

Kilka dowodów twierdzenia Fermata o sumie dwóch kwadratów

* Studentka, University of Cambridge

Maryna SPEKTROVA *

Zacznijmy od zaprezentowania dobrze znanego nam twierdzenia – bohatera tego artykułu.

Twierdzenie (o sumie dwóch kwadratów, Fermat). *Nieparzysta liczba pierwsza p może być przedstawiona jako suma dwóch kwadratów wtedy i tylko wtedy, gdy $p \equiv 1 \pmod{4}$.*

Zapis $n \equiv l \pmod{k}$ oznacza, że $n - l$ jest podzielne przez k . Innymi słowy, n i l dają tę samą resztę z dzielenia przez k .

Jednym z najbardziej interesujących aspektów tego twierdzenia jest duża liczba jego różnych dowodów, które łączą się z różnymi dziedzinami matematyki. W szczególności można je udowodnić za pomocą liczb Gaussa, skończonych i nieskończonych ciągów ułamków, lematu Thuego, metody nieskończonego schodzenia itd. W artykule zaprezentujemy trzy dowody: dwie interpretacje geometryczne oraz jeden dowód oparty na teorii aproksymacji.

Uwaga. Ponieważ dla każdej liczby całkowitej x zachodzi $x^2 \equiv 0, 1 \pmod{4}$, żadna liczba dająca resztę 3 z dzielenia przez 4 nie może być przedstawiona jako suma dwóch kwadratów, dlatego dalej będziemy dowodzić tylko jednej implikacji.

Pierwszy dowód będzie oparty na lemacie dotyczącym krat równoległobocznych. Nie będziemy definiować tego pojęcia formalnie, zamiast tego pokażemy przykład takiej kraty na marginesie.

Lemat (Minkowski). *Rozważmy kratę równoległoboczną oraz figurę wypukłą Φ , która jest symetryczna względem początku układu współrzędnych. Przypuśćmy, że $S(\Phi) > 4S_0$, gdzie S_0 jest polem podstawowego równoległoboku, a $S(\Phi)$ polem figury Φ . Wówczas w Φ leży pewien punkt tej kraty, różny od początku układu współrzędnych.*

Rozważmy kratę równoległoboczną daną wektorami $\vec{u} = (1, m)$, $\vec{v} = (0, p)$, gdzie m jest liczbą spełniającą $p \mid m^2 + 1$ (dowód istnienia takiej liczby dla $p = 4k + 1$ znajduje się na marginesie).

Zastosujmy teraz lemat Minkowskiego do koła o środku w początku układu współrzędnych i promieniu $\sqrt{2p}$. Ponieważ $S(\Phi) = 2\pi p > 4p = 4S_0$, w tym kole istnieje punkt kratowy A , czyli punkt postaci $x_0 \cdot \vec{u} + y_0 \cdot \vec{v} = (x_0, mx_0 + py_0)$ dla pewnych całkowitych x_0, y_0 . Kwadrat odległości punktu A od początku układu współrzędnych wynosi

$$x_0^2 + (mx_0 + py_0)^2 = x_0^2 + m^2 x_0^2 + 2mx_0 py_0 + p^2 y_0^2 = x_0^2(m^2 + 1) + p(2mx_0 y_0 + py_0^2),$$

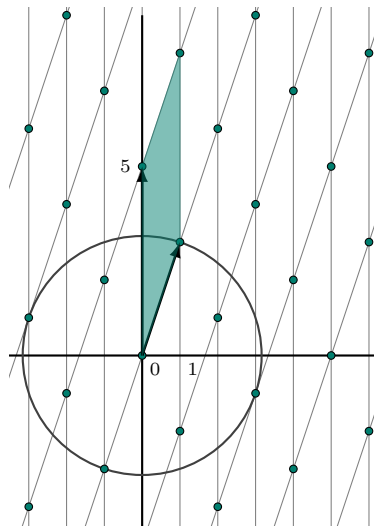
co jest podzielne przez p . Ponieważ A znajduje się wewnątrz koła o promieniu $\sqrt{2p}$, liczba ta mieści się w przedziale otwartym od 0 do $2p$, więc jest równa p . Dlatego poszukiwane dwa kwadraty to x_0^2 oraz $(mx_0 + py_0)^2$. $\square$

Drugi dowód jest nazywany „jednozdanowym dowodem Dona Zagiera” i rzeczywiście, gdy w 1990 roku został po raz pierwszy opublikowany, zawierał tylko jedno zdanie. Dla poprawienia czytelności przedstawiamy go tutaj nieco bardziej szczegółowo.

Rozważmy zbiór S zawierający wszystkie trójki (x, y, z) liczb naturalnych, które spełniają $x^2 + 4yz = p$. Oczywiście jest, że dla każdego p zbiór S jest skończony. Rozważmy następującą funkcję $f: \mathbb{N}^3 \rightarrow \mathbb{N}^3$:

$$f(x, y, z) = \begin{cases} (x + 2z, z, y - x - z), & \text{gdy } x < y - z, \\ (2y - x, y, x - y + z), & \text{gdy } y - z < x < 2y, \\ (x - 2y, x - y + z, y), & \text{gdy } 2y < x. \end{cases}$$

Trójki pojawiające się w poszczególnych przypadkach funkcji f nazwiemy odpowiednio trójkami pierwszego, drugiego i trzeciego rodzaju.



Przykład kraty danej przez wektory $(0, 5)$ i $(1, 3)$.

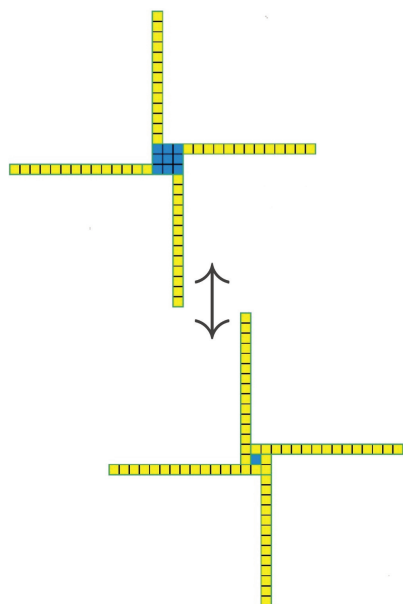
Wskażemy liczbę naturalną m , spełniającą $p \mid m^2 + 1$. Użyjemy twierdzenia Wilsona, które mówi, że $(p-1)! \equiv -1 \pmod{p}$ dla każdej liczby pierwszej p .

Z tego twierdzenia wynika, że $-1 \equiv (p-1)!$

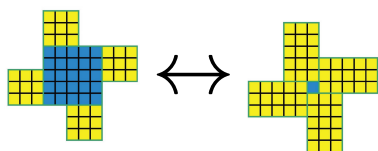
$$\begin{aligned} &= 1 \cdot 2 \cdot \dots \cdot \left(\frac{p-1}{2}\right) \cdot \left(\frac{p+1}{2}\right) \cdot \dots \cdot (p-1) \\ &\equiv 1 \cdot 2 \cdot \dots \cdot \left(\frac{p-1}{2}\right) \cdot \left(-\frac{p-1}{2}\right) \cdot \dots \cdot (-1) \\ &= \left(\left(\frac{p-1}{2}\right)!\right)^2 \cdot (-1)^{\frac{p-1}{2}} \\ &= \left(\left(\frac{p-1}{2}\right)!\right)^2 \pmod{p}. \end{aligned}$$

Ostatni krok jest prawdziwy, ponieważ $p \equiv 1 \pmod{4}$. Stąd $p \mid m^2 + 1$ dla $m = \left(\frac{p-1}{2}\right)!$.

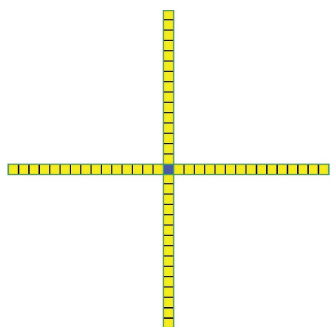
O dowodzie Zagiera pisał już Wojciech Czerwiński w Δ_{17} .



Przykład inwolucji dla $p = 61$
 $((3, 1, 13) \leftrightarrow (1, 15, 1))$



Przykład inwolucji dla $p = 61$
 $((5, 3, 3) \leftrightarrow (1, 3, 5))$



Jedynym punktem stałym f dla $p = 61$ jest $(1, 1, 15)$. Zobacz, jak ta konfiguracja różni się od $(1, 15, 1)$

O twierdzeniu Dirichleta o aproksymacji pisał Wojciech Czerwiński, w *Delta* 3/21.

Wykonując proste obliczenia, możemy łatwo sprawdzić, że:

- $f(S) \subseteq S$;
- f jest bijekcją, a nawet inwolucją, czyli $f(f(x, y, z)) = (x, y, z)$;
- Trójki pierwszego, drugiego i trzeciego rodzaju przechodzą odpowiednio na trójki trzeciego, drugiego i pierwszego rodzaju.

Teraz zbadajmy, które trójki $t \in S$ są punktami stałymi, tj. spełniają $f(t) = t$. Trójka może być punktem stałym tylko, jeśli jest drugiego rodzaju, co prowadzi do następującego układu równań:

$$\begin{cases} 2y - x = x, \\ y = y, \\ x - y + z = z, \end{cases}$$

który upraszcza się do $x = y$. Wiemy, że $x^2 + 4yz = p$, więc $x^2 + 4xz = p$. Ponieważ p jest liczbą pierwszą, łatwo otrzymujemy $x = y = 1, z = \frac{p-1}{4}$, zatem w S istnieje dokładnie jeden punkt stały, co oznacza, że S ma nieparzystą liczbę elementów.

Zauważmy, że elementy S możemy dobrać w pary w taki sposób, że trójka (x, y, z) jest w parze z (x, z, y) . Ponieważ S ma nieparzystą liczbę elementów, więc co najmniej jedna trójka jest sparowana sama ze sobą. Dla tej trójki (x_0, y_0, z_0) mamy $y_0 = z_0$, więc $p = x_0^2 + 4y_0^2 = x_0^2 + (2y_0)^2$. $\square$

Bardzo interesująca jest geometryczna interpretacja funkcji f – co zaskakujące, opublikowana dopiero w 2007 roku. Dla każdej trójki z S możemy rozważyć kwadrat o boku x i dobudować do niego 4 prostokąty o bokach y i z , jak na rysunku. Każda taka figura („wiatrak” lub „kwadrat ze skrzydłami”) może być uzyskana w wyniku dwóch różnych cięć: jednego z mniejszym kwadratem i drugiego z większym. W ten sposób otrzymujemy inwolucję między wszystkimi trójkami w S (w rzeczywistości ta bijekcja to właśnie f), której punktami stałymi są krzyże. Ponieważ dla figury w kształcie krzyża p musi być podzielne przez długość boku głównego kwadratu, więc bok ten musi mieć długość 1. Istnieje zatem dokładnie jeden punkt stały. Teraz to samo rozumowanie co powyżej kończy dowód.

Trzeci dowód opiera się na następującym twierdzeniu. Intuicyjnie określa ono, jak dobrze możemy przybliżyć daną liczbę rzeczywistą liczbami wymiernymi o mianowniku nieprzekraczającym zadanej liczby całkowitej N .

Twierdzenie (o aproksymacji, Dirichlet). *Dla każdych $\alpha \in \mathbb{R}$ oraz $N \in \mathbb{N}$ istnieją $r, q \in \mathbb{Z}$ takie, że:*

- $1 \leq q \leq N$ oraz
- $|\alpha - \frac{r}{q}| \leq \frac{1}{Nq}$.

Teraz rozważmy takie m , że $p \mid m^2 + 1$, i zastosujmy twierdzenie Dirichleta o aproksymacji dla $\alpha = \frac{m}{p}$ oraz $N = \lceil \sqrt{p} \rceil$. Otrzymujemy, że istnieją $q \in \mathbb{N}, r \in \mathbb{Z}$ takie, że $1 \leq q \leq N$ oraz $|\frac{m}{p} - \frac{r}{q}| \leq \frac{1}{Nq}$, czyli $|mq - rp| \leq \frac{p}{N} < \sqrt{p}$.

Zdefiniujmy M jako $(mq - rp)^2 + q^2$. Łatwo sprawdzić, że $M \equiv m^2 q^2 + q^2 \equiv 0 \pmod{p}$.

- Przypadek 1: $q \neq N$
Ponieważ zarówno $(mq - rp)^2$, jak i q^2 są mniejsze od p , więc $0 < M < 2p$. Ale M jest podzielne przez p , więc $M = p$, a zatem $p = (mq - rp)^2 + q^2$.
- Przypadek 2: $q = N$
Łatwo jest wykazać, że $M < 3p$, więc $M \in \{p, 2p\}$. Jeśli $M = p$, to twierdzenie zostało już udowodnione, więc rozważmy przypadek $M = 2p$. Wtedy $(mq - rp)^2 > 2p - (\sqrt{p} + 1)^2 = (\sqrt{p} - 1)^2 - 2$, więc $|mq - rp| \in \{\lfloor \sqrt{p} \rfloor - 1, \lfloor \sqrt{p} \rfloor\} = \{q - 2, q - 1\}$. Ponadto w tym przypadku $|mq - rp|$ oraz q mają tę samą parzystość, więc $|mq - rp| = q - 2$, zatem $(q - 2)^2 + q^2 = 2p$, czyli $(q - 1)^2 + 1 = p$, co oznacza, że przedstawiliśmy p jako sumę dwóch kwadratów. $\square$

Kiedy spojrzymy na pierwszy i trzeci dowód, możemy zauważyć kilka podobieństw. Czy to tylko przypadek? Okazuje się, że twierdzenie Dirichleta można wywnioskować z twierdzenia Minkowskiego. W tym celu rozważmy zbiór

$$S = \left\{ (x, y) \in \mathbb{R}^2 : |x| < N + \frac{1}{2}, |y - \alpha x| \leq \frac{1}{N} \right\}.$$

Zbiór S jest równoległobokiem o środku w początku układu współrzędnych i polu równym $2(N + \frac{1}{2}) \cdot \frac{2}{N} > 4$, więc możemy zastosować twierdzenie Minkowskiego dla S na zwykłej kracie kartezjańskiej. Dostajemy, że S zawiera punkt całkowity (q, r) (z symetrii możemy wybrać q dodatnie), a te q, r spełniają warunek z twierdzenia Dirichleta.

Twierdzenie Jacobiego o sumie dwóch kwadratów. Nie tylko twierdzenie Fermata ma kilka zaskakujących dowodów. To samo jest prawdą również dla jednego z jego uogólnień. Następujące twierdzenie można udowodnić, używając form kwadratowych, funkcji eliptycznych lub liczb Gaussa, ale przedstawimy tutaj jedynie dowód przy użyciu funkcji tworzących.

Twierdzenie (o sumie dwóch kwadratów, Jacobi). *Niech $r_2(n)$ oznacza liczbę sposobów wyrażenia liczby naturalnej jako sumy kwadratów dwóch liczb całkowitych. Ponadto niech $d_k(n)$ oznacza liczbę naturalnych dzielników n , które dają resztę k modulo 4. Wtedy dla każdej liczby naturalnej n zachodzi:*

$$r_2(n) = 4(d_1(n) - d_3(n)).$$

Uwaga. Ponieważ dla każdej liczby pierwszej postaci $p = 4k + 1$ mamy $d_1(p) = 2$ oraz $d_3(p) = 0$, twierdzenie Fermata jest bezpośrednim wnioskiem z tego twierdzenia.

Nasz dowód tego twierdzenia (właściwie jedynie bardzo krótki szkic) będzie oparty na następującym twierdzeniu:

Twierdzenie (potrójny iloczyn Jacobiego). *Dla dowolnych liczb zespolonych q i z spełniających $|q| < 1$ oraz $z \neq 0$ zachodzi następująca równość:*

$$\prod_{m=1}^{\infty} (1 - q^{2m}) (1 + q^{2m-1} z^2) (1 + q^{2m-1} z^{-2}) = \sum_{n=-\infty}^{\infty} q^{n^2} z^{2n}.$$

Korzystając z tego twierdzenia kilka razy, można udowodnić, że:

$$\begin{aligned} k^{-1} \prod_{m=1}^{\infty} (1 - x^{4m}) (1 - x^{4m-3} k^2) (1 - x^{4m-1} k^{-2}) \\ + \prod_{m=1}^{\infty} (1 - x^{4m}) (1 - x^{4m-3} k^{-2}) (1 - x^{4m-1} k^2) \\ = \prod_{m=1}^{\infty} (1 - (-x)^m) (1 + (-x)^m k) (1 + (-x)^{m-1} k^{-1}). \end{aligned}$$

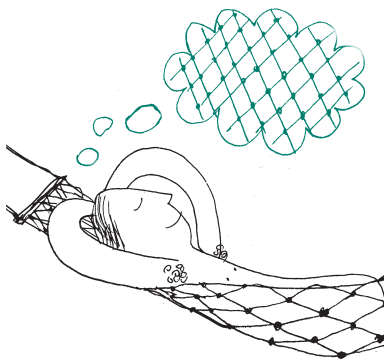
Różniczkując obie strony tej równości względem k i podstawiając $k = -1$, otrzymujemy następującą równość:

$$\left(\sum_{n=-\infty}^{\infty} x^{n^2} \right)^2 = 1 + 4 \sum_{n=1}^{\infty} \left(\frac{x^{4n-3}}{1 - x^{4n-3}} - \frac{x^{4n-1}}{1 - x^{4n-1}} \right).$$

W rzeczywistości ta równość jest równoważna twierdzeniu Jacobiego. Aby to zobaczyć, przekształćmy najpierw prawą stronę. Standardowy wzór na sumę szeregu geometrycznego implikuje:

$$\begin{aligned} \sum_{n=1}^{\infty} \left(\frac{x^{4n-3}}{1 - x^{4n-3}} - \frac{x^{4n-1}}{1 - x^{4n-1}} \right) \\ = \sum_{n=1}^{\infty} \left(\sum_{k=1}^{\infty} x^{(4n-3)k} - \sum_{k=1}^{\infty} x^{(4n-1)k} \right) \\ = \sum_{n \in \mathbb{N} \setminus \{0\}, k \in \mathbb{N}} (x^{(4n-3)k} - x^{(4n-1)k}) = \sum_{k=1}^{\infty} (d_1(k) - d_3(k)) x^k. \end{aligned}$$

Istnieje wiele kierunków, w których twierdzenie Fermata może być uogólnione lub rozwinięte: można badać liczby pierwsze postaci $x^2 + ny^2$; uzyskać kryterium dla wszystkich liczb naturalnych, by mogły być wyrażone jako suma dwóch kwadratów (twierdzenie o sumie dwóch kwadratów); rozważać większą liczbę kwadratów (twierdzenie Legendre'a o sumie trzech kwadratów, twierdzenie Lagrange'a o sumie czterech kwadratów) i wiele innych.



Teraz przekształćmy lewą stronę:

$$\left(\sum_{n=-\infty}^{\infty} x^{n^2}\right)^2 = (\dots + x^{(-1)^2} + x^{0^2} + x^{1^2} + \dots)^2 = 1 + \sum_{k=1}^{\infty} r_2(k)x^k.$$

Zatem udowodniliśmy, że

$$1 + \sum_{k=1}^{\infty} r_2(k)x^k = 1 + \sum_{k=1}^{\infty} (4d_1(k) - 4d_3(k))x^k,$$

co kończy dowód twierdzenia Jacobiego. $\square$

Paul Erdős, jeden z najwybitniejszych matematyków XX wieku, często mówił o Księdze, w której Bóg miałby przechowywać idealne dowody twierdzeń matematycznych. Powszechnie znane jest jego powiedzenie: „Nie musisz wierzyć w Boga, ale jako matematyk powinieneś wierzyć w Księgę”. Jeśli taka Księga naprawdę istnieje, wierzę, że powyższe dowody, będące doskonałym przykładem głębokich powiązań między różnymi obszarami matematyki, z pewnością zasłużyłyby na szczególne w niej miejsce.

Gdzie tu jest graf?

Sylvia SAPKOWSKA*

* Studentka, Wydział Matematyki, Informatyki i Mechaniki, Uniwersytet Warszawski

Algorytmy przeszukiwania grafów pojawiają się wszędzie. Bez nich nie działałaby Twoja ulubiona wyszukiwarka internetowa, a swoje „ścieżki w grafie” prawdopodobnie optymalizujesz podświadomie, na przykład planując swój dzień. Połączenia między Tobą, Twoimi przyjaciółmi i rodziną również tworzą graf, który gorączkowo przeszukujemy podczas spotkań towarzyskich.

Skoro algorytmy przeszukiwania grafów mają wiele zastosowań w codziennym życiu, nie powinno nas dziwić, że wiele zadań olimpijskich z matematyki można również rozwiązać, wykorzystując tę ideę.

Przeszukiwanie w głąb

Zanim przejdziemy dalej, omówmy najprostszy algorytm przeszukiwania grafów – **przeszukiwanie w głąb**, w skrócie DFS (*depth first search*).

Możemy o nim myśleć jak o „spacerze” po grafie. Podczas wykonywania algorytmu przechowujemy informację o tym, czy dany wierzchołek został już odwiedzony. Zaczynamy od wybrania wierzchołka początkowego. Kiedy w trakcie wykonywania algorytmu znajdziemy się w jakimś wierzchołku v , wykonujemy następujące instrukcje:

1. Oznacz wierzchołek v jako odwiedzony.
2. Dla każdego sąsiedniego wierzchołka u aktualnego wierzchołka v , jeśli u **nie został** jeszcze odwiedzony – rekurencyjnie wykonaj tę procedurę dla u .
3. Po zbadaniu wszystkich sąsiednich wierzchołków v wróć do poprzedniego wierzchołka – tego, z którego przyszlismy do v .

O algorytmie DFS można myśleć analogicznie do zwiedzania. Wyobraź sobie, że jesteś turystą w tętniącym życiem mieście, trzymasz w rękach mapę i chcesz zobaczyć każdą ukrytą perełkę, jaką to miasto ma do zaoferowania. Zaczynasz od znanego muzeum, chłonąc sztukę i historię, po czym rozglądasz się za kolejnym, jeszcze nieodwiedzonym miejscem. Jeśli znajdziesz nową kawiarnię, urokliwą alejkę lub inny zabytek, udajesz się tam, ciesząc się nowymi widokami i dodając je do swojej „listy odwiedzonych miejsc”. Gdy dojdiesz do punktu, w którym wszystko w okolicy zostało zwiedzone, czas wrócić – cofnąć się do ostatniego miejsca, które miało jeszcze niezbadane ścieżki. Stamtąd kontynuujesz przygodę, odkrywając nowe ciekawostki i poruszając się naprzód, aż w pełni zwiedzisz wszystkie możliwe miejsca. Ta metoda zwiedzania zapewnia, że nie ominiesz żadnej lokalizacji. W końcu odwiedźsz całe miasto, nie pozostawiając żadnego zabytku nieodkrytym ani żadnego zakątka niezbadanym! W przypadku

Pierwsza wersja przeszukiwania w głąb została wykorzystana w XIX wieku przez francuskiego matematyka Charlesa Pierre’a Trémaux jako strategia rozwiązywania labiryntów, jak ten poniżej. Komórki siatki można interpretować jako wierzchołki grafu. Krawędzie w grafie reprezentują wtedy możliwość ruchu, więc łączymy wierzchołki wtedy i tylko wtedy, gdy komórki reprezentowane przez te wierzchołki sąsiadują ze sobą bokiem.

