Własności funkcji ciągłych, funkcje wypukłe

Ostatnie zmiany wprowadzono 20 marca 2014 r, godz. 1:55

Przykład 7.1 Rozważymy sumę tzw. szeregu potęgowego, tj. szeregu postaci $\sum_{n=0}^{\infty} a_n x^n$. Wykażemy, że jeśli dla pewnej liczby $x_0 \neq 0$ szereg $\sum_{n=0}^{\infty} a_n x_0^n$ jest zbieżny i $|x| < |x_0|$, to szereg $\sum_{n=0}^{\infty} a_n n^k x^n$ jest bezwzględnie zbieżny dla każdej liczby naturalnej k. Mamy bowiem $\sum_{n=0}^{\infty} |a_n n^k x^n| = \sum_{n=0}^{\infty} |a_n x_0^n| \cdot n^k \big|\frac{x}{x_0}\big|^n$. Ostatni szereg jest zbieżny bo szereg $\sum_{n=0}^{\infty} n^k \big|\frac{x}{x_0}\big|^n$ jest zbieżny, a ciąg $\big(|a_n x_0^n|\big)$ jest ograniczony, bo jest przecież zbieżny do 0 (warunek konieczny zbieżności szeregu).

Wykażemy, że funkcja przypisująca liczbie x liczbę $\sum_{n=0}^{\infty} a_n x^n$ spełnia warunek Lipschitza w zbiorze $\{x\colon\ |x| \le r\}$, gdzie $r \in (0, |x_0|)$. Załóżmy, że $|x|, |y| \le r$. Mamy wtedy

$$\Big|\sum_{n=0}^{\infty} a_n x^n - \sum_{n=0}^{\infty} a_n y^n\Big| = \Big|\sum_{n=1}^{\infty} a_n\big(x^n - y^n\big)\Big| =$$

$$= |x-y|\Big|\sum_{n=1}^{\infty} a_n\big(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + xy^{n-2} + y^{n-1}\big)\Big| \le$$

$$\le |x-y| \sum_{n=1}^{\infty} n|a_n| r^{n-1} = |x-y| \cdot \frac{1}{r} \cdot \sum_{n=1}^{\infty} n|a_n| r^n .$$

Wykazaliśmy więc, że spełniony jest warunek Lipschitza ze stałą $\frac{1}{r} \cdot \sum_{n=1}^{\infty} n|a_n| r^n$ (zależną od r !). Nie twierdzimy, że jest to najmniejsza stała. Z tego, co udowodniliśmy wynika, że funkcja przypisująca liczbie x liczbę $\sum_{n=0}^{\infty} a_n x^n$ jest ciągła w każdym punkcie zbioru $\{x\colon\ |x| < |x_0|\}$

Funkcje nieciągłe pojawiają się w różnego rodzaju modelach matematycznych. Nie będziemy się nimi zajmować prawie wcale. Pierwszym naszym celem jest zaznajomienie się z podstawowymi własnościami funkcji ciągłych określonych na porządnych dziedzinach. Z naszego punktu widzenia najporządniejszymi możliwymi dziedzinami są przedziały. Rozpoczniemy od intuicyjnie oczywistego twierdzenia nazywanego często mylnie twierdzeniem Darboux. Wydaje się, że pierwszymi, którzy je udowodnili, zresztą niezależnie, byli Bolzano i Cauchy.

Twierdzenie 7.1 (o przyjmowaniu wartości pośrednich)

Jeśli f jest funkcją ciągłą w każdym punkcie pewnego przedziału P i dla pewnych punków x, z przedziału P zachodzi nierówność $f(x) < C < f(z)$, to między punktami x i z znajduje się taki punkt y, że $C = f(y)$.

Dowód. Dla ustalenia uwagi przyjmijmy, że $x < z$. Niech

$$y = \sup\{t \in [x, z]: \quad f(t) < C\}.$$

Oczywiście $x \in \{t \in [x, z]: \quad f(t) < C\}$ i $z \notin \{t \in [x, z]: \quad f(t) < C\}$. Wobec tego liczba y została zdefiniowana poprawnie. Udowodnimy, że $C = f(y)$. Załóżmy, że tak nie jest. Wtedy albo $f(y) < C$ albo $f(y) > C$. Z ciągłości funkcji f w punkcie y wynika, że istnieje taka liczba $\delta > 0$, że jeśli $|t-y| < \delta$, to $|f(t)-f(y)| < |f(y)-C|$. W pierwszym przypadku oznacza to, że

$$f(t) - f(y) \le |f(t) - f(y)| < |f(y) - C| = C - f(y),$$

więc $f(t) < C$, ale stąd wynika, że y nie jest ograniczeniem górnym rozpatrywanego zbioru $\{t \in [x, z]: \quad f(t) < C\}$. W drugim przypadku mamy

$$f(y) - f(t) \le |f(t) - f(y)| < |f(y) - C| = f(y) - C,$$

zatem $C < f(t)$, ale to oznacza, że każda liczba $M \in (y - \delta, y)$ jest ograniczeniem górnym zbioru $\{t \in [x, z]: \quad f(t) < C\}$, więc również w tym przypadku liczba y nie jest jego kresem górnym. Dowód został zakończony. ■

Typowym zastosowaniem twierdzenia o przyjmowaniu wartości pośrednich jest wykazywanie, że funkcja ciągła w każdym punkcie przedziału, przyjmująca w pewnym punkcie tego przedziału wartość dodatnią, a w innym – ujemną, ma między tymi punktami pierwiastek. Można go przybliżać skracając przedział dwukrotnie: sprawdzamy jaki znak ma wartość funkcji w środku przedziału i zastępujemy wyjściowy przedział dwa razy krótszym, na którego końcach funkcja przyjmuje wartości różnych znaków. Daje to w miarę rozsądną metodę przybliżania pierwiastków.*

Ponieważ mówimy od czasu do czasu o wielomianach, więc należy przypomnieć definicję funkcji wielomianowej.

Definicja 7.2 (wielomianu)

Wielomianem nazywamy funkcję $w: \mathbb{R} \longrightarrow \mathbb{R}$ (lub $w: \mathbb{C} \longrightarrow \mathbb{C}$), dla której istnieją takie liczby $a_0, a_1, \ldots, a_n$, że równość $w(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ ma miejsce dla każdej liczby rzeczywistej (zespolonej) x. ■

Lemat 7.3 (o współczynnikach wielomianowej funkcji zerowej)

Jeśli dla każdej liczby x zachodzi równość $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0$, to $a_0 = a_1 = a_2 = \ldots = a_n = 0$.

*Istnieją lepsze, ale bardziej skomplikowane.

Dowód. Załóżmy, że $s_n \neq 0$. Niech $|x| > 1 + \frac{|a_0|+|a_1|+|a_2|+\cdots+|a_{n-1}|}{|a_n|}$. Mamy wtedy $1 < |x| < |x|^2 < |x|^3 < \ldots < |x|^{n-1} < |x|^n$. Stąd i z nierówności trójkąta wynika, że $|a_0 + a_1x + a_2x^2 + \cdots + a_nx^n| \geq |a_nx^n| - |a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1}| \geq$
$\geq |a_nx^n| - \big(|a_0| + |a_1| \cdot |x| + |a_2| \cdot |x|^2 + \cdots + |a_{n-1}| \cdot |x|^{n-1}\big) \geq$
$\geq |a_nx^n| - \big(|a_0| + |a_1| + |a_2| + \cdots + |a_{n-1}|\big) \cdot |x|^{n-1} =$
$= |a_n| \cdot |x|^{n-1}\big(|x| - \frac{|a_0|+|a_1|+|a_2|+\cdots+|a_{n-1}|}{|a_n|}\big) > |a_n| \cdot |x|^{n-1} > 0$, wbrew założeniu. Wobec tego $a_n = 0$. Jednak wtedy $a_{n_1} = 0$ itd. (indukcja). ■

Wniosek 7.4 (o jednoznaczności współczynników)
Jeśli dla każdego x zachodzi równość

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = b_0 + b_1x + b_2x^2 + \cdots + b_mx^m,$$

to $a_0 = b_0$, $a_1 = b_1$, $a_2 = b_2$, ... (przyjmujemy, że $0 = a_{n+1} = a_{n+2} = a_{n+3} = \ldots$ i $0 = b_{m+1} = b_{m+2} = b_{m+3} = \ldots$)

Dowód. Przyjmujemy $c_0 = a_0 - b_0$, $c_1 = a_1 - b_1$, $c_2 = a_2 - b_2$, ... Po przeniesieniu wszystkiego na lewą stronę otrzymujemy równość $c_0 + c_1x + c_2x^2 + \cdots + c_nx^n = 0$, która zachodzi dla każdego x. Wynika stąd, że $0 = c_0 = c_1 = c_2 = \cdots$, a to stwierdzenie jest równoważne tezie. ■

Uwaga 7.5 (o słabszych założeniach)
Na GAL-u pojawi się (lub już pojawił się) tzw. wyznacznik Vandermonde'a. Wtedy będzie można udowodnić, że z tego, że równość $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0$ zachodzi dla $n+1$ różnych liczb x wynika, że $a_0 = a_1 = a_2 = \ldots = a_n = 0$.

Podamy teraz inny dowód właśnie wykazanego lematu przy założeniu, że równość $a_0 + a_1x_j + a_2x_j^2 + \cdots + a_nx_j^n = 0$ zachodzi dla liczb $x_0, x_1, \ldots, x_n$, o których wiemy, że $x_i \neq x_j$ dla $i \neq j$. Nie użyjemy ani wyznaczników, ani granic. Zaczniemy od dowodu dla $n = 1$. Wiemy, że $a_0 + a_1x_0 = 0$ i $a_0 + a_1x_1 = 0$. Odjąwszy te równości stronami otrzymujemy $a_1(x_0 - x_1) = 0$, więc $a_1 = 0$. Wobec tego $0 = a_0 + a_1x_0 = a_0$. Udowodniliśmy twierdzenie w tym wypadku. Załóżmy teraz, że teza zachodzi dla dowolnych liczb $\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_{n-1}, \hat{x}_0, \hat{x}_1, \ldots, \hat{x}_{n-1}$ przy założeniu, że liczby $\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_{n-1}$ są różne ($\hat{x}_i \neq \hat{x}_j$ dla $i \neq j$). Załóżmy teraz, że dane są takie liczby $a_0, a_1, \ldots, a_{n-1}, a_n$ i różne liczby $x_0, x_1, \ldots, x_n$, że

$$a_0 + a_1x_0 + a_2x_0^2 + a_3x_0^3 + \ldots + a_{n-1}x_0^{n-1} + a_nx_0^n = 0,$$
$$a_0 + a_1x_1 + a_2x_1^2 + a_3x_1^3 + \ldots + a_{n-1}x_1^{n-1} + a_nx_1^n = 0,$$
$$a_0 + a_1x_2 + a_2x_2^2 + a_3x_2^3 + \ldots + a_{n-1}x_2^{n-1} + a_nx_2^n = 0,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a_0 + a_1x_n + a_2x_n^2 + a_3x_2^3 + \ldots + a_{n-1}x_n^{n-1} + a_nx_n^n = 0.$$

Odejmiemy teraz pierwsze równanie od drugiego:

$$0 = a_1(x_1 - x_0) + a_2(x_1^2 - x_0^2) + a_3(x_1^3 - x_0^3) + a_{n-1}(x_1^{n-1} - x_0^{n-1}) + a_n(x_1^n - x_0^n) =$$
$$= (x_1 - x_0)\big(a_1 + a_2(x_1 + x_0) + a_3(x_1^2 + x_1x_0 + x_0^2) + \ldots +$$
$$+ a_{n-1}(x_1^{n-2} + x_1^{n-3}x_0 + \ldots + x_0^{n-2}) + a_n(x_1^{n-1} + x_1^{n-2}x_0 + \ldots + x_0^{n-1})\big)\,.$$

Ponieważ $x_-x_0 \neq 0$, więc

$$(a_1 + a_2x_0 + a_3x_0^2 + \ldots + a_{n-1}x_0^{n-2} + a_nx_0^{n-1}) + (a_2 + a_3x_0 + \ldots + a_nx_0^{n-2})x_1 +$$
$$+ (a_3 + a_4x_0 + \ldots + a_nx_0^{n-3})x_1^2 + \ldots + (a_{n-1} + a_nx_0)x_1^{n-2} + a_nx_1^{n-1}\,.$$

W taki sam sposób otrzymujemy równości:

$$(a_1 + a_2x_0 + a_3x_0^2 + \ldots + a_{n-1}x_0^{n-2} + a_nx_0^{n-1}) + (a_2 + a_3x_0 + \ldots + a_nx_0^{n-2})x_2 +$$
$$+ (a_3 + a_4x_0 + \ldots + a_nx_0^{n-3})x_2^2 + \ldots + (a_{n-1} + a_nx_0)x_2^{n-2} + a_nx_2^{n-1}\,,$$

....

$$(a_1 + a_2x_0 + a_3x_0^2 + \ldots + a_{n-1}x_0^{n-2} + a_nx_0^{n-1}) + (a_2 + a_3x_0 + \ldots + a_nx_0^{n-2})x_n +$$
$$+ (a_3 + a_4x_0 + \ldots + a_nx_0^{n-3})x_n^2 + \ldots + (a_{n-1} + a_nx_0)x_n^{n-2} + a_nx_n^{n-1}\,.$$

Przyjmujemy $\hat{a}_0 = a_1 + a_2x_0 + a_3x_0^2 + \ldots + a_{n-1}x_0^{n-2} + a_nx_0^{n-1}$,

$\hat{a}_1 = a_2 + a_3x_0 + \ldots + a_{n-1}x_0^{n-3} + a_nx_0^{n-2}$,

$\hat{a}_2 = a_3 + \ldots + a_{n-1}x_0^{n-4} + a_nx_0^{n-3}$,

....

$\hat{a}_{n-2} = a_{n-1} + a_nx_0$,

$\hat{a}_{n-1} = a_n$,

Mamy więc $\hat{a}_0 + \hat{a}_1x_j\hat{a}_2x_j^2 + \ldots + \hat{a}_{n-1}x_j^{n_1} = 0$ dla $j = 1, 2, \ldots, n$. Z założenia indukcyjnego wynika, że $\hat{a}_0 = 0$, $\hat{a}_1 = 0$, $\hat{a}_2 = 0$, ..., $\hat{a}_{n-1} = 0$. Stąd kolejno wnioskujemy, że zachodzą równości $a_n = 0$, $a_{n-1} = 0$, ..., $a_1 = 0$. Stąd i np. z równości $a_0 + a_1x_0 + a_2x_0^2 + a_3x_0^3 + \ldots + a_{n-1}x_0^{n-1} + a_nx_0^n = 0$ wynika, że również $a_0 = 0$. ■

Po tych twierdzeniach możemy już zdefiniować stopień wielomianu.

Definicja 7.6 (stopnia wielomianu)

Jeśli $w(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ i $a_n \neq 0$, to mówimy, że stopniem wielomianu w jest liczba naturalna n, piszemy $\deg w = n$. Przyjmujemy, że stopniem wielomianu zerowego jest $-\infty$. ■

Czytelnik zechce sprawdzić, że w tej sytuacji prawdziwe są wzory

$$\deg(w_1 \cdot w_2) = \deg w_1 + \deg w_2 \quad \text{oraz} \quad \deg(w_1 + w_2) \leq \max(\deg w_1, \deg w_2)\,.$$

Współczynnik a_n (przy najwyższej potędze zmiennej) nazywamy współczynnikiem kierującym wielomianu.

Twierdzenie 7.7 (o istnieniu pierwiastków wielomianów stopnia nieparzystego)

Każdy wielomian stopnia nieparzystego, tj. funkcja $w : \mathbb{R} \longrightarrow \mathbb{R}$ postaci

$$w(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n\,,$$

gdzie symbole $a_0,\ a_1,\ a_2,\ldots,a_n$ oznaczają liczby rzeczywiste, przy czym $a_n \neq 0$, zaś n jest liczbą naturalna nieparzystą, ma pierwiastek rzeczywisty, tzn. istnieje liczba x_0, taka że $w(x_0) = 0$.

Dowód. $\lim\limits_{x\to\pm\infty} \frac{1}{x} = 0$, zatem zachodzi równość

$$\lim_{x\to\pm\infty} \tfrac{w(x)}{x^n} = \lim_{x\to\pm\infty} \left(\tfrac{a_0}{x^n} + \tfrac{a_1}{x^{n-1}} + \cdots + \tfrac{a_{n-1}}{x} + a_n\right) = a_n\,.$$

Załóżmy, że $a_n > 0$ — przypadek $a_n < 0$ można sprowadzić do poprzedniego przez zastąpienie wielomianu w wielomianem przeciwnym $-w$. Stosując twierdzenia o granicach stwierdzamy, że z jednej strony zachodzi

$$\lim_{x\to\infty} w(x) = \lim_{x\to\infty} x^n \cdot \lim_{x\to\infty} \frac{w(x)}{x^n} = +\infty \cdot a_n = +\infty\,,$$

a z drugiej strony

$$\lim_{x\to-\infty} w(x) = \lim_{x\to-\infty} x^n \cdot \lim_{x\to-\infty} \frac{w(x)}{x^n} = -\infty \cdot a_n = -\infty\,.$$

Z tego wnioskujemy, że wielomian w przyjmuje zarówno wartości dodatnie jak i ujemne: jeśli x jest dostatecznie dużą liczbą dodatnią, to $w(x) > 0$, jeśli $|x|$ jest dostatecznie dużą liczbą dodatnią i $x < 0$, to $w(x) < 0$. Stąd zaś wynika, że wielomian ten przyjmuje w pewnym punkcie wartość 0, czyli że ma pierwiastek. Dowód został zakończony. ■

Powyższe twierdzenie nie oznacza, że umiemy znajdować pierwiastki takiego wielomianu w sposób podobny do stosowanego w szkołach dla wielomianów kwadratowych. Znaleziono w XVI wieku wzory na pierwiastki wielomianów stopnia trzeciego i czwartego, są one znacznie bardziej skomplikowane od wzorów na pierwiastki równania kwadratowego. Na przełomie osiemnastego i dziewiętnastego wieku udowodniono* (Ruffini 1799, Abel 1824, Galois 1830), że nie istnieją wzory na pierwiastki równań stopnia piątego i wyższego. Jest to wynik negatywny, teoretyczny, ale metody rozwinięte dla jego osiągnięcia znalazły znacznie później zastosowania również poza matematyką, np. w fizyce i w chemii. Z punktu widzenia tego wykładu nie ma to większego znaczenia. Mówimy o tym jedynie po to, by uświadomić czytelnikom, że w wielu przypadkach wypisywanie dokładnych wzorów jest niemożliwe, czasem jest możliwe, ale mało sensowne, bo wzory są tak zawiłe, że ich wypisanie niewiele daje, natomiast można używać wzorów przybliżonych, które w wielu przypadkach dają wystarczające rezultaty.

Teraz wykażemy twierdzenie, które właściwie wszyscy uważają za oczywiste. Jego dowód jest bardzo prosty, ale też dłuższy niż można spodziewać się. Zachęcamy

* Mało kto rozumiał wówczas te wtedy nowatorskie prace. Wśród tych, którzy je docenili był A.Cauchy.

do uważnego przyjrzenia mu się i ewentualnego skrócenia, jeśli się da.

Twierdzenie 7.8 (o monotoniczności różnowartościowej funkcji ciągłej)
Jeżeli f jest różnowartościową funkcją ciągłą określoną na przedziale P, to f jest funkcją ściśle monotoniczną.

Dowód. Wykażemy najpierw, że jeśli $x, y, z \in P$ oraz $x < y < z$, to $f(y)$ leży między punktami $f(x)$ i $f(z)$. Są dwie możliwości $f(x) < f(z)$ i $f(x) > f(z)$. Drugą możliwość można sprowadzić do pierwszej przez zastąpienie funkcji f funkcją przeciwną $-f$. Wystarczy więc zająć się pierwszą. Jeśli $f(y)$ nie leży między $f(x)$ i $f(z)$, to albo $f(y) < f(x)$, albo $f(z) < f(y)$. W pierwszym przypadku, na mocy twierdzenia o przyjmowaniu wartości pośrednich, istnieje punkt x' leżący między y i z, taki że $f(x) = f(x')$. Przeczy to różnowartościowości funkcji f. W drugim przypadku między x i y znajduje się punkt z', taki że $f(z) = f(z')$, co znów przeczy różnowartościowości funkcji f.
Teraz przejdziemy do właściwego dowodu. Załóżmy, że dla pewnych punktów r, s przedziału P zachodzą nierówności $r < s$ oraz $f(r) < f(s)$. Udowodnimy, że jeśli $u < v$, to również $f(u) < f(v)$ Z tego co już udowodniliśmy wynika, że jeśli $u < r$, to $f(r) < f(u)$ (dla dowodu rozważamy trójkę $x = u$, $y = r$, $z = s$), jeśli $r < u < s$, to $f(r) < f(u) < f(s)$ (tym razem $x = r$, $y = u$, $z = s$) i wreszcie jeśli $s < u$, to $f(s) < f(u)$. To samo dotyczy oczywiście $f(s)$. Punkty r, s dzielą przedział P na trzy podprzedziały. Jeśli u, v znajdują się w różnych podprzedziałach, to teza wynika z tego, co już stwierdziliśmy. Jeśli np. $u < v < r$, to ponieważ $f(u) < f(r)$ i $f(v)$ leży między $f(u)$ i $f(r)$, to $f(u) < f(v) < f(r)$. Pozostałe przypadki rozpatrujemy w identyczny sposób. Dowód został zakończony. ■

Udowodnimy teraz twierdzenie pozwalające stwierdzać ciągłość funkcji odwrotnej.

Twierdzenie 7.9 (o ciągłości funkcji odwrotnej)
Jeśli f jest funkcją ściśle monotoniczną określoną na pewnym przedziale P, to funkcja odwrotna f^{-1} przekształcająca obraz przedziału P na przedział P jest ciągła.

Dowód. Twierdzenie to wynika od razu z twierdzenia o ciągłości funkcji monotonicznej, które udowodniliśmy już wcześniej: funkcja monotoniczna, której obraz jest przedziałem jest ciągła i tego, że funkcja odwrotna do funkcji monotonicznej jest monotoniczna. Dowód został zakończony. ■

Z tego twierdzenia wynikają udowodnione już poprzednio twierdzenia o ciągłości logarytmu, funkcji arcsin i funkcji arctan, pierwiastków jako funkcji odwrotnych do

funkcji wykładniczej, funkcji sinus ograniczonej do przedziału $[-\frac{\pi}{2}, \frac{\pi}{2}]$, funkcji tangens ograniczonej do przedziału $(-\frac{\pi}{2}, \frac{\pi}{2})$ i funkcji potęgowych ograniczonych w razie potrzeby do zbioru liczb nieujemnych.
Zauważmy, że w twierdzeniu tym nie występuje założenie ciągłości funkcji f ! Ono nie jest potrzebne, zamiast niego występuje monotoniczność wyjściowej funkcji. W sytuacji ogólnej, gdy dziedzina funkcji nie jest przedziałem funkcja odwrotna ciągła być nie musi.

Następne twierdzenie okaże się bardzo przydatne do znajdowania najmniejszych i największych wartości funkcji. Szczególnie duże znaczenie mieć ono będzie w przypadku funkcji rzeczywistych wielu zmiennych. Będziemy je stosować w przypadku funkcji jednej zmiennej rzeczywistej między innymi po to, by później, w przypadku większej liczby zmiennych, łatwiej można było prześledzić rozumowania wykorzystujące pozornie całkowicie abstrakcyjne twierdzenia.

Twierdzenie 7.10 (Weierstrassa o przyjmowaniu kresów)
Załóżmy, że funkcja f jest ciągła w każdym punkcie przedziału domkniętego $[a,b]$. Wtedy w przedziale [a,b] znajdują się takie punkty p,q, że dla każdego punktu $x \in [a,b]$ zachodzi nierówność $f(p) \le f(x) \le f(q)$, tzn. $f(p)$ jest najmniejszą wartością funkcji f na przedziale [a,b], zaś $f(q)$ jest największą wartością funkcji f.

Dowód. Niech M będzie kresem górnym funkcji f na przedziale $[a,b]$. Istnieje ciąg (x_n) punktów przedziału $[a,b]$, taki że $\lim\limits_{n\to\infty} f(x_n) = M$. Z twierdzenia Bolzano – Weierstrassa wynika, że z ciągu (x_n) można wybrać podciąg zbieżny (x_{k_n}). Niech $q = \lim\limits_{n\to\infty} x_{k_n}$. Ponieważ dla każdej liczby naturalnej n zachodzi nierówność $a \le x_{k_n} \le b$, więc w granicy otrzymujemy $a \le q \le b$. Funkcja f jest ciągła w każdym punkcie przedziału $[a,b]$, w szczególności w punkcie q. Wynika stąd, że $f(q) = \lim\limits_{n\to\infty} f(x_{k_n}) = \lim\limits_{n\to\infty} f(x_n) = M$. Wykazaliśmy, więc że $\sup f = M = f(q)$, co oznacza, że $f(q)$ jest największą wartością funkcji f na przedziale $[a,b]$. Istnienie punktu, w którym funkcja f przyjmuje swą najmniejszą wartość, wnioskujemy stosując twierdzenie o wartości największej do funkcji $-f$. Dowód został zakończony. ■

Twierdzenie 7.11 (Cantora-Heine'go o jednostajnej ciągłości)
Jeśli funkcja f jest ciągła w każdym punkcie przedziału domkniętego $[a,b]$, to jest ona ciągła jednostajnie na tym przedziale.

Dowód. Załóżmy, że twierdzenie nie jest prawdziwe. Istnieje wtedy liczba $\varepsilon > 0$, taka że dla każdej liczby $\delta > 0$ istnieją takie liczby $x,y \in [a,b]$, że $|x-y| < \delta$ i jednocześnie $|f(x) - f(y)| \ge \varepsilon$. Niech x_n, y_n będą takimi liczbami z przedziału

$[a,b]$, że $|x_n - y_n| < \delta = \frac{1}{n}$ i jednocześnie $|f(x) - f(y)| \geq \varepsilon$. Z twierdzenia Bolzano – Weierstrassa wynika, że z ciągu (x_n) można wybrać podciąg zbieżny (x_{k_n}). Oznaczmy jego granicę przez g. Mamy więc $g = \lim\limits_{n\to\infty} x_{k_n}$, a ponieważ $|x_n - y_n| < \frac{1}{n}$, więc również $g = \lim\limits_{n\to\infty} y_{k_n}$. Oczywiście $g \in [a,b]$. Wobec tego funkcja f jest ciągła w punkcie g, zatem $f(g) = \lim\limits_{n\to\infty} f(x_{k_n}) = \lim\limits_{n\to\infty} f(y_{k_n})$, wbrew temu, że $|f(x_{k_n}) - f(y_{k_n})| \geq \varepsilon > 0$. Dowód został zakończony. ■

Twierdzenia o przyjmowaniu wartości pośrednich, o przyjmowaniu kresów i o jednostajnej ciągłości wyrażają najważniejsze własności funkcji ciągłej określonej na przedziale. W pierwszym przypadku jest to przedział absolutnie dowolny, w dwóch następnych domknięty i ograniczony. W dowodach dwóch ostatnich twierdzeń korzystaliśmy z twierdzenia Bolzano–Weierstrassa. Twierdzenia te można uogólnić nie zmieniając ich dowód zakładając nieco mniej o dziedzinie funkcji. Podamy definicję.

Definicja 7.12 (zbioru zwartego)
Zbiór $K \subseteq \mathbb{C}$ jest zwarty wtedy i tylko wtedy, gdy z każdego ciągu punktów zbioru K można wybrać podciąg zbieżny do punktu $p \in K$. ■

Zbiorem zwartym jest każdym przedział domknięty — to właściwie treść twierdzenia Bolzano–Weierstrassa. Przedział otwarty $(0,7)$ nie jest zwarty, jeśli bowiem $a_n = 7 - \frac{1}{n}$, to $a_n \in (0,7)$ i $\lim\limits_{n\to\infty} a_n = 7 \notin (0,7)$. Wynika stąd, że każdy podciąg tego ciągu jest zbieżny do $7 \notin K$. Czytelnik bez trudu stwierdzi, że każdy zbiór skończony jest zwarty, że suma **skończenie** wielu przedziałów domkniętych jest zwarta.

Przykład 7.2 Zbiór Cantora jest zwarty, bo jest ograniczony, więc z każdego ciągu punktów ze zbioru Cantora można wybrać podciąg zbieżny do granicy skończonej. Ta granica nie może znajdować się poza zbiorem Cantora, bo jego dopełnienie $\mathbb{R} \setminus C$ do całej prostej to suma przedziałów otwartych:

$$\mathbb{R} \setminus C = (-\infty, 0) \cup (1, +\infty) \cup \left(\tfrac{1}{3}, \tfrac{2}{3}\right) \cup \left(\tfrac{1}{9}, \tfrac{2}{9}\right) \cup \left(\tfrac{7}{9}, \tfrac{8}{9}\right) \cup \left(\tfrac{1}{27}, \tfrac{2}{27}\right) \cup \left(\tfrac{4}{27}, \tfrac{5}{27}\right) \cup \ldots.$$

Jeśli granica ciągu znajduje się w pewnym przedziale otwartym, to poza tym przedziałem otwartym jest jedynie skończenie wiele wyrazów tego ciągu. Granica każdego ciągu punktów zbioru Cantora musi być punktem zbioru Cantora. ■

Przykład 7.3 Prostokąt jest zbiorem zwartym. Wykażemy prawdziwość tego stwierdzenia w przypadku prostokąta, którego boki są równoległe do osi układu współrzędnych. Taki prostokąt można opisać za pomocą pary nierówności podwójnych: $P = \{(x,y)\colon\ \ a \leq x \leq b \ \text{ i } \ c \leq x \leq d\}$. Punkt (a,b) to lewy dolny wierzchołek tego prostokąta, punkt (c,d) to prawy górny. Wykażemy, że zbiór P jest zwarty. Niech $\big((x_n, y_n)\big)$ będzie dowolnym ciągiem punktów prostokąta P.

Ciąg $\big((x_n)\big)$ jest ograniczony, więc na mocy twierdzenia Bolzano–Weierstrassa można z niego wybrać podciąg (x_{n_k}) zbieżny do pewnej granicy p. Z tego, że $\lim\limits_{k\to\infty} x_{n_k} = p$ i z tego, ze dla każdego numeru n zachodzi nierówność $a \le x_n \le b$ wynika, że $a \le p \le b$. Ciąg $\big((y_{n_k})\big)$ jest ograniczony, więc można zeń wybrać podciąg $y_{n_{k_l}}$ zbieżny do pewnej liczby q. Oczywiście $c \le q \le d$. Ciąg $\big(x_{n_{k_l}}\big)$ jest zbieżny do p jako podciąg ciągu zbieżnego do p. Ponieważ $\lim\limits_{l\to\infty} x_{n_{k_l}} = p$ i $\lim\limits_{l\to\infty} y_{n_{k_l}} = q$, więc $\lim\limits_{l\to\infty}\big(x_{n_{k_l}}, y_{n_{k_l}}\big) = (p,q) \in P$. Zakończyliśmy dowód zwartości prostokąta P. Można w zasadzie w taki sam sposób wykazać, że dowolny prostokąt jest zwarty. Nie robimy tego od razu tylko dlatego, by nie zaciemniać dowodu formalnym opisem dowolnego prostokąta. ■

Przykład 7.4 Koło (domknięte) o środku (a,b) i promieniu $r > 0$, czyli zbiór $K = \{(x,y)\colon\quad (x-a)^2 + (y-b)^2 \le r^2\}$ jest zbiorem zwartym. Wykażemy to. Niech $\big((x_n, y_n)\big)$ będzie dowolnym ciągiem punktów koła K. Dla każdej liczby naturalnej n spełniona jest więc nierówność $(x_n - a)^2 + (y_n - b)^2 \le r^2$. Wobec tego $|x_n - a| \le r$, a z tego wynika, że ciąg $\big(x_n\big)$ jest ograniczona. Możemy więc wybrać zeń podciąg zbieżny $\big(x_{n_k}\big)$. Niech jego granicą będzie liczba p. Również ciąg $\big(y_{n_k}\big)$ jest ograniczony, więc z tego ciągu też można wybrać podciąg zbieżny $\big(y_{n_{k_l}}\big)$. Oznaczmy jego granicę przez q. Mamy więc $\lim\limits_{l\to\infty}\sqrt{(x_{n_{k_l}} - p)^2 + (y_{n_{k_l}} - q)^2} = 0$, a to oznacza, że $\lim\limits_{l\to\infty}(x_{n_{k_l}}, y_{n_{k_l}}) = (p,q)$. Ponieważ dla każdego numeru l spełniona jest nierówność $(x_{n_{k_l}} - a)^2 + (y_{n_{k_l}} - b)^2 \le r^2$, więc $(p-a)^2 + (q-b)^2 \le r^2$, tzn. $(p,q) \in K$. ■

Uwaga 7.13 (o zbiorach zwartych na płaszczyźnie)
Czytelnik bez trudu może uogólnić rozumowania z dwóch ostatnich przykładów i wykazać, że zbiór $C \subseteq \mathbb{R}^2$ jest zwarty, gdy spełnia następujące dwa warunki:

- **(i)** zbiór C jest ograniczony, tzn. istnieje liczba $c > 0$ taka, że odległość dowolnych dwóch punktów zbioru C nie przekracza liczby d;
- **(ii)** zbiór C jest domknięty, tzn. jeśli ciąg (p_n) punktów zbioru C ma granicę p, to również ta granica jest punktem zbioru C.

Dowód tego twierdzenia to nie jest trudny, ale nie będziemy go używać w tym roku w sytuacjach różnych od opisanych w przykładach poprzedzających tę uwagę. ■

Zadanie dla miłośników teorii mnogości Wykazać, że różnych podzbiorów zwartych prostej jest tyle, ile wszystkich liczb rzeczywistych. ■

Jest jasne, że w twierdzeniach Weierstrassa o osiąganiu kresów i w twierdzeniu Cantora–Heinego o jednostajnej ciągłości można zakładać, że dziedziną funkcji jest zbiór zwarty, niekoniecznie przedział domknięty. Dowody nie ulegają żadnym zmianom poza kosmetycznymi. Pokażemy teraz nieco inny dowód twierdzenia o ciągłości funkcji odwrotnej do danej funkcji ciągłej.

Twierdzenie 7.14 (**o ciągłości funkcji odwrotnej do różnowartościowej funkcji ciągłej określonej na zbiorze zwartym**)
Jeśli K jest zbiorem zwartym a $f\colon K \longrightarrow \mathbb{R}$ (albo $f\colon K \longrightarrow \mathbb{C}$) funkcją ciągłą różnowartościową, to funkcja odwrotna $f^{-1} : f(K) \longrightarrow K$ jest ciągła.

Dowód. Załóżmy, że tak nie jest. Wtedy istnieje taki ciąg (y_n) i $y \in K$, że $y_n \in f(K)$ dla każdej liczby naturalnej n, $\lim\limits_{n\to\infty} y_n = y$ i nie zachodzi równość $\lim\limits_{n\to\infty} f^{-1}(y_n) = f^{-1}(y)$ (granica może nie istnieć, a jeśli istnieje, to nie jest równa $f^{-1}(y)$). Niech $x_n = f^{-1}(y_n)$, $x = f^{-1}(y)$. Ponieważ x nie jest granicą ciągu (x_n), więc istnieje taka liczba $\varepsilon > 0$ i taki ściśle rosnący ciąg (n_k) liczb naturalnych, że $|x_{n_k} - x| \geq \varepsilon$ dla $k = 1, 2, 3, \ldots$. Z definicji zwartości wynika, że z ciągu (x_{n_k}) można wybrać podciąg zbieżny $(x_{n_{k_l}})$. Niech $\tilde{x} = \lim\limits_{l\to\infty} x_{n_{k_l}}$. Oczywiście $\tilde{x} \in K$. Ponieważ $|x_{n_{k_l}} - x| \geq \varepsilon$ dla każdego l, więc $|\tilde{x} - x| \geq \varepsilon > 0$. Ponieważ funkcja f jest ciągła w punkcie $\tilde{x}$, więc $f(\tilde{x}) = \lim\limits_{l\to\infty} f(x_{n_{k_l}}) = \lim\limits_{l\to\infty} y_{n_{k_l}} = \lim\limits_{n\to\infty} y_n = y = f(x)$, zatem $\tilde{x} = x$, wbrew wykazanej nierówności $|\tilde{x} - x| \geq \varepsilon > 0$. ■

Czytelnik z pewnością zauważył, że w tej wersji twierdzenia o ciągłości funkcji odwrotnej nie ma w ogóle mowy o monotoniczności. Dotyczy to też dowodu. To twierdzenie stosuje się również do zwartych podzbiorów płaszczyzny! Gdy K jest np. kołem, o monotoniczności w ogóle nie może być mowy, bo w zbiorze punktów płaszczyzny (liczb zespolonych) nierówność zdefiniowana nie została!

Przykład 7.5 Funkcja odwrotna do funkcji lipschitzowskiej nie musi być jednostajnie ciągła. Funkcja $\sqrt{x}$ spełnia na półprostej $[1, \infty)$ warunek Lipschitza ze stałą $\frac{1}{2}$, bo jeśli $1 \leq x < y$, to $\sqrt{y} - \sqrt{x} = \frac{y-x}{\sqrt{y}+\sqrt{x}} < \frac{y-x}{2}$. Przekształca ona półprostą $[1, \infty)$ na siebie. Funkcja odwrotna do niej, to x^2, która jak to wcześniej wykazaliśmy, nie jest jednostajnie ciągła na tej półprostej, w rzeczywistości na żadnej półprostej. ■

Przykład 7.6 Każdy rzeczywisty wielomian parzystego stopnia o dodatnim współczynniku kierującym przyjmuje najmniejszą wartość. Niech n będzie stopniem wielomianu. Jeśli $n = 0$, to wielomian jest funkcją stałą, więc każda jego wartość jest największa (i jednocześnie najmniejsza), więc dowodzić nie ma czego. Załóżmy, że $n > 0$ jest liczbą parzystą i że $a_n > 0$. Załóżmy, że $|x| > 1 + \frac{2|a_0|+|a_1|+|a_2|+\cdots+|a_{n-1}|}{|a_n|} =: m$. Wtedy, podobnie jak w dowodzie lematu o współczynnikach wielomianowej funkcji zerowej, mamy

$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \geq a_nx^n - |a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1}| \geq$
$\geq |x^{n-1}| \cdot \big[a_n \cdot |x| - \big(|a_0| + |a_1| + |a_2| + \cdots + |a_{n-1}|\big)\big] > |a_0| = |w(0)| \geq w(0)\,.$

Na przedziale domkniętym $[-m, m]$ funkcja ciągła w przyjmuje swą najmniejszą wartość. Załóżmy, że w punkcie $p \in [-m, m]$. Mamy więc $w(p) \leq w(x)$ dla każdego $x \in [-m, m]$. Jeśli $|x| > m$, to $w(x) > w(0) \geq w(p)$. Wykazaliśmy, że wielomian w przyjmuje swą najmniejszą (na całej prostej) wartość w punkcie p. ■

Przykład 7.7 Niech w będzie wielomianem o współczynnikach zespolonych i niech $\deg w \geq 1$. Wykażemy, że funkcja przypisująca liczbie z liczbę $|w(z)|$ ma najmniejszą wartość w pewnym punkcie płaszczyzny $\mathbb{C}$.

Załóżmy, że $|z| > 1 + \frac{2|a_0|+|a_1|+|a_2|+\cdots+|a_{n-1}|}{|a_n|} =: m$. Wtedy, podobnie jak w poprzednim przykładzie, mamy
$|a_0 + a_1z + a_2z^2 + \cdots + a_nz^n| \geq |a_nz^n| - |a_0 + a_1z + a_2z^2 + \cdots + a_{n-1}z^{n-1}| \geq$
$\geq |z^{n-1}| \cdot \big[|a_n| \cdot |z| - \big(|a_0| + |a_1| + |a_2| + \cdots + |a_{n-1}|\big)\big] > |a_0| = |w(0)|\,.$

Koło o środku w punkcie 0 jest zbiorem zwartym, więc funkcja ciągła $|w|$ przyjmuje swą najmniejszą wartość w pewnym punkcie p. Nierówność $|w(p)| \leq |w(z)|$ zachodzi zatem dla każdego z, dla którego $|z| \leq m$. Jeśli $|z| > m$, to $|w(z)| > |w(0)| \geq |w(p)|$, a to oznacza, że liczba $|w(p)|$ jest najmniejszą wartością funkcji $|w|$ nie tylko w kole o promieniu m i środku w punkcie 0, ale też w całej płaszczyźnie. ■

Twierdzenie 7.15 (Zasadnicze twierdzenie algebry)
Każdy wielomian o współczynnikach zespolonych, stopnia większego (ostro) od 0, ma co najmniej jeden pierwiastek zespolony.

Dowód. Niech $|w(z_0)|$ będzie najmniejszą wartością funkcji $z \mapsto |w(z)|$ — jej istnienie wykazaliśmy w poprzednim przykładzie. Wykażemy, że $w(z_0) = 0$. Przyjmijmy, że $z = z_0 + h$. Wtedy piszemy $w(z) = w(z_0 + h) = b_0 + b_1h + b_2h^2 + \cdots + b_nh^n$, gdzie $b_0 = =a_0 + a_1z_0 + \cdots + a_nz_0^n = w\,(z_0)$, $b_1 = a_1 + 2a_2z_0 + \cdots + na_nz_0^{n-1} = =w'\,(z_0)$, …, $b_n = a_n = \frac{1}{n!}w^{(n)}(z_0)$. Ponieważ stopień wielomianu równy jest n, więc $0 \neq a_n = b_n$. Niech $m \geq 1$ będzie najmniejszą taką liczbą, że $b_m \neq 0$. Załóżmy, że $w(z_0) \neq 0$.

Wtedy można napisać $w(z_0) = b_0 = |b_0| \cdot e^{i\varphi}$ dla pewnego $\varphi \in \mathbb{R}$. Mamy dalej $|w(z)| = |b_0 + b_mh^m + b_{m+1}h^{m+1} + \cdots + b_nh^n|$. Niech $\varrho < 1$ będzie liczbą dodatnią mniejszą niż $\frac{1}{2}|b_0|$ i niech $h = \varrho \cdot e^{i\frac{\varphi+\pi}{m}}$. Wtedy
$\big|b_0 + b_mh^m\big| = \big||b_0| \cdot e^{i\varphi} + \varrho^m e^{i(\varphi+\pi)}\big| = \big||b_0|e^{i\varphi} - \varrho^m e^{i\varphi}\big| = \big||b_0| - \varrho^m\big|\big|e^{i\varphi}\big| = |b_0| - \varrho^m\,.$
Załóżmy dodatkowo, że $\varrho\big(|b_{m+1}| + |b_{m+2}| + \cdots + |b_n|\big) < \frac{1}{2}$. Mamy wtedy
$|w(z)| = \big|b_0 + b_mh^m + b_{m+1}h^{m+1} + \cdots + b_nh^n\big| \leq$
$\leq \big|b_0 + b_mh^m\big| + \big|b_{m+1}h^{m+1} + \cdots + b_nh^n\big| = |b_0| - \varrho^m + \big|b_{m+1}h^{m+1} + \cdots + b_nh^n\big| \leq$

$\leq |b_0| - \varrho^m + \big(|b_{m+1}||h|^{m+1} + \cdots + |b_n||h|^n\big) \leq |b_0| - \varrho^m + |h|^{m+1}\big(|b_{m+1}| + \cdots + |b_n|\big) =$
$= |b_0| - \varrho^m + \varrho^{m+1}\big(|b_{m+1}| + \cdots + |b_n|\big) \leq |b_0| - \varrho^m + \frac{1}{2}\varrho^m = |b_0| - \frac{1}{2}\varrho^m < |b_0|$.
Okazało się, że wbrew założeniu liczba $|w(z_0)| = |b_0|$ *nie* jest najmniejszą wartością funkcji $|w|$. To kończy dowód tego, że $w(z_0) = 0$. Twierdzenie zostało więc wykazane. ■

Ważną klasę funkcji stanowią tzw. funkcje wypukłe. Przypomnijmy, że zbiór nazywany jest wypukłym wtedy i tylko wtedy, gdy wraz z każdymi dwoma punktami zawiera odcinek, który je łączy. Zbiorami wypukłymi są proste, płaszczyzny, cała przestrzeń trójwymiarowa, koło (ale nie okrąg), kula (ale nie jej powierzchnia zwana sferą), kwadrat (ale nie jego brzeg), trójkąt (ale nie jego brzeg). Czytelnicy zapewne pamiętają ze szkoły średniej, że wielokąt jest wypukły, jeśli jego kąty wewnętrzne są mniejsze niż $180°$, czyli π radianów. Jest jasne, że jedynymi podzbiorami wypukłymi prostej są przedziały, ewentualnie zdegenerowane do punktu. Mogą to być przedziały otwarte, domknięte, otwarto-domknięte, domknięto-otwarte, skończone lub nieskończone.

Definicja funkcji wypukłej

Funkcję f określoną na zbiorze wypukłym P nazywamy wypukłą, jeśli dla dowolnych punktów $x, y \in P$ i dowolnej liczby $t \in (0,1)$ zachodzi nierówność

$$f\big(tx + (1-t)y\big) \leq tf(x) + (1-t)f(y)\,.^*$$

Jeżeli nierówność ta jest ostra w przypadku $x \neq y$, to mówimy, że funkcja jest ściśle wypukła. Jeśli funkcja $-f$ jest wypukła, to mówimy, że funkcja f jest wklęsła, jeśli funkcja $-f$ jest ściśle wypukła, to funkcja f nazywana jest ściśle wklęsłą. ■

Przykład 7.8 Jeśli $f(x) = ax + b$, to funkcja f jest jednocześnie wypukła i wklęsła, nie jest ściśle wypukła. Stwierdzenie to wynika natychmiast z definicji:

$$f\big(tx+(1-t)y\big) = a\big(tx+(1-t)y\big)+b = t\big(ax+b\big)+(1-t)\big(ay+b\big) = tf(x)+(1-t)f(y)\,,$$

więc w przypadku funkcji liniowej nierówność występująca w definicji funkcji wypukłej staje się równością. ■

Przykład 7.9 Jeśli $f(x) = x^2$, to f jest funkcją ściśle wypukłą na całej prostej. Uzasadnimy to stwierdzenie. Dla $0 < t < 1$ mamy

$$tf(x) + (1-t)f(y) - f\big(tx + (1-t)y\big) = tx^2 + (1-t)y^2 - \big(tx + (1-t)y\big)^2 =$$
$$= t(1-t)(x-y)^2 \geq 0\,,$$

przy czym równość zachodzi wtedy i tylko wtedy, gdy $x = y$. ■

Przykład 7.10 Funkcja $f(x) = \sqrt{x}$ jest ściśle wklęsła — wynika to łatwo ze ścisłej

*Definicję tę stosuje się w niezmienionej formie również w przypadku funkcji wielu zmiennych.

wypukłości funkcji kwadratowej: nierówność $\sqrt{tx+(1-t)y} > t\sqrt{x}+(1-t)\sqrt{y}$ jest równoważna nierówności
$(tu+(1-t)v)^2 < tu^2+(1-t)v^2$, gdzie $u=\sqrt{x}$, $v=\sqrt{y}$. ■

Przed podaniem następnych przykładów skomentujemy definicję funkcji wypukłej i podamy kryterium pozwalające stwierdzać wypukłość niektórych funkcji. Funkcja jest wypukła jeśli połączywszy dwa punkty jej wykresu otrzymujemy odcinek, którego wszystkie punkty leżą nad wykresem funkcji lub na jej wykresie. Funkcja jest ściśle wypukła, jeśli wszystkie punkty wewnętrzne odcinka łączącego dwa punkty wykresu leżą nad wykresem funkcji. Jest tak dlatego, że w przypadku $0<t<1$, $x<y$ zachodzi nierówność $x<tx+(1-t)y<y$. W przykładzie pierwszym pokazaliśmy, że punkt $\big(tx+(1-t)y,\, tf(x)+(1-t)f(y)\big)$ leży na wykresie funkcji liniowej, której wykres przechodzi przez punkty $\big(x,f(x)\big)$ oraz $\big(y,f(y)\big)$, przyjmujemy $a=\frac{f(y)-f(x)}{y-x}$ oraz $b=f(x)$. Nierówność $f\big(tx+(1-t)y\big)\le tf(x)+(1-t)f(y)$, która występuje w definicji funkcji wypukłej, to stwierdzenie, że punkt $\big(tx+(1-t)y, f\left(tx+(1-t)y\right)\big)$ znajduje się pod punktem $\big(tx+(1-t)y, tf(x)+(1-t)f(y)\big)$. Oznacza to, funkcja jest wypukła wtedy i tylko wtedy, gdy zbiór punktów znajdujących się nad jej wykresem jest wypukły.

Twierdzenie 7.16 (o wypukłości funkcji ciągłej)
Funkcja f ciągła w każdym punkcie zbioru wypukłego P jest wypukła wtedy i tylko wtedy, gdy dla dowolnych $x,y\in P$ zachodzi nierówność $f\left(\frac{x+y}{2}\right)\le\frac{f(x)+f(y)}{2}$, ściśle wypukła, gdy ta nierówność jest ostra w każdym przypadku, w którym $x\ne y$.

Dowód. Jeśli f jest wypukła, to przyjmując w definicji wypukłości $t=\frac12$ otrzymujemy warunek podany w tym twierdzeniu, co kończy dowód konieczności tego warunku. Zajmiemy się teraz dowodem w „drugą” stronę.
Niech x,y będą dowolnymi punktami zbioru P. Mamy $f\left(\frac{x+y}{2}\right)\le\frac{f(x)+f(y)}{2}$. Ponieważ nierówność ta zachodzi dla dowolnych punktów x,y zbioru P, więc możemy zastąpić punkt y środkiem odcinka łączącego punkty x,y. Mamy $\frac12\left(x+\frac{x+y}{2}\right)=\frac34x+\frac14y$. Wobec tego mamy też

$$f\left(\tfrac34x+\tfrac14y\right)\le\tfrac12\left(f(x)+f\left(\tfrac{x+y}{2}\right)\right)\le\tfrac12\left(f(x)+\tfrac{f(x)+f(y)}{2}\right)=\tfrac34f(x)+\tfrac14f(y).$$

Wykazaliśmy więc, że nierówność definiująca wypukłość ma miejsce, gdy $t=\frac34$. Stosując to samo rozumowanie do punktów $\frac{x+y}{2}$ oraz y otrzymujemy nierówność $f\big(\frac14x+\frac34y\big)\le\frac14f(x)+\frac34f(y)$, a więc nierówność z definicji wypukłości w przypadku $t=\frac14$. Rozważając kolejno pary punktów x i $\frac34x+\frac14y$, $\frac34x+\frac14y$ i $\frac12(x+y)$, $\frac12(x+y)$ i $\frac14x+\frac34y$ oraz $\frac14x+\frac34y$ i y otrzymujemy nierówność kolejno dla $t=\frac78$, $t=\frac58$, $t=\frac38$ i $t=\frac18$. Otrzymaliśmy nierówność dla 7 wartości t: $\frac18$, $\frac28$, $\frac38$, $\frac48$, $\frac58$, $\frac68$,

$\frac{7}{8}$. W taki sam sposób możemy otrzymać nierówność w przypadku $t=\frac{k}{16}$, potem w przypadku $t=\frac{k}{32}$ itd.

Teraz skorzystamy z ciągłości funkcji f. Każda liczba $t\in(0,1)$ jest granicą ciągu (t_n) liczb postaci $\frac{k}{2^m}\in(0,1)$. Dla tych liczb nierówność jest już udowodniona.Mamy więc $f\left(t_nx+(1-t_n)y\right)\le t_nf(x)+(1-t_n)f(y)$. Przechodząc do granicy (wolno, bo f jest ciągła w każdym punkcie, w szczególności w $tx+(1-t)y$) otrzymujemy $f\left(tx+(1-t)y\right)\le tf(x)+(1-t)f(y)$, a to kończy dowód wypukłości funkcji f.

Należy jeszcze wykazać, że w przypadku ostrych nierówności funkcja f jest ściśle wypukła. Załóżmy, że tak nie jest. Wtedy istnieją takie liczby $x,y\in P$ oraz $t\in(0,1)$, że $x<y$ i $f(tx+(1-t)y)=tf(x)+(1-t)f(y)$. Załóżmy, że $0<s<t<1$. Wtedy*

$$tf(x)+(1-t)f(y)=f(tx+(1-t)y)=f\big(\tfrac{t-s}{1-s}x+\tfrac{1-t}{1-s}(sx+(1-s)y)\big)\le$$
$$\le\tfrac{t-s}{1-s}f(x)+\tfrac{1-t}{1-s}f(sx+(1-s)y)\le\tfrac{t-s}{1-s}f(x)+\tfrac{1-t}{1-s}\big(sf(x+(1-s)f(y))\big)=$$
$$=tf(x)+(1-t)f(y).$$

Wobec tego, że ten ciąg nierówności zaczyna się i kończy tym samym wyrażeniem, wszystkie nierówności są równościami, w tym $f(sx+(1-s)y)=sf(x)+(1-s)f(x)$, a to przeczy założeniu, bo oczywiście s może być liczbą postaci $\frac{k}{2^m}$. ■

Ostatni fragment tego dowodu może wyglądać nieco sztucznie, ale stanie się jaśniejszy po zapoznaniu się z twierdzeniem charakteryzującym funkcje wypukłe. W tej chwili wypada stwierdzić jedynie, że jeśli trzy punkty leżące na wykresie funkcji wypukłej leżą na jednej prostej, to wykres tej funkcji zawiera najmniejszy odcinek domkniętym, który je zawiera, a ostatni fragment dowodu w istocie rzeczy to pokazuje. By to dobrze zrozumieć trzeba pojąć, że jeśli $0<s<t<1$, to punkt $tx+(1-t)y$ leży bliżej punktu x niż punkt $sx+(1-s)y$, następnie narysować sobie to wszystko biorąc pod uwagę to, że żaden punkt wykresu funkcji wypukłej nie może się znaleźć nad odcinkiem łączącym dwa punkty tego wykresu.

Bez założenia ciągłości powyższe twierdzenie nie jest prawdziwe, ale przykłady o tym świadczące są bardzo nienaturalne — wymagają użycia pewnika wyboru, o którym coś zapewne studenci usłyszą na wstępie do matematyki.

Przykład 7.11 Funkcja wykładnicza o podstawie dodatniej i różnej od 1 jest ściśle wypukła. Wykażemy, że ma miejsce nierówność $a^{(x+y)/2}\le\frac{1}{2}\left(a^x+a^y\right)$, przy czym staje się ona równością jedynie wtedy, gdy $x=y$, bowiem

$$a^x+a^y-2a^{(x+y)/2}=\left(a^{x/2}-a^{y/2}\right)^2.$$

Stąd teza wynika natychmiast. ■

* Mamy więc $x<tx+(1-t)y<sx+(1-s)y<y$.

Przykład 7.12 Funkcja ln jest ściśle wklęsła. Dla dowodu wystarczy wykazać, że $\ln\left(\frac{x+y}{2}\right) \geq \frac{1}{2}(\ln x + \ln y)$ oraz że równość ma miejsce wtedy i tylko wtedy, gdy $x = y$. Nierówność ta jest równoważna następującej:

$$\left(e^{\ln x} + e^{\ln y}\right)/2 = \tfrac{x+y}{2} \geq e^{(\ln x + \ln y)/2},$$

która wynika natychmiast ze ścisłej wypukłości funkcji wykładniczej o podstawie e. ■

Przykład 7.13 Funkcja sinus jest ściśle wypukła na przedziale $[-\pi, 0]$ i ściśle wklęsła na przedziale $[0, \pi]$. Dla dowodu wystarczy wykazać, że

$$\text{jeśli } -\pi \leq x < y \leq 0\text{, to } \sin\tfrac{x+y}{2} < \tfrac{1}{2}(\sin x + \sin y)$$

oraz że

$$\text{jeśli } 0 \leq x < y \leq \pi\text{, to } \sin\tfrac{x+y}{2} > \tfrac{1}{2}(\sin x + \sin y).$$

Ponieważ $\sin(-x) = -\sin x$, więc wystarczy wykazać jedną z tych nierówności. Załóżmy, że $0 \leq x < y \leq \pi$. Mamy $\frac{1}{2}(\sin x + \sin y) = \sin\frac{x+y}{2}\cos\frac{x-y}{2} < \sin\frac{x+y}{2}$ — ostatnia nierówność wynika z tego, że $-\frac{\pi}{2} \leq \frac{x-y}{2} < 0$, więc $0 \leq \cos\frac{x-y}{2} < 1$. ■

Przykład 7.14 Funkcja tangens jest ściśle wypukła na przedziale $\left(-\frac{\pi}{2}, 0\right]$ i ściśle wklęsła na przedziale $\left[0, \frac{\pi}{2}\right)$. Podobnie jak w przypadku funkcji sinus wystarczy zająć się jednym z tych dwóch przedziałów. Załóżmy , że $0 \leq x < y < \frac{\pi}{2}$. Wykorzystamy znany wzór: $\operatorname{tg}\alpha - \operatorname{tg}\beta = \frac{\sin(\alpha-\beta)}{\cos\alpha\cos\beta}$. Mamy

$$\tfrac{1}{2}(\operatorname{tg}x + \operatorname{tg}y) - \operatorname{tg}\tfrac{x+y}{2} = \tfrac{1}{2}\left\{\left(\operatorname{tg}y - \operatorname{tg}\tfrac{x+y}{2}\right) - \left(\operatorname{tg}\tfrac{x+y}{2} - \operatorname{tg}x\right)\right\} =$$

$$= \frac{1}{2}\left\{\frac{\sin\frac{y-x}{2}}{\cos y\cos\frac{x+y}{2}} - \frac{\sin\frac{y-x}{2}}{\cos x\cos\frac{x+y}{2}}\right\} = \frac{\sin\frac{y-x}{2}(\cos x - \cos y)}{2\cos x\cos y\cos\frac{x+y}{2}} > 0$$

— ostatnia nierówność wynika z tego, że funkcja kosinus maleje na przedziale $[0, \frac{\pi}{2}]$. ■

Przykład 7.15 Niech $f(x) = |x|$. Wykażemy, że f jest funkcją wypukłą, ale nie ściśle. Tym razem skorzystamy bezpośrednio z definicji. Jeśli x, y są liczbami rzeczywistymi i $0 < t < 1$, to skorzystawszy z nierówności trójkąta otrzymujemy $|tx + (1-t)y| \leq t|x| + (1-t)|y|$, przy czym równość zachodzi wtedy i tylko wtedy, gdy $xy \geq 0$. ■

Przykład 7.16 Niech $f(x) = e^{|x|}$. Wykażemy, że funkcja ta jest ściśle wypukła. Ponieważ jest ciągła, więc można zajmować się jedynie przypadkiem $t = \frac{1}{2}$. Załóżmy, że $x \neq y$. Mamy w tej sytuacji $e^{|x+y|/2} \leq e^{(|x|+|y|)/2} \leq \frac{1}{2}\left(e^{|x|} + e^{|y|}\right)$, przy czym jeśli pierwsza nierówność staje się równością, to $xy \geq 0$ i wobec tego, że $x \neq y$, ma miejsce nierówność $|x| \neq |y|$ i wobec tego druga nierówność musi być ostra (funkcja wykładnicza jest ściśle wypukła). Dowód został zakończony. ■

Przykład 7.17 Funkcja $|x| + |x-1| + |x-2| + |x-3|$ jest wypukła jako suma czterech funkcji wypukłych. Nie jest ona ściśle wypukła, bo na przedziale $[1, 2]$ jest stała, zresztą na każdym z przedziałów $(-\infty,0]$, $[0,1]$, $[1,2]$, $[2,3]$, $[3,+\infty)$ jest liniowa, wykres tej funkcji składa się z trzech odcinków i dwu półprostych.

Przykład 7.18 Niech $f(x)=-\sqrt{|x|}$. Bez trudu sprawdzamy, że funkcja ta nie jest wypukła na całej prostej:

$$f\left(\tfrac{1}{2}\cdot(-1)+\tfrac{1}{2}\cdot 1\right)=f(0)>-1=\tfrac{1}{2}\left(f(-1)+f(1)\right).$$

Jest ona wypukła na każdej z półprostych $(-\infty,0]$, $[0,+\infty)$ — wynika to łatwo z tego, że — jak pokazaliśmy wcześniej — funkcja $\sqrt{\ }$ jest ściśle wklęsła. ■

Przykład 7.19 Wykażemy, że jeśli $a>1$ lub $a<0$, to funkcja x^a, zmiennej x, jest ściśle wypukła na półprostej $(0,\infty)$. Jeśli $0<a<1$, to funkcja x^2 jest ściśle wklęsła na półprostej $[0,\infty)$.

To rozumowanie pokazujemy tylko po to, by dowodząc później to samo stwierdzenie za pomocą rachunku różniczkowego studentom było łatwiej zrozumieć siłę metod analizy matematycznej. Można tego nie czytać, wystarczy sprawdzić jego długość, ewentualnie pomyśleć, czy przypadkiem nie można tego istotnie skrócić.

Wykażemy najpierw, że dla dowolnych liczb dodatnich $a\neq b$ i dowolnej liczby naturalnej dodatniej n zachodzi nierówność

$$\left(\tfrac{a^n+b^n}{2}\right)^{1/n}<\left(\tfrac{a^{n+1}+b^{n+1}}{2}\right)^{1/(n+1)}.$$

Dowód będzie indukcyjny. Dla $n=1$ mamy udowodnić, że $\frac{a+b}{2}<\left(\frac{a^2+b^2}{2}\right)^{1/2}$. Podnosząc ją stronami do kwadratu i mnożąc przez 4 otrzymujemy nierówność równoważną: $(a+b)^2<2(a^2+b^2)$, czyli $0<(a-b)^2$, więc prawdziwą.

Załóżmy, że nierówność $\left(\frac{a^n+b^n}{2}\right)^{1/n}<\left(\frac{a^{n+1}+b^{n+1}}{2}\right)^{1/(n+1)}$ zachodzi dla pewnej liczby naturalnej n. Założenie to jest równoważne (podnosimy obie strony do potęgi $n(n+1)$) temu, że $\left(\frac{a^n+b^n}{2}\right)^{n+1}<\left(\frac{a^{n+1}+b^{n+1}}{2}\right)^{n}$, tzn.

$$(a^n+b^n)^{n+1}<2(a^{n+1}+b^{n+1})^n. \tag{i}$$

Wykażemy, że $\left(\frac{a^{n+1}+b^{n+1}}{2}\right)^{1/(n+1)}<\left(\frac{a^{n+2}+b^{n+2}}{2}\right)^{1/(n+2)}$ czyli, że

$$\left(a^{n+1}+b^{n+1}\right)^{n+2}<2\left(a^{n+2}+b^{n+2}\right)^{n+1}. \tag{ii}$$

Dla dowodu wystarczy udowodnić, że

$$\tfrac{(a^{n+1}+b^{n+1})^{n+2}}{(a^n+b^n)^{n+1}}<\tfrac{(a^{n+2}+b^{n+2})^{n+1}}{(a^{n+1}+b^{n+1})^n}. \tag{iii}$$

Wtedy nierówność (ii) otrzymamy jako iloczyn nierówności (i) — założenie indukcyjne i nierówności (iii). Ostatnia nierówność jest równoważna nierówności

$$(a^{n+1}+b^{n+1})^{2n+2}<(a^n+b^n)^{n+1}(a^{n+2}+b^{n+2})^{n+1},$$

a ta nierówności

$$(a^{n+1}+b^{n+1})^2<(a^n+b^n)(a^{n+2}+b^{n+2}),$$

czyli $2a^{n+1}b^{n+1}<a^nb^{n+2}+a^{n+2}b^n$. Ostatnią nierówność otrzymujemy mnożąc oczy-

wistą nierówność $2ab < a^2 + b^2$ przez $a^n b^n$. Zakończyliśmy dowód nierówności
$$\left(\tfrac{a^n+b^n}{2}\right)^{1/n} < \left(\tfrac{a^{n+1}+b^{n+1}}{2}\right)^{1/(n+1)}.$$
Z niej wynika, że jeśli $1 \le n < m$ są liczbami naturalnymi, to
$$\left(\tfrac{a^n+b^n}{2}\right)^{1/n} < \left(\tfrac{a^m+b^m}{2}\right)^{1/m}. \qquad (*)$$
Podstawiając $a = x^{1/n}$ i $b = y^{1/n}$ w (*) otrzymujemy $\left(\frac{x+y}{2}\right)^{1/n} < \left(\frac{x^{m/n}+y^{m/n}}{2}\right)^{1/m}$, czyli $\left(\frac{x+y}{2}\right)^{m/n} < \frac{x^{m/n}+y^{m/n}}{2}$, a to oznacza, że funkcja $x^{m/n}$ jest ściśle wypukła. Podstawiając $a = \sqrt[m]{x}$ i $b = \sqrt[m]{y}$ w (*) otrzymujemy $\left(\frac{x^{n/m}+y^{n/m}}{2}\right)^{1/n} < \left(\frac{x+y}{2}\right)^{1/m}$, czyli $\frac{x^{n/m}+y^{n/m}}{2} < \left(\frac{x+y}{2}\right)^{n/m}$, a to oznacza, że funkcja $x^{n/m}$ jest ściśle wklęsła.

Wiemy już więc,że funkcja potęgowa o wykładniku wymiernym dodatnim jest ściśle wklęsła, gdy wykładnik jest mniejszy niż 1 i ściśle wypukła, gdy wykładnik jest większy od 1.

Jeśli $a > 1$, to istnieje taki ciąg (w_n), że $a = \lim\limits_{n\to\infty} w_n$ i $w_n > 1$ dla każdej liczby naturalnej $n \ge 1$. Wynika stąd, że
$$\left(\tfrac{x+y}{2}\right)^a = \lim_{n\to\infty}\left(\tfrac{x+y}{2}\right)^{w_n} \le \lim_{n\to\infty}\tfrac{x^{w_n}+y^{w_n}}{2} = \tfrac{x^a+y^a}{2},$$
a to oznacza, że funkcja x^a jest wypukła.

Analogicznie dowodzimy wklęsłość funkcji x^a dla $a \in (0,1)$.

Teraz załóżmy, że $a < 0$ oraz $0 < x < y$. Wtedy
$$\left(\tfrac{x+y}{2}\right)^a = e^{a\ln[(x+y)/2]} < e^{[a(\ln x)+a(\ln y)]/2} < \tfrac12\left[e^{a\ln x} + e^{a\ln y}\right] = \tfrac12\left[x^a + y^a\right]$$
— pierwsza nierówność wynika z tego, że funkcja $a\ln x$ jest ściśle wypukła, bo $a < 0$ i z tego, że funkcja wykładnicza o podstawie e jest ściśle rosnąca, druga nierówność — z tego, że funkcja wykładnicza o podstawie e jest ściśle rosnąca.

Jak widać ostatni rozważany przypadek był bardzo łatwy, skorzystaliśmy po prostu z tego, że złożenie dwu funkcji ściśle wypukłych jest ściśle wypukłe, jeśli zewnętrzna funkcja jest ściśle monotoniczna. ■

Przykład 7.20 Cena biletu kolejowego w ustalonej klasie jest funkcją wklęsłą odległości na jaką jest wystawiany.* Uzasadnimy to tak: przyrost ceny biletu spowodowany wydłużeniem się odległości jaką zamierzamy przejechać o ustaloną wielkość jest tym mniejszy im dłuższy dystans zamierzamy przebyć. Zapiszemy to za pomocą symboli matematycznych. Niech $p(x)$ oznacza cenę biletu pozwalającego na przejechanie x kilometrów. Niech h oznacza dowolną liczbę dodatnią i niech $x < y$. Wtedy $p(x+h) - p(x) \ge p(y+h) - p(y)$. Wykażemy, że ten warunek, w przypadku funk-

* Zakładamy, że bilet może być wystawiony na dowolną odległość, co w rzeczywistości nie jest prawdą. W rzeczywistości funkcja ta nie jest wklęsła, bo dziedzina nie jest przedziałem, lecz składa się wyłącznie z liczb całkowitych i w dodatku funkcja jest przedziałami stała: w większości przypadków wydłużenie podróży o 1 km nie zmienia ceny biletu. My rozpatrujemy pewną idealizację sytuacji rzeczywistej.

cji ciągłej określonej na przedziale, jest równoważny wklęsłości funkcji. Zakładamy oczywiście, że nierówność ma miejsce dla dowolnych liczb x, y, h przy założeniu, że $h > 0$ i $x < y$. Niech $r < s$ i $x = r$, $h = \frac{1}{2}(s - r)$, $y = \frac{1}{2}(s + r)$.

Nierówność $p(x + h) - p(x) \geq p(y + h) - p(y)$ przepisać można w postaci $p\big(\frac{1}{2}(s + r)\big) - p(r) \geq p(s) - p\big(\frac{1}{2}(s + r)\big)$, czyli $p\big(\frac{1}{2}(s + r)\big) \geq \frac{1}{2}\big(p(r) + p(s)\big)$, co pociąga za sobą wklęsłość funkcji *ciągłej* p.

Teraz załóżmy, że funkcja p jest wklęsła. Niech $u < v < w$ będą trzema punktami dziedziny funkcji p. Mamy $v = \frac{w-v}{w-u}u + \frac{v-u}{w-u}w$ oraz $0 < \frac{w-v}{w-u} < 1$ i $\frac{w-v}{w-u} + \frac{v-u}{w-u} = 1$, zatem $p(v) \geq \frac{w-v}{w-u}p(u) + \frac{v-u}{w-u}p(w)$. Tę ostatnią nierówność możemy przepisać na trzy różne sposoby:
$\frac{p(v)-p(u)}{v-u} \geq \frac{p(w)-p(u)}{w-u}$, $\frac{p(u)-p(v)}{u-v} \geq \frac{p(w)-p(v)}{w-v}$ i $\frac{p(u)-p(w)}{u-w} \geq \frac{p(v)-p(w)}{v-w}$.
Stosując te nierówności wnioskujemy, że $\frac{p(x+h)-p(x)}{x+h-x} \geq \frac{p(y+h)-p(y)}{y+h-y}$ — jeśli np. $x < y < x+h$, to stosujemy najpierw nierówność trzecią: $\frac{p(x+h)-p(x)}{x+h-x} \geq \frac{p(x+h)-p(y)}{x+h-y}$, a potem — drugą: $\frac{p(x+h)-p(y)}{x+h-y} \geq \frac{p(y+h)-p(y)}{y+h-y}$. Dowód został zakończony. ■

Końcówka ostatniego przykładu wymaga wyjaśnienia. Udowodniliśmy tam, że w przypadku funkcji wklęsłej p i trzech punktów jej dziedziny $u < v < w$ zachodzą nierówności:
$\frac{p(v)-p(u)}{v-u} \geq \frac{p(w)-p(u)}{w-u}$, $\frac{p(u)-p(v)}{u-v} \geq \frac{p(w)-p(v)}{w-v}$ i $\frac{p(u)-p(w)}{u-w} \geq \frac{p(v)-p(w)}{v-w}$.
Każda z nich może być potraktowana jako formalna interpretacja stwierdzenia: *iloraz* $\frac{p(v)-p(u)}{v-u}$ *jest funkcją nierosnącą*, w pierwszym przypadku zmiennej v, w drugim zmiennej u, w trzecim chodzi o wyrażenie $\frac{p(u)-p(w)}{u-w}$ jako funkcję zmiennej u. Każde z tych trzech stwierdzeń jest równoważne wklęsłości funkcji p. Wyrażenie $\frac{p(u)-p(v)}{u-v}$ nazywane jest ilorazem różnicowym funkcji p. Pokazuje ono jaka była względna zmiana wartości funkcji p. Stwierdzenie, że funkcja jest wklęsła oznacza więc, że rośnie ona coraz wolniej. Analogicznie funkcja wypukła rośnie coraz szybciej. Rezultaty te są ważne, więc zapiszmy je raz jeszcze w formie twierdzenia tym razem sformułowanego w przypadku funkcji wypukłej.

Twierdzenie 7.17 (charakteryzujące funkcje wypukłe)
Niech f będzie funkcją określoną na zbiorze wypukłym P. Następujące warunki są równoważne:

(i) funkcja f jest wypukła;
(ii) jeśli $x < y < z$ są punktami zbioru P, to $\frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x}$;
(iii) jeśli $x < y < z$ są punktami zbioru P, to $\frac{f(x)-f(y)}{x-y} \leq \frac{f(z)-f(y)}{z-y}$;
(iv) jeśli $x < y < z$ są punktami zbioru P, to $\frac{f(x)-f(z)}{x-z} \leq \frac{f(y)-f(z)}{y-z}$. ■

W przypadku funkcji ściśle wypukłych nierówności występujące w warunkach (ii) – (iv) są ostre.

Twierdzenie to będziemy stosować później, gdy będziemy badać wypukłość funkcji za pomocą pochodnych. Zakończymy rozważania o funkcjach wypukłych nierównością Jensena. Ma ona ważne zastosowania, jest to dobre narzędzie do uzyskiwania różnych oszacowań. Ma ważne zastosowania w rachunku prawdopodobieństwa. Rozpoczniemy od średnich ważonych.

Definicja 7.18 (średniej ważonej)
Średnią ważoną liczb $x_1,\ x_2,\ x_3,\dots,\ x_n$ z wagami $p_1,\ p_2,\ p_3,\dots,\ p_n$ nazywamy liczbę

$$p_1x_1+p_2x_2+p_3x_3+\cdots+p_nx_n$$

pod warunkiem, że $0\le p_1\,,\ 0\le p_2\,,\ 0\le p_3\,,\dots,0\le p_n$ i $p_1+p_2+p_3+\cdots+p_n=1$. ■

W przypadku, gdy wagi są równe, więc równe $\frac{1}{n}$, średnia ważona zwana jest średnią arytmetyczną, a czasem po prostu średnią. Jeśli np. policzono średnie płace dla różnych grup ludności i mamy policzyć średnią płacę w kraju, to ze względu na to, że np. ministrów jest istotnie mniej niż pielęgniarek (przynajmniej w chwili pisania tego tekstu), to ich płaca zostanie uwzględniona z mniejszą wagą niż płaca pielęgniarek. W obu przypadkach wagą będzie iloraz liczby członków danej grupy przez liczbę wszystkich zatrudnionych w kraju. Inny przykład sytuacji, w której pojawia się średnia ważona, to próba przewidywania swej wygranej przez uczestnika gra hazardowej. Wie on, że za uzyskanie wyniku j otrzymuje on kwotę x_j (ta liczba może być ujemna, wtedy hazardzista płaci). Jeśli wynik j uzyskiwany jest z prawdopodobieństwem p_j, to należy spodziewać się wygranej $p_1x_1+p_2x_2+p_3x_3+\cdots+p_nx_n$, tzn. grając wielokrotnie średnio uzyskiwać będziemy kwotę $p_1x_1+p_2x_2+p_3x_3+\cdots+p_nx_n$. Przykłady można mnożyć, ale nie będziemy tego robić.

Twierdzenie 7.19 (Nierówność Jensena)
Jeśli funkcja f jest wypukła, to dla dowolnych jej argumentów $x_1,\ ,x_2,\ x_3,\dots,x_n$ i dowolnych wag $p_1,p_2,\ ,\dots,\ p_n$ zachodzi nierówność:

$$f(p_1x_1+p_2x_2+p_3x_3+\cdots+p_nx_n)\le p_1f(x_1)+p_2f(x_2)+p_3f(x_3)+\cdots+p_nf(x_n)\,.$$

Nierówność ta w przypadku funkcji ściśle wypukłej, dodatnich wag $p_1,p_2,\dots,p_n$ i przynajmniej dwóch różnych argumentów spośród $x_1\,,\ x_2\,,\ \dots,\ x_n$ jest ostra.

Dowód Dla $n=1$ musi być $p_1=1$ i nierówność staje się oczywistą równością. Dla $n=2$ mamy $p_2=1-p_1$ i nierówność jest tą nierównością, która występuje w definicji funkcji wypukłej. Załóżmy, że dowodzone twierdzenie jest prawdziwe dla wszystkich możliwych wyborów n argumentów funkcji f i n wag. Niech liczby x_1,

x_2, ..., x_n, x_{n+1} będą dowolnymi argumentami funkcji f, a $p_1, p_2, \ldots, p_n, p_{n+1}$ dowolnym układem $n+1$ wag, tj. liczb nieujemnych, których suma równa jest 1. Jeśli którakolwiek z wag jest równa 0, to nierówność z $n+1$ argumentami i $n+1$ wagami jest prawdziwa na mocy uczynionego założenia (argument odpowiadający zerowej wadze jest nieistotny), bo w nierówności faktycznie nie występuje. Załóżmy teraz, że wszystkie wagi $p_1, p_2, \ldots, p_n, p_{n+1}$ są dodatnie. Niech $p'_n = p_n + p_{n+1}$ i $x'_n = \frac{p_n x_n + p_{n+1} x_{n+1}}{p_n + p_{n+1}} = \frac{p_n}{p'_n} x_n + \frac{p_{n+1}}{p'_n} x_{n+1}$. Zachodzi równość

$$p_1x_1 + p_2x_2 + \ldots + p_nx_n + p_{n+1}x_{n+1} = p_1x_1 + p_2x_2 + \ldots + p_{n-1}x_{n-1} + p'_nx'_n .$$

Z założenia, które uczyniliśmy wynika, że

$$\begin{aligned}
&f(p_1x_1 + p_2x_2 + \ldots + p_{n-1}x_{n-1} + p'_nx'_n) \le \\
&\quad \le p_1f(x_1) + p_2f(x_2) + \ldots + p_{n-1}f(x_{n-1}) + p'_nf(x'_n) \le \\
&\qquad \le p_1f(x_1) + p_2f(x_2) + \ldots + p_{n-1}f(x_{n-1}) + p'_n\left(\tfrac{p_n}{p'_n}f(x_n) + \tfrac{p_{n+1}}{p'_n}f(x_{n+1})\right) = \\
&\qquad\qquad = p_1f(x_1) + p_2f(x_2) + \ldots + p_{n-1}f(x_{n-1}) + p_nf(x_n) + p_{n+1}f(x_{n+1})
\end{aligned}$$

— druga z tych nierówności jest bezpośrednim wnioskiem z wypukłości funkcji f. Zakończyliśmy indukcyjny dowód nierówności Jensena. ■

Pokażemy teraz jej najprostsze zastosowania. Rozpoczniemy od klasycznej nierówności między średnimi.

Twierdzenie 7.20 (Nierówność Cauchy'ego między klasycznymi średnimi)
Dla dowolnych liczb dodatnich $a_1, a_2, \ldots, a_n$ zachodzi:

$$\sqrt[n]{a_1a_2\ldots a_n} \le \tfrac{a_1+a_2+\cdots+a_n}{n} .$$

Nierówność ta staje się równością wtedy i tylko wtedy, gdy $a_1 = a_2 = \ldots = a_n$.

Dowód. Zastosujemy nierówność Jensena do funkcji wypukłej $-\ln$, której wypukłość wykazaliśmy wcześniej. Mamy

$$\begin{aligned}
&-\ln\tfrac{a_1+a_2+\cdots+a_n}{n} = -\ln\left(\tfrac{1}{n}a_1 + \tfrac{1}{n}a_2 + \cdots + \tfrac{1}{n}a_n\right) \le \\
&\quad \le \tfrac{1}{n}(-\ln)(a_1) + \tfrac{1}{n}(-\ln)(a_2) + \cdots + \tfrac{1}{n}(-\ln)(a_n) = -\tfrac{1}{n}(\ln a_1 + \ln a_2 + \ldots + \ln a_n) .
\end{aligned}$$

Stąd $\ln\frac{a_1+a_2+\cdots+a_n}{n} \ge \frac{1}{n}(\ln a_1 + \ln a_2 + \ldots + \ln a_n)$ i wobec tego

$$\frac{a_1+a_2+\cdots+a_n}{n} = e^{\ln\left((a_1+a_2+\cdots+a_n)/n\right)} \ge e^{(\ln a_1+\ln a_2+\ldots+\ln a_n)/n} = \sqrt[n]{a_1a_2\ldots a_n} .$$

Równość w tych nierównościach ma miejsce wtedy i tylko wtedy, gdy wszystkie liczby $a_1, a_2, \ldots, a_n$ są równe, bowiem funkcja $-\ln$ jest ściśle wypukła.Dowód został zakończony. ■

Z kilku dowodów nierówności o średniej arytmetycznej i geometrycznej znanych autorowi podany wyżej jest najkrótszy.

Twierdzenie 7.21 (Nierówność Höldera)

Dla dowolnych liczb nieujemnych $a_1, a_2, \ldots, a_n$, $b_1, b_2, \ldots, b_n$ i dowolnych liczb dodatnich p, q takich, że $\frac{1}{p} + \frac{1}{q} = 1$ zachodzi nierówność:

$$a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \le \left(a_1^p + a_2^p + \cdots + a_n^p\right)^{1/p} \left(b_1^q + b_2^q + \cdots + b_n^q\right)^{1/q}$$

Dowód. Z równości $\frac{1}{p} + \frac{1}{q} = 1$ wynika, że $p > 1$, $q > 1$ oraz $\frac{1}{q} = 1 - \frac{1}{p} = \frac{p-1}{p}$. Ponieważ $p > 1$, więc funkcja x^p jest ściśle wypukła, jak to wykazaliśmy wcześniej. Możemy więc zastosować nierówność Jensena do tej funkcji. Bez straty ogólności można założyć, że wszystkie liczby b_1, b_2,$\ldots, b_n$ są dodatnie, gdyby dla pewnego j było $b_j = 0$ wykazalibyśmy nierówność dla $n-1$ liczb $a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n$, $b_1, \ldots, b_{j-1}, b_{j+1}, \ldots, b_n$, a więc z tą samą lewą stroną a prawą być może mniejszą (gdy $a_j > 0$) niż docelowa. Przyjmijmy $p_j = \frac{b_j^{p/(p-1)}}{\sum b_i^{p/(p-1)}}$. Mamy wtedy

$$\frac{\left(\sum_{j=1}^{n} a_j b_j\right)^p}{\left(\sum_{j=1}^{n} b_j^q\right)^{p/q}} = \left(\frac{\sum_{j=1}^{n} \frac{a_j}{b_j^{1/(p-1)}} b_j^{p/(p-1)}}{\sum_{i=1}^{n} b_i^{p/(p-1)}}\right)^p \cdot \sum_{i=1}^{n} b_i^{p/(p-1)} \le$$

$$\le \frac{\sum_{j=1}^{n} \frac{a_j^p}{b_j^{p/(p-1)}} b_j^{p/(p-1)}}{\sum_{i=1}^{n} b_i^{p/(p-1)}} \cdot \sum_{i=1}^{n} b_i^{p/(p-1)} = \sum_{j=1}^{n} a_j^p,$$

a ta nierówność jest równoważna dowodzonej. Dla $p = q = 2$ otrzymujemy nierówność Schwarza, tzn. stwierdzenie, że iloczyn skalarny dwóch wektorów $(a_1, a_2, \ldots, a_n)$ i $(b_1, b_2, \ldots, b_n)$ jest nie większy niż iloczyn ich długości. ■

Przykład 7.21 Wykażemy, że spośród 5-kątów wpisanych w okrąg o promieniu 1 największy obwód ma pięciokąt foremny.

Niech $2\alpha_1$, $2\alpha_2$, $2\alpha_3$, $2\alpha_4$, $2\alpha_5$ będą kątami środkowymi opartymi na bokach pięciokąta.*. Wtedy bokami są liczby $2\sin\alpha_1$, $2\sin\alpha_2$, $2\sin\alpha_3$, $2\sin\alpha_4$, $2\sin\alpha_5$ — wynika to z definicji sinusa, wobec tego połowa obwodu pięciokąta równa jest $\sin\alpha_1 + \sin\alpha_2 + \sin\alpha_3 + \sin\alpha_4 + \sin\alpha_5$. Oczywiście spełnione są kolejne nierówności $0 < \alpha_1 < \pi$, $0 < \alpha_2 < \pi$, $0 < \alpha_3 < \pi$, $0 < \alpha_4 < \pi$, $0 < \alpha_5 < \pi$. Na przedziale $[0, \pi]$ sinus jest funkcją ściśle wklęsłą, więc możemy zastosować nierówność Jensena:

$$\sin\alpha_1 + \sin\alpha_2 + \sin\alpha_3 + \sin\alpha_4 + \sin\alpha_5 =$$
$$= 5\left(\tfrac{1}{5}\sin\alpha_1 + \tfrac{1}{5}\sin\alpha_2 + \tfrac{1}{5}\sin\alpha_3 + \tfrac{1}{5}\sin\alpha_4 + \tfrac{1}{5}\sin\alpha_5\right) \le$$
$$\le 5\sin\left(\tfrac{1}{5}\alpha_1 + \tfrac{1}{5}\alpha_2 + \tfrac{1}{5}\alpha_3 + \tfrac{1}{5}\alpha_4 + \tfrac{1}{5}\alpha_5\right) = 5\sin\tfrac{\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5}{5} = 5\sin\tfrac{\pi}{5}.$$

* Wierzchołek kąta jest środkiem okręgu, ramiona przechodzą przez końce boku.

Wielkość $5\sin\frac{\pi}{5}$ to połowa obwodu pięciokąta foremnego wpisanego w okrąg, więc twierdzenie jest udowodnione. Wypada dodać, że ponieważ funkcja sinus na przedziale $[0,\pi]$ jest ściśle wklęsła, więc pięciokąty nieforemne mają obwody mniejsze niż pięciokąt foremny wpisany w ten sam okrąg. W ten sam sposób można wykazać, że długość n–kąta wpisanego w okrąg jest nie większa niż długość n–kąta foremnego wpisanego w ten sam okrąg, a stąd i z równości $\lim\limits_{n\to\infty} 2n\sin\frac{\pi}{n} = 2\pi$ oraz nierówności $\sin\frac{\pi}{n} < \frac{\pi}{n}$ wynika, że kresem górnym łamanych wpisanych w okrąg o promieniu 1 jest liczba 2π. Ona jest więc długością tego okręgu. ■

7.01 Wykazać, że spośród pięciokątów opisanych na okręgu o promieniu 1 najmniejszy obwód ma pięciokąt foremny.

7.02 Wyjaśnić, czy istnieje taka funkcja $f\colon\mathbb{R}\longrightarrow\mathbb{R}$, że dla każdego $p\in\mathbb{R}$ zachodzi równość $\lim\limits_{x\to p} f(x)=\infty$

7.03 Niech c_1, c_2, ... będą kolejnymi cyframi rozwinięcia trójkowego liczby $x\in C$, gdzie C oznacza zbiór Cantora, przy czym w ciągu (c_n) nie występuje cyfra 1. Niech $f(x)$ będzie liczbą, której kolejnymi cyframi w układzie trójkowym są: $2-c_1$, $2-c_2$, ... Wykazać, że funkcja f jest dobrze zdefiniowana na całym zbiorze Cantora. Czy f jest ciągła? Czy f jest monotoniczna?

7.04 Dane są takie koła K_1, K_2, K_3, ..., że dla każdej liczby naturalnej n koła K_1, K_2,..., K_n można ułożyć w kwadracie Q tak, by nie miały wspólnych punktów wewnętrznych. Dowieść, że w kwadracie Q można ułożyć wszystkie koła K_1, K_2, K_3, ... tak, by nie miały wspólnych punktów wewnętrznych.

7.05 Dowieść, że jeśli zbiór $A\subseteq$ nie jest zwarty, to istnieje

- **a** funkcja ciągła $f: A\longrightarrow\mathbb{R}$, która nie jest ograniczona z góry;
- **b** funkcja ciągła $f: A\longrightarrow\mathbb{R}$, która nie jest ograniczona ani z góry ani z dołu;
- **c** funkcja ciągła $f: A\longrightarrow\mathbb{R}$, która nie jest jednostajnie ciągła.

7.06 Wykazać, że jeśli $a,b\in\mathbb{R}$ i funkcja $f:(a,b)\longrightarrow\mathbb{R}$ jest jednostajnie ciągła, to istnieje taka funkcja ciągła $\tilde{f}:[a,b]\longrightarrow\mathbb{R}$, że dla każdego $x\in(a,b)$ zachodzi równość $f(x)=\tilde{f}(x)$, tzn. funkcję jednostajnie ciągłą na przedziale otwartym możemy przedłużyć do funkcji ciągłej na przedziale domkniętym o tych samych końcach.

7.07 Dowieść, że dla każdego wielokąta wypukłego istnieje prosta, która dzieli jednocześnie obwód i pole tego wielokąta na połowy.

7.08 Niech $f:\mathbb{R}\longrightarrow\mathbb{R}$ będzie funkcją ciągłą i niech $f^n=\underbrace{f\circ f\circ f\circ\ldots\circ f}_{\text{n razy } f}$. Wykazać, że jeśli istnieje taka liczba $c\in(0,1)$, że $|f^n(x)-f^n(y)|\le c|x-y|$, to istnieje dokładnie jedna liczba rzeczywista p taka, że $f(p) = p$.

7.09 Wyjaśnić, czy istnieje jednostajnie ciągła funkcja $f : [0, \infty) \longrightarrow \mathbb{R}$, która przekształca półprostą $[0, \infty)$ *na* całą prostą. Czy taka funkcja $f : [0, \infty) \longrightarrow \mathbb{R}$ może być różnowartościowa?

7.10 Wykazać, że dla dowolnych liczb dodatnich a, b, c, z których przynajmniej dwie są różne, zachodzi nierówność

$$\left(\tfrac{a^2+b^2+c^2}{a+b+c}\right)^{a+b+c} > a^a b^b c^c > \left(\tfrac{a+b+c}{3}\right)^{a+b+c}.$$

7.11 Wykazać, że $(3x + 2y) \cdot \operatorname{tg} \frac{3x+2y}{5} \le 3x \cdot \operatorname{tg} x + 2y \cdot \operatorname{tg} y$ dla $0 \le x, y < \frac{\pi}{2}$ oraz wyjaśnić, kiedy zachodzi równość.

7.12 Wykazać, że jeśli $a, b > 0$, to $(2 - \sqrt{3})a^{2+\sqrt{3}} + (2 + \sqrt{3})b^{2-\sqrt{3}} \ge 4\sqrt[4]{ab}$, dla jakich a, b zachodzi równość?

7.13 Dowieść, że jeśli prosta ma trzy różne punkty wspólne z wykresem funkcji wypukłej f, to ma z nim wspólny odcinek i f nie jest funkcją ściśle wypukłą.

7.14 Wykazać, że dla dowolnych parametrów $a, b \in \mathbb{R}$ równanie $\operatorname{tg} x = ax + b$ ma co najwyżej trzy różne rozwiązania w przedziale $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.

7.15 Wykazać, że jeśli funkcja ściśle wypukła jest ciągła i *nie* jest monotoniczna, to ma wartość najmniejszą i ta najmniejsza wartość jest przyjmowana w dokładnie jednym punkcie, przy czym jest to punkt wewnętrzny dziedziny funkcji.

7.16 Wykazać, że jeśli funkcje f i g są wypukłe, funkcja g jest niemalejąca, to funkcja $g \circ f$ jest *wypukła*, jeśli natomiast g jest nierosnąca, to złożenie $g \circ f$ *może* być funkcją *wklęsłą*, *wypukłą* lub nawet mieć punkty przegięcia.

7.17 Wykazać, że jeśli funkcja f jest wypukła na każdym z przedziałów $[a, b]$ i $[b, c]$ oraz różniczkowalna w punkcie b, to jest wypukła na $[a, c]$. Podać przykład świadczący o tym, że bez założenia różniczkowalności teza nie jest prawdziwa.