Analiza 1, część piąta

Jest tu „trochę” przykładów, których na wykładzie nie było, ale które warte są obejrzenia. Niektóre dowody są przeprowadzone w nieco inny sposób, ale student nie jest zobowiązany do powtarzania tekstów z wykładu – ma podawać poprawne rozumowania wychodzące z tych samych założeń! Tekst jest długawy, ale może część z Was przynajmniej skorzysta z niego.
W razie znalezienia błędów proszę o informację, z góry przepraszam za pomyłki.

F U N K C J E C I Ą G Ł E

1. Definicja funkcji.

Jednym z najważniejszych pojęć w matematyce jest pojęcie funkcji. Przypomnimy definicję.

Definicja funkcji, wartości, obrazu, dziedziny i przeciwdziedziny

Przyporządkowanie f elementom zbioru A elementów zbioru B w taki sposób, że każdemu elementowi zbioru A przypisany jest dokładnie jeden element zbioru B nazywamy funkcją ze zbioru A w zbiór B. Jeśli a jest elementem zbioru A, symbolicznie $a \in A$, czyli argumentem funkcji f, to przypisany mu element zbioru B oznaczamy symbolem $f(a)$ i nazywamy wartością funkcji f w punkcie a lub obrazem punktu a.* Zbiór A nazywamy dziedziną funkcji f, zbiór B – przeciwdziedziną, zbiór $f(A)$ złożony ze wszystkich wartości funkcji f, czyli elementów zbioru B postaci $f(a)$, gdzie $a \in A$ nazywamy obrazem zbioru A (przez funkcję f) lub zbiorem wartości funkcji f. Jeśli f przekształca zbiór A w zbiór B, to piszemy $f\colon A \to B$. Jeśli zbiór $f(A)$ wartości funkcji f pokrywa się z przeciwdziedziną B funkcji f, to mówimy, że f przekształca zbiór A *na* zbiór B i piszemy czasem $f\colon A \xrightarrow{na} B$. ■

Przykładem funkcji jest ciąg: jest to funkcja określona np. na zbiorze $\mathbb{N} = \{0,\ 1,\ 2,\ \ldots\}$. Innym przykładem, dobrze znanym ze szkoły, jest funkcja liniowa: $f(x) = ax + b$, gdzie $a,\ b$ są ustalonymi liczbami rzeczywistymi, x jest elementem zbioru wszystkich liczb rzeczywistych $\mathbb{R}$, na którym funkcja f jest określona, $f(x)$ jest elementem przeciwdziedziny $\mathbb{R}$; jeśli $a \neq 0$, to funkcja f przekształca zbiór $\mathbb{R}$ *na* siebie; jeśli $a = 0$, to jedyną wartością funkcji f jest liczba b. Jeszcze innym przykładem jest funkcja kwadratowa: $f(x) = ax^2 + bx + c$, gdzie a, b, c są liczbami rzeczywistymi, przy czym $a \neq 0$, funkcja ta jest określona na zbiorze wszystkich liczb rzeczywistych $\mathbb{R}$, przeciwdziedziną jest również $\mathbb{R}$, zbiorem wartości jest półprosta $\left[\frac{4ac-b^2}{4a}, +\infty\right)$ w przypadku $a > 0$, zaś w przypadku $a < 0$ zbiorem wartości jest półprosta $\left(-\infty, \frac{4ac-b^2}{4a}\right]$. Inny przykład funkcji znany ze szkoły to permutacje zbioru n–elementowego, można je traktować jako funkcje przekształcające zbiór $\{1,\ 2,\ ,n\}$ na dany zbiór złożony z n elementów: mamy ustawić elementy danego zbioru w kolejności, pierwszy w tym ustawieniu element to wartość permutacji w punkcie 1, drugi – wartość w punkcie $2, \ldots, n$–ty – wartość w punkcie n. Zadanie na ile sposobów może 10 osób wsiąść do trzech wind to pytanie: ile jest funkcji ze zbioru 10–elementowego w zbiór trójelementowy (osobie przypisujemy windę, do której ta osoba wsiada). Przykłady można mnożyć, ale nie będziemy tego robić teraz. Na razie będziemy zajmować się funkcjami rzeczywistymi jednej zmiennej rzeczywistej, co oznacza, że wartościami funkcji będą liczby rzeczywiste i dziedziną funkcji będzie jakiś zbiór złożony z liczb rzeczywistych. W rzeczywistości dziedzinami będą albo przedziały, albo sumy skończenie wielu lub nieskończenie wielu przedziałów, np dziedziną funkcji

* Czasem będziemy mówić: „ f–obrazem”, choć to nie brzmi dobrze, ale czasem należy wyraźnie zaznaczyć o jaką funkcję chodzi.

tg jest zbiór złożony z tych wszystkich liczb rzeczywistych, które nie są postaci $(2n+1)\frac{\pi}{2}$, czyli jest to suma przedziałów postaci $\big(-(2n+1)\frac{\pi}{2},(2n+1)\frac{\pi}{2}\big)$, gdzie n oznacza dowolną liczbę całkowitą. w przypadku funkcji zdefiniowanej wzorem $f(x)=\dfrac{x^2}{(x-1)(x+2)}$ można powiedzieć, że jej dziedziną jest zbiór wszystkich liczb rzeczywistych z wyjątkiem -2 i 1, czyli zbiór $(-\infty,-2)\cup(-2,1)\cup(1,+\infty)$.

Z punktu widzenia formalnego dopóki nie powiemy na jakim zbiorze funkcja ma być zdefiniowana, to nie została ona określona. W szczególności z formalnego punktu widzenia zadania: znaleźć dziedzinę funkcji określonej wzorem $\dots$, nie mają sensu. Pytanie o dziedzinę należy traktować jako pytanie o maksymalny zbiór, na którym można zdefiniować funkcję w sposób zaproponowany przez autora zadania. Nawet przy takiej interpretacji mogą powstawać wątpliwości: np. czy funkcja określona wzorem $f(x)=\dfrac{x^2}{x}$ może tym wzorem być zdefiniowana na całej prostej, czy też w punkcie 0 tym akurat wzorem nie da się jej zdefiniować. Autorowi tego tekstu wydaje się, że specjaliści od tak formułowanych zadań w większości przypadków uznają, że ta definicja w punkcie 0 nie działa, ale nie wydaje mu się, by ten problem wart był dyskusji – można po prostu takich zadań nie dawać, a jeśli się je daje, to unikać wieloznaczności. Będziemy jednak mówić np. o funkcji $\dfrac{4x^2-13x-167}{x^3-4x+3}$, zakładając przy tym, że jej dziedziną jest zbiór wszystkich tych liczb rzeczywistych, dla których mianownik jest różny od 0. Funkcja $\sqrt{1-e^x}$ będzie automatycznie zdefiniowana na zbiorze złożonym z liczb rzeczywistych niedodatnich. W przypadku jakichkolwiek wieloznaczności będziemy wyraźnie określać dziedzinę. Czasem też dziedzina z jakichś przyczyn będzie mniejsza niż maksymalna, np. zmienna będzie mieć jakieś pozamatematyczne znaczenie i wtedy interpretacja będzie żródłem ograniczeń dziedziny. Np. pytanie o maksymalne pole prostokąta o obwodzie 4 prowadzi do rozpatrywania funkcji $x(2-x)$ na przedziale otwartym $(0,2)$: x oznacza tu jeden wymiar prostokąta, a $2-x$ – drugi. Funkcję $x(2-x)$ można rozpatrywać nie tylko na przedziale $(0,2)$, ale z punktu widzenia zadanego pytania, nie ma to sensu.

W dalszej części zajmiemy się również funkcjami określonymi na podzbiorach płaszczyzny (czyli zbioru wszystkich liczb zespolonych $\mathbb{C}$), przestrzeni trójwymiarowej i ogólnie n–wymiarowej. Wartościami tych funkcji będą zazwyczaj liczby rzeczywiste, ale wystąpią również funkcje przekształcające pewne podzbiory płaszczyzny w płaszczyznę. Takie funkcje będą nazywane na ogół przekształceniami lub odwzorowaniami. Nie oznacza to, że funkcji z $\mathbb{R}$ *na* $\mathbb{R}$ danej wzorem $f(x)=x+1$ nie można nazwać odwzorowaniem – często termin ten jest używany, zwłaszcza wtedy, gdy mówimy o geometrii związanej z tą funkcją – jest przesunięcie o 1 w prawo.

2. Funkcje różnowartościowe, funkcja odwrotna

Ważną klasą funkcji są funkcje różnowartościowe, tj. takie które różnym punktom dziedziny przypisują różne wartości: $(x\neq y)\Rightarrow(f(x)\neq f(y))$. Jeśli f jest funkcją różnowartościową przekształcającą zbiór A *na* zbiór B, to można określić funkcję f^{-1} odwrotną do danej funkcji f: $f^{-1}(b)=a \iff b=f(a)$. Jeśli $f(x)=x^3$ dla każdej liczby rzeczywistej x, to funkcja f przekształca różnowartościowo zbiór $\mathbb{R}$ *na* siebie, więc można określić funkcję odwrotną: $f^{-1}(x)=\sqrt[3]{x}$. Jeśli $f(x)=e^x$ dla każdej liczby rzeczywistej x, to zbiorem wartości funkcji f jest zbiór wszystkich liczb dodatnich i wobec tego $f^{-1}(x)=\ln x$ dla każdej dodatniej liczby x. Jeśli $f(x)=x^2$ dla nieujemnych liczb x, to $f^{-1}(x)=\sqrt{x}$ dla każdej liczby nieujemnej x. Jeśli $f(x)=x^2$ dla każdej liczby niedodatniej x, to funkcja f przekształca zbiór wszystkich liczb niedodatnich *na* zbiór wszystkich liczb nieujemnych. Funkcja odwrotna do niej dana jest wzorem $f^{-1}(x) = -\sqrt{x}$. W ostatnich dwóch przykładach wzór był identyczny, ale dziedziny były różne. W związku z tym wzory na funkcję odwrotne też były różne.

W dalszym ciągu będziemy używać jeszcze dwu funkcji zdefiniowanych jako odwrotne do funkcji sinus i tangens. Oczywiście funkcje sinus i tangens jako okresowe nie są różnowartościowe, więc nie mają funkcji odwrotnych. Można więc postąpić tak, jak w przypadku pierwiastka kwadratowego, który jest zdefiniowany jako funkcja odwrotna do funkcji x^2 rozpatrywanej nie na całej dziedzinie, lecz na zbiorze, na którym funkcja x^2 jest różnowartościowa, i to możliwie najprostszym o tej własności.* Wybieramy możliwe najbardziej naturalne dziedziny. W przypadku sinusa ograniczamy się do przedziału $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, a w przypadku tangensa – do przedziału $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Zbiory wartości to odpowiednio Przedział domknięty $[-1, 1]$ i cała prosta $(-\infty, +\infty)$. Tradycyjnie zamiast pisać $\sin^{-1}$ piszemy $\arcsin$, a zamiast tg^{-1} piszemy arctg**, co zresztą pozwala na uniknięcie dwuznaczności związanej z oznaczeniami $\sin^{-1}$ i tg^{-1}. Podamy teraz definicje tych funkcji w jawny sposób.

Definicja funkcji arcsin i arctg

Jeśli $x \in [-1, 1]$, to $\arcsin x$ jest jedyną liczbą z przedziału $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, dla której zachodzi równość $\sin(\arcsin x) = x$.

Jeśli x jest liczbą rzeczywistą, to $\mathrm{arctg}\, x$ jest jedyną liczbą rzeczywistą z przedziału $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, dla której zachodzi równość $\mathrm{tg}(\mathrm{arctg}\, x) = x$. ■

Podamy przykłady $\arcsin 1 = \frac{\pi}{2}$, $\arcsin \frac{1}{2} = \frac{\pi}{6}$, $\arcsin\left(-\frac{\sqrt{2}}{2}\right) = -\frac{\pi}{4}$, $\mathrm{arctg}\sqrt{3} = \frac{\pi}{3}$, $\mathrm{arctg}(-1) = -\frac{\pi}{4}$, $\mathrm{arctg}\, 0 = 0$.

3. Granica funkcji

Wprowadzimy oznaczenie: $\overline{\mathbb{R}} = [-\infty, +\infty]$ oznacza zbiór złożony ze wszystkich liczb rzeczywistych uzupełniony symbolami nieskończonymi $-\infty$ i $+\infty$. Można myśleć, że $\overline{\mathbb{R}}$ to prosta z końcami. Podkreślić wypada, że symboli nieskończonych nie traktujemy jak liczb, bo np. nie wszystkie działania z ich użyciem są wykonalne.

Definicja punktu skupienia

Punkt $p \in \overline{\mathbb{R}}$ jest punktem skupienia zbioru $A \subset \mathbb{R}^n$ wtedy i tylko wtedy, gdy istnieje ciąg (a_n) punktów zbioru A, o wyrazach różnych od a, zbieżny do p. ■

$+\infty$ jest punktem skupienia zbioru wszystkich liczb naturalnych $\mathbb{N}$ – by się o tym przekonać wystarczy przyjąć $a_n = n$. Innych punktów skupienia zbiór $\mathbb{N}$ nie ma. W grę mogłyby wchodzić jedynie liczby nieujemne, bo granica ciągu liczb naturalnych jest albo równa $+\infty$, albo też jest liczbą nieujemną. Jeśli ciąg liczb naturalnych ma skończoną granicę, to ze względu na warunek Cauchy'ego odległości między wyrazami tego ciągu, których numery są dostatecznie duże, są mniejsze niż 1, a ponieważ są to liczby całkowite, więc te odległości są równe 0. Wykazaliśmy, że ciąg liczb naturalnych, który ma skończoną granicę musi być od pewnego miejsca stały, a więc granica jest równa pewnym wyrazom ciągu. Jest niezgodne z definicją punktu skupienia.

Każda liczba z przedziału domkniętego $[0, 1]$ jest punktem skupienia przedziału otwartego $(0, 1)$. Innych

---

* Zbiorów, na których funkcja x^2 jest różnowartościowa, jest bardzo dużo, np, $[-1,0]\cup(1,+\infty)$, $(-\infty,-2)$, $(-\infty,0]$, $[0,+\infty)$, zbiór złożony ze wszystkich liczb wymiernych dodatnich oraz ujemnych liczb niewymiernych i wiele innych.

** W niektórych krajach $\arctan$.

punktów skupienia przedział $(0,1)$ nie ma. To drugie zdanie jest prawdziwe w oczywisty sposób – granica ciągu liczb z przedziału $(0,1)$ musi się znajdować w przedziale $[0,1]$. Jest też jasne, że dla każdej liczby p z przedziału $[0,1]$ istnieje ciąg (a_n) liczb z przedziału $(0,1)$, taki że $p = \lim\limits_{n\to\infty} a_n$ oraz $a_n \neq p$ dla każdego n.

Każda liczba rzeczywista i oba symbole nieskończone są punktami skupienia dziedziny funkcji tangens, tj. zbioru tych liczb rzeczywistych, które nie są nieparzystymi wielokrotnościami liczby $\frac{\pi}{2}$. Łatwe uzasadnienie tego stwierdzenia pozostawiamy czytelnikom.

Teraz możemy już zdefiniować granicę funkcji.

Definicja granicy funkcji w punkcie.*

Niech p oznacza dowolny punkt skupienia dziedziny funkcji f. Mówimy, że $g \in \overline{\mathbb{R}}$ jest granicą funkcji f w punkcie p wtedy i tylko wtedy, gdy dla każdego ciągu (x_n) zbieżnego do p, którego wszystkie wyrazy są różne od p, ma miejsce równość $\lim\limits_{n\to\infty} f(x_n) = g$. Granicę funkcji f w punkcie p oznaczamy symbolem $\lim\limits_{x\to p} f(x)$. ■

Zwrócić należy uwagę na to, że wśród wyrazów ciągu zbieżnego do p, występującego w definicji granicy, nie ma p. Oznacza to w szczególności, że nawet wtedy, gdy p jest argumentem funkcji f, to wartość w tym punkcie nie ma wpływu na istnienie granicy w punkcie p, ani na jej wartość – można dowolnie zmieniać wartość funkcji w punkcie p nie zmieniając granicy w tym punkcie. Oznacza to, że jeśli funkcja ma granicę w punkcie p, to w dostatecznie bliskich punktach x wartość $f(x)$ jest bliska granicy g, pod warunkiem jednak, że $x \neq p$. Ponieważ ciąg ma co najwyżej jedną granicę, więc również funkcja może mieć tylko jedną granicę w jednym punkcie. Pojęcie granicy funkcji jest bardzo ważne, jest rozszerzeniem pojęcia granicy ciągu. Podamy teraz kilka przykładów.

Przykłady granic.

1. $\lim\limits_{x\to 0} \frac{\sin x}{x} = 1$. Równość ta została udowodniona wcześniej. ■

2. $\lim\limits_{x\to 0} \frac{\ln(1+x)}{x} = 1$. Również ta równość została udowodniona wcześniej. ■

3. $\lim\limits_{x\to 0} \frac{e^x - 1}{x} = 1$. Tę równość wykazaliśmy poprzednio.■

4. $\lim\limits_{x\to\infty} \left(1 + \frac{1}{x}\right)^x = e$. Tę równość wykażemy teraz. Trzeba wykazać, że dla każdego ciągu (x_n), którego granicą jest ∞ zachodzi równość $\lim\limits_{n\to\infty} \left(1 + \frac{1}{x_n}\right)^{x_n} = e$. Wiemy, że jest tak w przypadku $x_n = n$ – z definicji liczby e. Przypomnijmy też, że ciąg $\left(1 + \frac{1}{n}\right)^n$ jest rosnący. Stąd wynika, że jeśli $k > n$ jest liczbą naturalną, to $\left(1 + \frac{1}{n}\right)^n < \left(1 + \frac{1}{k}\right)^k < e$. Stąd i z definicji granicy wynika, że jeśli $\lim\limits_{n\to\infty} k_n = +\infty$, to $\lim\limits_{n\to\infty} \left(1 + \frac{1}{k_n}\right)^{k_n} = e$ – jeśli bowiem m jest jakąkolwiek liczbą naturalną, to dla dostatecznie dużych liczb naturalnych n, zachodzi nierówność $k_n > m$, zatem $\left(1 + \frac{1}{m}\right)^m < \left(1 + \frac{1}{k_n}\right)^{k_n} < e$. Teraz możemy przejść do właściwego dowodu. Niech $\lim\limits_{n\to\infty} x_n = +\infty$. Bez straty ogólności rozważań można przyjąć, że dla każdego n zachodzi nierówność $x_n \geq 1$, bo jest tak dla dostatecznie dużych n. Niech k_n będzie taką liczbą całkowitą, że $k_n \leq x_n < k_n + 1$

* Ta definicja jest nazywana ciągową lub definicją Heinego

– taka liczba k_n istnieje dokładnie jedna. Ponieważ $x_n - 1 < k_n$, więc $\lim\limits_{n\to\infty} k_n = +\infty$. Stąd i z tego, co wykazaliśmy poprzednio, wynika, że $\lim\limits_{n\to\infty}\left(1+\dfrac{1}{1+k_n}\right)^{k_n} = e = \lim\limits_{n\to\infty}\left(1+\dfrac{1}{k_n}\right)^{1+k_n}$.
Mamy również

$$\left(1+\frac{1}{1+k_n}\right)^{k_n} \leq \left(1+\frac{1}{1+k_n}\right)^{x_n} < \left(1+\frac{1}{x_n}\right)^{x_n} \leq \left(1+\frac{1}{k_n}\right)^{x_n} < \left(1+\frac{1}{k_n}\right)^{1+k_n} .$$

Z tej nierówności i twierdzenia o trzech ciągach wynika dowodzona przez nas teza. ■

5. Funkcja $\dfrac{1}{x}$, określona dla $x \neq 0$, nie ma granicy w punkcie 0, bowiem $\lim\limits_{n\to\infty}\dfrac{1}{1/n} = +\infty$ i jednocześnie $\lim\limits_{n\to\infty}\dfrac{1}{-1/n} = -\infty$, udało się nam więc wskazać dwa ciągi argumentów zbieżne do 0, takie że odpowiadające im ciągi wartości mają różne granice. ■

6. Funkcja $\sin\dfrac{1}{x}$, określona dla $x \neq 0$, nie ma granicy w punkcie 0, bowiem $\sin\dfrac{1}{1/(2n\pi)} = 0$ oraz $\sin\dfrac{1}{1/(2n\pi+\pi/2)} = 1$. Wskazaliśmy więc dwa ciągi argumentów, takie że odpowiadające im ciągi wartości są stałe i różne. ■

Oprócz granicy funkcji rozpatrywane są granice jednostronne funkcji w punkcie. Zdefiniujemy granicę lewostronną, definicja granicy prawostronnej jest analogiczna.

Definicja granicy lewostronnej

g jest granicą lewostronna funkcji f w punkcie p wtedy i tylko wtedy, gdy można znaleźć w dziedzinie ciąg (x_n) o wyrazach mniejszych (ściśle!) niż p, zbieżny do p i gdy dla każdego takiego ciągu odpowiadający mu ciąg wartości $(f(x_n))$ ma granicę g. Stosujemy oznaczenie $\lim\limits_{x\to p^-} f(x)$. ■

Łatwo można udowodnić, że funkcja $\dfrac{1}{x}$ ma jednostronne granice w punkcie 0: prawostronna jest równa $+\infty$, zaś lewostronną jest $-\infty$. Funkcja $\sin\dfrac{1}{x}$ nie ma granicy prawostronnej w punkcie 0 – wykazaliśmy to w przykładzie 6, wskazując dwa ciągi *dodatnich* argumentów tej funkcji zbieżne do 0, takie że odpowiadające im ciągi wartości mają różne granice.

Bez trudu można udowodnić „funkcyjną” wersję twierdzenia o scalaniu.

Twierdzenie o scalaniu

Funkcja f określona na zbiorze zawierającym ciąg liczb mniejszych niż p, zbieżny do p oraz ciąg liczb większych niż p, zbieżny do p, ma granicę w punkcie p wtedy i tylko wtedy, gdy ma obie granice jednostronne i są one równe.

Dowód. Jest jasne, że z istnienia granicy wynika istnienie granic jednostronnych – zamiast wszystkich ciągów zbieżnych do p, których wyrazy są różne od p, rozpatrujemy jedynie ich część. Jeśli natomiast wiemy, że istnieją granice jednostronne, to ciąg o wyrazach różnych od p możemy rozbić na podciąg o wyrazach mniejszych niż p i na podciąg o wyrazach większych niż p. Odpowiadające im ciągi wartości mają tę samą granicę, więc ciąg wartości odpowiadający naszemu ciągowi ma granicę i to równą wspólnej wartości obu granic jednostronnych. Oczywiście jeśli ciąg argumentów zawiera jedynie skończenie wiele wyrazów większych niż p, to nie możemy rozpatrywać granicy prawostronnej, ale to niczemu nie przeszkadza, bo w tym przypadku wystarczy skorzystać z istnienia granicy lewostronnej.

Podobnie jak w przypadku twierdzenia o scalaniu, można przenieść inne twierdzenia dotyczące granic ciągów na ogólniejszy przypadek granicy funkcji.

Twierdzenie o arytmetycznych własnościach granicy

A1. Jeśli istnieją granice $\lim\limits_{x\to p} f(x)$, $\lim\limits_{x\to p} g(x)$ i określona jest ich suma, to istnieje granica $\lim\limits_{x\to p}(f(x)+g(x))$ i zachodzi wzór: $\lim\limits_{x\to p}(f(x)+g(x)) = \lim\limits_{x\to p} f(x) + \lim\limits_{x\to p} g(x)$.

A2. Jeśli istnieją granice $\lim\limits_{x\to p} f(x)$, $\lim\limits_{x\to p} g(x)$ i określona jest ich różnica, to istnieje granica $\lim\limits_{x\to p}(f(x)-g(x))$ i zachodzi wzór: $\lim\limits_{x\to p}(f(x)-g(x)) = \lim\limits_{x\to p} f(x) - \lim\limits_{x\to p} g(x)$.

A3. Jeśli istnieją granice $\lim\limits_{x\to p} f(x)$, $\lim\limits_{x\to p} g(x)$ i określony jest ich iloczyn, to istnieje granica $\lim\limits_{x\to p}(f(x)\cdot g(x))$ i zachodzi wzór: $\lim\limits_{x\to p}(f(x)\cdot g(x)) = \lim\limits_{x\to p} f(x) \cdot \lim\limits_{x\to p} g(x)$.

A4. Jeśli istnieją granice $\lim\limits_{x\to p} f(x)$, $\lim\limits_{x\to p} g(x)$ i określony jest ich iloraz, to istnieje granica $\lim\limits_{x\to p} \dfrac{f(x)}{g(x)}$ i zachodzi wzór $\lim\limits_{x\to p} \dfrac{f(x)}{g(x)} = \dfrac{\lim\limits_{x\to p} f(x)}{\lim\limits_{x\to p} g(x)}$.

Dowód tego twierdzenia jest natychmiastową konsekwencją twierdzenia o arytmetycznych własnościach granicy ciągu. ■

Przed podaniem następnego twierdzenia przypomnijmy, że operujemy terminem *dla dostatecznie dużych* n. Oznacza to, że interesują nas liczby naturalne większe od pewnej liczby. Właściwie chodzi o to, by były one bliskie $+\infty$. W przypadku funkcji argument, którym w przypadku ciągu jest numer wyrazu, czyli n, ma być bliski punktowi p, który może – lecz nie musi – być równy $+\infty$. Wymaga więc zmiany sposób mówienia. Mówiąc x *jest dostatecznie bliski* p będziemy mieć na myśli, że:

$(+\infty)$ $x > M$ dla pewnej liczby rzeczywistej M, gdy $p = +\infty$,

$(-\infty)$ $x < M$ dla pewnej liczby rzeczywistej M, gdy $p = -\infty$,

$(\mathbb{R})$ $|x-p| < \delta$ dla pewnej **dodatniej** liczby δ, gdy $p \in \mathbb{R}$.

Twierdzenie o szacowaniu

N1. Jeśli $C < \lim\limits_{x\to p} f(x)$, to dla $x \neq p$, dostatecznie bliskich p zachodzi nierówność $C < f(x)$.

N2. Jeśli $C > \lim\limits_{x\to p} f(x)$, to dla $x \neq p$, dostatecznie bliskich p zachodzi nierówność $C > f(x)$.

N3. Jeśli $\lim\limits_{x\to p} g(x) < \lim\limits_{x\to p} f(x)$, to dla $x \neq p$, dostatecznie bliskich p zachodzi nierówność $g(x) < f(x)$.

N4. Jeśli $g(x) \le f(x)$ dla x dostatecznie bliskich p, to zachodzi nierówność $\lim\limits_{x\to p} g(x) \le \lim\limits_{x\to p} f(x)$.

Dowód. Zakładamy cały czas, że p jest punktem skupienia dziedziny funkcji. Zauważmy najpierw, że zaprzeczeniem zdania: *Dla wszystkich* $x \neq p$ *dostatecznie bliskich* p *spełniony jest warunek* W jest zdanie: *Istnieje ciąg* (x_n) *zbieżny do* p*, taki że* $x_n \neq p$ *dla każdego* n *i warunek* W *nie zachodzi dla żadnego wyrazu ciągu* (x_n).

Jeśli np. $p = +\infty$ i nie jest prawdą, że warunek W spełniony jest dla wszystkich x dostatecznie bliskich $p = +\infty$, to dla każdej liczby rzeczywistej M istnieje liczba $x > M$, dla której warunek W nie zachodzi. By otrzymać ciąg (x_n), którego granicą jest ∞, złożony z liczb, dla których warunek W nie zachodzi, wystarczy przyjąć, że $M = n$. Jeśli natomiast istnieje ciąg (x_n), którego granicą jest $+\infty$, taki że warunek W nie jest spełniony dla żadnego (x_n), to warunek W nie jest spełniony dla wszystkich dostatecznie dużych x, czyli nie jest spełniony dla wszystkich x dostatecznie bliskich $+\infty$. Analogicznie postępujemy w przypadku $p = -\infty$.

Jeśli $p \in \mathbb{R}$, to dla każdego $\delta > 0$ istnieje x, takie że $x \neq p$ i $|x-p| < \delta$, dla którego warunek W

nie zachodzi. By zdefiniować x_n przyjmujemy, że $\delta = \frac{1}{n}$. Z istnienia ciągu (x_n) złożonego z liczb, dla których warunek W nie zachodzi, wynika od razu, że nie jest możliwe, by warunek W był spełniony dla wszystkich x dostatecznie bliskich p.
Teraz możemy zająć się właściwym dowodem. Załóżmy, że $\lim\limits_{x\to p} f(x) < C$ oraz że nie jest prawdą, że dla x dostatecznie bliskich p zachodzi nierówność $f(x) < C$. Wynika stąd, że istnieje ciąg (x_n), taki że dla każdego n zachodzi nierówność $f(x_n) \geq C$. Stąd jednak wynika, że $\lim\limits_{x\to p} f(x_n) \geq C$, wbrew założeniu. Dowód w tym przypadku został zakończony. Stwierdzenie N2 dowodzimy analogicznie lub wnioskujemy z N1 zastępując funkcję f funkcją przeciwną $-f$. Stwierdzenie N3 wynika ze stwierdzeń poprzednich: starczy użyć liczby C leżącej między $\lim\limits_{x\to p} f(x)$ oraz $\lim\limits_{x\to p} g(x)$. Ostatni fragment twierdzenia to prosta konsekwencja tego, że ciąg o mniejszych wyrazach ma mniejszą granicę. Dowód został zakończony. ■

Podamy teraz inną definicję granicy funkcji. Z poprzednią można wiązać takie stwierdzenie (nieścisłe, ale ważne): *niezależnie od tego w jaki sposób argument dąży do p, to wartość funkcji zbliża się do g*. Z tą która pojawi się niebawem wiążemy stwierdzenie *jeśli argument funkcji jest dostatecznie bliski p, ale różny od p, to wartość funkcji jest bliska g*. Sformułujemy zapowiedzianą definicję bardzo dokładnie, bez żadnych skrótów. Ma ona dziewięć części, ale na ogół po przeczytaniu dwóch – trzech pierwszych nie ma potrzeby czytać dalej, bo można to samodzielnie napisać.

Definicja granicy funkcji*

1. $g, p \in \mathbb{R}$. Wtedy $g = \lim\limits_{x\to p} f(x)$ wtedy i tylko wtedy, gdy dla każdej liczby $\varepsilon > 0$ istnieje liczba $\delta > 0$ taka, że jeśli $0 < |x - p| < \delta$, to $|f(x) - g| < \varepsilon$.

2. $g \in \mathbb{R}$, $p = +\infty$. Wtedy $g = \lim\limits_{x\to p} f(x)$ wtedy i tylko wtedy, gdy dla każdej liczby $\varepsilon > 0$ istnieje liczba rzeczywista M, taka że jeśli $x > M$, to $|f(x) - g| < \varepsilon$.

3. $g \in \mathbb{R}$, $p = -\infty$. Wtedy $g = \lim\limits_{x\to p} f(x)$ wtedy i tylko wtedy, gdy dla każdej liczby $\varepsilon > 0$ istnieje liczba rzeczywista M, taka że jeśli $x < M$, to $|f(x) - g| < \varepsilon$.

4. $g = +\infty$, $p \in \mathbb{R}$. Wtedy $g = \lim\limits_{x\to p} f(x)$ wtedy i tylko wtedy, gdy dla każdej liczby M istnieje liczba rzeczywista $\delta > 0$, taka że jeśli $0 < |x - p| < \delta$, to $f(x) > M$.

5. $g = +\infty$, $p = +\infty$. Wtedy $g = \lim\limits_{x\to p} f(x)$ wtedy i tylko wtedy, gdy dla każdej liczby M istnieje liczba rzeczywista K, taka że jeśli $x > K$, to $f(x) > M$.

6. $g = +\infty$, $p = -\infty$. Wtedy $g = \lim\limits_{x\to p} f(x)$ wtedy i tylko wtedy, gdy dla każdej liczby M istnieje liczba rzeczywista K, taka że jeśli $x < K$, to $f(x) > M$.

7. $g = -\infty$, $p \in \mathbb{R}$. Wtedy $g = \lim\limits_{x\to p} f(x)$ wtedy i tylko wtedy, gdy dla każdej liczby M istnieje liczba rzeczywista $\delta > 0$, taka że jeśli $0 < |x - p| < \delta$, to $f(x) < M$.

8. $g = -\infty$, $p = +\infty$. Wtedy $g = \lim\limits_{x\to p} f(x)$ wtedy i tylko wtedy, gdy dla każdej liczby M istnieje liczba rzeczywista K, taka że jeśli $x > K$, to $f(x) < M$.

9. $g = -\infty$, $p = -\infty$. Wtedy $g = \lim\limits_{x\to p} f(x)$ wtedy i tylko wtedy, gdy dla każdej liczby M istnieje liczba rzeczywista K, taka że jeśli $x < K$, to $f(x) < M$.

Dowód. Dowód podamy w dwóch wybranych przypadkach: pierwszym i ósmym. Resztę czytelnik powinien uzupełnić samodzielnie, być może nie wszystko – tyle tylko, by w miarę swobodnie przepro-

* ta definicja nazywana jest definicją Cauchy'ego lub definicją otoczeniową, czasem, to już bełkot matematyczny, – epsilonowo–deltową.

wadzić dowód w którymś przypadku.

Założymy najpierw, że g, p są liczbami rzeczywistymi oraz że $g = \lim_{x\to p} f(x)$ w sensie definicji ciągowej. Jeśli istnieje liczba $\varepsilon > 0$, taka że dla każdej liczby $\delta > 0$ istnieje x, takie że $0 < |x-p| < \delta$ i jednocześnie $|f(x) - g| \geq \varepsilon$, to przyjmując, że x_n jest dobrane do $\frac{1}{n}$, tzn. $0 < |x_n - p| < \frac{1}{n}$ i $|f(x_n) - g| \geq \varepsilon$, otrzymujemy ciąg (x_n) zbieżny do p, o wyrazach różnych od p i taki że odpowiadający mu ciąg wartości funkcji nie jest zbieżny do liczby g, bowiem wszystkie wyrazy tego ciągu wartości pozostają w odległości nie mniejszej niż ε od g. Twierdzenie zostało udowodnione w jedna stronę. Teraz założymy, że $g = \lim_{x\to p} f(x)$ w sensie definicji otoczeniowej. Niech (x_n) będzie dowolnym ciągiem argumentów funkcji f zbieżnym do p, o wyrazach różnych od p i niech ε oznacza dowolną liczbę dodatnią. Z definicji otoczeniowej granicy funkcji wynika, że istnieje liczba $\delta > 0$, taka że jeśli $0 < |x-p| < \delta$, to $|f(x) - g| < \varepsilon$. Z definicji granicy ciągu wnioskujemy, że dla dostatecznie dużych n zachodzi nierówność $|x_n - p| < \delta$ i oczywiście $x_n \neq p$, zatem $0 < |x_n - p| < \delta$, a stąd wynika, że $|f(x_n) - g| < \varepsilon$. Stąd i z definicji granicy ciągu wynika, że $\lim_{n\to\infty} f(x_n) = g$, a wobec tego, że (x_n) jest dowolnym ciągiem, możemy stwierdzić, że g jest granicą w sensie definicji ciągowej.

Teraz, zgodnie z obietnicą, zajmiemy się przypadkiem 8, tj. założymy, że $g = -\infty$ oraz że $p = +\infty$. Zakładamy, że dla każdego ciągu (x_n) argumentów funkcji f, którego granicą jest $+\infty$ zachodzi równość $\lim_{n\to\infty} f(x_n) = -\infty$. Mamy wykazać, że dla każdej liczby rzeczywistej M istnieje liczba rzeczywista K, taka że jeśli $x > K$, to $f(x) < M$. Załóżmy, że tak nie jest. Istnieje więc liczba M taka, że dla każdej liczby K istnieje argument x funkcji f, taki że $x > K$ i jednocześnie $f(x) \geq M$. Przyjmując $K = n$ otrzymujemy argument x_n, taki że $x_n \geq n$ i $f(x_n) \geq M$. Stąd jednak wynika, że $-\infty$ nie jest granicą ciągu $(f(x_n))$, wbrew założeniu, kończy to dowód w jedną stronę. Teraz założymy, że dla każdej liczby rzeczywistej M istnieje liczba rzeczywista K, taka że jeśli $x > K$, to $f(x) < M$. Jeśli $\lim_{n\to\infty} x_n = +\infty$, to dla dostatecznie dużych n zachodzi nierówność $x_n > K$ i wobec tego $f(x_n) < M$. Wobec dowolności M, oznacza to, że $\lim_{n\to\infty} f(x_n) = -\infty$. Dowód został zakończony. ■

Z twierdzenia o trzech ciągach wynika analogiczne twierdzenie dla granic funkcji.

Twierdzenie o trzech funkcjach

Jeśli dla wszystkich argumentów x dostatecznie bliskich punktowi p zachodzi następująca nierówność podwójna $f(x) \leq g(x) \leq h(x)$ i istnieją granice $\lim_{x\to p} f(x)$, $\lim_{x\to p} h(x)$ oraz $\lim_{x\to p} f(x) = \lim_{x\to p} h(x)$, to również funkcja g ma granicę w punkcie p i zachodzi równość $\lim_{x\to p} f(x) = \lim_{x\to p} g(x) = \lim_{x\to p} h(x)$. ■

Z następnego twierdzenia w zasadzie nie będziemy korzystać, podajemy je tylko po to, by pokazać, pełną analogię pojęcia granicy ciągu i granicy funkcji, więc łatwy dowód pozostawiamy czytelnikom w charakterze zadania.

Twierdzenie Cauchy'ego o istnieniu granicy skończonej

Funkcja f ma granicę skończoną w punkcie p wtedy i tylko wtedy, gdy spełniony jest następujący warunek Cauchy'ego:

dla każdego $\varepsilon > 0$, dla wszystkich $x, y \neq p$ dostatecznie bliskich p
zachodzi nierówność $|f(x) - f(y)| < \varepsilon$. (w.C.) ■

Twierdzenie, które znajduje się poniżej ma bardzo prosty dowód, ale jest bardzo często stosowane.

Twierdzenie o granicy złożenia dwu funkcji

Załóżmy, że dziedzina funkcji f zawiera zbiór wartości funkcji g, że funkcja g ma granicę G w punkcie

p, że granica G jest punktem skupienia dziedziny funkcji f i funkcja f ma granicę H w punkcie G oraz że wartości funkcji g w punktach dostatecznie bliskich p są różne od G. Przy tych założeniach funkcja $f \circ g$ określona wzorem $(f \circ g)(x) = f(g(x))$ ma w punkcie p granicę, ta granica jest równa H. Założenia tego twierdzenia są tak dobrane, że dowód wynika od razu z definicji ciągowej granicy funkcji w punkcie. ■

Przed podaniem twierdzenia o istnieniu granic jednostronnych funkcji monotonicznej omówimy pojęcie kresu zbioru i kresu funkcji. Rozpoczniemy od definicji.

Definicja kresów

gz. Kresem górnym zbioru niepustego $A \subseteq \mathbb{R}$ nazywamy taki element M zbioru $\overline{\mathbb{R}}$, że dla każdego $a \in A$ zachodzi nierówność $a \leq M$ oraz że jeśli $M' < M$, to istnieje $a \in A$, dla którego $a > M'$. Innymi słowy: M jest najmniejszym ograniczeniem górnym zbioru A. Piszemy $\sup A$.

gf. Kresem górnym M funkcji f nazywamy kres górny zbioru jej wartości, tj. najmniejszą liczbę M, taką że $f(x) \leq M$ dla każdego argumentu x funkcji f. Piszemy $\sup f$.

dz. Kresem dolnym zbioru niepustego $A \subseteq \mathbb{R}$ nazywamy taki element M zbioru $\overline{\mathbb{R}}$, że dla każdego $a \in A$ zachodzi nierówność $a \geq M$ oraz że jeśli $M' > M$, to istnieje $a \in A$, dla którego $a < M'$. Innymi słowy: M jest największym ograniczeniem dolnym zbioru A. Piszemy $\inf A$.

df. Kresem dolnym M funkcji f nazywamy kres dolny zbioru jej wartości, tj. największą liczbę M, taką że $f(x) \geq M$ dla każdego argumentu x funkcji f. Piszemy $\inf f$.

Definicja granicy górnej

$M \in \overline{\mathbb{R}}$ jest granicą górną funkcji f przy $x \longrightarrow p$ wtedy i tylko wtedy, gdy

(i) dla każdego ciągu (x_n) o granicy p, wyrazach *różnych* od p, dla którego istnieje $\lim\limits_{n\to\infty} f(x_n)$ zachodzi nierówność $\lim\limits_{n\to\infty} f(x_n) \leq L$ oraz

(ii) M jest najmniejszym elementem $\overline{\mathbb{R}}$, dla którego spełniony jest warunek (i).

Piszemy wtedy $M = \limsup\limits_{x\to p} f(x)$. ■

Analogicznie definiujemy granicę dolną, którą oznaczamy przez $M = \liminf\limits_{x\to p} f(x)$. Warunek (ii) tej definicji można zastąpić stwierdzeniem: istnieje ciąg (x_n) o granicy p i wyrazach różnych od p, dla którego $\lim\limits_{n\to\infty} f(x_n) = L$. Oznacza to, że granica górna jest kresem górnym granic postaci $\lim\limits_{n\to\infty} f(x_n)$, gdzie (x_n) oznacza ciąg o wyrazach różnych od p, którego granicą jest p. Definicja granicy dolnej jest analogiczna. Można bez trudu wykazać, że jeśli p jest liczbą rzeczywistą, D dziedziną funkcji f, to

$$\left[\liminf_{x\to p} f(x), \limsup_{x\to p} f(x)\right] = \bigcap_{\delta>0} \left[\inf f\left(D \cap (p-\delta, p+\delta) \setminus \{p\}\right), \sup f\left(D \cap (p-\delta, p+\delta) \setminus \{p\}\right)\right].$$

Oznacza to, że rozpatrujemy otoczenie punktu p, znajdujemy najmniejszy przedział domknięty (być może nieskończony) zawierający obraz tej części dziedziny, która znalazła się w rozpatrywanym otoczeniu punktu p, z wyjątkiem punktu samego p. Następnie zmniejszamy to otoczenie (czyli mówiąc nieformalnie $\delta \to 0$), lewy koniec otrzymanego przedziału, być może zdegenerowanego do jednego punktu, to $\liminf\limits_{x\to p} f(x)$, a prawy – to $\limsup\limits_{x\to p} f(x)$. Zachęcam studentów do wykazania równoważności tych określeń oraz do samodzielnego ich sformułowania w przypadku $p = \pm\infty$.

Podamy teraz kilka przykładów kresów funkcji.

5. Niech $f(x) = \dfrac{x}{\sqrt{1+x^2}}$. Jest jasne, że $-1 < f(x) < 1$ dla każdej liczby rzeczywistej x. Czytelnik sprawdzi z łatwością, że jeśli $0 \le a < 1$ i $x > \dfrac{a}{\sqrt{1-a^2}}$, to $a < \dfrac{x}{\sqrt{1+x^2}} = f(x)$. Wykazaliśmy więc, że 1 jest ograniczeniem górnym funkcji f oraz że żadna liczba dodatnia mniejsza niż 1 nie jest ograniczeniem górnym funkcji f. Stąd wynika, że $\sup f = 1$. Ponieważ funkcja f jest nieparzysta ($f(-x) = -x$ dla każdego x), więc $\inf f = -1$.
6. Kresem górnym funkcji sin jest liczba 1, a kresem dolnym funkcji sinus – liczba -1.
7. Kresem górnym funkcji wykładniczej o podstawie e jest $+\infty$, a dolnym – liczba 0.
8. Kresem górnym logarytmu naturalnego jest $+\infty$, a kresem dolnym jest $-\infty$.
9. Kresem górnym funkcji liniowej niestałej jest $+\infty$, a kresem dolnym tej funkcji jest $-\infty$.
10. Kresem górnym funkcji f, danej wzorem $f(x) = x^2 + 2x - 2 = (x+1)^2 - 3$, jest $+\infty$, a kresem dolnym tej funkcji jest liczba -3.

4. Funkcje monotoniczne

Rozpoczniemy od przypomnienia definicji funkcji niemalejących i nierosnących, ściśle malejących i ściśle rosnących.

Definicja funkcji monotonicznych i ściśle monotonicznych

Funkcja f określona na zbiorze, którego elementami są liczby rzeczywiste, jest

1. ściśle rosnąca wtedy i tylko wtedy, gdy dla dowolnych argumentów $x,\ y$ konsekwencją nierówności $x < y$ jest nierówność $f(x) < f(y)$;

2. ściśle malejąca wtedy i tylko wtedy, gdy dla dowolnych argumentów $x,\ y$ konsekwencją nierówności $x < y$ jest nierówność $f(x) > f(y)$;

3. nierosnąca wtedy i tylko wtedy, gdy dla dowolnych argumentów $x,\ y$ konsekwencją nierówności $x < y$ jest nierówność $f(x) \ge f(y)$;

4. niemalejąca wtedy i tylko wtedy, gdy dla dowolnych argumentów $x,\ y$ konsekwencją nierówności $x < y$ jest nierówność $f(x) \le f(y)$. $\blacksquare$

Znów mamy do czynienia z rozszerzeniem pojęcia z ciągów na funkcje. Podamy tylko jeden przykład, bo licznych przykładów dostarczyliśmy już wcześniej w postaci ciągów monotonicznych, a i w przyszłości ich nie zabraknie. Niech $f(x) = \dfrac{1}{x}$ dla $x \ne 0$. Zauważmy, że w całej dziedzinie funkcja nie jest monotoniczna: $-1 < 1$ i jednocześnie $f(-1) = -1 < 1 = f(1)$, więc funkcja f nie może być nierosnąca, w szczególności nie może być malejąca; $1 < 2$ i jednocześnie $f(1) = 1 > \dfrac{1}{2} = f(2)$, zatem f nie może być funkcją niemalejącą, tym bardziej rosnącą. Natomiast czytelnik stwierdzi bez trudu, że f jest funkcją malejącą na każdej z dwu półprostych $(-\infty, 0)$ i $(0, +\infty)$.

Twierdzenie o istnieniu granic funkcji monotonicznej

Jeśli f jest funkcją monotoniczną, p jest punktem skupienia jej dziedziny, to jeśli istnieje ciąg (x_n) argumentów funkcji f mniejszych niż p, zbieżny do p, to f ma granicę lewostronną w punkcie p, jeśli istnieje ciąg argumentów funkcji f większych niż p, zbieżny do p, to f ma granicę prawostronną w punkcie p.

Dowód. Wystarczy oczywiście udowodnić to twierdzenie przy założeniu, że funkcja f jest niemalejąca. Jeśli bowiem f jest nierosnąca, to funkcja przeciwna $-f$ jest niemalejąca. Niech g będzie kresem górnym zbioru złożonego z wartości funkcji f osiąganych w punktach $x < p$ $\quad (g = \sup\{f(x)\colon \quad x <$

$p\}$). Trzeba wykazać, że $g = \lim_{x \to p^-} f(x)$. Jeśli $x < p$, to $f(x) \leq g$. Jeśli $m < g$ jest liczbą rzeczywistą, to ponieważ m nie jest ograniczeniem górnym zbioru tych wartości funkcji f, które są osiągane w punktach $x < p$, więc istnieje argument $x' < p$, taki że $f(x') > m$. Stąd wynika, że jeśli $x' < x < p$ to $m < f(x') \leq f(x) \leq g$, zatem: jeśli $x < p$ jest dostatecznie bliskie p, to $f(x)$ jest dostatecznie bliskie g, co dowodzi tego, że $g = \lim_{x \to p^-} f(x)$. Analogicznie dowodzimy, że jeśli istnieje ciąg argumentów większych niż p, zbieżny do p, to kres dolny tych wartości funkcji, które są przyjmowane w punktach większych niż p jest prawostronną granicą funkcji f w punkcie p. Dowód został zakończony. ■

5. Funkcje ciągłe

Przechodzimy teraz do najważniejszego tematu tego rozdziału – do ciągłości. Rozpoczniemy od definicji.

Definicja funkcji ciągłej

Funkcja f jest ciągła w punkcie p wtedy i tylko wtedy, gdy p jest argumentem funkcji i zachodzi jedna z dwu możliwości:

(i) p nie jest punktem skupienia dziedziny funkcji f;

(ii) p jest punktem skupienia dziedziny funkcji f, która ma granicę w punkcie p i ta granica jest równa wartości funkcji w punkcie p: $\lim_{x \to p} f(x) = f(p)$. ■

Twierdzenie charakteryzujące ciągłość

Funkcja f jest ciągła w punkcie p wtedy i tylko wtedy, gdy dla każdej liczby dodatniej ε istnieje liczba $\delta > 0$, taka że jeśli $|x - p| < \delta$, to $|f(x) - f(p)| < \varepsilon$.

Dowód. Jeżeli p nie jest punktem skupienia dziedziny funkcji f, to istnieje liczba $\delta > 0$, taka że jedynym punktem x dziedziny funkcji f, dla którego $|x - p| < \delta$ jest punkt p – w tym przypadku $|f(x) - f(p)| = |f(p) - f(p)| = 0 < \varepsilon$, niezależnie od wyboru liczby dodatniej ε. Pozostała część twierdzenia może być otrzymana natychmiast z definicji otoczeniowej granicy funkcji. Dowód został zakończony. ■

Z poznanych twierdzeń o granicach funkcji wynika od razu następujące twierdzenie.

Twierdzenie o operacjach na funkcjach ciągłych

Załóżmy, że funkcje f i g określone na wspólnej dziedzinie są ciągłe w punkcie p. Wtedy następujące funkcje są ciągłe w punkcie p: $f + g$, $f - g$, $f \cdot g$ oraz $\frac{f}{g}$ pod warunkiem $g(p) \neq 0$. ■

Ważną operacją jest składanie (superponowanie) funkcji. Polega ono na „wykonaniu" po kolei dwu funkcji: $(f \circ g)(x) = f(g(x))$. Okazuje się, że składając funkcje ciągłe otrzymujemy w rezultacie funkcję ciągłą.

Twierdzenie o ciągłości złożenia dwu funkcji

Jeżeli funkcja g jest ciągła w punkcie p, funkcja f określona na zbiorze zawierającym zbiór wartości funkcji g jest ciągła w punkcie $g(p)$, to złożenie $f \circ g$ jest funkcją ciągłą w punkcie p.

Dowód Wynika to od razu z otoczeniowej definicji ciągłości: jeśli $\varepsilon > 0$, to istnieje $\delta > 0$, takie że jeśli $|y - g(p)| < \delta$, to $|f(y) - f(g(p))| < \varepsilon$, istnieje też $\eta > 0$, takie że jeśli $|x - p| < \eta$, to $|g(x) - g(p)| < \delta$, a wobec tego $|f(g(x)) - f(g(p))| < \varepsilon$. Dowód został zakończony. ■

Nie jest natomiast prawdą, że funkcja odwrotna do funkcji f ciągłej w punkcie p musi być ciągła w punkcie $f(p)$. Zachęcamy czytelników do samodzielnego skonstruowania przykładu. Musi on być nieco dziwaczny, bowiem jeśli założymy, że funkcja f jest ciągła w całej dziedzinie, która jest przedziałem, to wtedy funkcja odwrotna musi być ciągłą, mówimy o funkcji, której wartościami są liczby rzeczywiste. Tego twierdzenia jednak nie udowodnimy teraz, bowiem jego dowód stanie się łatwiejszy później.

Przykłady

1. Funkcja stała jest ciągła w każdym punkcie. ■

2. Funkcja *identyczność*, czyli funkcja, której wartością w punkcie x jest liczba x jest ciągła w każdym punkcie prostej – wynika to natychmiast z definicji ciągłości. Zamiast mówić funkcja *identyczność* będziemy mówić funkcja x, rozumiejąc, że jest ona określona na całej prostej. ■

3. Funkcje x^2, x^3, ... są ciągłe w każdym punkcie prostej. Wynika to natychmiast z twierdzenia o ciągłości iloczynu funkcji ciągłych i poprzedniego przykładu. ■

4. Każdy wielomian, czyli funkcja postaci $a_0 + a_1x + \cdots + a_nx^n$, gdzie $a_0, a_1, \ldots, a_n$ są dowolnymi liczbami rzeczywistymi, jest ciągła w każdym punkcie prostej. Wynika to z poprzednich przykładów oraz twierdzenia o ciągłości iloczynu i sumy funkcji: funkcja postaci a_jx^j jest iloczynem funkcji stałej o wartości a_j oraz funkcji x^j, wielomian jest sumą takich funkcji. ■

5. Suma szeregu potęgowego, tj. szeregu postaci $\sum\limits_{n=0}^{\infty} a_nx^n$ jest funkcją ciągłą w każdym punkcie swej dziedziny, tj. w każdym punkcie x, w którym szereg potęgowy jest zbieżny. To twierdzenie nie jest takie proste jak poprzednie, ale udowodniliśmy je poprzednio. ■

6. Funkcja $\frac{x-1}{x+3}$, której dziedziną jest zbiór złożony ze wszystkich liczb rzeczywistych z wyjątkiem liczby -3 jest ciągła w każdym punkcie swej dziedziny, bo jest ilorazem funkcji ciągłych. ■

7. Funkcja wykładnicza e^x jest ciągła. Wykazaliśmy to w rozdziale 1. punkt 19. Wynika to również z wzoru udowodnionego tamże: $e^x = \sum\limits_{n=0}^{\infty} \dfrac{x^n}{n!}$ i przykładu poprzedniego. ■

8. Logarytm naturalny (o podstawie e) jest funkcją ciągłą. Można wywnioskować to z ciągłości funkcji wykładniczej w taki sam sposób jak w przykładzie trzynastym wnioskowana jest ciągłość funkcji $\arcsin$ z ciągłości funkcji sinus.■

9. Dla każdej liczby rzeczywistej a funkcja potęgowa x^a o wykładniku a jest ciągła w każdym punkcie półprostej $(0, +\infty)$. Wynika to z ciągłości logarytmu naturalnego, ciągłości funkcji wykładniczej o podstawie e i ciągłości iloczynu oraz złożenia funkcji ciągłych: $x^a = e^{a \ln x}$. ■

10. Jeśli $a > 0$, to funkcja x^a jest ciągła w punkcie 0, jej wartość w punkcie 0 definiujemy w tym przypadku jako 0. Trzeba udowodnić, że jeśli $\lim\limits_{n\to\infty} x_n = 0$, to również $\lim\limits_{n\to\infty} x_n^a = 0$. Jest tak dla $a = \dfrac{1}{k}$, k – dowolna liczba całkowita większa niż 1, bo $x^{1/k} = \sqrt[k]{x}$ – wykazaliśmy to w rozdziale 1, punkt 17, przykład i. W przypadku dowolnego a znajdujemy najpierw dodatnią liczbę całkowitą $k > \dfrac{1}{a}$. Dla każdej liczby nieujemnej $x < 1$ mamy wtedy $0 \le x^a \le x^{1/k}$. Teza wynika teraz z twierdzenia o trzech ciągach. ■

11. Jeśli $a = \dfrac{p}{q}$, gdzie q jest nieparzystą liczbą całkowitą dodatnią, zaś p liczbą całkowitą ujemną, to funkcja $x^a = \sqrt[q]{x^p}$ jest ciągła w każdym punkcie półprostej $(-\infty, 0)$. Wynika to od razu z ciągłości funkcji pierwiastek q–tego stopnia, ciągłości wielomianu i ciągłości ilorazu funkcji ciągłych oraz twierdzenia o ciągłości złożenia. ■

W ostatnich trzech przykładach wykazaliśmy, że funkcja potęgowa jest ciągła wszędzie tam, gdzie jest określona.

12. Funkcje sinus i kosinus są ciągłe w każdym punkcie prostej. Jest to konsekwencja nierówności $|\sin x - \sin y| \leq |x-y|$ oraz $|\cos x - \cos y| \leq |x-y|$. ■

13. Funkcja arcsin jest ciągła na przedziale $[-1,1]$. Załóżmy, że tak nie jest. Oznacza to, że istnieje ciąg (x_n) punktów przedziału $[-1,1]$ zbieżny do pewnej liczby g, taki że ciąg $(\arcsin x_n)$ nie jest zbieżny do $\arcsin g$. Oznacza, że z ciągu $(\arcsin x_n)$ można wybrać podciąg zbieżny $\arcsin x_{k_n}$ do granicy $G \neq \arcsin g$. Stąd i z ciągłości funkcji sinus wynika, że $g = \lim\limits_{n\to\infty} x_{k_n} = \lim\limits_{n\to\infty} \sin(\arcsin(x_{k_n})) = \sin(G) \neq \sin(\arcsin g) = g$. Otrzymaliśmy sprzeczność $g \neq g$. Oznacza to , że każdy podciąg zbieżny ciągu $(\arcsin x_n)$ ma granicę $\arcsin g$, a to oznacza, że $\lim\limits_{n\to\infty} \arcsin x_n = \arcsin g$. ■

14. Funkcja arctg jest ciągła na całej prostej. Dowód, który można przeprowadzić podobnie do podanego w poprzednim przykładzie dowodu ciągłości funkcji arcsin pozostawiamy czytelnikom, by mogli sprawdzić, na ile zrozumieli metodę. Oczywiście w dowodzie należy skorzystać z ciągłości funkcji tangens, ale to ostatnie stwierdzenie wynika z twierdzenia o ciągłości ilorazu. ■

15. Dla każdej liczby rzeczywistej $a > 0$ funkcja wykładnicza a^x jest ciągła w każdym punkcie prostej rzeczywistej. Wynika to z tego, że $a^x = e^{x \ln a}$, twierdzeń o ciągłości iloczynu i złożenia oraz ciągłości funkcji wykładniczej o podstawie e i ciągłości identyczności oraz funkcji stałej. ■

Z tych przykładów wynika, że każda funkcja, którą można zdefiniować „wzorem" używając standardowych funkcji, jest ciągła w całej swojej dziedzinie, np. $\exp\left(\dfrac{\sin(x^2-12x+2)}{\operatorname{tg}(\cos x + \ln x)}\right) - \sin\sqrt{x^4-113}$. Wynika to z wielokrotnego stosowania twierdzeń o ciągłości złożenia, sumy, różnicy, iloczynu i ilorazu. Mogłoby więc powstać wrażenie, że wszystkie funkcje są ciągłe. Tak jednak nie jest. Podamy poniżej kilka przykładów.

Przykłady funkcji nieciągłych

16. $\operatorname{sgn}(x) = \dfrac{|x|}{x}$ dla $x \neq 0$ oraz $f(0) = 0$, ta funkcja jest ciągła w każdym punkcie $p \neq 0$, bo wtedy jest stała w pewnym przedziale otwartym zawierającym p, w punkcie 0 ta funkcja jest nieciągła, bowiem jej granica prawostronna jest w tym punkcie równa 1, lewostronna jest równa -1, więc funkcja sgn (znak liczby) nie ma granicy w punkcie 0. ■

17. Niech $f(x) = \sin\dfrac{1}{x}$ dla $x \neq 0$, $f(0) = 0$. Funkcja tak zdefiniowana nie ma granicy w punkcie 0, więc nie jest w tym punkcie ciągła. We wszystkich innych punktach jest ciągła jako złożenie funkcji ciągłej sinus z funkcją ciągłą $\dfrac{1}{x}$. ■

18. Niech $f(x) = 1$ dla $x \neq 0$ i $f(0) = 0$. Funkcja ta jest nieciągła w punkcie 0, choć ma w tym punkcie granicę, jednak ta granica nie jest równa wartości funkcji w punkcie 0. W innych punktach p funkcja jest ciągła, bo jest stała na pewnym przedziale otwartym zawierającym punkt p. Oczywiście można uznać ten przykład za sztuczny. ■

19. Niech $V(t)$ oznacza objętość jednego kilograma wody w temperaturze t, ciśnienie jest stałe, tzw. normalne i niezależne od temperatury. Ze szkolnych lekcji fizyki wiadomo, że funkcja V ma nieciągłość w punkcie 0 tj. w temperaturze, w której następuje przejście ze stanu ciekłego w stały lub odwrotnie, zresztą w punkcie 0 funkcja jest z punktu widzenia fizyki niezdefiniowana, ze względu na zmianę stanu skupienia. Granice jednostronne istnieją: prawostronna jest mniejsza niż lewostronna (dlatego lód pływa w wodzie wystając z niej). Przykład ten podajemy po to, by czytelnicy tego tekstu zdawali sobie sprawę, że w niektórych sytuacjach pojawiają się funkcje nieciągłe w naturalnych sposób. ■

20. Niech $f(x) = 1$, jeśli liczba x jest wymierna, tj. x jest ilorazem dwu liczb całkowitych i niech $f(x) = 0$, jeśli x jest liczbą niewymierną, np. jeśli $x = \frac{m}{n}\sqrt{2}$, gdzie m, n są liczbami całkowitymi różnymi od 0. Funkcja ta nie ma granicy w żadnym punkcie, bo każda liczba rzeczywista jest granicą ciągu liczb wymiernych, np. swoich przybliżeń dziesiętnych oraz granicą ciągu liczb niewymiernych. W pierwszym przypadku ciąg wartości funkcji dąży do 1, a w drugim do 0. Wobec tego funkcja ta jest nieciągła w każdym punkcie! Można udowodnić, to jest łatwe!, że $f(x) = \lim\limits_{k\to\infty}\left(\lim\limits_{n\to\infty}(\cos(k!\pi x))^{2n}\right)$, więc ta dziwna funkcja może być otrzymana w wyniku podwójnego przejścia granicznego z funkcji uważanych za podstawowe. ■

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 o przyjmowaniu wartości pośrednich

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

Dowód. Niech $x_0 = x$, $z_0 = z$. Niech c_0 będzie środkiem odcinka o końcach x_0 i z_0. Mamy $c_0 = \frac{1}{2}(x_0 + z_0)$. Są trzy możliwości $f(c_0) = C$, $f(c_0) < C$ i $f(c_0) > C$. W pierwszym przypadku przyjmujemy $y = c_0$ i kończymy dowód. W drugim przypadku przyjmujemy $x_1 = c_0$ i $z_1 = z_0$. W trzecim przypadku przyjmujemy $x_1 = x_0$ i $z_1 = c_0$. W drugim i trzecim przypadku spełnione są zależności: $f(x_1) < f(z_1)$ oraz $|z_1 - x_1| = \frac{1}{2}|z_0 - x_0|$. Powtórzymy konstrukcję. Niech $c_1 = \frac{1}{2}(x_1 + z_1)$. Jeśli $f(c_1) = C$, to kończymy dowód przyjmując $y = c_1$. W przypadku przeciwnym przyjmujemy $x_2 = c_1$ i $z_2 = z_1$, jeżeli $f(c_1) < C$ oraz $x_2 = x_1$ i $z_2 = c_1$, jeżeli $f(c_1) > C$. W obu przypadkach $f(x_2) < C < f(z_2)$ i $|z_2 - x_2| = \frac{1}{2}|z_1 - x_1| = \left(\frac{1}{2}\right)^2 |z_0 - x_0|$. Postępując w dalszym ciągu w ten sposób natrafiamy na punkt y, taki że $C = f(y)$ lub otrzymujemy dwa ciągi (x_n) i (z_n) punktów przedziału P, takie że dla każdego n spełnione są związki $|z_n - x_n| = \left(\frac{1}{2}\right)^n |z_0 - x_0|$ oraz $f(x_n) < C < f(z_n)$. Z konstrukcji wynika, że dla każdego n punkty x_{n+1} i z_{n+1} znajdują się w przedziale domkniętym o końcach x_n i z_n. Stąd wynika, że jeżeli $k > m > n$, to punkty x_m, z_m, x_k, z_k znajdują się w przedziale o końcach x_n, z_n i wobec tego odległości między każdymi dwoma z nich są mniejsze niż $|z_n - x_n| = \left(\frac{1}{2}\right)^n |z_0 - x_0|$, w szczególności $|x_k - x_m| < \left(\frac{1}{2}\right)^n |z_0 - x_0|$ i $|z_k - z_m| < \left(\frac{1}{2}\right)^n |z_0 - x_0|$. Z tego, że $\lim\limits_{n\to\infty}\left(\frac{1}{2}\right)^n = 0$, wynika, że oba ciągi (x_n) i (z_n) spełniają warunek Cauchy'ego, więc każdy z nich ma skończoną granicę. Ponieważ $\lim\limits_{n\to\infty}|z_n - x_n| = 0$, więc te granice są równe. Niech $y = \lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} z_n$. Z ciągłości funkcji f w punkcie y wynika, że $f(y) = \lim\limits_{n\to\infty} f(x_n) \le C$ oraz $C \le \lim\limits_{n\to\infty} f(z_n) = f(y)$. Z nierówności $f(y) \le C \le f(y)$ wynika, że $C = f(y)$. 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. Naśladując dowód twierdzenia o przyjmowaniu wartości pośrednich, można skracać dwukrotnie w kolejnych krokach przedział, co daje rozsądną metodę przybliżania pierwiastków.*

Twierdzenie o istnieniu pierwiastków wielomianów stopnia nieparzystego

Każdy wielomian stopnia nieparzystego, tj. funkcja 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 $a_n \neq 0$, a 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 $\lim\limits_{x\to\pm\infty} \frac{w(x)}{x^n} = \lim\limits_{x\to\pm\infty} \Big(\frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + \cdots + \frac{a_{n-1}}{x} + a_n\Big) = a_n$. 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 $\lim\limits_{x\to\infty} w(x) = \lim\limits_{x\to\infty} x^n \cdot \lim\limits_{x\to\infty} \frac{w(x)}{x^n} = +\infty \cdot a_n = +\infty$, a z drugiej strony $\lim\limits_{x\to-\infty} w(x) = \lim\limits_{x\to-\infty} x^n \cdot \lim\limits_{x\to-\infty} \frac{w(x)}{x^n} = -\infty \cdot a_n = -\infty$. 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 początku wieku dziewiętnastego udowodniono (Ruffini, Abel, Galois), 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 w fizyce i 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.

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 Weierstrassa o przyjmowaniu kresów

Załóżmy, że funkcja f jest ciągła w każdym punkcie przedziału domkniętego $[a, b]$. Wtedy w przedziale [a,b] znajdują się punkty $p,\ q$, takie że dla każdego punktu x z tego przedziału zachodzi nierówność $f(p) \leq f(x) \leq f(q)$, tzn. $f(p)$ jest najmniejszą wartością funkcji f na przedziale [a,b], zaś $f(q)$ jest

* Istnieją lepsze, ale bardziej skomplikowane.

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 $\lim_{n\to\infty} f(x_n) = M$. Z twierdzenia Bolzano – Weierstrassa wynika, że z ciągu (x_n) można wybrać podciąg zbieżny (x_{k_n}). Niech $q = \lim_{n\to\infty} x_{k_n}$. 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. Wobec tego $f(q) = \lim_{n\to\infty} f(x_{k_n}) = \lim_{n\to\infty} f(x_n) = M$. 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. ■

Następne twierdzenie poprzedzimy dwiema definicjami.

Definicja jednostajnej ciągłości

Funkcja f jest jednostajnie ciągła na zbiorze A wtedy i tylko wtedy, gdy dla każdej liczby $\varepsilon > 0$ istnieje liczba $\delta > 0$, taka że jeśli $x, y \in A$ oraz $|x-y| < \delta$, to $|f(x) - f(y)| < \varepsilon$. ■

Definicja warunku Lipschitza

Funkcja f spełnia warunek Lipschitza na zbiorze A ze stałą dodatnią $L < +\infty$ wtedy i tylko wtedy, gdy dla dowolnych $x, y \in A$ zachodzi nierówność $|f(x) - f(y)| \le L|x-y|$. ■

Jest jasne, że funkcja spełniająca warunek Lipschitza jest jednostajnie ciągła: wystarczy przyjąć, że $\delta = \frac{\varepsilon}{L}$. Jest też jasne, że funkcja jednostajnie ciągła jest ciągła. Istnieją funkcje ciągłe, które nie są jednostajnie ciągłe. Przykładem jest funkcja x^2 rozpatrywana na całej prostej. Załóżmy bowiem, że $\varepsilon = 1$ oraz że istnieje liczba $\delta > 0$, taka że jeżeli $|x-y| < \delta$, to $|x^2 - y^2| < 1 = \varepsilon$. Niech $x = \frac{1}{\delta}$ i niech $y = \frac{1}{\delta} + \frac{\delta}{2}$. Wtedy $|x^2 - \mathbf{y}^2| = 1 + \left(\frac{\delta}{2}\right)^2 > 1 = \varepsilon$, więc wbrew przypuszczeniu nie istnieje liczba δ spełniająca żądany warunek. Można wykazać, że funkcje e^x na całej prostej, $\ln x$ na półprostej $(0,+\infty)$ itp. nie są jednostajnie ciągłe, choć zmniejszenie dziedziny może zmienić sytuację. Funkcja e^x spełnia warunek Lipschitza na półprostej postaci $(-\infty, a]$ dla każdej liczby rzeczywistej a. Jeśli bowiem $y < x \le a$, to $|e^x - e^y| = e^x(1 - e^{y-x}) \le e^x(1 - (1 + (y-x))) = e^x(x-y) \le e^a|x-y|$. Funkcja liniowa $ax+b$ spełnia warunek Lipschitza ze stałą $|a|$, co wynika od razu z równości $|(ax+b) - (ay+b)| = |a||x-y|$. Funkcja x^2 rozpatrywana nie na całej prostej, lecz na przedziale $[-M, M]$, gdzie $M > 0$, spełnia warunek Lipschitza ze stałą $2M$, bowiem $|x^2 - y^2| = |x-y||x+y| \le |x-y|(|x| + |y|) \le 2M|x-y|$. Funkcja $\sqrt{x}$ rozpatrywana na całej prostej jest jednostajnie ciągła ale nie spełnia warunku Lipschitza: jeśli $x > y$, to $0 < \sqrt{x} - \sqrt{y} < \sqrt{x-y}$, co można wykazać przenosząc $\sqrt{y}$ na prawą stronę nierówności, a następnie podnosząc obie strony nierówności do kwadratu – z tej nierówności wynika jednostajna ciągłość, starczy przyjąć, że $\delta = \varepsilon^2$, z warunku Lipschitza wynikałoby, że $L \ge \frac{\sqrt{1/n} - \sqrt{0}}{1/n - 0} = \sqrt{n} \xrightarrow[n\to\infty]{} +\infty$, co oczywiście przeczy temu, że $L < +\infty$. Wykazaliśmy więc, że

$$\textit{warunek Lipschitza} \quad \Longrightarrow \quad \textit{jednostajna ciągłość} \quad \Longrightarrow \quad \textit{ciągłość}$$

oraz że żadna z tych implikacji nie może być zastąpiona równoważnością. Warto jeszcze dodać, że w definicji (otoczeniowej) ciągłości funkcji w punkcie żąda się istnienia liczby $\delta > 0$ i że takie samo żądanie występuje w definicji ciągłości jednostajnej. Różnica polega na tym, że w definicji ciągłości liczba $\delta > 0$ jest dopasowywana do punktu, w którym badana jest ciągłość i do liczby $\varepsilon > 0$, natomiast w definicji ciągłości jednostajnej $\delta > 0$ zależy tylko od ε. Podamy teraz ważne twierdzenie, które wykorzystamy w rozdziale poświęconym całkom.

Twierdzenie 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 $|f(x)-f(y)| \geq \varepsilon$. Niech x_n, y_n będą takimi liczbami z przedziału $[a, b]$, że $|x_n - y_n| < \delta = \dfrac{1}{n}$ 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 $|x_n - y_n| < \dfrac{1}{n}$, 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. ■

Teraz zajmiemy się monotonicznymi funkcjami ciągłymi. Rozpoczniemy od twierdzenia gwarantującego ciągłość funkcji monotonicznej.

Twierdzenie o ciągłości funkcji monotonicznej

Jeśli funkcja monotoniczna f określona na zbiorze $A \subset \mathbb{R}$ przekształca zbiór A *na* przedział, to jest ciągła w każdym punkcie zbioru A.

Dowód. Dla ustalenia uwagi załóżmy, że funkcja f jest niemalejąca. Jeśli $p \in A$ jest granicą ciągu (a_n) punktów zbioru A mniejszych niż p, to istnieje granica $\lim\limits_{x\to p^-} f(x) \leq f(p)$. Ponieważ dla $x \geq p$ zachodzi nierówność $f(x) \geq f(p)$, a dla $x < p$ zachodzi nierówność $f(x) \leq \lim\limits_{x\to p^-} f(x)$, więc z tego, że obrazem zbioru A jest przedział, wynika, że $\lim\limits_{x\to p^-} f(x) = f(p)$: gdyby było $\lim\limits_{x\to p^-} f(x) < f(p)$, to punkty przedziału $\big(\lim\limits_{x\to p^-} f(x), f(p)\big)$ byłyby poza obrazem zbioru A, więc nie byłby on przedziałem. Analogicznie: jeśli istnieje ciąg (a'_n) większych niż p zbieżny do p, to $f(p) = \lim\limits_{x\to p^+} f(x)$. Stąd wynika, że f jest ciągła w punkcie p. Dowód został zakończony. ■

Wniosek

Jeśli f jest funkcją monotoniczną określoną na przedziale P, to f jest ciągła w każdym punkcie wtedy i tylko wtedy, gdy f przekształca przedział P *na* pewien przedział (zdegenerowany do punktu w przypadku, gdy f jest funkcją stałą). ■

Twierdzenie 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, z są punktami przedziału 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 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 przypadku ogólnym, gdy dziedzina funkcji nie jest przedziałem funkcja odwrotna ciągła być nie musi.

6. Funkcje wypukłe

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 $f\big(tx+(1-t)y\big) \le tf(x)+(1-t)f(y)$.* 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. ■

---

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

Przykłady

1. 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. ■

2. 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 $x = y$. ■

3. Funkcja $f(x) = \sqrt{x}$ jest ściśle wklęsła – wynika to łatwo ze ścisłej wypukłości funkcji kwadratowej: nierówność $\sqrt{tx+(1-t)y} > t\sqrt{x}+(1-t)\sqrt{y}$ jest równoważna nierówności $\big(tu+(1-t)v\big)^2 < tu^2+(1-t)v^2$, 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 $a = \dfrac{f(y)-f(x)}{y-x}$ oraz $b = f(x)$. Nierówność występująca w definicji funkcji wypukłej: $f\big(tx+(1-t)y\big) \leq tf(x)+(1-t)f(y)$ to po prostu 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 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 $f\left(\dfrac{x+y}{2}\right) \leq \dfrac{f(x)+f(y)}{2}$, ściśle wypukła, gdy ta nierówność jest ostra w każdym przypadku, w którym $x \neq y$.

Dowód. Jeśli f jest wypukła, to przyjmując w definicji wypukłości $t = \dfrac{1}{2}$ 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 $f\left(\dfrac{x+y}{2}\right) \leq \dfrac{f(x)+f(y)}{2}$. Ponieważ nierówność ta zachodzi dla dowolnych punktów x, y zbioru P, więc możemy zastąpić punkt y środkiem odcinka łączącego punkty x, y. Mamy $\dfrac{1}{2}\left(x+\dfrac{x+y}{2}\right) = \dfrac{3}{4}x+\dfrac{1}{4}y$. Wobec tego mamy też

$$f\left(\frac{3}{4}x+\frac{1}{4}y\right) \leq \frac{1}{2}\left(f(x)+f\big(\frac{x+y}{2}\big)\right) \leq \frac{1}{2}\left(f(x)+\frac{f(x)+f(y)}{2}\right) = \frac{3}{4}f(x)+\frac{1}{4}f(y).$$

Wykazaliśmy więc, że nierówność definiująca wypukłość ma miejsce w przypadku $t = \dfrac{3}{4}$. Stosując to samo rozumowanie do punktów $\dfrac{x+y}{2}$ oraz y otrzymujemy nierówność $f\big(\frac{1}{4}x+\frac{3}{4}y\big) \leq \frac{1}{4}f(x)+\frac{3}{4}f(y)$, a więc nierówność z definicji wypukłości w przypadku $t = \dfrac{1}{4}$. Rozważając kolejno pary punktów x i

$\dfrac{3}{4}x+\dfrac{1}{4}y$, $\dfrac{3}{4}x+\dfrac{1}{4}y$ i $\dfrac{1}{2}(x+y)$, $\dfrac{1}{2}(x+y)$ i $\dfrac{1}{4}x+\dfrac{3}{4}y$ oraz $\dfrac{1}{4}x+\dfrac{3}{4}y$ i y otrzymujemy nierówność kolejno dla $t=\dfrac{7}{8}$, $t=\dfrac{5}{8}$, $t=\dfrac{3}{8}$ i $t=\dfrac{1}{8}$. Otrzymaliśmy nierówność dla 7 wartości t: $\dfrac{1}{8}$, $\dfrac{2}{8}$, $\dfrac{3}{8}$, $\dfrac{4}{8}$, $\dfrac{5}{8}$, $\dfrac{6}{8}$, $\dfrac{7}{8}$. W taki sam sposób możemy otrzymać nierówność w przypadku $t=\dfrac{k}{16}$, potem w przypadku $t=\dfrac{k}{32}$ itd. Teraz skorzystamy z ciągłości funkcji f. Każda liczba $t\in(0,1)$ jest granicą ciągu (t_n) liczb postaci $\dfrac{k}{2^m}\in(0,1)$. 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 liczby $x,y\in P$ oraz $t\in(0,1)$, takie że $f(tx+(1-t)y)=tf(x)+(1-t)f(y)$. Załóżmy, że $0<s<t<1$. Mamy wtedy
$tf(x)+(1-t)f(y)=f(tx+(1-t)y)=f\left(\frac{t-s}{1-s}x+\frac{1-t}{1-s}(sx+(1-s)y)\right)\le$

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

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

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 i ich omówienie wykracza znacznie poza ramy tej książki.

Przykłady cd.

4. Funkcja wykładnicza o podstawie dodatniej i różnej od 1 jest ściśle wypukła. Wykażemy, że ma miejsce nierówność $a^{(x+y)/2}\le\dfrac{1}{2}\left(a^x+a^y\right)$, przy czym staje się ona równością jedynie wtedy, gdy $x=y$.Wynika to stąd, że $a^x+a^y-2a^{(x+y)/2}=\left(a^{x/2}-a^{y/2}\right)^2$. Stąd teza wynika natychmiast. ■

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

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

która wynika natychmiast ze ścisłej wypukłości funkcji wykładniczej o podstawie e. (zob. przykład 4.) ■

6. Funkcja sinus jest ściśle wypukła na przedziale $[-\pi,0]$ i ściśle wklęsła na przedziale $[0,\pi]$. Dla dowodu wystarczy wykazać, że jeśli $-\pi\le x<y\le 0$, to $\sin\dfrac{x+y}{2}<\dfrac{1}{2}\left(\sin x+\sin y\right)$ oraz że jeśli $0 \le x < y \le \pi$, to $\sin\frac{x+y}{2} > \frac{1}{2}(\sin x + \sin y)$. Ponieważ $\sin(-x) = -\sin x$, więc wystarczy wykazać jedną z tych nierówności. Załóżmy, że $0 \le x < y \le \pi$. 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 $-\frac{\pi}{2} \le \frac{x-y}{2} < 0$, więc $0 \le \cos\frac{x-y}{2} < 1$. ■

7. 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 $0 \le x < y < \frac{\pi}{2}$. Wykorzystamy znany wzór: $\operatorname{tg}\alpha - \operatorname{tg}\beta = \frac{\sin(\alpha-\beta)}{\cos\alpha\cos\beta}$. Mamy

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

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

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

8. 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 $|tx + (1-t)y| \le t|x| + (1-t)|y|$, przy czym równość zachodzi wtedy i tylko wtedy, gdy $xy \ge 0$. ■

9. 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 $x \ne y$. Mamy w tej sytuacji $e^{|x+y|/2} \le e^{(|x|+|y|)/2} \le \frac{1}{2}\left(e^{|x|} + e^{|y|}\right)$, przy czym jeśli pierwsza nierówność staje się równością, to $xy \ge 0$ i wobec tego, że $x \ne y$, ma miejsce nierówność $|x| \ne |y|$ i wobec tego druga nierówność musi być ostra (funkcja wykładnicza jest ściśle wypukła). Dowód został zakończony. ■

10. Funkcja $|x| + |x-1| + |x-2| + |x-3|$ jest wypukła jako suma czterech funkcji wypukłych, zob. przykład 8. 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.

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

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

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 (por. przykład 3.). ■

12. 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

* 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.

na przejechanie x kilometrów. Niech h oznacza dowolną liczbę dodatnią i niech $x < y$. Wtedy $p(x+h)-p(x) \geq p(y+h)-p(y)$. Wykażemy, że ten warunek, w przypadku funkcji 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 $x < y$. Jeśli $r < s$, to przyjmujemy $x = r, h = \frac{1}{2}(s-r), y = \frac{1}{2}(s+r)$. Nierówność $p(x+h)-p(x) \geq p(y+h)-p(y)$ to nierówność $p\big(\frac{1}{2}(s+r)\big) - p(r) \geq p(s) - p\big(\frac{1}{2}(s+r)\big)$, czyli $p\big(\frac{1}{2}(s+r)\big) \geq \frac{1}{2}\,(p(r)+p(s))$, 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:

$$\frac{p(v)-p(u)}{v-u} \geq \frac{p(w)-p(u)}{w-u}, \qquad \frac{p(u)-p(v)}{u-v} \geq \frac{p(w)-p(v)}{w-v} \qquad \text{i} \qquad \frac{p(u)-p(w)}{u-w} \geq \frac{p(v)-p(w)}{v-w}.$$

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. Wykazaliśmy tam, że w przypadku funkcji wklęsłej p i trzech punktów jej dziedziny $u < v < w$ zachodzą nierówności:

$$\frac{p(v)-p(u)}{v-u} \geq \frac{p(w)-p(u)}{w-u}, \qquad \frac{p(u)-p(v)}{u-v} \geq \frac{p(w)-p(v)}{w-v} \qquad \text{i} \qquad \frac{p(u)-p(w)}{u-w} \geq \frac{p(v)-p(w)}{v-w}.$$

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 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 w następnym rozdziale, 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 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 $\dfrac{1}{n}$, 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.

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 $x_1,x_2,\dots,x_n,x_{n+1}$ będą dowolnymi argumentami funkcji f , a $p_1,p_2,\dots,p_n,p_{n+1}$ 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 $p_1,p_2,\dots,p_n,p_{n+1}$ są dodatnie. Niech $p'_n=p_n+p_{n+1}$ i $x'_n=\dfrac{p_nx_n+p_{n+1}x_{n+1}}{p_n+p_{n+1}}=\dfrac{p_n}{p'_n}x_n+\dfrac{p_{n+1}}{p'_n}x_{n+1}$. 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

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

– 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.

Nierówność między klasycznymi średnimi – nierówność Cauchy'ego

Dla dowolnych liczb dodatnich $a_1, a_2, \ldots, a_n$ zachodzi: $\sqrt[n]{a_1 a_2 \ldots a_n} \le \dfrac{a_1 + a_2 + \cdots + a_n}{n}$.
Nierówność ta staje się równością wtedy i tylko wtedy, gdy $a_1 = a_2 = \ldots = a_n$.

Dowód. Zastosujemy nierówność Jensena do funkcji wypukłej $-\ln$ (przykład 5.). Mamy

$$-\ln \frac{a_1 + a_2 + \cdots + a_n}{n} = -\ln\left(\frac{1}{n}a_1 + \frac{1}{n}a_2 + \cdots + \frac{1}{n}a_n\right) \le$$
$$\le \frac{1}{n}(-\ln)(a_1) + \frac{1}{n}(-\ln)(a_2) + \cdots + \frac{1}{n}(-\ln)(a_n) = -\frac{1}{n}\left(\ln a_1 + \ln a_2 + \ldots + \ln a_n\right).$$

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

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

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.

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 $\dfrac{1}{p} + \dfrac{1}{q} = 1$ 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 $\dfrac{1}{p} + \dfrac{1}{q} = 1$ wynika, że $p > 1$, $q > 1$ oraz $\dfrac{1}{q} = 1 - \dfrac{1}{p} = \dfrac{p-1}{p}$. Z tego, że $p > 1$ wnioskujemy od razu, że funkcja x^p jest ściśle wypukła, bo jej pochodna $(x^p)' = px^{p-1}$ jest ściśle rosnąca. 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 $p_j = \dfrac{b_j^{p/(p-1)}}{\sum b_i^{p/(p-1)}}$. Mamy wtedy

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

a ta nierówność jest równoważna dowodzonej. Dla $p = q = 2$ otrzymujemy nierówność Schwarza, czyli 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 12

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 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:

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

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

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. ■

Definicja pochodnej

Załóżmy, że funkcja f jest określona w dziedzinie zawierającej przedział otwarty o środku p oraz że istnieje granica $\lim\limits_{h\to 0}\frac{f(p+h)-f(p)}{h}$. Granicę tę nazywamy pochodną funkcji f w punkcie p i oznaczamy symbolem $f'(p)$ lub $\frac{df}{dx}(p)$. Jeśli pochodna jest skończona, to mówimy, że funkcja f jest różniczkowalna w punkcie p. Funkcję liniową przypisującą liczbie h liczbę $f'(p)h$ nazywamy różniczką funkcji f w punkcie p i oznaczamy symbolem $df(p)$, a wartość tej funkcji liniowej w punkcie h oznaczamy przez $df(p)(h)$ lub $df(p)h$.■

Definicja prostej stycznej do wykresu funkcji

Załóżmy, że funkcja f ma pochodną w punkcie p oraz że jest ciągła w punkcie p.** Jeśli pochodna $f'(p)$ jest skończona, to mówimy, że prostą styczną do wykresu funkcji f w punkcie $(p,f(p))$ jest prosta, której współczynnik kierunkowy jest równy $f'(p)$ przechodząca przez punkt $(p,f(p))$. Jeśli $f'(p)=\pm\infty$, to mówimy, że styczną do wykresu w punkcie $(p,f(p))$ jest prosta pionowa przechodząca przez ten punkt, czyli prosta o równaniu $x=p$. ■

Jest jasne, że jeśli funkcja f jest różniczkowalna w punkcie p, to prosta styczna do wykresu tej funkcji w punkcie $(p,f(p))$ ma równanie $y=f'(p)(x-p)+f(p)$. Później przekonamy się, że próby przenoszenia definicji stycznej do okręgu na przypadek stycznej do wykresu funkcji nie mają większego sensu, bo prowadzą do wyników niezgodnych z intuicją. Motywy wprowadzenia podanej przez nas definicji są następujące. Jeśli $|h|\neq 0$ jest niedużą liczbą, to współczynnik kierunkowy prostej przechodzącej przez punkty $(p,f(p))$ oraz $(p+h,f(p+h))$ jest równy *ilorazowi różnicowemu* $\frac{f(p+h)-f(p)}{h}$, który jest w przybliżeniu równy $f'(p)$. Prosta styczna jest więc „granicą prostych” przechodzących przez punkt $(p,f(p))$ i jeszcze jeden punkt wykresu leżący blisko wymienionego. Nie zamierzamy tu precyzować pojęcia „granicy prostych”, bo używamy go jedynie w tym miejscu i to jedynie w celu wyjaśnienia,

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

** Wykażemy później, że jeśli pochodna $f'(p)$ funkcji f w punkcie p jest skończona, czyli że f jest różniczkowalna w punkcie p, to funkcja f jest ciągła w punkcie p, więc w tym przypadku nie ma potrzeby dodatkowo zakładać ciągłości funkcji w punkcie p.

skąd się taka definicja stycznej bierze. Mówiąc jeszcze mniej dokładnie: prosta styczna ma przylegać możliwie ściśle do wykresu w pobliżu punktu $(p, f(p))$, daleko od tego punktu wykres i styczna mogą się rozchodzić. Podamy teraz kilka przykładów.

Przykłady

1. Niech $f(x) = ax + b$. W tym przypadku iloraz różnicowy $\frac{f(p+h)-f(p)}{h} = \frac{a(p+h)-ap}{h} = a$ jest niezależny od h, zresztą również od p. Wobec tego pochodna funkcji liniowej $ax + b$ jest równa a. Z tego wynika, że prostą styczną do prostej $y = ax + b$ jest ona sama, co nie powinno dziwić, bo ona sama do siebie przylega najlepiej ze wszystkich prostych. Często stosowany jest zapis $(ax + b)' = a$. ■

2. Niech $f(x) = x^2$ i niech $p \in \mathbb{R}$. Bez trudu stwierdzamy, że $\frac{f(p+h)-f(p)}{h} = 2p + h \to_{h\to 0} 2p$, co oznacza, że pochodną funkcji f w punkcie p jest $2p$. Zwykle piszemy $(x^2)' = 2x$. Ponieważ $f'(0) = 0$, więc styczna do wykresu tej funkcji w punkcie $(0, 0)$ jest pozioma. Jeśli natomiast $p = 10$, to współczynnik kierunkowy stycznej do wykresu jest równy 20, więc styczna w punkcie $(20, 400)$ jest prawie pionowa. ■

3. Niech $f(x) = x^3$. Mamy $\frac{f(p+h)-f(p)}{h} = 3p^2 + 3ph + h^2 \to_{h\to 0} 3p^2$, co oznacza, że pochodną funkcji f w punkcie p jest $3p^2$, tzn. $(p^3)' = 3p^2$. I tym razem $f'(0) = 0$, więc styczna do wykresu funkcji f w punkcie $(0, f(0)) = (0, 0)$ jest pozioma. Jednak w tym przypadku wykres nie leży po jednej stronie stycznej, lecz przechodzi z jednej strony tej prostej na drugą. Pochodna jest dodatnia z jednym wyjątkiem: $f'(0) = 0$. Bez trudu można stwierdzić, że styczna do wykresu tej funkcji w każdym punkcie, z wyjątkiem punktu $(0, 0)$, przecina wykres w jeszcze jednym punkcie*, więc również w tym przypadku nie jest prawdą, że styczna ma z wykresem funkcji dokładnie jeden punkt wspólny. ■

4. Teraz zajmiemy się funkcją $f(x) = |x|$. Jeśli $p > 0$ i $|h| < p$, to $\frac{f(p+h)-f(p)}{h} = \frac{p+h-p}{h} = =1 \xrightarrow[h\to 0]{} 1$, co oznacza, że pochodną funkcji f w punkcie p jest 1. W taki sam sposób pokazać można, że $f'(p) = -1$ dla każdej liczby $p < 0$. Pozostał jeszcze jeden przypadek do rozważenia, mianowicie $p = 0$. Jeśli $h > 0$, to $\frac{f(0+h)-f(0)}{h} = 1$ i wobec tego $\lim\limits_{x\to 0^+} \frac{f(0+h)-f(0)}{h} = 1$. Analogicznie $\lim\limits_{x\to 0^-} \frac{f(0+h)-f(0)}{h} = -1$. Z tych dwu równości wynika od razu, że nie istnieje granica $\lim\limits_{x\to 0} \frac{f(0+h)-f(0)}{h}$, czyli że funkcja $|x|$ pochodnej w punkcie 0 nie ma, chociaż jest ciągła – ma ona w tym punkcie pochodne jednostronne, ale są one różne. Na wykresie funkcji jest to widoczne, w punkcie $(0, 0)$ wykres się załamuje, można powiedzieć, że wykres ma w tym punkcie „ostrze". Zauważmy, że rezultaty tych rozważań można opisać wzorem $(|x|)' = \frac{|x|}{x}$. ■

5. Podamy teraz przykład świadczący o tym, że istnieją funkcje ciągłe, które przynajmniej w niektórych punktach nie mają pochodnych jednostronnych. Tym studentom, których te przykłady męczą, zalecamy pominięcie tego punktu w pierwszym czytaniu i ewentualny powrót do niego później. Warto też spróbować sporządzić szkic wykresu funkcji, co może ułatwić zrozumienie sytuacji. Przechodzimy do szczegółów. Niech $f(x) = x \sin \frac{1}{x}$ dla $x \neq 0$ oraz $f(0) = 0$. Z oczywistej nierówności $|f(x)| \leq |x|$ wynika, że $\lim\limits_{x\to 0} f(x) = 0 = f(0)$, a to znaczy, że funkcja f jest ciągła w punkcie 0. Ciągłość w innych punktach jest oczywistym wnioskiem z twierdzenia o operacjach na

* Czytelnik zechce sprawdzić w jakim, to pomaga w zrozumieniu tekstu!

funkcjach ciągłych i twierdzenia o ciągłości złożenia dwu funkcji. Z twierdzeń, które udowodnimy niedługo wyniknie, że funkcja ta ma pochodną skończoną w każdym punkcie z wyjątkiem punktu 0. Wykażemy teraz, że funkcja ta nie ma pochodnej w punkcie 0, dokładniej, że w tym punkcie funkcja nie ma pochodnej prawostronnej. Jeśli $h > 0$, to $\dfrac{f(x+h)-f(x)}{h} = \sin\dfrac{1}{h}$. Wykazaliśmy wcześniej (poprzedni rozdział, przykłady granic, punkt 6), że funkcja ta nie ma granicy prawostronnej: $f\big(\frac{1}{2n\pi}\big) = 0$ oraz $f\big(\frac{1}{2n\pi+\pi/2}\big) = 1$. Widzimy, więc że dla każdej liczby naturalnej n punkt $\big(\frac{1}{2n\pi}, 0\big)$ leży na wykresie funkcji, co oznacza, że styczną do wykresu funkcji w punkcie $(0,0)$ powinna być pozioma oś układu współrzędnych. Jednakże dla każdej liczby naturalnej n punkt $\big(\frac{1}{2n\pi+\pi/2}, \frac{1}{2n\pi+\pi/2}\big)$ leży na wykresie funkcji, więc styczną powinna być prosta, na której te punkty leżą, czyli prosta o równaniu $y = x$ – styczną ma być prosta najdokładniej „przylegająca" do wykresu. Podobnie można uzasadniać, że styczną do wykresu tej funkcji w punkcie $(0,0)$ powinna być prosta o równaniu $y = kx$, gdzie k jest dowolną liczbą z przedziału $[-1,1]$ – na każdej takiej prostej znajdują się punkty leżące na wykresie funkcji f, tworzące ciąg zbieżny do 0. Można powiedzieć, że wykres funkcji $x\sin\dfrac{1}{x}$ oscyluje między prostymi $y = x$ oraz $y = -x$ i do żadnej z nich ani do żadnej leżącej w kącie przez nie wyznaczonym w punkcie $(0,0)$ nie „przylega". ■

6. Obliczymy teraz pochodną funkcji wykładniczej. Niech $f(x) = e^x$. Przypomnieć wypada, że $\lim\limits_{h\to 0}\dfrac{e^{x+h}-e^x}{h} = e^x$ – wykazaliśmy, że ta równość ma miejsce w rozdziale I, punkt 19, podpunkt m. Wobec tego pochodną w punkcie x funkcji wykładniczej o podstawie e jest liczba e^x, czyli $(e^x)' = e^x$. ■

7. Następną bardzo ważną funkcją jest logarytm naturalny. Znajdziemy jej pochodną. Niech $f(x) = \ln x$ dla każdej liczby dodatniej x. Przypomnijmy, że $\lim\limits_{x\to 0}\frac{\ln(1+x)}{x} = 1$ – wzór ten wykazaliśmy w rozdziale I, punkt 20, wzór (20.2). Mamy więc dla $x > 0$ następującą równość*: $\lim_{h\to 0}\frac{\ln(x+h)-\ln x}{h} = \lim_{h\to 0}\frac{\ln(1+x/h)}{x/h}\cdot\frac{1}{x} = 1\cdot\frac{1}{x} = \frac{1}{x}$. Znaczy to, że pochodną logarytmu naturalnego w punkcie x jest liczba $\frac{1}{x}$, czyli $(\ln x)' = \frac{1}{x}$. ■

8. Ostatnią z krótkiego cyklu „najważniejszych" funkcji elementarnych jest sinus. Przypomnijmy, że $\lim\limits_{x\to 0}\frac{\sin x}{x} = 1$ – ta równość została wykazana poprzednio. Z niej wynika, że $\lim\limits_{h\to 0}\frac{\sin(x+h)-\sin x}{h} = \lim\limits_{h\to 0}\frac{2\sin\frac{h}{2}\cos(x+\frac{h}{2})}{h} = \lim\limits_{h\to 0}\frac{\sin(h/2)}{h/2}\cdot\cos(x+\frac{h}{2}) = \cos x$. Udało się więc nam wykazać, że pochodną funkcji sinus w punkcie x jest liczba $\cos x$, czyli że zachodzi wzór $(\sin x)' = \cos x$. ■

Następne wzory wyprowadzimy po podaniu reguł, według których obliczane są pochodne. Nie będziemy w tym przypadku zajmować się pochodnymi nieskończonymi, bowiem w zastosowaniach będą nam potrzebne na ogół pochodne skończone.

Twierdzenie o arytmetycznych własnościach pochodnej

Załóżmy, że funkcje f i g są różniczkowalne w punkcie p. Wtedy funkcje $f \pm g$, $f\cdot g$ i, jeśli $g(p) \neq 0$, to również $\dfrac{f}{g}$ są różniczkowalne w punkcie p i zachodzą wzory:

$(f+g)'(x) = f'(x) + g'(x)$, $\qquad (f-g)'(x) = f'(x) - g'(x)$,

$(f\cdot g)' = f'(x)g(x) + f(x)g'(x)$, $\qquad \left(\dfrac{f}{g}\right)'(x) = \dfrac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$.

* Przypomnijmy, że $\ln(x+h)-\ln x = \ln\frac{x+h}{x} = \ln\big(1+\frac{h}{x}\big)$

Dowód twierdzenia o arytmetycznych własnościach pochodnej

Mamy $f'(p) = \lim\limits_{h\to 0} \dfrac{f(p+h)-f(p)}{h}$ oraz $g'(p) = \lim\limits_{h\to 0} \dfrac{g(p+h)-g(p)}{h}$ i wiemy, że te pochodne są skończone. Stąd i z twierdzenia o arytmetycznych własnościach granicy funkcji wynika, że

$$\lim_{h\to 0} \frac{f(p+h)+g(p+h)-f(p)-g(p)}{h} = \lim_{h\to 0} \frac{f(p+h)-f(p)}{h} + \lim_{h\to 0} \frac{g(p+h)-g(p)}{h} = f'(p)+g'(p)\,.$$

Udowodniliśmy więc twierdzenie o pochodnej sumy dwu funkcji różniczkowalnych. W identyczny sposób dowodzimy twierdzenie pochodnej różnicy funkcji różniczkowalnych. Zajmiemy się teraz iloczynem funkcji różniczkowalnych. Tym razem skorzystamy z udowodnionego wcześniej twierdzenia o ciągłości funkcji różniczkowalnej. Mamy

$$\lim_{h\to 0} \frac{f(p+h)g(p+h)-f(p)g(p)}{h} = \lim_{h\to 0} \frac{\big(f(p+h)-f(p)\big)\cdot g(p+h) + f(p)\big(g(p+h)-g(p)\big)}{h} =$$
$$= \lim_{h\to 0} \frac{f(p+h)-f(p)}{h} \cdot \lim_{h\to 0} g(p+h) + f(p)\cdot \lim_{h\to 0} \frac{g(p+h)-g(p)}{h} = f'(p)g(p)+f(p)g'(p)\,.$$

Teraz kolej na iloraz. Mamy teraz dodatkowe założenie: $g(p)\neq 0$. Wynika stąd, że istnieje liczba $\delta>0$, taka że jeśli $|h|<\delta$, to $|g(p+h)-g(p)|<|g(p)|$. Wnioskujemy stąd, że $|g(p+h)|<|g(p)|$, czyli że liczby $g(p)$ i $g(p+h)$ leżą po tej samej stronie zera, w szczególności $g(p+h)\neq 0$. Mamy zatem

$$\lim_{h\to 0} \frac{\dfrac{f(p+h)}{g(p+h)} - \dfrac{f(p)}{g(p)}}{h} = \lim_{h\to 0} \frac{f(p+h)g(p)-f(p)g(p+h)}{hg(p+h)g(p)} =$$
$$= \lim_{h\to 0} \frac{f(p+h)g(p)-f(p)g(p)-\big(f(p)g(p+h)-f(p)g(p)\big)}{hg(p+h)g(p)} =$$
$$= \lim_{h\to 0} \frac{\dfrac{f(p+h)-f(p)}{h}g(p) - f(p)\dfrac{g(p+h)-g(p)}{h}}{g(p+h)g(p)} = \frac{f'(p)g(p)-f(p)g'(p)}{g(p)^2}\,.$$

Dowód został zakończony. ■

Twierdzenie o pochodnej złożenia

Załóżmy, że funkcja g jest różniczkowalna w punkcie p, zaś funkcja f, określona na zbiorze zawierającym wszystkie wartości funkcji g, jest różniczkowalna w punkcie $g(p)$. Wtedy złożenie tych funkcji $f\circ g$ jest różniczkowalne w punkcie p i zachodzi wzór:

$$(f\circ g)'(x) = f'(g(x))g'(x)\,.$$

Wprowadzimy oznaczenie $y=g(x)$. Stosując to oznaczenie można napisać $(f\circ g)'(x) = f'(y)g(x)$ lub $\dfrac{d(f\circ g)}{dx}(x) = \dfrac{df}{dy}(g(x))\cdot\dfrac{dg}{dx}(x) = \dfrac{df}{dy}(y)\cdot\dfrac{dg}{dx}(x)$ lub krócej $\dfrac{d(f\circ g)}{dx} = \dfrac{df}{dy}\cdot\dfrac{dg}{dx}$. Często wzór ten zapisywany jest w postaci $\dfrac{d(f\circ g)}{dx} = \dfrac{df}{dg}\cdot\dfrac{dg}{dx}$ lub, po oznaczeniu $z=f(y)$, jako $\dfrac{dz}{dx} = \dfrac{dz}{dy}\cdot\dfrac{dy}{dx}$. W literaturze anglojęzycznej nosi nazwę „the Chain Rule", czego oczywistym motywem jest jego ostatnia postać, zwłaszcza jeśli zastosujemy go nie w przypadku złożenia dwu funkcji, lecz większej ich liczby – wtedy łańcuch staje się bardziej widoczny.

Dowód twierdzenia o pochodnej złożenia

Znów mamy do czynienia z dwiema funkcjami różniczkowalnymi: f w punkcie $q=g(p)$ oraz g w punkcie p. Niech $r_g(h) = \dfrac{g(p+h)-g(p)-g'(p)h}{h}$ i niech $r_g(0)=0$. Różniczkowalność funkcji g w punkcie p równoważna jest ciągłości funkcji r_g w punkcie 0. Prawdziwa jest zatem równość: $g(p+h) = g(p)+g'(p)h+r_g(h)h$. Przyjmijmy teraz, że $r_f(H) = \dfrac{f(g(p)+H)-f(g(p))}{H}$ oraz $r_f(0)=0$. Tak jak w przypadku funkcji g funkcja f jest różniczkowalna w punkcie $g(p)$ wtedy i tylko wtedy, gdy funkcja r_f jest ciągła w punkcie 0. Również w tym przypadku zachodzi też wzór: $f(g(p)+H) = f(g(p)) + f'(g(p))H + r_f(H)H$. Gotowi jesteśmy do „wydzielenia części liniowej złożenia" $f\circ g$ w otoczeniu punktu p: $f(g(p+h)) = f(g(p)+g'(p)h + r_g(h)h) =$

$$= f(g(p)) + f'(g(p))\Big(g'(p)h + r_g(h)h\Big) + r_f\Big(g'(p)h + r_g(h)h\Big)\Big(g'(p)h + r_g(h)h\Big) =$$

$$= f(g(p)) + f'(g(p))g'(p)h + h\cdot\Big[r_g(h) + r_f\Big(g'(p)h + r_g(h)h\Big)\Big(g'(p) + r_g(h)\Big)\Big].$$

Jasne jest, że granicą wyrażenia znajdującego się w nawiasie kwadratowym przy $x\to 0$ jest liczba 0. Stąd zaś wynika od razu, zob. twierdzenie charakteryzujące pochodną jako współczynnik wielomianu stopnia ≤ 1 najlepiej przybliżającego funkcję, że pochodną funkcji $f\circ g$ w punkcie p jest liczba $f'(g(p))g'(p)$. Dowód został zakończony. ■

Twierdzenie o pochodnej funkcji odwrotnej

Załóżmy, że funkcja f jest różniczkowalna w punkcie p, że $f'(p)\neq 0$, że funkcja f ma funkcję odwrotną oraz że funkcja f^{-1} odwrotna do f jest ciągła w punkcie $q=f(p)$. Wtedy funkcja f^{-1} jest różniczkowalna w punkcie q i zachodzi wzór $\left(f^{-1}\right)'(q) = \dfrac{1}{f'(p)}$.

Wzór na pochodną funkcji odwrotnej można zapisać również w postaci $\left(f^{-1}\right)'(f(p)) = \dfrac{1}{f'(p)}$ lub w postaci $\left(f^{-1}\right)'(q) = \dfrac{1}{f'(f^{-1}(q))}$. Piszemy też $\dfrac{dy}{dx} = \dfrac{dx}{dy}$, oznaczywszy uprzednio $y=f(x)$. Ten ostatni zapis, zwłaszcza w połączeniu z wzorem $\dfrac{dz}{dx} = \dfrac{dz}{dy}\dfrac{dy}{dx}$ sugeruje, że symbol $\dfrac{dy}{dx}$ można traktować jak ułamek. Trzeba jednak uważać, bo nie oznacza on ułamka, lecz pochodną i posługiwać się analogiami z ilorazem jedynie w zakresie dopuszczonym podawanymi twierdzeniami. Można np napisać wzór

$$\tfrac{dg}{dx} + \tfrac{dh}{dx} = \tfrac{d(g+h)}{dx}$$

– oznacza on, że pochodna sumy dwu funkcji względem zmiennej x jest równa sumie ich pochodnych względem tej samej zmiennej x. Natomiast wzoru $\dfrac{df}{dy} + \dfrac{dg}{dx} = \dfrac{df\cdot dx + dg\cdot dy}{dy\cdot dx}$ napisać *nie można* np. dlatego, że jego prawa strona nie ma sensu, nie jest zdefiniowana. Później rozważać będziemy pochodne wyższych rzędów i tam sytuacja będzie jeszcze bardziej skomplikowana.

Dowód twierdzenia o pochodnej funkcji odwrotnej

Tym razem wiemy, ze funkcja f jest różniczkowalna w punkcie p, że $f'(p)\neq 0$ oraz że funkcja f^{-1} odwrotna do funkcji f jest ciągła w punkcie $q=f(p)$. Wystarczy teraz wykazać, że

$$\lim_{h\to 0}\frac{f^{-1}(q+h) - f^{-1}(q)}{h} = \frac{1}{f'(p)}.$$

Oznaczmy $H = f^{-1}(q+h) - f^{-1}(q)$. Oczywiście H zależy od h. Z ciągłości funkcji f^{-1} w punkcie q wynika od razu, że $\lim\limits_{h\to 0} H = 0$. Zachodzi też równość

$$h = q+h-q = f(f^{-1}(q+h)) - f(f^{-1}(q)) = f(f^{-1}(q)+H) - f(f^{-1}(q)) = f(p+H) - f(p).$$

Z tej równości i z poprzednich wynika, że $\lim\limits_{h\to 0}\dfrac{f^{-1}(q+h) - f^{-1}(q)}{h} = \lim\limits_{H\to 0}\dfrac{H}{f(p+H) - f(p)} = \dfrac{1}{f'(p)}$.

Dowód został zakończony. ■

Ostatnim z tego cyklu twierdzeń służących do obliczania pochodnych jest

Twierdzenie o pochodnej szeregu potęgowego

Jeśli szereg potęgowy $\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ ma dodatni promień zbieżności, to *wewnątrz* przedziału zbieżności suma tego szeregu jest funkcją różniczkowalną i zachodzi wzór:

$$\left(\sum_{n=0}^{\infty} a_n(x-x_0)^n\right)' = \sum_{n=1}^{\infty} na_n(x-x_0)^{n-1} .$$

Wypada przestrzec, że szeregów na ogół nie wolno różniczkować w taki sposób, jak się różniczkuje sumy skończone. K.Weierstrass wykazał, że np. funkcja zdefiniowana jako suma szeregu $\sum\limits_{n=1}^{\infty} \frac{1}{2^n}\cos(7^n\pi x)$ jest ciągła na całej prostej i nie ma skończonej pochodnej w żadnym punkcie, chociaż każdy wyraz tego szeregu ma pochodną.

Dowód twierdzenia o pochodnej szeregu potęgowego

Wykazaliśmy wcześniej, że dla każdego ciągu liczb rzeczywistych (a_n) istnieje $r \geq 0$, takie że jeśli $|x| < r$, to szereg potęgowy $\sum\limits_{n=0}^{\infty} n^k a_n x^n$ jest zbieżny bezwzględnie dla każdej liczby $k = 0, 1, 2, 3, \ldots$ zaś w przypadku $|x| > 0$ szereg $\sum\limits_{n=0}^{\infty} a_n x^n$ jest rozbieżny. Obiecaliśmy wykazać, że funkcja s przypisująca liczbie x sumę szeregu $s(x) = \sum\limits_{n=0}^{\infty} a_n x^n$ jest różniczkowalna w każdym punkcie $x \in (-r, r)$ oraz że zachodzi równość $s'(x) = \left(\sum\limits_{n=0}^{\infty} a_n x^n\right)' = \sum\limits_{n=1}^{\infty} na_n x^{n-1}$ Zakładamy dalej, że $|x| < r$, że $d > 0$ jest liczbą mniejszą niż $r - |x|$ oraz że $0 < |h| < d$. Stąd wynika, że $|x+h| \leq |x| + |h| < r$, więc szeregi $\sum\limits_{n=0}^{\infty} na_n x^{n-1}$, $\sum\limits_{n=0}^{\infty} a_n x^n$ i $\sum\limits_{n=0}^{\infty} a_n(x+h)^n$ są zbieżne i to bezwzględnie. Mamy więc

$$\left|\frac{s(x+h)-s(x)}{h} - \sum_{n=0}^{\infty} na_n x^{n-1}\right| = \left|\sum_{n=0}^{\infty} a_n\left(\frac{(x+h)^n - x^n}{h} - nx^{n-1}\right)\right| =$$

$$= \left|\sum_{n=2}^{\infty} a_n \sum_{k=2}^{n} \binom{n}{k} x^{n-k} h^{k-1}\right| \leq |h| \cdot \sum_{n=2}^{\infty} |a_n| \sum_{k=2}^{n} \binom{n}{k} |x|^{n-k} |h|^{k-2} \leq$$

$$\leq |h| \cdot \sum_{n=2}^{\infty} |a_n| n^2 \sum_{k=2}^{n} \binom{n-2}{k-2} |x|^{(n-2)-(k-2)} |h|^{k-2} =$$

$$= |h| \cdot \sum_{n=2}^{\infty} n^2 |a_n| \left(|x| + |h|\right)^{n-2} \leq |h| \cdot \sum_{n=2}^{\infty} n^2 |a_n| \left(|x| + d\right)^{n-2} .$$

Przedostatnia nierówność wynika z tego, że jeśli $n \geq k \geq 2$, to $\binom{n}{k} = \frac{n\cdot(n-1)}{(k-1)\cdot k} \cdot \binom{n-2}{k-2} \leq n^2 \binom{n-2}{k-2}$.
Oczywiście $\lim\limits_{h\to 0} |h| \cdot \sum\limits_{n=2}^{\infty} n^2 |a_n| \left(|x|+d\right)^{n-2} = 0$, zatem również $\lim\limits_{h\to 0} \frac{s(x+h)-s(x)}{h} - \sum\limits_{n=1}^{\infty} na_n x^{n-1} = 0$, a stąd od razu wynika, że $s'(x) = \lim\limits_{h\to 0} \frac{s(x+h)-s(x)}{h} = \sum\limits_{n=1}^{\infty} na_n x^{n-1}$. Dowód został zakończony. ■

Pokażemy teraz, jak podane przed chwilą twierdzenia można stosować.

Przykłady

9. Znajdziemy pochodną funkcji kosinus. Mamy $\cos x = \sin\left(\frac{\pi}{2} - x\right)$. Skorzystamy z wzoru wynikającego z wzoru wykazanego w przykładzie pierwszym: $\left(\frac{\pi}{2} - x\right)' = \left((-1)x + \frac{\pi}{2}\right)' = -1$. Teraz skorzystamy z twierdzenia o pochodnej złożenia:

$$(\cos x)' = \left(\sin\left(\frac{\pi}{2} - x\right)\right)' = \cos\left(\frac{\pi}{2} - x\right) \cdot (-1) = -\sin x$$

– tutaj rolę funkcji f z wzoru na pochodną złożenia pełni sinus, którego pochodną jest kosinus, zaś rolę funkcji g odgrywa funkcja $\frac{\pi}{2} - x$, której pochodną jest -1. ■

10. Zastosujemy wzór na pochodną ilorazu dla uzyskania wzoru na pochodną funkcji tangens. Mamy $(\operatorname{tg} x)' = \left(\frac{\sin x}{\cos x}\right)' = \frac{(\sin x)'\cos x - (\cos x)'\sin x}{(\cos x)^2} = \frac{\cos x \cdot \cos x - (-\sin x)\sin x}{\cos^2 x} = \frac{1}{\cos^2 x} = \quad = 1 + \operatorname{tg}^2 x$.

11. Teraz kolej na kotangens. Wzór ten można uzyskać na różne sposoby, np. modyfikując nieznacznie wyprowadzenie wzoru na pochodną funkcji tangens. Można też zastosować metodę znaną już z wyprowadzenia wzoru na pochodną funkcji kosinus i właśnie tak postąpimy:
$(\operatorname{ctg} x)' = \left(\operatorname{tg}\left(\frac{\pi}{2} - x\right)\right)' = \frac{1}{\cos^2\left(\frac{\pi}{2} - x\right)} \cdot (-1) = -\frac{1}{\sin^2 x} = -1 - \operatorname{ctg}^2 x$.

12. Przypomnijmy, że funkcją odwrotną do funkcji tangens ograniczonej do przedziału $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ jest funkcja arctg, która przekształca zbiór wszystkich liczb rzeczywistych $\mathbb{R}$ na przedział $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Zachodzi zatem wzór: $\operatorname{tg}(\operatorname{arctg} x) = x$. Funkcja arctg jest ciągła. Pochodna funkcji tangens nie jest w żadnym punkcie mniejsza od 1, więc jest różna od 0. Wobec tego z twierdzenia o pochodnej funkcji odwrotnej wynika, że funkcja arctg ma pochodną w każdym punkcie. Z twierdzenia o pochodnej złożenia wynika, że musi zachodzić wzór:
$1 = (x)' = (\operatorname{tg}(\operatorname{arctg} x))' = \left(1 + \operatorname{tg}^2(\operatorname{arctg} x)\right) \cdot (\operatorname{arctg} x)' = \left(1 + x^2\right) \cdot (\operatorname{arctg} x)'$.
Stąd wnioskujemy, że $(\operatorname{arctg} x)' = \frac{1}{1+x^2}$. ■

13. Wyprowadzimy wzór na pochodną funkcji arcsin, czyli funkcji odwrotnej do funkcji sinus ograniczonej do przedziału $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Funkcja arcsinus jest ciągła i przekształca przedział $[-1, 1]$ na przedział $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Na tym ostatnim przedziale funkcja kosinus przyjmuje nieujemne wartości. Stąd wynika, że jeśli $-\frac{\pi}{2} \le y \le \frac{\pi}{2}$, to $\cos y = \sqrt{1 - \sin^2 y}$. Ponieważ pochodna funkcji sinus jest różna od 0 w punktach przedziału otwartego $(-\frac{\pi}{2}, \frac{\pi}{2})$, więc funkcja arcsin jest różniczkowalna w punktach odpowiadających punktom przedziału $(-\frac{\pi}{2}, \frac{\pi}{2})$, czyli w punktach przedziału otwartego $(-1, 1)$. Mamy więc

$$1 = (x)' = (\sin(\arcsin(x)))' = \cos(\arcsin(x)) \cdot (\arcsin(x))' = \sqrt{1 - \sin^2(\arcsin(x))} \cdot (\arcsin(\mathrm{x}))' = \\ = \sqrt{1 - x^2} \cdot (\arcsin(\mathrm{x}))'.$$

Stąd już łatwo wynika, że zachodzi wzór: $(\arcsin(\mathrm{x}))' = \frac{1}{\sqrt{1-x^2}}$. Wyprowadziliśmy więc wzór na pochodną funkcji arcsin w punktach wewnętrznych jej dziedziny. W punktach leżących na jej brzegu, czyli w punktach -1 i 1 można by mówić jedynie o pochodnych jednostronnych. Pozostawiamy czytelnikom wykazanie tego, że w obu końcach przedziału $[-1, 1]$ funkcja arcsin ma pochodną jednostronną i że ta pochodna jednostronna równa jest $+\infty$. Warto naszkicować sobie wykres funkcji arcsin – jest on oczywiście symetryczny do wykresu funkcji sinus, ograniczonej do przedziału $[-\frac{\pi}{2}, \frac{\pi}{2}]$, względem prostej o równaniu $y = x$. ■

14. Niech $f(x) = x^a$, gdzie a jest dowolną liczbą rzeczywistą, zaś x jest liczbą dodatnią. Wykażemy, że $(x^a)' = ax^{a-1}$.* Z definicji wynika, że $x^a = e^{a\ln x}$. Korzystając z twierdzenia o pochodnej złożenia dwu funkcji oraz poprzednio wyprowadzonych wzorów na pochodne funkcji wykładniczej, logarytmu i funkcji liniowej otrzymujemy: $(x^a)' = \left(e^{a\ln x}\right)' = e^{a\ln x} \cdot a \cdot \frac{1}{x} = ax^{a-1}$. Dodać wypada, że potęgę x^a można zdefiniować również w przypadku $x = 0$ i $a > 0$ oraz w przypadku $x < 0$, jeśli a jest liczbą wymierną, której mianownik jest całkowitą liczbą nieparzystą, a licznik – liczbą całkowitą, po ewentualnym skróceniu. Pozostawiamy czytelnikom uzasadnienie tego, że w obu tych przypadkach podany przez nas wzór na pochodną funkcji potęgowej pozostaje w mocy, oczywiście w przypadku pierwszym mowa jest jedynie o pochodnej prawostronnej, chyba że a jest wykładnikiem dodatnim, wymiernym o mianowniku nieparzystym (mowa o zapisie liczby wymiernej w postaci ułamka nieskracalnego, którego licznik i mianownik są liczbami całkowitymi). ■

15. Zajmiemy się teraz przez chwilę funkcją wykładniczą o dowolnej podstawie. Niech a będzie dowolną liczbą dodatnią, x – dowolną liczbą rzeczywistą. Postępując tak jak w przypadku funkcji potęgowej otrzymujemy wzór: $(a^x)' = \left(e^{x\ln a}\right)' = e^{x\ln a} \cdot \ln a = a^x \ln a$. ■

Na tym zakończymy krótki przegląd najbardziej podstawowych wzorów na pochodne. Pochodne będziemy obliczać wielokrotnie. Przekonamy się niebawem, że można ich używać w celu rozwiązywania rozlicznych problemów, np. znajdowania największych i najmniejszych wartości funkcji. Do tego potrzebne będą nam jednak twierdzenia pozwalające na wiązanie własności funkcji z własnościami jej pochodnej. Warto nadmienić, że z twierdzeń, które już podaliśmy, wynika, że funkcje zdefiniowane za pomocą „jednego wzoru", mają pochodną we wszystkich punktach swej dziedziny z wyjątkiem nielicznych punktów wyjątkowych, np. wzór $(\sqrt[3]{x})' = \left((x^{1/3})\right)' = \frac{1}{3}x^{-2/3} = \frac{1}{3\sqrt[3]{x^2}}$ ma miejsce dla wszystkich $x \neq 0$. Istnieją, co prawda, funkcje ciągłe określone na całej proste, które nie mają pochodnej w żadnym punkcie. Przykład pojawi się później **

Twierdzenie o ciągłości funkcji różniczkowalnej

Jeśli funkcja f jest różniczkowalna w punkcie p, to jest w tym punkcie ciągła.

Dowód.

$$\lim_{x\to p} f(x) = f(p) + \lim_{h\to 0}(f(p+h) - f(p)) = f(p) + \lim_{h\to 0} h \cdot \frac{f(p+h) - f(p)}{h} = f(p) + 0 \cdot f'(p) = f(p).$$

Dowód został zakończony. ■

2. Badanie funkcji za pomocą pochodnych: ekstrema i monotoniczność

Następne twierdzenie było używane przez Fermata (1601–1665) w odniesieniu do wielomianów jeszcze przed wprowadzeniem przez Newtona i Leibniza rachunku różniczkowego i całkowego. Fermat zajmował się znajdował między innymi znajdowaniem wartości największych i najmniejszych wielomianów na przedziałach domkniętych. Doprowadziło go to w gruncie rzeczy do pojęcia pochodnej, choć nie stworzył on teorii. Tym nie mniej odkrył twierdzenie, którego wagę trudno przecenić, choć zarówno twierdzenie jak i jego dowód są niesłychanie proste.

Twierdzenie o zerowaniu się pochodnej w punktach lokalnego ekstremum

Jeśli f jest funkcją różniczkowalną w punkcie p i przyjmuje w punkcie p wartość najmniejszą lub

* Dla $a=\frac{1}{2}$ jest to znany wielu czytelnikom z nauki w szkole wzór na pochodną pierwiastka kwadratowego, dla $a=2$ oraz $a=3$ otrzymaliśmy wzory wcześniej, zob, przykład 2,3.

** Ten stan rzeczy może ulec zmianie, w fizyce rozpatrywany jest tzw. ruch Browna, w którego modelu matematycznym tego rodzaju dziwactwa pojawiają się. Związany z ruchem Browna proces Wienera znajduje zastosowania również w modelach ekonomicznych.

największą, to $f'(p)=0$, podkreślić wypada, że zakładamy tu, że p jest środkiem pewnego przedziału otwartego zawartego w dziedzinie funkcji.

Dowód.

Załóżmy, że funkcja f ma w punkcie p wartość największą. Znaczy to, że dla każdego punktu x z dziedziny funkcji f zachodzi nierówność $f(x)\leq f(p)$, zatem dla $h>0$ mamy $\frac{f(p+h)-f(p)}{h}\leq 0$, wobec tego $f'(p)=\lim\limits_{h\to 0^+}\frac{f(p+h)-f(p)}{h}\leq 0$. Mamy też $f'(p)=\lim\limits_{h\to 0^-}\frac{f(p+h)-f(p)}{h}\geq 0$ dla $h<0$. Obie te nierówności mogą zachodzić jednocześnie jedynie w przypadku $f'(p)=0$. Jeśli f przyjmuje w punkcie p wartość najmniejszą, to funkcja przeciwna $-f$ przyjmuje w tym punkcie wartość największą, więc $0=(-f)'(p)=-f'(p)$. Dowód został zakończony. ■

Wypada podkreślić, że jeśli funkcja określona na przedziale przyjmuje wartość największą w jego końcu, to nawet w przypadku, gdy jest w tym końcu jednostronnie różniczkowalna, to jej pochodna nie musi być równa 0, funkcja x rozpatrywana na przedziale $[7,13]$ przyjmuje swą największą wartość w punkcie 13, w którym jej pochodną jest liczba 1.

Uwaga o wartościach funkcji w pobliżu punktu, w którym pochodna jest dodatnia

Jeśli funkcja f jest różniczkowalna w punkcie p oraz $f'(p)>0$, to istnieje liczba $\delta>0$ taka, że jeśli $0<h<\delta$, to $f(p-h)<f(p)<f(p+h)$, tzn. dostatecznie blisko punktu p, na lewo od niego wartości funkcji są mniejsze niż w wartość punkcie p, zaś na prawo od tego punktu, w jego pobliżu wartości funkcji są większe niż wartość w punkcie p.

Dowód.

Iloraz różnicowy $\frac{f(p+h)-f(p)}{h}$ jest dodatni dla dostatecznie małych h, bowiem ma dodatnią granicę przy $h\to 0$, zatem licznik i mianownik tego ułamka mają taki sam znak. ■

Twierdzenie Rolle'a

Jeżeli funkcja f jest ciągła w przedziale domkniętym $[a,b]$ i ma pochodną we wszystkich jego punktach wewnętrznych oraz $f(a)=f(b)$, to istnieje punkt $c\in(a,b)$, taki że $f'(c)=0$.

Dowód.

Załóżmy, że $f(a)=f(b)$ nie jest największą wartością funkcji f. Niech c będzie punktem, w którym funkcja f przyjmuje wartość największą spośród przyjmowanych na tym przedziale. Oczywiście $a<c<b$. Wobec tego f jest różniczkowalna w punkcie c i na mocy twierdzenia Fermata zachodzi równość $f'(c)=0$. Jeśli funkcja f nie przyjmuje wewnątrz przedziału $[a,b]$ wartości większych niż $f(a)=f(b)$, to albo przyjmuje mniejsze i możemy zamiast niej rozważyć funkcję przeciwną $-f$, albo funkcja f jest stała na przedziale $[a,b]$. W tym drugim przypadku c może być dowolnym punktem przedziału otwartego (a,b). Dowód został zakończony. ■

Interpretacja fizyczna tego twierdzenia może być np. taka: po prostoliniowej drodze porusza się pojazd, który rozpoczyna i kończy przemieszczanie się w tym samym punkcie $\big(f(a)=f(b)\big)$, ponieważ kończymy podróż w punkcie startu, więc w którymś punkcie musieliśmy zawrócić, w momencie zmiany kierunku jazdy nasza prędkość była równa 0.

Na wykresie funkcji punkty, o których jest mowa w dowodzie twierdzenia Rolle'a to te w otoczeniu, których wykres wygląda tak, jak wykres funkcji $-x^2$ w otoczeniu punktu 0. Oczywiście to nie są jedyne punkty, w których pochodna przyjmuje wartość 0. Niech $f(x)=\sin^3 x$. Wtedy $f'(x)=3\sin^2 x\cos x$, zatem $f'(0) = 0$, chociaż w punkcie 0 funkcja f nie ma lokalnego maksimum ani lokalnego minimum, w każdym przedziale postaci (δ, δ), gdzie $0 < \delta < \frac{\pi}{2}$, funkcja f jest ściśle rosnąca. Ma ona lokalne ekstrema, ale w innych punktach, np. w punktach $\pm\frac{\pi}{2}$.

Przejdziemy teraz do najważniejszego twierdzenia w rachunku różniczkowym, twierdzenia o wartości średniej.

Twierdzenie Lagrange'a o wartości średniej

Jeśli funkcja f jest ciągła w każdym punkcie przedziału domkniętego $[a, b]$ i ma pochodną we wszystkich punktach przedziału otwartego (a, b), to istnieje punkt $c \in [a, b]$, taki że $f'(c) = \frac{f(b) - f(a)}{b - a}$.

Dowód.

Niech $g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$ – od funkcji f odejmujemy funkcję $\frac{f(b) - f(a)}{b - a}(x - a)$, więc liniową, której przyrost na przedziale $[a, b]$ jest równy $f(b) - f(a)$, czyli jest równy przyrostowi funkcji f na tym przedziale. Mamy więc $g(a) = f(a) = g(b)$. Poza tym z definicji funkcji g wynika od razu, że jest to funkcja ciągła, jako różnica funkcji ciągłych. Taki sam argument przekonuje nas o istnieniu pochodnej g' we wszystkich punktach przedziału otwartego (a, b). Wobec tego dla funkcji g spełnione są założenia twierdzenia Rolle'a. Istnieje wobec tego taki punkt $c \in (a, b)$, że $0 = g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}$, a to właśnie mieliśmy wykazać. Dowód został zakończony. ■

Każdy czytelnik z pewnością zauważył, że twierdzenie Rolle'a jest przypadkiem szczególnym twierdzenia Lagrange'a o wartości średniej. Można też zinterpretować „fizycznie" twierdzenie Lagrange'a. Jeśli $f(x)$ oznacza położenie w chwili x obiektu poruszającego się po prostej, to $f'(c)$ oznacza prędkość w chwili c, natomiast $\frac{f(b) - f(a)}{b - a}$ to prędkość średnia w okresie od a do b. Wg. tej interpretacji twierdzenie o wartości średniej mówi, że prędkość chwilowa w pewnej chwili c równa jest prędkości średniej, co wygląda na stwierdzenie zupełnie oczywiste. Geometrycznie twierdzenie to oznacza, że jeśli poprowadzimy prostą przez dwa punkty leżące na wykresie funkcji f, to styczna do wykresu f w *pewnym* punkcie leżącym między wybranymi punktami jest równoległa do wybranej prostej.

Widzimy więc, że twierdzenie Lagrange'a ma krótki dowód, prosto można je zinterpretować na różne sposoby. Pokażemy teraz, że ma ono liczne i ważne konsekwencje. Najpierw jednak pokażemy jeszcze jedno jego uogólnienie.

Twierdzenie Cauchy'ego o wartości średniej

Jeśli funkcje f i g są ciągłe w każdym punkcie przedziału domkniętego $[a, b]$ i mają pochodne we wszystkich punktach przedziału otwartego (a, b) przy czym $g'(x) \neq 0$, to istnieje punkt $c \in [a, b]$, taki że $\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}$.

Dowód. Nie różni się niczym istotnym od dowodu twierdzenia Lagrange'a o wartości średniej. Rozpatrujemy pomocniczą funkcję h zdefiniowaną wzorem $h(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}\big(g(x) - g(a)\big)$ i stosujemy do niej twierdzenie Rolle'a. Zauważmy jeszcze, że z twierdzenia Rolle'a wynika, że funkcja g jest różnowartościowa na przedziale $[a, b]$, więc wszystkie rozpatrywane ilorazy w tym twierdzeniu i jego dowodzie maja sens. ■

Twierdzenie o monotoniczności funkcji różniczkowalnych

Załóżmy, że funkcja f jest ciągła w każdym punkcie przedziału P i że jest różniczkowalna we wszystkich jego punktach wewnętrznych. Przy tych założeniach funkcja f jest:

– niemalejąca ($x < y \Rightarrow f(x) \le f(y)$) wtedy i tylko wtedy, gdy jej pochodna f' jest nieujemna,

– nierosnąca ($x < y \Rightarrow f(x) \ge f(y)$) wtedy i tylko wtedy, gdy jej pochodna f' jest niedodatnia.

Dowód.

Jeśli funkcja jest niemalejąca, to iloraz różnicowy $\dfrac{f(x+h)-f(x)}{h}$ jest nieujemny, bo licznik i mianownik tego ułamka mają taki sam znak. Granica funkcji nieujemnej, jeśli istnieje, to jest nieujemna. Z tego zdania wynika natychmiast, że pochodna we wszystkich tych punktach przedziału P, w których istnieje, jest nieujemna. Załóżmy teraz, że pochodna w punktach wewnętrznych przedziału P jest nieujemna. Załóżmy, że $x, y \in P$ i że $x < y$. Z twierdzenia o wartości średniej zastosowanego do przedziału $[x, y]$ wynika, że $\dfrac{f(y)-f(x)}{y-x} = f'(z) \ge 0$ dla pewnego punktu $z \in (x, y)$. Ponieważ mianownik ułamka $\dfrac{f(y)-f(x)}{y-x}$ jest dodatni, a sam ułamek jest nieujemny, więc licznik tego ułamka, czyli różnica $f(y) - f(x)$, też jest nieujemny, zatem $f(y) \ge f(x)$, co dowodzi tego, że funkcja f jest niemalejąca. Drugi przypadek sprowadzamy jak zwykle do pierwszego zastępując funkcję f funkcją przeciwną $-f$. Dowód został zakończony. ■

Wniosek*

Funkcja ciągła na przedziale P, różniczkowalna we wszystkich jego punktach wewnętrznych jest stała wtedy i tylko wtedy, gdy $f'(x) = 0$ dla każdego punktu wewnętrznego przedziału P.

Dowód.

Funkcja stała jest jednocześnie niemalejąca i nierosnąca, zatem jej pochodna jest jednocześnie nieujemna i niedodatnia, czyli zerowa. Jeśli natomiast pochodna jest zerowa, czyli jednocześnie nieujemna i niedodatnia, to funkcja jest zarówno niemalejąca, jak i nierosnąca, więc jest stała. Dowód został zakończony. ■

Twierdzenie o ścisłej monotoniczności funkcji różniczkowalnych

Zakładamy jak poprzednio, że funkcja f jest ciągła w każdym punkcie przedziału P oraz ze jest różniczkowalna w każdym punkcie wewnętrznym przedziału P. Przy tych założeniach funkcja f jest:

– ściśle rosnąca wtedy i tylko wtedy, gdy jej pochodna jest nieujemna oraz między każdymi dwoma punktami przedziału P znajduje się punkt, w którym pochodna f' jest dodatnia,

– ściśle malejąca wtedy i tylko wtedy, gdy jej pochodna jest niedodatnia oraz między każdymi dwoma punktami przedziału P znajduje się punkt, w którym pochodna f' jest ujemna.

Dowód.

Załóżmy, że funkcja f jest ściśle rosnąca. Jest więc również niemalejąca, więc na podstawie poprzedniego twierdzenia jej pochodna jest nieujemna. Jeśli $x, y \in P$, $x < y$, to w pewnym punkcie wewnętrznym z przedziału $[x, y]$ zachodzi nierówność $f'(z) > 0$, bowiem gdyby pochodna równa była 0 w każdym punkcie przedziału $[x, y]$, to funkcja f byłaby stała na tym przedziale, więc nie byłaby *ściśle* rosnąca. Zajmiemy się dowodem implikacji przeciwnej. Zakładamy teraz, że f jest funkcją ciągłą, której pochodna jest nieujemna. Z poprzedniego twierdzenia wnioskujemy, że f jest funkcją niemalejącą. Jeśli nie jest ona ściśle rosnąca, to istnieją punkty $x, y \in P$, takie że $x < y$ i $f(x) = f(y)$. Jeśli $x < z < y$, to $f(x) \le f(z) \le f(y) = f(x)$, co oznacza, że $f(x) = f(z)$, a to z kolei oznacza, że f jest funkcją stałą na przedziale $[x, y]$, a z tego wynika, że $f'(z) = 0$ dla każdego punktu $z \in [x, y]$, wbrew założeniu.

* Można z łatwością ten wniosek udowodnić bezpośrednio, bez powoływania się na właśnie wykazane twierdzenie.

Druga część twierdzenia może być uzyskana z pierwszej przez rozważenie funkcji $-f$ zamiast funkcji f. Dowód został zakończony. ■

Twierdzenie o lipschitzowskości funkcji różniczkowalnej

Zakładamy jak w twierdzeniach poprzednich, że funkcja f jest określona na pewnym przedziale P, że jest na nim ciągła i że jest różniczkowalna we wszystkich punktach wewnętrznych tego przedziału. Przy tych założeniach funkcja f spełnia warunek Lipschitza ze stałą $L \geq 0$ wtedy i tylko wtedy, gdy $L \geq \sup\{|f'(t)|: \quad t \in \mathrm{int}P\}$.*

Dowód.

Jeśli $x, y \in P$, to na mocy twierdzenia Lagrange'a o wartości średniej istnieje punkt z leżący między x i y, taki że $|f(x) - f(y)| = |f'(z)(x-y)| \leq \sup\{|f'(t)|: \quad t \in \mathrm{int}P\} \cdot |x-y|$, co kończy dowód twierdzenia w jedną stronę. Dowód w drugą stronę wynika natychmiast z tego, że jeśli funkcja spełnia warunek Lipschitza ze stałą L, to dla dowolnych $x, y \in P$ zachodzi nierówność $\left|\dfrac{f(x) - f(y)}{x - y}\right| \leq L$, zatem $|f'(x)| = \lim\limits_{y \to x} \left|\dfrac{f(y) - f(x)}{y - x}\right| \leq L$, zatem $\sup\{|f'(t)|: \quad t \in \mathrm{int}P\} \leq L$. Dowód został zakończony. ■

Przykłady

16. Niech $f(x) = \dfrac{1}{x}$. Mamy $f'(x) = -\dfrac{1}{x^2} < 0$. Funkcja ma więc ujemną pochodna w każdym punkcie swej dziedziny $(-\infty, 0) \cup (0, \infty)$. Mamy też $f(-1) = -1 < 1 = f(1)$, wobec tego funkcja ta nie jest nierosnąca, tym bardziej nie jest malejąca. Przyczyną tego zjawiska jest to, że dziedzina tej funkcji *nie* jest przedziałem – malutka, raptem jednopunktowa dziura w dziedzinie, powoduje, że teza przestaje być prawdziwa! . Na każdym *przedziale*, na którym jest zdefiniowana, funkcja ta jest nierosnąca, a nawet ściśle malejąca. ■

17. Niech $f(x) = \sin x - \left(x - \dfrac{x^3}{6}\right)$. Mamy $f'(x) = \cos x - \left(1 - \dfrac{x^2}{2}\right)$ i wobec tego również $(f')'(x) = -\sin x + x$. Udowodniliśmy poprzednio, że jeśli $x > 0$, to $\sin x < x$. Z tej nierówności wynika, że $(f')'(x) > 0$ dla każdego $x > 0$. Wobec tego funkcja f' jest ściśle rosnąca na półprostej domkniętej $[0, \infty)$. Stąd wynika, że $f'(x) > f'(0) = \cos 0 - \left(1 - \dfrac{0^2}{2}\right) = 0$ dla każdego $x > 0$. Wobec tego funkcja f, której pochodna jest dodatnia na półprostej otwartej $(0, \infty)$, jest ściśle rosnąca na półprostej domkniętej $[0, \infty)$, zatem dla $x > 0$ zachodzi nierówność $f(x) > f(0) = 0 - \left(0 - \dfrac{0^3}{6}\right) = 0$. Wykazaliśmy w ten sposób, że $\sin x > x - \dfrac{x^3}{6}$ dla każdej liczby dodatniej x. ■

18. Zajmując się funkcja wykładniczą o podstawie e w rozdziale pierwszym wykazaliśmy, że dla każdej liczby rzeczywistej x zachodzi nierówność $e^x \geq 1 + x$, później zresztą wzmocniona. Wiemy, że pochodną funkcji e^x jest ta sama funkcja. Wartością tej pochodnej w punkcie 0 jest liczba $e^0 = 1$. Wobec tego równanie stycznej do wykresu funkcji wykładniczej w punkcie $(0, 1)$ ma postać $y = 1 \cdot (x - 0) + e^0 = x + 1$. Wobec tego wspomniana nierówność oznacza, że wykres funkcji wykładniczej o podstawie e znajduje się nad styczną do siebie w punkcie $(0, 0)$. Przekonamy się później, że jest to związane z wypukłością funkcji wykładniczej. ■

19. Wspomnieliśmy w przykładzie siedemnastym nierówność $\sin x < x$, która zachodzi dla $x > 0$. Pochodną funkcji sinus jest funkcja kosinus. W punkcie 0 wartość pochodnej to $\cos 0 = 1$. Wynika

* $\mathrm{int}\,P$ oznacza zbiór złożony ze wszystkich punktów wewnętrznych przedziału P, czyli przedział otwarty, którego końce pokrywają się z końcami przedziału P.

stąd, że równanie stycznej do wykresu funkcji sinus w punkcie $(0,0)$ przybiera postać

$$y = 1\cdot(x-0)+\sin 0 = x\,.$$

Wobec tego nierówność $x>\sin x$ oznacza, że na półprostej $(0,\infty)$ wykres funkcji sinus znajduje się pod styczną do tegoż wykresu w punkcie $(0,0)$. Przekonamy się później, że jest to związane z wklęsłością funkcji sinus na przedziale $[0,\pi]$, dla $x\ge\pi$ nierówność zachodzi, bo wartości funkcji sinus są mniejsze niż $1<\pi$. ■

20. Niech $f(x)=e^x-(1+x+\frac12x^2)$. Mamy $f'(x)=e^x-(1+x)\ge 0$. Stąd $(f')'(x)=e^x-1>0$, dla $x>0$ oraz $(f')'(x)=e^x-1<0$ dla $x<0$. Wynika stąd, że funkcja f' jest ściśle rosnąca na półprostej $[0,\infty)$ oraz ściśle malejąca na półprostej $(-\infty,0]$. Wobec tego najmniejszą wartością funkcji f' jest $f'(0)=e^0-1=0$. Oznacza to, że dla $x\ne 0$ zachodzi nierówność $f'(x)>0$, czyli $e^x>1+x$. Wobec tego, że funkcja f' przyjmuje wartości dodatnie na całej prostej z wyjątkiem jednego punktu, funkcja f jest ściśle rosnąca na całej prostej.
Mamy więc $f(x)>f(0)=e^0-(1+0+\frac12 0^2)=0$ dla $x>0$ oraz $f(x)<f(0)=0$ dla $x<0$, zatem dla $x>0$ zachodzi nierówność $e^x>1+x+\frac12x^2$, zaś dla $x<0$ – nierówność $e^x<1+x+\frac12x^2$. Rozumując w ten sam sposób można wykazać, że dla każdej liczby rzeczywistej x zachodzi nierówność $e^x\ge 1+x+\frac{1}{2!}x^2+\frac{1}{3!}x^3$, przy czym nierówność jest ostra dla $x\ne 0$. Uogólnienie pozostawiamy czytelnikom w charakterze prostego ćwiczenia. Zachęcamy też do porównania z rozumowaniami przeprowadzanymi w rozdziale pierwszym: bez trudu można zauważyć, że uzyskujemy teraz z łatwością nierówności, których wykazanie bez użycia pochodnych było dosyć trudne. ■

21. *Ten przykład będzie nieco dłuższy. Należy go przestudiować z uwagą. Stosowana tu metoda będzie używana później również w odniesieniu do funkcji wielu zmiennych, pokazuje ona też, że twierdzenia o istnieniu mogą się przydawać również w rozwiązywaniu problemów konkretnych.*
Niech $a\ge b>0$ będą liczbami rzeczywistymi. P niech oznacza prostokąt, którego jeden bok ma długość a, a drugi – b. Z prostokąta P wycinamy cztery kwadraty o boku $x\in\left(0,\frac b2\right)$ zawierające cztery wierzchołki P , tak że pole P zmniejsza się o $4x^2$. Następnie zaginamy „wystające" części powstałego dwunastokąta (niewypukłego) tak, by powstało pudełko o wymiarach $a-2x,\ b-2x,\ x$. Dla jakiego x pojemność otrzymanego pudełka będzie największa?

Rozwiązanie. Niech $V(x)=x(a-2x)(b-2x)$ będzie pojemnością pudełka. V jest funkcją ciągła, a nawet różniczkowalną w każdym punkcie swej dziedziny. Z punktu widzenia pojemności pudełka dziedziną funkcji V jest przedział $\left(0,\frac b2\right)$, ale można tę funkcję rozpatrywać na przedziale domkniętym $\left[0,\frac b2\right]$. Na przedziale $\left[0,\frac b2\right]$ funkcja V , jako ciągła , przyjmuje wartość najmniejszą oraz wartość największą. Ponieważ $V(0)=V\left(\frac b2\right)=0$ i $V(x)>0$ dla $x\in\left(0,\frac b2\right)$, więc najmniejsza wartość przyjmowana jest w końcach przedziału $\left[0,\frac b2\right]$, zaś największa – w pewnym punkcie wewnętrznym x_0 tego przedziału. Ponieważ funkcja V jest różniczkowalna w x_0, więc $V'(x_0)=0$. Wystarczy zatem znaleźć punkty w przedziale $\left(0,\frac b2\right)$, w których pochodna funkcji V przyjmuje wartość 0 i stwierdzić, w którym z nich V ma największą wartość – takie punkty są co najwyżej dwa, bo V jest wielomianem trzeciego stopnia, więc V' jest wielomianem kwadratowym. $V'(x)=12x^2-4(a+b)x+ab$. Wiemy, że ten wielomian ma co najmniej jeden pierwiastek w przedziale $\left(0,\frac b2\right)$ (nie ma potrzeby sprawdzać, że jego wyróżnik jest dodatni, bo to wynika z istnienia x_0 !).* Możemy teraz zastosować to samo rozumowanie do badania funkcji V na przedziale $\left[\frac{b}{2},\frac{a}{2}\right]$. Wewnątrz tego przedziału funkcja V przyjmuje wartości ujemne, na końcach – zero. Wobec tego swą najmniejszą wartość na $\left[\frac{b}{2},\frac{a}{2}\right]$ funkcja V przyjmuje wewnątrz przedziału i wobec tego jej pochodna V' przyjmuje wartość 0 w co najmniej jednym punkcie tego przedziału. Wynika z tego rozumowania, że w każdym z przedziałów $\left(0,\frac{b}{2}\right)$, $\left(\frac{b}{2},\frac{a}{2}\right)$ pochodna V' funkcji V ma co najmniej jeden pierwiastek, a ponieważ V' ma dokładnie dwa pierwiastki, więc w każdym z wymienionych przedziałów ma dokładnie jeden pierwiastek. Tak się dzieje w przypadku $a > b$. W przypadku $a = b$ sytuacja jest nieco inna: $V'\left(\frac{b}{2}\right) = 0$, co sprawdzamy bezpośrednim rachunkiem (ogólnie: jeśli liczba x_1 jest podwójnym pierwiastkiem funkcji f , tzn. $f(x) = (x - x_1)^2 g(x)$ dla pewnej funkcji g różniczkowalnej w x_1, to $f(x_1) = 0 = f'(x_1)$) i wobec tego również w tym przypadku w przedziale $\left(0,\frac{b}{2}\right)$ funkcja V' może mieć co najwyżej jeden pierwiastek, więc ma dokładnie jeden. Udowodniliśmy w ten sposób, że w przedziale $\left(0,\frac{b}{2}\right)$ funkcja V' ma dokładnie jeden pierwiastek x_0, którym jest mniejszy z dwóch pierwiastków tej funkcji, a liczba $V(x_0)$ jest największą wartością funkcji V przyjmowaną na przedziale $\left(0,\frac{b}{2}\right)$. Oczywiście zachodzi równość

$$x_0 = \frac{4(a+b)-\sqrt{\left(4(a+b)\right)^2-4\cdot 12ab}}{2\cdot 12} = \frac{a+b-\sqrt{a^2+b^2-ab}}{6}.$$

Uwaga: nie zajmowaliśmy się znakiem pochodnej V', bo nie było potrzeby ustalać na jakich przedziałach funkcja V rośnie, a na jakich maleje. Oczywiście można było postąpić inaczej: stwierdzić, że na przedziale $(0, x_0)$ pochodna V' funkcji V jest dodatnia, więc V na tym przedziale rośnie, a na przedziale $\left(x_0,\frac{b}{2}\right)$ pochodna V' jest ujemna, więc na tym przedziale funkcja V maleje. Z naszego rozumowania to też wynika, bo na przedziale $(0, x_0)$ pochodna V' nie przyjmuje wartości 0 , ma zatem ten sam znak we wszystkich punktach tego przedziału, zatem funkcja V jest na tym przedziale ściśle monotoniczna, nie może być malejąca, bo $V(x_0) > 0 = V(0)$, więc jest ściśle rosnąca, więc jej niezerująca się pochodna jest dodatnia. ■

22. Znaleźć maksimum objętości brył powstałych w wyniku obrotu trójkąta prostokątnego o obwodzie 1 wokół jego przeciwprostokątnej.

Rozwiązanie: Niech a, b, c oznaczają boki trójkąta, przy czym c oznacza przeciwprostokątną. Bryła, która powstaje w wyniku obrotu trójkąta wokół boku c to dwa stożki złączone podstawami. Promień tej wspólnej podstawy to wysokość trójkąta prostopadła do przeciwprostokątnej, więc równa $\dfrac{ab}{c}$ (pole trójkąta jest równe $\frac{1}{2}ab = \frac{1}{2}ch_c$, gdzie h_c jest wysokością trójkąta prostopadłą do przeciwprostokątnej c). Suma wysokości tych stożków jest równa c. Wobec tego suma ich objętości jest równa $V = \dfrac{\pi}{3}\cdot\left(\dfrac{ab}{c}\right)^2\cdot c = \dfrac{\pi(ab)^2}{3c}$. Wiadomo, że $a^2 + b^2 = c^2$ i $a + b + c = 1$. Stąd wynika, że $2ab = (a + b)^2 - (a^2 + b^2) = (1 - c)^2 - c^2 = 1 - 2c$. Wobec tego zachodzi wzór $V = V(c) = \dfrac{\pi(1-2c)^2}{12c} = \dfrac{\pi}{12}\left(\dfrac{1}{c} - 4 + 4c\right)$. Stąd $V'(c) = \dfrac{\pi}{12}\left(\dfrac{-1}{c^2} + 4\right)$. Stąd wnioskujemy z łatwością, że $V'(c) = 0$ wtedy i tylko wtedy, gdy $c = \pm\dfrac{1}{2}$, zatem kandydatami na punkt, w którym funkcja V przyjmuje swą największą wartość są $\dfrac{1}{2}$ oraz $-\dfrac{1}{2}$. Liczba c jest długością boku trójkąta, zatem jest dodatnia, bo jest długością odcinka, więc nie może być równa $-\dfrac{1}{2}$. Liczba $\dfrac{1}{2}$ też nie

* Drugi pierwiastek wielomianu V' też jest dodatni, bo iloczyn pierwiastków tego wielomianu jest równy $\frac{ab}{12}$, jest więc dodatni, zatem oba pierwiastki maja ten sam znak, ale z tego korzystać nie będziemy.

wchodzi w grę, bo wtedy musiałoby być $a + b = 1 - \frac{1}{2} = \frac{1}{2} = c$, co przeczyłoby nierówności trójkąta. Oznacza to, że na każdym przedziale zawartym w dziedzinie funkcji V jest ona ściśle monotoniczna, zatem kresy, jeśli w ogóle są przyjmowane, to w końcach przedziału. Musimy więc znaleźć dziedzinę funkcji V. Oczywistym warunkiem koniecznym na to, by liczby a, b, c były bokami trójkąta prostokątnego o obwodzie 1, jest, aby były dodatnimi rozwiązaniami układu równań: $a^2 + b^2 = c^2$, $a + b = 1 - c$. Warunek ten jest też dostateczny: jeśli $a, b > 0$ i $a^2 + b^2 = c^2$, to $(a+b)^2 > a^2+b^2 = c^2$, zatem $a+b > c$ i oczywiście $a+c > c > b$ oraz $b+c > c > a$. Oznacza to, że z odcinków o długościach a, b, c można zbudować trójkąt, oczywiście prostokątny. Ten układ równań równoważny jest następującemu: $a + b = 1 - c$, $ab = \frac{(1-c)^2 - c^2}{2} = \frac{1}{2} - c$. Wobec tego liczby a i b są pierwiastkami równania kwadratowego $t^2 - (1-c)t + \frac{1}{2} - c = 0$. Warunkiem koniecznym i dostatecznym na to, by to równanie to miało dodatnie pierwiastki dla dodatniej wartości parametru c, jest $0 < c < \frac{1}{2}$ i $0 \le \Delta = (1-c)^2 - 4(\frac{1}{2} - c) = -1 + 2c + c^2 = (c+1)^2 - 2$ czyli $\sqrt{2} - 1 \le c < \frac{1}{2}$. Ponieważ $V\left(\frac{1}{2}\right) = 0$, więc maksymalna wartość V jest równa $V\left(\sqrt{2}-1\right)$ – oczywiście maksymalna na przedziale $\left[\sqrt{2}-1, \frac{1}{2}\right)$. Łatwo zauważyć, że dla $c = \sqrt{2} - 1$ otrzymujemy trójkąt równoramienny (bo $\Delta = 0$, więc pierwiastki równania kwadratowego $x^2 - (1-c)x + \frac{1}{2} - c = 0$, czyli liczby a i b są równe).

Komentarz: Ten przykład powinien przekonać studentów o konieczności zwracania uwagi na dziedzinę funkcji. Omawiałem to zadanie wielokrotnie na ćwiczeniach, jeszcze się nie zdarzyło, by studenci chcieli, aby objętość V potraktować jako np. funkcję zmiennej a. Gdyby tak się stało, byłoby $V = V(a) = \frac{\pi a^2(1-2a)^2}{6(1-a)(1-2a+2a^2)}$ i maksimum osiągane byłoby w punkcie wewnętrznym dziedziny funkcji V, czyli przedziału $\left(0, \frac{1}{2}\right)$, mianowicie w punkcie $\frac{2-\sqrt{2}}{2}$, zatem w punkcie zerowania się pochodnej funkcji V. Byłoby znacznie mniej kłopotu z dziedziną funkcji, za to więcej z obliczeniami. Często też studenci nie potrafili stwierdzić, że ponieważ funkcja ma niezerową pochodną na przedziale, to jest na nim monotoniczna. Wydawało im się, że popełnili błąd w obliczeniach, bo skoro w jakimś punkcie ma być maksimum, to pochodna tam musi się zerować – zapominali więc o tym, że to twierdzenie mówi o punktach *wewnętrznych* dziedziny, końców nie dotyczy. ■

23. Znajdziemy teraz kres górny iloczynu trzech liczb nieujemnych, których suma jest równa 3. Oznaczmy te liczby przez x, y, z. Mamy więc $x \ge 0$, $y \ge 0$, $z \ge 0$ oraz $x + y + z = 3$. Mamy znaleźć kres górny wyrażenia $xy(3 - x - y)$, przy założeniu, że $x, y \ge 0$ oraz $x + y \le 3$. Niech $s = x + y \le 3$. Chwilowo traktować będziemy wielkość s jako stałą. Przy ustalonym s nasze wyrażenie to $x(s-x)(3-s)$. Mamy znaleźć jego kres górny zakładając, że $0 \le x$, $0 \le y = s - x$, czyli $0 \le x \le s$. Mamy więc do czynienia z funkcją kwadratową zmiennej x: $(3-s)(-x^2 + sx)$. Większość studentów pamięta z nauki szkolnej, że funkcja kwadratowa, której współczynnik przy x^2 jest ujemny, przyjmuje swą wartość największą w środku odcinka, w którego funkcja ta przyjmuje równe wartości (np. 0, wtedy końcami odcinka są pierwiastki funkcji). W naszym przypadku tym punktem jest $x = \frac{1}{2}(0+s) = \frac{s}{2}$.* By zakończyć zadanie należy znaleźć maksymalną wartość wyrażenia $(3-s)\frac{s^2}{4}$ na przedziale $[0,3]$. Mamy $\left((3-s)\frac{s^2}{4}\right)' = -\frac{s^2}{4} + (3-s)\frac{s}{2} = 3 - \frac{3}{4}s^2$. Ponieważ funkcja $(3-s)\frac{s^2}{4}$ zmiennej s jest ciągła na przedziale domkniętym $[0,3]$, więc osiąga w jakimś punkcie swój kres górny. Ponieważ w końcach przedziału przyjmuje wartość 0, a wewnątrz jest dodatnia, więc kres górny jest przyjmowany w jakimś punkcie wewnętrznym tego przedziału. Jedynym punktem w przedziale $(0,3)$, w którym pochodna funkcji $(3-s)\frac{s^2}{4}$ przyjmuje wartość 0, jest 2. Wartość funkcji $(3-s)\frac{s^2}{4}$ w tym punkcie równa jest 1. Odpowiednie wartości wyjściowych zmiennych to $x = y = z = 1$. Zadanie zostało rozwiązane.

Pokażemy teraz inne rozwiązanie tego samego problemu. Przypomnijmy, że w poprzednim rozdziale wykazaliśmy nierówność o średniej arytmetycznej i geometrycznej, która w przypadku trzech liczb nieujemnych x, y, z przybiera postać $\sqrt[3]{xyz} \le \frac{x+y+z}{3}$, przy czym staje się ona równością wtedy i tylko wtedy, gdy $x = y = z$. W naszym przypadku oznacza to, że $\sqrt[3]{xyz} \le 1$, przy czym nierówność staje się równością wtedy i tylko wtedy, gdy $x = y = z = 1$. Wobec tego największa wartością iloczynu trzech liczb nieujemnych, których suma jest równa 3 jest liczba 1. To drugie rozwiązanie jest krótsze, ale wymaga pewnego pomysłu. Później pokażemy jeszcze dwa inne rozwiązania tego zadania w oparciu o twierdzenia dotyczące funkcji dwu lub trzech zmiennych rzeczywistych. ■

Zanim pokażemy następne przykłady zauważmy, że z definicji pochodnej wynika następująca równość przybliżona $f'(p) \approx \frac{f(p+h)-f(p)}{h}$ dla $h \approx 0$. Nie troszcząc się przesadnie o precyzję rozumowania przepisać ją można w postaci $f(p+h) \approx f(p) + f'(p)h$. Można się spodziewać, że jest to przybliżenie dokładniejsze dla h dostatecznie bliskich 0 niż przybliżenie $f(p+h) \approx f(p)$, które jest konsekwencją ciągłości funkcji f w punkcie p. Tak jest w rzeczywistości, bowiem błąd przybliżenia $f(p+h) \approx f(p) + f'(p)h$ jest mały w porównaniu z $|h|$, bowiem

$$\lim_{h\to 0} \tfrac{f(p+h)-\left(f(p)+f'(p)h\right)}{h} = \lim_{h\to 0}\left(\tfrac{f(p+h)-f(p)}{h} - f'(p)\right) = 0.$$

Zauważmy jeszcze, że zachodzi następujące :

Twierdzenie (charakteryzujące pochodną jako współczynnik wielomianu stopnia ≤ 1 najlepiej przybliżającego funkcję.)

Załóżmy, że f jest funkcją ciągłą w punkcie p. Wtedy równość $\lim\limits_{h\to 0} \frac{f(p+h)-(ah+b)}{h} = 0$ zachodzi wtedy i tylko wtedy, gdy funkcja f jest różniczkowalna w punkcie p oraz $a = f'(p)$ i $b = f(p)$.

Dowód.

Jeżeli $\lim\limits_{h\to 0} \frac{f(p+h)-(ah+b)}{h} = 0$, to $\lim\limits_{h\to 0} \frac{f(p+h)-b}{h} - a = 0$, więc $a = \lim\limits_{h\to 0} \frac{f(p+h)-b}{h}$, zatem $0 = \lim\limits_{h\to 0} ah = \lim\limits_{h\to 0}(f(p+h)-b)$, czyli $b = \lim\limits_{h\to 0} f(p+h) = f(p)$. Z ostatniej równości wynika, że $a = \lim\limits_{h\to 0} \frac{f(p+h)-b}{h} = \lim\limits_{h\to 0} \frac{f(p+h)-f(p)}{h}$, a to oznacza, że f jest różniczkowalna w punkcie p i zachodzi równość $a = f'(p)$, co kończy dowód twierdzenia w jedną stronę. Przed sformułowaniem

* Tym, którzy akurat zapomnieli, że tak jest, podajemy uzasadnienie w oparciu o twierdzenia z tego rozdziału. Mamy $(x(s-x)(3-s))' = (3-s)(-2x+s)$. Ta pochodna jest dodatnia na półprostej $(-\infty, \frac{s}{2})$, a na półprostej $(0,\infty)$ jest ujemna. Wobec tego funkcja jest ściśle rosnąca na półprostej $(-\infty, 0]$, a na półprostej $[0,\infty)$ jest ściśle malejąca, więc liczba $(3-s)\cdot\frac{s}{2}\cdot(s-\frac{s}{2}) = (3-s)\frac{s^2}{4}$ jest jej największą wartością.

twierdzenia wykazaliśmy prawdziwość implikacji przeciwnej. Dowód został zakończony. ■

Z twierdzenia tego wynika, że spośród wszystkich wielomianów stopnia ≤ 1 zmiennej x najlepiej przybliża funkcję f w otoczeniu punktu p wielomian $f(p)+f'(p)(x-p)$. Żadne z twierdzeń do tej pory sformułowanych nie daje jawnego oszacowania błędu przybliżenia, ale pokazywaliśmy już jak można dowodzić nierówności, a to stwarza szanse na szacowanie błędu. Pokażemy, teraz kilka przykładów.

Przykłady cd.

24. $\sqrt{50} = \sqrt{49+1} \approx \sqrt{49} + \dfrac{1}{2\sqrt{49}} \cdot 1 = 7 + \dfrac{1}{14}$ – przyjęliśmy tu $h = 1$, $f(x) = \sqrt{x}$, zatem $f'(x) = \dfrac{1}{2\sqrt{x}}$, $p = 49$. Chociaż 1 nie jest małą liczbą, jednak przybliżenie, które uzyskaliśmy jest dosyć dobre. Rzeczywiście, $\left(7+\dfrac{1}{14}\right)^2 = 49 + 2\cdot 7\cdot \dfrac{1}{14} + \left(\dfrac{1}{14}\right)^2 = 50 + \dfrac{1}{196}$. Widzimy więc, że po podniesieniu do kwadratu przybliżonej wartości pierwiastka, otrzymaliśmy liczbę nieco tylko większą od 50. Mamy $7,07 < 7+\dfrac{1}{14} < 7,08$ oraz $7,07^2 = 49,9849$, co oznacza, że nasze przybliżenie pozwoliło nam znaleźć dwie cyfry po przecinku liczby $\sqrt{50}$ bez wykonania trudnych obliczeń! Wartość przybliżona jest w tym przypadku większa niż rzeczywista, bo styczna do wykresu pierwiastka kwadratowego leży nad wykresem. ■

25. $50^2 = (49+1)^2 \approx 49^2 + 2\cdot 49\cdot 1 = 2499$. Tym razem $f(x) = x^2$, zatem $f'(x) = 2x$, $p = 49$ i $h = 1$. W rzeczywistości $50^2 = 2500$, więc tym razem błąd, który popełniamy stosując wzór przybliżony zamiast dokładnego jest równy 1, więc jest ponad 100 razy większy niż w poprzednim przykładzie. ■

26. $e^{50} = e^{49+1} \approx e^{49} + e^{49}\cdot 1 = 2\cdot e^{49}$. W tym przykładzie $f(x) = e^x = f'(x)$, $p = 49$ i $h = 1$. Zatem błąd„ który popełniamy w tym przypadku, jest równy $e^{50} - 2\cdot e^{49} = (e-2)\cdot e^{49} > 0,7\cdot e^{49}$, jest więc ogromny i to nie tylko w porównaniu z $h = 1$, ale wręcz porównywalny z wartością funkcji. Liczba e^{50} jest równa w przybliżeniu $5,184705485\cdot 10^{21}$, $e^{49} \approx 1,907346557\cdot 10^{21}$, zaś $e^{50} - 2\cdot e^{49} \approx 1,370012371\cdot 10^{21}$ – to rezultaty uzyskane za pomocą odpowiedniego programu komputerowego (Maple V). Widzimy więc, że w tym ostatnim przypadku przybliżanie za pomocą wzoru $f(p+h) \approx f(p)+f'(p)h$ w ogóle nie ma sensu, w przypadku funkcji x^2 dawało przybliżenie gorsze niż w przypadku pierwiastka kwadratowego. Można dosyć prosto wyjaśnić, co jest tego przyczyną. Otóż z twierdzenia Lagrange'a o wartości średniej wynika, że dla każdego $h \neq 0$ istnieje co najmniej jedna liczba $\theta_h \in (0,1)$, taka że $f(p+h) - f(p) = f'(p+\theta_h\cdot h)h$, zatem $f(p+h) - (f(p)+f'(p)h) = \big(f'(p+\theta_h\cdot h) - f(p)\big)h$. O liczbie θ_h nic więcej nie wiemy ponad to, że znajduje się ona w przedziale $(0,1)$, oznacza to, że liczba $p+\theta_h\cdot h$ leży między p i $p+h$. W przypadku funkcji $\sqrt{x}$ i przedziału $(49,50)$ pochodna zmienia się bardzo nieznacznie: maleje od wartości $\dfrac{1}{14}$ do wartości $\dfrac{1}{2\sqrt{50}}$. W przypadku funkcji x^2 rośnie od wartości $2\cdot 49 = 98$ przyjmowanej w punkcie 49 do wartości $2\cdot 50 = 100$ przyjmowanej w punkcie 50, w tym przypadku zmiana wartości pochodnej jest istotnie większa. W przypadku funkcji e^x pochodna zmienia się od wartości e^{49} do wartości e^{50}, czyli o $(e-1)\cdot e^{49}$, czyli o wielkość ogromną. Sama zmiana pochodnej jeszcze o niczym nie świadczy, bo zmiana mogłaby być skoncentrowana na bardzo krótkim przedziale kończącym się w punkcie 50. Tak jednak w tym przypadku nie jest. I właśnie dlatego widoczne są różnice w dokładności. W przypadku funkcji wykładniczej pochodna rośnie od wartości e^{49} do wartości e^{50}, tj. o wielkość ogromną $(e-1)e^{49} > 1,7\cdot e^{49}$. Można się więc było spodziewać, że w tym przypadku wzór $f(p+h) \approx f(p) + f'(p)h$ będzie bardzo niedokładny: funkcja wykładnicza zagina się mocno ku górze odchodząc szybko od stycznej do siebie w jakimś punkcie, np. w $(49, e^{49})$. W przypadku funkcji kwadratowe x^2 pochodna wzrasta od wartości 98 do wartości 100, a więc zmiana jej wartości jest znacznie mniej spektakularna, niemniej i w tym przypadku wykres funkcji oddala się od stycznej w widoczny sposób, też ku górze. W przypadku funkcji $\sqrt{x}$ pochodna maleje, ale bardzo powoli, więc wykres odchyla się od stycznej ku dołowi, ale efekt ten jest nieznaczny: wykres nieomal pokrywa się ze styczną, więc przybliżenie liniowe działa bardzo dobrze. ■

27. Przy różnych okazjach na lekcjach fizyki w szkołach wykorzystywana jest równość przybliżona $\sin x \approx x$, np. w optyce przy wyprowadzaniu równania soczewki lub zwierciadła, przy wyprowadzania wzoru na okres wahań wahadła matematycznego. Jest to zastosowanie omawianej przez nas równości przybliżonej $f(p+h) \approx f(p) + f'(p)h$ w przypadku funkcji $\sin$, $p=0$, $h=x$. W tym przypadku $f(0) = \sin 0 = 0$ i $f'(0) = \cos 0 = 1$ i wobec tego $f(p) + f'(p)h = x$. Wykazaliśmy poprzednio (przykład 17.), że dla $x > 0$ zachodzi nierówność $x - \frac{x^3}{6} < \sin x < x$, więc błąd przybliżenia $\sin x \approx x$ jest mniejszy niż $\frac{x^3}{6}$, więc jeśli kąt x jest mały, to ten błąd jest bardzo mały, np. jeśli $x = \frac{1}{10}$, to błąd jest mniejszy niż $\frac{1}{6000}$. Wypada przypomnieć, że mowa o wielkości kąta wyrażonej w radianach, 1 radian to nieco ponad $57°$. Człowieka o wzroście 2 m, więc niższego niż np. Małgorzata Dydek (koszykarka z Gdańska, jedna z najwyższych na świecie) widać z odległości 200 m pod kątem około $0,01$ radiana, więc mówimy o rzeczywiście istniejących kątach, małych ale nie o znikomo małych, występujących niezwykle rzadko. Rachunek różniczkowy pozwala oszacować błąd nie tylko z góry, ale również z dołu. W tym przypadku można posłużyć się metodą zastosowaną we wspomnianym przykładzie 17 w celu wykazania nierówności $\sin x < x - \frac{x^3}{6} + \frac{x^5}{120}$, która zachodzi dla $x > 0$. Z tej nierówności wynika natychmiast, że $\frac{x^3}{6} - \frac{x^5}{120} < x - \sin x < \frac{x^3}{6}$, a więc błąd przybliżenia jest większy niż $\frac{x^3}{6} - \frac{x^5}{120}$ i mniejszy niż $\frac{x^3}{6}$. Gdybyśmy zainteresowali się błędem względnym, tj. wielkością $\frac{x - \sin x}{x}$, to okazałoby się, że w przypadku $0 < x < 0,1$ jest on mniejszy niż $\frac{1}{6}(0,1)^2 = \frac{1}{600}$, czyli mniejszy niż $\frac{1}{6}\%$. To całkiem dobra dokładność. ■

28. Niech $f(x) = e^x$, $p = 0$. Mamy $f'0 = e^0 = 1$ oraz $f(0) = e^0 = 1$, zatem $e^x \approx 1 + x$. Zbadamy dokładność tego przybliżenia dla $x > 0$. W przykładzie 20 wykazaliśmy, że dla $x > 0$ zachodzi nierówność $e^x > 1 + x + \frac{1}{2}x^2$. Wobec tego błąd przybliżenia jest *większy* niż $\frac{1}{2}x^2$. Oszacujemy go teraz z góry. Pokażemy trzy metody.

Metoda pierwsza. Znajdziemy liczbę $a > 0$, taką że dla wszystkich $x \in (0,3)$ zachodzi nierówność $e^x - 1 - x < ax^2$. Przyjmijmy $f(x) = e^x - 1 - x - ax^2$. Mamy $f'(x) = e^x - 1 - 2ax$ oraz $(f')'(x) = e^x - 2a$. Jeśli $2a \geq e^3$, np. $a \geq \frac{1}{2} \cdot 21,952 = \frac{1}{2} \cdot 2{,}8^3 > \frac{1}{2} \cdot e^3$, to $(f')'$ przyjmuje na przedziale $(0,3)$ wartości ujemne, więc f' jest funkcją malejącą na przedziale $[0,3]$, a ponieważ $f'(0) = e^0 - 1 - 2a \cdot 0 = 0$, więc również f' przyjmuje na przedziale $(0,3)$ jedynie wartości ujemne. Stąd wnioskujemy, że funkcja f maleje na przedziale $[0,3]$. Ponieważ $f(0) = 0$, więc wartości funkcji f na przedziale $(0,3)$ są liczbami ujemnymi. Wykazaliśmy więc, że jeśli $a \geq \frac{1}{2}e^3$, to $e^x - 1 - x < ax^2$ dla $x \in (0,3)$, np. $e^x - 1 - x < 11 \cdot x^2$. Czytelnik bez trudu stwierdzi, że jeśli zastąpimy przedział $(0,3)$ przedziałem $(0,2)$, to otrzymamy rezultat nieco dokładniejszy: $e^x - 1 - x < ax^2$ dla $a \geq \frac{1}{2} \cdot e^2$, np. $a \geq 4 > 3{,}92 = \frac{1}{2} \cdot 2{,}8^2 > \frac{1}{2} \cdot e^2$.

Metoda druga. Jeśli $3 > x > 0$, to zachodzi nierówność $e^x - 1 - x = \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \ldots <$ $< \frac{1}{2!}x^2\left(1 + \left(\frac{x}{3}\right) + \left(\frac{x}{3}\right)^2 + \ldots\right) = \frac{x^2}{2!}\frac{1}{1-\frac{x}{3}}$ – skorzystaliśmy tu z tego, że $1 > \frac{x}{3} > \frac{x}{4} > \ldots$ i z wzoru na sumę szeregu geometrycznego.

Metoda trzecia. Wykażemy , że jeśli $3 > x > 0$, to $e^x - 1 - x < \frac{x^2}{2(1-\frac{x}{3})}$. Nie użyjemy stosowanych poprzednio szeregów. Niech $g(x) = e^x - 1 - x - \frac{x^2}{2(1-\frac{x}{3})} = e^x - 1 - x - \frac{3}{2} \cdot \frac{x^2}{(3-x)}$. Mamy wobec tego

$$g'(x) = e^x - 1 - \frac{3}{2} \cdot \frac{2x(3-x) + x^2}{(3-x)^2} = e^x - 1 - \frac{3}{2} \cdot \frac{6x - x^2}{(3-x)^2}.$$

Kontynuując obliczenia otrzymujemy

$$(g')'(x) = e^x - \frac{3}{2} \cdot \frac{(6-2x)(3-x)^2 + 2(6x - x^2)(3-x)}{(3-x)^4} = e^x - \frac{3}{2} \cdot \frac{18}{(3-x)^3} = e^x - \frac{1}{(1-\frac{x}{3})^3}.$$

Dla każdego $x > 0$ mamy $e^{-x/3} > 1 - \frac{x}{3}$, więc jeśli $0 < x < 3$, to zachodzi nierówność $\left(\frac{1}{1-\frac{x}{3}}\right)^3 > \left(e^{x/3}\right)^3 = e^x$. Z tej nierówności wynika, że dla $0 < x < 3$ zachodzi $(g')'(x) < 0$, więc na przedziale $[0,3]$ funkcja g' jest nierosnąca, więc dla $0 < x \leq 3$ zachodzi nierówność $g'(x) \leq g'(0) = 0$. Wobec tego, że funkcja g ma ujemną pochodną na przedziale $[0,3]$, jest ona ściśle malejąca na tym przedziale, zatem $g(x) < g(0) = 0$ dla $x \in (0,3]$, a to właśnie chcieliśmy wykazać. Wobec tego dla $0 < x < \frac{3}{2}$ zachodzi nierówność $e^x - 1 - x < x^2$, bo w tym przypadku $2\left(1 - \frac{x}{3}\right) > 1$.

W przykładzie 20 wykazaliśmy, że dla $x < 0$ zachodzi nierówność $e^x < 1 + x + \frac{1}{2}x^2$. Stąd wynika, że dla $x < 0$ zachodzi nierówność $0 < e^x - (1+x) < \frac{1}{2}x^2$, zaś dla $3 > x > 0$ – nierówność $0 < e^x - (1+x) < x^2$. Stąd już łatwo wynika, że dla $x < \frac{3}{2}$ zachodzi nierówność $0 \leq e^x - (1+x) \leq x^2$, przy czym równość ma miejsce wtedy i tylko wtedy, gdy $x = 0$. W rozdziale pierwszym rozważaliśmy, jaką kwotę powinien wypłacić bank osobie, która wpłaciła kwotę k, jeśli oprocentowanie jest równe $100x\%$ w skali rocznej, a procenty są doliczane w sposób ciągły. Okazało się, że tą kwotą jest ke^x. Jeśli np. $x = 0,1$, czyli oprocentowanie w skali rocznej równe jest 10%, to różnica między wzorem liniowym (wypłacana jest kwota $k(1+x) = 1,1 \cdot k$) a dokładnym (wypłacana jest kwota $ke^x = k \cdot e^{0,1}$) jest mniejsza niż $k \cdot .0,1^2 = 0,01 \cdot k$. Z nierówności $e^x - (1+x) > \frac{1}{2}x^2$, która ma miejsce dla liczb $x > 0$, wynika, że różnica ta jest większa niż $\frac{1}{2} \cdot 0,1^2 \cdot k = 0,005 \cdot k$. Oczywiście przy małych kwotach różnica taka nie ma żadnego znaczenia praktycznego, jednak przy dużych jest inaczej, bo choć procentowo nie ulega to zmianie, to kwota może być znacząca. Efekt ten staje się bardziej widoczny, gdy rozpatrywany jest dłuższy okres czasu, np. 2 lata. Wtedy wzór liniowy daje wypłatę $k(1+2x)$, zaś nieliniowy – wypłatę ke^{2x}. W przypadku $x = 0,1$ różnica między tymi kwotami staje się większa niż $k\cdot\dfrac{0,2^2}{2}=k\cdot 0,02$, co oznacza, że błąd wzrósł w istotny sposób. Wspominaliśmy wcześniej, że podobne rozważania można prowadzić w fizyce przy dyskusji wzoru na długość np. pręta żelaznego w zależności od jego temperatury. Prowadzi to do wzoru $l(t)=l(t_0)e^{\lambda(t-t_0)}$, gdzie przez $l(t)$ oznaczyliśmy długość pręta w temperaturze t, zaś λ oznacza współczynnik rozszerzalności cieplnej, w przypadku żelaza $\lambda\approx 0,00000115=1,15\cdot 10^{-5}$. Jeśli zmiana temperatury jest niezbyt duża, np. mniejsza niż 50° C, to wykładnik jest mniejszy niż $0,00006$, więc jego kwadrat jest mniejszy niż $0,00000004$, co oznacza, że błąd, który popełnimy zastępując $e^{\lambda(t-t_0)}$ przez $1+\lambda(t-t_0)$ będzie mniejszy niż $0,00000004\cdot l(t_0)$, więc w przypadku np. szyny kolejowej – mniejszy od dokładności pomiaru jej długości. Inaczej jest w przypadku rozpadu promieniotwórczego. W wyniku rozważań analogicznych do tych, które doprowadziły nas do wzoru $e^x=\lim\limits_{n\to\infty}\left(1+\dfrac{x}{n}\right)^n$ otrzymujemy wzór $m(t)=m(t_0)e^{-\lambda(t-t_0)}$, gdzie $m(t)$ oznacza masę substancji promieniotwórczej w chwili t, a λ – stałą rozpadu. Rzecz w tym, iż w tym przypadku interesuje nas np. czas połowicznego rozpadu, to znaczy czas, w którym masa substancji zmniejsza się o połowę. W tym przypadku $t-t_0$ musi być tak duże, by zachodził wzór $\lambda(t-t_0)=\ln 2\approx 0,6931$, więc błąd spowodowany stosowaniem przybliżenia liniowego funkcji wykładniczej funkcją liniową byłby większy niż $\dfrac12\cdot 0,6931^2\approx 0,24$, więc w zasadzie niedopuszczalny jako za duży (24%).* Przykład powinien uświadomić studentom, że przed stosowaniem wzorów przybliżonych warto zastanowić się nad tym, czy wolno je stosować. ■

3. Badanie funkcji za pomocą pochodnych: wypukłość

W poprzednim rozdziale zdefiniowaliśmy funkcje wypukłe i wklęsłe. Pokazaliśmy jak można dowodzić, że funkcja ciągła jest wypukła. Teraz pokażemy, jak można to robić w przypadku funkcji różniczkowalnej. Powiążemy też wyraźnie pojęcie wypukłości funkcji z pojęciem stycznej do jej wykresu. Przypomnijmy, że funkcją wypukłą nazywaliśmy funkcję określoną na zbiorze wypukłym (jedynymi wypukłymi podzbiorami prostej są przedziały, zbiory jednopunktowe oraz zbiór pusty), taką że dla dowolnych punktów x,y z jej dziedziny i dowolnej liczby $f\in(0,1)$ zachodzi nierówność $f(tx+(1-t)y)\le tf(x)+(1-t)f(y)$, co oznacza, że punkty odcinka o końcach $(x,f(x))$ i $(y,f(y))$ leżą *nad* wykresem funkcji f lub na tym wykresie, niezależnie od wyboru punktów x i y. Przypomniana właśnie definicja jest równoważna temu, że spełniony jest jeden (którykolwiek) z trzech warunków:

(a) $\dfrac{f(y)-f(x)}{y-x}<\dfrac{f(z)-f(x)}{z-x}$ dla każdych x,y,z z dziedziny funkcji f, dla których $x<y<z$,

(b) $\dfrac{f(y)-f(x)}{y-x}<\dfrac{f(z)-f(y)}{z-y}$ dla każdych x,y,z z dziedziny funkcji f, dla których $x<y<z$,

(c) $\dfrac{f(x)-f(z)}{x-z}<\dfrac{f(y)-f(z)}{y-z}$ dla każdych x,y,z z dziedziny funkcji f, dla których $x<y<z$.

Udowodnimy teraz twierdzenie, które charakteryzuje funkcje wypukłe w terminach pochodnych. Przed sformułowaniem go wprowadzimy oznaczenia: $f'_-(x)=\lim\limits_{h\to 0^-}\dfrac{f(x+h)-f(x)}{h}$ dla oznaczenia *lewostronnej pochodnej* funkcji f w punkcie x oraz $f'_+(x)=\lim\limits_{h\to 0^+}\dfrac{f(x+h)-f(x)}{h}$ dla *pochodnej*

* W szkołach wzór na zmianę długości w wyniku podgrzania występuje w innej klasie niż wzór na zmianę masy pierwiastka promieniotwórczego w czasie, więc liczba uczniów, którzy zauważają niekonsekwencję w stosowaniu w jednym przypadku funkcji liniowej, a w drugim funkcji wykładniczej jest zaniedbywalnie mała. Można podejrzewać, że nie wszyscy nauczyciele maja czas i ochotę wyjaśniać, dlaczego w jednym przypadku stosowany jest jeden wzór, a w drugim – inny.

prawostronnej.

Twierdzenie o pochodnej funkcji wypukłej

Jeśli f jest funkcją wypukłą określoną na przedziale otwartym P, to

W1. w każdym punkcie $x \in P$ istnieją pochodne jednostronne $f'_-(x)$ i $f'_+(x)$ i $f'_-(x) \leq f'_+(x)$;

W2. jeśli $x, y \in P$ i $x < y$, to $f'_+(x) \leq f'_-(y)$, przy czym jeśli f jest ściśle wypukła, to nierówność jest ostra;

W3. funkcja f jest ciągła w każdym punkcie przedziału *otwartego* P.

Dowód.

Niech D_x, gdzie $D_x(t) = \dfrac{f(t) - f(x)}{t - x}$ dla dowolnego punktu $t \in P \setminus \{x\}$, oznacza iloraz różnicowy funkcji f w punkcie x. Załóżmy, że $u < v < x < r < s$ są punktami przedziału P. Z własności (c) wynika, że $D_x(u) \leq D_x(v)$. Z własności (b) wynika z kolei, że $D_x(v) \leq D_x(r)$, zaś z własności (a) wynika, że $D_x(r) \leq D_x(s)$. Mamy więc $D_x(u) \leq D_x(v) \leq D_x(r) \leq D_x(s)$. Oznacza to, że funkcja D_x jest niemalejąca w całym zbiorze $P \setminus \{x\}$. Ma więc granice jednostronne w każdym punkcie przedziału P, w tym w punkcie x. Zachodzą oczywiste równości: $\lim\limits_{t \to x^-} D_x(t) = f'_-(x)$ oraz $\lim\limits_{t \to x^+} D_x(t) = f'_+(x)$, przy czym $f'_-(x) \leq D_x(r)$, i wobec tego $f'_-(x) \leq f'_+(x)$. Pierwsza część twierdzenia została udowodniona.
Załóżmy teraz, że $x < r < y$. Z własności (b) wynika, że $D_x(r) \leq D_y(r)$, a z tego co udowodniliśmy dotychczas wynikają nierówności $f'_+(x) \leq D_x(r)$ oraz $D_y(r) \leq f'_-(y)$. Z trzech otrzymanych nierówności wynika, że $f'_+(x) \leq f'_-(y)$. Uzyskaliśmy więc drugą część tezy.
Z istnienia jednostronnych pochodnych *skończonych* w punkcie x wynika, że funkcja f jest w tym punkcie lewo– i prawostronnie ciągła, więc jest ciągła. Stwierdzenie tego, że w przypadku funkcji ściśle wypukłej nierówności stają się ostre wynika od razu z tego, że w przypadku funkcji ściśle wypukłej nierówności w (a), (b), (c) są ostre. Dowód został zakończony. ■

Wniosek z dowodu twierdzenia

Jeśli f jest funkcją wypukłą określoną na przedziale otwartym P, to dla dowolnego $h > 0$, takiego że $x + h \in P$ zachodzi nierówność $f(x+h) \geq f(x) + f'_+(x)h$. Jeśli $x - h \in P$, to zachodzi nierówność $f(x-h) \geq f(x) - f'_-(x)h$. W przypadku funkcji ściśle wypukłej nierówności te są ostre.

Dowód.

Wynika to natychmiast z tego, że $f'_+(x) \leq D_x(x+h) = \dfrac{f(x+h) - f(x)}{h}$ w pierwszym przypadku. W drugim przypadku z tego, że $f'_-(x) \geq D_x(x-h) = \dfrac{f(x-h) - f(x)}{-h}$. ■

Wykazane twierdzenie oznacza, że pochodna różniczkowalnej funkcji wypukłej jest niemalejąca. Wniosek to po prostu stwierdzenie, że wykres funkcji wypukłej leży nad styczną do siebie w dowolnym punkcie wewnętrznym przedziału–dziedziny. Przy okazji okazuje się, że funkcja wypukła może być nieróżniczkowalna w pewnych punktach, np. $|x|$, $|x+1| + |x| + |x-1|$ lub $e^{|x|}$, ale w punktach wewnętrznych dziedziny ma skończone pochodne jednostronne, więc jest „niedaleka” od funkcji różniczkowalnej. Wypada nadmienić, że te uwagi nie dotyczą końców przedziału–dziedziny, w których funkcja wypukła może nie być ciągła, np. jeśli $f(x) = x^2$ dla $x > 0$ i $f(0) = 1$, to f jest ściśle wypukła na półprostej domkniętej $[0, \infty)$, choć jest nieciągła w punkcie 0, więc tym bardziej nie ma w tym punkcie pochodnej. Takimi funkcjami nie będziemy się jednak zajmować, bo skłonni jesteśmy przyznać, że są one nieco sztuczne.

W przykładzie 18 pojawiła się nierówność $e^x > 1 + x$ dla $x \neq 0$. Teraz możemy ją wywnioskować ze ścisłej wypukłości funkcji e^x na przedziale $(-\infty,\infty)$. Podobnie nierówność $\sin x < x$ dla $0 < x < \pi$ jest konsekwencją ścisłej wklęsłości funkcji sinus na przedziale $[0,\pi]$. Jeśli $0 < x \neq 1$, to $\ln x < x-1$, co wynika z tego, że funkcja ln jest ściśle wklęsła na $(0,\infty)$, co wykażemy niebawem. Widzimy więc, że również w ten sposób można uzyskiwać różne oszacowania. Warto więc umieć wyjaśnić, czy funkcja na określonym przedziale jest wypukła, wklęsła, czy też ani wypukła, ani wklęsła. Okazuje się, że w wielu przypadkach można to wyjaśnić badając pochodną interesującej nas funkcji.

Twierdzenie o wypukłości funkcji, której pochodna jest niemalejąca

Jeśli funkcja f jest zdefiniowana na przedziale otwartym P i ma w punktach wewnętrznych tego przedziału jednostronne pochodne f'_+ i f'_-, dla których zachodzą warunki:

W1. dla każdego $x \in P$ zachodzi nierówność $f'_-(x) \leq f'_+(x)$,

W2. jeśli $x < y$ i $x, y \in P$, to $f'_+(x) \leq f'_-(y)$,

to funkcja f jest wypukła na przedziale P. Jeżeli nierówność w warunku W2 jest ostra, to funkcja f jest ściśle wypukła. W szczególności:

funkcja różniczkowalna f jest wypukła wtedy i tylko wtedy, gdy jej pochodna f' jest niemalejąca, ściśle wypukła – wtedy i tylko wtedy, gdy jej pochodna f' jest ściśle rosnąca.

Dowód.

Udowodnimy to twierdzenie dla funkcji różniczkowalnych, bo w tym przypadku dowód jest bardzo prosty, dowód wersji ogólnej przedstawimy na końcu paragrafu. Z wypukłości funkcji wynika, że jej pochodna jest niemalejąca – jest to wniosek z poprzedniego twierdzenia. Zakładamy więc, że funkcja f jest różniczkowalna, a jej pochodna f' jest niemalejąca: $x < y \Longrightarrow f'(x) \leq f'(y)$. By udowodnić, że funkcja f jest wypukła, wystarczy wykazać, że jeśli $x < y < z$, to $\frac{f(x)-f(y)}{x-y} \leq \frac{f(y)-f(z)}{y-z}$. Z twierdzenia o wartości średniej wynika, że istnieją punkty $r \in (x,y)$ oraz $s \in (y,z)$, takie że $\frac{f(x)-f(y)}{x-y} = f'(r)$ oraz $\frac{f(y)-f(z)}{y-z} = f'(s)$. Ponieważ $r < y < s$, więc $r < s$ i wobec tego $f'(r) \leq f'(s)$, co kończy dowód twierdzenia w tym przypadku. ■

Dowód w przypadku ogólnym pozostawiam w charakterze ćwiczenia, bardzo zachęcam do jego przeprowadzenia!

Przykłady cd.

29. Funkcja x^a jest ściśle wypukła na półprostej $(0,+\infty)$ dla $a > 1$ oraz dla $a < 0$, natomiast w przypadku $0 < a < 1$ jest ściśle wklęsła. Wynika to natychmiast z twierdzenia o wypukłości funkcji o niemalejącej pochodnej, bowiem $(x^a)' = ax^{a-1}$ i wobec tego $\big((x^a)'\big)' = a(a-1)x^{a-2}$, więc funkcja $\big((x^a)'\big)'$ jest dodatnia na półprostej $(0,+\infty)$ w przypadku $a > 1$ oraz $a < 0$, natomiast w przypadku $0 < a < 1$ – funkcja ta jest ujemna, z czego wynika, że pochodna $(x^a)'$ rośnie w pierwszych dwóch przypadkach, natomiast w trzecim – maleje. ■

30. *Uogólniona nierówność Bernoulliego.* Jeśli zachodzą nierówności $a > 1$ lub $a < 0$ i $-1 < x \neq 0$, to $(1+x)^a > 1 + ax$. Jeśli natomiast $0 < a < 1$ oraz $-1 < x \neq 0$, to $(1+x)^a < 1 + ax$.

Wynika to od razu z wyników poprzedniego przykładu, z tego że pochodną funkcji $(1+x)^a$ w punkcie 0 jest liczba a oraz z tego, że wykres funkcji ściśle wypukłej leży nad styczną mając z nią dokładnie jeden punkt wspólny, zaś wykres funkcji ściśle wklęsłej leży pod styczną mając z nią dokładnie jeden punkt wspólny. ■

31. Funkcja wykładnicza a^x o podstawie dodatniej $a \neq 1$ jest ściśle wypukła. Mamy bowiem $(a^x)' = \big(e^{x\ln a}\big)' = e^{x\ln a} \cdot \ln a = a^x \cdot \ln a$. Wobec tego $\big((a^x)'\big)' = a^x \cdot (\ln a)^2 > 0$ dla każdego x, więc funkcja $(a^x)'$ jest ściśle rosnąca na całej prostej, a wobec tego funkcja a^x jest ściśle wypukła. Wynika stąd, między innymi, że wykres funkcji wykładniczej leży nad styczną (w dowolnym punkcie), np. $a^x > 1 + x \cdot \ln a$ dla $x \neq 0$ i $0 < a \neq 1$. ■

32. Funkcja $\log_a x$ jest ściśle wklęsła na półprostej $(0, +\infty)$ w przypadku $a > 1$, natomiast w przypadku $0 < a < 1$ funkcja $\log_a x$ jest ściśle wypukła. Wynika to z tego, że $\log_a x = \dfrac{\ln x}{\ln a}$, wobec czego $(\log_a x)' = \dfrac{1}{x \ln a}$, więc pochodna ta jest ściśle malejąca w przypadku $\ln a > 0$, czyli w przypadku $a > 1$ oraz – ściśle rosnąca w przypadku $\ln a < 0$, czyli $0 < a < 1$. ■

33. Funkcja sinus jest ściśle wklęsła na każdym przedziale postaci $[2n\pi, (2n+1)\pi]$, zaś na przedziałach postaci $[(2n-1)\pi, 2n\pi]$ jest ona ściśle wypukła, n oznacza tu dowolną liczbę całkowitą. Wynika to stąd, że $(\sin x)' = \cos x$ i tego że funkcja kosinus maleje na przedziałach postaci $[2n\pi, (2n+1)\pi]$ i rośnie na przedziałach postaci $[(2n-1)\pi, 2n\pi]$. Ze ścisłej wklęsłości funkcji sinus na przedziale $\big[0, \frac{\pi}{2}\big]$ wynika, że jeśli $0 < x < \frac{\pi}{2}$, to $\dfrac{2x}{\pi} < \sin x < x$ – wykres leży nad sieczną (odcinkiem o końcach $(0,0)$ i $(\frac{\pi}{2}, 1)$) i pod styczną (w punkcie $(0,0)$). Druga z tych nierówności znamy już od dawna, ale warto raz jeszcze podkreślić jej związek z wklęsłością funkcji sinus. ■

34. Funkcja tangens jest ściśle wypukła na każdym przedziale postaci $[n\pi, n\pi + \frac{\pi}{2})$ zaś na każdym przedziale postaci $(n\pi - \frac{\pi}{2}, n\pi]$ jest ściśle wklęsła, n oznacza tu dowolną liczbę całkowitą. Wynika to, z tego że na przedziałach postaci $[n\pi, n\pi + \frac{\pi}{2})$ pochodna funkcji tangens, czyli funkcja $1 + \mathrm{tg}^2\, x$ rośnie, zaś na przedziałach postaci $(n\pi - \frac{\pi}{2}, n\pi]$ – maleje. ■

Analiza przykładów wskazuje na to, że zdarzają się funkcje, które w całej swej dziedzinie nie są ani wypukłe ani wklęsłe. W podanych przykładach zdarzało się tak, że po jednej stronie pewnego punktu mieliśmy do czynienia z funkcją wypukłą a po drugiej – z wklęsłą. Przy szkicowaniu wykresów funkcji rozsądnie jest znaleźć takie punkty zawczasu. Mają one swą nazwę.

Definicja punktu przegięcia

Punkt p jest jest punktem przegięcia funkcji f wtedy i tylko wtedy, gdy funkcja f jest różniczkowalna w punkcie p i istnieje liczba $\delta > 0$, taka że:

przedział $(p - \delta, p + \delta)$ jest zawarty w dziedzinie funkcji f,

na jednym z przedziałów $(p - \delta, p]$, $[p, p + \delta)$ funkcja f jest wypukła, a na drugim wklęsła,

na żadnym z przedziałów $(p - \eta, p]$, $[p, p + \eta)$, gdzie $\eta \in (0, \delta)$, funkcja f nie jest liniowa. ■

Wypada dodać, że w literaturze istnieje kilka nierównoważnych definicji punktu przegięcia, jednak wszystkie one pokrywają się w przypadku najprostszych funkcji. Przykładowym określeniem punktu przegięcia nierównoważnym podanemu wyżej jest: p jest punktem przegięcia funkcji f wtedy i tylko wtedy, gdy wykresu funkcji f ma styczną w punkcie $(p, f(p))$ przy czym z jednej strony tego punktu wykres znajduje się pod tą styczną, a z drugiej – nad nią. Czytelnik może sprawdzić, że 0 jest punktem przegięcia funkcji f zdefiniowanej wzorami $f(0) = 0$, $f(x) = x^2(2 + \sin\frac{1}{x})$ dla $x > 0$ oraz $f(x) = -x^2(2 + \sin\frac{1}{x})$ dla $x < 0$ w sensie drugiego określenia, ale nie jest punktem przegięcia w sensie definicji, którą podaliśmy wcześniej. Natomiast 0 byłoby punktem przegięcia funkcji f zdefiniowanej wzorami $f(x)=-x^2$ dla $x<0$ oraz $f(x)=x+x^2$ dla $x\geq 0$ gdybyśmy w definicji nie założyli różniczkowalności funkcji f w punkcie przegięcia. Bez założenie różniczkowalności 0 byłoby punktem przegięcia funkcji g zdefiniowanej wzorami $g(x)=-x^2$ dla $x\leq 0$ i $g(x)=x(x-2)$ dla $x\geq 0$, a to już nie wyglądałoby dobrze. Niestety, matematycy nie ustalili tej definicji na tyle sztywno, by jedna jej wersja została przyjęta przez wszystkich, więc czytając różne podręczniki można spotykać się z istotnie różnymi definicjami, które jednak w przypadku funkcji zdefiniowanych „za pomocą jednego wzoru"* dają ten sam rezultat.

Jest jasne, że punkty postaci $n\pi$ są punktami przegięcia funkcji sinus oraz funkcji tangens, że 0 jest punktem przegięcia funkcji x^{2n+1} dla $n=1,2,3,\ldots$, że funkcja postaci $ax+b$ nie ma punktów przegięcia, że funkcja x^{2n} dla $n=1,2,3,\ldots$ nie ma punktów przegięcia, bowiem jest ściśle wypukła. Funkcja $\sqrt[7]{x}$ zdefiniowana na całej prostej ma punkt przegięcia w 0, choć nie jest w tym punkcie różniczkowalna (ma pochodną, ale równą $+\infty$), bo jest ściśle wypukła na półprostej $(-\infty,0]$ zaś na półprostej $[0,+\infty)$ jest ściśle wklęsła. Te przykłady można mnożyć, ale nie będziemy tego robić, bo to proste pojęcie nie przysparza większych problemów studentom.

4. Symbole nieoznaczone, reguła de l'Hospitala

Często zachodzi potrzeba obliczenia granicy ilorazu dwu funkcji, gdy granica każdej z nich równa jest 0 lub ∞. Zdarza się, że trzeba obliczyć granicę iloczynu dwu funkcji, z których jedna ma granicę 0, a druga – ∞. Ten drugi przypadek można sprowadzić do pierwszego: $fg=\dfrac{f}{1/g}=\dfrac{g}{1/f}$. Bywa, że interesuje nas granica wyrażenia f^g przy czym granicą f jest 1, a granicą g jest ∞. Wzór $f^g=e^{g\cdot\ln f}$ pozwala problem zredukować do obliczania granicy iloczynu, więc w dalszym ciągu do obliczania granicy ilorazu. Zdarzają się też inne sytuacje, w których nie są spełnione założenia dotychczas sformułowanych twierdzeń o granicach. Podobnie jak w przypadku ciągów istnieje twierdzenie, które w wielu sytuacjach ułatwia znalezienie granicy. Jest to tzw. reguła de l'Hospitala, francuskiego markiza, który po wysłuchaniu wykładów Jana Bernoulliego wydał drukiem notatki z nich pod tytułem *Analyse des infiniment petites***, co spowodowało protesty rzeczywistego autora tekstu, ale wtedy nie istniało jeszcze pojęcie praw autorskich. Twierdzenie, które znajduje się niżej, pochodzi z tej właśnie książki (i – według historyków matematyki – powinno mieć inną nazwę).

Reguła de l'Hospitala

Załóżmy, że funkcje $f,g\colon(a,b)\longrightarrow\mathbb{R}$ są różniczkowalne w każdym punkcie przedziału (a,b), że $g(x)\neq 0\neq g'(x)$ dla każdego $x\in(a,b)$, że istnieje granica $\lim\limits_{x\to a}\frac{f'(x)}{g'(x)}=G\in\overline{\mathbb{R}}=\mathbb{R}\cup\{-\infty,+\infty\}$ oraz że spełniony jest jeden z dwóch warunków:

$$1^\circ\ \lim_{x\to a}f(x)=0=\lim_{x\to a}g(x),\qquad\qquad 2^\circ\ \lim_{x\to a}|g(x)|=+\infty.$$

Wtedy iloraz $\frac{f(x)}{g(x)}$ ma granicę przy $x\to a$ i zachodzi równość $\lim\limits_{x\to a}\frac{f(x)}{g(x)}=G=\lim\limits_{x\to a}\frac{f'(x)}{g'(x)}$.

Dowód w najprostszym przypadku.

Udowodnimy najpierw to twierdzenie przy bardzo mocnych założeniach. Chodzi nam o to, by wyjaśnić jego sens. Dowód w przypadku ogólnym podamy nieco później. Założymy mianowicie, że $a>-\infty$, że zachodzi warunek 1° oraz że istnieją *skończone* granice $\lim\limits_{x\to a}f'(x)$ i $\lim\limits_{x\to a}g'(x)$, przy czym ta druga jest różna od 0. W tej sytuacji można dookreślić funkcje f,g w punkcie a przyjmując $f(a)=0=g(a)$.

* Chodzi o tzw. funkcje analityczne, których definicję podamy w następnym rozdziale.

** Analiza nieskończenie małych

Z twierdzenia Lagrange'a o wartości średniej zastosowanego do funkcji f rozpatrywanej na przedziale $[a,x]$ wynika, że $\frac{f(x)-f(a)}{x-a} = f'(c_x)$ dla pewnego punktu $c_x \in (a,x)$. Stąd wynika natychmiast, że $\lim\limits_{x\to a} \frac{f(x)-f(a)}{x-a} = \lim\limits_{x\to a} f'(x)$. Wykazaliśmy więc, że w tym przypadku funkcję f można potraktować jako określoną w punkcie a i to w taki sposób, że $f'(a) = \lim\limits_{x\to a} f'(x)$. To samo dotyczy oczywiście funkcji g. Oczywiście w obu przypadkach mamy na myśli różniczkowalność prawostronną. Niech $r(h) = \dfrac{f(a+h)-f(a)-f'(a)h}{h}$ dla $h \neq 0$ oraz $r(0)=0$. Jest oczywiście $\lim\limits_{h\to 0} r(h) = 0$. Analogicznie niech $\rho(h) = \dfrac{g(a+h)-g(a)-g'(a)h}{h}$. Wtedy $\lim\limits_{h\to 0} \rho(h) = 0$. Stąd możemy wywnioskować, że

$$\frac{f(x)}{g(x)} = \frac{f(x)-0}{g(x)-0} = \frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f'(a)(x-a)+(x-a)r(x-a)}{g'(a)(x-a)+(x-a)\rho(x-a)} = \frac{f'(a)+r(x-a)}{g'(a)+\rho(x-a)} \xrightarrow[x\to a]{} \frac{f'(a)}{g'(a)}.$$

Ostatnie przejście graniczne jest wykonalne, bo założyliśmy, że $g'(a) \neq 0$.

Zaraz od tego i innych zbędnych założeń uwolnimy się.

Zauważmy, że ponieważ pochodna funkcji g jest różna od 0 w każdym punkcie przedziału (a,b), więc funkcja g jest różnowartościowa – wynika to z twierdzenia Rolle'a. Ponieważ jest to funkcja ciągła i różnowartościowa, więc jest ściśle monotoniczna. Ponieważ zamiast ilorazu $\frac{f}{g}$ można rozważać iloraz $\frac{-f}{-g}$ i jedna z funkcji g, $-g$ jest ściśle rosnąca, a druga – ściśle malejąca, więc możemy przyjąć, że g jest ściśle rosnąca. W dalszym ciągu zakładamy, więc że g jest ściśle rosnąca, więc dla każdego $x \in (a,b)$ musi być $g'(x) > 0$ (przypominamy, że $g'(x) \neq 0$ w całym przedziale (a,b)).*

Niech m, M będą dwiema liczbami rzeczywistymi, takimi że $m < G < M$. Jeśli $G = -\infty$, to oczywiście nie rozpatrujemy m, jeśli $G = +\infty$, to nie rozpatrujemy M. Niech $\widetilde{m}, \widetilde{M}$ będą takimi liczbami, że $m < \widetilde{m} < G < \widetilde{M} < M$. Ponieważ $\lim\limits_{x\to a^+} \dfrac{f'(x)}{g'(x)} = G$, więc istnieje liczba $c \in (a,b)$, taka że $\widetilde{m} < \dfrac{f'(x)}{g'(x)} < \widetilde{M}$ dla $x \in (a,c)$. Wobec tego na przedziale (a,c) funkcje $f' - \widetilde{m}g'$ oraz $\widetilde{M}g' - f'$ są dodatnie, zatem funkcje $f - \widetilde{m}g$ oraz $\widetilde{M}g - f$ są na tym przedziale rosnące.

Rozważymy przypadek pierwszy: $\lim\limits_{x\to a^+} f(x) = 0 = \lim\limits_{x\to a^+} g(x)$. Ponieważ g jest ściśle rosnąca i ma granicę 0 w lewym końcu dziedziny, więc jest dodatnia na przedziale (a,b). Funkcje $f - \widetilde{m}g$ oraz $\widetilde{M}g - f$ mają granicę (prawostronną) 0 w punkcie a, więc są dodatnie. Mamy więc dla $x \in (a,c)$ nierówność podwójną $\widetilde{m} < \dfrac{f(x)}{g(x)} < \widetilde{M}$, więc również $m < \dfrac{f(x)}{g(x)} < M$. Ponieważ m oznacza tu dowolna liczbę mniejszą niż G, a M – dowolna większą niż G, więc $\lim\limits_{x\to a^+} \dfrac{f(x)}{g(x)} = G$, co kończy dowód w przypadku pierwszym.

Teraz zakładamy, że $\lim\limits_{x\to a^+} |g(x)| = +\infty$. Ponieważ zakładamy, że $g'(x) > 0$ w przedziale (a,b), więc $\lim\limits_{x\to a^+} g(x) = -\infty$. Ponieważ funkcje $f - \widetilde{m}g$ oraz $\widetilde{M}g - f$ są rosnące na przedziale (a,c), więc dla $x \in (a,c)$ mamy $f(x) - \widetilde{m}g(x) < f(c) - \widetilde{m}g(c)$ oraz $\widetilde{M}g(x) - f(x) < \widetilde{M}g(c) - f(c)$. Ponieważ $\lim\limits_{x\to a^+} g(x) = -\infty$, więc możemy przyjąć, po ewentualnym zmniejszeniu c, że $g(x) < 0$ dla $x \in (a,c)$.

* Nie zakładamy ciągłości funkcji g', więc nie wolno nam skorzystać z własności Darboux. W istocie rzeczy może się zdarzyć, że pochodna funkcji określonej na przedziale ma punkty nieciągłości, jednak nawet w takiej sytuacji przysługuje jej własność Darboux, zresztą właśnie ten matematyk udowodnił, że pochodna funkcji różniczkowalnej ma własność przyjmowania wartości pośrednich. Nietrudny w istocie rzeczy dowód tego stwierdzenia można podać przyjrzawszy się właśnie przeprowadzonemu rozumowaniu.

Otrzymujemy więc nierówność podwójną:

$$\widetilde{m}+\frac{f(c)-\widetilde{m}g(c)}{g(x)}<\frac{f(x)}{g(x)}<\widetilde{M}+\frac{\widetilde{M}g(c)-f(c)}{g(x)}\,.$$

Ponieważ $\lim\limits_{x\to a^+}\frac{f(c)-\widetilde{m}g(c)}{g(x)}=0=\lim\limits_{x\to a^+}\frac{\widetilde{M}g(c)-f(c)}{g(x)}$, więc istnieje $d\in(a,c)$, takie że dla $x\in(a,d)$ zachodzi nierówność $m<\frac{f(x)}{g(x)}<M$. Konsekwencją ostatniego stwierdzenia jest równość $\lim\limits_{x\to a^+}\frac{f(x)}{g(x)}=G$. Dowód został zakończony. ■

W dowodzie tym wykorzystaliśmy w istotny sposób założenia $f(a)=g(a)=0$. Oczywiście bez tych założeń teza może być w konkretnej sytuacji prawdziwa jedynie przypadkiem – pochodne decydują o wielkości funkcji w otoczeniu punktu, w którym wartością funkcji jest 0, jeśli $f(a)\neq 0$, to „w pierwszym przybliżeniu" $f(x)\approx f(a)$!

Zauważmy jeszcze, że twierdzenie pozostaje prawdziwe dla granicy $\lim\limits_{x\to b}\frac{f(x)}{g(x)}$ po dokonaniu odpowiednich kosmetycznych zmian w założeniach i w tezie. Z tego zdania wynika, że można je też stosować w przypadku granic dwustronnych

Warto zauważyć, że istnieje analogia między regułą de l'Hospitala i twierdzeniem Stolza. Te rozważania nie będą ścisłe, bo mówić tu będziemy raczej o intuicjach. Ciąg można traktować jako funkcję określoną na zbiorze wszystkich liczb naturalnych. Wtedy $b=+\infty$. Niestety dziedzina nie jest w tym przypadku przedziałem, więc nie można mówić o pochodnej. Można jednak spojrzeć na zagadnienie nieco inaczej. Pochodna była nam potrzebna do oszacowania różnicy $f(x)-f(a)$, przy czym interesowała nas minimalna możliwa zmiana argumentu. Pisaliśmy przy odpowiednich założeniach, że $\frac{f(x)-f(a)}{g(x)-g(a)}\approx\frac{f'(a)}{g'(a)}$. W przypadku ciągu minimalna możliwa zmiana argumentu to 1. Wobec tego zamiast ilorazu pochodnych $\frac{f'(x)}{g'(x)}$, który przybliża interesujący nas iloraz różnicowy $\frac{\frac{f(x+h)-f(x)}{h}}{\frac{g(x+h)-g(x)}{h}}=\frac{f(x+h)-f(x)}{g(x+h)-g(x)}$ rozpatrujemy iloraz $\frac{a_{n+1}-a_n}{b_{n+1}-b_n}$. W twierdzeniu Stolza zakładaliśmy, że ciąg (b_n) jest ściśle monotoniczny. W regule de l'Hospitala też występuje to założenie, zakładamy mianowicie, że pochodna funkcji g nie przyjmuje wartości 0*, z czego wynika, że jest ona albo dodatnia, albo ujemna, a to pociąga za sobą ścisłą monotoniczność funkcji g.

Pokażemy teraz na kilku przykładach, jak można stosować regułę de l'Hospitala. Niektóre z podanych rezultatów zostały uzyskane wcześniej lub można je było uzyskać używając twierdzeń wykazanych wcześniej.

Przykłady cd.

35. $\lim\limits_{x\to\infty}\dfrac{x^a}{e^x}=0$. Możemy próbować zastosować regułę de l'Hospitala, bo mianownik ma granicę nieskończoną i jego pochodna, e^x, jest różna od 0 wszędzie. Nie jest istotne jaka jest granica licznika, a nawet czy licznik ma granicę. Iloraz pochodnych to $\dfrac{ax^{a-1}}{e^x}$, więc jest to wyrażenie tego samego typu co wyjściowe. Istotną zmianą jest pojawienie się w wykładniku $a-1$ w miejsce a. Jeśli $a\leq 1$, to licznik jest ograniczony z góry na półprostej $[1,+\infty)$, a mianownik dąży do $+\infty$, więc iloraz dąży do 0. Jeśli $a>1$, to stosujemy regułę de l'Hospitala $k\geq a$ razy. Omówimy to dokładniej. Po k–krotnym zróżniczkowaniu w liczniku pojawia się wyrażenie $a(a-1)(a-2)\cdot\ldots\cdot(a-k+1)x^{a-k}$, w mianowniku natomiast mamy e^x. Funkcja $a(a-1)(a-2)\cdot\ldots\cdot(a-k+1)x^{a-k}$ jest ograniczona,

* Co prawda pochodna nie musi być ciągła, ale ma własność przyjmowania wartości pośrednich, co zostało wykazane przez Darboux. Zresztą monotoniczność funkcji g wynika też z twierdzenia Rolle'a.

bo $k \geq a$, zatem $\lim\limits_{x\to+\infty} \dfrac{a(a-1)(a-2)\cdot\ldots\cdot(a-k+1)x^{a-k}}{e^x} = 0$. Dzięki regule de l'Hospitala możemy stwierdzić, że $\lim\limits_{x\to+\infty} \dfrac{a(a-1)(a-2)\cdot\ldots\cdot(a-k+2)x^{a-k+1}}{e^x} = 0$. Stosując twierdzenie jeszcze $k-1$ razy dochodzimy w końcu do granicy $\lim\limits_{x\to+\infty} \dfrac{x^a}{e^x} = 0$. Oczywiście wynik ten można otrzymać stosując jedynie elementarne metody: wykładnik a można zastąpić liczbą naturalną m większą od a, następnie skorzystać z nierówności $e^x > \left(1+\dfrac{x}{n}\right)^n$ prawdziwej dla każdej liczby naturalnej n i każdej liczby $x > 0$, następnie skorzystać z tego, że granicą ilorazu wielomianu stopnia m przez wielomian stopnia $n > m$ przy $x \to \infty$ jest liczba 0. Pokazaliśmy tu po prostu jak można wykorzystać twierdzenie de l'Hospitala, ta metoda pozwala na obliczanie granic w wielu sytuacjach, metody elementarne bywają trudne w zastosowaniach – trzeba mieć dobry pomysł! ■

36. $\lim\limits_{x\to+\infty} \dfrac{\ln x}{x^a} = 0$ dla każdego $a > 0$ – ten wynik już był omawiany w rozdziale pierwszym, ale pokażemy jak można go uzyskać za pomocą reguły de l'Hospitala. Ponieważ mianownik jest funkcja ściśle rosnącą o granicy $+\infty$, więc można spróbować obliczyć granicę ilorazu pochodnych: $\lim\limits_{x\to+\infty} \dfrac{1/x}{ax^{a-1}} = \lim\limits_{x\to+\infty} \dfrac{1}{ax^a} = 0$. Wobec tego istnieje również granica ilorazu funkcji i również jest równa 0. ■

37. $\lim\limits_{x\to0^+} x^x = 1$. Mamy bowiem: $x^x = e^{x\ln x}$. Funkcja wykładnicza jest ciągła, więc starczy wykazać, że $\lim\limits_{x\to0^+} x\ln x = 0$. Mamy $\lim\limits_{x\to0^+} x\ln x = \lim\limits_{x\to+\infty} \frac{1}{y}\ln\frac{1}{y} = -\lim\limits_{y\to+\infty} \frac{\ln y}{y} = 0$ – ostatnia równość wynika z rezultatu uzyskanego w poprzednim przykładzie dla $a = 1$, przedostatnia zaś — z tego, że $\ln\frac{1}{y} = -\ln y$. ■

38. $\lim\limits_{x\to0}(1+x)^{1/x} = e$ – to wzmocnienie wyniku $\lim\limits_{n\to\infty}\left(1+\dfrac{1}{n}\right)^n = e$. Dzięki oczywistej równości $(1+x)^{1/x} = e^{(1/x)\cdot\ln(1+x)}$ wiemy, że wystarczy wykazać równość $\lim\limits_{x\to0} \dfrac{\ln(1+x)}{x} = 1$. Ta równość została już wcześniej udowodniona, zresztą $\lim\limits_{x\to0} \dfrac{\ln(1+x)}{x} = \lim\limits_{x\to0} \dfrac{\ln(1+x)-\ln 1}{x}$, więc granica ta jest równa pochodnej logarytmu naturalnego w punkcie 1 (to wniosek z definicji pochodnej, reguła de l'Hospitala jest tu zbędna), czyli $\dfrac{1}{1} = 1$. ■

39. Pokażemy teraz, w znacznie prostszy sposób niż w rozdziale pierwszym, że ciąg $\left((1+\dfrac{1}{n})^n\right)$ jest wolno zbieżny do liczby e, więc nie należy go używać do jej przybliżonego obliczania. Obliczymy mianowicie granicę

$\lim\limits_{n\to\infty} \dfrac{e-\left(1+\dfrac{1}{n}\right)^n}{\dfrac{1}{n}}$. W tym celu wystarczy obliczyć granicę $\lim\limits_{x\to0} \dfrac{e-(1+x)^{1/x}}{x}$. Z istnienia tej ostatniej wynika oczywiście istnienie poprzedniej (definicja granicy wg. Heinego), odwrotne wynikanie nie zachodzi. Z rezultatu z poprzedniego przykładu wynika, że zarówno licznik jak i mianownik dążą do 0 przy $x \to 0$. Zbadamy więc iloraz pochodnych. Jest on równy pochodnej licznika, czyli

$$\left(e-(1+x)^{1/x}\right)' = \left(-e^{\ln(1+x)/x}\right)' = -e^{\ln(1+x)/x}\cdot\frac{\frac{1}{1+x}x-\ln(1+x)}{x^2} =$$
$$= -(1+x)^{1/x}\cdot\frac{x-(1+x)\ln(1+x)}{x^2(1+x)} = -\frac{(1+x)^{1/x}}{1+x}\cdot\frac{x-(1+x)\ln(1+x)}{x^2}.$$

Z tego co już wiemy wynika, że $\lim\limits_{x\to0} -\dfrac{(1+x)^{1/x}}{1+x} = -e$. Wystarczy więc obliczyć granicę drugiego

czynnika, czyli $\lim\limits_{x\to 0}\dfrac{x-(1+x)\ln(1+x)}{x^2}$. Jest jasne, że licznik i mianownik dążą do 0 przy $x\to 0$. Zajmiemy się więc ilorazem pochodnych. Jest on równy
$\dfrac{1-\left\{\ln(1+x)+(1+x)\cdot\dfrac{1}{1+x}\right\}}{2x}=-\dfrac{\ln(1+x)}{2x}=-\dfrac{1}{2}\cdot\dfrac{\ln(1+x)}{x}$, więc ma przy $x\to 0$ granicę $-\dfrac{1}{2}$. Wobec tego $\lim\limits_{x\to 0}\dfrac{e-(1+x)^{1/x}}{x}=(-e)\cdot\left(-\dfrac{1}{2}\right)=\dfrac{e}{2}$.

Z otrzymanej równości $\lim\limits_{n\to\infty}\dfrac{e-\left(1+\dfrac{1}{n}\right)^n}{\dfrac{1}{n}}=\dfrac{e}{2}$ wynika, że dla „dużych” n zachodzi równość przybliżona $e-\left(1+\dfrac{1}{n}\right)^n\approx\dfrac{e}{2n}$, więc dla uzyskania dobrej dokładności przybliżenia trzeba używać dużej liczby naturalnej n, co w zasadzie czyni przybliżenie bezużytecznym, więcej komentarzy znajdzie czytelnik w rozdziale pierwszym, punkt 19.n oraz w punkcie 13 rozdziału pierwszego.

Komentarz: w końcowej fazie obliczeń, przed zastosowaniem reguły de l'Hospitala, przedstawiliśmy ułamek w postaci iloczynu dwóch ułamków. Gdybyśmy tego nie uczynili obliczenia wyglądałyby o wiele poważniej. ■

40. W ostatnim przykładzie pokazaliśmy, że $e-\left(1+\frac{1}{n}\right)^n\approx\frac{e}{2n}$ dla dostatecznie dużych n. Podamy teraz konkretne oszacowanie. Warto porównać to rozumowanie z szacowaniami przeprowadzanymi w punkcie 19.n rozdziału pierwszego. Wykażemy, że zachodzi nierówność podwójna $\frac{e}{2n+2}<e-\left(1+\frac{1}{n}\right)^n<\frac{e}{2n+1}$ dla wszystkich liczb naturalnych $n\geq 1$. * Z nierówności tej wynika, że jeśli np. chcemy znaleźć trzy miejsca po przecinku dziesiętnego rozwinięcia liczby e stosując wzór $e=\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n$, to musimy wybrać n tak duże, by $\frac{e}{2n+1}<\frac{1}{1000}$, czyli $n>\frac{1000e-1}{2}>1359$. Z nierówności $e-\left(1+\frac{1}{n}\right)^n>\frac{e}{2n+2}$ wynika, że dla $n=1358$, błąd jest większy niż 0,001.

Zajmiemy się najpierw nierównością $\frac{e}{2n+2}<e-\left(1+\frac{1}{n}\right)^n$. Jest ona równoważna następującej nierówności $\left(1+\frac{1}{n}\right)^{n+1}<e\left(1+\frac{1}{2n}\right)$ – przenieśliśmy składniki zawierające liczbę e na prawą stronę, składniki bez e – na lewą, następnie pomnożyliśmy obie strony nierówności przez $1+\frac{1}{n}=\frac{n+1}{n}$. Teraz zastąpimy $\frac{1}{n}$ przez x. Mamy dowieść, że $(1+x)^{1+1/x}<e\left(1+\frac{x}{2}\right)$. Ponieważ logarytm naturalny jest funkcją ściśle rosnącą, więc nierówność ta równoważna jest następującej $\left(\frac{1}{x}+1\right)\ln(1+x)<1+\ln\left(1+\frac{x}{2}\right)$. Wystarczy rozpatrywać $x>0$. Po pomnożeniu obu stron przez $x>0$ i przeniesieniu wszystkich składników na jedną stronę otrzymujemy $x+x\ln\left(1+\frac{x}{2}\right)-(1+x)\ln(1+x)>0$. Oznaczywszy lewą stronę przez $f(x)$ stwierdzamy, że $f(0)=0$ oraz $f'(x)=1+\ln\left(1+\frac{x}{2}\right)+x\frac{\frac{1}{2}}{1+\frac{x}{2}}-\ln(1+x)-(1+x)\frac{1}{1+x}=\frac{x}{2+x}-\ln\frac{1+x}{1+\frac{x}{2}}=\frac{x}{2+x}-\ln\left(1+\frac{x}{2+x}\right)>0$ – ostatnia nierówność wynika z tego, że $\ln(1+y)<y$ dla $-1<y\neq 0$, co wynika np. z tego, że funkcja ln jest ściśle wklęsła na $(0,+\infty)$, a wykres funkcji ściśle wklęsłej leży pod styczną do tego wykresu. Wykazaliśmy, że $f'(x)>0$ dla $x>0$, więc funkcja f rośnie na półprostej $[0,+\infty)$, a ponieważ $f(0)=0$, więc dla $x>0$ przyjmuje jedynie wartości dodatnie. W szczególności $f\left(\frac{1}{n}\right)>0$, a to właśnie chcieliśmy udowodnić.

Teraz zajmiemy się nierównością $e-\left(1+\frac{1}{n}\right)^n<\frac{e}{2n+1}$. Jest ona równoważna nierówności $e<\left(1+\frac{1}{n}\right)^n\left(1+\frac{1}{2n}\right)$ — przenieśliśmy oba wyrazy zawierające liczbę e na lewą stronę nierówności, resztę — na prawą stronę, następnie podzieliliśmy nierówność przez $1-\frac{1}{2n+1}=\frac{2n}{2n+1}$. Te-

* zob. G.Pólya, G.Szegö, Aufgaben und Lehrsätze aus der Analysis, Springer 1964, wyd. 3, t.1.

raz oznaczymy $\frac{1}{n}$ przez x i zlogarytmujemy obie strony nierówności . Otrzymujemy nierówność $1 < \frac{1}{x}\ln(1+x) + \ln\left(1+\frac{x}{2}\right)$. Niech $g(x) = \frac{1}{x}\ln(1+x) + \ln\left(1+\frac{x}{2}\right) - 1$. Wykażemy, że g jest funkcją ściśle rosnącą. Mamy $g'(x) = -\frac{1}{x^2}\cdot\ln(1+x) + \frac{1}{x(1+x)} + \frac{1}{2}\cdot\frac{1}{1+x/2}$. Wykażemy, że dla $0 < x$ zachodzi $g'(x) > 0$. Niech $h(x) = x^2 g'(x) = \frac{x}{1+x} + \frac{x^2}{2+x} - \ln(1+x)$. Wystarczy wykazać, że $h(x) > 0$ dla $0 < x$. Mamy $h(0) = 0$ oraz $h'(x) = \frac{1}{(1+x)^2} + \frac{4x+x^2}{(2+x)^2} - \frac{1}{1+x} = \frac{x^2(x^2+5x+5)}{(1+x)^2(2+x)^2} > 0$ dla $x > 0$. Wobec tego funkcja h rośnie na półprostej $[0,+\infty)$, więc $h(x) > h(0) = 0$, co dowodzi tego, że $g'(x) > 0$. Z definicji funkcji g wynika natychmiast, że $\lim\limits_{x\to 0} g(x) = 0$, a wobec tego, że funkcja ta jest rosnąca na półprostej $(0,+\infty)$, zachodzi nierówność $g(x) > 0$ dla $x > 0$. W ten sposób udowodniliśmy drugą nierówność. ■

5. Szeregi potęgowe II

Wiele funkcji można przedstawiać w postaci sum szeregów potęgowych. Udało nam się już przedstawić w takiej postaci funkcję wykładniczą. W tym punkcie przekonamy się, że nie jest ona żadnym wyjątkiem – praktycznie wszystkie funkcje, które są zdefiniowane „za pomocą jednego wzoru", można tak zapisać, co ułatwia w licznych przypadkach poznanie ich własności. Z twierdzenia o różniczkowaniu szeregu potęgowego wynika, że wewnątrz dziedziny mają one skończoną pochodną i ta pochodna jest również sumą szeregu potęgowego, więc i ona ma skończoną pochodną wewnątrz swej dziedziny. Zaczniemy od logarytmu naturalnego. Udowodnimy mianowicie, że

41. $\ln(1+x) = \sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}\,x^n = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \ldots$ dla każdej liczby $x \in (-1,1]$. Jeśli $|x| < 1$, to $(\ln(1+x))' = \frac{1}{1+x} = 1 - x + x^2 - x^3 + \ldots = \left(x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \ldots\right)' = \left(\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}\,x^n\right)'$, zatem pochodna funkcji $\ln(1+x) - \sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}\,x^n$, określonej i ciągłej na przedziale $(-1,1]$ i różniczkowalnej w jego punktach wewnętrznych jest równa 0. Stąd wynika, że funkcja $\ln(1+x) - \sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}\,x^n$ jest stała na przedziale domknięto–otwartym $(-1,1]$. Wobec tego dla każdej liczby $x \in (-1,1]$ zachodzi równość

$$\ln(1+x) - \sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}\,x^n = \ln(1+0) - \sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}\,0^n = 0\,.$$

Przedstawiliśmy więc funkcję $\ln(1+x)$ w postaci sumy szeregu potęgowego o środku w punkcie 0. Z tego wzoru wynika, że jeśli $x_0 > 0$ i $0 < x \le 2x_0$, to $\ln x = \ln\left(x_0\left(1+\frac{x-x_0}{x_0}\right)\right) =$
$= \ln x_0 + \ln\left(1+\frac{x-x_0}{x_0}\right) = \ln x_0 + \sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}\left(\frac{x-x_0}{x_0}\right)^n$, a więc przedstawiliśmy funkcję $\ln x$ w postaci sumy szeregu potęgowego o środku w dowolnie wybranym punkcie x_0 i promieniu zbieżności x_0, więc maksymalnym o jakim można myśleć ($\ln 0$ nie jest zdefiniowany, więc jest naturalnym, że przedział zbieżności nie zawiera punktu 0).

Otrzymany wzór zastosujemy podstawiając $x = 1$ we wzorze $\ln(1+x) = \sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}\,x^n$. Rezultat to $\ln 2 = \ln(1+1) = \sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$. Znaleźliśmy więc sumę szeregu, którego zbieżność stwierdziliśmy już dawno. Przykład ten świadczy, że innym problemem jest wykazanie zbieżności szeregu, a innym znalezienie jego sumy. Dodajmy jeszcze, że szereg ten jest wolno zbieżny i nie warto znajdować przybliżeń dziesiętnych liczby $\ln 2$ za jego pomocą. Można natomiast np. zauważyć, że $\ln 2 = -\ln\frac{1}{2} = -\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}\left(-\frac{1}{2}\right)^n = \sum\limits_{n=1}^{\infty}\frac{1}{n}\left(\frac{1}{2}\right)^n$.

W tym przypadku błąd jaki popełniamy przybliżając liczbę $\ln 2$ sumą $\sum\limits_{n=1}^{k}\frac{1}{n}\left(\frac{1}{2}\right)^n$ jest równy $\sum\limits_{n=k+1}^{\infty}\frac{1}{n}\left(\frac{1}{2}\right)^n = \frac{1}{(k+1)2^{k+1}} + \frac{1}{(k+2)2^{k+2}} + \frac{1}{(k+3)2^{k+3}} + \cdots$, więc jest mniejszy niż suma

$$\frac{1}{(k+1)2^{k+1}} + \frac{1}{(k+1)2^{k+2}} + \frac{1}{(k+1)2^{k+3}} + \cdots = \frac{1}{(k+1)2^{k+1}}\frac{1}{1-\frac{1}{2}} = \frac{1}{(k+1)2^k}.$$

W przypadku szeregu anharmonicznego $1 - \dfrac{1}{2} + \dfrac{1}{3} - \dfrac{1}{4} + \cdots$ wartość bezwzględna błędu, który popełniamy zastępując liczbę $\ln 2$ sumą $\sum\limits_{n=1}^{k}(-1)^{n-1}\dfrac{1}{n}$ jest równa

$$\frac{1}{k+1} - \frac{1}{k+2} + \frac{1}{k+3} - \frac{1}{k+4} + \cdots = \frac{1}{(k+1)(k+2)} + \frac{1}{(k+3)(k+4)} + \cdots >$$
$$> \frac{1}{2}\left(\frac{1}{(k+1)(k+2)} + \frac{1}{(k+2)(k+3)} + \frac{1}{(k+3)(k+4)} + \frac{1}{(k+4)(k+5)} + \cdots\right) =$$
$$= \frac{1}{2}\left(\frac{1}{k+1} - \frac{1}{k+2} + \frac{1}{k+2} - \frac{1}{k+3} + \frac{1}{k+3} - \frac{1}{k+4} + \frac{1}{k+4} - \frac{1}{k+5} + \cdots\right) = \frac{1}{2(k+1)}.$$

Jeśli przyjmiemy $k = 4$, to w pierwszym przypadku błąd będzie *mniejszy* niż $\dfrac{1}{80}$, a w drugim – *większy* niż $\dfrac{1}{10}$. Dla $k = 9$ w pierwszym przypadku błąd jest *mniejszy* od $\dfrac{1}{5120}$, a w drugim – *większy* niż $\dfrac{1}{20}$. Jasne jest więc, że w razie konieczności przybliżenia liczby $\ln 2$ za pomocą ułamka dziesiętnego użyć należy szeregu $\sum\limits_{n=1}^{\infty}\dfrac{1}{n}\left(\dfrac{1}{2}\right)^n$, a nie szeregu anharmonicznego $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\dfrac{1}{n}$. ∎

42. Teraz zajmiemy się funkcją arctg. Mamy $(\operatorname{arctg} x)' = \dfrac{1}{1+x^2}$. Wobec tego dla $x \in (-1,1)$ zachodzi równość $(\operatorname{arctg})' = \dfrac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots = \left(x - \dfrac{x^3}{3} + \dfrac{x^5}{5} - \dfrac{x^7}{7} + \cdots\right)'$. Jasne jest, że szereg $x - \dfrac{x^3}{3} + \dfrac{x^5}{5} - \dfrac{x^7}{7} + \cdots = \sum\limits_{n=0}^{\infty}(-1)^n\dfrac{x^{2n+1}}{2n+1}$ jest zbieżny dla $x = \pm 1$, przy czym nie jest to zbieżność bezwzględna. Wobec tego jego przedziałem zbieżności jest $[-1,1]$ – szereg potęgowy jest *wewnątrz* przedziału zbieżności *bezwzględnie* zbieżny. Wykazaliśmy więc, że funkcja ciągła $\operatorname{arctg} x - \sum\limits_{n=0}^{\infty}(-1)^n\dfrac{x^{2n+1}}{2n+1}$ jest ciągła na przedziale $[-1,1]$ oraz że jej pochodna jest równa 0 w punktach wewnętrznych tego przedziału. Stąd wynika, że ta funkcja jest stała na przedziale $[-1,1]$. Wobec tego dla każdej liczby $x \in [-1,1]$ zachodzi równość:
$\operatorname{arctg} x - \sum\limits_{n=0}^{\infty}(-1)^n\dfrac{x^{2n+1}}{2n+1} = \operatorname{arctg} 0 - \sum\limits_{n=0}^{\infty}(-1)^n\dfrac{0^{2n+1}}{2n+1} = 0$. Stąd wynika, że dla każdej liczby $x \in [-1,1]$ zachodzi równość: $\operatorname{arctg} x = \sum\limits_{n=0}^{\infty}(-1)^n\dfrac{x^{2n+1}}{2n+1}$. Podobnie jak w poprzednim przykładzie tak i tu możemy uzyskać konkretne rezultaty. Np. podstawiając $x = 1$ do otrzymanego wzoru otrzymujemy $\dfrac{\pi}{4} = \operatorname{arctg} 1 = 1 - \dfrac{1}{3} + \dfrac{1}{5} - \dfrac{1}{7} + \cdots = \sum\limits_{n=0}^{\infty}(-1)^n\dfrac{1}{2n+1}$ – ta równość nazywana jest zwykle wzorem Leibniza. Można wykazać, że jeśli chcielibyśmy za pomocą tego wzoru znajdować przybliżenia dziesiętne liczby π, to musielibyśmy wykonać wiele obliczeń, co nawet w przypadku

komputerów ma istotne znaczenie – konkretnie: dla każdej liczby naturalnej n zachodzi nierówność podwójna $\frac{1}{4(n+1)} < \frac{1}{2n+1} - \frac{1}{2n+3} + \frac{1}{2n+5} - \frac{1}{2n+7} + \dots < \frac{1}{4n}$ (nie jest ona oczywista!), więc błąd popełniany przy zastępowaniu liczby $\frac{\pi}{4}$ liczbą $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots + (-1)^{n-1}\frac{1}{2n-1}$ jest zawarty między $\frac{1}{4(n+1)}$ oraz $\frac{1}{4n}$. Stosując wzór z lepiej dobranym x otrzymać można bez trudu szeregi „szybciej" zbieżne.* ■

43. Wykażemy teraz, że $\sin x = \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n+1}}{(2n+1)!}$. Wykazaliśmy poprzednio, że dla każdej liczby rzeczywistej $x > 0$ zachodzi nierówność podwójna $x - \frac{x^3}{3!} < \sin x < x$. Teraz wykażemy, że $\sin x < x - \frac{x^3}{3!} + \frac{x^5}{5!}$. Niech $f(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \sin x$. Mamy $f'(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cos x$ i wobec tego możemy napisać, że $(f')'(x) = -x + \frac{x^3}{3!} + \sin x$. Z przypomnianej nierówności wynika, że dla każdej liczby $x > 0$ zachodzi $(f')'(x) > 0$, a stąd wynika, że funkcja f' jest ściśle rosnąca na półprostej $[0, +\infty)$. Ponieważ $f'(0) = 0$, więc jeśli $x > 0$, to $f'(x) > f'(0) = 0$. Stąd wynika, że funkcja f jest ściśle rosnąca na półprostej $[0, +\infty)$ i wobec tego jeśli $x > 0$, to $f(x) > f(0) = 0$, czyli $x - \frac{x^3}{3!} + \frac{x^5}{5!} > \sin x$, a to właśnie chcieliśmy wykazać. Rozumując w taki sam sposób wnioskujemy, że dla $x > 0$ zachodzi nierówność $x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} < \sin x$ – obliczamy pochodną różnicy prawej i lewej strony tej nierówności, potem jej pochodną tej pochodnej, w wyniku otrzymujemy funkcję dodatnią na półprostej $(-\infty, 0)$ itd. Powtarzając to rozumowanie, czyli stosując indukcję zupełną, dochodzimy do wniosku, że dla każdej liczby rzeczywistej $x > 0$ i każdej całkowitej liczby nieujemnej n zachodzi nierówność podwójna:

$$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + (-1)^{2n-1}\frac{x^{4n-1}}{(4n-1)!} < \sin x < x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + (-1)^{2n}\frac{x^{4n+1}}{(4n+1)!}$$

Różnica skrajnych sum równa jest $\frac{x^{4n+1}}{(4n+1)!}$. Mamy też $\lim_{n\to\infty}\frac{x^{4n+1}}{(4n+1)!} = 0$ – można wywnioskować z kryterium ilorazowego d'Alemberta, że szereg $\sum_{n=0}^{\infty}\frac{x^{4n+1}}{(4n+1)!}$ jest zbieżny dla $x > 0$, więc jego wyraz ma granicę 0 albo zauważyć, że jeśli $n > x > 0$, to $\frac{x}{4n+1} < \frac{1}{4}$, a stąd wynika, że dla dostatecznie dużych n wyraz o numerze $n+1$ ciągu $\left(\frac{x^{4n+1}}{(4n+1)!}\right)$ jest mniejszy niż ćwierć wyrazu n–tego. Niech $s_{4n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + (-1)^{2n}\frac{x^{4n+1}}{(4n+1)!}$ i niech $s_{4n-1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + (-1)^{2n-1}\frac{x^{4n-1}}{(4n-1)!}$. Z tego że $s_{4n-1} < \sin x < s_{4n+1}$ i $\lim_{n\to\infty}(s_{4n+1} - s_{4n-1}) = 0$ wynika oczywiście, że $\lim_{n\to\infty} s_{4n-1} = \sin x = \lim_{n\to\infty} s_{4n+1}$. Oznacza to, że sumą szeregu $\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!}$ jest $\sin x$, czyli że $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots$. Na razie wiemy, że równość ta ma miejsce dla $x > 0$, ale w rzeczywistości, dzięki temu że prawa strona to szereg potęgowy o środku w punkcie 0, wiemy już, że promieniem zbieżności prawej strony jest

* Można o tym przeczytać np. w rachunku różniczkowym i całkowym G.M.Fichtenholza, t.2, rozdział XI, $ 8, punkt 410, książce wielokrotnie wznawianej przez PWN.

$+\infty$. Ponieważ obie strony lewa i prawa są funkcjami nieparzystymi zmiennej x równymi w przypadku $x>0$, więc są one równe dla każdej liczby rzeczywistej x. Wykazaliśmy więc, że dla każdej liczby rzeczywistej x zachodzi równość

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \tag{43}$$

„Po drodze" wykazaliśmy też, że skończona suma $\sum_{n=0}^{k} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$ przybliża sumę nieskończoną z błędem mniejszym niż $\frac{|x|^{2(k+1)+1}}{(2(k+1)+1)!}$. Wynika stąd np. że jeśli x jest miarą kąta ostrego, czyli $0 < x < \frac{\pi}{2} < \frac{3{,}16}{2} = 1{,}58$, to różnica między liczbami $x - \frac{x^3}{3!}$ i $\sin x$ jest mniejsza niż $\frac{x^5}{5!} < \frac{1{,}58^5}{5!} < \frac{10}{120} < 0{,}1$. Zwiększenie liczby składników o 1, tj. przybliżanie liczby $\sin x$ liczbą $x - \frac{x^3}{3!} + \frac{x^5}{5!}$ powoduje zmniejszenie błędu do wartości mniejszej niż $\frac{x^7}{7!} < \frac{1{,}58^7}{7!} < 0{,}005 = \frac{1}{200}$. Widzimy więc, że możemy w miarę dokładnie znajdować wartości liczbowe sinusów stosunkowo niewielkim kosztem, przy czym można dokładność istotnie zwiększyć zwiększając liczbę składników nieznacznie. Oczywiście tego typu oszacowania są przydatne nie tylko do rachowania, ale również w przypadku rozważań teoretycznych. Zwróćmy jeszcze uwagę na to, że jeśli interesujemy się liczbami x bliskimi 0, to błąd maleje, również błąd względny zmniejsza się. ■

44. Z wzoru $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$ wynika natychmiast, że

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

– po prostu obliczamy pochodne obu stron wzoru (43) , co wolno zrobić dzięki twierdzeniu o pochodnej szeregu potęgowego. ■

45. Zajmiemy się teraz dwumianem Newtona. Nie chodzi przy tym o wzór na obliczanie n-tej potęgi sumy dwu liczb, bo ten był z pewnością znany przed Newtonem, na pewno znał go Pascal, a prawdopodobnie (wg. N.Bourbaki, *Elementy historii matematyki*,PWN, Warszawa, 1980) był znany Arabom w wieku XIII, a Chińczykom w wieku XIV. Chodzi o wzór na $(1+x)^a$, gdzie wykładnik a nie musi być liczbą naturalną – może być dowolną liczbą rzeczywistą, liczba x zaś w przypadku dowolnego wykładniki nie może być dowolna, musi mieć wartość bezwzględną mniejszą niż 1.

Rozpoczniemy od zdefiniowania *symbolu Newtona.* Jeśli $a \in \mathbb{R}$ oraz $n \in \{1, 2, \ldots\}$, to przyjmujemy, że $\binom{a}{n} = \frac{a(a-1)(a-2)\ldots(a-n+1)}{n!}$. Dodatkowo $\binom{a}{0} = 1$ dla każdej liczby rzeczywistej a. Z definicji tej wynika natychmiast, że jeśli a jest liczbą naturalną nie mniejszą niż n, to $\binom{a}{n} = \frac{a!}{n!(a-n)!}$. Oznacza to, że nasza definicja jest po prostu rozszerzeniem definicji znanej ze szkoły. Przy okazji wypada stwierdzić, że jeżeli a jest liczbą naturalną *mniejszą* niż n, to $\binom{a}{n} = 0$, bowiem w tym przypadku w liczniku ułamka definiującego symbol Newtona występuje liczba $a - a = 0$.

Z definicji symbolu Newtona i tego, że $1 + \frac{a-n}{n+1} = \frac{a+1}{n+1}$ wynika od razu, że

$$\binom{a}{n} = \frac{a}{n} \cdot \binom{a-1}{n-1} \qquad \text{oraz} \qquad \binom{a}{n} + \binom{a}{n+1} = \binom{a+1}{n+1}. \tag{45}$$

Wykażemy teraz, że jeśli $|x| < 1$, to dla każdej liczby rzeczywistej a zachodzi wzór Newtona:

$$(1+x)^a = \binom{a}{0} + \binom{a}{1}x + \binom{a}{2}x^2 + \binom{a}{3}x^3 + \dots = \sum_{n=0}^{\infty} \binom{a}{n} x^n . \tag{45.N}$$

Rozpoczniemy od znalezienia promienia zbieżności szeregu potęgowego występującego po prawej stronie równości (45.N). Skorzystamy z kryterium ilorazowego d'Alemberta. Obliczamy granicę ilorazu zakładając, że a *nie* jest liczbą naturalną: $\lim\limits_{n\to\infty} \left| \frac{\binom{a}{n+1}x^{n+1}}{\binom{a}{n}x^n} \right| = \lim\limits_{n\to\infty} \left| \frac{(a-n)x}{n+1} \right| = |x|$. Stąd wnioskujemy, że w przypadku $|x| < 1$ szereg jest bezwzględnie zbieżny, natomiast w przypadku $|x| > 1$ szereg jest rozbieżny, bowiem jego wyraz nie dąży do 0 przy $n \longrightarrow \infty$. Obliczymy pochodną ilorazu $\sum\limits_{n=0}^{\infty} \binom{a}{n} x^n \Big/ (1+x)^a$. Mamy

$$\left(\sum_{n=0}^{\infty} \binom{a}{n} x^n \Big/ (1+x)^a \right)' = \frac{\sum\limits_{n=1}^{\infty} n \binom{a}{n} x^{n-1} \cdot (1+x)^a - a(1+x)^{a-1} \sum\limits_{n=0}^{\infty} \binom{a}{n} x^n}{(1+x)^{2a}} =$$

$$= a\,(1+x)^{-a-1} \left((1+x) \sum_{n=1}^{\infty} \binom{a-1}{n-1} x^{n-1} - \sum_{n=0}^{\infty} \binom{a}{n} x^n \right) =$$

$$= a(1+x)^{-a-1} \left(\sum_{n=1}^{\infty} \binom{a-1}{n-1} x^{n-1} + \sum_{n=1}^{\infty} \binom{a-1}{n-1} x^n - \sum_{n=0}^{\infty} \binom{a}{n} x^n \right) =$$

$$= a(1+x)^{-a-1} \left(\sum_{n=0}^{\infty} \binom{a-1}{n} x^n + \sum_{n=1}^{\infty} \binom{a-1}{n-1} x^n - \sum_{n=0}^{\infty} \binom{a}{n} x^n \right) =$$

$$= a(1+x)^{-a-1} \left(\sum_{n=1}^{\infty} \binom{a-1}{n} x^n + \sum_{n=1}^{\infty} \binom{a-1}{n-1} x^n - \sum_{n=1}^{\infty} \binom{a}{n} x^n \right) =$$

$$= a(1+x)^{-a-1} \sum_{n=1}^{\infty} \left(\binom{a-1}{n} + \binom{a-1}{n-1} - \binom{a}{n} \right) x^n = a(1+x)^{-a-1} \sum_{n=1}^{\infty} 0 \cdot x^n = 0 .$$

Wykazaliśmy, więc że pochodna ilorazu równa jest 0 w każdym punkcie przedziału $(-1,1)$. Stąd wynika, że na tym przedziale ten iloraz jest funkcją stałą, więc jego wartość w każdym punkcie x jest taka sama jak wartość w punkcie 0, a w tym punkcie wartość tego ilorazu jest równa $\sum\limits_{n=0}^{\infty} \binom{a}{n} 0^n \Big/ (1+0)^a = \binom{a}{0} \Big/ (1+0)^a = 1$. Wykazaliśmy więc, że dla każdego $x \in (-1,1)$ zachodzi równość $(1+x)^a = \sum\limits_{n=0}^{\infty} \binom{a}{n} x^n$, czyli zrealizowaliśmy nasz plan.

Zauważmy jeszcze, że dla $a = -1$ otrzymaliśmy wzór $\frac{1}{1+x} = \sum\limits_{n=0}^{\infty} \binom{-1}{n} x^n$. Czytelnik bez trudu stwierdzi, że $\binom{-1}{n} = \frac{(-1)(-2)\dots(-n)}{n!} = (-1)^n$, zatem otrzymaną równość możemy zapisać jako $\frac{1}{1+x} = \sum\limits_{n=0}^{\infty} (-1)^n x^n = \sum\limits_{n=0}^{\infty} (-x)^n$. Widzimy więc, że wzór (45.N) możemy potraktować nie tylko jako uogólnienie wzoru dwumianowego pozwalającego na zapisywanie w postaci sumy skończonej potęgi o wykładniku naturalnym sumy 2 składników, ale również jako uogólnienie wzoru na sumę nieskończonego ciągu geometrycznego (o ilorazie $-x$). ■

46. Wzór Newtona możemy zastosować do wyrażenia $\frac{1}{\sqrt{1-x^2}} = (1-x^2)^{-1/2}$. Otrzymujemy

$$\frac{1}{\sqrt{1-x^2}} = 1 + \binom{-1/2}{1} \cdot (-x^2) + \binom{-1/2}{2} \cdot (-x^2)^2 + \binom{-1/2}{3} \cdot (-x^2)^3 + \dots .$$

Mamy też

$$\binom{-1/2}{n}\cdot(-x^2)^n=\frac{(-1/2)\cdot(-3/2)\cdot(-5/2)\cdot\ldots\cdot\big(-(2n-1)/2\big)}{1\cdot 2\cdot 3\cdot\ldots\cdot n}\cdot(-x^2)^n=$$
$$=\frac{(1/2)\cdot(3/2)\cdot(5/2)\cdot\ldots\cdot\big((2n-1)/2\big)}{1\cdot 2\cdot 3\cdot\ldots\cdot n}\cdot(x^2)^n=\frac{1\cdot 3\cdot 5\cdot\ldots\cdot(2n-1)}{2\cdot 4\cdot 6\cdot\ldots\cdot(2n)}\cdot x^{2n}.$$

Stąd wynika, że $\binom{-1/2}{n}\cdot(-x^2)^n=\left(\frac{1}{2n+1}\cdot\frac{1\cdot 3\cdot 5\cdot\ldots\cdot(2n-1)}{2\cdot 4\cdot 6\cdot\ldots\cdot(2n)}\cdot x^{2n+1}\right)'$. Konsekwencją tej równości, twierdzenia o różniczkowaniu szeregu potęgowego, równości $\arcsin 0=0$ oraz wzoru $(\arcsin x)'=\frac{1}{\sqrt{1-x^2}}$ jest równość

$$\arcsin x=x+\frac{1}{3}\cdot\frac{1}{2}\cdot x^3+\frac{1}{5}\cdot\frac{1\cdot 3}{2\cdot 4}\cdot x^5+\cdots=\sum_{n=0}^{\infty}\frac{1}{2n+1}\cdot\frac{1\cdot 3\cdot 5\cdot\ldots\cdot(2n-1)}{2\cdot 4\cdot 6\cdot\ldots\cdot 2n}\cdot x^{2n+1},$$

która zachodzi dla $x\in(-1,1)$, bo dla tych x prawa strona jest szeregiem zbieżnym – wynika to łatwo z kryterium ilorazowego d'Alemberta, iloraz dwóch kolejnych wyrazów ma granicę x^2. Trochę trudniejszy dowód zbieżności tego szeregu dla $x=\pm 1$ opuszczamy. Ponieważ dla $x\notin[-1,1]$ granica ilorazu dwóch kolejnych wyrazów rozpatrywanego szeregu jest większa niż 1, więc w tym przypadku szereg jest rozbieżny. Stąd wynika, że równość w rzeczywistości zachodzi dla wszystkich $x\in[-1,1]$ i żadnych innych. Możemy więc zastosować ją w przypadku $x=\frac{1}{2}$. Mamy więc $\frac{\pi}{6}=\arcsin\frac{1}{2}=\frac{1}{2}+\frac{1}{3}\cdot\frac{1}{2}\cdot\left(\frac{1}{2}\right)^3+\frac{1}{5}\cdot\frac{1\cdot 3}{2\cdot 4}\cdot\left(\frac{1}{2}\right)^5+\cdots$. Jest oczywiste, że przybliżając liczbę $\frac{\pi}{6}$ liczbą $\frac{1}{2}+\frac{1}{3}\cdot\frac{1}{2}\cdot\left(\frac{1}{2}\right)^3+\frac{1}{5}\cdot\frac{1\cdot 3}{2\cdot 4}\cdot\left(\frac{1}{2}\right)^5$ popełniamy błąd mniejszy niż suma szeregu geometrycznego, którego pierwszym wyrazem jest $\frac{1}{7}\cdot\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\left(\frac{1}{2}\right)^7$, a ilorazem – liczba $\frac{1}{4}$, czyli $\frac{1}{7}\cdot\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\left(\frac{1}{2}\right)^7\cdot\frac{4}{3}=\frac{5}{10752}<0{,}0005$. Ponieważ mamy do czynienia z szeregiem zbieżnym „szybciej" niż szereg geometryczny o ilorazie $\frac{1}{4}$, więc wydłużenie sumy o jeden składnik spowoduje przynajmniej czterokrotne zmniejszenie się błędu jaki popełniamy zastępując sumę nieskończonego szeregu jego sumą częściową. Nie jest to rezultat rewelacyjny, ale jednak znacznie lepszy niż w przypadku szeregu Leibniza $\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots$. ■

POCHODNE WYŻSZYCH RZĘDÓW

1. Podstawowe definicje i twierdzenia

Z pochodnymi wyższych rzędów w istocie rzeczy już spotkaliśmy się. Po prostu w kilku przypadkach obliczaliśmy pochodną pochodnej. To oczywiście zdarza się często, gdy trzeba ustalić jakie własności ma funkcja. Przyjmuje się następujące określenie.

Definicja pochodnej wyższego rzędu

Niech f będzie funkcją określoną na zbiorze zawierającym przedział otwarty I zawierający punkt p. Niech $f^{(0)}(x) = f(x)$ dla każdego x z dziedziny funkcji f. Załóżmy, że funkcja f ma pochodną $(n-1)$–ego rzędu $f^{(n-1)}$ w każdym punkcie przedziału I. Jeśli funkcja $f^{(n-1)}$ ma w punkcie p pochodną $\left(f^{(n-1)}\right)'(p)$, to tę pochodną nazywamy *pochodną n–tego rzędu funkcji f w punkcie p* i oznaczamy symbolem $f^{(n)}(p)$. Jeśli pochodna n–tego rzędu jest *skończona*, to mówimy, że funkcja f jest *n–krotnie różniczkowalna* w tym punkcie. ■

Jest jasne, że $f' = f^{(1)}$. Zamiast pisać $f^{(2)}$ piszemy na ogół f''. Niektórzy matematycy zamiast $f^{(3)}$ piszą f'''.

Przykłady

1. Niech $f(x) = ax+b$. Wtedy dla każdego x mamy $f'(x) = a$, więc $f''(x) = 0$ dla każdej liczby rzeczywistej x. Wobec tego również $f^{(3)}(x) = 0$, a stąd wynika, że również $f^{(n)}(x) = 0$ dla każdej liczby naturalnej $n > 1$ i każdej liczby rzeczywistej x. ■

2. Niech $f(x) = ax^2+bx+c$. Wtedy $f'(x) = 2ax+b$, wobec tego $f''(x) = 2a$ i wobec tego dla każdej liczby naturalnej $n > 2$ i każdej liczby rzeczywistej x zachodzi równość $f^{(n)}(x) = 0$. ■

3. Niech f będzie wielomianem stopnia m, tzn. istnieją liczby rzeczywiste a_0, $a_1, \dots, a_m$, przy czym $a_m \neq 0$, takie że dla każdego $x \in \mathbb{R}$ zachodzi równość $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m$. Wtedy $f^{(m)}(x) = m!a_m$ dla każdego $x \in \mathbb{R}$ oraz $f^{(n)}(x) = 0$ dla każdej liczby naturalnej $n > m$ i każdej liczby rzeczywistej x.

Twierdzenie to wykazaliśmy już w przypadku $m = 1, 2$. Załóżmy, że twierdzenie jest prawdziwe dla wszystkich wielomianów stopnia *mniejszego* niż m. Wynika stąd, że dla wszystkich liczb rzeczywistych x zachodzi równość $f'(x) = a_1+2a_2+\cdots+ma_mx^{m-1}x$. Ponieważ f' jest wielomianem stopnia $m-1$, więc $(f')^{(m-1)}(x) = (m-1)! \cdot ma_m$ dla każdej liczby rzeczywistej x. Ponieważ $(f')^{(m-1)} = f^{(m)}$ oraz $(m-1)! \cdot m = m!$, więc dla każdej liczby rzeczywistej x mamy $f^{(m)}(x) = m!a_m$. Stąd oczywiście wynika, że jeśli $n > m$ jest liczbą naturalną, to $f^{(n)}(x) = 0$ dla każdego $x \in \mathbb{R}$. ■

4. Niech $f(x) = e^x$. Wtedy $f^{(1)}(x) = f'(x) = e^x$. Wobec tego dla każdej liczby naturalnej n i każdej rzeczywistej x zachodzi równość $f^{(n)}(x) = e^x$.

5. Niech $f(x) = \sin x$. Wtedy $f^{(1)}(x) = f'(x) = \cos x$. Zatem $f^{(2)}(x) = f''(x) = -\sin x = -f(x)$. Stąd wnioskujemy z łatwością, że $f^{(3)}(x) = -f'(x) = -\cos x$ i $f^{(4)}(x) = -f''(x) = \sin x$. Jasne jest, że od tego momentu będą się kolejno pojawiać, $\cos x$, $-\sin x$, $-\cos x$ i znów $\sin x$ itd. Można więc napisać $f^{(2n)}(x) = (-1)^n \sin x$ oraz $f^{(2n+1)}(x) = (-1)^n \cos x$ dla dowolnego $n \in \{0, 1, 2, \dots\}$ i $x \in \mathbb{R}$. ■

6. Podobnie jak w poprzednim przykładzie możemy wykazać, że $(\cos x)^{(2n)} = (-1)^n \cos x$ oraz $(\cos x)^{(2n+1)} = (-1)^{n+1} \sin x$. ■

7. Niech $f(x) = \ln x$. Mamy więc następującą równość $f^{(1)}(x) = f'(x) = \dfrac{1}{x} = x^{-1}$. Wobec tego $f^{(2)}(x) = f''(x) = \left(x^{-1}\right)' = (-1)x^{-1-1} = -x^{-2}$. Dalej $f^{(3)}(x) = \left(-x^{-2}\right) = 2x^{-3}$. Stąd wnioskujemy, że $f^{(4)}(x) = 2(-3)x^{-4} = -3!x^{-4}$. Analogicznie $f^{(5)}(x) = 4!x^{-5}$ itd. Ogólnie możemy napisać $f^{(n)}(x) = (\ln(x))^{(n)} = (-1)^{n-1}(n-1)!x^{-n}$ dla każdej liczby całkowitej $n \geq 1$ i każdej liczby rzeczywistej x. ■.

8. Obliczymy kilka pochodnych funkcji tangens. Mamy $(\operatorname{tg} x)' = 1 + \operatorname{tg}^2 x$. Wobec tego zachodzi równość $(\operatorname{tg} x)'' = \left(1 + \operatorname{tg}^2 x\right)' = 2\operatorname{tg} x(1 + \operatorname{tg}^2 x) = 2(\operatorname{tg} x + \operatorname{tg}^3 x)$ – skorzystaliśmy z wzoru na pochodną funkcji złożonej. Stąd $(\operatorname{tg} x)^{(3)} = 2(1 + 3\operatorname{tg}^2 x)(1 + \operatorname{tg}^2 x) = 2(1 + 4\operatorname{tg}^2 x + 3\operatorname{tg}^4 x)$, a stąd $(\operatorname{tg} x)^{(4)} = 2(8\operatorname{tg} x + 12\operatorname{tg}^3 x)(1 + \operatorname{tg}^2 x) = 8(2\operatorname{tg} x + 5\operatorname{tg}^3 x + 3\operatorname{tg}^5 x)$. Te obliczenia można kontynuować, jednak w tym przypadku nie da się napisać równie prosto jak w poprzednich przypadkach ogólnego wzoru na n–tą pochodna funkcji. ■

9. Znajdziemy teraz wzór na n–tą pochodną funkcji $\dfrac{x}{x^2+5x+6} = \dfrac{3}{x+3} - \dfrac{2}{x+2}$. W tym celu starczy znaleźć n–tą pochodną funkcji postaci $\dfrac{1}{x+c}$. Mamy $\left(\dfrac{1}{x+c}\right)' = -(x+c)^{-2}$. Stąd $\left(\dfrac{1}{x+c}\right)'' = -(-2)(x+c)^{-2-1} = 2(x+c)^{-3}$. Rozumując dalej w ten sam sposób otrzymujemy $\left(\dfrac{1}{x+c}\right)^{(3)} = -6(x+c)^{-4} = -6!(x+c)^{-4}$. Bez żadnych trudności piszemy wzór ogólny na n–tą pochodną tej funkcji: $\left(\dfrac{1}{x+c}\right)^{(n)} = (-1)^n n!(x+c)^{-n-1}$. Stąd wynika już od razu, że $\left(\dfrac{x}{x^2+5x+6}\right)^{(n)} = (-1)^n n!\left(3(x+3)^{-n-1} - 2(x+2)^{-n-1}\right)$. Wypada jednak zaznaczyć, że bez rozłożenia na czynniki mianownika nasze szanse na sukces byłyby znikome. ■

10. Wykazaliśmy poprzednio, że jeśli funkcja jest różniczkowalna na pewnym *przedziale* i jej pochodna jest na tym przedziale równa 0, to funkcja ta jest stała. Załóżmy teraz, że $f''(x) = 0$ dla wszystkich $x \in (a,b)$ dla pewnych $a, b \in \overline{\mathbb{R}}$. Wtedy na mocy poprzednio wykazanego stwierdzenia funkcja f' jest stała na przedziale (a,b). Niech $f'(x) = A$ dla wszystkich $x \in (a,b)$. Niech $g(x) = f(x) - Ax$. Zachodzi oczywista równość $g'(x) = 0$ dla każdej liczby $x \in (a,b)$. Wobec tego g jest funkcją stałą. Oznaczając jej jedyną wartość przez B otrzymujemy równość $B = g(x) = f(x) - Ax$. Stąd od razu wynika, że $f(x) = Ax + B$ dla każdej liczby $x \in (a,b)$. Wykazaliśmy więc, że jeśli druga pochodna jest tożsamościowo równa 0, to funkcja jest wielomianem stopnia nie większego niż 1. Podobnie można wykazać, że jeśli trzecia pochodna jest tożsamościowo równa 0 na pewnym przedziale, to funkcja jest na tym przedziale wielomianem stopnia nie większego niż 2. Jeśli bowiem $f^{(3)}(x) = 0$ dla każdego $x \in (a,b)$, to na mocy poprzedniego stwierdzenia funkcja f' jest wielomianem postaci $Ax + B$. Bez trudu zgadujemy, że $\left(\dfrac{1}{2}Ax^2 + Bx\right)' = Ax + B$. Stąd wynika, że $\left(f(x) - \dfrac{1}{2}Ax^2 - Bx\right)' = 0$ dla wszystkich $x \in (a,b)$. Teraz możemy stwierdzić, że funkcja $f(x) - \dfrac{1}{2}Ax^2 - Bx$ jest stała, co kończy dowód tego, że f jest wielomianem, którego stopień jest mniejszy niż 3. Jest całkowicie jasne, że kontynuując to rozumowanie wykażemy, że jeśli n–ta pochodna pewnej funkcji jest równa 0 w każdym punkcie pewnego przedziału, to funkcja ta na tym przedziale jest wielomianem, którego stopień jest mniejszy niż n. ■

11. Załóżmy, że f jest funkcją różniczkowalną na pewnym przedziale oraz że dla pewnej liczby rzeczywistej k równość $f'(x) = kf(x)$ zachodzi dla wszystkich x. Wykażemy, że w tej sytuacji istnieje stała $C \in \mathbb{R}$, taka że dla każdej liczby rzeczywistej x zachodzi równość $f(x) = Ce^{kx}$. W celu uzyskania tej równości starczy wykazać, że iloraz $\frac{f(x)}{e^{kx}}$ jest funkcja stałą, czyli że pochodna tego ilorazu jest wszędzie równa 0. Mamy $\left(\frac{f(x)}{e^{kx}}\right)' = \frac{f'(x)e^{kx} - ke^{kx}f(x)}{e^{2kx}} = \frac{f'(x) - kf(x)}{e^{kx}} = 0$ – ostatnia równość wynika z założenia o funkcji f. Wykazaliśmy więc, że iloraz jest funkcją stałą. Tę stałą oznaczamy przez C. Jasne jest, że $f(x) = Ce^{kx}$.

Rozważymy teraz nieco bardziej skomplikowaną zależność. Mianowicie założymy, f jest funkcją dwukrotnie różniczkowalną w każdym punkcie prostej * oraz że dla każdej liczby rzeczywistej x zachodzi równość $f''(x) = f(x)$. Bez trudu można podać dwa przykłady funkcji spełniających to równanie: $g(x) = e^x$ oraz $h(x) = e^{-x}$. Mając dwa, można ich podać o nieskończenie wiele. Jeśli c, d są dowolnymi liczbami rzeczywistymi, to funkcja $cg(x) + dh(x) = ce^x + de^{-x}$ również spełnia to równanie. Jasne jest, że również funkcja $u(x) = f(x) - cg(x) - dh(x)$ spełnia to równanie. Liczby c i d można dobrać w ten sposób, że $u(0) = 0 = u'(0)$ – wystarczy rozwiązać układ równań: $f(0) = c + d$, $f'(0) = c - d$ traktując c i d jako niewiadome, a $f(0)$ i $f'(0)$ jako dane liczby. Otrzymujemy $c = \frac{1}{2}(f(0) + f'(0))$ oraz $d = \frac{1}{2}(f(0) - f'(0))$. Poszukujemy więc dwukrotnie różniczkowalnej funkcji u, takiej że dla każdego x zachodzi równość $u''(x) = u(x)$ oraz $u'(0) = 0 = u(0)$. Wykażemy, że u jest funkcją zerową. Zauważmy najpierw, że $u''u' = uu'$, i wobec tego $\left(\frac{1}{2}(u')^2\right)' = \left(\frac{1}{2}u^2\right)'$. Stąd wynika, że funkcja $(u')^2 - u^2$ ma zerową pochodną, więc jest stała. Ponieważ $(u'(0))^2 - u(0)^2 = 0$, więc funkcja $(u')^2 - u^2$ jest zerowa, czyli $u'(x)^2 = u(x)^2$ dla każdej liczby rzeczywistej x. Załóżmy, że funkcja u przyjmuje w pewnym punkcie p wartość różną od 0. Są dwie możliwości: $u'(p) = u(p) \neq 0$ lub $u'(p) = -u(p) \neq 0$. Ponieważ obie funkcje u i u' są ciągłe, więc w pierwszym przypadku równość $u'(x) = u(x)$ zachodzi dla wszystkich x dostatecznie bliskich p, zaś w drugim przypadku dla wszystkich x dostatecznie bliskich p zachodzi równość $u'(x) = -u(x)$. *Dostatecznie bliskich* oznacza w tym przypadku dla wszystkich x z dowolnego przedziału przedziału I zwierającego punkt p, na którym funkcja u nie ma pierwiastków. Z pierwszej równości wynika, że istnieje stała C, taka że $u(x) = Ce^x$ dla wszystkich x z przedziału I. Z drugiej równości wynika istnienie stałej C, takiej że dla wszystkich x z przedziału I zachodzi równość $u(x) = Ce^{-x}$. Można założyć, że I jest maksymalnym przedziałem, który zawiera punkt p i w którym funkcja u nie ma pierwiastków. Oczywiście 0 nie leży w przedziale I. Wobec tego między p i 0 leży koniec q przedziału I, drugi koniec przedziału I znajduje się po przeciwnej stronie punktu p i nie jest wykluczone, że jest nieskończonością. Jest jasne, że $u(q) = 0$ – gdyby tak nie było, to przedział I sięgałby poza q. Ponieważ funkcja u jest ciągła i na przedziale I obowiązuje wzór $u(x) = Ce^x$ lub wzór $u(x) = Ce^{-x}$, więc w punkcie q mamy $u(q) = Ce^{\pm x}$. Jednocześnie $u(q) = 0$. Z dwóch ostatnich stwierdzeń wynika, że $C = 0$, a to oznacza, że wbrew uczynionemu założeniu $u(p) = Ce^{\pm p} = 0$. Wykazaliśmy więc, że u jest funkcją zerową, a to oznacza, że funkcja f jest postaci $ce^x + de^{-x}$. ■

12. Wykazaliśmy poprzednio , że równości $f^{(n)}(x) = 0$, $f'(x) = kf(x)$, $f''(x) = f(x)$ spełnione w każdym punkcie przedziału wymuszają, by funkcja f wyrażała się prostym wzorem. Omówimy jeszcze jeden przykład tego typu. Załóżmy mianowicie, że dla wszystkich punktów pewnego prze-

* Nie jest istotne, że dziedziną jest prosta, może być dowolny przedział.

działu I spełniona jest zależność $f''(x) = -f(x)$.** Wykażemy, że w tej sytuacji istnieją liczby $a, b \in \mathbb{R}$, takie że dla każdej liczby $x \in I$ zachodzi równość $f(x) = a\cos x + b\sin x$. Niech p oznacza dowolny punkt przedziału I. Jasne jest, że w każdym punkcie przedziału I zachodzi równość $(a\cos x + b\sin x)'' = -(a\cos x + b\sin x)$, tzn. funkcja postaci $a\cos x + b\sin x$ spełnia rozpatrywane równanie. Wybierzemy liczby a i b tak, by miały miejsce równości $f(p) = a\cos p + b\sin p$ oraz $f'(p) = -a\sin p + b\cos p$, tzn. $a = f(p)\cos p - f'(p)\sin p$ oraz $b = f(p)\sin p + f'(p)\cos p$. Niech $u(x) = f(x) - a\cos x - b\sin x$. Jest jasne, że $u''(x) = -u(x)$ dla każdej liczby $x \in I$ oraz że $u(p) = 0 = u'(p)$. Stąd wynika, że $\big((u'(x))^2 + (u(x))^2\big)' = 2\,(u''(x)u'(x) + u'(x)u(x)) = 0$, więc funkcja $(u'(x))^2 + (u(x))^2$ jest stała na przedziale I, zatem $(u'(x))^2 + (u(x))^2 = (u'(p))^2 + (u(p))^2 = 0$ dla każdego $x \in \mathbb{R}$. Suma *kwadratów* liczb rzeczywistych jest równa 0 wtedy i tylko wtedy, gdy obie te liczby są zerami. Wobec tego dla każdego $x \in \mathbb{R}$ zachodzi równość $u(x) = 0$, a zatem $f(x) = a\cos x + b\sin x$ dla każdego $x \in \mathbb{R}$. Okazało się, że również w tym przypadku można łatwo opisać wszystkie funkcje spełniające równanie $f'' = -f$. Tego typu równania nazywane są równaniami różniczkowymi zwyczajnymi. Istnieje obszerna teoria równań różniczkowych zwyczajnych. Nie mamy tu możliwości omawiania jej. Jest ona stosowana również w wielu dziedzinach poza matematyką, przede wszystkim w fizyce i w technice. Również w ekonomii. Równaniom różniczkowym poświęcony jest oddzielny wykład na drugim roku studiów. ■

Teraz zauważmy, że liczenie pochodnych wyższego rzędu polega na obliczaniu pochodnych rzędu pierwszego, więc właściwie już się z tym zapoznaliśmy. Jeśli chodzi o wzory ogólne, to oczywistym – i w zasadzie nie wartym wspomnienia – jest wzór na n–tą pochodną sumy dwu funkcji różniczkowalnych n–krotnie: $(f+g)^{(n)} = f^{(n)}+g^{(n)}$. Leibniz zauważył, że jeśli funkcje f i g są n–krotnie różniczkowalne, to zachodzi wzór bardzo podobny do wzoru dwumianowego Newtona:

$$(f \cdot g)^{(n)} = \sum_{j=0}^{n} \binom{n}{j} f^{(n-j)} g^{(j)} \qquad \text{(Leibniz)}$$

Prosty dowód tego wzoru wykorzystujący wzór na pochodną iloczynu dwu funkcji i znaną równość $\binom{n}{j} + \binom{n}{j+1} = \binom{n+1}{j+1}$, dzięki której współczynniki dwumianowe można obliczać za pomocą trójkąta Pascala, pozostawiamy czytelnikom w charakterze bardzo prostego ćwiczenia. Wzory na n–tą pochodną złożenia i funkcji odwrotnej są na tyle skomplikowane, że właściwie w ogóle nieprzydatne, zresztą trudno je znaleźć w literaturze.

Przejdziemy teraz do sformułowania jednego z najważniejszych wzorów analizy matematycznej, tzw. wzoru Taylora. Pierwszą pochodną funkcji wprowadziliśmy po to, by móc przybliżyć funkcję w pobliżu interesującego nas punktu wielomianem stopnia pierwszego. Drugie pochodne i pochodne wyższych rzędów pojawiły się w kilku miejscach w związku z bardziej szczegółowym badaniem funkcji . Okazuje się, że definicję pochodnej, związaną z przybliżaniem funkcji wielomianem stopnia pierwszego lub zerowego, można uogólnić. Tym zajmiemy się teraz. Efektem będzie zapowiadany wzór Taylora.

Poprzednio błąd przybliżenia miał być mały w porównaniu z pierwszą potęgą zmiany argumentu. Teraz zażądamy, by był mały w porównaniu z wyższymi potęgami h. Niestety nie będzie to możliwe

** Taka zależność, a dokładniej $f''=-\frac{g}{l}f$ pojawia się przy analizowaniu ruchu wahadła matematycznego o długości l przy założeniu, że amplituda jest tak mała, że przybliżenie $f \approx \sin f$ jest dostatecznie dokładne, g to przyspieszenie ziemskie.

przy użyciu wielomianów stopnia nie przekraczającego 1 – będziemy zmuszeni do użycia wielomianów stopnia wyższego.

Załóżmy, że $0 < |h| < 1$. Wobec tego $|h| > h^2 > |h|^3 > h^4 > \ldots$. Jasne jest też, że jeśli h jest bardzo blisko 0, to h^2 jest znacznie bliżej zera niż h, h^3 znacznie bliżej niż h^2 itd. Jest tak, bo $\lim\limits_{h\to 0}\dfrac{h^2}{h} = 0$ i ogólnie, jeśli $m > n$, to $\lim\limits_{h\to 0}\dfrac{h^m}{h^n} = 0$. Można myśleć o tym tak: jeżeli h jest bardzo małe i $m > n$, to liczba h^m stanowi znikomą część liczby h^n, oczywiście obie są wtedy bardzo małe, ale jedna jest istotnie mniejsza niż druga.

Wobec tego, z naszego punktu widzenia, różnica między dwiema funkcjami f i g będzie mała, jeśli będzie jeśli będzie dążyć do 0 *po podzieleniu przez* h^n, gdzie oznacza liczbę naturalną. Następujący lemat podaje warunek konieczny i dostateczny na to, by dwie funkcje były w tym sensie jedna drugiej.

Lemat o funkcjach ściśle przylegających

Jeśli funkcje f i g są n–krotnie różniczkowalne w punkcie 0, to $\lim\limits_{x\to 0}\dfrac{f(x)-g(x)}{x^n} = 0$ wtedy i tylko wtedy, gdy pochodne funkcji f i g w punkcie 0 są równe do n–tego rzędu włącznie: $f^{(j)}(0) = g^{(j)}(0)$ dla $j \in \{0, 1, \ldots, n\}$.

Dowód. Załóżmy, że $\lim\limits_{x\to 0}\dfrac{f(x)-g(x)}{x^n} = 0$. Niech $r(x) = f(x) - g(x)$. Trzeba udowodnić, że $r(0) = r'(0) = r''(0) = \ldots = r^{(n)}(0) = 0$. Załóżmy najpierw, że $0 \le j \le n$. Mamy wtedy $\lim\limits_{x\to 0}\dfrac{r(x)}{x^j} = \lim\limits_{x\to 0}\dfrac{r(x)}{x^n}\cdot\lim\limits_{x\to 0} x^{n-j} = 0$, bo pierwsza granica jest równa 0, a druga 0 lub 1 w zależności od tego, czy $j < n$ czy też $j = n$. Mamy $\lim\limits_{x\to 0} r(x) = 0$. Stąd i z tego, że funkcja r jest ciągła w punkcie 0, jako różniczkowalna, wynika, że $r(0) = 0$. Mamy $0 = \lim\limits_{x\to 0}\dfrac{r(x)}{x} = \lim\limits_{x\to 0}\dfrac{r(x)-r(0)}{x} = r'(0)$. Wobec tego $r'(0) = 0$. Teraz wykażemy, że $r''(0) = 0$ (zakładamy oczywiście, że$n \ge 2$). Stosujemy teraz regułę de l'Hospitala: $0 = \lim\limits_{x\to 0}\dfrac{r(x)}{x^2} = \lim\limits_{x\to 0}\dfrac{r'(x)}{2x} = \dfrac{1}{2}\lim\limits_{x\to 0}\dfrac{r'(x)-r'(0)}{x} = \dfrac{1}{2}r''(0)$. Wykażemy teraz w taki sam sposób, że również trzecia pochodna równa jest 0:

$$0 = \lim_{x\to 0}\frac{r(x)}{x^3} = \lim_{x\to 0}\frac{r'(x)}{3x^2} = \lim_{x\to 0}\frac{r''(x)}{6x} = \frac{1}{6}\lim_{x\to 0}\frac{r''(x)-r''(0)}{x} = \frac{1}{6}r^{(3)}(0).$$

Jasne jest, że tę procedurę można kontynuować.

Wykażemy teraz, że jeśli $r(0) = r'(0) = r''(0) = \ldots = r^{(n)}(0) = 0$, to $\lim\limits_{x\to 0}\dfrac{r(x)}{x^n} = 0$. Stosujemy regułę de l'Hospitala: $\lim\limits_{x\to 0}\dfrac{r(x)}{x^n} = \lim\limits_{x\to 0}\dfrac{r'(x)}{nx^{n-1}} = \lim\limits_{x\to 0}\dfrac{r''(x)}{n(n-1)x^{n-2}} = \ldots = \lim\limits_{x\to 0}\dfrac{r^{(n-1)}(x)}{n(n-1)\ldots 2x}$. Mamy dalej $\lim\limits_{x\to 0}\dfrac{r^{(n-1)}(x)}{x} = \lim\limits_{x\to 0}\dfrac{r^{(n-1)}(x)-r^{(n-1)}(0)}{x} = r^{(n)}(0) = 0$. Dowód lematu został zakończony. ■

Wniosek z dowodu.

Jeśli funkcja r jest n–krotnie różniczkowalna w punkcie 0 i $r(0) = r'(0) = r''(0) = r^{(n-1)}(0) = 0$, to $\lim\limits_{x\to 0}\dfrac{r(x)}{x^n} = \dfrac{r^{(n)}(0)}{n!}$. ■

Z lematu o funkcjach ściśle przylegających wynika, że jeśli chcemy przybliżyć funkcję w otoczeniu punktu p wielomianem w tak, by błąd przybliżenia był mały w porównaniu z h^n, to pochodne, do n–tego rzędu włącznie, tego wielomianu w punkcie 0 muszą być równe odpowiednim pochodnym funkcji f w punkcie p: $f^{(j)}(p) = w^{(j)}(0)$. Jeżeli $w(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ dla każdego $x \in \mathbb{R}$, to $w^{(j)}(0) = j!a_j$ dla $j = 0, 1, 2, \ldots, n$. Stąd wynika, że powinno być $a_j = \dfrac{f^{(j)}(p)}{j!}$. To motywuje wprowadzenie następującego określenia.

Definicja wielomianu Taylora i reszty

Załóżmy, że funkcja f ma w punkcie p pochodną n–tego rzędu. n–tym wielomianem Taylora funkcji f w punkcie p nazywamy wielomian $f(p)+f'(p)h+\frac{f''(p)}{2!}h^2+\frac{f^{(3)}(p)}{3!}h^3+\cdots+\frac{f^{(n)}(p)}{n!}h^n$ zmiennej h. n–tą resztą nazywamy różnicę

$$r_n(h)=f(p+h)-\left(f(p)+f'(p)h+\frac{f''(p)}{2!}h^2+\frac{f^{(3)}(p)}{3!}h^3+\cdots+\frac{f^{(n)}(p)}{n!}h^n\right)\ \blacksquare$$

Oczywiście wielomian Taylora określony jest dla wszystkich liczb h, natomiast reszta tylko dla takich h, dla których punkt $p+h$ znajduje się w dziedzinie funkcji f. Jasne jest też, że po to, by móc mówić o pochodnej $f^{(n)}(p)$ trzeba założyć istnienie pochodnej $f^{(n-1)}$ oraz wszystkich pochodnych niższego rzędu w pewnym otoczeniu punktu p. Zachodzi następujące

Twierdzenie G.Peano

Jeśli f jest funkcją n–krotnie różniczkowalną w punkcie p, to $\lim\limits_{h\to 0}\frac{r_n(h)}{h^n}=0$.

Równość $f(p+h)=f(p)+f'(p)h+\frac{f''(p)}{2!}h^2+\frac{f^{(3)}(p)}{3!}h^3+\cdots+\frac{f^{(n)}(p)}{n!}h^n+r_n(h)$ nazywana bywa wzorem Taylora z resztą Peano, jeśli dodamy informację zawartą w twierdzeniu Peano.

Wynika ono natychmiast z lematu o funkcjach ściśle przylegających. $\blacksquare$

Również z tego lematu wynika, że innego wyboru nie ma, jeśli chcemy mieć tak dokładne przybliżenie i nie chcemy zwiększać stopnia wielomianu ponad niezbedne minimum.

Twierdzenie o jednoznaczności wielomianu Taylora

Jeśli funkcja f jest n–krotnie różniczkowalna w punkcie p i w jest wielomianem stopnia nie większego niż n, tzn. istnieją liczby a_0, a_1,…, a_n, takie że dla każdej liczby rzeczywistej x zachodzi równość $w(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$ oraz $\lim\limits_{h\to 0}\frac{f(p+h)-w(p)}{h^n}=0$, to dla każdego $j\in\{0,1,2,\ldots,n\}$ zachodzi wzór $f^{(j)}(p)=j!a_j$, a więc w jest wielomianem Taylora funkcji f w punkcie p. $\blacksquare$

Nadmienić wypada, że Taylor był współczesny Newtonowi, wzór Taylora znaleziony został od razu. Idea przybliżania dokładniejszego niż liniowe była obecna w omawianej teorii od samego początku! Również współczesny Newtonowi był Szkot o nazwisku Maclaurin, którego nazwiskiem opatrywany jest wzór Taylora w przypadku $p=0$. Zaznaczmy jeszcze, że z wzorem Taylora związane jest szereg Taylora funkcji: $\sum\limits_{n=0}^{\infty}\frac{f^{(n)}(p)}{n!}h^n$. Szereg ten może mieć dodatni promień zbieżności lub zerowy. Po to, by w ogóle można było o nim mówić trzeba założyć, że funkcja ma w punkcie p pochodne wszystkich rzędów. Jednak nawet wtedy może mieć on zerowy promień zbieżności lub mieć sumę różną od $f(p+h)$. W przypadku $p=0$ mówi się zazwyczaj o szeregu Maclaurina. Czytelnik poznał już rozwinięcia w szereg Maclaurina funkcji wykładniczej o podstawie e: $e^x=\sum\limits_{n=0}^{\infty}\frac{x^n}{n!}$, funkcji sinus: $\sin x=\sum\limits_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!}$, funkcji kosinus: $\cos x=\sum\limits_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!}$, funkcji arkus tangens: $\operatorname{arctg} x=\sum\limits_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{2n+1}$ oraz rozwinięcie w szereg Taylora funkcji ln wokół punktu $p=1$: $\ln(1+x)=\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{x^n}{n}$, funkcji potęgowej o wykładniku $a\in\mathbb{R}$ wokół punktu $p=1$: $(1+x)^a=\sum\limits_{n=0}^{\infty}\binom{a}{n}x^n$, również funkcji arkus sinus i funkcji $\frac{x}{x^2+5x+6}$.

Definicja lokalnego ekstremum

Mówimy, że funkcja f określona na zbiorze zawierającym przedział I o środku w punkcie p ma w tym punkcie lokalne maksimum wtedy i tylko wtedy, gdy istnieje przedział $J \subset I$ o środku w punkcie p, taki że jeśli $x \in J$, to $f(x) \leq f(p)$. Jeśli nierówność jest ostra dla $x \neq p$, to mówimy, że lokalne maksimum jest właściwe. Analogicznie określamy lokalne minimum oraz lokalne minimum właściwe. Jeśli funkcja ma w punkcie p lokalne maksimum lub lokalne minimum, to mówimy, że ma w tym punkcie lokalne ekstremum. ∎

Jasne jest, że funkcje x^2, x^4, x^6 ... mają w punkcie 0 minima, natomiast funkcje przeciwne $-x^2$, $-x^4, -x^6$ ... mają w punkcie 0 maksima. Funkcje x, x^3, x^5... nie mają w punkcie 0 ekstremów, nawet lokalnych. Udowodnimy teraz twierdzenie pozwalające w licznych przypadkach łatwo stwierdzić, czy funkcja $n-$krotnie różniczkowalna w punkcie p ma w nim lokalne ekstremum.

Twierdzenie o lokalnych ekstremach

Załóżmy, że funkcja f jest $n-$krotnie różniczkowalna w punkcie p oraz że zachodzą równości $0 = f'(p) = f''(p) = \ldots = f^{(n-1)}(p)$ i nierówność $f^{(n)}(p) \neq 0$. Wtedy jeśli n jest liczbą nieparzystą, to funkcja f nie ma w punkcie p lokalnego ekstremum – w dowolnie małym otoczeniu punktu p przyjmuje zarówno wartości większe niż w punkcie p oraz wartości większe niż w punkcie p, jeśli natomiast n jest liczbą parzystą, funkcja to f ma w punkcie p lokalne ekstremum właściwe: minimum, gdy $f^{(n)}(p) > 0$, maksimum – w przypadku $f^{(n)}(p) < 0$.

Dowód. Skorzystamy z wzoru Taylora:

$$f(p+h) = f(p) + \frac{f'(p)}{1!}h + \frac{f''(p)}{2!}h^2 + \cdots + \frac{f^{(n-1)}(p)}{(n-1)!}h^{n-1} + \frac{f^{(n)}(p)}{n!}h^n + r_n(h)$$

Wobec założeń o pochodnych funkcji f w punkcie p możemy napisać

$$f(p+h) = f(p) + \frac{f^{(n)}(p)}{n!}h^n + r_n(h)) = f(p) + h^n\left(\frac{f^{(n)}(p)}{n!} + \frac{r_n(h)}{h^n}\right)$$

Ponieważ $\lim\limits_{h\to 0} \dfrac{r_n(h)}{h^n} = 0$, więc istnieje $\delta > 0$, taka że jeśli $0 < |h| < \delta$, to $\left|\dfrac{r_n(h)}{h^n}\right| < \left|\dfrac{f^{(n)}(p)}{n!}\right|$. Znak sumy dwu liczb jest taki sam jak znak tej z nich, której wartość bezwzględna jest większa. W przypadku sumy $\dfrac{f^{(n)}(p)}{n!} + \dfrac{r_n(h)}{h^n}$ jest on więc, przy założeniu, że $0 < |h| < \delta$ taki jak znak liczby $f^{(n)}(p)$ ($n!$ nie ma wpływu znak). Jeśli n jest liczbą nieparzystą, to znak iloczynu $h^n\left(\dfrac{f^{(n)}(p)}{n!} + \dfrac{r_n(h)}{h^n}\right)$ zmienia się wraz ze zmianą znaku h. Jeśli n jest liczbą parzystą, to znak ten jest niezależny od znaku h: w przypadku $f^{(n)}(p) < 0$ liczba $h^n\left(\dfrac{f^{(n)}(p)}{n!} + \dfrac{r_n(h)}{h^n}\right)$ jest ujemna, zaś w przypadku $f^{(n)}(p) > 0$ – dodatnia. Stąd teza wynika od razu. ∎

Podany przed chwilą dowód ilustruje jak stosowany jest wzór Taylora: Pewna własność przysługuje wielomianowi Taylora, reszta nie jest w stanie jej zmienić, bo jest za mała. Oczywiście istotnym założeniem jest $f^{(n)}(p) \neq 0$ – bez niego nie mamy podstaw do twierdzenia, że reszta jest mała w porównaniu z wielomianem Taylora funkcji $f(x) - f(p)$, przeciwnie w takim przypadku wszystkie informacje o zachowaniu się funkcji w pobliżu punktu p zawarte są w reszcie, o której niewiele wiemy! Po drugie wypada podkreślić, że mówimy tu jedynie o zachowaniu się funkcji w pobliżu punktu p, na nic więcej nie możemy liczyć, bo założenia, które uczyniliśmy dotyczą jedynie pochodnych w tym jednym punkcie! O wielkości liczby δ również nic nie możemy powiedzieć, jeśli w konkretnej sytuacji musimy coś konkretnego o niej powiedzieć, to wymaga to dalszego badania konkretnej funkcji.*

Przykłady cd.

13. Niech $f(x) = 3x^4 - 28x^3 + 84x^2 - 96x$. Mamy $f'(x) = 12x^3 - 84x^2 + 168x - 96 = 12(x^3 - 7x^2 + 14x - 8) = 12(x-1)(x-2)(x-4)$. Pochodna f' zeruje się jedynie w punktach 1,2,4. Druga pochodna jest równa $f''(x) = 36x^2 - 168x + 168 = 12(3x^2 - 14x + 14)$. wobec tego $f''(1) > 0$, $f''(2) < 0$ i $f''(4) > 0$, więc z twierdzenia o lokalnych ekstremach wynika, że w punktach 1 i 4 funkcja f ma lokalne minima, a w punkcie 2 ma lokalne maksimum. Z tego twierdzenia już więcej nic nie jesteśmy w stanie wywnioskować. Natomiast z twierdzenia o monotoniczności funkcji różniczkowalnych wynika, że na każdym z przedziałów $(-\infty, 1]$, $[1,2]$, $[2,4]$ oraz $[4, +\infty)$ funkcja f jest ściśle monotoniczna, bowiem w ich punktach wewnętrznych pochodna f' funkcji f nie zeruje się. Mamy $f(1) = -37$, $f(2) = -32$ oraz $f(4) = -64$. Wiemy więc, że funkcja f na przedziale $[1,2]$ rośnie, na przedziale $[2,4]$ maleje. Obliczywszy $f(0) = 0 > -37 = f(1)$ stwierdzamy, że na przedziale $(-\infty, 1]$ ta funkcja maleje (wcześniej już stwierdziliśmy, że f jest na tej półprostej ściśle monotoniczna!). Analogicznie z tego, że $f(5) = -5 > -64 = f(4)$ wynika, że na półprostej $[4, +\infty)$ funkcja f jest ściśle rosnąca. Z tego wszystkiego wynika, że $f(4) = -64$ jest najmniejszą wartością funkcji f na całej prostej, $f(1) = -37$ jest najmniejszą wartością funkcji f na przedziale $(-\infty, 2]$ (to nie jest maksymalny przedział, na którym ta wartość jest najmniejsza, ale ustalenie maksymalnego wymagałoby dalszych rozumowań, np. rozwiązania równania $f(x) = f(1)$ w przedziale $[2,4]$). $\blacksquare$

14. Zajmiemy się tą samą funkcją, którą badaliśmy w przykładzie 13: $f(x) = 3x^4 - 28x^3 + 84x^2 - 96x$. Teraz ustalimy jaka jest największa wartość tej funkcji na przedziale $[1,5]$. Pochodna w tym przedziale zeruje się w punktach 1, 2, 4, w punktach 1 i 4 druga pochodna jest dodatnia, więc funkcja ma w nich lokalne minima *właściwe*, więc na pewno nie ma tam wartości największej. Ponieważ f jest ciągła i rozpatrujemy ją na przedziale domkniętym i ograniczonym, więc w pewnym punkcie tego przedziału przyjmuje wartość największą (spośród przyjmowanych na tym przedziale). *standardowy błąd polega na stwierdzeniu, że ponieważ jedynym punktem zerowania się pochodnej oprócz punktów, w których funkcja ma lokalne minima jest 2, więc $f(2) = -32$ jest wartością największą funkcji f na przedziale* $[1,5]$. W rzeczywistości po ustaleniu, gdzie pochodna się zeruje należy rozważyć jeszcze końce przedziału oraz punkty, w których pochodna nie istnieje (w tym przypadku istnieje wszędzie). Wartość największa musi być przyjmowana w jednym z tych punktów. Wobec tego w naszym przypadku jest jeszcze jedna możliwość $x = 5$, drugi koniec przedziału już został rozważony, bo w punkcie $x = 1$ pochodna jest równa 0. Mamy $f(5) = -5 > -32 = f(2)$, więc największą wartością funkcji f na przedziale $[1,5]$ jest liczba $-5 = f(5)$. Dodajmy, że ten przykład jest bardzo prosty, bo chodzi jedynie o przedstawienie roli poszczególnych twierdzeń w badaniu funkcji. $\blacksquare$

15. Pokażemy teraz jak można stosować wzór Taylora do obliczania granic funkcji. Obliczymy mianowicie granicę ilorazu $\frac{(\ln(\cos x))^5}{(x^2 - \sin^2 x)(\operatorname{tg}^2 x)(\cos x - \cos 2x)^2}$ przy $x \longrightarrow 0$. Oczywiście licznik i mianownik dążą do 0, więc można spróbować zastosować regułę de l'Hospitala. Jednak licznik i mianownik wyglądają

* W wielu podręcznikach słowa maksimum, minimum, ekstremum oznaczają lokalne maksimum, lokalne minimum, lokalne ekstremum. Zdecydowaliśmy się na nieco dłuższe terminy, by uniknąć częstych nieporozumień związanych z krótszymi, wielu studentów, zwłaszcza słabiej przygotowanych, myli np. lokalne maksima z globalnymi, co może prowadzić do zupełnie bezsensownych wniosków.

dosyć nieprzyjemnie i można spodziewać się, że po zróżniczkowaniu nie będą wyglądać lepiej. Wobec tego należy zadać sobie pytanie: jak szybko licznik dąży do 0. Potem to samo pytanie należy odnieść do mianownika. Dokładniej: dla jakiej liczby naturalnej n granica $\lim\limits_{x\to 0}\frac{(\ln(\cos x))^5}{x^n}$ jest skończona i różna od 0. Jeśli taka liczba istnieje, to będziemy mówić, że licznik dąży do 0 tak szybko jak x^n. Wiemy, że $\ln(1+y) = y+r(y)$, gdzie r jest taką funkcją, że $\lim\limits_{y\to 0}\frac{r(y)}{y} = 0$ – wynika to z wzoru Taylora zastosowanego do funkcji $\ln$ w punkcie $p = 1$ i $n = 1$, czyli z wzoru na pochodną logarytmu. Jednocześnie zachodzi równość $\cos x = 1 - \frac{1}{2}x^2 + \varrho(x)$, gdzie ϱ jest taką funkcją, że $\lim\limits_{x\to 0}\dfrac{\varrho(x)}{x^2} = 0$ – znów stosujemy wzór Taylora, tym razem chodzi o funkcję kosinus w punkcie 0, $n = 2$.* Stąd wnioskujemy, że $\ln(\cos x) = \ln\left(1-\frac{1}{2}x^2+\varrho(x)\right) = \ln\left(1+\left(-\frac{1}{2}x^2+\varrho(x)\right)\right) = -\frac{1}{2}x^2+\varrho(x)+r\left(-\frac{1}{2}x^2+\varrho(x)\right)$. Mamy $\lim\limits_{x\to 0}\frac{\varrho(x)}{x^2} = 0$ i $\lim\limits_{x\to 0}\frac{r\left(-\frac{1}{2}x^2+\varrho(x)\right)}{x^2} = \lim\limits_{x\to 0}\frac{r\left(-\frac{1}{2}x^2+\varrho(x)\right)}{-\frac{1}{2}x^2+\varrho(x)}\frac{-\frac{1}{2}x^2+\varrho(x)}{x^2} = 0\cdot\left(-\frac{1}{2}\right) = 0$, więc $\lim\limits_{x\to 0}\frac{\ln(\cos x)}{x^2} = -\frac{1}{2}$. Wykazaliśmy więc, że licznik zachowuje się jak $(x^2)^5 = x^{10}$. Teraz zajmiemy się kolejno poszczególnymi członami mianownika. Zaczniemy od $x^2-\sin^2 x$. Mamy $\sin x = x-\frac{x^3}{3!}+\tilde{r}(x)$, gdzie $\dfrac{\tilde{r}(x)}{x^3}\longrightarrow 0$, gdy $x\longrightarrow 0$. Wobec tego zachodzi równość $x^2-\sin^2 x = x^2-\left(x-\frac{x^3}{3!}+\tilde{r}(x)\right)^2 = 2x\frac{x^3}{3!}+\widehat{r}(x)$, gdzie przez $\widehat{r}(x)$ oznaczyliśmy sumę wszystkich pozostałych (niezredukowanych) składników tj. $-\left(\frac{x^3}{3!}\right)^2-(\tilde{r}(x))^2-2x\tilde{r}(x)+2\frac{x^3}{3!}\tilde{r}(x)$. Jasne jest, że $\lim\limits_{x\to 0}\dfrac{\widehat{r}(x)}{x^4} = 0$. Stąd łatwo wynika, że $\lim\limits_{x\to 0}\dfrac{x^2-\sin^2 x}{x^4} = 2\dfrac{1}{3!} = \dfrac{1}{3}$. Jasne jest, że $\lim\limits_{x\to 0}\dfrac{\operatorname{tg} x}{x} = 1$, więc $\lim\limits_{x\to 0}\dfrac{\operatorname{tg}^2 x}{x^2} = 1$. Pozostał ostatni czynnik mianownika. Zastosujemy wzór Maclaurina dla funkcji kosinus i $n = 2$. Mamy $\cos x = 1-\frac{x^2}{2!}+\tilde{\varrho}(x)$, gdzie $\tilde{\varrho}$ jest taką funkcją, że $\lim\limits_{x\to 0}\dfrac{\tilde{\varrho}(x)}{x^2} = 0$ (w rzeczywistości dzięki temu, że wiemy jak przedstawić można funkcję kosinus w postaci sumy szeregu potęgowego, możemy napisać, że $\tilde{\varrho}(x) = \frac{x^4}{4!}-\frac{x^6}{6!}+\cdots$). Wobec tego $\cos x-\cos(2x) = 1-\frac{x^2}{2!}+\tilde{\varrho}(x)-\left(1-\frac{(2x)^2}{2!}+\tilde{\varrho}(2x)\right) = \frac{3}{2}x^2+\widehat{\varrho}(x)$, gdzie $\widehat{\varrho}$ jest taką funkcją, że $\lim\limits_{x\to 0}\dfrac{\widehat{\varrho}(x)}{x^2} = 0$. Wobec tego $\lim\limits_{x\to 0}\dfrac{\cos x-\cos(2x)}{x^2} = \dfrac{3}{2}$, zatem $\lim\limits_{x\to 0}\dfrac{(\cos x-\cos(2x))^2}{x^4} = \dfrac{9}{4}$. Pozostało stwierdzić, że szukana granica to $\frac{\left(-\frac{1}{2}\right)^5}{\frac{1}{3}\cdot 1\cdot\frac{9}{4}} = -\frac{1}{24}$. Postępowanie nasze polegało tu na tym, że zastępowaliśmy funkcje sinus, kosinus, logarytm naturalny wielomianami *odpowiedniego* stopnia, co ułatwiało obliczanie granicy. Można zastosować regułę de l'Hospitala zamiast wzoru Taylora, ale wzór Taylora jest wygodniejszy i nie wymaga większego namysłu. ■

W ostatnim przykładzie pojawiały się w dużych ilościach funkcje, których dokładne definicje nie miały żadnego znaczenia: r, ϱ, $\tilde{r}$, $\tilde{\varrho}$, $\widehat{r}$, $\widehat{\varrho}$. Istotne było jedynie to, że po podzieleniu przez odpowiednią potęgę funkcji x granicą każdej z nich przy $x\longrightarrow 0$ była liczba 0. Zwykle nie wprowadza się tylu oznaczeń. Stosowany jest symbol o. Przyjmujemy mianowicie następującą umowę: piszemy $f(x) = o(g(x))$ przy $x\longrightarrow p$ wtedy i tylko wtedy, gdy $\lim\limits_{x\to p}\dfrac{f(x)}{g(x)} = 0$. Można więc napisać np. $\ln(1+y) = y+o(y)$ przy $y\longrightarrow 0$, bo $\lim\limits_{y\to 0}\dfrac{\ln(1+y)-y}{y} = 0$. Można też napisać, że $x^{10} = o(e^x)$ przy $x\longrightarrow +\infty$, bo $\lim\limits_{x\to\infty}\dfrac{x^{10}}{e^x} = 0$. Mamy również $\sin x = x-\dfrac{x^3}{3!}+o(x^3)$, $\cos x = 1-\dfrac{x^2}{2!}+o(x^2)$ – to są oczywiste wnioski z wzoru Taylora. Przy użyciu właśnie wprowadzonego oznaczenia można zapisać twierdzenie Peano w

* Ponieważ $\ln(1+y)=y-\frac{1}{2}y^2+\cdots$, więc $r(y)=-\frac{y^2}{2}+\frac{y^3}{3}-\cdots$. Analogicznie stosując wzór $\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\cdots$ otrzymujemy równość $\varrho(x)=\frac{x^4}{4!}-\frac{x^6}{6!}+\cdots$.

następujący sposób:

$$f(p+h)=f(p)+\frac{f'(p)}{1!}h+\frac{f''(p)}{2!}h^2+\frac{f^{(3)}(p)}{3!}h^3+\cdots+\frac{f^{(n)}(p)}{n!}h^n+o(h^n)\quad \text{przy}\ \ h\longrightarrow 0\,.$$

Jasne jest, że jeśli $f(x)=o(x^k)$ przy $x\longrightarrow 0$, to $x^nf(x)=o(x^{n+k})$ przy $x\longrightarrow 0$. Jeśli $f(x)=0(x^k)$ przy $x\longrightarrow 0$ i $g(x)=o(x^n)$ przy $x\longrightarrow 0$, to $f(x)g(x)=o(x^{k+n})$ przy $x\longrightarrow 0$ oraz $f(x)+g(x)=o(x^l)$ przy $x\longrightarrow 0$, gdzie $l=\min(k,n)$. W przypadku sumy rezultat nie jest oczywiście „dokładny". Może się zdarzyć, że suma dąży do 0 „szybciej", bo człony decydujące o prędkości zbieżności mogą się zredukować przy dodawaniu lub odejmowaniu. Pokażemy teraz jak przy użyciu symbolu o można opisać rozwiązanie zadania przedstawione w przykładzie 15.

15'. Mamy znaleźć granicę $\lim\limits_{x\to 0}\dfrac{(\ln(\cos x))^5}{(x^2-\sin^2 x)\left(\operatorname{tg}^2 x\right)(\cos x-\cos 2x)^2}$. Skorzystamy z następujących równości $\ln(1+x)=x+o(x)$, $\operatorname{tg} x=x+o(x)$, $\sin x=x-\dfrac{x^3}{3!}+o(x^3)$ i $\cos x=1-\dfrac{x^2}{2!}+o(x^2)$ przy $x\longrightarrow 0$. Z tych równości wynika, że przy $x\longrightarrow 0$ zachodzi wzór $\ln(\cos x)=\ln(1-\dfrac{x^2}{2!}+o(x^2))=$
$=-\dfrac{x^2}{2}+o(x^2)+o\left(-\dfrac{x^2}{2}+o(x^2)\right)=-\dfrac{x^2}{2}+o(x^2)$ – ostatni wzór wynika stąd, że $\lim\limits_{x\to 0}\dfrac{-\frac{x^2}{2}+o(x^2)}{x^2}=$
$=-\dfrac{1}{2}\neq\pm\infty$. Rozumując dalej w taki sam sposób otrzymujemy $x^2-\sin^2 x=$
$=x^2-\left(x-\dfrac{x^3}{3!}+o(x^3)\right)^2=x^2-\left(x^2-2x\dfrac{x^3}{3!}+o(x^4)\right)=\dfrac{x^4}{3}+o(x^4)$ – nie jest oczywiście istotne, czy piszemy $o(x^4)$ czy też $-o(x^4)$. Ponieważ $\operatorname{tg} x=x+o(x)$, więc $\operatorname{tg}^2 x=(x+o(x))^2=$
$=x^2+2xo(x)+(o(x))^2=x^2+o(x^2)$. Przejdźmy do ostatniego etapu: $\cos x=1-\dfrac{x^2}{2!}+o(x^2)$, zatem $\cos 2x=1-\dfrac{(2x)^2}{2!}+o((2x)^2)=1-2x^2+o(x^2)$. Odejmując dwie ostatnie równości stronami otrzymujemy: $\cos x-\cos 2x=\left(-\dfrac{1}{2}+2\right)x^2+o(x^2)=\dfrac{3}{2}x^2+o(x^2)$. Stąd zaś wynika, że
$(\cos x-\cos 2x)^2=\left(\dfrac{3}{2}x^2+o(x^2)\right)^2=\dfrac{9}{4}x^4+3x^2o(x^2)+(o(x^2))^2=\dfrac{9}{4}x^4+o(x^4)$. Wobec tego

$$\lim_{x\to 0}\frac{(\ln(\cos x))^5}{(x^2-\sin^2 x)\left(\operatorname{tg}^2 x\right)(\cos x-\cos 2x)^2}=\lim_{x\to 0}\frac{\left(-\frac{x^2}{2}+o(x^2)\right)^5}{\left(\frac{x^4}{3}+o(x^4)\right)(x^2+o(x^2))\left(\frac{9}{4}x^4+o(x^4)\right)}=$$

$$=\lim_{x\to 0}\frac{x^{10}\left(-\frac{1}{2}+\frac{o(x^2)}{x^2}\right)^5}{\left(x^4\left(\frac{1}{3}+\frac{o(x^4)}{x^4}\right)\right)\left(x^2\left(1+\frac{o(x^2)}{x^2}\right)\right)\left(x^4\left(\frac{9}{4}+\frac{o(x^4)}{x^4}\right)\right)}=$$

$$=\lim_{x\to 0}\frac{\left(-\frac{1}{2}+\frac{o(x^2)}{x^2}\right)^5}{\left(\frac{1}{3}+\frac{o(x^4)}{x^4}\right)\left(1+\frac{o(x^2)}{x^2}\right)\left(\frac{9}{4}+\frac{o(x^4)}{x^4}\right)}=\frac{\left(-\frac{1}{2}\right)^5}{\frac{1}{3}\cdot 1\cdot\frac{9}{4}}=-\frac{1}{24}\,.\ \blacksquare$$

W ten sposób łatwiej jest operować wzorem Taylora, obliczać granice itp. Jednak trzeba pamiętać o tym, że symbol o nie jest normalnym symbolem oznaczającym funkcję – to jest skrót zdania mówiącego, że iloraz dwu wielkości jest zbieżny do 0, np. z równości (prawdziwej) $x-\sin x=o(x^2)$ przy $x\longrightarrow 0$ wynika równość $x-\sin x=o(x)$ przy $x\longrightarrow 0$, ale z tej drugiej równości pierwsza nie wynika: pierwsza równość oznacza bowiem, że $\lim\limits_{x\to 0}\dfrac{x-\sin x}{x^2}=0$ i z niej wynika oczywiście, że $\lim\limits_{x\to 0}\dfrac{x-\sin x}{x}=0$, bo $\lim\limits_{x\to 0}\dfrac{x-\sin x}{x}=\lim\limits_{x\to 0}\left(\dfrac{x-\sin x}{x^2}\cdot x\right)=0$. W przeciwną stronę wnioskować nie można. Trzeba równość $\lim\limits_{x\to 0}\dfrac{x-\sin x}{x^2}=0$ uzasadnić inaczej, można np. skorzystać z wzoru Maclaurina dla funkcji $x-\sin x$ i $n=2$, druga pochodna tej funkcji w punkcie 0 jest równa 0, więc pierwszy wielomian Taylora w punkcie 0 pokrywa się z drugim wielomianem Taylora w punkcie 0. Ogólnie jeśli $f(x)=o(x^2)$ przy $x\longrightarrow 0$, to również $f(x)=o(x)$ przy $x\longrightarrow 0$. Na odwrót być nie musi: $\ln(1+x)-x=o(x)$, natomiast nie jest prawdą, że $\ln(1+x)-x=o(x^2)$!

Warto stosować symbol o, ale trzeba umieć się nim posługiwać, więc studentom którzy mają kłopoty z analizą matematyczną polecam go z dużymi zastrzeżeniami, ci którzy dobrze zrozumieli pojęcie granicy nie powinni mieć z nim problemów, pod warunkiem starannego prześledzenie kilku rozumowań.

16. Znajdziemy raz jeszcze granicę $\lim\limits_{x\to 0}\dfrac{e-(1+x)^{1/x}}{x}$. Mamy $(1+x)^{1/x}=e^{(1/x)\cdot\ln(1+x)}=$
$=e^{(1/x)(x-x^2/2+o(x^2))}=e^{1-x/2+o(x)}=e\cdot e^{-x/2+o(x)}=e\Big(1-\dfrac{x}{2}+o(x)+o\Big(-\dfrac{x}{2}+o(x)\Big)\Big)=$
$=e\Big(1-\dfrac{x}{2}+o(x)\Big)$. Wobec tego $\lim\limits_{x\to 0}\dfrac{e-(1+x)^{1/x}}{x}=\lim\limits_{x\to 0}\dfrac{e-e\Big(1-\dfrac{x}{2}+o(x)\Big)}{x}=$
$=\lim\limits_{x\to 0}\left(\dfrac{e}{2}+\dfrac{o(x)}{x}\right)=\dfrac{e}{2}$, napisaliśmy $o(x)$ zamiast $-e\cdot o(x)$, ale ta operacja jest dozwolona, bo po pomnożeniu funkcji, której granicą jest 0, przez liczbę, otrzymujemy znów funkcję, której granicą jest 0. ■

Zajmiemy się teraz przez chwilę wypukłością funkcji dwukrotnie różniczkowalnych. Przypomnijmy, że funkcja różniczkowalna jest wypukła wtedy i tylko wtedy, gdy jej pochodna jest niemalejąca, ściśle wypukła - wtedy i tylko wtedy, gdy jej pochodna jest ściśle rosnąca. Korzystając z twierdzenia o monotoniczności funkcji różniczkowalnych stwierdzamy natychmiast prawdziwość następującego twierdzenia:

Twierdzenie o wypukłości funkcji dwukrotnie różniczkowalnych
Jeśli funkcja f jest określona na przedziale otwartym i jest dwukrotnie różniczkowalna w każdym punkcie tego przedziału, to jest wypukła wtedy i tylko wtedy, gdy jej druga pochodna f'' przyjmuje jedynie wartości nieujemne.
Dwukrotnie różniczkowalna funkcja określona na przedziale otwartym jest ściśle wypukła wtedy i tylko wtedy, gdy jej druga pochodna f'' jest nieujemna i w każdym przedziale zawartym w jej dziedzinie znajduje się co najmniej jeden punkt, w którym druga pochodna f'' jest dodatnia. ■

W istocie rzeczy badając wypukłość funkcji w poprzednim rozdziale już stosowaliśmy to twierdzenie. W niektórych przypadkach uzasadnienia wypukłości mogłyby zostać nieznacznie skrócone, gdybyśmy powoływali się wprost na twierdzenie o wypukłości funkcji dwukrotnie różniczkowalnych. Zachęcamy czytelnika do ponownego prześledzenia podanych wcześniej przykładów. Teraz natomiast sformułujemy twierdzenie, które w wielu przypadkach pozwala na łatwe znajdowanie punktów przegięcia funkcji wielokrotnie różniczkowalnej.

Twierdzenie o punktach przegięcia funkcji wielokrotnie różniczkowalnych

1. Jeśli p jest punktem przegięcia funkcji f, która jest dwukrotnie różniczkowalna w tym punkcie, to $f''(p) = 0$.

2. Jeśli funkcja f jest n–krotnie różniczkowalna w punkcie p, $n > 2$ i $0 = f''(p) = f^{(3)}(p) = \ldots = = f^{(n-1)}(p)$ i $f^{(n)}(p) \neq 0$, to jeśli n jest liczbą nieparzystą, to p jest punktem przegięcia funkcji f, jeśli natomiast liczba n jest parzysta, to p nie jest punktem przegięcia funkcji f.

Dowód. 1. Z definicji punktu przegięcia wynika, że istnieje liczba $\delta > 0$, taka że na jednym z przedziałów $(p - \delta, p]$, $[p, p + \delta)$ f jest funkcją wypukłą, a na drugim – wklęsłą. Dla ustalenia uwagi przyjmijmy, że na przedziale $(p - \delta, p]$ funkcja f jest wypukła, a na przedziale $[p, p + \delta)$ – wklęsła. Ponieważ f jest dwukrotnie różniczkowalna w punkcie p, więc jest różniczkowalna w punktach pewnego przedziału o środku w punkcie p. Bez straty ogólności można przyjąć, że tym przedziałem jest $(p - \delta, p + \delta)$. Wobec tego na przedziale $(p - \delta, p]$ pochodna f' funkcji f jest niemalejąca i wobec tego jej pochodna, czyli f'', jest nieujemna w każdym punkcie, w którym jest określona, w szczególności $f''(p) \geq 0$. Na przedziale $[p, p + \delta)$ funkcja f jest wklęsła i wobec tego $f''(p) \leq 0$. Ponieważ $f''(p) \leq 0 \leq f''(p)$, więc $f''(p) = 0$.

2. Zastosujemy wzór Taylora do funkcji f'' w punkcie p. Mamy $f''(p + h) = f''(p) + \frac{f^{(3)}(p)}{1!}h + + \cdots + \frac{f^{(n)}(p)}{(n-2)!}h^{n-2} + r_{n-2}(h) = h^{n-2}\left(\frac{f^{(n)}(p)}{(n-2)!} + \frac{r_{n-2}(h)}{h^{n-2}}\right)$. Niech $\delta > 0$ będzie taką liczbą dodatnią, że jeśli $0 < |h| < \delta$, to $\left|\frac{r_{n-2}(h)}{h^{n-2}}\right| < \frac{\left|f^{(n)}(p)\right|}{(n-2)!}$. Liczby $\frac{f^{(n)}(p)}{(n-2)!} + \frac{r_{n-2}(h)}{h^{n-2}}$ i $\frac{f^{(n)}(p)}{(n-2)!}$ mają więc taki sam znak. Jeśli liczba n jest nieparzysta, to liczba h^{n-2} jest dodatnia dla dodatnich h i ujemna dla h ujemnych. Wobec tego liczba $f''(p+h) = h^{n-2}\left(\frac{f^{(n)}(p)}{(n-2)!} + \frac{r_{n-2}(h)}{h^{n-2}}\right)$ jest na jednym z przedziałów $(-\delta, 0)$, $(0, \delta)$ dodatnia, a na drugim - ujemna. Wobec tego na jednym z przedziałów $(p - \delta, p]$, $[p, p + \delta)$ funkcja f jest ściśle wklęsła, a na drugim – ściśle wypukła. Wynika stąd, że p jest punktem przegięcia funkcji f. Jeżeli natomiast liczba n jest parzysta, to wtedy funkcja f'' ma w punkcie p lokalne ekstremum właściwe, więc albo na całym przedziale $(p - \delta, p + \delta)$ z wyjątkiem punktu p funkcja f'' jest dodatnia, albo na całym przedziale $p - \delta, p + \delta$ funkcja f'' jest ujemna. W pierwszym przypadku funkcja f jest ściśle wypukła na całym przedziale $(p - \delta, p + \delta)$, a w drugim – ściśle wklęsła. W żadnym z tych przypadków p nie jest punktem przegięcia funkcji f. Dowód został zakończony. ■

Również to twierdzenie dobrze ilustruje schemat rozumowania przedstawiany w tym rozdziale: funkcja f zachowuje się w dostatecznie małych otoczeniu punktu p tak jak funkcja $\frac{f^{(n)}(p)}{n!}h^n$ w otoczeniu 0 (reszta jest za mała, by mieć istotny wpływ na zachowanie się funkcji!).

17. Niech $f(x) = x^2(x - 6)^2$. Sporządzimy wykres funkcji f. w tym celu ustalimy, na jakich przedziałach funkcja rośnie, na jakich maleje, na jakich jest wypukła, na jakich jest wklęsła, gdzie ma lokalne ekstrema, gdzie punkty przegięcia i znajdziemy asymptoty – oczywiście część wymienionych obiektów może nie istnieć. Obliczymy pochodne: $f'(x) = 2x(x - 6)(x + x - 6) = 4(x^3 - 9x^2 + 18x)$, $f''(x) = 12(x^2 - 6x + 6)$ i $f^{(3)}(x) = 24(x - 3)$. Pierwiastkami pierwszej pochodnej są liczby: 0, 3, 6; drugiej: $3 - \sqrt{3}$ i $3 + \sqrt{3}$ i wreszcie trzeciej: 3. Widać od razu, że w punktach zerowania się pierwszej pochodnej druga przyjmuje wartości różne od 0, wobec tego we wszystkich tych punktach f ma lokalne ekstrema właściwe: w 0 i w 6 – lokalne minima właściwe (bo $f''(0), f''(6) > 0$), a w 3 – lokalne maksimum właściwe (bo $f''(3) = -3 < 0$). Ponieważ na przedziałach $(-\infty, 0]$, $[0,3]$, $[3,6]$ i $[6,\infty)$ funkcja f jest ściśle monotoniczna, bo w ich punktach wewnętrznych pierwsza pochodna jest różna od 0, więc na przedziałach $(-\infty, 0]$ i $[3,6]$ funkcja f maleje, a na przedziałach $[0,3]$ i $[6,\infty)$ – rośnie. Podkreślmy, że choć to bardzo łatwe, to jednak nie badaliśmy znaku pierwszej pochodnej, bo układ lokalnych ekstremów wymusza stwierdzenia na temat wzrostu i spadku wartości funkcji. Oczywiście w ostatecznym rozrachunku wiemy, jaki jest ten znak (pochodna jest różna od 0 w punktach przedziału $(-\infty, 0)$, funkcja maleje na tym przedziale, więc pochodna musi być ujemna, ale ten wniosek wyciągnęliśmy stosując ogólne twierdzenia o zachowaniu się funkcji). Oczywiście z definicji funkcji wynika natychmiast, bez oblicznia pochodnych, że wszystkie jej wartości są nieujemne, więc $0 = f(0) = f(6)$ jest nie tylko lokalnie najmniejszą wartością funkcji, ale również najmniejszą ze wszystkich w ogóle. Inaczej jest z liczbą $81 = f(3)$. W tym przypadku mamy do czynienia z minimum lokalnym: w $f(9) = 81 \cdot 9 > 81$, zatem 81 nie jest największą wartością funkcji, jest nią jeśli ograniczymy dziedzinę do dostatecznie krótkiego przedziału zawierającego 3, zachęcamy do sprawdzenia, że największym przedziałem, na którym funkcja f przyjmuje swą największą wartość w punkcie 3 i w żadnym innym jest $(3-3\sqrt{2}, 3+3\sqrt{2})$. Ponieważ w punktach zerowania się drugiej pochodnej trzecia przyjmuje wartości różne od 0, więc punkty $3-\sqrt{3}$ oraz $3+\sqrt{3}$ są punktami przegięcia funkcji f. Oczywiście pierwszy z nich znajduje się między 0 i 3, a drugi między 3 i 6. Na półprostej $(-\infty, 3-\sqrt{3}]$ funkcja f jest ściśle wypukła, na przedziale $[3-\sqrt{3}, 3+\sqrt{3}]$ – ściśle wklęsła, a na półprostej $[3+\sqrt{3}, +\infty)$ znów ściśle wypukła. jasne jest, że funkcja nie ma asymptot pionowych (jest ciągła w każdym punkcie prostej). Nie ma też ani poziomych ani ukośnych, bo $\lim\limits_{x\to\pm\infty} \big(x^2(x-6)^2 - ax - b\big) = +\infty$ niezależnie od wyboru liczb a i b. Zakończyliśmy badanie funkcji i jesteśmy już w stanie narysować jej wykres.

18. Teraz zbadamy funkcję $\sqrt{1-e^{-x^2}}$. Wzór ten określa ją na całej prostej, więc jest ona ciągła. Jest też różniczkowalna we wszystkich punktach $x \neq 0$, bo dla takich punktów zachodzi nierówność $1 - e^{-x^2} > 0$, a na półprostej $(0, +\infty)$ funkcja pierwiastek kwadratowy jest różniczkowalna. Pierwsza pochodna jest równa $\dfrac{xe^{-x^2}}{\sqrt{1-e^{-x^2}}}$. Jasne jest, że ten wzór nie działa w przypadku $x = 0$. Spróbujmy obliczyć pochodną w punkcie 0 korzystając bezpośrednio z jej definicji. Mamy

$$\lim_{x\to 0^+} \tfrac{f(x)-f(0)}{x} = \lim_{x\to 0^+} \sqrt{\tfrac{1-e^{-x^2}}{x^2}} = \sqrt{\lim_{x\to 0^+} \tfrac{e^{-x^2}-1}{-x^2}} = 1,$$

bo pierwiastek kwadratowy jest funkcją ciągłą (przedostatnia równość) oraz $\lim\limits_{x\to 0} \frac{e^x-1}{x} = 1$ (ostatnia równość). W taki sam sposób stwierdzamy, że $\lim\limits_{x\to 0^-} \frac{f(x)-f(0)}{x} = -1$. Oznacza to, że funkcja nie ma pochodnej w punkcie 0, bowiem jednostronne pochodne są różne. Oznacza to, że w punkcie 0 wykres „załamuje się”, lub też: „ma ostrze”. Znajdziemy drugą pochodną:

$$f''(x) = \left(\tfrac{xe^{-x^2}}{\sqrt{1-e^{-x^2}}}\right)' = \left(xe^{-x^2}\left(1-e^{-x^2}\right)^{-1/2}\right)' =$$
$$= e^{-x^2}\left(1-e^{-x^2}\right)^{-1/2} - 2x^2e^{-x^2}\left(1-e^{-x^2}\right)^{-1/2} - x^2e^{-2x^2}\left(1-e^{-x^2}\right)^{-3/2} =$$
$$= e^{-2x^2}\left(1-e^{-x^2}\right)^{-3/2}\left(e^{x^2}\left(1-2x^2\right) - 1 + x^2\right).$$

Wykażemy, że dla każdego $x \neq 0$ zachodzi nierówność $f''(x) < 0$. Wystarczy wykazać, że jeśli $y > 0$, to $e^y(1-2y) - 1 + y < 0$, piszemy y zamiast x^2. Mamy $(e^y(1-2y) - 1 + y)' =$
$= e^y(1-2y) - 2e^y + 1 = -2ye^y - e^y + 1 < 0$ dla $y > 0$, bo $-e^y + 1 < 0$ w przypadku $y > 0$. Ponieważ pochodna funkcji $e^y(1-2y) - 1 + y$ jest ujemna na półprostej $(0, +\infty)$, więc funkcja ta jest malejąca na

półprostej $[0,\infty)$, a ponieważ jej wartością w punkcie 0 jest 0, więc jej wartości w punktach dodatnich są ujemne.* Z tego, że druga pochodna jest ujemna na każdej z półprostych $(-\infty,0)$ oraz $(0,+\infty)$ wynika, że na każdej z półprostych $(-\infty,0]$, $[0,+\infty)$ funkcja f jest ściśle wklęsła. Nie jest jednak ona ściśle wklęsła na całej prostej, choć jest ciągła w punkcie 0! Jeśli $\delta>0$, to odcinek łączący punkty $\left(-\delta,\sqrt{1-e^{-\delta^2}}\right)$ i $\left(\delta,\sqrt{1-e^{-\delta^2}}\right)$ leży *nad* wykresem (z wyjątkiem końców) funkcji f zamiast pod wykresem. Musiałoby być odwrotnie, gdyby funkcja była ściśle wklęsła lub wklęsła na całej prostej lub choćby na przedziale $[-\delta,\delta]$.

19. Naszkicujemy teraz wykres funkcji f zdefiniowanej wzorem $f(x)=\sqrt[5]{\dfrac{7x^2-3}{9x^2-4}}$ zdefiniowanej dla $x\neq\pm\dfrac{2}{3}$. Zadanie to mieli rozwiązać studenci zaoczni we wrześniu 1996. Poza definicją funkcji podane były wzory na pierwszą i drugą pochodną tej funkcji:

$$f'(x)=-0{,}4x\left(9x^2-4\right)^{-6/5}\left(7x^2-3\right)^{-4/5},$$

$$f''(x)=0{,}24(315x^4-91x^2-20)\left(9x^2-4\right)^{-11/5}\left(7x^2-3\right)^{-9/5}$$

Studenci zostali poinformowani, że druga pochodna przyjmuje wartość 0 wtedy i tylko wtedy, gdy $x=\pm\sqrt{\dfrac{91+\sqrt{33481}}{630}}\approx 0{,}659458$. Jasne jest, że f jest funkcją parzystą, tzn. $f(-x)=f(x)$ dla każdej liczby x z dziedziny funkcji f. Wobec tego jej wykres jest symetryczny względem pionowej osi układu współrzędnych. Wystarczy więc badać f na jednej z półprostych $\left(-\infty,\dfrac{2}{3}\right)$, $\left(\dfrac{2}{3},\infty\right)$ oraz na jednym z przedziałów $\left(-\dfrac{2}{3},0\right]$, $\left[0,\dfrac{2}{3}\right)$. Trzeba więc ustalić na jakich przedziałach funkcja f rośnie, na jakich przedziałach maleje, na jakich przedziałach jest wypukła, a na jakich – wklęsła. Wyjaśnić, w jakich punktach dziedziny funkcja ma pochodną, a w jakich jej nie ma oraz obliczyć granice funkcji f, f', f'' w końcach przedziałów składających się na ich dziedziny. Również ustalić, gdzie są lokalne ekstrema i punkty przegięcia. Z wzoru na pierwszą pochodną wynika, że jest ona określona dla $x\neq\pm\dfrac{2}{3},\ \pm\sqrt{\dfrac{3}{7}}$, przy czym nieistnienie pochodnej w punktach $\pm\dfrac{2}{3}$ wynika z tego, że te punkty są poza dziedziną funkcji f i już to wystarcza, by nie miało sensu różniczkowanie funkcji w tych punktach. W punktach $\pm\sqrt{\dfrac{3}{7}}$ funkcja jest określona, więc teoretycznie nie ma przeszkód dla istnienia pochodnej, jednak wzór nie działa, bo nie można podnieść liczby 0 do potęgi o *wykładniku ujemnym*. Z wzoru na pochodną wynika jednak od razu, że $\lim\limits_{x\to-\sqrt{3/7}}f'(x)=+\infty$ oraz $\lim\limits_{x\to\sqrt{3/7}}f'(x)=-\infty$. Stąd, z definicji pochodnej i twierdzenia Lagrange'a o wartości średniej wynika, że $f'(\pm\sqrt{3/7})=\mp\infty$. Znaczy to, że funkcja f *nie* jest w tych punktach różniczkowalna, bo choć pochodna istnieje, to jest nieskończona. W szczególności w tych punktach wykres ma styczną, tyle że – pionową. Funkcja f jest więc ściśle rosnąca na półprostej $\left(-\infty,-\frac{2}{3}\right)$ oraz na przedziale $\left(-\frac{2}{3},0\right]$. Na przedziale $\left[0,\frac{2}{3}\right)$ oraz na półprostej $\left(\frac{2}{3},\infty\right)$ funkcja f jest ściśle malejąca. Podkreślmy: funkcja rośnie na każdym z dwóch przedziałów, ale nie na ich sumie o czym przekonamy się za chwilę. Zacznijmy od oczywistego stwierdzenia: $\dfrac{3}{7}<\dfrac{4}{9}$. Wobec tego funkcja f przyjmuje wartości dodatnie na półprostej $(-\infty,-\frac{2}{3})$ oraz na przedziale $(-\sqrt{\frac{3}{7}},0)$. Mamy

* Można udowodnić, że $e^y(1-2y)-1+y<0$ dla $y>0$ innymi metodami. Np. można wykorzystać wzór $e^y=\sum_0^\infty\frac{y^n}{n!}$ – otrzymamy po prostych rachunkach szereg, którego wszystkie wyrazy w przypadku $y>0$ są ujemne. Inna metoda to stwierdzenie, że w przypadku $y<1$ zachodzi nierówność $e^y<\frac{1}{1-y}$, zatem dla wszystkich $y\in\mathbb{R}$ zachodzi nierówność $e^y(1-y)<1$, wobec tego $e^y(1-2y)-1+y=e^y(1-y)-y(e^y-1)-1<-y(e^y-1)<0$ dla $y\neq0$.

$\lim\limits_{x\to-\infty}\sqrt[5]{\dfrac{7x^2-3}{9x^2-4}} = \lim\limits_{x\to-\infty}\sqrt[5]{\dfrac{7-3/x^2}{9-4/x^2}} = \sqrt[5]{\dfrac{7}{9}}$. Dalej $\lim\limits_{x\to-\frac{2}{3}^-} f(x) = +\infty$, bo $f(x) > 0$ na półprostej $(-\infty, -\frac{2}{3})$ i licznik dąży do liczby różnej od 0, zaś mianownik do 0. Następnie $\lim\limits_{x\to-\frac{2}{3}^+} f(x) = -\infty$, bo tym razem funkcja jest ujemna, a licznik dąży do liczby różnej od 0, podczas gdy mianownik – do 0. Zajmiemy się teraz wypukłością funkcji f. W tym celu ustalimy, gdzie jej druga pochodna f'' jest dodatnia, gdzie – ujemna. We wzorze na f'' wyrażenia $7x^2-3$ oraz $9x^2-4$ podnoszone są do nieparzystych potęg, następnie z otrzymanych wyników wyciągany jest pierwiastek stopnia nieparzystego. Wynika stąd od razu, że na każdym z kolejnych przedziałów $\left(-\infty, -\dfrac{2}{3}\right)$, $\left(-\dfrac{2}{3}, -\sqrt{\dfrac{91+\sqrt{33481}}{630}}\right)$, $\left(-\sqrt{\dfrac{91+\sqrt{33481}}{630}}, -\sqrt{\dfrac{3}{7}}\right)$ i $\left(-\sqrt{\dfrac{3}{7}}, 0\right)$ pochodna f'' ma inny znak: na pierwszym z wymienionych przedziałów jest dodatnia, na drugim – ujemna, na trzecim – dodatnia i wreszcie na czwartym ujemna. Wynika stąd, że na półprostej $\left(-\infty, -\dfrac{2}{3}\right)$ i na przedziale $\left[-\sqrt{\dfrac{91+\sqrt{33481}}{630}}, -\sqrt{\dfrac{3}{7}}\right]$ funkcja f jest wypukła, a na każdym z przedziałów $\left(-\dfrac{2}{3}, -\sqrt{\dfrac{91+\sqrt{33481}}{630}}\right]$ i $\left[-\sqrt{\dfrac{3}{7}}, 0\right]$ – wklęsła. Wobec tego punkty $-\sqrt{\dfrac{91+\sqrt{33481}}{630}}$ i $-\sqrt{\dfrac{3}{7}}$ są punktami przegięcia funkcji f, $-\dfrac{2}{3}$ punktem przegięcia nie jest, bo leży poza dziedziną funkcji f (ponieważ nie istnieje granica $\lim\limits_{x\to-\frac{2}{3}} f(x)$, więc nie można sensownie dookreślić funkcji w tym punkcie!). Innych punktów przegięcia nie ma: jedynym punktem, który jeszcze nie został zbadany jest 0 – można by pomyśleć, że z wklęsłości funkcji f na przedziale $\left[-\sqrt{\frac{3}{7}}, 0\right]$ oraz z parzystości wynika wklęsłość funkcji na przedziale $\left[-\sqrt{\frac{3}{7}}, \sqrt{\frac{3}{7}}\right]$, tak jest, ale trzeba się jeszcze wyraźnie powołać na różniczkowalność funkcji f w punkcie 0 (por. przykład poprzedni, czyli 18.), jednak takiego twierdzenie nie udowodniliśmy i prościej jest skorzystać z tego, że na przedziale otwartym $\left(-\sqrt{\frac{3}{7}}, \sqrt{\frac{3}{7}}\right)$ druga pochodna f'' funkcji f jest ujemna. Z tego, co do tej pory udało się nam ustalić, wynika, że prosta pozioma $y = \sqrt[5]{\dfrac{7}{9}}$ jest asymptotą poziomą funkcji f przy $x \longrightarrow \pm\infty$, zaś proste pionowe $x = \pm\dfrac{2}{3}$ są obustronnymi asymptotami pionowymi funkcji f przy $x \longrightarrow \pm\dfrac{2}{3}$.

Uwaga: w istocie rzeczy nie trzeba w tym zadaniu wykonywać żadnych obliczeń świadczących o tym, że $\dfrac{2}{3} > \sqrt{\dfrac{91+\sqrt{33481}}{630}} > \sqrt{\dfrac{3}{7}}$ *– ta nierówność wynika z tego, że* $\lim\limits_{x\to-\frac{2}{3}^+} f'(x) = +\infty$ *i* $\lim\limits_{x\to-\sqrt{\frac{3}{7}}^+} f'(x) = +\infty$*, więc na przedziale* $\left(-\dfrac{2}{3}, -\sqrt{\dfrac{3}{7}}\right)$ *pochodna* f' *musi najpierw maleć, a potem rosnąć, co oznacza, że druga pochodna musi przynajmniej w jednym punkcie tego przedziału przyjąć wartość 0, jedynym kandydatem jest punkt* $-\sqrt{\dfrac{91+\sqrt{33481}}{630}}$*. Zachodzi przybliżona równość* $\sqrt{\dfrac{3}{7}} \approx 0{,}654653$*, więc różnica między punktami* $\sqrt{\dfrac{3}{7}}$ *i* $\sqrt{\dfrac{91+\sqrt{33481}}{630}}$ *jest mniejsza niż* $0{,}01$*, zatem może być przeoczona przez program komputerowy rysujący wykresy funkcji (jeśli nie zażądamy odpowiedniej dokładności w tej okolicy!)* . $f(0{,}67) \approx 1{,}288$, $f(0{,}66) = -0{,}908$, więc w tym przypadku zmiana wartości argumentu o $0{,}01$ powoduje zmianę wartości funkcji o około $2{,}196$, więc ponad 200 razy większą niż zmiana argumentu. Rysując wykres na papierze, przyjmując np. że jednostka to 1 cm musimy zwracać uwagę na przedziały długości 0,1 mm, co jest mało realne ze względu na grubość ołówka, linie na rysunku komputerowym też muszą mieć jakąś grubość, więc jedyna rada, to obserwować okolice punktów przegięcia w dużym powiększeniu i zastosowaniu odcinków jednostkowych o różnych długościach na osiach: na osi argumentów odcinek jednostkowy może być np. około 200 razy dłuższy niż odcinek jednostkowy na osi wartości funkcji. ■

W ostatnio prezentowanych przykładach widać było, że w licznych przypadkach można omijać różne obliczenia stosując odpowiednie twierdzenia o charakterze ogólnym. Dużą rolę w tych rozumowaniach odgrywa wzór Taylora. Nasuwa się naturalne pytanie: czy nie można powiedzieć czegoś więcej o reszcie r_n przynajmniej w sytuacji, z którą często mamy do czynienia, mianowicie w przypadku funkcji, która ma więcej pochodnych w otoczeniu punktu p niż n. Okazuje się, że coś powiedzieć można, ale jednak niezbyt dużo. Podamy przykład twierdzenia tego typu, jest ich oczywiście więcej. Stwierdzić jednak wypada, że pożytek z nich na ogół nie jest zbyt wielki, zasadniczo rzecz biorąc twierdzenie Peano to wszystko, co w przypadku ogólnym powiedzieć można.

Twierdzenie Lagrange'a o reszcie we wzorze Taylora

Niech f będzie funkcją, która ma pochodną rzędu $n+1$ w każdym punkcie pewnego przedziału otwartego zawierającego p. Wtedy dla każdego punktu x z tego przedziału istnieje punkt y_x leżący między punktami x i p, dla którego zachodzi równość $r_n(x-p) = \dfrac{f^{(n+1)}(y_x)}{(n+1)!}(x-p)^{n+1}$.

Dowód. Niech $h(t) = \left(f(x) - f(p) - \dfrac{f'(p)}{1!}(x-p) - \cdots - \dfrac{f^{(n)}(p)}{n!}(x-p)^n\right)(x-t)^{n+1} -$

$- (x-p)^{n+1}\left(f(x) - f(t) - \dfrac{f'(t)}{1!}(x-t) - \cdots - \dfrac{f^{(n)}(t)}{n!}(x-t)^n\right).$

Mamy $h(x) = 0 = h(p)$. Ponieważ funkcja f ma w przedziale o końcach x, p $n+1$–ą pochodną, więc funkcja h jest różniczkowalna na tym przedziale, a ponieważ przyjmuje równe wartości w jego końcach, więc w pewnym punkcie wewnętrznym y_x tego przedziału zachodzi równość $h'(y_x) = 0$. Mamy wzór:

$h'(t) = -(n+1)(x-t)^n\left(f(x) - f(p) - \dfrac{f'(p)}{1!}(x-p) - \cdots - \dfrac{f^{(n)}(p)}{n!}(x-p)^n\right) +$

$+(x-p)^{n+1}\dfrac{f^{(n+1)}(t)}{n!}(x-t)^n$. Z tego wzoru wynika od razu, że dla $t = y_x$ zachodzi równość:

$$f(x) = f(p) + \frac{f'(p)}{1!}(x-p) + \cdots + \frac{f^{(n)}(p)}{n!}(x-p)^n + \frac{f^{(n+1)}(y_x)}{(n+1)!}(x-p)^{n+1},$$

czyli właśnie ta, którą chcieliśmy otrzymać. ■

Podkreślmy raz jeszcze: pozornie dzięki temu wzorowi wiemy coś więcej o reszcie. Kłopot polega jednak na tym, że o punkcie y_x występujacym we wzorze Lagrange'a nie wiemy nic, oprócz tego, że leży między p i x. To bardzo ogranicza możliwość wyciągania wniosków idących dalej niż te, które wynikają z wzoru Peano. Oczywiście czasem jest to możliwe. Jeśli np. $f(x) = \sin x$, to dla dowolnego $n \in \mathbb{N}$ mamy $|f^{(n+1)}(x)| \leq 1$, bowiem z dokładnością do znaku pochodna dowolnego rzędu to sinus lub kosinus. Stąd w szczególności wynika, że w przypadku funkcji sinus zachodzi nierówność $|r_n(h)| \leq \dfrac{1}{(n+1)!}|h|^{n+1}$. Otrzymaliśmy więc konkretne oszacowanie, jakiego z pewnością nie da się uzyskać z wzoru Peano. Przeczy to ostrzeżeniom wypowiadanym przed chwilą, ale tylko pozornie. W tym konkretnym przypadku istotą była dodatkowa wiedza o pochodnych dowolnego rzędu badanej funkcji i to ona w połączeniu z wzorem Lagrange'a pozwoliła na wyciągnięcie dalej idących wniosków.

20. Zdefiniujmy funkcję f wzorami

$$f(x) = \begin{cases} e^{-1/x}, & \text{if } x > 0; \\ 0, & \text{if } x \le 0. \end{cases}$$

Jest jasne, że w każdym punkcie, być może z wyjątkiem punktu 0, funkcja ta ma pochodne wszystkich rzędów, czyli jest różniczkowalna nieskończenie wiele razy. W punkcie 0 sytuacja nie jest już jasna, bo z prawej jego strony funkcja jest zdefiniowana inaczej niż z lewej, co mogłoby powodować kłopoty z ciągłością lub różniczkowalnością. Wykażemy poniżej, że w rzeczywistości funkcja f również w punkcie 0 jest różniczkowalna nieskończenie wiele razy oraz że $f^{(n)}(0) = 0$ dla każdej liczby naturalnej n. Wyniknie stąd, że wielomiany Maclaurina tej funkcji są funkcjami zerowymi, a więc dla każdego naturalnego n zachodzi równość $f(x) = r_n(x)$, oczywiście chodzi tu o resztę we wzorze Maclaurina, czyli $r_n(x) = f(x) - f(0) - \frac{f'(0)}{1!}x - \frac{f^{(2)}(0)}{2!}x^2 - \cdots - \frac{f^{(n)}(0)}{n!}x^n$. Oznacza to, że w tym konkretnym przypadku pomijanie reszty może być pozbawione sensu, bo w niej są zawarte wszystkie informacje o funkcji f! Przykład ten omawiamy po to tylko, by przestrzec czytelników, że każda metoda ma swoje ograniczenia, że stosując twierdzenia poprawnie, tj. wtedy, gdy ich założenia są spełnione możemy dochodzić do dziwnych wniosków lub ma wniosków mało interesujących. Po tych pesymistycznych uwagach zajmiemy się wykazaniem równości $f^{(n)}(0) = 0$.

Dla $n = 0$ równość ta jest bezpośrednim wnioskiem z określenia funkcji f w punkcie 0: $f^{(0)}(0) = f(0) = 0$. Jest też jasne, że $\lim\limits_{x\to 0^-} f(x) = 0$ – dla $x < 0$ jest $f(x) = 0$. Mamy też $\lim\limits_{x\to 0^+} f(x) = \lim\limits_{x\to 0^+} e^{-1/x} = 0$. Wykazaliśmy, że funkcja f jest ciągła w punkcie 0. Zauważmy teraz, że dla dowolnego $x < 0$ i dowolnej liczby naturalnej n zachodzi równość $f^{(n)}(x) = 0$, pochodna funkcji stałej jest równa 0, wartość pochodnej w punkcie zależy jedynie od zachowania się funkcji w otoczeniu tego punktu, w naszym przypadku rozpatrujemy chwilowo funkcję f na półprostej $(-\infty, 0)$.Teraz przeniesiemy się na półprostą $(0, \infty)$. W tym przypadku mamy $f(x) = e^{-1/x}$, $f'(x) = \frac{1}{x^2}e^{-1/x}$, $f''(x) = \left(\frac{1}{x^4} - \frac{2}{x^3}\right)e^{-1/x}$, $f^{(3)}(x) = \left(\frac{1}{x^6} - \frac{6}{x^5} + \frac{6}{x^4}\right)e^{-1/x}$, ... Jasne jest, że dla każdej liczby naturalnej n istnieje wielomian w_n stopnia $2n$, taki że $f^{(n)}(x) = w_n\left(\frac{1}{x}\right)e^{-1/x}$, np. $w_1(y) = y^2$, $w_2(y) = y^4 - 2y^3$, $w_3(y) = y^6 - 6y^5 + 6y^4$. Stąd od razu wynika, że $\lim\limits_{x\to 0^+} f^{(n)}(x) = \lim\limits_{y\to\infty} \frac{w_n(y)}{e^y} = 0$. Z definicji pochodnej i twierdzenia o wartości średniej wynika więc od razu, że $f'(0) = \lim\limits_{x\to 0} \frac{f(x) - f(0)}{x} = \lim\limits_{x\to 0} f'(c_x) = 0$, gdzie c_x jest punktem leżącym między 0 oraz x, w szczególności $|c_x| < |x|$. Analogicznie korzystając z już otrzymanego wyniku wnioskujemy, że $f''(0) = 0$ itd. Dowód został zakończony. ■

Uwaga. Funkcja opisana w przykładzie 20 może wydawać się nieco dziwna. Warto zaznaczyć, że takie zachowania się funkcji nie są możliwe w przypadku tzw. funkcji analitycznych, tj. takich funkcji nieskończenie wiele razy różniczkowalnych , które w pewnym otoczeniu dowolnie wybranego punktu dziedziny są równe sumie swego szeregu Taylora. W takim przypadku zerowanie się wszystkich pochodnych w pewnym punkcie powoduje, że funkcja jest stała w otoczeniu tego punktu, co jak widać z poprzedniego przykładu nie musi mieć miejsca w przypadku funkcji różniczkowalnej nieskończenie wiele razy. Istnienie takich funkcji nieskończenie wiele razy różniczkowalnych zauważone zostało nie od razu, stały się one istotnym narzędziem współczesnej matematyki. ■

Na tym kończymy przegląd zagadnień związanych z wielokrotnym różniczkowaniem funkcji.

Informacja: wzór z resztą ogólniejszej postaci (Schlö milcha–Rocha) znajduje się w książce G.M.Fichtenholza, „Rachunek różniczkowy i całkowy”, tom 1.

Zasadnicze twierdzenie algebry

Każdy wielomian o współczynnikach zespolonych, stopnia nie mniejszego od 1, ma co najmniej jeden pierwiastek zespolony.

Dowód. Niech $w(z) = a_0 + a_1z + \cdots + a_nz^n$, przy czym $n \geq 1$ i $a_n \neq 0$. Istnieje liczba $r > 0$, taka że jeśli $|z| \geq r$, to $|w(z)| > |a_0| = |w(0)|$, np. $r = 2 + \dfrac{2|a_0| + |a_1| + |a_2| + \cdots + |a_{n-1}|}{|a_n|}$. Jeśli bowiem $|z| \geq r$, to $|z| \geq 2 > 1$ i wobec tego

$$|w(z)| = |a_0 + a_1z + \cdots + a_nz^n| \geq |a_nz^n| - |a_0 + a_1z + \cdots + a_{n-1}z^{n-1}| \geq$$
$$\geq |a_n||z|^n - \big(|a_0| + |a_1||z| + \cdots + |a_{n-1}||z|^{n-1}|\big) \geq |a_n||z|^n - |z|^{n-1}\big(|a_0| + |a_1| + \cdots + |a_{n-1}||\big) =$$
$$= |z|^{n-1}\Big(|a_n||z| - \big(|a_0| + |a_1| + \cdots + |a_{n-1}|\Big) >$$
$$> \Big((2|a_0| + |a_1| + |a_2| + \cdots + |a_{n-1}|) - (|a_0| + |a_1| + \cdots + |a_{n-1}|)\Big) = |a_0|.$$ Z twierdzenia Bolzano–Weierstrassa wynika, że z ciągu liczb zespolonych (z_n) o module $\leq r$ można wybrać podciąg zbieżny do pewnej granicy g i wtedy oczywiście $|g| \leq r$. Stąd wynika od razu, że istnieje $|z_0|$, takie że $|z_0| \leq r$, $|z| \leq r \Rightarrow |w(z_0)| \leq |w(z)|$, czyli $|w(z_0)|$ jest najmniejszą wartością funkcji $|w|$ na kole o promieniu r i środku w punkcie 0. W szczególności $|w(z_0)| \leq |w(0)| = |a_0|$ i wobec tego również dla $|z| \geq r$ zachodzi nierówność $|w(z)| \geq |a_0| \geq |w(z_0)|$. Oznacza to, że $|w(z_0)|$ jest najmniejszą wartością funkcji $|w|$ na *całej* płaszczyźnie. Wykażemy, że $w(z_0) = 0$. Przyjmijmy, że $z = z_0 + h$. Wtedy możemy napisać $w(z) = w(z_0 + h) = b_0 + b_1h + b_2h^2 + \cdots + b_nh^n$, gdzie $b_0 = a_0 + a_1z_0 + \cdots + a_nz_0^n = w(z_0)$, $b_1 = a_1 + 2a_2z_0 + \cdots + na_nz_0^{n-1} = w'(z_0)$, ..., $b_n = a_n = \frac{1}{n!}w^{(n)}(z_0)$. Z założenia, że stopień wielomianu równy jest n wynika, że $0 \neq a_n = b_n$. Niech m będzie najmniejszą liczbą ≥ 1 taką, 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

$$|w(z)| = \big|b_0 + b_mh^m + b_{m+1}h^{m+1} + \cdots + b_nh^n\big| \leq \big|b_0 + b_mh^m\big| + \big|b_{m+1}h^{m+1} + \cdots + b_nh^n\big| =$$
$$= |b_0| - \varrho^m - \big|b_{m+1}h^{m+1} + \cdots + b_nh^n\big| \leq |b_0| - \varrho^m - \big(|b_{m+1}||h|^{m+1} + \cdots + |b_n||h|^n\big) \leq$$
$$\leq |b_0| - \varrho^m + |h|^{m+1}\big(|b_{m+1}| + \cdots + |b_n|\big) = |b_0| - \varrho^m + \varrho^{m+1}\big(|b_{m+1}| + \cdots + |b_n|\big) \leq |b_0| - \varrho^m + \tfrac{1}{2}\varrho^m =$$
$$= |b_0| - \tfrac{1}{2}\varrho^m < |b_0|.$$

Okazało się, że wbrew założeniu $|w(z_0)|$ *nie* jest najmniejszą wartością funkcji $|w|$. To kończy dowód tego, że $w(z_0) = 0$. Twierdzenie zostało więc wykazane. ■

Wniosek z zasadniczego twierdzenia algebry

Każdy wielomian o współczynnikach rzeczywistych, którego stopień nie jest mniejszy od 1, może być przedstawiony w postaci iloczynu wielomianów rzeczywistych stopnia pierwszego i drugiego.

Dowód. Jeśli współczynniki wielomianu w są liczbami rzeczywistymi, to $\overline{w(z)} = w(\overline{z})$ – prosty dowód tej równości wykorzystujący jedynie najprostsze własności sprzężenia czytelnik przeprowadzi samodzielnie. Z tej równości wynika, że jeśli liczba zespolona z_0 jest pierwiastkiem wielomianu o współczynnikach rzeczywistych, to również liczba zespolona $\overline{z_0}$ jest pierwiastkiem tego wielomianu. Wobec tego jeśli $z_0 \notin \mathbb{R}$, to wielomian w jest podzielny przez wielomian $(z - z_0)(z - \overline{z_0}) = z^2 - (z_0 + \overline{z_0})z + |z_0|^2$. Współczynniki tego wielomianu są liczbami rzeczywistymi, więc w ten sposób sprowadzamy problem do wielomianu stopnia o 2 mniejszego od w. Jeśli z_0 jest liczbą rzeczywistą, to wielomian w jest podzielny przez wielomian $z - z_0$, więc w tym przypadku redukujemy problem do wielomianu stopnia o 1 mniejszego od w. W obu przypadkach zmniejszamy stopnień interesującego nas wielomianu. Gdyby dowodzone twierdzenie nie było prawdziwe moglibyśmy założyć, że w jest wielomianem *najmniejszego* stopnia, dla którego nie zachodzi teza. Po podzieleniu go przez wielomian stopnia 1 lub otrzymujemy wielomian stopnia mniejszego, więc rozkładalny, a to dowodzi, że każdy wielomian o współczynnikach rzeczywistych może być przedstawiony w postaci iloczynu wielomianów stopnia pierwszego i drugiego o współczynnikach rzeczywistych. ■