1 Introducció i breu història

Definició. Una *equació diofàntica* és una equació algebraica a coeficients enters plantejada sobre el conjunt dels nombres enters $\mathbb{Z}$. És a dir, una equació en què es restringeixen les incògnites (i, per tant, les solucions) a $\mathbb{Z}$.

Aquestes equacions reben el seu nom de Diofant d'Alexandria (s. III), en ser el primer en fer-ne un estudi tot introduint un simbolisme en l'àlgebra. Han tingut un paper fonamental al llarg de la història de les Matemàtiques i avui dia segueixen sent un camp obert i molt actiu de la Teoria de Nombres. Efectivament, només n'hi ha uns pocs tipus que estiguin perfectament estudiades i caracteritzades, això és, que se'n conegui un mètode eficient per determinar-ne la resolubilitat i obtenir-ne totes les solucions en cas que hi hagi.

És habitual d'estudiar per separat cada tipus d'equació diofàntica en funció de les seves característiques principals (el seu grau, el nombre d'incògnites, etc). El desè problema de Hilbert, d'entre els 23 que va enunciar a finals del s. XIX deia: *donada una equació diofàntica amb qualsevol nombre d'incògnites i amb coeficients enters, trobar un procés mitjançant el qual es pugui determinar, en un nombre finit de passos, si l'equació és resoluble en el conjunt dels enters.* Aquest problema fou resolt el 1970 pel matemàtic Yuri Matiyasévich de forma negativa: a partir de conceptes de la teoria de nombres i de lògica matemàtica va demostrar que no pot existir cap algorisme general com el que Hilbert demanava.

En aquesta exposició ens centrarem en les equacions diofàntiques lineals i en alguns tipus concrets d'equacions quadràtiques.

2 Equacions diofàntiques lineals

2.1 D'una incògnita: $ax = b$

Es tracta del cas més elemental. Hi haurà solució si i només si $a|b$ i, en tal cas, $x = \frac{b}{a}$.

2.2 De dues incògnites: $ax + by = c$

Per al seu estudi necessitarem un resultat previ important:

Identitat de Bézout. *Si $a, b \in \mathbb{Z} - \{0\}$ i $d = \text{mcd}(a, b)$, llavors $\exists\, x, y \in \mathbb{Z}$ tals que $ax + by = d$.*

DEMOSTRACIÓ. Considerem el conjunt $S = \{ax + by \mid x, y \in \mathbb{Z},\ ax + by > 0\}$, que trivialment és no buit. Pel principi del bon ordre a $\mathbb{N}$, n'existeix un element mínim que anomenarem d. Per tant, $d = ax + by$ per uns certs x, y. Aplicant l'algorisme de la divisió entera a a i d obtenim $a = qd + r$ amb $0 \leq r < d$, i podem escriure $r = a - qd = a - q(ax + by) = a(1 - qx) + b(-qy)$, de manera que si fos $r > 0$ tindríem que $r \in S$ en contradicció amb el fet que d era el mínim. Per tant, $r = 0$ i $d|a$. El mateix raonament fent la divisió entera entre b i d demostra que $d|b$. Finalment, si d' és un altre divisor comú de a i b, llavors també $d'|ax + by = d$, i per tant $d = \text{mcd}(a, b)$. □

Teorema. *L'equació diofàntica $ax + by = c$ té solució si i només si $d = \text{mcd}(a, b)|c$. En tal cas, si x', y' és una solució particular, llavors la solució general és*

$$x = x' - tb'$$
$$y = y' + ta'$$

on $t \in \mathbb{Z}$, i a', b' són tals que $a = da'$, $b = db'$.

DEMOSTRACIÓ. Si x, y una solució i $d = \text{mcd}(a, b)$, llavors també $d|ax + by = c$. Recíprocament, sigui $d = \text{mcd}(a, b)$ amb $d|c$, és a dir, $c = dc'$. Per la identitat de Bézout, existeixen x, y tals que $ax + by = d$. Multiplicant aquesta igualtat per c' obtenim $axc' + byc' = dc' = c$, de manera que xc', yc' és una solució.

Suposem ara que l'equació és resoluble, i sigui x', y' una solució particular. Si x, y és la solució general tindrem que $ax + by = c$, $ax' + by' = c$. Restant ambdues expressions tenim $a(x - x') + b(y - y') = 0$, i dividint aquesta expressió per d, $a'(x - x') + b'(y - y') = 0$. Com $\text{mcd}(a', b') = 1$, pel teorema d'Euclides podem afirmar que $a'|(y - y')$ i $b'|(x - x')$, això és, $y = y' + ta'$, $x = x' + ub'$ amb $t, u \in \mathbb{Z}$. Ara bé,

$$c = ax + by = a(x' + ub') + b(y' + ta') = ax' + by' + ab'u + a'bt = c + ab'u + a'bt,$$

per tant $0 = ab'u + a'bt = a'db'u + a'b'dt = da'b'(u + t)$, de manera que $u = -t$ i obtenim l'expressió que volíem per a la solució general: $y = y' + ta'$, $x = x' - tb'$. □

A la pràctica, per trobar una solució particular de $ax + by = c$, comencem dividint per $d = \text{mcd}(a, b)$, i obtenim l'equació reduïda $a'x + b'y = c'$ amb les mateixes solucions que la inicial i $\text{mcd}(a', b') = 1$. Llavors es calculen r, s tals que $a'r + b's = 1$ (identitat de Bézout) mitjançant l'algorisme d'Euclides, i multiplicant-les per c' obtenim una solució particular de l'equació reduïda: $a'(rc') + b'(sc') = c'$.

L'algorisme d'Euclides és un mètode pràctic per calcular $\text{mcd}(a, b)$, que aplica recursivament

el fet que $\mathrm{mcd}(a,b) = \mathrm{mcd}(b,r)$, on r és el residu de la divisió entera de a per b:

$$\begin{array}{lll} a = bq + r & \mathrm{mcd}(a,b) = \mathrm{mcd}(b,r) & r < |b| \\ b = rq_1 + r_1 & \mathrm{mcd}(b,r) = \mathrm{mcd}(r,r_1) & r_1 < r \\ r = r_1q_2 + r_2 & \mathrm{mcd}(r,r_1) = \mathrm{mcd}(r_1,r_2) & r_2 < r_1 \\ \ldots & \ldots & \ldots \end{array}$$

Obtenim una successió decreixent de residus, per tant arribarà un moment que tindrem un residu igual a 0:

$$\begin{array}{lll} r_{k-2} = r_{k-1}\, q_k + r_k & \mathrm{mcd}(r_{k-2}, r_{k-1}) = \mathrm{mcd}(r_{k-1}, r_k) & r_k < r_{k-1} \\ r_{k-1} = r_k\, q_{k+1} + 0 & \mathrm{mcd}(r_{k-1}, r_k) = \mathrm{mcd}(r_k, 0) = r_k. & \end{array}$$

Així doncs $r_k = \mathrm{mcd}(a,b)$. Un cop desenvolupat l'algorisme d'Euclides, fent una substitució enrere podem obtenir la identitat de Bézout. En efecte, $d = r_k = r_{k-2} - r_{k-1}\, q_k$. Però r_{k-1} també es pot escriure com una suma de múltiples de r_{k-2} i r_{k-3}; r_{k-2} és suma de múltiples de r_{k-3} i r_{k-4}, i així successivament fins a obtenir la identitat de Bézout: $d = ar + bs$.

2.3 De més de dues incògnites: $a_1x_1 + \ldots + a_nx_n = c$

La identitat de Bézout és fàcilment generalitzable al cas de n enters, per tant aquí és vàlid el mateix resultat que en el cas de dues variables: l'equació general $a_1x_1 + \ldots + a_nx_n = c$ serà resoluble si i només si $d = \mathrm{mcd}(a_1, \ldots, a_n) | c$.

Equival a resoldre $n-1$ equacions de 2 incògnites, disminuint una incògnita a cada pas i finalment fent una substitució enrere: si l'equació és $a_1x_1 + a_2x_2 + \ldots + a_nx_n = c$, escrivim $d_2 = \mathrm{mcd}(a_1, a_2)$ i trobem la solució general $\bar{x_1}, \bar{x_2}$ de $a_1x_1 + a_2x_2 = d_2$. Llavors introduïm una nova variable y_2 i podem escriure $d_2y_2 + a_3x_3 + \ldots + a_nx_n = c$. Aquesta equació, ara de $n-1$ incògnites, segueix sent resoluble perquè $\mathrm{mcd}(d_2, a_3, \ldots, a_n) = \mathrm{mcd}(a_1, a_2, \ldots, a_n) | c$, i posem $d_3 = \mathrm{mcd}(d_2, a_3)$. Novament trobem la solució general $\bar{y_2}, \bar{x_3}$ de $d_2y_2 + a_3x_3 = d_3$ i escrivim $d_3y_3 + a_3x_3 + \ldots + a_nx_n = c$, i així successivament fins arribar a l'equació $d_{n-1}y_{n-1} + a_nx_n = c$, amb solució general $\bar{y_{n-1}}, \bar{x_n}$. A partir d'aquí fem substitució enrere per eliminar tots els d_i i fer que c només depengui de $a_1, \ldots, a_n$ i de les solucions generals intermèdies que hem trobat:

$$c = d_{n-1}\bar{y_{n-1}} + a_n\bar{x_n} = (d_{n-2}\bar{y_{n-2}} + a_{n-1}\bar{x_{n-1}})\, \bar{y_{n-1}} + a_n\bar{x_n} = \ldots$$

3 Equacions diofàntiques no lineals

3.1 L'equació $x^2 - y^2 = a$

Aquest és el cas més senzill d'equacions diofàntiques quadràtiques. Podem considerar $a \geq 0$, perquè en cas contrari canviem de signe tota l'equació i n'obtenim una del mateix tipus amb $a \geq 0$. També cal observar que si (x, y) n'és una solució, també ho és $(\pm x, \pm y)$, per tant podem restringir la cerca de solucions a $x \geq y \geq 0$. Observant que $x^2 - y^2 = (x+y)(x-y)$, i que $x + y$ i $x - y$ tenen la mateixa paritat (ja que $1 \equiv -1 \pmod 2$), queda clar que resoldre l'equació equival a trobar factoritzacions $a = bc$ amb b, c de la mateixa paritat. Si b, c són senars, a també ho és, i si b, c són parells, a és múltiple de 4. Cada factorització amb $b \geq c$ ens proporciona una solució $x + y = b$, $x - y = c$, això és, $x = \frac{b+c}{2}$, $y = \frac{b-c}{2}$, amb $x \geq y \geq 0$. Així doncs,

Teorema. *L'equació diofàntica $x^2 - y^2 = a$ té solució si i només si a és senar o múltiple de 4. En tal cas, cada factorització $a = bc$ amb $b \geq c$ i b, c de la mateixa paritat ens proporciona una solució $x = \frac{b+c}{2}$, $y = \frac{b-c}{2}$, amb $x \geq y \geq 0$.* □

3.2 L'equació $x^2 + y^2 = a$

Tot i la similitud d'aquesta equació amb l'anterior, el seu estudi resulta molt menys evident. Enunciem sense demostració el resultat que dóna la condició necessària i suficient per a la resolubilitat; les proves més directes treballen sobre l'aritmètica dels enters de Gauss (l'anell $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z},\ i^2 = -1\}$) i amb resultats sobre residus quadràtics mòdul p, amb p primer.

Teorema. *L'equació diofàntica $x^2 + y^2 = a$ té solució si i només si tots els factors primers de a de la forma $4m + 3$ apareixen amb un exponent parell en la seva factorització.*

3.3 L'equació de Pell: $x^2 - dy^2 = N$

Sovint s'anomena *equació de Pell* al cas particular $N = 1$, mentre que la forma general $x^2 - dy^2 = N$ se sol anomenar *equació de Pell generalitzada*. Fou Euler qui atribuí erròniament l'equació a John Pell, ja que el primer matemàtic europeu en trobar-ne una solució general fou Lord Brouncker, tot i que l'equació ja havia estat àmpliament estudiada en els temps de l'Índia antiga. Concretament, Brahmagupta (s. VII) descobrí la identitat

$$(x_1^2 - dy_1^2)(x_2^2 - dy_2^2) = (x_1x_2 + dy_1y_2)^2 - d(x_1y_2 + x_2y_1)^2,$$

que li va permetre construir un nombre il·limitat de solucions de l'equació a partir de solucions particulars. Bashkara (s. XII) va estendre els mètodes de Brahmagupta per trobar mètodes generals de resolució per al cas $N = 1$. No fou fins al XVIII que Lagrange completà totalment l'estudi dels casos $N = \pm 1$, en què donà condicions necessàries i suficients per a la resolubilitat, i un mètode per trobar totes les solucions mitjançant el desenvolupament en forma de fracció contínua de $\sqrt{d}$.

Tot seguit veiem unes definicions prèvies i els principals resultats (molts sense demostració donada la seva extensió).

Definició. Anomenem *fracció contínua simple* a l'expressió

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ldots + \cfrac{1}{a_n}}}}$$

on $a_i \in \mathbb{Z}^+$ per a $i \geq 1$, $a_0 \in \mathbb{Z}$. La notem com $\langle a_0, a_1, \ldots, a_n \rangle$. Si $\{a_n\}_{n \geq 0}$ és una successió d'enters que compleixen les condicions citades, definim la *fracció contínua simple infinita* $\langle a_0, a_1, a_2, \ldots \rangle$ com $\lim_{n \to \infty} \langle a_0, a_1, \ldots, a_n \rangle$. Cada $\langle a_0, a_1, \ldots, a_n \rangle$ defineix un nombre racional p_n/q_n que anomenem *n-èsim convergent* de la fracció contínua.

Teorema. *Tot nombre irracional ξ té una única representació en forma de fracció contínua infinita, i recíprocament, tota fracció contínua infinita defineix un únic nombre irracional. A més, l'irracional ξ és quadràtic si i només si la seva fracció contínua és periòdica, això és, de la forma $\langle b_0, \ldots, b_k, \overline{a_1, \ldots, a_r} \rangle$. Concretament, la fracció contínua de $\sqrt{d}$, amb $d \in \mathbb{Z}$ no quadrat, és de la forma $\langle a_0, \overline{a_1, \ldots, a_r} \rangle$.*

Teorema. *Totes les solucions positives de $x^2 - dy^2 = \pm 1$ són de la forma $x = p_n$, $y = q_n$, on p_n/q_n és un convergent de l'expressió de $\sqrt{d}$ com a fracció contínua. Si, en la notació del teorema anterior, r és el període de la fracció contínua, i r és parell, llavors $x^2 - dy^2 = -1$ no té solució, i totes les solucions positives de $x^2 - dy^2 = 1$ vénen donades per $x = p_{nr-1}$, $y = q_{nr-1}$ amb $n = 1, 2, 3, \ldots$ D'altra banda, si r és senar, llavors $x = p_{nr-1}$, $y = q_{nr-1}$ donen totes les solucions positives de $x^2 - dy^2 = -1$ amb $n = 1, 2, 3, \ldots$, i totes les de $x^2 - dy^2 = 1$ amb $n = 2, 4, 6, \ldots$*

Aquest és un potent resultat que assegura la resolubilitat de l'equació per a $N = 1$ i permet decidir-la per a $N = -1$, a més de proporcionar un mètode per trobar totes les solucions en ambdós casos. Tanmateix, si ja hem trobat la mínima solució positiva per a $N = \pm 1$, en podem trobar totes les altres amb mètodes més simples:

Teorema. *Si x_1, y_1 és la mínima solució positiva de $x^2 - dy^2 = 1$, llavors totes les solucions positives vénen donades per x_n, y_n amb $n = 1, 2, 3 \ldots$ on x_n, y_n són els enters definits per $x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n$.*

Demostració. Primer comprovem que x_n, y_n és una solució. Com el conjugat d'un producte és el producte de conjugats, tenim que $x_n - y_n\sqrt{d} = (x_1 - y_1\sqrt{d})^n$, i llavors podem escriure

$$x_n^2 - dy_n^2 = (x_n - y_n\sqrt{d})(x_n + y_n\sqrt{d}) = (x_1 - y_1\sqrt{d})^n(x_1 + y_1\sqrt{d})^n = (x_1^2 - dy_1^2)^n = 1.$$

Ara mostrem que tota solució positiva s'obté d'aquesta manera. Suposem que existeix una solució positiva s, t que no es troba entre els $\{x_n, y_n\}$. Llavors existirà un enter m tal que $(x_1 + y_1\sqrt{d})^m < s + t\sqrt{d} < (x_1 + y_1\sqrt{d})^{m+1}$. Com $(x_1 + y_1\sqrt{d})^{-m} = (x_1 - y_1\sqrt{d})^m$, multipliquem tota la desigualtat per aquesta expressió, i obtenim

$$1 < (s + t\sqrt{d})(x_1 - y_1\sqrt{d})^m < x_1 + y_1\sqrt{d}.$$

Si anomenem $a + b\sqrt{d} = (s + t\sqrt{d})(x_1 - y_1\sqrt{d})^m$, comprovem que

$$a^2 - db^2 = (s^2 - dt^2)(x_1^2 - dy_1^2)^m = 1,$$

de manera que a, b és una solució de l'equació tal que $1 < a + b\sqrt{d} < x_1 + y_1\sqrt{d}$, i si provem que és positiva, això és, $a, b > 0$, haurem arribat a una contradicció perquè, per hipòtesi, x_1, y_1 era la mínima. I efectivament, com $1 < a + b\sqrt{d}$ deduïm que $0 < a - b\sqrt{d} < 1$ (perquè el producte de tots dos és 1), i llavors $a = \frac{1}{2}(a + b\sqrt{d}) + \frac{1}{2}(a - b\sqrt{d}) > \frac{1}{2} + 0 > 0$, i també $b\sqrt{d} = \frac{1}{2}(a + b\sqrt{d}) - \frac{1}{2}(a - b\sqrt{d}) > \frac{1}{2} - \frac{1}{2} = 0$, per tant $b > 0$, com volíem veure. □

Teorema. *Si x_1, y_1 és la mínima solució positiva de $x^2 - dy^2 = -1$, llavors x_2, y_2, definits per $x_2 + y_2\sqrt{d} = (x_1 + y_1\sqrt{d})^2$, és la mínima solució positiva de $x^2 - dy^2 = 1$. A més, totes les solucions positives de $x^2 - dy^2 = -1$ vénen donades per x_n, y_n definits per $x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n$, amb $n = 1, 3, 5, \ldots$, i totes les de $x^2 - dy^2 = 1$ amb $n = 2, 4, 6, \ldots$*

La demostració que x_2, y_2 és la mínima solució positiva per a $N = 1$ segueix la mateixa línia que en el teorema anterior, això és, suposar que no ho és i arribar a una contradicció construint una solució per a $N = -1$ inferior a x_1, y_1. La resta és conseqüència del teorema de caracterització de les solucions a partir dels convergents de la fracció contínua de $\sqrt{d}$.

Exemple. Volem resoldre l'equació $x^2 - 82y^2 = 1$. Amb una simple inspecció veiem que $x = 9$, $y = 1$ és la mínima solució positiva de $x^2 - 82y^2 = -1$. Llavors calculem $(9 + \sqrt{82})^2 = 163 + 18\sqrt{82}$, i en virtut del darrer resultat podem concloure que $x = 163$, $y = 18$ és la mínima solució positiva de $x^2 - 82y^2 = 1$.

Quan $N \neq \pm 1$ es poden provar certs resultats per a $x^2 - dy^2 = N$, però no són tan complets com els que hem vist. Per exemple, si $r^2 - ds^2 = N$, i x_1, y_1 és la mínima solució

de $x^2 - dy^2 = 1$, llavors r_n, s_n definits per $r_n + s_n\sqrt{d} = (r + s\sqrt{d})(x_1 + y_1\sqrt{d})^n$ és una solució de $x^2 - dy^2 = N$ per a tot n. Tanmateix, res no assegura que amb aquest mètode trobem totes les solucions partint d'un r, s. Pel que fa a la resolubilitat, si $|N| < \sqrt{d}$ existeix un mètode senzill per decidir si l'equació és resoluble a partir de càlculs amb els a_i de la fracció contínua de $\sqrt{d}$. Si ho és, totes les solucions positives són de la forma $x = p_n$, $y = q_n$, on p_n/q_n és un convergent. Si $|N| > \sqrt{d}$, els procediments són molt més complexos.

3.4 L'equació pitagòrica: $x^2 + y^2 = z^2$

S'anomena així per la seva similitud amb el teorema de Pitàgores. Les solucions no trivials d'aquesta equació (això és, aquelles on cap de les incògnites val 0) s'anomenen *ternes pitagòriques*, i les més conegudes són $(3, 4, 5)$ i $(5, 12, 13)$.

Ens proposem cercar totes les solucions x, y, z positives. Observem que si $d|x$ i $d|y$, també $d|z$ i llavors $\frac{x}{d}, \frac{y}{d}, \frac{z}{d}$ és una solució reduïda. Per tant, ens limitarem a cercar solucions en què $\mathrm{mcd}(x, y) = 1$, anomenades ternes pitagòriques *primitives*. D'altra banda, si x, y són senars, llavors $z^2 \equiv 2 \pmod 4$, que és impossible perquè els únics residus quadràtics mòdul 4 són 0 o 1. Per tant, x, y han de tenir diferent paritat; sense pèrdua de generalitat, suposarem que x és parell i y senar.

Teorema. *La solució general de l'equació $x^2 + y^2 = z^2$ que satisfà les condicions $x > 0$, $y > 0$, $z > 0$, $\mathrm{mcd}(x, y) = 1$ i x parell és*

$$x = 2ab, \quad y = a^2 - b^2, \quad z = a^2 + b^2,$$

amb $a > b > 0$ enters de diferent paritat i $\mathrm{mcd}(a, b) = 1$.

Demostració. Com x ha de ser parell i $\mathrm{mcd}(x, y) = 1$, y i z han de ser senars amb $\mathrm{mcd}(y, z) = 1$, de manera que $\frac{1}{2}(z - y)$ i $\frac{1}{2}(z + y)$ són enters i primers entre ells. Llavors podem reescriure l'equació pitagòrica de la següent manera:

$$\left(\frac{x}{2}\right)^2 = \frac{z + y}{2} \cdot \frac{z - y}{2},$$

i com els dos factors són coprimers, han de ser quadrats. Anomenem $a^2 = \frac{1}{2}(z + y)$ i $b^2 = \frac{1}{2}(z - y)$, i llavors $a > b > 0$ i $\mathrm{mcd}(a, b) = 1$. A més, com $a^2 + b^2 = z$ senar, deduïm que a^2 i b^2 tenen diferent paritat, i per tant el mateix passa amb a i b. De manera que $z = a^2 + b^2$, $y = a^2 - b^2$ i $x^2 = 4a^2b^2$, això és, $x = 2ab$, i a, b verifiquen les condicions de l'enunciat.

Recíprocament, donats a, b amb les condicions de l'enunciat, es comprova fàcilment que $x^2 + y^2 = (2ab)^2 + (a^2 - b^2)^2 = (a^2 + b^2)^2 = z^2$, amb $x, y, z > 0$ i x parell. Per veure que

$\mathrm{mcd}(x, y) = 1$ suposem que $d|x$ i $d|y$; llavors també $d|z$, això és, $d|a^2 - b^2$ i $d|a^2 + b^2$, per tant $d|2a^2$ i $d|2b^2$. Com a, b són coprimers, també ho són a^2, b^2, de manera que d només pot ser 1 o 2, i la segona possibilitat queda descartada perquè y, z són senars en ser a, b de diferent paritat. □

3.5 L'equació de Fermat: $x^n + y^n = z^n$ amb $n > 2$

Pierre de Fermat (s. XVII) va enunciar que no existia cap solució x, y, z no trivial a l'equació $x^n + y^n = z^n$ si $n > 2$. Afirmava tenir una demostració, però mai no la va fer pública i aquest problema, conegut com l'*últim teorema de Fermat*, es convertí en un dels més grans reptes per als matemàtics posteriors. L'any 1995, el matemàtic anglès Andrew Wiles va aconseguir provar-lo fent servir eines modernes de teoria de nombres i de geometria algebraica, d'una gran complexitat. De fet, Wiles no demostrà el teorema en si, sinó l'anomenada conjectura de Taniyama-Shimura, enunciada a la dècada dels 50, que afirmava que tota corba el·líptica tenia associada una forma modular (un determinat tipus de funció analítica). A la dècada dels 80, Gerhard Frey i Jean-Pierre Serre es van adonar que la conjectura de Taniyama-Shimura implicava el teorema de Fermat: partint d'una eventual solució de l'equació de Fermat era possible construir una corba el·líptica que semblava que no era modular. Ken Ribet ho va acabà provant el 1986, i a partir d'aquí tots els esforços es van encaminar a demostrar la conjectura de Taniyama-Shimura.

És evident que n'hi ha prou en provar el teorema de Fermat per a $n = 4$ i per a $n = p$ primer. En efecte, si sabem que és cert en aquests dos casos i $n > 4$ és un nombre compost, podem escriure $n = ab$ i $x^n + y^n = z^n$ com $(x^a)^b + (y^a)^b = (z^a)^b$, amb $b = 4$ o b primer, i per tant tota possible solució per a n ens en proporcionaria una per a b.

Els únics casos "senzills" del teorema de Fermat són per a $n = 4$ i $n = 3$. En el cas $n = 4$ es parteix d'una solució x, y, z mínima, i mitjançant la tècnica del descens infinit s'arriba a construir una d'inferior, en contradicció amb allò que s'havia suposat. Per al cas $n = 3$, es comença factoritzant $x^3 + y^3 = (x + y)(x + \rho y)(x + \rho^2 y)$, on ρ és l'arrel cúbica primitiva complexa de 1, i s'estudia l'estructura dels factors en l'anell $\mathbb{Z}[\rho]$ per arribar, també, a una contradicció. Lamé i altres matemàtics del s. XIX van provar aquesta mateixa tècnica començant amb una factorització de $x^p + y^p$ que involucra arrels p-èsimes complexes de la unitat, però posteriorment es va veure que no era correcte perquè hi ha molts casos en què l'anell $\mathbb{Z}[\rho]$ no és un domini de factorització única, fet imprescindible per aplicar raonaments similars al cas $n = 3$.