Articles

Incloem la traducció que Anna Río Doval, del Departament de Matemàtica Aplicada II, ha fet de l'article sobre la prova del teorema de Fermat que Gerd Faltings va publicar recentment en els *Notices* de la AMS. Agraïm a la traductora el seu esforç i generositat.

La prova de R. Taylor i A. Wiles del teorema de Fermat

GERD FALTINGS
Max-Planck-Institut für Mathematik
Bonn, Alemanya.

La prova de la conjectura esmentada al títol va ser completada finalment el setembre de 1994. A. Wiles va anunciar aquest resultat l'estiu de 1993; això no obstant, hi havia un forat al seu treball. L'article de Taylor i Wiles no tapa aquest forat, sinó que l'esquiva. Aquest escrit és una adaptació de diverses xerrades que he donat sobre aquest tema i no tracta de cap manera sobre el meu propi treball. He intentat presentar les idees bàsiques per a una audiència matemàtica més àmplia i en el procés he deixat de banda alguns detalls que, segons la meva opinió, no són de massa interès per als no-especialistes. Els especialistes poden alleujar el seu avorriment trobant aquests errors i corregint-los.

Corbes el·líptiques

Per als nostres propòsits, una corba el·líptica E és donada com el conjunt de solucions $\{x, y\}$ d'una equació $y^2 = f(x)$, on $f(x) = x^3 + \ldots$ és un polinomi de grau tres. Normalment E està definida sobre els nombres racionals $\mathbf{Q}$; és a dir, els coeficients de f estan a $\mathbf{Q}$. Demanem també que els tres zeros de f siguin diferents (E és "no singular"). Podem considerar E com les solucions a $\mathbf{Q}$, $\mathbf{R}$ o $\mathbf{C}$, que es denoten per $E(\mathbf{Q})$, $E(\mathbf{R})$ i $E(\mathbf{C})$, respectivament. Normalment s'inclou en aquest conjunt un punt infinitament distant, denotat per ∞. Amb aquest afegitó, el conjunt de solucions té estructura de grup abelià, amb ∞ com a element neutre. L'invers de (x, y) és $(x, -y)$ i la suma de tres punts és nul·la si estan alineats. L'operació d'addició és donada per funcions algebraiques. Com a grup, $E(\mathbf{Q})$ és finitament generat (Teorema de Mordell), $E(\mathbf{R})$ és isomorf a $\mathbf{R}/\mathbf{Z}$ o a $\mathbf{R}/\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ i $E(\mathbf{C}) \simeq \mathbf{C}/\text{ret}$ (per exemple, $y^2 = x^3 - x$ dóna lloc a la ret $\mathbf{Z} \oplus \mathbf{Z}i$).

Per a un enter n, denotem per $E[n]$ els punts de n-torsió, és a dir, el nucli de la multiplicació per n. Sobre $\mathbf{C}$ aquests són isomorfs a $(\mathbf{Z}/n\mathbf{Z})^2$ i les coordenades són nombres algebraics. Per exemple, els punts de 2-torsió són exactament ∞ i els tres zeros de f (on $y = 0$). El grup de Galois absolut $\text{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ opera en ells, ja que les equacions que els determinen tenen coeficients a $\mathbf{Q}$. Això determina una representació $\text{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \text{GL}_2(\mathbf{Z}/n\mathbf{Z})$. Mitjançant un canvi de coordenades hom pot aconseguir que f tingui coeficients enters. Llavors, en reduir mòdul un nombre primer p s'obté un polinomi sobre el cos finit $\mathbf{F}_p$. Si els zeros del polinomi reduït són diferents, aleshores això dóna lloc a una corba el·líptica sobre $\mathbf{F}_p$. Això és cert per a tots els nombres primers p excepte per a un nombre finit: els divisors primers del discriminant de f. A més, l'elecció de f no és única. Diem que E té *bona reducció en p* si podem trobar un f tal que els seus zeros mòdul p siguin diferents. (Aquestes observacions no són del tot certes si $p = 2$, a causa del terme y^2.) En cas contrari, E té *mala reducció en p*. En aquest cas, si només hi ha coincidència entre dos dels zeros de f mòdul p, diem que E té *mala reducció semiestable*. E s'anomena *semiestable* si en tot p té reducció bona o semiestable. La corba $y^2 = x^3 - x$ no és semiestable en $p = 2$ (cap corba amb multiplicació complexa és semiestable).

Un exemple (que al final no existirà) de corba semiestable és la corba de Frey. A una solució de l'equació de Fermat $a^l + b^l = c^l$ (on a, b, c són relativament primers i $l \geq 3$ és primer) s'hi associa la corba

$$E: \; y^2 = x(x - a^l)(x - c^l).$$

Aquesta corba té mala reducció exactament en els divisors primers de abc. Té la notable propi-

etat següent: considerem al representació galoisiana associada $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{F}_l)$. Aquesta representació és no ramificada (l'anàleg de "bona reducció") en tots els nombres primers p en els quals E té bóna reducció. Aquí potser s'hauria de dir "cristal·lina" en lloc de "no ramificada" si $p = l$. A causa de la forma particular de l'equació de E, això també és cert per a tots els divisors primers $p > 2$ de abc. Per consegüent, els punts de l-torsió es comporten com si E tingués bona reducció en tot $p > 2$. Però, com veurem, no hi ha corbes el·líptiques sobre $\mathbf{Q}$ semiestables amb aquesta propietat, i aquesta és la contradicció desitjada.

Per tal d'assolir l'objectiu per aquesta via, cal reemplaçar corbes el·líptiques per formes modulars. Que això es pugui fer es desprèn de la conjectura de Taniyama-Weil (que es deu essencialment a Shimura). Si E satisfà la conclusió d'aquesta conjectura, és a dir, si E és "modular", aleshores, segons un teorema de K. Ribet, hom pot trobar una forma modular per a $\Gamma_0(2)$ que correspon a la representació de $E[l]$. Tanmateix, no existeixen unes tals formes modulars. El contingut dels articles de R. Taylor i A. Wiles és exactament la demostració de la conjectura de Taniyama-Weil per a les corbes el·líptiques sobre $\mathbf{Q}$ semiestables. Per tal d'explicar això ens calen uns quants resultats bàsics sobre formes modulars.

Formes modulars

Sigui $\mathbf{H} = \{\tau \in \mathbf{C} \mid \mathrm{Im}(\tau) > 0\}$ el semiplà superior, en el qual opera $\mathrm{SL}(2, \mathbf{R})$ segons la regla usual $(a\tau + b)/(c\tau + d)$. El subgrup $\Gamma_0(N)$ de $\mathrm{SL}(2, \mathbf{Z})$ està format per les matrius

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

amb $c \equiv 0 \pmod{N}$. Una forma modular (de pes 2) per a $\Gamma_0(N)$ és una funció $f(\tau)$, holomorfa a $\mathbf{H}$, tal que

$$f\big((a\tau + b)/(c\tau + d)\big) = (c\tau + d)^2 f(\tau)$$

per a

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$$

i $f(\tau)$ és "holomorfa a les puntes". Aquesta darrera condició implica, en particular, que a la sèrie de Fourier (ja que $f(\tau + 1) = f(\tau)$)

$$f(\tau) = \sum_{n \in \mathbf{Z}} a_n e^{2\pi i n \tau}$$

són nuls tots els coeficients a_n amb $n < 0$. Si, a més, $a_0 = 0$, aleshores f s'anomena *forma parabòlica*. L'àlgebra de Hecke $\mathbf{T}$ opera a l'espai de formes parabòliques. Està generada pels operadors T_p ($p \nmid N$ primer) i U_p ($p \mid N$). Per als coeficients de Fourier es té

$$a_n(T_p f) = a_{np}(f) + p a_{n/p}(f),$$
$$a_n(U_p f) = a_{np}(f).$$

Una forma pròpia és un vector propi comú de tots els operadors de Hecke. Hom sempre pot normalitzar-la de manera que $a_1(f) = 1$ i llavors $a_p(f)$ és el valor propi corresponent de T_p o U_p. Les equacions anteriors permeten determinar tots els a_n recursivament i, per tant, es pot determinar la forma pròpia f. Recíprocament, per a un sistema de valors propis $\{a_p\}$ donat es pot construir una sèrie de Fourier $f(\tau) = \sum a_n e^{2\pi i n \tau}$. Segons un teorema de A. Weil, això és una forma modular si, i només si, la L-sèrie $L(s, f) = \sum a_n n^{-s}$ té prolongació holomorfa a tot el pla complex i satisfà una determinada equació funcional. (Això també ha de ser cert per als torcements per caràcters de Dirichlet.)

En cas que tots els a_p pertanyin a $\mathbf{Q}$, la forma pròpia f té associada una corba el·líptica E amb bona reducció fora dels divisors primers de N. Si $p \nmid N$, el nombre de punts $\mathbf{F}_p$-racionals $E(\mathbf{F}_p)$ és igual a $\sharp E(\mathbf{F}_p) = p + 1 - a_p$. Recíprocament, per a cada corba el·líptica E sobre $\mathbf{Q}$ es pot definir una L-sèrie de Hasse-Weil $L(E, s)$ i es conjectura que aquesta té les bones propietats anteriors. Per tant, d'acord amb un teorema de A. Weil hauria de correspondre a una forma pròpia amb valors propis racionals. Aquest és el contingut de la conjectura de Taniyama-Weil.

Fins i tot si els coeficients no pertanyen a $\mathbf{Q}$, es pot construir una representació de Galois associada a la forma pròpia.

L'àlgebra de Hecke $\mathbf{T}$ és un $\mathbf{Z}$-mòdul finitament generat. La reemplacem ara per la seva completació $\hat{\mathbf{T}}$ en un ideal maximal $\mathbf{m}$ convenient (un "ideal no d'Eisenstein"), i denotem $\kappa = \mathbf{T}/\mathbf{m}$ el seu cos residual, de característica l. Aleshores, existeix una representació de

Galois 2-dimensional

$$\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\widehat{\mathbf{T}}),$$

que és no ramificada (o cristal·lina, respectivament) a $p \nmid N$, amb

$$\text{traça}(\rho(\mathrm{Frob}_p)) = T_p,$$
$$\det(\rho(\mathrm{Frob}_p) = p.$$

Una forma pròpia amb coeficients racionals dóna lloc a un homomorfisme $\widehat{\mathbf{T}} \to \mathbf{Z}_l$ i ρ indueix la representació l-àdica que descriu l'acció galoisiana a tots els punts de l^n-torsió de la corba el·líptica associada E. Recíprocament, és possible provar que E és modular si, i només si, la representació l-àdica associada es pot construir d'aquesta manera.

Deformacions

La representació l-àdica es construeix per a $l = 3$, començant amb la representació en els punts de 3-torsió. És conegut que aquesta és congruent a una representació modular, i llavors es prova que l'aixecament universal d'aquesta representació és modular, la qual cosa és el nucli de la demostració. El primer 3 és aquí molt especial. Per tant, es comença considerant $l = 3$.

Hom es pot restringir al cas que els punts de 3-torsió donen lloc a una aplicació exhaustiva

$$\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, \mathbf{F}_3)$$

(en aquest argument s'utilitzen també una vegada els punts de 5-torsió). Atès que $\mathrm{PGL}(2, \mathbf{F}_3) \simeq S_4$ (el grup simètric en els quatre elements de $\mathbf{P}^1(\mathbf{F}_3)$) és resoluble, la representació en els punts de 3-torsió és ja modular, segons teoremes ("d'aixecament") de Langlands i Tunnell. Això utilitza intensament les propietats especials del nombre primer $l = 3$. Per a $l = 2$, la teoria general no funciona bé per raons diverses, i per a $l = 5$, aquest començament és impossible. Busquem ara un argument de deformació perquè les representacions mòdul 9, 27, 81, 243, 729, etc., siguin reconegudes successivament com a modulars. Per a això s'utilitza la deformació universal de la representació mòdul 3: existeix una $\mathbf{Z}_3$-àlgebra $\mathcal{R}$ de la forma $\mathcal{R} = \mathbf{Z}_3[[T_1, \ldots, T_r]]/I$ (I és un ideal) i una representació galoisiana "universal"

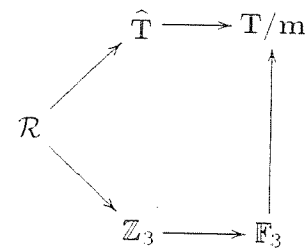
$$\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathcal{R})$$

amb les propietats següents:

- ρ és no ramificada (o cristal·lina, respectivament) a $p \nmid N$ (és a dir, E té bona reducció a p);
- ρ té certes propietats locals a $p \nmid N$ ("certes" no serà tractat aquí);
- $\det(\rho(\mathrm{Frob}_p)) = p$ per a $p \nmid N$;
- $\rho \mod (3, T_1, \ldots, T_r)$ és la nostra representació donada en $E[3]$;
- qualsevol altra representació $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathcal{A})$ amb les propietats 1)-4) prové, de manera única, d'un homomorfisme $\mathcal{R} \to \mathcal{A}$.

La construcció de $\mathcal{R}$ segueix principis generals. Bàsicament, es pren un conjunt de generadors $\{\sigma_1, \ldots, \sigma_s\}$ del grup de Galois, es considera l'anell de sèries de potències en $4s$ variables i es divideix pel més petit ideal I tal que mòdul I s'obté una representació amb les propietats 1), ..., 4), sempre que s'assigni a σ_i la matriu 2×2 que té com a coeficients les 4 variables corresponents a σ_i.

Un cop feta la construcció obtenim el diagrama commutatiu següent:



on les dues aplicacions de l'esquerra provenen de la representació galoisiana modular i de la representació associada a E. La idea de Wiles es mostrar ara que $\mathcal{R}$ és isomorf a $\widehat{\mathbf{T}}$, perquè llavors la representació galoisiana el·líptica és automàticament modular.

Naturalment, per a això es necessita informació sobre $\mathcal{R}$ que no proveeix la construcció general. Sigui W_n la representació adjunta de $\mathrm{sl}(2, \mathbf{Z}/3^n\mathbf{Z})$ (matrius 2×2 amb traça zero). Aleshores, per exemple, el nombre minimal de generadors $r(\mathcal{R} = \mathbf{Z}_3[[T_1, \ldots, T_r]]/I)$ és donat per $\dim_{\mathbf{F}_3} H^1_f(\mathbf{Q}, W_1)$, on H^1_f denota un grup de cohomologia que satisfà certes condicions locals corresponents a 1), 2) d'abans. Aquest també

s'anomena un grup de Selmer. Es veu això prenent $\mathcal{A} = \mathbf{F}_3[T]/(T^2)$ a les definicions. Es pot provar (M. Flach) que els ordres de $H^1_f(\mathbf{Q}, W_n)$ estan uniformement fitats en n. Aquests ordres apareixen en el següent criteri numèric per a la igualtat $\mathcal{R} = \widehat{\mathbf{T}}$: existeix un $\mathbf{Z}_3$-homomorfisme $\widehat{\mathbf{T}} \to \mathcal{O}$, on $\mathcal{O}$ és la clausura entera de $\mathbf{Z}_3$ en una extensió finita de $\mathbf{Q}_3$. Per simplificar, suposarem que $\mathcal{O} = \mathbf{Z}_3$. És conegut que $\widehat{\mathbf{T}}$ és Gorenstein; és a dir, $\mathrm{Hom}_{\mathbf{Z}_3}(\widehat{\mathbf{T}}, \mathbf{Z}_3)$ és un $\widehat{\mathbf{T}}$-mòdul lliure. Llavors, l'epimorfisme $\widehat{\mathbf{T}} \to \mathbf{Z}_3$ té un adjunt $\mathbf{Z}_3 \to \widehat{\mathbf{T}}$ i la composició d'aquests dos morfismes és la multiplicació per un element $\eta \in \mathbf{Z}_3$, que està ben definit llevat d'unitats. A més, $\eta \neq 0$. D'altra banda, sigui $\mathfrak{p} \subseteq \mathcal{R}$ el nucli de l'epimorfisme $\mathcal{R} \to \widehat{\mathbf{T}} \to \mathbf{Z}_3$. Aleshores es té ("$\sharp$"=ordre) $\sharp\mathfrak{p}/\mathfrak{p}^2 \geq \sharp\mathbf{Z}_3/\eta \cdot \mathbf{Z}_3$ amb igualtat si, i només si, $\mathcal{R} = \widehat{\mathbf{T}}$ i aquest és, a més, intersecció completa (I pot ser generat per r elements). L'expressió de l'esquerra $\sharp\mathfrak{p}/\mathfrak{p}^2$ és idèntica a l'ordre del grup de Selmer $H^1_f(\mathbf{Q}, W_n)$, per a $n > 0$. El primer intent tractava d'establir la igualtat utilitzant sistemes d'Euler (inventats per Kolyvagin). Tanmateix, només va ser possible demostrar que $\mathfrak{p}/\mathfrak{p}^2$ és anul·lat per η. Aquest és el contingut del teorema de M. Flach. Això no obstant, els nivells superiors del sistema d'Euler no van poder ser construïts.

La prova

Es prova primer el "cas minimal" i després hom es redueix a aquest. Amb *cas minimal* volem dir que tots els primers de mala reducció ja en tenen mòdul 3 (i no només mòdul potències més altes). Segons el teorema de Ribet i altres (utilitzat per a $l = 3$ i no per a $l =$ exponent de l'equació de Fermat), la representació galoisiana que pertany a la corba mòdul 3 és modular de nivell 3. En el cas minimal el càlcul de característiques d'Euler (Poitou-Tate) prova que $H^1_f(\mathbf{Q}, W_1)$ i $H^2_f(\mathbf{Q}, W_1)$ tenen la mateixa dimensió r. Per a cada n es trien r nombres primers $q_1, \ldots, q_r \equiv 1 \mod 3^n$. Aleshores es procedeix a l'utilització d'un subgrup de $\Gamma_0(N)$. Aquest subgrup conté la intersecció amb $\Gamma_1(q_1 \ldots q_r)$ i el quocient és isomorf a $G = (\mathbf{Z}/3^n\mathbf{Z})^r$. L'àlgebra de Hecke associada $\widehat{\mathbf{T}}_1$ és un $\mathbf{Z}$-mòdul lliure sobre $\mathbf{Z}_3[G]$, amb G-coinvariants $\widehat{\mathbf{T}}$, i és el quocient d'un anell de representacions $\mathcal{R}_1 = \mathbf{Z}_3[[T_1, \ldots, T_r]]/I_1$, el qual, novament, pot ser generat per r elements. L'ideal I_1 és petit a causa de l'acció lliure del grup G. Ara es pren límit $n \to \infty$ i, en el límit, $\mathcal{R}_1$ i $\widehat{\mathbf{T}}_1$ esdevenen anells de sèries de potències i són iguals. A més, s'obté $\mathcal{R}$ de $\mathcal{R}_1$ i $\widehat{\mathbf{T}}$ de $\widehat{\mathbf{T}}_1$ en ambdós casos posant-hi les relacions addicionals "$\sigma_i = 1$", on $\sigma_1, \ldots, \sigma_r$ són generadors de G. Finalment, $\mathcal{R} = \widehat{\mathbf{T}}$ i és intersecció completa.

Per reduir-se al cas minimal s'estima com canvien els dos costats de la desigualtat

$$\sharp\mathfrak{p}/\mathfrak{p}^2 \geq \sharp\mathbf{Z}_3/\eta \cdot \mathbf{Z}_3$$

quan es passa de nivell M a un nivell més gran N ($M \mid N$). Per al terme de l'esquerra $\sharp H^1_f(\mathbf{Q}, W_n)$ s'afebleixen certes condicions locals i s'obté una fita superior. En el terme de la dreta es dóna un fenomen de "fusió", és a dir, de congruències entre formes velles i formes noves. Aquí, una fita inferior ha estat construïda per Ribet i Ihara. Afortunadament les dues fites coincideixen i, per tant, tot queda demostrat.