Ułamki łańcuchowe a sumy dwóch kwadratów

Marcin MAZUR

O rozkładaniu liczb pierwszych na sumę kwadratów można przeczytać w *Delcie* 12/1990.

Opiszemy w tym artykule zadziwiające zastosowanie ułamków łańcuchowych: do dowodu twierdzenia Fermata o rozkładzie na sumę kwadratów.

Twierdzenie 1. *Liczba pierwsza p jest sumą dwóch kwadratów liczb całkowitych wtedy i tylko wtedy, gdy $p = 2$ lub $4|(p-1)$.*

Jak wykazał Lagrange, *każda* liczba naturalna jest sumą kwadratów co najwyżej czterech liczb naturalnych.

Dowód, opublikowany przez H.J.S. Smitha w roku pańskim 1855, należy do najpiękniejszych perełek elementarnej teorii liczb, pomyślałem więc, że warto podzielić się nim z Czytelnikami *Delty*.

Zauważmy na początek, że $2 = 1^2 + 1^2$. Ponadto, gdy $n = 2k+1$ jest sumą dwóch kwadratów liczb całkowitych, wówczas $4|(n-1)$. Wynika to z prostej obserwacji, że kwadrat liczby całkowitej daje z dzielenia przez 4 resztę 0 lub 1. Pozostaje więc wykazać, że liczby pierwsze $p \equiv 1 \pmod 4$ są sumami dwóch kwadratów.

Dalszą część dowodu na moment odłóżmy, by przypomnieć podstawowe fakty dotyczące ułamków łańcuchowych. Dla dowolnego ciągu liczb całkowitych a_0, $a_1 > 0$, $a_2 > 0, \ldots$ symbolem $[a_0; a_1; \ldots; a_n]$ oznaczamy liczbę

$$[a_0; a_1; \ldots; a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_n}}},$$

zwaną skończonym ułamkiem łańcuchowym o wyrazach $a_0, \ldots, a_n$. Oczywiście $[a_0; a_1; \ldots; a_n]$ jest liczbą wymierną, więc można ją zapisać jednoznacznie w postaci p_n/q_n, gdzie liczba q_n jest dodatnia i ułamek p_n/q_n jest nieskracalny.

Ogólnie, każdą liczbę wymierną $w \neq 0$ można jednoznacznie przedstawić w postaci p/q, gdzie $q > 0$ i liczby p, q są względnie pierwsze. Liczbę p nazywamy *licznikiem*, a liczbę q *mianownikiem* liczby wymiernej w.

Kluczowe znaczenie będzie dla nas miała tożsamość

$$[a_0; a_1; \ldots; a_n; a_{n+1}; \ldots; a_N] = \frac{p_n r + p_{n-1}}{q_n r + q_{n-1}}, \tag{1}$$

gdzie $r = [a_{n+1}; \ldots; a_N]$. Dla wygody Czytelnika podamy szkic jednego z licznych jej dowodów.

♠ Rozpatrzmy ciągi $\tilde{p}_n$, $\tilde{q}_n$ określone następująco: $\tilde{p}_{-1} = 1$, $\tilde{p}_0 = a_0$, $\tilde{p}_{n+1} = a_{n+1}\tilde{p}_n + \tilde{p}_{n-1}$ oraz $\tilde{q}_{-1} = 0$, $\tilde{q}_0 = 1$, $\tilde{q}_{n+1} = a_{n+1}\tilde{q}_n + \tilde{q}_{n-1}$. Oczywista indukcja dowodzi, że $\tilde{q}_n\tilde{p}_{n-1} - \tilde{p}_n\tilde{q}_{n-1} = (-1)^n$. W szczególności, liczby $\tilde{p}_n$ i $\tilde{q}_n$ są względnie pierwsze dla każdego $n \geq 0$. Ponadto $\tilde{q}_n > 0$ dla $n \geq 0$. Rozpatrzmy teraz funkcję

$$f_n(x) = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_n + \cfrac{1}{x}}}}.$$

Przez indukcję łatwo dowodzimy, że

$$f_n(x) = \frac{\tilde{p}_n x + \tilde{p}_{n-1}}{\tilde{q}_n x + \tilde{q}_{n-1}}. \tag{2}$$

Wobec tego, $\lim\limits_{x\to\infty} f_n(x) = \tilde{p}_n/\tilde{q}_n$. Z drugiej strony, wprost z definicji f_n wynika, że $\lim\limits_{x\to\infty} f_n(x) = [a_0; a_1; \ldots; a_n] = p_n/q_n$. Zatem $\tilde{p}_n = p_n$ i $\tilde{q}_n = q_n$ dla każdego $n \geq 0$. By zakończyć dowód tożsamości (1), pozostaje zauważyć, że obie jej strony są równe $f_n(r)$. ♠

Poniższe fakty są prostymi konsekwencjami tożsamości (1) i jej dowodu:

1. *Zachodzą równości $p_{n+1} = a_{n+1}p_n + p_{n-1}$ i $q_{n+1} = a_{n+1}q_n + q_{n-1}$. W szczególności, jeśli $a_0 \geq 1$, to $p_n \geq 2$ dla każdego $n > 0$.*
2. *Mamy $q_n p_{n-1} - p_n q_{n-1} = (-1)^n$.*
3. *Jeśli $a_0 > 0$, to wówczas $[a_n; a_{n-1}; \ldots; a_0] = p_n/p_{n-1}$. W szczególności liczby $[a_0; a_1; \ldots; a_n]$ i $[a_n; a_{n-1}; \ldots; a_0]$ mają takie same liczniki.*

Ponadto,

4. *Jeśli $a_0 \geq 2$ i $a > 0$, to licznik liczby $[a_0; a_1; \ldots; a_n; a; a_n; a_{n-1}; \ldots; a_0]$ jest liczbą złożoną.*

♠ Istotnie, z własności 3 wynika, że

$$[a; a_n; a_{n-1}; \ldots; a_0] = a + 1/[a_n; a_{n-1}; \ldots; a_0] = a + p_{n-1}/p_n.$$

Na mocy tożsamości (1) otrzymujemy równości

$$[a_0; a_1; \ldots; a_n; a; a_n; a_{n-1}; \ldots; a_0] = \frac{p_n(a + \frac{p_{n-1}}{p_n}) + p_{n-1}}{q_n(a + \frac{p_{n-1}}{p_n}) + q_{n-1}} = \frac{p_n(ap_n + 2p_{n-1})}{ap_nq_n + q_np_{n-1} + q_{n-1}p_n}.$$

Ponieważ liczby p_n i q_np_{n-1} są względnie pierwsze (własność 2!), więc liczby p_n i $ap_nq_n + q_np_{n-1} + q_{n-1}p_n$ też są względnie pierwsze. Wobec własności 2 mamy

$$\begin{aligned} ap_nq_n + q_np_{n-1} + q_{n-1}p_n &= ap_nq_n + 2q_np_{n-1} - (-1)^n = \\ &= q_n(ap_n + 2p_{n-1}) - (-1)^n, \end{aligned}$$

przeto liczby $ap_n + 2p_{n-1}$ i $ap_nq_n + q_np_{n-1} + q_{n-1}p_n$ także są względnie pierwsze. Zatem licznik rozważanego ułamka łańcuchowego jest równy $p_n(ap_n + 2p_{n-1})$. Ponieważ $a_0 \geq 2$, więc z własności 1 wynika, że liczby p_n i $ap_n + 2p_{n-1}$ są większe niż 1, a zatem licznik nasz jest liczbą złożoną. ♠

5. *Jeśli $a_0 > 0$, to licznik liczby $[a_0; a_1; \ldots; a_n; a_n; a_{n-1}; \ldots; a_0]$ jest równy $p_n^2 + p_{n-1}^2$.*

♠ W samej rzeczy, wobec własności 3, mamy $[a_n; a_{n-1}; \ldots; a_0] = p_n/p_{n-1}$, a zatem tożsamość (1) daje

$$[a_0; a_1; \ldots; a_n; a_n; a_{n-1}; \ldots; a_0] = \frac{p_n\frac{p_n}{p_{n-1}} + p_{n-1}}{q_n\frac{p_n}{p_{n-1}} + q_{n-1}} = \frac{p_n^2 + p_{n-1}^2}{q_np_n + q_{n-1}p_{n-1}}.$$

Ponadto, wobec własności 2, otrzymujemy

$$\begin{aligned} p_{n-1}(q_np_n + q_{n-1}p_{n-1}) = p_n(p_{n-1}q_n) + q_{n-1}p_{n-1}^2 = p_n(p_nq_{n-1} + (-1)^n) + q_{n-1}p_{n-1}^2 = \\ = q_{n-1}(p_n^2 + p_{n-1}^2) + (-1)^np_n. \end{aligned}$$

Zatem każdy wspólny dzielnik d liczb $p_n^2 + p_{n-1}^2$ i $q_np_n + q_{n-1}p_{n-1}$ dzieli p_n i w konsekwencji również p_{n-1}^2, a że NWD $(p_n, p_{n-1}) = 1$, więc $d = 1$. ♠

Własność 5 niesie następującą pokusę: dla danej liczby pierwszej $p \equiv 1 \pmod 4$ spróbujmy znaleźć taką liczbę naturalną i niepodzielną przez p, że $p/i = [a_0; \ldots; a_n; a_n; \ldots; a_0]$, a udowodnimy tym samym, że p jest sumą dwóch kwadratów. Na pierwszy rzut oka pomysł nie wygląda zbyt obiecująco. W szczególności, nie widać, jaką rolę miałby odgrywać warunek $p \equiv 1 \pmod 4$. Niebawem jednakże wszystko stanie się jasne, a niejasny i zwariowany pomysł zmieni się w precyzyjny dowód.

Przede wszystkim, warto sobie zdać sprawę, że – na szczęście! – każdą liczbę wymierną można przedstawić w postaci ułamka łańcuchowego i to na dokładnie dwa różne sposoby. Ściślej biorąc, każda liczba wymierna w może być jednoznacznie zapisana w postaci $[a_0; a_1; \ldots; a_n]$ dla pewnych liczb całkowitych $a_0, a_1 > 0, \ldots, a_{n-1} > 0$ i $a_n = 1$, gdzie $n > 0$. Przedstawienie takie będziemy nazywali *długim*. Drugim przedstawieniem liczby w jest $[a_0; a_1; \ldots; a_{n-1} + 1]$, które będziemy dalej zwać *krótkim*. Jeśli w nie jest liczbą całkowitą, to $n > 1$ i krótkie przedstawienie w jest jedynym przedstawieniem, kończącym się liczbą większą niż 1. Uzasadnienie omawianych faktów nie jest trudne i opiera się na następującej, prostej obserwacji: ciąg liczb $a_0, \ldots, a_{n-2}, a_{n-1} + 1$ jest równy ciągowi części całkowitych liczb b_k określonych w sposób rekurencyjny wzorami $b_0 = w$, $b_{k+1} = 1/\{b_k\}$, gdzie przez $\{x\}$ oznaczamy część ułamkową liczby x. Czytelnik z łatwością uzupełni szczegóły.

Weźmy teraz liczbę pierwszą p i rozpatrzmy zbiór $S_p = \{p/2, p/3, \ldots, p/\frac{p-1}{2}\}$. Wszystkie liczby z tego zbioru mają liczniki równe p i są większe od 2. Jeśli więc $w \in S_p$ ma krótkie rozwinięcie na ułamek łańcuchowy, równe $[a_0; \ldots; a_n]$, to mamy $a_0 \geq 2$ i $a_n \geq 2$. Zatem, liczba $\Phi(w) = [a_n; \ldots; a_0]$ jest też większa od 2 i (wobec własności 3) ma licznik równy p. Tym samym $\Phi(w) \in S_p$ i oczywiście $[a_n; \ldots; a_0]$ jest krótkim rozwinięciem liczby $\Phi(w)$ (nie może to być rozwinięcie długie, gdyż $a_0 \geq 2$). Zatem $\Phi(\Phi(w)) = w$, a więc funkcja $\Phi : S_p \to S_p$ jest inwolucją, tzn. $\Phi \circ \Phi = \mathrm{Id}$. Teraz widać, jakie znaczenie ma warunek $p \equiv 1 \pmod 4$. Otóż, jeśli $4|(p-1)$, to zbiór S_p ma nieparzystą liczbę elementów. Czytelnik bez trudu uzasadni, że jeśli pewien zbiór S ma nieparzystą liczbę elementów i funkcja $f : S \to S$ jest inwolucją, to $f(s) = s$ dla pewnego $s \in S$. W naszym przypadku dla pewnego $1 < j \leq (p-1)/2$ mamy $\Phi(p/j) = p/j$.

Innymi słowy, krótkie rozwinięcie liczby $p/j = [a_0; \ldots; a_n]$ jest symetryczne: $a_i = a_{n-i}$. Gdyby $n = 2k$, to mielibyśmy $p/j = [a_0; \ldots; a_{k-1}; a_k; a_{k-1}, \ldots; a_0]$. Jest to jednak niemożliwe: wobec własności 4 licznik ułamka łańcuchowego z prawej strony jest liczbą złożoną. Zatem n jest liczbą nieparzystą i $p/j = [a_0; \ldots; a_k; a_k; \ldots; a_0]$. Wobec własności 5 licznik p tego ułamka łańcuchowego jest sumą dwóch kwadratów. Twierdzenie 1 zostało więc dowiedzione.

Dla Czytelników mamy dwie propozycje:

Zadanie. Udowodnić Twierdzenie 1 rozważając długie rozwinięcia i zbiór $\{p/\frac{p+1}{2}, \ldots, p/(p-2)\}$.

Problem. Łatwo zauważyć, że definicja funkcji Φ ma sens dla dowolnej liczby wymiernej większej od 2. Warto zbadać własności otrzymanej w ten sposób inwolucji.

Teraz łatwo już uzyskać następującą charakteryzację sum dwóch kwadratów:

Twierdzenie 2. *Liczba naturalna n jest sumą dwóch kwadratów liczb całkowitych wtedy i tylko wtedy, gdy każda liczba pierwsza postaci $4k+3$ występuje w rozwinięciu liczby n na czynniki pierwsze w parzystej potędze.*

Dowód opiera się na dwóch obserwacjach. Po pierwsze, jeśli $n = a^2 + b^2$ i liczba pierwsza p postaci $4k+3$ dzieli n, to $p|a$ i $p|b$. Po drugie, jeśli $n = a^2 + b^2$ i $m = c^2 + d^2$, to wówczas $mn = (ac - bd)^2 + (ad + bc)^2$. Szczegóły tradycyjnie pozostawiamy Czytelnikowi.

Na zakończenie – uwaga natury ogólnej. W matematyce często się zdarza, że na pozór zupełnie nie związane z rozważanym problemem gałęzie są źródłem kluczowego pomysłu, który prowadzi do długo poszukiwanego rozwiązania. Powyższe rozważania stanowią miniaturowy przykład tej prawidłowości. Większość istotnych przełomów w matematyce została osiągnięta na tej właśnie drodze. Ważne jest więc, by Czytelnik poważnie zainteresowany jakąkolwiek dziedziną matematyki nie ograniczał się jedynie do problemów bezpośrednio z nią związanych, lecz wręcz przeciwnie, starał się zgłębiać przeróżne matematyczne teorie. Być może pewnego dnia zdoła je połączyć w swych rozważaniach i w nagrodę otrzyma wspaniały, niespodziewany wynik, czego szczerze i gorąco życzę.

Zadania

Przygotował Paweł STRZELECKI

M 877. Dane są dwie trójki różnych od zera liczb rzeczywistych, (x, y, z) oraz (a, b, c), o tej własności, że $a + b + c = x + y + z = 0$. Udowodnić, że

$$\frac{a^3 + b^3 + c^3}{x^3 + y^3 + z^3} = \frac{abc}{xyz}.$$

Rozwiązanie na str. 4

M 878. Dane są dwie trójki liczb dodatnich, (x, y, z) oraz (a, b, c). Wiadomo, że

$$\min(a, b, c) \leq \min(x, y, z) \quad \text{oraz} \quad \max(x, y, z) \leq \max(a, b, c)$$

Ponadto, $a + b + c = x + y + z$ i $abc = xyz$. Udowodnić, że zbiory $\{x, y, z\}$ i $\{a, b, c\}$ są równe.
Rozwiązanie na str. 7

M 879. Dziesięciu widzów ogląda w kinie nudny film. Wszyscy zajmują miejsca w tym samym rzędzie. Cierpliwość każdego z widzów wyczerpuje się, ale w losowej kolejności (każdą kolejność uznajemy za równoprawdopodobną). Widz zniecierpliwiony nudnym filmem od razu wychodzi z kina na świeże powietrze. Jakie jest prawdopodobieństwo tego, że któryś z widzów będzie zmuszony przeszkodzić innemu widzowi, żeby dostać się do wyjścia?
Rozwiązanie na str. 15

Redaguje Ewa CZUCHRY

F 497. Jaka jest rzeczywista głębokość rzeki, jeśli przy określaniu „na oko", w kierunku pionowym, jej głębokość wydaje się wynosić 2 m?
Rozwiązanie na str. 15

F 498. Jakie najmniejsze wymiary powinno mieć zwierciadło płaskie i jak je należy ustawić, żeby można się było w nim w całości obejrzeć?
Rozwiązanie na str. 10