Három út a Fermat–Euler-hegyre

Lagrange, Zagier és Minkowski vezetésével

Vlagyimir Tikhomirov

Tekintsük az első néhány 2-nél nagyobb prímet: 3, 5, 7, 11, 13, 17, 19, ... Az 5, a 13 és a 17 felírható két négyzetszám összegeként:

$$5 = 1^2 + 2^2,$$
$$13 = 2^2 + 3^2,$$
$$17 = 1^2 + 4^2,$$

míg a másik négy (3, 7, 11, 19) nem. (Ellenőrizze le!) Van-e mód rá, hogy megállapítsuk, melyik prím bontható fel és melyik nem, anélkül, hogy az összes lehetséges esetet végig kéne néznünk? És miben mások a felbontható prímek? Ezekre a kérdésekre ad választ a következő tétel.

Tétel. Egy kettőnél nagyobb prímszám akkor és csak akkor bontható fel két négyzetszám összegére, ha néggyel osztva egy maradékot ad.

(Valóban: $5 = 1 \cdot 4 + 1$, $13 = 3 \cdot 4 + 1$, $17 = 4 \cdot 4 + 1$, viszont $3 = 0 \cdot 4 + 3$, $7 = 1 \cdot 4 + 3$, $11 = 2 \cdot 4 + 3$, ...)

1. feladat. *Bizonyítsuk be a tétel „csak akkor" részét: ha egy páratlan prím felbontható két négyzetszám összegére, akkor 1 maradékot ad 4-gyel osztva.*

Kinek tűnt fel először ez? Bizonyítékaink vannak rá, hogy nemrég ünnepelhettük volna a felfedezés 350. évfordulóját. 1640 karácsonyán a nagy Pierre Fermat (1601–1665) levelet írt a híres Mersenne-nek, Descartes hű barátjának, a kor legfőbb közvetítőjének a tudományos levelezésben. Fermat megírta Mersennenek, hogy „minden prím, mely egy maradékot ad néggyel osztva, egyértelműen felírható két négyzetszám öszegeként."[1] A korban nem voltak szakfolyóiratok, ezért a matematikusok eredmlnyeiket egymással levelekben közölték. Általában csak a tételek állításait írták meg bizonyítás nélkül.

Ennek ellenére majdnem 20 évvel a Mersenne-nek írt levél után Fermat közölte a bizonyítás tervét is. 1659 augusztusában Carcavy-nak[2] írt levelében leírja, hogy bizonyítása a végtelen leszállás módszerén alapul. Felteszi, hogy az állítás nem igaz valamely $4n + 1$ alakú prímre, majd ebből levezeti, hogy ekkor valamely kisebb értékre is hamis kell legyen, és ezt folytatja egészen az 5-ig – így ellentmondáshoz érkezve, hiszen 5-re igaz az állítás.

Az első teljes bizonyítás Leonhard Eulertől (1707–1783) származik, az 1742 és 1747 közötti időből való. Euler nagyra becsülte Fermat-t, és elődjének való tiszteletadásként a fenti ötletre építette bizonyítását. Az utókor mindkét nagy

1. Ez a cikk nem foglalkozik az egyértelműséggel, melyet már jóval Fermat előtt felfedeztek. Lásd a Quantum ezen számának (1994. május/június) M115-ös feladatát. (A szerk.)

2. Mersenne halála után a tudományos közvetítői feladatot Pierre Carcavy († 1684.) vette át, aki királyi könyvtáros, amatőr matematikus és Fermat barátja volt. (A szerk.)

matematikus elismeréseképp nevezi a tételt Fermat–Euler-tételnek[3].

Majdnem minden szép matematikai eredmény (és hegycsúcs) rendelkezik egy közös tulajdonsággal: több úton is elérhető. A különböző oldalakról való megközelítés tiszta élvezetet nyújt azoknak, akik nem félnek belevágni.

A Fermat–Euler-tétel remek példája a több oldalról való bizonyíthatóságnak. A következőkben ezt mutatjuk be.

A 17. században felfedezett csúcsra három úton jutunk el. Az utak egyikét a 18. században lelték meg, a másik a 19. század terméke, a harmadik pedig a nemrég lezárult 20. századból való.

Lagrange bizonyítása

Az első bizonyítás (némi módosítással) majdnem minden számelméleti tankönyvben szerepel. A bizonyítás a következő lemmán alapul.

Wilson-lemma.[4] Minden p prímszámra $(p-1)!+1$ osztható p-vel.

Annak érdekében, hogy ne időzzünk túl sokat e csupán kiegészítő – mindazonáltal roppant hasznos – tétel bizonyításánál, a főgondolatot egy példán keresztül mutatom be: legyen $p = 13$. Keressük meg 2-től 11-ig minden egészhez azt a szorzót, amivel alkotott szorzata 1 maradékot ad 13-mal osztva, majd bontsuk fel $(13-1)!$-t.

$$(13-1)! = 12! = (2\cdot 7)\cdot(3\cdot 9)\cdot(4\cdot 10)\cdot(5\cdot 8)\cdot(6\cdot 11)\cdot 12$$

(A párok valóban 1 maradékot adnak: $2\cdot 7 = 14 = 13+1$, $3\cdot 9 = 27 = 2\cdot 13+1$, $4\cdot 10 = 5\cdot 8 = 40 = 3\cdot 13+1$, $6\cdot 11 = 66 = 5\cdot 13+1$.) Látható, hogy $12!$ maradéka 13-mal osztva 12, vagyis $12!+1$ osztható 13-mal! A lemma általános p prímre hasonló gondolatmenettel igazolható.

2. feladat. *Bizonyítsuk be, hogy bármely p prímszámra a 2, 3, …, $p-2$ egészek párosíthatók úgy, hogy a párosított a és b számok szorzata 1 legyen. (Azaz teljesüljön az $ab \equiv 1 \pmod p$ kongruencia.) Ezen állítás felhasználásával bizonyítsuk be a Wilson-lemmát.*

A Wilson-lemmából levezethetjük a következő következményt.

Következmény. Ha $p = 4n+1$ prímszám, akkor $((2n)!)^2+1$ osztható p-vel.

Bizonyítás. A bizonyításhoz $(p-1)!+1$ helyett írjunk $(4n)!+1$-et.

$$\begin{aligned}(4n)!+1 &= (1\cdot 2\cdot \ldots\cdot(2n)\cdot(2n+1)\cdot(2n+2)\cdot\ldots(4n))+1\\ &= (1\cdot 2\cdot\ldots\cdot(2n)\cdot(p-2n)\cdot(p-(2n-1))\cdot\ldots(p-1))+1\\ &\equiv \big((2n)!\cdot(-1)^{2n}\cdot(2n)!\big)+1 \pmod p\\ &\equiv ((2n)!)^2+1 \pmod p\end{aligned}$$

Mivel $(4n)!+1$ a Wilson-lemma értelmében osztható p-vel, következik, hogy a moduló p vele kongruens $((2n)!)^2+1$ is, ezzel a következményt bebizonyítottuk.

Legyen $N = (2n)!$. Ekkor a következmény azt állítja, hogy $N^2 \equiv -1 \pmod p$. Most következik a bizonyítás fő nehézsége.

Tekintsük a $\sqrt{p}$-nél nem nagyobb nemnegatív egészekből álló $(k;m)$ párokat. A k (vagy m) számok legnagyobbja nagyobb, mint $\sqrt{p}-1$, ezért a párok száma

3. Nem tévesztendő össze az Euler-tételként is ismert, szintén számelméleti Euler–Fermat-tétellel. (A ford.)

4. Itt csak segédtételként alkalmazzuk, egyébként Wilson-tétel néven ismeretes. (A ford.)

nagyobb, mint $((\sqrt{p}-1)+1)^2 = p$. A skatulyaelv[5] szerint (a moduló p maradékosztályokat tekintve skatulyáknak), legalább két olyan $(k_1; m_1)$ és $(k_2; m_2)$ pár van, amire a $k_1 + Nm_1$ és $k_2 + Nm_2$ számok p-vel osztva megegyező maradékot adnak. Ekkor $a + Nb$ (ahol $a = k_1 - k_2$, $b = m_1 - m_2$) oszható p-vel. Jegyezzük még meg, hogy $|a| < \sqrt{p}$ és $|b| < \sqrt{p}$.

Most $a^2 - N^2b^2 = (a - Nb)(a + Nb)$ osztható p-vel, és mivel $N^2 \equiv -1 \pmod{p}$, $a^2 + b^2$ is osztható p-vel – tehát $a^2 + b^2 = rp$ valamely pozitív r-re ($r \neq 0$, mert akkor a két $(k; m)$ pár azonos lenne). Már csak annyit kell megállapítani, hogy (az előbbi egyenlőtlenségek szerint) $a^2 + b^2 < 2p$, ezért $r = 1$, így $a^2 + b^2 = p$, ezzel befejezve a bizonyítást.

Zagier bizonyítása

A következő bizonyítás, melyet a kortárs D. Zagier adott, teljesen elkábított: az eredmény varázslatosan tűnik elő a semmiből.

A zagieri bizonyításban azt kell belátnunk, hogy minden $p = 4n + 1$ prímre az

$$x^2 + 4yz = p$$

egyenletnek van $(x; y; z)$ megoldása, ahol x, y és z pozitív egészek, továbbá $y = z$. Ebből következne, hogy $p = x^2 + 4y^2 = x^2 + (2y)^2$, ami igazolná az állítást.

Ezen megoldás létét egy kissé bizarr módon fogjuk bizonyítani: megmutatjuk, hogy a (nyilvánvalóan véges számú) *megoldások száma páratlan.* Hogyan kapcsolódik a páratlanság a szükséges alakú megoldásokhoz? Minden megoldás, amiben $y \neq z$, párba állítható y és z felcserélésével: ha (x, y, z) megoldás, akkor (x, z, y) is az. Ezért a nem megfelelő megoldások száma páros, tehát az összes megoldás száma csak akkor lehet páratlan, ha van olyan megoldás, amiben $y = z$.

A modern mód ezen gondolat kifejtésére a következő: tekintsük az egyenletet kielégítő hármasok S halmazán értelmezett J transzformációt, ami felcseréli y-t és z-t – $J(x, y, z) = (x, z, y)$ – és jegyezzük meg először, hogy J *involúció* – azaz kétszer egymás után alkalmazva az eredeti állapotot adja vissza; másodszor, hogy $(x, y, z) = J(x, y, z)$ *fixpontjai* p négyzetszámokra való felbontását jelentik (hiszen ekkor $y = z$); harmadszor pedig, hogy a nem fixpontok száma páros, lévén párokba rendezhetők úgy, hogy egymás képei legyenek. Természetesen ez utóbbi megállapítás *bármely* involúcióra igaz *bármely* véges halmazban.

Most tekintsük az (x, y, z) számhármasok B transzformációját, melyet a következőképpen definiálunk: $B(x, y, z) = (x', y', z')$, ahol

1. ha $x < y - z$, akkor $x' = x + 2z$, $y' = z$, $z' = y - x - z$;
2. ha $y - z \leq x \leq 2y$, akkor $x' = 2y - x$, $y' = y$, $z' = x - y + z$;
3. ha $x > 2y$, akkor $x' = x - 2y$, $y' = x - y + z$, $z' = y$.

J-hez hasonlóan S-en értelmezve ez a transzformáció is involúció S-en, a következőkben ezt mutatjuk meg. Először is B önmagára képezi le S-et. Valóban,

5. A Dirichlet-féle skatulyaelvről lásd a „Pigeons in every Pigeonhole” c. cikket a Quantum 1990 januári számában.

tekintsük például az 1. esetet:

$$\begin{aligned} x'^2 + 4y'z' &= (x+2z)^2 + 4z\cdot(y-x-z) \\ &= x^2 + 4xz + 4z^2 + 4yz - 4xz - 4z^2 \\ &= x^2 + 4yz \end{aligned}$$

A bizonyítás a másik két esetben is ugyanilyen egyszerű. Továbbá igaz, hogy ha $(x', y', z') = B(x, y, z)$, akkor $B(x', y', z') = (x, y, z)$. Ez szintén egyszerű számolással bizonyítható. Például legyen $x > y - z$, ekkor az 1. eset képleteit alkalmazzuk: $x' = x + 2z > 2z = 2y'$, ezért az $(x'', y'', z'') = B(x', y', z')$ kiszámításához a 3. eset képleteit használjuk:

$$\begin{aligned} x'' &= x' - 2y' = x + 2z - 2z = x \\ y'' &= x' - y' + z' = x + 2z - z + (y - x - z) = y \\ z'' &= y' = z \end{aligned}$$

A másik két eset vizsgálatát az Olvasóra hagyjuk. A fentiek igazolása után kijelenthetjük, hogy B involúció S-en.

Mi a helyzet B fixpontjaival? A definíció szerint az 1., illetve a 3. esetben $x' > x$ vagy $x' < x$, ezért fixpont csak a 2. esetben fordulhat elő, amikor is $x = x' = 2y - x$, azaz $x = y$.

Ennek a megfordítása is igaz: bármely $(x; x; z)$ hármast B önmagába viszi. (Ellenőrizze!) Viszont az ilyen számhármasok közül csak egy eleme S-nek: ha $p = x^2 + 4xz = x\cdot(x+4z)$, akkor $x = 1$ (hiszen p prím), és $z = n$ (emlékezzünk, hogy $p = 4n + 1$). Így az S halmaz B involúciójának egyedüli fixpontja az $(1; 1; n)$, emiatt tehát S-nek páratlan számú eleme van, amit bizonyítani kellett.

Minkowski bizonyítása

A következő, Hermann Minkowski (1864–1909) által adott bizonyítás még inkább próbára teszi a képzeletet.[6]

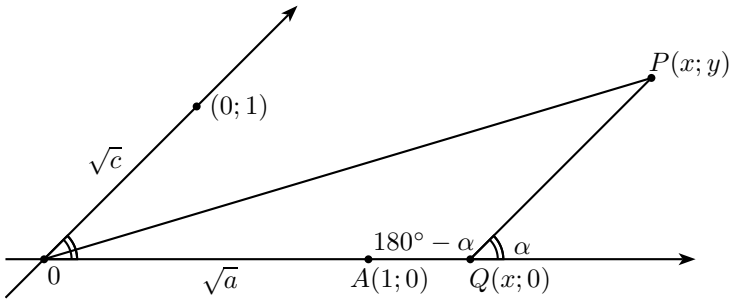
Minkowski bizonyítása egy olyan tétellel kezdődik, aminek látszólag semmi köze a Fermat–Euler-tételhez.

Tétel. Legyenek a, b és c egészek úgy, hogy $a > 0$ és $ac - b^2 = 1$ teljesüljön. Ekkor az $ax^2 + 2bxy + cy^2 = 1$ egyenletnek van egészekből álló $(x; y)$ megoldása.

Bizonyítás. Az $ax^2 + 2bxy + cy^2$ kifejezés tekinthető az O középpont és a $P(x; y)$ pont távolságnégyzetének egy megfelelő (nem feltétlenül derékszögű) koordináta-rendszerben. Ilyen koordináta-rendszert kapunk, ha a tengelyek α szöget zárnak be, ahol $\cos\alpha = \frac{b}{\sqrt{ac}}$ (ez lehetséges, mert $ac > 0$ és $\left|\frac{b}{\sqrt{ac}}\right| < 1$, lévén $ac = b^2 + 1 > b^2$). Az x- és y-tengelyeken az egységek legyenek rendre $\sqrt{a}$ és $\sqrt{c}$ (lásd az 1. ábrát). Ekkor az OP távolság négyzete az OQP háromszög felhasználásával számolható, ahol $Q(x; 0)$; ebben a háromszögben $OQ = |x|\sqrt{a}$, $QP = |y|\sqrt{c}$, és a Q-nál lévő szög α vagy $180° - \alpha$ x és y előjeleitől függően. Ugyanakkor a koszinusztétel az előjelektől függetlenül minden esetben ugyanazt adja (ellenőrizze!):

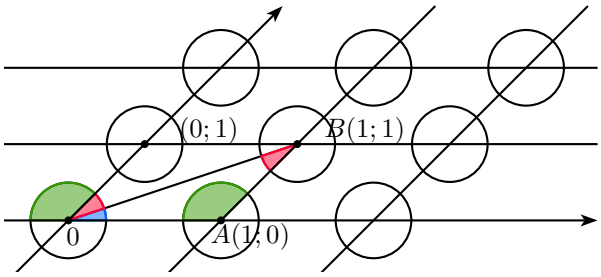
$$\begin{aligned} OP^2 &= OQ^2 - 2\cdot OQ\cdot QP\cdot\cos(Q\angle) + QP^2 \\ &= ax^2 + 2bxy + cy^2 \end{aligned}$$

6. A cikk a bizonyítást kissé módosított formában közli.



1. ábra.

Azon pontok, melyeknek mindkét koordinátája egész, rácsot alkotnak. Azt kell bizonyítani, hogy van rácspont az origótól 1 távolságra.



2. ábra.

Jelölje d az origó és a hozzá legközelebbi rácspont távolságát. Legyen $(m;n)$ rácspont O-tól d távolságra. Mivel a koordináta-rendszerben az $(x;y)$ és $(x_1;y_1)$ pontok távolsága egyenlő a $(0;0)$ és az $(x_1-x;y_1-y)$ pontok távolságával, bármely két rácspont között a távolság legalább d. Ebből következően a $\frac{d}{2}$ sugarú, rácspont középpontú körlapok nem fedik egymást: ha van két, A, illetve B középpontú kör, ami fedi egymást, akkor van egy közös C belső pontjuk. Erre a C pontra $AB < AC + CB < \frac{d}{2} + \frac{d}{2} = d$, ami ellentmondás.[7] A 2. ábrán látható, hogy az $O(0;0)$, $A(1;0)$, $B(1;1)$ pontok alkotta háromszög és a körök metszete a kör területének fele, tehát $\frac{d^2}{8}\pi$.[8] És ez mindössze egy része a háromszög terü-

7. Az, hogy két kör érintse egymást, nem kizárt, a bizonyítás menetében azonban ennek nincs szerepe. (A ford.)

8. Előfordulhat, hogy az ábrán a $(0;1)$, illetve az $(1;0)$ középpontú körök átnyúlnak a háromszög OB oldalán. Ez azonban nem módosítja a felírt kifejezés igazságát, mivel a két körnek nincs közös belső pontja, és a középpontos szimmetria miatt „kilógó” és „belógó” körszeletek egybevágók. (A ford.)

letének, ami a következő:

$$\begin{aligned}\frac{1}{2}\cdot OA\cdot AB\cdot\sin(A\angle) &= \frac{1}{2}\cdot\sqrt{a}\cdot\sqrt{c}\cdot\sin(A\angle)\\ &= \sqrt{ac\cdot(1-\cos^2)}\\ &= \frac{1}{2}\sqrt{ac-b^2}\\ &= \frac{1}{2}.\end{aligned}$$

Tehát $\frac{d^2}{8}\pi < \frac{1}{2}$, vagyis $d^2 < \frac{4}{\pi} < 2$. Mivel d^2 egész szám (hiszen $d^2 = am^2 + 2bmn + cn^2$), $d = 1$, ami igazolja Minkowski tételét.

De mi köze van ennek a bámulatos eredménynek a Fermat–Euler-tételhez? A lehető legtöbb!

A Wilson-lemma korábban bizonyított következménye szerint $N^2 + 1$ osztható p-vel, ahol $N = \left(\frac{p-1}{2}\right)!$. Most alkalmazzuk Minkowski tételét az $a = p$, $b = N$, $c = \frac{b^2+1}{a}$ értékekre ($a > 0$ és $ac - b^2 = 1$ valóban teljesül, továbbá a Wilson-lemma értelmében c is egész szám). A tétel szerint bizonyos m és n egészekre

$$1 = am^2 + 2bmn + cn^2,$$

mindkét oldalt a-val szorozva pedig

$$\begin{aligned}p = a &= a^2m^2 + 2abmn + \left(b^2+1\right)n^2\\ &= (am+bn)^2 + n^2,\end{aligned}$$

tehát p két négyzetszám összege. És a tétel ismét be van bizonyítva!

Megoldások

1. feladat. Ha $a^2 + b^2$ páratlan prímszám, akkor a két négyzetszám egyike – mondjuk a^2 – páros. Tehát $a = 2k$, míg b páratlan: $b = 2l+1$. Ekkor $a^2+b^2 = 4k^2 + 4l^2 + 4l + 1 = 4n + 1$, ahol $n = k^2 + l^2 + l$.

2. feladat. Valamely a egészre, $1 \le a \le p-1$, tekintsük az $1\cdot a, 2\cdot a, \ldots, (p-1)\cdot a$ szorzatokat. Ezen szorzatok p-vel osztva valamennyien különböző maradékot adnak (hiszen ha $ka \equiv la \pmod p$, $k > l$, akkor $(k-l)a$ osztható p-val, hiszen $k-l$ és a is pozitívok és p-nél kisebbek, és p prím). Mivel egyik szorzat sem osztható p-vel, a maradékok 1-től $p-1$-ig minden értéket felvesznek (redukált maradékrendszert alkotnak). Ezért minden a-hoz van olyan b ($1 \le b \le p-1$), amire $ab \equiv 1 \pmod p$. Az $a = 1$ és $a = p-1$ értékekre $a = b$. Minden más esetben ($2 \le a \le p-2$) $a \ne b$, mert $a^2 \equiv 1 \pmod p$-ből következik, hogy $a^2 - 1 = (a+1)(a-1)$ osztható p-vel, ez csak az $a - 1 = 0$ és az $a + 1 = p$ esetekben teljesül. Ez bizonyítja az első állítást. A Wilson-lemma bizonyítása ennél annyival több, hogy a $2 \cdot \ldots \cdot (p-2)$ szorzat tényezőit párosítjuk úgy, hogy a párok szorzata 1 maradékot adjon p-vel osztva. Ezt $p-1$-gyel (és 1-gyel) szorozva kapjuk $(p-1)!$-t, ami ezért $p-1$-gyel, azaz -1-gyel kongruens.