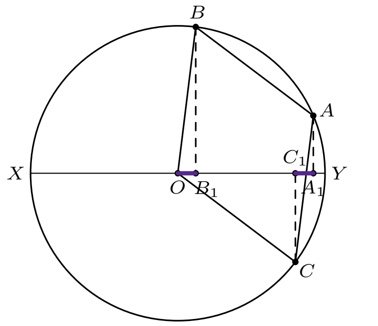
Rozwiązanie zadania M 1753.
Rozważmy konfigurację taką, jak na rysunku.



Czworokąt $ABOC$ jest rombem, więc rzuty prostokątne odcinków AC i BO na prostą XY są równe, czyli $OB_1 = A_1C_1$. Ponieważ $OX = OY$, dostajemy

$$XB_1 = XO + OB_1 = \\ = OY + A_1C_1 = OA_1 + C_1Y,$$

co kończy rozwiązanie.

Zbiór E ma $6m$ elementów, gdyż każdemu trójkątowi w grafie odpowiada $3! = 6$ elementów W^3. Rzut prostokątny zbioru E na płaszczyznę OXY (tzn. zbiór E_{xy}) zawiera wyłącznie pary wierzchołków połączonych krawędzią. Ponieważ każda krawędź wyznacza dwa wierzchołki, więc liczba wierzchołków w zbiorze E_{xy} nie przekracza $2k$, gdzie k jest liczbą krawędzi. Podobnie jest w każdym z pozostałych rzutów E_{xz} oraz E_{yz} zbioru E. Zatem zgodnie z nierównością Loomisa–Whitneya $(6m)^2 \leqslant 2k \cdot 2k \cdot 2k$, skąd wynika teza. □

Oszacowanie liczby trójkątów w grafie ma pewien związek z teorią baz danych w kontekście *zapytań trójkątnych* (*triangle queries*), które często pojawiają się w sytuacjach praktycznych. Ograniczenie z twierdzenia pomaga oszacować złożoność algorytmów, które realizują te zapytania. Więcej na ten temat można przeczytać na przykład w rozdziale 1.2 pracy Hunga Ngo *Worst-Case Optimal Join Algorithms: Techniques, Results, and Open Problems*, dostępnej w serwisie `arxiv.org`. A ponieważ trójkąty w grafach są takie przydatne, przedstawimy jeszcze jedno klasyczne twierdzenie, które jest z nimi związane.

Twierdzenie 3 (W. Mantel, 1906). *Jeśli graf nie zawiera trójkątów, to ma co najwyżej $\frac{n^2}{4}$ krawędzi.*

Dowód. Niech $G = \{v_1, v_2, \ldots, v_n\}$ będzie grafem o n wierzchołkach, niezawierającym trójkąta. Przez d_i oznaczmy liczbę krawędzi wychodzących z wierzchołka v_i. Niech A będzie największym w G zbiorem wierzchołków, w którym żadne dwa wierzchołki nie są połączone krawędzią. Zauważmy, że zbiór $B = G \setminus A$ zawiera przynajmniej jeden z końców każdej krawędzi, zatem $|E| \leqslant \sum_{v_i \in B} d_i$ (w sumie po prawej stronie liczone są wszystkie krawędzie, przy czym dwukrotnie liczone są te między dwoma wierzchołkami z B). Ponadto jeśli weźmiemy dowolny wierzchołek, to żaden z jego sąsiadów nie może być połączony (inaczej powstałby trójkąt), w związku z czym liczba tych sąsiadów nie może być większa od $|A|$ (z definicji zbioru A). Stosując nierówność między średnią arytmetyczną i geometryczną, otrzymujemy zatem

$$|E| \leqslant \sum_{v_i \in B} d_i \leqslant |B| \cdot |A| \leqslant \left(\frac{|A| + |B|}{2}\right)^2 = \frac{n^2}{4},$$

co kończy dowód i niniejszy artykuł. □

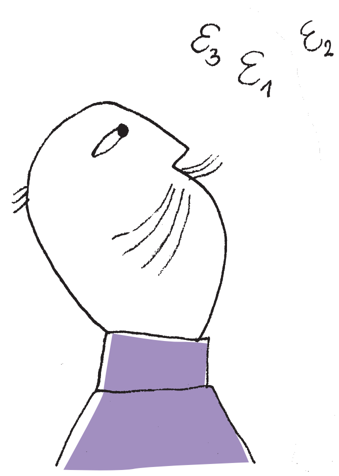
Twierdzenie Tomaszewskiego

Aleksander PAWLEWICZ

W kwietniu 1986 roku na łamach czasopisma „The American Mathematical Monthly” Richard Guy zaprezentował pewną hipotezę, wskazując jako jej autora Boguslava Tomaszewskiego. Hipoteza brzmiała tak:

Hipoteza. Rozważmy n liczb rzeczywistych $a_1, a_2, \ldots, a_n$ spełniających równość $a_1^2 + \ldots + a_n^2 = 1$. Czy wśród 2^n wyrażeń postaci $|\varepsilon_1 a_1 + \ldots + \varepsilon_n a_n|$, gdzie każda z liczb $\varepsilon_1, \ldots, \varepsilon_n$ jest równa $+1$ lub -1, wyrażeń o wartości > 1 może być więcej niż tych o wartości $\leqslant 1$?

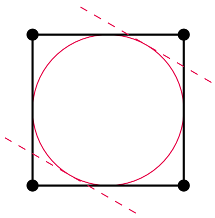
Przez 34 lata problem pozostawał otwarty. Pomimo licznych prób nie był znany żaden dowód potwierdzający hipotezę ani ją obalający. Dopiero w 2020 roku na portalu arxiv.org pojawił się artykuł *Proof of Tomaszewski's Conjecture on Randomly Signed Sums* zawierający negatywną odpowiedź na postawione wyżej pytanie. Autorami artykułu są Nathan Keller i Ohad Klein. Praca ta po pewnych modyfikacjach liczy 76 stron, a w skróconej wersji ukazała się w „Advances in Mathematics”. Już sama objętość świadczy o dużej trudności dowodu.



W niniejszym artykule zwięźle przedstawimy historię problemu, a także pewne wnioski wynikające z hipotezy, którą aktualnie możemy już nazywać twierdzeniem: spośród wyrażeń postaci $|\varepsilon_1 a_1 + \ldots + \varepsilon_n a_n|$ co najmniej połowa musi być nie większa od 1. Sami autorzy dowodu preferują sformułowanie probabilistyczne:

Twierdzenie. Niech $X = \sum_{i=1}^n a_i x_i$, gdzie a_i to liczby spełniające $\sum_{i=1}^n a_i^2 = 1$, zaś x_i są niezależnymi zmiennymi losowymi przyjmującymi wartości $+1$ i -1 z równym prawdopodobieństwem. Wówczas $\mathbb{P}\left(|X| \leqslant 1\right) \geqslant \frac{1}{2}$.

Istotnie, w opisanym modelu każdy z 2^n możliwych ciągów $\varepsilon_1, \ldots, \varepsilon_n$ jest równie prawdopodobny jako wynik losowania $x_1, \ldots, x_n$. Prawdopodobieństwo $\mathbb{P}(|X| \leqslant 1)$ opisuje więc proporcję liczby *dobrych* ciągów (czyli takich, dla których omawiane wyrażenie jest nie większe od 1) do liczby wszystkich ciągów.



Niezależnie od wyboru kierunku stycznych co najmniej połowa wierzchołków leży między nimi

To samo twierdzenie można równoważnie sformułować jeszcze inaczej: jeśli w n-wymiarową kostkę wpisano n-wymiarową kulę, a następnie wybrano dwie równoległe hiperpłaszczyzny styczne do kuli, to przynajmniej połowa wierzchołków kostki leży pomiędzy tymi hiperpłaszczyznami (lub na nich). Ilustrację dla $n = 2$ zamieszczamy obok.

Historia. Zanim uzyskano nierówność sformułowaną w powyższym twierdzeniu, przez kilkanaście lat otrzymywano stopniowo coraz lepsze oszacowania. Pierwszym znaczącym wynikiem na tym polu była nierówność $\mathbb{P}(|X| \leqslant 1) \geqslant \mathbf{\frac{1}{3}}$ otrzymana w 2002 roku przez A. Ben-Tala, A. Nemirovskiego i C. Roosa. Kolejny rezultat uzyskał w 2012 roku I. Shnurnikov – udowodnił mianowicie nierówność $\mathbb{P}(|X| \leqslant 1) \geqslant \mathbf{0{,}36}$. Pierwszymi, którzy przekroczyli barierę 3/8, byli R. B. Boppana i R. Holzman. W 2017 roku wykazali oni, że wartość $\mathbb{P}(|X| \leqslant 1)$ może być oszacowana z dołu przez $\mathbf{0{,}406}$. Kolejne dwa ograniczenia dolne ukazały się w podobnym czasie. Najpierw R. B. Boppana i H. Hendriks wykazali, że $\mathbb{P}(|X| \leqslant 1) \geqslant \mathbf{0{,}428}$, a następnie V. Dvořák, P. van Hintum i M. Tiba pokazali, że po prawej stronie powyższej nierówności można wstawić $\mathbf{0{,}46}$. Dwa ostatnie wyniki zostały opublikowane w latach 2020 i 2021.

W czasie gdy dowodzono coraz lepsze oszacowania dolne, omówione powyżej, pojawiały się także coraz lepsze wyniki dowodzące pożądanej nierówności $\mathbb{P}(|X| \leqslant 1) \geqslant \frac{1}{2}$ w szczególnych przypadkach. Najpierw w roku 1992 pojawił się dowód twierdzenia, gdy liczba zmiennych x_i i liczb a_i nie przekracza dziewięciu. Wynik ten zawdzięczamy R. Holzmanowi i D. J. Kleitmanowi. Innym kierunkiem dowodzenia hipotezy było wykazanie jej dla wszystkich a_i równych. Dokonał tego M. C. A. van Zuijlen w roku 2011. Kilka lat później, w roku 2018 T. Toufar udowodnił sformułowaną w twierdzeniu nierówność przy założeniu, że wszystkie poza jednym a_i są równe. Nieco innym kierunkiem uogólnień było wykazanie w roku 2015 przez V. K. Bentkusa i D. Dzindzalieta, że nierówność jest prawdziwa, gdy największa z liczb $|a_i|$ nie przekracza 0,16.

Te i wiele innych cennych informacji o Hipotezie i próbach jej dowodzenia można znaleźć we wspomnianym już artykule:

N. Keller, O. Klein, *Proof of Tomaszewski's Conjecture on Randomly Signed Sums*, Adv. Math. 407 (2022). Artykuł dostępny również pod adresem `https://arxiv.org/abs/2006.16834`

Wnioski płynące z twierdzenia. Aby zilustrować twierdzenie, omówimy teraz jego szczególny przypadek: pewną nierówność dwumianową. Przyjmijmy, że wszystkie współczynniki $a_1, \ldots, a_n$ są równe, czyli niech $a_i = \frac{1}{\sqrt{n}}$. Na mocy twierdzenia Tomaszewskiego wiemy, że

$$\mathbb{P}\big(|x_1 + \ldots + x_n| \leqslant \sqrt{n}\big) \geqslant \frac{1}{2},$$

gdzie x_i są niezależnymi zmiennymi losowymi o rozkładzie $\mathbb{P}(x_i = \pm 1) = \frac{1}{2}$.

Oszacowanie obok warto porównać z centralnym twierdzeniem granicznym: według niego rozkład zmiennej losowej $\frac{x_1+\ldots+x_n}{\sqrt{n}}$ zbiega przy $n \to \infty$ do rozkładu normalnego, w szczególności prawdopodobieństwo $\mathbb{P}\left(|x_1 + \ldots + x_n| \leqslant \sqrt{n}\right)$ zbiega do $\mathbb{P}(|\mathcal{N}(0,1)| \leqslant 1) \approx 0{,}68$.

Jest to lepszy wynik, ale prawdziwy jedynie *w granicy*. Dla ustalonego n prawdopodobieństwo może być istotnie mniejsze.

Wróćmy do sformułowania deterministycznego. Powyższa nierówność oznacza, że spośród wszystkich 2^n wyborów wartości ± 1 w ciągu $\varepsilon_1, \ldots, \varepsilon_n$ mamy

$$|\varepsilon_1 + \ldots + \varepsilon_n| \leqslant \sqrt{n}$$

dla co najmniej połowy z nich, a więc co najmniej 2^{n-1}. Odpowiedzmy teraz w inny sposób na pytanie o liczbę takich wyborów. Jeśli w wybranym ciągu jest k liczb $+1$ i $n-k$ liczb -1, to suma $\varepsilon_1 + \ldots + \varepsilon_n$ wynosi $k-(n-k)$, czyli $2k-n$. Warunek $|\varepsilon_1 + \ldots + \varepsilon_n| \leqslant \sqrt{n}$ sprowadza się więc do

$$\frac{n-\sqrt{n}}{2} \leqslant k \leqslant \frac{n+\sqrt{n}}{2}.$$

Dla ustalonego k spełniającego te warunki możliwych ustawień ± 1 jest $\binom{n}{k}$, co pozwala przeformułować naszą nierówność w postaci:

$$\sum_{k=\lceil \frac{n-\sqrt{n}}{2} \rceil}^{\lfloor \frac{n+\sqrt{n}}{2} \rfloor} \binom{n}{k} \geqslant 2^{n-1}.$$

Wniosek: około $\sqrt{n}$ największych wyrazów ciągu o wyrazach $\binom{n}{k}$ daje w sumie więcej, niż pozostałe wyrazy tego ciągu.