Sander R. Dahmen
Afdeling Wiskunde
Vrije Universiteit Amsterdam
s.r.dahmen@vu.nl

Arno Kret
Korteweg-de Vries Instituut
Universiteit van Amsterdam
a.l.kret@uva.nl

Evenement Abelprijs 2016

Andrew Wiles en de Abelprijs

Op 24 mei 2016 ontving Sir Andrew J. Wiles de Abelprijs 'for his stunning proof of Fermat's Last Theorem by way of the modularity conjecture for semistable elliptic curves, opening a new era in number theory'. In dit artikel schetsen Sander Dahmen en Arno Kret wat Wiles bewezen heeft en laten zien wat voor prachtig onderzoek er uit zijn werk voortgekomen is.

De Laatste Stelling van Fermat luidt als volgt.

Stelling 1. *Laat x, y en z gehele getallen ongelijk 0 zijn en $n \geq 3$ geheel. Dan geldt*

$$x^n + y^n \neq z^n.$$

Wiles leverde het finale ingrediënt voor deze stelling door de modulariteit van grote klassen van elliptische krommen te laten zien, wat feitelijk de oplossing is van een deel van de Langlands-vermoedens. Hiermee werd de weg geopend tot het verkrijgen van nog veel verder reikende modulariteitsresultaten en bijbehorende gevolgen, inclusief het oplossen van vele andere diofantische vergelijkingen. In dit artikel zullen we ons vooral richten op de wiskundige kant van het verhaal. Zie ook de referenties [2, 5, 13] die gericht zijn op een breed publiek, en voor het originele werk zie de artikelen [29, 30]. We gaan eerst uitleggen wat modulariteit van elliptische krommen betekent. Hiervoor komen modulaire vormen en elliptische krommen uitgebreid aan bod. Daarna geven we verbanden met andere problemen in de wiskunde, zoals het vermoeden van Birch en Swinnerton-Dyer, de Langlands-vermoedens en problemen in de theorie van diofantische vergelijkingen.

In 1967 schreef Langlands een brief aan André Weil [16], waarin hij een aantal vermoedens formuleerde. Zijn vermoedens, nu bekend als de *Langlands-vermoedens*, brengen de getaltheorie en representatietheorie met elkaar in verband. Bovendien zijn ze gerelateerd aan en kunnen ze gecombineerd worden met andere stellingen en vermoedens uit de getaltheorie en algebraïsche meetkunde, zoals klassenlichamentheorie, de Riemann-hypothese, het Sato–Tate-vermoeden, het Hasse–Weil-vermoeden, het Fontaine–Mazur-vermoeden en het eerder genoemde Birch–Swinnerton-Dyer-vermoeden.

Foto: Klaus Barner / CC BY-SA

Andrew Wiles voor het standbeeld van Pierre de Fermat in Beaumont-de-Lomagne, 1995

Voordat Wiles in 1994 zijn stelling bewees, was buiten het geval van eendimensionale representaties weinig van de vermoedens bekend in hogere dimensies. Men kon zien dat de vermoedens van Langlands consistent waren met andere vermoedens (in de zin dat er geen tegenspraak is), maar men kon niet veel bewijzen of narekenen. Hoewel het eendimensionale geval volgt uit klasselichamentheorie, is dit geval ook enorm gedegenereerd. Dus eigenlijk was het helemaal niet duidelijk of je Langlands' vermoedens moest geloven of niet. Misschien was dit ook de reden dat Weil nooit gereageerd heeft op de brief van Langlands. Het werk van Wiles heeft grote invloed gehad op de Langlands-vermoedens, omdat het een substantiële klasse van bewezen gevallen toevoegt, die daarvoor als onbereikbaar werden gezien

met de toenmalige wiskundige technieken. Na Wiles hebben veel wiskundigen het argument van Wiles uitgebreid, en op deze manier meer delen van de Langlandsvermoedens bewezen.

Wiskundige objecten: de spelers

We zullen eerst twee basisobjecten introduceren, namelijk elliptische krommen en modulaire vormen. We beperken ons hier tot enkele basale definities en eigenschappen. Gedegen introducties zijn op veel plekken te vinden, bijvoorbeeld de tekstboeken [25] en [8]. Vervolgens gaan we in op de relatie tussen de twee spelers in het beschrijven van modulariteit van elliptische krommen. Ten slotte bespreken we nog even kort de rol hiervan in het beroemde vermoeden van Birch en Swinnerton-Dyer.

Elliptische krommen

Laat K een lichaam zijn, zoals bijvoorbeeld $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ of $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ met p een priemgetal. Een *Weierstrass-vergelijking* over K is een vergelijking van de vorm

$$y^2 + dxy + ey = x^3 + ax^2 + bx + c \qquad (1)$$

met $a,b,c,d,e \in K$. Zo'n vergelijking definieert een algebraïsche kromme E over K. Als de karakteristiek van K niet 2 is (wat betekent dat $1+1 \neq 0$ in K, waarmee bijvoorbeeld $K = \mathbb{F}_2$ afvalt), kun je een dergelijke vergelijking altijd omschrijven naar een vergelijking waarbij $d = e = 0$. Laten we dit nu voor het gemak aannemen, dan worden de formules eenvoudiger. We hebben nu een vergelijking voor E van de vorm $y^2 = f(x)$, waarbij f een monisch kubisch polynoom $f = x^3 + ax^2 + bx + c$ is. Hoewel de vergelijking $y^2 = f(x)$ suggereert dat we gewoon in het affiene vlak K^2 werken, beschouwen we de kromme E ingebed in het projectieve vlak (en compleet, wat concreet betekent dat we één formeel extra punt aan de kromme toevoegen: het punt op oneindig). De K-rationale punten op E, genoteerd als $E(K)$, zijn de punten $(x,y) \in K^2$ in het vlak K^2 die aan bovenstaande vergelijking voldoen, samen met het punt op oneindig O. De niet-singuliere punten in

$$E(K) = \{(x,y) \in K^2 : y^2 = x^3 + ax^2 + bx + c\} \cup \{O\}$$

vormen op natuurlijke wijze een groep, met O als neutraal punt, door middel van de beroemde 'koorde- en raaklijn-constructie'. Dit is ook de reden om het punt O toe te voegen.

Als E niet-singulier is, spreken we van een *elliptische kromme* over K. We definiëren de discriminant $\Delta(E)$ van de Weierstrass-vergelijking door $\Delta(E) = 2^4 \cdot \Delta(f)$. Hierin is $\Delta(f)$ de discriminant van het kubische polynoom f, welke in het geval dat $a = 0$ simpelweg gegeven wordt door $-4b^3 - 27c^2$. Er geldt dat E niet-singulier is (en dus een elliptische kromme geeft) dan en slechts dan als $\Delta(E) \neq 0$. Twee krommen $E : y^2 = f(x)$ en $E' : y'^2 = g(x')$ zijn isomorf dan en slechts dan als ze gerelateerd zijn via een variabelentransformatie van de vorm

$$x = u^2 x' + r, \quad y = u^3 y' \qquad (2)$$

waarbij $r,u \in K$ met $u \neq 0$. In dat geval zijn de discriminanten gerelateerd via $\Delta(E) = u^{12}\Delta(E')$. Als E singulier is, dat wil zeggen dat $\Delta(E) = 0$, dan liggen de singuliere punten op de x-as (nog steeds aangenomen dat $d = e = 0$ en de karakteristiek van K niet 2 is), en heeft f meervoudige nulpunten, welke automatisch in K liggen als K een deellichaam van $\mathbb{C}$ of een eindig lichaam is. Er zijn nu twee mogelijkheden. Als $f(x)$ twee verschillende nulpunten heeft waarvan één dus een dubbel nulpunt is, dan zeggen we dat E een *knooppunt* heeft. Als $f(x)$ één nulpunt van orde drie heeft, dan zeggen we dat E een *spits* heeft.

Laat nu E een elliptische kromme over $\mathbb{Q}$ zijn, gegeven door een Weierstrass-vergelijking. Er zijn oneindig veel andere Weierstrass-vergelijkingen (1) die door variabelentransformaties, waarvan (2) speciale gevallen zijn, uit de oorspronkelijke vergelijking voor E verkregen kunnen worden. Een Weierstrass-vergelijking uit deze familie wordt *minimaal* genoemd als alle coëfficiënten in $\mathbb{Z}$ liggen en de absolute waarde van de discriminant minimaal is (onder de Weierstrass-vergelijkingen met coëfficiënten in $\mathbb{Z}$ in de familie). Voor zo'n minimale vergelijking kunnen we de coëfficiënten van E reduceren modulo een priemgetal p en zo een Weierstrass-vergelijking voor de reductie modulo p krijgen: de kromme $\tilde{E}$ over $\mathbb{F}_p$. Het reductietype en het aantal punten $\#\tilde{E}(\mathbb{F}_p)$ hangt niet van de gekozen minimale Weierstrass-vergelijking af.

Een belangrijke invariant voor elliptische krommen E over $\mathbb{Q}$ is de *conductor* van E, genoteerd als $N(E)$. De volledige definitie is behoorlijk subtiel (zie Hoofdstuk 4, §10 van [26]). We geven hier een gedeeltelijke definitie. Voor elk priemgetal p beschouwen we de conductor-exponent

$$e_p(E) := \begin{cases} 0 & \text{als } E \text{ niet singulier is mod } p \\ 1 & \text{als } E \text{ een knooppunt heeft mod } p \\ 2+\delta_p & \text{als } E \text{ een spits heeft mod } p \end{cases}$$

waarbij $\delta_p \in \mathbb{Z}_{\geq 0}$. Verder geldt $\delta_2 \leq 6$, $\delta_3 \leq 3$ en $\delta_p = 0$ als $p \geq 5$. Nu is de conductor $N(E) := \prod_p p^{e_p(E)}$. In het bijzonder, als de reducties van E modulo 2 en modulo 3 geen spits oplevert, dan legt bovenstaande $N(E)$ volledig vast. Als voor geen enkel priemgetal p de reductie van E modulo p een spits geeft, dan noemen we E *semi-stabiel*; dit komt op hetzelfde neer als dat $N(E)$ kwadraatvrij is.

Voor een elliptische kromme $E/\mathbb{Q}$ definiëren we $a_p(E) := p + 1 - \#\tilde{E}(\mathbb{F}_p)$ voor elk priemgetal p. De L-functie van E is nu

$$L_E(s) := \prod_{p \nmid N(E)} (1 - a_p(E)p^{-s} + p^{1-2s})^{-1} \cdot \prod_{p \mid N(E)} (1 - a_p(E)p^{-s})^{-1}$$

waar de producten over priemgetallen p lopen. Het is klassiek bekend dat $|a_p(E)| \leq 2\sqrt{p}$ voor alle priemen p en dat de Dirichlet-reeks corresponderend met $L_E(s)$ convergeert naar een holomorfe functie voor $s \in \mathbb{C}$ met reëel deel strikt groter dan $3/2$.

Veel van het bovenstaande voor $K = \mathbb{Q}$ kunnen we generaliseren naar getallenlichamen K. Dit betekent dat K een eindige lichaamsuitbreiding van $\mathbb{Q}$ is, zoals bijvoorbeeld $\mathbb{Q}(i) = \{a + bi : a,b \in \mathbb{Q}\}$, $\mathbb{Q}(\sqrt[3]{2}) = \{a + b\sqrt[3]{2} + c\sqrt[3]{2}^2 : a,b,c \in \mathbb{Q}\}$ of $\mathbb{Q}(\alpha) = \{a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}\}$ waarbij α een nulpunt (laten we zeggen in $\mathbb{C}$) is van een polynoom van graad n met rationale coëfficiënten dat irreducibel over $\mathbb{Q}$ is. Dit laatste is het generieke voorbeeld en omvat alle getallenlichamen (ingebed in $\mathbb{C}$). Binnen zo'n getallenlichaam K definieert men de *ring van gehelen* $\mathcal{O}_K$ als die elementen van K die nulpunt zijn van een monisch polynoom met coëfficiënten in $\mathbb{Z}$. Om in te zien dat $\mathcal{O}_K$ inderdaad een ring is, moet men nog een beetje werk verzetten. De ring van gehelen in $\mathbb{Q}$ is $\mathbb{Z}$ en de rol die $\mathbb{Z}$ binnen $\mathbb{Q}$ speelt, is analoog aan de rol die $\mathcal{O}_K$ binnen K speelt. Voor bijvoorbeeld $K = \mathbb{Q}(i)$ hebben we eenvoudig dat de ring van gehelen $\mathcal{O}_K = \mathbb{Z}[i] = \{a + bi : a,b \in \mathbb{Z}\}$ de ring van Gaussische gehelen is. Voor een elliptische kromme E over een getallenlichaam K kunnen we weer een Weierstrass-vergelijking opschrijven, maar dan met coëfficiënten in $\mathcal{O}_K$. Voor elk priem-

ideaal $\wp \subset \mathcal{O}_K$ ongelijk 0 kunnen we nu de reductie van de Weierstrass-vergelijking nemen modulo $\wp$ om zo een kromme $\tilde{E}$ over een eindig lichaam $\mathbb{F}_\wp := \mathcal{O}_K/\wp$ te krijgen. Is de vergelijking gekozen zodat de macht van $\wp$ minimaal is in (het ideaal voortgebracht door) de discriminant van de vergelijking (weer onder alle relevante variabelentransformaties), dan is het reductietype weer uniek bepaald evenals het gehele getal $a_\wp(E) := N(\wp)+1-\#\tilde{E}(\mathbb{F}_\wp)$, waarbij $N(\wp) := \#\mathbb{F}_\wp$. De elliptische kromme E heeft, net als in het geval waar $K=\mathbb{Q}$, een *conductor* $\mathcal{N}$, die nu een ideaal is in $\mathcal{O}_K$. De kromme E heeft niet-singuliere reductie modulo precies die priemidealen $\wp \subset \mathcal{O}_K$ die $\mathcal{N}$ niet delen. De L-functie gekoppeld aan E is

$$L_E(s) := \prod_{\wp \nmid \mathcal{N}} (1-a_\wp(E)N(\wp)^{-s}+N(\wp)^{1-2s})^{-1} \cdot \prod_{\wp \mid \mathcal{N}} (1-a_\wp(E)N(\wp)^{-s})^{-1}.$$

Modulaire vormen

Klassieke (of elliptische) modulaire vormen zijn holomorfe functies op het complexe bovenhalfvlak $\mathbb{H} := \{x+yi \in \mathbb{C} : y>0\}$ met bepaalde transformatie- en groei-eigenschappen. Spitsvormen zijn modulaire vormen met nog een wat sterkere groei-eigenschap en we beperken ons in dit artikel tot deze klasse. *In de rest van deze subparagraaf zullen k en N staan voor positieve gehele getallen*. Om alles precies te maken, introduceren we eerst een aantal groepen. Namelijk

$$\mathrm{GL}_2^+(\mathbb{R}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| a,b,c,d \in \mathbb{R},\ ad-bc>0 \right\},$$

een discrete ondergroep hiervan

$$\mathrm{SL}_2(\mathbb{Z}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| a,b,c,d \in \mathbb{Z},\ ad-bc=1 \right\}$$

en een ondergroep (van eindige index) van laatstgenoemde groep

$$\Gamma(N) := \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) \middle| \gamma \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} (\mathrm{mod}\ N) \right\}.$$

Een *congruentieondergroep van* $\mathrm{SL}_2(\mathbb{Z})$ is een ondergroep Γ van $\mathrm{SL}_2(\mathbb{Z})$ die $\Gamma(N)$ bevat voor een N. De kleinste N voor zo'n Γ waarvoor $\Gamma(N) \subset \Gamma$ heet het *niveau* van Γ. Men gaat eenvoudig na dat voor $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2^+(\mathbb{R})$ en $\tau \in \mathbb{H}$ de formule

$$\gamma\tau := \frac{a\tau+b}{c\tau+d} \tag{3}$$

een *actie* van $\mathrm{GL}_2^+(\mathbb{R})$ op $\mathbb{H}$ definieert (oftewel, $\gamma\tau \in \mathbb{H}$, $(\gamma\gamma')\tau = \gamma(\gamma'\tau)$ en $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\tau = \tau$). Dezelfde formule geeft dus ook een actie van elke ondergroep van $\mathrm{GL}_2^+(\mathbb{R})$, zoals de congruentieondergroepen, op $\mathbb{H}$.

Laat Γ een congruentieondergroep van $\mathrm{SL}_2(\mathbb{Z})$ van niveau N zijn. We definiëren nu een *spitsvorm voor* Γ *van gewicht* k als een holomorfe (dat wil zeggen complex differentieerbare) functie $f : \mathbb{H} \to \mathbb{C}$ met de transformatie-eigenschap (voor alle $\tau \in \mathbb{H}$)

$$f\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\tau\right) = (c\tau+d)^k f(\tau) \quad \text{voor alle} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \tag{4}$$

en groei-eigenschap

$$\tau = x+yi \mapsto y^{k/2}\,|f(\tau)| \text{ is begrensd op } \mathbb{H}. \tag{5}$$

Men gaat gemakkelijk na dat de spitsvormen voor Γ van gewicht k een $\mathbb{C}$-vectorruimte vormen (onder de gebruikelijke operaties), die we noteren als $S_k(\Gamma)$. Met wat meer werk is ook aan te tonen dat $S_k(\Gamma)$ eindigdimensionaal is.

Laat $f : \mathbb{H} \to \mathbb{C}$ nu een holomorfe functie zijn met transformatie-eigenschap (4). Voor $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & N \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma$ specialiseert (4) tot $f(\tau+N) = f(\tau)$ voor alle $\tau \in \mathbb{H}$, oftewel f is periodiek met periode M een deler van N. Samen met de holomorfie van f geeft dit direct dat f een Fourier-ontwikkeling heeft: $f(\tau) = \sum_{n=-\infty}^{\infty} a_n \exp(2\pi i\tau n/M)$ met alle $a_n \in \mathbb{C}$ uniek bepaald. De groei-eigenschap (5) impliceert dat $f(x+yi) \to 0$ als $y \to \infty$ (aangezien $k>0$ per aanname), wat equivalent is met $a_n = 0$ voor alle $n \leq 0$. Voor $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ is dit laatste equivalent met (5), maar voor algemene Γ is dit niet noodzakelijk waar. Namelijk, de formule (3) definieert een actie van Γ op $\mathbb{P}^1(\mathbb{Q}) = \mathbb{Q} \cup \{\infty\}$, met een eindig aantal banen, de zogenaamde *spitsen* van Γ. Bij elke spits kan men een bepaalde Fourier-ontwikkeling in $\exp(2\pi i\tau/N)$ kiezen, en (5) is equivalent met dat al deze Fourier-ontwikkelingen enkel positieve machten van $\exp(2\pi i\tau/N)$ bevatten. Algemene modulaire vormen mogen ook nog een constante term in zulke Fourier-ontwikkelingen hebben, maar voor spitsvormen moet per definitie bij elke spits de constante dus nul zijn. Hier komt de naam spitsvorm vandaan.

De belangrijkste congruentieondergroepen om het werk van Wiles te beschrijven, zijn gegeven door

$$\Gamma_0(N) := \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) \middle| \gamma \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} (\mathrm{mod}\ N) \right\},$$

hierbij betekent $*$ dat elke mogelijke waarde toegestaan is, dus de conditie op $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is simpelweg dat $N|c$. Voor later gebruik introduceren we ook

$$\Gamma_1(N) := \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) \middle| \gamma \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} (\mathrm{mod}\ N) \right\}.$$

Merk op dat $\Gamma(N) \subset \Gamma_1(N) \subset \Gamma_0(N) \subset \mathrm{SL}_2(\mathbb{Z})$ en verder dat alle inclusies gelijkheden zijn dan en slechts dan als $N=1$. Laat $f \in S_k(\Gamma_i(N))$ (met $i \in \{0,1\}$). Aangezien $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_i(N)$ is f periodiek met periode 1. Schrijven we q voor de functie gegeven door $q(\tau) = \exp(2\pi i\tau)$, dan krijgen we zo de zogenaamde q-expansie

$$f = \sum_{n=1}^{\infty} a_n(f)\, q^n. \tag{6}$$

Bovenstaande notatie $a_n(f)$ voor de (uniek bepaalde) Fourier-coëfficiënten van $f \in S_k(\Gamma_i(N))$ zullen we in de rest van dit artikel aanhouden.

Laten we eens naar het eenvoudigste geval $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ kijken. Het kleinste gewicht k waarvoor $S_k(\mathrm{SL}_2(\mathbb{Z})) \neq \{0\}$, is $k=12$. In dat geval is de ruimte eendimensionaal en voortgebracht door de discriminant functie

$$\Delta := q\prod_{n=1}^{\infty}(1-q^n)^{24} = q-24q^2+252q^3-1472q^4+4830q^5-6048q^6+\cdots$$

Deze functie kan worden verkregen als discriminant van een op natuurlijke wijze geparametriseerde familie van elliptische krommen, vandaar de naam. De coëfficient $a_n(\Delta)$ van q^n in bovenstaande q-expansie van Δ wordt traditioneel als $\tau(n)$ genoteerd. De functie $n \mapsto \tau(n)$ heet *Ramanujans tau-functie* en in 1916 vermoedde Ramanujan een aantal opmerkelijke eigenschappen ervoor, namelijk

- τ is multiplicatief, oftewel $\tau(nm) = \tau(n)\tau(m)$ voor alle $n,m \in \mathbb{Z}_{>0}$ met $\mathrm{ggd}(m,n)=1$;
- $\tau(p^r) = \tau(p)\tau(p^{r-1}) - p^{11}\tau(p^{r-2})$ voor alle priemgetallen p en $r \in \mathbb{Z}_{\geq 2}$;
- $|\tau(p)| \leq 2p^{11/2}$ voor alle priemgetallen p.

De eerste twee eigenschappen zijn al in 1917 bewezen door Mordell en de laatste is in 1974 door Deligne bewezen als (niet-triviaal) gevolg van zijn bewijs van de beroemde Weil-vermoedens. Voor dit laatste kreeg Deligne in 1978 de Fieldsmedaille, verder heeft hij onder meer ook de Abel-

Foto: abelprize.no, Audun Braastad

Andrew Wiles tijdens de prijsuitreiking van de Abelprijs op 24 mei 2016 in Oslo

prijs ontvangen in 2013; voor meer info zie [18]. Er is ook nog veel onbekend over Ramanujans tau-functie. Zo weten we bijvoorbeeld niet of $\tau(n)$ gelijk aan nul kan zijn voor een $n \in \mathbb{Z}_{>0}$. Men vermoedt van niet, dit staat ook wel bekend als Lehmers vermoeden, zie [17]. Het is momenteel bekend dat $\tau(n) \neq 0$ voor $n < 8 \times 10^{23}$, maar een bewijs dat dit voor alle n geldt, lijkt voorlopig nog niet in zicht.

De eerste twee eigenschappen kunnen het best verklaard worden met behulp van *Hecke-operatoren*. Voor elke priem p hebben we een lineaire operator

$$T_p : S_k(\Gamma) \to S_k(\Gamma).$$

Een natuurlijke manier om deze in te voeren is met behulp van representanten $\gamma_1, \gamma_2, \ldots, \gamma_{d_p}$ van nevenklassen van Γ in een decompositie

$$\Gamma \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \Gamma = \coprod_{j=1}^{d_p} \Gamma \gamma_j$$

(waarbij de eindigheid van d_p niet geheel triviaal is). Voor $f \in S_k(\Gamma)$ definiëren we nu

$$\begin{aligned} T_p(f) : \mathbb{H} &\to \mathbb{C} \\ \tau &\mapsto \sum_{j=1}^{d_p} \frac{\det(\gamma_j)^{k-1}}{(c\tau + d)^k} f(\gamma_j \tau). \end{aligned}$$

Er kan eenvoudig aangetoond worden dat dit niet van de keuze van de representanten $\{\gamma_j\}_{j=1}^{d_p}$ afhangt en dat $T_p(f) \in S_k(\Gamma)$. Dat T_p lineair is, volgt direct uit de definities.

Het *Petersson-inproduct* op $S_k(\Gamma)$ wordt gegeven door

$$\langle f, g \rangle := \int_{\Gamma \backslash \mathbb{H}} f(\tau) \overline{g(\tau)} y^k \, d\nu(\tau)$$

waarbij $\tau = x + yi$ en $d\nu(\tau) := dxdy/y^2$ de hyperbolische maat op $\mathbb{H}$ is. Deze is namelijk invariant onder de actie van $\mathrm{GL}_2^+(\mathbb{R})$ op $\mathbb{H}$. Convergentie van de integraal volgt in essentie uit de groei-eigenschap (5).

We zullen ons vanaf nu beperken tot $\Gamma = \Gamma_0(N)$. Het mooie is dat alle T_p's onderling commuteren en dat het Hermitische operatoren zijn voor $p \nmid N$ ten aanzien van het Petersson-inproduct op $S_k(\Gamma_0(N))$. Voor elke positieve echte deler M van N en positieve deler d van N/M geldt voor $f \in S_k(\Gamma_0(M))$ dat $g \in S_k(\Gamma_0(N))$ met $g(\tau) := f(d\tau)$. Zo'n g wordt een *oudvorm* van $S_k(\Gamma_0(N))$ genoemd. De deelruimte van $S_k(\Gamma_0(N))$ omspannen door de oudvormen heet de *oudruimte*, genoteerd $S_k(\Gamma_0(N))^{\mathrm{oud}}$. Het orthogonaal complement met betrekking tot het Petersson-inproduct heet de *nieuwruimte*, genoteerd

$$S_k(\Gamma_0(N))^{\mathrm{nieuw}} := \left(S_k(\Gamma_0(N))^{\mathrm{oud}}\right)^{\perp}.$$

De Hecke-operatoren T_p behouden de beide ruimtes en op de nieuwruimte zijn alle T_p's Hermitisch. De spectraalstelling uit de lineaire algebra geeft nu dat $S_k(\Gamma_0(N))^{\mathrm{nieuw}}$ een gemeenschappelijke basis van eigenvectoren, genaamd *Hecke-eigenvormen*, heeft ten aanzien van alle T_p's. Het blijkt dat voor zulke eigenvormen f automatisch $a_1(f) \neq 0$ en we kunnen dus normaliseren tot $a_1(f) = 1$. Zo'n genormaliseerde Hecke-eigenvorm in de nieuwruimte noemen we een *nieuwvorm*. Deze nieuwvormen spelen een belangrijke rol in de beschrijving van Wiles' resultaat. We noemen nog dat voor een nieuwvorm f geldt dat $T_p(f) = a_p(f) f$ voor alle priemen p en dat f volledig vast ligt door alle $a_p(f)$ met p priem. De nieuwvormen in $S_k(\Gamma_0(N))$ vormen zo een canonieke basis voor $S_k(\Gamma_0(N))^{\mathrm{nieuw}}$. Aan elke $f \in S_k(\Gamma_0(N))$ kunnen we een L-reeks koppelen middels de Dirichlet-reeks

$$L_f(s) := \sum_{n=1}^{\infty} \frac{a_n(f)}{n^s}. \tag{7}$$

Deze convergeert voor $s \in \mathbb{C}$ met $\Re s > 1 + k/2$ tot een holomorfe functie. Voor een nieuwvorm $f \in S_k(\Gamma_0(N))$ wordt de L-reeks ook gegeven door het Euler-product

$$\begin{aligned} L_f(s) = &\prod_{p \nmid N} (1 - a_p(f) p^{-s} + p^{k-1-2s})^{-1} \\ &\cdot \prod_{p \mid N} (1 - a_p(f) p^{-s})^{-1} \end{aligned}$$

waarbij p over de priemgetallen loopt. Vergelijking met (7) geeft weer hoe alle Fourier-coëfficiënten van f recursief bepaald worden door de $a_p(f)$ met p priem. In het bijzonder geeft dit zo voor $\Delta \in S_{12}(\Gamma_0(1))$ de eerste twee eerder genoemde eigenschappen van τ. Laat verder voor een nieuwvorm f,

$$\xi_f(s) := N^{s/2} (2\pi)^{-s} \Gamma(s) L_f(s).$$

Deze functie voldoet aan de functionaalvergelijking

$$\xi_f(s) = \pm \xi_f(k - s)$$

voor alle $s \in \mathbb{C}$ en een teken $\pm$ dat alleen van f afhangt (de eigenwaarde van de bijbehorende Fricke-involutie). In het bijzonder heeft $L_f(s)$ een analytische voortzetting tot het hele complexe vlak.

Voor $\Gamma_1(N)$ geldt een soortgelijk verhaal als boven, alleen moeten we dan nog een extra type (eenvoudige) Hecke-operatoren introduceren en de Hermitische eigenschap wordt vervangen door 'normaal', wat volstaat voor de benodigde spectraalstelling. Uiteindelijk krijgen we dan ook *nieuwvormen* voor $S_k(\Gamma_1(N))$ die op een canonieke wijze een basis voor $S_k(\Gamma_1(N))^{\mathrm{nieuw}}$ vormen.

Modulariteit

Definitie 2. Een elliptische kromme E over $\mathbb{Q}$ heet *modulair* als er een nieuwvorm $f \in S_2(\Gamma_0(N(E)))$ is zodat $a_p(E) = a_p(f)$ voor alle priemgetallen p.

Merk op dat dit laatste direct equivalent is met $L_E(s) = L_f(s)$. Er zijn vele andere equivalente definities, ook meer meetkundig georiënteerde, maar daar gaan we nu niet verder op in. Het vermoeden dat alle elliptische krommen over $\mathbb{Q}$ modulair zijn, stond bekend als het vermoeden van Shimura–Taniyama–Weil (en onder vele andere namen, inclusief de meeste permutaties van niet lege deelverzamelingen van de drie namen). Inmiddels is het volledig bewezen.

Stelling 3 (Modulariteit). *Alle elliptische krommen over $\mathbb{Q}$ zijn modulair.*

Modulariteit van alle semi-stabiele elliptische krommen over $\mathbb{Q}$ is in 1994 bewezen door Wiles [30], met hulp van Taylor [29]. Dit volstond voor het voltooien van een bewijs voor de Laatste Stelling van Fermat. Vervolgens werden de methoden van Wiles en Taylor gegeneraliseerd, waarbij een aantal gecompliceerde technische problemen overwonnen werden, totdat uiteindelijk in 1999 de volledige modulariteitsstelling hierboven werd bewezen door Breuil, Conrad, Diamond en Taylor [3].

Beschouw als voorbeeld van modulariteit de elliptische kromme E gegeven door de minimale Weierstrass-vergelijking

$$y^2 + y = x^3 - x^2.$$

De kromme heeft enkel voor $p = 11$ een singuliere reductie, waar het een knooppunt heeft. Dit geeft voor de conductor $N(E) = 11$. Verder hebben we heel concreet voor alle priemen p,

$$\#\tilde{E}(\mathbb{F}_p) = \#\{(x,y) \in \mathbb{F}_p^2 : y^2 + y = x^3 - x^2\} + 1.$$

De extra $+1$ komt van het 'punt op oneindig'. We kunnen een klein tabelletje maken van $\#\tilde{E}(\mathbb{F}_p)$ en dus ook van $a_p(E) = p + 1 - \#\tilde{E}(\mathbb{F}_p)$:

p	2	3	5	7	11	13	17	19	$\cdots$	2017
$\#\tilde{E}(\mathbb{F}_p)$	5	5	5	10	11	10	20	20	$\cdots$	2035
$a_p(E)$	-2	-1	1	-2	1	4	-2	0	$\cdots$	-17

Volgens de modulariteitstelling is er een nieuwvorm $f \in S_2(\Gamma_0(11))$ zodanig dat $a_p(E) = a_p(f)$ voor alle priemen p. Met standaard theorie kan men checken dat $S_2(\Gamma_0(11))^{\text{nieuw}} = S_2(\Gamma_0(11))$ eendimensionaal is en dat $f := q\prod_{n=1}^{\infty}(1-q^n)^2(1-q^{11n})^2$ een genormaliseerde vorm in deze ruimte definieert. Dus dit moet de nieuwvorm f uit de modulariteitstelling zijn. Inderdaad geeft (formeel) expanderen van het product dat

$$\begin{aligned} f = {} & q - 2q^2 - q^3 + 2q^4 + q^5 + 2q^6 \\ & - 2q^7 - 2q^9 - 2q^{10} + q^{11} - 2q^{12} \\ & + 4q^{13} + 4q^{14} - q^{15} - 4q^{16} - 2q^{17} \\ & + 4q^{18} + 2q^{20} + \cdots - 17q^{2017} + \cdots. \end{aligned}$$

Hier is een klein tabelletje voor $a_p(f)$:

p	2	3	5	7	11	13	17	19	$\cdots$	2017
$a_p(f)$	-2	-1	1	-2	1	4	-2	0	$\cdots$	-17

We zien dat voor alle priemen p in de voorgaande twee tabellen we inderdaad hebben dat $a_p(E) = a_p(f)$.

Vermoeden van Birch en Swinnerton-Dyer
Laat E een elliptische kromme over $\mathbb{Q}$ zijn. Het is een stelling dat de groep van rationale punten $E(\mathbb{Q})$ eindig voortgebracht is, oftewel

$$E(\mathbb{Q}) \simeq T \times \mathbb{Z}^r$$

voor een eindige groep T en $r \in \mathbb{Z}_{\geq 0}$ die de *rang* van E wordt genoemd en genoteerd als $\operatorname{rang}(E(\mathbb{Q}))$. Aangezien dankzij de modulariteit van E geldt dat $L_E(s) = L_f(s)$ voor een nieuwvorm f en $L_f(s)$ een analytische voortzetting tot $\mathbb{C}$ heeft, volgt nu direct dat $L_E(s)$ een analytische voortzetting tot $\mathbb{C}$ heeft. Evenzo voldoet de functie

$$\xi_E(s) := N(E)^{s/2}(2\pi)^{-s}\Gamma(s)L_E(s)$$

aan de functionaalvergelijking

$$\xi_E(s) = \pm\xi_E(2-s)$$

voor alle $s \in \mathbb{C}$ en een zeker teken $\pm$ dat enkel van E afhangt, welke vermoedelijk $(-1)^{\operatorname{rang}(E(\mathbb{Q}))}$ is. Merk op dat $N(E)^{s/2}(2\pi)^{-s}\Gamma(s) \neq 0$ als $s = 1$, dus $L_E(s)$ is analytisch in een omgeving van $s = 1$. Vermoedelijk reflecteert $L_E(s)$ bij $s = 1$ diepe aritmetische informatie over E. Introduceer de *analytische rang* $r_{\text{an}}(E) := \operatorname{ord}_{s=1}(L_E(s))$, oftewel de nulpuntsorde van $L_E(s)$ bij $s = 1$. Introduceer ook de *algebraïsche rang* $r_{\text{al}}(E) := \operatorname{rang}(E(\mathbb{Q}))$.

Vermoeden 4 (Zwak Birch–Swinnerton-Dyer-vermoeden). *Voor alle elliptische krommen E over $\mathbb{Q}$ geldt*

$$r_{\text{an}}(E) = r_{\text{al}}(E).$$

Er bestaat ook een sterkere versie van dit vermoeden, die de eerste coëfficiënt ongelijk 0 in de Taylor-expansie van $L_E(s)$ rond $s = 1$ relateert aan andere (aritmetische) invarianten van E. Met name de orde van zijn zogenaamde *Shafarevich–Tate-groep*, waarvan men *vermoedt* dat die eindig is. Verder zijn er natuurlijke generalisaties waarbij $\mathbb{Q}$ wordt vervangen door een algemeen getallenlichaam, en eventueel tevens E wordt vervangen door een zogenaamde 'abelse variëteit' van willekeurige dimensie. Vermoeden 4 is echter al moeilijk genoeg en staat, naast andere grote problemen zoals de Riemann-hypothese, bekend als één van de zeven millenniumprijsproblemen van het Clay Mathematics Institute (CMI), waarvan de oplossing naast eeuwige roem ook nog een miljoen dollar oplevert. (Eén zo'n probleem, bekend als het Poincaré-vermoeden, is inmiddels opgelost.) De officiële beschrijving voor het CMI van Vermoeden 4 als millenniumprobleem is trouwens door Wiles gegeven.

Hij heeft ook, samen met Coates, in de jaren zeventig één van de eerste resultaten in de richting van het Birch–Swinnerton-Dyer-vermoeden voor een grote klasse van elliptische krommen verkregen [6]. In de jaren tachtig volgde nog belangrijke resultaten van Gross–Zagier en Kolyvagin. Hiermee was toen bekend dat voor *modulaire* elliptische krommen E over $\mathbb{Q}$ geldt dat

$$r_{\text{an}}(E) \leq 1 \Rightarrow r_{\text{an}}(E) = r_{\text{al}}(E). \qquad (8)$$

Samenvattend heeft de modulariteit van elliptische krommen dus (ten minste) twee directe belangrijke implicaties voor het Birch–Swinnerton-Dyer-vermoeden. Voor elke elliptische krommen E over $\mathbb{Q}$ geldt ten eerste dat de L-functie $L_E(s)$ een analytische continuatie tot een omgeving van $s = 1$ (zelfs heel $\mathbb{C}$) heeft, waarmee $r_{\text{an}}(E)$ überhaupt gedefinieerd kan worden. En ten tweede is (8) nu onvoorwaardelijk waar.

Generalisaties van Wiles' resultaten
We bespreken eerst enkele generalisaties van Wiles' modulariteitsresultaat. Deze maken het (onder andere) mogelijk dat som-

mige generalisaties van de Laatste Stelling van Fermat verkregen kunnen worden. Dit zullen we ook aanstippen.

Modulariteitsresultaten
Er zijn veel richtingen waarin modulariteit gegeneraliseerd is (en nog kan worden). We zullen ons hier voornamelijk beperken tot enige belangrijke klassen van elliptische krommen over getallenlichamen, maar maken ook een klein uitstapje naar Galois-representaties.

Alle elliptische krommen over $\mathbb{Q}$. De methoden van [30] en [29] werden snel uitgebreid om modulariteit aan te tonen van een veel grotere klasse van elliptische krommen over $\mathbb{Q}$ dan de semi-stabiele. De verwachting was verder dat de modulariteit van *alle* elliptische krommen over $\mathbb{Q}$ binnen handbereik was. Hoewel natuurlijk zeker niet eenvoudig, duurde het niet lang voordat de modulariteit van alle elliptische krommen over $\mathbb{Q}$ met conductor niet deelbaar door 3^3 bewezen was. Dit laatste komt neer op $\delta_3 = 0$. Om 'technische problemen bij reductie modulo 3' te boven te komen als $\delta_3 > 0$ moest zeker nog verder hard gewerkt worden. Dit lukte inderdaad door Breuil, Conrad, Diamond en Taylor [3], waarmee Stelling 3 uiteindelijk bewezen was. Zie [9] voor een uitgebreid overzicht.

Serres modulariteitsvermoeden. Dit is een vergaand vermoeden van Serre [23] dat zegt dat bepaalde Galois-representaties isomorf zijn aan (reducties van) Galois representaties die komen van nieuwvormen voor $\Gamma_1(N)$. Het is uiteindelijk volledig bewezen door Khare, Wintenberger en Kisin, zie [14] en [15].

We beginnen met een elementair voorbeeld waarin we in eerste instantie de Galois-representaties even onder tafel vegen. Beschouw het polynoom met gehele coëfficiënten $g = x^3 - x - 1$. We merken op dat g geen nulpunten in $\mathbb{Q}$ heeft en dat de discriminant -23 is. Voor elk priemgetal p noteren we met n_p het aantal nulpunten (zonder op de multipliciteit te letten) van g modulo p, oftewel

$$n_p := \#\{x \in \mathbb{F}_p : x^3 - x - 1 = 0\}.$$

Voor $p \neq 23$ kan g modulo p irreducibel blijven, splitsen in een (irreducibel) kwadratisch en een lineair stuk, of splitsen in drie verschillende (monische) lineaire stukken. Voor zulke p hebben we dus $n_p \in \{0,1,3\}$. Verder $g \equiv (x-10)^2(x-3) \pmod{23}$, dus $n_{23} = 2$. Hieronder staat een klein tabelletje voor n_p:

p	2	3	5	7	11	13	17	19	23	29	⋯	59
n_p	0	0	1	1	1	0	1	1	2	0	⋯	3

We brengen de nieuwvorm

$$\Delta := q\prod_{n=1}^{\infty}(1-q^n)^{24} = \sum_{n=1}^{\infty}\tau(p)q^p \in S_{12}(\Gamma_0(1))$$

in herinnering. Noteren we met $\overline{\tau(p)} \in \mathbb{F}_{23} = \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{22}\}$ de reductie van $\tau(p)$ modulo 23, dan hebben we hiervoor het volgende tabelletje:

p	2	3	5	7	11	13	17	19	23	29	⋯	59
$\overline{\tau(p)}$	$\overline{22}$	$\overline{22}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{22}$	$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{22}$	⋯	$\overline{2}$

Bij vergelijking van de tabellen voor n_p en $\overline{\tau(p)}$ vindt de lezer waarschijnlijk snel het opzienbarende patroon

$$n_p - 1 \equiv \tau(p) \pmod{23}. \qquad (9)$$

Men kan bewijzen dat dit inderdaad voor alle priemgetallen p geldt!

We willen dit resultaat nog even herschrijven in termen van een *Galois-representatie*. Laat K het kleinste deellichaam van $\mathbb{C}$ zijn waarin de drie nulpunten $\alpha_1, \alpha_2, \alpha_3$ van g liggen. Dan is K een getallenlichaam van graad 6 over $\mathbb{Q}$. De Galoisgroep $\mathrm{Gal}(K/\mathbb{Q})$ is per definitie de groep van lichaamsautomorfismen $\sigma : K \to K$ (met samenstelling als operatie), waarbij lichaamsautomorfisme betekent dat σ bijectief is en voldoet aan $\sigma(x+y) = \sigma(x) + \sigma(y)$ en $\sigma(xy) = \sigma(x)\sigma(y)$ voor alle $x, y \in K$. Elke $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ induceert een permutatie van de nulpunten $\alpha_1, \alpha_2, \alpha_3$ en wordt hier volledig door bepaald. In dit geval is elke permutatie van de nulpunten van g voort te zetten tot een $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ (want g is irreducibel en $\mathrm{Discriminant}(g) = -23$ is geen kwadraat in $\mathbb{Q}$). Derhalve kunnen we $\mathrm{Gal}(K/\mathbb{Q})$ identificeren met de volledige permutatiegroep op drie elementen. We schrijven $\mathrm{Gal}(K/\mathbb{Q}) = \{e, (12), (13), (23), (123), (132)\}$, waarbij bijvoorbeeld een permutatie zoals (123) staat voor het lichaamsautomorfisme $\sigma : K \to K$ met $\sigma(\alpha_1) = \alpha_2$, $\sigma(\alpha_2) = \alpha_3$, $\sigma(\alpha_3) = \alpha_1$, en e staat voor de identiteit. We definiëren nu het groepshomomorfisme

$$\rho : \mathrm{Gal}(K/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{F}_{23})$$

door middel van

$$(123) \mapsto R := \begin{pmatrix} \overline{11} & \overline{8} \\ \overline{-8} & \overline{11} \end{pmatrix},$$
$$(12) \mapsto S := \begin{pmatrix} \overline{1} & \overline{0} \\ \overline{0} & \overline{-1} \end{pmatrix}.$$

Voor elk priemgetal $p \neq 23$ definiëren we nu een *Frobenius-element* $\mathrm{Frob}_p \in \mathrm{Gal}(K/\mathbb{Q})$ als volgt

$$\mathrm{Frob}_p := \begin{cases} (123), & \text{als } n_p = 0, \\ (12), & \text{als } n_p = 1, \\ e, & \text{als } n_p = 3. \end{cases}$$

Merk op dat voor de sporen (de som van de diagonaalelementen, genoteerd als Tr) van S, T en I (de identiteit in $\mathrm{GL}_2(\mathbb{F}_{23})$) we hebben $\mathrm{Tr}(R) = \overline{22}$, $\mathrm{Tr}(S) = \overline{0}$, $\mathrm{Tr}(I) = \overline{2}$. Voor all $p \neq 23$ zien we nu dat (9) equivalent is met $\mathrm{Tr}(\rho(\mathrm{Frob}_p)) = \overline{\tau(p)}$. Het bovenstaande kunnen we als volgt generaliseren. Laat h een irreducibel polynoom van graad $n > 0$ met rationale coëfficiënten zijn en laat L het kleinste deellichaam van $\mathbb{C}$ zijn waarin alle n nulpunten $\alpha_1, \alpha_2, \ldots, \alpha_n$ van h liggen. De groep $\mathrm{Gal}(L/\mathbb{Q})$ van lichaamsautomorfismen van L kunnen we identificeren met een ondergroep van de permutatiegroep op n elementen door te kijken wat de acties van de automorfismen $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$ op de n nulpunten van h zijn. Laat l een priemgetal en $\mathbb{F}$ een eindige lichaamsuitbreiding van $\mathbb{F}_l$ zijn en beschouw een Galois-representatie

$$\rho : \mathrm{Gal}(L/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{F}),$$

oftewel ρ is een groepshomomorfisme. Laat $c \in \mathrm{Gal}(L/\mathbb{Q})$ complexe conjugatie beperkt tot L zijn, dan noemen we ρ *oneven* als $\det(\rho(c)) = -1$. Verder noemen we ρ *absoluut irreducibel* als er geen $M \in \mathrm{GL}_2(\overline{\mathbb{F}})$ is (met $\overline{\mathbb{F}}$ een algebraïsche afsluiting van $\mathbb{F}$) zodat $M\mathrm{Beeld}(\rho)M^{-1}$ bestaat uit bovendriehoeksmatrices. In het voorgaande voorbeeld is eenvoudig in te zien dat de Galois-representatie oneven is, met ietsje meer werk kan ook aangetoond worden dat deze absoluut irreducibel is. Serres modulariteitsvermoeden zegt nu in essentie dat voor een Galois-representatie ρ zoals hierboven die oneven en absoluut irreducibel is er een nieuwvorm $f \in S_k(\Gamma_1(N))$ bestaat voor zekere $k \geq 2$ en $N \geq 1$ en een priemideaal λ (in de ring van gehelen van het getallenlichaam voortgebracht door de Fourier-coëfficiënten van f) met $l \in \lambda$ zodat voor alle priemgetallen $p \nmid Nl$,

$$\mathrm{Tr}(\rho(\mathrm{Frob}_p)) = \overline{a_p(f)}$$

waarbij $\overline{a_p(f)}$ de reductie van a_p modulo

λ is (op een consistente wijze ingebed in $\mathbb{F}$). Hierbij is een *Frobenius-element* $\mathrm{Frob}_p \in \mathrm{Gal}(L/\mathbb{Q})$ een afbeelding waarvoor er een priemideaal $\wp \subset \mathcal{O}_L$ bestaat met $p \in \wp$, zó dat $\mathrm{Frob}_p(\wp) = \wp$ en $\mathrm{Frob}_p(x) \equiv x^p \pmod{\wp}$ voor alle $x \in \mathcal{O}_L$. Voor $p \nmid Nl$ is zo'n Frobenius-element Frob_p uniek op conjugatie na, zodat $\mathrm{Tr}(\rho(\mathrm{Frob}_p))$ dus uniek bepaald is.

$\mathbb{Q}$-krommen. Bovenstaande heeft als gevolg dat bepaalde hogerdimensionale generalisaties van elliptische krommen over $\mathbb{Q}$, namelijk zogenaamde GL_2-type abelse variëteiten over $\mathbb{Q}$, modulair zijn. Wat dit precies inhoudt, gaan we hier verder niet bespreken, maar nauw hiermee samenhangend is de modulariteit van $\mathbb{Q}$-krommen. Dat dit een gevolg van Serres modulariteitsvermoeden is, was al door Ribet bewezen in [21]. Om uit te leggen wat een $\mathbb{Q}$-kromme is, beginnen we met een willekeurige elliptische kromme E over een getallenlichaam K. Zonder verlies van algemeenheid kunnen we K indien nodig iets groter kiezen zodat K een eindige Galois-uitbreiding van $\mathbb{Q}$ is, dat wil zeggen dat er een polynoom g met rationale coëfficienten is zodat K het kleinste deellichaam van $\mathbb{C}$ is dat alle nulpunten van g bevat (zo'n g kan trouwens altijd irreducibel en niet-constant gekozen worden). Voor elke $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ krijgen we een elliptische kromme E^σ door σ simpelweg los te laten op alle coëfficiënten van een Weierstrass-vergelijking voor E. Is E bijvoorbeeld gegeven door $y^2 = x^3 + ax^2 + bx + c$ (met $a,b,c \in K$), dan wordt E^σ gegeven door $y^2 = x^3 + \sigma(a)x^2 + \sigma(b)x + \sigma(c)$. Voor een niet-triviale σ kan men afvragen of E^σ *isogeen* is aan E. Dit laatste houdt grosso modo in dat er rationale functies (quotiënten van polynomen) in x en y met coefficiënten in K zijn die E afbeelden op E^σ. Dit is equivalent met de uitspraak dat $a_\wp(E) = a_\wp(E^\sigma)$ voor alle priemidealen $\wp \subset \mathcal{O}_K$ met $\wp \neq 0$, maar deze equivalentie is hoogst niet-triviaal. Als voor elke $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ nu geldt dat E^σ isogeen is aan E, dan noemen we E een $\mathbb{Q}$-kromme. We merken op dat alle elliptische krommen over $\mathbb{Q}$ $\mathbb{Q}$-krommen zijn, maar er zijn ook oneindig veel essentieel verschillende andere $\mathbb{Q}$-krommen. Als voorbeeld geven we

$$E : y^2 = x^3 + (156\sqrt{-3} - 135)x + 2(546\sqrt{-3} + 41).$$

De Galoisgroep van $K := \mathbb{Q}(\sqrt{-3})$ bestaat uit de identiteitsafbeelding op K en complexe conjugatie c (beperkt tot K). Door in de vergelijking voor E elke $\sqrt{-3}$ te vervangen voor $-\sqrt{-3}$ krijgen we

$$E^c : y^2 = x^3 + (-156\sqrt{-3} - 135)x + 2(-546\sqrt{-3} + 41).$$

Dat E een $\mathbb{Q}$-kromme is, volgt nu omdat we een afbeelding (een 3-isogenie) met behulp van rationale functies van E naar E^c hebben, expliciet:

$$x \mapsto \frac{-x^3/3 - 6x^2 - (104\sqrt{-3}+99)x - (520\sqrt{-3}+4216/3)}{x^2+18x+81}$$
$$y \mapsto \frac{\sqrt{-3}x^3y/9 + 3\sqrt{-3}x^2y + (104+3\sqrt{-3})xy + (104-5759\sqrt{-3}/9)y}{x^3+27x^2+243x+729}.$$

Het begin van de q-expansies van de nieuwvormen in $S_2(\Gamma_0(63))$ worden gegeven door

$$\begin{aligned} f &= q + q^2 - q^4 + 2q^5 - q^7 - 3q^8 \\ &\quad + 2q^{10} - 4q^{11} + \cdots \\ g_1 &= q + \sqrt{3}\,q^2 + q^4 - 2\sqrt{3}\,q^5 + q^7 \\ &\quad - \sqrt{3}\,q^8 - 6q^{10} + 2\sqrt{3}\,q^{11} + \cdots \\ g_2 &= q - \sqrt{3}\,q^2 + q^4 + 2\sqrt{3}\,q^5 + q^7 \\ &\quad + \sqrt{3}\,q^8 - 6q^{10} - 2\sqrt{3}\,q^{11} + \cdots. \end{aligned}$$

Nu is E modulair in de zin dat

$$L_E(s) = L_{g_1}(s)L_{g_2}(s).$$

Voor algemene $\mathbb{Q}$-krommen over een getallenlichaam K houdt modulariteit in dat na het eventueel vervangen van K voor een groter getallenlichaam (ook al was K Galois en alle relevante isogenieën gedefinieerd over K) dat hun L-functie gelijk is aan een product van L-functies van nieuwvormen in $S_2(\Gamma_1(N))$ voor zekere waardes van N.

Hilbert-modulariteit. Voor bepaalde elliptische krommen over getallenlichamen groter dan $\mathbb{Q}$ hebben we in het voorgaande een vorm van modulariteit gezien. Echter, $\mathbb{Q}$-krommen zijn relatief 'zeldzaam'. De hoop is dat alle elliptische krommen over alle getallenlichamen modulair zijn op één of andere manier (wat een speciaal geval van de Langlands-vermoedens is). In het geval dat het getallenlichaam totaal reëel is (zie hieronder), is er aardige progressie geboekt op dit terrein en kan de modulariteit uitgedrukt worden met behulp van zogenaamde Hilbert-modulaire vormen.

Laat g een irreducibel polynoom van graad $n > 0$ met rationale coëfficiënten zijn. Stel verder dat alle (a priori complexe) nulpunten $\alpha := \alpha_1, \alpha_2, \ldots, \alpha_n$ van g reëel zijn. Een getallenlichaam van de vorm $K := \mathbb{Q}(\alpha)$ noemen we *totaal reëel*. Elk element in K is van de vorm $a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}$, met unieke $a_0, a_1, \ldots, a_{n-1} \in \mathbb{Q}$. We beschouwen voor $i = 1, 2, \ldots, n$ de inbedding

$$\begin{aligned} &\sigma_i : K \to \mathbb{R} \\ &a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1} \\ &\quad \mapsto a_0 + a_1\alpha_i + \cdots + a_{n-1}\alpha_i^{n-1}. \end{aligned}$$

Nemen we bijvoorbeeld het totaal reële getallenlichaam $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a,b \in \mathbb{Q}\}$, dan hebben we naast de identiteitsafbeelding $K \to \mathbb{R} : x \mapsto x$ nog de afbeelding die een element $a + b\sqrt{2}$ naar $a - b\sqrt{2}$ stuurt.

Voor het bespreken van Hilbert-modulaire vormen beschouwen we om te beginnen de groepen

$$\mathrm{GL}_2(\mathcal{O}_K) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a,b,c,d \in \mathcal{O}_K,\ ad - bc \in \mathcal{O}_K^* \right\},$$

$$\mathrm{GL}_2^+(\mathcal{O}_K) := \{\gamma \in \mathrm{GL}_2(\mathcal{O}_K) : \sigma_i(\det(\gamma)) > 0 \text{ voor } i = 1, \ldots, n\}.$$

We kunnen $\mathrm{GL}_2^+(\mathcal{O}_K)$ inbedden in $(\mathrm{GL}_2^+(\mathbb{R}))^n$ via

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \left(\begin{pmatrix} \sigma_1(a) & \sigma_1(b) \\ \sigma_1(c) & \sigma_1(d) \end{pmatrix}, \ldots, \begin{pmatrix} \sigma_n(a) & \sigma_n(b) \\ \sigma_n(c) & \sigma_n(d) \end{pmatrix} \right).$$

Via de actie (3) van $\mathrm{GL}_2^+(\mathbb{R})$ op $\mathbb{H}$ verkrijgen we zo een actie van $\mathrm{GL}_2^+(\mathcal{O}_K)$ op het n-voudig bovenhalfvlak $\mathbb{H}^n$. Merk op voor $K = \mathbb{Q}$ dat $\mathrm{GL}_2^+(\mathcal{O}_K) = \mathrm{SL}_2(\mathbb{Z})$ en dat bovenstaande actie reduceert tot de 'standaard' actie van $\mathrm{SL}_2(\mathbb{Z})$ op $\mathbb{H}$. Voor elk ideaal $0 \neq \mathcal{N} \subset \mathcal{O}_K$ introduceren we, analoog aan $\Gamma_0(N)$ in het $\mathrm{SL}_2(\mathbb{Z})$ geval, de groep

$$\Gamma_0(\mathcal{N}) := \left\{ \gamma \in \mathrm{GL}_2^+(\mathcal{O}_K) \,\middle|\, \gamma \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod{\mathcal{N}} \right\}.$$

Evenzo kunnen we $\Gamma_1(N)$ generaliseren (evenals andere congruentieondergroepen). Globaal gezegd is een *Hilbert-modulaire spitsvorm* voor $\Gamma_0(\mathcal{N})$ van (parallel) gewicht k een holomorfe functie $f : \mathbb{H}^n \to \mathbb{C}$ die voor alle $z = (z_1, \ldots, z_n) \in \mathbb{H}^n$ en $\gamma \in \Gamma_0(\mathcal{N})$ transformeert als

$$f(\gamma z) = \left(\prod_{i=1}^{n} \frac{(\sigma_i(c)z_i + \sigma_i(d))^k}{\sigma_i(\det(\gamma))^{k-1}} \right) f(z)$$

en verder bepaald (begrensd) groeigedrag heeft (dit wordt ook wel omschreven als 'nul zijn op de spitsen'). Deze functies vormen weer een eindigdimensionale $\mathbb{C}$-vectorruimte, genoteerd als $S_k(\Gamma_0(\mathcal{N}))$.

Aan de hand hiervan kan men uiteindelijk weer bepaalde $\mathbb{C}$-vectorruimtes $S_k(\mathcal{N})^{\text{nieuw}} \subset S_k(\mathcal{N})$ introduceren (met $S_k(\mathcal{N}) = S_k(\Gamma_0(\mathcal{N}))$ als het zogenaamde strikte klassegetal h gelijk aan 1 is, maar $S_k(\mathcal{N})$ is ingewikkelder voor $h > 1$), en daarop werkende Hecke-operatoren $T_{\wp}$ voor priemidealen $0 \neq \wp \subset \mathcal{O}_K$. Dit geeft het analogon van $S_k(\Gamma_0(N))^{\text{nieuw}} \subset S_k(\Gamma_0(N))$ en Hecke-operatoren T_p voor p priem. Een *Hilbert-nieuwvorm* is dan een $f \in S_k(\mathcal{N})^{\text{nieuw}}$ die simultaan eigenvector is voor alle $T_{\wp}$'s (+ een keuze van normalisatie van f). Voor zo'n f definiëren we $a_{\wp}(f)$'s als bijbehorende eigenwaardes, oftewel

$$T_{\wp} f = a_{\wp}(f) f.$$

De $a_{\wp}(f)$'s kunnen trouwens ook weer gerelateerd worden aan coëfficiënten van een (n-dimensionale) Fourier-ontwikkeling van f. Er is ook weer een L-functie,

$$L_f(s) := \prod_{\wp \nmid \mathcal{N}} (1 - a_{\wp}(f) N(\wp)^{-s} + N(\wp)^{1-2s})^{-1} \cdot \prod_{\wp \mid \mathcal{N}} (1 - a_{\wp}(f) N(\wp)^{-s})^{-1}$$

die convergeert voor $s \in \mathbb{C}$ met reëel deel strikt groter dan $3/2$, een analytische continuatie heeft tot heel $\mathbb{C}$ en aan een natuurlijke functionaalvergelijking voldoet.

Het *Hilbert-modulair* zijn van een elliptische kromme E over een totaal reëel getallenlichaam K is nu per definitie dat er een Hilbert-nieuwvorm $f \in S_2(\mathcal{N})$ is, met $\mathcal{N}$ de conductor van E en voor alle priemidealen $0 \neq \wp \subset \mathcal{O}_K$ geldt $a_{\wp}(E) = a_{\wp}(f)$. Merk op dat dit laatste equivalent is aan $L_E(s) = L_f(s)$. Of alle elliptische krommen over alle totaal reële getallenlichamen Hilbert-modulair zijn, is een groot open probleem en vergaande generalisatie van het Shimura–Taniyama–Weil-vermoeden. In een aantal gevallen is die modulariteit echter recentelijk bewezen. Een spectaculair voorbeeld is het bewijs door Freitas, Siksek en Le Hung [11] van Hilbert modulariteit voor alle elliptische krommen over alle reëel kwadratische getallenlichamen, oftewel lichamen van de vorm $\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} : a, b \in \mathbb{Q}\}$ met d positief geheel en geen kwadraat.

Trouwens, stel K is een totaal reëel lichaam en $F = K(\sqrt{x})$ is een kwadratische uitbreiding verkregen door een kwadratische wortel van een element $x \in K$ toe te voegen, onder de extra voorwaarde dat $\sigma(x) < 0$ voor elke inbedding $\sigma : K \to \mathbb{R}$. Dan kondigen Allen, Calegari, Caraiani, Gee, Helm, Le Hung, Newton, Scholze, Taylor en Thorne [27] aan dat elliptische krommen over K modulair worden (in een zin die we niet nader zullen omschrijven hier) na eventueel een eindige uitbreiding K'/K, alhoewel er op het moment van het schrijven van dit artikel nog geen preprint beschikbaar is. Dit nieuwe bewijs is mogelijk door nieuwe technieken die in de laatste jaren zijn ontwikkeld door Scholze [22] en Caraiani en Scholze [4].

Diofantische resultaten

De oorspronkelijke motivatie voor Wiles om de modulariteit van (semi-stabiele) elliptische krommen over $\mathbb{Q}$ te bewijzen, was om een bewijs voor de Laatste Stelling van Fermat te verkrijgen. Dit modulariteitsbewijs en generalisaties hiervan vinden hun toepassingen ook weer in het oplossen van andere diofantische vergelijkingen. Een belangrijke klasse zijn de zogenaamde gegeneraliseerde Fermat-vergelijkingen: laat $p, q, r \in \mathbb{Z}_{\geq 2}$ en beschouw de diofantische vergelijking

$$x^p + y^q = z^r, \quad x, y, z \in \mathbb{Z}, \quad \gcd(x, y, z) = 1. \tag{10}$$

De extra ggd-eis staat er om allerlei 'flauwe' oplossingen uit te sluiten: voor de exponenten $(p, q, r) = (3, 5, 7)$ bijvoorbeeld vindt men eenvoudig de identiteit

$$(r^{12} s^{28} (r+s)^{30})^3 + (r^7 s^{17} (r+s)^{18})^5 = (r^5 s^{12} (r+s)^{13})^7.$$

Verder noemen we een oplossing $(x, y, z) = (a, b, c)$ van (10) *niet-triviaal* als $abc \neq 0$.

Voor het bewijs van de Laatste Stelling van Fermat begint men met het construeren van een zogenaamde Frey-kromme: aan een hypothetische niet-triviale oplossing $a^l + b^l = c^l$ met $l \geq 5$ priem kent men toe de elliptische kromme over $\mathbb{Q}$,

$$E : y^2 = f(x), \quad f(x) := x(x - a^l)(x + b^l).$$

Voor de discriminant van f geldt

$$\Delta(f) = a^{2l} b^{2l} (a^l + b^l)^2 = (abc)^{2l}.$$

Om tot een tegenspraak te komen, gebruikt men zogenaamde niveauverlaging à la Ribet [20]. Hiervoor heeft men naast een technisch 'irreducibiliteits-ingrediënt' (à la Mazur) nodig dat E modulair is. De verdere aritmetische eigenschappen van E die gebruikt worden, zijn ten eerste dat $\Delta(f) = \alpha \beta^l$ met $\alpha, \beta \in \mathbb{Z}$ en α een getal dat niet van de (hypothetische) oplossing afhangt (in dit geval is simpelweg $\alpha = 1$). Ten tweede moeten de priemgetallen p waarvoor $f(x)$ modulo p een drievoudig nulpunt krijgt ook niet van de (hypothetische) oplossing afhangen. In dit geval krijgen we dat voor elkaar om zonder verlies van algemeenheid aan te nemen dat $\gcd(a, b, c) = 1$, zodat voor geen enkel priemgetal p alle drie de nulpunten 0, a^l en $-b^l$ van $f(x)$ gelijk worden modulo p. Na het samensmeden van alle ingrediënten komt men er in dit geval op uit dat er bij een oplossing een nieuwvorm in $S_2(\Gamma_0(2))$ hoort. Laatstgenoemde ruimte is nuldimensionaal, zodat de geconstrueerde nieuwvorm niet kan bestaan, een tegenspraak. Samen met de exponentengevallen $l = 3, 4$, volgt de Laatste Stelling van Fermat.

In andere gevallen kan men soms ook een bijbehorende Frey-kromme construeren. Om zo de diofantische vergelijking op te lossen is in ieder geval modulariteit van de kromme nodig. We geven een paar voorbeelden.

Neem $(p, q, r) = (l, l, 2)$ met $l \geq 7$ priem in (10), dus we beschouwen een (hypothetische) oplossing $a^l + b^l = c^2$. In [7] is bewezen dat er dan geen niet-triviale oplossingen zijn. De Frey-kromme is

$$E : y^2 = f(x), \quad f(x) := x^3 + 2cx^2 + a^p x,$$

met discriminant

$$\Delta(f) = 4a^{2l}(c^2 - a^l) = 4(a^2 b)^l.$$

We merken op dat E niet semi-stabiel is, maar de modulariteit van E is een relatief kleine uitbreiding van het semi-stabiele geval, welke vrijwel na Wiles' resultaat beschikbaar kwam. Verder kan men samen met een paar gevallen voor kleine exponenten aantonen dat voor $(p, q, r) = (l, l, 2)$ er geen niet-triviale oplossingen zijn voor alle gehele $l \geq 4$.

Nemen we nu het geval $(p, q, r) = (l, l, 3)$, ook beschouwd in [7], dan heeft de bijbehorende Frey-kromme E een conductor N die deelbaar is door 3^3. Bij publicatie was de modulariteit van zulke E nog niet bekend, waardoor het diofantisch resultaat nog voorwaardelijk was. Dankzij de volle kracht van de modulariteit van alle elliptische krommen over $\mathbb{Q}$, samen met wat gevallen voor kleine exponenten, is nu onvoorwaardelijk bewezen dat er in dit geval geen niet-triviale oplossingen zijn voor gehele $l \geq 3$.

Voor $(p, q, r) = (4, 2, l)$ met $l \geq 211$ wordt in [10] de Frey-kromme

$$E: y^2 = f(x),$$
$$f(x) := x^3 + 2(1+i)ax^2 + (b+ia^2)x,$$

beschouwd (met $i = \sqrt{-1}$), waarbij

$$\Delta(f) = -4i(a^2-ib)^2(a^2+ib).$$

De factorisatie $(a^2-ib)(a^2+ib) = c^l$ geeft dat $\Delta(f)$ van de gewenste vorm is. Nu is E een $\mathbb{Q}$-kromme over $\mathbb{Q}(i)$ en derhalve modulair. Een isogenie van E naar de Galois-geconjugeerde van E (gegeven door de vergelijking van E met overal i vervangen door $-i$) wordt expliciet gegeven door

$$(x,y) \mapsto (\tfrac{1}{2}i(y^2/x^2), \\ -\tfrac{1}{4}(1-i)y(b+ia^2-x^2)/x^2).$$

Samen met gevallen voor kleine exponenten is er inmiddels bekend dat voor $(p,q,r) = (4,2,l)$ er geen niet-triviale oplossingen zijn voor gehele $l \geq 4$.

Er zijn trouwens ook voorbeelden waarin er wel niet-triviale oplossingen zijn en een aanpak via modulariteit van Frey-krommen, naast andere methoden, de diofantische vergelijking kan oplossen. Voor $(p,q,r) = (2,3,l)$ kan men bijvoorbeeld de Frey-kromme

$$E: y^2 = f(x), \quad f(x) = x^3 + 3bx - 2a$$

beschouwen met

$$\Delta(f) = -2^2 \cdot 3^3(a^2+b^3) = -2^2 \cdot 3^3 c^l.$$

In [19] wordt het geval $l = 7$ volledig opgelost; er zijn 5 paren niet-triviale oplossingen. In dit geval is trouwens weer de modulariteit van alle elliptische krommen over $\mathbb{Q}$ nodig.

Voor sommige diofantische problemen over $\mathbb{Z}$ of $\mathbb{Q}$ kan het soms handig zijn om een Frey-kromme te construeren over een groter getallenlichaam K. Maar dit kan ook simpelweg het geval zijn voor generalisaties van diofantische problemen over $\mathbb{Q}$ naar problemen over K. Men kan bijvoorbeeld kijken naar een generalisatie van de Laatste Stelling van Fermat waarbij de vergelijking $x^n + y^n = z^n$ moet worden opgelost in $x,y,z \in O_K$. Hiervoor kan men de 'standaard' Frey-kromme gebruiken, maar dit is nu een elliptische kromme over K geworden. Als K reëel kwadratisch is, hebben we Hilbert-modulariteit van de Frey-kromme. In dat geval zijn Freitas en Siksek in [12] erin geslaagd om de rest van de benodigde ingrediënten vaak werkzaam te krijgen en bewijzen onder andere dat voor oneindig veel verschillende reëel kwadratische getallenlichamen K er een constante B_K is, zodat voor alle priemgetallen $l \geq B_K$ er geen niet-triviale oplossingen zijn van $x^l + y^l = z^l$ in $x,y,z \in O_K$ (of K, dat maakt hier niet uit).

Bovenstaande vormt slechts een kleine selectie van diofantische vergelijkingen waarbij modulariteit een belangrijke rol speelt. Het is de verwachting dat er in de toekomst nog vele bij zullen komen, zowel voor de types van modulariteit die we besproken hebben als voor andere types, en zowel voor bewezen gevallen van modulariteit als voor nog te bewijzen gevallen.

De Langlands-vermoedens

Zoals eerder uitgelegd, het resultaat van Wiles kan gezien worden als een speciaal geval van de Langlands-vermoedens. Deze vermoedens zijn nog open, maar dankzij het werk van Wiles zijn er in de afgelopen jaren grote stappen gezet. In deze paragraaf willen we graag een tipje van de sluier oplichten en een idee geven waar de Langlands-vermoedens over gaan. Een probleem met de vermoedens is dat ze niet eenvoudig te formuleren zijn. Hierom geven we in dit overzichtsartikel alleen een speciaal geval, dat minder algemeen is dan de originele vermoedens. Dit speciale geval heeft een eenvoudigere formulering en is toch (hopelijk) nog steeds interessant.

Systemen van diofantische vergelijkingen

Met een *diofantisch systeem* bedoelen we een collectie van een eindig aantal vergelijkingen

$$\begin{cases} F_1(x_1,x_2,\ldots,x_b) = 0 \\ F_2(x_1,x_2,\ldots,x_b) = 0 \\ \vdots \qquad\qquad \vdots \\ F_a(x_1,x_2,\ldots,x_b) = 0 \end{cases} \tag{11}$$

waarbij a, b positieve gehele getallen zijn en $F_1,\ldots,F_a$ polynomen zijn met gehele coefficiënten in variabelen $x_1,\ldots,x_b$. Sinds de oudheid zijn wiskundigen geïnteresseerd in het oplossen van dergelijke systemen van diofantische vergelijkingen. Het voorbeeld van een elliptische kromme correspondeert met het geval waar $a = 1$, $b = 2$ en $F(x_1,x_2) = x_1^2 - (x_2^3 + Ax_2 + B)$.

Gegeven een diofantisch systeem van vergelijkingen X kunnen we voor alle priemgetallen p het aantal oplossingen $\#X(\mathbb{Z}/p\mathbb{Z})$ van X modulo p tellen. Merk op dat $(\mathbb{Z}/p\mathbb{Z})^b$ een eindige verzameling is, en voor elk element $(x_i) \in (\mathbb{Z}/p\mathbb{Z})^b$ kunnen we kijken of hij wel of niet aan de vergelijkingen in het systeem (11) modulo p voldoet. Bijvoorbeeld kunnen we kijken naar de vergelijking X gegeven door $x^4 + y^3 = z^2$. Het aantal oplossingen $\#X(\mathbb{Z}/p\mathbb{Z})$ voor priemgetallen p kunnen we weer in een tabel zetten:

p	2	3	5	7	11	13	$\cdots$
$\#X(\mathbb{F}_p)$	4	18	100	1210	1210	2028	$\cdots$

(12)

De Langlands-vermoedens voorspellen dat deze reeks getallen (op eindig veel priemgetallen p na) overeenkomt met 'Heckeeigenwaarden' van 'automorfe vormen'.

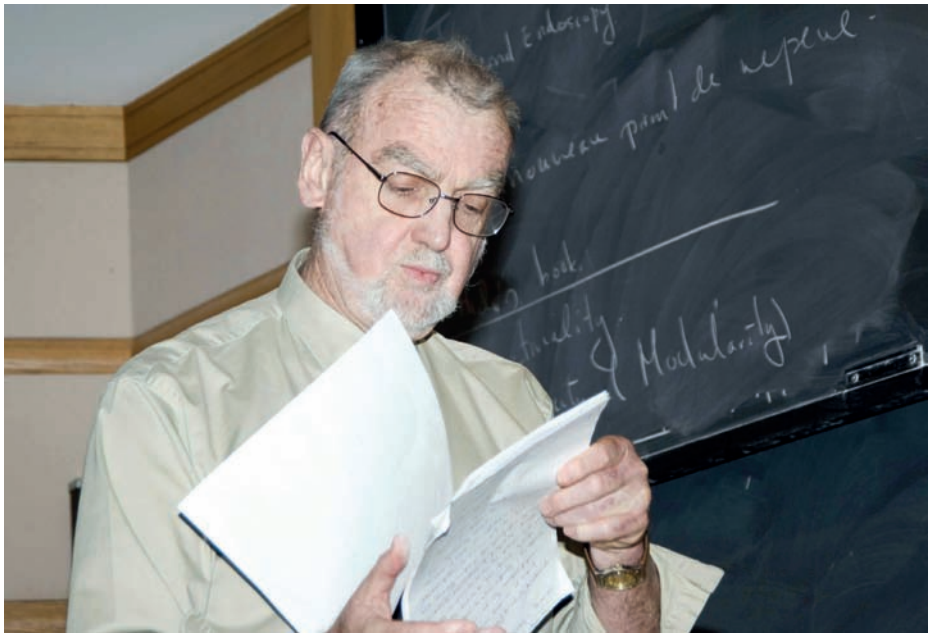

Foto: C.J. Mozzochi, Princeton, NJ

Robert Langlands tijdens de conferentie 'The $\mathcal{L}$-Group at 40', Institute for Advanced Study, Princeton, 2006

Een belangrijke vraag is, gegeven een diofantisch systeem X, in hoeverre het rijtje $\#X(\mathbb{Z}/p\mathbb{Z})$ 'reducibel' is, oftewel ontbonden kan worden als lineaire combinatie van meer eenvoudige rijtjes (die bijvoorbeeld afkomen van eenvoudigere diofantische systemen). Rond de tijd dat Langlands zijn brief aan Weil schreef, was de meetkundegroep van Alexander Grothendieck bezig met een hypothetische theorie van *motieven* van variëteiten. Elk motief M zou aanleiding geven tot een rijtje getallen $\#M(\mathbb{Z}/p\mathbb{Z})$ (mogelijk niet in $\mathbb{Z}$!) waar p varieert over de priemgetallen, en voor elk diofantisch systeem X, zou het rijtje $\#X(\mathbb{Z}/p\mathbb{Z})$ moeten ontbinden als een lineaire combinatie van rijtjes die bij motieven horen. In deze zin zouden motieven de bouwstenen vormen van algebraïsche vergelijkingen. Wat motieven precies zijn, is tot nu toe nog onbekend en het vinden van de juiste definitie is een open probleem. Langlands formuleerde zijn originele vermoeden in termen van deze hypothetische theorie van motieven van Grothendieck.

Automorfe vormen

We gaan nu de analytische kant bekijken, en definiëren wat discrete automorfe vormen zijn. Kies positieve gehele getallen $n,N\in\mathbb{Z}_{\geq 1}$. We noteren $\mathrm{GL}_n(\mathbb{Z})$ voor de groep van alle inverteerbare $n\times n$-matrices g met gehele coëfficiënten zó dat $\det(g)=\pm 1$ (dan heeft de inverse matrix g^{-1} ook gehele coëfficiënten). We beginnen met de *congruentieondergroepen* $\Gamma(N)\subset\mathrm{GL}_n(\mathbb{Z})$, gedefinieerd door

$$\Gamma(N)=\{g\in\mathrm{GL}_n(\mathbb{Z})\mid g\equiv 1_{n\times n}(\mathrm{mod}\ N)\}.$$

We schrijven $\mathrm{GL}_n(\mathbb{R})$ voor de groep van reële inverteerbare $n\times n$-matrices en $\mathrm{PGL}_n(\mathbb{R})$ voor het quotiënt $\mathrm{GL}_n(\mathbb{R})/\mathbb{R}^\times$, waar we $\mathbb{R}^\times$ zien in $\mathrm{GL}_n(\mathbb{R})$ als de scalaire matrices.

Kies een continu groepsmorfisme $\omega:\mathbb{R}^\times\to\mathbb{C}^\times$. We schrijven $H_\omega=L^2(\mathrm{GL}_n(\mathbb{R})/\Gamma(N),\omega)$ voor de ruimte van meetbare functies $f:\mathrm{GL}_n(\mathbb{R})\to\mathbb{C}$, zó dat

(A1) Voor zekere $N\in\mathbb{Z}_{\geq 1}$ groot genoeg (het *niveau* van f) geldt $f(g\gamma)=f(g)$ voor alle $\gamma\in\Gamma(N)$, en alle $g\in\mathrm{GL}_n(\mathbb{R})$.

(A2) Er geldt $f(zg)=\omega(z)f(g)$.

(A3) De integraal

$$\int_{\mathrm{PGL}_n(\mathbb{R})}|\det(g)|^{-2w/n}\,|f(g)|^2\,\mu(g)$$

convergeert. Hier is μ de Haar-maat op $\mathrm{PGL}_n(\mathbb{R})$, en $w\in\mathbb{R}_{>0}$ is het unieke reele getal zó dat $|\cdot|^{-w}\omega$ beeld heeft in de eenheidscirkel $S^1\subset\mathbb{C}^\times$.

Per definitie zijn in H_ω twee functies $f,h\in H_\omega$ equivalent als het verschil $f-h$ support heeft op een verzameling van maat 0. Modulo deze equivalentierelatie is H_ω een Hilbertruimte. De groep $\mathrm{GL}_n(\mathbb{R})$ werkt op H_ω door translaties: Als $f\in H_\omega$ en $g\in\mathrm{GL}_n(\mathbb{R})$, dan hebben we de getransleerde functie ${}^gf\in H$ met ${}^gf(x)=f(gx)$ voor alle $x\in\mathrm{GL}_n(\mathbb{R})$.

De ruimte van discrete automorfe vormen $A_{N,\omega}$ is een bepaalde deelruimte van $H_{N,\omega}$. Het punt is dat er ontzettend veel L^2-functies $f:\mathrm{GL}_n(\mathbb{R})\to\mathbb{C}$ zijn, en je wilt de theorie algebraïsch maken. Dit kan door te beperken naar functies die voldoen aan bepaalde eindigheidscondities (K-eindig, $\mathfrak{Z}$-eindig) en groeicondities die we niet allemaal expliciet zullen maken (zie [1, 1.3(a,b,c,d)]), maar je kan ook representatietheorie gebruiken om de discrete automorfe vormen te definiëren. Neem eerst $H^{\mathrm{disc}}_{N,\omega}\subset H_{N,\omega}$, de complexe deelvectorruimte van H_ω opgespannen door alle deelruimten $V\subset H_{N,\omega}$, die $\mathrm{GL}_n(\mathbb{R})$-stabiel zijn, ofwel ${}^gf\in V$ voor alle $g\in\mathrm{GL}_n(\mathbb{R})$ en alle $f\in V$, en bovendien irreducibel, ofwel de enige gesloten $\mathrm{GL}_n(\mathbb{R})$-stabiele deelruimten W van V zijn 0 en V zelf. De ruimte $H^{\mathrm{disc}}_{N,\omega}$ is geen Hilbertruimte meer, dit komt omdat de ruimte niet compleet is. Het orthogonale complement van de afsluiting van $H^{\mathrm{disc}}_{N,\omega}$ in $H_{N,\omega}$ is het continue spectrum. Deze ruimte is vooral belangrijk voor de theorie van Eisenstein-reeksen. De ruimte van discrete automorfe vormen $A_{N,\omega}$ is nu de ruimte van functies $f\in H^{\mathrm{disc}}_{N,\omega}$ die 'K-eindig' zijn, ofwel zó dat de vectorruimte opgespannen door de getransleerde functies ${}^gf$ eindigdimensionaal is waarbij g varieert over de orthogonale matrices:

$$\dim_{\mathbb{C}}\langle\,{}^gf\mid g\in\mathrm{GL}_n(\mathbb{R}),g^{\mathrm{t}}g=1\,\rangle<\infty.$$

De ruimte $A_{N,\omega}$ ligt dicht in $H^{\mathrm{disc}}_{N,\omega}$. Als laatste: een discrete automorfe vorm is een functie $f:\mathrm{GL}_n(\mathbb{R})\to\mathbb{C}$ zó dat voor N voldoende groot, en een $\omega:\mathbb{R}^\times\to\mathbb{C}^\times$ geldt $f\in H_{N,\omega}$. De ruimte van automorfe vormen $A_{N,\omega}$ ligt dicht in $H^{\mathrm{disc}}_{N,\omega}$.

We zijn vooral geïnteresseerd in *eigenfuncties*. Voor elke automorfe vorm $f:\mathrm{GL}_n(\mathbb{R})\to\mathbb{C}$, definiëren we voor elk priemgetal p copriem met N een nieuwe automorfe vorm $T_p(f)$ op $\mathrm{GL}_n(\mathbb{R})$ door de formule

$$T_p(f)(x)=\sum_{i=1}^{\deg(T_p)}f(x\cdot\gamma_i)\,. \tag{13}$$

Hier is $\deg(T_p)$ een natuurlijk getal en zijn $\gamma_i\in\mathrm{GL}_n(\mathbb{Z})$ matrices zó dat geldt

$$\Gamma(N)\cdot\begin{pmatrix}1&&\\&\ddots&\\&&p\end{pmatrix}\cdot\Gamma(N)=\coprod_{i=1}^{\deg(T_p)}\Gamma(N)\cdot\gamma_i\subset\mathrm{GL}_n(\mathbb{Z})\,.$$

Merk op dat de uitdrukking (13) onafhankelijk is van de keuze van de matrices γ_i.

Hier is een misschien meer intuïtieve manier om over formule (13) na te denken. Stel f is een automorfe vorm. Stel $g\in\mathrm{GL}_n(\mathbb{Z})$ is de diagonale matrix $\mathrm{diag}(1,\ldots,1,p)$. Definieer de functie ${}^gf(x)$ door simpelweg met g te transleren: ${}^gf(x):=f(xg)$ op $\mathrm{GL}_n(\mathbb{R})$. Dan is ${}^gf$ niet $\Gamma(N)$-invariant, maar $g^{-1}\Gamma(N)g$-invariant; om een $\Gamma(N)$-invariante functie te krijgen, wil je het gemiddelde nemen van de $\Gamma(N)$-translaties van ${}^gf$ (dit zijn er eindig veel). Als je dit op de juiste manier doet, dan kom je op de formule in (13) uit.

De automorfe vorm f is een eigenvorm als het een eigenvector is voor de operatoren T_p voor alle priemgetallen p met $p\nmid N$. De corresponderende eigenwaarden $\lambda_p\in\mathbb{C}$ met $T_pf=\lambda_pf$ zullen we de *Hecke-eigenwaarden* noemen. Voor elk priemgetal $p\nmid N$ hebben we dus zo'n eigenwaarde. Behalve de operatoren T_p kan je ook kijken naar andere Hecke-operatoren gegeven door een matrix verschillend van $\mathrm{diag}(1,\ldots,1,p)$. De reden voor onze keuze is dat de eigenwaarden van deze operator oplossingsaantallen van diofantische systemen geeft (zie (14)).

Voorbeeld 5. Laten we kijken hoe het voorbeeld van modulaire vormen in dit kader past. Stel $f:\mathbb{H}\to\mathbb{C}$ is een genormaliseerde modulaire eigenvorm van gewicht $k\geq 0$ voor $\Gamma_0(N)$. Laat $\mathbb{H}^\pm=\{x+yi\in\mathbb{C}\mid y\neq 0\}$ het complexe dubbelhalfvlak zijn. Dan kunnen we de functie f voortzetten naar een functie $\tilde{f}$ op $\mathbb{H}^\pm$ door de formule $\tilde{f}(\tau)=\tau^{-k}f(1/\tau)$ voor $\tau\in\mathbb{H}^\pm$ met negatief imaginair deel. We definiëren de functie $h:\mathrm{GL}_2(\mathbb{R})\to\mathbb{C}$ door

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}\mapsto\frac{(ad-bc)^{k-1}}{(ci+d)^k}\cdot\tilde{f}\Big(\frac{ai+b}{ci+d}\Big).$$

Dan is h een automorfe eigenvorm en er geldt $T_p\cdot h=a_p\cdot h$, waarbij a_p de p-de Fourier-coëfficiënt van f is, voor alle $p\nmid N$.

We zijn nu in staat het voorspelde verband tussen de automorfe vormen en de diofantische vergelijkingen te formuleren.

Vermoeden 6 (Langlands). *Stel X is een diofantisch systeem van vergelijkingen als in* (11). *Er bestaan positieve gehele getallen $n_1,\ldots,n_r \in \mathbb{Z}_{\geq 1}$, eigenvormen $f_1,\ldots,f_r$ van voldoende groot niveau N op de groepen $\mathrm{GL}_{n_1}(\mathbb{R})$, $\mathrm{GL}_{n_2}(\mathbb{R}),\ldots,$ $\mathrm{GL}_{n_r}(\mathbb{R})$, en gehele getallen $m_1,\ldots,m_r \in \mathbb{Z}$ zó dat voor alle $p \nmid N$, geldt*

$$\#X(\mathbb{F}_p) = \sum_{i=1}^{r} m_i \cdot a_p(f_i), \qquad (14)$$

waarbij $a_p(f_i)$ de eigenwaarde is van de operator T_p werkend op f_i.

Voor de experts refereren we naar [24, Theorem 6.1] en de vermoedens in [28, Section 4]. De oplettende lezer ziet dat Vermoeden 6 slechts één richting op gaat: van diofantisch naar automorf. De originele vermoedens van Langlands gaan ook in de omgekeerde richting.

Verder, in bovenstaand vermoeden hebben we een eindige hoeveelheid priemgetallen uitgesloten. Neem als voorbeeld een affiene Weierstrass-vergelijking X van een elliptische kromme die niet minimaal is. Dan kunnen we nog steeds punten tellen op X, alleen bij de (eindig veel) priemgetallen waar de vergelijking niet minimaal is, hoeven aantallen niet gelijk te zijn aan $p - a_p(f)$ waar f de corresponderende eigenvorm is. Verder merken we op dat de functie $g(x) = |x|^{-1}$ op $\mathrm{GL}_1(\mathbb{R}) = \mathbb{R}^\times$ een automorfe eigenvorm is met eigenwaarde p voor T_p op GL_1. We zien: de getallenrij $\#X(\mathbb{Z}/p\mathbb{Z})$ wordt beschreven door de twee automorfe vormen f en g. Wegens Stelling 3, een directe generalisatie van het resultaat van Wiles, weten we nu dus dat bovenstaand vermoeden waar is voor X.

Referenties

1 Armand Borel en Herve Jacquet, Automorphic forms and automorphic representations, *Automorphic forms, representations and L-functions*, Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, OR, 1977, Part 1, 1979.

2 Alex van den Brandhof, Abelprijs gaat naar Andrew Wiles, voor oplossing Fermat, *NRC Handelsblad*, 15 maart 2016.

3 Christophe Breuil, Brian Conrad, Fred Diamond en Richard Taylor, On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises, *J. Amer. Math. Soc.* 14(4) (2001), 843–939.

4 Ana Caraiani en Peter Scholze, On the generic part of the cohomology of compact unitary Shimura varieties, *Annals of Math.*, arXiv:1511.02418.

5 Davide Castelvecchi, Fermat's last theorem earns Andres Wiles the Abel prize, *Nature*, March 15 2016.

6 John H. Coates en Andrew J. Wiles, On the conjecture of Birch and Swinnerton-Dyer, *Invent. Math.* 39(3) (1977), 223–251.

7 Henri Darmon en Loïc Merel, Winding quotients and some variants of Fermat's last theorem, *J. Reine Angew. Math.* 490 (1997), 81–100.

8 Fred Diamond en Jerry Shurman, *A First Course in Modular Forms*, Graduate Texts in Mathematics, No. 228, Springer, 2005, xvi+436 pp.

9 Bas Edixhoven, Rational elliptic curves are modular, *Séminaire N. Bourbaki, 1999–2000*, exp. no. 871, pp. 161–188.

10 Jordan S. Ellenberg, Galois representations attached to $\mathbb{Q}$-curves and the generalized Fermat equation $A^4 + B^2 = C^p$, *Amer. J. Math.* 126(4) (2004), 763–787.

11 Nuno Freitas, Bao V. Le Hung en Samir Siksek, Elliptic curves over real quadratic fields are modular, *Invent. Math.* 201(1) (2015), 159–206.

12 Nuno Freitas en Samir Siksek, The asymptotic Fermat's Last Theorem for five-sixths of real quadratic fields, *Compositio Mathematica* 151(8) (2015), 1395–1415.

13 Rishi Iyengar, British professor gets math's top prize for proving a 350-year-old theorem, *Time Magazine*, March 18 2016.

14 Chandrashekhar Khare en Jean-Pierre Wintenberger, Serre's modularity conjecture. I, II, *Invent. Math.* 178(3) (2009), 485–504, 505–586.

15 Mark Kisin, Modularity of 2-adic Barsotti-Tate representations, *Invent. Math.* 178(3) (2009), 587–634.

16 Robert Langlands, Letter to André Weil, available at publications.ias.edu/rpl/section/21.

17 Derrick H. Lehmer, The vanishing of Ramanujan's function $\tau(n)$, *Duke Math. J.* 14 (2017), 429–433.

18 Frans Oort, The Weil conjectures, *Nieuw Archief Wisk.* 5/15(3) (2014), 211–219.

19 Bjorn Poonen, Edward F. Schaefer en Michael Stoll, Twists of $X(7)$ and primitive solutions to $x^2 + y^3 = z^7$, *Duke Math. J.* 137(1) (2007), 103–158.

20 Kenneth A. Ribet, On modular representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ arising from modular forms, *Invent. Math.* 100(2) (1990), 431–476.

21 Kenneth A. Ribet, Abelian varieties over $\mathbb{Q}$ and modular forms, *Algebra and Topology 1992 (Taejon)*, Korea Adv. Inst. Sci. Tech., Taejon, 1992, pp. 53–79.

22 Peter Scholze, On torsion in the cohomology of locally symmetric varieties, *Ann. of Math. (2)* 182(3) (2015), 945–1066.

23 Jean-Pierre Serre, Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, *Duke Math. J.* 54(1) (1987), 179–230.

24 Jean-Pierre Serre, *Lectures on $N_X(p)$*, CRC Press, 2011.

25 Joseph H. Silverman, *The Arithmetic of Elliptic Curces*, Graduate Text of Mathematics, No. 106, Springer, 2009, 2nd edition.

26 Joseph H. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*, Graduate Text of Mathematics, No. 151, Springer, 1994.

27 Richard L. Taylor, Potential automorphy of some non-self dual Galois representations, www.math.princeton.edu/events/seminars/princeton-universityias-number-theory-seminar/potential-automorphy-some-non-self.

28 Richard L. Taylor, Galois representations, *Annales de la Faculte des Sciences de Toulouse 13* (2004), 73–119.

29 Richard L. Taylor en Andrew J. Wiles, Ring-theoretic properties of certain Hecke algebras, *Ann. of Math. (2)* 141(3) (1995), 553–572.

30 Andrew J. Wiles, Modular elliptic curves and Fermat's last theorem, *Ann. of Math. (2)* 141(3) (1995), 443–551.