Analiza matematyczna dla informatyków

wykład dla pierwszego roku informatyki na Wydziale Matematyki, Informatyki i Mechaniki Uniwersytetu Warszawskiego

(skrypt wykładu w roku akademickim 2008/2009)

Marcin Moszyński

Skład w systemie $\TeX$

w wersji 2007/8:

- Tomasz Idziaszek
- Tomasz Kazana
- Piotr Stańczyk

w bieżącej wersji (ulepszenia, dodatki i ogólny „nadzór”):

- Tomasz Kazana

Oznaczenia „edytorskie”:

$\square$	— koniec dowodu (ewentualnie jego szkicu)
B.D.	— bez dowodu (choć czasem brak dowodu jest sygnalizowany inaczej)
$\forall$	— (w spisie zadań po każdym rozdziale) zadanie „obowiązkowe”, tj. do zrobienia we wszystkich grupach ćwiczeniowych

O wykładzie i o skrypcie

Niniejszy skrypt zawiera moje wykłady wygłaszane dla studentów pierwszego roku informatyki na Wydziale Matematyki, Informatyki i Mechaniki Uniwersytetu Warszawskiego w roku akademickim 2008/2009 (semestr zimowy i letni). Obecna wersja jest sukcesywnie rozszerzana o kolejne rozdziały w miarę postępu wykładu. Ściślej — to najpierw powstają kolejne rozdziały skryptu (przede wszystkim w oparciu o mój skrypt z ubiegłego roku) i z jego pomocą prowadzę wykłady.

Po każdym z rozdziałów dołączony jest dotyczący go zestaw zadań z zaznaczonymi zadaniami do zrobienia na ćwiczeniach we wszystkich grupach.

Semestr zimowy (ok. 15 wykładów po 90 minut) to rozdziały I — VI. Obejmuje on kilka podstawowych działów analizy ujętych w sposób dosyć skrótowy, choć zawierających najważniejsze pojęcia i twierdzenia każdego z nich. Te działy to kolejno: szkic teorii aksjomatycznej liczb rzeczywistych, teoria ciągów i szeregów liczbowych, funkcje jednej zmiennej — granica, ciągłość, rachunek różniczkowy oraz zbieżność ciągów i szeregów funkcyjnych.

Rozdziały VII — XI dotyczą (/będą dotyczyły) semestru letniego (ok. 21 wykładów). Poza rachunkiem całkowym jednej zmiennej (z całką Riemanna), stanowiącym uzupełnienie klasycznej tematyki „Analizy I" z semestru zimowego, jest (/będzie) tu przegląd kilku dalszych ważnych, związanych z analizą, działów matematyki. Z konieczności, w tej części wykładu bardzo wiele twierdzeń musi być formułowanych bez dowodów. Omawiamy (/omówimy) tu przestrzenie metryczne, funkcje wielu zmiennych — ciągłość i rachunek różniczkowy, teorię miary (z całką Lebesgue'a) użytą do całkowania funkcji wielu zmiennych oraz równania różniczkowe zwyczajne.

Wykład ten jest w zasadzie samowystarczalny, choć Słuchacz/Czytelnik może z powodzeniem korzystać także z wielu pozycji bogatej literatury obejmującej powyższe działy analizy. Spośród zwięzłych ujęć tematyki o nieco zbliżonym zakresie polecam np.:

- (ad. rozdziały I — VII) Kazimierz Kuratowski, *Rachunek różniczkowy i całkowy*, PWN (Biblioteka matematyczna, tom 22)
- (ad. rozdziały VIII — XI) wybrane fragmenty książki Witolda Kołodzieja *Analiza matematyczna*, PWN (Matematyka dla politechnik)

Szanowny Czytelniku!

Będę wdzięczny za wszelkie uwagi dotyczące skryptu. Można je np. przesyłać na mój adres e-mailowy: `mmoszyns@mimuw.edu.pl`

Autor

Spis treści

I Liczby rzeczywiste — szkic teorii aksjomatycznej, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, potęga rzeczywista

[około $2\frac{1}{2}$ wykładu]

1. Nieco oznaczeń

Punktem wyjścia do całej właściwie matematyki jest teoria mnogości (tj. zbiorów) i logika. Potrzebujemy ich więc także w analizie matematycznej. Sporo elementów powyższych teorii poznacie Państwo na wykładzie „Podstawy matematyki". Oczekuję, że nie jest Panstwu obca podstawowa symbolika logiczna (np. sens symboli $\Rightarrow$, $\vee$, $\wedge$) i rachunku zbiorów (np. sens $\cup$, $\cap$). Teraz wyjaśnię zatem tylko kilka potrzebnych nam symboli — liczę, że przynajmniej częściowo znajomych.

- **kwantyfikatory**:
  $\forall$ — „dla każdego" (od ang. **A**LL; wersja „szkolna" — $\bigwedge$),
  $\exists$ — „istnieje" (od **E**XISTS; wersja „szkolna" — $\bigvee$);
- **„indeksowane" działania na zbiorach** (uogólnienia $\cap$ i $\cup$):
  jeśli I to pewien zbiór („indeksów") oraz dla każdego $i \in I$ zbiór X_i jest podzbiorem zbioru X, to
  $\bigcap_{i\in I} X_i := \{x \in X\colon \forall_{i\in I}\ x \in X_i\}$,[2)]
  $\bigcup_{i\in I} X_i := \{x \in X\colon \exists_{i\in I}\ x \in X_i\}$;
- **funkcje**: $f\colon A \to B$ — funkcja ze zbioru A w zbiór B.

2. Aksjomaty liczb rzeczywistych

Co to są liczby rzeczywiste, tj. jak się nimi posługiwać, jakie obowiązują dla nich reguły — to dość dobrze każdy z Państwa wie; przynajmniej macie już Państwo wyrobione nawyki i rozwinięte intuicje ich dotyczące. Dla matematyka (i dla informatyka...) to jednak za mało. My potrzebujemy ścisłych reguł rozumowania i narzędzi weryfikowania hipotez. Zapewni nam to *teoria aksjomatyczna*. Najpierw przyjmujemy więc kilka podstawowych pojęć (tzw. *pojęć pierwotnych*), takich, które w naszej teorii przyjmujemy bez definicji. Są to: $\mathbb{R}$ — zbiór liczb rzeczywistych, dwie operacje $+$ i $\cdot$, dwa wyróżnione elementy zbioru $\mathbb{R}$ — mianowicie 0 i 1 oraz relację (porządku) $\leqslant$. Wszystkie pozostałe obiekty będziemy musieli zdefiniować.

Drugi „fundament" to *aksjomaty* (inaczej *pewniki*), czyli te własności dotyczące powyższych pojęć pierwotnych, które przyjmujemy za punkt wyjścia w naszej teorii. Przyjmujemy je zatem bez żadnego dowodu, jako fakty niepodważalne. Natomiast wszystkie inne twierdzenia (dla niektórych z nich będziemy używali też innych nazw: lemat, własność, wniosek, fakt itp.) będą już wymagały dowodu, który będzie musiał być ścisłym logicznie rozumowaniem, wykorzystującym wyłącznie aksjomaty (które właściwe także są twierdzeniami, tyle że niezbyt

[2)] Symbol $:=$ lub $=:$ z formalnego punktu widzenia to to samo co $=$, natomiast będziemy go używać raczej tylko wtedy, gdy wprowadzamy (definiujemy) jakieś nowe oznaczenie; dwukropek „:" jest wówczas po stronie definiowanego obiektu.

„trudnymi"...) lub twierdzenia wcześniej udowodnione[3)] . Oczywiście aksjomaty będą własnościami w pełni zgodnymi z naszą intuicją. Będzie ich na tyle dużo, by „wszystko co trzeba" dało się przy ich pomocy udowodnić. Ponadto (co już znacznie mniej ważne) na tyle mało, by jedne z drugich nie wynikały (tzw. *niezależność aksjomatów*).

Oto one (jest ich kilkanaście, podajemy je „po trochu"). Pierwsze cztery mówią, że trójka $(\mathbb{R}, 0, +)$ jest *grupą przemienną*, tzn.:

(D1.) (*łączność* $+$) $\forall_{x,y,z\in\mathbb{R}}\ (x+y)+z = x+(y+z)$;

(D2.) (*neutralność 0*) $\forall_{x\in\mathbb{R}}\ x+0 = 0+x = x$;

(D3.) (*istnienie elementu przeciwnego*) $\forall_{x\in\mathbb{R}}\ \exists_{y\in\mathbb{R}}\ x+y = y+x = 0$;

(D4.) (*przemienność* $+$) $\forall_{x,y\in\mathbb{R}}\ x+y = y+x$.

Uwaga. Same aksjomaty (D1.)–(D3.) stanowią de facto definicję *grupy*. Jako sprawdzian zrozumienia powyższej uwagi, proponuję samodzielne dokończenie poniższej definicji.

Definicja. *Trójka* $(G, e, \odot)$*, gdzie* G *— zbiór,* $e \in G$*,* $\odot$ *— operacja w* G *(tzn.* $\odot\colon G\times G \to G$*) jest* ***grupą*** *wtedy i tylko wtedy, gdy*[4)] *... Grupa ta jest* ***przemienna*** *(inaczej* ***abelowa****) wtw ...*

Kolejne aksjomaty dotyczą mnożenia i liczby 1. Dla wygody oznaczymy $\mathbb{R}^* := \mathbb{R}\setminus\{0\}$.

(M1.) (*łączność* $\cdot$) $\forall_{x,y,z\in\mathbb{R}}\ (x\cdot y)\cdot z = x\cdot(y\cdot z)$;

(M2.) (*neutralność 1*) $\forall_{x\in\mathbb{R}}\ x\cdot 1 = 1\cdot x = x$;

(M3.) (*istnienie elementu odwrotnego*) $\forall_{x\in\mathbb{R}^*}\ \exists_{y\in\mathbb{R}}\ x\cdot y = y\cdot x = 1$;

(M4.) (*przemienność* $\cdot$) $\forall_{x,y\in\mathbb{R}}\ x\cdot y = y\cdot x$.

Analogia (M1.)–(M4.) do (D1.)–(D4.) narzuca się sama, choć widać pewną różnicę w (M3.) (jaką?). Gdy dodamy następny aksjomat, a mianowicie

(01.) $0 \neq 1$,

łatwo będzie dowieść (zachęcam), że $\forall_{x,y\in\mathbb{R}^*}\ x\cdot y \neq 0$, a zatem, że mnożenie $\cdot$ można „obciąć" do mnożenia $\tilde{\cdot}$ w $\mathbb{R}^*$ (tj. $\tilde{\cdot}\colon (\mathbb{R}^*)\times(\mathbb{R}^*) \to \mathbb{R}^*$ i $x\tilde{\cdot}y := x\cdot y$ dla $x, y \in \mathbb{R}^*$) i $(\mathbb{R}^*, 1, \tilde{\cdot})$ jest grupą (także przemienną).

Kolejny aksjomat opisuje ważną własność dotyczącą jednocześnie dodawania i mnożenia:

(DM.) (*rozdzielność mnożenia względem dodawania*) $\forall_{x,y,z\in\mathbb{R}}\ x\cdot(y+z) = x\cdot y + x\cdot z$.

Wymienione dotąd aksjomaty stanowią razem definicję *ciała* (a dokładniej — ciała przemiennego — gdyż niektórzy wyłączają przemienność mnożenia z definicji ciała). Następne dwa aksjomaty wiążą ze sobą działania i relację $\leqslant$:

(DP.) $\forall_{x,y,z\in\mathbb{R}}\ x \leqslant y \Rightarrow x+z \leqslant y+z$,

(MP.) $\forall_{x,y,z\in\mathbb{R}}\ (x \leqslant y \wedge 0 \leqslant z) \Rightarrow xz \leqslant yz$.

Gdy dołożymy jeszcze cztery aksjomaty dotyczące samej relacji $\leqslant$:

(P1.) (*zwrotność*) $\forall_{x\in\mathbb{R}}\ x \leqslant x$,

(P2.) (*słaba antysymetria*) $\forall_{x,y\in\mathbb{R}}\ (x \leqslant y \wedge y \leqslant x) \Rightarrow x = y$,

(P3.) (*przechodniość*) $\forall_{x,y,z\in\mathbb{R}}\ (x \leqslant y \wedge y \leqslant z) \Rightarrow x \leqslant z$,

(P4.) (*spójność*) $\forall_{x,y\in\mathbb{R}}\ x \leqslant y \vee y \leqslant x$,

to otrzymamy *ciało uporządkowane*.

Pytanie, czy to już wszystkie potrzebne aksjomaty. Nie, bo zbiór liczb wymiernych $\mathbb{Q}$ (ściśle zdefiniujemy go wkrótce) ze zwykłymi działaniami i nierównością wszystkie powyższe

---

[3)] Uwaga! Ten idealistyczny program będziemy realizowali niestety z licznymi odstępstwami — niektóre dowody będziemy na wykładzie pomijali lub skracali, a niektóre znane, czy oczywiste dla Państwa twierdzenia (w tym pewne analogi sformułowanych już twierdzeń) będziemy przemilczali. A to, by Państwa nie zanudzić i by zdążyć na czas z obszernym programem.

[4)] Dalej „wtedy i tylko wtedy, gdy" skracamy do wtw.

aksjomaty spełnia, a przecież $\mathbb{Q}$ i $\mathbb{R}$ różnią się między sobą wieloma własnościami (np. jaką?). Na szczęście, to czego brakuje to tylko jeden aksjomat, choć już nie tak intuicyjny, jak wcześniejsze. By go zgrabnie sformułować, przyjmijmy następujące definicje:

Definicja. *Niech $A \subseteq \mathbb{R}$, $b \in \mathbb{R}$.*

- *b jest **ograniczeniem górnym (dolnym)** [5] zbioru A wtw $\forall_{a\in A}\ a \leqslant b$ ($b \leqslant a$).*
- *Zbiór A jest **ograniczony z góry (z dołu)** wtw istnieje $b \in \mathbb{R}$, będące ograniczeniem górnym (dolnym) A.*
- *Zbiór A jest **ograniczony** wtw jest ograniczony z góry i z dołu.*
- *b jest **elementem największym (najmniejszym)** zbioru A wtw $b \in A \wedge \forall_{a\in A}\ a \leqslant b$ ($b \leqslant a$).*
- *Niech A będzie ograniczony z góry (z dołu), wówczas b jest **kresem górnym**, czyli **supremum** (**kresem dolnym**, czyli **infimum**) zbioru A wtw b jest elementem najmniejszym (największym) zbioru wszystkich ograniczeń górnych (dolnych) A.*

Teraz możemy sformułować ostatni z aksjomatów:

(Z.) (*aksjomat zupełności* [6]) Dla każdego niepustego, ograniczonego z góry zbioru $A \subseteq \mathbb{R}$ istnieje $b \in \mathbb{R}$ taki, że b jest kresem górnym A.

A zatem „nasza” teoria liczb rzeczywistych opiera się na zestawie siedemnastu aksjomatów. Ale w skrócie $\mathbb{R}$ to po prostu ciało (przemienne) uporządkowane, zupełne.

Tu należy się Państwu jedna ważna uwaga. Poza regułami wynikającymi z podanych już aksjomatów, w całej matematyce, a zatem także w uprawianej przez nas teorii, obowiązują dodatkowo pewne ogólne reguły. Są to przede wszystkim reguły teorii mnogości (czyli teorii zbiorów) oraz logiki. Zawierają one np. zasady posługiwania się formułami matematycznymi, kwantyfikatorami, czy relacją równości (choćby to, że jeśli $a = b$, to $b = a$; albo że jeśli $a = b$ i $b = c$, to $a = c$). Oczywiście nie jesteśmy w stanie ich wszystkich tu omówić, ale dla uspokojenia wspomnę, że są one dość naturalne i intuicyjne oraz że więcej na ten temat dowiecie się Państwo na wykładach przedmiotu Podstawy Matematyki[7] .

Nie da się ukryć, że uprawianie teorii aksjomatycznej, szczególnie na samym początku, bywa dość żmudne. Ograniczymy się więc tylko do paru przykładów pokazujących „jak to działa”, a inne znane nam dobrze elementarne własności liczb rzeczywistych przyjmiemy bez dowodu, choć zachęcam do samodzielnego uzupełniania tych luk.

Twierdzenie I.1. $\forall_{x\in\mathbb{R}}\ \exists!_{y\in\mathbb{R}}$ [8] $x + y = 0$

Dowód.

Istnienie jakiegoś $y \in \mathbb{R}$ takiego, że $x + y = 0$ gwarantuje nam (D3.). By wykazać jednoznaczność, załóżmy, że $y, y' \in \mathbb{R}$ są takie, że $x + y = 0$ i $x + y' = 0$. Zatem dodając do drugiej równości y, dostajemy $y + (x + y') = y + 0$, zatem z (D1.) i (D2.) $(y + x) + y' = y$, skąd na mocy naszego pierwszego założenia i (D2.) $y' = y$. $\square$

Powyższe twierdzenie pozwala nam zatem zdefiniować[9] jednoznacznie *element przeciwny* do x jako taki $y \in \mathbb{R}$, że $x + y = 0$. Oznaczamy go $-x$. To z kolei pozwala zdefiniować operację *odejmowania* jako $a - b := a + (-b)$.

---

[5] W ten sposób będziemy często zapisywać dwie analogiczne definicje „za jednym zamachem”.

[6] Bywa on często, a nawet częściej, nazywany aksjomatem *ciągłości.*

[7] Można też przejrzeć podręczniki o tytule zbliżonym do ”Wstęp do matematyki”, bądź inne, dotyczące teorii mnogości i logiki matematycznej.

[8] $\exists!$ — „istnieje dokładnie jeden”.

[9] Nie zawsze definicję poprzedzam tytułem „Definicja” — robię to jedynie przy „bardziej uroczystych” okazjach.

Analogicznie postępujemy w przypadku mnożenia i dla $x \neq 0$ uzyskujemy *element odwrotny* do x (oznaczany oczywiście $\frac{1}{x}$ lub x^{-1}), a następnie operację dzielenia („——" lub „:") przez liczby $\neq 0$.

Inny prosty przykład elementarnego twierdzenia to dualna wersja aksjomatu zupełności (Z). Proszę ją sformułować samodzielnie zastępując ograniczoność z góry przez ograniczoność z dołu, a kres górny — dolnym, a następnie proszę pomyśleć nad ścisłym dowodem. Uzyskamy więc:

Twierdzenie I.2. ...

Dowód.

... $\square$

Dla wygody powinniśmy jeszcze przyjąć między innymi następujące definicje:

- $a \geqslant b \Leftrightarrow b \leqslant a$
- $a < b \Leftrightarrow (a \leqslant b \wedge a \neq b)$
- $a > b \Leftrightarrow b < a$

Definiujemy także *moduł* (inaczej *wartość bezwzględną*) liczby $x \in \mathbb{R}$ wzorem:

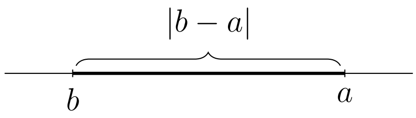
$$|x| := \begin{cases} x & \text{dla } x \geqslant 0 \\ -x & \text{dla } x < 0. \end{cases}$$

Nietrudny dowód poniższego faktu pozostawiam Państwu.

Fakt (nierówność trójkąta).

$$\forall_{x,y\in\mathbb{R}}\ |x+y| \leqslant |x| + |y|.$$

B.D.



Rysunek 1: Długość odcinka

Moduł będzie nam służył między innymi do mierzenia „odległości" pomiędzy liczbami. Tę odległość pomiędzy a oraz b wyrażamy wzorem $|b-a|$ — geometrycznie interpretujemy ją jako długość odcinka łączącego a z b na osi liczbowej będącej z kolei geometryczną interpretacją zbioru $\mathbb{R}$ (patrz rys. 1).

3. Nieco uwag o kresach

Element największy zbioru $A \subset \mathbb{R}$, o ile takowy istnieje, jest na mocy aksjomatu (P2.) wyznaczony jednoznacznie. Oznaczamy go $\max A$. Podobnie jest z elementem najmniejszym; oznaczamy go $\min A$. Oczywiście $\max A$ jest jednocześnie kresem górnym A (a $\min A$ — kresem dolnym), ale np. przedział $(0;1)$ [10] ma kres górny równy 1, a elementu największego nie posiada. Zatem kresy zbioru to coś w rodzaju prawego i lewego „końca" zbioru, które do tego zbioru mogą należeć lub nie. Także kresy, gdy istnieją są oczywiście wyznaczone jednoznacznie

[10] Zakładam, że definicje przedziałów otwartych, domkniętych, otwarto-domkniętych (ograniczonych i nie) są znane ze szkoły. Używamy notacji $(a;b)$, $[a;b]$, $(a;b]$ i $[a;b)$.

(dlaczego?). Kres górny zbioru A oznaczamy symbolem $\sup A$, a kres dolny $\inf A$. Gdy A nie jest ograniczony z góry, to fakt ten oznaczamy $\sup A = +\infty$. Analogicznie dla A nieograniczonego z dołu umownie piszemy $\inf A = -\infty$. Są to jednak na razie tylko oznaczenia, tzn. samo $\pm\infty$ na razie nie jest jeszcze[11)] żadnym „matematycznym” obiektem!

Odwołując się do przykładu z liczbami wymiernymi, który motywował nieco wcześniej dołączenie aksjomatu zupełności, warto tę motywację uzupełnić o uwagę, że w zbiorze liczb wymiernych nie ma zupełności. Znów wyprzedzając nieco ścisłą definicję zbioru liczb wymiernych, można tu podać przykład zbioru $A = \{x \in \mathbb{Q} : x \cdot x \leqslant 2\}$, który jest niepusty i ograniczony w $\mathbb{Q}$ z góry, ale kresu górnego w $\mathbb{Q}$ nie posiada.

4. Liczby naturalne, całkowite, wymierne

Zbiór liczb naturalnych $\mathbb{N}$ to podzbiór zbioru $\mathbb{R}$, który często jest określany jako

$$\{1,\ 1+1,\ 1+1+1,\ \ldots\}.$$

Matematycy do tego zbioru dorzucają jeszcze chętnie 0. Na tym wykładzie tego nie zrobimy (tj. $0 \notin \mathbb{N}$), jednak to jedynie kwestia umowy. Tymczasem dużo ważniejszy problem to sprawa ścisłości powyższej „definicji”, a właściwie — braku ścisłości. Jak bowiem rozumieć ów trzykropek „...”? Aby to uściślić postąpimy następująco.

Definicja. *Niech $B \subset \mathbb{R}$. Zbiór B* ***jest induktywny*** *wtw $1 \in B$ oraz $\forall_{x \in B}\ x + 1 \in B$.*

Jak widać z tej definicji, zbiorów induktywnych jest wiele — np. $\mathbb{R}$, $(-1; +\infty)$, $[1; +\infty)$[12)], ale zgodnie z naszą intuicją induktywne powinny być także inne, niezdefiniowane dotąd zbiory — przede wszystkim $\mathbb{N}$. Ta sama intuicja podpowiada nam, że $\mathbb{N}$ powinien być zawarty w każdym zbiorze induktywnym. Stąd poniższa definicja.

Definicja.

$$\mathbb{N} := \bigcap_{B \in \mathbb{I}} B,$$

gdzie $\mathbb{I}$ oznacza zbiór wszystkich induktywnych podzbiorów $\mathbb{R}$.

Wśród wielu własności zbioru $\mathbb{N}$ znaczenie zasadnicze ma dla nas twierdzenie (być może znane Państwu) zwane *zasadą indukcji zupełnej* (w skr. ZIZ). Sformułujemy je tak:

Twierdzenie I.3 (ZIZ). *Jeżeli $A \subset \mathbb{N}$ spełnia warunki*

1. („warunek początkowy”) $1 \in A$

2. („krok indukcyjny”) $\forall_{n \in \mathbb{N}}\ (n \in A \Rightarrow n + 1 \in A)$,

to $A = \mathbb{N}$.

Dowód.
Zauważmy, że A jest induktywny, bowiem gdy $n \in A$, to na mocy faktu, że $A \subset \mathbb{N}$ oraz zał. 2. otrzymujemy $n + 1 \in A$. Czyli $A \in \mathbb{I}$, a stąd $A \supset \bigcap_{B \in \mathbb{I}} B = \mathbb{N}$. Zatem $A = \mathbb{N}$. □

ZIZ bywa bardzo przydatna przy dowodzeniu wielu matematycznych faktów związanych z liczbami naturalnymi. Dowody używające ZIZ noszą nazwę *dowodów indukcyjnych*. Przykład takiego dowodu pojawi się jeszcze w tym rozdziale (patrz nierówność Bernoulli'ego).

Oto przykłady innych elementarnych własności $\mathbb{N}$. Podajemy je bez dowodu.

[11)] Ale wkrótce również same symbole $+\infty$ i $-\infty$ będą miały swój matematyczny sens.

[12)] Znów użyliśmy niezdefiniowanego symbolu $+\infty$, ale to nie szkodzi, bo nawet nie mając tej definicji możemy zdefiniować od razu cały „napis” oznaczający przedział nieskończony — mianowicie, dla $a \in \mathbb{R}$, oznaczamy (jak było chyba Państwu wiadomo...) $(a; +\infty) := \{x \in \mathbb{R} : x > a\}$; $(-\infty; a] := \{x \in \mathbb{R} : x \leqslant a\}$ i analogicznie dla $[a; +\infty)$ i $(-\infty; a)$.

Twierdzenie I.4 (zamkniętość $\mathbb{N}$ względem + i $\cdot$).

$$\forall_{m,n\in\mathbb{N}}\ m+n, m\cdot n\in\mathbb{N}.$$

Twierdzenie I.5 (zasada Archimedesa). *Dla dowolnego $x\in\mathbb{R}$ oraz $a>0$ istnieje $n\in\mathbb{N}$ takie, że $n\cdot a>x$. W szczególności $\mathbb{N}$ jest nieograniczony* [13] *z góry.*

Potrzebujemy jeszcze parę definicji i oznaczeń istotnych dla nas zbiorów:

- zbiór liczb *całkowitych* $\mathbb{Z}:=\mathbb{N}\cup\{0\}\cup\{x\in\mathbb{R}:-x\in\mathbb{N}\}$;
- $\mathbb{N}_0:=\mathbb{N}\cup\{0\}$, dla $k\in\mathbb{Z}$ $\quad\mathbb{N}_k:=\{n+k:n\in\mathbb{N}_0\}$;
- zbiór liczb *wymiernych* $\mathbb{Q}:=\{\frac{m}{n}:m\in\mathbb{Z},n\in\mathbb{N}\}$;
- zbiór *cyfr*(zapis dziesiętny) $C_{10}:=\{0,1,2,3,4,5,6,7,8,9\}\subset\mathbb{N}_0$, gdzie $2:=1+1$, $3:=2+1,\ \ldots,\ 9:=8+1$.

Niech $n\in\mathbb{N}$ oraz $c_1,\ldots,c_n\in C_{10}$, gdzie $c_1\neq 0$. Zdefiniujemy *rekurencyjnie*[14] liczbę, którą zapisywać będziemy $c_1c_2\ldots c_n$. Gdy $n=1$, to liczba ta jest po prostu równa liczbie c_1. Ponadto dla dowolnego $n\in N$, $c_1c_2\ldots c_nc_{n+1}:=c_{n+1}+(9+1)\cdot c_1c_2\ldots c_n$. ZIZ dowodzi, że tym sposobem liczba n-cyfrowa została zdefiniowana dla dowolnego $n\in\mathbb{N}$. Zachodzi także:

Twierdzenie I.6. *Każda liczba naturalna ma jednoznaczny zapis w powyższej postaci.*

Dowód pomijamy, ograniczając się do wskazówki, że część dotycząca samego istnienia zapisu może być łatwo wykazana przez indukcję. Dlaczego przyjął się akurat zapis dziesiętny? To pytanie raczej z historii matematyki. Jednak czasem przydają się też inne typy zapisu — np. informatykowi bliski powinien być *zapis dwójkowy*, a także *zapis szestnastkowy*. Jak ogólnie, dla dowolnego $k\in\mathbb{N}_2$, zdefiniować zapis „k–tkowy” proszę wymyślić samodzielnie! Proszę przy tym zauważyć, że **pojawi się tu pewien problem** dla $k>10$, jeżeli jako elementów zbioru cyfr C_k zdecydujemy się użyć liczb zapisanych przy użyciu zapisu dziesiętnego (jaki to problem?).

Teraz pewna ważna własność $\mathbb{Z}$.

Twierdzenie I.7 (zasada maksimum). *Każdy niepusty, ograniczony z góry podzbiór $\mathbb{Z}$ posiada element największy.*

Dowód znów pomijamy. Oczywiście można także wykazać analogiczną „zasadę minimum”.

Powyższe twierdzenie pozwala na sformułowanie następującej definicji:

Definicja. ***Część całkowita*** *liczby $x\in\mathbb{R}$ to* $\max\{m\in\mathbb{Z}:m\leqslant x\}$. *Oznaczamy ją $[x]$.* [15]

Lemat.

$$\forall_{x,y\in\mathbb{R}}\ \Big(y-x\geqslant 1\Rightarrow\exists_{m\in\mathbb{Z}}\ x\leqslant m\leqslant y\Big).$$

Dowód.
Wystarczy wziąć $m:=[y]$. Mamy wtedy $m\leqslant y$ z definicji części całkowitej. Przypuśćmy, że $m<x$. Wówczas $m+1<x+1\leqslant y$ oraz $m+1\in\mathbb{Z}$, a zatem $m\neq[y]$ — sprzeczność, więc $m\geqslant x$. $\square$

Na koniec tego „podrozdziału” bardzo ważna własność $\mathbb{Q}$.

---

[13] „jest nieograniczony = nie jest ograniczony”, choć — uwaga! — nie zawsze w matematyce dodanie do pojęcia „nie” daje pojęcie będące zaprzeczeniem wyjściowego – np. niemalejący ...

[14] Tzn. opisując co należy zrobić dla n początkowego (tu np. dla $n=1$) oraz sposób przejścia od dowolnego n do $n+1$. Jak widać idea definicji rekurencyjnej przypomina ideę zawartą w ZIZ i stąd niektórzy nazywają ten rodzaj definicji definicją *indukcyjną*.

[15] Bywają też w użyciu inne oznacznia, ponadto niestety „[ ]” używamy też czasem jako nawiasu... — liczę na Państwa domyślność...

Twierdzenie I.8 (o gęstości $\mathbb{Q}$).

$$\forall_{x,y\in\mathbb{R}}\ \Big(y > x \Rightarrow \exists_{q\in\mathbb{Q}}\ x \leqslant q \leqslant y\Big).$$

Dowód.
Korzystając z zasady Archimedesa wybierzmy $n \in \mathbb{N}$ takie, że $n > \frac{1}{y-x}$. Niech $x' := n \cdot x$, $y' := n \cdot y$. Mamy $y' - x' > 1$, zatem z lematu istnieje $m \in \mathbb{Z}$ takie, że $x' \leqslant m \leqslant y'$, skąd $x \leqslant \frac{m}{n} \leqslant y$. □

5. Potęga rzeczywista

W tej ostatniej części rozdziału I naszkicujemy definicję potęgi x^y dla dowolnych $x > 0$ i $y \in \mathbb{R}$. Zrobimy to etapami, znów z pominięciem wielu dowodów...

Etap 1. *x^n dla $n \in \mathbb{N}$, $x \in \mathbb{R}$.*
Definiujemy rekurencyjnie: $x^1 := x$, $x^{n+1} := x \cdot x^n$ dla $n \in \mathbb{N}$.

Etap 2. *x^n dla $n \in \mathbb{Z}$, $x \neq 0$.*
Dla $n \in \mathbb{N}$ definicja była już w poprzednim etapie. Pozostają przypadki z $n \leqslant 0$. Definiujemy więc $x^0 := 1$ oraz $x^{-n} := \frac{1}{x^n}$ dla $n \in \mathbb{N}$.

Uwaga. Nie zdefiniowaliśmy 0^n dla $n \leqslant 0$. Dla $n < 0$ nie zrobimy tego, jednak **niekiedy**, dla wygody przyjmuje się, że $0^0 = 1$. Np. przyjmuje się tak we wzorze Newtona sformułowanym niżej. Ogólnie należy jednak z tą umową uważać (o istotnych tego powodach przekonacie się Państwo w przyszłości).

Fakt I.1. *Dla dowolnych $x, y \neq 0$ oraz $m, n \in \mathbb{Z}$ zachodzi:*

1. $x^{m+n} = x^m \cdot x^n$,
2. $x^{m\cdot n} = (x^m)^n$,
3. $(x \cdot y)^m = x^m \cdot y^m$.

B.D.

Fakt I.2 (wzór Newtona). *Dla $a, b \in \mathbb{R}$, $n \in \mathbb{N}_0$ zachodzi*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{(n-k)}.$$ [16)]

B.D.

Etap 3. *Definicja $\sqrt[n]{a}$ dla $a \geqslant 0, n \in \mathbb{N}$.*
Chcielibyśmy zdefiniować $\sqrt[n]{a}$ (tj. pierwiastek n–tego stopnia z a) jako liczbę nieujemną, która daje a po podniesieniu do potęgi n-tej. Ale tu pojawia się problem — skąd bowiem gwarancja, że taka liczba w ogóle istnieje? Aby się o tym przekonać, postąpimy nieco ostrożniej — i tu znów przyda się aksjomat zupełności.

Definicja. $\sqrt[n]{a} := \sup A$, *gdzie* $A = \{x \geqslant 0 : x^n \leqslant a\}$.

Zauważmy, że to poprawna definicja — ten kres istnieje, bo powyższy zbiór A jest niepusty (0 do niego należy) oraz ograniczony z góry — gdy $a \leqslant 1$, to np. przez 1, a gdy $a > 1$, to np. przez a. Zgodnie z naszą intencją zachodzi:

[16)] Zakładam, że symbol Newtona $\binom{n}{k}$ jest znany ze szkoły. Symbol „skróconego sumowania" $\sum_{k=m}^{n} a_k$ „definiuje" się nieformalnie jako $a_m + a_{m+1} + \cdots + a_n$, a ścisłą, rekurencyjną definicję pozostawiam do wymyślenia Państwu.

Twierdzenie I.9. $\forall_{a\geqslant 0,\, n\in\mathbb{N}}\ (\sqrt[n]{a})^n = a$. *Ponadto* $\sqrt[n]{a}$ *dla* $a \geqslant 0$ *jest* ***jedyną*** *taką liczbą nieujemną, której* n*-ta potęga to* a.

B.D.

Uwaga. Dodatkowo przyjmujemy, że jeśli $n \in \mathbb{N}$ oraz n jest nieparzyste i $c < 0$, to $\sqrt[n]{c} := -\sqrt[n]{-c}$. Oczywiście wówczas także $(\sqrt[n]{c})^n = c$.

Zgodnie ze znanym zwyczajem często poszemy $\sqrt{a}$ zamiast $\sqrt[2]{a}$.

A oto ważny wynik dotyczący niewymierności pierwiastków.

Twierdzenie I.10. *Niech* $m, n \in \mathbb{N}$. *Jeżeli* $\sqrt[n]{m} \notin \mathbb{N}$, *to* $\sqrt[n]{m} \notin \mathbb{Q}$. **B.D.**

Wniosek. $\mathbb{Q} \neq \mathbb{R}$, *bo* $\sqrt[2]{2} \notin \mathbb{Q}$ *na mocy twierdzenia I.10.*

Wspomnijmy jeszcze, że pominięty przez nas dowód tw. I.9 nie jest bardzo trudny, ale wymaga więcej czasu. Zachęcam do samodzielnego udowodnienia. Przyda się m. in. następująca wygodna własność potęgi naturalnej.

Fakt (nierówność Bernoulli'ego). *Jeżeli* $a > -1$, $n \in \mathbb{N}_0$, *to*

$$(1+a)^n \geqslant 1 + na. \tag{I.1}$$

Dowód.

Przeprowadzimy tu dowód przez indukcję. W tym celu ustalmy najpierw dowolnie $a > -1$. Zauważmy, że (I.1) zachodzi przy $n = 1$ (jest nawet równość). Teraz załóżmy, że (I.1) zachodzi dla pewnego n. Mnożąc obie strony nierówności przez $(1 + a)$ (a ściślej — korzystając z aksjomatu (MP.)) otrzymujemy

$$(1+a)^{n+1} = (1+a)^n(1+a) \geqslant (1+na)(1+a) = 1 + na + a + na^2 \geqslant 1 + (n+1)a.$$

W efekcie uzyskujemy więc odpowiednik nierówności (I.1) dla $n+1$ zamiast dla n (oczywiście powyżej użyliśmy w rzeczywistości także wielu innych aksjomatów, nie jedynie (MP.) — jakich?). W takim momencie na ogół zwyczajowo kończy się dowód indukcyjny. Jednak pytanie: gdzie tu ZIZ?? Aby więc było całkiem ściśle, tym razem dokładnie to wyjaśnimy. Mianowicie niech $A := \{n \in \mathbb{N} :\ \text{zachodzi (I.1)}\}$. To co dotychczas wykazaliśmy oznacza „w języku zbioru A", że $1 \in A$ oraz, że dla każdego $n \in \mathbb{N}$ jeśli $n \in A$, to $n + 1 \in A$. A zatem na mocy ZIZ otrzymujemy $A = \mathbb{N}$, a to oznacza właśnie, że (I.1) zachodzi dla każdego $n \in \mathbb{N}$. By całość zakończyć wystarczy więc jeszcze sprawdzić, że (I.1) zachodzi dla $n = 0$ (bo $0 \notin \mathbb{N}$, ale $0 \in \mathbb{N}_0$), co jest oczywiste. □

Twierdzenie I.10 można z kolei wykazać w oparciu o teorię podzielności, na którą jednak niestety czasu nam brak.

Etap 4. x^q dla $x > 0$, $q \in \mathbb{Q}$.

Potrzebny nam będzie

Lemat. *Jeżeli* $x > 0$ *oraz* $n, n' \in \mathbb{N}, m, m' \in \mathbb{Z}$ *spełniają* $\frac{m}{n} = \frac{m'}{n'}$, *to* $\sqrt[n]{x^m} = \sqrt[n']{x^{m'}}$.

Dowód.

W tym dowodzie przyda się

Lemacik. *Jeżeli* $a, b > 0$ *oraz* $N \in \mathbb{N}$, *to* $a^N = b^N \Leftrightarrow a = b$.

Prosty dowód lemaciku zostawiam Państwu. By wykazać tezę lematu, wystarczy sprawdzić „równość po podniesieniu do potęgi $N = n \cdot n'$", która na mocy twierdzenia I.9 i faktu I.1 pkt. 2. równoważna jest $x^{m \cdot n'} = x^{m' \cdot n}$ — co zachodzi z założenia. □

Przyjmujemy następującą definicję:

Definicja. *Dla* $x > 0$ *oraz* $q \in \mathbb{Q}$

$$x^q := \sqrt[n]{x^m},$$

gdzie $n \in \mathbb{N}$, $m \in \mathbb{Z}$ *są takie, że* $q = \frac{m}{n}$.

Ta definicja jest poprawna dzięki powyższemu lematowi, gdyż gwarantuje on, że wartość $\sqrt[n]{x^m}$ nie zależy od wyboru n i m spełniających $\frac{m}{n} = q$. Zauważmy też, że dla $q \in \mathbb{Z}$ ta definicja pokrywa się z def. z etapu 2.

Etap 5. *x^y dla $x > 0$, $y \in \mathbb{R}$.*

Definicja.

1. *Dla $x \geqslant 1$, $y \in \mathbb{R}$ $\quad x^y := \sup \{x^q : q \in \mathbb{Q}, q \leqslant y\}$;*
2. *dla $0 < x < 1$ i $y \in \mathbb{R}$, korzystając z 1. mamy zdefiniowane $(\frac{1}{x})^y$ i definiujemy $x^y := \frac{1}{(\frac{1}{x})^y}$.*

Nietrudno wykazać, że powyższa definicja jest poprawna, tj. że zbiór, którego kres pojawia się w 1. jest ograniczony z góry. Łatwo też wykazać, że dla $y \in \mathbb{Q}$ tak zdefiniowana potęga pokrywa się z tą z poprzedniego etapu. Jednak tak naprawdę żmudna i nietrywialna praca, to wykazanie, że tak zdefiniowana potęga rzeczywista posiada wszelkie „potrzebne" własności. Z braku czasu poniższy fakt podajemy znów bez dowodu.

Fakt („algebraiczne" własności potęgowania). *Dla $a, b > 0$ oraz $x, y \in \mathbb{R}$ zachodzi:*

1. $a^{x+y} = a^x \cdot a^y$,
2. $a^{x \cdot y} = (a^x)^y$,
3. $(a \cdot b)^x = a^x \cdot b^x$.

Przyjmujemy następującą terminologię dotyczącą funkcji określonych na podzbiorach $\mathbb{R}$ o wartościach w $\mathbb{R}$. Niech $f : X \longrightarrow \mathbb{R}$, gdzie $X \subset \mathbb{R}$.

Definicja. *Funkcja f jest*

- ***dodatnia*** *wtw* $\forall_{x \in X}\ f(x) > 0$;
- ***nieujemna*** *wtw* $\forall_{x \in X}\ f(x) \geqslant 0$;
- ***rosnąca*** *wtw* $\forall_{x,y \in X}\ (x \leqslant y \Rightarrow f(x) \leqslant f(y))$;
- ***ściśle rosnąca*** *wtw* $\forall_{x,y \in X}\ (x < y \Rightarrow f(x) < f(y))$

W analogiczny sposób (proszę samodzielnie wypisać...) określa się, że f jest ***ujemna****,* ***niedodatnia****,* ***malejąca****,* ***ściśle malejąca*** [17] *. Ponadto f jest*

- ***monotoniczna*** *wtw f jest rosnąca lub malejąca;*
- ***ściśle monotoniczna*** *wtw f jest ściśle rosnąca lub ściśle malejąca.*

Wnioski.

(i) (o funkcji wykładniczej) *Niech $a > 0$. Funkcja wykładnicza o podstawie a, tj. funkcja $W_a : \mathbb{R} \to \mathbb{R}$ zadana dla $x \in \mathbb{R}$ wzorem $W_a(x) = a^x$ jest dodatnia. Dla $a > 1$ $\quad W_a$ jest ściśle rosnąca, a dla $a < 1$ ściśle malejąca.*

(ii) (o funkcji potęgowej) *Niech $\alpha \in \mathbb{R}$. Funkcja potęgowa o wykładniku α, tj. funkcja $P_\alpha : (0; +\infty) \to \mathbb{R}$ zadana wzorem $P_\alpha(x) = x^\alpha$ dla $x > 0$ jest dodatnia. Dla $\alpha > 0$ funkcja P_α jest ściśle rosnąca, a dla $\alpha < 0$ ściśle malejąca.*

[17] Ale proszę o ostrożność! Niektórzy stosują inną terminologię, w której „nasza" rosnąca nazywa się *niemalejąca*, a „nasza" ściśle rosnąca nazywa się rosnąca (i analogicznie dla malejącej i ściśle malejącej)... Proszę jednak trzymać się terminologii tu przyjętej.

Dowód.
Dla $x \geqslant 1$, $q \in \mathbb{Q}$ zachodzi $x^q > 0$ (patrz etapy 1 — 4), stąd kres górny z punktu 1. definicji w etapie 5 jest dodatni, czyli $x^y > 0$. Zatem dla $x \in (0;1)$ także $x^y > 0$ na mocy 2. definicji. Stąd dodatniość obu funkcji W_a i P_α.

Teraz zajmiemy się ścisłą monotonicznością dla W_a. Niech $a > 1$ oraz $x < y$. Na mocy dodatniości oraz powyższego faktu (pkt. 1.) zachodzi

$$\frac{a^y}{a^x} = a^{y-x} = \sup A,$$

gdzie $A = \{a^q : q \in \mathbb{Q}, q \leqslant y - x\}$. Ponieważ $y - x > 0$, więc korzystając z tw. I.8 (o gęstości $\mathbb{Q}$) wybierzemy $q_0 \in \mathbb{Q}$ takie, że $\frac{y-x}{2} \leqslant q_0 \leqslant y - x$. A zatem $q_0 > 0$ oraz $a^{q_0} \in A$, więc $\sup A \geqslant a^{q_0}$. Jednak z definicji potęgi dla wykładników wymiernych (etapy 1 — 4) z faktu, że $q_0 > 0$ i $a > 1$ dostajemy łatwo[18] , że $a^{q_0} > 1$, czyli w efekcie $\sup A > 1$, skąd $a^y > a^x$. Dla $a < 1$ — dowód łatwy z punktu 2. definicji i z powyższego już wykazanego. Dowód ścisłej monotoniczności dla P_α — analogiczny, ale zamiast punktu 1. pow. faktu należy użyć punktu 3. $\square$

[18] Zachęcam do ścisłego wykazania tego przy użyciu podanych definicji.

Zadania do Rozdziału I

1. Wykaż następujące tożsamości i nierówności:

   $\forall$ (a) $\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}$ dla $n \in \mathbb{N}$ oraz $q \in \mathbb{R}\setminus\{1\}$;

   (b) $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$ dla $n \in \mathbb{N}_0$ oraz $a, b \in \mathbb{R}$ (wzór Newtona, fakt str. 12);

   $\forall$ (c) $|a| + |b| \geqslant |a - b| \geqslant ||a| - |b||$ dla $a, b \in \mathbb{R}$;

   $\forall$ (d) $|\sum_{k=1}^{n} x_k| \leqslant \sum_{k=1}^{n} |x_k|$ dla $x_1, \ldots, x_n \in \mathbb{R}$;

   (e) $\sum_{k=1}^{n} \frac{1}{\sqrt{k}} > \sqrt{n}$ dla $n \in \mathbb{N}, n \geqslant 2$;

   $\forall$ (f) $n! \leqslant (\frac{n+1}{2})^n$ dla $n \in \mathbb{N}$;

   Uwaga: w a), b) przyjmujemy $0^0 = 1$.

2. Wykaż, że

   (a) $\forall_{q>1}\ \forall_{k\in\mathbb{N}}\ \exists_{c>0}\ \forall_{n\in\mathbb{N}}\ q^n \geqslant cn^k$;

   (b) $\forall_{q>1}\ \forall_{\alpha>0}\ \exists_{c>0}\ \forall_{x\geqslant 1}\ q^x \geqslant cx^\alpha$.

3. Niech p_n oznacza n-tą z kolei liczbę pierwszą. Wykaż, że $\forall_{n\in\mathbb{N},\ n\geqslant 12}\ p_n \geqslant 3n$.

   Uwaga: tu można użyć wiedzy „szkolnej”, a nie tylko tej z wykładu. Np. zakładam, że każdy student orientuje się co to jest liczba pierwsza (a zatem $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7$, $p_5 = 11$ itd).

$\forall$ 4. Wykaż w sposób całkowicie ścisły (wskazując na każdym kroku rozumowania z jakiego aksjomatu lub uprzednio wykazanego twierdzenia należy skorzystać) **kilka** elementarnych własności liczb rzeczywistych — **np.**: te poniższe:

   (a) $\forall_{a\in\mathbb{R}}\ a \cdot 0 = 0$;

   (b) $\forall_{a\in\mathbb{R}}\ (-1) \cdot a = -a$;

   (c) $\forall_{a,b\in\mathbb{R}}\ -(a+b) = -a - b$;

   (d) $\forall_{a\in\mathbb{R}}\ a^2 \geqslant 0$;

   lub inne wskazane własnoręcznie (nie z tych wykazanych na wykładzie).

5. Niech $A, B \subset \mathbb{R}$ będą niepuste i ograniczone. Czy istnieje wzór wyrażający:

   (a) $\sup(A \cup B)$,

   (b) $\inf(A \cup B)$,

   (c) $\sup(A \cap B)$,

   (d) $\inf(A \cap B)$

   przy pomocy kresów zbiorów A i B? Jeśli tak, to znajdź taki wzór(i udowodnij), a jeśli nie, to wykaż, że nie istnieje.

$\forall$ 6. Niech I będzie pewnym niepustym zbiorem („indeksów”) oraz dla dowolnego $i \in I$ niech $A_i \subset \mathbb{R}$ będzie niepusty i ograniczony z góry. Udowodnij, że $\sup(\bigcup_{i\in I} A_i) = \sup\{\sup A_i : i \in I\}$.

Uwaga: proszę to zrobić przynajmniej przy dodatkowym założeniu, że zbiór z prawej strony jest ograniczony z góry; bez tego założenia istotna staje się umowa o „$+\infty$” z wykładu.

Dla $A, B \subset \mathbb{R}$ określamy działania algebraiczne (na zbiorach) $+$ i $\cdot$ następująco:
$A + B := \{a + b \in \mathbb{R} : a \in A, b \in B\}$, $A \cdot B := \{a \cdot b \in \mathbb{R} : a \in A, b \in B\}$.
Oznaczmy też:

- $-A := \{-1\} \cdot A$,
- $A \leqslant B$ wtw $\forall_{a \in A,\ b \in B}\ a \leqslant b$

i gdy $c \in \mathbb{R}$

- $c \leqslant A$ wtw $\forall_{a \in A}\ c \leqslant a$ i analogicznie $c < A$, $c > A$, $c \geqslant A$,
- $c + A := \{c\} + A$.

7. Wykaż, że jeśli $A, B \subset \mathbb{R}$ są niepuste i ograniczone z góry, to:
   (a) $\sup(A + B) = \sup(A) + \sup(B)$;
   (b) $\sup(A \cdot B) = \sup A \cdot \sup(B)$ przy dodatkowym założeniu, że $A, B > 0$;
   (c) $\inf(-A) = -\sup(A)$.

   Uwaga: w dowodzie pkt. a) można np. wykorzystać wynik z zadania I.6 oraz szczególną wersję pkt. a) dla $A = \{a\}$. Dla b) — analogicznie.

8. Wykaż, że jeśli $\emptyset \neq A \subset \mathbb{R}$ oraz $\inf A = \sup A$, to A jest zbiorem jednoelementowym.

9. Wykaż, że jeśli A, B są niepustymi podzbiorami $\mathbb{R}$, to $A \leqslant B \Rightarrow \sup A \leqslant \inf B$.

10. Znajdź oba kresy zbiorów:
    (a) $\{a^2 - ab : a, b \in (0; 1)\}$;
    (b) $\{|\frac{1}{n} - \frac{1}{m}| : n, m \in \mathbb{N}, n \neq m\}$;
    $\forall$ (c) $\{\frac{n-k}{n+k} : n, k \in \mathbb{N}\}$.

11. Znajdź dowód (pominięty na wykładzie) dla szczególnego przypadku $n = 2$ w twierdzeniu I.9., tj. wykaż, że $(\sqrt[2]{a})^2 = a$ dla $a \geqslant 0$.

II Ciągi liczbowe, granica

[około $2\frac{1}{2}$ wykładu]

1. Podstawowe pojęcia i oznaczenia

Niech $n_0 \in \mathbb{Z}$. *Ciągiem (indeksowanym od n_0)* nazywamy funkcję określoną na $\mathbb{N}_{n_0}$. Jej wartości nazywamy *wyrazami ciągu*. Gdy wyrazy są liczbami rzeczywistymi, mówimy o *ciągu liczbowym* (ew. *rzeczywistym*)[19] .

Najczęściej będziemy mieli do czynienia z sytuacją, gdy *indeks początkowy* n_0 równy jest 0 lub 1. Ciąg będziemy oznaczali jedną literą, np. a lub tak: $\{a_n\}_{n \geqslant n_0}$ — ta druga możliwość pozwala na wyraźne zaznaczenie początkowego indeksu. Choć czasem (np. przy większym pośpiechu...) może skrócimy to nieco do $\{a_n\}$. Aby jednak nie przypominać przy każdej okazji jaki jest indeks początkowy rozważanych ciągów przyjmijmy, że na ogół (przynajmniej w tym rozdziale, choć nie tylko) będzie to właśnie n_0. Indeks „ogólny" dla ciągu będziemy najczęściej oznaczali przez n, ale nie zawsze — oczywiście zamiast $\{a_n\}_{n \geqslant n_0}$ możemy równie dobrze pisać np. $\{a_k\}_{k \geqslant n_0}$.

Aby wyraźnie podkreślić zasadniczą różnicę pomiędzy n-tym wyrazem a_n ciągu $a = \{a_k\}_{k \geqslant n_0}$ a „całym" ciągiem a, będziemy na wykładzie unikali często stosowanego żargonowego sformułowania „... ciąg a_n", pisząc zamiast tego „... ciąg $\{a_n\}_{n \geqslant n_0}$" lub „... ciąg a".

Reasumując, ciąg liczbowy a to (dla pewnego $n_0 \in \mathbb{Z}$) funkcja $a\colon \mathbb{N}_{n_0} \to \mathbb{R}$. Inaczej tylko niż przy typowym zapisie dla funkcji oznaczamy wartość tej funkcji w punkcie $n \in \mathbb{N}_{n_0}$, czyli n-ty wyraz ciągu a. Piszemy bowiem a_n zamiast $a(n)$, choć ten drugi zapis też jest czasem stosowany.

Dla ciągów, podobnie jak ogólnie dla funkcji, określa się działania dodawania, mnożenia, odejmowania i dzielenia przez liczbę. Oznacza się je tymi samymi symbolami, co odpowiednie działania dla liczb, choć formalnie są to oczywiście zupełnie inne działania. Np. gdy $r \in \mathbb{R}$ oraz a, b są ciągami liczbowymi o tym samym indeksie początkowym n_0, to $(r \cdot a)_n := r \cdot a_n$, $(a + b)_n := a_n + b_n$ dla $n \geqslant n_0$. Analogicznie („punktowo") określamy pozostałe działania, przy czym dzielić „wolno" tylko przez ciąg o wszystkich wyrazach $\neq 0$.

Będziemy też używali symbolu nierówności $\leqslant, \geqslant, <, >$ dla ciągów (znów, to małe nadużycie...) np. $a \leqslant b$ wtw $\forall_{n \geqslant n_0}\ a_n \leqslant b_n$ [20] oraz $r \leqslant a$ wtw $\forall_{n \geqslant n_0}\ r \leqslant a_n$ i analogicznie przy pozostałych typach nierówności.

Ciąg a jest *ograniczony z góry (z dołu)* wtw $a \leqslant r$ ($r \leqslant a$) dla pewnego $r \in \mathbb{R}$; a jest *ograniczony* tzn., że jest ograniczony z góry i z dołu.

Ważna klasa ciągów to *ciągi monotoniczne*, tj. *rosnące* i *malejące* (nie jednocześnie...). Obowiązuje tu terminologia wprowadzona już w rozdziale I ogólnie dla wszystkich funkcji. Dla takich szczególnych funkcji, jakimi są ciągi, warunki z odpowiednich definicji można zapisać prościej. A więc np. $\{a_n\}_{n \geqslant n_0}$ jest rosnący wtw $\forall_{n \geqslant n_0}\ a_{n+1} \geqslant a_n$ (przypominam, że nierówność jest „$\geqslant$", nie „$>$", a ta ostra pojawia się w definicji ciągu ściśle rosnącego).

Przy okazji teorii zbieżności ciągów przydatna bywa poniższa terminologia:

- gdy φ jest pewną formułą zdaniową ze zmienną ze zbioru $\mathbb{Z}$ lub z pewnego $\mathbb{N}_k$, to zdanie: *$\varphi(n)$ dla dostatecznie dużych n* oznacza to samo co

$$\exists_{N \in \mathbb{Z}}\ \forall_{n \geqslant N}\ \varphi(n).$$

[19] Na tym wykładzie liczby są w zasadzie tylko rzeczywiste, ale ogólniej ciągi liczbowe mogą mieć wyrazy będące dowolnymi liczbami zespolonymi.

[20] Formalnie powinniśmy napisać tu „$\forall_{n \in \mathbb{N}_{n_0}}$" lub „$\forall_{n \in \mathbb{Z}, n \geqslant n_0}$", ale taki skrótowy zapis „$\forall_{n \geqslant n_0}$" z domyślnym wyborem $n \in \mathbb{Z}$, a nie wszystkich $n \in \mathbb{R}$, stosować będziemy często.

Dla skrócenia będziemy pisać: d.d.d. zamiast „dla dostatecznie dużych”.

- gdy W jest pewną własnością ciągu lub ciągów, to mówimy, że W zachodzi *od pewnego miejsca*, gdy istnieje $N \in \mathbb{Z}$ takie, że W zachodzi dla tych ciągów „obciętych” do $\mathbb{N}_N$ (tzn. w miejsce ciągu $\{a_n\}_{n \geqslant n_0}$ bierzemy $\{a_n\}_{n \geqslant N}$).

Np. Ciąg a jest rosnący od pewnego miejsca wtw $\exists_{N \in \mathbb{Z}}\ \{a_n\}_{n \geqslant N}$ jest rosnący wtw

$$\exists_{N \in \mathbb{Z}}\ \forall_{n \geqslant N}\ a_{n+1} \geqslant a_n$$

wtw d.d.d. n $a_{n+1} \geqslant a_n$. Jednak **nie** powinniśmy używać sformułowań „a jest rosnący d.d.d. n” **ani** „a_n jest rosnący d.d.d. n” (dlaczego?).

Zajmijmy się wreszcie najważniejszym tu dla nas pojęciem granicy ciągu liczbowego $a = \{a_n\}_{n \geqslant n_0}$.

Definicja.

- *Niech $g \in \mathbb{R}$. Ciąg a jest* ***zbieżny do*** *g wtw*

$$\forall_{\epsilon > 0}\ \exists_{N \geqslant n_0}\ \forall_{n \geqslant N}\ |a_n - g| < \epsilon$$

  (tj. $\forall_{\epsilon > 0}$ ($|a_n - g| < \epsilon$ d.d.d. n)).

- *Ciąg a jest* ***zbieżny*** *wtw a jest zbieżny do g dla pewnego $g \in \mathbb{R}$.*

- *Ciąg a jest* ***rozbieżny*** *wtw a nie jest zbieżny.*

- *Ciąg a jest* ***rozbieżny do*** *$+\infty$* ***($-\infty$)*** *wtw $\forall_{C \in \mathbb{R}}\ \exists_{N \geqslant n_0}\ \forall_{n \geqslant N}\ a_n > C$ ($a_n < C$).*

- *Niech $g \in \mathbb{R} \cup \{+\infty, -\infty\} =: \overline{\mathbb{R}}$. Ciąg a* ***ma granicę*** *g lub inaczej: g* ***jest granicą*** *a wtw [$g \in \mathbb{R}$ i a jest zbieżny do g] lub [$g = \pm\infty$ i a jest rozbieżny do g]. Zapisujemy to symbolem $a_n \to g$, ewentualnie $a_n \xrightarrow[n \to +\infty]{} g$, a samą granicę ciągu a oznaczamy $\lim_{n \to +\infty} a_n$.*

Uwaga 1. Nie możemy więc powiedzieć „a jest zbieżny do $+\infty$”, ale możemy „a ma granicę $+\infty$”, co znaczy tyle co „a jest rozbieżny do $+\infty$”.

Uwaga 2. Wbrew tej tradycyjnej notacji, jak widać z definicji, granica i zbieżność są związane z „całym” ciągiem, a nie z „jakimś jego n-tym wyrazem”. Może lepsza byłaby więc notacja „$\lim a$”, „$a \to g$”, ale to wbrew tradycji...

Fakt II.1 (oczywisty). *Jeśli a posiada granicę, to jest ona wyznaczona jednoznacznie.* **B.D.**

A zatem symbol $\lim\limits_{n \to +\infty} a_n$ jest poprawnie określony w przypadku ciągu posiadającego granicę.

Fakt II.2 (też oczywisty). *Jeżeli $a_n = b_n$ d.d.d. n, to $a_n \to g \Leftrightarrow b_n \to g$.* **B.D.**

Przykłady (najbardziej elementarne).

1. Ciąg stały. Oznaczmy: $a \equiv r$ wtw $\forall_{n \geqslant n_0}\ a_n = r$. Oczywiście gdy $a \equiv r$, to $a_n \to r$.

2. $n \xrightarrow[n \to +\infty]{} +\infty$ — to natychmiastowy wniosek z zasady Archimedesa (tw. I.5).

3. $\frac{1}{n} \to 0$ — to jak wyżej (to „klasyczny przykład szkolny”).

4. $(-1)^n$ — ciąg o wyrazach zadanych takim wzorem jest rozbieżny i (co gorsza...) nie ma granicy żadnej.

Dalsze „elementarne” przykłady wygodniej będzie badać po rozwinięciu choć trochę teorii zbieżności.

2. Własności arytmetyczne granicy

Zaczniemy od twierdzenia o zachowaniu się granicy przy dokonywaniu podstawowych operacji algebraicznych na ciągach. Najpierw jednak częściowo[21] rozszerzymy działania określone w $\mathbb{R}$ na tzw. *rozszerzony zbiór liczb rzeczywistych* $\overline{\mathbb{R}}$, który został zdefiniowany przy okazji definicji granicy ciągu. Rozszerzenia te zdefiniowane są w następujących tabelach (wiersz odpowiada lewemu argumentowi a kolumna prawemu). Znak „$\times$” oznacza, że dane działanie nie jest *określone*. O a i b zakładamy, że należą do $\mathbb{R}$.

$+$	b	$+\infty$	$-\infty$
a	$a+b$	$+\infty$	$-\infty$
$+\infty$	$+\infty$	$+\infty$	$\times$
$-\infty$	$-\infty$	$\times$	$-\infty$

$-$	b	$+\infty$	$-\infty$
a	$a-b$	$-\infty$	$+\infty$
$+\infty$	$+\infty$	$\times$	$+\infty$
$-\infty$	$-\infty$	$-\infty$	$\times$

$\cdot$	b	$+\infty$	$-\infty$
a	$a \cdot b$	$+\infty \quad a>0$ $\times \quad a=0$ $-\infty \quad a<0$	$-\infty \quad a>0$ $\times \quad a=0$ $+\infty \quad a<0$
$+\infty$	$+\infty \quad b>0$ $\times \quad b=0$ $-\infty \quad b<0$	$+\infty$	$-\infty$
$-\infty$	$-\infty \quad b>0$ $\times \quad b=0$ $+\infty \quad b<0$	$-\infty$	$+\infty$

$:$	b	$+\infty$	$-\infty$
a	$a:b \quad b\neq 0$ $\times \quad b=0$	0	0
$+\infty$	$+\infty \quad b>0$ $\times \quad b=0$ $-\infty \quad b<0$	$\times$	$\times$
$-\infty$	$-\infty \quad b>0$ $\times \quad b=0$ $+\infty \quad b<0$	$\times$	$\times$

Możemy też rozszerzyć działania jednoargumentowe (elementy przeciwny i odwrotny):

z	$-z$
a	$-a$
$+\infty$	$-\infty$
$-\infty$	$+\infty$

z	z^{-1}
a	$a^{-1} \quad a\neq 0$ $\times \quad a=0$
$+\infty$	0
$-\infty$	0

choć w poniższym twierdzeniu nie będzie nam to potrzebne.

Twierdzenie II.1 (o rachunkowych własnościach granicy). *Niech $\circledast$ oznacza jedno z działań $+$, $-$, $\cdot$, $:$. Załóżmy, że $a_n \to g$, $b_n \to h$, gdzie $g, h \in \overline{\mathbb{R}}$ są takie, że działanie $g \circledast h$ jest określone*[22] *oraz, w przypadku gdy $\circledast$ jest dzieleniem, że wszystkie wyrazy ciągu $\{b_n\}$ są różne od 0. Wówczas $(a \circledast b)_n \to g \circledast h$.*

Ponadto jeśli $\{a_n\} > 0$ oraz $a_n \to 0$, to $\frac{1}{a_n} \to +\infty$.

Udowodnimy tylko jeden z przypadków: gdy $\circledast$ jest mnożeniem oraz $g, h \in \mathbb{R}$. Najpierw jednak:

Lemat. *Ciąg zbieżny jest ograniczony.*

Dowód.

[21] Tylko częściowo, bo można wykazać, że całkiem się nie da, jeśli chcielibyśmy przy tym rozszerzeniu zachować np. aksjomaty ciała.

[22] Tzn. w tabeli dla $\circledast$ w kratce odpowiadającej parze (g, h) **nie** występuje „$\times$”.

Niech $a_n \to g \in \mathbb{R}$. Z warunku z definicji „dla $\epsilon = 1$" dostajemy: istnieje $N \geqslant n_0$ takie, że $\forall_{n\geqslant N}\ |a_n - g| < 1$ czyli $\forall_{n\geqslant N}\ g-1 < a_n < g+1$. Stąd, jeśli oznaczymy $A := \{a_k\colon k \in \mathbb{Z}, n_0 \leqslant k \leqslant N\}$, to

$$\forall_{n\geqslant n_0}\ \min(\{g-1\} \cup A) \leqslant a_n \leqslant \max(\{g+1\} \cup A).$$

□

Dowód (części twierdzenia II.1).
Niech $g, h \in \mathbb{R}$. Na mocy lematu ciąg a jest ograniczony, czyli[23] dla pewnego $C > 0$ zachodzi $\forall_{n\geqslant n_0}\ |a_n| \leqslant C$. A zatem na mocy własności modułu (w tym nierówności trójkąta) mamy dla $n \geqslant n_0$:

$$\begin{aligned} |a_nb_n - gh| &= |a_nb_n - a_nh + a_nh - gh| \leqslant |a_n| \cdot |b_n - h| + |a_n - g| \cdot |h| \leqslant \\ &\leqslant C \cdot |b_n - h| + |a_n - g| \cdot |h|. \end{aligned} \tag{II.1}$$

Teraz, aby posługując się definicją wykazać, że $a_nb_n \to gh$, rozważmy dowolne $\epsilon > 0$. Na mocy definicji istnieją takie $N_a, N_b \in \mathbb{Z}$, że

$$|a_n - g| < \frac{\epsilon}{C + |h|} \text{ dla } n \geqslant N_a,$$

$$|b_n - h| < \frac{\epsilon}{C + |h|} \text{ dla } n \geqslant N_b.$$

Biorąc więc $N := \max\{N_a, N_b\}$ mamy

$$\forall_{n\geqslant N}\ |a_nb_n - gh| < (C + |h|)\frac{\epsilon}{C + |h|} = \epsilon.$$

□

Warto zwrócić baczną uwagę na powyższy dowód, choć jest to dowód faktu znanego Państwu z pewnością ze szkoły. Zawiera on dwa elementy dość typowe dla wielu dowodów, które będziemy przeprowadzać. Pierwszy to chwyt użyty w formule (II.1) polegający na odjęciu i dodaniu tej samej liczby (a_nh) przed użyciem nierówności trójkąta. Drugi to dobór N jako większego spośród N_a i N_b.

Użyteczność twierdzenia II.1 dla znajdowania granic rozmaitych ciągów zadanych zawiłymi wzorami wydaje się oczywista. Czasem jest ono nawet skuteczniejsze niż mogłoby się to wydawać na pierwszy rzut oka.

Przykład (szkolny).

$$a_n = \frac{n^2 - 7n + 3}{2n^2 + 3n + 1}.$$

Z twierdzenia II.1 widać, że „mianownik" ma granicę $+\infty$, ale nie daje ono nic dla licznika... („$+\infty - (+\infty)$"). Nie możemy też więc użyć bezpośrednio tw. II.1 dla ilorazu... Stosując jednak standardowy chwyt z dzieleniem licznika i mianownika przez n^2 przekształcamy a_n do postaci:

$$a_n = \frac{1 - \frac{7}{n} + \frac{3}{n^2}}{2 + \frac{3}{n} + \frac{1}{n^2}},$$

skąd już $a_n \to \frac{1}{2}$ na mocy tw. II.1.

Jednak aby naprawdę móc skutecznie wykorzystać tw. II.1 potrzeba nam więcej zbadanych granic dla elementarnych przykładów ciągów — te kilka z przykładu ze str. 19 to z pewnością zbyt mało. Poniżej podamy więcej przykładów. Część z nich zostanie zbadana wkrótce na wykładzie, część na ćwiczeniach, a część zostanie bez dowodu (do ewentualnego samodzielnego sprawdzenia).

[23] Zachęcam do samodzielnego wykazania, że „ograniczoność to to samo, co ograniczoność modułu z góry".

Przykłady (też elementarne).

a. $n^\alpha \to +\infty$ dla $\alpha > 0$;

b. $a^n \to \begin{cases} 0 & \text{dla } |a| < 1 \\ 1 & \text{dla } a = 1 \\ +\infty & \text{dla } a > 1, \end{cases}$ ponadto $\{a^n\}$ nie posiada granicy gdy $a \leqslant -1$;

c. $\frac{n^\alpha}{C^n} \to 0$ przy dowolnych $\alpha \in \mathbb{R}$ oraz $C > 1$;

d. $\frac{a^n}{n!} \to 0$ przy dowolnym $a \in \mathbb{R}$;

e. $\sqrt[n]{a} \to 1$ dla $a > 0$;

f. $\sqrt[n]{n} \to 1$;

g. $(1+\frac{1}{n})^n \to e$ dla pewnego $e \in (2;3]$. Uwaga: to właśnie przyjmujemy jako definicję liczby e (tj. „e — to granica ciągu...”).

Wykazanie, że zachodzą zbieżności/rozbieżności z powyższego przykładu przy użyciu samej tylko definicji granicy byłoby zadaniem dość trudnym. Nieco prościej można sobie z nimi poradzić przy użyciu kilku użytecznych twierdzeń, które wkrótce Państwo poznacie. Wcześniej jednak ostrzeżenie związane z ostatnim przykładem. Poniższe „rozumowanie” „oparte” na tw. II.1 jest dość częste i typowe dla wielu **nieostrożnych** studentów:

$$„\left(1+\frac{1}{n}\right)^n = \underbrace{\left(1+\frac{1}{n}\right)\cdot\ldots\cdot\left(1+\frac{1}{n}\right)}_{n \text{ razy}} \to \underbrace{1\cdot\ldots\cdot 1}_{n \text{ razy}} = 1.”$$

Jednak $1 \neq e$ ... — gdzie tkwi błąd?

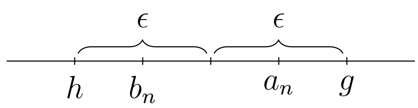
3. Granica a nierówności

Zacznijmy od następującego ważnego, choć prostego twierdzenia:

Twierdzenie II.2 (o zachowaniu nierówności przy przejściu granicznym). *Jeżeli $a_n \leqslant b_n$ d.d.d. n, oraz $a_n \to g$, $b_n \to h$, to $g \leqslant h$.*

Oczywiście przyjmujemy umowę, że $s \leqslant +\infty$ oraz $-\infty \leqslant s$ dla dowolnego $s \in \overline{\mathbb{R}}$. Przyjmujemy także: $-\infty < +\infty$ oraz $-\infty < s < +\infty$ dla dowolnego $s \in \mathbb{R}$.

Dowód (twierdzenia II.2).
Twierdzenie wykażemy tylko dla przypadku $g, h \in \mathbb{R}$. Przypuśćmy, że $h < g$ i niech $\epsilon := \frac{g-h}{2}$.



Rysunek 2: Tak np. mogłoby być, gdyby $h < g$

Wówczas, ponieważ $\epsilon > 0$, d.d.d. n zachodzi (odpowiednie „N” dobieramy najpierw dla ciągu a, potem dla b, a następnie bierzemy większe z nich — jak w dowodzie tw. II.1):

$$b_n < g - \epsilon = h + \epsilon < a_n,$$

co jest z kolei sprzeczne z założeniem, że $a_n \leqslant b_n$ d.d.d. n. (patrz rys. 2) $\square$

Zauważmy, że nie można w pow. twierdzeniu zmienić w obu miejscach „$\leqslant$” na „$<$”. Wystarczy np. rozważyć przykład $a_n := 1$ oraz $b_n := 1 + \frac{1}{n}$.

Kolejne twierdzenie bywa bardzo użyteczne przy badaniu granic wielu ciągów, dla których użycie twierdzenia o rachunkowych własnościach granicy wydaje się niewystarczające.

Twierdzenie II.3 (o trzech ciągach). *Jeżeli*

$$a_n \leqslant b_n \leqslant c_n$$

d.d.d. n*, oraz* $\lim_{n\to+\infty} a_n = \lim_{n\to+\infty} c_n = g$*, to* $\lim_{n\to+\infty} b_n = g$.

Zanim przystąpimy do dowodu zauważmy, że gdybyśmy już wiedzieli, że ciąg środkowy $\{b_n\}$ posiada granicę, to dowód mielibyśmy natychmiast dzięki tw. II.2. Jednak tego nie wiemy...

Dowód.
Znów rozważamy tylko przypadek $g \in \mathbb{R}$. Niech $\epsilon > 0$. Biorąc d.d.d. n mamy z definicji zbieżności oraz z nierówności z założenia twierdzenia:

$$g - \epsilon < a_n \leqslant b_n \leqslant c_n < g + \epsilon,$$

a zatem w szczególności $|b_n - g| < \epsilon$. □

Uwaga. W przypadku, gdy $g = +\infty$ lub $-\infty$ założenia można osłabić — wystarczy jedna nierówność d.d.d. n. Lewa dla $+\infty$, a prawa dla $-\infty$. Tak uproszczone twierdzenie nazywane bywa „twierdzeniem o dwóch ciągach”.

Wniosek. *Jeżeli ciąg* c *jest ograniczony oraz ciąg* a *zbieżny do* 0*, to ich iloczyn* $c \cdot a$ *jest zbieżny do* 0.

Dowód.
Z ograniczoności c mamy dla pewnego $M \in \mathbb{R}$

$$\forall_{n\geqslant n_0}\ |c_n| \leqslant M,$$

skąd

$$\forall_{n\geqslant n_0}\ 0 \leqslant |c_n a_n| \leqslant M \cdot |a_n|.$$

Jednak skoro $a_n \to 0$, zatem $|a_n| \to 0$, więc (z tw. II.1 dla ciągu stałego i $\{|a_n|\}$) $M \cdot |a_n| \to M \cdot 0 = 0$. To pozwala użyć tw. II.3 z „lewym” ciągiem stale równym 0. Tak więc $|c_n a_n| \to 0$, a stąd także $c_n a_n \to 0$. □

W powyższym dowodzie został użyty (dwukrotnie) następujący oczywisty i użyteczny

Fakcik. $a_n \to 0$ *wtw* $|a_n| \to 0$.

Dowód.
Patrz definicja zbieżności... □

Kolejne twierdzenie także związane z nierównością, ale już w całkiem inny sposób, to twierdzenie dotyczące ciągów monotonicznych.

Twierdzenie II.4 (o granicy ciągu monotonicznego). *Jeżeli ciąg* $a = \{a_n\}_{n\geqslant n_0}$ *jest monotoniczny, to posiada granicę. Jest ona równa* $\sup\{a_n : n \geqslant n_0\}$ *gdy* a *jest rosnący, natomiast gdy* a *jest malejący, to równa jest* $\inf\{a_n : n \geqslant n_0\}$ [24]. *W szczególności ciąg rosnący ograniczony z góry, a także ciąg malejący ograniczony z dołu jest zbieżny.*

Dowód.
Ograniczymy się do przypadku ciągu rosnącego ograniczonego z góry. Niech $g := \sup\{a_n : n \geqslant n_0\}$. Wykażemy, że $a_n \to g$. Niech $\epsilon > 0$. Ponieważ a jest ograniczony z góry, zatem $g \in \mathbb{R}$ (aksjomat ciągłości!), a stąd $g - \epsilon < g$. Z definicji sup liczba $g - \epsilon$ nie jest zatem ograniczeniem

[24] Przypominam o umowie z podrozdziału 3., dotyczącej kresów zbiorów nieograniczonych.

górnym zbioru $\{a_n : n \geqslant n_0\}$, czyli istnieje $N \geqslant n_0$ takie, że $a_N > g - \epsilon$. Jednak ponieważ a jest rosnący, zatem dla dowolnego $n \geqslant N$ mamy

$$a_n \geqslant a_N > g - \epsilon,$$

a jednocześnie $g + \epsilon > g \geqslant a_n$, skąd $|a_n - g| < \epsilon$. $\square$

Przykłady. Powyższych twierdzeń użyjemy do zbadania paru spośród przykładów a) — g) ze strony 22.

a) $n^\alpha \to +\infty$ dla $\alpha > 0$. Z zasady Archimedesa znajdziemy $k \in \mathbb{N}$ takie, że $k \geqslant \frac{1}{\alpha}$. Stąd $\frac{1}{k} \leqslant \alpha$ i z własności potęgi $\forall_{n\in\mathbb{N}}\ b_n := \sqrt[k]{n} = n^{\frac{1}{k}} \leqslant n^\alpha$. Na mocy twierdzenia o 2 ciągach wystarczy wykazać, że $b_n \to +\infty$. Ale znów z własności potęgi b jest ciągiem rosnącym, zatem b ma pewną granicę g i na dodatek, z tw. II.2, $g \geqslant 0$. Ponadto, z tw. II.1 dla mnożenia (plus indukcja „po k") mamy $b_n^k \to g^k$, ale $b_n^k = n \to +\infty$, stąd $g^k = +\infty$. Więc skoro $g \geqslant 0$, to $g = +\infty$.

c) $\frac{n^\alpha}{c^n} \to 0$ dla $\alpha \in \mathbb{R}$, $c > 1$. Wybierzmy jakieś $k \in \mathbb{N}$ takie, że $k > \alpha$. Zauważmy, że $\sqrt[k]{c} > 1$ zatem $r := \sqrt[k]{c} - 1 > 0$. Wyraz ogólny naszego ciągu zapiszemy w następujący sposób:

$$a_n = \frac{n^\alpha}{c^n} = \frac{n^\alpha}{((\sqrt[k]{c})^n)^k} = \frac{n^\alpha}{((1+r)^n)^k}.$$

Teraz z nierówności Bernoulli'ego mamy

$$0 \leqslant a_n \leqslant \frac{n^\alpha}{(1+nr)^k} \leqslant \frac{n^\alpha}{n^k r^k} = \frac{r^{-k}}{n^{k-\alpha}} =: c_n.$$

Ponieważ $k > \alpha$ zatem na mocy punktu a) oraz tw. II.1 (dla dzielenia) mamy $c_n \to 0$, skąd $a_n \to 0$ z tw. o trzech ciągach.

f) $\sqrt[n]{n} \to 1$. Wykażemy najpierw, że $c_n := \sqrt[n]{\sqrt[2]{n}} \to 1$. Z nierówności Bernoulli'ego otrzymujemy łatwo, że $\sqrt[n]{1+a} \leqslant 1 + \frac{a}{n}$ dla $a > -1$. Stąd mamy

$$1 \leqslant \sqrt[n]{\sqrt[2]{n}} = \sqrt[n]{1 + (\sqrt[2]{n} - 1)} \leqslant 1 + \frac{\sqrt[2]{n} - 1}{n} = 1 + \frac{1}{\sqrt[2]{n}} - \frac{1}{n},$$

więc $c_n \to 1$ na mocy tw. o 3 ciągach. Ale $\sqrt[n]{n} = c_n \cdot c_n \to 1 \cdot 1 = 1$.

e) $\sqrt[n]{a} \to 1$ dla $a > 0$. Gdy $a \geqslant 1$, to d.d.d. n mamy $1 \leqslant \sqrt[n]{a} \leqslant \sqrt[n]{n}$, więc wystarczy użyć f) i tw. o 3 ciągach. Teraz dla $0 < a < 1$ mamy $\sqrt[n]{a} = \frac{1}{\sqrt[n]{\frac{1}{a}}}$, więc wystarczy skorzystać z poprzedniego przypadku i z tw. II.1.

g) Wykażemy najpierw, że $\{a_n\}_{n\geqslant 1}$ o wyrazach zadanych wzorem $a_n := (1 + \frac{1}{n})^n$ jest rosnący (a nawet ściśle rosnący). Nierówność $a_{n+1} > a_n$ równoważna jest nierówności

$$\frac{1}{1 + \frac{1}{n+1}} < \left(\frac{1 + \frac{1}{n+1}}{1 + \frac{1}{n}}\right)^n,$$

czyli $1 - \frac{1}{n+2} < (1 - \frac{1}{n^2+2n+1})^n$, a tę ostatnią nierówność łatwo dowieść z nierówności Bernoulli'ego. Następnie dowodzimy, że $\{a_n\}$ jest ograniczony z góry przez 3 rozpisując a_n przy użyciu wzoru Newtona:

$$a_n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} \leqslant \sum_{k=0}^{n} \frac{1}{k!} \leqslant 1 + \sum_{k=1}^{n} \frac{1}{2^{k-1}} = 1 + \sum_{k=0}^{n-1} \frac{1}{2^k} \leqslant 1 + 2,$$

przy czym wyjaśnienie pow. nierówności pozostawiam Państwu. Zbieżność $\{a_n\}$ wynika zatem z tw. II.4 i stąd też $e \leqslant 3$. A z tw. II.2, z tego, że $a_2 > 2$ i że a — rosnący, dostajemy $e > 2$.

4. Podciągi

Definicja. $\{a'_n\}_{n\geqslant n_0}$ ***jest podciągiem*** *ciągu* $\{a_n\}_{n\geqslant n_0}$ *wtw istnieje ściśle rosnący ciąg* $\{k_n\}_{n\geqslant n_0}$ *o wyrazach ze zbioru* $\mathbb{N}_{n_0}$ *taki, że* $\forall_{n\geqslant n_0}\ a'_n = a_{k_n}$.

Przykład. $\left\{\frac{1}{n^2}\right\}_{n\geqslant 1}$ jest podciągiem ciągu $a = \left\{\frac{1}{n}\right\}_{n\geqslant 1}$ (wystarczy wziąć $k_n := n^2$) ale $\left\{\frac{1}{\sqrt{n}}\right\}_{n\geqslant 1}$ nie jest podciągiem ciągu a. Także ciąg $b = \left\{\frac{1}{[\sqrt{n}]}\right\}_{n\geqslant 1}$ nie jest podciągiem ciągu a, choć każdy wyraz b_n jest pewnym wyrazem ciągu a.

Trochę w związku z ostatnim przykładem, a trochę z powodów, które wyjaśnią się za moment, przyjmiemy jeszcze drugą definicję.

Definicja. $\{a'_n\}_{n\geqslant n'_0}$ ***jest uogólnionym podciągiem*** *ciągu* $\{a_n\}_{n\geqslant n_0}$ *wtw istnieje ciąg* $\{k_n\}_{n\geqslant n'_0}$ *o wyrazach ze zbioru* $\mathbb{N}_{n_0}$ *taki, że* $k_n \to +\infty$ *oraz* $\forall_{n\geqslant n'_0}\ a'_n = a_{k_n}$.

Uwaga. Podciąg jest uogólnionym podciągiem, bo skoro $\{k_n\}$ jest ściśle rosnący i $\forall_{n\in\mathbb{N}_{n_0}}\ k_n \in \mathbb{N}_{n_0}$, to $k_n \to +\infty$ (dlaczego?). Przykład powyższy z ciągiem b pokazuje, że uogólniony podciąg nie musi być podciągiem.

Twierdzenie II.5 (o granicy uogólnionego podciągu). *Jeżeli a' jest uogólnionym podciągiem ciągu a oraz $a_n \to g \in \overline{\mathbb{R}}$, to $a'_n \to g$.*

Dowód.
Rozważymy tylko przypadek gdy $g \in \mathbb{R}$ (pozostałe są analogiczne). Niech $a'_n = a_{k_n}$, $k_n \to +\infty$. Niech $\epsilon > 0$ i niech M takie, że $|a_n - g| < \epsilon$ dla $n \geqslant M$. Ponieważ $k_n \to +\infty$ zatem dobierzemy N takie, że $k_n \geqslant M$ dla $n \geqslant N$. A zatem $|a_{k_n} - g| < \epsilon$ dla $n \geqslant N$. Dla dowolnego $\epsilon > 0$ dobraliśmy N takie jakie było wymagane definicją i stąd $a'_n = a_{k_n} \to g$. $\square$

Przykład. Rozważmy jeden z przypadków przykładu b) ze strony 22, mianowicie ciąg geometryczny $\{q^n\}_{n\geqslant 1}$ dla $1 > q \geqslant 0$. Oczywiście jest on malejący i ograniczony z dołu przez 0, zatem zbieżny na mocy twierdzenia II.4 do pewnego $g \in \mathbb{R}$. Pytanie tylko czym jest to g? Rozważmy ciąg $\left\{q^{(n+1)}\right\}_{n\geqslant 1}$ — oczywiście to podciąg pierwszego ciągu, zatem z tw. II.5 (i uwagi) mamy $q^{(n+1)} \to g$. Z drugiej strony, z tw. II.1,

$$q^{(n+1)} = q \cdot q^n \to q \cdot g,$$

zatem (patrz fakt II.1 str. 19) szukana granica g spełnia równanie

$$g = q \cdot g.$$

. Łatwo je rozwiązać: ponieważ $q \neq 1$, zachodzi $g = 0$. W efekcie $q^n \to 0$.

Twierdzenia II.5 nie daje się oczywiście odwrócić. Przykładem jest tu np. ciąg o wyrazach $(-1)^n$, który ma podciąg zbieżny do 1, ale sam nie jest zbieżny do 1 (w ogóle nie ma granicy). Inaczej mówiąc jeden ciąg a może mieć wiele rozmaitych granic swoich podciągów.

Niech

$$GP(a) = \left\{g \in \overline{\mathbb{R}} : \text{istnieje podciąg } a' \text{ ciągu } a \text{ taki, że } a'_n \to g\right\}.$$

Już wkrótce wykażemy, że zawsze $GP(a)$ jest niepusty. Gdy $a_n \to g$, to $GP(a)$ jest oczywiście (na mocy tw. II.5) zbiorem jednopunktowym — równym $\{g\}$. Można też dowieść, że $GP(a)$ posiada element największy i najmniejszy[25)] .

[25)] Uwaga: ale w nieco szerszym rozumieniu — przez rozszerzenie (w oczywisty sposób) pojęcia max i min na podzbiory zbioru $\overline{\mathbb{R}}$. $GP(a)$ może nie być już bowiem podzbiorem $\mathbb{R}$. Jeżeli $+\infty \in GP(a)$, to przyjmujemy zatem $+\infty = \max GP(a)$, a gdy $GP(a) = \{-\infty\}$ to $\max GP(a) = -\infty$. I analogicznie z min.

Warto wiedzieć o istnieniu następujących uogólnień na **wszystkie** ciągi pojęcia granicy ciągu (która **nie dla każdego** ciągu istnieje). Mianowicie:

$$\limsup_{n\to+\infty} a_n := \max GP(a)$$

i analogicznie

$$\liminf_{n\to+\infty} a_n := \min GP(a)$$

Można wykazać, że zachodzi

Fakt. $\liminf_{n\to+\infty} a_n = \limsup_{n\to+\infty} a_n = g$ *wtw* $\lim_{n\to+\infty} a_n = g$. **B.D.**

Zajmiemy sie teraz ważnymi rezultatami dotyczącymi istnienia dla danego ciągu pewnych jego szczególnych podciągów.

Lemat. *Każdy ciąg rzeczywisty posiada podciąg monotoniczny.*

Dowód.
Dla „ustalenia uwagi" przyjmujemy, że $n_0 = 1$. Dla ciągu a określamy

$$A := \left\{ k \in \mathbb{N} : \forall_{l>k}\ a_l \geqslant a_k \right\}.$$

Zachodzi jeden z dwóch poniższych przypadków. Dla obu zdefiniujemy rekurencyjnie taki ciąg indeksów $\{k_n\}_{n\geqslant 1}$ (ściśle rosnący), że $\{a_{k_n}\}_{n\geqslant 1}$ jest monotoniczny.

1°. A **jest nieograniczony z góry.** Wówczas niech $k_1 := \min A$, $k_{n+1} := \min\{k \in A : k > k_n\}$ dla $n \in \mathbb{N}$ (inaczej mówiąc, k_n to „n-ty z kolei (co do wielkości) element A") — to poprawna definicja, bo zbiór powyższy jest niepusty skoro A — nieograniczony z góry i dzięki zasadzie minimum to min istnieje. Jednocześnie z definicji mamy $k_{n+1} > k_n$, $k_n \in A \subset \mathbb{N}$ dla dowolnego $n \in \mathbb{N}$, a ponadto z definicji A

$$\forall_{l>k_n}\ a_l \geqslant a_{k_n},$$

w szczególności zatem $a_{k_{n+1}} \geqslant a_{k_n}$. Czyli $\{a_{k_n}\}_{n\geqslant 1}$ jest podciągiem ciągu a i to podciągiem rosnącym.

2°. A **jest ograniczony z góry.** Wybierzmy zatem (zasada Archimedesa) $M \in \mathbb{N}$ takie, że $\forall_{k\in A}\ k < M$. Dla $l \in \mathbb{N}_M$ oznaczmy

$$B(l) := \{k \in \mathbb{N} : k > l, a_k < a_l\}.$$

Zauważmy, że $B(l) \neq \emptyset$. Gdyby bowiem $B(l) = \emptyset$, to $\forall_{k>l}\ a_k \geqslant a_l$, czyli $l \in A$, co jest niemożliwe, bo $l \geqslant M$. Znów więc dzięki zasadzie minimum $B(l)$ posiada min i co więcej zachodzi $\min B(l) > l \geqslant M$. Zatem możemy określić $f : \mathbb{N}_M \to \mathbb{N}_M$ wzorem

$$f(l) = \min B(l)$$

i mamy z definicji $f(l) \in B(l)$, skąd dla $l \in \mathbb{N}_M$ zachodzi

a) $f(l) > l$,

b) $a_{f(l)} < a_l$.

Teraz definiując $k_1 := M$, $k_{n+1} := f(k_n)$ dla $n \in \mathbb{N}$ uzyskujemy z a) $k_{n+1} > k_n$ oraz z b) $a_{k_{n+1}} = a_{f(k_n)} < a_{k_n}$ dla dowolnego $n \in \mathbb{N}$. Czyli $\{a_{k_n}\}_{n\geqslant 1}$ jest podciągiem malejącym ciągu a (i to nawet ściśle).

□

Powyższy lemat i twierdzenie II.4 dają natychmiast

Wniosek. *Każdy ciąg posiada podciąg, który ma granicę w* $\overline{\mathbb{R}}$.

A przy mocniejszych założeniach otrzymujemy znane i bardzo ważne

Twierdzenie II.6 (Bolzano - Weierstrassa). *Każdy ciąg ograniczony posiada podciąg zbieżny.*

Dowód.
Na mocy lematu ograniczony ciąg a posiada podciąg monotoniczny a' — ale ten podciąg jest ciągiem ograniczonym, skoro a jest ograniczony. Zatem z tw. II.4 ciąg a' — zbieżny. □

Jak widać, w tym podejściu do tw. B.-W. cały ciężar dowodu spadł na dowód lematu.

5. Zupełność (trochę inna)

O zupełności była już mowa przy okazji aksjomatu zupełności. Teraz ta nazwa pojawi się w kontekście ciągów w innym, choć powiązanym ściśle z poprzednim, znaczeniu.

Definicja. *Ciąg a jest* ***ciągiem Cauchy'ego*** *wtw*

$$\forall_{\epsilon>0}\ \exists_{N\geqslant n_0}\ \forall_{m,n\geqslant N}\ |a_m - a_n| < \epsilon \tag{II.2}$$

Jak widać, nieco to przypomina definicję granicy ciągu, ale w ogóle w tej definicji granica się nie pojawia. Mowa jest jedynie o tym, że „dla dużych indeksów wyrazy są sobie bliskie". Jednak to pierwsze podobieństwo okazuje się nie być przypadkowe! Zachodzi bowiem:

Twierdzenie II.7 (o zupełności $\mathbb{R}$). *Ciąg liczbowy jest ciągiem Cauchy'ego wtw jest ciągiem zbieżnym.*

Dowód.
„$\Leftarrow$"
Niech $a_n \to g \in \mathbb{R}$ i niech $\epsilon > 0$. Z definicji zbieżności do g istnieje N takie, że $|a_n - g| < \frac{\epsilon}{2}$ dla dowolnego $n \geqslant N$. W szczególności zatem, gdy $m, n \geqslant N$, to $|a_n - g| + |a_m - g| < \epsilon$, skąd z nierówności trójkąta

$$|a_n - a_m| = |a_n - g + g - a_m| \leqslant |a_n - g| + |a_m - g| < \epsilon.$$

„$\Rightarrow$"
Teraz załóżmy, że a jest ciągiem Cauchy'ego.

Wykażemy najpierw, że a jest ograniczony. Korzystając z (II.2) („z $\epsilon = 1$") wybierzmy $N' \geqslant n_0$ takie, że

$$\forall_{m,n\geqslant N'}\ |a_m - a_n| < 1.$$

A zatem jeżeli $n \geqslant N'$, to (bierzemy $m = N'$) $|a_{N'} - a_n| < 1$ skąd $|a_n| \leqslant |a_{N'}| + 1$. A jeżeli $n \leqslant N'$, to $|a_n| \leqslant \max\{|a_k| : n_0 \leqslant k \leqslant N',\ k \in \mathbb{Z}\}$. Stąd a — ciąg ograniczony.

Na mocy twierdzenia Bolzano - Weierstrassa istnieje podciąg uogólniony a' ciągu a, który jest zbieżny. Niech więc $a'_n = a_{k_n} \to g \in \mathbb{R}$, gdzie $k_n \to +\infty$.

Wykażemy, że również $a_n \to g$. Niech $\epsilon > 0$. Korzystając znów z (II.2) wybierzmy $N \geqslant n_0$ takie, że $\forall_{m,n\geqslant N}\ |a_m - a_n| < \frac{\epsilon}{2}$. Ponieważ $k_n \to +\infty$ oraz $a_{k_n} \to g$, zatem możemy wybrać takie s, że $|a_{k_s} - g| < \frac{\epsilon}{2}$ i jednocześnie $k_s \geqslant N$. W szczególności („bierzemy $m = k_s$")

$$\forall_{n\geqslant N}\ |a_{k_s} - a_n| < \frac{\epsilon}{2},$$

skąd dla dowolnego $n \geqslant N$

$$|a_n - g| = |a_n - a_{k_s} + a_{k_s} - g| \leqslant |a_n - a_{k_s}| + |a_{k_s} - g| < \epsilon.$$

To dowodzi, że $a_n \to g$. □

Warto wspomnieć, że powyższe twierdzenie, podobnie jak twierdzenie II.4 (o granicy ciągu monotonicznego) może służyć do dowodu istnienia granicy w takich sytuacjach, gdy nawet nie mamy pomysłu jaka mogłaby być ta ewentualna granica ciągu.

6. Informacja o dalszych twierdzeniach dotyczących granicy ciągu

Na koniec rozdziału II podaję tu bez dowodu kilka użytecznych czasem twierdzeń.

Twierdzenie II.8 (o granicy średniej arytmetycznej i geometrycznej). *Niech* $n_0 = 1$. *Jeżeli* $a_n \to g \in \mathbb{R}$, *to*

$$\frac{\sum_{k=1}^{n} a_k}{n} \to g.$$

Jeżeli ponadto $\forall_{n \geqslant 1}\ a_n > 0$, *to*

$$\sqrt[n]{\prod_{k=1}^{n} a_k} \to g.$$

Twierdzenie II.9 (Stolza). *Niech* $\{a_n\}_{n \geqslant n_0}$ *i* $\{b_n\}_{n \geqslant n_0}$ *będą takie, że*

$$\frac{a_{n+1} - a_n}{b_{n+1} - b_n} \to g \in \overline{\mathbb{R}},$$

gdzie $\{b_n\}_{n \geqslant n_0}$ *jest ciągiem ściśle monotonicznym o wyrazach niezerowych oraz zachodzi jeden z warunków:*

- $a_n \to 0,\ b_n \to 0$
- $b_n \to +\infty$
- $b_n \to -\infty$

Wówczas $\dfrac{a_n}{b_n} \to g$.

Zadania do Rozdziału II

1. Znajdź granicę (lub wykaż jej brak) dla ciągów o wyrazach zadanych wzorami podanymi niżej (o ile „się da" — **bez** użycia twierdzenia o granicy średniej arytmetycznej i geometrycznej i twierdzenia Stolza).

   (a) $\dfrac{n^{\left(9^{\left(9^9\right)}\right)}}{1,0000001^n}$

   (b) $\dfrac{(n-7)^{100}}{(n+7)^{101}}$

   (c) $\dfrac{7^n+6^n-n^{1000}}{(7,1)^n-7^n+n^{1001}}$

   (d) $\dfrac{100^n+n!-\sqrt{n!}}{n!-200^n}$

   (e) $(1,00001-\frac{1}{n})^n$

   (f) $(1+\frac{1}{n^2})^n$

   (g) $(0,9999+\frac{1}{n})^n$

   (h) $(1+\frac{1}{n})^{n^2}$

   (i) $\sqrt[n]{n^{100}}$

   (j) $\sqrt[n]{7^n+3^n}$

   (k) $\sqrt[n]{7^n-3^n}$

   (l) $\sqrt[n]{n!}$

   (m) $\sqrt[n]{n!-100^n}$

   (n) $\sqrt{n^2+n}-n$

   (o) $\frac{n^n}{n!}$

   (p) $\frac{10}{1}\cdot\frac{11}{3}\cdot\ldots\cdot\frac{n+9}{2n-1}$ $(=\prod_{k=1}^{n}\frac{k+9}{2k-1})$

   (q) $\prod_{k=2}^{n}(1-\frac{1}{k^2})$.

$\forall$ **Uwaga.** Do zrobienia w każdej grupie przynajmniej: 2 przykłady spośród (a) - (d), 3 przykłady spośród (e) - (h) oraz przykład (k) lub (m).

2. Znajdź granicę (lub wykaż jej brak) dla następujących ciągów zadanych rekurencyjnie wzorami

   (a) $a_1=\alpha,\ \ a_{n+1}=\frac{1}{2}(a_n+\frac{5}{a_n})$ dla $n\geqslant 1$, dla $\alpha=2$ oraz dla $\alpha=3$;

   (b) $a_1=x,\ \ a_{n+1}=\frac{3}{4}a_n\cdot(2-a_n)$ dla $n\geqslant 1$, w zależności od $x\in[0;2]$.

$\forall$ **Uwaga.** Należy rozwiązać co najmniej jeden z podpunktów.

3. Zbadaj, czy ciągi zadane wzorami poniżej są zbieżne

   (a) $\frac{1}{n}+\frac{1}{n+1}+\cdots+\frac{1}{2n}$ $(=\sum_{k=0}^{n}\frac{1}{n+k})$;

   (b) $\sum_{k=0}^{n}\frac{c_k}{10^k}$, gdzie $\{c_n\}_{n\geqslant 0}$ jest pewnym ciągiem cyfr.

4. Wykaż, że jeśli $\forall_{n\geqslant n_0}\ a_n>0$ oraz $\frac{a_{n+1}}{a_n}\to g$, to $\sqrt[n]{a_n}\to g$.

5. Wykaż, że dla dowolnego $p\in\mathbb{N}_0$ zachodzi
$$\frac{\sum_{k=1}^{n}k^p}{n^{p+1}}\to\frac{1}{p+1}.$$

6. Przeprowadź samodzielnie pominięty na wykładzie dowód twierdzenia II.1 (o rachunkowych własnościach granicy) dla przypadku ilorazu.

7. Wykaż, że jeżeli $k\in\mathbb{N}$ oraz $\forall_{n\geqslant n_0}\ a_n\geqslant 0$, to $a_n\to g\Rightarrow\sqrt[k]{a_n}\to\sqrt[k]{g}$ (rozważ kolejno przypadki: $g=1$, $g>0$, $g=0$)

8. Znajdź przykłady do kilku „nieoznaczoności", tj. sytuacji gdy operacja nie jest określona w rozumieniu ze strony 20, wykazując, że nie dałoby się danej operacji dla tej sytuacji określić w żaden sposób z zachowaniem tezy twierdzenia II.1 (np. dla „$+\infty-(+\infty)$", „$0\cdot(+\infty)$", ...)

9. Wykaż, że $a_n \to 0$ wtw $|a_n| \to 0$ oraz, że niezależnie od wartości g, $a_n \to g \Rightarrow |a_n| \to |g|$ (z umową, że $|\pm\infty| = +\infty$).

$\forall$ 10. Niech $\emptyset \neq A \subset \mathbb{R}$ i niech c będzie ograniczeniem górnym A. Wykaż, że $c = \sup A$ wtw istnieje ciąg $\{a_n\}$ złożony z elementów zbioru A taki, że $a_n \to c$.

11. Znajdź $\sup A$ oraz $\inf A$ dla $A =$

    (a) $\{\frac{n}{m} + \frac{m}{n} : n, m \in \mathbb{N}\}$;

    (b) $\{\frac{a}{a+b} + \frac{b}{b+c} + \frac{c}{c+a} : a, b, c > 0\}$.

12. Wykaż, że jeśli $a_{2n} \to g$ oraz $a_{2n+1} \to g$, to $a_n \to g$

13. Podaj ogólniejsze niż powyżej (jak najogólniejsze ...) warunki na ciągi indeksów $\{k_n\}$, $\{l_n\}$ gwarantujące, że jeśli $a_{k_n} \to g$ i $a_{l_n} \to g$, to zachodzi $a_n \to g$ (oczywiście także wykaż tak utworzone twierdzenie).

14. Wykaż, że jeżeli $a_{2n} \to g$, $a_{2n+1} \to h$, $a_{3n} \to f$ dla pewnych $f, g, h \in \overline{\mathbb{R}}$, to a_n ma granicę $(= f = g = h)$.

15. Wykaż **Twierdzenie**: *Jeżeli ciąg liczbowy nie posiada granicy, to istnieją dwa jego podciągi posiadające różne granice.*

$\forall$ 16. Wykorzystując twierdzenie o 3 ciągach (twierdzenie II.3) oraz twierdzenie o granicy uogólnionego podciągu (twierdzenie II.5) wykaż, że

$$\left(1 + \frac{1}{x_n}\right)^{x_n} \to e$$

jeżeli $x_n \to +\infty$

17. Korzystając z zadań II.16 i II.7 wykaż, że jeżeli $a \in \mathbb{Q}$ oraz $x_n \to +\infty$, to $(1+\frac{a}{x_n})^{x_n} \to e^a$.

18. Wykaż, że dla dowolnego $x \in \mathbb{R}$ zachodzi $\frac{[nx]}{n} \to x$. Mamy w ten sposób przykład „jednolicie zadanego" ciągu liczb wymiernych zbieżnego do dowolnego $x \in \mathbb{R}$. Posługując się powyższym ciągiem znajdź analogiczny ciąg złożony z liczb niewymiernych.

19. Udowodnij twierdzenie Stolza (tw. II.9).

20. Udowodnij wybraną część twierdzenia o granicy średniej arytmetycznej i geometrycznej (tw. II.8).

21. Wykaż, że jeśli $(a_{n+1} - a_n) \to g$, to $\frac{a_n}{n} \to g$

22. Wykaż, że jeśli $(a_{n+1} + a_n) \to 0$, to $\frac{a_n}{n} \to 0$. Czy można tu 0 zastąpić przez dowolne g?

III Szeregi liczbowe

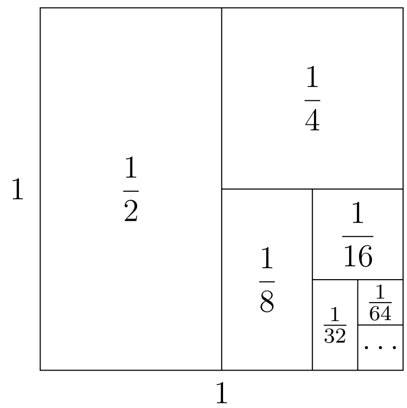
[około 2 wykładów]

1. Definicja „sumy nieskończonej”

Zapewne wielu spośród Państwa posługiwało się nieskończonym sumowaniem jeszcze na długo przed poznaniem pojęcia sumy szeregu — czyli matematycznego uściślenia pojęcia takiej właśnie „nieskończonej sumy”. Typowy przykład to równość

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots = 1,$$

którą można „udowadniać” na wiele sposobów — „algebraicznie” i „geometrycznie” (np. „krojenie” kwadratu 1×1 — patrz rys. 3).



Rysunek 3: Takimi prostokątami „wypełniamy” cały kwadrat o boku 1.

Pojęcie ciągu (z poprzedniego rozdziału) pozwala zrealizować następujący pomysł prowadzący do wspomnianego uściślenia:

> *Zamiast mówić o dodawaniu nieskończenie wielu składników, mówimy o ciągu złożonym z coraz „dłuższych” zwykłych skończonych sum.*

Stąd poniższa definicja. Niech $\{a_n\}_{n \geqslant n_0}$ — ciąg liczbowy.

Definicja. ***Szeregiem*** *o wyrazach a_n dla $n \geqslant n_0$ nazywamy ciąg $\{S_n\}_{n \geqslant n_0}$ zadany wzorem*

$$S_n := \sum_{k=n_0}^{n} a_k, \quad n \geqslant n_0.$$

Oznaczamy go symbolem

$$\sum_{n=n_0}^{+\infty} a_n.$$

Ciąg $\{S_n\}_{n \geqslant n_0}$ nazywamy także ***ciągiem sum częściowych*** *szeregu $\sum_{n=n_0}^{+\infty} a_n$* [26]*. Tym samym symbolem*

$$\sum_{n=n_0}^{+\infty} a_n.$$

[26] Prowadzi to do dość dziwnej sytuacji, że $\{S_n\}_{n \geqslant n_0}$ (czyli $\sum_{n=n_0}^{+\infty} a_n$) jest sam swoim ciągiem sum częściowych — ale cóż — taka jest tradycja ...

oznaczamy również granicę ciągu $\{S_n\}_{n\geqslant n_0}$, o ile granica ta istnieje — nazywamy ją wówczas ***sumą szeregu***[27)]

A zatem np. „napis" $\sum_{n=1}^{+\infty} a_n$ to ciąg o kolejnych wyrazach $S_1 = a_1$, $S_2 = a_1 + a_2$, $S_3 = a_1 + a_2 + a_3, \ldots$ lub granica tego ciągu — wybór znaczenia zależy od kontekstu.

W związku z powyższą definicją, dla szeregów obowiązuje wprowadzona w rozdziale II cała terminologia związana z granicą ciągu. Mówimy więc np., że szereg $\sum_{n=n_0}^{+\infty} a_n$ ma granicę $g \in \overline{\mathbb{R}}$ wtw $S_n \to g$. Są jednak pewne (uświęcone tradycją) wyjątki: nie używa się symbolu „$\sum_{n=n_0}^{+\infty} a_n \to g$" — zamiast tego pisze się

$$\sum_{n=n_0}^{+\infty} a_n = g.$$

Jednak cały czas, tak jak było ogólnie dla wszystkich ciągów, *szereg jest zbieżny* wtw ma granicę (sumę) **skończoną** (tzn. istnieje granica g ciągu sum częściowych i $g \in \mathbb{R}$).

Podsumujmy zatem: owa nieściśle określona dotąd „suma nieskończona" to zwyczjnie granica (o ile istnieje) ciągu sum częściowych, czyli suma szeregu. Natomiast sam szereg (tj. ciąg sum częściowych), to po prostu szczególny rodzaj ciągu. A właściwie może lepiej byłoby powiedzieć: ciąg utworzony w pewien szczególny sposób z ciągu swoich wyrazów $\{a_n\}$. To wcześniejsze stwierdzenie było nienajlepsze, gdyż ma miejsce następujący fakt:

Fakt. *Każdy ciąg jest szeregiem.*

Dowód pozostawiam Państwu (jest nietrudny ...).

2. Ogólne twierdzenia i podstawowe przykłady

Główne twierdzenia, które będą nas interesowały w omawianej tu teorii szeregów to tzw. „kryteria zbieżności", czyli twierdzenia gwarantujące, że przy pewnych założeniach o ciągu wyrazów $\{a_n\}_{n\geqslant n_0}$ szereg $\sum_{n=n_0}^{+\infty} a_n$ jest zbieżny (na ogół nie będziemy wnikali w to jaka jest suma szeregu). Wcześniej jednak sformułujemy kilka twierdzeń o bardziej ogólnym charakterze. Warto zwrócić uwagę na to, że większość z nich to proste konsekwencje, czy wręcz przeformułowania twierdzeń uzyskanych uprzednio w rozdziale dotyczącym ciągów.

Twierdzenie III.1 (o warunku Cauchy'ego dla szeregów). $\sum_{n=n_0}^{+\infty} a_n$ *jest zbieżny wtw*

$$\forall_{\epsilon>0}\ \exists_{N\geqslant n_0}\ \forall_{m\geqslant n\geqslant N}\ \left|\sum_{k=n}^{m} a_k\right| < \epsilon. \tag{III.1}$$

Dowód.
Jeżeli $S_n := \sum_{k=n_0}^{n} a_k$, to $\sum_{k=n}^{m} a_k = S_m - S_{n-1}$ zatem (III.1) można zapisać

$$\forall_{\epsilon>0}\ \exists_{N\geqslant n_0}\ \forall_{m\geqslant n\geqslant N}\ |S_m - S_{n-1}| < \epsilon,$$

co nie tylko wygląda „podobnie" do warunku Cauchy'ego dla ciągu $\{S_n\}_{n\geqslant n_0}$, ale, jak bardzo łatwo się przekonać, [28)] jest mu równoważne. Teza wynika zatem z twierdzenia o zupełności $\mathbb{R}$ (tw. II.7). □

Twierdzenie III.2 (o warunku koniecznym zbieżności szeregu). *Jeżeli* $\sum_{n=n_0}^{+\infty} a_n$ *— zbieżny, to* $a_n \to 0$.

[27)] W związku z tym symbol $\sum_{n=n_0}^{+\infty} a_n$ przestaje niestety mieć swój jednoznaczny sens — oprócz znaczenia szeregu (czyli pewnego ciągu) symbol ten może też oznaczać jego granicę, tj. sumę (czyli pewien element $\overline{\mathbb{R}}$), o ile ta istnieje.

[28)] Choć rzeczywiście łatwo, zachęcam by oba warunki wypisać i by szczegółowo samodzielnie prześledzić dlaczego zachodzą implikacje w obydwie strony.

Dowód.
Można np. użyć twierdzenia III.1, ale można inaczej. Załóżmy, że $\sum_{n=n_0}^{+\infty} a_n = g \in \mathbb{R}$. Dla $n \geqslant n_0+1$ zachodzi $a_n = S_n - S_{n-1} \to g - g = 0$ (korzystamy z twierdzeń o granicy podciągu uogólnionego oraz o granicy różnicy). □

Twierdzenie III.3. *Szereg o wyrazach nieujemnych posiada sumę rzeczywistą lub równą $+\infty$. Jest on zbieżny wtw jego ciąg sum częściowych jest ograniczony z góry.*

Dowód.
To jasne, bo $\{S_n\}_{n\geqslant n_0}$ jest rosnący, zatem wystarczy użyć twierdzenia II.4. □

Uwaga dotycząca oznaczeń. Dla szeregów o wyrazach nieujemnych[29)] piszemy często „$\sum_{n=n_0}^{+\infty} a_n < +\infty$" zamiast „$\sum_{n=n_0}^{+\infty} a_n$ jest zbieżny".

Przykład (szereg geometryczny). Na ogół szeregi zbieżne o nawet dość „prosto wyglądających" wyrazach mają granice nie będące żadnymi „znanymi" liczbami (w tym — wymiernymi). Jednak dla *szeregu geometrycznego* $\sum_{n=0}^{+\infty} q^n$ sprawa jest prosta:

- dla $q \in (-1;1)$ ma sumę $\frac{1}{1-q}$
- dla $q \geqslant 1$ ma sumę $+\infty$
- dla $q \leqslant -1$ nie posiada granicy

Wynika to natychmiast ze znanego (indukcja ...) wzoru na S_n:

$$S_n = \sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q} \text{ dla } q \neq 1.$$

Przed kolejnym twierdzeniem wprowadźmy następującą definicję.

Definicja. $\sum_{n=n_0}^{+\infty} a_n$ ***jest bezwzględnie zbieżny*** *wtw* $\sum_{n=n_0}^{+\infty} |a_n| < +\infty$.

Twierdzenie III.4 (o zbieżności bezwzględnej). *Szereg bezwzględnie zbieżny jest zbieżny.*

Dowód.
Skoro $\sum_{n=n_0}^{+\infty} |a_n|$ jest zbieżny zatem na mocy tw. III.1 zachodzi

$$\forall_{\epsilon>0}\ \exists_{N\geqslant n_0}\ \forall_{m\geqslant n\geqslant N}\ \sum_{k=n}^{m} |a_k| < \epsilon.$$

Ponieważ jednak (nierówność trójkąta uogólniona na $(m-n+1)$ składników)

$$|\sum_{k=n}^{m} a_k| \leqslant \sum_{k=n}^{m} |a_k|,$$

zatem w efekcie otrzymujemy warunek Cauchy'ego także dla $\sum_{n=n_0}^{+\infty} a_n$, więc z tw. III.1 wynika jego zbieżność. □

Jak się już wkrótce przekonamy, twierdzenie odwrotne do twierdzenia III.4 nie zachodzi. Może się więc zdarzyć szereg zbieżny, który nie jest bezwzględnie zbieżny. O takim szeregu mówimy, że jest *zbieżny warunkowo*.

Ostatnie z twierdzeń „ogólnych" to prosty wniosek z twierdzenia o własnościach rachunkowych granicy ciągu.

Twierdzenie III.5. *Załóżmy, że $\sum_{n=n_0}^{+\infty} a_n = A$, $\sum_{n=n_0}^{+\infty} b_n = B$, gdzie $A, B \in \overline{\mathbb{R}}$ oraz że $r \in \mathbb{R}$. Wówczas:*

- *jeżeli $A \pm B$ jest określone, to $\sum_{n=n_0}^{+\infty}(a_n \pm b_n) = A \pm B$*

[29)] I tylko dla takich raczej.

- *jeżeli $r \cdot A$ jest określone, to $\sum_{n=n_0}^{+\infty} r \cdot a_n = r \cdot A$*

Dowód.
To oczywiste z tw. II.1. □

Szereg $\sum_{n=n_0}^{+\infty}(a_n+b_n)$ to suma szeregów $\sum_{n=n_0}^{+\infty} a_n$ i $\sum_{n=n_0}^{+\infty} b_n$ [30]. Z tw. III.5 wynika więc w szczególności, że suma szeregów zbieżnych jest szeregiem zbieżnym i analogicznie dla różnicy, czy mnożenia szeregu przez liczbę.

Uwaga. $\sum_{n=n_0}^{+\infty}(a_n \cdot b_n)$ **nie** jest (na ogół ...) iloczynem szeregów $\sum_{n=n_0}^{+\infty} a_n$ i $\sum_{n=n_0}^{+\infty} b_n$ (rozumianych wciąż jako odpowiednie ciągi sum częściowych) — dlaczego?

Aby w tym podrozdziale powiększyć nieco zasób (dotąd ubogi) przykładów szeregów zbieżnych i rozbieżnych sformułujemy jeszcze następujący fakt.

Lemat (o zagęszczaniu). *Jeżeli $\{a_n\}_{n\geqslant 1}$ jest malejący i nieujemny, to*

$$\sum_{n=1}^{+\infty} a_n < +\infty \quad \text{wtw} \quad \sum_{n=1}^{+\infty} 2^n \cdot a_{2^n} < +\infty.$$

Dowód pozostawiamy jako zadanie.

Przykład. Niech $\alpha \in \mathbb{R}$. $\sum_{n=1}^{+\infty} \frac{1}{n^\alpha} < +\infty$ wtw $\alpha > 1$.
Mamy bowiem $2^n \cdot \frac{1}{(2^n)^\alpha} = \frac{1}{2^{n\cdot(\alpha-1)}} = q^n$, dla $q = \frac{1}{2^{(\alpha-1)}}$. Ponieważ $q \in (-1;1)$ wtw $\alpha > 1$, zatem wystarczy użyć lematu o zagęszczaniu oraz przykładu z szeregiem geometrycznym („zagęszczanie" dokonało tu „cudownej" przemiany badanych szeregów na proste już dla nas szeregi geometryczne).

Warto zapamiętać, że szereg $\sum_{n=1}^{+\infty} \frac{1}{n}$ (tzw. szereg *harmoniczny*) jest „jeszcze" rozbieżny. Tu $\alpha = 1$, czyli to przypadek „graniczny" pomiędzy zbieżnością (dla $\alpha > 1$) i rozbieżnością (dla $\alpha \leqslant 1$).

3. Kryteria zbieżności bezwzględnej

Zaczniemy od bardzo prostego, ale w pewnym sensie także najważniejszego (i w praktyce bardzo użytecznego) twierdzenia pomagającego badać zbieżność konkretnych szeregów o wyrazach nieujemnych.

Kryterium III.1 (porównawcze). *Jeżeli $0 \leqslant a_n \leqslant b_n$ d.d.d. n oraz $\sum_{n=n_0}^{+\infty} b_n$ jest zbieżny, to $\sum_{n=n_0}^{+\infty} a_n$ — zbieżny.* [31]

Dowód poprzedzimy następującym oczywistym faktem (którego dowód zostawiam Państwu).

Lemacik. *Dla dowolnego $n_0' \geqslant n_0$ $\sum_{n=n_0}^{+\infty} a_n$ jest zbieżny wtw $\sum_{n=n_0'}^{+\infty} a_n$ jest zbieżny.*

Dowód (kryterium III.1).
Załóżmy, że $0 \leqslant a_n \leqslant b_n$ dla $n \geqslant n_0'$. Na mocy lemaciku oraz tw. III.3 wystarczy wykazać, że ciąg sum częściowych szeregu $\sum_{n=n_0'}^{+\infty} a_n$ jest ograniczony z góry. Ale z założenia oraz z tw. III.3 istnieje $M \in \mathbb{R}$ takie, że

$$\sum_{k=n_0'}^{n} a_k \leqslant \sum_{k=n_0'}^{n} b_k \leqslant M$$

dla dowolnego $n \geqslant n_0'$. □

Wnioski.

[30] Tzn. ciąg sum częściowych dla $\sum_{n=n_0}^{+\infty}(a_n + b_n)$ jest sumą ciągów sum częściowych dla $\sum_{n=n_0}^{+\infty} a_n$ i dla $\sum_{n=n_0}^{+\infty} b_n$.

[31] Na ogół bez przypominania przyjmujemy, że n_0 jest początkowym indeksem dla rozważanych szeregów.

1) *Jeżeli* $0 \leqslant a_n \leqslant b_n$ *d.d.d.* n *oraz* $\sum_{n=n_0}^{+\infty} a_n$ *— rozbieżny, to* $\sum_{n=n_0}^{+\infty} b_n$ *— rozbieżny.*
A zatem mamy też proste w użyciu kryterium rozbieżności.

2) *Jeżeli* $|a_n| \leqslant b_n$ *d.d.d.* n *oraz* $\sum_{n=n_0}^{+\infty} b_n$ *— zbieżny, to* $\sum_{n=n_0}^{+\infty} a_n$ *— zbieżny bezwzględnie, a stąd również — zbieżny.*
Kryterium porównawcze można więc de facto traktować jako kryterium dotyczące zbieżności bezwzględnej szeregów.

Przed sformułowaniem kolejnego kryterium przyjmijmy następującą definicję i oznaczenie.

Definicja. *Załóżmy, że* $a_n, b_n \neq 0$ *d.d.d.* n*. Ciągi* a *i* b *są* ***asymptotycznie podobne*** *wtw* $\frac{a_n}{b_n} \to g$ *dla pewnego* $g \in \mathbb{R} \setminus \{0\}$*. Oznaczmy to w skrócie* $a \sim b$ *bądź* $a_n \sim b_n$ [32] .

Kryterium III.2 (asymptotyczne). *Jeżeli* $a_n \sim b_n$ *oraz* $\{b_n\}_{n \geqslant n}$ *ma wyrazy stałego znaku od pewnego miejsca, to* $\sum_{n=n_0}^{+\infty} a_n$ *jest zbieżny wtw* $\sum_{n=n_0}^{+\infty} b_n$ *jest zbieżny.*

Dowód.
Zauważmy najpierw, że gdy $a_n \sim b_n$ oraz $\{b_n\}_{n \geqslant n}$ ma wyrazy stałego znaku od pewnego miejsca, to także $\{a_n\}_{n \geqslant n}$ ma wyrazy stałego znaku od pewnego miejsca (patrz np. rozumowanie poniżej). Ponadto $a_n \sim b_n$ wtw $b_n \sim a_n$. A zatem z tej „symetrii" wynika, że wystarczy wykazać implikację w jedną stronę. Ponieważ przemnożenie szeregów przez (-1) nie wpływa na ich zbieżność, możemy założyć, że $b_n > 0$ d.d.d. n oraz że $g = \lim_{n \to +\infty} \frac{a_n}{b_n} > 0$. Z definicji granicy („$\epsilon = \frac{g}{2}$") mamy $\frac{a_n}{b_n} > g - \frac{g}{2} = \frac{g}{2}$ d.d.d. n, więc $\frac{2}{g} \cdot a_n > b_n > 0$ d.d.d. n (w szczególności $\{a_n\}_{n \geqslant n_0}$ ma stały znak od pewnego miejsca). Zatem jeżeli $\sum_{n=n_0}^{+\infty} a_n$ jest zbieżny, to $\sum_{n=n_0}^{+\infty} \frac{2}{g} a_n$ — też (tw. III.5), więc z kryterium porównawczego uzyskujemy zbieżność $\sum_{n=n_0}^{+\infty} b_n$. □

Uwaga. Zbieżność szeregów z kryterium III.2 jest oczywiście równoważna (przy założeniach tego kryterium) ich bezwzględnej zbieżności. Zatem nie warto „próbować" tego kryterium, gdy spodziewamy się zbieżności warunkowej.

Jak widać z dowodu, kryterium asymptotyczne jest formalnie słabsze od kryterium porównawczego, które stanowiło istotę jego dowodu. Jednak w praktyce, kryterium asymptotyczne bywa często dużo wygodniejsze w użyciu. Idea jego użycia jest taka: jeżeli wyraz szeregu $\sum_{n=n_0}^{+\infty} a_n$ jest zadany dość skomplikowanym wzorem, to znajdujemy „prostszy" ciąg $\{b_n\}$ o stałym znaku i asymptotycznie podobny do $\{a_n\}$ (a jak taki znaleźć: często wystarczy w formule na a_n zostawić tylko „to co najistotniejsze"). Tym sposobem problem badania zbieżności sprowadza nam się do badania tylko „prostszego" szeregu $\sum_{n=n_0}^{+\infty} b_n$.

Przykład. Niech
$$a_n = \frac{\frac{1}{n^2} - \frac{70}{n^3}}{\frac{1}{n} + \frac{17}{n^2}}.$$
Czy $\sum_{n=1}^{+\infty} a_n$ jest zbieżny? Łatwo odgadnąć „wygodny" $\{b_n\}$. Niech mianowicie
$$b_n := \frac{\frac{1}{n^2} + 0}{\frac{1}{n} + 0} = \frac{1}{n} > 0.$$
Banalnie sprawdzamy, że $a_n \sim b_n$ (a nawet $\frac{a_n}{b_n} \to 1$) i z kryterium asymptotycznego $\sum_{n=1}^{+\infty} a_n$ — rozbieżny, bo $\sum_{n=1}^{+\infty} \frac{1}{n}$ — rozbieżny. Oczywiście dało się też bezpośrednio szacować a_n z dołu (jak?) tak by użyć „zwykłego" kryterium porównawczego, ale użyte kryterium III.2 wydaje się tu szybsze i bardziej „automatyczne".

Czasami (choć zazwyczaj nie aż tak często, jak chcieliby tego studenci...) przydają się następujące dwa kryteria, także będące konsekwencjami kryterium porównawczego.

[32] Ta druga wersja oznaczenia (choć wygodna) jest nieco nieformalna (podobnie jak oznaczenie $a_n \to g$), bo chodzi tu przecież o własność **ciągów**, a nie własności ich n-tego wyrazu ... Ponadto — uwaga: ani nazwa „asymptotycznie podobne", ani symbol „$\sim$" nie są zbyt powszechnie używane.

Kryterium III.3 (d'Alemberta) oraz III.4 (Cauchy'ego). *Niech*

$$c_n := \begin{cases} |\frac{a_{n+1}}{a_n}| & \textit{dla d'Alemberta}^{33)} \\ \sqrt[n]{|a_n|} & \textit{dla Cauchy'ego} \end{cases}$$

i załóżmy, że $c_n \to g \in \overline{\mathbb{R}}$. Jeżeli $g < 1$, to $\sum_{n=n_0}^{+\infty} a_n$ jest zbieżny bezwzględnie, a jeżeli $g > 1$, to $\sum_{n=n_0}^{+\infty} a_n$ jest rozbieżny.

Dowód (dla d'Alemberta). (dla C. — jeszcze łatwiej...) Jeżeli $0 \leqslant g < 1$, to $g = 1 - 2\epsilon$ dla pewnego $\epsilon > 0$, zatem z definicji granicy dla ciągu $\{c_n\}$ istnieje N takie, że $c_n \leqslant g + \epsilon = 1 - \epsilon$ o ile $n \geqslant N$. Zatem dla $n \geqslant N + 1$

$$(1-\epsilon)^{(n-N)} \geqslant \prod_{k=N}^{n-1} c_k = \left|\frac{a_n \cdot \ldots \cdot a_{N+1}}{a_{n-1} \cdot \ldots \cdot a_N}\right| = \frac{|a_n|}{|a_N|},$$

czyli $|a_n| \leqslant \frac{|a_N|}{(1-\epsilon)^N} \cdot (1-\epsilon)^n$ d.d.d. n.

Ponieważ szereg, którego n-ty wyraz jest po prawej stronie pow. nierówności to zbieżny szereg geometryczny pomnożony przez stałą, zatem $\sum_{n=n_0}^{+\infty} a_n$ jest zbieżny bezwzględnie na mocy wniosku 2 z kryterium porównawczego.

Jeżeli $g > 1$ to zapisując $g = 1+2\epsilon$ z $\epsilon > 0$ i biorąc N takie, że $c_n \geqslant g - \epsilon = 1+\epsilon$ dostajemy analogicznie $|a_n| \geqslant \frac{|a_N|}{(1+\epsilon)^N}(1+\epsilon)^n$ d.d.d. n. Zatem z twierdzenia „o dwóch ciągach" $|a_n| \to +\infty$, czyli $a_n \nrightarrow 0$. Stąd $\sum_{n=n_0}^{+\infty} a_n$ — rozbieżny na mocy tw. III.2 (o warunku koniecznym). □

Przykłady.

1. Dla $\sum_{n=1}^{+\infty} \frac{1}{n^\alpha}$ te kryteria nie działają dla żadnego $\alpha \in \mathbb{R}$, bo dostajemy $g = 1$ (choć nawet nie dla wszystkich $\alpha \in \mathbb{R}$ jest to przy naszej obecnej wiedzy takie jasne; patrz np. zadanie II.7)

2. Dla $\sum_{n=0}^{+\infty} \frac{1}{n!}$ dostajemy przy użyciu kryterium d'Alemberta $c_n = \frac{1}{n+1} \to 0$, zatem szereg ten jest zbieżny. Co więcej można wykazać (i nie jest to bardzo trudne, choć na nasze wykłady — zbyt czasochłonne, ale zachęcam do samodzielnych prób[34)]) następujący fakt.

Fakt. $\sum_{n=0}^{+\infty} \frac{1}{n!} = e$.

Uwaga. Z twierdzenia o granicy średniej geometrycznej (tw. II.8) nietrudno dowieść, że ciąg spełniający założenia kryterium d'Alemberta musi także spełniać założenia kryterium Cauchy'ego. A zatem z formalnego punktu widzenia kryterium Cauchy'ego jest „mocniejsze" (tzn. działa dla niemniejszej klasy przypadków niż kryt. d'Alemberta). Mimo to czasem wygodniej jest użyć kryt. d'Alemberta niż Cauchy'ego ze względów czysto rachunkowych.

4. Kryteria zbieżności „niekoniecznie bezwzględnej"

Kryteria z poprzedniego podrozdziału nie nadawały się do bezpośredniego dowodzenia zbieżności takiego szeregu, który nie byłby zbieżny bezwzględnie. Wszystkie one opierały się bowiem na kryterium porównawczym. Do badania zbieżności szeregów, które nie muszą być jednak bezwzględnie zbieżne przydaje się dość często następujące kryterium.

Kryterium III.5 (Dirichleta). *Jeżeli $\{a_n\}_{n \geqslant n_0}$ jest monotoniczny od pewnego miejsca i $a_n \to 0$ oraz $\sum_{n=n_0}^{+\infty} b_n$ ma ograniczony ciąg sum częściowych, to*

$$\sum_{n=n_0}^{+\infty} a_n \cdot b_n$$

[33)] A zatem w kryt. III.3 zakładamy, że $a_n \neq 0$ d.d.d. n.

[34)] Patrz zadanie III.18.

jest zbieżny.

W dowodzie wykorzystamy następującą formułę.

Lemat (przekształcenie Abela). *Dla dowolnych liczb $u_1, \ldots, u_n, v_1, \ldots, v_n$ zachodzi*

$$\sum_{k=1}^{n} u_k \cdot v_k = \sum_{k=1}^{n-1} (\sum_{l=1}^{k} u_l) \cdot (v_k - v_{k+1}) + (\sum_{l=1}^{n} u_l) \cdot v_n.$$ [35)]

Dowód.
Prosta indukcja. □

Dowód (kryterium Dirichleta).
Możemy założyć, że $\{a_n\}_{n \geqslant n_0}$ jest monotoniczny (patrz lemacik ze str. 34), a co więcej, że jest malejący (gdyby był rosnący, to zamiast wyrazów a_n, b_n można rozważyć $-a_n$, $-b_n$ i skorzystać z wersji „malejącej”).

Wykażemy, że szereg $\sum_{n=n_0}^{+\infty} a_n b_n$ spełnia war. Cauchy'ego dla szeregów, co dzięki tw. III.1 da nam jego zbieżność. Niech $T_n := \sum_{k=n_0}^{n} b_k$ dla $n \geqslant n_0$. Niech $M > 0$ będzie takie, że $\forall_{n \geqslant n_0}\ |T_n| \leqslant M$. To, że $\{a_n\}_{n \geqslant n_0}$ — malejący i zbieżny do 0 daje nam, że $\forall_{n \geqslant n_0}\ a_n \geqslant 0$. Zatem dla $\epsilon > 0$ można dobrać $N > n_0$ takie, że

$$\forall_{n \geqslant N}\ 0 \leqslant a_n < \frac{\epsilon}{2M}.$$

Korzystając teraz kolejno z lematu (o przekształceniu Abela), z nierówności trójkąta, z „malenia” $\{a_n\}$ dostajemy dla dowolnych $m \geqslant n \geqslant N$:

$$\begin{aligned} |\sum_{k=n}^{m} b_k a_k| \leqslant \sum_{k=n}^{m-1} |\sum_{l=n}^{k} b_l| \cdot (a_k - a_{k+1}) + |\sum_{l=n}^{m} b_l| \cdot a_m &= \sum_{k=n}^{m-1} |T_k - T_{n-1}| \cdot (a_k - a_{k+1}) + \\ +|T_m - T_{n-1}| \cdot a_m \leqslant 2M \sum_{k=n}^{m-1} (a_k - a_{k+1}) + 2M a_m &= 2M(a_n - a_m + a_m) = 2M a_n < \epsilon \end{aligned}$$

□

W udowodnionym właśnie nowym kryterium nieco „dziwaczne” może się wydawać założenie dotyczące ograniczoności ciągu sum częściowych dla $\sum_{n=n_0}^{+\infty} b_n$. Jednak ciągów b_n spełniających ten warunek jest bardzo wiele. Wybór każdego konkretnego takiego ciągu daje nam automatycznie jakieś kryterium, będące pewnym szczególnym przypadkiem kryt. Dirichleta. Np., gdy $b_n := (-1)^n$ to odpowiedni ciąg sum częściowych jest ograniczony, gdyż przyjmuje tylko dwie możliwe wartości: 0 oraz 1 względnie -1 (w zależności od parzystości n_0). Zatem uzyskujemy

Wniosek (kryterium Leibnitza). *Jeżeli $\{a_n\}_{n \geqslant n_0}$ jest monotoniczny od pewnego miejsca i $a_n \to 0$, to $\sum_{n=n_0}^{+\infty} (-1)^n a_n$ jest zbieżny.*

To kryterium pozwala nam podać zapowiadany wcześniej przykład.

Przykład. $\sum_{n=1}^{+\infty} (-1)^n \frac{1}{n}$ jest warunkowo zbieżny. Ogólniej: $\sum_{n=1}^{+\infty} (-1)^n \frac{1}{n^\alpha}$ dla $\alpha \in (0; 1]$ jest warunkowo zbieżny.

5. Zmiana kolejności sumowania

Można zadać sobie ogólne pytanie:

35) Gdy $n = 1$, to z prawej strony wzoru pojawia się „$\sum_{k=1}^{0} \ldots$” Stosujemy umowę, że zawsze gdy $q < p$, to $\sum_{k=p}^{q} \ldots = 0$.

Jakie własności zwykłego („skończonego") sumowania przenoszą się na „sumowanie nieskończone"?

Dość chyba naturalne oczekiwanie, że przenoszą się „wszystkie" własności okazuje się jednak być złudne. Należy zachować daleko posuniętą ostrożność przy próbach przenoszenia własności sum skończonych na przypadek nieskończony. Dość często okazuje się, że w miarę „bezboleśnie" takiego przeniesienia można dokonać przy dodatkowym założeniu o bezwzględnej zbieżności szeregu. Zilustrujemy to na przykładzie problemu przemienności dodawania. Aby to uściślić przypomnijmy najpierw, że $p : \mathbb{N}_{n_0} \to \mathbb{N}_{n_0}$ jest *permutacją* $\mathbb{N}_{n_0}$ wtw p jest „na" i „1-1"[36)].

Problem. *Jaki jest związek* $\sum_{n=n_0}^{+\infty} a_n$ *z* $\sum_{n=n_0}^{+\infty} a_{p(n)}$ *(istnienia granic, wartości sumy)?*

Zacznijmy od twierdzenia „pozytywnego".

Twierdzenie III.6 (o przemienności szeregów bezwzględnie zbieżnych). *Jeżeli* $\sum_{n=n_0}^{+\infty} a_n$ *jest bezwzględnie zbieżny, to przy dowolnej permutacji p zbioru $\mathbb{N}_{n_0}$ dla sum szeregów zachodzi*

$$\sum_{n=n_0}^{+\infty} a_n = \sum_{n=n_0}^{+\infty} a_{p(n)}.$$

B.D.

Okazuje się, że założenie o bezwzględnej zbieżności jest tu bardzo istotne — wręcz niezbędne, o ile obracamy się w kręgu szeregów zbieżnych. Co jeszcze ciekawsze, prawdziwy jest poniższy, dość zaskakujący na pierwszy rzut oka wynik.

Twierdzenie III.7 (Riemanna). *Załóżmy, że* $\sum_{n=n_0}^{+\infty} a_n$ *jest zbieżny warunkowo. Wówczas dla dowolnego* $G \in \overline{\mathbb{R}}$ *istnieje taka permutacja p zbioru* $\mathbb{N}_{n_0}$*, że*

$$\sum_{n=n_0}^{+\infty} a_{p(n)} = G.$$

Istnieje także taka permutacja, że $\sum_{n=n_0}^{+\infty} a_{p(n)}$ *nie posiada granicy.* **B.D.**

Podobne „kłopoty" mogą pojawiać się także dla własności innych niż przemienność. Np. dla grupowania wyrazów poprzez „dopisywanie nawiasów", co wiąże się z własnością łączności dodawania (patrz — zadania do tego rozdziału).

6. Mnożenie szeregów

Jak mnożyć szeregi? Właściwie — wiadomo: skoro szereg to po prostu ciąg sum częściowych, to można zwyczajnie brać iloczyn ciągów sum częściowych. Jednak tak określone działanie w zbiorze szeregów nie jest zbyt interesujące i nie ma zbyt istotnych zastosowań. Zamiast powyższego „zwykłego" iloczynu szeregów zdefiniujemy inne — dość popularne działanie zwane *iloczynem Cauchy'ego*. Zrobimy to tylko dla szeregów o indeksie początkowym $n_0 = 0$. Iloczyn Cauchy'ego będziemy tu oznaczać symbolem $\odot$ (raczej niespotykanym gdzie indziej...).

Definicja. $\sum_{n=0}^{+\infty} a_n \odot \sum_{n=0}^{+\infty} b_n := \sum_{n=0}^{+\infty} c_n$, *gdzie*

$$c_n := \sum_{k=0}^{n} a_k \cdot b_{n-k}$$

dla $n \in \mathbb{N}_0$.

[36)] Podobnie definiuje się permutację w przypadku zbiorów skończonych zamiast $\mathbb{N}_{n_0}$. Oczywiście napis „1-1" nie oznacza tu liczby 0 lecz **różnowartościowość** funkcji p (patrz np. rozdział IV).

Łatwo sprawdzić, że $\odot$ posiada sporo właściwości analogicznych do własności zwykłego iloczynu dla liczb rzeczywistych, takich jak np. łączność, czy przemienność. Nas jednak przede wszystkim interesować będzie związek pomiędzy mnożeniem szeregów a ich sumami. Sformułujemy bez dowodu następujące twierdzenie dotyczące tej kwestii.

Twierdzenie III.8 (tw. Mertensa + tw. Cesaro[37]). *Rozważmy dwa zbieżne szeregi takie, że* $\sum_{n=0}^{+\infty} a_n = A$, $\sum_{n=0}^{+\infty} b_n = B$ *i niech* $\sum_{n=0}^{+\infty} c_n = \sum_{n=0}^{+\infty} a_n \odot \sum_{n=0}^{+\infty} b_n$. *Wówczas, jeżeli zachodzi* ***któryś*** *z poniższych warunków:*

1. *(Mertens) przynajmniej jeden z szeregów* $\sum_{n=0}^{+\infty} a_n$, $\sum_{n=0}^{+\infty} b_n$ *jest bezwzględnie zbieżny,*

2. *(Cesaro)* $\sum_{n=0}^{+\infty} c_n$ *posiada granicę,*

to

$$\sum_{n=0}^{+\infty} c_n = A \cdot B.$$

Przykłady. Określimy funkcję $\exp\colon \mathbb{R} \to \mathbb{R}$ wzorem

$$\exp(x) := \sum_{n=0}^{+\infty} \frac{x^n}{n!} \qquad \text{dla } x \in \mathbb{R}.$$

Nietrudno zauważyć, że powyższy szereg jest bezwzględnie zbieżny dla każdego $x \in \mathbb{R}$. Ponadto można też wykazać (polecam jako zadanie „rachunkowe"), że

$$\sum_{n=0}^{+\infty} \frac{x^n}{n!} \odot \sum_{n=0}^{+\infty} \frac{y^n}{n!} = \sum_{n=0}^{+\infty} \frac{(x+y)^n}{n!}\ ^{38)}.$$

W takim razie, dzięki twierdzeniu Mertensa prawdziwy jest

Fakt III.1. $\forall_{x,y\in\mathbb{R}}\ \exp(x+y) = \exp(x)\cdot\exp(y)$.

W przyszłości okaże się, że $\exp(x)$ to to samo co e^x, jednak na razie brak nam jeszcze narzędzi, by to wykazać.

Teraz kolej na *funkcje trygonometryczne:* sin i cos. Definiujemy je tak:

$$\sin(x) := \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \quad \cos(x) := \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!} \quad \text{dla } x \in \mathbb{R}.$$

I znów dzięki twierdzeniu Mertensa, rozumując jak wyżej, można wykazać

Fakt III.2. *Dla dowolnych* $x, y \in \mathbb{R}$ *zachodzi:*

1. $\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y)$;

2. $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$;

3. $(\sin(x))^2 + (\cos(x))^2 = 1$.

B.D.

[37] ściślej — to tylko wniosek z tw. Cesaro zwany też twierdzeniem Abela.

[38] Uwaga: tu symbol „$\sum$ ..." ma oznaczać szereg, w odróżnieniu od „$\sum$ ..." w definicji exp, gdzie oznacza on sumę odpowiedniego szeregu.

Zadania do Rozdziału III

1. Wykaż, że każdy ciąg liczbowy jest szeregiem (fakt ze str. 32).

2. Znajdź sumy poniższych szeregów:

   (a) $\sum_{n=1}^{+\infty} \frac{1}{n(n+1)}$

   (b) $\sum_{n=2}^{+\infty} \frac{1}{n^2-1}$

   $\forall$ (c) $\sum_{n=0}^{+\infty} \frac{(2^n+(-1)^n)^2}{11^n}$

   (d) $\sum_{n=0}^{+\infty} \frac{n}{17^n}$ (najpierw wyprowadź wzór na wyraz S_n ciągu sum częściowych szeregu $\sum_{n=0}^{+\infty} nq^n$ dla $q \neq 1$, zapisując S_{n+1} przy pomocy S_n na dwa istotnie różne sposoby).

$\forall$ 3. [39] Zbadaj zbieżność i bezwzględną zbieżność poniższych szeregów:

   (a) $\sum_{n=1}^{+\infty} \frac{n^2-3}{n^4+3n}$

   (b) $\sum_{n=1}^{+\infty} \frac{n^2-3}{n^3+3n}$

   (c) $\sum_{n=1}^{+\infty} \frac{4^n+3^n}{3^n-2^n}$

   (d) $\sum_{n=1}^{+\infty} \frac{-n^2+3n+1}{\sqrt{n^7-1}}$

   (e) $\sum_{n=0}^{+\infty} \frac{n^{10000}}{(1,000001)^n}$

   (f) $\sum_{n=1}^{+\infty} (\sqrt[n]{5}-1)$

   (g) $\sum_{n=1}^{+\infty} (\sqrt[n]{2}-1)$

   (h) $\sum_{n=1}^{+\infty} (\sqrt[n]{n}-1)\cdot(-1)^n$

   (i) $\sum_{n=1}^{+\infty} (\sqrt[n]{3}-1)^\alpha$ w zależności od $\alpha \in \mathbb{R}$

   (j) $\sum_{n=1}^{+\infty} \frac{n!}{n^n}$

   (k) $\sum_{n=0}^{+\infty} (-1)^n \frac{\sqrt{n+2}}{n+3}$

   (l) $\sum_{n=1}^{+\infty} \frac{(-1)^n}{\sqrt[n]{n}}$.

4. Korzystając z wiedzy z GAL-u: postaci trygonometrycznej liczby zespolonej i jej n-tej potęgi (wzór de Moivre'a) oraz ze wzoru na $\sum_{k=0}^{n} z^k$ [40], wykaż, że szereg $\sum_{n=1}^{+\infty} \frac{\sin(nx)}{n}$ jest zbieżny przy dowolnym $x \in \mathbb{R}$.

5. Wykaż, że jeżeli $\{a_n\}_{n\geqslant 0}$ jest ściśle malejący i $a_n \to 0$, to $\sum_{n=0}^{+\infty} (-1)^n a_n > 0$.

6. Wykaż, że $\cos(2) < 0$.

7. Wykaż, że jeżeli zachodzi któryś z warunków:

   (a) $\{na_n\}_{n\geqslant 1}$ jest ograniczony

   (b) $\forall_{n\in\mathbb{N}}\ a_n \geqslant 0$ i $\sum_{n=1}^{+\infty} a_n$ jest zbieżny,

   to $\sum_{n=1}^{+\infty} (a_n)^2$ jest zbieżny. Czy założenie o nieujemności w b) jest istotne?

[39] Przynajmniej 2 szt. spośród a)–d) i 4 szt. spośród pozostałych.
[40] Patrz zadanie I.1.

8. Czy suma szeregów rozbieżnych jest zawsze szeregiem rozbieżnym?

9. Wykaż, że jeżeli $\{a_n\}_{n\geqslant 1}$ jest malejący oraz $\sum_{n=1}^{+\infty} a_n$ jest zbieżny, to $na_n \to 0$. Czy założenie, że ciąg jest malejący jest istotne?

10. Czy prawdziwe jest następujące „twierdzenie o trzech szeregach”:
Jeżeli $a_n \leqslant b_n \leqslant c_n$ dla $n \geqslant n_0$ oraz szeregi $\sum_{n=n_0}^{+\infty} a_n$ i $\sum_{n=n_0}^{+\infty} c_n$ są zbieżne, to $\sum_{n=n_0}^{+\infty} b_n$ jest zbieżny?

11. Wykaż, że jeżeli $\sum_{n=n_0}^{+\infty} a_n$ jest bezwzględnie zbieżny, to $|\sum_{n=n_0}^{+\infty} a_n| \leqslant \sum_{n=n_0}^{+\infty} |a_n|$.

12. Wykaż „twierdzenie o reszcie szeregu zbieżnego”:
Jeżeli $\sum_{n=n_0}^{+\infty} a_n$ jest zbieżny, to $\forall_{\epsilon>0}\ \exists_{N\geqslant n_0}\ |\sum_{n=N}^{+\infty} a_n| < \epsilon$.

$\forall$ 13. Wykaż „lemat o zagęszczaniu” (patrz str. 34).

14. Udowodnij następujące „drugie kryterium porównawcze”: *Jeżeli ciągi a, b o wyrazach dodatnich spełniają $\frac{a_{n+1}}{a_n} \leqslant \frac{b_{n+1}}{b_n}$ d.d.d. n oraz $\sum_{n=n_0}^{+\infty} b_n$ jest zbieżny, to $\sum_{n=n_0}^{+\infty} a_n$ jest też zbieżny.*

15. Wykaż, że w kryterium „asymptotycznym” (kryterium III.2) nie można zrezygnować z założenia o stałym znaku (od pewnego miejsca).

16. Znajdź przykład takiego $\{a_n\}_{n\geqslant 1}$ i $g \in \mathbb{R}$, że $\forall_{n\in\mathbb{N}}\ a_n > 0$, $\sqrt[n]{n} \to g$, ale $\frac{a_{n+1}}{a_n} \not\to g$.

17. Wykaż, że jeżeli $\sum_{k=k_0}^{+\infty} a_k$ jest bezwzględnie zbieżny, $\sum_{k=k_0}^{+\infty} a_k = g$ oraz liczby $c_{k,n}$ określone dla dowolnych $k \geqslant k_0$, $n \geqslant n_0$ spełniają:

    (a) $\exists_{M>0}\ \forall_{k\geqslant k_0, n\geqslant n_0}\ |c_{k,n}| \leqslant M$,

    (b) $\forall_{k\geqslant k_0}\ \lim_{n\to+\infty} c_{k,n} = 1$,

    to ciąg określony wzorem $L_n := \sum_{k=k_0}^{+\infty} a_k c_{k,n}$ $(n \geqslant n_0)$ jest zbieżny do g („dyskretna” wersja tw. Lebesgue’a o zbieżności majoryzowalnej).

18. Korzystając z zadania III.17 wykaż, że $\sum_{n=0}^{+\infty} \frac{1}{n!} = e$ (gdzie e to zdefiniowana w rozdziale II liczba równa $\lim_{n\to+\infty}(1+\frac{1}{n})^n$).

19. Przy użyciu kryterium Dirichleta (kryterium III.5) udowodnij poniższe kryterium Abela:
Jeżeli $\{a_n\}_{n\geqslant n_0}$ jest monotoniczny od pewnego miejsca i jest ograniczony oraz $\sum_{n=n_0}^{+\infty} b_n$ jest zbieżny, to $\sum_{n=n_0}^{+\infty} a_n b_n$ jest zbieżny.

    Poniższe dwa zadania dotyczą dwóch sposobów grupowania wyrazów szeregu.

20. (a) Wykaż, że jeżeli $\sum_{n=0}^{+\infty} a_{2n}$ i $\sum_{n=0}^{+\infty} a_{2n+1}$ są zbieżne, to zbieżny jest także $\sum_{n=0}^{+\infty} a_n$.

    (b) Czy zachodzi odwrotna implikacja?

    (c) Sformułuj i wykaż uogólnienie twierdzenia z punktu a) takie, by obejmowało ono możliwie ogólne rozkłady zbioru indeksów na dwa podzbiory.

$\forall$ 21. [41] Niech $\{a_n\}_{n\geqslant n_0}$ będzie ciągiem liczbowym, $\{p_n\}_{n\geqslant 1}$ niech będzie ściśle rosnącym ciągiem indeksów z $\mathbb{N}_{n_0}$ takim, że $p_1 = n_0$ (p_n będziemy interpretować jako „początek n-tej

[41] Przynajmniej z punktem c) w wersji (i).

grupy" przy grupowaniu[42] wyrazów szeregu $\sum_{n=n_0}^{+\infty} a_n$). Niech

$$A_n := \sum_{k=p_n}^{p_{n+1}-1} a_k$$

dla $n \in \mathbb{N}$ oraz niech $G \in \overline{\mathbb{R}}$. Wykaż, że

(a) $\sum_{n=n_0}^{+\infty} a_n = G \Rightarrow \sum_{n=1}^{+\infty} A_n = G$,

(b) może nie zachodzić „$\Leftarrow$" powyżej,

(c) „$\Leftarrow$" powyżej zachodzi, o ile zachodzi **któreś** z poniższych założeń:

i. $\forall_{n\in\mathbb{N}}\ p_{n+1} - p_n = 2$ oraz $a_n \to 0$,

ii. $\{p_{n+1} - p_n\}_{n\geqslant 1}$ jest stały oraz $a_n \to 0$,

iii. $\{p_{n+1} - p_n\}_{n\geqslant 1}$ jest ograniczony oraz $a_n \to 0$,

iv. dla dowolnego $n \in \mathbb{N}$ wszystkie liczby w zbiorze $\{a_k \colon p_n \leqslant k \leqslant p_{n+1} - 1\}$ mają ten sam znak (tj. wszystkie są $\geqslant 0$ lub wszystkie są $\leqslant 0$; ale oczywiście ten znak może zależeć od n),

v. szereg $\sum_{n=n_0}^{+\infty} a_n$ posiada granicę.

$\forall$ 22. Zbadaj zbieżność szeregów

(a) $\displaystyle\sum_{n=1}^{+\infty} (-1)^n \frac{1}{n + (-1)^{n+1}}$;

(b) $\displaystyle\sum_{n=1}^{+\infty} (-1)^n \frac{1}{\sqrt{n} + (-1)^{n+1}}$.

23. Udowodnij twierdzenie o przemienności szeregu bezwzględnie zbieżnego (tw. III.6, dowód pominięty na wykładzie).

24. Wykaż, że jeśli $\forall_{n\geqslant n_0}\ a_n \geqslant 0$ oraz p jest permutacją $\mathbb{N}_{n_0}$, to $\sum_{n=n_0}^{+\infty} a_n = \sum_{n=n_0}^{+\infty} a_{p(n)}$ (niezależnie od tego, czy zachodzi zbieżność, czy nie).

25. Zbadaj, czy iloczyn Cauchy'ego $\odot$ jest operacją: przemienną, łączną. Jakie są szeregi neutralne dla $\odot$? Jakie szeregi posiadają elementy odwrotne względem $\odot$?

26. Znajdź wzór opisujący zwykły iloczyn $\cdot$ szeregów $\sum_{n=0}^{+\infty} a_n$ i $\sum_{n=0}^{+\infty} b_n$ (tzn. działanie „$\cdot$" jest takie, że $(\sum_{n=0}^{+\infty} a_n) \cdot (\sum_{n=0}^{+\infty} b_n) = \sum_{n=0}^{+\infty} d_n$, gdzie $\sum_{k=0}^{n} d_n = (\sum_{k=0}^{n} a_n)(\sum_{k=0}^{n} b_n)$ przy dowolnym $n \in \mathbb{N}_0$).

27. Wykaż, że iloczyn Cauchy'ego szeregów bezwzględnie zbieżnych jest szeregiem bezwzględnie zbieżnym.

$\forall$ 28. Wypisz szczegółowe dowody (pominięte na wykładzie) dla formuł „algebraicznych" dotyczących iloczynów Cauchy'ego odpowiednich szeregów potrzebnych przy dowodach **przynajmniej jednej** spośród poniższych formuł (patrz fakty III.1 i III.2 ze str. 39):

(a) $\exp(x + y) = \exp(x) \cdot \exp(y)$,

(b) $\sin(x + y) = \sin(x) \cdot \cos(y) + \cos(x) \cdot \sin(y)$,

(c) $\cos(x + y) = \cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y)$,

(d) $(\sin(x))^2 + (\cos(x))^2 = 1$.

29. Wykaż, że $\forall_{x\in\mathbb{R}}\ \exp(x) > 0$. (Wskazówka: użyj zad. III.28 (a).)

---

[42] Opisywany tu rodzaj grupowania nazywany bywa *rozstawianiem* (albo *dopisywaniem*) *nawiasów*.

IV Granica i ciągłość funkcji

[około 3 wykładów]

1. Granica funkcji

Pojęcie granicy ciągu wprowadzone w rozdziale II rozszerzymy na znacznie ogólniejsze sytuacje. To uogólnienie pójdzie w dwóch kierunkach. Po pierwsze, będziemy rozważać szerszą klasę funkcji niż tylko ciągi. Po drugie, granica będzie mogła być rozważana nie tylko „w $+\infty$”, ale także „w innych punktach”. Dla funkcji $f : D \to \mathbb{R}$, gdzie dziedzina D funkcji f jest podzbiorem $\mathbb{R}$, granicę będziemy mogli (ewentualnie, bo nie musi ona istnieć) rozważać jedynie w takich punktach z $\overline{\mathbb{R}}$, które są tzw. *punktami skupienia D*.

Definicja. *Niech $a \in \overline{\mathbb{R}}, D \subset \mathbb{R}$. Wówczas a jest* ***punktem skupienia*** *D (będziemy to skracać: p.s.) wtw istnieje ciąg $\{x_n\}_{n\geqslant n_0}$ w. $D \setminus \{a\}$* [43] *taki, że $x_n \to a$.*

Rozważane przez nas funkcje będą określone najczęściej na dziedzinach będących pewnymi przedziałami[44] . Wtedy oczywiście sprawa jest prosta: a jest punktem skupienia przedziału niezerowej długości wtw a należy do tego przedziału, bądź jest jego prawym lub lewym końcem (niezależnie od tego, czy te końce do przedziału należą, czy nie). Ale np. zbiór skończony nie ma punktów skupienia, $\mathbb{N}_{n_0}$ ma tylko $+\infty$, dla $\{0\} \cup [1;2)$ zbiorem punktów skupienia jest $[1;2]$, a dla $\mathbb{R}$ zbiorem tym jest $\overline{\mathbb{R}}$. Przejdźmy zatem do samej definicji granicy funkcji.

Definicja (Heinego). *Niech $f : D \to \mathbb{R}$, a — p.s. D oraz niech $g \in \overline{\mathbb{R}}$. Wówczas g jest* ***granicą** f **w punkcie*** *a wtw dla dowolnego $\{x_n\}_{n\geqslant n_0}$ w. $D \setminus \{a\}$ takiego, że $x_n \to a$ zachodzi*

$$f(x_n) \to g.$$

Granicę tę oznaczamy przez

$$\lim_{x\to a} f(x),$$

ponadto fakt, że g jest granicą f w punkcje a oznaczamy także przez

$$\lim_{x\to a} f(x) = g \qquad \textit{lub} \quad f(x) \xrightarrow[x\to a]{} g$$

(ewentualnie $f(x) \to g$, gdy a jest jedynym p.s. D).

Oczywiście (jak wynika z faktu 1 ze str. 19), jeśli istnieje granica f w punkcie a, to jest ona wyznaczona jednoznacznie.

Przykłady (bardzo proste). Niech $f, g, h : \mathbb{R} \to \mathbb{R}$ będą zadane dla dowolnego $x \in \mathbb{R}$ wzorami

$$f(x) = c, \qquad g(x) = \begin{cases} c & \text{dla } x \neq 0 \\ d & \text{dla } x = 0 \end{cases}, \qquad h(x) = x,$$

gdzie c, d — ustalone liczby. Wówczas dla dowolnego $a \in \overline{\mathbb{R}}$ mamy oczywiście

$$\lim_{x\to a} f(x) = c = \lim_{x\to a} g(x), \quad \lim_{x\to a} h(x) = a.$$

Warto zwrócić uwagę na fakt, że wybór liczby d nie miał w powyższym przykładzie **żadnego** wpływu na wartość $\lim_{x\to a} g(x)$ nawet wtedy, gdy $a = 0$. Ogólnie bowiem, jak widać natychmiast z definicji granicy, wybór wartości $f(a)$ (gdy $a \in D$, co ogólnie nie musi wcale zachodzić) nie wpływa ani na fakt istnienia granicy f w punkcie, ani na jej wartość.

---

[43] Będziemy używać skrótu: $\{x_n\}_{n\geqslant n_0}$ w. X, który oznacza, że $\{x_n\}_{n\geqslant n_0}$ jest ciągiem o **w**yrazach w zbiorze X (tzn. funkcją z $\mathbb{N}_{n_0}$ w X). Inaczej: w. = „o wyrazach w”.

[44] Patrz str. 50.

Poniższa uwaga wyjaśnia sprawę ewentualnej niejednoznaczności pojęcia granicy w przypadku ciągu, który przecież także jest funkcją...

Uwaga. Jeżeli $f : \mathbb{N}_{n_0} \to \mathbb{R}$, to pojęcie granicy funkcji f w $+\infty$ pokrywa się z pojęciem granicy ciągu $f = \{f(n)\}_{n \geqslant n_0}$ z rozdziału II (zatem nie ma też kolizji oznaczeń „lim" i „$\to$").

By tę uwagę wykazać wystarczy skorzystać z twierdzenia o granicy uogólnionego podciągu (twierdzenie II.5).

Notacja związana z pojęciem granicy funkcji może jednak sprawiać pewne kłopoty. Związane jest to z faktem, że granica, o ile istnieje, jest jednoznacznie wyznaczona przez funkcję oraz punkt „w którym granica jest rozważana". Zatem optymalnym oznaczeniem byłoby np. „$\lim_a f$" w miejsce tradycyjnego „$\lim_{x\to a} f(x)$". A tradycyjny zapis, przez to, że pojawia się tam nie samo f, ale „jakieś $f(x)$", w naturalny sposób zachęca do zastępowania owego „$f(x)$" konkretnym wzorem, którym funkcja może być zadana. Ale co oznacza np. napis:

$$\lim_{m\to+\infty}(m - [m])\,?$$

Czy chodzi tu o granicę ciągu? Wtedy jest to ciąg zerowy i granica równa 0. Ale może chodzi o granicę funkcji $f : \mathbb{R} \to \mathbb{R}$ zadanej wzorem $f(m) = m - [m]$ da dowolnego $m \in \mathbb{R}$? A ta „niestety" nie istnieje! (dlaczego?). Problem polega więc na tym, że podany jest wzór, zamiast funkcji. Wiemy więc tylko, jakim wzorem funkcja ta jest zadana, ale nie wiemy na jakiej dziedzinie. Aby zbytnio nie odchodzić od tradycyjnej notacji, możemy w takich wieloznacznych sytuacjach pisać:

$$\lim_{x\to a,\, x\in D} \mathit{wz\acute{o}r}(x).$$

Jednak tak będziemy postępować tylko sporadycznie. Raczej będziemy liczyli na to, że znaczenie tego typu symbolu będzie jasne z kontekstu jego użycia. Jednocześnie będziemy się starali sprawę wyboru pomiędzy ciągiem a „nieciągiem" [45] rozstrzygnąć poprzez użycie odpowiedniej zmiennej: dla ciągu będziemy raczej rezerwować litery: n, m, k, a dla innych funkcji: x, y, z.

Przyjęta przez nas definicja granicy funkcji odwołuje się do zdefiniowanej już wcześniej granicy ciągu. Jednak nie było to konieczne, można było użyć tzw. definicji Cauchy'ego — pewnego warunku nie odwołującego się wcale do ciągów, równoważnego warunkowi z definicji Heinego.

Twierdzenie IV.1. *Niech $f : D \to \mathbb{R}$, a — p.s. D, $g \in \overline{\mathbb{R}}$. Następujące trzy warunki są równoważne:*

(i) $\lim_{x\to a} f(x) = g$

(ii) *dla dowolnego ściśle monotonicznego $\{x_n\}_{n\geqslant n_0}$ w $D \setminus \{a\}$ takiego, że $x_n \to a$ zachodzi $f(x_n) \to g$.*

(iii) *(„definicja" Cauchy'ego)*
przypadek 1. — dla $a, g \in \mathbb{R}$:

$$\forall_{\epsilon>0}\ \exists_{\delta>0}\ \forall_{x\in D\setminus\{a\}}\ (|x-a| < \delta \Rightarrow |f(x) - g| < \epsilon)$$

przypadek 2. — dla $a \in \mathbb{R}, g = +\infty$:

$$\forall_{M\in\mathbb{R}}\ \exists_{\delta>0}\ \forall_{x\in D\setminus\{a\}}\ (|x-a| < \delta \Rightarrow f(x) > M)$$

[45] Choć dla „nieciągu" wybór dziedziny pozostaje nieraz wystarczająco duży by niejednoznaczność nadal miała miejsce. Np. w powyższym przykładzie zupełnie co innego uzyskujemy dla dwóch różnych dziedzin $D_2 := \{\frac{1}{2} + n : n \in \mathbb{N}\}$, $D_3 := \{\frac{1}{3} + n : n \in \mathbb{N}\}$.

przypadek 3. — dla $a = -\infty, g \in \mathbb{R}$:

$$\forall_{\epsilon>0}\ \exists_{M\in\mathbb{R}}\ \forall_{x\in D}\ (x < M \Rightarrow |f(x) - g| < \epsilon)$$

... itd. — w sumie należałoby wypisać 9 przypadków obejmujących wszystkie możliwości dla par a, g będących (niezależnie) w $\mathbb{R}$ lub $+\infty$ lub $-\infty$. Liczę, że na podstawie tych trzech przypadków, każdy ze Słuchaczy będzie w stanie wypisać dowolny z pominiętych.

Dowód tego twierdzenia najłatwiej przeprowadzić dowodząc implikacji $(ii) \Rightarrow (iii)$ oraz $(iii) \Rightarrow (i)$, co dzięki oczywistości $(i) \Rightarrow (ii)$ da potrzebne równoważności. Szczegóły dowodu pomijam i zostawiam jako zadanie.

Odpowiednikiem twierdzenia o rachunkowych własnościach granicy ciągu (tw. II.1) jest twierdzenie poniższe:

Twierdzenie IV.2 (o rachunkowych własnościach granicy funkcji). *Niech $f_1, f_2 : D \to \mathbb{R}$, a — p.s. D oraz niech $\circledast$ oznacza jedno z działań $+, -, \cdot, :$. Załóżmy, że $\lim_{x\to a} f_j(x) = g_j$ dla $j = 1, 2$, gdzie $g_1, g_2 \in \overline{\mathbb{R}}$ i że działanie $g_1 \circledast g_2$ jest określone oraz, w przypadku gdy $\circledast$ jest dzieleniem, że $\forall_{x\in D}\ f_2(x) \neq 0$. Wówczas $\lim_{x\to a}(f_1 \circledast f_2)(x) = g_1 \circledast g_2$.*

Dla porządku należy jeszcze wyjaśnić, że działanie $\circledast$ „na funkcjach" określone jest naturalnym wzorem $(f_1 \circledast f_2)(x) := f_1(x) \circledast f_2(x)$ dla $x \in D$.

Dowód.
Wystarczy użyć definicji Heinego granicy i twierdzenia II.1. □

W powyższym twierdzeniu nie wspomnieliśmy o jeszcze jednej ważnej operacji dla funkcji, a mianowicie o *złożeniu funkcji* oznaczanym przy pomocy symbolu $\circ$. Przypominamy, że jeżeli $f : X \to Y$, $g : Y \to Z$, to $g \circ f : X \to Z$ zadana jest wzorem

$$(g \circ f)(x) := g(f(x))$$

dla dowolnego $x \in X$. Odpowiednie twierdzenie „o granicy złożenia" (zwane też twierdzeniem „o podstawianiu") ma treść nietrudną do odgadnięcia, choć (uwaga!) pojawia się tam pewien „haczyk". Sprawy związane z tym twierdzeniem można odnaleźć w zadaniach do rozdziału IV. Warto tu wspomnieć, że samo twierdzenie można traktować jako uogólnienie twierdzenia II.5 (o granicy uogólnionego podciągu).

Operacją nieco przypominającą branie podciągu danego ciągu jest w ogólnym przypadku funkcji operacja *obcięcia* [46] Przypomnijmy, że *obcięcie funkcji $f : X \to Y$ do zbioru $X' \subset X$* oznaczamy symbolem $f|_{X'}$ oraz że $f|_{X'}: X' \to Y$, przy czym dla $x \in X'$ po prostu

$$(f|_{X'})(x) := f(x).$$

Za analog (choć nie uogólnienie) twierdzenia „o granicy podciągu" można by więc uznać fakt następujący, całkiem oczywisty z definicji granicy.

Fakt. *Niech $f : D \to \mathbb{R}$, $D' \subset D$ oraz a — p.s. D'. Jeżeli $\lim_{x\to a} f(x) = g$, to $\lim_{x\to a}(f|_{D'})(x) = g$.* [47]

Znacznie jednak ważniejsze jest następujące wzmocnienie powyższego faktu.

Twierdzenie IV.3 (o „scalaniu" [48]). *Jeżeli $f : D \to \mathbb{R}$, $D_1, D_2 \subset D$, a — p.s. D_1 i D_2 oraz $D \setminus \{a\} = (D_1 \cup D_2) \setminus \{a\}$, to $\lim_{x\to a} f(x) = g$ wtw dla $j = 1$ i dla $j = 2$ zachodzi $\lim_{x\to a}(f|_{D_j})(x) = g$.*

---

[46] Nie ma tu pełnej analogii — np. przy braniu podciągu dziedzina nie zmienia się, przy obcinaniu — owszem.

[47] Zgodnie z wcześniejszym ustaleniem moglibyśmy pisać $\lim_{x\to a,\, x\in D'} f(x)$ zamiast $\lim_{x\to a}(f|_{D'})(x)$.

[48] Nazwę tę zapożyczyłem od pana Michała Krycha (patrz też zadanie II.13).

Dowód.
Oczywisty, jeśli użyć definicji Cauchy'ego (tzn. tw. IV.1)... □

Oprócz zdefiniowanego już pojęcia granicy funkcji rozważa się także tzw. *granice jednostronne funkcji.* Można je zdefiniować powtarzając z odpowiednimi modyfikacjami definicję dla „zwykłej" granicy, albo któryś z warunków jej równoważnych z twierdzenia IV.1. My jednak postąpimy inaczej. Dla $D \subset \mathbb{R}$ oraz $a \in \overline{\mathbb{R}}$ oznaczmy

$$D^a_{+(-)} := \{x \in D : x > (<) a\}$$

(dla $a = 0$ upraszczamy to do D_+, D_-). Inaczej mówiąc, D^a_+ to część zbioru D położona na prawo od punktu a, a D^a_- — na lewo od a. W szczególności $D^{+\infty}_+ = D^{-\infty}_- = \emptyset$, $D^{+\infty}_- = D^{-\infty}_+ = D$ oraz $\mathbb{R}_+ = (0; +\infty)$, $\mathbb{R}_- = (-\infty; 0)$.

Definicja. *Niech $f : D \to \mathbb{R}$, oraz a — p.s. $D^a_{+(-)}$. Jeżeli istnieje $\lim_{x\to a}(f\,|_{D^a_{+(-)}})(x)$, to nazywamy ją* ***granicą prawostronną (lewostronną)*** *f w punkcie a i oznaczamy*

$$\lim_{x\to a+(-)} f(x).$$

Obie nazywamy ***granicami jednostronnymi*** *f w punkcie a.*

A zatem granice jednostronne to szczególne przypadki zdefiniowanej na początku rozdziału granicy funkcji, tyle, że rozważanej na ewentualnie zmniejszonej dziedzinie. Nie ma zatem potrzeby dowodzenia osobnych analogów „jednostronnych" wszystkich formułowanych wcześniej lub dopiero w przyszłości twierdzeń dot. „zwykłych" granic. Po prostu należy te „zwykłe" twierdzenia zastosować do funkcji obciętych do odpowiednich zbiorów D^a_+ lub D^a_-. Szczególnym przypadkiem twierdzenia o „scalaniu" jest

Wniosek. *Jeżeli $f : D \to \mathbb{R}$ oraz a — p.s. D^a_+ i D^a_-, to $\lim_{x\to a} f(x) = g$ wtw $\lim_{x\to a+} f(x) = g = \lim_{a\to a-} f(x)$.*

W przypadku, gdy posługiwaliśmy się „zmienną całkowitą n" często używaliśmy skrótu d.d.d. n. Jak to przenieść na przypadek ogólniejszy "zmiennej x ze zbioru D" i punktu skupienia a zbioru D? Zrobimy to następująco. Termin: *dla $x \in D$ dostatecznie bliskich a* (w skrócie zapiszemy: d. $x \in D$ d.b. a) będzie odtąd tym samym co:

- $\exists_{\delta>0}\ \forall_{x\in D,\, 0<|x-a|<\delta}$, gdy $a \in \mathbb{R}$,
- $\exists_{M\in\mathbb{R}}\ \forall_{x\in D,\, x>(<)M}$, gdy $a = +\infty\ (-\infty)$.

Dopuszczamy tu jednak dowolność szyku zdania, podczas gdy w wersji z kwantyfikatorami, kwantyfikatory muszą być zawsze na początku zdania.

Jak widzieliśmy np. w przypadku twierdzenia IV.2 (o rachunkowych własnościach granicy funkcji), twierdzenia dotyczące granic ciągów miewają nierzadko swe naturalne uogólnienia obowiązujące dla granic funkcji. Tak jest również w przypadku kilku innych twierdzeń z rozdziału II.

Na użytek poniższych twierdzeń przyjmujemy, że $D \subset \mathbb{R}$, $a \in \overline{\mathbb{R}}$, a — p.s. D, $f, g, h : D \to \mathbb{R}$ oraz $c, d \in \overline{\mathbb{R}}$.

Twierdzenie IV.4 (o trzech (ew. dwóch) funkcjach). *Jeżeli*

$$f(x) \leqslant g(x) \leqslant h(x)$$

d. $x \in D$ d.b. a oraz

$$\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = c,$$

to $\lim_{x\to a} g(x) = c$. Gdy $c = +\infty$ $(-\infty)$, to założenia dotyczące funkcji h (funkcji f) można pominąć.

Twierdzenie IV.5 (o zachowaniu nierówności przy przejściu granicznym). *Jeśli* $f(x) \leqslant g(x)$ *d.* $x \in D$ *d.b.* a *oraz* $\lim_{x\to a} f(x) = c$ *i* $\lim_{x\to a} g(x) = d$*, to* $c \leqslant d$*.*

Twierdzenie IV.6 (o warunku Cauchy'ego dla funkcji). *Funkcja* f *posiada skończoną*[49] *granicę w* a *wtw*

- $\forall_{\epsilon>0}\ \exists_{\delta>0}\ \forall_{x,y\in D\setminus\{a\}}\ (|x-a|, |y-a| < \delta \Rightarrow |f(x)-f(y)| < \epsilon)$*, gdy* $a \in \mathbb{R}$*;*
- $\forall_{\epsilon>0}\ \exists_{M\in\mathbb{R}}\ \forall_{x,y\in D\setminus\{a\}}\ (x, y > (<) M \Rightarrow |f(x)-f(y)| < \epsilon)$*, gdy* $a = +\infty$ $(-\infty)$*.*

Fakcik. $f(x) \xrightarrow[x\to a]{} 0$ *wtw* $|f(x)| \xrightarrow[x\to a]{} 0$*.*

Twierdzenie IV.7 (o granicach jednostronnych funkcji monotonicznej). *Jeżeli* f *jest monotoniczna oraz* a *— p.s.* D_+^a (D_-^a)*, to istnieje prawostronna (lewostronna) granica* f *w punkcie* a*.*

Dowody.

Powyższe twierdzenia sprowadzają się łatwo do odpowiednich twierdzeń z rozdziału II. Dla twierdzeń IV.4, IV.5 i fakciku jest to całkiem proste. W przypadku twierdzenia IV.6 do implikacji „$\Rightarrow$" łatwo użyć po prostu twierdzenia IV.1 (definicji Cauchy'ego). Natomiast przy dowodzie „$\Leftarrow$", używając twierdzenia II.7 (o zupełności $\mathbb{R}$) łatwo możemy wykazać, że dla dowolnego $\{x_n\}$ w. $D \setminus \{a\}$ takiego, że $x_n \to a$, ciąg $\{f(x_n)\}$ jest zbieżny. Pozostaje wykazać, że granica $\{f(x_n)\}$ jest taka sama dla wszystkich rozważanych $\{x_n\}$. Jak to wykazać? — pozostawiam to Państwu ...(nietrudne!). Dla dowodu twierdzenia IV.7 można najpierw skorzystać z warunku „równoważnego" (ii) z twierdzenia IV.1, dzięki czemu będziemy mieli do czynienia jedynie z ciągami $\{f(x_n)\}$, które sa monotoniczne (dlaczego?). Gdy zatem skorzystamy z twierdzenia II.4 (o granicy ciągu monotonicznego), do zakończenia dowodu pozostanie rozwiązanie podobnego problemu, co przy dowodzie twierdzenia IV.6. □

Na zakończenie podrozdziału dotyczącego granicy funkcji podamy przykłady kilku ważnych granic funkcji. Wykazanie części z podanych tu równości będzie zadaniem dla Państwa (m. in. na ćwiczenia).

Przykłady.

1. $\lim_{x\to+\infty}(1+\frac{1}{x})^x = e$. Nie chodzi tu oczywiście o granicę ciągu o wyrazach $(1+\frac{1}{n})^n$ znanego Państwu z rozdziału II, ale o granicę funkcji określonej np. na $\mathbb{R}_+$. Jednak dość łatwo wykazać powyższą równość, korzystając właśnie z tego, że e jest granicą powyższego ciągu.

2. $\lim_{x\to+\infty} \frac{x^\alpha}{c^x} = 0$ dla $\alpha \in \mathbb{R}$, $c > 1$. Tu, podobnie jak w poprzednim przykładzie, można wykorzystać zbieżność ciągu $\frac{n^\alpha}{c^n} \to 0$ (str. 22). Pomocne będą też informacje o monotoniczności funkcji wykładniczej i potęgowej (patrz wniosek str. 14).

3. $\lim_{x\to 0} a^x = 1$ dla $a > 0$. To oczywiście uogólnienie znanego nam faktu: $\sqrt[n]{a} \to 1$ (patrz przykład e) strona 22). Dzięki twierdzeniu IV.7 wiemy, że istnieją granice jednostronne w punkcie 0 — oznaczmy je odpowiednio: g_-, g_+. Zatem z definicji granicy funkcji mamy $\sqrt[n]{a} = a^{\frac{1}{n}} \to g_+$, skąd $g_+ = 1$ oraz $\sqrt[n]{\frac{1}{a}} = a^{-\frac{1}{n}} \to g_-$, czyli $g_+ = g_- = 1$. Zatem z wniosku ze strony 46 uzyskujemy potrzebną równość.

4. $\lim_{x\to 0} \frac{\exp(x)-1}{x} = 1$

5. $\lim_{x\to 0} \frac{\sin(x)}{x} = 1$

[49] Tzn. rzeczywistą (należącą do $\mathbb{R}$).

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

7. $\lim_{x\to 1} \frac{x^\alpha-1}{x-1} = \lim_{x\to 0} \frac{(1+x)^\alpha-1}{x} = \alpha$ dla dowolnego $\alpha \in \mathbb{R}$. Ten ostatni przykład można uzyskać z przykładu 4. — a jak? To zostawiam Państwu jako ćwiczenie do zrobienia już po przejściu przez podrozdział 5.1 (patrz też zadanie IV.3).

Równości z przykładów 4, 5 i 6 zostaną wkrótce wyjaśnione w oparciu o tzw. *szeregi potęgowe.*

2. Ciągłość funkcji w punkcie

Niech $f : D \to \mathbb{R}$ i niech $a \in D \subset \mathbb{R}$.

Definicja (Heinego). *Funkcja f* ***jest ciągła w (punkcie)*** *a wtw dla dowolnego $\{x_n\}$ w. D takiego, że $x_n \to a$ zachodzi $f(x_n) \to f(a)$.*

Zauważmy, że ta definicja bardzo przypomina definicję (Heinego) granicy. Ale istotne różnice są takie: tu zamiast granicy g jest $f(a)$ oraz tu $a \in D$, ale za to a nie musi koniecznie być punktem skupienia D. Ponadto może tu zachodzić $x_n = a$ dla pewnego n.

Czasami używa się alternatywnej definicji — tzw. definicji Cauchy'ego, która z kolei przypomina „definicję" Cauchy'ego granicy funkcji (patrz tw. IV.1 war. (iii)). Skoro jednak my zdecydowaliśmy się na definicję ciągłości w wersji Heinego, ta alternatywna definicja będzie dla nas już twierdzeniem.

Twierdzenie IV.8. *Następujące warunki są równoważne:*

(i) *f jest ciągła w a,*

(ii) *jeżeli a — p.s. D, to $\lim_{x\to a} f(x) = f(a)$,*

(iii) *(„definicja" Cauchy'ego)*

$$\forall_{\epsilon>0}\, \exists_{\delta>0}\, \forall_{x\in D}\ (|x-a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon).$$

Dowód.
(i) $\Rightarrow$ *(ii)* wynika natychmiast z obu definicji (Heinego) dla ciągłości i dla granicy. *(iii)* $\Rightarrow$ *(i)* — to z kolei bardzo łatwo wykazać z definicji granicy ciągu. Wystarczy więc wykazać, że *(ii)* $\Rightarrow$ *(iii)*. Załóżmy więc *(ii)*. Jeżeli a — p.s. D, to *(iii)* wynika bezpośrednio z twierdzenia IV.1 (patrz przypadek 1 w (iii)). Jeżeli a — nie jest p.s. D, to istnieje takie $\delta > 0$, że w przedziale $(a-\delta; a+\delta)$ nie ma żadnego elementu zbioru D poza a. Stąd by wykazać *(iii)* wystarczy dla dowolnego $\epsilon > 0$ dobrać tę właśnie liczbę δ. □

Wniosek. *Jeżeli $a \in D$, ale a nie jest p.s. D, to f jest ciągła w a.*

W szczególności np. dowolny ciąg $\{x_n\}_{n\geqslant n_0}$ jest funkcją ciągłą w każdym z punktów $n \in \mathbb{N}_{n_0}$. Oczywiście jednak, tak naprawdę, ciągłość w punkcie jest interesująca wyłącznie dla tych punktów $a \in D$, które są p.s. dziedziny funkcji.

Informacja o tym, że jakaś funkcja f jest ciągła w punkcie a bywa bardzo wygodna. Każdy taki przypadek ciągłości pozwala nam bowiem sformułować następujący „fakcik", będący po prostu przeformułowaniem definicji ciągłości w punkcie:

Jeżeli $x_n \in D$ d.d.d. n, to $x_n \to a \Rightarrow f(x_n) \to f(a)$.

A więc, inaczej mówiąc, w takiej sytuacji możemy „mechanicznie" obie strony symbolu $x_n \to a$ „obłożyć" funkcją f. To pozwala nam na poważne rozszerzenie zasobu ciągów, dla których będziemy w stanie znaleźć granicę. Jednak pod jednym warunkiem — musimy znać jakieś funkcje ciągłe w pewnych punktach. I to możliwie dużo ... Tą praktyczną sprawą zajmiemy się w następnych podrozdziałach.

Teraz natomiast przykład całkiem negatywny ...

Przykład (funkcja Dirichleta). Niech $f : \mathbb{R} \to \mathbb{R}$ zadana będzie wzorem

$$f(x) = \begin{cases} 1 & \text{dla } x \in \mathbb{Q} \\ 0 & \text{dla } x \notin \mathbb{Q}. \end{cases}$$

Łatwo dla dowolnego $a \in \mathbb{R}$ skonstruować dwa ciągi: $\{x_n\}$ — o wyrazach wymiernych i $\{x'_n\}$ — o wyrazach niewymiernych takie, że $x_n, x'_n \to a$ (patrz zadanie II.18). Ponieważ

$$1 \equiv f(x_n) \to 1 \neq 0 \leftarrow f(x'_n) \equiv 0$$

zatem w każdym punkcie $a \in \mathbb{R}$ funkcja f jest nieciągła (tj. nie jest ciągła w a).

3. Funkcje ciągłe

Zdefiniowaliśmy już ciągłość w punkcie. Teraz zajmiemy się ciągłością funkcji „w ogóle". Dość często, gdy mowa o funkcjach ciągłych, można usłyszeć następującą „intuicyjno-geometryczną definicję":

Funkcja ciągła to taka funkcja, której wykres można narysować bez odrywania ołówka.

Jeśli chodzi o intuicję związaną z ciągłością, to powyższe stwierdzenie bywa czasem użyteczne, choć nawet pomijając kwestię ścisłości, trudno uznać je za stwierdzenie prawdziwe. Istnieją bowiem tak dziwne funkcje ciągłe, których wykresu z pewnością nie dałoby się naszkicować nawet z grubsza... Tymczasem ścisła definicja jest taka:

Definicja. *Funkcja jest ciągła wtw jest ciągła w każdym punkcie swojej dziedziny.*

Natychmiastowym wnioskiem z twierdzenia o rachunkowych własnościach granicy funkcji (tw. IV.2) oraz z tw. IV.8 jest następujący wynik.

Fakt IV.1. *Suma, iloczyn, różnica i iloraz funkcji ciągłych (w przypadku ilorazu zakładamy, że funkcja przez którą dzielimy ma wszystkie wartości różne od* 0*) jest ciągła.*

Bezpośrednio z definicji ciągłości wynika natomiast „zamkniętość" klasy funkcji ciągłych na jeszcze jedną operację.

Fakt IV.2. *Złożenie funkcji ciągłych jest funkcją ciągłą.*

Inną operacją na funkcjach „nie psującą" ciągłości jest np. obcinanie funkcji do mniejszej dziedziny.

Uwaga. Analogiczne fakty dotyczące ciągłości w ustalonym punkcie są oczywiście także prawdziwe.

Przykłady.

1. Wielomian to dowolna funkcja zadana wzorem

$$w(x) := \sum_{k=0}^{n} a_k x^k$$

dla $x \in D \subset \mathbb{R}$. ($a_0, \ldots, a_n$ — ustalone liczby). Gdy $a_n \neq 0$ to n nazywa się *stopniem wielomianu* w. Ponieważ z definicji funkcja identycznościowa $\mathbb{x} : D \to \mathbb{R}$ ($\mathbb{x}(x) = x$ dla $x \in D$) jest oczywiście ciągła i podobnie funkcja stała, zatem z faktu 1 wynika ciągłość dowolnego wielomianu.

2. Funkcja wymierna to dowolny iloraz dwóch wielomianów zadanych na wspólnej dziedzinie D, na której wielomian z mianownika nie osiąga wartości 0. Z faktu 1 taka funkcja też jest ciągła.

3. Moduł: $|\cdot| : \mathbb{R} \to \mathbb{R}$, tzn. f taka, że $f(x) = |x|$ też jest funkcją ciągłą na mocy tw. IV.8 i nierówności
$$||x| - |y|| \leqslant |x - y|,$$
wynikającej łatwo z nierówności trójkąta.

Inne przykłady funkcji ciągłych podamy już wkrótce. Wcześniej sformułujemy kilka istotnych ogólnych twierdzeń opisujących własności funkcji ciągłych. Jednak najpierw wprowadźmy oznaczenie przedziałów o końcach a, b wygodne w sytuacji, gdy nie wiemy który z nich jest początkiem, a który końcem przedziału: dla $a, b \in \mathbb{R}$

$$(a?b) := \begin{cases} (a; b) \text{ gdy } a \leqslant b \\ (b; a) \text{ gdy } b < a \end{cases}$$

i analogicznie dla innego typu przedziałów: $[a?b)$, $[a?b]$ i $(a?b]$.

Dowodami poniższych trzech twierdzeń zajmiemy się po sformułowaniu trzeciego z nich. We wszystkich zakładamy, że $a, b \in \mathbb{R}$, $a \leqslant b$.

Twierdzenie IV.9 (Bolzano o własności Darboux; o osiąganiu wartości pośrednich). *Jeżeli* $f : [a; b] \to \mathbb{R}$ *jest ciągła oraz* $y \in (f(a)?f(b))$, *to istnieje* $x \in (a; b)$ *takie, że* $f(x) = y$.

Twierdzenie IV.10 (Weierstrassa; o osiąganiu kresów). *Jeżeli* $f : [a; b] \to \mathbb{R}$ *jest ciągła to istnieją* $m; M \in [a; b]$ *takie, że* $f(m) = \inf f([a; b])$ *oraz* $f(M) = \sup f([a; b])$.[50)] *W szczególności* f *jest ograniczona.*

Z tw. IV.9 i IV.10 uzyskujemy natychmiast wniosek dotyczący „obrazu ciągłego" dowolnego przedziału. Sprecyzujmy tu, że $I \subset \mathbb{R}$ nazywamy *przedziałem* wtw

$$\forall_{a,b \in I}\ [a?b] \subset I.$$

Przedziałami są zatem np.: $\emptyset$, $\{7\}$, $\mathbb{R}$, $(143; +\infty)$. *Przedziałem domkniętym* nazywamy zbiór postaci $[a; b]$ gdzie $a, b \in \mathbb{R}$, $a \leqslant b$, a *przedziałem niezdegenerowanym* każdy przedział różny od $\emptyset$ i od przedziału jednopunktowego (postaci $\{x\}$).

Wniosek. *Jeżeli* I *— przedział oraz* $f : I \to \mathbb{R}$ *jest ciągła, to* $f(I)$ *— przedział. Jeśli ponadto* I *— przedział domknięty, to* $f(I)$ *— także przedział domknięty.*

Dla celów kolejnego twierdzenia przyjmujemy następującą definicję.

Definicja. *Funkcja* $f : D \to \mathbb{R}$ *jest* ***jednostajnie ciągła*** *wtw*

$$\forall_{\epsilon > 0}\ \exists_{\delta > 0}\ \forall_{x,y \in D}\ (|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon).$$

Powyższą definicję warto porównać ze zwykłym warunkiem ciągłości, który na mocy „definicji" Cauchy'ego można zapisać równoważnie tak:

$$\forall_{x \in D}\ \forall_{\epsilon > 0}\ \exists_{\delta > 0}\ \forall_{y \in D}\ (|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon).$$

Różnica jest zrozumiała: w warunku na jednostajną ciągłość δ dobrać trzeba uniwersalnie dla wszystkich $x \in D$, w sposób zależny jedynie od ϵ, a w warunku ciągłości δ mogła być dobierana w sposób zależny i od ϵ i od x. Stąd „jednostajność" w nazwie („**jedno** wspólne

[50)] Gdy $f : D \to X$ oraz $A \subset D$, to $f(A) := \{f(a) \in X : a \in A\}$. Zbiór ten nazywamy *obrazem* A (*przy pomocy*, ew. *względem* f).

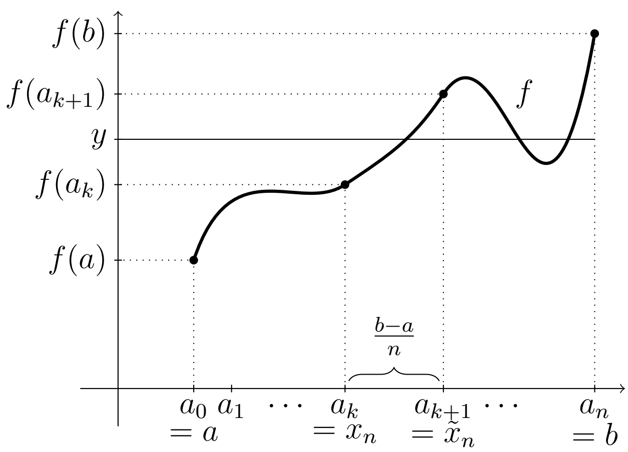
δ dla wszystkich x"). W szczególności funkcja jednostajnie ciągła jest też ciągła. Przykłady funkcji jednostajnie ciągłych to funkcja $|\cdot| : \mathbb{R} \to \mathbb{R}$, funkcja identycznościowa $\mathbb{x} : \mathbb{R} \to \mathbb{R}$, a także $f : [0; +\infty) \to \mathbb{R}$, $f(x) = \sqrt{x}$. Natomiast $g : \mathbb{R} \to \mathbb{R}$, $g(x) = x^2$, choć jest ciągła, ale jednostajnie ciągła nie jest (sprawdzenie tych własności f i g to zadania dla Państwa).

Wobec wcześniejszego wywodu „o wyższości ciągłości jednostajnej nad «zwykłą» ciągłością", następne twierdzenie może być pewnym zaskoczeniem...

Twierdzenie IV.11 (o jednostajnej ciągłości). *Jeżeli $f : [a; b] \to \mathbb{R}$ jest ciągła, to f jest jednostajnie ciągła.*

Warto jednak zwrócić uwagę na to, że ten „zaskakujący" wynik dotyczy tylko funkcji określonych na przedziałach domkniętych — zdefiniowane przed chwilą nowe pojęcie nie jest więc zapewne całkiem niepotrzebne...

Dowody (tw. IV.9, IV.10, IV.11).
Głównym „chwytem" użytym w dowodzie każdego z tych trzech twierdzeń będzie twierdzenie Bolzano - Weierstrassa (tw. II.6). Zacznijmy od dowodu tw. IV.9. Przypuśćmy, że f nie osiąga wartości y. Weźmy dowolne $n \in \mathbb{N}$ i „podzielmy" przedział $[a; b]$ na n-części o równych długościach $d_n := \frac{b-a}{n}$ punktami $a_0 = a, a_1, \dots, a_n = b$, tzn. $a_k := a + kd_n$ dla $k \in \{0, \dots, n\}$. Wśród dwóch liczb $f(a_0)$, $f(a_n)$ jedna jest większa od y, a druga mniejsza od y — w takiej sytuacji mówimy, że liczby te są *po przeciwnych stronach* y. Zatem musi istnieć takie $k \leqslant n - 1$, że liczby $f(a_k)$, $f(a_{k+1})$ są po przeciwnych stronach y (patrz rys. 4).



Rysunek 4: Liczby $f(a_k)$, $f(a_{k+1})$ są po przeciwnych stronach y.

Oznaczmy przez x_n tę spośród liczb a_k, a_{k+1} dla której wartość f jest mniejsza niż y, a przez $\tilde{x}_n$ tę, dla której dla której wartość f jest większa od y. Tym sposobem określiliśmy dwa ciągi $\{x_n\}_{n \geqslant 1}$ i $\{\tilde{x}_n\}_{n \geqslant 1}$ o wyrazach w $[a; b]$ (więc ograniczone), które spełniają

$$\forall_{n \in \mathbb{N}}\ f(x_n) < y < f(\tilde{x}_n), \tag{IV.1}$$

$$\forall_{n \in \mathbb{N}}\ |x_n - \tilde{x}_n| = d_n. \tag{IV.2}$$

Korzystamy teraz z tw. Bolzano - Weierstrassa i wybieramy podciąg zbieżny $\{x_{k_n}\}_{n \geqslant 1}$ ciągu $\{x_n\}_{n \geqslant 1}$. Mamy zatem $x_{k_n} \to g$ dla pewnego $g \in \mathbb{R}$. Ponadto $g \in [a; b]$ na mocy tw. II.2. Ponieważ na mocy (IV.2) $x_n - \tilde{x}_n \to 0$, zatem mamy także $\tilde{x}_{k_n} \to g$. Dzięki ciągłości f mamy zatem

$$f(x_{k_n}) \to f(g), \qquad f(\tilde{x}_{k_n}) \to f(g),$$

skąd na mocy (IV.1), stosując ponownie tw. II.2, dostajemy $f(g) \leqslant y \leqslant f(g)$, czyli $y = f(g)$, co jest sprzeczne z naszym założeniem.

Aby wykazać tw. IV.10 wystarczy dowód części dotyczącej „sup" (część dot. „inf" uzyskamy wtedy stosując część dot. „sup" do funkcji $-f$). Niech $c \in \overline{\mathbb{R}}$ będzie równe $\sup f([a;b])$. Istnieje zatem ciąg $\{x_n\}_{n\geqslant 1}$ w. $[a;b]$ taki, że $f(x_n) \to c$ (patrz np. zadanie II.10). Wybierzmy podciąg $\{x_{k_n}\}_{n\geqslant 1}$ zbieżny, tzn. $x_{k_n} \to M$ dla pewnego $M \in [a;b]$. Stąd $f(x_{k_n}) \to f(M)$ i jednocześnie $f(x_{k_n}) \to c$ czyli $c = f(M)$.

Aby wykazać tw. IV.11 przypuśćmy, że jego teza jest fałszywa. A zatem niech $\epsilon > 0$ będzie takie, że dla dowolnego $\delta > 0$ istnieją $x, y \in [a;b]$ dla których zachodzi

$$|x-y| < \delta \text{ oraz } |f(x) - f(y)| \geqslant \epsilon.$$

W szczególności dla dowolnego $n \in \mathbb{N}$ możemy wybrać x_n, $y_n \in [a;b]$ takie, że (w powyższym bierzemy „$\delta = \frac{1}{n}$") $|x_n - y_n| < \frac{1}{n}$ oraz

$$|f(x_n) - f(y_n)| \geqslant \epsilon. \tag{IV.3}$$

Znów wybieramy zbieżny podciąg $\{x_{k_n}\}_{n\geqslant 1}$ ciągu $\{x_n\}_{n\geqslant 1}$, tzn. $x_{k_n} \to g \in [a;b]$ i wówczas (analogicznie jak w dowodzie tw. IV.9) mamy też $y_{k_n} \to g$, a stąd $f(x_{k_n}) \to f(g) \leftarrow f(y_{k_n})$, czyli $|f(x_{k_n}) - f(y_{k_n})| \to 0$ — sprzeczność z (IV.3).

Jedną z ważnych operacji na funkcjach, o której jeszcze dotąd nie wspominaliśmy jest odwracanie funkcji. Zacznijmy od przypomnienia, że $f : X \to Y$ jest *odwracalna* wtw jest „na" (tzn. $\forall_{y\in Y}\ \exists_{x\in X}\ f(x) = y$) oraz „1–1", tj. *różnowartościowa* (tzn. $\forall_{x_1,x_2\in X}\ (x_1 \neq x_2 \Longrightarrow f(x_1) \neq f(x_2))$). Dla odwracalnej funkcji $f : X \to Y$ określamy *funkcję odwrotną* $f^{-1} : Y \to X$ warunkiem

$$\forall_{x\in X, y\in Y}\ f^{-1}(y) = x \Leftrightarrow y = f(x).$$

Okazuje się, że pod pewnymi warunkami ciągłość zachowuje się także przy operacji odwracania funkcji.

Twierdzenie IV.12 (o ciągłości funkcji odwrotnej). *Jeżeli $I, Y \subset \mathbb{R}$, I — przedział oraz $f : I \to Y$ jest ciągła i odwracalna, to f^{-1} też jest ciągła.*

A zatem owym potrzebnym w założeniu warunkiem jest to, że dziedzina f jest przedziałem. W dowodzie tw. IV.12 wykorzystamy dwa lematy, które są także interesujące same w sobie. W obu z nich załóżmy, że I — przedział oraz $f : I \to \mathbb{R}$.

Lemat IV.1. *Jeżeli f jest ciągła, to f jest „1–1" wtw f jest ściśle monotoniczna.*[51)]

Dowód.
Trzeba dowieść tylko „$\Longrightarrow$". Przypuśćmy, że f nie jest ściśle monotoniczna. Nietrudno wówczas wykazać (nie potrzeba tu korzystać jeszcze z ciągłości f, zostawiam ten krok Państwu), że istnieją takie $x_1 < x_2 < x_3$ w I, że zachodzi $f(x_1) > f(x_2) < f(x_3)$ lub $f(x_1) < f(x_2) > f(x_3)$ (rys. 5).

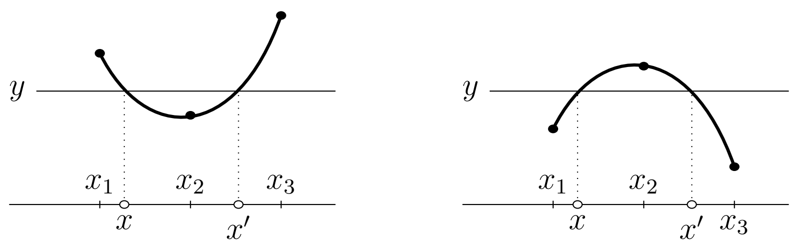
„Dla ustalenia uwagi" załóżmy, że zachodzi pierwszy z tych przypadków oraz że $f(x_1) < f(x_3)$. Wówczas niech $y \in (f(x_2); f(x_1)) \subset (f(x_2); f(x_3))$. Na mocy tw. IV.9 (własność Darboux) istnieją $x \in (x_1; x_2), x' \in (x_2; x_3)$ takie, że $f(x) = y = f(x')$, co przeczy różnowartościowości f. □

Każdy z Państwa bez trudu wskaże przykład pokazujący istotność założenia, że powyżej dziedzina f jest przedziałem. Podobna sytuacja jest też poniżej.

Lemat IV.2. *Jeżeli f jest monotoniczna, to f jest ciągła wtw $f(I)$ jest przedziałem.*

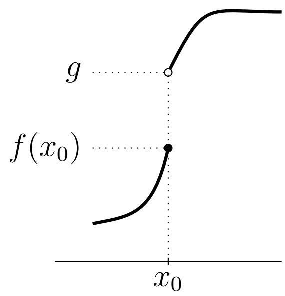
---

51) Ogólnie dla funkcji stosujemy terminologię związania z monotonicznością analogiczną do tej dla ciągów. Tzn. funkcja $f : D \to \mathbb{R}$ jest *ściśle rosnąca* (*malejąca*) wtw $\forall_{x,y\in D}\ (x < y \Longrightarrow f(x) < (>) f(y))$, a po prostu *rosnąca* (*malejąca*) gdy ostrą nierówność z prawej strony „$\Longrightarrow$" zastąpimy odpowiednią nieostrą. Łącznie na funkcje rosnące i malejące mówimy, że są *monotoniczne*. Analogicznie funkcje ściśle rosnące i ściśle malejące nazywamy *ściśle monotonicznymi*.



Rysunek 5: To typowe możliwości, gdy brak monotoniczności...

Dowód.
Implikacja „$\Longrightarrow$" wynika z wniosku po tw. IV.9 i IV.10. Załóżmy więc, że $f(I)$ — przedział. Wykażemy, że dla dowolnego $x_0 \in I$ takiego, że x_0 nie jest prawym końcem I, dla $g := \lim_{x \to x_0^+} f(x)$ (zauważmy, że ta granica istnieje na mocy tw. IV.7) zachodzi $g = f(x_0)$. Możemy założyć, że f jest rosnąca (gdy jest malejąca, rozumowanie „przejdzie" dla $-f$...).



Rysunek 6: Jeśli $f(x_0) \neq g$, to y nie będzie należało do $f(I)$.

Niech więc $x' > x_0$, $x' \in I$. Dla dowolnego x takiego, że $x_0 < x < x'$ zachodzi $f(x_0) \leqslant f(x) \leqslant f(x')$, zatem na mocy twierdzenia IV.5 mamy $f(x_0) \leqslant g \leqslant f(x')$. Gdyby zachodziło $f(x_0) \neq g$, to mielibyśmy $f(x_0) < g \leqslant f(x')$ dla wszystkich $x' \in I$ t. że $x' > x_0$ (a takie istnieją dzięki założeniu o x_0). Rozważmy wówczas jakikolwiek $y \in (f(x_0); g)$ (patrz rys. 6). Należy on do $f(I)$ (bo $f(I)$ — przedział). Z drugiej strony, skoro f rosnąca, to dla $x \in I$ takich, że $x \leqslant x_0$ mamy $f(x) \leqslant f(x_0) < y$, a z kolei dla $x' \in I$ takich, że $x' > x_0$ mamy $f(x') \geqslant g > y$. Zatem y nie może być równe $f(x)$ dla żadnego $x \in I$, tzn. $y \notin f(I)$ — sprzeczność!

Analogicznie uzyskujemy odpowiednią równość dotyczącą granicy lewostronnej, skąd łącznie uzyskamy ciągłość f. $\square$

Dowód (twierdzenia IV.12).
Z lematu 1 f jest ściśle monotoniczna zatem f^{-1} też (dlaczego? ...). Ponadto $J := f(I)$ — przedział (wniosek z tw. IV.9 i IV.10). Zatem $f^{-1} : J \to I$ oraz $f^{-1}(J) = I$ — przedział. Zatem f^{-1} jest ciągła na mocy lematu 2. $\square$

4. Szeregi potęgowe

Poznaliśmy już sporo twierdzeń opisujących własności funkcji ciągłych, ale wciąż mamy niewiele przykładów takich funkcji. Nie wiemy np. nic o ciągłości takich elementarnych funkcji jak funkcje trygonometryczne, wykładnicze, czy potęgowe. Wciąż główny nasz „pozytywny" przykład to wielomian. W tym podrozdziale, w pewnym sensie, uogólnimy pojęcie wielomianu. Mianowicie zamiast mówić o funkcji będącej „skończoną" sumą jednomianów „$a_k x^k$" (— taka właśnie była definicja wielomianu) zajmiemy się szerszą klasą funkcji zadanych „nieskończonymi" sumami jednomianów.

Definicja. ***Szeregiem potęgowym*** *nazywamy rodzinę (tzn. zbiór) wszystkich szeregów liczbowych danych wzorem*

$$\sum_{n=0}^{+\infty} a_n(x-x_0)^n \text{ [52]}$$

dla $x \in \mathbb{R}$, gdzie $\{a_n\}_{n\geqslant 0}$ jest zadanym ciągiem liczb zwanych ***współczynnikami szeregu potęgowego*** *oraz x_0 jest zadaną liczbą zwaną* ***środkiem szeregu potęgowego****. Zbiór Z złożony z tych $x \in \mathbb{R}$ dla których szereg powyższy jest zbieżny to* ***zbiór zbieżności****, a funkcja $S : Z \to \mathbb{R}$ zadana dla $x \in Z$ wzorem $S(x) := \sum_{n=0}^{+\infty} a_n(x-x_0)^n$ to* ***suma*** *tego szeregu potęgowego.*

Przykłady. Oczywiście każdy wielomian określony na $\mathbb{R}$ jest sumą szeregu potęgowego (można wziąć $x_0 = 0$ i współczynniki $a_n = 0$ d.d.d. n). Funkcje exp, sin, cos zdefiniowane pod koniec rozdziału III są również sumami szeregów potęgowych o środku 0 i zbiorze zbieżności $\mathbb{R}$. Dla exp — to jasne z definicji. Dla sin i cos sprawa jest trochę delikatniejsza, bo np. szereg $\sum_{n=0}^{+\infty}(-1)^n\frac{x^{2n}}{(2n)!}$ **nie jest** szeregiem potęgowym w rozumieniu powyższej definicji. Dlaczego więc mimo to twierdzimy, że cos jest sumą szeregu potęgowego? Odpowiedź jest prosta: można z łatwością wykazać, że zachodzi

$$\forall_{x\in\mathbb{R}}\ \cos x \text{ [53]} = \sum_{m=0}^{+\infty} a_m x^m,$$

gdzie

$$a_m := \begin{cases} 0 & \text{dla } m \text{ — nieparzystych} \\ \frac{(-1)^n}{(2n)!} & \text{dla } m = 2n \end{cases}$$

Analogicznie można postąpić dla funkcji sin.

A zatem wspomniane wcześniej uogólnienie klasy wszystkich wielomianów, to klasa wszystkich sum szeregów potęgowych.

Na ogół $Z \neq \mathbb{R}$, ale zawsze $\{x_0\} \subset Z$ — choć czasem Z to „tylko” $\{x_0\}$, jak np. dla szeregu potęgowego $\sum_{n=0}^{+\infty} n!x^n$. Pytanie zatem — na ile „dziwny” może być zbiór Z? Okazuje się, że nie może być dziwny wcale!

Twierdzenie IV.13. *Zbiór zbieżności szeregu potęgowego o środku w x_0 jest przedziałem*[54] *postaci $Z = Z_0 \cup Z_1$, gdzie $Z_0 = (x_0 - R; x_0 + R)$ dla pewnego $R \in [0; +\infty] := [0; +\infty) \cup \{+\infty\}$ oraz $Z_1 \subset \mathbb{R} \cap \{x_0 - R, x_0 + R\}$. Dla $x \in Z_0$ szereg liczbowy dany przez ten parametr x jest bezwzględnie zbieżny.*

A zatem Z to przedział, który — jeśli pominąć jego końce — jest symetryczny względem x_0. Powyższe $R \in [0; +\infty]$, będące połową długości tego przedziału, nazywane jest *promieniem zbieżności*. Gdy rozważamy szeregi potęgowe, dla których $0 < R < +\infty$, to może się zdarzyć każda z „wersji końców” przedziału zbieżności: brak końców, tylko lewy, tylko prawy, oba (zachęcam do samodzielnych poszukiwań odpowiednich przykładów). Zbiór Z_0 z powyższego twierdzenia będziemy nazywali *otwartym przedziałem zbieżności.*

Dowód (twierdzenia IV.13).
Wystarczy ograniczyć się do przypadku, gdy $x_0 = 0$. Zauważmy, że teza będzie wykazana, o ile wykażemy następujący lemat.

---

[52] Tu umowa, że $0^0 = 1$. Dla uproszczenia zapisu, gdy mowa o szeregu potęgowym $\{\sum_{n=0}^{+\infty} a_n(x-x_0)^n : x \in \mathbb{R}\}$, to najczęściej piszemy po prostu $\sum_{n=0}^{+\infty} a_n(x-x_0)^n$.

[53] W przypadku niektórych funkcji elementarnych np. sin, cos, tg, ctg (a także $\log_a$, ln, które pojawią się już wkrótce) zwyczajowo można pomijać nawias przy pisaniu argumentu. Zatem np. $\sin x = \sin(x)$.

[54] W analizie matematycznej ważną rolę odgrywają ogólniejsze szeregi potęgowe, w których zarówno x, x_0 jak i a_n mogą być liczbami zespolonymi. Wówczas Z to pewien podzbiór $\mathbb{C}$ złożony z koła otwartego $\{z \in \mathbb{C} : |z - x_0| < R\}$ i pewnego podzbioru okręgu tego koła.

Lemat. *Jeżeli szereg liczbowy* $\sum_{n=0}^{+\infty} a_n x^n$ *jest zbieżny dla pewnego* $x \in \mathbb{R}$*, to dla dowolnego* $x' \in \mathbb{R}$ *takiego, że* $|x'| < |x|$ *szereg* $\sum_{n=0}^{+\infty} a_n (x')^n$ *jest bezwzględnie zbieżny.*

Dowód.
Niech $|x'| < |x|$. Ze zbieżności $\sum_{n=0}^{+\infty} a_n x^n$ mamy $a_n x^n \to 0$, więc w szczególności dla pewnego $M \in \mathbb{R}_+$ dla dowolnego $n \geqslant 0$ zachodzi

$$|a_n x^n| \leqslant M$$

skąd $0 \leqslant |a_n (x')^n| = |a_n x^n| \cdot |\frac{x'}{x}|^n \leqslant M q^n$ dla $q := |\frac{x'}{x}| < 1$. A zatem teza lematu wynika z kryterium porównawczego (kryt. III.1). □

□

Okazuje się, że rozszerzając w taki właśnie sposób pojęcie wielomianu, jak uczyniliśmy to tutaj, nie wyszliśmy na szczęście poza klasę funkcji ciągłych.

Twierdzenie IV.14 (o ciągłości sumy szeregu potęgowego). *Suma szeregu potęgowego jest funkcją ciągłą.*

Udowodnimy tylko część powyższego twierdzenia — pominiemy trudniejszą sprawę — ciągłości w ewentualnych końcach przedziału zbieżności.

Dowód (ciągłości w punktach z Z_0).
Znów możemy założyć, że $x_0 = 0$ (dla innych x_0 trzeba złożyć sumę szeregu pot. $\sum_{n=0}^{+\infty} a_n x^n$ z funkcją „$x \longmapsto x - x_0$"). Niech R będzie promieniem zbieżności szeregu potęgowego $\sum_{n=0}^{+\infty} a_n x^n$ oraz dla $x \in (-R; R)$ niech $S(x) := \sum_{n=0}^{+\infty} a_n x^n$. Niech $t_0 \in (-R; R)$ (zatem $R > 0$). Wykażemy, że S jest ciągła w punkcie t_0. Możemy wybrać R' takie, że $|t_0| < R' < R$. Na mocy twierdzenia IV.13 szereg $\sum_{n=0}^{+\infty} a_n (R')^n$ jest zbieżny bezwzględnie, tj.

$$\sum_{n=0}^{+\infty} |a_n||R'|^n < +\infty. \tag{IV.4}$$

Niech $\epsilon > 0$. Dowód będzie zakończony (na mocy tw. IV.8 — „def. Cauchy'ego") jeżeli znajdziemy $\delta > 0$ taką, że gdy $|t - t_0| < \delta$, to $|S(t) - S(t_0)| < \epsilon$. Wskażemy takie δ, które będzie dodatkowo spełniać warunek:

$$[t_0 - \delta, t_0 + \delta] \subset (-R', R'). \tag{IV.5}$$

Najpierw korzystając z (IV.4) dobierzemy $N \in \mathbb{N}$ takie, że

$$\sum_{n=N+1}^{+\infty} |a_n||R'|^n < \frac{\epsilon}{3}. \tag{IV.6}$$

Jeśli δ spełnia (IV.5) oraz $|t - t_0| < \delta$, to $t_0, t \in (-R'; R')$ i na mocy (IV.6) mamy (zachęcam do samodzielnego **szczegółowego** uzasadnienia poniższych nierówności)

$$\begin{aligned} |S(t) - S(t_0)| &\leqslant \left|\sum_{n=0}^{N} a_n t^n - \sum_{n=0}^{N} a_n t_0^n\right| + \sum_{n=N+1}^{+\infty} |a_n||t|^n + \sum_{n=N+1}^{+\infty} |a_n||t_0|^n \\ &< \left|\sum_{n=0}^{N} a_n t^n - \sum_{n=0}^{N} a_n t_0^n\right| + \frac{2}{3} \cdot \epsilon. \end{aligned} \tag{IV.7}$$

Ponieważ $S_N : \mathbb{R} \to \mathbb{R}$ zadane wzorem $S_N(x) := \sum_{n=0}^{N} a_n x^n$ jest wielomianem, a zatem funkcją ciągłą, zatem w szczególności możemy dobrać $\delta > 0$ tak, że (IV.5) zachodzi, oraz że jeżeli $|t - t_0| < \delta$, to $|S_N(t) - S_N(t_0)| < \frac{\epsilon}{3}$. A zatem na mocy (IV.7) tak dobrane δ spełnia nasze warunki. □

Z powyższego twierdzenia uzyskujemy w szczególności ciągłość funkcji sin, cos, exp.

Przykład. Pokażemy jak można użyć twierdzenia IV.14 do obliczenia granic z przykładów 4, 5 i 6 ze strony 47. Zrobimy to na przykładzie $\lim_{x\to 0}\frac{\exp(x)-1}{x}$. Zauważmy, że dla dowolnego $x\neq 0$ zachodzi

$$\frac{\exp(x)-1}{x}=\frac{1}{x}\left(\sum_{n=0}^{+\infty}\frac{x^n}{n!}-1\right)=\frac{1}{x}\sum_{n=1}^{+\infty}\frac{x^n}{n!}=\sum_{n=1}^{+\infty}\frac{x^{n-1}}{n!}=\sum_{n=0}^{+\infty}\frac{x^n}{(n+1)!}.$$

Szereg potęgowy $\sum_{n=0}^{+\infty}\frac{x^n}{(n+1)!}$ ma zatem zbiór zbieżności równy $\mathbb{R}$, w szczególności jego suma S jest funkcją ciągłą w 0 oraz $\lim_{x\to 0}\frac{\exp(x)-1}{x}=\lim_{x\to 0}S(x)=S(0)=\frac{0^0}{1!}=1$.

Na zakończenie wspomnijmy jeszcze, że zapis f w postaci sumy szeregu potęgowego nazywany jest ogólnie *rozwinięciem f w szereg potęgowy*. Do tej sprawy będziemy jeszcze powracać w dalszych rozdziałach.

5. O kilku funkcjach elementarnych

Zajmiemy się tu kilkoma często używanymi funkcjami (niektórymi tzw. *funkcjami elementarnymi* — ten termin na ogół nie jest definiowany w jakiś jednoznaczny sposób...). Podamy niektóre ich ważne własności (część z nich bez dowodu), z których będziemy korzystali w dalszych częściach wykładu.

5.1. Funkcja wykładnicza i logarytm

Niech $a>0$. Funkcja wykładnicza (o podstawie a) zadana wzorem $W_a(x):=a^x$ dla $x\in\mathbb{R}$ zdefiniowana została już w rozdziale I.

Fakt. *W_a jest ciągła; gdy $a>1$ jej granica w $+\infty$ równa jest $+\infty$, a granica w $-\infty$ równa jest 0 i odwrotnie, gdy $a<1$. W obu przypadkach $W_a(\mathbb{R})=(0;+\infty)$ i $W_a\colon\mathbb{R}\to(0;+\infty)$ jest odwracalna.*

Dowód.
Ciągłość wynika łatwo z przykładu 3 str. 47, bowiem $\lim_{x\to x_0}W_a(x)=\lim_{h\to 0}W_a(x_0+h)=\lim_{h\to 0}a^{x_0+h}=a^{x_0}\lim_{h\to 0}a^h=a^{x_0}=W_a(x_0)$. Granice w $\pm\infty$: istnienie wynika z twierdzenia IV.7, a konkretną ich wartość uzyskamy z faktu, że $a^n\to 0$ dla $0<a<1$. Ponieważ obraz W_a musi być przedziałem, zatem ścisła monotoniczność (patrz wniosek ze strony 14) i wyliczone granice w $\pm\infty$ dają nam $W_a(\mathbb{R})=(0;+\infty)$ i odwracalność. □

Definicja. *Niech $a>0$, $a\neq 1$.* ***Logarytmem o podstawie*** *a nazywamy funkcję odwrotną do funkcji wykładniczej $W_a\colon\mathbb{R}\to(0;+\infty)$. Oznaczamy ją symbolem $\log_a$ (gdy $a=e$, to $\log_a$ nazywamy też* ***logarytmem naturalnym*** *i oznaczamy przez $\ln$).*

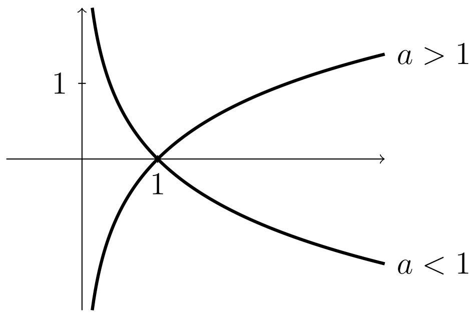
Z twierdzenia o ciągłości funkcji odwrotnej (tw. IV.12), z definicji f^{-1} i z odpowiednich własności potęgi rzeczywistej łatwo uzyskujemy następujący wynik.

Fakt. *Dla $a>0$, $a\neq 1$ funkcja $\log_a\colon(0;+\infty)\to\mathbb{R}$ jest ściśle monotoniczna i ciągła. Gdy $a>1$ jej granica w $+\infty$ równa jest $+\infty$, a w 0 równa jest $-\infty$ i odwrotnie gdy $a<1$. Ponadto dla $a,b\in(0;+\infty)\setminus\{1\}$, $x,y>0$, $\alpha\in\mathbb{R}$ zachodzi:*

(i) $\log_a(x\cdot y)=\log_a x+\log_a y$;

(ii) $\log_a(x^\alpha)=\alpha\log_a x$;

(iii) $\log_a b\cdot\log_b x=\log_a x$.



Rysunek 7: Dwa warianty wykresu funkcji logarytmicznej.

Ostatni ze wzorów pokazuje, że zamiast logarytmów o różnych podstawach, można śmiało używać jednej tylko funkcji logarytmicznej — np. ln.

Na koniec tego podrozdziału zajmiemy się zdefiniowaną w rozdziale III i rozważaną też w tym rozdziale funkcją exp. Okazuje się, że ona również jest funkcją wykładniczą.

Fakt. $\forall_{x\in\mathbb{R}}\ \exp(x) = e^x$ *(tzn.* $\exp = W_e$*).*

Dowód.
Wykazaliśmy, że zarówno funkcja exp jak i W_e są ciągłe. Ponadto obie spełniają tożsamość

$$\forall_{x,y\in\mathbb{R}}\ f(x+y) = f(x)\cdot f(y) \tag{IV.8}$$

i mają dodatnie wartości. Można też wykazać, że mają tę samą wartość w punkcie 1, tzn. że $\exp(1) = e$ — patrz. np. zadania III.17 i III.18. Reszta dowodu wynika z następującego lematu.

Lemat. *Dla każdego* $c > 0$ *istnieje dokładnie jedna funkcja ciągła* $f\colon \mathbb{R} \to (0;+\infty)$ *spełniająca* (IV.8) *oraz warunek* $f(1) = c$ *(mianowicie* W_c*).*

Dowód.
Istnienie jest jasne, bo $f = W_c$ spełnia te warunki. Wykażemy jednoznaczność. Przez indukcję „po n" łatwo dowodzimy, że dla dowolnego $x \in \mathbb{R}$, $n \in \mathbb{N}_0$ zachodzi

$$f(nx) = (f(x))^n.$$

Biorąc $\frac{x}{n}$ zamiast x w ostatniej formule, dostajemy

$$f\left(\frac{x}{n}\right) = (f(x))^{\frac{1}{n}}$$

dla $n \geqslant 1$.

Mamy też $f(-x) = (f(x))^{-1}$. Stąd

$$f\left(\frac{m}{n}\right) = f(1)^{\frac{m}{n}} = c^{\frac{m}{n}} = W_c(\frac{m}{n})$$

dla dowolnego $m \in \mathbb{Z}$, $n \in \mathbb{N}$, a zatem $f(q) = W_c(q)$ dla wszystkich $q \in \mathbb{Q}$. Dla $x \in \mathbb{R}$ weźmy ciąg $\{q_n\}$ w. $\mathbb{Q}$ taki, że $q_n \to x$ (patrz np. zadanie II.18). Wówczas z ciągłości funkcji f i W_c mamy $f(x) = \lim_{n\to+\infty} f(q_n) = \lim_{n\to+\infty} W_c(q_n) = W_c(x)$. □

□

5.2. Funkcja potęgowa

Rozważamy funkcję potęgową z potęgą $\alpha > 0$. Dla takiego α określmy $P_\alpha\colon [0;+\infty)$ [55] $\to [0;+\infty)$, $P_\alpha(x) := x^\alpha$ dla $x \geqslant 0$.

[55] W rozdziale I dziedziną P_α było zawsze $(0;+\infty)$. Tu dziedzinę tę troszkę powiększamy. Formalnie jest to więc już inna funkcja, choć używamy tego samego oznaczenia.

Fakt. *Dla $\alpha > 0$ funkcja P_α jest ciągła i ściśle rosnąca,* $\lim_{x\to+\infty} P_\alpha(x) = +\infty$*,* $P_\alpha([0; +\infty)) = [0; +\infty)$*.*

Dowód.
Ścisły wzrost dla P_α był wykazany w rozdziale I. Ciągłość w punktach $x_0 > 0$ wynika z twierdzenia o ciągłości złożenia funkcji ciągłych (fakt 2 ze strony 49) i wzoru $x^\alpha = e^{\ln(x^\alpha)} = e^{\alpha \ln x}$ dla $x > 0$ (używamy wykazanej w podrozdziale 5.1. ciągłości funkcji wykładniczej i logarytmicznej). Ciągłość w 0 wynika z informacji dotyczącej granicy ln w 0 oraz granicy W_e w $-\infty$, a granicę w $+\infty$ wyliczamy wykorzystując granice obu powyższych funkcji w $+\infty$. Z własności Darboux dostajemy zatem $f([0; +\infty)) = [0; +\infty)$. □

5.3. Funkcje trygonometryczne (sin, cos, tg, ctg)

Tu podamy tylko kilka informacji o funkcjach trygonometrycznych związanych z niezdefiniowaną dotąd przez nas liczbą π. Poniższy fakt podamy bez dowodu.

Fakt. *Istnieje liczba $\pi > 0\pi$ taka, że $\sin(\pi) = 0$ oraz $\sin x > 0$ dla $x \in (0; \pi)$.*

Uwaga. Powyższe warunki wyznaczają liczbę π jednoznacznie. Można więc uznać je za definicję liczby π. Dowód faktu nie jest trudny (patrz zadanie IV.23). Nieco trudniej jest wskazać jakieś dość precyzyjne oszacowanie liczby π przy użyciu konkretnych liczb wymiernych. Chwilowo bez dowodu przyjmiemy, że

$$3 < \pi < 4.$$

W przyszłości może uda nam się wyliczyć coś więcej na ten temat...

W oparciu o powyższy fakt oraz o wzory na sin i cos sumy i „jedynkę” trygonometryczną (patrz fakt ze strony 39) można wykazać praktycznie wszystkie znane wzory trygonometryczne (w tym tzw. wzory „redukcyjne”) dotyczące funkcji sin i cos.

W szczególności nietrudno wykazać, że $\sin(\frac{\pi}{2}) = 1$, $\cos\frac{\pi}{2} = 0$, $\cos\pi = -1$, a także, że sin i cos są funkcjami okresowymi[56] o okresie 2π.

Funkcje tg („tangens”) i ctg („kotangens”) określa się następująco: $\mathrm{tg}\colon \mathbb{R} \setminus \{k\pi + \frac{\pi}{2}\colon k \in \mathbb{Z}\} \to \mathbb{R}$,

$$\mathrm{tg}(x) := \frac{\sin x}{\cos x};$$

$\mathrm{ctg}\colon \mathbb{R} \setminus \{k\pi\colon k \in \mathbb{Z}\} \to \mathbb{R}$,

$$\mathrm{ctg}(x) := \frac{\cos x}{\sin x}.$$

Definicje te są poprawne, można bowiem łatwo sprawdzić, że zbiór zer[57] sinusa, to zbiór $\{k\pi\colon k \in \mathbb{Z}\}$, a cosinusa $\{k\pi + \frac{\pi}{2}\colon k \in \mathbb{Z}\}$. Co więcej tg i ctg są ciągłe, jako ilorazy funkcji ciągłych. Są okresowe o okresie (nawet) π (uwaga — tu dziedzina nie jest $\mathbb{R}$, więc owa okresowość jest w nieco innym znaczeniu, niż z definicji okresowości z przypisu — w jakim?) oraz $\lim_{x\to\frac{\pi}{2}\pm} \mathrm{tg}\, x = \mp\infty = \lim_{x\to 0\mp} \mathrm{ctg}\, x$.

[56] Przypomnijmy, że $f\colon \mathbb{R} \to \mathbb{R}$ jest *okresowa* wtw istnieje $T \neq 0$ takie, że $f(x + T) = f(x)$ dla dowolnego $x \in \mathbb{R}$ oraz że każda liczba $T \neq 0$ o powyższej własności nazywa się *okresem* funkcji f.

[57] *Zero funkcji* f to każdy taki x z dziedziny f, że $f(x) = 0$.

Zadania do Rozdziału IV

1. Sformułuj pominięte na wykładzie przypadki „definicji" Cauchy'ego granicy (tw. IV.1). Dla jednego wybranego (spośród dziewięciu) przypadku udowodnij twierdzenie IV.1.

$\forall$ 2.

   (a) Wykaż, że poniższe „twierdzenie" jest **fałszywe**:

   Załóżmy, że $A, B \subset \mathbb{R}$, $a, b, c \in \overline{\mathbb{R}}$ i a — p.s. A, b — p.s. B oraz, że $f\colon A \to B$ i $g\colon B \to \mathbb{R}$. Wówczas, jeżeli

   (i) $\lim_{x\to a} f(x) = b$,

   (ii) $\lim_{y\to b} g(y) = c$,

   to $\lim_{x\to a}(g \circ f)(x) = c$.

   (b) Wykaż twierdzenie „o granicy złożenia" (inaczej „o podstawianiu") powstałe przez dołożenie powyżej jeszcze jednego założenia:

   (iii) zachodzi przynajmniej jeden z warunków:
   - $b \in B$ i g jest ciągła w punkcie b;
   - $f(x) \neq b$ d. $x \in A$ d.b. a (np. gdy $b = \pm\infty$...).

3. Wykorzystując twierdzenie z zadania IV.2 (b), policzoną na wykładzie (przykład 4 ze str. 47) granicę $\lim_{x\to 0} \frac{\exp(x)-1}{x}$ oraz wiedzę z podrozdz. 5.1 wykaż, że $\lim_{x\to 1} \frac{x^\alpha-1}{x-1} = \lim_{x\to 0} \frac{(1+x)^\alpha-1}{x} = \alpha$ (przykład 7 ze str. 48).

4. Wykaż (ze szczegółami), że $\lim_{x\to+\infty} \left(1+\frac{1}{x}\right)^x = e$ (przykład 1 ze str. 47) oraz $\lim_{x\to+\infty} \frac{x^\alpha}{c^x} = 0$ dla $\alpha \in \mathbb{R}$, $c > 1$ (przykład 2 ze str. 47).

$\forall$ 5. Znajdź[58] poniższe granice, lub wykaż, że nie istnieją:

   (a) $\lim_{x\to a} \frac{x^2+x-2}{x^2-3x+2}$ dla $a = \pm\infty; \pm 2; \pm 1; 0$;

   (b) $\lim_{x\to 0} \frac{\sin(7x)}{\sin(13x)}$;

   (c) $\lim_{x\to \pi} \frac{\sin x}{x^2-\pi^2}$;

   (d) $\lim_{x\to 1} \frac{x^e-1}{1-x^\pi}$;

   (e) $\lim_{x\to 0} \frac{\sqrt[2007]{1+x}-1}{2x+x^2}$;

   (f) $\lim_{x\to+\infty}(\sin\sqrt{1+x} - \sin\sqrt{x})$;

   (g) $\lim_{x\to+\infty}(\sin(1+x) - \sin x)$;

   (h) $\lim_{x\to+\infty} \frac{\ln x}{x^\alpha}$ dla $\alpha > 0$;

   (i) $\lim_{x\to+\infty} \frac{\ln(x^{2007}+x^{2009})}{\sqrt[9999]{x}}$;

   (j) $\lim_{x\to 0} \sin\frac{1}{x}$;

   (k) $\lim_{x\to 0} x\sin\frac{1}{x}$;

   (l) $\lim_{x\to 0+} x^x$;

   (m) $\lim_{x\to 0} \frac{(\cos x)^{\sqrt{2}}-1}{x^2}$;

   (n) $\lim_{x\to 0} \frac{2^{\sin x}-1}{x}$;

   (o) $\lim_{x\to 0} \frac{a^x-1}{x}$ dla $a > 0$.

6. Dla jakich parametrów $a, b \in \mathbb{R}$ jest ciągła funkcja $f\colon \mathbb{R} \to \mathbb{R}$ zadana wzorem

   (a) $f(x) = \begin{cases} x & \text{dla } |x| \leqslant 1 \\ x^2+ax+b & \text{dla } |x| > 1, \end{cases}$

   (b) $f(x) = \begin{cases} a & \text{dla } x = 0 \\ \sin\frac{b}{|x|} & \text{dla } x \neq 0. \end{cases}$

7. Znajdź przykład funkcji $f\colon \mathbb{R} \to \mathbb{R}$, która jest ciągła w

   (a) dokładnie jednym punkcie;

[58] Obowiązkowo przynajmniej 8 szt. spośród (b)–(e), (h)–(o).

(b) dokładnie dwóch punktach;

(c) dokładnie n punktach ($n \in \mathbb{N}$, ustalone);

(d) każdym punkcie zbioru $\mathbb{Z}$, a w pozostałych punktach jest nieciągła.

8. Uzupełnij szczegóły dowodu twierdzenia „o granicach jednostronnych funkcji monotonicznej” (tw. IV.7).

9. Rozważamy funkcję $f\colon \mathbb{R} \to \mathbb{R}$ zadaną wzorem $f(x) = \begin{cases} 0 & \text{dla } x \notin \mathbb{Q} \\ \frac{1}{\operatorname{mian}(x)} & \text{dla } x \in \mathbb{Q}, \end{cases}$ gdzie $\operatorname{mian}(x) := \min\{n \in \mathbb{N}\colon \exists_{m\in\mathbb{Z}}\ x = \frac{m}{n}\}$ dla $x \in \mathbb{Q}$. Wykaż, że f jest ciągła w punkcie x wtw $x \notin \mathbb{Q}$.

10. Wykaż, że każde z poniższych równań ma co najmniej dwa pierwiastki (tzn. rozwiązania) w $\mathbb{R}$:

(a) $e^x = 1 + 2x$;

(b) $2^x = 4x$;

(c) $e^{-(x^2+\sin x)} = \frac{1}{2}$.

∀ 11. Znajdź pewną liczbę wymierną będącą przybliżeniem jakiegoś pierwiastka podanego równania z podaną dokładnością d (tzn. takie $y \in \mathbb{Q}$, że istnieje pierwiastek p równania taki, że $|y - p| \leqslant d$):

(a) $x^3 - 3x = -1$, $d = \frac{1}{10}$;

(b) $x^3 + x^2 + 2x + 1 = 0$, $d = \frac{1}{8}$.

12. Wykaż, że wielomian stopnia nieparzystego posiada pierwiastek rzeczywisty.

13. Twierdzenie o punkcie stałym , to każde twierdzenie postaci:

Jeżeli $f\colon X \to X$ oraz zachodzi Z, to istnieje $x \in X$ takie, że $f(x) = x$,

gdzie Z to pewne założenie dotyczące funkcji f i zbioru X. Wykaż twierdzenie o punkcie stałym dla każdego z poniższych założeń Z:

∀ (a) f jest ciągła, X — przedział domknięty;

(b) f jest ciągła i malejąca, $X = \mathbb{R}$;

(c) f jest zwężająca, tzn. $\exists_{c<1}\ \forall_{x,y\in X}\ |f(x) - f(y)| \leqslant c|x - y|$, $X = \mathbb{R}$.

14. Wykaż, że funkcja $f\colon \mathbb{R} \to \mathbb{R}$ zadana wzorem

$$f(x) = \frac{1000x^{14} - 7x^{11} + 12x + 7}{(x^7 - 1)^2 + 1}$$

jest ograniczona.

∀ 15. Wykaż następujące twierdzenie „o osiąganiu jednego kresu”:

Jeżeli I to niepusty przedział, $f\colon I \to \mathbb{R}$ jest funkcją ciągłą oraz istnieje $x_0 \in I$ taki, że dla dowolnego a będącego końcem przedziału I nienależącym do I zachodzi $\lim_{x\to a} f(x) > f(x_0)$ ($< f(x_0)$), to f osiąga swój kres dolny (górny).[59)]

[59)] Przez dolny lub górny kres funkcji rozumiemy oczywiście odpowiedni kres jej zbioru wartości, tzn. tu inf lub sup $f(I)$.

16. Wykaż, że funkcja f z zadania IV.14 osiąga obydwa swe kresy.

17. Funkcja $f\colon D \to \mathbb{R}$ jest *lipschitzowska* wtw $\exists_{c\in\mathbb{R}} \forall_{x,y\in D}\ |f(x)-f(y)| \leqslant c|x-y|$. Znajdź **wszystkie** te implikacje, które zachodzą pomiędzy parami zdań utworzonymi spośród: „f jest lipschitzowska”, „f jest ciągła”, „f jest jednostajnie ciągła”, niezależnie od wyboru funkcji f.

18. Zbadaj jednostajną ciągłość i lipschitzowskość funkcji zadanych poniższymi wzorami:

    (a) $\sqrt{x}$ dla $x \geqslant 0$;

    (b) x^2 dla $x \in \mathbb{R}$;

    (c) $|x|$ dla $x \in \mathbb{R}$;

    $\forall$ (d) $\ln x$ dla $x > 0$;

    $\forall$ (d') $\ln x$ dla $x \geqslant a$, gdzie $a > 0$ jest ustalone.

19. Znajdź przykład funkcji $f\colon A \to B$, gdzie $A, B \subset \mathbb{R}$, która jest ciągła i odwracalna, ale $f^{-1}\colon B \to A$ nie jest ciągła. (Wskazówka: weź B — przedział, ale A — nie).

20. Znajdź zbiór zbieżności i promień zbieżności następujących szeregów potęgowych:

    (a) $\sum_{n=0}^{+\infty} 2^n x^n$;

    (b) $\sum_{n=0}^{+\infty} \frac{1}{(n+1)^2}(x+1)^n$;

    (c) $\sum_{n=0}^{+\infty} \frac{1}{\sqrt{n+1}}(x-1)^n$;

    (d) $\sum_{n=0}^{+\infty} (2+(-1)^n)^n x^n$;

    (e) $\sum_{n=0}^{+\infty} \frac{n^{1000}}{\sqrt{n!}} x^n$;

    (f) $\sum_{n=0}^{+\infty} \frac{\sqrt{n!}}{n^{1000}} x^n$;

    (g) $\sum_{n=0}^{+\infty} \frac{1}{\ln(2+n)} x^n$.

21. Wykaż, że jeśli dwa szeregi potęgowe o środku w 0 i o dodatnich promieniach zbieżności mają sumy równe w pewnym przedziale $(-r; r)$, $r > 0$, to szeregi te są identyczne (tzn. mają te same ciągi współczynników). Powyższy fakt, to tzw. twierdzenie o jednoznaczności rozwinięcia w szereg potęgowy i jest ono uogólnieniem znanego faktu dotyczącego jednoznaczności współczynników wielomianu. Wskazówka: użyj sprytnie twierdzenia o ciągłości sumy szeregu potęgowego (tw. IV.14).

22. Wykaż, że każda funkcja $f\colon \mathbb{R} \to \mathbb{R}$ ciągła i *addytywna*, tj. spełniająca warunek

    $$\forall_{x,y\in\mathbb{R}}\ f(x+y) = f(x) + f(y),$$

    jest funkcją liniową, tzn. zadaną wzorem $f(x) = ax$ dla dowolnego $x \in \mathbb{R}$, przy pewnym (ustalonym) $a \in \mathbb{R}$.

23. Wykaż fakt o istnieniu liczby π (ze strony 58). Wskazówka: wykorzystaj zadanie III.5.

24. W oparciu o wiedzę z wykładu (tzn. definicję sin i cos, wzory z rozdziału III (fakt 2 ze str. 39) oraz fakt z zadania IV.23) wykaż następujące własności sin i cos:

    (a) $\cos \pi = -1$;

    (b) $\cos \frac{\pi}{3} = \frac{1}{2}$;

(c) sin i cos są okresowe o okresie 2π;

(d) $\sin x = 0$ wtw $x \in \{k\pi \colon k \in \mathbb{Z}\}$.

25. Wykaż, że funkcja Dirichleta (przykład ze strony 49) jest okresowa. Znajdź zbiór wszystkich jej okresów.

26. Wykaż, że jeżeli $f \colon \mathbb{R} \to \mathbb{R}$ jest ciągła i okresowa o okresie T_n dla dowolnego $n \in \mathbb{N}$, przy czym $\lim_{n\to+\infty} T_n = 0$ oraz $\forall_{n\in\mathbb{N}}\ T_n \neq 0$, to f jest stała. Czy ciągłość jest tu istotnym założeniem?

27. Wykaż, że istnieje $M \in \mathbb{R}$ takie, że dla $x \geqslant M$ zachodzi

    (a) $x^{10000} < \frac{1}{10000!}x^{10001} - 10000!\sqrt{x}x^{10000}$;

    (b) $1000! \ln x < x^{\frac{1}{1000!}}$.

28. Wykaż, że $\lim_{x\to+\infty}\left(1+\frac{y}{x}\right)^x = e^y$ dla dowolnego $y \in \mathbb{R}$ (uwaga: użyj tu **świadomie** ciągłości odpowiedniej funkcji w odpowiednim punkcie...).

29. Wykorzystując fakt, że $\lim_{x\to0}\frac{(1+x)^\alpha-1}{x} = \alpha$ (przykład 7 ze str. 48) udowodnij następujące kryterium zbieżności szeregów (Raabego): *Jeżeli $a_n > 0$ dla $n \geqslant n_0$ oraz $\lim_{n\to+\infty} n(1-\frac{a_{n+1}}{a_n}) = \alpha$,* [60] *to $\sum_{n=n_0}^{+\infty} a_n$ jest zbieżny, gdy $\alpha > 1$ oraz jest rozbieżny, gdy $\alpha < 1$.* Wskazówka: użyj również „drugiego kryterium porównawczego" z zadania III.14.

30. Wykorzystując informacje o granicach odpowiednio dobranych funkcji (w odpowiednich punktach) zbadaj zbieżność szeregów

    (a) $\sum_{n=1}^{+\infty}(\sqrt[n]{3}-1)^\alpha$

    (b) $\sum_{n=1}^{+\infty}(\sqrt[n]{n}-1)^\alpha$

    $\forall$ (c) $\sum_{n=1}^{+\infty}(1-\cos\frac{1}{n})^\alpha$

    w zależności od wartości parametru $\alpha \in \mathbb{R}$. Wskazówka: użyj kryterium asymptotycznego (kryterium III.2).

[60] Można zamiast tej granicy wziąć także $\lim_{n\to+\infty} n(\frac{a_n}{a_{n+1}} - 1)$.

V Rachunek różniczkowy

[około $3\frac{1}{2}$ wykładu]

W rozdziale IV rozważaliśmy funkcje ciągłe, czyli takie, o których można powiedzieć, że „lokalnie zachowują się w sposób dosyć regularny". Obecnie zajmiemy się funkcjami „jeszcze bardziej regularnymi", a mianowicie *różniczkowalnymi.* Popularna geometryczna „definicja" (podobnie jak w przypadku ciągłości, mocno nieścisła, ale sugestywna...) jest następująca:

Funkcja jest różniczkowalna, gdy jej wykres nie posiada „kantów".

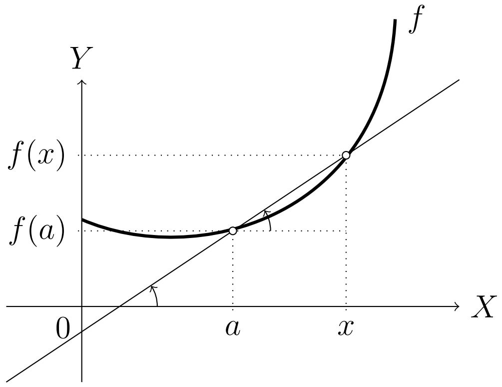
A zatem chodzi o klasę tych funkcji, z którymi w praktyce mamy do czynienia najczęściej. Rzeczywiście — ogromna większość funkcji pojawiających się przy próbach matematycznego opisu zjawisk z otaczającego nas świata — to funkcje różniczkowalne. Również samo pojęcie *pochodnej*, bezpośrednio związane z różniczkowalnością, bardzo często pojawia się przy takich opisach — np. dla wyrażenia szybkości zmian pewnych wielkości w czasie.

1. Pochodna funkcji

Rozważmy funkcję $f : D \to \mathbb{R}$, gdzie $D \subset \mathbb{R}$ oraz niech $a \in D$ i a — p.s. D. *Ilorazem różnicowym* dla funkcji f i punktu a nazywamy funkcję określoną na $D \setminus \{a\}$, zadaną dla $x \neq a$ wzorem

$$\frac{f(x) - f(a)}{x - a}.$$

Geometryczny sens powyższej wielkości jest taki: to tangens kąta utworzonego przez oś OX oraz przez prostą wyznaczoną punktami wykresu f dla argumentów a i x (patrz rys. 8). Powyższą prostą nazywa się często *sieczną* do wykresu funkcji f.



Rysunek 8: Wartość ilorazu różnicowego dla f i a w punkcie x to po prostu tangens zaznaczonego tu kąta.

Definicja.

- *Jeżeli istnieje granica w punkcie a ilorazu różnicowego dla f i a, tzn.*

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a},$$

to nazywamy ją ***pochodną*** *f* ***w (punkcie)*** *a i oznaczamy $f'(a)$.*

- ***Pochodna lewostronna (prawostronna)*** *f w punkcie a to*

$$\lim_{x\to a-(+)} \frac{f(x)-f(a)}{x-a}.\ ^{61)}$$

Oznaczamy ją $f'_-(a)$ $(f'_+(a))$.

- *Funkcja f jest* ***różniczkowalna w (punkcie)*** *a wtw $f'(a)$ istnieje i jest skończona (tj. $f'(a) \in \mathbb{R}$).*

- *Funkcja f jest* ***różniczkowalna*** *wtw $\forall_{a\in D}$ f jest różniczkowalna w punkcie a. W tej sytuacji funkcję $D \ni x \rightsquigarrow f'(x)$ [62] nazywamy* ***pochodną*** *f i oznaczamy symbolem f'.*

Wspomniany we wstępnie do niniejszego rozdziału brak „kantów" w wykresie funkcji różniczkowalnej najlepiej chyba uściślić jako istnienie *prostej stycznej* do wykresu f, rozumianej jako prosta „graniczna" prostych siecznych do wykresu „przy $x \to a$".

Definicja. *Załóżmy, że f posiada pochodną w a.* ***Prosta styczna do wykresu*** *f dla a (lub w punkcie $(a, f(a))$) to*

- *w przypadku, gdy $f'(a) \in \mathbb{R}$ zbiór*

$$\{(x, y) \in \mathbb{R} \times \mathbb{R} : y = f'(a)(x-a) + f(a)\},$$

- *gdy $f'(a) = \pm\infty$ zbiór*

$$\{(x, y) \in \mathbb{R} \times \mathbb{R} : x = a\}.$$

A zatem powinniśmy jeszcze trochę poprawić „definicję" geometryczną różniczkowalności wykluczając w niej nie tylko „kanty", ale także pionowe styczne ...

Uwagi.

1. Zamiast granicy $\lim_{x\to a} \frac{f(x)-f(a)}{x-a}$ w definicji pochodnej możemy równoważnie rozważać granicę $\lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$.

2. Zamiast oznaczenia f' na pochodną często używane jest tradycyjne oznaczenie

$$\frac{df}{dx}$$

   może dość niewygodne jako symbol funkcji, ale za to odzwierciedlające z grubsza sens pojęcia pochodnej (d oznacza tu "przyrost").

3. Często przy definicji pochodnej robione jest nieco silniejsze założenie dotyczące dziedziny D i punktu $a \in D$. Zakłada się mianowicie, że dla pewnego $\delta > 0$ zbiór

$$D_{a,\delta} := \{x \in D : |x-a| < \delta\} \tag{V.1}$$

   jest jednym z przedziałów $(a-\delta; a+\delta)$, $[a; a+\delta)$ lub $(a-\delta; a]$. Mówimy wówczas, że a *ma w D otoczenie będące przedziałem*. Założenie to jest oczywiście spełnione, jeżeli np. D jest dowolnym niezdegenerowanym przedziałem, ewentualnie skończoną sumą takich przedziałów, oraz $a \in D$. A zatem spełnione jest w przypadkach najczęściej tu przez nas rozważanych (zachęcam jednak do podania jakiegoś przykładu „negatywnego", tj. dziedziny D oraz $a \in D$ nie spełniających tego warunku).

[61] Oczywiście, by w ogóle mówić o tych *jednostronnych* pochodnych, a musi być p.s. D^a_- lub, odpowiednio, D^a_+.

[62] Symbol: $D \ni x \rightsquigarrow$ wzór(x) oznacza funkcję g określoną na D zadaną dla wszystkich $x \in D$ jako $g(x) =$ wzór(x).

4. Gdy a jest *obustronnym* punktem skupienia D, tzn. a — p.s. D^a_+ oraz D^a_-, to $f'(a)$ istnieje wtw istnieją $f'_+(a)$ i $f'_-(a)$ i są sobie równe.

5. Podobnie jak granica funkcji w punkcie, tak i pochodna funkcji w punkcie to pojęcia *lokalne*, tzn. dla dowolnego $\delta > 0$ istnienie i wartość $f'(a)$ jest tym samym co istnienie i wartość $f'_\delta(a)$, gdzie $f_\delta := f\,|_{D_{a,\delta}}$ (patrz oznaczenie z powyższego punktu 3). Mówiąc bardziej obrazowo (ale zupełnie nie ściśle ...) punkty „dalekie od a" nie mają wpływu na $f'(a)$ [63)]

Przykłady (najprostsze).

1. Funkcja stała ma w każdym punkcie skupienia swej dziedziny pochodną równą 0, bo iloraz różnicowy jest stale równy 0.

2. Jeżeli $f : \mathbb{R} \to \mathbb{R}$ jest funkcją *afiniczną* (zwaną czasem *liniową*, choć to trochę mylące) tzn. zadaną dla $x \in \mathbb{R}$ wzorem $f(x) = ax+b$ (a, b — ustalone liczby), to iloraz różnicowy dla f i dowolnego $x_0 \in \mathbb{R}$ jest funkcją stale równą a, zatem też f' jest funkcją stale równą a, co więcej wykres f jest jednocześnie prostą styczną do „siebie samego" dla x_0, niezależnie od wyboru x_0.

3. Niech $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$. Wówczas $f'(x) = 1$ dla $x > 0$, $f'(x) = -1$ dla $x < 0$, a $f'(0)$ nie istnieje, ponieważ $f'_-(0) = -1$, $f'_+(0) = 1$. Zatem to przykład funkcji ciągłej, która nie w każdym punkcie posiada pochodną.

4. Rozważmy funkcję signum $\mathrm{sgn} : \mathbb{R} \to \mathbb{R}$ zadaną wzorem [64)]

$$\mathrm{sgn}(x) = \begin{cases} 1 & \text{dla } x > 0 \\ 0 & \text{dla } x = 0 \\ -1 & \text{dla } x < 0 \end{cases}$$

   Zachodzi: $f'(x) = 0$, dla $x \neq 0$, $f'(0) = +\infty$. Jest to więc przykład funkcji nieciągłej, która w każdym punkcie posiada pochodną.

5. Niech $f : [0; +\infty) \to \mathbb{R}$, $f(x) = \sqrt{x}$. Wówczas po standardowych przekształceniach wzoru na iloraz różnicowy tej funkcji uzyskujemy $f'(x) = \frac{1}{2\sqrt{x}}$ dla $x > 0$ oraz $f'(0) = +\infty$. Jest to więc przykład funkcji ciągłej, która nie jest różniczkowalna, ale w każdym punkcie posiada pochodną.

Uzupełnieniem powyższych przykładów 3, 4 i 5 może być następujący ważny rezultat dotyczący związków pochodnej z ciągłością.

Fakt. *Jeżeli f jest różniczkowalna w punkcie x_0, to jest też w tym punkcie ciągła.*

Dowód.
Gdy $f'(x_0) \in \mathbb{R}$, to $\lim_{x\to x_0} f(x) = \lim_{x\to x_0}(\frac{f(x)-f(x_0)}{x-x_0}(x - x_0) + f(x_0)) = f'(x_0) \cdot 0 + f(x_0) = f(x_0)$. □

[63)] Pytania dla Słuchaczy/Czytelników: A zatem, czy w ogóle jakiś punkt dziedziny poza samym a ma wpływ na $f'(a)$? Może wystarczy znać tylko a i $f(a)$? ... Ostrzegam, że są to pytania dotyczące bardziej mankamentów logicznych naszego potocznego języka, niż matematyki.

[64)] Funkcja oznaczana tym symbolem nie zawsze jest tą tu właśnie zdefiniowaną funkcją. Rozmaicie bywa wybierana wartość sgn dla argumentu 0.

2. Różniczkowanie funkcji elementarnych

Zajmiemy się tu wzorami umożliwiającymi obliczanie pochodnych funkcji elementarnych. Dokładniej — zajmiemy się tymi funkcjami, które można uzyskać, wychodząc od znanych nam kilku ich podstawowych typów, przy pomocy znanych nam operacji na funkcjach. Zaczniemy od poniższych czterech wzorów na pochodne już wcześniej przez nas badanych funkcji.

Fakt.

1. *Jeśli* $\alpha \in \mathbb{R}$ *oraz* $f : D \to \mathbb{R}$ *jest zadana wzorem* $f(x) = x^\alpha$ *przy czym a)* $\alpha \in \mathbb{N}$ *i* $D = \mathbb{R}$ *lub b)* $\alpha \in \mathbb{Z}$ *i* $D = \mathbb{R} \setminus \{0\}$ *lub c)* $D = \mathbb{R}_+$, *to* $f'(x) = \alpha x^{\alpha-1}$ *(tu ewentualne „*0^0*”, mogące się pojawić w a), uznajemy za równe 1).*
2. *Jeżeli* $a > 0$ *oraz* $f : \mathbb{R} \to \mathbb{R}$ *zadana jest wzorem* $f(x) = a^x$, *to* $f'(x) = a^x \ln a$. *W szczególności* $\exp' = \exp$.
3. $\sin' = \cos$.
4. $\cos' = -\sin$.

Dowód.
Sprawdzimy wszystkie wzory na pochodne w punkcie x_0. Dla pierwszego wzoru można skorzystać z tego, że $\lim_{x\to1} \frac{x^\alpha - 1}{x-1} = \alpha$ (patrz przykład IV.7 strona 47) oraz zapisać dla $x_0 \neq 0$ i $\frac{x}{x_0} > 0$

$$\frac{x^\alpha - x_0^\alpha}{x - x_0} = x_0^{\alpha-1} \frac{(\frac{x}{x_0})^\alpha - 1}{\frac{x}{x_0} - 1},$$

co daje natychmiast potrzebny wynik, o ile $x_0 \neq 0$ (przypadek $x_0 = 0$, możliwy tylko dla a), jest oczywisty). Dla dowodu punktu 2 zauważmy, że dla $a \neq 1$ i $x \neq x_0$

$$\frac{a^x - a^{x_0}}{x - x_0} = a^{x_0} \frac{a^{x-x_0} - 1}{x - x_0} = a^{x_0} \frac{e^{(x-x_0)\ln a} - 1}{(x - x_0)\ln a} \ln a \xrightarrow[x\to x_0]{} a^{x_0} \ln a$$

ponieważ

$$\lim_{x\to0} \frac{e^x - 1}{x} = 1$$

(patrz przykład IV.4 strona 47). Dla $a = 1$ ten wzór na pochodną jest oczywisty. Wreszcie wzory 3 i 4 to prosta konsekwencja wzorów na sin i cos od sumy argumentów (patrz fakt III.2 ze str. 39) oraz przykładów 5 i 6 ze strony 47. Np.

$$\frac{\cos(x_0 + h) - \cos x_0}{h} =$$

$$\cos x_0 \frac{(\cos h - 1)}{h^2} h - \sin x_0 \frac{\sin h}{h} \xrightarrow[h\to0]{} \cos x_0 \cdot \frac{1}{2} \cdot 0 - \sin x_0 \cdot 1 = -\sin x_0.$$

□

Poniższy rezultat pozwoli nam badać funkcje zadane bardziej skomplikowanymi wzorami.

Twierdzenie V.1 (o własnościach rachunkowych pochodnej).

a. *Jeżeli* $f, g : D \to \mathbb{R}$ *są różniczkowalne w punkcie* $a \in D$, *to* $f \pm g$ *i* $f \cdot g$ *także są różniczkowalne w* a *oraz*

$$(f \pm g)'(a) = f'(a) \pm g'(a),$$

$$(f \cdot g)'(a) = f'(a) \cdot g(a) + f(a) \cdot g'(a) \qquad \textit{(wzór Leibnitza).}$$

Jeżeli ponadto $\forall_{x\in D}\ g(x)\neq 0$*, to* $\frac{f}{g}$ *jest różniczkowalna w* a *oraz*

$$\left(\frac{f}{g}\right)'(a)=\frac{f'(a)g(a)-f(a)g'(a)}{(g(a))^2}. \tag{V.2}$$

b. *Jeżeli* $A,B\subset\mathbb{R}$*,* $f:A\to B$*,* $g:B\to\mathbb{R}$ *oraz* f *jest różniczkowalna w* $a\in A$ *i* g *jest różniczkowalna w* $f(a)$*, to* $g\circ f$ *jest różniczkowalna w* a *oraz*

$$(g\circ f)'(a)=g'(f(a))f'(a).$$

c. *Jeżeli* $I,Y\subset\mathbb{R}$*,* I *— przedział oraz* $f:I\to Y$ *jest odwracalna i różniczkowalna, przy czym* $\forall_{x\in I}\ f'(x)\neq 0$*, to* f^{-1} *też jest różniczkowalna, oraz dla dowolnego* $y\in Y$

$$(f^{-1})'(y)=\frac{1}{f'(f^{-1}(y))}.$$

Dowód.
Dla $f+g$ teza wynika natychmiast z faktu, że iloraz różnicowy dla $f+g$ i a to odpowiednia suma ilorazów różnicowych. Z kolei wzór Leibnitza uzyskamy stosując standardowy „chwyt" podobny do tego, który został użyty w dowodzie twierdzenia o granicy iloczynu ciągów (patrz twierdzenie II.1): dla $x\in D\setminus\{a\}$

$$\begin{aligned}\frac{f(x)g(x)-f(a)g(a)}{x-a} &= \frac{f(x)g(x)-f(x)g(a)+f(x)g(a)-f(a)g(a)}{x-a}\\ &= f(x)\frac{g(x)-g(a)}{x-a}+\frac{f(x)-f(a)}{x-a}g(a)\end{aligned}$$

(trzeba tu też skorzystać z tego, że f jest ciągła w a — patrz fakt ze strony 65).

Przed dowodem części dotyczącej ilorazu, wykażemy punkt b). W tym celu oznaczmy $D_1:=\{x\in A: f(x)\neq f(a)\}$, $D_2=A\setminus D_1$. Niech $x\in A\setminus\{a\}$ oraz rozważmy iloraz różnicowy dla złożenia

$$i(x):=\frac{(g\circ f)(x)-(g\circ f)(a)}{x-a}.$$

Jeżeli $x\in D_2$, to mamy oczywiście $f(x)=f(a)$, skąd $i(x)=0$. Jeżeli natomiast $x\in D_1$, to

$$i(x)=\frac{g(f(x))-g(f(a))}{f(x)-f(a)}\cdot\frac{f(x)-f(a)}{x-a}.$$

Załóżmy, że a jest p.s. obydwu zbiorów D_1 i D_2. Ponieważ f jest różniczkowalna w a, zatem jest ciągła w a, więc $\lim_{x\to a}f(x)=f(a)$. A zatem mamy

$$\lim_{x\to a}(i|_{D_1})(x)=g'(f(a))f'(a) \tag{V.3}$$

oraz

$$\lim_{x\to a}(i|_{D_2})(x)=0 \tag{V.4}$$

Jednocześnie, skoro a jest p.s. D_2, to

$$f'(a)=\lim_{\substack{x\to a\\ x\in D_2}}\frac{f(x)-f(a)}{x-a}=0.$$

Stąd na mocy twierdzenia o scalaniu (tw. IV.3) z (V.3) i (V.4) wynika, że $\lim_{x\to a} i(x) = g'(f(a))f'(a)$, co kończy dowód punktu b) w tym przypadku. Jeżeli a jest p.s. tylko jednego spośród zbiorów D_1, D_2, to dowód jest prosty i de facto zawarty w powyższych rozważaniach.

Powróćmy do ilorazu. Zauważmy, że iloraz mozna zapisać przy pomocy mnożenia i składania:

$$\left(\frac{f}{g}\right) = f \cdot (h \circ g),$$

gdzie $h : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ zadana jest wzorem $h(x) = \frac{1}{x}$. A zatem w tym przypadku teza wynika natychmiast z udowodnionych już części twierdzenia dotyczących iloczynu i złożenia oraz faktu ze strony 66 pkt. 1 dla $\alpha = -1$.

Udowodnimy c). Możemy założyć, że przedział I jest niezdegenerowany. Ponieważ f jest w szczególności ciągła, zatem Y jest przedziałem (i to niezdegenerowanym — dlaczego?) i f^{-1} też jest ciągła (patrz tw. IV.12). Zatem gdy $y_0 \in Y$, to y_0 — p.s. Y oraz dla $y \in Y \setminus \{y_0\}$, dzięki odwracalności, mamy $f^{-1}(y) \neq f^{-1}(y_0)$, skąd

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{1}{\frac{f(f^{-1}(y)) - f(f^{-1}(y_0))}{f^{-1}(y) - f^{-1}(y_0)}} \xrightarrow[y \to y_0]{} \frac{1}{f'(f^{-1}(y_0))},$$

przy czym powyższa zbieżność jest konsekwencją ciągłości f^{-1} w y_0 oraz różniczkowalności f w $f^{-1}(y_0)$. $\square$

Powyższe twierdzenie wraz z udowodnionym wcześniej faktem pozwala wyliczyć pochodne kolejnych funkcji elementarnych.

Przykład.

1. Niech $1 \neq a > 0$. Na mocy pkt. c) twierdzenia mamy dla $x > 0$

$$(\log_a)'(x) = \frac{1}{\ln a \cdot a^{\log_a x}} = \frac{1}{x \cdot \ln a},$$

   w szczególności $\ln'(x) = \frac{1}{x}$.

2. Dla x należącego do dziedziny tangensa, na mocy (V.2), mamy

$$(\mathrm{tg})'(x) = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = 1 + \mathrm{tg}^2\, x = \frac{1}{\cos^2 x}\ ^{65)}\ .$$

3. Podobnie dla x z dziedziny kotangensa

$$(\mathrm{ctg})'(x) = \frac{-(\sin^2 x + \cos^2 x)}{\sin^2 x} = -1 - \mathrm{ctg}^2\, x = \frac{-1}{\sin^2 x}.$$

Oczywistą konsekwencją twierdzenia V.1 jest następujący wniosek, dotyczący zachowania różniczkowalności przy podstawowych operacjach na funkcjach.

Wniosek. *Suma, iloczyn, iloraz i złożenie (o ile mają sens) funkcji różniczkowalnych jest funkcją różniczkowalną. Funkcja odwrotna do funkcji różniczkowalnej z niezerującą się pochodną, określonej na przedziale jest funkcją różniczkowalną.*

Na zakończenie tego podrozdziału warto podkreślić, że w efekcie udało nam się osiągnąć zapowiadany cel związany z praktyczną „wyliczalnością" wzorów na pochodne wszystkich funkcji elementarnych. Jest to bardzo komfortowa sytuacja — jak przekonamy się w rozdziale VII całkiem odmienna od tej, jaką będziemy mieli przy *całkowaniu*, czyli operacji odwrotnej (w nieco nieścisłym sensie) do różniczkowania.

---

65) Stosujemy tu popularną (choć niestety czasem mylącą) konwencję pisania $f^2(x)$ zamiast $(f(x))^2$ dla pewnych funkcji f — szczególnie trygonometrycznych.

3. Pochodna i ekstrema lokalne

Jak wskazywałaby na to geometryczna interpretacja pochodnej związana ze styczną do wykresu, powinny istnieć łatwe do opisania relacje pomiędzy rozmaitymi własnościami funkcji a własnościami jej pochodnej. W następnym podrozdziale przekonamy się, że tak jest np. z monotonicznością funkcji. Tu natomiast przyjrzyjmy się tego typu związkom, które mają miejsce dla innej własności: posiadania przez funkcję *ekstremum lokalnego*.

Definicja. *Niech $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$ i $x_0 \in D$. Funkcja f posiada* ***maksimum (minimum) lokalne*** *w x_0 wtw dla pewnego $\delta > 0$ zachodzi*

$$f(x_0) = \max(\min)\{f(x) : x \in D, |x - x_0| < \delta\}.$$

Funkcja f posiada ***ekstremum lokalne*** *w x_0 wtw f posiada maksimum lub minimum lokalne w x_0.*

Ekstremum lokalne posiada zatem np. funkcja cos w 0, funkcja stała — w każdym punkcie, funkcja $f : [0; 1] \to \mathbb{R}$ zadana wzorem $f(x) = x$ — w punkcie 0 i 1. Nie widać jednak z tych przykładów by fakt posiadania ekstremum lokalnego wpływał w jakiś jednolity sposób na własności pochodnej. Dlatego ograniczymy się do rozważania tylko niektórych punktów dziedziny funkcji. Mówimy mianowicie, że x_0 jest *punktem wewnętrznym* zbioru D wtw $(x_0 - \delta; x_0 + \delta) \subset D$ dla pewnego $\delta > 0$.

Twierdzenie V.2 (o ekstremach lokalnych). *Jeżeli f posiada ekstremum lokalne w punkcie wewnętrznym x_0 swojej dziedziny oraz f jest różniczkowalna w x_0* [66]*, to $f'(x_0) = 0$.*

Dowód.
Niech D — dziedzina f i niech $\delta > 0$ będzie takie, że $(x_0 - \delta; x_0 + \delta) \subset D$ oraz równocześnie $\forall_{x \in (x_0-\delta; x_0+\delta)}\ f(x) \leqslant f(x_0)$ (a zatem zakładamy, że w x_0 jest maksimum lokalne, gdyby było to jednak minimum lokalne, wystarczy rozważać $-f$ zamiast f). W efekcie dla $x \in (x_0 - \delta; x_0)$ mamy $\frac{f(x)-f(x_0)}{x-x_0} \geqslant 0$, a dla $x \in (x_0; x_0 + \delta)$ mamy $\frac{f(x)-f(x_0)}{x-x_0} \leqslant 0$. Stąd $f'_-(x_0) \geqslant 0$ oraz $f'_+(x_0) \leqslant 0$. Ponieważ jednak $f'_+(x_0) = f'_-(x_0) = f'(x_0)$, zatem $f'(x_0) = 0$. □

Klasyczny przykład funkcji $f : \mathbb{R} \to \mathbb{R}$ zadanej wzorem $f(x) = x^3$, która choć ma w zerze pochodną równą 0, jednak nie posiada w tym punkcie ekstremum, pokazuje, że nie zachodzi twierdzenie odwrotne do twierdzenia V.2. A zatem twierdzenie to daje tylko pewien warunek konieczny na „posiadanie ekstremum lokalnego w x_0”. Niemniej w wielu zadaniach bywa to bardzo przydatne.

Przykład (znajdowanie kresów funkcji — sposób I). Znajdziemy kres górny i dolny zbioru wartości funkcji $f : [0; 3] \to \mathbb{R}$ zadanej wzorem

$$f(x) = 2x^3 - 9x^2 + 12x.$$

Po pierwsze zauważmy, że f jest ciągła, a zatem osiąga w pewnych punktach z $[0; 3]$ swój kres[67] górny i dolny. W każdym z tych punktów f posiada zatem ekstremum lokalne. Jeśli taki punkt x jest punktem wewnętrznym przedziału, to $f'(x) = 6x^2 - 18x + 12 = 6(x^2 - 3x + 2) = 6(x - 1)(x - 2) = 0$, czyli x to 1 lub 2. W efekcie wiemy, że kresy osiągane są w jednym z punktów 0, 1, 2, 3 (na początku wiedzieliśmy tylko, że są osiągane gdzieś w $[0; 3]$ — zatem zbiór „punktów podejrzanych” udało nam się solidnie zmniejszyć ...). Mamy $f(0) = 0$, $f(1) = 5$, $f(2) = 4$, $f(3) = 9$, a zatem kres górny f to 9, a dolny to 0.

---

[66] Wystarczy zakładać, że f posiada pochodną w x_0.

[67] Kresem górnym (dolnym) funkcji nazywamy odpowiedni kres jej zbioru wartości.

4. Twierdzenia o wartości średniej

Poznamy tu trzy tzw. *twierdzenia o wartości średniej* dla pochodnej funkcji. Są to twierdzenia fundamentalne dla rachunku różniczkowego funkcji jednej zmiennej — pozwolą nam one uzyskać wiele ważnych własności różniczkowania.

Niech $a, b \in \mathbb{R}$, $a < b$. Dwa pierwsze twierdzenia dotyczą jednej funkcji $f : [a; b] \to \mathbb{R}$.

Twierdzenie V.3 (Rolle'a). *Jeżeli f jest ciągła w a i w b, różniczkowalna w każdym punkcie przedziału $(a; b)$* [68] *oraz $f(b) = f(a)$, to*

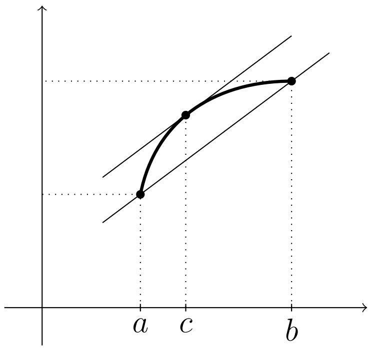
$$\exists_{c\in(a;b)}\ f'(c) = 0.$$

Teza twierdzenia Rolle'a ma prostą interpretację geometryczną: w pewnym punkcie wewnątrz przedziału styczna do wykresu jest pozioma, co dzięki naszym intuicjom związanym z różniczkowalnością funkcji wydaje się całkiem naturalne przy przyjętym założeniu, że wartości funkcji są równe na końcach. Następne twierdzenie jest uogólnieniem poprzedniego — rezygnujemy w nim z założenia $f(b) = f(a)$.

Twierdzenie V.4 (Lagrange'a). *Jeżeli f jest ciągła w a i w b oraz jest różniczkowalna w $(a; b)$, to*

$$\exists_{c\in(a;b)}\ f'(c) = \frac{f(b) - f(a)}{b - a}.$$

To twierdzenie także wydaje się być zgodne z naszą intuicją — styczna do wykresu ma być równoległa do prostej siecznej odpowiadającej argumentom a i b (patrz rys. 9).



Rysunek 9: Styczna dla punktu c jest równoległa do siecznej dla a i b.

Trzecie twierdzenie dotyczy już dwóch funkcji $f, g : [a; b] \to \mathbb{R}$ i jest uogólnieniem obu poprzednich twierdzeń [69] .

Twierdzenie V.5 (Cauchy'ego). *Jeżeli f i g są ciągłe w a i w b oraz różniczkowalne w $(a; b)$, to*

$$\exists_{c\in(a;b)}\ (f(b) - f(a))g'(c) = (g(b) - g(a))f'(c)$$

Dowody.

Zacznijmy od twierdzenia Rolle'a. Zauważmy najpierw, że f — ciągła, zatem z twierdzenia Weiestrassa (tw. IV.10) istnieją $m, M \in [a; b]$ takie, że $\forall_{x\in[a;b]}\ f(m) \leqslant f(x) \leqslant f(M)$. Jeżeli $f(m) = f(M)$, to f jest stała, więc $f'(c) = 0$ dla **każdego** $c \in (a; b)$. Jeśli natomiast $f(m) \neq f(M)$, to jedna z liczb m, M musi byś różna od a i od b, gdyż $f(a) = f(b)$. Biorąc tę właśnie liczbę jako c, z twierdzenia V.2 uzyskujemy tezę, gdyż f posiada w szczególności ekstremum lokalne w c i c jest punktem wewnętrznym $[a; b]$.

[68] Zamiast „różniczkowalna w każdym punkcie zbioru X" będziemy też mówić *różniczkowalna w X*.

[69] Dlaczego?

Teraz pozostałe twierdzenia uzyskamy natychmiast, stosując twierdzenie Rolle'a do odpowiednio dobranych funkcji „pomocniczych" $\tilde{f} : [a; b] \to \mathbb{R}$. Dla twierdzenia Lagrange'a $\tilde{f}$ definiujemy wzorem

$$\tilde{f}(x) := f(x) - \left[\frac{f(b) - f(a)}{b - a}(x - a) + f(a)\right]$$

Dla twierdzenia Cauchy'ego bierzemy natomiast

$$\tilde{f}(x) := f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)),$$

o ile $g(b) \neq g(a)$, a gdy $g(b) = g(a)$ teza wynika natychmiast z twierdzenia Rolle'a

Przykładem bardzo ważnej konsekwencji twierdzenia Lagrange'a jest następujący wynik dotyczący najprostszego *równania różniczkowego*: $f'(x) = 0$.

Wniosek. *Jeżeli I — przedział oraz $f : I \to \mathbb{R}$ spełnia $f'(x) = 0$ dla dowolnego $x \in I$, to f jest funkcją stałą.*

Dowód.
Dla dowolnych $x, y \in I$, $x < y$ istnieje $c \in (x; y)$ takie, że $\frac{f(x)-f(y)}{x-y} = f'(c) = 0$, skąd $f(x) = f(y)$. □

Należy jednak **koniecznie** pamiętać, że powyższy wniosek dotyczy wyłącznie funkcji określonych na przedziale.

Kolejnym ważnym wnioskiem jest kryterium monotoniczności funkcji. Podobnie jak przed chwilą, istotne jest tu, że dziedzina funkcji to przedział.

Twierdzenie V.6 (o monotoniczności). *Jeżeli I — przedział oraz $f : I \to \mathbb{R}$ jest różniczkowalna, to*

1. *f jest rosnąca (malejąca) wtw $\forall_{x \in I}\ f'(x) \geqslant 0$ ($\leqslant 0$);*

2. *Jeżeli $\forall_{x \in I}\ f'(x) > 0$ (< 0), to f jest ściśle rosnąca (ściśle malejąca).*

Dowód.
Implikacja „$\Rightarrow$" w pkt. 1. to natychmiastowy wniosek z definicji pochodnej, a pozostała część tezy twierdzenia wynika (też natychmiastowo) z tw. Lagrange'a. □

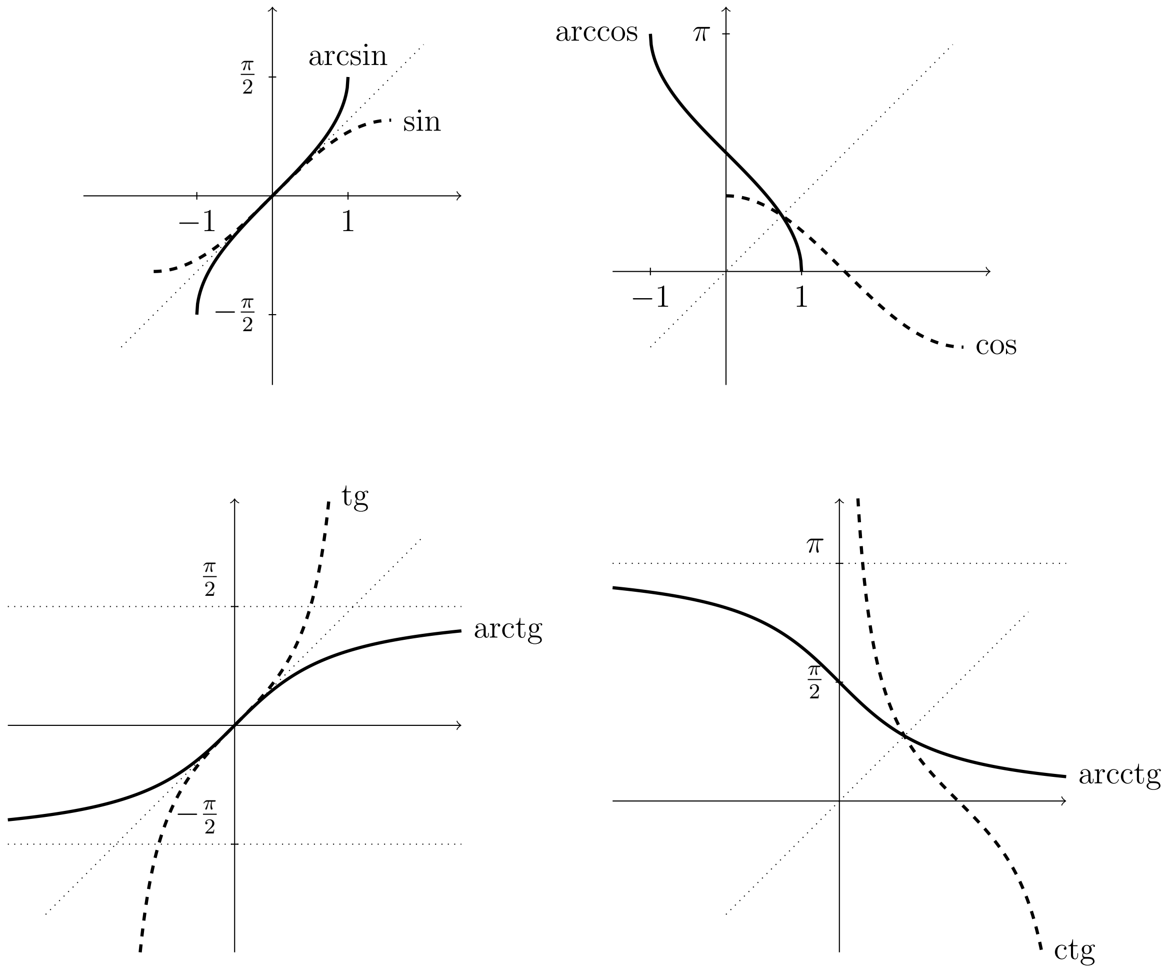
Wspomniany niedawno przykład funkcji $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$ pokazuje, że w pkt. 2. powyżej implikacja „$\Leftarrow$" **nie zachodzi**, bowiem $f'(0) = 0$.

Twierdzenie o monotoniczności wykorzystamy w poniższym przykładzie, gdzie zdefiniujemy kolejne funkcje „elementarne": arcsin, arccos, arctg i arcctg.

Przykład 1 (funkcje „arkus..."). Rozważmy następujące cztery funkcje, będące obcięciami znanych nam funkcji trygonometrycznych do pewnych podzbiorów ich dziedzin.

$s : \left[-\frac{\pi}{2}; \frac{\pi}{2}\right] \to [-1; 1]$, $\quad s(x) = \sin x$;
$c : [0; \pi] \to [-1; 1]$, $\quad c(x) = \cos x$;
$t : \left(-\frac{\pi}{2}; \frac{\pi}{2}\right) \to \mathbb{R}$, $\quad t(x) = \operatorname{tg} x$;
$ct : (0; \pi) \to \mathbb{R}$, $\quad ct(x) = \operatorname{ctg} x$.

Na mocy tw. V.6 pkt.2 oraz na mocy wzorów z podrozdziału V.2 wszystkie te funkcje są różnowartościowe [70)] (s i t są ściśle rosnące, c i ct — ściśle malejące). Korzystając z twierdzenia o własności Darboux (tw. IV.9) oraz badając granice, względnie wartości powyższych funkcji w końcach ich dziedzin uzyskujemy też, że funkcje te są „na". Funkcje arcsin, arccos, arctg i arcctg to funkcje odwrotne do s, c, t i ct. Ich wykresy są zatem symetryczne odpowiednio do wykresów funkcji s, c, t, ct względem prostej o równaniu $y = x$ (patrz rys. 10).



Rysunek 10: Wykresy „fragmentów” funkcji trygonometrycznych i odpowiadających im „arkusów”.

Z twierdzenia V.1 uzyskujemy też różniczkowalność arctg i arcctg oraz różniczkowalność w $(-1;1)$ arcsin i arccos oraz wzory:

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}, \qquad \arccos'(x) = \frac{-1}{\sqrt{1-x^2}} \qquad \text{dla} \quad x \in (-1;1);$$

$$\operatorname{arctg}'(x) = \frac{1}{1+x^2}, \qquad \operatorname{arcctg}'(x) = \frac{-1}{1+x^2} \qquad \text{dla} \quad x \in \mathbb{R}.$$

Twierdzenie o monotoniczności pozwala też rozwiązywać zadania na znajdowanie kresów funkcji przy użyciu metody alternatywnej do tej użytej w przykładzie ze str. 69 (wykorzystującej twierdzenie o ekstremach lokalnych).

Przykład 2 (znajdowanie kresów funkcji — sposób II). Rozważmy tę samą funkcję co poprzednio. Najpierw znajdziemy możliwie duże przedziały zawarte w dziedzinie f, po obcięciu do których f jest monotoniczna (tzw. *maksymalne przedziały monotoniczności*). Dzięki tw. V.6 sprowadza się to do rozwiązania nierówności $f'(x) \leqslant 0$ lub $f'(x) \geqslant 0$. Ponieważ $f'(x) = 6(x-1)(x-2)$ zatem bez trudu uzyskujemy, że na $[1;2]$ f jest malejąca, a na $[0;1]$ i na $[2;3]$ jest rosnąca (choć na sumie: $[0;1] \cup [2;3]$ już **nie** — dlaczego?). Oczywiście kres górny f może być osiągany jedynie w prawym końcu któregoś przedziału, gdzie f rośnie lub lewym takiego, gdzie

[70] Należy tu czasem użyć też faktu, że pochodna funkcji stałej na pewnym przedziale ma w **całym** tym przedziale pochodną równą 0.

f maleje, czyli w 1 lub 3. Ponieważ $f(1) = 5 < 9 = f(3)$, więc kres górny to 9. Podobnie kres dolny może być osiągany jedynie w prawym końcu przedziału, gdzie f maleje lub lewym, gdzie rośnie, czyli w 2, 0, 3. A zatem kres dolny to $f(0) = 0$, bo $f(3) = 9 > f(2) = 4 > 0 = f(0)$.

Jak widać z czysto rachunkowego punktu widzenia, ta metoda jest bardzo podobna do metody I. Zamiast równania $f'(x) = 0$ rozwiązujemy nierówność $f'(x) \geqslant 0$ lub $\leqslant 0$, a to na ogół robi się bardzo podobnie. Główna różnica polega na sposobie argumentacji. Metoda II ma oczywiście swoje ograniczenia: cała dziedzina musi dać się rozbić na sumę przedziałów monotoniczności. Ma też jednak pewną wyższość nad metodą I — można ją bez większego trudu uogólnić na przypadek funkcji określonych na innych przedziałach niż tylko domknięte, z czym dla metody I mogą być pewne kłopoty (zachęcam do znalezienia stosownego przykładu).

A oto jeszcze jedno zastosowanie twierdzenia o monotoniczności.

Przykład 3 (dowodzenie nierówności). Wykażemy nierówność

$$\sin x \leqslant x \qquad \text{dla} \;\; x \geqslant 0.$$

Rozważmy $f : [0; +\infty) \longrightarrow \mathbb{R}$ zadaną wzorem $f(x) = \sin x - x$. Mamy

$$f'(x) = \cos x - 1 \leqslant 0$$

dla każdego $x \geqslant 0$. Funkcja f jest więc malejąca — w szczególności $\sin x - x = f(x) \leqslant f(0) = 0$, czyli $\sin x \leqslant x$ dla dowolnego $x \geqslant 0$. A zatem, w dużym skrócie, sprowadziliśmy dowód nierówności dla funkcji do pewnej nierówności na jej pochodną. Całkiem podobnie dowodzimy nierówność

$$\ln(1 + x) \leqslant x \qquad \text{dla} \;\; x > -1.$$

należy tylko badać monotoniczność funkcji zadanej wzorem $\ln(1 + x) - x$ osobno na lewo i osobno na prawo od 0.

Ważną konsekwencją twierdzenia Cauchy'ego (tw. V.5) jest tzw. reguła de l'Hospitala, pomocna niekiedy przy obliczaniu granic funkcji (choć niestety także często jest nieprzydatna, albo bywa używana wtedy, gdy można się łatwo obyć bez niej...).

Twierdzenie V.7 (reguła de l'Hospitala). *Niech $-\infty \leqslant a < b \leqslant +\infty$. Załóżmy, że funkcje f i g określone w $(a; b)$ są różniczkowalne oraz że $g'(x) \neq 0$ dla $x \in (a; b)$. Niech $x_0 = a$ lub b. Jeżeli istnieje granica*

$$\lim_{x \to x_0} \frac{f'(x)}{g'(x)} \tag{V.5}$$

oraz zachodzi któreś z założeń

wersja 1. *(„$\frac{0}{0}$"):* $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0$

wersja 2. *(„$\frac{?}{\pm\infty}$"):* $\lim_{x\to x_0} g(x) = +\infty$ *lub* $-\infty$,

to

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)}.$$

Dowód.

Przedstawimy tu tylko dowód dla wersji 1. założeń z dodatkowymi założeniami, że $x_0 = a \in \mathbb{R}$. Funkcje f i g „dookreślimy" w punkcie a biorąc $f(a) = g(a) = 0$ tzn., nieco ściślej, zdefiniujemy $\tilde{f}, \tilde{g} : [a; b) \to \mathbb{R}$ wzorami

$$\tilde{f}(x) = \begin{cases} 0 & \text{dla } x = a \\ f(x) & \text{dla } x \in (a; b), \end{cases} \qquad \tilde{g}(x) = \begin{cases} 0 & \text{dla } x = a \\ g(x) & \text{dla } x \in (a; b). \end{cases}$$

Niech $x \in (a; b)$. Oczywiście $\tilde{f}$ i $\tilde{g}$ są ciągłe w a i w x oraz różniczkowalne w przedziale $(a; x)$. W szczególności zatem, z twierdzenia Rolle'a, mamy $g(x) \neq 0$ na mocy założenia, że pochodna g' jest niezerowa. Ponadto z tw. Cauchy'ego dla pewnego $c_x \in (a;x)$ zachodzi

$$\frac{f(x)}{g(x)} = \frac{\tilde{f}(x) - \tilde{f}(a)}{\tilde{g}(x) - \tilde{g}(a)} = \frac{f'(c_x)}{g'(c_x)}. \tag{V.6}$$

Skoro $a < c_x < x$, zatem na mocy twierdzenia o trzech funkcjach (tw. IV.4) mamy $\lim_{x\to a} c_x = a$ i przy tym $c_x \neq a$. Stąd, na mocy (V.6), dzięki istnieniu granicy (V.5), otrzymujemy tezę twierdzenia.[71] □

Zauważmy jeszcze, że granice pojawiające się w twierdzeniu V.7 to granice de facto jednostronne (choć nie zostały użyte symbole $\lim_{x\to x_0\pm}$, ale x_0 jest „na końcu" dziedziny). A zatem tylko do granic jednostronnych można reguły de l'Hospitala używać w sposób bezpośredni. W przypadku granic „obustronnych" trzeba właściwie użyć jej dwukrotnie — dla „każdej ze stron" osobno. I jeszcze jedna sprawa. Stosując regułę de l'Hospitala **nie można** zapomnieć, że istnienie granicy (V.5) jest jednym z założeń twierdzenia!

Przykład („nieoznaczoność" typu „$1^{\pm\infty}$"). Znajdziemy $\lim_{x\to 0}(\cos x)^{\frac{1}{x}}$. Ponieważ granica podstawy, tj. cos równa jest 1, a granica wykładnika co prawda nie istnieje, ale istnieją granice jednostronne równe odpowiednio $\pm\infty$, zatem taką sytuację określamy mianem *nieoznaczoności*[72] *typu „$1^{+\infty}$" lub odpowiednio „$1^{-\infty}$".* Dla $0 \neq x \in (-\frac{\pi}{2}; \frac{\pi}{2})$ mamy:

$$(\cos x)^{\frac{1}{x}} = e^{\ln((\cos x)^{\frac{1}{x}})} = e^{\frac{1}{x}\ln\cos x}. \tag{V.7}$$

Policzmy więc najpierw $\lim_{x\to 0\pm} \frac{\ln\cos x}{x}$ — użyjemy regułę de l'Hospitala (wersję 1) mamy bowiem $\lim_{x\to 0\pm} x = \lim_{x\to 0\pm} \ln\cos x = \ln 1 = 0$. Ponieważ iloraz pochodnych to

$$\frac{\frac{1}{\cos x}\cdot(-\sin x)}{1} = -\operatorname{tg} x$$

i posiada on (obustronną) granicę w 0 równą 0, zatem także $\lim_{x\to 0} \frac{\ln\cos x}{x} = 0$, stąd, dzięki ciągłości funkcji wykładniczej, na mocy (V.7), mamy

$$\lim_{x\to 0}(\cos x)^{\frac{1}{x}} = e^0 = 1.$$

5. Wyższe pochodne

Będziemy tu mówili o pochodnych n-tego rzędu (inaczej: n-tych pochodnych), gdzie $n \in \mathbb{N}_0$. Dla $n = 0$ n-ta pochodna funkcji f to po prostu sama funkcja f — określona jest więc ona w każdym punkcie dziedziny. Z kolei dla $n = 1$ *pierwsza pochodna* f w punkcie x to po prostu pochodna, czyli $f'(x)$, o ile istnieje. Oznaczmy więc $f^{(0)}(x) = f(x)$, $f^{(1)}(x) = f'(x)$. Ogólnie, n-tą pochodną funkcji f w punkcie x_0 będziemy oznaczać przez $f^{(n)}(x_0)$. Zdefiniujemy ją przy użyciu rekursji „po n", startując np. od przyjętej już definicji dla $n = 0$. Mówiąc niezbyt dokładnie, $(n+1)$-sza pochodna będzie po prostu pochodną n-tej pochodnej, musimy to jednak doprecyzować. Dla zbioru $D \subset \mathbb{R}$, punktu $x_0 \in \mathbb{R}$ oraz $\delta > 0$ przyjmijmy (a właściwie przypomnijmy — patrz str. 64) oznaczenie

$$D_{x_0,\delta} := D \cap (x_0 - \delta; x_0 + \delta),$$

które przyda nam się poniżej.

[71] Patrz też ew.: twierdzenie „o granicy złożenia" z zadania IV.2 do rozdziału IV.

[72] Ściślej, słowo „nieoznaczoność" wyraża niemożność sensownego zdefiniowania odpowiedniego działania — tu potęgowania „$1^{+\infty}$" ani „$1^{-\infty}$".

Definicja (przejście od n do $n+1$ w definicji n-tej pochodnej). *Niech $f : D \to \mathbb{R}$, $x_0 \in D$. Jeżeli $n \in \mathbb{N}_0$, to $f^{(n+1)}(x_0)$ **istnieje** wtw dla pewnego $\delta > 0$ dla dowolnego $x \in D_{x_0,\delta}$ $f^{(n)}(x)$ istnieje i jest skończona oraz funkcja $D_{x_0,\delta} \ni x \rightsquigarrow f^{(n)}(x)$ posiada pochodną w x_0 [73]. W takiej sytuacji tę pochodną w x_0 oznaczamy $f^{(n+1)}(x_0)$ i nazywamy $n+1$–**szą pochodną** f **w** x_0 (ew. pochodną $n+1$-szego rzędu w x_0).*

Będziemy także mówić, że f jest *n–krotnie różniczkowalna w (punkcie) x_0* wtw $f^{(n)}(x_0)$ istnieje i jest **skończona** oraz, że f jest *n–krotnie różniczkowalna* wtw f jest n–krotnie różniczkowalna w x dla dowolnego x z dziedziny funkcji f. W tej ostatniej sytuacji funkcję $D \ni x \rightsquigarrow f^{(n)}(x)$ nazywamy *n-tą pochodną f* i (oczywiście) oznaczamy przez $f^{(n)}$. Czasami na $f^{(n)}$ używa się też tradycyjnego oznaczenia $\frac{d^n f}{dx^n}$.

Zwróćmy jeszcze uwagę na pewną subtelność związaną z definicją wyższych pochodnych. W przypadku 1–szej pochodnej, aby mogła być ona określona w punkcie x_0, z punktu widzenia własności samej dziedziny wystarczało w zasadzie by x_0 był jej elementem oraz punktem skupienia. Mógł więc to być np. punkt $x_0 = 0$ dziedziny $D = \{-\frac{1}{n} : n \in \mathbb{N}\} \cup [0;1]$. Jednak już dla $n = 2$ podobna sytuacja nie jest możliwa, bowiem w żadnym z punktów postaci $-\frac{1}{n}$ pierwsza pochodna funkcji określonej na D nie istnieje (bo punkty te nie są punktami skupienia D). Zatem zgodnie z definicją, $f^{(2)}(0)$ nie istnieje, niezależnie od tego jak „regularną” funkcję f rozważamy na tej dziedzinie.

Dzięki znalezionym już przez nas wzorom na pierwsze pochodne możemy teraz bez trudu (np. indukcyjnie) dowieść, że wiele spośród funkcji elementarnych to funkcje n-krotnie różniczkowalne dla **dowolnego** n. Tak jest np. z funkcjami: wykładniczymi, potęgowymi (określonymi na $\mathbb{R}_+$), wielomianami (określonymi na $\mathbb{R}$), logarytmami, sin oraz cos. Co więcej, katalog takich funkcji można bardzo rozszerzyć dzięki poniższemu rezultatowi będącemu wnioskiem (choć może nie we wszystkich punktach trywialnym[74]) z twierdzenia o własnościach rachunkowych pochodnej (tw. V.1).

Twierdzenie V.8 (własności rachunkowe n-krotnego różniczkowania). *Suma, iloczyn i iloraz funkcji f i g różniczkowalnych n-krotnie w punkcie x_0 są funkcjami n-krotnie różniczkowalnymi w x_0 oraz zachodzi*

1. $(f+g)^{(n)}(x_0) = f^{(n)}(x_0) + g^{(n)}(x_0)$, $(\alpha \cdot f)^{(n)}(x_0) = \alpha \cdot f^{(n)}(x_0)$, *dla* $\alpha \in \mathbb{R}$;
2. *(wzór Leibnitza rzędu n)* $(f \cdot g)^{(n)}(x_0) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x_0) g^{(n-k)}(x_0)$.

Złożenie $g \circ f$ funkcji f różniczkowalnej n-krotnie w x_0 z funkcją g różniczkowalną n-krotnie w $f(x_0)$ jest n-krotnie różniczkowalne w x_0.

Jeżeli I — przedział oraz $f : I \to f(I) \subset \mathbb{R}$ jest odwracalna i n-krotnie różniczkowalna oraz $f'(x) \neq 0$ dla dowolnego $x \in I$, to f^{-1} jest n-krotnie różniczkowalna. **B.D.**

Przy użyciu tego twierdzenia otrzymujemy m. in. n-krotną różniczkowalność dla dowolnego n funkcji: tg, ctg, arctg, arcctg, a także funkcji arcsin i arccos obciętych do przedziału $(-1;1)$.

Dziwić może nieco brak w powyższym twierdzeniu wzorów na n-tą pochodną ilorazu oraz złożenia. Dla ilorazu wzór taki można by jeszcze ewentualnie wypisać, choć byłby on dość skomplikowany. Natomiast wzór na n-tą pochodną złożenia jest już tak makabrycznie skomplikowany, że zapisanie go w zwięzłej formie jest nie lada sztuką! Zachęcam do wypisania go tylko dla $n = 3$ (i sądzę, że to wystarczy, by powyższą opinię podzielić...).

Na koniec tego podrozdziału — dwa często spotykane oznaczenia: klasa $C^n(D)$ to zbiór wszystkich tych funkcji określonych na D, które są n-krotnie różniczkowalne oraz ich n-ta pochodna $f^{(n)}$ jest funkcją ciągłą, a klasa $C^\infty(D)$ — tych, które są n-krotnie różniczkowalne dla każdego $n \in \mathbb{N}$. Używa się też sformułowania: f **jest klasy** C^n (odp. **klasy** C^∞). Zamiast C^0 piszemy na ogół C, czyli $C(D)$, to po prostu zbiór wszystkich ciągłych funkcji $f : D \to \mathbb{R}$.

---

[73] W szczególności zatem x_0 musi być punktem skupienia i elementem D.

[74] Choć osiągalnym przy tak już dalece rozwiniętej przez nas teorii — patrz — Zadania.

6. Druga pochodna i wypukłość

Spośród n-tych pochodnych dwie — mianowicie pierwsza i druga wyróżniają się ze względu na ich liczne zastosowania i czytelną interpretację geometryczną. O pierwszej już powiedzieliśmy nieco. Teraz w formie bardzo skrótowej zajmiemy się drugą pochodną. Najpopularniejsze zastosowanie ma ona chyba w fizyce — np. określa wartość przyśpieszenia (podczas gdy pierwsza — prędkości) punktu poruszającego się „jednowymiarowo", ale też każdej ze współrzędnych punktu poruszającego w wielu wymiarach (wtedy różniczkowana funkcja określa odpowiednią współrzędną położenia punktu, a zmienna to czas).

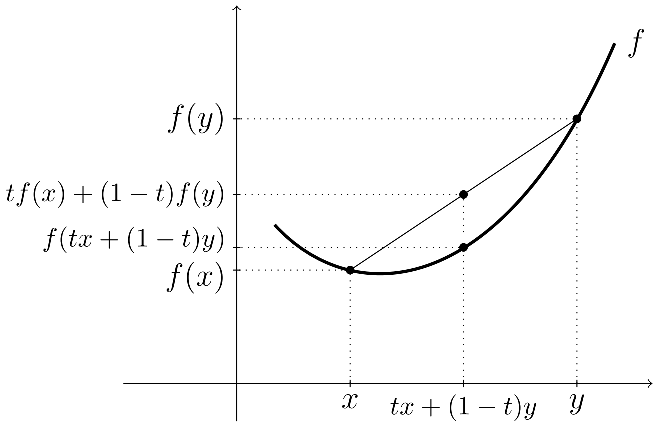
Znak pierwszej pochodnej ma ścisły związek z dość „geometryczną" własnością funkcji jaką jest monotoniczność. Tymczasem, jak zaraz zobaczymy, znak drugiej pochodnej wiąże się z inną, też bardzo geometryczną własnością — mianowicie z wypukłością. Przypomnijmy tu, że podzbiór $A \subset \mathbb{R}^k$ [75] jest *wypukły* wtw dla dowolnych $a, b \in A$ odcinek łączący a i b zawarty jest w A. Zdefiniujmy pojęcie *wypukłości funkcji* $f : I \to \mathbb{R}$ [76], gdzie I — przedział (I oznacza przedział w całym tym podrozdziale). Niech N_f oznacza zbiór punktów położonych „nieostro" **nad** wykresem f, tzn. $N_f := \{(x, y) \in \mathbb{R}^2 : x \in I, y \geqslant f(x)\}$.

Definicja. *$f : I \to \mathbb{R}$ jest* ***wypukła*** *wtw N_f jest zbiorem wypukłym; f jest* ***wklęsła*** *wtw $(-f)$ jest wypukła.*

Uwaga. Oczywiście, w definicji wypukłości funkcji wystarczy zakładać, że każda *cięciwa wykresu*, tzn. odcinek łączący dwa punkty wykresu, zawiera się w N_f, a zatem zapisując ten fakt w formie analitycznej uzyskujemy, że f jest wypukła wtw

$$\forall_{x,y \in I}\ \forall_{t \in [0;1]}\ f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y) \tag{V.8}$$

(patrz rys. 11).



Rysunek 11: Przy wypukłości cięciwa leży nad wykresem.

Analogiczny warunek, tyle że z nierównością w stronę przeciwną, równoważny jest wklęsłości funkcji.

Powyższy analityczny warunek wyrażający wypukłość funkcji można łatwo uogólnić do warunku dotyczącego n punktów z odcinka I zamiast tylko dwóch punktów x i y.

Fakt (nierówność Jensena). *Niech $n \in \mathbb{N}_2$. Funkcja $f : I \to \mathbb{R}$ jest wypukła wtw*

$$\forall_{x_1,\ldots,x_n \in I}\ \forall_{\substack{t_1,\ldots,t_n \in [0;1],\\ t_1+\ldots+t_n=1}}\ f\Big(\sum_{i=1}^{n} t_i x_i\Big) \leqslant \sum_{i=1}^{n} t_i f(x_i). \tag{V.9}$$

[75] Tu standardowo oznaczamy przez X^k iloczyn kartezjański k-egzemplarzy zbioru X.

[76] Uwaga! Z formalnego punktu widzenia taka funkcja to to samo co jej wykres, a więc pewien podzbiór $\mathbb{R}^2$. Jednak wypukłość f jako takiego właśnie zbioru jest zupełnie **czym innym** niż wypukłość f jako funkcji, o czym przekonamy się za chwilę.

Dowód.
„$\Longleftarrow$" — oczywisty z uwagi powyżej (wystarczy rozważyć $t_1 = t$, $t_2 = 1 - t$, $t_3 = \ldots = t_n = 0$ oraz $x_1 = x$, $x_2 = y$, a pozostałe x_i — dowolne),
„$\Longrightarrow$" — przy założeniu wypukłości łatwo wykazać indukcyjnie, że dla każdego $n \in \mathbb{N}_2$ zachodzi (V.9). □

Jeżeli zastosować nierówność Jensena do pewnych odpowiednio dobranych funkcji f (oraz odpowiednich x_i, t_i), można uzyskać wiele ciekawych, ważnych i znanych nierówności. Nieco przykładów zostało umieszczonych w zadaniach.

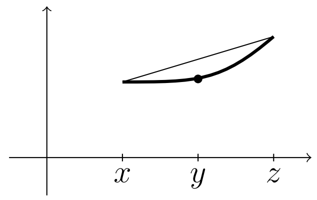
Jednak zasadnicze pytania, na które należałoby odpowiedzieć zanim zacznie się stosować powyższy fakt, są następujące:

Jak rozpoznać, czy dana funkcja jest wypukła? Czy można to zrobić prościej niż poprzez bezpośrednie sprawdzenie warunku (V.8)*?*

Okazuje się, że w przypadku tych funkcji, dla których umiemy „wyliczyć" pochodną odpowiedź jest nietrudna.

Twierdzenie V.9. *Jeżeli $f : I \to \mathbb{R}$ jest różniczkowalna, to f jest wypukła (wklęsła) wtw f' jest rosnąca (malejąca).*

Dowód.
W oparciu o (V.8) nietrudno wykazać charakteryzację wypukłości w terminach „wzrostu" ilorazów różnicowych zawartą w poniższym lemacie (patrz rys. 12).



Rysunek 12: Prawa cięciwa jest bardziej (nie mniej) stroma od lewej.

Lemat. *Funkcja f jest wypukła wtw dla dowolnych x, y, $z \in I$ takich, że $x < y < z$ zachodzi*

$$\frac{f(x) - f(y)}{x - y} \leqslant \frac{f(y) - f(z)}{y - z}.$$

B.D.

Tezę twierdzenia łatwo uzyskać teraz z lematu, wykorzystując definicję pochodnej jako granicy ilorazu różnicowego oraz twierdzenie Lagrange'a o wartości średniej (tw. V.4). □

To twierdzenie pozwala nam uzyskać wypukłość bądź wklęsłość wielu funkcji elementarnych obciętych do odpowiednich przedziałów. Np. funkcje wykładnicze są wypukłe, $\log_a$ jest wklęsły przy $a > 1$ oraz wypukły dla $0 < a < 1$, funkcja potęgowa (określona na $[0; +\infty)$) z wykładnikiem $\alpha \geqslant 1$ jest wypukła, a z wykładnikiem $\alpha \in [0; 1]$ — wklęsła[77], sinus obcięty do $[0; \pi]$ jest wklęsły.

Gdy funkcja jest dwukrotnie różniczkowalna, charakteryzacja wypukłości sprowadza się na mocy twierdzenia V.9 jedynie do badania znaku drugiej pochodnej. Uwaga: zamiast $f^{(2)}(x)$ używa się często oznaczenia

$$f''(x).$$

Wniosek. *Jeżeli $f : I \to \mathbb{R}$ jest dwukrotnie różniczkowalna, to f jest wypukła (wklęsła) wtw*

$$\forall_{x \in I}\ f''(x) \geqslant 0 \ (\leqslant 0).$$

[77] Co prawda, gdy $\alpha < 1$, to brak różniczkowalności w 0, ale wtedy mamy wypukłość po obcięciu funkcji do $(0; +\infty)$, skąd na całej dziedzinie łatwo (jak?) uzyskać wypukłość dzięki ciągłości (patrz też zadanie V.42).

7. Wzór Taylora

Gdy znamy wartość funkcji f w punkcie x_0 i wiemy, że f jest w tym punkcie ciągła, to możemy powiedzieć, że mamy jakąś informację o wartościach tej funkcji f w punktach „bliskich x_0" — wiemy mianowicie, że są one „bliskie $f(x_0)$". Ściślej, mamy

$$f(x) = f(x_0) + R_0(x), \qquad \text{gdzie} \quad R_0(x) \xrightarrow[x\to x_0]{} 0.$$

Gdy założymy nieco więcej — różniczkowalność w x_0, to fakt, że $f'(x_0)$ jest odpowiednią granicą ilorazu różnicowego można równoważnie zapisać w taki sposób:

$$f(x) = f(x_0) + f'(x_0)\cdot(x - x_0) + R_1(x), \qquad \text{gdzie} \quad \frac{R_1(x)}{x - x_0} \xrightarrow[x\to x_0]{} 0.$$

Oczywiście mamy w szczególności także $R_1(x) \xrightarrow[x\to x_0]{} 0$, ale informacja, że $\frac{R_1(x)}{x-x_0} \xrightarrow[x\to x_0]{} 0$ jest znacznie mocniejsza (dlaczego?). Inaczej mówiąc, wydaje się, że przybliżenie f „w pobliżu x_0" przez funkcję zadaną wzorem

$$f(x_0) + f'(x_0)(x - x_0),$$

którą znamy, o ile tylko znamy wartości $f(x_0)$ i $f'(x_0)$, jest „lepsze" niż poprzednie przybliżenie funkcją stale równą $f(x_0)$. Powstaje naturalne pytanie, czy znając $f^{(k)}(x_0)$ dla $0 \leqslant k \leqslant n$ będziemy w stanie uzyskać coraz lepsze przybliżenia, w podobnym rozumieniu. Okazuje się, że odpowiedź jest nietrudna. Funkcja przybliżająca f „w pobliżu x_0" jest tym razem pewnym wielomianem wyznaczonym przez liczby $f(x_0)$,...,$f^{(n)}(x_0)$. Nazywamy go *wielomianem Taylora*, a dokładniej, n-tym *wielomianem Taylora funkcji f w punkcie x_0* i oznaczamy przez T_{n,f,x_0}, albo krócej przez T_n, gdy f i x_0 są ustalone. Wielomian $T_n : D \to \mathbb{R}$, gdzie D — dziedzina f, zadany jest wzorem[78)] :

$$T_n(x) := \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k. \tag{V.10}$$

W szczególności, uprzednio wypisana przez nas funkcja przybliżająca funkcję f była w obu przypadkach $n = 0$ i $n = 1$ równa właśnie odpowiedniemu T_n. Tak jak już zapowiedzieliśmy, wielomian T_n „dość dokładnie" przybliża f „w pobliżu" x_0. I czasami właśnie taka całkiem nieścisła informacja

$$„f \approx T_n"$$

nazywana bywa *wzorem Taylora* („$\approx$" to: „równa się w przybliżeniu"). Inny zapis tego samego, to

$$f = T_n + R_n, \tag{V.11}$$

gdzie R_n — „małe w pobliżu x_0". Oczywiście sama formuła (V.11) nie jest żadnym matematycznym twierdzeniem — to nic więcej niż po prostu definicja funkcji R_n, tzn.

$$R_n := f - T_n,$$

gdzie T_n zadane jest przez (V.10) (gdy potrzeba zaznaczyć zależność od f i x_0 piszemy R_{n,f,x_0} zamiast R_n). Funkcję R_n nazywa się *n-tą resztą Taylora* (funkcji f w punkcie x_0). Istnieje wiele uściśleń w.w. wzoru Taylora, mogących w jakimś sensie wyrażać „małość" reszty Taylora. Poznamy tu dwa z nich. Pierwsze to twierdzenie Peano, będące uogólnieniem przytoczonych na wstępie wyników dla $n = 0$ i $n = 1$.

[78)] T_n można też traktować jako funkcję określoną np. na całym $\mathbb{R}$.

Twierdzenie V.10 (Peano o postaci reszty Taylora). *Jeżeli f jest n-krotnie różniczkowalna w x_0 oraz x_0 ma otoczenie w dziedzinie f będące przedziałem*[79] *, to*

$$\frac{R_n(x)}{(x-x_0)^n} \xrightarrow[x\to x_0]{} 0.$$

Dowód.
Stosując $(n-1)$-krotnie regułę de l'Hospitala sprowadzamy badanie granicy $\lim_{x\to x_0}\frac{R_n(x)}{(x-x_0)^n}$ do badania granicy $\lim_{x\to x_0}\frac{1}{n!}\left[\frac{f^{(n-1)}(x)-f^{(n-1)}(x_0)}{x-x_0}-f^{(n)}(x_0)\right]$ [80] , równej 0 na mocy definicji $f^{(n)}(x_0)$. □

Tezę twierdzenia Peano wygodniej niekiedy zapisać w postaci takiej:

$$f(x) = T_n(x) + (x-x_0)^n \cdot r(x), \quad \text{gdzie } \lim_{x\to x_0} r(x) = 0.$$

Bardzo często to twierdzenie jest znacznie zgrabniejszym narzędziem pomocnym przy liczeniu granic funkcji niż sama reguła de l'Hospitala. Pozwala ono bowiem de facto zastąpić wielomianem nawet dość skomplikowaną funkcję (zastępujemy odpowiednio dobranym wielomianem Taylora tej funkcji + „nieistotną" resztą). Problem sprowadza się więc najczęściej do trywialnego zadania polegającego na obliczeniu granicy ilorazu dwóch wielomianów. Jednocześnie stosując tę metodę, chyba lepiej rozumiemy rozwiązanie niż używając reguły de l'Hospitala.

Przykład. Obliczmy

$$\lim_{x\to 0}\frac{\sqrt{1+x}-1-\frac{1}{2}x}{x^2}.$$

Weźmy $f(x) := \sqrt{1+x}$, $x > -1$. Mamy $f(0) = 1$, $f'(0) = \frac{1}{2}$, $f''(0) = -\frac{1}{4}$. Stąd dla $x_0 = 0$

$$T_2(x) = 1 + \frac{1}{2}x - \frac{1}{8}x^2$$

i z twierdzenia Peano $\sqrt{1+x} = T_2(x) + x^2 \cdot r(x)$, gdzie $r(x) \xrightarrow[x\to 0]{} 0$. Stąd

$$\lim_{x\to 0}\frac{\sqrt{1+x}-1-\frac{1}{2}x}{x^2} = \lim_{x\to 0}\left(\frac{-\frac{1}{8}x^2}{x^2} + r(x)\right) = -\frac{1}{8}.$$

Warto jeszcze wspomnieć w kontekście tezy twierdzenia Peano o tzw. notacji *o-małe* często stosowanej dla skrócenia zapisu rozmaitych formuł, czy rachunków. Mianowicie napis „$f(x) = o(g(x))$ przy $x \to x_0$" oznacza po prostu, że $\lim_{x\to x_0}\frac{f(x)}{g(x)} = 0$. Często, gdy wiadomo o jakie chodzi x_0, pisze się tylko „$f(x) = o(g(x))$". Co więcej, używany również bywa zapis typu „$u(x) = h(x)+o(g(x))$", oznaczający to samo, co „$u(x)-h(x) = o(g(x))$". Np. tezę twierdzenia Peano można by zapisać:

$$f(x) = T_n(x) + o((x-x_0)^n).$$

Taka notacja bywa wygodna, ale należy zachować ostrożność. Np. dwa „o" nie muszą być sobie równe, choć są zapisane tym samym symbolem. A zatem z tego, że $e^x - 1 - \frac{x^2}{2} = x + o(x^2)$ oraz $\sin x = x + o(x^2)$ **nie wynika**, że $\sin x = e^x - 1 - \frac{x^2}{2}$. Dlatego dla początkujących polecam jednak raczej całkiem ścisły zapis w stylu: $\sin x = x + r(x)\cdot x^2$, gdzie $r(x)$ ma granicę 0 w 0. Dla różnych funkcji, a co za tym idzie — różnych „r-ów", można wtedy, dla ich odróżnienia, zastosować numerację r_1, r_2 itd.

Zapowiadana, druga wersja wzoru Taylora umożliwi znacznie konkretniejsze szacowanie „błędu" (czyli reszty Taylora) pomiędzy funkcją a jej wielomianem Taylora. Takie szacowanie

[79] Patrz uwaga 3 str.64.

[80] Zachęcam do samodzielnego szczegółowego prześledzenia.

nie daje się uzyskać w oparciu o twierdzenie Peano, zawierające tylko informację o pewnej granicy związanej z tym błędem. Niestety jednak nie dostaniemy nic „za darmo". Będziemy musieli przyjąć mocniejsze założenia o funkcji f.

Twierdzenie V.11 (Lagrange'a o postaci reszty Taylora). *Jeżeli f jest $(n+1)$-krotnie różniczkowalna w przedziale $(x_0?x)$ oraz n-ta pochodna f jest ciągła w punktach x_0 i x, to istnieje $c \in (x_0?x)$ taki, że*

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}. \tag{V.12}$$

Dowód.
Rozważmy dwie pomocnicze funkcje $\varphi, \psi \colon [x_0?x] \to \mathbb{R}$ zadane wzorami (t jest zmienną):

$$\varphi(t) := f(x) - T_{n,f,t}(x), \quad \psi(t) := (x-t)^{n+1}.$$

Do tych funkcji zastosujemy twierdzenie Cauchy'ego (tw. V.5). Istnieje zatem $c \in (x_0?x)$ takie, że

$$(\varphi(x_0) - \varphi(x)) \cdot \psi'(c) = (\psi(x_0) - \psi(x)) \cdot \varphi'(c). \tag{V.13}$$

Uwzględniając teraz, że dla $t \in (x_0?x)$

$$\varphi'(t) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n, \quad \psi'(t) = -(n+1)(x-t)^n$$

(rachunki prowadzące do wzoru na $\varphi'(t)$ pozostawiam Państwu...) oraz

$$\varphi(x) = f(x) - f(x) = 0, \quad \psi(x) = 0,$$

$$\varphi(x_0) = R_{n,f,x_0}(x) = R_n(x), \quad \psi(x_0) = (x-x_0)^{n+1},$$

po podstawieniu do (V.13) otrzymujemy (V.12). □

Uwagi.

1. Nieco może zawiłe założenia twierdzenia można oczywiście nieco wzmocnić i zakładać po prostu, że f jest $(n+1)$-krotnie różniczkowalna w $[x_0?x]$.

2. Gdy $n = 0$, to $T_0(x) = f(x_0)$, więc uzyskujemy dokładnie twierdzenie Lagrange'a o wartości średniej (tw. V.4).

3. Należy pamiętać o tym, że liczba c z tezy twierdzenia, nie jest żadną „uniwersalną" stałą, ale może zależeć dosłownie od wszystkiego, tj. od f, x_0, x oraz n.

Najbardziej chyba typowy przykład zastosowania twierdzenia Lagrange'a o postaci reszty Taylora to znajdowanie przybliżeń wymiernych rozmaitych liczb z kontrolą wielkości błędu przybliżenia.

Przykład. Dotąd niezbyt wiele wiedzieliśmy na temat wartości liczby e. Właściwie jedynie, że $2 < e < 3$ (choć dzięki definicji e oszacowanie z dołu łatwo można było poprawić). Obecnie bez trudu możemy np. wykazać, że „$e = 2,7...$"[81]. Rozważmy bowiem $f = \exp$ oraz $x_0 = 0$, $x = 1$. Mamy $e = f(1) = T_n(1) + R_n(1)$, gdzie $T_n(1) = \sum_{k=0}^{n} \frac{\exp(0)}{k!} 1^k = \sum_{k=0}^{n} \frac{1}{k!}$ oraz z twierdzenia V.11 $R_n(1) = \frac{\exp(c_n)}{(n+1)!}$ dla pewnego $c_n \in (0;1)$. W szczególności zatem $0 < R_n(1) < \frac{e}{(n+1)!} < \frac{3}{(n+1)!}$. Biorąc zatem $n = 5$ uzyskujemy przybliżenie

$$e \approx 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} = 2,71(6)$$ [82]

[81] Te „..." oznaczają jakieś dalsze cyfry w *rozwinięciu dziesiętnym*, którego nie definiowaliśmy dotąd i niestety nie zdefiniujemy (z braku czasu). Zachęcam do samodzielnego zdefiniowania i wykazania istnienia dla dowolnej liczby rzeczywistej.

[82] Zapis dziesiętny „z okresem" (x) uważam (z konieczności) za znany.

z błędem mniejszym niż $\frac{3}{6!} = \frac{3}{720} = \frac{1}{240}$, a zatem, zgodnie z obietnicą, „$e = 2,7...$". Przy odrobinie większej pracowitości możemy też uzyskać „$e = 2,71...$", co pozostawiam Czytelnikom/Słuchaczom.

Inne ważne zastosowanie twierdzenia V.11 to rozwijanie pewnych funkcji w szeregi potęgowe, o czym wspominaliśmy już nieco w podrozdziale IV.4. Niech f będzie funkcją różniczkowalną dowolną liczbę razy w punkcie x_0. Wówczas szereg potęgowy

$$\sum_{n=0}^{+\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

ma poprawnie zdefiniowane współczynniki. Nazywamy go *szeregiem Taylora funkcji* f o środku w x_0 (a w szczególnym przypadku $x_0 = 0$ używa się też nazwy *szereg Maclaurina*). Naturalne pytanie:

czy szereg ten jest zbieżny do $f(x)$*?*

nie ma oczywiście jednoznacznej ogólnej odpowiedzi. W każdym razie często (niestety?), **nie jest prawdą**, że

$$f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

dla wszystkich x, mimo iż oczekiwalibyśmy być może, że taka równość zachodzi. Odpowiedź pozytywna jest na pewno dla $x = x_0$, i czasem (tj. dla pewnych f) **tylko** wtedy! Dla pewnych „dobrych" funkcji równość zachodzi dla wszystkich x z pewnego otoczenia $(x_0 - \delta, x_0 + \delta)$ punktu x_0 ($\delta > 0$ oczywiście) — mówimy wtedy, że f jest *analityczna* w otoczeniu punktu x_0. Pytanie powyższe może być trudniejsze niż pytanie o samą zbieżność powyższego szeregu (którym zajmowaliśmy się w podrozdziale IV.4.) — tu ważna jest nie tylko zbieżność, ale również to, by suma była równa właśnie $f(x)$.

Przykładami „w pełni pozytywnymi", tj. takimi, dla których zbieżność szeregu Taylora o środku x_0 do $f(x)$ ma miejsce przy każdym $x \in \mathbb{R}$ są między innymi funkcje exp, sin, cos przy $x_0 = 0$. Dla wymienionych funkcji dowód tego faktu jest dość oczywisty, bowiem jak łatwo wyliczyć, szeregi Taylora, które otrzymamy w tych przykładach to znane nam już dobrze wcześniej (z rozdziału IV) szeregi potęgowe dające rozwinięcia funkcji exp, sin i cos. Nie jest to wcale sprawa przypadku — jest to związane z następującym ogólnym wynikiem.

Fakt. *Jeżeli f posiada rozwinięcie w szereg potęgowy o środku w x_0 zbieżne do $f(x)$ dla $x \in (x_0 - r, x_0 + r)$ przy pewnym $r > 0$, to tym szeregiem potęgowym jest szereg Taylora funkcji f o środku w x_0.*

Na razie pominiemy dowód — wrócimy do niego jeszcze w następnym rozdziale.

W szczególności z powyższego faktu wynika różniczkowalność w x_0 dowolną liczbę razy funkcji zadanej szeregiem potęgowym o środku w x_0. Jest to też wzmocnienie faktu dotyczącego jednoznaczności rozwijania funkcji w szereg potęgowy — patrz np. zadanie IV.21. Jeżeli więc znamy już jakieś rozwinięcie funkcji f w szereg potęgowy, to odpowiedź na zadane wcześniej pytanie o zbieżność szeregu Taylora do $f(x)$ jest pozytywna (dla odpowiednich x). Tak jest zatem np. dla funkcji $f(x) := \frac{1}{1-x}$ dla $x \in (-1; 1)$ bo $f(x) = \sum_{n=0}^{+\infty} x^n$ dla takich x (szereg geometryczny).

Na zakończenie naszych rozważań dotyczących wzoru Taylora zajmijmy się jeszcze jedną z pominiętych tu dotąd funkcji — mianowicie funkcją potęgową z dowolnym wykładnikiem $\alpha \in \mathbb{R}$ (powyżej mieliśmy jedynie $\alpha = -1$), dla której nie znamy jak dotąd żadnego ogólnego rozwinięcia w szereg potęgowy. Pewne, na razie tylko częściowe, informacje o takim rozwinięciu uzyskamy właśnie z twierdzenia Lagrange'a o postaci reszty Taylora.

Przykład. Niech $\alpha \in \mathbb{R}$ oraz $f\colon (-1;+\infty) \to \mathbb{R}$, $f(x) = (1+x)^\alpha$. Łatwo wyliczyć, że $f^{(n)}(0) = \alpha \cdot (\alpha-1) \cdot \ldots \cdot (\alpha-(n-1))$, a zatem szereg Taylora dla f o środku w 0 ma postać

$$\sum_{n=0}^{+\infty} \binom{\alpha}{n} x^n,$$

gdzie

$$\binom{\alpha}{n} := \frac{\alpha \cdot \ldots \cdot (\alpha-(n-1))}{n!}$$

jest uogólnieniem znanego symbolu Newtona na przypadek dowolnego $\alpha \in \mathbb{R}$ (dla $\alpha \in \mathbb{N}$, $\alpha \geqslant n$ obie definicje pokrywają się oczywiście). Nietrudno tu wykazać w oparciu o twierdzenie V.11, że gdy $1 > x > 0$, to $R_n(x) \xrightarrow[n\to+\infty]{} 0$ (szczegóły zostawiam jako zadanie — patrz zadanie V.29). Niestety informacje zawarte w tym twierdzeniu okazują się za słabe, by wykazać tego typu zbieżność dla $-1 < x < 0$ (przy dowolnym α), choć zbieżność taka także ma miejsce. Potrzebne są tu jednak inne metody... Warto zauważyć, że uzyskana tu (częściowo) równość

$$(1+x)^\alpha = \sum_{n=0}^{+\infty} \binom{\alpha}{n} x^n \quad \text{dla } |x| < 1$$

ma ścisły związek z tzw. dwumianem Newtona, tj. wzorem na $(a+b)^\alpha$ dla $\alpha \in \mathbb{N}$. Oczywiście, gdy $\alpha \in \mathbb{N}$, to powyższa suma „kończy się" de facto na $n = \alpha$, bowiem wtedy przy $n \geqslant \alpha + 1$ zachodzi $\binom{\alpha}{n} = 0$.

Zadania do Rozdziału V

∀ 1. [83] Znajdź wzory na pochodne funkcji zadanych poniższymi wzorami:

(a) x^x dla $x > 0$;

(b) $x^{(x^7)}$ dla $x > 0$;

(c) $(x^x)^7$ dla $x > 0$;

(d) $\log_{(2+x^2)}(1 + x^2)$ dla $x \in \mathbb{R}$.

2. Rozważamy tzw. *funkcje hiperboliczne* sinh, cosh i tgh będące swego rodzaju analogami funkcji trygonometrycznych, zadane wzorami:

$$\sinh x = \frac{e^x - e^{-x}}{2}, \quad \cosh x = \frac{e^x + e^{-x}}{2}, \quad \operatorname{tgh} x = \frac{\sinh x}{\cosh x} \quad \text{dla } x \in \mathbb{R}.$$

Wykaż wzory:

(a) $(\cosh x)^2 - (\sinh x)^2 = 1$;

(b) $\sinh' = \cosh$;

(c) $\cosh' = \sinh$;

(d) $\operatorname{tgh}'(x) = \frac{1}{(\cosh x)^2} = 1 - (\operatorname{tgh} x)^2$ dla $x \in \mathbb{R}$.

Wykaż, że funkcje $\sinh\colon \mathbb{R} \to \mathbb{R}$, $(\cosh|_{\mathbb{R}_+})\colon \mathbb{R}_+ \to (1; +\infty)$ oraz $\operatorname{tgh}\colon \mathbb{R} \to (-1; 1)$ są odwracalne, oraz znajdź wzory na pochodne funkcji do nich odwrotnych **w oparciu o twierdzenie V.1**.

∀ 3. [84] Znajdź maksymalne przedziały monotoniczności, ekstrema lokalne oraz kresy dla funkcji zadanych poniższymi wzorami:

(a) x^x dla $x > 0$;

(b) $\frac{x^2+1}{x^2+x+1}$ dla $x \in \mathbb{R}$;

(c) $x^{1000} \cdot e^{-x}$ dla $x \in \mathbb{R}$;

(d) $\frac{x^4}{(1+x)^3}$ dla $x \in \mathbb{R} \setminus \{-1\}$;

(e) $|x^2 + 2x - 3| + \frac{3}{2} \ln|x|$ dla $x \in \mathbb{R} \setminus \{0\}$;

(f) $\sin(\sin x)$ dla $x \in \mathbb{R}$;

(g) $\sqrt[3]{\frac{x^2}{x+1}}$ dla $x \in \mathbb{R} \setminus \{-1\}$.

4. Niech $A = \{a_n \colon n \in \mathbb{N}\}$. Znajdź $\sup A$ i $\inf A$, gdy $a_n =$

(a) $\sqrt[n]{n}$;

(b) $n^5 \cdot 2^{-n}$.

5. Dla poniższych $f\colon \mathbb{R} \to \mathbb{R}$ zbadaj, czy f jest różniczkowalna oraz czy f' jest ciągła (w przypadku różniczkowalności f)

(a) $f(x) = |x|^\alpha$ w zależności od parametru $\alpha > 0$;

[83] Przynajmniej 2 przykłady.

[84] Przynajmniej 2 przykłady.

(b) $f(x) = \begin{cases} e^x & \text{dla } x \leqslant 0 \\ ax+b & \text{dla } x > 0 \end{cases}$ w zależności od parametrów $a, b \in \mathbb{R}$;

(c) $f(x) = \begin{cases} 0 & \text{dla } x = 0 \\ |x|^\alpha \sin\frac{1}{x} & \text{dla } x \neq 0 \end{cases}$ w zależności od parametru $\alpha > 0$.

$\forall$ 6. Niech $f\colon (a;b) \to \mathbb{R}$ oraz $x_0 \in (a;b)$. Definiujemy *pochodną symetryczną* w x_0

$$f'_{sym}(x_0) := \lim_{h\to 0} \frac{f(x_0+h) - f(x_0-h)}{2h},$$

o ile granica ta istnieje.

(a) Wykaż, że jeśli $f'(x_0)$ istnieje, to istnieje też $f'_{sym}(x_0)$ i $f'_{sym}(x_0) = f'(x_0)$.

(b) Czy z istnienia $f'_{sym}(x_0)$ wynika istnienie $f'(x_0)$?

(c) Czy z istnienia i skończoności $f'_{sym}(x_0)$ wynika ciągłość f w x_0?

7. Dla $c \in \mathbb{R}$ definiujemy $f_c\colon \mathbb{R} \to \mathbb{R}$ wzorem $f_c(x) = x^2(\varphi(x) + c)$, gdzie

$$\varphi(x) = \begin{cases} 0 & \text{dla } x = 0 \\ \sin\frac{1}{x} & \text{dla } x \neq 0. \end{cases}$$

Wykaż, że jeżeli $c \geqslant 1$, to f_c posiada minimum lokalne w 0, ale przy dowolnym $r > 0$ funkcja $(f_c)|_{[0;r]}$ ani $(f_c)|_{[-r;0]}$ nie jest monotoniczna. Dla jakich c minimum lokalne w 0 jest ściśle[85)] ? Czy f_c jest różniczkowalna?

$\forall$ 8. [86)] Wykaż poniższe nierówności, wykorzystując twierdzenia rachunku różniczkowego (tj. dotyczące pochodnych):

(a) $(x+y)^\alpha \leqslant (\geqslant) x^\alpha + y^\alpha$ dla $x, y \geqslant 0$ i $\alpha \leqslant (\geqslant) 1$;

(b) $xe^{-x^2} + ye^{-y^2} + ze^{-z^2} \leqslant \sqrt{\frac{9}{2e}}$ dla $x, y, z \in \mathbb{R}$;

(c) $\ln(1+x) > (<) x - \frac{x^2}{2}$ dla $x > 0$ $(-1 < x < 0)$;

(d) $\sqrt[3]{1+x} > 1 + \frac{1}{3}x - \frac{1}{9}x^2$ dla $x > 0$.

9. Znajdź pewne (ewentualnie wersja troszkę trudniejsza: wszystkie) takie $\alpha \in \mathbb{R}$, że dla dowolnego $x > -1$ zachodzi $\ln(1+x) \leqslant x - \frac{x^2}{2} + \alpha x^3$.

$\forall$ 10. [87)] Ile pierwiastków (tzn. rozwiązań) posiada równanie (x jest "niewiadomą", $x \in \mathbb{R}$):

(a) $x^{11} - 11x + 1 = 0$;

(b) $6\ln(x^2+1) = e^x$;

(c) $a^x = x$, w zależności od parametru $a > 0$.

11. Zbadaj dla jakich $a \in \mathbb{R}$ funkcja $f_a\colon \mathbb{R} \to \mathbb{R}$, $f_a(x) = x^5 - 10x^2 + ax$ jest

(a) różnowartościowa,

(b) monotoniczna,

[85)] Minimum lokalne f w x_0 jest *ściśle* (inna nazwa: *istotne*) wtw $\exists_{\delta>0} \forall_{\substack{x\in D \\ 0<|x-x_0|<\delta}} f(x) > f(x_0)$. Analogicznie dla maksimum.

[86)] Przynajmniej 2 przykłady.

[87)] Przynajmniej 1 przykład.

(c) ściśle monotoniczna.

12. Zbadaj dla jakich $a \in \mathbb{R}$ funkcja $g_a \colon \mathbb{R} \to \mathbb{R}$, $g_a(x) = ax + \sin x$ jest
   (a) rosnąca;
   (b) ściśle rosnąca;
   (c) malejąca;
   (d) ściśle malejąca.

∀ 13. Niech $f \colon \mathbb{R} \to \mathbb{R}$ będzie różniczkowalna. Wykaż, że $f' = f$ wtw istnieje $c \in \mathbb{R}$ takie, że $f = c \cdot \exp$. Uwaga: to już nieco „poważniejsze” równanie różniczkowe, niż $f' = 0$...

14. Wykaż, że dla dowolnego $x > -1$ zachodzi

$$\operatorname{arctg} x + \operatorname{arctg} \frac{1-x}{1+x} = \frac{\pi}{4}.$$

   Czy podobnego typu rezultat ma miejsce dla $x < -1$?

15. Rozważamy wszystkie trójkąty wpisane w okrąg o promieniu 1. Czy wśród nich istnieje taki, którego obwód jest największy? Jeżeli tak, to jaki jest jego obwód?

16. Trójkąty prostokątne o obwodzie 1 obracamy wokół przeciwprostokątnej. Czy dla jakiegoś z nich objętość otrzymanej bryły obrotowej jest największa? Jeśli tak, to znajdź tę największą objętość.

   Uwaga: w zadaniach V.15 i V.16 oczywiście można stosować znane ze szkoły wzory geometryczne.

17. Niech $f \colon [a; b) \to \mathbb{R}$ będzie ciągła w a i różniczkowalna w $(a; b)$. Wykaż, że jeżeli istnieje granica $\lim_{x \to a+} f'(x) =: g$, to także istnieje $f'_+(a)$ i równa jest g. Uwaga! Wynika z tego, że ewentualna nieciągłość pochodnej funkcji różniczkowalnej (na przedziale) nie może być „zbyt trywialna” — nie mogą pojawiać się zwykłe „skoki”, tj. sytuacje, gdy granica istnieje, ale jest różna od wartości w punkcie granicznym.

18. Niech I będzie przedziałem oraz $f \colon I \to \mathbb{R}$ — funkcją różniczkowalną. Wykaż, że f jest lipschitzowska (patrz zadanie IV.17) wtw f' jest ograniczona.

19. Wykaż jednostajną ciągłość funkcji zadanych wzorami:
   (a) $\sqrt{x^2 + x}$ dla $x \geqslant 0$;
   (b) $\sin(\ln x)$ dla $x \geqslant 1$.

   Wskazówka: wykaż najpierw, że jeśli f jest jednostajnie ciągła na dwóch przedziałach, które nie są rozłączne, to jest też jednostajnie ciągła na ich sumie.

20. Wykaż, że pochodna funkcji różniczkowalnej na przedziale posiada własność Darboux, tzn. wykaż, że jeżeli $f \colon [a; b] \to \mathbb{R}$ jest różniczkowalna oraz $c \in (f'(a)?f'(b))$, to istnieje $x \in (a; b)$ takie, że $f'(x) = c$.

∀ 21. [88] Wykorzystując twierdzenie Lagrange'a o wartości średniej (twierdzenie V.4) zbadaj zbieżność następujących szeregów:

[88] Przynajmniej 1 przykład.

(a) $\sum_{n=1}^{+\infty} \frac{\ln(n+1)-\ln(n)}{\sqrt{n}}$;

(b) $\sum_{n=1}^{+\infty} \frac{1}{n}\sqrt{e^{\frac{1}{n}} - e^{\frac{1}{n+1}}}$;

(c) $\sum_{n=1}^{+\infty} \sqrt{n}(\operatorname{arctg}(n+1) - \operatorname{arctg}(n))$.

$\forall$ 22. [89] Znajdź poniższe granice. Każdy z przykładów **spróbuj** zbadać korzystając z reguły de l'Hospitala i **odrębnie**, korzystając ze wzoru Taylora.

(a) $\lim_{x\to 0+} x\ln x$;

(b) $\lim_{x\to 1} x^{\frac{1}{1-x}}$;

(c) $\lim_{x\to 1+} \left(\frac{1}{\ln x} - \frac{1}{x-1}\right)$;

(d) $\lim_{x\to \frac{\pi}{2}} \frac{\operatorname{tg} 3x}{\operatorname{tg} x}$;

(e) $\lim_{x\to +\infty} \frac{x+\sin x}{x-\sin x}$;

(f) $\lim_{x\to 0} \frac{e^x-x-1}{1-\cos x}$;

(g) $\lim_{x\to +\infty} x - x^2\ln(1+\frac{1}{x})$;

(h) $\lim_{x\to 0} \left(\frac{1}{x} - \frac{1}{\sin x}\right)$;

(i) $\lim_{x\to 0} \frac{\sin x - x}{x^3}$;

(j) $\lim_{x\to 0} \frac{2x\ln(1+x)+x^3}{\sqrt{1-e^{-x^4}}}$;

(k) $\lim_{x\to 0} \frac{\sin x + e^x - (1+2x+\frac{x^2}{2})}{x^4}$.

$\forall$ 23. [90] Znajdź przybliżenia wymierne poniższych liczb z podaną dokładnością d:

(a) $\sqrt{e}$, $d = 0,001$;

(b) $\cos^2 1$, $d = 0,001$;

(c) $\ln(\frac{3}{2})$, $d = \frac{1}{20}$.

24. Poniższe liczby zapisz w postaci sum szeregów o wyrazach wymiernych:

(a) $\frac{\sin\sqrt{3}}{\sqrt{3}}$;

(b) $\ln(\frac{8}{3})$;

(c) $\frac{1}{\sqrt[3]{2}}$.

25. Udowodnij pominięte w dowodzie z wykładu przypadki w regule de l'Hospitala (twierdzenie V.7).

26. Niech $f\colon D \to \mathbb{R}$ będzie n-krotnie różniczkowalna w $x_0 \in D$, przy czym x_0 ma otoczenie w D będące przedziałem. Wykaż, że jeśli w jest wielomianem stopnia $\leqslant n$ takim, że $f(x) = w(x) + o((x-x_0)^n)$, przy $x \to x_0$, to $w = T_{n,f,x_0}$.

27. Wykorzystując wzór Taylora zbadaj zbieżność poniższych szeregów:

(a) $\sum_{n=1}^{+\infty} \left(e^{\frac{1}{\sqrt{n}}} - 1 - \frac{1}{\sqrt{n}} - \frac{1}{2n}\right)$;

(b) $\sum_{n=1}^{+\infty} \left(\ln(n+1) - \ln(n) - \frac{1}{n}\right)$;

(c) $\sum_{n=1}^{+\infty} |\operatorname{tg}(\frac{1}{\sqrt{n}}) - \frac{1}{\sqrt{n}}|^\alpha$ w zależności od $\alpha > 0$;

(d) $\sum_{n=1}^{+\infty} \left(\sqrt[n]{a} - \frac{\sqrt[n]{b}+\sqrt[n]{c}}{2}\right)$ w zależności od $a, b, c > 0$.

[89] Przynajmniej 3 przykłady.
[90] Przynajmniej 1 przykład.

28. Wykaż, że jeśli $f(x) = a_0 + a_1 x + \ldots + a_n x^n$ dla $x \in \mathbb{R}$, to dla dowolnego $x_0 \in \mathbb{R}$ istnieją $b_0, b_1, \ldots, b_n \in \mathbb{R}$ takie, że

$$\forall_{x\in\mathbb{R}}\ f(x) = b_0 + b_1(x - x_0) + \ldots + b_n(x - x_0)^n.$$

Jakim wzorem zadane są powyższe współczynniki b_k?

29. Wykaż, że $(1+x)^\alpha = \sum_{n=0}^{+\infty} \binom{\alpha}{n} x^n$ dla $x \in [0;1)$ w oparciu o twierdzenie Lagrange'a o postaci reszty Taylora (twierdzenie V.11).

30. Niech $f\colon (a;b) \to \mathbb{R}$ będzie n-krotnie różniczkowalna w $c \in (a;b)$, $n \geqslant 2$. Wykaż następujące kryterium na ekstrema.

Jeżeli $f^{(k)}(c) = 0$ dla dowolnego $k = 1, \ldots, n-1$ oraz $\alpha := f^{(n)}(c) \neq 0$, to

- *jeżeli n jest parzyste i $\alpha > 0 (< 0)$, to f posiada ścisłe minimum (maksimum) lokalne w c;*
- *jeżeli n jest nieparzyste, to f nie posiada ekstremum lokalnego w c.*

31. Wykaż część twierdzenia V.8 dotyczącą iloczynu oraz złożenia funkcji n-krotnie różniczkowalnych w punkcie (dowody pominięte na wykładzie).

32. Znajdź wzory na $f^{(n)}(x)$ dla

    (a) $f(x) = xe^x$, $n = 1000$;

    (b) $f(x) = x^2 \sin(5x)$, $n = 100$.

33. Znajdź $\sup\{n \in \mathbb{N} \colon f \in C^n(\mathbb{R})\}$ dla następujących $f\colon \mathbb{R} \to \mathbb{R}$:

    (a) $f(x) = \begin{cases} x^7 & \text{dla } x > 0 \\ x^5 & \text{dla } x \leqslant 0 \end{cases}$;

    (b) $f(x) = \begin{cases} e^{\frac{1}{x}} & \text{dla } x > 0 \\ 0 & \text{dla } x \leqslant 0 \end{cases}$.

34. Wykaż, że jeżeli $f\colon (a;b) \to \mathbb{R}$ jest wypukła, to f jest ciągła. Znajdź przykład pokazujący, że funkcja wypukła może nie być ciągła w końcu przedziału określoności (gdy koniec ten do przedziału należy).

35. Wykaż, że jeżeli $f\colon I \to \mathbb{R}$ jest ciągła (I jest przedziałem) oraz $\forall_{x,y\in I}\ f\left(\frac{x+y}{2}\right) \leqslant \frac{f(x)+f(y)}{2}$, to f jest wypukła.

36. Rozstrzygnij, które spośród operacji: dodawanie, odejmowanie, mnożenie, dzielenie, składanie, różniczkowanie (oczywiście, przy założeniu, że dana operacja jest wykonalna) zachowują wypukłość funkcji.

37. Przedstaw szczegóły dowodu faktu o nierówności Jensena (str. 76).

$\forall$ 38. [91] Wykaż następujące nierówności w oparciu o nierówność Jensena:

    (a) $\sqrt[n]{x_1 \cdot \ldots \cdot x_n} \leqslant \frac{x_1+\ldots+x_n}{n}$ dla $x_1, \ldots, x_n \geqslant 0$, $n \in \mathbb{N}$;

    (b) $\left(\sum_{k=1}^n x_k\right)^\alpha \leqslant (\geqslant) n^{\alpha-1} \cdot \sum_{k=1}^n x_k^\alpha$ dla $x_1, \ldots, x_n \geqslant 0$, $n \in \mathbb{N}$ i $\alpha \geqslant 1$ ($0 < \alpha \leqslant 1$);

[91] Przynajmniej 1 przykład.

(c) $n \cdot \left(\frac{n+1}{2}\right)^{\left(\frac{n+1}{2}\right)} \leqslant \sum_{k=1}^{n} k^k$ dla $n \in \mathbb{N}$.

39. Wykaż, że jeśli $f\colon (a;b) \to \mathbb{R}$ jest wypukła oraz różniczkowalna w punkcie $x_0 \in (a;b)$, to wykres f „leży nad” styczną do wykresu f dla x_0, tzn. $\forall_{x\in(a;b)}\ f(x) \geqslant f(x_0)+f'(x_0)(x-x_0)$.

40. Znajdź wszystkie funkcje $f\colon [a;b] \to \mathbb{R}$, które są wypukłe i wklęsłe jednocześnie.

41. Wykaż, że jeżeli f jest wypukła i odwracalna, to f^{-1} jest wypukła lub wklęsła (wyjaśnij od od czego to zależy).

42. Wykaż, że jeżeli $f\colon [a;b] \to \mathbb{R}$ jest ciągła oraz $f|_{(a;b)}$ jest wypukła, to f też jest wypukła.

VI Zbieżność ciągów i szeregów funkcji

[około $1\frac{1}{2}$ wykładu]

1. O różnych pojęciach zbieżności ciągu funkcji

W II i III rozdziale zajmowaliśmy się zbieżnością ciągów i szeregów liczbowych. Ale czy można mówić o zbieżności w przypadku ciągów, których wyrazami są nie liczby lecz funkcje (takie ciągi nazywamy *ciągami funkcyjnymi*)? No cóż, o tym że można, świadczy choćby tytuł tego rozdziału. Co więcej, w odróżnieniu od sytuacji jaką mieliśmy dla ciągów liczbowych, poznamy nie jeden, ale dwa, a właściwie nawet trzy rodzaje zbieżności ciągów funkcyjnych.

Niech $f, f_n : D \to \mathbb{R}$ [92] dla $n \geqslant n_0$. Naturalne wydaje się określenie, że ciąg funkcyjny $\{f_n\}$ jest zbieżny do funkcji f wtw

$$\forall_{x\in D}\ f_n(x) \to f(x). \tag{VI.1}$$

Taki rodzaj zbieżności nazywamy *zbieżnością punktową* i oznaczamy symbolem[93]

$$f_n \to f,$$

a zatem $f_n \to f$ wtw zachodzi (VI.1). Gdy taka zbieżność zachodzi, to funkcję f nazywamy *granicą* ciągu $\{f_n\}$ (mówimy też *granica punktowa*).

Przyjrzyjmy się nieco głębiej takiej zbieżności. Gdy skorzystamy z definicji granicy ciągu liczbowego $\{f_n(x)\}_{n\geqslant n_0}$, to powyższą definicję możemy w sposób równoważny zapisać w postaci

$$\forall_{x\in D}\ \forall_{\epsilon>0}\ \exists_{N\geqslant n_0}\ \forall_{n\geqslant N}\ |f_n(x) - f(x)| < \epsilon \tag{VI.2}$$

W warunku (VI.2) indeks N możemy zatem dobierać w sposób zależny zarówno od ϵ jak i od $x \in D$. Gdy przypomnimy sobie pojęcie ciągłości jednostajnej (patrz podrozdział IV.3) oraz to, co odróżnia jej definicję od definicji „zwykłej" ciągłości, naturalny wyda nam się pomysł, by zmodyfikować warunek (VI.2) i dopuścić jedynie „jednostajny po x" (tzn., taki sam dla wszystkich x, niezależny od x) dobór N do ϵ. Otrzymamy wtedy warunek następujący [94]

$$\forall_{\epsilon>0}\ \exists_{N\geqslant n_0}\ \forall_{n\geqslant N}\ \forall_{x\in D}\ |f_n(x) - f(x)| < \epsilon \tag{VI.3}$$

(patrz rys. 13 — wykres f_n dla $n \geqslant N$ jest zawarty cały w „pasie" pomiędzy $f-\epsilon$ a $f+\epsilon$). I takim właśnie warunkiem definiujemy drugi rodzaj zbieżności — *zbieżność jednostajną*, którą oznaczamy symbolem

$$f_n \rightrightarrows f.$$

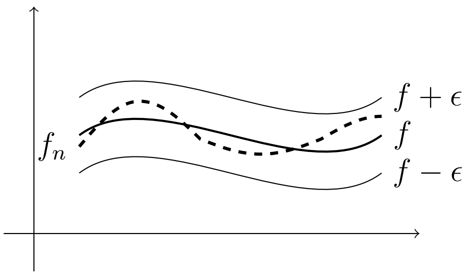
Tzn., $f_n \rightrightarrows f$ wtw zachodzi (VI.3).

Uwagi (proste, ale ważne).

[92] Na ogół w tym rozdziale D oznacza dziedzinę rozważanych funkcji; zazwyczaj będziemy tak przyjmować bez przypominania.

[93] Ten zapis przy pomocy „$\to$" jest nieco dwuznaczny, bo tego samego symbolu używaliśmy przy rozważaniu granicy ciągu liczbowego (choćby przed chwilą, w (VI.1)).

[94] Pamiętajmy o tym, że sąsiadujące ze sobą kwantyfikatory ogólne „$\forall$ " możemy przestawiać — dotyczy to zarówno (VI.2) jak i (VI.3). Zmianą **istotną** jest dopiero przestawienie „$\forall$ " i „$\exists$ ".



Rysunek 13: Wykres f_n jest cały zawarty w pasie pomiędzy $f-\epsilon$ a $f+\epsilon$.

1. Zbieżność jednostajna to „lepszy” rodzaj zbieżności, tzn.,
$$f_n \rightrightarrows f \Rightarrow f_n \to f.$$

2. Przy obu rodzajach zbieżności granica wyznaczona jest jednoznacznie. W przypadku zbieżności punktowej wynika to z analogicznego faktu dla granicy ciągu liczbowego, a zatem w przypadku zbieżności jednostajnej wystarczy użyć uwagę 1.

3. W warunku (VI.3), ze względu na „dowolność” $\epsilon > 0$, nierówność „$< \epsilon$” można oczywiście zastąpić przez „$\leqslant \epsilon$”. Korzystając teraz z definicji kresu górnego możemy ten warunek zapisać równoważnie w postaci
$$\forall_{\epsilon>0}\ \exists_{N\geqslant n_0}\ \forall_{n\geqslant N}\ \sup_{x\in D}|f_n(x)-f(x)| \leqslant \epsilon,$$ [95]
co (analogicznie jak przed chwilą) równoważne jest warunkowi z „$< \epsilon$”, a to z kolei oznacza dokładnie, że
$$\lim_{n\to+\infty}(\sup_{x\in D}|f_n(x)-f(x)|) = 0$$
czyli, że ciąg **liczbowy**[96] $\{\sup_{x\in D}|f_n(x)-f(x)|\}_{n\geqslant n_0}$ ma granicę 0.
Jeżeli więc dla $g : D \to \mathbb{R}$ oznaczymy
$$\|g\| := \sup_{x\in D}|g(x)|,$$
to mamy:
$$f_n \rightrightarrows f \quad \text{wtw} \quad \|f_n - f\| \to 0.$$
Jest to wygodna „alternatywna definicja” zbieżności jednostajnej, bowiem sprowadza ona problem do badania zbieżności pewnego ciągu liczbowego. Symbol $\|\cdot\|$ używany jest do oznaczania *normy*, czyli wielkości wyrażającej w jakimś sensie długość wektorów — tu tymi wektorami są funkcje o wartościach w $\mathbb{R}$ [97] . Można definiować rozmaite normy — ta konkretna tu zdefiniowana bywa nazywana „normą supremum” i czasem oznacza się ją przez $\|\cdot\|_\infty$. Każda norma musi spełniać kilka warunków (o tym wspomnimy jeszcze w przyszłości...) i wybierając jakąś normę zawsze możemy w sposób taki jak wyżej zdefiniować pewien „nowy” rodzaj zbieżności. My jednak teraz zadowolimy się tą jedną[98] normą.

[95] Symbol $\sup_{x\in X} g(x)$ to skrót (wygodny) od $\sup\{g(x)\in\mathbb{R} : x \in X\}$.

[96] Ściślej, ciąg ten ma wyrazy w $\overline{\mathbb{R}}$ (może zdarzyć się $+\infty$), ale definicja granicy dla tego typu ciągów przenosi się w sposób oczywisty.

[97] Wektory — to po prostu elementy przestrzeni liniowej, w naszym wypadku chodzi o przestrzeń funkcji $f : D \to \mathbb{R}$ z naturalnymi działaniami.

[98] Tak naprawdę nie jedną, bo każdy zbiór D wyznacza inną normę supremum, „mierzącą” funkcje określone na zbiorze D. W razie potrzeby odróżnienia możemy np. używać oznaczenia $\|\cdot\|_D$.

Przykład. Niech $D \subset R$ i $f_n : D \to \mathbb{R}$ niech będą zadane wzorem

$$f_n(x) = \frac{x}{n}$$

dla $x \in D$ i $n \in \mathbb{N}$. Rozważymy parę rozmaitych dziedzin D. Jednak ponieważ $\forall_{x\in\mathbb{R}}\ \frac{x}{n} \to 0$, zatem niezależnie od wyboru D mamy $f_n \to 0$, gdzie tym razem 0 nie oznacza liczby 0 lecz **funkcję** stałą równą 0 (przypominam też dwuznaczny sens „$\to$” użytej tu przed chwilą w dwóch różnych znaczeniach ...). A zatem jeżeli również $f_n \rightrightarrows f$ dla pewnej funkcji f, to na mocy uwagi 1 i 2 jedynym „kandydatem” na f jest także $f = 0$. Niech $D = \mathbb{R}$. Mamy wtedy

$$\|f_n - 0\| = \|f_n\| = \sup_{x\in\mathbb{R}} |\frac{x}{n}| = +\infty \not\to 0,$$

zatem $f_n \not\rightrightarrows 0$, czyli f_n w ogóle nie jest ciągiem funkcyjnym zbieżnym jednostajnie! Teraz rozważmy $D = [-5; 7]$. Wówczas

$$\|f_n - 0\| = \frac{1}{n} \sup_{x\in[-5;7]} |x| = \frac{7}{n} \to 0,$$

zatem $f_n \rightrightarrows 0$ w tym przypadku. Nietrudno uogólnić to na przypadek dowolnego ograniczonego zbioru D — wówczas również $f_n \rightrightarrows 0$, gdyż $M_D := \sup_{x\in D} |x| < +\infty$, skąd $\|f_n - 0\| = \frac{1}{n} M_D \to 0$. A zatem by nasz ciąg $\{f_n\}_{n\geqslant 1}$ był zbieżny jednostajnie zbiór D nie może być „zbyt duży”. Np. $D = \mathbb{R}$ był „za duży” na zbieżność jednostajną, ale ciąg $\{f_n\}_{n\geqslant 1}$ był jednostajnie zbieżny dla D będącego dowolnym przedziałem $[a; b]$.

Powyższy przykład sugeruje wprowadzenie jeszcze jednego rodzaju zbieżności. Będzie on dotyczył tylko funkcji określonych na przedziałach [99] . Będzie to tzw. *zbieżność niemal jednostajna*, którą będziemy oznaczać symbolem

$$f_n \mathrel{\substack{\to \\ \dashrightarrow}} f.$$

Jeśli $f, f_n : I \to \mathbb{R}$, gdzie I — przedział, to $f_n \mathrel{\substack{\to \\ \dashrightarrow}} f$ wtw

$$\forall_{a,b\in I}\ (f_n|_{[a;b]}) \rightrightarrows (f|_{[a;b]}).$$

Już samo oznaczenie sugeruje, że ten rodzaj zbieżności jest gdzieś „pomiędzy” zbieżnością punktową a jednostajną, tzn. że

$$f_n \rightrightarrows f \Rightarrow f_n \mathrel{\substack{\to \\ \dashrightarrow}} f \Rightarrow f_n \to f$$

(np. dla dowodu drugiej implikacji wystarczy rozważać sytuację gdy $a = b$). Przykład takiej sytuacji, że $\{f_n\}_{n\geqslant 1}$ jest zbieżny niemal jednostajnie, ale nie jednostajnie uzyskamy biorąc $D = \mathbb{R}$ w przykładzie powyżej. Oczywiście, także przy tym rodzaju zbieżności granica jest zdefiniowana jednoznacznie i może być nią jedynie granica punktowa. Może się zdarzyć, że to nowe pojęcie zbieżności nie wnosi jednak nic naprawdę nowego. Tak będzie np. wtedy, gdy I samo jest już przedziałem domkniętym — wtedy zbieżność jednostajna i niemal jednostajna są tym samym.

Dla wprowadzonych tu różnych rodzajów zbieżności ciągów funkcyjnych można sformułować nieco twierdzeń analogicznych do odpowiednich twierdzeń dotyczących ciągów liczbowych — np. do twierdzenia o rachunkowych własnościach granicy, w przypadku zbieżności punktowej. Dla zbieżności jednostajnej, jedną z takich ważnych analogii jest odpowiednik twierdzenia o zupełności II.7, w którym zamiast „zwykłego” warunku Cauchy'ego pojawia się „jednostajny” warunek Cauchy'ego. Sprawę tę jednak odkładamy do zadań (patrz — zadanie VI.5).

[99] Można to też uogólnić na funkcje określone na innych zbiorach, ale tu nie będziemy się tym zajmować.

2. Szeregi funkcyjne

Niech $f_n : D \to \mathbb{R}$ dla $n \geqslant n_0$. *Szereg funkcyjny* $\sum_{n=n_0}^{+\infty} f_n$ określamy zupełnie analogicznie jak w przypadku szeregów liczbowych. Utożsamiamy go bowiem z *ciągiem sum częściowych* $\{S_n\}_{n\geqslant n_0}$, który jest w tej sytuacji ciągiem funkcyjnym, przy czym

$$S_n := \sum_{k=n_0}^{n} f_k \quad \text{dla } n \geqslant n_0.$$

Każdy z poznanych tu trzech rodzajów zbieżności w odniesieniu do szeregu funkcyjnego $\sum_{n=n_0}^{+\infty} f_n$ oznacza więc po prostu odpowiednią zbieżność dla $\{S_n\}_{n\geqslant n_0}$. *Sumą szeregu funkcyjnego* nazywa się granicę **punktową** ciągu $\{S_n\}_{n\geqslant n_0}$ (o ile istnieje).

Poznane w IV rozdziale szeregi potęgowe też można utożsamiać z pewnymi szeregami funkcyjnymi. W rozdziale IV szeregiem potęgowym nazywaliśmy rodzinę szeregów liczbowych postaci $\sum_{n=0}^{+\infty} a_n(x-x_0)^n$ rozważanych dla wszystkich $x \in \mathbb{R}$. Jednak zamiast mówić o rodzinie szeregów liczbowych, możemy wyrazy tego szeregu potraktować jako funkcje zmiennej x. Tzn. będziemy mieli tu do czynienia z szeregiem funkcyjnym $\sum_{n=0}^{+\infty} f_n$, gdzie funkcje $f_n : \mathbb{R} \to \mathbb{R}$ są dla $n \geqslant 0$ zadane wzorami $f_n(x) = a_n(x-x_0)^n$ dla $x \in \mathbb{R}$. Tak właśnie będziemy rozumieli pojęcie szeregu potęgowego w tym rozdziale.

Punktowa zbieżności szeregu funkcyjnego $\sum_{n=n_0}^{+\infty} f_n$ to, jak natychmiast widać, po prostu zbieżność szeregu liczbowego $\sum_{n=n_0}^{+\infty} f_n(x)$ dla wszystkich x z dziedziny funkcji f_n (wspólnej dla wszystkich n). Np., gdy „obetniemy" szereg potęgowy do jego zbioru zbieżności, to otrzymamy szereg funkcyjny zbieżny punktowo.

Sformułujemy teraz parę ogólnych wyników dla znacznie "trudniejszej" zbieżności — jednostajnej.

Twierdzenie VI.1 (warunek konieczny zbieżności jednostajnej). *Jeżeli* $\sum_{n=n_0}^{+\infty} f_n$ *jest jednostajnie zbieżny, to* $f_n \rightrightarrows 0$.

Jak widać, jest to analog twierdzenia o warunku koniecznym zbieżności szeregu liczbowego — zresztą dowód także jest analogiczny ...

Dowód.
Niech $S_n := \sum_{k=n_0}^{n} f_k$ dla $n \geqslant n_0$. Dla pewnej funkcji F zachodzi

$$\|S_n - F\| \to 0,$$

a zatem także $\|S_{n-1} - F\| \to 0$. Stąd mamy dla $n \geqslant n_0 + 1$

$$0 \leqslant \|f_n\| = \|S_n - S_{n-1}\| = \|(S_n - F) + (F - S_{n-1})\| \leqslant \|S_n - F\| + \|F - S_{n-1}\|, \quad \text{(VI.4)}$$

przy czym ostatnia nierówność to konsekwencja następującego lematu.

Lemat (nierówność trójkąta dla $\|\cdot\|$).

$$\|f+g\| \leqslant \|f\| + \|g\|.$$

Dowód lematu.
Wystarczy użyć zwykłej nierówności trójkąta (dla modułu $|\cdot|$) oraz definicji kresu górnego. □

Teraz z (VI.4) i z twierdzenia o trzech ciągach otrzymujemy $\|f_n\| \to 0$, czyli $f_n \rightrightarrows 0$. □

Oczywiście istnieje wiele innych warunków koniecznych zbieżności jednostajnej szeregu funkcyjnego. Np. takim warunkiem koniecznym jest oczywiście jego punktowa zbieżność.

Kolejne twierdzenie daje pewien wygodny warunek dostateczny. Przypomina ono nieco twierdzenie o zbieżności szeregu liczbowego bezwzględnie zbieżnego.

Twierdzenie VI.2 (warunek dostateczny zbieżności jednostajnej). *Jeżeli szereg liczbowy $\sum_{n=n_0}^{+\infty} \|f_n\|$ jest zbieżny, to $\sum_{n=n_0}^{+\infty} f_n$ jest jednostajnie zbieżny.*

Dowód pominiemy, ale warto wiedzieć, że nietrudno go uzyskać w oparciu o wspomniany niedawno jednostajny warunek Cauchy'ego (zachęcam do samodzielnych prób dowodu).

Zauważmy, że sformułowany wcześniej warunek konieczny, tzn. $\|f_n\| \to 0$ jest też „warunkiem koniecznym dla warunku dostatecznego" tzn. dla $\sum_{n=n_0}^{+\infty} \|f_n\| < +\infty$.

Uwaga. Przyjrzyjmy się trochę praktycznej skuteczności obu powyższych twierdzeń. Rozważmy więc ciąg $\{\|f_n\|\}_{n \geqslant n_0}$. Są trzy możliwości:

1. $\|f_n\| \nrightarrow 0$ — wówczas $\sum_{n=n_0}^{+\infty} f_n$ **nie** jest jednostajnie zbieżny na mocy twierdzenia VI.1.
2. $\|f_n\| \to 0$, ale $\sum_{n=n_0}^{+\infty} \|f_n\| = +\infty$ — wówczas powyższe twierdzenia **nie dają nic**.
3. $\sum_{n=n_0}^{+\infty} \|f_n\| < +\infty$ — wtedy $\sum_{n=n_0}^{+\infty} f_n$ **jest** jednostajnie zbieżny na mocy twierdzenia VI.2.

A zatem mamy poważną lukę w praktycznej użyteczności powyższych twierdzeń, opisaną możliwością 2. Co robić gdy na taką sytuację natrafimy? Ogólnej recepty nie ma — trzeba każdy taki przypadek badać indywidualnie. Czasem jednak „ratunek" jest banalny: warto sprawdzić czy $\sum_{n=n_0}^{+\infty} f_n$ jest w ogóle punktowo zbieżny — jeśli nie, to tym bardziej nie może być zbieżny jednostajnie. Przykład takiej sytuacji, to $\sum_{n=1}^{+\infty} \frac{x}{n}$, rozważany dla $x \in [0;1]$ (tzn. $f_n : [0;1] \to \mathbb{R}, f_n(x) = \frac{x}{n}$). Mamy wówczas $\|f_n\| = \frac{1}{n} \to 0$, ale $\sum_{n=1}^{+\infty} \|f_n\| = \sum_{n=1}^{+\infty} \frac{1}{n} = +\infty$. Jednak tu brak zbieżności punktowej, bo szereg liczbowy $\sum_{n=1}^{+\infty} \frac{x}{n}$ jest zbieżny jedynie dla $x = 0$.

Sformułujemy teraz często stosowany wniosek z twierdzenia VI.2.

Wniosek (kryterium Weierstrassa). *Jeżeli istnieje ciąg liczbowy $\{c_n\}_{n \geqslant n_0}$ taki, że*

$$\forall_{n \geqslant n_0,\ x \in D}\ |f_n(x)| \leqslant c_n \tag{VI.5}$$

oraz $\sum_{n=n_0}^{+\infty} c_n$ jest zbieżny, to $\sum_{n=n_0}^{+\infty} f_n$ jest jednostajnie zbieżny

Dowód.
Na mocy (VI.5) i definicji normy supremum oraz definicji kresu górnego, dla dowolnego $n \geqslant n_0$ mamy

$$0 \leqslant \|f_n\| \leqslant c_n,$$

a zatem z kryterium porównawczego $\sum_{n=n_0}^{+\infty} \|f_n\|$ — zbieżny. Teza wynika więc z twierdzenia VI.2. □

Jednym z zastosowań powyższego kryterium jest ważny wynik dotyczący szeregów potęgowych. Niech R będzie promieniem zbieżności szeregu potęgowego $\sum_{n=0}^{+\infty} a_n(x - x_0)^n$. Poniżej funkcje będące jego wyrazami rozważamy jedynie na otwartym przedziale jego zbieżności tzn. na $Z_0 := (x_0 - R; x_0 + R)$ (czyli funkcje te obcinamy do Z_0).

Fakt. *Szereg potęgowy jest niemal jednostajnie zbieżny w swoim otwartym przedziale zbieżności*[100] .

Dowód.
Oczywiście możemy ograniczyć się do przypadku $x_0 = 0$. Dla dowodu niemal jednostajnej zbieżności powinniśmy dowodzić zbieżność jednostajną szeregu powstałego po obcięciu odpowiednich funkcji do dowolnego przedziału $[a;b]$ zawartego w $(-R;R)$. Wystarczy więc rozważać $a = -r$ i $b = r$, gdzie $0 \leqslant r < R$. Ale dla dowolnego $x \in [-r;r]$ i $n \geqslant 0$ mamy

[100] Zbieżność **niemal** jednostajna w „całym" przedziale zbieżności też zachodzi, ale dowód tego jest już subtelniejszy . . .

$|a_n x^n| \leqslant |a_n| r^n = |a_n r^n|$, a szereg $\sum_{n=0}^{+\infty} |a_n r^n|$ jest zbieżny, gdyż $r \in (-R, R)$ a szereg potęgowy jest bezwzględnie zbieżny w otwartym przedziale zbieżności (patrz np. lemat ze strony 55). A zatem potrzebna nam zbieżność jednostajna wynika z kryterium Weierstrassa. $\square$

3. Własności granic ciągów i szeregów funkcyjnych

Zajmiemy się tu pytaniem:

Jakie własności wyrazów ciągu (ewentualnie szeregu) funkcyjnego przenoszą się na jego granicę?

Ściślej — zajmiemy się głównie ciągłością i różniczkowalnością. Nietrudno znaleźć przykłady pokazujące, że żadna z tych własności nie zachowuje się przy zbieżności punktowej (zadanie VI.7). Zajmiemy się więc zbieżnością jednostajną i problemem ciągłości.

Twierdzenie VI.3 (o ciągłości granicy). *Jeżeli funkcje f_n są ciągłe dla $n \geqslant n_0$ oraz $f_n \rightrightarrows f$, to f jest ciągła (tzn. granica jednostajnie zbieżnego ciągu funkcji ciągłych jest ciągła).*

Dowód.
Wykażemy ciągłość f w dowolnym punkcie $x \in D$. Niech $x_n \in D$, $x_n \to x$. Musimy wykazać, że $f(x_n) \to f(x)$. Mamy

$$\begin{aligned} |f(x_n) - f(x)| &= |(f(x_n) - f_k(x_n)) + (f_k(x_n) - f_k(x)) + (f_k(x) - f(x))| \leqslant \\ &\leqslant |f(x_n) - f_k(x_n)| + |f_k(x_n) - f_k(x)| + |f_k(x) - f(x)| \leqslant \\ &\leqslant 2\|f - f_k\| + |f_k(x_n) - f_k(x)| \end{aligned} \tag{VI.6}$$

dla dowolnych n i k. Niech $\epsilon > 0$. Ponieważ $\|f_n - f\| \to 0$, zatem dobierzmy $k \geqslant n_0$ takie, że $\|f_k - f\| < \frac{\epsilon}{3}$. Ponieważ f_k jest ciągła w x_0, zatem dobierzmy takie N, że $|f_k(x_n) - f_k(x)| < \frac{\epsilon}{3}$ dla dowolnego $n \geqslant N$. Wówczas korzystając z (VI.6) dla dowolnego $n \geqslant N$ mamy $|f(x_n) - f(x)| < \frac{2}{3}\epsilon + \frac{\epsilon}{3} = \epsilon$. $\square$

Ponieważ ciągłość jest własnością „lokalną”, zatem prostą konsekwencją powyższego twierdzenia jest podobny wynik dotyczący zbieżności niemal jednostajnej.

Wniosek. *Jeżeli funkcje f_n są ciągłe dla $n \geqslant n_0$ oraz $f_n \overset{\to}{\dashrightarrow} f$, to f jest ciągła.*

Uwaga. Wniosek ten pozwala nam m.in. przedstawić alternatywny dowód części twierdzenia o ciągłości sumy szeregu potęgowego (twierdzenie IV.14) — części dotyczącej ciągłości w otwartym przedziale zbieżności. Wystarczy bowiem skorzystać z udowodnionego niedawno faktu o niemal jednostajnej zbieżności dla takiego szeregu (strona 93).

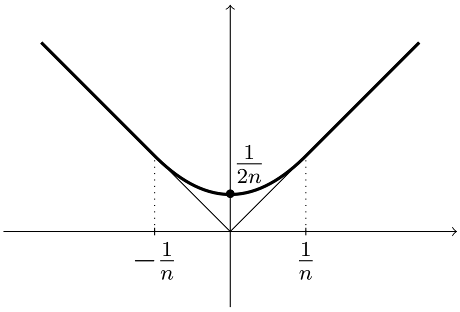
Niestety sprawa różniczkowalności funkcji okazuje się być bardziej złożona niż sprawa ciągłości. Analog twierdzenia VI.3 dla różniczkowalności nie jest bowiem prawdziwy. Łatwo się o tym przekonać konstruując odpowiednio przybliżenia funkcji $|\cdot|$ (nieróżniczkowalnej w 0) funkcjami różniczkowalnymi — proszę samodzielnie wykonać tę konstrukcję w oparciu o rysunek 14.

Sprawa różniczkowalności granicy nie jest jednak całkiem beznadziejna, można bowiem wykazać twierdzenie następujące.

Twierdzenie VI.4 (o różniczkowalności granicy). *Jeżeli $f_n, f, g \colon I \to \mathbb{R}$, gdzie I jest przedziałem oraz funkcje f_n są różniczkowalne i spełnione są warunki:*

1. $f_n \to f$,
2. $f_n' \rightrightarrows g$,

to f też jest różniczkowalna oraz $f' = g$. **B.D.**



Rysunek 14: Nieróżniczkowalną funkcję $|\cdot|$ można łatwo „przybliżyć” jednostajnie funkcjami różniczkowalnymi poprzez „zaokrąglanie kantu”...

A zatem dla różniczkowalności granicy potrzebna jest nie tyle jednostajna zbieżność samego ciągu $\{f_n\}$, co raczej ciągu pochodnych: $\{f_n'\}$.

Uwaga 1. W powyższym twierdzeniu w punkcie 2. wystarczy zakładać zbieżność niemal jednostajną. Wynika to (podobnie jak w wypadku kwestii ciągłości — patrz uwaga po twierdzeniu VI.3) z tego, że różniczkowalność jest pojęciem „lokalnym”.

Uwaga 2. Zarówno twierdzenie VI.3 jak i twierdzenie VI.4 mają swoje odpowiedniki dla szeregów funkcyjnych (proszę je sformułować samodzielnie, jako proste ćwiczenie). Związane jest to z faktem, że zarówno ciągłość jak i różniczkowalność zachowują się przy dodawaniu funkcji, a zatem odpowiednie ciągi sum częściowych będą się składać z funkcji ciągłych, ewentualnie różniczkowalnych, o ile to samo założymy o wyrazach szeregu funkcyjnego. W przypadku różniczkowania takie twierdzenie dotyczące szeregów nazywane jest twierdzeniem „o różniczkowaniu szeregu wyraz po wyrazie” i jego teza zapisywana bywa w formie (nieco nieścisłej...)

$$\left(\sum_{n=n_0}^{+\infty} f_n\right)' = \sum_{n=n_0}^{+\infty} f_n'.$$

Wniosek. *Szereg potęgowy można w otwartym przedziale zbieżności różniczkować „wyraz po wyrazie”. Tzn., jeżeli S jest sumą szeregu potęgowego $\sum_{n=0}^{+\infty} a_n(x-x_0)^n$ oraz Z_0 jest jego otwartym przedziałem zbieżności, to dla $x \in Z_0$ funkcja S jest różniczkowalna w x oraz*

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

W szczególności S jest klasy C^∞ w otwartym przedziale zbieżności.

Dowód.
Na mocy uwagi 1 wystarczy tu dowieść, że szereg potęgowy $\sum_{n=0}^{+\infty} (n+1)a_{n+1}(x-x_0)^n$ jest w Z_0 zbieżny niemal jednostajnie. A to z kolei wynika z faktu o niemal jednostajnej zbieżności szeregu potęgowego (ze str. 93) oraz z poniższego prostego lematu, który pozostawiam do dowodu Czytelnikom/Słuchaczom. □

Lemat. *Promienie zbieżności szeregów $\sum_{n=0}^{+\infty} a_n(x-x_0)^n$ i $\sum_{n=0}^{+\infty} (n+1)a_{n+1}(x-x_0)^n$ są równe.*

Uwaga. W oparciu o powyższy wniosek łatwo można udowodnić fakt ze strony 81 (mówiący o tym, że rozwinięcie w szereg potęgowy jest szeregiem Taylora dla sumy tego szeregu potęgowego).

Na zakończenie jeszcze jeden przykład zastosowania różniczkowania „wyraz po wyrazie”.

Przykład (funkcja ζ [101] Riemanna). Jak wiemy, dla dowolnego $x > 1$ szereg $\sum_{n=1}^{+\infty} \frac{1}{n^x}$ jest zbieżny. Zdefiniujemy zatem funkcję $\zeta\colon (1;+\infty) \to \mathbb{R}$ wzorem

$$\zeta(x) := \sum_{n=1}^{+\infty} \frac{1}{n^x}.$$

Nietrudno wykazać (zadanie VI.10), że funkcja ta jest n-krotnie różniczkowalna przy dowolnych n, tzn. że $\zeta \in C^\infty((1;+\infty))$.

4. Aproksymacja[102] funkcji ciągłych

W matematyce i jej zastosowaniach często zamiast danej funkcji f wygodnie jest rozważać jakieś jej przybliżenia funkcjami należącymi do pewnej określonej klasy funkcji.

Jaki to rodzaj przybliżenia i jaka klasa funkcji aproksymujących — to zależy już od konkretnej sytuacji. Możliwość znajdowania tego typu przybliżeń gwarantują różne tzw. *twierdzenia o aproksymacji*, czyli po prostu twierdzenia, które mówią, że dla funkcji f istnieje ciąg funkcyjny $\{f_n\}$ złożony z funkcji odpowiedniej klasy zbieżny w odpowiednim sensie do f. Sformułuję tu tylko jedno takie twierdzenie — dotyczące aproksymacji funkcji ciągłej wielomianami.

Twierdzenie VI.5 (Weierstrassa). *ażda funkcja ciągła na **przedziale domkniętym** jest granicą jednostajnie zbieżnego ciągu wielomianów.* **B.D.**

Twierdzenie to jest szczególnym przypadkiem dużo bardziej abstrakcyjnego twierdzenia Stone'a–Weierstrassa. Twierdzeniem podobnym do VI.5 jest twierdzenie o aproksymacji funkcji ciągłych funkcjami *kawałkami liniowymi* (ściślej: afinicznymi...) — patrz zadanie VI.13. Warto też wiedzieć o tym, że jest bardzo wiele różnych twierdzeń o aproksymacji, które dotyczą rozmaitych zbieżności (niekoniecznie spośród trzech rodzajów tu poznanych) oraz rozmaitych funkcji (niekoniecznie ciągłych). Np. dla wielu zastosowań ważne są rozmaite wyniki dotyczące aproksymacji tzw. *wielomianami trygonometrycznymi*, co jest ściśle związane z nieobecną w tym wykładzie teorią *szeregów Fouriera*.

---

[101] Czytaj „dzeta".
[102] Aproksymacja = przybliżanie.

Zadania do Rozdziału VI

$\forall$ 1. [103] Zbadaj zbieżność punktową, jednostajną, niemal jednostajną ciągów funkcyjnych $\{f_n\}$ zadanych poniższymi wzorami:

(a) x^n **(1.)** dla $x \in [0;1]$; **(2.)** dla $x \in [0;1)$;

(b) $x^n - x^{n+1}$ dla $x \in [0;1]$;

(c) $x^n - x^{2n}$ dla $x \in [0;1]$;

(d) $\frac{1}{n+x}$ dla $x \in (0;+\infty)$;

(e) $\frac{nx}{1+n+x}$ dla $x \in [0;+\infty)$;

(f) $\sin(\frac{x}{n})$ dla $x \in \mathbb{R}$;

(g) $\operatorname{arctg}(nx)$ dla $x \in \mathbb{R}$;

(h) $x \operatorname{arctg}(nx)$ dla $x \in \mathbb{R}$.

$\forall$ 2. [104] Zbadaj zbieżność punktową, jednostajną, niemal jednostajną szeregów funkcyjnych zadanych następującymi wzorami:

(a) $\sum_{n=1}^{+\infty} \frac{x}{n^2}$ dla $x \in \mathbb{R}$;

(b) $\sum_{n=1}^{+\infty} \frac{x^2}{n^4+x^4}$ dla $x \in \mathbb{R}$;

(c) $\sum_{n=1}^{+\infty} \frac{\sin(n^2x)}{n^2+x^2}$ dla $x \in \mathbb{R}$;

(d) $\sum_{n=0}^{+\infty} x^2e^{-nx}$ dla $x \in (0;+\infty)$;

(e) $\sum_{n=0}^{+\infty} xe^{-nx}$ dla $x \in (0;+\infty)$;

(f) $\sum_{n=0}^{+\infty} e^{-nx}$ dla $x \in (0;+\infty)$;

(g) $\sum_{n=1}^{+\infty} \frac{1}{n}\sin(\frac{x}{n})$ dla $x \in [-100;100]$;

(h) $\sum_{n=1}^{+\infty} \frac{(-1)^n}{n+x}$ dla $x \in [0;+\infty)$.

3. Zbadaj, które z poniższych „twierdzeń” dotyczących zbieżności jednostajnej są rzeczywiście **twierdzeniami** (tu $f, f_n, g, g_n \colon D \to \mathbb{R}$):

(a) $f_n \rightrightarrows f$ oraz $A \subset D$, to $f_n|_A \rightrightarrows f|_A$.

(b) $A \cup B = D$ oraz $f_n|_A \rightrightarrows f|_A$ i $f_n|_B \rightrightarrows f|_B$, to $f_n \rightrightarrows f$.

(c) $f_n \rightrightarrows f$ oraz $g_n \rightrightarrows g$, to $(f_n + g_n) \rightrightarrows f + g$.

(d) $f_n \rightrightarrows f$ oraz $g_n \rightrightarrows g$, to $(f_n \cdot g_n) \rightrightarrows f \cdot g$.

Zbadaj analogiczne „twierdzenia” dotyczące „$\to$” oraz „$\dashrightarrow\!\!\!\!\rightrightarrows$”.

4. Wykaż, że jeśli $f_n \rightrightarrows f$, to $\|f_n\| \to \|f\|$ (także, gdy $\|f\| = +\infty$).

5. Po uważnej lekturze odpowiednich fragmentów wykładu odgadnij, napisz i zapisz w równoważnej postaci z użyciem $\|\cdot\|$ „jednostajny warunek Cauchy'ego” dla ciągów funkcyjnych. Następnie sformułuj i udowodnij „jednostajny” odpowiednik twierdzenia II.7 (o zupełności...).

6. W oparciu o twierdzenie wykazane w zadaniu powyższym udowodnij twierdzenie o warunku dostatecznym zbieżności jednostajnej dla szeregu funkcyjnego (tw. VI.2).

[103] Przynajmniej 2 przykłady z a)–e) i jeden z f)–h).

[104] Przynajmniej 2 przykłady.

$\forall$ 7. Znajdź przykłady ciągów funkcyjnych pokazujące, że przy zbieżności punktowej ani ciągłość ani różniczkowalność nie muszą się zachowywać (przy „przejściu do granicy”).

8. Wykaż, że wypukłość funkcji zachowuje się przy zbieżności punktowej (!!).

$\forall$ 9. [105] Zbadaj ciągłość i różniczkowalność funkcji f, w przypadku różniczkowalności zbadaj znak $f'(0)$.

   (a) $f(x) = \sum_{n=1}^{+\infty} \frac{\sin(nx)}{n^3}$ dla $x \in \mathbb{R}$;
   (b) $f(x) = \sum_{n=1}^{+\infty} \operatorname{arctg}(\frac{x}{n^2})$, $x \in \mathbb{R}$;
   (c) $f(x) = \sum_{n=1}^{+\infty} (\cos \frac{x}{n} - 1)$, $x \in \mathbb{R}$.

10. Wykaż, że (patrz przykład ze strony 95):

   (a) $\zeta \in C^1((1; +\infty))$;
   (b) $\zeta \in C^\infty((1; +\infty))$.

$\forall$ 11. „Oblicz” sumy szeregów:

   (a) $\sum_{n=0}^{+\infty} \frac{n}{2^n}$;
   (b) $\sum_{n=1}^{+\infty} \frac{1}{3^n \cdot n}$;
   (c) $\sum_{n=0}^{+\infty} \frac{n^2}{7^n}$.

12. Znajdź $f^{(n)}(0)$:

   (a) $f(x) = \frac{1}{2+3x^2}$ dla $x \in \mathbb{R}$, $n = 1001$.
   (b) $f(x) = \operatorname{arctg} x$ dla $x \in \mathbb{R}$; $n = 999, n = 1000$.
   (c) $f(x) = \frac{x}{(x-2)(x-3)}$ dla $x \in (-1; 1)$; $n = 100$. Wskazówka: zapisz $f(x)$ jako $\frac{A}{x-2} + \frac{B}{x-3}$ dla pewnych $A, B \in \mathbb{R}$.

13. Funkcja $g\colon [a; b] \to \mathbb{R}$ jest *kawałkami liniowa* wtw istnieją liczby $a_0 \leqslant a_1 \leqslant \cdots \leqslant a_k$ takie, że $a_0 = a, a_k = b$ oraz $f\,|_{[a_{j-1}; a_j]}$ jest wielomianem stopnia $\leqslant 1$ dla dowolnego $j = 1, \ldots, k$. Wykaż, że każda funkcja ciągła określona na przedziale domkniętym jest granicą jednostajną ciągu funkcji kawałkami liniowych.

[105] Przynajmniej jeden przykład.

VII Rachunek całkowy

[około 3 wykładów]

1. Całka nieoznaczona

W tym podrozdziale zajmiemy się „operacją” odwrotną do różniczkowania. Niech $f : D \to \mathbb{R}$, $D \subset \mathbb{R}$. Każdą funkcję $F : D \to \mathbb{R}$ taką, że $F' = f$ nazywamy *funkcją pierwotną* funkcji f. Druga nazwa na funkcję pierwotną to *całka nieoznaczona*. Oczywiście nie każda funkcja f posiada funkcję pierwotną, np. łatwo sprawdzić, że nie posiada jej funkcja $f : \mathbb{R} \to \mathbb{R}$ zadana wzorem

$$f(x) = \begin{cases} 0 & x \leqslant 0 \\ 1 & x > 0 \end{cases}$$

(dlaczego?). Co więcej, nawet jeśli istnieje funkcja pierwotna jakiejś funkcji, to nie jest ona wyznaczona jednoznacznie. Na szczęście, dla funkcji określonych na przedziale ta niejednoznaczność nie jest „duża”.

Fakt. *Jeżeli I — przedział, $f : I \to \mathbb{R}$ oraz F jest funkcją pierwotną funkcji f, to $F_1 : I \to \mathbb{R}$ jest funkcją pierwotną f wtw $F_1 = F + C$ dla pewnej funkcji stałej C.*

Dowód.
Oczywiście $(F + C)' = F' = f$. Jeżeli $F_1' = f$, to $(F_1 - F)' = f - f = 0$ zatem $F_1 - F$ jest stała, bo I — przedział (patrz wniosek ze strony 71). □

Oczywiście, gdy dziedzina funkcji nie jest przedziałem, to ta niejednoznaczność może być „większa”, np. dla funkcji zadanej wzorem $\frac{1}{x}$ okreśonej na $\mathbb{R} \setminus \{0\}$ funkcje pierwotne to wszystkie funkcje postaci:

$$f(x) = \ln |x| + \begin{cases} c_1 & \text{dla } x < 0 \\ c_2 & \text{dla } x > 0, \end{cases}$$

gdzie $c_1, c_2 \in \mathbb{R}$. Tradycyjne oznaczenie na funkcję pierwotną, czyli całkę nieoznaczoną, to

$$\int f(x)dx.$$

To oznaczenie ma bardzo wiele wad. Np. napis $\int f(x)dx$ nie oznacza jednej funkcji, tylko całą ich klasę — nie bardzo wiadomo zatem jak się tym posługiwać. Można by to na siłę uściślić i wprowadzić rozmaite operacje — np. dodawanie — dla tego typu klas funkcji. Nie będziemy jednak tu tego robić i zgodnie z tradycją będziemy raczej traktować całkę nieoznaczoną „prawie tak jak funkcję” pamiętając, że to tak naprawdę cała ich klasa. Pomimo wad, ta notacja posiada też nieco zalet, o których wspomnimy później.

Pojawia się naturalne pytanie:

„dla jakich funkcji całka nieoznaczona w ogóle istnieje?”

W następnym podrozdziale wykażemy, że odpowiedź jest pozytywna np. dla wszystkich funkcji ciągłych określonych na przedziale. To, jak się wydaje, już całkiem nieźle. Jednak z praktycznego — rachunkowego — punktu widzenia ważne jest inne pytanie:

„jak obliczyć całkę nieoznaczoną z funkcji zadanej elementarnym wzorem?”

Czy można zrobić to tak samo łatwo, jak w przypadku różniczkowania? Odpowiedź brzmi: **NIE**! Przypomnijmy, że w przypadku różniczkowania mieliśmy — po pierwsze — wzory na pochodną podstawowych funkcji elementarnych — a po drugie — wzory rachunkowe na pochodną sumy, złożenia i iloczynu. I to właśnie gwarantowało nam możliwość praktycznego różniczkowania dowolnie skomplikowanych funkcji elementarnych. Co zatem mamy do dyspozycji w przypadku całkowania? Właściwie tylko to, co da się wywnioskować z wyżej wspomnianych wzorów dotyczących różniczkowania, niejako poprzez ich „odwrócenie". A zatem np. konsekwencją wzorów na pochodną podstawowych funkcji elementarnych oraz faktu ze strony 99 są umieszczone w poniższej tabeli wzory na całki. Zawiera ona wzory opisujące funkcję i (obok) ogólną postać jej funkcji pierwotnej (n, k oznaczają tu dowolną liczbę całkowitą, natomiast C, C', C_k — dowolną liczbę rzeczywistą („stałą")).

$f(x)$	$\int f(x)dx$
x^α; $(x>0, \alpha\neq -1)$ lub $(x\in\mathbb{R}, \alpha\in\mathbb{N})$	$\frac{1}{\alpha+1}x^{\alpha+1}+C$
x^α; $x\neq 0$, $\alpha\in\mathbb{Z}_-$	$\begin{cases}\frac{1}{\alpha+1}x^{\alpha+1} & \text{gdy } \alpha\neq -1\\ \ln\lvert x\rvert & \text{gdy } \alpha=-1\end{cases} + \begin{cases} C_1 & \text{dla } x<0\\ C_2 & \text{dla } x>0\end{cases}$
e^x; $x\in\mathbb{R}$	e^x+C
a^x; $x\in\mathbb{R}$, $a>0$	$\frac{a^x}{\ln a}+C$
$\sin x$; $x\in\mathbb{R}$	$-\cos x+C$
$\cos x$; $x\in\mathbb{R}$	$\sin x+C$
$\operatorname{tg}^2 x+1=\frac{1}{\cos^2 x}$; $x\neq n\cdot\pi+\frac{\pi}{2}$	$\operatorname{tg} x+C_k$ dla $x\in(k\cdot\pi-\frac{\pi}{2}; k\cdot\pi+\frac{\pi}{2})$
$\operatorname{ctg}^2 x+1=\frac{1}{\sin^2 x}$; $x\neq n\cdot\pi$	$-\operatorname{ctg} x+C_k$ dla $x\in(k\cdot\pi;(k+1)\cdot\pi)$
$\frac{1}{\sqrt{1-x^2}}$; $x\in(-1;1)$	$\arcsin x+C$ oraz $-\arccos x+C'$
$\frac{1}{1+x^2}$; $x\in\mathbb{R}$	$\operatorname{arctg} x+C$ oraz $-\operatorname{arcctg} x+C'$

Niech teraz F, G będą odpowiednio funkcjami pierwotnymi funkcji f i g. Jako konsekwencję wzoru na pochodną sumy otrzymujemy oczywiście, że $F+G$ jest funkcją pierwotną $f+g$. Podobnie, gdy $a\in\mathbb{R}$, to $a\cdot F$ jest funkcją pierwotną $a\cdot f$. W zapisie tradycyjnym fakty te przedstawia się tak:

$$\int (f+g)(x)dx=\int f(x)dx+\int g(x)dx, \quad \int (a\cdot f)(x)dx=a\int f(x)dx.$$

Podkreślam jednak znów, że powyższe wzory wymagają uściśleń (a ich całkiem ścisła wersja, to właśnie zdanie je poprzedzające). Z kolei natychmiastową konsekwencją wzoru na pochodną iloczynu dwóch funkcji jest fakt poniższy.

Fakt VII.1 (o całkowaniu „przez części"). *Jeśli f oraz g są różniczkowalne oraz H jest funkcją pierwotną funkcji $f'\cdot g$, to $f\cdot g-H$ jest funkcją pierwotną $f\cdot g'$.*

Tradycyjny — nieformalny zapis tego faktu ma postać:

$$\int f(x)\cdot g'(x)dx=f(x)\cdot g(x)-\int f'(x)\cdot g(x)dx.$$

Wreszcie, „odwrócenie" [106] twierdzenia o pochodnej złożenia (twierdzenie V.1 punkt b)) ma postać następującą:

Fakt VII.2 (o całkowaniu „przez podstawienie"). *Jeżeli g jest różniczkowalna i F jest funkcją pierwotną f, to $F\circ g$ jest funkcją pierwotną do $(f\circ g)\cdot g'$ [107].*

W zapisie tradycyjnym można by to przedstawić tak:

$$\int f(g(x))\cdot g'(x)dx=\int f(y)dy, \quad \text{gdzie } y=g(x). \qquad \text{(VII.1)}$$

[106] To inne „odwrócenie" niż wtedy, gdy mowa o *twierdzeniu odwrotnym* (czyli związanym z implikacją w stronę przeciwną niż w danym twierdzeniu).
[107] Oczywiście zakładamy, że złożenie $f\circ g$ (a zatem automatycznie też $F\circ g$) jest określone.

I właśnie przy tym wzorze objawia się główna zaleta owego dziwacznego tradycyjnego oznaczenia funkcji pierwotnej, a szczególnie jego tajemniczego zakończenia „dx". Ma to bezpośredni związek z innym tradycyjnym oznaczeniem — na pochodną. Wspominaliśmy już o zapisie $g' = \frac{dg}{dx}$. Można pójść jeszcze dalej — skoro $y = g(x)$, to zapiszmy nieformalnie

$$g'(x) = \frac{dy}{dx}$$

i potraktujmy powyższy zapis tak jak ułamek. Wówczas pod „$\int$" z lewej strony wzoru ((VII.1)) uzyskamy: „$f(y)\frac{dy}{dx}dx = f(y)dy$", czyli wyrażenie znajdujące się pod „$\int$" ze strony prawej. Ta mocno podejrzana manipulacja prowadzi na szczęście do całkiem ścisłego i prawdziwego wyniku opisanego w sformułowanym przed chwilą fakcie VII.2. Tak więc przy praktycznych rachunkach można posługiwać się tego typu „skracaniem dx-ów", pod warunkiem jednak, że **zachowuje się pełną świadomość** jak to skracanie zastąpić ścisłą argumentacją z faktu o całkowaniu przez podstawienie.

Nauką praktycznego posługiwania się poznanymi tu wzorami zajmiecie się Państwo na ćwiczeniach. Teraz powrócimy natomiast do pytania o praktyczne całkowanie funkcji elementarnych. To, czego nam brakuje najbardziej dotkliwie, to chyba wzory na całkę iloczynu i na całkę złożenia. Zamiast tego mamy pewne szczególne namiastki. Wzór na całkowanie przez części pozwala nam na scałkowanie tylko iloczynu postaci $f \cdot g'$ i to tylko wtedy, gdy umieliśmy scałkować $f' \cdot g$. Z kolei wzór na całkowanie przez podstawienie umożliwia scałkowanie nie samego złożenia $f \circ g$ ale tylko funkcji $(f \circ g) \cdot g'$ (ale za to nie musimy umieć całkować g — wystarczy, że umiemy zrobić to dla f). Wszystko to sprawia, że całkowanie jest w praktyce znacznie trudniejsze niż różniczkowanie. Ale też **znacznie ciekawsze**! To trochę jak rozwiązywanie łamigłówek ... No i tak jak należało się spodziewać, czasem — nawet dla dość „prostych" funkcji, praktyczne scałkowanie poprzez zapisanie całki jako funkcji elementarnej bywa po prostu **niewykonalne**... Jeden z bardziej znanych przykładów to całka

$$\int e^{(x^2)}dx \quad ^{108)} .$$

Nie da się więc scałkować wszystkiego, co chcieliśmy. Ale coś jednak scałkować się da. Przez kilkaset (około 200) lat wymyślono wiele metod radzenia sobie z różnymi typami całek. Jeden z najważniejszych takich typów to całki z funkcji wymiernych. Możliwość ich wyliczenia jest ważna nie tylko „sama dla siebie". Wiele innych typów całek można sprowadzić właśnie do całek z funkcji wymiernych.

Rozważmy więc funkcję wymierną $f : D \to \mathbb{R}$, $f(x) = \frac{w(x)}{v(x)}$, gdzie w i v są wielomianami, $\deg v \geqslant 1$ oraz $D = \mathbb{R} \setminus D_0$, gdzie D_0 jest zbiorem (skończonym) wszystkich pierwiastków rzeczywistych wielomianu v. Aby wyliczyć $\int f(x)dx$ postępujemy następująco.

Etap 1. (dzielenie z resztą) Zapisujemy $f(x)$ jako $u(x) + \frac{r(x)}{v(x)}$, gdzie $w(x) = u(x) \cdot v(x) + r(x)$, u, r — wielomiany i $\deg r < \deg v$. Ponieważ wyliczenie $\int u(x)dx$ jest proste (patrz tabela) zatem dalej wystarczy zająć się $\int \frac{r(x)}{v(x)}dx$.

Etap 2. (rozkład na ułamki proste) Wielomian v można (jak wiadomo z algebry) rozłożyć na iloczyn tzw. *wielomianów nierozkładalnych* (stopnia 1 lub 2), tzn.

$$v(x) = \alpha \cdot (x - x_1)^{k_1} \cdot \dots \cdot (x - x_s)^{k_s} \cdot (p_1(x))^{l_1} \cdot \dots \cdot (p_t(x))^{l_t}, \tag{VII.2}$$

gdzie $\alpha \in \mathbb{R}$, $x_1, \dots, x_s$ — różne parami pierwiastki wielomianu v, $p_1, \dots, p_t$ — różne parami wielomiany 2-go stopnia postaci

$$p_j(x) = (x - y_j)^2 + z_j^2, \quad j = 1, \dots, t,$$

108) Dowód, że $\int e^{(x^2)}dx$ nie jest „funkcją" elementarną to rzecz zupełnie **nie** trywialna ...

gdzie $y_j, z_j \in \mathbb{R}$ i $z_j \neq 0$ oraz $k_1, \ldots, k_s, l_1, \ldots, l_t \in \mathbb{N}$. [109)]

Znany algebraiczny fakt mówi, że w tej sytuacji funkcja wymierna zadana wzorem $\frac{r(x)}{v(x)}$ (dla $\deg r < \deg v$) jest sumą pewnej liczby *ułamków prostych*, tzn. funkcji wymiernych, z których każda opisana jest wzorem postaci

$$\frac{A}{(x-x_j)^k} \quad \text{lub} \quad \frac{Ax+B}{(p_i(x))^l}$$

gdzie $A, B \in \mathbb{R}$, $j = 1, \ldots, s$, $i = 1, \ldots, t$ oraz $k, l \in \mathbb{N}$, $k \leqslant k_j$ natomiast $l \leqslant l_j$. W praktyce znalezienie rozkładu na taką sumę sprowadza się łatwo do rozwiązania pewnego układu równań („liniowych").

Etap 3. (całkowanie ułamków prostych) Pozostaje więc scałkować każdy z ułamków prostych. W pierwszym wystarczy po „podstawieniu $y = x - x_j$" zajrzeć do tabeli całek (funkcja potęgowa z wykładnikiem $-k$). Z drugim jest troszkę trudniej. Najpierw zauważmy, że

$$\frac{Ax+B}{(p_i(x))^l} = \frac{\frac{A}{2} \cdot p_i'(x)}{(p_i(x))^l} + \frac{\tilde{B}}{(p_i(x))^l}$$

dla pewnego $\tilde{B} \in \mathbb{R}$. Pierwszy z tych składników łatwo scałkować przez „podstawienie $y = p_i(x)$". Drugi przez pewne podstawienie „afiniczne" (tzn. „$y = ax + b$", ale z jakimi a, b?) łatwo sprowadzić do całki z funkcji zadanej wzorem

$$\frac{1}{(x^2+1)^l},$$

na którą można znaleźć wzór rekurencyjny po l (zadanie VII.3), a dla $l = 1$ całkowanie daje funkcję arctg (patrz tabela).

Przykład. Znajdź $\displaystyle\int \frac{x^5 + 2x^4 + 4x^3 + 7x^2 + 3x + 2}{x^4 + 2x^3 + 3x^2 + 4x + 2} dx$ (na dziedzinie $D = \mathbb{R} \setminus D_0$, gdzie D_0 — zbiór pierwiastków rzeczywistych wielomianu z mianownika). Niech f oznacza powyższą funkcję podcałkową. Mamy dla $x \in D$

$$f(x) = x + \frac{x^3 + 3x^2 + x + 2}{(x^3 + x^2 + 2x + 2)(x+1)} = x + \frac{x^3 + 3x^2 + x + 2}{(x^2+2)(x+1)^2}.$$

A zatem możliwe są następujące postaci ułamków prostych w rozkładzie drugiego składnika:

$$\frac{Ax+B}{x^2+2}, \; \frac{C}{(x+1)^2}, \; \frac{C'}{x+1}$$

Zachodzi zatem $D_0 = \{-1\}$ oraz dla $x \in D$

$$\begin{aligned}\frac{x^3 + 3x^2 + x + 2}{(x^2+2)(x+1)^2} &= \frac{Ax+B}{x^2+2} + \frac{C}{(x+1)^2} + \frac{C'}{x+1} = \\ &= \frac{(Ax+B)(x+1)^2 + C(x^2+2) + C'(x^2+2)(x+1)}{(x^2+2)(x+1)^2},\end{aligned}$$

a stąd dla $x \in D$ licznik prawej i lewej strony muszą być równe, zatem równe muszą być kolejne współczynniki przy x^3, x^2, x^1, x^0, tzn.

$$\begin{aligned}1 &= A + C' \\ 3 &= B + 2A + C + C' \\ 1 &= A + 2B + 2C' \\ 2 &= B + 2C + 2C'.\end{aligned}$$

[109)] Przy czym oczywiście w rozkładzie (VII.2) część z iloczynem $(x - x_j)^{k_j}$ lub część z iloczynem $(p_j^{(x)})^{l_j}$ może się okazać „pusta".

Otrzymaliśmy więc układ czterech równań liniowych z czterema niewiadomymi, który łatwo rozwiązujemy i otrzymujemy rozwiązanie:

$$A = 1,\ B = 0,\ C = 1,\ C' = 0$$

Stąd ostatecznie

$$\int f(x)dx = \int xdx + \int \frac{x}{x^2+2}dx + \int \frac{1}{(x+1)^2}dx = \frac{1}{2}x^2 + \frac{1}{2}\int \frac{2x}{x^2+2}dx + \int \frac{1}{(x+1)^2}dx =$$
$$= \frac{1}{2}x^2 + \frac{1}{2}\ln(x^2+2) - \frac{1}{x+1} + \begin{cases} c_1 \text{ dla } x < -1 \\ c_2 \text{ dla } x > -1 \end{cases}$$

gdzie c_1 i c_2 są dowolnie dobranymi stałymi.

Podsumowując, jesteśmy w stanie wyliczyć całkę z każdej funkcji wymiernej, dla której potrafimy znaleźć explicite rozkład postaci (VII.2) dla jej mianownika.

Jednym z typów całek, które sprowadzają się do całkowania funkcji wymiernych są całki postaci

$$\int W(\sin x, \cos x)dx,$$

gdzie W jest ilorazem dwóch wielomianów dwóch zmiennych[110)] . Wówczas dla ustalonego $n \in \mathbb{Z}$ dla $x \in ((2n-1)\pi; (2n+1)\pi)$ można użyć podstawienia

$$„t = \mathrm{tg}(\frac{x}{2})".$$

Nietrudno ze wzorów trygonometrycznych wyliczyć, że wówczas

$$\cos x = \frac{1-t^2}{1+t^2},\ \sin x = \frac{2t}{1+t^2}.$$

Ponadto

$$\frac{dt}{dx} = \frac{1+t^2}{2}$$

(t traktujemy tu jako „funkcję zmiennej x"). W efekcie użycie całkowania przez podstawienie sprowadzi więc problem do obliczenia pewnej całki z funkcji wymiernej („zmiennej t"). Zachęcam zarówno do sprawdzenia podanych wyżej wzorów (zad. VII.4) oraz do szczegółowego prześledzenia tego, jak należy tu użyć **ścisłego** faktu o całkowaniu przez podstawienie.

Inny przykład zastosowania całek z funkcji wymiernej to całki zawierające tzw. *niewymierności stopnia drugiego*, tzn. wyrażenia postaci $\sqrt{ax^2+bx+c}$. Warto wiedzieć, że istnieją różne podstawienia, w tym tzw. *podstawienia Eulera*, które mogą sprowadzić takie całki także do całek z pewnych funkcji wymiernych.

Na zakończenie tego podrozdziału zajmiemy się (wbrew jego tytułowi) całką **oznaczoną**. Nazwa „nieoznaczona" dla rozważanej tu dotąd całki związana była oczywiście z niejednoznacznością wyboru funkcji pierwotnej. W przypadku jednak gdy funkcja f jest określona na przedziale, niejednoznaczność ta — jak widzieliśmy (fakt VII.1, strona 100) — jest niewielka. Jeśli zatem $f : I \to \mathbb{R}$, gdzie I — przedział, posiada funkcję pierwotną F, oraz $a, b \in I$, to liczbę $F(b) - F(a)$ nazywamy *całką oznaczoną* od a do b z funkcji f i zapisujemy ją symbolem

$$\int_a^b f(x)dx.$$

Istotne jest to, że ta definicja jest poprawna — w tym sensie, że rzeczywiście liczba powyższa zależy jedynie od f, a oraz b natomiast **nie** zależy od wyboru samej funkcji pierwotnej. Dodanie bowiem ew. stałej do funkcji F nie wpływa na wartość obliczanej różnicy.

110) Wielomian dwóch zmiennych to suma skończonej liczby funkcji zadanych wzorami postaci Cx^ky^l, gdzie $C \in \mathbb{R}$, $k, l \in \mathbb{N}_0$ (x, y — oznaczają zmienne).

Uwaga. Dla dowolnego $a \in I$ funkcja $\varphi : I \to \mathbb{R}$ zadana wzorem $\varphi(x) = \int_a^x f(s)ds$ jest zatem funkcją pierwotną f. Ponadto $\varphi(a) = 0$.

Czasami jest wygodniej posługiwać się taką konkretną „zaczepioną w punkcie a" funkcją pierwotną f niż bliżej nie sprecyzowaną całką nieoznaczoną z f.

Wzory na całkowanie przez części i podstawienie, które zapisane przy użyciu symbolu całki nieoznaczonej były niezbyt ścisłe, mają swoje odpowiedniki — tym razem ścisłe „w 100%" — także dla całek oznaczonych. Bezpośrednio z definicji całki oznaczonej, z faktów VII.1 i VII.2 ze strony 100 otrzymujemy następujące wzory (każdy z nich obowiązuje przy założeniach odpowiedniego faktu oraz dodatkowo f i g muszą być określone na przedziale do którego należą też a i b):

Wzór na całkowanie przez części dla całki oznaczonej:

$$\int_a^b f(x) \cdot g'(x)dx = [f(x) \cdot g(x)]_a^b - \int_a^b f'(x) \cdot g(x)dx, \qquad \text{(VII.3)}$$

gdzie symbol $[h(x)]_a^b$ definiujemy jako $h(b) - h(a)$.

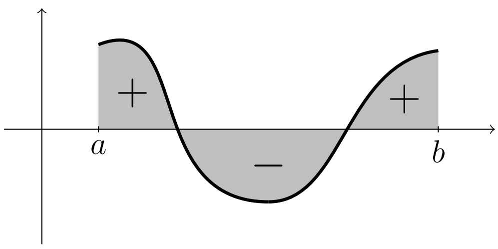
Wzór na całkowanie przez podstawienie dla całki oznaczonej:

$$\int_a^b f(g(x)) \cdot g'(x)dx = \int_{g(a)}^{g(b)} f(x)dx \text{ [111]} \qquad \text{(VII.4)}$$

Jak się już wkrótce okaże — całka oznaczona jest nie tylko wygodna, ale ma też bardzo ważną interpretację geometryczną.

2. Całka Riemanna

Zajmiemy się tu pojęciem całki zupełnie innej (przynajmniej na poziomie definicji) niż całka oznaczona i nieoznaczona zdefiniowane w poprzednim podrozdziale. Naszym celem będzie określenie dla funkcji $f : [a; b] \to \mathbb{R}$ takiej liczby, której wartość w przypadku funkcji nieujemnej można by interpretować jako pole powierzchni obszaru pomiędzy wykresem f a osią „X", a w przypadku ogólnym pole to byłoby liczone z uwzględnieniem znaku „$-$" dla tych fragmentów wykresu, które są poniżej osi „X" (patrz rys. 15).



Rysunek 15: Całka Riemanna to pole między wykresem a osią X „z uwzględnieniem znaku".

Gdyby z góry założyć np. ciągłość f, sprawa byłaby dość łatwa. My jednak będziemy nieco ambitniejsi — spróbujemy podać odpowiednią definicję, która mogłaby mieć szersze zastosowanie.

Potrzebne nam będzie zatem parę pomocniczych definicji i oznaczeń. *Podziałem* przedziału $[a; b]$ nazwiemy dowolny ciąg skończony $(x_0, \ldots, x_m)$ taki, że $x_0 = a; x_m = b$ oraz $x_{j-1} \leqslant x_j$ dla

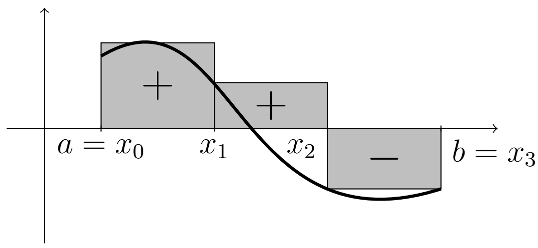
[111] Może niektórych dziwi lub wręcz bulwersuje użycie po prawej stronie wzoru zmiennej x, podczas gdy w wersji „nieoznaczonej" było tam y, gdzie „$y = g(x)$" (a zatem na ogół „$y \neq x$"...), jednak tu dla całki oznaczonej ten problem już nie istnieje. Możemy użyć jako zmienną x, y, s, t i inne litery. To nie ma **żadnego** wpływu na wartość liczby, którą w efekcie otrzymujemy z prawej strony wzoru.

$j = 1, \ldots, m$. Takie podziały będziemy oznaczać jedną literą, np. P, a zbiór wszystkich możliwych podziałów P przedziału $[a; b]$ oznaczmy przez $\mathcal{P}$. Dla funkcji ograniczonej $f : [a; b] \to \mathbb{R}$ oraz dla $P = (x_1, \ldots, x_m) \in \mathcal{P}$ definiujemy *sumę górną* i *sumę dolną* dla f i P odpowiednio wzorami:

$$\hat{S}(f, P) := \sum_{j=1}^{m} (x_j - x_{j-1}) \cdot \sup_{t \in [x_{j-1}; x_j]} f(t);$$

$$\check{S}(f, P) := \sum_{j=1}^{m} (x_j - x_{j-1}) \cdot \inf_{t \in [x_{j-1}; x_j]} f(t).$$

Dzięki ograniczoności f obie sumy są poprawnie zdefiniowanymi liczbami rzeczywistymi i mają sens geometryczny najlepszego przybliżenia szukanego pola „od góry” lub odpowiednio „od dołu” przez sumę pól prostokątów (z uwzględnieniem znaku) o podstawach wyznaczonych przez podział P (patrz rysunek 16).



Rysunek 16: Suma górna.

Nietrudno zauważyć, że biorąc „drobniejszy podział” (tj. dokładając dodatkowe punkty do danego podziału) ewentualnie możemy zmniejszyć sumę górną, a sumę dolną zwiększyć (lub pozostaną one niezmienione). Wydaje się więc, że sensownie byłoby określić szukane przez nas pole jako

$$\inf_{P \in \mathcal{P}} \hat{S}(f, P), \tag{VII.5}$$

jednak czemu nie wziąć „równie dobrej” liczby

$$\sup_{P \in \mathcal{P}} \check{S}(f, P)? \tag{VII.6}$$

Ponieważ obie metody wydają się dobre, ale nie wiemy, czy prowadzą do tego samego wyniku zatem postąpimy ostrożnie: liczbę określoną wzorem (VII.5) nazwijmy *całką górną*, a wzorem (VII.6) — *całką dolną* z funkcji f. Oznaczmy je odpowiednio symbolami

$$\hat{\int}_{[a;b]} f, \quad \check{\int}_{[a;b]} f.$$

Z pewnością zachodzi nierówność

$$\check{\int}_{[a;b]} f \leqslant \hat{\int}_{[a;b]} f \tag{VII.7}$$

Jeśli bowiem rozważymy dowolne podziały P_1 i P_2 przedziału $[a; b]$ to biorąc podział $\tilde{P}$ powstały przez „połączenie” P_1 i P_2 (ścisłą definicję tego „połączenia” pozostawiam Państwu ...), a więc podział „drobniejszy” niż P_1 i P_2, dostajemy

$$\check{S}(f, P_1) \leqslant \check{S}(f, \tilde{P}) \leqslant \hat{S}(f, \tilde{P}) \leqslant \hat{S}(f, P_2)$$

skąd (VII.7) wynika łatwo z definicji kresów (patrz np. zad. I.9). Nie ma jednak powodu, by w (VII.7) zachodziła równość bez jakiś dodatkowych założeń o f.

Zatem, dla dowolnej ograniczonej funkcji $f : [a;b] \to \mathbb{R}$ przyjmujemy następującą definicję.

Definicja. *Funkcja f jest* ***całkowalna w sensie Riemanna*** *wtw $\hat{\int}_{[a;b]} f = \check{\int}_{[a;b]} f$. Jeśli f jest całkowalna w sensie Riemanna, to wspólną wartość jej całki górnej i dolnej nazywamy* ***całką Riemanna****unkcji f (na $[a;b]$) i oznaczamy symbolem $\int_{[a;b]} f$ lub $\int_{[a;b]} f(x)dx$.*

Klasę wszystkich funkcji całkowalnych w sensie Riemanna będziemy tu oznaczać przez $\Re$, a gdy będzie nam zależało na podkreśleniu, że chodzi o funkcję określoną na $[a;b]$, będziemy używać symbolu $\Re([a;b])$.

Zanim zajmiemy się ogólnymi wynikami dotyczącymi całkowalności i całki przyjrzyjmy się następującym — skrajnie różnym z punktu widzenia tej teorii — sytuacjom.

Przykłady.

1. Funkcja stała: $f \equiv c$, $c \in \mathbb{R}$. Wówczas niezależnie od wyboru P zachodzi $\hat{S}(f,P) = c \cdot (b-a) = \check{S}(f,P)$, więc całka górna i dolna równe są $c \cdot (b-a)$. Zatem $f \in \Re$ i $\int_{[a;b]} f(x)dx = c \cdot (b-a)$.
2. Funkcja Dirichleta. Gdy f jest obcięciem funkcji Dirichleta do przedziału $[a;b]$ (patrz przykład ze strony 49), to dla dowolnego P mamy $\hat{S}(f,P) = (b-a)$ oraz $\check{S}(f,P) = 0$. Zatem $f \notin \Re$, o ile $b > a$.

Czas wreszcie na jakieś „pozytywne” twierdzenie o całkowalności w sensie Riemanna.

Twierdzenie VII.1 (o całkowalności funkcji ciągłych). *Funkcja ciągła określona na przedziale domkniętym jest całkowalna w sensie Riemanna (tzn. $C([a;b]) \subset \Re([a;b])$).*

Zanim przystąpimy do dowodu, wykażemy pomocny lemat. Najpierw dla dowolnego podziału $P = (x_0, \ldots, x_m)$ przedziału $[a;b]$ zdefiniujmy *średnicę podziału* P, oznaczaną przez $|P|$, jako długość najdłuższego odcinka z podziału P, tzn.

$$|P| := \max_{j=1,\ldots,m} (x_j - x_{j-1}).$$

Lemat. *Jeżeli $f : [a;b] \to \mathbb{R}$ jest ciągła oraz $\{P_n\}$ jest ciągiem przedziałów $[a;b]$, dla którego $|P_n| \to 0$, to $\hat{S}(f,P_n) - \check{S}(f,P_n) \to 0$*

Dowód (lematu).
Niech $b > a$ i $\epsilon > 0$. Na mocy jednostajnej ciągłości f (patrz tw. IV.11) wybierzmy $\delta > 0$ taką, że jeżeli $y, z \in [a;b]$ oraz $|y - z| < \delta$, to

$$|f(y) - f(z)| < \frac{\epsilon}{b-a}. \tag{VII.8}$$

Niech N będzie takie, że dla $n \geqslant N$ zachodzi $|P_n| < \delta$. Ustalmy dowolne $n \geqslant N$. W szczególności więc (VII.8) zachodzi dla dowolnych y i z leżących w przedziale pomiędzy sąsiednimi punktami podziału P_n. Jeżeli $P_n = (x_0, \ldots, x_m)$ oraz $j \in \{1, \ldots, m\}$, to z twierdzenia Weierstrassa o osiąganiu kresów (tw. IV.10) $\sup_{x \in [x_{j-1};x_j]} f(x) = f(y_j)$ oraz $\inf_{x \in [x_{j-i};x_j]} f(x) = f(z_j)$ dla pewnych $y_j, z_j \in [x_{j-1}, x_j]$, przy czym ponieważ $n \geqslant N$, zatem na mocy powyższych rozważań

$$f(y_j) - f(z_j) < \frac{\epsilon}{b-a}.$$

A zatem mamy

$$0 \leqslant \hat{S}(f,P_n) - \check{S}(f,P_n) = \sum_{j=1}^{m} (f(y_j) - f(z_j))(x_j - x_{j-1}) < \frac{\epsilon}{b-a} \sum_{j=1}^{m} (x_j - x_{j-1}) = \frac{\epsilon}{b-a} \cdot (b-a) = \epsilon.$$

□

Dowód (twierdzenia VII.1).
Rozważmy jakikolwiek ciąg podziałów $\{P_n\}$ taki, że $|P_n| \to 0$ (np. P_n może być podziałem na równe części o długości $\frac{b-a}{n}$). Na mocy definicji całki górnej i dolnej oraz na mocy (VII.7) mamy oczywiście

$$\check{S}(f,P_n) \leqslant \check{\int}_{[a;b]} f \leqslant \hat{\int}_{[a;b]} f \leqslant \hat{S}(f,P_n) \tag{VII.9}$$

dla dowolnego n. A stąd

$$0 \leqslant \hat{\int}_{[a;b]} f - \check{\int}_{[a;b]} f \leqslant \hat{S}(f,P_n) - \check{S}(f,P_n),$$

więc na mocy lematu oraz twierdzenia o trzech ciągach $\hat{\int}_{[a;b]} f = \check{\int}_{[a;b]} f$ □

Twierdzenie odwrotne do twierdzenia VII.1 nie jest prawdziwe, tzn. nie każda funkcja całkowalna w sensie Riemanna musi być ciągła (patrz np. zad. VII.5). Jednak można wykazać twierdzenie, które mówi, że $f \in \Re$ wtw f jest ograniczona i jej zbiór punktów nieciągłości jest „mały” [112] (przypominam także, że całkowalność w sensie Riemanna dotyczy wyłącznie funkcji określonych na przedziałach domkniętych). Przy czym „mały” jest np. dowolny zbiór skończony, ale także dowolny zbiór przeliczalny.

Udowodnimy teraz przydatne twierdzenie dotyczące aproksymowania całki z funkcji ciągłej tzw. „sumami Riemanna”. Jeżeli $P = (x_0, \ldots, x_m)$ jest podziałem przedziału $[a;b]$, to *sumą Riemanna* dla f i P nazywamy dowolną liczbę, którą można przedstawić w postaci

$$\sum_{j=1}^{m} f(y_j)(x_j - x_{j-1})$$

dla pewnych $y_j \in [x_{j-1}; x_j]$, $j = 1, \ldots, m$.

Twierdzenie VII.2 (o sumach Riemanna). *Jeżeli $f : [a;b] \to \mathbb{R}$ jest ciągła, $\{P_n\}$ jest ciągiem podziałów $[a;b]$ o średnicach zbieżnych do 0 oraz dla dowolnego n S_n jest pewną sumą Riemanna dla f i P_n, to $S_n \to \int_{[a;b]} f$.*

Dowód.
Na mocy definicji sumy górnej i dolnej dla f i P_n mamy oczywiście

$$\check{S}(f,P_n) \leqslant S_n \leqslant \hat{S}(f,P_n). \tag{VII.10}$$

Jednocześnie na mocy całkowalności f (z twierdzenia VII.1) i z definicji całki oraz całki górnej i dolnej (patrz np. (VII.9)) mamy też

$$\check{S}(f,P_n) \leqslant \int_{[a;b]} f \leqslant \hat{S}(f,P_n) \tag{VII.11}$$

Z (VII.10) i (VII.11) otrzymujemy $|\int_{[a;b]} f - S_n| \leqslant \hat{S}(f,P_n) - \check{S}(f,P_n)$, skąd na mocy lematu i twierdzenia o trzech ciągach uzyskujemy tezę. □

Uwaga. Twierdzenie powyższe pozostanie prawdziwe, jeśli ciągłość f zastąpimy jej całkowalnością w sensie Riemanna (dowód jednak będzie trudniejszy ...).

Warto zwrócić uwagę, że twierdzenie VII.2 daje możliwość znajdowania przybliżonych wartości całek. Niestety, bez żadnych „gwarancji” dotyczących wielkości błędu. Warto też wiedzieć, że istnieją twierdzenia, które podają oszacowania błędu przy przybliżaniu sumami Riemanna lub innymi sumami podobnego rodzaju, przy dodatkowych założeniach dotyczących np. regularności funkcji.

[112] Ściślej, „mały” to zbiór o mierze Lebesgue'a równej 0, ale o tym pojęciu powiemy dopiero w rozdziale X.

Całka Riemanna posiada kilka intuicyjnie naturalnych własności, które zostały zebrane poniżej. Przyjmijmy jeszcze dla wygody następującą notację: jeżeli $f : D \to \mathbb{R}$ oraz $[a;b] \subset D$ i $f_{|[a;b]} \in \Re$, to

$$\int_{[a;b]} f := \int_{[a;b]} f\,|_{[a;b]}\ .$$

Twierdzenie VII.3 (o własnościach całki Riemanna).

1. *(liniowość) Jeżeli $f, g \in \Re([a;b])$, i $c \in \mathbb{R}$, to $c \cdot f$, $f + g \in \Re([a;b])$ oraz*

$$\int_{[a;b]} (f+g) = \int_{[a;b]} f + \int_{[a;b]} g,$$

$$\int_{[a;b]} (c \cdot f) = c \cdot \int_{[a;b]} f.$$

2. *(monotoniczność) Jeżeli $f, g \in \Re([a;b])$ oraz $\forall_{x \in [a;b]}\ f(x) \leqslant g(x)$, to*

$$\int_{[a;b]} f \leqslant \int_{[a;b]} g.$$

3. *(addytywność względem przedziału) Jeżeli $f \in \Re([a;b])$ oraz $a \leqslant c \leqslant b$, to $f\,|_{[a;c]}, f\,|_{[c;b]} \in \Re$ oraz*

$$\int_{[a;c]} f + \int_{[c;b]} f = \int_{[a;b]} f.$$

Punkt 2 wynika natychmiast z definicji całki górnej (lub dolnej) oraz z własności kresów. Pozostałe punkty są może nie tyle trudne, ale żmudne w dowodzie i dlatego, z konieczności, ich dowód pomijamy.

Twierdzenie zwane „podstawowym twierdzeniem rachunku całkowego" (p.t.r.c.), które teraz sformułujemy, ma kluczowe znaczenie dla teorii całki Riemanna. Wiąże ono bowiem całkę Riemanna z całką oznaczoną, a jednocześnie, dzięki temu związkowi, daje praktyczną możliwość obliczania wielu całek Riemanna, bez konieczności posługiwania się definicją (czy ewentualnie twierdzeniem VII.2).

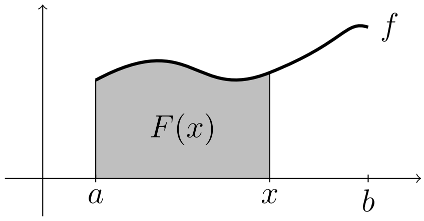
Twierdzenie VII.4 (p.t.r.c.). *Niech $f : [a;b] \to \mathbb{R}$ będzie funkcją ciągłą i zdefiniujmy $F : [a;b] \to \mathbb{R}$ wzorem*

$$F(x) = \int_{[a;x]} f(t)dt$$

dla $x \in [a;b]$. Wówczas F jest funkcją pierwotną funkcji f.

Wniosek. *Jeżeli $f : [a;b] \to \mathbb{R}$ — ciągła, to $\int_{[a;b]} f(x)dx = \int_a^b f(x)dx$.*

Dowód (wniosku).
Z twierdzenie VII.4 i z definicji całki oznaczonej mamy $\int_a^b f(x)dx = F(b) - F(a) = \int_{[a;b]} f - \int_{[a;a]} f = \int_{[a;b]} f(x)dx$. □



Rysunek 17: Sens geometryczny liczby $F(x)$ z tw. VII.4.

Dowód (twierdzenia VII.4).
Niech $x_0 \in [a;b)$ i niech $\epsilon > 0$. Ponieważ f jest ciągła w x_0, zatem dobierzemy $\delta > 0$ takie, że $b > x_0 + \delta$ oraz dla $t \in [a;b]$ spełniających $|t - x_0| < \delta$ zachodzi

$$f(x_0) - \epsilon < f(t) < f(x_0) + \epsilon$$

Jeżeli zatem $0 < h < \delta$, to dla $t \in [x_0; x_0 + h]$ nierówności powyższe zachodzą, więc na mocy twierdzenia VII.3 pkt. 2 możemy „scałkować je stronami" i korzystając z przykładu 1 ze strony 106 otrzymujemy

$$h(f(x_0) - \epsilon) \leqslant \int_{[x_0;x_0+h]} f(t)dt \leqslant h(f(x_0) + \epsilon). \tag{VII.12}$$

Z drugiej strony, iloraz różnicowy dla F można na mocy twierdzenia VII.3 pkt. 3 zapisać następująco:

$$\frac{F(x_0 + h) - F(x_0)}{h} = \frac{1}{h} \int_{[x_0;x_0+h]} f(t)dt.$$

Zatem dzięki (VII.12), dla $0 < h < \delta$ mamy

$$f(x_0) - \epsilon \leqslant \frac{F(x_0 + h) - F(x_0)}{h} \leqslant f(x_0) + \epsilon.$$

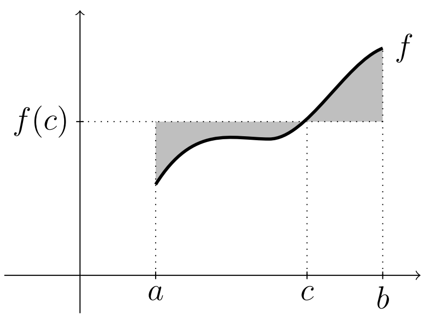
Wykazaliśmy więc, że $F'_+(x_0) = f(x_0)$ i analogicznie dowodzi się, że $F'_-(x_0) = f(x_0)$ dla $x \in (a;b]$. □

Pewna grupa twierdzeń dotyczących całkowania nosi wspólną nazwę „twierdzenia o wartości średniej" (tym razem — dla całek). Zajmiemy się tu tylko najprostszym z nich. Zakładamy, że $b > a$.

Twierdzenie VII.5 (o wartości średniej). *Jeżeli* $f : [a;b] \to \mathbb{R}$ *jest ciągła, to istnieje* $c \in (a;b)$ *takie, że*

$$\int_{[a;b]} f(x)dx = f(c) \cdot (b - a).$$

Uwaga. Geometryczny sens tego twierdzenia jest taki, że dla pewnej wartości $f(c)$ funkcji f pole prostokąta o podstawie na odcinku $[a;b]$ i wysokości $f(c)$ pokrywa się z polem pomiędzy osią X a wykresem f (patrz rysynek 18).



Rysunek 18: Dobór c w tw. VII.5 powinien być taki, by zaznaczone pola miały równe wartości.

Dowód.
Jeśli F — funkcja z twierdzenia VII.4, to na mocy twierdzenia Lagrange'a o wartości średniej (tw. V.4) oraz na mocy twierdzenia VII.4 mamy

$$\frac{\int_{[a;b]} f(x)dx}{b - a} = \frac{F(b) - F(a)}{b - a} = F'(c) = f(c)$$

dla pewnego $c \in (a;b)$. □

3. Całki niewłaściwe

Całkowanie w sensie Riemanna było „z definicji" wykonalne tylko dla funkcji całkowalnych w sensie Riemanna. A zatem, w szczególności, dziedzina całkowanej funkcji musiała być przedziałem domkniętym (więc — między innymi — o skończonej długości), a sama funkcja — ograniczona. Obecnie pokażemy jak można rozszerzyć pojęcie całki tak, by objąć także niektóre przypadki, gdy powyższe warunki spełnione nie są. Posłuży do tego właśnie pojęcie całki niewłaściwej[113] . Sam pomysł jest bardzo prosty i w swojej istocie bardzo przypomina pomysł, który posłużył przy definicji sumy szeregu. Niech $f : [a; b) \to \mathbb{R}$ gdzie $b > a$, przy czym $b = +\infty$ lub $b \in \mathbb{R}$. W obu sytuacjach całka Riemanna z f nie jest poprawnie zdefiniowana. Załóżmy jednak, że

$$\forall_{r\in[a;b)}\ f|_{[a;r]} \in \Re. \tag{VII.13}$$

Definicja. *Jeżeli istnieje* $\lim_{r\to b-} \int_{[a;r]} f$*, to granicę tę nazywamy* ***całką niewłaściwą*** *z* f *(po przedziale* $[a; b)$*) i oznaczamy ją symbolem*

$$\int_a^b f(x)dx\ ^{114)}\ .$$

Mówimy, że powyższa całka (niewłaściwa) jest ***zbieżna*** *wtw granica ta istnieje i jest skończona* [115] .

Tradycyjnie, gdy $b = +\infty$ mówi się o całce niewłaściwej *I rodzaju*, a gdy $b \in \mathbb{R}$ — *II rodzaju*. Opisana tu sytuacja dotyczy „niewłaściwości prawostronnej", tzn. sytuacji gdy f nie jest określona w punkcie b — prawym końcu dziedziny. Całkiem analogicznie postępuje się w przypadku „niewłaściwości lewostronnej", tzn. gdy $f : (b; a] \to \mathbb{R}$ (należy wówczas zastąpić „$b-$" przez „$b+$", „$[a; r]$" przez „$[r; a]$", „$\int_a^b$" przez „$\int_b^a$" oraz „$[a; b)$" przez „$(b; a]$").

Uwagi.

1. Jeżeli $b \in \mathbb{R}$ i f przedłuża się do funkcji $\tilde{f} : [a; b] \to \mathbb{R}$ takiej, że $\tilde{f} \in \Re$, to łatwo sprawdzić (zad. VII.13), że $\int_a^b f(x)dx$ jest zbieżna, a przy tym $\int_a^b f(x)dx = \int_{[a;b]} \tilde{f}(x)dx$. Z tego (i nie tylko tego) względu lepszym może oznaczeniem na całkę nieoznaczoną byłoby „$\int_{[a;b)} f(x)dx$", jednak pozostaniemy przy oznaczeniu tradycyjnym.

2. Można też zdefiniować pojęcie całek niewłaściwych „mieszanych", służących obliczaniu „całki" z funkcji określonych na przedziałach obustronnie otwartych, albo — co gorsza — przedziałów pozbawionych pewnych punktów (skończonej liczby). Tu tej definicji nie podamy (patrz jednak zad. VII.14).

Przykłady. Przy użyciu wniosku z p.t.r.c. (tw. VII.4) bez trudu wyliczamy, że

1. $\int_0^1 \frac{1}{x^\alpha}dx$ („lewostronnie niewłaściwa") jest dla $\alpha > 0$ zbieżna wtw $\alpha < 1$ oraz wówczas równa jest $\frac{1}{1-\alpha}$ (patrz rys. 19)

2. $\int_1^{+\infty} \frac{1}{x^\alpha}dx$ („prawostronnie niewłaściwa") jest dla $\alpha > 0$ zbieżna wtw $\alpha > 1$ oraz wówczas równa jest $\frac{1}{\alpha-1}$ (patrz rys. 20).

3. $\int_{-\infty}^0 e^x dx = -\int_0^1 \ln(x)dx = 1$. Pierwsza z równości jest intuicyjnie jasna ze względu na symetrię wykresów obu rozważanych funkcji (patrz rys. 21).

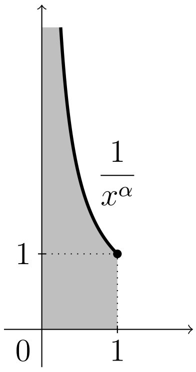
Teoria całek niewłaściwych posiada wiele analogii do teorii szeregów (czego można było się spodziewać już na podstawie samej definicji). Zilustrujemy to przy pomocy dwóch kryteriów zbieżności całek niewłaściwych.

---

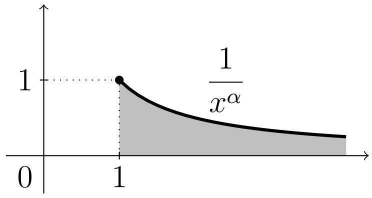
[113] Proszę nie mylić z całką nieoznaczoną!

[114] A zatem tak jak całkę oznaczoną, co jednak nie powinno prowadzić do nieporozumień.

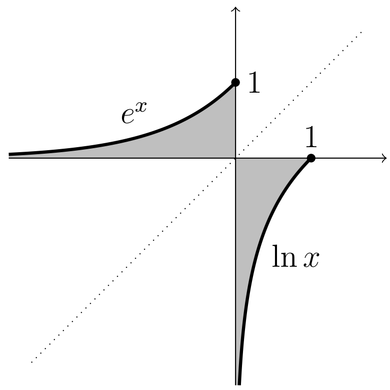
[115] W wypadku przeciwnym (granica nie istnieje lub jest równa $+\infty$ lub $-\infty$) mówimy, że całka jest *rozbieżna*.



Rysunek 19: Geometryczny sens całki niewłaściwej $\int_0^1 \frac{1}{x^\alpha}dx$ to pole nieograniczonego obszaru częściowo tylko zaznaczonego tu...



Rysunek 20: ... i podobnie dla $\int_1^{+\infty} \frac{1}{x^\alpha}dx$.



Rysunek 21: Zakreskowane pola powierzchni są równe.

Kryterium 1 (porównawcze). *indexkryterium!porównawcze Załóżmy, że* $f_1, f_2 : [a; b) \to \mathbb{R}$ *i obie funkcje spełniają warunek* (VII.13). *Jeśli dla dowolnego* $x \in [a; b)$

$$0 \leqslant f_1(x) \leqslant f_2(x),$$

oraz $\int_a^b f_2(x)dx$ *jest zbieżna, to* $\int_a^b f_1(x)$ *też jest zbieżna.*

Dowód.
Niech $F_j(r) := \int_{[a;r]} f_j(x)dx$ dla $r \in [a; b)$, $j = 1, 2$. Na mocy twierdzenia VII.3 funkcje F_1 i F_2 są rosnące oraz $0 \leqslant F_1(r) \leqslant F_2(r)$ dla dowolnego r. Zatem teza wynika natychmiast z twierdzenia o granicach jednostronnych funkcji monotonicznych (tw. IV.7) i z twierdzenia o zachowaniu nierówności przy przejściu granicznym (tw. IV.5). □

Drugie kryterium dotyczy jedynie całek niewłaściwych I-go rodzaju.

Kryterium 2 (Dirichleta). *indexkryterium!Dirichleta Jeżeli* $f, g : [a; +\infty) \to \mathbb{R}$ *oraz*

1. f *spełnia* (VII.13) *(z* $b = +\infty$*) i istnieje* $M \in \mathbb{R}$ *takie, że*

$$\forall_{r \in [a;+\infty)} \left| \int_{[a;r)} f(x)dx \right| < M;$$

2. g *jest malejąca i* $\lim_{x \to +\infty} g(x) = 0$,

to $\int_a^{+\infty} f(x)g(x)dx$ *jest zbieżna.* **B.D.**

Oczywiście analogiczne twierdzenia zachodzą także dla całek „lewostronnie niewłaściwych". Przykłady zastosowań powyższych kryteriów do badania zbieżności konkretnych całek niewłaściwych pozostawiamy na ćwiczenia (patrz — zadania do tego rozdziału).

Zadania do Rozdziału VII

$\forall$ 1. Oblicz poniższe całki oznaczone i nieoznaczone. W punktach b), d), f), g) przy stosowaniu twierdzenia o całkowaniu przez podstawienie przedstaw „dwa” rozwiązania: jedno z użyciem nieformalnego chwytu „$dy = \frac{dy}{dx} \cdot dx$” i drugie — całkowicie formalne i ścisłe (ze szczegółowym wypisaniem postaci funkcji, do których stosowane jest twierdzenie o c.p.p.).

(a) $\int x^3 \ln x dx$, $x > 0$

(b) $\int \operatorname{tg} x dx$, $x \neq k\pi + \frac{\pi}{2}$ $(k \in \mathbb{Z})$

(c) $\int_{-100}^{1000} e^{(2x+1)} dx$

(d) $\int_0^1 x^3 \sqrt{7 + x^4} dx$

(e) $\int_0^1 \operatorname{arctg} x dx$

(f) $\int \frac{x^2}{1+x^3} dx$, $x \neq -1$

(g) $\int \frac{x^4}{1+x^2}$, $x \in \mathbb{R}$

(h) $\int \frac{1}{1+x^2+x^4} dx$, $x \in \mathbb{R}$

(i) $\int_0^1 \frac{e^t}{e^{2t}+e^t+1} dt$

(j) $\int \frac{1}{1+\cos x}$, $x \in (-\pi; \pi)$

(k) $\int_0^x e^s \sin(3s) ds$

(l) $\int_0^{2\pi} \frac{1}{5-3\cos x} dx$ (Uwaga: wynik musi być $> 0 \ldots$)

(m) $\int \frac{1}{x^{\frac{3}{2}}+x^{\frac{1}{2}}} dx$, $x > 0$

(n) $\int_1^e \ln x dx$

(o) $\int |x| dx$, $x \in \mathbb{R}$

(p) $\int_0^1 \frac{x}{x^2+1} dx$

(q) $\int_0^\pi \sin x e^x dx$

(r) $\int \frac{2y}{(y-3)^2(y^2+3)} dy$, $y > 3$

(s) $\int_0^{\frac{\pi}{2}} \sin^2 x dx$

(t) $\int_0^{\frac{\pi}{2}} \cos^2 x dx$

(u) $\int_{-1}^1 \sin(x^3) dx$

(v) $\int_4^9 \frac{t-1}{\sqrt{t+1}} dt$

(w) $\int_1^2 \frac{e^{\frac{1}{x}}}{x^2} dx$

(x) $\int_0^1 \frac{s}{1+s^4} ds$ (Wskazówka: podstaw „$y = s^2$”).

$\forall$ 2. [116] Dla $f : [a; b] \to \mathbb{R}$ określamy:

- *długość wykresu* f, o ile f jest klasy C^1, wzorem

$$\int_a^b \sqrt{1 + (f'(x))^2} dx;$$

[116] Przynajmniej po jednym przykładzie na każdy z trzech wzorów

- *pole powierzchni obrotowej* powstałej przez obrót wykresu f (zawartego w płaszczyźnie XY) wokół osi X w przestrzeni XYZ, o ile f jest klasy C^1, wzorem
$$2\pi \int_a^b f(x)\sqrt{1+(f'(x))^2}dx;$$
- *objętość bryły obrotowej* ograniczonej powyższą powierzchnią obrotową i płaszczyznami „$x = a$", „$x = b$", o ile f jest ciągła, wzorem
$$\pi \int_a^b (f(x))^2 dx.$$

W oparciu o powyższe wzory oblicz:

(a) długość okręgu o promieniu r,

(b) objętość kuli o promieniu r,

(c) pole powierzchni sfery o promieniu r,

(d) objętość walca obrotowego o promieniu podstawy r i wysokości h,

(e) pole powierzchni bocznej walca obrotowego o promieniu podstawy r i wysokości h,

(f) objętość stożka obrotowego o promieniu podstawy r i wysokości h,

(g) pole powierzchni bocznej stożka obrotowego o promieniu podstawy r i wysokości h.

$\forall$ 3. Niech $I_n(x) := \int_0^x \frac{1}{(t^2+1)^n} dt$ dla $n \in \mathbb{N}$. Znajdź wzór rekurencyjny na funkcje I_n (nie wymagający użycia całek).

4. Wyprowadź wzory związane z podstawieniem trygonometrycznym „$t = tg(\frac{x}{2})$" podane na wykładzie (strona 103).

$\forall$ 5. Wykaż, że funkcja $\chi : [-1; 1] \to \mathbb{R}$ zadana wzorem
$$\chi(x) = \begin{cases} 0 \text{ dla } -1 \leqslant x \leqslant 0 \\ 1 \text{ dla } 0 < x \leqslant 1 \end{cases}$$
jest całkowalna w sensie Riemanna (choć nie jest ciągła ...).

$\forall$ 6. Znajdź granice ciągów zadanych poniższymi wzorami, wykorzystując twierdzenie o sumach Riemanna (twierdzenie VII.2).

(a) $\frac{1}{n^{\alpha+1}} \sum_{k=1}^n k^\alpha$, dla $\alpha \geqslant 0$,

(b) $\sum_{k=n+1}^{2n} \frac{1}{k}$.

7. Niech $f : I \to \mathbb{R}$, gdzie I — przedział oraz f posiada funkcję pierwotną. Wykaż, że dla dowolnych $a, b, c \in I$ zachodzi $\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx$.

8. Wykaż, że jeżeli $f, g \in C([a; b])$, $b > a$ oraz $f(x) < g(x)$ przy dowolnym $x \in (a; b)$, to $\int_{[a;b]} f(x)dx < \int_{[a;b]} g(x)dx$.

9. Znajdź $\lim_{n\to+\infty} \int_n^{n+7} \frac{\sin x}{x} dx$.

$\forall$ 10. Wykaż, że jeżeli $f_n, f \in \Re([a; b])$ oraz $f_n \rightrightarrows f$, to
$$\int_{[a;b]} f_n \to \int_{[a;b]} f.$$

$\forall$ 11. Zbadaj zbieżność poniższych całek niewłaściwych:

(a) $\int_0^{+\infty} \frac{1}{\frac{1}{2}+x^2} dx$

(b) $\int_0^1 \frac{1}{\sqrt{1-x}} dx$

(c) $\int_2^{+\infty} \frac{1}{\ln x} dx$

(d) $\int_{-\infty}^0 \frac{x}{x^2+\frac{1}{2}} dx$

(e) $\int_0^{+\infty} x \cdot e^{-x} dx$

(f) $\int_0^{+\infty} x^{17} e^{-\sqrt{x}} dx$

(g) $\int_0^1 (\ln x)^2 dx$

(h) $\int_1^{+\infty} \frac{\sin x}{x} dx$

(i) $\int_1^{+\infty} \frac{\cos x}{\sqrt[3]{x}} dx$

(j) $\int_0^1 \frac{\sin x}{x^{\frac{3}{2}}} dx$

(k) $\int_0^1 \frac{1}{\sqrt[3]{1-x^3}} dx$.

12. Oblicz $\int_0^{+\infty} x^n e^{-x} dx$ dla każdego $n \in \mathbb{N}_0$.

13. Wykaż, że jeżeli $\tilde{f} : [a;b] \to \mathbb{R}$ i $\tilde{f} \in \Re$ oraz $f = \tilde{f}\,|_{[a;b)}$ to $\int_a^b f(x)dx$ jest zbieżna oraz równa się $\int_{[a;b]} \tilde{f}(x)dx$.

14. Zdefiniujemy całkę niewłaściwą *mieszaną* dla $f : (a,b) \to \mathbb{R}$ ($a, b \in \overline{\mathbb{R}}$, $a < b$). Zakładamy, że dla dowolnych a', b' takich, że $a < a' < b' < b$ zachodzi $f\,|_{[a';b']} \in \Re$. Niech $c \in (a;b)$. Mówimy, że $\int_a^b f(x)dx$ istnieje wtw istnieją $\int_a^c f(x)dx$ oraz $\int_c^b f(x)dx$ oraz ich suma jest określona. W tej sytuacji $\int_a^b f(x)dx$ określamy jako powyższa sumę. Wykaż, że definicja ta (istnienie i wartość całki) nie zależy od wyboru $c \in (a,b)$.

15. Zbadaj zbieżność (tzn. istnienie i skończoność) całki niewłaściwej mieszanej (patrz zad. VII.14):

(a) $\int_a^{+\infty} \frac{1}{(x-a)^\alpha} dx$, w zależności od $a, \alpha \in \mathbb{R}$

(b) $\int_{-\infty}^a \frac{1}{(a-x)^\alpha} dx$, w zależności od $a, \alpha \in \mathbb{R}$

(c) $\int_0^\pi \frac{\sin x}{x^{\frac{3}{2}} \cdot (\pi-x)^\alpha} dx$, w zależności od $\alpha \in \mathbb{R}$.

16. Zaproponuj sformułowanie „kryterium asymptotycznego dla całek niewłaściwych” wzorując się na sytuacji znanej z teorii szeregów. Udowodnij tak sformułowane kryterium.

17. Udowodnij następujące „kryterium całkowe zbieżności szeregów”: *Jeżeli* $f : [n_0; +\infty) \to [0; +\infty)$ *jest malejąca i ciągła, to* $\int_{n_0}^{+\infty} f(x)dx$ *jest zbieżna wtw* $\sum_{n=n_0}^{+\infty} f(n)$ *jest zbieżny.*

$\forall$ 18. [117] Wykorzystaj powyższe kryterium jako alternatywną metodę badania zbieżności szeregów $\sum_{n=1}^{+\infty} \frac{1}{n^\alpha}$ oraz $\sum_{n=2}^{+\infty} \frac{1}{n(\ln n)^\alpha}$ dla $\alpha > 0$.

[117] Przynajmniej jeden z dwóch przykładów.

VIII Ciągłość funkcji wielu zmiennych. Przestrzenie metryczne

[około 2 wykładów]

1. Przestrzenie metryczne

Gdy wprowadzaliśmy we wcześniejszych rozdziałach pojęcie granicy ciągu liczbowego oraz związane z tym pojęcia takie jak granica funkcji, czy ciągłość, widoczne było, że zasadniczą rolę umożliwiającą sformułowanie i zrozumienie tych pojęć pełniło tu coś co nazywaliśmy *odległością* pomiędzy liczbami, mającą dla liczb x i y wartość równą

$$|x-y|.$$

Wydaje się więc, że moglibyśmy bez większych trudności uogólnić powyższe pojęcia na przypadki ciągów o wyrazach z innych zbiorów niż tylko $\mathbb{R}$ lub funkcji działających pomiędzy takimi zbiorami, o ile potrafilibyśmy w jakiś sposób mierzyć odległość pomiędzy ich elementami.

Zacznijmy zatem od problemu mierzenia odległości. I tak jak to się typowo czyni w matematyce, podejdziemy do tego w sposób abstrakcyjny. Tzn. nie będziemy na razie rozważać konkretnych zbiorów i sposobów mierzenia w nich odległości, lecz założymy, że mamy pewien zbiór X oraz funkcję

$$\rho : X \times X \to [0; +\infty)$$

mającą pełnić rolę odległości (tzn. dla $x, y \in X$ odległość pomiędzy x a y ma być równa $\rho(x, y)$) i sformułujemy minimalne warunki jakie funkcja ta powinna spełniać, aby się mogła do celu takiego nadawać. Taka abstrakcyjna matematyczna odległość nosi nazwę *metryki.*

Definicja. *ρ jest* ***metryką*** *wtw*

1) *(niezdegenerowanie)* $\forall_{x,y \in X}\ (\rho(x, y) = 0 \iff x = y)$*;*

2) *(symetria)* $\forall_{x,y \in X}\ \rho(x, y) = \rho(y, x)$*;*

3) *(nierówność trójkąta)* $\forall_{x,y,z \in X}\ \rho(x, z) \leqslant \rho(x, y) + \rho(y, z)$*.*

Gdy ρ jest metryką, to parę (X, ρ) nazywamy *przestrzenia metryczną* (czasem mówimy tak też na sam zbiór X, gdy jasne jest jaka metryka w X została wybrana).

Mimo, iż takie abstrakcyjne pojęcie odległości dopuszcza rozmaite „dziwne" przykłady, nie zanadto zgodne z przyzwyczajeniami niektórych osób do tego, co potocznie odległością bywa nazywane, okazuje się, że powyższa definicja zawiera akurat te warunki, które wystarczają do sensownego uogólnienia wspomnianych przez nas wcześniej pojęć. Tymi uogólnieniami zajmiemy się w dalszych podrozdziałach, teraz natomiast przyjrzyjmy się kilku przykładom.

Przykład 1 (metryka euklidesowa w $\mathbb{R}^d$). Najważniejszy dla nas w tym semestrze przykład metryki to standardowo stosowana, znana Państwu ze szkoły, *metryka euklidesowa* w $\mathbb{R}^d$ (znana przynajmniej dla $d = 2$ i 3). Dla $x, y \in \mathbb{R}^d$ zadana jest ona wzorem

$$\rho(x, y) = \left(\sum_{k=1}^{d} (x_k - y_k)^2\right)^{\frac{1}{2}} \tag{VIII.1}$$

gdzie $x = (x_1, \dots, x_d)$, $y = (y_1, \dots, y_d)$ [118]. W szczególności dla $d = 1$ otrzymujemy tu wspomnianą wcześniej odległość $\rho(x, y) = |x - y|$.

Taka odległość jest nam „najbliższa”, bo wzór (VIII.1) opisuje zgodnie z twierdzeniem Pitagorasa (użytym $d - 1$ razy) po prostu „zwykłą” (tzn. właśnie euklidesową ...) długość odcinka (lub inaczej — długość wektora) łączącego x i y. Inaczej mówiąc zachodzi

$$\rho(x, y) = \|x - y\|,$$

gdzie $\| \ \|$ jest *normą* euklidesową w $\mathbb{R}^d$ (oznaczoną na GAL-u jako $\| \ \|_2$) zadaną wzorem

$$\|x\| = \left(\sum_{k=1}^{d} x_k^2 \right)^{\frac{1}{2}}. \tag{VIII.2}$$

I choć, jak Państwo wiecie, jest wiele różnych norm zadanych w $\mathbb{R}^d$, ten symbol $\| \ \|$ będzie na naszym wykładzie z Analizy Matematycznej na ogół oznaczał właśnie normę euklidesową (niezależnie od wartości d). Pozostaje nam zatem sprawdzić, że ρ tu zdefiniowane jest metryką. Warunki 1) i 2) definicji są w tym przypadku oczywiste. Pozostaje do sprawdzenia 3). Ale na mocy wykazanej na GAL-u nierówności trójkąta dla normy euklidesowej

$$\rho(x, z) = \|x - z\| = \|x - y + y - z\| \leqslant \|x - y\| + \|y - z\| = \rho(x, y) + \rho(y, z)$$

czyli 3) zachodzi. A zatem $(\mathbb{R}^d, \rho)$ jest przestrzenią metryczną. Warto też wspomnieć, że gdy rozważamy nie cały zbiór $\mathbb{R}^d$, ale dowolny $D \subset \mathbb{R}^d$ i rozważymy $\tilde{\rho} := \rho|_{D \times D}$, to $\tilde{\rho}$ będzie oczywiście metryką w D (mówimy wtedy, że $(D, \tilde{\rho})$ jest podprzestrzenią metryczną[119] przestrzeni metrycznej $(\mathbb{R}^d, \rho)$).

Przykład 2 (metryka indukowana przez normę). Sytuację opisaną w przykładzie 1 można uogólnić na dowolną *przestrzeń unormowaną* $(X, \| \ \|)$, tzn. przestrzeń liniową X (nad $K = \mathbb{R}$ lub $\mathbb{C}$) z zadaną normą $\| \ \|$. Przypominam (znów z GAL-u) — funkcja $\| \ \| : X \to [0; +\infty)$ jest *normą*[120] wtw

i) (niezdegenerowanie) $\forall_{x \in X} \ (\|x\| = 0 \Rightarrow x = 0)$,

ii) (jednorodność) $\forall_{\lambda \in K} \ \forall_{x \in X} \ \|\lambda x\| = |\lambda| \cdot \|x\|$,

iii) (nierówność trójkąta) $\forall_{x, y \in X} \ \|x + y\| \leqslant \|x\| + \|y\|$.

A zatem możemy określić $\rho : X \times X \to [0; +\infty)$ wzorem

$$\rho(x, y) = \|x - y\| \tag{VIII.3}$$

i ρ tak zadane będzie metryką w X: warunek 1) wynika z i), warunek 2) z ii) (dla $\lambda = -1$) a warunek 3) z iii) (tak jak w przykładzie 1).

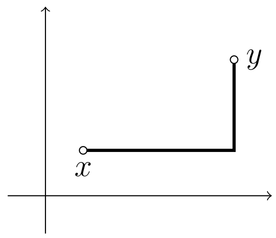
Np. gdy rozważymy $\mathbb{R}^2$ z normą $\| \ \|_1$ zadaną wzorem $\|x\|_1 = |x_1| + |x_2|$ to otrzymamy tzw. *metrykę miejską*, której nazwa bierze się stąd, że odległość „w niej” liczona odpowiada

---

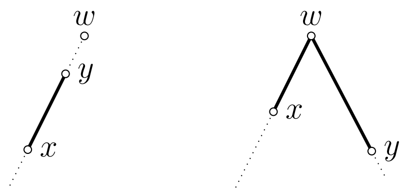
[118] Uwaga! Jak widać stosujemy tu nieco inną notację od tej stosowanej w semestrze zimowym na wykładzie z GAL-u. $\mathbb{R}^d$ nie oznacza teraz zbioru macierzy o 1 kolumnie i d-wierszach lecz po prostu iloczyn kartezjański $\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$, tj. d-„egzemplarzy” zbioru $\mathbb{R}$. Elementy $\mathbb{R}^d$ oznaczamy tu jedną literą (bez strzałki nad nią) i na ogół dla $x \in \mathbb{R}^d$ oraz $j \in \{1, \dots, d\}$ symbol x_j oznacza j-tą współrzędną x. $\mathbb{R}^d$ będziemy też traktować jako przestrzeń liniową (nad $\mathbb{R}$, z naturalnymi działaniami „po współrzędnych”) i jej elementy będą wektorami, co nie zmieni sposobu naszej notacji.

[119] Analogicznie konstrukcję metryki ograniczonej do podzbioru zbioru X można oczywiście przeprowadzić dla każdej przestrzeni metrycznej (X, ρ).

[120] Czasami słowa „norma” (nad)używa się także w odniesieniu do funkcji, które mogą osiągnąć wartość $+\infty$. Tak było np. w rozdziale VII gdy określaliśmy normę funkcji — dla funkcji f nieograniczonej zachodziło $\|f\| = +\infty$. Przy takim rozszerzeniu pojęcia normy podane tu warunki definicji wymagałyby jednak pewnej modyfikacji (warunek 2)).



Rysunek 22: Metryka miejska



Rysunek 23: Metryka kolejowa

najkrótszej drodze pomiędzy punktami, którą trzeba przebyć poruszając się wyłącznie wzdłuż kierunków prostopadłych osi współrzędnych (patrz rys. 22), co odpowiada sieci ulic w „idealnym" mieście.

Inny przykład metryki indukowanej przez normę, to metryka zadana wzorem (VIII.3) przez normę „supremum" z rozdziału VI rozważaną w przestrzeni $l^\infty(S)$ złożonej ze wszystkich funkcji f ograniczonych, $f : S \to \mathbb{R}$ (ew. $\mathbb{C}$), gdzie S jest pewnym ustalonym zbiorem.

Istnieją jednak także inne metryki, czasem dość „dziwne", które nie pochodzą od zadanej normy.

Przykład 3 (metryka kolejowa). Rozważmy płaszczyznę $\mathbb{R}^2$ i ustalony jej punkt w — „węzeł kolejowy". Odległość pomiędzy x a y liczona jest następująco: jeżeli oba punkty leżą na tej samej półprostej o początku w punkcie w, to $\rho(x, y)$ jest długością („euklidesową") odcinka xy, w przeciwnym przypadku $\rho(x, y)$ jest sumą długości odcinków xw i wy. Odpowiada to oczywiście drodze przebywanej przy użyciu „idealnej" promienistej sieci kolejowej z punktem węzłowym w. Sprawdzenie, że jest to metryka pozostawiam Państwu (patrz zad. VIII.1).

Przykład 4 (metryka dyskretna). Niech X będzie dowolnym zbiorem. Metrykę dyskretną w X określamy wzorem

$$\rho(x, y) = \begin{cases} 1 & \text{gdy } x \neq y \\ 0 & \text{gdy } x = y. \end{cases}$$

I tu dowód (prościutki), że mamy do czynienia z metryką pozostaje dla Państwa (patrz zadanie VIII.2).

2. Zbiory otwarte i domknięte. Zbieżność ciągów

Wprowadzone przez nas pojęcie metryki pozwala na wyróżnienie dwóch podstawowych typów zbiorów — zbiorów otwartych i domkniętych. Niech ρ będzie metryką w zbiorze X. Wzorując się na geometrycznej terminoligii zaczerpniętej z $\mathbb{R}^3$ określamy najpierw *kulę* (otwartą) o środku $a \in X$ i promieniu $r \geqslant 0$ w X wzorem:

$$K(a, r) := \{x \in X : \rho(a, x) < r\}.$$

Definicja. *Zbiór $U \subset X$ jest* ***otwarty*** *(w X) wtw $\forall_{x \in U}\ \exists_{r>0}\ K(x, r) \subset U$. Zbiór $F \subset X$ jest* ***domknięty*** *(w X) wtw $X \setminus F$ jest otwarty.*

Podkreślmy od razu, że nie każdy zbiór musi być otwarty lub domknięty oraz że otwartość i domkniętość nie wykluczają się wzajemnie. Np. X oraz $\emptyset$ są zarówno otwarte jak i domknięte. Pojęcie metryki pozwala też mówić o ograniczoności zbiorów. Mianowicie $B \subset X$ jest *ograniczony* wtw jest zawarty w pewnej kuli, tzn.

$$\exists_{a \in X, r \in \mathbb{R}_+}\ B \subset K(a, r).$$

W przypadku znanej nam dobrze przestrzeni metrycznej $\mathbb{R}$ zbiorami otwartymi są np. wszystkie przedziały otwarte, a domkniętymi — wszystkie przedziały domknięte, choć istnieje wiele różnych zbiorów otwartych, czy domkniętych nie będących przedziałami. Przedziały „otwarto-domknięte” nie są (wbrew nazwie) ani otwarte, ani domknięte. Każda kula (otwarta) jest w przestrzeni metrycznej zbiorem otwartym (patrz zad. VIII.3). W skrajnie „dziwnej” przestrzeni dyskretnej każdy podzbiór jest jednocześnie otwarty i domknięty (patrz zad. VIII.2). W każdej przestrzeni metrycznej zbiór jednopunktowy jest domknięty (patrz zad. VIII.3). Na szczęście w bliskiej naszym sercom przestrzeni euklidesowej $\mathbb{R}^d$ sytuacja jest całkiem odmienna.

Fakt VIII.1. *Jedynymi podzbiorami $\mathbb{R}^d$, które są jednocześnie otwarte i domknięte* [121] *są $\emptyset$ i $\mathbb{R}^d$* [122].

Obie klasy — zbiorów otwartych i domkniętych mają ważne algebraiczne (w znaczeniu działań na zbiorach) własności.

Fakt VIII.2. *Suma dowolnej rodziny i przecięcie skończonej rodziny zbiorów otwartych jest zbiorem otwartym. Suma skończonej rodziny i przecięcie dowolnej rodziny zbiorów domkniętych jest zbiorem domkniętym.*

Dowód.
Niech $U_i \subset X$, U_i — otwarte dla wszystkich $i \in I$, gdzie I — pewien zbiór indeksów. A zatem gdy $x \in \bigcup_{i\in I} U_i$, to $x \in U_{i_0}$ dla pewnego $i_o \in I$, więc $K(x,r) \subset U_{i_0} \subset \bigcup_{i\in I} U_i$ dla pewnego $r > 0$ — stąd $\bigcup_{i\in I} U_i$ — otwarty.

Gdy $x \in \bigcap_{i\in I} U_i$, to dla wszystkich $i \in I$ mamy: $x \in U_i$, a zatem $K(x, r_i) \subset U_i$ dla pewnego $r_i > 0$. Skoro I — skończony, to $r := \min\{r_i : i \in I\} > 0$. Jednocześnie $K(x,r) \subset K(x,r_i) \subset U_i$ dla każdego i, skąd $K(x,r) \subset \bigcap_{i\in I} U_i$, co dowodzi otwartości zbioru $\bigcap_{i\in I} U_i$. Teza dla zbiorów domkniętych wynika teraz natychmiast ze wzorów de Morgana. □

W szczególności zatem dowolny zbiór skończony jest domknięty (w każdej przestrzeni metrycznej) jako skończona suma zbiorów jednopunktowych.

Warto jeszcze wspomnieć, że w matematyce rozważa się także ogólniejsze przestrzenie niż przestrzenie metryczne, mianowicie *przestrzenie topologiczne*. Tam pojęciem „pierwotnym” jest właśnie rodzina zbiorów otwartych (zwana topologią).

Tak jak już zapowiadaliśmy w przestrzeniach metrycznych można zdefiniować pojęcie granicy ciągu:

Definicja. *Ciąg $\{x_n\}_{n\geqslant n_0}$ elementów X jest* ***zbieżny*** *do $g \in X$ wtw*

$$\forall_{\epsilon>0}\ \exists_{N\geqslant n_0}\ \forall_{n\geqslant N}\ \rho(x_n, g) < \epsilon.$$

Jak więc widać uogólnienie tego pojęcia z przypadku ciągów liczbowych polega po prostu na zastąpieniu napisu „$|x_n - g|$” napisem „$\rho(x_n, g)$”. Podobnie jak było to dla ciągów liczbowych będziemy używali wymiennie symboli $x_n \to g$ oraz $\lim_{n\to+\infty} x_n = g$ i g będziemy nazywali *granicą* ciągu $\{x_n\}_{n\geqslant n_0}$. Podkreślmy, że rozważamy tu tylko $g \in X$, nie mamy bowiem ogólnie rzecz biorąc żadnych odpowiedników $+\infty$ i $-\infty$. Tak jak wcześniej, ciąg *zbieżny* oznacza zbieżny do pewnego $g \in X$. Łatwo zauważyć, że zachodzi

$$x_n \to g \quad wtw \quad \rho(x_n, g) \to 0 \tag{VIII.4}$$

Powyższy fakt pozwala więc sprowadzić badanie zbieżności w przestrzeni metrycznej X do badania zbieżności zwykłego ciągu liczbowego $\{\rho(x_n, g)\}$ (o ile wiemy, jaka miałaby być ewent. granica g). Jednak w najważniejszym dla nas przypadku — przestrzeni euklidesowej sprawa jest jeszcze prostsza. Umówmy się najpierw co do notacji: gdy rozważamy ciąg

[121] W sensie metryki euklidesowej — taki wybór uznajemy za umowny w tym i dalszych rozdziałach i nie będziemy o tym przypominać. Wszelkie odstępstwa od tej umowy będą wyraźnie zaznaczane.
[122] Przestrzenie metryczne o tej własności co opisana tu własność $\mathbb{R}^d$ noszą nazwę *spójnych*.

$\{x_n\}_{n\geqslant n_0}$ o wyrazach w $\mathbb{R}^d$, to $(x_n)_j$ będzie oznaczać j-tą współrzędną punktu x_n, czyli $x_n = ((x_n)_1, \ldots, (x_n)_d)$.

Fakt VIII.3. *Zbieżność w $\mathbb{R}^d$ jest zbieżnością „po współrzędnych", tzn. $x_n \to g$ wtw dla każdego $j = 1, \ldots, d$ zachodzi*

$$(x_n)_j \to g_j.$$

Dowód.
Mamy $\|x_n - g\| \geqslant |(x_n)_j - g_j| \geqslant 0$ więc „$\Rightarrow$" wynika z twierdzenie o 3 ciągach i z (VIII.4). Z drugiej strony (VIII.4) daje też „$\Leftarrow$", gdyż $\|x_n - g\| = (\sum_{j=1}^{d}((x_n)_j - g_j)^2)^{\frac{1}{2}}$. □

Przykład.

$$(\frac{1}{n}, \sqrt[n]{2}, \sqrt[n]{n}) \to (0, 1, 1).$$

Po tych wstępnych rozważaniach warto chyba zadać następujące pytanie związane z definicją zbieżności:

> *Po co zakładaliśmy, że metryka ρ musi spełniać aż tyle warunków, skoro przy definicji zbieżności nie odegrały one żadnej roli?*

Otóż przy samej definicji może nie odegrały, ale aby ta definicja spełniała nasze naturalne oczekiwania, granica (o ile istnieje) powinna być wyznaczona jednoznacznie.

Fakt VIII.4. *Ciąg zbieżny w przestrzeni metrycznej ma tylko jedną granicę.*

Dowód.
Załóżmy, że $x_n \to g$ i $x_n \to g'$. Zatem na mocy warunków 2) i 3) definicji metryki oraz (VIII.4) mamy

$$0 \leqslant \rho(g, g') \leqslant \rho(g, x_n) + \rho(x_n, g') = \rho(x_n, g) + \rho(x_n, g') \to 0$$

skąd $\rho(g, g') = 0$ na mocy twierdzenia o trzech ciągach. Czyli dzięki warunkowi 1) definicji metryki $g = g'$. □

Zauważmy, że w powyższym dowodzie zostały użyte wszystkie trzy warunki z definicji metryki!

Okazuje się, że pojęcie zbieżności ciągów może posłużyć do sformułowania tzw. „ciągowej definicji" domkniętości zbioru — odzwierciedlającej chyba lepiej niż definicja pierwotna związek z nazwą „domknięty".

Fakt VIII.5. *Zbiór $F \subset X$ jest domknięty wtw dla dowolnego ciągu $\{x_n\}_{n\geqslant n_0}$ o wyrazach w F, jeżeli dla $x \in X$ zachodzi $x_n \to x$, to $x \in F$.* **B.D.**

Dowód tego faktu pozostawiam jako zadanie dla Państwa (zad. VIII.9). Oznacza on, że domkniętość zbioru to „możliwość przechodzenia do granicy z przynależnością do tego zbioru".

Przyjęte przez nas bardzo ogólne warunki z definicji metryki pozwoliły co prawda wykazać jednoznaczność granicy, ale nie pozwalają ogólnie uzyskać bardzo wielu analogów twierdzeń znanych nam z teorii ciągów liczbowych. Na szczęście jednak spora ich część jest prawdziwa w przestrzeniach euklidesowych $\mathbb{R}^d$. Oto dwa z nich.

Twierdzenie VIII.1 (Bolzano - Weiestrassa). *Każdy ciąg ograniczony*[123] *w $\mathbb{R}^d$ posiada podciąg zbieżny.*

Dowód.
Jeżeli $\{x_n\}_{n\geqslant n_0}$ jest ograniczony, to oczywiście każdy z ciągów $\{(x_n)_j\}_{n\geqslant n_0}$ dla $j = 1, \ldots, d$ jest także ograniczony. Dzięki tw. Bolzano - Weiestrassa dla ciągów liczbowych znajdziemy zatem taki ściśle rosnący ciąg indeksów $\{k_n^{(1)}\}$, że $(x_{k_n^{(1)}})_1 \to g_1$ dla pewnego $g_1 \in \mathbb{R}$. Znaleźliśmy więc podciąg ciągu $\{x_n\}$, który jest „zbieżny na pierwszej współrzędnej". Możemy więc postąpić

[123] Ciąg o wyrazach w przestrzeni metrycznej jest ograniczony wtw zbiór jego wyrazów jest ograniczony.

analogicznie, wybierając podciąg tego znalezionego już podciągu „zbieżny na drugiej współrzędnej” [124] i nie stracimy przy tym zbieżności na współrzędnej 1-szej. Jednocześnie podciąg podciągu jest też podciągiem, zatem kontynuując to postępowanie, po d-krokach uzyskamy podciąg zbieżny na każdej współrzędnej, a więc zbieżny na mocy faktu VIII.3. □

Z drugiej strony, w dowolnej przestrzeni metrycznej ciąg zbieżny musi być ograniczony (patrz zad. VIII.7).

Twierdzenie VIII.2 (o zupełności). *Ciąg $\{x_n\}_{n\geqslant n_0}$ o wyrazach w $\mathbb{R}^d$ jest zbieżny wtw*

$$\forall_{\epsilon>0}\ \exists_{N\geqslant n_0}\ \forall_{m,n\geqslant N}\ \rho(x_m,x_n)<\epsilon \text{ [125]}.$$

Dowód pominiemy, gdyż jest on niemal identyczny z dowodem tego twierdzenia dla przypadku skalarnego (tw. II.7). Warto tylko zauważyć, że prostsza implikacja, tj. „$\Rightarrow$” zachodzi nie tylko w $\mathbb{R}^d$, ale w dowolnej przestrzeni metrycznej.

Zdefiniujemy jeszcze tzw. zbiory *zwarte*. W przypadku podzbiorów $\mathbb{R}^d$ zbiór jest *zwarty* wtw jest domknięty i ograniczony. Ogólnie definicja jest inna (patrz przypis niżej). Z faktu VIII.5 i twierdzenia VIII.1 otrzymujemy następujący wynik.

Wniosek. *Jeżeli $K\subset\mathbb{R}^d$ jest zwarty, to każdy ciąg o wyrazach w K posiada podciąg zbieżny do granicy z K [126].*

Co więcej nie trudno również wykazać implikację przeciwną (patrz zad. VIII.10).

Przykładem zbioru zwartego jest więc każdy przedział domknięty w $\mathbb{R}$. Zbiorami, które można uznać za wielowymiarowe analogi przedziałów domkniętych są „kostki domknięte”, czyli $\{x\in\mathbb{R}^d : \forall_{j=1,\ldots,d}\ a_j\leqslant x_j\leqslant b_j\}$, gdzie $a_1,\ldots,a_d,b_1\ldots,b_d\in\mathbb{R}$ oraz „kule domknięte”, czyli $\{x\in\mathbb{R}^d : \|a-x\|\leqslant r\}$, gdzie $a\in\mathbb{R}^d$, $r\in[0;+\infty)$. Jedne i drugie są zbiorami zwartymi.

3. Granica i ciągłość funkcji wielu zmiennych

Obecnie zajmiemy się głównie granicą i ciągłością funkcji określonych w $D\subset\mathbb{R}^k$ o wartościach w $\mathbb{R}^l$. Ponieważ jednak zarówno definicje, jak i część wyników dają się sformułować ogólnie — dla przestrzeni metrycznych, zatem załóżmy, że mamy dwie przestrzenie metryczne (X,ρ_X), (Y,ρ_Y), podzbiór $D\subset X$ oraz funkcję $f : D\to Y$. Poniższa definicja jest przeniesieniem na bardziej abstrakcyjny grunt znanych Państwu definicji z rodziału IV. Jest to więc zarazem uogólnienie jak i przypomnienie.

Definicja.

- *Punkt $a\in X$ jest* ***punktem skupienia*** *zbioru D (przypominam skrót „p. s.”) wtw istnieje ciąg $\{x_n\}_{n\geqslant n_0}$ elementów zbioru $D\setminus\{a\}$ taki, że $x_n\to a$.*
- *Niech a — p. s. D oraz niech $g\in Y$. Wówczas $\lim_{x\to a}f(x)=g$, czyli g jest* ***granicą*** *f* ***w punkcie*** *a wtw dla dowolnego ciągu $\{x_n\}_{n\geqslant n_0}$ elementów $D\setminus\{a\}$ takiego, że $x_n\to a$ (w sensie ρ_X) zachodzi $f(x_n)\to g$ (w sensie ρ_Y).*
- *Niech $a\in D$. Funkcja f jest* ***ciągła w (punkcie)*** *a wtw dla dowolnego ciągu $\{x_n\}_{n\geqslant n_0}$ elementów D takiego, że $x_n \underset{\rho_X}{\longrightarrow} a$ [127] zachodzi $f(x_n)\underset{\rho_Y}{\longrightarrow} f(a)$.*

[124] Pozostawiam Słuchaczom / Czytelnikom sformułowanie pojęcia „zbieżności na j-tej współrzędnej”.

[125] Jak nietrudno się domyślić warunek ten stanowi definicję *ciągu Cauchy'ego* w dowolnej przestrzeni metrycznej. Jednocześnie te przestrzenie metryczne, w których zbieżność jest równoważna warunkowi Cauchy'ego, tak jak w tw. VIII.2, nazywane są przestrzeniami *zupełnymi*.

[126] Teza tego wyniku stanowi jednocześnie definicję zwartości w przypadku ogólnym.

[127] Zamiast pisać, że chodzi o zbieżność w sensie metryki ρ, w przypadku gdyby mogła powstać ewentualnie niejasność o jaką chodzi metrykę, będziemy czasem pisać „$\underset{\rho}{\to}$” zamiast samej strzałki „$\to$”.

- *Funkcja f jest* **ciągła** *wtw jest ciągła w każdym $a \in D$.*

Podane tu definicje granicy i ciągłości w punkcie to tzw. *wersje Heinego*. Podobnie jak dla przypadku skalarnego równoważne im są odpowiednie *wersje Cauchy'ego*. Np. dla ciągłości w punkcie a ma ona postać:

$$\forall_{\epsilon>0} \exists_{\delta>0} \forall_{x\in D} \ (\rho_X(x,a) < \delta \Rightarrow \rho_Y(f(x), f(a)) < \epsilon).$$

Również związek ciągłości z granicą jest analogiczny jak w przypadku skalarnym, tzn. gdy $a \in D$ i a jest p. s. D, to f jest ciągła w a wtw

$$\lim_{x\to a} f(x) = f(a).$$

W przypadku funkcji określonych na podzbiorach $\mathbb{R}^d$ można oprócz „zwykłej" (zdefiniowanej w tym podrozdziale) granicy funkcji rozważać także tzw. *granice iterowane*. Np. dla funkcji dwóch zmiennych będą to granice następujące

$$\lim_{x_2\to a_2} \lim_{x_1\to a_1} f(x_1, x_2) \quad \text{oraz} \quad \lim_{x_1\to a_1} \lim_{x_2\to a_2} f(x_1, x_2)$$

(o ile mają one sens i istnieją). Należy tu podkreślić, że związki pomiędzy istnieniem (też wartością) dla poszczególnych granic iterowanych oraz dla „zwykłej" granicy f w $a = (a_1, a_2)$ są bardzo luźne — patrz zadania VIII.13 i VIII.14.

Oczywiście funkcje stałe są zawsze ciągłe. W przypadku funkcji wielu zmiennych ważny przykład funkcji ciągłych to funkcje „współrzędne", tzn. $\mathbb{x}_j : \mathbb{R}^d \to \mathbb{R}$ dla $j = 1, \ldots, d$ zadane dla $x \in \mathbb{R}^d$ wzorem

$$\mathbb{x}_j(x) = x_j$$

(ich ciągłość wynika np. z faktu VIII.3 str. 119).

Oczywiste jest też, że obcięcie funkcji ciągłej do podzbioru jej dziedziny jest funkcją ciągłą.

Aby móc w praktyce sprawdzać ciągłość rozmaitych funkcji, z którymi będziemy mieli do czynienia, bez konieczności każdorazowego odwoływania się do definicji ciągłości, wygodnie będzie nam używać poniższych dwóch faktów. Pierwszy z nich dotyczy funkcji pomiędzy dowolnymi przestrzeniami metrycznymi i jest natychmiastową konsekwencją definicji (Heinego).

Fakt VIII.1. *Złożenie funkcji ciągłych jest funkcją ciągłą.*

Zanim sformułujemy kolejny wynik przyjmijmy następujące oznaczenie dotyczące funkcji o wartościach w $\mathbb{R}^d$. Jeżeli $f : X \to \mathbb{R}^d$ to dla $j = 1, \ldots, d$ przez f_j będziemy oznaczać „j-tą funkcję współrzędną" funkcji f, tj. funkcję z X w $\mathbb{R}$ zadaną dla $x \in X$ wzorem

$$f_j(x) = (f(x))_j,$$

a zatem $f(x) = (f_1(x), \ldots, f_d(x))$.

Fakt VIII.2.
(i) Funkcja $f : X \to \mathbb{R}^d$ jest ciągła wtw dla każdego $j = 1, \ldots, d$ funkcja f_j jest ciągła.
(ii) Suma, różnica, iloczyn, iloraz[128)] *funkcji ciągłych o wartościach rzeczywistych jest funkcją ciągłą.*

Dowód.
Część (i) wynika natychmiast z faktu VIII.3 str. 119, a część (ii) z twierdzenia o rachunkowych własnościach granicy (dla ciągów liczbowych — tw. II.1). □

[128)] Oczywiście przy założeniu, że iloraz ten ma sens, tzn., że nie występuje „dzielenie przez 0".

Przykład. Funkcja $f : \mathbb{R}^3 \to \mathbb{R}^2$ zadana wzorem

$$f(x) = (x_1x_2 + x_3, (|x_1| + 1)^{(x_2-x_3)})$$

jest ciągła. Mamy bowiem $f_1 = \mathbb{x}_1 \cdot \mathbb{x}_2 + \mathbb{x}_3$, $f_2 = \exp \circ((\mathbb{x}_2 - \mathbb{x}_3) \cdot (g \circ \mathbb{x}_1))$, gdzie $g : \mathbb{R} \to \mathbb{R}$ zadana jest wzorem $g(t) = \ln(|t|+1)$, czyli ciągłość f wynika z faktu 1 i faktu 2 oraz z ciągłości funkcji exp, g i $\mathbb{x}_j$.

Okazuje się, że ciągłość jest bardzo blisko związana z pojęciem otwartości (także domkniętości) zbiorów. Mianowicie dla funkcji określonych na przestrzeni metrycznej ciągłość oznacza dokładnie zachowywanie otwartości (równoważnie — domkniętości) zbiorów przy braniu przeciwobrazów zbiorów. Przypomnijmy tu, że dla $f : D \to Y$ *przeciwobraz* zbioru $A \subset Y$ względem funkcji f oznaczany jest przez $f^{-1}(A)$ (uwaga! nie należy tego mylić z analogicznie oznaczanym obrazem A względem funkcji odwrotnej do f, która może nawet w ogóle nie istnieć...) oraz

$$f^{-1}(A) := \{x \in D : f(x) \in A\}.$$

Twierdzenie VIII.3. *Dla funkcji $f : D \to Y$ następujące warunki są równoważne:*
(i) f jest ciągła;
(ii) dla dowolnego zbioru otwartego U w Y istnieje zbiór otwarty V w X taki, że $f^{-1}(U) = D \cap V$;
(iii) dla dowolnego zbioru domkniętego F w Y istnieje zbiór domknięty H w X taki, że $f^{-1}(F) = D \cap H$.[129)]

Nie będziemy tu dowodzić implikacji (ii) $\Rightarrow$ (i) ani (iii) $\Rightarrow$ (i). Ograniczymy się do dowodu (i) $\Rightarrow$ (ii) skąd (i) $\Rightarrow$ (iii) wynika natychmiast ze wzorów de Morgana oraz zachowania działań na zbiorach przy braniu przeciwobrazu względem funkcji.

Dowód ((i) $\Rightarrow$ (ii)).
Zakładamy, że f — ciągła i niech U — otwarty w Y. Dla dowolnego $x \in f^{-1}(U) \subset D$ mamy $f(x) \in U$ zatem wybierzmy $\epsilon_x > 0$ taki, że

$$K(f(x), \epsilon_x) \subset U.$$

Stąd korzystając z ciągłości f w x wybierzmy takie $\delta_x > 0$, że dla dowolnego $x' \in D$ jeżeli $\rho_X(x, x') < \delta_x$, to $f(x') \in U$. Powyższe oznacza dokładnie, że

$$K(x, \delta_x) \cap D \subset f^{-1}(U). \tag{VIII.5}$$

Niech zatem

$$V := \bigcup_{x \in f^{-1}(U)} K(x, \delta_x)$$

— jest to zbiór otwarty w X jako suma kul otwartych. Ponadto mamy $f^{-1}(U) \subset V \cap D$, bo dla dow. $x \in f^{-1}(U)$ oczywiście $x \in K(x, \delta_x)$. Z drugiej strony na mocy (VIII.5) mamy też

$$V \cap D = \bigcup_{x \in f^{-1}(U)} (K(x, \delta_x) \cap D) \subset f^{-1}(U).$$

$\square$

Wniosek. *Niech $f : D \to Y$ będzie funkcją ciągłą. Wówczas*
a) jeśli D jest otwarty w X oraz U otwarty w Y, to $f^{-1}(U)$ jest otwarty w X;
b) jeżeli D jest domknięty w X oraz F domknięty w Y, to $f^{-1}(F)$ jest domknięty w X.

[129)] Zbiory postaci $D \cap V$ i $D \cap H$ gdzie V — otwarty w X, H — domknięty w X, to dokładnie wszystkie zbiory otwarte lub — odpowiednio — domknięte w przestrzeni metrycznej D z metryką ρ_X ograniczoną do D. A zatem w punkcie (ii) można po prostu mówić o otwartości $f^{-1}(U)$ **w** D a w (iii) o domkniętości $f^{-1}(F)$ **w** D.

Dowód.
Wynika to natychmiast z tw. VIII.1 oraz z tego, że przecięcie dwóch zbiorów otwartych (domkniętych) jest zbiorem otwartym (domkniętym). □

Powyższy wniosek przydaje się często jako szybki sposób dowodzenia otwartości lub domkniętości zbiorów zadanych przy pomocy pewnych równości bądź nierówności.

Przykład. Zbiór $A_1 = \{x \in \mathbb{R}^2 \setminus \{(0,1)\} : x_1^2 < x_2\}$ jest otwarty (w $\mathbb{R}^2$) ponieważ $A_1 = f^{-1}(U)$ dla $f : D \to \mathbb{R}$, gdzie $D = \mathbb{R}^2 \setminus \{(0,1)\}$, $f(x) = x_1^2 - x_2$, $U = (-\infty; 0)$.

Zbiór $A_2 = \{x \in \mathbb{R}^2 : x_1 \geqslant 0 \text{ i } x_1^2 - x_2^2 = 7\}$ jest domknięty (w $\mathbb{R}^2$) ponieważ $A_2 = f^{-1}(F_1) \cap g^{-1}(F_2)$, gdzie $f, g : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = x_1$, $g(x) = x_1^2 - x_2^2$, $F_1 := [0; +\infty)$, $F_2 = \{7\}$.

Jednym z ważnych zadań, którymi się wkrótce zajmiemy będzie znajdowanie kresów funkcji wielu zmiennych (będzie to też jeden z celów rachunku różniczkowego wielu zmiennych, który jest tematem następnego rozdziału). Podobnie jak było to w przypadku jednej zmiennej, zadanie takie jest stosunkowo łatwe, gdy funkcja swój kres osiąga. Stąd niezwykle ważną rolę pełni następujące uogólnienie twierdzenia IV.10.

Twierdzenie VIII.4 (Weierstrassa, o osiąganiu kresów). *Jeżeli $K \subset \mathbb{R}^d$ jest zbiorem zwartym oraz $f : K \to \mathbb{R}$ jest ciągła, to f osiąga swoje kresy, tzn. istnieją $m, M \in K$ takie, że $f(m) = \inf_{x \in K} f(x)$ oraz $f(M) = \sup_{x \in K} f(x)$. W szczególności f jest ograniczona.*

Dowód.
Dowód ten jest analogiczny do dowodu tw. IV.10 — bierzemy $x_n \in K$ takie, że $f(x_n) \to \inf_{x \in K} f(x)$ i korzystając z wniosku ze strony 120 [130] wybieramy podciąg $\{x'_n\}$ ciągu $\{x_n\}$ taki, że $x'_n \to m$ dla pewnego $m \in K$ mamy zatem $f(x'_n) \to f(m)$ z ciągłości f, a jednocześnie $f(x'_n) \to \inf_{x \in K} f(x)$ skąd $f(m) = \inf_{x \in K} f(x)$. Analogicznie postępujemy dla sup. □

Na zakończenie tego rozdziału sformułujmy definicję *ekstremów lokalnych* dla ogólnej sytuacji funkcji skalarnej f określonej na podzbiorze D przestrzeni metrycznej X.

Definicja. *Funkcja $f : D \to \mathbb{R}$ posiada w $a \in D$* ***maksimum (minimum)*** *lokalne wtw*

$$\exists_{r>0} \forall_{x \in K(a,r) \cap D} \; f(x) \leqslant f(a) \quad (f(x) \geqslant f(a)).$$

Jest to zatem całkiem naturalne uogólnienie pojęcia, które znaliśmy dla funkcji zmiennej rzeczywistej. Badanie ekstremów lokalnych będzie jednym z naszych głównych celów w następnym rozdziale.

[130] Jak widać z tego dowodu twierdzenie to można uogólnić na funkcje ciągłe określone na zbiorach zwartych w dowolnych przestrzeniach metrycznych — patrz przypis do przywołanego tu wniosku.

Zadania do Rozdziału VIII

$\forall$ 1. [131] Wykaż, że określona w przykładzie 3 ze strony 117 metryka kolejowa spełnia warunki metryki. Naszkicuj kulę o środku a i promieniu 1 w przypadku gdy a) $a = w$; b) $a \neq w$.

$\forall$ 2. Wykaż, że określona w przykładzie 4 ze strony 117 metryka dyskretna spełnia warunki metryki. Opisz $K(a, r)$ w przestrzeni metrycznej z tą metryką w zależności od r. Wykaż, że każdy podzbiór jest w tej przestrzeni metrycznej otwarty i domknięty zarazem.

$\forall$ 3. Wykaż, że w dowolnej przestrzeni metrycznej każda kula otwarta o promieniu $r > 0$ jest zbiorem otwartym, a każdy zbiór jednopunktowy zbiorem domkniętym.

4. Wykaż, że podzbiór U przestrzeni metrycznej jest otwarty wtw U jest sumą pewnej rodziny kul otwartych w tej przestrzeni.

$\forall$ 5. Znajdź przykład pokazujący, że a) suma dowolnej rodziny zbiorów domkniętych nie musi być zbiorem domkniętym; b) przecięcie dowolnej rodziny zbiorów otwartych nie musi być zbiorem otwartym.

6. Przy użyciu wzorów de Morgana (dla rachunku zbiorów) uzupełnij szczegółami dowód drugiej części faktu VIII.2 ze strony 118.

$\forall$ 7. Wykaż, że w przestrzeni metrycznej każdy ciąg zbieżny jest ograniczony.

$\forall$ 8. Zbadaj zbieżność (i ew. znajdź granicę) ciągu o wyrazach $(\sqrt[n]{n^{100} + 7^n - 6^n}, (1 - \frac{1}{n})^{(n^2)}) \in \mathbb{R}^2$ w następujących przestrzeniach metrycznych

   (a) $\mathbb{R}^2$ (czyli „zwykłej $\mathbb{R}^2$”, czyli $\mathbb{R}^2$ z metryką euklidesową);
   (b) $\mathbb{R}^2$ z metryką miejską;
   (c) $\mathbb{R}^2$ z metryką kolejową z węzłem $w = (0, 0)$;
   (d) $\mathbb{R}^2$ z metryką kolejową z węzłem $w = (7, 0)$;
   (e) $\mathbb{R}^2$ z metryką dyskretną.

$\forall$ 9. Wykaż fakt VIII.5 ze str. 119, tzn. równoważność definicji domkniętości i „ciągowej definicji” domkniętości.

10. Wykaż, że jeżeli $K \subset \mathbb{R}^d$ jest taki, że każdy ciąg o wyrazach w K posiada podciąg zbieżny do elementu z K, to K jest domknięty i ograniczony (a więc zwarty).

11. Niech X będzie przestrzenią metryczną z metryką ρ i niech $a \in X$. Określmy $g : X \to \mathbb{R}$ wzorem $g(x) = \rho(a, x)$. Czy g musi być ciągła?

$\forall$ 12. [132] Znajdź wszystkie punkty skupienia poniższych zbiorów w $\mathbb{R}^d$:

   (a) $\{x \in \mathbb{R}^2 : x_1, x_2 \in \mathbb{Q}\}$;
   (b) $\{x \in \mathbb{R}^2 : x_1 \in \mathbb{Z}\}$;
   (c) $\{x \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 < 1\}$;
   (d) $\{x \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 > 1\}$;
   (e) $\{(\frac{1}{n}, \sqrt[n]{n}, \sqrt[n]{n!}) \in \mathbb{R}^3 : n \in \mathbb{N}\}$.

---

[131] Jako domowe.

[132] Przynajmniej 2 przykłady.

13. Wykaż, że jeżeli $a \in \mathbb{R}^2$, $D \subset \mathbb{R}^2$ i istnieje $r > 0$ takie, że $K(a,r) \setminus \{a\} \subset D$ oraz $f : D \to \mathbb{R}$ i $\lim_{x\to a} f(x) = c$, dla x_1 d.b. a_1 istnieje $\lim_{x_2\to a_2} f(x_1,x_2)$ oraz dla x_2 d.b. a_2 istnieje $\lim_{x_1\to a_1} f(x_1,x_2)$, to

$$\lim_{x_2\to a_2}\lim_{x_1\to a_1} f(x_1,x_2) = \lim_{x_1\to a_1}\lim_{x_2\to a_2} f(x_1,x_2) = c.$$

∀ 14. Zbadaj istnienie i ew. wartość granicy oraz obu granic iterowanych funkcji f w punkcie $(0,0)$ dla poniższych przykładów $f : \mathbb{R}^2 \to \mathbb{R}$

(a) $f(x) = \begin{cases} \frac{x_1x_2}{x_1^2+x_2^2} & \text{dla } x \neq (0,0) \\ 0 & \text{dla } x = (0,0); \end{cases}$

(b) $f(x) = x_1 \cdot \mathbb{D}(x_2)$, gdzie $\mathbb{D}$ oznacza funkcję Dirichleta.

∀ 15. [133] Wyjaśnij szczegółowo w oparciu o odpowiednie fakty z wykładu, dlaczego funkcje zadane poniższymi wzorami są ciągłe:

(a) $\left(\log_{(x_1^2+2)}(x_2^2+2), (x_1^2+1)^{\left((x_2^2+1)^{x_1}\right)}\right)$, $x \in \mathbb{R}^2$;

(b) $(x_1+x_2+x_3, x_1\cdot x_2\cdot x_3, x_1^{(x_2^{x_3})}, \log_{x_1}(x_2+x_3), 0)$, $x \in (1;+\infty)^3$.

∀ 16. [134] Zbadaj ciągłość funkcji $f : \mathbb{R}^d \to \mathbb{R}$ zadanej wzorem

$$f(x) = \begin{cases} w(x) & \text{dla } x \neq 0 \\ 0 & \text{dla } x = 0, \end{cases}$$

gdy $w(x)$ dla $x \neq 0$ zadane jest wzorem

(a) $\frac{x_1^2x_2^2}{x_1^2x_2^2+(x_1-x_2)^2}$, $\quad d = 2$;

(b) $\frac{x_1x_3+x_2x_3}{x_1^2+x_2^2+x_3^2}$, $\quad d = 3$;

(c) $\frac{x_1^2x_2}{x_1^2x_2^2+x_1^2+x_2^2}$, $\quad d = 2$;

(d) $\frac{x_1^2x_2}{x_1^4+x_2^2}$, $\quad d = 2$.

17. Wykaż, że każda funkcja $f : X \to Y$ jest ciągła, gdy w X rozważamy metrykę dyskretną.

∀ 18. [135] Dla każdego z poniższych podzbiorów $\mathbb{R}^d$ rozstrzygnij czy jest on domknięty, otwarty, ograniczony, zwarty (w dwóch pierwszych sprawach użyj najlepiej wniosku ze strony 122 oraz faktu VIII.1 ze strony 118)

(a) $\{x \in \mathbb{R}^2 : x_1 > x_2\}$;

(b) $\{x \in \mathbb{R}^2 : 2x_1 + 3 \geqslant e^{(x_1-x_2)}\}$;

(c) $\{x \in \mathbb{R}^2 : e^{(x_1+x_2^2)} = \ln \frac{1}{1+x_1^2+x_2^2}\}$;

(d) $\{x \in \mathbb{R}^3 : \sin(\sin(\cos(x_1\cdot x_2))) = \sin(\sin(x_1+x_2))\cdot x_2;\ x_1 \geqslant -1,\ x_2 \leqslant 1\}$;

(e) $\{x \in \mathbb{R}^2 \setminus \{(0,1),(1,0)\} : \frac{1}{(x_1-1)^2+x_2^2} < \frac{1}{(x_2-1)^2+x_1^2},\ x_1 < x_2\}$.

∀ 19. Niech $A = \{(x,y) \in \mathbb{R}^2 : x \neq 0, y = \frac{1}{x}\}$. Udowodnij, że A jest domknięty korzystając z „ciągowej definicji” domkniętości (fakt VIII.5 str. 119). Wyjaśnij dlaczego nie można tu w sposób bezpośredni wykorzystać wniosku ze str. 122.

---

[133] Przynajmniej 1 przykład.

[134] Przynajmniej 2 przykłady, w tym c).

[135] Przynajmniej 2 przykłady.

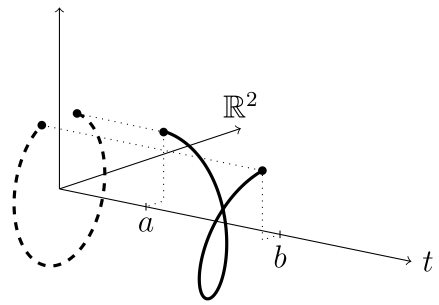
IX Rachunek różniczkowy funkcji wielu zmiennych

[około 5 wykładów]

W tym rozdziale zobaczymy, w jaki sposób można uogólnić pojęcie pochodnej funkcji jednej zmiennej o wartościach rzeczywistych na przypadek funkcji wielu zmiennych o wartościach wektorowych, tzn. funkcji z $D \subset \mathbb{R}^m$ w $\mathbb{R}^k$. Sporo uwagi poświęcimy funkcjom skalarnym (tj. $k = 1$) — a szczególnie problemom związanym ze znajdowaniem ich ekstremów. Wbrew tytułowi rozdziału zaczniemy jednak od sytuacji niejako przeciwnej, mianowicie od funkcji jednej zmiennej ($m = 1$) ale o wartościach wektorowych ($k \geqslant 1$).

1. Pochodna funkcji wektorowej 1-nej zmiennej

Funkcje zmiennej rzeczywistej o wartościach wektorowych to obiekty, z którymi miewamy do czynienia bardzo często w różnych zastosowaniach matematyki — szczególnie w fizyce. Wówczas zmienna to najczęściej czas „t", a wartość funkcji to np. położenie poruszającego się punktu w przestrzeni — wtedy $f(t) \in \mathbb{R}^d$, gdzie $d = 3$ (ewentualnie też 2 lub 1, gdy chodzi o ruch płaski lub na prostej). Gdy będziemy chcieli opisać ruch nie jednego, ale n punktów, wartości funkcji będą już wektorami z $\mathbb{R}^{n \cdot d}$. W przypadku, gdy dziedziną f jest przedział domknięty $[a; b]$ i f jest ciągła, funkcja taka bywa często nazywana **krzywą**. Jednak nie należy mylić obrazu $f([a; b])$ funkcji f, który też potocznie bywa nazywany „krzywą", z samą funkcję f, która jest oczywiście czymś „więcej" — dwie różne funkcje (krzywe) mogą mieć ten sam obraz. Jak więc wyobrażać sobie taką funkcję? Gdy $d = 2$, można ewentualnie posługiwać się wykresem (patrz rysunek 24).



Rysunek 24: Wykres krzywej (linia ciągła) i jej obraz (linia przerywana).

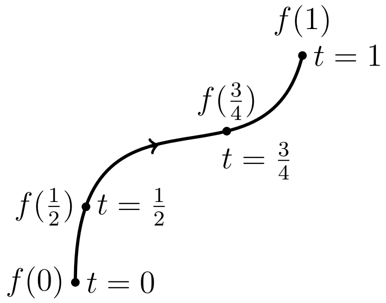
Inny, nieco prostszy sposób, to traktowanie f jako „ruchu wzdłuż jej obrazu" (patrz rysunek 25).

Uogólnienie pojęcia pochodnej na tego typu funkcje jest sprawą bardzo prostą. Niech $f : D \to \mathbb{R}^k$, czyli $f(x) = (f_1(x), \ldots, f_k(x))$, gdzie wszystkie funkcje współrzędne $f_j : D \to \mathbb{R}$ to już dobrze nam znane funkcje skalarne jednej zmiennej oraz niech $a \in D \subset \mathbb{R}$. Wówczas, jak łatwo sie domyślić, f jest *różniczkowalna* w a wtw każda ze współrzędnych funkcji f_j jest różniczkowalna w a i gdy tak jest, to *pochodną* f *w punkcie* a nazywamy wektor

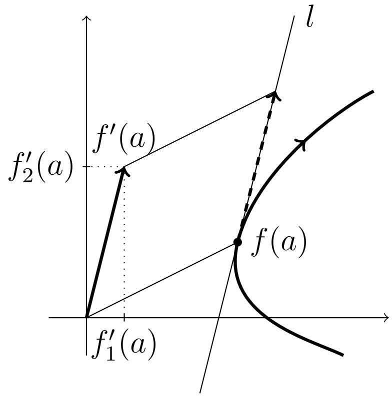
$$(f_1'(a), \ldots, f_k'(a)),$$

który oznaczamy $f'(a)$, czyli tak samo jak w przypadku funkcji skalarnych. Inne oznaczenia to tradycyjne $\frac{df}{dx}(a)$ oraz $\dot{f}(a)$ (to ostatnie szczególnie często używane jest w sytuacjach fizycznych, gdy zmienna ma interpretację czasu). A więc pochodna w punkcie jest wektorem z $\mathbb{R}^k$.

Tak rozumiana pochodna ma swój dobrze chyba znany sens geometryczno-fizyczny. Mianowicie po „zaczepieniu" wektora $f'(a)$ w punkcie $f(a)$ uzyskamy punkt leżący na tzw. *prostej stycznej* do *trajektorii* — czyli do obrazu f — w punkcie $f(a)$. Ponadto długość tego wektora,



Rysunek 25:



Rysunek 26: Wektor styczny i prosta styczna

czyli $\|f'(a)\|$ to skalarna prędkość poruszania się wzdłuż trajektorii (czyli ta wartość, którą odczytujemy na liczniku, gdy poruszamy się np. autem i $f(t)$ oznacza np. nasze położenie w chwili t).

Sam wektor $f'(a)$ to tzw. „wektor prędkości". A zatem, jeśli tylko $f'(a) \neq 0$, to l — prosta styczna wspomniana wyżej jest zbiorem punktów zadanym następująco:

$$l = \{f(a) + tf'(a) \in \mathbb{R}^k : t \in \mathbb{R}\}.$$

Można oczywiście pytać o to jakie twierdzenia dotyczące pochodnej przenoszą się z przypadku skalarnego na przypadek wektorowy. Zauważmy jednak, że bardzo wiele własności znanych nam dla funkcji skalarnych po prostu nie ma sensu w przypadku $k > 1$. Tak np. jest z monotonicznością, czy z pojęciem ekstremum lokalnego. Ale czy jest np. prawdziwe twierdzenie Lagrange'a o wartości średniej w wersji „wektorowej" (tu zarówno założenia, jak i teza mają sens)? Moglibyśmy odpowiedniego punktu „pośredniego" c szukać osobno dla każdej spośród funkcji współrzędnych f_j. Do nich przecież twierdzenie Lagrange'a stosuje się, jako że są one skalarne. Jednak nie ma żadnego powodu by to c było takie samo dla wszystkich j ... Nie powinno więc dziwić, że na to pytanie odpowiedź jest negatywna (choć nie był to dowód, jednak odpowiedni przykład nietrudno znaleźć — patrz zad. IX.2). Na szczęście, można udowodnić rozmaite „namiastki" tw. Lagrange'a, które pozwalają uzyskiwać różne oszacowania dotyczące przyrostu funkcji na podstawie oszacowań „wielkości" pochodnej. Oto jedna z nich.

Fakt. *Jeżeli* $f : [a; b] \to \mathbb{R}^k$ *jest klasy* C^1 [136] *, to*

$$\|f(b) - f(a)\| \leqslant \int_{[a;b]} \|f'(t)\| dt \leqslant (b-a) \cdot \sup_{t \in [a;b]} \|f'(t)\|.$$

Aby to udowodnić posłużmy się lematem. Najpierw oznaczenie: dla $\varphi : [a; b] \to \mathbb{R}^k$ takiej, że $\varphi_j \in \Re$ dla wszystkich $j = 1, \ldots, k$ przyjmujemy

$$\int_{[a;b]} \varphi(t) dt := \left(\int_{[a;b]} \varphi_1(t) dt, \ldots, \int_{[a;b]} \varphi_k(t) dt\right).$$

Lemat. *Dla* φ *j.w. zachodzi*

$$\left\|\int_{[a;b]} \varphi(t) dt\right\| \leqslant \int_{[a;b]} \|\varphi(t)\| dt.$$

B.D.

[136] Tzn. każda z funkcji f_j jest klasy C^1.

Dowód (faktu).
Na mocy twierdzenia VII.4 mamy $f_j(b) - f_i(a) = \int_{[a;b]} f_j'(t)dt$ dla $j = 1, \ldots, m$, skąd

$$f(b) - f(a) = \int_{[a;b]} f'(t)dt,$$

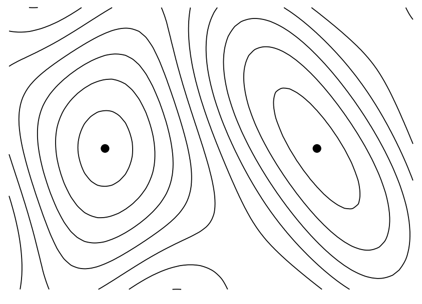
a zatem pierwsza nierówność wynika z lematu, a druga z własności „monotoniczności" całki Riemanna. $\square$

Zauważmy, że pierwsza nierówność z powyższego faktu ma bardzo dobrze zrozumiały sens fizyczny. Całka $\int_{[a;b]} \|f'(t)\|dt$ to bowiem po prostu „długość trasy" przebytej przez punkt poruszający się w sposób opisany funkcją f („droga = prędkość »skalarna«$\cdot$ czas", gdy ta prędkość skalarna jest stała, ogólnie jednak trzeba użyć całki ...). A jest przecież dość oczywiste, że długość przebytej trasy jest nie mniejsza niż odległość położenia końcowego $f(b)$ od początkowego $f(a)$, czyli właśnie $\|f(b) - f(a)\|$.

2. Metody różniczkowania funkcji wielu zmiennych

Dziedziny rozważanych tu funkcji będą podzbiorami $\mathbb{R}^m$, wprowadzimy więc dla niech kilka pojęć i oznaczeń analogicznych niektórym wcześniej już poznanym pojęciom dotyczącym podzbiorów $\mathbb{R}$. Jeśli $D \subset \mathbb{R}^m$, to a jest *punktem wewnętrznym* D wtw $K(a, \epsilon) \subset D$ dla pewnego $\epsilon > 0$. Zbiór wszystkich punktów wewnętrznych zbioru D nazywamy *wnętrzem* D i oznaczamy $\operatorname{Int} D$. *Otoczeniem* punktu $a \in \mathbb{R}^m$ nazywamy dowolny otwarty podzbiór $U \subset \mathbb{R}^m$ taki, że $a \in U$. Odcinek łączący punkt a z punktem b, tzn. zbiór $\{a + t(b - a) : t \in [0; 1]\}$ oznaczać będziemy przez $[a; b]$ [137).

Jak wyobrażać sobie funkcje wielu zmiennych? Jeśli ograniczymy się do przypadku funkcji skalarnej dwóch zmiennych, można odwołać się do przedstawień graficznych. Jedna możliwość to wykres, będący w tym przypadku czymś w rodzaju 2 wymiarowej powierzchni w $\mathbb{R}^3$ (oczywiście dla odpowiednio „regularnych" funkcji), albo „mapy plastycznej", w której wysokość n.p.m., odpowiada wartości funkcji w punkcie z płaszczyzny $\mathbb{R}^2$ będącym rzutem na „podstawę" mapy. Inny — płaski sposób — to odwołanie się do mapy poziomic funkcji, czyli obrazu z zaznaczonymi zbiorami, na których funkcja przyjmuje pewne ustalone wartości (rys. 25).



Rysunek 27:

2.1. Pochodne cząstkowe

W tym podrozdziale powiemy o najprostszym chyba sposobie różniczkowania funkcji wielu zmiennych. Polega on na tym, że jedną ze zmiennych traktujemy jako zmienną po której

[137) Niestety w przypadku $m = 1$ mamy tu pewną kolizję oznaczeń, bowiem np. $[1; 0]$ to wg. wcześniejszych oznaczeń $\emptyset$, a nie odcinek łączący liczby 0 i 1. Stosowniejsze byłoby więc używane wcześniej w $\mathbb{R}$ oznaczenie $[a?b]$. Jednak liczę, że uda się uniknąć nieporozumień — to oznaczenie stosować będziemy jedynie dla konkretnych $m > 1$ lub dla „ogólnych" m.

będziemy różniczkować, a pozostałe traktujemy jak parametry (stałe). Niech więc $f : D \to \mathbb{R}$, gdzie $D \subset \mathbb{R}^m$ i $a \in D$. Oznaczmy $D_a^j := \{t \in \mathbb{R} : (a_1, \dots t_j \dots, a_m) \in D\}$ [138] oraz

$$f_a^j : D_a^j \to \mathbb{R}, \qquad f_a^j(t) = f(a_1, \dots, t_j \dots, a_m).$$

W szczególności $a_j \in D_a^j$. Jeżeli f_a^j posiada pochodną w punkcie a_j, to nazywamy ją *pochodną cząstkową f w punkcie a po j-tej zmiennej* (ew. po „x_j”, lub inaczej, w zależności od przyjętych oznaczeń zmiennych). Tradycyjne oznaczenie to

$$\frac{\partial f}{\partial x_j}(a),$$

choć zamiast $\frac{\partial f}{\partial x_j}$ będziemy często używać też krótszych oznaczeń: $\partial_{x_j} f$ lub po prostu $\partial_j f$. Tak więc $\frac{\partial f}{\partial x_j}(a) = \partial_{x_j} f(a) = \partial_j f(a) = (f_a^j)'(a_j)$.

Przykład. Niech $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) = e^{x_1} \cdot x_2^3$. Wówczas $\partial_1 f(x) = \frac{\partial f}{\partial x_1}(x) = e^{x_1} \cdot x_2^3$, $\partial_2 f(x) = \frac{\partial f}{\partial x_2}(x) = 3e^{x_1} x_2^2$.

Oczywiście, jeżeli funkcja f posiada skończoną pochodną cząstkową po x_j w każdym punkcie a podzbioru $\tilde{D} \subset D$, to możemy rozważać funkcję $\partial_j f = \frac{\partial f}{\partial x_j} : \tilde{D} \to \mathbb{R}$. Zwróćmy uwagę na to, że mimo iż w definicji pochodnej cząstkowej różniczkuje się „po jednej zmiennej”, to $\partial_j f$ jest funkcją m–zmiennych, tak samo jak funkcja f (nie należy mylić funkcji $\partial_j f$ z $(f_a^j)'$, argumentem funkcji $\partial_j f$ jest bowiem punkt $a \in \mathbb{R}^m$, w którym ta pochodna cząstkowa jest liczona, a nie tylko jego j-ta współrzędna a_j).

Przykład. Niech

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x) = \begin{cases} 1 & \text{gdy } x_1 = 0 \text{ lub } x_2 = 0 \\ 0 & \text{wpp.}^{\,139)} \end{cases}.$$

A zatem funkcja ta nie jest ciągła w punkcie 0 ($= (0,0)$). Jednak oczywiście $\partial_1 f(0) = \partial_2 f(0) = 0$. A zatem z istnienia i skończoności nawet obu pochodnych cząstkowych f w jakimś punkcie nie wynika wcale ciągłość f w tym punkcie ...

Jak więc widać z powyższego przykładu, różniczkowanie cząstkowe może nie być zbyt dobrym analogiem różniczkowania funkcji jednej zmiennej ... Mimo to jednak okazuje się ono całkiem wystarczającym narzędziem w pewnych sytuacjach. Tak jest np. dla warunku koniecznego na ekstrema (tzn. maksima lub minima) lokalne.

Twierdzenie IX.1 (o ekstremach lokalnych). *Jeżeli $a \in \operatorname{Int} D$ oraz $f : D \to \mathbb{R}$ posiada w punkcie a ekstremum lokalne i istnieje $\partial_j f(a)$, to $\partial_j f(a) = 0$.*

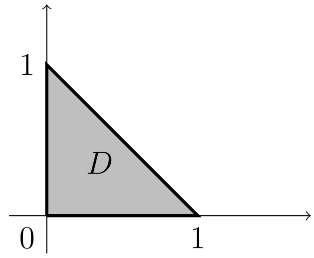
Dowód.
Z założeń twierdzenia wynika, że dla wszystkich $i = 1, \dots, m$ liczba a_i jest punktem wewnętrznym D_a^i oraz f_a^i posiada ekstremum lokalne w a_i. Tak jest więc w szczególności dla $i = j$, ale f_a^j posiada pochodną w a_j i $(f_a^j)'(a_j) = \partial_j f(a)$. Zatem na mocy twierdzenia o ekstremach lokalnych dla funkcji jednej zmiennej (tw. V.2) mamy $\partial_j f(a) = 0$. □

Podobnie jak w przypadku funkcji jednej zmiennej, twierdzenie to pozwala w wielu sytuacjach znajdować kresy funkcji.

Przykład. Niech $f(x,y) = xy(1-x-y)$ i niech $D = \{(x,y) \in \mathbb{R}^2 : x, y \geqslant 0, x+y \leqslant 1\}$ (patrz rysunek 28). Znajdziemy kresy funkcji f rozważanej w zbiorze D. Zauważmy najpierw, że na mocy wniosku ze strony 122 zbiór D jest domknięty i oczywiście jest ograniczony, a zatem jest zwarty. Ponadto f jest ciągła, zatem oba kresy są osiągane w D na mocy twierdzenia

[138] Gdzie $(a_1, \dots t_j \dots, a_m)$ to punkt powstały z a przez zastąpienie a_j przez t.

Weiestrassa (tw. VIII.2). W każdym z punktów, gdzie są one osiągane funkcja f posiada w szczególności ekstremum lokalne. Taki punkt może albo należeć do $\operatorname{Int} D$, albo do „brzegu” D, tj. do $D \setminus \operatorname{Int} D$ [140]. Łatwo sprawdzić, że $\operatorname{Int} D = \{(x, y) \in \mathbb{R}^2 : x, y > 0, x + y < 1\}$.



Rysunek 28:

Na mocy twierdzenia IX.1, punkt (x, y) z wnętrza D, w którym f posiada ekstremum musi spełniać dwa warunki:

$$\partial_x f(x, y) = y(1 - x - y) - xy = 0,$$

$$\partial_y f(x, y) = x(1 - x - y) - xy = 0,$$

czyli $1 - 2x - y = 0 = 1 - 2y - x$ (bo $x, y \neq 0$), skąd $x = y = \frac{1}{3}$. Czyli kres, jeżeli osiągany jest we wnętrzu D, może być tylko w punkcie $(\frac{1}{3}, \frac{1}{3})$, w którym f ma wartość $\frac{1}{27}$. Natomiast w każdym punkcie brzegu D funkcja f ma wartość $0 < \frac{1}{27}$. Stąd kres górny nie może być osiągany na brzegu, jest więc osiągany w $(\frac{1}{3}, \frac{1}{3})$ i podobnie kres dolny nie może być osiągany w $(\frac{1}{3}, \frac{1}{3})$, jest więc osiągany na brzegu. A stąd $\sup_{x \in D} f(x) = \frac{1}{27}$, $\inf_{x \in D} f(x) = 0$.

Należy jednak zwrócić uwagę na to, że w istotny sposób skorzystaliśmy tu z wiedzy, że oba te kresy są w D osiągane!

Oczywiście dla pochodnych cząstkowych obowiązują analogiczne jak dla pochodnej jednej zmiennej własności „arytmetyczne” dotyczące sumy, różnicy, iloczynu i ilorazu funkcji. Co będzie natomiast, gdy złożymy funkcję m zmiennych f z funkcją (też ew. wielu zmiennych) wewnętrzną g o wartościach w $\mathbb{R}^m$ i założymy, że zarówno f jak i funkcje współrzędne g_j posiadają wszystkie pochodne cząstkowe odpowiednio w $g(a)$ i a? Czy to pozwala stwierdzić istnienie pochodnych cząstkowych złożenia $f \circ g$ w a? I jaki ew. wzór tu obowiązuje? Niestety, bez dodatkowych założeń, odpowiedź na pytanie o istnienie jest negatywna! To kolejny znak wskazujący na „słabość” różniczkowania cząstkowego. Wzór na pochodne cząstkowe złożenia, który można uzyskać przy dodatkowych założeniach nosi nazwę *reguły łańcuchowej*. Zajmiemy się nim dopiero przy omawianiu trzeciego z kolei sposobu różniczkowania funkcji wielu zmiennych ...

Ważną klasę stanowią te funkcje f, dla których $\partial_i f$ istnieją i mają skończone wartości w każdym punkcie dziedziny D oraz $\partial_i f$, jako funkcje określone na D, są ciągłe dla wszystkich $i = 1, \ldots, m$. Nazywamy je funkcjami klasy C^1 (lub piszemy $f \in C^1(D)$). Można (trochę nieściśle) stwierdzić, że wszystkie funkcje, które zadają się „jednolitymi” wzorami przy użyciu zwykłych działań i składania z użyciem wyłącznie różniczkowalnych funkcji „elementarnych” jednej zmiennej oraz funkcji współrzędnych $\mathbb{x}_j$ to funkcje klasy C^1. Np. $f : \mathbb{R}^3 \to \mathbb{R}$ zadana wzorem:

$$\frac{\ln(x_1^2 + x_2^2 + x_3^2 + 1) \cdot \sin(x_1 \cdot x_2 \cdot \cos(x_3))}{e^{x_1} + e^{x_2} + e^{x_3}}.$$

W szczególności, przy pewnych dodatkowych założeniach o dziedzinie D, np. dla D — otwartych, funkcje klasy C^1 są ciągłe. Dowód nie jest trudny, choć nie jest aż tak prosty jak dla 1-nej zmiennej (patrz też podrozdział 2.3).

[140] Nie jest to ogólna definicja tzw. *brzegu* zbioru!

Różniczkowanie cząstkowe można też w analogiczny sposób zdefiniować dla funkcji o wartościach wektorowych (tj. w $\mathbb{R}^d$). Należy wówczas jedynie pochodną $(f_a^j)'(a_j)$ funkcji (wektorowej tym razem) f_a^j w punkcie a_j, która pojawia się w definicji $\partial_j f(a)$, rozumieć tak jak to zostało określone w podrozdziale IX.1. A zatem wprowadzone tu definicje pochodnych cząstkowych, czy klasy C^1 oraz odpowiednie oznaczenia rozszerzamy w ten właśnie sposób na przypadek funkcji z $D \subset \mathbb{R}^m$ w $\mathbb{R}^k$. Oczywiście ma miejsce wzór $\partial_j f(a) = (\partial_j f_1(a), \ldots, \partial_j f_k(a))$.

2.2. Pochodna kierunkowa

Kolejny, nieco bardziej „zaawansowany" sposób różniczkowania funkcji wielu zmiennych, to *różniczkowanie kierunkowe*, tzn. „w kierunku ustalonego wektora". Załóżmy, że $f : D \to \mathbb{R}^k$, gdzie $D \subset \mathbb{R}^m$, natomiast o punkcie a, w którym będziemy „różniczkować", założymy dla wygody więcej niż przy okazji pochodnych cząstkowych: będziemy zakładać, że a należy do wnętrza D, tj. $a \in \operatorname{Int} D$. Ponadto niech v będzie dowolnym wektorem z $\mathbb{R}^m$ — będzie to „kierunek różniczkowania" [141]. Zauważmy, że aby opisać zachowanie się funkcji f obciętej do prostej przechodzącej przez a o kierunku wyznaczonym przez v [142] (ściślej, do przecięcia tej prostej z dziedziną D) można posłużyć się pomocniczą funkcją φ_v jednej zmiennej, zadaną wzorem

$$\varphi_v(t) := f(a + tv).$$

Ponieważ $a \in \operatorname{Int} D$, zatem φ_v jest poprawnie określona w pewnym otoczeniu $t = 0$ i jest to funkcja o wartościach w $\mathbb{R}^k$, tak jak f (oczywiście, φ_v wyznaczona jest nie tylko przez v, lecz także przez f oraz a).

Definicja. *Jeżeli istnieje $\varphi_v'(0)$, to nazywamy ją* ***pochodną kierunkową*** *funkcji f* ***w punkcie*** *a* ***w kierunku*** *v oraz oznaczamy przy pomocy dowolnego spośród symboli*

$$\frac{\partial f}{\partial \vec{v}}(a), \quad \partial_{\vec{v}} f(a) \text{ lub } \partial_v f(a).$$

Nietrudno zauważyć, że pochodna kierunkowa w kierunkach „osi współrzędnych" ma bliski związek z pochodnymi cząstkowymi. Oznaczamy przez e_j j-ty wektor bazy kanonicznej w $\mathbb{R}^m$, tj. $(e_j)_i = 1$ dla $i = j$ oraz $(e_j)_i = 0$ dla $i \neq j$.

Fakt. *Gdy $a \in \operatorname{Int} D$, wówczas o ile istnieje jedna z pośród pochodnych $\frac{\partial f}{\partial \vec{e_j}}(a)$, $\frac{\partial f}{\partial x_j}(a)$, to istnieje też druga i zachodzi*

$$\frac{\partial f}{\partial \vec{e_j}}(a) = \frac{\partial f}{\partial x_j}(a).$$

Dowód.
Mamy $\varphi_{e_j}(t) = f(a + te_j) = f(a_1, \ldots a_j + t \ldots, a_m) = f_a^j(t + a_j)$, stąd $\frac{\partial f}{\partial \vec{e_j}}(a) = \varphi_{e_j}'(0) = (f_a^j)'(a_j) = \frac{\partial f}{\partial x_j}(a)$. □

Jednak z istnienia wszystkich pochodnych cząstkowych nie wynika wcale istnienie pochodnych kierunkowych we wszystkich kierunkach.

Przykład. Niech $f : \mathbb{R}^2 \to \mathbb{R}$ będzie taka, jak w przykładzie 2 ze strony 129. Wówczas $\frac{\partial f}{\partial x_1}(0) = \frac{\partial f}{\partial x_2}(0) = 0$, ale dla $v \neq re_j$, $r \in \mathbb{R}$, $j = 1, 2$, nie istnieje skończona pochodna kierunkowa $\frac{\partial f}{\partial \vec{v}}(0)$ ponieważ φ_v jest nieciągła w 0.

A zatem rzeczywiście „różniczkowalność we wszystkich kierunkach" to coś istotnie lepszego, niż tylko „różniczkowalność cząstkowa". Okazuje się jednak, że to ciągle jeszcze dość mało ... W szczególności ta „lepsza" różniczkowalność wciąż jeszcze nie gwarantuje nawet ciągłości funkcji! Odpowiedni przykład znajdziecie Państwo w zadaniach (patrz zad. IX.6).

---

[141] Choć tak naprawdę, ta nazwa byłaby dobra gdybyśmy ograniczyli się do wektorów długości 1, które można utożsamiać z „kierunkowymi", bez tego założenia to coś więcej niż tylko kierunek ...

[142] Gdy $v = 0$, to otrzymujemy tylko sam punkt a zamiast prostej.

2.3. Różniczka

Jak widzieliśmy, wprowadzone dotąd dwie metody uogólnienia pojęcia pochodnej funkcji jednej zmiennej na przypadek wielu zmiennych, choć dość naturalne i łatwo zrozumiałe, okazały się jednak „słabe” z analitycznego punktu widzenia. Np., istnienie i skończoność w danym punkcie tak uogólnionych pochodnych nie gwarantowało nawet ciągłości funkcji w tym punkcie.

Poznamy tu pojęcie tzw. *różniczki* [143] funkcji wielu zmiennych, które można uznać za najwłaściwsze uogólnienie pochodnej funkcji jednej zmiennej — nie ma ono bowiem mankamentu wspomnianego wyżej, a ponadto teoria z nim związana zawiera uogólnienie wielu ważnych elementów teorii znanej nam z „wymiaru 1”. Niestety jednak pojęcie to jest nieco trudniejsze od dwóch poprzednio omawianych. Dlatego, aby definicja różniczki stała się bardziej zrozumiała, zacznijmy od następującego przeformułowania definicji skończonej pochodnej funkcji skalarnej 1-zmiennej: liczba p jest pochodną funkcji f w punkcie a wtw

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

W liczniku, oprócz „przyrostu funkcji” $f(a+h)-f(a)$ pojawia się wyrażenie $p\cdot h$, czyli — z punktu widzenia „przyrostu argumentów” h — wyrażenie liniowe od h. Zamiast więc myśleć o samej liczbie p (pochodnej f w punkcie a) możemy równie dobrze myśleć o funkcji liniowej[144] (inaczej: przekształceniu liniowym) $L:\mathbb{R}\to\mathbb{R}$ zadanej wzorem

$$L(h) = p\cdot h$$

— to przekształcenie liniowe L nazywamy w przypadku funkcji skalarnej *różniczką* funkcji f *w punkcie* a. Inaczej mówiąc różniczka f w punkcie a, to takie przekształcenie liniowe $L:\mathbb{R}\to\mathbb{R}$, że

$$\lim_{h\to 0} \frac{1}{|h|}[f(a+h)-f(a)-L(h)] = 0, \tag{IX.1}$$

tzn. „w pobliżu” $h=0$ wartość $L(h)$ przybliża przyrost $f(a+h)-f(a)$ z dokładnością „wyższego rzędu” niż h [145] . Warunek (IX.1) stanowi dla nas podpowiedź jak należy zdefiniować różniczkę dla funkcji wielu zmiennych.

Rozważmy $f:D\to\mathbb{R}^k$, gdzie $D\subset\mathbb{R}^m$ oraz $a\in\operatorname{Int} D$. Te założenia obowiązywać będą w całym bieżącym rozdziale.

Definicja. *Przekształcenie liniowe $L:\mathbb{R}^m\to\mathbb{R}^k$ jest* ***różniczką*** *f w punkcie a wtw*

$$\lim_{h\to 0} \frac{1}{\|h\|}[f(a+h)-f(a)-L(h)] = 0. \tag{IX.2}$$

Różniczkę f w punkcie a oznaczamy przez $Df(a)$ [146] . O funkcji f mówimy, że jest **różniczkowalna w punkcie** a wtw $Df(a)$ istnieje, oraz **różniczkowalna**, gdy jest różniczkowalna w każdym punkcie $x\in D$.

Uwagi.

[143] Używana bywa też nazwa *różniczka zupełna*, a także, po prostu, *pochodna*.

[144] Proszę nie mylić z potoczną nazwą „liniowa” oznaczającą de facto funkcję afiniczną, tj. zadaną wzorem $cx+d$.

[145] Tzn. z dokładnością do $o(h)$ przy $h\to 0$.

[146] Niektórzy oznaczają ją też $f'(a)$ zamiast $Df(a)$, co prowadzi do pewnego (niedużego) zamieszania w przypadku jednej zmiennej. Liczę, że stosowane tu przez nas oznaczenie różniczki „D” nie będzie się myliło z „D” oznaczającym tu także dziedzinę funkcji ...

1. Czasami wygodnie jest zapisać warunek (IX.2) w równoważny mu następujący sposób:

$$f(x) = f(a) + L(x-a) + r(x-a) \tag{IX.3}$$

dla pewnej funkcji r określonej w pewnym otoczeniu 0, spełniającej

$$\lim_{h\to 0} \frac{1}{\|h\|} r(h) = 0 \tag{IX.4}$$

(tak jak dla $m = 1$ powyższą równość oznaczamy w skrócie $r(h) = o(h)$ przy $h \to 0$). A zatem można powiedzieć, że L jest różniczką f w punkcie a wtedy, gdy f w pobliżu a przybliża się z dokładnością do $o(x-a)$ przez $f(a) + L(x-a)$, czyli w szczególności przez pewną funkcję afiniczną, której część liniową stanowi L.

2. By powyższe oznaczenie $Df(a)$ było sensowne, należałoby najpierw sprawdzić, że istnieje co najwyżej jedna różniczka f w punkcie a. Załóżmy więc, że L i $\tilde{L}$ są obie takimi różniczkami wówczas na mocy (IX.2) dla L i $\tilde{L}$ otrzymujemy

$$\frac{1}{\|h\|}(L-\tilde{L})(h) = \frac{1}{\|h\|}(L(h) - \tilde{L}(h)) \xrightarrow[h\to 0]{} 0$$

W szczególności dla dowolnego $x \in \mathbb{R}^m \setminus \{0\}$ możemy użyć powyższego dla h postaci tx, gdzie $t \in \mathbb{R}_+$, a zatem z liniowości L i $\tilde{L}$

$$\frac{1}{t\|x\|}(L-\tilde{L})(tx) = \frac{1}{\|x\|}(L-\tilde{L})(x) \xrightarrow[t\to 0]{} 0$$

skąd $\frac{1}{\|x\|}(L-\tilde{L})(x) = 0$ (gdyż nie zależy od t), czyli $(L-\tilde{L})(x) = 0$ dla dowolnego $x \in \mathbb{R}^m$, tzn. $L = \tilde{L}$.

3. Ponieważ $Df(a)$ jest przekształceniem (czyli funkcją) z $\mathbb{R}^m$ w $\mathbb{R}^k$, zatem trzeba zdecydować jak oznaczać jego wartość na wektorze $h \in \mathbb{R}^m$. Niestety, nie jest to zbyt wygodne, ale użyjemy w tej sytuacji oznaczenia $Df(a)(h)$ lub $(Df(a))(h)$. Tymczasem samo Df (bez „(a)") może być używane do oznaczania różniczki f traktowanej jako nowa funkcja określona w tych punktach a, w których $Df(a)$ istnieje. Wartości tej funkcji to już elementy nie $\mathbb{R}^k$, lecz zbioru przekształceń liniowych z $\mathbb{R}^m$ w $\mathbb{R}^k$!

4. Jak widzieliśmy ze wstępu przed definicją, w przypadku funkcji jednej zmiennej (oraz punktów wewnętrznych dziedziny) różniczkowalność w rozumieniu wcześniejszym (z rozdziału V) jest tym samym, co tu zdefiniowana. Ponadto dla f różniczkowalnej w a zachodzi

$$\forall_{h\in\mathbb{R}}\ Df(a)(h) = f'(a) \cdot h.$$

Przykłady.

1. Funkcja stała f jest oczywiście różniczkowalna w każdym punkcie wewnętrznym a swej dziedziny i $Df(a) = 0$ (tu 0 rozumiemy jako zerowe przekształcenie liniowe z $\mathbb{R}^m$ w $\mathbb{R}^k$ dla odp. m i k).

2. Niech $A : \mathbb{R}^m \to \mathbb{R}^k$ będzie przekształceniem liniowym. Wówczas $(DA)(x) = A$ dla każdego $x \in \mathbb{R}^m$, bowiem z liniowości A mamy

$$A(x+h) - A(x) - A(h) = 0$$

dla dowolnych $x, h \in \mathbb{R}^m$.

3. Powyższy przykład można nieco rozszerzyć. Jeżeli $f : \mathbb{R}^m \to \mathbb{R}^k$ jest przekształceniem afinicznym, tzn. zadanym wzorem $f(x) = Ax + b$, gdzie $b \in \mathbb{R}^k$ oraz A — jak wyżej, to $(Df)(x) = A$ dla dowolnego $x \in \mathbb{R}^m$.

4. Niech $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = x_1 x_2$. Spróbujmy odgadnąć jaka jest wartość $Df(a)$ dla ustalonego $a \in \mathbb{R}^2$. Mamy:

$$f(a+h) - f(a) = (a_1 + h_1)(a_2 + h_2) - a_1 a_2 = a_1 h_2 + a_2 h_1 + h_1 h_2.$$

Wiemy, że różniczka $Df(a)$ powinna być taką funkcją liniową, że po jej odjęciu od powyższego przyrostu pozostaje część „rzędu wyższego" niż h. Musimy więc przyrost ten zapisać jako sumę „części liniowej" (różniczki) i wyrażenia typu $o(h)$. W naszym przypadku naturalnym kandydatem na tę częśc liniową, czyli różniczkę, wydaje sie być składnik zadany wzorem $a_1 h_2 + a_2 h_1$. Spróbujmy zatem sprawdzić, czy funkcja liniowa $\mathbb{R}^2 \ni h \to a_1 h_2 + a_2 h_1 \in \mathbb{R}$ „nadaje się" jako $Df(a)$, tzn. sprawdźmy czy $h_1 h_2$ będące ową „pozostałą częścią przyrostu"jest typu $o(h)$. Mamy:

$$|\frac{1}{\|h\|} h_1 h_2| = \frac{|h_1 h_2|}{\sqrt{h_1^2 + h_2^2}} = |h_1| \cdot \frac{|h_2|}{\sqrt{h_1^2 + h_2^2}} \leqslant |h_1|,$$

skąd $\lim_{h \to 0} \frac{1}{\|h\|} h_1 h_2 = 0$, a zatem rzeczywiście $Df(a)(h) = a_1 h_2 + a_2 h_1$.

Tak jak zapowiadaliśmy, tak rozumiana różniczkowalność gwarantuje automatycznie ciągłość funkcji.

Twierdzenie IX.2 (o ciągłości funkcji różniczkowalnej). *Jeżeli f jest różniczkowalna w a, to f jest ciągła w a.*

Dowód.
Na mocy (IX.3), (IX.4) mamy $f(x) = f(a) + Df(a)(x-a) + r(x-a)$, gdzie $\lim_{x \to a} r(x-a) = \lim_{x \to a} \|x-a\| \left(\frac{1}{\|x-a\|} r(x-a) \right) = 0$. Dla $x \in D \setminus \{a\}$ mamy ponadto

$$\lim_{x \to a} Df(a)(x-a) = \lim_{h \to a} Df(a)(h) = Df(a)(0)) = 0,$$

gdyż $Df(a)$ jest funkcją ciągłą jako przekształcenie liniowe (wynika to natychmiast np. z faktu 2 strona 121 i znanej Państwu z GAL-u ogólnej postaci funkcji liniowych). Stąd $\lim_{x \to a} f(x) = f(a)$, czyli f jest ciągła. $\square$

Okazuje się także, że wprowadzone przez nas pojęcie różniczkowalności jest „silniejsze" niż obie nasze wcześniejsze próby takiej definicji. A ściślej, ma miejsce wynik następujący.

Fakt. *Jeżeli f jest różniczkowalna w a, to dla dowolnego $v \in \mathbb{R}^m$ pochodna kierunkowa f w punkcie a w kierunku v istnieje oraz*

$$\frac{\partial f}{\partial \vec{v}}(a) = Df(a)(v), \tag{IX.5}$$

w szczególności istnieją wszystkie pochodne cząstkowe $\frac{\partial f}{\partial x_j}(a)$ dla $j = 1, \ldots, m$ oraz

$$\frac{\partial f}{\partial x_j}(a) = Df(a)(e_j). \tag{IX.6}$$

Dowód.
Musimy wykazać, że $\varphi_v'(0)$ istnieje i jest równa $Df(a)(v)$, gdzie $\varphi_v(t) := f(a + tv)$ dla $t \in \mathbb{R}$

z pewnego otoczenia liczby 0. Gdy $v = 0$ jest to oczywiste (mamy 0 po obu stronach). Niech więc $v \neq 0$ i $t \neq 0$. Wówczas na mocy (IX.3) oraz liniowości różniczki

$$\frac{1}{t}(\varphi_v(t)-\varphi_v(0)) = \frac{1}{t}[Df(a)(tv)+r(tv)] = \frac{1}{t}[tDf(a)(v)+|t|\|v\|\frac{r(tv)}{\|tv\|}] = Df(a)(v)+\frac{|t|}{t}\|v\|\frac{r(tv)}{\|tv\|},$$

skąd $\varphi_v'(0) = Df(a)(v) + \lim_{t\to 0} \frac{|t|}{t}\|v\|\frac{r(tv)}{\|tv\|} = Df(a)(v)$, gdyż $\frac{|t|}{t}$ jest ograniczoną funkcją zmiennej $t \neq 0$ oraz zachodzi (IX.4). Dowodzi to pierwszej części faktu, a druga część wynika z pierwszej oraz z faktu ze strony 131. □

Uwaga. Z powyższego faktu wynika w szczególności, że pochodna kierunkowa w a funkcji różniczkowalnej w punkcie a zależy w sposób liniowy od kierunku różniczkowania, tzn.

$$\mathbb{R}^m \ni v \rightsquigarrow \frac{\partial f}{\partial \vec{v}}(a) \in \mathbb{R}^k$$

jest przekształceniem liniowym (równym po prostu $Df(a)$). Gdy zamiast różniczkowalności, założymy tylko, że f posiada w każdym kierunku pochodną kierunkową w punkcie a, to takiej liniowej zależności od kierunku nie mamy powodu oczekiwać.

W przypadku jednej zmiennej pochodna była obiektem bardzo prostym — liczbą. Tymczasem różniczka funkcji wielu zmiennych to obiekt dość skomplikowany — przekształcenie liniowe ... Jak się tym w praktyce posługiwać? Na szczęście, jak zapewne pamiętają Państwo jeszcze z wykładu z GAL-u, każde przekształcenie liniowe L z $\mathbb{R}^m$ w $\mathbb{R}^k$ posiada swoją bardzo wygodną reprezentację w formie znacznie chyba bardziej „namacalnej" niż samo L. Chodzi oczywiście o reprezentację macierzową — w postaci macierzy przekształcenia L. Tu, mówiąc o macierzy przekształcenia L zawsze będziemy mieli na myśli jego macierz w bazach standardowych. Taka reprezentacja pozwala na jednoznaczne zakodowanie L przy pomocy $k \cdot m$ liczb — wyrazów tej macierzy. Macierz ta ma m kolumn i k wierszy i jak wiadomo, jej j-ta kolumna utworzona jest z wektora $L(e_j)$, gdzie e_j jest j-tym wektorem bazy standardowej w $\mathbb{R}^m$. A zatem wzór (IX.6) pozwala na sformułowanie następującego wniosku.

Wniosek. *Jeżeli f jest różniczkowalna w a, to macierz przekształcenia liniowego $Df(a)$ jest taką macierzą $k \times m$, której j-ta kolumna utworzona jest z wektora $\frac{\partial f}{\partial x_j}(a)$, tzn. jej miejsce w i-tym wierszu oraz j-tej kolumnie równe jest $\frac{\partial f_i}{\partial x_j}(a)$ [147], dla $i = 1,\ldots,k$, $j = 1,\ldots,m$.*

Macierz przekształcenia $Df(a)$ nazywana bywa **macierzą Jakobiego** (f w punkcie a). Oznaczać ją będziemy $MJf(a)$. A zatem $MJf(a) = (\frac{\partial f_i}{\partial x_j}(a))_{\substack{i=1,\ldots,k \\ j=1,\ldots,m}}$, tzn.

$$MJf(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1}(a) & \cdots & \frac{\partial f_k}{\partial x_m}(a) \end{pmatrix}.$$

Gdy $m = k$, to $MJf(a)$ jest macierzą kwadratową i definiujemy **jakobian** f w punkcie a wzorem

$$Jf(a) := \det Df(a) = \det MJf(a).$$

Szczególne przypadki to:

- **k=1**, czyli f jest funkcją o wartościach liczbowych — wówczas $MJf(a)$ ma jeden wiersz, który utożsamiamy po prostu z wektorem z $\mathbb{R}^m$. Wektor ten nazywany jest **gradientem**

[147] Przypomnijmy, że f_i to i-ta funkcja współrzędna funkcji f, tzn. $f(x) = (f_1(x),\ldots,f_k(x)) \in \mathbb{R}^k$.

(f w punkcie a) i oznaczany jest najczęściej $\operatorname{grad} f(a)$, lub $\nabla f(a)$ [148]. A zatem w tym przypadku:

$$MJf(a) = \operatorname{grad} f(a) = (\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_m}(a)) \in \mathbb{R}^m \approx \mathbb{R}^{1,m} \;\; ^{149)} .$$

- **m=1**, czyli f jest funkcją jednej zmiennej (o wartościach wektorowych — w $\mathbb{R}^k$) — wówczas $MJf(a)$ ma jedną kolumnę, która jest po prostu transpozycją wektora pochodnej $f'(a)$ wprowadzonego w podrozdziale 1. A zatem wówczas

$$MJf(a) = (f'(a))^T = \begin{pmatrix} f_1'(a) \\ \vdots \\ f_k'(a) \end{pmatrix}.$$

- **m=k=1**, czyli f jest zwykłą funkcją skalarną jednej zmiennej — wówczas mamy macierz 1×1:
$$MJf(a) = (f'(a)).$$

W przypadku ogólnym można myśleć o $MJf(a)$ jako o macierzy, której kolejne wiersze to gradienty kolejnych funkcji $f_1, \ldots, f_k$ w punkcie a.

Przykład. $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x) = (x_1x_2, x_1 + x_2)$. Wówczas

$$MJf(x) = \begin{pmatrix} x_2 & x_1 \\ 1 & 1 \end{pmatrix}, \quad Jf(x) = x_2 - x_1.$$

Zauważmy, że według przyjętej przez nas definicji, by mówić o macierzy Jakobiego f w punkcie a powinniśmy zakładać, że f jest w a różniczkowalna. Jednak właściwie dla wypisania tej macierzy wystarczą nam same pochodne cząstkowe f w a, które mogą istnieć nawet bez różniczkowalności.

Inna konsekwencja wzoru (IX.6) dotyczy ekstremów lokalnych funkcji.

Wniosek. *Jeżeli $a \in \operatorname{Int} D$ oraz $f : D \to \mathbb{R}$ jest różniczkowalna w a i posiada w a ekstremum lokalne, to $Df(a) = 0$, czyli* $\operatorname{grad} f(a) = 0$.

Dowód.
Z twierdzenia o ekstremach lokalnych (tw. IX.1) i z (IX.6) mamy $\operatorname{grad} f(a) = MJf(a) = 0$, skąd $Df(a) = 0$. □

Jak widzieliśmy, praktyczne sprawdzanie dla konkretnych funkcji, że są one różniczkowalne, wcale nie jest sprawą łatwą, gdy ma się do dyspozycji wyłącznie definicję. A zatem, choć różniczkowalność jest, jak powiedzieliśmy, pojęciem „znacznie lepszym” niż istnienie pochodnych cząstkowych, za to wydaje się, że jest znacznie mniej praktyczna. Istnienie pochodnych cząstkowych sprawdzało się bowiem w „typowych” sytuacjach bardzo prosto. Na szczęście, wcale jednak nie jest tak źle, jak mogło się wydawać. W „typowych” sytuacjach można się bowiem posłużyć następującym ważnym rezultatem.

Twierdzenie IX.3 (o różniczkowalności dla klasy C^1). *Jeśli $f : D \to \mathbb{R}^k$ jest klasy C^1 oraz $a \in \operatorname{Int} D$, to f jest różniczkowalna w punkcie a.* **B.D.**

A zatem choć samo istnienie pochodnych cząstkowych do różniczkowalności nie wystarczało, to już jednak ich ciągłość tę różniczkowalność gwarantuje! W efekcie pozwala nam to bez trudu uzyskać różniczkowalność wszystkich funkcji zadanych elementarnymi wzorami:

[148] Proszę nie mylić jednak ∇ z Δ, który to znaczek używany jest m.in. do oznaczania tzw. Laplasjanu …

[149] Tym razem „$\approx$” oznacza utożsamienie $\mathbb{R}^m$ z macierzami rzeczywistymi $1 \times m$.

Przykład. Niech $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = x_1 e^{x_2}$. Mamy $\partial_1 f(x) = e^{x_2}$, $\partial_2 f(x) = x_1 e^{x_2}$ — a zatem $\partial_1 f$ i $\partial_2 f$ są ciągłe (patrz fakty 1 i 2 ze strony 121). Czyli f jest klasy C^1, co dzięki powyższemu twierdzeniu daje różniczkowalność f.

Jak należało się spodziewać, klasa funkcji różniczkowalnych posiada także w przypadku funkcji wielu zmiennych naturalne własności algebraiczne.

Twierdzenie IX.4 (o własnościach rachunkowych różniczki).

1. *Suma funkcji różniczkowalnych jest różniczkowalna. Co więcej, jeżeli $f, g : D \to \mathbb{R}^k$ są różniczkowalne w $a \in D$, to $f + g$ jest różniczkowalna w a oraz $D(f + g)(a) = D(f)(a) + D(g)(a)$.*

2. *Iloczyn funkcji różniczkowalnej skalarnej i funkcji różniczkowalnej wektorowej jest różniczkowalny. Jeśli $f : D \to \mathbb{R}$ i $g : D \to \mathbb{R}^k$ są różniczkowalne w $a \in D$, to $f \cdot g$ jest różniczkowalna w a* [150] *.*

3. *Złożenie funkcji różniczkowalnych jest różniczkowalne. Co więcej jeśli $f : D \to D' \subset \mathbb{R}^k$ jest różniczkowalna w $a \in D$ oraz $g : D' \to \mathbb{R}^l$ jest różniczkowalna w $f(a)$, to $g \circ f$ jest różniczkowalna w a oraz*

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a). \tag{IX.7}$$

Dowód powyższego twierdzenia pozostawiam jako zadanie. Najlepiej zacząć od p. 3. — patrz zadanie IX.24. Punkty 1. i 2. można dowodzić niezależnie, ale ciekawszy jest dowód z użyciem punktu 3. — patrz zadanie IX.25.

Twierdzenie to ma także swój analog dotyczący klasy C^1 (patrz zadanie IX.26). Bardzo ważnym wnioskiem z p. 3. twierdzenia IX.4 jest wzór na pochodne cząstkowe złożenia, zwany też *regułą łańcuchową*.

Wniosek. *Przy założeniach punktu 3. twierdzenia IX.4 zachodzi*

$$MJ(g \circ f)(a) = MJg(f(a)) \cdot MJf(a), \tag{IX.8}$$

gdzie „·" po prawej stronie powyżej oznacza mnożenie macierzy. W szczególności, gdy $l = 1$, ma miejsce następujący wzór

reguła łańcuchowa:

$$\frac{\partial(g \circ f)}{\partial x_s}(a) = \sum_{j=1}^{k} \frac{\partial g}{\partial y_j}(f(a)) \cdot \frac{\partial f_j}{\partial x_s}(a),$$

czyli

$$\partial_s(g \circ f)(a) = \sum_{j=1}^{k} \partial_j g(f(a)) \cdot \partial_s f_j(a), \tag{IX.9}$$

gdzie $s = 1, \ldots, m$.

Dowód.

Wzór (IX.8) to natychmiastowa konsekwencja wzoru (IX.7) oraz faktu, że macierz złożenia przekształceń liniowych, to iloczyn ich macierzy. Z kolei wzór (IX.9) otrzymamy z (IX.8) stosując wzór na wyrazy iloczynu macierzy. □

[150] Tu wzór na $D(f \cdot g)(a)$ pomijamy, ale patrz zadanie IX.25.

Podkreślmy, że na ogół nie wolno zamieniać kolejności mnożenia macierzy po prawej stronie (IX.8) (nawet gdyby miało to sens ...).

Wzór (IX.9) może być łatwiejszy do zapamiętania, gdy będziemy myśleć o funkcji wewnętrznej f jako o pewnej „zamianie zmiennych" z x na y i użyjemy nieformalnego tradycyjnego zapisu „$f_j(x) = y_j$" w związku z czym zamiast $\frac{\partial f_j}{\partial x_s}$ napiszemy „$\frac{\partial y_j}{\partial x_s}$". Jeżeli jeszcze pominiemy „oczywiste" punkty w których różniczkujemy, to otrzymamy tradycyjny zapis tego wzoru

$$\frac{\partial g}{\partial x_s} = \sum_{j=1}^{k} \frac{\partial g}{\partial y_j} \cdot \frac{\partial y_j}{\partial x_s}.$$

Traktując napisy z prawej strony jak ułamki, po prawej stronie otrzymamy właśnie to co po lewej, choć (niestety ...) już k razy, a nie tylko jeden raz ...

Przykład. Niech $g(y) = y_1 \cdot y_2^2$, $y \in \mathbb{R}^2$, $f(x) = (x^2, 3x)$, $x \in \mathbb{R}$. Mamy

$$(g \circ f)(x) = x^2 \cdot (3x)^2 = 9x^4.$$

W szczególności $(g \circ f)'(1) = 36$. Ten sam wynik możemy uzyskać stosując regułę łańcuchową:

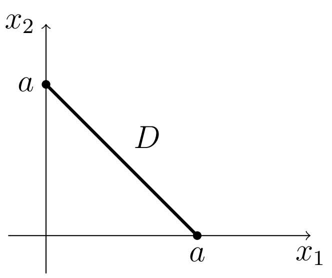
$$\begin{aligned}(g \circ f)'(1) = \frac{\partial (g \circ f)}{\partial x}(1) = \sum_{j=1}^{2} \frac{\partial g}{\partial y_j}(f(1)) \cdot \frac{\partial f_j}{\partial x}(1) \quad &= \quad y_2^2\,|_{y=(1,3)} \cdot 2 + (2y_1 \cdot y_2)\,|_{y=(1,3)} \cdot 3 \\ &= \quad 9 \cdot 2 + 2 \cdot 1 \cdot 3 \cdot 3 = 4 \cdot 9 = 36.\end{aligned}$$

3. Ekstrema warunkowe[151)]

W poprzednim podrozdziale sformułowaliśmy (i to dwukrotnie) pewne warunki konieczne na ekstremum funkcji wielu zmiennych (patrz twierdzenie IX.1 i wniosek 2 strona IX.18). Były to warunki, które (podobnie jak w przypadku funkcji jednej zmiennej) działały jedynie w przypadku, punktów wewnętrznych. W praktyce jednak często mamy do czynienia z sytuacją, gdy funkcja jest określona na zbiorze, który w ogóle nie posiada punktów wewnętrznych. Rozważmy dla przykładu następujący prosty problem:

> *Jaki jest największy możliwy iloczyn dwóch liczb nieujemnych, których suma jest równa a ($a > 0$)?*

Spróbujmy problem ten potraktować, jako zadanie dotyczące funkcji dwóch zmiennych x_1 i x_2 odpowiadających powyższym dwóm liczbom. A zatem nasza funkcja to $f : D \to \mathbb{R}$ zadana wzorem $f(x) = x_1 \cdot x_2$, gdzie $D := \{x \in \mathbb{R}^2 : x_1, x_2 \geqslant 0, x_1 + x_2 = a\}$.



Rysunek 29:

Oczywiście, pierwsza — podstawowa obserwacja jest taka, że f jest ciągła, a D — zwarty (jest to odcinek „z końcami" — patrz rys. 29), zatem dzięki twierdzeniu Weierstrassa (tw.

151) Inna nazwa *ekstrema związane*.

VIII.2) mamy pewność, że f osiąga swą największą wartość. Oznacza to, że problem powyższy jest w ogóle poprawnie postawiony (posiada rozwiązanie). Ponadto, w punkcie, w którym f osiąga tę największą wartość funkcja posiada w szczególności ekstremum lokalne. Jednak $\operatorname{Int} D = \emptyset$! Wcześniejsze kryteria są więc tu całkowicie bezużyteczne (także dlatego, że tu nie mamy pochodnych cząstkowych funkcji f mimo, że wzór na f jest „całkiem elegancki" — jednak to dziedzina jest „zła"...). Jak zatem rozwiązać ten problem? Jeden ze sposobów jest Państwu dobrze znany i w tym przypadku wiąże się z wyborem prostszego modelu matematycznego dla tego problemu: po co bowiem rozważać funkcję 2 zmiennych, skoro można było od razu użyć funkcji jednej zmiennej ...? Poniższy przykład dotyczy właśnie takiego prostego pomysłu. Opisany tam sposób potraktujemy jako ilustrację ogólnej metody postępowania w podobnych sytuacjach (na ogół bardziej złożonych).

Przykład (metoda parametryzacji). Chcemy znaleźć jakąś „parametryzację"(czyli opis) zbioru D przy pomocy mniejszej liczby zmiennych, tzn. tu funkcję ciągłą $\varphi : \tilde{D} \to D$, gdzie $\varphi(\tilde{D}) = D$ oraz $\tilde{D}$ ma już „duże" wnętrze, ale w przestrzeni niższego wymiaru (— niższego niż liczba zmiennych dla f). Dla naszego przykładu wymiar ten będzie niższy niż 2, czyli $\tilde{D} \subset \mathbb{R}^1$. Ogólnie, w opisany sposób może się to nie udać „globalnie", ale może dać się podzielić D na takie „kawałki", na których parametryzację uda się znaleźć. Tu jednak łatwo wskazać nawet wiele różnych „globalnych" parametryzacji D, np. „$x_2 = a - x_1$", czyli ściślej: $\varphi : [0; a] \to D$, $\varphi(t) = (t, a - t)$, dla $t \in \tilde{D} = [0; a]$ — takie φ można nazwać „parametryzacją przy pomocy pierwszej współrzędnej". Gdy mamy jakąś parametryzację $\varphi : \tilde{D} \to D$ oraz f posiada ekstremum lokalne w $x_0 \in D$ i $\varphi(t_0) = x_0$, to dzięki ciągłości φ funkcja

$$\tilde{f} := f \circ \varphi$$

posiada ekstremum lokalne (tego samego typu) w t_0 (dlaczego? — proszę samodzielnie sprawdzić z definicji ekstremum lokalnego). W naszym wypadku $\tilde{f}(t) = f(t, a - t) = t \cdot (a - t)$ — to funkcja jednej zmiennej, łatwa do zbadania — osiąga ekstremum lokalne na końcach przedziału: $\tilde{f}(0) = \tilde{f}(a) = 0$ oraz wewnątrz — w punkcie $\frac{a}{2}$, $\tilde{f}(\frac{a}{2}) = \frac{a^2}{4}$. Stąd rozwiązaniem naszego zadania jest wartość $\frac{a^2}{4}$.

Idea tego postępowania polega na tym, że zamiast rozważać f określoną na D, które jako podzbiór $\mathbb{R}^m$ jest zbiorem „cienkim" (w każdym razie o pustym wnętrzu) — intuicyjnie — zbiorem „wymiaru" niższego niż m, rozważamy funkcję $\tilde{f} := f \circ \varphi$ określoną już na $\tilde{D} \subset \mathbb{R}^l$, gdzie l — to właśnie ten „wymiar" zbioru D, $l < m$ i $\tilde{D}$ ma już „duże" wnętrze. Do $\tilde{f}$ możemy zatem próbować stosować poznane wcześniej metody szukania ekstremów przy pomocy pochodnych cząstkowych (o ile istnieją). W tym prostym przykładzie pochodna cząstkowa to po prostu zwykła pochodna — funkcja bowiem miała tylko jedną zmienną.

Opiszemy teraz inną metodę, która pozwala rozwiązać nasze zadanie bez potrzeby szukania parametryzacji. Metoda ta będzie oparta na twierdzeniu, które za chwilę sformułujemy. Aby jednak treść tego twierdzenia była od razu bardziej zrozumiała, zapowiemy pewne rzeczy, które będą się w nim pojawiać. Zauważmy najpierw, że w rozważanym przez nas przypadku funkcję f określiliśmy od razu na zbiorze D, choć jak widzieliśmy można było rozszerzyć ją do bardzo „eleganckiej" funkcji zadanej na znacznie większym zbiorze — np. na całym $\mathbb{R}^2$ (zachowując wzór $x_1 \cdot x_2$). W poniższym twierdzeniu funkcja f będzie zatem określona na pewnym zbiorze otwartym $U \subset \mathbb{R}^m$, jednak ekstrema będziemy badać nie dla niej, lecz — tak jak w rozważanym przykładzie — dla jej obcięcia do pewnego „cienkiego" podzbioru M zbioru U. Zauważmy, że badany wcześniej zbiór D można rozbić na sumę dwóch następujących podzbiorów: $D_1 := \{x \in \mathbb{R}^2 : x_1, x_2 > 0, x_1 + x_2 = a\}$ oraz $D_2 = \{(a, 0), (0, a)\}$. Zbiór D_2 to „brzeg" [152], który badać można osobno. Jeśli ograniczymy się do badania funkcji na D_1,

[152] Uwaga — **nie** jest to jednak tzw. *brzeg topologiczny* zbioru, o którym my co prawda nie mówiliśmy, ale z którym to pojęciem możecie się Państwo spotkać ...

to w tym konkretnym przypadku wygodnie byłoby rozważyć $U := \{x \in \mathbb{R}^2 : x_1, x_2 > 0\}$ oraz $M := \{x \in U : x_1 + x_2 = a\}$. Powyższe M jest podzbiorem zbioru U opisanym pewnym równaniem — czyli *warunkiem* (— stąd tytułowe ekstremum *warunkowe*). W języku fizycznym mówi się też, że zostały nałożone pewne *więzy* (— stąd ekstremum *związane*). W praktyce tych równań może być więcej — powiedzmy k, gdzie $1 \leqslant k < m$. Każde równanie można zapisać używając osobnej funkcji skalarnej, albo można wszystkie razem opisać jedną funkcją wektorową $F : U \to \mathbb{R}^k$ (jej składowe F_j opisują poszczególne równania). W naszym przypadku $k = 1$ i biorąc F zadane wzorem $F(x) = x_1 + x_2$ możemy zapisać, że $M = \{x \in U : F(x) = c\}$, gdzie tu $c = a$. Zbiór tej postaci to *poziomica* funkcji wektorowej F, czyli przecięcie k poziomic funkcji skalarnych $F_1, \dots, F_k$. Sformułujmy zatem zapowiadane twierdzenie.

Twierdzenie IX.5 (o mnożnikach Lagrange'a). *Załóżmy, że U jest otwartym podzbiorem $\mathbb{R}^m$, funkcje $f : U \to \mathbb{R}$, $F : U \to \mathbb{R}^k$, gdzie $1 \leqslant k < m$, są klasy C^1 oraz $c \in \mathbb{R}^k$. Niech $M := \{x \in U : F(x) = c\}$. Jeżeli funkcja $f|_M$ posiada ekstremum lokalne w $x_0 \in M$ oraz*

$$\operatorname{rank} MJF(x_0) = k, \quad [153] \tag{IX.10}$$

to istnieją $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ takie, że

$$\operatorname{grad} f(x_0) = \sum_{j=1}^{k} \lambda_j \operatorname{grad} F_j(x_0). \tag{IX.11}$$

B.D.

Tezę powyższego twierdzenia można sformułować inaczej tak:

wektor $\operatorname{grad} f(x_0)$ *jest kombinacją liniową wektorów* $\operatorname{grad} F_1(x_0), \dots, \operatorname{grad} F_k(x_0)$.

Współczynniki $\lambda_1, \dots, \lambda_k$ będące współczynnikami w tej kombinacji liniowej nazywane są *mnożnikami Lagrange'a* (choć czasem tą nazwą określa się liczby $-\lambda_j$). Założenie (IX.10), które może Państwa nieco niepokoić, to warunek, który gwarantuje, że „w pobliżu x_0" M ma „wymiar" równy $m - k$, albo inaczej, że równania opisane funkcjami $F_1, \dots, F_k$ są „niezależne w pobliżu x_0". Należy pamiętać o tym warunku! Zobaczmy teraz jak działa to twierdzenie w naszym konkretnym zadaniu.

Przykład (metoda mnożników Lagrange'a). Już przed twierdzeniem określiliśmy U i F, M i c, przypominamy jeszcze, że $f : U \to \mathbb{R}$, $f(x) = x_1 \cdot x_2$. Oczywiście U jest otwarty oraz F i f klasy C^1. Mamy $MJF(x) = \operatorname{grad} F(x) = (1, 1)$ zatem (IX.9) zachodzi dla dowolnego $x_0 \in M$. Załóżmy teraz, że w $x \in M$ funkcja $f|_M$ posiada ekstremum lokalne. Zatem z powyższego twierdzenia istnieje $\lambda_1 = \lambda \in \mathbb{R}$ takie, że spełnione są równania

$$\begin{cases} x_2 = \lambda \\ x_1 = \lambda \\ x_1 + x_2 = a. \end{cases} \tag{IX.12}$$

Dwa pierwsze, to równania skalarne uzyskane „na obu współrzędnych" wektorowej równości (IX.11), a trzecie to równanie opisujące przynależność x do M. Mamy więc układ 3 równań z 3 niewiadomymi: x_1, x_2, λ. Ponadto mamy nierówności

$$\begin{cases} x_1 > 0 \\ x_2 > 0, \end{cases} \tag{IX.13}$$

związane z tym, że $x \in U$. Rozwiązujemy to łatwo: uzyskujemy z równań $2\lambda = a$ skąd $x_1 = x_2 = \lambda = \frac{a}{2}$ (nierówności są spełnione, bo $a > 0$), czyli jedynie w punkcie $(\frac{a}{2}, \frac{a}{2})$ funkcja $f|_M$

[153] Przypominam (z GAL-u), że rank oznacza *rząd* — w tym wypadku rząd macierzy Jakobiego funkcji F w punkcie a.

może (choć nie musi) posiadać ekstremum. Aby więc zakończyć rozwiązanie wystarczy osobno zbadać wartości f w obu punktach zbioru D_2 — tam jednak f osiąga wartość 0 i korzystając z tego co już wykazaliśmy przedtem, tj. że f osiąga gdzieś w D swoje kresy, widzimy, że kres górny osiągany jest w $(\frac{a}{2}, \frac{a}{2})$ i wynosi $\frac{a^2}{4}$.

Przy bardziej złożonych zadaniach rozwiązywanie układu równań może być trudniejsze i często trzeba w istotny sposób wykorzystać także nierówności, które mogą się pojawić w związku z warunkami opisującymi zbiór otwarty U. Jednak zawsze mamy tę samą liczbę równań co niewiadomych! Równań jest bowiem $m + k$ (m współrzędnych gradientów z tezy twierdzenia IX.5, oraz k równań na przynależność do M) i niewiadomych $m+k$ (m — liczba współrzędnych punktu x oraz k — liczba mnożników Lagrange'a).

4. Różniczkowanie a odwracalność

W twierdzeniu „o własnościach rachunkowych różniczki" (tw. IX.4) nic nie wspomnieliśmy o różniczkowalności funkcji odwrotnej. Okazuje się, że dla funkcji wielu zmiennych prawdziwe jest twierdzenie podobne do twierdzenia znanego nam dla jednej zmiennej, kiedy to dla różniczkowalności funkcji f^{-1} wystarczała de facto różniczkowalność f, ciągłość f^{-1} [154] oraz fakt, że $f'(x) \neq 0$. W pewnym sensie jednak zasadniczy był tu ostatni warunek, który zamienia się w przypadku wielu zmiennych w swoje naturalne uogólnienie[155]

$$\det Df(x) \neq 0, \tag{IX.14}$$

czyli warunek odwracalności przekształcenia liniowego $Df(x)$ (lub, w „języku macierzowym", $Jf(x) \neq 0$, czyli istnienia $(MJf(x))^{-1}$). Zanim jednak sformułujemy twierdzenie o różniczkowaniu funkcji odwrotnej zajmiemy się niejako odwróceniem tego problemu, tzn. odwracalnością funkcji spełniającej warunek (IX.14). I choć spełnienie warunku (IX.14) nawet na całej dziedzinie i nawet z dodatkowymi założeniami dotyczącymi regularności funkcji **nie** gwarantuje wcale odwracalności funkcji (patrz np. zad. IX.30), to jednak można uzyskać coś, co nazywane bywa *lokalną odwracalnością*.

Twierdzenie IX.6 (o lokalnym odwracaniu[156]). *Jeżeli U jest otwartym podzbiorem $\mathbb{R}^d$, $f : U \to \mathbb{R}^d$ jest funkcją klasy C^1 oraz $x_0 \in U$ jest taki, że*

$$\det Df(x_0) \neq 0, \tag{IX.15}$$

to istnieje $r > 0$, takie, że $V := f(K(x_0, r))$ jest zbiorem otwartym, $\tilde{f} := f|_{K(0,r)}$ jest odwracalna jako funkcja z $K(0, r)$ w V oraz $\tilde{f}^{-1}$ jest klasy C^1. **B.D.**

W szczególności $\tilde{f}$ z tezy twierdzenia jest *dyfeomorfizmem* (klasy C^1), tzn. odwracalną funkcją klasy C^1 taką, że odwrotna do niej też jest klasy C^1. Zapowiadane twierdzenie o różniczkowaniu funkcji odwrotnej jest właśnie konsekwencją powyższego twierdzenia. Warto też wspomnieć, że inną jego ważną konsekwencją jest tzw. *twierdzenie o funkcji uwikłanej*, którego tu nie będziemy szczegółowo formułować. Zamiast tego wspomnę jedynie, że dotyczy ono rozwiązywania równań (także układów równań) typu

$$F(a, x) = 0,$$

gdzie x — niewiadoma (lub niewiadome) oraz a — parametr (lub parametry) w taki sposób, aby rozwiązanie $x(a)$ zadane było „lokalnie" jako „dobra" funkcja (np. funkcja kl. C^1) od parametru a.

[154] Patrz — ostatnie zdanie dowodu tw. V.1.c).

[155] Choć nie jedyne uogólnienie. Innym byłby np. warunek $Df(x) \neq 0$, który jednak nie jest tu dobrym wyborem...

[156] Używana bywa też nazwa „twierdzenie o funkcji odwrotnej", jednak ta sama nazwa używana jest do wielu zupełnie różnych twierdzeń...

Twierdzenie IX.7 (o różniczkowaniu funkcji odwrotnej). *Jeżeli U jest otwartym podzbiorem $\mathbb{R}^d$ oraz $f : U \to W \subset \mathbb{R}^d$ jest odwracalną funkcją klasy C^1, taką że dla dowolnego $x \in U$ zachodzi* (IX.14)*, to f^{-1} jest klasy C^1 oraz dla dowolnego $x \in W$ zachodzi*

$$Df^{-1}(x) = (Df(f^{-1}(x)))^{-1}. \tag{IX.16}$$

Ponadto W jest otwarty.

Dowód.

Rozważmy dowolny $a \in W$. Niech $x_0 = f^{-1}(a)$ i niech r, V, $\tilde{f}$ będzie dobrane do x_0 tak jak w tw. IX.6. W szczególności zatem $a \in V = f(K(0,r)) \subset W$, zatem ze względu na dowolność a i otwartość V zbiór W też jest otwarty. Ponadto dla dowolnego $y \in V$ mamy $f^{-1}(y) = \tilde{f}^{-1}(y)$, czyli $f^{-1}|_V = \tilde{f}^{-1}$, zatem $f^{-1}|_V$ jest klasy C^1, skąd f^{-1} jest klasy C^1, dzięki otwartości V i dowolności a. Pozostaje wykazać (IX.16). Mamy $f \circ f^{-1} = I_W$, gdzie I_W — funkcja identycznościowa na W, tzn. $I_W(x) = x$ dla $x \in W$. Dla dowolnego $x \in W$ funkcja f^{-1} jest różniczkowalna w x, jak już wykazaliśmy. Także (z założenia) f jest różniczkowalna w $f^{-1}(x)$, a zatem z tw. IX.4 pkt. 3 mamy

$$Df(f^{-1}(x)) \circ Df^{-1}(x) = DI_W(x) = i,$$

gdzie i — przekształcenie identycznościowe na $\mathbb{R}^d$, skąd natychmiast dostajemy (IX.16). □

Na zakończenie tego podrozdziału zauważmy tylko, że w tym twierdzeniu przyjęliśmy nieco inne założenia niż w przypadku jednowymiarowym, mianowicie założenie o tym, że f jest klasy C^1. Jak się okazuje jest to silniejsze[157)] niż ciągłość f^{-1}, którą zakładaliśmy przy $d = 1$. Zrobiliśmy tak jednak ze względu na łatwość dowodu, choć prawdziwe jest także twierdzenie z ciągłością f^{-1} zamiast założenia, że f jest klasy C^1.

5. Pochodne cząstkowe wyższych rzędów

Jak różniczkować wielokrotnie funkcje wielu zmiennych?

Pytanie to wydaje się trudniejsze niż w przypadku jednej zmiennej z następujących powodów. Pochodna funkcji jednej zmiennej była funkcją także jednej zmiennej o wartościach tego samego typu, co wartości funkcji wyjściowej. Gdy f jest funkcją m zmiennych o wartościach w $\mathbb{R}^k$, to jej różniczka w każdym punkcie x nie jest już elementem $\mathbb{R}^k$. A zatem Df można traktować jako funkcję zależną od x (czyli też m zmiennych), ale już o zupełnie innych wartościach — w zbiorze przekształceń liniowych z $\mathbb{R}^m$ w $\mathbb{R}^k$. Dwukrotne różniczkowanie f musiałoby więc polegać na jednokrotnym różniczkowaniu takiej dziwnej dość funkcji Df. Wydaje się to trudne do pojęcia, szczególnie gdy pomyślimy o jeszcze wyższych różniczkach niż druga...[158)]

Na ogół jednak postępuje się trochę inaczej — zamiast różniczkować tak zawiłą funkcję Df rozpatruje się funkcje $x \rightsquigarrow Df(x)(h) \in \mathbb{R}^k$ przy ustalonych h. Dla wszystkich $h \in \mathbb{R}^m$ są to funkcje m zmiennych, ale już o wartościach w $\mathbb{R}^k$, tak jak wyjściowa funkcja f. Jeśli przy każdym wyborze h taka funkcja — oznaczmy ją jako $D_h f$ — jest w punkcie x różniczkowalna, to mówimy, że f jest *dwukrotnie różniczkowalna* w x, a *drugą różniczką* w x nazywamy przekształcenie $D^2 f(x) : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^k$ zadane dla $h^{(1)}$, $h^{(2)} \in \mathbb{R}^m$ wzorem

$$D^2 f(x)(h^{(1)}, h^{(2)}) := D(D_{h^{(1)}} f)(x)(h^{(2)}).$$

Kontynuując takie postępowanie można zdefiniować rekurencyjnie n–krotną różniczkowalność i n–tą różniczkę $D^n f(x)$ funkcji f w punkcie x, która — jak łatwo się przekonać — będzie

157) Oczywiście — przy spełnieniu pozostałych założeń.
158) Choć jest wykonalne.

określona na $(\mathbb{R}^m)^n$ i będzie przekształceniem n–liniowym (tzn. liniowym ze względu na każdą ze zmiennych wektorowych $h^{(1)}, \ldots, h^{(n)}$ przy ustalonych pozostałych). Jak widać jest to trochę skomplikowane, ale na szczęście nie będziemy tym pojęciem posługiwać się dalej — podaję to jedynie w celach „informacyjnych".

Znacznie prostsze pojęciowo są tzw. *wyższe pochodne cząstkowe*. Jak pamiętamy pochodna cząstkowa $\partial_{x_j} f$ była funkcją tego samego typu co f (tyle samo zmiennych, taki sam typ wartości). A zatem wyższe pochodne cząstkowe można łatwo określić rekurencyjnie, podobnie jak określało się wyższe pochodne funkcji jednej zmiennej, a jedyna różnica jest taka, że kolejne różniczkowania mogą być „po rozmaitych zmiennych". Jeżeli $f : D \to \mathbb{R}^k$, gdzie $D \subset \mathbb{R}^m$ to jej pochodną rzędu n w punkcie x po zmiennych kolejno: $x_{j_1}, \ldots, x_{j_n}$, gdzie $j_1, \ldots, j_n \in \{1, \ldots, m\}$, oznaczamy (o ile istnieje) którymkolwiek z symboli

$$\partial_{x_{j_n} \ldots x_{j_1}} f(x), \quad \partial_{j_n, \ldots, j_1} f(x), \quad \frac{\partial^n f}{\partial x_{j_n} \ldots \partial x_{j_1}}(x).$$

A ścisła definicja rekurencyjna jest taka:

Jeżeli $a \in D$ oraz dla pewnego $r > 0$ pochodna cząstkowa $\partial_{j_n, \ldots, j_1} f(x)$ istnieje i jest skończona przy dowolnym $x \in D_{a,r} := K(a,r) \cap D$ oraz $j_{n+1} \in \{1, \ldots, m\}$, to

$$\partial_{j_{n+1}, j_n, \ldots, j_1} f(a) := \partial_{j_{n+1}} (\partial_{j_n, \ldots, j_1} f)(a),$$

o ile pochodna cząstkowa po j_{n+1} z prawej strony powyższego wzoru istnieje.

A zatem może być aż m^n różnego typu pochodnych cząstkowych rzędu n. Np. gdy $m = 2$ i $n = 2$ mogą pojawiać się 4 takie pochodne:

$$\frac{\partial^2 f}{\partial x_1 \partial x_1}(a), \quad \frac{\partial^2 f}{\partial x_2 \partial x_1}(a), \quad \frac{\partial^2 f}{\partial x_1 \partial x_2}(a), \quad \frac{\partial^2 f}{\partial x_2 \partial x_2}(a),$$

a przy $m = 3$ i $n = 2$ będzie ich już aż 9.

Przykład. Niech $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = x_1 e^{x_2}$. Wówczas

$\partial_{x_1} f(x) = e^{x_2}$, $\partial_{x_2} f(x) = x_1 e^{x_2}$, $\partial_{x_1 x_1} f(x) = 0$, $\partial_{x_2 x_2} f(x) = x_1 e^{x_2}$, $\partial_{x_2 x_1} f(x) = e^{x_2}$, $\partial_{x_1 x_2} f(x) = e^{x_2}$.

W powyższym przykładzie zwraca uwagę fakt, że w każdym punkcie x zachodzi równość

$$\partial_{x_1 x_2} f(x) = \partial_{x_2 x_1} f(x).$$

Okazuje się, że nie jest to przypadek. I choć taka sytuacja nie zachodzi zawsze, to jednak, jak zobaczymy za chwilę, tak jest dla funkcji „dostatecznie regularnych". Najpierw musimy określić czym ma być owa regularność. Będzie to uogólnienie na wyższe rzędy pojęcia klasy C^1. Mianowicie $f : D \to \mathbb{R}$ jest *klasy* C^n (zapisujemy to także: $f \in C^n(D)$) wtw istnieją jej wszystkie[159] pochodne cząstkowe rzędu n i są one funkcjami ciągłymi na D.

Twierdzenie IX.8. *Jeżeli f jest klasy C^n to dla dowolnego $x \in \operatorname{Int} D$ wartość $\partial_{j_1, \ldots, j_n} f(x)$ nie zależy od kolejności liczb $j_1, \ldots, j_n$.*

B.D.

A zatem w przypadku takim jak w twierdzeniu można uprościć nieco notację służącą do zapisu pochodnych cząstkowych — nie jest bowiem ważne w jakiej kolejności różniczkujemy, ważne jedynie ile razy różniczkujemy po danej zmiennej. W takiej sytuacji stosowany jest zatem najczęściej jeden spośród następujących zapisów:

$$\partial^\alpha f(x), \quad \partial_{x_1^{\alpha_1} \ldots x_m^{\alpha_m}} f(x), \quad \frac{\partial^n f}{\partial x_1^{\alpha_1} \ldots x_m^{\alpha_m}}(x) \text{ [160]},$$

---

159) Tzn. wszystkie ze wspomnianych wcześniej m^n możliwych.

160) Ponadto najczęściej pomija się w drugim i trzecim z tych zapisów te zmienne, dla których $\alpha_j = 0$ np. $\frac{\partial^3 f}{\partial x_1 \partial x_3^2}(x) = \frac{\partial^3 f}{\partial x_1^1 \partial x_2^0 \partial x_3^2}(x)$.

gdzie α to tzw. *multiindeks*, czyli element $\mathbb{N}^m$, $\alpha = (\alpha_1, \ldots, \alpha_m)$, $n = \alpha_1 + \ldots + \alpha_m$ oraz α_j jest liczbą różniczkowań po zmiennej x_j.

Jednak aby korzystać w pełni z powyższej wygody trzeba najpierw wiedzieć, że funkcja jest klasy C^n... Korzystanie do tego celu tylko z definicji byłoby tu mało sensowne. By sprawdzić bowiem np. że $\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}$, musielibyśmy i tak wyliczyć m. in. obie te pochodne cząstkowe i sprawdzić ciągłość każdej z nich, więc ich równość byłaby zapewne jasna i tak, bez twierdzenia IX.8... Na szczęście prawdziwy jest poniższy fakt dotyczący operacji na funkcjach klasy C^n.

Fakt. *Jeżeli f i g są funkcjami klasy C^n określonymi na zbiorach otwartych, to ich suma, różnica, iloczyn, iloraz oraz złożenie (o ile dana operacja jest określona) jest także klasy C^n. Jeżeli ponadto f jest odwracalna i ma różniczkę $Df(x)$ odwracalną w każdym punkcie dziedziny funkcji f, to f^{-1} jest klasy C^n.*

Nietrudny dowód indukcyjny tego faktu pomijamy. Dzięki niemu, by sprawdzić np., że funkcja 3-ech zmiennych zadana wzorem $f(x) = e^{x_1 \ln(1+x_2^2+x_3^2)}$ jest klasy C^n wystarczy sprawdzić, że są takimi funkcje 3-ech zmiennych: $\mathbb{x}_1$, $\mathbb{x}_2$, $\mathbb{x}_3$ oraz jednej zmiennej: exp i ln — a to jest oczywiste przy dowolnym n.

Warto jeszcze wspomnieć, że jeżeli f jest klasy C^n, to jest różniczkowalna n–krotnie (w sensie, o którym mówiliśmy na początku tego rozdziału) w każdym punkcie z wnętrza swej dziedziny. Stanowi to uogólnienie znanego nam już wcześniej wyniku dla $n = 1$ (tw. IX.3).

6. Pochodne cząstkowe rzędu 2 i ekstrema lokalne

Spośród pochodnych cząstkowych wyższych rzędów najważniejsze będą dla nas pochodne rzędu 2. Powodem tego jest pewne ważne kryterium dające warunki dostateczne na ekstrema lokalne funkcji. Zanim sformułujemy to kryterium, opiszemy sytuację, jaka ma miejsce dla jednej zmiennej.

Fakt. *Jeżeli $f : D \to \mathbb{R}$, $D \subset \mathbb{R}$, $a \in \operatorname{Int} D$ oraz f jest dwukrotnie różniczkowalna w a, przy czym $f'(a) = 0$ i $f''(a) \neq 0$, to*

(i) *jeżeli $f''(a) > 0$, to f ma minimum lokalne w a;*

(ii) *jeżeli $f''(a) < 0$, to f ma maksimum lokalne w a.*

Fakt ten nietrudno udowodnić w oparciu o wzór Taylora z resztą Peano (patrz np. zadanie V.30). Nie pojawił się on na wykładzie, ponieważ nie ma zbyt istotnego praktycznego zastosowania. W przypadku bowiem funkcji jednej zmiennej, w „typowych" problemach, rozpoznaje się ekstrema łatwiej poprzez badanie przedziałów monotoniczności, do czego wystarcza na ogół 1-sza pochodna. Dla wielu zmiennych jednak nie mamy pojęcia monotoniczności, a zatem wynik w stylu powyższego faktu mógłby okazać się narzędziem przydatnym. Powstaje zatem pytanie:

Jak powinny wyglądać analogi założeń „$f'(a) = 0$" oraz „$f''(a) < (>) 0$"?

Z pierwszym z nich sprawa jest prosta — zastąpimy go warunkiem $Df(a) = 0$, czyli $\operatorname{grad} f(a) = 0$. By poradzić sobie z drugim warunkiem wprowadzimy najpierw następujące oznaczenia. Załóżmy, że $f : D \to \mathbb{R}$, $a \in \operatorname{Int} D$, gdzie $D \subset \mathbb{R}^m$ i f jest klasy C^2 w pewnym otoczeniu punktu a. Zdefiniujmy macierz $Hf(a)$ o wymiarach $m \times m$:

$$Hf(a) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{i,j=1,\ldots,m}$$

(przypominam, że wg. pow. notacji i to numer wiersza, a j — kolumny). Macierz ta nazywana jest *macierzą Hessego* (funkcji f w punkcie a) i, jak można wykazać, jest to macierz drugiej różniczki $D^2f(a)$ [161]) . Na mocy twierdzenia IX.8 macierz ta jest symetryczna, a zatem hermitowska (bo ma wyrazy rzeczywiste). Dla takich macierzy znacie Państwo (oczywiście dzięki wykładom z GAL-u) pojęcia *dodatniej* i *ujemnej określoności*, a także *nieokreśloności*, odpowiadające analogicznym pojęciom dla formy dwuliniowej wyznaczonej przez tę macierz. I właśnie te dwa pierwsze pojęcia użyte w odniesieniu do macierzy Hessego posłużą nam za analogi warunków $f''(a) > 0$ i $f''(a) < 0$.

Twierdzenie IX.9. *Załóżmy, że f jest klasy C^2 w pewnym otoczeniu punktu $a \in \operatorname{Int} D$ oraz że* $\operatorname{grad} f(a) = 0$. *Wówczas*

(i) *jeżeli $Hf(a)$ jest dodatnio określona, to f ma minimum lokalne w a;*

(ii) *jeżeli $Hf(a)$ jest ujemnie określona, to f ma maksimum lokalne w a;*

(iii) *jeżeli $Hf(a)$ jest nieokreślona, to f nie posiada ekstremum lokalnego w a.*

B.D.

Dowód tego twierdzenia można przeprowadzić (podobnie jak dla $m = 1$) przy użyciu odp. wersji wzoru Taylora dla wielu zmiennych. My jednak nie będziemy nawet podawać postaci wielomianu Taylora dla m zmiennych, choć warto wiedzieć, że ta teoria poznana dla jednej zmiennej ma także swoje wielowymiarowe uogólnienie.

Uwagi.

1. Sytuacja (iii) nie miała swego odpowiednika dla $m = 1$. A to dlatego, że macierz 1×1 nie może być nieokreślona.

2. Ekstrema lokalne z p. (i) i (ii) twierdzenia są nawet *ścisłe* (inaczej *właściwe*), tzn. w pewnym otoczeniu punktu a wartość $f(a)$ osiągana jest **tylko** w punkcie a.

3. Pamiętajmy, że przypadki (i), (ii), (iii) nie wyczerpują wszystkich możliwości — macierz $Hf(a)$ może być jeszcze *półokreślona* (dodatnio lub ujemnie)[162]) i w tym przypadku twierdzenie nie rozstrzyga o istnieniu lub nieistnieniu ekstremum. Odpowiada to sytuacji, gdy $f''(a) = 0$ dla $m = 1$.

4. Twierdzenie to daje nam pewne warunki **dostateczne** na ekstrema lokalne. Stanowi ono zatem ważne uzupełnienie twierdzenia IX.1, dającego jedynie warunki konieczne (przy odpowiednich założeniach), które są zresztą częścią pojawiających się tu warunków dostatecznych.

Aby ułatwić Państwu praktyczne korzystanie z twierdzenia IX.9 podam tu jeszcze kilka algebraicznych faktów pomocnych przy badaniu określoności macierzy.

Niech A będzie macierzą rzeczywistą symetryczną o wymiarach $m \times m$.

Fakt IX.1 (kryterium Sylwestera).

- *A jest dodatnio określona wtw* $\forall_{k=1,\dots,m}\ \det A^{(k)} > 0$,
- *A jest ujemnie określona wtw* $\forall_{k=1,\dots,m}\ (-1)^k \det A^{(k)} > 0$,

gdzie $A^{(k)}$ jest macierzą powstałą z A przez skreślenie kolumn i wierszy o numerach $> k$.

161) $D^2f(a)$ jest w szczególności formą dwuliniową, a pojęcie macierzy formy dwuliniowej poznaliście Państwo na wykładach z GAL-u.

162) Tzn. określona nieujemnie (odpow.: niedodatnio) ale **nie** określona dodatnio (odpow.: ujemnie).

Powyższy wynik poznaliście Państwo już, być może, na wykładzie z GAL-u. Do sformułowania kolejnego wyniku potrzebne nam będzie pojęcie wartości własnej macierzy. Liczba $\lambda \in \mathbb{C}$ jest *wartością własną* macierzy A (tu nie musi być ona symetryczna) wtw λ jest pierwiastkiem równania

$$\det(A - \lambda I) = 0.$$

Łatwo wykazać, że lewa strona powyższego równania jest wielomianem m-tego stopnia względem λ, zatem macierz może mieć conajwyżej m różnych wartości własnych zespolonych, a licząc je „z krotnościami” jako pierwiastków powyższego wielomianu jest ich dokładnie m. Można też wykazać, że dla A — hermitowskich wszystkie one są liczbami rzeczywistymi.

Fakt IX.2.

- *A jest dodatnio określona wtw wszystkie wartości własne A są dodatnie*[163] *,*
- *A jest ujemnie określona wtw wszystkie wartości własne A są ujemne,*
- *A jest nieokreślona wtw istnieje dodatnia wartość własna A i ujemna wartość własna A.*

Nietrudno wykazać, że iloczyn wszystkich wartości własnych A liczonych z w.w. krotnościami równy jest $\det A$. Otrzymujemy stąd wniosek, będący bezpośrednią konsekwencją ostatniej części powyższego faktu.

Wniosek. *Jeżeli m jest parzyste oraz $\det A < 0$, to A jest nieokreślona.*

Dowód.
Niech $\lambda_1, \ldots, \lambda_m$ będą wartościami własnymi A wypisanymi z krotnościami. Zatem $\det A = \lambda_1 \cdot \ldots \cdot \lambda_m$, czyli $\lambda_j \neq 0$ dla każdego j. Gdyby wszystkie λ_j były dodatnie, to $\det A$ byłby dodatni. Gdyby wszystkie były ujemne, to także mielibyśmy $\det A > 0$ ze względu na parzystość m. Stąd istnieją λ_i i λ_j różnych znaków. $\square$

Zakończymy ten rozdział konkretnym przykładem.

Przykład. Niech $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = x_1^2 + x_2^2 + ax_1x_2$. Spróbujemy wyznaczyć wszystkie ekstrema lokalne tej funkcji w zależności od parametru $a \in \mathbb{R}$. Sprawdzimy więc warunek konieczny $\operatorname{grad} f(x) = 0$ na to, by w x było ekstremum lokalne. Mamy

$$\partial_{x_1} f(x) = 2x_1 + ax_2, \;\; \partial_{x_2} f(x) = 2x_2 + ax_1,$$

otrzymujemy zatem układ równań

$$\begin{cases} 2x_1 + ax_2 & = 0 \\ ax_1 + 2x_2 & = 0. \end{cases}$$

Ponieważ macierz tego układu (lewej strony) to $B = \begin{pmatrix} 2 & a \\ a & 2 \end{pmatrix}$ i $\det B = 4 - a^2$ zatem, dla $a \neq \pm 2$ układ ten ma dokładnie jedno rozwiązanie $x = 0$ (tj. $x_1 = x_2 = 0$). Gdy $a = 2$, to rozwiązaniem jest cała prosta l_2 o równaniu $x_2 = -x_1$, a gdy $a = -2$, to rozwiązaniem jest prosta l_{-2} orównaniu $x_2 = x_1$.

Teraz policzmy $Hf(x)$. Mamy

$$Hf(x) = \begin{pmatrix} 2 & a \\ a & 2 \end{pmatrix} = B$$

[163] Przypomnijmy, że dodatnie tzn. > 0, nie tylko $\geqslant 0$.

(tę ostatnią równość należy raczej traktować jako przypadek, ponadto zauważmy, że tu wyszła nam macierz niezależna w ogóle od x). Dzięki kryterium Sylwestera macierz ta jest dodatnio określona wtw $\det B > 0$ wtw $|a| < 2$, ale nie jest ona nigdy ujemnie określona. Z kolei z wniosku z faktu 2, gdy $|a| > 2$, to $Hf(x)$ jest nieokreślona. To samo możemy uzyskać bezpośrednio z faktu 2, bowiem $\det(B - \lambda I) = (2-\lambda)^2 - a^2 = (2-a-\lambda)(2+a-\lambda)$, skąd wartościami własnymi $Hf(x)$ są liczby $2-a$ i $2+a$. W efekcie, twierdzenia IX.1 i IX.9 dają nam częściowe rozwiązanie problemu:

- jeżeli $|a| < 2$, to jedyne ekstremum lokalne jest w $x = 0$ i jest to ścisłe minimum lokalne;
- jeżeli $|a| > 2$, to f nie posiada ekstremów lokalnych.

Pozostaje przypadek $a = \pm 2$. Na szczęście daje się on zbadać „ręcznie" (zresztą powyższe przypadki także dawały się zbadać bez całej tej teorii — jak?). Wówczas bowiem $f(x) = (x_1 \pm x_2)^2$, a zatem na całej prostej „$x_1 \pm x_2$" osiągana jest wartość najmniejsza funkcji f wynosząca 0. W szczególności zatem w każdym punkcie prostej $l_{\pm 2}$ funkcja f osiąga minimum lokalne (nieścisłe) i są to jedyne ekstrema lokalne f.

Zadania do Rozdziału IX

$\forall$ 1. Wyznacz prostą styczną do obrazu krzywej f w punkcie a:

(a) $f : [0; 2N\pi] \to \mathbb{R}^3$, $f(t) = (R\cos t, R\sin t, \frac{\alpha t}{2\pi})$ gdzie $N \in \mathbb{N}$, $R, \alpha > 0$, $a = f(\pi)$;

(b) $f : [0; 2N\pi] \to \mathbb{R}^2$, $f(t) = (R + \alpha t)(\cos t, \sin t)$, gdzie $N \in \mathbb{N}$, $R, \alpha > 0$, $a = f(\pi)$.

Ponadto zinterpretuj sens geometryczny obrazów tych krzywych oraz rolę parametrów N, R, α.

2. Znajdź przykład takiej krzywej $f : [a; b] \to \mathbb{R}^d$ klasy C^1, dla której nie zachodzi teza twierdzenia Lagrange'a o wartości średniej, tzn. takiej, że

$$\forall_{c\in[a;b]}\ f(b) - f(a) \neq (b - a) \cdot f'(c).$$

$\forall$ 3. Rozważamy następującą sytuację dotyczącą funkcji d–zmiennych, gdzie $d = 1$ lub 2. Funkcja $f : \bar{K}(0, 1) \to \mathbb{R}$ jest ciągła, po obcięciu do $K(0, 1)$ jest klasy C^1 oraz:

1. 0 jest jedynym punktem $x \in K(0, 1)$ gdzie $\operatorname{grad} f(x) = 0$,

2. f ma maksimum lokalne w 0,

3. f **nie** osiąga wartości największej w punkcie 0.

(a) Wykaż, że taka sytuacja nie jest możliwa gdy $d = 1$.

(b) Przedstaw przekonywującą ilustrację mapy poziomic takiej funkcji f, która jest przykładem powyższej sytuacji przy $d = 2$.

$\forall$ 4. Oblicz pochodne cząstkowe (rzędu 1) po wszystkich zmiennych (w każdym punkcie dziedziny) funkcji f i sprawdź, że f jest funkcją klasy C^1:

(a) $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x) = x_1 e^{(x_1+2x_2)}$;

(b) $f : (0; +\infty)^3 \to \mathbb{R}$, $f(x) = x_1^{(x_2^{x_3})}$;

(c) $f : (1; +\infty) \times (0; +\infty) \to \mathbb{R}$, $f(x) = \log_{x_1} x_2$;

(d) $f : \mathbb{R}^2 \to \mathbb{R}^3$, $f(x) = (\operatorname{arctg}(x_1 e^{x_1+x_2}), x_1 \sin(x_1 x_2), (x_1 x_2)^2)$.

5. Dla poniższych funkcji $f : \mathbb{R}^2 \to \mathbb{R}$ zbadaj w jakich punktach x istnieje pochodna cząstkowa $\partial_1 f(x)$, a w jakich $\partial_2 f(x)$:

(a) $f(x) = |x_1|x_2$;

(b) $f(x) = \sqrt{x_1^2 + x_2^4} - |x_1|$.

6. Dla funkcji z zadania VIII.16 zbadaj

(i) istnienie każdej z pochodnych cząstkowych (1–go rzędu) w 0;

(ii) istnienie pochodnych kierunkowych w kierunku każdego wektora w 0. Wyniki skonfrontuj z wynikami zadania VIII.16 (dot. ciągłości f w 0).

7. Wykaż, że jeżeli funkcja skalarna f określona na „kostce" $D := [a_1; b_1] \times \ldots \times [a_m; b_m]$ posiada wszystkie pochodne cząstkowe (rzędu 1) w każdym punkcie dziedziny i są one funkcjami ograniczonymi, to f jest funkcją Lipshitzowską, tzn.

$$\exists_{c\in\mathbb{R}}\ \forall_{x,y\in D}\ |f(x) - f(y)| \leqslant c\|x - y\|.$$

8. Wykaż, że jeżeli U jest otwartym podzbiorem $\mathbb{R}^d$ oraz $f : U \to \mathbb{R}$ spełnia warunki takie jak w zadaniu IX.7, to f jest ciągła.

   Wskazówka: użyj wyniku z zadania IX.7.

9. Wykaż, że jeżeli istnieje pochodna kierunkowa $\partial_v f(a)$, to dla dowolnego $\alpha \in \mathbb{R}$ istnieje także $\partial_{\alpha v} f(a)$ oraz $\partial_{\alpha v} f(a) = \alpha \partial_v f(a)$.

10. Niech $f : \mathbb{R}^d \to \mathbb{R}$ będzie funkcją jednorodną, tzn. taką, że dla dowolnego $t \in \mathbb{R}$ i $x \in \mathbb{R}^d$ zachodzi $f(tx) = tf(x)$. Wykaż, że f jest różniczkowalna w 0 wtw f jest funkcjonałem liniowym (tzn. skalarną funkcją liniową).

$\forall$ 11. Zbadaj różniczkowalność funkcji $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x) = x_1^2 x_2 + x_1 x_2 x_3 + 7x_2^2 - 9x_3$ dwoma sposobami:

   (a) bezpośrednio z definicji różniczki;

   (b) w oparciu o twierdzenie „o różniczkowalności dla klasy C^1” (tw. IX.3).

12. Niech $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = |x_1| \cdot x_2$ (patrz zad. IX.5 a). Zbadaj różniczkowalność f.

$\forall$ 13. [164] Znajdź macierz Jakobiego funkcji f w punkcie a:

   (a) $f : \mathbb{R}^2 \to \mathbb{R}^3$, $f(x) = (x_1, x_1 + x_2, x_1 \cdot x_2)$, $a = (1, 7)$;

   (b) $f : \mathbb{R}^3 \to \mathbb{R}^2$, $f(x) = (x_1 + 2x_2 + 3x_3, x_1 \cdot x_2^2 \cdot x_3^3)$, $a = (0, 1, 2)$;

   (c) $f : \mathbb{R}^4 \to \mathbb{R}$, $f(x) = (e^{x_1 x_2 x_3 x_4})$, $a = (1, 0, 1, 0)$;

   (d) $f : \mathbb{R} \to \mathbb{R}^4$, $f(t) = (0, t, t^2, t^3)$, $a = 1$;

   (e) $f : \mathbb{R} \to \mathbb{R}$, $f(t) = \operatorname{arctg}(t^2 + 1)$, $a = 1$.

$\forall$ 14. Oblicz pochodną kierunkową funkcji f w punkcie a w kierunku $v = (0, \frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$ dla przykładu z zadania IX.13 c) dwoma metodami:

   (a) przy użyciu definicji $\frac{\partial f}{\partial v}(a)$,

   (b) wykorzystując wynik z zadania IX.13 c) i odpowiedni (jaki?) wynik z wykładu.

$\forall$ 15. [165] W oparciu o twierdzenie o ekstremach lokalnych (tw. IX.1)[166] znajdź kresy (górny i dolny) poniższych funkcji $f : D \to \mathbb{R}$. Zbadaj, czy kresy te są przez f osiągane.

   (a) $D = \{x \in \mathbb{R}^2 : x_2 \geqslant 0,\ x_1^2 + x_2^2 \leqslant 1\}$, $f(x) = x_1^2 + (x_2 - \frac{1}{2})^2$,

   (b) $D = \{(x, y, z) \in \mathbb{R}^3 : x, y, z \geqslant 0,\ x + y + z \leqslant 4\}$, $f(x, y, z) = xyz(4 - x - y - z)$,

   (c) $D = \{x \in \mathbb{R}^2 : x_1, x_2 \leqslant 0\}$, $f(x) = (x_1 + 3x_2)e^{2x_1 + x_2}$,

   (d) $D = \mathbb{R}^2$, $f(x) = x_1 x_2 e^{-(x_1^2 + x_2^2)}$.

   Wskazówka do c) i d): udowodnij następujące twierdzenie „o osiąganiu jednego kresu” (por. także zad. IV.15): *Gdy istnieje zbiór zwarty $K \subset D$ oraz $a \in K$ takie, że $\forall_{x \in D \setminus K}\ f(x) \geqslant f(a)$ ($\leqslant f(a)$), to f, jeśli jest ciągła, osiąga swój kres dolny (górny).*

---

[164] Choć 2 przykłady.

[165] Przynajmniej jeden z a), b) i jeden z c), d).

[166] Tu proszę nie stosować jeszcze twierdzenia o mnożnikach Lagrange'a.

16. Wyznacz wymiary takiego prostopadłościennego akwarium bez „górnej przykrywki” o objętości 100 litrów, na którego zbudowanie potrzeba zużyć najmniejszej powierzchni szyb.

17. Wyznacz wymiary prostopadłościennego 100 litrowego akwarium o szkielecie zbudowanym z prętów (wzdłuż wszystkich krawędzi), na zbudowaniu którego potrzeba najmniejszej długości prętów.

18. Bez użycia pochodnych cząstkowych rzędu 2 znajdź ekstrema lokalne funkcji $f : \mathbb{R}^2 \to \mathbb{R}$ zadanej wzorem $f(x, y) = x^2 y^3 (6 - x - y)$.

    Wskazówka: naszkicuj poziomicę „zerową” dla f (tzn. $f^{-1}(\{0\})$) oraz zbadaj znaki f poza tą poziomicą.

19. Przy użyciu „reguły łańcuchowej” (wniosek ze str. 137) wyprowadź wzory:

    $\forall$ (a) na pochodną iloczynu dwóch skalarnych funkcji różniczkowalnych jednej zmiennej,

    (b) na pochodną iloczynu trzech skalarnych funkcji różniczkowalnych jednej zmiennej,

    (c) na pochodną funkcji danej wzorem $f(t) = (g(t))^{h(t)}$, gdzie $g(t) > 0$ i g oraz h — różniczkowalne funkcje skalarne jednej zmiennej.

20. Korzystając z „reguły łańcuchowej” wykaż następujące twierdzenie:

    Jeżeli $f : U \to \mathbb{R}$ jest funkcją różniczkowalną, taką że $\forall_{x \in U}$ $\operatorname{grad} f(x) = 0$ oraz U jest taki, że każde dwa jego punkty można połączyć łamaną zawartą w U, to f jest stała.

$\forall$ 21. Niech $f : \mathbb{R}^2 \to \mathbb{R}$ będzie różniczkowalna oraz $\forall_{x \in \mathbb{R}^2}$ $\frac{\partial f}{\partial x_2}(x) = 2\frac{\partial f}{\partial x_1}(x)$. Wykaż, że na każdej prostej o równaniu $2x_2 + x_1 = c$ funkcja f jest stała.

22. Niech $f : \mathbb{R}^2 \to \mathbb{R}$ będzie różniczkowalna oraz $\forall_{x,y \in \mathbb{R}}$ $x\frac{\partial f}{\partial y}(x, y) = y\frac{\partial f}{\partial x}(x, y)$. Wykaż, że jeśli $\forall_{x>0}$ $f(x, x) \geqslant 0$, to $\forall_{x,y \in \mathbb{R}}$ $f(x, y) \geqslant 0$.

    Wskazówka: znajdź najpierw pewną rodzinę podzbiorów $\mathbb{R}^2$, na których f jest stała.

$\forall$ 23. Niech $A \subset \mathbb{R}^d$ oraz $v \in \mathbb{R}^d$. Będziemy mówić, że v jest *prostopadły* do A *w punkcie* $x_0 \in A$ wtw dla dowolnej krzywej $\gamma : [a; b] \to A$ różniczkowalnej w $(a; b)$ zachodzi $\gamma'(t_0) \perp v$ dla każdego $t_0 \in (a; b)$ takiego, że $\gamma(t_0) = x_0$. Wykaż twierdzenie:

    Jeżeli $f : U \to \mathbb{R}$ jest różniczkowalna, U — otwarty w $\mathbb{R}^d$, to $\operatorname{grad} f(x_0)$ *jest prostopadły do poziomicy funkcji f wyznaczonej wartością $f(x_0)$ (tzn. do $f^{-1}(\{f(x_0)\})$) w punkcie x_0, dla dowolnego $x_0 \in U$.*

24. Wykaż p. 3 twierdzenia IX.4 (czyli punkt o różniczkowaniu złożenia).

25. Wykaż, że jeśli $f : D \to \mathbb{R}$ i $g : D \to \mathbb{R}^k$ są różniczkowalne w $a \in D \subset \mathbb{R}^m$ to $f \cdot g$ jest różniczkowalna w a oraz dla $h \in \mathbb{R}^m$ zachodzi

    $$D(f \cdot g)(a)(h) = Df(a)(h) \cdot g(a) + f(a) \cdot Dg(a)(h).$$

    Dowód przeprowadź w oparciu o różniczkowanie złożenia.

26. Wykaż, że suma, różnica, iloczyn, iloraz i złożenie funkcji klasy C^1 określonych na zbiorach otwartych należą do klasy C^1.

$\forall$ 27. [167] Znajdź kresy funkcji f zadanych poniższymi wzorami na zbiorze M, zbadaj czy są one osiągane.

(a) $f(x) = x_1^2 + x_2^2$, $\quad M = \{x \in \mathbb{R}^2 : 2x_1 + 3x_2 = 7\}$;

(b) $f(x) = \sqrt{(x_1 - 2)^2 + x_2^2}$, $\quad M = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$;

(c) $f(x, y, z) = xyz$, $\quad M = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = x^2 + y^2 + z^2 = 1\}$;

(d) $f(x) = Ax_1 + Bx_2 + C$, $\quad M = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$;

(e) $f(x, y) = x^2 + 12xy + 2y^2$, $\quad M = \{4x^2 + y^2 \leqslant 25\}$;

(f) $f(x, y, z) = x - 2y + 2z$, $\quad M = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1,\ x + y + z = 0\}$;

(g) $f(x) = \|x\|^2$, $\quad M = \left\{x \in \mathbb{R}^3 : x_1^2 + \frac{x_2^2}{4} + \frac{x_3^2}{9} = 1\right\}$;

(h) $f(x) = x_1 x_2^2 x_3^3$, $\quad M = \{x \in \mathbb{R}^3 : x_1, x_2, x_3 > 0, x_1 + 2x_2 + 3x_3 = 6\}$;

(i) $f(x) = x_1^2 + 2x_2^2 + 3x_3^2$, $\quad M = \{x \in \mathbb{R}^3 : \|x\|^2 \leqslant 100\}$;

(j) $f(x, y, z) = x + y + z$, $\quad M = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \leqslant z \leqslant 1\}$;

(k) $f(x, y, z) = x + 2y$, $\quad M = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1, 2y^2 + 9z^2 = 1\}$;

(l) $f(x, y, z) = x + 2y$, $\quad M = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1, x^2 + 4y^2 + 9z^2 = 1\}$.

28. Wykaż nierówności:

(a) $\frac{x^n + y^n}{2} \geqslant (\frac{x+y}{2})^n$ dla $n \geqslant 1$, $x, y \geqslant 0$;

(b) (nier. Höldera) $\sum_{k=1}^n x_k y_k \leqslant (\sum_{k=1}^n |x_k|^p)^{\frac{1}{p}} \cdot (\sum_{k=1}^n |y_k|^q)^{\frac{1}{q}}$ dla $x_k, y_k \in \mathbb{R}$, $p, q \geqslant 1$, $\frac{1}{p} + \frac{1}{q} = 1$.

Wskazówka: Szukaj kresów jednej (odpowiedniej...) ze stron nierówności traktowanej jako funkcja określona na zbiorze, na którym druga ze stron jest stała.

29. Czy istnieje punkt z płaszczyzny w $\mathbb{R}^3$ o równaniu $3x_1 - 2x_3 = 0$, dla którego suma kwadratów odległości od punktów $(1, 1, 1)$ i $(2, 3, 4)$ jest najmniejsza? Jeśli tak, to znajdź wszystkie takie punkty.

$\forall$ 30. Niech I będzie otwartą d-wymiarową kostką, tzn. zbiorem postaci $I = (a_1; b_1) \times \ldots \times (a_d; b_d) \subset \mathbb{R}^d$ ($I \neq \emptyset$). Rozważamy funkcje $f : I \to \mathbb{R}^d$ klasy C^1 spełniające $\forall_{x \in I}\ \det Df(x) \neq 0$.

(a) Sprawdź, że gdy $d = 2$ i $f(x) = (x_1 \cos x_2, x_1 \sin x_2)$ dla $x \in I = I_R = (0, R) \times \mathbb{R}$, gdzie $R \in (0; +\infty)$, to warunki powyższe są spełnione, ale f nie jest różnowartościowa.

(b) Wykaż, że gdy I z pkt. a) zmienimy na $(0, R) \times (0; 2\pi)$ nie zmieniając wzoru na f, to f będzie dyfeomorfizmem na $f(I)$. Wyznacz $f(I)$. Wyznacz dla $R > 1$ wektor $Df^{-1}((-1, 0))((1, 2))$ bez użycia wzoru na funkcję f^{-1}.

(c) Wykaż, że gdy $d = 1$, to $f : I \to f(I)$ jest odwracalna, a nawet jest dyfeomorfizmem.

$\forall$ 31. Znajdź wszystkie pochodne cząstkowe wszystkich możliwych rzędów dla:

$\forall$ (a) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = x_1^2 x_2$,

[167] Przynajmniej 3 przykłady, w tym conajmniej jeden z „$\leqslant$”.

(b) $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x) = x_1 e^{(x_1+2x_2+3x_3)}$

Dla jakich $n \in \mathbb{N}$ w powyższych przykładach f jest funkcją klasy C^n?

$\forall$ 32. Oblicz $\frac{\partial f}{\partial x_2 \partial x_1}(0)$ i $\frac{\partial f}{\partial x_1 \partial x_2}(0)$ dla $f : \mathbb{R}^2 \to \mathbb{R}$ zadanej wzorem

$$f(x) = \begin{cases} 0 & \text{dla } x = 0 \\ x_1 x_2 \frac{x_1^2 - x_2^2}{\|x\|^2} & \text{dla } x \neq 0. \end{cases}$$

Czy ta funkcja jest klasy C^2?

$\forall$ 33. [168] Znajdź ekstrema lokalne poniższych funkcji:

(a) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = (x_1 - 1)^2 + 2x_2^2$;

(b) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = 2x_1x_2 - x_1^2 - 2x_2^2$;

(c) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + 4xy$;

(d) $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x) = x_1^2 + x_2^2 + x_3^2 + 14x_2 - 10x_3$;

(e) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = x_1x_2(1 - x_1)(2 - x_2)$;

(f) $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = y^2 + z^2 + 2xy$;

(g) $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = -2x^2 + 3y^2 + z^2 + 2yz$;

(h) $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x) = 2x_1^2 + x_2^2 + x_3^2 + x_1x_2$;

(i) $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x) = -3x_1^2 - 2x_2^2 - \frac{1}{8}x_3^2 + x_1x_3$;

(j) $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x) = -4x_1^2 - 5x_2^2 - \frac{1}{8}x_3^2 + 2x_1x_3$;

(k) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = e^{2x_1+3x_2}(8x_1^2 - 6x_1x_2 + 3x_2^2)$;

(l) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = e^{-(x+y)^2}$.

[168] Przynajmniej 3 przykłady, w tym conajmniej 1 dla $\mathbb{R}^3$.

X Teoria miary i całki. Rachunek całkowy wielu zmiennych

[około $5\frac{1}{2}$ wykładu]

W tym rozdziale poznamy ogólniejszy i „wygodniejszy” rodzaj całki niż poznana w rozdziale VII całka Riemanna, a mianowicie *całkę Lebesgue'a*. Użyjemy jej tu do całkowania funkcji wielu zmiennych. Zaczniemy jednak od ujęcia abstrakcyjnego.

1. Miara i całka względem miary.

Omówimy pokrótce abstrakcyjne podejście do całkowania, czyli tzw. *teorię Lebesgue'a miary i całki*. Przy tym podejściu można będzie całkować funkcje względem pewnej wybranej przez nas *miary*, czyli pewnego wybranego sposobu mierzenia zbiorów. Musimy zatem zacząć nasze rozważania od definicji miary. Opiszemy tam własności takiego sposobu mierzenia „rozmiaru” zbiorów, który by niejako uogólniał znane dobrze (choć nie na ścisłym poziomie) pojęcia długości, pola powierzchni i objętości. Będzie to Państwu nieco przypominać abstrakcyjną definicję metryki — tam uogólnialiśmy pojęcie odległości, tu natomiast uogólnimy pojęcie pola powierzchni, czy objętości zbioru. Jest jednak pewna ważna różnica pomiędzy tymi dwoma rodzajami uogólnień. Metryka pozwalała na mierzenie odległości pomiędzy **każdymi** dwoma punktami zbioru, tymczasem miara dość często będzie mierzyła tylko **niektóre** podzbiory danego zbioru. Rodzinę (czyli zbiór) tych podzbiorów będziemy nazywali rodziną *zbiorów mierzalnych*. Powodem takiej sytuacji jest najczęściej nie tyle chęć niemierzenia niektórych podzbiorów, co raczej niemożność ich zmierzenia, choć może to się wydawać zaskakujące...

1.1. Sigma–ciała

Rodzina zbiorów mierzalnych nie może być całkiem dowolną rodziną podzbiorów ustalonego zbioru X. Musi to być rodzina zamknięta ze względu na wszystkie przeliczalne operacje na zbiorach i $\emptyset$ musi do niej należeć. Taką rodzinę zbiorów nazywamy *σ–ciałem*[169)]. Precyzyjna definicja jest następująca.

Definicja. *Rodzina $\mathbb{M}$ podzbiorów zbioru X jest σ–**ciałem** (podzbiorów X) wtw*

1. $\emptyset \in \mathbb{M}$,
2. $\forall_{A\in\mathbb{M}}\ X \setminus A \in \mathbb{M}$,
3. *Dla każdego ciągu zbiorów* $\{A_n\}_{n\geqslant 1}$ *takich, że* $\forall_{n\in\mathbb{N}}\ A_n \in \mathbb{M}$ *zachodzi* $\displaystyle\bigcup_{n\in\mathbb{N}} A_n \in \mathbb{M}$.

Korzystając ze znanych własności działań na zbiorach nietrudno wykazać poniższy wniosek (patrz zad. X.6).

Wniosek. *Jeżeli $\mathbb{M}$ jest σ–ciałem oraz $A, B, A_n \in \mathbb{M}$ dla każdego $n \in \mathbb{N}$, to*

1. $A \setminus B, A \cup B, A \cap B \in \mathbb{M}$,
2. $\displaystyle\bigcap_{n\in\mathbb{N}} A_n \in \mathbb{M}$.

Oczywiście do σ–ciała musi zawsze należeć także zbiór X (na mocy punktów 1. i 2. definicji).

Przykłady (najprostsze).

[169)] Czyt.: „sigma–ciało”; niektórzy używają zamiast tego nazwy σ–*algebra*.

1. 2^X — rodzina wszystkich podzbiorów zbioru X jest σ–ciałem — to największe σ–ciało podzbiorów X.

2. Druga skrajność, to $\{\emptyset, X\}$ — to najmniejsze σ–ciało podzbiorów X.

3. Niech $X = \{1, 2, 3\}$. rodzina $\mathbb{M} = \{\emptyset, \{1\}, \{2, 3\}, X\}$ jest σ–ciałem podzbiorów X (różnym od dwóch poprzednich przykładów).

Jednak tak naprawdę najważniejsze przykłady σ–ciał są na ogół trudne do opisania w jawnej formie. Dlatego wygodnie jest używać pojęcia *σ–ciała generowanego* przez pewną rodzinę $\mathbb{A}$ podzbiorów X (tzn. $\mathbb{A} \subset 2^X$). To σ–ciało będziemy oznaczać przez $\sigma(\mathbb{A})$ i definiujemy jako najmniejsze (w sensie zawierania zbiorów) σ–ciało podzbiorów X zawierające rodzinę $\mathbb{A}$. To, że dla dowolnej rodziny $\mathbb{A}$ takie najmniejsze σ–ciało istnieje jest łatwe do wykazania. Niech bowiem $M_{\mathbb{A}}$ oznacza zbiór wszystkich σ–ciał podzbiorów X zawierających $\mathbb{A}$ [170)]. Wówczas

$$\bigcap_{\mathbb{M} \in M_{\mathbb{A}}} \mathbb{M}$$

jest rodziną podzbiorów X, która zawiera $\mathbb{A}$ (bo każde $\mathbb{M} \in M_{\mathbb{A}}$ zawierało $\mathbb{A}$) i jest σ–ciałem, bo przecięcie dowolnego zbioru σ–ciał jest σ–ciałem (wynika to natychmiast z definicji σ–ciała). Na dodatek ta rodzina zawiera się w każdym $\mathbb{M} \in M_{\mathbb{A}}$ na mocy swej definicji — jest to więc właśnie $\sigma(\mathbb{A})$.

Przykłady.

1. $X = \mathbb{N}$, $\mathbb{A} = \{\{n\} : n \in \mathbb{N}\}$ — rodzina wszystkich jednoelementowych podzbiorów zbioru $\mathbb{N}$. Wówczas $\sigma(\mathbb{A}) = 2^{\mathbb{N}}$, gdyż każdy podzbiór zbioru $\mathbb{N}$ jest skończoną lub przeliczalną sumą elementów rodziny $\mathbb{A}$.

2. $X = \mathbb{R}^d$, $\mathbb{A}$ — rodzina wszystkich otwartych podzbiorów $\mathbb{R}^d$. Wówczas $\sigma(\mathbb{A})$ nazywane jest rodziną *zbiorów borelowskich* w $\mathbb{R}^d$ (a jego elementy to *zbiory borelowskie*) i oznaczane przez $\mathcal{B}\left(\mathbb{R}^d\right)$. Można wykazać, choć wcale nie jest to tak bardzo proste, że $\mathcal{B}\left(\mathbb{R}^d\right) \neq 2^{\mathbb{R}^d}$.

Każda rodzina zbiorów $\mathbb{A}$ o tej własności, że $\sigma(\mathbb{A}) = \mathbb{M}$ dla danego σ–ciała $\mathbb{M}$ nosi nazwę rodziny *generatorów* σ–ciała $\mathbb{M}$. Rodzina generatorów pełni rolę nieco podobną do roli bazy przestrzeni liniowej. Nie jest ona też na ogół jednoznacznie wyznaczona przez $\mathbb{M}$.

Przykład. Poza rodziną wszystkich otwartych podzbiorów $\mathbb{R}$, także każda z poniższych rodzin generuje (tzn. jest rodziną generatorów dla) $\mathcal{B}\left(\mathbb{R}^d\right)$ (patrz zad. X.7):

(i) rodzina wszystkich domkniętych podzbiorów $\mathbb{R}$,

(ii) $\{(a; b) : a, b \in \mathbb{R}\}$;

(iii) $\{[a; b] : a, b \in \mathbb{R}\}$;

(iv) $\{(-\infty; a) : a \in \mathbb{R}\}$;

(v) $\{(a; +\infty) : a, b \in \mathbb{Q}\}$.

Oczywiście tego typu rodzin generatorów jest znacznie więcej (patrz zad. X.7).

170) Oczywiście $M_{\mathbb{A}} \neq \emptyset$, bo np. $2^X \in M_{\mathbb{A}}$.

1.2. Miary.

Niech $\mathbb{M}$ będzie pewnym σ–ciałem podzbiorów zbioru X. $\mathbb{M}$ będziemy traktować jako rodzinę zbiorów mierzalnych, czyli możliwych do zmierzenia (przy użyciu miary). Ale tak jak zapowiadaliśmy, nie będziemy rozważać tu na razie żadnej konkretnej miary, lecz podamy definicję abstrakcyjną, tj. określającą warunki (aksjomaty), które uznajemy za niezbędne do tego by dany sposób pomiaru nazwać miarą.

Definicja. *Funkcja* $\mu : \mathbb{M} \to [0; +\infty]$ [171] *jest* ***miarą*** *wtw*

1. $\mu(\emptyset) = 0$,
2. *(przeliczalna addytywność) dla każdego ciągu* $\{A_n\}_{n\geqslant 1}$ *parami rozłącznych (tj.* $\forall_{\substack{m,n\in\mathbb{N}\\ m\neq n}} A_m \cap A_n = \emptyset$*) zbiorów mierzalnych (tj.* $\forall_{n\in\mathbb{N}} A_n \in \mathbb{M}$*) zachodzi*

$$\mu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n=1}^{+\infty} \mu(A_n) \text{ [172]}. \tag{X.1}$$

Zauważmy, że zbiór, którego miara znajduje się po lewej stronie (X.1) jest mierzalny, dzięki założeniu, że $\mathbb{M}$ jest σ–ciałem.

Zanim zajmiemy się przykładami miar, wypiszemy kilka podstawowych własności każdej miary, będących konsekwencją powyższej definicji. Zakładamy tu, ze μ jest *miarą w* X, tzn. że μ jest miarą określoną na pewnym σ–ciele $\mathbb{M}$ podzbiorów X.

Fakt.

1. *(σ–podaddytywność) Jeżeli* $A_n \in \mathbb{M}$ *dla każdego* $n \in \mathbb{N}$*, to*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n=1}^{+\infty} \mu(A_n).$$

2. *(skończona addytywność) Jeżeli* $A_1, \ldots, A_n \in \mathbb{M}$ *i są to zbiory parami rozłączne, to*

$$\mu\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_k).$$

3. *(monotoniczność) Jeżeli* $A, B \in \mathbb{M}$ *i* $A \subset B$*, to* $\mu(A) \leqslant \mu(B)$.

Nietrudny dowód powyższego faktu zastawiam Państwu jako zadanie (patrz zad. X.15).

Przykłady (najprostsze).

1. (miara zerowa) Niech $\mathbb{M}$ będzie dowolnym σ–ciałem podzbiorów X. Funkcja $\mu : \mathbb{M} \to \mathbb{R}$ stale równa zero jest oczywiście miarą (ale inne funkcje stałe na $\mathbb{M}$ na miarę się nie nadają...).
2. (delta Diraca) Niech $x_0 \in X$ i niech $\mathbb{M}$ będzie pewnym σ–ciałem podzbiorów X (czasami w tym przypadku zakłada się dodatkowo, że $\{x_0\} \in \mathbb{M}$, choć nie jest to niezbędne). Funkcja $\delta_{x_0} : \mathbb{M} \to \mathbb{R}$ zadana dla $A \in \mathbb{M}$ wzorem

$$\delta_{x_0}(A) = \begin{cases} 1 & \text{gdy } x_0 \in A \\ 0 & \text{gdy } x_0 \notin A \end{cases}$$

nazywa się *deltą Diraca* w x_0 i jest ona miarą (patrz zad. X.13). Najczęściej używa się tej miary dla przypadku gdy $X = \mathbb{R}^d$, $x_0 = 0$ i $\mathbb{M} = \mathcal{B}(\mathbb{R}^d)$.

---

[171] Ten „przedział" jest „domknięty" z prawej strony tzn. chodzi o zbiór $[0; +\infty) \cup \{+\infty\}$.

[172] Jeśli choć jedna spośród miar $\mu(A_n)$ równa jest $+\infty$, to sumę tę definiujemy jako $+\infty$, w przeciwnym wypadku jest to zwykła suma szeregu — posiada on sumę (skończoną lub nie), gdyż $\mu(A_n) \geqslant 0$.

3. (miara licząca) Rozważamy znów dowolne X oraz $\mathbb{M} = 2^X$. Niech $\# : 2^X \to \mathbb{N}_0 \cup \{+\infty\}$ będzie zadana dla $A \subset X$ następująco:

$$\#(A) = \begin{cases} \text{liczba elementów zbioru } A, \text{ gdy } A \text{ — zbiór skończony} \\ +\infty, \text{ gdy } A \text{ — zbiór nieskończony.} \end{cases}$$

Funkcja ta jest miarą, tzw. *miarą liczącą* (patrz zad. X.13).

Wyróżnia się rozmaite szczególne rodzaje miar. Na razie określimy dwa z nich.

Miara μ jest *skończona* wtw $\mu(X) < +\infty$. Dwa pierwsze spośród powyższych przykładów to miary skończone, natomiast miara licząca jest skończona wtw X jest zbiorem skończonym.

Miara μ jest *zupełna* wtw każdy podzbiór zbioru miary zerowej jest zbiorem mierzalnym[173]. Oczywiście każda miara określona na σ–ciele 2^X jest zupełna, ale nie jest to wcale warunek konieczny zupełności. Można wykazać, że każda miara daje się uzupełnić, tj. rozszerzyć do pewnego, być może większego σ–ciała, tak, że to rozszerzenie będzie już miarą zupełną (patrz zad. X.22).

Zbiory miary zero pełnią szczególną rolę w teorii miary — to takie zbiory, które w pewnym sensie są „nieistotne". Ściślej wyraża to poniższy rezultat.

Fakt.

1. *Skończone i przeliczalne operacje sumy, przecięcia oraz — dla dwóch zbiorów — różnicy na zbiorach miary zero dają zbiór miary zero.*

2. *Jeżeli A i Z są mierzalne oraz $\mu(Z) = 0$, to $\mu(A \cup Z) = \mu(A \setminus Z) = \mu(A)$.*

Dowód (część).
Wykażemy tylko, że jeżeli $\mu(A_n) = 0$ dla wszystkich $n \in \mathbb{N}$, to $\mu(\bigcup_{n\in\mathbb{N}} A_n) = 0$. Na mocy σ–podaddytywności miary μ mamy

$$0 \leqslant \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n=1}^{+\infty} \mu(A_n) = \sum_{n=1}^{+\infty} 0 = 0.$$

□

1.3. Miara Lebesgue'a.

Wspomniane wcześniej przykłady miar były dość proste i prawdę powiedziawszy, dla nas niezbyt ważne. Główna miara jakiej potrzebujemy to taka, która miałaby coś wspólnego z mierzeniem długości w przypadku podzbiorów $\mathbb{R}^1$, powierzchni — w przypadku $\mathbb{R}^2$ oraz objętości — w przypadku $\mathbb{R}^3$. Ogólnie, w przypadku $\mathbb{R}^d$, będzie to tzw. d–wymiarowa *miara Lebesgue'a.* Podanie ścisłej definicji takiej miary jest niestety sprawą dość trudną. Każdy znany dotąd sposób jej (a właściwie — ich, ze względu na dowolność wymiaru d) definiowania wymaga jednoczesnego dowodzenia wielu nietrywialnych faktów. A zatem my ograniczymy się tylko do samego sformułowania twierdzenia o istnieniu potrzebnej nam miary, przy dowolnie wybranym wymiarze d.

Twierdzenie X.1 (o istnieniu miary Lebesgue'a). *Niech $d \in \mathbb{N}$. Istnieje dokładnie jedna miara μ, spełniająca poniższe warunki:*

1. *μ jest określona na pewnym σ–ciele $\mathbb{M}$ podzbiorów $\mathbb{R}^d$ takim, że $\mathcal{B}\ (\mathbb{R}^d) \subset \mathbb{M}$;*

2. *$\mu([a_1; b_1] \times \ldots \times [a_d; b_d]) = (b_1 - a_1) \cdot \ldots \cdot (b_d - a_d)$ dla dowolnych $a_1 \leqslant b_1, \ldots, a_d \leqslant b_d$;*

[173] A zatem podzbiór ten musi też mieć automatycznie zerową miarę na mocy monotoniczności miary.

3. *(przesuwalność)* $\forall_{A\in\mathbb{M}}\ \forall_{x\in\mathbb{R}^d}\ \mu(x+A)=\mu(A)$ [174] *;*

4. *μ jest zupełna oraz każdy zbiór z $\mathbb{M}$ jest sumą pewnego zbioru borelowskiego i podzbioru pewnego zbioru borelowskiego o mierze μ równej zero.*[175]

B.D.

Miarę spełniającą warunki tego twierdzenia nazywamy d–wymiarową *miarą Lebesgue'a*, a przy $d=1$ także, po prostu, miarą Lebesgue'a. Będziemy ją oznaczać przez λ^d, a gdy $d=1$, także przez λ.

W pewnym sensie najważniejszą spośród własności miary λ^d jest własność 2. z powyższego twierdzenia. Dzięki przeliczalnej addytywności miary pozwala ona na „wyliczenie" miary każdego zbioru, który jest przeliczalną sumą parami rozłącznych „kostek" tego typu (tj. o krawędziach równoległych do osi). Zauważmy przy tym, że tak naprawdę nie jest istotne, by kostki te były domknięte — ich miara nie zależy od tego, czy któraś ze „ścian" zawiera się w nich, czy nie. Np., dla $d=2$ mamy

$$\lambda^2([0;1]\times[0;1])=\lambda^2((0;1)\times(0;1))=1,$$

gdyż

$$[0;1]\times[0;1]=((0;1)\times(0;1))\cup(\{0\}\times[0;1])\cup(\{1\}\times[0;1])\cup([0;1]\times\{0\})\cup([0;1]\times\{1\})$$

przy czym każdy z czterech zbiorów dodanych do $(0;1)\times(0;1)$ jest na mocy pkt. 2. twierdzenia X.1 zbiorem miary 0. Większość zbiorów, z którymi mamy w praktyce do czynienia to zbiory, które są równe wcześniej wspomnianej przeliczalnej sumie kostek, z dokładnością ewentualnie do pewnego zbioru miary zero.

Możliwość zmierzenia przy pomocy miary Lebesgue'a rzeczywiście pokaźnej liczby zbiorów zapisana jest w warunku 1 twierdzenia X.1 — mierzalne bowiem są tu przynajmniej wszystkie zbiory borelowskie. A znalezienie zbioru nieborelowskiego — to prawdziwa sztuka (choć jest to możliwe, jak już wcześniej wspomnieliśmy, patrz zad. X.23). Sigma ciało, na którym określona jest d–wymiarowa miara Lebesgue'a, zwane też *σ–ciałem zbiorów mierzalnych w sensie Lebesgue'a (w $\mathbb{R}^d$)* — oznaczamy je $\mathcal{L}(\mathbb{R}^d)$, jest „nieco" większe od σ–ciała $\mathcal{B}(\mathbb{R}^d)$, ale jak mówi warunek 4 twierdzenia X.1, każdy taki zbiór mierzalny różni się od pewnego zbioru borelowskiego o zbiór miary (miary λ^d) zerowej.

1.4. Funkcje mierzalne

Podobnie jak nie każdy podzbiór $\mathbb{R}^d$ da się zmierzyć miarą Lebesgue'a, tak nie każdą funkcję da się scałkować względem miary λ^d. Zresztą z tego typu trudnością spotkaliśmy się już przy teorii całki Riemanna. Jak już wkrótce przekonamy się, przy całkowaniu względem miary λ^1 ograniczenia dotyczące całkowanych funkcji są jednak dużo mniejsze niż przy całkowaniu w sensie Riemanna (to jeden z powodów, dla których teoria całki względem miary Lebesgue'a okazuje się „lepsza" niż teoria całki Riemanna). Pewnych ograniczeń jednak uniknąć się nie da. Dla całki Lebesgue'a ograniczenia te są dwóch typów: „jakościowe" i „ilościowe". To pierwsze związane jest z potrzebą tzw. *mierzalności funkcji*, o drugim powiemy już przy samej definicji całki.

Niech $\mathbb{M}$ będzie pewnym σ–ciałem podzbiorów X oraz niech $f: D\to\mathbb{R}$, gdzie $D\subset X$.

Definicja. *Funkcja f jest* ***mierzalna*** *wtw*

$$\forall_{A\in\mathcal{B}(\mathbb{R})}\ f^{-1}(A)\in\mathbb{M}.^{176)} \tag{X.2}$$

[174] $x+A:=\{x+a\in\mathbb{R}^d : a\in A\}$.

[175] Oznacza to, że μ jest wspomnianym wcześniej uzupełnieniem obcięcia μ do $\mathcal{B}(\mathbb{R}^d)$.

Jak widać, w definicji tej istotny jest wybór σ–ciała $\mathbb{M}$, natomiast w ogóle nie jest jeszcze potrzebna miara (stąd tu mówimy o, jedynie, „jakościowym”, a nie „ilościowym” ograniczeniu). Czasami, aby wyraźnie zaznaczyć o jaki wybór σ–ciała chodzi w danym przypadku, będziemy zamiast „mierzalna” mówić $\mathbb{M}$–*mierzalna*, a dla $\mathbb{M} = \mathcal{L}\ (\mathbb{R}^d)$ także *mierzalna w sensie Lebesgue'a*.

Choć całkować będziemy jedynie funkcje o wartościach liczbowych (ewentualnie o wartościach w $\overline{\mathbb{R}}$, ale o tym później), z kilku względów warto jest pojęcie mierzalności uogólnić też na przypadek funkcji o innych wartościach. Załóżmy mianowicie, że $f : D \to X'$, $D \subset X$, przy czym zadane jest także pewne σ–ciało $\mathbb{M}'$ podzbiorów zbioru X'. Wówczas f jest *mierzalna*, albo precyzyjniej: $\mathbb{M}, \mathbb{M}'$ – *mierzalna* wtw $\forall_{A \in \mathbb{M}'}\ f^{-1}(A) \in \mathbb{M}$. A zatem o poprzednim, szczególnym, przypadku mierzalności funkcji skalarnej możemy teraz powiedzieć, że była to $\mathbb{M}$, $\mathcal{B}(\mathbb{R})$ – mierzalność w sensie tej ogólniejszej terminologii.

Definicja mierzalności była co prawda krótka, ale dość niemiła do praktycznego użytku. Trzeba bowiem sprawdzać pewien warunek dla wszystkich zbiorów $A \in \mathbb{M}'$, podczas gdy opisanie postaci dowolnego elementu σ–ciała $\mathbb{M}'$ może być, jak wiemy, bardzo trudne. Na szczęście sprawa się bardzo upraszcza, gdy umiemy opisać postać generatorów σ–ciała $\mathbb{M}'$.

Fakt. *Jeżeli* $\mathbb{M}' = \sigma(\mathbb{A})$ *i* $D \in \mathbb{M}$, *to* f *jest* $\mathbb{M}$, $\mathbb{M}'$ *– mierzalna wtw*

$$\forall_{A \in \mathbb{A}}\ f^{-1}(A) \in \mathbb{M}. \tag{X.3}$$

Dowód.
Oczywiście implikacja „$\Rightarrow$” zachodzi, gdyż $\mathbb{A} \subset \mathbb{M}'$. Wykażemy „$\Leftarrow$”. Załóżmy zatem (X.3) i zdefiniujmy $\mathbb{M}_f := \{A \subset X' : f^{-1}(A) \in \mathbb{M}\}$. Dzięki własnościom przeciwobrazu funkcji łatwo wykazać, że $\mathbb{M}_f$ jest σ–ciałem podzbiorów zbioru X'. Np. jeżeli $\forall_{n\in\mathbb{N}}\ A_n \in \mathbb{M}_f$, to $f^{-1}(\bigcup_{n\in\mathbb{N}} A_n) = \bigcup_{n\in\mathbb{N}} f^{-1}(A_n) \in \mathbb{M}$, bo $\mathbb{M}$ jest σ–ciałem oraz $f^{-1}(A_n) \in \mathbb{M}$ dla każdego n. A zatem $\bigcup_{n\in\mathbb{N}} A_n \in \mathbb{M}_f$. Ponadto (X.3) oznacza, że $\mathbb{A} \subset \mathbb{M}_f$. Stąd $\sigma(\mathbb{A}) \subset \mathbb{M}_f$ gdyż $\sigma(\mathbb{A})$ jest najmniejszym w sensie „$\subset$” σ–ciałem podzbiorów X' zawierającym $\mathbb{A}$, a $\mathbb{M}_f$ — „tylko” pewnym σ–ciałem podzbiorów X' zawierającym $\mathbb{A}$. A więc $\mathbb{M}' \subset \mathbb{M}_f$, co oznacza właśnie, że f jest $\mathbb{M}$, $\mathbb{M}'$ mierzalna. □

Warto zapamiętać rozumowanie użyte w powyższym dowodzie, gdyż jest ono charakterystyczne dla sytuacji, w których pojawiają się σ–ciała generowane przez pewne rodziny zbiorów.

Przykłady.

1. Funkcje ciągłe. Rozważamy $f : D \to \mathbb{R}$, gdzie $D \subset \mathbb{R}^d$, dla $\mathbb{R}^d$ rozważamy σ–ciało $\mathcal{L}\ (\mathbb{R}^d)$ zbiorów mierzalnych w sensie Lebesgue'a i $D \in \mathcal{L}\ (\mathbb{R}^d)$. Jeżeli f jest ciągła, to f jest mierzalna (tzn. $\mathcal{L}\ (\mathbb{R}^d)$, $\mathcal{B}\ (\mathbb{R})$ – mierzalna). Wynika to z tego, że dla $\mathcal{B}(\mathbb{R})$ (z definicji) rodzina zbiorów otwartych jest zbiorem generatorów, a skoro f — ciągła, to dla U — otwartego w $\mathbb{R}$ zbiór $f^{-1}(U)$ ma na mocy twierdzenia VIII.3 postać $V \cap D$, gdzie V — otwarty w $\mathbb{R}^d$. A zatem $V \in \mathcal{B}\ (\mathbb{R}^d) \subset \mathcal{L}\ (\mathbb{R}^d)$, skąd $f^{-1}(U) = V \cap D \in \mathcal{L}\ (\mathbb{R}^d)$. Jak z tego dowodu widać, gdy f — ciągła i $D \in \mathcal{B}(\mathbb{R}^d)$, to f jest nawet $\mathcal{B}(\mathbb{R}^d)$ – mierzalna.

2. Funkcje charakterystyczne zbiorów mierzalnych. Niech $\mathbb{M}$ będzie σ–ciałem podzbiorów X, $A \in \mathbb{M}$ i rozważmy χ_A – *funkcję charakterystyczną* zbioru A, tj. funkcję z X w $\mathbb{R}$ zadaną wzorem:
$$\chi_A(x) = \begin{cases} 1 & \text{dla } x \in A \\ 0 & \text{dla } x \notin A. \end{cases}$$

[176] A zatem, m. in., od dziedziny f wymagamy jedynie mierzalności ($D = f^{-1}(\mathbb{R})$). W przypadku miary λ^1 to znaczenie słabszy warunek niż dla całkowalności w sensie Riemanna — tam dziedzina musiała być przedziałem domkniętym.

I choć takiej funkcji na ogół daleko do ciągłości, to jest ona mierzalna. Niech bowiem B będzie jakimkolwiek podzbiorem $\mathbb{R}$ (nie koniecznie nawet borelowskim). Wówczas $\chi_A^{-1}(B)$ jest jednym z następujących zbiorów: $\emptyset$, A, $X \setminus A$, X (pierwsza sytuacja — gdy $0, 1 \notin B$, druga — gdy $1 \in B$ ale $0 \notin B$, trzecia — gdy $0 \in B$, ale $1 \notin B$ i ostatnia — gdy $0, 1 \in B$). A zatem χ_A jest mierzalna (tzn. $\mathcal{B}(\mathbb{R})$ – mierzalna, tj. $\mathbb{M}$, $\mathcal{B}(\mathbb{R})$ – mierzalna), bo każdy z tych czterech zbiorów należy do $\mathbb{M}$. Ale co więcej χ_A jest też $2^{\mathbb{R}}$ – mierzalna (tj. $\mathbb{M}$, $2^{\mathbb{R}}$ – mierzalna). Analogiczny wynik uzyskamy rozważając χ_A jako funkcję określoną nie na całym X, ale na dowolnym $D \in \mathbb{M}$, gdy $A \subset D$.

Załóżmy teraz, że $\mathbb{M}, \mathbb{M}', \mathbb{M}''$ są σ–ciałami podzbiorów odpowiednio X, X', X'' oraz $D \in \mathbb{M}$, $D' \in \mathbb{M}'$. Jak się okazuje, mierzalność względem odpowiednich σ–ciał zachowuje się przy składaniu funkcji.

Fakt. *Jeżeli $f : D \to D'$ jest $\mathbb{M}$, $\mathbb{M}'$ – mierzalna oraz $g : D' \to X''$ jest $\mathbb{M}'$, $\mathbb{M}''$ – mierzalna, to $g \circ f$ jest $\mathbb{M}$, $\mathbb{M}''$ – mierzalna (patrz rys.* **??***).*

$$\begin{tikzcd}
X \supset D \arrow[r, "f"'] \arrow[rr, bend left, "g \circ f"] & D' \arrow[r, "g"'] & X'' \\
\mathbb{M} & \cap & \mathbb{M}'' \\
& X' & \\
& \mathbb{M}' &
\end{tikzcd}$$

Rysunek 30:

Dowód.
Niech $A \in \mathbb{M}''$. Mamy $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A))$, ale $g^{-1}(A) \in \mathbb{M}'$ na mocy mierzalności g, a zatem $f^{-1}(g^{-1}(A)) \in \mathbb{M}$ na mocy mierzalności f. $\square$

Uwaga. Wbrew pozorom jednak ogólnie nie jest prawdą, że złożenie dwóch funkcji mierzalnych jest funkcją mierzalną, jeżeli składamy funkcje rzeczywiste zmiennej rzeczywistej oraz mierzalność **obu** funkcji rozumiemy jako $\mathcal{L}(\mathbb{R})$, $\mathcal{B}(\mathbb{R})$–mierzalność. By mieć pewność, że złożenie takie będzie mierzalne, powinniśmy zakładać, że funkcja zewnętrzna jest $\mathcal{B}(\mathbb{R})$, $\mathcal{B}(\mathbb{R})$–mierzalna[177)] , co jest nieco silniejszym warunkiem niż ten wcześniejszy (gdyż $\mathcal{B}(\mathbb{R}) \subset \mathcal{L}(\mathbb{R})$, ale równość nie zachodzi). Na szczęście jest to też warunek „prawie zawsze" spełniony w „typowych sytuacjach".

Wiemy już jak zachowuje się własność mierzalności przy składaniu funkcji — czas teraz na inne operacje.

Fakt. *Suma, iloczyn, różnica i iloraz funkcji skalarnych mierzalnych (tzn. $\mathbb{M}$–mierzalnych) są mierzalne (w przypadku ilorazu zakładamy oczywiście, że dzielimy przez funkcję o wartościach różnych od 0).*

Dowód.
Niech $f_1, f_2 : D \to \mathbb{R}$ będą mierzalne, gdzie $D \subset X$ oraz w X rozważamy σ–ciało $\mathbb{M}$. Wykażemy, ze $f_1 + f_2$ jest mierzalna. W tym celu rozważmy $f : D \to \mathbb{R}^2$ zadaną wzorem $f(x) = (f_1(x), f_2(x))$.

Zauważmy najpierw, że f jest $\mathbb{M}$, $\mathcal{B}$ $(\mathbb{R}^2)$–mierzalna. Wynika to z faktu, że w $\mathcal{B}$ $(\mathbb{R}^2)$ można jako zbiór generatorów wybrać rodzinę wszystkich prostokątów postaci $[a; b] \times [c; d]$ (patrz zad. X.10). Gdy to już wiemy, na mocy faktu ze str. 158 wystarczy sprawdzić, że $f^{-1}([a; b] \times [c; d]) \in \mathbb{M}$. Ale $f^{-1}([a; b] \times [c; d]) = \{x \in X_1 : f_1(x) \in [a; b] \text{ i } f_2(x) \in [c; d]\} = f_1^{-1}([a; b]) \cap f_2^{-1}([c; d]) \in \mathbb{M}$. Potrzebne było też by $D \in \mathbb{M}$, ale to mamy zagwarantowane np. przez mierzalność f_1.

177) Takie funkcje noszą nazwę *borelowskich.* Np. każda funkcja ciągła określona na zbiorze borelowskim (tj. z $\mathcal{B}(\mathbb{R})$) jest oczywiście borelowska — patrz przykład 1.

Teraz zauważmy, że $f_1 + f_2 = p \circ f$, gdzie $p : \mathbb{R}^2 \to \mathbb{R}$ to funkcja opisująca dodawanie liczb, tzn. $p(x) = x_1 + x_2$. Ponieważ p jest ciągła, zatem w szczególności jest też oczywiście $\mathcal{B}(\mathbb{R}^2)$, $\mathcal{B}(\mathbb{R})$–mierzalna (patrz przykład 1.). A zatem na mocy faktu dotyczącego składania funkcja $f_1 + f_2$ jest mierzalna.

Określając teraz odpowiednio inaczej funkcję p uzyskamy dowody dla pozostałych działań. □

Czasami, oprócz funkcji czysto skalarnych, wygodnie jest rozważać funkcje o wartościach w $\overline{\mathbb{R}}$. Nietrudno jest rozszerzyć pojęcie mierzalności na takie funkcje. Mianowicie $f : D \to \overline{\mathbb{R}}$ jest *mierzalna* wtw zachodzi (X.2) oraz $f^{-1}(\{+\infty\}), f^{-1}(\{-\infty\}) \in \mathbb{M}$.

Sformułowany przed chwilą fakt dotyczący podstawowych działań na funkcjach pozostaje prawdziwy także dla tak rozszerzonego pojęcia mierzalności, pod warunkiem, że działania na rozważanych funkcjach będą określone (tzn. dla każdego x odpowiednie działanie dla elementów $f_1(x)$ i $f_2(x)$ z $\overline{\mathbb{R}}$ będą określone w sensie zdefiniowanym na stronie 20).

Sformułujemy jeszcze jeden fakt, dotyczący granicy ciągu funkcji mierzalnych. Dowód (nietrudny!) jest zadaniem dla Państwa (patrz zad. X.27).

Fakt. *Jeżeli ciąg funkcyjny $\{f_n\}_{n \geqslant n_0}$ jest punktowo zbieżny do funkcji f i wszystkie jego wyrazy f_n są funkcjami mierzalnymi, to f też jest mierzalna.*

Także ten fakt pozostaje prawdziwy dla funkcji o wartościach w $\overline{\mathbb{R}}$, o ile tylko rozszerzymy (w oczywisty chyba sposób) pojęcie zbieżności punktowej na ciągi funkcyjne tego rodzaju.

1.5. Całka względem miary

Określimy tu wreszcie zapowiadane już od dłuższego czasu pojęcie całki względem dowolnej miary. Taki typ całki nosi też nazwę całki Lebesgue'a.[178)] Definicja całki będzie składała się z trzech etapów. Najpierw zdefiniujemy całkę dla tzw. *funkcji prostych* (tych, dla których będzie ona określona), następnie dla funkcji mierzalnych nieujemnych i wreszcie ogólniej, dla wszystkich funkcji mierzalnych, dla których będzie to możliwe.

Zanim przystąpimy do definicji, przyjmiemy najpierw jedną ważną umowę, która uprości nam zapis rozmaitych formuł przy uprawianiu tej teorii. Przyjmujemy mianowicie, że **na użytek teorii miary**:

$$0 \cdot \pm\infty = \pm\infty \cdot 0 = 0. \tag{X.4}$$

Ta umowa może Państwa nieco bulwersować — pamiętają Państwo zapewne, że już w rozdziale II działania typu „0 razy nieskończoność" zostały uznane za nieokreślone (zgodnie zresztą z doświadczeniem płynącym z teorii granicy ciągu — patrz np. zadanie II.7). I to się nie zmienia — nadal uznajemy nieokreśloność tego działania, a wzór (X.4) będziemy traktować jedynie jako **umowę**, która parokrotnie skróci nam po prostu zapis formuł.

Załóżmy, że $(X, \mathbb{M}, \mu)$ jest *przestrzenią mierzalną*, tzn. że $\mathbb{M}$ jest σ–ciałem podzbiorów (mierzalnych) X oraz μ — miarą określoną na $\mathbb{M}$.

Etap 1. Funkcję $f : D \to \mathbb{R}$ nazywamy *funkcją prostą* wtw f jest mierzalna i jej zbiór wartości jest skończony. Załóżmy, że zbiorem wartości funkcji f jest $\{c_1, \ldots, c_n\}$, gdzie $c_i \neq c_j$ dla $i \neq j$, i niech $A_i := \{x \in D : f(x) = c_i\}$, tzn. $A_i = f^{-1}(\{c_i\})$ — w szczególności więc A_i są mierzalne i tworzą rozbicie D na rozłączne zbiory, tzn. $A_i \cap A_j = \emptyset$ dla $i \neq j$ oraz $\bigcup_{i=1}^{n} A_i = D$.

[178)] Jest pewna niekonsekwencja w tym nazewnictwie. Nazwisko Lebesgue'a pojawia się zarówno przy omawianej tu ogólnej teorii miary i całki, jak i przy samym pojęciu całki, ale w ogólnym przypadku (dla dowolnej miary). Tymczasem nazwa *miara Lebesgue'a* odnosi się jedynie do poznanych już przez nas konkretnych miar λ^d określonych dla podzbiorów $\mathbb{R}^d$.

Jest więc chyba naturalne, że całkę z f powinniśmy określić wzorem:

$$S(f)\ ^{179)} := \sum_{i=1}^{n} c_i \cdot \mu(A_i),$$

o ile suma powyższa jest *określona*, tzn. nie występuje w niej jednocześnie $+\infty$ i $-\infty$. W szczególności na pewno $S(f)$ jest określone, gdy dodatkowo $f \geqslant 0$. Zauważmy też, że tu po raz pierwszy może zaistnieć konieczność użycia umowy (X.4), gdy zdarzy się, że $c_i = 0$ oraz $\mu(A_i) = +\infty$ dla pewnego i (na odwrót zdarzyć się nie może, bo założyliśmy, że f ma wartości w $\mathbb{R}$ a nie w $\overline{\mathbb{R}}$). Oczywiście może się także zdarzyć, że wartość $S(f)$ będzie określona, ale nieskończona.

Warto jeszcze wspomnieć o tym, że choć funkcje proste wydają się być (jak sama ich nazwa na to wskazuje) mało skomplikowane, to jednak każda funkcja mierzalna g jest granicą pewnego zbieżnego punktowo ciągu funkcji prostych $\{g_n\}$, a jeśli dodatkowo $g \geqslant 0$, to $\{g_n\}$ można wybrać tak by ciąg ten był rosnący (tzn. $\forall_{x\in D}\ \{g_n(x)\}$ — rosnący). Dowód tego faktu pozostawiam jako zadanie (patrz zad. X.31).

Etap 2. Niech $f : D \to \overline{\mathbb{R}}$ będzie mierzalna i $f \geqslant 0$. Wówczas *całkę z f względem miary μ* definiujemy wzorem

$$\int_D f d\mu := \sup_{g\in P_f} S(g), \tag{X.5}$$

gdzie P_f oznacza zbiór wszystkich funkcji prostych g takich, że $0 \leqslant g \leqslant f$ (w szczególności $P_f \neq \emptyset$, bo funkcja zerowa na D należy do P_f). Oczywiście, jak wynika ze wzoru (X.5), gdy $f \equiv 0$, to $\int_D f d\mu = 0$. Podkreślmy, że dla przypadku funkcji mierzalnej $\geqslant 0$ całka jest **zawsze określona**, choć może być równa $+\infty$.

Etap 3. Niech $f : D \to \overline{\mathbb{R}}$ będzie mierzalna. Zdefiniujemy tzw. część dodatnią i część ujemną funkcji f (obie będą $\geqslant 0$), które będziemy oznaczać f^+ i f^-, odpowiednio. Dla $x \in D$ określamy

$$f^+(x) = \max(0, f(x)),\ f^-(x) = -\min(0, f(x)). \tag{X.6}$$

Nietrudno sprawdzić, że f^+ i f^- są obie mierzalne (zadanie X.30) oraz zachodzi

$$f = f^+ - f^-,\ |f| = f^+ + f^-. \tag{X.7}$$

W szczególności, gdy $f \geqslant 0$, to $f = f^+$ oraz $f^- = 0$, zatem aby zdefiniować określoność i wartość całki z f rozszerzając jednocześnie definicję z poprzedniego etapu, postąpimy następująco (wciąż zakładając, oczywiście, mierzalność f).

Definicja.

- *Całka z f względem μ **jest nieokreślona** wtw $\int_D f^+ d\mu = \int_D f^- d\mu = +\infty$. W pozostałych przypadkach całka z f względem μ **jest określona** i wartość jej zadana jest wzorem*

$$\int_D f d\mu := \int_D f^+ d\mu - \int_D f^- d\mu.$$

- *Funkcja f jest **całkowalna względem** μ (ew. **całkowalna w sensie Lebesgue'a** wzgl. μ)*[180] *wtw*

$$\int_D f^+ d\mu,\ \int_D f^- d\mu < +\infty.$$

  O takiej funkcji będziemy też mówić, że jest klasy L^1, a zbiór wszystkich funkcji określonych na D i całkowalnych względem μ będziemy oznaczać $L^1(D, \mu)$.

[179] Na razie jeszcze nie używamy tu symbolu całki „$\int$”, ale jak się wkrótce przekonamy $S(f)$ będzie rzeczywiście równe całce z f, o ile $S(f)$ jest określone.

[180] Często będziemy mówić w skrócie: *całkowalna*.

Należy podkreślić istotną różnicę pomiędzy sytuacją gdy całka z f jest określona a sytuacją gdy f jest całkowalna. Dla funkcji całkowalnej całka jest oczywiście określona, ale odwrotnej zależności nie ma (— całkować można więc też niektóre funkcje niecałkowalne, wbrew nazwie...). To pełna analogia do znanej nam sytuacji z teorii granicy ciągu. Istnienie granicy (odpowiednik określonej całki) to warunek ogólniejszy niż zbieżność (odpowiednik całkowalności).

Z definicji powyższej dostajemy natychmiast:

Wniosek. *Funkcja mierzalna f jest całkowalna wtw całka z f jest określona i ma skończoną wartość (tzn. jest liczbą rzeczywistą).*

Sformułowane w definicji warunki dotyczące określoności całki i całkowalności stanowią właśnie owe „ograniczenia ilościowe” dla całkowanych funkcji, o których mówiliśmy w podrozdziale 1.4. Pamiętajmy, że aby te warunki można było w ogóle wyrazić, potrzebna nam była mierzalność funkcji (uznana wcześniej za „ograniczenie jakościowe”).

Przykłady.

1. Funkcja stała. Jeżeli $f \equiv c$ gdzie $c \in \overline{\mathbb{R}}$, to $\int_D f d\mu = c \cdot \mu(D)$, z zastosowaniem umowy (X.4) dotyczącej mnożenia 0 i $\pm\infty$. Łatwo to sprawdzić bezpośrednio z etapu 2 definicji, gdy $c \in [0; +\infty]$, a gdy $c < 0$ — z etapu 3. W szczególności zachodzi

$$\mu(D) = \int_D 1 d\mu \tag{X.8}$$

(tu zwyczajowo utożsamiamy liczbę 1 z funkcją stale równą 1 na D, którą tu całkujemy, oznaczając ją też przez 1).

2. Całka dla miary Diraca. Niech δ_{x_0} będzie miarą Diraca w x_0 określoną na σ–ciele wszystkich podzbiorów (2^X) zbioru X i rozważmy dowolną funkcję $f : X \to \overline{\mathbb{R}}$. Wówczas $\int_X f d\delta_{x_0} = f(x_0)$. Inaczej mówiąc całkowanie względem tej miary polega po prostu na braniu wartości funkcji w punkcie x_0 (patrz zad. X.32).

3. Całka względem miary liczącej. Rozważmy miarę liczącą # dla podzbiorów zbioru liczb naturalnych $\mathbb{N}$ i niech $f : \mathbb{N} \to \mathbb{R}$ (tzn. f jest po prostu ciągiem, $f = \{f(n)\}_{n \geqslant 1}$). Nietrudno sprawdzić, że gdy $f \geqslant 0$, to

$$\int_{\mathbb{N}} f d\# = \sum_{n=1}^{+\infty} f(n)$$

oraz że wzór ten zachodzi także dla funkcji całkowalnych, co w tym przypadku równoważne jest po prostu bezwzględnej zbieżności szeregu z prawej strony wzoru (zad. X.33).

Opisaliśmy powyżej tylko najprostsze przykłady całkowania. Jednak oczywiście najważniejsza jest dla nas sprawa całkowania względem miary Lebesgue'a λ^d, szczególnie dla $d > 1$. Zajmiemy się tym nieco dalej.

Wcześniej sformułujemy kilka ogólnych własności całki względem miary. Dla wygody przyjmijmy notację podobną do tej użytej przy całce Riemanna: jeżeli $f : D \to \overline{\mathbb{R}}$ oraz $D' \subset D$, to

$$\int_{D'} f d\mu := \int_{D'} f\,|_{D'}\, d\mu,$$

o ile całka z prawej strony jest określona.

Twierdzenie X.2 (o własnościach całki Lebesgue'a).

1. *Jeżeli $f \in L^1(D, \mu)$, to $\mu(\{x \in D : |f(x)| = +\infty\}) = 0$.*

2. *(jednorodność) Jeśli* $f \in L^1(D,\mu)$ *i* $a \in \mathbb{R}$, *to* $af \in L^1(D,\mu)$ [181] *oraz* $\int_D af d\mu = a \cdot \int_D f d\mu$.

3. *(addytywność) Jeżeli* $f, g \in L^1(D,\mu)$, *to* $f+g \in L^1(D,\mu)$ [182] *oraz*
$$\int_D (f+g)d\mu = \int_D f d\mu + \int_D g d\mu.$$

4. *(monotoniczność) Jeżeli* $f, g : D \to \overline{\mathbb{R}}$ *są mierzalne, całki* f *i* g *są określone oraz* $g \leqslant f$, *to*
$$\int_D g d\mu \leqslant \int_D f d\mu.$$

5. *(addytywność względem zbioru) Jeżeli zbiory mierzalne* A_n *są parami rozłączne dla* $n \in \mathbb{N}$ *oraz* $D = \bigcup_{n\in\mathbb{N}} A_n$, f *jest funkcją mierzalną i całka z* f *jest określona, to każda z całek z* $f|_{A_n}$ *też jest określona, oraz*
$$\int_D f d\mu = \sum_{n=1}^{+\infty} \int_{A_n} f d\mu.$$
W szczególności, gdy $f \in L^1(D,\mu)$ *oraz* D' *jest mierzalnym podzbiorem* D, *to* $f|_{D'} \in L^1(D',\mu)$.

6. *Jeżeli* $f : D \to \mathbb{R}$ *jest funkcją prostą, to* $S(f)$ *jest określona wtw całka z* f *względem* μ *jest określona i wówczas* $\int_D f d\mu = S(f)$.

7. *Jeżeli* $f : D \to \overline{\mathbb{R}}$ *jest mierzalna i* $\mu(D) = 0$, *to* $\int_D f d\mu = 0$.

8. *Jeżeli* $f : D \to \overline{\mathbb{R}}$ *jest mierzalna i* $f \geqslant 0$ *oraz* $\int_D f d\mu = 0$, *to* $\mu(\{x \in D : f(x) \neq 0\}) = 0$.

B.D.

Wniosek. *Jeżeli* $f : D \to \overline{\mathbb{R}}$ *jest mierzalna, to* $f \in L^1(D,\mu)$ *wtw* $\int_D |f| d\mu < +\infty$.

Dowód.
Wynika to natychmiast z własności 2 i 3 powyżej oraz ze wzoru (X.7). □

Własność 7 oznacza, że zbiory miary zero są „nieistotne” przy całkowaniu. W praktyce, w różnych sytuacjach będziemy mieli do czynienia z funkcjami, które chcielibyśmy scałkować „po zbiorze D”, choć nie będą one określone w pewnych jego punktach. Jeżeli zbiór tych punktów nieokreśloności zawiera się w pewnym zbiorze miary zero, to o funkcji takiej będziemy mówić że jest *określona prawie wszędzie na* D (ew. μ*–prawie wszędzie*, jeśli będziemy chcieli wyraźnie zaznaczyć o jaką chodzi miarę). Z tego typu sytuacją spotkaliśmy się już w sformułowaniu własności 2 z twierdzenia X.2 (patrz przypis do tej własności). Pytanie tylko — jak należy określić całkę z funkcji „po D” w tej sytuacji? Otóż przypuśćmy, że $Z \subset D$ jest wspomnianym wcześniej zbiorem miary 0 (dziedzina f zawiera $D \setminus Z$) — oczywiście nie ma powodu by Z był jednoznacznie wyznaczony, jednak jeżeli $f|_{D\setminus Z}$ jest mierzalna i całka z niej jest określona, to w oparciu o twierdzenie X.2 nietrudno wykazać, że wartość $\int_{D\setminus Z} (f|_{D\setminus Z})\, d\mu$ daje przy każdym takim wyborze zbioru Z ten sam wynik. Dlatego w powyższej sytuacji określamy
$$\int_D f d\mu := \int_{D\setminus Z} (f|_{D\setminus Z}) d\mu \text{ [183]} .$$

---
[181] Uwaga: definiując $a \cdot f$ stosujemy tu konwencję (X.4) w tym znaczeniu, żę $(a \cdot f)(x) = a \cdot f(x)$, zatem gdy $a = 0$ i $f(x_0) = \pm\infty$, to $(a \cdot f)(x_0) = 0$.
[182] Jeżeli dla pewnego $x_0 \in D$ zachodzi $f(x_0) = +\infty$ i $g(x_0) = -\infty$ lub odwrotnie, to $f+g$ nie jest w ogóle określona w punkcie x_0. Jednak na mocy punktu 1 zbiór takich x_0 ma miarę μ zerową i jako $f+g$ przyjmujemy tu jakąkolwiek funkcję mierzalną, której wartość wynosi $f(x) + g(x)$ dla pozostałych $x \in D$.
[183] Własności 5 i 7 gwarantują ponadto, że jeśli $\tilde{f}$ jest jakimkolwiek mierzalnym przedłużeniem $f|_{D\setminus Z}$ do całej dziedziny D, to przy takim określeniu $\int_D f d\mu = \int_D \tilde{f} d\mu$.

Jeżeli dla pewnego zbioru Z miary zero takiego, że dziedzina f zawiera $D \setminus Z$ funkcja $f|_{D\setminus Z}$ jest mierzalna, to mówimy, że funkcja f określona prawie wszędzie na D jest *mierzalna*[184] i podobnie dla innego typu własności funkcji, np. dla całkowalności.

Podobnego określenia *prawie wszędzie*, czy *dla prawie wszystkich...* (ew. *μ–prawie wszystkich...*) będziemy używać ogólnie w sytuacjach, w których rozważana własność zachodzi poza pewnym zbiorem miary zero.

Nietrudno wykazać poniższy rezultat dotyczący *równości prawie wszędzie.*

Fakt. *Jeżeli $f, g : D \to \mathbb{R}$ są mierzalne i równe μ–prawie wszędzie (tzn. $f(x) = g(x)$ dla μ–prawie wszystkich $x \in D$), to całka z f jest określona wtw całka z g jest określona oraz wówczas $\int_D f d\mu = \int_D g d\mu$. W szczególności $f \in L^1(D, \mu)$ wtw $g \in L^1(D, \mu)$.*

Na koniec omawiania ogólnych własności całki Lebesgue'a sformułujemy dwa twierdzenia „o przechodzeniu do granicy pod znakiem całki", czyli takie, których teza mówi, że „całka granicy równa jest granicy całek".

Twierdzenie X.3. *Załóżmy, że funkcje $f_n : D \to \overline{\mathbb{R}}$ są mierzalne dla $n \geqslant n_0$ oraz że $f_n \to f$* [185] *. Jeżeli spełnione jest **któreś** z poniższych założeń:*

(M) (tw. „o zbieżności monotonicznej") $\{f_n\}_{n \geqslant n_0}$ jest rosnącym ciągiem funkcji nieujemnych, tzn.

$$\forall_{\substack{n \geqslant n_0 \\ x \in D}} \ 0 \leqslant f_n(x) \leqslant f_{n+1}(x),$$

(L) (tw. Lebesgue'a „o zbieżności majoryzowalnej") $\{f_n\}_{n \geqslant n_0}$ jest majoryzowalny przez funkcję całkowalną, tzn. istnieje $F \in L^1(D, \mu)$ taka, że

$$\forall_{\substack{n \geqslant n_0 \\ x \in D}} \ |f_n(x)| \leqslant F(x),$$

to

$$\lim_{n \to +\infty} \int_D f_n d\mu = \int_D f d\mu \ ^{186)} .$$

B.D.

Warto zwrócić uwagę na fakt, że w twierdzeniu o przechodzeniu do granicy pod znakiem całki dla całki Riemanna (patrz zad. VII.10) zakładaliśmy **aż** zbieżność jednostajną. Powyższe twierdzenia to zatem kolejny „dowód" na wyższość całki Lebesgue'a nad całką Riemanna...

Uwaga. Nie wolno zapominać o założeniach powyższego twierdzenia. Aby sie o tym przekonać wystarczy rozważyć funkcje $f_n : \mathbb{R} \to \mathbb{R}$, $f_n = \chi_{[n;+\infty)}$. Wówczas $f_n \longrightarrow 0$ (funkcja zerowa na $\mathbb{R}$), ale $\forall_n \ \int_{\mathbb{R}} f_n \, d\lambda^1 = +\infty$, choć $\int_{\mathbb{R}} 0 \, d\lambda^1 = 0$... Nie pomaga tu wcale nawet to, że mamy do czynienia z monotonicznym ciągiem funkcji mierzalnych nieujemnych („rosnący" z wersji (M) twierdzenia nie może być zastąpione przez „malejący", wbrew sugestii z nazwy tego twierdzenia...).

2. Całka względem miary Lebesgue'a

Zajmiemy się wreszcie najważniejszą dla nas całką — względem miary Lebesgue'a. Ponieważ stanowi to nasz podstawowy obiekt zainteresowań, uprościmy nieco notację dla takich całek. Zamiast $\int_D f d\lambda^d$, czy $L^1(D, \lambda^d)$, będziemy więc często pisać

$$\int_D f(x)dx \ ^{187)} , \quad L^1(D),$$

[184] Tu już wybór zbioru Z może być istotny, np. wtedy gdy miara μ nie jest zupełna.

[185] Przypominam, że symbol „$\longrightarrow$" oznacza zbieżność **punktową** ciągu funkcyjnego.

[186] Określoność całki dla f (w szczególności jej mierzalność), a dla (L) nawet całkowalność, wynika z założeń twierdzenia (dlaczego?).

[187] Oczywiście zamiast x może być inny symbol.

odpowiednio. Często też będziemy mówić po prostu „całka Lebesgue'a” mając na myśli całkę względem miary Lebesgue'a λ^d (i mając nadzieję, że to nie doprowadzi do żadnych nieporozumień). Ponieważ w przypadku $d = 1$ oraz $D = [a; b] \subset \mathbb{R}$ symbol $\int_D f(x)dx$ oznaczał już wcześniej całkę Riemanna, zatem aby z tego „nadużycia” się wytłumaczyć zacznijmy nasze rozważania właśnie od $d = 1$.

2.1. Całka Lebesgue'a względem jednowymiarowej miary Lebesgue'a. Porównanie z całką Riemanna.

Na szczęście dla funkcji całkowalnych w sensie Riemanna obydwie tytułowe całki pokrywają się ze sobą. Ściślej — ma miejsce następujące twierdzenie.

Twierdzenie X.4 (całka Lebesgue'a kontra całka Riemanna). *Niech $f : [a; b] \to \mathbb{R}$. Wówczas $f \in \mathcal{R}\ ([a; b])$ wtw f jest ograniczona i zbiór jej punktów nieciągłości ma miarę λ^1 równą zero. Jeżeli $f \in \mathcal{R}\ ([a; b])$, to $f \in L^1([a; b])$ [188] oraz całka Riemanna z f równa jest całce względem miary λ^1 z f.*

Dowód (fragmenty).
Wykażemy tylko implikację $f \in \mathcal{R}\ ([a; b]) \Rightarrow f \in L^1([a; b])$ przy założeniu, że prawdziwa jest pierwsza część twierdzenia. Niech zatem $f \in \mathcal{R}([a; b])$ i niech $Z \subset [a; b]$ będzie taki, że $\lambda^1(Z) = 0$ oraz $\tilde{f} := f|_{[a;b]\setminus Z}$ jest ciągła. Wykażemy, że f jest mierzalna. Wystarczy dowieść, że dla dowolnego $A \in \mathcal{B}(\mathbb{R})$ zachodzi $f^{-1}(A) \in \mathcal{L}(\mathbb{R})$. Ale $f^{-1}(A) = \{x \in [a; b] \setminus Z : f(x) \in A\} \cup \{x \in Z : f(x) \in A\} = \tilde{f}^{-1}(A) \cup Z'$, gdzie $Z' \subset Z$.

Jednak $\tilde{f}$ jest ciągła i określona na zbiorze mierzalnym, zatem jest mierzalna (patrz — przykład 1 ze strony 158), czyli $\tilde{f}^{-1}(A) \in \mathcal{L}(\mathbb{R})$. Ponadto $Z' \in \mathcal{L}(\mathbb{R})$, bo miara λ^1 jest zupełna, stąd $f^{-1}(A) \in \mathcal{L}(\mathbb{R})$. Pozostaje wykazać, że $\int_{[a;b]} |f| d\lambda^1 < +\infty$. Ale $|f|$ jest ograniczona, tzn. $|f| \leqslant c$ dla pewnej stałej $c \in \mathbb{R}$, skąd $\int_{[a;b]} |f| d\lambda^1 \leqslant \int_{[a;b]} c\ d\lambda^1 = c \cdot \lambda^1([a; b]) = c(b - a) < +\infty$. □

A zatem całkowanie w sensie Lebesgue'a można traktować dla $d = 1$ po prostu jako uzupełnienie teorii całkowania w sensie Riemanna o pewne przypadki, gdy to drugie całkowanie było niewykonalne. Oto prosty przykład.

Przykład. Rozważmy funkcję Dirichleta na przedziale $[0; 1]$, tzn. $f : [0; 1] \to \mathbb{R}$,

$$f(x) = \begin{cases} 1 & \text{gdy } x \in \mathbb{Q} \\ 0 & \text{gdy } x \notin \mathbb{Q}. \end{cases}$$

Jak wiemy, $f \notin \mathcal{R}([0; 1])$ (zbiór punktów nieciągłości to $[0; 1]$, więc wynika to choćby z twierdzenia X.4, ale ten przykład badaliśmy też w rozdziale VII). Z drugiej strony $f = \chi_{\mathbb{Q}\cap[0;1]}$, więc jest to funkcja mierzalna (patrz przykład 2 ze strony 158) ponieważ $\mathbb{Q} \cap [0; 1]$ jest zbiorem mierzalnym, jako zbiór przeliczalny. Na dodatek $\lambda^1(\mathbb{Q} \cap [0; 1]) = 0$ ze względu na tę przeliczalność. Mamy więc:

$$\int_{[0;1]} |f| d\lambda^1 = \int_{[0;1]} f d\lambda^1 = \int_{[0;1]\cap\mathbb{Q}} 1d\ \lambda^1 + \int_{[0;1]\setminus\mathbb{Q}} 0d\ \lambda^1 = 1 \cdot 0 + 0 \cdot 1 = 0.$$

A zatem $f \in L^1([0; 1])$. Ale zamiast powyższego rachunku można po prostu użyc tego, że f jest funkcją równą zero prawie wszędzie (i powołać się na fakt ze strony 164).

Powyższy przykład chyba dobrze ilustruje istotę sprawy. Nawet jeśli ograniczymy się do rozważania tylko funkcji określonych na przedziale $[a; b]$, całka Lebesgua umożliwia całkowanie

[188] Uwaga: „1” pojawiające się w oznaczeniu „L^1” nie ma nic wspólnego z tym, że rozważamy akurat miarę λ^1. Ta jedynka pochodzi od ogólniejszego oznaczenia $L^p(X, \mu)$ związanego z klasą funkcji, dla których $\int_X |f|^p d\mu < +\infty$.

większej niż całka Riemanna klasy funkcji, mimo iż na pierwszy rzut oka definicje obu całek nie różniły się tak bardzo, poza szczegółami „technicznymi”. W obu sytuacjach podstawą definicji było wyliczenie całki najpierw dla funkcji posiadającej skończony zbiór wartości. W przypadku całki Riemanna były to funkcje stałe na przedziałach utworzonych przez wybrany podział odcinka $[a; b]$, a w przypadku całki Labegue'a — były to funkcje proste. I tu okazuje się, że różnica jest ogromna — ten pierwszy rodzaj funkcji to też funkcje proste, ale bardzo szczególnej postaci! Ograniczenie się do funkcji „stałych na odcinkach” nie pozwala „wnikać” w subtelności funkcji „rozgrywające się” na bardziej zawiłych zbiorach takich jak np. $\mathbb{Q}$ na co pozwalają dużo lepiej ogólne funkcje proste.

Jak pamiętamy, część ograniczeń związanych z całką Riemanna udało się znieść poprzez rozważanie całek niewłaściwych. Warto zatem porównać także ten sposób całkowania z całkowaniem w sensie Lebesgue'a. Niech $I = [a; b)$, $-\infty < a < b \leqslant +\infty$ lub $I = (a; b]$, $-\infty \leqslant a < b < +\infty$ oraz $f : I \to \mathbb{R}$.

Twierdzenie X.5. *Jeżeli istnieje całka niewłaściwa $\int_a^b f(x)dx$ oraz całka Lebesgue'a $\int_I f(x)dx$ jest określona, to obie te całki są równe. W szczególności, jeżeli f jest mierzalna i całka $\int_a^b |f(x)|dx$ jest zbieżna, to $f \in L^1(I)$.* **B.D.**

Twierdzenie to nietrudno wykazać w oparciu o twierdzenie X.3 oraz o twierdzenie X.4 (patrz zadanie X.37).

Należy jednak pamiętać o tym, że w pierwszej części powyższego twierdzenia zakładamy nie tylko istnienie całki niewłaściwej, ale też określoność całki Lebesgue'a. Czasami spełnione jest tylko to pierwsze założenie i w takiej sytuacji teoria całki Lebesgue'a jest bezradna, a jedynym ratunkiem pozostaje całka niewłaściwa. Ilustruje to poniższy przykład.

Przykład. Niech $f : [1; +\infty) \to \mathbb{R}$, $f(x) = \frac{\sin(x)}{x}$. Całka niewłaściwa $\int_1^{+\infty} f(x)dx$ jest zbieżna (na mocy kryterium Dirichleta — patrz np. zadanie VII.10 (h)). Można jednak wykazać, że całka Lebesgue'a z f nie jest określona (zadanie X.38).

Warto też podkreślić, że druga część twierdzenia X.5 bywa wygodnym narzędziem do badania całkowalności (w sensie Lebesgue'a) wielu funkcji jednej zmiennej. Na przykład tą drogą uzyskujemy natychmiast całkowalność funkcji $f : [1; +\infty) \longrightarrow \mathbb{R}$ zadanych wzorem $f(x) = \frac{1}{x^\alpha}$ dla wszystkich $\alpha > 1$ (i brak całkowalności dla pozostałych).

3. Całkowanie w wielu wymiarach i twierdzenie Fubiniego

Opiszemy teraz w jaki sposób da się w praktyce całkować względem miary λ^d dla $d > 1$. Dzięki wzorowi (X.8) umożliwi to nam także obliczanie miar λ^d rozmaitych zbiorów. Jak zaraz zobaczymy, całki wielowymiarowe dadzą się sprowadzić do całek względem miary λ^1, które dzięki poprzedniemu podrozdziałowi nie stanowią już dla nas problemu. Twierdzenie, które pozwala sprawdzić całkę wielowymiarową do jednowymiarowej to twierdzenie Fubiniego. Można je formułować dla funkcji określonych na dziedzinie $D \subset \mathbb{R}^d$, jednak nieco prostsze sformułowanie otrzymamy gdy D będzie całym $\mathbb{R}^d$. Dlatego zacznijmy od poniższej uwagi dotyczącej nie koniecznie tylko miary λ^d, ale abstrakcyjnej miary μ określonej na σ-ciele podzbiorów X.

Uwaga. Niech $D \subset X$, $D \in \mathbb{M}$ i niech $f : D \to \overline{\mathbb{R}}$. Określimy $\tilde{f} : X \to \overline{\mathbb{R}}$ wzorem

$$\tilde{f}(x) = \begin{cases} f(x) & \text{dla } x \in D \\ 0 & \text{dla } x \notin D. \end{cases}$$

Wówczas każda z własności: mierzalność, określoność całki, całkowalność dla f jest równoważna tej własności dla $\tilde{f}$. Ponadto, gdy jedna z tych całek jest określona, to druga jest określona także i są sobie równe.

Dzięki temu będziemy mogli się ograniczyć do rozważania funkcji określonych na $\mathbb{R}^d$.

Stosowanie twierdzenia Fubiniego powoduje zmniejszenie „wymiaru miary” po której się całkuje. Wiąże się to każdorazowo z podziałem zmiennych „skalarnych” $x_1, \ldots, x_d$ na dwie rozłączne „grupy” i w efekcie najpierw całkuje się „po zmiennych” z jednej z nich, a potem z drugiej. Nie musi to być koniecznie podział typu „$x_1, \ldots, x_k$” i „$x_{k+1}, \ldots, x_d$” — może być to zupełnie dowolny podział. Ustalmy więc podział na grupę d_1 i d_2 elementową, gdzie $d_1+d_2 = d$ i $d_1, d_2 \geqslant 1$, przy czym w pierwszej grupie są zmienne o numerach $i_1, \ldots, i_{d_1}$ wypisanych tu w kolejności rosnącej, a w drugiej o pozostałych numerach $j_1, \ldots j_{d_2}$ też wypisanych w kolejności rosnącej. Przyjmujemy teraz taką notację, która pozwoli nam traktować ten podział na dwie grupy zmiennych analogicznie jak przy rozważaniu iloczynu kartezjańskiego $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Mianowicie dla $s \in \mathbb{R}^{d_1}$ oraz $t \in \mathbb{R}^{d_2}$ niech symbol (s, t) oznacza tu nie jak dotąd „parę s, t” lecz taki element $x \in \mathbb{R}^d$, że $x_{i_k} = s_k$ dla $k = 1, \ldots, d_1$ oraz $x_{j_k} = t_k$ dla $k = 1, \ldots d_2$. Jeżeli komuś wydaje się zbyt zawiłe, to może z powodzeniem myśleć o najprostszym przypadku gdy $d = 2$, $d_1 = d_2 = 1$, $i_1 = 1$ oraz $j_1 = 2$. Wówczas napis (x_1, x_2) oznacza to co wcześniej — punkt z $\mathbb{R}^2$ o pierwszej współrzędnej x_1 a drugiej x_2. Dla funkcji $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ i dla $s \in \mathbb{R}^{d_1}$ nich f_s oznacza funkcję z $\mathbb{R}^{d_2}$ w $\overline{\mathbb{R}}$ zadaną dla $t \in \mathbb{R}^{d_2}$ wzorem

$$f_s(t) = f((s, t)).$$

Ponadto niech $\mathcal{C}_2 f$ oznacza funkcję „całka po drugiej zmiennej”[189], której dziedziną jest

$$D_2 := \{s \in \mathbb{R}^{d_1} : \text{ całka z funkcji } f_s \text{ względem miary } \lambda^{d_2} \text{ jest określona}\}$$

i która dla $s \in D_2$ zadana jest wzorem

$$\mathcal{C}_2 f(s) = \int_{\mathbb{R}^{d_2}} f_s d\lambda^{d_2}.$$

Teraz możemy juz wreszcie sformułować zapowiadane twierdzenie (i nie wykluczone, że samo sformułowanie jest mniej skomplikowane, niż poprzedzające je oznaczenia ...)

Twierdzenie X.6 (Fubiniego). *Załóżmy, że $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ jest funkcją mierzalną nieujemną*[190] *lub funkcją całkowalną względem miary λ^d. Wówczas funkcja $\mathcal{C}_2 f$ jest określona prawie wszędzie na $\mathbb{R}^{d_1}$*[191]*, całka z $\mathcal{C}_2 f$ względem λ^{d_1} jest określona oraz zachodzi:*

$$\int_{\mathbb{R}^{d_1}} \mathcal{C}_2 f d\lambda^{d_1} = \int_{\mathbb{R}^d} f d\lambda^d. \tag{X.9}$$

B.D.

Uwagi.

1. Bardziej tradycyjny, znacznie milszy w czytaniu i bardziej chyba czytelny (choć jednak mniej ścisły) zapis wzoru (X.9) ma postać następującą:

$$\int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f((x_1, x_2)) dx_2 \right) dx_1 = \int_{\mathbb{R}^d} f(x) dx. \tag{X.10}$$

   Co więcej w tradycyjnej wersji można spotkać też zapis prawej strony powyższego wzoru w formie następującej:

$$\int \int_{\mathbb{R}^d} \cdots \int f(x) dx_1 dx_2 \ldots dx_d \text{ , } (d \text{ symboli całki } „\textstyle\int”)$$

[189] Ściślej — chodzi o całkę (stąd ozn.$\mathcal{C}$) „po drugiej grupie zmiennych”, a pierwsza grupa pełni rolę zmiennych dla tej już „wycałkowanej po drugich zmiennych” funkcji $\mathcal{C}_2 f$.

[190] Wersja dla funkcji mierzalnej nieujemnej bywa wiązana także z nazwiskiem Tonellego.

[191] Czyli $\lambda^{d_1}(\mathbb{R}^{d_1} \setminus D_2) = 0$.

gdzie „$\int\int_{\mathbb{R}^d}\cdots\int$" i „$dx_1dx_2\dots dx_d$" mają zastępować odpowiednio „$\int_{\mathbb{R}^d}$" i „dx" i jednocześnie wyraźnie podkreślić „wielowymiarowość" całkowania, które tam się odbywa. W tradycyjnym nazewnictwie całka z prawej strony to, w zależności od d, tzw. *całka podwójna*, *potrójna* itd. („po$d$-tna" ...), a wyrażenie z lewej strony — *całka iterowana.* Oczywiście użycie twierdzenia Fubiniego $d-1$ razy pozwala zastąpić całkę „po$d$-tną" całką iterowaną „w pełni", tj. taką, że występuje tam d całek po $\mathbb{R}$. Jednak kolejność wykonywania poszczególnych całkowań może być rozmaita — zależy to od tego jakiego podziału zmiennych na dwie grupy dokonamy na każdym kroku. Np., gdy $d=3$, to dzieląc x_1,x_2,x_3 najpierw na x_1 oraz x_2,x_3 a następnie tę drugą grupę na x_2 i x_3 uzyskujemy ostatecznie całkę iterowaną

$$\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(x_1,x_2,x_3)dx_3\right)dx_2\right)dx_1.$$

Taką samą całką iterowaną uzyskamy przy podziale najpierw na x_1,x_2 oraz x_3, a potem — dla pierwszej grupy — na x_1 i x_2. Z kolei gdy najpierw podzielimy na x_1,x_3 oraz x_2 a potem pierwszą grupę na x_3 i x_1, to uzyskamy najpierw

$$\int_{\mathbb{R}^2}\left(\int_{\mathbb{R}} f(x_1,x_2,x_3)dx_2\right)d(x_1,x_3)$$

a następnie

$$\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(x_1,x_2,x_3)dx_2\right)dx_1\right)dx_3.$$

Ogólnie, sprowadzając do postaci całek iterowanych po $\mathbb{R}$, możemy uzyskać $d!$ rozmaitych kolejności całkowania. Warto podkreślić, że wybór tej kolejności może być bardzo istotny w praktycznych rachunkach — niektóre z kolejności mogą prowadzić do łatwiejszych, inne do trudniejszych (czy wręcz — „niewykonalnych") rachunków.

2. Sformułowaliśmy założenia w dwóch wersjach. Sprawdzenie, że funkcja jest mierzalna i nieujemna bywa na ogół rzeczą nietrudną. Co jednak zrobić, gdy funkcja nie jest nieujemna? Aby spróbować drugiej wersji twierdzenia, trzeba wiedzieć, że funkcja jest całkowalna. Czasem to można stwierdzić od razu (np. funkcja jest mierzalna i ograniczona oraz zeruje się poza zbiorem ograniczonym), ale czasem wydaje się, że jedynym sposobem byłoby stwierdzenie jaką wartość ma całka (dla całkowalności wartość ta musi być skończona). Trzeba by więc ją policzyć, a do tego trzeba użyć twierdzenia Fubiniego ...Czy to nie błędne koło? Na szczęście nie! Tak naprawę, by stwierdzić całkowalność należy wiedzieć nie tyle jaka jest wartość $\int_{\mathbb{R}^d} f d\lambda^d$ (nie wiemy nawet, czy to jest określone), ale czy $\int_{\mathbb{R}^d}|f|d\lambda^d$ jest skończona. A do obliczenia $\int_{\mathbb{R}^d}|f|d\lambda^d$ można zastosować pierwszą wersję twierdzenia — dla funkcji mierzalnych nieujemnych. Tak więc często postępowanie jest takie: najpierw używamy wersji pierwszej (do $|f|$ — mierzalnej, gdy f była mierzalna) by sprawdzić całkowalność i jeśli f okaże się całkowalna, to do obliczenia całki stosujemy wersję drugą.

3. Oczywiście jeśli założenia twierdzenia są spełnione, to wszystkie całki iterowane są równe. Jednak bez tych założeń całki iterowane mogą być różne. Z drugiej strony, równość całek iterowanych nie gwarantuje, że założenia twierdzenia są spełnione!

Przykłady.

1. Obliczymy całkę Lebesgue'a z funkcji $f:\mathbb{R}^2\to\mathbb{R}$ zadanej wzorem $f(x)=x_1x_2e^{-(x_1^2+x_2^2)}$. Jednak najpierw musimy się upewnić, czy całka ta w ogóle jest określona. Oczywiście

f jest mierzalna jako funkcja ciągła, nie jest jednak nieujemna, a zatem spróbujmy postąpić zgodnie z uwagą 2. Mamy

$$\begin{aligned}\int_{\mathbb{R}^2}|f|d\lambda^2 &\overset{192)}{=} \\ \int_{\mathbb{R}^2}|x_1x_2|e^{-(x_1^2+x_2^2)}dx &\overset{193)}{=} \\ \int_{\mathbb{R}}(\int_{\mathbb{R}}|x_1|e^{-x_1^2}\cdot|x_2|e^{-x_2^2}dx_1)dx_2 &= \\ \int_{\mathbb{R}}|x_2|e^{-x_2^2}(\int_{\mathbb{R}}|x_1|e^{-x_1^2}dx_1)dx_2 &= \\ (\int_{\mathbb{R}}|x_2|e^{-x_2^2}dx_2)\cdot(\int_{\mathbb{R}}|x_1|e^{-x_1^2}dx_1) &= \\ (\int_{\mathbb{R}}|t|e^{-t^2}dt)^2 = (C_1+C_2)^2, &\end{aligned}$$

gdzie $C_1 = \int_{[-\infty;0]}|t|e^{-t^2}dt$, $C_2 = \int_{[0;+\infty)}|t|e^{-t^2}dt$. Te dwie całki Lebesgue'a są na mocy twierdzenia X.5 równe odpowiednim całkom niewłaściwym, dla których możemy już stosować metody z rozdziału VII. W szczególności bez trudu zauważamy, że $C_1 = C_2$ (dzięki parzystości funkcji zadanej wzorem $|t|e^{-t^2}$ [194]). Mamy też $C_2 = \lim_{r\to+\infty}\int_0^r te^{-t^2}dt = \lim_{r\to+\infty} -\frac{1}{2}\int_0^{-r^2}e^s ds = \frac{1}{2}\int_{-\infty}^0 e^s ds = \frac{1}{2}$, na mocy wzoru na całkowanie przez podstawienie dla całek oznaczonych. W efekcie zatem wykazaliśmy całkowalność f, gdyż $\int_{\mathbb{R}^2}|f|d\lambda^2 = 1 < +\infty$. Możemy więc „prawomocnie" zająć się obliczaniem całki z f przy użyciu drugiej wersji twierdzenia Fubiniego — tej dla funkcji całkowalnych[195]. Rachunek nieco skrócimy, bo kolejne przejścia są prawie te same co wyżej. Uzyskujemy więc:

$$\int_{\mathbb{R}^2} fd\lambda^2 = (\int_{\mathbb{R}} te^{-t^2}dt)^2 = (\int_{-\infty}^0 te^{-t^2}dt + \int_0^{+\infty} te^{-t^2}dt)^2 = (-C_1+C_2)^2 = 0.$$

2. Obliczymy pole powierzchni koła o promieniu $r > 0$, ściślej obliczymy $\lambda^2(K_r)$, gdzie $K_r := \{x\in\mathbb{R}^2 : \|x\| < r\}$ (a więc rozważamy koło otwarte).

   Po pierwsze zauważmy, że $\lambda^2(K_r) = \int_{K_r} 1 d\lambda^2$ (patrz (X.7)). Ponieważ funkcja stale równa 1 na K_r jest mierzalna i nieujemna, zatem możemy użyć twierdzenia Fubiniego, ale najpierw zgodnie z wymogami założeń twierdzenia Fubiniego (wg. uwagi ze strony 166) musimy rozszerzyć tę funkcję do określonej na $\mathbb{R}^2$ i równej po prostu χ_{K_r}. Mamy więc

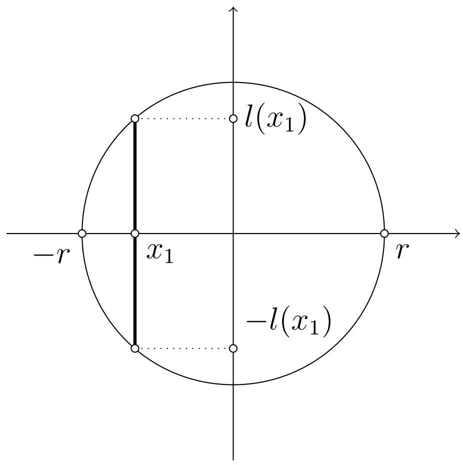
$$\lambda^2(K_r) = \int_{\mathbb{R}^2}\chi_{K_r}d\lambda^2 = \int_{\mathbb{R}}(\int_{\mathbb{R}}\chi_{K_r}(x_1,x_2)dx_2)dx_1$$

   Przyjrzyjmy się wewnętrznej całce. Przy ustalonym $x_1\in\mathbb{R}$ funkcja $\lambda_{K_r}(x_1,x_2)$ zeruje się gdy $x_1^2+x_2^2 \geqslant r^2$, a dla pozostałych przypadków ma wartość 1. W efekcie (patrz rys. 31), jeżeli $|x_1|\geqslant r$, to wewnętrzna całka wynosi 0, jako całka z funkcji zerowej. Natomiast gdy $|x_1| < r$, to funkcja pod wewnętrzną całką równa jest 1 wtw $|x_2| < l(x_1) := \sqrt{r^2 - x_1^2}$. W efekcie mamy (stosujemy addytywność całki względem zbioru)

$$\lambda^2(K_r) = 0 + \int_{(-r;r)}\left(\int_{(-l(x_1);l(x_1))} 1dx_2\right)dx_1 = \int_{(-r;r)}\lambda^1\Big(\big(-l(x_1);l(x_1)\big)\Big)d_{x_1} =$$
$$= \int_{(-r;r)} 2\sqrt{r^2-x_1^2}dx_1.$$

[194] Aby zrobić to całkiem ściśle, trzeba użyć definicji całki niewłaściwej i tw. o całkowaniu przez podstawienie dla całki oznaczonej — zachęcam do wypisania szczegółów ...

[195] Chodzi tu oczywiscie o całkowalność względem miary λ^2 (w sensie Lebesgue'a). Jednak coraz częściej będziemy mówić tylko „całkowalna", pozostawiając resztę w domyśle. Natomiast, aby nie było nieporozumień, w przypadku całkowalności w sensie Riemanna zawsze będziemy używać pełnej nazwy.



Rysunek 31:

Jak poradzić sobie z tą ostatnią całką Lebesgue'a? Gdybyśmy całkowali funkcję określoną na przedziale domkniętym, a nie otwartym, nie byłoby problemu, bo na mocy twierdzenia X.4 sprawa sprowadzałaby się do policzenia całki Riemanna (równej całce oznaczonej). Szczęśliwie miara λ^1 zbiorów jednopunktowych jest zerowa, zatem ta całka jest równa całce po $[-r;r]$. A zatem

$$\begin{aligned}\lambda^2(K_r) = 2\int_{-r}^{r}\sqrt{r^2-t^2}dt &= 2\int_{-r}^{r} r\sqrt{1-\left(\frac{t}{r}\right)^2}dt = {}^{196)}\\ 2r^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\sqrt{1-\sin^2(u)}\cos(u)du &= 2r^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos^2 u\, du\end{aligned}$$

Do całki $I := \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos^2 u$ zastosujemy całkowanie przez części:

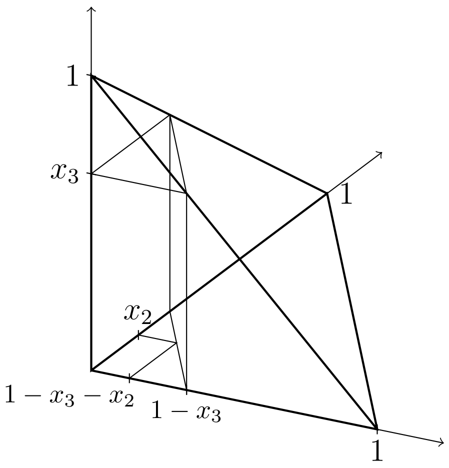
$$I = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\sin'(u)\cos(u)du = 0 - \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\sin(u)(-\sin(u)) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\sin^2(u)du$$

A zatem $2I = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos^2 u + \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\sin^2(u) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 1 = \pi$, czyli $I = \frac{\pi}{2}$. Ostatecznie więc pole powierzchni K_r wynosi πr^2, co (miejmy nadzieję) nikogo z Państwa nie zaskoczyło ...

Zauważmy jeszcze, że w tych wyliczeniach wybór „kolejności całkowania" (tzn. podziału zmiennych na pierwszą i drugą grupę) nie miał żadnego znaczenia dla samych rachunków, ze względu na symetrię rozważanego zbioru (podobnie było w przykładzie poprzednim).

3. Obliczymy $\int_C x_3 d\lambda^3$, gdzie C jest czworościanem w $\mathbb{R}^3$ o wierzchołkach $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(0,0,0)$ a ściślej $C = \{x \in \mathbb{R}^3 : x_1, x_2, x_3 \geqslant 0 \text{ oraz } x_1 + x_2 + x_3 \leqslant 1\}$.

   Całkowana funkcja jest więc mierzalna i nieujemna — możemy zatem postępować podobnie jak w przypadku 2, tyle że w trzech wymiarach zamiast w dwóch. Skrócimy zatem tym razem rozumowanie, pomijając niektóre szczegóły i pamiętając, że najważniejsze jest wyznaczenie faktycznego zakresu całkowania przy całkach iterowanych uzyskanych po zastosowaniu tw. Fubiniego. Tzn. musimy zastąpić całki po całym $\mathbb{R}$ całkami po odpowiednich zbiorach związanych z geometrią wyjściowego zbioru C, na którym zadana jest nasza funkcja. Ustalmy, że stosując twierdzenie Fubiniego dwukrotnie, kolejność całkowania (licząc od najbardziej wewnętrznego) w całkach iterowanych będzie taka: x_1, x_2, x_3. Ustalając zmienną najbardziej zewnętrzną, tzn. x_3 widzimy, że wystarczy ją rozważać w przedziale $[0;1]$. Ponadto, gdy $x_3 \in [0;1]$, współrzędna x_2 punktu $x \in C$



Rysunek 32:

musi spełniać $x_2 \in [0; 1-x_3]$, gdyż zachodzi $x_1, x_2 \geqslant 0,\ x_1+x_2 \leqslant 1-x_3$. A gdy ustalimy x_2 w tym przedziale, to zachodzi $x_1 \in [0; 1-x_3-x_2]$ (patrz rysunek 32).

W efekcie mamy:

$$\begin{aligned}\int_C x_3 d\lambda^3 &= \int_C x_3 dx = \int_0^1 \left(\int_0^{1-x_3} \left(\int_0^{1-x_3-x_2} x_3 dx_1 \right) dx_2 \right) dx_3 = \\ &= \int_0^1 \left(\int_0^{1-x_3} x_3(1-x_3-x_2) dx_2 \right) dx_3 = \int_0^1 x_3 \left((1-x_3)^2 - \frac{1}{2}(1-x_3)^2 \right) dx_3 = \\ &= \frac{1}{2} \int_0^1 x_3(1-x_3)^2 dx_3 = \frac{1}{2} \int_0^1 t^3 - 2t^2 + t dt = \frac{1}{2} \left(\frac{1}{4} - \frac{2}{3} + \frac{1}{2} \right) = \frac{1}{24}.\end{aligned}$$

Na zakończenie tego podrozdziału pokażemy pewne wygodne zastosowanie twierdzenia Fubiniego związane ze zbiorami miary zero. Wykażemy mianowicie, że podzbiór $\mathbb{R}^{d+1}$, który jest wykresem funkcji mierzalnej określonej na podzbiorze $\mathbb{R}^d$ ma miarę λ^{d+1} równą zero. By wygodniej było sobie wyobrazić odpowiednią sytuację najlepiej rozważyć $d = 1$ (wykres funkcji jednej zmiennej jako podzbiór $\mathbb{R}^2$) lub $d = 2$ (wykres funkcji dwóch zmiennych jako podzbiór $\mathbb{R}^3$). Ogólnie, dla $f : D \to \mathbb{R}$, gdzie $D \subset \mathbb{R}^d$, wykres f to zbiór

$$\Gamma_f := \{(x_1, \ldots, x_d, f(x_1, \ldots, x_d)) \in \mathbb{R}^{d+1} : (x_1, \ldots, x_d) \in D\}.$$

Fakt. *Jeśli $f : D \to \mathbb{R}$ jest funkcją mierzalną, to $\lambda^{d+1}(\Gamma_f) = 0$.*

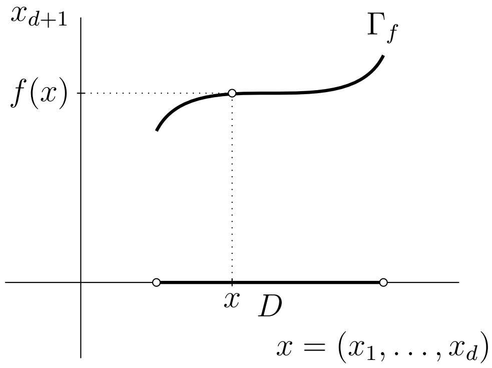
Dowód.
Zauważmy na początek, że Γ_f jest mierzalny. Niech bowiem $F : D \times \mathbb{R} \to \mathbb{R}$ będzie zadana wzorem $F(x_1, \ldots, x_{d+1}) = f(x_1, \ldots, x_d) - x_{d+1}$. Można wykazać — choć nie jest to wcale aż tak bardzo oczywiste [197] — że F jest mierzalna. Ponieważ zachodzi

$$\Gamma_f = F^{-1}(\{0\}),$$

zatem $\Gamma_f \in \mathcal{L}(\mathbb{R}^{d+1})$. Użyjemy twierdzenia Fubiniego dzieląc zmienne $x_1, \ldots, x_d, x_{d+1}$ na dwie grupy: $x_1, \ldots, x_d$ oraz x_{d+1} (dla zmiennej z $\mathbb{R}^d$ związanej z pierwszą grupą użyjemy oznaczenia x):

$$\lambda^{d+1}(\Gamma_f) = \int_{\mathbb{R}^{d+1}} \chi_{\Gamma_f} d\lambda^{d+1} = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} \chi_{\Gamma_f}(x, x_{d+1}) dx_{d+1} \right) dx$$

[197] Patrz zadanie X.40.



Rysunek 33:

Zauważmy, że dla ustalonego $x \in \mathbb{R}^d$ funkcja χ_{Γ_f} ma w punkcie (x, x_{d+1}) wartość różną od zera (mianowicie 1) wtw $(x, x_{d+1}) \in \Gamma_f$ wtw $x \in D$ oraz $x_{d+1} = f(x)$ — patrz rys. 33. Mamy zatem

$$\begin{aligned}\lambda^{d+1}(\Gamma_f) &= \int_D \Big(\int_{\{f(x)\}} 1 dx_{d+1}\Big) dx \\ &= \int_D \lambda^1(\{f(x)\}) dx = \int_D 0 dx = 0.\end{aligned}$$

□

4. Całkowanie przez podstawienie. Współrzędne biegunowe i sferyczne.

Przedstawimy tu twierdzenie pozwalające na tzw. „zamianę zmiennych" w całce względem miary Lebesgue'a λ^d. Będzie to nieco podobne do zamian zmiennych (podstawiania) związanych z całkami oznaczonymi (a przy $d = 1$, w typowych sytuacjach, będzie to wręcz identyczne). Stosowanie tego twierdzenia często znacznie upraszcza rachunki towarzyszące całkowaniu w wielu wymiarach.

Niech U będzie otwartym podzbiorem $\mathbb{R}^d$ oraz niech $\Phi : U \to \mathbb{R}^d$ będzie funkcją różniczkowalną (może lepiej użyć tu słowa „przekształcenie" zamiast funkcja, Φ ma opisywać właśnie ową „zamianę zmiennych"). Przypomnijmy, że w takiej sytuacji w każdym punkcie x zbioru U określany jest jakobian $J\Phi(x)$ funkcji Φ, będący wyznacznikiem różniczki $D\Phi(x)$. Pełni on rolę podobną do pochodnej funkcji wewnętrznej (zamiany zmiennych) pojawiającej się we wzorze na całkowanie przez podstawienie dla całki oznaczonej.

Twierdzenie X.7 (o całkowaniu przez podstawienie). *Niech Φ będzie dyfeomorfizmem zbioru otwartego $U \subset \mathbb{R}^d$ o wartościach w $\mathbb{R}^d$, niech $A \subset U$ oraz $f : \Phi(A) \to \overline{\mathbb{R}}$ niech będzie funkcją mierzalną nieujemną lub funkcją całkowalną. Określmy $f_\Phi : A \to \overline{\mathbb{R}}$ dla $x \in A$ wzorem*

$$f_\Phi(x) = f(\Phi(x)) \cdot |J\Phi(x)|.$$

Wówczas f_Φ jest mierzalna, a w przypadku całkowalności f — funkcja f_Φ jest całkowalna i w obu przypadkach

$$\int_A f_\Phi d\lambda^d = \int_{\Phi(A)} f d\lambda^d,$$

tzn.

$$\int_A f(\Phi(x)) \cdot |J\Phi(x)| dx = \int_{\Phi(A)} f(x) dx. \tag{X.11}$$

B.D.

Uwagi.

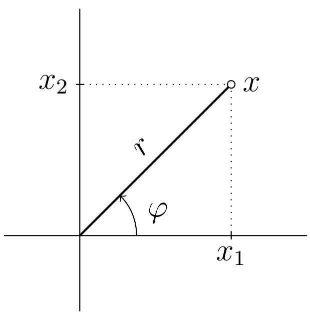
1. Dziwić może nieco, że we wzorze (X.11) na całkowanie przez podstawienie jakobian $J\Phi(x)$ pojawia się w module, podczas gdy modułu tego nie było przy $g'(x)$ we wzorze dla całki oznaczonej:
$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(x)dx.$$
Wyjaśnienie tego problemu tkwi w notacji „$\int_a^b$", która nie zawsze oznacza to samo co „$\int_{[a;b]}$". Nawet jeśli $a \leqslant b$, to nie gwarantuje to, że $g(a) \leqslant g(b)$. Gdy g jest malejąca, to $g'(x) \leqslant 0$ i jednocześnie $g(a) \geqslant g(b)$. Stąd, po przemnożeniu obu całek przez -1, powyższy wzór będzie miał następującą postać z użyciem zapisu dla całek Riemanna „po zbiorach" zamiast całek oznaczonych „od... do...":
$$\int_{[a;b]} f(g(x))|g'(x)|dx = \int_{g([a;b]} f(x)dx,$$
czyli tak jak w (X.11).

2. Warto zwrócić uwagę, że w twierdzeniu X.7 zakładamy, że Φ jest dyfeomorfizmem. W szczególności Φ jest różnowartościowa. Jednak wzór dla całki oznaczonej działał także bez założenia o różnowartościowości g.

Najczęściej stosujemy twierdzenie X.7 w ten sposób, że dla danej funkcji $f : D \to \mathbb{R}$ próbujemy przedstawić D w postaci $\Phi(A)$ dla tak dobranego Φ i A, że lewa strona wzoru (X.11) jest możliwa do wyliczenia w praktyce (zazwyczaj, już wtedy, poprzez użycie twierdzenia Fubiniego). Często bywa tak, że udaje się uzyskać jedynie $D =$ „ prawie $\Phi(A)$", tzn. $\lambda^d(D \setminus \Phi(A)) = 0$, ale to w zupełności wystarcza, bo całka po zbiorze miary zerowej równa jest zero.

Opiszemy teraz dwa najczęściej chyba stosowane podstawienia — jedno dla $d = 2$, drugie dla $d = 3$.

Współrzędne biegunowe w $\mathbb{R}^2$. Pomysł polega tu na tym, by punkt $x \in \mathbb{R}^2$ opisywać nie za pomocą jego współrzędnych kartezjańskich $x_1, x_2 \in \mathbb{R}$, ale za pomocą tzw. *współrzędnych biegunowych*, tzn. kąta pomiędzy osią poziomą a wektorem x, oznaczonego jako φ oraz długości wektora x, oznaczanej przez r.



Rysunek 34:

Zachodzi zatem
$$\begin{aligned} x_1 &= r\cos\varphi \\ x_2 &= r\sin\varphi \end{aligned}$$
przy czym $\varphi \in (0; 2\pi]$, $r \geqslant 0$ (patrz rys. 34). Jednak ponieważ w twierdzeniu X.7 musimy mieć zbiór otwarty U i dyfeomorfizm Φ, zatem musimy zakres rozważanych φ i r nieco ograniczyć.

Definiujemy mianowicie przekształcenie $\Phi : U \to \mathbb{R}^2$, gdzie $U = (0; 2\pi) \times (0; +\infty)$, zadane wzorem

$$\Phi(\varphi, r) = (r\cos\varphi, r\sin\varphi).$$

Łatwo wykazać, że Φ jest dyfeomorfizmem na swój obraz $V := \mathbb{R}^2 \setminus Z$ [198)], gdzie $Z = [0; +\infty) \times \{0\}$, zatem $\lambda^2(Z) = 0$. Zatem $\mathbb{R}^2 =$„prawie $\Phi(U)$”... Policzmy jeszcze jakobian. Mamy

$$J\Phi(\varphi, r) = \det\begin{pmatrix} -r\sin\varphi & \cos\varphi \\ r\cos\varphi & \sin\varphi \end{pmatrix} = -r\sin^2 - r\cos^2 = -r,$$

a stąd $|J\Phi(\varphi, r)| = r$ dla dowolnych $(\varphi, r) \in U$.

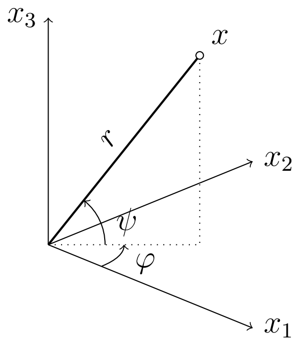
Przykład. Obliczmy ponownie pole powierzchni koła o promieniu $R > 0$ (patrz przykład 2 ze strony 32). Koło $K_R = \{x \in \mathbb{R}^2 : \|x\| < R\}$ jest postaci $\Phi(A)$ „z dokładnością do zbioru miary zero” dla $A = (0; 2\pi) \times (0; R)$ (tzn. $K_r \setminus Z = \Phi(A)$). A zatem na mocy twierdzenia X.7, a następnie twierdzenia X.6 mamy:

$$\lambda^2(K_R) = \lambda^2(K_R \setminus Z) = \int_{K_R\setminus Z} 1dx = \int_{\Phi(A)} 1dx = \int_A 1 \cdot rd(\varphi, r) =$$
$$= \int_0^{2\pi}\left(\int_0^R rdr\right)d\varphi = \int_0^R rdr \cdot \int_0^{2\pi} 1dy = \frac{1}{2}R^2 \cdot 2\pi = \pi R^2.$$

A zatem rachunki okazały się tu dużo prostsze niż przy bezpośrednim stosowaniu twierdzenia Fubiniego.

Ogólnie — warto z pewnością próbować stosować to podstawienie w sytuacji, gdy zbiór ma symetrię obrotową wokół zera, choć nie tylko wtedy ...

Współrzędne sferyczne w $\mathbb{R}^3$. Sens geometryczny tych współrzędnych jest dość podobny do sensu zmiennych biegunowych z tym, że użyty jest jeszcze jeden kąt, a sam pomysł jest dość dobrze znany z geografii. Spotkać się można nawet z określeniem „współrzędne geograficzne”, odnoszacym się do omawianego tu podstawienia. Tu bowiem punkt x z $\mathbb{R}^3$ opisany jest przy pomocy długości geograficznej φ i szerokości geograficznej ψ wyrażonych w odpowiednich kątach, oraz przy pomocy długości r wektora x (patrz rys. 35). Tym razem rozważamy $\varphi \in (-\pi, \pi]$, $\psi \in [-\frac{\pi}{2}; \frac{\pi}{2}]$ oraz $r \geqslant 0$. Dla celów twierdzenia X.7 określamy zatem $\Phi : U \to \mathbb{R}^3$,



Rysunek 35:

gdzie $U := (-\pi, \pi) \times (-\frac{\pi}{2}; \frac{\pi}{2}) \times (0; +\infty)$ oraz

$$\Phi(\varphi, \psi, r) = (r\cos\psi \cdot \cos\varphi, r\cos\psi \cdot \sin\psi, r\sin\psi).$$

Tym razem Φ jest też dyfeomorfizmem na swój obraz $V := \mathbb{R}^3 \setminus Z$, gdzie Z ma miarę λ^3 zerową, przy czym $Z = (-\infty; 0] \times \{0\} \times \mathbb{R}$. Jak nietrudno wyliczyć (zad. X.38), jakobian

198) Najlepiej znaleźć wzór na Φ^{-1}, co nie jest trudne.

zadaje się tym razem wzorem

$$J\Phi(\varphi,\psi,r) = r^2\cos\psi = |J\Phi(\varphi,\psi,r)|$$

bo $\psi \in (-\frac{\pi}{2};\frac{\pi}{2}))$ dla $(\varphi,\psi,r) \in U$.

Przykład. Obliczymy objętość kuli o promieniu $R > 0$, tzn. miarę zbioru $V_R := \{x \in \mathbb{R}^3 : \|x\| < R\}$. Postępując podobnie jak w przykładzie poprzednim otrzymujemy $\lambda^3(V_R) = \lambda^3(\Phi(A))$ dla $A = (-\pi;\pi) \times (-\frac{\pi}{2};\frac{\pi}{2}) \times (0;R)$, skąd

$$\lambda^3(V_R) = \int_{\Phi(A)} 1dx = \int_A 1\cdot r^2\cos\psi d\,(\varphi,\psi,r) = \int_{-\pi}^{\pi}(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}(\int_0^R r^2\cos\psi dr)d\psi)d\varphi =$$
$$= \int_0^R r^2dr\cdot\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos\psi d\,\psi\cdot\int_{-\pi}^{\pi}1d\varphi = \frac{1}{3}R^3\cdot 2\cdot 2\pi = \frac{4}{3}\pi R^3.$$

Zadania do Rozdziału X

$\forall$ 1.

(a) Znajdź wszystkie σ-ciała podzbiorów X, gdy $X =$

i. $\{1, 2\}$

ii. $\{1, 2, 3\}$

(b) Znajdź wszystkie elementy σ-ciała generowane przez $\{\{1, 2\}, \{2, 3\}\}$ dla $X = \{1, 2, 3, 4\}$.

$\forall$ 2. Wykaż, że przecięcie dowolnej rodziny σ-ciał podzbiorów zbioru X jest też σ-ciałem (podzbiorów X). Znajdź przykład pokazujący, że analog powyższego dla sumy nie jest prawdziwy.

3. σ-ciało $\mathbb{M}$ podzbiorów X nazywamy *cegiełkowym* wtw istnieje rodzina $\{A_i\}_{i\in I}$ podzbiorów X taka, że I jest skończony lub przeliczalny, A_i są parami rozłączne, $\bigcup_{i\in I} A_i = X$ oraz $\mathbb{M} = \sigma(\{A_i\}_{i\in I})$ [199]. Rodzina $\{A_i\}_{i\in I}$ nazywa się wówczas zbiorem (rodziną) *cegiełek* dla $\mathbb{M}$ (a każdy A_i jest *cegiełką*).

$\forall$ (a) Wykaż, że jeśli $\{A_i\}_{i\in I}$ jest zbiorem cegiełek dla $\mathbb{M}$, to

$$\mathbb{M} = \{\bigcup_{i\in I'} A_i : I' \subset I\}.$$

(b) Wykaż, że zbiór cegiełek σ-ciała cegiełkowego jest jednoznacznie wyznaczony.

(c) Wykaż, że jeśli X jest zbiorem skończonym, to każde σ-ciało jego podzbiorów jest cegiełkowe.

(d) Oblicz liczbę wszystkich σ-ciał podzbiorów zbioru $\{1, \ldots, n\}$ (może być dla $n = 10$).

(e) Wykaż, że $B(\mathbb{R})$ nie jest σ-ciałem cegiełkowym.

4. Niech $X \cap Y = \emptyset$ i niech $\mathcal{A} \subset 2^x$, $\mathcal{B} \subset 2^Y$ oraz niech $Z = X \cup Y$. Wykaż, że $\sigma_Z(\sigma_X(\mathcal{A}) \cup \sigma_Y(\mathcal{B})) = \sigma_Z(\mathcal{A} \cup \mathcal{B})$, gdzie $\sigma_W(\cdot)$ oznacza odpowiednie σ-ciało podzbiorów zbioru W, przy $W = X, Y, Z$.

$\forall$ 5. Udowodnij wniosek ze strony 153.

6. Rozważamy następujące rodziny podzbiorów $\mathbb{R}$:

- $\mathcal{A}_1 = \{(a; b) : a, b \in \mathbb{R}\}$,
- $\mathcal{A}_2 = \{[a; b] : a, b \in \mathbb{R}\}$,
- $\mathcal{A}_3 = \{(-\infty; a] : a \in \mathbb{R}\}$,
- $\mathcal{A}_4 = \{(a; +\infty) : a \in \mathbb{Q}\}$,
- $\mathcal{A}_5 = \{(a; b] : a, b \in \mathbb{Q}\}$,
- $\mathcal{A}_6 = \{U \subset \mathbb{R} : U$ — otwarty$\}$,
- $\mathcal{A}_7 = \{K \subset \mathbb{R} : K$ — zwarty$\}$.

Wykaż, że:

$\forall$ (a) $\sigma(\mathcal{A}_1) = \sigma(\mathcal{A}_2)$,

[199] Czyli krócej: $\mathbb{M}$ jest generowane przez pewne co najwyżej przeliczalne rozbicie zbioru X.

∀ (b) $\sigma(\mathcal{A}_2) = \sigma(\mathcal{A}_3)$,

(c) $\sigma(\mathcal{A}_2) = \sigma(\mathcal{A}_4)$,

∀ (d) $\sigma(\mathcal{A}_1) = \sigma(\mathcal{A}_5)$,

(e) $\sigma(\mathcal{A}_1) = \sigma(\mathcal{A}_6)$, ($= \mathcal{B}(\mathbb{R})$ z definicji),

(f) $\sigma(\mathcal{A}_6) = \sigma(\mathcal{A}_7)$.

7. (a) Wykaż, że dla dowolnego $A \in \mathcal{B}(\mathbb{R})$ zachodzi $\mathbb{R}^k \times A \times \mathbb{R}^l \in \mathcal{B}(\mathbb{R}^{k+1+l})$ (ew. tylko przypadek $k = 0, l = 1$);

   (b) Wykaż analogiczny wynik jak w a) dla $\mathcal{L}(\mathbb{R})$ oraz $\mathcal{L}(\mathbb{R}^{k+1+l})$.

   Wskazówka do a): wykaż najpierw, że rodzina $\{A \in \mathcal{B}(\mathbb{R}) : \mathbb{R}^k \times A \times \mathbb{R}^l \in B(\mathbb{R}^{k+l+1})\}$ jest σ-ciałem podzbiorów $\mathbb{R}$.

8. Wykaż, że jeżeli $\mathcal{A}$ i $\mathcal{B}$ są rodzinami generatorów dla $\mathcal{B}(\mathbb{R})$, to $\{A \times B \subset \mathbb{R}^2 : A \in \mathcal{A}, B \in \mathcal{B}\}$ jest rodziną generatorów dla $\mathcal{B}(\mathbb{R}^2)$.

9. Korzystając z zadania X.6 i X.8 wykaż, że rodzina $\{[a;b] \times [c;d] : a, b, c, d \in \mathbb{R}\}$ jest rodziną generatorów dla $\mathcal{B}(\mathbb{R}^2)$.

∀ 10. Wykaż, że są miarami:

   (a) δ_{x_0} z przykładu 2 strona 155;

   (b) # z przykładu 3 strona 156;

   (c) $\mu : \mathbb{M} \to [0; +\infty]$ zadana na σ-ciele $\mathbb{M} = \{\emptyset, \{1\}, \{2,3\}, \{1,2,3\}\}$ podzbiorów zbioru $\{1,2,3\}$ w sposób następujący: $\mu(\emptyset) = 0$, $\mu(\{1\}) = a$, $\mu(\{2,3\}) = b$, $\mu(\{1,2,3\}) = a + b$, gdzie a, b są dowolnymi elementami z $[0; +\infty]$.

∀ 11. labelzadX.10d Opisz wszystkie miary określone na σ-ciele cegiełkowym (patrz zadanie X.3).

∀ 12. Wykaż ogólne własności miary z faktu ze strony 155.

∀ 13. Czy to prawda, że miara μ określona na σ-ciele generowanym przez rodzinę $\mathcal{A}$ jest jednoznacznie wyznaczona przez swe wartości na wszystkich generatorach?

14. Czy powyżej zmieniłoby coś ograniczenie się do miar spełniających warunek $\mu(X) = 1$ (czyli *probabilistycznych*)?

    Wskazówka: rozważ miary określone na σ-ciele z zadania X.1 b).

∀ 15. Niech $\mathbb{M}$ będzie σ-ciałem podzbiorów X oraz $X' \in \mathbb{M}$ i niech μ będzie miarą określoną na $\mathbb{M}$. Wykaż, że $\mathbb{M}' := \{A \in \mathbb{M} : A \subset X'\}$ jest σ-ciałem podzbiorów X', oraz że $\mu' : \mathbb{M}' \to [0; +\infty]$ zadana jako $\mu' = \mu|_{\mathbb{M}'}$ jest miarą (w X').

∀ 16. Niech $\mathbb{M}$ będzie σ-ciałem podzbiorów X, μ — miarą na $\mathbb{M}$, Y niech będzie dowolnym zbiorem oraz niech $f : X \to Y$. Określamy $\mathbb{M}_f := \{A \subset Y : f^{-1}(A) \in \mathbb{M}\}$ oraz $\mu_f : \mathbb{M}_f \to [0; +\infty]$ zadajemy wzorem

$$\mu_f(A) = \mu(f^{-1}(A)) \text{ dla } A \in \mathbb{M}_f.$$

Wykaż, że $\mathbb{M}_f$ jest σ-ciałem podzbiorów Y, f — funkcją $\mathbb{M}$, $\mathbb{M}_f$–mierzalną oraz μ_f — miarą w Y.

17. Wykaż, że suma (a także kombinacja liniowa z nieujemnymi współczynnikami) miar określonych na $\mathbb{M}$ jest miarą.

$\forall$ 18. Czy poniższe własnosci są prawdziwe dla wszystkich miar $\mu : \mathbb{M} \to [0; +\infty]$?

 (a) Jeżeli $\forall_{n\in\mathbb{N}}$ $(A_n \in \mathbb{M}$ oraz $A_{n+1} \subset A_n)$, to $\mu(\bigcap_{n\in N} A_n) = \lim_{n\to+\infty} \mu(A_n)$

 (b) Jeżeli $\forall_{n\in\mathbb{N}}$ $(A_n \in \mathbb{M}$ oraz $A_n \subset A_{n+1})$, to $\mu(\bigcup_{n\in N} A_n) = \lim_{n\to+\infty} \mu(A_n)$

 Czy sytuacja zmieni się, gdy założymy dodatkowo, że $\forall_{n\in\mathbb{N}}$ $\mu(A_n) < +\infty$?

19. Niech μ będzie miarą określoną na σ-ciele $\mathbb{M}$ podzbiorów X. Definiujemy $\mathcal{Z} := \{Z \subset X : \exists_{B\in\mathbb{M}}\ (\mu(B) = 0 \text{ i } Z \subset B)\}$ oraz $\mathbb{M}_u := \sigma(\mathbb{M} \cup \mathcal{Z})$ (tzw. *σ-ciało uzupełnione względem μ*).

 (a) Wykaż, że σ-ciało $\mathbb{M}_u$ ma postać

$$\mathbb{M}_u = \{A \cup Z \subset X : A \in \mathbb{M} \text{ oraz } Z \in \mathcal{Z}\}$$

 (b) Definiujemy $\mu_u : \mathbb{M}_u \to [0; +\infty]$ wzorem

$$\mu_u(A \cup Z) = \mu(A)$$

 dla dowolnych $A \in \mathbb{M}$, oraz $Z \in \mathcal{Z}$. Wykaż, że definicja ta jest poprawna, tzn. że nie zależy od wyboru A i Z jak wyżej, ale jedynie od $A \cup Z$.

 (c) Wykaż, że zdefiniowana wyżej funkcja μ_u jest miarą zupełną, oraz że $\mu_{u|\mathbb{M}} = \mu$.

20. W przedziale $[-1; 1]$ rozważamy relację $\sim$ zdefiniowaną następująco: $x \sim y$ wtw $x - y \in \mathbb{Q}$.

 (a) Wykaż, że $\sim$ jest relacją równoważności.

 $\forall$ (b) Niech $K = \{[x] : x \in [0; 1]\}$ — zbiór klas abstrakcji relacji $\sim$ i niech $w : K \to [0; 1]$ będzie pewną „funkcją wyboru" dla zbioru K, tzn. dla dowolnej klasy $k \in K$ $w(k) \in k$ (inaczej mówiąc „w wybiera z każdej klasy abstrakcji po jednym jej elemencie" — korzystamy tu z pewnika wyboru, by mieć gwarancję, że taka w istnieje). Wykaż, że $A := \{w(k) : k \in K\}$ nie jest mierzalny w sensie Lebesgue'a (w szczególności $A \notin \mathcal{B}(\mathbb{R})$).

21. Określamy pewien ciąg przedziałów otwartych zawartych w $[0; 1]$ w sposób następujący (rekurencyjnie). W pierwszym kroku określamy jeden przedział o środku $\frac{1}{2}$ i długości $r_1 \in (0; 1)$. W $n + 1$–szym kroku rozważamy najpierw wszystkie przedziały domknięte, których rozłączną sumę stanowi różnica przedziału $[0; 1]$ i sumy wszystkich przedziałów otwartych uzyskanych w krokach $1, \ldots, n$. Następnie tworzymy przedziały otwarte biorąc jako ich środki wszystkie środki powyższych przedziałów domkniętych oraz wybierając wspólną długość r_{n+1}, tak by tworzone przedziały zawierały się w tych przedziałach domkniętych. Niech $\mathcal{C}$ będzie uzupełnieniem w $[0; 1]$ sumy wszystkich uzyskanych tak przedziałów. Gdy $r_n = \frac{1}{3^n}$, to uzyskujemy tzw. *zbiór Cantora*.

 $\forall$ (a) Wykaż, że zbiór Cantora jest nieprzeliczalnym zbiorem (i nieskończonym) o mierze Lebesgue'a zerowej.

 (b) Czy można tak dobrać ciąg $\{r_n\}_{n\geqslant1}$, by uzyskać $\lambda^1(\mathcal{C}) > 0$?

22. Znajdź przykład funkcji $f : X \to \mathbb{R}$ niemierzalnej

    (a) przy dowolnie przez siebie wybranym X i σ-ciele $\mathbb{M}$ podzbiorów X;

    (b) dla $X = \mathbb{R}$ i $\mathbb{M} = \mathcal{L}(\mathbb{R})$ (Wskazówka: użyj zadania X.20).

23. Wykaż, że funkcja $f : D \to \overline{\mathbb{R}}$ jest mierzalna (przy zadanym σ-ciele $\mathbb{M}$ w X, $D \subset X$) wtw $\forall_{r\in\mathbb{R}}\ \{x \in D : f(x) < r\} \in \mathbb{M}$ oraz $D \in \mathbb{M}$.

$\forall$ 24. Wykaż fakt ze strony 160 o mierzalności punktowej granicy ciągu funkcji mierzalnych.

25. Wykaż, że fakt ze strony 159 (o działaniach na funkcjach mierzalnych) można rozszerzyć na funkcje mierzalna o wartościach w $\overline{\mathbb{R}}$, pod warunkiem określoności działań.

26. Niech μ będzie miarą określoną na σ-ciele $\mathbb{M}$ podzbiorów X. Określimy tzw. zbieżność *μ–prawie wszędzie* ciągu funkcyjnego $\{f_n\}_{n\geqslant n_0}$ o wyrazach $f_n : D \to \overline{\mathbb{R}}$, gdzie $D \subset X$, do funkcji $f : D \to \overline{\mathbb{R}}$; będziemy to oznaczać symbolem $f_n \xrightarrow[\text{p.w.}]{} f$. Mianowicie $f_n \xrightarrow[\text{p.w.}]{} f$ wtw $\exists_{\substack{Z\in\mathbb{M}\\ \mu(Z)=0}}\ f_n|_{D\setminus Z} \to f|_{D\setminus Z}$.

    (a) Czy granica przy takiej zbieżności jest funkcją wyznaczoną jednoznacznie?

    (b) Wykaż, że jeśli μ jest miarą zupełną, $f_n \xrightarrow[\text{p.w.}]{} f$ oraz f_n są wszystkie mierzalne, to f też jest mierzalna.

    (c) Wyjaśnij dlaczego założenie zupełności jest istotne w b), nawet przy założeniu, że $D \in \mathbb{M}$.

$\forall$ 27. Wykaż, że jeśli f jest mierzalną funkcją, to f^+ i f^- także są mierzalne.

28. Wykaż, że każda funkcja mierzalna jest granicą punktową ciągu funkcji prostych. Wykaż, że ciąg ten można wybrać rosnącym, jeśli wyjściowa funkcja jest $\geqslant 0$.

$\forall$ 29. Wykaż informacje o całkowaniu względem miary Diraca δ_{x_0} zawarte w przykładzie 2 ze strony 162.

30. Wykaż informacje o całkowaniu względem miary liniowej $\#$ w $\mathbb{N}$ zawarte w przykładzie 3 ze strony 162.

$\forall$ 31. Wykaż, że jeśli $f \in L^1(D,\mu)$, to $|\int_D f d\mu| \leqslant \int_D |f| d\mu$.

32. Rozwiąż zadanie III.17 wykorzystując twierdzenie Lebesgue'a o zbieżności majoryzowalnej (tw. X.3 (L)).

$\forall$ 33. Wykaż, że jeżeli $f \in L^1(D,\mu)$, $\forall_{n\in\mathbb{N}}\ (A_{n+1} \supset A_n$ i A_n jest mierzalny), oraz $\bigcup_{n\in\mathbb{N}} A_n = D$, to

$$\lim_{n\to+\infty} \int_{A_n} f d\mu = \int_D f d\mu.$$

34. Wykaż twierdzenie X.5 (o związku całki niewłaściwej z całką względem λ^1).

$\forall$ 35. Wykaż, ża całka Lebesgue'a dla funkcji f z przykładu ze strony 166 jest nieokreślona.

$\forall$ 36. Znajdź przykład pokazujący, że w twierdzeniu Fubiniego (tw. X.6) funkcja $\mathcal{C}_2 f$ może nie być określona na całym zbiorze $\mathbb{R}^{d_1}$.

37. Wykaż mierzalność funkcji F z dowodu faktu ze strony 171. Wskazówka: użyj wyniku z zadania X.8 b).

38. Sprawdź wzór na jakobian „dla współrzędnych sferycznych” ze strony 175.

39. Oblicz:

$\forall$ (a) $\int_{[0;1]\times[0;2]} e^{x_1+x_2} dx$

(b) $\int_{[0;1]\times[1;\sqrt{3}]} \frac{x_1^2}{1+x_2^2} dx$

$\forall$ (c) $\int_{[-1;1]\times[0;1]} (x_1+x_2)^{2222} dx$

(d) $\int_{[1;2]\times[3;4]} \ln(x_1 x_2)\cdot x_2 dx$

$\forall$ (e) $\lambda^2(T)$, gdzie T — trójkąt „pełny” na płaszczyźnie $\mathbb{R}^2$ o wierzchołkach 0; $(a,0)$; (c,h), gdzie $a,h,c>0$,

(f) $\lambda^2(A)$, gdzie $A=\{x\in\mathbb{R}^2 : x_1\geqslant 0, x_1^2\leqslant x_2\leqslant\sqrt{x_1}\}$,

$\forall$ (g) $\int_R x_1 dx$, gdzie R — równoległobok „pełny” na płaszczyźnie $\mathbb{R}^2$ o wierzchołkach 0; $(2,1)$; $(1,1)$; $(3,2)$. Wskazówka: użyj całkowania przez (odpowiednie) podstawienie (równoległobok jest afinicznym obrazem kwadratu $[0;1]\times[0;1]\ldots$),

(h) jak powyżej, ale dla wierzchołków: $(2,2)$; $(4,3)$, $(3,3)$, $(5,4)$,

(i) $\int_T \cos(x_1+x_2)dx$, gdzie T—„pełen” trójkąt ograniczony prostymi o równaniach: $x_1=0$, $x_2=\pi$, $x_1=x_2$,

(j) $\int_{K(0,1)} x_1x_2 dx$,

(k) $\int_{K_{+++}} x_1x_2 dx$, gdzie $K_{+++}=\{x\in\mathbb{R}^3 : x_1,x_2,x_3>0, \|x\|<1\}$,

$\forall$ (l) $\int_{K_{+-}} x_3 dx$, gdzie $K_{+-}=\{x\in\mathbb{R}^3 : x_1>0, x_2<0, \|x\|<1\}$,

(m) $\int_{P_{1,2}} x_1^2+x_2^2 dx$, gdzie $P_{1,2}$ — pierścień kołowy („pełny”) na płaszczyźnie, o środku 0 i promieniu 1 (wewnętrznym) oraz 2 (zewnętrznym),

$\forall$ (n) $\lim_{n\to+\infty}\int_W \sqrt[n]{x_1x_2}dx$, gdzie $W=\{x\in\mathbb{R}^2 : \|x\|\leqslant 1, 0\leqslant x_1<x_2\}$,

(o) $\lambda^d(T_d(a))$, gdzie $T_d(a)=\{x\in\mathbb{R}^d : \sum_{k=1}^d x_k\leqslant a, \forall_{k=1,\ldots,d}\ x_k\geqslant 0\}, a>0$. Wskazówka: użyj indukcji „po d”,

(p) objętość elipsoidy $E=\{x\in\mathbb{R}^3 : \frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}+\frac{x_3^2}{c^2}\leqslant 1\}$, $a,b,c>0$,

$\forall$ (q) Objętość walca obrotowego o wysokości h i promieniu podstawy r,

(r) $\int_W x_1 dx$, gdzie $W=\{x\in\mathbb{R}^3 : x_1^2+x_2^2\leqslant 1,\ 0\leqslant x_3\leqslant 1\}$,

$\forall$ (s) Objętość stożka obrotowego o wysokości h i promieniu podstawy r,

(t) $\int_S |x_1x_2x_3|dx$, gdzie $S=\{x\in\mathbb{R}^3 : \rho\geqslant x_3\geqslant 3\sqrt{x_1^2+x_2^2}\}$,

(u) Objętość „pełnego” torusa powstałego przez obrót koła w płaszczyźnie „x_1, x_3” o środku $(R,0,0)$ i promieniu r wokół osi x_3 $(0<r<R)$,

$\forall$ (v) $\int_{\mathbb{R}^2} e^{-\|x\|^2}dx$,

$\forall$ (w) $\int_{\mathbb{R}^3} e^{-\|x\|^2}dx$. Wskazówka: wykorzystaj rezultat (nie metodę ...) dla przypadku $\mathbb{R}^2$ powyżej.

(x) $\int_B x_1x_2 dx$, gdzie $B=\{x\in\mathbb{R}^2 : x_1,x_2\leqslant 0,\ 1\leqslant 4x_1^2+x_2^2\leqslant 4,\ x_2\leqslant 2x_1\}$,

$\forall$ (y) $\int_{D_\alpha} x_1-x_2 dx$, gdzie $D_\alpha=\{x\in\mathbb{R}^2 : x_1\geqslant 1, x_1-\frac{1}{x_1^\alpha}\leqslant x_2\leqslant x_1+\frac{1}{x_1^\alpha}\}$ dla $\alpha>0$. Uwaga! Czy całka ta jest określona?

40. Wyprowadź wzór z zadania VII.2 na objętość bryły obrotowej.

XI Spis symboli i skrótów

- exp funkcja exponencjalna, 39
- sin sinus, 39
- cos kosinus, 39
- p.s. punkt skupienia, 43
- $\lim_{x\to a}$ granica funkcji w punkcie, 43
- $\xrightarrow[x\to a]{}$ ma granicę ... (funkcja w punkcie a), 43
- w. o wyrazach w, 43
- $\lim_{x\to a,\, x\in D}$ granica (funkcji) w punkcie a dla funkcji określonej na D, 44
- $\circ$ złożenie funkcji, 45
- $|_X$ obcięcie funkcji, 45
- D_+ część zbioru D położona na prawo od 0, 46
- D_- część zbioru D położona na lewo od 0, 46
- D_+^a część zbioru D położona na prawo od punktu a, 46
- D_-^a część zbioru D położona na lewo od punktu a, 46
- $\lim_{x\to a+}$ granica prawostronna w punkcie a, 46
- $\lim_{x\to a-}$ granica lewostronna w punkcie a, 46
- d. ... d.b. dla ... dostatecznie bliskich, 46
- $\mathbb{x}$ funkcja identycznościowa, 49
- $(a?b)$, $[a?b)$, $[a?b]$, $(a?b]$ przedziały „o nieznanej kolejności końców”, 50
- „na” funkcja „na”, 52
- „1–1” funkcja różnowartościowa, 52
- f^{-1} funkcja odwrotna, 52
- $[a;+\infty]$ przedział domknięty wraz z nieskończonością, 54
- $\log_a$ logarytm o podstawie a, 56
- ln logarytm naturalny, 56
- π liczba π, 58
- tg tangens, 58
- ctg kotangens, 58
- mian najmniejszy mianownik, 60
- $f'(a)$ pochodna funkcji w punkcie, 63

XII *

Skorowidz