Analiza Matematyczna I

Notatki do wykładu dla studentów Matematyki Komputerowej

Marian Mrozek
Uniwersytet Jagielloński

25 maja 2013

Spis treści

1 Wprowadzenie do analizy matematycznej **13**

1.1 Rozwój pojęcia liczby . . . 13
1.1.1 Liczby naturalne . . . 13
1.1.2 Ułamki . . . 13
1.1.3 Liczby ujemne, całkowite i wymierne . . . 13
1.1.4 Liczby niewymierne . . . 14
1.1.5 Liczby rzeczywiste . . . 15
1.1.6 Liczby urojone i zespolone . . . 15
1.1.7 Liczby w komputerze . . . 15
1.2 Ciągi . . . 17
1.2.1 Przybliżanie pierwiastka . . . 17
1.2.2 Ciągi przybliżeń . . . 17
1.2.3 Granice . . . 19
1.2.4 Obliczanie granic ciągu. . . . 20
1.2.5 Liczba π . . . 21
1.3 Szeregi . . . 22
1.3.1 Pole koła . . . 22
1.3.2 Liczba π jeszcze raz. . . . 22
1.3.3 Szeregi . . . 23
1.4 Funkcje . . . 23
1.4.1 Koncepcja funkcji . . . 23
1.4.2 Miejsca zerowe funkcji . . . 24
1.4.3 Wykres funkcji . . . 24
1.4.4 Geometryczne konstrukcje funkcji . . . 24
1.4.5 Rozwijanie funkcji w szereg . . . 27
1.5 Pola figur i całkowanie . . . 27
1.5.1 Całka oznaczona . . . 27
1.5.2 Całka nieoznaczona . . . 28
1.6 Styczna do krzywej i granica funkcji . . . 30
1.6.1 Styczna do krzywej . . . 30
1.6.2 Granica funkcji . . . 32
1.6.3 Funkcje ciągłe . . . 32
1.6.4 Obliczanie granic. . . . 33
1.7 Pochodna funkcji . . . 33
1.7.1 Lokalne maksima i minima . . . 33
1.7.2 Pochodna funkcji . . . 35
1.7.3 Podstawowe twierdzenie rachunku różniczkowego i całkowego . . . 36

2 Teoria mnogości **39**
2.1 Problemy z nieskończonością . . . 39
2.2 (*)Uproszczony formalizm . . . 40
2.3 Zbiory . . . 43
2.3.1 Operacje mnogościowe . . . 44
2.3.2 Podzbiory . . . 45
2.3.3 Elementarne własności operacji mnogościowych . . . 47
2.3.4 Sumy i iloczyny nieskończone . . . 48
2.4 Iloczyn kartezjański . . . 48
2.4.1 Para . . . 48
2.4.2 Iloczyn kartezjański . . . 50
2.5 Relacje . . . 51
2.5.1 Relacje . . . 51
2.5.2 Własności relacji . . . 53
2.5.3 Złożenie relacji . . . 54
2.5.4 Charakterystyka własności relacji . . . 54
2.6 Relacje równoważności . . . 54
2.6.1 Relacja równoważności . . . 54
2.6.2 Równoliczność . . . 55
2.7 Funkcje . . . 55
2.7.1 Relacje jednolistne . . . 55
2.7.2 Funkcje częściowe . . . 55
2.7.3 Funkcje . . . 57
2.7.4 (*)Zawieranie się funkcji . . . 57
2.7.5 Sklejenie funkcji . . . 58
2.7.6 (*)Zestawienie i iloczyn kartezjański funkcji . . . 58
2.7.7 Obrazy i przeciwobrazy . . . 59
2.7.8 Złożenie funkcji . . . 59
2.7.9 Injekcje, surjekcje, i bijekcje. . . . 60
2.7.10 (*)Iloczyn kartezjański n zbiorów . . . 60
2.7.11 (*)Sumy i iloczyny indeksowane . . . 60
2.8 Struktury porządkowe . . . 61
2.8.1 Częściowy i liniowy porządek . . . 61
2.8.2 Majoranty i minoranty . . . 62
2.9 Przestrzenie ciągłe . . . 63
2.9.1 Kresy . . . 63
2.9.2 Przedziały . . . 64
2.9.3 Własność L . . . 65
2.9.4 Przestrzenie ciągłe . . . 65
2.10 Funkcje monotoniczne . . . 66
2.10.1 Funkcje rosnące i malejące . . . 66

3 Liczby naturalne, całkowite i wymierne **67**
3.1 Liczby naturalne . . . 67
3.1.1 Aksjomatyka liczb naturalnych . . . 67
3.1.2 (*)Konstrukcja von Neumanna . . . 68
3.1.3 (*)Jednoznaczność . . . 69
3.2 Ciągi i podciągi . . . 69
3.2.1 Definicja ciągu i podciągu . . . 69

3.2.2 Ciąg przybliżeń pierwiastka jeszcze raz. . . . 70
3.2.3 Monotoniczność ciągu. . . . 71
3.2.4 Rekurencyjne definiowanie ciągu. . . . 72
3.2.5 Pierwiastek poprzez ciąg rekurencyjny. . . . 72
3.2.6 (**)Arytmetyka liczb natualnych 73
3.2.7 Zbiory skończone i przeliczalne 74
3.2.8 Rozszerzona zasada indukcji 74
3.2.9 Silnia i symbol Newtona 75
3.3 Liczby całkowite. . . . 76
3.3.1 Działania 76
3.3.2 Grupy 78
3.3.3 Zbiór liczb całkowitych 79
3.3.4 Sumy i iloczyny 80
3.4 Liczby wymierne 80
3.4.1 Ciała 80
3.4.2 Funkcja potęgowa i wielomian 81
3.4.3 Zbiór liczb wymiernych 81
3.4.4 Wzory skróconego mnożenia 81
3.5 Liczby reprezentowalne 82

4 Liczby rzeczywiste i liczby zespolone 87
4.1 Zbiór liczb rzeczywistych 87
4.1.1 Ciała uporządkowane 87
4.1.2 Zbiór liczb rzeczywistych 88
4.1.3 (**)Konstrukcja Cauchy'ego 88
4.1.4 Podzbiory $\mathbb{R}$ 89
4.1.5 Funkcja potęgowa w $\mathbb{R}$ 89
4.1.6 Nierówność Bernoullego 90
4.1.7 Wartość bezwzględna 90
4.2 Twierdzenie o istnieniu pierwiastka. . . . 91
4.2.1 Własność Archimedesa 91
4.2.2 Gęstość zbioru w zbiorze 91
4.2.3 Pierwiastki 92
4.3 Rozszerzony zbiór liczb rzeczywistych 93
4.3.1 $\mathbb{R}$ uzupełniony o $-\infty$ i $+\infty$ 93
4.3.2 Kresy funkcji 94
4.4 Ciało liczb zespolonych 96
4.4.1 Liczby zespolone 96
4.4.2 Ciało $\mathbb{C}$ 96
4.4.3 Zespolona funkcja kwadratowa 98
4.4.4 Liczby zespolone o module jeden. . . . 98
4.4.5 Odległość w $\mathbb{R}$ i $\mathbb{C}$. . . . 99
4.5 Granice ciągów w $\mathbb{R}$ 99
4.5.1 Eksperymenty numeryczne 99
4.5.2 Granice pozorne 103
4.5.3 Definicja Cauchy'ego granicy ciągu 103
4.5.4 Granica ciągu, a struktura ciała w $\mathbb{R}$ 106
4.5.5 Granica ciągu, a struktura porządkowa w $\mathbb{R}$ 107
4.6 Przykłady i zastosowania granic ciągów w $\mathbb{R}$ 109

4.6.1 Granice kilku ważnych ciągów . . . 109
4.6.2 Procent składany . . . 110
4.6.3 Liczba e . . . 112
4.6.4 Funkcja wykładnicza . . . 113

5 Topologia przestrzeni metrycznych 115
5.1 Przestrzenie metryczne . . . 115
5.1.1 Przestrzeń metryczna . . . 115
5.1.2 Przykłady metryk w $\mathbb{R}^n$ . . . 116
5.1.3 Metryka indukowana i podprzestrzenie przestrzeni metrycznej . . . 116
5.1.4 Kule . . . 117
5.1.5 Zbiory ograniczone . . . 118
5.2 Średnica zbiorów . . . 118
5.2.1 Średnica zbioru . . . 118
5.3 Zbiory otwarte i domknięte, wnętrze, domknięcie i brzeg zbioru . . . 119
5.3.1 Wnętrze i domknięcie . . . 120
5.3.2 Własności wnętrza i domknięcia . . . 120
5.4 Otoczenia, punkty skupienia . . . 121
5.4.1 Rodzina otoczeń punktu . . . 121
5.4.2 Punkty skupienia zbioru . . . 121
5.5 Metryki równoważne, iloczyn kartezjański przestrzeni metrycznych . . . 122
5.5.1 Równoważność metryk. . . . 122
5.5.2 Iloczyn kartezjański przestrzeni metrycznych . . . 122
5.6 (*)Topologia przestrzeni metrycznej . . . 122
5.6.1 Fundamentalne własności zbiorów otwartych i domkniętych . . . 122
5.7 (**)Bazy otoczeń . . . 123
5.7.1 Baza otoczeń punktu . . . 123
5.7.2 Własność Hausdorffa . . . 124
5.7.3 Metryka w $\bar{\mathbb{R}}$ . . . 124
5.8 (**)Przeliczalne bazy otoczeń punktu . . . 125
5.9 Zupełność . . . 125
5.9.1 Warunek Cauchy'ego . . . 125
5.9.2 Przestrzenie zupełne . . . 126

6 Granica funkcji 127
6.1 Granica funkcji w przestrzeniach metrycznych . . . 127
6.1.1 Eksperymenty numeryczne . . . 127
6.1.2 Definicja granicy funkcji w przestrzeniach metrycznych . . . 127
6.1.3 Jednoznaczność granicy . . . 128
6.1.4 Granica złożenia funkcji . . . 130
6.2 (**)Przypadki szczególne definicji granicy . . . 130
6.2.1 Kryterium na granicę w bazach . . . 131
6.2.2 Granica ciągu . . . 131
6.2.3 Granica funkcji w przestrzeni metrycznej . . . 131
6.2.4 Granica nieskończona w $\bar{\mathbb{R}}$ . . . 131
6.2.5 Granica w nieskończoności . . . 132
6.3 Granice jednostronne . . . 132
6.4 Punkty graniczne . . . 132
6.4.1 Ciągowa charakterystyka granicy . . . 132

6.4.2 Zbiór punktów granicznych . . . 133
6.4.3 Ciągowa charakterystyka granicy ciągu . . . 134
6.5 Granica dolna i górna . . . 134
6.5.1 Granica dolna i górna . . . 134
6.5.2 Własności granicy dolnej i górnej . . . 135
6.6 Granica, a struktura algebraiczna . . . 135
6.6.1 Sumy, różnica, iloczyn i iloraz funkcji . . . 135
6.6.2 Granica sumy, różnicy i iloczynu . . . 136
6.6.3 Granica ilorazu . . . 136
6.7 Granica, a struktura porządkowa . . . 137
6.7.1 Zachowywanie nierówności w granicy . . . 137
6.7.2 Twierdzenie o trzech funkcjach . . . 137
6.7.3 Granica funkcji monotonicznej . . . 138
6.8 Granice w produkcie kartezjańskim . . . 138

7 Ciągłość funkcji 141
7.1 Koncepcja ciągłości . . . 141
7.1.1 Definicja ciągłości . . . 141
7.1.2 Kryteria ciągłości . . . 142
7.1.3 Ciągłość złożenia . . . 142
7.1.4 Ciągłość operacji arytmetycznych . . . 142
7.1.5 Twierdzenie o lokalnym zachowywaniu znaku . . . 142
7.1.6 (*)Nieciągłości funkcji $f : \mathbb{R} \to \mathbb{R}$ . . . 143
7.1.7 Ciągłość bijekcji monotonicznej . . . 144
7.2 Ciągłość, a zwartość . . . 144
7.2.1 Zwartość podprzestrzeni $\mathbb{R}^d$. . . . 144
7.2.2 (*)Intuicja zwartości i ogólna definicja zwartości. . . . 145
7.2.3 Przestrzenie zwarte . . . 147
7.2.4 Obraz zbioru zwartego przez odwzorowanie ciągłe. . . . 147
7.2.5 Funkcja ciągła na przedziale zwartym osiąga swoje kresy . . . 148
7.2.6 Funkcja odwrotna do funkcji ciągłej . . . 148
7.3 Jednostajna ciągłość . . . 149
7.3.1 Funkcje Lipschitza . . . 149
7.4 Ciągłość, a spójność . . . 150
7.4.1 Spójne podzbiory $\bar{\mathbb{R}}$. . . . 150
7.4.2 (*)Przestrzenie spójne . . . 150
7.4.3 Obraz zbioru spójnego przez odwzorowanie ciągłe. . . . 151
7.4.4 Własność Darboux. . . . 151

8 Szeregi 153
8.1 Eksperymenty numeryczne . . . 153
8.1.1 Przybliżanie liczby π polami wielokątów wpisanych w koło. . . . 153
8.1.2 Warunek konieczny zbieżności szeregu . . . 153
8.1.3 Rozbieżność do nieskończoności . . . 154
8.2 Własności podstawowe . . . 157
8.2.1 Pojęcie szeregu . . . 157
8.2.2 Szereg geometryczny . . . 158
8.3 Kryteria zbieżności . . . 159
8.3.1 Szereg $\sum_{n=1}^{\infty} \frac{1}{n^p}$. . . . 159

8.3.2 Kryterium Leibniza . . . 160
8.3.3 Kryterium Cauchy'ego i kryterium d'Alemberta . . . 160
8.3.4 Szeregi potęgowe . . . 161
8.4 Operacje na szeregach . . . 162
8.4.1 Suma i iloczyn szeregów . . . 162
8.4.2 (**)Permutacja szeregu i zbieżność bezwarunkowa . . . 163
8.4.3 (**)Twierdzenie Riemanna . . . 164
8.4.4 Szeregi w programie Mathematica . . . 164

9 Funkcje wykładnicza, logarytmiczna i trygonometryczne 165
9.1 Funkcja wykładnicza i logarytmiczna . . . 165
9.1.1 Zespolona funkcja wykładnicza . . . 165
9.1.2 Funkcja wykładnicza e^x . . . 166
9.1.3 Logarytm naturalny . . . 167
9.1.4 Funkcja wykładnicza a^x . . . 168
9.2 Funkcje trygonometryczne . . . 168
9.2.1 Nawijanie prostej na okrąg . . . 168
9.2.2 Funkcje sin i cos . . . 169
9.2.3 Liczba π . . . 172
9.2.4 Okresowość . . . 172
9.2.5 Wzory trygonometryczne . . . 173

10 Przestrzenie wektorowe 175
10.1 Definicja przestrzeni wektorowej . . . 175
10.1.1 Przestrzeń wektorowa . . . 175
10.2 Przestrzenie wektorowe z iloczynem skalarnym . . . 176
10.2.1 Iloczyn skalarny . . . 176
10.2.2 Nierówność Cauchy-Schwartza . . . 177
10.3 Przestrzenie wektorowe unormowane . . . 177
10.3.1 Norma . . . 178
10.3.2 Norma euklidesowa . . . 178
10.3.3 Przestrzenie Hilberta i Banacha . . . 178

11 Pochodne i różniczki. 181
11.1 Pochodna funkcji . . . 181
11.1.1 Definicja pochodnej . . . 181
11.1.2 Pochodna funkcji stałej i identycznościowej. . . . 182
11.1.3 Pochodna lewo- i prawostronna. . . . 183
11.1.4 Ciągłość funkcji różniczkowalnej . . . 184
11.2 Różniczka funkcji . . . 185
11.2.1 Aproksymacja liniowa . . . 185
11.2.2 Relacja "o małe" . . . 187
11.2.3 Płaskość w zerze . . . 188
11.2.4 Różniczkowalność . . . 188
11.3 Różniczkowanie sumy, różnicy, iloczynu i ilorazu . . . 190
11.4 Różniczkowanie funkcji złożonej . . . 192
11.5 Różniczkowanie funkcji wykładniczej i funkcji trygonometrycznych . . . 194
11.6 Nieciągłości pochodnej i funkcje klasy C^1. . . . 196
11.7 Metoda Newtona . . . 197

11.8 Ekstrema lokalne, a różniczkowalność. . . . 200
11.8.1 Maksima i minima lokalne . . . 200
11.8.2 Twierdzenie Rolle'a. . . . 200
11.9 Twierdzenie o wartości średniej . . . 200
11.9.1 Uogólnione twierdzenie o wartości średniej . . . 200
11.9.2 Twierdzenie Lagrange'a o wartości średniej . . . 201
11.9.3 Twierdzenie o przyrostach skończonych . . . 202
11.10 Zastosowania twierdzenia o wartości średniej . . . 202
11.10.1 Pochodna, a monotoniczność . . . 202
11.10.2 Twierdzenie o wartościach pośrednich dla pochodnej . . . 203
11.10.3 Reguła de l'Hospitala . . . 203
11.11 Różniczkowanie funkcji odwrotnej . . . 206
11.11.1 Wzór na pochodną funkcji odwrotnej . . . 206
11.11.2 Pochodna funkcji logarytmicznej . . . 207
11.11.3 Funkcje kołowe (odwrotne do trygonometrycznych) . . . 207
11.11.4 Pochodna funkcji potęgowej . . . 208
11.12 Pochodne wyższych rzędów . . . 208
11.12.1 n-ta pochodna . . . 208
11.12.2 Wielomian Taylora . . . 210
11.12.3 n-płaskość w zerze. . . . 211
11.13 Wzory Taylora i szereg Taylora. . . . 212
11.13.1 Wzór Taylora z resztą Peano. . . . 212
11.13.2 Wzór Taylora z resztą Lagrange'a . . . 213
11.13.3 Punkty przegięcia . . . 215
11.13.4 Szereg Taylora . . . 215

12 Całka oznaczona **219**
12.1 Całka Riemanna funkcji jednej zmiennej . . . 219
12.1.1 Pole obszaru. . . . 219
12.1.2 Przedziały w $\mathbb{R}$ i ich podziały. . . . 219
12.1.3 Całkowalność w sensie Riemanna . . . 221
12.1.4 Epsilonowe kryterium całkowalności . . . 222
12.2 Własności całki Riemanna . . . 224
12.2.1 Monotoniczność całki . . . 224
12.2.2 Przestrzeń wektorowa funkcji całkowalnych w sensie Riemanna. . . . 224
12.2.3 Miarowe kryterium całkowalności. . . . 225
12.2.4 Całkowalność złożenia i iloczynu . . . 225
12.2.5 Całka Riemanna na podprzedziale . . . 225
12.2.6 Całka oznaczona . . . 226
12.2.7 Twierdzenie Riemanna . . . 226

13 Funkcja pierwotna i całka nieoznaczona **231**
13.1 Podstawowe twierdzenie rachunku różniczkowego i całkowego . . . 231
13.1.1 Pochodna całki . . . 231
13.1.2 Funkcja pierwotna i całka nieoznaczona. . . . 232
13.1.3 Podstawowe twierdzenie rachunku różniczkowego i całkowego . . . 232
13.2 Podstawowe techniki całkowania. . . . 233
13.2.1 Własności całki nieoznaczonej. . . . 233
13.2.2 Całkowanie przez części . . . 234

13.2.3 Całkowanie przez podstawienie . . . 236
13.2.4 Twierdzenia o wartości średniej dla całek . . . 237
13.3 Całkowanie funkcji wektorowych . . . 238
13.4 Zastosowania geometryczne całek . . . 239
13.4.1 Długość krzywej . . . 239
13.4.2 Bryła obrotowa . . . 240
13.4.3 Obętość bryły obrotowej . . . 241
13.4.4 Pole powierzchni bocznej bryły obrotowej . . . 241

14 Ciągi i szeregi funkcyjne **243**
14.1 Zbieżność punktowa i jednostajna . . . 243
14.1.1 Zbieżność punktowa . . . 243
14.1.2 Zbieżność jednostajna . . . 244
14.1.3 Kryterium Cauchy'ego zbieżności jednostajnej . . . 245
14.2 Zbieżność jednostajna, a ciągłość, różniczka i całka . . . 246
14.2.1 Granica, a zbieżność jednostajna . . . 246
14.2.2 Całkowanie i różniczkowanie, a zbieżność jednostajna . . . 247
14.3 Aproksymacja funkcji ciągłych . . . 248
14.4 Szeregi potęgowe . . . 248
14.4.1 Twierdzenie o szeregu Taylora . . . 249
14.4.2 Funkcje analityczne . . . 249
14.5 Szeregi Fouriera . . . 251
14.5.1 Wielomiany trygonometryczne . . . 251
14.5.2 Zespolony iloczyn skalarny . . . 252
14.5.3 Przestrzeń wektorowa z zespolonym iloczynem skalarnym . . . 252
14.5.4 Układ ortonormalny . . . 253
14.5.5 Szereg Fouriera . . . 255
14.5.6 Rzut ortonormalny . . . 257
14.5.7 Jądro Dirichleta . . . 258
14.5.8 Zbieżność punktowa szeregu Fouriera . . . 259

Wstęp

Niniejsze notatki do kursu Analiza Matematyczna I przeznaczonego dla studentów kierunku Matematyka Komputerowa stanowią roboczą wersję i w trakcie kursu mogą podlegać aktualizacji. Punktem wyjścia tych notatek jest rękopis moich materiałów do wykładu z analizy matematycznej prowadzonego w latach 1984-1990 dla studentów informatyki UJ. W owym czasie był to bardzo obszerny kurs obejmujący 240h wykładów i 240h ćwiczeń. Wykładało i studiowało się wtedy zupełnie inaczej, bo wykładowca miał do dyspozycji tylko tablicę i kredę, a studenci to co zdołali w trakcie wykładu zanotować.

Dziś możliwości są zupełnie inne, stąd i te notatki są inne, choć w dużym stopniu wykorzystują jeszcze mój dawny rękopis. Już wtedy starałem się o to, by część odnosząca się do intuicji była obszerna i równocześnie wyraźnie oddzielona od części formalnej. Wydaje się, że takie podejście jest szczególnie ważne dla studentów specjalności Matematyka Komputerowa, dlatego planuję je kontynuować również w ramach obecnego wykładu. Część materiału poświęcona intuicjom ma charakter całkowicie nieformalny. Dla wyraźnego odróżnienia od części formalnej jest złożona brązową, pochyloną czcionką. W szczególności cały pierwszy rozdział jest nieformalny. Nieformalne mówienie o matematyce rodzi często więcej pytań niż daje odpowiedzi. Ale takie właśnie jest jej zadanie: sprowokować do myślenia. Dociekliwy czytelnik, który czyta najpierw nieformalny tekst matematyczny wyrabia sobie pewne idee, które wymagają dopracowania. W moim przekonaniu takie podejście powoduje, że studiowanie tekstu formalnego, który to dopracowanie dostarcza, staje się w efekcie bardziej naturalne i znacznie łatwiejsze. Najważniejsze pytania jakie rodzi tekst nieformalny zostały w tekście wyróżnione poprzez umieszczenie w ramce o różowym tle. Pytań jest oczywiście dużo więcej niż tych, które w tekście zostały jawnie postawione.

Różne kolory tła tekstu oznaczają: blado pomarańczowy - twierdzenia, blado zielony - definicje, różowy - problemy, blado niebieski (C++) i blado fioletowy (Mathematica) - listingi programów komputerowych, żółłty - ćwiczenia tablicowe, niebieski - ćwiczenia kokmputerowe.

W notatkach odręcznych

- znak i oznacza początek dowodu indukcyjnego,
- znak s oznacza początek dowodu nie wprost,
- znak x oznacza sprzeczność,
- znak p oznacza początek dowodu przez przypadki,
- znak □ oznacza koniec fragmentu dowodu,
- znak ■ oznacza koniec dowodu,
- wytłuszczony wykrzyknik zastępuje słowo "niech",
- dwie pionowe kreski po lewej stronie tekstu oznaczają materiał nieformalny.

Aspekty obliczeniowe analizy matematycznej, ilustrowane programami komputerowymi w języku C++ oraz instrukcjami programu Mathematica złożone są czcionką w kolorze niebieskim.

Znajomość rozdziałów, podrozdziałów i dowodów oznaczonych (*) obowiązuje jedynie studentów mierzących w ocenę bdb. Rozdziały i podrozdziały oznaczone (**) są nieobowiązkowe.

Życzę wszystkim uczestnikom kursu udanego studiowania matematyki komputerowej na Wydziale Matematyki i Informatyki UJ.

Rozdział 1

Wprowadzenie do analizy matematycznej

Nim zabierzemy się do studiowania analizy matematycznej w szczegółach, warto sobie zadać pytanie czym analiza matematyczna się zajmuje. Nie da się na nie dać dobrej odpowiedzi bez przyjrzenia się jak koncepcje, które ukształtowały współczesną analizę matematyczną stopniowo kształtowały się na przestrzeniach dziejów. Niniejszy rozdział wprowadzający stanowi krótki przegląd podstawowych problemów analizy matematycznej oraz ich umiejscowienia w historii matematyki.

1.1 Rozwój pojęcia liczby

1.1.1 Liczby naturalne

Umiejętność posługiwania się liczbami rozwijała się u człowieka stopniowo. Pierwotny zbiór liczebników ograniczał się do trzech: jeden, dwa i wiele. Wraz z rozwojem ludzkiej cywilizacji pojawiła się konieczność dokładnego rozeznania w liczebności oddziału wojowników, stada bydła, czy zestawu strzał. Tak zrodziła się umiejętność porównywania liczności zbiorów, która doprowadziła do pojawienia się abstrakcyjnego pojęcia liczby naturalnej. Zapewne niedługo potem człowiek nauczył się te liczby dodawać, odejmować i mnożyć. Operacje te wymagały opanowania stosownych algorytmów postępowania, a ich efektywność silnie zależała od stosowanego sposobu zapisu liczb. Używany przez nas dziś dziesiątkowy system pozycyjny wywodzi się z Indii, gdzie pojawił się około VIII w.

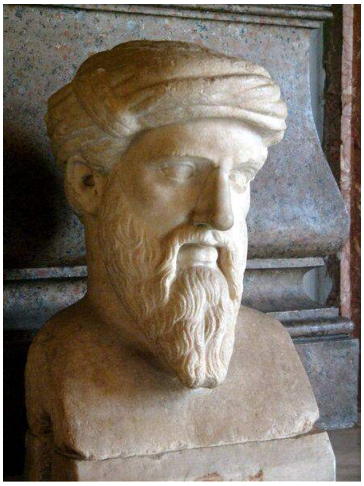
1.1.2 Ułamki

Pojawienie się ułamków wiąże się z rozwojem handlu, który wymagał doskonalenia umiejętności mierzenia. Do mierzenia używano miary. Na przykład w przypadku płynów i materiałów sypkich mogła to być gliniana amfora o ustalonej objętości. Początkowo zaniedbywano powstające przy odmierzaniu części ułamkowe, ale gdy z biegiem czasu rosła potrzeba dokładności, wprowadzano miary pomocnicze będące częścią wyjściowej miary, na przykład połową, jedną trzecią czy jedną czwartą. Co charakterystyczne, były to tzw. ułamki proste, czyli ułamki o liczniku jeden. Gdy handel zaczął przekraczać granice pojedynczych plemion pojawiła się konieczność wprowadzania przelicznika dla różnych systemów miar obowiązujących w różnych rejonach. To zapewne przyczyniło się do powstania abstrakcyjnej postaci ułamka.

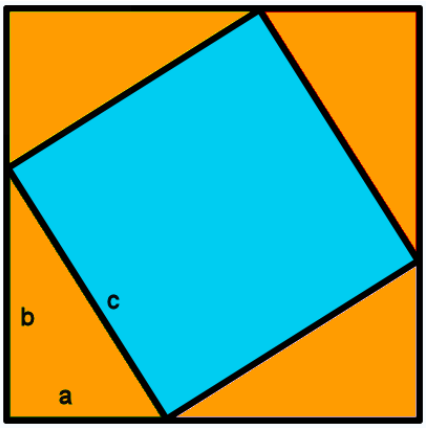
1.1.3 Liczby ujemne, całkowite i wymierne

Dwie liczby naturalne i dwa ułamki można zawsze do siebie dodać i w wyniku powstaje inna liczba naturalna lub inny ułamek. Jednakże od liczby a nie można odjąć liczby b gdy $b \geqslant a$ bez rozszerzenia zbioru liczb o liczby ujemne i zero. Jak wiemy z każdą liczbą a powiązać można liczbę do niej przeciwną, oznaczaną $-a$, tak że ich suma daje zero. Zbiór liczb naturalnych rozszerzony o zero i liczby przeciwne do liczb naturalnych nazywamy zbiorem liczb całkowitych, a analogicznie rozszerzony zbiór ułamków nazywamy zbiorem liczb wymiernych. Zwyczajowo stosujemy oznaczenia $\mathbb{N}$ dla liczb naturalnych, $\mathbb{Z}$ dla liczb całkowitych oraz $\mathbb{Q}$ dla liczb wymiernych. Natomiast zbiór liczb wymiernych dodatnich, czyli po prostu zbiór klasycznych ułamków oznaczamy $\mathbb{Q}^+$.

Liczby ujemne w pewnym stopniu były znane już w starożytnych Chinach i Indiach, gdzie interpretowane były w kategoriach długu. Jednak nie były specjalnie doceniane i na codzień radzono sobie bez nich. Matematyka europejska zaakceptowała liczby ujemne dopiero w XVII w, ale nawet dziś w życiu codziennym jesteśmy w stanie radzić sobie bez liczb ujemnych. Największy pożytek z liczb ujemnych jest przy ujednolicaniu algorytmów, gdyż dzięki nim nie musimy rozważać różnych przypadków, by zagwarantować, że wszystkie rozważane w trakcie pracy algorytmu liczby są dodatnie.



Rysunek 1.1: Pitagoras z Samos, VI w. przed Chrystusem. Źródło: Wikipedia



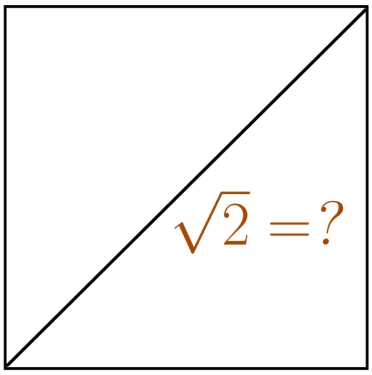
Rysunek 1.2: $c^2 = (a+b)^2 - 4\frac{ab}{2} = a^2 + b^2$

1.1.4 Liczby niewymierne

Odkrycie, że ułamki nie wystarczają do mierzenia dowolnych wielkości przypisuje się szkole pitagorejskiej, wywodzącej się od żyjącego w VI w. przed Chrystusem w Grecji Pitagorasa z Samos (rys.1.1). Pitagorejczycy święcie wierzyli, że ułamki wystarczają do zmierzenia dowolnych długości. Dlatego dokonane przez nich odkrycie było dla nich ciosem, tym większym, że było konsekwnecją słynnego twierdzenia udowodnionego przez Pitagorasa (rys.1.2).

Twierdzenie 1.1.1 *(Pitagoras z Samos) W trójkącie prostokątnym suma kwadratów przyprostokątnych jest równa kwadratowi przeciwprostokątnej.*

Z twierdzenia Pitagorasa wynika, że przekątna kwadratu o boku długości jeden jest liczbą, która podniesiona do kwadratu wynosi dwa, czyli tak zwanym pierwiastkiem z dwóch (rys. 1.3). Pitagorejczycy pierwsi zauważyli, że żaden ułamek podniesiony do kwadratu nie może dać liczby dwa. Było to szokujące spostrzeżenie, z którym niełatwo było się pogodzić. Tak pojawiły się liczby niewymierne, a wraz nimi problem ich przybliżania ułamkami.



Rysunek 1.3: Przekątna kwadratu to $\sqrt{2}$. Ale ile to jest $\sqrt{2}$?

Ćwiczenie 1.1.2 • Rozkładając na czynniki pierwsze licznik i mianownik ułamka spróbuj uzasadnić, dlaczego $\sqrt{2}$ nie może być ułamkiem.

• Czy podobne rozumowanie może być przeprowadzone dla $\sqrt[3]{2}$? dla innych pierwiastków?

Problem 1.1.3 Czy wiemy czym naprawdę jest liczba $\sqrt{2}$? Czy ona rzeczywiście ma sens? Jakich liczb jeszcze brakuje wśród ułamków, by wszystkie długości dało się wymierzyć? W jaki sposób je zdefiniować?

1.1.5 Liczby rzeczywiste

Precyzyjna definicja liczby niewymiernej jest skomplikowana. Stwierdzenie, że liczby niewymierne, to wszystkie liczby, które nie są wymierne jest dobrą definicją tylko wtedy gdy wiemy co oznacza zwrot "wszystkie liczby". Intuicyjnie chodzi nam o taki zbiór liczb, który pozwala mierzyć długości tak odcinków jak i krzywych. Zbiór takich liczb został precyzyjnie zdefiniowany dopiero w XIX wieku. Określa się go mianem zbioru liczb rzeczywistych i oznacza symbolem $\mathbb{R}$. Natomiast zbiór liczb rzeczywistych dodatnich oznacza się symbolem $\mathbb{R}^+$.

1.1.6 Liczby urojone i zespolone

Określenie liczby rzeczywiste *zrodziło się w kontraście do tzw. liczb urojonych. Liczba urojona to liczba, której kwadrat jest liczbą ujemną. Kwadrat liczby rzeczywistej nie może być ujemny, tak więc liczb urojona nie jest liczbą rzeczywistą. Sensowność takich liczb długo budziła wątpliwości, co wyjaśnia ich nazwę.*

Liczby zespolone, będące sumami liczb rzeczywistych i urojonych, mają wiele cech podobnych do liczb rzeczywistych, a także wiele zalet, których liczby rzeczywiste nie mają. Dzięki temu pozwalają rozwiązywać pewne problemy "na skróty". Jednym z pierwszych pożytków z liczb zespolonych było podanie w XVI wieku algorytmów wyznaczania miejsc zerowych równania trzeciego i czwartego stopnia.

1.1.7 Liczby w komputerze

Rachunki liczbowe na papierze wykonywane są zazwyczaj na ułamkach zwykłych, a częściej na ułamkach dziesiętnych.

Wraz ze zbudowaniem w połowie XX w. pierwszych komputerów stworzono też koncepcję przechowywania i przetwarzania liczb w komputerze. Było to konieczne, bo wszystko co przetwarza współczesny komputer na poziomie procesora sprowadza się do wykonywania obliczeń na liczbach. Typowy użytkownik komputera ogląda filmy, słucha muzyki, przegląda strony internetowe albo po prostu gra na komputerze i nie uświadamia sobie, że wszystko to sprowadza się do przetwarzania liczb. Zazwyczaj są to liczby całkowite, ale wiele pozamatematycznych zastosowań wymaga też obliczeń na liczbach wymiernych.

Komputer jest oczywiście równie niezastąpiony jeśli naszym celem jest przeprowadzenie mniej lub bardziej skomplikowanych obliczeń matematycznych, inżynierskich bądź naukowych. Jeśli obliczenia są bardzo specjalistyczne, jest ich szczególnie dużo, a nam zależy na efektywności, piszemy własny lub zamawiamy stosowny program w kompilowanym języku wysokiego poziomu. Jeśli obliczenia są typowe, a ich ilość jest umiarkowana, sięgamy po specjalistyczne oprogramowanie takie jak Mathematica, Maple, Matlab czy Scilab.

W ramach niniejszego kursu aspekty numeryczne analizy matematycznej demonstrować będziemy programami napisanymi dla środowiska Mathematica, a także programami napisanymi w języku C++.

Mathematica to bogaty pakiet oprogramowania bardzo ułatwiający wszelkie obliczenia matematyczne. Umożliwia on tak prostą pracę interaktywną jak i pisanie własnego oprogramowania. Na przykład, aby pomnożyć przez siebie liczby 123 i 456 wystarczy napisać

```
123*456
```

i przycisnąć rówcznocześnie klawisze 'Shift' i 'Enter'. Przyciśnięcie 'Shift' wraz 'Enter' jest sygnałem, że oczekujemy przeprowadzenia obliczeń. Natychmiast otrzymujemy wynik

$$56088$$

W programie Mathematica operacje arytmetyczne dodawania, odejmowania, mnożenia, dzielenia i potęgowania oznaczamy odpowiednio znakami `+`, `-`, `*`, `/`, `^`. *Tak więc pisząc*

```
3^4
```

otrzymujemy

$$81$$

W podobny sposób przeprowadzamy obliczenia na ułamkach dziesiętnych. Na przykład

```
3.14 + 2.72
```

daje wynik 5.86. Trzeba jedynie pamiętać, że zgodnie z anglosaskim zwyczajem używamy kropki dziesiętnej, a nie przecinka. Obliczenia na liczbach dziesiętnych związane są niestety z błędami zaokrągleń. Na przykład

```
1.0/3.0
```

daje 0.333333 co jest wynikiem przybliżonym. Błędu zaokrągleń można uniknąć przeprowadzając obliczenia na liczbach wymiernych. Jeśli napiszemy

```
1/3
```

Mathematica zwróci ułamek

$$\frac{1}{3},$$

a pisząc

```
7/3+3/11
```

otrzymamy jako wynik ułamek

$$\frac{86}{33}.$$

co , w przeciwieństwie do obliczeń na liczbach dziesiętnych, jest wynikiem dokładnym. Jeśli potrzebujemy jednak przybliżenia dziesiętnego, zawsze możemy je otrzymać. Pisząc

```
N[1/23]
```

otrzymujemy przybliżenie ułamka $\frac{1}{23}$

$$0.0434783.$$

1.2 Ciągi

1.2.1 Przybliżanie pierwiastka

Choć wśród ułamków brak jest takiego, który podniesiony do kwadratu daje liczbę dwa, to rozważając wszystkie ułamki o mianowniku m znajdziemy taki, postaci $\frac{i}{m}$, że

$$\left(\frac{i}{m}\right)^2 < 2 < \left(\frac{i+1}{m}\right)^2.$$

Dla danego m licznik taki jest dokładnie jeden. Sam pierwiastek z dwóch, czymkolwiek jest, leży gdzieś pomiędzy $\frac{i}{m}$, a $\frac{i+1}{m}$, zatem $\frac{i}{m}$ jest przybliżeniem $\sqrt{2}$ z dokładnością co najmniej $\frac{1}{m}$. Na przykład, dla $m = 5$ mamy

$$\left(\frac{7}{5}\right)^2 = \frac{49}{25} < 2 < \frac{64}{25} = \left(\frac{8}{5}\right)^2$$

zatem $\frac{7}{5}$ jest przybliżeniem $\sqrt{2}$ z dokładnoscią większą niż $\frac{1}{5}$. Aby dostać lepszą dokładność wystarczy wziąć odpowiednio duży mianownik m.

1.2.2 Ciągi przybliżeń

W praktyce wygodnie ograniczyć się do mianowników postaci $m = 10^n$. Niestety ręczne poszukiwanie stosownego licznika jest bardzo żmudne, a dla dużych n praktycznie niewykonalne. Potrzebujemy algorytmu, który dla zadanego mianownika 10^n pozwoli nam wyznaczyć stosowny licznik i_n oraz komputera, który sprawnie wykona obliczenia w tym algorytmie. Łatwo zauważyć, że

$$i_n = \max\{\, i \mid i * i < 2 * 100^n \,\}.$$

Sugeruje to algorytm, który przy zadanym n sprawdza nierówność $i * i < 2 * 100^n$ dla kolejnych liczb naturalnych i i zwraca i_n jako ostatnie i, dla którego nierówność ta zachodzi. Wtedy $x_n := \frac{i_n}{10^n}$ jest przybliżeniem $\sqrt{2}$, które, jak łatwo sprawdzić, spełnia

$$|x_n - \sqrt{2}| \leqslant \frac{1}{10^n}.$$

Kolejne liczby x_n tworzą tzw. ciąg liczbowy, w tym przypadku tzw. ciąg przybliżeń. Ciąg przybliżeń, to ciąg skonstruowany w celu przybliżania pewnej liczby niewymiernej liczbami wymiernymi. Ciągi przybliżeń to nie jedyne ciągi rozważane w matematyce, choć ich rola praktyczna jest bardzo istotna. Badanie ciągów to jeden z głównych obiektów zainteresowań analizy matematycznej.

Wprowadźmy w programie Mathematica rozdzielony średnikami ciąg instrucji

```
n = 3; i = 1; While[i^2 < 2*100^n, i = i + 1]; {(i-1)/10^n, i/10^n}
```

Pojawiają się tu zmienne `i` *oraz* `n`. *Instrukcja* `While` *to instrukcja pętli. Jak długo spełniony jest warunek* `i^2 < 2*100^n` *wykonywana jest instrukcja* `i = i + 1`. *Ponieważ przed wykonaniem instrukcji* `While` *ustawiliśmy zmienną* `n` *na 3, a* `i` *na 1, pętlę wykonujemy dla kolejnych wartości* `i` $= 1, 2, \ldots$ *aż zajdzie nierówność* `i^2` $\geqslant 2000000$. *Mathematica wyprowadza jako wynik ostatnie obliczone wyrażenie, którym jest para ułamków* `{(i-1)/10^n,i/10^n}`. *W rozważanym przypadku otrzymamy*

$$\{\frac{707}{500}, \frac{283}{200}\}.$$

jako oszacowanie od dołu i od góry $\sqrt{2}$ z dokładnością do $\frac{1}{1000}$.

Podstawiając w powyższym listingu za `n` *kolejne liczby naturalne i wykonując obliczenia otrzymamy przybliżenia $\sqrt{2}$ od dołu i od góry o coraz lepszej*

dokładności:

$$\left\{\frac{7}{5}, \frac{3}{2}\right\}$$
$$\left\{\frac{141}{100}, \frac{71}{50}\right\}$$
$$\left\{\frac{707}{500}, \frac{283}{200}\right\}$$
$$\left\{\frac{7071}{5000}, \frac{14143}{10000}\right\}$$
$$\left\{\frac{141421}{100000}, \frac{70711}{50000}\right\}$$
$$\left\{\frac{1414213}{1000000}, \frac{707107}{500000}\right\}$$
$$\ldots$$

Obliczenia te, przynajmniej teoretycznie, możemy kontynuować w nieskończoność. Otrzymujemy w ten sposób ciąg przybliżeń od dołu i ciągu przybliżeń od góry liczby $\sqrt{2}$.

W tym miejscu trzeba zrobić uwagę na temat stosowanej w programie Mathematica notacji. Mathematica używa nawiasów wąsiastych $\{,\}$ do oznaczania list. Zwyczajowo w matematyce nawiasów wąsiastych używa się do oznaczania zbiorów. Różnica ta ma znaczenie, bo przy wypisywaniu elementów listy kolejność jest istotna, a przy wypisywaniu elementów zbioru nie jest. Innymi słowy, dwie listy, które różnią się tylko kolejnością elementów są różne, a dwa zbiory, które różnią się tylko kolejnością w jakiej wymieniono ich elementy są traktowane jako identyczne. Korzystając z programu Mathematica musimu o tej różnicy pamiętać.

Zaproponowana tu metoda liczenia pierwiastka należy do najprostszych, ale oczywiście jest daleka od doskonałości. Mathematica radzi sobie doskonale nie tylko z liczeniem pierwiastków, ale również innych funkcji elementarnych, w szczególności funkcji trygonometrycznych i funkcji logarytmicznej. Trzeba jednak pamiętać, że jeśli jako argument podamy liczbę dokładną, na przykład naturalną lub ułamek, a wynik nie jest liczbą naturalną bądź ułamkiem, aby zachować dokładność Mathematica nie przeprowadzi obliczeń. Na przykład pisząc

```
Sqrt[2]
```

otrzymamy jako wynik $\sqrt{2}$. Ale pisząc

```
Sqrt[2.]
```

otrzymamy 1.41421.

Do zagadnienia obliczania pierwiastka będziemy jeszcze wracać w dalszych rozdziałach.

Oczywiście ciągi nie muszą być zadawane w tak skomplikowany sposób jak nasz ciąg przybliżeń. Najprościej zadać ciąg podając wzór na jego n-ty wyraz. Na przykład ciąg dany wzorem

$$x_n = \frac{1}{n} \tag{1.1}$$

to tak zwany ciąg harmoniczny.

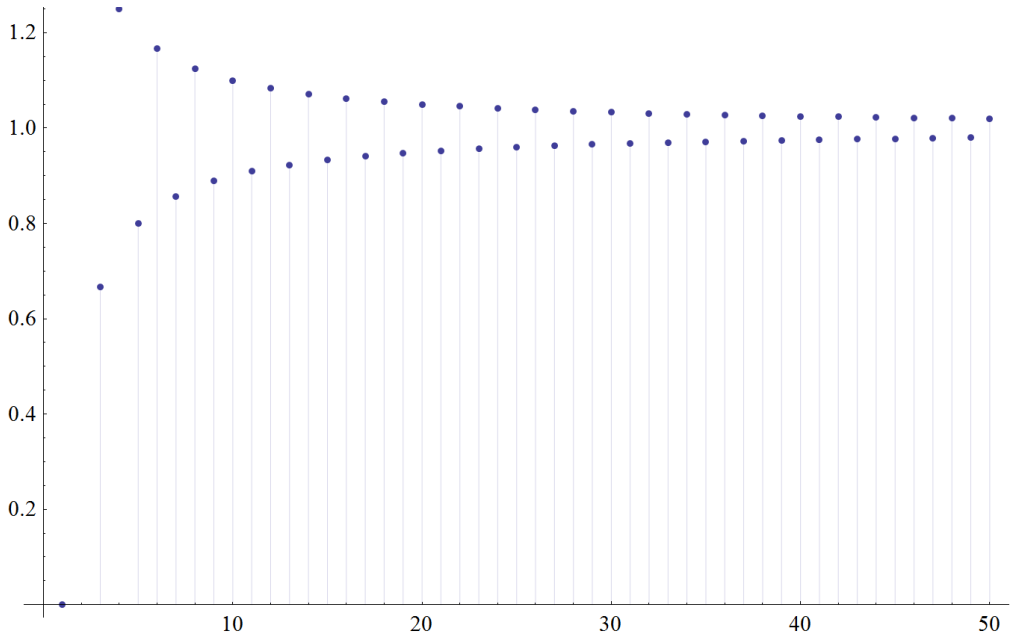
Problem 1.2.1 Czy ciąg harmoniczny przybliża jakąś liczbę? Jeśli tak to jaką?

Gdy ciąg określony jest wzorem, na przykład wzorem (1.1), wystarczy napisać

```
Table[1/n, {n, 1., 10.}]
```

a otrzymamy przybliżenia początkowych wyrazów tego ciągu

$$\{1., 0.5, 0.333333, 0.25, 0.2, 0.166667, 0.142857, 0.125, 0.111111, 0.1\}$$



Rysunek 1.4: Wykres początkowych wyrazów ciągu $1+\frac{(-1)^n}{n}$.

Można też łatwo uzyskać wizualizację (wykres) ciągu. Na przykład wykres początkowych wyrazów ciągu danego wzorem

$$u_n = 1 + \frac{(-1)^n}{n}$$

uzyskamy pisząc

```
DiscretePlot[1 + (-1)^n/n, {n, 1, 50}]
```

Uzyskamy w ten sposób wykres przedstawiony na rys. 1.4.

1.2.3 Granice

Obrazowo mówimy, że ciąg ten osiąga liczbę $\sqrt{2}$ w granicy, albo krócej, że zmierza do $\sqrt{2}$. Właśnie w tym celu go skonstruowaliśmy. Ale czy możemy mieć pewność, że nasz ciąg dla dużych wartości n nie zmieni zdania i na przykład zamiast zmierzać do $\sqrt{2}$ zacznie uciekać do nieskończoności? Jesteśmy w stanie policzyć tylko skończoną ilość wyrazów tego ciągu. Zawsze zostanie nieskończenie wiele wyrazów, których nie policzyliśmy. Czy możemy mieć pewność, że nasz ciąg przybliżeń nigdy nas nie zawiedzie?

Problem 1.2.2 Co to znaczy, że ciąg osiąga jakąś liczbę w granicy (zmierza do granicy)? Jakie ciągi mają granicę? Czy, kiedy i jak granicę daję się policzyć?

Odpowiedź na te pytania to jedno z fundamantalnych zadań analizy matematycznej. Pojęciem granicy w takim bądź innym sensie posługiwano się już od czasów starożytnych, jednak aż do wieku XIX matematycy nie umieli sobie poradzić z postawieniem precyzyjnej definicji. Dopiero czeski matematyk Bernard Bolzano (rys.1.5) oraz niezależnie francuski matematyk, Augustin-Louis Cauchy (rys.1.5) pierwsi postawili precyzyjną definicję. Ponieważ prace Bolzano długo pozostawały mało znane, przyjęło się mówić o definicji Cauchy'ego.

Według definicji Cauchy'ego ciąg liczb x_n jest zbieżny do granicy g jeżeli

$$\forall \epsilon > 0 \; \exists N \in \mathbb{N} \; \forall n > N \quad |x_n - g| < \epsilon.$$





Rysunek 1.5: Augustin Louis Cauchy (1789 - 1857) i Bernard Bolzano (1781 - 1844) Źródło: Wikipedia

Piszemy wtedy

$$\lim_{n\to\infty} x_n = g.$$

Ćwiczenie komputerowe 1.2.3 Dla każdego z poniższych ciągów postaw hipotezę, czy ma on granicę. Jeśli uważasz, że ma granicę, spróbuj odgadnąć jaką. Wcześniej przeprowadź eksperymenty wykorzystując instrukcję Table w Mathematica.

1. $\frac{n}{n+1}$
2. $1+(-1)^n$
3. $1+\frac{(-1)^n}{n}$
4. (*) $\cos n\pi$
5. (*) $\sin n$
6. (*) $\frac{\sin n}{n}$

1.2.4 Obliczanie granic ciągu.

W praktyce niewiele granic liczy się wprost z definicji, choć najprostsze granice właśnie tak się wyznacza. Znając już granice jakichś ciągów liczymy granice innych ciągów korzystając z różnych twierdzeń. Do najważniejszych należą

Twierdzenie 1.2.4 *Jeśli ciągi a_n, b_n są zbieżne odpowiednio do a i b to ich suma, różnia, iloczyn oraz iloraz są zbieżne odpowiednio do $a+b$, $a-b$, ab, $\frac{a}{b}$, przy czym w przypadku ilorazu trzeba dodatkowo założyć, że $b\neq 0$.*



Rysunek 1.6: Archimedes z Syrakuz (ok. 287–212 pne.) (medal Fieldsa). Źródło: Wikipedia

Twierdzenie 1.2.5 *Jeśli ciągi a_n, b_n, c_n spełniają nierówność*

$$a_n \leqslant b_n \leqslant c_n$$

i ciągi a_n oraz c_n są zbieżne i to do tej samej granicy, to ciąg b_n też jest zbieżny i też do tej granicy.

Ćwiczenie 1.2.6
- Posługując się definicją granicy oblicz granice ciągów $\frac{1}{n}$, $\frac{1}{\sqrt{n}}$, $\frac{1}{2^n}$ oraz ciągu, którego wszystkie wyrazy są stale równe stałej wielkości a.
- (*) Oblicz granice ciągów $\frac{2-3n+5n^2}{n^2}$, $\frac{1}{1+\sqrt{n}}$.

1.2.5 Liczba π

Wśród ważnych problemów rozważanych już w starożytności było wyznaczanie obwodu i pola koła o zadanym promieniu. Kluczowy w tych obliczeniach jest obwód koła o średnicy jeden, zwyczajowo oznaczany grecką literą π. Koło o innej średnicy ma proporcjonalnie większy bądź mniejszy obwód, w szególności koło o promieniu r, a więc średnicy $2r$ ma obwód $2\pi r$.

Liczba π, podobnie jak $\sqrt{2}$, jest liczbą niewymierną, ale o tym starożytni nie wiedzieli. Udowodniono to dopiero w XVIII w. Co więcej, w wieku XIX udowodniono, że π jest liczbą transcendentną, to znaczy, że nie jest pierwiastkiem żadnego wielomianu o współczynnikach wymiernych. Oznacza to w szczególności, że nawet jeśli zbiór ułamków rozszerzymy o wszystkie możliwe pierwiastki, to nadal nie wystarczy on do mierzenia wszelkich długości.

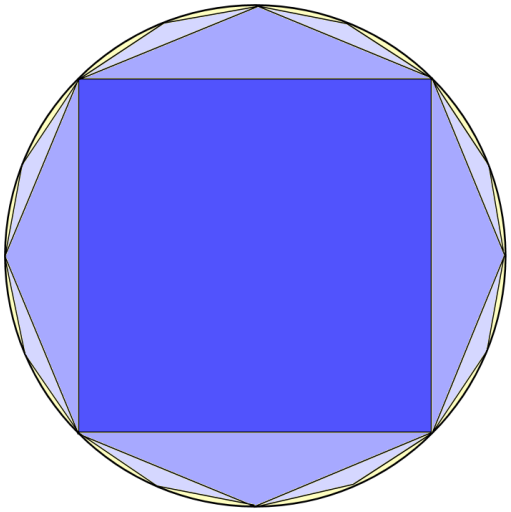
W starożytności przyjmowano różne liczby pomiędzy 3 i 3.16 jako przybliżenie liczby π bez przywiązywania szczególnej wagi do popełnianego błędu. Dopiero grecki matematyk Archimedes (rys. 1.6) zaproponował algorytmiczną metodę wyznaczania π, która przez ponad 1000 lat dostarczała coraz lepszych przybliżeń. Jest to metoda geometryczna, oparta o wpisywanie w okrąg i opisywanie na okręgu wielokątów (rys.1.7). Posługując się 96-kątem Archimedes oszacował, że $223/71 < \pi < 22/7$ ($3.1408 < \pi < 3.1429$), co daje dobre przybliżenie π do dwóch miejsc po przecinku. Co ciekawe, w 1596 roku, ciągle w oparciu o metodę metodę Archimedesa, holenderski matematyk Ludolph van Ceulen policzył π z dokładnością do 35 cyfr.

Oznaczmy przez a_n bok 2^n-kąta foremnego wpisanego w okrąg o średnicy jeden. W oparciu o twierdzenie Pitagorasa można pokazać, że

$$a_2 = \frac{\sqrt{2}}{2}, \qquad a_{n+1} = \frac{1}{2}\sqrt{2 - \sqrt{4 - 4a_n^2}}. \tag{1.2}$$

Zatem obwód tego 2^n-kąta wynosi $A_n := 2^n a_n$.

Można się spodziewać, że ciąg A_n powinien zmierzać do liczby π, bo tyle wynosi obwód koła o średnicy jeden. W rozdziale 4.5.1 przyjrzymy się bliżej zachowaniu ciągu (1.2).



Rysunek 1.7: Przybliżanie obwodu i pola koła wielokątami

Ćwiczenie 1.2.7 • (*) Opierając się na swojej wiedzy z geometrii spróbuj uzasadnić wzór (1.2).

• (*) Uzasadnij stwierdzenie, że bok 2^n-kąta foremnego opisanego na okręgu o średnicy jeden wyraża się wzorem

$$b_n := \frac{a_n}{\sqrt{1-a_n^2}}.$$

W programie Mathematica liczbę π zapisuje się jako `Pi`.

1.3 Szeregi

1.3.1 Pole koła

Archimedesowi zawdzięczamy również sposób liczenia pola koła. Zastosował on tu również metodę wielokątów. Najpierw zauważył, że pole n-kąta foremnego wpisanego w okrąg można policzyć sumując pola n trójkątów równoramiennych. Mają one wszystkie jednakową podstawę x_n i jednakową wyskość h_n, zatem pole wielokąta wynosi $P_n := \frac{nx_nh_n}{2}$ co równe jest $\frac{L_nh_n}{2}$, gdzie $L_n := nx_n$ jest obwodem wielokąta. Przy n zmierzającym do nieskończoności wysokości h_n zmierzają do promienia r koła, a obwody L_n zmierzają do obwodu koła, o którym wiemy już, że wynosi $2\pi r$. Zatem pola P_n zmierzają do $\frac{2\pi rr}{2}$ skąd wnosimy, że pole koła o promieniu r to

$$P := \pi r^2. \tag{1.3}$$

1.3.2 Liczba π jeszcze raz.

Powróćmy do obliczania liczby π. Widzieliśmy, że metoda oparta o przybliżanie obwodu koła daje niezbyt zadowalający wynik. Często jednak, jeśli jeden algorytm zastąpimy innym, to możemy uzyskać lepsze wyniki przy zastosowaniu tej samej arytmetyki komputerowej. Wiemy, że liczba π jest nie tylko obwodem koła o średnicy jeden, ale również polem koła o promieniu jeden. Zastępując pole tego koła polami wpisanych w nie 2^n-kątów foremnych również otrzymamy ciąg przybliżeń liczby π.

Aby wykorzystać wzory z podrozdziału 1.2.5, zamiast koła o promieniu jeden rozważymy ponownie koło o średnicy jeden, a więc o promieniu $1/2$. Zatem ze wzoru (1.3) jego pole to $\frac{\pi}{4}$. Oznaczmy przez P_n pole 2^n-kąta foremnego wpisanego w koło o średnicy jeden oraz przyjmijmy $s_n := 4P_n$. Oczekujemy, że ciąg s_n będzie zmierzał do czterokrotnego pola koła o średnicy jeden, czyli do π. Łatwo zauważyć, że P_{n+1} jest większe od P_n o pole 2^n trójkątów równoramiennych o podstawie równej bokowi 2^n-kąta, a ramionach równych bokowi 2^{n+1}-kąta. Można policzyć, że jeden taki trójkąt ma pole $\frac{a_n(1-\sqrt{1-a_n^2})}{4}$, gdzie a_n, jak poprzednio, oznacza bok 2^n-kąta foremnego wpisanego w koło o średnicy jeden. Zatem

$$P_{n+1} = P_n + 2^n \frac{a_n(1-\sqrt{1-a_n^2})}{4}$$

oraz

$$s_{n+1} = s_n + 2^n a_n (1-\sqrt{1-a_n^2}). \tag{1.4}$$

1.3.3 Szeregi

Warto zwrócić uwagę, że ciąg s_n jest specyficzny. Jego m-ty wyraz może zostać zapisany w postaci

$$s_m = \sum_{n=1}^{m} c_n, \tag{1.5}$$

gdzie

$$c_n = \begin{cases} 0 & \text{dla } n = 1, \\ 2 & \text{dla } n = 2, \\ 2^n a_n (1-\sqrt{1-a_n^2}) & \text{dla } n > 2. \end{cases}$$

Ciąg takiego typu nazywamy szeregiem o wyrazach c_n.

Zazwyczaj (choć bardzo nieściśle) używamy zapisu

$$\sum_{n=1}^{\infty} c_n$$

na oznaczenie szeregu o wyrazach c_n, a s_m dane wzorem (1.5) nazywamy m-tą sumą częściową szeregu.

Mówimy, że szereg jest zbieżny, jeżeli ciąg s_n jest zbieżny. W takim przypadku granicę ciągu s_n nazywamy sumą szeregu i oznaczamy tym samym symbolem co sam szereg.

Szeregi odgrywają ważną rolę w wielu zastosowaniach, dlatego badaniu ich zbieżności poświęca się w analizie matematycznej szczególną uwagę. Szczególnie ważne są narzędzia, które pozwalają określić, czy szereg jest zbieżny, czy nie. Jeśli nie wiemy czy szereg jest zbieżny, to nie możemy z pożytkiem wykorzystywać w obliczeniach numerycznych.

1.4 Funkcje

1.4.1 Koncepcja funkcji

Koncepcja funkcji, jedna z fundamentalnych idei współczesnej matematyki, jako samodzielny obiekt badań pojawiła się w drugiej połowie XVII w. W tym pierwotnym ujęciu funkcja była rozumiana jako wartość, tzw. zmienna zależna, którą można obliczyć rachunkiem algebraicznym bądź poprzez konstrukcje geometryczne w oparciu o inną wartość, tzw. zmienną niezależną oraz pewne stałe. Jako przykład pierwszej kategorii może tu posłużyć funkcja kwadratowa

$$x \mapsto x^2,$$

funkcja wyznaczająca pole koła o promieniu r

$$r \mapsto \pi r^2,$$

czy ogólnie wielomian

$$x \mapsto a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

1.4.2 Miejsca zerowe funkcji

Wśród ważnych zagadnień praktycznych jakie pojawiają się w kontekście funkcji jest pytanie czy i dla jakich wartości zmiennej niezależnej, albo inaczej argumentu, funkcja przyjmuje wartość zero. Argument taki nazywa się miejscem zerowym funkcji.

Jedną z prostych metod poszukiwania miejsca zerowego funkcji jest tzw. metoda bisekcji. Wymaga ona dysponowania dwoma argumentami $a < b$, w których funkcja przyjmuje wartości o znakach przeciwnych. Konstruujemy wtedy rekurencyjnie dwa ciągi

$$\begin{aligned} x_1 &:= a \\ y_1 &:= b \\ x_{n+1} &:= \begin{cases} x_n & \text{if } f(x_n)f(\frac{x_n+y_n}{2}) < 0 \\ \frac{x_n+y_n}{2} & \text{otherwise.} \end{cases} \\ y_{n+1} &:= \begin{cases} y_n & \text{if } f(y_n)f(\frac{x_n+y_n}{2}) < 0 \\ \frac{x_n+y_n}{2} & \text{otherwise.} \end{cases} \end{aligned}$$

W wielu przypadkach, choć niestety nie zawsze, ciągi te dążą do wspólnej granicy będącej miejscem zerowym funkcji f.

Ćwiczenie komputerowe 1.4.1 (*) Zaimplementuj metodę bisekcji w języku Mathematica. Przetestuj ją dla funkcji

(1) $x \mapsto x^2 - 3$ dla $a = 1$ i $b = 3$

(2) $x \mapsto \frac{x+\cos x}{\sin x} - 1$ dla $a = -1$ i $b = 1$.

Postaw hipotezy na temat zbieżności ciągów x_n i y_n. Jeśli uznasz, że są one zbieżne do wspólnej granicy, zastanów się, czy jest ona poszukiwanym miejscem zerowym, a jeśli nie jest, to co jest tego przyczyną.

1.4.3 Wykres funkcji

Jeśli zadawala nas zgrubne przybliżenie miejsc zerowych funkcji, wystarczy narysować wykres funkcji. Z wykresu, obok miejsc zerowych możemy też odczytać tzw. ekstrema, czyli największe i jakie najmniejsze wartości osiąga funkcja w zadanym zbiorze. Rys. 1.8 przedstawia wykres wielomianu $Q(x) := -16 + 16x^2 - 4x^3 - 3x^4 + x^5$, z którego na przykład odczytać możemy, że wartością najmniejszą w przedziale $[-1, 1]$ jest -16. Analizowanie funkcji poprzez narysowanie wykresu wystarcza jednak tylko do pobieżnego zapoznania się z funkcją. Jak zobaczymy, analiza matematyczna dostarcza znacznie lepsze metody do analizowania własności funkcji.

Wykresy funkcji jednej zmiennej bardzo łatwo uzyskać przy pomocy programu Mathematica. Na przykład wykres wielomianu $Q(x) := -16 + 16x^2 - 4x^3 - 3x^4 + x^5$ na rys. 1.8 uzyskano instrukcją

```
Plot[-16+16x^2-4x^3-3x^4+x^5, {x, -2.5, 3.5}]
```

1.4.4 Geometryczne konstrukcje funkcji

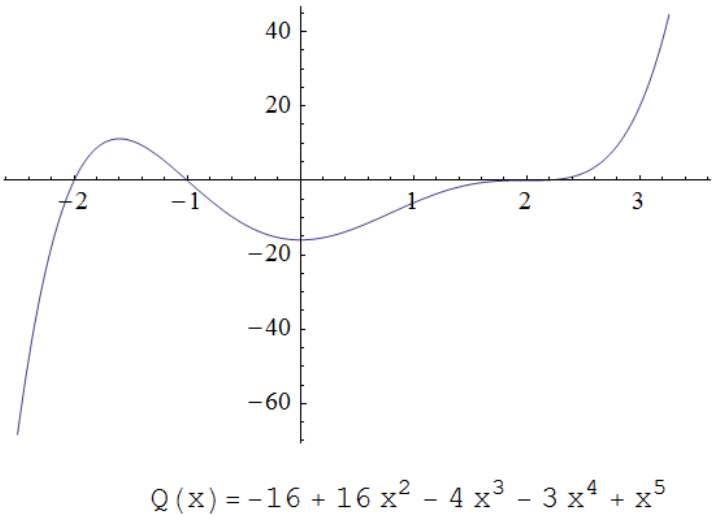
Funkcje pierwotnie definiowane poprzez konstrukcje geometryczne, to przede wszystkim funkcje trygonometryczne, na przykład

$$x \mapsto \sin x.$$

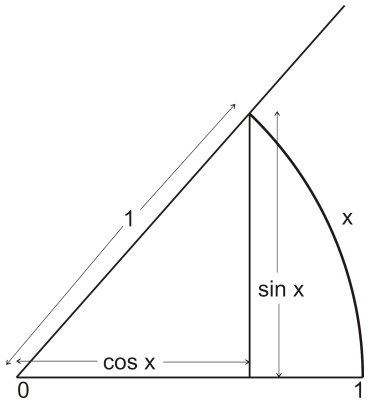
Konstrukcję geometryczną $\sin x$ i $\cos x$ pokazano na rys. 1.9. W konstrukcji tej $|x|$ jest interpretowane jako długość łuku wycinka koła o promieniu jeden, a znak x rozstrzyga o wzajemnym położeniu początku i końca łuku. Przyjęte jest, że znak plus oznacza iż początek łuku występuje przed końcem łuku przy ruchu przeciwnym do ruchu wskazówek zegara. Wykres funkcji $\sin$ przedstawiono na rys. 1.10.

W wielu zastosowaniach bardziej od funkcji $\sin$ i $\cos$ przydaje się funkcja tangens

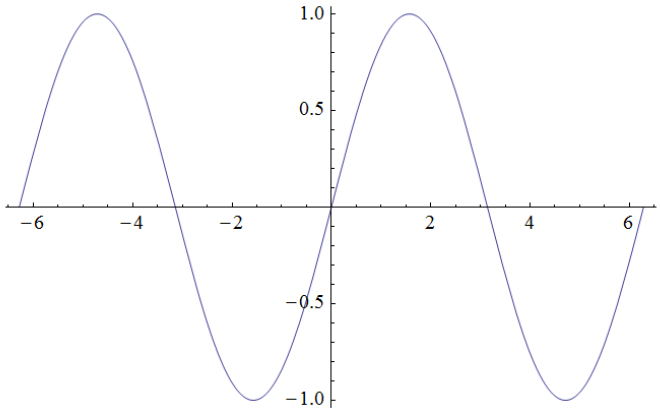
$$\mathrm{tg} : x \mapsto \frac{\sin x}{\cos x}$$



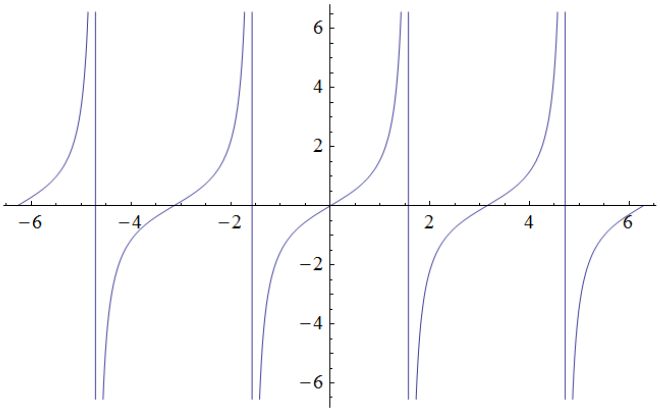
Rysunek 1.8: Wykres wielomianu $Q(x) := -16 + 16x^2 - 4x^3 - 3x^4 + x^5$.



Rysunek 1.9: Geometryczna definicja funkcji sin i cos: w wycinku koła o promieniu 1 i długości łuku x wartość $\sin x$ definiuje się jako długość odcinka spuszczonego z końca łuku prostopadle do ramienia wycinka koła ze znakiem plus, gdy koniec łuku jest ponad ramieniem, a znakiem minus, gdy koniec łuku jest poniżej ramienia.



Rysunek 1.10: Wykres funkcji sin.



Rysunek 1.11: Wykres funkcji tg.

oraz funkcja cotangens

$$\operatorname{ctg} : x \mapsto \frac{\cos x}{\sin x}.$$

Ponieważ dzielenie przez zero nie ma sensu, funkcja tangens nie ma określonej wartości dla argumentów x, dla których funkcja $\cos$ przyjmuje wartość zero, a funkcja cotangens dla argumentów, dla których funkcja sinus się zeruje. Wykres funkcji tg przedstawiono na rys. 1.11.

W programie Mathematica zwyczajowo nazwy funkcji rozpoczyna się od dużej litery. Sin, Cos, Tan, Cot to nazwy funkcji sinus, cosinus, tangens i cotangens. Wykres funkcji sinus na rys. 1.10 uzyskano instrukcją

```
Plot[Sin[x], {x, -2 Pi, 2 Pi}]
```

Możemy też definiować własne funkcje, na przykład

```
f[x_] := x^2
```

definiuje funkcję, która zmiennej x przypisuje wartość x^2. Teraz pisząc

```
f[3]
```

otrzymamy wartość 9.

Co ciekawe, możemy też definiować funkcje, które w różnych przypadkach zdefiniowane są różnymi wzorami. Na przykład

```
g[x_] := If[x^2 > 1, x^2, -1]
```

definiuje funkcję, którą przy zastosowaniu tradycyjnej matematycznej notacji zapisalibyśmy wzorem

$$x \mapsto \begin{cases} x^2 & \text{gdy } x^2 > 1, \\ -1 & \text{w przeciwnym razie.} \end{cases}$$

Instrukcja `If` *języka Mathematica to tak zwana instrukcja warunkowa. Ma ona trzy argumenty. Zwraca drugi, gdy pierwszy ewaluuje się do prawdy. W przeciwnym razie zwraca argument trzeci.*

1.4.5 Rozwijanie funkcji w szereg

Jednym z ważniejszych osiągnięć analizy matematycznej było odkrycie, że funkcje trygonometryczne nie muszą być obliczane metodą geometryczną, ale mogą być wyznaczone przez przybliżanie wartości szeregu. Na przykład

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

gdzie $m! = 1 \cdot 2 \cdot 3 \cdots n$*, czytane* m *silnia, to iloczyn kolejnych liczb naturalnych od* 1 *do* m*. Wzór ten oraz podobne wzory dla innych funkcji trygonometrycznych odkrył w wieku XIV hinduski matematyk Madhavan of Sangamagramam.*

Dzisiejsza definicja funkcji jest nieporównanie szersza od pierwotnej koncepcji i obejmuje ciągi jako szczególny przypadek. W stosowanej dzisiaj definicji nie wnikamy w jaki sposób (algebraiczny, geometryczny czy jakikolwiek inny) określono jak zmiennej niezależnej przypisano zmienną zależną, traktując ją po prostu jako skądś daną. Taką jest na przykład funkcja Dirichleta

$$x \mapsto \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q}, \end{cases}$$

w której definicji nie interesujemy się jak w praktyce sprawdzić, czy liczba jest wymierna czy nie jest.

1.5 Pola figur i całkowanie

1.5.1 Całka oznaczona

Widzieliśmy, że koncepcja granicy jest pomocna przy obliczaniu pola koła. W podobny sposób możemy postępować z innymi figurami. Problem można sprowadzić do liczenia pola pod wykresem funkcji. Rozważmy na przykład figurę o soczewkowatym kształcie przedstawioną na rys. 1.12.

Jak łatwo zauważyć pole tej figury jest różnicą pól pod wykresami funkcji

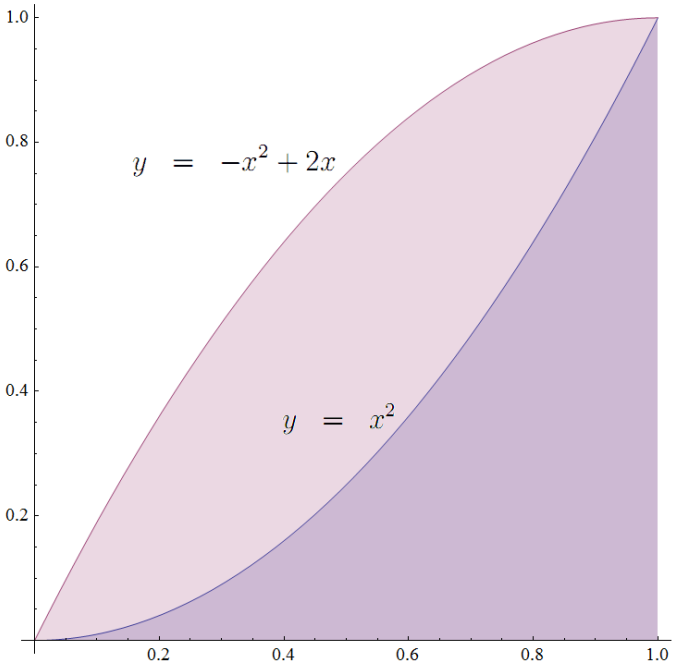
$$\begin{aligned} y &= -x^2 + 2x, \\ y &= x^2. \end{aligned}$$

Pole pod wykresem funkcji f *na przedziale* $[a, b]$ *określane jest mianem* całki oznaczonej *i oznaczane dość tajemniczym symbolem*

$$\int_a^b f(x)dx.$$

Do policzenia całki oznaczonej z funkcji f *na przedziale* $[a, b]$ *też możemy zastosować technikę przejścia granicznego. W tym celu dzielimy przedział* $[a, b]$ *na* n *równych przedziałów i nad każdym z nich budujemy prostokąt o wysokości równej wartości funkcji* f *w prawym końcu przedziału. Sumując pola prostokątów otrzymujemy przybliżenie pola pod wykresem równe:*

$$\sum_{i=1}^{n} \frac{1}{n} f(x_i), \tag{1.6}$$



Rysunek 1.12: Pole soczewkowatej figury jest różnicą pól pod wykresami funkcji.

gdzie

$$x_i = a + i\frac{b-a}{n}.$$

Przykładowy efekt takiego postępowania dla funkcji

$$y = \sqrt{1-x^2}$$

na przedziale $[0,1]$ *przy* $n = 5$ *przedstawiono na rys. 1.13.*

Podobnie jak poprzednio naturalne jest oczekiwanie, że takie postępowanie w granicy da poszukiwane pole pod wykresem funkcji.

Ten sposób liczenia pól stosowano już w starożytności, choć zrozumienie przyczyn jego poprawności bardzo długo pozostawało mgliste. Proces przejścia granicznego opisywano jako zastąpienie skończonego podziału podziałem na nieskończoną ilość przedziałów o nieskończenie małej lecz dodatniej długości dx*. Ta tajemnicza liczba* dx *dawała po przemnożeniu przez wartość funkcji na tym przedziale kolejną nieskończenie małą lecz dodatnią wartość* $f(x)dx$*. Dopiero wysumowanie nieskończonej ilości tych nieskończenie małych wartości dawało poszukiwane pole. Opis ten oczywiście wyjaśnia zapis i nazwę całki (całkowita suma, powstała przez wysumowanie wszystkich nieskończenie małych* $f(x)dx$*), ale w żadnym przypadku nie może być uznany za definicję całki oznaczonej.*

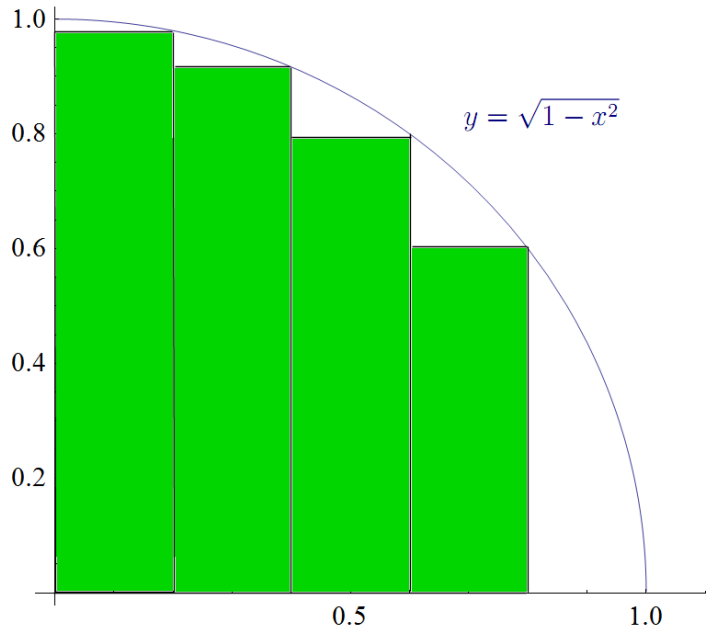
Dopiero niemiecki matematyk, Georg Friedrich Bernhard Riemann (rys. 1.14) podał jako pierwszy precyzyjną definicję całki oznaczonej oraz warunki, kiedy ta definicja ma sens. Dziś wyrażenie (1.6) *określa się mianem sumy Riemanna.*

1.5.2 Całka nieoznaczona

Zakładając, że całka oznaczona z funkcji f *jest dobrze zdefiniowana (do czego potrzebujemy sum Riemanna i granicy), możemy zdefiniować funkcję*

$$F(u) := \int_0^u f(x)dx,$$

dla której zachodzi wzór



Rysunek 1.13: Pole pod wykresem ćwiartki okręgu.



Rysunek 1.14: Georg Friedrich Bernhard Riemann, 1826-1866. Źródło: Wikipedia

$$\int_a^b f(x)dx = F(b) - F(a).$$

Funkcja F spełniająca powyższy wzór nosi nazwę całki nieoznaczonej funkcji f. Oznaczamy ją (nieprecyzyjnie) symbolem

$$\int f(x)dx$$

Jeśli potrafilibyśmy znaleźć analityczny sposób na znalezienie całki nieoznaczonej w oparciu o funkcję f, liczenie pól znacznie ułatwiłoby się. Żyjący na przełomie X i XI w. naukowiec arabski znany w Europie jako Alhacen potrafił pokazać (oczywiście w ówczesnym języku matematyki), że jeśli f jest wielomianem stopnia drugiego

$$f(x) := a_0 + a_1 x + a_2 x^2,$$

to jego całką nieoznaczoną jest

$$F(x) := a_0 x + \frac{a_1 x^2}{2} + \frac{a_2 x^3}{3}.$$

Alhacen potrafił też podać wzór na całkę nieoznaczoną dla wielomianów stopnia trzeciego i czwartego. Wykorzystując ten wzór, możemy bardzo szybko policzyć pole naszej soczewki z rys.1.12, bo przyjmując wygodne oznaczenie

$$[F(u)]_a^b := F(b) - F(u)$$

otrzymujemy

$$\begin{aligned} \int_0^1 x^2 dx &= \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3} \\ \int_0^1 (-x^2 + 2x)dx &= \left[-\frac{x^3}{3} + \frac{2x^2}{2}\right]_0^1 = \frac{2}{3}. \end{aligned}$$

Płyną z tego dwa ważne wnioski. Po pierwsze, granica skomplikowanego ciągu wcale nie musi być trudną w wyliczeniu liczbą niewymierną. Po drugie i ważniejsze, pojęcie granicy może być pomocne nie tylko w obliczeniach liczb niewymiernych. Nie mniej, a może nawet więcej, przydaje się ono do definiowania nowych pojęć, trudnych do zdefiniowania inaczej.

1.6 Styczna do krzywej i granica funkcji

1.6.1 Styczna do krzywej

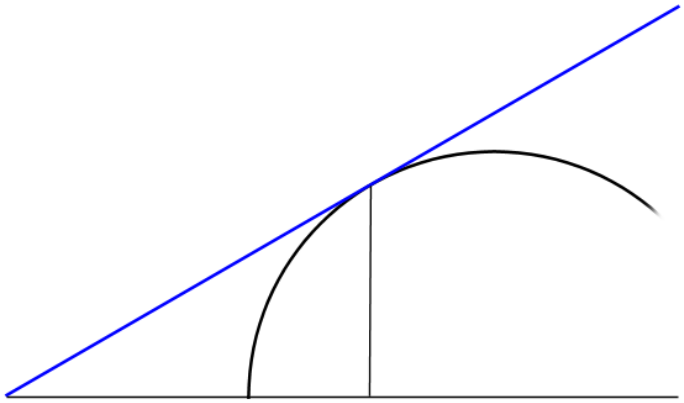
Problem wyznaczenia stycznej do krzywej w zadanym punkcie to też jeden z problemów rozwiązywanych już w starożytności przy wykorzystaniu przejścia granicznego. W uproszczeniu, styczna do krzywej w danym punkcie krzywej, to prosta przechodząca przez zadany punkt, która w jego pobliżu nie przecina krzywej w żadnym innym punkcie. Rys. 1.15 przedstawia styczną do łuku okręgu.

Styczną można wyznaczać geometrycznie, przykładając do krzywej linijkę, ale nie jest to metoda dokładna. Lepiej jest wyliczyć kąt jaki styczna tworzy z prostą poziomą, albo tak zwany współczynnik nachylenia prostej, czyli tangens tego kąta.

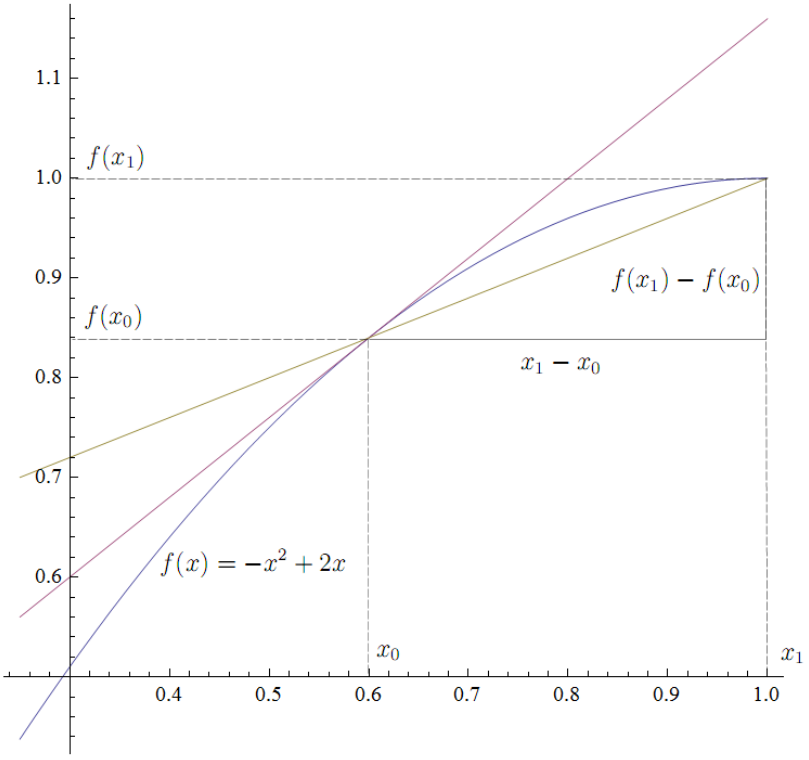
Jeśli krzywa zadana jest funkcją f, to współczynnik nachylenia stycznej w punkcie wykresu o współrzędnych $(x_0, f(x_0))$ można przybliżyć, analizując sieczną poprowadzoną z punktu $(x_0, f(x_0))$. Jest to prosta, która przecina krzywą w punkcie $(x_0, f(x_0))$ oraz dodatkowo w punkcie $(x_1, f(x_1))$, gdzie $x_1 = x_0 + h$ jest bliskie x_0.

Rysując stosowny trójkąt prostokątny (rys. 1.16), łatwo zauważyć, że tangens kąta, jaki sieczna tworzy z osią x wynosi

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{f(x_0 + h) - f(x_0)}{h}.$$



Rysunek 1.15: Styczna do okręgu



Rysunek 1.16: Styczna i sieczna do krzywej.



Rysunek 1.17: Karol Weierstrass (1815-1897). Źródło: Wikipedia

Zmierzając z h do zera zbliżamy sieczną do stycznej, a więc w granicy dostać powinniśmy współczynnik nachylenia stycznej. I rzeczywiście dostajemy, jeśli tylko funkcja jest odpowiednio porządna.

Należy zwrócić uwagę, że mamy tu do czynienia z innego typu granicą niż rozważane do tej pory, bo nie rozważamy ciągu, lecz funkcję

$$h \mapsto \frac{f(x_0+h)-f(x_0)}{h} \tag{1.7}$$

zmiennej h (x_0 traktujemy jako parametr). Jak definiuje się granicę ciągu już mniej więcej wiemy. A jak się to robi w przypadku granicy funkcji?

1.6.2 Granica funkcji

Chociaż pojęciem granicy funkcji operowano co najmniej od XVII w., to pierwszą precyzyjną definicję postawił czeski matematyk Bernard Bolzano dopiero w 1817 roku, a używaną współcześnie definicję podał niemiecki matematyk Karol Weierstrass (rys. 1.17).

Według Weierstrassa funkcja f ma w punkcie x_0 granicę g jeżeli

$$\forall \epsilon > 0 \, \exists \delta > 0 \, \forall x \neq x_0 \quad |x - x_0| < \delta \;\Rightarrow\; |f(x) - g| < \epsilon.$$

Piszemy wtedy

$$\lim_{x \to x_0} f(x) = g.$$

1.6.3 Funkcje ciągłe

Warto sobie od razu powiedzieć, że w przeciwieństwie do ciągów, typowa funkcja, z jaką spotykamy się na co dzień, jeśli jest określona w punkcie x_0, to ma w nim granicę równą swojej wartości w x_0. Może to robić wrażenie, że teoria granicy funkcji jest mniej użyteczna od granicy ciągu. Ale właśnie ten fakt, którego bez pojęcia granicy funkcji nie da się zaobserwować, jest interesujący sam w sobie. Funkcje o tej własności określamy mianem ciągłych. Obrazowo mówimy, że funkcja ciągła to taka funkcja, której wykres można narysować bez odrywania ołówka od papieru. Nie jest to do końca prawdą, bo dotyczy jedynie funkcji, których dziedzina jest odcinkiem domkniętym.

1.6.4 Obliczanie granic.

Granica funkcji może być liczona w punkcie x_0, w którym funkcja jest nieokreślona, ale jest określona dla wartości różnych od x_0. Wtedy ciągłość funkcji w punktach różnych od x_0 nie wystarcza, by wyznaczyć granicę. Mimo to często można ją łatwo policzyć. Na przykład funkcja

$$f : x \to \frac{2x + x^2}{x}$$

jest nieokreślona w zerze, ale określona dla wartości większych i mniejszych od zera. Dla wartości tych jest ona równa funkcji

$$x \to 2 + x,$$

która jest ciągła i określona w zerze, a zatem ma w zerze granicę dwa. Zatem

$$\lim_{x\to 0} f(x) = 2.$$

Przykład ten może wydawać się sztuczny, ale docenimy jego wartość, gdy uświadomimy sobie, że dokładnie taka granica pojawia się przy obliczaniu współczynnika kierunkowego stycznej do wykresu funkcji $x \mapsto x^2$ w punkcie 1, bo zgodnie ze wzorem (1.7) jest to funkcja

$$h \mapsto \frac{(x_0 + h)^2 - x_0^2}{h} = \frac{2h + h^2}{h},$$

czyli funkcja f. Oczywiście nie wszystkie granice funkcji dają się tak łatwo policzyć. I tu, jak w przypadku ciągów, korzystamy z różnych twierdzeń. W szczególności dla funkcji można udowodnić odpowiedniki twierdzenia o trzech ciągach oraz o granicy sumy, różnicy, iloczynu i ilorazu.

Problem 1.6.1 Jak uogólnić pojęcie granicy, tak by granica ciągu i funkcji wpadały "do jednego worka" i odpowiednie twierdzenia obejmowały tak ciągi jak i funkcje? Jak dołączyć do takiej definicji przypadek ciągu lub funkcji zmierzającej do nieskończoności?

Oczywiście granica funkcji nie musi istnieć. Jest tak na przykład w przypadku funkcji

$$x \mapsto \sin \frac{1}{x},$$

która jest określona i ciągła dla wszystkich x z wyjątkiem zera, ale w zerze nie ma granicy.

Z kolei funkcja

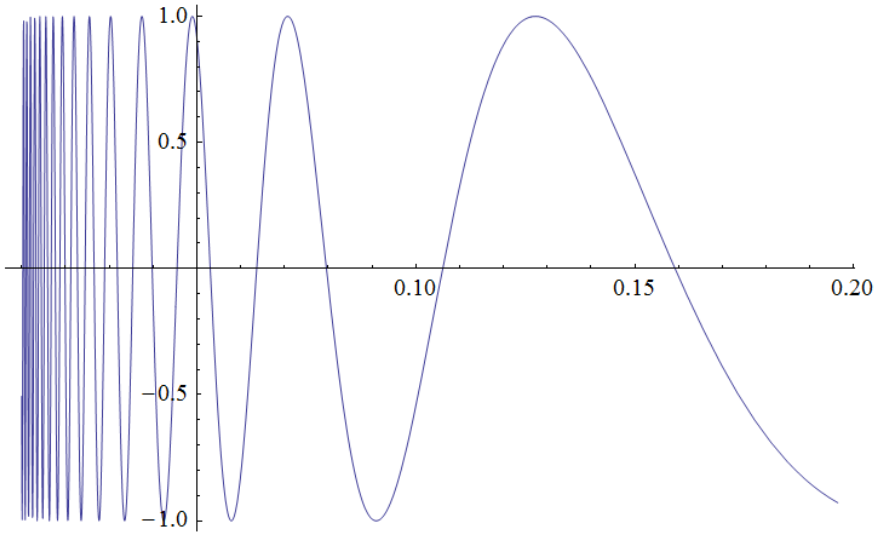
$$x \mapsto \frac{\sin x}{x},$$

która też jest określona i ciągła dla wszystkich x z wyjątkiem zera, w zerze ma granicę wynoszącą jeden, ale nie jest oczywiste jak to pokazać. Jednak z z wykresu (rys. 1.19) widać, że należy się tego spodziewać.

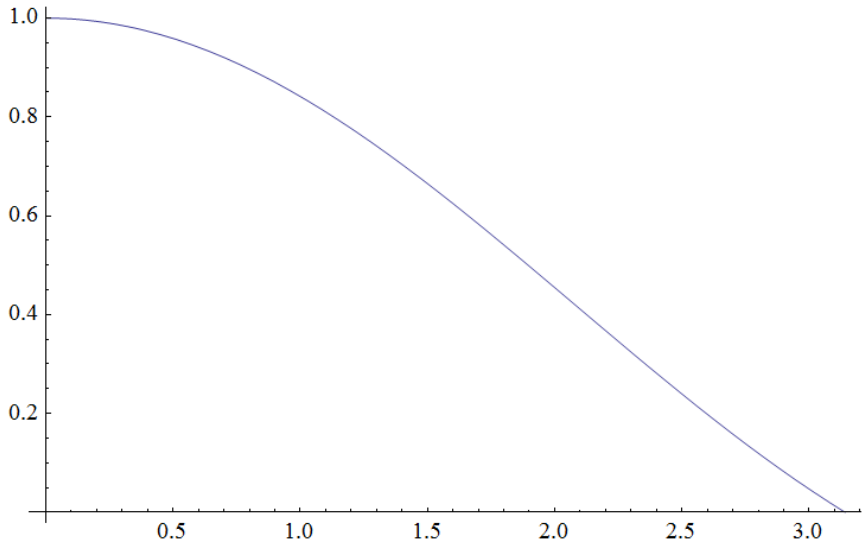
1.7 Pochodna funkcji

1.7.1 Lokalne maksima i minima

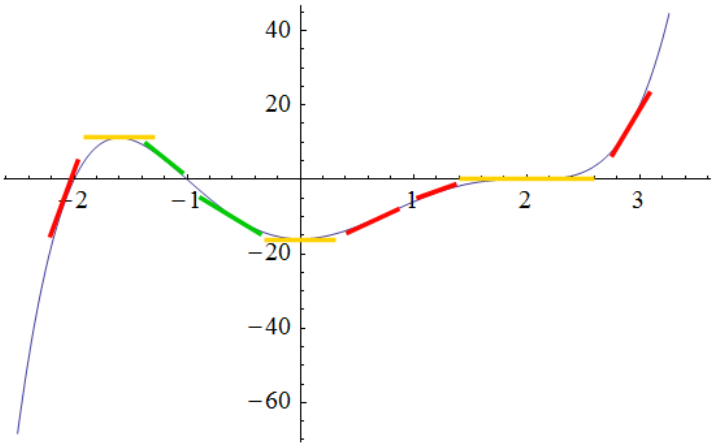
Współczynnik nachylenia stycznej do wykresu funkcji niesie w sobie informacje o samej funkcji. Tam gdzie ten współczynnik jest dodatni funkcja (umownie) rośnie, a gdzie jest ujemny maleje. Zerowa wartość współczynnika nachylenia oznacza, że styczna jest pozioma. Może to świadczyć o lokalnej górce (lokalne maksimum) lub lokalnym dołku (lokalne minimum). Pozioma styczna może też wystąpić w miejcu, chwilowo przestaje rosnąć bądź maleć.



Rysunek 1.18: Wykres funkcji $x \mapsto \sin \frac{1}{x}$. Granica tej funkcji w zerze nie istnieje



Rysunek 1.19: Wykres funkcji $x \mapsto \frac{\sin x}{x}$. Granica tej funkcji w zerze wynosi jeden.



$$Q(x) = -16 + 16\,x^2 - 4\,x^3 - 3\,x^4 + x^5$$

Rysunek 1.20: Wybrane styczne do wykresu funkcji. Na czerwono zaznaczono styczne w miejscach gdzie wykres się wznosi, na zielono gdzie opada, a na żółto miejsca gdzie styczna jest pozioma.

1.7.2 Pochodna funkcji

Dlatego dla danej funkcji f interesujące jest poszukiwanie miejsc zerowych nowej funkcji, która punktowi x przypisuje współczynnik nachylenia stycznej do wykresu funkcji f w punkcie x. Funkcja taka nazywana jest pochodną funkcji f i oznaczana f'. Wiemy, że współczynnik kierunkowy stycznej obliczamy jako granicę przy $h \to 0$ wyrażenia

$$\frac{f(x+h) - f(x)}{h}$$

zwanego ilorazem różnicowym. Zatem

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{1.8}$$

oczywiście przy założeniu, że ta granica istnieje. Wprowadzając oznaczenia na różnice

$$\begin{aligned} \Delta x &:= (x+h) - x = h \\ \Delta y &:= f(x+h) - f(x) \end{aligned}$$

możemy iloraz różnicowy napisać w postaci

$$\frac{\Delta y}{\Delta x},$$

co wyjaśnia jego nazwę. Używając nieprecyzyjnego języka nieskończenie małych można powiedzieć, że pochodną funkcji otrzymujemy zastępując w ilorazie różnicowym róznice Δy i Δx nieskończenie małymi różnicami oznaczanymi dy i dx i zwanymi różniczkami. Wyjaśnia to używaną do dziś nazwę teorii pochodnych: rachunek różniczkowy. Sama koncepcja różniczki doczekała się precyzyjnej definicji i choć odbiega ona od pierwotnej idei, pozwala zapis

$$f'(x) = \frac{dy}{dx}$$

uczynić precyzyjnym.

Jeśli potrafimy ze wzoru na funkcję f wyprowadzić wzór na funkcję f' oraz potrafimy znaleźć miejsca zerowe funkcji f', to potrafimy wskazać miejsca podejrzane o występowanie lokalnych minimów i lub maksimów. Okazuje się, że choć w definicji pochodnej występuje granica, to w wielu przypadkach jest to granica, którą możemy policzyć czysto analitycznie bez konieczności stosowania przybliżeń. Na przykład dla funkcji $x \mapsto f(x) := x^2 - 4x + 5$ mamy

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{(x+h)^2 - 4(x+h) + 5 - (x^2 - 4x + 5)}{h} \\ &= \lim_{h\to 0} (2x + h - 4) = 2x - 4, \end{aligned}$$

bo funkcja $x \mapsto 2x - 4$ jest ciągła. Widać, że $f'(2) = 0$, zatem funkcja f, jeśli ma gdzieś minimu lokalne to w punkcie 2.

1.7.3 Podstawowe twierdzenie rachunku różniczkowego i całkowego

Perski matematyk Sharaf al-Din al-Tusi już z końcem XII w. potrafił pokazać, że (używając współczesnego języka) pochodną wielomianu trzeciego stopnia

$$Q(x) := b_0 + b_1 x + b_2 x^2 + b_3 x^3$$

jest

$$Q'(x) = b_1 + 2b_2 x + 3b_3 x^2.$$

Przypomnijmy, że około 200 lat wcześniej, Alhacen policzył, że całką nieoznaczoną

$$f(x) := a_0 + a_1 x + a_2 x^2,$$

jest

$$F(x) = \int f(x)dx = a_0 x + \frac{a_1 x^2}{2} + \frac{a_2 x^3}{3}.$$

Podstawiając

$$\begin{aligned} b_1 &= a_0 \\ 2b_2 &= a_1 \\ 3b_3 &= a_2 \end{aligned}$$

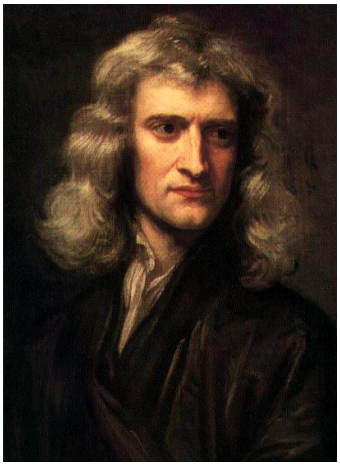
możemy zaobserwować, że jest związek między między liczeniem pochodnej i całkowaniem:

$$\begin{aligned} \int Q'(x)dx &= b_1 x + b_2 x^2 + b_3 x^3 = Q(x) - b_0 \\ \left(\int f(x)dx\right)' &= F'(x) = f(x). \end{aligned}$$

Zatem już w XII w. były podstawy do odkrycia tego związku, ale udało się to dopiero 500 lat później. Pod koniec XVII w. Anglik Isaac Newton oraz Niemiec Gottfried Leibniz (rys. 1.21) pokazali następujące twierdzenie

Twierdzenie 1.7.1 *Jeśli f jest pochodną funkcji F, to F jest całką nieoznaczoną funkcji f. Jeśli F jest całką nieoznaczoną funkcji f, to pochodną funkcji F jest f.*

Twierdzenie to jest dzisiaj określane mianem podstawowego twierdzenia rachunku różniczkowego i całkowego. Do czasu odkrycia tego twierdzenia teorie pochodnej i całki rozwijane były niezależnie, przy czym w liczeniu pochodnych notowano większe postępy niż w liczeniu całek. Odkrycie Newtona i Leibniza pokazało, że w istocie jest jedna teoria, a każdy wzór na pochodną jest równocześnie wzorem na całkę. Umożliwiło to liczenie wielu nowych całek, dało impuls do intensywnego rozwoju rachunku różniczkowego i całkowego i przyczyniło się do szybkiego rozwoju nauk technicznych.





Rysunek 1.21: Isaac Newton (1643 - 1727) i Gottfried Leibniz (1646 - 1716). Źródło: Wikipedia

Rozdział 2

Teoria mnogości

2.1 Problemy z nieskończonością

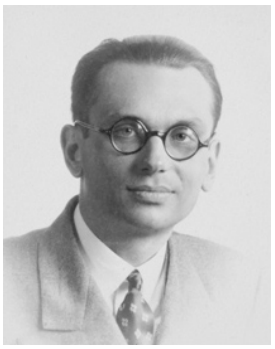
Niewymierność liczby $\sqrt{2}$ uświadomiła matematykom, że czasami konieczne jest rozważanie bytów nieskończonych. W przypadku liczby $\sqrt{2}$ jest to niekończony ciąg przybliżeń $\sqrt{2}$ liczbami wymiernymi. Stwierdzenia o ciągach skończonych, przynajmniej teoretycznie, można sprawdzić analizując po kolei wszystkie wyrazy ciągu. Ponieważ jest ich skończenie wiele, da się to zrobić w skończonym czasie. Rozważania na temat ciągów nieskończonych i nieskończonych bytów przez całe tysiąclecia sprawiały matematykom trudność. Używano pojęć "nieskończenie mały" lub "nieskończenie wielki" choć widać było, że pojęcia te są kłopotliwe: dało się z nich wywieść tak wnioski pożądane jak i nonsensowne.

W XIX wieku, wraz ze wzrostem precyzji w fromułowaniu idei rachunku różniczkowego i całkowego i pojawiającymi się coraz częściej sprzecznościami związanymi z nieodpowiednim posługiwaniem się nieskończonością, zaczęła narastać potrzeba usystematyzowania metod badawczych matematyki. W 1920 roku niemiecki matematyk Dawid Hilbert (rys. 2.1) zaproponował swój program, dziś znany jako program Hilberta, formalnego podejścia do matematyki. Zuważył on, że choć analiza matematyczna wymaga operowania nieskończonością, to język, w którym mówi się o matematyce jest skończony, bo kłada się ze skończonej ilości słów.

Hilbert postulował, że wszystkie twierdzenia matematyki należy wyprowadzić drogą formalnego dowodu z dobrze dobranego zestawu aksjomatów, czyli możliwie oczywistych stwierdzeń przyjmowanych bez dowodu. Nie był to pomysł nowy: w III stuleciu p.n.e takie podejście do geometrii zaproponował grecki matematyk Euklides (rys. 2.2). Zasługą Hilberta była obserwacja, że takie podejście może rozwiązać problemy z nieskończonością. Każde twierdzenie daje się zapisać jako skończony ciąg słów, a każdy dowód jest skończonym ciągiem formuł, również zbudowanych ze skończonej ilości słów. Zatem poprawność dowodu można sprawdzić drogą skończonej inspekcji i to również w przypadku gdy samo twierdzenie dotyczy bytów nieskończonych. Matematycznym ujęciem formalnego opisu teorii matematycznej, w tym dowodów zajmuje się logika matematyczna oraz tzw. metamatematyka.

By jednak podejście takie było przekonujące wypada najpierw udowodnić, że zestaw przyjętych bez dowodu aksjomatów nie jest sprzeczny sam w sobie.





Rysunek 2.1: Dawid Hilbert, 1862-1943 oraz Kurt Gödel, 1906-1978. Źródło: Wikipedia



Rysunek 2.2: Euklides, IV-III w. p.n.e., przedstawiony na posągu w Muzeum Historii Naturalnej w Oksfordzie. Źródło: Wikipedia

Hilbert gorąco wierzył, że jego program da się zrealizować. Jednak już 11 lat później austriacki matematyk Kurt Gödel (rys. 2.1) pokazał, że nie istnieje niesprzeczny układ aksjomatów, z którego dałoby się wyprowadzić wszystkie twierdzenia dotyczące liczb naturalnych. Innymi słowy, jeśli teoria obejmująca arytmetykę liczb naturalnych jest niesprzeczna (wywiedziona z niesprzecznego zestawu aksjomatów) to jest niezupełna (są w niej zdania p, dla których nie ma dowodu ani dla p, ani dla nieprawda, że p).

Pomimo twierdzenia Gödla jak i trudności w realizacji postulatu sprawdzenia niesprzeczności program Hilberta szybko przyjął się i pozostaje do dziś powszechnie przyjętą metodą uprawiania matematyki.

2.2 (*)Uproszczony formalizm

Jedną z metod sprawdzania niesprzeczności aksjomatów teorii jest pokazanie, że aksjomaty te są twierdzeniami innej teorii, której niesprzeczność już udowodniono, albo przynajmniej teorii, w której niesprzeczność aksjomatów budzi mniej wątpliwości.

W pierwszym dwudziestoleciu XX wieku teoria zbiorów, w języku polskim najczęściej określana mianem teoria mnogości, stała się fundamentem głównego nurtu współczesnej matematyki, jako najbardziej elementarna teoria, w której weryfikować można aksjomaty bardziej zaawansowanych teorii.

Twórcą teorii mnogości był niemiecki matematyk Georg Cantor (rys. 2.3).

Do całkowicie formalnego przedstawienia matematycznej teorii używa się języka zbudowanego na bazie pewnego alfabetu. W przypadku teorii mnogości a wraz z nią praktycznie całej matematyki wystarcza (z nadmiarem) całkiem ubogi język zbudowany na bazie alfabetu obejmującego

- Symbole logiczne
  - $\neg$ - negacja
  - $\vee$ - alternatywa
  - $\wedge$ - koniunkcja
  - $\Rightarrow$ - implikacja
  - $\Leftrightarrow$ - równoważność
  - $\forall$ - kwantyfikator "dla każdego"
  - $\exists$ - kwantyfikator "istnieje"
- symbole relacyjne i funkcyjne



Rysunek 2.3: Georg Cantor, 1845-1918. Źródło: Wikipedia

- przeliczalny zestaw nazw (zmiennych): $\square, \square|, \square||, \ldots$

- nawiasy (i)

W przypadku teorii mnogości nie są potrzebne symbole funkcyjne i jest tylko jeden symbol relacyjny: $\in$ - symbol przynależności do zbioru.

W praktyce postępować będziemy znacznie swobodniej, przede wszystkim stosując bogatszy zestaw zmiennych opartych o różne czcionki ($a, b, c, ..., A, B, C, ..., \mathbb{A}, \mathbb{B}, \mathbb{C}, ..., \mathcal{A}, \mathcal{B}, \mathcal{C}, ..., \alpha, \beta, \gamma, ..., \Gamma, \Delta, \Lambda...$) a także nazw wieloznakowych jak np. $\max, \sup, \sin, \ln, ...$ oraz różnych zamienników symboli logicznych.

Formułami (wyrażeniami) języka są skończone ciągi znaków alfabetu, które spełniają pewne formalne reguły. Przykładowo $\exists x \in X$ jest formułą, a $x \exists X \in$ formułą nie jest. Nie będziemy formalizować reguł wymaganych od ciągów znaków, by były formułami języka teorii mnogości, zdając się w tej materii na intuicję.

Ponieważ przyjmujemy nieskończony zestaw nazw w alfabecie, za każdym razem możemy teoretycznie użyć nowej nazwy, by uniknąć kolizji nazw. Jest to jednak niewygodne, a w praktyce niewykonalne. Dlatego wygodnie operować pojęciem zakresu występowania nazwy. Rozróżniać będziemy:

- związane występowanie nazwy - ograniczone do formuły, w której nazwa jest użyta. Przykładem jest występowanie zmiennej x w formułach $\exists x\ \varphi(x)$, $\int_0^1 f(x)dx$.

- niezwiązane występowanie nazwy - występowanie rozpoczęte zwrotem "Niech x będzie takie, że ..." lub $x := ...$ czytane "Zdefiniujmy x jako ...". Zakres występowania takiej nazwy obejmuje wszsytkie formuły od użycia tego zwrotu aż do ponownego użycia tego zwrotu, które to rozpoczyna nowy zakres.

Formuła, w której wszystkie występujące nazwy są związane jest *zdaniem* języka, w przeciwnym razie jest *formą zdaniową*. Formę zdaniową o niezwiązanych zmiennych $x_1, x_2, \ldots x_n$ zapisujemy często w postaci $\varphi(x_1, x_2, \ldots x_n)$.

Definicje rozumiemy jako skróty. Do definiowania używać będziemy symbolu $:\Leftrightarrow$ lub $\Leftrightarrow :$, w którym po stronie dwukropka umieszczać będziemy to co definiujemy, a po stronie drugiej to jak definiujemy. Oto pierwsze definicje:

$$
\begin{array}{rrcl}
 & y \notin z & :\Leftrightarrow & \neg \quad y \in z \\
 & y \subset z & :\Leftrightarrow & \forall x \quad x \in y \;\Rightarrow\; x \in z \\
 & y = z & :\Leftrightarrow & y \subset z \quad \wedge \quad z \subset y \\
 & y \neq z & :\Leftrightarrow & \neg \quad y = z \\
 & y \subsetneq z & :\Leftrightarrow & y \subset z \quad \wedge \quad y \neq z \\
\exists! x & \varphi(x) & :\Leftrightarrow & \exists x \quad \varphi(x) \quad \wedge \quad (\forall y \forall z \quad \varphi(y) \wedge \varphi(z) \;\Rightarrow\; y = z) \\
\forall x \in A & \varphi(x) & :\Leftrightarrow & \forall x \quad x \in A \;\Rightarrow\; \varphi(x) \\
\exists x \in A & \varphi(x) & :\Leftrightarrow & \exists x \quad x \in A \wedge \varphi(x)
\end{array}
$$

Trzeba podkreślić, że w praktyce nawet tak bardzo uproszczony formalizm jest zbyt zawiły i bardzo utrudnia rozumienie tekstu. Na codzień wymagamy od czytelnika interpretacji nieformalnego tekstu, to znaczy domyślenia się w jaki sposób należy tekst nieformalny lub nie całkiem formalny zastąpić poprawnym tekstem formalnym. Uważny czytelnik szybko zauważy stosowanie niezdefiniowanego zapisu skrótowego wymagającego interpretacji bądź równoległe stosowanie tego samego oznaczenia (tej samej nazwy) w dwóch różnych znaczeniach, to znaczy wymagającego stosowania odmiennych nazw w zapisie formalnym. Ustalenie właściwego znaczenia odbywa się zazwyczaj na podstawie kontekstu.

W każdej matematycznej teorii przyjmuje się jako oczywisty pewien zestaw zdań, tzw. *aksjomatów*. *Wywodem* zdania F z przesłanek (zdań pomocniczych) P_1, P_2,... P_k jest ciąg formuł języka spełniający trzy warunki:

- pierwszą formułą dowodu jest formuła $\neg F$,
- każda formuła jest jedną z przesłanek P_i, aksjomatem, już udowodnionym twierdzeniem bądź wynika z formuł poprzednich na mocy przyjętych reguł wnioskowania,
- w ciągu tym jest para formuł A, $\neg A$.

Dowodem, a dokładniej *dowodem nie wprost* zdania F nazywamy jego wywód nie wymagający użycia żadnej przesłanki. Dowód nie wprost jest oparty o koncepcję, że ze zdania prawdziwego nie można wywnioskować jakiegoś innego zdania i równocześnie zaprzeczenia tego zdania.

Twierdzeniem jest zdaniem, dla których można podać dowód. Jeśli dowód twierdzenia jest stosunkowo prosty to samo twierdzenie określa się mianem *uwaga*. Twierdzenia, których rola jest głównie pomocnicza określa się mianem *lematu*. Twierdzenia wynikające bezpośrednio z wcześniej udowodnionego innego twierdzenia określa się mianem *wniosku*.

Każda reguła wnioskowania ma postać

$$A_1, A_2, \ldots A_n \quad \vdash \quad B,$$

gdzie $A_1, A_2, \ldots A_n, B$ są formułami, a symbol $\vdash$ oznacza zwrot "pozwana wywnioskować, że". W sformalizowanej matematyce reguły wnioskowania są ściśle wyprecyzowane i ograniczone do absolutnego minimum. Dla naszej wygody przyjmiemy dość bogaty zestaw reguł wnioskowania. Przede wszystkim będzie to następujący zestaw reguł opartych o tautologie logiczne (zdania, prawdziwe bez względu na prawdziwość zdań składowych):

$$
\begin{array}{cc}
A, A \Rightarrow B \;\vdash\; B & \\
\neg\neg A \;\vdash\; A & A \;\vdash\; \neg\neg A \\
A \wedge B \;\vdash\; A & A \wedge B \;\vdash\; B \\
A \vee B, \neg A \;\vdash\; B & A \vee B, \neg B \;\vdash\; A \\
\neg(A \wedge B) \;\vdash\; \neg A \vee \neg B & \neg(A \vee B) \;\vdash\; \neg A \wedge \neg B \\
\neg(A \Rightarrow B) \;\vdash\; A \wedge \neg B & A \Rightarrow B \;\vdash\; \neg A \vee B \\
A \Leftrightarrow B \;\vdash\; (A \Rightarrow B) \wedge (B \Rightarrow A) & (A \Rightarrow B) \wedge (B \Rightarrow A) \;\vdash\; A \Leftrightarrow B \\
A, B \;\vdash\; A \wedge B & A \vee B \;\vdash\; \neg A \Rightarrow B.
\end{array}
$$

Ponadto przyjmiemy następujące reguły wnioskowania związane z kwantyfikatorami

$$\begin{array}{rcll} \neg\forall x\,\varphi(x) & \vdash & \exists x\,\neg\varphi(x) & \\ \neg\exists x\,\varphi(x) & \vdash & \forall x\,\neg\varphi(x) & \\ \forall x\,\varphi(x) & \vdash & \varphi(a) & \text{dla dowolnej nazwy } a \\ \exists x\,\varphi(x) & \vdash & \varphi(a) & \text{dla dowolnej jeszcze nie używanej (nowej) nazwy } a \\ \varphi(z) & \vdash & \forall x\,\varphi(x) & \text{jeśli } \varphi(z) \text{ zostało udowodnione bez żadnych założeń o } z \\ \varphi(z) & \vdash & \exists x\,\varphi(x) & \text{jeśli o } z \text{ pokazano wcześniej, że istnieje} \end{array}$$

Ponieważ naszym celem jest analiza matematyczna, a nie teoria mnogości, jako zbiór aksjomatów przyjmiemy wszystkie intuicyjnie oczywiste aksjomaty i twierdzenia teorii mnogości, odnośnie szczegółów odsyłając czytelnika do podręczników teorii mnogości.

W codziennej praktyce matematycznej wypisywanie kompletnych formalnych dowodów jest niewykonalne. Najczęściej zamiast kompletnych dowodów podajemy mniej lub bardziej dokładne szkice dowodów, opierając się o różnego rodzaju schematy dowodowe. Schemat dowodowy to pewien sposób rozumowania, o którym wiadomo, że może zostać zastąpiony formalnym dowodem. Do najważniejszych należy zaliczyć

Dowód wprost. Aby pokazać formułę A wystarczy znaleźć ciąg formuł zakończony formułą A taki, że każda formuła jest aksjomatem, udowodnionym twierdzeniem lub wynika z formuł wcześniejszych w ciągu na mocy reguł wnioskowania

Dowód implikacji. Jeśli formuła ma postać $A \Rightarrow B$ wystarczy znaleźć ciąg formuł rozpoczynający się formułą A u zakończony formułą B taki, że każda formuła jest aksjomatem, udowodnionym twierdzeniem lub wynika z formuł wcześniejszych w ciągu na mocy reguł wnioskowania

Dowód przez przypadki. Jeśli formuła C posiada wywód z przesłanki A oraz wywód z przesłanki B, a formuła $A \vee B$ posiada dowód, to formuła C posiada dowód.

Dowód formuły $\forall x \in A\,\varphi(x)$. Aby udowodnić taką formułę wystarczy podać wywód formuły $\varphi(x)$ z przesłanki $x \in A$.

2.3 Zbiory

Dla zadanej formy zdaniowej $\varphi(x)$ naturalna jest chęć utworzenia zbioru takich elementów u, dla których $\varphi(u)$ zachodzi. Innymi słowy, tak zwane *wyrażenie klasowe*, czyli wyrażenie postaci

$$\{\, x \mid \varphi(x) \,\} \text{ lub } \{\, x : \varphi(x) \,\}$$

chcielibyśmy móc uważać za zbiór.

Jako przykład, na razie nieformalny, niech nam posłuży forma zdaniowa

$$x \textit{ jest parą liczb } (x_1, x_2) \textit{ takich, że } x_1^2 + x_2^2 <= 1.$$

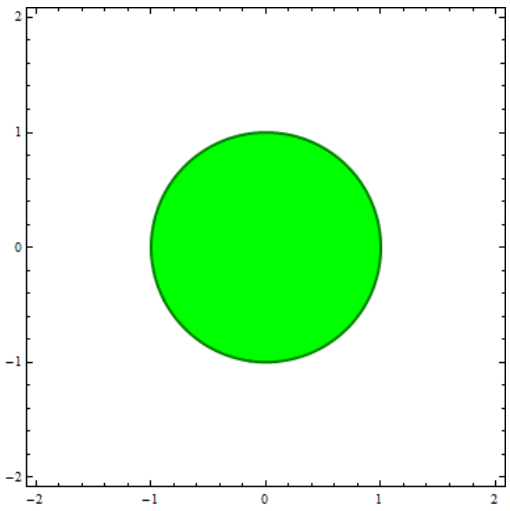
Zaletą takiego przykładu jest, że związany z tą formą zbiór

$$\{\, x \mid x = (x_1, x_2) \textit{ i } x_1^2 + x_2^2 <= 1 \,\}$$

łatwo zwizualizować rysunkiem na płaszczyźnie (rys. 2.4).

Niestety bezkrytyczne tworzenie zbiorów związanych z formami zdaniowymi czasami może prowadzić do sprzeczności. Stosowny dobór aksjomatów teorii mnogości gwarantuje jednak, że w praktycznie wszystkich rozsądnych przypadkach jest to rzeczywiście zbiór. W szczególności

$$\varnothing := \{\, x \mid x \neq x \,\}$$



Rysunek 2.4: Zbiór $\{\, x \mid x = (x_1, x_2) \text{ i } x_1^2 + x_2^2 <= 1 \,\}$ zwizualizowany w układzie współrzędnych na płaszczyźnie.

jest zbiorem, zwanym *zbiorem pustym*. Natomiast

$$\mathbb{V} := \{\, x \mid x = x \,\}$$

nie jest zbiorem, lecz tak zwaną *klasą pełną*.

Pojęcie klasy, ogólniejsze od pojęcia zbioru, pozwala operować (w ograniczony sposób) wyrażeniem klasowym bez wywoływania sprzeczności. Nie będziemy jednak zagłębiać się w te subtelności, pozostawiając je teoretykom teorii mnogości.

2.3.1 Operacje mnogościowe

Posługując się wyrażeniem klasowym, zdefiniujemy teraz kilka podstawowych operacji, które stosujemy do zbiorów i klas.

Definicja 2.3.1 Niech A, B będą zbiorami lub klasami. Definiujemy ich *sumę* (*unię*) oraz *iloczyn* (*przecięcie*), *różnicę* i *różnicę symetryczną* odpowiednio jako

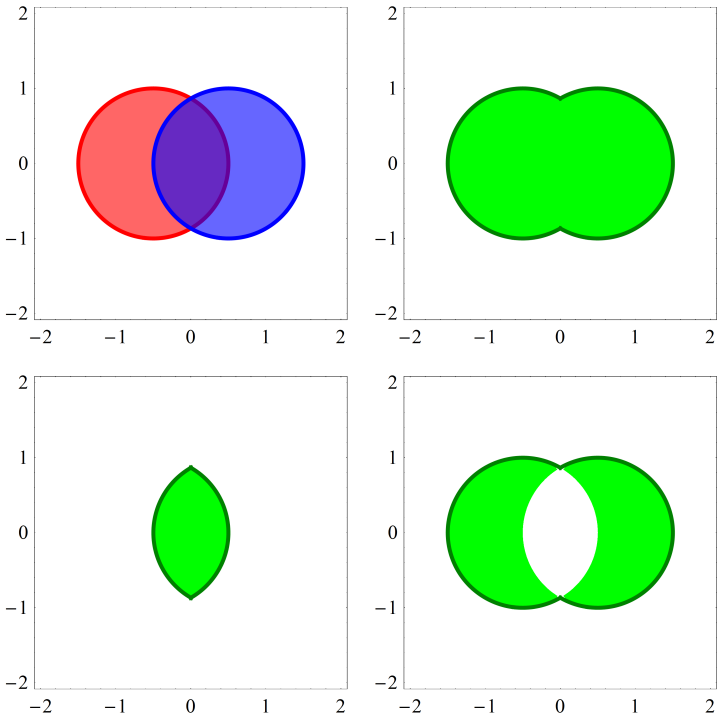
$$\begin{aligned} A \cup B &:= \{\, x \mid x \in A \vee x \in B \,\}, \\ A \cap B &:= \{\, x \mid x \in A \wedge x \in B \,\}, \\ A \setminus B &:= \{\, x \mid x \in A \wedge x \notin B \,\}, \\ A \ominus B &:= (A \setminus B) \cup (B \setminus A). \end{aligned}$$

Rozważmy zbiory na płaszczyźnie

$$A = \{\, (x_1, x_2) \mid (x_1 + 0.5)^2 + x_2^2 <= 1 \,\} \quad B = \{\, (x_1, x_2) \mid (x_1 - 0.5)^2 + x_2^2 <= 1 \,\}.$$

Łatwo sprawdzamy, że

$$\begin{aligned} A \cup B &= \{\, (x_1, x_2) \mid (x_1 + 0.5)^2 + x_2^2 <= 1 \textit{ lub } (x_1 - 0.5)^2 + x_2^2 <= 1 \,\}, \\ A \cap B &= \{\, (x_1, x_2) \mid (x_1 + 0.5)^2 + x_2^2 <= 1 \textit{ i } (x_1 - 0.5)^2 + x_2^2 <= 1 \,\}, \\ A \ominus B &= \{\, (x_1, x_2) \mid \big((x_1 + 0.5)^2 + x_2^2 <= 1 \textit{ i } (x_1 - 0.5)^2 + x_2^2 > 1\big) \textit{ lub } \big((x_1 + 0.5)^2 + x_2^2 > 11 \textit{ i } (x_1 - 0.5)^2 + x_2^2 <= 1\big) \,\}. \end{aligned}$$



Rysunek 2.5: Wierszami od lewego górnego rogu: zbiory A i B, zbiór $A \cup B$, zbiór $A \cap B$ i zbiór $A \ominus B$.

Wizualizację tych zbiorów na płaszczyźnie przedstawiono na rys. 2.5).

Mówimy, że A, B są *rozłączne* jeżeli $A \cap B = \varnothing$. Zbiór

$$\{a\} := \{\, x \mid x = a \,\}$$

ma dokładnie jeden element, mianowicie a. Zbiór taki to *singleton*. Podobnie definiujemy *dubleton*

$$\{a, b\} := \{\, x \mid x = a \vee x = b \,\}.$$

W zasadzie definicję tę stosujemy tylko wtedy gdy $a \neq b$, gdyż w przeciwnym razie dubleton pokrywa się z singletonem. Piszemy też

$$\{a_1, a_2 \ldots a_n\} := \{\, x \mid x = a_1 \vee x = a_2 \ldots a_n \,\}.$$

2.3.2 Podzbiory

Definicja 2.3.2 Mówimy, że zbiór A *zawiera się* w zbiorze B lub A jest *podzbiorem* B jeżeli $A \subset B$ co, na mocy naszej definicji, oznacza, że

$$x \in A \;\Rightarrow\; x \in B.$$

Mówimy, że A jest *podzbiorem właściwym* B jeżeli $A \subsetneq B$, co znow na mocy definicji oznacza, że $A \subset B$ oraz $A \neq B$.

Definicja 2.3.3 *Zbiór podzbiorów* zbioru A definiujemy jako zbiór postaci

$$\mathcal{P}(A) := \{\, U \mid U \subset A \,\}.$$

Definicja 2.3.4 Jeśli $A \subset X$ to zbiór $X \setminus A$ nazywamy *uzupełnieniem* A do X i oznaczamy $C_X A$. Gdy X jest ustalony i znany z kontekstu, to często nazywamy go *przestrzenią* a uzupełnienie A do X oznaczamy po prostu CA lub $\setminus A$.

Rysunki 2.4 oraz 2.5 uzyskano przy pomocy pakietu oprogramowania `PSets` napisanego w języku Mathematica. Pakiet ten pozwala na definiowanie zbiorów na płaszczyźnie zadanych formami zdaniowymi zbudowanymi z nierówności algebraicznych. Aby móc korzystać z pakietu, trzeba go najpierw wczytać instrukcją

```
<< PSets.m
```

Teraz wprowadzając

```
sA = PSet[{x, y}, (x + 0.5)^2 + y^2 <= 1];
sB = PSet[{x, y}, (x - 0.5)^2 + y^2 <= 1];
Print[sA \[Union] sB];
Print[sA \[Intersection] sB];
Print[sA \[Backslash] sB];
Print[sA \[CirlceMinus] sB];
```

obliczymy sumę, iloczyn, różnicę i różnicę symetryczną zbiorów

$$\begin{aligned} A: &= \{\,(x,y) \mid (x+0.5)^2 + y^2 \leqslant 1\,\} \\ B: &= \{\,(x,y) \mid (x-0.5)^2 + y^2 \leqslant 1\,\}. \end{aligned}$$

Możemy też sprawdzić, czy konkretny punkt, na przykład $(1,0)$ należy do zbioru, na przykład A, pisząc

```
{1,0} \[Element] sA
```

Użycie operatorów `\back[Subset]`, `\back[Superset]` i `\back[equal]` pozwala na sprawdzanie zawierania się i równości zbiorów. Warto wiedzieć, że różnego rodzaju operatory, w tym operatory mnogościowe w Mathematica można wprowadzać używając skrótów klawiaturowych. Wybrane skróty przedstawiono w tabeli poniżej. W tabeli tej znakiem ⁞ oznaczono przyciśnięcie klawisza `Esc` w lewym górnym rogu klawiatury.

operator	skrót	forma graficzna
\[Union]	⁞un⁞	$\cup$
\[Intersection]	⁞inter⁞	$\cap$
\[Backslash]	⁞\⁞	$\setminus$
\[CirlceMinus]	⁞c-⁞	$\ominus$
\[Subset]	⁞sub⁞	$\subset$
\[Superset]	⁞sup⁞	$\supset$
\[Equal]	⁞==⁞	$=$
\[Element]	⁞el⁞	$\in$
\[SmallCirlce]	⁞sc⁞	$\circ$
\[CenterDot]	⁞.⁞	$\cdot$
\[CircleDot]	⁞c.⁞	$\odot$
\[CirclePlus]	⁞c+⁞	$\oplus$
\[Cross]	⁞cross⁞	$\times$
\[Times]	⁞*⁞	$\times$

Pisząc

```
SetPlot[(sA \[Union] sB), {-2, 2}, {-2, 2}]
```

otrzymamy rysunek zbioru $A \cup B$ w domyślnych kolorach. Pisząc

```
SetPlot[(sA \[Backslash] sB), {-2, 2}, {-2, 2}, InteriorStyle -> Green,
  BoundaryStyle -> {RGBColor[0, 0.5, .0], Thick}]
```

uzyskamy rysunek zbioru $A \setminus B$ w kolorze i stylu dobranym przez nas. Pisząc

```
SetPlot[(sA \[CirlceMinus] sB), {-2, 2}, {-2, 2}, Grid->True]
```

uzyskamy rysunek zbioru $A \ominus B$ z naniesioną siatką poziomych i pionowych linii o współrzędnych całkowitych. Rysunek taki pozwala w szczególności szybko policzyć ilość zawartych w zbiorze tzw. punktów kratowych, tj. punktów o obu współrzędnych całkowitych.

Ćwiczenie komputerowe 2.3.5 Wykorzystując pakiet Sets oblicz i narysuj zbiory A, B, $A \cup B$, $A \cap B$, $A \setminus B$, $B \setminus A$ i $A \ominus B$ dla różnych, zaproponowanych przez Ciebie podzbiorów płaszczyzny. Eksperymentując, ustal znacznie parametrów instrukcji SetPlot. Zastanów się, czy otrzymane rysunki można zawsze uznać za w pełni poprawne.

Czasami warto kilka zbiorów umieścić razem na jednym wykresie (rysunku). Pisząc

```
pa = SetPlot[sA, {-2, 2}, {-2, 2},
  InteriorStyle -> {Red, Opacity[0.6]},
  BoundaryStyle -> {Red, Thick}];
pb = SetPlot[sB, {-2, 2}, {-2, 2},
  InteriorStyle -> {Blue, Opacity[0.6]},
  BoundaryStyle -> {Blue, Thick}];
Show[pa,pb]
```

uzyskamy oba zbiory A i B na jednym wykresie, tak jak to przedstawiono w lewym gónym rogu rys. 2.5.

Ćwiczenie komputerowe 2.3.6 Wykorzystaj instrukcję Show, by zilustrować: rozłączność zbiorów, zawieranie się zbiorów jak również brak rozłączności i brak zawierania się.

Możemy również eksperymentować z pakietem `FSets`, który implementuje operacje mnogościowe na zbiorach skończonych. Na przykład pisząc

```
sA=FSet[a,b,c]
sB=FSet[b,c,d]
sC=sA\[Union]sB
```

obliczymy, że sumą mnogościową zbiorów a, b, c i b, c, d jest a, b, c, d.

2.3.3 Elementarne własności operacji mnogościowych

Uwaga 2.3.7 *Niech A, B, C będą zbiorami. Operacje mnogościowe mają następujące własności.*

$$A \subset B,\ B \subset C \ \Rightarrow\ A \subset C,$$
$$A \setminus A = \varnothing, \quad A \setminus \varnothing = A, \quad A \cup A = A, \quad A \cap A = A,$$
$$A \cup B = B \cup A, \quad A \cap B = B \cap A,$$
$$A \subset B \ \Leftrightarrow\ A \cup B = B \ \Leftrightarrow\ A \cap B = A,$$
$$A \subset A \cup B, \quad B \subset A \cup B, \quad A \cap B \subset A, \quad A \cap B \subset B,$$
$$A \cup (B \cup C) = (A \cup B) \cup C, \quad A \cap (B \cap C) = (A \cap B) \cap C,$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Uwaga 2.3.8 *Niech $A, B \subset X$ będą zbiorami w przestrzeni X. Uzupełnienie do przestrzeni X ma następujące własności.*

$$A \setminus B = A \cap (\setminus B),$$
$$\setminus(\setminus A) = A, \quad \setminus(A \cup B) = (\setminus A) \cap (\setminus B),$$
$$\setminus(A \cap B) = (\setminus A) \cup (\setminus B),$$
$$A \subset B \;\Leftrightarrow\; \setminus A \supset \setminus B.$$

Ćwiczenie komputerowe 2.3.9 Wykorzystaj pakiet FSets, by zilustrować wybrane elementarne własności operacji mnogościowych.

2.3.4 Sumy i iloczyny nieskończone

Definicja 2.3.10 Dla zbioru lub klasy $\mathcal{A}$ definiujemy

$$\bigcup \mathcal{A} := \{\, x \mid \exists A \in \mathcal{A}\; x \in A \,\},$$
$$\bigcap \mathcal{A} := \{\, x \mid \forall A \in \mathcal{A}\; x \in A \,\}.$$

Nieformalnie o $\mathcal{A}$ mówimy wtedy, że jest *rodziną* zbiorów, a więc zbiorem lub klasą, której elementy są zbiorami.

Twierdzenie 2.3.11 *Zachodzą następujące równości:*

$$\bigcup \varnothing = \varnothing, \quad \bigcap \varnothing = \mathbb{V}, \quad \bigcup \mathbb{V} = \mathbb{V}, \quad \bigcap \mathbb{V} = \varnothing.$$

2.4 Iloczyn kartezjański

Wśród kluczowych pojęć analizy matematycznej jest pojęcie ciągu i funkcji. Funkcję najprościej zdefiniować jak pewien zbiór par. Pierwszy element pary to argument, a drugi to wartość funkcji dla tego argumentu. Zatem jeśli chcemy definicję funkcji podać na gruncie teorii mnogości, musimy zdefiniować parę elementów a i b, zwyczajowo oznaczaną (a, b). W pojęciu pary istotne jest, by było wiadomo, który element pary jest pierwszy, a który drugi. Tak więc roli pary nie może pełnić dubleton, bo w dubletonie kolejność nie ma znaczenia. Mamy bowiem $\{a, b\} = \{b, a\}$. Nie jest oczywiste jak postawić definicję pary na gruncie teorii mnogości. Nie możemy zdefiniować pary jako dwuelementowego ciągu, gdyż ciąg jest szczególnym przypadkiem funkcji, a na gruncie teorii mnogości funkcję definiuje się jako pewien zbiór par (x, y) o stosownych własnościach.

2.4.1 Para

Parą elementów a i b nazywamy zbiór $\{\{a, b\}, \{b\}\}$ oznaczany krótko (a, b).

Powyższą, powszechnie dziś używaną definicję pary podał w 1921 roku polski matematyk Kazimierz Kuratowski (rys. 2.6). Wcześniejsze, bardziej skomplikowane definicje się nie przyjęły.

Kluczową własnością pary, której nie ma dubleton, jest

Uwaga 2.4.1

$$(a, b) = (c, d) \;\Leftrightarrow\; a = c \wedge b = d.$$



Rysunek 2.6: Kazimierz Kuratowski, 1896-1980

W poniższym dowodzie trzykrotnie zastosowano schemat dowodowy zwany dowodem przez przypadki. W dowodzie przez przypadki zamiast jednego dowodu przedstawiamy dwa lub więcej dowodów, każdy z pewnym dodatkowym nieudowodnionym założeniem. Dowód jest kompletny, jeśli dodatkowo potrafimy udowodnić, że przynajmniej jedno z tych dodatkowych założeń musi być prawdziwe. Ten ostatni fakt często sprawdzić jest bardzo łatwo, bo rozważane przypadki wzajemnie się dopełniają: na przykład jeden przypadek to $x > 0$, a drugi $x \leqslant 0$. Zaletą takiego podejścia jest możliwość odmiennego poprowadzenia dowodu w każdym z przypadków.

Dowód: Choć własność wydaje się oczywista, to jednak jej w miarę formalny dowód jest pracochłonny. Aby udowodnić, że

$$(a, b) = (c, d) \;\Leftrightarrow\; a = c \wedge b = d.$$

wystarczy podać oddzielnie dowód implikacji

$$(a, b) = (c, d) \;\Rightarrow\; a = c \wedge b = d$$

oraz implikacji

$$a = c \wedge b = d \;\Rightarrow\; (a, b) = (c, d).$$

Podamy dowód pierwszej z tych implikacji. Dowód drugiej jest znacznie łatwiejszy i zostawiamy go jako ćwiczenie.

Załóżmy zatem, że $(a, b) = (c, d)$, co zgodnie z naszą definicją pary oznacza, że $\{\{a, b\}, \{b\}\} = \{\{c, d\}, \{d\}\}$. Ponieważ $\{b\}$ jest elementem zbioru $\{\{a, b\}, \{b\}\}$, więc jest on też elementem zbioru $= \{\{c, d\}, \{d\}\}$. Zatem albo $\{b\} = \{\{c, d\}\}$ albo $\{b\} = \{d\}$. Dalszą część dowodu możemy zatem poprowadzić przez przypadki.

Załóżmy najpierw, że $\{b\} = \{\{c, d\}\}$. Ponieważ tak c jak i d należą do dubletona $\{\{c, d\}\}$, a dubleton $\{\{c, d\}\}$ jest w tym przypadku równy singletonowi $\{b\}$, więc c i d należą do singletona $\{b\}$. Zgodnie z definicją singletona jest to możliwe tylko wtedy gdy $c = b$ i $d = b$. Zatem również $c = d$ co oznacza w szczególności, że dubleton $\{c, d\}$ pokrywa się z singletonem $\{d\}$, a w konsekwencji mamy

$$(c, d) = \{\{c, d\}, \{d\}\} = \{\{d\}, \{d\}\} = \{\{d\}\}.$$

Zatem $\{\{a, b\}, \{b\}\} = \{\{d\}\}$. Ponieważ $\{a, b\}$ należy do lewej strony tej równości, więc należy też do prawej strony. Stąd $\{a, b\} = \{d\}$. To oznacza, że $d = a$ i $d = b$. Pokazaliśmy zatem, że $a = b = c = d$, więc w szczególności $a = c$ i $b = d$. Kończy to dowód w przypadku gdy $\{b\} = \{\{c, d\}\}$.



Rysunek 2.7: Kartezjusz, 1596-1650

By rozważyć drugi przypadek przyjmijmy, że $\{b\} = \{d\}$. Wtedy $b = d$ i pozostaje pokazać, że $a = c$. Z $b = d$ wynika, że $(a, b) = (c, b)$. Ponieważ $\{a, b\} \in (a, b)$, więc również $\{a, b\} \in (c, b)$. Mamy zatem do rozpatrzenia dwa podprzypadki: albo $\{a, b\} = \{c, b\}$ albo $\{a, b\} = \{b\}$. W pierwszym podprzypadku dostajemy $a = c$ lub $a = b$. Podpodprzypadek $a = c$ nie wymaga już dalszego rozumowania, bo chcieliśmy właśnie pokazać, że $a = c$. Natomiast jeśli $a = b$, to $\{a\} = \{c, b\}$, skąd również wynika, że $a = c$.

Pozostaje jeszcze rozważyć podprzypadek $\{a, b\} = \{b\}$. Wtedy $a = b$, skąd $(a, b) = \{\{a\}\}$. Zatem $\{c, d\} \in \{\{a\}\}$, skąd $a = c = d$. Tak więc w tym podprzypadku znów dostajemy, że wszsytkie cztery elementy są równe, skąd w szczególności $a = c$ i $b = d$. □

2.4.2 Iloczyn kartezjański

Zbiór wszystkich możliwych par (x, y), gdzie $x \in X$, a $y \in Y$ tworzy tzw.iloczyn kartezjański zbiorów X i Y. Nazwa wywodzi się od francuskiego matematyka Kartezjusza (René Descartes, rys. 2.7), który jako pierwszy zaproponował algebraiczne podejście do geometrii poprzez potraktowanie płaszczyzny jako zbioru par liczb.

Iloczynem kartezjańskim zbiorów X i Y nazywamy zbiór

$$X \times Y := \{\, (x, y) \mid x \in X, y \in Y \,\}.$$

Iloczyn kartezjański $X \times X$ oznaczamy krótko X^2.

Twierdzenie 2.4.2 *Iloczyn kartezjański ma następujące własności.*

(i) Jeśli $A \neq \varnothing \neq B$, *to*

$$A \times B \subset X \times Y \;\Leftrightarrow\; A \subset X \text{ i } B \subset Y,$$

(ii) Dla dowolnych zbiorów A_1, A_2, B *mamy*

$$\begin{aligned}(A_1 \times B) \cup (A_2 \times B) &= (A_1 \cup A_2) \times B,\\ (A_1 \times B) \cap (A_2 \times B) &= (A_1 \cap A_2) \times B,\end{aligned}$$

(iii) Dla dowolnych zbiorów A, B, C, D *mamy*

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D).$$

W pakiecie `FSets` iloczyn kartezjański zbiorów skończonych obliczamy instrukcją

```
sC=sA \[Cross] sB
```

2.5 Relacje

Pojęcie relacji znane jest nam dobrze z języka codziennego. Mówi się na przykład, że dwie osoby pozostają w relacji pokrewieństwa. Albo, że dwa państwa pozostają w dobrych relacjach. Relacja oznacza pewien związek pomiędzy dwoma elementami. Z abstrakcyjnego punktu widzenia relacja pomiędzy elementami zbiorów X *i* Y *to w pewien sposób wyróżniony podzbiór iloczynu kartezjańskiego* $X \times Y$. *Wyróżnienie może nastąpić poprzez podanie warunku wymaganego, by para była w relacji, ale w przypadku zbiorów skończonych często ograniczamy się do wypisania par pozostających w relacji bez merytorycznego wnikania co te pary połączyło.*

Pojęcie relacji jest punktem wyjścia do szeregu podstawowych pojęć współczesnej matematyki: relacji równoważności, relacji częściowego i liniowego porządku. Funkcje to też szczególny przypadek relacji.

2.5.1 Relacje

Relacją w zbiorach X i Y nazywamy dowolny podzbiór $R \subset X \times Y$. Gdy $X = Y$, mówimy krótko o relacji w X. Tradycyjnie piszemy xRy zamiast $(x, y) \in R$. Zbiór pusty i $X \times X$ to odpowiednio *relacja pusta* i *relacja pełna* w X. *Relacją identycznościową* w X jest

$$\Delta_X := \mathrm{id}_X := \{\, (x, x) \mid x \in X \,\}.$$

Rozważmy relację w zbiorze $X := \{\, 1, 2, 3, 4 \,\}$ daną wzorem

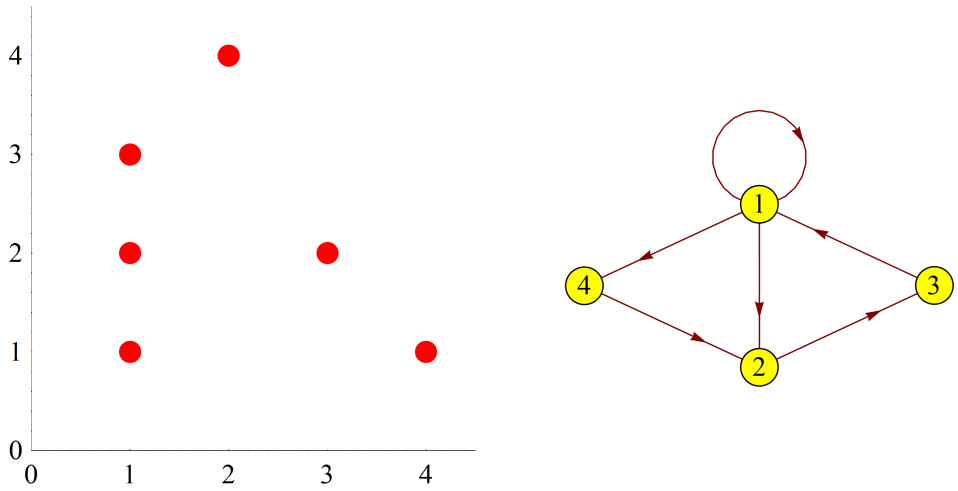
$$R := \{\, (1,1), (1,2), (1,4), (2,3), (3,1), (4,2) \,\}. \tag{2.1}$$

Relację tę można zobrazować rysując jej wykres oraz graf (rys. 2.8). Wykres relacji rysujemy podobnie jak wykres funkcji, zaznaczając w układzie współrzędnych te punkty, których współrzędne odpowiadają parom pozostającym w relacji. Rysowanie wykresu relacji jest możliwe tylko wtedy gdy zbiory X *i* Y *są podzbiorami prostej lub w jakiś sposób dają się utożsamić z pewnymi podzbiorami prostej. Same zbiory* X *i* Y *nie muszą być skończone.*

Narysowanie grafu relacji jest możliwe tylko wtedy gdy zbiór X *(oba zbiory* X *i* Y *gdy relacja jest w zbiorach* X *i* Y*) są skończone. Każdy z elementów zbioru* X *(lub zbiorów* X *i* Y*) zaznaczamy jako punkt na płaszczyźnie. Punkty łączymy strzałką, gdy odpowiadające im elementy pozostają w relacji.*

Pakiet `Relations` w języku Mathematica pozwala na rysowanie wykresów i grafów relacji w zbiorze skończonym X. Na przykład instrukcja

```
R = Relation[{1,2,3,4},{1->1,1->2,1->4,2->3,3->1,4->2}]
```



Rysunek 2.8: Wykres i graf relacji R w X

tworzy relację w zbiorze $X := \{1, 2, 3, 4\}$ daną wzorem (2.1). Rysunek wykresu tej relacji uzyskamy pisząc

```
RChart[re]
```

Natomiast graf uzyskamy pisząc

```
RGraph[re]
```

Aby sprawdzić, czy dana para (a, b) jest w relacji R wystarczy napisać

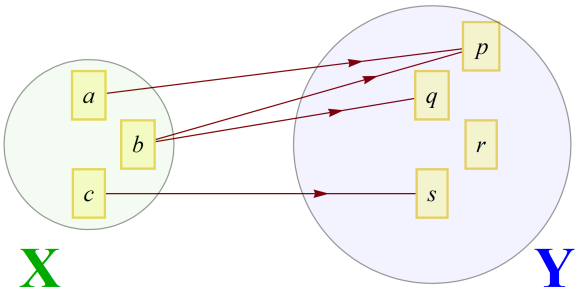
```
a \[CenterDot] R \[CenterDot] b
```

Rozważmy relację w $X := \{a, b, c\}$ i $Y := \{p, q, r, s\}$ daną wzorem

$$R := \{(a, p), (b, p), (b, q), (c, s)\}.$$

Relację tę można zobrazować rysując jej graf (rys. 2.9). W przypadku relacji w zbiorach X i Y gdy $X \neq Y$ rysowanie grafu jest bardziej skomplikowane, bo nie można się zdać na automatyczne rozlokowanie wierzchołków. Rysunek 2.9 uzykano następującym ciągiem instrukcji

```
sx = 0; sy = 7;
g1 = GraphPlot[{a -> p, b -> p, b -> q, c -> s, r -> r},
   VertexCoordinateRules -> {a -> {sx, 1}, b -> {sx + 1, 0},
   c -> {sx, -1}, p -> {sy + 1, 2}, q -> {sy, 1},
   r -> {sy + 1, 0}, s -> {sy, -1}},
   VertexLabeling -> True, DirectedEdges -> True,
   SelfLoopStyle -> False, PlotRange -> {{-2, 12}, {-4, 4}}];
g2 = RegionPlot[x^2 + y^2 <= 3, {x, -2, 12}, {y, -4, 4},
   PlotRange -> {{-2, 12}, {-4, 4}}, AspectRatio -> Automatic,
   PlotStyle -> {Opacity[0.05], Darker[Green]}];
g3 = RegionPlot[(x - 7)^2 + y^2 <= 8, {x, -2, 12}, {y, -4, 4},
   PlotRange -> {{-2, 12}, {-4, 4}}, AspectRatio -> Automatic,
   PlotStyle -> {Opacity[0.05], Blue}];
g4 = Graphics[{
   Text[Style["X", Darker[Green], Bold, 24], {-1, -2.5}],
```



Rysunek 2.9: Graf relacji R w X i Y

```
    Text[Style["Y", Blue, Bold, 24], {sy + 2.5, -2.5}]},
    PlotRange -> {{-2, 12}, {-4, 4}}];
u = Show[g1, g2, g3, g4]
```

2.5.2 Własności relacji

Niech $R \subset X \times X$ będzie relacją w X.

Definicja 2.5.1 Wyróżniamy następujące własności relacji R.

$$\begin{aligned}
R \text{ zwrotna} &:\Leftrightarrow \quad \forall x \in X \quad x \,\mathrm{R}\, x,\\
R \text{ symetryczna} &:\Leftrightarrow \quad \forall x, y \in X \quad x \,\mathrm{R}\, y \;\Rightarrow\; y \,\mathrm{R}\, x,\\
R \text{ antysymetryczna} &:\Leftrightarrow \quad \forall x, y \in X \quad x \,\mathrm{R}\, y \wedge y \,\mathrm{R}\, x \;\Rightarrow\; x = y,\\
R \text{ przechodnia} &:\Leftrightarrow \quad \forall x, y, z \in X \quad x \,\mathrm{R}\, y \wedge y \,\mathrm{R}\, z \;\Rightarrow\; x \,\mathrm{R}\, z,\\
R \text{ spójna} &:\Leftrightarrow \quad \forall x, y \in X \quad x \,\mathrm{R}\, y \vee y \,\mathrm{R}\, x.
\end{aligned}$$

Pakiet `Relations` umożliwia proste sprawdzanie, czy dana relacja ma powyższe własności. Wprowadzając

```
IsReflexive[re]
IsSymmetric[re]
IsAntisymmetric[re]
IsTransitive[re]
IsTotal[re]
```

dowiadujemy się, że relacja zadana wzorem (2.1) jest antysymetryczna, ale pozostałych rozważanych własności nie posiada.

Ćwiczenie komputerowe 2.5.2 Dla każdej z powyższych pięciu własności podaj przykład relacji w zbiorze pięcioelementowym, która ma dokładnie tę i, o ile to możliwe, żadną inną z powyższych własności. Przykład zilustruj rysunkiem wykresu i grafu opracowanym w programie Mathematica. Zastanów się, które z powyższych własności łatwo sprawdzić oglądając wykres, a które oglądając graf.

2.5.3 Złożenie relacji

Relacje, podobnie jak funkcje można składać. Można też brać relację odwrotną do dowolnej relacji.

Definicja 2.5.3 Dla relacji R w $X \times Y$ i relacji S w $Y \times Z$ definiujemy *obłożenie relacji R przez relację S* zwane również *podstawieniem relacji R w relacji S* jako

$$S \circ R := \{\, (x,z) \in X \times Z \mid \exists y \in Y \; xRy \text{ i } ySz \,\}.$$

Definiujemy również *relację odwrotną do relacji R* jako

$$R^{-1} := \{\, (y,x) \in Y \times X \mid xRy \,\}.$$

Twierdzenie 2.5.4 *Niech Q będzie relacją w $W \times X$, R relacją w $X \times Y$, a S relacją w $Y \times Z$. Wtedy*

$$(S \circ R) \circ Q = S \circ (R \circ Q),$$
$$(S \circ R)^{-1} = R^{-1} \circ S^{-1}, \quad (R^{-1})^{-1} = R,$$
$$\Delta_X^{-1} = \Delta_X.$$

Pakiet `Relations.m` umożliwia obliczanie złożenia relacji przy użyciu operatora `\[SmallCirlce]`.

2.5.4 Charakterystyka własności relacji

Następujące twierdzenie opisuje własności relacji w języku zbiorów.

Twierdzenie 2.5.5 *Niech $R \subset X \times X$ będzie relacją. Wtedy:*

$$\begin{aligned} R \text{ zwrotna} &\Leftrightarrow \Delta_X \subset R, \\ R \text{ symetryczna} &\Leftrightarrow R^{-1} = R, \\ R \text{ anysymetryczna} &\Leftrightarrow R \cap R^{-1} \subset \Delta_X, \\ R \text{ przechodnia} &\Leftrightarrow R \circ R \subset R, \\ R \text{ spójna} &\Leftrightarrow R \cup R^{-1} = X \times X. \end{aligned}$$

2.6 Relacje równoważności

Relacja równoważności dokonuje utożsamienia obiektów pod pewnym względem przy zachowaniu różnic pod innym względem. Relacją równoważności jest na przykład bycie tej samej narodowości. Dwie osoby tej samej narodowości mogą się różnić pod wieloma względami, ale pod względem narodowości są identyczne.

W matematyce relacje równoważności służą jako narzędzie do definiowania nowych pojęć. Posłużymy się nim by z liczb naturalnych skonstruować liczby całkowite, z liczb całkowitych liczby wymierne, a z liczb wymiernych liczby rzeczywiste.

2.6.1 Relacja równoważności

Definicja 2.6.1 Relację w klasie X, która jest zwrotna, symetryczna i przechodnia nazywamy *relacją równoważności*.

Dla $x \in X$ klasę

$$[x]_R := \{\, y \in X \mid xRy \,\}$$

nazywamy *klasą równoważności* elementu x.

Twierdzenie 2.6.2 *Niech R będzie relacją równoważności w X. Jeśli $x, y \in X$ oraz xRy, to $[x]_R = [y]_R$. Innymi słowy, klasy równoważności równoważnych elementów są identyczne. Ponadto dla dowolnych $x, y \in X$ albo $[x]_R = [y]_R$, albo $[x]_R \cap [y]_R = \varnothing$, to znaczy klasy równoważności dwóch elementów są albo równe, albo rozłączne.*

Pakiet `Relations.m` umożliwia sprawdzenie czy relacja jest relacją równoważności. Na przykład pisząc

```
rR = Relation[{a, b, c}, {a -> a, b -> b, c -> c, a -> b, b -> a}]
IsEquivalence[rR]
```

otrzymamy odpowiedź `True`.

2.6.2 Równoliczność

Definicja 2.6.3 Zbiory X, Y nazywamy *równolicznymi* jeśli istnieje bijekcja $f : X \to Y$.

Twierdzenie 2.6.4 *Relacja równoliczności zbiorów jest relacją równoważności w klasie pełnej.*

Twierdzenie 2.6.5 *(G. Cantor) Żaden zbiór nie jest równoliczny ze zbiorem swoich podzbiorów.*

2.7 Funkcje

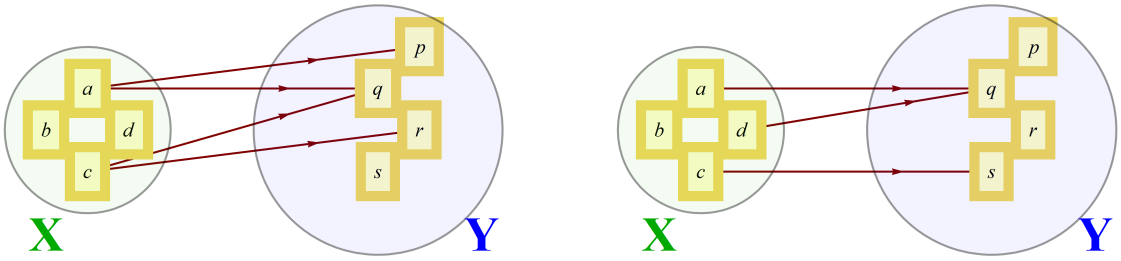
2.7.1 Relacje jednolistne

Definicja 2.7.1 Niech $R \subset X \times Y$ będzie relacją. Mówimy, że R jest *jednolistna* jeżeli

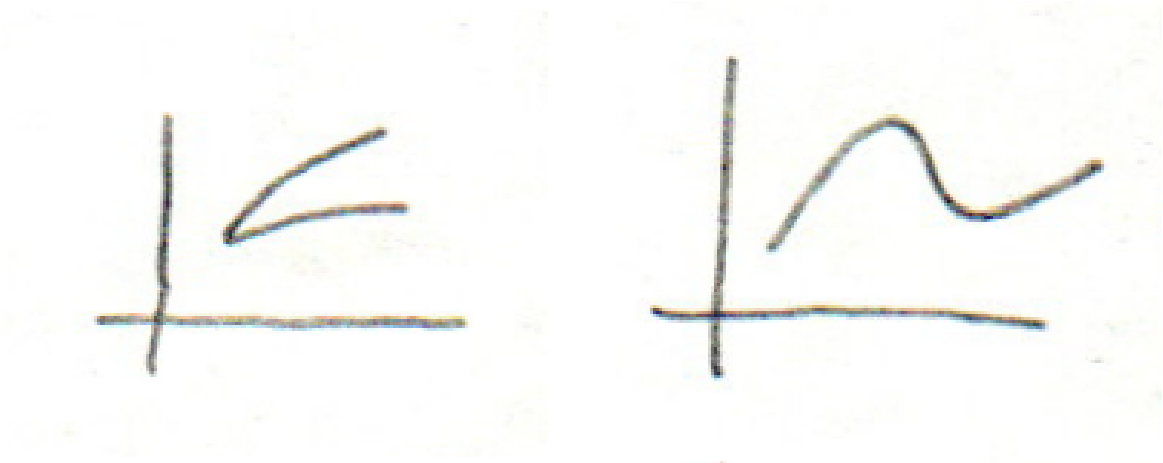
$$(x, y) \in R \wedge (x, y') \in R \;\Rightarrow\; y = y'.$$

Przykładowy graf relacji niejednolistnej (po lewej) i jednolistnej (po prawej) przedstawiono na rys. 2.10. Przykładowy wykres relacji niejednolistnej (po lewej) i jednolistnej (po prawej) przedstawiono na rys. 2.11.

2.7.2 Funkcje częściowe



Rysunek 2.10: Graf relacji niejednolistnej i jednolistnej



Rysunek 2.11: Wykres relacji niejednolistnej i jednolistnej

Definicja 2.7.2 Mówimy, że trójka (X, f, Y) jest *funkcją częściową z X do Y* jeżeli f jest relacją jednolistną w $X \times Y$. Tradycyjnie piszemy $f : X \multimap Y$ na oznaczenie funkcji częściowej z X do Y. Jeśli para (x, y) jest elementem funkcji częściowej f z X do Y, to element y jest wyznaczony jednoznacznie przez element x. Oznaczamy go $f(x) := y$ i nazywamy *wartością funkcji częściowej f w x*. *Dziedzinę* funkcji częściowej f definiujemy jako

$$\operatorname{dom} f := \{\, x \in X \mid \exists y \in Y\ (x, y) \in f \,\}.$$

Obraz funkcji częściowej f definiujemy jako

$$\operatorname{im} f := \{\, y \in Y \mid \exists x \in X\ (x, y) \in f \,\}.$$

Uwaga 2.7.3 *Dwie funkcje częściowe (X_1, f_1, Y_1) i (X_2, f_2, Y_2) są sobie równe wtedy i tylko wtedy gdy $X_1 = X_2$ i $f_1 = f_2$ i $Y_1 = Y_2$.*

Pakiet `Relations.m` umożliwia sprawdzenie czy relacja jest funkcją częściową. Na przykład pisząc

```
rR = Relation[{a, b, c}, {b -> b, c -> c, b -> a}]
IsPartialFunction[rR]
```

otrzymamy odpowiedź `False`.

2.7.3 Funkcje

Definicja 2.7.4 Jeżeli dla funkcji częściowej (X, f, Y) spełniony jest warunek $\operatorname{dom} f = X$, to piszemy

$$f : X \to Y$$

i mówimy, że (X, f, Y) jest *funkcją* z X do Y lub *odwzorowaniem* z X do Y lub *przekształceniem* X w Y. Często, od strony formalnej niepoprawnie, mówimy, że to f jest funkcją, rozumiejąc, że jest jasne o jakie X i Y chodzi.

Pakiet `Relations.m` umożliwia sprawdzenie czy relacja jest funkcją. Na przykład pisząc

```
rR = Relation[{a, b, c}, {a->c, b -> b, c -> a}]
IsFunction[rR]
```

otrzymamy odpowiedź `True`.

Na zbiór funkcji z X do Y stosujemy oznaczenie

$$Y^X := \{\, f \mid f : X \to Y \text{ - funkcja z } X \text{ do } Y \,\}$$

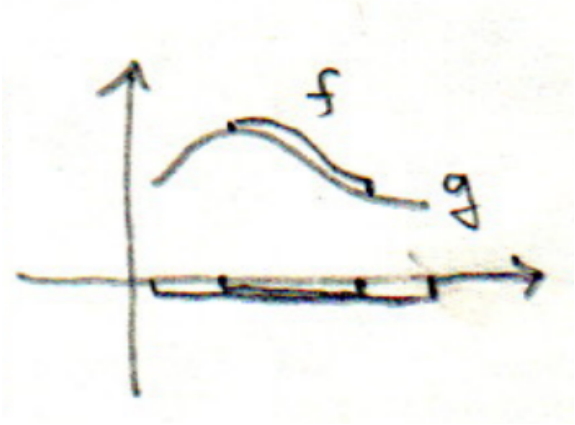
Przykład 2.7.5

1. *Niech $c \in Y$. Wtedy $\{\, (x, c) \mid x \in X \,\}$ jest funkcją z X do Y. Nazywamy ją* funkcją stałą.
2. *Relacja identycznościowa Δ_X jest funkcją $\Delta_X : X \to X$.*
3. *Relacja pełna w przestrzeni X nie jest funkcją z wyjątkiem przypadku gdy X jest pusty lub jest singletonem.*

2.7.4 (*)Zawieranie się funkcji

Uwaga 2.7.6 *Niech $f, g : X \multimap Y$. Wtedy warunek $f \subset g$ jest równoważny koniunkcji warunków*

$$\operatorname{dom} f \subset \operatorname{dom} g \quad i \quad \forall x \in \operatorname{dom} f \quad f(x) = g(x).$$



Rysunek 2.12: Zawieranie funkcji

Przykład zawierania się funkcji przedstawiono na rys. 2.12

Definicja 2.7.7 Niech f, g będą fukcjami częściowymi z X do Y. Mówimy, że f jest *zawężeniem* g lub g jest *rozszerzeniem* f jeżeli $f \subset g$.

Niech $f : X \to Y$, $A \subset X$. *Zawężeniem* funkcji f do A nazywamy funkcję

$$f_{|A} := f \cap A \times Y.$$

2.7.5 Sklejenie funkcji

Uwaga 2.7.8 *Niech f, g będą fukcjami częściowymi z X do Y. Suma $f \cup g$ jest funkcją częściową z X do Y wtedy i tylko wtedy gdy*

$$\forall x \in \operatorname{dom} f \cap \operatorname{dom} g \; f(x) = g(x).$$

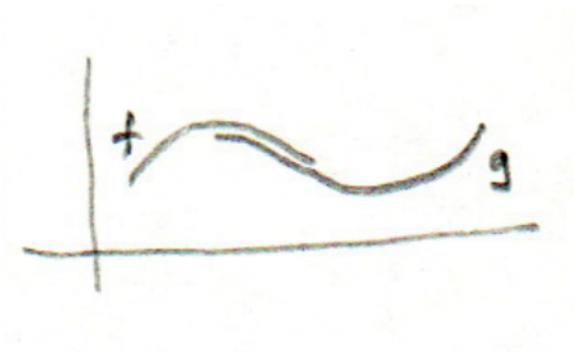
Definicja 2.7.9 Jeśli $f, g : X \multimap Y$ są funkcjami częściowymi z X do Y i $f \cup g$ jest funkcją to nazywamy ją *sklejeniem* funkcji f i g.

Przykładowe sklejenie funkcji przedstawiono na rys. 2.13

2.7.6 (*)Zestawienie i iloczyn kartezjański funkcji

Definicja 2.7.10 Niech $f : X \to Y$ i $g : X \to Z$. *Zestawieniem* funkcji f i g nazywamy funkcję

$$(f, g) : X \ni x \mapsto (f(x), g(x)) \in X \times Y.$$



Rysunek 2.13: Sklejenie funkcji

Definicja 2.7.11 Niech $f : X_1 \to Y_1$, a $g : X_2 \to Y_2$. *Iloczynem kartezjańskim* funkcji f i g nazywamy funkcję

$$f \times g : X_1 \times X_2 \ni (x_1, x_2) \to (f(x_1), g(x_2)) \in Y_1 \times Y_2.$$

Oznaczenie na zestawienie i iloczyn kartezjański funkcji koliduje z parą funkcji i iloczynem kartezjańskim funkcji potraktowanych jako zbiory. Jednak w tym drugim znaczeniu oznaczeń tych praktycznie się nie używa, więc ryzyko błędnej interpretacji jest minimalne.

2.7.7 Obrazy i przeciwobrazy

Niech $f : X \to Y$, $A \subset X$, $B \subset Y$. *Obrazem* zbioru A nazywamy

$$f(A) := \{\, f(x) \mid x \in A \,\}.$$

Przeciwobrazem zbioru B nazywamy

$$f^{-1}(B) := \{\, x \in X \mid f(x) \in B \,\}.$$

Uwaga 2.7.12 *Niech $f : X \to Y$, a $A, A_1, A_2 \subset X$ i $B, B_1, B_2 \subset Y$ będą zbiorami. Wtedy*

$$\begin{gathered}
A_1 \subset A_2 \;\Rightarrow\; f(A_1) \subset f(A_2), \\
B_1 \subset B_2 \;\Rightarrow\; f^{-1}(B_1) \subset f^{-1}(B_2), \\
f(A_1 \cup A_2) = f(A_1) \cup f(A_2), \\
f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2), \\
f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2), \\
f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2), \\
f(f^{-1}(B)) = B \cap \operatorname{im} f, \\
f^{-1}(f(A)) \supset A.
\end{gathered}$$

2.7.8 Złożenie funkcji

Twierdzenie 2.7.13 *Niech $f : X \multimap Y$, $g : Y \multimap Z$ będą funkcjami częściowymi. Wtedy $g \circ f$ jest funkcją częściową z X do Z,*

$$\operatorname{dom} g \circ f = f^{-1}(\operatorname{dom} g),$$

oraz dla $x \in \operatorname{dom} g \circ f$

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

Co więcej, jeśli f i g są funkcjami, to $g \circ f$ jest funkcją.

2.7.9 Injekcje, surjekcje, i bijekcje.

Relacja odwrotna do relacji jednolistnej nie musi być jednolistna. Zatem relacja odwrotna do funkcji częściowej nie musi być funkcją częściową.

Funkcję $f : X \to Y$ nazywamy *injekcją*, jeżeli dla dowolnych $x_1, x_2 \in X$

$$f(x_1) = f(x_2) \;\Rightarrow\; x_1 = x_2.$$

Funkcję f nazywamy *surjekcją*, jeżeli dla dowolnego $y \in Y$ istnieje $x \in X$ takie, że $y = f(x)$. Funkcję f nazywamy *bijekcją* jeśli jest injekcją i surjekcją.

Twierdzenie 2.7.14 *Niech $f : X \to Y$ będzie funkcją. Relacja odwrotna f^{-1} jest jednolistna wtedy i tylko wtedy gdy f jest injekcją. Relacja odwrotna f^{-1} jest funkcją wtedy i tylko wtedy gdy f jest bijekcją.*

Relację odwrotną do bijekcji nazywamy *funkcją odwrotną*.

Pakiet `Relations.m` umożliwia sprawdzenie injektywności, surjektywności i bijektywności funkcji przy pomocy metod `IsInjective`, `IsSurjective` i `IsBijective`.

2.7.10 (*)Iloczyn kartezjański n zbiorów

Iloczyn kartezjański nie jest łączny:

$$(X \times Y) \times Z \neq X \times (Y \times Z),$$

bo $((a, b), c) \neq (a, (b, c))$. Jest jednak naturalna bijekcja

$$(X \times Y) \times Z \ni ((a, b), c) \mapsto (a, (b, c)) \in X \times (Y \times Z),$$

która pozwala utożsamiać te zbiory. Pozwala to mówić o iloczynie kartezjańskim trzech (i analogicznie) więcej zbiorów bez podawania nawiasów. Dla n-krotnego iloczynu kartezjańskiego zbioru X przez siebie stosujemy krótką notację X^n.

2.7.11 (*)Sumy i iloczyny indeksowane

Definicja 2.7.15 Często zdarza się, że rodzina podzbiorów zbioru X zadana jest jako kodziedzina odwzorowania, t.zw. odwzorowania indeksującego $A : J \to \mathcal{P}(X)$. Wtedy rodzinę tę zapisujemy często jako $\{A_j\}_{j \in J}$ oraz wprowadzamy oznaczenia na sumę i przecięcie tej rodziny

$$\bigcup_{j \in J} A_j := \bigcup \{ A_j \mid j \in J \}, \quad \bigcap_{j \in J} A_j := \bigcap \{ A_j \mid j \in J \}.$$

Twierdzenie 2.7.16 *Niech $A : J \to \mathcal{P}(X)$ będzie odwzorowaniem indeksującym rodzinę podzbiorów X. Jeśli $J = J_1 \cup J_2$, to*

$$\bigcup_{j\in J} A_j = \left(\bigcup_{j\in J_1} A_j\right) \cup \left(\bigcup_{j\in J_2} A_j\right), \quad \bigcap_{j\in J} A_j = \left(\bigcap_{j\in J_1} A_j\right) \cap \left(\bigcap_{j\in J_2} A_j\right).$$

Ponadto

$$\setminus \left(\bigcup_{j\in J} A_j\right) = \bigcap_{j\in J} \setminus A_j, \quad \setminus \left(\bigcap_{j\in J} A_j\right) = \bigcup_{j\in J} \setminus A_j.$$

Jeśli dodatkowo $B : K \to \mathcal{P}(Y)$ jest odwzorowaniem indeksującym rodzinę podzbiorów Y, to

$$\left(\bigcup_{j\in J} A_j\right) \times \left(\bigcup_{k\in K} B_k\right) = \bigcup_{(j,k)\in J\times K} A_j \times B_k,$$
$$\left(\bigcup_{j\in J} A_j\right) \cap \left(\bigcup_{k\in K} B_k\right) = \bigcup_{(j,k)\in J\times K} A_j \cap B_k,$$
$$\left(\bigcap_{j\in J} A_j\right) \cup \left(\bigcap_{k\in K} B_k\right) = \bigcap_{(j,k)\in J\times K} A_j \cup B_k.$$

2.8 Struktury porządkowe

2.8.1 Częściowy i liniowy porządek

Definicja 2.8.1 Relację R w X nazywamy *częściowym porządkiem* lub krótko *porządkiem* jeżeli R jest relacją zwrotną, antysymetryczną i przechodnią, a *liniowym porządkiem* jeżeli dodatkowo R jest spójna. Parę (X, R) nazywamy *przestrzenią częściowo uporządkowaną* lub krótko *przestrzenią uporządkowaną* jeżeli R jest częściowym porządkiem w X, a *przestrzenią liniowo uporządkowaną* jeżeli R jest porządkiem liniowym.

Klasę przestrzeni częściowo uporządkowanych oznaczać będziemy Ord, a liniowo uporząkowanych LOrd. Jeśli R jest częściowym porządkiem zwyczajowo będziemy pisać $\leqslant$ w miejsce R, a $\geqslant$ w miejsce R^{-1}. Ponadto mówić będziemy, że x *jest mniejsze* od y i pisać $x < y$ jeżeli

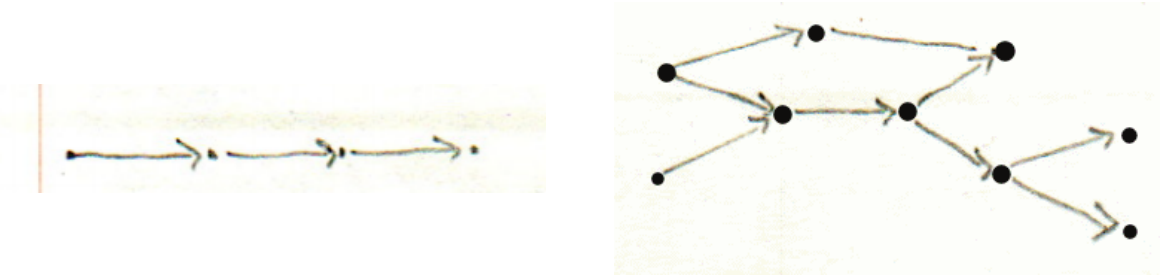
$$x \leqslant y \text{ i } x \neq y.$$

Analogicznie mówić będziemy, że x *jest większe* od y i pisać $x > y$ jeżeli

$$x \geqslant y \text{ i } x \neq y.$$

W rozdziale 3.3.3 podamy formalną definicję porządku linowego w zbiorze liczb całkowitych $\mathbb{Z}$ (wzór (3.6)), a w rozdziale 3.4.3 porządku liniowego w zbiorze liczb wymiernych $\mathbb{Q}$ (wzór (3.7)). W obu przypadkach będą to porządki zgodne z codzienną intuicją związaną z porównywaniem liczb.

Uwaga 2.8.2 *Jeśli R jest porządkiem w X, a $Y \subset X$, to $R_Y := R \cap (Y \times Y)$ jest porządkiem w Y. Jeśli R jest liniowy, to R_Y jest liniowy.*



Rysunek 2.14: Graf liniowego i częściowego porządku

Porządek R_Y nazywamy *porządkiem indukowanym*, a (Y, R_Y) nazywamy podprzestrzenią przestrzeni uporządkowanej (X, R).

Porządek w $\mathbb{Z} \subset \mathbb{Q}$ jest przykładem porządku indukowanego z $\mathbb{Q}$.

Ćwiczenie 2.8.3 Udowodnić uwagę 2.8.2

Aby zwiększyć przejrzystość rysunku, rysując graf częściowego bądź liniowego porządku pomijamy krawędzie, które są konsekwencją własności przechodniości i własności zwrotności. Tak otrzymany graf częściowego porząku nazywamy grafem zredukowanym. Jest on bardziej przejrzysty, bo zawiera mniej krawędzi, które w wyobraźni łatwo uzupełnić wykorzystując własność zwrotności i przechodniości,

Rys. 2.14 przedstawia graf zredukowany przestrzeni liniowo uporządkowanej (po lewej) oraz graf zredukowany przestrzeni częściowo uporządkowanej (po prawej).

2.8.2 Majoranty i minoranty

Definicja 2.8.4 Niech $(X, \leqslant)$ będzie przestrzenią częściowo uporządkowaną i niech A będzie podzbiorem X. Mówimy, że $x \in X$ jest *majorantą* zbioru A jeżeli $a \leqslant x$ dla każdego elementu $a \in A$. Jeżeli $x \in A$ i x jest majorantą A, to x nazywamy *elementem największym* w A. Z kolei mówimy, że x jest *elementem maksymalnym* w A jeżeli $x \in A$ oraz w A nie istnieje element y większy od x. Analogicznie, odwracając kierunek nierówności, definiujemy *minorantę*, *element najmniejszy* i *element minimalny*. Jeśli w podzbiorze A przestrzeni częściowo uporządkowanej istnieje element największy (najmniejszy) to oznaczamy go $\max A$, $(\min A)$.

Rys. 2.15 przedstawia przestrzeń częściowo uporządkowaną X oraz jej podzbiór A. W zbiorze A są dwa elementy maksymalne e oraz f, ale nie ma w nim elementu największego. Nie ma też majorant. Natomiast elementem najmniejszym zbioru A jest d. Jest to też jedyny element minimalny tego zbioru. Minorantami zbioru A są elementy a, b, c, d.

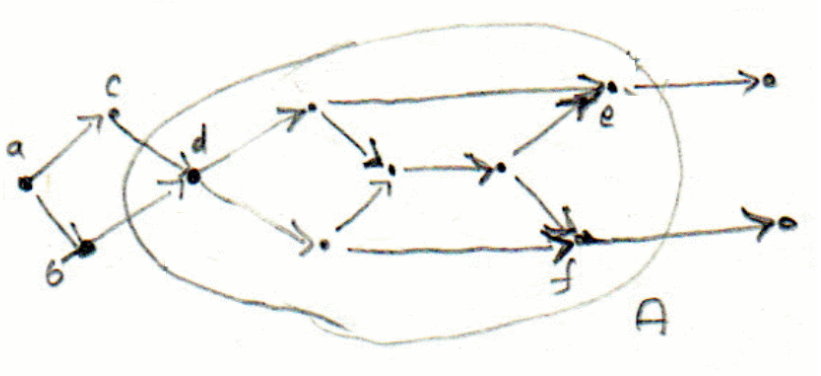
Twierdzenie 2.8.5 *Niech $(X, \leqslant)$ będzie przestrzenią częściowo uporządkowaną i niech A będzie podzbiorem X.*

(i) Tak element największy jak i najmniejszy w zbiorze A istnieje co najwyżej jeden.

(ii) Podzbiór skończony niepusty A przestrzeni częściowo uporząkowanej posiada co najmniej jeden element minimalny i co najmniej jeden element maksymalny w A.

(iii) Element maksymalny (minimalny) zbioru liniowo uporządkowanego jest jego elementem największym (najmniejszym).

(iv) Podzbiór skończony niepusty przestrzeni liniowo uporządkowanej posiada element największy i element najmniejszy.



Rysunek 2.15: Przetrzeń częściowo uporządkowana i jej podzbiór A

Dowód: Aby pokazać własnośc (i) przyjmijmy najpierw, że z_1 i z_2 są elementami największymi w zbiorze A. Z definicji elementu największego wiemy, że z_1 jest majorantą A, zatem

$$\forall x \in A \;\; x \leqslant z_1. \tag{2.2}$$

Ponieważ z_2 też jest największy, więc również

$$\forall x \in A \;\; x \leqslant z_2. \tag{2.3}$$

Z definicji elementu największego wiemym że $z_1, z_2 \in A$. Podstawiając z_2 za x w (2.2) dostajemy $z_2 \leqslant z_1$. Podobnie z (2.3) dostajemy $z_1 \leqslant z_2$. Wobec własności antysymetrii relacji porządku dostajemy $z_1 = z_2$. Zatem element największy jest co najwyżej jeden. Podobnie pokazujemy, że element najmniejszy jest co najwyżej jeden.

Własność (ii) pokażemy nie wprost. Przypuśćmy, że w A nie ma elementu maksymalnego. Zatem dla każdego $x \in A$ znajdziemy $x^* \in A$ taki, że $x^* > x$. Ponieważ $A \neq \varnothing$, możemy wybrać element $a \in A$. Rozważmy elementy

$$a < a^* < a^{**} < \cdots$$

Ponieważ zbiór A jest skończony, więc w tym ciągu nierówności muszą znaleźć się dwa wyrazy równe. Ale równe być nie mogą, bo nierówności są silne. Sprzeczność ta dowodzi, że przypuszcznie iż w A nie ma elementu maksymalnego było fałszywe. Podobnie dowodzimy, że w A musi być element minimalny.

Aby pokazać własność (iii) przypuśćmy, że a jest maksymalny w liniowo uporządkowanym zbiorze A i równocześnie nie jest on największy. Zatem istnieje $b \in A$ taki, że nie jest prawdą iż $b \leqslant a$. Ponieważ A jest liniowo uporządkowany, z własności spójności wynika, że $b > a$. Ale to przeczy maksymalności a. Zatem a jest największy Podobnie pokazujemy, że element minimalny zbioru liniowo uporządkowanego jest jego elementem najmniejszym.

Aby zakończyć dowód wystarczy zauważyć, że własność (iv) wynika natychmiast z własności (ii) i (iii). □

2.9 Przestrzenie ciągłe

Wiemy, że zbiór liczb wymiernych $\mathbb{Q}$ jest dziurawy. Brakuje w nim wielu liczb potrzebnych do mierzenia długości: na przykład długości przekątnej kwadratu lub obwodu okręgu. Potrzebujemy "zakleić" te dziury. W tym rozdziale podamy formalną definicję przestrzeni liniowo uporządkowanej bez dziur, tzw. przestrzeni ciągłej. Zbiór $\mathbb{Q}$ nie jest przestrzenią ciągła. W rozdziale 4.1.3 skonstruujemy z liczb wymiernych zbiór liczb rzeczywistych $\mathbb{R}$ i pokażemy, że w przeciwieństwie do $\mathbb{Q}$ jest on przestrzenią ciągłą.

2.9.1 Kresy

Niech X będzie przestrzenią częściowo uporządkowaną i niech $A \subset X$.

Definicja 2.9.1 Mówimy, że zbiór A jest *ograniczony od góry (od dołu)* jeżeli posiada majorantę (minorantę).

Definicja 2.9.2 Mówimy, że x jest *kresem górnym (supremum)* zbioru A i piszemy

$$x = \sup A,$$

jeżeli A jest ograniczony od góry, a x jest jego najmniejszą majorantą.

Definicja 2.9.3 Mówimy, że x jest *kresem dolnym (infimum)* zbioru A i piszemy

$$x = \inf A,$$

jeżeli A jest ograniczony od dołu, a x jest jego największą minorantą.

Ćwiczenie 2.9.4
- Podaj przykład ilustrujący, że w zbiorze skończonym częściowo uporządkowanym X może być podzbiór A, którego kres dolny (górny) nie istnieje, albo istnieje, ale do A nie należy
- Pokaż, że jeśli X jest skończony, liniowo uporządkowany, to dla dowolnego $A \subset X$ kres dolny (górny) istnieje i zawsze należy do A.

Zobaczymy, że w zbiorach nieskończonych liniowo uporządkowanych mogą być podzbiory ograniczone, których kresy nie istnieją, bądź istnieją, ale do podzbioru nie należą.

2.9.2 Przedziały

Definicja 2.9.5 Niech X będzie przestrzenią częściowo uporządkowaną i niech $a, b \in X$. *Przedział otwarty, lewostronnie otwarty, prawostronnie otwarty oraz przedział domknięty o końcach a, b* definiujemy odpowiednio wzorami

$$\begin{aligned}
(a,b) &:= \{\, x \in X \mid a < x \text{ i } x < b \,\},\\
(a,b] &:= \{\, x \in X \mid a < x \text{ i } x \leqslant b \,\},\\
[a,b) &:= \{\, x \in X \mid a \leqslant x \text{ i } x < b \,\},\\
[a,b] &:= \{\, x \in X \mid a \leqslant x \text{ i } x \leqslant b \,\}.
\end{aligned}$$

Podobnie definiujemy *lewy półprzedział otwarty*, *prawy półprzedział otwarty*, *lewy półprzedział domknięty*, *prawy półprzedział domknięty* o końcu a odpowiednio wzorami wzorami

$$\begin{aligned}
(-\infty,a) &:= \{\, x \in X \mid x < a \,\},\\
(a,\infty) &:= \{\, x \in X \mid a < x \,\},\\
(-\infty,a] &:= \{\, x \in X \mid x \leqslant a \,\},\\
[a,\infty) &:= \{\, x \in X \mid a \leqslant x \,\}.
\end{aligned}$$

Należy zwrócić uwagę, że oznaczenie (a, b) na przedział otwarty jest nieodróżnialne od oznaczenia na parę (a, b). Jednak z kontekstu zawsze można będzie ustalić, o którą interpretację chodzi.

2.9.3 Własność L

Mówimy, że niepusty podzbiór A przestrzeni liniowo uporządkowanej X ma *własność (L)* jeżeli A jest ograniczony od góry oraz

$$x \in A \text{ i } y \leqslant x \;\Rightarrow\; y \in A.$$

Analogicznie mówimy, że A ma *własność (R)* jeżeli A jest ograniczony od dołu oraz

$$x \in A \text{ i } y \geqslant x \;\Rightarrow\; y \in A.$$

Uwaga 2.9.6 *Każdy lewy półprzedział ma własność (L), a każdy prawy półprzedział ma własność (R). Niepusty, właściwy podzbiór A przestrzeni liniowo uporządkowanej X ma własność (L) wtedy i tylko wtedy gdy $X \setminus A$ ma własność (R).*

Pojawia się naturalne pytanie czy w jakiejś przestrzeni liniowo uprządkowanej mogą istnieć zbiory, które mają własność (L) lub (R), a nie są przedziałami. Odpowiedź jest twierdząca. Zbiór

$$A := \{\, q \in \mathbb{Q} \mid q \leqslant 0 \text{ albo } q^2 < 2 \,\}$$

ma własność (L), ale nie jest przedziałem, bo ewentualnym końcem tego przedziału musiałby być $\sqrt{2}$, a wiemy już, że w zbiorze liczb wymiernych takiej liczby nie ma. Zbiór ten jest równocześnie przykładem zbioru ograniczonego od góry, który nie ma (w $\mathbb{Q}$) kresu górnego. Ma on jednak kres górny, do tego zbioru nie należący, jeśli A potraktujemy jako podzbiór zbioru liczb rzeczywistych. Kresem tym jest w takim przypadku $\sqrt{2}$.

2.9.4 Przestrzenie ciągłe

Definicja 2.9.7 Mówimy, że przestrzeń liniowo uporządkowana X jest *przestrzenią ciągłą* jeżeli dla każdego niepustego, właściwego podzbioru A tej przestrzeni, który ma własność (L) istnieje $a \in X$ takie, że

$$A = (-\infty, a) \text{ lub } A = (-\infty, a]$$

co jest równoważne założeniu, że dla każdego niepustego, właściwego podzbioru A tej przestrzeni, który ma własność (R) istnieje $a \in X$ takie, że

$$A = (a, \infty) \text{ lub } A = [a, \infty).$$

Od zbioru liczb rzeczywistych, jakkolwiek go zdefiniujemy, oczekujemy, żeby był przestrzenią liniowo uporządkowaną ciągłą.

Twierdzenie 2.9.8 *Następujące trzy warunki są parami równoważne.*

(i) X jest przestrzenią ciągłą.

(ii) Każdy niepusty, ograniczony od góry podzbiór X posiada kres górny.

(iii) Każdy niepusty, ograniczony od dołu podzbiór X posiada kres dolny.

2.10 Funkcje monotoniczne

2.10.1 Funkcje rosnące i malejące

Definicja 2.10.1 Niech X i Y będą przestrzeniami uporządkowanymi, a $f : X \to Y$ niech będzie funkcją. Mówimy, że f jest *słabo rosnąca*, jeżeli

$$\forall x_1, x_2 \in X \quad x_1 < x_2 \;\Rightarrow\; f(x_1) \leqslant f(x_2),$$

a *silnie rosnąca*, jeżeli

$$\forall x_1, x_2 \in X \quad x_1 < x_2 \;\Rightarrow\; f(x_1) < f(x_2).$$

Podobnie, mówimy, że f jest *słabo malejąca*, jeżeli

$$\forall x_1, x_2 \in X \quad x_1 < x_2 \;\Rightarrow\; f(x_1) \geqslant f(x_2),$$

a *silnie malejąca*, jeżeli

$$\forall x_1, x_2 \in X \quad x_1 < x_2 \;\Rightarrow\; f(x_1) > f(x_2).$$

Twierdzenie 2.10.2 *Niech X będzie przestrzenią liniowo uporządkowaną i niech $f : X \to X$ będzie odwracalna. Wtedy jeśli f jest rosnąca (malejąca) to f^{-1} jest rosnąca (malejąca).*

Rozdział 3

Liczby naturalne, całkowite i wymierne

3.1 Liczby naturalne

Naszym celem w tym i następnym rozdziale jest naszkicowanie konstrukcji zbiorów liczbowych na gruncie teorii mnogości. Czasami, jak to zrobimy w przypadku zbioru liczb całkowitych, wymiernych, czy zespolonych wygodnie jest przeprowadzić bezpośrednie konstrukcje. Robimy tak zazwyczaj wtedy, gdy jest jakaś szczególnie naturalna, narzucająca się konstrukcja. Jeśli możliwych, istotnie różnych konstrukcji jest wiele, preferujemy podanie ogólnej definicji wraz ze stwierdzeniem pokazującym, że stosowny zbiór da się skonstruować i to w jedyny, z dokładnością do pewnego utożsamienia, sposób. Tak postąpimy w przypadku liczb naturalnych i liczb rzeczywistych.

Ponadto w rozdziale tym zdefiniujemy pojęcie działania, grupy oraz ciała. Pojęcia te pozwalają uogólnić koncepcję dodawania i mnożenia liczb i dowodzić twierdzenia modelowane na własnościach liczb, a prawdziwe w znacznie szerszym zakresie. W przypadku takich ogólnych definicji często używamy pojęcia aksjomatyka. Należy jednak podkreślić, że użycie słowa aksjomatyka w odniesieniu do definicji nie oznacza, że rozszerzamy zbiór aksjomatów. W matematyce budowanej w oparciu o terię mnogości jedyne potrzebne aksjomaty to nie omówione przez nas aksjomaty teorii mnogości. Użycie zwrotu aksjomatyka w odniesieniu do definicji służy jedynie podkreśleniu, że wszystkie twierdzenia, które wywieść można z danej definicji stanowią ważną teorię, która może być rozwijana niezależnie od teorii mnogości.

3.1.1 Aksjomatyka liczb naturalnych

Naturalne jest przyjąć, że pierwotny pasterz, który wiedział, że ma stado owiec liczące n sztuk, był świadom, że może też mieć stado o jedną sztukę większe. "O jeden więcej" to bardzo naturalne i bardzo elementarne pojęcie w matematyce określane mianem następnika. Jest ono podstawą następującej definicji określanej mianem aksjomatyki liczb naturanych.

Definicja 3.1.1 *Zbiór liczb naturalnych* to trójka $(N, 0, S)$ spełniająca warunki

(i) N jest zbiorem (w skrócie: $N \in \mathbb{V}$),

(ii) $S : N \to N$ jest injekcją,

(iii) $0 \in N \setminus S(N)$

(iv) (*zasada indukcji*) dla dowolnego $K \subset N$ jeśli

- $0 \in K$,
- $n \in K \;\Rightarrow\; S(n) \in K$

to $K = N$.



Rysunek 3.1: John von Neumann, 1903-1957, Źródło: Wikipedia

3.1.2 (*)Konstrukcja von Neumanna

Postawienie powyższej definicji oczywiście nie gwarantuje, że coś co spełnia wymogi tej definicji rzeczywiście istnieje i jest w jakimś rozsądnym sensie dokładnie jedno (nie chcielibyśmy mieć wielu istotnie różnych zbiorów liczb naturalnych). Dlatego potrzebna jest jakaś konstrukcja zbioru liczb naturalnych na gruncie teorii mnogości oraz uzasadnienie jednoznaczności. Formalną konstrukcję zbioru liczb naturalnych na gruncie teorii mnogości podał amerykański matematyk węgierskiego pochodzenia, John von Neumann (rys. 3.1). Rolę następnika pełni w niej funkcja S, bo na tym etapie nie mamy jeszcze formalnie zdefiniowanego dodawania. Dodawanie zdefiniujemy dopiero w oparciu o pojęcie następnika.

Dla dowolnego a połóżmy $s(a) := a \cup \{a\}$.

Twierdzenie 3.1.2 *Niech $\mathcal{A}$ będzie klasą wszystkich zbiorów A takich, że*

(i) $\varnothing \in A$,

(ii) $a \in A \Rightarrow s(a) \in A$.

Biorąc trójkę $(N, 0, S)$ gdzie

(i) $N := \bigcap \mathcal{A}$,

(ii) $0 := \varnothing$,

(iii) $S : N \ni a \mapsto s(a) \in N$,

otrzymujemy zbiór liczb naturalnych.

Warto się przyjrzeć jak wygląda kilka początkowych liczb naturalnych według von Neumanna. I tak, mamy

$$\begin{aligned} 0 &:= \varnothing, \\ 1 &:= s(0) = \{\varnothing\}, \\ 2 &:= s(1) = \{\varnothing\} \cup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\}, \\ 3 &:= s(2) = \{\varnothing, \{\varnothing\}\} \cup \{\{\varnothing, \{\varnothing\}\}\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}, \\ &\dots . \end{aligned}$$

3.1.3 (*)Jednoznaczność

Konstrukcja von Neumanna nie jest jedyną możliwą. Na przykład można przyjąć

$$\begin{aligned} 0 &:= \varnothing, \\ 1 &:= \{\varnothing\}, \\ 2 &:= \{\{\varnothing\}\}, \\ 3 &:= \{\{\{\varnothing\}\}\}, \\ &\dots . \end{aligned}$$

Dlatego potrzebne jest jeszcze poniższe twierdzenie, które mówi, że dowolne konstrukcje zbioru liczb naturalnych są identyczne z dokładnością do bijekcji.

Twierdzenie 3.1.3 *Jeśli $(N_1, 0_1, S_1)$ i $(N_2, 0_2, S_2)$ są dwoma zbiorami liczb naturalnych, to istnieje bijekcja $f : N_1 \to N_2$ taka, że $f(0_1) = (0_2)$ oraz dla każdego $n \in N_1$ zachodzi $f(S_1(n)) = S_2(f(n))$.*

Zatem zbiór liczb naturalnych istnieje i jest wyznaczony jednoznacznie z dokładnością do bijekcji.

W tym momencie widać pożytek z postawienia formalnej definicji (aksjomatyki) zbioru liczb naturalnych. Pozwala ona na codzień nie myśleć o skomplikowanej konstrukcji liczb naturalnych i badać ich własności jedynie poprzez odnoszenie się do własności podanych w definicji.

Szczególną rolę pełni własność 3.1.1(iv). Jest ona podstawą schematu dowodowego zwanego dowodem przez indukcję. Schemat ten mówi, że aby udowodnić iż pewna własność przysługuje wszystkim liczbom naturalnym, wystarczy pokazać, że przysługuje ona liczbie 0 oraz udowodnić implikację, że jeśli przysługuje liczbie naturalnej n to przysługuje liczbie następnej $S(n)$.

Tradycyjnie zbiór liczb naturalnych oznaczamy przez $\mathbb{N}$ i stosujemy notację $n^* := S(n)$ dopóki nie zdefiniujemy formalnie dodawania liczb naturalnych oraz liczby jeden.

3.2 Ciągi i podciągi

Jesteśmy wreszcie gotowi podać formalną definicję ciągu. Do tego potrzebowaliśmy definicji funkcji oraz definicji zbioru liczb naturalnych. Ciąg to po prostu funkcja ze zbioru liczb naturalnych do dowolnego innego zbioru.

3.2.1 Definicja ciągu i podciągu

Definicja 3.2.1 *Ciągiem* w zbiorze X nazywamy dowolne odwzorowanie $a : \mathbb{N} \to X$.

Dla ciągu a obok notacji funkcyjnej $a(n)$ dla wartości a w n stosujemy też tradycyjną notację ciągową $a_n := a(n)$. O a_n mówimy często, że jest *n-tym wyrazem ciągu a* lub krótko *wyrazem ciągu a*.

Definicja 3.2.2 Niech X będzie zbiorem, $c : \mathbb{N} \to X$ ciągiem o wartościach w X, a $k : \mathbb{N} \to \mathbb{N}$ ciągiem silnie rosnącym. Wtedy ciąg będący złożeniem $c \circ k$ nazywamy *podciągiem* ciągu c. Jego n-tym wyrazem jest $(c \circ k)(n)$. Tradycyjnie oznaczamy go c_{k_n}.

```
(* Definiujemy funkcję o nazwie SqrtTwoBf obliczającą pierwiastek z dwóch
   z jednym argumentem określającym żądaną dokładność *)
SqrtTwoBf[n_] := (* Znak podkreślenia po n informuje, że n jest argumentem *)
  Module[{i}, (* Moduł może zawierać zmienne lokalne.
                 Ich listę umieszczamy na początku *)
    i = 1; (* Inicjalizujemy zmienną i wartością jeden *)
    While[   (* Początek pętli *)
      i^2 < 2*100^n,(* Pętlę wykonujemy jak długo spełniony jest warunek *)
      i = i + 1 (* Jedyna instrukcja wewnątrz naszej pętli *)
    ]; (* Koniec pętli *)
    (* Ostatnie obliczone wyrażenie to wartość zwracana przez funkcję *)
    (* Jest nim para ułamków szacujących pierwiastek od dołu i góry *)
    {(i - 1)/10^n, i/10^n}
  ] (* Koniec modułu definiującego funkcję *)
```

Listing 3.1: Obliczania pierwiastka metodą brutalnej siły zademonstrowane w języku Mathematica. Tekst zamknięty w parach dwuznaków (* i *) oznacza komentarz.

Rozważmy ciągi

$$a : \mathbb{N} \ni n \to \frac{1}{1+n} \in \mathbb{Q},$$
$$k : \mathbb{N} \in n \to 2n \in \mathbb{N}.$$

Początkowe wyrazów ciągu a to

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$$

Początkowe wyrazy ciągu k to

$$0, 2, 4, 6, \dots$$

Ciąg k jest rosnącym ciągiem o wartościach w $\mathbb{N}$. Zatem możemy rozważać złożenie $b := a \circ k : \mathbb{N} \to \mathbb{Q}$. Złożenie to jest ciągiem

$$b : \mathbb{N} \ni n \to a(k(n)) = a_{k_n} = \frac{1}{1+2n} \in \mathbb{Q}.$$

Kilka początkowych wyrazów ciągu b to

$$1, \quad \frac{1}{1+2\times 1} = \frac{1}{3}, \quad \frac{1}{1+2\times 2} = \frac{1}{5}, \quad \frac{1}{1+2\times 3} = \frac{1}{7}, \dots.$$

Widać więc, że ciąg b jest zbudowany z wybranych wyrazów ciągu a, przy czym wyboru dokonujemy przy pomocy ciągu k. Wyjaśnia to nazwę podciąg.

3.2.2 Ciąg przybliżeń pierwiastka jeszcze raz.

Ciąg instrukcji, który wykorzystywaliśmy do liczenia przybliżeń pierwiastka wygodnie zapakować w postaci programu, z którego można korzystać wielokrotnie. Program ten przedstawiono na listingu 3.1. Wystarczy teraz napisać

```
SqrtTwo[3]
```

i otrzymamy

$$\{\frac{707}{500}, \frac{283}{200}\}.$$

```
(* Funkcja obliczająca pierwiastek z liczby a>0 z dokładnością do d cyfr *)
SqrtBfPlus[a_, d_] :=
  Module[{i, n},          (* Zmienne lokalne: licznik i mianownik *)
    i = 1; n = 1;
    While[True, (* Główną pętlę przerwiemy instrukcją Break poniżej *)
      (* Wewnętrzna pętla dobiera wartość licznika do aktualnego mianownika *)
      While[(i + 1)*(i + 1) < a*n*n, i = i + 1];
      (* Po osiągnięciu żądanej dokładności wychodzimy z pętli *)
      If[n >= 10^d, Break[]];
      (* Rozszerzamy reprezentację ułamka przed kolejnym przejściem pętli *)
      i = i*10;
      n = n*10;
    ];
    {i/n, (i + 1)/n}
  ]
```

Listing 3.2: Usprawnione podejście do obliczania pierwiastka przedstawione w programie Mathematica.

Ćwiczenie komputerowe 3.2.3 Adaptuj nasz program **SqrtTwoBf** do **SqrtBf** tak by liczył pierwiastek z dowolnej liczby naturalnej.

Ćwiczenie komputerowe 3.2.4 Sprawdź doświadczalnie jak rośnie czas obliczeń naszego algorytmu **SqrtTwo** przy zwiększaniu dokładności. Warto wiedzieć, że obliczenie wyrażenia, którego wykonanie w programie Mathematica trwa zbyt długo zawsze można przerwać wybierając w menu Evaluation opcję Abort evaluation.

Jak widać nasz algorytm liczenia pierwiastka jest daleki od doskonałości. Znacznie ulepszoną wersję przedstawiono na listingu 3.2 Jak zobaczymy niebawem są jeszcze lepsze algorytmy. Warto jednak uświadomić sobie, że, choć tego nie jesteśmy w stanie zaobserwować przy obliczaniu pojedynczego pierwiastka, jedną z przyczyn, dla których obliczenia przy dużej dokładnosci są dość wolne jest przekroczenie pojemności maszynowego typu **int** i konieczność programowego symulowania większej pojemności. Między innymi z tego powodu w tradycyjnych obliczeniach numerycznych zamiast dokładnych – ale na ogół bardzo wolnych – obliczeń na liczbach całkowitych lub ułamkach preferuje się znacznie szybsze, ale obarczone pewnym błędem, obliczenia na maszynowych liczbach zmiennopozycyjnych takich jak **double**, o których opowiemy w rozdziale 3.5.

Ćwiczenie komputerowe 3.2.5 Adaptuj program z listingu 3.1 lub listingu 3.2 tak by wyznaczał ciąg przybliżeń dodatniego pierwiastka równania

$$x^2 + x = 3.$$

Listing 3.2 prezentuje nieco udoskonaloną wersję liczenia pierwiastka metodą brutalnej siły. Zastosowano tu pętlę w pętli, by przy obliczeniach dla zwiększonej dokładności niepotrzebnie nie powtarzać tego co już jest wiadome na podstawie obliczeń wcześniejszych.

Ćwiczenie komputerowe 3.2.6 Zmodyfikuj program z listingu 3.2 tak, by liczył $\sqrt[3]{a}$ dla zadanej liczby a. Przetestuj go dla liczb, dla których znasz prawidłową odpowiedź.

3.2.3 Monotoniczność ciągu.

Twierdzenie 3.2.7 *Niech X będzie przestrzenią uporządkowaną. Ciąg $a : \mathbb{N} \to X$ jest słabo rosnący wtedy i tylko wtedy gdy*

$$\forall n \in \mathbb{N} \quad a_n \leqslant a_{n+1}.$$

Jest on silnie rosnący wtedy i tylko wtedy gdy

$$\forall n \in \mathbb{N} \quad a_n < a_{n+1}.$$

Analogicznie można scharakteryzować ciągi słabo i silnie malejące.

3.2.4 Rekurencyjne definiowanie ciągu.

Zobaczymy w rozdziale 3.2.5, że dla $a > 1$ funkcja

$$f : \mathbb{R}^+ \ni x \mapsto \frac{a + x^2}{2x} \in \mathbb{R}^+$$

ma tę sympatyczną własność, iż jeśli x jest przybliżeniem $\sqrt{a}$, to $f(x)$ jest jeszcze lepszym przybliżeniem $\sqrt{a}$. Zaczynając od liczby jeden możemy zatem wyprodukować ciąg coraz lepszych przybliżeń $\sqrt{a}$

$$1, f(1), f(f(1)), f(f(f(1))), \ldots.$$

Cechą tego ciągu jest to, że aby wyliczyć wyraz n-ty musimy już mieć policzony wyraz $(n-1)$-szy. Podanie ogólnego wzoru na n-ty wyraz jest jednak trudne o ile w ogóle możliwe. Może się zatem pojawić wątpliwość. Jeśli nie potrafimy podać wzoru na 4N4-ty wyraz, to czy rzeczywiście możemy mówić o dobrze zdefiniowanym ciągu? Poniższe twierdzenie o rekurencyjnym definiowaniu ciągu upewnia nas, że takie postępowanie prowadzi do dobrze zdefiniowanego ciągu.

Twierdzenie 3.2.8 *(o rekurencyjnym definiowaniu ciągu) Niech $f : X \to X$ będzie funkcją i niech $c \in X$. Wtedy istnieje dokładnie jeden ciąg $a : \mathbb{N} \to X$ taki, że*

$$\begin{aligned} a_0 &= c, \\ \forall n \in \mathbb{N} \quad a_{n^*} &= f(a_n). \end{aligned}$$

Powyższe twierdzenie dowodzi się przy pomocy zasady indukcji. Twierdzenie to pozwala definiować ciągi bez wypisywania jawnego wzoru na n-ty wyraz. Wystarczy podać wyraz początkowy oraz receptę jak z wyrazu n-tego dostać wyraz $(n+1)$. Mówimy wtedy o rekurencyjnym definiowaniu ciągu.

3.2.5 Pierwiastek poprzez ciąg rekurencyjny.

Jako ilustrację pożytków z ciągu zdefiniowanego rekurencyjnie rozważymy następującą, bardzo szybką metodę liczenia pierwiastka z liczby a. Używamy w tym celu ciągu zadanego rekurencyjnie wzorami

$$\begin{aligned} x_0 &:= 1 \\ x_n &:= \frac{a + x_n^2}{2x_n} \end{aligned}$$

Można pokazać, że ciąg ten jest zbieżny do $\sqrt{a}$, i to bardzo szybko.

Program liczący pierwiastek tą metodą napisany w języku Mathematica przedstawiono na listingu 3.3.

Problem 3.2.9 Jak ocenić prędkość zbieżności? Jak skonstruować ciąg, który będzie możliwie szybko zbieżny do poszukiwanej liczby?

```
(* Szybkie obliczanie pierwiastka z liczby a z dokladnoscia do epsilon *)
SqrtQuick[a_, d_] := Module[{x, eps},
  eps = 10^(-d); (* żądana dokładność *)
  x = 1; (* wartość początkowa ciągu rekurencyjnego *)
  While[Abs[x*x - a]/2 > eps,(* jeśli nie osiągnięto żądanej dokładności *)
   x = (x*x + a)/(2*x); (* nowa wartość ciągu rekurencyjnego *)
   Print[N[{x - (x*x - a)/2, x}, d + 1]] (* wydruk kontrolny *)
   ];
  {x - (x*x - a)/2, x} (* końcowy wynik *)
]
```

Listing 3.3: Szybkie wyznaczenie ciągu przybliżeń pierwiastka w języku Mathematica.

Odpowiedzi na pytania tego typu daje analiza numeryczna oraz teoria aproksymacji.

Rekurencyjne definiowanie ciągów dostępne jest również w językach programowania. Mówi się tam o rekurencyjnym wywoływaniu funkcji, ale chodzi o funkcje z argumentem będącym liczbą naturalną, a więc w terminologii przyjętej w matematyce o ciąg. W języku Mathematica definicję rekurencyjną podaje się jako parę wzorów. Na ogół bardzo upraszcza to pisanie programu, choć często odbija się to na jego efektywności. Na przykład wersja rekurencyjna szybkiego liczenia pierwiastka z listingu 3.3 to

```
QuickSqrt[a_, 0] = 1;
QuickSqrt[a_, n_] := Module[{x}, x = QuickSqrt[a, n - 1]; (x^2 + a)/(2*x)];
```

Pisząc

```
QuickSqrt[2.,4]
```

otrzymujemy przybliżenie $\sqrt{2}$ z dokładnością do 5 miejsc po przecinku. W rozwiązaniu tym jest jednak pewnym problemem jak dobrać wyraz ciągu by uzyskać żądaną dokładność.

Definicję rekurencyjna `QuickSqrt` *można zapisać jeszcze krócej, bez użycia modułu:*

```
QuickSqrt[a_, 0] = 1;
QuickSqrt[a_, n_] := (QuickSqrt[a, n - 1]^2 + a)/(2* QuickSqrt[a, n - 1]);
```

Ćwiczenie komputerowe 3.2.10 Sprawdź czy i jak to rozwiązanie działa. Zastanów się ile razy wywoływana jest funkcja QuickSqrt by obliczyć czwarty wyraz ciągu przybliżającego $\sqrt{2}$. Wyciągnij wnioski.

3.2.6 (**)Arytmetyka liczb natualnych

Jesteśmy teraz gotowi podać formalne definicje dodawania i mnożenia liczb naturalnych. Są to definicje rekurencyjne. Wykorzystujemy w nich aksjomatycznie daną funkcję następnika. Przypomnijmy, że następnik liczby n oznaczamy n^*.

Operacje dodawania liczb naturalnych definiujemy rekurencyjnie:

$$m + 0 := m \tag{3.1}$$
$$m + n^* := (m + n)^* \tag{3.2}$$

Liczbę jeden, oznaczaną 1, definiujemy jako następnik zera. Tak więc $1 := 0^*$. Podstawiając zero w miejsce n we wzorze 3.2 i wykorzystując 3.1 ortzymujemy

$$m + 0^* = (m + 0)^* = m^*.$$

Zatem definiując liczbę jeden jako następnik zera, to jest przyjmując $1 := 0^*$ otrzymujemy

$$m + 1 = m^*,$$

co pozwala nam zapisywać następnik n w tradycyjnej formie $n+1$.

Również rekurencyjnie definiujemy iloczyn:

$$\begin{aligned} m \cdot 0 &:= 0 \\ m \cdot n^* &:= m \cdot n + m \end{aligned}$$

Zwyczajowo znak mnożenia w iloczynie pomijamy, jeśli nie prowadzi to do nieporozumień.

Możemy teraz podać formalną definicję nierówności.

Definicja 3.2.11 Mówimy, że m jest *nie większe (mniejsze lub równe)* od n i piszemy $m \leqslant n$ jeżeli

$$\exists k \in \mathbb{N}\ m + k = n.$$

Mówimy, że m jest *mniejsze* niż n jeżeli $m \leqslant n$ i $m \neq n$.

Podobnie definiujemy podzielność.

Definicja 3.2.12 Mówimy, że m *dzieli* n i piszemy $m|n$ jeżeli

$$\exists k \in \mathbb{N}\ m \cdot k = n.$$

Można indukcyjnie pokazać, że w obu powyższych przypadkach jeśli k istnieje to jest jedyne. Nazywamy je odpowiednio *różnicą* (dla nierówności) oraz *ilorazem* (dla podzielności) i oznaczamy odpowiednio $n - m$ i $\frac{n}{m}$.

Dla $p \in N$ ędziemy stosować oznaczenia

$$\begin{aligned} \mathbb{N}_p &:= \{\, n \in \mathbb{N} \mid n \geqslant p \,\}, \\ I_p : &= \{\, n \in \mathbb{N} \mid 1 \leqslant n \leqslant p \,\}. \end{aligned}$$

Zauważmy, że w szczególności $\mathbb{N}_0 = \mathbb{N}$ oraz $I_0 = \varnothing$.

3.2.7 Zbiory skończone i przeliczalne

Definicja 3.2.13 Mówimy, że zbiór X jest skończony, jeśli istnieje liczba naturalna $n \in \mathbb{N}$ taka, że zbiór X jest równoliczny ze zbiorem I_n. Wtedy liczbę n nazywamy *mocą* zbioru X i piszemy $\operatorname{card} X = n$. Mówimy, że zbiór X jest *przeliczalny*, jeżeli X jest równoliczny ze zbiorem liczb naturalnych.

Definicja 3.2.14 *Ciągiem skończonym* w zbiorze X nazywamy dowolne odwzorowanie $a : I_p \to X$.

3.2.8 Rozszerzona zasada indukcji

Czasami wzór czy twierdzenie, które chcemy dowieść przy pomocy zasady indukcji nie ma sensu dla liczby zero, czy kilku początkowych liczb naturalnych. Innymi słowy odnosi się do zbioru $\mathbb{N}_p = \{p, p+1, p+2, \ldots\}$. Co więcej, przy dowodzeniu względnej prawdziwości twierdzenia dla n w drugim kroku indukcyjnym wygodnie jest skorzystać z założenia indukcyjnego nie tylko dla $n-1$, ale dla wszystkich $k < n$. Poprawność takiego postępowania uzasadnia następująca rozszerzona zasada indukcji.

Twierdzenie 3.2.15 *(Rozszerzona zasada indukcji) Niech $p \in \mathbb{N}$ oraz $Z \subset \mathbb{N}$. Jeżeli zachodzą warunki*

(i) $p \in Z$,

(ii) $\mathbb{N}_p \cap I_n \subset Z \Rightarrow n+1 \in Z$

to $Z = \mathbb{N}_p$.

Mamy też rozszerzone twierdzenie o rekurencyjnym definiowaniu ciągu.

Twierdzenie 3.2.16 *(rozszerzone o rekurencyjnym definiowaniu ciągu) Niech $f : X \to X$ będzie funkcją, niech $c \in X$ i $p \in \mathbb{N}$. Wtedy istnieje dokładnie jeden ciąg $a : \mathbb{N}_p \to X$ taki, że*

$$\begin{aligned} a_p &= c, \\ \forall n \in \mathbb{N}_p \quad a_{n+1} &= f(a_n). \end{aligned}$$

3.2.9 Silnia i symbol Newtona

Jednym z klasycznych pytań matematyki zbiorów skończonych jest pytanie na ile sposobów elementy zbioru n-elementowego X można ustawić w ciąg. Każde takie ustawienie to bijekcja z I_n do X. Pytamy więc ile jest możliwych takich bijekcji. Okazuje się, że jest ich $n!$. Dokładniej, można pokazać następujące twierdzenie.

Dla $n \in \mathbb{N}$ definiujemy rekurencyjnie

$$n! := \begin{cases} 1 & n = 0, \\ n(n-1)! & n > 0. \end{cases}$$

Wyrażenie $n!$ czytamy "n silnia".

Wyrażenie $n!$ pojawia się w kontekście klasycznego pytania na ile sposobów elementy zbioru n-elementowego X można ustawić w ciąg. Każde takie ustawienie to bijekcja z I_n do X. Pytamy więc ile jest możliwych takich bijekcji. Okazuje się, że jest ich $n!$. Dokładniej, można pokazać następujące twierdzenie.

Twierdzenie 3.2.17 *Niech X będzie zbiorem skończonym i niech $n := \operatorname{card} X$. Wtedy*

$$\operatorname{card} \{ f : I_n \to X \mid f \text{ jest bijekcją} \} = n!$$

Dla $n, k \in \mathbb{N}$ definiujemy tzw. symbol Newtona wzorem

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}.$$

Symbol Newtona ma również teoriomnogościową interpretację. Mamy bowiem następujące twierdzenie

Twierdzenie 3.2.18 *Niech X będzie zbiorem skończonym o mocy n. Wtedy*

$$\operatorname{card} \{ A \subset X \mid \operatorname{card} A = k \} = \binom{n}{k}.$$

Z twierdzenia tego wynika w szczególności, że symbol Newtona jest zawsze liczbą naturalną, tzn. że $k!(n-k)!$ dzieli $n!$.

W programie Mathematica silnię obliczamy stosując tradycyjną notację. Pisząc

```
50!
```

otrzymujemy

304140932017133780436126081660647688443776415689605120000000000000

Natomiast symbol Newtona obliczymy używając funkcji `Binomial`. Pisząc

```
Binomial[5, 2]
```

otrzymujemy $\binom{5}{2}$, czyli 10.

3.3 Liczby całkowite.

3.3.1 Działania

W rozdziale 4 podamy formalną definicję zbioru liczb rzeczywistych. Najpierw jednak przyjrzymy się oddzielnie tzw. strukturom algebraicznym formalizującym operacje dodawania i mnożenia.

Dodawanie i mnożenie liczb to operacje arytmetyczne znane nam dobrze ze szkoły. Z formalnego punktu widzenia operacje te to funkcje, które parom liczb rzeczywistych przyporządkowują odpowiednio ich sumę lub iloczyn.

Działaniem dwuargumentowym w zbiorze X lub krótko *działaniem* nazywamy dowolne odwzorowanie $d : X \times X \to X$. W przypadku działania dwuargumentowego zazwyczaj stosujemy tzw. notację infixową, w której symbol działania stawiamy pomiędzy argumentami

$$x \,\mathrm{d}\, y := d(x, y).$$

Mówimy o działaniach zamiast po prostu o funkcjach dwuargumentowych, bo w przypadku działań zazwyczaj oczekujemy spełnienia pewnych specjalnych własności, modelowanych na własnościach dodawania i mnożenia liczb.

Definicja 3.3.1 Działanie $d : X \times X \to X$ nazywamy *łącznym* jeżeli

$$\forall x, y, z \in X \quad (x \,\mathrm{d}\, y) \,\mathrm{d}\, z = x \,\mathrm{d}\, (y \,\mathrm{d}\, z),$$

a *przemiennym* jeżeli

$$\forall x, y \in X \quad x \,\mathrm{d}\, y = y \,\mathrm{d}\, x.$$

Element $e \in X$ nazywamy *elementem neutralnym* dla działania d, jeżeli

$$\forall x \in X \quad e \,\mathrm{d}\, x = x \,\mathrm{d}\, e = x.$$

Jeśli $d : X \times X \to X$ ma element neutralny e to dla $x \in X$ mówimy, że x' jest jego *elementem odwrotnym*, jeżeli

$$x \,\mathrm{d}\, x' = x' \,\mathrm{d}\, x = e.$$

Uwaga 3.3.2 *Jeśli działanie $d : X \times X \to X$ ma element neutralny, to jest on dokładnie jeden. Jeśli działanie d jest łączne i ma element neutralny, to element odwrotny do danego elementu może być co najwyżej jeden.*

Dowód: Aby pokazać, że element neutralny jest dokładnie jeden należy pokazać, że jeśli dwa elementy spełniają wymogi elementu neutralnego, to są one sobie równe. Załóżmy zatem, że e oraz f są elementami neutralnymi dla działania d. Mamy zatem

$$\forall x \in X \quad e \,\mathrm{d}\, x = x \,\mathrm{d}\, e = x \tag{3.3}$$

oraz

$$\forall x \in X \quad f \,\mathrm{d}\, x = x \,\mathrm{d}\, f = x. \tag{3.4}$$

Ponieważ równości w (3.3) zachodzą dla każdego x, więc w szczególności dla $x = f$ mamy

$$e \,\mathrm{d}\, f = f.$$

Podobnie, po podstawieniu $x = e$, z równości (3.4) dostajemy

$$e \,\mathrm{d}\, f = e.$$

Zatem $e = f$. Dowód drugiej części uwagi jest podobny i zostawiamy go jako ćwiczenie. □

Tak dodawanie jak i mnożenie liczb możemy traktować jako działania, jeśli odpowiednio dobierzemy zbiór liczb, na przykład są to działania w zbiorze liczb naturalnych, całkowitych, wymiernych, rzeczywistych, rzeczywistych dodatnich oraz w zbiorze ułamków (liczb wymiernych dodatnich). Mnożenie jest też działaniem w zbiorze ułamków właściwych, czyli liczb wymiernych dodatnich i mniejszych niż jeden. Jednak dodawania nie można już traktować jako działania w tym zbiorze, bo suma dwóch ułamków właściwych może nie być ułamkiem właściwym, np. $\frac{2}{3} + \frac{2}{3} = \frac{4}{3}$.

Tak dodawanie jak i mnożenie liczb są łączne oraz przemienne. Elementem neutralnym dla dodawania liczb jest zero, a dla mnożenia jest jeden. W przypadku mnożenia elementem odwrotnym do x *jest* $\frac{1}{x}$*, oczywiście o ile* $\frac{1}{x}$ *ma sens i jest elementem zbioru, w którym mnożenie rozważamy jako działanie. Zero nigdy nie posiada elementu odwrotnego względem mnożenia. Dla pozostałych liczb istnienie elementu odwrotnego względem mnożenia zależy od zbioru, w którym mnożenie rozważamy jako działanie. Na przykład liczba* 2 *posiada element odwrotny względem mnożenia w zbiorze liczb wymiernych, ale go nie posiada w zbiorze liczb całkowitych.*

W przypadku dodawania zwyczajowo zamiast o elemencie odwrotnym mówimy o elemecie przeciwnym. Dla liczby x *jest nim* $-x$*, oczywiście o ile* $-x$ *jest elementem zbioru, w którym dodawanie rozważamy jako działanie. Elementy przeciwne istnieją dla dodawania rozważanego jako działanie w zbiorze liczb rzeczywistych, wymiernych, całkowitych.*

Przykład 3.3.3 W zbiorze skończonym działanie można zadać przy pomocy tabelki. Na przykład poniższa tabelka

	a	b	c
a	b	a	c
b	a	b	c
c	b	c	a

(3.5)

zadaje działanie w zbiorze trzyelementowym $\{a, b, c\}$. Wartość $x \,\mathrm{d}\, y$ jest elementem znajdującym się w przecięciu wiersza oznaczonego x i kolumny oznaczonej y.

Pakiet `Operations` umożliwia definiowanie działań w zbiorach skończonych i sprawdzanie, czy działanie ma rozważane własności. Instrukcja `Operations` ma dwa argumenty. Pierwszy to lista elementów zbioru, w którym określono działanie. Drugi argument to tabelka zadająca działanie, zapisana jako lista wierszy. Wprowadzając

```
op = Operation[{a,b,c},{{b,a,c},{a,b,c},{b,c,a}}]
```

definiujemy działanie zadane w tabelce (3.5). Z kolei wprowadzając

```
IsAssociative[op]
HasNeutral[op]
HasInverses[op]
```

dowiadujemy się, że działanie to ma element neutralny, ale nie jest łączne i, przynajmniej dla pewnych elementów, nie ma ono elementów odwrotnych.

Ćwiczenie 3.3.4 (i) Ustal, które elementy w działaniu rozważanym w przykładzie 3.3.3 nie posiadają elementów odwrotnych.

(ii) W tabelce (3.5) wystarczy zamienić miejscami dwa elementy, by otrzymać działanie, które jest łączne, ma element neutralny i wszystkie elementy odwrotne. Ustal jaka to zamiana.



Rysunek 3.2: Évariste Galois, 1811-1832, Źródło: Wikipedia

3.3.2 Grupy

Wśród wielu różnych działań szczególną rolę pełnią działania łączne posiadające element neutralny oraz elementy odwrotne. Na przykład dodawanie w zbiorze liczb całkowitych oraz mnożenie w zbiorze liczb wymiernych dodatnich są takimi działaniami. Okazuje się, że jest bardzo wiele działań o takich własnościach, określonych w bardzo różnych zbiorach, w szczególności w zbiorach skończonych. Dlatego warto wyróżnić takie działania wprowadzając pojęcie grupy. Jednym z pierwszych, którzy pokazali wagę tego pojęcia był francuski matematyk Évariste Galois (rys. 3.2). Posłużył się on pojęciem grupy w celu określenia kiedy daje się wyrazić miejsca zerowe wielomianu przy użyciu czterech operacji arytmetycznych i pierwiastkowania.

Definicja 3.3.5 Trójkę (G, d, e) nazywamy *grupą* jeżeli $d : G \times G \to G$ jest działaniem spełniającym następujące warunki

G1) d jest łączne,

G2) e jest elementem neutralnym dla d,

G3) każdy element w G posiada element odwrotny.

W przypadku gdy spełniona jest jedynie własność G1) trójkę (G, d, e) nazywamy *półgrupą*. Jeśli spełnione są tylko własności G1) i G2), półgrupę nazywamy *półgrupą z jedynką* lub *monoidem*. Grupę (półgrupę) nazywamy *przemienną* albo *abelową*, jeżeli działanie d jest przemienne.

Przykładem grupy przemiennej jest $(\mathbb{Z}, +, 0)$ oraz $(\mathbb{R}, \cdot, 1)$. Grupą przemienną jest też $(\mathbb{R}^+, \cdot, 1)$, natomiast nie jest grupą $(\mathbb{N}, +, 0)$, bo dodawanie nie posiada w tym zbiorze elementu przeciwnego(odwrotnego). Z kolei $(\mathbb{N}, +, 0)$ jest przykładem półgrupy.

Ćwiczenie 3.3.6 (*)Niech X będzie ustalonym zbiorem. Pokazać, że zbiór bijekcji X w X ze składaniem odwzorowań jako działaniem jest grupą.

Jeśli rozważamy ustaloną grupę lub półgrupę tradycyjnie działanie d oznaczamy '$\cdot$', element neutralny oznaczamy 1, a

element odwrotny do x oznaczmy x^{-1}. W praktyce '$\cdot$' często w ogóle pomijamy. Ponadto dla $n \in \mathbb{N}$ stosujemy oznaczenia

$$\begin{aligned} x^n &:= \overbrace{x \cdot x \cdot \cdots \cdot x}^{n\text{-krotnie}} \\ x^{-n} &:= \left(x^{-1}\right)^n. \end{aligned}$$

Ten system notacji określamy mianem notacji multyplikatywnej.

W przypadku grupy lub półgrupy przemiennej tradycyjnie działanie d oznaczamy '$+$', a element neutralny oznaczamy 0. W przypadku tym zamiast o elemencie odwrotnym do x mówimy o elemencie przeciwnym do x i oznaczamy go $-x$. Ponadto dla $n \in \mathbb{N}$ stosujemy oznaczenia

$$\begin{aligned} nx &:= \overbrace{x + x + \cdots + x}^{n\text{-krotnie}} \\ (-n)x &:= n(-x). \end{aligned}$$

Ten system notacji określamy mianem notacji addytywnej.

Choć formalnie grupa jest trójką składającą się, ze zbioru, działania i elementu zbioru, w praktyce często o samym zbiorze mówimy, że jest grupą, zdając się na domyślność czytelnika co jest brane jako działanie, a co jako element neutralny.

3.3.3 Zbiór liczb całkowitych

Wiemy, że aby odejmować bez ograniczeń liczby naturalne, potrzebujemy rozszerzyć zbiór liczb naturalnych o liczby ujemne. Na przykład, żeby wykonać odejmowanie $3-5$ potrzebujemy liczby -2. Jak formalnie zdefiniować liczbę -2 na gruncie tego co już mamy, to jest liczb naturalnych? Można na przykład przyjąć, że liczba -2 to para $(3,5)$. Przy takiej definicji nie wychodzimy poza pojęcia, które już mamy zdefiniowane. Pojawia się jednak problem. Liczba -2 to również wynik odejmowania $7-9$. Zatem pary $(3,5)$ i $(7,9)$ trzeba jakoś utożsamić. Nie możemy postawić warunku $3-5=7-9$, bo jeszcze nie zdefiniowaliśmy odejmowania gdy wynik jest ujemny. Ale warunek $3-5=7-9$, po przeniesieniu 5 i 7 na drugą stronę równości, daje warunek równoważny $3+9=7+5$, który operuje jedynie dodawaniem liczb naturalnych, a to już mamy zdefiniowane. Sugeruje to, by parę (m,n) utożsamić z parą (m',n') gdy $m+n'=m'+n$ i prowadzi do następującej formalnej konstrukcji.

Zbiór liczb całkowitych definiujemy jako zbiór klas równoważności

$$\mathbb{Z} := \{\, [(m,n)]_R \;\mid\; (m,n) \in \mathbb{N}^2 \,\}$$

względem relacji

$$R := \{\, ((m,n),(m',n')) \in \mathbb{N}^2 \times \mathbb{N}^2 \;\mid\; m+n' = m'+n \,\}.$$

Dla $n \in \mathbb{N}$ wprowadzamy oznaczenie $-n := [(0,n)]_R$ oraz utożsamiamy klasę $[n,0]_R$ z liczbą n, co pozwala traktować $\mathbb{Z}$ jako rozszerzenie $\mathbb{N}$.

Ćwiczenie 3.3.7 (*)Rozszerzyć operację dodawania z $\mathbb{N}$ na $\mathbb{Z}$ tak by w $\mathbb{Z}$ uzyskać strukturę grupy przemiennej. Rozszerzyć operację mnożenia z $\mathbb{N}\setminus\{0\}$ na $\mathbb{Z}\setminus\{0\}$ tak by w $\mathbb{Z}\setminus\{0\}$ uzyskać strukturę półgrupy przemiennej.

Dla $m,n \in \mathbb{Z}$ definiujemy

$$m \leqslant n \;:\Leftrightarrow\; n-m \in \mathbb{N}. \tag{3.6}$$

Łatwo sprawdzić, że definicja ta rozszerza definicję 3.2.11 nierówności w $\mathbb{N}$. *Przedziałem* w $\mathbb{Z}$ o końcach $m,n \in \mathbb{Z}$ nazywamy zbiór

$$[m,n]_{\mathbb{Z}} := \{\, x \in \mathbb{Z} \mid m \leqslant x \leqslant n \,\}.$$

Łatwo zauważyć, że zbiór ten jest niepusty wtedy i tylko wtedy gdy $m \leqslant n$.

3.3.4 Sumy i iloczyny

Niech G będzie półgrupą (w notacji addytywnej lub multyplikatywnej), m, n elementami $\mathbb{Z}$, a

$$x : [m, n]_{\mathbb{Z}} \to G$$

ciągiem. W przypadku notacji addytywnej definiujemy rekurencyjnie *sumę* elementów ciągu x wzorem

$$\sum_{i=m}^{n} x_i := \begin{cases} 0 & m > n \\ \left(\sum_{i=m}^{n-1} x_i\right) + x_n & \text{w przeciwnym razie .} \end{cases}$$

Analogicznie dla notacji multyplikatywnej definiujemy rekurencyjnie *iloczyn* elementów ciągu x wzorem

$$\prod_{i=m}^{n} x_i := \begin{cases} 1 & m > n \\ \left(\prod_{i=m}^{n-1} x_i\right) \cdot x_n & \text{w przeciwnym razie .} \end{cases}$$

Często stosujemy też mniej formalną, ale bardziej intuicyjną notację kropkową:

$$\begin{aligned} x_m + x_{m+1} + \ldots + x_n &:= \sum_{i=m}^{n} x_i \\ x_m \cdot x_{m+1} \cdot \ldots \cdot x_n &:= \Pi_{i=m}^{n} x_i. \end{aligned}$$

Dla $n \in \mathbb{N}$ w przypadku n-krotnie wykonanego działania na elemencie x stosujemy oznaczenie

$$x^n := \begin{cases} 1 & n = 0, \\ x \cdot x^{n-1} & n > 0 \end{cases}$$

w przypadku notacji multyplikatywnej oraz oznaczenie

$$nx := \begin{cases} 0 & n = 0, \\ x + (n-1)x & n > 0 \end{cases}$$

w przypadku notacji addytywnej.

3.4 Liczby wymierne

Mamy już formalnie zdefiniowane liczby naturalne i całkowite. Po drodze wprowadziliśmy pojęcie grupy. Widzieliśmy, że choć liczby całkowite z dodawaniem tworzą grupę, to są również grupy, które niewiele mają wspólnego z liczbami całkowitymi. Naszym następnym celem są liczby wymierne. Liczby wymierne z dodawaniem też tworzą grupę, a z mnożeniem tworzą grupę gdy pominiemy zero. Ponadto, operacje dodawania i mnożenia liczb wymiernych są ze sobą powiązane: n-krotne dodadnie do siebie liczby a to to samo co pomnożenie liczby a przez n, czego konsekwencją jest rozdzielczość mnożenia względem dodawania $(a + b) \cdot c = a \cdot c + b \cdot c$. Powstaje pytanie, czy sformalizowanie tej koncepcji prowadzi już jednoznacznie do aksjomatyki liczb wymiernych, tak jak to przedstawiliśmy dla liczb naturalnych?

3.4.1 Ciała

Definicja 3.4.1 Piątkę $(F, +, \cdot, 0, 1)$ gdzie $0 \neq 1$ nazywamy *ciałem* jeżeli $(F, +, 0)$ oraz $(F \setminus \{0\}, \cdot, 1)$ są grupami przemiennymi, a ponadto zachodzi tzw. *prawo rozdzielności* mnożenia względem dodawania

$$\forall x, y, z \in F \quad (x + y) \cdot z = x \cdot z + y \cdot z.$$

Okazuje się, że są ciała istotnie różne od ciała liczb wymiernych. Jak zobaczymy mamy ciało liczb rzeczywistych, ciało liczb zespolonych, ale co ciekawe, podobnie jak w przypadku grup, są też ciała skończone.

Ćwiczenie 3.4.2 Pokazać, że zbiór $\{0,1\}$ z działaniami $+$ i $*$ danymi tabelami

$+$	0	1
0	0	1
1	1	0

$*$	0	1
0	0	0
1	0	1

jest ciałem. Ciało to zwyczajowo oznaczamy $\mathbb{Z}_2$.

3.4.2 Funkcja potęgowa i wielomian

Niech $(F,+,\cdot,0,1)$ będzie ciałem. Dla $n\in\mathbb{N}$ definiujemy *funkcję potęgową*

$$\cdot^n : F \ni x \mapsto x^n \in F,$$

a dla $a_0,a_1,\dots a_n\in F$ definiujemy *wielomian* A o współczynnikach $a_0,a_1,\dots a_n$

$$Q_{a_0,a_1,\dots a_n} : F \ni x \mapsto a_0+a_1x+a_2x^2+\dots+a_nx^n \in F.$$

Największe n takie, że $a_n\neq 0$ nazywamy stopniem wielomianu Q.

3.4.3 Zbiór liczb wymiernych

Konstrukcja zbioru liczb wymiernych ze zbioru liczb całkowitych bardzo przypomina konstrukcję liczb całkowitych z liczb naturalnych. Aby zapewnić sobie możliwość dzielenia m/n rozważamy pary (p,q) liczb całkowitych oraz relację utożsamiającą pary $(p,q)]$ i (p',q') gdy $pq'=qp'$.

Zbiór liczb wymiernych definiujemy jako zbiór klas równoważności

$$\mathbb{Q} := \{\, [(p,q)]_S \;\mid\; (p,q)\in\mathbb{Z}\times\mathbb{N}_1 \,\}$$

względem relacji

$$S := \{\, ((p,q),(p',q'))\in(\mathbb{Z}\times\mathbb{N}_1)\times(\mathbb{Z}\times\mathbb{N}_1) \;\mid\; pq'=qp' \,\}.$$

Dla liczb wymiernych $[(p,q)]_S$, $[(p',q')]_S$ definiujemy relację nierówności

$$[(p,q)]_S \leqslant [(p',q')]_S \;:\Leftrightarrow\; pq'\leqslant qp'. \tag{3.7}$$

Łatwo sprawdzić, że definicja ta nie zależy od wyboru reprezentantów i rozszerza definicję nierówności zadaną wzorem (3.6) dla liczb całkowitych.

Ćwiczenie 3.4.3 (*)Rozszerzyć operacje dodawania i mnożenia z $\mathbb{Z}$ na $\mathbb{Q}$ tak by w $\mathbb{Q}$ uzyskać strukturę ciała.

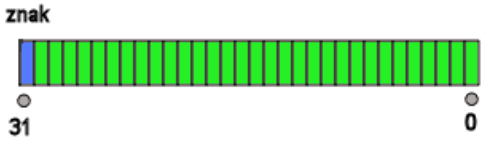
3.4.4 Wzory skróconego mnożenia

Niniejsze wzory obowiązujące w każdym ciele będą nam wielokrotnie pomocne. Ich indukcyjny dowód pomijamy.

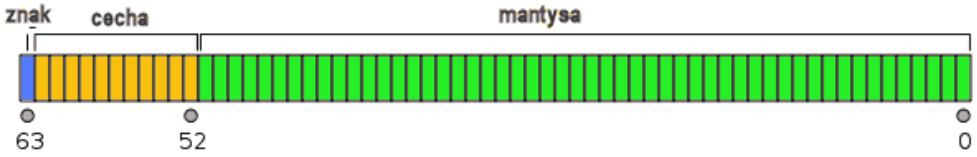
Twierdzenie 3.4.4 *Niech $(F,+,\cdot,0,1)$ będzie ciałem, $a,b\in F$ oraz $n\in\mathbb{N}_1$. Wtedy*

$$b^n-a^n = (b-a)\sum_{i=0}^{n-1} a^{n-i-1}b^i, \tag{3.8}$$

$$(a+b)^n = \sum_{i=0}^{n}\binom{n}{i}a^ib^{n-i}. \tag{3.9}$$



Rysunek 3.3: 32-bitowe słowo (4 bajty) używane do reprezentacji liczb typu **int**.



Rysunek 3.4: 64-bitowe słowo (8 bajtów) używane do reprezentacji liczb typu **double**.

3.5 Liczby reprezentowalne

W pamięci komputera możemy reprezentować jedynie pewien skończony podzbiór zbioru liczb. Używane są różne podzbiory. W przypadku liczb, na których operuje procesor są to pewne podzbiory zbioru liczb całkowitych oraz pewne podzbiory zbioru liczb wymiernych, najczęściej liczby całkowite typu **int** oraz wymierne typu **double**.

Do reprezentacji liczb typu **int** używa się zazwyczaj 32 cyfr dwójkowych, czyli inaczej bitów (rys.3.3). Sama liczba zapisana jest przy pomocy 31 bitów, a 32-gi bit używany jest do zapisania znaku. Pozwala to na zapisywanie liczb całkowitych z zakresu -2147483648 do $+2147483647$.

Liczby typu **double** przechowywane są zazwyczaj na 64 bitach, z czego jeden bit reprezentuje znak s, 11 bitów cechę N, a pozostałe 52 mantysę m będącą ułamkiem o mianowniku 2^{52} z przedziału $[1, 2)$ (rys.3.4). Umożliwa to zapisywanie liczb w postaci

$$x = sm2^N,$$

a także, przy pomocy specjalnego kodu, liczby zero. Wykorzystanie cechy i mantysy pozwala na przybliżanie liczb o bardzo dużym zakresie, od bardzo małych, rzędu 10^{-323}, do bardzo dużych, rzędu 10^{308}, przy zachowaniu w miarę stałej dokładności, wynoszącej około 15 cyfr znaczących.

Większość języków programowania zadowala się takim systemem liczb reprezentowalnych jaki jest dostępny na poziomie procesora. Tak jest na przykład w języku C++. Takie podejście jest wystarczające do większości zastosowań. Jeśli jednak z jakiegoś powodu zachodzi potrzeba stosowania większej dokładności w obliczeniach, zawsze można skorzystać z biblioteki udostępniającej liczby reprezentowalne reprezentowane na dowolnej ilości bitów, ograniczonej nie wielkością słowa maszynowego ale jedynie całkowitą ilością wolnej pamięci. Jedyne co przy takim rozwiązaniu tracimy to czas obliczeń, który jest istotnie wolniejszy niż w przypadku obliczeń wykonywanych na pojedynczych słowach maszynowych.

Specjalizowane programy do matematyki obliczeniowej, takie jak Mathematica mają tego typu biblioteki wbudowane. Na przykład, aby pomnożyć przez siebie dwie liczby 25-cyfrowe wystarczy napisać

```
1234567890123456789012345*1234567890123456789012345.
```

Natychmiast otrzymujemy wynik

$$1524157875323883675049533479957338669120562399025$$

Pisząc

```
2^200
```

otrzymujemy

160693804425899027554196209234116260252220299378279283530137 6

Możliwa jest też zamiana ułamka na liczbę dziesiętną o dowolnie wysokiej dokładności. Na przykład pisząc

```
N[1/23, 50]
```

otrzymujemy przybliżenie ułamka $\frac{1}{23}$

0.043478260869565217391304347826086956521739130434783

z dokładnością do 50 cyfr.

Tam gdzie liczy się prędkość obliczeń (na przykład, gdy musimy policzyć nie jeden pierwiastek, ale miliony pierwiastków jako część innego zadania) zamiast języków interpretowanych, takich jak Mathematica, używamy języków kompilowanych. Współcześnie wszędzie tam, gdzie kluczowa jest szybkość obliczeń prym wiedzie język C++, choć na przykład fizycy preferują jeszcze szybszy, ale nieco już sędziwy język Fortran. Najszybszy kod można uzyskać stosując bezpośrednio język maszynowy, ale pisanie w nim programów jest wyjątkowo żmudne.

Listingi 3.4 i 3.5 prezentują nasz algorytm liczenia pierwiastka metodą brutalnej siły napisany w języku C++. Działa on z dokładnością do 10^{-4}. Większej dokładności w oparciu o użyty w nim standardowy typ całkowitoliczbowy **int** nie da się uzyskać, bo dla $n >= 100000$ wyrażnie $p * n * n$ nie może być policzone poprawnie gdy reprezentowane jest jako typ **int**.

Odpowiednik programu z listingu 3.3 napisany w języku C++ przedstawono na listingu 3.6.

```
/*************** SqrtTwoBf.cpp *****************/
/*     Obliczanie przybliżenia pierwiastka z 2   */
/*     metodą brutalnej siły                     */
/***********************************************/

/* Tu włączamy standardową bibliotekę wejścia wyjścia */
#include <iostream>
/* I podpinamy tak zwaną standardową przestrzeń nazw.
   Bez tego nazwy takie jak cout poniżej musielibyśmy
   pisać jako std::cout */
using namespace std;

/* Właściwy początek programu.
   Program główny w C++ ma zawsze nazwę main.
   Pomiędzy nawiasami ( i ) może być lista ewentualnych
   argumentów (danych wejściowych). W naszym przypadku
   jest ona pusta. Słowo int przed main oznacza typ
   zwracanej wartości, który musi tu być, ale nas
   nie interesuje. Sam program mieści się pomiędzy
   nawiasami { i }                                  */
int main(){

  /* Deklarujemy i inicjalizujemy trzy zmienne typu int */
  int a=2,n=10000,i=n;

  /* Jak długo spełniony jest warunek zwiększamy zmienną i.
     Taka konstrukcja nazywana jest pętlą while.
     W tej pętli wykonywana jest tylko jedna instrukcja,
     zwiększenia zmiennej i o jeden. Znak = nie pełni tu
     roli porównania (wtedy dostalibyśmy sprzeczność 0=1)
     tylko przypisania: zawartość i zastąp przez i+1 */
  while(i*i<=a*n*n) i=i+1;

  /* Ostatnia dobra wartość zmiennej i
     była mniejsza o jeden. Poprawiamy  */
  i=i-1;

  /* Wyprowadzamy zmienną i, tekst "/", zmienną n
     oraz znak nowej linii na standardowe wyjście */
  cout << i << "/" << n << endl;
  /* Koniec programu */
}
```

Listing 3.4: Obliczania pierwiastka metodą brutalnej siły w C++

```
/*************** SqrtBfPlus.cpp ******************/
/* Usprawnienie programu SqrtTwoBf.cpp            */
/* Zamiast zaczynać dobieranie wartości zmiennej i */
/* od 1, zaczynamy od poprzedniej znalezionej     */
/* wartości                                       */
/**************************************************/

#include <iostream>
/* Włączenie dodatkowej biblioteki zawierającej
   m.in. setprecision wyjaśnione dalej */
#include <iomanip>
using namespace std;

/* Podprogram obliczający pierwiastek z liczby a.
   Przyjmuje on wartość jako argument typu double.
   Wartość zwracana jest też typu double */
double SqrtBfPlus(double a){
  /* Ustawiamy precyzję na 14 cyfr */
  double p=100000000000000.0;
  /* Deklarujemy dwie zmienne lokalne typu zmiennopozycyjnego */
  double n=1,i=1;
  /* Pętlę przerwiemy instrukcją break poniżej */
  while(true){
    /* Dopóki kwadrat z i/n nie przekracza a zwiększamy i */
    while(i*i<=a*n*n) i=i+1;
    /* Ostatnie dobre i bylo mniejsze o jeden */
    i=i-1;
    /* i/n jest naszym kolejnym przybliżeniem pierwiastka */
    cout << i/n << endl;
    /* Gdy osiągniemy żądaną precyzję, przerywamy pętlę */
    if(n>=p) break;
    /* Zwiększamy n oraz i 10-krotnie */
    n=n*10;
    i=i*10;
    /* Koniec bloku. Zawracamy i podejmujemy próbę
       kolejnego przejścia instrukcji w bloku.
       Ponieważ n w każdym kolejnym przejściu
       jest 10-krotnie większe, nie będziemy
       tego na szczęście robić w nieskończoność ... */
  }
  /* Zwracamy wynik */
  return i/n;
  /* Koniec podprogramu */
}
/* Program główny */
int main(){
  /* Ustalamy dokładność wyprowadzanych wyników na 15 cyfr */
  cout << setprecision(15);

  /* Deklarujemy jedną zmienną typu double */
  double a;
  /* Wyprowadzamy tekst na standardowe wyjście */
  cout << "Podaj liczbe naturalna: ";
  /* Wprowadzamy wartość zmiennej a z klawiatury */
  cin >> a;
  /* Uruchamiamy podprogram */
  cout << "Pierwiastek z " << a << " wynosi " << SqrtBfPlus(a) << "\n";
  /* Koniec programu */
}
```

Listing 3.5: Usprawnione podejście do obliczania pierwiastka w C++

```
/*************** SqrtBfQuick.cpp ******************/
/*      Szybki sposób liczenia pierwiastka        */
/*************************************************/
#include <iostream>
#include <iomanip>
using namespace std;

/* Podprogram obliczający wartość bezwzględną */
double abs(double x){
  if(x>=0) return x;
  else return -x;
}
/* Podprogram obliczający pierwiastek z liczby a */
double SqrtBfQuick(double a){
  double x=a;
  /* Dopóki wartość bezwzględna różnicy kwadratu przekracza 10
     w potędze -15 zapisane w postaci wykładniczej jako 1e-15   */
  while(abs(x*x-a)>1e-15){
    x=(x*x+a)/(2*x);
    cout << x << endl;
  }
  return x;
}
/* Początek głównego programu */
int main(){
  cout << setprecision(15);
  double a;
  cout << "Podaj liczbe dodatnia: ";
  cin >> a;
  cout << "Pierwiastek z " << a << " wynosi " << SqrtBfQuick(a) << "\n";
}
```

Listing 3.6: Szybkie wyznaczenie ciągu przybliżeń pierwiastka w C++.

Rozdział 4

Liczby rzeczywiste i liczby zespolone

4.1 Zbiór liczb rzeczywistych

Bardzo użyteczną własnością dodawania liczb jest zachowywanie nierówności przy operacjach arytmetycznych. Jeśli $x \leqslant y$ to $x + z \leqslant y + z$. Z mnożeniem jest podobnie. Nierówność $x \leqslant y$ implikuje $xz \leqslant yz$, ale pod dodatkowym warunkiem, że $z > 0$. Jeśli $z < 0$, to nie wszystko jest stracone: nierówność się nie zachowuje, ale pojawia się nierówność przeciwna: $xz \geqslant yz$. Przypadek $z = 0$ jest specyficzny: przemnożenie przez zero powoduje, że po obu stronach pojawia się zero. Te użyteczne własności są konsekwencją faktu, że definicja porządku w zbiorach liczbowych posługuje się operacją dodawania, a mnożenie jest powiązane z dodawaniem własnością rozdzielności. Jednak jeśli mamy jakieś ciało, a w nim jakiś porządek, to nie mamy gwarancji, że nierówności się zachowują przy wykonywaniu operacji dodawania lub mnożenia w ciele. Dlatego warto wyróżnić te ciała wyposażone w relację porządkującą, dla których zachowywanie nierówności ma miejsce.

4.1.1 Ciała uporządkowane

Definicja 4.1.1 Szóstkę $(F, +, \cdot, 0, 1, \leqslant)$ nazywamy ciałem uporządkowanym, jeżeli $(F, +, \cdot, 0, 1)$ jest ciałem, $(F, \leqslant)$ jest przestrzenią liniowo uporządkowaną oraz

(i) $\forall x, y, z \in F \quad x \leqslant y \;\Rightarrow\; x + z \leqslant y + z$,

(ii) $\forall x, y \in F \quad x > 0,\, y > 0 \;\Rightarrow\; xy > 0$.

Wprowadzamy oznaczenia

$$\mathbb{Q}^+ := \{\, q \in \mathbb{Q} \mid q > 0 \,\}, \quad \mathbb{Q}^- := \{\, q \in \mathbb{Q} \mid q < 0 \,\}.$$

Ćwiczenie 4.1.2 Pokazać, że zbiór liczb wymiernych $\mathbb{Q}$ wraz z określonymi wcześniej działaniami dodawania i mnożenia oraz relacją porządku jest ciałem uporządkowanym.

Ćwiczenie 4.1.3 Pokazać, że ciało z ćw. 3.4.2 z porządkiem $\{0 \leqslant 0, 0 \leqslant 1, 1 \leqslant 1\}$ nie jest ciałem uporządkowanym.

4.1.2 Zbiór liczb rzeczywistych

Jakkolwiek zdefiniujemy zbiór liczb rzeczywistych, oczekujemy by miał on strukturę ciała uporządkowanego. Wiemy, że ciało liczb wymiernych z naturalnym porządkiem jest ciałem uporządkowanym. Pamiętamy też, że od zbioru liczb rzeczywistych oczekujemy by, w przeciwieństwie do zbioru liczb wymiernych, był przestrzenią uporządkowaną ciągłą. Powstaje pytanie, czy nie żądamy za dużo, to znaczy czy koncepcja ciała uporządkowanego ciągłego nie jest sprzeczna sama w sobie, a jeśli nie jest, to czy mogą być ciała uporząkowane istotnie różne od liczb rzeczywistych. Odpowiedź zgodną z naszymi oczekiwaniami daje następujące twierdzenie.

Twierdzenie 4.1.4 *Istnieje ciało uporządkowane ciągłe. Jeśli F_1 i F_2 są dwoma takimi ciałami, to istnieje silnie rosnąca bijekcja $f : F_1 \to F_2$ taka, że*

$$\forall x, y \in F_1 \quad f(x+y) = f(x) + f(y) \text{ oraz } f(x \cdot y) = f(x) \cdot f(y).$$

Bijekcja f w powyższym twierdzeniu daje sposób utożsamienia dwóch formalnie różnych ciał uporządkowanych. Można więc powiedzieć, że z dokładnością do utożsamienia poprzez bijekcję f istnieje dokładnie jedno ciało uporządkowane ciągłe. Jest nim intuicyjnie rozważany od rozdziału wstępnego zbiór liczb rzeczywistych z naturalnymi działaniami dodawania i mnożenia oraz naturalnym porządkiem.

Definicja 4.1.5 Ciałem liczb rzeczywistych nazywamy ciało uporządkowane ciągłe. Ciało to oznaczamy $\mathbb{R}$.

4.1.3 (**)Konstrukcja Cauchy'ego

Dowód tw. 4.1.4 nie jest prosty. Ograniczymy się jedynie do podania szkicu konstrukcji liczb rzeczywistych pochodzącej od Cauchy'ego.

Pierwsze konstrukcje ciała liczb rzeczywistych podali Cauchy, Dedekind i Weierstrass. Konstrukcja Cauchy'ego opiera się na pojęciu ciągu Cauchy'ego.

Definicja 4.1.6 Ciąg $x : \mathbb{N} \to \mathbb{Q}$ nazywamy *ciągiem Cauchy'ego* jeżeli

$$\forall \epsilon \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \forall m, n \in \mathbb{N} \quad m, n \geqslant N \;\Rightarrow\; |x_m - x_n| < \epsilon.$$

Dla ciągów $x, y : \mathbb{N} \to \mathbb{Q}$ definiujemy sumę $x + y$ i iloczyn $x \cdot y$ wzorami

$$\begin{aligned} x + y &: \mathbb{N} \ni n \mapsto x_n + y_n \in \mathbb{Q}, \\ x \cdot y &: \mathbb{N} \ni n \mapsto x_n \cdot y_n \in \mathbb{Q} \end{aligned}$$

oraz nierówność

$$x \leqslant y \;:\Leftrightarrow\; \forall \epsilon \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \forall n \in \mathbb{N} \quad n \geqslant N \;\Rightarrow\; x_n - \epsilon \leqslant y_n.$$

W zbiorze ciągów Cauchy'ego wprowadzamy relację równoważności

$$x \sim y \;:\Leftrightarrow\; \forall \epsilon \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \forall n \in \mathbb{N} \quad n \geqslant N \;\Rightarrow\; |x_n - y_n| < \epsilon.$$

Zbór liczb rzeczywistych według Cauchy'ego definiujemy jako

$$\mathbb{R} := \{\, [x]_\sim \mid x : \mathbb{N} \to \mathbb{Q} \text{ ciąg Cauchy'ego.} \,\}$$

Liczbę wymierną $q \in Q$ utożsamiamy z klasą równoważności ciągu stale równego q.

4.1.4 Podzbiory $\mathbb{R}$

Zbiór $\mathbb{N}$ możemy utożsamiać ze zbiorem

$$\{\, x \in \mathbb{R} \mid \exists n \in \mathbb{N}\ x = n \cdot 1 \,\}.$$

Zbiór $\mathbb{Z}$ możemy utożsamiać ze zbiorem

$$\{\, x \in \mathbb{R} \mid x \in \mathbb{N} \text{ lub } -x \in \mathbb{N} \,\}.$$

Zbiór $\mathbb{Q}$ możemy utożsamiać ze zbiorem

$$\{\, x \in \mathbb{R} \mid \exists p \in \mathbb{Z}\, \exists q \in \mathbb{N}_1\ x = p \cdot q^{-1} \,\}.$$

Wprowadzamy oznaczenia

$$\begin{aligned} \mathbb{R}^+ &:= \{\, x \in \mathbb{R} \mid x > 0 \,\}, \\ \mathbb{R}^- &:= \{\, x \in \mathbb{R} \mid x < 0 \,\}, \\ \mathbb{R}^* &:= \{\, x \in \mathbb{R} \mid x \geqslant 0 \,\}. \end{aligned}$$

4.1.5 Funkcja potęgowa w $\mathbb{R}$

Uwaga 4.1.7 *Mamy następujące własności funkcji potęgowej w* $\mathbb{R}$

(i) $\forall x \in \mathbb{R} \quad x^2 \geqslant 0$,

(ii) $\forall x \in \mathbb{R} \quad x^2 = 0 \Leftrightarrow x = 0$,

(iii) $\forall x, y, z \in \mathbb{R} \quad x > 0,\ y < z \Rightarrow xy < xz$,

(iv) $\forall x, y, z \in \mathbb{R} \quad x < 0,\ y < z \Rightarrow xy > xz$,

(v) $\forall x \in \mathbb{R} \quad x > 0 \Rightarrow x^{-1} > 0$.

Dowód: ćwiczenie.

Twierdzenie 4.1.8 *Funkcja potęgowa na zbiorze* $\mathbb{R}^*$ *jest silnie rosnąca, tzn.*

$$\forall n \in \mathbb{N}_1 \forall x, y \in \mathbb{R}^* : \quad 0 \leqslant x < y \Rightarrow x^n < y^n$$

Dowód: Dowód poprowadzimy indukcyjnie względem n. Dla $n = 1$ teza pokrywa się z założeniem, więc implikacja jest oczywiście prawdziwa. Załóżmy zatem, że dla pewnego $k \in \mathbb{N}$ mamy

$$\forall x, y \in \mathbb{R}^* \ \ x < y \Rightarrow x^k < y^k. \tag{4.1}$$

Chcemy pokazać, że

$$\forall x, y \in \mathbb{R}^* \ \ x < y \Rightarrow x^{k+1} < y^{k+1}. \tag{4.2}$$

Rozważmy więc $x, y \in \mathbb{R}$ takie, że $x < y$. Z założenia indukcyjnego (4.1) wnosimy, że $x^k < y^k$. Stosując dwukrotnie własność (ii) ciała uporządkowanego dostajemy

$$x^{k+1} = x^k x < y^k x < y^k y = y^{k+1}.$$

Zatem $x^{k+1} < y^{k+1}$ co dowodzi, że implikacja (4.2) zachodzi. Zatem na mocy zasady indukcji twierdzenie zostało udowodnione. □

4.1.6 Nierówność Bernoullego

Poniższa nierówność przyda się nam w dowodzie twierdzenia 4.6.2, o zbieżności ciągu o wyrazach $\left(1+\frac{1}{n}^n\right)$. Jej autorem jest słynny matematyk szwajcarski Jacob Bernoulli(rys. 4.7).

Twierdzenie 4.1.9

$$\forall n \in \mathbb{N}_1 \forall x \in \mathbb{R}: \quad x > -1 \;\Rightarrow\; (1+x)^n \geqslant 1+nx$$

Dowód: (*) Dowód poprowadzimy indukcyjnie względem n. Dla $n=1$ nierówność jest oczywista. Załóżmy, że zachodzi ona dla pewnego $k \in \mathbb{N}$ i wszystkich $x > -1$. Ustalmy $y > -1$. Na mocy założenia

$$(1+y)^k \geqslant 1+ky. \tag{4.3}$$

Ponieważ $y+1 > 0$, więc mnożąc nierówność (4.3) obustronnie przez $y+1$ dostajemy

$$(1+y)^{k+1} \geqslant (1+ky)(1+y) = 1+(k+1)y+ky^2 \geqslant 1+(k+1)y,$$

co dowodzi, że teza zachodzi dla $n = k+1$ i kończy dowód indukcyjny. □

4.1.7 Wartość bezwzględna

Definicja 4.1.10 Dla $x \in \mathbb{R}$ definiujemy *wartość bezwzględną* x jako

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

Twierdzenie 4.1.11 *Niech $x, y, a \in \mathbb{R}$. Wtedy*

$$|x| = |-x|, \quad |x| \geqslant 0, \quad -|x| \leqslant x \leqslant |x|, \tag{4.4}$$

$$|x| = 0 \;\Leftrightarrow\; x = 0, \tag{4.5}$$

$$|x| \leqslant a \;\Leftrightarrow\; -a \leqslant x \leqslant a, \quad |x| \geqslant a \;\Leftrightarrow\; x \leqslant -a \text{ lub } x \geqslant a, \tag{4.6}$$

$$||x|-|y|| \leqslant |x \pm y| \leqslant |x|+|y|, \tag{4.7}$$

$$|xy| = |x||y|, \quad \sqrt{x^2} = |x|. \tag{4.8}$$

Dowód: Własności (4.4), (4.5), (4.6) i (4.8) są oczywiste gdy rozważy się oddzielnie przypadki $x \geqslant 0$ i $x < 0$. Aby pokazać (4.7) najpierw udowodnimy, że dla dowolnych $x, y \in \mathbb{R}$ mamy

$$|x+y| \leqslant |x|+|y|. \tag{4.9}$$

Z własności (4.4) mamy

$$-|x| \leqslant x \leqslant |x| \quad \text{i} \quad -|y| \leqslant y \leqslant |y|.$$

Z własności ciała uporządkowanego mamy zatem

$$-(|x|+|y|) \leqslant x+y \leqslant |x|+|y|.$$

Zatem z (4.6) dostajemy (4.9). Ponieważ (4.9) zachodzi dla dowolnych $x, y \in \mathbb{R}$, więc w szczególności możemy podstawić w (4.9) $x-y$ za x. W efekcie dostajemy

$$|x| = |(x-y)+y| \leqslant |x-y|+|y|.$$

Podobnie otrzymujemy

$$|y| \leqslant |y-x| + |x|,$$

skąd wnosimy, że

$$|x| - |y| \leqslant |x-y| \quad \text{i} \quad |y| - |x| \leqslant |y-x| = |x-y|.$$

Zatem z (4.6) dostajemy

$$||x| - |y|| \leqslant |x-y|,$$

a także

$$||x| - |y|| = ||x| - |-y|| \leqslant |x-(-y)| = |x+y|,$$

co kończy dowód własności (4.7).

4.2 Twierdzenie o istnieniu pierwiastka.

Naszym celem w tym rozdziale jest udowodnienie, że w zbiorze liczb rzeczywistych dodatnich pierwiaski istnieją. W szczególności ostatecznie zamkniemy sprawę sensowności $\sqrt{2}$. Nim jednak twierdzenie to udowodnimy potrzebujemy dwa pomocnicze twierdzenia: własność Archimedesa oraz gęstość $\mathbb{Q}$ w $\mathbb{R}$.

4.2.1 Własność Archimedesa

Twierdzenie 4.2.1 *Ciało liczb rzeczywistych ma własność Archimedesa, tzn.*

$$\forall a, b \in \mathbb{R}^+ \; \exists n \in \mathbb{N} \quad na > b.$$

Dowód: (*) Dowód poprowadzimy nie wprost. Załóżmy zatem, że istnieją liczby rzeczywiste dodatnie a, b takie, że dla każdej liczby naturalnej n zachodzi $na \leqslant b$. Oznacza to, że b jest majorantą zbioru

$$A := \{\, x \in \mathbb{R} \mid \exists n \in \mathbb{N}\; x = na \,\}.$$

Niech $c := \sup A$ będzie jego kresem gónym. Ponieważ $c - a < c$, więc $c - a$ nie jest majorantą zbioru A. Zatem znajdziemy $n_0 \in \mathbb{N}$ takie, że $n_0 a > c - a$. Ale wtedy $(n_0 + 1)a > c$, a w konsekwencji c nie jest majorantą A, a tym bardziej kresem górnym zbioru A. Otrzymana sprzeczność dowodzi twierdzenia. □

Wniosek 4.2.2 *Ciało liczb wymiernych ma wartość Archimedesa.*

Dowód: ćwiczenie.

4.2.2 Gęstość zbioru w zbiorze

Definicja 4.2.3 Niech X będzie przestrzenią liniowo uporządkowaną, a A niech będzie podzbiorem X. Mówimy, że A jest *gęsty* w X jeżeli

$$\forall x, y \in X \quad x < y \;\Rightarrow\; \exists z \in A \; x < z < y.$$

Gdy $A = X$, mówimy, że A jest *gęsty w sobie*.

Twierdzenie 4.2.4 *Zbiór liczb wymiernych jest gęsty w zbiorze liczb rzeczywistych.*

Dowód: Niech $x, y \in \mathbb{R}$ będą takie, że $x < y$. Możliwe są trzy przypadki

$$x < 0 < y \text{ lub } 0 \leqslant x < y \text{ lub } x < y \leqslant 0.$$

W pierwszym przypadku teza jest oczywista, bo zero jest liczbą wymierną. Rozważmy drugi przypadek. Niech $a := y - x > 0$. Z własności Archimedesa znajdziemy liczbę naturalną n taką, że $na > 1$. Wtedy $\frac{1}{n} < a$. Również z własności Archimedesa wniosimy, że zbiór

$$A := \{\, k \in \mathbb{N} \mid k\tfrac{1}{n} > x \,\}$$

jest niepusty. Niech $m := \min A$. Wtedy $m - 1 \notin A$. W szczególności $\frac{m-1}{n} \leqslant x$ oraz

$$\frac{m}{n} = \frac{m-1}{n} + \frac{1}{n} \leqslant x + \frac{1}{n} < x + a = y.$$

Zatem $x < \frac{m}{n} < y$, a ponieważ oczywiście $\frac{m}{n} \in \mathbb{Q}$, teza w tym przypadku zachodzi. Pozostaje rozważyć trzeci przypadek. Wtedy $0 \leqslant -y < -x$, zatem z rozważonego już przypadku znajdziemy $z \in \mathbb{Q}$ takie, że $-y < z < -x$. Ale wtedy $x < -z < y$ i oczywiście $-z \in \mathbb{Q}$ co kończy dowód twierdzenia. □

4.2.3 Pierwiastki

Teraz możemy już udowodnić twierdzenie o istnieniu n-tego pierwiastka

Twierdzenie 4.2.5 *(o istnieniu n-tego pierwiastka)*

$$\forall n \in \mathbb{N}_1 \forall x \in \mathbb{R}^* \exists! y \in \mathbb{R}^* : \quad y^n = x$$

Dowód: (*) Pokażemy najpierw, że pierwiastek istnieje co najwyżej jeden. Dla dowodu nie wprost, przyjmijmy, że istnieją $y_1, y_2 \in \mathbb{R}^*$ takie, że $y_1 \neq y_2$ oraz $y_1^n = y_2^n$. Dla ustalenia uwagi możemy przyjąć, że $y_1 < y_2$ (dowód gdy $y_1 > y_2$ jest analogiczny). Na mocy twierdzenia 4.1.8 mamy $y_1^n < y_2^n$ co przeczy $y_1^n = y_2^n$ i dowodzi jednoznaczności.

Dla dowodu istnienia zauważmy najpierw, że $0^n = 0$, zatem teza jest oczywista gdy $x = 0$. Możemy więc założyć, że $x > 0$. Rozważmy zbiór

$$E := \{\, u \in \mathbb{R}^* \mid u^n < x \,\}.$$

Oczekujemy, że kres górny tego zbioru jest poszukiwanym pierwiastkiem. Musimy jednak najpierw pokazać, że E jest zbiorem niepustym i ograniczonym od góry.

Pokażemy, że $E \neq \varnothing$. Niech $t := \frac{x}{x+1}$. Mamy $0 < t < 1$ oraz $t < x$. Mnożąc nierówności $0 < t < 1$ obustronnie przez $t > 0$ dostajemy z własności ciała uporządkowanego dostajemy $0 < t^2 < t$. Ale $t < 1$, zatem $0 < t^2 < 1$. Rozumując indukcyjnie stwierdzamy, że również $0 < t^{n-1} < 1$. Mnożąc tę nierówność obustronnie przez $t >$) dostajemy $t^n < t$. Zatem $t^n < x$ co oznacza, że $t \in E$. Tak więc $E \neq \varnothing$. Pokażemy z kolei, że $1 + x$ jest majorantą zbioru E. Gdyby tak nie było, znaleźlibyśmy $s \in E$, takie, że $s > 1 + x$. Wtedy $s^n > s > x$. Zatem $s \notin E$. Otrzymana sprzeczność pokazuje, że $1 + x$ rzeczywiście jest majorantą zbioru E, zatem E jest ograniczony od góry. Zatem ma kres górny.

Połóżmy $y := \sup E$. Pokażemy, że rzeczywiście $y^n = x$. Dla dowodu nie wprost załóżmy, że $y^n \neq x$. Mamy do rozważenia dwa przypadki: $y^n < x$ lub $y^n > x$. Rozważmy najpierw przypadek $y^n < x$.

Liczymy teraz, że da się skonstruować $y_1 > y$ takie, że $y_1^n < y$, co przeczyłoby $y = \sup E$. Naturalne jest poszukiwać y_1 w postaci $y_1 = y + h$. Pytanie jak dobrać h? Mamy mieć $(y + h)^n < x$ co jest równoważne

$$(y + h)^n - y^n < x - y^n. \tag{4.10}$$

Wiemy ze wzoru (3.8), że

$$(y + h)^n - y^n = h\left((y + h)^{n-1} + (y + h)^{n-2}y + \cdots y^{n-1}\right) < hn(y + h)^n.$$

Zatem, dla $h < 1$ mamy

$$(y + h)^n - y^n < hn(y + 1)^n.$$

Stąd, by zagwarantować (4.10), wystarczy zapewnić, że

$$hn(y+1)^n < x - y^n,$$

a w tym celu wystarczy wziąć $h < \frac{x-y^n}{n(y+1)^n}$. Jesteśmy teraz gotowi, by ze zrozumieniem wrócić do formalnego dowodu.

Dobierzmy $h \in \mathbb{Q}$ takie, że

$$0 < h < \min\{1, \frac{x-y^n}{n(y+1)^n}\}.$$

Istnienie takiego h zapewnia nam gęstość $\mathbb{Q}$ w $\mathbb{R}$ (twierdzenie 4.2.4). Łatwo sprawdzamy, że wtedy

$$(y+h)^n - y^n < hn(y+1)^n < x - y^n.$$

Zatem $(y+h) < x$, czyli $y+h \in E$. Ale $y+h > y = supE$, skąd $y+h \notin E$. Otrzymana sprzeczność wyklucza nierówność $y^n < x$.

Pozostaje pokazać, że nierówność $y^n > x$ też prowadzi do sprzeczności. *Tym razem liczymy, że uda się skonstruować $y_2 < y$, również będące majorantą zbioru E, co doprowadzi do sprzeczności z faktem, że y jako kres górny jest najmniejszą majorantą. Poszukujemy y_2 w postaci $y-k$. Chcemy mieć $(y-k)^n > x$ co jest równoważne postulatowi*

$$y^n - (y-k)^n < y^n - x$$

Rozpisując podobnie jak poprzednio mamy

$$y^n - (y-k)^n = k\left(y^{n-1} + y^{n-2}(y-k) + \cdots + (y-k)^{n-1}\right) < kny^{n-1}.$$

Podpowiada to jak dobrać k.

Dobierzmy $k \in \mathbb{Q}$ takie, że

$$0 < k < \frac{y^n - x}{ny^{n-1}}.$$

Łatwo sprawdzamy, że przy tak dobranym k mamy $(y-k)^n > x$. Twierdzimy, że $y-k$ jest majorantą zbioru E. Gdyby tak nie było, znaleźlibyśmy $v \in E$ takie, że $v > y-k$. Ale wtedy $v^n > (v-k)^n > x$, co przeczy $v \in E$. Zatem $y-k$ rzeczywiście jest majorantą zbioru E. Ale $y-k < y$, skąd y nie jest najmniejszą majorantą $E<$ zatem $y \neq \sup E$. Otrzymana sprzeczność dowodzi, że $y^n = x$. □

Nieujemną liczbę, której n-ta potęga jest równa liczbie $x \geqslant 0$ nazywamy n-tym *pierwiastkiem* z x i oznaczamy $\sqrt[n]{x}$.

Twierdzenie 4.2.6

$$\forall n \in \mathbb{N}_1 \forall x, y \in \mathbb{R}^* : \quad \sqrt[n]{xy} = \sqrt[n]{x}\sqrt[n]{y}$$

Dowód: ćwiczenie.

W programie Mathematica pierwiastek kwadratowy z x zapisujemy jako `Sqrt[x]`. Nie ma natomiast podobnej notacji dla pierwiastków stopnia wyższego niż dwa. Aby policzyć $\sqrt[n]{x}$ piszemy `x^(1/n)`. *Notacja ta bierze się stąd, że, jak wkrótce zobaczymy, funkcję potęgową można rozszerzyć na wykładniki rzeczywiste i wtedy $\sqrt[n]{x} = x^{\frac{1}{n}}$.*

4.3 Rozszerzony zbiór liczb rzeczywistych

4.3.1 $\mathbb{R}$ uzupełniony o $-\infty$ i $+\infty$

Aby brać kres górny lub dolny podzbioru $\mathbb{R}$ trzeba sprawdzać czy jest on ograniczony odpowiednio od góry lub dołu. Jest to niewygodnie, dlatego przydaje się zbiór liczb rzeczywistych rozszerzony o plus i minus nieskończoność. W takim rozszerzonym zbiorze każdy podzbiór ma kres górny lub dolny. Jest nim zwykła liczba gdy zbiór jest ograniczony, a plus lub minus nieskończoność jeśli jest nieograniczony od góry lub nieograniczony od dołu.

Definicja 4.3.1 Definiujemy rozszerzony zbiór liczb rzeczywistych jako

$$\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}.$$

Ponadto przyjmujemy, że dla każdego $x \in \mathbb{R} \quad -\infty < x < \infty$.

Uwaga 4.3.2 *Każdy niepusty podzbiór $\bar{\mathbb{R}}$ posiada kresy.*

Definicja 4.3.3 W $\bar{\mathbb{R}}$ definiujemy następujące działania:

(i) Dla $x \in \mathbb{R} \quad x + \infty := \infty,\ x - \infty := -\infty,\ \frac{x}{\infty} := \frac{x}{-\infty} := 0.$

(ii) Dla $x \in \mathbb{R}^+ \quad x \cdot \infty := \infty,\ x \cdot (-\infty) := -\infty.$

(iii) Dla $x \in \mathbb{R}^- \quad x \cdot \infty := -\infty,\ x \cdot (-\infty) := \infty.$

(iv) Ponadto $\infty \cdot \infty := (-\infty) \cdot (-\infty) := \infty,\ \infty \cdot (-\infty) := (-\infty) \cdot \infty := -\infty.$

(v) Oraz $\infty + \infty := \infty,\ (-\infty) + (-\infty) := -\infty.$

Uwaga 4.3.4 *Jeśli $A \subset \bar{\mathbb{R}}$ jest zbiorem niepustym, to A jest ograniczony od góry (od dołu) wtedy i tylko wtedy gdy* $\sup A < \infty$ *(*$\inf A > -\infty$*).*

4.3.2 Kresy funkcji

Definicja 4.3.5 Niech X będzie zbiorem, $A \subset X$ jego podzbiorem, a $f : X \to \bar{\mathbb{R}}$ funkcją. Definiujemy *kres górny funkcji f na zbiorze A* jako

$$\sup_{x \in A} f(x) := \sup f(A) = \sup \{\, f(x) \mid x \in A \,\}.$$

Definiujemy *kres dolny funkcji f na zbiorze A* jako

$$\inf_{x \in A} f(x) := \inf f(A) = \inf \{\, f(x) \mid x \in A \,\}.$$

Mówimy, że f jest *ograniczona od góry (od dołu)* na zbiorze A gdy $f(A)$ jest ograniczony od góry (od dołu) co jest równoważne warunkowi $\sup_{x \in A} f(x) < \infty$ $(\inf_{x \in A} f(x) > -\infty)$.

Twierdzenie 4.3.6 *Mamy następujące własności kresów funkcji:*

(i) Niech $f : X \to \bar{\mathbb{R}}$ będzie funkcją, a $A \subset X$. Wtedy

$$\begin{aligned} \inf_{x\in A}(-f(x)) &= -\sup_{x\in A} f(x), \\ \sup_{x\in A}(-f(x)) &= -\inf_{x\in A} f(x). \end{aligned}$$

(ii) Niech $f_i : X \to \bar{\mathbb{R}}$ dla $i \in \{1, 2\}$ będą funkcjami, a $A \subset X_i$. Wtedy

$$\begin{aligned} \sup_{x\in A}(f_1(x) + f_2(x)) &\leqslant \sup_{x\in A} f_1(x) + \sup_{x\in A} f_2(x), \\ \inf_{x\in A}(f_1(x) + f_2(x)) &\geqslant \inf_{x\in A} f_1(x) + \inf_{x\in A} f_2(x). \end{aligned}$$

Dowód: (*)Aby udowodnić własność (i) połóżmy $a := \sup_{x\in A} f(x)$ oraz $b := \inf_{x\in A}(-f(x))$. Wystarczy pokazać, że $-a \leqslant b$ i $-a \geqslant b$. Niech $x \in A$. Wtedy z definicji kresu górnego $f(x) \leqslant a$. Zatem $-a \leqslant -f(x)$ dla wszystkich elementów $x \in A$. Zatem $-a$ jest minorantą zbioru $\{ -f(x) \mid x \in A \}$, skąd $-a \leqslant \inf_{x\in A}(-f(x)) = b$. Potrzebujemy jeszcze pokazać, że $-a \geqslant b$. Ustalmy ponownie $x \in A$. Z definicji kresu dolnego mamy $b \leqslant -f(x)$, skąd $f(x) \leqslant -b$. Zatem $-b$ jest majorantą zbioru $\{ f(x) \mid x \in A \}$, co oznacza, że $-b \geqslant a$, czyli $-a \geqslant b$. Dowodzi to pierwszej równości własności (i). Drugą dowodzi się analogicznie.

Dla dowodu pierwszej nierówności własności (ii) połóżmy $a_1 := \sup_{x\in A} f_1(x)$, $a_2 := \sup_{x\in A} f_2(x)$ oraz ustalmy $x \in A$. Z definicji kresu górnego mamy

$$f_1(x) \leqslant a_1 \quad f_2(x) \leqslant a_2.$$

Zatem

$$f_1(x) + f_2(x) \leqslant a_1 + a_2.$$

Ponieważ $x \in A$ jest dowolnie ustalone, wnosimy, że $a_1 + a_2$ jest majorantą zbioru $\{ f_1(x) + f_2(x) \mid x \in A \}$. Ponieważ $\sup_{x\in A}(f_1(x) + f_2(x))$ jest najmniejszą majorantą tego zbioru, dostajemy

$$\sup_{x\in A}(f_1(x) + f_2(x)) \leqslant a_1 + a_2,$$

co dowodzi pierwszej nierówności we własności (ii). Drugą nierówność dowodzi się analogicznie. □

W podobny sposób dowodzi się następujących twierdzeń.

Twierdzenie 4.3.7 *Niech $f : X_1 \times X_2 \to \bar{\mathbb{R}}$ będzie funkcją, a $A_i \subset X_i$ dla $i \in \{1, 2\}$. Wtedy*

$$\begin{aligned} \sup_{(x_1,x_2)\in A_1\times A_2} f(x_1, x_2) &= \sup_{x_1\in A_1}\left(\sup_{x_2\in A_2} f(x_1, x_2)\right), \\ \inf_{(x_1,x_2)\in A_1\times A_2} f(x_1, x_2) &= \inf_{x_1\in A_1}\left(\inf_{x_2\in A_2} f(x_1, x_2)\right). \end{aligned}$$

Twierdzenie 4.3.8 *Niech $f_i : X_i \to \bar{\mathbb{R}}$ oraz $A_i \subset X_i$ dla $i \in \{1, 2\}$. Jeśli $\inf_{x_i\in A_i} f_i(x_i) > -\infty$ oraz $\sup_{x_i\in A_i} f_i(x_i) < \infty$ to*

$$\begin{aligned} \sup_{(x_1,x_2)\in A_1\times A_2}(f_1(x_1) + f_2(x_2)) &= \sup_{x_1\in A_1} f_1(x_1) + \sup_{x_2\in A_2} f_2(x_2), \\ \inf_{(x_1,x_2)\in A_1\times A_2}(f_1(x_1) + f_2(x_2)) &= \inf_{x_1\in A_1} f_1(x_1) + \inf_{x_2\in A_2} f_2(x_2). \end{aligned}$$

Twierdzenie 4.3.9 *Niech $f : X \to \bar{\mathbb{R}}$ będzie funkcją, $A \subset X$, $a\ c \in \mathbb{R}$. Wtedy*

$$\begin{aligned} \sup_{x\in A}(f(x)+c) &= \sup_{x\in A} f(x) + c, \\ \inf_{x\in A}(f(x)+c) &= \inf_{x\in A} f(x) + c. \end{aligned}$$

4.4 Ciało liczb zespolonych

4.4.1 Liczby zespolone

Chcemy rozszerzyć ciało liczb rzeczywistych tak, by pozostało ciałem, ale była w nim liczba i o własności $i^2 = -1$. Widać, że jeśli ma to być ciało to musi w nim też być liczba $a + bi$ dla dowolnych rzeczywistych a, b. Jednak suma i iloczyn liczb tej postaci nie wymaga już dalszego rozszerzania, bo

$$\begin{aligned} (a+bi)+(c+di) &= (a+c)+(b+d)i, \\ (a+bi)(c+di) &= ac+adi+bci+bdi^2 = (ac-bd)+(ad+bc)i. \end{aligned}$$

Niestety nie ma szans, by ciało to było uporządkowane, bo wtedy albo $i > 0$, albo $-i > 0$, co na mocy własności ciała uporządkowanego prowadzi do $i^2 > 0$ sprzecznego z $i^2 = -1$.

Pozostaje jednak pytanie, czy sam pomysł liczby i jako elementu ciała rozszerzającego ciało liczb rzeczywistych też nie prowadzi do jakiejś sprzeczności, której jeszcze nie widzimy. Aby przekonać się, że tak nie jest, trzeba zbudować formalny model liczb zespolonych.

4.4.2 Ciało $\mathbb{C}$

W $\mathbb{C} := \mathbb{R}^2$ wprowadzamy działania

$$\begin{aligned} (a,b)+(c,d) &:= (a+c, b+d), \\ (a,b)(c,d) &:= (ac-bd, ad+bc). \end{aligned}$$

Twierdzenie 4.4.1 *Zbiór $\mathbb{C} := \mathbb{R}^2$ z tak określonymi działaniami oraz parami $(0,0)$ i $(1,0)$ jako elementami neutralnymi odpowiednio dla dodawania i mnożenia jest ciałem.*

Definicja 4.4.2 Ciało to nazywamy *ciałem liczb zespolonych*.

Dla dowolnych $a, b \in \mathbb{R}$ mamy $(a,0)+(b,0) = (a+b,0)$ oraz $(a,0)(b,0) = (ab,0)$, zatem parę $(a,0)$ możemy utożsamiać z liczbą rzeczywistą a.

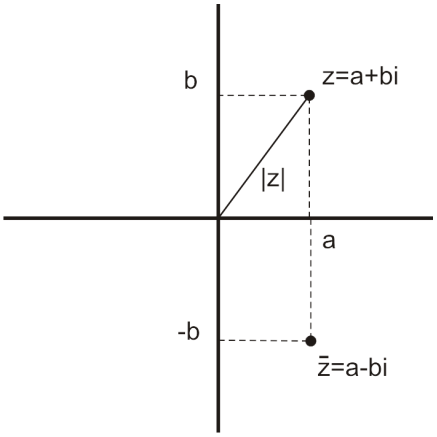
Zwyczajowo kładziemy $i := (0,1)$. Jest to tak zwana *jednostka urojona*. Łatwo sprawdzamy, że $i^2 = (-1,0) = -1$ oraz $(a,b) = (a,0)+(b,0)i = a+bi$.

Dla liczby zespolonej $z := a + bi$ definiujemy jej *moduł* $|z| := \sqrt{a^2+b^2}$ oraz jej *sprzężenie* $\bar{z} := a - bi$.

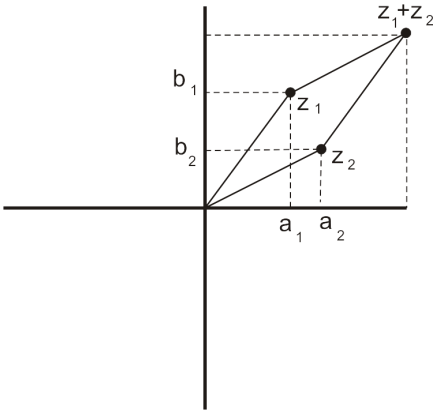
Dla liczby zespolonej $z = a+bi$ liczby $a, b \in \mathbb{R}$ nazywamy odpowiednio *częścią rzeczywistą* i *częścią urojoną* z i oznaczamy $\operatorname{re} z$ i $\operatorname{im} z$.

W programie Mathematica jednostkę urojoną i oznaczamy I, a liczbę zespoloną $a + bi$ zapisujemy jako `a + b I`.

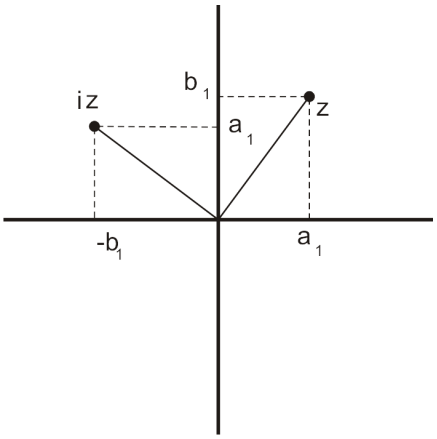
Interpretacja geometryczna liczby zespolonej przedstawiona jest na rys. 4.1, interpretacja sumy na rys. 4.2, a interpretacja iloczynu liczby zespolonej przez liczbę i na rys. 4.3.



Rysunek 4.1: Interpretacja geometryczna liczby zespolonej $z := a + bi$, jej modułu oraz sprzężenia.



Rysunek 4.2: Interpretacja geometryczna sumy liczb zespolonych.



Rysunek 4.3: Interpretacja geometryczna iloczynu liczby zespolonej przez liczbę i.

4.4.3 Zespolona funkcja kwadratowa

Wiemy ze szkoły, że miejsca zerowe funkcji kwadratowej

$$x \mapsto ax^2 + bx + c$$

zależą od tzw. wyróżnika $\Delta := b^2 - 4ac$*, w szczególności miejsc zerowych nie ma, gdy wyróżnik jest ujemny. Sytuacja ulega zmianie, gdy w funkcji tej dopuścimy argument zespolony.*

Twierdzenie 4.4.3 *Rozważmy zespoloną funkcją kwadratową o współczynnikach rzeczywistych tj. funkcję*

$$Q : \mathbb{C} \ni z \mapsto az^2 + bz + c \in \mathbb{C},$$

gdzie $a, b, c \in \mathbb{R}$ *są zadanymi współczynnikami rzeczywistymi. Niech* $\Delta := b^2 - 4ac$*. Jeśli* $\Delta > 0$*, to* Q *ma dwa rzeczywiste miejsca zerowe*

$$z_{1,2} := \frac{-b \pm \sqrt{\Delta}}{2a}.$$

Gdy $\Delta = 0$ *funkcja* Q *ma jeden pierwiatek rzeczywisty*

$$z := \frac{-b}{2a}.$$

Natomiast dla $\Delta < 0$ *mamy dwa zespolone, wzajemnie sprzężone miejsca zerowe*

$$z_{1,2} := \frac{-b \pm i\sqrt{-\Delta}}{2a}.$$

4.4.4 Liczby zespolone o module jeden.

Półprosta poprowadzona z początku układu współrzędnych pod kątem φ *do osi* X *przecina okrąg jednostkowy*

$$\mathbb{S}^1 := \{\, z \in \mathbb{C} \mid \ |z| = 1 \,\}$$

(rys. 4.4) w punkcie o współrzędnych

$$z_\varphi(\cos\varphi, \sin\varphi) \in \mathbb{C}.$$

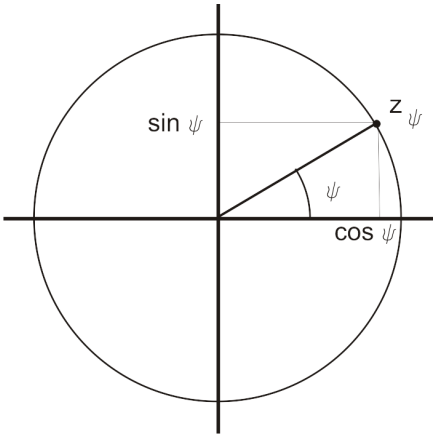
Wykorzystując tożsamości trygonometryczne dla sumy kątów można pokazać, że

$$\begin{aligned} z_{\varphi+\psi} &= \cos(\varphi+\psi) + i\sin(\varphi+\psi) \\ &= (\cos\varphi\cos\psi - \sin\varphi\sin\psi) + i(\sin\varphi\cos\psi + \cos\varphi\sin\psi) \\ &= (\cos\varphi + i\sin\varphi)(\cos\psi + i\sin\psi) \\ &= z_\varphi z_\psi. \end{aligned}$$

To przypomina fundamentalną własność funkcji wykładniczej. Matematyk szwajcarski Leonard Euler pokazał, że rzeczywistą funkcję wykładniczą można rozszerzyć na argumenty zespolone

$$e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

i przy takiej definicji dla $x \in \mathbb{R}$ *zachodzi słynny wzór o nawijaniu prostej na okrąg*



Rysunek 4.4: Liczby zespolone o module jeden.

$$e^{ix} = \cos x + i \sin x,$$

który jest wygodną podstawą tak formalnej definicji jak i własności funkcji trygonometrycznych. Aby pójść tą drogą potrzebujemy rozszerzyć pojęcie granicy ciągu i zbieżności szeregu na liczby zespolone.

4.4.5 Odległość w $\mathbb{R}$ i $\mathbb{C}$.

Występująca w definicji Cauchy'ego granicy ciągu wartość bezwzględna $|a-b|$ dla $a, b \in \mathbb{R}$ może być interpretowana jako odległość pomiędzy liczbami a i b. Uogólniając pojęcie odległości możemy się pokusić o uogólnienie pojęcia granicy.

Moduł liczby zespolonej pod wieloma względami przypomina wartość bezwzględną:

Twierdzenie 4.4.4 *Moduł liczby zespolonej ma następujące własności*

(i) $\forall z \in \mathbb{C} \quad |z| = 0 \Leftrightarrow z = 0,$

(ii) $\forall z_1, z_2 \in \mathbb{C} \quad |z_1 z_2| = |z_1||z_2|,$

(iii) $\forall z_1, z_2 \in \mathbb{C} \quad |z_1 + z_2| \leqslant |z_1| + |z_2|.$

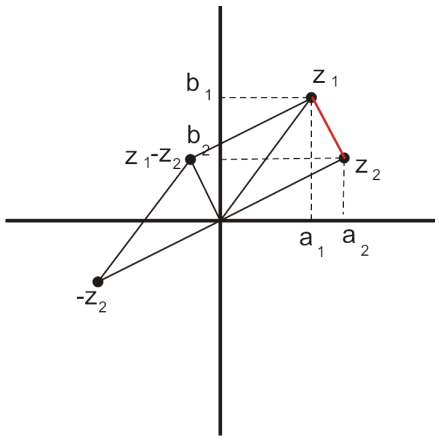
Dlatego $|z_1 - z_2|$ można interpretować jako odległość pomiędzy liczbami zespolonymi z_1 i z_2. Zilustrowano to na rys. 4.5

4.5 Granice ciągów w $\mathbb{R}$

W rozdziale 1.2 przedstawiliśmy ciągi jako narzędzie do przybliżania liczb, które nie są liczbami wymiernymi i analizowaliśmy prosty algorytm pozwalający przybliżać liczbę $\sqrt{2}$. Przedstawiliśmy też ideę granicy ciągu jako liczby, którą ciąg przybliża. Teraz przyjrzymy się temu pojęciu dokładniej, ale zaczniemy od dwóch dodatkowych eksperymentów numerycznych.

4.5.1 Eksperymenty numeryczne

Listing 4.1 prezentuje program w języku Mathematica obliczający wyrazy ciągu (1.2) rozważanego w rozdziale 1.2.5 jako narzędzie do przybliżania liczby π. Jego odpowiednik w języku C++ przedstawiono na listingu 4.2.



Rysunek 4.5: Odległość liczb zespolonych z_1 i z_2.

```
(* Przybliżanie liczby pi obwodem wpisanych wielokątów *)
PiByCircumference[d_] :=
  Module[{a, m, n, A},
    (* Zaczynamy od wpisanego kwadratu.
       Jego bok a wynosi pół pierwistka z dwóch.
       Do policzenia  pierwiastka wykorzystujemy funkcję wbudowaną Sqrt *)
    a = Sqrt[2.]/2;
    m = 4; (* Ilośc boków wielokąta, na początek cztery *)
    n = 1;       (* Numer kolejnego wyrazu ciągu, na początek jeden *)
    Do[(*Początek pętli Do*)
      A=m*a;     (* Obliczamy długość obwodu *)
      (* Domyślnie dla liczb maszynowych Mathematica wypisuje tylko 6 cyfr.
         NumberForm pozwala wydobyć więcej cyfr *)
      Print[n, " ", NumberForm[A, 16]];  (* Wydruk kontrolny. *)
      n = n + 1;         (* Kolejny wyraz ciągu *)
      m = m*2; (* Kolejny wielobok będzie miał dwa razy więcej boków *)
      a = Sqrt[2 - Sqrt[4 - 4*a*a]]/2, (* o takiej długości *)
      {d} (* Ilość powtórzeń pętli Do *)
    ];(*Koniec pętli Do*)
    a*m
  ]
```

Listing 4.1: Wyznaczenie ciągu przybliżeń liczby π w języku Mathematica w oparciu o obwody wielokątów

```
/****************** PiByCircumference.cpp *****************/
/* Obliczanie przybliżenia liczby pi poprzez obwody       */
/* wielokątów foremnych wpisanych w okrąg o średnicy jeden */
/**********************************************************/
#include <iostream>
#include <iomanip>
/* Biblioteka matematyczna, zawiera standardowe implementacje
   wielu matematycznych, funkcji. My skorzystamy poniżej
   z funkcji sqrt liczącej pierwiastek kwadratowy */
#include <cmath>
using namespace std;

int main(){
  cout << setprecision(15);
  cout << fixed << right;
  /* Zaczynamy od wpisanego kwadratu, którego
     bok a wynosi pół pierwiastka z 2. Do policzenia
     pierwiastka wykorzystujemy biblioteczny
     podprogram sqrt */
  double a=sqrt(2.0)/2;

  /* Ilośc boków wielokąta, na początek cztery */
  int m=4;
  /* Numer kolejnego wyrazu ciągu, ilość wyrazów do policzenia */
  int n=1,nmax=40;

  /* Dla wszystkich n od 1 do nmax */
  while(n<nmax){
    /* Obliczamy długość obwodu */
    double A=m*a;
    cout << "   A[" << setw(2) << n << "] =";
    cout << setw(18) << A << endl;
    /* Kolejny wielokąt będzie miał dwa razy więcej boków */
    m=m*2;
    /* A tyle wyniesie długość jego boku */
    a=sqrt(2-sqrt(4-4*a*a))/2;
    /* Kolejny wyraz ciągu */
    n=n+1;
  }
}
```

Listing 4.2: Wyznaczenie ciągu przybliżeń liczby π w C++ w oparciu o obwody wielokątów

Uruchamiając program w C++ z listingu 4.2 otrzymujemy następujący ciąg liczb (mogą występować drobne różnice dla różnych platform), z których pierwszych czternaście to:

$$\begin{aligned}
A_1 &= 2.828427124746190 \\
A_2 &= 3.061467458920719 \\
A_3 &= 3.121445152258053 \\
A_4 &= 3.136548490545940 \\
A_5 &= 3.140331156954754 \\
A_6 &= 3.141277250932759 \\
A_7 &= 3.141513801144285 \\
A_8 &= 3.141572940367515 \\
A_9 &= 3.141587725275636 \\
A_{10} &= 3.141591421514008 \\
A_{11} &= 3.141592345563526 \\
A_{12} &= 3.141592576625775 \\
A_{13} &= 3.141592634298169 \\
A_{14} &= 3.141592650272843
\end{aligned}$$

Do tego miejsca ciąg ten zachowuje się bardzo przyzwoicie, a jego kolejne wyrazy rosną. Jednak dalsze wyrazy pokazują, że nie do końca spełnia on nasze oczekiwania.

$$\begin{aligned}
A_{15} &= 3.141592648934190 \\
A_{16} &= 3.141592684156058 \\
A_{17} &= 3.141592607672169 \\
A_{18} &= 3.141592910939673 \\
A_{19} &= 3.141591696683685 \\
A_{20} &= 3.141596553704820 \\
A_{21} &= 3.141596553704820 \\
A_{22} &= 3.141518840465548 \\
A_{23} &= 3.141207968282266 \\
A_{24} &= 3.142451272494134 \\
A_{25} &= 3.142451272494134 \\
A_{26} &= 3.162277660168380 \\
A_{27} &= 3.162277660168380 \\
A_{28} &= 2.828427124746190 \\
A_{29} &= 0.000000000000000
\end{aligned}$$

Poczynając od wyrazu A_{15} ciąg przestaje być rosnący i zamiast przybliżać, zaczyna oddalać się od π, a wyraz A_{29} wynosi wręcz zero! I w tym przypadku przyczyną ograniczeń algorytmu jest skończona arytmetyka komputera. Efekt obliczeń z precyzją maszynową w środowisku Mathematica jest analogiczny.

Wartość liczby π z dokładnością do 15-tu miejsc dziesiętnych wynosi

$$\pi = 3.141592653589793\ldots.$$

Zatem przy zastosowaniu 64-bitowego typu **double** *przedstawiona metoda pozwala policzyć π z dokładnością do około 8 cyfr znaczących.*

Warto podkreślić, że sama obserwacja numeryczna zachowania tego ciągu może prowadzić do błędnego wniosku iż ciąg ten jest zbieżny do zera. Aby mieć praktyczny pożytek z ciągu jako narzędzia do przybliżania liczb, musimy umieć od strony teoretycznej przeanalizować jego zachowanie.

Mathematica radzi sobie łatwo z wyliczeniem π z dużą dokładnością, ale wymaga to obliczeń wysokiej precyzji, znacznie kosztowniejszych od obliczeń z precyzją maszynową. Na przykład pisząc

```
N[Pi, 1000]
```

otrzymujemy

$$3.14159265358979323846264338327950288419716939937510582097494459230781\backslash$$
$$64062862089986280348253421170679821480865132823066470938446095505822 31\backslash$$
$$72535940812848111745028410270193852110555964462294895493038196442881 09\backslash$$
$$75665933446128475648233786783165271201909145648566923460348610454326 64\backslash$$
$$82133936072602491412737245870066063155881748815209209628292540917153 64\backslash$$
$$36789259036001133053054882046652138414695194151160943305727036575959 19\backslash$$
$$53092186117381932611793105118548074462379962749567351885752724891227 93\backslash$$
$$81830119491298336733624406566430860213949463952247371907021798609437 02\backslash$$
$$77053921717629317675238467481846766940513200056812714526356082778577 13\backslash$$
$$42757789609173637178721468440901224953430146549585371050792279689258 92\backslash$$
$$35420199561121290219608640344181598136297747713099605187072113499999 98\backslash$$
$$37297804995105973173281609631859502445945534690830264252230825334468 50\backslash$$
$$35261931188171010003137838752886587533208381420617177669147303598253 49\backslash$$
$$04287554687311595628638823537875937519577818577805321712268066130019 27\backslash$$
$$87661119590921642019 9$$

jako przybliżenie liczby π z dokładnością do 1000 cyfr.

Ćwiczenie komputerowe 4.5.1 Zmodyfikuj program z listingu 4.2 tak by przybliżał π wykorzystując 2^n-kąty foremne opisane na okręgu (patrz cw. 1.2.7). Wykorzystaj ten program do prześledzenia zachowania tego ciągu przybliżeń.

4.5.2 Granice pozorne

W poprzednim podrozdziale zaobserwowaliśmy, że wyciąganie wniosków na temat granicy ciągu z samego numerycznego eksperymentu może być niebezpieczne. Okazuje się, że ciąg może robić wrażenie zbieżnego, a mimo wszystko zbieżnym nie być. Rozważmy ciąg dany wzorem rekurencyjnym

$$x_0 = 136.462554878752 \tag{4.11}$$
$$x_1 = 5.39650717129623 \tag{4.12}$$
$$x_n = 0.3x_0 + 1 - 1.4x_1^2 \tag{4.13}$$

Program w języku Mathematica wyznaczający wyrazy tego ciągu przedstawiony jest na listingu 4.3. Wersja w C++ jest na listingu 4.4.

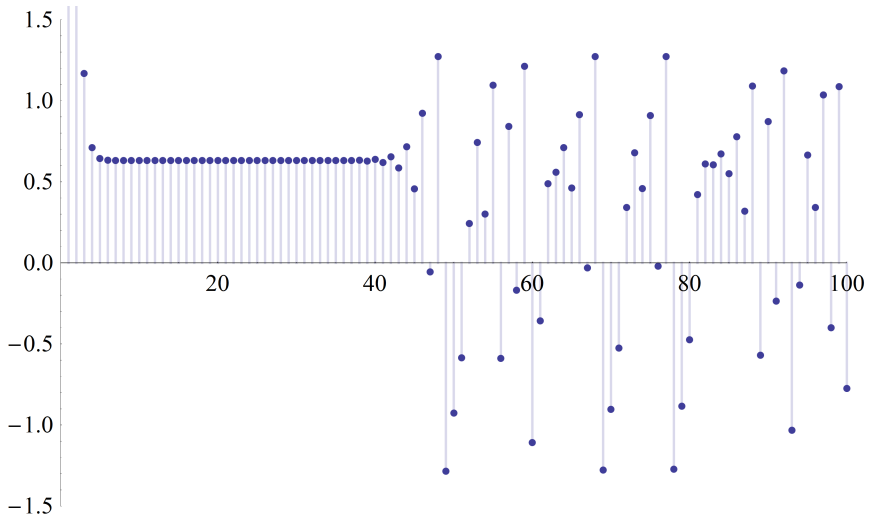
Przejrzenie pierwszych 15-tu wyrazów tego ciągu może robić wrażenie, że ciąg ten jest zbieżny do granicy wynoszącej około 0.63135. Przejrzenie 50-ciu wyrazów zaciera to pierwsze wrażenie. Jednakże, po doświadczeniach z ciagiem przybliżającym π możemy być skłonni podejrzewać, że przyczyną niestabilnego zachowania dalszych wyrazów tego ciągu znów jest skończona arytmetyka komputera. Okazuje się jednak, że obserwowana numerycznie zbieżność tego ciągu jest pozorna. Wykres pierwszych 100 wyrazów tego ciągu przedstawiono na rys. 4.6. Przykłady tego typu zaobserwowane zostały przez matematyków po raz pierwszy dopiero w XX wieku. Rozważany przez nas przykład wywodzi się z tzw. odwzorowania Hénona.

4.5.3 Definicja Cauchy'ego granicy ciągu

Niech $x : \mathbb{N} \to \mathbb{R}$ będzie ciągiem liczb rzeczywistych.

```
(* Przykład z ciągiem pozornie zbieżnym, zadanym rekurencyjnie *)
Henon[n_] := Module[{a, b, x0, x1, x2, k},
  a = 1.4; b = 0.3; (* Pomocnicze parametry a i b *)
  k = 0; (* Licznik wyrazów ciągu *)
  (* Dwa początkowe wyrazy ciągu *)
  x0 = 136.462554878752;
  x1 = 5.39650717129623;
  Do[k = k + 1;
    (* Kolejny wyraz ciągu wyliczamy w oparciu o dwa poprzednie *)
    x2 = b*x0 + 1 - a*x1*x1;
    (* Nowe wartości dwóch porzednich wyrazów ciągu *)
    x0 = x1;
    x1 = x2;
    Print[{k, NumberForm[x2, 16]}],
    {n}
  ];
]
```

Listing 4.3: Przykład ciągu pozornie zbieżnego w języku Mathematica.



Rysunek 4.6: Wykres pierwszych 100 wyrazów ciągu danego wzorami (4.11)-(4.13)

```
/************************** Henon.cpp ************************/
/* Przykład z zadanym rekurencyjnie ciągiem, pozornie zbieżnym. */
/*************************************************************/
#include <iostream>
#include <iomanip>
using namespace std;

int main(){
  cout << setprecision(15);
  /* Dla większej czytelności wyników narzucamy stałopozycyjny,
     wyrównywany w prawo zapis liczb */
  cout << fixed << right;

  /* Pomocnicze parametry a i b */
  double a=1.4,b=0.3;
  /* Dwa początkowe wyrazy ciągu */
  double x0=136.462554878752;
  double x1=5.39650717129623;
  /* Ilość wyrazów ciągu do policzenia */
  int nmax=15;
  /* Inna forma pętli. Instrukcje w bloku {...} wykonujemy dla wartości n
     od jeden do nmax zwiększając n za każdym razem o jeden */
  for(int n=1;n<=nmax;n=n+1){
    /* Kolejny wyraz ciągu wyliczamy w oparciu o dwa poprzednie  */
    double x2=b*x0+1-a*x1*x1;
    /* Wyprowadzamy wyniki na standardowe wyjście. setw(p) ustala ilość pól
       dla wyprowadzanej kolejnej zmiennej. Pozwala to na lepsze
       formatowanie wypisywanych danych */
    cout << "x[" << setw(2) << n << "]= ";
    cout << setw(20) << x0 << endl;
    /* Nowe wartości dwóch porzednich wyrazów ciągu */
    x0=x1;
    x1=x2;
  }
}
```

Listing 4.4: Przykład ciągu pozornie zbieżnego w C++.

Definicja 4.5.2 Mówimy, że ciąg x jest *zbieżny do granicy* $g \in \mathbb{R}$ jeżeli

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \forall n > N \quad |x_n - g| < \epsilon.$$

Piszemy wtedy

$$\lim_{n\to\infty} x_n = g.$$

Mówimy, że ciąg x jest *zbieżny*, jeżeli istnieje $g \in \mathbb{R}$ takie, że x jest zbieżny do g.

Twierdzenie 4.5.3 *Niech $p \in \mathbb{N}_1$. Wtedy*

$$\lim_{n\to\infty} \frac{1}{n^p} = 0.$$

Dowód:
Niech $\epsilon \in \mathbb{R}^+$. Weźmy $n \in \mathbb{N}$ takie, że $N > \frac{1}{\sqrt[p]{\epsilon}}$ Niech n będzie większe od N. Wtedy $n > \frac{1}{\sqrt[p]{\epsilon}}$ i z monotoniczności funkcji potęgowej $n^p > \frac{1}{\epsilon}$ i w konsekwencji $\frac{1}{n^p} < \epsilon$. Zatem dla $n > N$ mamy

$$|\frac{1}{n^p} - 0| = \frac{1}{n^p} < \epsilon.$$

□

Mówimy, że ciąg $a : \mathbb{N} \to \mathbb{R}$ jest *ograniczony*, jeśli istnieje $M \in \mathbb{R}$ takie, że $|a_n| \leqslant M$ dla wszystkich $n \in \mathbb{N}$.

Twierdzenie 4.5.4 *Ciąg zbieżny w $\mathbb{R}$ jest ograniczony.*

Dowód:
Niech $a_* := \lim_{n\to\infty} a_n$. Dla $\epsilon = 1$ znajdziemy $N \in \mathbb{N}$ takie, że dla $n > N$ mamy $|a_n - a_*| < 1$. Zatem dla takich n jest

$$|a_n| = |a_n - a_* + a_*| \leqslant |a_n - a_*| + |a_*| < 1 + |a_*|.$$

W zbiorze skończonym $\{\, |a_n| \mid n \in \mathbb{N} \text{ i } n \leqslant N \,\}$ jest element największy. Oznaczmy go M_0 i weźmy

$$M := \max(1 + |a_*|, M_0).$$

Wtedy dla wszystkich $n \in \mathbb{N}$ mamy $a_n \leqslant M$. □

4.5.4 Granica ciągu, a struktura ciała w $\mathbb{R}$

Niech $a, b : \mathbb{N} \to \mathbb{R}$ będą ciągami liczb rzeczywistych. Niech $\diamond \in \{ +, -, \cdot \}$ będzie operatorem dodawania, odejmowania lub mnożenia w $\mathbb{R}$. Definiujemy nowy ciąg

$$a \diamond b : \mathbb{N} \ni n \mapsto a_n \diamond b_n \in \mathbb{R}.$$

Twierdzenie 4.5.5 *(i) Jeśli ciągi a, b są zbieżne to ciąg $a \diamond b$ jest zbieżny oraz*

$$\lim_{n\to\infty} (a_n \diamond b_n) = \lim_{n\to\infty} a_n \diamond \lim_{n\to\infty} b_n.$$

(ii) Jeśli ciągi a, b są zbieżne i $\lim_{n\to\infty} b_n \neq 0$ to ciąg $a \,/\, b$ jest zbieżny oraz

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{\lim_{n\to\infty} a_n}{\lim_{n\to\infty} b_n}.$$

Dowód:
Niech $\lim_{n\to\infty} a_n = a_*$ i $\lim_{n\to\infty} b_n = b_*$ Przedstawimy dowód dla przypadku gdy $\diamond$ jest iloczynem. Pozostałe dowody są albo łatwiejsze (suma, różnica), albo wyglądają podobnie (iloraz).

Naszym celem jest oszacowanie różnicy

$$|a_n b_n - a_* b_*|$$

dla dużych n, przy wykorzystaniu oszacowań na różnice $|a_n - a_*|$ i $|b_n - b_*|$, które możemy uzyskać ze zbieżności ciągów a i b. Sztuczką powszechnie stosowaną przy takich oszacowaniach jest dodanie i odjęcie wyrazu pomocniczego i skorzystanie z własności wartości bezwzględnej.

$$|a_n b_n - a_* b_*| = |a_n b_n - a_* b_n + a_* b_n - a_* b_*| = |(a_n - a_*)b_n + a_*(b_n - b_*)| \leqslant |a_n - a_*||b_n| + |a_*||b_n - b_*|$$

Aby suma dwóch ostatnich składników była mniejsza od ϵ wystarczy, by każdy z nich był mniejszy od $\frac{\epsilon}{2}$. Rozwiązując nierówność

$$|a_*||b_n - b_*| < \frac{\epsilon}{2},$$

widzimy, że wystarczy, by $|b_n - b_*| < \frac{\epsilon}{2|a_*|}$, a to mamy zagwarantowane dla $n > N_1$ dobranego dla ciągu b do stałej $\frac{\epsilon}{2|a_*|}$. Wyjątkiem jest $|a_*| = 0$, bo wtedy nie można dzielić przez $|a_*|$, ale z tym nie ma problemu, bo przy $|a_*| = 0$ lewa strona nierówności jest zero, więc nierówność jest spełniona dla wszystkich n.

Trochę bardziej kłopotliwy jest przypadek nierówności

$$|a_n - a_*||b_n| < \frac{\epsilon}{2},$$

bo tu n występuje w dwóch miejscach. Jednak wiemy, że ciąg zbieżny jest ograniczony, więc możemy znaleźć stałą $M > 0$ taką, że $|b_n| < M$ dla wszystkich n. Zatem wystarczy, by

$$|a_n - a_*| < \frac{\epsilon}{2M},$$

a to mamy zagwarantowane dla $n > N_2$ dobranego dla ciągu a do stałej $\frac{\epsilon}{2M}$.

Aby uniknąć oddzielnego rozważania przypadku $a_* = 0$ wystarczy zauważyć, że $|a_*| < |a_*| + 1$, a $|a_*| + 1 \neq 0$, więc można od razu dobrać N_1 tak, by $|b_n - b_*| < \frac{\epsilon}{2(|a_*|+1)}$.

Weźmy ustalone $\epsilon > 0$. Niech M będzie takie, że $|b_n| < M$ dla wszystkich $n \in \mathbb{N}$. Niech $N_1 \in \mathbb{N}$ oraz $N_2 \in \mathbb{N}$ będą takie, że

$$\begin{aligned} n > N_1 \quad &\Rightarrow \quad |b_n - b_*| < \frac{\epsilon}{2(|a_*|+1)}, \\ n > N_2 \quad &\Rightarrow \quad |a_n - a_*| < \frac{\epsilon}{2M}. \end{aligned}$$

Niech $N := \max\{N_1, N_2\}$. Ustalmy $n > N$. Wtedy

$$|a_n b_n - a_* b_*| \leqslant |a_n - a_*||b_n| + |a_*||b_n - b_*| \leqslant |a_n - a_*|M + (|a_*|+1)|b_n - b_*| < \frac{\epsilon}{2M}M + (|a_*|+1)\frac{\epsilon}{2(|a_*|+1)} = \epsilon,$$

co dowodzi, że

$$\lim_{n\to\infty} a_n b_n = a_* b_*.$$

□

4.5.5 Granica ciągu, a struktura porządkowa w $\mathbb{R}$

Niech $a, b, c : \mathbb{N} \to \mathbb{R}$ będą ciągami liczb rzeczywistych.

Twierdzenie 4.5.6 *Jeśli istnieje $N \in \mathbb{N}$ takie że dla wszystkich $n > N$ ciągi a_n, b_n, c_n spełniają nierówność*

$$a_n \leqslant b_n \leqslant c_n$$

i ciągi a_n oraz c_n są zbieżne do granicy g, to ciąg b_n też jest zbieżny do granicy g.

Dowód:
Ustalmy $\epsilon > 0$. Ze zbieżności ciągów a i c do g wynika, że możemy dobrać stałą $N > 0$ tak, by dla $n > N$ zachodziły nierówności

$$|a_n - g| < \epsilon,$$
$$|c_n - g| < \epsilon.$$

Oznacza to, że dla $n \in \mathbb{N}$ zachodzi

$$-\epsilon < a_n - g < \epsilon,$$
$$-\epsilon < c_n - g < \epsilon.$$

Zatem mamy dla $n > N$

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

co oznacza, że dla takich n

$$|b_n - g| < \epsilon$$

i dowodzi, że

$$\lim_{n\to\infty} b_n = g.$$

□

Twierdzenie 4.5.7 *Niech $a : \mathbb{N} \to \mathbb{R}$ będzie ciągiem liczb rzeczywistych.*

(i) Jeśli a jest ciągiem rosnącym i ograniczonym od góry, to ciąg ten jest zbieżny oraz

$$\lim_{n\to\infty} a_n = \sup\{\, a_n \mid n \in \mathbb{N}\,\}.$$

(ii) Jeśli a jest ciągiem malejącym i ograniczonym od dołu, to ciąg ten jest zbieżny oraz

$$\lim_{n\to\infty} a_n = \inf\{\, a_n \mid n \in \mathbb{N}\,\}.$$

Dowód:
Udowodnimy własność (i). Drugą własność dowodzi się analogicznie. Oznaczmy

$$a_* := \sup\{\, a_n \mid n \in \mathbb{N}\,\}.$$

Ustalmy $\epsilon > 0$. Z definicji kresu górnego wynika, że $a_* - \epsilon$ nie jest majorantą zbioru $\{\, a_n \mid n \in \mathbb{N}\,\}$, zatem istnieje takie $N \in \mathbb{N}$, że $a_N > a_* - \epsilon$. Zatem dla $n > N$ mamy

$$a_* - \epsilon < a_N \leqslant a_n \leqslant a_* < a_* + \epsilon,$$

czyli

$$|a_n - a_*| < \epsilon,$$

co dowodzi, że ciąg a jest zbieżny i jego granicą jest a_*.

□

4.6 Przykłady i zastosowania granic ciągów w $\mathbb{R}$

4.6.1 Granice kilku ważnych ciągów

Twierdzenie 4.6.1

$$p \in \mathbb{R}^+ \quad \Rightarrow \quad \lim_{n\to\infty} \sqrt[n]{p} = 1 \tag{4.14}$$

$$\lim_{n\to\infty} \sqrt[n]{n} = 1 \tag{4.15}$$

$$\alpha \in \mathbb{Q}, a > 1 \quad \Rightarrow \quad \lim_{n\to\infty} \frac{n^\alpha}{a^n} = 0 \tag{4.16}$$

$$p \in \mathbb{R} \quad \Rightarrow \quad \lim_{n\to\infty} p^n = \begin{cases} 1 & p = 1 \\ 0 & p \in (-1,1) \\ \textit{nie istn. w } \mathbb{R} & p \leqslant -1 \textit{ lub } p > 1. \end{cases} \tag{4.17}$$

Dowód:

Dla pokazania własności (4.14) rozważmy najpierw przypadek $p \geqslant 1$. Niech $x_n := \sqrt[n]{p} - 1$. Wtedy $x_n \geqslant 0$ oraz na mocy nierówności Bernoulliego mamy

$$1 + nx_n \leqslant (1 + x_n)^n = p.$$

Zatem dla wszystkich $n \in \mathbb{N}$ mamy

$$0 \leqslant x_n \leqslant \frac{p-1}{n}.$$

Z twierdzenia o trzech ciągach wynika, że ciąg $\lim_{n\to\infty} x_n = 0$. Tak więc własność (4.14) dla $p \geqslant 1$ otrzymujemy z twierdzenia o granicy sumy ciągów. Gdy $p \in (0,1)$ wystarczy zauważyć, że dla $q := \frac{1}{p}$ mamy z twierdzenia o granicy ilorazu ciągów

$$\lim_{n\to\infty} \sqrt[n]{p} = \frac{1}{\lim_{n\to\infty} \sqrt[n]{q}} = \frac{1}{1} = 1.$$

Aby udowodnić (4.15) oznaczmy $x_n := \sqrt[n]{n} - 1$. Wystarczy pokazać, że ciąg o wyrazach x_n zmierza do zera. Z rozwinięcia dwumianu Newtona mamy

$$n = (1 + x_n)^n \geqslant \frac{1}{2} n(n-1) x_n^2,$$

skąd wynika, że dla $n \geqslant 2$ mamy

$$0 < x_n \leqslant \sqrt{\frac{2}{n-1}} \leqslant \frac{2}{\sqrt{n}},$$

a zatem $\lim_{n\to\infty} x_n = 0$.

Dla dowodu (4.16) oznaczmy $b := a - 1$ wybierzmy $k \in \mathbb{N}$, takie że $k > \alpha + 1$. Mamy dla $n > 2k$

$$a^n = (1+b)^n > \binom{n}{k} b^k = \frac{n(n-1)\cdots(n-k+1)}{1 \cdot 2 \cdots k} b^k > \frac{(\frac{n}{2})^k}{k!} b^k,$$

zatem

$$0 < \frac{n^\alpha}{(1+b)^n} < \frac{2^k k!}{b^k} n^{\alpha - k} < \frac{2^k k!}{b^k} \frac{1}{n}$$

i tezę znów dostajemy z twierdzenia o trzech ciągach.

Własność (4.17) jest prostym wnioskiem z (4.16) dla $p \in (0,1)$. Dla $p \in \{0,1\}$ własność jest oczywista. Dla $p \in (-1,0)$ wynika z nierówności

$$-|p|^n \leqslant p^n \leqslant |p|^n.$$

W pozostałym przypadku ciąg ma wartości dowolnie wielkie dla wykładników parzystych, jest więc nieograniczony, a więc nie może być zbieżny.

□

Aby policzyć granicę ciągu $a : \mathbb{N} \to \mathbb{R}$ w programie Mathematica należy użyć instrukcji

```
Limit[a[n], n -> Infinity, Assumptions -> Element[n, Integers]]
```

Na przykład pisząc

```
Limit[Sqrt[(1 + n)/(1 + 3 n)], n -> Infinity, Assumptions -> Element[n, Integers]]
```

otrzymujemy $\frac{1}{\sqrt{3}}$. Często ten sam wynik otrzymamy używając krótszej instrukcji

```
Limit[a[n], n -> Infinity]
```

która służy do liczenia granicy funkcji. Granica funkcji jest szerszym pojęciem, ale w przypadku argumentu zmierzającego do nieskończoności często daje ten sam wynik. Jednak nie jest to dokładnie to samo, bo na przykład w przypadku ciągu

$$\lim_{n\to\infty} \cos 2\pi n = 1,$$

gdyż jest to ciąg stale równy jeden, ale granica funkcji

$$\lim_{x\to\infty} \cos 2\pi x$$

nie istnieje. Mathematica dla instrukcji

```
Limit[Cos[2 Pi n], n -> Infinity, Assumptions -> Element[n, Integers]]
```

zwraca liczbę 1, natomiast instrukcja

```
Limit[Cos[2 Pi n], n -> Infinity]
```

zwraca

```
Interval[{-1, 1}]
```

co oznacza, że granica nie istnieje, a zbiór punktów granicznych, czyli granic wszystkich możliwych podciągów, to przedział $[-1, 1]$. Pojęciem zbioru punktów granicznych zajmiemy się w rozdziale 6.

4.6.2 Procent składany

Matematyk szwajcarski, Jakob Bernoulli (rys. 4.7) rozważał problem procentu składanego, który przedstawić można skrótowo następująco. Wpłacamy do banku 1zł oprocentowane na 100%. Po roku odbieramy $1 + 1 = 2$ zł. Jeśli bank jest gotowy wypłacać proporcjonalne oprocentowanie za dowolnie krótki okres, to po pół roku wypłacimy $1 + \frac{1}{2} = 1.5$ zł. Wpłacając tę kwotę na kolejne pół roku odbierzemy $(1 + \frac{1}{2})(1 + \frac{1}{2}) = 2.25$ zł. Zadając sobie trud wpłat i wypłat n-krotnie w ciągu roku na koniec roku wypłacimy kwotę

$$\left(1 + \frac{1}{n}\right)^n \text{ zł.}$$

Powstaje pytanie: czy możemy zyskać w ten sposób dowolnie wiele, jeśli n będzie odpowiednio duże? Programy w języku Mathematica oraz w C++ realizujące stosowny eksperyment numeryczny przedstawiono na listingach 4.5 i 4.6.

```
(* Obliczanie przybliżeń liczby e według idei procentu składanego
   (ang.compound interest) *)
CompoundInterest[d_] := Module[{n},
  n = 1;
  Do[
    Print["e[", n, "]= ", NumberForm[(1. + 1./n)^n, 16]];
    n = n + 1,
    {d}
  ];
]
```

Listing 4.5: Eksperyment numeryczny z procentem składanym w języku Mathematica

```
/******************** CompoundInterest.cpp ******************/
/* Obliczanie przybliżeń liczby e według idei procentu       */
/* składanego (ang. compound interest)                       */
/************************************************************/
#include <iostream>
#include <cmath>
#include <iomanip>
using namespace std;

int main(){

  cout << fixed << right;
  cout << setprecision(15);

  // Policzymy 1000 wyrazów ciągu
  int nmax=1000;
  for(int n=1;n<=nmax;n=n+1){
    // Czynnik, który trzeba przemnożyć n razy przez siebie
    // Piszemy 1.0/n, a nie 1/n, bo w arytmetyce
    // całkowitoliczbowej 1/n daje zero.
    double x=1+1.0/n;
    // W C++ brak wbudowanego potęgowania,
    // więc liczymy potęgę przy pomocy pętli
    double prod=1;
    for(int i=1;i<=n;i=i+1) prod=prod*x;
    cout << "   s[" << setw(3) << n << "] =" << setw(18) << prod;
    cout << endl;
  }
}
```

Listing 4.6: Eksperyment numeryczny z procentem składanym w C++

4.6.3 Liczba e

Twierdzenie 4.6.2 *Ciąg $\left(1+\frac{1}{n}\right)^n$ jest rosnący i ograniczony, a więc ma skończoną granicę.*

Dowód: Pokażemy najpierw, że ciąg ten jest rosnący. Podstawiając w nierówności Bernoullego (tw. 4.1.9) $-\frac{1}{n^2}$ za x i przekształcając uzyskaną nierówność otrzymujemy dla $n>1$

$$\begin{aligned}\left(1-\frac{1}{n^2}\right)^n &\geqslant 1-\frac{1}{n}\\ \left(1-\frac{1}{n}\right)^n\left(1+\frac{1}{n}\right)^n &\geqslant 1-\frac{1}{n}.\end{aligned}$$

Z kolei dzieląc obustronnie przez $1-\frac{1}{n}$, mnożąc przez $\left(\frac{n}{n-1}\right)^{n-1}$ i przekształcając dostajemy

$$\begin{aligned}\left(1-\frac{1}{n}\right)^{n-1}\left(1+\frac{1}{n}\right)^n &\geqslant 1\\ \left(1+\frac{1}{n}\right)^n &\geqslant \left(\frac{n}{n-1}\right)^{n-1}\\ \left(1+\frac{1}{n}\right)^n &\geqslant \left(1+\frac{1}{n-1}\right)^{n-1}\end{aligned}$$

do dowodzi, że rozważany ciąg jest rosnący.

Aby pokazać, że jest on ograniczony, najpierw zauważmy, że podstawiając 1 za b oraz $\frac{1}{2}$ za a we wzorze (3.8) otrzymujemy

$$1-\frac{1}{2^n}=\frac{1}{2}\left(\frac{1}{2^{n-1}}+\frac{1}{2^{n-2}}+\cdots+1\right). \tag{4.18}$$

Zatem podstawiając we wzorze (3.9) 1 za a oraz $\frac{1}{n}$ za b, przekształcając i na koniec wykorzystując (4.18) otrzymujemy

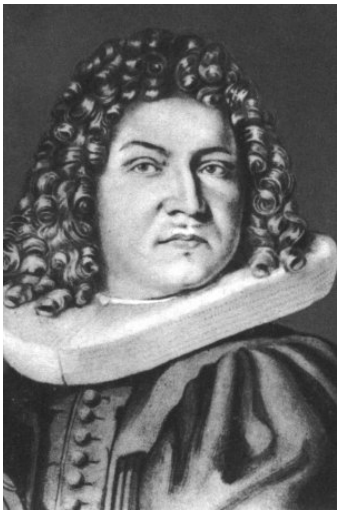
$$\begin{aligned}\left(1+\frac{1}{n}\right)^n &= 1+\binom{n}{1}\frac{1}{n}+\binom{n}{2}\frac{1}{n^2}+\binom{n}{3}\frac{1}{n^3}+\cdots+\binom{n}{n}\frac{1}{n^n}\\ &= 1+\frac{n}{1}\frac{1}{n}+\frac{n(n-1)}{1\cdot 2}\frac{1}{n^2}+\frac{n(n-1)(n-2)}{1\cdot 2\cdot 3}\frac{1}{n^3}+\cdots+\frac{n(n-1)\cdots 1}{1\cdot 2\cdots n}\frac{1}{n^n}\\ &= 1+1+\frac{1(1-\frac{1}{n})}{1\cdot 2}+\frac{1(1-\frac{1}{n})(1-\frac{2}{n})}{1\cdot 2\cdot 3}+\cdots+\frac{1(1-\frac{1}{n})\cdots(1-\frac{n-1}{n})}{1\cdot 2\cdots n}\\ &< 1+1+\frac{1}{1\cdot 2}+\frac{1}{1\cdot 2\cdot 3}+\cdots+\frac{1}{1\cdot 2\cdots n}\\ &< 1+1+\frac{1}{2}+\frac{1}{2^2}+\cdots+\frac{1}{2^{n-1}}\\ &= 1+2\left(1-\frac{1}{2^n}\right)<1+2=3.\end{aligned}$$

Zatem na mocy twierdzenia 4.5.7 ciąg $\left(1+\frac{1}{n}\right)^n$ jest zbieżny. □

Choć Bernoulli pierwszy zauważył, że ciąg o wyrazach $\left(1+\frac{1}{n}\right)^n$ jest zbieżny, jego granicę nazywamy liczbą Eulera *i oznaczamy literą e, bo to Leonard Euler (rys. 4.7), inny szwajcarski matematyk, pokazał jak ważna jest ona w matematyce. On pierwszy oznaczał ją literą e i tak pozostało do dziś.*

Liczba e w przybliżeniu wynosi

$$e=2.71828182845904523536\ldots$$





Rysunek 4.7: Jacob Bernoulli (1654 - 1705) i Leonard Euler (1707 - 1783)

Co ciekawe, ciąg $\left(1+\frac{1}{n}\right)^n$ nadaje liczbie e ważną interpretację, jednak zupełnie nie nadaje się do obliczania jej przybliżeń, bo jest bardzo wolno zbieżny. W rozdziale 8 poznamy nieporównanie szybszą metodę wyznaczania liczby e.

W programie Mathematica liczba e jest oznaczana E.

4.6.4 Funkcja wykładnicza

Wpłacając kwotę 1zł na rachunek z oprocentowaniem nie 100%, tylko x (potraktowane jako liczba, 100%=1) po roku wypłacamy kwotę

$$\exp(x) := \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n \text{ zł.}$$

Jest to przykład funkcji ciągłej, której wzrost (lub spadek) jest proporcjonalny do jej wartości.

Można to zapisać w postaci równania różniczkowego

$$f'(x) = af(x).$$

Rozwiązaniem tego równania jest funkcja wykładnicza i jej wielokrotności. Bardzo wiele zjawisk w fizyce (np. rozpad pierwiastków promieniotwórczych, biologii (np. wzrost populacji przy nieograniczonych zasobach), ekonomii (np. piramidy finansowe) i innych dziedzinach opisywanych jest przy pomocy funkcji wykładniczej.

Można pokazać, że funkcja wykładnicza spełnia wzór

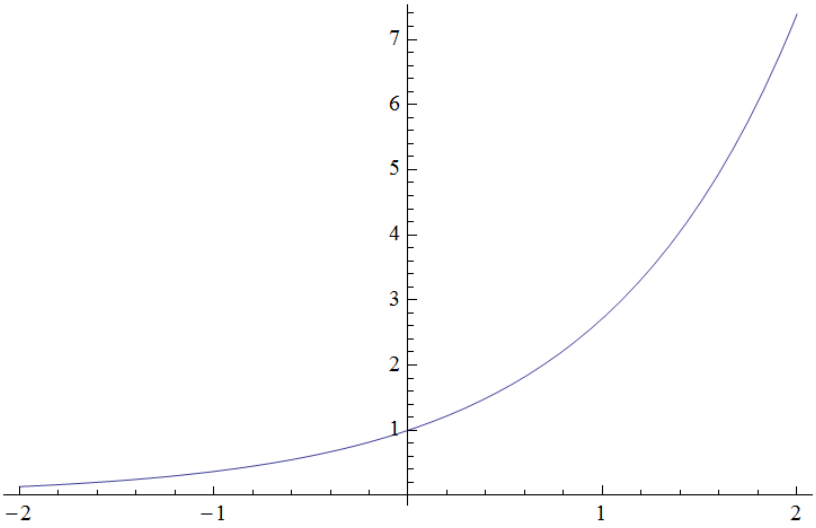
$$f(x+y) = f(x)f(y),$$

a stąd, że jest postaci

$$\mathbb{R} \ni x \mapsto a^x \in \mathbb{R}^+.$$

W wyrażeniu tym $a \in \mathbb{R}^+$, a a^x jest rozszerzeniem wyrażenia a^n dla $n \in \mathbb{N}$ na dowolny wykładnik rzeczywisty. W szczególności łatwo stąd wywnioskować, że dla $\frac{p}{q} \in \mathbb{Q}^+$

$$a^{\frac{p}{q}} = \sqrt[q]{a^p},$$



Rysunek 4.8: Wykres funkcji exp.

$$a^0 = 1,$$

oraz dla $r \in \mathbb{Q}^-$

$$a^r = \frac{1}{a^{-r}}.$$

Przyjmując powyższe wzory jako definicje, można zdefiniować funkcję wykładniczą dla dowolnego $x \in \mathbb{R}$ *wzorem*

$$a^x := \sup \{\, a^r \mid r \in \mathbb{Q}, r < x \,\}$$

i pokazać, że

$$\exp(x) = e^x.$$

W programie Mathematica funkcja $\exp$ *jest oznaczana* `Exp`. *Natomiast wyrażenie* a^x *występujące w funkcji wykładniczej o podstawie* a *w programie Mathematica zapisywane jest w postaci* `a^x`.

Wykres funkcji wykładniczej przedstawiono na rys. 4.8

Od strony obliczeniowej wygodniejszy jest jednak wzór

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

który nam posłuży jako podstawa do formalnej definicji funkcji wykładniczej.

Rozdział 5

Topologia przestrzeni metrycznych

Jak widzieliśmy w poprzednim rozdziale jest potrzeba rozszerzenia pojęcia granicy tak, by przynajmniej obejmowało ono granice ciągów liczb zespolonych. W rzeczywistości, jak już mówiliśmy, koncepcja granicy jest kluczowa dla całej analizy. Dlatego spróbujemy zastanowić się co jest rzeczywiście niezbędne do zdefiniowania granicy.

Intencją zwrotu $|x_n - g| < \epsilon$ w definicji Cauchy'ego granicy ciągu jest stwierdzenie, że x_n leży nie dalej niż ϵ od g. Widzieliśmy, że również w przypadku liczb zespolonych możemy mierzyć odległość między $u, w \in \mathbb{C}$ rozważając funkcję, która parze liczb zespolonych przyporządkowuje moduł ich różnicy.

$$\mathbb{C} \ni (u, w) \mapsto |u - w| \in \mathbb{R}^*.$$

Ma ona własności analogiczne do funkcji, która parze liczb rzeczywistych przyporządkowuje wartość bezwzględną ich różnicy. W tym rozdziale przedstawimy ideę odległości, która uogólnia obie te funkcje. Jak zobaczymy później jest to jednak tylko krok pośredni do uzyskania ogólnego pojęcia granicy.

5.1 Przestrzenie metryczne

5.1.1 Przestrzeń metryczna

Definicja 5.1.1 *Odległością (metryką)* w zbiorze X nazywamy odwzorowanie

$$\rho : X \times X \to \mathbb{R}^*$$

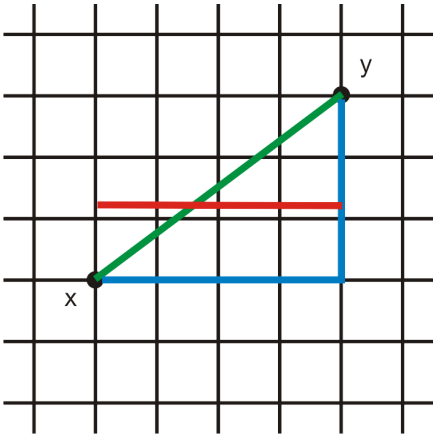
spełniające własności

(i) $\rho(x, y) = 0 \Leftrightarrow x = y$,

(ii) $\rho(x, y) = \rho(y, x)$,

(iii) $\rho(x, z) \leqslant \rho(x, y) + \rho(y, z)$.

Parę (X, ρ) nazywamy *przestrzenią metryczną*.



Rysunek 5.1: Odległość punktów $x, y \in \mathbb{R}^2$ w metryce euklidesowej (zielony), Manhattan (niebieski) i maksimum (czerwony).

5.1.2 Przykłady metryk w $\mathbb{R}^n$

Przestrzeń $\mathbb{R}^n$ to iloczyn kartezjański n kopii zbioru liczb rzeczywistych $\mathbb{R}$. Element $x \in \mathbb{R}^n$ ma postać n-tki liczb: $x = (x_1, x_2, \ldots, x_n)$, którą krótko zapisujemy $x = (x_i)_{i=1,n}$. Dla $x = (x_i)_{i=1,n}, y = (y_i)_{i=1,n} \in \mathbb{R}^n$ definiujemy

$$\begin{aligned} d_e(x,y) &:= \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}, \\ d_{\max}(x,y) &:= \max\{\, |x_i - y_i| \mid i = 1, 2, \ldots n \,\}, \\ d_1(x,y) &:= \sum_{i=1}^{n} |x_i - y_i|. \end{aligned}$$

Jak łatwo zauważyć, pierwsza z tych definicji uogólnia na przypadek $\mathbb{R}^n$ metrykę zadaną w $\mathbb{R}$ poprzez wartość bezwzględną oraz metrykę w $\mathbb{C}$ zadaną poprzez moduł.

Twierdzenie 5.1.2 *(i) (R^n, d_e), $(R^n, d_{\max})$ i (R^n, d_1) są przestrzeniami metrycznymi.*

(ii) d_e w $\mathbb{R}^2 = \mathbb{C}$ pokrywa się z modułem, a w $\mathbb{R}$ z wartością bezwzględną różnicy argumentów.

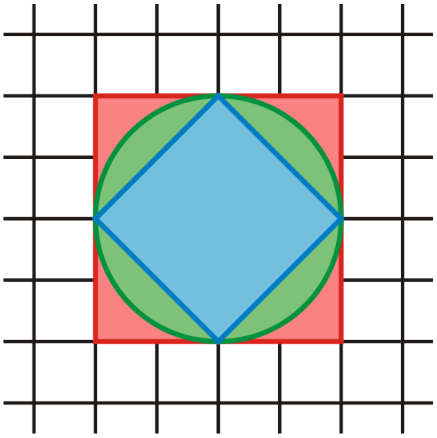
Metryka d_e to *metryka euklidesowa*, metrykę $d_{\max}$ określa się mianem *metryka maksimum*, a metrykę d_1 mianem *metryka Manhattan*. Łatwo sprawdzić, że dla $n = 1$ wszystkie te trzy metryki są sobie równe, dane wzorem $\rho(x, y) = |x - y|$ dla $x, y \in \mathbb{R}$.

Przykład liczenia odległości w $\mathbb{R}^2$ w różnych metrykach przedstawiono na rys. 5.1.

5.1.3 Metryka indukowana i podprzestrzenie przestrzeni metrycznej

Twierdzenie 5.1.3 *Niech (X, ρ) będzie przestrzenią metryczną, a $Y \subset X$ podzbiorem. Wtedy $\rho_Y := \rho_{|Y \times Y}$ jest metryką w Y.*

Dowód: ćwiczenie.



Rysunek 5.2: Kule w $\mathbb{R}^2$ w metryce euklidesowej (zielony), Manhattan (niebieski) i maksimum (czerwony). Odpowiednie sfery zaznaczono ciemniejszym kolorem.

Definicja 5.1.4 Metrykę ρ_Y nazywamy *metryką indukowaną* z X w Y, a przestrzeń metryczną (Y, ρ_Y) nazywamy *podprzestrzenią* przestrzeni (X, ρ).

Zatem każdy podzbiór przestrzeni metrycznej jest przestrzenią metryczną. Pokazuje to, że przestrzenie metryczne mogą być bardzo różnorodne.

5.1.4 Kule

Definicja 5.1.5 Niech (X, ρ) będzie przestrzenią metryczną, x punktem w X, a r liczbą rzeczywistą dodatnią. Przez *kulę otwartą w X o środku w x i promieniu r* rozumiemy zbiór

$$K(x, r) := \{\, y \in X \mid \rho(y, x) < r \,\}.$$

Podobnie definiujemy *kulę domkniętą* jako zbiór

$$\bar{K}(x, r) := \{\, y \in X \mid \rho(y, x) \leqslant r \,\},$$

a *sferę* jako zbiór

$$S(x, r) := \{\, y \in X \mid \rho(y, x) = r \,\}.$$

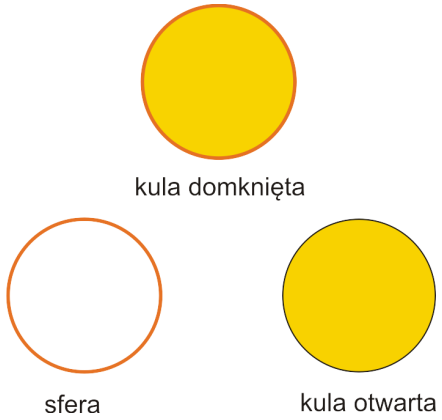
Dla ustalonej przestrzeni metrycznej (X, ρ) oraz $x \in X$ będziemy używać oznaczenia

$$\mathrm{Ball}(x) := \mathrm{Ball}(x, X) := \mathrm{Ball}(x, X, \rho) := \{\, K(x, r) \mid r \in \mathbb{R}^+ \,\}.$$

na rodzinę kul otwartych o środku w x.
Przykład kul domkniętych i sfer w $\mathbb{R}^2$ w różnych metrykach przedstawiono na rys. 5.2.

Różnica między kulą domkniętą, a otwartą jest analogiczna do różnicy między przedziałem domkniętym, a przedziałem otwartym. Na rysunku można ją przedstawić tylko umownie, na przykład zaznaczając sferę ograniczającą kulę otwartą linią cienką, a sferę ograniczającą kulę domkniętą linią grubą (rys.5.3).

Posługując się metryką, można pojęcie granicy ciągu w przestrzeni metrycznej sformułować tak



Rysunek 5.3: Umowne przedstawienie różnicy między kulą domknięta, sferą i kulą otwartą.

$$\lim_{n\to\infty} x_n = g \quad :\Leftrightarrow \quad \forall \epsilon > 0\, \exists N \in \mathbb{N} \forall n \in \mathbb{N} \quad n \geqslant N \;\Rightarrow\; \rho(x_n, g) < \epsilon,$$

a posługując się pojęciem kuli tak

$$\lim_{n\to\infty} x_n = g \quad :\Leftrightarrow \quad \forall K \in \mathrm{Ball}(g)\, \exists N \in \mathbb{N} \forall n \in \mathbb{N} \quad n \geqslant N \;\Rightarrow\; x_n \in K.$$

Docelowo chcielibyśmy je przeformułować tak

$$\lim_{n\to\infty} x_n = g \quad :\Leftrightarrow \quad \forall V \in \mathrm{Nb}(g)\, \exists U \in \mathrm{Nb}(\infty) \forall n \in \mathbb{N} \quad n \in U \;\Rightarrow\; x_n \in V,$$

gdzie $\mathrm{Nb}(g)$ to pewna rodzina zbiorów "zadająca coraz większe wymagania bliskości" do g, a $\mathrm{Nb}(\infty)$ to pewna rodzina zbiorów "zadająca coraz większe wymagania bliskości" do ∞. W naszym przypadku ta pierwsza rodzina to rodzina kul o środku w g, a ta druga to rodzina przedziałów $[N, \infty)$. Tego typu rodziny to tzw. bazy otoczeń. Sformalizujemy to pojęcie pod koniec tego rozdziału.

5.1.5 Zbiory ograniczone

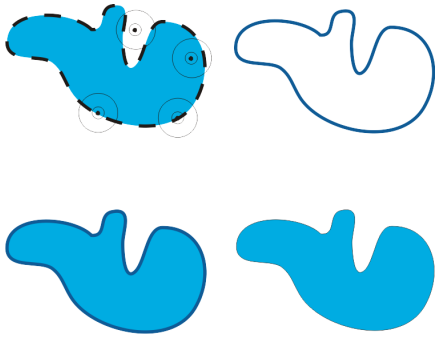
W dalszym ciągu przyjmujemy, że X jest przestrzenią metryczną z metryką ρ.

Definicja 5.1.6 Podzbiór $A \subset X$ nazywamy *ograniczonym* jeśli istnieje kula $K(x, r)$ o środku w x i promieniu r taka, że $A \subset K(x, r)$.

Twierdzenie 5.1.7 *Suma skończonej ilości zbiorów ograniczonych jest zbiorem ograniczonym.*

5.2 Średnica zbiorów

5.2.1 Średnica zbioru



Rysunek 5.4: Poszukiwanie punktów brzegowych (na górze po lewej) zbioru A, brzeg zbioru A (na górze po prawej), domknięcie zbioru A jako przykład zbioru domkniętego (na dole po lewej) i otwarcie zbioru A jako przykład zbioru otwartego (na dole po prawej).

Definicja 5.2.1 *Średnicę* zbioru $A \subset X$ definiujemy jako

$$\rho(A) := \sup \{ \rho(x, y) \mid x, y \in A \}$$

gdy A jest zbiorem niepustym, a jako zero w przeciwnym razie.

Zauważmy, że dla metryki euklidesowej d_e w $\mathbb{R}$ oraz $x, y \in \mathbb{R}$ takich, że $x \leqslant y$ średnica przedziału otwartego (x, y) oraz domkniętego $[x, y]$ wynosi $y - x$.

Uwaga 5.2.2 *Średnica zbioru ma następujące własności.*

(i) Zbiór A jest ograniczony wtedy i tylko wtedy, gdy jego średnica jest skończona.

(ii) Średnica zbioru A wynosi zero wtedy i tylko wtedy, gdy zbiór A ma nie więcej niż jeden element.

(iii) Jeśli $A \subset B$, to $\rho(A) \leqslant \rho(B)$.

(iv) Jeśli $A : \mathbb{N} \to \mathcal{P}(X)$ jest ciągiem zbiorów takich, że $\rho(A_n) < \frac{1}{n}$, to $\rho(\bigcap_{n=1}^{\infty} A_n) = 0$.

5.3 Zbiory otwarte i domknięte, wnętrze, domknięcie i brzeg zbioru

Pojęcia zbioru otwartego i domkniętego uogólniają pojęcia kuli otwartej i domkniętej. Intuicyjnie zbiór otwarty to taki, który jest rozłączny ze swoim brzegiem, a zbiór domknięty to taki, który zawiera swój brzeg. Punkt x jest punktem brzegowym zbioru A jeżeli każda kula $K \in \mathrm{Ball}(x)$ zawiera tak punkty należące do A jak i nie należące do A. Ideę tę przedstawiono na rys. 5.4.

Definicja 5.3.1 Niech $A \subset X$. *Brzegiem* zbioru A w X nazywamy zbiór

$$\mathrm{bd}\, A := \{ x \in X \mid \forall K \in \mathrm{Ball}(x) \quad K \cap A \neq \varnothing \text{ i } K \setminus A \neq \varnothing \}.$$

Mówimy, że A jest *otwarty*, jeżeli $A \cap \mathrm{bd}\, A = \varnothing$. Mówimy, że A jest *domknięty*, jeżeli $\mathrm{bd}\, A \subset A$.

Twierdzenie 5.3.2 *A jest otwarty wtedy i tylko wtedy gdy*

$$\forall x \in A\, \exists K \in \mathrm{Ball}(x) \quad K \subset A.$$

A jest domknięty wtedy i tylko wtedy gdy

$$\forall x \notin A\, \exists K \in \mathrm{Ball}(x) \quad K \cap A = \varnothing.$$

5.3.1 Wnętrze i domknięcie

Definicja 5.3.3 Niech $A \subset X$. *Wnętrzem* zbioru A nazywamy zbiór

$$\operatorname{int} A := A \setminus \operatorname{bd} A$$

Domknięciem zbioru A nazywamy zbiór

$$\operatorname{cl} A := A \cup \operatorname{bd} A$$

Twierdzenie 5.3.4 *Niech $A \subset X$. Wtedy*

$$\operatorname{int} A = \{\, x \in X \mid \exists K \in \mathrm{Ball}(x)\ K \subset A. \,\}$$

oraz

$$\operatorname{cl} A = \{\, x \in X \mid \forall K \in \mathrm{Ball}(x)\ K \cap A \neq \varnothing. \,\}$$

Niech $X \in \mathrm{Metr}$, $x \in X$, $r > 0$.

Twierdzenie 5.3.5 *Kula otwarta $K(x,r)$ jest zbiorem otwartym, a kula domknięta $\bar{K}(x,r)$ jest zbiorem domkniętym.*

Rodzinę zbiorów otwartych w przestrzeni metrycznej (X,ρ) oznaczać będziemy $\mathrm{Open}(X,\rho)$, a zbiorów domkniętych $\mathrm{Clos}(X,\rho)$. W notacji tej będziemy pomijać metrykę i pisać $\mathrm{Open}(X)$, $\mathrm{Clos}(X)$ gdy metryka będzie znana z kontekstu.

5.3.2 Własności wnętrza i domknięcia

Uwaga 5.3.6 *Niech $A \subset X$. Wtedy*

(i) $\operatorname{int} A \subset A \subset \operatorname{cl} A$,

(ii) *A jest otwarty wtedy i tylko wtedy gdy* $A = \operatorname{int} A$,

(iii) *A jest domknięty wtedy i tylko wtedy gdy* $A = \operatorname{cl} A$,

(iv) $\setminus \operatorname{int} A = \operatorname{cl}(\setminus A)$, $\quad \setminus \operatorname{cl} A = \operatorname{int}(\setminus A)$,

(v) $\operatorname{bd} A = \operatorname{cl} A \cap \operatorname{cl}(\setminus A)$ *jest zbiorem domkniętym.*

Twierdzenie 5.3.7 *Operacje wnętrza i domknięcia są monotoniczne względem relacji inkluzji, tzn. jeśli* $A \subset B \subset X$, *to* $\operatorname{int} A \subset \operatorname{int} B$ *oraz* $\operatorname{cl} A \subset \operatorname{cl} B$.

Twierdzenie 5.3.8 *Niech* $A \subset X$. *Wnętrze* A *jest zbiorem otwartym. Jest to największy zbiór otwarty zawarty w* A. *Podobnie domknięcie* A *jest zbiorem domkniętym i jest to najmniejszy zbiór domknięty zawierający* A.

Wniosek 5.3.9 *Dla każdego* $A \subset X$ *mamy*

$$\operatorname{int}(\operatorname{int} A) = \operatorname{int} A, \quad \operatorname{cl}(\operatorname{cl} A) = \operatorname{cl} A.$$

Twierdzenie 5.3.10 *Dla dowolnych* $A, B \subset X$ *mamy*

$$\begin{aligned} \operatorname{int}(A \cap B) &= \operatorname{int} A \cap \operatorname{int} B, \\ \operatorname{int}(A \cup B) &\supset \operatorname{int} A \cup \operatorname{int} B, \\ \operatorname{cl}(A \cap B) &\subset \operatorname{cl} A \cap \operatorname{cl} B, \\ \operatorname{cl}(A \cup B) &= \operatorname{cl} A \cup \operatorname{cl} B. \end{aligned}$$

5.4 Otoczenia, punkty skupienia

5.4.1 Rodzina otoczeń punktu

Definicja 5.4.1 Mówimy, że zbiór A jest *otoczeniem* punktu $x \in X$ jeżeli $x \in \operatorname{int} A$. Mówimy, że A jest *otoczeniem otwartym* punktu x jeżeli A jest otwarty i $x \in A$. Zbiór otoczeń punktu x oznaczamy $\operatorname{Nb}(x)$, a zbiór otoczeń otwartych punktu x oznaczamy $\operatorname{ONb}(x)$.

5.4.2 Punkty skupienia zbioru

Definicja 5.4.2 Punkt $x \in X$ nazywamy *punktem skupienia* zbioru $A \subset X$, jeżeli

$$\forall V \in \operatorname{Nb}(x)\, \exists y \in X \quad y \neq x \text{ i } y \in V \cap A.$$

Zbiór punktów skupienia zbioru A oznaczamy A'.

Definicja 5.4.3 Mówimy, że $x \in X$ jest *punktem izolowanym* zbioru $A \subset X$, jeżeli $x \in A$ i x nie jest punktem skupienia zbioru A.

Dla zbioru

$$A := \{ \tfrac{1}{n} \mid n \in \mathbb{N}_1 \} \cup \{ \tfrac{n}{n+1} \mid n \in \mathbb{N}_1 \}$$

jedynymi punktami skupienia są 0 i 1, z czego tylko 1 jest elementem A. Wszystkie pozostałe punkty zbioru A to jego punkty izolowane.

5.5 Metryki równoważne, iloczyn kartezjański przestrzeni metrycznych

5.5.1 Równoważność metryk.

Definicja 5.5.1 Niech ρ_1 i ρ_2 będą metrykami w zbiorze X. Mówimy, że ρ_1 i ρ_2 są *równoważne* wtedy i tylko wtedy, gdy zachodzą następujące dwa warunki

$$\mathrm{Open}(X,\rho_1) = \mathrm{Open}(X,\rho_2).$$

Twierdzenie 5.5.2 *Metryki ρ_1, ρ_2 na X są rówoważne wtedy i tylko wtedy gdy*

$$\begin{aligned} \forall x \in X\, \forall K_1 \in \mathrm{Ball}(x,X,\rho_1)\, \exists K_2 \in \mathrm{Ball}(x,X,\rho_2) \quad K_2 &\subset K_1, \\ \forall x \in X\, \forall K_2 \in \mathrm{Ball}(x,X,\rho_2)\, \exists K_1 \in \mathrm{Ball}(x,X,\rho_1) \quad K_1 &\subset K_2. \end{aligned}$$

Twierdzenie 5.5.3 *Metryki euklidesowa, maksimum i Manhattan w $\mathbb{R}^d$ są wzajemnie równoważne.*

5.5.2 Iloczyn kartezjański przestrzeni metrycznych

Twierdzenie 5.5.4 *Niech (X_i,ρ_i) dla $i = 1,2,\dots d$ będą przestrzeniami metrycznymi. Dla $x := (x_1,x_2,\dots,x_d) \in X_1 \times X_2 \times \cdots \times X_d$ i $y := (y_1,y_2,\dots,y_d) \in X_1 \times X_2 \times \cdots \times X_d$ połóżmy*

$$\begin{aligned} \rho_e(x,y) &:= \sqrt{\sum_{i=1}^{n} \rho_i(x_i,y_i)^2}, \\ \rho_{\max}(x,y) &:= \max\{\, \rho_i(x_i,y_i) \mid i = 1,2,\dots n \,\}, \\ \rho_1(x,y) &:= \sum_{i=1}^{n} \rho_i(x_i,y_i). \end{aligned}$$

Wtedy ρ_e, $\rho_{\max}$ i $\rho_1(x,y)$ są wzajemnie równoważnymi metrykami w $X_1 \times X_2 \times \cdots \times X_d$.

5.6 (*)Topologia przestrzeni metrycznej

5.6.1 Fundamentalne własności zbiorów otwartych i domkniętych

Twierdzenie 5.6.1 *Niech (X,ρ) będzie przestrzenią metryczną. Wtedy*

$$\begin{aligned} &\varnothing, X \in \mathrm{Open}(X), \\ &U_1, U_2 \in \mathrm{Open}(X) \;\Rightarrow\; U_1 \cap U_2 \in \mathrm{Open}(X), \\ &\{U_\iota\}_{\iota\in I} \subset \mathrm{Open}(X) \;\Rightarrow\; \bigcup_{\iota\in I} U_\iota \in \mathrm{Open}(X). \end{aligned}$$

Twierdzenie 5.6.2 *Niech (X,ρ) będzie przestrzenią metryczną. Wtedy*

$$\begin{aligned}&\varnothing, X \in \mathrm{Clos}(X),\\ &U_1, U_2 \in \mathrm{Clos}(X) \;\Rightarrow\; U_1 \cup U_2 \in \mathrm{Clos}(X),\\ &\{U_\iota\}_{\iota\in I} \subset \mathrm{Clos}(X) \;\Rightarrow\; \bigcap_{\iota\in I} U_\iota \in \mathrm{Clos}(X).\end{aligned}$$

Definicja 5.6.3 Rodzinę zbiorów otwartych przestrzeni metrycznej nazywamy *topologią* tej przestrzeni.

Widzimy, że różne metryki mogą prowadzić do tej samej rodziny zbiorów otwartych, czyli topologii. I właśnie topologia jest tym co nas w tym wykładzie interesuje bardziej niż metryka. Topologię można studiować bez metryki. Wtedy przyjmuje się jako aksjomaty własności rodziny zbiorów otwartych zawarte w tezie twierdzenia 5.6.1. Przestrzeń z zadaną rodziną takich zbiorów nazywa się przestrzenią topologiczną. Okazuje się, że istnieją przestrzenie topologiczne, których topologii nie da się zadać przy pomocy metryki. Dziedzinę matematyki, która zajmuje się badaniem przestrzeni topologicznych też określa się mianem topologia.

5.7 (**)Bazy otoczeń

5.7.1 Baza otoczeń punktu

Niech (X,ρ) będzie ustaloną przestrzenią metryczną.

Definicja 5.7.1 Mówimy, że rodzina $\mathcal{B} \subset X$ jest *bazą otoczeń punktu $x \in X$* jeżeli zachodzą następujące dwa warunki

(i) $\forall B \in \mathcal{B} \quad x \in B$,

(ii) $\forall U \in \mathrm{Open}(X) \quad x \in U \;\Rightarrow\; \exists B \in \mathcal{B}\; B \subset U$.

Rodzinę rodzin zbiorów $\{\mathcal{B}_x\}_{x\in X}$ nazywamy *bazą otoczeń* w X jeżeli dla każdego $x \in X$ rodzina $\mathcal{B}_x$ jest bazą otoczeń punktu x.

Dla $x \in X$ zdefiniujmy

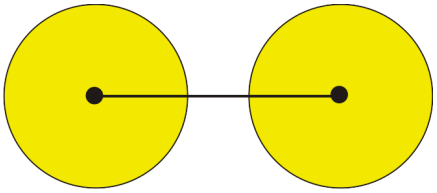
$$\begin{aligned}\mathrm{ONb}(x) &:= \{\, V \in \mathrm{Open}(X) \mid x \in V \,\},\\ \mathrm{Nb}(x) &:= \{\, A \subset X \mid \exists V \in \mathrm{Open}(X)\; x \in V \subset A \,\},\\ \mathrm{Ball}_0(x) &:= \{\, K(x, \tfrac{1}{n}) \mid n \in \mathbb{N} \,\}.\end{aligned}$$

Dla każdego $x \in X$ mamy

$$\mathrm{Ball}_0(x) \subset \mathrm{Ball}(x) \subset \mathrm{ONb}(x) \subset \mathrm{Nb}(x).$$

Twierdzenie 5.7.2 *Rodziny rodzin zbiorów $\{\mathrm{Ball}_0(x)\}_{x\in X}$, $\{\mathrm{Ball}(x)\}_{x\in X}$, $\{\mathrm{ONb}(x)\}_{x\in X}$, $\{\mathrm{Nb}(x)\}_{x\in X}$ są bazami otoczeń w X.*

Bazy otoczeń są przydatne, bo wiele warunków wystarczy sprawdzać jedynie na bazach. W szczególności mamy następujące twierdzenie.



Rysunek 5.5: Rozdzielenie punktów kulami w przestrzeni metrycznej.

Twierdzenie 5.7.3 *Niech $\{\mathcal{B}_x\}_{x\in X}$ będzie bazą otoczeń w X. Wtedy*

(i) Zbiór A jest otwarty w X wtedy i tylko wtedy gdy dla każdego $x \in A$ istnieje $B \in \mathcal{B}_x$ takie, że $B \subset A$.

(ii) Zbiór A jest domknięty w X wtedy i tylko wtedy gdy dla każdego $x \notin A$ istnieje $B \in \mathcal{B}_x$ takie, że $B \cap A = \varnothing$.

5.7.2 Własność Hausdorffa

Twierdzenie 5.7.4 *Niech $\{\mathcal{B}_x\}_{x\in X}$ będzie bazą otoczeń w X. Jeśli $x_1, x_2 \in X$ i $x_1 \neq x_2$, to istnieją $B_1 \in \mathcal{B}_{x_1}$ oraz $B_2 \in \mathcal{B}_{x_2}$ takie, że $B_1 \cap B_2 = \varnothing$.*

5.7.3 Metryka w $\bar{\mathbb{R}}$

Twierdzenie 5.7.5 *Istnieje silnie rosnąca bijekcja $f : \bar{\mathbb{R}} \to [-1, 1]$. W szczególności jest taką funkcja zadana wzorem*

$$f(x) := \begin{cases} -1 & x = -\infty, \\ \frac{x}{1+|x|} & x \in \mathbb{R}, \\ 1 & x = \infty. \end{cases}$$

Jeśli $f : \bar{\mathbb{R}} \to [-1, 1]$ jest silnie rosnącą bijekcją, to funkcja $\rho_f : \bar{\mathbb{R}} \times \bar{\mathbb{R}} \to \mathbb{R}^*$ dana wzorem

$$\rho_f(x, y) := |f(x) - f(y)|$$

jest metryką w $\bar{\mathbb{R}}$. Co więcej, jeśli $f_1, f_2 : \bar{\mathbb{R}} \to [-1, 1]$ są dwoma silnie rosnącymi bijekcjami, to metryki ρ_{f_1} i ρ_{f_2} są równoważne.

W dalszym ciągu zbiór $\bar{\mathbb{R}}$ traktować będziemy jako przestrzeń metryczną z metryką ρ_f.

Twierdzenie 5.7.6 *Metryka indukowana z $\bar{\mathbb{R}}$ w $\mathbb{R}$ jest równoważna metryce euklidesowej w $\mathbb{R}$.*

Zbiór $\bar{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$ jest podzbiorem $\bar{\mathbb{R}}$, jest więc przestrzenią metryczną z metryką indukowaną z $\bar{\mathbb{R}}$. Łatwo sprawdzić, że w przestrzeni tej wszystkie punkty są izolowane, z wyjątkiem punktu ∞, który jest punktem skupienia zbioru $\mathbb{N} \subset \bar{\mathbb{N}}$.

Ponieważ kule w metryce ρ_f nie mają prostego opisu, wygodnie jest posługiwać się następującym twierdzeniem.

Twierdzenie 5.7.7 *Dla $x \in \bar{\mathbb{R}}$ połóżmy*

$$\mathcal{B}_x := \begin{cases} \{\,[-\infty, -n) \mid n \in \mathbb{N}\,\} & x = -\infty, \\ \{\,(x - \frac{1}{n}, x + \frac{1}{n}) \mid n \in \mathbb{N}\,\} & x \in \mathbb{R}, \\ \{\,(n, \infty] \mid n \in \mathbb{N}\,\} & x = \infty. \end{cases}$$

Wtedy $\{\mathcal{B}_x\}_{x \in \bar{\mathbb{R}}}$ stanowi bazę otoczeń w $\bar{\mathbb{R}}$.

5.8 (**)Przeliczalne bazy otoczeń punktu

Rodzina $\mathrm{Ball}_0(x)$ jest przeliczalna. Pozwala to w szczególności udowodnić następujące dwa twierdzenia.

Twierdzenie 5.8.1 *Zbiór $A \subset X$ jest domknięty wtedy i tylko wtedy gdy dla każdego ciągu $x : \mathbb{N} \to A$ jeśli x jest zbieżny do $x_0 \in X$ to $x_0 \in A$.*

Twierdzenie 5.8.2 *Punkt $x_0 \in X$ jest punktem skupienia zbioru $A \subset X$ wtedy i tylko wtedy gdy istnieje ciąg $a : \mathbb{N} \to A \setminus \{x_0\}$ zbieżny do x_0.*

5.9 Zupełność

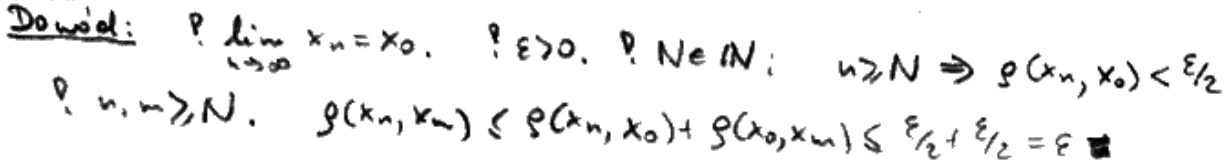
5.9.1 Warunek Cauchy'ego

Definicja 5.9.1 Niech (X, ρ) będzie przestrzenią metryczną. Mówimy, że ciąg $X : \mathbb{N} \to X$ spełnia *warunek Cauchy'ego* jeżeli

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \forall n, m \geqslant N \quad \rho(x_n, x_m) < \epsilon.$$

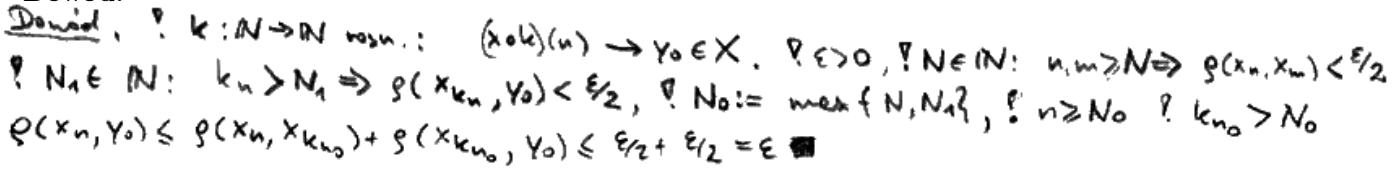
Twierdzenie 5.9.2 *Każdy ciąg zbieżny jest ciągiem Cauchy'ego.*

Dowód:



Twierdzenie 5.9.3 *(*)Jeśli ciąg Cauchy'ego posiada podciąg zbieżny to jest zbieżny.*

Dowód:



Twierdzenie 5.9.4 *(*)Każdy ciąg Cauchy'ego jest ograniczony*

Dowód:

Dowód. $\exists N \in \mathbb{N}: n,m \geq N \Rightarrow \rho(x_n, x_m) < 1$, $\{x_n\}_1^\infty \subset \bigcup_{i=1}^{N} K(x_i, 1) \cup K(x_N, 1)$ ■

Wniosek 5.9.5 *Każdy ciąg Cauchy'ego w $\mathbb{R}^d$ jest zbieżny.*

Dowód:

Dowód. Każdy ciąg ogr. w $\mathbb{R}^n$ posiada podc. zb. ■

(*)

5.9.2 Przestrzenie zupełne

Definicja 5.9.6 Przestrzeń metryczną nazywamy zupełną, jeżeli każdy ciąg Cauchy'ego w tej przestrzeni jest zbieżny.

Wniosek 5.9.7 *$\mathbb{R}^d$ jest przestrzenią zupełną.*

Przykład. $\mathbb{Q}$ nie jest przestrzenią zupełną.

Rozdział 6

Granica funkcji

6.1 Granica funkcji w przestrzeniach metrycznych

6.1.1 Eksperymenty numeryczne

Przy obliczeniach numerycznych granicy $\lim_{x\to x_0} f(x)$ nie możemy rozważyć wszystkich wartości argumentu x, więc w praktyce wybieramy ciąg x_n zmierzający do x_0 i obserwujemy zachowanie $f(x_n)$. Program w języku Mathematica obliczający przybliżenie granicy funkcji poprzez obliczanie wartości funkcji w ciągu punktów przedstawiono na listingu 6.1. Analogiczny przykład w C++ w odniesieniu do granicy $\lim_{x\to 0} \frac{\sin x}{x}$ i ciągu $x_n := \frac{1}{n}$ przedstawiono w programie z listingu 6.2.

Ćwiczenie komputerowe 6.1.1

- Sprawdź działanie programu z listingu 6.1 dla funkcji $x \mapsto \frac{\sin x}{x}$ i dla ciągu $x_n := \frac{1}{n^p}$ przy różnych wartosciach p.
- Sprawdź działanie programu z listingu 6.1 dla funkcji $x \mapsto \sin\frac{1}{x}$ i dla różnych ciągów.
- W programie z listingu 6.2 zmień ciąg x_n na $x_n := \frac{1}{n^p}$ dla różnych wartości p i zaobserwuj zachowanie się programu.
- Zmodyfikuj program z listingu 6.2 zastępując funkcję $x \mapsto \frac{\sin x}{x}$ funkcją $x \mapsto \sin\frac{1}{x}$ i zaobserwuj zachowanie się programu dla różnych ciągów.

Eksperymenty numeryczne pokazują, że definicję granicy funkcji lepiej byłoby postawić uniezależniając się od ciągów, bo zachowanie funkcji może być różne, przy podstawieniu w miejsce argumentów różnych ciągów. Taką definicję rozważaliśmy już w rozdziale wstępnym.

6.1.2 Definicja granicy funkcji w przestrzeniach metrycznych

Niech X, Y będą przestrzeniami metrycznymi odpowiednio z metrykami ρ_X i ρ_Y. Niech A będzie podzbiorem X. Rozważmy punkt skupienia x_0 zbioru A, punkt $y_0 \in Y$ oraz funkcję $f : A \to Y$.

Twierdzenie 6.1.2 *Następujące warunki są równoważne:*

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in A \setminus \{x_0\} \quad \rho_X(x, x_0) < \delta \ \Rightarrow\ \rho_Y(f(x), y_0) < \epsilon, \tag{6.1}$$

$$\forall V \in \mathrm{Nb}(y_0)\ \exists U \in \mathrm{Nb}(x_0)\ \forall x \in U \cap A \setminus \{x_0\} \quad f(x) \in V. \tag{6.2}$$

Dowód: Załóżmy, że warunek (6.1) zachodzi. Niech V będzie otoczeniem punktu y_0. Z definicji otoczenia znajdziemy kulę $K(y_0, r)$ w Y zawartą w V. Ponieważ $r > 0$, na mocy założenia (6.1) znajdziemy $\delta > 0$ takie, że dla wszystkich $x \in A \setminus \{x_0\}$

$$\rho_X(x, x_0) < \delta \ \Rightarrow\ \rho_Y(f(x), y_0) < r.$$

```
(* Przybliżanie granicy funkcji f granicą ciągu n wartości funkcji
   w punktach ciągu x *)
FunctionLimit[f_, x_, n_] :=
  Module[{k},
  k = 1;
  Do[
    Print[f[x[k]]];
    k = k + 1,
    {n}
  ];
]

(* Funkcja, której granicę w zerze liczymy *)
g[x_] := Sin[x]/x;

(* Ciąg o wyrazach zmierzających do zera *)
x[n_] := 1./n

(* Test przybliżania wartości funkcji *)
FunctionLimit[g, x, 100]
```

Listing 6.1: Przybliżanie granicy funkcji w języku Mathematica

Zatem dla $x \in A \setminus \{x_0\}$ mamy

$$x \in K(x_0, \delta) \;\Rightarrow\; y \in K(y_0, r) \subset V.$$

Ponieważ $K(x_0, \delta)$ jest otoczeniem x_0, więc dobraliśmy do V otoczenie U, którego istnienie jest postulowane w (6.2). Dowodzi to, że warunek (6.1) implikuje warunek (6.2). Dowód implikacji przeciwnej jest podobny. Zostawiamy go jako ćwiczenie. □

Twierdzenie 6.1.2 pozwala postawić następującą ogólną definicję granicy w przestrzeniach metrycznych.

Definicja 6.1.3 Mówimy, że $y_0 \in Y$ jest *granicą* funkcji f w punkcie x_0 i piszemy

$$\lim_{x \to x_0} f(x) = y_0$$

jeżeli zachodzi którykolwiek z równoważnych warunków (6.1) i (6.2).

6.1.3 Jednoznaczność granicy

Twierdzenie 6.1.4 *(o jednoznaczności granicy) Niech X, Y będą przestrzeniami metrycznymi, $A \subset X$, $f : X \to Y$ funkcją, a x_0 punktem skupienia A w X. Jeśli dla pewnych $y_1, y_2 \in Y$ mamy*

$$\lim_{x \to x_0} f(x) = y_1 \;\; i \;\; \lim_{x \to x_0} f(x) = y_2$$

to $y_1 = y_2$.

Dowód: Dla dowodu nie wprost załóżmy, że $y_1 \neq y_2$. Wtedy $r := \rho(y_1, y_2) > 0$. Z warunku (6.1) zastosowanego dla y_1 i y_2 do $\epsilon := \frac{r}{2}$ znajdziemy odpowiednio δ_1 i δ_2 takie, że dla $x \in A \setminus \{x_0\}$

$$\rho(x, x_0) < \delta_1 \quad \Rightarrow \quad \rho(f(x), y_1) < \epsilon \tag{6.3}$$

$$\rho(x, x_0) < \delta_2 \quad \Rightarrow \quad \rho(f(x), y_2) < \epsilon. \tag{6.4}$$

```
/********************* FunctionLimit.cpp*******************/
/* Przybliżanie granicy funkcji granicą ciągu wartości    */
/* funkcji                                                */
/**********************************************************/
#include <iostream>
#include <cmath>
#include <iomanip>
using namespace std;

/* Funkcja, której granicę w zerze liczymy */
double f(double x){
  return sin(x)/x;
}

/* Ciąg o wyrazach zmierzających do zera */
double x(int n){
  return 1.0/n;
}

/* W programie głównym wypisujemy ciąg wartości funkcji
   w zadanym ciągu punktów */
int main(){
  cout << fixed << right;
  cout << setprecision(15);

  for(int n=1;n<=20;n=n+1){
    cout << "x[" << setw(2) << n << "] =" << setw(18) << x(n);
    cout << "    f(x[" << setw(2) << n << "]) =";
    cout << setw(18) << f(x(n)) << endl;
  }
}
```

Listing 6.2: Przybliżanie granicy funkcji w C++

Niech $\delta := \min(\delta_1, \delta_2)$. Ponieważ x_0 jest punktem skupienia zbioru A w X, a $K(x_0, \delta)$ jest otoczeniem x_0 w X, więc znajdziemy punkt $x_* \in K(x_0, \delta) \cap A$. Ponieważ w szczególności $\rho(x_*, x_0) < \delta_1$ i $\rho(x_*, x_0) < \delta_2$, więc na mocy (6.3) i (6.4) mamy $\rho(f(x_*), y_1) < \epsilon$ i $\rho(f(x_*), y_2) < \epsilon$. Zatem z nierówności trójkąta

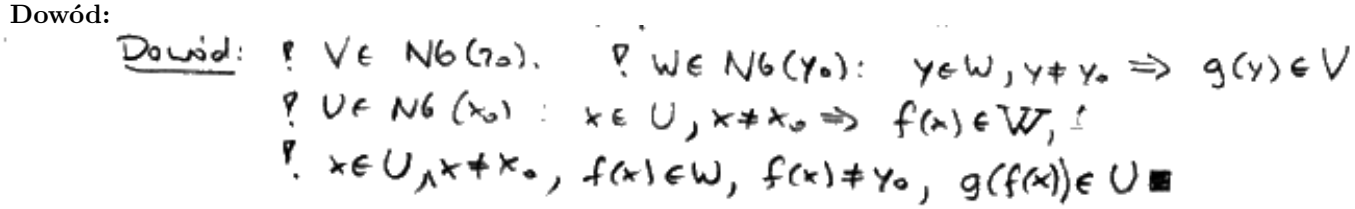
$$\rho(y_1, y_2) \leqslant \rho(y_1, f(x_*)) + \rho(f(x_*), y_2) < \epsilon + \epsilon = \rho(y_1, y_2).$$

Otrzymana sprzeczność dowodzi tezy. □

6.1.4 Granica złożenia funkcji

Twierdzenie 6.1.5 *(O granicy złożenia funkcji) Niech X, Y, Z będą przestrzeniami metrycznymi, $x_0 \in X'$, $y_0 \in Y'$, $z_0 \in Z$. Jeśli $f : X \to Y$ i $g : Y \to Z$ są takie, że $y_0 \notin \operatorname{im} f$ oraz $\lim_{x\to x_0} f(x) = y_0$ i $\lim_{y\to y_0} g(y) = z_0$ to $\lim_{x\to x_0}(g \circ f)(x) = z_0$.*

Dowód:



W programie Mathematica granicę $\lim_{x\to u} f(x)$ dla funkcji $f : \bar{R} \to \bar{\mathbb{R}}$ i $u \in \bar{\mathbb{R}}$ liczymy przy pomocy instrukcji

```
Limit[f[x], x -> u]
```

Na przykład

```
Limit[1/(1 + x), x -> Infinity]
```

zwraca 0.

6.2 (**)Przypadki szczególne definicji granicy

Przypomnijmy, że ciąg $x : \mathbb{N} \to \mathbb{R}$ ma granicę x_* wtedy i tylko wtedy gdy

$$\forall \epsilon > 0\, \exists N \in \mathbb{N} \quad n > N \;\Rightarrow\; |x_n - x_*| < \epsilon.$$

Natomiast funkcja $f : \mathbb{R} \to \mathbb{R}$ ma granicę f_* przy x zmierzającym do x_* wtedy i tylko wtedy gdy

$$\forall \epsilon > 0\, \exists \delta > 0 \quad |x - x_*| < \delta,\, x \neq x_* \;\Rightarrow\; |f(x) - f_*| < \epsilon.$$

Funkcja f może też mieć granicę nieskończoną. Mówimy, że funkcja $f : \mathbb{R} \to \mathbb{R}$ ma granicę ∞ przy x zmierzającym do x_* wtedy i tylko wtedy gdy

$$\forall A > 0\, \exists \delta > 0 \quad |x - x_*| < \delta,\, x \neq x_* \;\Rightarrow\; f(x) > A.$$

Powstaje naturalne pytanie, czy te trzy formalnie różne definicje nie dają się uogólnić do jednej definicji uniwersalnej. Podpowiedź jak to zrobić zawarta jest w następującym przeformułowaniu powyższych definicji.

$$\begin{aligned}
\lim_{n\to\infty} x_n = x_* \quad &:\Leftrightarrow \quad \forall \epsilon > 0\, \exists N \in \mathbb{N}\, \forall n \in (N, \infty] \setminus \{\infty\} \quad x_n \in (x_* - \epsilon, x_* + \epsilon),\\
\lim_{x\to x_*} f(x) = f_* \quad &:\Leftrightarrow \quad \forall \epsilon > 0\, \exists \delta > 0\, \forall x \in (x - \delta, x + \delta) \setminus \{x_*\} \quad f(x) \in (f_* - \epsilon, f_* + \epsilon),\\
\lim_{x\to x_*} f(x) = \infty \quad &:\Leftrightarrow \quad \forall A > 0\, \exists \delta > 0\, \forall x \in (x - \delta, x + \delta) \setminus \{x_*\} \quad f(x) \in (A, \infty].
\end{aligned}$$

Widać, że w każdym przypadku pojawiają się otoczenia: albo postaci $(x - r, x + r)$ dla $x \in \mathbb{R}$, albo postaci $(r, \infty]$ w przypadku punktu w nieskończoności. Tak więc zastępując te konkretne otoczenia w $\bar{\mathbb{R}}$ ogólnymi otoczeniami w przestrzeni metrycznej dostajemy dość ogólną definicję granicy.

6.2.1 Kryterium na granicę w bazach

Niech $\mathcal{B}_1$ będzie bazą otoczeń x_0, a $\mathcal{B}_2$ bazą otoczeń y_0.

Twierdzenie 6.2.1 *Warunek*

$$\lim_{x\to x_0} f(x) = y_0$$

jest równoważny warunkowi

$$\forall V \in \mathcal{B}_2\ \exists U \in \mathcal{B}_1\ \forall x \in U \cap A \setminus \{x_0\} \quad f(x) \in V.$$

6.2.2 Granica ciągu

Zobaczmy jak z naszego ogólnego kryterium na granicę w bazach można odzyskać naszą definicję granicy ciągu liczb. Jako przestrzeń X bierzemy $\bar{\mathbb{N}}$. Jako przestrzeń Y bierzemy $\mathbb{R}$. Jako metrykę w $\mathbb{R}$ bierzemy metrykę euklidesową. Jako zbiór A, na którym jest określona funkcja $f : A \to Y$ bierzemy $\mathbb{N}$. Funkcja określona na $\mathbb{N}$ to przecież ciąg, więc w miejsce f pisać będziemy bardziej tradycyjnie a, a jako argumentu zamiast x używać będziemy n. Jako punkt x_0, w którym liczymy granicę bierzemy ∞. Punkt ∞ jest rzeczywiście punktem skupienia zbioru $A = \mathbb{N}$ w przestrzeni $X = \bar{\mathbb{N}}$. Jako bazę otoczeń dla $x_0 = \infty$ bierzemy

$$\mathcal{B}_1 := \{B(N) := (N, \infty] \mid N \in \mathbb{N}\}.$$

Jako bazę otoczeń dla y_0 bierzemy $\mathcal{B}_2 := \mathrm{Ball}(y_0)$. Podstawiając to wszystko do naszej ogólnej definicji dostajemy

$$\begin{aligned}\lim_{n\to\infty} a_n = y_0 \quad &:\Leftrightarrow \quad \forall K = K(y_0,\epsilon) \in \mathrm{Ball}(y_0)\ \exists B = B(N) \in \mathcal{B}_1\ \forall n \in B \cap A \setminus \{\infty\} \quad a_n \in K \\ &\Leftrightarrow \quad \forall \epsilon > 0\ \exists N \in \mathbb{N}\ \forall n > N \quad |a_n - y_0| < \epsilon,\end{aligned}$$

a więc dokładnie taki warunek jaki postawiliśmy jako definicję granicy ciągu.

6.2.3 Granica funkcji w przestrzeni metrycznej

Definicję Cauchy'ego granicy funkcji $\mathbb{R}$ w $\mathbb{R}$ odzyskujemy biorąc $A = X = Y = \mathbb{R}$, jako metrykę biorąc metrykę euklidesową w $\mathbb{R}$, a jako bazy otoczeń $\mathcal{B}_1$ i $\mathcal{B}_2$ rodziny kul w x_0 i y_0. Dostajemy

$$\begin{aligned}\lim_{x\to x_0} f(x) = y_0 \quad &:\Leftrightarrow \quad \forall K = K(y_0,\epsilon) \in \mathrm{Ball}(y_0)\ \exists B = K(x_0,\delta) \in \mathrm{Ball}(x_0)\ \forall x \in B \cap A \setminus \{x_0\} \quad f(x) \in K \\ &\Leftrightarrow \quad \forall \epsilon > 0\ \exists \delta > 0\ \forall x \neq x_0 \quad |x - x_0| < \delta \ \Rightarrow\ |f(x) - y_0| < \epsilon,\end{aligned}$$

6.2.4 Granica nieskończona w $\bar{\mathbb{R}}$

Przypadek nieskończonej wartości granicy funkcji uzyskujemy biorąc $Y = \bar{\mathbb{R}}$, $y_0 = \infty$, $\mathcal{B}_1 = \mathrm{Ball}(x_0)$ i $\mathcal{B}_2 = \{(E, \infty] \mid E \in \mathbb{R}^+\}$. Dostajemy

$$\begin{aligned}\lim_{x\to x_0} f(x) = \infty \quad &:\Leftrightarrow \quad \forall B = (E,\infty] \in \mathcal{B}_1\ \exists K = K(x_0,\delta) \in \mathrm{Ball}(x_0)\ \forall x \in K \cap A \setminus \{x_0\} \quad f(x) \in B \\ &\Leftrightarrow \quad \forall E > 0\ \exists \delta > 0\ \forall x \in A \setminus \{x_0\} \quad \rho(x, x_0) < \delta \ \Rightarrow\ f(x) > E,\end{aligned}$$

Podobnie wygląda przypadek granicy wynoszącej $-\infty$. Za bazę otoczeń $\mathcal{B}_2$ bierzemy jedynie $\{[-\infty, -E) \mid E \in \mathbb{R}^+\}$

6.2.5 Granica w nieskończoności

Przypadek granicy w nieskończoności dla funkcji $f : X \to \mathbb{R}$ uzyskujemy biorąc $X = \bar{\mathbb{R}}$, $A = \mathbb{R}$, $x_0 = \infty$, $\mathcal{B}_1 = \{ (E, \infty] \mid E \in \mathbb{R}^+ \}$, $\mathcal{B}_2 = \mathrm{Ball}(y_0)$. Dostajemy

$$\begin{aligned} \lim_{x\to\infty} f(x) = y_0 \quad &:\Leftrightarrow \quad \forall K = K(y_0, \epsilon) \in \mathrm{Ball}(y_0)\ \exists B = (E, \infty] \in \mathcal{B}_1\ \forall x \in B \cap A \setminus \{\infty\} \quad f(x) \in K \\ &\Leftrightarrow \quad \forall \epsilon > 0\ \exists E > 0\ \forall x \in \mathbb{R} \quad x > E \ \Rightarrow\ |f(x) - y_0| < \epsilon, \end{aligned}$$

Podobnie wygląda przypadek granicy w $-\infty$. Za bazę otoczeń $\mathcal{B}_1$ bierzemy jedynie $\{[-\infty, -E) \mid E \in \mathbb{R}^+\}$

6.3 Granice jednostronne

Niech Y będzie przestrzenią metryczną. Rozważmy $f : \mathbb{R} \to Y$ i punkt $x_0 \in \mathbb{R}$.

Definicja 6.3.1 Mówimy, że f ma w x_0 granicę lewostronną równą $y \in Y$ i piszemy

$$\lim_{x\to x_0^-} f(x) = y,$$

jeżeli $f_{|(-\infty,x_0]} : (-\infty, x_0] \to Y$ ma w x_0 granicę y. Analogicznie, mówimy, że f ma w x_0 granicę prawostronną równą $y \in Y$ i piszemy

$$\lim_{x\to x_0^+} f(x) = y,$$

jeżeli $f_{|[x_0,\infty)} : [x_0, \infty) \to Y$ ma w x_0 granicę y.

Twierdzenie 6.3.2 *Funkcja f ma w x_0 granicę $y \in Y$ wtedy i tylko wtedy gdy obie granice jednostronne istnieją i są sobie równe.*

W programie Mathematica granicę jednostronną $\lim_{x\to u^-} f(x)$ liczymy przy pomocy instrukcji

```
Limit[f[x], x -> u, Direction -> 1]
```

a granicę jednostronną $\lim_{x\to u^+} f(x)$ liczymy przy pomocy instrukcji

```
Limit[f[x], x -> u, Direction -> -1]
```

Na przykład

```
Limit[1/x, x -> 0, Direction -> 1]
```

zwraca $-\infty$.

6.4 Punkty graniczne

6.4.1 Ciągowa charakterystyka granicy

Ustalmy dwie przestrzenie metryczne X, Y oraz funkcję $f : X \to Y$ i punkt $x_0 \in X'$.

Twierdzenie 6.4.1 *Następujące dwa warunki są sobie równoważne*

(i) *istnieje granica*

$$\lim_{x\to x_0} f(x) = y_0,$$

(ii) *dla każdego ciągu* $x : \mathbb{N}_1 \to X \setminus \{x_0\}$ *zachodzi*

$$\lim_{n\to\infty} x_n = x_0 \;\Rightarrow\; \lim_{n\to\infty} f(x_n) = y_0.$$

Dowód:

Dowód: $\Rightarrow$ wniosek z tw. o gran. złożenia

$\Leftarrow$ ① $\neg \lim_{x\to x_0} f(x) = g$, $\exists\, V \in \mathcal{N}_6(g)\ \forall\, U \in \mathcal{N}_6(x_0)\ \exists\, x \in U : f(x) \notin V$

Niech $\mathcal{B} = \{B_n : n \in \mathbb{N}\}$ będzie przeliczalną, malejącą bazą ot. X.

Do B_n znajdę $x_n \in B_n$: $f(x_n) \notin V$, $x : \mathbb{N} \ni n \to x_n \in X$.

? $\lim_{n\to\infty} x_n = x_0$? ! $U \in \mathcal{N}_6(x_0)$! $N \in \mathbb{N}$: $B_N \subset U$, ! $n > N$, $x_n \in B_n \subset B_N \subset U$ □

Pon. $\forall_n\ f(x_n) \notin V \in \mathcal{N}_6(y_0)$ więc $\neg\, f(x_n) \to g$ □■

6.4.2 Zbiór punktów granicznych

Definicja 6.4.2 Niech $f : X \to Y$ będzie funkcją odwzorowującą przestrzeń metryczną X w przestrzeń metryczną Y. Niech $x_0 \in X$ będzie punktem skupienia X w X. Definiujemy *zbiór punktów granicznych* funkcji f w x_0 jako

$$\operatorname{Lim}_{x\to x_0} f(x) := \{\, y \in Y \mid \exists a : \mathbb{N} \to X \ \lim_{n\to\infty} a_n = x_0 \text{ i } \lim_{n\to\infty} f(a_n) = y \,\}.$$

Twierdzenie 6.4.3 *Zbiór* $\operatorname{Lim}_{x\to x_0} f(x)$ *jest domknięty.*

Dowód:

Dowód: ! $A := \operatorname{Lim}_{x\to x_0} f(x)$, ! $y \in \operatorname{cl} A$, ! $\mathcal{B} = \{B_n : n \in \mathbb{N}\}$ zstępująca baza ot. otw. y

! $\mathcal{C} = \{C_n : n \in \mathbb{N}\}$ zstęp. baza ot. otw. x_0. ! $n \in \mathbb{N}$, $\exists\, y_n \in A \cap B_n$ ∧ $\exists\, (x_n : \mathbb{N} \to Y)$ (n-ty ciąg)

$\lim_{m\to\infty} f(x_n(m)) = y_n \wedge \lim_{m\to\infty} x_n(m) = x_0$. Dla dost. dużych m: $x_n(m) \in C_n$ ∧ $f(x_n(m)) \in B_n$

Mogę więc skonstruować ciąg $m : \mathbb{N} \to \mathbb{N}$ taki, że

$x_n(m(n)) \in C_n \wedge f(x_n(m(n)) \in B_n$. ! $w_n = x_n(m(n))$. $\lim_{n\to\infty} w_n = x_0$, $\lim_{n\to\infty} f(w_n) = y$ ■

Uwaga $\operatorname{Lim}_{x\to x_0} f(x) = \{y\}$ nie gwarantuje istnienia granicy, bo:

n.p. $a : \mathbb{N} \to \mathbb{R}$, $a_n = (1 + (-1)^n)n$ nie jest zbieżny, ale $\operatorname{Lim}_{n\to\infty} a(n) = \{0\}$

6.4.3 Ciągowa charakterystyka granicy ciągu

Twierdzenie 6.4.4 *Niech $a : \mathbb{N} \to Y$ będzie ciągiem w przestrzeni metrycznej Y, a $y \in Y$. Następujące warunki są równoważne.*

(i) $lim_{n\to\infty} a_n = y$,

(ii) $\forall k : \mathbb{N} \to \mathbb{N} \quad \lim_{n\to\infty} k_n = \infty \;\Rightarrow\; \lim_{n\to\infty}(a \circ k)(n) = y$,

(iii) $\forall k : \mathbb{N} \to \mathbb{N} \quad k$ *silnie rosnący* $\;\Rightarrow\; \lim_{n\to\infty}(a \circ k)(n) = y$,

(iv) Dla każdego a' podciągu a jest $\lim_{n\to\infty} a'_n = y$.

Twierdzenie 6.4.5 *Niech $a : \mathbb{N} \to Y$ będzie ciągiem w przestrzeni metrycznej Y. Wtedy*

$$\operatorname*{Lim}_{n\to\infty} a_n = \{\, y \in Y \mid \exists k \mathbb{N} \to \mathbb{N} \text{ silnie rosnący i taki, że } \lim a_{k_n} = y \,\}.$$

6.5 Granica dolna i górna

6.5.1 Granica dolna i górna

Niech X będzie przestrzenią metryczną, $Y \subset \bar{\mathbb{R}}$ oraz $f : X \to Y$. Załóżmy ponadto, że dla pewnego $x_0 \in X'$ zbiór $\operatorname{Lim}_{x\to x_0} f(x)$ jest niepusty.

Definicja 6.5.1 W przypadku gdy $\operatorname{Lim}_{x\to x_0} f(x)$ jest ograniczony od góry, granicę *górną* definiujemy wzorem

$$\limsup_{x\to x_0} f(x) := \sup \operatorname*{Lim}_{x\to x_0} f(x).$$

Analogicznie, gdy $\operatorname{Lim}_{x\to x_0} f(x)$ jest ograniczony od dołu, granicę *dolną* definiujemy wzorem

$$\liminf_{x\to x_0} f(x) := \inf \operatorname*{Lim}_{x\to x_0} f(x).$$

Zobaczymy później, że w przypadku $Y = \bar{\mathbb{R}}$ zbiór punktów granicznych jest zawsze niepusty, a ponieważ każdy zbiór w $\bar{\mathbb{R}}$ jest ograniczony tak od góry jak i od dołu, więc granica górna i dolna są w tym przypadku zawsze określone.

Twierdzenie 6.5.2 *Niech $f : X \to \bar{\mathbb{R}}$ oraz $x_0 \in X'$. Wtedy $a = \lim_{x\to x_0} f(x)$ jest równoważne równości*

$$\liminf_{x\to x_0} f(x) = \limsup_{x\to x_0} f(x).$$

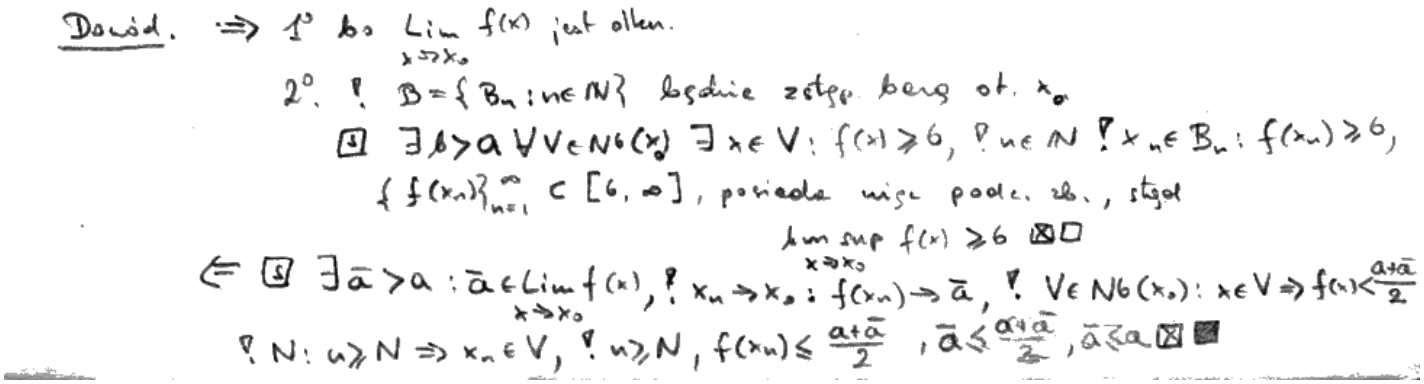
W przypadku gdy $Y = \bar{\mathbb{R}}$ możemy podać przydatne kryterium na granicę górną i dolną, a także twierdzenie o zachowywaniu nierówności przez te granice. W dowodach wykorzystujemy nieudowodniony jeszcze fakt, że każdy ciąg w $\bar{\mathbb{R}}$ posiada podciąg zbieżny. Fakt ten udowodnimy w rozdziale 7.

6.5.2 Własności granicy dolnej i górnej

Twierdzenie 6.5.3 *Niech $f : X \to \bar{\mathbb{R}}$ oraz $x_0 \in X'$. Wtedy $a = \limsup_{x\to x_0} f(x)$ jest równoważne koniunkcji następujących dwóch warunków.*

$$a \in \operatorname*{Lim}_{x\to x_0} f(x),$$
$$\forall b > a\ \exists V \in \mathrm{Nb}(x_0) \quad x \in V \ \Rightarrow\ f(x) < b.$$

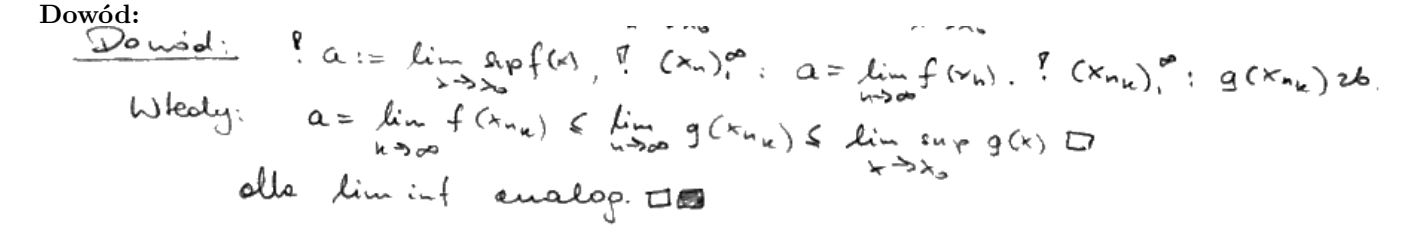
Dowód:



Twierdzenie 6.5.4 *Niech $f, g : X \to \bar{\mathbb{R}}$ oraz $x_0 \in X'$. Jeśli istnieje $V \in \mathrm{Nb}(x_0)$ takie, że dla każdego $x \in V \setminus \{x_0\}$ mamy $f(x) \leqslant g(x)$ to*

$$\begin{aligned} \liminf_{x\to x_0} f(x) &\leqslant \liminf_{x\to x_0} g(x), \\ \limsup_{x\to x_0} f(x) &\leqslant \limsup_{x\to x_0} g(x). \end{aligned}$$

Dowód:



6.6 Granica, a struktura algebraiczna

6.6.1 Sumy, różnica, iloczyn i iloraz funkcji

Niech X będzie przestrzenią metryczną, a x_0 punktem skupienia X. Niech Y będzie ciałem $\mathbb{R}$ lub ciałem $\mathbb{C}$.

Załóżmy, że dane są funkcje

$$f_1, f_2 : X \to Y.$$

Niech $\diamond : Y \times Y$ będzie działaniem w ciele Y. Definiujemy funkcję

$$f_1 \diamond f_2 : X \ni x \mapsto f_1(x) \diamond f_2(x) \in Y.$$

Dla $\diamond \in \{+, -, \cdot, /\}$ nazywamy ją odpowiednio *sumą, różnicą, iloczynem i ilorazem* funkcji f_1 i f_2.

6.6.2 Granica sumy, różnicy i iloczynu

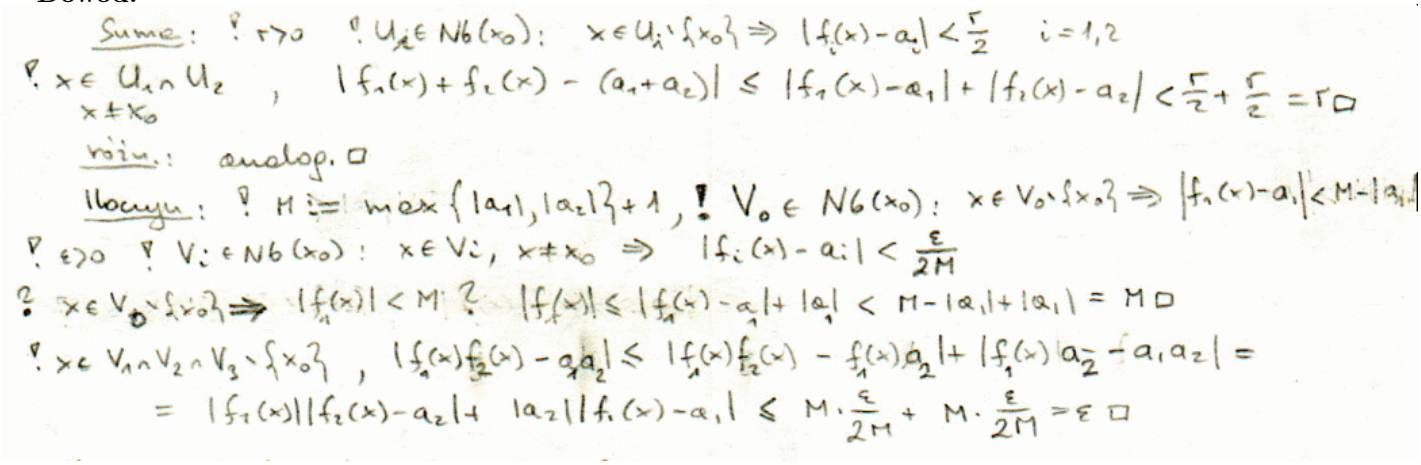
Twierdzenie 6.6.1 *Załóżmy, że f_1, f_2 mają w x_0 odpowiednio granice $a_1, a_2 \in Y$, tzn.*

$$\lim_{x \to x_0} f_i(x) = a_i.$$

Wtedy istnieją granice sumy, różnicy i iloczynu funkcji f_1, f_2 w x_0 i wynoszą odpowiednio

$$\begin{aligned} \lim_{x \to x_0} (f_1(x) + f_2(x)) &= a_1 + a_2, \\ \lim_{x \to x_0} (f_1(x) - f_2(x)) &= a_1 - a_2, \\ \lim_{x \to x_0} (f_1(x) f_2(x)) &= a_1 a_2. \end{aligned}$$

Dowód:



6.6.3 Granica ilorazu

Twierdzenie 6.6.2 *Przy założeniach poprzedniego twierdzenia, jeśli ponadto $a_2 \neq 0$, to istnieje też granica ilorazu i wynosi*

$$\lim_{x \to x_0} \frac{f_1(x)}{f_2(x)} = \frac{a_1}{a_2}.$$

Dowód:

Uwaga. Wyst. pokazać dla $f_1(x)=1$ i skorzystać z włas. iloczynu.
$\exists\, c\in(0,|a_2|)$, $\exists\, U\in N6(x_0)$: $x\in U\setminus\{x_0\} \Rightarrow |f_2(x)-a_2|<|a_2|-c$
$\exists\, x\in U, x\neq x_0 \Rightarrow |f_2(x)|>c$? $|f_2(x)| = |a_2-(f_2(x)-a_2)| > |a_2|-|f_2(x)-a_2| > |a_2|-|a_2|+c=c$
$\forall\,\varepsilon>0$ $\exists\, V\in N6(x_0)$: $x\in V_2, x\neq x_0 \Rightarrow |f_2(x)-a_2|<\varepsilon c^2$
$\exists\, x\in U\cap V$. $\left|\frac{1}{f_2(x)}-\frac{1}{a_2}\right| = \frac{|a_2-f_2(x)|}{|a_2||f_2(x)|} < \frac{\varepsilon c^2}{c^2} = \varepsilon$ □

Powyższe twierdzenia zachodzą też w przypadku $Y=\bar{\mathbb{R}}$, ale tylko wtedy, gdy odpowiednie działania mają sens.

6.7 Granica, a struktura porządkowa

6.7.1 Zachowywanie nierówności w granicy

Twierdzenie 6.7.1 *Niech $f,g: X\to\bar{\mathbb{R}}$, $x\in X'$. Jeśli istnieje otoczenie V punktu x_0 takie, że dla każdego $x\in X\setminus\{x_0\}$ mamy $f(x)\leqslant g(x)$, to*

$$\lim_{x\to x_0} f(x) \leqslant \lim_{x\to x_0} g(x).$$

Dowód:

Dowód: ⊡ $a>b$, $\exists\, U_1\in Nb(a), U_2\in Nb(b)$: $U_1\cap U_2=\emptyset$. Możemy przyjąć, że $U_1=(p,q)$, $U_2=(r,s)$, $r<s\leq p<q$
$\exists\, V_1\in N6(x_0)$: $x\in V_1, x\neq x_0 \Rightarrow f(x)\in U_1$
$\exists\, V_2\in N6(x_0)$: $x\in V_2, x\neq x_0 \Rightarrow g(x)\in U_2$
$\exists\, x\in V\cap V_1\cap V_2$, $f(x)\in U_1$, $g(x)\in U_2$, $g(x)<s\leq p<f(x)$ ⊠

6.7.2 Twierdzenie o trzech funkcjach

Twierdzenie 6.7.2 *Niech $f,g,h: X\to\bar{\mathbb{R}}$, $x\in X'$. Jeśli istnieje otoczenie V punktu x_0 takie, że dla każdego $x\in X\setminus\{x_0\}$ mamy $f(x)\leqslant g(x)\leqslant h(x)$ oraz istnieją granice*

$$\lim_{x\to x_0} f(x), \quad \lim_{x\to x_0} h(x)$$

i obie są równe pewnemu $a\in\bar{\mathbb{R}}$, to istnieje granica $\lim_{x\to x_0} g(x)$ i też wynosi a.

Dowód:

Dowód: $\forall U \in Nb(x_0)$. Możemy przyjąć że $U=(p,q) \vee U=(p,\infty] \vee U=[-\infty,q)$ i $p<x_0<q$

1. $U=(p,q)$ $\exists V_1 \in Nb(x_0): x \in V_1 \setminus \{x_0\} \Rightarrow f(x) \in U$

$\exists V_2 \in Nb(x_0): x \in V_2 \setminus \{x_0\} \Rightarrow h(x) \in U$

$\exists x \in V \cap V_1 \cap V_2$, $p<f(x) \leq g(x) \leq h(x)<q$, $g(x) \in U$ □

pozost. przypadki podobnie □■

6.7.3 Granica funkcji monotonicznej

Twierdzenie 6.7.3 *Niech $A \subset \bar{\mathbb{R}}$ i $\omega := \sup A$. Jeśli ω jest punktem skupienia zbioru A, a $f : A \to \bar{\mathbb{R}}$ jest funkcją rosnącą i ograniczoną od góry, to istnieje granica $\lim_{x\to\omega} f(x)$ i jest równa*

$$\sup\{\, f(x) \mid x \in A, x < \omega \,\}.$$

Dowód:

Dowód: $\exists \mu := \sup_{x \in A \setminus \{\omega\}} f(x)$, $\forall V \in Nb(\mu)$. Możemy przyjąć że $V=(a,b) \ni \mu$,

$\exists \bar{x} \in A \setminus \{\omega\}: f(\bar{x})>a$, $\exists U := (\bar{x}, \omega)$

$\forall x \in U$, $x>\bar{x}$, $f(x) \geq f(\bar{x})>a$, $x<\omega$, $f(x) \leq \mu < b$, $f(x) \in V$ □■

6.8 Granice w produkcie kartezjańskim

Twierdzenie 6.8.1 *Niech X oraz $Y_1, Y_2, \ldots Y_n$ będą przestrzeniami metrycznymi, a $f = (f_1, f_2, \ldots f_n) : X \to Y_1 \times Y_2 \times \cdots \times Y_n$ niech będzie zestawieniem funkcji $f_i : X \to Y_i$ dla $i = 1, 2, \ldots, n$. Niech x_0 będzie punktem skupienia przestrzeni X oraz niech $y := (y_1, y_2, \ldots, y_n) \in Y_1 \times Y_2 \times \cdots \times Y_n$. Wtedy*

$$\lim_{x\to x_0} f(x) = y \;\Leftrightarrow\; \lim_{x\to x_0} f_i(x) = y_i \text{ dla } i = 1, 2, \ldots n.$$

Dowód: (**)Dowód przedstawimy dla przypadku $n = 2$.

Dowód: $\Rightarrow$ $\forall V_i \in Nb(y_i)$ $i=1,2$, $V_1 \times V_2 \in Nb(y)$, $\exists U \in Nb(x_0)$ $x \in U \setminus \{x_0\} \Rightarrow f(x) \in V_1 \times V_2$

$\forall x \in U \setminus \{x_0\}$, $f(x) \in V_1 \times V_2$, $f_1(x) \in V_1$, $f_2(x) \in V_2$ $\square$

$\Leftarrow$ Wyst. zadał otoczenie post. $V_1 \times V_2$ gdzie $V_i \in Nb(y_i)$ $i=1,2$.

$\forall U_i \in Nb(x_0)$: $x \in U_i \setminus \{x_0\} \Rightarrow f_i(x) \in V_i$

$\forall x \in (U_1 \cap U_2) \setminus \{x_0\}$, $f_i(x) \in V_i$; $i=1,2$, $f(x) \in V_1 \times V_2$ $\square$■

Rozdział 7

Ciągłość funkcji

7.1 Koncepcja ciągłości

7.1.1 Definicja ciągłości

Niech $f : X \to Y$ będzie odwzorowaniem przestrzeni metrycznych.

Definicja 7.1.1 Mówimy, że f jest *ciągła* w punkcie $x_0 \in X$ jeżeli dla każdego otoczenia $V \in \mathrm{Nb}(f(x_0))$ istnieje otoczenie $U \in \mathrm{Nb}(x_0)$ takie, że

$$x \in U \;\Rightarrow\; f(x) \in V.$$

Mówimy, że f jest *ciągła* jeśli jest ciągła w każdym punkcie $x \in X$.

Twierdzenie 7.1.2 *Funkcja f jest ciągła w x_0 wtedy i tylko wtedy gdy zachodzi jeden z następujących dwóch warunków*

$$x_0 \notin X' \quad \textit{lub} \quad x_0 \in X' \ i \ \lim_{x \to x_0} f(x) = f(x_0).$$

Dowód:

Dowód: $\Rightarrow$ ① $x_0 \notin A' \square$, $x_0 \in A' \square$

$\Leftarrow$ ① $x_0 \notin A'$ $\exists U \in Nb(x_0) : U \cap A = \{x_0\}$: U dobre dla wszystkich V

$x_0 \in A'$ $\forall V \in Nb(x_0)$ $\exists U \in Nb(x_0)$: $x \in U \cap A$, $x \neq x_0$ $\Rightarrow f(x) \in V$

ale dla $x = x_0$ oczywiste $\square$■

Twierdzenie 7.1.3 *Niech $\mathcal{B}_1$ będzie bazą otoczeń x_0, a $\mathcal{B}_2$ bazą otoczeń $f(x_0)$. Wtedy ciągłość $f : X \to Y$ w x_0 jest równoważna każdemu z trzech poniższych warunków*

$$\forall V \in \mathcal{B}_2 \; \exists U \in \mathcal{B}_1 \quad f(U) \subset V,$$
$$\forall V \in \mathcal{B}_2 \; \exists U \in \mathcal{B}_1 \quad U \subset f^{-1}(V),$$
$$\forall V \in \mathcal{B}_2 \quad f^{-1}(V) \in \mathrm{Nb}(x_0).$$

Dowód:

Dow.: $\Rightarrow$ $\forall V \in \mathcal{B}$ $\exists U \in Nb(x_0)$: $x \in U \Rightarrow f(x) \in V$, $\exists U' \in \mathcal{B}_1$: $U' \subset U$, $f(U') \subset V$ □

$\Leftarrow$ $\forall V \in Nb(f(x_0))$ $\exists V' \in \mathcal{B}_2$: $V' \subset V$, $\exists U \in \mathcal{B}_1$: $f(U) \subset V'$, $f(U) \subset V$ □■

7.1.2 Kryteria ciągłości

Twierdzenie 7.1.4 *Funkcja $f : X \to Y$ jest ciągła wtedy i tylko wtedy gdy dla każdego zbioru otwartego V w Y zbiór $f^{-1}(V)$ jest otwarty w X.*

Dowód:

Dowód: $\Rightarrow$ $\forall V \in Open(Y)$, $\forall x \in f^{-1}(V)$, $V \in Nb(f(x))$, $f^{-1}(V) \in Nb(x)$, $f^{-1}(V) \in Open(X)$ □

$\Leftarrow$ $\forall x \in X$ $\forall V \in ONb(f(x))$, $f^{-1}(V) \in Open(X)$, $x \in f^{-1}(V)$, $f^{-1}(V) \in Nb(x)$ □■

Twierdzenie 7.1.5 *Funkcja $f : X \to Y$ jest ciągła wtedy i tylko wtedy gdy dla każdego zbioru domkniętego K w Y zbiór $f^{-1}(K)$ jest domknięty w X.*

7.1.3 Ciągłość złożenia

Twierdzenie 7.1.6 *Niech X, Y, Z będą przestrzeniami metrycznymi, a $f : X \to Y$ i $g : Y \to Z$ odwzorowaniami. Jeśli f jest ciągłe w $x_0 \in X$, a g jest ciągłe w $f(x_0)$, to $g \circ f$ jest ciągłe w x_0. Zatem jeśli f i g są ciągłe, to $g \circ f$ jest ciągłe.*

Dowód: łatwe ćwiczenie.

7.1.4 Ciągłość operacji arytmetycznych

Niech X będzie przestrzenią metryczną, a Y niech będzie ciałem liczb rzeczywistych lub ciałem liczb zespolonych.

Twierdzenie 7.1.7 *Jeśli $f, g : X \to Y$ są ciągłe w $x_0 \in X$ to $f + g$, $f - g$, $f \cdot g$ są ciągłe w x_0. Jeśli dodatkowo $g(x_0) \neq 0$, to f/g jest ciągłe w x_0.*

Dowód: łatwe ćwiczenie.

7.1.5 Twierdzenie o lokalnym zachowywaniu znaku

Twierdzenie 7.1.8 *Niech $f : X \to \mathbb{R}$ będzie funkcją ciągłą. Jeśli $f(x_0) > 0$ dla pewnego $x_0 \in X$, to istnieje U otoczenie punktu x_0 takie, że dla każdego $x \in U$ mamy $f(x) > 0$.*

Dowód:

Dow. $(0,\infty) \in \mathrm{Open}(\mathbb{R})$, $U := f^{-1}((0,\infty))$, $U \in \mathcal{N}_6(x_0)$ ■

Twierdzenie 7.1.9 *Niech X, Y będą przestrzeniami metrycznymi.*

(i) *Niech $c \in Y$. Funkcja stała $c_X : X \ni x \mapsto c \in Y$ jest ciągła.*

(ii) *Niech $X \subset Y$. Inkluzja $j : X \ni x \mapsto x \in Y$ jest ciągła.*

(iii) *Identyczność $\mathrm{id}_X : X \ni x \mapsto x \in X$ jest ciągła.*

(iv) *Odwzorowania rzutowania*

$$p_X : X \times Y \ni (x,y) \mapsto x \in X$$
$$q_X : X \times Y \ni (x,y) \mapsto y \in Y$$

są ciągłe.

(v) *Zawężenie $f_{|A}$ odwzorowania ciągłego $f : X \to Y$ do zbioru $A \subset X$ jest ciągłe.*

Dowód: łatwe ćwiczenie.

Twierdzenie 7.1.10 *Niech $P : \mathbb{C} \to \mathbb{C}$ będzie wielomianem. Wtedy P jest ciągłe.*

Dowód: łatwe ćwiczenie.

Niech $P, Q : \mathbb{C} \to \mathbb{C}$ będą wielomianami. *Funkcją wymierną* jest funkcja

$$\frac{P}{Q} : \mathbb{C} \setminus Q^{-1}(0) \ni x \mapsto \frac{P(x)}{Q(x)} \in \mathbb{C}.$$

Twierdzenie 7.1.11 *Funkcja wymierna jest ciągła.*

Dowód: łatwe ćwiczenie.

7.1.6 (*)Nieciągłości funkcji $f : \mathbb{R} \to \mathbb{R}$

Niech $f : (a,b) \to \mathbb{R}$ i niech $x_0 \in (a,b)$.

Definicja 7.1.12 Mówimy, że f ma w x_0 nieciągłość pierwszego rodzaju, jeżeli f nie jest ciągła w x_0, ale istnieją granice jednostronne $\lim_{x \to x_0^-} f(x)$ i $\lim_{x \to x_0^+} f(x)$. W przeciwnym razie mówimy, że f ma w x_0 nieciągłość drugiego rodzaju.

Twierdzenie 7.1.13 *Funkcja f ma w x_0 nieciągłość pierwszego rodzaju wtedy i tylko wtedy gdy $\lim_{x \to x_0^-} f(x) \neq \lim_{x \to x_0^+} f(x)$ lub $\lim_{x \to x_0^-} f(x) = \lim_{x \to x_0^+} f(x) \neq f(x_0)$.*

Twierdzenie 7.1.14 *Jeśli $f : (a, b) \to \mathbb{R}$ jest funkcją słabo rosnącą, to każdy jej punkt nieciągłości jest pierwszego rodzaju. Ponadto*

$$\forall x_0 \in (a,b) \quad \lim_{x \to x_0^-} f(x) \leqslant f(x_0) \leqslant \lim_{x \to x_0^+} f(x),$$
$$\forall x_1, x_2, x_3 \in (a,b) \quad x_1 < x_2 < x_3 \;\Rightarrow\; \lim_{x \to x_1^+} f(x) \leqslant f(x_2) \leqslant \lim_{x \to x_3^-} f(x).$$

Analogiczne własności ma funkcja słabo malejąca.

Dowód:

Dowód: Z tw. o granicy funkcji monotonicznej istnieją granice jednostronne, a stąd każdy punkt jest punktem nieciągłości I rodzaju. Własności (*), (**) są konsekwencją tw. o przejściu do granicy w nierówności. ■

7.1.7 Ciągłość bijekcji monotonicznej

Twierdzenie 7.1.15 *Niech (a, b) i (c, d) będą otwartymi przedziałami w $\mathbb{R}$ i niech $f : (a, b) \to (c, d)$ będzie bijekcją rosnącą. Wtedy f jest ciągła.*

Dowód:

Dowód: ☐ $\exists x_0 \in (a,b)$: $\alpha := \lim_{x \to x_0^-} f(x) < \lim_{x \to x_0^+} f(x) =: \beta$. ! $\gamma \in (\alpha, \beta) \subset (c,d)$, $\exists x_1 \in (a,b)$: $\gamma = f(x_1)$

? $x_1 = x_0$? ☐ ☐ $x_1 < x_0$! $\bar{x}$: $x_1 < \bar{x} < x_0$ $\gamma = f(x_1) < f(\bar{x}) \leqslant \lim_{x \to x_0^-} f(x) = \alpha$ ☒

$x_0 < x_1$ analog. ☒☐

! $\gamma_1, \gamma_2 \in (\alpha, \beta)$: $\gamma_1 \neq \gamma_2$, $\gamma_1 = f(x_0) = \gamma_2$ ☒■

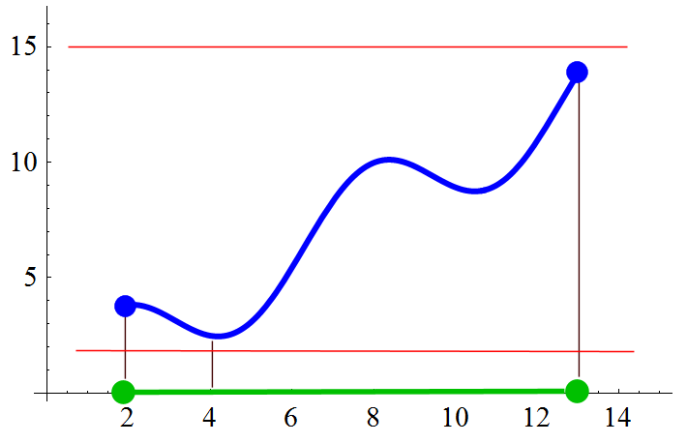
7.2 Ciągłość, a zwartość

Funkcje ciągłe mają wiele pożytecznych własności. Jednakże wiele z tych własności wymaga, by zbiory na których funkcje są określone miały pewne sprzyjające tym własnościom cechy. Do cech takich należą rozważane w tym rozdziale zwartość, spójność oraz zupełność.

7.2.1 Zwartość podprzestrzeni $\mathbb{R}^d$.

Przyczyny braku zbieżności mogą być różne. Ciąg

$$a : \mathbb{N}_1 \ni n \mapsto \frac{1}{n} + (-1)^n \to \mathbb{R}$$



Rysunek 7.1: Funkcja ciągła na zbiorze domkniętym i ograniczonym jest ograniczona

nie jest zbieżny, bo nie może się zdecydować czy zbiegać do 1 czy do -1. Inaczej: można z niego wybrać podciągi zbieżne do dwóch różnych granic. Z kolei ciąg

$$a : \mathbb{N}_1 \ni n \mapsto 2^n \to \mathbb{R},$$

w ogóle nie posiada podciągu zbieżnego. Ale temu można zaradzić zmieniając przestrzeń $\mathbb{R}$ na $\bar{\mathbb{R}}$. Wtedy ten ciąg staje się zbieżny do $+\infty$. Można pokazać, że w $\bar{\mathbb{R}}$ każdy ciąg posiada podciąg zbieżny. Przestrzenie topologiczne, które mają taką własność są z wielu powodów mile widziane. Do takich zaliczają się przestrzenie zwarte. Ponieważ ich ogólna definicja jest nieco skomplikowana, podamy ją oddzielnie w rozdz. 7.2.3. Jednak w następującym ważnym przypadku szczególnym definicja jest prosta.

Definicja 7.2.1 Niech X będzie podprzestrzenią przestrzeni metrycznej $\mathbb{R}^d$ z metryką euklidesową. Mówimy, że przestrzeń X jest zwarta jeśli X jako podzbiór $\mathbb{R}^d$ jestdomknięty i ograniczony.

Domknięte i ograniczone podzbiory w $\mathbb{R}^n$ mają też inne przydatne własności. Rysunek 7.1 pokazuje, że należy oczekiwać iż funkcja ciągła f na domkniętym i ograniczonym podzbiorze $\mathbb{R}^n$ jest ograniczona.

Rysunek 7.2 z kolei pokazuje, że założenie iż zbiór jest ograniczony jest istotne.

Co więcej, sama domkniętość przedziału nie wystarcza. Przedział $[0, \infty)$, też jest domknięty, ale funkcja

$$[0, \infty) \ni x \mapsto x^2 \in \mathbb{R}$$

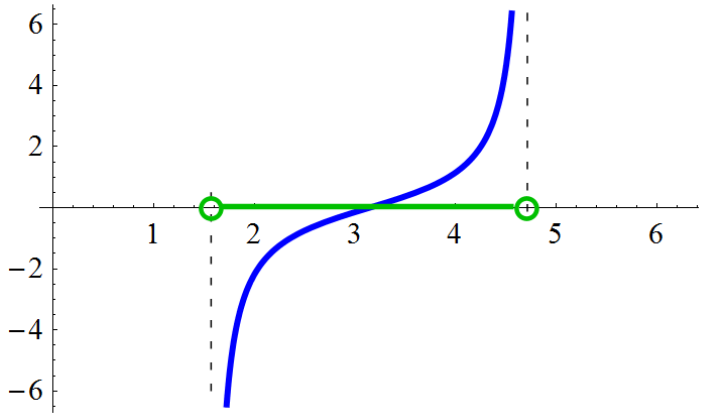
jest nieograniczona. Zatem istotne jest, by rozważać przedział domknięty i ograniczony. Rozważania te sugerują, że prawdziwe powinno być następujące twierdzenie (formułujemy je dla najprostszego przypadku).

Twierdzenie 7.2.2 *Niech $f : [a, b] \to \mathbb{R}$ będzie funkcją ciągłą. Wtedy f jest funkcją ograniczoną.*

7.2.2 (*)Intuicja zwartości i ogólna definicja zwartości.

Potencjalny dowód tw. 7.2.2 mógłby przebiegać tak: Do $y \in \operatorname{im} f$ dobieram $x \in [a, b]$ takie, że $f(x) = y$ oraz $V_x \in \mathrm{ONb}(x)$ takie, że

$$u \in V_x \;\Rightarrow\; f(u) \in [y - 1, y + 1].$$



Rysunek 7.2: Funkcja ciągła na przedziale otwartym (a, b) nie musi być ograniczona

Mamy

$$[a, b] \subset \bigcup \{ V_x \mid x \in [a, b] \}.$$

Mówimy, że rodzina zbiorów $\{ V_x \mid x \in [a, b] \}$ stanowi pokrycie otawrte przedziału $[a, b]$. Gdyby było możliwe zastąpienie tego pokrycia podpokryciem skończonym, to znaczy wybranie skończonej ilości punktów $x_1, x_2, \ldots x_k \in [a, b]$, tak, by

$$[a, b] \subset \bigcup_{i=1}^{k} V_{x_i},$$

to otrzymalibyśmy

$$f([a, b]) \subset \bigcup_{i=1}^{k} [f(x_i) - 1, f(x_i) + 1]$$

a stąd $f([a, b]) \subset [c, d]$, gdzie

$$c := \min \{ f(x_i) - 1 \mid i = 1, 2, \ldots k \} \quad d := \max \{ f(x_i) + 1 \mid i = 1, 2, \ldots k \}$$

i dowód byłby zakończony. Okazuje się, że jedynymi podzbiorami w $\mathbb{R}^n$, które mają własność istnienia podpokrycia skończonego dla dowolnego pokrycia otwartego są zbiory domknięte i ograniczone. Niestety nie jest to prawdą w dowolnej przestrzeni metrycznej. Ponieważ pomysł polegający na wybraniu z nieskończonego pokrycia zbioru pewnego podpokrycia skończonego przydaje się w wielu dowodach, dlatego warto wyróżnić klasę zbiorów mających taką własność. Zbiory takie nazywane są zbiorami zwartymi.

Co więcej, okazuje się, że zwartość zbioru nie zależy od topologii przestrzeni, w której zbiór jest zanurzony, a tylko od topologii indukowanej w zbiorze. Dlatego mówimy też o przestrzeniach zwartych.

Niech X będzie przestrzenią metryczną i $A \subset X$.

Definicja 7.2.3 Mówimy, że zbiór A jest *zwarty* jeżeli dla każdej rodziny $\mathcal{C} \subset \mathrm{Open}(X)$ takiej, że $A \subset \bigcup \mathcal{C}$ istnieje skończona podrodzina $\mathcal{C}' \subset \mathcal{C}$ taka, że $A \subset \bigcup \mathcal{C}'$.

Definicja 7.2.4 Mówimy, że przestrzeń metryczna X jest zwarta, jeżeli X traktowane jako zbiór w przestrzeni X jest zbiorem zwartym.

Definicja 7.2.5 Klasę przestrzeni zwartych oznaczać będziemy Comp.

7.2.3 Przestrzenie zwarte

Niech X będzie ustaloną przestrzenią metryczną. Poniższe twierdzenie pokazuje, że zwartość jest własnością zbioru zależną jedynie od metryki indukowanej w tym zbiorze, a nie od metryki w całej przestrzeni.

Twierdzenie 7.2.6 *Niech $Y \subset X$ będzie podprzestrzenią X, a $A \subset Y$. Zbiór A jest zwarty w Y wtedy i tylko wtedy gdy jest zwarty w X.*

Twierdzenie 7.2.7 *Każdy zwarty zbiór $A \subset X$ jest domknięty.*

Twierdzenie 7.2.8 *Podzbiór domknięty zbioru zwartego jest zbiorem zwartym.*

Twierdzenie 7.2.9 *Niech $K \subset X$ będzie zbiorem zwartym, a $E \subset K$ jego nieskończonym podzbiorem. Wtedy zbiór punktów skupienia $E' \neq \varnothing$.*

Twierdzenie 7.2.10 *Niech $E \subset \mathbb{R}^d$. Następujące warunki są równoważne.*

(i) E jest domknięty i ograniczony.

(ii) E jest zwarty.

(iii) Każdy nieskończony podzbiór zbioru E ma punkt skupienia należący do E.

Zatem ogólna definicja zwartości w przypadku podzbiorów $\mathbb{R}^d$ jest równoważna definicji, którą podaliśmy wcześniej.

Twierdzenie 7.2.11 *Przestrzeń $\bar{\mathbb{R}}$ jest zwarta.*

7.2.4 Obraz zbioru zwartego przez odwzorowanie ciągłe.

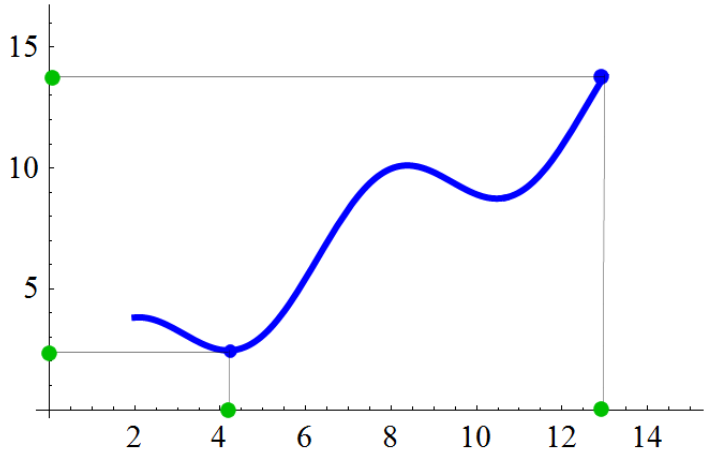
Ustalmy przestrzenie metryczne X, Y oraz funkcję $f : X \to Y$.

Twierdzenie 7.2.12 *Jeśli f jest ciągła, a $A \subset X$ jest zwarty, to $f(A)$ jest zwarty.*

Dowód: (*)

Dowód: ! $\{C_\alpha : \alpha \in A\}$ pokrycie otw. $f(K)$, $\{f^{-1}(C_\alpha) : \alpha \in A\}$ pokr. otw. K,
$K \subset \bigcup_{i=1}^{n} f^{-1}(C_{\alpha_i})$, $f(K) \subset \bigcup_{i=1}^{n} C_{\alpha_i}$ ■

□



Rysunek 7.3: Funkcja ciągła na przedziale zwartym osiąga swoje kresy

Wniosek 7.2.13 *Zachodzą następujące własności.*

(i) Jeśli $f : X \to \mathbb{R}^d$ jest ciągła, a $K \subset X$ jest zwarty, to $f(K)$ jest domknięty i ograniczony.

(ii) Jeśli przestrzeń X jest zwarta, a $f : X \to \mathbb{R}^d$ jest ciągła, to f jest ograniczona.

(iii) Jeśli $f : [a, b] \to \mathbb{R}^d$ jest ciągła, to jest ograniczona.

7.2.5 Funkcja ciągła na przedziale zwartym osiąga swoje kresy

Twierdzenie 7.2.14 *Jeśli X jest przestrzenią zwartą, a $f : X \to \mathbb{R}$ jest ciągła oraz $m := \inf f(X)$, $M := \sup f(X)$, to istnieją $p, q \in X$ takie, że $f(p) = m$ i $f(q) = M$.*

Dowód:

Dowód. $f(X)$ jako zwarta jest dom., ohpod $m, M \in f(X)$ ■

7.2.6 Funkcja odwrotna do funkcji ciągłej

Twierdzenie 7.2.15 *Jeśli $f : X \to Y$ jest ciągłą bijekcją, a X jest zwarty, to f^{-1} jest ciągłe.*

Dowód:

Dowód: $\forall A$ dom $\subset X$, $(f^{-1})^{-1}(A) = f(A)$ jest zwarty, więc dom. ■

Przykład. $\mathbb{R}^+ \ni x \longrightarrow \sqrt[n]{x} \in \mathbb{R}^+$ jest ciągłe

Dowód: $f: \mathbb{R}^+ \ni y \to y^n \in \mathbb{R}^+$ jest ciągłe. $\forall K_a := [0,a]$, $f^{-1}|_{K_a}$ ciągłe

$f^{-1} = \bigcup_{a \in \mathbb{R}^+} f^{-1}|_{K_a}$, f^{-1} cg., bo cg. w ot. każdego punktu

7.3 Jednostajna ciągłość

Definicja 7.3.1 Funkcja $f : X \to Y$ jest *jednostajnie ciągła* wtedy i tylko wtedy gdy

$$\forall \epsilon > 0 \; \exists \delta > 0 \; \forall x, y \in X \quad \rho_X(x,y) < \delta \;\Rightarrow\; \rho_Y(f(x), f(y)) < \epsilon.$$

Przykład $\mathbb{R} \ni x \to x^2 \in \mathbb{R}$ nie jest jednost. cg., bo:

$\forall \epsilon = 1$, $\forall \delta$: $|x-y| < \delta \Rightarrow |x^2 - y^2| < 1$

$\forall x_n := \sqrt{n} + \frac{1}{\sqrt{n}}$, $y_n = \sqrt{n}$, $|x_n - y_n| = \frac{1}{\sqrt{n}}$ $\forall n_0$: $\frac{1}{\sqrt{n_0}} < \delta$

$|x_n^2 - y_n^2| = (\sqrt{n} + \frac{1}{\sqrt{n}})^2 - (\sqrt{n})^2 = n + 2 + \frac{1}{n} - n = 2 + \frac{1}{n} > 1$ ⊠

Twierdzenie 7.3.2 *Jeśli $f : X \to Y$ jest jednostajnie ciągła, to f jest ciągła.*

Twierdzenie 7.3.3 *Jeśli $f : X \to Y$ jest ciągła, a X jest przestrzenią zwartą, to f jest jednostajnie ciągła.*

Dowód:

Dowód: $\forall \varepsilon > 0$ $\forall x \in X$, $\forall \delta_x \in \mathbb{R}^+$: $x' \in K(x, \delta_x) \Rightarrow f(x') \in K(f(x), \varepsilon/2)$,

$X \subset \bigcup_{x \in X} K(x, \delta_x/2)$, $X \subset \bigcup_{i=1}^{n} K(x_i, \delta_{x_i}/2)$, $\forall \delta := \min\{\delta_{x_i}\}_1^n$,

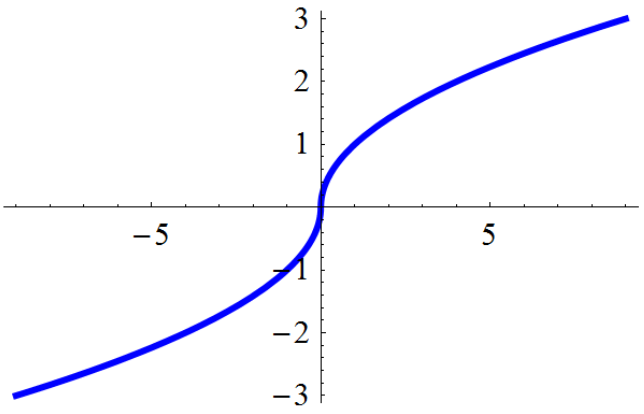
$\forall x, y \in X : \rho(x,y) < \delta/2$, $\exists i : x \in K(x_i, \delta/2)$, $y \in K(x_i, \delta)$, $\rho(f(x), f(x_i)) < \varepsilon/2$,

$\rho(f(y), f(x_i)) < \varepsilon/2$, $\rho(f(x), f(y)) < \varepsilon$ ■

7.3.1 Funkcje Lipschitza

Definicja 7.3.4 $f : X \to Y$ jest *funkcją Lipschitza* jeżeli istnieje stała $L > 0$ taka, że

$$\forall x_1, x_2 \in X \quad \rho_Y(f(x_1), f(x_2)) \leqslant L \rho_X(x_1, x_2).$$



Rysunek 7.4: Funkcja jednostajnie ciągła, która nie jest funkcją Lipschitza

Twierdzenie 7.3.5 *Jeśli $f : X \to Y$ jest funkcją Lipschitza, to jest funkcją jednostajnie ciągłą.*

Przykład. $!\, X \in NVect$, $f: X \ni x \to \|x\| \in \mathbb{R} \Longrightarrow$ f lipschitzowa ze st. 1

Dow.: $\|x\| = \|x-y+y\| \leq \|x-y\|+\|y\|$, $\|x\|-\|y\| \leq \|x-y\|$, $\big|\|x\|-\|y\|\big| \leq \|x-y\|$ □

7.4 Ciągłość, a spójność

7.4.1 Spójne podzbiory $\bar{\mathbb{R}}$.

Intuicja przestrzeni spójnej jest bardzo prosta. Jest to taka przestrzeń, która nie rozpada się na dwa lub więcej kawałków. Formalna definicja jest jednak nieco skomplikowana. Dlatego najpierw podamy definicję spójności dla przestrzeni metrycznych będących podzbiorami $\bar{\mathbb{R}}$

Definicja 7.4.1 Niech $X \subset \bar{\mathbb{R}}$ będzie podprzestrzenią. Mówimy, że X jest przestrzenią spójną jeżeli X jest przedziałem w $\bar{\mathbb{R}}$.

7.4.2 (*)Przestrzenie spójne

Niech X będzie przestrzenią metryczną.

Definicja 7.4.2 Mówimy, że zbiór $E \subset X$ jest *niespójny*, jeżeli istnieją rozłączne zbiory $A, B \in \mathrm{Open}(X)$ takie, że $A \cap E \neq \varnothing \neq B \cap E$ oraz $E \subset A \cup B$. Jeśli takie zbiory nie istnieją, mówimy, że E jest *spójny*.

Twierdzenie 7.4.3 *Niech $E \subset \bar{\mathbb{R}}$. E jest zbiorem spójnym, wtedy i tylko wtedy gdy*

$$\forall x, y, z \in E \quad z \in (x, y) \;\Rightarrow\; z \in E.$$

Wniosek 7.4.4 *Zbiory spójne w $\bar{\mathbb{R}}$ to przedziały.*

Zatem ogólna definicja spójności w przypadku podzbiorów $\bar{\mathbb{R}}$ jest równoważna definicji, którą już podaliśmy.

7.4.3 Obraz zbioru spójnego przez odwzorowanie ciągłe.

Twierdzenie 7.4.5 *Niech $f : X \to Y$ będzie odwzorowaniem ciągłym. Jeśli $E \subset X$ jest spójny, to $f(E)$ jest zbiorem spójnym.*

Dowód:

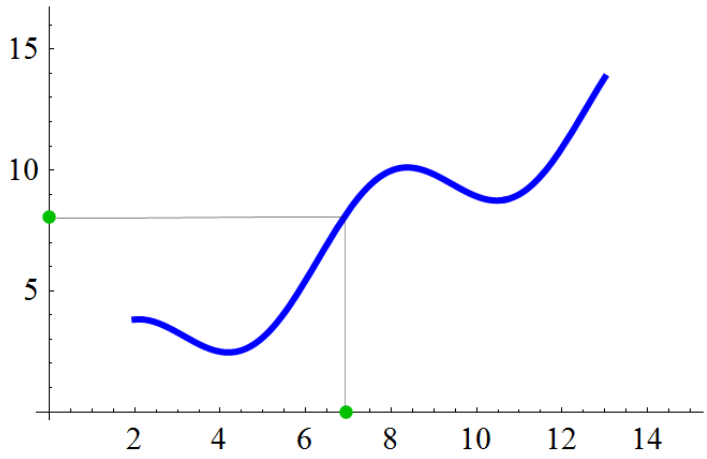
Dowód: $\boxed{\cdot}$! $A, B \in \mathrm{Open}(Y)$: $A \cap B = \emptyset \wedge f(E) \subset A \cup B \wedge A \cap f(E) \neq \emptyset \wedge B \cap f(E) \neq \emptyset$.
! $A' := f^{-1}(A) \in \mathrm{Open}(X)$, ! $B' := f^{-1}(B) \in \mathrm{Open}(X)$, $A' \cap B' = \emptyset$, $E \subset A' \cup B'$,
$A' \cap E \neq 0$, $B' \cap E \neq \emptyset$, E niesp. $\boxtimes$ ■

Wniosek 7.4.6 *Jeśli $a, b \in \bar{\mathbb{R}}$, $a < b$ i $f : [a, b] \to \bar{\mathbb{R}}$ jest odwzorowaniem ciągłym, to $f([a, b])$ jest przedziałem domkniętym.*

7.4.4 Własność Darboux.

Wniosek 7.4.7 *(własność Darboux) Jeśli $a, b \in \bar{\mathbb{R}}$, $a < b$, $f : [a, b] \to \bar{\mathbb{R}}$ jest odwzorowaniem ciągłym oraz $f(a) < f(b)$ to dla każdego $c \in [f(a), f(b)]$ istnieje $\xi \in [a, b]$ takie, że $f(\xi) = c$.*

Własność Darboux zobrazowano na rys. 7.5



Rysunek 7.5: Własność Darboux.

Rozdział 8

Szeregi

8.1 Eksperymenty numeryczne

8.1.1 Przybliżanie liczby π polami wielokątów wpisanych w koło.

W rozdziale 1.3.2 zaprezentowaliśmy metodę przybliżania liczby π poprzez pola wielokątów wpisanych w koło i zauważyliśmy, że stosowny ciąg (1.4) jest szeregiem. Przyjrzyjmy się teraz zachowaniu tego szeregu od strony numerycznej Program wyznaczający wyrazy ciągu (1.4) w Mathematica jest przedstawiony na listingu 8.1, a w języku C++ na listingu 8.2

Uruchamiając program łatwo sprawdzić, że metoda oparta o pola przy zastosowaniu tego samego typu **double** *daje znacznie dokładniejsze wyniki niż metoda poprzednia, bo już 26-ta suma częściowa daje wartość*

$$s_{26} = 3.141592653589793,$$

która jest poprawnym przybliżeniem liczby π do ostatniej cyfry włącznie.

8.1.2 Warunek konieczny zbieżności szeregu

Twierdzenie 8.1.1 *Jeśli szereg $\sum_{n=1}^{\infty} c_n$ jest zbieżny, to ciąg c_n zmierza do zera.*

Nie jest to jednak warunek wystarczający, bo na przykład o szeregu

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

zwanym szeregiem harmonicznym, można pokazać, że nie jest zbieżny w sensie definicji Cauchy'ego.

Pogrupujmy wyrazy szeregu harmonicznego w porcje kolejno po jeden, dwa, cztery itd.

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \left(\frac{1}{2} + \frac{1}{3}\right) + \left(\frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}\right) + \dots$$

```
(* Obliczanie przybliżenia liczby pi poprzez pola wielokątów foremnych
   wpisanych w okrąg o średnicy jeden                                    *)
PiByField[d_] :=
  Module[{s, m, n, a},
    n = 2;
    m = 4; (* Ilość boków wielokąta, na początek kwadrat *)
    a = Sqrt[2.]/2;      (* Bok kolejnych wielokątów  *)
    s = 2; (*  czterokrotne pole kolejnego wielokąta *)
    Do[
      Print["s[", n, "]=", NumberForm[s, 16]];
      s = s + m*a*(1 - Sqrt[1 - a*a]);
      a = Sqrt[2 - Sqrt[4 - 4*a*a]]/2;
      m = m*2;
      n = n + 1,
      {d}
    ];
    s
  ]
```

Listing 8.1: Wyznaczenie ciągu przybliżeń liczby π w języku Mathematica w oparciu o pola wielokątów

Tak więc sumy kolejnych porcji to

$$\begin{aligned} T_0 &:= 1 \\ T_1 &:= \frac{1}{2}+\frac{1}{3} \\ T_2 &:= \frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7} \\ T_3 &:= \frac{1}{8}+\frac{1}{9}+\frac{1}{10}+\frac{1}{11}+\frac{1}{12}+\frac{1}{13}+\frac{1}{14}+\frac{1}{15} \\ \dots &\quad \dots \\ T_k &:= \sum_{n=2^k}^{2^{k+1}-1}\frac{1}{n}. \end{aligned}$$

Na listingach 8.3 i 8.4 przedstawiono program odpowiednio w Mathematica i w C++, który liczy sumy kolejnych porcji T_k dla szeregu harmonicznego.

Wynik działania programu sugeruje, że sumy te są większe od pewnej stałej $c > 0$, co w dalszym toku udowodnimy formalnie. Na razie zaobserwujmy, że konsekwencją tego faktu jest

$$\sum_{n=1}^{2^m}\frac{1}{n} \geqslant mc.$$

Zatem sumy częściowe szeregu harmonicznego mogą być większe od dowolnie zadanej liczby.

8.1.3 Rozbieżność do nieskończoności

Mówimy, że ciąg x_n jest rozbieżny do nieskończoności, jeżeli zachodzi warunek

$$\forall M > 0 \exists N > 0 \forall n \in \mathbb{N} \quad n > N \;\Rightarrow\; x_n > M.$$

Możemy więc powiedzieć, że sumy częściowe szeregu $\sum_{n=1}^{\infty}\frac{1}{n}$ są rozbieżne do nieskończoności.

```
/************************* PiByField.cpp ****************************/
/* Obliczanie przybliżenia liczby pi poprzez pola wielokątów foremnych */
/* wpisanych w okrąg o promieniu jeden                                 */
/********************************************************************/
#include <iostream>
#include <cmath>
#include <iomanip>
using namespace std;

int main(){
  cout << fixed << right;
  cout << setprecision(15);
  int m=4; /* Ilość boków wielokąta, na początek kwadrat. */
  /* Bok kolejnych wielokątów wpisanych w koło o średnicy jeden.
     Na początek bok kwadratu */
  double a=sqrt(2.0)/2;
  /* Zmienna s zawiera czterokrotne pole kolejnego wielokąta.
     Zaczynamy od kwadratu, który ma pole 1/2. */
  double s=2;
  int nmax=30;
  for(int n=2;n<nmax;n=n+1){
    cout << "   s[" << setw(2) << n << "] =" << setw(18) << s;
    cout << endl;
    /* Czterokrotne pole kolejnego wielokąta
       przyrasta o czterokrotne pole m trókątów  równobocznych */
    s=s+m*a*(1-sqrt(1-a*a));
    m=m*2; /* Podwajamy ilość boków */
    a=sqrt(2-sqrt(4-4*a*a))/2;  /* I wyznaczamy kolejny bok */
  }
}
```

Listing 8.2: Wyznaczenie ciągu przybliżeń liczby π w C++ w oparciu o pola wielokątów

```
(* Analiza zachowania szeregu harmonicznego *)
HarmonicSeries[d_] := Module[{s, n, k},
  k = 0;
  Do[
    s = 0.;
    n = 2^k;
    (* Obliczamy sumę wyrazów szeregu od 2^k do 2^{k+1}-1 *)
    Do[
      s = s + 1./n;
      n = n + 1,
      {2^k}
    ];
  Print["T[", k, "]= ", NumberForm[s, 16]];
  k = k + 1,
  {d}
  ];
]
```

Listing 8.3: Obliczanie sumy $\sum_{n=2^k}^{2^{k+1}-1} \frac{1}{n}$ w Mathematica

```
/********************** HarmonicSeries.cpp ***************************/
/* Analiza zachowania szeregu harmonicznego                          */
/*********************************************************************/
#include <iostream>
#include <cmath>
#include <iomanip>
using namespace std;

int main(){
  cout << fixed << right;
  cout << setprecision(15);

  double m=1; /* Będziemy sumować w porcjach po m dla m=1,2,4,... */
  for(int k=0;k<25;k=k+1){
    double s=0;
    /* Obliczamy sumę wyrazów szeregu od m=2^k do 2m-1=2^{k+1}-1 */
    for(double n=m;n<m*2;++n){
      s=s+1/n;
    }
    cout << "T[" << k << "]= " << s << endl;
    m=m*2; /* Kolejna porcja będzie dwa razy większa */
  }
}
```

Listing 8.4: Obliczanie sumy $\sum_{n=2^k}^{2^{k+1}-1} \frac{1}{n}$ w C++

Co ciekawe, szereg

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

jest zbieżny.

Ćwiczenie komputerowe 8.1.2 Adaptuj program z listingu 8.3 i/lub listingu 8.4 do prześledzenia zachowania szeregów

$$\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}} \quad \text{i} \quad \sum_{n=1}^{\infty}\frac{1}{n^2}.$$

Co sądzisz na temat ich zbieżności?

8.2 Własności podstawowe

8.2.1 Pojęcie szeregu

Niech X będzie albo ciałem $\mathbb{R}$ albo ciałem $\mathbb{C}$.

Definicja 8.2.1 Dla $k \in \mathbb{N}$ parę ciągów $x, s : \mathbb{N}_k \to X$ nazywamy *szeregiem* jeżeli

$$s_n = \sum_{i=k}^{n} x_i.$$

Element x_n nazywamy n-tym *wyrazem* szeregu, a element s_n nazywamy n-tą *sumą częściową* szeregu. Mówimy, że szereg jest *zbieżny*, jeżeli ciąg jego sum częściowych jest zbieżny. Jeśli ciąg sum częściowych jest zbieżny, to jego granicę nazywamy *sumą szeregu* i oznaczamy

$$\sum_{i=k}^{\infty} x_i.$$

W praktyce symbolem $\sum_{i=k}^{\infty} x_i$ zazwyczaj oznaczamy nie tylko sumę szeregu, co jest sensowne tylko jeśli szereg jest zbieżny, ale także sam szereg. To czy mamy na myśli szereg, czy jego sumę, rozstrzygamy na podstawie kontekstu.

Twierdzenie 8.2.2 *Szereg $\sum_{i=k}^{\infty} x_i$ jest zbieżny wtedy i tylko wtedy gdy*

$$\forall \epsilon > 0 \exists N \in \mathbb{N}_k \forall m \geqslant n \geqslant N \quad |\sum_{i=n}^{m} x_i| < \epsilon.$$

Dowód:
Własność ta oznacza, że ciąg sum częściowych szeregu spełnia warunek Cauchy'ego. Zatem teza wynika z zupełności przestrzeni $\mathbb{R}$ lub $\mathbb{C}$.

□

Wniosek 8.2.3 *Jeśli szereg $\sum_{n=1}^{\infty} x_n$ jest zbieżny, to $\lim_{n\to\infty} x_n = 0$.*

Definicja 8.2.4 Szereg $\sum_{i=k}^{\infty} x_i$ nazywamy *bezwzględnie zbieżnym*, jeżeli szereg $\sum_{i=k}^{\infty} |x_i|$ jest zbieżny.

Twierdzenie 8.2.5 *Jeśli szereg $\sum_{i=k}^{\infty} x_i$ jest bezwzględnie zbieżny to jest zbieżny oraz*

$$|\sum_{i=k}^{\infty} x_i| \leqslant \sum_{i=k}^{\infty} |x_i|.$$

Dowód:
Ponieważ ciąg sum częściowych dla szeregu wartości bezwzględnych spełnia warunek Cauchy'ego, więc, jak łatwo sprawdzić, również ciąg sum częściowych wyjściowego szeregu spełnia ten warunek i teza wynika z zupełności $\mathbb{R}$ lub $\mathbb{C}$.

□

8.2.2 Szereg geometryczny

Definicja 8.2.6 Niech $a, q \in \mathbb{C}$, $a \neq 0$. Szereg $\sum_{n=0}^{\infty} aq^n$, nazywamy szeregiem geometrycznym.

Twierdzenie 8.2.7 *Szereg geometryczny jest zbieżny gdy $|q| < 1$. W przeciwnym razie jest on rozbieżny.*

Dowód: Wiemy, że

$$s_n = a\frac{1-q^{n+1}}{1-q}.$$

Zatem, jeśli $|q| < 1$, to

$$\lim_{n\to\infty} s_n = \frac{a}{1-q},$$

czyli szereg jest zbieżny. Natomiast jeśli $|q| \geqslant 1$, to ciąg aq^n na pewno nie jest zbieżny do zera, więc nie jest spełniony warunek konieczny zbieżności. □

Twierdzenie 8.2.8 *(Kryterium porównawcze) Jeśli ciągi $a, b : \mathbb{N} \to \mathbb{R}$ mają wyrazy nieujemne i istnieje $N \in \mathbb{N}$ takie, że $a_n \leqslant b_n$ dla $n \geqslant N$, to:*

- *jeśli $\sum_{n=1}^{\infty} b_n$ jest zbieżny, to $\sum_{n=1}^{\infty} a_n$ jest zbieżny,*
- *jeśli $\sum_{n=1}^{\infty} a_n$ jest rozbieżny, to $\sum_{n=1}^{\infty} b_n$ jest rozbieżny.*

Dowód:
Jest to wniosek z zupełności przestrzeni $\mathbb{R}$.

□

Wniosek 8.2.9 *Niech $\sum_{n=k}^{\infty} z_n$ będzie szeregiem liczb zespolonych, a $\sum_{n=k}^{\infty} x_n$ zbieżnym szeregiem liczb rzeczywistych. Jeśli istnieje $n \in \mathbb{N}$ takie, że $|z_n| \leqslant x_n$ dla $n \geqslant N$, to szereg $\sum_{n=k}^{\infty} z_n$ jest bezwzględnie zbieżny.* □

8.3 Kryteria zbieżności

8.3.1 Szereg $\sum_{n=1}^{\infty} \frac{1}{n^p}$.

Twierdzenie 8.3.1 *Niech $a : \mathbb{N} \to \mathbb{R}$ będzie ciągiem o wyrazach nieujemnych. Szereg $\sum_{n=1}^{\infty} a_n$ jest zbieżny wtedy i tylko wtedy gdy jego ciąg sum częściowych jest ograniczony.*

Dowód:

Dow.: $S_n = S_{n-1} + a_n \geqslant S_{n-1}$ jest monotoniczny □

Twierdzenie 8.3.2 *Niech ciąg $a : \mathbb{N}_1 \to \mathbb{R}^+$ będzie słabo malejący. Rozważmy ciąg*

$$b : \mathbb{N}_0 \ni n \mapsto 2^n a_{2^n} \in \mathbb{R}^+.$$

Wtedy zbieżność szeregu $\sum_{n=1}^{\infty} a_n$ jest równoważna zbieżności szeregu $\sum_{n=0}^{\infty} b_n$.

Dowód:

Dow.: ! $S_n := a_1 + \ldots + a_n$, $t_n := a_1 + 2a_2 + 4a_4 + \ldots 2^n a_{2^n}$

Pokażemy, że:

1° $S_n \leqslant t_n$

2° $t_n \leqslant 2 S_n$

ad 1°. $S_n = a_1 + \ldots + a_n \leqslant a_1 + \ldots + a_{2^{n+1}-1} = a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + \ldots$

$+ (a_{2^n} + \ldots + a_{2^{n+1}-1}) \leqslant a_1 + 2a_2 + 4a_4 + \ldots 2^n a_{2^n} = t_n$ □

ad 2°: $S_{2^n} = a_1 + \ldots + a_{2^n} = a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \ldots + (a_{2^{n-1}+1} + \ldots + a_{2^n}) \geqslant$

$\geqslant \frac{1}{2} a_1 + a_2 + 2a_4 + 4a_8 + \ldots 2^{n-1} a_{2^n} = \frac{1}{2}(a_1 + 2a_2 + 4a_4 + \ldots + 2^n a_{2^n}) = \frac{1}{2} t_n$ □

Stąd S_n i t_n są równocześnie ogran. lub nieogr., a więc odp. te szer. równ. zbieżne lub rozbieżne ■

Wniosek 8.3.3 *Szereg $\sum_{n=1}^{\infty} \frac{1}{n^p}$ jest zbieżny dla $p > 1$, a rozbieżny dla $p \leqslant 1$.*

Dowód:

Dowód: $\sum_{n=0}^{\infty} 2^n \cdot \frac{1}{(2^n)^p} = \sum_{n=0}^{\infty} \left(\frac{1}{2^{p-1}}\right)^n$

zbieżny gdy $\frac{1}{2^{p-1}} < 1$ tj. gdy $p - 1 > 0$

rozbieżny gdy $\frac{1}{2^{p-1}} \geqslant 1$ tj. gdy $p - 1 \leqslant 0$ □

8.3.2 Kryterium Leibniza

Twierdzenie 8.3.4 *(Leibniza o szeregu naprzemiennym) Niech $a : \mathbb{N}_0 \to \mathbb{R}^+$ będzie ciągiem słabo malejącym i zbieżnym do zera. Wtedy szereg $\sum_{n=0}^{\infty}(-1)^n a_n$ jest zbieżny oraz dla dowolnego $m \in \mathbb{N}_1$*

$$\sum_{n=0}^{2m-1}(-1)^n a_n \leqslant \sum_{n=0}^{\infty}(-1)^n a_n \leqslant \sum_{n=0}^{2m}(-1)^n a_n.$$

Dowód:

Dowód: $S_{2n-1} = (a_0 - a_1) + \ldots + (a_{2n-2} - a_{2n-1}) \geqslant 0$

$S_{2n} = a_0 - (a_1 - a_2) - (a_3 - a_4) - \ldots - (a_{2n-1} - a_{2n}) \leqslant a_0$

S_{2n-1} - rosn., S_{2n} - malej.

$0 \leqslant S_{2n-1} \leqslant S_{2n} \leqslant a_0$, więc obie ograniczone, stąd obie zbieżne

! $\lim_{n\to\infty} S_{2n-1} = s^*$, $\lim_{n\to\infty} S_{2n} = s^{**}$, $0 = \lim_{n\to\infty} a_n = \lim_{n\to\infty}(S_{2n} - S_{2n-1}) = s^* - s^{**}$, $s^* = s^{**}$ ■

Przykład. $\sum_{n=1}^{\infty}(-1)^n \frac{1}{n}$ jest zbieżny (i równocześnie nie jest zbieżny bezwzględnie)

8.3.3 Kryterium Cauchy'ego i kryterium d'Alemberta

Badaniem zbieżności szeregów zajmowali się w szczególności matematycy francuscy: najpierw Jean le Rond d'Alembert (rys. 8.1), a następnie Augustin Louis Cauchy (rys. 1.5). Podali oni następujące kryteria zbieżności szeregów.

Twierdzenie 8.3.5 *(Kryterium Cauchy'ego i d'Alemberta) Niech $a : \mathbb{N} \to \mathbb{C}$ będzie ciągiem oraz*

$$\alpha := \limsup_{n\to\infty} \sqrt[n]{|a_n|}, \quad \beta := \limsup_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}.$$

Wtedy:

- *jeśli $\alpha < 1$ to $\sum_{n=0}^{\infty} a_n$ jest zbieżny,*
- *jeśli $\beta < 1$ to $\sum_{n=0}^{\infty} a_n$ jest zbieżny,*
- *jeśli $\alpha > 1$ to $\sum_{n=0}^{\infty} a_n$ jest rozbieżny,*
- *jeśli istnieje $N \in \mathbb{N}$ takie, że $\left|\frac{a_{n+1}}{a_n}\right| > 1$ dla $n \geqslant N$, to $\sum_{n=0}^{\infty} a_n$ jest rozbieżny.*

Dowód:



Rysunek 8.1: Jean le Rond d'Alembert, 1717-1783, Źródło: Wikipedia

Dowód. ad 1. $\alpha<1$, $\exists\gamma$: $\alpha<\gamma<1$, $\exists N\in\mathbb{N}$ $\forall n\geq N$ $\sqrt[n]{|a_n|}<\gamma$, $|a_n|<\gamma^n$, $\sum\gamma^n$ zb.: stosujemy kryt. porównawcze ☐

ad 2. Zn. podciąg a_{n_k}: $\lim\sqrt[n]{|a_{n_k}|}=\alpha>1$, skąd: $|a_{n_k}|>1$; nie jest spełn. w kon.

ad 3. $\exists\gamma$: $\alpha<\gamma<1$ $\exists N\in\mathbb{N}$ $\forall n\geq N$ $\left|\frac{a_{n+1}}{a_n}\right|<\gamma$, $|a_{n+1}|<\gamma|a_n|\leq\gamma^2|a_{n-1}|<\ldots$

$\ldots<\gamma^{n-N+1}|a_N|$, $|a_n|<\gamma^n\frac{|a_N|}{\gamma^N}$; stos. kryt. porów.

ad 4. $n\geq N\Rightarrow|a_n|\geq|a_N|$; nie jest spełn. w kon. zbieżn. ☐■

Przykład $\sum_{n=0}^{\infty}\frac{z^n}{n!}$ zbieżny dla dowolnego $z\in\mathbb{C}$

8.3.4 Szeregi potęgowe

Definicja 8.3.6 *Szeregiem potęgowym* nazywamy szereg postaci $\sum_{n=0}^{\infty}c_nz^n$, gdzie $z\in\mathbb{C}$ jest ustalonym parametrem, a ciąg $c:\mathbb{N}_0\to\mathbb{C}$ jest tzw. *ciągiem współczynników szeregu potęgowego*.

Twierdzenie 8.3.7 *Niech* $\sum_{n=0}^{\infty} c_n z^n$ *będzie szeregiem potęgowym,* $\alpha := \limsup_{n\to\infty} \sqrt[n]{|c_n|}$, *a*

$$R := \begin{cases} \frac{1}{\alpha} & gdy\ 0 < \alpha < \infty, \\ 0 & gdy\ \alpha = \infty, \\ \infty & gdy\ \alpha = 0. \end{cases}$$

Wtedy szereg $\sum_{n=0}^{\infty} c_n z^n$ *jest zbieżny dla* $|z| < R$, *a rozbieżny dla* $|z| > R$. *Kryterium nie rozstrzyga zbieżności gdy* $|z| = R$.

Dowód:
Jest to wniosek z kryterium Cauchy'ego dla szeregów.

□

Przykłady.

$\sum n^n z^n$ $\quad R = 0$

$\sum \frac{z^n}{n!}$ $\quad R = \infty$

$\sum z^n$ $\quad R = 1$ $\quad |z| = 1$ rozb.

$\sum \frac{z^n}{n}$ $\quad R = 1$

$\sum \frac{z^n}{n^2}$ $\quad R = 1$ $\quad |z| = 1$ zbieżny

8.4 Operacje na szeregach

8.4.1 Suma i iloczyn szeregów

Twierdzenie $a: \mathbb{N} \to \mathbb{C}$, $b: \mathbb{N} \to \mathbb{C}$, $c \in \mathbb{C}$

$\sum a_n$ zb, $\sum b_n$ zb $\Rightarrow$ $\sum (a_n + b_n)$ zb : $\sum (a_n + b_n) = \sum a_n + \sum b_n$

$\sum a_n$ zb $\Rightarrow$ $\sum c a_n$ zb : $\sum c a_n = c \sum a_n$.

Dowód:
Stosujemy do sum częściowych tw. granicy sumy i iloczynu ciągów.

□

Definicja. (Iloczyn Cauchy'ego szeregów)

Iloczynem Cauchy'ego szeregów $\sum_0^\infty a_n$, $\sum_0^\infty b_n$ ($\in \mathbb{C}$) nazywamy szereg o wyrazie ogólnym

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

Twierdzenie. Jeśli $\sum a_n$, $\sum b_n$ zbieżne przy czym choć jeden bezwzględnie oraz $c_n = \sum_{k=0}^{n} a_k b_{n-k}$ to $\sum_{n=0}^{\infty} c_n$ zbieżny oraz

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right)$$

Dowód:

Dowód: ! $A_n = \sum_{i=0}^{n} a_i$, $B_n = \sum_{i=0}^{n} b_i$, $C_n = \sum_{i=0}^{n} c_i$, $\alpha := \sum_{i=0}^{\infty} |a_i| < \infty$ (zb. bezwzgl.)

! $A = \lim A_n$, $B = \lim B_n$, $\beta_n = B_n - B$

$C_n = a_0 b_0 + (a_0 b_1 + a_1 b_0) + (a_0 b_2 + a_1 b_1 + a_2 b_0) + \ldots + (a_n b_0 + a_{n-1} b_1 + \ldots + a_0 b_n) =$

$= a_0 B_n + a_1 B_{n-1} + \ldots a_n B_0 = a_0(B + \beta_n) + a_1(B + \beta_{n-1}) + \ldots + a_n(B + \beta_0) =$

$= A_n B + a_0 \beta_n + a_1 \beta_{n-1} + \ldots + a_n \beta_0$

! $\gamma_n := a_0 \beta_n + a_1 \beta_{n-1} + a_2 \beta_{n-2} + \ldots + a_n \beta_0$. Wyst. pokazać, że $\gamma_n \to 0$

? $\varepsilon > 0$, $\beta_n \to 0$ więc $\exists N \in \mathbb{N}$: $n \geq N \Rightarrow |\beta_n| < \varepsilon$

$|\gamma_n| \leq |a_n \beta_0 + a_{n-1} \beta_1 + \ldots + a_{n-N} \beta_N| + |a_{n-N-1} \beta_{N+1} + \ldots + a_0 \beta_n| \leq$

$\leq |a_n||\beta_0| + |a_{n-1}||\beta_1| + \ldots + |a_{n-N}||\beta_N| + \alpha\varepsilon$. Po obu stronach przechodzimy do lim sup

Stąd $\limsup |\gamma_n| \leq \alpha\varepsilon$, bo $a_k \to 0$, czyli $\limsup |\gamma_n| = 0$, bo ε dowolne ■

8.4.2 (**)Permutacja szeregu i zbieżność bezwarunkowa

! $a: \mathbb{N} \to \mathbb{C}$, $k: \mathbb{N} \to \mathbb{N}$ bijekcja

Def. Szereg $\sum_{n=0}^{\infty} (a \circ k)(n)$ nazywam permutacją szeregu $\sum_{n=0}^{\infty} a_n$

Def. Szereg $\sum_{n=0}^{\infty} a_n$ zbieżny bezwarunkowo $:\Leftrightarrow$ dowolna permutacja jest zbieżna

Przykład. $S = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} < 1 - \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$

Permutacja: $1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} - \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \ldots$

$\frac{1}{4k-3} + \frac{1}{4k-1} - \frac{1}{2k} > 0$ zatem $s_3' < s_6' < s_9' \ldots$; $\limsup s_n' > s_3' = \frac{5}{6} = S$

8.4.3 (**)Twierdzenie Riemanna

Twierdzenie (Riemanna).
$\forall\ a: \mathbb{N} \to \mathbb{R}$: $\sum a_n$ - zbieżny ale nie jest bezwzgl. zbieżny
$\exists\ \alpha, \beta \in \mathbb{R}$: $\alpha < \beta \Rightarrow \exists\ k: \mathbb{N} \to \mathbb{N}$ bij, że
szereg spermutowany $\sum_{n=0}^{\infty} (a \circ k)(n)$ o sumie częśc. S_n:
$\liminf S_n = \alpha$ $\limsup S_n = \beta$.

Twierdzenie. Jeśli $\sum a_n$ jest szeregiem o wyrazach zespolonych, który jest zbieżny bezwzględnie, to dowolna permutacja tego szeregu jest zbieżna i to zawsze do tej samej sumy.

8.4.4 Szeregi w programie Mathematica

Program Mathematica dość dobrze radzi sobie z liczeniem sum szeregów, a w przypadku braku zbieżności sygnalizuje ten fakt. Do policzenia sumy szeregu $\sum_{i=p}^{\infty} a_n$ używamy instrukcji

```
Sum[a[n],{n,p,Infinity}]
```

Na przykład

```
Sum[1/n!, {n, 2, Infinity}]
```

zwraca $\sum_{n=2}^{\infty} \frac{1}{n!} = -2 + e$. Oznacza to, że Mathematica, w miarę swoich umiejętności, stara się zwrócić wynik dokładny, używając symbolicznych oznaczeń na stałe matematyczne.

Rozdział 9

Funkcje wykładnicza, logarytmiczna i trygonometryczne

9.1 Funkcja wykładnicza i logarytmiczna

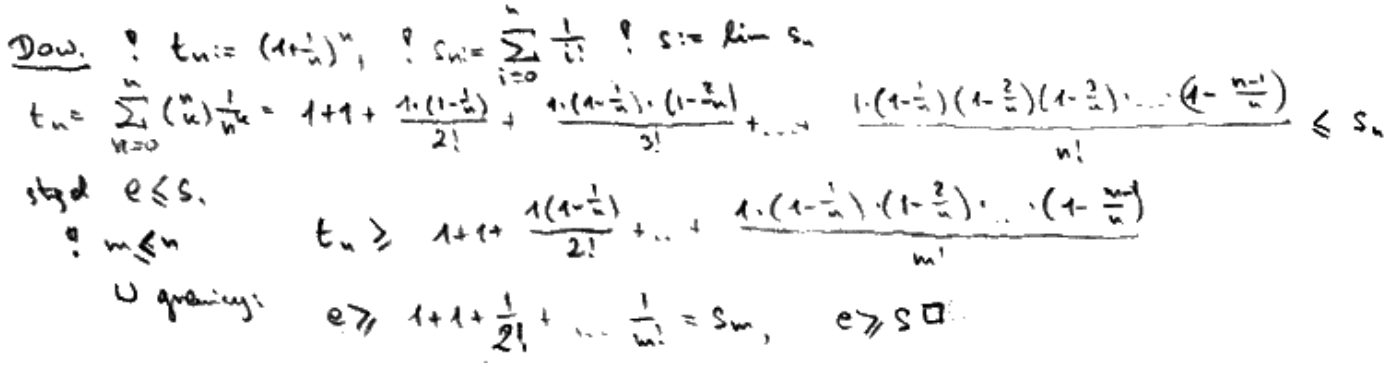
9.1.1 Zespolona funkcja wykładnicza

Następujące twierdzenie pozwala przybliżać liczbę e za pomocą szeregów i to znacznie szybciej niż w przypadku ciągu, którym posłużyliśmy się przy jej definiowaniu

Twierdzenie 9.1.1

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

Dowód:



Łatwo sprawdzić za pomocą kryterium d'Alemberta, że dla dowolnego $z \in \mathbb{C}$ szereg

$$\sum_{n=0}^{\infty} \frac{z^n}{n!}$$

jest zbieżny. Rozważać zatem możemy funkcję

$$E : \mathbb{C} \ni z \mapsto E(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!} \in \mathbb{C}.$$

Twierdzenie 9.1.2 *Funkcja E ma następujące własności:*

(i) $E(0) = 1$ *oraz* $E(1) = e$,

(ii) dla dowolnych $z_1, z_2 \in \mathbb{C}$

$$E(z_1 + z_2) = E(z_1)E(z_2),$$

(iii) dla dowolnego $z \in \mathbb{C}$ *mamy* $E(z) \neq 0$ *oraz*

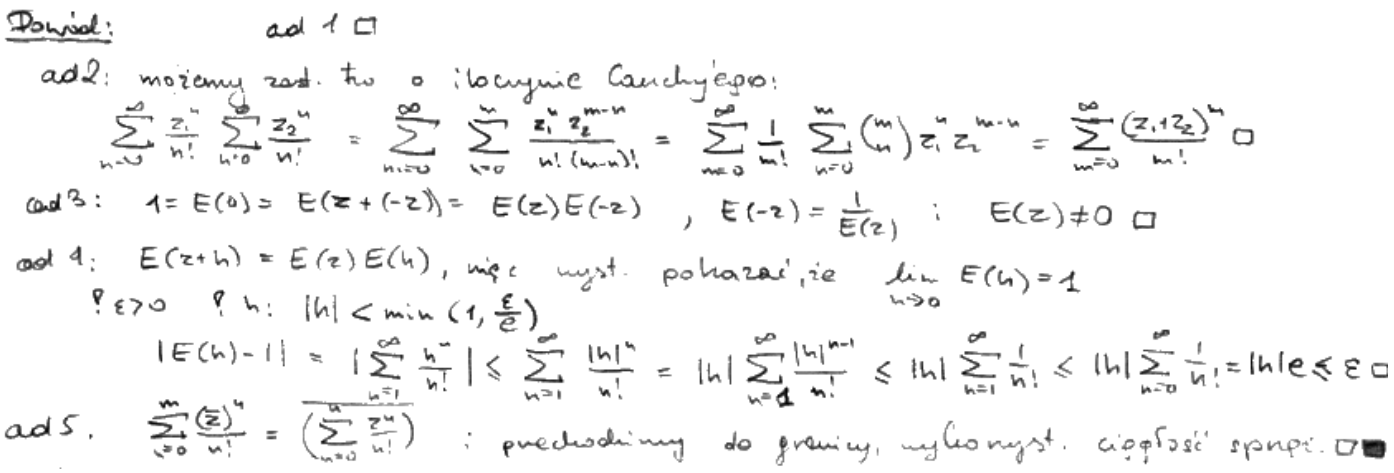
$$E(-z) = \frac{1}{E(z)},$$

(iv) E jest funkcją ciągłą,

(v) dla dowolnego $z \in \mathbb{C}$

$$E(\bar{z}) = \overline{E(z)}.$$

Dowód:



Definicja 9.1.3 Funkcję E nazywamy *zespoloną funkcją wykładniczą*. Stosujemy dla niej również tradycyjną notację

$$e^z := E(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!} \in \mathbb{C}.$$

9.1.2 Funkcja wykładnicza e^x

Twierdzenie 9.1.4 *Istnieje dokładnie jedna funkcja ciągła $f : \mathbb{R} \to \mathbb{R}^+$ taka, że $f(1) = e$ oraz dla dowolnych $x, y \in \mathbb{R}$ zachodzi $f(x+y) = f(x)f(y)$. Jest ona zwana funkcją wykładniczą o podstawie e. Powszechnie stosowaną notacją dla $f(x)$ jest e^x, a sama funkcja f jest oznaczana*

$$\exp : \mathbb{R} \ni x \mapsto e^x \in \mathbb{R}^+.$$

Funkcja ta ma również następujące własności:

(i) dla dowolnych $p \in \mathbb{Z}$ oraz $q \in \mathbb{N}_1$

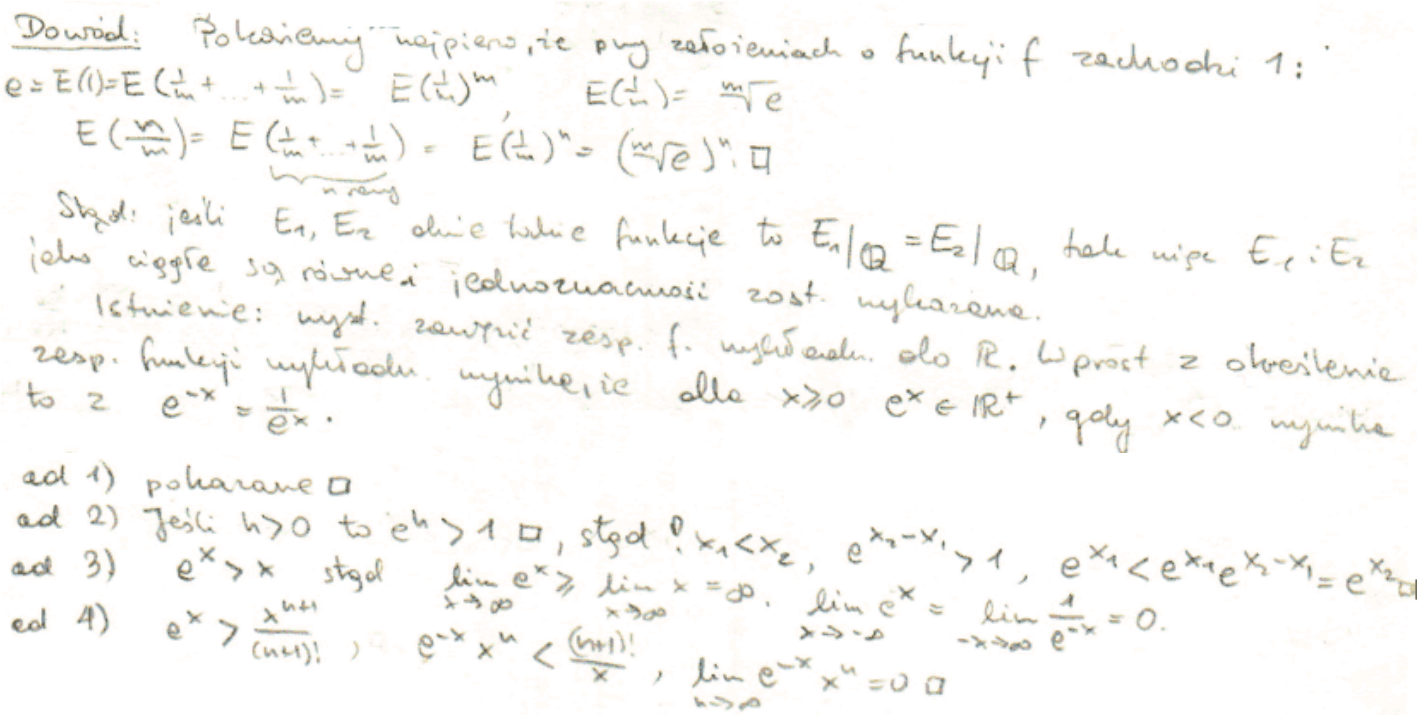
$$e^{\frac{p}{q}} = \sqrt[q]{e^p},$$

(ii) jest silnie rosnąca,

(iii) $\lim_{x\to\infty} e^x = \infty$, $\lim_{x\to-\infty} e^x = 0$,

(iv) dla dowolnego $n \in \mathbb{N}$ mamy $\lim_{x\to\infty} x^n e^{-x} = 0$.

Dowód:



9.1.3 Logarytm naturalny

Funkcja exp, jako bijekcja silnie rosnąca z $\mathbb{R}$ do $\mathbb{R}^+$, ma funkcję odwrotną, która jest bijekcją silnie rosnącą z $\mathbb{R}^+$ do $\mathbb{R}$.

Definicja 9.1.5 Funkcję tę nazywamy *logarytmem naturalnym* i oznaczamy

$$\ln : \mathbb{R}^+ \ni x \mapsto \ln x \in \mathbb{R}.$$

Z definicji tej wynika w szczególności, że dla dowlnych $x \in \mathbb{R}^+$ i $y \in \mathbb{R}$ równość $y = \ln x$ jest równoważna równości $x = e^y$.

Twierdzenie 9.1.6 *Funcja* ln *ma następujące własności:*

(i) jest silnie rosnąca,

(ii) $\ln 1 = 0$, $\ln e = 1$,

(iii) dla dowolnych $x, y \in \mathbb{R}^+$, $n \in \mathbb{N}$

$$\ln(xy) = \ln x + \ln y, \quad \ln \frac{x}{y} = \ln x - \ln y, \quad \ln x^n = n \ln x, \quad \ln \sqrt[n]{x} = \frac{\ln x}{n},$$

(iv) $\lim_{x\to\infty} \ln x = \infty$, $\lim_{x\to 0} \ln x = -\infty$,

(v) dla dowolnych $p \in \mathbb{N}$, $q \in \mathbb{Z}^+$, $x \in \mathbb{R}^+$

$$e^{\frac{p}{q} \ln x} = \sqrt[q]{x^p} = x^{\frac{p}{q}}.$$

Dowód:

Dowód: ad 1) trzeba pokazać jedynie ciągłość, ale pokonaliśmy swego czasu, że bijekcja rosnąca przedziału otw. w przedział otw. jest ciągłe □
ad 2) □
ad 3) ? $u = \ln x$, $v = \ln y$, $x = e^u$, $y = e^v$, $xy = e^{u+v}$, $u+v = \ln(xy)$ □
dla $\ln(\frac{x}{y})$ analog; $\ln x^n = \ln x + \ldots + \ln x = n \ln x$, $n \ln \sqrt[n]{x} = \ln(\sqrt[n]{x})^n = \ln x$ □
ad 4.) wniosek z tw o granicy funkcji monoton. □
ad 5) $e^{p/q \ln x} = e^{\ln \sqrt[q]{x^p}} = \sqrt[q]{x^p}$ □■

W programie Mathematica logarytm naturalny oznaczany jest `Log`.

9.1.4 Funkcja wykładnicza a^x

Przy pomocy logarytmu naturalnego możemy zdefiniować funkcję wykładniczą o dowolnej podstawie $a \in \mathbb{R}^+$.

Definicja 9.1.7 Funkcję wykładniczą o podstawie $a \in \mathbb{R}^+$ definiujemy wzorem

$$\mathbb{R} \ni x \mapsto a^x := e^{x \ln a} \in \mathbb{R}^+.$$

9.2 Funkcje trygonometryczne

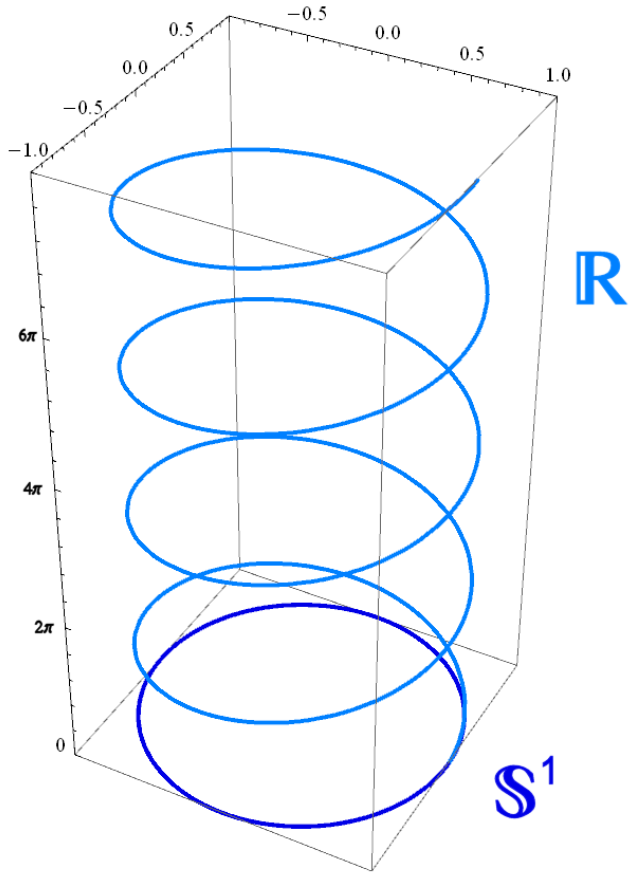
9.2.1 Nawijanie prostej na okrąg

Zespolona funkcja wykładnicza pozwoli nam też podać formalną i czysto analityczną definicję funkcji trygonometrycznych. Zacznijmy od następującej uwagi:

Uwaga 9.2.1 *Funkcja*

$$\mathbb{R} \ni x \mapsto e^{ix} \in S^1 \tag{9.1}$$

jest dobrze określona i ciągła.



Rysunek 9.1: Wykres funkcji $\mathbb{R} \ni t \mapsto e^{it} \in \mathbb{S}^1$.

Dowód:

Dowód: $e^{ix} \in S^1$ □, ciągłość: złożenie funkcji wykładn

Funkcję daną wzorem (9.1) określa się często mianem nawijania prostej na okrąg. Terminologię tę wyjaśnia rys. 9.1

9.2.2 Funkcje sin i cos

Możemy teraz podać czysto analityczną definicję funkcji sin oraz cos. Analizując własności tych funkcji stopniowo przekonamy się, że pokrywają się one z funkcjami sin i cos zdefiniowanymi geometrycznie.

Definicja 9.2.2 Funkcje cos i sin definiujemy wzorami

$$\begin{aligned} \cos &: \mathbb{R} \ni x \mapsto \operatorname{re} e^{ix} = \frac{e^{ix} + e^{-ix}}{2} \in [-1, 1] \\ \sin &: \mathbb{R} \ni x \mapsto \operatorname{im} e^{ix} = \frac{e^{ix} - e^{-ix}}{2i} \in [-1, 1]. \end{aligned}$$

```
(* Obliczanie funkcji sinus za pomocą rozwinięcia w szereg *)
SinSeries[x_, d_] :=
  Module[{s, znak, silnia, y, n},
  s = 0; znak = 1; silnia = 1; y = x; n = 0;
  Do[
    s = s + znak*y/silnia;
    znak = -znak;
    silnia = silnia*(2*n + 2)*(2*n + 3);
    y = y*x*x;
    n = n + 1,
    {d}
  ];
  s
]
```

Listing 9.1: Wyznaczenie wartości funkcji sin w programie Mathematica w oparciu o rozwinięcie w szereg

Uwaga 9.2.3 *Funkcje* sin *i* cos *są ciągłe. Ponadto dla każdego* $x \in \mathbb{R}$ *zachodzą wzory:*

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \qquad \sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$$

$$e^{ix} = \cos x + i \sin x$$

$$\cos^2 x + \sin^2 x = 1.$$

Dowód:

Dowód:
1) - wsp. z ciągłości $x \to e^{ix}$
2) $\cos x = \frac{e^{ix} + e^{-ix}}{2} = \frac{1}{2}\left(\sum_{n=0}^{\infty} \frac{(ix)^n}{n!} + \sum_{n=0}^{\infty} \frac{(-ix)^n}{n!}\right) = \sum_{n=0}^{\infty} \frac{(ix)^n + (-ix)^n}{(2n)!} = \sum_{k=0}^{\infty} i^{2k} \frac{x^{2k}}{(2k)!}$ □
$\sin x$ - podobnie
3) □
4) $\cos^2 x + \sin^2 x = |e^{ix}|^2 = 1$ □■

Program, który przybliża wartości funkcji sin w oparciu o szereg w uwadze 9.2.3 przedstawiono na listingach 9.1 (Mathematica) i 9.2 (C++).

```
/********************* SinSeries.cpp ************************/
/* Obliczanie funkcji sinus za pomocą rozwinięcia w szereg */
/************************************************************/
#include <iostream>
#include <cmath>
#include <iomanip>
using namespace std;

int main(){

  cout << fixed << right;
  cout << setprecision(15);

  double pi=3.141592653589793;
  double x=pi/6;
  double y=x;

  int nmax=10,znak=1,silnia=1;
  double s=0;
  for(int n=0;n<nmax;n=n+1){
    s=s+znak*y/silnia;
    cout << "   s[" << setw(2) << n << "] =" << setw(18) << s;
    cout << endl;
    silnia=silnia*(2*n+2)*(2*n+3);
    y=y*x*x;
    znak=-znak;
  }
}
```

Listing 9.2: Wyznaczenie wartości funkcji sin w C++ w oparciu o rozwinięcie w szereg

Ćwiczenie komputerowe 9.2.4 • Sprawdź zachowanie się programu z listingu 9.2 dla zmiennej x inicjalizowanej wartościami $\frac{\pi}{4}$, $\frac{\pi}{3}$, $\frac{\pi}{2}$, π, 2π. Co obserwujesz w zachowaniu się ciągu przybliżeń? Jakie rodzą się pytania?

• Adaptuj program z listingów 9.1 i 9.2 do liczenia wartości funkcji cos wykorzystując fakt, że

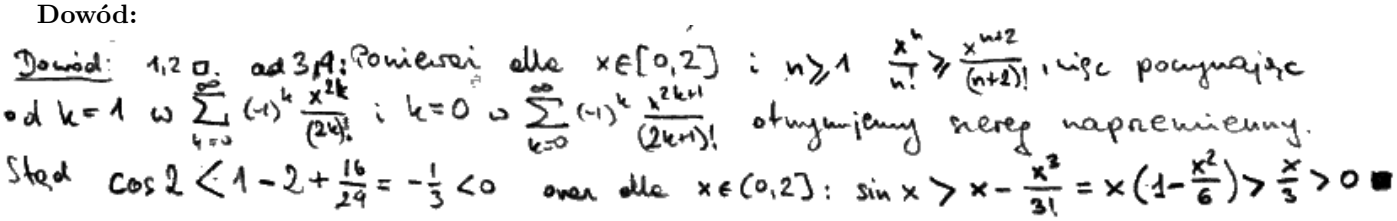
$$\cos x = 1 - \frac{x^2}{1\cdot 2} + \frac{x^4}{1\cdot 2\cdot 3\cdot 4} - \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.$$

Oceń poprawność i przydatność implementacji wykorzystując argument, w którym wartość funkcji cos jest Ci znana.

Lemat 9.2.5 *Zachodzą następujące wzory*

$$\cos 0 = 1, \quad \sin 0 = 0,$$
$$\cos 2 < 0,$$
$$x \in (0,2] \;\Rightarrow\; \sin x > 0.$$

Dowód:



9.2.3 Liczba π

Dysponujemy teraz wystarczającą maszynerią, by podać analityczną, a nie geometryczną definicję liczby π.

Wniosek 9.2.6 *Zbiór dodatnich miejsc zerowych funkcji* cos *jest domknięty i niepusty. W szczególności jego kres dolny jest dodatni i też jest miejscem zerowym funkcji* cos.

Definicja 9.2.7 Liczbę π definiujemy jako podwojone najmniejsze dodatnie miejsce zerowe funkcji cos. Innymi słowy

$$\pi := 2 \inf \{\, x \in \mathbb{R}^+ \mid \cos x = 0 \,\}.$$

9.2.4 Okresowość

Definicja 9.2.8 Niech F oznacza ciało liczb rzeczywistych lub ciało liczb zespolonych. Mówimy, że funkcja $f : F \multimap F$ jest okresowa wtedy i tylko wtedy gdy istnieje takie $t \in F$, że dla każdego $x \in \operatorname{dom} f$ mamy $x + t \in \operatorname{dom} f$ oraz $f(x + t) = f(x)$. Liczbę t nazywamy wtedy okresem funkcji f. Jeśli istnieje najmniejszy okres dodatni funkcji f nazywamy go okresem podstawowym funkcji f

Twierdzenie 9.2.9 *Mamy następujące własności:*

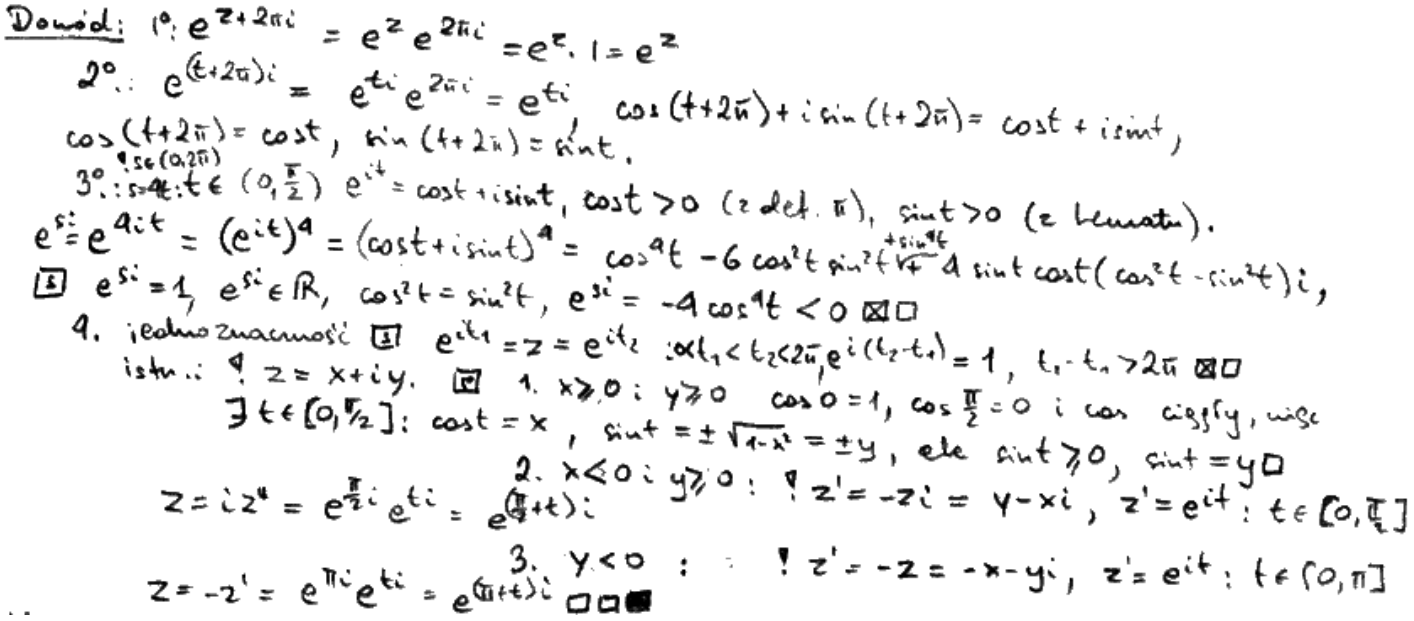
(i) funkcja $\exp : \mathbb{C} \to \mathbb{C}$ *jest okresowa o okresie* $2\pi i$,

(ii) funkcje $\sin, \cos : \mathbb{R} \to [-1, 1]$ *są okresowe o okresie* 2π,

(iii) $t \in (0, 2\pi) \Rightarrow e^{it} \neq 1$,

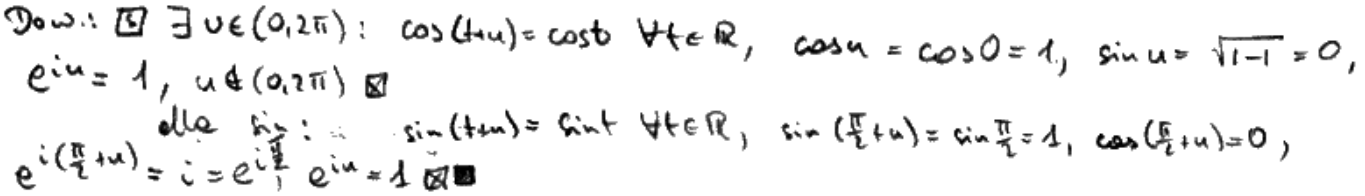
(iv) dla każdego $z \in S^1$ *istnieje dokładnie jedno* $t \in [0, 2\pi)$ *takie, że* $e^{it} = z$.

Dowód:



Wniosek 9.2.10 *Okresem podstawowym funkcji* sin *i* cos *jest* 2π.

Dowód:



9.2.5 Wzory trygonometryczne

Twierdzenie 9.2.11 *Dla dowolnych* $x, y \in \mathbb{R}$ *zachodzą wzory*

$$\begin{aligned} \sin(x+y) &= \sin x \cos y + \cos x \sin y, \\ \cos(x+y) &= \cos x \cos y - \sin x \sin y. \end{aligned}$$

Dowód:

Dowód: $\cos(x+y)+i\sin(x+y) = e^{i(x+y)} = e^{ix}e^{iy} = (\cos x + i\sin x)(\cos y + i\sin y) =$
$= (\cos x \cos y - \sin x \sin y) + i(\sin x \cos y + \cos x \sin y)$ ■

Twierdzenie 9.2.12 *(wzór de Moivre'a) Dla dowolnego $x \in \mathbb{R}$ oraz $n \in \mathbb{N}$ zachodzi wzór*

$$(\cos x + i \sin x)^n = \cos nx + i \sin nx.$$

Dowód:

Dow.: $(\cos x + i\sin x)^n = e^{inx} = \cos nx + i\sin nx$ ■

Rozdział 10

Przestrzenie wektorowe

10.1 Definicja przestrzeni wektorowej

Zaproponowana przez Kartezjusza algebraizacja geometrii poprzez wprowadzenie układu współrzędnych na płaszczyźnie oraz geometryczna interpretacja liczb zespolonych zaowocowały wprowadzeniem w XIXw. koncepcji wektorów jako zaczepionych w początku układu n-wymiarowej przestrzeni zorientowanych odcinków, które można dodawać podobnie jak dodaje się liczby rzeczywiste i zespolone. Intuicyjne wyobrażenie geometryczne wektora możliwe jest jedynie w przestrzeni dwu- i trójwymiarowej, bo jedynie z takimi przestrzeniami mają do czynienia nasze zmysły. Od strony algebraicznej wektor to po prostu układ n liczb zwanych współrzędnymi wektora. Dodawanie wektorów metodą geometryczną sprowadza się do konstrukcji przekątnej rozpiętego na wektorach równoległoboku. W przypadu płaszczyzny pokrywa się to z geometryczną interpretacją sumy liczb zespolonych przedstawioną na rysunku 4.2. Wektory wraz z operacją dodawania stanowią grupę. Poza wymiarem dwa wektory nie da się jednak zdefiniować mnożenia wektorów tak, by uzyskać strukturę ciała. Możliwe jest jednak mnożenie wektora przez liczbę, które sprowadza się do stosownej zmiany jego długości i, w przypadku mnożenia liczby ujemnej, symetrycznego odbicia względem początku układu.

Powyższe idee prowadzą do koncepcji przestrzeni wektorowej nad zadanym ciałem.

10.1.1 Przestrzeń wektorowa

Definicja 10.1.1 Niech F będzie ciałem i niech 1 oznacza element neutralny względem mnożenia w ciele F. *Przestrzenią wektorową nad F* nazywamy piątkę $(X, F, +, \cdot, 0)$ taką, że $(X, +, 0)$ jest grupą abelową, a

$$\cdot : F \times X \to X$$

jest działaniem (tzw. *działaniem zewnętrznym*) spełniającym dla dowolnych $\alpha, \beta \in F$ oraz $x, y \in X$ warunki

(W1) $(\alpha + \beta)x = \alpha x + \beta x$,

(W2) $\alpha(x + y) = \alpha x + \alpha y$,

(W3) $\alpha(\beta x) = (\alpha\beta)x$,

(W4) $1.x = x$.

Elementy ciała F określa się często mianem *skalarów*, a o działaniu zewnętrznym mówi się, że jest to *mnożenie przez skalary*. Klasę przestrzeni wektorowych nad ciałem F oznaczać będziemy $\mathrm{Vect}(F)$.

Przykład 10.1.2 Klasycznym przykładem przestrzeni wektorowej jest $\mathbb{R}^n$ nad ciałem $\mathbb{R}$. Elementami tej przestrzeni są n-tki (zwane również krotkami lub po prostu wektorami) postaci $x = (x_1, x_2, \dots x_n)$, gdzie $x_i \in \mathbb{R}$ są liczbami rzeczywistymi. Sumę $x = (x_1, x_2, \dots x_n)$ i $y = (y_1, y_2, \dots y_n)$ definiuje się wzorem

$$x + y := (x_1 + y_1, x_2 + y_2, \dots x_n + y_n),$$

a mnożenie wektora x przez skalar $\alpha \in \mathbb{R}$ wzorem

$$\alpha x := (\alpha x_1, \alpha x_2, \dots \alpha x_n).$$

Łatwo sprawdzić, że jest to rzeczywiście przestrzeń wektorowa. W podobny sposób konstruujemy przestrzeń wektorową F^n dla dowolnego ciała F, na przykład dla ciała liczb zespolonych.

Przykład 10.1.3 Niech A będzie dowolnym ustalonym zbiorem. Niech $F(A, \mathbb{R})$ oznacza zbiór wszystkich funkcji $f : A \to \mathbb{R}$. Dla $f, g \in F(A, \mathbb{R})$ definiujemy sumę $f + g$ jako funkcję

$$f + g : A \ni x \mapsto f(x) + g(x) \in \mathbb{R},$$

a mnożenie funkcji f przez skalar $\alpha \in \mathbb{R}$ jako funkcję

$$\alpha f : A \ni x \mapsto \alpha f(x) \in \mathbb{R}.$$

Łatwo sprawdzić, że jest to rzeczywiście przestrzeń wektorowa. Jeśli X jest przestrzenią topologiczną, to w podobny sposób konostruujemy przestrzeń funkcji ciągłych z X do $\mathbb{R}$ oznaczaną jako $C(A, \mathbb{R})$

Przykład z przestrzenią wektorową funkcji ograniczonych sygnalizuje, że pożytki ze stosowania tego pojęcia mogą wykraczać poza wektory na płaszczyźnie i ich uogólnienia na $\mathbb{R}^n$. W istocie, można pokazać, że przestrzeń $B(A)$ nie jest izomorficzna z przestrzenią $\mathbb{R}^n$ dla żadnego n.

10.2 Przestrzenie wektorowe z iloczynem skalarnym

Poprzez interpretację wektorów na płaszczyźnie jako liczb zespolonych wiemy, że można je nie tylko dodawać, ale i mnożyć z zachowaniem własności ciała. Uzyskanie struktury ciała dla wektorów w przestrzeni trówymiarowej i wyższych wymiarach nie jest możliwe. Natomiast bardzo pożyteczny jest iloczyn skalarny. Jest to odwzorowanie, które parze wektorów (v, w) przyporządkowuje liczbę $v \cdot w$ będącą iloczynem długości wektora w przez długość rzutu prostopadłego wektora v na prostą, w której zawiera się wektor w, wzięta ze znakiem plus, jeśli kąt między wektorami jest ostry i znakiem minus w przeciwnym razie. Łatwo sprawdzić, że tak zdefiniowany iloczyn skalarny jest przemienny i wyraża się wzorem

$$v \cdot w = |v||w| \cos \alpha$$

gdzie $|v|$ i $|w|$ oznaczają długości wektorów v i w, a α jest kątem pomiędzy nimi. Ze wzoru tego wynikają od razu najważniejsze własności iloczynu skalarnego: wektory v i w są prostopadłe wtedy i tylko wtedy gdy $v \cdot w = 0$ oraz $v \cdot v = |v|^2$.

W praktyce wygodne jest przyjęcie aksjomatycznej definicji iloczynu skalarnego.

10.2.1 Iloczyn skalarny

Definicja 10.2.1 Niech X będzie przestrzenią wektorową nad $\mathbb{R}$. Odwzorowanie $\varphi : X \times X \to \mathbb{R}$ nazywamy *iloczynem skalarnym* jeżeli dla wszystkich $x, y, z \in X$ oraz $\alpha \in \mathbb{R}$ spełnione są następujące warunki:

(IS1) $\varphi(x + y, z) = \varphi(x, z) + \varphi(y, z)$,

(IS2) $\varphi(\alpha x, y) = \alpha\varphi(x, y)$,

(IS1) $\varphi(x, y) = \varphi(y, x)$,

(IS1) $\varphi(x, x) \geqslant 0$ i $\varphi(x, x) = 0 \Leftrightarrow x = 0$.

Iloczynem skalarnym w $\mathbb{R}^d$ jest odwzorowanie

$$(x_1, x_2, \ldots, x_d), (y_1, y_2, \ldots, y_d) \mapsto \sum_{i=1}^{d} x_i y_i.$$

Wzór na iloczyn skalarny w $\mathbb{R}^d$ jest podstawą definiowania ilocznu skalarnego w innych przestrzeniach. W szczególności w pewnych przestrzeniach ciągów iloczynem skalarnym jest $\sum_{i=1}^{\infty} x_i y_i$, a w pewnych przestrzeniach funkcyjnych iloczynem skalarnym jest $\int_a^b f(x)g(x)dx$.

10.2.2 Nierówność Cauchy-Schwartza

Twierdzenie 10.2.2 *(**Nierówność Cauchy-Schwartza**) Niech X będzie przestrzenią wektorową nad $\mathbb{R}$ z iloczynem skalarnym φ. Dla dowolnych $x, y \in X$ mamy*

$$|\varphi(x, y)| \leqslant \sqrt{\varphi(x, x)}\sqrt{\varphi(y, y)}.$$

Dowód: Rozważmy funkcję

$$\mu : \mathbb{R} \ni t \mapsto \varphi(x + ty, x + ty) \in [0, \infty).$$

Wykorzystując własności iloczynu skalarnego dostajemy dla dowolnego $t \in [0, \infty)$

$$0 \leqslant \mu(t) = t^2\varphi(x, x) + 2t\varphi(x, y) + \varphi(y, y).$$

Zatem wyróżnik w tej nierówności kwadratowej musi być niedodatni, tj.

$$4\varphi(x, y)^2 - 4\varphi(x, x)\varphi(y, y) \leqslant 0$$

co dowodzi tezy. □

Wniosek 10.2.3 *Dla $(x_1, x_2, \ldots, x_d), (y_1, y_2, \ldots, y_d) \in \mathbb{R}^d$ mamy*

$$|\sum_{i=1}^{d} x_i y_i| \leqslant \sqrt{\sum_{i=1}^{d} x_i^2}\sqrt{\sum_{i=1}^{d} y_i^2}.$$

10.3 Przestrzenie wektorowe unormowane

W przestrzeniach wektorowych przydałoby się pojęcie odległości, tak byśmy mogli nadać jej strukturę przestrzeni metrycznej i rozważać zbiory otwarte i domknięte oraz granice. Jeśli mamy do dyspozycji koncepcję długości wektora, to odległość dwóch wektorów można zdefiniować jako długość ich różnicy. Dlatego przydałaby się aksjomatyczna definicja długości wektora, która gwarantowałaby, że wydobyć z niej można odległość.

10.3.1 Norma

Definicja 10.3.1 Niech X będzie przestrzenią wektorową nad $\mathbb{R}$. Odwzorowanie $||\cdot|| : X \to \mathbb{R}^*$ nazywamy *normą* w X jeżeli dla wszystkich $x, y \in X$ oraz $\alpha \in \mathbb{R}$

(N1) $||x|| = 0 \Leftrightarrow x = 0$,

(N2) $||\alpha x|| = |\alpha|\,||x||$,

(N3) $||x + y|| \leqslant ||x|| + ||y||$.

Twierdzenie 10.3.2 *Jeśli φ jest iloczynem skalarnym w przestrzeni wektorowej X, to*

$$||x||_\varphi := \sqrt{\varphi(x, x)}$$

zadaje w tej przestrzeni normę.

Dowód: Własności (N1) i (N2) są oczywiste. Aby sprawdzić własność (N3) zauważmy, że

$$||x + y||_\varphi^2 = \varphi(x + y, x + y) = \varphi(x, x) + 2\varphi(x, y) + \varphi(y, y) \leqslant \varphi(x, x) + 2\sqrt{\varphi(x, y)}\sqrt{\varphi(x, y)} + \varphi(y, y) = (||x||_\varphi + ||y||_\varphi)^2$$

□

10.3.2 Norma euklidesowa

Definicja 10.3.3 Normę w $\mathbb{R}^d$ pochodzącą od iloczynu skalarnego $\sum_{i=1}^d x_i y_i$, tj. normę daną wzorem $||x|| := \sqrt{\sum_{i=1}^d x_i^2}$ nazywamy *normą euklidesową*.

W $\mathbb{R}^d$ zadać można również normy nie pochodzące od iloczynu skalarnego, na przykład normą jest $||x||_1 : \sum_{i=1}^d |x_i|$.

Definicja 10.3.4 Przestrzeń wektorową nad ciałem F z zadaną normą nazywamy *przestrzenią wektorową unormowaną*. Klasę takich przestrzeni oznaczać będziemy $\mathrm{NVect}(F)$.

10.3.3 Przestrzenie Hilberta i Banacha

Twierdzenie 10.3.5 *Niech $||.||$ będzie normą w przestrzeni wektorowej X. Wtedy wzór*

$$\rho(x, y) := ||x - y||$$

określa w X metrykę.

□



Rysunek 10.1: Stefan Banach, 1892 - 1945

Definicja 10.3.6 Przestrzeń wektorową z iloczynem skalarnym, zupełną nazywamy *przestrzenią Hilberta*. Przestrzeń wektorową unormowaną, zupełną nazywamy *przestrzenią Banacha*.

Pojęcie przestrzeni Hilberta i przestrzeni Banacha pozwala metody opracowane dla przestrzeni $\mathbb{R}^n$ stosować w przestrzeniach funkcyjnych. Są to różnego rodzaju przestrzenie, w których rolę wektorów pełnią funkcje.

Rozdział 11

Pochodne i różniczki.

11.1 Pochodna funkcji

11.1.1 Definicja pochodnej

W rozdziale wstępnym wprowadziliśmy intuicyjną koncepcję pochodnej funkcji $f : \mathbb{R} \to \mathbb{R}$ w punkcie x_0 jako współczynnika kierunkowego nachylenia stycznej do wykresu poprowadzonej w punkcie $(x_0, f(x_0))$. Współczynnik kierunkowy prostej w układzie współrzędnych x i y to tangens kąta jaki tworzy ta prosta z osią x. Łatwo jest policzyć ten współczynnik dla siecznej poprowadzonej przez punkty $(x_0, f(x_0))$ i $(x_1, f(x_1))$. Z rysunku 11.1 wnosimy, że tangens ten to iloraz

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0},$$

zwany ilorazem różnicowym. *Ustalając x_0 i przybliżając x_1 do x_0 przybliżamy sieczną do stycznej w $(x_0, f(x_0))$. Można zatem oczekiwać, że współczynnik kierunkowy siecznej zmierza do współczynnika kierunkowego stycznej. Niech $h := x_1 - x_0$ będzie* przyrostem zmiennej x *względem punktu x_0. Iloraz różnicowy można wtedy zapisać jako*

$$\frac{f(x_0 + h) - f(x_0)}{h},$$

a współczynnik kierunkowy stycznej w $(x_0, f(x_0))$ jako

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \tag{11.1}$$

co prowadzi do definicji pochodnej funkcji f w punkcie x_0.

Choć powyższa interpretacja dotyczy funkcji o wartościach w zbiorze liczb rzeczywistych, granica (11.1) ma sens również wtedy, gdy funkcja f ma wartości w przestrzeni wektorowej $\mathbb{R}^m$, zatem nasza formalna definicja będzie nieco ogólniejsza.

Niech $f : (a, b) \to \mathbb{R}^m$ będzie funkcją określoną na przedziale otwartym $(a, b) \subset \mathbb{R}$ o wartościach w $\mathbb{R}^m$. Ustalmy punkt $x_0 \in (a, b)$

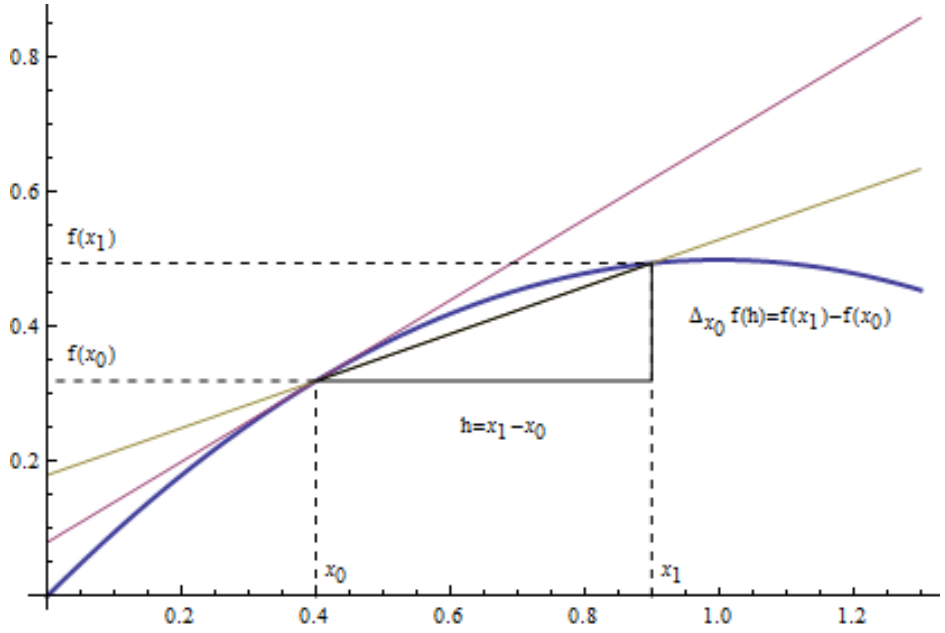
Definicja 11.1.1 Mówimy, że funkcja f jest *różniczkowalna w punkcie x_0* jeżeli w $\mathbb{R}^m$ istnieje granica

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Granicę tę nazywamy *pochodną* funkcji f w punkcie x_0 i oznaczamy $f'(x_0)$. Jeśli f ma pochodną w każdym punkcie $x \in (a, b)$ to mówimy, że f jest różniczkowalna na przedziale (a, b) i definiujemy również funkcję

$$f' : (a, b) \ni x \mapsto f'(x) \in \mathbb{R}^m,$$

którą nazywamy *funkcją pochodną* funkcji f.



Rysunek 11.1: Styczna i sieczna do krzywej.

Pochodzenie określenia "różniczkowalna" wyjaśnimy w rozdziale 11.2.4.

11.1.2 Pochodna funkcji stałej i identycznościowej.

Twierdzenie 11.1.2 (o pochodnej funkcji stałej) *Funkcja stała*

$$c_{\mathbb{R}} : (\infty, \infty) \ni x \mapsto c \in \mathbb{R}.$$

ma pochodną w każdym punkcie oraz

$$(c_{\mathbb{R}})' = 0_{\mathbb{R}}$$

co czasami zapisujemy mniej formalnie, ale krócej jako

$$c' = 0.$$

Dowód: Rzeczywiście, dla funkcji stałej iloraz różnicowy wynosi

$$\frac{c_{\mathbb{R}}(x+h) - c_{\mathbb{R}}(x)}{h} = \frac{c-c}{h} = 0,$$

więc granica ilorazu różnicowego jako granica funkcji stale równej zero wynosi zero. Dowodzi to naszej tezy. □

Twierdzenie 11.1.3 (o pochodnej identyczności) *Funkcja identycznościowa*

$$\mathrm{id}_{\mathbb{R}} : \mathbb{R} \ni x \mapsto x \in \mathbb{R}.$$

ma pochodną w każdym punkcie oraz

$$(\mathrm{id}_{\mathbb{R}})' = 1_{\mathbb{R}}$$

co czasami zapisujemy mniej formalnie, ale krócej jako

$$x' = 1.$$

Dowód: Dla funkcji identycznościowej iloraz różnicowy wynosi

$$\frac{\mathrm{id}_{\mathbb{R}}(x+h) - \mathrm{id}_{\mathbb{R}}(x)}{h} = \frac{(x+h)-x}{h} = \frac{h}{h} = 1,$$

więc granica ilorazu różnicowego jako granica funkcji stale równej jeden wynosi jeden. Dowodzi to naszej tezy. □

Przykład 11.1.4 Rozważmy funkcję

$$f : \mathbb{R} \ni x \mapsto |x| \in \mathbb{R}.$$

Jeśli x jest dodatnie, to dla dostatecznie małych h również $x+h$ jest dodatnie, zatem wtedy $|x| = x$, $|x+h| = x+h$ czyli w tym przypadku iloraz różnicowy wynosi

$$\frac{f(x+h)-f(x)}{h} = \frac{|x+h|-|x|}{h} = \frac{(x+h)-x}{h} = \frac{h}{h} = 1.$$

Jeśli x jest ujemne, to dla dostatecznie małych h również $x+h$ jest ujemne, zatem wtedy $|x| = -x$, $|x+h| = -x-h$ czyli w tym przypadku iloraz różnicowy wynosi

$$\frac{f(x+h)-f(x)}{h} = \frac{|x+h|-|x|}{h} = \frac{(-x-h)-(-x)}{h} = \frac{-h}{h} = -1.$$

Dla $x = 0$ iloraz różnicowy może wynosić tak jeden jak i minus jeden, w zależności od znaku h, więc nie ma granicy. Podsumowując,

$$f'(x) = \begin{cases} -1 & \text{gdy } x < 0, \\ \text{nie istnieje} & \text{gdy } x = 0, \\ 1 & \text{gdy } x > 0. \end{cases}$$

11.1.3 Pochodna lewo- i prawostronna.

Ostatni przykład podpowiada, że, podobnie jak w przypadku granicy, warto też rozważać pochodne lewo- i prawostronne. Niech $f : [a,b] \to \mathbb{R}^m$ będzie funkcją określoną na przedziale domkniętym $[a,b]$.

Definicja 11.1.5 Niech $x_0 \in (a,b]$. Mówimy, że f ma pochodną lewostronną w x_0 jeśli w $\mathbb{R}^m$ istnieje granica lewostronna

$$\lim_{h\to 0^-} \frac{f(x_0+h)-f(x_0)}{h}. \tag{11.2}$$

Granicę tę nazywamy pochodną lewostronną funkcji f w x_0 i oznaczamy $f'_-(x_0)$. Podobnie, jeśli $x_0 \in [a,b)$ to mówimy, że f ma pochodną prawostronną w x_0 jeśli w $\mathbb{R}^m$ istnieje granica prawostronna

$$\lim_{h\to 0^+} \frac{f(x_0+h)-f(x_0)}{h}. \tag{11.3}$$

Granicę tę nazywamy pochodną prawostronną funkcji f w x_0 i oznaczamy $f'_+(x_0)$.

Przykład 11.1.6 Funkcja f w przykładzie 11.1.4 ma w zerze pochodną lewostronną równą -1 oraz pochodną prawostronną równą 1.

Ćwiczenie 11.1.7 Pokaż, że funkcja $f:(a,b)\to\mathbb{R}^m$ ma w punkcie $x_0\in(a,b)$ pochodną wtedy i tylko wtedy gdy ma w tym punkcie tak pochodną lewostronną jak i prawostronną i są one sobie równe.

W dalszym ciągu wygodnie będzie również rozważać różniczkowalność funkcji określonej na przedziale domkniętym.

Definicja 11.1.8 Mówimy, że funkcja $f:[a,b]\to\mathbb{R}^m$ jest *różniczkowalna* jeśli ma pochodną w każdym punkcie przedziału (a,b), pochodną prawostronną w a i pochodną lewostronną w b.

W przypadku funkcji różniczkowalnej $f:[a,b]\to\mathbb{R}^m$ wygodnie jest mówić krótko o pochodnej f w a i w b rozumiejąc przez to odpowiednio pochodną prawostronną w a i lewostronną w b.

11.1.4 Ciągłość funkcji różniczkowalnej

W dalszych rozdziałach podamy szereg twierdzeń gwarantujących istnienie pochodnej, co pozwoli nam konstruować znacznie bardziej zaawansowane przykłady. Na koniec tego rozdziału podajmy następujące twierdzenie określające warunek konieczny istnienia pochodnej.

Twierdzenie 11.1.9 *Jeśli funkcja $f:(a,b)\to\mathbb{R}^m$ ma w punkcie $x_0\in(a,b)$ pochodną to jest w tym punkcie ciągła.*

Dowód: Aby pokazać ciągłość f w x_0 ustalmy $\epsilon>0$. Na mocy definicji pochodnej istnieje granica

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}.$$

Z definicji granicy dla $\epsilon'=1$ znajdziemy $\delta'>0$ takie, że jeśli $h\in(-\delta',\delta')$ to

$$\left|\frac{f(x_0+h)-f(x_0)}{h}-f'(x_0)\right|<1.$$

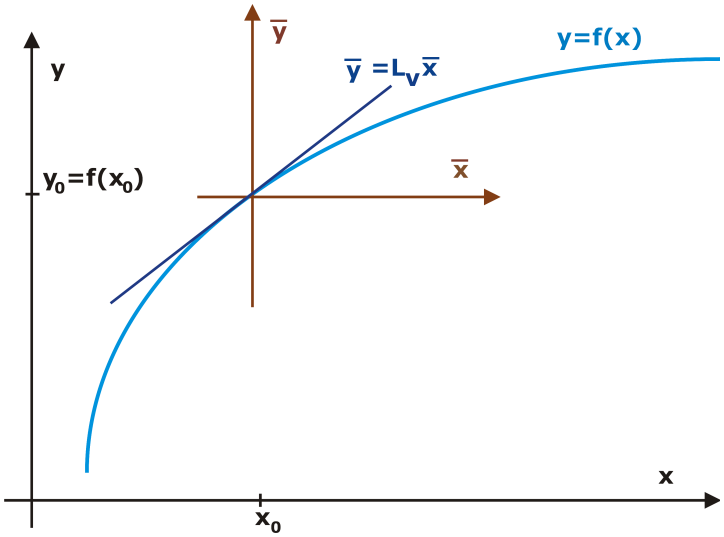
W szczególności

$$\left|\frac{f(x_0+h)-f(x_0)}{h}\right| \leqslant \left|\frac{f(x_0+h)-f(x_0)}{h}-f'(x_0)\right| + |f'(x_0)| < 1+|f'(x_0)|. \tag{11.4}$$

Weźmy $\delta := \min(\delta', \frac{\epsilon}{1+|f'(x_0)|})$. Z (11.4) dla $h\in(-\delta,\delta)$ mamy

$$|f(x_0+h)-f(x_0)| < (1+|f'(x_0)|)|h| \leqslant \epsilon$$

co pokazuje, że $\lim_{h\to 0} f(x) = f(x_0)$ i dowodzi, że f jest ciągła w x_0. □



Rysunek 11.2: Aproksymacja liniowa funkcji f w punkcie $(x_0, f(x_0))$ w lokalnym układzie współrzędnych $(\bar{x}, \bar{y})$.

Ćwiczenie 11.1.10 Uogólnić twierdzenie 11.1.9 na przypadek funkcji $f : [a, b] \to \mathbb{R}^m$.

11.2 Różniczka funkcji

11.2.1 Aproksymacja liniowa

Interpretacja pochodnej jako współczynnika kierunkowego nachylenia stycznej w danym punkcie do wykresu nie jest jedyną użyteczną intuicją w tym kontekście. Pojęcie pochodnej wiąże się również z ideą przybliżania funkcji funkcją liniową.

Definicja 11.2.1 *Funkcja liniowa* $\mathbb{R} \to \mathbb{R}^n$ jednej zmiennej to funkcja postaci

$$\mathbb{R} \ni t \mapsto tv \in \mathbb{R}^n$$

dla ustalonego $v \in \mathbb{R}^n$. Wektor v (liczbę w przypadku $n = 1$) nazywać będziemy *współczynnikiem* funkcji liniowej jednej zmiennej.

Funkcję taką oznaczać będziemy L_v, lub krótko L, jeśli współczynnik v jest nieistotny. Łatwo sprawdzić, że funkcja liniowa jest ciągła i w zerze ma wartość zero.

Uwaga 11.2.2 *Jeśli $L : \mathbb{R} \to \mathbb{R}^m$ jest funkcją liniową, to $L = L_v$ dla $v = L(1)$. Innymi słowy wartość funkcji liniowej dla argumentu jeden jest równa jej współczynnikowi.*

Funkcja liniowa daje się bardzo prosto wyliczać, więc jest dobrym kandydatem na przybliżanie (lokalne) innych funkcji. Idea jest następująca. Wiemy, że pewna funkcja f ma w punkcie x_0 wartość $y_0 := f(x_0)$. Interesuje nas wartość funkcji f w punkcie x_1 bliskm x_0, czyli $y_1 := f(x_1)$. Czy dysponując przyrostem zmiennej x, czyli różnicą $x_1 - x_0$, możemy przy pomocy pewnej funkcji liniowej wyliczyć przybliżenie przyrostu zmiennej y, czyli różnicę $y_1 - y_0 = f(x_1) - f(x_0)$? Podkreślmy, że chodzi nam o funkcję liniową w układzie współrzędnych przesuniętym do punktu (x_0, y_0), a więc w układzie $(\bar{x}, \bar{y})$, gdzie $\bar{x} := x - x_0$, $\bar{y} := y - y_0$, zwanym lokalnym układem współrzędnych względem punktu (x_0, y_0). Zwróćmy uwagę, że punkt o współrzędnych (x_0, y_0) w lokalnym układzie współrzędnych ma współrzędne $(0, 0)$. Nasze pytanie można więc sformułować następująco (rys. 11.2):

Czy istnieje funkcja liniowa L w lokalnym układzie współrzędnych $(\bar{x}, \bar{y})$, taka że

$$y_1 - y_0 \approx \bar{y} = L(\bar{x}) = L(x_1 - x_0)?$$

Zwróćmy uwagę, że to pytanie należy uściślić. Jeśli przybliżenie rozumieć będziemy jako wymóg, by w granicy, przy $x \to x_0$, popełniony błąd, czyli wyrażenie

$$y_1 - y_0 - L(x_1 - x_0) = f(x_1) - f(x_0) - L(x_1 - x_0)$$

zmierzało do zera, to wymóg ten spełni każda funkcja liniowa, bo jako ciągła ma w zerze granicę równą wartości, czyli zero. Naturalne jest jednak oczekiwanie, że im bliżej znajdziemy się x_0, tym przybliżenie będzie lepsze.

Miarą jakości przybliżenia jest błąd względem przyrostu $\bar{x} = x_1 - x_0$ mierzony ilorazem

$$\frac{y_1 - y_0 - L(x_1 - x_0)}{x_1 - x_0} = \frac{f(x_0 + \bar{x}) - f(x_0) - L(\bar{x})}{\bar{x}}.$$

Chcemy, by ten błąd zmierzał do zera przy $\bar{x} \to 0$. Niech $L = L_v$. Przy użyciu bardziej tradycyjnego oznaczenia h w miejsce $\bar{x}$, nasz wymóg można zapisać jako

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - hv}{h} = 0,$$

co jak łatwo sprawdzić, sprowadza się do żądania, by

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = v.$$

Tak więc, jeśli stosowne przybliżenie liniowe istnieje, to współczynnik v jest pochodną funkcji w punkcie x_0.

Czasami wygodnie jest patrzeć na problem sprowadzony do zera. Jak widzieliśmy wygodnie jest przesunąć układ współrzędnych tak, by interesujący nas punkt był w początku układu. Wygodnie jest też sprowadzić definiowania pochodnej do definiowania pochodnej równej zero. Z naszych wcześniejszych rozważań wynika, że pochodna równa zero oznacza, że przybliżająca funkcja liniowa ma współczynnik v równy zero. Funkcja liniowa z $\mathbb{R}$ do $\mathbb{R}$ o współczynniku zero ma wykres będący linią poziomą przechodzącą przez początek układu współrzędnych. Z drugiej strony, jeśli pochodna wynosi zero, to mamy

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = 0.$$

Zwrócmy uwagę, że funkcją, której granicę tu rozważamy jest $\varphi : h \mapsto f(x_0 + h) - f(x_0)$, gdzie x_0 traktujemy jako parametr. Przy tych oznaczeniach mamy

$$\lim_{h \to 0} \frac{\varphi(h)}{h} = 0.$$

Z definicji granicy wiemy, że wtedy dla dowolnego $\epsilon > 0$ istnieje $\delta > 0$ taka, że dla $|h| < \delta$ mamy

$$\left| \frac{\varphi(h)}{h} \right| < \epsilon$$

lub równoważnie

$$|\varphi(h)| < \epsilon |h|.$$

Funkcję o tej własności nazywamy płaską w zerze. Pojęcie to wyjaśnia rysunek 11.3.

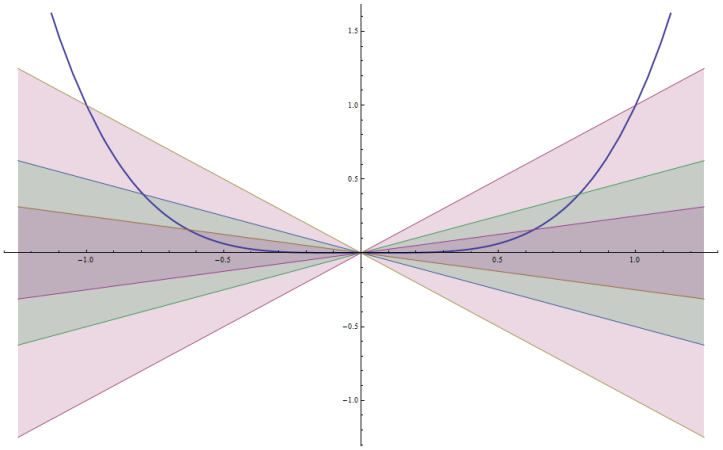
Używając języka funkcji płaskich w zerze możemy powiedzieć, że $f'(x) = a$ wtedy i tylko wtedy gdy funkcja

$$h \mapsto f(x + h) - f(x) - ah$$

jest płaska w zerze. Jest tak, bo dla $h = 0$ mamy

$$f(x + 0) - f(x) - a \cdot 0 = 0,$$

a, jak niebawem będziemy w stanie policzyć formlnie, pochodna tej funkcji w zerze wynosi $f'(x) - a$.



Rysunek 11.3: Funkcja płaska w zerze. Każdy ze stożków po zawężeniu do stosownego otocznia zera zawiera fragment wykresu.

11.2.2 Relacja "o małe"

Aby sformalizować opisane w poprzednim podrozdziale intuicje zaczniemy od podania definicji relacji "o małe", która przyda się nie tylko w podaniu formalnej definicji płaskości w zerze, ale również w następnych rozdziałach.

Rozważmy $f : (a, b) \to \mathbb{R}^m$ i $g : (a, b) \to \mathbb{R}$. Załóżmy ponadto, że istnieje V otoczenie x_0 takie, że $g(x) \neq 0$ dla $x \in V \setminus \{x_0\}$. Stawiamy definicję ogólną, ale warto zaznaczyć, że na początku interesować nas będzie głównie przypadek szczególny gdy g jest funkcją identycznościową, t.j. gdy $g(x) = x$.

Definicja 11.2.3 Mówimy, że funkcja f jest *o małe* od funkcji g w otoczeniu punktu x_0 i piszemy

$$f(x) = o(g(x)) \text{ przy } x \to x_0,$$

jeżeli

$$\forall \epsilon > 0 \; \exists \delta > 0 \quad |x - x_0| < \delta \;\Rightarrow\; ||f(x)|| < \epsilon |g(x)|.$$

Następującą uwagę można wywnioskować wprost z definicji granicy.

Uwaga 11.2.4 *Relacja $f(x) = o(g(x))$ przy $x \to x_0$ zachodzi wtedy i tylko wtedy gdy*

$$f(x_0) = 0 \; i \; \lim_{x \to x_0} \frac{||f(x)||}{|g(x)|} = 0.$$

□

Z uwagi tej łatwo wyprowadzić następujący wniosek

Wniosek 11.2.5 *Relacja $f(x) = o(g(x))$ przy $x \to x_0$ zachodzi wtedy i tylko wtedy gdy istnieje funkcja ciągła $\varphi : V \to \mathbb{R}^m$ taka, że $\varphi(x_0) = 0$ oraz $f(x) = \varphi(x)g(x)$ dla $x \in V$.*

Dowód: Wystarczy przyjąć

$$\varphi(x) := \begin{cases} \frac{f(x)}{g(x)} & \text{dla } x \neq 0, \\ 0 & \text{dla } x = 0. \end{cases}$$

□Z kolei z twierdzenia o granicy sumy i różnicy natychmiast dostajemy następujący wniosek.

Wniosek 11.2.6 *Jeśli istnieje V otoczenie x_0 takie, że $g(x) \neq 0$ dla $x \in V \setminus \{x_0\}$ oraz funkcje $f_i : \mathbb{R} \multimap \mathbb{R}$ spełniają $f_i(x) = o(g(x))$ przy $x \to x_0$ to również $(f_1 + f_2)(x) = o(g(x))$ i $(f_1 - f_2)(x) = o(g(x))$ przy $x \to x_0$* □

11.2.3 Płaskość w zerze

Podamy teraz formalną definicję płaskości funkcji w zerze. Niech $f : (a, b) \to \mathbb{R}^m$ będzie funkcją o wartościach w $\mathbb{R}^m$, określoną na przedziale otwartym $(a, b) \subset \mathbb{R}$ takim, że $0 \in (a, b)$.

Definicja 11.2.7 Mówimy, że f jest płaska w zerze jeżeli $f(x) = o(x)$ przy $x \to 0$.

Następująca uwaga wynika wprost z uwagi 11.2.4.

Uwaga 11.2.8 *Funkcja f jest płaska w zerze wtedy i tylko wtedy gdy $f(0) = 0$ oraz*

$$\lim_{x \to 0} \frac{f(x)}{x} = 0.$$

Jeśli f jest płaska w zerze, to funkcja $\varphi : (a, b) \to \mathbb{R}^m$ zdefiniowana wzorem

$$\varphi(x) := \begin{cases} \frac{f(x)}{x} & \text{gdy } x \neq 0, \\ 0 & \text{dla } x = 0. \end{cases}$$

jest ciągła w zerze oraz $f(x) = x\varphi(x)$. Łatwo sprawdzić, że również na odwrót, jeśli funkcja φ o takich własnościach istnieje, to f jest płaska w zerze. Mamy zatem następującą uwagę.

Uwaga 11.2.9 *Funkcja f jest płaska w zerze wtedy i tylko wtedy gdy istnieje funkcja $\varphi : (a, b) \to \mathbb{R}^m$ ciągła w zerze taka, że $\varphi(0) = 0$ oraz $f(x) = x\varphi(x)$ dla $x \in (a, b)$.*

11.2.4 Różniczkowalność

Wróćmy teraz do idei przybliżania przyrostu funkcji funkcją liniową. Wprowadźmy tradycyjne oznaczenia na przyrost wartości argumentu i przyrost wartości funkcji (różnice) w punkcie x_0:

$$\begin{aligned} \Delta x &:= x - x_0, \\ \Delta y &:= f(x) - f(x_0). \end{aligned}$$

Niech $L = L_{f'(x_0)}$ będzie funkcją liniową przybliżającą w loklanym układzie współrzędnych przyrost funkcji f. Wtedy

$$\Delta y \approx L(\Delta x) = f'(x_0)\Delta x, \tag{11.5}$$

przy czym popełniony błąd względny jest tym mniejszy im mniejsze jest Δx. Leibniz nie dysponował precyzyjnym aparatem i mówił, że jeśli Δx jest nieskończenie małe, to przybliżona równość (11.5) staje się równością. Takie nieskończenie małe różnice Δx i Δy nazywał różniczkami (malusieńkimi różnicami) i oznaczał dy i dx. Używając jego języka napisalibyśmy, że różniczki dx i dy spełniają równanie

$$dy = f'(x)dx. \tag{11.6}$$

Problem w tym, że różniczki jako wielkości nieskończenie małe trudno jest precyzyjnie zdefiniować. Pojęcie różniczki odczarował A. Cauchy, traktując równanie (11.6) jako równanie odwzorowania liniowego przybliżającego przyrost funkcji, a samą różniczkę definiując jako to przybliżające odwzorowanie liniowe. I tak do dziś rozumiemy różniczkę. Natomiast termin różniczkowalność, ze względu na bezpośredni związek pomiędzy różniczką i pochodną, jest odnoszony zarówno do istnienia pochodnej jak i różniczki, bo, jak zobaczymy, są to fakty równoważne.

Jak poprzednio, niech $f : (a, b) \to \mathbb{R}^m$ będzie funkcją określoną na przedziale otwartym $(a, b) \subset \mathbb{R}$ o wartościach w $\mathbb{R}^m$. Rozważmy *przyrost funkcji f w punkcie $x_0 \in (a, b)$*, t.j. funkcję $\Delta_{x_0} f : (a - x_0, b - x_0) \to \mathbb{R}^m$ określoną wzorem

$$\Delta_{x_0} f(h) := f(x_0 + h) - f(x_0).$$

Definicja 11.2.10 Funkcję f nazywamy różniczkowalną w punkcie x_0 jeżeli istnieje funkcja liniowa $L : \mathbb{R} \to \mathbb{R}^m$ taka, że funkcja $\Delta_{x_0} f - L$, czyli różnica funkcji $\Delta_{x_0} f$ i funkcji L jest płaska w zerze. Odwzorowanie L nazywamy wtedy *różniczką* funkcji f w punkcie x_0 i oznaczamy $d_{x_0} f$.

Aby powyższa definicja miała sens, musimy wiedzieć, że różniczka jest wyznaczona jednoznacznie. W istocie różniczka jest wyznaczona przez pochodną, jak pokazuje następujące rozumowanie. Jeśli L jest różniczką fukcji f w punkcie x_0 to z płaskości w zerze funkcji $\Delta_{x_0} f - L$ wnosimy, że

$$\lim_{h \to 0} \frac{\Delta_{x_0} f(h) - L(h)}{h} = 0.$$

Ponieważ $L(h) = L(1)h$, jest to równoważne warunkowi

$$\lim_{h \to 0} \frac{\Delta_{x_0} f(h)}{h} - L(1) = 0$$

oraz warunkowi

$$\lim_{h \to 0} \frac{\Delta_{x_0} f(h)}{h} = L(1).$$

Ponieważ $\Delta_{x_0} f(h) = f(x_0 + h) - f(x_0)$, powyższa granica pokrywa się z granicą (11.1). Pokazaliśmy zatem następującą uwagę.

Uwaga 11.2.11 *Funkcja liniowa $L : \mathbb{R} \to \mathbb{R}^m$ jest różniczką funkcji f w punkcie x_0 wtedy i tylko wtedy gdy $L(1)$ jest pochodną tej funkcji w punkcie x_0. W szczególności f jest różniczkowalna w puncie x_0 wtedy i tylko wtedy gdy f ma pochodną w punkcie x_0.*

Tak więc, jeśli odwzorowanie liniowe L spełniające wymogi definicji 11.2.10 istnieje, to współczynnik $L(1)$, a zatem i samo odwzorowanie L są wyznaczone jednoznacznie.

Z uwagi 11.2.11 wynika, że stwierdzenia "f jest różniczkowalna" i "f ma pochodną" mogą być używane zamiennie. Jednak pojęcia różniczki i pochodnej nie są tożsame, choć są ściśle powiązane: wartość różniczki dla argumentu jeden jest pochodną. Różniczka jest odwzorowaniem liniowym, które przemnaża argument przez pochodną. Koncepcja różniczki bazuje na aproksymacji liniowej funkcji w punkcie. Koncepcja pochodnej bazuje na stycznej do wykresu funkcji w punkcie. Związek między różniczką, a pochodną można więc też wyrazić stwierdzając, że wykres aproksymacji liniowej funkcji f w zadanym punkcie jest prostą, która pokrywa się ze styczną do wykresu funkcji f w tym punkcie.

Zauważmy, że jeśli funkcja jest płaska w zerze, to ma pochodną w zerze równą zero. Rzeczywiście, na mocy uwagi 11.2.8 jeśli f jest płaska w zerze, to $f(0) = 0$ oraz $\lim_{x \to 0} \frac{f(x)}{x} = 0$. Tę ostanią granicę można przepisać jako

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

co oznacza, że f ma w zerze pochodną zero. Łatwo zauważyć, że można przeprowadzić również odwrotne wnioskowanie. Prowadzi to do następującej uwagi.

Uwaga 11.2.12 *Niech $f : \mathbb{R} \multimap \mathbb{R}^m$ będzie funkcją określoną w otoczeniu zera. Funkcja f jest płaska w zerze wtedy i tylko wtedy gdy $f(0) = 0$ i $f'(0) = 0$.* □

11.3 Różniczkowanie sumy, różnicy, iloczynu i ilorazu

Naturalne jest pytanie czy znając pochodną dwóch funkcji możemy wyznaczyć pochodną sumy, różnicy, iloczynu i ilorazu tych funkcji. Jest tak w istocie, jak pokazują następujące twierdzenia.

Twierdzenie 11.3.1 (o pochodnej sumy i różnicy) *Niech $f, g : (a, b) \to \mathbb{R}^m$ będą różniczkowalne w punkcie $x_0 \in (a, b)$. Wtedy funkcje $f + g$ i $f - g$ są różniczkowalne w x_0 oraz*

$$(f + g)'(x_0) = f'(x_0) + g'(x_0) \quad i \quad (f - g)'(x_0) = f'(x_0) - g'(x_0).$$

Dowód: Iloraz różnicowy dla sumy $f + g$ to

$$\frac{(f + g)(x_0 + h) - (f + g)(x_0)}{h} = \frac{f(x_0 + h) - f(x_0)}{h} + \frac{g(x_0 + h) - g(x_0)}{h}. \tag{11.7}$$

Ponieważ

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0) \quad \text{i} \quad \lim_{h\to 0} \frac{g(x_0 + h) - g(x_0)}{h} = g'(x_0),$$

więc z (11.7) wnosimy, że

$$\lim_{h\to 0} \frac{(f + g)(x_0 + h) - (f + g)(x_0)}{h} = f'(x_0) + g'(x_0),$$

co dowodzi tezy dla sumy. Dla różnicy dowód jest analogiczny. □

Twierdzenie o pochodnej iloczynu przedstawimy dla funkcji o wartościach w zbiorze liczb zespolonych $\mathbb{C}$. Oczywiście twierdzenie to obejmuje również przypadek funkcji o wartościach rzeczywistych, bo zbiór liczb rzeczywistych traktujemy jako podzbiór zbioru liczb zespolonych. Jako ćwiczenie zostawimy przypadek, gdy jedna funkcja przyjmuje wartości skalarne, a druga wektorowe, a iloczyn funkcji jest rozumiany jako mnożenie skalara przez wektor oraz przypadek, gdy obie funkcje są wektorowe, a iloczyn jest rozumiany jako iloczyn skalarny wektorów.

Twierdzenie 11.3.2 (o pochodnej iloczynu) *Niech $f, g : (a, b) \to \mathbb{C}$ będą różniczkowalne w punkcie $x_0 \in (a, b)$. Wtedy funkcja fg jest różniczkowalna w x_0 oraz*

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

Dowód: Iloraz różnicowy dla iloczynu fg to

$$\begin{aligned}
\frac{(fg)(x_0 + h) - (fg)(x_0)}{h} &= \frac{f(x_0 + h)g(x_0 + h) - f(x_0)g(x_0)}{h} \\
&= \frac{f(x_0 + h)g(x_0 + h) - f(x_0 + h)g(x_0) + f(x_0 + h)g(x_0) - f(x_0)g(x_0)}{h} \\
&= f(x_0 + h)\frac{g(x_0 + h) - g(x_0)}{h} + \frac{f(x_0 + h) - f(x_0)}{h}g(x_0)
\end{aligned}$$

Funkcja f jako różniczkowalna w x_0 jest również ciągła w x_0, więc

$$\lim_{h\to 0} \left(f(x_0 + h)\frac{g(x_0 + h) - g(x_0)}{h} + \frac{f(x_0 + h) - f(x_0)}{h}g(x_0) \right) = f(x_0)g'(x_0) + f'(x_0)g(x_0).$$

Zatem z otrzymanego wyżej przedstawienia ilorazu różnicowego dla iloczynu również

$$(fg)'(x_0) = \lim_{h\to 0} \frac{(fg)(x_0 + h) - (fg)(x_0)}{h} = f(x_0)g'(x_0) + f'(x_0)g(x_0)$$

co dowodzi naszej tezy. □

Z twierdzenia o pochodnej iloczynu i z faktu, że pochodna funkcji stałej wynosi zero dostajemy natychmiast następujący wniosek.

Wniosek 11.3.3 *Niech $f : (a, b) \to \mathbb{C}$ będzie różniczkowalne w punkcie $x_0 \in (a, b)$ i niech $c \in \mathbb{C}$ będzie stałą. Wtedy funkcja cf jest różniczkowalna w x_0 oraz*

$$(cf)'(x_0) = cf'(x_0).$$

Z twierdzenia o pochodnej iloczynu wyprowadzimy też następujące twierdzenie

Twierdzenie 11.3.4 (o pochodnej funkcji potęgowej) *Niech $n \in \mathbb{N}_0$ będzie ustaloną liczbą naturalną. Funkcja potęgowa*

$$\mathbb{R} \ni x \mapsto x^n \in \mathbb{R}$$

jest różniczkowalna. Jej pochodną jest funkcja

$$\mathbb{R} \ni x \mapsto nx^{n-1} \in \mathbb{R}$$

co zapisujemy mniej formalnie, ale krócej jako

$$(x^n)' = nx^{n-1}.$$

Dowód: Dowód poprowadzimy indukcyjnie względem n. Dla $n \in \{0, 1\}$ twierdzenie wynika natychmiast z twierdzeń 11.1.2 i 11.1.3. Załóżmy, że $n > 1$ oraz że twierdzenie zachodzi dla $k < n$. Funkcja $x \mapsto x^n$ jest wtedy iloczynem funkcji identycznościowej oraz funkcji $x \mapsto x^{n-1}$. Z założenia indukcyjnego mamy $'x^{n-1})' = (n-1)x^{n-2}$, bo $n - 1 < n$. Zatem z twierdzenia o pochodnej iloczynu mamy

$$(x^n)' = (xx^{n-1})' = x'x^{n-1} + x(n-1)x^{n-2} = nx^{n-1}$$

co dowodzi tezy. □

Ćwiczenie 11.3.5 Niech $\alpha : (a, b) \to \mathbb{R}$ i $f : (a, b) \to \mathbb{R}^m$ będą różniczkowalne w $x_0 \in (a, b)$. Pokaż, że funkcja

$$\alpha f : (a, b) \ni x \mapsto \alpha(x)f(x) \in \mathbb{R}^m$$

jest różniczkowalna w x_0 oraz

$$(\alpha f)'(x_0) = \alpha'(x_0)f(x_0) + \alpha(x_0)f'(x_0).$$

Wyprowadź stosowny wniosek gdy funkcja α lub funkcja f jest stała.

Ćwiczenie 11.3.6 Niech $f, g : (a, b) \to \mathbb{R}^m$ będą różniczkowalne w $x_0 \in (a, b)$. Pokaż, że funkcja

$$f \cdot g : (a, b) \ni x \mapsto f(x) \cdot g(x) \in \mathbb{R}$$

jest różniczkowalna w x_0 oraz

$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0).$$

Wyprowadź stosowny wniosek gdy funkcja f lub funkcja g jest stała.

Twierdzenie 11.3.7 (o pochodnej ilorazu) *Niech $f, g : (a,b) \to \mathbb{C}$ będą różniczkowalne w $x_0 \in (a,b)$ i niech $g(x_0) \neq 0$. Wtedy funkcja*

$$x \mapsto \frac{f(x)}{g(x)}$$

jest dobrze określona w pewnym otoczeniu x_0 i różniczkowalna w x_0 oraz

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}.$$

Dowód: Zauważmy, że funkcja g jako różniczowalna w x_0 jest ciągła w x_0, skąd wnosimy, że własność $g(x_0) \neq 0$ przenosi się na pewne otoczenie x_0. Zatem iloraz funkcji f i g jest dobrze określony w pewnym otoczeniu punktu x_0. Iloraz różnicowy dla ilorazu funkcji f i g to

$$\begin{aligned}
\frac{\left(\frac{f}{g}\right)(x_0+h) - \left(\frac{f}{g}\right)(x_0)}{h} &= \frac{\frac{f(x_0+h)}{g(x_0+h)} - \frac{f(x_0)}{g(x_0)}}{h} \\
&= \frac{f(x_0+h)g(x_0) - f(x_0)g(x_0+h)}{hg(x_0)g(x_0+h)} \\
&= \frac{f(x_0+h)g(x_0) - f(x_0)g(x_0) + f(x_0)g(x_0) - f(x_0)g(x_0+h)}{hg(x_0)g(x_0+h)} \\
&= \frac{(f(x_0+h) - f(x_0))\,g(x_0) - f(x_0)\,(g(x_0+h) - g(x_0))}{hg(x_0)g(x_0+h)} \\
&= \frac{\frac{f(x_0+h)-f(x_0)}{h}g(x_0) - f(x_0)\frac{g(x_0+h)-g(x_0)}{h}}{g(x_0)g(x_0+h)}
\end{aligned}$$

Przechodząc w otrzymanej równości z h do granicy w zerze dostajemy tezę. □

Twierdzenie 11.3.8 (o pochodnej zestawienia) *Niech $f_i : (a,b) \to \mathbb{C}$ będą funkcjami różniczkowalnymi w $x_0 \in (a,b)$ dla $i = 1, 2, \ldots m$. Wtedy zestawienie $f := (f_1, f_2, \ldots f_m) : (a,b) \to \mathbb{C}^m$ jest różniczkowalne w x_0 oraz*

$$f'(x_0) = (f_1'(x_0), f_2'(x_0), \ldots, f_m'(x_0))$$

Dowód: Teza wynika natychmiast z twierdzenia o granicy zestawienia funkcji (tw. 6.8.1), gdyż

$$\frac{f(x_0+h) - f(x_0)}{h} = \left(\frac{f_i(x_0+h) - f_i(x_0)}{h}\right)_{i=1}^{m}$$

□

11.4 Różniczkowanie funkcji złożonej

Bardzo przydatna w liczeniu pochodnych jest informacja jak pochodna złożenia funkcji wyraża się poprzez pochodne funkcji składanych. Nim przedstawimy stosowne twierdzenie, zajmiemy się twierdzeniem pomocniczym, które przyda się nam później w dowodzie twierdzenia o różniczkowaniu złożenia.

Zauważmy, że jeśli $f : \mathbb{R} \multimap \mathbb{R}^m$ jest różniczkowalna w $x_0 \in \mathbb{R}$, to funkcja $h \mapsto \frac{\Delta_{x_0} f(h)}{h} - f'(x_0)$ ma w zerze granicę zero, choć sama w zerze nie jest zdefiniowana, bo ułamek $\frac{\Delta_{x_0} f(h)}{h}$ dla $h = 0$ nie ma sensu. Możemy jednak rozszerzyć tę funkcję do funkcji $r : \mathbb{R} \multimap \mathbb{R}^m$, która z definicji w zerze przyjmuje wartość zero. Będzie to funkcja ciągła i będzie spełniać wzór

$$\Delta_{x_0} f(h) - hv = hr(h).$$

Okazuje się, że implikacja odwrotna również zachodzi, co prowadzi do następującego twierdzenia.

Twierdzenie 11.4.1 *Niech $f : \mathbb{R} \multimap \mathbb{R}^m$ będzie funkcją określoną w otoczeniu $x_0 \in \mathbb{R}$. Wtedy różniczkowalność f w x_0 jest równoważna istnieniu funkcji $r : \mathbb{R} \multimap \mathbb{R}^m$ określonej i ciągłej w otoczeniu zera oraz wektora $v \in \mathbb{R}^m$ takich, że*

$$\begin{aligned} r(0) &= 0 \\ \Delta_{x_0} f(h) - hv &= hr(h). \end{aligned}$$

Dowód: Załóżmy najpierw, że f jest różniczkowalna w x_0. Połóżmy $v := f'(x_0)$ oraz zdefiniujmy funkcję r wzorem

$$r(h) := \begin{cases} \frac{\Delta_{x_0} f(h)}{h} - v & \text{gdy } x \neq 0 \\ 0 & \text{w przeciwnym razie.} \end{cases}$$

Oczywiście r jest określona w otoczeniu zera i spełnia wzór (11.8) Ponieważ $f'(x_0) = v$, więc

$$\lim_{h \to 0} \frac{\Delta_{x_0} f(h)}{h} - v = 0$$

co oznacza, że r jest również ciągła. Na odwrót, jeśli funkcja o postulowanych własnościach istnieje, to z jej ciągłości w zerze wynika, że istnieje granica

$$\lim_{h \to 0} \frac{\Delta_{x_0} f(h)}{h},$$

a tym samym f jest różniczkowalna w x_0.

Twierdzenie 11.4.2 (o pochodnej złożenia funkcji) *Niech $f : (a, b) \to (c, d)$ będzie funkcją różniczkowalną w $x_0 \in (a, b)$, a $g : (c, d) \to \mathbb{R}^m$ funkcją różniczkowalną w $f(x_0)$. Wtedy złożenie $g \circ f$ jest różniczkowalne w x_0 oraz*

$$(g \circ f)'(x_0) = g'(f(x_0)) f'(x_0).$$

Dowód: Z założenia o różniczkowalności f w x_0 i g w $f(x_0)$ oraz z twierdzenia 11.4.1 wnosimy, że istnieje V otoczenie zera oraz ciągłe, zerujące się w zerze funkcje $\varphi, \psi : V \to \mathbb{R}$ takie, że dla $h, k \in V$

$$\Delta_{x_0} f(h) = f(x_0 + h) - f(x_0) = (f'(x_0) + \varphi(h))\, h \tag{11.8}$$

$$\Delta_{f(x_0)} g(k) = g(f(x_0) + k) - g(f(x_0)) = (g'(f(x_0)) + \psi(k))\, k. \tag{11.9}$$

Ponieważ funkcja $h \mapsto (f'(x_0) + \varphi(h))\, h$ jest ciągła i zeruje się w zerze, więc istnieje W otoczenie zera takie, że dla $h \in W$ mamy

$$(f'(x_0) + \varphi(h))\, h \in V.$$

Zatem wzór zachodzi również dla $k = (f'(x_0) + \varphi(h))\, h$ przy $h \in V$. Mamy zatem dla $h \in V$

$$g(f(x_0) + (f'(x_0) + \varphi(h))\, h) - g(f(x_0)) = (g'(f(x_0)) + \psi(f'(x_0) + \varphi(h))\, h)\, (f'(x_0) + \varphi(h))\, h. \tag{11.10}$$

Wykorzystując (11.8) oraz (11.10) otrzymujemy

$$\begin{aligned} \Delta_{x_0}(g \circ f)(h) &= g(f(x_0 + h)) - g(f(x_0)) = g(f(x_0) + (f'(x_0) + \varphi(h))\, h) - g(f(x_0)) \\ &= (g'(f(x_0)) + \psi(f'(x_0) + \varphi(h))\, h)\, (f'(x_0) + \varphi(h))\, h. \end{aligned}$$

Stąd

$$\frac{\Delta_{x_0}(g \circ f)(h)}{h} = (g'(f(x_0)) + \psi(f'(x_0) + \varphi(h))\, h)\, (f'(x_0) + \varphi(h)).$$

Ponieważ tak ψ jak i φ zmierzają do zera przy h zmierzającym do zera, prawa strona powyższej równości zmierza do $g'(f(x_0))f'(x_0)$. Zatem również lewa strona jest zbieżna do $g'(f(x_0))f'(x_0)$. Dowodzi to tezy, bo lewa strona to iloraz różnicowy dla funkcji $g \circ f$ w x_0. □

11.5 Różniczkowanie funkcji wykładniczej i funkcji trygonometrycznych

W rozdziale tym podamy wzory na pochodną funkcji wykładniczej oraz funkcji trygonometrycznych. Wzory te uzyskamy badając pochodną funkcji

$$\mathbb{R} \ni t \mapsto e^{tz} \in \mathbb{C} \tag{11.11}$$

dla ustalonej liczby zespolonej $z \in \mathbb{C}$. Stosowny wzór podamy dla dowolnego zespolonego z, ale interesować nas będą głównie przypadki gdy z jest liczbą rzeczywistą bądź liczbą urojoną i.

Zauważmy, że iloraz różnicowy dla funkcji (11.11) ma postać

$$\frac{e^{(t+h)z} - e^{tz}}{h} = e^{tz}\frac{e^{hz}-1}{h},$$

więc interesować nas będzie granica $\lim_{h\to 0}\frac{e^{hz}-1}{h}$.

Lemat 11.5.1 *Funkcja*

$$\mathbb{C}\setminus\{0\} \ni u \mapsto \frac{e^u - 1}{u} \in \mathbb{C}$$

ma w zerze granicę równą jeden.

Dowód: Przypomnijmy, że zgodnie z definicją 9.1.3 mamy

$$e^u = \sum_{n=0}^{\infty}\frac{u^n}{n!},$$

skąd

$$e^u - 1 = \sum_{n=1}^{\infty}\frac{u^n}{n!},$$

bo $\frac{u^0}{0!} = 1$. Zatem dla dowolnego $u \in \mathbb{C}\setminus\{0\}$ mamy

$$\begin{aligned}\left|\frac{e^u-1}{u} - 1\right| &= \left|\frac{\sum_{n=0}^{\infty}\frac{u^n}{n!} - 1}{u} - 1\right| = \left|\frac{\sum_{n=1}^{\infty}\frac{u^n}{n!}}{u} - 1\right| = \left|\sum_{n=1}^{\infty}\frac{u^{n-1}}{n!} - 1\right| \\ &= \left|\sum_{n=2}^{\infty}\frac{u^{n-1}}{n!}\right| \leqslant \sum_{n=2}^{\infty}\frac{|u|^{n-1}}{n!} = |u|\sum_{n=2}^{\infty}\frac{|u|^{n-2}}{n!}\end{aligned}$$

Ponieważ interesuje nas tylko otoczenie zera, możemy założyć, że $|u| < 1$. Dla takich u dostajemy

$$\left|\frac{e^u-1}{u} - 1\right| \leqslant |u|\sum_{n=2}^{\infty}\frac{1}{n!} \leqslant |u|e.$$

Tak więc dla $|u| < 1$ mamy

$$0 \leqslant \left|\frac{e^u-1}{u} - 1\right| \leqslant |u|e.$$

Ponieważ prawa strona tej nierówności zmierza do zera przy u zmiarzającym do zera, z twierdzenia o trzech funkcjach wnosimy, że

$$\lim_{u\to 0}\left|\frac{e^u-1}{u} - 1\right| = 0,$$

co dowodzi tezy. □

Twierdzenie 11.5.2 *Funkcja*

$$\mathbb{R} \ni t \mapsto e^{tz} \in \mathbb{C}$$

jest różniczkowalna, a jej pochodną jest funkcja

$$\mathbb{R} \ni t \mapsto ze^{tz} \in \mathbb{C}$$

Dowód: Dla $z = 0$ rozważana funkcja jest stale równa jeden, zatem jest różniczkowalna, a jej pochodną jest funkcja stale równa zero, co jest zgodne z tezą twierdzenia. Załóżmy zatem, że $z \neq 0$. Pochodną rozważanej funkcji w punkcie z jest

$$\lim_{h\to 0} \frac{e^{(t+h)z} - e^z}{h} = e^{tz} \lim_{h\to 0} \frac{e^{hz} - 1}{h} = ze^{tz} \lim_{h\to 0} \frac{e^{hz} - 1}{hz} = ze^{tz},$$

bo z lematu 11.5.1 wiemy, że

$$\lim_{h\to 0} \frac{e^{hz} - 1}{hz} = \lim_{u\to 0} \frac{e^u - 1}{u} = 1.$$

□

Wniosek 11.5.3 (o różniczkowaniu rzeczywistej funkcji wykładniczej) *Rzeczywista funkcja wykładnicza*

$$\mathbb{R} \ni x \mapsto e^x \in \mathbb{R}^+$$

jest różniczkowalna, a jej pochodną jest ona sama, co nieprecyzyjnie, ale zwięźle zapisujemy jako

$$(e^x)' = e^x.$$

Dowód: Tezę dostajemy natychmiast z twierdzenia 11.5.2 po podstawieniu za z stałej jeden. □

Twierdzenie 11.5.4 (o różniczkowaniu funkcji trygonometrycznych) *Funkcje trygonometryczne:* sin, cos, tg, ctg *są różniczkowalne w swojej dziedzinie i zachodzą wzory*

$$(\sin t)' = \cos t, \quad (\cos t)' = -\sin t, \quad (\operatorname{tg} t)' = \frac{1}{\cos^2 x}, \quad (\operatorname{ctg} t)' = -\frac{1}{\sin^2 x}.$$

Dowód: Z twierdzenia 11.5.2 dla $z = i$ oraz dla $z = -i$ mamy

$$(\sin t)' = \left(\frac{e^{it} - e^{-it}}{2i}\right)' = \frac{1}{2i}\left((e^{it})' - (e^{-it})'\right) = \frac{1}{2i}\left(ie^{it} + ie^{-it}\right) = \frac{e^{it} + e^{-it}}{2} = \cos t.$$

Podobnie

$$(\cos t)' = \left(\frac{e^{it} + e^{-it}}{2}\right)' = \frac{1}{2}\left((e^{it})' + (e^{-it})'\right) = \frac{1}{2}\left(ie^{it} - ie^{-it}\right) = -\frac{e^{it} - e^{-it}}{2i} = -\sin t.$$

Z twierdzenia o granicy ilorazu mamy

$$(\operatorname{tg} t)' = \left(\frac{\sin t}{\cos t}\right)' = \frac{(\sin t)'\cos t - \sin t(\cos t)'}{\cos^2 t} = \frac{\cos t \cos t - \sin t(-\sin t)}{\cos^2 t} = \frac{1}{\cos^2 t}.$$

Podobnie

$$(\operatorname{ctg} t)' = \left(\frac{\cos t}{\sin t}\right)' = \frac{(\cos t)'\sin t - \cos t(\sin t)'}{\sin^2 t} = \frac{-\sin t \sin t - \cos t(\cos t)}{\sin^2 t} = -\frac{1}{\sin^2 t}.$$

□

11.6 Nieciągłości pochodnej i funkcje klasy C^1.

W przykładzie 11.1.4 pokazaliśmy, że funkcja $x \mapsto |x|$ nie posiada w zerze pochodnej. Oczywiście poza zerem jest różniczkowalna. Jej nieróżniczkowalność w zerze bierze się stąd, że pochodne lewostronna i prawostronna, choć istnieją, nie są sobie równe. Poniższy przykład pokazuje, że nawet pochodne jednostronne mogą w pewnym punkcie nie istnieć.

Przykład 11.6.1 Rozważmy funkcję

$$\mathbb{R} \ni x \mapsto \begin{cases} x \sin \frac{1}{x} & \text{gdy } x \neq 0, \\ 0 & \text{dla } x = 0. \end{cases}$$

Funkcja ta jest ciągła, a dla $x \neq 0$ różniczkowalna na podstawie twierdzenia o różniczkowaniu iloczynu, funkcji złożonej i ilorazu. Dla $x = 0$ iloraz różnicowy ma postać

$$\frac{h \sin \frac{1}{h}}{h} = \sin \frac{1}{h}.$$

Wiemy, że granica $\lim_{h\to 0} \sin \frac{1}{h}$ nie istnieje. Nie istnieją nawet stosowne granice jednostronne. Zatem rozważana funkcja nie jest różniczowalna w zerze i nie posiada w zerze pochodnych jednostronnych.

Interesujące jest pytanie, czy jeśli funkcja jest różniczkowalna w każdym punkcie, to jej pochodna musi być funkcją ciągłą. Okazuje się, że niekoniecznie, jak pokazuje poniższy przykład.

Przykład 11.6.2 Rozważmy funkcję

$$\mathbb{R} \ni x \mapsto \begin{cases} x^2 \sin \frac{1}{x} & \text{gdy } x \neq 0, \\ 0 & \text{dla } x = 0. \end{cases}$$

Funkcja ta jest ciągła, a dla $x \neq 0$ różniczkowalna na podstawie twierdzenia o różniczkowaniu iloczynu, funkcji złożonej i ilorazu. Dla $x = 0$ iloraz różnicowy ma postać

$$\frac{h^2 \sin \frac{1}{h}}{h} = h \sin \frac{1}{h}.$$

Granica $\lim_{h\to 0} h \sin \frac{1}{h}$ istnieje i wynosi zero. Zatem również w zerze nasza funkcja jest różniczkowalna. Jej pochodną jest funkcja

$$\mathbb{R} \ni x \mapsto \begin{cases} 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right) & \text{gdy } x \neq 0, \\ 0 & \text{dla } x = 0. \end{cases}$$

Pokażemy, że pochodna ta nie jest ciągła w zerze. Gdyby była ciągła w zerze, to dla każdego ciągu x_n zbieżnego do zera ciąg wartości byłby zbieżny do pochodnej w zerze, czyli do zera. Ale ciąg $x_n = \frac{1}{2\pi n}$ jest zbieżny do zera, a ciąg wartości pochodnej w punktach tego ciągu wynosi stale -1, więc jest zbieżny do -1, a nie do zera.

Definicja 11.6.3 Funkcję $f : (a, b) \to \mathbb{R}^m$, która jest różniczkowalna i ma ciągłą pochodną nazywamy *funkcją klasy* C^1. Zbiór takich funkcji oznaczamy $C^1((a, b), \mathbb{R}^m)$.

Nazwa "klasa C^1" wyjaśni się później, gdy zdefiniujemy również funkcje klasy C^n.

W programie Mathematica pochodną funkcji f zmiennej x obliczamy używając instrukcji

```
D[f[x],x]
```

Na przykład pochodną funkcji

$$f(x) := \frac{1}{4}\log\left(\frac{x^2-1}{x^2+1}\right)$$

obliczymy pisząc

```
D[(1/4) Log[(x^2 - 1)/(x^2 + 1)], x]
```

Jako wynik program Mathematica zwraca

$$\frac{\left(x^2+1\right)\left(\frac{2x}{x^2+1}-\frac{2x\left(x^2-1\right)}{\left(x^2+1\right)^2}\right)}{4\left(x^2-1\right)}.$$

Nie jest to czytelna postać pochodnej. Instrukcja `Simplify` pozwala uzyskać najbardziej zwartą postać wyrażenia. Jako argumentu można użyć znaku %, który odnosi się do ostatniego wyrażenia obliczonego przez program Mathematica.

```
Simplify[%]
```

daje w wyniku

$$\frac{x}{x^4-1}.$$

11.7 Metoda Newtona

Rozwiązywanie równań interesowało matematyków już od starożytności. Pierwotnie stosowano głównie metody geometryczne, które pozwalały rozwiązywać równania liniowe i wybrane równania wyższych stopni, głównie kwadratowe. Po raz pierwszy podejście algebraiczne w rozwiązywaniu równania kwadratowego zastosowano w VII wieku w Indiach. W XVIw. metodę rozwiązywania równania trzeciego i czwartego stopnia znaleźli włoscy matematycy Niccolo Fontana Tartaglia i Lodovico Ferrari. Liczne próby ogólnego rowiązania równań wyższych stopni spełzły na niczym, a na początku XIXw. norwerski matematyk Niels Henrik Abel udowodnił, że nie da się podać ogólnego algorytmu rozwiązania równania stopnia większego niż cztery za pomocą skończonej ilości operacji arytmetycznych $+, -, *, /$ oraz pierwiastkowania.

Pozostały zatem metody przybliżone. Wsród nich prym wiedzie wywodząca się z prac Izaaka Newtona i nazywana jego imieniem metoda iteracyjna wykorzystująca przybliżanie funkcji styczną do wykresu.

Załóżmy, że chcemy rozwiązać równanie postaci

$$f(x) = 0,$$

gdzie $f : [a, b] \to \mathbb{R}$ jest funkcją różniczkowalną. Rozwiązania takiego równania nazywane są również miejscami zerowymi funkcji f.

Załóżmy, że dany jest punkt $x_0 \in [a, b]$, a $y = cx + d$ jest równaniem prostej stycznej do wykresu f w punkcie x_0.

Wiemy, że c, czyli współczynnik kierunkowy tej prostej jest równy $f'(x_0)$. Ponieważ punkt $(x_0, f(x_0))$ należy do stycznej, więc $f(x_0) = f'(x_0)x_0 + d$, skąd wyliczamy d co pozwala przepisać nam równanie stycznej w postaci

$$y = f(x_0) + f'(x_0)\left(x - x_0\right).$$

Przyjmując $y = 0$ i rozwiązując względem x znajdujemy

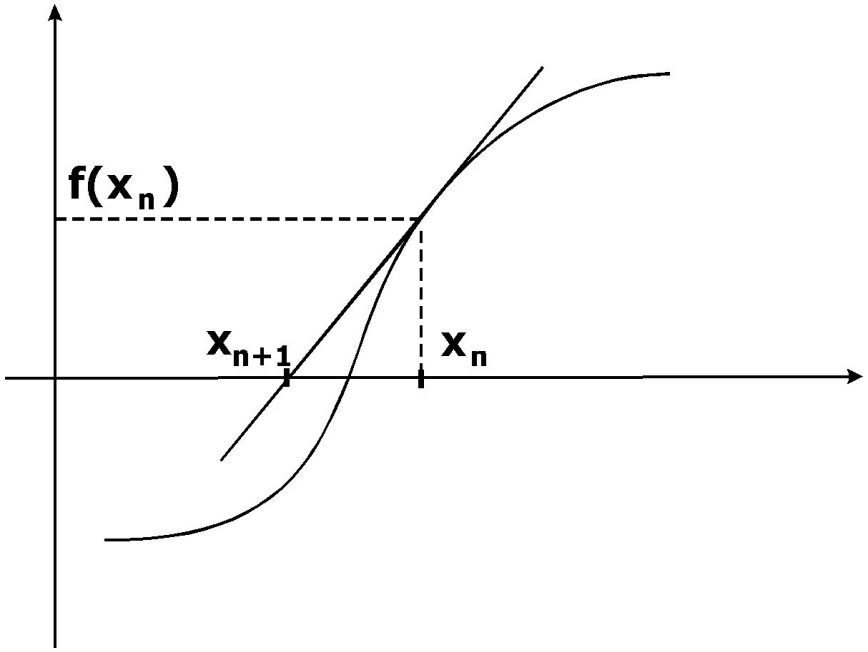
$$x = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Jest to przybliżenie miejsca zerowego funkcji f tym lepsze im bliżej x_0 było miejsca zerowego. Dlatego iterując ten proces, to znaczy licząc wyrazy ciągu danego rekurencyjnie wzorem

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$$

```
/********** newton.cpp *************/
/*      Pierwiastek wielomianu      */
/*      metodą Newtona              */
/************************************/
#include <iostream>
using namespace std;
/* Stopień wielomianu */
const int degree=2;
/* Tablica współczynników wielomianu */
double coef[degree+1];
/* maksymalna liczba iteracji */
int limit=30;
/* oczekiwana dokładność */
double epsilon=0.001;
/* Wyliczenie wartości wielomianu w x */
double f(double x){
  double s=coef[degree];
  for(int i=degree-1;i>=0;i--) s=coef[i]+x*s;
  return s;
}
/* Wyliczenie pochodnej wielomianu w x */
double derf(double x){
  double s=degree*coef[degree];
  for(int i=degree-1;i>=1;i--) s=i*coef[i]+x*s;
  return s;
}
/* Program główny */
int main(){
  /* Wczytujemy współczynniki wielomianu */
  for(int i=0;i<degree+1;++i){
    cout << "c[" << i << "]="; cin >> coef[i];
  }
  /* oraz wyjściowe przybliżenie pierwiastka */
  double x;
  cout << "x0="; cin >> x;
  /* Pętla główna */
  for(int i=0;i<limit;++i){
    double val=f(x);
    cout << "   x   =" <<  x   << "\n";
    cout << "   f(x)=" <<  val << "\n";
    /* Jeśli osiągnięto żądaną dokładność można przerwać pętlę */
    if(val<epsilon && val>-epsilon) break;
    /* Nowe przybliżenie */
    x=x-val/derf(x);
  }
}
```

Listing 11.1: Wyznaczanie miejsc zerowych wielomianu w C++.



Rysunek 11.4: Interpretacja geometryczna metody Newtona

możemy się spodziewać, że będzie on zbieżny do miejsca zerowego funkcji f w przedziale $[a, b]$. Nie zawsze tak jest, ale dzieje się tak na tyle często, że metoda Newtona jest do dziś bardzo powszechnie stosowanym narzędziem. Interpretacja geometryczna związku x_n i x_{n+1} z pochodną funkcji f jest przedstawiona na rys. (11.4).

Program w C++ znajdujący miejsca zerowe wielomianu przedstawiono na listingu 11.1.

W programie Mathematica rozwiązanie równania metodą Newtona umożliwia insturkcja `FindRoot`*. Ponieważ program Mathematica sam potrafi znaleźć pochodną, jedyne co musimy zrobić to zapisać równanie i podać wartość, od której należy rozpocząć poszukiwanie pierwiastka. Na przykład pisząc*

```
FindRoot[Cos[x] == x^3, {x, 1}]
```

otrzymujemy jako rozwiązanie

```
{x -> 0.865474}
```

Program Mathematica przekształca równanie $\cos x = x^3$ do postaci $f(x) = 0$, gdzie

$$f(x) := \cos x - x^3,$$

znajduje pochodną $f'(x) = -\sin x - 3x^2$ i następnie liczy wyrazy ciągu danego wzorem rekurencyjnym

$$\begin{aligned} x_0 &:= 1 \\ x_{n+1} &:= x_n - \frac{\cos x_n - x_n^3}{-\sin x_n - 3x_n^2} \end{aligned}$$

aż do osiągnięcia przybliżenia granicy o domyślnie ustawionej dokładności.

Ćwiczenie 11.7.1 Pokaż, że zaprezentowana w rozdziale 3.2.5 metoda liczenia pierwiastka kwadratowego $\sqrt{a}$ jest w istocie metodą Newtona zastosowaną do równania

$$x^2 - a = 0.$$

11.8 Ekstrema lokalne, a różniczkowalność.

11.8.1 Maksima i minima lokalne

Pochodna funkcji przydaje się przy poszukiwaniu lokalnych minimów i maksimów funkcji.

Definicja 11.8.1 Mówimy, że funkcja $f : (a,b) \to \mathbb{R}$ ma w $x_0 \in (a,b)$ *minimum lokalne* jeżeli istnieje otoczenie $U \subset (a,b)$ punktu x_0 takie, że dla wszystkich $x \in U$ zachodzi $f(x) \geqslant f(x_0)$. Podobnie mówimy, że f ma w $x_0 \in (a,b)$ *maksimum lokalne* jeżeli istnieje otoczenie $U \subset (a,b)$ punktu x_0 takie, że dla wszystkich $x \in U$ zachodzi $f(x) \leqslant f(x_0)$. Mówimy, że f ma w $x_0 \in (a,b)$ *ekstremum lokalne* jeśli f ma w x_0 minimum lokalne lub maksimum lokalne.

Twierdzenie 11.8.2 (o ekstremum lokalnym funkcji różniczkowalnej) *Jeśli $f : (a,b) \to \mathbb{R}$ jest różniczkowalna w $x_0 \in (a,b)$ i ma w x_0 ekstremum lokalne, to $f'(x_0) = 0$.*

Dowód: Załóżmy, że f ma w x_0 maksimum lokalne. Zatem istnieje $\delta > 0$ taka, że $(x_0-\delta, x_0+\delta) \subset (a,b)$ oraz $f(x) \leqslant f(x_0)$ dla $x \in (x_0-\delta, x_0+\delta)$. W szczególności dla $h \in (-\delta,\delta)$ mamy $f(x_0+h) - f(x_0) \leqslant 0$. Zatem dla $h \in (-\delta, 0)$ jest

$$\frac{f(x_0+h)-f(x_0)}{h} \geqslant 0,$$

skąd wnosimy, że pochodna lewostronna f w x_0 jest nieujemna. Z drugiej strony dla $h \in (0,\delta)$ jest

$$\frac{f(x_0+h)-f(x_0)}{h} \leqslant 0,$$

skąd wnosimy, że pochodna prawostronna f w x_0 jest niedodatnia. Ale pochodna f w x_0 istnieje, więc jest równa tak pochodnej lewostronnej jak i prawostronnej. Zatem $f'(x_0) = 0$. □

11.8.2 Twierdzenie Rolle'a.

Francuski matematyk Michel Rolle (1652-1719) uważany jest za autora pierwszego w miarę współczesnego dowodu następującego twierdzenia.

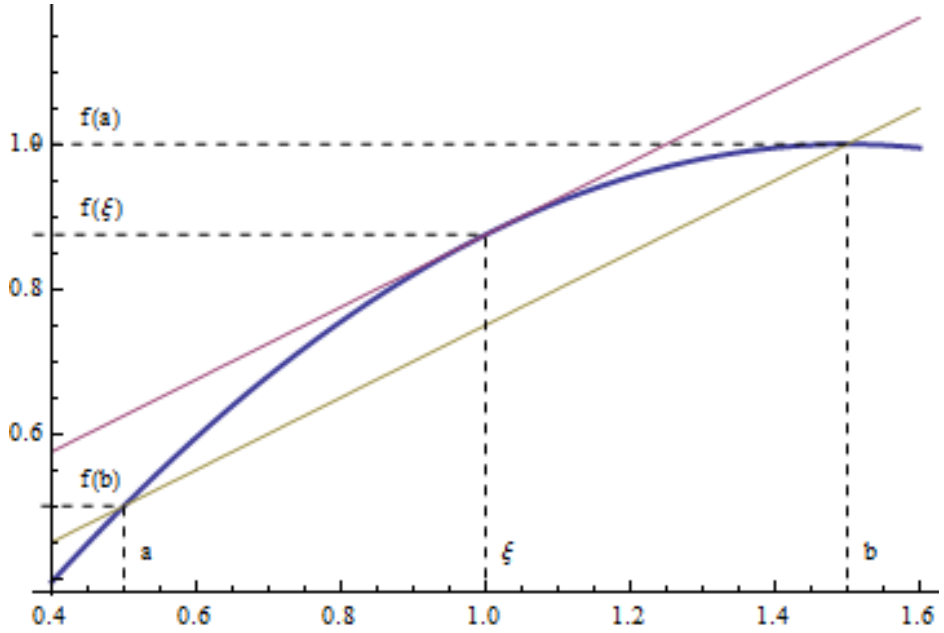
Twierdzenie 11.8.3 (Rolle'a) *Załóżmy, że $f : [a,b] \to \mathbb{R}$ jest ciągła oraz, że jest różniczkowalna dla $x \in (a,b)$. Jeśli $f(a) = f(b)$ to istnieje $\xi \in (a,b)$ takie, że $f'(\xi) = 0$.*

Dowód: Funkcja f, jako funkcja ciągła na przedziale zwartym, osiąga wartość największą oraz najmniejszą. Jeśli tak wartość największa jak i najmniejsza są osiągane na końcach przedziału $[a,b]$, to założenie $f(a) = f(b)$ implikuje, że f jest stała, a więc ma pochodną zero w każdym punkcie przedziału (a,b). Załóżmy zatem, że wartość największa lub wartość najmniejsza jest osiągnięta w pewnym punkcie $\xi \in (a,b)$. Wtedy f ma w ξ ekstremum lokalne, zatem z twierdzenia o ekstremum lokalnym funkcji różniczkowalnej wynika, że $f'(\xi) = 0$. □

11.9 Twierdzenie o wartości średniej

11.9.1 Uogólnione twierdzenie o wartości średniej

W twierdzeniu Rolle'a sieczna do wykresu poprowadzona przez punkty $(a, f(a))$ i $(b, f(b))$ jest pozioma. Styczna do wykresu w punkcie osiągania ekstremum też jest pozioma. Zatem można stwierdzić, że istnieje punkt wewnątrz przedziału (a,b) taki, że styczna do wykresu w tym punkcie jest równoległa do siecznej



Rysunek 11.5: Sieczna i równoległa do niej styczna.

przez punkty $(a, f(a))$ i $(b, f(b))$. Okazuje się, że to stwierdzenie jest prawdziwe również wtedy gdy $f(a) \neq f(b)$. Fakt ten, znany jako twierdzenie o wartości średniej, ilustruje rysunek 11.5. W polskiej literaturze z twierdzeniem tym kojarzony jest francuski matematyk Joseph Louis Lagrange (1736-1813), Udowodnimy najpierw jeszcze ogólniejsze twierdzenie, pochodzące od Augustina Cauchy'ego, z którego twierdzenie o wartości średniej wyprowadzimy jako wniosek.

Twierdzenie 11.9.1 (uogólnione o wartości średniej) *Niech $f, g : [a, b] \to \mathbb{R}$ będą ciągłymi funkcjami, różniczkowalnymi w przedziale (a, b). Wtedy istnieje punkt $\xi \in (a, b)$ taki, że*

$$(f(b) - f(a))\, g'(\xi) = (g(b) - g(a))\, f'(\xi).$$

Dowód: Rozważmy funkcję $\varphi : [a, b] \to \mathbb{R}$ daną wzorem $\varphi(x) := (f(b) - f(a))\, g(x) - (g(b) - g(a))\, f(x)$. Funkcja ta jest ciągła na przedziale $[a, b]$ i różniczkowalna na przedziale (a, b). Co więcej, prostym rachunkiem sprawdzamy, że

$$\varphi(a) = f(b)g(a) - f(a)g(b) = \varphi(b),$$

zatem stosując do φ twierdzenie Rolle'a znajdziemy punkt $\xi \in (a, b)$ taki, że $\varphi'(\xi) = 0$. Ale $\varphi'(\xi) = (f(b) - f(a))\, g'(\xi) - (g(b) - g(a))\, f'(\xi)$. Zatem $(f(b) - f(a))\, g'(\xi) - (g(b) - g(a))\, f'(\xi) = 0$, co dowodzi naszej tezy. □

11.9.2 Twierdzenie Lagrange'a o wartości średniej

Przyjmując w twierdzeniu 11.9.1 za funkcję g identyczność otrzymujemy jako wniosek następujące klasyczne twierdzenie o wartości średniej.

Twierdzenie 11.9.2 (Lagrange'a o wartości średniej) *Niech $f : [a, b] \to \mathbb{R}$ będzie funkcją ciągłą, różniczkowalną w przedziale (a, b). Wtedy istnieje punkt $\xi \in (a, b)$ taki, że*

$$(f(b) - f(a)) = (b - a)\, f'(\xi).$$

□

11.9.3 Twierdzenie o przyrostach skończonych

Twierdzenia o wartości średniej nie dają się przenieść w dosłownej formie na przypadek funkcji o wartościach wektorowych, na przykład zespolonych, jak pokazuje następujący przykład.

Przykład 11.9.3 Rozważmy funkcję

$$f : [0,1] \ni t \mapsto e^{2t\pi i} \in \mathbb{C}.$$

Mamy $f[0] = f[1] = 1$, zatem gdyby twierdzenie o wartości średniej zachodziło, to istniałby punkt $\xi \in (0,1)$ taki, że $f'(\xi) = 0$. Taki punkt jednak nie istnieje, bo z twierdzenia 11.5.2 wiemy, że pochodną jest $f'(t) = 2\pi i e^{2t\pi i} \neq 0$.

Mamy jednak pokrewne twierdzenie.

Twierdzenie 11.9.4 (o przyrostach skończonych) *Niech $f : [a,b] \to \mathbb{R}^m$ będzie funkcją ciągłą, różniczkowalną na (a,b). Wtedy istnieje $\xi \in (a,b)$ takie, że*

$$||f(b) - f(a)|| \leqslant ||f'(\xi)||(b-a).$$

Dowód: Oznaczmy $z := f(b) - f(a)$. Jeśli $z = 0$, teza jest oczywista. Zatem załóżmy, że $z \neq 0$ i rozważmy funkcję

$$\varphi : [a,b] \ni t \mapsto z \cdot f(t) \in \mathbb{R}.$$

Funkcja φ jest funkcją o wartościach rzeczywistych, jest ciągła, a na podstawie ćwiczenia 11.3.6 wiemy też, że jest różniczkowalna i $\varphi'(t) = z \cdot f'(t)$. Możemy więc do niej zastosować twierdzenie o wartości średniej. Zatem istnieje $\xi \in (a,b)$ takie, że

$$\varphi(b) - \varphi(a) = \varphi'(\xi)(b-a) = (b-a)z \cdot f'(\xi).$$

Z drugiej strony

$$\varphi(b) - \varphi(a) = z \cdot (f(b) - f(a)) = z \cdot z = ||z||^2.$$

Zatem z nierówności Schwartza dostajemy

$$||z||^2 = (b-a)z \cdot f'(\xi) = |(b-a)z \cdot f'(\xi)| \leqslant (b-a)||z||||f'(\xi)||.$$

Dzieląc powyższą nierówność obustronnie przez $||z||$ otrzymujemy tezę. □

11.10 Zastosowania twierdzenia o wartości średniej

11.10.1 Pochodna, a monotoniczność

Monotoniczność funkcji różniczkowalnej znajduje odwzierciedlenie w znaku jej pochodnej. Dokładniej, mamy następujące twierdzenie.

Twierdzenie 11.10.1 *Niech $f : (a,b) \to \mathbb{R}$ będzie funkcją różniczkowalną. Mają miejsce następujące równoważności*

(i) $f' \geqslant 0$ *wtedy i tylko wtedy gdy f jest słabo rosnąca,*

(ii) $f' > 0$ *wtedy i tylko wtedy gdy f jest silnie rosnąca,*

(iii) $f' \leqslant 0$ *wtedy i tylko wtedy gdy f jest słabo malejąca,*

(iv) $f' < 0$ *wtedy i tylko wtedy gdy f jest silnie malejąca.*

Dowód: Załóżmy najpierw, że $f' \geqslant 0$. Niech $x_1, x_2 \in (a,b)$ oraz $x_1 < x_2$. Z twierdzenia o wartości średniej wiemy, że $f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1)$ dla pewnego $\xi \in (x_1, x_2)$. Ale z założenia $f'(\xi) \geqslant 0$, skąd $f(x_2) - f(x_1) \geqslant 0$, a to dowodzi, że f jest słabo rosnąca. Z kolei jeśli f jest słabo rosnąca, to iloraz różnicowy w każdym punkcie przedziału (a,b) jest nieujemny. Zatem i pochodna jest nieujemna, jako granica ilorazu różnicowego. To dowodzi równoważności (i). Dowód pozostałych równoważności jest podobny, więc zostawiamy go jako ćwiczenie. □

11.10.2 Twierdzenie o wartościach pośrednich dla pochodnej

Widzieliśmy w przykładzie 11.6.2, że istnienie pochodnej w każdym punkcie dziedziny funkcji f nie gwarantuje ciągłości pochodnej. Tym niemniej, pochodna, nawet jeśli jest nieciągła, ma własność przyjmowania wartości pośrednich, podobnie jak funkcja ciągła.

Twierdzenie 11.10.2 *Niech $f:(a,b)\to\mathbb{R}$ będzie funkcją różniczkowalną. Niech $c,d\in(a,b)$ będą takie, że $c<d$ oraz $f'(c)<f'(d)$. Wtedy dla każdego $\lambda\in(f'(c),f'(d))$ istnieje $\xi\in(c,d)$ takie, że $f'(\xi)=\lambda$.*

Dowód: Rozważmy funkcję $g:(a,b)\to\mathbb{R}$ daną wzorem $g(t):=f(t)-\lambda t$. Jest ona ciągła różniczkowalna oraz $g'(c)=f'(c)-\lambda<0$ i $g'(d)=f'(d)-\lambda>0$. Zatem $g(c)>g(c+h_1)$ dla pewnego $h_1\in(0,d-c)$, bo w przeciwnym razie iloraz różnicowy $\frac{g(c+h)-g(c)}{h}>=0$ dla $h>0$ co prowadzi do sprzeczności z $g'(c)<0$. Podobnie stwierdzamy, że $g(d)>g(d-h_2)$ dla pewnego $h_2\in(0,d-c)$. Wynika stąd, że maksium, które g na przedziale $[c,d]$ musi osiągać na mocy tw. 7.2.14 , nie może być osiągnięte na końcach tego przedziału lecz w pewnym punkcie $\xi\in(c,d)$. Zatem jest to maksimum lokalne. Z tw. 11.8.2 $g'(\xi)=f'(\xi)-\lambda=0$, skąd $f'(\xi)=\lambda$. □

Z udowodnionego twierdzenia wynika następujący ważny wniosek.

Wniosek 11.10.3 *Niech $f:(a,b)\to\mathbb{R}$ będzie funkcją różniczkowalną. Jeśli $f'(x)\neq0$ dla wszystkich $x\in(a,b)$, to f' przyjmuje wyłącznie wartości dodatnie lub wyłącznie wartości ujemne.*

Ćwiczenie 11.10.4 Wywnioskuj z tw. 11.10.2, że jeśli $f:(a,b)\to\mathbb{R}$ jest funkcją różniczkowalną, to f' nie posiada punktów nieciągłości pierwszego rodzaju.

11.10.3 Reguła de l'Hospitala

Jeśli $f:\mathbb{R}\to\mathbb{R}$ zeruje się w zerze, to

$$\frac{f(x)}{x}=\frac{f(x)-f(0)}{x-0}$$

może być traktowane jako przybliżenie $f'(0)$, oczywiście przy założeniu, że pochodna ta istnieje. Zatem jeśli mamy dwie takie funkcje, $f,g:\mathbb{R}\to\mathbb{R}$ i obie zerują się w zerze, to mamy

$$\frac{f(x)}{g(x)}=\frac{\frac{f(x)}{x}}{\frac{g(x)}{x}}$$

Prawą stronę tego wyrażenia można traktować jako przybliżenie $\frac{f'(0)}{g'(0)}$. Lewa strona nie jest przybliżeniem $\frac{f(0)}{g(0)}$, bo jest to pozbawiony sensu ułamek o zerze w liczniku i mianowniku. Jest to jednak przybliżenie granicy ułamków $\frac{f(x)}{g(x)}$ przy x zmierzającym do zera o ile ta granica istnieje. Zatem, przy odpowiednich założeniach, możemy się spodziewać wzoru

$$\lim_{x\to0}\frac{f(x)}{g(x)}=\frac{f'(0)}{g'(0)}.$$

Udowodnimy teraz nawet nieco ogólniejsze twierdzenie. Z twierdzeniem tym kojarzony jest francuski matematyk Guillaume de l'Hospital (rys. 11.6), który pierwszy je opublikował. Jednak są argumenty za tym, że autorem twierdzenia jest szwajcarski matematyk Johann Bernoulli (rys. 11.6), brat Jacoba Bernoullego (rys. 4.7). Okazuje się ono bardzo pomocne przy wyznaczaniu granic ilorazów, do których nie możemy zastosować twierdzenia o granicy ilorazu. Przypomnijmy bowiem, że twierdzenie o granicy ilorazu wymaga, by mianownik miał granicę różną od zera.




Rysunek 11.6: Guillaume de l'Hospital (1661 - 1704) i Johann Bernoulli (1667 - 1748) Źródło: Wikipedia

Twierdzenie 11.10.5 (reguła de l'Hospitala). *Niech $f, g : (a, b) \to \mathbb{R}$ będą różniczkowalne oraz $g'(x) \neq 0$ dla $x \in (a, b)$. Załóżmy, że*

$$\lim_{x\to a} f(x) = 0 = \lim_{x\to a} g(x) \tag{11.12}$$

lub

$$\lim_{x\to a} g(x) \in \{-\infty, \infty\}. \tag{11.13}$$

Jeśli istnieje granica

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} \tag{11.14}$$

to istnieje granica

$$\lim_{x\to a} \frac{f(x)}{g(x)} \tag{11.15}$$

i obie te granice są sobie równe.

Twierdzenie 11.10.5 wyprowadzimy z następującego lematu

Lemat 11.10.6 *Przy założeniach twierdzenia 11.10.5 mamy nierówności*

$$\limsup_{x\to a} \frac{f(x)}{g(x)} \leqslant \limsup_{x\to a} \frac{f'(x)}{g'(x)} \tag{11.16}$$

oraz

$$\liminf_{x\to a} \frac{f(x)}{g(x)} \geqslant \liminf_{x\to a} \frac{f'(x)}{g'(x)} \tag{11.17}$$

Dowód: (twierdzenia 11.10.5) Ponieważ granica dolna jest niewiększa od granicy górnej, z (11.16) i (11.17) mamy ciąg nierówności

$$\liminf_{x\to a} \frac{f'(x)}{g'(x)} \leqslant \liminf_{x\to a} \frac{f(x)}{g(x)} \leqslant \limsup_{x\to a} \frac{f(x)}{g(x)} \leqslant \limsup_{x\to a} \frac{f'(x)}{g'(x)}. \tag{11.18}$$

W twierdzeniu 11.10.5 zakładamy istnienie granicy (11.14). Zatem stosowna granica dolna i granica górna są sobie równe. Zatem pierwsze i ostatnie wyrażenie w ciągu nierówności (11.18) są sobie równe. Jest to możliwe tylko wtedy, gdy wszystkie

wyrażenia są sobie równe. W szczególności drugie i trzecie wyrażenie są sobie równe. Wynika stąd istnienie granicy (11.15) i równość obu granic. □

Dowód: (lematu 11.10.6) Rozważmy dowód wzoru (11.16). Oznaczmy

$$A := \limsup_{x\to a} \frac{f'(x)}{g'(x)}.$$

Jeśli $A = \infty$, to wzór (11.16) oczywiście zachodzi, załóżmy więc, że $A < \infty$. Wystarczy pokazać, że

$$\forall\, q > A \quad \limsup_{x\to a} \frac{f(x)}{g(x)} \leqslant q. \tag{11.19}$$

Ustalmy zatem liczbę $q > A$. Z definicji granicy znajdziemy $c \in (a, b)$ takie, że

$$\forall\, x \in (a, c) \quad \frac{f'(x)}{g'(x)} < q. \tag{11.20}$$

Niech $x, y \in (a, c)$ oraz $x < y$. Zwężenia $f_{|[x,y]}$ i $g_{|[x,y]}$ spełniają założenia uogólonionego twierdzenia o wartości średniej (tw. 11.9.1) zatem znajdziemy $\xi \in (x, y)$ takie, że

$$(f(y) - f(x))\, g'(\xi) = (g(y) - g(x))\, f'(\xi).$$

Wykorzystując założenie o niezerowaniu się pochodnej funkcji g przepisujemy powyższą równość w postaci

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f(y) - f(x)}{g(y) - g(x)} = \frac{f'(\xi)}{g'(\xi)}.$$

Zatem z (11.20) wnosimy, że dla $x, y \in (a, c)$

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f(y) - f(x)}{g(y) - g(x)} < q. \tag{11.21}$$

Rozważmy przypadek założenia (11.12). Przechodząc ze zmienną y w (11.21) do granicy w a dostajemy

$$\frac{f(x)}{g(x)} \leqslant q$$

dla $x \in (a, c)$, skąd natychmiast wynika (11.19).

Przypadek (11.13) jest nieco bardziej zagmatwany. Dowód zaprezentujemy przy założeniu $\lim_{x\to a} g(x) = \infty$. Rozumowanie dla $\lim_{x\to a} g(x) = -\infty$ jest analogiczne. Zauważmy najpierw, że z $\lim_{x\to a} g(x) = \infty$ znajdziemy $c' \in (a, c)$ takie, że $g(x) > 0$ dla $x < c'$. Założenie o niezerowaniu się pochodnej funkcji g implikuje, że g jest funkcją monotoniczną, a w konsekwencji, ze względu na założenie $\lim_{x\to a} g(x) = \infty$, jest funkcją silnie malejącą. Zatem $g(x) - g(y) > 0$. Mnożąc (11.21) przez $g(x) - g(y)$ i przenosząc $f(y)$ na prawą stronę dostajemy

$$f(x) < f(y) + q\,(g(x) - g(y)) = qg(x) + f(y) - qg(y).$$

Dla $x \in (a, c')$ mamy $g(x) > 0$, więc dzieląc ostatnią nierówność przez $g(x)$ dostajemy

$$\frac{f(x)}{g(x)} < q + \frac{f(y)}{g(x)} - q\frac{g(y)}{g(x)}.$$

Ze względu na założenie $\lim_{x\to a} g(x) = \infty$, dla zadanego $\epsilon > 0$ i przy ustalonym $y \in (a, c)$ znajdziemy $c'' \in (a, c')$ takie, że dla $x \in (a, c'')$

$$\frac{f(y)}{g(x)} - q\frac{g(y)}{g(x)} < \epsilon.$$

Zatem z ostatnich dwóch nierówności dla $x \in (a, c'')$ i dowolnego $\epsilon > 0$ mamy

$$\frac{f(x)}{g(x)} < q + \epsilon.$$

Ze względu na dowolność $\epsilon > 0$ dla $x \in (a, c'')$ mamy

$$\frac{f(x)}{g(x)} \leqslant q,$$

skąd wnosimy (11.19) również dla przypadku założenia (11.13).

Dowód własności (11.17) jest podobny. Pozostawiamy go jako ćwiczenie. □

Regułę de l'Hospitala zaprezentowaliśmy dla granicy w lewym końcu przedziału. Oczywiście analogiczna reguła obowiązuje również dla prawego końca przedziału, a tym samym dla granicy lewostronnej, prawostronnej i zwykłej granicy.

Przykład 11.10.7 Granica $\lim_{x\to 0} \frac{\sin x}{x}$ nie daje się policzyć z twierdzenia o granicy ilorazu, bo mianownik zmierza do zera. Ponieważ jednak również licznik zmierza do zera oraz tak licznik jak i mianownik są różniczkowalne w otoczeniu zera, a mianownik ma pochodną stale równą jeden, więc możemy zastosować regułę de l'Hospitala. Dostajemy

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

11.11 Różniczkowanie funkcji odwrotnej

11.11.1 Wzór na pochodną funkcji odwrotnej

Twierdzenie 11.11.1 *Załóżmy, że $f : (a, b) \to \mathbb{R}$ jest funkcją różniczkowalną oraz $f'(x) \neq 0$ dla wszystkich $x \in (a, b)$. Wtedy f jest bijekcją przedziału (a, b) w pewien przedział (c, d). Co więcej, bijekcja odwrotna $f^{-1} : (c, d) \to (a, b)$ jest różniczkowalna oraz*

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)} \tag{11.22}$$

dla wszystkich $x \in (a, b)$.

Dowód: Z wniosku 11.10.3 wnosimy, że f' jest wszędzie dodatnia lub wszędzie ujemna, zatem z tw. 11.10.1 funkcja jest silnie rosnąca lub silnie malejąca. W szczególności jest zatem injekcją. Niech $\Gamma := \operatorname{im} f$ będzie obrazem przedziału (a, b) w funkcji f. Ponieważ zbiory spójne w $\mathbb{R}$ pokrywają się z przedziałami (wn. 7.4.4), z tw. 7.4.5 wnosimy, że Γ jest również przedziałem. Pokażemy, że przedział ten jest otwarty. Rzeczywiście, jeśli d jest lewym lub prawym końcem przedziału Γ i $d \in \Gamma$, to d jest elementem największym lub najmniejszym w Γ. Równocześnie $d = f(t)$ dla pewnego $t \in (a, b)$. Wybierzmy $t', t'' \in (a, b)$ takie, że $t' < t < t''$. Z monotoniczności f mamy $f(t') < d < f(t'')$ lub $f(t') > d > f(t'')$. Ponieważ $f(t'), f(t'') \in \Gamma$, wykluczone jest by d było elementem największym lub najmniejszym w Γ. Zatem Γ jest przedziałem otwartym. Przyjmijmy, że $\Gamma = (c, d)$.

Ustalmy $x \in (a, b)$ oraz połóżmy $y := f(x)$. Wtedy $x = f^{-1}(y)$. Niech k będzie takie, że $y + k \in (c, d)$. Połóżmy $h := f^{-1}(y + k) - f^{-1}(y)$. Wtedy $f^{-1}(y + k) = f^{-1}(y) + h = x + x$, zatem $y + k = f(x + h)$. Mamy

$$\frac{f^{-1}(y+k) - f^{-1}(y)}{k} = \frac{f^{-1}(y+k) - f^{-1}(y)}{(y+k) - y} = \frac{h}{f(x+h) - f(x)} = \frac{1}{\frac{f(x+h)-f(x)}{h}}.$$

Zauważmy, że jeśli $k \to 0$ to również $h \to 0$. Zatem przechodząc w powyższej równości z k do granicy w zerze dostajemy

$$(f^{-1})'(y) = \frac{1}{f'(x)}$$

co dowodzi różniczkowalności f^{-1} oraz równości (11.22). □

11.11.2 Pochodna funkcji logarytmicznej

Stosując twierdzenie 11.11.1 do funkcji wykładniczej dostajemy twierdzenie o różniczkowaniu funkcji logarytmicznej.

Twierdzenie 11.11.2 *Funkcja* $\ln : \mathbb{R}^+ \to \mathbb{R}$ *jest funkcją różniczkowalną oraz*

$$(\ln x)' = \frac{1}{x}. \tag{11.23}$$

Dowód: Z wniosku 11.5.3 wiemy pochodną funkcji $x \to e^x$ jest ona sama. Ponieważ funkcja wykładnicza nie przyjmuje wartości zero, możemy do niej zastosować twierdzenie 11.11.1. Funkcją odwrotną do funkcji wykładniczej jest funkcja logarytmiczna $x \to \ln x$, zatem funkcja ta jest różniczkowalna oraz

$$(\ln x)'(e^x) = \frac{1}{(e^x)'} = \frac{1}{e^x}.$$

Przyjmując $y := e^x$ dostajemy zatem

$$(\ln y)' = \frac{1}{y}$$

co pokrywa się ze wzorem (11.23), gdyż w obu wzorach występujące w nich zmienne są związane. □

11.11.3 Funkcje kołowe (odwrotne do trygonometrycznych)

Funkcje sin, cos i tg nie są bijekcjami, więc nie mają funkcji odwrotnych. Jednak po zawężeniu do odpowiednio małego przedziału stają się bijekcjami. Zwyczajowo rozważamy zawężenia $\sin_{|[-\pi/2,\pi/2]}$, $\cos_{|[0,\pi]}$, $\mathrm{tg}_{|[-\pi/2,\pi/2]}$. Pochodne tych funkcji na wnętrzach tych przedziałów są niezerowe, zatem zawężenia do wnętrz są odwracalne. Łatwo pokazać, że w konsekwencji odwracalne są również zawężenia do przedziałów domkniętych.

Definicja 11.11.3 Funkcje odwrotne do funkcji $\sin_{|[-\pi/2,\pi/2]}$, $\cos_{|[0,\pi]}$, $\mathrm{tg}_{|[-\pi/2,\pi/2]}$ oznaczamy odpowiednio $\arcsin$, $\arccos$ i $\mathrm{arc\,tg}$. Nazywamy je funkcjami *kołowymi* lub *cyklometrycznymi*.

Jako prosty wniosek z tw. 11.11.1 dostajemy następujące twierdzenie

Twierdzenie 11.11.4 *Funkcja* $\arcsin : [-1,1] \to [-\pi/2, \pi/2]$ *jest różniczkowalna dla* $x \in (-1,1)$ *oraz*

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}.$$

Funkcja $\arccos : [-1,1] \to [0,\pi]$ *jest różniczkowalna dla* $x \in (-1,1)$ *oraz*

$$(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}.$$

Funkcja $\mathrm{arc\,tg} : \mathbb{R} \to (-\pi/2, \pi/2)$ *jest różniczkowalna dla* $x \in \mathbb{R}$ *oraz*

$$(\mathrm{arc\,tg}\, x)' = \frac{1}{1+x^2}.$$

Funkcje kołowe w programie Mathematica oznaczane są `ArcSin`, `ArcCos`, `ArcTan`.

11.11.4 Pochodna funkcji potęgowej

Wiemy już, że dla $n \in \mathbb{N}_1$ mamy $(x^n)' = nx^{n-1}$. Wzór ten uogólnić można na przypadek funkcji $x \to x^\alpha$ dla dowolnego $\alpha \in \mathbb{R}$.

Twierdzenie 11.11.5 *Dla dowolnego $\alpha \in \mathbb{R}$ funkcja potęgowa*

$$\mathbb{R}^+ \ni x \mapsto x^\alpha \in \mathbb{R}^+$$

jest różniczkowalna oraz

$$(x^\alpha)' = \alpha x^{\alpha-1}.$$

Dowód: Ponieważ $x^\alpha = e^{\alpha \ln x}$, więc funkcja $x \mapsto x^\alpha$ jest złożeniem funkcji $x \mapsto \alpha \ln x$ oraz funkcji $u \mapsto e^u$. Zatem z twierdzenia o różniczkowaniu złożenia funkcji mamy

$$(x^\alpha)' = e^{\alpha \ln x} \frac{\alpha}{x} = x^\alpha \frac{\alpha}{x} = \alpha x^{\alpha-1}.$$

□

11.12 Pochodne wyższych rzędów

11.12.1 n-ta pochodna

Jeśli pochodna f' funkcji f sama ma pochodną, to nazywamy ją drugą pochodną funkcji f i oznaczamy f''. Proces brania kolejnych pochodnych możemy kontynuować. Precyzyjniej, definicję różniczkowalności i pochodnej rozszerzamy rekurencyjnie w następujący sposób.

Definicja 11.12.1 Mówimy, że funkcja $f : (a, b) \to \mathbb{R}^m$ jest *1-różniczkowalna* jeśli jest różniczkowalna. Jej pochodną f' nazywamy *pierwszą pochodną* i zapisujemy jako $f^{(1)}$. Mówimy, że funkcja f jest *n-różniczkowalna* jeśli jej $(n-1)$-sza pochodna $f^{(n-1)}$ jest różniczkowalna. Pochodną $f^{(n-1)}$ nazywamy *n-tą pochodną* funkcji f i zapisujemy jako $f^{(n)}$.

Dla małych n zamiast $f^{(n)}$ piszemy czasem f'', f''' itd. Dodatkowo wygodnie jest samą funkcję nazywać *zerową pochodną* i zapisywać jako $f^{(0)} := f$.

Definicja 11.12.2 Mówimy, że funkcja $f : (a, b) \to \mathbb{R}^m$ jest klasy C^n jeżeli jest n-różniczkowalna we wszystkich punktach przedziału (a, b) oraz $f^{(n)}$ jest ciągła. Zbiór takich funkcji oznaczamy $C^n((a, b), \mathbb{R}^m)$.

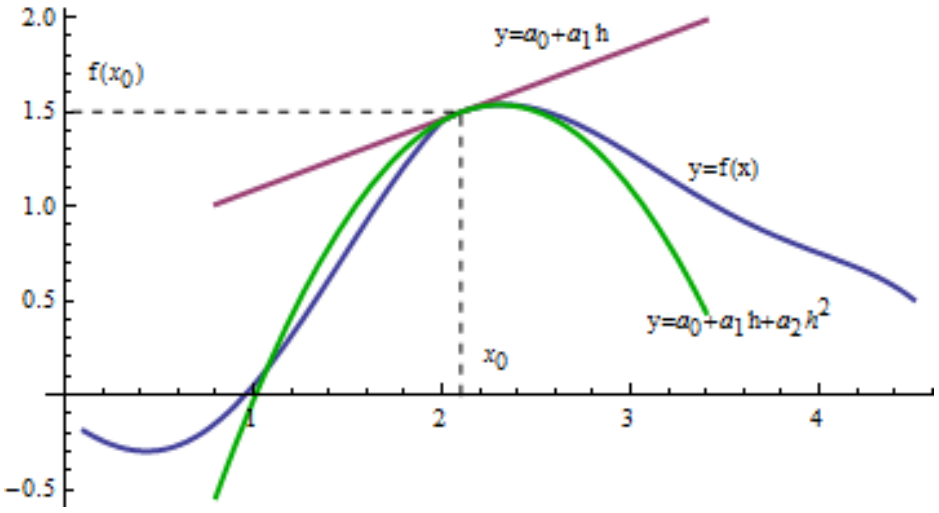
W programie Mathematica n-tą pochodną funkcji f zmiennej x obliczamy używając instrukcji

```
D[f[x],{x,n}]
```

Na przykład drugą pochodną funkcji

$$f(x) := \frac{1}{4} \ln\left(\frac{x^2-1}{x^2+1}\right)$$

obliczymy pisząc



Rysunek 11.7: Wykres funkcji f (niebieski), prostej stycznej do wykresu f w punkcie x_0 (fioletowy) i paraboli stycznej do wykresu w punkcie x_0 (zielony).

```
D[(1/4) Log[(x^2 - 1)/(x^2 + 1)], {x,2}]
```

Jako wynik program Mathematica zwraca skomplikowane wyrażenie, ale po zastosowaniu do niego `Simplify` dostajemy

$$\frac{-3x^4-1}{(x^4-1)^2}.$$

Aproksymacja liniowa poprzez różniczkę jest prosta, ale często nie jest wystarczająco dokładna. Możemy wtedy rozważać aproksymację krzywą drugiego stopnia (parabolą) zadaną wzorem $y = a_0 + a_1h + a_2h^2$ gdzie $h = x - x_0$ jest przyrostem zmiennej niezależnej (rys. 11.7). Możemy też przybliżać wielomianem n-ego stopnia postaci

$$W(h) := a_0 + a_1h + a_2h^2 + \ldots + a_nh^n.$$

Różnica

$$f(x) - W[h] = f(x_0 + h) - \left(a_0 + a_1h + a_2h^2 + \ldots + a_nh^n\right)$$

powinna być $o(h^n)$, bo oczekujemy błędu, który szybciej zmierza do zera niż wielomian aproksymacyjny. Pokażemy, że jeśli funkcja r jest n-różniczkowalna w zerze, to własność $r(h) = o(h^n)$ jest równoważna zerowaniu się pochodnych r do n-tej włącznie, tzn. warunkowi

$$r(0) = r'(0) = r''(0) = \ldots = r^{(n)}(0) = 0.$$

W odniesieniu do naszego problemu aproksymacyjnego oznacza to, że funkcja

$$h \mapsto f(x_0 + h) - W(h)$$

powinna mieć zerowe pochodne w zerze do n-tej włącznie co jest równoważne postulatowi by funkcje $h \mapsto f(x_0 + h)$ i $h \mapsto W(h)$ miały te same pochodne do n-tej włącznie. Jak zobaczymy pochodne wielomianu jest stosunkowo łatwo policzyć: $W^{(i)}(0) = i!a_i$. Zatem tyle muszą wynosić pochodne funkcji $h \mapsto f(x_0 + h)$ w zerze. Są one równe pochodnym $x \mapsto f(x)$ w x_0, zatem poszukiwana aproksymacja ma postać

$$f(x_0 + h) = f(x_0) + \frac{f'(x_0)}{1!}h + \frac{f''(x_0)}{2!}h^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}h^n + r(h)$$

gdzie $r(h) = o(h^n)$. Otrzymany w ten sposób wielomian aproksymacyjny nazywany jest wielomianem Taylora funkcji f w x_0, bo pochodzi od angielskiego matematyka Brooka Taylora (rys. 11.8).



Rysunek 11.8: Brook Taylor, 1685-1731. Źródło: Wikipedia

Twierdzenie 11.12.3 *Wielomian*

$$W : \mathbb{R} \ni x \mapsto a_0 + a_1 h + a_2 h^2 + \ldots + a_n h^n \in \mathbb{R}. \tag{11.24}$$

jest funkcją i-różniczkowalną dla wszystkich $i \in \mathbb{N}$ oraz

$$W^{(i)}(0) = i! a_i \tag{11.25}$$

gdzie przyjmujemy $a_i = 0$ dla $i > n$.

Dowód: Pokażemy indukcyjnie, że dla $k \in \mathbb{N}$ zachodzi ogólniejszy wzór

$$W^{(k)}(x) = \sum_{i=k}^{n} \frac{i!}{(i-k)!} a_i x^{i-k}. \tag{11.26}$$

Rzeczywiście, dla $k = 0$ wzór (11.26) pokrywa się z definicją (11.24) wielomianiu W. Natomiast jeśli (11.26) zachodzi dla pewnego k, to biorąc pochodną obu stron dostajemy

$$\begin{aligned} W^{(k+1)}(x) &= \left(W^{(k)}\right)'(x) = \left(k! a_k + \sum_{i=k+1}^{n} \frac{i!}{(i-k)!} a_i x^{i-k}\right)' = (k! a_k)' + \sum_{i=k+1}^{n} \frac{i!}{(i-k)!} (a_i x^{i-k})' \\ &= \sum_{i=k+1}^{n} \frac{i!}{(i-k)!} (i-k) a_i x^{i-k-1} = \sum_{i=k+1}^{n} \frac{i!}{(i-(k+1))!} a_i x^{i-(k+1)}. \end{aligned}$$

Podstawiając we wzorze (11.26) za x zero dostajemy (11.25). □

11.12.2 Wielomian Taylora

Niech $f : (a, b) \to \mathbb{R}$ będzie n-różniczkowalna w $x_0 \in (a, b)$.

Definicja 11.12.4 *Wielomianem Taylora funkcji f* w punkcie x_0 stopnia n nazywamy wielomian

$$T^n_{x_0} f : \mathbb{R} \ni h \mapsto f(x_0) + \frac{f'(x_0)}{1!}h + \frac{f''(x_0)}{2!}h^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}h^n \in \mathbb{R}.$$

Z twierdzenia 11.12.3 łatwo wyprowadzamy następujący wniosek.

Wniosek 11.12.5 *Funkcja f i jej wielomian Taylora $T^n_{x_0} f$ mają takie same pochodne do n-tej włącznie odpowiednio w zerze i w x_0, tzn. dla $i = 0, 1, 2, \ldots n$ zachodzi*

$$\left(T^n_{x_0} f\right)^{(i)}(0) = f^{(i)}(x_0).$$

□

11.12.3 n-płaskość w zerze.

Uogólnieniem pojęcia płaskości w zerze jest n-płaskość w zerze.

Definicja 11.12.6 Mówimy, że funkcja $r : \mathbb{R} \multimap \mathbb{R}$ określona w otoczeniu zera jest n-płaska w zerze, jeżeli $r(h) = o(h^n)$ przy $h \to 0$.

Lemat 11.12.7 *Jeśli $W : \mathbb{R} \to \mathbb{R}$ jest wielomianem stopnia co najwyżej n oraz*

$$\lim_{h\to 0} \frac{W(h)}{h^n} = 0 \tag{11.27}$$

to $W = 0$.

Dowód: Dowód poprowadzimy indukcyjnie względem n. Rozważmy najpierw przypadek $n = 0$. Wtedy $W(h) = a_0$ jest wtedy funkcją stałą. Z (11.27) mamy

$$0 = \frac{W(h)}{h^0} = \lim_{h\to 0} a_0 = a_0,$$

skąd $W = 0$.

Przyjmijmy z kolei, że lemat jest spełniony dla $m < n$ i rozważmy wielomian

$$W : \mathbb{R} \ni h \mapsto a_0 + a_1 h + a_2 h^2 + \ldots + a_n h^n \in \mathbb{R}$$

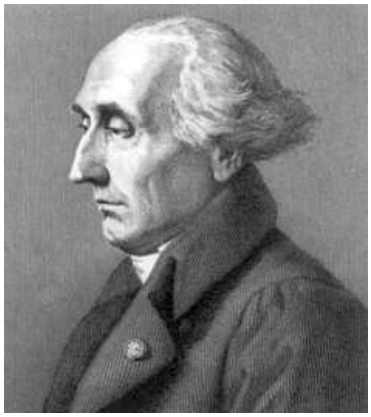
spełniający (11.27). Mamy

$$a_0 = \lim_{h\to 0} W(h) = \lim_{h\to 0} \frac{W(h)}{h^n} h^n = \lim_{h\to 0} \frac{W(h)}{h^n} \lim_{h\to 0} h^n = 0.$$

Zatem $W(h) = a_1 h + a_2 h^2 + \ldots + a_n h^n = hW_1(h)$, gdzie $W_1(h) := a_1 + a_2 h + \ldots + a_n h^{n-1}$. Ale W_1 jest wielomianem stopnia co najwyżej $(n-1)$ oraz

$$\lim_{h\to 0} \frac{W_1(h)}{h^{n-1}} = \lim_{h\to 0} \frac{hW_1(h)}{h^n} = \lim_{h\to 0} \frac{W(h)}{h^n} = 0,$$

zatem z założenia inducyjnego $W_1 = 0$, skąd $W = 0$. □





Rysunek 11.9: Joseph Louis Lagrange (1736 - 1813) i Giuseppe Peano (1858 - 1932). Źródło: Wikipedia

11.13 Wzory Taylora i szereg Taylora.

11.13.1 Wzór Taylora z resztą Peano.

Przedstawimy teraz dwa, powiązane ze sobą twierdzenia. Pierwsze uogólnia uwagę 11.2.12 na przypadek n-płaskości w zerze. Drugie precyzuje jak wielomian Talora funkcji n-różniczkowalnej przybliża samą funkcję. Pochodzą one od włoskiego matematyka Giuseppe Peano (rys. 11.9).

Twierdzenie 11.13.1 (Peano o funkcji n-płaskiej) *Niech $r : \mathbb{R} \multimap \mathbb{R}$ będzie funkcją n-różniczkowalną w zerze. Funkcja ta jest n-płaska w zerze wtedy i tylko wtedy gdy*

$$r^{(i)}(0) = 0 \text{ dla } i = 0, 1, 2, \dots n. \tag{11.28}$$

Twierdzenie 11.13.2 (wzór Taylora z resztą Peano) *Niech $f : (a, b) \to \mathbb{R}$ będzie funkcją $(n-1)$-różniczkowalną na przedziale (a, b) oraz n-różniczkowalną w $x_0 \in (a, b)$. Wtedy istnieje dokładnie jeden wielomian $W : \mathbb{R} \to \mathbb{R}$ stopnia co najwyżej n taki, że dla $h \in (a - x_0, b - x_0)$*

$$f(x_0 + h) = W(h) + r(h)$$

dla pewnej n-płaskiej w zerze funkcji $r : \mathbb{R} \multimap \mathbb{R}$. Co więcej, wielomianem tym jest $T^n_{x_0} f$, czyli wielomian Taylora funkcji f w x_0 stopnia n. W szczególności $h \mapsto f(x_0 + h) - T^n_{x_0} f(h)$ jest funkcją n-płaską w zerze.

Dowód: (twierdzenia 11.13.1 i twierdzenia 11.13.2) Najpierw pokażemy indukcyjnie, że warunek (11.28) implikuje n-płaskość w zerze funkcji r. Dla $n = 1$ wynika to wprost z uwagi 11.2.12. Załóżmy, że implikacja zachodzi dla $m < n$ oraz warunek (11.28) jest spełniony. Ponieważ $r(0) = 0$, z uwagi 11.2.4 wnosimy, że wystarczy pokazać, iż

$$\lim_{h \to 0} \frac{r(h)}{h^n} = 0. \tag{11.29}$$

Wtedy $(r')^{(i)}(0) = 0$ dla $i = 0, 1, 2, \dots n - 1$, a zatem z założenia indukcyjnego zastosowanego do funkcji r' wnosimy, że

$r'(h) = o(h^{n-1})$ przy $h \to 0$. Zatem również z uwagi 11.2.4 wynika, iż

$$\lim_{h\to 0} \frac{r'(h)}{h^{n-1}} = 0,$$

a stąd, przy wykorzystaniu reguły de l'Hospitala

$$\lim_{h\to 0} \frac{r(h)}{h^n} = \lim_{h\to 0} \frac{r'(h)}{nh^{n-1}} = \frac{1}{n} \lim_{h\to 0} \frac{r'(h)}{h^{n-1}} = 0,$$

co dowodzi (11.29).

Rozważmy z kolei funkcję f spełniającą założenia twierdzenia 11.13.2. Niech $r(h) := f(x_0 + h) - T^n_{x_0} f(h)$. Wtedy $r^{(i)}(0) = 0$ dla $i = 0, 1, 2, \ldots n$. Zatem na mocy już udowodnionej części twierdzenia 11.13.1 funkcja r jest n-płaska w zerze. Dowodzi to, że za postulowany w twierdzeniu 11.13.2 wielomian W można wziąć wielomian Taylora $T^n_{x_0} f$. W szczególności zatem taki wielomian W istnieje, choć nie możemy jeszcze twierdzić, że jest wyznaczony jednoznacznie.

Aby pokazać jednoznaczność załóżmy, że istnieją wielomiany W_1 i W_2 stopnia co najwyżej n oraz n-płaskie w zerze funkcje r_1 i r_2 takie, że

$$f(x_0 + h) = W_i(h) + r_i(h) \text{ przy } h \to 0.$$

W szczególności

$$W_1(h) + r_1(h) = W_2(h) + r_2(h).$$

Wielomian $W(h) := W_1(h) - W_2(h)$ spełnia zatem $W(h) = r_2(h) - r_1(h)$. Z wniosku 11.2.6 wynika, że $W(h) = o(h^n)$ przy $h \to 0$. W szczególności z uwagi 11.2.4 wnosimy, że

$$\lim_{h\to 0} \frac{W(h)}{h^n} = 0,$$

a więc lemat 11.12.7 implikuje, że $W = 0$.

Pozostaje jeszcze uzupełnić dowód twierdzenia 11.13.1. Niech $r : \mathbb{R} \multimap \mathbb{R}$ będzie funkcją n-różniczkowalną i n-płaską w zerze. Z udowodnionej już części twierdzenia 11.13.2 wiemy, że $h \mapsto r(h) - T^n_0 r(h)$ jest funkcją n-płaską w zerze. Zatem z wniosku 11.2.6 wnosimy, że wielomia $T^n_0 r = r - (r - T^n_0 r)$ jest n-płaska w zerze. Z uwagi 11.2.4 oraz lematu 11.12.7 wnosimy, że $T^n_0 r =$, a stąd, że zachodzi (11.28).

□

11.13.2 Wzór Taylora z resztą Lagrange'a

Szczególnie interesujący jest przypadek, gdy funkcja $f : (a, b) \to \mathrm{R}$ jest n-różniczkowalna dla każdego $n \in \mathbb{N}$. Wtedy możemy zbudować wielomian Taylora $T^n_{x_0} f$ dowolnie wysokiego stopnia. Kolejne wielomiany Taylora tworzą więc sumy częściowe szeregu

$$\sum_{i=0}^{\infty} \frac{f^{(i)}(x_0)}{i!} h^i. \tag{11.30}$$

Naturalne jest zatem pytanie czy i kiedy ten szereg jest zbieżny do $f(x_0 + h)$. W tym celu trzeba zbadać kiedy

$$r_n(h) := f(x_0 + h) - \sum_{i=0}^{n} \frac{f^{(i)}(x_0)}{i!} h^i = f(x_0 + h) - T^n_{x_0} f(h),$$

zwane n-tą resztą, zmierza do zera przy n zmierzającym do nieskończoności. W pierwszej chwili może się wydawać, że wzór Taylor z resztą Peano pozwoli odpowiedzieć na to pytanie, bo ze wzoru tego wiemy, że

$$r_n(h) = f(x_0 + h) - T^n_{x_0} f(h) = o(h^n) \text{ przy } h \to 0,$$

zatem na podstawie wniosku 11.2.5

$$r_n(h) = f(x_0 + h) - T^n_{x_0} f(h) = \varphi(h)h^n \tag{11.31}$$

dla pewnej ciągłej i zerującej się w zerze funkcji φ. Ponieważ dla $|h| < 1$ wyrażenie h^n zmierza do zera przy n zmierzającym do nieskończoności, może się wydawać, że dowodzi to iż n-ta reszta $r_n(h)$ zmierza do zera. Problem w tym, że funkcja φ we wzorze (11.31) zależy od n, zatem wzór ten powinien być zapisany jako

$$r_n(h) = f(x_0 + h) - T^n_{x_0} f(h) = \varphi_n(h)h^n.$$

Widać teraz, że chociaż h^n zmierza do zera przy n dążącym do nieskończoności, to o granicy $\lim_{n\to\infty} \varphi_n(h)$ nic nie potrafimy powiedzieć i nasze rozumowanie się rozsypuje. Aby je uratować musimy coś więcej wiedzieć o reszcie $r_n(h)$.

Niech $f : (c, d) \to \mathbb{R}$ będzie funcją $(n + 1)$-różniczkowalną. Ustalmy $a, b \in (c, d)$ takie, że $a < b$.

Zaczniemy od lematu, który uogólnia twierdzenie Rolle'a (tw. 11.8.3).

Lemat 11.13.3 *Jeśli f jest n-płaska w a oraz $f(b) = 0$ to dla każdego $i = 1, 2, \ldots, n + 1$ istnieje $\xi_i \in (a, b)$ takie, że $f^{(i)}(\xi_i) = 0$.*

Dowód: Z twierdzenia Peano o funkcji płaskiej wnosimy, że $f^{(i)}(a) = 0$ dla $i = 1, 2, \ldots n$. Dowód poprowadzimy indukcyjnie względem n. Dla $n = 0$ lemat wynika wprost z twierdzenia Rolle'a (tw. 11.8.3). Załóżmy zatem, że twierdzenie zachodzi dla $m < n$. Stosując je dla $m = n$ otrzymujemy istnienie potrzebnych ξ_i dla $i = 1, 2, \ldots, n$. W szczególności $f^{(n)}(\xi_n) = 0$. Ponieważ również $f^{(n)}(a) = 0$, więc stosując twierdzenie Rolle'a do $f^{(n)}$ na przedziale $[a, \xi_n]$ znajdziemy $\xi_{n+1} \in (a, \xi_n)$ takie, że $\left(f^{(n)}\right)'(\xi_{n+1}) = 0$. Zatem $f^{(n+1)}(\xi_{n+1}) = 0$. □

Uogólnimy teraz twierdzenie o wartości średniej.

Twierdzenie 11.13.4 *Jeśli f jest n-płaska w a to istnieje $\xi \in (a, b)$ takie, że*

$$f(b) = \frac{f^{(n+1)}(\xi)}{(n + 1)!}(b - a)^{n+1}.$$

Dowód: Połóżmy $M := \frac{f(b)}{(b-a)^{n+1}}$ i rozważmy funkcję $g : t \mapsto f(t) - M(t - a)^{n+1}$. Prostym rachunkiem sprawdzamy, że $g(a) = g(b) = 0$ oraz $g^{(i)}(a) = 0$ dla $i = 1, 2, \ldots n$. Zatem z twierdzenia Peano o funkcji płaskiej g spełnia założenia lematu 11.13.3. Istnieje więc $\xi \in (a, b)$ takie, że $g^{(n+1)}(\xi) = 0$. Ale $g^{(n+1)}(t) = f^{(n+1)}(t) - (n + 1)!M$, więc $f^{(n+1)}(\xi) = (n + 1)!M$, skąd wynika teza. □

Jesteśmy teraz gotowi podać wzór Taylora ze zmodyfikowaną resztą, zwaną resztą Lagrange'a. Twierdzenie to pochodzi od francuskiego matematyka włoskiego pochodzenia, J.L. Lagrange'a (rys. 11.9).

Twierdzenie 11.13.5 (wzór Taylora z resztą Lagrange'a) *Niech $f : (c, d) \to \mathbb{R}$ będzie funcją $n + 1$-różniczkowalną. Niech $a, b \in (c, d)$ spełniają $a < b$. Wtedy istnieje $\xi \in (a, b)$ takie, że*

$$f(b) = T^n_a f(b - a) + \frac{f^{(n+1)}(\xi)}{(n + 1)!}(b - a)^{n+1}. \tag{11.32}$$

Przyjmując oznaczenie $h := b - a$ oraz rozpisując wielomian Taylora możemy też powyższą tezę zapisać jako

$$f(a + h) = f(a) + \frac{f'(a)}{1!}h + \frac{f''(a)}{2!}h^2 + \ldots + \frac{f^{(n)}(a)}{n!}h^n + \frac{f^{(n+1)}(\xi)}{(n + 1)!}h^{n+1}.$$

Dowód: Rozważmy funkcję $r : x \mapsto f(x) - T^n_a f(x - a)$. Funkcja ta jest n-płaska w a, bo jej pochodne w a do n-tej włącznie się zerują. Zatem z twierdzenia 11.13.4 istnieje $\xi \in (a, b)$ takie, że

$$r(b) = \frac{r^{(n+1)}(\xi)}{(n + 1)!}(b - a)^{n+1}.$$

Ale $r^{(n+1)}(x) = f^{(n+1)}(x)$, bo $(n+1)$-sza pochodna wielomianu stopnia co najwyżej n zeruje się, Zatem

$$\frac{f^{(n+1)}(\xi)}{(n+1)!}(b-a)^{n+1} = r(b) = f(b) - T_a^n f(b-a),$$

skąd wynika (11.32). □

W programie Mathematica n-ty wielomian Taylora funkcji f w punkcie u wyliczony dla argumentu $(x-u)$ uzyskamy instrukcją

```
Series[f[x],{x,u,n}]
```

Na przykład pisząc

```
Series[Exp[x], {x, 0, 6}]
```

otrzymujemy

$$1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \frac{x^6}{720} + O\left[x\right]^7$$

Wyrażenie $O\left[x\right]^7$ jest tutaj niezbyt fortunnym, choć dającym się wytłumaczyć, sposobem zapisania reszty Lagrange'a.

11.13.3 Punkty przegięcia

Definicja 11.13.6 Mówimy, że funkcja $f : (a,b) \to \mathbb{R}$ ma punkt przegięcia w $x_0 \in (a,b)$ jeżeli f jest różniczkowalna w x_0 oraz istnieje $\delta > 0$ taka, że dla $0 < |h| < \delta$ funkcja

$$h \mapsto \operatorname{sgn}\left[\Delta_{x_0} f(h) - f'(x_0)h\right] \operatorname{sgn} h$$

jest stała, to znaczy stale równa jeden lub stale równa minus jeden.

Twierdzenie 11.13.7 *Niech $f : (a,b) \to \mathbb{R}$ będzie klasy C^n oraz niech $f^{(i)}(x_0) = 0$ dla $i = 1, 2, \ldots n-1$ i $f^{(n)}(x_0) \neq 0$. Jeśli n jest parzyste, to f ma w x_0 maksimum lokalne gdy $f^{(n)}(x_0) < 0$ bądź minimum lokalne gdy $f^{(n)}(x_0) > 0$. Natomiast jeśli n jest nieparzyste, to f ma w x_0 punkt przegięcia.*

Dowód: Ponieważ f jest klasy C^n, więc $f^{(n)}$ jest funkcją ciągłą, zatem w pewnym otoczeniu punktu x_0 ma stały znak. Wybierzmy $\delta > 0$ tak, by $\operatorname{sgn} f^{(n)}(x)$ było stałe dla $0 < |x - x_0| < \delta$. Załóżmy, że $0 < |h| < \delta$. Ze wzoru Taylora z resztą Lagrange'a dla pewnego $\theta \in (0,1)$

$$f(x_0 + h) - f(x_0) = \frac{f^{(n)}(x_0 + \theta h)}{n!} h^n,$$

skąd otrzymujemy tezę, ponieważ znak wyrażenia $f^{(n)}(x_0 + \theta h)$ jest stały. □

11.13.4 Szereg Taylora

Dysponujemy teraz narzędziami do badania zbieżności szeregu (11.30). Zacznijmy od formalnej definicji.

Definicja 11.13.8 Mówimy, że funkcja $f : (a,b) \to \mathbb{R}$ jest klasy C^∞ jeżeli jest klasy C^n dla każdego $n \in \mathbb{N}$.

Definicja 11.13.9 Niech $f : (a, b) \to \mathbb{R}$ będzie funkcją klasy C^∞ i niech $x_0 \in (a, b)$. Szereg formalny (bez pytania o zbieżność)

$$\sum_{i=0}^{\infty} \frac{f^{(i)}(x_0)}{i!} h^i$$

nazywamy *szeregiem Taylora* funkcji f w punkcie x_0 dla argumentu h i oznaczamy $T_{x_0} f(h)$.

Przykład 11.13.10 Pochodną rzeczywistej funkcji wykładniczej $\exp : \mathbb{R} \ni x \to e^x \in \mathbb{R}$ jest ona sama. Zatem wszystkie jej pochodne są jej równe. W szczególności funkcja wykładnicza jest klasy C^∞ oraz $\exp^{(n)}(0) = \exp(0) = 1$. Zatem łatwo sprawdzamy, że szeregiem Taylora rzeczywistej funkcji wykładniczej w zerze jest

$$\sum_{i=0}^{\infty} \frac{1}{i!} h^i,$$

a więc ten sam szereg, który posłużył nam do zdefiniowania funkcji wykładniczej. O szeregu tym wiemy, że jest zbieżny do $\exp(h)$.

Wzór Taylora z resztą Lagrange'a podpowiada nam jak zagwarantować zbieżność szeregu Taylora.

Twierdzenie 11.13.11 *Niech $f : (a, b) \to \mathbb{R}$ będzie klasy C^∞ i niech $x_0 \in (a, b)$. Jeżeli istnieją stałe $\epsilon > 0$ oraz $M > 0$ takie, że dla wszystkich $n \in \mathbb{N}$*

$$|f^{(n)}(x)| \leqslant M \text{ dla } x \in (x_0 - \epsilon, x_0 + \epsilon),$$

to dla $|h| < \epsilon$ szereg Taylora $T_{x_0} f(h)$ jest zbieżny do $f(x_0 + h)$, co możemy też zapisać w postaci

$$f(x_0 + h) = \sum_{i=0}^{\infty} \frac{f^{(i)}(x_0)}{i!} h^i.$$

Dowód: Niech $|h| < \epsilon$ i niech $s_n := T^n_{x_0} f(h)$ oraz $r_n := f(a+h) - s_n$. Ze wzoru Taylora z resztą Lagrange'a (tw. 11.13.5) dla pewnego $\xi \in (a, b)$ mamy

$$|r_n| = \left| \frac{f^{(n+1)}(\xi)}{n!} h^{n+1} \right| \leqslant \frac{M}{n!} |h|^{n+1}.$$

Prawa strona powyższej nierówności zmierza do zera przy n dążącym do nieskończoności. Zatem $\lim_{n\to\infty} r_n = 0$, skąd $\lim_{n\to\infty} s_n = f(a + h)$. □

Nie w każdej sytuacji powyższe twierdzenie działa. Czasami, tak jak w poniższym przykładzie, musimy sobie radzić na piechotę.

Przykład 11.13.12 Skonstruujemy szereg Taylora funkcji ln w otoczeniu jedynki. Indukcyjnie sprawdzamy, że dla $n \geqslant 1$

$$\ln^{(n)}(x) = (-1)^{n-1}\frac{(n-1)!}{x^n}.$$

Zatem $\ln^{(n)}(1) = (-1)^n(n-1)!$. Stąd

$$(T_1 \ln)(h) = \ln 1 + \sum_{i=1}^{\infty} \frac{(-1)^{i-1}(i-1)!}{i!}h^i = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{i}h^i.$$

Ze wzoru Taylora z resztą Lagrange'a mamy dla $|h| < 1$ i pewnego ξ pomiędzy 1 i $1+h$

$$\left|\ln(1+h) - \sum_{i=1}^{n} \frac{(-1)^{i-1}}{i}h^i\right| = \left|\frac{(-1)^n}{(n+1)(1+\xi)^{n+1}}h^{n+1}\right| \leqslant \frac{1}{n+1}\frac{|h|^{n+1}}{(1-|\xi|)^{n+1}} \leqslant \frac{1}{n+1}\frac{1}{(1-|h|)^{n+1}}.$$

Ponieważ prawa strona otrzymanej nierówności zmierza do zera przy n zmierzającym do nieskończoności otrzymujemy

$$\lim_{n\to\infty}\left|\ln(1+h) - \sum_{i=1}^{n} \frac{(-1)^{i-1}}{i}h^i\right| = 0$$

Zatem dla $|h| < 1$ szereg $T_1 \ln(h)$ jest zbieżny oraz

$$\ln(1+h) = \frac{h}{1} - \frac{h^2}{2} + \frac{h^3}{3} - \dots.$$

Rozdział 12

Całka oznaczona

12.1 Całka Riemanna funkcji jednej zmiennej

12.1.1 Pole obszaru.

Gdy potrzebujemy pokryć pewien obszar (na przykład plac lub podłogę łazienki) kwadratowymi płytkami o boku jednostkowej długości (na przykład 10cm), musimy umieć policzyć ile płytek do tego celu potrzebujemy. Zakładając, że obszar ma kształt prostokąta o bokach długości $m, n \in \mathbb{N}$ łatwo stwierdzamy, że potrzebujemy mn płytek. Jeśli boki prostokąta nie są liczbami całkowitymi lub obszar ma kształt inny niż prostokątny, sprawa robi się bardziej skommplikowana, bo bez cięcia płytek sobie nie poradzimy. Zawsze jednak możemy zapytać ile maksymalnie płytek daje się ułożyć w granicach obszaru tak, by nie zachodziły one na siebie, a także ile minimalnie potrzeba płytek, by pokryć cały obszar (rys. 12.1). Te dwie liczby stanowią oszacowanie od dołu i od góry tego co intuicyjnie określamy jako powierzchnię obszaru. Lepsze przybliżenia powierzchni dostaniemy posługując się płytkami o mniejszym boku. Kontynuując proces szacowania powierzchni przy użyciu płytek o coraz krótszych bokach otrzymujemy dwa ciągi szacujące powierzchnię obszaru: jeden rosnący i szacujący powierzchnię od dołu, drugi malejący i szacujący powierzchnię od góry. Wiemy, że takie ciągi, jako monotoniczne i ograniczone, mają granice. Intuicja podpowiada nam, że oba te ciągi mają tę samą granicę i jest nią pole powierzchni rozważanego obszaru. Intuicja jest tu zawodna, bo można podać przykład zbiorów, dla których granice te są różne, co oznacza, że intuicyjne pojęcie powierzchni obszaru nie zawsze ma sens.

Do zagadnień związanych z definicją pola powierzchni podzbioru $\mathbb{R}^2$ (ogólniej miary podzbioru $\mathbb{R}^d$) jeszcze wrócimy, a na razie ograniczymy się do szczególnego przypadku obszaru postaci

$$\{\,(x,y) \in \mathbb{R}^2 \mid a \leqslant x \leqslant b,\ 0 \leqslant y \leqslant f(x)\,\},$$

gdzie $a < b$ są liczbami rzeczywistymi, a $f : [a,b] \to \mathbb{R}^*$ jest pewną funkcją. Pole powierzchni takiego obszaru można oszacować dzieląc przedział $[a,b]$ na mniejsze przedziały i sumując pola prostokątów zbudowanych nad tymi przedziałami tak, by mieściły się pod wykresem funkcji lub by pokrywały obszar. Otrzymujemy w ten sposób odpowiednio przybliżenie pola powierzchini od dołu lub od góry (rys. 12.2).

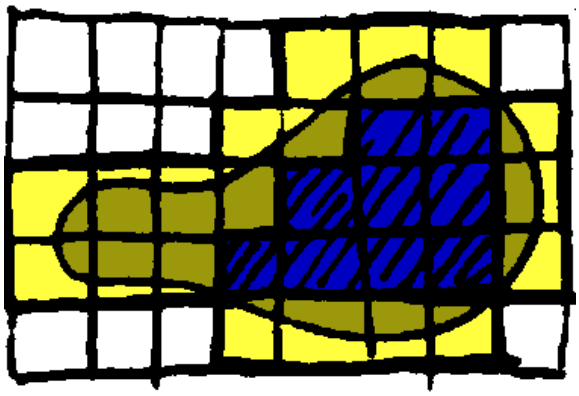
12.1.2 Przedziały w $\mathbb{R}$ i ich podziały.

W niniejszym rozdziale rozważać będziemy przedziały zwarte w $\mathbb{R}$, a więc przedziały postaci $[a,b]$, gdzie $a, b \in \mathbb{R}$ i $a < b$. Zbiór takich przedziałów oznaczać będziemy $\mathrm{Przedz}(\mathbb{R})$. Liczby a oraz b nazywamy *końcami* przedziału $[a,b]$. *Długością* przedziału $I = [a,b] \in \mathrm{Przedz}(\mathbb{R})$ nazywamy liczbę $b - a$. Jest to liczba dodatnia, którą oznaczać będziemy $\operatorname{len} I$.

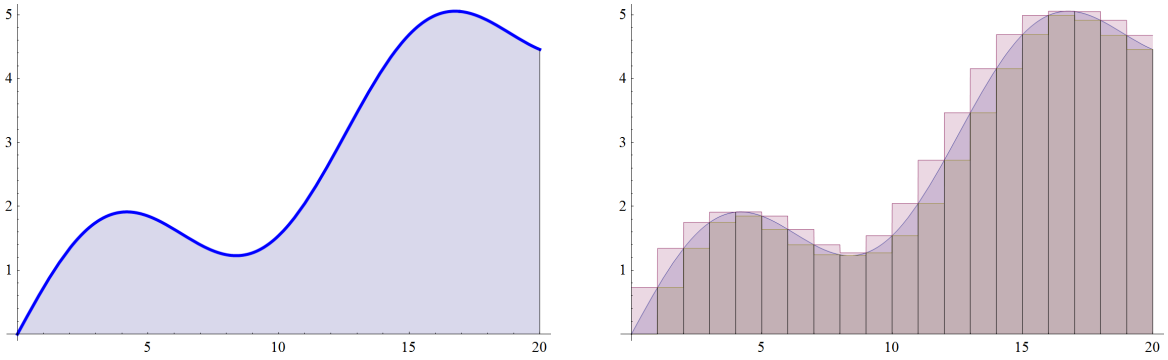
Mówimy, że dwa przedziały $I, J \in \mathrm{Przedz}(\mathbb{R})$ są *rozgraniczone* jeżeli mają rozłączne wnętrza. Skończony zbiór przedziałów $\mathcal{P} \subset \mathrm{Przedz}(\mathbb{R})$ nazywamy *podziałem* przedziału $I \in \mathrm{Przedz}(\mathbb{R})$, jeżeli dowolne dwa różne elementy $\mathcal{P}$ są rozgraniczone oraz $I = \bigcup \mathcal{P}$. *Średnicą podziału $\mathcal{P}$* nazywamy liczbę

$$\delta(\mathcal{P}) := \max\{\,\operatorname{len} P \mid P \in \mathcal{P}\,\}.$$

Zbiór wszystkich podziałów przedziału I oznaczać będziemy $\mathrm{Podz}(I)$.



Rysunek 12.1: Przybliżanie pola figury poprzez zliczanie kwadratów wpisanych w figurę i pokrywających figurę w sieci kwadratowej na płaszczyźnie.



Rysunek 12.2: Po lewej obszar pod wykresem funkcji. Po prawej aproksymacja pola powierzchni tego obszaru od dołu sumą pól prostokątów zawartych w obszarze oraz aproksymacja od góry sumą pól prostokątów pokrywających obszar.

Niech $x : \{ 0, 1, 2, \ldots n \} \to [a, b]$ będzie silnie rosnącym, skończonym ciągiem takim, że $x_0 = a$ i $x_n = b$. Łatwo sprawdzić, że wtedy rodzina

$$\mathcal{P}_x := \{ [x_{i-1}, x_i] \mid i = 1, 2, \ldots n \}$$

jest podziałem przedziału $[a, b]$, zwanym *podziałem zadanym ciągiem x*.

Łatwo zauważyć, że każdy podział $\mathcal{P} \in \mathrm{Podz}(I)$ przedziału $I = [a, b] \in \mathrm{Przedz}(\mathbb{R})$ jest zadany pewnym ciągiem $x : \{ 0, 1, 2, \ldots n \} \to [a, b]$. W istocie ciąg x zadający podział $\mathcal{P}$ otrzymamy biorąc zbiór końców przedziałów w $\mathcal{P}$ ustawiony w ciąg rosnący. Ciąg ten nazywać będziemy *ciągiem zadającym podział $\mathcal{P}$* i oznaczać $x_{\mathcal{P}}$.

Zatem podziały przedziału $I = [a, b] \in \mathrm{Przedz}(\mathbb{R})$ możemy utożsamiać ze skończonymi, silnie rosnącymi ciągami liczb rozpoczynającymi się od liczby a i kończącymi się na liczbie b.

Uwaga 12.1.1 *Jeśli $\mathcal{P}$ jest podziałem przedziału I, to*

$$\operatorname{len} I = \sum_{P \in \mathcal{P}} \operatorname{len} P.$$

Definicja 12.1.2 Niech $\mathcal{P}_1, \mathcal{P}_2 \in \mathrm{Podz}(I)$ będą podziałami przedziału $I \in \mathrm{Przedz}(\mathbb{R})$. Mówimy, że podział $\mathcal{P}_2$ jest *podpodziałem* podziału $\mathcal{P}_1$ i piszemy $\mathcal{P}_2 \sqsubset \mathcal{P}_1$ jeżeli dla każdego przedziału $P_2 \in \mathcal{P}_2$ istnieje przedział $P_1 \in \mathcal{P}_1$ taki, że $P_2 \subset P_1$.

Łatwo zaobserwować, że podział $\mathcal{P}_2 \in \mathrm{Podz}(I)$ jest podpodziałem podziału $\mathcal{P}_1 \in \mathrm{Podz}(I)$ wtedy i tylko wtedy gdy ciąg zadający podział $\mathcal{P}_1$ jest podciągiem ciągu zadającego podział $\mathcal{P}_2$.

Uwaga 12.1.3 *Niech $\mathcal{P}_1, \mathcal{P}_2$ będą dwoma dowolnymi podziałami przedziału $I \in \mathrm{Przedz}(\mathbb{R})$. Istnieje podział $\mathcal{P}_3$ przedziału I taki, że $\mathcal{P}_3 \sqsubset \mathcal{P}_1$ i $\mathcal{P}_3 \sqsubset \mathcal{P}_2$.*

Dowód: Jako podział $\mathcal{P}_3$ wystarczy wziąć podział zadany przez zbiór końców przedziałów w $\mathcal{P}_1 \cup \mathcal{P}_2$ posortowany w ciąg rosnący.

□

12.1.3 Całkowalność w sensie Riemanna

Niech $I = [a, b] \in \mathrm{Przedz}(\mathbb{R})$ będzie ustalonym przedziałem, a $f : [a, b] \to \mathbb{R}$ niech będzie funkcją ograniczoną. Dla podzbioru $P \subset [a, b]$ przyjmujemy oznaczenia

$$\begin{aligned} m(f, P) &:= \inf \{ f(x) \mid x \in P \}, \\ M(f, P) &:= \sup \{ f(x) \mid x \in P \}. \end{aligned}$$

Zwróćmy uwagę, że ograniczoność funkcji f gwarantuje iż liczby $m(f, P)$ i $M(f, P)$ są dobrze określone.

Definicja 12.1.4 Niech $\mathcal{P}$ będzie podziałem przedziału $[a, b]$. Wyrażenia

$$\begin{aligned} L(f, \mathcal{P}) &:= \sum_{P \in \mathcal{P}} m(f, P) \operatorname{len} P, \\ U(f, \mathcal{P}) &:= \sum_{P \in \mathcal{P}} M(f, P) \operatorname{len} P \end{aligned}$$

nazywamy odpowiednio *dolną i górną sumą Darboux* dla funkcji f i podziału $\mathcal{P}$.

Łatwo zauważyć, że jeżeli $x : \{0, 1, 2, \dots n\} \to [a, b]$ jest ciągiem zadającym podział $\mathcal{P}$, to mamy

$$\begin{aligned} L(f, \mathcal{P}) &= \sum_{i=1}^{n} m(f, [x_{i-1}, x_i])(x_i - x_{i-1}), \\ U(f, \mathcal{P}) &= \sum_{i=1}^{n} M(f, [x_{i-1}, x_i])(x_i - x_{i-1}). \end{aligned}$$

Jeśli funkcja f ma wartości nieujemne, to iloczyny $m(f, [x_{i-1}, x_i])(x_i - x_{i-1})$ i $M(f, [x_{i-1}, x_i])(x_i - x_{i-1})$ odpowiadają polu prostokątów o podstawie $[x_{i-1}, x_i]$ i wysokości równej odpowiednio kresowi dolnemu i kresowi górnemu funkcji f na przedziale $[x_{i-1}, x_i]$. Zatem suma dolna Darboux i suma górna Darboux dają odpowiednio przybliżenie od dołu i od góry pola pod wykresem funkcji f.

Definicja 12.1.5 Wyrażenia

$$\begin{aligned} L(f) &:= \sup\{ L(f, \mathcal{P}) \mid \mathcal{P} \in \mathrm{Podz}([a, b]) \}, \\ U(f) &:= \inf\{ L(f, \mathcal{P}) \mid \mathcal{P} \in \mathrm{Podz}([a, b]) \}. \end{aligned}$$

nazywamy odpowiednio *całką dolną i całką górną* funkcji f. Jeżeli $L(f) = U(f)$, to funkcję f nazywamy *całkowalną w sensie Riemanna*, a liczbę $L(f) = U(f)$ nazywamy *całką Riemanna* funkcji f i oznaczamy $\int_a^b f$. Zbiór funkcji ograniczonych $f : [a, b] \to \mathbb{R}$, które są całkowalne w sensie Riemanna oznaczamy $\mathrm{Riem}([a, b])$.

Przykład 12.1.6 Najprostszym przykładem funkcji całkowalnej w sensie Riemanna jest funkcja stała $c_{[a,b]}$. Rzeczywiście, dla dowolnego przedziału $P \subset [a, b]$ mamy $m(c_{[a,b]}, P) = M(c_{[a,b]}, P) = c$, zatem dla dowolnego podziału $\mathcal{P} \in \mathrm{Podz}([a, b])$ jest

$$L(c_{[a,b]}, \mathcal{P}) = U(c_{[a,b]}, \mathcal{P}) = \sum_{P \in \mathcal{P}} c \operatorname{len} P = c \sum_{P \in \mathcal{P}} \operatorname{len} P = c(b - a).$$

Zatem $L(c_{[a,b]}) = U(c_{[a,b]}) = c(b - a)$ i w konsekwencji

$$\int_a^b c_{[a,b]} = c(b - a).$$

Przykład 12.1.7 Rozważmy funkcję Dirichleta

$$f_D : [0, 1] \ni x \mapsto \begin{cases} 0 & x \in \mathbb{Q}, \\ 1 & x \notin \mathbb{Q}. \end{cases}$$

Jeśli $P \subset [0, 1]$ jest przedziałem o niepustym wnętrzu, to $m(f_D, P) = 0$ i $M(f_D, P) = 1$. Stąd dla dowolnego $\mathcal{P} \in \mathrm{Podz}([0, 1])$ mamy $L(f_D, \mathcal{P}) = 0$ i $U(f_D, \mathcal{P}) = 1$. W konsekwencji $L(f_D) = 0 \neq 1 = U(f_D)$. Zatem funkcja Dirichleta nie jest całkowalna w sensie Riemanna.

12.1.4 Epsilonowe kryterium całkowalności

Lemat 12.1.8 *Jeśli* $\mathcal{P}_1, \mathcal{P}_2 \in \mathrm{Podz}([a,b])$ *oraz* $\mathcal{P}_2 \sqsubset \mathcal{P}_1$, *to*

$$L(f,\mathcal{P}_1) \leqslant L(f,\mathcal{P}_2) \text{ i } U(f,\mathcal{P}_2) \leqslant U(f,\mathcal{P}_1).$$

Dowód: Niech $P \in \mathcal{P}_1$. Z uwagi 12.1.1 mamy

$$m(f,P)\operatorname{len} P = m(f,P) \sum_{P' \subset P,\, P' \in \mathcal{P}_2} \operatorname{len} P' = \sum_{P' \subset P,\, P' \in \mathcal{P}_2} m(f,P)\operatorname{len} P' \leqslant \sum_{P' \subset P,\, P' \in \mathcal{P}_2} m(f,P')\operatorname{len} P'.$$

Sumując lewą i prawą stronę powyższych nierówności dla $P \in \mathcal{P}_1$ dostajemy

$$L(f,\mathcal{P}_1) \leqslant L(f,\mathcal{P}_2).$$

Dowód nierówności dla sum górnych jest analogiczny. □

Z lematu 12.1.8 i uwagi 12.1.3 dostajemy następujący wniosek

Wniosek 12.1.9 *Dla dowolnych podziałów* $\mathcal{P}_1, \mathcal{P}_2 \in \mathrm{Podz}([a,b])$ *mamy*

$$L(f,\mathcal{P}_1) \leqslant U(f,\mathcal{P}_2).$$

Z powyższgo wniosku dostajemy natychmiast następujący wniosek

Wniosek 12.1.10 *Dla dowolnej funkcji ograniczonej* $f : [a,b] \to \mathbb{R}$ *mamy*

$$L(f) \leqslant U(f).$$

Jesteśmy teraz przygotowani, by pokazać następującą charakteryzację funkcji całkowalnych w sensie Riemanna.

Twierdzenie 12.1.11 *Funkcja ograniczona* $f : [a,b] \to \mathbb{R}$ *jest całkowalna w sensie Riemanna wtedy i tylko wtedy gdy dla każdego* $\epsilon > 0$ *istnieje podział* $\mathcal{P} \in \mathrm{Podz}([a,b])$ *przedziału* $[a,b]$ *taki, że*

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) < \epsilon. \tag{12.1}$$

Dowód: Przyjmijmy najpierw, że f jest całkowalna w sensie Riemanna. Ustalmy $\epsilon > 0$. Z definicji całki dolnej i górnej do $\epsilon/2$ znajdziemy podziały $\mathcal{P}_1, \mathcal{P}_2 \in \mathrm{Podz}([a,b])$ takie, że

$$L(f,\mathcal{P}_1) > L(f) - \frac{\epsilon}{2} \text{ oraz } U(f,\mathcal{P}_2) < U(f) + \frac{\epsilon}{2}.$$

Na mocy uwagi 12.1.3 możemy wybrać podział $\mathcal{P}_3 \in \mathrm{Podz}([a,b])$ taki, że $\mathcal{P}_3 \sqsubset \mathcal{P}_1$ oraz $\mathcal{P}_3 \sqsubset \mathcal{P}_2$. Z lematu 12.1.8 oraz z (12.1.4) dostajemy

$$U(f,\mathcal{P}_3) - L(f,\mathcal{P}_3) \leqslant U(f,\mathcal{P}_2) - L(f,\mathcal{P}_1) < U(f) + \frac{\epsilon}{2} - (L(f) - \frac{\epsilon}{2}) = \epsilon,$$

bo z założenia $L(f) = U(f)$. Zatem podział $\mathcal{P}_3$ spełnia wymóg (12.1).

Dla dowodu przeciwnej implikacji przyjmijmy nie wprost, że f nie jest całkowalna w sensie Riemanna. Wtedy $L(f) \neq U(f)$, zatem z wniosku 12.1.10 wynika, że $L(f) < U(f)$. Z definicji sum dolnych i górnych wnosimy, że dla dowolnego $\mathcal{P} \in \mathrm{Podz}([a,b])$ mamy

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) \geqslant U(f) - L(f).$$

W szczególności wynika stąd, że dla $\epsilon := \frac{U(f)-L(f)}{2}$, które na mocy przyjętego założenia jest dodatnie, nie znajdziemy podziału $\mathcal{P}$ spełniającego warunek (12.1). A to dowodzi, że implikacja przeciwna również zachodzi. □

12.2 Własności całki Riemanna

12.2.1 Monotoniczność całki

Twierdzenie 12.2.1 *Niech $f, g : [a, b] \to \mathbb{R}$ będą funkcjami ograniczonymi. Jeśli $f \leqslant g$ to*

$$L(f) \leqslant L(g) \text{ oraz } U(f) \leqslant U(g).$$

W szczególności, jeśli tak f jak i g są całkowalne w sensie Riemanna, to

$$\int_a^b f \leqslant \int_a^b g.$$

12.2.2 Przestrzeń wektorowa funkcji całkowalnych w sensie Riemanna.

Twierdzenie 12.2.2 *Jeśli $f, g : [a, b] \to \mathbb{R}$ są całkowalne w sensie Riemanna to $f + g$ oraz $f - g$ są całkowalne w sensie Riemanna oraz*

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g \qquad i \qquad \int_a^b (f-g) = \int_a^b f - \int_a^b g.$$

Ponadto, dla dowolnego $\alpha \in \mathbb{R}$ funkcja αf jest całkowalna w sensie Riemanna oraz

$$\int_a^b (\alpha f) = \alpha \int_a^b f.$$

Uwaga 12.2.3 *Jeśli $f : [a, b] \to \mathbb{R}$ jest całkowalna w sensie Riemanna, to*

$$\left| \int_a^b f \right| \leqslant (b-a) \sup_{x \in [a,b]} |f(x)|.$$

Twierdzenie 12.2.4 *Przestrzeń $\mathrm{Riem}([a, b])$ funkcji całkowalnych w sensie Riemanna na przedziale zwartym $[a, b]$ z działaniem dodawania funkcji oraz możenia funkcji przez liczbę jest przestrzenią wektorową. Ponadto, odwzorowanie*

$$\mathrm{Riem}([a,b]) \ni f \mapsto ||f|| := \sup_{x \in [a,b]} |f(x)| \in \mathbb{R}^*$$

jest normą, która zadaje w $\mathrm{Riem}([a, b])$ strukturę przestrzeni wektorowej unormowanej, a tym samym strukturę przestrzeni metrycznej. Co więcej, odwzorowanie

$$\mathrm{Riem}([a,b]) \ni f \mapsto \int_a^b f \in \mathbb{R}$$

jest odwzorowaniem liniowym oraz Lipschitzowskim ze stałą $\mathrm{len}[a, b]$, a więc w szczególności odwzorowaniem ciągłym.

12.2.3 Miarowe kryterium całkowalności.

Definicja 12.2.5 *Miarą zewnętrzną* podzbiorów $\mathbb{R}$ nazywamy funkcję

$$\mu^* : \mathcal{P}(\mathbb{R}) \ni E \mapsto \inf\{\textstyle\sum_{A\in\mathcal{A}} \operatorname{len} A \mid \mathcal{A} \subset \operatorname{Przedz}(\mathbb{R}) \text{ oraz } E \subset \bigcup \mathcal{A}\}.$$

Mówimy, że zbiór $E \subset \mathbb{R}$ jest *miary zero* jeżeli $\mu^*(E) = 0$.

Twierdzenie 12.2.6 *Miara zewnętrzna ma następujące własności:*

(i) $\mu^*(\varnothing) = 0$,

(ii) $E \subset F \;\Rightarrow\; \mu^*(E) \leqslant \mu^*(F)$,

(iii) $\mu^*(\bigcup_{i=1}^{\infty} E_i) \leqslant \sum_{i=1}^{\infty} \mu^*(E_i)$,

Przykład 12.2.7 Zbiorem miary zero jest każdy skończony zbiór punktów.

Wniosek 12.2.8 *Suma co najwyżej przeliczalnej rodziny zbiorów miary zero jest zbiorem miary zero.*

Twierdzenie 12.2.9 *Funkcja ograniczona $f : [a, b] \to \mathbb{R}$ jest całkowalna w sensie Riemanna wtedy i tylko wtedy gdy zbiór punktów, w których nie jest ciągła ma miarę zero. W szczególności, jeśli f jest ciągła, to jest całkowalna w sensie Riemanna.*

12.2.4 Całkowalność złożenia i iloczynu

Twierdzenie 12.2.10 *Całka Riemanna ma następujące własności:*

(i) jeśli $f \in \operatorname{Riem}(I)$ jest całkowalna w sensie Riemanna, a $\varphi : \mathbb{R} \to \mathbb{R}$ jest ciągła, to $\varphi \circ f \in \operatorname{Riem}(I)$,

(ii) jeśli $f, g \in \operatorname{Riem}(I)$, to ich ich iloczyn $f \cdot g \in \operatorname{Riem}(I)$,

(i) jeśli $f \in \operatorname{Riem}(I)$, to $|f| \in \operatorname{Riem}(I)$.

12.2.5 Całka Riemanna na podprzedziale

Definicja 12.2.11 Niech $f : \mathbb{R} \multimap \mathbb{R}$ będzie funkcją ograniczoną i niech $[a, b] \subset \operatorname{dom} f$. Mówimy, że *$f$ jest całkowalna w sensie Riemanna na przedziale $[a, b]$*, jeżeli zwężenie $f_{|[a,b]}$ jest funkcją całkowalną w sensie Riemanna. W takim przypadku całkę $\int_a^b f_{|[a,b]}$ nazywamy całką Riemanna funkcji f na przedziale $[a, b]$ lub całką oznaczoną funkcji f od a do b i oznaczamy

$$\int_a^b f := \int_a^b f_{|[a,b]}.$$

Twierdzenie 12.2.12 *Niech $f : \mathbb{R} \multimap \mathbb{R}$ będzie funkcją ograniczoną i niech $[a, b]$ będzie przedziałem zwartym zawartym w dziedzinie funkcji f. Jeśli $\mathcal{P}$ jest podziałem przedziału $[a, b]$ zadanym ciągiem $x : \{0, 1, 2, \ldots k\} \to [a, b]$, to f jest całkowalna w sensie Riemanna na przedziale $[a, b]$ wtedy i tylko wtedy gdy dla każdego $i = 1, 2, \ldots k$ funkcja f jest całkowalna w sensie Riemanna na przedziale $[x_{i-1}, x_i]$. Co więcej, gdy którykolwiek z tych warunków zachodzi, to*

$$\int_a^b f = \sum_{i=1}^{k} \int_{x_{i-1}}^{x_i} f.$$

12.2.6 Całka oznaczona

Definicja 12.2.13 Niech $f : \mathbb{R} \multimap \mathbb{R}$ będzie funkcją ograniczoną całkowalną w sensie Riemanna na przedziale $[a, b]$. Całkę Riemanna z f na przedziale $[a, b]$ określa się też mianem *całki oznaczonej* z f na przedziale $[a, b]$ i oznacza się bardziej tradycyjnym symbolem

$$\int_a^b f(x)dx := \int_a^b f := \int_{[a,b]} f.$$

Wniosek 12.2.14 *Jeśli f jest całkowalna w sensie Riemanna na przedziale $[a, b]$ oraz $c \in (a, b)$, to f jest całkowalna w sensnie Riemanna na przdziałach $[a, c]$ i $[c, b]$ oraz*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

W programie Mathematica całkę oznaczoną z funkcji f na przedziale $[a, b]$ obliczamy używając instrukcji

```
Integrate[f[x], {x, a, b}]
```

W miarę swoich możliwości Mathematica stara się zwrócić dokładny wynik. Na przykład

```
Integrate[Sqrt[1 - x^2], {x, 0, 1}]
```

zwraca $\frac{\pi}{4}$.

12.2.7 Twierdzenie Riemanna

Niech $f : \mathbb{R} \multimap \mathbb{R}$ będzie funkcją ograniczoną i niech $[a, b]$ będzie przedziałem zwartym zawartym w dziedzinie funkcji f. Niech $\mathcal{P}$ będzie podziałem przedziału $[a, b]$ zadanym ciągiem $x : \{0, 1, 2, \ldots k\} \to [a, b]$.

Definicja 12.2.15 Niech $\Xi := \{\xi_1, \xi_2, \ldots \xi_k\} \subset [a, b]$ będzie zbiorem punktów takich, że $\xi_i \in [x_{i-1}, x_i]$ dla $i = 1, 2, \ldots k$. Parę $(\mathcal{P}, \Xi)$ nazywamy *parą aproksymacyjną* dla f na przedziale $[a, b]$, a wyrażenie

$$S(f, \mathcal{P}, \Xi) := \sum_{i=1}^{k} f(\xi_i)(x_i - x_{i-1})$$

nazywamy *sumą aproksymacyjną* dla f na przedziale $[a, b]$. Ciąg par aproksymacyjnych $(\mathcal{P}_n, \Xi_n)_{n=1}^{\infty}$ nazywamy ciągiem normalnym, jeżeli ciąg średnic podziałów $\mathcal{P}_n$ zmierza do zera.

Twierdzenie 12.2.16 (Riemanna). *Niech $f : [a, b] \to \mathbb{R}$ będzie funkcją ograniczoną. Funkcja ta jest całkowalna w sensie Riemanna wtedy i tylko wtedy gdy istnieje liczba $g \in \mathbb{R}$ taka, że dla każdego $\epsilon > 0$ istnieje $\delta > 0$ taka, że dla każdej pary aproksymacyjnej $(\mathcal{P}, \Xi)$ dla f na przedziale $[a, b]$ mamy*

$$\delta(\mathcal{P}) < \delta \;\Rightarrow\; |S(f, \mathcal{P}, \Xi) - g| < \epsilon.$$

Liczba g jest wyznaczona jednoznacznie i jest równa całce $\int_a^b f$.

Program, który liczy przybliżenie całki

$$\int_0^x f(t)dt$$

w oparciu o sumy Riemanna

$$s_n := \sum_{i=1}^{n} \frac{1}{n} f(\frac{i}{n}x)$$

przedstawia listing 12.1. Podobny program w języku C++ przedstawia listing 12.2.

Wyrazy ciągu s_n powinny przybliżać pole ćwiartki koła, czyli

$$\frac{\pi}{4} = 0.7853981633974483...,$$

natomiast 20-ty wyraz ciągu s_n wyliczony przez program to

$$s_{20} = 0.785397208948704.$$

Widzimy więc, że z punktu widzenia liczenia liczby π ta technika jest gorsza niż zastosowana wcześniej technika szeregów. Jednak pożytki z niej leżą gdzie indziej.

Ćwiczenie komputerowe 12.2.17 Adaptuj program z listingu 12.2 do policzenia przybliżenia pola soczewkowatej figury na rys. 1.12.

```
(* Przybliżanie całki oznaczonej z funkcji f od 0 do x poprzez
   pola m prostokątów pod wykresem *)
RiemIntegral[f_, x_, m_] :=
  Module[{s, k, y},
  s = 0.; k = 0;
  (* Sumujemy w zmiennej s wartości funkcji w punktach x*k/m *)
  Do[
    y = x*k/m;
    s = s + f[y];
    k = k + 1,
    {m}
  ];
  (* Zwracamy sumę pomnożoną przez wspólną długość podstawy wynoszącą 1/m *)
  s/m
]

(* Wykres poniższej funkcji nad przedziałem[0,1] to ćwiartka okręgu *)
arc[x_] := Sqrt[1 - x^2];

(* Przybliżymy pole pod ćwiartką okręgu danego funkcją arc nad[0,1]
Wykonamy 20 prób za każdym razem podwajając ilość przedziałów *)
m = 1; d = 1;
Do[
  p = RiemIntegral[arc, 1., m];
  (* Wydruk kontrolny: całka powinna być bliska pi/4 *)
  Print["s[", d, "]=", NumberForm[p,16]];
  m = m*2; d = d + 1,
  {20}
]
```

Listing 12.1: Przybliżanie całki oznaczonej w języku Mathematica

```
/******************** RiemIntegral.cpp ********************/
/* Przybliżanie całki oznaczonej poprzez pola prostokątów  */
/* pod wykresem                                            */
/*********************************************************/
#include <iostream>
#include <cmath>
#include <iomanip>
using namespace std;

/* Wykres poniższej funkcji nad przedziałem [0,1] to ćwiartka okręgu */
double arc(double x){
  return sqrt(1-x*x);
}

/* Przybliżanie całki Riemanna funkcji f od 0 do x sumą m prostokątów
   Zapis double (*f)(double) oznacza, że argumentem f jest funkcja */
double RiemIntegral(double (*f)(double), double x,int m){
  double s=0;
  /* Sumujemy w zmiennej s wartości funkcji w punktach x*k/m */
  for(int k=0;k<m;k=k+1){
    double y=x*k/m;
    s=s+f(y);
  }
  /* Zwracamy sumę pomnożoną przez wspólną długość podstawy wynoszącą 1/m */
  return s=s/m;
}

int main(){
  cout << fixed << right;
  cout << setprecision(15);

  int m=1;
  /* Przybliżymy pole pod ćwiartką okręgu danego funkcją arc nad [0,1]
     Wykonamy 20 prób za każdym razem podwajając ilość przedziałów */
  for(int d=1;d<=20;d=d+1){
    double p=RiemIntegral(arc,1.,m);
    /* Wydruk kontrolny: całka powinna być bliska pi/4 */
    cout << "s[" << setw(2) << d << "] =" << setw(18) << p;
    cout << endl;
    m=m*2;
  }
}
```

Listing 12.2: Przybliżanie całki oznaczonej w C++

Rozdział 13

Funkcja pierwotna i całka nieoznaczona

Jak wspomnieliśmy w rozdziale wstępnym Newton i Leibniz odkryli, że pojęcie pochodnej i pojęcie całki są ze sobą powiązane. Wzięcie całki z pochodnej oraz pochodnej z całki prowadzą do wyjściowej funkcji. W niniejszym rozdziale doprecyzujemy co to oznacza oraz zajmiemy się pożytkami jakie z tego faktu płyną.

13.1 Podstawowe twierdzenie rachunku różniczkowego i całkowego

13.1.1 Pochodna całki

Twierdzenie 13.1.1 *Niech $f : [a, b] \to \mathbb{R}$ będzie funkcją całkowalną w sensie Riemanna. Rozważmy funkcję*

$$F : [a, b] \ni x \mapsto \int_a^x f(t)dt.$$

Wtedy

(i) funkcja F jest funkcją Lipschitza,

(ii) jeśli f jest ciągła w punkcie $x_0 \in [a, b]$ to F jest różniczkowalna w x_0 oraz $F'(x_0) = f(x_0)$.

Dowód: Aby pokazać własność (i) rozważmy $x, y \in (a, b)$ spełniające $x < y$. Niech $M := \sup_{x\in[a,b]} f(x)|$. Mamy

$$\left|\int_x^y f\right| \leqslant \int_x^y |f| \leqslant M(y - x) = M|y - x|.$$

Zatem F rzeczywiście jest funkcją Lipschitza.

Dla dowodu własności (ii) ustalmy $\epsilon > 0$. Z ciągłości f w x_0 znajdziemy $\delta > 0$ takie, ze

$$|x - x_0| < \delta \;\Rightarrow\; |f(x) - f(x_0| < \epsilon.$$

Zatem dla $|h| < \delta$

$$\left|\frac{F(x_0 + h) - F)x_0)}{h} - f(x_0)\right| = \frac{1}{|h|}\left|\int_{x_0}^{x_0+h} f(t)dt - \int_{x_0}^{x_0+h} f(x_0)dt\right| \leqslant \frac{1}{|h|}\int_{x_0}^{x_0+h} |f(t) - f(x_0)| \leqslant \frac{1}{|h|}\int_{x_0}^{x_0+h} \epsilon dt = \left|\frac{h\epsilon}{h}\right| = \epsilon,$$

co dowodzi, że F jest różniczkowalna w x_0 oraz $F'(x_0) = f(x_0)$. □

13.1.2 Funkcja pierwotna i całka nieoznaczona.

Definicja 13.1.2 Niech $f : [a, b] \to \mathbb{R}$. Mówimy, że $F : [a, b] \to \mathbb{R}$ jest funkcją pierwotną funkcji f jeżeli F jest różniczkowalna oraz $F'(x) = f(x)$ dla każdego $x \in [a, b]$.

Z twierdzenia 13.1.1 łatwo wyprowadzamy następujacy wniosek.

Wniosek 13.1.3 *Jeśli $f : [a, b] \to \mathbb{R}$ jest ciągła to funkcja*

$$[a, b] \ni x \mapsto \int_a^x f(t)dt \in \mathbb{R}$$

jest funkcją pierwotna funkcji f.

Naturalne jest pytanie czy funkcja f moze mieć więcej niż jedną pierwotną. Dodając do pierwotnej funkcję stałą dostajemy inną pierwotną tej samej funkcji, bo pochodna funkcji stałej jest zero. Okazuje się, że więcej pierwotnych już nie ma jak wynika z następującego twierdzenia.

Twierdzenie 13.1.4 *Jeśli F_1 i F_2 są pierwotnymi funkcji f, to różnica $F_1 - F_2$ jest funkcja stałą.*

Dowód: Niech $F := F_1 - F_2$. Funkcja F jest różniczkowalna oraz $F' = F_1' - F_2' = 0$. Niech $x, y \in [a, b]$ i $x < y$. Z twierdzenia o wartości średniej wiemy, że istnieje $\xi \in (x, y)$ takie, że

$$F(x) - F(y) = F'(\xi)(x - y) = 0.$$

Zatem $F(x) = F(y)$, a ponieważ x, y są dowolne, otrzymujemy tezę. □

Wygodnie jest zebrać wszystkie pierwotne razem.

Definicja 13.1.5 Niech $f : [a, b] \to \mathbb{R}$ będzie funkcją całkowalną w sensie Riemanna. Zbiór wszystkich pierwotnych funkcji f nazywamy całką nieoznaczoną f i oznaczamy $\int f$.

W szczególności zapis $F \in \int f$ oznacza, że F jest fukcją pierwotną f. Stosowany jest również tradycyjny, choć mało precyzyjny zapis

$$F(x) = \int f(x)dx + C,$$

który należy interpretować tak, że wybierając jakąś pierwotną z całki nieoznaczonej i dodając do niej stałą dostaniemy inną pierwotną.

13.1.3 Podstawowe twierdzenie rachunku różniczkowego i całkowego

Leibniz i Newton odkryli, że znajomość pierwotnej pozwala wygodnie obliczać pole pod wykresem funkcji. Jest to nieporównalnie łatwiejsze od liczenia sum aproksymacyjnych Riemanna. Natomiast maszynką dostarczające liczne (choć niestety nie wszystkie potrzebne) wzory na funkcje pierwotne są wszelkie wzory na pochodne.

Twierdzenie 13.1.6 (podstawowe twierdzenie rachunku różniczkowego i całkowego) *Niech $f : [a, b] \to \mathbb{R}$ będzie całkowalna w sensie Riemanna. Jeśli F jest pierwotną funkcji f to*

$$F(x) = \int_a^x f(t)dt + F(a). \tag{13.1}$$

W szczególności

$$\int_a^b f(t)dt = F(b) - F(a). \tag{13.2}$$

Dla wygody w dalszym ciągu stosować będziemy zapis $F|_a^b := F(b) - F(a)$.

Dowód: Pokażemy najpierw (13.2). Dla $k \in \mathbb{N}_1$ wybierzmy podział $\mathcal{P}_k$ przedziału $[a, b]$ o średnicy nie większej niż $\frac{1}{k}$, zadanym ciągiem $x^k : \{0, 1, \dots n_k\} \to \mathbb{R}$. Z twierdzenia o wartości średniej zastosowanego do F mamy

$$F(b) - F(a) = \sum_{i=1}^{n} F'(\xi_i^k)(x_i^k - x_{i-1}^k) \tag{13.3}$$

dla pwenych punktów $\xi_i^k \in (x_{i-1}^k, x_i^k)$. Ponieważ $F'(x_i^k) = f(x_i^k)$, więc prawa strona równości (13.3) jest sumą aproksymacyjną dla f na $[a, b]$. Zatem z twierdzenia Riemanna (tw. 12.2.16) prawa strona (13.3) zmierza do $\int_a^b f$ przy $k \to \infty$, bo podziały $\mathcal{P}_k$ dobraliśmy tak, by ich średnice zmierzały do zera. Ale na mocy równości (13.3) wszystkie te sumy aproksymacyjne są równe $F(b) - F(a)$, zatem i granica jest równa $F(b) - F(a)$ co dowodzi (13.2).

Aby pokazać (13.1) wystarczy zastosować pokazany już wzór (13.2) do zawężenia $f_{|[a,x]}$. □

13.2 Podstawowe techniki całkowania.

13.2.1 Własności całki nieoznaczonej.

Podamy teraz kilka podstawowych wzorów dla całki nieoznaczonej. Pojawiające się w nich funkcje rozważać będziemy na przedziale $[a, b]$. Symbolem Const oznaczać będziemy rodzinę funkcji stałych na przedziale $[a, b]$.

Twierdzenie 13.2.1 *(liniowość całki nieoznaczonej). Niech $f, g : [a, b] \to \mathbb{R}$ będą całkowalne w sensie Riemanna. Wtedy $f + g$ jest całkowalna w sensie Riemanna oraz*

$$\int (f + g) = \int f + \int g. \tag{13.4}$$

Ponadto αf jest całkowalna w sensie Riemanna dla dowolnego $\alpha \in \mathbb{R}$ oraz

$$\int (\alpha f) = \alpha \int f. \tag{13.5}$$

Zauważmy, że wzory (13.4) i (13.5) należy rozumieć jako równość zbiorów, ponieważ całka nieoznaczona jest zbiorem, a dokładniej zbiorem funkcji pierwotnych.

Dowód: Całkowalność funkcji $\alpha f + \beta g$ i funkcji αf dostajemy wprost z twierdzenia 12.2.2. Zauważmy, że jeżeli F jest pierwotną f, a G jest pierwotną g, to $F+G$ jest pierwotną $f+g$ na podstawie twierdzenia o pochodnej sumy funkcji. Dowodzi to zawieranie się strony prawej w lewej we wzorze (13.4). Aby pokazać inkluzję przeciwną załóżmy, że H jest elementem lewej strony równości (13.4). Wtedy H jest pierwotną funkcji $f + g$. Niech F będzie dowolnie ustaloną pierwotną funkcji f. Połóżmy $G : H - F$. Mamy $G' = H' - F' = (f + g) - f = g$. Zatem G jest pierwotną funkcji g, a ponieważ $H = F + G$, więc H jest sumą pierwotnej funkcji f i pierwotnej funkcji g. Zatem H należy do prawej strony (13.4).

Dowód (13.5) jest podobny i zostawiamy go jako ćwiczenie. □

Następujące twierdzenia są natychmiastową konsekwencją stosownych wzorów dla pochodnych i definicji całki nieoznaczonej.

Twierdzenie 13.2.2 *Zachodzą następujące wzory*

$$\int 0_{[a,b]}dx = \text{Const}, \qquad \int 1_{[a,b]}dx = x + \text{Const},$$
$$\int \frac{1}{x}dx = \ln x + \text{Const}, \qquad \int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + \text{Const } dla\ \alpha \neq -1,$$
$$\int e^x dx = e^x + \text{Const}\,.$$

□

Twierdzenie 13.2.3 *Zachodzą następujące wzory*

$$\int \sin x dx = -\cos x + \text{Const}, \qquad \int \cos x dx = \sin x + \text{Const},$$
$$\int \frac{1}{\cos^2 x}dx = \operatorname{tg} x + \text{Const}, \qquad \int \frac{1}{\sqrt{1-x^2}}dx = \operatorname{arc\,sin} x + \text{Const},$$
$$\int \frac{1}{1+x^2}dx = \operatorname{arc\,tg} x + \text{Const}\,.$$

□

13.2.2 Całkowanie przez części

Widzieliśmy, że całkowanie sumy dwóch funkcji jest proste. Jest tak, bo pochodna sumy jest sumą pochodnych i w efekcie również całka sumy jest sumą całek. Niestety pochodna iloczynu nie jest iloczynem pochodnych, a w efekcie całka iloczynu nie jest iloczynem całek. Dlatego nie jest oczywiste jak liczyć całki iloczynu dwóch funkcji. Jednak wzór na pochodną iloczynu jest w tym przypadku pomocny, chociaż nie bezpośrednio i nie automatycznie. Całkując wzór na pochodną iloczynu dostajemy

$$\int (fg)' = \int f'g + \int fg'.$$

Ponieważ fg jest pierwotną $(fg)'$, dostajemy

$$fg + \text{Const} = \int f'g + \int fg'.$$

Jeśli jedna z dwóch całek po prawej stronie jest łatwiejsza do policzenia niż druga, to wzór ten mozemy potraktować jako metodę sprowadzającą liczenie całki $\int f'g$ do liczenia całki $\int fg'$. Jak zobaczymy, metoda ta może być użyteczna. Określana jest mianem całkowania przez części.

Twierdzenie 13.2.4 *Załóżmy, że $f, g : [a, b] \to \mathbb{R}$ są klasy C^1. Wtedy*

$$\int f'g = fg - \int fg' \tag{13.6}$$

W szczególności mamy też wzór dla całek oznaczonych:

$$\int_a^b f'(x)g(x)dx = f(x)g(x)|_a^b - \int_a^b f(x)g'(x)dx. \tag{13.7}$$

Dowód: Niech F będzie elementem lewej strony wzoru (13.6). Wtedy $F' = f'g$. Mamy

$$F = fg - (fg - F).$$

Zatem aby pokazać, że F jest elementem prawej strony wzoru (13.6), wystarczy pokazać, że $fg - F$ jest pierwotną funkcji fg'. Jest tak w istocie, bo

$$(fg - F)' = (fg)' - F' = f'g + fg' - f'g = fg'.$$

Rozważmy teraz element prawej strony równości (13.6). Ma on postać $fg - G$, gdzie G jest pierwotną fg'. Musimy pokazać, że $fg - G$ jest pierwotną $f'g$. Rzeczywiście tak jest, bo

$$(fg - G)' = f'g + fg' - fg' = f'g.$$

Aby pokazać (13.7) wybierzmy F pierwotną dla $f'g$. Ze wzoru (13.6) wiemy, że $F = fg - G$, gdzie G jest pewną pierwotną fg'. W szczególności

$$\begin{aligned} F(b) &= f(b)g(b) - G(b) \\ F(a) &= f(a)g(a) - G(a). \end{aligned}$$

Odejmując te równości stronami dostajemy

$$F(b) - F(a) = f(b)g(b) - f(a)g(a) - (G(b) - G(a)). \tag{13.8}$$

W oparciu o przyjętą konwencję zapisywania różnic mamy

$$f(b)g(b) - f(a)g(a) = f(x)g(x)|_a^b.$$

Z podstawowego twierdzenia rachunku różniczkowego i całkowego (tw. 13.1.6) oraz z definicji F i G jako funkcji pierwotnych odpowiednio $f'g$ oraz fg' otrzymujemy

$$\begin{aligned} F(b) - F(a) &= \int_a^b f'(x)g(x)dx \\ G(b) - G(a) &= \int_a^b f(x)g'(x)dx. \end{aligned}$$

Wstawiając otrzymane trzy równości do (13.8) otrzymujemy (13.7). □

Przykład 13.2.5 Metodę całkowania przez części zastosujemy do policzenia całki

$$\int x \sin x dx.$$

Musimy w tym celu przedstawić funkcję $x \mapsto x \sin x$ jako iloczyn pewnej funkcji przez pochodną innej funkcji. Mamy dwie możliwości:

$$x \sin x = x(-\cos x)'$$

lub

$$x \sin x = \left(\frac{x^2}{2}\right)' \sin x$$

i nie jest oczywiste, z której lepiej skorzystać. Pierwsza sprowadzi problem do policzenia całki

$$\int (x)'(-\cos x)dx,$$

a druga do całki

$$\int \left(\frac{x^2}{2}\right)(\sin x)'dx.$$

Z tych dwóch całek zdecydowanie łatwiejsza jest pierwsza, bo

$$\int (x)'(-\cos x)dx = -\int \cos x dx = -\sin x + \text{Const}\,.$$

Zatem wstawiając do wzoru (13.6) za f funkcję $x \mapsto -\cos x$, a za g funkcję identycznościową $x \mapsto x$ dostajemy

$$\int x \sin x dx = \int (-\cos x)' x dx = -x\cos x - \int (-\cos x)x'dx = -x\cos x + \int \cos x dx = -x\cos x + \sin x + \text{Const}\,.$$

13.2.3 Całkowanie przez podstawienie

Podobnie jak w przypadku iloczynu nie jest oczywiste jak całkować złożenie funkcji. Pochodna złożenia nie jest złożeniem pochodnych zatem i całka złożenia nie jest złożeniem całek. Jednak całkując wzór na pochodną złożenia dostajemy wzór dostarczający metodę liczenia całek przydatną w przypadku złożeń i zwaną całkowaniem przez podstawienie.

Twierdzenie 13.2.6 *Załóżmy, że $f : [a,b] \to \mathbb{R}$ jest funkcją ciągłą, a $g : [\alpha, \beta] \to [a,b]$ jest klasy C^1. Jeśli F jest pierwotną funkcji f to $F \circ g$ jest pierwotną funkcji $(f \circ g)g'$. W konsekwencji*

$$\int (f \circ g)g' = \left(\int f\right) \circ g. \tag{13.9}$$

W szczególności

$$\int_\alpha^\beta f(g(t))g'(t)dt = \int_{g(\alpha)}^{g(\beta)} f(x)dx. \tag{13.10}$$

Dowód: Jeśli F jest pierwotną funkcji f to oczywiście $F \circ g$ jest pierwotną funkcji $(f \circ g)g'$, zatem prawa strona zawiera się w lewej w (13.9). Aby pokazać, że lewa strona zawiera się w prawej przyjmijmy, że G jest pierwotną funkcji $(f \circ g)g'$. Wybierzmy też pewną pierwotną F funkcji f. Wiemy już, że wtedy $F \circ g$ jest pierwotną $(f \circ g)g'$, zatem na mocy

twierdzenia 13.1.4 istnieje stała C taka, że $G = F \circ g + C$. Mamy

$$G = F \circ g + C = F \circ g + C \circ g = (F + C) \circ g = \bar{F} \circ g,$$

gdzie $\bar{F} = F + C$ jest inną pierwotną f. Zatem G jako podstawienie g w pierwotnej $\bar{F}$ funkcji f należy do prawej strony równości (13.9).

Pozostaje pokazać (13.10). Wiemy, że jeśli F jest pierwotną f to $F \circ g$ jest pierwotną $(f \circ g)g'$. Zatem z podstawowego twierdzenia rachunku różniczkowego i całkowego (tw. 13.1.6) mamy

$$\int_\alpha^\beta f(g(t))g'(t)dt = (F \circ g)(\alpha) - (F \circ g)(\beta) = F(g(\alpha)) - F(g(\beta)) = \int_{g(\alpha)}^{g(\beta)} f(x)dx.$$

□

Przykład 13.2.7 Policzymy całkę nieoznaczoną $\int a^x dx$. Mamy $\int a^x dx = \int e^{x \ln a} dx$. Chcielibyśmy sprowadzić policzenie całki $\int e^{x \ln a} dx$ do całki $\int e^u du$. Funkcja $x \mapsto e^{x \ln a}$ jest podstawieniem funkcji $g : x \mapsto x \ln a$ w funkcji $f : u \mapsto e^u$. Ponieważ pochodną g jest funkcja stała $\ln a$, więc z twierdzenia 13.2.6 mamy

$$\begin{aligned}\int a^x dx &= \int e^{x \ln a} dx = \frac{1}{\ln a} \int e^{x \ln a} \ln a dx = \frac{1}{\ln a} \left(\int e^u du \right) \circ (x \mapsto x \ln a) = \\ &\frac{1}{\ln a} (e^u + \text{Const}) \circ (x \mapsto x \ln a) = \frac{1}{\ln a} e^{x \ln a} + \text{Const} = \frac{a^x}{\ln a} + \text{Const}\,.\end{aligned}$$

13.2.4 Twierdzenia o wartości średniej dla całek

Twierdzenie 13.2.8 (o wartości średniej dla całek) *Niech $f : [a, b] \to \mathbb{R}$ będzie funkcją klasy C^1, a $g : [a, b] \to \mathbb{R}$ funkcją całkowalną w sensie Riemanna. Jeśli g ma stały znak na $[a, b]$ to istnieje $\xi \in [a, b]$ takie, że*

$$\int_a^b fg = f(\xi) \int_a^b g. \tag{13.11}$$

Dowód: Rozważmy przypadek $g \geqslant 0$. Połóżmy $m := \inf \{ f(x) \mid x \in [a, b] \}$, $M := \sup \{ f(x) \mid x \in [a, b] \}$ i ustalmy $x \in [a, b]$. Wtedy $m \leqslant f(x) \leqslant M$, a zatem również $mg(x) \leqslant f(x)g(x) \leqslant Mg(x)$. Całkując dostajemy

$$m \int_a^b g(t)dt \leqslant \int_a^b f(t)g(t)dt \leqslant \int_a^b Mg(t)dt.$$

Z otrzymanej nierówności wynika, że jeśli $\int_a^b g = 0$, to również $\int_a^b fg = 0$ i w konsekwencji (13.11) zachodzi dla dowolnie dobranego ξ. Pozostaje zatem rozważyć przypadek $\int_a^b g \neq 0$. Wtedy

$$m \leqslant \frac{\int_a^b fg}{\int_a^b g} \leqslant M.$$

Zatem z twierdzenia 7.2.14 i wniosku 7.4.7 znajdziemy $\xi \in [a, b]$ takie, że

$$f(\xi) = \frac{\int_a^b fg}{\int_a^b g}$$

co dowodzi (13.11) dla $g \geqslant 0$. Rozumowanie dla $g \leqslant 0$ jest analogiczne. □

Przyjmując w twierdzeniu 13.2.8 za g funkcję stale równą jeden dostajemy następujący wniosek.

Wniosek 13.2.9 *Jeśli $f : [a, b] \to \mathbb{R}$ jest klasy C^1, to istnieje $\xi \in [a, b]$ takie, że*

$$\int_a^b f = f(\xi)(b - a).$$

W programie Mathematica całkę nieoznaczoną (a dokładniej jedną z funkcji pierwotnych) funkcji f obliczamy używając instrukcji

```
Integrate[f[x], x]
```

Na przykład

```
Integrate[Sin[3 x]^7, x]
```

zwraca

$$-\frac{35}{192}\cos(3x) + \frac{7}{192}\cos(9x) - \frac{7}{960}\cos(15x) + \frac{\cos(21x)}{1344}$$

13.3 Całkowanie funkcji wektorowych

Rozważmy funkcję $f : [a, b] \to \mathbb{R}^m$. Funkcję taką określać będziemy mianem *krzywej*. Wtedy $f = (f_1, f_2, \ldots, f_m)$ gdzie $f_i : [a, b] \to \mathbb{R}$ dla $i = 1, 2, \ldots, m$.

Definicja 13.3.1 Mówimy, że f jest całkowalna w sensie Riemanna jeżeli funkcje f_i są całkowalne w sensie Riemanna dla wszystkich $i = 1, 2, \ldots, m$. W takim przypadku jako całkę Riemanna f przyjmujemy

$$\int_a^b f := (\int_a^b f_1, \int_a^b f_2, \ldots, \int_a^b f_m).$$

Twierdzenie 13.3.2 *Niech $f : [a, b] \to \mathbb{R}^m$ będzie całkowalna w sensie Riemanna. Wtedy funkcja*

$$||f|| : [a, b] \ni x \mapsto ||f(x)|| \in \mathbb{R}$$

jest całkowalna w sensie Riemanna oraz

$$|| \int_a^b f|| \leqslant \int_a^b ||f||.$$

Dowód: Niech $f = (f_1, f_2, \ldots f_m)$. Z założenia funkcje f_i są całkowalne w sensie Riemanna. Ponieważ funkcja

$$\mathbb{R}^m \ni x \mapsto ||x|| \in \mathbb{R}$$

jest ciągła, zbiór punktów nieciągłości funkcji $||f||$ zawiera się w sumie mnogościowej zbiorów punktów nieciągłości funkcji f_i. Na mocy twierdzenia 12.2.9 całkowalność funkcji f_i implikuje, że zbiór punktów nieciągłości funkcji f_i jest miary zero. Zatem również zbiór punktów nieciągłości funkcji $||f||$ jest miary zero, a więc z twierdzenia 12.2.9 wnosimy, że funkcja ta jest całkowalna w sensie Riemanna.

Połóżmy $v_i := \int_a^b f_i$ i $v = (v_1, v_2, \ldots v_n)$. Jeśli $v = 0$ to teza jest oczywista. Załóżmy zatem, że $v \neq 0$. Dla $x \in [a, b]$ mamy z nierówności Schwartza

$$|v \cdot f(x)| \leqslant ||v||\ ||f(x)||.$$

Funkcja

$$[a,b] \ni x \mapsto v \cdot f(x) = \sum_{i=1}^{m} v_i f_i(x) \in \mathbb{R}$$

jest całkowalna w sensie Riemanna. Zatem

$$||v||^2 = \sum_{i=1}^{m} v_i v_i = \sum_{i=1}^{m} v_i \int_a^b f_i = \int_a^b v \cdot f \leqslant \int_a^b ||v||\ ||f(x)||dx = ||v|| \int_a^b ||f(x)||dx.$$

Dzieląc otrzymaną nierówność obustronnie przez $||v||$ otrzymujemy

$$||v|| \leqslant \int_a^b ||f(x)||dx.$$

Ponieważ $v = \int_a^b f$, dowodzi to tezy. □

13.4 Zastosowania geometryczne całek

13.4.1 Długość krzywej

Niech $\gamma : [a,b] \to \mathbb{R}^m$ będzie krzywą i niech $\mathcal{P}$ będzie podziałem przedziału $[a,b]$ zadanym ciągiem punktów $(x_0, x_1, \ldots x_n)$. Liczba

$$\Lambda(\gamma, \mathcal{P}) := \sum_{i=1}^{m} ||\gamma(x_i) - \gamma(x_{i-1})||$$

określa długość łamanej rozpiętej na punktach $\gamma(x_i)$.

Definicja 13.4.1 Mówimy, że krzywa γ jest *prostowalna* jeżeli istnieje kres górny

$$\Lambda(\gamma) := \sup \{\, \Lambda(\gamma, \mathcal{P}) \mid \mathcal{P} \text{ podział przedziału [a,b]} \,\}.$$

Mówimy wtedy, że liczba $\Lambda(\gamma)$ jest *długością* krzywej γ.

Twierdzenie 13.4.2 *Niech $\gamma : [a,b] \to \mathbb{R}^m$ będzie krzywą klasy C^1. Wtedy γ jest prostowalna oraz*

$$\Lambda(\gamma) = \int_a^b ||\gamma'(t)||dt. \tag{13.12}$$

Dowód: Pokażemy najpierw, że w (13.12) lewa strona jest niewiększa od prawej. Niech $\mathcal{P}$ będzie podziałem przedziału $[a,b]$ wyznaczonym przez punkty $x_0, x_1, \ldots x_n$. Wtedy

$$||\gamma(x_i) - \gamma(x_{i-1})|| = ||\int_{x_{i-1}}^{x_i} \gamma'(t)dt|| \leqslant \int_{x_{i-1}}^{x_i} ||\gamma'(t)dt||$$

dla $i = 1, 2, \ldots n$. Sumując po i dostajemy

$$\Lambda(\gamma, \mathcal{P}) \leqslant \int_a^b ||\gamma'(t)dt||,$$

a ponieważ $\mathcal{P}$ jest dowolnym podziałem $[a,b]$ otrzymujemy

$$\Lambda(\gamma) \leqslant \int_a^b ||\gamma'(t)||dt.$$

Dla dowodu nierówności przeciwnej ustalmy $\epsilon > 0$. Założyliśmy, że γ jest klasy C^1, więc γ' jest funkcją ciągłą. Jako funkcja ciągła na zbiorze zwartym jest ona jednostajnie ciągła. Zatem znajdziemy $\delta > 0$ takie, że dla $s,t \in [a,b]$

$$|s-t| < \delta \;\Rightarrow\; ||\gamma'(s) - \gamma'(t)|| < \epsilon' := \frac{\epsilon}{2(b-a)}. \tag{13.13}$$

Niech $\mathcal{P}$ będzie podziałem przedziału $[a,b]$ wyznaczonym przez punkty $x_0, x_1, \ldots x_n$ o średnicy nie przekraczającej δ. Z (13.13) wnosimy, że dla $t \in [x_{i-1}, x_i]$

$$||\gamma'(t)|| \leqslant ||\gamma'(t) - \gamma'(x_i)|| + ||\gamma'(x_i)|| \leqslant ||\gamma'(x_i)|| + \epsilon'. \tag{13.14}$$

Niech $\Delta x_i := x_i - x_{i-1}$. Z (13.14) mamy

$$\begin{aligned}\int_{x_{i-1}}^{x_i} ||\gamma'(t)dt|| \;\leqslant\; & \int_{x_{i-1}}^{x_i} (||\gamma'(x_i)|| + \epsilon')\, dt = ||\gamma'(x_i)||\Delta x_i + \epsilon'\Delta x_i = || \int_{x_{i-1}}^{x_i} \gamma'(x_i)dt|| + \epsilon'\Delta x_i = \\ & || \int_{x_{i-1}}^{x_i} (\gamma'(t) + \gamma'(x_i) - \gamma'(t))\, dt|| + \epsilon'\Delta x_i \leqslant \\ & || \int_{x_{i-1}}^{x_i} \gamma'(t)dt|| + || \int_{x_{i-1}}^{x_i} (\gamma'(x_i) - \gamma'(t))\, dt|| + \epsilon'\Delta x_i \leqslant \\ & ||\gamma(x_i) - \gamma(x_{i-1})|| + 2\epsilon'\Delta x_i.\end{aligned}$$

Sumując po i dostajemy

$$\int_a^b ||\gamma'(t)||dt \leqslant \Lambda(\gamma, \mathcal{P}) + 2\epsilon'(b-a) = \Lambda(\gamma, \mathcal{P}) + \epsilon \leqslant \Lambda(\gamma) + \epsilon.$$

Ponieważ otrzymana nierówność zachodzi dla każdego $\epsilon > 0$, otrzymujemy

$$\int_a^b ||\gamma'(t)||dt \leqslant \Lambda(\gamma).$$

□

Przykład 13.4.3 Rozważmy krzywą

$$\gamma : [0, 2\pi] \ni t \mapsto e^{it} \in \mathbb{R}^2.$$

Krzywa ta jest klasy C^1 oraz $\gamma'(t) = ie^{it}$. Zatem $||\gamma'(t)|| = 1$. Stąd

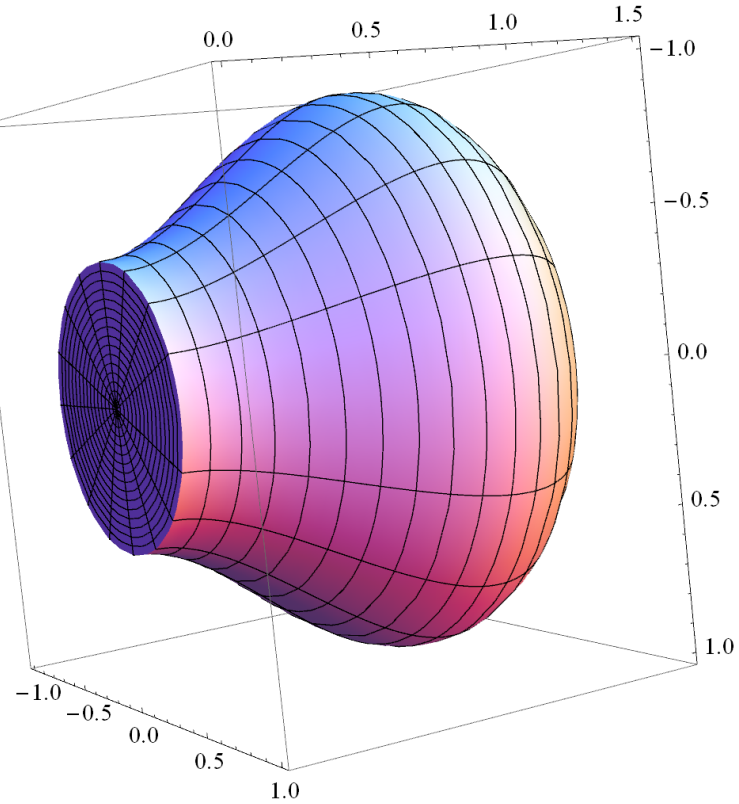
$$\Lambda(\gamma) = \int_0^{2\pi} ||\gamma'(t)||dt = \int_0^{2\pi} dt = 2\pi - 0 = 2\pi.$$

13.4.2 Bryła obrotowa

Niech $f : [a,b] \to \mathbb{R}^+$ będzie funkcją ciągłą. Rozważmy bryłę

$$V := \{\, (x,y,z) \in \mathbb{R}^3 \mid a \leqslant x \leqslant b,\; y^2 + z^2 \leqslant f(x)^2 \,\}$$

Nazywamy ją *bryłą obrotową* zadaną przez funkcję f.



Bryła ta powstaje poprzez obrót trapezu krzywoliniowego pod wykresem funkcji f wokół osi x.

13.4.3 Obętość bryły obrotowej

Twierdzenie 13.4.4 *Objętość bryły obrotowej dana jest wzorem*

$$\pi \int_a^b f(x)^2 dx.$$

13.4.4 Pole powierzchni bocznej bryły obrotowej

Powierzchnią boczną bryły obrotowej jest zbiór

$$S := \{\, (x,y,z) \in \mathbb{R}^3 \mid a \leqslant x \leqslant b,\ y^2 + z^2 = f(x)^2 \,\}$$

Twierdzenie 13.4.5 *Jeśli funkcja f jest klasy C^1, to pole powierzchni bocznej bryły obrotowej dane jest wzorem*

$$2\pi \int_a^b f(x)\sqrt{1+f'(x)^2}dx.$$

Rozdział 14

Ciągi i szeregi funkcyjne

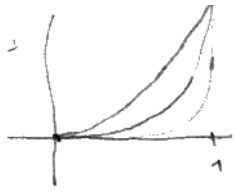
14.1 Zbieżność punktowa i jednostajna

14.1.1 Zbieżność punktowa

Definicja. - ! $X \in \mathbb{V}$, $Y \in$ Metr, $(s_n)_{n=1}^{\infty}$ ciąg funkcji, t.j. $s_n : X \to Y$. Niech $s : X \to Y$.
Mówimy, że s_n zmierza do s punktowo $:\Leftrightarrow$
$\forall x \in X \ \lim_{n\to\infty} s_n(x) = s(x) \iff \forall x \in X \ \forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall n \geq N \ \rho(s_n(x), s(x)) < \varepsilon$

Przykład 1. ! $f_n : [0,1] \ni x \to x^n \in [0,1]$, f_n ciągła, ale

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1 \end{cases}$$ – nieciągła



Przykład 2. $f_n : [0,1] \ni x \to \lim_{m\to\infty} (\cos(n!\pi x))^{2m} \in \mathbb{R}$. Wtedy

$$f_n(x) = \begin{cases} 1 & n!x \in \mathbb{N} \\ 0 & n!x \notin \mathbb{N} \end{cases}$$ czyli $f_n \in \mathrm{Riem}([0,1])$, ale

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0 & x \notin \mathbb{Q} \\ 1 & x \in \mathbb{Q} \end{cases}$$ bo $x \notin \mathbb{Q}$ to $\forall n \in \mathbb{N} \ n!x \notin \mathbb{N}$: $f_n(x) = 0$
$x \in \mathbb{Q}$ to $\exists N \in \mathbb{N} \ \forall n \geq N \ n!x \in \mathbb{N}$: $f_n(x) = 1$

W granicy dostajemy więc funkcję Dirichleta, która nie jest całkowalna w sensie Riemanna.

Przykład 3. $f_n : \mathbb{R} \ni x \to \frac{\sin nx}{\sqrt{n}} \in 0$, $\lim_{n\to\infty} f_n(x) = 0$
$f_n'(x) = \sqrt{n} \cos nx$: nie jest zbieżny do zera.

14.1.2 Zbieżność jednostajna

Definicja Metryki ρ, ρ' na Y są jednostajnie równoważne :$\Leftrightarrow$
$\forall \varepsilon>0\ \exists \delta>0 \quad \rho(x,y)<\delta \Rightarrow \rho'(x,y)<\varepsilon$: $\forall \varepsilon>0\ \exists \delta>0\ \rho'(x,y)<\delta \Rightarrow \rho(x,y)<\varepsilon$

Definicja 14.1.1 Niech (Y,ρ) będzie przestrzenią metryczną. Mówimy, że ρ jest *metryką ograniczoną*, jeżeli istnieje stała $M>0$ taka, że dla dowolnych $x,y\in Y$ mamy $\rho(x,y)<M$.

Lemat. $\forall (Y,\rho)\in$ Metr. Istnieje metryka $\bar{\rho}$ jednost. równoważna metryce ρ i ograniczona

Definicja. $\forall X\in \mathbb{V}$, $\forall (Y,\rho)\in$ Metr, możemy przyjąć, że ρ - ograniczona.
Niech $F(X,Y) := \{f: X\to Y\}$ || Inne ozn.: Y^X
Na $F(X,Y)$ wprowadzamy metrykę wzorem
$\rho(f,g) := \sup\{\rho(f(x),g(x)) : x\in X\}$ - dobrze określ., bo ρ - ogr.
Topologię daną tą metryką nazywamy topologią jednostajną na $F(X,Y)$, a bieżność w tej topologii - zbieżnością jednostajną.

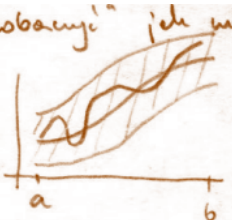
Przykład. $\{f_n\}_1^\infty \subset F(X,Y)$, $f\in F(X,Y)$

$\lim\limits_{n\to\infty} f_n = f$ (jednost.) $\Leftrightarrow \forall \varepsilon>0\ \exists N\in\mathbb{N}\ \forall n\geq N \quad \rho(f_n,f)<\varepsilon$
$\Leftrightarrow \forall \varepsilon>0\ \exists N\in\mathbb{N}\ \forall n\geq N \quad \sup\limits_{x\in X} \rho(f_n(x),f(x))<\varepsilon$
$\Leftrightarrow \forall \varepsilon>0\ \exists N\in\mathbb{N}\ \forall n\geq N\ \forall x\in X \quad \rho(f_n(x),f(x))<\varepsilon$
$\Leftrightarrow \lim\limits_{n\to\infty} \sup\limits_{x\in X} \rho(f_n(x),f(x)) = 0$

Wniosek 1. f_n zbieżny jednost. to f_n zbieżny punktowo.

Wniosek 2. $\forall \{f_n\}_1^\infty \subset F(X,Y)$, $f\in F(X,Y)$, $\varphi: Z\to X$. Wtedy
$\lim\limits_{n\to\infty} f_n = f$ (jednost.) $\Rightarrow \lim\limits_{n\to\infty} f_n\circ\varphi = f\circ\varphi$ (jednost.)

Warto „zobaczyć" jak wygląda kula w metryce top. jednost.:



Warto zauważyć, że zb. jednost. tym się różni od punktowej, że żądamy możliwości uniwersalnego doboru δ do ε niezależnie od x.

Jeśli $V \in NVect$ to $F(X,V)$ ma naturalną strukturę przestrzeni wektorowej jednakże normy nie opór nie da się przenieść w sposób prosty z V na $F(X,V)$, dlatego, że zawsze musimy się dla podprzestrzeni funkcji ograniczonych

Definicja $\forall V \in NVect$, $B(X,V) := \{f : X \to V : f \text{ - ograniczona}\}$.
Wtedy $B(X,V)$ jest podprzestrzenią $F(X,V)$ oraz
$\|f\| := \sup_{x \in X} \|f(x)\|$ jest normą na $B(X,V)$, otrzymujemy więc przestrzeń wektorową unormowaną.

Twierdzenie. Jeśli Y zupełne to $F(X,Y)$ zupełne.

Oznaczenie $ZMetr := \{X : X \text{ - metryczne, zupełne}\}$.

Twierdzenie. $\forall X \in ZMetr$, $Y \subset X$. Wtedy:
$Y \in ZMetr \iff Y \in Clos(X)$.

14.1.3 Kryterium Cauchy'ego zbieżności jednostajnej

Twierdzenie (WKW Cauchy'ego zbieżności jednostajnej)
$\forall X \in \mathbb{V}$, $Y \in ZMet$ $\{f_n\} \subset F(X,Y)$. Wtedy
f_n zbieżny jednost. $\iff \forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n,m \geq N \; \rho(f_n, f_m) < \varepsilon$
$\iff \forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n,m \geq N \; \forall x \in X \; \rho(f_n(x), f_m(x)) < \varepsilon$ ∎

Twierdzenie. ! $X \in \mathbb{V}$. ! $V \in NVect$ i V zupełne. Wtedy $B(X,V)$ jest zupełne (Inaczej: granica ciągu jedn. zbieżnego funkcji ograniczonych jest ograniczona). ■

Wniosek. (kryterium porównawcze (Weierstrassa) zbieżn. jednostajnej). ! $\{f_n\} \subset B(X,V)$, ! $M: \mathbb{N} \to \mathbb{R}$. Jeśli $\forall x \in X$ $\|f_n(x)\| \leq M_n$ i szereg $\sum M_n$ zbieżny to $\sum f_n$ zbieżny jednostajnie.

Uwaga. $\{f_n\}, \{g_n\} \subset B(X,V)$ $\alpha, \beta \in \mathbb{R}$, $f_n \to f$, $g_n \to g$ jedn. to $\alpha f_n + \beta g_n \to \alpha f + \beta g$ jedn.

14.2 Zbieżność jednostajna, a ciągłość, różniczka i całka

14.2.1 Granica, a zbieżność jednostajna

! $X \in Top$. ! $Y \in ZMetr$

Twierdzenie. ! $\{f_n\} \subset F(X,Y)$. $f \in F(X,Y)$ i $f_n \to f$ jedn. ! $x_0 \in X'$ i $\lim_{x \to x_0} f_n(x) = a_n$. Wtedy:

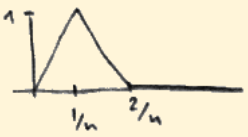
$\{a_n\}_1^\infty$ zbieżny, istnieje $\lim_{x \to x_0} f(x)$ oraz $\lim_{n \to \infty} a_n = \lim_{x \to x_0} f(x)$

Inaczej: $\lim_{n \to \infty} \lim_{x \to x_0} f_n(x) = \lim_{x \to x_0} \lim_{n \to \infty} f_n(x)$.

Wniosek. $\{f_n\} \subset C(X,Y)$ i $f = \lim_{n \to \infty} f_n$ (jedn.). Wtedy $f \in C(X,Y)$

Inaczej: $C(X,Y)$ jest dkn. (zup.) podpn. $F(X,Y)$

Uwaga. Jest to jedynie warunek konieczny ciągłości granicy:
W pewnym szczeg. przypadku prawdziwe jest twierdzenie odwrotne:



$f_n \to 0$ nie jednost.

Twierdzenie. $\forall K \in \mathrm{Comp}$, $\{f_n\} \subset C(K,Y)$, $f_n \geq f_{n+1}$, $f \in C(K,Y)$:
$\forall x \in K \ \lim_{n\to\infty} f_n(x) = f(x)$ (zbieżn. punktowa).
$\Longrightarrow$ $\lim f_n = f$ jednost.

14.2.2 Całkowanie i różniczkowanie, a zbieżność jednostajna

Twierdzenie. $\forall Q \in \mathrm{Prodz}(\mathbb{R}^m)$, $\forall \{f_n\}_1^\infty \subset \mathrm{Riem}(Q,\mathbb{R}^k)$
Jeśli $f_n \to f \in F(Q,\mathbb{R}^k)$ jednost. to $f \in \mathrm{Riem}(Q,\mathbb{R}^k)$: $\int f = \lim_{n\to\infty} \int f_n$

Wniosek. $\{f_n\}_0^\infty \subset \mathrm{Riem}(Q,\mathbb{R}^k)$: $f(x) = \sum_{n=0}^{\infty} f_n(x)$ (zbieżność jednost.) to
$f \in \mathrm{Riem}(Q,\mathbb{R}^k)$: $\int f = \sum_{n=0}^{\infty} \int f_n$ ■

Twierdzenie. $\forall \{f_n\}_1^\infty$ gdzie $f_n : [a,b] \to \mathbb{R}^k$ będzie ciągiem funkcji różniczkowalnych takim, że dla pewnego $x_0 \in [a,b]$ $(f_n(x_0))_1^\infty$ zbieżny oraz ciąg pochodnych f_n' jest zbieżny jednostajnie. Wtedy sam ciąg f_n jest zbieżny jednostajnie do pewnej funkcji różniczkowalnej f oraz
$f'(x) = \lim_{n\to\infty} f_n'(x)$

14.3 Aproksymacja funkcji ciągłych

Twierdzenie (Weierstrassa).
$\forall f \in C([a,b], \mathbb{R}) \Rightarrow \exists P_n : [a,b] \to \mathbb{R}$ ciąg wielomianów taki, że
$\lim_{n\to\infty} P_n|_{[a,b]} = f$ (jednost.)

14.4 Szeregi potęgowe

Definicja Szeregiem potęgowym o środku w $z_0 \in \mathbb{C}$ nazywamy szereg $\sum_{n=0}^{\infty} c_n (z-z_0)^n$ (uogólnienie pojęcia definicyjnego)

Uwaga. Jeśli $R = 1/\limsup \sqrt[n]{|c_n|}$ to dla
$z \in K(z_0, R)$ - szereg pot. zb.
$z \notin \overline{K}(z_0, R)$ - - - rozb.
$z \in S(z_0, R)$ - problem zbieżności nie rozstrzygnięty ■

Twierdzenie. Niech promień zbieżności szeregu $\sum_{n=0}^{\infty} c_n (x-x_0)^n$ wynosi $R > 0$ i niech
$f(x) = \sum_{n=0}^{\infty} c_n (x-x_0)^n$ dla $x \in (x_0-R, x_0+R)$
Wtedy: 1) $\forall r \in (0,R)$ pow. szereg jest zbieżny jednost. w $[x_0-r, x_0+r]$
2) f ciągła i różn. na (x_0-R, x_0+R) oraz $f'(x) = \sum_{n=0}^{\infty} n c_n (x-x_0)^{n-1}$

Wniosek. Przy założeniach poprzedniego twierdzenia:
$f \in C^\infty((x_0-R, x_0+R), \mathbb{R})$ oraz $f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1) c_n (x-x_0)^{n-k}$
(szereg potęgowy można różniczkować wyraz po wyrazie)
W szczególności: $f^{(n)}(x_0) = n!\, c_n$ ■

Twierdzenie. $\forall a : \mathbb{N}\times\mathbb{N} \to \mathbb{R}$, $\exists \sum_{j=1}^{\infty} |a_{ij}| = b_i$ i niech $\sum_{i=1}^{\infty} b_i$ zbieżny. Wtedy
$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}$.

14.4.1 Twierdzenie o szeregu Taylora

Twierdzenie (Taylora). Niech szereg $\sum_{n=0}^{\infty} c_n (x-x_0)^n$ będzie zbieżny w (x_0-R, x_0+R) i niech $f(x)$ oznacza sumę tego szeregu. Jeśli $x_1 \in (x_0-R, x_0+R)$ to funkcję f można przedstawić w postaci sumy szeregu o środku w x_1, który jest zbieżny dla $x \in (x_1-r, x_1+r)$ gdzie $r = R-|x_1-x_0|$.

14.4.2 Funkcje analityczne

Twierdzenie. Załóżmy, że $\sum_{n=0}^{\infty} a_n (x-x_0)^n$ i $\sum_{n=0}^{\infty} b_n (x-x_0)^n$ zbieżne w (x_0-R, x_0+R). Połóżmy: $E := \{x \in (x_0-R, x_0+R) : \sum_{n=1}^{\infty} a_n (x-x_0)^n = \sum_{n=1}^{\infty} b_n (x-x_0)^n\}$. Niech $E' \neq \emptyset$. Wtedy $\forall n \in \mathbb{N}$ $a_n = b_n$, t.j. $E = (x_0-R, x_0+R)$.

Definicja. Funkcję $f: (a,b) \to \mathbb{R}$ nazywamy analityczną $:\Leftrightarrow$ $\forall x_0 \in (a,b)$ f daje się przedstawić w postaci sumy szeregu potęgowego o pewnym otoczeniu x_0.

Uwaga. Funkcją analityczną w przedziale (x_0-R, x_0+R) jednoznacznie wyznaczają pochodne tej funkcji w punkcie x_0.

Przykład. $f(x) := \begin{cases} e^{-\frac{1}{x}} & x \neq 0 \\ 0 & x = 0. \end{cases}$

Można sprawdzić, że $\forall n \in \mathbb{N}_0$ $f^{(n)}(0) = 0$. Zatem gdyby f była analityczna, musiałaby znikać w pewnym otoczeniu zera.

Uwaga. Rozwinięcie funkcji analitycznej w szereg w punkcie x_0 jest szeregiem Taylora tej funkcji w x_0, t.j.

f analit. $\Rightarrow$ $\forall x_0 \in \operatorname{dom} f$ $f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x-x_0)^n$ $\square$



Rysunek 14.1: Joseph Fourier (1768-1830)

Uwaga. Zbiór zer funkcji analitycznej nie posiada punktów skupienia, a dwie f. analityczne równe w pewnym zb. mającym punkt skupienia są równe □.

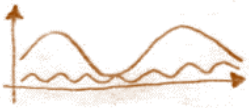
Uwaga. Jeśli f daje się przedstawić w postaci szeregu potęgowego w x_0 zbieżnego w (x_0-R, x_0+R) to f analityczna w tym przedziale □.

Wniosek. Wielomiany, funkcje wykładnicze, funkcje trygonometryczne, to funkcje analityczne w całym $\mathbb{R}$ □

14.5 Szeregi Fouriera

W teorii drgań już w XVIII w. pojawia się zagadnienie rozłożenia danej funkcji okresowej na sumę sinusów i cosinusów (drgania proste).





Bernoulli i Euler podawali przykłady takich rozkładów i podejrzewali, że rozkład taki jest zawsze możliwy. Dowód tego faktu pochodzi od G. Fouriera.

14.5.1 Wielomiany trygonometryczne

Definicja. Wielomianem trygonometrycznym nazywamy funkcję postaci

$$T: \mathbb{R} \ni x \longrightarrow \sum_{n=0}^{N} (a_n \cos nx + b_n \sin nx) = a_0 + \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx) \in \mathbb{C} \qquad a_n, b_n \in \mathbb{C}$$

Uwaga 1. Wielomian trygonometryczny można przedstawić w postaci

$$T(x) = \sum_{n=-N}^{N} c_n e^{inx} \qquad \text{gdzie} \qquad c_n = \begin{cases} \frac{a_n - ib_n}{2} & n > 0 \\ a_0 & n = 0 \\ \frac{a_{-n} + ib_{-n}}{2} & n < 0 \end{cases} \qquad a_n = c_n + c_{-n} \qquad b_n = \frac{c_{-n} - c_n}{i}$$

Uwaga 2. Współczynniki wielomianu trygonometrycznego T wyrażają się wzorem:

$$c_m = \frac{1}{2\pi} \int_0^{2\pi} T(x) e^{-imx} dx$$

Uwaga 3. Wielomian T jest rzeczywisty $\Leftrightarrow \forall m = 0, 1, \ldots N \quad c_{-m} = \overline{c_m}$

14.5.2 Zespolony iloczyn skalarny

Definicja. $V \in \mathrm{Vect}(\mathbb{C})$. Odwzorowanie $\varphi: V \times V \to \mathbb{C}$ nazywamy iloczynem skalarnym (zespolonym) w V jeżeli zachodzą następujące warunki

ISZ1) $\varphi(x+y,z) = \varphi(x,z) + \varphi(y,z)$

ISZ2) $\varphi(\alpha x, y) = \alpha\, \varphi(x,y)$

ISZ3) $\varphi(x,y) = \overline{\varphi(y,x)}$

ISZ4) $\varphi(x,x) \geq 0$; $\varphi(x,x) = 0 \iff x = 0$

Wniosek. 1) $\varphi(z, x+y) = \varphi(z,x) + \varphi(z,y)$

2) $\varphi(x, \alpha y) = \overline{\alpha}\, \varphi(x,y)$

Przykład. Określamy odwzorowanie $\varphi: \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ wzorem

$$\varphi(x,y) := \sum_{i=1}^{n} x_i \overline{y_i} \qquad x = (x_1, \dots x_n) \in \mathbb{C}^n; \quad y = (y_1, \dots y_n) \in \mathbb{C}^n$$

Warunek ISZ3 wywodzi się wprost z tego przykładu. Po to wstawiamy $\overline{y_i}$, by w szczególnym przypadku dostać:

$\varphi(x,x) = \sum_{i=1}^{n} x_i \overline{x_i} = \sum_{i=1}^{n} |x_i|^2$, czyli to samo co w przypadku nezwistym.

14.5.3 Przestrzeń wektorowa z zespolonym iloczynem skalarnym

Twierdzenie $|\varphi(x,y)| \leq \sqrt{\varphi(x,x)}\,\sqrt{\varphi(y,y)}$

Wniosek. Przestrzeń wektorowa z zespolonym iloczynem skalarnym jest przestrzenią wektorową unormowaną z normą daną przez

$$\|x\|_\varphi := \sqrt{\varphi(x,x)}$$

Przykład. $Riem([a,b],\mathbb{C})$ jest przestrzenią wektorową nad $\mathbb{C}$. Określony odwzorowanie

$$\varphi: Riem([a,b],\mathbb{C}) \times Riem([a,b],\mathbb{C}) \ni (f,g) \longrightarrow \beta \int_a^b f\bar{g} \quad , \text{ gdzie } \beta \in \mathbb{R}^+$$

Spełnia ono własności ISZ1 - ISZ3, a w ISZ4 nie spełnia jedynie implikacji „$\Rightarrow$"

‖ **Przykład.**

14.5.4 Układ ortonormalny

Lekarstwem jest zmodyfikowanie przestrzeni. Określamy dla $f,g \in Riem([a,b],\mathbb{C})$:

$$f \sim g :\Leftrightarrow \mu^*(\{x : f(x) \neq g(x)\}) = 0$$

Uwaga 1. $\sim$ jest relacją równoważności w $\mathrm{Riem}([a,b],\mathbb{C})$

2. $([f],[g]) \to [f]+[g] := [f+g]$
   $(c,[f]) \to c[f] := [cf]$
   są dobrze określonymi działaniami w przestrzeni ilorazowej i zadają w niej strukturę przestrzeni wektorowej (dlaczego?).
   Oznaczmy ją $\mathrm{Riem}_0([a,b],\mathbb{C})$

3. $\tilde{\varphi}: ([f],[g]) \to \int_a^b f\bar{g}$ jest dobrze określonym iloczynem skalarnym (spełnia różnicę ISZ4) ■

Uwaga. Poniższy fragment teorii dotyczy przestrzeni $\mathrm{Riem}_0([a,b],\mathbb{C})$. Nie będziemy jednak explicite wypisywać klas $[f],[g]$, lecz całe klasy oznaczać literami f,g i będziemy o nich mówić po prostu funkcje f,g, chyba że wyraźnie zaznaczymy, że chodzi o funkcję w zwykłym sensie.

Normę pochodzącą od iloczynu skalarnego $\tilde{\varphi}$ oznaczać będziemy $\|\cdot\|_2$, tak więc:

$$\|f\|_2 := \int_a^b f\bar{f} = \int_a^b |f|^2.$$

Sam iloczyn skalarny $\tilde{\varphi}(f,g)$ oznaczać będziemy dla krótkości $(f|g)$. Tak więc:

$$(f|g) := \int_a^b f\bar{g}$$

Definicja. Ciąg $\{\psi_n\}_1^\infty \subset \mathrm{Riem}_0([a,b],\mathbb{C})$ nazywamy układem ortonormalnym $:\Leftrightarrow$

$$(\psi_n|\psi_m) = \begin{cases} 0 & n \neq m \\ 1 & n = m \end{cases}$$

Przykład. Ciągami ortonormalnymi na $[-\pi, \pi]$ są:

$$\left\{\left(x \to \frac{1}{\sqrt{2\pi}} e^{inx}\right)\right\}_{n=-\infty}^{\infty} \text{ oraz } \left\{\left(x \to \frac{1}{2\pi}\right), \left(x \to \frac{1}{\sqrt{\pi}} \cos x\right), \left(x \to \frac{1}{\sqrt{\pi}} \sin x\right), \left(x \to \frac{1}{\sqrt{\pi}} \cos 2x\right), \ldots\right\}$$

14.5.5 Szereg Fouriera

Definicja. $f \in \mathrm{Riem}_0([a,b], \mathbb{C})$ $\{\varphi_n\} \subset \mathrm{Riem}_0([a,b], \mathbb{C})$ – ciąg ortonormalny. Wtedy

$$c_n := (f|\varphi_n) = \int_a^b f\bar{\varphi}_n$$

nazywamy n-tym współczynnikiem Fouriera funkcji f względem układu $\{\varphi_n\}$, a szereg formalny (funkcyjny)

$$\sum c_n \varphi_n$$

szeregiem Fouriera funkcji f względem układu $\{\varphi_n\}$.

Twierdzenie. Niech $\{\varphi_n\}$ będzie układem ortonormalnym i niech

$$s_n = \sum_{m=0}^{n} c_m \varphi_m \quad \text{gdzie } c_m \text{ – } m\text{-ty współ. Fouriera elle } f.$$

Jeśli $t_n = \sum_{m=0}^{n} \gamma_m \varphi_m$ i $\gamma_m \in \mathbb{C}$ to

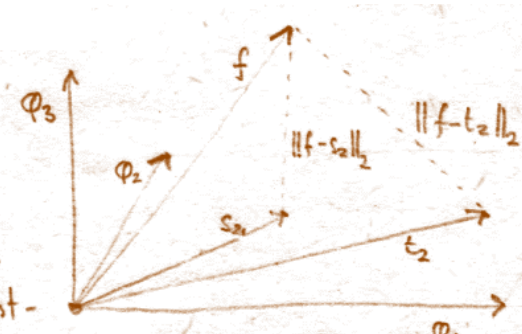
$$\int_a^b |f - s_n|^2 \leq \int_a^b |f - t_n|^2 \quad \text{tj.} \quad \|f - s_n\|_2 \leq \|f - t_n\|_2$$

oraz równość zachodzi tylko wtedy gdy $\gamma_m = c_m$ dla $m = 0, 1, 2, \ldots$

Ponadto:

$$\sum_{m=1}^{\infty} |c_m|^2 \leq \int_a^b |f|^2 = \|f\|_2^2 \quad \text{(nierówność Bessela)}$$

W szczególności $\lim_{n\to\infty} c_n = 0$.



Interpretacja geometryczna: s_n jest rzutem prostokątnym f na hiperpł. wym. przez $\varphi_1, \dots \varphi_n$. Rzut ten jest położony najbliżej f spośród wszystkich wektorów hiperpłaszczyzny. Nierówność Bessela stanowi „niekompletne twierdzenie Pitagorasa (Brak równości, bo ortonormalność układu $\{\varphi_n\}$ jeszcze nie gwarantuje, iż wyczerpuje on wszystkie możliwe kierunki.

Uwaga. W szczeg. pokaz., że: $\|s_n\|_2^2 = \sum_{k=0}^{n} |c_k|^2 \leq \|f\|_2^2$ □

W dalszym ciągu rozważać będziemy przypadek szczególny: $\mathrm{Riem}_0([-\pi,\pi],\mathbb{C})$ oraz iloczyn skalarny dany przez:

$$(f|g) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f\bar{g}$$

Układem ortonormalnym jest wtedy (n.p.) rodzina $\{E_n : n \in \mathbb{Z}\}$, gdzie

$$E_n : [-\pi,\pi] \ni x \longrightarrow e^{inx} \in \mathbb{C}.$$

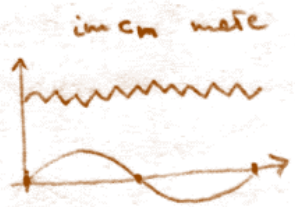
Funkcje okresowe o okresie 2π i całkowalne w sensie Riemanna na $[-\pi,\pi]$ można utożsamiać (przez zawężenie do $[-\pi,\pi]$) z elementami $\mathrm{Riem}_0([-\pi,\pi],\mathbb{C})$. Będziemy to często robić nie wspominając o tym wyraźnie.

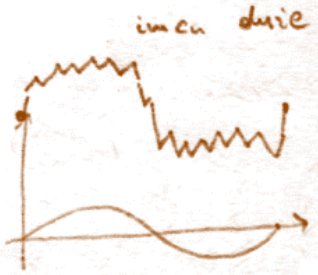
Szereg Fouriera względem układu $\{E_n\}$ będziemy nazywali krótko szeregiem Fouriera.

Tak więc n-ty współczynnik Fouriera funkcji okresowej o okresie 2π dany jest wzorem:

$$c_n(f) := c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt - \frac{i}{2\pi}\int_{-\pi}^{\pi} f(t)\sin nt\,dt$$

Interpretacje wzoru:
im c_n wyjdzie duże jeżeli f ma kształt zbliżony do kształtu wagi.





14.5.6 Rzut ortonormalny

Oznaczenie. $C_n(f): [-\pi,\pi] \ni x \to c_n e^{inx} \in \mathbb{C}$ — rzut prostokątny funkcji f w kierunku E_n.

Uwaga 1: Jeśli f jest funkcją rzeczywistą, to

$$C_n(f) + C_{-n}(f) = a_n \cos nx + b_n \sin nx$$

jest funkcją rzeczywistą. Jest to rzut prostokątny f na „płaszczyznę" wyznaczoną przez „wektory" $(x \to \sin nx), (x \to \cos nx)$, t.j. składnik postaci $a_n \cos nx + b_n \sin nx$ w trygonometrycznej postaci szeregu Fouriera.

Wtedy $a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx$ $\quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx$

Uwaga 2. $C_n(f)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-s)e^{ins}ds$

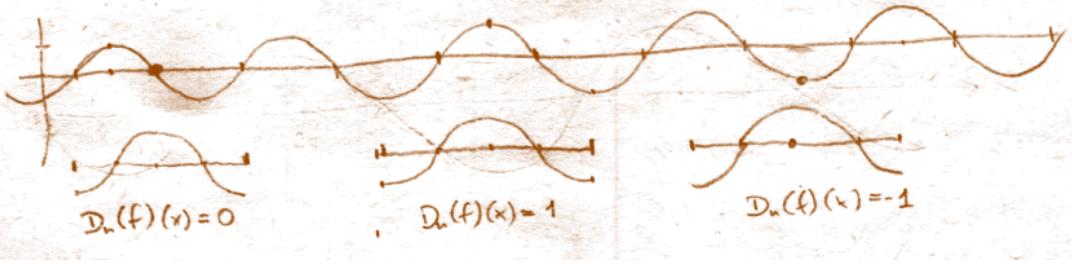
W szczególności:

$$\underbrace{[C_n(f)+C_{-n}(f)]}_{D_n(f)}(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-s)\left[e^{ins}+e^{-ins}\right]ds = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x-s)\cos ns\, ds = \frac{1}{\pi} f * \cos_n(x)$$

splot

Oznacza to, że rzut f na „płaszczyznę" częstotliwości n (t.j. wyznaczoną przez $\sin nx, \cos nx$) w punkcie x jest duży jeśli funkcja „współgra" z ($s \to \cos ns$) przesuniętym do x:

$a_n \sin nx + b_n \cos nx$



Oznaczenie.

$F_n(f): [-\pi,\pi] \ni x \longrightarrow \sum_{n=-N}^{N} c_n e^{inx} = \sum_{n=-N}^{N} C_n(f)(x) \in \mathbb{C}$ n-ta suma częśc. sz. F.

$F(f): [-\pi,\pi] \ni x \longrightarrow \sum_{n=-\infty}^{\infty} c_n e^{inx} \in \mathbb{C}$ – szereg Fouriera dla f.

14.5.7 Jądro Dirichleta

Definicja. N-tym jądrem Dirichleta nazywamy funkcję:

$$D_N: [-\pi,\pi] \ni x \longrightarrow \sum_{n=-N}^{N} e^{inx} \in \mathbb{R} \quad (!)$$

Uwaga 3.

$$F_N(f)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-s)\, D_N(s)\, ds = f * D_N(x)$$

splot

Uwaga 4. a) $D_N(x) = \begin{cases} \dfrac{\sin\left(N+\frac{1}{2}\right)x}{\sin\frac{1}{2}x} & x \neq 0 \quad (x \neq 2k\pi) \\ 2N+1 & x = 0 \quad (x = 2k\pi) \end{cases}$

14.5.8 Zbieżność punktowa szeregu Fouriera

Twierdzenie. (O zbieżności punktowej szeregu Fouriera)
! $f: \mathbb{R} \to \mathbb{C}$ okresowa o okresie 2π i całk. w sensie Riem. na $[-\pi, \pi]$, ! $x \in \mathbb{R}$: $\exists \delta > 0 \; \exists M > 0$:

$$|f(x+t) - f(x)| < M|t| \quad \text{dla } t \in (-\delta, \delta)$$

Wtedy: $\lim_{N \to \infty} F_N(f)(x) = f(x)$

Założenie to oznacza, że w punkcie x wykres musi się mieścić w pewnym stożku:



a nie odpadają przypadki





Wniosek 1. Jeśli $f(x) = 0$ dla $x \in (a,b) \subset [-\pi, \pi]$ to $F(f)(x) = 0$ dla $x \in (a,b)$

Wniosek 2. $f(x) = g(x)$ dla $x \in (a,b)$ to $\lim_{n \to \infty} [F_n(f) - F_n(g)](x) = 0$ dla $x \in (a,b)$

Przykład. Rozwińmy w szereg Fouriera w przedziale $[-\pi, \pi]$ funkcję $y = x$:

$c_0 = \int_{-\pi}^{\pi} x\,dx = 0$. Dla $n \neq 0$ jest:

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} x e^{-inx}dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(\cos nx - i\sin nx)dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} x\cos nx\,dx - i\frac{1}{2\pi}\int_{-\pi}^{\pi} x\sin nx\,dx$$

Ale:

$$\int_{-\pi}^{\pi} x\cos nx\,dx = \int_{-\pi}^{\pi} x\frac{d\sin nx}{n} = x\frac{\sin nx}{n}\Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi}\frac{\sin nx}{n}dx = 0 - \frac{1}{n^2}\int_{-\pi}^{\pi}\sin nx\,d(nx) = \frac{1}{n^2}[\cos nx]_{-\pi}^{\pi} = 0$$

$$\int_{-\pi}^{\pi} x\sin nx\,dx = \int_{-\pi}^{\pi} x\,d\left(\frac{-\cos nx}{n}\right) = \frac{-x\cos nx}{n}\Big|_{-\pi}^{\pi} + \int_{-\pi}^{\pi}\frac{\cos nx}{n^2}d(nx) = \frac{-[\pi\cos n\pi + \pi\cos n\pi]}{n} + \left[\frac{\sin nx}{n^2}\right]_{-\pi}^{\pi} =$$

$$= \frac{-2\pi\cos n\pi}{n} = \frac{-2\pi(-1)^{|n|}}{n}$$

Zatem: $c_n = -i\frac{1}{2\pi}\cdot\frac{-2\pi(-1)^{|n|}}{n} = \frac{(-1)^{|n|}i}{n}$

Mamy więc $F(f)(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} = \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty}\frac{(-1)^{|n|}i}{n}(\cos nx + i\sin nx) =$

$$\sum_{n=1}^{\infty}\left(\frac{(-1)^n i}{n}(\cos nx + i\sin nx) + \frac{(-1)^n i}{-n}(\cos(-nx) + i\sin(-nx))\right) = \sum_{n=1}^{\infty}\frac{2(-1)^{n+1}}{n}\sin nx =$$

$$= 2\left(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \dots\right)$$

Z ostatniego twierdzenia wnosimy, że szereg ten jest zbieżny dla $x \in (-\pi, \pi)$

Zatem $\frac{x}{2} = \sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \dots$ dla $x \in (-\pi, \pi)$

Wszczególności dla $x = \frac{\pi}{2}$ otrzymamy: $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \dots$ □

Twierdzenie Parsevala. Niech $f, g: \mathbb{R} \to \mathbb{C}$ okresowe o okresie 2π i całkowalne w sensie Riemanna. Wtedy:

1) $\lim_{n\to\infty} \|F_n(f) - f\|_2 = 0$ tj. $F_n(f) \to f$ w sensie normy $\|\cdot\|_2$

2) $(f|g) = \sum_{n=-\infty}^{\infty} c_n(f)\overline{c_n(g)}$

3) $\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |c_n(f)|^2$ (wzór Lapunowa)

Bibliografia

[1] F. Leja, Rachunek różniczkowy i całkowy, Państwowe Wydawnictwo Naukowe, Warszawa 1969.

[2] W. Rudin, Podstawy analizy matematycznej, Państwowe Wydawnictwo Naukowe, Warszawa 1982.

[3] W. Rudnicki, Wykłady z analizy matematycznej, Państwowe Wydawnictwo Naukowe, Warszawa 2001.

[4] G.M. Fichtenholz, Rachunek różniczkowy i całkowy, tom I, II i III. Państwowe Wydawnictwo Naukowe, Warszawa 1978.

[5] W. Krysicki, L. Włodarski, Analiza matematyczna w zadaniach, część I i II, Państwowe Wydawnictwo Naukowe, Warszawa 1986.

[6] J. Banaś, S. Wędrychowicz, Zbiór zadań z analizy matematycznej, Wydawnictwo Naukowo-Techniczne, Warszawa 2001.