Analiza Matematyczna

Adrian Madajewski

Niniejszy plik jest w całości bazowany na wykładach

prof. dr hab. Dariusza Bugajewskiego

z przedmiotu

Analiza Matematyczna

na Uniwersytecie im. Adama Mickiewicza w Poznaniu

UNIWERSYTET
IM. ADAMA MICKIEWICZA
W POZNANIU

8 Całka Riemanna

8.1 Definicja i podstawowe własności całki

Definicja 68. Niech $[a, b]$ będzie danym przedziałem. Przez podział P przedziału $[a, b]$ będziemy nazywali skończony zbiór punktów $x_0, x_1, \dots, x_n$, gdzie

$$a = x_0 < x_1 < \dots < x_n = b$$

Będziemy pisać $\Delta x_i = x_i - x_{i-1}$ $(i = 1, \dots, n)$. Długość największego z odcinków $[x_{i-1}, x_i]$ nazywać będziemy średnicą podziału P i oznaczamy ją symbolem $\delta(P)$. $\delta(P) = \max\limits_{1 \leqslant i \leqslant n} \Delta x_i$. Niech f będzie ograniczoną funkcją rzeczywistą określoną na $[a, b]$. W każdym z przedziałów $[x_{i-1}, x_i]$ wybierzmy dowolny punkt ξ_i $(i = 1, \dots n)$ i utwórzmy sumę $R = \sum_{i=1}^{n} f(\xi_i)\Delta x_i$. Sumę te nazywamy sumą Riemanna odpowiadającą podziałowi P, przy ustalonym wyborze punktów ξ_i. Przez $\mathfrak{R}(f, P)$ oznaczać będziemy zbiór wszystkich możliwych sum Riemanna odpowiadających podziałowi P. Utwórzmy teraz ciąg (P_k) podziałów przedziału $[a, b]$:

$$a = x_0^{(k)} < x_1^{(k)} < \dots < x_{n(k)}^{(k)} = b;$$

$$\Delta_i^{(k)} = x_i^{(k)} - x_{i-1}^{(k)};$$

$$\delta(P_k) = \max_{1 \leqslant i \leqslant n(k)} \Delta x_i^{(k)}, k = 1, 2, \dots$$

Ciąg (P_k) nazywamy ciągiem normalnym podziałów, jeśli $\delta(P_k) \to 0$ przy $k \to \infty$. Oznaczmy przez $\mathfrak{R}(f, P_k)$ zbiór wszystkich sum Riemanna odpowiadających podziałowi P_k.

Definicja 69. Jeśli dla dowolnego ciągu normalnego podziałów (P_k) i dla dowolnych sum Riemanna $R_k \in \mathfrak{R}(f, P_k)$ istnieje skończona granica $I = \lim\limits_{k \to \infty} R_k$, to tę granicę nazywamy całką Riemanna funkcji f na przedziale $[a, b]$ i oznaczamy ją symbolem

$$\int_a^b f dx \text{ lub } \int_a^b f(x) dx$$

O funkcji f mówimy wówczas, że jest całkowalna w sensie Riemanna na przedziale $[a, b]$, lub że jest ona R-całkowalna na tym przedziale.

Powyższą definicję mozna sformułować w następujący równoważny sposób.

Definicja 70. Mówimy, że funkcja f jest całkowalna w sensie Riemanna na przedziale $[a, b]$, jeśli istnieje liczba $I \in \mathbb{R}$ taka, że

$$\forall_{\varepsilon > 0} \exists_{\delta > 0} \forall_P \forall_{R \in \mathfrak{R}(f,P)} \delta(P) < \delta \implies |R - I| < \varepsilon$$

Piszemy wówczas $I = \int_a^b f(x)dx = \lim\limits_{\delta(P) \to 0} R$.

Równoważność definicji 69 i 70 można pokazać analogicznie jak w dowodzie twierdzenia 35.

Przykład 25. (a) Funkcja stała $f(x) = c, c \in \mathbb{R}, x \in [a, b]$ jest całkowalna w sensie Riemanna na tym przedziale. Niech P będzie dowolnym podziałem przedziału $[a, b]$:

$$a = x_0 < x_1 < \cdots < x_n = b$$

Dowolna suma Riemanna odpowiadającą podziałowi P ma postać:

$$R = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) = \sum_{i=1}^{n} c(x_i - x_{i-1}) = c(b - a),$$

$$(\xi_i \in [x_{i-1}, x_i], i = 1, \dots, n)$$

Stąd wynika, że $\int_a^b f(x)dx = c(b - a)$.

(b) Roważmy ponownie funkcję Dirichleta z Przykładu 18 (a), zawężoną do przedziału $[a, b]$. Dla każdego podziału P przedziału $[a, b]$ można utworzyć sumę Riemanna równą zeru, jeśli wszystkie punkty ξ_i będą liczbami niewymiernymi, lub równą $(b - a)$, jeśli wszystkie punkty ξ_i będą liczbami wymiernymi. Jest więc jasne, że dla każdego ciągu normalnego podziałów (P_k) granica $\lim\limits_{k\to\infty} R_k$, gdzie $R_k \in \mathfrak{R}(f, P_k)$, $k \in \mathbb{N}$, nie istnieje.

Definicja 71. Niech f będzie ograniczoną funkcją rzeczywistą określona na $[a, b]$. Każdemu podziałowi P przedziału $[a, b]$ odpowiadają liczby:

$$M_i = \sup_{x_{i-1} \leqslant x \leqslant x_i} f(x) \quad m_i = \inf_{x_{i-1} \leqslant x \leqslant x_i} f(x)$$

$$U(f, P) = \sum_{i=1}^{n} M_i \Delta x_i \quad L(f, P) = \sum_{i=1}^{n} m_i \Delta x_i$$

Liczby $U(f, P)$ i $L(f, P)$ nazywamy odpowiednio sumą górną i dolną lub sumami Darboux funkcji f przy podziale P przedziału $[a, b]$. Dalej,

$$\overline{\int_a^b f(x)dx} = \inf_P U(f, P), \tag{36}$$

$$\underline{\int_a^b f(x)dx} = \sup_P L(f, P), \tag{37}$$

gdzie kres górny i dolny są brane ze względu na wszystkie podziały P przedziału $[a, b]$. Lewe strony równości (36) i (37) nazywają się odpowiednio górną i dolną całką Darboux funkcji f na przedziale $[a, b]$.

Ponieważ funkcja f jest ograniczona, więc istnieją liczby rzeczywistę m i M takie, że

$$m \leqslant f(x) \leqslant M \quad \text{dla} \quad x \in [a,b]$$

Oznacza to, że przy dowolnym podziale P przedziału $[a,b]$ mamy

$$m(b-a) \leqslant L(f,P) \leqslant U(f,P) \leqslant M(b-a)$$

a zatem zbiory $\{L(f,P) : P\}$ i $\{U(f,P) : P\}$ są ograniczone. Wynika stąd, że całki górna i dolna są określone przy dowolnej funkcji ograniczonej f.

Definicja 72. Mówimy, że podział P^* przedziału $[a,b]$ jest rozdrobnieniem (lub zagęszczeniem) podziału P tego przedziału, jeśli $P \subset P^*$, to znaczy, jeśli każdy punkt przedziału P jest także punktem przedziału P^*. Jeśli dane są dwa podziały P_1,P_2, to podział $P^* = P_1 \cup P_2$ nazywać będziemy ich wspólnym rozdrobnieniem (lub wspólnym zagęszczeniem).

Twierdzenie 74. *Jeśli P^* jest rozdrobnieniem podziału P, to*

$$L(f,P) \leqslant L(f,P^*) \quad U(f,P) \leqslant U(f,P^*)$$

Dowód. Załóżmy wpierw, że P^* zawiera tylko o jeden punkt więcej niż P. Niech tym dodatkowym punktem będzie x^* i niech $x_{i-1} < x^* < x_i$, gdzie x_{i-1}, x_i są dwoma kolejnymi punktami przedziału P. Przyjmijmy

$$\omega_1 = \inf_{x_{i-1} \leqslant x \leqslant x^*} f(x), \quad \omega_2 = \inf_{x^* \leqslant x \leqslant x_i} f(x)$$

Wtedy $\omega_1 \geqslant m_i$ i $\omega_2 \geqslant m_i$, gdzie $m_i = \inf\limits_{x_{i-1} \leqslant x \leqslant x_i} f(x)$. Mamy więc

$$\begin{aligned} L(f,P^*) - L(f,P) &= \omega_1(x^* - x_{i-1}) + \omega_2(x_i - x^*) - m_i(x_i - x_{i-1}) \\ &= (\omega_1 - m_i)(x^* - x_{i-1}) + (\omega_2 - m_i)(x_i - x^*) \geqslant 0 \end{aligned}$$

Jeśli P^* zawiera o k punktów więcej niż P, to powtarzając powyższe rozumowanie k razy otrzymamy pierwszą nierówność tezy. Dowód drugiej przebiega analogicznie.

Twierdzenie 75. *Jeśli f jest funkcją ograniczoną na przedziale $[a,b]$, to*

$$\underline{\int_a^b f(x)dx} \leqslant \overline{\int_a^b f(x)dx}$$

Dowód. Niech P^* będzie wspólnym rozdrobnieniem podziałów P_1 i P_2 przedziału $[a,b]$. Z Twierdzenia 75 wynika, że

$$L(f,P_1) \leqslant L(f,P^*) \leqslant U(f,P^*) \leqslant U(f,P_2)$$

Stąd $L(f,P_1) \leqslant U(f,P_2)$. Traktując P_2 jako ustalone i obliczając kres górny ze względu na wszystkie podziały P_1, wobec poprzedniej nierówności otrzymujemy

$$\underline{\int_a^b f(x)dx} \leqslant U(f,P_2)$$

Przechodząc do kresu dolnego ze względu na wszystkie podziały P_2 otrzymujemy tezę dowodzonego twierdzenia.

Udowodnimy teraz dwa kryteria całkowalności funkcji w sensie Riemanna. W oparciu o drugie z tych kryteriów podamy równoważną definicję całki w sensie Riemanna.

Twierdzenie 76. *Na to, aby ograniczona funkcja f była całkowalna w sensie Riemanna na przedziale $[a, b]$ potrzeba i wystarcza, aby dla dowolnego $\varepsilon > 0$ istniał taki podział P przedziału $[a, b]$, że*

$$U(f, P) - L(f, P) \leqslant \varepsilon \tag{38}$$

Dowód. Załóżmy wpierw, że funkcja f jest całkowalna w sensie Riemanna na przedziale $[a, b]$. Wówczas dla każdego danego $\varepsilon > 0$ istnieje taki podział P przedziału $[a, b]$, że nierówność

$$|R - \int_a^b f(x)dx| < \frac{\varepsilon}{2}, \text{ czyli}$$
$$\int_a^b f(x)dx - \frac{\varepsilon}{2} < R < \int_a^b f(x)dx + \frac{\varepsilon}{2}$$

jest spełniona przy dowolnym wyborze punktów ξ_i w każdym z przedziałów podziału. Ponieważ sumy Darboux są — przy danym podziale przedziału — odpowiednio kresem górnym i dolnym sum całkowych, zatem spełniają one nierówności

$$\int_a^b f(x)dx - \frac{\varepsilon}{2} \leqslant L(f, P) \leqslant U(f, P) \leqslant \int_a^b f(x)dx + \frac{\varepsilon}{2}$$

a więc $U(f, P) - L(f, P) < \varepsilon$. Załóżmy teraz, że (38) zachodzi. Dla dowolnego podziału P mamy

$$L(f, P) \leqslant \underline{\int_a^b f(x)dx} \leqslant \overline{\int_a^b f(x)dx} \leqslant U(f, P)$$

Jeśli $U(f, P) - L(f, P) < \varepsilon$, to wówczas

$$0 \leqslant \overline{\int_a^b f(x)dx} - \underline{\int_a^b f(x)dx} < \varepsilon$$

Z dowolności $\varepsilon > 0$ wynika, że $\underline{\int_a^b f(x)dx} = \overline{\int_a^b f(x)dx}$. Oznaczając ponadto $\underline{\int_a^b f(x)dx} = \overline{\int_a^b f(x)dx} = I$ mamy $L(f, P) \leqslant I \leqslant U(f, P)$. Ustalmy $\varepsilon > 0$ i niech P bedzie danym podziałem przedziału $[a, b]$, dla którego (38) zachodzi. Jeśli przez R oznaczymy jedną z wartości sum Riemanna odpowiadającej podziałowi P, to

$$L(f, P) \leqslant R \leqslant U(f, P)$$

Ponieważ liczby R oraz I znajdują się w przedziale $[L(f,P),U(f,P)]$, zatem

$$|R-I| \leqslant \varepsilon$$

Wobec Twierdzenia 74 oraz Definicji 70 wnioskujemy, że $I = \int_a^b f(x)dx$

Jako wniosek z powyższego twierdzenia otrzymujemy następujące

Twierdzenie 77. *Na to by ograniczona funkcja f była całkowalna w sensie Riemanna na przedziale $[a,b]$ potrzeba i wystarcza, by*

$$\underline{\int_a^b} f dx = \overline{\int_a^b} f dx \tag{39}$$

Dowód. W dowodzie Twierdzenia 76 pokazaliśmy, że (38) implikuje (39). Załóżmy teraz, że (39) zachodzi. Dla danej liczby $\varepsilon > 0$ istnieją podziały P_1 i P_2 przedziału $[a,b]$ takie, że

$$\underline{\int_a^b} f dx - \frac{\varepsilon}{2} < L(f,P_1), \quad U(f,P_2) < \overline{\int_a^b} f dx + \frac{\varepsilon}{2}$$

Jeśli podział P jest wspólnym rozdrobniemiem podziałów P_1 i P_2, to na mocy Twierdzenia 74 otrzymujemy

$$U(f,P) \leqslant U(f,P_2) < \overline{\int_a^b} f dx + \frac{\varepsilon}{2} = \underline{\int_a^b} f dx + \frac{\varepsilon}{2} < L(f,P_1) + \varepsilon \leqslant L(f,P) + \varepsilon$$

Stąd $U(f,P) - L(f,P) \leqslant \varepsilon$, a zatem warunek (38) jest spełniony. Wobec Twierdzenia 76 dowód jest zakończony.

Definicja 73. Mówimy, że ograniczona funkcja f jest całkowalna w sensie Riemanna, jeśli

$$\overline{\int_a^b} f dx = \underline{\int_a^b} f dx$$

Wspólną wartość określoną powyższą równością nazywamy całką Riemanna funkcji f na przedziale $[a,b]$.

Zbadamy teraz całkowalność w sensie Riemanna pewnych klas funkcji.

Twierdzenie 78. *Funkcja ciągła na przedziale $[a,b]$ jest na tym przedziale całkowalna w sensie Riemanna.*

Dowód. Funkcja f jest jednostajnie ciągła na $[a,b]$ (por. Tw. 51), a zatem dla dowolnego $\varepsilon > 0$ istnieje $\delta > 0$ taka, że $|f(x) - f(t)| < \frac{\varepsilon}{b-a}$ dla wszystkich $x, t \in [a,b]$, dla których $|x-t| < \delta$. Niech P będzie podziałem przedziału $[a,b]$, dla którego $\delta(P) < \delta$. Wtedy mamy $M_i - m_i \leqslant \frac{\varepsilon}{b-a}$ dla $i = 1, \ldots, n$ i wobec tego

$$U(f,P) - L(f,P) = \sum_{i=1}^{n}(M_i - m_i)\Delta x_i \leqslant \frac{\varepsilon}{b-a}\sum_{i=1}^{n}\Delta x_i = \varepsilon$$

Na mocy Twierdzenia 76 funkcja f jest całkowalna w sensie Riemanna na $[a, b]$.

Udowodnimy teraz następujące uogólnienie powyższego twierdzenia.

Twierdzenie 79. *Jeśli f jest funkcją ograniczoną i mającą tylko skończoną liczbę punktów nieciągłości na przedziale $[a, b]$, to jest ona całkowalna w sensie Riemanna na tym przedziale.*

Dowód. Ponieważ funkcja f jest ograniczona, więc istnieją liczby rzeczywiste m, M takie, że $m \leqslant f(x) \leqslant M$ dla wszystkich $x \in [a, b]$. Załóżmy, że f ma k punktów nieciągłości na przedziale $[a, b]$. Weźmy dowolne $\varepsilon > 0$ i $\delta_1 < \frac{\varepsilon}{8(M-m)k}$ (oczywiście $M \neq m$). Rozważmy przedziały otwarte $(x_l - \delta_1, x_l + \delta_1)$, $l = 1, \ldots, k$, gdzie x_l są punktami nieciągłości funkcji f. Dopełnienie sumy tych przedziałów do przedziału $[a, b]$ składa się ze skończonej liczby przedziałów domkniętych, na których funkcja f jest ciągła, a więc i jednostajnie ciągła. Ponieważ tych przedziałów jest skończenie wiele, więc dla danego $\varepsilon > 0$ istnieje liczba $\delta_2 > 0$ taka, że dla dowolnych punktów x, t należacych do jednego z tych przedziałów, na których funkcja f jest ciągła i spełniająca nierówność $|x - t| < \delta_2$ mamy $|f(x) - f(t)| < \frac{\varepsilon}{2(b-a)}$. Weźmy teraz liczbe $\delta = \min(\delta_1, \delta_2)$. Niech $P = \{x_0, \ldots, x_n\}$ będzię dowolnym podziałem przedziału $[a, b]$, dla którego $\delta(P) < \delta$. Ponadto rozbijmy zbiór indeksów $\{1, \ldots, n\}$ na dwa rozłącznę zbiory A i B w następujący sposób: do zbioru A zaliczymy te liczby i, dla których przedział $[x_{i-1}, x_i]$ nie ma punktów wspólnych z żadnym z skontruowanych powyżej otoczeń punktów x_l, $l = 1, \ldots, k$, a do zbioru B pozostałe przedziały powstające z podziału P przedziału $[a, b]$. Wówczas

$$U(f,P) - L(f,P) = \sum_{i=1}^{n}(M_i - m_i)\Delta x_i = \sum_{i\in A}(M_i - m_i)\Delta x_i + \sum_{i\in B}(M_i - m_i)\Delta x_i$$

Ponadto

$$\sum_{i\in A}(M_i - m_i)\Delta x_i \leqslant \frac{\varepsilon}{2(b-a)}\sum_{i\in A}\Delta x_i \leqslant \frac{\varepsilon}{2(b-a)}(b-a) = \frac{\varepsilon}{2}$$

Suma długości podprzedziałów przedziału $[a, b]$ indeksowanych przez liczby ze zbioru B jest nie większa niż

$$(\delta + 2\delta_i + \delta)k < 4\frac{\varepsilon}{8(M-m)k}k = \frac{\varepsilon}{2(M-m)}$$

Dlatego

$$\sum_{i\in B}(M_i - m_i)\Delta x_i \leqslant (M-m)\sum_{i\in B}\Delta x_i < (M-m)\frac{\varepsilon}{2(M-m)} = \frac{\varepsilon}{2}$$

Dla podziału P o średnicy mniejszej niż δ otrzymujemy zatem

$$U(f,P) - L(f,P) = \sum_{i=1}^{n} (M_i - m_i)\Delta x_i < \varepsilon$$

co kończy dowód.

Uwaga 38. Twierdzenie 78 można istotnie uogólnić. Mianowicie dowodzi się, że jeśli f jest ograniczoną funkcją na przedziale $[a,b]$, to jest ona całkowalna w sensie Riemanna na tym przedziale wtedy i tylko wtedy, gdy jest ona ciągła prawie wszędzie na $[a,b]$, to znaczy zbiór punktów nieciągłości funkcji f ma miarę Lebesgue'a równą zeru. (por. [7], s. 270). Przykładów takich funkcji dostarcza następujące

Twierdzenie 80. *Funkcja monotoniczna na przedziale $[a,b]$ jest na tym przedziale całkowalna w sensie Riemanna.*

Dowód. Załóżmy, że f jest funkcją niemalejącą. Niech będzie dane dowolne $\varepsilon > 0$. Weźmy podział P przedziału $[a,b]$ na n równych części o długości $\frac{b-a}{n}$. Ponieważ f jest niemalejącą zatem $M_i = f(x_i)$ oraz $m_i = f(x_{i-1})$ dla $i = 1,\ldots,n$. Mamy więc

$$U(f,P) - L(f,P) = \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))\frac{b-a}{n} = (f(b) - f(a))\frac{b-a}{n}$$

Biorąc n tak duże, aby $(f(b) - f(a))\frac{b-a}{n} < \varepsilon$ i stosując twierdzenie 76 otrzymujemy tezę. W przypadku funkcji nierosnącej dowód jest analogiczny.

Twierdzenie 81. *Jeśli f jest całkowalna w sensie Riemanna na przedziale $[a,b]$, $m \leqslant f(x) \leqslant M$ dla $x \in [a,b]$ oraz ϕ jest funkcją ciągłą na $[m,M]$, to funkcja złożona $h = \phi \circ f$ jest R-całkowalna na $[a,b]$.*

Dowód. Ustalmy $\varepsilon > 0$. Ponieważ funkcja ϕ jest jednostajnie ciągła na $[M,m]$, więc istnieje $\delta > 0$ taka, że $\delta < \varepsilon$ i $|\phi(s) - \phi(t)| < \varepsilon$, jeśli $|s-t| < \delta$. Ponieważ f jest R-całkowalna na $[a,b]$, więc istnieje podział $P = \{x_0,\ldots,x_n\}$ przedziału $[a,b]$ taki, że $U(f,P) - L(f,P) < \delta^2$. Niech

$$M_i = \sup_{x_{i-1} \leqslant x \leqslant x_i} f(x), \quad m_i = \inf_{x_{i-1} \leqslant x \leqslant x_i} f(x),$$

$$M_i^* = \sup_{x_{i-1} \leqslant x \leqslant x_i} h(x), \quad m_i^* = \inf_{x_{i-1} \leqslant x \leqslant x_i} h(x)$$

dla $i = 1,\ldots,n$. Podzielmy zbiór $\{1,\ldots,n\}$ na dwa rozłącznę zbiory A i B w taki sposób, że $i \in A$, jeśli $M_i - m_i < \delta$ oraz $i \in B$ w przypadku przeciwnym. Wówczas wobec powyższego wyboru δ mamy $M_i^* - m_i^* < \varepsilon$ dla $i \in A$. Natomiast dla $i \in B$ mamy $M_i^* - m_i^* \leqslant 2K$, gdzie $K = \sup\{|\phi(t)| : m \leqslant t \leqslant M\}$. Stąd otrzymujemy

$$\delta \sum_{i \in B} (x_i - x_{i-1}) \leqslant \sum_{i \in B} (M_i - m_i)(x_i - x_{i-1}) < \delta^2, \quad \text{zatem} \sum_{i \in B} (x_i - x_{i-1}) < \delta$$

Mamy więc

$$U(h,P) - L(h,P) = \sum_{i\in A}(M_i^* - m_i^*)(x_i - x_{i-1}) + \sum_{i\in B}(M_i^* - m_i^*)(x_i - x_{i-1})$$

a zatem

$$U(h,P) - L(h,P) \leqslant \varepsilon(a+b+2K)$$

Ponieważ ε było dowolne, zatem na mocy twierdzenia 76 funkcja h jest R-całkowalna.

Następujące twierdzenie opisuję związek całki Riemanna z operacjami arytmetycznymi.

Twierdzenie 82. *Jeśli funkcje f i g są R-całkowalne na przedziale $[a,b]$, to również R-całkowalne są funkcje $f+g$, λf (λ jest dowolną stałą rzeczywistą) i fg oraz prawdziwe są równości:*

$$\int_a^b (f+g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx, \tag{40}$$

$$\int_a^b (\lambda f)(x)dx = \lambda \int_a^b f(x)dx \tag{41}$$

Dowód. Jest jasne, że dla dowolnego $R \in \mathfrak{R}(f+g,P)$ mamy $R = R_f + R_g$, gdzie $R_f \in \mathfrak{R}(f,P)$, $R_g \in \mathfrak{R}(g,P)$. Niech $I_1 = \int_a^b f(x)dx$, $I_2 = \int_a^b f(x)dx$ oraz $I = I_1 + I_2$. Mamy

$$\forall_{\varepsilon>0}\exists_{\delta>0}\forall_P\forall_{R_f\in\mathfrak{R}(f,P)}\delta(P) < \delta \implies |R_f - I_1| < \frac{\varepsilon}{2} \quad \text{oraz}$$

$$\forall_{\varepsilon>0}\exists_{\delta>0}\forall_P\forall_{R_g\in\mathfrak{R}(g,P)}\delta(P) < \delta \implies |R_g - I_2| < \frac{\varepsilon}{2} \quad \text{Stąd}$$

$$\forall_{\varepsilon>0}\exists_{\delta>0}\forall_P\forall_{R\in\mathfrak{R}(f+g,P)}\delta(P) < \delta \implies |R - I| \leqslant |R_f - I_1| + |R_g - I_2| < \varepsilon$$

Wobec powyższego jest jasne, że funkcja $f+g$ jest R-całkowalna na przedziale $[a,b]$ oraz, że spełniony jest wzór 40. Dowód wzoru 41 jest analogiczny.

Dalej przyjmując $\phi(t) = t^2$ oraz stosując do ϕ poprzednie twierdzenie (81) otrzymujemy R-całkowalność funkcji f^2.

R-całkowalność iloczynu funkcji fg wynika z tożsamości

$$fg = \frac{1}{4}[(f+g)^2 - (f-g)^2].$$

Twierdzenie 83. *(a) Jeśli funkcje f i g są R-całkowalne na przedziale $[a,b]$ oraz $f(x) \leqslant g(x)$ dla każdego $x \in [a,b]$, to*

$$\int_a^b f(x)dx \leqslant \int_a^b g(x)dx$$

(b) Jeśli funkcja f jest R-całkowalna na przedziale $[a, b]$, to funkcja $|f|$ jest również R-całkowalna na tym przedziale oraz:

$$\left| \int_a^b f(x)dx \right| \leqslant \int_a^b |f(x)|dx$$

Dowód. (a) Jeśli $m \leqslant f(x) \leqslant M$ dla $x \in [a, b]$, to

$$m(b-a) \leqslant \int_a^b f(x)dx \leqslant M(b-a)$$

Stąd, jeśli $f(x) \geqslant 0$ dla $x \in [a, b]$, to $\int_a^b f(x)dx \geqslant 0$. Wobec tego nierówność $f(x) \leqslant g(x)$ dla $x \in [a, b]$ implikuje

$$\int_a^b f(x)dx \leqslant \int_a^b g(x)dx$$

(b) Biorąc $\phi(t) = |t|$ w Twierdzeniu 81 otrzymujemy całkowalność funkcji $|f|$. Ponieważ $-|f(x)| \leqslant f(x) \leqslant |f(x)|$ dla $x \in [a, b]$, zatem na mocy (a) otrzymujemy

$$\left| \int_a^b f(x)dx \right| \leqslant \int_a^b |f(x)|dx$$

Uwaga 39. (a) Punkt (a) Twierdzenia 83 można udowodnić bezpośrednio w oparciu o definicję całki Riemanna (Def. 69) oraz Wniosek 3 (b).

(b) Twierdzenie odwrotne do Twierdzenia 83 (b) nie jest prawdziwe, to znaczy z R-całkowalności $|f|$ nie wynika R-całkowalność funkcji f. Dla przykładu niech

$$f(x) = \begin{cases} 1 & \text{dla } x \in \mathbb{Q} \cap [a, b], \\ -1 & \text{dla } x \in (\mathbb{R} \setminus \mathbb{Q}) \cap [a, b]. \end{cases}$$

Oczywiście funkcja $|f|$ jest R-całkowalna na przedziale $\underline{[a, b]}$ oraz $\int_a^b |f(x)|dx = b - a$. Z kolei $underline \int_a^b f(x)dx = -(b-a)$ oraz $\overline{\int_a^b f(x)dx} = b - a$, a zatem wobec Twierdzenia 77 funkcja f nie jest R-całkowalna na przedziale $[a, b]$.

Twierdzenie 84. *Jeśli dwie funkcje f i g są równe na przedziale $[a, b]$ z wyjątkiem skończonego zbioru punktów $\{x_1, \ldots, x_k\}$ i jedna z nich, na przykład g jest R-całkowalna na tym przedziale, to druga też jest na nim R-całkowalna i zachodzi równość*

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

Dowód. Ponieważ $f = g + (f - g)$, więc wystarczy udowodnić, że funkcja $\phi = f - g$ jest R-całkowalna na $[a, b]$ i $\int_a^b \phi(x)dx = 0$. Oznaczmy $N = \max\{|\phi(x_1), \ldots, \phi(x_k)\}$. Niech P będzie podziałem przedziału $[a, b]$ o średnicy

δ. Funkcja ϕ na co najwyżej $2k$ przedziałach podziału P nie jest tożsamościowo równa zeru. Dlatego mamy $U(\phi, P) \leqslant 2Nk\delta$ i $L(\phi, P) = 0$, zatem $U(\phi, P) - L(\phi, P) \leqslant 2Nk\delta$. Biorąc ϕ odpowiednio małe możemy uczynić różnicę $U(\phi, P) - L(\phi, P)$ dowolnie małą. To oznacza, że funkcja ϕ jest R-całkowalna. Ponadto jasne jest, że $\int_a^b \phi(x)dx = 0$.

Wniosek 18. *Niech funkcja f będzie określona i ograniczona na przedziale otwartym (a, b). Jeśli po nadaniu jej pewnych wartości $f(a)$ i $f(b)$ stanie się ona R-całkowalna na przedziale domkniętym $[a, b]$ — to taką pozostanie — gdy liczby $f(a)$ i $f(b)$ zmienimy w sposób dowolny. Wartość całki nie ulegnie przy tym zmianie.*

Następujący lemat pozwala przy przybliżaniu całki Riemanna sumami całkowymi ograniczyć się tylko do podziałów zawierających z góry ustalony punkt.

Lemat 2. *Niech $c \in [a, b]$ i niech Π^* oznacza zbiór wszystkich podziałów przedziału $[a, b]$ spełniających warunek:*

$$P = \{x_0, \dots, x_n\} \in \Pi^* \text{ wtedy i tylko wtedy } x_j = c \text{ dla pewnego } j.$$

Wówczas dla dowolnej funkcji f, ograniczonej na $[a, b]$ zachodzą równości:

$$\sup_{P \in \Pi^*} L(f, P) = \sup_P L(f, P), \quad \inf_{P \in \Pi^*} U(f, P) = \inf_P U(f, P)$$

Dowód. Ponieważ Π^* jest podzbiorem zbioru wszystkich podziałów przedziału $[a, b]$, więc

$$\sup_{P \in \Pi^*} L(f, P) \leqslant \sup_P L(f, P), \quad \inf_{P \in \Pi^*} U(f, P) \geqslant \inf_P U(f, P) \tag{42}$$

Zauważmy, że dla dowolnego podziału P przedziału $[a, b]$ istnieje podział od niego drobniejszy $P^* \in \Pi^*$. Istotnie, jeśli $P \in \Pi^*$, to przyjmujemy $P^* = P$. Jeśli natomiast $P \notin \Pi^*$, to przez dołączenie punktu c do układu punktów wyznaczających P otrzymujemy podział P^* o żądanych własnościach. Mamy więc

$$L(f, P) \leqslant L(f, P^*), \quad U(f, P) \geqslant U(f, P^*),$$

skąd otrzymujemy

$$L(f, P) \leqslant \sup_{P^* \in \Pi^*} L(f, P^*), \quad U(f, P) \geqslant \inf_{P^* \in \Pi^*} U(f, P^*).$$

Wobec dowolności podziału P mamy

$$\sup_P L(f, P) \leqslant \sup_{P^* \in \Pi^*} L(f, P), \quad \inf_P U(f, P) \geqslant \inf_{P^* \in \Pi^*} U(f, P)$$

Z powyższych nierówności i z (42) otrzymujemy tezę.

Twierdzenie 85. *Niech $a < c < b$. Funkcja f jest R-całkowalna na przedziale $[a, b]$ wtedy i tylko wtedy, gdy jest ona R-całkowalna na przedziałach $[a, c]$ i $[c, b]$. Zachodzi przy tym równość*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

(addytywność całki ze względu na przedział)

Dowód. Załóżmy, że funkcja f jest R-całkowalna na przedziale $[a, b]$. Na mocy powyższego lematu możemy ograniczyć się do podziałów przedziału $[a, b]$ zawierających punkt c. Jeśli P jest takim podziałem, to wówczas $P = P_1 \cup P_2$, gdzie P_1 jest podziałem przedziału $[a, c]$, a P_2 — podziałem przedziału $[c, b]$ oraz mamy

$$U(f, P) = U(f, P_1) + U(f, P_2), \quad L(f, P) = L(f, P_1) + L(f, P_2).$$

Niech będzie dane dowolne $\varepsilon > 0$ i niech

$$U(f, P) - L(f, P) < \frac{\varepsilon}{2}.$$

Stąd $U(f, P_1) - L(f, P_1) < \frac{\varepsilon}{2}$ i $U(f, P_2) - L(f, P_2) < \frac{\varepsilon}{2}$. Funkcja f jest więc całkowalna na przedziałach $[a, c]$ i $[c, b]$ oraz zachodzą nierówności

$$U(f, P_1) < \int_a^c f(x)dx + \frac{\varepsilon}{2}, \quad \int_a^c f(x)dx < L(f, P_1) + \frac{\varepsilon}{2},$$

$$U(f, P_2) < \int_c^b f(x)dx + \frac{\varepsilon}{2}, \quad \int_c^b f(x)dx < L(f, P_2) + \frac{\varepsilon}{2},$$

Wobec powyższego otrzymujemy $U(L, P) < \int_a^c f(x)dx + \int_c^b f(x)dx + \varepsilon$ i w konsekwencji $\int_a^b f(x)dx < \int_a^c f(x)dx + \int_c^b f(x)dx + \varepsilon$. Ponieważ $\varepsilon > 0$ było dowolne, zatem

$$\int_a^b f(x)dx \leqslant \int_a^c f(x)dx + \int_c^b f(x)dx. \tag{43}$$

Analogicznie $\int_a^c f(x) + \int_c^b f(x)dx < L(f, P) + \varepsilon$, skąd

$$\int_a^c f(x)dx + \int_c^b f(x)dx \leqslant \int_a^b f(x)dx, \tag{44}$$

bowiem $\varepsilon > 0$ jest dowolne. Z nierówności (43) i (44) otrzymujemy żądaną równość. Uzasadnienie implikacji odwrotnej jest analogiczne.

Rozszerzymy teraz zasięg Definicji 69.

Definicja 74. W przypadku gdy $b < a$ lub $b = a$, to całkę Riemanna z funkcji f określamy wzorami

$$\int_a^b f(x)dx = -\int_b^a f(x)dx \quad \text{lub odpowiednio} \quad \int_a^b f(x)dx = 0.$$

W całce $\int_a^b f(x)dx$ liczbę a nazywamy dolną granicą całkowania, liczbę b – górną granicą całkowania, bez względu na to, czy $b \geqslant a$, czy też $b < a$.

Wniosek 19. *(a) Niech $a, b, c \in \mathbb{R}$ i niech f będzie funkcją R-całkowalna na najwiekszym z przedziałów domkniętych o końcach we wskazanych punktach. Wówczas obcięcie funkcji f do każdego z dwóch pozostałych przedziałów domkniętych jest funkcja R-całkowalną na odpowiednim przedziale oraz zachodzi równość*

$$\int_a^b f(x)dx + \int_b^c f(x)dx + \int_c^a f(x)dx = 0. \tag{45}$$

Dowód. Wobec symetrii równość (45) względem a, b, c możemy bez straty ogólności założyć, że $a = \min\{a, b, c\}$. Jeśli $\max\{a, b, c\} = c$ oraz $a < b < c$, to na mocy Twierdzenia 85 mamy

$$\int_a^b f(x)dx + \int_c^b f(x)dx - \int_a^c f(x)dx = 0,$$

zatem wobec Definicji 74 otrzymujemy równość (45).
Jeśli $\max\{a, b, c\} = b$ oraz $a < c < b$, to ponownie na mocy Twierdzenia 85 mamy

$$\int_a^c f(x)dx + \int_c^b f(x)dx - \int_a^b f(x)dx = 0,$$

stąd wobec Definicji 74 wynika równość (45).
W końcu, jeśli jakiekolwiek dwa z punktów a, b, c lub wszystkie trzy pokrywają się, to (45) jest bezpośrednią konsekwencją Definicji 74.

(b) Jeśli funkcja f jest R-całkowalna na przedziale $[a, b]$ i $a \leqslant c < d \leqslant d$, to jest ona również R-całkowalna na przedziale $[c, d]$.

Definicja 75. Niech każdej uporządkowanej parze (α, β) punktów α, β przedziału $[a, b]$ odpowiada dokładnie jedna liczba $I(\alpha, \beta)$, przy czym dla dowolnej trójki punktów $\alpha, \beta, \gamma \in [a, b]$ zachodzi równość

$$I(\alpha, \gamma) = I(\alpha, \beta) + I(\beta, \gamma).$$

Wówczas funkcja $I(\alpha, \beta)$ nazywa się addytywną funkcją przedziału zorientowanego (dla $\alpha = \gamma$ wobec powyższej równości otrzymujemy $I(\alpha, \beta) = -I(\beta, \alpha)$), określoną na odcinkach zawartych w przedziale $[a, b]$.

Wniosek 20. *Jeśli funckja f jest R-całkowalna na przedziale $[a,b]$ oraz $\alpha,\beta,\gamma\in[a,b]$, to kładąc $I(\alpha,\beta)=\int_\alpha^\beta f(x)dx$, na mocy równości 45 otrzymujemy*

$$\int_\alpha^\gamma f(x)dx=\int_\alpha^\beta f(x)dx+\int_\beta^\gamma f(x)dx,\quad \textit{czyli}\quad I(\alpha,\gamma)=I(\alpha,\beta)+I(\beta,\gamma),$$

czyli całka Riemanna jest adytywną funkcją przedziału zorientowanego.

Udowodnimy teraz ważne twierdzenie o funkcji górnej granicy całkowania.

Twierdzenie 86. *Niech f będzie funkcja R-całkowalną na przedziale $[a,b]$. Dla dowolnego punktu $x\in[a,b]$ określamy*

$$F(x)=\int_a^x f(t)dt.$$

Wówczas funkcja F jest ciągła na przedziale $[a,b]$. Ponadto, jeśli funkcja f jest ciągła w punkcie $x_0\in[a,b]$, to funkcja F jest różniczkowalna w tym punkcie oraz $F'(x_0)=f(x_0)$.

Dowód. Niech M będzie takie, że $|f(t)|\leqslant M$ dla $t\in[a,b]$. Wówczas, jeśli $a\leqslant x\leqslant y\leqslant b$, to

$$|F(y)-F(x)|=\left|\int_x^y f(t)dt\right|\leqslant M(y-x).$$

Stąd wynika natychmiast, że dla dowolnego $\varepsilon>0$ mamy $|F(y)-F(x)|<\varepsilon$, jeśli tylko $|y-x|<\frac{\varepsilon}{M}$, a zatem funkcja F jest ciągła.
Załóżmy, że f jest ciągła w punkcie x_0. Dla danego $\varepsilon>0$ wybierzmy $\delta>0$ tak, aby $|f(t)-f(x_0)|<\varepsilon$ jeśli tylko $|t-x_0|<\delta$ i $a\leqslant t\leqslant b$. Wówczas dla $s,t\in(x_0-\delta,x_0+\delta)$, $s,t\in[a,b]$, $s\neq t$ mamy

$$\left|\frac{F(t)-F(s)}{t-s}-f(x_0)\right|=\left|\frac{1}{t-s}\int_s^t(f(u)-f(x_0))du\right|<\varepsilon,$$

Stąd wynika, że $F'(x_0)=f(x_0)$, co kończy dowód.

Oznaczmy $F(x)=I(a,x)$ dla $x\in[a,b]$, gdzie I oznacza addytywną funkcję przedziału zorientowanego. Mamy

$$I(\alpha,\beta)=I(a,\beta)-I(a,\alpha)=F(\beta)-F(\alpha)$$

dla każdej uporządkowanej pary punktów (α,β) z przedzialu $[a,b]$. W ten sposób każda addytywna funkcja przedziału zorientowanego ma postać

$$I(\alpha,\beta)=F(\beta)-F(\alpha), \tag{46}$$

gdzie $x\mapsto F(x)$ jest funkcją określoną na przedziale $[a,b]$. Można łatwo sprawdzić, że jest również na odwrót, to znaczy, że z dowolnej funkcji $x\mapsto F(x)$

określonej na przedziale $[a, b]$ można przy pomocy (46) otrzymać addytywną funkcję przedziału zorientowanego.

Wniosek 21. *Jeśli* f jest funkcją R-całkowalną na przedziale $[a, b]$, to na mocy (46) funkcja $F(x) = \int_a^x f(t)dt$ generuje addytywną funkcję

$$I(\alpha, \beta) = \int_\alpha^\beta f(t)dt.$$

8.2 Całka nieoznaczona

Definicja 76. Niech f będzie funkcją określoną na pewnym przedziale I. Każdą funkcję F różniczkowalną na tym przedziale i spełniającą w każdym punkcie $x \in I$ równość

$$F'(x) = f(x)$$

nazywamy funkcją pierwotną funkcji f. Funkcję pierwotną nazywamy również całką nieoznaczoną danej funkcji i oznaczamy symbolem $\int f(x)dx$ (symbol ten należy również rozumieć jako oznaczenie dowolnej funkcji pierwotnej funkcji f na tym przedziale). W symbolu tym znak $\int$ nazywa się znakiem całki nieoznaczonej, f — funkcją podcałkową, a $f(x)dx$ — wyrażeniem podcałkowym.

Uwaga 40. Jeśli F jest funkcją pierwotną funkcji f, to suma $F + c$, gdzie c jest dowolną stałą, jest również funkcją pierwotną funkcji f, bowiem $(F + c)' = F' = f$.
Na odwrót, dwie dowolne fukcje pierwotne F i G tej samej funkcji f różnią się o stałą, bowiem $(F - G)' = f - f = 0$.
Jeśli F jest więc konkretną funkcją pierwotną funkcji f na przedziale I, to na tym przedziale

$$\int f(x)dx = F(x) + c,$$

to znaczy dowolna inna funkcja pierwotna funkcji f może być otrzymana z danej funkcji F przez dodanie stałej.

Bezpośrednio z Twierdzenia 86 otrzymujemy następujący

Wniosek 22. *Każda funkcja ciągła* f *na przedziale* $[a, b]$ *ma na nim funkcję pierwotną.*

Dowód wniosku 22 można uzyskać bez pojęcia całki Riemanna. Jest on jednak dość długi.
Istnieją również funkcję nieciągłe, które posiadają funkcje pierwotne.

Przykład 26. Niech

$$f(x) = \begin{cases} 2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x} & \text{dla } x \neq 0 \\ 0 & \text{dla } x = 0. \end{cases}$$

Ponieważ nie istnieje granica funkcji f w zerze, zatem f nie jest funkcją ciągłą. Można łatwo sprawdzić, że funkcja

$$F(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & \text{dla } x \neq 0 \\ 0 & \text{dla } x = 0, \end{cases}$$

jest funkcją pierwotną funkcji f (por. [7], s. 90-91).

Podamy teraz przykłady funkcji całkowalnej w sensie Riemanna, która nie posiada funkcji pierwotnej.

Przykład 27. Niech $f : (1, 3) \mapsto \mathbb{R}$ będzie określona wzorem $f(x) = [x]$. Jest jasne, że dla $f(x) = F'(x)$ dla $x \in (1, 2) \cup (2, 3)$, gdzie

$$F(x) = \begin{cases} x + c_1 & \text{dla } x \in (1, 2) \\ 2x + c_2 & \text{dla } x \in (2, 3), \end{cases}$$

gdzie, $c_1, c_2 \in \mathbb{R}$. Funkcja pierwotna funkcji f na przedziale $(1, 3)$ (z dokładnością do stałej musiałaby mieć postać)

$$F(x) = \begin{cases} x + c_1 & \text{dla } x \in (1, 2] \\ 2x + c_1 - 2 & \text{dla } x \in (2, 3), \end{cases}$$

gdzie $c_1 \in \mathbb{R}$. Istotnie, aby funkcja F była ciągła dla $x = 2$, to $\lim\limits_{x \to 2^-} F(x) = \lim\limits_{x \to 2^+} F(x)$, czyli $2 + c_1 = 4 + c_2$, zatem $c_2 = c_1 - 2$. Można łatwo sprawdzić, że $F'_-(2) = 1$ oraz $F'_+(2) = 2$, czyli F nie jest funkcją pierwotną funkcji f na przedziale $(1, 3)$.

Tablica 1. Całki nieoznaczona podstawowych funkcji elementarnych

$f(x)$	$F(x)$	Ograniczenia ze względu na argument $x \in \mathbb{R}$
0	$c = \text{const.}$	
$a = \text{const.}$	$ax + c$	
x^p	$\frac{1}{p+1}x^{p+1} + c$	$p \neq -1, x > 0$ $(p \in \mathbb{R})$, $x \neq 0$ $(p \in \mathbb{Z})$, $x \in \mathbb{R}$ $(p \in \mathbb{N})$
$\frac{1}{x}$	$\ln\lvert x\rvert + c$	$x \neq 0$
a^x	$\frac{a^x}{\ln a} + c$	$x \in \mathbb{R}$ $(a > 0, a \neq 1)$
e^x	$e^x + c$	
$\sin x$	$-\cos x + c$	
$\cos x$	$\sin x + c$	
$\frac{1}{\cos^2 x}$	$\operatorname{tg} x + c$	$x \neq \frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}$
$\frac{1}{\sin^2 x}$	$-\operatorname{ctg} x + c$	$x \neq k\pi,\ k \in \mathbb{Z}$
$\frac{1}{1+x^2}$	$\operatorname{arctg} x + c$ $\operatorname{arcctg} x + \hat{c}$	
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x + c$ $-\arccos x + \hat{c}$	$\lvert x\rvert < 1$

Uwaga 41. (a) Wzory zawarte w Tablicy 2 otrzymujemy przez bezpośrednie różniczkowanie funkcji $F(x)$ (zob. Tablica 1).

(b) Jeżeli zakres argumentów, dla których spełniona jest równość $F(x) = f'(x)$ nie jest przedziałem (skończonym lub nieskończonym), to nie można twierdzić, że wyrażenie $F(x) + c$ obejmuje wszystkie funkcje pierwotne funkcji f w tym zakresie argumentów. Dla przykładu funkcja

$$G(x) = \begin{cases} \ln(-x) & \text{dla } x < 0 \\ \ln x & \text{dla } x > 0, \end{cases}$$

jest funkcją pierwotną funkcji $x \mapsto \frac{1}{x}$ $(x \neq 0)$, mimo, że nie podpada pod wzór $\ln|x| + c$.

Następujące twierdzenie podaje reguły obliczani całek nieoznaczonych.

Twierdzenie 87. *(a) Jeśli istnieją całki nieoznaczone funkcji $u, v : P \mapsto \mathbb{R}$, gdzie P jest przedziałem oraz $\alpha, \beta \in \mathbb{R}$, to istnieje całka nieoznnaczona funkcji $\alpha u + \beta v$ oraz zachodzi wzór*

$$\int (\alpha u(x) + \beta v(x))dx = \alpha \int u(x)dx + \beta \int v(x)dx + c \text{ dla } x \in P. \tag{47}$$

(b) Przy założeniach punktu (a) oraz przy założeniu, że funkcje u, v są różniczkowalne oraz jedna z całek występujących w poniższym wzorze istnieje, prawdziwy

jest następujący wzór zwany wzorem na całkowanie przez części:

$$\int u(x)v'(x)dx = u(x)v(x) - \int u'(x)v(x)dx + c \text{ dla } x \in P. \tag{48}$$

(c) Jeśli na przedziale I, $\int f(x)dx = F(x) + c$ oraz $\phi : P \mapsto I$ jest odwzorowaniem klasy C^1, to

$$\int f(\phi(t))\phi'(t)dt = F(\phi(t)) + c \text{ dla } t \in P. \tag{49}$$

Dowód. (a) Wzór (47) sprawdza się bezpośrednio przez różniczkowanie lewej i prawej strony z wykorzystaniem liniowości różniczkowania (zob. Tw. 54 (a), (b)).

(b) Załóżmy, że $\int u(x)v'(x)dx = \Phi(x)$ dla $x \in P$. Ponieważ $(u(x)v(x))' = u'(x)v(x) + u(x)v'(x)$, $(u(x)v(x) - \Phi(x))' = u'(x)v(x)$ dla $x \in P$, a więc $uv - \Phi$ jest funkcją pierwotną funkcji $u'v$ na przedziale P. Ponadto mamy

$$\int u(x)v'(x)dx = \int (u(x)v(x))'dx - \int u'(x)v(x)dx, \text{ a więc}$$

$$\int u(x)v'(x)dx = u(x)v(x) - \int u'(x)v(x)dx + c.$$

(c) Wzór (49) jest bezpośrednią konsekwencją reguły różniczkowania funkcji złożonej (zob. Tw. 55).

Uwaga 42. (a) Wzór (49) pokazuje, że chcąc uzyskać funkcję pierwotną funkcji $t \mapsto f(\phi(t))\phi'(t)$ można postąpić w następujacy sposób:

$$\int f(\phi(t))\phi'(t)dt = \int f(x)dx = F(x) + c = F(\phi(t)) + c,$$

to znaczy najpierw dokonać zmiany $\phi(t) = x$ i przejść do nowej zmiennej x, a następnie przejść do poprzedniej zmiennej podstawiając $x = \phi(t)$. Wzór ten nazywa się wzorem na całkowanie przez podstawienie.

(b) Jeśli $\phi : P \mapsto I$ jest bijekcją, to aby obliczyć całkę nieoznaczoną $\int f(x)dx$ można obliczyć całke nieoznaczoną $\int f(\phi(t))\phi'(t)dt$, a następnie dokonać podstawienia $t = \phi^{-1}(x)$.

(c) W szczególności wzór (48) jest prawdziwy jeśli $u, v \in C^1$.

Przykład 28. (a) Obliczmy $\int \arcsin x dx$. Najpierw zastosujmy wzór (48) przyjmując $u(x) = \arcsin x$ i $v'(x) = 1$. Wówczas $u'(x) = \frac{1}{\sqrt{1-x^2}}$ (dla $x \in (-1, 1)$) i $v(x) = x$ oraz $\int \arcsin x dx = x \arcsin x - \int \frac{x}{\sqrt{1-x^2}}dx$. Aby obliczyć tę ostatnią całkę zastosujmy podstawienie $\sqrt{1-x^2}=t$. Mamy wówczas $\frac{-xdx}{\sqrt{1-x^2}}=dt$ i stąd $\int \frac{x}{\sqrt{1-x^2}}dx = -\int dt = -t+c = -\sqrt{1-x^2}+c$, a zatem

$$\int \arcsin x dx = x \arcsin x + \sqrt{1-x^2} + c.$$

(b) Obliczmy $\int \operatorname{tg} t dt$. Niech np. $P=(-\frac{\pi}{2},\frac{\pi}{2})$, $I=(0,1]$, $\phi : P \mapsto I$ będzie określone wzorem $\phi(t)=cost$, natomiast $f(x)=\frac{1}{x}$ dla $x\in(0,1]$. Mamy

$$\int \operatorname{tg} t dt = \int \frac{1}{\cos t}\sin t dt = -\int \frac{1}{x}dx = -\ln x + c = \ln\cos t + c.$$

W ogólności podstawiamy $\cos t = x$. Wówczas $-\sin t dt = dx$ oraz

$$\int \operatorname{tg} t dt = -\int \frac{dx}{x} = -\ln|x| + c = -\ln|\cos t| + c.$$

Wobec Uwagi 41 (b) powyższy wzór nie objemuje wszystkich funkcji pierwotnych funkcji $t \mapsto \operatorname{tg} t$, gdzie $t \neq \frac{\pi}{2}+k\pi$, $k\in\mathbb{Z}$.

(c) Obliczmy $\sqrt{1-x^2}$. Niech $I=(-1,1)$, $P=(-\frac{\pi}{2},\frac{\pi}{2})$, $f(x)=\sqrt{1-x^2}$ dla $x\in I$ oraz $\phi(t)=sint$ dla $t\in P$. Oczywiście $\phi : P\mapsto I$ jest bijekcją oraz $\phi^{-1}(x)=\arcsin x$ dla $x\in I$. Mamy

$$\begin{aligned}\int \sqrt{1-x^2}dx &= \int \sqrt{1-\sin^2 t}\cos t dt = \int \cos^2 t dt \int \frac{dt}{2} + \frac{1}{2}\int \cos 2t dt \\ &= \frac{1}{2}t + \frac{1}{4}\sin 2t + c = \frac{1}{2}\arcsin x + \frac{1}{4}\sin(2\arcsin x) + c \\ &= \frac{1}{2}\arcsin x + \frac{1}{2}\sin(\arcsin x)\cos(\arcsin x) + c \\ &= \frac{1}{2}\arcsin x + \frac{1}{2}x\sqrt{1-\sin^2(\arcsin x)} + c \\ &= \frac{1}{2}\arcsin x + \frac{1}{2}x\sqrt{1-x^2} + c.\end{aligned}$$

Uwaga 43. Funkcjami elementarnymi nazywamy funkcje potęgowe, wykładnicze, trygonometryczne, funkcje odwrotne do nich, ich superpozycje oraz funkcje powstałe przez wykonanie skończonej ilości działań na nich: dodawanie, odejmowanie, mnożenie, dzielenie oraz składanie. Podstawowe metody całkowania funkcji elementarnych oraz pewne tzw. wzory rekurencyjne można znaleźć np. w książce [5], s. 260-273.

8.3 Rachunek całek (oznaczonych) w sensie Riemanna

Udowodnimy wpierw następujące twierdzenie nazywane podstawowym twierdzeniem rachunku całkowego lub wzorem Newtona-Leibniza.

Twierdzenie 88. *Jeśli funkcja f jest R-całkowalna na przedziale $[a, b]$ oraz F jest funkcją pierwotną funkcji f na tym przedziale, to*

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

Dowód. Wybierzmy dowolne $\varepsilon > 0$. Istnieje wówczas taki podział P przedziału $[a, b]$, że $U(f, P) - L(f, P) < \varepsilon$. Wówczas

$$U(f, P) < \varepsilon + \int_a^b f(x)dx, \quad \int_a^b f(x)dx < \varepsilon + L(f, P).$$

Na mocy Twierdzenie Lagrange'a o Wartości Średniej istnieją punkty ξ_j ($j = 1, \ldots, n$) takie, że $x_{j-1} < \xi_j < x_j$ ($P = \{x_0, x_1, \ldots, x_n\}$) oraz $F(x_j) - F(x_{j-1}) = f(\xi_j)(x_j - x_{j-1})$, a zatem

$$\begin{aligned} F(b) - F(a) &= \sum_{j=1}^{n} (F(x_j) - F(x_{j-1})) \\ &= \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1}) \leqslant U(f, P) < \varepsilon + \int_a^b f(x)dx, \\ F(b) - F(a) &\geqslant L(f, P) > -\varepsilon + \int_a^b f(x)dx. \end{aligned}$$

Stąd

$$\left| F(b) - F(a) - \int_a^b f(x)dx \right| < \varepsilon$$

i wobec dowolności $\varepsilon > 0$ otrzymujemy żądany wzór.

Przykład 29. Istnieją funkcje, które nie są całkowalne w sensie Riemanna, ale posiadają funkcje pierwotne. Niech

$$f(x) = \begin{cases} 2x \cos \frac{\pi}{x^2} - \frac{2\pi}{x} \sin \frac{\pi}{x^2} & \text{dla } x \neq 0, \\ 0 & \text{dla } x = 0 \end{cases}$$

oraz

$$F(x) = \begin{cases} x^2 \cos \frac{\pi}{x^2} & \text{dla } x \neq 0, \\ 0 & \text{dla } x = 0 \end{cases}$$

Można łatwo sprawdzić, że $F'(x) = f(x)$ dla każdego $x \in [0, 1]$. Funkcja f nie jest jednak R-całkowalna na przedziale $[0, 1]$, ponieważ nie jest na nim ograniczona. Niech bowiem $x_k = \frac{1}{\sqrt{\frac{1}{2}+2k}}$, $k \in \mathbb{N}$. Mamy

$$\lim_{k\to\infty} \frac{2\pi}{x_k} \sin\frac{\pi}{x_k^2} = \lim_{k\to\infty} \frac{2\pi}{\frac{1}{\sqrt{\frac{1}{2}+2k}}} \sin\frac{\pi}{\frac{1}{\frac{1}{2}+2k}}$$

$$= \lim_{k\to\infty} 2\pi\sqrt{\frac{1}{2}+2k}\,\sin(\frac{\pi}{2}+2k\pi) = +\infty.$$

Udowodnimy teraz reguły obliczania całek oznaczonych.

Twierdzenie 89. *(o całkowaniu przez części). Niech pochodne funkcji u i v będą R-całkowalne na przedziale $[a,b]$. Wówczas zachodzi wzór (zwany wzorem na całkowanie przez części.)*

$$\int_a^b u(x)v'(x)dx = u(x)v(x)\Big|_a^b - \int_a^b u'(x)v(x)dx,$$

gdzie $u(x)v(x)\Big|_a^b = u(b)v(b) - u(a)v(a)$.

Dowód. Istotnie, ponieważ $(uv)'(x) = u(x)v'(x) + u'(x)v(x)$, więc

$$\int_a^b (uv)'(x)dx = \int_a^b u(x)v'(x)dx + \int_a^b u'(x)v(x)dx,$$

a zatem na mocy wzoru Newtona-Leibniza otrzymujemy

$$u(x)v(x)\Big|_a^b = \int_a^b u(x)v'(x)dx + \int_a^b u'(x)v(x).$$

Wniosek 23. *Jeśli funkcja f ma na przedziale o końcach x_0 i x ciągłe pochodne do rzędu $n+1$ włącznie to*

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \ldots \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + r_n(x_0,x),$$

gdzie $r_n(x_0,x) = \frac{1}{n!}\int_{x_0}^x f^{(n+1)}(t)(x-t)^n dt$.
(Wzór Taylora dla funkcji f z resztą w postaci całkowej).

Dowód. Stosując Wzór Newtona-Leibniza i wzór na całkowanie przez części wykonujemy następujący ciąg przekształceń, w którym wszystkie różniczkowania i podstawienia wykonywane są względem t:

$$
\begin{aligned}
f(x) - f(x_0) &= \int_{x_0}^{x} f'(t)dt = -\int_{x_0}^{x} f'(t)(x-t)'dt \\
= -f'(t)(x-t)\big|_{x_0}^{x} &+ \int_{x_0}^{x} f''(t)(x-t)dt = f'(x_0)(x-x_0) - \frac{1}{2}\int_{x_0}^{x} f''(t)((x-t)^2)'dt \\
&= f'(x_0)(x-x_0) - \frac{1}{2}f''(t)(x-t)^2\big|_{x_0}^{x} + \frac{1}{2}\int_{x_0}^{x} f''(t)(x-t)^2 dt \\
&= f'(x_0)(x-x_0) + \frac{1}{2}f''(x_0)(x-x_0)^2 - \frac{1}{2\cdot 3}\int_{x_0}^{x} f'''(t)((x-t)^3)'dt \\
&= f'(x_0)(x-x_0) + \frac{1}{2}f''(x_0)(x-x_0)^2 + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n \\
&\qquad + \frac{1}{n!}\int_{x_0}^{x} f^{(n+1)}(t)(x-t)^n dt.
\end{aligned}
$$

Twierdzenie 90. *(o całkowaniu przez podstawienie). Jeśli funkcja $f : [a,b] \mapsto \mathbb{R}$ jest ciągła, a $\phi : [\alpha,\beta] \mapsto [a,b]$ ma ciągłą pochodną na przedziale $[\alpha,\beta]$ oraz $\phi(\alpha) = a$, $\phi(\beta) = b$, to*

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\phi(t))\phi'(t)dt.$$

Dowód. Ponieważ funkcje podcałkowe są ciągłe, zatem całki bo obu stronach powyższej równości istnieją. Jeśli F jest funkcją pierwotną funkcji f na przedziale $[a,b]$, to $\Phi = F \circ \phi$ jest funkcją pierwotną funkcji $(f \circ \phi)\phi'$ na przedziale $[\alpha,\beta]$. Na mocy Wzoru Newtona-Leibniza mamy zatem

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

oraz

$$\int_\alpha^\beta f(\phi(t))\phi'(t)dt = F(\phi(\beta)) - F(\phi(\alpha)) = F(b) - F(a),$$

co kończy dowód twierdzenia.

Powyższe twierdzenie jest wystarczające dla wielu zastosowań. Można udowodnić następujące jego uogólnienie:

Twierdzenie 91. *Niech $\phi : [\alpha,\beta] \mapsto [a,b]$ będzie ściśle monotinicznym przekształceniem przedziału $[\alpha,\beta]$ na przedział $[a,b]$ i niech pochodna ϕ' będzie R-całkowalna na $[a,b]$. Wówczas dla dowolnej funkcji f R-całkowalnej na przedziale $[a,b]$ funkcja $(f \circ \phi)\phi'$ jest R-całkowalna na $[\alpha,\beta]$ oraz zachodzi równość*

$$\int_{\phi(\alpha)}^{\phi(\beta)} f(x)dx = \int_\alpha^\beta f(\phi(t))\phi'(t)dt.$$

Dowód powyższego twierdzenia można znaleźć w książce [9].

Przykład 30. Obliczmy $\int_{\frac{1}{\pi}}^{\frac{2}{\pi}} \frac{1}{x^2} \sin \frac{1}{x} dx$. Niech $f(x) = \frac{1}{x^2} \sin \frac{1}{x}$, $x \in [\frac{1}{\pi}, \frac{2}{\pi}]$ oraz niech $\phi(t) = \frac{1}{t}$, $t \in [\frac{\pi}{2}, \pi]$. Na mocy twierdzenia 90 otrzymujemy

$$\int_{\frac{1}{\pi}}^{\frac{2}{\pi}} \frac{1}{x^2} \sin \frac{1}{x} dx = \int_{\pi}^{\frac{\pi}{2}} t^2 \sin t(-\frac{1}{t^2}) dt = \int_{\frac{\pi}{2}}^{\pi} \sin t dt = -\cos t\Big|_{\frac{\pi}{2}}^{\pi} = 1.$$

Udowodnimy teraz dwa twierdzenia całkowe o wartości średniej.

Twierdzenie 92. *(I Twierdzenie Całkowe o Wartości Średniej). Niech funkcje f, g będą R-całkowalne na przedziale $[a, b]$ oraz niech $m = \inf\{f(x) : x \in [a, b]\}$, $M = \sup\{f(x) : x \in [a, b]\}$. Jeśli funkcja g jest nieujemna lub niedodatnia na $[a, b]$, to*

$$\int_a^b (fg)(x) dx = \mu \int_a^b g(x) dx, \tag{50}$$

gdzie $\mu \in [m, M]$. Jeśli ponadto funkcja f jest ciągła na przedziale $[a, b]$, to istnieje punkt $\xi \in [a, b]$ taki, że

$$\int_a^b (fg)(x) dx = f(\xi) \int_a^b g(x) dx. \tag{51}$$

Dowód. Dla ustalenia uwagi załóżmy, że $g(x) \geqslant 0$ dla $x \in [a, b]$. Wówczas $mg(x) \leqslant f(x)g(x) \leqslant Mg(x)$ dla $x \in [a, b]$, a zatem

$$m \int_a^b g(x) dx \leqslant \int_a^b (fg)(x) dx \leqslant M \int_a^b g(x) dx. \tag{52}$$

Jeśli $\int_a^b g(x) = 0$, to z (52) wynika natychmiast (50). Natomiast gdy $\int_a^b g(x) \neq 0$, to biorąc $\mu = (\int_a^b g(x) dx)^{-1} \int_a^b (fg) dx$ i uwzględniając (52) otrzymujemy $m \leqslant \mu \leqslant M$. Równość (51) wynika z (50) i z tego, że funkcja ciągła na przedziale domkniętym osiąga swoje kresy i ma Własność Darboux.

Wniosek 24. *Jeśli $g(x) = 1$ dla $x \in [a, b]$, to wzór (51) przyjmuje postać*

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

Twierdzenie 93. *(II Twierdzenie Całkowe o Wartości Średniej). Jeśli funkcje f i g są R-całkowalne na przedziale $[a, b]$ i ponadto funkcja g jest monotoniczna,* *to istnieje punkt* $\xi \in [a, b]$ *taki, że*

$$\int_a^b (fg)(x)dx = g(a) \int_a^\xi f(x)dx + g(b) \int_\xi^b f(x)dx. \tag{53}$$

Dowód. Dowód przeprowadzimy w szczególnym przypadku, gdy funkcja f jest ciągła, a funkcja g jest klasy C^1. Dowód w przypadku ogólnym można znaleźć np. w [2], t. II, s. 101-102 lub w [11], s. 359-363.
Niech F będzie dowolną funkcją pierwotną funkcji f, czyli $F' = f$. Całkując przez części otrzymujemy

$$\int_a^b f(x)g(x)dx = \int_a^b F'(x)g(x)dx - F(x)g(x)\Big|_a^b - \int_a^b F(x)g'(x)dx.$$

Ponieważ funkcja g jest monotoniczna, zatem jej pochodna na przedziale $[a, b]$ ma stały znak, zatem na mocy Twierdzenia 92 mamy

$$\int_a^b F(x)g'(x) = F(\xi) \int_a^b g'(x)dx$$

dla pewnego $\xi \in [a, b]$. Stąd

$$\begin{aligned}\int_a^b f(x)g(x)dx &= F(b)g(b) - F(a)g(a) - F(\xi)g(b) + F(\xi)g(a) \\ &= g(a)(F(\xi) - F(a)) + g(b)(F(b) - F(\xi)) \\ &= g(a) \int_a^\xi f(x)dx + g(b) \int_\xi^b f(x)dx.\end{aligned}$$

Uwaga 44. II Twierdzenie Całkowe o Wartości Średniej bywa podawane w różnych postaciach.

(a) Jeśli w przedziale $[a, b]$ funkcja g jest nierosnąca i nieujemna, a funckja f jest R-całkowalna to

$$\int_a^b (fg)(x)dx = g(a) \int_a^\xi f(x)dx, \tag{54}$$

gdzie ξ jest pewnym punktem z przedziału $[a, b]$.

(b) Analogicznie, jeśli funkcja g jest niemalejąca i nieujemna, to zachodzi wzór

$$\int_a^b (fg)(x)dx = g(b) \int_\xi^b f(x)dx, \tag{55}$$

gdzie $\xi \in [a, b]$.

Wzory (53), (54), (55) nazywają się wzorami Bonneta.

8.4 Całka z funkcji o wartościach zespolonych

Definicja 77. Niech f_1, f_2 będą funkcjami rzeczywistymi określonymi na przedziale $[a, b]$ i niech $f = (f_1, f_2)$ będzie odwzorowaniem przedziału $[a, b]$ w zbiór $\mathbb{C}$. Mówimy, że f jest całkowalna w sensie Riemanna na przedziale $[a, b]$, jeśli funkcje f_1, f_2 są R-całkowalne na tym przedziale. W tym wypadku określamy

$$\int_a^b f(x)dx = \left(\int_a^b f_1(x)dx, \int_a^b f_2(x)dx \right),$$

lub równoważnie

$$\int_a^b f(x)dx = \int_a^b \mathrm{Re} f(x)dx + i \int_a^b \mathrm{Im} f(x)dx.$$

Jest oczywiste, że Twierdzenie 82 (dla sumy oraz iloczynu przez liczbę rzeczywistą funkcji R-całkowalnych) jest prawdziwe także dla funkcji o wartościach w $\mathbb{C}$. To samo dotyczy twierdzeń 85, 86, 88, 90 (por. [1], s. 272-273). Aby się o tym przekonać, należy jedynie zastosować poprzednie rezultaty do poszczególnych współrzędnych. Dla przykłady sformułujemy Podstawowe Twierdzenie Rachunku Całkowego.

Twierdzenie 94. *Niech f i F będą funkcjami określonymi na przedziale $[a, b]$ o wartościach w $\mathbb{C}$. Jeśli odwzorowanie f jest R-całkowalne na tym przedziale oraz $F'(x) = f(x)$ dla $x \in [a, b]$, to*

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

Prawdziwy jest również analog Twierdzenia 83 (b), jednakże jego dowód jest bardziej subtelny.

Twierdzenie 95. *Jeśli odwzorowanie $f : [a, b] \mapsto \mathbb{C}$ jest R-całkowalne, to funkcja $|f|$ jest również R-całkowalna oraz*

$$\left| \int_a^b f(x)dx \right| \leqslant \int_a^b |f(x)|dx.$$

Dowód. Niech f_1, f_2 będą składowymi odwzorowania f (to znaczy $f = (f_1, f_2)$). Wówczas $|f| = \sqrt{f_1^2 + f_2^2}$. Każda z funkcji f_1^2, f_2^2 jest R-całkowalna, więc z ciągłości pierwiastka i Twierdzenia 81 wynika, że funkcja $|f|$ jest również całkowalna. Niech $y = (y_1, y_2)$, gdzie $y_i = \int_a^b f_i(x)dx$ dla $i = 1, 2$. Wówczas $y = \int_a^b f(x)dx$ oraz

$$|y|^2 = y_1^2 + y_2^2 = y_1 \int_a^b f_1(x)dx + y_2 \int_a^b f_2(x)dx = \int_a^b (y_1 f_1(x) + y_2 f_2(x))dx.$$

Na podstawie nierówności Schwarza mamy

$$y_1 f_1(x) + y_2 f_2(x) \leqslant |y||f(x)| \text{ dla } x \in [a,b],$$

a zatem

$$|y|^2 \leqslant |y| \int_a^b |f(x)|dx.$$

Dzieląc ostatnią nierówność przez $|y| \neq 0$, otrzymujemy tezę (dla $y = 0$ twierdzenie jest oczywiste).

W przypadku funkcji R-całkowalnych określonych na przedziale $[a,b]$ o wartościach w $\mathbb{C}$, Wniosek 24 nie zachodzi (por. Rozdział 7.6). Prawdziwe jest natomiast następujące

Twierdzenie 96. *Jeśli $f : [a,b] \mapsto \mathbb{C}$ jest funkcja R-całkowalną oraz $f([a,b]) \subset B(x^0, r)$, gdzie $B(x^0, r)$ oznacza kule domkniętą o środku w punkcie $x^0 \in \mathbb{C}$ i promieniu r, to*

$$\frac{1}{b-a}\int_a^b f(x)dx \in B(x^0, r). \tag{56}$$

Dowód. Rozważmy funkcję $g(t) = f(t) - x^0$, $t \in [a,b]$. Mamy

$$\begin{aligned}\left|\frac{1}{b-a}\int_a^b f(t)dt - x^0\right| &= \frac{1}{b-a}\left|\int_a^b (f(t) - x^0)dt\right| \\ &\leqslant \frac{1}{b-a}\int_a^b |f(t) - x^0|dt \leqslant \frac{1}{b-a}(b-a)r = r,\end{aligned}$$

a zatem (56) zachodzi.

8.5 Zastosowania całki Riemanna

Wiele zastosowań całki Riemanna opiera się na następującym twierdzeniu, które podaje warunek na to, aby addytywna funkcja przedziału była generowana przez całkę.

Twierdzenie 97. *Jeśli dla addytywnej funkcji $J(\alpha, \beta)$ określonej dla punktów $\alpha, \beta \in [a,b]$ istnieje funkcja R-całkowalna na $[a,b]$ i taka, że*

$$\inf_{x\in[\alpha,\beta]} f(x)(\beta - \alpha) \leqslant J(\alpha,\beta) \leqslant \sup_{x\in[\alpha,\beta]} f(x)(\beta-\alpha)$$

dla dowolnych $a \leqslant \alpha < \beta \leqslant b$, to

$$J(a,b) = \int_a^b f(x)dx.$$

Dowód. Niech $P = \{x_0, \ldots, x_n\}$ będzie dowolnym podziałem przedziału $[a, b]$ i niech $m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}$ oraz $M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}$ $(i = 1, \ldots, n)$. Dla dowolnego przedziału $[x_{i-1}, x_i]$ mamy

$$m_i \Delta x_i \leqslant J(x_{i-1}, x_i) \leqslant M_i \Delta x_i.$$

Sumując powyższe nierówności i korzystając z addytywności funkcji $J(\alpha, \beta)$ otrzymujemy

$$L(f, P) = \sum_{i=1}^{n} m_i \Delta x_i \leqslant J(a, b) \leqslant \sum_{i=1}^{n} M_i \Delta x_i = U(f, P).$$

Niech będzie dane dowolne $\varepsilon > 0$. Wówczas istnieje taki podział P przedziału $[a, b]$, że $U(f, P) - L(f, P) < \varepsilon$. Mamy zatem

$$\int_a^b f(x)dx - \varepsilon < L(f, P) \leqslant J(a, b) \leqslant U(f, P) < \int_a^b f(x)dx + \varepsilon,$$

czyli $\left|J(a, b) - \int_a^b f(x)dx\right| < \varepsilon$, a zatem wobec dowolności $\varepsilon > 0$ otrzymujemy tezę.

Omówimy teraz kilka geometrycznych zastosowań całki Riemanna. Zajmiemy się wpierw zagadnieniem długości krzywej.

Definicja 78. Ciągłe odwzorowanie γ przedziału $[a, b]$ w zbiór $\mathbb{R}$ lub $\mathbb{C}$ nazywamy krzywą lub drogą w $\mathbb{R}$ lub $\mathbb{C}$. Punkty $A = \gamma(a)$, $B = \gamma(b)$, nazywają się odpowiednio początkiem i końcem drogi. Jeśli $\gamma(a) = \gamma(b)$, to powiemy, że γ jest krzywą zamkniętą. Zbiór $\gamma([a, b])$ nazywamy obrazem krzywej γ.

Uwaga 45. Zauważmy, że jeden i ten sam zbiór może być obrazem wielu różnych krzywych. Ponadto może on okazać się nie tym, co w naszym potocznym wyobrażeniu jest linią. Istnieją przykłady krzywych, które wypełniają cały kwadrat jednostkowy (tak zwane „Krzywe Peano”, zob. R. Eugelking, K. Sieklucki, Geometria i topologia, cz-II, s. 134-135).

Definicja 79. Krzywą $\gamma : [a, b] \mapsto \mathbb{R}$ (lub $\mathbb{C}$) nazywamy łukiem, jeśli odwzorowanie γ jest wzajemnie jednoznacznie. Krzywą zamkniętą $\gamma : [a, b] \mapsto \mathbb{R}$ (lub $\mathbb{C}$) nazywamy łukiem zamkniętym, jeśli funkcja γ jest wzajemnie jednoznaczna na przedziale $[a, b)$.

Definicja 80. Z każdym podziałem $P = \{x_0, x_1, \ldots x_n\}$ przedziału $[a, b]$ i z dowolną krzywą γ w $\mathbb{R}$ lub $\mathbb{C}$ wiążemy liczbę

$$V(P, \gamma) = \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})|,$$

gdzie i-ty składnik powyższej sumy oznacza odległość pomiędzy puntkami $\gamma(x_{i-1})$ oraz $\gamma(x_i)$. Zatem $V(P,\gamma)$ jest długością łamanej o wierzchołkach $\gamma(x_1),\dots\gamma(x_n)$ występujacych w takim porządku.
Długością krzywej γ nazywamy liczbę

$$L(\gamma) = l(a,b) = \sup_P V(P,\gamma),$$

gdzie kres górny jest wzięty po zbiorze wszytkich podziałów przedziału $[a,b]$. Jeśli $L(\gamma) < +\infty$, to powiemy, że γ jest krzywą prostowalną.

W pewnych przypadkach możemy obliczać $L(\gamma)$ jako całkę Riemanna.

Definicja 81. Krzywa $\gamma : [a,b] \mapsto \mathbb{R}$ (lub $\mathbb{C}$) nazywa się krzywą danej klasy gładkości, jeśli funkcja γ należy do tej klasy (to znaczy do klasy $C^{(n)}$ przy pewnym n). Krzywą klasy C^1 nazywamy krzywą gładką. Natomiast mówimy, że krzywa jest kawałkami gładka, jeśli przedział $[a,b]$ można podzielić na skończoną liczbę przedziałów tak, że na każdym z nich funkcja γ jest klasy C^1.

Twierdzenie 98. *Jeśli krzywa $\gamma : [a,b] \mapsto \mathbb{R}$ (lub $\mathbb{C}$) jest gładka, to γ jest prostowalna oraz*

$$L(\gamma) = \int_a^b |\gamma'(t)|dt. \tag{57}$$

Dowód. Jest jasne, że dla $a \leqslant \alpha < \beta \leqslant b$ zachodzi równość $L(\alpha,\gamma) = L(\alpha,\beta) + L(\beta,\gamma)$. Ponadto wobec Twierdzenia Lagrange'a (Tw. 62) mamy

$$\inf_{t\in[\alpha,\beta]} |\gamma'(t)|(\beta-\alpha) \leqslant L(\alpha,\beta) \leqslant \sup_{t\in[\alpha,\beta]} |\gamma'(t)|(\beta-\alpha).$$

Wzór (57) jest więc konsekwencją Twierdzenia 97 (to, że krzywa γ jest prostowalna wynika z Twierdzenia Lagrange'a i założenia, że $\gamma \in C^1$; mamy bowiem $L(\gamma) \leqslant \sup\limits_{t\in[a,b]} \gamma'(t)(b-a) < +\infty$).

Uwaga 46.

(a) Niech będzie dana krzywa na płaszczyźnie (w $\mathbb{C}$) o równaniach $x = x(t)$, $y = y(t)$ $(\gamma(t) = (x(t), y(t))), t \in [a,b]$. Jeśli krzywa γ jest gładka, to na mocy 57 jej długość wyrażą się wzorem

$$L(\gamma) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2}dt. \tag{58}$$

(b) Długość wykresu funkcji $y = f(x)$, $x \in [a,b]$, klasy C^1 wyraża się wzorem

$$L(a,b) = \int_a^b \sqrt{1 + f'(x)^2}dx.$$

Wynika to ze wzoru (58) bowiem $\{(x, f(x)) : x \in [a, b]\} = \gamma([a, b])$, gdzie $\gamma(t) = (t, f(t))$ dla $t \in [a, b]$.

(c) **Współrzędne biegunowe**
Przez dowolny punkt O zwany biegunem poprowadzimy oś S, mającą początek w tym punkcie. Współrzędnymi biegunowymi punktu P nazywamy liczbe r będącą długościa wektora $\overrightarrow{OP}$ i liczbę $\phi \in [0, 2\pi)$ będącą miarą łukową kąta skierowanego $\angle(S, \overrightarrow{OP})$. Liczbę ϕ nazywamy amplitudą punktu P (biegunowi, to znaczy punktowi O, można przyporządkować dowolną amplitudę), natomiast r - promieniem wodzącym punktu P.
Obierzmy układ kartezjański prostokątny i układ biegunowy tak, by biegun leżał w początku układu kartezjańskiego, a oś biegunowa pokrywała się z osią odciętych. Oznaczmy przez (x, y) współrzędne prostokątne oraz przez (r, ϕ) - współrzędne biegunowe tego samego punktu w obu układach. Wówczas zachodzą następujące związki:

$$\begin{aligned}
&x = r\cos\phi \\
&y = r\sin\phi \\
&r = \sqrt{x^2 + y^2} \text{ dla } x^2 + y^2 > 0 \\
&\cos\phi = \frac{x}{\sqrt{x^2 + y^2}}, \qquad\qquad \sin\phi = \frac{y}{\sqrt{x^2 + y^2}}.
\end{aligned}$$

Załóżmy, że funkcja g klasy C^1 jest określona we współrzędnych biegunowych, to znaczy $r = g(\phi)$, gdzie $\phi_1 \leqslant \phi \leqslant \phi_2$. Za pomocą wzorów $x = r\cos\phi = g(\phi)\cos\phi$, $y = r\sin\phi = g(\phi)\sin\phi$ otrzymujemy przedstawienie parametryczne funkcji g, zatem

$$L(\phi_1, \phi_2) = \int_{\phi_1}^{\phi_2} \sqrt{g(\phi)^2 + g'(\phi)^2} d\phi$$

(d) Gdy krzywa jest kawałkami gładka, to aby obliczyć jej długość dzielimy ją na skończoną liczbę krzywych gładkich i do każdej z nich stosujemy wzór (57), a następnie dodajemy otrzymane liczby.

Zajmiemy się teraz zastosowaniem całek Riemanna do obliczania pola powierzchni.

Definicja 82. Rozpatrzmy na płaszczyźnie dowolną figurę P, która jest obszarem ograniczonym. Załóżmy, że brzeg figury P jest obrazem gładkiej krzywej zamkniętej $\gamma : [a, b] \mapsto \mathbb{C}$ (lub składa się z obrazów kilku takich krzywych). Zakładamy, że $\gamma'(t) \neq 0$ dla każdego $t \in [a, b]$. Jordan udowodnił, że rozważana krzywa zamknięta rozcina płaszczyznę na dwa obszary, wewnętrzny i zewnętrzny, dla których jest ona wspólnym brzegiem. Rozważmy wszystkie możliwe wielokąty A, całkowicie zawarte w P i wielokąty B, całkowicie zawierające obszar P. Jeśli $|A|$ i $|B|$ oznaczają pola tych wielokątów, to $|A| \leqslant |B|$. Zbiór

$\{|A| : A \subset P\}$ jest ograniczony z góry przez którąkolwiek z liczb $|B|$, a zatem na mocy Aksjomatu Dedekinda posiada on kres górny $|P_*|$. Natomiast zbiór $\{|B| : P \subset B\}$ jest ograniczony z dołu przez liczbę $|P_*|$ i posiada kres dolny $|P^*|$. Kres górny $|P_*|$ nazywamy wewnętrzną, a kres dolny $|P^*|$ — zewnętrzną miarą Jordana figury P.
Jeśli $|P_*| = |P^*|$, to mówimy, że figura P jest mierzalna w sensie Jordana, a wspólną wartość tych kresów nazywamy miarą Jordana lub polem figury P i oznaczamy symbolem $|P|$.
Będziemy mówili, że figura ma pole równe zeru, jeśli można pokryć ją obszarem wielokątnym o dowolnie małym polu.

Uwaga 47.

(a) Umieśćmy rozpatrywaną figurę P wewnątrz prostokąta R o bokach równoległych do osi układu współrzędnych. Prostokąt ten rozbijamy na części za pomocą pewnej liczby prostych równoległych do jego boków. Oznaczmy przez $\tilde{A}$ figurę złożoną z prostokątów całkowicie zawartych w P, natomiast przez $\tilde{B}$ — figurę złożoną z prostokątów mających punkty wspólne z P. Przez d oznaczmy długość najdłuższej z przekątnych prostokątów. Można udowodnić, że figura P jest mierzalna w sensie Jordana wtedy i tylko wtedy, gdy przy $d \to 0$ obydwa pola $|\tilde{A}|$ i $|\tilde{B}|$ dążą do wspólnej granicy $|P|$; jeśli ten warunek jest spełniony, to wspólna granica $|P|$ jest równa polu figury P ([2], t. II, s. 163 — 164).

(b) Załóżmy, że figura P jest rozcięta na dwie figury P_1 i P_2 (to znaczy figury P_1 i P_2 nie mają punktów wewnętrznych wspólnych). Można udowodnić, że mierzalność dwóch pośród trzech figur P, P_1, P_2, pociąga za sobą mierzalność trzeciej, przy czym $|P| = |P_1| + |P_2|$, to znaczy pole figury ma własność addytywności. ([2], t. II, s. 162).

Można nietrudno udowonić ([2], t. II, s. 161, 164 — 165) następujące kryteria mierzalności figur w sensie Jordana.

Twierdzenie 99.

(a) Na to, by figura P była mierzalna w sensie Jordana, potrzeba i wystarcza, żeby dla każdego $\varepsilon > 0$ można było znaleźć takie dwa wielokąty $A \subset P$ i $P \subset B$, że $|B| - |A| < \varepsilon$.

(b) Na to, żeby figura P była mierzalna w sensie Jordana potrzeba i wystarcza, żeby jej brzeg miał pole równe zeru.

(c) Jeśli figura P jest ograniczona wykresami kilku funkcji ciągłych, z których każda jest w postaci $y = f(x)$, $x \in [a, b]$ lub $x = g(y)$, $y \in [c, d]$, to figura jest mierzalna w sensie Jordana.

Niech f będzie nieujemną funkcją ciągłą określoną na przedziale $[a,b]$. Niech $A=f(a)$, $B=f(b)$. Rozważmy figurę $aABb$ zwaną trapezem krzywolinowym. Jest ona ograniczona pionowymi odcinkami aA, bB, odcinkiem $[a,b]$ na osi OX oraz wykresem funkcji f. Dla $a \leqslant \alpha < \beta \leqslant b$ oznaczmy przez $S(\alpha,\beta)$ pole trapezu krzywolinowego $\alpha f(\alpha)f(\beta)\beta$. Ponadto przyjmijmy $S(\beta,\alpha)=-S(\alpha,\beta)$. Wiemy, że pole jest funkcją addytywną, to znaczy jeśli $a \leqslant \alpha < \beta < \gamma \leqslant b$, to

$$S(\alpha,\beta)+S(\beta,\gamma)=S(\alpha,\gamma).$$

Stąd wynika, że $S(\alpha,\beta)$ jest addytywną funkcją przedziału zorientowanego. Ponieważ pole figury zawartej w innej figurze nie może być większe od pola tej „obejmującej" figury, więc

$$\inf_{x\in[\alpha,\beta]} f(x)(\beta-\alpha) \leqslant S(\alpha,\beta) \leqslant \sup_{x\in[\alpha,\beta]} f(x)(\beta-\alpha).$$

Na mocy Twierdzenia 97 otrzymujemy następujący wzór na pole trapezu krzywolinowego

$$S(a,b)=\int_a^b f(x)dx. \tag{59}$$

Jest jasne, że jeśli $f(x) \leqslant 0$ dla $x\in[a,b]$, to $S(a,b)=-\int_a^b f(x)dx$. Ponadto, jeśli trapez krzywolinowy $CABD$ jest ograniczony z góry i z dołu wykresami krzywych o równaniach $y=f(x)$ i $y=g(x)$ ($a\leqslant x\leqslant b$), to znaczy $f(x)\geqslant g(x)$ dla $x\in[a,b]$, to oznaczając przez $|P|$ pole tego trapezu mamy

$$|P|=\int_a^b |f(x)-g(x)|dx.$$

Niech teraz będzie dany wycinek AOB ograniczony promieniami OA, OB (każdy z promieni OA, OB może być punktem) i wykresem funkcji ciągłej o równaniu biegnowym $r=g(\phi)>0$ dla $\phi_1\leqslant\phi\leqslant\phi_2$ gdzie $\phi_2-\phi_1<2\pi$ ($A=g(\phi_1)$, $B=g(\phi_2)$). Oznaczmy przez $S(\phi_1,\phi_2)$ pole wycinka AOB. Mamy

$$\inf_{\phi\in[\phi_\alpha,\phi_\beta]} \frac{1}{2}g^2(\phi)(\phi_\beta-\phi_\alpha) \leqslant S(\phi_\alpha,\phi_\beta) \leqslant \sup_{\phi\in[\phi_\alpha,\phi_\beta]} \frac{1}{2}g^2(\phi)(\phi_\beta-\phi_\alpha)$$

(przypomnijmy, że pole wycinka kołowego o promieniu r i kącie środkowym $\Delta\phi$ wyraża się wzorem $\frac{1}{2}r^2\Delta\phi$). Stąd otrzyjmujemy

$$S(\phi_1,\phi_2)=\frac{1}{2}\int_{\phi_1}^{\phi_2} r^2 d\phi.$$

Roważmy, jeszcze przypadek, gdy mamy krzywą gładką o równaniach $x=x(t)$, $y=y(t)$, $t\in[a,b]$. Jeśli $y(t)\geqslant 0$ oraz $x'(t)>0$ w przedziale $[a,b]$, to pole obszaru D, zawartego między obrazem tej krzywej, osią OX i rzędnymi w punktach końcowych krzywej, wyraża się wzorem

$$|D|=\int_a^b y(t)x'(t)dt.$$

Istotnie funkcja $x = x(t)$ jest rosnąca, bowiem $x'(t) > 0$, a zatem posiada funkcję odwrotną $t = t(x)$ w przedziale $[\alpha, \beta]$, gdzie $\alpha = x(a)$, $\beta = x(b)$. Obraz danej krzywej możemy więc traktować jako wykres funkcji $y = y(t(x))$, $\alpha < x < \beta$, przy czym na mocy wzoru (59)

$$|D| = \int_\alpha^\beta y(t(x))dx.$$

Wykonując w powyższej całce podstawienie $x = x(t), a \leqslant t \leqslant b$, otrzymujemy żądany wzór (por. Tw. 90).

Jeśli natomiast $x'(t) < 0$ w przedziale $[a, b]$, to $|D| = -\int_a^b y(t)x'(t)dt$.

Całka Riemanna posiada również zastosowania do obliczania objętności pewnych brył oraz pól powierzchni obrotowych. Niech będzie dana bryła V, to znaczy obszar ograniczony w przestrzeni trójwymiarowej. Załóżmy, że brzeg S bryły V jest powierzchnią zamkniętą (lub składa się z kilku takich powierzchni).

Definicja 83. Rozpatrzmy wielościany X o objętości $|X|$, zawarte całkowicie w bryle V oraz wielościany Y o objętości $|Y|$ zawierjące w sobie całą bryłe V. Istnieje kres górny $|V_*|$ liczb $|X|$ i kres dolny $|V^*|$ liczb $|Y|$, przy czym $|V_*| \leqslant |V^*|$; kresy te nazywamy odpowiednio wewnętrzną i zewnętrzną objetością bryły V. Jeśli oba kresy są równe, to ich wspólną wartość nazywamy objętością bryły V. Bryła ma objętość równą zeru, jeśli można umieścić ją w bryle wielościennej o dowolnie małej objętości.

Również w tym przypadku są prawdziwe odpowiedniki Uwagi 47 i Twierdzenia 99, w których odpowiednio zastąpimy figury — bryłami, wielokąty — wielościanami, prostokąty — prostopadłościanami, pola — objętościami, funkcje ciągłe jednej zmiennej — funkcjami ciągłymi dwóch zmiennych.
W szczególności można udowodnić (por. [2], t. II, s. 177 — 178), że jeśli trape krzywolinowy $aABb$ „obrócimy" względem osi OX, to otrzymana w ten sposób bryła posiada objętość. Istotnie ponieważ wykres funkcji ciągłej $y = f(x)$, $x \in [a, b]$ można pokryć prostokątami o dowolnie małym polu, zatem brzeg rozważanej bryły można pokryć bryłami (pierścieniami walcowymi) o dowolnie małej objętości (to znaczy suma objętości tych brył jest dowolnie mała).
Oznaczmy przez $V(\alpha, \beta)$ objętość bryły otrzymanej przez obrót trapezu krzywolinowego $\alpha f(\alpha) f(\beta) \beta$ dookoła osi OX. Z własności objętości brył otrzymujemy następujące związki: jeśli $a \leqslant \alpha < \beta < \gamma \leqslant b$, to

$$V(\alpha, \gamma) = V(\alpha, \beta) + V(\beta, \gamma)$$

oraz

$$\pi\left(\inf_{x\in[\alpha,\beta]} f(x)\right)^2 (\beta - \alpha) \leqslant V(\alpha, \beta) \leqslant \pi\left(\sup_{x\in[\alpha,\beta]} f(x)\right)^2 (\beta - \alpha),$$

przy czym w powyższych nierównościach zastosowaliśmy wzór na objętość walca, łatwo wynikający z Definicji 83 (zob. [2], t. II, s. 176). Na mocy Twierdzenia 97 otrzymujemy

$$V(a,b) = \pi \int_a^b f(x)^2 dx.$$

Jest jasne, że jeśli trapez krzywoliniowy jest ograniczony z góry i z dołu wykresami funkcji ciągłych $y = f_1(x), y = f_2(x)$, $f_1(x) \geqslant f_2(x) \geqslant> 0$ $(x \in [a,b])$, to

$$V(a,b) = \pi \int_a^b |f_1(x)^2 - f_2(x)^2| dx.$$

Niniejszy paragraf zakończymy krótką informacją o polu powierzchni obrotowej. Ze względu na niewystarczjacy aparat pojęciowy, nie możemy jeszcze zdefiniować pola powierzchni zakrzywionej.
Niech będzie dana krzywa w $\mathbb{C} : \gamma(t) = (x(t), y(t))$, $t \in [a,b]$. Jeśli krzywa γ jest gładka, to pole powierzchni obrotowej powstałej przez obrót obrazu krzywej γ dookoła osi OX istnieje i wyraża się wzorem

$$|P| = 2\pi \int_a^b y(t)\sqrt{x'(t)^2 + y'(t)^2} dt.$$

W szczególności jeśli „obracamy” wykres funkcji $y = f(x)$ $(a \leqslant x \leqslant b)$ klasy C^1 dookoła osi OX, to

$$|P| = 2\pi \int_a^b f(x)\sqrt{1 + f'(x)^2} dx.$$

8.6 Całki niewłaściwe

Definicja 84. Niech f będzie funkcją rzeczywistą określoną na przedziale $[a,b)$ i R-całkowalną w każdym przedziale domkniętym $[a,\beta]$, gdzie $a < \beta < b$. Wobec tego dla każdego $a < \beta < b$ istnieje całka

$$J(\beta) = \int_a^\beta f(x) dx. \tag{60}$$

Punkt b nazywać będziemy punktem osobliwym funkcji f, jeśli albo $b = +\infty$, albo funkcja f nie jest ograniczona na przedziale $[a,b)$.
Jeśli b jest punktem osobliwym funkcji f i całka (60) dąży do skończonej granicy, gdy $\beta \to b$, to tę granicę nazywamy całką niewłaściwą z funkcji f na przedziale $[a,b)$ i oznaczamy symbolem

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

O funkcji f mówimy wówczas, że jest całkowalna w sensie niewłaściwym na przedziale $[a,b)$. Mamy zatem $\int_a^b f(x)dx = \lim_{\beta\to b} \int_a^\beta f(x)dx$. Jeśli ta granica nie istnieje, to mówimy, że całka niewłaściwa jest rozbieżna. Analogicznie określamy całkę niewłaściwą z funkcji rzeczywistej określonej na przedziale $(a, b]$, gdy a jest punktem osobliwym funkcji f, to znaczy $a = -\infty$ albo funkcja f nie jest ograniczona w otoczeniu punktu a. Zakładając, że funkcja f jest R-całkowalna na każdym przedziale $[\alpha, b]$, gdzie $a < \alpha < b$, przyjmujemy

$$\int_a^b f(x)dx = \lim_{\alpha \to a} \int_\alpha^b f(x)dx$$

(o ile ta granica istnieje).
W przypadku, gdy $b < a$ przyjmujemy

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

Jeśli funkcja f ma w przedziale $[a, b]$ skończenie wiele punktów osobliwych, to dzielimy ten przedział na przedziały mające po jednym punkcie osobliwym na początku lub na końcu przedziału i wówczas sumę całek niewłaściwych odpowiadających tym podprzedziałom nazywamy całką niewłaściwą funkcji f na przedziale $[a, b]$.

Przykład 31.

(a) Zbadamy, dla jakich $p > 0$ istnieje całka niewłaściwa $I_p = \int_a^b \frac{dx}{(b-x)^p}$. Badamy całki

$$J_p(\beta) = \int_a^\beta \frac{dx}{(b-x)^p} = \begin{cases} -\ln(b-x)\Big|_a^\beta & \text{dla } p = 1, \\ -\frac{(b-x)^{-p+1}}{-p+1}\Big|_a^\beta & \text{dla } p \neq 1. \end{cases}$$

$$= \begin{cases} -\frac{(b-\beta)^{1-p}}{1-p} + \frac{(b-a)^{1-p}}{1-p} & \text{dla } 0 < p < 1, \\ \ln\frac{b-a}{b-\beta} & \text{dla } p = 1, \\ -\frac{1}{(1-p)(b-\beta)^{p-1}} + \frac{1}{(1-p)(b-a)^{p-1}} & \text{dla } p > 1. \end{cases}$$

Stąd $\lim_{\beta \to b} J_p(\beta) = \begin{cases} (b-a)^{1-p} & \text{dla } 0 < p < 1, \\ +\infty & \text{dla } p \geqslant 1, \end{cases}$ czyli całka niewłaściwa $I_p = \int_a^b \frac{dx}{(b-x)^p}$ istnieje dla $0 < p < 1$ i nie istnieje dla $p \geqslant 1$. W szczególności $\int_0^1 \frac{dx}{\sqrt{x}}$ istnieje, natomiast $\int_0^1 \frac{dx}{x}$ nie istnieje.

(b) Zbadamy teraz, dla jakich $p > 0$ istnieje całka niewłaściwa $K_p = \int_a^\infty \frac{dx}{x^p}$. Badamy całki

$$K_p(\beta) = \int_a^\beta \frac{dx}{x^p} = \begin{cases} \ln x\Big|_a^\beta & \text{dla } p = 1, \\ \frac{x^{-p+1}}{-p+1}\Big|_a^\beta & \text{dla } p \neq 1, \end{cases}$$

$$= \begin{cases} \frac{\beta^{1-p}}{1-p} - \frac{a^{1-p}}{1-p} & \text{dla } 0 < p < 1, \\ \ln\frac{\beta}{a} & \text{dla } p = 1, \\ \frac{1}{(1-p)\beta^{p-1}} - \frac{1}{(1-p)a^{p-1}} & \text{dla } p > 1. \end{cases}$$

Zatem $\lim\limits_{\beta\to+\infty} K_p(\beta) = \begin{cases} +\infty & \text{dla } 0 < p \leqslant 1, \\ \frac{1}{(p-1)a^{p-1}} & \text{dla } p > 1, \end{cases}$ a więc całka niewłaściwa K_p istnieje dla $p > 1$ i nie istnieje, gdy $0 < p \leqslant 1$.

Następujące twierdzenie podaje podstawowe własności całek niewłaściwych.

Twierdzenie 100.

(a) Jeśli funkcja f jest ograniczona na przedziale skończonym $[a,b)$ i jeśli całka $I(\beta) = \int_a^\beta f(x)dx$ istnieje dla każdego $a < \beta < b$, to granica $\lim\limits_{\beta\to b} \int_a^\beta f(x)dx$ istnieje.

(b) Jeśli funkcja f jest R-całkowalna na przedziale $[a,b]$, to całka Riemanna i całka niewłaściwa funkcji f (na przedziałach $[a,b]$ oraz $[a,b)$, odpowiednio) pokrywają się.

(c) Jeśli funkcje $f, g : [a,b] \mapsto \mathbb{R}$ są całkowalne (w sensie niewłaściwym) na przedziale $[a,b)$, to dla dowolnych $\lambda_1, \lambda_2 \in \mathbb{R}$ funkcja $(\lambda_1 f + \lambda_2 g)(x)$ jest całkowalna (w sensie całki niewłaściwej) oraz

$$\int_a^b (\lambda_1 f + \lambda_2 g)(x)dx = \lambda_1 \int_a^b f(x)dx + \lambda_2 \int_a^b g(x)dx.$$

Ponadto, jeśli $c \in [a,b]$, to $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$.

(d) Załóżmy, że funkcja f jest całkowalna w sensie niewłaściwym na przedziale $[a,b)$ oraz $\phi : [\alpha,\gamma) \mapsto [a,b)$ jest ściśle monotoniczną funkcją klasy C^1, przy czym $\phi(\alpha) = a$ oraz $\phi(\beta) \to b$ jeśli $\beta \to \gamma$, $\beta \in [\alpha,\gamma)$. Wówczas całka niewłaściwa funkcji $t \to (f \circ \phi)(t)\phi'(t)$ na przedziale $[\alpha,\gamma)$ istnieje oraz

$$\int_a^b f(x)dx = \int_a^\gamma (f \circ \phi)(t)\phi'(t)dt$$

(zamiana zmiennych w całkach niewłaściwych).

(e) Jeśli $f, g \in C^1[a,b)$, to

$$\int_a^b (fg')(x)dx = (fg)(x)\Big|_a^b - \int_a^b (f'g)(x)dx,$$

gdzie $(fg)(x)\big|_a^b = \lim\limits_{x\to b}(fg)(x) - (fg)(x)$, o ile dwa spośród trzech występujących w równości wyrażeń mają sens. Stąd już wynika istnienie trzeciego (całkowanie przez części dla całek niewłaściwych).

Dowód.

(a) Istotnie, jeśli $a < \beta < \beta' < b$, to $|I(\beta)-I(\beta')| = |\int_\beta^{\beta'} f(x)dx| \leqslant M(\beta'-\beta)$, a zatem jeśli $\varepsilon > 0$, $b-\beta < \frac{\varepsilon}{M}$, $b-\beta' < \frac{\varepsilon}{M}$, to $|I(\beta) - I(\beta')| < \varepsilon$. Na mocy Kryterium Cauchy'ego (Tw. 36) wnioskujemy, że $\lim\limits_{\beta \to b} I(\beta)$ istnieje.

(b) Wynika z ciągłości funkcji $I(\beta) = \int_a^\beta f(x)dx$, na przedziale $[a,b]$, na którym funkcja f jest R-całkowalna.

(c) Pierwsza część tezy wynika, z tego, że dla $\beta \in [a,b)$ mamy

$$\int_a^\beta (\lambda_1 f + \lambda_2 g)(x)dx = \lambda_1 \int_a^\beta f(x)dx + \lambda_2 \int_a^\beta g(x)dx.$$

Druga część tezy wynika z równości

$$\int_a^\beta f(x)dx = \int_a^c f(x)dx + \int_c^\beta f(x)dx,$$

która zachodzi dla dowlnych $c, \beta \in [a,b)$.

(d) Na podstawie Twierdzenia (91) otrzymujemy

$$\int_{a=\phi(\alpha)}^{\phi(\beta)} f(x)dx = \int_\alpha^\beta (f \circ \phi)(t)\phi'(t)dt \text{ dla } \beta \in [a,\gamma),$$

a następnie przechodzimy do granicy przy $\beta \to \gamma$.

Uwaga 48.

(a) Jeżeli przy założeniach punktu (a) powyższego twierdzenia nadamy funkcji f dowolną wartość, to otrzymamy funkcję R-całkowalną na przedziale $[a,b]$ oraz całka Riemanna z tej funkcji na przedziale $[a,b]$ jest równa $\lim\limits_{\beta \to b} \int_a^\beta f(x)dx$. Aby to udowodnić, wystarczy dla przykładu wykorzystać Uwagę 38, a następnie własności funkcji górnej granicy całkowania (dla całki Riemanna).

(b) Iloczyn dwóch funkcji $f, g : [a,b] \mapsto \mathbb{R}$ całkowalnych w sensie niewłaściwym nie musi być funkcją całkowalną w sensie niewłaściwym. Dla przykładu $\int_0^1 \frac{dx}{\sqrt{1-x}}$ istnieje natomiat $\int_0^1 \frac{dx}{\sqrt{1-x}\sqrt{1-x}}$ nie istnieje.

(c) Dla przykładu wobec Tw. 100 (a) całka niewłaściwa $\int_0^1 \frac{\sin x}{x} dx$ istnieje, bowiem funkcja $f(x) = \frac{\sin x}{x}$ jest ograniczona na przedziale $(0,1]$ oraz R-całkowalna na każdym przedziale $[\alpha, 1]$, $0 < \alpha < 1$.

Definicja 85. Niech funkcja f ma w przedziale $[a,b]$ skończoną liczbę punktów osobliwych. Jeśli całka niewłaściwa $\int_a^b f(x)dx$ istnieje, to mówimy, że jest ona zbieżna, natomiast gdy istnieje całka niewłaściwa $\int_a^b |f(x)|dx$, to o całce $\int_a^b f(x)dx$ mówimy, że jest zbieżna bezwzględnie. Całka niewłaściwa, która jest zbieżna, ale nie jest zbieżna bezwzględnie, nazywa się całką warunkowo zbieżną.

Udowodnimy teraz podstawowe kryteria zbieżności całki niewłaściwej.

Twierdzenie 101. *(Kryterium Cauchy'ego) Niech b będzie jedynym punktem osobliwym funkcji f w przedziale $[a,b]$. Całka niewłaściwa $\int_a^b f(x)dx$ jest zbieżna wtedy i tylko wtedy, gdy dla każdego $\varepsilon > 0$ istnieje taka liczba $\beta_0 \in (a,b)$, że*

$$\left| \int_\beta^{\beta'} f(x)dx \right| < \varepsilon$$

dla dowolnych liczb β, β', spełniających nierówności $\beta_0 < \beta < \beta' < b$.

Dowód. Zbieżność całki niewłaściwej funkcji f oznacza, że $\int_a^b f(x)dx = \lim\limits_{\beta \to b} I(\beta)$, gdzie $I(\beta) = \int_a^\beta f(x)dx$. Na mocy Twierdzenia 36 jest to równoważne następującemu warunkowi:
Dla dowolnego $\varepsilon > 0$ istnieje $\beta_0 \in (a,b)$ takie, że $|I(\beta) - I(\beta')| = |\int_\beta^{\beta'} f(x)dx| < \varepsilon$ dla dowolnych β, β' spełniających nierówności $\beta_0 < \beta < \beta' < b$, co kończy dowód.

Odpowiednik Twierdzenia 36 jest prawdziwy również wtedy, gdy $p = +\infty$ lub $p = -\infty$, wtedy warunek Cauchy'ego przyjmuje postać:
dla dowolnego $\varepsilon > 0$ istnieje $M \in \mathbb{R}$ takie, że dla dowolnych $x', x'' \in E$, jeśli $x' > M$ oraz $x'' > m$, to $|f(x') - f(x'')| < \varepsilon$.

Twierdzenie 102. *(Kryterium Porównawcze) Jeśli funkcje f i F spełniają nierówność $|f(x)| \leqslant F(x)$ dla $a \leqslant x < b$ i b jest jedynym punktem osobliwym dla obu funkcji w przedziale $[a,b]$ oraz jeśli funkcja f jest R-całkowalna na każdym przedziale $[a,\beta], \beta < b$ i istnieje całka niewłaściwa $\int_a^b F(x)dx$, to istnieje również całka niewłaściwa $\int_a^b f(x)dx$ i jest ona bezwzględnie zbieżna.*

Dowód. Z Kryterium Cauchy'ego wynika, że dla dowolnego $\varepsilon > 0$ istnieje $\beta_0 \in (a,b)$ takie, że

$$\left| \int_\beta^{\beta'} F(x)dx \right| < \varepsilon \text{ dla } \beta_0 < \beta < \beta' < b.$$

Mamy więc

$$\left| \int_\beta^{\beta'} f(x)dx \right| \leqslant \int_\beta^{\beta'} \big|f(x)\big|dx \leqslant \int_\beta^{\beta'} F(x)dx = \left| \int_\beta^{\beta'} F(x)dx \right| < \varepsilon.$$

Z powyższych równości i z Kryterium Cauchy'ego wynika istnienie całek niewłaściwych $\int_a^b f(x)dx$, $\int_a^b |f(x)|dx$.

Wniosek 25.

(a) *Jeśli całka niewłaściwa* $\int_a^b |f(x)|dx$ *jest zbieżna, to zbieżna jest również całka niewłaściwa* $\int_a^b f(x)dx$ *(czyli bezwględna zbieżność pociąga zbieżność).*

(b) *Jeśli w przedziale* $[a, b)$ *funkcje* f *i* F *spełniają nierówność* $0 \leqslant f(x) \leqslant F(x)$ *i całka* $\int_a^b f(x)dx$ *nie istnieje, to całka niewłaściwa* $\int_a^b F(x)dx$ *jest rozbieżna.*

Twierdzenie 103. *(Kryterium Dirichleta) Niech punkt* b *będzie jedynym punktem osobliwym iloczynu funkcji* f *i* g *w przedziale* $[a, b]$. *Jeśli funkcja* $F(x) = \int_a^x f(t)dt$ *istnieje i jest ograniczona na przedziale* $[a, b)$, *a funkcja* $g(x)$ *dąży monotonicznie do zera, gdy* $x \to b$, *to całka niewłaściwa* $\int_a^b (fg)(x)dx$ *jest zbieżna.*

Dowód. Na mocy II Twierdzenia Całkowego o Wartości Średniej mamy

$$\int_\beta^{\beta'} (fg)(x)dx = g(\beta)\int_\beta^\xi f(x)dx + g(\beta')\int_\xi^{\beta'} f(x)dx,$$

dla dowolnych $a < \beta < \beta' < b$, gdzie ξ jest pewnym punktem leżącym między β i β'. Wobec przyjętych założeń dla dowolnego $\varepsilon > 0$ istnieje $\beta_0 \in (a, b)$ takie, że

$$\left| \int_\beta^{\beta'} (fg)(x)dx \right| < \varepsilon \text{ dla } \beta_0 < \beta < \beta' < b.$$

Na mocy kryterium Cauchy'ego całka niewłaściwa $\int_a^b (fg)(x)dx$ jest zbieżna.

Twierdzenie 104. *(Kryterium Abela) Niech* b *będzie jedynym punktem osobliwym dla iloczynu funkcji* f *i* g *w przedziale* $[a, b]$. *Jeśli całka niewłaściwa* $\int_a^b f(x)dx$ *jest zbieżna, a funkcja* g *jest monotoniczna i ograniczona, to całka niewłaściwa* $\int_a^b (fg)(x)dx$ *jest zbieżna.*

Dowód. Analogiczny do dowodu Kryterium Dirichleta.

Uwaga 49. Nietrudno zauważyć, że kryterium Abela wynika z kryterium Dirichleta. Niech funkcje f i g spełniają bowiem założenia kryterium Abela. Wówczas istnieje granica $\lim_{x \to b} g(x) = g(b)$. Mamy $f(x)g(x) = f(x)g(b) + f(x)[g(x) - g(b)]$. Jest jasne, że drugi składnik tej sumy spełnia założenia kryterium Dirichleta. Stąd całka niewłaściwa $\int_a^b (fg)(x)dx$ jest zbieżna.

Przykład 32.

(a) Rozważmy całkę niewłaściwą $\int_0^1 \frac{\cos x}{x} dx$. Ponieważ $\frac{\cos x}{x} \geqslant \frac{\cos 1}{x} \geqslant 0$ dla $x \in (0,1]$ oraz $\cos 1 \int_0^1 \frac{dx}{x}$ jest całką niewłaściwą rozbieżną, zatem na mocy Kryterium Porównawczego całka niewłaściwa $\int_0^1 \frac{\cos x}{x} dx$ jest rozbieżna.

(b) Rozważmy całkę niewłaściwą $\int_0^\infty \frac{\sin x}{x} dx$. Niech $g(x) = \frac{1}{x}$ oraz $f(x) = \sin x$ dla $x > 0$. Mamy $\int_0^\infty \frac{\sin x}{x} dx = \int_0^1 \frac{\sin x}{x} dx + \int_1^\infty \frac{\sin x}{x} dx$, zatem wobec Uwagi 48 (c) wystarczy zbadać zbieżność całki niewłaściwej $\int_1^\infty \frac{\sin x}{x} dx$. Ponieważ $\int_1^x \sin t dt = \cos 1 - \cos x$ jest funkcją ograniczoną na przedziale $[1, +\infty)$ oraz $\frac{1}{x} \to 0$ monotonicznie przy $x \to \infty$, zatem na mocy Kryterium Dirichleta całka niewłaściwa $\int_1^\infty \frac{\sin x}{x} dx$ jest zbieżna.
Pokażemy teraz, że całka niewłaściwa $\int_0^\infty \frac{\sin x}{x} dx$ nie jest zbieżna bezwzględnie. Mamy

$$\begin{aligned}
\lim_{N\to\infty} \int_0^{N\pi} \Big|\frac{\sin x}{x}\Big| dx &= \lim_{N\to\infty} \sum_{k=1}^N \int_{(k-1)\pi}^{k\pi} \frac{|\sin x|}{x} dx \\
&\geqslant \lim_{N\to\infty} \sum_{k=1}^N \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} |\sin x| dx \\
&= \lim_{N\to\infty} \sum_{k=1}^N \frac{1}{k\pi} \int_0^{\pi} \sin x dx \\
&= \frac{2}{\pi} \lim_{N\to\infty} \sum_{k=1}^N \frac{1}{k} = +\infty,
\end{aligned}$$

co kończy dowód.

8.7 Całka Riemanna-Stieltjesa (względem funkcji monotonicznej)

Całka Riemanna-Stieltjesa jest bezpośrednim uogólnieniem całki Riemanna.

Definicja 86. Niech $f, g : [a,b] \mapsto \mathbb{R}$ będą funkcjami ograniczonymi. Dla dowolnego podziału $P = \{x_0, \ldots, x_n\}$ przedziału $[a,b]$ tworzymy następujące sumy

$$S = \sum_{i=1}^n f(\xi_i)(g(x_i) - g(x_{i-1})), \text{ gdzie } \xi_i \in [x_{i-1}, x_i], i = 1, \ldots, n.$$

Sumy te nazywamy sumami Riemanna-Stieltjesa odpowiadającymi podziałowi P przy ustalonym wyborze punktów ξ_i. Przez $S(f,P)$ oznaczać będziemy zbiór wszystkich możliwych sum Riemanna-Stieltjesa odpowiadających podziałowi P.

Jeśli dla dowolnego ciągu normalnego podziałów (P_k) i dla dowolnych sum Riemanna-Stieltjesa $S_k \in S(f, P_k)$ istnieje skończona granica $I = \lim_{k\to\infty} S_k$, to tę granicę nazywamy całką Riemanna-Stieltjesa funkcji f względem funkcji g na przedziale $[a, b]$. O funkcji f mówimy wówczas, że jest całkowalna w sensie Riemanna-Stieltjesa (lub krótko: (R-S) całkowalna) względem funkcji g na przedziale $[a, b]$. Całkę Riemanna-Stieltjesa oznaczamy symbolem.

$$\int_a^b f dg \text{ lub } \int_a^b f(x)dg(x).$$

Definicja 87. Niech g będzie monotonicznie rosnącą funkcją określoną na przedziale $[a, b]$ (ponieważ $g(a)$ i $g(b)$ są skończone, więc funkcja jest ograniczona na $[a, b]$). Jeśli P jest jakimś podziałem przedziału $[a, b]$, to określamy $\Delta g_i = g(x_i) - g(x_{i-1})$ (oczywiście $\Delta g_i \geqslant 0$). Dla ograniczonej funkcji $f : [a, b] \mapsto \mathbb{R}$ piszemy

$$U(f, g, P) = \sum_{i=1}^{n} M_i \Delta g_i, \quad L(f, g, P) = \sum_{i=1}^{n} m_i \Delta g_i,$$

gdzie m_i oraz M_i mają ten sam sens, co w Definicji 71. Liczby $U(f, g, P)$ i $L(f, g, P)$ nazywać będziemy górną i dolną sumą Darboux-Stieltjesa. Dalej

$$\overline{\int_a^b} f dg = \inf_P U(f, g, P), \quad \underline{\int_a^b} f dg = \sup_P L(f, g, P),$$

gdzie kres górny i dolny są wzięte ze względu na wszystkie możliwe podziały przedziału $[a, b]$.

Rozumując analogicznie jak w dowodzie Twierdzeń 76 i 77 otrzymujemy dwa kryteria całkowalności funkcji w sensie Riemanna-Stieltjesa.

Twierdzenie 105. *Na to, by ograniczona funckja f była (R-S)-całkowalna względem funkcji rosnącej g na przedziale $[a, b]$ potrzeba i wysarcza, aby był spełniony jeden z następujących warunków:*

(a) dla dowolnego $\varepsilon > 0$ istnieje taki podział P przedziału $[a, b]$, że

$$U(f, g, P) - L(f, g, P) < \varepsilon.$$

(b)

$$\overline{\int_a^b} f dg = \underline{\int_a^b} f dg.$$

Zbadamy teraz klasy funkcji (R-S)-całkowalnych. Załóżmy wpierw, że g jest funkcją rosnącą na $[a, b]$. Wówczas rozumując podobnie jak w dowodzie Twierdzeń 78 oraz 80 wnioskujemy, że jeśli f jest funkcją ciągła lub monotoniczną (w tym przypadku zakładamy dodatkowo, że g jest funkcją ciągła na $[a, b]$), to jest ona (R-S)-całkowalna względem funkcji g na $[a, b]$.

Twierdzenie 106. *Niech $f : [a, b] \mapsto \mathbb{R}$ będzie funkcją ograniczoną i majacą tylko skończoną ilość punktów nieciągłości na przedziale $[a, b]$ i niech funkcja rosnąca g będzie ciągła w każdym z punktów, w których nieciągła jest funkcja f. Wtedy f jest (R-S)-całkowalna względem funkcji g.*

Dowód. Niech będzie dane $\varepsilon > 0$. Ponadto niech $M = \sup\limits_{x \in [a,b]} |f(x)|$ oraz niech E oznacza zbiór punktów nieciągłości f. Ponieważ E jest zbiorem skończonym i g jest ciągła w każdym z punktów E, więc możemy pokryć zbiór E skonczoną liczbą przedziałów rozłącznych $[u_j, v_j] \subset [a, b]$ tak, że suma różnic $g(v_j) - g(u_j) < \varepsilon$. Możemy poza tym tak umieścić te przedziały, aby każdy z punktów zbioru $E \cap (a, b)$ leżał we wnętrzu któregoś z przedziałów $[u_j, v_j]$. Usuńmy przedziały (u_j, v_j) z odcinka $[a, b]$. Pozostały zbiór K jest skończoną sumą przedziałów domkniętych. Wobec tego na mocy Twierdzenia Cantora funkcja f jest jednostajnie ciągła na K, a więc istnieje $\delta > 0$ takie, że $|f(s)-f(t)| < \varepsilon$, jeśli tylko $s, t \in K$ oraz $|s - t| < \delta$.
Utwórzmy teraz podział $P = \{x_0, \ldots, x_n\}$ przedziału $[a, b]$ tak, aby każdy z punktów u_j, v_j występował w P oraz aby żaden z punktów przedziału (u_j, v_j) nie należał do P. Dalej jeśli x_{i-1} nie jest żadnym z punktów u_j, to ma być $\Delta x_i < \delta$. Zauważmy, że $M_i - m_i \leqslant 2M$ dla dowolnego i oraz, że $M_i - m_i \leqslant \varepsilon$ o ile x_{i-1} nie jest żadnym z punktów u_j. Wobec tego mamy

$$U(f, g, P) - L(f, g, P) \sum_{i=1}^{n} (M_i - m_i) \Delta g_i \leqslant [g(b) - g(a)]\varepsilon + 2M\varepsilon.$$

Ponieważ $\varepsilon > 0$ było dowolne, zatem na mocy Twierdzenia 105 (a) dowód jest skończony.

Uwaga 50. Jeśli funkcje f i g posiadają w przedziale $[a, b]$ wspólny punkt nieciągłości to może się zdarzyć, że f nie jest (R-S)-całkowalna względem funkcji g (zob. Przykład 33).

Powtarzając rozumowanie z dowodu Twierdzenia 81 otrzymujemy następujące.

Twierdzenie 107. *Niech f będzie funkcją (R-S)-całkowalną względem funkcji rosnącej g na przedziale $[a, b]$, $m \leqslant f(x) \leqslant M$ i niech funkcja ϕ będzie ciągła na przedziale $[m, M]$. Wówczas funkcja $h(x) : [a, b] \mapsto \mathbb{R}$ określona wzorem $h(x) = \phi(f(x))$ dla $x \in [a, b]$ jest funkcją (R-S)-całkowalną względem funkcji g.*

Twierdzenie 108. *Jeśli* funkcja $f : [a, b] \mapsto \mathbb{R}$ jest R-całkowalna, a funkcja $g : [a, b] \mapsto \mathbb{R}$ spełnia warunek Lipschitza ze stałą $L > 0$, to funkcja f jest (R-S)-całkowalna względem funkcji g.

Dowód. Załóżmy wpierw dodatkowo, że funkcja g jest rosnąca. Wówczas $\Delta g_i \leqslant \Delta x_i$, skąd
$$U(f,g,P) - L(f,g,P) \leqslant L(U(f,P) - L(f,P)).$$
Ponieważ funkcja f jest R-całkowalna, zatem wobec Twierdzenia 105 (a) funkcja f jest (R-S)-całkowalna względem funkcji g.
Niech teraz g będzie dowolną funkcją spełniającą warunek Lipschitza ze stałą $L > 0$. Przedstawmy funkcję g w postaci

$$g(x) = Lx - (Lx - g(x)) = g_1(x) - g_2(x).$$

Funkcja $g_1(x) = Lx$, $x \in [a,b]$ spełnia oczywiście Warunek Lipschitza, a jednocześnie jest rosnąca. Te same własności ma funkcja $g_2(x) = Lx - g(x)$, $x \in [a,b]$, bowiem dla $a \leqslant x < x' \leqslant b$ mamy

$$|g_2(x') - g_2(x)| = L(x' - x) + |g(x') - g(x)| \leqslant 2L(x' - x).$$

Stąd
$$U(f,g_2,P) - L(f,g_2,P) \leqslant 2L(U(f,P) - L(f,P))$$
i dalej rozumujemy jak wyżej.

Można również udowodnić (zob. [2], t. III, s. 74—75) następujące

Twierdzenie 109. *Jeśli funkcja $f : [a,b] \mapsto \mathbb{R}$ jest R-całkowalna, a funkcję $g : [a,b] \mapsto \mathbb{R}$ można przedstawić w następującej postaci:*

$$g(x) = c + \int_a^x \phi(t)dt,$$

gdzie c jest stałą, a ϕ — funkcją bezwzględnie całkowalną (w sensie Riemanna lub w sensie niewłaściwym) w przedziale $[a,b]$ (lub $[a,b)$), to funkcja f jest (R-S)-całkowalna względem funkcji g.

Twierdzenie 110. *Niech $f, g, g_1, g_2, f_1, f_2 : [a,b] \mapsto \mathbb{R}$ będą funkcjami ograniczonymi. Wówczas*

(a) $\int_a^b dg = g(b) - g(a)$;

(b) $\int_a^b [f_1 \pm f_2] dg = \int_a^b f_1 dg \pm \int_a^b f_2 dg$;

(c) $\int_a^b f d[g_1 \pm g_2] = \int_a^b f dg_1 \pm \int_a^b f dg_2$;

(d) $\int_a^b k f d[lg] = kl \int_a^b f dg$ *$(k, l = const.)$;*

(w punktach (b), (c), (d) zakładamy istnienie całek Riemanna-Stieltjesa po prawych stronach nierówności)

(e) *$\int_a^b f dg = \int_a^c f dg + \int_c^b f dg$, przy założenium że $a < c < b$ oraz, że istnieje całka po lewej stronie nierówności (zob. [2], t. III, s.75—76).*

Przykład 33. Z istnienia całek $\int_a^c f dg$ i $\int_c^b f dg$ nie wynika na ogół istnienie całki $\int_a^b f dg$. Niech w przedziale $[-1, 1]$ funkcje g i g będą określone następująco:

$$f(x) = \begin{cases} 0 \text{ dla } & -1 \leqslant x \leqslant, 0 \\ 1 \text{ dla } & 0 < x \leqslant 1, \end{cases} \qquad g(x) = \begin{cases} 0 \text{ dla } & -1 \leqslant x < 0, \\ 1 \text{ dla } & 0 \leqslant x \leqslant 1, \end{cases}$$

Można łatwo sprawdzić, że obie całki $\int_{-1}^0 f dg$, $\int_0^1 f dg$ istnieją i są równe zeru, bowiem odpowiadające im sumy Riemanna-Stieltjesa są równe zeru, bowiem $f(x) = 0$ dla $x \in [-1, 0]$; druga, bowiem dla $x \in [0, 1]$ funkcja g jest stała, czyli $\Delta g_i = 0$. Natomiast $\int_{-1}^1 f dg$ nie istnieje. Istotnie, dokonajmy podziału przedziału $[-1, 1]$ na podprzedziały tak, żeby punkt 0 nie był punktem podziału i utwórzmy sumę

$$S = \sum_{i=1}^{n} f(\xi_i)\Delta g_i.$$

Jeśli punkt 0 należy do przedziału $[x_{k-1}, x_k]$, to przy $x_{k-1} < 0 < x_k$ w sumie S jest różny tylko k-ty składnik bowiem $\Delta g_i = 0$ dla $i \neq k$. W takim razie $S = f(\xi_k)(g(x_k) - g(x_{k-1})) = f(\xi_k)$. W zależności od tego czy $\xi_k \leqslant 0$ czy $\xi_k > 0$ mamy $S = 0$ lub $S = 1$, zatem $\int_{-1}^1 f dg$ nie istnieje.

Twierdzenie 111. *Niech $g : [a, b] \mapsto \mathbb{R}$ będzie funkcją rosnącą. Wówczas*

(a) *Jeżeli funkcje $f_1, f_2 : [a, b] \mapsto \mathbb{R}$ są (R-S)-całkowalne względem funckji g oraz $f_1(x) < f_2(x)$ dla każdego $x \in [a, b]$, to*

$$\int_a^b f_1 dg \leqslant \int_a^b f_2 dg.$$

(b) *Jeśli $f : [a, b] \mapsto \mathbb{R}$ jest funkcją (R-S)-całkowalną względem funkcji g, to $|f|$ jest funkcją (R-S)-całkowalna względem funkcji g oraz*

$$\left| \int_a^b f dg \right| \leqslant \int_a^b |f| dg.$$

Ponadto jeśli $|f(x)| \leqslant M$ na $[a, b]$, to $\left| \int_a^b f dg \right| \leqslant M(g(b) - g(a))$.

(c) *Jeśli $f, h : [a, b] \mapsto \mathbb{R}$ są funkcjami (R-S)-całkowalnymi względem g, to iloczyn fh jest funkcją (R-S)-całkowalną względem funkcji g.*

Dowód powyższego twierdzenia jest analogiczny do dowodów twierdzeń 82 (dla iloczynu funkcji) oraz 83.

Udowodnimy teraz twierdzenia o zamianie zmiennych oraz o całkowaniu przez części dla całek Riemanna-Stieltjesa.

Twierdzenie 112. *Niech ϕ będzie funkcją ściśle rosnącą odwzorowującą przedział $[\alpha, \beta]$ na przedział $[a, b]$. Niech g będzie również funkcją rosnącą na $[a, b]$ i niech f będzie funkcją (R-S)-całkowalną względem funkcji g na $[a, b]$. Określmy na przedziale $[\alpha, \beta]$ funckje F i G wzorami*

$$G(y) = g(\phi(y)), \quad F(y) = f(\phi(y)),$$

Wówczas F jest funkcją (R-S)-całkowalną względem funkcji G na $[\alpha, \beta]$ oraz

$$\int_\alpha^\beta F dG = \int_a^b f dg.$$

Dowód. Każdemu podziałowi $P = \{x_0, \ldots, x_n\}$ przedziału $[a, b]$ odpowiada podział $Q = \{y_0, \ldots, y_n\}$ przedziału $[\alpha, \beta]$ taki, że $y_i = \phi(x_i)$, $i = 0, \ldots, n$. Dowolny podział przedziału $[\alpha, \beta]$ możemy otrzymać w ten sposób. Ponieważ wartości przyjmowane przez f na przedziale $[x_{i-1}, x_i]$ są dokładnie takie same, jak wartości przyjmowane przez F na $[y_{i-1}, y_i]$, zatem

$$U(f, g, P) = U(F, G, Q), \quad L(f, G, P) = L(F, G, Q). \tag{61}$$

Ponieważ funkcja f jest (R-S)-całkowalna względem g, więc możemy tak wybrać P, aby zarówno $U(f, g, P)$ i $L(f, g, P)$ były bliskie $\int_a^b f dg$. Wówczas (61) w połączeniu z Twierdzeniem 105 (a) pokazuje, że funkcja F jest (R-S)-całkowalna względem funkcji G i że zachodzi wzór z tezy twierdzenia.

Twierdzenie 113. *Niech $f, g : [a, b] \mapsto \mathbb{R}$ będą funkcjami ograniczonymi. Wówczas*

$$\int_a^b f dg = (fg)(x)\Big|_a^b - \int_a^b g df, \tag{62}$$

przy założeniu, że przynajmniej jedna z tych całek istnieje.

Dowód. Niech istnieje całka $\int_a^b g df$ i niech $P = \{x_0, \ldots, x_n\}$ będzie podziałem przedziału $[a, b]$. Sumę Riemanna-Stieltjesa odpowiadającą podziałowi P przy ustalonym wyborze punktów ξ_i możemy przedstawić w postaci

$$S = \sum_{i=1}^n f(x_i)(g(x_i) - g(x_{i-1})) = f(b)g(b) - f(a)g(a) -$$
$$\left\{ g(a)(f(\xi_1) - f(a)) + \sum_{i=2}^n g(x_{i-1})[f(\xi_i) - f(\xi_{i-1})] + g(b)(f(b) - f(\xi_n)) \right\}.$$

Wyrażenie w nawiasach sześciennych przedstawia pewną sumę Riemanna-Stieltjesa dla całki $\int_a^b g df$, której istnienie zakładamy. Odpowiada ona podziałowi $\tilde{P} =$

$\{a, \xi_1, \ldots, \xi_n, b\}$ przedziału $[a, b]$. Ponadto $\delta(\tilde{P}) \leqslant 2\delta(P)$. Jeśli $\delta(P) \to 0$ to $\delta(\tilde{P}) \to 0$, a zatem istnieje również granica dla sum S, to jest $\int_a^b f dg$ i całka ta dana jest wzorem (62).

Definicja 88. Jednostkową funkcją schodkową naywamy funkcję

$$f(x) = \begin{cases} 0 \text{ dla } & x \leqslant 0, \\ 1 \text{ dla } & x > 0. \end{cases}$$

Twierdzenie 114. *Jeśli $a \leqslant x \leqslant b$, f jest funkcją ograniczoną na $[a, b]$ oraz ciągłą w punkcie s, a $g(x) = I(x - s)$, to*

$$\int_a^b f dg = f(s).$$

Dowód. Niech $P = \{x_0, \ldots, x_n\}$ będzie podziałem przedziału $[a, b]$. Niech s należy na przykład do k-tego podprzedziału; jest więc $x_{k-1} \leqslant s \leqslant x_k$. Wówczas $\Delta I_k = 1$, a przy $i \neq k$ mamy $\Delta I_i = 0$. Suma Riemanna-Stieltjesa S sprowadza się więc do jednego składnika: $S = f(\xi_k)$. Niech teraz $\delta(P) \to 0$. Na mocy ciągłości funkcji f w punkcie s, $f(\xi_k) \to f(s)$, czyli

$$\int_a^b f dg = f(s).$$

Przykład 34.

$$\int_1^\pi \ln x d[x] = (\ln 2) \cdot 1 + (\ln 3) \cdot 1 = \ln 6.$$

Twierdzenie 115. *Jeśli $f : [a.b] \mapsto \mathbb{R}$ jest funkcja R-całkowalną oraz pochodna funkcji $g : [a, b] \mapsto \mathbb{R}$ jest całkowalna w sensie Riemanna, to*

$$\int_a^b f dg = \int_a^b f g' dx. \tag{63}$$

Dowód. Na mocy twierdzeń 82 i 108 obydwie całki występujące we wzorze (63) istnieją. Niech teraz $P = \{x_0, \ldots, x_n\}$ będzie podziałem przedziału $[a, b]$. Na mocy Twierdzenia Lagrange'a o Wartośći Średniej sumę Riemanna-Stieltjesa odpowiadającą podziałowi P możemy zapisać w postaci

$$S = \sum_{i=1}^n f(\xi_i)(g(x_i) - g(x_{i-1})) = \sum_{i=1}^n f(\xi_i) g'(\mu_i)(x_i - x_{i-1}),$$

gdzie $x_{i-1} < \mu_i < x_i$, $i = 1, \ldots, n$. Podstawiając w powyższej równości $\xi_i = \mu_i$ oraz zakładając, że $\delta(P) \to 0$ otrzymujemy wzór (63).

Wniosek 26. *Niech $g(x) = x$ dla $x \in [a,b]$ i niech ϕ będzie ściśle rosnącą funkcją odwzorowującą przedział $[\alpha, \beta]$ na przedział $[a,b]$ i taką, że funkcja ϕ' jest funkcją R-całkowalną na $[\alpha, \beta]$. Wobec twierdzeń 112 i 115 otrzymujemy*

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\phi(y))\phi'(y)dy;$$

powyższy wzór jest treścią twierdzenia 91, jeśli ϕ jest funkcją rosnącą. Dobierając $\phi(y) = \phi(-y + \alpha + \beta)$ możemy uzyskać tezę twierdzenia dla funkcji malejącej.

Podamy jeszcze trzy twierdzenia dotyczące całek Riemanna-Stieltjesa. Dowody tych twierdzeń można znaleźć w [2], t. III, s. 79—82.

Twierdzenie 116. *Jeśli $f : [a,b] \mapsto \mathbb{R}$ jest R-całkowalna, a funkcję $g : [a,b] \mapsto \mathbb{R}$ można przedstawić w postaci*

$$g(x) = c + \int_a^x \phi(t)dt, \quad c = const, \quad x \in [a,b],$$

gdzie funkcja ϕ jest bezwzględnie całkowalna w $[a,b]$, to

$$\int_a^b f dg = \int_a^b f\phi dx$$

Twierdzenie 117. *Jeśli $f : [a,b] \mapsto \mathbb{R}$ jest R-całkowalna, a funkcja $g : [a,b] \mapsto \mathbb{R}$ jest ciągła w całym przedziale $[a,b]$ i ma w nim, poza conajmniej skończoną liczbą punktów pochodną g' bewględnie całkowalną w $[a,b]$, to*

$$\int_a^b f dg = \int_a^b f g' dx.$$

Twierdzenie 118. *Niech $f : [a,b] \mapsto \mathbb{R}$ będzie funkcją ciągła,a funkcja $g : [a,b] \mapsto \mathbb{R}$ niech ma w przedziale $[a,b]$ z pominięciem co najwyżej skończonej liczby punktów pochodną g', bezwzględnie całkowalną w tym przedziale. Niech ponadto funkcja g ma w skończonej liczbie punktów*

$$a = c_0 < c_1 < \cdots < c_n = b$$

nieciągłość pierwszego rodzaju. Wówczas istnieje całka Riemanna-Stieltjesa funkcji f względem funkcji g i wyraża się wzorem

$$\begin{aligned}\int_a^b f dg = & \int_a^b f g' dx + f(a)[g(a+0) - g(a)] \\ & + \sum_{k=1}^{n-1} f(c_k)(g(c_k+0) - g(c_k-0)) + f(b)[g(b) - g(b-0)].\end{aligned}$$

8.8 Całka krzywoliniowa skierowana

Definicja 89. Niech będzie dana krzywa $\gamma : [a, b] \mapsto \mathbb{C}$ $(\gamma(t) = (\phi(t), \psi(t)))$, $t \in [a, b]$ i niech będą dane wzdłuż obrazu tej krzywej pewne funkcje $f(x, y)$, $g(x, y)$ $(f, g : \gamma([a, b]) \mapsto \mathbb{R})$. Niech $P = \{t_0, \ldots, t_n\}$ będzie podziałem przedziału $[a, b]$. Tworzymy sumy

$$\sigma = \sum_{i=1}^{n} [f(\phi(\xi_i), \psi(\xi_i))(\phi(t_i) - \phi(t_{i-1})) + g(\phi(\xi_i), \psi(\xi_i))(\psi(t_i) - \psi(t_{i-1}))],$$

gdzie $\xi_i \in [t_{i-1}, t_i]$, $i = 1, \ldots, n$.
Jeśli dla dowolnego ciągu normalnego podziałów (P_k) i dla dowolnych sum σ_k odpowiadających podziałowi P_k istnieje skończona granica $I = \lim_{k\to\infty} \sigma_k$, to tę granicę nazywamy całką krzywoliniową skierowaną $fdx + gdy$ po drodze γ i oznaczamy ją symbolem

$$\int_\gamma fdx + gdy.$$

Bezpośrednio z definicji wynika następujące

Twierdzenie 119.

(a) *Niech* $-\gamma(t) = (\phi(-t), \psi(-t))$, $t \in [-b, -a]$. *Wówczas, jeśli istnieje całka krzywoliniowa skierowana dla* $fdx + gdy$ *po krzywej* γ, *to istnieje też całka po krzywej* $-\gamma$ *i oraz*

$$\int_{-\gamma} fdx + gdy = -\int_\gamma fdx + gdy.$$

(b) *Niech* $\gamma = \gamma_1 + \gamma_2$ *(to znaczy istnieje punkt* $c \in [a, b]$ *taki, że* $\gamma(t) = \gamma_1(t)$ *dla* $t \in [a, c]$ *oraz* $\gamma_2(t) = \gamma(t)$ *dla* $t \in [c, b]$*). Wówczas całka dla* $fdx + gdy$ *po krzywej* γ *istnieje wtedy i tylko wtedy, gdy istnieją całki dla* $fdx + gdy$ *po krzywych* γ_1, γ_2 *przy czym*

$$\int_\gamma fdx + gdy = \int_{\gamma_1} fdx + gdy + \int_{\gamma_2} fdx + gdy$$

Udowodnimy teraz twierdzenie o zamianie całki krzywoliniowej skierowanej na całkę Riemanna.

Twierdzenie 120. *Jeżeli funkcje* f, g *są ciągłe na obrazie krzywej gładkiej* $\gamma : [a, b] \mapsto \mathbb{C}$ *to* $\int_\gamma fdx + gdy$ *istnieje oraz zachodzi wzór*

$$\int_\gamma fdx + gdy = \int_a^b [f(\phi(t), \psi(t))\phi'(t) + g(\phi(t), \psi(t))\psi'(t)]dt. \tag{64}$$

Dowód. Sumę σ występującą w Definicji 89 możemy zapisać w postaci

$$\sigma = \sum_{i=1}^{n} [f(\phi(\xi_i), \psi(\xi_i))\phi'(\nu_i) + g(\phi(\xi_i), \psi(\xi_i))\psi'(\mu_i)](t_i - t_{i-1}),$$

gdzie $\nu_i, \mu_i \in [t_{i-1}, t_i]$, bowiem do różnic $\phi(t_i) - \phi(t_{i-1})$ oraz $\psi(t_i) - \psi(t_{i-1})$ można zastosować Twierdzenie Lagrange'a o Wartości Średniej. Całka po prawej stronie równości (64) istnieje, bowiem funkcja podcałkowa jest ciągła oraz jest granicą ciagu sum postaci

$$\tilde{\sigma} = \sum_{i=1}^{n} [f(\phi(\xi_i), \psi(\xi_i))\phi'(\xi_i) + g(\phi(\xi_i), \psi(\xi_i))\psi'(\xi_i)](t_i - t_{i-1})$$

(dla dowolnego ciągu normalnego podziałów (P_k)). Wystarczy zatem wykazać, że sumy σ i $\tilde{\sigma}$ różnią się o dowolnie mało, jeśli średnica podziału jest dostatecznie mała, lub, że różnice $|\phi'(\nu_i) - \phi'(\xi_i)|$, $|\psi'(\mu_i) - \psi'(\xi_i)|$ są dowolne małe. To zaś wynika z jednostajnej ciągłości pochodnych ϕ', ψ' w przedziale $[a, b]$.

Przykład 35. Korzystając ze wzoru 64 obliczmy wartość całki krzywoliniowej $H = \int_L 2xydx + x^2dy$ wziętej po drodze L łączącej punkty $O(0,0)$ i $A(1,1)$ jeśli droga L jest:

(a) prostą $y = x$ ($\gamma(x) = (x, x)$, $x \in [0, 1]$)

$$\int_L 2xydx + x^2dy = \int_0^1 3x^2dx = 1,$$

(b) parabolą $y = x^2$ ($\sigma(x) = (x, x^2)$, $x \in [0, 1]$)

$$\int_L 2xydx + x^2dy = \int_0^1 4x^3dx = 1.$$

Twierdzenie 131. *(Całkowe Kryterium Zbieżności Szeregów). Jeśli f jest funkcją dodatnią i nierosnącą na przedziale $[1, +\infty)$ oraz $f(n) = a_n$ dla $n \in \mathbb{N}$, to:*

(a) szereg $\sum_{n=1}^{\infty} a_n$ jest zbieżny wtedy i tylko wtedy, gdy zbieżna jest całka niewłaściwa $I = \int_1^{+\infty} f(x)dx$;

(b) ciąg $(s_n - I_n)$, gdzie $s_n = \sum_{k=1}^{n} a_k$ i $I_n = \int_1^n f(x)dx$ jest zbieżny i jego granica należy do przedziału $[0, a_1]$.

Dowód. Dla $k \leqslant x \leqslant k+1$ mamy $a_{k+1} \leqslant f(x) \leqslant a_k$, a więc $a_{k+1} \leqslant \int_k^{k+1} f(x)dx \leqslant a_k$, skąd

$$a_2 + \cdots + a_n \leqslant \int_1^n f(x)dx \leqslant a_1 + \cdots + a_{n-1},$$

czyli

$$s_n - a_1 \leqslant I_n \leqslant s_{n-1}.$$

Jeśli całka I istnieje, to ciąg (I_n) jest zbieżny, więc i ograniczony. Z powyższej nierówności wynika, że ograniczony jest również ciąg (s_n). Stąd na mocy Twierdzenia 127 wnioskujemy, że szereg $\sum_{n=1}^{\infty} a_n$ jest zbieżny.
Załóżmy teraz, że całka I jest rozbieżna. Ponieważ ciąg (I_n) jest niemalejący, więc $\lim_{n\to\infty} I_n = +\infty$. Wówczas również $\lim_{n\to\infty} s_n = +\infty$, czyli $\sum_{n=1}^{\infty} a_n$ jest rozbieżny.
Ciąg $(s_n - I_n)$ jest nierosnący, bowiem

$$(s_{n-1} - I_{n-1}) - (s_n - I_n) = (I_n - I_{n-1}) - (s_n - s_{n-1}) = \int_{n-1}^{n} f(x)dx - a_n \geqslant 0.$$

Ponadto jest on ograniczony, bowiem

$$0 < a_n \leqslant s_n - I_n \leqslant a_1.$$

Wobec tego jest on ciągiem zbieżnym $\lim_{n\to\infty} (s_n - I_n) \in [0, a_1]$.

10 Ciągi i szeregi funkcyjne

10.1 Zbieżność punktowa i jednostajna

Definicja 98. Niech (f_n), $n \in \mathbb{N}$, będzie ciągiem funkcji o wartościach rzeczywistych lub zespolonych, określonych na zbiorze $E \subset \mathbb{R}$ (lub $E \subset \mathbb{C}$). Jeśli ciąg liczb $(f_n(x))$ jest zbieżny dla każdego $x \in E$, to mówimy że ciąg (f_n) jest zbieżny lub zbieżny punktowo na zbiorze E do funkcji f, gdzie

$$f(x) = \lim_{n\to\infty} f_n(x) \quad (x \in E).$$

Analogicznie, jeśli ciąg (s_n) sum częsciowych szeregu $\sum\limits_{n=1}^{\infty} f_n$ ($s_n = \sum\limits_{k=1}^{n} f_k$, $k \in \mathbb{N}$) jest punktowo zbieżny na zbiorze E do funkcji s ($s(x) = \lim\limits_{n\to\infty} s_n(x)$, $x \in E$), to mówimy, że szereg $\sum\limits_{n=1}^{\infty} f_n$ jest punktowo zbieżny na zbiorze E do funkcji s. Funkcję s nazywamy wówczas sumą szeregu $\sum\limits_{n=1}^{\infty} f_n$ i piszemy

$$s(x) = \sum_{n=1}^{\infty} f_n(x) \quad (x \in E).$$

Definicja 99. Mówimy, że ciąg funkcji (f_n), $n \in \mathbb{N}$, jest zbieżny jednostajnie na zbiorze E do funkcji f, jeżeli dla dowolnego $\varepsilon > 0$ istnieje takie $N \in \mathbb{N}$, że dla $n \geqslant N$ zachodzi

$$|f_n(x) - f(x)| < \varepsilon, \quad \text{dla każdego } x \in E. \tag{73}$$

Analogicznie, jeśli ciąg (s_n) sum częsciowych szeregu $\sum\limits_{n=1}^{\infty}$ jest zbieżny jednostajnie na zbiorze E do funkcji s to mówimy, że szereg $\sum\limits_{n=1}^{\infty}$ jest zbieżny jednostajnie na zbiorze E do funkcji s.

Uwaga 57. Jest oczywiste, że każdy ciąg zbieżny jednostajnie jest także zbieżny punktowo; odwortnie zachodzić nie musi (zob. Przykład 45 (b)).
Róznica między zbieżnością punktową, a jednostajną polega na tym, że w pierwszym przypadku dla dowolnego $\varepsilon > 0$ i dowolnego ustalonego $x \in E$ można dobrać takie N (które zależy i od ε i od x), że dla $n \geqslant N$ będzie spełniona nierówność (73); jeżeli ciąg (f_n) jest zbieżny jednostajnie, to przy każdym $\varepsilon > 0$ można dobrać jedną wspólną dla wszystkich punktów $x \in E$ liczbę N.

Przykład 41.

(a) Niech $f_n(x) = \frac{x}{1+n^2x^2}$, $x \in \mathbb{R}$.
Wówczas $\lim\limits_{n\to\infty} f_n(x) = 0$ dla każdego $x \in \mathbb{R}$.
Dalej, ponieważ

$$(1-nx)^2 \geqslant 0 \quad \text{oraz} \quad (1+nx)^2 \geqslant 0$$

zatem

$$\frac{2nx}{1+n^2x^2} \leqslant 1 \quad \text{oraz} \quad \frac{2nx}{1+n^2x^2} \geqslant -1.$$

Stąd

$$-\frac{1}{2n} \leqslant f_n(x) = \frac{1}{2n}\frac{2nx}{1+n^2x^2} \leqslant \frac{1}{2n},$$

zatem, aby nierówność $|f_n(x)| \leqslant \varepsilon$ była spełniona dla wszystkich x wystarczy przyjąć $n > \frac{1}{2\varepsilon}$. Tak, więc liczba $N = [\frac{1}{2\varepsilon}] + 1$ nadaje się dla wszystkich x jednocześnie.

(b) Niech $f_n(x) = n^2(1-x^2)^n$, $x \in [0,1]$, $n \in \mathbb{N}$.
Jeśli $0 < x \leqslant 1$, to $\lim\limits_{n\to\infty} f_n(x) = 0$, na mocy Twierdzenia 34 (c).
Ponieważ $f_n(0) = 0$ dla dowolnego $n \in \mathbb{N}$, zatem $\lim\limits_{n\to\infty} f_n(x) = 0$ dla $x \in [0,1]$.
Ponieważ $f_n(\frac{1}{n}) = n^2\frac{1}{n}(1-(\frac{1}{n})^2)^n = n\frac{(n+1)^n}{n^n}\frac{(n-1)^n}{n^n}$, a więc nie możemy osiągnąc tego, by $|f_n(x)| < 1$ dla każdego $x \in [0,1]$ i prawie wszystkich n.

Następujące twierdzenie jest przeniesiem kryterium Cauchy'ego na przypadek zbieżnośći jednostajnej.

Twierdzenie 144. *Ciąg funkcji (f_n) określonych na zbiorze E, jest na tym zbiorze jednostajnie zbieżny wtedy i tylko wtedy, gdy dla dowolnego $\varepsilon > 0$ istnieje liczba naturalna N taka, że przy $m, n \geqslant N$, $x \in E$ mamy*

$$|f_n(x) - f_m(x)| < \varepsilon. \tag{74}$$

Dowód. Niech ciąg (f_n) będzie zbieżny jednostajnie na zbiorze E do funkcji f. Istnieje wtedy liczba naturalna N, taka, że jeżeli $n \geqslant N$ i $x \in E$, to zachodzi nierówność

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2}.$$

Wtedy

$$|f_n(x) - f_m(x)| \leqslant |f_n(x) - f(x)| + |f(x) - f_m(x)| < \varepsilon$$

jeżeli $n, m \geqslant N$, $x \in E$.
Na odwrót, niech będzie spełniony warunek Cauchy'ego. Na mocy Twierdzenia

33 ciąg $(f_n(x))$ jest przy każdym x zbieżny do granicy, którą oznaczmy przez $f(x)$. Pokażemy, że w tym przypadku zbieżność jest jednostajna.
Niech będzie dana liczba $\varepsilon > 0$. Niech $N \in \mathbb{N}$ będzie takie, aby spełniona była nierówność (74). Ustalmy n i przejdźmy w nierówności (74) z m do granicy $(m \to \infty)$. Ponieważ $f_m(x) \to f(x)$ przy $m \to \infty$, zatem

$$|f_n(x) - f(x)| < \varepsilon$$

dla dowolnego $n > N$ i dowolnego $x \in E$. Dowód jest zakończony.

Następujące kryterium jest bezpośrednią konsekwencją Definicji 99.

Twierdzenie 145. *Niech* $\lim\limits_{n\to\infty} f_n(x) = f(x)$ *dla* $x \in E$. *Określmy* $M_n = \sup\limits_{x \in E} |f_n(x) - f(x)|$. *Wówczas* $f_n \to f(x)$ *jednostajnie na* E *wtedy i tylko wtedy, gdy* $M_n \to 0$ *przy* $n \to \infty$.

Podamy teraz wygodnie kryterium zbieżności jednostajnej szeregów funkcyjnych pochodzące od Weierstrassa.

Twierdzenie 146. *(Twierdzenie o Majorancie). Niech* (f_n) *będzie ciągiem funkcyjnym określonym na zbiorze* E *i niech*

$$|f_n(x)| \leqslant M_n \quad (x \in E, n \in \mathbb{N}).$$

Wtedy szereg $\sum\limits_{n=1}^{\infty} f_n$ *jest zbieżny jednostajnie na* E*, jeśli zbieżny jest szereg liczbowy* $\sum\limits_{n=1}^{\infty} M_n$.

Dowód. Jeżeli $\sum\limits_{n=1}^{\infty} M_n$ jest zbieżny to przy dowolnym $\varepsilon > 0$

$$\left| \sum_{i=n}^{m} f_i(x) \right| \leqslant \sum_{i=n}^{m} M_i < \varepsilon \quad (x \in E),$$

jeżeli tylko m i n są dostatecznie duże (zob. Tw. 121). Wobec Twierdzenia 144 wnosimy, więc, że szereg funkcyjny $\sum\limits_{n=1}^{\infty} f_n$ jest zbieżny jednostajnie.

Uwaga 58. Jest jasne, że jeżeli do danego szeregu funkcyjnego można zastosować Kryterium Weierstrassa, to szereg ten jest zbieżny bezwzględnie. Możliwe są jednak przypadki gdy dany szereg jest zbieżny jednostajnie, nie będąc zbieżnym bezwzględnie. (np. $\sum\limits_{n=1}^{\infty} \frac{(-1)^{n-1}}{x^2+n}$ ([2], t. II, przykład 7/367)). Możliwe są również przypadki gdy szereg funkcyjny $\sum\limits_{n=1}^{\infty} f_n$ jest zbieżny bezwzględnie i jednostajnie, a szereg $\sum\limits_{n=1}^{\infty} |f_n|$ jest zbieżny niejednostajnie (np. $\sum\limits_{n=1}^{\infty} \frac{x^2(-1)^{n-1}}{(1+x^2)^n}$ ([2], t. II. przykład. 8 s. 367)).

Rozumując analogicznie jak w dowodzie Twierdzenia 134 otrzymujemy nastepujące kryterium zbieżności jednostajnej szeregów funkcyjnych zwane **Kryterium Dirichleta**.

Twierdzenie 147. *Niech sumy częściowe s_n szeregu funkcyjnego $\sum_{n=1}^{\infty} f_n$ będą wspólnie ograniczone dla dowolnych x i n: $|s_n(x)| \leqslant M$, a funkcje $g_n(x)$ (dla każdego x) tworzą ciąg monotoniczny zbieżny do 0 jednostajnie na zbiorze E. Wówczas szereg funkcyjny $\sum_{n=1}^{\infty} f_n g_n$ jest także zbieżny jednostajnie w tym zbiorze.*

Twierdzenie 148. *(Kryterium Abela). Niech szereg funkcyjny $\sum_{n=1}^{\infty} f_n$ będzie zbieżny jednostajnie na zbiorze E i niech funkcje $g_n(x)$ tworzą (dla każdego x) ciąg monotoniczny i są wspólnie ograniczone dla dowolnych x i n: $|g_n(x)| \leqslant K$. Wówczas szereg funkcyjny $\sum_{n=1}^{\infty} f_n g_n$ jest zbieżny jednostajnie w zbiorze E.*

Dowód. Z uwagi na jednostajną zbieżność szeregu funkcyjnego $\sum_{n=1}^{\infty} f_n$ znajdziemy wskaźnik N, niezależny od x taki, że dla $N \leqslant p \leqslant q$ mamy

$$\left| \sum_{n=p}^{q} f_n(x) \right| < \frac{\varepsilon}{3^k}.$$

Wówczas dla $N \leqslant p \leqslant q$, stosując przekształcenie Abela (Lemat 3, w miejsce A_n wstawiamy $s_{p,n} = f_p + f_{p+1} + \cdots + f_n$; $s_{p,p-1} = 0$) mamy

$$\left| \sum_{n=p}^{\infty} f_n(x) g_n(x) \right| = \left| \sum_{n=p}^{q-1} s_{p,n}(x)(g_n(x) - g_{n+1}(x)) + s_{p,q}(x) g_q(x) - s_{p,p-1}(x) g_p(x) \right|$$
$$< |g_p(x)| \frac{\varepsilon}{3^k} + \frac{\varepsilon}{3^k}(|g_p(x) + g_q(x)|) \leqslant \varepsilon$$

dla każdego $x \in E$. Wobec Twierdzenia 144 dowód jest zakończony.

Lemat 3. Niech $a_m, a_{m+1}, \ldots, a_n$ i $b_m, b_{m+1}, \ldots, b_n$ będą liczbami rzeczywistymi, $m \leqslant n$. Oznaczmy

$$s_{m,k} = b_m + b_{m+1} + \cdots + b_k \quad \text{dla} \quad k = m, m+1, \ldots, n, \quad s_{m,m-1} = 0.$$

Wówczas
$$\sum_{k=m}^{n} a_k b_k = a_n s_{m,n} + \sum_{k=m}^{n-1} (a_k - a_{k+1}) s_{m,k}$$

(zob. [6], s.65-66).

10.2 Zbieżność jednostajna a ciągłość, różniczkowalność i całkowalność

Powstaje naturalne pytanie, czy przy założeniu, że funkcje f_n określone na zbiorze E są ciągłe lub różniczkowalne, lub całkowalne w sensie Riemanna, to analogiczną własność będzie miała funkcja graniczna?. Jaki związek zachodzi na przykład pomiędzy $f'_n i f'$ lub pomiędzy całkami Riemanna z f_n a całką Riemanna z f?
Pytanie czy granica ciągu funkcyjnego jest funkcją ciągłą jest pytaniem czy zachodzi równość

$$\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t) = \lim_{n\to\infty} f_n(x) \quad (\text{z ciągłośći } f_n; \quad (t, x \in E)),$$

tj. czy istotna jest kolejnośc, w jakiej dokunuje się przejść granicznych.

Przykład 42. Niech $f_n(x) = \frac{x^2}{(1+x^2)^n}$, $x \in \mathbb{R}$, $n \in \mathbb{N}$ i niech $f(x) = \sum\limits_{n=0}^{\infty} f_n(x) = \sum\limits_{n=0}^{\infty} \frac{x^2}{(1+x^2)^n}$. Ponieważ $f_n(0) = 0$, więc i $f(0) = 0$. Suma rozważanego szeregu jako szeregu geometrycznego jest równa $1 + x^2$ dla $x \neq 0$.
Zatem $f(x) = \begin{cases} 0 & \text{dla } x = 0 \\ 1 + x^2 & \text{dla } x \neq 0 \end{cases}$, a więc szereg, którego wyrazami są funkcję ciągłe, może mieć sumę nieciągłą.

Twierdzenie 149. *Niech $f_n \to f$ jednostajnie na zbiorze E (zawartym w $\mathbb{R}$ lub $\mathbb{C}$). Niech x będzie punktem skupienia zbioru E i niech*

$$\lim_{t\to x} f_n(t) = A_n \quad (n \in \mathbb{N}).$$

Wówczas ciąg (A_n) jest zbieżny i $\lim\limits_{t\to x} f(t) = \lim\limits_{n\to\infty} A_n$
(tj. $\lim\limits_{t\to x}\lim\limits_{n\to\infty} f_n(t) = \lim\limits_{n\to\infty}\lim\limits_{t\to x} f_n(t)$).

Dowód. Niech będzie dana liczba $\varepsilon > 0$. Ponieważ ciąg (f_n) jest zbieżny jednostajnie, więc istnieje takie N, że jeżeli $m, n \geqslant N$, $t \in E$ to zachodzi nierówność

$$|f_n(t) - f_m(t)| < \varepsilon.$$

Przechodząc w ostatniej nierówności do granicy z t $(t \to x)$, otrzymujemy $|A_n - A_m| \leqslant \varepsilon$ dla $m, n \geqslant N$. Stąd wynika, że ciąg (A_n) jest ciągiem Cauchy'ego, a zatem ma granicę. Oznaczmy tę granicę przez A. Wówczas mamy

$$|f(t) - A| \leqslant |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n + A|. \tag{75}$$

Wybierzmy n tak duże, żeby nierówność

$$|f(t) - f_n(t)| < \frac{\varepsilon}{3} \tag{76}$$

była spełniona dla każdego $t \in E$ (jest to możliwe dzięki jednostajnej zbieżności ciągu (f_n)) i żeby

$$|A_n - A| < \frac{\varepsilon}{3}. \tag{77}$$

Następnie dla tego n weźmy takie otoczenie V punktu x, aby dla $t \in V \cap E$, $t \neq x$ zachodziło

$$|f_n(t) - A_n| < \frac{\varepsilon}{3}. \tag{78}$$

Podstawiając zależności (76), (77), (78) do nierówności (75) otrzymujemy

$$|f(t) - A| < \varepsilon \text{ dla } t \in V \cap E, \quad t \neq x.$$

Stąd $\lim\limits_{t \to x} f(t) = \lim\limits_{n \to \infty} A_n$.

Twierdzenie 150. *Jeżeli (f_n) jest ciągiem funkcji ciągłych na zbiorze E i jeżeli $f_n \to f$ jednostajnie na E, to funkcja f jest ciągła na zbiorze E.*

Dowód. Jest to bezpośrednia konsekwencja Twierdzenia 149.

Uwaga 59. Twierdzenie odwrotne do Twierdzenia 150 nie jest prawdziwe, to znaczy ciąg funkcji ciągłych może być niejednostajnie zbieżny do funkcji ciągłej (zob. Przykład 41 (b)).

Wniosek 30. *Jeżeli (f_n) jest ciągiem funkcji ciągłych na zbiorze E i jeżeli ciąg (f_n) jest zbieżny jednostajnie na zbiorze E, to jego suma $s(x) = \sum\limits_{n=1}^{\infty} f_n(x)$, $x \in E$ jest funkcją ciągłą na tym zbiorze.*

Przykład 43. Roważmy ponownie funkcje z przykładu 41 (b). Mamy $\int\limits_0^1 f_n(x)dx = n^2 \int\limits_0^1 x(1-x^2)^n dx = n^2 \int\limits_0^1 \frac{t^n}{2} dt = \frac{n^2 t^{n+1}}{2(n+1)}\Big|_0^1 = \frac{n^2}{2n+2}$ $(1 - x^2 = t)$. Wobec tego $\lim\limits_{n \to \infty} \int\limits_0^1 f_n(x)dx = \lim\limits_{n \to \infty} \frac{n^2}{2n+2} = +\infty$, natomiast $\int\limits_0^1 \lim\limits_{n \to \infty} f_n(x)dx = 0$.

Twierdzenie 151. *Jeżeli (f_n) jest ciągiem funkcji R-całkowalnych na przedziale $[a, b]$ oraz $f_n \to f$ jednostajnie na $[a, b]$, to f jest funkcją R-całkowalną na $[a, b]$ oraz*

$$\int_a^b f dx = \lim_{n \to \infty} \int_a^b f_n dx.$$

Dowód. Wystarczy udowodnić twierdzenie dla funkcji f_n o wartościach rzeczywistych (analogicznie dla zespolonych).

Niech
$$\varepsilon_n = \sup_{x\in[a,b]} |f_n(x) - f(x)|.$$
Wówczas $f_n(x) - \varepsilon_n \leqslant f(x) \leqslant f_n(x) + \varepsilon_n$ dla $x \in [a, b]$, zatem wobec Definicji 71 oraz Twierdzenia 77 mamy

$$\int_a^b (f_n - \varepsilon_n)dx \leqslant \underline{\int_a^b} f dx \leqslant \overline{\int_a^b} f dx \leqslant \int_a^b (f_n + \varepsilon_n)dx/ \tag{79}$$

Stąd $0 \leqslant \overline{\int_a^b} f dx - \underline{\int_a^b} f dx \leqslant 2\varepsilon_n(b-a)$. Ponieważ na mocy Twierdzenia 145, $\varepsilon_n \to 0$ przy $n \to \infty$, a więc całki górna i dolna funkcji f są równe. Wobec tego funkcja f jest R-całkowalna na $[a, b]$. Przekształcając (79) inaczej otzymujemy
$$\left| \int_a^b f dx - \int_a^b f_n dx \right| \leqslant \varepsilon_n(b-a).$$
więc $\lim_{n\to\infty} \int_a^b f_n dx = \int_a^b f dx$.

Wniosek 31. *Jeżeli funkcje f_n są R-całkowalne na przedziale $[a, b]$ i szereg $\sum_{n=1}^{\infty} f_n$ jest jednostajnie zbieżny na $[a, b]$, to*
$$\int_a^b s dx = \sum_{n=1}^{\infty} \int_a^b f_n dx, \ \textit{gdzie}\ s(x) = \sum_{n=1}^{\infty} f_n(x) \ \textit{dla}\ x \in [a, b].$$

Inaczej mówiąc jednostajnie zbieżny szereg funkcyjny można całkować wyraz po wyrazie.

Przykład 44. Niech $f_n(x) = \frac{\sin(nx)}{\sqrt{n}}$ $(x \in \mathbb{R}, n \in \mathbb{N})$. Wówczas $f(x) = \lim_{n\to\infty} f_n(x) = 0$, czyli $f'(x) = 0$ dla $x \in \mathbb{R}$. Z drugiej strony $f_n'(x) = \sqrt{n}\cos(nx)$, więc (f_n') nie jest zbieżny do f'. Na przykład $f_n'(0) = \sqrt{n} \to +\infty$ przy $n \to \infty$, podczas gdy $f'(0) = 0$. Zbieżność jednostajna ciągu (f_n) nie pociąga za soba zbieżności punktowej ciągu $(f'n)$.

Twierdzenie 152. *Niech (f_n) będzie ciągiem funkcji różniczkowalnych na $[a, b]$ takim, że ciąg $f_n(x_0)$ jest zbieżny dla pewnego punktu $x_0 \in [a, b]$. Jeżeli ciąg (f_n') jest zbieżny jednostajnie na $[a, b]$, to także ciąg (f_n) jest zbieżny jednostajnie do pewnej funkcji f i zachodzi równość*
$$f'(x) = \lim_{n\to\infty} f_n'(x). \quad x \in [a, b].$$

Dowód. Niech będzie dana liczba $\varepsilon > 0$. Niech $N \in \mathbb{N}$ bedzie takie, aby dla $n, m \in \mathbb{N}$ zachodziło

$$|f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2} \tag{80}$$

oraz

$$|f_n'(t) - f_m'(t)| < \frac{\varepsilon}{2(b-a)} \quad (a \leqslant t \leqslant b). \tag{81}$$

Jeżeli do funkcji $f_n - f_m$ zastosujemy Twierdzenie 73 o Wartości Średniej, to dzięki (81) mamy

$$|f_n(x) - f_m(x) - f_n(t) + f_m(t)| \leqslant \frac{|x-t|\varepsilon}{2(b-a)} \leqslant \frac{\varepsilon}{2} \tag{82}$$

dla dowolnych wartośći x i t z przedziału $[a, b]$ oraz $n, m \geqslant N$. Z nierówności

$$|f_n(x) - f_m(x)| \leqslant |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)|$$

wynika, na mocy (80) i (82), że

$$|f_n(x) - f_m(x)| < \varepsilon \quad (a \leqslant x \leqslant b, \quad n, m \geqslant N)$$

i wobec tego ciąg (f_n) jest zbieżny jednostajnie na $[a, b]$. Niech

$$f(x) = \lim_{n\to\infty} f_n(x) \quad (a \leqslant x \leqslant b).$$

Ustalmy punkt x z przedziau $[a, b]$ i określmy

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t - x}, \quad \phi(t) = \frac{f(t) - f(x)}{t - x} \text{ dla } t \in [a, b], \quad t \neq x. \tag{83}$$

Wtedy

$$\lim_{t\to x} \phi_n(t) = f_n'(x) \quad (n \in \mathbb{N}) \tag{84}$$

Z pierwszej nierówności (82) wynika, że

$$|\phi_n(t) - \phi_m(t)| \leqslant \frac{\varepsilon}{2(b-a)} \quad (n, m \geqslant N)$$

i wobec tego ciąg (ϕ_n) jest jednostajnie zbieżny przy $t \neq x$. Ponieważ (f_n) jest zbieżny do f, zachodzi

$$\lim_{n\to\infty} \phi_n(t) = \phi(t) \tag{85}$$

jednostajnie na zbiorze tych t, że $a \leqslant t \leqslant b$, $t \leqslant x$. Stosując do ciągu (ϕ_n) Twierdzenie 149 wnioskujemy, że na podstawie (84) i (85), że

$$\lim_{t\to x} \lim_{n\to\infty} \phi_n(t) = \lim_{n\to\infty} \lim_{t\to x} \phi_n(t) = \lim_{n\to\infty} f_n'(x),$$

a zatem $f(x) = \lim_{n\to\infty} f'_n(x)$.

Wniosek 32. *Niech (f_n) będzie ciągiem funkcji rózniczkowalnych na $[a,b]$ takim, że szereg $\sum_{n=1}^{\infty} f_n(x_0)$ jest zbieżny dla pewnego punktu $x_0 \in [a,b]$. Jeżeli szereg $\sum_{n=1}^{\infty} f'_n$ jest zbieżny jednostajnie na $[a,b]$, to także szereg $\sum_{n=1}^{\infty} f_n$ jest zbieżny jednostajnie na $[a,b]$ oraz*

$$\left(\sum_{n=1}^{\infty} f_n(x)\right)' = \sum_{n=1}^{\infty} f'_n(x), \quad x \in [a,b].$$

Twierdzenie 153. *Istnieje funkcja rzeczywista określona na prostej rzeczywistej, która jest ciągła, lecz w żadnym punkcie nie posiada pochodnej.*

Dowód. Niech $\phi(x) = |x|$ dla $-1 \leqslant x \leqslant 1$ i rozszerzmy ϕ na zbiór wszystkich liczb rzeczywistych kładąc $\phi(x+2) = \phi(x)$. Wtedy dla dowolnych s i t

$$|\phi(s) - \phi(t)| \leqslant |s-t| \tag{86}$$

Ponieważ $\phi(\mathbb{R}) \subset [0,1]$ w uzasadnieniu (86) ze względu na okresowość funkcji ϕ wystarczy ograniczyć się do przedziałów $[-1,1]$ i $[0,2]$. Jeżeli $s,t \in [-1,1]$, to $|\phi(s) - \phi(t)| = ||s| - |t|| \leqslant |s-t|$. Niech teraz $s \in [0,1]$, $t \in [1,2]$. Ponieważ

$$\phi(x) = \begin{cases} x, & x \in [0,1], \\ -x+2, & x \in [1,2], \end{cases}$$

zatem $|\phi(s)-\phi(t)| = |s+t-2| = (\text{dla } s+t<2) = -s-t+2 \leqslant t-s = |s-t|$ — w pozostałych przypadkach rozmieszczenia punktów $s,t \in [0,2]$ rozumowanie jest analogiczne.
W szczególności funkcja ϕ jest ciągła na $\mathbb{R}$. Określmy

$$f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \phi(4^n x). \tag{87}$$

Ponieważ $0 \leqslant \phi \leqslant 1$, zatem na mocy Kryterium Weierstrassa (Twierdzenie 146) szereg (87) jest zbieżny jednostajnie na $\mathbb{R}$. Z twierdzenia 150 wynika więc, że jego suma jest funkcją ciągła na $\mathbb{R}$.
Ustalmy liczbę rzeczywistą x i liczbę naturalną m. Niech $\delta_m = \pm\frac{1}{2}4^{-m}$, przy czym znak wybieramy tak, aby pomiędzy $4^m x$ i $4^m(x+\delta_m)$ nie znajdowała się żadna liczba całkowita. Można to zrobić, bowiem $4^m|\delta_m| = \frac{1}{2}$. Określamy

$$\gamma_n = \frac{\phi(4^n(x+\delta_m)) - \phi(4^n x)}{\delta_m}.$$

Jeżeli $n > m$, to $4^n\delta_m$ jest liczbą parzystą i wobec tego $\gamma_n = 0$ (okresem funkcji ϕ jest liczba 2). Jeżeli $0 \leqslant n \leqslant m$, to (86) implikuje, że $|\gamma_n| \leqslant 4^n$. Ponieważ $|\delta_m| =$

$\frac{|\phi(4^m(x+\delta_m))-\phi(4^m x)|}{|\delta_m|} = 4^m$ (między liczbami 4^m i $4^m(x+\delta_m)$) nie znajduję się żadna liczba całkowita, $4^m|\delta_m| = \frac{1}{2}$, zatem $|\phi(4^m(x+\delta_m)) - \phi(4^m x)| = \frac{1}{2}$, a zatem

$$
\begin{aligned}
\left|\frac{f(x+\delta_m)-f(x)}{\delta_m}\right| &= \left|\frac{1}{\delta_m}\left(\sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n \phi(4^n(x+\delta_m)) - \sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n \phi(4^n x)\right)\right| \\
&= \left|\frac{1}{\delta_m}\sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n (\phi(4^n(x+\delta_m)) - \phi(4^n x))\right| \\
&= \left|\sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n \gamma_n\right| \\
&= \left|\left(\frac{3}{4}\right)^m \gamma_m + \sum_{n=0}^{m-1}\left(\frac{3}{4}\right)^n \gamma_n\right| \\
&\geqslant 3^m - \sum_{n=0}^{m-1} 3^n \\
&= 3^m - \frac{1-3^m}{1-3} \\
&= \frac{1}{2}(3^m+1).
\end{aligned}
$$

Przy $m \to \infty$, $\delta_m \to 0$, natomiast $\frac{1}{2}(3^m+1) \to +\infty$, Wynika stąd, że funkcja f nie jest różniczkowalna w punkcie x.

10.3 Szeregi potęgowe

Definicja 100. Niech będzie dany ciąg (c_n) liczb zespolonych. Szereg

$$\sum_{n=0}^{\infty} c_n z^n \quad \text{lub ogólniej} \quad \sum_{n=0}^{\infty} c_n (z-a)^n, \quad z, a \in \mathbb{C}$$

nazywamy szeregiem potęgowym. Liczby c_n nazywamy współczynnikami tego szeregu. Funkcje, które można przedstawić w postaci sumy szeregu potęgowego nazywamy funkcjami analitycznymi.

Twierdzenie 154. *Dla każdego szeregu potęgowego* $\sum\limits_{n=0}^{\infty} c_n z^n$ *istnieje dokładnie jedna liczba* $R \in \mathbb{R}_+ \cup \{0\}$ *o tej własności, że:*

(i) jeśli $|z| < R$*, to szereg potęgowy jest zbieżny bezwględnie,*

(ii) jeśli $|z| > R$*, to szereg potęgowy jest rozbieżny*

Liczbę R nazywamy promieniem zbieżnośći szeregu potęgowego.

Dowód. Niech $A = \left\{ |z| \in \mathbb{R}_+ \cup \{0\} : \lim\limits_{n\to\infty} c_n z^n = 0 \right\}$. Zbiór A jest niepusty, bowiem $0 \in A$. Kładziemy $R = \sup A$ (jeśli zbiór A jest nieograniczony to przyjmujemy $R = +\infty$). Jeśli $|z| > R$, to $z \notin A$, czyli $\lim\limits_{n\to\infty} c_n z^n \neq 0$, a zatem szereg $\sum\limits_{n=0}^{\infty} c_n z^n$ nie może być zbieżny, co dowodzi (ii). Zauważmy, że jeśli $R = 0$, to z (ii) wynika, że jeśli $|z| > 0$, to szereg potęgowy $\sum\limits_{n=0}^{\infty} c_n z^n$ będzie rozbieżny (w tym przypadku $\{z \in \mathbb{C} : |z| < 0\} = \emptyset$, ale każdy szereg potęgowy jest zbieżny dla $z = 0$.).
Załóżmy teraz, że $R > 0$. Wówczas jeśli $|z| < R$, to istnieje takie $z_0 \in A$, że $|z| < |z_0| < R$. Ponieważ $\lim\limits_{n\to\infty} c_n z^n = 0$, zatem istnieje $M > 0$ takie, że $|c_n z_0^n| \leqslant M$ dla każdego $n \in \mathbb{N}$. Mamy $|c_n z^n| = |c_n z_0^n| \cdot \left|\frac{z}{z_0}\right|^n \leqslant M q^n$, gdzie $q = \left|\frac{1}{z_0}\right| < 1$. Szereg $\sum\limits_{n=0}^{\infty} c_n z^n$ jest więc zbieżny bezwględnie na mocy Kryterium Porównawczego. Jednoznaczność liczby R jest oczywista.

Następne dwa twierdzenia podają wzory na obliczanie promienia zbieżności szeregu potęgowego.

Twierdzenie 155. *(Cauchy-Hadamard) Promień zbieżnośći szeregu potęgowego* $\sum\limits_{n=0}^{\infty} c_n z^n$ *obliczamy według wzory* $R = \frac{1}{\alpha}$*, gdzie* $\alpha = \overline{\lim\limits_{n\to\infty}} \sqrt[n]{|c_n|}$ *(jeśli* $\alpha = 0$*, to* $R = +\infty$*, jeśli* $\alpha = +\infty$*, to* $R = 0$*).*

Dowód. Przyjmijmy $a_n = c_n z^n$, Mamy

$$\overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|} = |z| \overline{\lim_{n\to\infty}} \sqrt[n]{|c_n|} = |z|\alpha.$$

Niech $\alpha \in (0, +\infty)$. Wówczas na mocy Kryterium Pierwiastkowego (Tw. 132), jeśli $|z|\alpha < 1$ czyli $|z| < \frac{1}{\alpha}$, to szereg potęgowy $\sum_{n=0}^{\infty} c_n z^n$ jest zbieżny, natomiast jeśli $|z| > \frac{1}{\alpha}$, to szereg potęgowy jest rozbieżny. Stąd $R = \frac{1}{\alpha}$.
Dalej jeśli $\alpha = 0$, to $\overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|} = 0$, zatem szereg potęgowy jest zbieżny dla każdego $z \in \mathbb{C}$, czyli $R = +\infty$.
W końcu jeśli $\alpha = +\infty$, to $\overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|} = +\infty$, czyli szereg potęgowy jest rozbieżny dla każdego $z \neq 0$, a zatem $R = 0$.

Twierdzenie 156. *Jeżeli ciąg $\left(\frac{c_{n+1}}{c_n}\right)$ ma granicę g, to promień zbieżności szeregu potęgowego $\sum_{n=0}^{\infty} c_n z^n$ jest równy $\frac{1}{g}$ (jeśli $g = 0$, to $R = +\infty$; jeśli $g = +\infty$, to $R = 0$).*

Dowód. Mamy $\lim_{n\to\infty} \left|\frac{c_{n+1}z^{n+1}}{c_n z^n}\right| = |z| \lim_{n\to\infty} \left|\frac{c_{n+1}}{c_n}\right| = |z|g$. Jeśli $g \in (0, +\infty)$, to dla $|z| < \frac{1}{g}$ rozważana granica jest mniejsza o 1, a zatem na mocy Kryterium d'Alamberta szereg potęgowy $\sum_{n=0}^{\infty} c_n z^n$ jest zbieżny. Stąd $R \geqslant \frac{1}{g}$. Gdyby jednak przypuścić, że $R > \frac{1}{g}$, to dla z takich, że $\frac{1}{g} < |z| < R$ szereg potęgowy byłby bezwględnie zbieżny, co jest niemożliwe ze względu na to, że $|z|g > 1$, a więc $R = \frac{1}{g}$. Dalej rozumujemy analogicznie jak w dowodzie Twierdzenia 155.

Przykład 45.

(a) Dla szeregu $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ mamy $\lim_{n\to\infty} \frac{\frac{1}{(n+1)!}}{\frac{1}{n!}} = \frac{1}{+\infty} = 0$, a zatem na mocy Twierdzenia 156, $R = +\infty$.

(b) Dla szeregu $\sum_{n=0}^{\infty} z^n$ mamy, wobec Twierdzenia 155, $R = 1$. Jeśli $|z| = 1$, to szereg jest rozbieżny ponieważ z^n nie dązy do zera przy $n \to \infty$.

(c) Dla szeregu $\sum_{n=0}^{\infty} \frac{z^n}{n}$ mamy $R = 1$. Dla $z = 1$ dany szereg jest oczywiście rozbieżny. Sprawdzimy, że dla wszystkich pozostałych punktów okręgu koła zbieżności (kołem zbieżności szeregu $\sum_{n=0}^{\infty} c_n(z-a)^n$ nazywamy koło otwarte o środku w punkcie a i promieniu R, gdzie R jest promieniem zbieżności danego szeregu potęgowego) szereg ten jest zbieżny. Ponieważ $a_n = \frac{1}{n}$ dąży monotonicznie do zera przy $n \to \infty$ oraz $\left|\sum_{m=0}^{n} z^m\right| =$

$\left|\frac{1-z^{n+1}}{1-z}\right| \leqslant \frac{2}{|1-z|}$, jeśli $|z| = 1$, $z \neq 1$, zatem na mocy Kryterium Dirichleta (Tw. 134), dany szereg jest zbieżny w każdym punkcie okręgu $|z| = 1$ z wyjątkiem punktu $z = 1$.
Ponadto $\sum\limits_{m=0}^{n} z^m \leqslant \frac{2}{\delta} = M$ dla $|z| \leqslant 1$ oraz $|1-z| \geqslant \delta > 0$, zatem rozważany szereg jest zbieżny w każdym zbiorze $A_\delta = \{z \in \mathbb{C} : |z| \leqslant 1 \wedge |1-z| \geqslant \delta\}$.

(d) Dla szeregu $\sum\limits_{n=0}^{\infty} \frac{z^n}{n^2}$ mamy $R = 1$. Szereg ten jest zbieżny we wszystkich punktach okręgu koła zbieżności, bowiem $\left|\frac{z^n}{n^2}\right| = \frac{1}{n^2}$ jeśli $|z| = 1$.

Z przykładów (b), (c) oraz (d) wynika, że zachowanie sie szeregu na okręgu koła zbieżności jest róznorodne i wymaga specjalnego badania.

Twierdzenie 157. *Niech szereg* $\sum\limits_{n=0}^{\infty} c_n z^n$ *będzie zbieżny przy* $|z| < R$ *i niech*

$$f(z) = \sum_{n=0}^{\infty} c_n z^n \quad (|z| < R). \tag{88}$$

Wówczas szereg potęgowy jest zbieżny jednostajnie w każdym kole domkniętym zawartym w kole zbieżności. Ponadto funkcja f jest ciągła i różniczkowalna w kole zbieżnośći oraz

$$f'(z) = \sum_{n=1}^{\infty} n c_n z^{n-1} \quad (|z| < R). \tag{89}$$

Innymi słowy szereg potęgowy można różniczkować “wyraz po wyrazie” wewnątrz koła zbieżności. O szeregu potęgowym mówimy ponadto, że jest on *niemial jednostajnie zbieżny* w kole zbieżności.

Dowód. Niech $0 < r < R$. Jeżeli $|z| \leqslant r$, $|c_n z^n| \leqslant |c_n r^n|$, a ponieważ $\sum\limits_{n=0}^{\infty} c_n r^n$ jest bewzględnie zbieżny, więc z Kryterium Weierstrassa wynika jednostajna zbieżność szeregu potęgowego na kole $\overline{B}(0, r)$ (koło domknięte o środku 0 i promieniu r).
Ponieważ $\sqrt[n]{n} \to 1$ przy $n \to \infty$, zatem $\overline{\lim\limits_{n\to\infty}} \sqrt[n]{n|c_n|} = \overline{\lim\limits_{n\to\infty}} \sqrt[n]{|c_n|}$, a więc szeregi (88) i (89) mają ten sam promień zbieżności.
Szereg (89) jako szereg potęgowy jest zbieżny jednostajnie na kole $\overline{B}(0, r)$ dla każdego $0 < r < R$ i wobec tego możemy stosować Wniosek 32. W takim razie (89) jest spełnione, jeśli tylko $|z| < r$. Jednakże dla każdego z takiego, że $|z| < R$, istnieje $0 < r < R$ takie, że $|z| < r$. Stąd (89) jest spełnione jeśli tylko $|z| < R$. Ciagłość funkcji f wynika z istnienia pochodnej f'.

Wniosek 33. *Jeżeli spełnione są założenia Twierdzenia 157, to funkcja f ma pochodne wszystkich rzędów w kole zbieżności $\overline{B}(0,r)$, przy czym*

$$f^{(k)}(z) = \sum_{n=k}^{\infty} n(n-1)(n-2)\dots(n-k+1)c_n z^{n-k}, \tag{90}$$

a w szczególności

$$f^{(k)}(0) = k!c_k \quad (k = 0, 1, 2, \dots) \tag{91}$$

($f^{(0)}$ oznacza tutaj funkcję, a $f^{(k)}$ - k-tą pochodną funkcji f przy $k \in \mathbb{N}$).

Dowód. Równość (90) otrzymamy stosując kolejno Twierdzenie 157 do funkcji f, a następnie do f', f'' itd., wykorzystując Twierdzenie 34 (b) ($\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$). Następnie podstawiając w (90), $x = 0$ otrzymujemy (91).

Rozważmy teraz szereg potęgowy $\sum\limits_{n=0}^{\infty} a_n x^n$, gdzie $x \in \mathbb{R}$, $a_n \in \mathbb{R}$ dla każdego $n \in \mathbb{N} \cup \{0\}$. Jeśli szereg ten jest zbieżny na końcu przedziału zbieżności, na przykład w punkcie $z = r$ ($r < +\infty$, oczywiście), to jego suma jest funkcją ciągłą nie tylko na przedziale $(-R, R)$, lecz również prawostronnie w punkcie $x = R$. Wynika to następującego **Twierdzenia Abela**.

Twierdzenie 158. *(Abela). Niech szereg potęgowy $\sum\limits_{n=0}^{\infty} a_n x^n$ ma promień zbieżności $R > 0$. Jeżeli szereg ten jest zbieżny na końcu przedziału zbieżności (to znaczy dla $x = R$ lub $x = -R$) to suma tego szeregu jest ciągła jednostajnie w tym końcu.*

Dowód. Załóżmy, że szereg $\sum\limits_{n=0}^{\infty} a_n R^n$ jest zbieżny. Oznaczmy n-tą resztę tego szeregu przez R_n:

$$R_n = a_{n+1}R^{n+1} + a_{n+2}R^{n+2} + \dots$$

Mamy więc $\lim\limits_{n\to\infty} R_n = 0$. Niech będzie dane $\varepsilon > 0$. Istnieje zatem takie $k \in \mathbb{N}$, że dla $i > n > k$ mamy

$$\left| a_{n+1}R^{n+1} + \dots + a_i R^i \right| < \varepsilon.$$

Oznaczmy ogólnie n-tą resztę szeregu $\sum\limits_{n=0}^{\infty} a_n x^n$ przez $R_n(x)$:

$$R_n(x) = a_{n+1}x^{n+1} + a_{n+2}x^{n+2} + \dots$$

W szczególności $R_n(R) = R_n$. Zauważmy, że

$$R_n(x) = a_{n+1}R^{n+1}\left(\frac{x}{R}\right)^{n+1} + a_{n+2}R^{n+2}\left(\frac{x}{R}\right)^{n+2} + \dots$$

i zastosujmy Kryterium Dirichleta (Tw. 134), przyjmując $a_n = (\frac{x}{R})^{n+1}$, $n \in \mathbb{N}$ oraz $b_n = a_{n+1}R^{n+1}$, $n \in \mathbb{N}$. Sumy częściowe szeregu $\sum\limits_{k=n+1}^{\infty} a_k R^k$ tworzą ciąg ograniczony (poprzez liczbę ε) oraz

$$1 > \left(\frac{x}{R}\right)^{n+1} > \left(\frac{x}{R}\right)^{n+2} > \dots \quad \text{i} \quad \lim_{n\to\infty} \left(\frac{x}{R}\right)^n = 0 \quad \text{dla} \quad 0 \leqslant x < R.$$

Ponadto z dowodu Kryterium Dirichleta wynika, że $|R_n(x)| < 2\varepsilon$, $0 \leqslant x < R$. Nierówność ta jest również spełniona dla $x = R$. Doszliśmy więc do wniosku, że jeśli $n > k$, to nierówność

$$\left| \sum_{m=0}^{\infty} a_m x^m - \sum_{m=0}^{n} a_m x^m \right| < 2\varepsilon$$

jest spełniona przez każde x takie, że $0 \leqslant x < R$. Oznacza to, że szereg $\sum\limits_{n=0}^{\infty} a_n x^n$ jest w tym przedziale jednostajnie zbieżny. Wobec Twierdzenia 150, jego suma jest ciągła lewstronnie dla $x = R$.
W przypadku, gdy szereg $\sum\limits_{n=0}^{\infty} a_n(-R)^n$ jest zbieżny, rozumowanie jest analogiczne.

Dowód Twierdzenia Abela dla zespolonych szeregów potęgowych można znaleźć w książce F. Leji, Funkcje Zespolone, PMN, Warszawa. 1973, s. 51 - 52.

Wniosek 34. *Jeżeli szeregi liczbowe rzeczywiste* $\sum\limits_{n=0}^{\infty} a_n$, $\sum\limits_{n=0}^{\infty} b_n$, $\sum\limits_{n=0}^{\infty} c_n$ *są zbieżne odpowiednio do* A, B *i* C *i jeżeli* $c_n = a_0 b_n + \dots + a_n b_0$, *to* $C = AB$.

Dowód. Określmy $f(x) = \sum\limits_{n=0}^{\infty} a_n$, $g(x) = \sum\limits_{n=0}^{\infty} b_n$, $h(x) = \sum\limits_{n=0}^{\infty} c_n$ dla $0 \leqslant x <\leqslant 1$.
Z Twierdzenia 154 wynika, że R dla tych szeregów nie może być mniejsze od 1. Gdyby $R < 1$, to na przykład $\sum\limits_{n=0}^{\infty} a_n$ byłby rozbieżny. Jeżeli $x < 1$, to szeregi te są zbieżne bewzględnie, to można je mnożyć według Definicji 95.
Po wymnożeniu otrzymujemy

$$f(x)g(x) = h(x) \quad (0 \leqslant x < 1). \tag{92}$$

Z Twierdzenia 154 wynika, że $f(x) \to A$, $g(x) \to B$, $h(x) \to C$, przy $x \to 1$. Wobec równości (92) mamy $AB = C$.

Zajmiemy sie teraz zagadnieniem rozwijania funkcji w szereg potęgowy.

Definicja 101. Jeżeli szereg $\sum\limits_{n=0}^{\infty} c_n z^n$ jest zbieżny dla każdego $z \in B(0, R)$ przy pewnym $R > 0$, to powiemy, że funkcja $f(z) = \sum\limits_{n=0}^{\infty} c_n z^n$, $z \in B(0, R)$ pozwala rozwinąc się w szereg potęgowy w otoczeniu punktu $z = 0$. Analogicznie jeżeli szereg $\sum\limits_{n=0}^{\infty} c_n(z-a)^n$ jest zbieżny dla z spełniających nierówność $|z-a| < R$, to powiemy, że szereg potęgowy jest zbieżny w otoczeniu punktu $z = a$.

Definicja 102. Załóżmy, że funkcja f jest klasy $C^{(\infty)}$ w pewnym otoczeniu punktu z_0. Szereg $\sum\limits_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z-z_0)^n$ dla z należacych do tego otoczenia, nazywamy szeregiem Taylora funkcji f. Dla $z = 0$ szereg Taylora nazywa się szeregiem Maclaurina funkcji f.

Dla funkcji f o wartościach zespolonych dowodzi się, że jeśli jest ona różcznikowalna w pewnym otoczeniu punktu z_0, to jest ona klasy $C^{(\infty)}$ w tym otoczeniu.

Twierdzenie 159. *Jeżeli funkcja f posiada rozwinięcie w szereg potęgowy w otoczeniu punktu z_0, to jest to jedyne rozwinięcie i to w szereg Taylora.*

Dowód. Załóżmy, że $f(z) = \sum\limits_{n=0}^{\infty} c_n(z-z_0)^n$ dla $x \in B(0,R)$, $R > 0$. Wobec Twierdzenia 157, $n!c_n = f^{(n)}(z_0)$ dla każdego $n \in \mathbb{N} \cup \{0\}$. Stąd $c_n = \frac{f^{(n)}(z_0)}{n!}$, co kończy dowód.

Zajmiejmy się teraz funkcjami o argumentach i wartościach rzeczywistych. Powstaje pytanie, jakie funkcję są sumami szeregów potęgowych. Z Twierdzenia 158 wynika, że funkcje takie muszą mieć pochodnie wszystkich rzędów. Następujący przykład pokazuje, że nie jest to warunek wystarczający rozwijalności funkcji w szereg potęgowy (dla funkcji rzeczywistych).

Przykład 46. Niech $f(x) = \begin{cases} e^{\frac{-1}{x^2}} & \text{jeśli } x \neq 0 \\ 0 & \text{jeśli } x = 0 \end{cases}$. W oparciu o Regułe de l'Hospitala można łatwo sprawdzić, że $\frac{1}{x^n} e^{\frac{-1}{x^2}} \to 0$, gdy $x \to 0$ ($n \in \mathbb{N}$ jest ustalone). W tym celu wystarczy wyliczyć granicę $\lim\limits_{y \to +\infty} \frac{y^{\frac{n}{2}}}{e^y}$. Oznaczmy $f(y) = y^{\frac{n}{2}}$ i $g(y) = e^y$. Wówczas $g^{(m)}(y) = e^y \neq 0$ ($y \in \mathbb{R}$, $m = 1, 2, \ldots$). Biorąc $m > \frac{n}{2}$ i różniczkując n-razy otrzymujemy

$$\frac{f^{(m)}(y)}{g^{(m)}(y)} = \frac{\frac{n}{2}(\frac{n}{2}-1)\cdots(\frac{n}{2}-m+1)y^{\frac{n}{2}-m}}{e^y} = \frac{m!\binom{\frac{n}{2}}{m}}{y^{m-\frac{n}{2}}e^y} \to 0 \text{ jeśli } y \to +\infty.$$

Stąd

$$0 = \lim_{y \to +\infty} \frac{f^{(m)}(y)}{g^{(m)}(y)} = \cdots = \lim_{y \to +\infty} \frac{f'(y)}{g'(y)} = \lim_{y \to +\infty} \frac{y^{\frac{n}{2}}}{e^y}.$$

Pokażemy, że funkcja f posiada pochodnie wszystkich rzędów w punkcie $x = 0$ i, że $f^{(n)}(0) = 0$. Dokładniej udowodnimy, że

$$f^{(n)}(x) = P_{3n}\Big(\frac{1}{x}\Big)e^{\frac{-1}{x^2}} \quad x \neq 0$$

$$f^{(n)}(0) = 0,$$

gdzie $P_{3n}(x)$ jest wielomianem stopnia $3n$.
Dla $n = 1$ mamy $f'(x) = \frac{2}{x^3} e^{\frac{-1}{x^2}}$, gdy $x \neq 0$. Natomiast dla $x \neq 0$ mamy $\frac{f(x)-f(0)}{x} = \frac{e^{\frac{-1}{x^2}}}{x} \to 0$, gdy $x \to 0$. Stąd otrzymujemy, że $f'(0) = 0$.
Przypuśćmy teraz, że teza jest prawdziwa dla liczby n. Wówczas

$$\begin{aligned} f^{(n+1)}(x) =& (f^{(n)})'(x) \\ =& -\frac{1}{x^2} P'_{3n}\left(\frac{1}{x}\right) e^{-\frac{1}{x^2}} + P_{3n}\left(\frac{1}{x}\right) \frac{2}{x^3} e^{-\frac{1}{x^2}} \\ =& \left[\frac{2}{x^3} P_{3n}\left(\frac{1}{x}\right) - \frac{1}{x^2} P'_{3n}\left(\frac{1}{x}\right) \right] e^{-\frac{1}{x^2}} \\ =& P_{3(n+1)}\left(\frac{1}{x}\right) e^{-\frac{1}{x^2}} \end{aligned}$$

oraz

$$\frac{f^{(n)}(x) - f^{(n)}(0)}{x} = \frac{P_{3n}(\frac{1}{x}) e^{-\frac{1}{x^2}}}{x} = Q_{3n+1}\left(\frac{1}{x}\right) e^{-\frac{1}{x^2}} \to 0, \text{ gdy } x \to 0.$$

gdzie Q_{3n+1} jest wielomianem stopnia $3n + 1$.
Na mocy Zasady Indukcji Matematycznej teza jest prawdziwa dla dowolnej liczby naturalnej n. W szczególności więc funkcja f ma w zerze wszystkie pochodnie równe zeru. Stąd $\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = 0$ podczas gdy $f(x) = e^{-\frac{1}{x^2}} \neq 0$.

Następujące twierdzenie wynikające natychmiast z Twierdzenia 65 podaje warunek rozwijalności w szereg potęgowy.

Twierdzenie 65. *(Taylor). Jeśli na przedziale domkniętym o końcach x_0, x funkcja f jest ciągła razem ze swoimi pochodnymi do rzędu n włącznie, a w wewnętrznych punktach przedziału posiada ona pochodną rzędu $n + 1$, to dla dowolnej funkcji ϕ ciągłej na tym przedziale i mającej różną od zera pochodną w jego wewnętrznych punktach, można znaleźć taki punkt $\xi \in (x_0, x)$, że*

$$r_n(x_0, x) = \frac{\phi(x) - \phi(x_0)}{\phi'(\xi)n!} f^{(n+1)}(\xi)(x - \xi)^n$$

Twierdzenie 160. *Jeśli na przedziale domkniętym o końcach x_0, x funkcja f ma pochodne wszystkich rzędów oraz reszta $r_n(x_0, x)$ we Wzorze Taylora dąży* *do* 0 *przy* $k \to \infty$*, to*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

Wniosek 35. *Jeżeli na przedziale domkniętym o końcach* x_0*,* x *funkcja* f *ma pochodne wszystkich rzędów oraz istnieje stała* $K > 0$ *taka, że* $|f^{(n)}(t)| \leqslant K$ *dla wszystkich* t *należących do tego przedziału i* $n = 0, 1, 2, \ldots$*, to*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

Wniosek 14. *Kładąc do wzoru z Twierdzenia 65* $\phi(t) = (x - t)^{n+1}$ *otrzymujemy*

$$r_n(x_0, x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x - x_0)^{n+1}$$

Dowód. Ponieważ dla Reszty Lagrange'a (Wn. 14) mamy

$$|r_n(x_0, x)| = \left| \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x - x_0)^{n+1} \right| \leqslant K \frac{|x - x_0|^{n+1}}{(n+1)!} \to 0 \text{ przy } n \to \infty$$

Podamy jeszcze jedno twierdzenie dotyczące rozwijalności funkcji w szereg Taylora.

Twierdzenie 161. *Niech szereg* $\sum\limits_{n=0}^{\infty} c_n x^n$ *będzie zbieżny dla* $|x| < R$ *i niech* $f(x)$ *oznacza sumę tego szeregu na przedziale* $(-R, R)$*. Jeżeli* $-R < a < R$*, to funkcje* f *można rozwinąc w punkcie* a *w szereg potęgowy, który jest zbieżny dla* $|x - a| < R - |a|$*:*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x - a)^n \quad (|x - a| < R - |a|).$$

Przykład 47. Powróćmy do Przykładu 22. Otrzymane w nim wyniki możemy zapisać następująco:

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

oraz

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n,$$

gdzie

$$\alpha \in \mathbb{R}, |x| < 1, \binom{\alpha}{n} = \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}, \binom{\alpha}{0} = 1.$$

Otrzymaliśmy zatem rozwinięcie funkcji wykładniczej, sinusa, cosinusa w szereg Maclaurina...

Niniejszy paragraf zakończymy krótką informacją o funkcjach e^z, $\cos z$, $\sin z$ w dziedzinie zespolonej. Dla $z \in \mathbb{C}$ definiujemy

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad \sin(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}, \quad \cos(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}.$$

Twierdzenie 162.

(a) $e^a e^b = e^{a+b}$ *dla dowolnych* $a, b \in \mathbb{C}$.

(b) Funkcja e^z *nie przyjmuje nigdzie wartości* 0.

(c) Między funkcjami e^z, $\cos z$, $\sin z$ *zachodzą następujące związki:*

$$e^{iz} = \cos z + i \sin z, \quad e^{-iz} = \cos z - i \sin z. \tag{93}$$

(d) $(e^z)' = e^z$, $(\cos z)' = -\sin z$, $(\sin z)' = \cos z$.

(e) $e^z = 1$ *wtedy i tylko wtedy, gdy* $z = 2k\pi i$, $k = 0, \pm 1, \pm 2, \dots$.

Dowód.

(a) Szeregi $\sum\limits_{n=0}^{\infty} \frac{a^n}{n!}$ i $\sum\limits_{n=0}^{\infty} \frac{b^n}{n!}$ są bezwzględnie zbieżnie dla dowolnych wartości $a, b \in \mathbb{C}$. Wobec Twierdzenia Mertensa mamy

$$\left(\sum_{n=0}^{\infty} \frac{a^n}{n!}\right)\left(\sum_{n=0}^{\infty} \frac{b^n}{n!}\right) = \sum_{n=0}^{\infty} \frac{(a+b)^n}{n!},$$

bowiem

$$\frac{a^n}{n!}+\frac{a^{n-1}}{(n-1)!}\frac{b}{1!}+\cdots+\frac{b^n}{n!}=\frac{1}{n!}\left[a^n+\binom{n}{1}a^{n-1}b+\cdots+\binom{n}{n}b^n\right]=\frac{(a+b)^n}{n!}$$

stąd $e^a e^b = e^{a+b}$.

(b) Gdyby bowiem dla pewnej wartości $z = a$ było $e^a = 0$, to ponieważ liczba e^{-a} jest skończona, zatem iloczyn $e^a e^{-a}$ miałby wartość 0. Na mocy (a) mamy $e^a e^{-a} = e^0 = 1$, co daje sprzeczność, zatem $e^a \neq 0$.

(c) Mamy

$$e^{iz} = \sum_{n=0}^{\infty} i^n \frac{z^n}{n!} = \sum_{k=0}^{\infty} i^{2k} \frac{z^{2k}}{(2k!)} + i \sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!} = \cos z + i \sin z,$$

bowiem $i^{2k} = (-1)^k$.
Drugi z wzorów (93) otrzymamy zastępując w pierwszym z przez $-z$, bowiem z określenia funkcji $\sin z$, $\cos z$ wynika, że $\cos(-z) = \cos z$, $\sin(-z) = -\sin z$.

(d) Na mocy Twierdzenia 157 mamy

$$(e^z)' = \left(\sum_{n=0}^{\infty} \frac{z^n}{n!}\right) = \sum_{n=1}^{\infty} n \frac{z^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{z^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{z^n}{n!} = e^z.$$

Pozostałe dwa wzory uzasadniamy analogicznie.

(e) Istotnie stosując (a) i (c) otrzymujemy $e^z = e^x e^{iy} = e^x(\cos y + i \sin y)$, a więc równość $e^z = 1$ zachodzi wtedy i tylko wtedy, gdy $e^x \cos y = 1$ oraz $e^x \sin y = 0$. Stąd wynika, że $\sin y = 0$, bowiem $e^x > 0$, a zatem $y = n\pi$, gdzie $n = 0, \pm 1, \pm 2, \ldots$. Gdy n jest liczbą nieparzystą, wówczas $\cos y = \cos(n\pi)$ jest liczbą ujemną; aby więc zachodziło $e^x \cos y = 1$, n musi być liczbą parzystą, stąd $y = 2k\pi$ oraz $e^x = 1$, gdzie $k = 0, \pm 1, \pm 2, \ldots$. Funkcja e^x zmiennej rzeczywistej przyjmuje wartość 1 tylko dla $x = 0$, zatem $e^z = 1$, gdy $z = x + iy = 2k\pi i$, co należało dowieść.

Wniosek 36.

(a) Mamy

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad \sin z = \frac{e^{iz} - e^{-iz}}{2i}$$

Wzory te nazywamy wzorami Eulera. Wzory te otrzymujemy natychmiast dodając i odejmując równania (93).

(b) *Z wzorów Eulera wynika następujący związek*

$$\sin^2 z + \cos^2 z = 1 \quad z \in \mathbb{C}.$$

Mamy bowiem

$$\sin^2 z + \cos^2 z = \frac{e^{2iz} + 2e^{iz}e^{-iz} + e^{-2iz}}{4} - \frac{e^{2iz} - 2e^{iz}e^{-iz} + e^{-2iz}}{4} = 1.$$

(c) *Nierówność* $|\cos x| \leqslant 1$ *i* $|\sin x| \leqslant 1$ *prawdziwe dla rzeczywistych* x*, przestają na ogół być prawdziwe dla wartości zespolonych. Na przykład dla* $z = i$ *mamy*

$$\cos i = \frac{e^{-1} + e}{2} = 1,532\ldots, \quad \sin i = \frac{e^{-1} - e}{2i} = i \cdot 1,175\ldots.$$

Mimo, to $\sin^2 i + \cos^2 i = 1$.

(d) *Funkcja* e^z *jest okresowa. Okresem tej funkjci jest każda wielokrotność liczby* $2\pi i$*. Istotnie z Twierdzenia 162 (c) wynika, że jeżeli* $z' = z + 2k\pi i$*,* $k \in \mathbb{Z}$*, to*

$$e^{z'} = e^z e^{2k\pi i} = e^z$$

odwrotnie, jeżeli $e^{z'} = e^z$ *wtedy i tylko wtedy, gdy* $z' = z + 2k\pi i$*,* $k \in \mathbb{Z}$*.*

10.4 Szeregi Fouriera

Definicja 103. Szeregiem trygonometrycznym nazywamy szereg funkcyjny w postaci

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx) \quad (x \in \mathbb{R}) \tag{94}$$

gdzie a_n, b_n są stałymi ($a_n, b_n \in \mathbb{C}$).

Aby ustalić możliwość rozwinięcia w szereg trygonometryczny (94) dla danej funkcji f o okresie 2π należy rozpocząć od ustalenia ciągu współczynników $a_0, a_1, b_1, a_2, b_2, \ldots$. Pokażemy teraz metodę wyznaczania tych współczynników pochodzącą od Eulera i Fouriera. Załóżmy, że funkcja f (o wartościach rzeczywistych lub zespolonych) jest R-całkowalna na przedziale $[-\pi, \pi]$. Ponadto załóżmy, że funkcja f ma rozwinięcie w szereg trygonometryczny (94) jednostajnie zbieżny. Możemy go zatem całkować wyraz po wyrazie na tym przedziale. Mamy

$$\int_{-\pi}^{\pi} f(x)dx = \pi a_0 + \sum_{n=1}^{\infty}(a_n \int_{-\pi}^{\pi} \cos nxdx + b_n \int_{-\pi}^{\pi} \sin nxdx).$$

Ponadto

$$\int_{-\pi}^{\pi} \cos nxdx = \frac{\sin nx}{n}\Big|_{\pi}^{\pi} = 0,$$

$$\int_{-\pi}^{\pi} \sin nxdx = -\frac{\cos nx}{n}\Big|_{\pi}^{\pi} = 0,$$

dlatego wszystkie wyrazy pod znakiem sumy są równe zero, czyli

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)dx.$$

Aby ustalić wartość współczynników a_m ($m \in \mathbb{N}$ — ustalone) pomnóżmy obie strony równości (94) przez $\cos mx$. Wówczas szereg ten pozostaje w dalszym ciągu jednostajnie zbieżny (jest to bezpośrednia konsekwencja Definicji 99) i możemy go całkować wyraz po wyrazie na tym samym przedziale:

$$\begin{aligned}\int_{-\pi}^{\pi} f(x)\cos mxdx =& \frac{a_0}{2}\int_{-\pi}^{\pi} \cos mxdx \\ &+ \sum_{n=1}^{\infty}(a_n \int_{-\pi}^{\pi} \cos nx \cos mxdx + b_n \int_{-\pi}^{\pi} \sin nx \cos mxdx).\end{aligned}$$

Ponieważ

$$\int_{-\pi}^{\pi} \sin nx \cos mxdx = \frac{1}{2}\int_{-\pi}^{\pi} \sin(n+m)x + \sin(n-m)xdx = 0,$$

$$\int_{-\pi}^{\pi} \cos nx \cos mx dx = \frac{1}{2}\int_{-\pi}^{\pi} \cos(n+m)x + \cos(n-m)x dx = 0, \text{ jeśli } n \neq m$$

oraz (dla $n = m$)

$$\int_{-\pi}^{\pi} \cos^2 mx dx = \int_{-\pi}^{\pi} \frac{1+\cos 2mx}{2} dx = \frac{1}{2}(x + \frac{1}{2m}\sin 2mx)\Big|_{-\pi}^{\pi} = \pi,$$

zatem pod znakiem sumy pozostaje jedynie całka mnożona przez współczynnik a_m. Stąd

$$a_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos mx dx, \quad m \in \mathbb{N}.$$

Analogicznie mnożąc szereg (94) przez $\sin mx$ i całkując wyraz po wyrazie wyznaczamy współczynnik przy sinusie:

$$b_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin mx dx, \quad m \in \mathbb{N}.$$

Tym razem wykorzystujemy równości:

$$\int_{-\pi}^{\pi} \sin nx \sin mx dx = \begin{cases} 0 & \text{jeśli } n \neq m, \\ \pi & \text{jeśli } n = m. \end{cases}$$

Definicja 104. Wzory na współczynniki a_n i b_n noszą nazwę wzorów Eulera-Fouriera, a same współczynniki nazywają się współczynnikami Fouriera danej funkcji.

Definicja 105. Niech f bedzię funkcją R-całkowalną na przedziale $[-\pi, \pi]$. Wówczas można obliczyć współczynniki Fouriera funkcji f i zbadać szereg (94). Szereg trygonometryczny o tak dobranych współczynnikach nazywamy szeregiem Fouriera funkcji f i zapisujemy

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx).$$

Pisząc powyżej znak “$\sim$” nie zakładamy niczego na temat zbieżności szeregu po jego prawej stronie. Szereg ten nie musi być zbieżny, a jeśli jest zbieżny to nie znaczy to, by jego suma była równa $f(x)$.

Wyprowadzając wzory Eulera-Fouriera w istocie udowodniliśmy następujące

Twierdzenie 163. *Jeśli funkcja f R-całkowalna na przedziale $[-\pi, \pi]$ i okresowa o okresie 2π daje się rozwinąć w szereg trygonometryczny jednostajnie zbieżny, to ten szereg jest jej szeregiem Fouriera.*

Uwaga 60. Wzory Eulera-Fouriera przybierają prostszą postać, gdy rozważana funkcja f jest albo parzysta, albo nieparzysta. Jeśli f jest parzysta to wówczas funkcje $f(x)\cos nx$ są również parzystę, podczas gdy $f(x)\sin nx$ są nieparzyste. Wynika stąd, że wzory na współczynniki szeregu Fouriera tej funkcji przybierają następującą postać:

$$a_0 = \frac{2}{\pi}\int_0^\pi f(x)dx, \quad a_n = \frac{2}{\pi}\int_0^\pi f(x)\cos nx dx, \quad b_n = 0, \quad n \in \mathbb{N}.$$

Natomiast w przypadku gdy, f jest funkcją nieparzystą, nieparzyste są również funkcje $f(x)\cos nx$. Z kolei $f(x)\sin nx$ jest funkcją parzystą. Tym razem mamy więc

$$a_0 = 0, \quad a_n = 0, \quad b_n = \frac{2}{\pi}\int_0^\pi f(x)\sin nx dx, \quad n \in \mathbb{N}.$$

Pokażemy teraz, że szereg trygonometryczny (94) można przedstawić w innej postaci, która jest w wielu sytuacjach dogodniejsza.

Niech f będzie funkcją 2π-okresową, R-całkowalną na przedziale $[-\pi, \pi]$. Niech

$$\frac{a_0}{2} + \sum_{n=1}^\infty (a_n \cos nx + b_n \sin nx)$$

będzie jej szeregiem Fouriera. Przekształcimy wyraz ogólny tego szeregu za pomocą wzorów Eulera (Wn. 36 (a)). Mamy

$$\begin{aligned} a_n \cos nx + b_n \sin nx &= a_n \frac{e^{inx} + e^{-inx}}{2} + b_n \frac{e^{inx} - e^{-inx}}{2i} \\ &= \frac{a_n - ib_n}{2} e^{inx} + \frac{a_n + ib_n}{2} e^{-inx} \\ &= c_n e^{inx} + c_{-n} e^{-inx} \end{aligned}$$

gdzie

$$c_n = \frac{a_n - ib_n}{2}, \quad c_{-n} = \frac{a_n + ib_n}{2}.$$

Oznaczając $c_0 = \frac{a_0}{2}$ otrzymamy dla sum częsciowych szeregu Fouriera następujące wyrażenie

$$\frac{a_0}{2} + \sum_{n=1}^N (a_n \cos nx + b_n \sin nx) = c_0 + \sum_{n=1}^N (c_n e^{inx} + c_{-n} e^{-inx}) = \sum_{n=-N}^N c_n e^{inx}.$$

Dla nowych współczynników Fouriera c_n otrzymamy wzór

$$\begin{aligned} c_n = \frac{a_n - ib_n}{2} &= \frac{1}{2}\Big(\frac{1}{\pi}\int_{-\pi}^\pi f(x)\cos nx dx - i\frac{1}{\pi}\int_{-\pi}^\pi f(x)\sin nx dx\Big) \\ &= \frac{1}{2\pi}\int_{-\pi}^\pi f(x)(\cos nx - i \sin nx) dx \\ &= \frac{1}{2\pi}\int_{-\pi}^\pi f(x) e^{-inx} dx. \end{aligned}$$

Powyższy wzór wyprowadziliśmy dla $n > 0$. Bezpośrednio widać, że jest on prawdziwy dla $n = 0$. Można bez trudny sprawdzić, że jest on również prawdziwy dla $n < 0$. Widzimy więc, że zbieżność szeregu Fouriera funkcji f oznacza, że istnieje granica

$$\begin{aligned} \lim_{N\to\infty}\Big(\frac{a_0}{2}+\sum_{n=1}^{N}(a_n\cos nx+b_n\sin nx)\Big) &= \lim_{N\to\infty}\sum_{n=-N}^{N}c_ne^{inx} \\ &= \sum_{n=-\infty}^{\infty}c_ne^{inx}. \end{aligned} \tag{95}$$

Definicja 106. Szereg (95) nazywamy szeregiem Fouriera w postaci zespolonej. Współczynniki tego szeregu wyrażają się wzorami:

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{inx}dx \quad (n = 0, \pm 1, \pm 2, \ldots).$$

Zajmiemy się teraz badaniem zbieżności szeregu Fouriera. Udowodnimy wpierw

Lemat 4. *(Riemann-Lebesgue) Jeśli funkcja g jest R-całkowalna na przedziale $[a, b]$, to*

$$\lim_{p\to\infty}\int_a^b g(t)\sin pt dt = 0 \qquad \textit{oraz} \qquad \lim_{p\to\infty}\int_a^b g(t)\cos pt dt = 0.$$

Dowód. Wystarczy przeprowadzić dowód dla pierwszej z tych granic i dla funkcji rzeczywistych. Zauważmy wpierw, że na dowolnym przedziale ograniczonym $[\alpha, \beta]$ prawdziwe jest oszacowanie

$$\left|\int_\alpha^\beta \sin pt dt\right| = \left|\frac{\cos p\alpha - \cos p\beta}{p}\right| \leqslant \frac{2}{p}.$$

Wybierzmy dowolne $\varepsilon > 0$. Istnieje wówczas taki podział $P = \{t_0, t_1, \ldots, t_n\}$ przedziału $[a, b]$, że $U(f, P) - L(f, P) < \frac{\varepsilon}{2}$. Oznaczmy $M_j = \sup\{g(t) : t_{j-1} \leqslant t \leqslant t_j\}$ i $m_j = \inf\{g(t) : t_{j-1} \leqslant t \leqslant t_j\}$ dla $j = 1, 2, \ldots, n$. Niech $p > \frac{4}{\varepsilon}\sum_{j=1}^{n}|m_j|$. Wówczas mamy

$$\begin{aligned} \int_a^b g(t)\sin pt dt &= \sum_{j=1}^{n}\int_{t_{j-1}}^{t_j} g(t)\sin pt dt \\ &= \sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}(g(t)-m_j)\sin pt dt + \sum_{j=1}^{n}m_j\int_{t_{j-1}}^{t_j}\sin pt dt. \end{aligned}$$

Ponieważ $0 \leqslant g(t) - m_j \leqslant M_j - m_j$ dla $t_{j-1} \leqslant t \leqslant t_j$, więc

$$\begin{aligned}
\left| \int_a^b g(t) \sin pt dt \right| &\leqslant \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (g(t) - m_j) |\sin pt| dt + \sum_{j=1}^{n} |m_j| \left| \int_{t_{j-1}}^{t_j} \sin pt dt \right| \\
&\leqslant \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (M_j - m_j) dt + \frac{2}{p} \sum_{j=1}^{n} |m_j| \\
&= U(f, P) - L(f, P) + \frac{2}{p} \sum_{j=1}^{n} |m_j| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,
\end{aligned}$$

a to oznacza, że zachodzi pierwsza nierówność z tezy.

Wniosek 37. *Współczynniki Fouriera a_n i b_n funkcji całkowalnej dążą do zera przy $n \to \infty$.*

Aby zbadać zbieżność szeregu (94) w punkcie x musimy zbadać, jak zachowują się sumy częsciowe tego szeregu w tym punkcie. Mamy

$$\begin{aligned}
s_n(x) &= \frac{a_0}{2} + \sum_{m=1}^{n} (a_m \cos mx + b_m \sin mx) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(u) du + \sum_{m=1}^{n} \frac{1}{\pi} \int_{-\pi}^{\pi} f(u) (\cos mu \cos mx + \sin mu \sin mx) du \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} f(u) \left(\frac{1}{2} + \sum_{m=1}^{n} \cos m(u - x) \right) du.
\end{aligned}$$

Ponieważ

$$\begin{aligned}
\frac{1}{2} + \sum_{m=1}^{n} \cos m\alpha &= \frac{1}{2 \sin \frac{\alpha}{2}} \left(\sin \frac{\alpha}{2} + \sum_{m=1}^{n} (\sin(m + \frac{1}{2}\alpha) - \sin(m - \frac{1}{2})\alpha) \right) \\
&= \frac{\sin(2n+1)\frac{\alpha}{2}}{2 \sin \frac{\alpha}{2}},
\end{aligned}$$

dla $\alpha = u - x \neq 2k\pi, k \in \mathbb{Z}$, więc otrzymujemy

$$s_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(u) \frac{\sin(2n+1)\frac{u-x}{2}}{2 \sin \frac{u-x}{2}} du \tag{96}$$

(oczywiście $\frac{1}{2} + \sum_{m=1}^{n} \cos m2k\pi = \frac{1}{2} + n$ oraz $\lim_{\alpha \to 2k\pi} \frac{\sin(2n+1)\frac{\alpha}{2}}{2 \sin \frac{\alpha}{2}} = n + \frac{1}{2}$ dla $k \in \mathbb{Z}$).

Definicja 107. Całkę (96) nazywamy całką Dirichleta, a funkcję

$$D_n(u - x) = \frac{\sin(2n+1)\frac{u-x}{2}}{2 \sin \frac{u-x}{2}} \quad \text{dla} \quad u - x \neq 2k\pi$$

oraz

$$D_n(2k\pi) = \frac{1}{2} + n, \quad k \in \mathbb{Z},$$

jadręm Dirichleta.

Lemat 5. *Jeśli f jest funkcją 2π-okresową i R-całkowalną na przedziale $[-\pi, \pi]$, to*

$$\int_{\alpha}^{\alpha+2\pi} f(u)du = \int_{-\pi}^{\pi} f(u)du,$$

dla dowolnego $\alpha \in \mathbb{R}$.

Dowód. Z okresowości funkcji f wynika, że jest ona R-całkowalna na dowolnym przedziale i ponadto mamy

$$\int_{\alpha}^{\alpha+2\pi} f(u)du = \int_{\alpha}^{-\pi} f(u)du + \int_{-\pi}^{\pi} f(u)du + \int_{\pi}^{\alpha+2\pi} f(u)du.$$

Jeśli w ostatniej całce zamienimy zmienną podstawiając $u = t + 2\pi$, to otrzymamy

$$\begin{aligned}\int_{\alpha}^{\alpha+2\pi} f(u)du &= \int_{\alpha}^{-\pi} f(u)du + \int_{-\pi}^{\pi} f(u)du + \int_{-\pi}^{\alpha} f(t+2\pi)dt \\ &= \int_{\alpha}^{-\pi} f(u)du + \int_{-\pi}^{\pi} f(u)du - \int_{\alpha}^{-\pi} f(t)dt \\ &= \int_{-\pi}^{\pi} f(u)du.\end{aligned}$$

Stosując Lemat 5 możemy zapisać

$$s_n(x) = \frac{1}{\pi}\int_{x-\pi}^{x+\pi} f(u)\frac{\sin(2n+1)\frac{u-x}{2}}{2\sin\frac{u-x}{2}}du \quad (\alpha = x - \pi).$$

Zamieniając w całce zmienną przy pomocy wzoru $t = u - x$, otrzymujemy

$$s_n(x) = \frac{1}{\pi}\int_{\pi}^{\pi} f(x+t)\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}dt.$$

Następnie rozbijając te całkę na dwie $(\int_{-\pi}^{0} + \int_{0}^{\pi})$ i sprowadzając pierwszą z nich do całki po przedziale $[0, \pi]$ otrzymujemy następujący wzór na sumę częsciową szeregu Fouriera:

$$s_n(x) = \frac{1}{\pi}\int_{0}^{\pi} (f(x+t) + f(x-t))\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}dt \tag{97}$$

Badanie zbieżności szeregu Fouriera funkcji f w punkcie x sprowadza się więc do zbadania całki (97). Zauważmy, że funkcja podcałkowała w tej całce w ogóle nie ma granicy, gdy $n \to \infty$.

Twierdzenie 164. *(Zasada Lokalizacji) Jeśli f jest 2π-okresową funkcją $R-$ całkowalną na przedziale $[-\pi,\pi]$ oraz $0<\delta<\pi$, to*

$$\lim_{n\to\infty}\left(\int_{-\pi}^{\delta}+\int_{\delta}^{\pi}\right)\left(f(x+t)\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}\right)dt=0.$$

Dowód. Ustalmy punkt x i niech $\delta<|t|<\pi$, $g(t)=\frac{f(x+t)}{2\sin\frac{1}{2}t}$. Jest to funkcja 4π-okresowa i R-całkowalna, gdyż funkcja $\sin\frac{1}{2}t$ nie zeruje się na tych przedziałach. Na mocy Lematu Riemanna-Lebesgue'a całki:

$$\int_{-\pi}^{-\delta}g(t)\sin(n+\frac{1}{2})tdt,\qquad\int_{\delta}^{\pi}g(t)\sin(n+\frac{1}{2})tdt$$

dążą do zera przy $n\to\infty$.

Uwaga 61. Z powyższego twierdzenia wynika, że zachowanie się ciągu $(s_n(x))$ jeśli chodzi o zbieżność, zależy tylko od wartości funkcji f w jakimś dowolnie małym otoczeniu punktu x.
Ponadto badanie ciągu $(s_n(x))$ sprowadza się do zbadania całki

$$\frac{1}{\pi}\int_{-\delta}^{\delta}f(x+t)\frac{\sin(n+\frac{1}{2}t)}{2\sin\frac{1}{2}t}dt,$$

przy dowolnie małej liczbie $\delta>0$.

Udowodnimy teraz twierdzenie, które podaje warunki dostateczne zbieżności szeregu Fouriera. Inne tego typu twierdzenia można znaleźć w ksiażce [2], t. 3. Podamy wpierw następującą definicję.

Definicja 108. Funkcja f nazywamy się przedziałami lub kawałkami rózniczkowalna na przedziale $[a,b]$, jeśli przedział ten można podzielić na skończoną liczbę przedziałów domkniętych, wewnątrz których jest ona różniczkowalna, a na końcach ma granice i „pochodne jednostronne", to znaczy granice jednostronne ilorazów różnicowych, w których zastępujemy wartości funkcji odpowiednimi granicami jednostronnymi.

Twierdzenie 165. *Jeśli funkcja f jest kawałkami rózniczkowalna na przedziale $[-\pi,\pi]$, to jej szereg Fouriera jest zbieżny w każdym punkcie x i jego suma wynosi $s=\frac{f(x+0)+f(x-0)}{2}$. (Ta suma jest oczywiście równa $f(x)$, jeśli w tym punkcie funkcja f jest ciągła).*

Dowód. Podstawiając w równości (97) $f=1$ i uwzględniając, że dla tej funkcji $s_n(x)=1$ w każdym punkcie x (bowiem $a_0=2$ oraz $a_n=b_n=0$ dla $n\geqslant 1$), otrzymujemy

$$1=\frac{2}{\pi}\int_0^{\pi}\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}dt,\quad\text{czyli}\quad s=\frac{2s}{\pi}\int_0^{\pi}\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}dt.$$

Stąd

$$s_n(x) - s = \frac{1}{\pi}\int_0^\pi \left(f(x+t) + f(x-t) - f(x+0) - f(x-0)\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} \right) dt$$
$$= \frac{1}{\pi}\int_0^\pi g(t)\sin(n+\frac{1}{2})t dt.$$

gdzie $g(t) = \begin{cases} \left(\frac{f(x+t)-f(x+0)}{t} - \frac{f(x-t)-f(x-0)}{-t}\right)\frac{\frac{1}{2}t}{\sin\frac{1}{2}t} & \text{dla } 0 < t \leqslant \pi, \\ 0 & \text{dla } t = 0. \end{cases}$

Na mocy Lematu Riemanna-Lebesgue'a, aby zakończyć dowód wystarczy stwierdzić, że g jest funkcją R-całkowalną. Na przedziale $(0, \pi]$ funkcja g jest ciagła wszędzie z wyjątkiem tylko skończonej liczby punktów, w których funkcje $f(x+t)$ i $f(x-t)$ nie są ciagłe. Wystarczy zatem zbadać zachowanie się funkcji g przy $t \to 0^+$. Wykażemy, że istnieje skończona granica $\lim\limits_{t\to 0^+} g(t) = K$. To będzie oznaczało, że funkcja g jest R-całkowalna na przedziale $[0, \pi]$ jako funkcja ograniczona i ciągła, która posiada skończoną liczbę punktów nieciągłości (Tw. 79).

Niech wpierw x leży wewnątrz przedziału, na którym funkcja f jest różniczkowalna. Wówczas $f(x^+) = f(x^-) = f(x)$ oraz

$$\lim_{t\to 0^+} \frac{f(x+t) - f(x)}{t} = f'_+(x) = f'(x), \quad \lim_{t\to 0^+} \frac{f(x-t) - f(x)}{t} = f'_-(x) = f'(x),$$

a zatem w tym przypadku $\lim\limits_{t\to 0^+} g(t) = 0$.

Jeśli x jest punktem ciągłości funkcji f, w którym nie ma pochodnej (ale istnieją pochodne jednostronne), to

$$\lim_{t\to 0^+} g(t) = f'_+(x) - f'_-(x).$$

Jeśli w końcu x jest punktem nieciągłości funkcji f, to wiemy, że istnieją skończone granice

$$\lim_{t\to 0^+} \frac{f(x+t) - f(x^+)}{t}, \quad \lim_{t\to 0^+} \frac{f(x-t) - f(x^-)}{t},$$

zatem również w tym przypadku granica $\lim\limits_{t\to 0^+} g(t)$ istnieje i jest skończona.

W 1876 roku Du Bois-Reymond podał przykład funkcji ciągłej, o szeregu Fouriera rozbieżnym w pewnych punktach. Lebesgue w 1906 roku podał przykład funkcji ciągłej, do której szereg Fouriera jest zbieżny wszędzie, ale nie jest zbieżny jednostajnie (zob. [2], t. 3, s. 419-422). W roku 1966 matematyk szwedzki L. Carlson udowodnił, że jeśli f jest funkcją ciągłą i 2π-okresową, to jej szereg Fouriera jest do niej prawie wszędzie zbieżny, to znaczy wszędzie z wyjątkiem punktów, których miara Lebesgue'a jest równa zero.

Sytuacja ta ulega jednak zmianie, gdy zamiast sum częsciowych szeregu Fouriera rozważymy ciąg ich średnich arytmetycznych:

$$\sigma_n(x) = \frac{s_0(x) + s_1(x) + \cdots + s_n(x)}{n+1}, \quad n = 0, 1, 2, \ldots$$

(średnie Cesáro). Znajdziemy teraz reprezentację całkową średniej σ_n. Mamy

$$\sigma_n(x) = \frac{1}{2(n+1)\pi} \int_{-\pi}^{\pi} f(x+t) \left(\sum_{k=0}^{n} \frac{\sin(k + \frac{1}{2}t)}{\sin \frac{1}{2}t} \right) dt.$$

Ponieważ

$$\sum_{k=0}^{n} 2 \sin(k + \frac{1}{2})t \sin \frac{1}{2}t = -\sum_{k=0}^{n} \Big(\cos(k+1)t - \cos kt \Big)$$
$$= 1 - \cos(n+1)t = 2 \sin^2 \frac{n+1}{2} t,$$

więc

(98) $$\sigma_n(x) = \frac{1}{2(n+1)\pi} \int_{-\pi}^{\pi} f(x+t) \frac{\sin^2 \frac{1}{2}(n+1)t}{\sin^2 \frac{1}{2}t} dt.$$

Definicja 109. Całkę (98) nazywamy całką Fejéra. Ponadto oznaczamy

$$\kappa_n(t) = \frac{\sin^2 \frac{1}{2}(n+1)t}{2(n+1)\sin^2 \frac{1}{2}t}$$
$$\kappa_n(2l\pi) = \frac{1}{n+1} \sum_{k=0}^{n} \kappa_n(2l\pi) = \frac{1}{n+1} \sum_{k=0}^{n} \Big(\frac{1}{2} + k \Big) = \frac{1}{2}(n+1),$$

$n = 0, 1, 2, \ldots, l \in \mathbb{N}$, nazywamy jądrem Fejéra. Mamy więc

(99) $$\sigma_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x+t) \kappa_n(t) dt.$$

Zauważmy, ponadto że $\kappa_n(t) \geqslant 0$ dla wszystkich t.

Twierdzenie 166. *(Fejéra) Jeśli funkcja f jest ciągła i okresowa o okresie 2π i jeśli (σ_n) jest ciągiem średnich arytmetycznych sum częściowych jej szeregu Fouriera, to $\lim_{n\to\infty} \sigma_n(x) = f(x)$ jednostajnie na $\mathbb{R}$.*

Dowód. Zauważmy wpierw, że jeśli f jest funkcją tożsamościowo równa 1, to dla niej $\sigma_n(x) = 1$ dla wszystkich x, zatem na mocy wzoru (99) otrzymujemy

$$1 = \frac{1}{\pi} \int_{-\pi}^{\pi} k_n(t) dt.$$

Stąd

$$f(x) - \sigma_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi}(f(x) - f(x+t)\kappa_n(t))dt. \tag{100}$$

Ponieważ funkcja f jest ciągła na przedziale $[-\pi, \pi]$, zatem jest na nim jednostajnie ciągła. Również jest ona jednostajnie ciągła na przedziale $[0, 2\pi]$ (możemy przy tym założyć, że stała $\delta > 0$ występująca w warunku jednostajnej ciągłości, jest w obydwu przypadkach mniejsza od π). Stąd, wobec okresowości funkcji f wnioskujemy, że jest ona jednostajnie ciągła na $\mathbb{R}$.
Niech będzie dane $\varepsilon > 0$. Wybierzmy $\delta > 0$ tak, aby $|f(x) - f(x+t)| < \frac{\varepsilon}{3}$ dla wszystkich x i wszystkich t spełniających nierówność $|t| < \delta$. Możemy przy tym założyć, że $\delta < \pi$. Istnieje również $M > 0$ takie, że $|f(x)| \leqslant M$ dla wszystkich x, czyli $|f(x) - f(x+t)| \leqslant 2M$ dla dowolnych x i t. Całkę (100) rozdzielamy na trzy całki w następujący sposób:

$$\begin{aligned}
&\frac{1}{\pi}\int_{-\pi}^{\pi}((f(x) - f(x+t))\kappa_n(t))dt = \\
&= \frac{1}{\pi}\Big(\int_{-\pi}^{-\delta} + \int_{-\delta}^{\delta} + \int_{\delta}^{\pi}\Big)(f(x) - f(x+t))\kappa_n(t)dt \\
&= I_1 + I_2 + I_3.
\end{aligned}$$

Szacujemy kolejno wartości powyżej otrzymanych całek. Mamy

$$|I_1| = \frac{1}{\pi}\left|\int_{-\pi}^{-\delta}(f(x) - f(x+t))\kappa_n(t)dt\right| \leqslant \frac{2m}{\pi}\int_{-\pi}^{-\delta}\kappa_n(t)dt.$$

Ponieważ dla $t \in [-\pi, -\delta]$ prawdziwa jest nierówność

$$\kappa_n(t) = \frac{\sin^2\frac{1}{2}(n+1)t}{2(n+1)\sin^2\frac{1}{2}t} \leqslant \frac{1}{2(n+1)\sin^2\frac{1}{2}\delta},$$

więc

$$|I_1| \leqslant \frac{M}{(n+1)\pi\sin^2\frac{1}{2}\delta}\int_{-\pi}^{0}dt = \frac{M}{(n+1)\sin^2\frac{1}{2}\delta}$$

Analogicznie otrzemujemy

$$|I_3| \leqslant \frac{M}{(n+1)\pi\sin^2\frac{1}{2}\delta}\int_{-\pi}^{0}dt = \frac{M}{(n+1)\sin^2\frac{1}{2}\delta}$$

Natomiast

$$|I_2| = \frac{1}{\pi}\left|\int_{-\delta}^{\delta}(f(x) - f(x+t)\kappa_n(t))dt\right| \leqslant \frac{\varepsilon}{3\pi}\int_{-\pi}^{\pi}\kappa_n(t)dt = \frac{\varepsilon}{3}.$$

Otrzymujemy więc

$$|f(x) - \sigma_n(x)| < \frac{\varepsilon}{3} + \frac{2M}{(n+1)\sin^2\frac{1}{2}t}.$$

Biorąc n_0 tak duże, aby

$$\frac{2M}{(n+1)\sin^2 \frac{1}{2}t} < \frac{2\varepsilon}{3} \quad \text{dla} \quad n \geqslant n_0$$

otrzymujemy nierówność $|f(x) - \sigma_n(x)| < \varepsilon$ prawdziwą dla każdego $n \geqslant n_0$ i wszystkich $x \in \mathbb{R}$, co kończy dowód.

Wniosek 38. *Jeśli dwie funkcje f i g ciagłe i okresowe o okresie 2π mają ten sam szereg Fouriera, to $f(x) = g(x)$ dla dowolnego x.*

Wniosek 39. *Jeśli funkcja f jest ciągła i 2π okresowa oraz*

$$\int_{-\pi}^{\pi} f(x)dx = 0, \quad \int_{-\pi}^{\pi} f(x)\cos nx dx = \int_{-\pi}^{\pi} f(x)\sin nx dx = 0 \quad \textit{dla} \quad n \in \mathbb{N},$$

to $f(x) = 0$ dla każdego $x \in \mathbb{R}$.

Definicja 110. Funkcja postaci

$$\frac{A_0}{2} + \sum_{k=1}^{n}(A_K \cos kx + B_k \sin kx), \quad {A_n}^2 + {B_n}^2 > 0$$

nazywa się wielomianem trygonometrycznym stopnia n ($n = 0, 1, 2, \ldots$).

Wniosek 40. *(II Twierdzenie Aproksymacyjne Weierstrassa) Dla każdej funkcji ciągłej i 2π-okresowej istnieje ciąg wielomianów algebraicznych zbieżny do niej jednostajnie na $\mathbb{R}$.*

Wniosek 41. *(I Twierdzenie Aprosymacyjne Weierstrassa) Dla każdej funkcji ciągłej na przedziale $[a, b]$ istnieje ciąg wielomianów algebraicznych zbieżny do niej jednostajnie na tym przedziale.*

Dowód. Wystarczy pokazać, że dla każdego $\varepsilon > 0$ istnieje wielomian P taki, że $|f(x) - P(x)| < \varepsilon$ dla $x \in [a, b]$. Odwzorujemy przedział $[0, \pi]$ na przedział $[a, b]$ poprzez przekształcenie

$$x = a + \frac{b-a}{\pi}t, \quad t \in [0, \pi]$$

i defniujemy $f^*(t) = f(a + \frac{b-a}{\pi}t)$. Funkcja f^* jest określona na przedziale $[0, \pi]$. Przedłużamy ją do funkcji parzystej na przedzial $[-\pi, 0]$ to znaczy określamy $f^*(t) = f^*(-t)$ jeśli $t \in [-\pi, 0]$. Otrzymana w ten sposób funkcja f^* jest ciągła na przedziale $[-\pi, \pi]$ oraz $f^*(\pi) = f^*(-\pi)$. Posiada ona więc ciagłe i 2π-okresowe rozszerzenia na całą prostą $\mathbb{R}$. Na mocy Twierdzenia Fejéra (166) dla dowolnego $\varepsilon > 0$ istnieje wielomian trygonometryczny T taki, że

$$|f^*(t) - T(t)| < \frac{\varepsilon}{2} \quad \text{dla wszystkich} \quad t.$$

Ponieważ $\cos kt$ i $\sin kt$, $k \in \mathbb{N}$ rozwijają się w szeregi potęgowe zbieżne, a więc i wielomian trygonometryczny T rozwija się w szeregi potęgowe zbieżne na całej prostej, a zatem jednostajnie zbieżne na każdym przedziale domkniętym. Niech więc

$$T(t) = \sum_{n=0}^{\infty} c_k t^k.$$

Jeśli $P_n(t)$ są sumami częsciowymi tego szeregu, to na mocy jego jednostajnej zbieżności na przedziale $[-\pi, \pi]$ istnieje wkaźnik n_0, że dla $n \geqslant n_0$ i $t \in [-\pi, \pi]$ mamy

$$|T(t) - P_n(t)| < \frac{\varepsilon}{2}.$$

Biorąc $n = n_0$ i $P(t) = P_{n_0}(t)$ otrzymujemy

$$|f^*(t) - P(t)| \leqslant |f^*(t) - T(t)| + |T(t) - P(t)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

dla wszystkich $t \in [-\pi, \pi]$ (a w szczególności $t \in [0, \pi]$). Powracając do zmiennej x to znaczy podstawiając, otrzymujemy

$$|f(x) - P(\pi \frac{x-a}{b-a})| < \varepsilon \quad \text{dla} \quad x \in [a, b],$$

gdzie $P(\pi \frac{x-a}{b-a})$ jest oczywiście wielomianem zmiennej x.

Twierdzenie 167. *(Własność Minimum Współczynników Fouriera) Jeśli* $s_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n}(a_k \cos(kx) + b_k \sin(kx))$ *jest* n*-tą sumą częsciową szeregu Fouriera funkcji* f *R-całkowalnej na przedziale* $[-\pi, \pi]$ *i* $T_n(x) = \frac{A_0}{2} + \sum_{k=1}^{n}(A_k \cos(kx) + B_k \sin(kx))$ *jest wielomianem stopnia nie większego niż* n*, to*

$$\int_{-\pi}^{\pi} |f(x) - s_n(x)|^2 dx \leqslant \int_{-\pi}^{\pi} |f(x) - T_n(x)|^2 dx$$

i równośc ma miejsce wtedy i tylko wtedy, gdy

$$a_0 = A_0, \quad a_k = A_k, \quad b_k = B_k \quad \textit{dla} \quad k \in \mathbb{N}.$$

Dowód. ...

Wniosek 42. *(Nierówność Bessela) Jeśli* a_0*,* a_n *i* b_n *(*$n \in \mathbb{N}$*) są współczynnikami Fouriera funkcji R-całkowalnej na przedziale* $[-\pi, \pi]$*, to prawdziwa jest nierówność*

$$\frac{|a_0|}{2} + \sum_{n=1}^{\infty}(|a_n|^2 + |b_n|^2) \leqslant \frac{1}{\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx.$$

Dowód. Wiemy, że

$$0 \leqslant \int_{-\pi}^{\pi} |f(x) - s_m(x)|^2 dx = \int_{-\pi}^{\pi} |f(x)|^2 dx - \pi \Big\{ \frac{|a_0|}{2} + \sum_{n=1}^{\infty} (|a_n|^2 + |b_n|^2) \Big\}$$

dla $m \in \mathbb{N}$. Stąd otrzymujemy

$$\frac{|a_0|}{2} + \sum_{n=1}^{m} (|a_n|^2 + |b_n|^2) \leqslant \frac{1}{\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx.$$

a następnie przechodzimy do granicy przy $m \to \infty$.

Twierdzenie 168. *(Tożsamość Parsevala) Jeśli f jest funkcją ciągła i 2π-okresową a (s_n) jest ciągiem sum częsciowych jej szeregu Fouriera, to*

$$\lim_{n \to \infty} \int_{-\pi}^{\pi} |f(x) - s_n(x)|^2 dx = 0$$

oraz zachodzi równość (zwana Tożsamościa Parsevala)

$$\frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} (|a_n|^2 + |b_n|^2) = \frac{1}{\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx.$$

Dowód. Niech będzie dana liczba $\varepsilon > 0$. Z Twierdzenia Fejéra wynika, że istnieje wskaźnik n_0 taki, że dla $n \geqslant n_0$ i wszystkich x (rzeczywistych lub zespolonych) mamy

$$|f(x) - \sigma_n(x)| < \frac{\sqrt{\varepsilon}}{\sqrt{2\pi}}.$$

Z własności Minimum Współczynników Fouriera wynika, że

$$\int_{-\pi}^{\pi} |f(x) - s_n(x)|^2 dx \leqslant \int_{-\pi}^{\pi} |f(x) - \sigma_n(x)|^2 dx \leqslant \varepsilon$$

dla $n > n_0$, a zatem

$$\lim_{n \to \infty} \int_{-\pi}^{\pi} |f(x) - s_n(x)|^2 dx = 0$$

Ponieważ

$$\int_{-\pi}^{\pi} |f(x)|^2 dx - \pi \Big(\frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} (|a_n|^2 + |b_n|^2) \Big) = \int_{-\pi}^{\pi} |f(x) - s_m(x)|^2 dx,$$

więc przechodząc do granicy przy $m \to \infty$ otrzymujemy drugą część tezy.

11 Przestrzenie metryczne

11.1 Pojęcie przestrzeni metrycznej. Zbiory w przestrzeniach metrycznych

Definicja 111. Niepusty zbiór X nazywamy przestrzenią metryczną, jeśli określona jest w nim funkcja $d : X \times X \mapsto \mathbb{R}_+ \cup \{0\}$ spełniająca następujące warunki:

(a) $d(x, y) = 0$ wtedy i tylko wtedy, gdy $x = y$;

(b) $d(x, y) = d(y, x)$ — symetria;

(c) $d(x, y) \leqslant d(x, z) + d(z, y)$ — Nierówność Trójkąta.

Warunki (b) i (c) muszą być spełnione dla dowolnych $x, y, z \in X$.
Elementy przestrzeni metrycznej nazywać będziemy punktami,funkcję d — metryką lub odległością.
Przestrzeń metryczna X z metryką d oznaczamy symbolem (X, d).

Przykład 48.

(a) Zbiór $\mathbb{R}$ oraz zbiór $\mathbb{C}$ są przestrzeniami metrycznymi z metryką określoną następująco: $d(x, y) = |x - y|$ dla $x, y \in \mathbb{R}$ (lub $\mathbb{C}$). Nierówność Trójkąta w tym przypadku wynika z Tw. 12 lub z Tw. 17 (f).

(b) Przy dowolniej liczbie naturalnej k niech $\mathbb{R}^k$ będzie zbiorem uporządkowanych k-wyrazowych ciągów

$$x = (x_1, x_2, \dots, x_k),$$

gdzie $x_1, x_2, \dots, x_k$ są liczbami rzeczywistymim zwanymi współrzędnymi elementu x. Elementy $\mathbb{R}^k$ nazywamy często wektorami, szczególnie w przypadku $k > 1$.
Jeżeli $y = (y_1, y_2, \dots, y_k)$ i $\alpha \in \mathbb{R}$, to definiujemy

$$x + y = (x_1 + y_1, \dots, x_k + y_k) \quad \alpha x = (\alpha x_1, \dots, \alpha x_k).$$

Oczywiście $x + y \in \mathbb{R}^k$ oraz $\alpha x \in \mathbb{R}^k$. Powyższe wzory definiują dodawanie wektorów oraz mnożenie wektora przez liczbe rzeczywistą (skalar). Obie te operacje spełniają prawa przemienności i łączności oraz rozdzielność mnożenia względem dodawania. Ponadto istnieje elementy neutralny dodawania $0 = (0, 0, \dots, 0)$. zwany wektorem zerowym. Każdy wektor $x \in \mathbb{R}^k$ posiada element przeciwny oraz $1x = x$ dla każdego $x \in \mathbb{R}^k$. Zbiór $\mathbb{R}^k$ z tak określonymi działaniami jest więc przestrzenią liniową nad ciałem

liczb rzeczywistych.
W zbiorze $\mathbb{R}^k$ definiujemy

$$d_{\mathbb{R}^k}(x,y) = \left(\sum_{j=1}^{k} (x_j - y_j)^2 \right)^{1/2},$$

gdzie $x = (x_1, \ldots, x_k)$, $y = (y_1, \ldots, y_k) \in \mathbb{R}^k$.
Wobec nierówności Schwarza dla $x, y, z \in \mathbb{R}^k$ otrzymujemy

$$\begin{aligned}
\sum_{j=1}^{k} (x_j + y_j)^2 &= \sum_{j=1}^{k} x_j^2 + 2\sum_{j=1}^{k} x_j y_j + \sum_{j=1}^{k} y_j^2 \\
&\leqslant \sum_{j=1}^{k} x_j^2 + 2\left(\sum_{j=1}^{k} x_j^2\right)^{\frac{1}{2}} \left(\sum_{j=1}^{k} y_j^2\right)^{\frac{1}{2}} + \sum_{j=1}^{k} y_j^2 \\
&= \left(\Big(\sum_{j=1}^{k} x_j^2\Big)^{\frac{1}{2}} + \Big(\sum_{j=1}^{k} y_j^2\Big)^{\frac{1}{2}} \right)^2
\end{aligned}$$

Powyższa nierównośc nazywa się Nierównościa Minkowskiego. Wobec nierówności Minkowskiego jest jasne, że powyżej określona funkcja d spełnia nierówność Trójkata. Oczywiście spełnia ona również warunki (a) i (b) Definicji 111, a zatem $(\mathbb{R}^k, d)$ jest przestrzenią metryczną.

(c) Niepusty podzbiór Y przestrzeni metrycznej X jest również przestrzenią metryczną z tą samą metryką.

(d) Dowolny niepusty zbiór X wraz z funkcją $d : X \times X \mapsto \mathbb{R}_+ \cup \{0\}$ określonej następująco: $d(x,y) = 0$, gdy $x = y$, $d(x,y) = 1$, gdy $x \neq y$ jest przestrzenią metryczną.

Definicja 112. Niech X będzie przestrzenią metryczną z metryką d. Wszystkie występujące poniżej punkty i zbiory są elementami i podzbiorami X.

(a) Kulą otwartą o środku w punkcie x i promieniu $r > 0$ nazywamy zbiór wszystkich punktów y takich, że $d(x,y) < r$ (oznaczamy ją symbolem $B(x,r)$).

(b) Otoczeniem punktu x nazywamy każdy zbiór U, który zawiera pewną kulę o środku w punkcie x i promieniu $\varepsilon > 0$.

(c) Punkt x nazywamy punktem wewnętrznym zbioru E, jeśli istnieje otoczenie U punktu x takie, że $U \subset E$.

(d) Zbiór E nazywamy otwartym, jeśli każdy punkt zbioru E jest jego punktem wewnętrznym lub jeśli E jest zbiorem pustym.

(e) Dopełnienie zbioru E nazywamy zbiór wszystkich punktów $x \in X$ takich, że $x \notin E$ (oznaczamy je symbolem E^c).

(f) Zbiór E nazywamy domkniętym, jeśli jego dopełnienie jest zbiorem otwartym.

(g) Punkt x nazywamy punktem skupienia zbioru E, jeśli każde otoczenie punktu x zawiera punkt $y \neq x$ taki, że $y \in E$.

(h) Jeśli $x \in E$ i x nie jest punktem skupienia zbioru E, to x nazywamy punktem izolowanym zbioru E.

(i) Zbiór E nazywamy doskonałym, jeśli jest domknięty i każdy punkt zbioru E jest jego punktem skupienia.

(j) Zbiór E jest ograniczony, jeśli zawiera się on w pewnej kuli (inaczej mówiąc gdy istnieje liczba $M > 0$ i punkt $y \in X$ takie, że $d(x, y) \leqslant M$ dla wszystkich $x \in E$).

(k) Zbiór E nazywamy gęstym w X, jeśli każdy punkt zbioru X jest punktem skupienia zbioru E lub należy do E.

Przykład 49.

(a) Kula otwarta w przestrzeni metrycznej X jest zbiorem otwartym. Weźmy kulę $B(x_0, r) \subset X$ i niech $x \in B(x_0, r)$. Niech $\varepsilon = \frac{1}{2}(r - d(x, x_0))$. Wówczas z Nierówności Trójkąta wynika, że $B(x, \varepsilon) \subset B(x_0, r)$.

(b) W przestrzeni $\mathbb{R}$ przedział otwarty jest zbiorem otwartym. Ten sam przedział otwarty rozważany jako podzbiór przestrzeni $\mathbb{R}^2$ nie jest zbiorem otwartym, bowiem nie zawiera żadnej kuli otwartej w $\mathbb{R}^2$.

(c) Cała przestrzeń metryczna X jest jednocześnie zbiorem otwartym i domkniętym ($X^c = \emptyset$). Zbiór pusty jest również zbiorem domkniętym ($\emptyset^c = X$).

(d) Przedział domknięty $[a, b]$ w przestrzeni $\mathbb{R}$ jest zbiorem domkniętym, bowiem jego dopełnienie jest zbiorem otwartym. Ten sam przedział domknięty rozważany jako podzbiór przestrzeni $\mathbb{R}^2$ jest również zbiorem domkniętym. Przedział otwarty (a, b) rozważany jako podzbiór przestrzeni $\mathbb{R}^2$ nie jest zbiorem domkniętym.

Twierdzenie 169. *Niech X będzie przestrzenią metryczną z metryką d. Wszystkie występujące poniżej zbiory są podzbiorami przestrzeni X.*

(a) Dla rodziny $\{G_\alpha\}$ zbiorów otwartych, zbiór $\bigcup_\alpha G_\alpha$ jest również otwarty.

(b) Jeśli zbiory $G_1, \ldots, G_n$ są otwarte, to zbiór $\bigcap\limits_{j=1}^{n} G_j$ jest otwarty.

(c) Jeśli $\{F_\alpha\}$ jest rodziną zbiorów domniętych, to $\bigcap\limits_{\alpha} F_\alpha$ jest także domknięty.

(d) Jeśli zbiory $F_1, \ldots, F_n$ są domkniętę, to ich suma $\bigcup\limits_{j=1}^{n} F_j$ jest także zbiorem domkniętym.

Dowód.

(a) Niech $G = \bigcup\limits_{\alpha} G_\alpha$. Jeśli $x \in G$, to $x \in G_\alpha$ przy pewnym α. Ponieważ x jest punktem wewnętrznym zbioru G_α, wieć jest on również punktem wewnętrznym zbioru G, bowiem $G_\alpha \subset G$, czyli zbiór G jest otwarty.

(b) Niech $H = \bigcap_{j=1}^{n} G_j$ i niech $x \in H$. Istnieją wówczas kule B_j o środku w punkcie x takie, że $B_j \subset G_j$ dla $j = 1, \ldots, n$. Niech r_j oznacza promień kuli B_j. Oznaczmy $r = \min r_1, \ldots, r_j$ i niech B będzie kulą domkniętą o środku w punkcie x i promieniu r. Wówczas $B \subset B_j \subset G_j$ dla $1 \leqslant j \leqslant n$. Stąd $B \subset H$, co oznacza, że zbiór H jest otwarty.

(c) Mamy $\left(\bigcap\limits_{\alpha} F_\alpha\right)^c = \bigcup\limits_{\alpha} F_\alpha^c$. Ponieważ zbiory F_α^c są otwarte, zatem na podstawie punktu (a) otwarty jest również zbiór $\bigcap\limits_{\alpha} F_\alpha^c$, a więc $\bigcap\limits_{\alpha} F_\alpha$ jest domknięty.

(d) Wystarczy ponownie przejść do dopełnień i zastosować punkt (b).

Przykład 50. W częsciach (b) i (d) powyższego twierdzenia założenie o skończoności rodzin G_j i F_j jest koniecznie. Istotnie niech $G_n = (-\frac{1}{n}, \frac{1}{n})$. Wówczas G_n jest otwartym podzbiorem przestrzeni $\mathbb{R}$, ale $G = \bigcap\limits_{n=1}^{\infty} G_n = \{0\}$. Przekrój nieskończonej rodziny zbiorów otwartych nie musi być otwarty. Analogicznie niech $F_n = [\frac{1}{n}, 1 - \frac{1}{n})$ dla $n = 2, 3, \ldots$. Wówczas zbiór F_n jest domknięty, ale $F = \bigcup\limits_{n=2}^{\infty} F_n = (0, 1)$ nie jest zbiorem domkniętym.

Twierdzenie 170. *Podzbiór E przestrzeni metrycznej (X, d) jest domknięty wtedy i tylko wtedy, gdy E zawiera wszystkie swoje punkty skupienia.*

Dowód. Niech zbiór E będzie domknięty i niech x będzie punktem skupienia zbioru E. Punkt $x \in E$, bowiem w przeciwnym razie istniałoby takie otoczenie U tego punktu, że $E \cap U = \emptyset$, co jest niemożliwe, bowiem x jest punktem skupienia zbioru E. Na odwrót, załóżmy, że zbiór E zawiera wszystkie swoje punkty skupienia. Wówczas jeśli $x \in E^c$, to $x \notin E$, czyli x nie jest punktem skupienia zbioru E. Istnieje więc otoczenie V punktu x takie, że $V \cap E = \emptyset$. Stąd wynika, że $V \subset E^c$. Wobec tego punkt x jest punktem wewnętrznym zbioru E^c, co oznacza, że jest on otwarty, a więc zbiór E jest domknięty.

Lemat 6. *Jeśli* E jest niepustym, domkniętym podzbiorem zbioru $\mathbb{R}$, ograniczonym z góry, to $y = \sup E \in E$ (analogicznie dla niepustych zbiorów domkniętych, ograniczonych z dołu oraz ich infimów).

Dowód. Przypuśćmy, że $y \notin E$. Dla każdego $h > 0$ istnieje punkt $x \in E$ taki, że $y - h \leqslant x \leqslant y$, gdyż w przeciwnym razie liczba $y - h$ byłaby kresem górnym zbioru E. Każde otoczenie punktu y zawiera więc pewien punkt x ze zbioru E, przy czym $x \neq y$, ponieważ $y \notin E$. To oznacza, że y jest punktem skupienia zbioru E nienależącym do E, czyli wobec Twierdzenia 170 zbiór E nie jest domnięty. Sprzeczność.

Definicja 113. Pokryciem otwartym zbioru E przestrzeni metrycznej X nazywamy rodzinę $\{G_\alpha\}$ podzbiorów otwartych przestrzeni X, dla której

$$E \subset \bigcup_{\alpha} G_\alpha$$

Definicja 114. Podzbiór K przestrzeni metrycznej X nazywamy zwartym, jeśli każde pokrycie otwarte zbioru K zawiera podpokrycie skończone, to znaczy, jeśli $\{G_\alpha\}$ jest pokryciem otwartym zbioru K, to istnieje skończona liczba wskaźników $\alpha_1, \ldots, \alpha_n$ takich, że

$$K \subset G_{\alpha_1} \cup \ldots G_{\alpha_n}.$$

Jest oczywiste, że każdy zbiór skończony jest zwarty.

Twierdzenie 171. *Z*warty podzbiór przestrzeni metrycznej jest ograniczony i domknięty.

Dowód. Niech K będzie zbiorem zwartym. Rozważmy rodzinę kul otwartych $\{B(x,1)\}_{x \in K}$. Ponieważ jest ona pokryciem otwartym zbioru K, więc istnieją punkty $x_1, \ldots, x_k \in K$ takie, że $K \subset B(x_1, 1) \cup \ldots \cup B(x_n, 1)$. Niech $r = \max\{1 + d(x_1, x_2), \ldots, 1 + d(x_1, x_n))\}$. Zauważmy, że $K \in B(x_1, r)$. Istotnie, jeśli $y \in K$, to $d(x_j, y) < 1$ dla pewnego $j \in \{1, \ldots, n\}$. Stąd

$$d(x_1, y) \leqslant d(x_1, x_j) + d(x_j, y) < 1 + d(x_1, x_j) \leqslant r,$$

zatem zbiór K jest ograniczony.
Pokażemy teraz, że dopełnienie zbioru K jest otwartym podzbiorem przestrzeni X. Niech $x \in K^c$. Dla dowolnego punktu $y \in K$ niech V_x^y i W_y będą kulami otwarymi odpowiednio o środkach w punktach x i y i o promieniach mniejszych niż $\frac{1}{2}d(x, y)$. Ponieważ K jest zbiorem zwartym, więc istnieje skończona liczba punktów $y_1, \ldots, y_n$ należacych do K i takich, że $K \subset W_{y_1} \cup \ldots \cup W_{y_n} = W$.

Jeśli $V = V^x_{y_1} \cap \ldots \cap V^x_{y_n}$, to V jest otoczeniem punktu x nie przecinającym się ze zbiorem W. Stąd $V \subset K^c$, czyli x jest punktem wewnętrznym zbioru K^c.

Twierdzenie 172. *Domknięty podzbiór zbioru zwartego jest zbiorem zwartym.*

Dowód. Załóżmy, że $F \subset K$ i niech F będzie zbiorem domkniętym, a K - zwartym. Niech V_α będzie otwartym pokryciem zbioru F. Jeśli do rodziny $\{V_\alpha\}$ dodamy zbiór F^c, to otrzymamy otwarte pokrycie zbioru K. Ponieważ zbiór K jest zwarty, więc istnieje skończona podrodzina $\{V_{\alpha_1}, \ldots, V_{\alpha_n}, F^c\}$, stanowiąca jego pokrycie. Odrzucając F^c, otrzymujemy otwarte i skończone pokrycie zbioru F.

Wniosek 43. *Jeśli zbiór F jest domknięty, a K — zwarty, to zwarty jest również zbiór $F \cap K$.*

Twierdzenie 173. *Niech K będzie zbiorem zwartym. Wówczas każdy jego nieskończony podzbiór ma punkt skupienia należacy do K.*

Dowód. Niech K będzie zbiorem zwartym i niech E będzie nieskończonym podzbiorem zbioru K. Gdyby żaden punkt zbioru K nie był punktem skupienia zbioru E, to każdy punkt $y \in K$ miałby otoczenie V_y, zawierające nie więcej niż jeden punkt zbioru E (właśnie punkt y, jeśli $y \in E$). Oczywiście żadna skończona podrodzina rodziny $\{V_y\}$ nie może pokryć zbioru E. Jest to także prawdziwe dla K, ponieważ $E \subset K$. Przeczy to jednak zwartości zbioru K.

Można również udowodnić twierdzenie odwrotne do powyższego (dla przestrzeni metrycznych!), mianowicie

Twierdzenie 174. *Jeśli każdy nieskończony i przeliczalny podzbiór zbioru K zawartego w przestrzeni metrycznej (X, d) ma punkt skupienia należacy do K, to K jest zbiorem zwartym.*

Dowód powyższego twierdzenia można znaleźć w książce K. Kuratowskiego, Wstęp do teorii mnogości i topologii. PWN, W-wa, 1980, s. 182–184.

Definicja 115. Jeśli $a_i, b_i \in \mathbb{R}$, $a_i < b_i$, $i = 1, \ldots, k$, to zbiór wszystkich punktów $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$ takich, że $a_i \leqslant x_i \leqslant b_i$, $i = 1, \ldots, k$, nazywamy kostką k-wymiarową (por. Def. 38).

Twierdzenie 175. *Kostka k-wymiarowa jest zwartym podzbiorem przestrzeni $\mathbb{R}^k$.*

Dowód. Porównaj z Uw. 11.

Twierdzenie 176. *(Heine-Borel) Podzbiór E przestrzeni $\mathbb{R}^k$ jest zwarty wtedy i tylko wtedy, gdy jest on domknięty i ograniczony.*

Dowód. Jeśli zbiór E jest zwarty, to na mocy Twierdzenia 171 jest on domknięty i ograniczony. Na odwrót, jeśli E jest domknięty i ograniczony, to zawiera się on w pewnej kostce, a zatem na mocy Twierdzenia 175 jest on zwarty, jako domknięty podzbiór zbioru zwartego.

Twierdzenie 177. *(Bolzano - Weierstrass) Nieskończony i ograniczony podzbiór przestrzeni R^k posiada punkt skupienia w R^k.*

Dowód. Ponieważ zbiór E jest ograniczony, więc jest on podzbiorem pewnej kostki k-wymiarowej. Z Twierdzenia 173 wynika, że ma on punkt skupienia należacy do tej kostki.

Lemat 7. *Każda przestrzeń metryczna zwarta jest ośrodkowa, to znaczy zawiera przeliczalny zbiór gęsty.*

Dowód powyższego lematu mozna znaleźć w książce K. Kuratowskiego, Wstęp do teorii mnogości i topologii, PWN W-wa, 1980, s. 183.

Definicja 116. Podzbiór E przestrzeni metrycznej X nazywamy spójnym, jeśli nie istnieją dwa otwarte podzbiory A i B przestrzeni X takie, że przekrój $A \cap B$ jest pusty, iloczyn $A \cap E$ i $B \cap E$ są niepuste oraz $E \subset A \cup B$.

Twierdzenie 178. *P*odzbiór E prostej rzeczywistej $\mathbb{R}$ jest spójny wtedy i tylko wtedy, gdy posiada on następującą własność: jeśli $x \in E$, $y \in E$ i $x < z < y$, to $z \in E$.

Dowód. Przypuścmy, że ten warunek nie jest spełniony. Wówczas dla pewnych punktów x, y, z mamy $x \in E$, $y \in E$, $x < z < y$ i $z \notin E$. Jeśli oznaczymy $A = \{\alpha : \alpha < z\}$ i $B = \{\beta : \beta > z\}$, to zbiory A i B są otwarte oraz $A \cap E \neq \emptyset$, $B \cup E \neq \emptyset$, $E \subset A \cup B$, $A \cap B = \emptyset$, a więc zbiór E nie jest spójny.
Załóżmy teraz, że zbiór E nie jet spójny. Istnieją wówczas punkty $x \in E$, $y \in E$, $x < y$ i rozłączne zbiory otwarte A i B takie, że $x \in A$, $y \in B$, i $E \subset A \cup B$. Niech $S = A \cap [x, y]$ i $z = \sup S$. Ponieważ $y \in B$ i zbiór B jest otwaty, więc $z < y$. Gdyby więc $z \in A$, to z tego, że A jest zbiorem otwartym wynika, że z nie jest kresem górnym zbioru S. Stąd $z \notin A$. Ponieważ $x \in A$, więc $x < z$. Następnie, gdyby $z \in B$, to z otwartości zbioru B wynikałoby, że z nie byłby kresem górnym zbioru S, a więc $z \notin B$. Ponieważ jednak $E \subset A \cup B$, zatem $z \notin E$, a więc warunek występujący w tezie twierdzenia nie jest spełniony.

Poniższe twierdzenie charakteryzuje zbiory spójne w $\mathbb{R}^k$. Jego dowód można znaleźć na przykład w książce [6], cz. I, s. 114-115.

Twierdzenie 179. *Niepusty zbiór otwarty $A \subset \mathbb{R}^k$ jest spójny wtedy i tylko wtedy, gdy dowolne dwa jego punkty można połączyć linią łamaną zawartą całkowicie w $\mathbb{R}^k$.*

Powyższe twierdzenie dla zbiorów domniętych nie jest prawdziwe. Wystarczy rozważyć dowolny okrąg zawarty w $\mathbb{R}^2$.

11.2 Pojęcie granicy i funkcji ciągłej w przestrzeni metrycznej

Definicja 117. Ciąg (p_n) w przestrzeni metrycznej X z metryką d nazywamy zbieżnym do punktu $p \in X$, jeśli posiada on następującą własność:
Dla każdego $\varepsilon > 0$ istnieje taka liczba naturalna n_0, że dla wszystkich $n \geqslant n_0$ spełniona jest nierówność $d(p_n, p) < \varepsilon$. W tym przypadku będziemy mówili, że p jest granicą ciągu (p_n) i będziemy pisali $p_n \to p$ lub $\lim_{n\to\infty} p_n = p$. Mówimy, że ciąg (p_n) jest zbieżny, jeśli jest zbieżny do jakiegoś punktu przestrzeni X. Jeśli ciąg (p_n) nie jest zbieżny, to mówimy, że jest on rozbieżny.

W przypadku przestrzeni metrycznych można również udowodnić (rozumowanie jest analogiczne) odpowiednik Twierdzenia 23, opisujący właności ciągów zbieżnych (z wyjątkiem punktu (e)).
Dla ciągów przestrzeni $\mathbb{R}^k$ możemy zbadać związki pomiędzy zbieżnościa, a operacjami algebraicznymi. Również w tym przypadku prawdziwy jest odpowiednik Twierdzenia 24 (a), (b), (c). Udowodnia się go w oparciu o Twierdzenie 24 i następujący odpowiednik Twierdzenia 25 charakteryzujący zbieżność ciągów w $\mathbb{R}^k$.

Twierdzenie 180. *Załóżmy, że $x_n \in \mathbb{R}^k$, $x_n = (\xi_1^n, \ldots, x_k^n)$, $n \in \mathbb{N}$. Ciąg (x_n) jest zbieżny do punktu $x = (\xi^1, \ldots, \xi^k)$ wtedy i tylko wtedy, gdy $\lim_{n\to\infty} \xi_n^j = \xi^j$ dla $j = 1, \ldots, k$.*

Dowód powyższego twierdzenia jest analogiczny do dowodu Twierdzenia 25.

W oparciu o twierdzenia 173 i 174 można podać następującą definicję zbioru zwartego w przestrzeni metrycznej.

Definicja 118. Podzbiór K przestrzeni metrycznej X nazywamy zwartym, jeśli z każdego ciągu (p_n) elementów tego zbioru można wyrwać podciąg (p_{n_k}) zbieżny do pewnego elementu $p \in K$.

Ponadto powtarzając dowód Twierdzenia 30 i wykorzystując Twierdzenie 171 otrzymujemy następujący odpowiednik Twierdzenia Bolzano-Weierstrassa dla $\mathbb{R}^k$.

Twierdzenie 181. *Każdy ograniczony ciąg w przestrzeni $\mathbb{R}^k$ zawiera podciąg zbieżny.*

Wprowadzimy teraz ogólną definicję granicy odwzorowania między dwiema przestrzeniami metrycznymi.

Definicja 119. Niech (X, d_X), (Y, d_Y) będą przestrzeniami metrycznymi, $E \subset$

X, $f : E \mapsto Y$, x_0 niech będzie punktem skupienia zbioru E oraz niech $y_0 \in Y$. Mówimy, że odwzorowanie f ma w punkcie x_0 granicę y_0, jeśli dla dowolnego $\varepsilon > 0$ istnieje $\delta > 0$ takie, że

$$d_Y(f(x), y_0) < \varepsilon$$

dla wszystkich punktów $x \in E$, dla których

$$0 < d_X(x, x_0) < \delta.$$

Będziemy wówczas pisali $f(x) \to y_0$ przy $x \to x_0$ lub $\lim\limits_{x \to x_0} = y_0$.

Powyższa definicja nosi nazwę definicji granicy funkcji w sensie Cauchy'ego. Również w tym przypadku (to znaczy jeśli X i Y są przestrzeniami metrycznymi) prawdziwy jest odpowiednik Twierdzenia 35, który orzeka równoważnośc definicji Cauchy'ego oraz definicji Heinego granicy funkcji w punkcie, która jest sformułowana w języku ciągów. Ponadto granica odwzorowania w punkcie jest wyznaczona jednoznacznie (por. Wn. 6). Prawdziwy jest również odpowiednik Twierdzenia 37 dotyczący własności arytmetycznych granicy w przypadku gdy X jest przestrzenią metryczną. natomiast $Y = \mathbb{C}$ (punkt (a) tego twierdzenia jest oczywiście prawdziwy dla $Y = \mathbb{R}^k$).

Definicja 120. Niech (X, d_X), (Y, d_Y) będą przestrzeniami metrycznymi $E \subset X$, $x_0 \in E$ i $f : E \mapsto Y$. Mówimy, że odwzorowanie f jest ciągłe w punkcie x_0 jeśli dla dowolnego $\varepsilon > 0$ istnieje $\delta > 0$ takie, że

$$d_Y(f(x), f(x_0)) < \varepsilon$$

dla wszystkich punktów $x \in E$, dla których

$$d_X(x, x_0) < \delta.$$

Jeśli funkcja f jest ciągła w każdym punkcie zbioru E, to f nazywamy ciągłą na E.

Zauważmy, że aby funkcja f była ciągła w punkcie x_0, to musi być określona w tym punkcie (por. Def. 119).
Można w szczególności udowodnić, że metryka $d : (X, d) \times (X, d) \mapsto \mathbb{R}$ jest funkcją ciągłą, to znaczy jeżeli $\lim\limits_{n\to\infty} x_n = x$, $\lim\limits_{n\to\infty} y_n = y$ ($x_n, y_n, x, y \in X$ dla $n \in \mathbb{N}$), to $\lim\limits_{n\to\infty} d(x_n, y_n) = d(x, y)$ (zob. [3], s. 65).
W sytuacji opisanej w Definicji 120 prawdziwe pozostają odpowiedniki Uwagi 24 (ciągłości funkcji w punkcie izolowanym), Twierdzenia 41 (warunek równoważny ciągłości odwzorowania f w punkcie x_0 będacym punktem skupienia zbioru E) oraz Twierdzenia 42 dla przypadku, gdy X jest przestrzenią metryczną natomiast $Y = \mathbb{C}$ (przypadek sumy dwóch odwzorowań jest oczywiście prawdziwy dla $Y = \mathbb{R}^k$).
Twierdzenie 43 dla odwzorowań o wartościach w $\mathbb{R}^k$ przybiera następującą postać.

Twierdzenie 182. *Dla funkcji rzeczywistych $f^1, \ldots, f^k$ określonych na przestrzeni metrycznej X zdefiniujemy odwzorowanie $f : X \mapsto \mathbb{R}^k$ wzorem*

$$f(x) = (f^1(x), \ldots, f^k(x)), \quad x \in X.$$

Wówczas odwzorowanie f jest ciągłe wtedy i tylko wtedy, gdy każda z funkcji $f^1, \ldots, f^k$ jest ciągła.

Dowód powyższego twierdzenia jest analogiczny do dowodu Twierdzenia 43. Przejdźmy teraz do funkcji złożonych. Sformułujemy odpowiednik Twierdzenia 44 (tylko punkt (c) o złożeniu funkcji ciągłych).

Twierdzenie 183. *Niech (X, d_x), (Y, d_Y) będą przestrzeniami metrycznymi, $f : X \mapsto Y$, $g : Y \mapsto Z$ i $h : X \mapsto Z$ będzie funkcją złożoną $h = g \circ f$. Jeśli funkcja f jest ciągła w punkcie $x_0 \in X$, a g jest ciągła w punkcie $f(x_0)$, to funckja h jest również ciągła w punkcie x_0.*

Dowód powyższego twierdzenia jest analogiczny do dowodu Twierdzenia 44.

Twierdzenie 184. *Odwzorowanie f przestrzeni metrycznej X w przestrzeń metryczną Y jest ciągłe na X wtedy i tylko wtedy gdy zbiór $f^{-1}(V)$ jest otwarty w X dla dowolnego zbioru otwartego V w Y.*

Dowód. Przypuśćmy, że f jest ciągłe na X. Niech V będzie zbiorem otwartym w Y. Musimy udowodnić, że każdy punkt zbioru $f^{-1}(V)$ jest jego punktem wewnętrznym. Załóżmy więc, że $p \in X$, $f(p) \in V$. Ponieważ V jest otwarty, więc istnieje takie $\varepsilon > 0$, że $y \in V$, jeśli $d_Y(f(p), y) < \varepsilon$, a ponieważ f jest ciagłe w punkcie p, zatem istnieje $\delta > 0$ takie, że $d_Y(f(x), f(p)) < \varepsilon$, jeśli $d_X(x, p) < \delta$. W ten sposób $x \in f^{-1}(V)$ jeśli tylko $d_X(x, p) < \delta$.
Odwrotnie przypuśćmy, że zbiór $f^{-1}(V)$ będzie otwarty w X dla dowolnego zbioru otwartego V w Y. Ustalmy $p \in X$ i $\varepsilon > 0$, i niech V będzie zbiorem wszystkich $y \in Y$ takich, że $d_Y(y, f(p)) < \varepsilon$. Wówczas V jest otwarty i dlatego $f^{-1}(V)$ jest otwarty. Istnieje więc takie $\delta > 0$, że $x \in f^{-1}(V)$, jeśli tylko $d_X(p, x) < \delta$. Warunek $x \in f^{-1}(V)$ oznacza, że $f(x) \in V$, czyli $d_Y(f(x), f(p)) < \varepsilon$, co kończy dowód.

Wniosek 44. *Odwzorowanie f przestrzeni metrycznej X w przestrzeń metryczną Y jest ciągłe wtedy i tylko wtedy, gdy $f^{-1}(C)$ jest dommnitym podzbiorem X dla dowolnego domkniętego zbioru C w Y.*

Dowód. Wystarczy zauważyć, że $f^{-1}(E^c) = [f^{-1}(E)]^c$.

Niech $X_1, \ldots, X_n, Y$ bedą zbiorami.

Definicja 121. Załóżmy, że każdej uporządkowanej n-ce $(x_1, \ldots, x_n)$ przyporządkowany jest dokładnie jeden element $y \in Y$. Oznaczamy $y = f(x_1, \ldots, x_n)$. Wtedy f nazywa się funkcją n-zmiennych z $X_1 \times \cdots \times X_n$ do Y. Elementy $x_1, \ldots, x_n$ nazywamy wówczas zmiennymi niezależnymi, natomiast y — zmienną zależną.

Definicja 122. Punkt $y_0 \in \mathbb{R}^n$ nazywa się granicą funkcji $f : E \mapsto \mathbb{R}^n$, $E \subset \mathbb{R}^m$, w punkcie x_0 bedącym punktem skupienia zbioru E, jeśli dla dowolnego $\varepsilon > 0$ istnieje takie $\delta > 0$, że

$$d_{\mathbb{R}^n}(f(x), y_0) < \varepsilon$$

dla wszystkich punktów $x \in E$, dla których

$$0 < d_{\mathbb{R}^m}(x, x_0) < \delta$$

Powyższa definicja jest oczywiście szczególnym przypadkiem definicji 119. Dla funkcji n-zmiennych można oczywiście wprowadzić pojęcie granicy iterowanej. Dla prostoty ograniczym się do funkcji dwóch zmiennych.

Definicja 123. Niech $f : E \mapsto \mathbb{R}^n$, $E \subset \mathbb{R}^2$. Dla ustalonego punktu x_0 i dowolnego punktu y z pewnego zbioru E niech istnieje granica $\phi(y) = \lim\limits_{x \to x_0} f(x, y)$, $(x, y) \in E$. Wówczas, jeśli istnieje

$$\lim_{y \to y_0} \phi(y) = \lim_{y \to y_0} \Big(\lim_{x \to x_0} f(x, y) \Big),$$

to element ten nazywamy granicą iterowaną funkcji $f(x, y)$ gdy najpierw $x \to x_0$, a następnie $y \to y_0$. W analogiczny sposób określamy

$$\lim_{x \to x_0} \Big(\lim_{y \to y_0} f(x, y) \Big).$$

Następujący przykład pokazuje, że istnienie granicy w punkcie (x_0, y_0) jest niezależne od istnienia granic iterowanych.

Przykład 51.

(a) Niech funkcja $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$ będzie określona wzorem

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{jeśli } x^2 + y^2 \neq 0, \\ 0 & \text{jeśli } x^2 + y^2 = 0. \end{cases}$$

Wówczas $f(0, y) = f(x, 0) = 0$ oraz $f(x, x) = \frac{1}{2}$ dla $x \neq 0$. W ten sposób funkcja f nie posiada granicy przy $(x, y) \to (0, 0)$. Ponadto, mamy $\lim\limits_{y \to 0} \Big(\lim\limits_{x \to 0} f(x, y) \Big) = \lim\limits_{y \to 0} 0 = 0$ oraz $\lim\limits_{x \to 0} \Big(\lim\limits_{y \to 0} f(x, y) \Big) = \lim\limits_{x \to 0} 0 = 0$.

(b) Niech funkcja $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$ będzie określona wzorem

$$f(x, y) = \begin{cases} x + y \sin \frac{1}{x} & \text{dla } x \neq 0, \\ 0 & \text{dla } x = 0. \end{cases}$$

Mamy $\lim\limits_{(x,y)\to(0,0)} f(x,y) = 0$ oraz $\lim\limits_{x\to 0}\Big(\lim\limits_{y\to 0} f(x,y)\Big) = 0$. Zauważmy, że granica iterowana $\lim\limits_{y\to 0}\Big(\lim\limits_{x\to 0} f(x,y)\Big)$ w ogóle nie istnieje.

Definicja 124. Mówimy, że funkcja $f : E \mapsto \mathbb{R}^n$, $E \subset \mathbb{R}^m$ jest ciągła w punkcie $x_0 \in E$, jeśli dla dowolnego $\varepsilon > 0$ istnieje $\delta > 0$, takie, że

$$d_{\mathbb{R}^n}(f(x), f(x_0)) < \varepsilon$$

dla wszystkich $x \in E$, dla których

$$d_{\mathbb{R}^m}(x, x_0) < \delta$$

Powyższa definicja jest oczywiście szczególnym przypadkiem Definicji 120.

Definicja 125. W sytuacji opisanej powyżej w Definicji 124 niech $x_0 = (x_0^1, \ldots, x_0^n)$. Mówimy, że funkcja f jest ciągła w punkcie x_0 ze względu na zmienną x_j $(1 \leqslant j \leqslant m)$, jeśli spełniony jest warunek występujący w Definicji 124 dla $x = (x_0^1, \ldots, x_0^{j-1}, x_j, x_0^{j+1}, \ldots, x_0^m)$.

Jest jasne, że w powyżej opisanej sytuacji, jeśli funkcja f jest ciągła w punkcie x_0 to f jest ciągła w x_0 ze względu na każdą zmienną z osobna. Odwrotnie zachodzić nie musi. Rozważmy ponownie funkcje f z Przykładu 51. Niech $x_0 = (0,0)$. Mamy $\lim\limits_{x_1\to 0} f(x_1, 0) = 0 = f(x_0)$, $\lim\limits_{x_2\to 0} f(0, x_2) = 0 = f(x_0)$. Funkcja f nie jest jednak ciągła w punkcie x_0, bowiem nie ma granicy w tym punkcie.

11.3 Własności funkcji ciągłych

Twierdzenie 185. *Niech f będzie odwzorowaniem ciągłym zwartej przestrzeni metrycznej X w przestrzeń metryczną Y. Wówczas zbiór $f(X)$ jest zwarty.*

Dowód. Niech $\{V_\alpha\}$ będzie pokryciem otwartym zbioru $f(X)$. Ponieważ f jest ciągłe, więc z Twierdzenia 181 wynika, że wszystkie zbiory $f^{-1}(V_\alpha)$ są otwarte. Poniewaz przestrzeń X jest zwarta, wieć istnieje skończona liczba wskaźników $\alpha_1, \ldots, \alpha_n$ takich, że

$$X \subset f^{-1}(V_{\alpha_1}) \cup \ldots \cup f^{-1}(V_{\alpha_n})$$

Ponieważ $f(f^{-1}(E)) = E$ dla każdego $E \subset Y$, zatem wobec powyższego

$$\begin{aligned} f(X) &\subset f(f^{-1}(V_{\alpha_1} \cup \ldots \cup f^{-1}(V_{\alpha_n}))) = \\ &= f(f^{-1}(V_{\alpha_1})) \cup \ldots \cup f(f^{-1}(V_{\alpha_n})) = V_{\alpha_1} \cup \ldots \cup V_{\alpha_n}. \end{aligned}$$

Definicja 126. Odwzorowanie f zbioru E w przestrzeń metryczną Y nazywamy ograniczonym, jeśli zbiór $f(E)$ jest ograniczony.

Twierdzenie 186. *Jeśli f jest odwzorowaniem ciągłym zwartej przestrzeni metrycznej X w przestrzeń $\mathbb{R}^k$, to zbiór $f(X)$ jest domknięty i ograniczony (odwzorowanie f jest więc ograniczone).*

Dowód. Teza twierdzenia wynika z charakteryzacji zbiorów zwartych w $\mathbb{R}^k$ (Tw. 186) i z Twierdzenia 185

Twierdzenie 187. *Niech f będzie pewną funkcją rzeczywistą określoną na zwartej przestrzeni metrycznej X i niech*

$$M = \sup_{p\in X} f(p), \quad m = \inf_{p\in X} f(p).$$

Wówczas istnieją punkty $p, q \in X$ takie, że $f(p) = M$ i $f(q) = m$.

Dowód. Z Twierdzenia 186 wynika, że $f(X)$ jest ograniczonym i domkniętym podzbiorem zbioru $\mathbb{R}$, a więc wobec Lematu 6, f osiąga swój kres górny i dolny.

Twierdzenie 187 uogólnia znane Twierdzenie Weierstrassa (Tw. 50). Udowodnimy teraz następujące twierdzenie o ciągłości funkcji odrwotnej (por. Tw. 49).

Twierdzenie 188. *Niech f będzie ciągłym i różnowartościowym odwzorowaniem zwartej przestrzeni metrycznej X na przestrzeń metryczną Y. Wówczas odwzorowanie f^{-1} jest ciągłym odwzorowaniem Y na X.*

Dowód. Stosując Twierdzenie 184 dla f^{-1} zamiast f widzimy, że wystarczy wykazać, że $f(V)$ jest zbiorem otwartym w Y dla każdego zbioru otwartego V w X. Ustalmy taki zbiór V. Dopełnienie V^c zbioru V jest domknięte w X i stąd zwarte (Tw. 171). Stąd $f(V^c)$ jest zwartym podzbiorem przestrzeni Y (Tw. 185), a więc domnięty w Y. (Tw. 172). Ponieważ f jest wzajemnie jednoznacznym odwzorowaniem X na Y, więc zbiór $f(V)$ pokrywa sie z dopełnieniem zbioru $f(V^c)$. Stąd $f(V)$ jest otwarty.

W książce [6] podany jest przykład świadczący o tym, że założenie zwartości w Twierdzeniu 188 jest konieczne.

Definicja 127. Niech f będzie odwzorowaniem przestrzeni metrycznej X w przestrzeń metryczną Y. Mówimy, że funkcja f jest jednostajnie ciągła na X, jeśli spełniony jest następujący warunek:
Dla dowolnego $\varepsilon > 0$ istnieje $\delta > 0$ takie, że

$$d_Y(f(x_1), f(x_2)) < \varepsilon$$

dla wszystkich punktów $x_1, x_2 \in X$, dla których

$$d_X(x_1, x_2) < \delta.$$

Następujący wynik jest uogólnieniem Twierdzenia Cantora (Tw. 51). Jego dowód jest analogiczny.

Twierdzenie 189. *Niech f będzie ciągłym przekształceniem zwartej przestrzeni metrycznej X w przestrzeń metryczną Y. Wówczas odwzorowanie f jest jednostajnie ciągłe na X.*

Twierdzenie 190. *Jeśli f jest ciagłym przekształceniem spójnej przestrzeni metrycznej X w przestrzeń metryczną Y, to zbiór $f(X)$ jest spójny.*

Dowód. Jeśli $f(X)$ nie jest spójny, to istnieją otwarte i rozłączne podzbiory V, W przestrzeni Y, obydwa przecinające się z $f(X)$ i takie, że $f(x) \subset V \cup W$. Ponieważ f jest ciągłe, więc zbiory $f^{-1}(V)$ i $f^{-1}(W)$ są otwarte. Ponadto są one niepuste i nierozłączne, a ich suma jest równa X. Oznacza to, że zbiór X nie jest spójny, wbrew założeniu.

Twierdzenie Darboux (Tw. 48) jest szczególnym przypadkiem powyższego twierdzenia. Istotnie, na mocy Twierdzenia 178 przedział $P \subset \mathbb{R}$ jest zbiorem spójnym. Wobec Twierdzenia 190, $f(P)$ jest spójnym podzbiorem przesztrzeni $\mathbb{R}$ i wystarczy jeszcze raz powołać się na Twierdzenie 178.
Zajmiemy się teraz badaniem rodziny funkcji jednakowo ciągłych. Najważniejszym wynikiem jest tutaj Twierdzenie Arzéli, posiadające istotne zastosowania na przykład w teorii równań różniczkowych.

Uwaga 62. Zauważmy, że definicje i twierdzenia z Paragrafu 10.1 oraz twierdzenia 149 i 150 pozostają oczywiście niezmienione, jesli E ędzie podzbiorem dowolnej przestrzeni metrycznej (faktycznie przestrzeń metryczna jest potrzebna w Tw. 149 oraz Tw. 150).

Definicja 128. Niech (f_n) będzie ciągiem funkcji określonych na zbiorze E. Mówimy, że ciąg (f_n) jest punktowo ograniczony na E, jeśli dla dowolnego $x \in E$, ciąg $(f_n(x))$ jest ograniczony, to znaczy jeśli istnieje funkcja ϕ przyjmująca wartości rzeczywiste i taka, że

$$|f_n(x)| \leqslant \phi(x) \quad (x \in E, n \in \mathbb{N}).$$

Mówimy, że ciąg (f_n) jest wspólnie ograniczony na E, jeśli istnieje liczba $M > 0$ taka, że

$$|f_n(x)| \leqslant M \quad (x \in E, n \in \mathbb{N}).$$

Definicja 129. Rodzinę $\mathcal{F}$ funkcji o wartościach zespolonych określonych na podzbiorze E przestrzeni metrycznej X nazywamy jednakowo ciągła na E, jeżeli dla dowolnej liczby $\varepsilon > 0$ istnieje $\delta > 0$ takie, że

$$|f(x) - f(y)| < \varepsilon$$

dla $x, y \in E$, $d(x, y) < \delta$ oraz $f \in \mathcal{F}$.

Zbadamy teraz związek między pojeciami jednakowej ciągłośći, a zbieznością jednostajną ciagu funkcji ciągłych.

Twierdzenie 191. *Jeśli (f_n) jest ciągiem funkcji zespolonych określonych na przeliczalnym zbiorze E i mającym własność, że zbiór wartości przyjmowanych w dowolnym ustalonym punkcie zbioru E przez funkcje f_n jest ciągiem ograniczonym, to istnieje podciąg (f_{n_k}) taki, że dla każdego $x \in E$, ciąg $(f_{n_k}(x))$ jest zbieżny.*

Dowód. Niech (x_i), $i = 1, 2, \ldots$ będzie ciągiem zawierającym wszystkie punkty zbioru E. Ponieważ ciąg $(f_n(x_1))$ jest ograniczony, więc istnieje podciąg, który oznaczymy przez (f_{1_k}), taki, że $(f_{1_k}(x_1))$ jest zbieżny przy $k \to \infty$. Rozważmy teraz ciągi $s_1, s_2, s_3, \ldots$,k które zapiszemy w postaci macierzy nieskończonej:

$$\begin{array}{cccccc} s_1: & f_{11} & f_{12} & f_{13} & f_{14} & \cdots \\ s_2: & f_{21} & f_{22} & f_{23} & f_{24} & \cdots \\ s_3: & f_{31} & f_{32} & f_{33} & f_{34} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{array}$$

Ciągi te mają następujące własności:

(a) ciąg (s_n) jest podciągiem ciągu (s_{n-1}), przy $n = 2, 3, \ldots$.

(b) $(f_{n_k}(x))$ jest przy $k \to \infty$ ciągiem zbieżnym (ograniczoność punktowa ciągu $(f_i(x))$ umożliwia wybrane podciągu o tej własności)

(c) zachowana jest kolejność funkcji, w której występują w następujących po sobie podciągach, to znaczy jeśli jakaś funkcja występuje wcześniej niż inna w ciągu s_1 to podobnie będzie we wszystkich ciągach s_n aż do miejsca, gdy któraś z funkcji zostanie przy wybieraniu podciągu pominięta.

Rozpatrzmy teraz ciąg

$$s: \quad f_{11} \quad f_{22} \quad f_{33} \quad f_{44} \quad \ldots$$

Zgodnie z (c) ciąg S jest (z wyjątkiem być może $n-1$ pierwszych wyrazów) podciągiem ciągu s_n. Z (b) wynika zatem, że ciąg $f_{nn}(x_i)$ jest przy $n \to \infty$ zbieżny dla każdego $x_i \in E$.

Twierdzenie 192. *(Arzéla) Niech K będzie zwartym podzbiorem przestrzeni metrycznej, $f_n : K \mapsto \mathbb{C}$ będą funkcjami ciągłymi dla $n \in \mathbb{N}$, ciąg (f_n) będzie ograniczony w każdym punkcie zbioru K i niech tworzy rodzinę jednakowo ciągłą na K. Wówczas*

(a) (f_n) jest wspólnie ograniczony na K,

(b) (f_n) zawiera podciąg jednostajnie zbieżny na K.

Dowód. Niech będzie dana liczba $\varepsilon > 0$. Na mocy Definicji 129 istnieje $\delta > 0$ takie, że warunek $d(x,y) < \delta$, $x, y \in K$ pociąga

$$|f_n(x) - f_n(y)| < \varepsilon \quad \text{dla każdego} \quad n \in \mathbb{N}. \tag{101}$$

Ponieważ K jest zbiorem zwartym, więc istnieje skończony zbiór $p_1, \ldots, p_n$ w K taki, że dla dowolnego $x \in K$ istnieje co najniej jeden punkt p_i, dla którego $d(x, p_i) < \delta$. Ponieważ ciąg (p_n) jest ograniczony punktowo, więc istnieje takie $M_i < +\infty$, że $|f_n(p_i)| < M_i$ dla wszystkich n. Jeżeli $M = \max\{M_1, \ldots, M_n\}$, to z (101) wynika, że $|f_n(x)| < M + \varepsilon$, dla dowolnego $x \in E$. To dowodzi części (a) tezy.
Udowodnimy teraz część (b). Na mocy Lematu 7 istnieje przeliczalny zbiór gęsty E w K. Z Twierdzenia 191 wynika, że istnieje podciąg (f_{n_i}) ciągu (f_n) zbieżny dla każdego $x \in E$. Niech $f_{n_i} = g_i$ dla uproszczenia zapisu. Wykażemy, że ciąg (g_i) jest zbieżny jednostajnie na K.
Niech będzie dana liczba $\varepsilon > 0$ i obierzmy $\delta > 0$ analogicznie jak na początku tego dowodu. Niech $V(x, \delta)$ będzie zbiorem wszystkich $y \in K$, dla których $d(x, y) < \varepsilon$. Ponieważ E jest gęsty w K, a K jest zwarty, więc istnieje skończony zbiór punktów $x_1, \ldots, x_m$ zbioru E taki, że

$$K \subset V(x_1, \delta) \cup \ldots \cup V(x_m, \delta). \tag{102}$$

Ponieważ $(g_i(x))$ jest zbieżny dla dowolnego $x \in E$, więc istnieje liczba naturalna N taka, że jeśli tylko $i \geqslant N$, $j \geqslant N$ oraz $1 \leqslant s \leqslant m$, to

$$|g_i(x_s) - g_j(x_s)| < \varepsilon. \tag{103}$$

Jeśli $x \in K$, to z (102) pokazuje, że $x \in V(x_s, \delta)$ dla pewnego s i wobec tego dla dowolnego i mamy

$$|g_i(x) - g_i(x_s)| < \varepsilon.$$

Jeśli ponadto $i \geqslant N$ oraz $j \geqslant N$, to z (103) wynika, że

$$|g_i(x) - g_j(x)| \leqslant |g_i(x) - g_i(x_s)| + |g_i(x_s) - g_j(x_s)| + |g_j(x_s) - g_j(x)| < 3\varepsilon.$$

Wobec Twierdzenia 144 dowód jest zakończony.

Następujące ciągi funkcjujne $f_n(x) = n$, $f_n(x) = \frac{x^2}{x^2+(1-nx)^2}$, $x \in [0,1]$, $n \in \mathbb{N}$ pokazują (por. [7], s. 133), że w Twerdzeniu Arzéli założenie o punktowej ograniczoności oraz o jednakowej ciągłości nie mogą być pominięte.

Twierdzenie 193. *Niech K będzie zwartą przestrzenią metryczną i niech $f_n : K \mapsto \mathbb{C}$ będą funkcjami ciągłymi, $n \in \mathbb{N}$. Jeżeli ciąg (f_n) jest zbieżny jednostajnie na K, to (f_n) jest rodziną jednakowo ciągłą na K.*

Dowód. Niech $\varepsilon > 0$. Istnieje więc liczba naturalna N taka, że

$$\sup_{x \in K} |f_n(x) - f_N(x)| < \varepsilon \quad \text{dla} \quad n > N.$$

Ponieważ na zbiorach zwartych funkcje ciągłe są jednostajnie ciągłe, więc istnieje $\delta > 0$ takie, że

$$(104) \qquad |f_i(x) - f_i(y)| < \varepsilon \quad (1 \leqslant i \leqslant N, \quad d(x,y) < \delta).$$

Jeśli $n > N$ oraz $d(x,y) < \delta$, to

$$(105) \quad |f_n(x) - f_n(y)| \leqslant |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < 3\varepsilon.$$

Wobec (104) i (105 dowód jest zakończony)

11.4 Przestrzenie metryczne zupełne

Definicja 130. Mówimy, że ciąg (p_n) elementów przestrzeni (X,d) jest ciągiem fundamentalnym (lub ciągiem Cauchy'ego), jeśli dla dowolnego $\varepsilon > 0$ istnieje liczba naturalna N taka, że $d(p_m, p_n) < \varepsilon$ dla $n \geqslant N$ i $m \geqslant N$.

Rozumując analogicznie jak w dowodzie Twierdzenia 32 otrzymujemy następujące.

Twierdzenie 194. *Każdy ciąg zbieżny (w dowolnej przestrzeni metrycznej) jest ciągiem Cauchy'ego.*

Definicja 131. Przestrzeń metryczna (X,d) jest przestrzenią zupełną, jeżeli każdy ciąg Cauchy'ego (p_n) elementów tej przestrzeni, jest zbieżny do pewnego elementu należącego do X.

Uwaga 63.

(a) Przypomnijmy, że zbiory $\mathbb{R}$ i $\mathbb{C}$ są przestrzeniami metrycznymi zupełnymi (zob. Tw. 33 i Uw. 17 (b)). Ponadto rozumując analogicznie jak w Uwadze 17 (c) i uwzględniając Twierdzenie 180 wnioskujemy, że przestrzeń $\mathbb{R}^k$ z metryką zdefiniowaną w Przykładzie 48 (c) jest przestrzenią metryczną zupełną.

(b) Dowolny podzbiór domknięty przestrzeni metrycznej zupełnej (X,d) (traktowany jako przestrzeń metryczna) jest przestrzenią metryczną zupełną. Istotnie, niech zbiór $E \subset X$ będzie domknięty i niech (p_n) będzie dowolnym ciągiem Cauchy'ego elementów zbioru E. Wobec zupełności przestrzeni (X,d) wnosimy, że istnieje element $p \in X$ taki, że $\lim_{n\to\infty} p_n = p$. Wobec Twierdzenia 170, $p \in E$, co kończy uzasadnienie.

Twierdzenie 195. *Każdy podzbiór zwarty przestrzeni metrycznej* (X,d) *(traktowany jako przestrzeń metryczna) jest przestrzenią zupełną.*

Dowód. Niech A będzie zwartym podzbiorem przestrzeni metrycznej (X,d) i niech (p_n) będzie ciągiem elementów zbioru A. Wobec Definicji 119 istnieje podciąg (p_{n_k}) ciągu (p_n) taki, że $\lim_{k\to\infty} p_{n_k} = p$ i $p \in A$. Z powyższego wnioskujemy, że dla dowolnego $\varepsilon > 0$ istnieją $N_1(\varepsilon), N_2(\varepsilon) \in \mathbb{N}$ takie, że $d(p_m, d_n) < \frac{\varepsilon}{2}$ dla $n \geqslant N_1(\varepsilon)$ i $m \geqslant N_1(\varepsilon)$ oraz $d(p_{n_k}, p) < \frac{\varepsilon}{2}$ dla $k \geqslant N_2(\varepsilon)$. Niech $N = \max\{N_1(\varepsilon), N_2(\varepsilon)\}$. Ponieważ $n_k \geqslant k$ dla $k \in \mathbb{N}$ (zob. Def. 42), zatem dla $k \geqslant N$ otrzymujemy

$$d(p_k, p) \leqslant d(p_{n_k}, p) + d(p_{n_k}, p_k) < \varepsilon,$$

czyli $\lim_{k\to\infty} p_k = p$ $(p \in A)$, a więc ciąg (p_n) jest zbieżny do pewnego elementu zbioru A.

Rozumując analogicznie jak w dowodzie Twierdzenia 36 otrzymujemy następujący warunek konieczny i wystarczający istnienia granicy w punkcie.

Twierdzenie 196. *Niech (X,d) będzie przestrzenią metryczną, (Y,ρ) przestrzenią metryczną zupełną, $E \subset X$. Ponadto niech f będzie odwzorowaniem E w Y, a p — punktem skupienia zbioru E. Na to, by istniała granica $\lim_{x\to p} f(x)$ potrzeba i wystarcza, by spełniony był następujący warunek Cauchy'ego:*
dla dowolnego $\varepsilon > 0$ istnieje $\delta > 0$ takie, że dla dowolnych $x', x'' \in E$, jeśli $0 < d(x',p) < \delta$ oraz $0 < d(x'',p) < \delta$, to $\rho(f(x'), f(x'')) < \varepsilon$

Przykład 52. Niech X będzie przestrzenią metryczną i niech $K = \mathbb{R}$ lub $\mathbb{C}$. Oznaczmy przez $C(X,K)$ zbiór wszystkich ciągłych i ograniczonych funkcji określonych na X, o wartościach w K, z metryką określoną następująco:

$$d(f,g) = \sup_{x\in X} |f(x) - g(x)|, \quad f, g \in C(X,K).$$

Wobec założenia, że funkcje $f, g : X \mapsto K$ są ograniczonem, liczba $d(f,g)$ jest skończona. Jest jasne, że warunki (a), (b) Definicji 111 są spełnione. Niech teraz $f, g, h \in C(X,K)$. Ponieważ dla dowolnego $x \in X$ mamy

$$|f(x) - g(x)| \leqslant |f(x) - h(x)| + |h(x) - g(x)|,$$

zatem

$$\begin{aligned} d(f,g) &= \sup_{x\in X} |f(x) - g(x)| \\ &\leqslant \sup_{x\in X} |f(x) - h(x)| + \sup_{x\in X} |h(x) - g(x)| \\ &= d(f,h) + d(h,g), \end{aligned}$$

a więc warunek (c) Definicji 111 jest również spełniony.
Wobec Twierdzenia 145 możemy powiedzieć, że ciąg (f_n) jest zbieżny do f w sensie metryki w $C(X,K)$ wtedy i tylko wtedy, gdy $f_n \to f$ jednostajnie na X. Sprawdzimy teraz, że $C(X,K)$ jest przestrzenią metryczną zupełną. Niech

(f_n) będzie ciągiem Cauchy'ego w $C(X, K)$. Oznacza to, że każdej liczbie $\varepsilon > 0$ odpowiada liczba naturalna N taka, że $d(f_n, f_m) < \varepsilon$ dla $n \geqslant N$ i $m \geqslant N$. Wobec tego na mocy Twierdzenia 144 istnieje funkcja f określona na X, do której ciąg (f_n) jest jednostajnie zbieżny. Na mocy Twierdzenia 150, funkcja f jest ciągła. Ponadto f jest ograniczona, bowiem istnieje $n \in \mathbb{N}$ takie, że

$$|f(x) - f_n(x)| < \varepsilon \quad \text{dla dowolnego} \quad x \in X,$$

a funkcja f_n jest ograniczona. Stąd $f \in C(X, K)$, a ponieważ $f_n \to f$ jednostajnie na X, zatem $d(f_n, f) \to 0$ przy $n \to \infty$, co dowodzi zupełności przestrzeni $C(X, K)$. Zauważmy ponadto, że jeśli X jest przestrzenią metryczną zwartą, to wobec Twierdzenia 186 nie musimy zakładać dodatkowo ograniczoności funkcji określonych na X.

Udowodnimy teraz pewne twierdzenie z teorii punktów stałych odwzorowań między przestrzeniami metrycznymi. Posiada ono ogromną ilość zastosowań zwłaszcza w teorii równań różniczkowych i całkowych. Wpierw jednak podamy kilka definicji.

Definicja 132. Niech A będzie dowolnym zbiorem. Punktem stałym odwzorowania $f : A \mapsto A$ nazywamy dowolny punkt $a \in A$ taki, że $f(a) = a$.

Definicja 133. Odwzorowanie $f : (X, d) \mapsto (Y, \rho)$ między przestrzeniami metrycznymi spełniające warunek $\rho(f(x), f(z)) < Md(x, z)$ dla dowolnych $x, z \in X$ gdzie M jest pewną stałą nieujemną, nazywamy lipschitzowskim; najmniejszą taką stałą M nazywamy stałą Lipschitza odwzorowania f i oznaczamy $L(f)$ (zob. Przykład 11). Jeżeli $L(f) < 1$, to odwzorowanie f nazywamy kontrakcją o stałej $L(f)$. Jeżeli $L(f) = 1$, to odwzrowanie nazywamy nierozszerzającym.

Zauważmy, że każde odwzorowanie lipschitzowskie jest oczywiście ciągłe.

Definicja 134. Niech A będzie dowolnym zbiorem i niech $f : A \mapsto A$ będzie odwzorowaniem zbioru A w siebie. Dla danego elementu $a \in A$ definiujemy $f^0(a) = a$ oraz $f^{n+1}(a) = f(f^n(a))$ dla $n \in \mathbb{N} \cup \{0\}$. Element $f^n(a)$ nazywamy n-tą iteracją elementu a odwzorowania f, natomiast zbiór $\{f^n(a) : n = 0, 1, 2, \ldots\} \subset A$ nazywamy orbitą elementu a wyznaczoną przez odwzorowanie f.

Twierdzenie 197. *(Zasada Kontrakcji Banacha) Niech (X, d) będzie przestrzenią metryczną zupełną i niech $f : X \mapsto X$ będzie kontrakcją. Wówczas odwzorowanie f posiada dokładnie jeden punkt stały u oraz $f^n(x) \to u$ dla każdego $x \in X$.*

Dowód. Niech $k < 1$ będzie stałą kontrakcji odwzorowania f (to znaczy $k = L(f)$). Odwzorowanie f posiada co najwyżej jeden punkt stały, bowiem jeśli $f(x_0) = x_0$ oraz $f(y_0) = y_0$, $x_0 \neq y_0$, to

$$d(x_0, y_0) = d(f(x_0), f(y_0)) \leqslant kd(x_0, y_0) < d(x_0, y_0),$$

co daje sprzeczność.
Aby udowodnić istnienie punktu stałego pokażemy, że dla danego $x \in X$ ciąg iteracji $(f^n(x))$ jest do niego zbieżny. Zauważmy wpierw, że

$$d(f(x), f^2(x)) \leqslant k(x, f(x)) \quad \text{oraz} \quad d(f^n(x), f^{n+1}(x)) \leqslant k^n d(x, f(x)),$$

co wynika z Zasady Indukcji Matematycznej. Dlatego dla dowolnych $n, p \in \mathbb{N}$ mamy

$$\begin{aligned} d(f^n(x), f^{n+p}(x)) &\leqslant \sum_{i=n}^{n+p-1} d(f^i(x), f^{i+1}(x)) \\ &\leqslant (k^n + \cdots + k^{n+p-1}) d(x, f(x)) \\ &= k^n(1 + \cdots + k^{p-1}) d(x, f(x)) \\ &\leqslant \frac{k^n}{1-k} d(x, f(x)), \end{aligned}$$

bowiem $0 < k < 1$. Ponieważ $k^n \to 0$, zatem z powyższych nierówności wynika, że ciąg $(f_n(x))$ jest ciągiem Cauchy'ego. Ponieważ przestrzeń X jest zupełna, zatem istnieje pewien element $u \in X$ taki, że $f^n(x) \to u$. Na mocy ciągłości odwzorowania f otrzymujemy $f(f^n(x)) \to f(u)$. Ponieważ jednak $(f^{n+1}(x))$ jest podciągiem ciągu $(f_n(x))$, zatem $f(u) = u$, czyli u jest punktem stałym odwzorowania f. Wykazaliśmy więc, że dla każdego $x \in X$ granica ciągu $(f_n(x))$ istnieje i jest punktem stałym odwzorowania f. Ponieważ f posiada co najwyżej jeden punkt stały, zatem każdy ciąg $(f_n(x))$ jest zbieżny do tego samego punktu.

11.5 Przestrzenie unormowane, przestrzenie Banacha i Hilberta

Definicja 135. Przestzeń wektorową (liniową) X nad ciałem K ($K = \mathbb{R}$ lub $K = \mathbb{C}$) nazywamy przestrzenią unormowaną (odpowiednio rzeczywistą lub zespoloną), jeśli każdemu wektorowi $x \in X$ przyporządkowana jest dokładnie jedna liczba nieujemna $||x||$ w taki sposób, że spełnione są następujące warunki:

(a) $||x + y|| \leqslant ||x|| + ||y||$ dla dowolnych $x, y \in X$ (Nierówność Trójkąta, subaddytywność),

(b) $||\alpha x|| = |\alpha| ||x||$ dla dowolnego $x \in X$ i $\alpha \in K$ (jednorodność),

(c) jeśli $||x|| = 0$, to $x = 0$.

Zauważmy, że z warunku (b) wynika, iż $||0|| = 0$.

Uwaga 64. Norma przestrzeni unormowanej wyznacza w naturalny sposób pewną metrykę tej przestrzeni. Przyjmuje się mianowicie:

$$d(x, y) = ||x - y|| \quad \text{dla dowolnych} \quad x, y \in X \tag{106}$$

Z warunku (c) oraz następującej po nim obserwacji wynika, że $d(x,y) = 0$ wtedy i tylko wtedy gdy $x = y$. Z (b) i z równości $x - y = (-1)(y - x)$ wnosimy, że $d(x,y) = d(y,x)$. Konsekwencją warunku (a) i równości $x - y = (x - z) + (z - y)$ jest Nierówność Trójkąta: $d(x,y) \leqslant d(x,z) + d(z,y)$. Stwierdzamy więc, że funkcja d spełnia aksjomaty metryki.

Definicja 136. Metrykę (106) nazywamy metryką wyznaczoną przez normę. Zbieżność określona tutaj przez nią:

$$\lim_{n\to\infty} x_n = x \iff \lim_{n\to\infty} ||x_n - x|| = 0$$

nosi nazwę zbieżności według normy.

Uwaga 65. Na ogół nie każda metryka jest wyznaczona przez pewną normę (zob. K. Sieklucki, Geometria i topologia, cz. 1, PWN, Warszawa, 1979, s. 265).

Definicja 137. Przestrzeń unormowana, zupełna ze względu na metrykę wyznaczoną przez normę nazywamy przestrzenią Banacha.

Przykład 53.

(a) W przestrzeni liniowej $\mathbb{R}^k$ określamy normę wzorem

$$||x|| = \left(\sum_{i=1}^{k} x_i^2\right)^{\frac{1}{2}}, \quad x = (x_1, \ldots, x_k) \in \mathbb{R}^k,$$

Warunku (b) i (c) w definicji normy są spełnione w sposób oczywisty. Wobec Nierówności Minkowskiego (zob. Przykład 48 (b)) nierówność trójkata jest spełniona. Tak zdefiniowana norma wyznacza metrykę rozważaną w Przykładzie 48 (b). Ponieważ przestrzeń $\mathbb{R}^k$ jest przestrzenią zupełną ze względu na te metrykę (Uwaga 63 (a)), zatem $\mathbb{R}^k$ z powyżej zdefiniowaną normą jest przestrzenią Banacha.

(b) W przestrzeni $\mathbb{R}^k$ można wprowadzić normę na wiele różnych sposobów, na przykład

$$||x|| = \max_{1\leqslant i\leqslant k} |x_i|, x = (x_1, \ldots, x_k) \in \mathbb{R}^k$$

(c) Rozważmy zbiór $C(X,K)$ (zob. Przykład 52). Zbiór $C(X,K)$ jest oczywiście przestrzenią liniową ze względu na działania dodawania i mnożenia przez skalar określone w Definicji 28. Określamy

$$||f|| = \sup_{x\in X} |f(x)|, \quad \text{dla} \quad f \in C(X,K).$$

Ponieważ f jest z założenia ograniczona, więc $||g|| < +\infty$. Jest jasne, że $||g|| = 0$ tylko wtedy gdy $f = 0$. Jeżeli $h_\alpha = \alpha f$, to $|h_\alpha(x)| = |\alpha||f(x)|$ dla każdego $x \in X$, zatem $||h_\alpha|| = |\alpha|||f||$. Jeśli natomiast $h = f + g$, to

$$|h(x)| \leqslant |f(x) + g(x)| \leqslant ||f|| + ||g|| \quad \text{dla dowolnego} \quad x \in X,$$

zatem

$$||f+g|| \leqslant ||f|| + ||g||.$$

Tak zdefiniowana norma wyznacza metrykę rozważaną w Przykładzie 52, zatem $C(X,K)$ z powyżej zdefiniowaną normą jest przestrzenią Banacha.

(d) Można udowodnić, że jeśli rozważyc zbiór $C(X,K)$ z normą

$$||f||_1 = \int_a^b |f(t)|dt, \quad f \in C([a,b],K),$$

to otrzymamy przestrzeń unormowaną, która nie jest przestrzenią Banacha (zob. Trenogin, V. A., Funtksjonal'nyj analiz, Moskwa, 1980, s. 49-50).

Uwaga 66. Analogicznie jak w Twierdzeniu 12 (7) można sprawdzić, że norma posiada następującą własność

$$||x|| - ||y|| \leqslant ||x-y|| \tag{107}$$

dla dowolnych x, y należacych do przestrzeni wektorowej X. Z (107) wynika od razu, że norma jest funkcją ciągłą. Innymi słowy

$$\lim_{n\to\infty} x_n = x \implies \lim_{n\to\infty} ||x_n|| = ||x||.$$

Twierdzenie 198. *W przestrzeni unormowanej działania algebraiczne są funkcjami ciągłymi, to znaczy*

$$\lim_{n\to\infty} x_n = x \ i \ \lim_{n\to\infty} y_n = y \implies \lim_{n\to\infty} (x_n + y_n) = x + y$$

$$\lim_{n\to\infty} \lambda_n = \lambda \ i \ \lim_{n\to\infty} x_n = x \implies \lim_{n\to\infty} \lambda_n x_n = \lambda x$$

Dowód. Ponieważ $(x_n + y_n) - (x + y) = (x_n - x) + (y_n - y)$, więc

$$0 \leq ||(x_n + y_n) - (x + y)|| \leq ||x_n - x|| + ||y_n - y||$$

dla $n \in \mathbb{N}$. Z założenia $\lim\limits_{n\to\infty} ||x_n - x|| = 0$ i $\lim\limits_{n\to\infty} ||y_n - y|| = 0$, zatem wobec powyższej nierówności otrzymujemy $\lim\limits_{n\to\infty} ||(x_n - y_n) - (x + y)|| = 0$ czyli $\lim\limits_{n\to\infty} (x_n - y_n) = x + y$. Podobnie, z równości $\lambda_n x_n - \lambda x = (\lambda_n - \lambda)x_n + \lambda(x_n - x)$ otrzymujemy

$$0 \leq ||\lambda_n x_n - \lambda x|| \leq |\lambda_n - \lambda| ||x_n|| + |\lambda| ||x_n - x||$$

dla $n \in \mathbb{N}$. Z założenia $\lim\limits_{n\to\infty} \lambda_n = \lambda$ i $\lim\limits_{n\to\infty} x_n = x$, a zatem wobec powyższej nierówności i Uwagi 66 $\lim\limits_{n\to\infty} \lambda_n x_n = \lambda x$.

Definicja 138. Niech X będzie przestrzenią liniową nad ciałem skalarów K ($K = \mathbb{R}$ lub $K = \mathbb{C}$). Iloczynem skalarnym w przestrzeni X nazywamy każdą funkcję

$$(x, y) \to (x|y)$$

spełniającą następujące warunki:

(a) $\forall_{x,y\in X}(x|y) = \overline{(x|y)}$ (sprzężenie)

(b) $\forall_{x,y,z\in X}(x + y|z) = (x|z) + (y|z)$

(c) $\forall_{\alpha\in K}\forall_{x,y\in X}(\alpha x|y) = \alpha(x|y)$

(d) $\forall_{x\in X}(x|x) \geq 0$

(e) $\forall_{x\in X}(x|x) = 0 \implies x = 0$

Liczbę $(x|y)$ nazywamy iloczynem skalarnym wektorów x, y.

Uwaga 67. Z warunków (a)-(e) Definicji 138 wynika od razu, że

$$\forall_{x\in X}(x|x) \in \mathbb{R}, \quad \forall_{x,y,z\in X}(z|x + y) = (z|x) + (z|y)$$

oraz

$$\forall_{\alpha\in K}\forall_{x,y\in X}(x|\alpha y) = \overline{\alpha}(x|y).$$

Twierdzenie 199. *(Nierówność Schwarza) Dla dowolnych $x, y \in X$ mamy*

$$|(x|y)| \leq \sqrt{(x|x)}\sqrt{(y|y)}. \tag{108}$$

Dowód. Istotnie, niech $y \neq 0$ i niech $\alpha = (x|y)/(y|y)$. Wtedy

$$\begin{aligned} 0 \leq (x - \alpha y|x - \alpha y) &= (x|x) - \overline{\alpha}(x|y) - \alpha\overline{(x|y)} + \alpha^2(y|y) \\ &= (x|x) - \frac{|(x|y)|^2}{(y|y)}. \end{aligned}$$

Stąd (108) (także dla $y = 0$, bowiem $(x|0) = (x|0x) = 0(x|x) = 0$).

Uwaga 68. Mając dany iloczyn skalarny w przestrzeni X możemy określić w niej normę za pomocą wzoru

$$||x|| = \sqrt{(x|x)} \text{ dla dowolnego } x \in X.$$

Z warunków (d) i (e) Defincji 138 wynika, że $x \to ||x||$ jest dobrze określoną funkcją nieujemną i że spełnia ona warunek (c) w definicji normy. Jej jednorodność wynika z następujących równości:

$$\begin{aligned} ||\alpha x|| &= \sqrt{(\alpha x|\alpha x)} = \sqrt{\alpha(x|\alpha x)} = \sqrt{\alpha\overline{\alpha}(x|x)} \\ &= \sqrt{|\alpha|^2(x|x)} = |\alpha|\sqrt{(x|x)} = |\alpha|||x||. \end{aligned}$$

Na mocy Nierówności Schwarza otrzymujemy

$$\|x+y\|^2 = \|x\|^2 + (x|y) + \overline{(x|y)} + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

a stąd $\|x + y\| \leq \|x\| + \|y\|$, to jest subaddytywność normy.

Definicja 139. Przestrzeń unormowaną o normie wyznaczonej przez iloczyn skalarny nazywamy przestrzenią unitarną. Przestrzeń unitarną zupełną nazywamy przestrzenią Hilberta.

Przykład 54.

(a) W przestrzeni $\mathbb{R}^k$ możemy określić iloczyn skalarny następująco:

$$(x|y) = \sum_{i=1}^{k} x_i y_i, \text{ gdzie } x = (x_1, \ldots, x_k), y = (y_1, \ldots, y_k) \in \mathbb{R}^k$$

Powyższy iloczyn skalarny oznacza normę rozważaną w Przykładzie 53 (a). Przestrzeń $\mathbb{R}^k$ jest przestrzenia Hilberta.

(b) Przy dowolnej liczbie naturalnej k niech $\mathbb{C}^k$ będzie zbiorem uporządkowanych k-wyrazowych ciągów $x = (x_1, \ldots, x_k)$, gdzie $x_1, \ldots, x_k$ są liczbami zespolonymi zwanymi współrzędnymi elementu x. Działania dodawania i mnożenia przez skalar określamy w zwykły sposób ("po współczynnikach"). Zbiór $\mathbb{C}^k$ z tak określonymi działaniami jest przestrzenią liniową nad ciałem $\mathbb{R}$ lub $\mathbb{C}$ (sprawdzenie elementarne). Przestrzeń $\mathbb{C}^k$ z iloczynem skalarnym określonym wzorem

$$(x|y) = \sum_{i=1}^{k} x_i \overline{y_i}, \text{ gdzie } x = (x_1, \ldots, x_k), y = (y_1, \ldots, y_k) \in \mathbb{C}^k$$

jest przestrzenią Hilbreta.
Nierówność Schwarza podana w Twierdzeniu 18 jest szczególnym przypadkiem nierówności (108) dla powyżej zdefiniowanego iloczynu skalarnego.

(c) Przestrzeń wektorowa $C([a,b],\mathbb{C})$ z iloczynem skalarnym określonym wzorem

$$(f|g) = \int_a^b f(t)\overline{g(t)}dt, \text{ gdzie } f, g \in C([a,b],\mathbb{C})$$

jest przestrzenią unitarną, ale nie jest przestrzenią Hilberta (por. Przykład 53 (d)).

(d) Przestrzeń $C([a,b],\mathbb{R})$ z normą określoną w Przykładzie 53 (c) jest przykładem przestrzeni Banacha, która nie jest przestrzenia Hilberta. Przestrzeń $C([a,b],\mathbb{R})$ nie jest refleksywna, a taką jest każda przestrzeń Hilberta (zob. A. Alexiewicz, Analiza funkcjonalna, PWN, Warszawa, 1969, s189, 423).

Twierdzenie 200. *Iloczyn skalarny w przestrzeni unitarnej jest funkcją ciągłą, to znaczy*

$$\lim_{n\to\infty} x_n = x \ i \ \lim_{n\to\infty} y_n = y \implies \lim_{n\to\infty} (x_n|y_n) = (x|y).$$

Dowód. Z założenia $\lim_{n\to\infty} \|x_n - x\| = 0$ i $\lim_{n\to\infty} \|y_n - y\| = 0$. Ponieważ

$$(x_n|y_n) - (x|y) = (x_n - x|y_n) + (x|y_n - y),$$

więc na mocy Nierówności Schwarza mamy

$$|(x_n|y_n) - (x|y)| \leq \|x_n - x\|\|y_n\| + \|x\|\|y_n - y\| \quad \text{dla} \quad n \in \mathbb{N}.$$

Stąd uwzględniając ciągłość normy wnosimy, że

$$\lim_{n\to\infty} |(x_n|y_n) - (x|y)| = 0$$

11.6 Operatory liniowe

Definicja 140. Odzworowanie liniowe A przestrzeni wektorowej X w przestrzeń wektorową Y nazywamy przekształceniem linowym, jeżeli:

(a) $A(x_1 + x_2) = Ax_1 + Ax_2$ dla dowolnych $x_1, x_2 \in X$ (addytywność)

(b) $A(\alpha x) = \alpha A(x)$ dla dowolnych $x \in X, \alpha \in K$ ($K = \mathbb{C}$ lub $\mathbb{R}$; jednorodność)

Zauważmy, że jeśli A jest liniowe, to często piszemy Ax zamiast $A(x)$. Ponadto zauważmy, że jeśli A jest linowe to $A0 = 0$.

Definicja 141. Przekształcenie liniowe przestrzeni wektorowej X w X nazywamy operatorem liniowym X. Operator liniowy A na X, który jest wzajemnie jednoznaczny oraz odwzorowuje X na X nazywa się odwracalny. W tym przypadku można określić na X operator A^{-1} przyjmując $A^{-1}(Ax) = x$.

Następujące twierdzenie podaje ważną własność operatorów liniowych na przestrzeniach skończenie wymiarowych

Twierdzenie 201. *Operator liniowy na skończenie wymiarowej przestrzeni X jest wzajemnie jednoznaczny wtedy i tylko wtedy, gdy zbiorem jego wartości jest cała przestrzeń X.*

Dowód powyższego twierdzenia można znaleźć na przykład w książce [6] s. 174.

Definicja 142. Rozważmy przekształcenie liniowe A przestrzeni unormowanej X na przestrzeń unormowaną Y. Określamy

$$\|A\| = \sup\left\{\frac{\|Ax\|}{\|x\|} : x \in X, x \neq 0\right\}. \tag{109}$$

Jeżeli $\|A\| < +\infty$, to mówimy, że A jest ograniczonym przekształceniem liniowym. W przypadku, gdy $Y = \mathbb{R}$ lub $\mathbb{C}$ to mówimy wtedy o ograniczonym funkcjonale liniowym (rzeczywistym lub zespolonym).
W równości 109 symbol $\|x\|$ oznacza normę wektora x w przestrzeni X, natomiast $\|Ax\|$ - normę wektora Ax w przestrzeni Y.

Uwaga 69. Zauważmy, że w 109 możemy ograniczyć się do wektorów jednostkowych, to jest do takich wektorów x, dla których $\|x\| = 1$. Nie zmienia się przy tym kres górny, ponieważ

$$\|A(\alpha x)\| = \|\alpha Ax\| = |\alpha|\|Ax\|.$$

Zauważmy też, że $\|A\|$ jest najmniejszą liczbą taką, że dla każdego $x \in X$ zachodzi

$$\|Ax\| \leqslant \|A\|\|x\|.$$

Twierdzenie 202. *Dla dowolnego przekształcenia liniowego A przestrzeni unormowanej X w przestrzeń unormowaną Y następujące warunki są równoważne:*

(a) przekształcenie A jest ograniczone

(b) przekształcenie A jest ciągłe

(c) przekształcenie A jest ciągłe tylko w jednym punkcie przestrzeni X

Dowód. Ponieważ $||A(x_1 - x_2)|| \leqslant ||A||||x_1 - x_2||$, zatem jest jasne, że warunek (a) implikuje (b). Warunek (c) wynika z (b) w sposób oczywisty.
Przypuśćmy teraz, że przekształcenie A jest ciągłe w punkcie $x_0 \in X$. Wówczas dla każdej liczby $\varepsilon > 0$ istnieje liczba $\delta > 0$ taka, że $||x - x_0|| < \delta$ implikuje $||Ax - Ax_0|| < \varepsilon$. Innymi słowy nierówność $||x|| < \delta$ implikuje nierówność

$$||A(x_0 + x) - Ax_0|| < \varepsilon.$$

Z liniowości przekształcenia A mamy wówczas $||Ax|| < \varepsilon$. Stąd dla dowolnego $a \in (0, \delta)$ otrzumujemy

$$\begin{aligned}\sup\{||Ax|| : x \in X, ||x|| = 1\} &= \sup\{||A(\frac{1}{\alpha}\alpha x)|| : x \in X, ||x|| = 1\}\\ &= \frac{1}{\alpha}\sup\{||A(\alpha x)|| : x \in X, ||x|| = 1\}\\ &< \frac{\varepsilon}{\alpha},\end{aligned}$$

zatem $||A|| \leqslant \frac{\varepsilon}{\delta}$, a więc warunek (a) wynika z (c).

Przykład 55.

(a) Dla $f \in C([a,b],\mathbb{C})$ określamy

$$A(f) = \int_a^b f(t)dt.$$

Jest jasne, że $A : C([a,b],\mathbb{C}) \mapsto \mathbb{C}$ jest zespolonym funkcjonałem liniowym. Mamy

$$||A|| = \sup_{||f||=1} \Big|\int_a^b f(t)dt\Big| \leqslant \sup_{||f||=1} \int_a^b |f(t)|dt \leqslant \int_a^b 1dt = b-a,$$

a zatem funkcjonał A jest ograniczony (a więc ciągły na mocy powyższego twierdzenia).

(b) Rozpatrzymy nieco ogólniejszą sytuację. Niech $K : [a,b] \mapsto \mathbb{C}$ będzie pewną funkcją ciągłą. Dla $f \in C([a,b],\mathbb{C})$ określamy

$$A(f) = \int_a^b K(t)f(t)dt.$$

Można łatwo sprawdzić, że $A : C([a,b],\mathbb{C}) \mapsto \mathbb{C}$ jest zespolonym funkcjonałem liniowym ponieważ

$$\begin{aligned} ||A|| &= \sup_{||f||=1} \Big|\int_a^b K(t)f(t)dt\Big| \\ &\leqslant \sup_{||f||=1} \int_a^b |K(t)f(t)|dt \\ &\leqslant \int_a^b |K(t)| \cdot 1dt \\ &< +\infty, \end{aligned}$$

zatem funkcjonał A jest ograniczony.

(c) Oznaczmy przez E podprzestrzeń przestrzeni unormowanej $C([a,b],\mathbb{R})$, której elementami są funkcje różcznikowalne w sposób ciągły. Określamy przekształcenie liniowe $A : E \mapsto C([a,b],\mathbb{R})$ wzorem

$$Af(t) = f'(t), \quad f \in E, t \in [a,b]$$

Przekształcenie A nie jest jednak ciągłe. To wynika na przykład z faktu, że ciąg funkcji $\phi_n(t) = \frac{\sin nt}{n}$, $n \in \mathbb{N}$, $t \in [a,b]$ jest zbieżny według normy przestrzeni $C([a,b],\mathbb{R})$ do funkcji tożsamościowo równej 0 na $[a,b]$, natomiast ciąg $A\phi_n(t) = \cos nt$, $n \in \mathbb{N}$, $t \in [a,b]$ nie jest zbieżny dla p.w. t (zob. [1], Zad. 5, s. 42).

(d) Przekształcenie liniowe A zdefiniowane w punkcie (c) można rozpatrywać jako przekształcenie działające z przestrzeni $C^1([a,b],\mathbb{R})$ funkcji różniczkowalnych na $[a,b]$ z normą

$$||f||_1 = \max_{t\in[a,b]}|f(t)| + \max_{t\in[a,b]}|f'(t)|, \quad f \in C^1([a,b],\mathbb{R})$$

w przestrzeń $C([a,b],\mathbb{R})$. W tym przypadku przekształcenie A jest „na" (każda funkcja ciągła posiada funkcję pierwotną) oraz jest ciągłe, bowiem

$$||A|| = \sup_{||f||_1=1}||Af|| = \sup_{||f||_1=1}(\sup_{t\in[a,b]}|f'(t)|) \leqslant 1.$$

Podziękowania

Pragnę serdecznie podziękować wszystkim, którzy dołożyli coś od siebie do stworzenia tego pliku. W szczególności niżej wymienionym:

- Tymon Tomczak (ostatni rozdział)
- Lidia Kopczyńska (dysk ze zdjęciami z wykładów)

Przykład 56. Oznaczmy przez $L(X,Y)$ przestrzeń wektorową wszystkich przekształceń liniowych i ciągłych przestrzeni unormowanej X w przestrzeń unormowaną Y ze zwykłymi działanami dodawania przekształceń i mnożenia przekształceń przez skalar z ciała K ($K = \mathbb{R}$ lub $K = \mathbb{C}$). Wobec Twierdzenia 202 dla dowolnego $A \in L(X,Y)$ mamy $||A|| < +\infty$. Dalej jeżeli $||A|| = 0$, to wobec Uwagi 69 mamy $Ax = 0$ dla każdego $x \in X$, to jest $A = 0$. Z kolei, ponieważ dla każdego $x \in X$ i $\alpha \in K$:

$$||(\alpha A)(x)|| = ||\alpha A(x)|| = |\alpha|||Ax|| \leqslant |\alpha|||A||||x||,$$

zatem $||\alpha A|| \leqslant |\alpha||A||$. Niech teraz $\alpha \neq 0$. Mamy $A = \frac{1}{\alpha}\alpha A$, a wobec tego $||A|| \leqslant \frac{1}{\alpha}||\alpha A||$, to jest $||\alpha A|| \leqslant ||\alpha A||$, co łącznie z poprzednio uzyskaną nierównością daje równość $|\alpha|||A|| = ||\alpha A||$.
W końcu dla każdego $x \in X$ mamy

$$||(A+B)x|| \leqslant ||Ax|| + ||Bx|| \leqslant (||A|| + ||B||)||x||,$$

a więc $||A+B|| \leqslant ||A|| + ||B||$. W ten sposób wyrażenie zdefiniowane równością (109) dla elementów przestrzeni wektorowej $L(X,Y)$ staje się normą.

Twierdzenie 203. *Jeżeli Y jest przestrzenią Banacha, to przestrzeń $L(X,Y)$ z normą określoną równością (109) jest też przestrzenią Banacha.*

Dowód. Niech (A_n) będzie ciągiem elementów przestrzeni $L(X, Y)$ spełniającym warunek Cauchy'ego. Dla $\varepsilon > 0$ istnieje więc taka liczba $n \in \mathbb{N}$, że $||A_n - A_m|| < \varepsilon$ dla $n, m \geqslant N$. Wobec Uwagi 69 jest jasne, że

$$||A_n x - A_m x|| \leqslant \varepsilon ||x|| \quad \text{dla} \quad m, m \geqslant N, x \in X. \tag{110}$$

Z (110) wnosimy, że dla każdego $x \in X$ ciąg $(A_n x)$ elementów przestrzeni Y spełnia warunek Cauchy'ego, a zatem jest on zbieżny. Niech $A(x) = \lim_{n\to\infty} A_n(x)$ dla wszystkich $x \in X$. Ponieważ każde przekształcenie A_n jest liniowe, więc z ciąłośći działań algebraicznych w przestrzeni Y (Tw. 198) wynika, że

$$A(x + y) = \lim_{n\to\infty} A_n(x + y) = \lim_{n\to\infty} A_n(x) + \lim_{n\to\infty} A_n(y) = A(x) + A(y)$$

oraz

$$A(\alpha x) = \lim_{n\to\infty} A_n(\alpha x) = \alpha \lim_{n\to\infty} A_n(x) = \alpha A x$$

dla dowolnych $x, y \in X, \alpha \in K$. Zatem przekształcenie $A : X \to Y$ jest liniowe. Przechodząc w nierówności (110) do granicy przy $m \to \infty$ dostajemy

$$||A_n x - Ax|| \leqslant \varepsilon ||x|| \quad \text{dla} \quad n \geqslant N, x \in X. \tag{111}$$

Wobec (111) dla każdego $x \in X$ otrzymujemy

$$||Ax|| \leqslant ||A_N x|| + \varepsilon ||x|| \leqslant (||A_N|| + \varepsilon)||x||.$$

Na mocy Twierdzenia 202 przekształcenie A jest ciągłe czyli $A \in L(X, Y)$. Ponadto (111) implikuje, że $||A_n - A|| \leqslant \varepsilon$ dla $n \geqslant N$. Stąd ciąg (A_n) jest zbieżny według normy w przestrzeni $L(X, Y)$ do przekształcenia A, co kończy dowód.

Twierdzenie 204. *Jeżeli* $A \in L(X, Y), B \in L(Y, Z)$, *to* $(B \circ A) \in L(X, Z)$ *oraz*

$$||B \circ A|| \leqslant ||B|| ||A||. \tag{112}$$

Dowód. Dla każdego $x \in X$ mamy

$$||(B \circ A)x|| = ||B(Ax)|| \leqslant ||B|| ||A|| ||x||,$$

a stąd otrzymujemy (112).

Następne twierdzenie orzeka, że każde przekształcenie liniowe między przestrzeniami skończenie wymiarowymi jest ciągłe (faktycznie jest ono jednostajnie ciągłe).

Twierdzenie 205. *Jeżeli* $A : \mathbb{R}^m \to \mathbb{R}^n$ *jest przekształceniem liniowym, to* $||A|| < +\infty$.

Dowód. Niech $\{e_1, \ldots, e_m\}$ będzie bazą standardową w $\mathbb{R}^m$ i załóżmy, że $x = \sum_{i=1}^m c_i e_i$, $||x|| \leqslant 1$, co oznacza, że $|c_i| \leqslant 1$ dla $i = 1, \ldots, m$. Wtedy

$$||Ax|| = \left\| \sum_{i=1}^m c_i A e_i \right\| \leqslant \sum_{i=1}^m |c_i| ||Ae_i|| \leqslant \sum_{i=1}^m ||Ae_i||,$$

zatem

$$||A|| \leqslant \sum_{i=1}^m ||Ae_i|| < +\infty$$

(jednostajna ciągłość przekształcenia A wynika z nierówności $||Ax - Ay|| \leqslant \|A\| \|x - y\|$ dla $x, y \in \mathbb{R}^m$).

Twierdzenie 206. *Niech Ω będzie zbiorem wszystkich odwracalnych operatorów liniowych na $\mathbb{R}^k$.*

(a) Jeżeli $A \in \Omega, B \in L(\mathbb{R}^k, \mathbb{R}^k)$ i $\|B - A\| \|A^{-1}\| < 1$ (kula o środku w punkcie A), to $B \in \Omega$.

(b) Ω jest otwartym podzbiorem przestrzeni $L(\mathbb{R}^k, \mathbb{R}^k)$ i odwzorowanie $A \mapsto A^{-1}$ jest ciągłe na Ω.

(Odwzorowanie z punktu (b) odwzorowuje wzajemnie jednoznacznie zbiór Ω na siebie i samo jest odwracalne)

Dowód. Niech $||A|| = \frac{1}{\alpha}, ||B - A|| = \beta$. Wobec założenia mamy $\beta < \alpha$. Dla dowolnego $x \in \mathbb{R}^k$ otrzymujemy

$$\begin{aligned}
\alpha ||x|| &= \alpha ||AA^{-1}x|| \\
&\leqslant \alpha \|A^{-1}\| \|Ax\| \\
&= ||Ax|| \\
&\leqslant ||(A - B)x|| + ||Bx|| \\
&\leqslant \beta ||x|| + ||Bx||,
\end{aligned}$$

a więc

$$(\alpha - \beta) ||x|| \leqslant ||\beta x|| \quad (x \in \mathbb{R}^k). \tag{113}$$

Ponieważ $\alpha - \beta > 0$, (113) pokazuje, że $Bx \neq 0$, jeśli $x \neq 0$, a więc, że B jest $1 : 1$ (bowiem jeśli $x \neq y$, to $Bx - By = B(x - y) \neq 0$, czyli $Bx \neq By$). Na mocy Twierdzenia 201, $B \in \Omega$. Zachodzi to dla każdego B, przy którym $||B - A|| < \alpha$. Wnioskujemy więc, że Ω jest zbiorem otwartym.
Zastepując w nierównośći (113) x przez $B^{-1}y$, otrzymujemy

$$(\alpha - \beta) ||B^{-1}y|| \leqslant ||BB^{-1}y|| = ||y||,$$

a więc $||B^{-1}|| \leqslant (\alpha - \beta)^{-1}$. Równość $B^{-1} - A^{-1} = B^{-1}(A - B)A^{-1}$ i Twierdzenie 204 pociągają za sobą

$$||B^{-1} - A^{-1}|| \leqslant \|B^{-1}\|\|A - B\|\|A^{-1}\| \leqslant \frac{\beta}{\alpha(\alpha - \beta)},$$

i tym samym twierdzenie o ciągłości ponieważ $\beta \to 0$, jeśli $B \to A$.

Przykład 57.

(a) Niech Y będzie przestrzenią unormowaną. Można łatwo sprawdzić, że $A \in L(K, Y)$ ($K = \mathbb{R}$ lub $\mathbb{C}$) wtedy i tylko wtedy, istnieje element $a \in Y$ taki, że dla każdego $x \in K$

$$Ax = ax \tag{114}$$

Element a jest przy tym wyznaczony jednoznacznie przez przekształcenie A oraz

$$||A|| = ||a||$$

(zob. [3], s. 113) Dla każdego $a \in Y$ oznaczmy przez A_a przekształcenie określone wzorem (114) Wtedy

$$A_{a+b} = A_a + A_b, \quad A_{\alpha a} = \alpha A_a \quad (\alpha \in K),$$

co oznacza, że $a \mapsto A_a$ jest przekształceniem liniowym. Ponadto $||A_a|| = ||a||$. Zauważmy jeszcze, że jeśli $\alpha \in K$, $b \in Y$, to

$$A_\alpha \in L(K, K), \quad A_b \in L(K, Y) \quad \text{i} \quad A_{b\alpha} = A_b \circ A_\alpha.$$

(b) Można udowodnić, że odwzorowanie $A : K^m \to Y$ jest liniowe wtedy i tylko wtedy, gdy dla każdego $x = (x_1, \ldots, x_m) \in K^m$

$$Ax = \sum_{i=1}^{m} a_i x_i,$$

gdzie $a_1, \ldots, a_m$ są elementami przestrzeni Y wyznaczonymi jednoznacznie przez przekształcenie A ([3], s. 113).

(c) Z poprzedniego przykładu wynika, że każde odwzorowanie $A : K^m \to K^n$, $A = (A_1, \ldots, A_m)$ jest liniowe wtedy i tylko wtedy, gdy istnieją liczby $a_{ij} \in K$, $(i = 1, \ldots, n; j = 1, \ldots, m)$ - wyznaczone jednoznacznie przez odwzorowanie A - takie, że dla każdego $x = (x_1, \ldots, x_m) \in K^m$

$$A_i x = \sum_{j=1}^{m} a_{ij} x_j, \quad i = 1, \ldots, n. \tag{115}$$

Układ (115) zapisuje się także, przy użyciu macierzy w postaci

$$\begin{bmatrix} A_1x \\ A_2x \\ \vdots \\ A_nx \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

Ponieważ macierz jednokolumnową utożsamia się zwykle z wektorem stanowiącym jej kolumnę, zatem

$$Ax = [A]x, \quad x \in K^m,$$

gdzie $[A] = [a_{ij}] \quad (i = 1, \ldots, n; j = 1, \ldots, m)$.
Weźmy teraz jakąkolwiek macierz liczbową $[A] = [a_{ij}] \quad (i = 1, \ldots, n; j = 1, \ldots, m)$. Wyznacza ona pewne przekształcenie $A_[A] \in L(K^m, K^n)$ wyznaczone wzorem (115). Z definicji sumy $[A] + [B]$ macierzy $[A]$ i $[B]$ oraz iloczynu $\alpha[A]$ macierzy $[A]$ przez pewną liczbe $\alpha \in K$ wynika od razu, że

$$A_{[A]+[B]} = A_{[A]} + A_{[B]}, \quad A_{\alpha[A]} = \alpha A_{[A]},$$

a zatem określona w ten sposób bijekcja

$$[A] \to A_{[A]}$$

zbioru wszystkich maceirzy rozważanego typu na $L(K^m, K^n)$ jest przekształceniem liniowym. Można łatwo sprawdzić, że

$$||A|| = \left(\sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij}^2 \right) \right)^{\frac{1}{2}}.$$

Zauważmy jeszcze, że jeśli $[A] = [a_{ij}]$ $(i = 1, \ldots, n; j = 1, \ldots, k)$ i $[B] = [b_{ij}]$ $(i = 1, \ldots, k; j = 1, \ldots, m)$, to odwzorowanie $A_{[A][B]} \in L(K^m, K^n)$ jest złożeniem odwzorowań $A_{[B]} \in L(K^m, K^k)$ i $A_{[A]} \in L(K^k, K^n)$:

$$A_{[A][B]} = A_{[A]} A_{[B]}.$$

12 Rachunek różniczkowy funkcji wielu zmiennych

12.1 Różniczkowanie

Określenie pochodnej funkcji rzeczywistej zmiennej rzeczywitej zostało poane w Definicji 54 (sem. I). Punktem wyjścia do następującego uogólnienia tej definicji na przypadek wielu zmiennych o wartościach w przestrzeni $\mathbb{R}^m$ jest Twierdzenie 53 i Wniosek 8 (b) (patrz sem. I)

Definicja 143. Niech E będzie zbiorem otwartym w $\mathbb{R}^n$, f - odwzorowanie E w $\mathbb{R}^m$ i niech $x \in E$. Jeśli istnieje przekształcenie liniowe przestrzeni $\mathbb{R}^n$ w $\mathbb{R}^m$ takie, że

$$\lim_{h \to 0} \frac{||f(x+h) - f(x) - Ah||}{||h||} = 0, \tag{116}$$

to mówimy, że przekształcenie f jest różniczkowalne w punkcie x i piszemy

$$f'(x) = A.$$

Przekształcenie liniowe A nazywamy wówczas pochodną odwzorowania f w punkcie x. Jeżeli odwzorowanie f jestróżcznikowalne dla każdego $x \in E$, to mówimy, że f jest różniczkowalna w E.

Uwaga 70.

(a) Jeśli we wzorze (116) $h \in \mathbb{R}^n$ jest dostatecznie małe, to punkt $x + h \in E$, bowiem E jest zbiorem otwartym. Wobec tego $f(x+h)$ ma sens i $f(x+h) \in \mathbb{R}^m$. Ponadto $Ah \in \mathbb{R}^m$, a więc $f(x+h) - f(x) - Ah \in \mathbb{R}^m$ i norma w liczniku ułamka (116) jest normą w przestrzeni $\mathbb{R}^m$. W mianowniku natomiast wystepuje norma wektora $h \in \mathbb{R}^n$.

(b) Zauważmy, że wobec Przykładu 57 (a), jeśli $n = m = 1$, to powyższa definicja jest zgodna z Definicją 54.

Następujące twierdzenie pokazuje, że pochodna odwzorowania jest wyznaczona jednoznacznie.

Twierdzenie 207. *Niech E, x, f będą takie jak w Definicji 143. Jeśli odwzorowanie f jest różcznikowalne w puncie x, to istnieje dokładnie jedno przekształcenie liniowe $A : \mathbb{R}^n \to \mathbb{R}^m$, dla którego zachodzi równość (116).*

Dowód. Załóżmy, że równość (116) zachodzi dla $A = A_1$ i $A = A_2$. Jeśli

$B = A_1 - A_2$, to z równości

$$\begin{aligned}||Bh|| &= ||-f(x+h) + f(x) + A_1 h + f(x+h) - f(x) - A_2 h|| \\ &\leqslant ||f(x+h) - f(x) - A_1 h|| + ||f(x+h) - f(x) - A_2 h||\end{aligned}$$

dla dostatecznie małych h wynika, że $\frac{||Bh||}{||h||} \to 0$, gdy $h \to 0$. Wynika stąd, że dla ustalonego $h \neq 0$

$$\frac{||Bh||}{||h||} \to 0, \quad \text{gdy} \quad t \to 0. \tag{117}$$

Ponieważ przekształcenie B jest liniowe, zatem lewa strona wzoru (117) nie zależy od t. Stąd $Bh = 0$ dla wszystkich $h \in \mathbb{R}^n$, a więc $B \equiv 0$.

Uwaga 71.

(a) Równość (116) można zapisać w postaci

$$f(x+h) - f(x) = Ah + r(h), \tag{118}$$

gdzie

$$\lim_{h\to 0} \frac{||r(h)||}{||h||} = 0.$$

Oczywiście (118) implikuje (116). Jeśli natomiast spełnione jest (116), to wystarczy przyjąc $r(h) = f(x+h) - f(x) - Ah$, by zachodziło (118).

(b) Niech f i E będa takie same jak w Definicji 143 i niech f będzie różczkowalna na E. Dla każdego $x \in E$, $f'(x)$ jest funkcją, mianowicie przekształceniem liniowym $\mathbb{R}^n$ w $\mathbb{R}^m$. Natomiast f' jest funkcją, która odwzorowuje E w przestrzeń wektorową przekształceń liniowych $\mathbb{R}^n \to \mathbb{R}^m$.

(c) Pochodna zdefiniowana przez (116) lub (118) jest często nazywana różniczką f w x lub pochodną zupełną f w x.

Wniosek 45. *Niech E będzie otwartym podzbiorem przestrzeni $\mathbb{R}^n$. Jeśli odwzorowanie $f : E \to \mathbb{R}^m$ jest różczkowalne w punkcie x, to jest ono ciągłe w tym punkcie.*

Dowód. Ponieważ odwzorowanie $f'(x)$ jest ciągłe (zob. Tw. 205), zatem wobec (118) mamy

$$\lim_{h\to 0} f(x+h) = \lim_{h\to 0}(f(x) + f'(x)h + r(h)) = f(x),$$

bowiem z równości

$$\lim_{h\to 0} \frac{||r(h)||}{||h||} = 0$$

wynika, że $\lim_{h\to 0} r(h) = 0$.
(Dla $||h|| < 1$ mamy $||r(h)|| \leqslant \frac{||r(h)||}{||h||}$, a więc $\lim_{h\to 0} r(h) = 0$, czyli wobec ciągłości normy $r(h) \to 0$ przy $h \to 0$).

Zajmiemy się teraz podstawowymi własnościami pochodnych funkcji wielu zmiennych. Następujące twierdzenie orzeka liniowość różniczkowania.

Twierdzenie 208. *Niech E będzie otwartym podzbiorem przestrzeni $\mathbb{R}^n$ i niech odwzorowania $f_1 : E \to \mathbb{R}^m$, $f_2 : E \to \mathbb{R}^m$ będą różniczkowalne w punkcie $x \in E$. Wówczas liniowa kombinacja $(\alpha_1 f_1 + \alpha_2 f_2) : E \to \mathbb{R}^m$ $(\alpha_1, \alpha_2 \in \mathbb{R})$ jest również odwzorowaniem różniczkowalnym w tym punkcie oraz zachodzi równość*

$$(\alpha_1 f_1 + \alpha_2 f_2)'(x) = (\alpha_1 f_1' + \alpha_2 f_2')(x).$$

Dowód. Dla dostatecznie małych $h \in \mathbb{R}^n$ mamy

tucos

Ponieważ

$$\lim_{h\to 0} \frac{||r_1(h)||}{||h||} = \lim_{h\to 0} \frac{||r_2(h)||}{||h||} = 0,$$

zatem

$$\lim_{h\to 0} \frac{||\alpha_1 r_1(h) + \alpha_2 r_2(h)||}{||h||} = 0.$$

Wobec liniowości przekształcenia $\alpha_1 f_1'(x) + \alpha_2 f_2'(x)$ oraz Uwagi 71 (a) dowód jest zakończony.

Dla funkcji rzeczywistych zachodzi następujące.

Twierdzenie 209. *Jeśli funkcje $f : E \to \mathbb{R}, g : E \to \mathbb{R}$ określone na zbiorze otwartym $E \subset \mathbb{R}^n$ są różniczkowalne w punkcie $x \in E$, to funkcje $fg, \frac{f}{g}$ (przy założeniu $g(x) \neq 0$) są różniczkowalne w punkcie x oraz*

(a) $(fg)'(x) = g(x)f'(x) + f(x)g'(x)$,

(b) $\left(\frac{f}{g}\right)'(x) = \frac{1}{g^2(x)}(g(x)f'(x) - f(x)g'(x))$.

Dowód. Tego ma nie być na egzaminie i brakuje mi zdjęcia. Pozdrawiam.

Następujące twierdzenie rozszerza na przypadek wielu zmiennych regułę różniczkowania funkcji złożonej (zob. Tw. 55).

Twierdzenie 210. *Niech E będzie niepustym podzbiorem otwartym przestrzeni $\mathbb{R}^n$ i niech $f : E \to \mathbb{R}^m$ będzie różniczkowalna w punkcie $x_0 \in E$. Ponadto* *załóżmy, że funkcja g odwzorowuje pewien zbiór otwarty zawierający $f(E)$ w przestrzeń $\mathbb{R}^k$ i jest różniczkowalna w punkcie $f(x_0)$. Wówczas funkcja złożona $g \circ f : E \to \mathbb{R}^k$ jest różniczkowalna w punkcie x_0 i zachodzi równość*

$$(g \circ f)'(x_0) = g'(f(x_0)) \circ f'(x_0).$$

Dowód. Niech $y_0 = f(x_0)$ i $A = f'(x_0)$ oraz $B = g'(y_0)$. Oznaczmy

$$u(h) = f(x_0 + h) - f(x_0) - Ah, \quad v(k) = g(y_0 + k) - g(y_0) - Bk$$

dla dowolnych $h \in \mathbb{R}^n$ i $k \in \mathbb{R}^m$, przy których $f(x_0 + h)$ oraz $g(x_0 + k)$ są określone. Wtedy na mocy (118)

$$||u(h)|| = ||\varepsilon(h)||||h||, \quad ||v(k)|| = ||\eta(k)||||k||, \tag{119}$$

gdzie $||\varepsilon(h)|| \to 0$ przy $h \to 0$ oraz $||\eta(k)|| \to 0$ przy $k \to 0$. Przy danym h, niech $k = f(x_0 + h) - f(x_0)$. Wtedy

$$||k|| = ||Ah + u(h)|| \leqslant (||A|| + ||\varepsilon(h)||)||h|| \tag{120}$$

oraz

$$\begin{aligned} g(f(x_0 + h)) - g(f(x_0)) - B(Ah) &= g(y_0 + k) - g(y_0) - B(Ah) \\ &= B(k - Ah) + v(k) \\ &= Bu(h) + v(k). \end{aligned}$$

Wobec tego na mocy (119) i (120) dla $h \neq 0$ otrzymujemy

$$\begin{aligned} \frac{||g(f(x_0 + h)) - g(f(x_0)) - B(Ah)}{||h||} &= \frac{||Bu(h) + v(k)||}{||h||} \\ &\leqslant ||B||||\varepsilon(h)|| + (||A|| + ||\varepsilon(h)||)||\eta(k)||. \end{aligned}$$

Jeżeli $h \to 0$, to $||\varepsilon(h)|| \to 0$, a także na mocy (120) $k \to 0$, więc też $||\eta(k)|| \to 0$. Wynika stąd, że $(g \circ f)'(x_0) = BA$, co należało dowieść. **Twierdzenie 211.** *Niech E będzie podzbiorem otwartym przestrzeni $\mathbb{R}^n$ i niech odwzorowanie $f : E \to \mathbb{R}^m$ ma składowe $f_1, \dots, f_m$ ($f = (f_1, \dots, f_m)$). Wówczas f jest różniczkowalne w punkcie $x_0 \in E$ wtedy i tylko wtedy, gdy każda z funkcji f_i jest różniczkowalna w tym punkcie. Ponadto prawdzimy jest wzór*

$$f'(x_0) = (f_1'(x_0), \dots, f_m'(x_0)). \tag{121}$$

Dowód. Równość (118) dla $x = x_0$ jest równoważna następujacemu układowi równości

$$f_i(x_0 + h) - f_i(x_0) = A_i h + r_i(h), \quad i = 1, \dots, m,$$

gdzie $A = (A_1, \dots, A_m), r = (r_1, \dots, r_m)$ oraz $x + h \in E$. Jest jasne że przekształcenie A jest liniowe wtedy i tylko wtedy, gdy każde przekształcenie A_i jest liniowe. Ponadto z nierówności

$$|r_i(h)| \leqslant \Big(\sum_{i=1}^m |r_i(h)|^2 \Big)^{1/2} = ||r(h)||$$

wynika, że

$$\lim_{h\to 0} \frac{||r(h)||}{||h||} = 0$$

wtedy i tylko wtedy gdy

$$\lim_{h\to 0} \frac{||r_i(h)||}{||h||} = 0$$

dla każdego $i = 1, \ldots, m$. Stąd przekształcenie f jest różniczkowalne w punkcie x_0 wtedy i tylko wtedy, gdy każda z funkcji f_i jest różniczkowalna w tym punkcie. Ponadto, jeśli A jest pochodną f w punkcie x_0, to A_i jest pochodną f_i w tym punkcie. Stąd (121), co kończy dowód.

Definicja 144. Niech E, x, i f będą takie same jak w Definicji 143. Jeżeli f jest różniczkowalna w punkcie x, to macierz odwzorowania $f'(x)$ względem baz standardowych w przesztrzeniach $\mathbb{R}^n$ i $\mathbb{R}^m$ nazywamy macierzą Jacobiego w punkcie x i oznaczamy $[f'(x)]$. Wyznacznik macierzy Jacobiego nazywamy jakobianem i oznaczamy symbolem $J_f(x)$ lub $\frac{\partial(y_1,\ldots,y_n)}{\partial(x_1,\ldots,x_n)}$, gdzie $(y_1, \ldots, y_m) = f(x_1, \ldots, x_n)$.

Przypomnijmy teraz definicje zbioru wypukłego.

Definicja 145. Podzbiór E przestrzeni liniowej X nazywamy wypukłym, jeżeli

$$\alpha x + (1-\alpha)y \in E$$

dla każdego $x \in E$, $y \in E$ i $0 < \alpha < 1$.

Jako wniosek z Twierdzenia 210 otrzymujemy następujące twierdzenie o wartości średniej.

Twierdzenie 212. *Niech $E \subset \mathbb{R}^n$ będzie zbiorem otwartym i wypukłym. Jeśli funkcja $f : E \to \mathbb{R}$ jest różniczkowalna, to dla dowolnych $x \in E$, $x + h \in E$ istnieje $0 < \theta < 1$ takie, że*

$$f(x+h) - f(x) = f'(x+\theta h)h.$$

Dowód. Niech $F(t) = f(x + th)$ dla $t \in [0, 1]$. Wówczas na mocy założenia oraz twierdzenia o różniczkowaniu funkcji złożonej stwierdzamy, że do funkcji F możemy zastosować Twierdzenie Lagrange'a o wartości średniej (Tw. 62). Istnieje zatem $0 < \theta < 1$ takie, że

$$F(1) - F(0) = F'(\theta).$$

Zauważmy, teraz że $F(1) = f(x + h)$, $F(0) = f(x)$ oraz $F'(\theta) = f'(x + \theta h)h$. Dla funkcji o wartościach w $\mathbb{R}^m$ twierdzenie o wartości średniej przyjmuje nastepującą postać.

Twierdzenie 213. *Niech f będzie funkcją określoną i różniczkowalna na otwartym i wypukłym zbiorze $E \subset \mathbb{R}^n$ odwzorowującą E w $\mathbb{R}^m$. Jeśli istnieje liczba* *nieujemna M taka, że $||f'(x)|| \leqslant M$ dla każdego $x \in E$, to dla dowolnych $a, b \in E$*

$$||f(b) - f(a)|| \leqslant M||b - a||.$$

Dowód. Ustalmy $a, b \in E$. Określmy $\gamma(t) = (1-t)a+tb$ dla $t \in [0,1]$. Ponieważ E jest zbiorem wypukłym, więc $\gamma(t) \in E$ dla $0 \leqslant t \leqslant 1$. Niech $F(t) = f(\gamma(t))$, $t \in [0,1]$. Wówczas

$$F'(t) = f'(\gamma(t))\gamma'(t) = f'(\gamma(t))(b-a)$$

i wobec tego dla $t \in [0,1]$ mamy

$$||F'(t)|| = ||f'(\gamma(t))||||b-a|| \leqslant M||b-a||.$$

Na mocy Twierdzenia 73, $||F(1) - F(0)|| \leqslant M||b-a||$, ale $F(0) = f(a)$ oraz $F(1) = f(b)$, co kończy dowód. (Twierdzenie 73 zostało udowodnione dla funkcji o wartościach w $\mathbb{C}$, ale analogiczne rozumowanie dowodzi jego prawdziwości dla funkcji o wartościach w $\mathbb{R}^n$).

Wniosek 46. *Jeśli w założeniach poprzedniego twierdzenia dodamy warunek $f'(x) = 0$ dla dowolnego $x \in E$, to f jest funkcją stałą. Wystarczy bowiem zauważyć, że założenia Twierdzenia 213 są spełnione dla $M = 0$.*
Można udowodnić, iż wniosek ten pozostaje prawdzimy jeśli założymy, że zbiór E jest obszarem, to znaczy jest on otwarty i spójny (zob. ćw). Wiadomo, że każdy zbiór wypukły jest spójny (zob. [3], Tw. 4, s. 93); odwrotne zachodzić nie musi.