Reszta jest dziełem człowieka, czyli Fermat i inni

Mariusz SKAŁBA[*]

[*]Instytut Matematyki, Wydział Matematyki, Informatyki i Mechaniki, Uniwersytet Warszawski

Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk.

Leopold Kronecker

Nie ma słynniejszego twierdzenia niż Wielkie Twierdzenie Fermata (WTwF) i tego nie zamierzam tu dowodzić. Zacznę po prostu od sformułowania faktu, który od 1995 roku jest rzeczywiście twierdzeniem za sprawą Andrew Wilesa, a wcześniej przez około trzy i pół wieku był hipotezą zajmującą głowy największych matematyków i rzesze amatorów.

Twierdzenie 1. *Jeżeli n jest liczbą naturalną większą od* 2, *to równanie*

$$x^n + y^n = z^n \tag{1}$$

nie ma rozwiązań w liczbach naturalnych x, y, z.

Jak powszechnie wiadomo, wykładnik $n = 2$ jest zupełnie wyjątkowy, gdyż wtedy równanie (1), zwane równaniem Pitagorasa, ma nieskończenie wiele nieproporcjonalnych rozwiązań, a nawet dużo więcej – ma rozwiązanie w wielomianach o współczynnikach całkowitych:

$$x = k(m^2 - n^2), \quad y = k \cdot 2mn, \quad z = k(m^2 + n^2).$$

Pokażemy przede wszystkim, że dla ustalonego $n > 2$ równanie (1) nie ma rozwiązań w wielomianach $x(t), y(t), z(t)$ o współczynnikach zespolonych, chyba że wszystkie wymienione wielomiany są stałe! Załóżmy nie wprost, że trójka wielomianów $x(t), y(t), z(t)$ jest hipotetycznym rozwiązaniem równania (1), przy czym nie wszystkie wielomiany $x(t), y(t), z(t)$ są stałe oraz (co możemy założyć) $x(t), y(t), z(t)$ nie mają wspólnego dzielnika wielomianowego dodatniego stopnia. Jeśli $\omega_n = \exp(2\pi i/n)$, to mamy tożsamość algebraiczną

$$\alpha^n - 1 = (\alpha - 1)(\alpha - \omega_n)(\alpha - \omega_n^2) \cdot \ldots \cdot (\alpha - \omega_n^{n-1}),$$

z której po podstawieniu $\alpha = z/y$ i pomnożeniu obu stron przez y^n otrzymujemy

$$z^n - y^n = (z - y)(z - \omega_n y)(z - \omega_n^2 y) \cdot \ldots \cdot (z - \omega_n^{n-1} y).$$

Nasze hipotetyczne wielomiany spełniają więc równanie

$$x(t)^n = (z(t) - y(t))(z(t) - \omega_n y(t))(z(t) - \omega_n^2 y(t)) \cdot \ldots \cdot (z(t) - \omega_n^{n-1} y(t)).$$

Ponieważ czynniki po prawej stronie są parami względnie pierwsze, więc z jednoznaczności rozkładu wielomianów na czynniki wynika, że istnieją takie wielomiany $u(t), v(t), w(t)$, że

$$z(t) - y(t) = u(t)^n, \quad z(t) - \omega_n y(t) = v(t)^n, \quad z(t) - \omega_n^2 y(t) = w(t)^n.$$

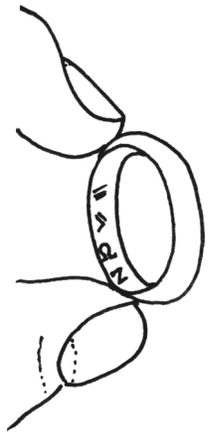
Po rozwiązaniu pierwszych dwóch z powyższych równań względem $z(t)$ oraz $y(t)$ i podstawieniu tych wartości do trzeciego równania otrzymujemy po uproszczeniach

$$-\omega_n u(t)^n + (\omega_n + 1)v(t)^n = w(t)^n.$$

Ponieważ z liczb zespolonych można wyciągać pierwiastki dowolnych stopni, więc powyższą równość można zapisać jako

$$x_1(t)^n + y_1(t)^n = z_1(t)^n,$$

gdzie największy ze stopni wielomianów $x_1(t), y_1(t), z_1(t)$ jest n razy mniejszy niż największy ze stopni wielomianów $x(t), y(t), z(t)$. To postępowanie można kontynuować, ale to oczywista sprzeczność! Przeprowadzone rozumowanie ilustruje słynną *metodę regresji*: za pomocą hipotetycznego rozwiązania konstruujemy rozwiązanie w pewnym sensie mniejsze i to prowadzi do sprzeczności.
W rozważonym przykładzie mielibyśmy nieskończony ciąg wielomianów o ściśle malejących dodatnich stopniach! Metodę tę stosowano z pewnym powodzeniem również do przypadku liczbowego, chociaż pojawiają się tu już dość szybko fundamentalne trudności i, jak pokazała historia, nie udało się ich do końca przezwyciężyć.



Aby udowodnić WTwF, wystarczy rozważać wykładniki pierwsze nieparzyste i $n = 4$.

Wielkie sukcesy uzyskał jednak Ernst Eduard Kummer w połowie XIX wieku. Rozważał on mianowicie pierścień $\mathbb{Z}[\omega_p]$, gdzie $n = p$ jest rozważanym wykładnikiem i zakłada się, że p jest liczbą pierwszą nieparzystą. Do pierścienia $\mathbb{Z}[\omega_p]$ należą liczby zespolone postaci $a_0 + a_1\omega_p + a_2\omega_p^2 + \cdots + a_{p-2}\omega_p^{p-2}$, gdzie $a_0, a_1, \ldots, a_{p-2}$ to zwykłe liczby całkowite. W przypadku, gdy w pierścieniu $\mathbb{Z}[\omega_p]$

Rozwiązanie zadania M 1480.
Zauważmy, że dla całkowitych x mamy równość. Ponadto dodanie do dowolnego x liczby całkowitej k zmienia każdą ze stron nierówności o nk, możemy zatem zakładać, że x należy do przedziału $(0,1)$. Niech $t_{k,l} = k/l$ dla takich względnie pierwszych liczb całkowitych dodatnich k i l, że $k < l \leqslant n$. Zauważmy, że obie strony nierówności są w przedziale $(0,1)$ niemalejącymi funkcjami x, przy czym prawa strona zmienia wartość tylko w punktach $t_{k,l}$. Wystarczy więc sprawdzić nierówność tylko dla tych punktów. Niech od tej pory $x = t_{k,l}$.

Dla każdego $i = 1, 2, \ldots, n$ zachodzi równość $ik = q_i l + r_i$ dla pewnych jednoznacznie wyznaczonych liczb całkowitych $0 \leqslant r_i \leqslant l-1$ i $q_i \geqslant 0$ (dzielimy ik z resztą przez l). Zauważmy, że liczby $r_1, \ldots, r_{l-1}$ są dodatnie: $r_i \neq 0$ dla $i < l$, bo inaczej l dzieliłoby ik. Ponadto te liczby są parami różne – gdy $r_i = r_j$ dla $i \leqslant j < l$, to l dzieli $(i-j)k$, a stąd $i = j$. W takim razie ciąg $r_1, \ldots, r_{l-1}$ jest pewną permutacją liczb $1, 2, \ldots, l-1$. Stąd i z nierówności między średnimi dostajemy

$$\frac{r_1}{1} + \ldots + \frac{r_{l-1}}{l-1} \geqslant l - 1.$$

Zatem

$$r_n \leqslant l - 1 \leqslant \sum_{i=1}^{n} \frac{r_i}{i} = nk - l\sum_{i=1}^{n} \frac{q_i}{i} =$$
$$= q_n l + r_n - l\sum_{i=1}^{n} \frac{q_i}{i},$$

a stąd

$$\sum_{i=1}^{n} \frac{\lfloor ix \rfloor}{i} = \sum_{i=1}^{n} \frac{\lfloor ik/l \rfloor}{i} = \sum_{i=1}^{n} \frac{q_i}{i} \leqslant$$
$$\leqslant q_n = \lfloor nk/l \rfloor = \lfloor nx \rfloor.$$

rozkład na czynniki nierozkładalne jest jednoznaczny, nierozwiązalność równania (1) można uzyskać w podobny sposób jak wyżej dla wielomianów. Pewne komplikacje związane ze skutecznym przeprowadzeniem regresji związane są z istnieniem w pierścieniu $\mathbb{Z}[\omega_p]$ nietrywialnych elementów *odwracalnych*, tzn. takich elementów α, że $\alpha \cdot \beta = 1$ dla pewnego $\beta \in \mathbb{Z}[\omega_p]$. Za trywialne uznajemy elementy postaci $\pm\omega_p^k$; oczywiście są one odwracalne. Dla $p \geqslant 5$ istnieją również elementy odwracalne γ, które nie są pierwiastkami z jedności. Na przykład dla $p = 7$ przyjmijmy $\gamma = 1 + \omega_7$. Mamy wówczas

$$(1+\omega_7)^{-1} = \frac{1-\omega_7}{1-\omega_7^2} = \frac{1-\omega_7^8}{1-\omega_7^2} = 1 + \omega_7^2 + \omega_7^4 + \omega_7^6 \in \mathbb{Z}[\omega_7],$$

a więc γ jest odwracalny i, co łatwo wykazać, nietrywialny. Niestety, dla $p > 19$ w pierścieniu $\mathbb{Z}[\omega_p]$ nie ma jednoznaczności rozkładu na elementy nierozkładalne, a oto przykład dla $p = 23$. Niech $\omega = \omega_{23}$. Mamy

$$\alpha = (1+\omega^2+\omega^4+\omega^5+\omega^6+\omega^{10}+\omega^{11})(1+\omega+\omega^5+\omega^6+\omega^7+\omega^9+\omega^{11}) =$$
$$= 2(\omega^5+\omega^6+\omega^7+\omega^9+\omega^{10}+3\omega^{11}+\omega^{12}+\omega^{13}+\omega^{15}+\omega^{16}+\omega^{17}).$$

Zatem liczba α dzieli się przez 2 mimo tego, że 2 nie dzieli żadnego z czynników poprzedniego iloczynu. Ponadto można wykazać, że liczba 2 jest nierozkładalna w $\mathbb{Z}[\omega_{23}]$, a więc w $\mathbb{Z}[\omega_{23}]$ nie ma jednoznaczności rozkładu na czynniki nierozkładalne! Brak jednoznaczności rozkładu na czynniki to największa trudność, którą należy pokonać, adaptując metodę regresji. Genialny pomysł Kummera polegał na wprowadzeniu do rozważań nowych obiektów, tzw. *liczb idealnych* i wykazaniu, że każda liczba z $\mathbb{Z}[\omega_p]$ rozkłada się w jednoznaczny sposób na iloczyn liczb idealnych. Używając współczesnego języka (wprowadzonego przez Dedekinda), zbiory liczb z $\mathbb{Z}[\omega_p]$ podzielnych przez daną liczbę idealną Kummera to po prostu ideały pierwsze pierścienia $\mathbb{Z}[\omega_p]$ – stąd zresztą ich nazwa! Miarą niejednoznaczności rozkładu na poziomie liczb jest tzw. *grupa klas ideałów* pierścienia $\mathbb{Z}[\omega_p]$, której definicji tu nie podamy. Nadmienimy tylko, że Kummer udowodnił, między innymi, następujące twierdzenie.

Twierdzenie 2. *Jeżeli $p > 2$ jest liczbą pierwszą oraz rząd grupy klas ideałów nie dzieli się przez p, to równanie* (1) *nie ma rozwiązań dla wykładnika $n = p$.*

Metody wypracowane przez Kummera, Dedekinda i innych dały początek *algebraicznej teorii liczb*, ważnemu działowi matematyki współczesnej. Grupa klas ideałów i grupa elementów odwracalnych są bardzo blisko spokrewnione z funktorami K_0 oraz K_1 *algebraicznej K-teorii* – nowoczesnego działu współczesnej matematyki, który próbuje łączyć algebrę i geometrię na wysokim i abstrakcyjnym poziomie.

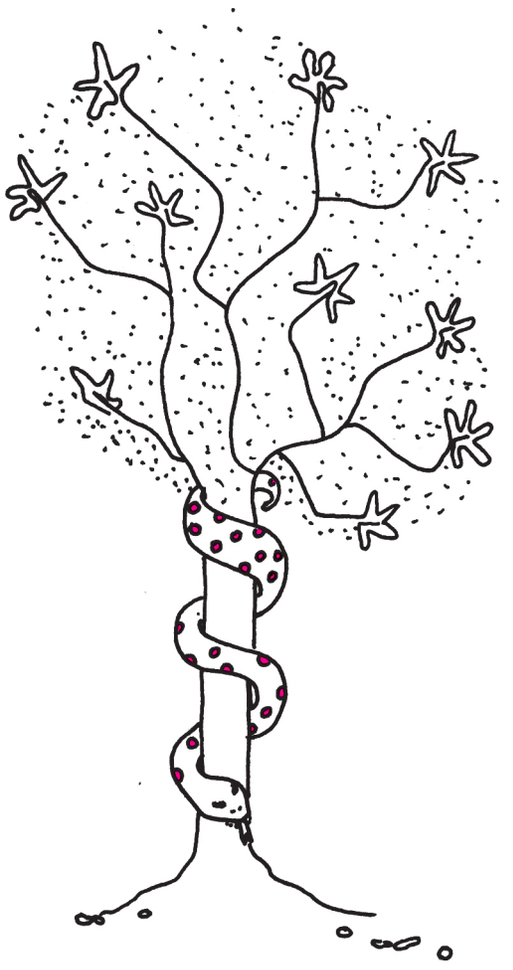
Trzeba zmierzać do końca tej historii i dlatego muszę przemilczeć wiele ciekawych częściowych rezultatów uzyskanych przed rokiem 1983. W tymże roku Gerd Faltings udowodnił hipotezę Mordella o tym, że na każdej krzywej rodzaju większego od 1 istnieje co najwyżej skończenie wiele punktów wymiernych. Wynika stąd, że dla $n \geqslant 4$ równanie (1) ma co najwyżej skończenie wiele *nieproporcjonalnych* rozwiązań.

To oczywiście wspaniały wynik, ale przecież nie tego oczekiwał Fermat. Na szczęście historia nabrała tempa, gdy w 1986 roku Gerhard Frey związał z hipotetycznym rozwiązaniem równania (1) $x = a$, $y = b$, $z = c$ następującą krzywą eliptyczną

$$y^2 = x(x + a^p)(x - b^p).$$

Wyróżnik tej krzywej wynosi $(abc)^{2p}$ i to nasunęło Freyowi przypuszczenie, że taka krzywa... w ogóle nie powinna istnieć – to oczywiście zakończyłoby dowód WTwF! Pomysł okazał się trafiony, ale szczegóły dopinało wielu wybitnych matematyków. I tak Jean-Pierre Serre sformułował precyzyjnie przypuszczenie Freya, a Ken Ribet wyprowadził je w tym samym roku ze słynnej hipotezy Taniyamy–Shimury–Weila o tym, że każda krzywa eliptyczna o współczynnikach wymiernych jest *modularna*. Tę właśnie hipotezę udowodnił Andrew Wiles w 1995 roku z pomocą swojego ucznia Richarda Taylora dla tzw. krzywych półstabilnych – klasa tych krzywych obejmuje krzywe typu Freya i to... kończy dowód WTwF. Artykuł zakończymy omówieniem pojęcia *krzywa eliptyczna modularna*. Jest wiele sformułowań tej własności – wybieramy wersję arytmetyczną. Z każdą krzywą eliptyczną

$$y^2 = x^3 + Ax^2 + Bx + C, \qquad \text{gdzie } A, B, C \in \mathbb{Z}$$



i liczbą pierwszą p można związać kongruencję

$$y^2 \equiv x^3 + Ax^2 + Bx + C \pmod{p}.$$

Niech N_p oznacza liczbę rozwiązań tej kongruencji modulo p.
Helmut Hasse udowodnił w 1933 roku hipotezę Artina głoszącą, że

$$|N_p - p| < 2\sqrt{p}.$$

Wynika z niej natychmiast, że rozwiązania kongruencji istnieją dla dostatecznie dużych p. Jest to satysfakcjonujący rezultat ilościowy, ale być może o liczbie $a_p := p - N_p$ można powiedzieć coś więcej, niż tylko to, że jest mniejsza od $2\sqrt{p}$? Okazuje się, że tak! Przyjrzyjmy się najpierw krzywej

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

Bardzo łatwo wykazać, że jeśli $p \equiv 3 \pmod 4$, to $a_p = 0$. Dużo trudniej zauważyć, że dla $p \equiv 1 \pmod 4$ też istnieje dość zwarty i dość jednolity wzór na liczbę a_p. Mianowicie liczbę p przedstawiamy w postaci $p = a^2 + b^2$, gdzie a, b są dodatnie, przy czym a jest nieparzysta (b zaś parzysta). Jak wiadomo, takie przedstawienie liczby pierwszej $p \equiv 1 \pmod 4$ zawsze istnieje i jest dokładnie jedno – jest to twierdzenie Fermata (ani małe, ani wielkie, ale wspaniałe). Wówczas a_p dane jest następującym wzorem

$$a_p = \begin{cases} 2a & \text{gdy } a \equiv 1 \pmod 4 \\ -2a & \text{gdy } a \equiv 3 \pmod 4. \end{cases}$$

Krzywa eliptyczna $y^2 = x^3 + x$ należy do dość wąskiej klasy krzywych eliptycznych z *mnożeniem zespolonym* – odwzorowanie $(x, y) \mapsto (-x, yi)$ przeprowadza punkty krzywej na punkty krzywej. Niewiele krzywych ma tego typu algebraiczne „symetrie".

Rozpatrzmy teraz typową krzywą eliptyczną (bez mnożenia zespolonego)

$$y^2 = x^3 - 4x^2 + 16$$

oraz, jak wyżej, odpowiednie kongruencje modulo różne liczby pierwsze p. Okazuje się, że ciąg liczb a_p można uzyskać w następujący sposób. Rozważmy iloczyn funkcyjny

$$\Theta(T) = T \prod_{j=1}^{\infty} \left((1 - T^j)(1 - T^{11j})\right)^2$$

przy czym dla uproszczenia pomińmy kwestię zbieżności. Jeśli zapiszemy teraz powyższy iloczyn formalny jako szereg formalny

$$\Theta(T) = \sum_{k=1}^{\infty} c_k T^k,$$

to okazuje się, że dla każdej liczby pierwszej $p \geqslant 3$ zachodzi równość $a_p = c_p$, a więc znowu otrzymaliśmy „wzór" na a_p. Hipoteza Taniyamy–Shimury–Weila przewidywała właśnie, że tego typu „wzór" na a_p można podać dla każdej krzywej eliptycznej o współczynnikach wymiernych. Znaczenie twierdzenia o modularności znacznie wykracza poza zastosowanie do dowodu WTwF. Daje ono, na przykład, fundament do ścisłego sformułowania hipotezy Bircha–Swinnertona-Dyera, która nadal jest jednym z problemów milenijnych. Do tych zastosowań trzeba przywołać sformułowania bardziej geometryczno-analityczne od stosowanych w tym tekście.

Polecam w tej sprawie znakomity artykuł Zbigniewa Marciniaka *O Wielkim Twierdzeniu Fermata*, *Matematyka-Społeczeństwo-Nauczanie* 22, `www.msn.uph.edu.pl/smp/msn/22/10-15.pdf`.

Na koniec odnieśmy się do słynnej myśli wypowiedzianej przez Leopolda Kroneckera, że liczby całkowite stworzył Bóg, a *reszta jest dziełem człowieka*. Historia WTwF doskonale ilustruje starą prawdę, że jeśli chcemy się czegoś dowiedzieć (pozornie prostego!) o bardzo prostych obiektach, musimy wyjść ze świata tychże obiektów i mieć nadzieję, że wzbogacona w ten sposób perspektywa pozwoli nam dostrzec te powiązania, które nie były widoczne z bliska. Do lat osiemdziesiątych XX wieku to wyjście z raju liczb naturalnych było bardzo ograniczone – cały czas obracano się w świecie liczb algebraicznych i ich rozkładów na czynniki. Pomysł Freya wyrwał badaczy WTwF z zaklętego kręgu liczb, a świat krzywych eliptycznych i form modularnych okazał się wystarczającym wzbogaceniem ubogiego i jakże błahego kontekstu WTwF.