Szanowni Państwo, poniżej jest następna porcja notatek z wykładu. Dowód warunku koniecznego i dostatecznego całkowalności w sensie Riemanna jest nieco prostszy niż ten z wykładu, mam nadzieję, że poprawny. Będę wdzięczny za informacje o zauważonych błędach, również tzw. głupich błędach, literówkach itp. Poprawię je w miarę szybko, co może ułatwić leturę następnym czytelnikom, np. kolegom z grupy. Mam nadzieję, że uda mi się w miarę szybko wyprodukować ciąg dalszy.

Pamiętnik z wykładu z AM1, część siódma

Całka Riemanna

Przypomnijmy definicję

Definicja funkcji całkowalnej w sensie Riemanna

Funkcja $f\colon [a,b] \longrightarrow \mathbb{R}$ jest całkowalna w sensie Riemanna wtedy i tylko wtedy, gdy istnieje liczba rzeczywista I, taka że dla każdej liczby $\varepsilon > 0$ istnieje liczba $\delta > 0$, taka że jeżeli $a = x_0 < x_1 < < x_2 < \ldots < x_{n-1} < x_n = b$, $x_{i-1} \leq t_i \leq x_i$ oraz $t_i - t_{i-1} < \delta$ dla $i = 1, 2, \ldots, n$, to zachodzi nierówność

$$\left| I - \Big(f(t_1)(x_1 - x_0) + f(t_2)(x_2 - x_1) + \ldots + f(t_n)(x_n - x_{n-1}) \Big) \right| < \varepsilon \,.$$

Liczba I nazywana jest wtedy całką Riemanna funkcji f na przedziale $[a,b]$ i oznaczana symbolem $\displaystyle\int_a^b f(x)\,dx$. ■

Z tej definicji wynika od razu, że jeżeli funkcja f jest całkowalna w sensie Riemanna, to jest ograniczona. Jeśli bowiem ustalimy punkty $a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$ dostatecznie drobnego podziału przedziału $[a,b]$, to zmieniając odpowiednio punkt t_i możemy dowolnie zwiększyć wartość bezwzględną składnika $f(t_i)(x_i - x_{i-1})$ zachowując jednocześnie wszystkie inne punkty (funkcja nieograniczona na całym przedziale $[x_0, x_n]$ musi być nieograniczona na co najmniej jednym z przedziałów $[x_0, x_1]$, $[x_1, x_2]$,$\ldots$, $[x_{n-1}, x_n]$). Wykazaliśmy więc, że

Twierdzenie

Funkcja całkowalna w sensie Riemanna jest ograniczona. ■

Niestety istnieją funkcje ograniczone, które całkowalne w sensie Riemanna nie są. Przyjąwszy np.

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

otrzymujemy funkcję niecałkowalną w sensie Riemanna, bo wybierając wymierne $t_1, \ldots, t_n$ otrzymujemy $f(t_1)(x_1 - x_0) + f(t_2)(x_2 - x_1) + \ldots + f(t_n)(x_n - x_{n-1}) = b - a$, zaś dla niewymiernych $t_1, \ldots, t_n$ otrzymujemy $f(t_1)(x_1 - x_0) + f(t_2)(x_2 - x_1) + \ldots + f(t_n)(x_n - x_{n-1}) = 0$ niezależnie od tego jak drobno podzielony został przedział $[a,b]$. Nie ma więc kandydata na całkę. Sformułujemy i udowodnimy twierdzenie charakteryzujące funkcje całkowalne w sensie Riemanna, ale musi to być poprzedzone definicją uogólniającą pojęcie długości przedziału i kilkoma twierdzeniami na ten temat. Zaczniemy od dowodu bardzo ważnego twierdzenia o pokryciach przedziału domkniętego przedziałami otwartymi.

Twierdzenie o liczbie Lebesgue'a

Jeśli $\{I_t\colon \ t \in T\}$ jest pewną rodziną przedziałów otwartych, która pokrywa przedział domknięty $[a,b]$, tzn. $\bigcup_{t\in T} I_t \supset [a,b]$, to istnieje liczba $\lambda > 0$ taka, że jeśli zbiór $A \subset [a,b]$ ma średnicę mniejszą lub równa niż λ (czyli odległość każdych dwóch punktów zbioru A jest mniejsza lub równa λ), to istnieje $t(A) \in T$ takie, że $A \subset I_{t(A)}$ (mały zbiór musi być zawarty w jakimś, niekoniecznie jednym, elemencie pokrycia przedziałami otwartymi).

Dowód.

Załóżmy, że teza nie jest prawdziwa. Wtedy dla każdej liczby naturalnej n istnieje zbiór A_n, który nie jest zawarty w żadnym z przedziałów I_t i którego średnica jest mniejsza niż $\frac{1}{n}$. Niech $a_n \in A_n$. Z ciągu (a_n) można wybrać podciąg zbieżny (a_{n_k}). Niech $p = \lim_{k\to\infty} a_{n_k}$. Niech $p \in I_{t(p)}$, indeks $t(p)$ istnieje, bo $\bigcup_{t\in T} I_t \supset [a,b] \ni p$. Ponieważ $I_{(t(p))}$ jest przedziałem *otwartym*, więc istnieje liczba $\delta > 0$ taka, że $(p-\delta, p+\delta) \subset I_{t(p)}$. Dla dostatecznie dużych k mamy $a_{n_k} \in (p - \frac{\delta}{2}, p + \frac{\delta}{2})$. Stąd jednak wynika, że dla każdego $x \in A_{n_k}$ zachodzi nierówność $|x-p| \le |x-a_{n_k}| + |a_{n_k} - p| < \dfrac{1}{n_k} + \dfrac{\delta}{2}$. Oczywiście dla dostatecznie dużych k zachodzi też nierówność $\dfrac{1}{n_k} < \dfrac{\delta}{2}$ i wobec tego $|x-p| < \delta$. To jednak oznacza, że $A_{n_k} \subset (p-\delta, p+\delta) \subset I_{t(p)}$ wbrew temu, że zbiór A_{n_k} nie jest zawarty w żadnym z przedziałów I_t. Dowód został zakończony. ■

Komentarz: *W dowodzie wykorzystywaliśmy jedynie jedną własność przedziału domkniętego, mianowicie to, że z każdego ciągu punktów przedziału domkniętego można wybrać podciąg zbieżny do granicy znajdującej się w* **tym** *przedziale. Własność ta nie przysługuje przedziałom otwartym, np. z ciągu* $(\frac{1}{n+10})$ *punktów przedziału* $(0,1)$ *nie da się wybrać podciągu zbieżnego do granicy leżącej w przedziale* $(0,1)$*, bowiem ciąg ten jest zbieżny do punktu* $0 \notin [0,1]$.

Definicja zbioru zwartego*

Zbiór $C \subset \mathbb{R}$ jest zwarty wtedy i tylko wtedy, gdy z każdego ciągu punktów (p_n) zbioru C można wybrać podciąg (p_{n_k}) zbieżny do pewnego punktu $p \in C$. ■

Łatwo można zauważyć, że suma skończenie wielu przedziałów domkniętych jest zbiorem zwartym. Zwarty jest też zbiór Cantora, czyli zbiór złożony z tych punktów przedziału domkniętego $[0,1]$, które można zapisać w układzie trójkowym bez użycia cyfry 1. Wynika to łatwo z oczywistego stwierdzenia: jeśli z przedziału domkniętego usuniemy sumę dowolnej rodziny przedziałów otwartych, to otrzymamy zbiór zwarty. Jasne jest też, że zbiór zwarty musi być ograniczony: w innym przypadku moglibyśmy znaleźć ciąg punktów tego zbioru, którego granica byłaby nieskończona, więc

* Nie jest to najbardziej ogólna definicja, ale wystarczająca dla naszych celów, z innymi studenci zapoznają się na zajęciach z topologii.

nie moglibyśmy z niego wybrać podciągu zbieżnego do punktu w **tym** zbiorze. Wobec tego każdy zbiór zwarty $C \subset \mathbb{R}$ musi być podzbiorem pewnego przedziału domkniętego $[a,b]$. Jasne jest też, że żaden punkt zbioru $[a,b] \setminus C$ nie może być granicą ciągu punktów zbioru C: jeśli $\lim\limits_{n\to\infty} c_n = p \notin C$ i $c_n \in C$, to wszystkie podciągi ciągu (c_n) są zbieżne do punktu p, wbrew temu, że z ciągu (c_n) można wybrać podciąg zbieżny do granicy należącej do C. W ten sposób wykazaliśmy, że prawdziwe jest

Twierdzenie charakteryzujące podzbiory zwarte $\mathbb{R}$

Zbiór $C \subset \mathbb{R}$ jest zwarty wtedy i tylko wtedy, gdy jest ograniczony i gdy jego dopełnienie $\mathbb{R} \setminus C$ jest sumą pewnej rodziny przedziałów otwartych. ■

Z tekstu zawartego w komentarzu po dowodzie twierdzenia o liczbie Lebesgue'a wynika od razu, że prawdziwe jest

Twierdzenie o liczbie Lebesgue'a dla zbioru zwartego

Jeśli $\{I_t\colon \quad t \in T\}$ jest pewną rodziną przedziałów otwartych, która pokrywa zbiór zwarty C, tzn. $\bigcup\limits_{t\in T} I_t \supset C$, to istnieje liczba $\lambda > 0$ taka, że jeśli zbiór $A \subset C$ ma średnicę mniejszą lub równa niż λ (czyli odległość każdych dwóch punktów zbioru A jest mniejsza lub równa λ), to istnieje $t(A) \in T$ takie, że $A \subset I_{t(A)}$ (mały zbiór musi być zawarty w jakimś, niekoniecznie jednym, elemencie pokrycia przedziałami otwartymi). ■

Definicja liczby Lebesgue'a

Liczba λ, o której mówi powyższe twierdzenie Lebesgue'a nazywana jest liczbą Lebesgue'a pokrycia $\{I_t\colon \quad t \in T\}$. ■

Twierdzenie Heinego-Borela o pokryciach skończonych

Jeśli $\{I_t\colon \quad t \in T\}$ jest pewną rodziną przedziałów otwartych, która pokrywa przedział domknięty $[a,b]$, tzn. $\bigcup\limits_{t\in T} I_t \supset [a,b]$, to istnieje zbiór *skończony* $T_0 \subset T$ taki, że

$$\bigcup_{t\in T_0} I_t \supset [a,b]$$

czyli z każdego pokrycia przedziału domkniętego przedziałami otwartymi można wybrać podpokrycie skończone tego przedziału.

Dowód.

Niech λ oznacza liczbę Lebesgue'a pokrycia $\{I_t\colon \quad t \in T\}$. Niech $n > \frac{b-a}{\lambda}$ oznacza liczbę naturalną. Niech $x_i = a + \frac{i}{n}(b-a)$ dla $i = 0, 1, \ldots, n$. Z twierdzenia o liczbie Lebesgue'a wynika, że dla każdego z przedziałów $[x_{i-1}, x_i]$ istnieje $t(i) \in T$ takie, że $[x_{i-1}, x_i] \subset I_{t(i)}$ dla $i = 1, 2, \ldots, n$. Stąd oczywiście wynika, że $[a,b] \subset \bigcup\limits_{i=1}^{n} I_{t(i)}$, co kończy dowód. ■

Definicja dowolnie długiej sumy liczb nieujemnych

Niech $\mathcal{A} = \{a_t\colon \quad t \in T\}$ oznacza pewien zbiór złożony z liczb nieujemnych. Piszemy

$$\sum \mathcal{A} = \sum_{t\in T} a_t = \sup\Big\{\sum_{t\in\tilde{T}} a_t\colon \quad \tilde{T}\subset T, \qquad \tilde{T} \text{ –zbiór skończony}\Big\} \ \blacksquare$$

Mówiąc słowami (niedokładnie!): suma dowolnego zbioru liczb dodatnich równa jest kresowi górnemu sum skończenie wielu jego elementów. Niedokładność polega na tym, że sumujemy nie elementy zbioru liczbowego, lecz indeksowane elementy. Jeśli jakaś liczba jest przypisana np. trzem różnym indeksom t, to liczymy ją trzy razy a nie raz jakby to miało miejsce w przypadku sumowania elementów zbioru.

Jasne jest, że jeśli T oznacza zbiór złożony z kolejnych liczb całkowitych większych od $k-1$, to $\sum\limits_{t\in T} a_t$ jest po prostu sumą szeregu $\sum\limits_{n=k}^{\infty} a_n$, więc obecna definicja ma jedynie rozszerzyć zakres poprzedniej, która wymagała uporządkowania zbioru numerów liczb dodawanych.

Stwierdzenie o liczbie składników sumy skończonej

Jeśli $\sum\limits_{t\in T} a_t < \infty$, to zbiór indeksów $t\in T\colon \quad a_t > 0$ jest skończony lub co najwyżej przeliczalny: $\mathrm{card}(\{t\in T\colon \quad a_t>0\}) \le \aleph_0$.*

Dowód.

Załóżmy, że tak nie jest. Wtedy dla pewnej liczby naturalnej n zbiór $T_n := \{t\in T\colon \quad a_t \ge \frac{1}{n}\}$ musi być nieprzeliczalny, więc tym bardziej nieskończony, bo $\{t\in T\colon \quad a_t>0\} = \bigcup_{n=1}^{\infty} T_n$ .Wtedy jednak $+\infty \le \sum_{t\in T_n} \le \sum_{t\in T} a_t$ wbrew założeniu. $\blacksquare$

Stwierdzenie banalne o długości przedziału

Jeśli dla każdego $t\in T$ symbol I_t oznacza przedział dodatniej długości o końcach a_t i b_t oraz $\bigcup\limits_{t\in T} I_t \supset [a,b]$, to $\sum\limits_{t\in T}(b_t - a_t) \ge b - a$.**

Dowód.

Wystarczy dowieść, że teza ma miejsce w przypadku $b - a < \infty$, bo półprosta i prosta mogą być przedstawione w postaci sumy wstępującego ciągu przedziałów skończonych. Możemy w tym przypadku założyć, że $\sum\limits_{t\in T}(b_t - a_t) < \infty$, bo w przypadku przeciwnym nic do dowodu nie ma. Jeśli ta suma jest skończona, to zbiór T jest co najwyżej przeliczalny (poprzednie stwierdzenie). Załóżmy że jest on zbiorem liczb naturalnych. Niech $\varepsilon > 0$ oznacza dowolną liczbę dodatnią. Niech $J_n = (a_n - \frac{\varepsilon}{2^{n+2}}, b_n + \frac{\varepsilon}{2^{n+2}})$. Jasne jest, że $\sum\limits_n \big(b_n + \frac{\varepsilon}{2^{n+2}} - (a_n - \frac{\varepsilon}{2^{n+2}})\big) = \varepsilon + \sum\limits_n (b_n - a_n)$.

* symbol $\mathrm{card}(A)$ oznacza liczbę elementów zbioru A, czyli jego moc, używane na innych przedmiotach oznaczenie $|A|$ oznaczać będzie już niedługo miarę zbioru.

** Nie precyzujemy czy przedziały są otwarte domknięte, czy otwarto–domknięte, czy skończone, czy nieskończone, a_t oznacza lewy koniec, b_t – prawy.

Wobec tego, że ε oznacza dowolną liczbę dodatnią wystarczy wykazać tezę dla rodziny $\{J_n\}$: jeśli $\varepsilon+\sum_n(b_n-a_n)\geq b-a$ dla każdego $\varepsilon>0$, to również $\sum_n(b_n-a_n)\geq b-a$. Przedziały J_n są otwarte, więc istnieje liczba $\lambda>0$ taka, że każdy przedział o długości $<\lambda$ jest zawarty w pewnym przedziale J_n. Niech $a=x_0<x_1<\ldots<x_{m-1}<x_m=b$ oznaczają takie punkty, że $x_i-x_{i-1}<\lambda$ dla $i=1,2\ldots,m$. Dla każdego i wybieramy numer $n(i)$ w taki sposób, że $[x_{i-1},x_i]\subset J_{n(i)}$. Wykażemy, że suma długości przedziałów $J_{n(1)}$, $J_{n(2)}$, $\ldots, J_{n(m)}$ jest większa niż $b-a$, z czego teza wyniknie od razu. Mamy $[x_0,x_1]\subset J_{n(1)}$. Niech i_1 będzie największym numerem takim, że $[x_0,x_{i_1}]\subset J_{n(1)}$. Ponieważ $J_{n(1)}$ jest przedziałem więc punkty x_i dla $i>i_1$ znajdują się poza $J_{n(1)}$. Oczywiście $[x_{i_1},x_{i_1+1}]\subset J_{n(i_1+1)}$. Niech teraz i_2 będzie największym numerem takim, że $[x_{i_1},x_{i_2}]\subset J_{n(i_1+1)}$. Tak jak poprzednio dla $i>i_2$ punkt x_i znajduje się poza przedziałem $J_{n(i_1+1)}$. Definiując kolejno numery $0=i_0$, i_1, i_2,$\ldots,i_k$ otrzymujemy ciąg numerów taki, że $[x_{i_{j-1}},x_{i_j}]\subset J_{n(i_{j-1}+1)}$, $i_k=m$, $x_{i_j+1}\notin J_{n(i_{j-1}+1)}$. Wobec tego numery $n(i_0+1)$, $n(i_1+1)$,$\ldots$, $n(i_{k-1}+1)$ są różne. Liczba $x_{i_j}-x_{i_{j-1}}$ jest mniejsza niż długość przedziału $J_{n(i_{j-1}+1)}$. Wobec tego $b-a=\sum(x_{i_j}-x_{i_{j-1}})$ jest liczbą mniejszą niż suma długości przedziałów $J_{n(i_0+1)}$, $J_{n(i_1+1)}$, $\ldots$, $J_{n(i_{k-1}+1)}$, tu korzystamy z tego, że przedziały $J_{n(i_0+1)}$, $J_{n(i_1+1)}$, $\ldots$, $J_{n(i_{k-1}+1)}$ są *różne* (przedziały $J_{n(1)}$, $J_{n(2)}$, $J_{n(m)}$ nie muszą być różne). Ta zadziwiająco skomplikowana – jak na tak oczywiste stwierdzenie – konstrukcja prowadzi do przypisania przedziałom $[x_{i_{j-1}},x_{i_j}]$ różnych przedziałów $J_{n(i_{j-1})}$, co pozwala na zastąpienie sumy długości tych pierwszych sumą długości tych drugich. Dowód został zakończony. ■

Osoby, którym wydaje się, że powyższe rozumowanie jest za długie, że to przerost formy nad treścią zapraszam do jego skrócenia (odrzucając jednak metodę polegającą na opuszczeniu kilku słów lub stwierdzenia, że jest to oczywiste!).

Teraz możemy wprowadzić zdefiniować miarę (zewnętrzną) zbioru $A\subset\mathbb{R}$.

Definicja miary zewnętrznej

$|A|=\inf_{t\in T}\{\sum|I_t|\colon\ \bigcup_{t\in T}I_t\supset A\}$, gdzie I_t oznacza przedział niezdegenerowany a $|I_t|$ jego długość, czyli różnicę końców. Liczbę $|A|$ nazywamy miarą zewnętrzną zbioru A.■

Dzięki *stwierdzeniu banalnemu o długości przedziału* definicja ta nie prowadzi do zamieszania w oznaczeniach: Jeśli zbiór A jest przedziałem, to $|A|$ jest jego długością!

Zupełnie oczywiste jest

Twierdzenie o monotoniczności miary zewnętrznej

Jeśli $A\subset B$, to $|A|\leq|B|$. ■

Twierdzenie o podaddytywności miary zewnętrznej

Dla dowolnych zbiorów $A_1,A_2,\ldots$ zachodzi nierówność: $\left|\bigcup_{n=1}^{\infty}A_n\right|\leq\sum_{n=1}^{\infty}|A_n|$.

Dowód.

Niech $\varepsilon>0$ będzie dowolną liczbą. Dla każdego n istnieje rodzina przedziałów $\{I_t\colon\ \ t\in T_n\}$ taka, że $\bigcup_{t\in T_n} I_t \supset A_n$ i $\sum_{t\in T_n} |I_t| \le |A_n| + \frac{\varepsilon}{2^n}$. Niech $T = \bigcup_{n=1}^{\infty} T_n$. Jasne jest, że $\bigcup_{t\in T} I_t \supset \bigcup_{n=1}^{\infty} A_n$ oraz że $\sum_{t\in T} |I_t| = \sum_{n=1}^{\infty}\Big(\sum_{t\in T_n} |I_t|\Big) \le \sum_{n=1}^{\infty}\Big(|A_n| + \frac{\varepsilon}{2^n}\Big) = \sum_{n=1}^{\infty} |A_n| + \varepsilon$. Ponieważ ta nierówność zachodzi dla dowolnej liczby $\varepsilon > 0$, więc $\left|\bigcup_{n=1}^{\infty} A_n\right| \le \sum_{n=1}^{\infty} |A_n|$, a to chcieliśmy wykazać. ■

Oczywiście chciałoby się stwierdzić, że jeśli zbiory $\{A_n\}$ są parami rozłączne, to $\left|\bigcup_{n=1}^{\infty} A_n\right| = \sum_{n=1}^{\infty} |A_n|$. Tego niestety nie można udowodnić bez dodatkowych założeń. W tym roku nie będzie nas ta kwestia interesować, zajmiemy się tym w roku przyszłym, bo wymaga to większej pracy i znalazło miejsce w programie roku drugiego, a do naszych aktualnych celów wystarczy twierdzenie o podaddytywności miary zewnętrznej. Jest natomiast prawdziwe stwierdzenie następujące

Twierdzenie o przeliczalnej addytywności miary zewnętrznej dla przedziałów

Jeśli przedziały I_1, I_2, $I_3, \ldots$ są parami rozłączne (tzn. $I_k \cap I_l = \emptyset$ dla $k \ne l$), to zachodzi równość: $\big|I_1 \cup I_2 \cup I_3 \cup \ldots\big| = \big|I_1\big| + \big|I_2\big| + \big|I_3\big| + \cdots$

Dowód.

Nierówność $\big|I_1 \cup I_2 \cup I_3 \cup \ldots\big| \le \big|I_1\big| + \big|I_2\big| + \big|I_3\big| + \cdots$ wynika wprost z definicji miary. Jasne jest również, że dla każdego n zachodzi nierówność

$$\big|I_1 \cup I_2 \cup \ldots \cup I_n\big| \le \big|I_1 \cup I_2 \cup I_3 \cup \ldots\big| \le \big|I_1\big| + \big|I_2\big| + \big|I_3\big| + \cdots. \tag{*}$$

Jednak ze stwierdzenia banalnego o długości przedziału wynika, że dla dowolnej liczby naturalnej n zachodzi nierówność $\big|I_1 \cup I_2 \cup \ldots \cup I_n\big| = \big|I_1\big| + \big|I_2\big| + \cdots + \big|I_n\big|$ *, a stąd $\big|I_1 \cup I_2 \cup I_3 \cup \ldots\big| \ge \big|I_1 \cup I_2 \cup \ldots \cup I_n\big| = \big|I_1\big| + \big|I_2\big| + \cdots + \big|I_n\big| \xrightarrow[n\to\infty]{} \big|I_1\big| + \big|I_2\big| + \big|I_3\big| + \cdots$.
Wobec tego $\big|I_1 \cup I_2 \cup I_3 \cup \ldots\big| \ge \big|I_1\big| + \big|I_2\big| + \big|I_3\big| + \cdots$, co w połączeniu z nierównością (*) dowodzi twierdzenia. ■

Definicja zbioru miary 0

Mówimy, że zbiór A jest zbiorem miary 0 wtedy i tylko wtedy, gdy $|A| = 0$. ■

Z twierdzenia o podaddytywności miary wynika, że suma przeliczalnej rodziny zbiorów miary 0 jest również zbiorem miary 0. Dla rodzin większej mocy to oczywiście nie jest prawdą: każdy niepusty zbiór jest sumą zbiorów jednopunktowych, a nie każdy ma miarę 0, np. $|[2,5]| = 3 > 0$. W szczególności każdy zbiór przeliczalny ma miarę 0, np. $\mathbb{Q}$. Istnieją też nieprzeliczalne zbiory miary 0.

Przykład (*zbiór Cantora*).

Opiszemy tzw. zbiór Cantora $\mathcal{C}$. Składa się z tych liczb z przedziału $[0,1]$, które można zapisać w układzie trójkowym bez użycia cyfry 1. Zbiór ten otrzymujemy usuwając z przedziału $[0,1]$ kolejno

---

* Jeśli rodzina przedziałów (I_t) pokrywa zbiór $I_1 \cup I_2 \cup \ldots \cup I_n$, to możemy zastąpić przedział I_t przedziałami $I_1 \cap I_t$, $I_2 \cap I_t$, ... $I_n \cap I_t$, oczywiście $|I_1 \cap I_t| + |I_2 \cap I_t| + \ldots + |(I_n \cap I_t)| \le |I_t|$.

przedziały otwarte $(\frac{1}{3},\frac{2}{3})$, $(\frac{1}{9},\frac{2}{9})$, $(\frac{7}{9},\frac{8}{9})$, $(\frac{1}{27},\frac{2}{27})$, $(\frac{7}{27},\frac{8}{27})$, $(\frac{19}{27},\frac{20}{27})$, $(\frac{25}{27},\frac{26}{27})$, $\ldots$: za każdym razem usuwamy z jakiegoś przedziału jego środkową część, której długość to $\frac{1}{3}$ długości przedziału, z którego ją usuwamy. Otrzymujemy zbiór, który nie zawiera żadnego przedziału. liczba $\frac{1}{3}$ jest jego elementem, chociaż w układzie trójkowym zwykle zapisujemy ją w postaci $0{,}1$. Można ją jednak zapisać w układzie trójkowym jako $0{,}02222222\ldots$, bowiem $\frac{2}{9}+\frac{2}{27}+\frac{2}{81}+\cdots=\frac{2/9}{1-1/3}=\frac{1}{3}$. Zbiór Cantora jest mocy kontinuum, bo ma dokładnie tyle elementów ile jest ciągów, których elementami są 0 i 2. Jest on miary 0, bo można go pokryć 2^n przedziałami domkniętymi o długościach $\frac{1}{3^n}$, więc jego miara nie przekracza liczby $\frac{2^n}{3^n}\xrightarrow[n\to\infty]{}0$. ■

Zadanie. Skonstruować funkcję $f\colon\mathbb{R}\longrightarrow\mathbb{R}$, która przekształca **każdy** przedział na zbiór wszystkich liczb rzeczywistych. ■

Teraz możemy już sformułować twierdzenie opisujące funkcje całkowalne w sensie Riemanna.

Twierdzenie charakteryzujące funkcje całkowalne w sensie Riemanna

Funkcja $f\colon[a,b]\longrightarrow\mathbb{R}$ jest całkowalna w sensie Riemanna wtedy i tylko wtedy, gdy jest ograniczona i jej zbiór punktów nieciągłości ma miarę 0. ☒

Przed dowodem tego twierdzenia sformułujemy warunek typu warunku Cauchy'ego dla zbieżności występującej w definicji całki Riemanna i wykażemy, że jest on konieczny i dostateczny dla jej istnienia. Będziemy potrzebować kilku terminów.

Oznaczenia:

$a=x_0<x_1<x_2<\ldots<x_{n-1}<x_n=b$, punkty $x_0,x_1,\ldots,x_n$ nazywamy węzłami podziału przedziału $[a,b]$, największą z liczb x_1-x_0, $x_2-x_1,\ldots,x_n-x_{n-1}$ nazywamy średnicą podziału;
$m_j=\inf\{f(t)\colon\quad x_{j-1}\le t\le \mathbf{x}_j\}$, $M_j=\sup\{f(t)\colon\quad x_{j-1}\le t\le \mathbf{x}_j\}$;
liczba $\omega_j=M_j-m_j$ oscylacją funkcji f na przedziale $[x_{j-1},x_j]$;
liczba $\omega_f(x)=\inf\limits_{\delta>0}\left(\sup\limits_{|t-x|\le\delta}f(t)-\inf\limits_{|t-x|\le\delta}f(t)\right)$ nazywana jest oscylacją funkcji f w punkcie x;
suma $\sum\limits_{j=1}^{n}M_j(x_j-x_{j-1})$ nazywana jest sumą górną Darboux funkcji f;
suma $\sum\limits_{j=1}^{n}m_j(x_j-x_{j-1})$ nazywana jest sumą dolną Darboux funkcji f. ■

Warto zauważyć, że jeśli podział $a=x_0<x_1<x_2<\ldots<x_{n-1}<x_n=b$ zastąpimy drobniejszym, tzn. do węzłów $x_0,x_1,x_2,\ldots,x_{n-1},x_n$ dodamy nowe, to górna suma Darboux zmniejszy się lub zachowa swą wartość, a dolna – przeciwnie – zwiększy lub zachowa, zatem oscylacja na przedziale ulegnie zmniejszeniu lub zachowa swą wartość (oscylacja na podprzedziale nie przekracza wartości oscylacji na przedziale). Wynika to od razu z tego, ze kres górny funkcji może jedynie zmaleć w wyniku zmniejszenia dziedziny, a dolny – wzrosnąć.

Warunek CICR *typu Cauchy'ego istnienia całki Riemanna*

Dla każdej liczby $\varepsilon>0$ istnieje liczba $\delta>0$ taka, że jeśli dla każdego $j=1,2,\ldots,n$ zachodzi nierówność $x_j - x_{j-1} < \delta$, to $\sum_{j=1}^{n} \omega_j (x_j - x_{j-1}) < \varepsilon$. ■

Twierdzenie o istnieniu całki Riemanna

Funkcja $f\colon [a,b] \longrightarrow \mathbb{R}$ jest całkowalna w sensie Riemanna wtedy i tylko wtedy, gdy zachodzi warunek CICR.

Dowód.

Jeśli funkcja f jest całkowalna w sensie Riemanna, to dla każdego $\varepsilon > 0$ istnieje $\delta > 0$ taka, że jeśli $x_{j-1} \le t_j \le x_j$, to $\left| \int_a^b f(x)\,dx - \sum_{j=1}^{n} f(t_j)(x_j - x_{j-1}) \right| < \frac{\varepsilon}{3}$. Stąd natychmiast wynika, że $\left| \int_a^b f(x)\,dx - \sum_{j=1}^{n} m_j(x_j - x_{j-1}) \right| \le \frac{\varepsilon}{3}$ oraz $\left| \int_a^b f(x)\,dx - \sum_{j=1}^{n} M_j(x_j - x_{j-1}) \right| \le \frac{\varepsilon}{3}$. Wobec tego $\sum_{j=1}^{n} \omega_j(x_j - x_{j-1}) = \left| \sum_{j=1}^{n} M_j(x_j - x_{j-1}) - \sum_{j=1}^{n} m_j(x_j - x_{j-1}) \right| \le 2\frac{\varepsilon}{3} < \varepsilon$, zatem funkcja f spełnia warunek CICR.

Załóżmy teraz, że funkcja f spełnia warunek CICR. W niejawny sposób zakładamy, tu że funkcja f jest ograniczona, bowiem można wykonać odejmowanie sum Darboux, więc nie mogą one być jednocześnie równe $+\infty$ ani też $-\infty$, muszą być skończone, bo ich różnica jest mniejsza np. od 1 dla dostatecznie drobnego podziału, oczywiście $\sup\limits_{t\in[a,b]} f(t) = \max\limits_{j=1,2,\ldots,n} M_j$, $\inf\limits_{t\in[a,b]} f(t) = \min\limits_{j=1,2,\ldots,n} m_j$.
Niech D_n oznacza górną sumę Darboux funkcji f dla podziału przedziału $[a,b]$ na 2^n równych podprzedziałów a d_n wartość dolnej sumy Darboux dla tego podziału. Ciąg (D_n) jest oczywiście nierosnący (rozdrabniamy podziały więc górna suma Darboux nie może wzrosnąć), a ciąg (d_n) jest niemalejący. Oba mają więc granice. Ponieważ dla każdego n zachodzi nierówność $d_n \le D_n$, więc obie te granice są skończone. Muszą one być równe, bo z warunku CICR wynika od razu, że $\lim\limits_{n\to\infty}(D_n - d_n) = 0$. Oznaczmy tę wspólną granicę przez I. Wykażemy, że I jest całką Riemanna funkcji f. Mamy oczywiście $D_n \ge I \ge d_n$ dla każdej liczby naturalnej n. Niech $\varepsilon > 0$ oznacza dowolną liczbę i niech n będzie tak dużą liczbą naturalną, że $D_n - d_n < \frac{\varepsilon}{3}$. Niech $\delta > 0$ będzie liczbą taką, że jeśli $x_j - x_{j-1} < \delta$, to $\sum_{j=1}^{n} \omega_j(x_j - x_{j-1}) < \frac{\varepsilon}{3}$, niech $x_{j-1} \le t_j \le x_j$. Oczywiście $\sum_{j=1}^{n} m_j(x_j - x_{j-1}) \le \sum_{j=1}^{n} f(t_j)(x_j - x_{j-1}) \le \sum_{j=1}^{n} M_j(x_j - x_{j-1})$. Utwórzmy z punktów $x_0, x_1, \ldots, x_n$ i węzłów podziału przedziału $[a,b]$ na 2^n równych podprzedziałów nowy podział przedziału $[a,b]$. Niech punkty $a = z_0, z_1, \ldots, z_m = b$ oznaczają węzły tego nowego podziału, d jego sumę dolną Darboux, D – sumę górną Darboux tego podziału. Oczywiście mamy $d_n \le d \le D \le D_n$ oraz $\sum_{j=1}^{n} m_j(x_j - x_{j-1}) \le d \le D \le \sum_{j=1}^{n} M_j(x_j - x_{j-1})$, bo w wyniku rozdrobnienia podziału suma górna może tylko zmaleć, a dolna wzrosnąć. Ponieważ liczby d, D, I leżą w przedziale $[d_n, D_n]$, więc

$|I-D| \leq |D_n - d_n| < \frac{\varepsilon}{3}$ (i $|I-d| \leq |D_n - d_n| < \frac{\varepsilon}{3}$). Analogicznie $\left|\sum_{j=1}^{n} M_j(x_j - x_{j-1}) - D\right| < \frac{\varepsilon}{3}$ oraz $\left|\sum_{j=1}^{n} M_j(x_j - x_{j-1}) - \sum_{j=1}^{n} f(t_j)(x_j - x_{j-1})\right| < \frac{\varepsilon}{3}$. Możemy więc napisać:

$$\left|I - \sum_{j=1}^{n} f(t_j)(x_j - x_{j-1})\right| \leq |I-D| + \left|D - \sum_{j=1}^{n} M_j(x_j - x_{j-1})\right| +$$

$$+ \left|\sum_{j=1}^{n} M_j(x_j - x_{j-1}) - \sum_{j=1}^{n} f(t_j)(x_j - x_{j-1})\right| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

Wykazaliśmy więc, że $I = \int_a^b f(x)\,dx$. ■

Dowód twierdzenia charakteryzującego funkcje całkowalne w sensie Riemanna

Wiemy już, że jeśli funkcja jest całkowalna w sensie Riemanna, to jest ograniczona. Trzeba jeszcze wykazać, że jej zbiór punktów nieciągłości ma miarę 0. Załóżmy, że $\sum_{j=1}^{n} \omega_j(x_j - x_{j-1}) < \varepsilon$. Niech $\sigma(\alpha)$ będzie sumą tych liczb $x_j - x_{j-1}$, dla których $\omega_j \geq \alpha$. Zachodzi nierówność $\alpha \cdot \sigma(\alpha) \leq \sum_{j=1}^{n} \omega_j(x_j - x_{j-1}) < \varepsilon$, zatem $\sigma(\alpha) < \frac{\varepsilon}{\alpha}$. Niech $\sigma_n(\alpha)$ oznacza sumę długości tych przedziałów z podziału przedziału $[a,b]$ na 2^n równych podprzedziałów, na których oscylacja funkcji f nie jest mniejsza niż α. Z otrzymanej nierówności wynika, że $\sigma_n(\alpha) \xrightarrow[n\to\infty]{} 0$. Stąd natychmiast wynika, że zbiór punktów x, dla których $\omega_f(x) \geq \alpha$ ma miarę 0: węzłów wszystkich podziałów przedziału $[a,b]$ na 2^n podprzedziałów, $n = 1, 2, \ldots$ jest przeliczalnie wiele, więc tworzą one zbiór miary 0. Inne punkty, w których oscylacja nie jest mniejsza niż α znajdują się oczywiście w przedziałach których suma długości jest mniejsza niż $\sigma_n(a) \xrightarrow[n\to\infty]{} 0$. W ten sposób zakończyliśmy dowód twierdzenia w jedną stronę.

Niech teraz $f\colon [a,b] \longrightarrow \mathbb{R}$ oznacza funkcję, C – zbiór punktów, w których jest ona ciągła, $D = [a,b] \setminus C$ – zbiór punktów, w których f jest nieciągła i niech dla każdego $x \in [a,b]$ będzie spełniona nierówność $|f(x)| \leq M < +\infty$. Wykażemy, że jeśli $|D| = 0$, to funkcja f jest całkowalna w sensie Riemanna. Możemy oczywiście zakładać, że $M > 0$, bo jeśli $M = 0$, to funkcja f jest tożsamościowo równa 0, więc jest oczywiście całkowalna w sensie Riemanna.

Niech $\varepsilon > 0$. Dla każdego $x \in C$ istnieje liczba $\delta_x > 0$ taka, że jeśli $|t - x| < \delta_x$ i $t \in [a,b]$, to $|f(t) - f(x)| < \frac{\varepsilon}{4(b-a)}$. Z definicji zbioru miary 0 wynika, że istnieją przedziały I_1, I_2, ... takie, że $\bigcup_{n=1}^{\infty} I_n \supset D$ i $\sum_{n=1}^{\infty} |I_n| < \frac{\varepsilon}{5M}$. Przedziały I_1, I_2, ... mogą być domknięte, otwarto-domknięte, itp. Niech $\tilde{I}_n$ oznacza przedział otwarty, którego środkiem jest środek przedziału I_n i który jest dłuższy od I_n o $\frac{\varepsilon}{20 \cdot M \cdot 2^n}$. Oczywiście $\tilde{I}_n \supset I_n$, więc

$\bigcup\limits_{n=1}^{\infty} \tilde{I}_n \supset D$ i $\sum\limits_{n=1}^{\infty} |\tilde{I}_n| = \sum\limits_{n=1}^{\infty} \Big(|I_n| + \dfrac{\varepsilon}{20 \cdot M \cdot 2^n}\Big) = \sum\limits_{n=1}^{\infty} |I_n| + \sum\limits_{n=1}^{\infty} \dfrac{\varepsilon}{20 \cdot M \cdot 2^n} < \dfrac{\varepsilon}{5M} + \dfrac{\varepsilon}{20M} = \dfrac{\varepsilon}{4M}$.

Rodzina $\mathcal{F} = \{(x - \delta_x, x + \delta_x\colon \quad x \in C)\} \cup \{\tilde{I}_n\colon \quad n \in \mathbb{N}\}$ składa się z przedziałów otwartych, jej suma zawiera przedział $[a,b]$. Wobec tego istnieje liczba $\lambda > 0$ taka, że każdy przedział $[c,d] \subset [a,b]$ długości $\leq \lambda$ jest zawarty w którymś elemencie rodziny $\mathcal{F}$. Niech $a = x_0 < x_1 < \ldots < x_n = b$ oznacza podział przedziału $[a,b]$ na przedziały długości $< \lambda$. Niech N_C oznacza zbiór złożony z tych numerów $j \in \{1, 2, \ldots, n\}$, dla których istnieje punkt $x(j) \in C$ taki, że zachodzi inkluzja $[x_{j-1}, x_j] \subset \big(x(j) - \delta_{x(j)}, x(j) + \delta_{x(j)}\big)$, zaś N_D – zbiór złożony z numerów pozostałych, tj. takich, dla których taki punkt $x(j)$ nie istnieje, zatem istnieć musi liczba $n(j)$ taka, że $[x_{j-1}, x_j] \subset \tilde{I}_{n(j)}$. Jasne jest, że $\bigcup\limits_{j \in N_D} [x_{j-1}, x_j] \subset \bigcup\limits_{j \in N_D} \tilde{I}_n$, zatem $\sum\limits_{j \in N_D} \big|[x_{j-1}, x_j]\big| = \Big|\bigcup\limits_{j \in N_D} [x_{j-1}, x_j]\Big| \leq \sum\limits_n |\tilde{I}_n| < \dfrac{\varepsilon}{4M}$.

Dla każdego j zachodzi nierówność $M_j - m_j \leq 2M$. Ma więc miejsce nierówność

$$\sum_{j \in N_D} \omega_j (x_j - x_{j-1}) \leq 2M \sum_{j \in N_D} (x_j - x_{j-1}) < 2M \cdot \frac{\varepsilon}{4M} = \frac{\varepsilon}{2}.$$

Jeśli $|t - x| < \delta_x$, $x \in C$, to $|f(t) - f(x)| < \dfrac{\varepsilon}{4(b-a)}$. Stąd wynika, że jeśli $|t - x| < \delta_x$ i $|s - x| < \delta_x$, to $|f(t) - f(s)| \leq |f(t) - f(x)| + |f(x) - f(s)| < \dfrac{\varepsilon}{2(b-a)}$. Stąd wnioskujemy, że dla $j \in N_C$ zachodzi nierówność $M_j - m_j \leq \dfrac{\varepsilon}{2(b-a)}$. Dzięki niej możemy napisać:

$$\sum_{j \in N_C} \omega_j (x_j - x_{j-1}) \leq \frac{\varepsilon}{2(b-a)} \cdot \sum_{j \in N_C} (x_j - x_{j-1}) \leq \frac{\varepsilon}{2(b-a)} \cdot (b-a) = \frac{\varepsilon}{2}.$$

Wobec tego możemy napisać:

$$\sum_{j=1}^{n} \omega_j (x_j - x_{j-1}) = \sum_{j \in N_D} \omega_j (x_j - x_{j-1}) + \sum_{j \in N_C} \omega_j (x_j - x_{j-1}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Wykazaliśmy więc, że różnica między sumą górną Darboux i sumą dolną Darboux funkcji f jest mniejsza od ε, jeśli tylko przedział $[a,b]$ został podzielony na dostatecznie krótkie podprzedziały, a to oznacza, że funkcja spełnia warunek CIRC, czyli że jest całkowalna w sensie Riemanna. ■

Z twierdzenia charakteryzującego funkcje całkowalne i tego, że suma i iloczyn funkcji ograniczonych są funkcjami ograniczonymi oraz z tego że suma dwóch zbiorów miary 0 jest zbiorem miary 0 wynika

Twierdzenie

Suma i iloczyn funkcji całkowalnych w sensie Riemanna są funkcjami całkowalnymi w sensie Riemanna. ■

Mamy również łatwe do wywnioskowania twierdzonka

1. $c \displaystyle\int_a^b f(x)\,dx = \int_a^b c f(x)\,dx$ dla dowolnej liczby $c \in \mathbb{R}$ i funkcji $f\colon [a,b] \longrightarrow \mathbb{R}$;

2. $\displaystyle\int_a^b \big(f(x)+g(x)\big)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$

dla dowolnych funkcji $f,g\colon[a,b]\longrightarrow\mathbb{R}$ całkowalnych w sensie Riemanna;

3. jeśli $f(x)\le g(x)$ dla $a\le x\le b$, to $\displaystyle\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx$

dla dowolnych funkcji $f,g\colon[a,b]\longrightarrow\mathbb{R}$ całkowalnych w sensie Riemanna;

4. jeśli $f(x)\le g(x)$ dla $a\le x\le b$ i $\displaystyle\int_a^b f(x)\,dx < \int_a^b g(x)\,dx$ ($f,g\colon[a,b]\longrightarrow\mathbb{R}$ są całkowalne w sensie Riemanna), to zbiór $\{x\in[a,b]\colon\quad f(x)<g(x)\}$ jest miary dodatniej;

5. jeśli $f(x)\le g(x)$ dla $a\le x\le b$ i zbiór $\{x\in[a,b]\colon\quad f(x)<g(x)\}$ jest miary dodatniej ($f,g\colon[a,b]\longrightarrow\mathbb{R}$ są całkowalne w sensie Riemanna), to $\displaystyle\int_a^b f(x)\,dx < \int_a^b g(x)\,dx$;

6. jeśli funkcja $f\colon[a,b]\longrightarrow\mathbb{R}$ jest całkowalna w sensie Riemanna i $a<c<b$, to jest całkowalna w sensie Riemanna na obu przedziałach $[a,c]$ i $[c,b]$ i $\displaystyle\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$;

7. jeśli funkcja $f\colon[a,b]\longrightarrow\mathbb{R}$ jest całkowalna w sensie Riemanna, to również funkcja $|f|$ jest całkowalna w sensie Riemanna i zachodzi nierówność $\displaystyle\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f(x)|\,dx$;

8. jeśli funkcja $f\colon[a,b]\longrightarrow\mathbb{R}$ jest monotoniczna, to jest całkowalna w sensie Riemanna;

9. jeśli funkcje $g,f\colon[a,b]\longrightarrow\mathbb{R}$ są całkowalne w sensie Riemanna i zbiór

$\{x\in[a,b]\colon\quad f(x)\ne g(x)\}$ jest miary 0, to $\displaystyle\int_a^b f(x)\,dx = \int_a^b g(x)\,dx$.

Wiedząc, że całka istnieje możemy we wszystkich przypadkach, w których występują dwie funkcje określone na tym samym przedziale rozpatrywać te same podziały tego przedziału i z nich wybierać te same punkty t_j. Wtedy własności 1,2,3,7 wynikają z odpowiednich twierdzeń o ciągach. Całkowalność funkcji $|f|$ wynika z całkowalności funkcji f, bo jeśli f jest ciągła w pewnym punkcie, to $|f|$ też, z ograniczoności f ograniczoność $|f|$ wynika natychmiast, stąd własność 7. Funkcja monotoniczna na przedziale *domkniętym* jest ograniczona, zbiór jej punktów nieciągłości jest przeliczalny, zatem jest miary 0. Stąd własność 8. Całkowalność na podprzedziale wynika z całkowalności na przedziale od razu, równość $\displaystyle\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$ wynika z tego, że można rozważać tylko te podziały przedziału $[a,b]$, w których c pojawia się jako węzeł. Jeśli zbiór D jest miary zero, to dla każdego przedziału $[c,d]$ zbiór $[c,d]\setminus D$ jest niepusty, bo ma miarę $\ge d-c$ – wynika to z podaddytywności miary. Wynika stąd od razu własność 9: w charakterze punktów t_j wystąpić mogą punkty, w których wartości f i g są takie same. Z tej własności wynika od razu własność czwarta. Pozostało udowodnić własność 5. Niech $D(f)$ oznacza zbiór punktów nieciągłości funkcji f, $D(g)$ – zbiór punktów nieciągłości funkcji g, A – zbiór tych punktów $x\in[a,b]$, dla których $f(x)<g(x)$. Zbiór A *nie* jest miary 0, zbiory $D(f)$ i $D(g)$ są miary 0. Zbiór $A\setminus(D(f)\cup D(g))$ *nie* jest miary 0, więc jest mocy kontinuum*. Niech $p\in A\setminus(D(f)\cup D(g))\cup\{a,b\}$. Mamy $f(p)<g(p)$.

* Gdyby zbiór $A\setminus(D(f)\cup D(g))$ był miary 0, to zbiór A byłby sumą 3 zbiorów miary 0, więc zbiór A byłby miary 0.

Niech α, β będą takimi liczbami, że $f(p) < \alpha < \beta < g(p)$. Ponieważ obie funkcje f i g są ciągłe, więc istnieje przedział $[c,d] \subset [a,b]$ taki, że dla $x \in [c,d]$ zachodzi nierówność $f(x) < \alpha < \beta < g(x)$. Mamy więc

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^d f(x)\,dx + \int_d^b f(x)\,dx \le \int_a^b f(x)\,dx + \alpha(d-c) + \int_d^b f(x)\,dx <$$
$$< \int_a^c g(x)\,dx + \beta(d-c) + \int_d^b g(x)\,dx \le \int_a^c g(x)\,dx + \int_c^d g(x)\,dx + \int_d^b g(x)\,dx = \int_a^b g(x)\,dx,$$

co kończy uzasadnienie własności 5 w oparciu o własności 3 i 6. ■

Twierdzenie o wartości średniej

Jeśli $f\colon [a,b \longrightarrow \mathbb{R}]$ jest ciągła, $g\colon [a,b \longrightarrow \mathbb{R}]$ – całkowalna w sensie Riemanna i nieujemna, to istnieje liczba $c \in [a,b]$ taka, że $\int_a^b f(x)g(x)\,dx = f(c)\int_a^b g(x)\,dx$.

Dowód.

Niech $m = \inf\{f(t)\colon \ \ t \in [a,b]\}$, $M = \sup\{f(t)\colon \ \ t \in [a,b]\}$. Dla każdego $x \in [a,b]$ zachodzi więc nierówność $mg(x) \le f(x)g(x) \le Mg(x)$. Wobec tego

$$m\int_a^b g(x)\,dx = \int_a^b mg(x)\,dx \le \int_a^b f(x)g(x)\,dx \le \int_a^b Mg(x)\,dx = M\int_a^b g(x)\,dx$$

Jeśli $\int_a^b g(x)\,dx = 0$, to przyjmujemy np. $c = \frac{1}{2}(a+b)$. Jeśli jednak $\int_a^b g(x)\,dx \ne 0$, czyli $\int_a^b g(x)\,dx > 0$, to otrzymujemy $m \le \frac{\int_a^b f(x)g(x)\,dx}{\int_a^b g(x)\,dx} \le M$, a ponieważ f ma własność Darboux, bo jest ciągła, więc istnieje $c \in [a,b]$ takie, że $f(c) = \frac{\int_a^b f(x)g(x)\,dx}{\int_a^b g(x)\,dx}$. Dowód został zakończony. ■

Łatwo można zauważyć, że w przypadku tej wersji twierdzenia o wartości średniej nie da się ominąć założenia ciągłości funkcji f – inaczej niż w wersji z poprzedniej części tego tekstu, gdzie własność Darboux i tak była (pochodna, jeśli istnieje w każdym punkcie przedziału, ma własność Darboux). Funkcja całkowalna w sensie Riemanna własności Darboux mieć nie musi, np. funkcja monotoniczna, która ma punkt nieciągłości.

Niech $f\colon [a,b] \longrightarrow \mathbb{R}$ oznacza funkcję całkowalną w sensie Riemanna. Niech $F(x) = \int_a^x f(t)\,dt$.

Twierdzenie o ciągłości i różniczkowalności całki

Funkcja F spełnia warunek Lipschitza na przedziale $[a,b]$. Jeśli funkcja f jest ciągła w punkcie $p \in [a,b]$, to F ma pochodną w punkcie p i zachodzi równość $F'(p) = f(p)$.

Dowód.

f jest ograniczona na $[a,b]$. Niech $M > 0$ będzie taką liczbą, że $|f(x)| \le M$ dla każdego $x \in [a,b]$. Niech $a \le x < y \le b$. Mamy

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

Musimy jeszcze wykazać, że funkcja F jest różniczkowalna w punktach ciągłości funkcji f. Załóżmy, że $p < b$, $0 < h < b - p$. Mamy $\left|\frac{1}{h}\big(F(p+h) - F(p)\big) - f(p)\right| = \left|\frac{1}{h}\left(\int_p^{p+h} f(t)\,dt\right) - f(p)\right| =$ $\left|\frac{1}{h}\int_p^{p+h}\Big(f(t) - f(p)\Big)\,dt\right| \leq$

$$\leq \frac{1}{h}\int_p^{p+h}\Big|f(t) - f(p)\Big|\,dt \leq \sup_{p\leq t\leq p+h}\Big|f(t) - f(p)\Big| \xrightarrow[h\to 0]{} 0\,.$$

Ostatnia równość (chodzi o granicę, znaku $=$ nie ma, ale jest strzałka, która go zastępuje) wynika z ciągłości funkcji f w punkcie p, to jedyny moment, w którym ciągłość jest wykorzystywana. Wykazaliśmy, że $f(p)$ jest prawostronną pochodną funkcji F w punkcie p. W taki sam sposób wykazać można, że jest to również pochodna lewostronna, gdy $a < p$. Dowód został zakończony. ■

Zauważmy, że z tego twierdzenia wynika, że funkcja ciągła na przedziale jest pochodną pewnej funkcji różniczkowalnej, tym razem to nie szkic dowodu, lecz dokładne rozumowanie. Przy okazji okazuje się, że w przypadku funkcji ciągłych oba podejścia do całkowania dają ten sam wynik: całka Riemanna równa jest różnicy wartości funkcji pierwotnej. Pozwala to znajdować liczne całki Riemanna. Podkreślić jednak wypada, że istnieją funkcje różniczkowalne, których pochodne są niecałkowalne w sensie Riemanna i oczywiście funkcje, które są całkowalne w sensie Riemanna i nie mają własności przyjmowania wartości pośrednich. Jeśli funkcja f jest całkowalna w sensie Riemanna i ma funkcję pierwotną F, to całka $\int_a^b f(x)\,dx$ równa jest różnicy wartości funkcji pierwotnej na końcach przedziału. Wynika to łatwo z twierdzenia o wartości średniej:

$$F(b) - F(a) = \sum_{j=1}^{n}\big(F(x_j) - F(x_{j-1})\big) = \sum_{j=1}^{n} F'(t_j)(x_j - x_{j-1}) = \sum_{j=1}^{n} f(t_j)(x_j - x_{j-1})\,.$$

Ostatnia suma jest bliska całce Riemanna funkcji f, bo założyliśmy, że ta funkcja jest całkowalna w sensie Riemanna. Dodajmy jeszcze, bez dowodu, że funkcja spełniająca warunek Lipschitza jest różniczkowalna w prawie każdym punkcie, tj. zbiór punktów nieróżniczkowalności funkcji lipschitzowskiej jest miary 0, ale to twierdzenie dosyć daleko wykracza poza program wykładu z analizy.

Twierdzenie o przybliżaniu funkcji całkowalnych funkcjami ciągłymi

Niech $f\colon [a,b] \longrightarrow \mathbb{R}$ będzie funkcją całkowalną w sensie Riemanna. Dla każdej liczby $\varepsilon > 0$ istnieje funkcja ciągła $g\colon [a,b] \longrightarrow \mathbb{R}$, taka że

$$\int_a^b |f(x) - g(x)|\,dx < \varepsilon \qquad \text{i} \qquad \inf_{t\in[a,b]} f(t) \leq \inf_{t\in[a,b]} g(t) \leq \sup_{t\in[a,b]} g(t) \leq \sup_{t\in[a,b]} f(t)\,. \tag{PFC}$$

Jeśli f jest monotoniczna, ale nie jest stała, to istnieje funkcja ściśle monotoniczna g, dla której spełniona jest powyższa nierówność (PFC).

Dowód.

Niech $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ będzie tak drobnym rozbiciem przedziału $[a,b]$, że $\left|\int_a^b f(x)\,dx - \sum_{j=1}^{n} f(t_j)(x_j - x_{j-1})\right| < \frac{\varepsilon}{2}$ dla dowolnego wyboru punktów $t_j \in [x_{j-1}, x_j]$ i niech

$M = \sup_{t\in[a,b]} |f(t)|$. Możemy założyć, że $M > 0$, dla funkcji zerowej twierdzenie jest oczywiste. Niech $\delta > 0$ będzie taką liczbą, że $4(n-1)M\delta < \varepsilon$ i $3\delta < x_j - x_{j-1}$ dla $j = 1, 2, \ldots n$. Niech P_j oznacza przedział domknięty o środku x_j, $j = 1, 2, \ldots, n-1$ i długości δ. Niech I_1, I_2, $\ldots, I_n$ będą kolejnymi przedziałami, z których składa się zbiór $[a,b] \setminus (P_1 \cup P_2 \cup \ldots \cup P_{n-1})$. Niech g będzie funkcją określoną na przedziale $[a,b]$, która

(1) jest ciągła,

(2) we wszystkich punktach przedziału I_j przyjmuje wartość m_j,

(3) jest postaci $ax+b$ na każdym przedziale P_1, P_2, $\ldots, P_{n-1}$.

Jest jasne, że na każdym z przedziałów I_1, I_2, $\ldots$, I_n zachodzi nierówność $f(x) \geq g(x)$, na przedziałach P_1, P_2, $\ldots$, P_{n-1} zachodzi nierówność $|f(x) - g(x)| < 2M$. Mamy więc

$$\frac{\varepsilon}{2} > \int_a^b f(x) - \sum_{j=1}^n m_j(x_j - x_{j-1}) = \sum_{j=1}^n \int_{x_{j-1}}^{x_j} \big(f(x) - m_j\big)\,dx \geq \sum_{j=1}^n \int_{I_j} \big(f(x) - m_j\big)\,dx =$$

$$= \sum_{j=1}^n \int_{I_j} \big(f(x) - g(x)\big)\,dx = \sum_{j=1}^n \int_{I_j} \big|f(x) - g(x)\big|\,dx$$

— jeśli $I = [\alpha, \beta]$, to $\int_I h(x)\,dx = \int_\alpha^\beta h(x)\,dx$.

Mamy też $\displaystyle\sum_{j=1}^{n-1} \int_{P_j} \big|f(x) - g(x)\big|\,dx \leq 2M(n-1)\delta < \frac{\varepsilon}{2}$. Z dwu ostatnich nierówności i z tego, że

$$\int_a^b \big|f(x) - g(x)\big|\,dx = \int_{I_1} \big|f(x) - g(x)\big|\,dx + \int_{P_1} \big|f(x) - g(x)\big|\,dx + \int_{I_2} \big|f(x) - g(x)\big|\,dx +$$

$$+ \int_{P_2} \big|f(x) - g(x)\big|\,dx + \cdots + \int_{P_{n-1}} \big|f(x) - g(x)\big|\,dx + \int_{I_n} \big|f(x) - g(x)\big|\,dx =$$

$$= \sum_{1}^{n-1} \int_{P_j} \big|f(x) - g(x)\big|\,dx + \sum_{j=1}^{n} \int_{I_j} \big|f(x) - g(x)\big|\,dx < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

wynika pierwsza część tezy. Z konstrukcji wynika natychmiast, że wszystkie wartości funkcji g znajdują się między kresami funkcji f. Jest również jasne, że jeśli f jest funkcją monotoniczną, to również g jest funkcją monotoniczną.

Wykażemy jeszcze, że jeśli funkcja f jest niemalejąca i nie jest stała, to można znaleźć funkcję ściśle rosnącą g spełniającą nierówności (PFC). W dalszej części rozumowania g oznacza funkcję, którą skonstruowaliśmy poprzednio. Ponieważ f nie jest stała, więc dla dostatecznie małych ε funkcja g również nie jest stała. Niech $c \in [a,b]$ będzie takim punktem, że $g(a) < g(c) < g(b)$. Niech $0 < k < 1$ i niech $g_k(x) = k^2\big((x-c)+g(c)\big)+(1-k)g(x)$. Mamy więc $\big|g_k(x)-g(x)\big| = k\big|k(x-c)+g(c)-g(x)\big| \leq k\big(\frac{b-a}{2} + M\big) \xrightarrow[k\to 0]{} 0$, $g_k(b) = k\big(k(b-c)+g(c)\big) + (1-k)g(b) = g(b) + k\Big(k(b-c) - \big(g(b) - g(c)\big)\Big)$. Z ostatniej równości i z tego, że $g(b) > g(c)$ wynika, że dla k dostatecznie bliskich 0 zachodzi nierówność $g_k(b) < g(b)$. Analogicznie $g_k(a) = k\big(k(a-c)+g(c)\big) + (1-k)g(a) = g(a) + k\Big(\big(g(c) - g(a)\big) - k(c-a)\Big)$. Wynika stąd, że dla k dostatecznie bliskich 0 zachodzi nierówność $g_k(a) > g(a)$. Wobec tego dla każdego $x \in [a,b]$ zachodzi nierówność $g(a) < g_k(a) \leq g_k(x) \leq g_k(b) < g(b)$.

Funkcja g_k jest rosnąca, bo jest sumą funkcji niemalejącej $(1-k)g$ i ściśle rosnącej $k^2\big((x-c)+g(c)\big)$. Ta obserwacja kończy dowód. $\blacksquare$

Twierdzenie o przybliżaniu funkcji całkowalnych funkcjami ciągłymi pozwoli nam w przyszłości na przeprowadzanie dowodów w przypadku funkcji ciągłych, a następnie wyciąganie ogólniejszych wniosków. Przykłady pojawią się wkrótce. Liczbę $\int_a^b |f(x)-g(x)|\,dx$ można traktować jako odległość między funkcjami całkowalnymi f i g. Z formalnego punktu widzenia nie jest to najwłaściwsze ze względu na to, że całka z różnicy funkcji przyjmujących te same wartości poza zbiorem miary 0 równa jest 0, więc należałoby najpierw wprowadzić relację równoważności: dwie funkcje są równoważne wtedy i tylko wtedy, gdy zbiór tych punktów, w których przyjmują one różne wartości ma miarę 0. Oznacza to, że z takiego punktu widzenia funkcje różniące się np. tylko w punktach wymiernych pewnego przedziału są tą samą funkcją. Po przyjęciu takiej umowy liczba $\int_a^b |f(x)-g(x)|\,dx$ może pełnić rolę odległości. Twierdzenie o przybliżaniu funkcji całkowalnych ciągłymi mówi, że zbiór funkcji ciągłych jest gęsty w zbiorze funkcji całkowalnych. Stwierdzenie to w pewnym sensie przypomina twierdzenie o gęstości liczb wymiernych w zbiorze wszystkich liczb rzeczywistych (w każdym przedziale znajduje się nieskończenie wiele liczb wymiernych!).

Lemacik o szacowaniu całki

Niech $f\colon [a,b] \longrightarrow \mathbb{R}$ oznacza funkcję całkowalną w sensie Riemanna. Niech $|f(x)| \le M$ dla każdego $x \in [a,b]$. Załóżmy, że $\big|\{x \in [a,b]\colon \quad |f(x)| > \varepsilon\}\big| < \delta$. Przy tych założeniach

$$\int_a^b |f(x)|\,dx < M\delta + \varepsilon(b-a)\,.$$

Dowód.

Załóżmy, że $a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$ jest na tyle drobnym podziałem przedziału $[a,b]$, że suma Riemanna z nim związana przybliża całkę $\int_a^b |f(x)|\,dx$ z błędem mniejszym niż $\eta > 0$. Niech $|f(t_j)| \le \varepsilon$, jeśli tylko w przedziale $[x_{j-1},x_j]$ znajduje się taki punkt. Jeśli musieliśmy wybrać t_j tak, że $|f(t_j)| > \varepsilon$, to mamy $|f(t)| > \varepsilon$ dla każdego $t \in x_{j-1},x_j]$. Wobec tego suma długości tych przedziałów jest $< \delta$ i wobec tego suma pozostałych jest jest $> b-a-\delta$. Stąd wynika, że $\int_a^b |f(x)|\,dx < M\delta + \varepsilon(b-a\delta) < M\delta + \varepsilon(b-a)$. Dowód został zakończony. $\blacksquare$

Twierdzenie Weierstrassa o przybliżaniu funkcji ciągłych wielomianami

Dla każdej liczby $\varepsilon > 0$ i dla każdej funkcji ciągłej $F\colon [a,b] \longrightarrow \mathbb{R}$ istnieje wielomian W taki, że $|F(x) - W(x) < \varepsilon$ dla każdego punktu $x \in [a,b]$.

Dowód. (Bernstein).

Istnieje wiele dowodów tego twierdzenia. Wybieramy ten, bo ma on swą oczywistą genezę w twierdzeniu z rachunku prawdopodobieństwa i jeśli ktoś do niego wtedy wróci, np. dlatego, że będzie on tam powtórzony przy okazji prawa wielkich liczb, to będzie mu łatwiej pojąć, o co w tym wszystkim chodzi.

Wystarczy udowodnić twierdzenie w przypadku $[a,b]=[0,1]$. By się z tym pogodzić wystarczy przyjąć, że $t=\frac{x-a}{b-a}$, $f(t)=F\big(a+t(b-a)\big)=F(x)$ i analogicznie $w(t)=W\big(a+t(b-a)\big)=W(x)$. Jasne jest, że wtedy $|f(t)-w(t)|=|F(x)-W(x)|$, przy czym $a\le x\le b$ wtedy i tylko wtedy, gdy $0\le t\le 1$.

Niech $b_n(t)=\sum\limits_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}t^k(1-t)^{n-k}$. Wielomian b_n nazywany jest n–tym wielomianem Bernsteina funkcji f. Wykażemy, że jeśli liczba n jest dostatecznie duża, to przyjęcie $w(t)=b_n(t)$ powoduje, że dla każdego $t\in[0,1]$ zachodzi nierówność $|f(t)-w(t)|=|f(t)-b_n(t)|<\varepsilon$.

Zaczniemy od pomocniczych równości.

$$\sum_{k=0}^{n}\binom{n}{k}t^k(1-t)^{n-k}=1 \tag{W1}$$

$$\sum_{k=0}^{n}k\binom{n}{k}t^k(1-t)^{n-k}=nt \tag{W2}$$

$$\sum_{k=0}^{n}\binom{n}{k}t^kk^2(1-t)^{n-k}=n(n-1)t^2+nt \tag{W3}$$

$$\forall_{\delta>0}\sum_{\left|\frac{k}{n}-t\right|\ge\delta}\binom{n}{k}t^k(1-t)^{n-k}\le\frac{1}{n}\cdot\frac{t(1-t)}{\delta^2}\le\frac{1}{4n\delta^2} \tag{W4}$$

Równość (W1) wynika natychmiast z tego, że $1=\big(t+(1-t)\big)^n=\binom{n}{k}t^k(1-t)^{n-k}$.

Równość (W2) podobnie: $\sum\limits_{k=0}^{n}k\binom{n}{k}t^k(1-t)^{n-k}=nt\sum\limits_{k=1}^{n}\binom{n-1}{k-1}t^{k-1}(1-t)^{n-1-(k-1)}=$

$$=nt\sum_{k=0}^{n-1}\binom{n-1}{k}t^k(1-t)^{n-1-k}=nt\big(t+(1-t)\big)^{n-1}=nt.$$

Kolej na (W3). $\sum\limits_{k=0}^{n}k^2\binom{n}{k}t^k(1-t)^{n-k}=\sum\limits_{k=2}^{n}k(k-1)\binom{n}{k}t^k(1-t)^{n-k}+\sum\limits_{k=1}^{n}k\binom{n}{k}t^k(1-t)^{n-k}=$

$$=n(n-1)t^2\sum_{k=2}^{n}\binom{n-2}{k-2}t^{k-2}(1-t)^{n-2-(k-2)}+nt=n(n-1)t^2\sum_{k=0}^{n-2}\binom{n-2}{k}t^k(1-t)^{n-2-k}+nt=$$

$$=n(n-1)t^2\big(t+(1-t)\big)^{n-2}+nt=n(n-1)t^2+nt.$$

Teraz kolej na najważniejszą z tych czterech równości, zwana nierównością Czebyszewa (w przypadku ogólniejszym, na omówienie którego tu nie ma miejsca).

$$n^2\delta^2\sum_{\left|\frac{k}{n}-t\right|\ge\delta}\binom{n}{k}t^k(1-t)^{n-k}\le\sum_{\left|\frac{k}{n}-t\right|\ge\delta}\big(k-nt\big)^2\binom{n}{k}t^k(1-t)^{n-k}<\sum_{k=0}^{n}\big(k-nt\big)^2\binom{n}{k}t^k(1-t)^{n-k}=$$

$$=\sum_{k=0}^{n}k^2\binom{n}{k}t^k(1-t)^{n-k}-2\sum_{k=0}^{n}knt\binom{n}{k}t^k(1-t)^{n-k}+\sum_{k=0}^{n}\big(nt\big)^2\binom{n}{k}t^k(1-t)^{n-k}=$$

$$=n(n-1)t^2+nt-2n^2t^2+n^2t^2=nt-nt^2=nt(1-t)$$

Z otrzymanej nierówności łatwo wynika, że

$$\sum_{\left|\frac{k}{n}-t\right|\geq\delta}\binom{n}{k}t^k(1-t)^{n-k}\leq\frac{nt(1-t)}{n^2\delta^2}=\frac{1}{n}\cdot\frac{t(1-t)}{\delta^2}\leq\frac{1}{4n\delta^2}\,.$$

Jesteśmy gotowi do dowodu. Ponieważ f jest ciągła na przedziale *domkniętym*, więc istnieje liczba $\delta>0$, taka że jeśli $|t-s|<\delta$, to $|f(t)-f(s)|<\frac{\varepsilon}{2}$. Dzięki (W1) mamy teraz

$$\big|f(t)-b_n(t)\big|=\Big|f(t)-\sum_{k=0}^{n}f\big(\tfrac{k}{n}\big)\tbinom{n}{k}t^k(1-t)^{n-k}\Big|=\Big|\sum_{k=0}^{n}\big(f(t)-f\big(\tfrac{k}{n}\big)\big)\tbinom{n}{k}t^k(1-t)^{n-k}\Big|\leq$$

$$\leq\sum_{\left|\frac{k}{n}-t\right|<\delta}\big|f(t)-f\big(\tfrac{k}{n}\big)\big|\tbinom{n}{k}t^k(1-t)^{n-k}+\sum_{\left|\frac{k}{n}-t\right|\geq\delta}\big|f(t)-f\big(\tfrac{k}{n}\big)\big|\tbinom{n}{k}t^k(1-t)^{n-k}\leq$$

$$\leq\frac{\varepsilon}{2}\sum_{\left|\frac{k}{n}-t\right|<\delta}\tbinom{n}{k}t^k(1-t)^{n-k}+2M\sum_{\left|\frac{k}{n}-t\right|\geq\delta}\tbinom{n}{k}t^k(1-t)^{n-k}\leq\frac{\varepsilon}{2}+2M\frac{1}{4n\delta^2}=\frac{\varepsilon}{2}+\frac{M}{2n\delta^2}$$

Jeśli $n>\dfrac{M}{\varepsilon\delta^2}$, to otrzymujemy $\big|f(t)-b_n(t)\big|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$. Dowód został zakończony. ■

Krótki komentarz probabilistyczny.

Załóżmy, że w Nibylandii (pozdrowienia od Piotrusia Pana) wyprodukowano monetę niecałkiem symetryczną: rzucając nią otrzymujemy reszkę z prawdopodobieństwem t a drugą stronę z mało czytelną podobizną jakiegoś fruwającego stworzenia – z prawdopodobieństwem $1-t$. Prawdopodobieństwo uzyskania w n rzutach tą monetą dokładnie k–reszek równe jest więc $\binom{n}{k}t^k(1-t)^{n-k}$. Liczba $\sum\limits_{k=0}^{n}k\binom{n}{k}t^k(1-t)^{n-k}\big(=nt\big)$ oznacza więc średnia liczbę reszek otrzymanych w k rzutach tą monetą. Oczekujemy więc, że rzucając tą monetą n razy otrzymamy nt, a raczej około nt, reszek. Wzór (W4) wyjaśnia, jakie jest prawdopodobieństwo tego, że liczba rzutów (k), w których wypadła reszka będzie różnić się od oczekiwanej (nt) o pewien ustalony procent liczby rzutów lub bardziej, dlatego zajmujemy się tam różnicą $\left|\frac{k}{n}-t\right|$ (nierówność $\left|\frac{k}{n}-t\right|\geq\delta$ równoważna jest temu, że $|k-nt|\geq\delta n$, ta ustalona część n to δn), prawdopodobieństwo to dąży do 0 – jest to tzw. słabe prawo wielkich liczb. Liczba $b_n(t)$ jest więc średnią liczb $f\big(\frac{k}{n}\big)$, ta średnia jest mniej więcej równa $f(t)$, bo na ogół $\frac{k}{n}\approx t$. Powinna więc mieć miejsce równość przybliżona $f(t)\approx b_n(t)$. W końcówce nie jesteśmy całkiem precyzyjni, ale wcześniej staraliśmy się wyjaśnić precyzyjnie, o co nam chodzi. ■