efficiënt door herhaald kwadrateren en vermenigvuldigen, en (tussen)resultaten blijven klein omdat we modulo N rekenen. Dit geeft daarom een efficiënte test, die garandeert dat als N er niet aan voldoet, N niet priem is. Zo'n a is dan een *getuige* voor de samengesteldheid van N. Zo is $n = 314921$ een niet-priem in ons lijstje, omdat $2^{314920} \equiv 227428 \bmod 314921$. De getuige 2 vertelt ons niets over de factoren van n. Hebben we de getuige in handen, dan is de correctheid van zijn aantijging eenvoudig te verifiëren. Maar hoe eenvoudig is het *vinden* van een getuige?

Quasi

De overgebleven getallen 314021 en 314821 uit ons lijstje hebben beide de eigenschap dat er geen Fermat-getuigen te vinden zijn die met het getal onderling ondeelbaar zijn. Volgens het adagium *if it looks like a prime and it behaves like a prime, it must be a prime* zouden beiden gemakkelijk voor priem uitgemaakt worden. Toch is één van deze getallen niet priem. Wiskunde is welhaast uitgevonden om dergelijke drogredeneringen te weerleggen, dus we gaan op zoek naar een methode om ook zulke quasi-priemgetallen te ontmaskeren.

Het probleem is dat niet elke a voldoet als Fermat-getuige voor een samengesteld getal. Sterker nog, er bestaan samengestelde getallen waarvoor zulke getuigen net zo moeilijk te vinden zijn als factoren: voor zulke *Carmichaelgetallen* voldoen alleen a die een deler met N gemeen hebben niet aan de Fermat-identiteit $a^{N-1} \equiv 1 \bmod N$.

Fermat-getuigen leveren eigenlijk een groepentheoretisch argument: de met N onderling ondeelbare restklassen modulo N vormen een groep $(\mathbf{Z}/N\mathbf{Z})^*$ onder vermenigvuldiging. Priemgetallen onderscheiden zich doordat de orde van deze groep $N-1$ is, want voor samengestelde N is de orde kleiner. In een eindige groep deelt de orde van elk element de groepsorde. Voor een Carmichaelgetal N geldt de curieuze eigenschap dat de orde van elk element $N-1$ deelt, hoewel $N-1$ *niet* de groepsorde is. Enkele jaren geleden werd aangetoond dat er oneindig veel Carmichaelgetallen bestaan.

Nauwkeuriger kijken

Het ligt voor de hand een verfijnder argument te zoeken om na te gaan of de orde van de groep $(\mathbf{Z}/N\mathbf{Z})^*$ kleiner dan $N-1$ is. Het gevaar doemt dan op dat men eigenschappen van $N-1$ wil gebruiken die slechts aan de factorisatie af te lezen zijn, maar $N-1$ zou wel eens moeilijk in factoren te ontbinden kunnen zijn! Gelukkig volstaat het om alleen gebruik te maken van factoren 2 in $N-1$: schrijf $N-1 = 2^k \cdot r$, met r oneven. In plaats van direct naar $a^{N-1} \bmod N$ te kijken, kan men eerst $a^r \bmod N$ bepalen. Als dit de restklasse van 1 of van $-1 \bmod N$ oplevert krijgen we geen nuttige informatie over de aard van N; vinden we een andere restklasse b, dan gaan we kwadrateren, om $b^2, b^{2^2}, \ldots, b^{2^k} = a^{N-1} \bmod N$ te bepalen. Is de laatste niet 1, dan is N samengesteld volgens het Fermat argument. Is de laatste wél 1, dan zijn we ergens een restklasse c tegengekomen die zelf niet gelijk aan die van 1 was maar het kwadraat wel; als c de restklasse van -1 was kan N priem zijn, maar is $c \neq -1 \bmod N$ dan moet N samengesteld zijn. De reden is dat $\mathbf{Z}/N\mathbf{Z}$ een lichaam vormt als N een priemgetal is, en in een lichaam kan de vergelijking $x^2 - 1 = 0$ slechts de twee oplossingen ± 1 hebben.

De primaliteitstest volgens Agrawal, Kayal, Saxena en Lenstra

Gegeven: een geheel getal N.

1. Test of er gehele getallen m, e groter dan 1 bestaan met $m^e = N$; als dat zo is dan is N samengesteld.
2. Laat v het dichtst bij $(^2 \log N)^2$ gelegen gehele getal zijn; bepaal de priemgetallen $2, 3, \ldots, p_i, \ldots$ maar stop zodra ofwel p_i een deler van N is ofwel p_i geen der getallen $N-1, N^2-1, \ldots, N^v-1$ deelt en $p_i \geq 5$. In het eerste geval is N priem als $p_i = N$ en anders samengesteld, in het tweede geval gaan we verder met $r = p_i$.
3. Test voor $a = 1, 2, \ldots, r-2$ of $(x+a)^N \equiv x^N + a \bmod (x^r - 1)/(x-1)$ in $\mathbf{Z}/N\mathbf{Z}[x]$; als dit voor een a niet geldt is N samengesteld. Als het voor alle a geldt is N priem.

Het blijkt dat $q = 314821$ snel doorslaat wanneer we hem op dergelijke wijze aan de tand voelen: voor $a = 2$ vinden we nog dat $a^{(q-1)/2} \equiv -1 \bmod q$ maar voor $a = 3$ is $c = a^{(q-1)/4} \equiv 290603 \bmod q$ en $c^2 \equiv 1 \bmod q$. We hebben dus een quasi-priemgetal q in de lijst gevonden!

Spelden te over

Het aardige is nu dat we kunnen bewijzen dat er voor elke samengestelde N heel veel a bestaan waarmee dit verfijndere argument werkt: minstens $3/4$ van alle restklassen modulo N doet het. We kunnen rustig spreken van *kroon*-getuigen. Onze quasi-priem q valt bij haast 85% van de bases door de mand, zoals bij $3, 5, 11$, maar niet bij 2 of 7.

Plotseling zijn we terechtgekomen in een hooiberg waar de *meeste* sprieten speld blijken. Door *willekeurige* restklassen a te 'ondervragen' kan de *kans* dat bij een samengestelde N geen getuige wordt gevonden klein gemaakt worden; maar, behalve voor piepkleine N, vinden we uit zo'n 'buurtonderzoek' nooit een bewijs voor de primaliteit omdat het testen van minstens 25% van de restklassen ondoenlijk is. Onder aanname van een gegeneraliseerde Riemann-hypothese is echter bewezen dat in de kleine verzameling van alle restklassen met representant tussen 0 en $2(\log N)^2$ altijd een getuige zit, voor samengestelde N. Dat levert dan een deterministische test op met polynomiale rekentijd, waarvan de correctheid helaas alleen bewezen is onder aanname van een onbewezen vermoeden.

Spelden zoeken

Het getal $p = 314021$ doorstaat alle genoemde tests glansrijk; en hoewel we misschien geneigd zijn dan te *geloven* in de onbezoedeldheid van p, nemen we niet met minder genoegen dan onomstotelijk *bewijs*.

Bij het toekennen van de predicaten priem en samengesteld moet als leidraad gelden, om maar eens een ander afgekloven angelsaksisch bon-mot te gebruiken, dat *justice not only has to be done, but also has to be* seen *to be done*. Er spelen dan dus drie kwesties: wat is de waarheid, welk argument toont dit aan, en hoe vind ik dat argument? Dat laatste is bij uitstek een algoritmische vraag. Het is de enige vraag die we voor samengegestelde N niet on-conditioneel in polynomiale tijd hebben opgelost.

Vormen (kroon)getuigen het eenvoudig te verifiëren argument voor samengesteldheid, datzelfde kunnen priemcertificaten betekenen voor priemgetallen. Het certificaat voor de primaliteit van N kan bijvoorbeeld bestaan uit een rij van alle priemdelers p van $N-1$ en bij elke p een restklasse a waarvoor a^{N-1} wel maar $a^{(N-1)/p}$ niet de restklasse van $1 \bmod N$ is. Dat toont aan dat de orde van $(\mathbf{Z}/N\mathbf{Z})^*$ precies $N-1$ is, mits we recursief bewijzen dat de priemdelers p inderdaad priem zijn. Voor verificatie moeten we laten zien dat $N-1$ opgebouwd is uit de gegeven priemdelers en dat de restklassen a aan de genoemde eisen voldoen. Het probleem met zo'n 'certificaat van onbesmet blazoen' is natuurlijk weer om het te *vinden* — we hebben bijvoorbeeld hier de factorisatie van $N-1$ nodig.

Priembewijzen

Dit is het moment om een *primaliteitsbewijzer* aan te roepen. Door computerpakketten worden soms eventuele speciale eigenschappen (herkenbare vorm, zoals bij Mersenne-getallen $2^m - 1$) in combinatie met een database gebruikt, waarna zonodig een algemene primaliteitsbewijzer wordt aangeroepen. Van dat laatste bestaan twee praktische methoden: het *Jacobisom*-algoritme of het *elliptische kromme*-algoritme. Het eerste maakt gebruik van eigenschappen van cyclotomische uitbreidingen van $\mathbf{Z}/N\mathbf{Z}$. Het tweede algoritme gebruikt het aantal punten op elliptische krommen over eindige lichamen. Van beide bestaan

Nitin Saxena, Neeraj Kayal en Manindra Agrawal, de ontdekkers van de polynomiale primaliteitstest.

deterministische versies (maar worden in de praktijk probabilistische versies gebruikt); van de eerste is een preciese rekentijdanalyse bekend (met $\log\log\log N$ in de exponent), de tweede is conceptueel iets eenvoudiger en kan een certificaat opleveren, dat in de praktijk onhandelbaar groot is. Beide methoden zijn gebruikt voor priemgetallen tot een omvang van zo'n 2500 decimale cijfers en leveren routinematig bewijzen voor priemgetallen van enkele honderden cijfers.[2]

Tenslotte is er nog een theoretisch resultaat, gebaseerd op een generalisatie naar abelse variëteiten van de elliptische krommen test. Dat is een probabilistisch algoritme met een *verwachte* polynomiale rekentijd. Hiervan is geen praktische versie bekend.

De bijdrage uit India

De preprint van de Indiërs toont het bestaan aan van een methode aan waarmee *deterministisch in polynomiale tijd* bepaald kan worden of een gegeven getal priem of samengesteld is. Het maakt dus in theorie het zoeken naar getuigen en het aanroepen van de eerder genoemde priembewijzers overbodig. Eén van de grote verrassingen is hoe simpel het uiteindelijke algoritme is: het kan aan middelbare scholieren uitgelegd worden, en de correctheid kan voor tweedejaars wiskundestudenten begrijpelijk gemaakt worden.

De elementaire gedachte achter het algoritme is een andere eigenschap van priemgetallen dan de stelling van Fermat te gebruiken, en wel dat in de polynoomring $\mathbf{Z}/N\mathbf{Z}[x]$ voor een a onderling ondeelbaar met N geldt dat $(x+a)^N = x^N + a$ dan en slechts dan als N priem is.

Dat de identiteit geldt als N priem is volgt direct uit het binomium van Newton, de deelbaarheidseigenschappen van binomiaalcoëfficiënten en de kleine stelling van Fermat. Dat híer ook de omkering geldt, is eenvoudig te bewijzen door voor samengestelde N een binomiaalcoëfficiënt aan te geven die niet door N deelbaar is.

Het algoritme

In de gegeven vorm is de polynoomidentiteit natuurlijk niet te gebruiken, omdat het op een of andere manier de evaluatie van de $N+1$ coëfficiënten inhoudt. In plaats daarvan wordt de vergelijking genomen modulo het cyclotomische polynoom $(x^r-1)/(x-1)$, voor een hulp-priemgetal r; anders gezegd, we verifiëren een identiteit in een uitbreiding van $\mathbf{Z}/N\mathbf{Z}$ met een r-de eenheidswortel (hetgeen lijkt op wat er gebeurt in de Jacobisomtest), en dat kan weer heel efficiënt. In het algoritme wordt r geschikt gekozen en wordt de identiteit gecontroleerd voor voldoende waarden van a.

Het correctheidsbewijs toont aan dat in de geconstrueerde ring de identiteiten voor de gekozen waarden van a alleen kunnen gelden voor N die priemmacht zijn; dat N een echte macht is kan snel worden uitgesloten, zoals we zagen. In de complexiteitsanalyse bestaat het werk er voornamelijk uit te laten zien dat de omvang van de kleinst mogelijke geschikte r (effectief) begrensd wordt door een macht van $\log N$.

Verdere ontwikkelingen

In de oorspronkelijke versie wordt een polynomiale grens op de rekentijd verkregen met behulp van een vrij diep resultaat uit de analytische getaltheorie. Door Hendrik Lenstra is aangetoond dat dit ook met elementaire resultaten uit de analytische getaltheorie mogelijk is, en dat daarmee bovendien een *effectief* resultaat verkregen kan worden: er is een grens op de benodigde rekentijd voor dit algoritme van de vorm $C(\log N)^{12}$ met berekenbare constante C, als men afziet van machten van $\log\log N$. Het hele resultaat is dan met tweedejaars-algebra verkregen!

Voor een scherpere exponent moet harder gewerkt worden. Wellicht is de best mogelijke exponent voor een variant van dit algoritme 6, maar aan zo'n resultaat wordt nog gewerkt.

Of het nieuwe algoritme ook in de praktijk getuigenzoekers en priembewijzers gaat vervangen is maar zeer de vraag: hoewel het algoritme asymptotisch beter is, garandeert dat nog niet dat het in praktische omstandigheden (met N van niet meer dan enkele honderden of op zijn hoogst duizenden cijfers) ook superieur is. Misschien is een combinatie met andere methoden wel van praktisch belang.

Noten

1 "The problem of distinguishing prime numbers from composite numbers and of resolving the latter into their prime factors is known to be one of the most important and useful in arithmetic. [...] Further, the dignity of the science itself seems to require that every possible means be explored for the solution of a problem so elegant and so celebrated.", Carl Friedrich Gauss, *Disquisitiones Arithmeticae*, article 329, 1801 translation by Arthur A. Clarke, 1966 for Yale University Press revised by William C. Waterhouse et al for Springer Verlag, 1986.

2 Op 15 februari 2003 werd bekend gemaakt dat met het elliptische krommen algoritme, middels een berekening die 22 weken duurde, de primaliteit werd bewezen van een getal van 5878 cijfers, een record voor priemgetallen die niet van een speciale vorm zijn.