Jaap Top
Johann Bernoulli Instituut voor Wiskunde en Informatica
Rijksuniversiteit Groningen
j.top@rug.nl

Vakantiecursus

Het vermoeden van Birch en Swinnerton-Dyer

Op de Vakantiecursus 2013 van het Platform Wiskunde Nederland besprak Jaap Top een van de zeven millenniumproblemen: het vermoeden van Birch en Swinnerton-Dyer. Welke vooruitgang is er geboekt sinds dit probleem uit de getaltheorie vijftig jaar geleden werd geformuleerd?

Precies vijftig jaar geleden, in 1963, publiceerden Bryan Birch en Peter Swinnerton-Dyer in het *Journal für die reine und angewandte Mathematik* het eerste gedeelte [1] van hun artikel 'Notes on Elliptic Curves'. De tweede helft [2] verscheen in 1965 in hetzelfde tijdschrift. In de artikelen doen beide wiskundigen uit Cambridge (Engeland) verslag van hun jarenlange onderzoek betreffende elliptische krommen over de rationale getallen $\mathbb{Q}$. Daarbij was een voor hun tijd nieuw middel ingezet: de EDSAC 2 [16], een computer die de universiteit van Cambridge in 1958 had aangeschaft en die tot 1965 dienst deed.

Een (over)versimpelde versie van de vraag die Birch en Swinnerton-Dyer probeerden te beantwoorden is de volgende. Begin met een polynoom

$$F(x) := ax^3 + bx^2 + cx + d,$$

waarbij $a \neq 0$ en b, c, d rationale getallen zijn. Bestaan er rationale getallen r zodat $F(r)$ een kwadraat is (dat wil zeggen, $F(r) = s^2$ voor zeker rationaal getal s). En zo ja, zijn er eindig veel, of zelfs oneindig veel zulke r?

Fermat (1601–1665) en ook Euler (1707–1783) hadden al een methode aangegeven waarmee uit een oplossing $x = r$ in veel gevallen een andere oplossing gevonden kon worden. We komen hier straks op terug. Op p. 171 van een artikel [11] van Henri Poincaré (1854–1912) staat een uitspraak die in feite betekent dat uit een eindig aantal oplossingen $x = r_1, x = r_2, \ldots, x = r_q$ door herhaald toepassen van methoden als die van Fermat en Euler, *alle* oplossingen geconstrueerd kunnen worden. Poincaré verzuimt echter een bewijs of zelfs maar een schets van een argumentatie bij zijn uitspraak te geven. L.J. Mordell (1888–1972) toonde overigens in 1922 aan dat de uitspraak van Poincaré correct was, zie [10].

Wiskundig gezien is het nieuwe idee van Birch en Swinnerton-Dyer vooral dat ze het polynoom $F(x)$ modulo zoveel mogelijk priemgetallen p gaan bekijken. Voor zo'n priemgetal p kan je tellen hoeveel verschillende $r \bmod p$ er zijn met de eigenschap dat $F(r) \equiv s^2 \bmod p$ voor een zekere gehele s. Een heel naïeve en niet erg precieze vorm van het vermoeden van Birch en Swinnerton-Dyer, helemaal in de stijl van de inleiding van hun eerste artikel, luidt dan:

Als er voor de meeste priemgetallen p onverwacht veel verschillende $r \bmod p$ bestaan waarvoor $F(r) \bmod p$ een kwadraat modulo p is, dan neemt $F(x)$ voor oneindig veel rationale waarden $x = r$ een kwadraat als waarde aan.

De verdienste van de twee artikelen van Birch en Swinnerton-Dyer is dat dit is omgesmeed tot een precieze formulering, waarvoor ze bovendien een enorme numerieke evidentie gegeven hebben. In de afgelopen vijftig jaar is flink wat vooruitgang geboekt, maar het vermoeden is nog altijd niet bewezen. In het jaar 2000 werd het vermoeden gekozen als een van de problemen in de beroemde Claylijst. Andrew Wiles schreef voor die gelegenheid een korte en heel leesbare introductie [15].

In de voor u liggende tekst gaan we vooral niet proberen de meest algemene en precieze formulering te geven. Wel hopen we wat handvatten te leveren waarmee u zelf kan rekenen aan en experimenteren met speciale gevallen van het vermoeden, en we geven een kort overzicht van tot nu toe behaalde resultaten.

Constructies van Fermat en Euler

Gegeven zijn rationale getallen a, b, c, d met $a \neq 0$, en de bijbehorende derdegraads veelterm

$$F(x) := ax^3 + bx^2 + cx + d.$$

Een methode waarmee zowel Fermat als Euler in diverse voorbeelden indrukwekkende rationale getallen r (met tellers en noemers die te groot zijn voor een moderne standaard grafische rekenmachine) vonden, zodat $F(r)$ een kwadraat is, werkt als volgt. We gaan ervan uit dat al zo'n oplossing $x = r$ gevonden is en we zoeken een nieuwe. Daartoe schrijven we $x = \xi + r$, zodat de gegeven oplossing correspondeert met $\xi = 0$. Als veelterm in ξ hebben we dan

$$\tilde{F}(\xi) = F(\xi + r) = a\xi^3 + \tilde{b}\xi^2 + \tilde{c}\xi + \tilde{d},$$

met rationale getallen $a \neq 0$ (dezelfde

waarmee we begonnen) en $\tilde{b}, \tilde{c}, \tilde{d}$. Merk op dat

$$\tilde{d} = \tilde{F}(0) = F(r),$$

en daarvan is gegeven dat het een kwadraat is. Dus we kunnen schrijven $\tilde{d} = e^2$ voor een zeker rationaal getal e. Dit zorgt ervoor, dat we een uit het wat ouderwetse middelbaar onderwijs beproefde rekentechniek kunnen toepassen: 'kwadraat afsplitsen'. In dit geval houdt dit in dat rationale getallen α, β gezocht worden zodat geldt

$$\tilde{F}(\xi) = (\alpha\xi^2 + \beta\xi + e)^2 + \gamma\xi^3 + \delta\xi^4$$

voor zekere rationale γ, δ. Door de coëfficiënten van de machten van ξ te vergelijken, zien we dat moet gelden

$$\begin{cases} 2\beta e = \tilde{c}, \\ 2\alpha e + \beta^2 = \tilde{b}. \end{cases}$$

Dit levert precies één oplossing (β, α), behalve in het speciale geval dat $e = 0$ geldt. Dat laatste komt overeen met de voorwaarde dat we waren begonnen met een nulpunt $x = r$ van de veelterm $F(x)$. Gaan we er vanuit dat r geen nulpunt is, dan hebben we nu een schrijfwijze

$$\tilde{F}(\xi) = (\alpha\xi^2 + \beta\xi + e)^2 + \gamma\xi^3 + \delta\xi^4$$

gevonden. Hoe helpt dit alles bij het vinden van een $r' \neq r$ zodat $F(r)$ een kwadraat is, of equivalent, een $r'' \neq 0$ zodat $\tilde{F}(r'')$ een kwadraat is? Wel, in de gevonden uitdrukking voor $\tilde{F}(\xi)$ is de eerste term zonder meer een kwadraat, wat ook voor ξ wordt ingevuld. Dus we hebben een voorbeeld als de resterende termen samen de waarde 0 opleveren. Het resterende stuk laat zich schrijven als

$$\xi^3(\gamma + \delta\xi).$$

Voor $\xi = 0$ levert dit de waarde 0, maar deze ξ hadden we al. Een andere waarde voor ξ die werkt, is te vinden in het geval dat $\delta \neq 0$.

Een voorbeeld: begin met

$$F(x) := x^3 + 2x + 1.$$

Deze is te schrijven als

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

Omdat $2x^3 - \frac{1}{4}x^4$ als nulpunt $x = 8$ heeft, geldt $F(8)$ is een kwadraat. Verder is

$$\begin{aligned} F(x+8) &= x^3 + 24x^2 + 194x + 529 \\ &= \left(-\frac{3287}{24334}x^2 - \frac{97}{23}x - 23\right)^2 \\ &\quad - \frac{38998}{279841}x^3 - \frac{10804369}{592143556}x^4. \end{aligned}$$

Een nulpunt van

$$-\frac{38998}{279841}x^3 - \frac{10804369}{592143556}x^4$$

is

$$x = -\frac{82519768}{10804369}$$

en door hier 8 bij op te tellen, volgt dat $F\left(\frac{3915184}{10804369}\right)$ een kwadraat is. Bij zowel Fermat als Euler komen aanzienlijk omvangrijkere rekenvoorbeelden voor...

Als $\delta = 0$, dan levert de methode geen nieuwe waarde r' zodat $F(r')$ een kwadraat is. Nagaande hoe de $\delta, \gamma, \beta, \alpha$ bepaald worden, blijkt dat $\delta = 0$ equivalent is met $\alpha = 0$, oftewel dat was begonnen met een veelterm

$$\tilde{F}(\xi) = a\xi^3 + \tilde{b}\xi^2 + \tilde{c}\xi + \tilde{d} = \gamma\xi^3 + (\beta\xi + s)^2.$$

Een meer analytische beschrijving van wat hier gebeurt, werkt dan als volgt. De vraag is waarden voor ξ te vinden zodat $\tilde{F}(\xi)$ een kwadraat is, oftewel zodat $\sqrt{\tilde{F}(\xi)}$ bestaat, en rationaal is. De functie

$$\xi \mapsto \alpha\xi^2 + \beta\xi + s$$

die hierboven bepaald wordt, is precies de tweede-orde-Taylorbenadering bij $\xi = 0$ van $\pm\sqrt{\tilde{F}(\xi)}$ (met teken + als $s > 0$ en teken $-$ als $s < 0$). En deze tweede-orde-benadering voldoet aan $\alpha = 0$ (oftewel, heeft slechts graad ≤ 1), precies dan als de functie

$$\xi \mapsto \sqrt{\tilde{F}(\xi)}$$

voor $\xi = 0$ een buigpunt heeft.

Ook als $\delta \neq 0$ kan er iets misgaan: namelijk, we zoeken naar een nulpunt van $\gamma + \delta\xi$, en dat is natuurlijk $\xi = -\gamma/\delta$. Dat zou de nieuwe waarde voor ξ moeten zijn zodat $\tilde{F}(\xi)$ een kwadraat is. Echter, als $\gamma = 0$ dan geeft dit gewoon weer de al bekende waarde $\xi = 0$.

Terugvertaald naar $F(x)$: gegeven een rationaal getal r zodat $F(r)$ een kwadraat is. Dan levert de door Fermat en Euler gebruikte methode een ander rationaal getal $r' \neq r$ waarvoor eveneens geldt $F(r')$ is een kwadraat, behalve in drie gevallen:

- $F(r) = 0$;
- voor $x = r$ heeft $\sqrt{F(x)}$ een buigpunt;
- de methode levert dezelfde r weer op omdat "$\gamma = 0$".

Het bovenstaande laat zich vrij eenvoudig vertalen in de moderne manier waarop 'elliptische krommen' behandeld worden. Daartoe eerst heel kort iets over de theorie. Een goed leesbare, elementaire inleiding in dit fascinerende onderwerp is het boek van Joe Silverman en John Tate [12]. Allereerst eisen we van de veelterm $F(x)$ dat deze geen meervoudige nulpunten heeft, dus $F(x)$ is niet te schrijven in de vorm $a(x-\lambda)^2(x-\mu)$ of zelfs $a(x-\lambda)^3$ (was dat wel het geval, dan is het oorspronkelijk gestelde probleem heel simpel op te lossen, en er zijn dan in alle gevallen oneindig veel verschillende rationale waarden $x = r$ waarvoor $F(r)$ een kwadraat is). Voldoet $F(x)$ aan de bovengenoemde eis, definieer dan een 'elliptische kromme' als de collectie van alle paren (x, y) die voldoen aan

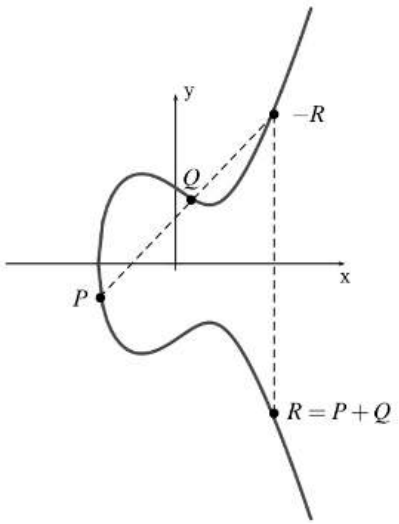
$$y^2 = F(x).$$

Hier voegt men nog een extra punt aan toe dat O genoemd wordt. Dat is niet een van de paren (x, y), maar een nieuw punt, waaraan gedacht mag worden als een oneindig ver van de oorsprong verwijderd punt. De aanname dat $F(x)$ geen meervoudige nulpunten heeft, betekent dat in elk punt (α, β) van de kromme een unieke raaklijn aan de kromme is gedefinieerd, namelijk de lijn met vergelijking

$$2\beta(y - \beta) = F'(\alpha)(x - \alpha).$$

Op de collectie van alle punten op de kromme laat zich dan als volgt een optelling definiëren:

1. $P + O = P$.
2. $P + Q = O$ als P en Q elkaars gespiegelde zijn ten opzichte van de x-as.
3. Als $P \neq Q$ en P en Q zijn niet elkaars gespiegelde ten opzichte van de x-as, dan snijdt de verbindingslijn van P en Q de kromme in drie punten (het derde punt kan samenvallen met P of met Q in het geval dat de verbindingslijn een raaklijn is). Nu geldt $P + Q = R$ waarbij R de gespiegelde in de x-as is van het derde gevonden punt.
4. Als $P = Q$ en dit punt ligt niet op de x-as (in dat geval was de optelling al in punt 2 hiervoor gedefinieerd), neem dan de raaklijn door P aan de kromme en het derde snijpunt van die raaklijn met de kromme. Dan geldt $P + P = R$ waarbij R de gespie-



Figuur 1

gelde in de x-as is van het derde snijpunt. Zie Figuur 1.

De zo gegeven optelling blijkt aan de gebruikelijke regels te voldoen, dus bijvoorbeeld geldt $(P + Q) + R = P + (Q + R)$ voor elk drietal P, Q, R, en elk punt P heeft een unieke tegengestelde Q (namelijk de gespiegelde in de x-as) zodat $P + Q = O$. Die tegengestelde wordt als $-P$ geschreven. Pas op: we zijn misschien een beetje gewend om bij $-(x, y)$ beide coördinaten van het punt van teken te veranderen, maar voor de hier gegeven regel geldt $-(x, y) = (x, -y)$. Tenslotte geldt per definitie dat de gemaakte optelling commutatief is: er geldt $P + Q = Q + P$ voor elk paar punten P, Q.

Als begonnen wordt met punten P, Q die beide coördinaten rationaal hebben, dan geldt hetzelfde voor $P + Q$ en voor $-P$ (voor zover die $\neq O$ is). Dus samen met O is de collectie van al zulke punten gesloten onder de bewerkingen optellen en tegengestelde nemen. Deze collectie wordt in de artikelen van Birch en Swinnerton-Dyer met de letter Γ aangeduid.

In termen van de optelling blijkt de constructiemethode van Fermat en van Euler het volgende te doen: gegeven een punt $P \in \Gamma$ wordt het 'nieuwe' punt

$$R := -3P = -(P + P + P)$$

bepaald. Als begonnen wordt met een P op de x-as, dan is $P + P = O$, en dus $R = P$. Die hadden we al. En als begonnen was met een buigpunt van de grafiek van $\sqrt{F(x)}$, dan geldt $R = O$, dus ook hier krijgen we geen 'echte' oplossing. Het derde geval waar geen nieuwe oplossing gekregen wordt is nu ook eenvoudig te verklaren: als geldt $4P = O$, dan is $R = P$, dus dan levert de methode hetzelfde punt op waarmee was begonnen. In termen van de gebruikte variabelen, gebeurt dit als $y = 0$, dus als geldt

$$\begin{aligned}\tilde{F}(\xi) &= (\alpha\xi^2 + \beta\xi + e)^2 - \alpha^2\xi^4 \\ &= 2\alpha\beta\xi^3 + (2\alpha e + \beta^2)\xi^2 + 2\beta e\xi + e^2.\end{aligned}$$

Deze veelterm heeft graad 3 en geen meervoudige nulpunten als aan de drie eisen

$$\begin{cases} \alpha \neq 0, \\ e \neq 0, \\ \beta^2 \neq 8\alpha e. \end{cases}$$

is voldaan. In dat geval voldoet $P = (0, e)$ aan $4P = O$ op de kromme gegeven door $y^2 = \tilde{F}(\xi)$.

Nog een voorbeeld: begin met

$$F(x) := x^3 + 9x^2 - 16x + 64.$$

Beginnend met $x = 0$ levert de methode van Fermat en Euler dan achtereenvolgend (reken maar na!) de waarden $x = 8$ en $x = -8$ op, gevolgd door (opnieuw) $x = 0$, enzovoort. Dus we krijgen een periodieke baan, met periode 3. Het punt $P = (0, 8)$ op de kromme met vergelijking $y^2 = x^3 + 9x^2 - 16x + 64$ blijkt in dit geval te voldoen aan $7P = O$.

Ten slotte: de methode zoals hier beschreven houdt zoals we zagen nauw verband met de tweede-orde-Taylorbenadering van de functie $\sqrt{\tilde{F}(\xi)}$ om $\xi = 0$. Maar we hadden ook kunnen volstaan met de eerste-orde-Taylorbenadering! Dat levert een schrijfwijze

$$\tilde{F}(\xi) = (\beta\xi + e)^2 + a\xi^3 + \epsilon\xi^2.$$

Door ξ zo te kiezen dat $a\xi^3 + \epsilon\xi^2$ de waarde 0 krijgt, wordt weer een oplossing verkregen. In termen van de gegeven optelling betekent dit, dat uitgaande van een punt P op de kromme, het nieuwe punt

$$R' = -2P = -(P + P)$$

wordt geconstrueerd. Oftewel: neem het derde snijpunt met de kromme van de raaklijn door P aan de kromme. Deze 'raaklijnmethode' staat in veel teksten over elliptische krommen vermeld, maar de 'paraboolmethode' die door Fermat en Euler veel vaker wordt toegepast, lijkt wat in de vergetelheid geraakt.

Modulorekenen

Het probleem van kwadraten als waarden $F(x)$ bij een rationale x te vinden, verandert niet als x met een vast rationaal getal $\neq 0$ wordt vermenigvuldigd, en evenmin als $F(x)$ met een vast kwadraat wordt vermenigvuldigd. Door $F(x) = ax^3 + bx^2 + cx + d$ te vermenigvuldigen met a^2 en vervolgens x te vervangen door $X = ax$, mogen we dus aannemen

$$F(X) = X^3 + bX^2 + cX + d$$

voor zekere (misschien nieuwe) rationale b, c, d. Vermenigvuldigen we dit met n^6 en schrijven we $x := n^2X$ (een nieuwe variabele x), dan staat er

$$x^3 + n^2bx^2 + n^4cx + n^6d.$$

Voor een geschikte keuze van n is dit een veelterm die alle coëfficiënten geheel heeft. Kortom, zonder verlies van algemeenheid kan worden aangenomen dat

$$F(x) = x^3 + bx^2 + cx + d$$

voor zekere gehele getallen b, c, d. Dit stelt ons in staat de veelterm te reduceren modulo ieder geheel getal, in het bijzonder modulo priemgetallen.

Daartoe eerst kort iets over het rekenen modulo een priemgetal. Kies een willekeurig priemgetal p. De 'getallen modulo p' zijn de uitdrukkingen $\overline{a} = a \bmod p$ waarin a een geheel getal is. Hierbij geldt de regel

$$\overline{a} = \overline{b} \quad \Leftrightarrow \quad p \text{ is een deler van } b - a.$$

Anders gezegd: $\overline{a}$ hangt alleen maar af van de rest die overblijft wanneer a door p gedeeld wordt, niet van het gehele getal a zelf. Er zijn dus in totaal precies p 'getallen modulo p', namelijk

$$\overline{0},\ \overline{1},\ \overline{2},\ \ldots,\ \overline{p-1}.$$

Met 'getallen modulo p' valt net zo te rekenen als met gewone gehele getallen:

$$\overline{a} + \overline{b} := \overline{a + b}$$

en

$$\overline{a} \cdot \overline{b} := \overline{a \cdot b}$$

blijken goed-gedefinieerde bewerkingen op te leveren.

Tot nu toe is niet gebruikt dat p wordt verondersteld een priemgetal te zijn. Die eigenschap levert iets speciaals voor de vermenig-

vuldiging: neem een willekeurige $\overline{a} \neq \overline{0}$. Met ook nog twee elementen $\overline{b}, \overline{c}$ geldt dan

$$\begin{aligned}\overline{a} \cdot \overline{b} = \overline{a} \cdot \overline{c} &\Leftrightarrow \overline{ab} = \overline{ac} \\ &\Leftrightarrow p \text{ is een deler van } ab - ac \\ &\Leftrightarrow p \text{ is een deler van } a(b-c) \\ &\Leftrightarrow p \text{ is een deler van } b - c \\ &\Leftrightarrow \overline{b} = \overline{c}.\end{aligned}$$

Immers, $\overline{a} \neq \overline{0}$ oftewel p is geen deler van a. Dus we zien: alle producten $\overline{a} \cdot \overline{b}$ verschillen, als $\overline{b}$ de collectie 'getallen modulo p' doorloopt. Er zijn p zulke producten en ook precies p 'getallen modulo p', dus kennelijk komt elk getal modulo p voor als zo'n product. In het bijzonder bestaat er een unieke $\overline{b}$ waarvoor geldt

$$\overline{a} \cdot \overline{b} = \overline{1},$$

deze $\overline{b}$ heet de 'inverse' van $\overline{a}$. Met de getallen modulo p kan men dus optellen, aftrekken, vermenigvuldigen en zelfs delen vergelijkbaar met gewone (rationale of reële) getallen. De collectie getallen modulo p, voorzien van deze optelling en vermenigvuldiging, wordt gewoonlijk genoteerd als $\mathbb{F}_p$.

Wij zijn met name geïnteresseerd in kwadraten in $\mathbb{F}_p$, oftewel elementen van de soort $\overline{a}^2 = \overline{a} \cdot \overline{a}$. Er geldt

$$\begin{aligned}\overline{a}^2 = \overline{b}^2 &\Leftrightarrow p \text{ is een deler van } a^2 - b^2 \\ &\Leftrightarrow p \text{ is een deler van } (a-b)(a+b) \\ &\Leftrightarrow p \text{ deelt } a-b \text{ of } p \text{ deelt } a+b \\ &\Leftrightarrow \overline{a} = \overline{b} \text{ of } \overline{a} = \overline{-b}.\end{aligned}$$

Voor $p = 2$ geldt bovendien $\overline{b} = \overline{-b}$ voor elke b, in dat geval is elk van de twee elementen van $\mathbb{F}_2$ zijn eigen kwadraat. Voor $p > 2$ geldt $\overline{b} = \overline{-b}$ alleen als $\overline{b} = \overline{0}$. Dus in dat geval zijn de $(p-1)/2$ kwadraten van de paren $\{\pm\overline{a}\}$ met $\overline{a} \neq \overline{0}$ allemaal verschillend, en bovendien verschillend van $\overline{0}^2 = \overline{0}$. Conclusie: voor oneven p zijn er precies $(p+1)/2$ kwadraten in $\mathbb{F}_p$ (inclusief $\overline{0}$). Kortom, ongeveer de helft van alle elementen van $\mathbb{F}_p$ is kwadraat.

Voorbeeld: de kwadraten in $\mathbb{F}_{17}$ zijn

$$\left\{\overline{0},\ \overline{1},\ \overline{2},\ \overline{4},\ \overline{8},\ \overline{9},\ \overline{13},\ \overline{15},\ \overline{16}\right\}$$

en die in $\mathbb{F}_{23}$ zijn

$$\left\{\overline{0},\ \overline{1},\ \overline{2},\ \overline{3},\ \overline{4},\ \overline{6},\ \overline{8},\ \overline{9},\ \overline{12},\ \overline{13},\ \overline{16},\ \overline{18}\right\}.$$

De veelterm

$$F(x) = x^3 + bx^2 + cx + d$$

met gehele b, c, d levert ook een veelterm met coëfficiënten in $\mathbb{F}_p$, namelijk

$$\overline{F}(x) := x^3 + \overline{b}x^2 + \overline{c}x + \overline{d}.$$

Voor elke $r \in \mathbb{F}_p$ is dan ook $\overline{F}(r) \in \mathbb{F}_p$, en men kan de vraag stellen hoe vaak het gebeurt dat $\overline{F}(r)$ een kwadraat is. Equivalent hiermee: neem de vergelijking

$$s^2 = \overline{F}(r),$$

en schrijf M_p voor het aantal oplossingen $(r, s) \in \mathbb{F}_p \times \mathbb{F}_p$ dat deze vergelijking heeft. Beter nog, omdat in de theorie van elliptische krommen er ook nog dat ene 'oneindige' punt O is, tellen we die er ook bij:

$$N_p := 1 + M_p.$$

Een triviale grens is dan $N_p \geq 1$, want minder dan 0 oplossingen kan de vergelijking niet hebben. Evenzo is er de bovengrens $N_p \leq 2p+1$: immers, die zou bereikt worden als elke $r \in \mathbb{F}_p$ een waarde $\overline{F}(r)$ zou opleveren die een kwadraat $\neq \overline{0}$ is. In 1933 bewees Helmut Hasse [8] veel sterkere grenzen voor N_p: er geldt

$$p + 1 - 2\sqrt{p} < N_p < p + 1 + 2\sqrt{p}.$$

Deze grenzen laten zien dat voor ieder priemgetal p, het getal N_p vrij dicht bij p ligt, dus voor ruwweg de helft van de mogelijke r zal $\overline{F}(r)$ een kwadraat zijn, voor de andere helft niet.

Als voorbeeld nemen we

$$F(x) := x^3 + x + 1.$$

Dat levert Tabel 1.

Het vermoeden

Laat

$$F(x) := x^3 + bx^2 + cx + d$$

een derdegraads veelterm zijn met gehele coëfficiënten b, c, d en zonder meervoudig nulpunt. Zoals in de vorige sectie gezien is, wordt dan voor elk priemgetal p een geheel getal N_p gedefinieerd. Definieer vervolgens, voor elke $P \geq 2$,

$$f(P) := \prod_{p \leq P} \frac{N_p}{p}.$$

Birch en Swinnerton-Dyer formuleren dan het volgende vermoeden.

$F(r)$ is voor oneindig veel rationale getallen r een kwadraat, precies dan als de functie $P \mapsto f(P)$ onbegrensd is.

De uitspraak is zelfs preciezer: voor $P \to \infty$ zou moeten gelden

$$f(P) \sim C \cdot \ln(P)^g,$$

voor een positieve constante C, waarin g de 'rang' (het maximale aantal onafhankelijke elementen) van de groep Γ is.

Als voorbeeld nemen we weer $F(x) = x^3 + x + 1$ en we tekenen de grafiek van $P \mapsto f(P)$ op drie verschillende intervallen. Zie Figuur 2. Zelfs in dit kleine voorbeeld blijkt al wel, dat het lastig te voorspellen is of de functie $P \mapsto f(P)$ begrensd is of niet. Om het vermoeden veel preciezer te maken, wordt de zogeheten L-reeks van de elliptische kromme gebruikt. De daartoe gebruikte manipulatie ziet er mogelijk nogal willekeurig uit, maar er zit een omvangrijke theorie achter die we hier laten voor wat het is.

Het polynoom $F(x)$ levert voor bijna alle priemgetallen p een polynoom $\overline{F}(x)$ over $\mathbb{F}_p$ op, dat eveneens geen meervoudige nulpunten heeft. De eindige verzameling priemgetallen waarvoor er wel een meervoudig nulpunt is, samen met nog het priemgetal 2, noemen we S. Het al eerder genoemde resultaat van Hasse zegt dat als $p \notin S$, dan is N_p ongeveer $p+1$. We schrijven daarom

$$N_p = p + 1 - a_p,$$

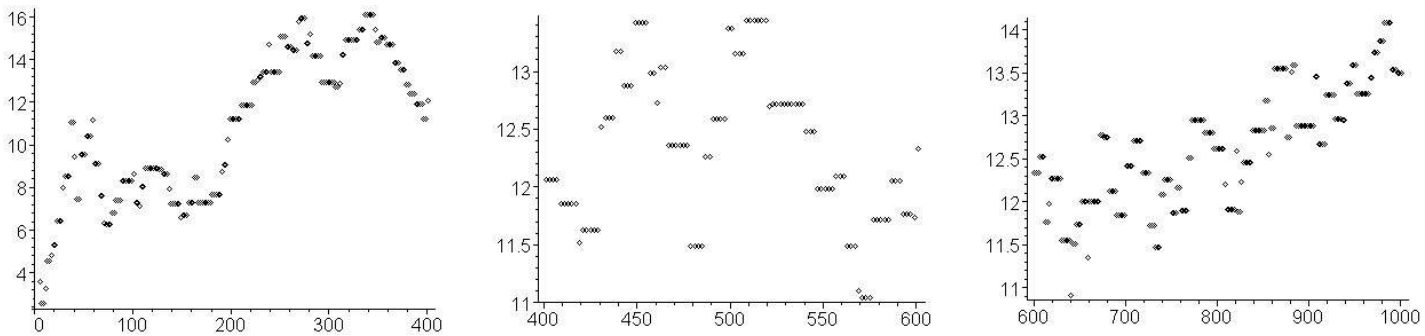
waarin dus a_p een geheel getal is dat voldoet aan $|a_p| < 2\sqrt{p}$. Dan is

$$\frac{N_p}{p} = 1 - \frac{a_p}{p} + p^{-1},$$

en dat is wat verkregen wordt door $s = 1$ in te

p	2	3	5	7	11	13	17	19	23	29	97	101	211	307	401
N_p	3	4	9	5	14	18	18	21	28	36	97	105	223	301	432

Tabel 1



Figuur 2

vullen in de uitdrukking

$$1 - a_p p^{-s} + p^{1-2s}.$$

In plaats van kijken of $f(P)$ willekeurig groot wordt als P groeit, kunnen we natuurlijk ook kijken of $\frac{1}{f(P)}$ naar 0 gaat als $P \to \infty$. Dat betekent, dat we een product nemen van factoren $\frac{p}{N_p}$, oftewel van

$$\frac{1}{1 - a_p p^{-s} + p^{1-2s}}$$

(en dan invullen $s = 1$). Dit leidt ertoe, te kijken naar het product

$$\prod_{p \notin S} \frac{1}{1 - a_p p^{-s} + p^{1-2s}}.$$

Voor $s > \frac{3}{2}$ blijkt dit product te convergeren (maar $s = 1$ is te klein!). De theorie voegt ook voor de resterende $p \in S$ een factor toe, en zo ontstaat een functie $s \mapsto L(s)$ die de "L-reeks" van de bijbehorende elliptische kromme heet.

Deze L-reeks geeft aanleiding tot de volgende versie van het vermoeden. Stel dat het mogelijk is, de functie $s \mapsto L(s)$ op een natuurlijke manier uit te breiden naar een op alle reële getallen gedefinieerde (willekeurig vaak differentieerbare) functie, die we ook $L(s)$ noemen. Zou dan gelden dat $L(1) = 0$, precies dan als er oneindig veel rationale r bestaan waarvoor $F(r)$ een kwadraat is? Beter nog, zou $L(s)$ voor $s = 1$ een nulpunt hebben waarvan de multipliciteit precies de "rang" van de groep Γ is? En zelfs hiermee eindigt het vermoeden niet; als die rang van de groep Γ gelijk is aan g, dan zou de g-de afgeleide van $L(s)$ voor $s = 1$ een door de uiteindelijke versie van het vermoeden voorspelde waarde moeten hebben. Maar het zou te ver voeren om de details daarvan hier te geven.

Rekenen met de functie $L(s)$ is voor iedereen mogelijk, bijvoorbeeld met behulp van programma Magma. Op http://magma.maths.usyd.edu.au/calc kan je ermee aan de slag. De regels

```
R<x>:=PolynomialRing(Rationals());
E:=EllipticCurve(x^3+x+1);
Evaluate(LSeries(E),1);
Evaluate(LSeries(E),1 : Derivative:=1);
```

leveren als uitvoer

```
0.00000000000000000000000000000000
1.78580949386920068702005538692
```

Dit geeft aan dat waarschijnlijk geldt $L(1) = 0$, dus als het vermoeden in dit geval klopt dan is $r^3 + r + 1$ een kwadraat voor oneindig veel verschillende rationale r. Ook zien we dat de afgeleide $L'(1) \neq 0$. En hadden we `Derivative:=5` geschreven, dan was de vijfde afgeleide in het punt $s = 1$ benaderd. Met het commando

```
MordellWeilShaInformation(E);
```

is overigens te zien dat er inderdaad oneindig veel rationale r zijn waarvoor $r^3 + r + 1$ een kwadraat is: de groep Γ van rationale punten op de kromme met vergelijking $y^2 = x^3 + x + 1$ blijkt te bestaan uit alle nP waarbij $P = (0, 1)$, en $nP \neq mP$ voor $n \neq m$.

En nu we toch met Magma bezig zijn: ook de voorbeelden in de voorgaande secties zijn daarmee uitstekend te verifiëren. Ter illustratie:

```
R<x>:=PolynomialRing(Rationals());
E:=EllipticCurve(x^3+2*x+1);
P:=E![0,1];
P;
(-3)*P;
(-3)^2*P;
```

Dit levert als antwoord de punten $P = (0, 1)$ en $-3P$ en $9P$, en het is gemakkelijk te controleren dat de eerste coördinaat daarvan overeenstemt met wat we al eerder hadden uitgerekend. Ook de eenvoudiger raaklijnmethode (dus $-2P$ en $4P$ en $-8P$, enzovoorts) uitrekenen is zo geen probleem.

Resultaten

Al is het vermoeden van Birch en Swinnerton-Dyer nog onbewezen, er is in de afgelopen vijftig jaar flink wat vooruitgang geboekt. Enkele van de meest spectaculaire ontdekkingen daarvan noemen we hier kort, al is dit overzicht verre van compleet.

Voor speciale $F(x)$, namelijk van de vorm $F(x) = x^3 + ax$ en ook de vorm $F(x) = x^3 + k$, konden Birch en Swinnerton-Dyer in hun artikelen al nagaan dat inderdaad een natuurlijke uitbreiding van de functie $L(s)$ tot alle getallen bestaat. Bovendien slaagden ze erin, in deze gevallen een formule voor $L(1)$ te vinden waarmee ze in concrete gevallen konden verifiëren of $L(1) = 0$ of $L(1) \neq 0$. Vrij snel na het bekendmaken van hun vermoeden werd het verregaand veralgemeniseerd. Allereerst werd (met hulp van Ian Cassels, Jean-Pierre Serre en John Tate) de L-functie in alle gevallen precies beschreven, inclusief factoren voor de 'slechte priemen'. Bovendien werd het vermoeden uitgebreid naar andere situaties dan alleen maar de rationale getallen, en ook naar hogere dimensies (niet alleen elliptische krommen).

In 1977 publiceerden John Coates en Andrew Wiles [4] (dezelfde die later onder meer de Laatste Stelling van Fermat in samenwerking met Richard Taylor zou bewijzen) een eerste flinke stap vooruit. Ze beperkten zich tot elliptische krommen met een extra eigenschap, namelijk dat ze 'complexe vermenigvuldiging' toelaten. De gevallen $F(x) = x^3 + ax$ en $F(x) = x^3 + k$ waar Birch en Swinnerton-Dyer al diverse voorbeelden van beschouwden, voldoen hieraan, maar ook bijvoorbeeld $F(x) = x^3 + 2kx^2 + 2k^2x$. Dankzij die complexe vermenigvuldiging was bekend dat de bijbehorende functie $L(s)$ inderdaad een natuurlijke uitbreiding tot alle getallen bezit. Het resultaat van Coates en Wiles is in dit geval

$$\Gamma \text{ heeft oneindig veel elementen } \Rightarrow L(1) = 0.$$

In 1982 gaf Dorian Goldfeld [6] een interpretatie van de constante C die optreedt in de eerste, naïeve versie $f(P) \sim C \ln(P)^g$ van het vermoeden. Ook Keith Conrad [5] schreef hierover een lang artikel.

Een groot probleem bij het vermoeden in het algemene geval was het uitbreiden van de functie $L(s)$. Dankzij ouder werk van Martin Eichler (1912–1992) en Goro Shimura (geboren in 1930) was dat echter geen probleem voor de 'modulaire elliptische krommen over $\mathbb{Q}$'. Onderzoek in de jaren tachtig ging daarom met name uit van de voorwaarde dat de elliptische kromme over $\mathbb{Q}$ die beschouwd werd, 'modulair' was. In 1986 publiceerden Dick Gross en Don Zagier een resultaat [7] over zogeheten 'Heegner punten' op modulaire elliptische krommen. Ze lieten zien dat zo'n punt P oneindig veel verschillende veelvouden nP heeft, precies dan als $L(1) = 0$ en $L'(1) \neq 0$. En dus

$$L(1) = 0 \text{ en } L'(1) \neq 0$$
$$\Rightarrow \Gamma \text{ heeft oneindig veel elementen}$$

voor modulaire elliptische krommen. Kort daarop completeerde de Russische wiskundige Victor Kolyvagin [9] dit werk van Gross en Zagier met de twee stellingen

$$L(1) \neq 0 \quad \Rightarrow \quad \Gamma \text{ is eindig}$$

en ook

$$L(1) = 0 \text{ en } L'(1) \neq 0 \quad \Rightarrow \quad \Gamma \text{ heeft rang } 1.$$

Ook hier onder de aanname dat het om modulaire elliptische krommen over $\mathbb{Q}$ gaat.

In 1998 (de publicatie volgde in 2001) vulden Christophe Breuil, Brian Conrad, Fred Diamond en Richard Taylor [3] het werk dat een paar jaar eerder was begonnen door Andrew Wiles [13,14]: het resultaat is, dat *elke* elliptische kromme over $\mathbb{Q}$ 'modulair' is. Daarmee gelden dus de stellingen van Gross en Zagier, en die van Kolyvagin, voor alle elliptische krommen over $\mathbb{Q}$. Toch is hiermee het originele vermoeden van Birch en Swinnerton-Dyer nog zeker niet bewezen. De Amerikaan Chris Skinner geeft de afgelopen maanden (zomer 2013) voordrachten over de omkering van Kolyvagins stellingen: als Γ eindig is, of rang 1 heeft, hoe zou je dan kunnen aantonen dat $L(1) \neq 0$ respectievelijk $L'(1) \neq 0$? Skinner heeft ideeën hoe je dit wellicht kan aanpakken. En zodra $L(s)$ in $s = 1$ een nulpunt heeft met multipliciteit minstens 2, dan is uit de tot nu toe bewezen stellingen nog niet te concluderen of Γ oneindig is. Ook omgekeerd, bevat Γ oneindig veel elementen, dan volgt dat $L(s)$ in $s = 1$ een nulpunt heeft, maar over de multipliciteit ervan valt nog niks te concluderen. Kortom, het laatste woord over het vermoeden van Birch en Swinnerton-Dyer is nog zeker niet gezegd!

Referenties

1 B.J. Birch en H.P.F. Swinnerton-Dyer, Notes on elliptic curves. I, *J. reine angew. Math.* 212 (1963), 7–25.

2 B.J. Birch en H.P.F. Swinnerton-Dyer, Notes on elliptic curves. II, *J. reine angew. Math.* 218 (1965), 79–108.

3 C. Breuil, B. Conrad, F. Diamond en R. Taylor, On the modularity of elliptic curves over **Q**: wild 3-adic exercises, *J. Amer. Math. Soc.* 14 (2001), 843–939.

4 J. Coates en A. Wiles, On the Conjecture of Birch and Swinnerton-Dyer, *Invent. Math.* 39 (1977), 223–252.

5 K. Conrad, Partial Euler products on the critical line, *Canad. J. Math.* 57 (2005), 267–297.

6 D. Goldfeld, Sur les produits eulériens attachés aux courbes elliptiques, *C.R.Acad. Sci. Paris Sér. I Math.* 294 (1982), 471–474.

7 B.H. Gross en D.B. Zagier, Heegner points and derivatives of L-series, *Invent. Math.* 84 (1986), 225–320

8 H. Hasse, Beweis des Analogons der Riemannschen Vermutung für die Artinschen und F.K. Schmidtschen Kongruenzzetafunktionen in gewissen elliptischen Fällen, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen* (1933), 253–262.

9 V.A. Kolyvagin, Finiteness of $E(Q)$ and Ш(E, Q) for a subclass of Weil curves (Russisch), *Izv. Akad. Nauk SSSR Ser. Mat.* 52 (1988), 522–540, 670–671; Engelse vertaling in *Math. USSR-Izv.* 32 (1989), 523–541.

10 L.J. Mordell, On the rational solutions of the indeterminate equations of the third and fourth degrees, *Proc. Cambridge Phil. Soc.* 21 (1922), 179–192

11 H. Poincaré, Sur les propriétés arithmétiques des courbes algébriques, *J. de math. pures et appl. 5e série 7* (1901), 161–234.

12 J.H. Silverman en J.T. Tate, *Rational Points on Elliptic Curves*, Undergraduate Texts in Mathematics, Springer, New York, 1992.

13 R. Taylor en A. Wiles, Ring-theoretic properties of certain Hecke algebras, *Ann. of Math.* 141 (1995), 553–572.

14 A. Wiles, Modular elliptic curves and Fermat's last theorem, *Ann. of Math.* 141 (1995), 443–551.

15 A. Wiles, The Birch and Swinnerton-Dyer Conjecture, www.claymath.org/millennium/Birch_and_Swinnerton-Dyer_Conjecture/birchswin.pdf, 2000.

16 M.V. Wilkes, Edsac 2, *IEEE Annals of the History of Computing* 14 (1992), 49–56.