Analiza matematyczna ISIM II

Ryszard Szwarc*

Spis treści

---

*Wykład prowadzony w semestrze letnim 2014 na podstawie notatek Magdaleny Świczewskiej z 2005-2006, opracowany na podstawie notatek Mateusza Wasylkiewicza

1 Całki niewłaściwe

Przykłady.

(a) $f(x) = \dfrac{1}{x}$, $0 < x \leqslant 1$. Dla $0 < a < 1$ mamy

$$\int\limits_a^1 \frac{1}{x}\, dx = \log x \Big|_a^1 = -\log a \underset{a \to 0^+}{\longrightarrow} \infty.$$

To oznacza, że pole obszaru pod wykresem funkcji $y = 1/x$, $0 < x \leqslant 1$, jest nieskończone.

(b) $f(x) = \dfrac{1}{\sqrt{x}}$, $0 < x \leqslant 1$. Wtedy dla $0 < a < 1$ mamy

$$\int\limits_a^1 \frac{1}{\sqrt{x}}\, dx = 2\sqrt{x} \Big|_a^1 = 2 - 2\sqrt{a} \underset{a \to 0^+}{\longrightarrow} 2.$$

Pole pod wykresem $y = \dfrac{1}{\sqrt{x}}$, $0 < x \leqslant 1$, jest skończone i równe 2 pomimo tego, że obszar pod wykresem jest nieograniczony.

(c) $f(x) = \dfrac{2}{x^3}$, $x \geqslant 1$. Dla $b > 1$ mamy

$$\int\limits_1^b \frac{2}{x^3}\, dx = -\frac{1}{x^2}\Big|_1^b = 1 - \frac{1}{b^2} \underset{b \to \infty}{\longrightarrow} 1.$$

Definicja 1.1. *Mówimy, że całka* $\int\limits_a^b f(x)\,dx$ *jest niewłaściwa z osobliwością w punkcie* b *jeśli*

1. *Funkcja* $f(x)$ *jest określona w przedziale* $[a, b)$.
2. $f(x)$ *jest całkowalna na każdym przedziale* $[a, b']$ *dla* $a < b' < b$ *(np.* $f(x)$ *jest ciągła na* $[a, b)$*).*
3. $b = \infty$ *albo* $b < \infty$ *i* $f(x)$ *jest nieograniczona w pobliżu* b.

Podobnie określa się całkę niewłaściwą $\int\limits_a^b f(x)\,dx$ z osobliwością w dolnej granicy całkowania a.

Przykłady.

Całka	Punkt osobliwy
$\int\limits_1^\infty \frac{1}{x}\,dx$	∞
$\int\limits_0^1 \frac{1}{\sqrt{1-x}}\,dx$	1
$\int\limits_0^1 \frac{1}{\log x}\,dx$	1
$\int\limits_0^\pi \frac{\sin x}{x}\,dx$	nie ma osobliwości
$\int\limits_0^\pi \frac{\sin x}{x^2}\,dx$	0
$\int\limits_\pi^\infty \frac{\operatorname{sgn}(\sin x)}{x}\,dx$	∞

Definicja 1.2. *Załóżmy, że dla całki* $\int\limits_a^b f(x)\,dx$ *z osobliwością w punkcie* b *istnieje granica* $\lim\limits_{b'\to b^-} \int\limits_a^{b'} f(x)\,dx$. *Mówimy wtedy, że całka* $\int\limits_a^b f(x)\,dx$ *jest zbież-*

na i piszemy

$$\int_a^b f(x)\,dx = \lim_{b'\to b^-} \int_a^{b'} f(x)\,dx.$$

Podobnie określamy zbieżność całki z osobliwością w punkcie a. W przeciwnym wypadku, gdy granica nie istnieje, mówimy, że całka jest rozbieżna.

Przykład.

$$\int_0^1 \log x\,dx = \lim_{a\to 0^+} \int_a^1 \log x\,dx = \lim_{a\to 0^+} (x\log x - x)\Big|_a^1$$
$$= \lim_{a\to 0^+} (-1 - a\log a + a) = -1,$$

bo $\lim\limits_{a\to 0^+} a\log a = 0$.

Twierdzenie 1.3 (warunek Cauchy'ego zbieżności całki)**.** *Załóżmy, że całka* $\int_a^b f(x)\,dx$ *ma osobliwość w punkcie* b*. Całka ta jest zbieżna wtedy i tylko wtedy, gdy dla dowolnej dodatniej liczby* ε *istnieje liczba* $a < b_0 < b$ *taka, że dla dowolnych* b' *i* b'' *z warunku* $b_0 < b' < b'' < b$ *wynika*

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

Dowód. Zbieżność całki oznacza z definicji istnienie granicy $\lim\limits_{b'\to b^-} F(b')$, gdzie $F(b') = \int_a^{b'} f(x)\,dx$. Z kolei granica ta istnieje wtedy i tylko wtedy, gdy spełniony jest warunek Cauchy'ego, czyli w zapisie kwantyfikatorowym

$$\forall\,\varepsilon > 0\ \exists\, b_0 < b\ \forall\, b',\, b''\ \ [b_0 < b' < b'' < b] \implies |F(b'') - F(b')| < \varepsilon.$$

Ale

$$F(b'') - F(b') = \int_a^{b''} f(x)\,dx - \int_a^{b'} f(x)\,dx = \int_{b'}^{b''} f(x)\,dx.$$

$\square$

Przykład. Sprawdzamy zbieżność całki $\int\limits_0^\infty \frac{\sin x}{x}\,dx$. Dla $0 < b' < b''$ mamy[1]

$$\left| \int\limits_{b'}^{b''} \frac{\sin x}{x}\,dx \right| \leqslant \frac{2}{b'}.$$

Dla $\varepsilon > 0$ przyjmijmy $b_0 = \frac{2}{\varepsilon}$. Wtedy dla $b'' > b' > \frac{2}{\varepsilon}$ otrzymujemy

$$\left| \int\limits_{b'}^{b''} \frac{\sin x}{x}\,dx \right| < \varepsilon.$$

Można udowodnić, że (por. 1.5.2)

$$\int\limits_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}.$$

Przypuśćmy, że całka $\int\limits_a^b f(x)\,dx$ ma osobliwość w b. Dla $a < c < b$ całki $\int\limits_c^b f(x)\,dx$ i $\int\limits_a^b f(x)\,dx$ są jednocześnie zbieżne lub jednocześnie rozbieżne, , bo warunki Cauchy'ego są dla nich identyczne Ponadto w przypadku zbieżności mamy

$$\int\limits_a^b f(x)\,dx = \int\limits_a^c f(x)\,dx + \int\limits_c^b f(x)\,dx.$$

Ostatni wzór otrzymujemy przez przejście graniczne $b' \to b^-$ w równości

$$\int\limits_a^{b'} f(x)\,dx = \int\limits_a^c f(x)\,dx + \int\limits_c^{b'} f(x)\,dx.$$

[1]Z drugiego twierdzenia o wartości średniej mamy $\int\limits_{b'}^{b''} \frac{\sin x}{x}\,dx = \frac{1}{b'}\int\limits_{b'}^{\xi} \sin x\,dx = \frac{\cos b' - \cos\xi}{b'}$, dla pewnego ξ, $b' < \xi < b''$.

Dla całek z osobliwością w punkcie b spełniony jest wzór

$$\int_a^b [Af(x) + Bg(x)]\,dx = A\int_a^b f(x)\,dx + B\int_a^b g(x)\,dx,$$

o ile całki po prawej stronie są zbieżne.

Definicja 1.4. *Mówimy, że całka* $\int_a^b f(x)\,dx$ *z osobliwością w* b *jest bezwzględnie zbieżna, jeśli zbieżna jest całka* $\int_a^b |f(x)|\,dx$.

Twierdzenie 1.5. *Całka bezwzględnie zbieżna jest zbieżna.*

Dowód. Dla $b' < b'' < b$ mamy

$$\left|\int_{b'}^{b''} f(x)\,dx\right| \leqslant \int_{b'}^{b''} |f(x)|\,dx.$$

Zatem z warunku Cauchy'ego dla całki $\int_a^b |f(x)|\,dx$ wynika ten warunek dla całki $\int_a^b f(x)\,dx$. □

Przykład. $\int_\pi^\infty \frac{\sin x}{x^2}\,dx$. Sprawdzamy warunek Cauchy'ego dla całki $\int_\pi^\infty \frac{|\sin x|}{x^2}\,dx$.

$$\int_{b'}^{b''} \frac{|\sin x|}{x^2}\,dx \leqslant \int_{b'}^{b''} \frac{1}{x^2}\,dx = \frac{1}{b'} - \frac{1}{b''} < \frac{1}{b'}.$$

Twierdzenie 1.6 (kryterium porównawcze)**.** *Niech* $0 \leqslant f(x) \leqslant g(x)$ *dla* $a \leqslant x < b$.

(i) *Ze zbieżności całki* $\int_a^b g(x)\,dx$ *wynika zbieżność* $\int_a^b f(x)\,dx$. *Ponadto*

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

(ii) Z rozbieżności całki $\int\limits_a^b f(x)\,dx$ *wynika rozbieżność* $\int\limits_a^b g(x)\,dx$.

Dowód. (i) Dla $a < b' < b'' < b$ mamy

$$0 \leqslant \int\limits_{b'}^{b''} f(x)\,dx \leqslant \int\limits_{b'}^{b''} g(x)\,dx.$$

Stąd otrzymujemy zbieżność całki z funkcji $f(x)$. Przechodzimy do granicy $b' \to b^-$ w nierówności

$$\int\limits_a^{b'} f(x)\,dx \leqslant \int\limits_a^{b'} g(x)\,dx$$

aby otrzymać drugą część tezy. $\square$

Uwaga 1.7. Jeśli całka $\int\limits_a^b f(x)\,dx$ z osobliwością w b jest bezwzględnie zbieżna, to

$$\left|\int\limits_a^b f(x)\,dx\right| \leqslant \int\limits_a^b |f(x)|\,dx.$$

Rzeczywiście, mamy $-|f(x)| \leqslant f(x) \leqslant |f(x)|$. Po scałkowaniu otrzymujemy

$$-\int\limits_a^{b'} |f(x)|\,dx \leqslant \int\limits_a^{b'} f(x)\,dx \leqslant \int\limits_a^{b'} |f(x)|\,dx.$$

Przechodzimy do granicy $b' \to b^-$ i otrzymujemy

$$-\int\limits_a^{b} |f(x)|\,dx \leqslant \int\limits_a^{b} f(x)\,dx \leqslant \int\limits_a^{b} |f(x)|\,dx.$$

Przykład. Czy całka $\int\limits_0^\infty \frac{\sin x}{x}\,dx$ jest bezwzględnie zbieżna ?

Dla $k \geqslant 1$ mamy

$$\int\limits_{(k-1)\pi}^{k\pi} \frac{|\sin x|}{x}\,dx \geqslant \frac{1}{k\pi}\int\limits_{(k-1)\pi}^{k\pi} |\sin x|\,dx = \frac{1}{k\pi}\int\limits_0^\pi \sin x\,dx = \frac{2}{k\pi}.$$

Zatem

$$\int\limits_0^{n\pi} \frac{|\sin x|}{x}\,dx = \sum_{k=1}^{n} \int\limits_{(k-1)\pi}^{k\pi} \frac{|\sin x|}{x}\,dx \;\geqslant\; \frac{2}{\pi}\sum_{k=1}^{n}\frac{1}{k} \;\xrightarrow[n]{}\; \infty.$$

Twierdzenie 1.8. *Jeśli funkcja $F(x)$ jest ciągła w przedziale $[a,b]$ i różniczkowalna w sposób ciągły w $[a,b)$ oraz $F'(x) = f(x)$ dla $a \leqslant x < b$, to*

$$\int\limits_a^b f(x)\,dx = F(b) - F(a).$$

Dowód.

$$\int\limits_a^{b'} f(x)\,dx = F(b') - F(a) \xrightarrow[b' \to b^-]{} F(b) - F(a).$$

$\square$

Twierdzenie 1.9. *Przy założeniach poprzedniego twierdzenia z $b = \infty$ i dodatkowym założeniu, że $L = \lim\limits_{x\to\infty} F(x)$ mamy*

$$\int\limits_a^b f(x)\,dx = L - F(a).$$

Przykłady.

(a) $\displaystyle\int\limits_0^1 \log x\,dx = (x\log x - x)\Big|_{0^+}^{1} = -1$. Rolę funkcji $F(x)$ spełnia

$$F(x) = \begin{cases} x\log x - x & 0 < x \leqslant 1, \\ 0 & x = 0. \end{cases}$$

(b) $\displaystyle\int\limits_1^\infty \frac{dx}{x\sqrt{x}} = -\frac{2}{\sqrt{x}}\Big|_1^\infty = 2.$

1.1 Całki niewłaściwe z funkcji nieujemnych

Przypuśćmy, że całka $\int\limits_a^b f(x)\,dx$ ma osobliwość w punkcie b oraz $f(x) \geqslant 0$ dla $a \leqslant x < b$. Wtedy funkcja $F(b') = \int\limits_a^{b'} f(x)\,dx$ jest rosnąca. Zatem całka $\int\limits_a^b f(x)\,dx$ jest zbieżna albo rozbieżna do ∞.

Przykłady.

(a) $\int\limits_0^1 \frac{dx}{\sqrt{x+x^4}}$. Mamy

$$0 < \frac{1}{\sqrt{x+x^4}} \leqslant \frac{1}{\sqrt{x}}, \qquad \int\limits_0^1 \frac{dx}{\sqrt{x}} = 2\sqrt{x}\Big|_0^1 = 2.$$

Zatem rozważana całka jest zbieżna.

(b) $\int\limits_1^\infty \frac{dx}{x+\sqrt{x}}$. Dla $x \geqslant 1$ mamy

$$\frac{1}{x+\sqrt{x}} \geqslant \frac{1}{2x}, \qquad \int\limits_1^\infty \frac{dx}{2x} = \frac{1}{2}\log x\Big|_1^\infty = \infty.$$

Zatem $\int\limits_1^\infty \frac{dx}{x+\sqrt{x}} = \infty$.

Uwaga 1.10. W kryterium porównawczym wystarczy, aby $0 \leqslant f(x) \leqslant g(x)$ dla $c \leqslant x < b$ dla pewnego punktu c, $a < c < b$. Rzeczywiście, całki $\int\limits_c^b f(x)\,dx$ oraz $\int\limits_a^b f(x)\,dx$ są jednocześnie zbieżne lub jednocześnie rozbieżne.

Twierdzenie 1.11 (kryterium graniczne)**.** *Załóżmy, że funkcje ciągłe $f(x)$ i $g(x)$ są określone i dodatnie na przedziale $[a, b)$ oraz*

$$\lim_{x\to b^-} \frac{f(x)}{g(x)} = A > 0.$$

Wtedy całki $\int\limits_a^b f(x)\,dx$ oraz $\int\limits_a^b g(x)\,dx$ są jednocześnie zbieżne lub jednocześnie rozbieżne.

Dowód. Z założenia można znaleźć punkt $a \leqslant c < b$ taki, że dla $c \leqslant x < b$ mamy

$$\frac{1}{2}A < \frac{f(x)}{g(x)} < \frac{3}{2}A.$$

Wtedy

$$\frac{1}{2}Ag(x) < f(x) < \frac{3}{2}Ag(x), \qquad c \leqslant x < b.$$

Z kryterium porównawczego i z Uwagi 1.10 otrzymujemy tezę twierdzenia. □

Przykłady.

(a) $\int\limits_0^1 \dfrac{dx}{x - \log(1+x)}$. Mamy

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

$$\int\limits_0^1 \frac{dx}{x^2} = -\frac{1}{x}\bigg|_{0^+}^{1} = \infty.$$

(b) $\displaystyle\int\limits_0^1 \frac{dx}{\log(1+\sqrt{x})}$.

$$\lim_{x\to 0^+} \frac{\frac{1}{\log(1+\sqrt{x})}}{\frac{1}{\sqrt{x}}} = \lim_{x\to 0^+} \frac{\sqrt{x}}{\log(1+\sqrt{x})} \underset{y=1+\sqrt{x}}{=} \lim_{y\to 1^+} \frac{y-1}{\log y}$$
$$= \frac{1}{\lim\limits_{y\to 1^+} \frac{\log y}{y-1}} = \frac{1}{(\log y)'\Big|_{y=1}} = 1.$$

$$\int\limits_0^1 \frac{dx}{\sqrt{x}} = 2\sqrt{x}\Big|_0^1 = 2.$$

Uwaga 1.12. Jeśli w założeniach twierdzenia $A = 0$, to ze zbieżności $\int\limits_a^b g(x)\,dx$ wynika zbieżność $\int\limits_a^b f(x)\,dx$. Jeśli $A = \infty$, to z rozbieżności $\int\limits_a^b g(x)\,dx$ wynika rozbieżność $\int\limits_a^b f(x)\,dx$. Rzeczywiście, jeśli $A = 0$, to

$$\frac{f(x)}{g(x)} \leqslant 1, \text{ dla } a < c \leqslant x < b.$$

Wtedy
$$0 \leqslant f(x) \leqslant g(x), \text{ dla } a < c \leqslant x < b.$$

Z kryterium porównawczego otrzymujemy tezę. Jeśli $A = \infty$, to

$$\lim_{x\to b^-} \frac{g(x)}{f(x)} = 0.$$

Korzystając z pierwszej części uzyskujemy żądaną konkluzję.

Przykłady.

(a) Dla $\alpha, \beta > 0$ rozważamy całkę $\int\limits_1^\infty x^\alpha e^{-x^\beta}\,dx$. Przyjmijmy $f(x) = x^\alpha e^{-x^\beta}$ oraz $g(x) = x^{-2}$. Wtedy

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{x^{\alpha+2}}{e^{x^\beta}} \underset{y=x^\beta}{=} \lim_{y\to\infty} \frac{y^\gamma}{e^y},$$

dla $\gamma = \frac{\alpha+2}{\beta}$. Niech $n = [\gamma] + 1$. Wtedy

$$0 \leqslant \frac{y^\gamma}{e^y} \leqslant \frac{y^n}{e^y} \leqslant \frac{y^n}{\frac{y^{n+1}}{(n+1)!}} = \frac{(n+1)!}{y}.$$

Zatem $\lim\limits_{x\to\infty} \dfrac{f(x)}{g(x)} = 0$. Całka funkcji $g(x) = x^{-2}$ jest zbieżna na półprostej $[1, \infty)$, zatem zbieżna jest też całka $\int\limits_1^\infty x^\alpha e^{-x^\beta}\,dx$.

(b) Obracamy wykres funkcji $y = x^{-1}$ dla $x \geqslant 1$ wokół osi OX. Otrzymujemy tzw. róg archanioła Gabriela (opisany przez włoskiego matematyka i fizyka Evangelistę Torricellego 1608-1647).

Obliczamy objętość obszaru ograniczonego przez róg przyjmując, że x jest liczone w metrach.

$$V = \pi \int\limits_1^\infty \frac{dx}{x^2} = -\frac{\pi}{x}\Big|_1^\infty = \pi\ (\mathrm{m}^3).$$

Obliczamy pole powierzchni rogu[2].

$$S = 2\pi \int\limits_1^\infty \frac{1}{x}\sqrt{1 + \frac{1}{x^4}}\,dx \geqslant 2\pi \int\limits_1^\infty \frac{dx}{x} = 2\pi \log x\Big|_1^\infty = \infty.$$

Zagadka. Wyobraźmy sobie, że róg wykonany jest z wsiąkliwej białej bibuły. Nalewamy do rogu π metrów sześciennych czarnego atramentu. Następnie wylewamy atrament. Wewnętrzna strona rogu zostanie zabarwiona na czarno. Czyli za pomocą skończonej ilości atramentu zabarwiliśmy nieskończoną powierzchnię. Jak wyjaśnić ten paradoks ?

1.2 Całki i szeregi

Rozważamy całkę niewłaściwą $\int\limits_a^b f(x)\,dx$ z osobliwością w $b \leqslant \infty$. Niech $a = b_0 < b_1 < b_2 < \ldots < b_n \ldots$, oraz $b_n \underset{n}{\longrightarrow} b$.

---

[2] $S = 2\pi \int\limits_a^b f(x)\sqrt{1 + f'(x)^2}\,dx$

Twierdzenie 1.13.

(i) Jeśli całka $\int\limits_a^b f(x)\,dx$ *jest zbieżna, to zbieżny jest szereg całek właściwych* $\sum\limits_{n=1}^{\infty} \int\limits_{b_{n-1}}^{b_n} f(x)\,dx$.

(ii) Jeśli $f(x) \geqslant 0$, *to implikacja odwrotna jest również prawdziwa.*

Dowód. (i) Obliczamy sumę częściową szeregu.

$$S_n = \sum_{k=1}^{n} \int\limits_{b_{k-1}}^{b_k} f(x)\,dx = \int\limits_a^{b_n} f(x)\,dx \underset{n}{\longrightarrow} \int\limits_a^b f(x)\,dx.$$

(ii) Niech $f(x) \geqslant 0$. Dla zbieżności całki $\int\limits_a^b f(x)\,dx$ wystarczy pokazać, że całki $\int\limits_a^{b'} f(x)\,dx$ są wspólnie ograniczone, niezależnie od $b' < b$. Niech $b' < b$. Ponieważ $b_n \to b$, to $b_{n_0} > b'$ dla pewnego wskaźnika n_0. Wtedy

$$\int\limits_a^{b'} f(x)\,dx \leqslant \int\limits_a^{b_{n_0}} f(x)\,dx = \sum_{k=1}^{n_0} \int\limits_{b_{k-1}}^{b_k} f(x)\,dx \leqslant \sum_{k=1}^{\infty} \int\limits_{b_{k-1}}^{b_k} f(x)\,dx.$$

□

Twierdzenie 1.14. *Załóżmy, że* $f(x)$ *jest dodatnią funkcją malejącą na przedziale* $[1,\infty)$. *Wtedy zbieżność całki* $\int\limits_1^{\infty} f(x)\,dx$ *jest równoważna zbieżności szeregu* $\sum\limits_{n=1}^{\infty} f(n)$. *Ponadto dla* $I_n = \int\limits_1^n f(x)\,dx$ *oraz* $S_n = \sum\limits_{k=1}^{n-1} f(k)$ *ciąg liczb* $S_n - I_n$ *jest zbieżny.*

Dowód. Z nierówności

$$f(k) \leqslant \int\limits_{k-1}^{k} f(x)\,dx \leqslant f(k-1) \tag{1.1}$$

wnioskujemy, że zbieżność szeregu $\sum\limits_{k=1}^{\infty} f(k)$ jest równoważna ze zbieżnością szeregu $\sum\limits_{k=1}^{\infty}\int\limits_{k-1}^{k} f(x)\,dx$. Z kolei zbieżność szeregu całek jest równoważna ze zbieżnością całki $\int\limits_{1}^{\infty} f(x)\,dx$.

Zsumujmy (1.1) dla $k = 2, 3, \ldots, n$. Wtedy

$$\underbrace{f(2) + f(3) + \ldots + f(n)}_{S_n - f(1) + f(n)} \leqslant \underbrace{\int\limits_{1}^{n} f(x)\,dx}_{I_n} \leqslant \underbrace{f(1) + f(2) + \ldots + f(n-1)}_{S_n}.$$

Otrzymujemy więc $0 \leqslant S_n - I_n \leqslant f(1) - f(n) \leqslant f(1)$. Ciąg $S_n - I_n$ jest rosnący. Rzeczywiście, mamy $f(n) > \int\limits_{n}^{n+1} f(x)\,dx$. To oznacza, że $S_{n+1} - S_n > I_{n+1} - I_n$, czyli $S_{n+1} - I_{n+1} > S_n - I_n$. Ciąg $S_n - I_n$ jest zatem zbieżny. $\square$

Przykłady.

(a) $f(x) = \dfrac{1}{x^\alpha}$, $\alpha > 0$. Mamy

$$\int\limits_{1}^{\infty} \frac{dx}{x^\alpha} = \begin{cases} \log x & \alpha = 1, \\ \frac{1}{1-\alpha} x^{1-\alpha} & \alpha \neq 1 \end{cases}\Bigg|_{1}^{\infty} = \begin{cases} \infty & 0 < \alpha \leqslant 1, \\ \frac{1}{\alpha-1} & \alpha > 1. \end{cases}$$

Zatem szereg $\sum\limits_{n=1}^{\infty} \dfrac{1}{n^\alpha}$ jest zbieżny tylko dla $\alpha > 1$.

(b) $f(x) = \dfrac{1}{x \log^\alpha x}$, $\alpha > 0$, $x \geqslant 2$. Mamy

$$\int\limits_{2}^{\infty} \frac{dx}{x \log^\alpha x} = \begin{cases} \log\log x & \alpha = 1, \\ \frac{1}{1-\alpha} (\log x)^{1-\alpha} & \alpha \neq 1 \end{cases}\Bigg|_{2}^{\infty} = \begin{cases} \infty & 0 < \alpha \leqslant 1, \\ \frac{1}{\alpha-1} (\log 2)^{1-\alpha} & \alpha > 1. \end{cases}$$

Wnioskujemy, że szereg

$$\sum_{n=2}^{\infty} \frac{1}{n \log^\alpha n}$$

jest zbieżny tylko dla $\alpha > 1$.

(c) $f(x) = \dfrac{1}{x}$. Mamy

$$S_n - I_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1} - \int\limits_1^n \frac{1}{x}\,dx$$
$$= 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1} - \log n \xrightarrow[n]{} c,$$

gdzie c jest stałą Eulera ($c = 0,57721\ldots$).

Twierdzenie 1.15. *Jeśli funkcja $g(x)$ jest nieujemna i malejąca w przedziale $[a,b)$ oraz $\lim\limits_{x\to b^-} g(x) = 0$, natomiast dla $a \leqslant b' < b$ całki (właściwe) $\int\limits_a^{b'} f(x)\,dx$ są wspólnie ograniczone, to całka $\int\limits_a^{b} f(x)g(x)\,dx$ jest zbieżna. W szczególności teza jest spełniona, gdy całka $\int\limits_a^{b} f(x)\,dx$ jest zbieżna.*

Dowód. Z założenia $\left|\int\limits_a^{b'} f(x)\,dx\right| \leqslant M$ dla pewnej stałej M i wszystkich $a \leqslant b' < b$. Sprawdzamy warunek Cauchy'ego zbieżności całki $\int\limits_a^{b} f(x)g(x)\,dx$. Dla $a \leqslant b' < b'' < b$, na podstawie twierdzenia o wartości średniej, mamy

$$\int\limits_{b'}^{b''} f(x)g(x)\,dx = g(b')\int\limits_{b'}^{\xi} f(x)\,dx$$

dla pewnego punktu pośredniego ξ, $b' < \xi < b''$. Zatem

$$\left|\int\limits_{b'}^{b''} f(x)g(x)\,dx\right| = g(b')\left|\int\limits_a^{\xi} f(x)\,dx - \int\limits_a^{b'} f(x)\,dx\right|$$
$$\leqslant g(b')\left\{\left|\int\limits_a^{\xi} f(x)\,dx\right| + \left|\int\limits_a^{b'} f(x)\,dx\right|\right\} \leqslant 2Mg(b').$$

Jeśli całka $\int\limits_a^b f(x)\,dx$ jest zbieżna, tzn. istnieje granica

$$\lim_{b'\to b^-} \int\limits_a^{b'} f(x)\,dx$$

to całki $\int\limits_a^{b'} f(x)\,dx$ są wspólnie ograniczone. $\square$

Przykłady.

(a) Badamy zbieżność całki Dirichleta $\int\limits_0^\infty \frac{\sin x}{x}\,dx$. Wystarczy zbadać zbieżność całki $\int\limits_{\pi/2}^\infty \frac{\sin x}{x}\,dx$. Przyjmijmy $g(x) = \frac{1}{x}$ oraz $f(x) = \sin x$. Wtedy

$$\left| \int\limits_{\pi/2}^{b} \sin x\,dx \right| = |\cos(\pi/2) - \cos b| \leqslant 1.$$

Zatem całka $\int\limits_0^\infty \frac{\sin x}{x}\,dx$ jest zbieżna. Można udowodnić, że wartość całki wynosi $\pi/2$.

(b) $\int\limits_0^\infty \sin(x^2)\,dx$ nosi nazwę całki Fresnela. Zbadamy zbieżność całki $\int\limits_{\sqrt{\pi/2}}^\infty \sin(x^2)\,dx$.

Przyjmujemy $f(x) = 2x\sin(x^2)$ oraz $g(x) = \frac{1}{2x}$. Wtedy

$$\left| \int\limits_{\sqrt{\pi/2}}^{b} 2x\sin(x^2)\,dx \right| = \left| -\cos(x^2)\Big|_{\sqrt{\pi/2}}^{b} \right| = |\cos(b^2)| \leqslant 1.$$

1.3 Całki niewłaściwe z osobliwością w kilku punktach

Definicja 1.16. *Mówimy, że całka* $\int\limits_a^b f(x)\,dx$ *ma osobliwość w punktach* a *i* b, *jeśli całki* $\int\limits_a^c f(x)\,dx$ *i* $\int\limits_c^b f(x)\,dx$ *mają osobliwości w punktach* a *i* b, *odpowiednio, dla* $a < c < b$. *Mówimy, że całka* $\int\limits_a^b f(x)\,dx$ *jest zbieżna, jeśli zbieżne są całki* $\int\limits_a^c f(x)\,dx$ *i* $\int\limits_c^b f(x)\,dx$. *Określamy wtedy*

$$\int\limits_a^b f(x)\,dx = \int\limits_a^c f(x)\,dx + \int\limits_c^b f(x)\,dx.$$

Wartość całki po lewej stronie nie zależy od wyboru punktu c.

Przykład.

$$\int\limits_0^\infty \frac{dx}{x^5 + x^2 + \sqrt{x}}.$$

Badamy całki

$$\int\limits_0^1 \frac{dx}{x^5 + x^2 + \sqrt{x}}, \qquad \int\limits_1^\infty \frac{dx}{x^5 + x^2 + \sqrt{x}}.$$

Mamy

$$0 < \frac{1}{x^5 + x^2 + \sqrt{x}} < \frac{1}{\sqrt{x}}, \qquad 0 < \frac{1}{x^5 + x^2 + \sqrt{x}} < \frac{1}{x^5}$$

oraz

$$\int\limits_0^1 \frac{dx}{\sqrt{x}} = 2, \qquad \int\limits_1^\infty \frac{dx}{x^5} = \frac{1}{4}.$$

Definicja 1.17. *Mówimy, że całka* $\int\limits_a^b f(x)\,dx$ *ma osobliwość w punktach* a, b *i* c, $a < c < b$ *jeśli całki* $\int\limits_a^c f(x)\,dx$ *i* $\int\limits_c^b f(x)\,dx$ *mają osobliwości w punktach*

a i c oraz w c i b, odpowiednio. Jeśli obie całki są zbieżne, to określamy

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

Przykład. $\displaystyle\int_{-\infty}^{\infty} \frac{dx}{\sqrt[3]{x^2}(x^2+1)}$. Mamy trzy punkty osobliwe $-\infty$, 0 oraz ∞. Funkcja podcałkowa jest parzysta, więc wystarczy zbadać całkę $\displaystyle\int_0^{\infty} \frac{dx}{\sqrt[3]{x^2}(x^2+1)}$. Mamy

$$0 < x \leqslant 1, \qquad \frac{1}{\sqrt[3]{x^2}(x^2+1)} \leqslant \frac{1}{\sqrt[3]{x^2}}, \qquad \int_0^1 \frac{dx}{\sqrt[3]{x^2}} = 3;$$

$$x \geqslant 1, \qquad \frac{1}{\sqrt[3]{x^2}(x^2+1)} \leqslant \frac{1}{x^2}, \qquad \int_1^{\infty} \frac{dx}{x^2} = 1.$$

Uwaga 1.18. Mamy $\displaystyle\int_{-a}^{a} \sin x\,dx = 0$. Ale granica całek $\displaystyle\int_a^b \sin x\,dx$, gdy $a \to -\infty$, $b \to \infty$ nie istnieje, bo

$$\int_a^b \sin x\,dx = \cos a - \cos b.$$

Określa się słabszą zbieżność całki $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx$ w sensie wartości głównej. Mówimy, że pv $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx$ jest zbieżna, jeśli istnieje granica $\displaystyle\lim_{a\to\infty} \int_{-a}^{a} f(x)\,dx$. Dla porównania, zbieżność całki w zwykłym sensie oznacza istnienie granicy $\displaystyle\lim_{\substack{a\to\infty \\ b\to\infty}} \int_{-a}^{b} f(x)\,dx$.

Rozważmy całkę $\displaystyle\int_{-1}^{1} f(x)\,dx$ z osobliwością w punkcie 0. Mówimy, że całka

$\mathrm{pv}\int\limits_{-1}^{1} f(x)\,dx$ jest zbieżna, jeśli istnieje granica

$$\lim_{\varepsilon\to 0^+}\left[\int\limits_{-1}^{-\varepsilon} f(x)\,dx + \int\limits_{\varepsilon}^{1} f(x)\,dx\right].$$

Zwykła zbieżność tej całki oznaczałaby istnienie granicy

$$\lim_{\substack{\varepsilon\to 0^+\\ \eta\to 0^+}}\left[\int\limits_{-1}^{-\eta} f(x)\,dx + \int\limits_{\varepsilon}^{1} f(x)\,dx\right].$$

Przykład. $\mathrm{pv}\int\limits_{-1}^{1}\frac{dx}{x} = 0$, bo $\int\limits_{-1}^{-\varepsilon}\frac{dx}{x} + \int\limits_{\varepsilon}^{1}\frac{dx}{x} = 0$. Całka nie jest zbieżna w zwykłym sensie, bo całki $\int\limits_{0}^{1}\frac{dx}{x}$ i $\int\limits_{-1}^{0}\frac{dx}{x}$ nie są zbieżne.

1.4 Całki z parametrem

Niech $f(x,y)$ będzie funkcją dwu zmiennych określoną na pewnym podzbiorze $U\subset\mathbb{R}^2$. Np. $U=\mathbb{R}^2$ lub $U=[a,b]\times[c,d]$.

Przykłady.

$$f(x,y) = xy,$$
$$g(x,y) = \sqrt{x^2+y}, \quad [0,1]\times[1,2].$$

Gdy zbiór U, na którym określona jest funkcja $f(x,y)$, nie jest podany, to przyjmujemy za U największy możliwy zakres (x,y), dla których zapis $f(x,y)$ ma sens. W przeciwnym wypadku, gdy zbiór U jest podany, to $f(x,y)$ jest określona tylko na U. W drugim przykładzie $g(1,1)=\sqrt{2}$, ale wartość funkcji w punkcie $(-1,1)$ nie jest określona.

Funkcję $f(x,y)$ określoną na $[a,b]\times[c,d]$ możemy traktować jako rodzinę funkcji jednej zmiennej

$$[a,b]\ni x\longmapsto f(x,y)$$

z parametrem y z przedziału $[c,d]$. Podobnie $f(x,y)$ możemy traktować jako rodzinę funkcji

$$[c,d]\ni y\longmapsto f(x,y)$$

z parametrem x z przedziału $[a,b]$.

Przykłady. Dla

$$f(x,y) = xy, \qquad g(x,y) = \sqrt{x^2+y},\ [0,1]\times[1,2]$$

mamy

$$\begin{array}{llll} f(x,0)=0 & f(x,1)=x & g(x,1)=\sqrt{x^2+1} & g(x,2)=\sqrt{x^2+2} \\ f(0,y)=0 & f(2.y)=2y & g(0,y)=\sqrt{y} & g(\frac{1}{2},y)=\sqrt{y+\frac{1}{4}}. \end{array}$$

Definicja 1.19. *Mówimy, że funkcja $f(x,y)$ określona na $U \subset \mathbb{R}^2$ jest ciągła w punkcie $(x_0,y_0) \in U$, jeśli dla dowolnych ciągów $x_n \to x_0$ i $y_n \to y_0$ takich, że $(x_n,y_n) \in U$, mamy $f(x_n,y_n) \to f(x_0,y_0)$.*

Intuicyjnie oznacza to, że dla punktów (x,y) z U położonych blisko punktu (x_0,y_0) wartości $f(x,y)$ leżą blisko $f(x_0,y_0)$. Zapis kwantyfikatorowy tego sformułowania ma postać:

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x,y\ \{(x,y) \in U,\ |x-x_0| < \delta,\ |y-y_0| < \delta \implies \\ |f(x,y) - f(x_0,y_0)| < \varepsilon\}.$$

Można udowodnić, że podany wyżej warunek jest równoważny z warunkiem z definicji.

Definicja 1.20. *Mówimy, że funkcja $f(x,y)$ jest ciągła na zbiorze $U \subset \mathbb{R}^2$, jeśli ta funkcja jest ciągła w każdym punkcie zbioru U.*

Przykład. Funkcja $f(x,y) = xy$ jest ciągła na $\mathbb{R}^2$.

Uwaga 1.21. Jeśli funkcja $f(x,y)$ jest ciągła na $[a,b] \times [c,d]$, to każda z funkcji

$$\begin{aligned} &[a,b] \ni x \longmapsto f(x,y), \qquad y \in [c,d], \\ &[c,d] \ni y \longmapsto f(x,y), \qquad x \in [a,b] \end{aligned}$$

jest ciągła.

Te własności oznaczają, że funkcja $f(x,y)$ jest ciągła ze względu na każdą zmienną z osobna i są słabsze niż ciągłość funkcji $f(x,y)$ określona w definicji. Porównajmy ciągłość funkcji $f(x,y)$ w punkcie (x_0,y_0) z połączonymi dwoma warunkami: ciągłością funkcji $x \mapsto f(x,y_0)$ w punkcie x_0 i ciągłością funkcji

$y \mapsto f(x_0, y)$ w punkcie y_0. Dla uproszczenia przyjmijmy, że punkt (x_0, y_0) leży wewnątrz prostokąta $[a, b] \times [c, d]$, czyli w $(a, b) \times (c, d)$.

Warunek ciągłości w (x_0, y_0) oznacza, że dla małych wartości $\delta > 0$, jeśli punkt (x, y) leży w kwadracie

$$|x - x_0| < \delta, \ |y - y_0| < \delta, \tag{1.2}$$

to wartość $f(x, y)$ leży blisko wartości $f(x_0, y_0)$.

Z kolei ciągłości funkcji $x \mapsto f(x, y_0)$ oraz $y \mapsto f(x_0, y)$ w punktach x_0 i y_0 (odpowiednio) oznacza, że jeśli (x, y) leży na jednym z odcinków,

$$\begin{aligned} &|x - x_0| < \delta, \ y = y_0, \quad \text{poziomy} \\ &x = x_0, \ |y - y_0| < \delta, \quad \text{pionowy} \end{aligned}$$

to wartość $f(x, y)$ leży blisko wartości $f(x_0, y_0)$. Te odcinki leżą w kwadracie opisanym przez (1.2); są jego osiami symetrii.

Zadanie. Znaleźć funkcję $f(x, y)$ określoną na $[-1, 1] \times [-1, 1]$ taką, że każda z funkcji $x \mapsto f(x, y)$ i $y \mapsto f(x, y)$ jest ciągła na przedziale $[-1, 1]$, ale funkcja $f(x, y)$ nie jest ciągła w punkcie $(0, 0)$.

Twierdzenie 1.22. *Funkcja $f(x, y)$ ciągła na prostokącie $U = [a, b] \times [c, d]$ jest jednostajnie ciągła, tzn. dla dowolnej liczby $\varepsilon > 0$ istnieje liczba $\delta > 0$ taka, że z warunków (x, y), $(x', y') \in U$ oraz $|x - x'| < \delta$, $|y - y'| < \delta$ wynika $|f(x, y) - f(x', y')| < \varepsilon$.*

Uwaga 1.23. Warunek w tezie oznacza, że jeśli punkty (x, y) i (x', y') leżą blisko siebie, to wartości funkcji f w tych punktach są bliskie siebie.

Dowód. (*nie wprost*). Załóżmy, że dla pewnej liczby $\varepsilon > 0$ oraz $\delta_n = \dfrac{1}{n}$ istnieją punkty (x_n, y_n), $(x'_n, y'_n) \in U$, dla których

$$|x_n - x'_n| < \frac{1}{n}, \ |y_n - y'_n| < \frac{1}{n}, \quad |f(x_n, y_n) - f(x'_n, y'_n)| \geqslant \varepsilon. \tag{1.3}$$

Z twierdzenia Bolzano-Weierstrassa można znaleźć rosnący ciąg liczb naturalnych n_k taki, że ciągi x_{n_k} oraz y_{n_k} są zbieżne. Niech $x_{n_k} \underset{k}{\to} x_0$ oraz $y_{n_k} \underset{k}{\to} y_0$. Wtedy $a \leqslant x_0 \leqslant b$ oraz $c \leqslant y_0 \leqslant d$, tzn. $(x_0, y_0) \in U$. Ponadto $x'_{n_k} \underset{k}{\to} x_0$ oraz

$y'_{n_k} \underset{k}{\to} y_0$, bo $x_{n_k} - x'_{n_k} \underset{k}{\to} 0$ i $y_{n_k} - y'_{n_k} \underset{k}{\to} 0$. Z założenia ciągłości funkcji f (w punkcie (x_0, y_0)) mamy

$$f(x_{n_k}, y_{n_k}) \underset{k}{\longrightarrow} f(x_0, y_0), \quad f(x'_{n_k}, y'_{n_k}) \underset{k}{\longrightarrow} f(x_0, y_0).$$

Zatem

$$f(x_{n_k}, y_{n_k}) - f(x'_{n_k}, y'_{n_k}) \underset{k}{\longrightarrow} 0.$$

Ale to przeczy warunkowi (1.3). □

Twierdzenie 1.24. *Każda funkcja ciągła na prostokącie $[a, b] \times [c, d]$ jest ograniczona.*

Dowód. (*nie wprost*). Załóżmy, że funkcja f nie jest ograniczona. Tzn. dla dowolnej liczby n można znaleźć punkt $(x_n, y_n) \in U$ spełniający $|f(x_n, y_n)| \geqslant n$, czyli $|f(x_n, y_n)| \underset{n}{\to} \infty$. Podobnie jak w dowodzie poprzedniego twierdzenia, można znaleźć ciąg wskaźników n_k taki, że $x_{n_k} \underset{k}{\to} x_0$ i $y_{n_k} \underset{k}{\to} y_0$. Wtedy z ciągłości funkcji f w punkcie (x_0, y_0) otrzymamy

$$f(x_{n_k}, y_{n_k}) \underset{k}{\longrightarrow} f(x_0, y_0)$$

co przeczy warunkowi

$$n_k \leqslant |f(x_{n_k}, y_{n_k})| \underset{k}{\longrightarrow} \infty.$$

□

Lemat 1.25. *Załóżmy, że funkcja $f(x, y)$ jest ciągła na prostokącie $[a, b] \times [c, d]$. Wtedy funkcje*

$$g(x) = \int_c^d f(x, y)\, dy \quad \text{i} \quad h(y) = \int_a^b f(x, y)\, dx$$

są ciągłe na przedziałach $[a, b]$ i $[c, d]$, odpowiednio.

Przykład. Dla funkcji $f(x, y) = x^2 y^{1/2} + x^3 y^3$ określonej na prostokącie $[0, 1] \times [0, 1]$ mamy

$$\begin{aligned} g(x) = \int_0^1 f(x, y)\, dy = \int_0^1 (x^2 y^{1/2} + x^3 y^3)\, dy \\ = \left(x^2\, \frac{2}{3} y^{3/2} + x^3\, \frac{1}{4} y^4 \right) \Bigg|_{y=0}^{y=1} = \frac{2}{3} x^2 + \frac{1}{4} x^3, \end{aligned}$$

$$h(y) = \int\limits_0^1 f(x,y)\,dx = \int\limits_0^1 (x^2 y^{1/2} + x^3 y^3)\,dx$$
$$= \left(\frac{x^3}{3} y^{1/2} + \frac{1}{4} x^4 y^3\right)\Big|_{x=0}^{x=1} = \frac{1}{3} y^{1/2} + \frac{1}{4} y^3.$$

Uwaga 1.26. Przy całkowaniu względem dy zmienną x traktujemy jako parametr (stałą). Podobnie, przy całkowaniu względem dx zmienna y jest parametrem.

Dowód. (tylko dla funkcji $g(x)$). Mamy

$$|g(x) - g(x')| = \left|\int\limits_c^d f(x,y)\,dy - \int\limits_c^d f(x',y)\,dy\right| = \left|\int\limits_c^d [f(x,y) - f(x',y)]\,dy\right|$$
$$\leqslant \int\limits_c^d |f(x,y) - f(x',y)|\,dy.$$

Z jednostajnej ciągłości funkcji f wiemy, że dla liczby $\varepsilon > 0$ istnieje liczba $\delta > 0$ taka, że jeśli $|x - x'| < \delta$ oraz $|y - y'| < \delta$, to $|f(x,y) - f(x',y')| < \dfrac{\varepsilon}{d-c}$. Załóżmy, że $|x - x'| < \delta$ oraz $y = y'$. Wtedy

$$|f(x,y) - f(x',y)| < \frac{\varepsilon}{d-c}.$$

Zatem

$$|g(x) - g(x')| < \frac{\varepsilon}{d-c} \cdot (d-c) = \varepsilon.$$

□

Przykład. Niech

$$f(x,y) = x \sin y, \qquad [0,1] \times [0,2\pi].$$

Przy ustalonej wartości y funkcja $x \mapsto f(x,y)$ jest różniczkowalna (jako funkcja liniowa). Mówimy wtedy, że istnieje pochodna cząstkowa funkcji $f(x,y)$ względem zmiennej x (zmienna y traktowana jest jako parametr).

Definicja 1.27. *Pochodna cząstkowa funkcji $f(x,y)$ względem zmiennej x w punkcie (x_0,y_0) oznaczana jest symbolem $\dfrac{\partial f}{\partial x}(x_0,y_0)$ i określona wzorem*

$$\frac{\partial f}{\partial x}(x_0,y_0) = \frac{d}{dx} f(x,y_0)\Big|_{x=x_0} = \lim_{h\to 0} \frac{f(x_0+h,y_0) - f(x_0,y_0)}{h}.$$

Podobnie określamy pochodną cząstkową funkcji $f(x,y)$ względem y w punkcie (x_0,y_0), czyli

$$\frac{\partial f}{\partial y}(x_0,y_0) = \frac{d}{dy} f(x_0,y)\Big|_{y=y_0} = \lim_{h\to 0} \frac{f(x_0,y_0+h)-f(x_0,y_0)}{h}.$$

Przykład. Dla $f(x,y)=\sqrt{x^2+y^4}$ mamy

$$\frac{\partial f}{\partial x}(x,y) = \frac{1}{2\sqrt{x^2+y^4}} \cdot 2x = \frac{x}{\sqrt{x^2+y^4}}.$$

$$\frac{\partial f}{\partial y}(x,y) = \frac{1}{2\sqrt{x^2+y^4}} \cdot 4y^3 = \frac{2y^3}{\sqrt{x^2+y^4}}.$$

Otrzymujemy

$$\frac{\partial f}{\partial x}(1,2) = \frac{1}{\sqrt{17}}, \qquad \frac{\partial f}{\partial y}(1,2) = \frac{16}{\sqrt{17}}.$$

Uwaga 1.28. Przy obliczaniu pochodnych cząstkowych w punkcie (x_0,y_0) wystarczy znać dwie funkcje $x \mapsto f(x,y_0)$ oraz $y \mapsto f(x_0,y)$ i zróżniczkować je w punktach x_0, y_0 (odpowiednio). Alternatywnie można obliczyć pochodne cząstkowe funkcji f w dowolnym punkcie i na końcu wykonać podstawienie $x=x_0$, $y=y_0$.

Przykłady. Obliczyć $\dfrac{\partial f}{\partial x}(\pi/2,0)$ dla funkcji

$$f(x,y) = \sin(y\cos x)\sqrt{4+e^{\operatorname{arctg}(x^2+1)+y^2\log\sin x}}.$$

Podstawiamy $y=0$ i otrzymujemy $f(x,0)=0$. Stąd

$$\frac{\partial f}{\partial x}(\pi/2,0) = \frac{d}{dx} f(x,0)\Big|_{x=\pi/2} = 0.$$

Twierdzenie 1.29. *Załóżmy, że funkcja $f(x,y)$ jest ciągła na prostokącie $[a,b]\times[c,d]$ oraz, że pochodna cząstkowa $\dfrac{\partial f}{\partial x}(x,y)$ jest ciągła na $[a,b]\times[c,d]$. Wtedy funkcja*

$$g(x) = \int_c^d f(x,y)\,dy$$

jest różniczkowalna na przedziale $[a,b]$ oraz

$$g'(x) = \frac{d}{dx}\left(\int_c^d f(x,y)\,dy\right) = \int_c^d \frac{\partial f}{\partial x}(x,y)\,dy.$$

Uwaga 1.30. Teza twierdzenia mówi, że można wejść z pochodną względem zmiennej x pod znak całki. Wewnątrz całki trzeba użyć pochodnej cząstkowej, bo funkcja pod całką zależy od dwu zmiennych.

Uwaga 1.31. Zwykle w przykładach występujące funkcje są określone i ciągłe oraz różniczkowalne na większym obszarze niż podany prostokąt. Jednak, gdy rygorystycznie ograniczamy się do obszaru $[a, b] \times [c, d]$, to pochodne funkcji g na końcach przedziału $x = a$, $x = b$ i pochodne cząstkowe funkcji $f(x, y)$ w punktach postaci $(x, y) = (a, y)$ i $(x, y) = (b, y)$ traktujemy jako pochodne jednostronne: prawostronne, gdy $x = a$ i lewostronne, gdy $x = b$.

Dowód. Obliczamy iloraz różnicowy

$$\begin{aligned}\frac{g(x+h) - g(x)}{h} &= \frac{1}{h}\left[\int_c^d f(x+h, y)\,dy - \int_c^d f(x, y)\,dy\right] \\ &= \frac{1}{h}\int_c^d [f(x+h, y) - f(x, y)]\,dy.\end{aligned}$$

Z twierdzenia Lagrange'a zastosowanego do każdej z funkcji $x \mapsto f(x, y)$ otrzymujemy

$$f(x+h, y) - f(x, y) = \frac{\partial f}{\partial x}(x + \theta h, y)\ h,$$

przy czym współczynnik θ leży w przedziale $(0, 1)$ i jego wartość zależy od x, y i h. Wtedy

$$\begin{aligned}\frac{g(x+h) - g(x)}{h} &= \int_c^d \frac{\partial f}{\partial x}(x + \theta h, y)\,dy \\ &= \int_c^d \frac{\partial f}{\partial x}(x, y)\,dy + \int_c^d \left[\frac{\partial f}{\partial x}(x + \theta h, y) - \frac{\partial f}{\partial x}(x, y)\right] dy. \qquad (1.4)\end{aligned}$$

Funkcja $\dfrac{\partial f}{\partial x}(x, y)$ jest ciągła, zatem jest też jednostajnie ciągła. Stąd wynika, że granica przy $h \to 0$ drugiego składnika wynosi 0 i uzyskujemy tezę twierdzenia.

Formalnie: dla ustalonej liczby $\varepsilon > 0$ można znaleźć liczbę $\delta > 0$ taką, że jeśli $|x - x'| < \delta$, $|y - y'| < \delta$, to

$$\left|\frac{\partial f}{\partial x}(x', y') - \frac{\partial f}{\partial x}(x, y)\right| < \frac{\varepsilon}{d - c}.$$

Załóżmy, że $|h| < \delta$ wtedy dla $x' = x + \theta h$ oraz $y' = y$ wnioskujemy, że drugi składnik w (1.4) jest mniejszy niż ε. Stąd dostajemy tezę. $\square$

Twierdzenie 1.32. *Dla funkcji $f(x, y)$ ciągłej na prostokącie $[a, b] \times [c, d]$ mamy*

$$\int_a^b \left(\int_c^d f(x, y)\, dy \right) dx = \int_c^d \left(\int_a^b f(x, y)\, dx \right) dy.$$

Uwaga 1.33. Pojedyncze całki są funkcjami ciągłymi zmiennych x i y odpowiednio, zatem można wykonać operację całkowania względem pozostałej zmiennej.

Dowód. Określmy funkcję $F(x)$ na przedziale $[a, b]$ wzorem

$$F(x) = \int_c^d \left(\int_a^x f(t, y)\, dt \right) dy = \int_c^d h(x, y)\, dy,$$

gdzie

$$h(x, y) = \int_a^x f(t, y)\, dt.^3$$

Funkcja $h(x, y)$ jest ciągła na $[a, b] \times [c, d]$ (zadanie).

Ponadto funkcja $x \mapsto h(x, y)$ jest funkcją pierwotną do funkcji $x \mapsto f(x, y)$ dla każdego ustalonego argumentu y z przedziału $[c, d]$. Tzn.

$$\frac{\partial h}{\partial x}(x, y) = f(x, y).$$

Z poprzedniego twierdzenia zastosowanego do funkcji $h(x, y)$ otrzymujemy

$$F'(x) = \frac{d}{dx} \int_c^d h(x, y)\, dy = \int_c^d \frac{\partial h}{\partial x}(x, y)\, dy = \int_c^d f(x, y)\, dy.$$

Z zasadniczego twierdzenia rachunku różniczkowego i całkowego otrzymujemy

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

[3] Por. ze wzorem $h(x) = \int_a^x f(t)\, dt$

czyli

$$\int_c^d \left(\int_a^b f(t,y)\,dt \right) dy - 0 = \int_a^b \left(\int_c^d f(x,y)\,dy \right) dx.$$

□

Uwaga 1.34. Dla nieujemnej funkcji $f(x,y)$ obie całki występujące w tezie można interpretować jako objętość obszaru pod wykresem funkcji $z = f(x,y)$, gdy $a \leqslant x \leqslant b$ i $c \leqslant y \leqslant d$. Obszar dzielimy na n plasterków równej szerokości płaszczyznami pionowymi $x = x_i$. Następnie każdy plasterek dzielimy na m słupków równej szerokości płaszczyznami pionowymi $y = y_j$. Każdy słupek ma objętość równą w przybliżeniu

$$f(x_i, y_j)\Delta x_i \Delta y_j.$$

Sumujemy objętości słupków i otrzymujemy

$$\begin{aligned}\sum_{i=1}^n \left[\sum_{j=1}^m f(x_i,y_j)\Delta y_j \right] \Delta x_i &\sim \sum_{i=1}^n \left(\int_c^d f(x_i,y)\,dy \right) \Delta x_i \\ &= \sum_{i=1}^n g(x_i)\Delta x_i \sim \int_a^b g(x)\,dx = \int_a^b \left(\int_c^d f(x,y)\,dy \right) dx.\end{aligned}$$

Możemy sumę objętości słupków zapisać inaczej i otrzymać

$$\begin{aligned}\sum_{j=1}^m \left[\sum_{i=1}^n f(x_i,y_j)\Delta x_i \right] \Delta y_j &\sim \sum_{j=1}^m \left(\int_a^b f(x,y_j)\,dx \right) \Delta y_j \\ &= \sum_{j=1}^m h(y_j)\Delta y_j \sim \int_c^d h(y)\,dy = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy.\end{aligned}$$

Przykłady.

1. Obliczyć $\displaystyle\int_0^1 \frac{x^b - x^a}{\log x}\,dx, \quad 0 < a < b.$

   Funkcja podcałkowa jest określona w $x = 0$ i w $x = 1$ poprzez swoje granice

   $$\lim_{x\to 0^+} \frac{x^b - x^a}{\log x} = 0, \qquad \lim_{x\to 1^-} \frac{x^b - x^a}{\log x} = \lim_{x\to 1^-} \frac{bx^{b-1} - ax^{a-1}}{x^{-1}} = b - a.$$

Z warunku $\dfrac{\partial}{\partial y}(x^y) = x^y \log x$ [4] otrzymujemy

$$\int_0^1 \frac{x^b - x^a}{\log x}\,dx = \int_0^1 \left(\int_a^b x^y\,dy \right) dx = \int_a^b \left(\int_0^1 x^y\,dx \right) dy$$

$$\int_a^b \frac{1}{y+1} x^{y+1}\Big|_{x=0}^{x=1} dy = \int_a^b \frac{1}{y+1}\,dy = \log(y+1)\Big|_{y=a}^{y=b} = \log\frac{b+1}{a+1}.$$

2. Obliczyć $\dfrac{d}{dx}\displaystyle\int_0^1 \operatorname{arctg}\frac{x}{y}\,dy, \quad x > 0$. Funkcja podcałkowa przy ustalonej wartości $x > 0$ ma granicę, gdy $y \to 0^+$, równą $\pi/2$. Jej pochodna cząstkowa względem x ma granicę równą 0, gdy $y \to 0^+$, co widać z obliczeń poniżej. Zatem

$$\frac{d}{dx}\int_0^1 \operatorname{arctg}\frac{x}{y}\,dy = \int_0^1 \frac{\partial}{\partial x}\left(\operatorname{arctg}\frac{x}{y}\right) dy = \int_0^1 \frac{\frac{1}{y}}{1+\frac{x^2}{y^2}}\,dy$$

$$= \int_0^1 \frac{y}{x^2+y^2}\,dy = \frac{1}{2}\log(x^2+y^2)\Big|_{y=0}^{y=1} = \frac{1}{2}\log(x^2+1) - \log x.$$

3. Obliczyć $\dfrac{d}{dx}\displaystyle\int_0^1 \log(x^2+y^2)\,dy$.

$$\frac{d}{dx}\int_0^1 \log(x^2+y^2)\,dy = \int_0^1 \frac{\partial}{\partial x}\log(x^2+y^2)\,dy$$

$$= \int_0^1 \frac{2x}{x^2+y^2}\,dy = 2\operatorname{arctg}\frac{y}{x}\Big|_{y=0}^{y=1} = 2\operatorname{arctg}\frac{1}{x}.$$

---

[4] $\frac{\partial}{\partial y}(x^y) = \frac{\partial}{\partial y}(e^{y\log x}) = e^{y\log x}\log x = x^y \log x.$

1.5 Dwie ważne całki niewłaściwe

1.5.1 Całka Eulera

Całka niewłaściwa

$$\int_0^\infty e^{-x^2}\,dx$$

jest zbieżna, co wynika np. z nierówności

$$0 \leqslant e^{-x^2} \leqslant \frac{1}{x^2+1}$$

i z

$$\int_0^\infty \frac{1}{x^2+1}\,dx = \operatorname{arctg} x\Big|_{x=0}^{\infty} = \frac{\pi}{2}.$$

Można też użyć oszacowania

$$\int_0^\infty e^{-x^2}\,dx \leqslant \sum_{n=0}^{\infty} e^{-n^2}.$$

Wartość całki ma istotne znaczenie w rachunku prawdopodobieństwa. Zaczniemy od nieformalnego wyznaczenia wartości, które będzie podstawą ścisłych obliczeń.

Oznaczmy $I = \int\limits_0^\infty e^{-x^2}\,dx$. Wtedy

$$\begin{aligned} I^2 &= \int_0^\infty e^{-x^2}\left(\int_0^\infty e^{-y^2}\,dy\right)dx \underset{y:=xy}{=} \int_0^\infty e^{-x^2}\left(\int_0^\infty xe^{-(xy)^2}\,dy\right)dx \\ &= \int_0^\infty\left(\int_0^\infty xe^{-x^2(y^2+1)}\,dx\right)dy = \int_0^\infty \frac{-1}{2(y^2+1)}e^{-x^2(y^2+1)}\Big|_{x=0}^{\infty}\,dy \\ &= \int_0^\infty \frac{1}{2(y^2+1)}\,dy = \frac{\pi}{4}. \end{aligned}$$

Zatem

$$\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}. \tag{1.5}$$

Przy obliczeniach zamiana kolejności całkowania wymaga uzasadnienia.

Przejdziemy do ścisłego wyprowadzenia wzoru (1.5). Oznaczmy

$$A_n = \int\limits_{1/n^2}^{n^2} e^{-x^2}\,dx \cdot \int\limits_{1/n^3}^{n^3} e^{-y^2}\,dy.$$

Mamy $A_n \underset{n}{\to} I^2$. Dalej

$$\begin{aligned}
A_n &= \int\limits_{1/n^2}^{n^2} e^{-x^2}\left[\int\limits_{1/n^3}^{n^3} e^{-y^2}\,dy\right]dx \underset{y:=xy}{=} \int\limits_{1/n^2}^{n^2} e^{-x^2}\left[\int\limits_{1/(n^3x)}^{n^3/x} xe^{-(xy)^2}\,dy\right]dx \\
&\geqslant \int\limits_{1/n^2}^{n^2} e^{-x^2}\left[\int\limits_{1/n}^{n} xe^{-(xy)^2}\,dy\right]dx = \int\limits_{1/n}^{n}\left[\int\limits_{1/n^2}^{n^2} xe^{-x^2(y^2+1)}\,dx\right]dy \\
&= \int\limits_{1/n}^{n} \frac{-1}{2(y^2+1)} e^{-x^2(y^2+1)}\Big|_{x=1/n^2}^{x=n^2}\,dy = \int\limits_{1/n}^{n} \frac{1}{2(y^2+1)}\left[e^{-(y^2+1)/n^4} - e^{-(y^2+1)n^4}\right]dy \\
&\geqslant \int\limits_{1/n}^{n} \frac{1}{2(y^2+1)}\left[e^{-(n^2+1)/n^4} - e^{-(1/n^2+1)n^4}\right]dy \\
&= \frac{1}{2}\left[e^{-\frac{1}{n^2}-\frac{1}{n^4}} - e^{-n^2-n^4}\right]\left[\operatorname{arctg} n - \operatorname{arctg}\frac{1}{n}\right] \underset{n}{\longrightarrow} \frac{\pi}{4}.
\end{aligned}$$

Z drugiej strony

$$\begin{aligned}
A_n &= \int\limits_{1/n^2}^{n^2} e^{-x^2}\left[\int\limits_{1/(n^3x)}^{n^3/x} xe^{-(xy)^2}\,dy\right]dx \leqslant \int\limits_{1/n^2}^{n^2} e^{-x^2}\left[\int\limits_{1/n^5}^{n^5} xe^{-(xy)^2}\,dy\right]dx \\
&= \int\limits_{1/n^5}^{n^5}\left[\int\limits_{1/n^2}^{n^2} xe^{-x^2(y^2+1)}\,dx\right]dy = \int\limits_{1/n^5}^{n^5} \frac{1}{2(y^2+1)}\left[e^{-(y^2+1)/n^4} - e^{-(y^2+1)n^4}\right]dy \\
&\leqslant \int\limits_{1/n^5}^{n^5} \frac{1}{2(y^2+1)}\,dy = \frac{1}{2}\left[\operatorname{arctg} n^5 - \operatorname{arctg}\frac{1}{n^5}\right] \underset{n}{\longrightarrow} \frac{\pi}{4}.
\end{aligned}$$

Z twierdzenia o trzech ciągach uzyskujemy

$$A_n \underset{n}{\longrightarrow} \frac{\pi}{4}.$$

Wzór

$$\int\limits_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$$

można wyprowadzić inaczej. Oznaczmy

$$I_n = \int\limits_{-n}^{n} e^{-x^2}\,dx.$$

$$I_n^2 = \int\limits_{-n}^{n}\left(\int\limits_{-n}^{n} e^{-x^2}e^{-y^2}\,dy\right)dx = \int\limits_{-n}^{n}\left(\int\limits_{-n}^{n} e^{-(x^2+y^2)}\,dy\right)dx.$$

Liczba I_n^2 jest równa objętości obszaru pod wykresem funkcji

$$z = e^{-(x^2+y^2)}, \quad -n \leqslant x, y \leqslant n$$

(por. Uwaga 1.34). Ten obszar zawiera w sobie obszar pod wykresem odpowiadający kołu $x^2+y^2 \leqslant n^2$. Z kolei jest zawarty w obszarze pod wykresem odpowiadający kołu $x^2+y^2 \leqslant 2n^2$, bo

$$\{(x,y)\ : x^2+y^2 \leqslant n^2\} \subset [-n,n]\times[-n,n] \subset \{(x,y)\ : x^2+y^2 \leqslant 2n^2\}.$$

Obszary nad kołami możemy uzyskać przez obrót obszarów płaskich pod wykresami jednej funkcji

$$\begin{aligned} z &= e^{-x^2}, \quad 0 \leqslant x \leqslant n, \\ z &= e^{-x^2}, \quad 0 \leqslant x \leqslant n\sqrt{2}. \end{aligned}$$

wokół osi pionowej z, czyli $x = y = 0$.

Ogólnie, gdy obracamy wokół osi z obszar płaski pod wykresem funkcji

$$z = f(x) \geqslant 0, \quad 0 \leqslant x \leqslant a,$$

to w wyniku otrzymujemy obszar pod wykresem funkcji

$$z = f(\sqrt{x^2+y^2}), \qquad \sqrt{x^2+y^2} \leqslant a.$$

Z powyższych rozważań wynika, że

$$I_n^2 \geqslant 2\pi \int\limits_0^n x e^{-x^2}\,dx = -\pi e^{-x^2}\Big|_{x=0}^{x=n} = \pi(1-e^{-n^2}) \underset{n}{\to} \pi.$$

Z drugiej strony

$$I_n^2 \leqslant 2\pi \int\limits_0^{n\sqrt{2}} xe^{-x^2}\,dx = -\pi e^{-x^2}\Big|_{x=0}^{x=n\sqrt{2}} = \pi(1-e^{-2n^2}) \underset{n}{\to} \pi.$$

Z twierdzenia o trzech ciągach otrzymujemy $I_n^2 \underset{n}{\to} \pi$.

1.5.2 Całka Dirichleta

Całka niewłaściwa

$$\int\limits_0^\infty \frac{\sin x}{x}\,dx$$

jest zbieżna. Ta wielkość pojawia się w kilku działach matematyki, np. w teorii szeregów i całek Fouriera, teorii funkcji zespolonych.

Jedynym punktem osobliwym jest ∞. Nieścisłe obliczenie wartości całki jest następujące.

$$\int\limits_0^\infty \frac{\sin x}{x}\,dx = \int\limits_0^\infty \sin x \left(\int\limits_0^\infty e^{-xy}\,dy\right)\,dx = \int\limits_0^\infty \left(\int\limits_0^\infty e^{-xy}\sin x\,dx\right)\,dy$$
$$\int\limits_0^\infty \left[-\frac{e^{-xy}(\cos x + y\sin x)}{y^2+1}\Big|_{x=0}^{\infty}\right]\,dy = \int\limits_0^\infty \frac{1}{y^2+1}\,dy = \frac{\pi}{2}.$$

Podobnie jak przy całce Eulera zamiana kolejności całkowania wymaga uzasadnienia. Całka Dirichleta, w odróżnieniu od całki Eulera, nie jest bezwzględnie zbieżna, przez co ścisłe obliczenie jej wartości jest bardziej kłopotliwe.

Oznaczmy

$$I_n = \int\limits_{1/n}^{n} \frac{\sin x}{x}\,dx.$$

Wtedy

$$I_n \underset{n}{\longrightarrow} \int\limits_0^\infty \frac{\sin x}{x}\,dx.$$

Dalej

$$I_n = \int_{1/n}^{n} \sin x \left[\int_{0}^{\infty} e^{-xy}\, dy \right] dx$$

$$= \underbrace{\int_{1/n}^{n} \sin x \left[\int_{0}^{n^2} e^{-xy}\, dy \right] dx}_{S_n} + \underbrace{\int_{1/n}^{n} \sin x \left[\int_{n^2}^{\infty} e^{-xy}\, dy \right] dx}_{R_n}.$$

Mamy

$$|R_n| \leqslant \int_{1/n}^{n} \left[\int_{n^2}^{\infty} e^{-xy}\, dy \right] dx = \int_{1/n}^{n} \frac{e^{-n^2 x}}{x}\, dx \leqslant ne^{-n} \cdot n \xrightarrow[n]{} 0.$$

$$S_n = \int_{0}^{n^2} \left[\int_{1/n}^{n} e^{-xy} \sin x\, dx \right] dy = \int_{0}^{n^2} \left[-\frac{e^{-xy}(\cos x + y \sin x)}{y^2+1} \Bigg|_{x=1/n}^{x=n} \right] dy$$

$$= \int_{0}^{n^2} \left[\frac{e^{-y/n}\left(\cos\frac{1}{n} + y \sin\frac{1}{n}\right)}{y^2+1} - \frac{e^{-ny}\,(\cos n + y \sin n)}{y^2+1} \right] dy$$

$$= \cos\frac{1}{n} \int_{0}^{n^2} \frac{1}{y^2+1} e^{-y/n}\, dy + \sin\frac{1}{n} \int_{0}^{n^2} \frac{y}{y^2+1} e^{-y/n}\, dy$$

$$- \cos n \int_{0}^{n^2} \frac{1}{y^2+1} e^{-ny}\, dy - \sin n \int_{0}^{n^2} \frac{y}{y^2+1} e^{-ny}\, dy$$

Suma wartości bezwzględnych dwu ostatnich składników jest mniejsza niż

$$\int_{0}^{n^2} \frac{y}{y^2+1} e^{-ny}\, dy + \int_{0}^{n^2} \frac{1}{y^2+1} e^{-ny}\, dy \leqslant \left(\frac{1}{2} + 1\right) \int_{0}^{n^2} e^{-ny}\, dy$$

$$= \frac{3}{2n}[1 - e^{-n^3}] \leqslant \frac{3}{2n} \xrightarrow[n]{} 0.$$

Drugi składnik można oszacować z góry następująco:

$$\sin\frac{1}{n}\int\limits_0^{n^2}\frac{y}{y^2+1}e^{-y/n}\,dy \leqslant \frac{1}{n}\int\limits_0^{n^2}\frac{y}{y^2+1}\,dy = \frac{\log(n^4+1)}{2n}$$
$$\leqslant \frac{\log(2n^4)}{2n} = \frac{4\log n+\log 2}{2n} \underset{n}{\longrightarrow} 0\,.$$

W pierwszym składniku $\cos\frac{1}{n}$ dąży do 1. Dalej

$$\int\limits_0^{n^2}\frac{1}{y^2+1}e^{-y/n}\,dy \leqslant \int\limits_0^{n^2}\frac{1}{y^2+1}\,dy = \operatorname{arctg} n^2 \leqslant \frac{\pi}{2}.$$

Z drugiej strony

$$\int\limits_0^{n^2}\frac{1}{y^2+1}e^{-y/n}\,dy \geqslant \int\limits_0^{\sqrt{n}}\frac{1}{y^2+1}e^{-y/n}\,dy \geqslant e^{-1/\sqrt{n}}\int\limits_0^{\sqrt{n}}\frac{1}{y^2+1}\,dy$$
$$= e^{-1/\sqrt{n}}\operatorname{arctg}\sqrt{n} \underset{n}{\longrightarrow} \frac{\pi}{2}.$$

Reasumując, otrzymujemy $S_n \underset{n}{\longrightarrow} \frac{\pi}{2}$, zatem $I_n \underset{n}{\longrightarrow} \frac{\pi}{2}$.

2 Charakteryzacja funkcji całkowalnych w sensie Riemanna

Definicja 2.1. *Zbiór $A \subset \mathbb{R}$ nazywamy zbiorem miary zero, jeśli dla dowolnej liczby $\varepsilon > 0$ istnieje pokrycie zbioru A skończoną lub przeliczalną sumą przedziałów otwartych (a_n, b_n) takich, że*

$$\sum_{n=1}^{N}(b_n-a_n) \leqslant \varepsilon,\ \ N\in\mathbb{N}, \text{lub } N=\infty.$$

Przykład. Zbiory skończone mają miarę zero, bo

$$\{x_1,\, x_2, \ldots,\, x_N\} \subset \bigcup_{n=1}^{N}\left(x_n-\frac{\varepsilon}{2N}, x_n+\frac{\varepsilon}{2N}\right).$$

Uwaga 1. Możemy przyjąć, że $N=\infty$, dołączając w razie potrzeby puste przedziały postaci (a_j, b_j), gdzie $b_j = a_j$, dla $j > N$.

Lemat 2.2. *Przeliczalna suma zbiorów miary zero jest zbiorem miary zero.*

Dowód. Zakładamy, że zbiory A_j, $j \geqslant 1$, mają miarę zero i rozważamy zbiór

$$A = \bigcup_{j=1}^{\infty} A_j.$$

Dla liczby $\varepsilon > 0$ i wskaźnika $j \geqslant 1$ istnieją przedziały (a_{nj}, b_{nj}) takie, że

$$A_j \subset \bigcup_{n=1}^{\infty} (a_{nj}, b_{nj}), \quad \sum_{n=1}^{\infty} (b_{nj} - a_{nj}) \leqslant \frac{\varepsilon}{2^j}.$$

Zatem

$$A \subset \bigcup_{j=1}^{\infty} \bigcup_{n=1}^{\infty} (a_{nj}, b_{nj})$$

oraz

$$\sum_{j=1}^{\infty} \sum_{n=1}^{\infty} (b_{nj} - a_{nj}) \leqslant \varepsilon \sum_{j=1}^{\infty} \frac{1}{2^j} = \varepsilon.$$

Uwaga. Na podstawie lematu można zmienić definicję zbioru miary zero dopuszczając pokrycie przedziałami domkniętymi $[a_n, b_n]$.

Wniosek 2.3. *Zbiór przeliczalny ma miarę zero. W szczególności zbiór liczb wymiernych* $\mathbb{Q}$ *ma miarę zero.*

Przykład. Zbiór miary zero może być nieprzeliczalny. Zbiór Cantora C jest mocy continuum i ma miarę zero. Zbiór Cantora składa się z liczb przedziału $[0, 1]$, które w nieskończonym rozwinięciu w systemie trójkowym mają cyfry 0 i 2. Dla ustalonej wartości n rozważmy liczby $x < 1$ zbioru Cantora, dla których $x_{n+1} = 0$. Wtedy

$$x \in [0.x_1 \ldots x_n, 0.x_1 \ldots x_n 1].$$

Długość przedziału wynosi 3^{-n-1}. Mamy 2^n takich przedziałów, więc opisane wyżej liczby są zawarte w skończonej sumie przedziałów o sumarycznej długości $2^n 3^{-n-1}$. Liczby zbioru Cantora, dla których $x_{n+1} = 2$ można uzyskać z liczb, dla których $x_{n+1} = 0$ poprzez dodanie liczby $0.0 \ldots 02 = 2 \cdot 3^{-n-1}$. Zatem zbiór C jest zawarty w sumie przedziałów o sumarycznej długości $2^{n+1} 3^{-n-1}$. Stąd wynika, że zbiór ten jest miary zero. $\square$

Twierdzenie 2.4. *Ograniczona funkcja $f(x)$ na przedziale $[a, b]$ jest całkowalna w sensie Riemanna wtedy i tylko wtedy, gdy zbiór punktów nieciągłości funkcji $f(x)$ ma miarę zero.*

Dowód. Dla liczby $\delta > 0$ i punktu $a \leqslant x \leqslant b$ określmy wielkość

$$\omega_\delta(f,x) = \sup_{\substack{|x'-x|<\delta \\ |x''-x|<\delta}} |f(x') - f(x'')| = \sup_{|x'-x|<\delta} f(x') - \inf_{|x'-x|<\delta} f(x').$$

Niech

$$\omega(f,x) = \lim_{\delta \to 0^+} \omega_\delta(f,x) = \inf_{\delta>0} \omega_\delta(f,x).$$

Liczbę $\omega(f,x)$ nazywamy oscylacją funkcji f w punkcie x.

Lemat 2.5. *Ograniczona funkcja f na przedziale jest ciągła w punkcie x wtedy i tylko wtedy, gdy $\omega(f,x) = 0$.*

Dowód. Załóżmy, że f jest ciągła w x. Wtedy dla liczby $\varepsilon > 0$ istnieje $\delta > 0$ taka, że za warunku $|x' - x| < \delta$ wynika $|f(x') - f(x)| < \varepsilon/2$. Wtedy dla $|x' - x| < \delta$, $|x'' - x| < \delta$ otrzymujemy

$$|f(x') - f(x'')| \leqslant |f(x') - f(x)| + |f(x'') - f(x)| < \varepsilon.$$

Stąd

$$\omega(f,x) \leqslant \omega_\delta(f,x) \leqslant \varepsilon,$$

czyli $\omega(f,x) = 0$.

Jeśli $\omega(f,x) = 0$ i $\varepsilon > 0$, to $\omega_\delta(f,x) < \varepsilon$ dla pewnej liczby $\delta > 0$. Wtedy dla $|x' - x| < \delta$ otrzymujemy $|f(x') - f(x)| < \varepsilon$. To oznacza, że funkcja f jest ciągła w punkcie x. □

Wracamy do dowodu twierdzenia. Załóżmy, że funkcja f jest całkowalna w sensie Riemanna na przedziale $[a,b]$. Niech X oznacza zbiór punktów nieciągłości funkcji f, czyli

$$X = \{x : \omega(f,x) > 0\}.$$

Oznaczmy

$$X_n = \left\{x : \omega(f,x) > \frac{1}{n}\right\}.$$

Wtedy

$$X = \bigcup_{n=1}^{\infty} X_n.$$

Wystarczy pokazać, że X_n jest miary zero dla dowolnej liczby n. Z całkowalności funkcji f istnieje podział $P = \{x_0, x_1, \ldots, x_N\}$ przedziału $[a,b]$ taki, że

$$\sum_{i=1}^{N}(M_i - m_i)\Delta x_i < \frac{\varepsilon}{n},$$

gdzie m_i i M_i oznaczają kresy dolny i górny funkcji f na przedziale $[x_{i-1}, x_i]$. Podzielimy liczby $1, 2, \ldots, N$ na dwa podzbiory A i B. Do A zaliczymy, te wskaźniki i, dla których X_n ma niepusty przekrój z przedziałem (x_{i-1}, x_i). W związku z tym B składa się z tych liczb i, dla których $X_n \cap (x_{i-1}, x_i) = \emptyset$. Wtedy

$$X_n \subseteq \bigcup_{i\in A}[x_{i-1}, x_i] \cup \{x_0, x_1, \ldots, x_N\}. \tag{2.1}$$

Zauważmy, że jeśli $i \in A$, to $M_i - m_i \geqslant \dfrac{1}{n}$. Stąd wynika, że

$$\frac{1}{n}\sum_{i\in A}\Delta x_i \leqslant \sum_{i\in A}(M_i - m_i)\Delta x_i \leqslant \sum_{i=1}^{N}(M_i - m_i)\Delta x_i < \frac{\varepsilon}{n},$$

czyli

$$\sum_{i\in A}\Delta x_i < \varepsilon.$$

Na podstawie (2.1) wnioskujemy, że X_n jest miary zero.

Odwrotnie, załóżmy, że zbiór X jest miary zero. Dla liczby $\varepsilon > 0$ istnieją przedziały (a_n, b_n) takie, że

$$X \subseteq \bigcup_{n=1}^{\infty}(a_n, b_n), \quad \sum_{n=1}^{\infty}(b_n - a_n) < \varepsilon.$$

Dla $x \in [a,b] \setminus X$ mamy $\omega(f,x) = 0$. Zatem istnieje liczba $\delta_x > 0$ taka, że $\omega_{\delta_x}(f,x) < \varepsilon$. To oznacza, że jeśli $V_x = (x-\delta_x, x+\delta_x)$, to $M_{V_x}(f) - m_{V_x}(f) < \varepsilon$. Mamy

$$[a,b] \subset \bigcup_{n=1}^{\infty}(a_n, b_n) \cup ([a,b] \setminus X) \subseteq \bigcup_{n=1}^{\infty}(a_n, b_n) \cup \bigcup_{x\in[a,b]\setminus X} V_x.$$

Lemat 2.6. *Jeśli przedział* $[a,b]$ *jest zawarty w rodzinie przedziałów* (c_λ, d_λ), $\lambda \in \Lambda$, *to*

$$[a,b] \subseteq (c_{\lambda_1}, d_{\lambda_1}) \cup (c_{\lambda_1}, d_{\lambda_1}) \cup \ldots \cup (c_{\lambda_K}, d_{\lambda_K})$$

dla pewnego skończonego podzbioru λ_1, $\lambda_2 \ldots, \lambda_K$.

Dowód. Dla ustalonego parametru λ istnieją dwa ciągi $c_{\lambda,n}$ i $d_{\lambda,n}$ o wyrazach wymiernych spełniające

$$c_\lambda \underset{n}{\swarrow} c_{\lambda,n} \leqslant d_{\lambda,n} \underset{n}{\nearrow} d_\lambda.$$

Wtedy

$$[a,b] \subseteq \bigcup_{\lambda\in\Lambda}(c_\lambda, d_\lambda) = \bigcup_{\lambda\in\Lambda}\bigcup_{n=1}^{\infty}(c_{\lambda,n}, d_{\lambda,n}).$$

Ta rodzina przedziałów otwartych jest przeliczalna, zatem

$$[a,b] \subseteq \bigcup_{k=1}^{\infty}(c_{\lambda_k,n_k}, d_{\lambda_k,n_k}) \subseteq \bigcup_{k=1}^{\infty}(c_{\lambda_k}, d_{\lambda_k}),$$

dla pewnego ciągu parametrów λ_k i pewnego ciągu liczb naturalnych n_k. Załóżmy, nie wprost, że

$$[a,b] \not\subseteq \bigcup_{k=1}^{K}(c_{\lambda_k}, d_{\lambda_k})$$

dla dowolnej wartości K. Tzn. istnieje punkt $x_K \in [a,b]$ leżący poza $(c_{\lambda_k}, d_{\lambda_k})$ dla $k = 1,2,\ldots,K$. Ciąg x_K zawiera podciąg zbieżny do punktu z $[a,b]$ z twierdzenia Bolzano-Weierstrassa. Oznaczmy symbolem x_0 granicę tego podciągu. Wtedy $x_0 \in (c_{\lambda_L}, d_{\lambda_L})$, dla pewnego L. Ale w przedziale $(c_{\lambda_L}, d_{\lambda_L})$ nie ma żadnego z punktów x_K dla $K \geqslant L$. Zatem x_0 nie leży w $(c_{\lambda_L}, d_{\lambda_L})$, co prowadzi do sprzeczności. □

Z lematu otrzymujemy

$$[a,b] \subset \bigcup_{n=1}^{N}(a_n, b_n) \cup \bigcup_{k=1}^{K} V_{x_k}.$$

Niech P będzie podziałem przedziału $[a,b]$ takim, że każdy jego domknięty podprzedział jest zawarty w jednym z przedziałow (a_n, b_n) dla $n = 1,2\ldots N$ lub V_{x_k} dla $k = 1,2,\ldots,K$. Wtedy

$$U(P,f) - L(P,f) < \varepsilon(b-a) + 2M\varepsilon,$$

gdzie $M = \sup\limits_{a\leqslant x\leqslant b}|f(x)|$. Ostatnia nierówność wynika z faktu, że jeśli przedział podziału jest zawarty w przedziale postaci (a_n, b_n), to rozpiętość wartości funkcji na tym przedziale jest nie większa niż $2M$, a suma długości takich przedziałów jest mniejsza niż ε. Z kolei jeśli przedział podziału jest zawarty w pewnym V_{x_k}, to oscylacja funkcji f na tym przedziale jest mniejsza niż ε a suma długości takich (wszystkich) przedziałów nie przekracza $b-a$. □

3 Całka Riemanna-Stieltjesa

Niech $\alpha : [a,b] \longrightarrow \mathbb{R}$ będzie funkcją rosnącą (niekoniecznie ściśle rosnącą) spełniającą $\alpha(a) < \alpha(b)$. Można traktować przedział $[a,b]$ jako pręt wykonany z niejednorodnego materiału, dla którego wielkość $\alpha(x)$ oznacza masę fragmentu pręta od punktu a do punktu x. Dla podziału $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ przedziału $[a,b]$ określamy

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}), \quad i = 1, 2, \dots, n.$$

Dalej dla ograniczonej funkcji $f : [a,b] \longrightarrow \mathbb{R}$ definiujemy sumy górne i dolne wzorami

$$U(\mathcal{P}, f, \alpha) = \sum_{i=1}^{n} M_i \, \Delta\alpha_i, \qquad L(\mathcal{P}, f, \alpha) = \sum_{i=1}^{n} m_i \, \Delta\alpha_i,$$

gdzie

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

Podobnie jak przy zwykłej całce Riemanna określamy całki górne i dolne

$$\int_{\underline{a}}^{b} f(x)\, d\alpha(x) = \sup_{\mathcal{P}} L(\mathcal{P}, f, \alpha), \quad \int_{a}^{\overline{b}} f(x)\, d\alpha(x) = \inf_{\mathcal{P}} U(\mathcal{P}, f, \alpha).$$

Jeśli $\displaystyle\int_{\underline{a}}^{b} f(x)\, d\alpha(x) = \int_{a}^{\overline{b}} f(x)\, d\alpha(x)$, to mówimy, że funkcja f jest całkowalna względem funkcji α i wspólną wartość oznaczamy symbolem $\displaystyle\int_{a}^{b} f(x)\, d\alpha(x)$. Całka nosi nazwę całki Riemanna-Stieltjesa. Całkowalność zapisujemy w skrócie poprzez $f \in \mathcal{R}(\alpha)$. Dla $\alpha(x) = x$ otrzymujemy zwykłe pojęcie całki Riemanna.

Przykłady

(a) $\alpha : [0,2] \longrightarrow \mathbb{R}, \qquad \alpha(x) = \begin{cases} 0 & 0 \leqslant x \leqslant 1, \\ 1 & 1 < x \leqslant 2. \end{cases}$.

(b) $\alpha : [0,1] \longrightarrow \mathbb{R}, \qquad \alpha(x) = x^2.$

W przykładzie (a) rozważmy podział

$$\mathcal{P}_n = \left\{0, 1 - \frac{1}{n}, 1 + \frac{1}{n}, 2\right\}.$$

Wtedy

$$L(\mathcal{P}_n, f.\alpha) = \inf_{1-\frac{1}{n} \leqslant x \leqslant 1+\frac{1}{n}} f(x), \quad U(\mathcal{P}_n, f.\alpha) = \sup_{1-\frac{1}{n} \leqslant x \leqslant 1+\frac{1}{n}} f(x).$$

Jeśli funkcja $f(x)$ jest ciągła w punkcie 1, to

$$L(\mathcal{P}_n, f, \alpha) \underset{n}{\longrightarrow} f(1), \quad U(\mathcal{P}_n, f, \alpha) \underset{n}{\longrightarrow} f(1).$$

Niech $\mathcal{P}^*$ będzie rozdrobnieniem podziału $\mathcal{P}$, tzn. $\mathcal{P} \subset \mathcal{P}^*$. Wtedy

$$L(\mathcal{P}, f, \alpha) \leqslant L(\mathcal{P}^*, f, \alpha), \quad U(\mathcal{P}^*, f, \alpha) \leqslant U(\mathcal{P}, f, \alpha).$$

Uzasadnienie jest identyczne jak dla zwykłej całki Riemanna. Z tych nierówności wynika, że dla dowolnych dwu podziałów $\mathcal{P}_1$ i $\mathcal{P}_2$ mamy

$$L(\mathcal{P}_1, f, \alpha) \leqslant U(\mathcal{P}_2, f, \alpha),$$

co pociąga

$$\underline{\int_{a}^{b}} f(x)\, d\alpha(x) \leqslant \overline{\int_{a}^{b}} f(x)\, d\alpha(x).$$

Podobne rozumowanie jak dla zwykłej całki Riemanna prowadzi do następnego twierdzenia.

Twierdzenie 3.1. *Warunek $f \in \mathcal{R}(\alpha)$ jest spełniony wtedy i tylko wtedy, gdy dla dowolnej dodatniej liczby ε istnieje podział $\mathcal{P}$ przedziału $[a, b]$ taki, że*

$$U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) < \varepsilon.$$

Twierdzenie 3.2. *Jeśli dla podziału $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ i liczby $\varepsilon > 0$ spełniony jest warunek $U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) < \varepsilon$, to dla dowolnego wyboru punktów s_i, t_i z przedziału $[x_{i-1}, x_i]$ mamy*

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \Delta\alpha_i < \varepsilon.$$

Ponadto jeśli $f \in \mathcal{R}(\alpha)$, to

$$\left| \sum_{i=1}^{n} f(t_i) \Delta\alpha_i - \int_{a}^{b} f(x)\, d\alpha(x) \right| < \varepsilon.$$

Dowód. Mamy $|f(s_i) - f(t_i)| \leqslant M_i - m_i$. Zatem

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)|\Delta\alpha_i \leqslant \sum_{i=1}^{n}(M_i - m_i)\Delta\alpha_i = U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) < \varepsilon.$$

Jeśli $f \in \mathcal{R}(\alpha)$, to

$$L(\mathcal{P}, f, \alpha) \leqslant \int\limits_a^b f(x)\, d\alpha(x) \leqslant U(\mathcal{P}, f, \alpha)$$

oraz

$$L(\mathcal{P}, f, \alpha) \leqslant \sum_{i=1}^{n} f(t_i)\Delta\alpha_i \leqslant U(\mathcal{P}, f, \alpha)$$

dla dowolnego wyboru punktów t_i z $[x_{i-1}, x_i]$. Zatem

$$\left|\sum_{i=1}^{n} f(t_i)\Delta\alpha_i - \int\limits_a^b f(x)\, d\alpha(x)\right| \leqslant U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) < \varepsilon.$$

□

Twierdzenie 3.3. *Jeśli $f(x)$ jest funkcją ciągłą na $[a, b]$, to $f \in \mathcal{R}(\alpha)$.*

Dowód. Przy dowodzie skorzystamy z Twierdzenia 3.1. Ustalmy liczbę $\varepsilon > 0$. Dla liczby $\eta > 0$ istnieje liczba $\delta > 0$, taka, że

$$|s - t| < \delta \implies |f(s) - f(t)| < \eta.$$

Niech $\mathcal{P}$ będzie podziałem przedziału $[a, b]$ na n równych części tak, aby $\dfrac{b-a}{n} < \delta$. Wtedy $M_i - m_i < \eta$ dla $i = 1, 2, \ldots, n$. Stąd

$$\begin{aligned} U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) &= \sum_{i=1}^{n}(M_i - m_i)\Delta\alpha_i < \eta \sum_{i=1}^{n} \Delta\alpha_i \\ &= \eta \sum_{i=1}^{n}[\alpha(x_i) - \alpha(x_{i-1})] = \eta[\alpha(b) - \alpha(a)]. \end{aligned}$$

Wystarczy teraz przyjąć $\eta = \dfrac{\varepsilon}{\alpha(b) - \alpha(a)}$. □

Twierdzenie 3.4. *Jeśli $\alpha(x)$ jest ciągła na przedziale $[a, b]$, natomiast funkcja $f(x)$ jest monotoniczna na $[a, b]$, to $f \in \mathcal{R}(\alpha)$.*

Dowód. Rozpatrzymy przypadek funkcji rosnącej $f(x)$. Dla ustalonej liczby naturalnej n rozważmy wielkości

$$\alpha(a) < \alpha(a) + i\frac{\alpha(b) - \alpha(a)}{n} < \alpha(b), \quad i = 1, \ldots, n-1.$$

Z własności Darboux istnieją punkty $x_i \in (a, b)$ takie, że

$$\alpha(x_i) = \alpha(a) + i\frac{\alpha(b) - \alpha(a)}{n}, \quad i = 1, \ldots, n-1.$$

Przyjmijmy $x_0 = a$ i $x_n = b$. Wtedy $x_0 < x_1 < \ldots < x_n$ oraz

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) = \frac{\alpha(b) - \alpha(a)}{n}.$$

Dalej dla $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ otrzymujemy

$$\begin{aligned} U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) &= \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} (M_i - m_i) \\ &= \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] = \frac{\alpha(b) - \alpha(a)}{n} [f(b) - f(a)]. \end{aligned}$$

Jeśli n jest dostatecznie duże, to prawa strona jest mniejsza niż z góry ustalona liczba $\varepsilon > 0$. □

Twierdzenie 3.3 można uogólnić dopuszczając skończenie wiele punktów nieciągłości całkowanej funkcji.

Twierdzenie 3.5. *Niech $f(x)$ będzie funkcją ograniczoną na $[a, b]$ mającą skończenie wiele punktów nieciągłości. Załóżmy, że funkcja $\alpha(x)$ jest ciągła w każdym punkcie nieciągłości funkcji $f(x)$. Wtedy $f \in \mathcal{R}(\alpha)$.*

Dowód. Ustalmy $\varepsilon > 0$. Oznaczmy $M = \sup\limits_{a \leqslant x \leqslant b} |f(x)|$. Niech k oznacza liczbę punktów nieciągłości funkcji $f(x)$. Każdy taki punkt nieciągłości leży w przedziale $[u_j, v_j] \subset [a, b]$, $i = 1, 2, \ldots, k$ takim, że

$$\alpha(v_j) - \alpha(u_j) < \frac{\varepsilon}{k}, \quad j = 1, 2, \ldots, k.$$

Można przyjąć, ewentualnie zwiększając u_j i zmniejszając v_j, że przedziały $[u_j, v_j]$ są rozłączne, tzn.

$$u_1 < v_1 < u_2 < v_2 < \ldots < u_k < v_k.$$

Ponadto możemy przyjąć (w razie potrzeby poszerzając przedziały), że każdy punkt nieciągłości funkcji $f(x)$ leżący w (a,b) leży w (u_j,v_j) dla pewnej wartości j. Punkty a i b mogą być też punktami nieciągłości funkcji $f(x)$. Wtedy a leży w $[a,v_1)$ oraz b leży w $(u_k,b]$.

Usuńmy przedziały (u_j,v_j) z $[a,b]$. Pozostały zbiór

$$K = [a,b] \setminus \bigcup_{j=1}^{k} (u_j,v_j)$$

składa się ze skończonej liczby domkniętych przedziałów. Funkcja $f(x)$ jest jednostajnie ciągła na zbiorze K. Istnieje więc liczba $\delta > 0$ taka, że jeśli $s,t \in K$ oraz $|s-t| < \delta$, to $|f(s)-f(t)| < \varepsilon$.

Niech $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ będzie podziałem przedziału $[a,b]$ o własnościach:

(a) $\mathcal{P}$ zawiera wszystkie punktu u_j i v_j.

(a) $\mathcal{P}$ nie zawiera punktów z (u_j,v_j) dla żadnego j.

(c) Jeśli x_{i-1} jest różny od u_j (czyli (x_{i-1},x_i) nie jest postaci (u_j,v_j)), to $x_i - x_{i-1} < \delta$.

To oznacza, że przedziały podziału $\mathcal{P}$ złożone są z przedziałów $[u_j,v_j]$ oraz z przedziałów pomiędzy nimi, o długości mniejszej niż δ.

Zauważmy, że $M_i - m_i \leqslant 2M$ dla każdego $i = 1,2,\ldots,n$ oraz $M_i - m_i \leqslant \varepsilon$ jeśli x_{i-1} jest różny od u_j. Niech A oznacza zbiór indeksów i, dla których x_{i-1} jest różny od u_j. Zbiór pozostałych indeksów oznaczymy symbolem B. Wtedy

$$\begin{aligned} U(\mathcal{P},f,\alpha) - L(\mathcal{P},f,\alpha) &= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \\ &= \sum_{i\in A} (M_i - m_i)\Delta\alpha_i + \sum_{i\in B} (M_i - m_i)\Delta\alpha_i \leqslant \varepsilon \sum_{i\in A} \Delta\alpha_i + 2M \sum_{i\in B} \Delta\alpha_i \\ &\leqslant \varepsilon \sum_{i=1}^{n} \Delta\alpha_i + 2Mk\frac{\varepsilon}{k} = \varepsilon[\alpha(b) - \alpha(a) + 2M]. \end{aligned}$$

□

Uwaga. Założenia w twierdzeniu są istotne. Niech

$$f(x) = \alpha(x) = \begin{cases} 0 & 0 \leqslant x < 1 \\ 1 & 1 \leqslant x \leqslant 2. \end{cases}$$

Rozważmy dowolny podział $\mathcal{P}$ przedziału $[0, 2]$. Niech x_j będzie punktem podziału o najmniejszym wskaźniku takim, że $x_j \geqslant 1$. Wtedy $x_{j-1} < 1$. Funkcja $f(x)$ jest stała na każdym przedziale $[x_{i-1}, x_i]$ dla $i \neq j$. Zatem

$$U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) = \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i = (M_j - m_j)[\alpha(x_j) - \alpha(x_{j-1})] = 1.$$

To oznacza, że $f \notin \mathcal{R}(\alpha)$, czyli $\alpha \notin \mathcal{R}(\alpha)$.

Określmy funkcję

$$g(x) = \begin{cases} 0 & 0 \leqslant x \leqslant 1 \\ 1 & 1 < x \leqslant 2. \end{cases}$$

Funkcja $g(x)$ jest równa funkcji $f(x)$ we wszystkich punktach $x \neq 1$. Mamy $g \in \mathcal{R}(\alpha)$. Rzeczywiście, niech $\mathcal{P}$ będzie podziałem zawierającym punkt $1 = x_j$. Funkcja $\alpha(x)$ jest stała na każdym przedziale $[x_{i-1}, x_i]$ dla $i \neq j$. Tzn. $\Delta\alpha_i = 0$ dla $i \neq j$. Jedynie na przedziale $[x_{j-1}, x_j] = [x_{j-1}, 1]$ funkcja $\alpha(x)$ nie jest stała. Ale $M_j - m_j = 0$, bo funkcja $g(x)$ stała na przedziale $[x_{j-1}, 1]$. Zatem

$$U(\mathcal{P}, g, \alpha) - L(\mathcal{P}, g, \alpha) = \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i = 0.$$

Wszystkie twierdzenia dotyczące arytmetycznych własności całki Riemanna są prawdziwe również dla całki Riemanna-Stieltjesa, bez zmiany dowodów. Ponadto

$$\left| \int_a^b f(x)\, d\alpha(x) \right| \leqslant M[\alpha(b) - \alpha(a)], \quad \text{gdzie } M = \sup_{a \leqslant x \leqslant b} |f(x)|. \tag{3.1}$$

Rzeczywiście

$$\begin{aligned} U(\mathcal{P}, f, \alpha) &= \sum_{i=1}^{n} M_i \Delta\alpha_i \leqslant M \sum_{i=1}^{n} \Delta\alpha_i = M[\alpha(b) - \alpha(a)], \\ L(\mathcal{P}, f, \alpha) &= \sum_{i=1}^{n} m_i \Delta\alpha_i \geqslant -M \sum_{i=1}^{n} \Delta\alpha_i = -M[\alpha(b) - \alpha(a)]. \end{aligned}$$

Ponadto jeśli $f \in \mathcal{R}(\alpha_1) \cap \mathcal{R}(\alpha_2)$, to $f \in \mathcal{R}(\alpha_1 + \alpha_2)$ oraz

$$\int_a^b f(x) d(\alpha_1 + \alpha_2)(x) = \int_a^b f(x) d\alpha_1(x) + \int_a^b f(x) d\alpha_2(x). \tag{3.2}$$

Istotnie teza i wzór wynikają z równości

$$\begin{aligned} U(\mathcal{P}, f, \alpha_1 + \alpha_2) &= U(\mathcal{P}, f, \alpha_1) + U(\mathcal{P}, f, \alpha_2), \\ L(\mathcal{P}, f, \alpha_1 + \alpha_2) &= L(\mathcal{P}, f, \alpha_1) + L(\mathcal{P}, f, \alpha_2). \end{aligned}$$

Podobnie można uzasadnić, że jeśli $f \in \mathcal{R}(\alpha)$ i $t > 0$, to $f \in \mathcal{R}(t\alpha)$ oraz

$$\int_a^b f(x) d(t\alpha)(x) = t \int_a^b f(x) d\alpha(x).$$

Twierdzenie 3.6. *Jeśli $a < c < b$ oraz funkcja $f(x)$ jest ograniczona i ciągła w punkcie c, to dla*

$$\alpha_c(x) = \begin{cases} 0 & x \leqslant c, \\ 1 & x > c, \end{cases}$$

mamy

$$\int_a^b f(x) d\alpha_c(x) = f(c).$$

Dowód. Rozważmy podział

$$\mathcal{P}_n = \left\{ a, c, c + \frac{1}{n}, b \right\}.$$

Wtedy

$$\begin{aligned} U(\mathcal{P}_n, f, \alpha_c) &= \sup_{c \leqslant x \leqslant c + \frac{1}{n}} f(x) \xrightarrow[n]{} f(c), \\ L(\mathcal{P}_n, f, \alpha_c) &= \inf_{c \leqslant x \leqslant c + \frac{1}{n}} f(x) \xrightarrow[n]{} f(c). \end{aligned}$$

□

Twierdzenie 3.7. *Niech $d_n > 0$ i $\sum_{n=1}^{\infty} d_n < \infty$, oraz c_n będzie ciągiem różnych punktów przedziału (a, b). Dla funkcji*

$$\alpha(x) = \sum_{n=1}^{\infty} d_n \alpha_{c_n}(x)$$

i funkcji $f(x)$ ciągłej na przedziale $[a, b]$ mamy

$$\int_a^b f(x) d\alpha(x) = \sum_{n=1}^{\infty} d_n f(c_n).$$

Dowód. Szereg określający funkcję $\alpha(x)$ jest zbieżny w każdym punkcie i $\alpha(x)$ jest funkcją rosnącą. Ponadto $\alpha(a) = 0$ i $\alpha(b) = \sum_{n=1}^{\infty} d_n$. Dla ustalonej liczby $\varepsilon > 0$ wybierzmy N tak duże, aby

$$\sum_{n=N+1}^{\infty} d_n < \varepsilon.$$

Niech

$$\alpha_1(x) = \sum_{n=1}^{N} d_n\, \alpha_{c_n}(x), \qquad \alpha_2(x) = \sum_{n=N+1}^{\infty} d_n\, \alpha_{c_n}(x).$$

Funkcje $\alpha_1(x)$ i $\alpha_2(x)$ są rosnące oraz $\alpha(x) = \alpha_1(x) + \alpha_2(x)$. Na podstawie Twierdzenia 3.6 i wzoru (3.2) otrzymujemy

$$\int_a^b f(x)\, d\alpha_1(x) = \sum_{n=1}^{N} d_n \int_a^b f(x)\, d\alpha_{c_n}(x) = \sum_{n=1}^{N} d_n f(c_n).$$

Ponieważ

$$\alpha_2(b) - \alpha_2(a) = \sum_{n=N+1}^{\infty} d_n < \varepsilon,$$

to

$$\left| \int_a^b f(x)\, d\alpha_2(x) \right| \leqslant M\varepsilon, \qquad M = \sup_{a \leqslant x \leqslant b} |f(x)|.$$

W rezultacie dostajemy

$$\begin{aligned} \left| \int_a^b f(x)\, d\alpha(x) - \sum_{n=1}^{N} d_n f(c_n) \right| &= \left| \int_a^b f(x)\, d\alpha(x) - \int_a^b f(x)\, d\alpha_1(x) \right| \\ &= \left| \int_a^b f(x)\, d\alpha_2(x) \right| \leqslant M\varepsilon. \end{aligned}$$

Gdy $N \to \infty$, to

$$\left| \int_a^b f(x)\, d\alpha(x) - \sum_{n=1}^{\infty} d_n f(c_n) \right| \leqslant M\varepsilon.$$

□

Uwaga. Dla $a = 0$, $b = 1$ oraz $d_n = 2^{-n}$ niech ciąg c_n składa się ze wszystkich liczb wymiernych z przedziału $(0, 1)$. Można pokazać, że $\alpha(x)$ jest nieciągła w każdym punkcie wymiernym przedziału $(0, 1)$. **Zadanie** Udowodnić, że ograniczona funkcja $f(x)$ na przedziale $[0, 1]$ jest całkowalna względem funkcji $\alpha(x)$ wtedy i tylko wtedy, gdy $f(x)$ jest ciągła w każdym punkcie wymiernym przedziału.

W pewnych przypadkach całkę Riemanna-Stieltjesa można zamienić na całkę Riemanna.

Twierdzenie 3.8. *Załóżmy, że funkcja $\alpha(x)$ jest różniczkowalna w przedziale $[a, b]$ oraz funkcja $\alpha'(x)$ jest całkowalna w sensie Riemanna na przedziale $[a, b]$, (np. $\alpha'(x)$ jest ciągła). Wtedy warunki $f \in \mathcal{R}(\alpha)$ oraz $f\alpha' \in \mathcal{R}$ są równoważne. Ponadto*

$$\int_a^b f(x)\,d\alpha(x) = \int_a^b f(x)\alpha'(x)\,dx.$$

Dowód. Z założenia dla liczby $\varepsilon > 0$ istnieje podział $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ przedziału $[a, b]$ taki, że

$$U(\mathcal{P}, \alpha') - L(\mathcal{P}, \alpha') < \varepsilon. \tag{3.3}$$

Z twierdzenia Lagrange'a otrzymujemy

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) = \alpha'(t_i)\Delta x_i,$$

dla pewnych punktów $t_i \in [x_{i-1}, x_i]$, $i = 1, 2, \ldots, n$. Dla dowolnego wyboru punktów $s_i \in [x_{i-1}, x_i]$ dostajemy

$$\begin{aligned}\sum_{i=1}^n |\alpha'(t_i) - \alpha'(s_i)|\Delta x_i \leqslant \sum_{i=1}^n [M_i(\alpha') - m_i(\alpha')]\,\Delta x_i \\ = U(\mathcal{P}, \alpha') - L(\mathcal{P}, \alpha') < \varepsilon.\end{aligned}$$

Oznaczmy $M = \sup\limits_{a \leqslant x \leqslant b} |f(x)|$. Wtedy

$$\begin{aligned}\left|\sum_{i=1}^n f(s_i)\Delta\alpha_i - \sum_{i=1}^n f(s_i)\alpha'(s_i)\Delta x_i\right| &= \left|\sum_{i=1}^n f(s_i)\alpha'(t_i)\Delta x_i - \sum_{i=1}^n f(s_i)\alpha'(s_i)\Delta x_i\right| \\ &\leqslant \sum_{i=1}^n |f(s_i)|\,|\alpha'(t_i) - \alpha'(s_i)|\,\Delta x_i \leqslant M\sum_{i=1}^n |\alpha'(t_i) - \alpha'(s_i)|\Delta x_i \leqslant M\varepsilon.\end{aligned}$$

Zatem

$$\sum_{i=1}^{n} f(s_i)\Delta\alpha_i \leqslant \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i + M\varepsilon \leqslant U(\mathcal{P}, f\alpha') + M\varepsilon.$$

Obliczamy kres górny lewej strony względem punktów s_i z $[x_{i-1}, x_i]$ i otrzymujemy

$$U(\mathcal{P}, f, \alpha) \leqslant U(\mathcal{P}, f\alpha') + M\varepsilon.$$

Podobne rozumowanie daje nierówność

$$U(\mathcal{P}, f\alpha') \leqslant U(\mathcal{P}, f, \alpha) + M\varepsilon.$$

Dostajemy więc

$$|U(\mathcal{P}, f, \alpha) - U(\mathcal{P}, f\alpha')| \leqslant M\varepsilon. \tag{3.4}$$

Warunek (3.3) pozostaje spełniony, gdy podział $\mathcal{P}$ zastąpimy rozdrobnieniem, czyli podziałem $\mathcal{P}' \supset \mathcal{P}$. Zatem warunek (3.4) będzie spełniony dla $\mathcal{P}' \supset \mathcal{P}$. Zatem

$$\left| \int_a^{\bar{b}} f(x)\, d\alpha(x) - \int_a^{\bar{b}} f(x)\alpha'(x)\, dx \right| \leqslant M\varepsilon.$$

Ponieważ ε jest dowolną liczbą dodatnią, to

$$\int_a^{\bar{b}} f(x)\, d\alpha(x) = \int_a^{\bar{b}} f(x)\alpha'(x)\, dx.$$

Podobnie otrzymujemy

$$\int_{\underline{a}}^{b} f(x)\, d\alpha(x) = \int_{\underline{a}}^{b} f(x)\alpha'(x)\, dx.$$

□

Przykład. Rozważmy pręt o długości 1. Moment bezwładności pręta przy obrocie wokół osi prostopadłej do pręta, przechodzącej przez punkt początkowy pręta ($x = 0$) wynosi

$$\int_0^1 x^2 dm(x),$$

gdzie $m(x)$ oznacza masę fragmentu pręta odpowiadającego $[0, x]$, Jeśli $m'(x) = \varrho(x)$, to

$$\int_0^1 x^2 dm(x) = \int_0^1 x^2 \varrho(x)\, dx.$$

Jeśli całkowita masa pręta jest umieszczona punktach x_n , $n = 1, 2, \ldots$, to

$$\int_0^1 x^2\, dm(x) = \sum_{n=1}^{\infty} x_n^2 m_n,$$

gdzie m_n oznacza masę umieszczoną w punkcie x_n.

4 Funkcje wielu zmiennych

Będziemy rozważać funkcje określone na podzbiorze $A \subseteq \mathbb{R}^n$ o wartościach rzeczywistych. Większość teorii dotyczy $n = 2$ lub $n = 3$. Punkty w $\mathbb{R}^2$, $\mathbb{R}^3$ lub $\mathbb{R}^n$ będziemy oznaczać odpowiednio przez

$$(x, y), \qquad (x, y, z), \qquad x = (x_1, x_2, \ldots, x_n).$$

Przykłady.

$f(x, y) = xy$	pole prostokąta o bokach $x, y > 0$,
$f(x, y, z) = xyz$	objętość prostopadłościanu,
$f(x, y, z) = \dfrac{c}{\sqrt{x^2 + y^2 + z^2}}$,	potencjał grawitacyjny, $(x, y, z) \neq 0$.

W przestrzeni $\mathbb{R}^n$ rozważamy metrykę

$$d(x, y) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2} = \|x - y\|, \quad \text{gdzie} \quad \|x\| = \sqrt{\sum_{k=1}^{n} x_k^2}.$$

Twierdzenie 4.1. $\|x + y\| \leqslant \|x\| + \|y\|$.

Dowód. Możemy założyć, że $x, y \neq 0$, tzn. $\|x\|,\ \|y\| > 0$. Mamy

$$\|x + y\|^2 = \sum_{k=1}^{n} (x_k + y_k)^2 = \sum_{k=1}^{n} x_k^2 + \sum_{k=1}^{n} y_k^2 + 2 \sum_{k=1}^{n} x_k y_k,$$

$$(\|x\| + \|y\|)^2 = \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| = \sum_{k=1}^{n} x_k^2 + \sum_{k=1}^{n} y_k^2 + 2\|x\|\,\|y\|.$$

Wystarczy udowodnić, że

$$\sum_{k=1}^{n} x_k y_k \leqslant \|x\|\,\|y\| = \sqrt{\sum_{k=1}^{n} x_k^2}\sqrt{\sum_{k=1}^{n} y_k^2}. \tag{4.1}$$

Skorzystamy z nierówności [5]

$$2ab \leqslant \lambda a^2 + \frac{b^2}{\lambda} \qquad \lambda > 0.$$

Wtedy

$$2\sum_{k=1}^{n} x_k y_k \leqslant \lambda \sum_{k=1}^{n} x_k^2 + \frac{1}{\lambda}\sum_{k=1}^{n} y_k^2 = \lambda\|x\|^2 + \frac{1}{\lambda}\|y\|^2.$$

Przyjmijmy $\lambda = \dfrac{\|y\|}{\|x\|}$. Wtedy prawa strona ostatniej nierówności będzie równa $2\|x\|\,\|y\|$. □

Uwaga 4.2. Można sprawdzić, analizując dowód, kiedy w (4.1) otrzymujemy równość. Obie strony zerują się, gdy jeden z wektorów x lub y jest zerowy. Dla niezerowych wektorów dla pewnej dodatniej liczby λ mamy $\sqrt{\lambda}x_i = \dfrac{y_i}{\sqrt{\lambda}}$, tzn. wektory x i y są równoległe i mają ten sam zwrot.

Wniosek 4.3 (Nierówność trójkąta)**.**

$$d(x,z) \leqslant d(x,y) + d(y,z).$$

Dowód.

$$d(x,z) = \|x-z\| = \|(x-y)+(y-z)\| \leqslant \|x-y\| + \|y-z\| = d(x,y)+d(y,z).$$

□

Uwaga 4.4. Z wniosku wynika, że

$$|d(x,z) - d(y,z)| \leqslant d(x,y). \tag{4.2}$$

Rzeczywiście, mamy

$$d(x,z) \leqslant d(x,y) + d(y,z), \qquad d(y,z) \leqslant d(y,x) + d(x,z).$$

[5] Po przekształceniu $\left(\sqrt{\lambda}a - \dfrac{b}{\sqrt{\lambda}}\right)^2 \geqslant 0$

Zatem

$$d(x,z) - d(y,z) \leqslant d(x,y), \qquad d(y,z) - d(x,z) \leqslant d(x,y).$$

Stąd otrzymujemy (4.2).

Definicja 4.5. *Podzbiór $A \subseteq \mathbb{R}^2$ nazywamy otwartym, jeśli dla każdego punktu (x_0, y_0) w A można znaleźć liczbę $\delta > 0$ taką, że jeśli $d((x,y),(x_0,y_0)) < \delta$, to (x,y) leży w A. Warunek $d((x,y),(x_0,y_0)) < \delta$, oznacza, że*

$$(x - x_0)^2 + (y - y_0)^2 < \delta^2.$$

Czyli koło otwarte o środku w (x_0, y_0) i promieniu δ leży w A.

Przykłady. Zbiory $A = \{(x,y) \,:\, x^2 + y^2 < 1\}$, $B = \{(x,y) \,:\, x^2 + y^2 > 1\}$ są otwarte. Rzeczywiście, jeśli $x_0^2 + y_0^2 < 1$, to możemy przyjąć $\delta = 1 - \sqrt{x_0^2 + y_0^2}$. Dla $x_0^2 + y_0^2 > 1$ przyjmujemy $\delta = \sqrt{x_0^2 + y_0^2} - 1$.

4.1 Granica funkcji wielu zmiennych

Przypuśćmy, że funkcja $f(x,y)$ jest określona w kole otwartym o środku w (x_0, y_0), być może z wyłączeniem punktu (x_0, y_0). Mówimy, że funkcja $f(x,y)$ ma granicę L w punkcie (x_0, y_0) jeśli wartości $f(x,y)$ leżą blisko wartości L, gdy punkt (x,y) leży blisko (x_0, y_0), ale $(x,y) \neq (x_0, y_0)$. Tzn.

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x, y \; \{\, d((x,y),(x_0,y_0)) < \delta \implies |f(x,y) - L| < \varepsilon \,\}$$

Piszemy wtedy $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = L$.

Przykłady.

(a) $\lim\limits_{(x,y)\to(x_0,y_0)} x = x_0$, $\lim\limits_{(x,y)\to(x_0,y_0)} y = y_0$.

(b) $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2 - y^2}{x^2 + y^2}$. Oznaczmy $f(x,y) = \dfrac{x^2 - y^2}{x^2 + y^2}$. Wtedy $f(x,0) = 1$, oraz $f(0,y) = -1$. Zatem granica nie istnieje.

(c) Niech $g(x,y) = [f(x,y)]^2$, dla $f(x,y)$ z przykładu (b). Wtedy $g(x,0) = g(0,y) = 1$, ale $g(x,x) = 0$. Stąd granica $\lim\limits_{(x,y)\to(0,0)} g(x,y)$ nie istnieje.

(d) $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^3+y^3}{x^2+y^2}$. Mamy

$$|x^3+y^3| \leqslant |x|x^2 + |y|y^2 \leqslant (|x|+|y|)(x^2+y^2).$$

Zatem

$$\frac{|x^3+y^3|}{x^2+y^2} \leqslant |x|+|y| \xrightarrow[(x,y)\to(0,0)]{} 0.$$

Można też przeprowadzić rozumowanie z użyciem współrzędnych biegunowych. Dla $x = r\cos\theta$ i $y = r\sin\theta$ warunek $(x,y) \to (0,0)$ jest równoważny warunkowi $r \to 0^+$, bo $r = \sqrt{x^2+y^2}$. Wtedy

$$\frac{x^3+y^3}{x^2+y^2} = \frac{r^3(\cos^3\theta + \sin^3\theta)}{r^2} = r(\cos^3\theta + \sin^3\theta) \xrightarrow[r\to 0^+]{} 0,$$

bo $|\cos^3\theta + \sin^3\theta| \leqslant 2$.

Zadanie. Wskazać funkcję $f(x,y)$ taką, że $\lim\limits_{t\to 0^+} f(ta,tb) = 0$ dla dowolnego wektora $(a,b) \neq (0,0)$, ale granica funkcji $f(x,y)$ w punkcie $(0,0)$ nie istnieje.

Uwaga 4.6. Zapis stosowany w niektórych podręcznikach

$$\lim_{\substack{x\to x_0\\ y\to y_0}} f(x,y) = L$$

jest mylący, bo sugeruje, że $x \neq x_0$ i $y \neq y_0$. Przy obliczaniu granicy

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y)$$

wymagamy, aby $(x,y) \neq (x_0,y_0)$.

Działania arytmetyczne na granicach są spełnione tak jak dla funkcji jednej zmiennej. Na przykład poniżej korzystamy ze wszystkich takich działań.

$$\lim_{(x,y)\to(-1,3)} \frac{x^3+y^3}{x^2+y^2} = \frac{(-1)^3+3^3}{(-1)^2+3^2} = 2,6.$$

Prawdziwe jest też twierdzenie o podstawianiu.

Twierdzenie 4.7. *Jeśli* $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = L$ *oraz funkcja* $g(t)$ *jest ciągła w punkcie* L, *to*

$$\lim_{(x,y)\to(x_0,y_0)} g(f(x,y)) = g(L) = g\left(\lim_{(x,y)\to(x_0,y_0)} f(x,y)\right).$$

Przykład.

$$\lim_{(x,y)\to(e,1)} \log\frac{y}{x} \underset{t=\frac{y}{x}}{=} \lim_{t\to\frac{1}{e}} \log t = \log\frac{1}{e} = -1.$$

Definicja 4.8. *Mówimy, że funkcja* $f(x,y)$ *jest ciągła w* (x_0,y_0), *jeśli jest określona w pewnym kole wokół* (x_0,y_0) *oraz* $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0,y_0)$.

Przykład. Funkcja $f(x,y) = \sin\dfrac{xy}{1+x^2+y^2}$ jest ciągła w każdym punkcie.

Dla zbioru $A \subseteq \mathbb{R}^2$ i punktu p zbioru A mogą zdarzyć się dwa przypadki.

(1) p leży w A z pewnym kołem wokół siebie. Tzn. p należy do wnętrza zbioru A.

(2) Każde koło o środku w p zawiera punkty ze zbioru A i spoza zbioru A. Tzn. p leży na brzegu zbioru A.

Wnętrze i brzeg zbioru A oznaczamy symbolami $\operatorname{int} A$ oraz $\operatorname{bd} A$, odpowiednio.

Przykład. $A = \{(x,y) : |x|+|y| \leqslant 1\}$. Wtedy $\operatorname{int} A = \{(x,y) : |x|+|y| < 1\}$ oraz $\operatorname{bd} A = \{(x,y) : |x|+|y| = 1\}$.

Mówimy, że funkcja $f(x,y)$ jest ciągła na zbiorze A jeśli f jest określona na A, ciągła w każdym punkcie wewnętrznym oraz dla punktów brzegowych (x_0,y_0) spełnia

$$\lim_{\substack{(x,y)\to(x_0,y_0)\\(x,y)\in A}} f(x,y) = f(x_0,y_0).$$

Przykłady.

1. $A = [0,1] \times [0,2]$ oraz
$$f(x,y) = \begin{cases} 4 - x - y & (x,y) \in A, \\ 0 & (x,y) \notin A. \end{cases}$$
Funkcja f jest ciągła na A, tzn. gdy rozważamy ją tylko na zbiorze A. Ale f nie jest ciągła na $\mathbb{R}^2$.

2. $B = \{(x,y) : x^2 + y^2 \leqslant 1\}$ oraz
$$f(x,y) = \begin{cases} 1 - x^2 - y^2 & (x,y) \in B, \\ 0 & (x,y) \notin B. \end{cases}$$
Funkcja f jest tym razem ciągła na $\mathbb{R}^2$. Można potraktować f jako złożenie $g(h(x,y))$, gdzie $h(x,y) = 1 - x^2 - y^2$ oraz $g(t) = \max(t,0)$. Obie funkcje są ciągłe. Z twierdzenia o podstawianiu wynika, że ich złożenie jest funkcją ciągłą.

5 Pochodne cząstkowe

Definicja 5.1. *Załóżmy, że funkcja $f(x,y)$ jest określona w otoczeniu punktu (x_0, y_0). Pochodną cząstkową względem x funkcji f w punkcie (x_0, y_0) określamy wzorem*
$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}.$$
Podobnie określamy pochodną cząstkową względem y
$$\frac{\partial f}{\partial y}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h}.$$

Aby obliczyć $\frac{\partial f}{\partial x}(x_0, y_0)$ i $\frac{\partial f}{\partial y}(x_0, y_0)$ wystarczy znać wartości $f(x,y)$ na fragmentach dwu prostych przechodzących przez punkt (x_0, y_0).

Uwaga 5.2.
$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{d}{dx} f(x, y_0)\Big|_{x = x_0}, \qquad \frac{\partial f}{\partial y}(x_0, y_0) = \frac{d}{dy} f(x_0, y)\Big|_{y = y_0}.$$

Przykład. $f(x,y) = 24xy - 5x^2y$. Chcemy obliczyć obie pochodne cząstkowe w punkcie $(1,2)$. Możemy to zrobić na dwa sposoby.

(a) Obliczamy $\dfrac{\partial f}{\partial x}$ i $\dfrac{\partial f}{\partial y}$ w dowolnym punkcie i po wykonaniu obliczeń podstawiamy $(1,2)$. Mamy

$$\frac{\partial f}{\partial x} = 24y - 10xy, \qquad \frac{\partial f}{\partial y} = 24x - 5x^2.$$

Zatem

$$\frac{\partial f}{\partial x}(1,2) = 24 \cdot 2 - 10 \cdot 2 = 28, \qquad \frac{\partial f}{\partial y}(1,2) = 24 \cdot 1 - 5 \cdot 1^2 = 19.$$

(b) Obliczamy $f(x,2)$ oraz $f(1,y)$.

$$f(x,2) = 48x - 10x^2, \qquad f(1,y) = 24y - 5y = 19y.$$

Dalej

$$\begin{aligned} \frac{\partial f}{\partial x}(1,2) &= \frac{d}{dx} f(x,2)\Big|_{x=1} = \frac{d}{dx}(48x - 10x^2)\Big|_{x=1} = (48 - 20x)\Big|_{x=1} = 28, \\ \frac{\partial f}{\partial y}(1,2) &= \frac{d}{dy} f(1,y)\Big|_{y=2} = \frac{d}{dy}(19y)\Big|_{y=2} = 19. \end{aligned}$$

Przykład.

$$f(x,y) = \begin{cases} \dfrac{x^3y - xy^3}{x^2+y^2} & (x,y) \neq (0,0), \\ 0 & x = y = 0. \end{cases} \tag{5.1}$$

Dla $(x,y) \neq (0,0)$ obliczamy $\dfrac{\partial f}{\partial x}$.

$$\frac{\partial f}{\partial x} = \frac{(3x^2y - y^3)(x^2+y^2) - (x^3y - xy^3)2x}{(x^2+y^2)^2} = \frac{x^4y + 4x^2y^3 - y^5}{(x^2+y^2)^2}. \tag{5.2}$$

Ze wzoru $f(y,x) = -f(x,y)$ otrzymujemy

$$\frac{\partial f}{\partial y} = -\frac{xy^4 + 4x^3y^2 - x^5}{(x^2+y^2)^2}. \tag{5.3}$$

Ponieważ wartość $f(0,0)$ jest określona osobno pochodne cząstkowe w punkcie $(0,0)$ musimy obliczyć inaczej. Mamy $f(x,0) = 0$ oraz $f(0,y) = 0$. Zatem

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0. \tag{5.4}$$

Twierdzenie 5.3. *Załóżmy, że funkcja $f(x,y)$ ma pochodną cząstkową względem x w prostokącie $(a,b)\times(c,d)$. Wtedy dla punktów (x_1,y) i (x_2,y) z tego prostokąta mamy*

$$f(x_2,y)-f(x_1,y)=\frac{\partial f}{\partial x}(\zeta,y)\,(x_2-x_1),$$

dla pewnej liczby ζ leżącej pomiędzy x_1 i x_2.

Podobnie jeśli funkcja $f(x,y)$ ma pochodną cząstkową względem y, to

$$f(x,y_2)-f(x,y_1)=\frac{\partial f}{\partial y}(x,\eta)\,(y_2-y_1),$$

dla pewnej liczby η leżącej pomiędzy y_1 i y_2.

Dowód. Rozważamy funkcję jednej zmiennej $x\mapsto f(x,y)$ na przedziale (a,b). Z twierdzenia Lagrange'a otrzymujemy

$$f(x_2,y)-f(x_1,y)=\frac{d}{dx}f(x,y)\Big|_{x=\zeta}(x_2-x_1)=\frac{\partial f}{\partial x}(\zeta,y)\,(x_2-x_1).$$

$\square$

5.1 Wyższe pochodne cząstkowe

Dla funkcji $f(x,y)$ pochodne cząstkowe $\dfrac{\partial f}{\partial x}$ i $\dfrac{\partial f}{\partial y}$ są znowu funkcjami dwu zmiennych. Możemy więc obliczać pochodne cząstkowe tych funkcji. Następne pochodne cząstkowe oznaczamy symbolami

$$\frac{\partial^2 f}{\partial x^2},\quad \frac{\partial^2 f}{\partial y\partial x},\quad \frac{\partial^2 f}{\partial x\partial y},\quad \frac{\partial^2 f}{\partial y^2},$$

przy czym w pochodnych mieszanych wykonujemy różniczkowanie w kolejności od prawej do lewej strony

Przykłady.

1. $f(x,y)=\sin(xy^2)$.

$$\frac{\partial f}{\partial x}=y^2\cos(xy^2),\qquad \frac{\partial f}{\partial y}=2xy\cos(xy^2),$$

$$\frac{\partial^2 f}{\partial x^2} = -y^4 \sin(xy^2), \qquad \frac{\partial^2 f}{\partial y \partial x} = 2y\cos(xy^2) - 2xy^3 \sin(xy^2),$$
$$\frac{\partial^2 f}{\partial x \partial y} = 2y\cos(xy^2) - 2xy^3 \sin(xy^2), \qquad \frac{\partial^2 f}{\partial y^2} = 2x\cos(xy^2) - 4x^2y^2 \sin(xy^2).$$

Zauważmy, że pochodne mieszane są w tym przypadku równe.

2. Rozważmy ponownie funkcję $f(x, y)$ z (5.1). Obliczymy pochodne mieszane w $(0, 0)$. Ze wzorów (5.2), (5.3) i (5.4) mamy

$$\frac{\partial f}{\partial x}(0, y) = -y, \qquad \frac{\partial f}{\partial y}(x, 0) = x.$$

Zatem
$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1, \qquad \frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1.$$

Uwaga 5.4. Z przykładu wynika, że pochodne mieszane nie muszą być sobie równe. Niedługo udowodnimy, że jeśli pochodne te są ciągłe w danym punkcie, to są w tym punkcie równe sobie.

5.2 Reguła łańcucha

Dla funkcji jednej zmiennej jeśli $y = g(u)$ oraz $u = f(x)$, to

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}.$$

Dla funkcji wielu zmiennych jest wiele możliwości złożenia funkcji.

(a) Niech $z = f(x, y)$ oraz $x = g_1(t)$, $y = g_2(t)$. Otrzymujemy $z = f(g_1(t), g_2(t))$. Wtedy

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dt},$$

gdzie $\frac{\partial z}{\partial x}$ i $\frac{\partial z}{\partial y}$ są obliczane w x i y a $\frac{dx}{dt}$, $\frac{dy}{dt}$ i $\frac{dz}{dt}$ są obliczane w t. Po wykonaniu obliczeń trzeba podstawić $x = g_1(t)$ oraz $y = g_2(t)$, tzn. wynik ma być zapisany w języku zmiennej t.

(b) $z = f(x,y)$ oraz $x = g_1(u,v)$, $y = g_2(u,v)$. Tzn. $z = f(g_1(u,v), g_2(u,v))$. Wtedy

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial u},$$
$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial v}.$$

Pochodne $\frac{\partial z}{\partial x}$ i $\frac{\partial z}{\partial y}$ są obliczane w (x,y) a pozostałe w (u,v). Wynik ma być zapisany w języku zmiennych u i v.

Przykład. $z = (u^2 - v^2)\log(u^2+v^2)$. Przyjmijmy $x = u^2 - v^2$, $y = u^2 + v^2$. Wtedy $z = x \log y$.

$$\begin{aligned}\frac{\partial z}{\partial u} &= \log y \cdot 2u + \frac{x}{y} \cdot 2u = 2u\log(u^2+v^2) + 2u\frac{u^2-v^2}{u^2+v^2},\\ \frac{\partial z}{\partial v} &= \log y \cdot (-2v) + \frac{x}{y} \cdot 2v = -2v\log(u^2+v^2) + 2v\frac{u^2-v^2}{u^2+v^2}.\end{aligned}$$

Dowód $\left(\text{reguły (b) dla } \frac{\partial z}{\partial u}\right)$. Zakładamy, że wszystkie pochodne cząstkowe po prawej stronie wzoru istnieją i że funkcje $\frac{\partial f}{\partial x}$ oraz $\frac{\partial f}{\partial y}$ są ciągłe. Mamy $z(u,v) = f(g_1(u,v), g_2(u,v))$. Przyjmujemy oznaczenia

$$\begin{array}{ll} x = g_1(u,v) & y = g_2(u,v), \\ s = g_1(u+h,v) - g_1(u,v), & t = g_2(u+h,v) - g_2(u,v). \end{array}$$

Wielkości s i t zależą od h.

$$\frac{\partial z}{\partial u}(u,v) = \lim_{h\to 0} \frac{f(g_1(u+h,v), g_2(u+h,v)) - f(g_1(u,v), g_2(u,v))}{h}$$
$$\lim_{h\to 0} \frac{f(x+s, y+t) - f(x, y+t) + f(x, y+t) - f(x,y)}{h}$$

Z Twierdzenia 5.3 mamy

$$\begin{aligned} f(x+s,y+t) - f(x,y+t) &= \frac{\partial f}{\partial x}(x+\theta_1 s, y+t)\, s,\\ f(x,y+t) - f(x,y) &= \frac{\partial f}{\partial y}(x, y+\theta_2 t)\, t,\end{aligned}$$

dla pewnych $0 < \theta_1, \theta_2 < 1$. Gdy $h \to 0$, to $s \to 0$ oraz $t \to 0$. Zatem

$$\frac{\partial f}{\partial x}(x+\theta_1 s, y+t) \underset{h\to 0}{\longrightarrow} \frac{\partial f}{\partial x}(x,y) = \frac{\partial z}{\partial x},$$
$$\frac{\partial f}{\partial y}(x, y+\theta_2 t) \underset{h\to 0}{\longrightarrow} \frac{\partial f}{\partial y}(x,y) = \frac{\partial z}{\partial y}.$$

Dalej

$$\frac{s}{h} = \frac{g_1(u+h,v) - g_1(u,v)}{h} \underset{h\to 0}{\longrightarrow} \frac{\partial g_1}{\partial u}(u,v) = \frac{\partial x}{\partial u},$$
$$\frac{t}{h} = \frac{g_2(u+h,v) - g_2(u,v)}{h} \underset{h\to 0}{\longrightarrow} \frac{\partial g_2}{\partial u}(u,v) = \frac{\partial y}{\partial u}.$$

Reasumując, w granicy otrzymamy $\dfrac{\partial z}{\partial x} \cdot \dfrac{\partial x}{\partial u} + \dfrac{\partial z}{\partial y} \cdot \dfrac{\partial y}{\partial u}$. □

Przykład. $w = \cos(xyz^2)$ oraz $x = \sin t$, $y = t^2$, $z = e^t$.

$$\begin{aligned}\frac{dw}{dt} &= \frac{\partial w}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial w}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial w}{\partial z} \cdot \frac{dz}{dt} \\ &= -\sin(xyz^2) yz^2 \cos t - \sin(xyz^2) xz^2\, 2t - \sin(xyz^2) 2xyz\, e^t \\ &= -\sin(t^2 e^{2t} \sin t)\,[t^2 e^{2t} \cos t + 2te^{2t} \sin t + 2t^2 e^{2t} \sin t].\end{aligned}$$

Przykład. Góra piasku w kształcie stożka rośnie w tempie 4 litry na sekundę, a promień podstawy r rośnie w tempie e^{-r} decymetrów na sekundę. W jakim tempie rośnie wysokość w momencie, gdy $V = 60\,\mathrm{l}$ oraz $r = 6\,\mathrm{dcm}$?

Mamy $V = \frac{1}{3}\pi r^2 h$. Zatem

$$h = \frac{3}{\pi}\frac{V}{r^2}.$$

Wielkości h, V i r są funkcjami czasu t. Otrzymujemy

$$\frac{dh}{dt} = \frac{\partial h}{\partial V}\frac{dV}{dt} + \frac{\partial h}{\partial r}\frac{dr}{dt} = \frac{3}{\pi}\frac{1}{r^2} \cdot 4 - \frac{3}{\pi}\frac{2V}{r^3} \cdot e^{-r}.$$

Niech t_0 oznacza moment czasu, gdy $V = 60\,\mathrm{l}$ oraz $r = 6\,\mathrm{dcm}$. Wtedy

$$\left.\frac{dh}{dt}\right|_{t=t_0} = \frac{3}{\pi}\frac{1}{6^2} \cdot 4 - \frac{3}{\pi}\frac{2\cdot 60}{6^3} \cdot e^{-6} \text{ (dcm/sek)}.$$

Ile wynosi r jeśli $\dfrac{dr}{dt} = e^{-r}$?

5.3 Różniczkowalność funkcji wielu zmiennych

Definicja 5.5. *Mówimy, że funkcja $f(x,y)$ jest różniczkowalna w punkcie (x_0,y_0), jeśli istnieją pochodne cząstkowe $a = \dfrac{\partial f}{\partial x}(x_0,y_0)$, $b = \dfrac{\partial f}{\partial y}(x_0,y_0)$ oraz*

$$\lim_{(h_1,h_2)\to(0,0)} \frac{|f(x_0+h_1,y_0+h_2) - f(x_0,y_0) - ah_1 - bh_2|}{\sqrt{h_1^2+h_2^2}} = 0.$$

Uwagi 5.6.

(a) Przyrost argumentu od (x_0,y_0) do (x_0+h_1,y_0+h_2) wynosi (h_1,h_2). Wielkość $\sqrt{h_1^2+h_2^2}$ jest więc długością tego przyrostu.

(b) Dla funkcji jednej zmiennej różniczkowalność oznacza, że

$$\frac{f(x_0+h)-f(x_0)}{h} \xrightarrow[h\to 0]{} f'(x_0) =: a.$$

Tzn.

$$\lim_{h\to 0} \frac{|f(x_0+h)-f(x_0)-ah|}{|h|} = 0.$$

Twierdzenie 5.7. *Załóżmy, że dla funkcji $f(x,y)$ pochodne cząstkowe $\dfrac{\partial f}{\partial x}$ i $\dfrac{\partial f}{\partial y}$ są ciągłe w punkcie (x_0,y_0). Wtedy funkcja $f(x,y)$ jest różniczkowalna w punkcie (x_0,y_0).*

Dowód.

$$\begin{aligned}
&f(x_0+h_1,y_0+h_2) - f(x_0,y_0) - ah_1 - bh_2\\
&= f(x_0+h_1,y_0) - f(x_0,y_0) - ah_1 + f(x_0+h_1,y_0+h_2) - f(x_0+h_1,y_0) - bh_2\\
&= \frac{\partial f}{\partial x}(x_0+\theta_1 h_1,y_0)h_1 - \frac{\partial f}{\partial x}(x_0,y_0)h_1 + \frac{\partial f}{\partial y}(x_0+h_1,y_0+\theta_2 h_2)h_2 - \frac{\partial f}{\partial y}(x_0,y_0)h_2\\
&= \underbrace{\left[\frac{\partial f}{\partial x}(x_0+\theta_1 h_1,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)\right]}_{A} h_1 + \underbrace{\left[\frac{\partial f}{\partial y}(x_0+h_1,y_0+\theta_2 h_2) - \frac{\partial f}{\partial y}(x_0,y_0)\right]}_{B} h_2
\end{aligned}$$

Mamy

$$\begin{aligned}|Ah_1+Bh_2| \leqslant |A||h_1|+|B||h_2| \leqslant |A|\sqrt{h_1^2+h_2^2}+|B|\sqrt{h_1^2+h_2^2}\\ =(|A|+|B|)\sqrt{h_1^2+h_2^2}.\end{aligned}$$

Zatem

$$\frac{|f(x_0+h_1,y_0+h_2)-f(x_0,y_0)-ah_1-bh_2|}{\sqrt{h_1^2+h_2^2}} \leqslant |A|+|B| \xrightarrow[(h_1,h_2)\to(0,0)]{} 0.$$

□

Przykład. $f(x,y)=\sin(xy)$. Mamy

$$\frac{\partial f}{\partial x}=y\cos(xy), \qquad \frac{\partial f}{\partial y}=x\cos(xy).$$

Pochodne cząstkowe są ciągłe. Zatem $f(x,y)$ jest różniczkowalna w każdym punkcie. W punkcie $(0,0)$ pochodne cząstkowe zerują się. Różniczkowalność oznacza więc, że

$$\frac{|\sin(xy)|}{\sqrt{x^2+y^2}} \xrightarrow[(x,y)\to(0,0)]{} 0.$$

5.3.1 Interpretacja geometryczna różniczkowalności

Wykres funkcji $z=f(x,y)$ jest podzbiorem przestrzeni $\mathbb{R}^3$. Rozważamy obraz prostej $y=y_0$ przez funkcję $f(x,y)$, czyli krzywą $(x,y_0)\mapsto f(x,y_0)$. Ta krzywa znajduje się w płaszczyźnie pionowej $y=y_0$. Chcemy znaleźć styczną do tej krzywej. Gdyby funkcja f zależała tylko od zmiennej x, to styczna do krzywej miałaby równanie $z-z_0=f'(x_0)(x-x_0)$. Ale rozważamy $(x,y_0)\mapsto f(x,y_0)$. Zatem równanie stycznej ma postać

$$\begin{cases} z-z_0=\dfrac{\partial f}{\partial x}(x_0,y_0)\,(x-x_0)\\ y=y_0\end{cases}$$

Przy oznaczeniu $a=\dfrac{\partial f}{\partial x}(x_0,y_0)$ równanie stycznej, to

$$\begin{cases} z-z_0=a(x-x_0)\\ y=y_0\end{cases}$$

Podobnie, rozważamy obraz prostej $x = x_0$ czyli funkcję $(x_0, y) \mapsto f(x_0, y)$. Równanie stycznej ma postać

$$\begin{cases} z - z_0 = b(y - y_0) \\ x = x_0 \end{cases}$$

gdzie $b = \dfrac{\partial f}{\partial y}(x_0, y_0)$. Te dwie styczne rozpinają płaszczyznę o równaniu

$$z - z_0 = a(x - x_0) + b(y - y_0).$$

Niech (x, y, z) oznacza punkt na płaszczyźnie odpowiadający punktowi (x, y), w zamyśle położonym blisko punktu (x_0, y_0). Zatem

$$z = z_0 + a(x - x_0) + b(y - y_0) = f(x_0, y_0) + ah_1 + bh_2,$$

przy oznaczeniach $h_1 = x - x_0$ i $h_2 = y - y_0$. Wtedy

$$|f(x, y) - z| = |f(x_0 + h_1, y_0 + h_2) - f(x_0, y_0) - ah_1 - bh_2|.$$

Reasumując, różniczkowalność oznacza, że iloraz odległości punktu wykresu $(x, y, f(x, y))$ i odpowiadającego punktu (x, y, z) na płaszczyźnie, przez odległość pomiędzy (x_0, y_0) i (x, y), jest mały, gdy $(x, y) \to (x_0, y_0)$. W takim wypadku mówimy, że płaszczyzna $z - z_0 = a(x - x_0) + b(y - y_0)$ jest styczna do wykresu w punkcie (x_0, y_0, z_0), gdzie $z_0 = f(x_0, y_0)$.

Podobnie określamy różniczkowalność funkcji wielu zmiennych. Funkcja n zmiennych $f(x_1, \ldots, x_n)$ jest różniczkowalna w punkcie $(x_1, \ldots, x_n)$, jeśli

$$\lim_{(h_1,\ldots,h_n)\to 0} \frac{|f(x_1 + h_1, \ldots, x_n + h_n) - f(x_1, \ldots, x_n) - a_1h_1 - \ldots - a_nh_n|}{\sqrt{h_1^2 + \ldots + h_n^2}} = 0,$$

gdzie

$$a_1 = \frac{\partial f}{\partial x_1}(x_1, \ldots, x_n), \ \ldots, \ a_n = \frac{\partial f}{\partial x_n}(x_1, \ldots, x_n).$$

Tzn.

$$a_i = \lim_{t\to 0} \frac{f(x_1, \ldots, x_i + t, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{t}.$$

Wygodnie będzie zastosować zapis wektorowy

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}, \quad A = (a_1, \ldots, a_n).$$

Wtedy
$$a_j = \lim_{t\to 0} \frac{f(x+te_j) - f(x)}{t} = \frac{\partial f}{\partial x_j}(x),$$
gdzie e_j jest j-tym wektorem bazowym, tzn.
$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \longleftarrow j\text{–te miejsce}$$
Warunek różniczkowalności ma postać
$$\lim_{h\to 0} \frac{|f(x+h) - f(x) - Ah|}{\|h\|} = 0, \text{ gdzie } \|h\| = \sqrt{h_1^2 + \ldots + h_n^2}.$$

Rozważmy m funkcji $f_1, \ldots, f_m$, każda zależna od n zmiennych $x_1, \ldots, x_n$. Możemy utworzyć jedną funkcję $f(x)$, ale o wartościach wektorowych wzorem
$$\mathbb{R}^n \ni x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \longmapsto f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} \in \mathbb{R}^m.$$
Odwrotnie, każda funkcja $f : \mathbb{R}^n \to \mathbb{R}^m$ składa się z rodziny m funkcji o wartościach rzeczywistych.

Przykład. Rozważmy odwzorowanie
$$\mathbb{R}^2 \ni \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \begin{pmatrix} x_1^2 + x_2^2 \\ x_1 x_2 \\ x_1 \sin x_2 \end{pmatrix} \in \mathbb{R}^3.$$
Wtedy $f_1(x_1, x_2) = x_1^2 + x_2^2$.

Mówimy, że funkcja $f : \mathbb{R}^n \to \mathbb{R}^m$ jest różniczkowalna w punkcie x, jeśli każda z funkcji $f_1, \ldots, f_m$ jest różniczkowalna w x. Tzn. dla $i = 1, 2, \ldots, m$ mamy
$$\lim_{h\to 0} \frac{|f_i(x+h) - f_i(x) - A_i h|}{\|h\|} = 0,$$
gdzie
$$A_i = (a_{i1}, \ldots, a_{in}), \quad a_{ij} = \frac{\partial f_i}{dx_j}(x).$$
Czy można to objąć jednym zapisem dla funkcji $f(x)$?

Definicja 5.8. *Mówimy, że funkcja $f : \mathbb{R}^n \to \mathbb{R}^m$ jest różniczkowalna w punkcie x, jeśli*

$$\lim_{h\to 0} \frac{\|f(x+h) - f(x) - Ah\|}{\|h\|} = 0, \quad \text{gdzie } A = (a_{ij}).$$

Sprawdzimy, że faktycznie warunek w definicji oznacza, że każda z funkcji f_i jest różniczkowalna. Skorzystamy z nierówności

$$|c_i| \leqslant \left(\sum_{j=1}^m c_j^2\right)^{1/2} \leqslant \sum_{j=1}^m |c_j|, \quad i = 1, 2, \ldots, m.$$

Otrzymujemy

$$|f_i(x+h)-f_i(x)-A_ih| \leqslant \|f(x+h)-f(x)-Ah\| \leqslant \sum_{j=1}^m |f_j(x+h)-f_j(x)-A_jh|.$$

Z pierwszej nierówności wynika, że jeśli f jest różniczkowalna w punkcie x według Definicji 5.8, to każda z funkcji f_i jest różniczkowalna w x. Z kolei z drugiej nierówności wynika, że jeśli każda z funkcji f_i jest różniczkowalna w x, to f jest różniczkowalna w punkcie x według Definicji 5.8.

Stosujemy oznaczenie $A = Df(x)$, tzn. $Df(x)$ jest macierzą wymiaru $m \times n$ złożoną z pochodnych cząstkowych. Numer wiersza odpowiada numerowi funkcji składowej, natomiast numer kolumny odpowiada numerowi zmiennej, względem której obliczana jest pochodna cząstkowa.

Przykład.

$$\mathbb{R}^2 \ni \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \begin{pmatrix} x_1 - x_2 \\ x_1x_2 \\ x_1^2 + x_2^2 \end{pmatrix} \in \mathbb{R}^3.$$

Wtedy

$$Df(x) = \begin{pmatrix} 1 & -1 \\ x_2 & x_1 \\ 2x_1 & 2x_2 \end{pmatrix}, \qquad Df(1,2) = \begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 2 & 4 \end{pmatrix}.$$

Lemat 5.9. *Załóżmy, że dla pewnej macierzy B wymiaru $m \times n$ mamy*

$$\lim_{h\to 0} \frac{\|f(x+h) - f(x) - Bh\|}{\|h\|} = 0.$$

Wtedy $B = Df(x)$.

Dowód. Mamy

$$|f_i(x+h) - f_i(x) - B_i h| \leqslant \|f(x+h) - f(x) - Bh\|,$$

gdzie B_i oznacza i-ty wiersz macierzy B. Zatem

$$\lim_{h\to 0} \frac{|f_i(x+h) - f_i(x) - B_i h|}{\|h\|} = 0.$$

Niech $h = te_j$. Wtedy

$$0 = \lim_{t\to 0} \frac{|f_i(x+te_j) - f_i(x) - tb_{ij}|}{|t|} = \lim_{t\to 0} \left| \frac{f_i(x+te_j) - f_i(x)}{t} - b_{ij} \right|.$$

Stąd otrzymujemy $b_{ij} = \dfrac{\partial f_i}{\partial x_j}(x)$. □

Lemat 5.10 (nierówność Schwarza)**.**

$$(a_1b_1 + a_2b_2 + \ldots + a_nb_n)^2 \leqslant (a_1^2 + a_2^2 + \ldots + a_n^2)(b_1^2 + b_2^2 + \ldots + b_n^2).$$

Równość zachodzi tylko wtedy, gdy wektory $a = (a_1, a_2, \ldots, a_n)$ i $b = (b_1, b_2, \ldots, b_n)$ są równoległe.

Dowód. [6]

$$\begin{aligned}
(a_1^2 + a_2^2 + \ldots + a_n^2)&(b_1^2 + b_2^2 + \ldots + b_n^2) - (a_1b_1 + a_2b_2 + \ldots + a_nb_n)^2 \\
&= \sum_{i\neq j} a_i^2 b_j^2 - 2\sum_{i<j} a_ib_ia_jb_j = \sum_{i<j} (a_i^2b_j^2 + a_j^2b_i^2) - 2\sum_{i<j} a_ib_ia_jb_j \\
&= \sum_{i<j} (a_ib_j - a_jb_i)^2 \geqslant 0.
\end{aligned}$$

Przypuśćmy, że w nierówności Schwarza mamy równość oraz, że $a_{i_0} \neq 0$. Wtedy $a_{i_0}b_j = a_jb_{i_0}$, czyli wektor b jest wielokrotnością wektora a ze współczynnikiem b_{i_0}/a_{i_0}. □

Lemat 5.11. *Dla macierzy A wymiaru $m \times n$ oraz wektora $h \in \mathbb{R}^n$ mamy*

$$\|Ah\| \leqslant \|A\|_{HS}\|h\|, \quad \text{gdzie } \|A\|_{HS} = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2}.$$

Wielkość $\|A\|_{HS}$ nazywamy normą Hilberta-Schmidta macierzy A.

[6] Inny dowód został podany na stronie 51

Dowód. Oznaczmy
$$Ah = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Wtedy $\|Ah\|^2 = \sum_{i=1}^{m} y_i^2$. Ale $y_i = \sum_{j=1}^{n} a_{ij}h_j$. Z nierówności Schwarza mamy
$$y_i^2 = \left(\sum_{j=1}^{n} a_{ij}h_j\right)^2 \leqslant \sum_{j=1}^{n} a_{ij}^2 \sum_{j=1}^{n} h_j^2 = \sum_{j=1}^{n} a_{ij}^2 \, \|h\|^2,$$

czyli
$$\|Ah\|^2 = \sum_{i=1}^{m} y_i^2 \leqslant \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2 \, \|h\|^2 = \|A\|_{HS}^2 \, \|h\|^2.$$

$\square$

Twierdzenie 5.12. *Jeśli funkcja $f(x)$ jest różniczkowalna w punkcie $a \in \mathbb{R}^n$, to $f(x)$ jest ciągła w a.*

Dowód. Trzeba udowodnić, że $\lim_{h\to 0} \|f(a+h) - f(a)\| = 0$. Oznaczmy
$$u(h) = f(a+h) - f(a) - Ah, \quad \text{gdzie } A = Df(a).$$

Dla $h \neq 0$ mamy
$$\begin{aligned} \|f(a+h) - f(a)\| &= \|u(h) + Ah\| \leqslant \|u(h)\| + \|Ah\| \\ &\leqslant \|u(h)\| + \|A\|_{HS} \, \|h\| = \left[\frac{\|u(h)\|}{\|h\|} + \|A\|_{HS}\right] \|h\| \end{aligned}$$

Z różniczkowalności pierwszy składnik w nawiasie kwadratowym dąży do zera, gdy $h \to 0$. Zatem całe wyrażenie dąży do zera, gdy $h \to 0$. $\square$

Twierdzenie 5.13. *Załóżmy, że funkcja $f : \mathbb{R}^n \to \mathbb{R}^m$ jest różniczkowalna w punkcie $a \in \mathbb{R}^n$ natomiast funkcja $g : \mathbb{R}^m \to \mathbb{R}^p$ jest różniczkowalna w punkcie $b = f(a)$. Wtedy funkcja złożona $g \circ f : \mathbb{R}^n \to \mathbb{R}^p$ jest różniczkowalna w punkcie a oraz*
$$D(g \circ f)(a) = Dg(b)\, Df(a).$$

Uwagi.

(a) Macierze $Dg(b)$ i $Df(a)$ mają wymiary $p \times m$ i $m \times n$ odpowiednio. Po pomnożeniu otrzymamy macierz wymiaru $p \times n$.

(b) Wzór na pochodną funkcji złożonej wielu zmiennych zgadza się ze wzorem dla jednej zmiennej:

$$(g \circ f)'(a) = g'(b)\, f'(a), \quad b = f(a).$$

(c) Można nieformalnie wyjaśnić wzór w twierdzeniu. Oznaczmy $A = Df(a)$, $B = Dg(b)$. Różniczkowalność oznacza, że

$$\begin{aligned} f(a+h) &\approx f(a) + Ah, \quad \text{gdy } h \to 0, \\ g(b+k) &\approx g(b) + Bk, \quad \text{gdy } k \to 0. \end{aligned}$$

Zatem

$$\begin{aligned} g(f(a+h)) &\approx g(f(a) + Ah) = g(b + Ah) \\ &\approx g(b) + BAh = g(f(a)) + BAh, \quad \text{gdy } h \to 0. \end{aligned}$$

Stąd $D(g \circ f)(a) = BA$.

Dowód. Posłużymy się oznaczeniami z Uwagi (c). Trzeba udowodnić, że

$$\lim_{h \to 0} \frac{\|g \circ f(a+h) - g \circ f(a) - BAh\|}{\|h\|} = 0.$$

Wtedy teza wynika z Lematu 5.9. Oznaczmy

$$u(h) = f(a+h) - f(a) - Ah, \qquad v(k) = g(b+k) - g(b) - Bk.$$

Przyjmijmy $k = f(a+h) - f(a) = u(h) + Ah$. Wtedy

$$\begin{aligned} g \circ f(a+h) - g \circ f(a) - BAh &= g(f(a+h)) - g(f(a)) - BAh \\ &= g(b+k) - g(b) - BAh = v(k) + Bk - BAh \\ &= v(k) + B(k - Ah) = v(k) + Bu(h). \end{aligned}$$

Wprowadzamy oznaczenie

$$\varphi(k) = \begin{cases} \dfrac{\|v(k)\|}{\|k\|} & k \neq 0, \\ 0 & k = 0 \end{cases}.$$

Wtedy

$$\frac{\|g\circ f(a+h)-g\circ f(a)-BAh\|}{\|h\|} \leqslant \frac{\|v(k)\|}{\|h\|}+\frac{\|Bu(h)\|}{\|h\|} \leqslant \varphi(k)\frac{\|k\|}{\|h\|}+\|B\|_{HS}\frac{\|u(h)\|}{\|h\|}$$

Z różniczkowalności funkcji f drugi składnik dąży do zera, gdy $h \to 0$. Ponadto, gdy $h \to 0$, to również $k \to 0$. Zatem $\varphi(k) \to 0$. Dalej

$$\frac{\|k\|}{\|h\|} \leqslant \frac{\|u(h)\|}{\|h\|}+\frac{\|Ah\|}{\|h\|} \leqslant \frac{\|u(h)\|}{\|h\|}+\|A\|_{HS} \underset{h\to 0}{\longrightarrow} \|A\|_{HS}.$$

□

Przykład.

$$\begin{aligned} f(x,y) &= (x^2+1, y^2), & a &= (2,1),\\ g(u,v) &= (u+v, u, v^2), & b &= f(2,1) = (5,1). \end{aligned}$$

Mamy

$$Df(x,y) = \begin{bmatrix} 2x & 0 \\ 0 & 2y \end{bmatrix}, \qquad Dg(u,v) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2v \end{bmatrix}.$$

Zatem

$$Df(2,1) = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}, \qquad Dg(5,1) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Otrzymujemy

$$D(g\circ f)(2,1) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 4 & 0 \\ 0 & 4 \end{bmatrix}.$$

Twierdzenie 5.14. *Jeśli funkcja $f : \mathbb{R}^n \to \mathbb{R}$ ma ciągłe pochodne cząstkowe w punkcie $a \in \mathbb{R}^n$, to f jest różniczkowalna w a.*

Dowód. Dla $h \in \mathbb{R}^n$ określamy ciąg wektorów v_i wzorami $v_0 = a$, $v_1 = v_0 + h_1e_1$, $v_2 = v_1 + h_2e_2$, … $v_n = v_{n-1} + h_ne_n = a + h$. Wtedy punkty v_j i v_{j-1} różnią się tylko na j-tej współrzędnej o h_j. Zatem

$$f(a+h) - f(a) = f(v_n) - f(v_0) = \sum_{j=1}^{n} [f(v_j) - f(v_{j-1})]$$
$$= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(w_j)\, h_j,$$

gdzie punkt w_j leży na odcinku pomiędzy v_{j-1} i v_j. Otrzymujemy

$$\left| f(a+h) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)\, h_j \right| = \left| \sum_{j=1}^{n} \left[\frac{\partial f}{\partial x_j}(w_j) - \frac{\partial f}{\partial x_j}(a) \right] h_j \right|$$
$$\leqslant \sum_{j=1}^{n} \left| \frac{\partial f}{\partial x_j}(w_j) - \frac{\partial f}{\partial x_j}(a) \right| |h_j| \leqslant \sum_{j=1}^{n} \left| \frac{\partial f}{\partial x_j}(w_j) - \frac{\partial f}{\partial x_j}(a) \right| \|h\|.$$

Zatem

$$\frac{\left| f(a+h) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)\, h_j \right|}{\|h\|} \leqslant \sum_{j=1}^{n} \left| \frac{\partial f}{\partial x_j}(w_j) - \frac{\partial f}{\partial x_j}(a) \right|.$$

Gdy $h \to 0$, to $v_j \to a$ dla $j = 1, 2, \ldots, n$. Wtedy $w_j \to a$. Reasumując, prawa strona ostatniego wzoru dąży do zera. □

Uwaga 5.15. Z twierdzenia otrzymujemy hierarchię własności funkcji: ciągłe pochodne cząstkowe dają różniczkowalność, z której wynika istnienie pochodnych cząstkowych.

Twierdzenie 5.16 (reguła łańcucha)**.** *Załóżmy, że funkcje* $f_1, f_2, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$ *mają ciągłe pochodne cząstkowe w punkcie* a*. Załóżmy też, że funkcja* $g : \mathbb{R}^m \to \mathbb{R}$ *ma ciągłe pochodne cząstkowe w punkcie* $b = (f_1(a), f_2(a), \ldots, f_m(a))$*. Przyjmijmy, że* $f_j(x) = f_j(x_1, x_2, \ldots, x_n)$ *oraz* $g(y) = g(y_1, y_2, \ldots, y_m)$*. Wtedy funkcja złożona* $G : \mathbb{R}^n \to \mathbb{R}$ *określona wzorem*

$$G(x_1, x_2, \ldots, x_n) = g(f_1(x), f_2(x), \ldots, f_m(x))$$

jest różniczkowalna w punkcie a *oraz*

$$\frac{\partial G}{\partial x_i}(a) = \frac{\partial g}{\partial y_1}(b)\frac{\partial f_1}{\partial x_i}(a) + \frac{\partial g}{\partial y_2}(b)\frac{\partial f_2}{\partial x_i}(a) + \ldots + \frac{\partial g}{\partial y_m}(b)\frac{\partial f_m}{\partial x_i}(a).$$

Dowód. Tworzymy funkcję $F : \mathbb{R}^n \to \mathbb{R}^m$ wzorem

$$F(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{pmatrix}.$$

Wtedy $G(x) = g(F(x))$. Zatem funkcja G jest różniczkowalna w punkcie a oraz $DG(a) = Dg(b)\, DF(a)$. Ale

$$DG(a) = \left(\frac{\partial G}{\partial x_1}(a), \frac{\partial G}{\partial x_2}(a), \dots, \frac{\partial G}{\partial x_n}(a) \right)$$
$$Dg(b) = \left(\frac{\partial g}{\partial y_1}(b), \frac{\partial g}{\partial y_2}(b), \dots, \frac{\partial g}{\partial y_m}(b) \right)$$

$$DF(a) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & \vdots & \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{bmatrix}.$$

Aby obliczyć $\dfrac{\partial G}{\partial x_i}(a)$ mnożymy skalarnie wiersz $Dg(b)$ przez i-tą kolumnę macierzy $DF(a)$. □

Twierdzenie 5.17. *Jeśli dla funkcji n zmiennych $f : \mathbb{R}^n \to \mathbb{R}$ i ustalonych $i \neq j$ pochodne cząstkowe $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ oraz $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ są ciągłe w punkcie a, to*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Dowód. Ustalmy $i < j$. Przy obliczaniu $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)$ i $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(a)$ w punkcie a używamy funkcji

$$g(x, y) = f(a_1, \dots, \underset{x}{\cancel{a_i}}, \dots, \underset{y}{\cancel{a_j}}, \dots, a_n).$$

Wystarczy zatem rozważyć przypadek funkcji $g : \mathbb{R}^2 \to \mathbb{R}$ zakładając, że pochodne mieszane $\dfrac{\partial^2 g}{\partial x \partial y}$ i $\dfrac{\partial^2 g}{\partial y \partial x}$ są ciągłe w punkcie (c, d). Dla przyrostu $h = (h_1, h_2)$ rozważmy wyrażenie

$$I = g(c + h_1, d + h_2) - g(c, d + h_2) - g(c + h_1, d) + g(c, d).$$

Oznaczmy $\varphi(y) = g(c + h_1, y) - g(c, y)$. Wtedy

$$\begin{aligned} I &= \varphi(d + h_2) - \varphi(d) = \varphi'(d + \theta_1 h_2)h_2 \\ &= \left[\frac{\partial g}{\partial y}(c + h_1, d + \theta_1 h_2) - \frac{\partial g}{\partial y}(c, d + \theta_1 h_2)\right] h_2 = \frac{\partial^2 g}{\partial x \partial y}(c+\theta_2 h_1, d+\theta_1 h_2)\, h_1 h_2. \end{aligned}$$

Zamieniając rolami x i y otrzymamy

$$I = \frac{\partial^2 g}{\partial y \partial x}(c + \theta_1' h_1, d + \theta_2' h_2)\, h_1 h_2.$$

Przyjmijmy, że $h_1, h_2 > 0$. Wtedy

$$\frac{\partial^2 g}{\partial x \partial y}(c + \theta_2 h_1, d + \theta_1 h_2) = \frac{\partial^2 g}{\partial y \partial x}(c + \theta_1' h_1, d + \theta_2' h_2).$$

Przechodzimy do granicy, gdy $h_1, h_2 \to 0$ i otrzymujemy

$$\frac{\partial^2 g}{\partial x \partial y}(c, d) = \frac{\partial^2 g}{\partial y \partial x}(c, d).$$

Zauważmy, że

$$\frac{\partial^2 g}{\partial x \partial y}(a_i, a_j) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a), \qquad \frac{\partial^2 g}{\partial y \partial x}(a_i, a_j) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

□

5.4 Geometria odwzorowań z $\mathbb{R}^n$ w $\mathbb{R}^m$

Rozważmy odwzorowanie $c : \mathbb{R} \to \mathbb{R}^3$. Takie odwzorowanie nazywamy krzywą

$$c(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \\ c_3(t) \end{pmatrix}.$$

Wektorem siecznym odpowiadajacym dwu momentom czasu t i $t+h$ jest

$$\frac{c(t+h)-c(t)}{h}.$$

Wektor styczny do krzywej w punkcie $c(t)$ otrzymujemy przez przejście graniczne

$$c'(t)=\lim_{h\to 0}\frac{c(t+h)-c(t)}{h}.$$

Mamy $c'(t)=Dc(t)$.

Przykład. Znaleźć wektor styczny do krzywej

$$c(t)=\begin{pmatrix}-\sin t\\ \cos t\\ t\end{pmatrix}$$

odpowiadający momentowi $t=\pi$. Chodzi więc o wektor styczny w punkcie $c(\pi)=\begin{pmatrix}0\\ -1\\ \pi\end{pmatrix}$

$$c'(t)=\begin{pmatrix}-\cos t\\ -\sin t\\ 1\end{pmatrix},\qquad c'(\pi)=\begin{pmatrix}1\\ 0\\ 1\end{pmatrix}$$

Niech $f:\mathbb{R}^3\to\mathbb{R}^3$. Wtedy $\sigma(t)=f(c(t))$ jest nową krzywą. Różniczkując otrzymamy

$$\sigma'(t)=Df(c(t))\,c'(t).$$

Funkcja f przekształca $c(t)$ na $\sigma(t)$. Odwzorowanie $Df(c(t))$ przekształca wektor styczny $c'(t)$ na wektor styczny $\sigma'(t)$. Podobna interpretacja dotyczy krzywych $c:\mathbb{R}\to\mathbb{R}^n$ i funkcji $f:\mathbb{R}^n\to\mathbb{R}^m$.

5.5 Gradient i poziomice funkcji

Pochodne cząstkowe w jednym punkcie nie dają pełnej informacji o zachowaniu się funkcji.

Definicja 5.18. *Załóżmy, że funkcja $f:\mathbb{R}^n\to\mathbb{R}$ jest różniczkowalna w punkcie $a\in\mathbb{R}^n$. Gradientem nazywamy wektor*

$$\nabla f(a)=\left(\frac{\partial f}{\partial x_1}(a),\ldots,\frac{\partial f}{\partial x_n}(a)\right).$$

Tzn. $\nabla f(a)=Df(a)$.

Ustalmy punkt x i wektor $v \in \mathbb{R}^n$. Chcemy zbadać tempo zmiany wartości funkcji $f : \mathbb{R}^n \to \mathbb{R}$ w punkcie x wzdłuż prostej $x + tv$, gdzie $t \in \mathbb{R}$. W tym celu obliczamy $\frac{d}{dt} f(x+tv)\Big|_{t=0}$. Z definicji mamy

$$\frac{d}{dt} f(x+tv)\Big|_{t=0} = \lim_{h\to 0} \frac{f(x+hv) - f(x)}{h}.$$

To wyrażenie nazywamy pochodną kierunkową w punkcie x w kierunku v. Dla $v = e_j$ otrzymamy w wyniku $\frac{\partial f}{\partial x_j}(x)$. Tzn. pochodne cząstkowe są również pochodnymi kierunkowymi w kierunkach równoległych do poszczególnych osi współrzędnych. Ze wzoru na pochodną funkcji złożonej otrzymujemy

$$\frac{d}{dt} f(x+tv) = \frac{\partial f}{\partial x_1}(x+tv)\, v_1 + \ldots + \frac{\partial f}{\partial x_n}(x+tv)\, v_n,$$

zatem

$$\frac{d}{dt} f(x+tv)\Big|_{t=0} = \nabla f(x) \circ v. \tag{5.5}$$

Przykład. $f(x,y,z) = e^{xy} \sin z$. Chcemy obliczyć pochodną kierunkową w punkcie $(0, 1, \frac{\pi}{2})$ w kierunku wektora $(2,1,1)$. Mamy

$$\nabla f(x,y,z) = (ye^{xy} \sin z, xe^{xy} \sin z, e^{xy} \cos z), \qquad \nabla f(0,1,\frac{\pi}{2}) = (1,0,0).$$

Zatem

$$\nabla f(0,1,\frac{\pi}{2}) \circ (2,1,1) = 2.$$

Przy porównywaniu pochodnych kierunkowych w różnych kierunkach wybiera się wektory v o długości 1. Liczbę $\nabla f(x) \circ v$ można interpretować jako tempo zmiany wartości funkcji f w kierunku v, gdy prędkość zmiany argumentu wynosi 1.

Twierdzenie 5.19. *Załóżmy, że $\nabla f(x) \neq 0$. Wtedy wektor $\nabla f(x)$ wskazuje kierunek najszybszego wzrostu wartości funkcji f, startując z punktu x. Z kolei wektor $-\nabla f(x)$ wskazuje kierunek najszybszego spadku wartości funkcji.*

Dowód. Niech v będzie dowolnym wektorem o długości 1. Tempo zmiany wartości funkcji w kierunku v wynosi

$$\nabla f(x) \circ v \leqslant \|\nabla f(x)\|\, \|v\| = \|\nabla f(x)\|.$$

Równość zachodzi tylko wtedy, gdy wektory v i $\nabla f(x)$ mają ten sam kierunek i zwrot, tzn. $v = \dfrac{\nabla f(x)}{\|\nabla f(x)\|}$. $\square$

Przykład. W którym kierunku od punktu $(0,1)$ funkcja $f(x,y) = x^2 - y^2$ rośnie najszybciej ? Mamy $\nabla f(x,y) = (2x, -2y)$. Zatem $\nabla f(0,1) = (0,-2)$. Funkcja rośnie najszybciej w kierunku $(0,-1)$.

Definicja 5.20. *Poziomicą funkcji $f : \mathbb{R}^n \to \mathbb{R}$ nazywamy zbiór postaci $\{x \in \mathbb{R}^n : f(x) = c\}$ dla ustalonej wartości c.*

Przykład. $f(x,y,z) = x^2 + y^2 + z^2$. Poziomica $\{(x,y,z) : x^2 + y^2 + z^2 = 1\}$ jest sferą. Mamy $\nabla f(x,y,z) = 2(x,y,z)$. Zatem gradient jest prostopadły do sfery.

Wartość funkcji, gdy argument porusza się po poziomicy nie zmienia się. Wydaje się, że gradient powinien być zawsze prostopadły do poziomicy.

Twierdzenie 5.21. *Załóżmy, że funkcja $f : \mathbb{R}^n \to \mathbb{R}$ jest różniczkowalna w punkcie $x_0 \in \mathbb{R}^n$. Wtedy gradient $\nabla f(x_0)$ jest prostopadły do poziomicy $S = \{x \in \mathbb{R}^n : f(x) = f(x_0)\}$ w następującym sensie: dla dowolnej krzywej $c(t)$ leżącej na poziomicy S, spełniającej $c(0) = x_0$ i $c'(0) = v$, mamy $\nabla f(x_0) \perp v$.*

Dowód. Funkcja $f(c(t))$ jest stała. Zatem

$$0 = \frac{d}{dt} f(c(t)) = Df(c(t))\, c'(t) = \nabla f(c(t)) \circ c'(t).$$

Dla $t = 0$ otrzymujemy $\nabla f(x_0) \circ v = 0$. $\square$

Definicja 5.22. *Niech $S = \{x \in \mathbb{R}^n : f(x) = c\}$. Przestrzenią styczną do poziomicy S w punkcie x_0 nazywamy hiperprzestrzeń określoną przez*

$$\nabla f(x_0) \circ (x - x_0) = 0, \quad \text{o ile } \nabla f(x_0) \neq 0.$$

Przykład. Znaleźć równanie płaszczyzny stycznej do powierzchni $3xy + z^2 = 4$ w punkcie $(1,1,1)$. Chodzi o poziomicę funkcji $f(x,y,z) = 3xy + z^2$. Obliczamy

$$\nabla f(x,y,z) = (3y, 3x, 2z), \qquad \nabla f(1,1,1) = (3,3,2).$$

Równanie ma postać $3(x-1) + 3(y-1) + 2(z-1) = 0$, po uproszczeniu $3x + 3y + 2z = 8$.

Uwaga. Określenie płaszczyzny stycznej do poziomicy zgadza sie z określeniem płaszczyzny stycznej do wykresu funkcji $z = f(x, y)$. Rzeczywiście, niech $z_0 = f(x_0, y_0)$. Rozważamy funkcję $F(x, y, z) = f(x, y) - z$. Wykres funkcji f można utożsamić z poziomicą funkcji F przy $c = 0$. Mamy

$$\nabla F(x_0, y_0, z_0) = \left(\underbrace{\frac{\partial f}{\partial x}(x_0, y_0)}_{a}, \underbrace{\frac{\partial f}{\partial y}(x_0, y_0)}_{b}, -1 \right).$$

Równanie ma postać $a(x - x_0) + b(y - y_0) - (z - z_0) = 0$, czyli $z - z_0 = a(x - x_0) + b(y - y_0)$.

Definicja 5.23. *Zbiór $D \subset \mathbb{R}^n$ nazywamy ograniczonym, jeśli*

$$D \subset \{x \in \mathbb{R}^n \, : \, \|x\| \leqslant K\}$$

dla pewnej stałej liczby K.

Uwagi.

(a) Dla $n = 3$ warunek w definicji oznacza, że zbiór D jest zawarty w kuli o środku w $(0, 0, 0)$ i promieniu K.

(b) Jeśli $\|x\| \leqslant K$, to $|x_i| \leqslant K$ dla $i = 1, 2, \ldots, n$. Zatem współrzędne punktów ze zbioru ograniczonego D są wspólnie ograniczone. Odwrotnie, jeśli współrzędne punktów z D są wspólnie ograniczone, tzn. $|x_i| \leqslant K$ dla $i = 1, 2, \ldots, n$ i $x \in D$, to

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \leqslant \sqrt{nK^2} = K\sqrt{n}.$$

Definicja 5.24. *Mówimy, że zbiór $D \subset \mathbb{R}^d$ jest domknięty, jeśli dla dowolnego ciągu $x^{(n)} \in D$ z warunku $\|x^{(n)} - x\| \xrightarrow[n]{} 0$ wynika $x \in D$.*

Przykład. Dla funkcji ciągłej $f : \mathbb{R}^d \to \mathbb{R}$ zbiór $\{x \in \mathbb{R}^d \, : \, f(x) \leqslant c\}$ jest domknięty. Rzeczywiście, niech $f(x^{(n)}) \leqslant c$ oraz $\|x^{(n)} - x\| \xrightarrow[n]{} 0$. Wtedy z ciągłości mamy $f(x) = \lim_n f(x^{(n)})$. Zatem $f(x) \leqslant c$. Również zbiór postaci $\{x \in \mathbb{R}^d \, : \, f(x) = c\}$ jest domknięty bo

$$\{x \in \mathbb{R}^d \, : \, f(x) = c\} = \{x \in \mathbb{R}^d \, : \, f(x) \leqslant c\} \cap \{x \in \mathbb{R}^d \, : \, -f(x) \leqslant -c\}.$$

Zbiory

$$B = \{(x, y, z) \,:\, x^2 + y^2 + z^2 \leqslant 1\}, \qquad S = \{(x, y, z) \,:\, x^2 + y^2 + z^2 = 1\}$$

są domknięte.

Definicja 5.25. *Zbiór $D \subset \mathbb{R}^d$ nazywamy zwartym, jeśli D jest domknięty i ograniczony.*

Twierdzenie 5.26. *Funkcja ciągła określona na zbiorze zwartym jest ograniczona i osiąga swoje kresy dolny i górny.*

Lemat 5.27. *Niech D będzie zwartym podzbiorem w $\mathbb{R}^d$. Każdy ciąg $x^{(n)}$ punktów z D zawiera podciąg zbieżny do punktu ze zbioru D.*

Dowód. Wiemy, że ciągi współrzędnych $x_1^{(n)}, x_2^{(n)}, \ldots, x_d^{(n)}$ są ograniczone. Z twierdzenia Bolzano-Weierstrassa istnieje ciąg wskaźników n_k taki, że ciągi $x_1^{(n_k)}, x_2^{(n_k)}, \ldots, x_d^{(n_k)}$ są zbieżne. Oznaczmy $\lim\limits_k x_i^{(n_k)} = x_i$. Wtedy dla $x = (x_1, x_2, \ldots, x_d)$ mamy $\|x^{(n_k)} - x\| \underset{k}{\longrightarrow} 0$. Z domkniętości zbioru D mamy $x \in D$. □

Dowód twierdzenia. Załóżmy nie wprost, że istnieje ciąg $x^{(n)} \in D$ taki, że $|f(x^{(n)})| > n$. Z ciągu $x^{(n)}$ wybieramy podciąg $x^{(n_k)}$ zbieżny np. do x. Z ciągłości mamy $f(x^{(n_k)}) \underset{k}{\longrightarrow} f(x)$. Ale $|f(x^{(n_k)})| > n_k \underset{k}{\longrightarrow} \infty$. Dowód drugiej części tezy można przeprowadzić tak samo jak dowód Twierdzenia 3.18 z części I. □

Dla funkcji $f : \mathbb{R}^n \to \mathbb{R}$ odwzorowanie $x \mapsto \nabla f(x)$ nazywamy gradientowym polem wektorowym. Dla $n = 3$ wykres pola leży w $\mathbb{R}^6$.

Przykład. W początku układu $\mathbb{R}^3$ umieszczamy dużą masę M. Na masę m umieszczoną w punkcie (x, y, z) działa siła przyciągania

$$F(x, y, z) = -\frac{GMm}{r^2}\,\vec{n}, \quad r = \sqrt{x^2 + y^2 + z^2},\ \vec{n} = \frac{(x, y, z)}{r}.$$

Określamy funkcję

$$V(x, y, z) = \frac{GMm}{r}.$$

Funkcję $V(x, y, z)$ nazywamy potencjałem grawitacyjnym. Mamy $\nabla V = F$. Rzeczywiście

$$\nabla\left(\frac{1}{r}\right) = \left(\frac{\partial}{\partial x}\left(\frac{1}{r}\right), \frac{\partial}{\partial y}\left(\frac{1}{r}\right), \frac{\partial}{\partial z}\left(\frac{1}{r}\right)\right)$$

oraz

$$\frac{\partial}{\partial x}\left(\frac{1}{\sqrt{x^2+y^2+z^2}}\right) = -\frac{1}{2}\frac{2x}{(\sqrt{x^2+y^2+z^2})^3} = -\frac{x}{r^3}.$$

Zatem

$$\nabla\left(\frac{1}{r}\right) = -\frac{1}{r^3}(x,y,z) = -\frac{1}{r^2}\,\vec{n}.$$

5.6 Ekstrema funkcji wielu zmiennych

Definicja 5.28. *Załóżmy, że funkcja ciągła f o wartościach rzeczywistych jest określona na podzbiorze $\mathbb{R}^n$. Mówimy, że punkt x_0 jest lokalnym minimum funkcji f, jeśli dla pewnej liczby $\delta > 0$ mamy $f(x) \geqslant f(x_0)$ dla $\|x - x_0\| < \delta$.*

Przykłady.

(a) $f(x,y) = x^2 + y^2$. W punkcie $(0,0)$ występuje minimum.

(b) $g(x,y) = x^2 - y^2$. W punkcie $(0,0)$ nie ma lokalnego minimum ani maksimum.

Twierdzenie 5.29. *Załóżmy, że $f : \mathbb{R}^n \to \mathbb{R}$ jest różniczkowalna i posiada lokalne ekstremum w punkcie x_0. Wtedy $\nabla f(x_0) = 0$.*

Dowód. Ustalmy wektor $v \in \mathbb{R}^n$. Rozważamy funkcję $g(t) = f(x_0 + tv)$. Jeśli x_0 jest lokalnym minimum funkcji $f(x)$, to funkcja $g(t)$ posiada lokalne minimum w punkcie $t = 0$. Zatem $g'(0) = 0$. Ale

$$0 = g'(0) = \frac{d}{dt} f(x_0 + tv)\bigg|_{t=0} = \nabla f(x_0) \circ v.$$

Ponieważ v jest dowolnym wektorem, to $\nabla f(x_0) = 0$. [7] □

Definicja 5.30. *Punkty x_0, dla których $\nabla f(x_0) = 0$ nazywamy stacjonarnymi.*

Definicja 5.31. *Mówimy, że funkcja $f : \mathbb{R}^n \to \mathbb{R}$ jest klasy C^2 jeśli jest dwukrotnie różniczkowalna i wszystkie pochodne cząstkowe drugiego rzędu są ciągłe. Dla funkcji klasy C^2 macierz*

$$Hf(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{i,j=1}^{n}$$

[7] Jeśli $a \circ v = 0$ dla wszystkich $v \in \mathbb{R}^n$, to $a = 0$. Rzeczywiście $\|a\|^2 = a \circ a = 0$.

nazywamy hessjanem. Hessjan jest macierzą symetryczną, bo

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x).$$

Definicja 5.32. *Macierz kwadratowa A wymiaru $n \times n$ jest dodatnio określona, jeśli jest symetryczna, tzn. $a_{ij} = a_{ji}$, oraz $Av \circ v > 0$ dla $v \in \mathbb{R}^n$, $v \neq 0$.*

Uwaga.

$$Av \circ v = \sum_{i=1}^{n} (Av)_i v_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} v_j \right) v_i = \sum_{i,j=1}^{n} a_{ij} v_i v_j.$$

Lemat 5.33. *Dla macierzy dodatnio określonej A istnieje liczba $\delta > 0$ taka, że jeśli $|b_{ij} - a_{ij}| < \delta$ oraz $b_{ij} = b_{ji}$, to macierz $B = (b_{ij})_{i,j=1}^n$ też jest dodatnio określona.*

Dowód. Wystarczy sprawdzić, że $Bv \circ v > 0$ dla $\|v\| = 1$. Rzeczywiście, każdy wektor $w \neq 0$ można zapisać jako $w = \lambda v$, gdzie $\|v\| = 1$, oraz $\lambda = \|w\|$. Wtedy $Bw \circ w = \lambda^2 \, Bv \circ v$. Zbiór $S = \{v \in \mathbb{R}^n \, : \, v_1^2 + v_2^2 + \ldots + v_n^2 = 1\}$ jest zwarty. Funkcja

$$S \ni v \longmapsto Av \circ v = \sum_{i,j}^{n} a_{ij} v_i v_j$$

jest ciągła. Ta funkcja osiąga minimum w pewnym punkcie v_0 na zbiorze S. Tzn.

$$Av \circ v \geqslant Av_0 \circ v_0 =: m > 0, \quad \text{dla } v \in S.$$

Załóżmy, że $|b_{ij} - a_{ij}| < \delta$. Wtedy dla $\|v\| = 1$ mamy

$$\begin{aligned} Bv \circ v = Av \circ v + (B - A)v \circ v \geqslant m + \sum_{i,j=1}^{n} (b_{ij} - a_{ij}) v_i v_j \\ \geqslant m - \sum_{i,j=1}^{n} |b_{ij} - a_{ij}| \, |v_i| \, |v_j| \geqslant m - \delta \sum_{i,j=1}^{n} |v_i| \, |v_j| \\ = m - \delta \left(\sum_{i=1}^{n} |v_i| \right)^2 \geqslant m - n^2 \delta. \end{aligned}$$

Przyjmijmy $\delta = \dfrac{m}{2n^2}$. Wtedy $Bv \circ v > 0$. □

Twierdzenie 5.34. *Załóżmy, że funkcja* $f : \mathbb{R}^n \to \mathbb{R}$ *jest klasy* C^2 *oraz* $\nabla f(x_0) = 0$ *dla pewnego punktu* $x_0 \in \mathbb{R}^n$.

(i) Jeśli macierz $Hf(x_0)$ *jest dodatnio określona, to w punkcie* x_0 *funkcja* $f(x)$ *posiada lokalne minimum.*

(ii) Jeśli macierz $-Hf(x_0)$ *jest dodatnio określona, to w punkcie* x_0 *funkcja* $f(x)$ *posiada lokalne maksimum.*

(iii) Jeśli macierz $Hf(x_0)$ *posiada wartości własne różnych znaków, to w punkcie* x_0 *nie ma lokalnego ekstremum.*

Dowód. (i) Zastosujemy oznaczenie $x := x_0$. Ustalmy wektor $0 \neq v \in \mathbb{R}^n$ i rozważmy funkcję jednej zmiennej $g(t) = f(x+tv)$. Z reguły łańcucha mamy

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x+tv)v_i = \nabla f(x+tv) \circ v.$$

Zatem $g'(0) = \nabla f(x) \circ v = 0$. Obliczamy $g''(t)$.

$$\begin{aligned} g''(t) = \frac{d}{dt}\left(\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x+tv)v_i\right) &= \sum_{i=1}^{n} \frac{d}{dt}\left(\frac{\partial f}{\partial x_i}(x+tv)\right) v_i \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(x+tv)v_jv_i = Hf(x+tv)\, v \circ v. \end{aligned}$$

Zatem $g''(0) = Hf(x)v \circ v > 0$. To oznacza, że funkcja $g(t)$ posiada ścisłe lokalne minimum w punkcie $t = 0$. Ze wzoru MacLaurina dla $n = 2$ otrzymujemy

$$g(t) = g(0) + g'(0)t + \frac{1}{2}g''(\theta t)t^2$$

dla pewnej wartości $0 < \theta < 1$. Dla $t = 1$ otrzymujemy

$$f(x+v) = f(x) + \frac{1}{2}\sum_{i,j=1}^{n} \widetilde{h}_{ij}v_iv_j,$$

gdzie

$$\widetilde{h}_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x+\theta v).$$

Jeśli wektor v ma odpowiednio małą normę, to liczby $\widetilde{h}_{ij}$ leżą blisko liczb $h_{ij} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(x)$, z założenia o ciągłości drugich pochodnych cząstkowych.

Wtedy z lematu macierz $\widetilde{H} = (\widetilde{h}_{ij})_{i,j=1}^n$ jest dodatnio określona, jeśli tylko norma $\|v\|$ jest odpowiednio mała, np. $\|v\| < \eta$ dla pewnej liczby $\eta > 0$. Wtedy

$$f(x+v) = f(x) + \frac{1}{2}\sum_{i,j=1}^n \widetilde{h}_{ij} v_i v_j > f(x)$$

dla $0 < \|v\| < \eta$.

Załóżmy, że macierz $Hf(x)$ ma wartości własne $\lambda_1 > 0$ i $\lambda_2 < 0$. Niech v_1 i v_2 oznaczają odpowiadające jednostkowe wektory własne. Rozważmy funkcje

$$g_1(t) = f(x + tv_1), \qquad g_2(t) = f(x + tv_2).$$

Wtedy pochodne tych funkcji w $t = 0$ zerują się. Ponadto

$$g_1''(0) = Hf(x)v_1 \circ v_1 = \lambda_1 v_1 \circ v_1 = \lambda_1 > 0,$$

$$g_2''(0) = Hf(x)v_2 \circ v_2 = \lambda_2 v_2 \circ v_2 = \lambda_2 < 0.$$

Zatem g_1 posiada ścisłe lokalne minimum w punkcie 0, natomiast g_2 posiada ścisłe lokalne maksimum w tym punkcie. To oznacza, że na prostej $t \mapsto x+tv_1$ funkcja f posiada minimum w punkcie x natomiast na prostej $t \mapsto x + tv_2$ ma w punkcie x lokalne maksimum. $\square$

Uwaga. Wektory v_1 i v_2 z ostatniej części dowodu są do siebie prostopadłe.

Zadanie. Znaleźć funkcję $f(x, y)$ taką, że funkcja $t \mapsto f(tu, tv)$ przyjmuje minimum dla $t = 0$ dla dowolnego wektora $(u, v) \neq 0$, ale $f(x, y)$ nie posiada minimum w punkcie $(0, 0)$.

5.7 Ekstrema warunkowe-metoda mnożników Lagrange'a

Często chcemy znaleźć maksimum i minimum funkcji wielu zmiennych, ale przy pewnych ograniczeniach.

Przykład. Firma sprzedaje produkty A i B. Zysk ze sprzedaży wynosi $f(x, y)$, gdzie x i y oznaczają ilości sprzedanych produktów A i B, odpowiednio. Ze względu na ograniczone zasoby finansowe musi być spełniony warunek $g(x, y) = c$.

Twierdzenie 5.35 (Lagrange)**.** *Załóżmy, że funkcje* $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ *i* $g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ *są klasy* C^1*. Niech* $S = \{x \in U : g(x) = c\}$*. Jeśli funkcja* $f\big|_S$ *przyjmuje minimum lub maksimum w punkcie* x_0 *oraz* $\nabla g(x_0) \neq 0$*, to* $\nabla f(x_0) = \lambda \nabla g(x_0)$ *dla pewnej stałej* λ*. Tzn. gradienty* $\nabla f(x_0)$ *i* $\nabla g(x_0)$ *są równoległe.*

Dowód nieścisły. Wiemy, że przestrzeń styczna do poziomicy S w punkcie x_0 składa się z wektorów prostopadłych do $\nabla g(x_0)$. Niech $\sigma(t) : (-1, 1) \to S$ będzie krzywą klasy C^1 przechodzącą przez x_0 w chwili $t = 0$, tzn. $\sigma(0) = x_0$. Wtedy funkcja złożona $f(\sigma(t))$ przyjmuje ekstremum w chwili $t = 0$. Zatem

$$0 = \frac{d}{dt} f(\sigma(t))\Big|_{t=0} = Df(\sigma(t))\sigma'(t)\Big|_{t=0} = Df(x_0)\sigma'(0) = \nabla f(x_0) \circ \sigma'(0).$$

$\sigma'(0)$ jest wektorem stycznym do S w punkcie x_0. Tzn. gradient $\nabla f(x_0)$ jest prostopadły do każdego wektora stycznego do S w punkcie x_0. Np. jeśli v jest takim wektorem stycznym, to istnieje krzywa $\sigma : (-1, 1) \to S$ taka, że $\sigma(0) = x_0$ i $\sigma'(0) = v$. Zatem $\nabla f(x_0)$ jest prostopadły do przestrzeni stycznej do S w punkcie x_0. Ale $\nabla g(x_0)$ jest też prostopadły do tej przestrzeni stycznej. To oznacza, że $\nabla f(x_0)$ i $\nabla g(x_0)$ są równoległe. □

Uwaga. Nieścisłość polega na tym, że dla wektora v z przestrzeni stycznej do S w x_0 trzeba znaleźć krzywą $\sigma(t)$ taką, że $\sigma(t) \in S$ oraz $\sigma(0) = x_0$, $\sigma'(0) = v$. Taką krzywą można łatwo znaleźć, gdy poziomica S jest wykresem funkcji $n - 1$ zmiennych.

Przykład. Załóżmy, że $S = \{(x, y, z) : g(x, y, z) = 0\}$. Przypuśćmy, że $g(x, y, z) = h(x, y) - z$. Wtedy S jest wykresem funkcji $z = h(x, y)$. Niech (x_0, y_0, z_0) będzie punktem z S, tzn. $z_0 = h(x_0, y_0)$. Rozważmy krzywą

$$\sigma(t) = (x_0 + at, y_0 + bt, h(x_0 + at, y_0 + bt)).$$

Wtedy

$$\sigma'(0) = (a, b, c), \quad \text{gdzie } c = \nabla h(x_0, y_0) \circ (a, b),$$

jest wektorem stycznym do S w punkcie (x_0, y_0, z_0). Przestrzeń złożona z takich wektorów ma wymiar 2. Ale przestrzeń styczna do S w punkcie (x_0, y_0, z_0) ma również wymiar 2. Zatem przestrzenie te są równe.

5.7.1 Stosowanie metody Lagrange'a

Trzeba znaleźć punkt $x \in U$ i stałą λ takie, że

$$\begin{aligned}
\frac{\partial f}{\partial x_1}(x_1, x_2, \ldots, x_n) &= \lambda \frac{\partial g}{\partial x_1}(x_1, x_2, \ldots, x_n) \\
&\vdots \\
\frac{\partial f}{\partial x_n}(x_1, x_2, \ldots, x_n) &= \lambda \frac{\partial g}{\partial x_n}(x_1, x_2, \ldots, x_n) \\
g(x_1, x_2, \ldots, x_n) &= c
\end{aligned}$$

Mamy układ $n+1$ równań z $n+1$ zmiennymi $x_1, x_2, \ldots, x_n$ i λ.

Przykłady.

(a) Niech S będzie prostą przechodzącą przez $(-1, 0)$ o nachyleniu $45°$. Chcemy znaleźć minimum funkcji $f(x, y) = x^2 + y^2$ na S. Prosta ma równanie $y = x + 1$. Możemy rozwiązać zadanie na dwa sposoby.

   (i) Podstawiamy $y = x + 1$ do funkcji $f(x, y)$ i obliczamy minimum funkcji kwadratowej.

   (ii) Stosujemy metodę Lagrange'a. Prosta S jest poziomicą funkcji $g(x, y) = x - y + 1 = 0$. Mamy $\nabla f(x, y) = (2x, 2y)$ oraz $\nabla g(x, y) = (1, -1)$. Gradienty są równoległe tylko wtedy, gdy $y = -x$. W połączeniu z równaniem $y = x + 1$ otrzymamy $x = -\frac{1}{2},\ y = \frac{1}{2}$.

(b) $f(x, y) = x^2 - y^2$, $S = \{(x, y) : x^2 + y^2 = 1\}$.

   (i) Podstawiamy $y^2 = 1 - x^2$. Wtedy $f(x, y) = 2x^2 - 1$. Obliczamy ekstrema na przedziale $[-1, 1]$.

   (ii) Parametryzujemy okrąg $x = \cos t$, $y = \sin t$ i obliczamy ekstrema funkcji $\cos 2t$.

   (iii) $\nabla f(x, y) = (2x, -2y)$ oraz $\nabla g(x, y) = (2x, 2y)$. Wektory są równoległe, gdy $x = 0$ lub $y = 0$. Otrzymujemy cztery rozwiązania $(\pm 1, 0)$ i $(0, \pm 1)$. Mamy $f(\pm 1, 0) = 1$ oraz $f(0, \pm 1) = -1$.

(c) $f(x, y, z) = x + z$, $S = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$. Mamy

$$\nabla f(x, y, z) = (1, 0, 1), \quad \nabla g(x, y, z) = (2x, 2y, 2z).$$

Wektory te są równoległe, gdy $y = 0$ oraz $z = x$. Zatem $2x^2 = 1$. Otrzymujemy dwa rozwiązania $\pm\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$ oraz

$$f\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) = \sqrt{2}, \qquad f\left(-\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right) = -\sqrt{2}.$$

(d) Rozważmy macierz symetryczną A wymiaru $n \times n$. Określamy

$$f(x) = (Ax, x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j, \quad x = (x_1, x_2, \ldots, x_n).$$

Chcemy znaleźć ekstrema funkcji $f(x)$ na

$$S = \{(x_1, x_2, \ldots, x_n) \,:\, g(x) := x_1^2 + x_2^2 + \ldots + x_n^2 = 1\}.$$

Określmy

$$F(x, y) = (Ax, y) = \sum_{i,j=1}^{n} a_{ij} x_i y_j.$$

Wtedy $f(x) = F(x, x)$. Obliczamy pomocniczo pochodne cząstkowe funkcji F.

$$\frac{\partial F}{\partial x_k}(x, y) = \sum_{j=1}^{n} a_{kj} y_j, \qquad \frac{\partial F}{\partial y_k}(x, y) = \sum_{i=1}^{n} a_{ik} x_i.$$

Mamy zatem

$$\begin{aligned}\frac{\partial f}{\partial x_k}(x) = \frac{\partial F}{\partial x_k}(x, x) + \frac{\partial F}{\partial y_k}(x, x) &= \sum_{j=1}^{n} a_{kj} x_j + \sum_{i=1}^{n} a_{ik} x_i \\ &= \sum_{j=1}^{n} a_{kj} x_j + \sum_{i=1}^{n} a_{ki} x_i = 2 \sum_{j=1}^{n} a_{kj} x_j.\end{aligned}$$

Dalej

$$\frac{\partial g}{\partial x_k}(x) = 2x_k.$$

Otrzymujemy więc układ równań

$$\begin{aligned}
\sum_{j=1}^{n} a_{1j}x_j &= \lambda x_1,\\
\sum_{j=1}^{n} a_{2j}x_j &= \lambda x_2,\\
&\vdots\\
\sum_{j=1}^{n} a_{nj}x_j &= \lambda x_n.
\end{aligned}$$

To oznacza, że $Ax = \lambda x$. Czyli x jest wektorem własnym o długości 1. Uporządkujmy wartości własne macierzy A według wielkości: $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$. Niech $v_1, v_2, \ldots, v_n$ oznaczają odpowiadające wektory własne o długości 1. Wtedy

$$f(v_k) = (Av_k, v_k) = \lambda_k(v_k, v_k) = \lambda_k.$$

Reasumując

$$\min_{\|x\|=1}(Ax, x) = \lambda_1, \qquad \max_{\|x\|=1}(Ax, x) = \lambda_n.$$

5.7.2 Procedura znajdowania wartości największej i najmniejszej funkcji na zbiorze zwartym

1. Znaleźć punkty krytyczne funkcji wewnątrz zbioru, tzn. punkty stacjonarne oraz punkty, w których nie można obliczyć pochodnych cząstkowych.

2. Znaleźć punkty krytyczne funkcji obciętej do brzegu zbioru, np. metodą mnożników Lagrange'a.

3. Obliczyć wartości funkcji w znalezionych punktach.

4. Wybrać wartość największą i najmniejszą.

5.7.3 Metoda mnożników Lagrange'a przy kilku warunkach

Załóżmy, że powierzchnia $S \subset \mathbb{R}^n$ jest określona przez k warunków

$$\begin{aligned} g_1(x_1, x_2, \ldots, x_n) &= c_1, \\ g_2(x_1, x_2, \ldots, x_n) &= c_2, \\ &\vdots \\ g_k(x_1, x_2, \ldots, x_n) &= c_k. \end{aligned}$$

Twierdzenie 5.36. *Jeśli funkcja $f\Big|_S$ posiada ekstremum w punkcie $x_0 \in S$ oraz gradienty $\nabla g_1(x_0), \nabla g_1(x_0), \ldots, \nabla g_k(x_0)$ są liniowo niezależne, to*

$$\nabla f(x_0) = \lambda_1 \nabla g_1(x_0) + \lambda_2 \nabla g_2(x_0) + \ldots + \lambda_k \nabla g_k(x_0)$$

dla pewnych stałych $\lambda_1, \lambda_2, \ldots, \lambda_k$.

Uwaga. Aby znaleźć punkt x_0 trzeba rozwiązać $n + k$ równań przy $n + k$ niewiadomych.

Przykład. Znaleźć ekstrema funkcji $f(x, y, z) = y + z$ przy warunkach $x^2 + z^2 = 1$ i $y^2 + z^2 = 4$. Możemy przyjąć $g_1(x, y, z) = x^2 + y^2$ oraz $g_2(x, y, z) = y^2 + z^2$. Rozwiązujemy równanie $\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2$. Otrzymujemy 3 równania

$$\begin{aligned} 0 &= 2\lambda_1 x, \\ 1 &= 2\lambda_2 y, \\ 1 &= 2\lambda_1 z + 2\lambda_2 z. \end{aligned}$$

Rozpatrzymy dwa przypadki.

(a) $x = 0$. Wtedy $z = \pm 1$ oraz $y = \pm\sqrt{3}$.

(b) $\lambda_1 = 0$. Wtedy $y = z$, zatem $z^2 = 2$. Otrzymujemy sprzeczność z warunkiem $x^2 + z^2 = 1$.

Wartość największa jest osiągnięta w punkcie $(0, \sqrt{3}, 1)$ a wartość najmniejsza w $(0, -\sqrt{3}, -1)$.

Nieścisły dowód twierdzenia. Niech $\sigma(t)$ będzie krzywą klasy C^1 leżącą w powierzchni S taką, że $\sigma(0) = x_0$. Mamy

$$g_j(\sigma(t)) = c_j, \quad \text{dla } j = 1, 2, \ldots, k.$$

Zatem

$$0 = \frac{d}{dt} g_j(\sigma(t)) = \nabla g_j(\sigma(t)) \circ \sigma'(t).$$

Dla $t = 0$ otrzymujemy

$$\nabla g_j(x_0) \circ \sigma'(0) = 0, \quad \text{dla } j = 1, 2, \ldots, k.$$

To oznacza, że wektor $\sigma'(0)$ jest prostopadły do wektorów

$$\nabla g_1(x_0),\ \nabla g_2(x_0),\ \ldots,\ \nabla g_k(x_0).$$

Wektor $\sigma'(0)$ jest styczny do powierzchni S w punkcie x_0. Wymiar przestrzeni liniowej V_1 rozpiętej przez wszystkie wektory styczne $\sigma'(0)$ wynosi $n - k$. Z kolei wymiar przestrzeni V_2 rozpiętej przez wektory $\nabla g_1(x_0)$, $\nabla g_2(x_0)$, $\ldots$, $\nabla g_k(x_0)$ wynosi k, bo gradienty są liniowo niezależne. Ale V_1 i V_2 są do siebie prostopadłe, zatem $V_1^\perp = V_2$. Rozważmy funkcję $t \mapsto f(\sigma(t))$. Funkcja ta osiąga ekstremum dla $t = 0$. Czyli

$$0 = \frac{d}{dt} f(\sigma(t))\Big|_{t=0} = \nabla f(x_0) \circ \sigma'(0),$$

dla dowolnej wyżej opisanej krzywej σ. Zatem $\nabla f(x_0) \in V_1^\perp = V_2$. □

5.8 Twierdzenie o funkcji uwikłanej

Z teorii funkcji jednej zmiennej $y = f(x)$ wiemy, że jeśli f jest klasy C^1 oraz $f'(x_0) \neq 0$, to równanie $f(x) = y$ dla y w pobliżu $y_0 = f(x_0)$ ma jednoznaczne rozwiązanie $x = f^{-1}(y)$ leżące w pobliżu x_0. Rzeczywiście, rozważmy przypadek $f'(x_0) > 0$. Zatem $f'(x) > 0$ dla x w pewnym przedziale wokół x_0, np. w $(x_0-\delta, x_0+\delta)$. Wtedy $f(x)$ jest ściśle rosnąca w $(x_0-\delta, x_0+\delta)$. Zatem posiada funkcję odwrotną $x = g(y)$. Proces odwracania jest ważny również dla funkcji wielu zmiennych.

Przykład. Współrzędne biegunowe na płaszczyźnie punktu (x, y) wyrażają się wzorami $x = r\cos\theta$, $y = r\sin\theta$. Dla $x, y > 0$ mamy

$$r = \sqrt{x^2 + y^2}, \qquad \theta = \operatorname{arctg} \frac{y}{x}.$$

Rozważmy równanie $F(x,y,z) = 0$. Przypuśćmy, że $F(x_0,y_0,z_0) = 0$. Interesuje nas obliczenie zmiennej z z równania w pobliżu (x_0,y_0,z_0). Tzn. chcemy, aby dla (x,y) blisko (x_0,y_0) znaleźć z blisko z_0 tak, aby $F(x,y,z) = 0$. Np. niech $F(x,y,z) = x^2+y^2+z^2-1$ oraz $F(0,0,1) = 0$. Wtedy

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

jest rozwiązaniem równania. Podobnie dla $F(0,0,-1)$ rozwiązaniem jest

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

Z kolei dla $F(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0)$ mamy dwa rozwiązania

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

lub brak rozwiązań, jeśli $x^2+y^2 > 1$.

Twierdzenie 5.37. *Załóżmy, że funkcja $F : \mathbb{R}^{n+1} \to \mathbb{R}$ jest klasy C^1. Będziemy stosować oznaczenie $(x,z) \in \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$. Załóżmy, że*

$$F(x_0,z_0) = 0, \quad \text{oraz} \quad \frac{\partial F}{\partial z}(x_0,z_0) \neq 0.$$

Wtedy równanie $F(x,z) = 0$ ma jednoznaczne rozwiązanie w pobliżu (x_0,z_0). Tzn. istnieje kula otwarta $U \subset \mathbb{R}^n$ o środku w x_0 oraz przedział otwarty V wokół z_0 takie, że dla dowolnego wyboru $x \in U$ istnieje jedyne rozwiązanie $z \in V$ takie, że $F(x,z) = 0$. Ponadto funkcja $z = g(x)$ jest klasy C^1 na U.

Przykład. Dla funkcji $F(x,y,z) = x^2+y^2+z^2-1$ mamy

$$\frac{\partial F}{\partial z}(0,0,\pm 1) = \pm 2, \qquad \frac{\partial F}{\partial z}\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0\right) = 0.$$

Uwaga. Przyjmijmy, że funkcja $F(x_1,x_2,\ldots,x_n,z)$ jest liniowa. Możemy obliczyć zmienną z z równania $F(x,z) = 0$, o ile współczynnik przy zmiennej z jest niezerowy. Tzn. $\dfrac{\partial F}{\partial z} \neq 0$.

Twierdzenie nabiera istotnego znaczenia, gdy nie jesteśmy w stanie obliczyć $z = g(x)$ jawnym wzorem. Okazuje się jednak, że wiele informacji o funkcji g można uzyskać mimo braku jawnego wzoru. Wiemy, że $z_0 = g(x_0)$ oraz $F(x, g(x)) = 0$ dla $x \in U$. Zatem

$$0 = \frac{\partial}{\partial x_i} F(x, g(x)) = \frac{\partial F}{\partial x_i}(x, g(x)) + \frac{\partial F}{\partial z}(x, g(x)) \frac{\partial g}{\partial x_i}(x).$$

Otrzymujemy

$$\frac{\partial g}{\partial x_i}(x) = -\frac{\frac{\partial F}{\partial x_i}(x, g(x))}{\frac{\partial F}{\partial z}(x, g(x))}.$$

Z założenia $\frac{\partial F}{\partial z}(x_0, z_0) \neq 0$, zatem $\frac{\partial F}{\partial z}(x, g(x)) \neq 0$, dla x w pobliżu x_0, bo funkcje F i g są klasy C^1. Podstawiamy $x = x_0$, aby otrzymać

$$\frac{\partial g}{\partial x_i}(x_0) = -\frac{\frac{\partial F}{\partial x_i}(x_0, z_0)}{\frac{\partial F}{\partial z}(x_0, z_0)}. \tag{5.6}$$

Przykłady.

(a) Rozważamy równanie $F(x, y, z) = xy + z + 3xz^5 = 4$ i rozwiązanie $(1, 0, 1)$. Wtedy

$$\begin{aligned} \frac{\partial F}{\partial z}(1, 0, 1) &= 1 + 15xz^4\Big|_{(1,0,1)} = 16, \\ \frac{\partial F}{\partial x}(1, 0, 1) &= y + 3z^5\Big|_{(1,0,1)} = 3, \\ \frac{\partial F}{\partial y}(1, 0, 1) &= x\Big|_{(1,0,1)} = 1. \end{aligned}$$

Na podstawie wzoru (5.6) otrzymujemy

$$\frac{\partial g}{\partial x}(1, 0) = -\frac{3}{16}, \quad \frac{\partial g}{\partial y}(1, 0) = -\frac{1}{16}.$$

(b) Niech $F(x, y, z) := x^3 + 3y^2 + 8xz^2 - 3yz^3 = 1$. W pobliżu jakich punktów powierzchnia zadana równaniem może być przedstawiona jako wykres funkcji $z = g(x, y)$? Obliczamy

$$\frac{\partial F}{\partial z} = 16xz - 9yz^2 \neq 0.$$

Zatem muszą być spełnione warunki $z \neq 0$ oraz $16x - 9yz \neq 0$.

Jeśli chcemy obliczyć $x = h(y, z)$, to

$$\frac{\partial F}{\partial x} = 3x^2 + 8z^2 \neq 0.$$

Wystarczy zatem, aby $x \neq 0$ lub $z \neq 0$.

Wniosek 5.38. *Jeśli funkcja $f(x_1, x_2, \ldots, x_n)$ spełnia $f(a_1, a_2, \ldots, a_n) = 0$ oraz $\nabla f(a_1, a_2, \ldots, a_n) \neq 0$, to z równania*

$$f(x_1, x_2, \ldots, x_n) = 0$$

można obliczyć jedną zmienną względem pozostałych w pobliżu $(a_1, a_2, \ldots, a_n)$.

Dowód. Oznaczmy $a = (a_1, a_2, \ldots, a_n)$. Z założenia $\dfrac{\partial f}{\partial x_i}(a) \neq 0$ dla pewnej wartości i. Przez zmianę numeracji możemy przyjąć, że $\dfrac{\partial f}{\partial x_n}(a) \neq 0$. Funkcja f zależy od $x_1, x_2, \ldots, x_{n-1}$ oraz od $z = x_n$. Z poprzedniego twierdzenia z równania

$$f(x_1, x_2, \ldots, x_{n-1}, z) = 0$$

można obliczyć z w zależności od $x_1, x_2, \ldots, x_{n-1}$. □

Dowód twierdzenia. Z założenia mamy $\dfrac{\partial F}{\partial z}(x_0, z_0) \neq 0$. Rozważymy przypadek $\dfrac{\partial F}{\partial z}(x_0, z_0) > 0$. Z ciągłości pochodnych cząstkowych można znaleźć liczby dodatnie a i b takie, że jeśli $\|x - x_0\| \leqslant a$ oraz $|z - z_0| \leqslant a$, to $\dfrac{\partial F}{\partial z}(x, z) > b$. Zbiór określony warunkami $\|x - x_0\| \leqslant a$, $|z - z_0| \leqslant a$ jest domknięty i ograniczony, zatem z ciągłości pochodnych cząstkowych mamy

$$\left|\frac{\partial F}{\partial x_i}(x, z)\right| \leqslant M, \quad \left|\frac{\partial F}{\partial z}(x, z)\right| \leqslant M \quad \text{dla } \|x - x_0\| \leqslant a,\ |z - z_0| \leqslant a.$$

Lemat 5.39. *Dla funkcji $f : \mathbb{R}^n \to \mathbb{R}$ klasy C^1 mamy*

$$f(x) - f(x_0) = \nabla f(x_0 + \theta(x - x_0)) \circ (x - x_0)$$

dla pewnej liczby θ, $0 < \theta < 1$.

Dowód lematu. Określamy funkcję $g(t) = f(x_0 + t(x - x_0))$ przy ustalonych punktach x i x_0. Wtedy z twierdzenia Lagrange'a otrzymujemy

$$f(x) - f(x_0) = g(1) - g(0) = g'(\theta) = \nabla f(x_0 + \theta(x - x_0)) \circ (x - x_0).$$

$\square$

Z lematu mamy

$$\begin{aligned} F(x, z) &= F(x, z) - F(x_0, z_0) \\ &= \nabla F(x_0 + \theta(x - x_0), z_0 + \theta(z - z_0)) \circ (x - x_0, z - z_0) \end{aligned}$$

Oznaczmy

$$x_\theta = x_0 + \theta(x - x_0),\ z_\theta = z_0 + \theta(z - z_0), \quad \nabla_x F = \left(\frac{\partial F}{\partial x_1}, \ldots, \frac{\partial F}{\partial x_n}\right).$$

Wtedy

$$F(x, z) = \nabla_x F(x_\theta, z_\theta) \circ (x - x_0) + \frac{\partial F}{\partial z}(x_\theta, z_\theta)(z - z_0). \tag{5.7}$$

Dla $\|x - x_0\| \leqslant a$, $|z - z_0| \leqslant a$ mamy $\|x_\theta - x_0\| \leqslant a$ oraz $|z_\theta - z_0| \leqslant a$. Stąd otrzymujemy

$$\|\nabla_x F(x_\theta, z_\theta)\| \leqslant M\sqrt{n}.$$

Zatem

$$|\nabla_x F(x_\theta, z_\theta) \circ (x - x_0)| \leqslant M\sqrt{n}\|x - x_0\|. \tag{5.8}$$

Rozważamy tylko $z = z_0 \pm a$. Wtedy

$$\left|\frac{\partial F}{\partial z}(x_\theta, z_\theta)(z - z_0)\right| = \left|\frac{\partial F}{\partial z}(x_\theta, z_\theta)(\pm a)\right| > ab.$$

Z (5.7) otrzymujemy

$$\left|F(x, z) - \frac{\partial F}{\partial z}(x_\theta, z_\theta)(z - z_0)\right| = |\nabla_x F(x_\theta, z_\theta) \circ (x - x_0)| \leqslant M\sqrt{n}\|x - x_0\|.$$

Wybierzmy liczbę $0 < \delta \leqslant a$ taką, że $M\sqrt{n}\delta < ab$. Niech $\|x - x_0\| < \delta$. Wtedy

$$\left|F(x, z_0 \pm a) - \frac{\partial F}{\partial z}(x_\theta, z_\theta)(\pm a)\right| \ < \ M\sqrt{n}\delta \ < \ ab \ < \ \left|\frac{\partial F}{\partial z}(x_\theta, z_\theta)(\pm a)\right|.$$

Lemat 5.40. *Jeśli* $|u-v|<|v|$, *to liczby* u *i* v *mają ten sam znak.*

Z lematu wynika, że $F(x,z_0+a)>0$ oraz $F(x,z_0-a)<0$. Z własności Darboux mamy $F(x,z)=0$ dla pewnej liczby z z przedziału (z_0-a,z_0+a). Takie rozwiązanie jest jedyne w tym przedziale, bo funkcja

$$(z_0-a,z_0+a)\ni z\mapsto F(x,z)$$

jest ściśle rosnąca, co wynika z dodatniości pochodnej cząstkowej względem z. Reasumując pokazaliśmy, że dla $\|x-x_0\|<\delta$ istnieje jedyne rozwiązanie z w przedziale (z_0-a,z_0+a) spełniające $F(x,z)=0$. W ten sposób otrzymujemy funkcję $z=g(x)$. Sprawdzimy, że g jest funkcją ciągłą. Załóżmy nie wprost, że $x_m\to x$, ale $g(x_m)\not\to g(x)$. Ciąg $g(x_m)$ jest ograniczony. Istnieje zatem podciąg $g(x_{m_k})$ zbieżny do liczby $\tilde{z}\neq g(x)$ z przedziału $[z_0-a,z_0+a]$. Mamy

$$0=F(x_{m_k},g(x_{m_k}))\xrightarrow[k]{}F(x,\tilde{z}).$$

Stąd $F(x,\tilde{z})=0$. Ale $\tilde{z}\neq z_0\pm a$, bo $F(x,z_0\pm a)\neq 0$. Czyli $\tilde{z}$ leży w przedziale (z_0-a,z_0+a). Ale $F(x,g(x))=0$, więc otrzymujemy sprzeczność z jednoznacznością rozwiązania.

Zbadamy różniczkowalność funkcji $g(x)$. Przyjmujemy $x=x_0+he_i$. Wtedy

$$\nabla_x F(x_\theta,z_\theta)\circ(x-x_0)=\frac{\partial F}{\partial x_i}(x_\theta,z_\theta)h.$$

We wzorze (5.7) podstawiamy $z=g(x)$. Lewa strona wzoru zeruje się. Otrzymujemy więc

$$\frac{g(x_0+he_i)-g(x_0)}{h}=-\frac{\frac{\partial F}{\partial x_i}(x_\theta,z_\theta)}{\frac{\partial F}{\partial z}(x_\theta,z_\theta)}.$$

Mamy

$$x_\theta=x_0+\theta(x-x_0)=x_0+\theta he_i\xrightarrow[h\to 0]{}x_0,$$
$$z_\theta=z_0+\theta(z-z_0)=g(x_0)+\theta[g(x_0+he_i)-g(x_0)]\xrightarrow[h\to 0]{}g(x_0)=z_0,$$

bo g jest ciągła. Zatem

$$\frac{\partial g}{\partial x_i}(x_0)=-\frac{\frac{\partial F}{\partial x_i}(x_0,z_0)}{\frac{\partial F}{\partial z}(x_0,z_0)}.$$

Ten sam dowód daje

$$\frac{\partial g}{\partial x_i}(x) = -\left.\frac{\frac{\partial F}{\partial x_i}(x,z)}{\frac{\partial F}{\partial z}(x,z)}\right|_{z=g(x)}.$$

Widzimy, że pochodne cząstkowe funkcji g są ciągłe, zatem g jest funkcją różniczkowalną. □

Uwaga. Jeśli wiemy, że funkcja $z = g(x)$ jest różniczkowalna, to jej pochodne cząstkowe można obliczyć stosując różniczkowanie niejawne. Mamy $F(x, g(x)) \equiv 0$. Różniczkujemy względem x_i aby otrzymać

$$\frac{\partial F}{\partial x_i}(x, g(x)) + \frac{\partial F}{\partial z}(x, g(x))\,\frac{\partial g}{\partial x_i}(x) = 0.$$

Chcemy obliczyć wielkości $z_1, z_2, \ldots, z_m$ z równań

$$\begin{array}{rcl} F_1(x_1, x_2, \ldots, x_n; z_1, z_2, \ldots, z_m) & = & 0, \\ F_2(x_1, x_2, \ldots, x_n; z_1, z_2, \ldots, z_m) & = & 0, \\ & \vdots & \\ F_m(x_1, x_2, \ldots, x_n; z_1, z_2, \ldots, z_m) & = & 0, \end{array} \tag{5.9}$$

i otrzymać rozwiązanie w postaci

$$\begin{array}{rcl} z_1 & = & g_1(x_1, x_2, \ldots, x_n), \\ z_2 & = & g_2(x_1, x_2, \ldots, x_n), \\ & \vdots & \\ z_m & = & g_m(x_1, x_2, \ldots, x_n). \end{array} \tag{5.10}$$

Będziemy stosować zapis

$$x = (x_1, x_2, \ldots, x_n), \quad z = (z_1, z_2, \ldots, z_m).$$

Załóżmy, że $(x_0, z_0) \in \mathbb{R}^n \times \mathbb{R}^m$ jest rozwiązaniem układu. Rozważamy wyznacznik

$$\Delta = \begin{vmatrix} \frac{\partial F_1}{\partial z_1}(x_0; z_0) & \frac{\partial F_1}{\partial z_2}(x_0; z_0) & \ldots & \frac{\partial F_1}{\partial z_m}(x_0; z_0) \\ \frac{\partial F_2}{\partial z_1}(x_0; z_0) & \frac{\partial F_2}{\partial z_2}(x_0; z_0) & \ldots & \frac{\partial F_2}{\partial z_m}(x_0; z_0) \\ \vdots & \vdots & \ldots & \vdots \\ \frac{\partial F_m}{\partial z_1}(x_0; z_0) & \frac{\partial F_m}{\partial z_2}(x_0; z_0) & \ldots & \frac{\partial F_m}{\partial z_m}(x_0; z_0) \end{vmatrix}$$

Twierdzenie 5.41 (o funkcji uwikłanej)**.** *Załóżmy, że funkcje F_1, F_2, $\dots$ F_m są klasy C^1. Niech punkt $(x_0; z_0)$ będzie rozwiązaniem układu równań (5.9) oraz $\Delta \neq 0$. Wtedy istnieją liczby $\delta > 0$ i $a > 0$ takie, że dla $\|x - x_0\| < \delta$ istnieje jedyny z spełniający $\|z - z_0\| < a$ taki, że (x, z) jest rozwiązaniem układu równań (5.9). Ponadto funkcje z (5.10) są klasy C^1.*

Przykład. Czy w pobliżu $(x, y; u, v) = (1, 1; 1, 1)$ można obliczyć u i v z równań

$$\begin{aligned} xu + yuv^2 &= 2, \\ xu^3 + y^2v^4 &= 2 \end{aligned}$$

jako funkcje zmiennych x i y ? Przyjmujemy

$$F_1(x, y; u, v) = xu + yuv^2 - 2,$$
$$F_2(x, y; u, v) = xu^3 + y^2v^4 - 2.$$

Mamy

$$\Delta = \left.\begin{vmatrix} x + yv^2 & 2yuv \\ 3xu^2 & 4y^2v^3 \end{vmatrix}\right|_{x=1,y=1,u=1,v=1} = \begin{vmatrix} 2 & 2 \\ 3 & 4 \end{vmatrix} = 2 \neq 0.$$

Chcemy obliczyć $\dfrac{\partial u}{\partial x}(1, 1)$ i $\dfrac{\partial v}{\partial x}(1, 1)$. Stosujemy różniczkowanie niejawne. Otrzymujemy

$$\begin{aligned} u + x\,\frac{\partial u}{\partial x} + yv^2\,\frac{\partial u}{\partial x} + 2yuv\,\frac{\partial v}{\partial x} &= 0, \\ u^3 + 3xu^2\,\frac{\partial u}{\partial x} + 4y^2v^3\,\frac{\partial v}{\partial x} &= 0. \end{aligned}$$

Podstawiamy $x = 1, y = 1, v = 1, u = 1$. Po uproszczeniu otrzymujemy

$$\begin{aligned} 2\,\frac{\partial u}{\partial x} + 2\,\frac{\partial v}{\partial x} &= -1, \\ 3\,\frac{\partial u}{\partial x} + 4\,\frac{\partial v}{\partial x} &= -1. \end{aligned}$$

Zatem

$$\frac{\partial u}{\partial x}(1, 1) = \frac{\begin{vmatrix} -1 & 2 \\ -1 & 4 \end{vmatrix}}{\begin{vmatrix} 2 & 2 \\ 3 & 4 \end{vmatrix}} = -1, \qquad \frac{\partial v}{\partial x}(1, 1) = \frac{\begin{vmatrix} 2 & -1 \\ 3 & -1 \end{vmatrix}}{\begin{vmatrix} 2 & 2 \\ 3 & 4 \end{vmatrix}} = \frac{1}{2}.$$

Dowód. Przyjmijmy oznaczenia $a = x_0$ i $b = z_0$. Wyznacznik Δ w punkcie (a, b) nie znika, zatem jedna z liczb $\dfrac{\partial F_j}{\partial z_m}(a; b)$ w ostatniej kolumnie jest niezerowa. Możemy przyjąć, że $\dfrac{\partial F_m}{\partial z_m}(a; b) \neq 0$, ewentualnie zmieniając numerację równań. Na podstawie Twierdzenia 5.37 możemy z równania

$$F_m(x_1, x_2, \ldots, x_n; z_1, z_2, \ldots, z_m) = 0$$

obliczyć

$$z_m = g(x_1, x_2, \ldots, x_n; z_1, z_2, \ldots, z_{m-1}) = g(x, \tilde{z}).$$

Ponadto $g(a, \tilde{b}) = b_m$. Po podstawieniu $z_m = g(x, \tilde{z})$ ostatnie równanie staje się tożsamością

$$F_m(x; \tilde{z}, g(x, \tilde{z})) \equiv 0. \tag{5.11}$$

Podstawiamy obliczoną wartość z_m do pierwszych $m-1$ równań. Otrzymamy układ

$$\begin{aligned} H_1(x, \tilde{z}) \quad &:= \quad F_1(x; \tilde{z}, g(x, \tilde{z})) = 0, \\ H_2(x, \tilde{z}) \quad &:= \quad F_2(x; \tilde{z}, g(x, \tilde{z})) = 0, \\ &\vdots \\ H_{m-1}(x, \tilde{z}) \quad &:= \quad F_{m-1}(x; \tilde{z}, g(x, \tilde{z})) = 0. \end{aligned}$$

Chcemy obliczyć $z_1, z_2, \ldots, z_{m-1}$ z nowego układu równań. Mamy rozwiązanie $x = a$, $\tilde{z} = \tilde{b}$, bo wtedy $g(a; \tilde{b}) = b_m$. Sprawdzamy, czy założenia twierdzenia o funkcji uwikłanej są spełnione dla nowego układu. Obliczamy

$$\frac{\partial H_i}{\partial z_j} = \frac{\partial F_i}{\partial z_j} + \frac{\partial F_i}{\partial z_m} \frac{\partial g}{\partial z_j}, \qquad j = 1, 2, \ldots, m-1. \tag{5.12}$$

Różniczkujemy tożsamość (5.11) względem z_j, aby otrzymać

$$\frac{\partial F_m}{\partial z_j} + \frac{\partial F_m}{\partial z_m} \frac{\partial g}{\partial z_j} = 0, \qquad j = 1, 2, \ldots, m-1. \tag{5.13}$$

Rozważamy wyznacznik

$$\Delta(x; \tilde{z}) := \det\left(\frac{\partial F_i}{\partial z_j}(x; \tilde{z}, g(x, \tilde{z})) \right)_{i,j=1,2,\ldots,m}.$$

Wiemy, że dla $x = a$, $\tilde{z} = \tilde{b}$ mamy $g(a, \tilde{b}) = b_m$. Zatem

$$\Delta(a; \tilde{b}) \neq 0.$$

W wyznaczniku $\Delta(x; \tilde{z})$ mnożymy ostatnią kolumnę przez liczbę $\dfrac{\partial g}{\partial z_j}(x; \tilde{z})$ i dodajemy do j-tej kolumny, dla wszystkich $j = 1, 2, \ldots, m-1$. Otrzymamy

$$\Delta(x, \tilde{z}) = \begin{vmatrix} \frac{\partial F_1}{\partial z_1} + \frac{\partial F_1}{\partial z_m}\frac{\partial g}{\partial z_1} & \ldots & \frac{\partial F_1}{\partial z_{m-1}} + \frac{\partial F_1}{\partial z_m}\frac{\partial g}{\partial z_{m-1}} & \frac{\partial F_1}{\partial z_m} \\ \vdots & & \vdots & \vdots \\ \frac{\partial F_{m-1}}{\partial z_1} + \frac{\partial F_{m-1}}{\partial z_m}\frac{\partial g}{\partial z_1} & \ldots & \frac{\partial F_{m-1}}{\partial z_{m-1}} + \frac{\partial F_{m-1}}{\partial z_m}\frac{\partial g}{\partial z_{m-1}} & \frac{\partial F_{m-1}}{\partial z_m} \\ \frac{\partial F_m}{\partial z_1} + \frac{\partial F_m}{\partial z_m}\frac{\partial g}{\partial z_1} & \ldots & \frac{\partial F_m}{\partial z_{m-1}} + \frac{\partial F_m}{\partial z_m}\frac{\partial g}{\partial z_{m-1}} & \frac{\partial F_m}{\partial z_m} \end{vmatrix}$$

Z (5.13) ostatni wiersz zeruje się poza ostatnim elementem. Z (5.12) otrzymujemy więc

$$\Delta(x, \tilde{z}) = \frac{\partial F_m}{\partial z_m} \begin{vmatrix} \frac{\partial H_1}{\partial z_1}(x; \tilde{z}) & \ldots & \frac{\partial H_1}{\partial z_{m-1}}(x; \tilde{z}) \\ \vdots & & \vdots \\ \frac{\partial H_{m-1}}{\partial z_1}(x; \tilde{z}) & \ldots & \frac{\partial H_{m-1}}{\partial z_{m-1}}(x; \tilde{z}) \end{vmatrix},$$

gdzie $\dfrac{\partial F_m}{\partial z_m}$ jest obliczone w $(x; \tilde{z}, g(x, \tilde{z}))$. Ponieważ $\Delta(a, \tilde{b}) \neq 0$, to wyznacznik nowego układu dla $x = a$, $\tilde{z} = \tilde{b}$ jest różny od zera.

Możemy zatem kontynuować obliczając kolejne zmienne

$$\begin{aligned} z_m &= g_1(x_1, \ldots, x_n; z_1, \ldots, z_{m-1}), \\ z_{m-1} &= g_2(x_1, \ldots, x_n; z_1, \ldots, z_{m-2}), \\ &\vdots \\ z_2 &= g_{m-1}(x_1, \ldots, x_n; z_1), \\ z_1 &= g_m(x_1, \ldots, x_n). \end{aligned}$$

Wykonujemy podstawienie wstecz, aby ostatecznie obliczyć zmienne $z_1, z_2, \ldots, z_m$ za pomocą $x_1, x_2, \ldots, x_n$. □

Szczególnym przypadkiem twierdzenia o funkcji uwikłanej jest twierdzenie o funkcji odwrotnej. Chcemy z układu równań

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= y_1, \\ f_2(x_1, x_2, \ldots, x_n) &= y_2, \\ &\vdots \\ f_n(x_1, x_2, \ldots, x_n) &= y_n, \end{aligned} \tag{5.14}$$

obliczyć $x_1, x_2, \ldots, x_n$, jako funkcje od $y_1, y_2, \ldots, y_n$. Załóżmy, że $x = a$ i $y = b$ jest rozwiązaniem układu. Rozważamy

$$\begin{aligned} F_1(x_1, \ldots, x_n; y_1, \ldots, y_n) &= f_1(x_1, \ldots, x_n) - y_1 = 0, \\ F_2(x_1, \ldots, x_n; y_1, \ldots, y_n) &= f_2(x_1, \ldots, x_n) - y_2 = 0, \\ &\vdots \\ F_n(x_1, \ldots, x_n; y_1, \ldots, y_n) &= f_n(x_1, \ldots, x_n) - y_n = 0. \end{aligned}$$

Z twierdzenia o funkcji uwikłanej badamy wyznacznik

$$\Delta = \left.\begin{vmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \end{vmatrix}\right|_{\substack{x=a \\ y=b}} = \left.\begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{vmatrix}\right|_{x=a}$$

Wyznacznik

$$\begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{vmatrix}$$

nazywamy jakobianem odwzorowań $f_1, f_2, \ldots, f_n$.

Twierdzenie 5.42 (o funkcji odwrotnej)**.** *Niech $U \subset$ będzie otwartym podzbiorem przestrzenia $\mathbb{R}^n$. Rozważamy funkcje $f_1, f_2, \ldots, f_n$ klasy C^1 na U. Załóżmy, że układ równań (5.14) ma rozwiązanie $x = a$, $y = b$ dla $a \in U$. Jeśli*

$$\Delta = \det\left(\frac{\partial f_i}{\partial x_j}(a)\right) \neq 0,$$

to układ ma jednoznaczne rozwiązanie dla y w pobliżu b i x w pobliżu a. Tzn. istnieje liczba $\delta > 0$ taka, że dla $\|y - b\| < \delta$ istnieje jedyny punkt $x \in U$ taki, *że* $\|x-a\| < \delta$ *oraz* x *i* y *są rozwiązaniem układu (5.14). Ponadto funkcje*

$$\begin{aligned} x_1 &= g_1(y_1, y_2, \ldots, y_n), \\ x_2 &= g_2(y_1, y_2, \ldots, y_n), \\ &\vdots \\ x_n &= g_n(y_1, y_2, \ldots, y_n) \end{aligned}$$

są klasy C^1.

Przykład. Rozważmy układ równań

$$\begin{aligned} \frac{x^4+y^4}{x} &= u, \\ \sin x + \cos y &= v. \end{aligned}$$

W pobliżu jakich punktów możemy obliczyć x i y względem u i v ? Obliczamy jakobian

$$\Delta = \begin{vmatrix} 3x^2 - \frac{y^4}{x^2} & \frac{4y^3}{x} \\ \cos x & -\sin y \end{vmatrix}$$

Powinien być spełniony warunek $\Delta \neq 0$. Wyznacznik jest niezerowy dla $x = \frac{\pi}{2}$ i $y = \frac{\pi}{2}$. Zatem można rozwiązać układ w pobliżu $u = \frac{\pi^3}{4}$ i $v = 1$. Rozwiązania będą leżały w pobliżu $x = \frac{\pi}{2}$, $y = \frac{\pi}{2}$.

Twierdzenie o funkcji odwrotnej można sformułować w postaci zbliżonej w zapisie do twierdzenia dla jednej zmiennej. Dla funkcji $f_1, f_2, \ldots, f_n : U \to \mathbb{R}$ tworzymy funkcję $f : U \to \mathbb{R}^n$ wzorem

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix} \in \mathbb{R}^n, \qquad x = (x_1, x_2, \ldots, x_n).$$

Wtedy układ równań w twierdzenia o funkcji odwrotnej ma postać $f(x) = y$, gdzie

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Zauważmy, że $\Delta = \det(Df(a)) \neq 0$. Załóżmy, że $f(a) = b$ dla $a \in U$. Wtedy dla y w pobliżu b istnieje jedyne rozwiązanie x w pobliżu a. Ponadto $x = g(y)$, gdzie g jest klasy C^1. Tzn. g jest funkcją odwrotną do funkcji f. Obliczmy $Dg(y)$. Mamy

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

Różniczkujemy obie strony. Wtedy

$$Dg(f(x))\, Df(x) = I,$$

czyli

$$Dg(y) = (Df(x))^{-1}, \quad y = f(x).$$

Dla funkcji jednej zmiennej wzory mają postać $y = f(x)$, $x = g(y)$ oraz

$$g'(y) = \frac{1}{f'(x)}.$$

Przykład. W pobliżu jakich punktów funkcja $f : \mathbb{R}^2 \to \mathbb{R}^2$

$$f(x, y) = (x - y, x^5 + y^5)$$

jest odwracalna ?

$$\det Df(x, y) = \begin{vmatrix} 1 & -1 \\ 5x^4 & 5y^4 \end{vmatrix} = 5(x^4 + y^4).$$

Funkcja jest odwracalna poza punktem $(0, 0)$.

5.9 Różniczka

Rozważmy funkcję różniczkowalną $f : \mathbb{R}^n \to \mathbb{R}$. Wtedy

$$f(x) \approx f(a) + \nabla f(a) \circ (x - a),$$

gdy $x, a \in \mathbb{R}^n$ są blisko siebie. Istotnie wiemy, że

$$\lim_{x \to a} \frac{|f(x) - f(a) - \nabla f(a) \circ (x - a)|}{\|x - a\|} = 0.$$

Wyrażenie $\nabla f(a) \circ (x - a)$ nazywamy różniczką odpowiadającą przyrostowi $x-a$. Podobnie dla $f : \mathbb{R}^n \to \mathbb{R}^n$ możemy zapisać $f(x) \approx f(a) + Df(a)\,(x-a)$. Oznaczmy $Df(a) = A$. Dla x blisko a mamy $y = f(x) \approx f(a) + A(x - a)$. Załóżmy, że $y = f(a) + A(x-a)$. Wtedy $x = a + A^{-1}(y-b)$. W rzeczywistości mamy $x \approx a + A^{-1}(y - b)$.

6 Całki podwójne

Niech R będzie prostokątem $[a,b]\times[c,d]$. Rozważamy nieujemną funkcję $f(x,y)$ określoną na R. Wykres ma postać powierzchni leżącej nad R. Powierzchnia $z=f(x,y)$ oraz cztery pionowe płaszczyzny $x=a$, $x=b$, $y=c$ i $y=d$ ograniczają obszar trójwymiarowy B. Chcemy obliczyć objętość tego obszaru. Załóżmy, że całka podwójna została określona tak, aby

$$\iint\limits_R f(x,y)\,dx\,dy=\int\limits_a^b\int\limits_c^d f(x,y)\,dx\,dy=\mathrm{vol}(B).$$

Przykłady.

(a) $f(x,y)=k$, $k\geqslant 0$. Obszar jest prostopadłościanem o wysokości k.

$$\int\limits_a^b\int\limits_c^d k\,dx\,dy=k(b-a)(d-c).$$

(b) $f(x,y)=1-x$, $0\leqslant x\leqslant 1$, $0\leqslant y\leqslant 1$. Obszar jest połową sześcianu o boku 1.

$$\int\limits_0^1\int\limits_0^1(1-x)\,dx\,dy=\frac{1}{2}.$$

6.1 Zasada Cavalieriego

Przy bardziej złożonych funkcjach $f(x,y)$ możemy zastosować zasadę Cavalieriego. Załóżmy, że bryła ma własność, że pola przekroju płaszczyznami równoległymi do ustalonej płaszczyzny, w odległości x od tej płaszczyzny, wynoszą $A(x)$. Bryła mieści się pomiędzy płaszczyznami $x=a$ i $x=b$. Wtedy

$$V=\int\limits_a^b A(x)\,dx.$$

Rozważmy nieujemną funkcję $f(x,y)$ na $[a,b]\times[c,d]$. Pole przekroju płaszczyzną pionową $x=x_0$ wynosi

$$A(x_0)=\int\limits_c^d f(x_0,y)\,dy.$$

Zatem objętość bryły wynosi

$$V = \int_a^b \left(\int_c^d f(x,y)\,dy \right) dx.$$

Można też zastosować cięcia płaszczyznami równoległymi do płaszczyzny pionowej $y = 0$. Wtedy

$$V = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy.$$

Przykład. $f(x,y) = x^2 + y^2$, $-1 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$.

$$V = \int_{-1}^{1} \left(\int_0^1 (x^2 + y^2)\,dy \right) dx = \int_{-1}^{1} \left(x^2 + \frac{1}{3} \right) dx = \frac{2}{3} + \frac{2}{3} = \frac{4}{3}.$$

6.2 Ścisłe określenie całki podwójnej Riemanna

Podziałem prostokąta $R = [a,b] \times [c,d]$ nazywamy parę $\mathcal{P} = (\mathcal{P}_1, \mathcal{P}_2)$, gdzie $\mathcal{P}_1$ jest podziałem przedziału $[a,b]$, a $\mathcal{P}_2$ podziałem przedziału $[c,d]$:

$$\mathcal{P}_1 = \{x_0, x_1, \ldots, x_n\}, \qquad \mathcal{P}_2 = \{y_0, y_1, \ldots, y_m\}.$$

Podprzedziałem nazywamy każdy z prostokątów

$$S_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j].$$

Rozważamy funkcję $f(x,y)$ określoną na R. Dla podprzedziału S niech

$$m_S(f) = \inf_{(x,y)\in S} f(x,y), \qquad M_S(f) = \sup_{(x,y)\in S} f(x,y).$$

Symbolem ΔS oznaczamy pole powierzchni prostokąta S. Sumy dolne i górne są zdefiniowane wzorami

$$L(\mathcal{P}, f) = \sum_{S\in\mathcal{P}} m_S(f)\Delta S, \qquad U(\mathcal{P}, f) = \sum_{S\in\mathcal{P}} M_S(f)\Delta S.$$

Uwaga. Jeśli $f(x,y) \geqslant 0$, to objętość obszaru pod wykresem mieści pomiędzy liczbami $L(\mathcal{P}, f)$ i $U(\mathcal{P}, f)$.

Podział $\mathcal{P}' = (\mathcal{P}_1, \mathcal{P}_2)$ nazywamy rozdrobnieniem podziału $\mathcal{P} = (\mathcal{P}_1, \mathcal{P}_2)$, jeśli $\mathcal{P}_1'$ jest rozdrobnieniem $\mathcal{P}$, a $\mathcal{P}_2'$ rozdrobnieniem $\mathcal{P}_2$.

Lemat 6.1. *Jeśli $\mathcal{P}'$ jest rozdrobnieniem $\mathcal{P}$, to*

$$L(\mathcal{P}, f) \leqslant L(\mathcal{P}', f), \qquad U(\mathcal{P}, f) \geqslant U(\mathcal{P}', f).$$

Określamy całki dolną i górną wzorami

$$\underline{\iint_R} f(x,y)\,dx\,dy = \sup_{\mathcal{P}} L(\mathcal{P}, f), \qquad \overline{\iint_R} f(x,y)\,dx\,dy = \inf_{\mathcal{P}} U(\mathcal{P}, f).$$

Mówimy, że funkcja $f(x,y)$ jest całkowalna jeśli

$$\underline{\iint_R} f(x,y)\,dx\,dy = \overline{\iint_R} f(x,y)\,dx\,dy.$$

Twierdzenie 6.2. *Funkcja ograniczona $f(x,y)$ na prostokącie R jest całkowalna wtedy i tylko wtedy, gdy dla dowolnej liczby $\varepsilon > 0$ można znaleźć podział $\mathcal{P}$ spełniający*

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

Uwaga. Dowód jest bardzo podobny do przypadku jednej zmiennej. Implikacja $\Leftarrow$ jest użyteczna.

Lemat 6.3. *Każda funkcja ciągła $f(x)$, o wartościach liczbowych, określona na zwartym podzbiorze $R \subset \mathbb{R}^2$ jest jednostajnie ciągła, tzn. gdy dwa argumenty funkcji są położone blisko siebie, to również wartości funkcji leżą blisko siebie. Czyli*

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x, y \in R\ \|x - y\| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Dowód nie wprost. Załóżmy, że istnieje (złośliwa) liczba $\varepsilon > 0$ taka, że dla $\delta_n = \frac{1}{n}$ istnieją punkty x_n i y_n w R spełniające

$$\|x_n - y_n\| < \frac{1}{n}, \qquad |f(x_n) - f(y_n)| \geqslant \varepsilon.$$

Z ciągu x_n można wybrać zbieżny podciąg x_{n_k}. Niech $x_{n_k} \xrightarrow[k]{} x_0$. Wtedy

$$\|y_{n_k} - x_0\| \leqslant \|y_{n_k} - x_{n_k}\| + \|x_{n_k} - x_0\| \leqslant \frac{1}{n_k} + \|x_{n_k} - x_0\| \xrightarrow[k]{} 0.$$

Czyli $y_{n_k} \xrightarrow[k]{} x_0$. Zatem $f(x_{n_k}) \xrightarrow[k]{} f(x_0)$ oraz $f(y_{n_k}) \xrightarrow[k]{} f(x_0)$. Otrzymujemy sprzeczność, bo $|f(x_{n_k}) - f(y_{n_k})| \geqslant \varepsilon$. $\square$

Twierdzenie 6.4. *Funkcja ciągła jest całkowalna na prostokącie.*

Dowód. Z jednostajnej ciągłości, jeśli podział $\mathcal{P}$ jest wystarczająco drobny, to $U(\mathcal{P}, f) - L(\mathcal{P}, f) < \varepsilon$. □

Twierdzenie 6.5. *Rozważmy dwie funkcje f i g, całkowalne na prostokącie R. Wtedy*

(i) $\displaystyle\iint\limits_R (f+g)\,dx\,dy = \iint\limits_R f\,dx\,dy + \iint\limits_R g\,dx\,dy.$

(ii) $\displaystyle\iint\limits_R cf\,dx\,dy = c\iint\limits_R f\,dx\,dy.$

(iii) *Jeśli $f(x,y) \leqslant g(x,y)$ na R, to*

$$\iint\limits_R f\,dx\,dy \leqslant \iint\limits_R g\,dx\,dy.$$

(iv) *Jeśli R_i, $i = 1, 2, \ldots, n$, są prostokątami o bokach równoległych do osi takimi, że f jest całkowalna na każdym z nich oraz $R = R_1 \cup \ldots \cup R_n$, to f jest całkowalna na R oraz*

$$\iint\limits_R f\,dx\,dy = \sum_{i=1}^{n} \iint\limits_{R_i} f\,dx\,dy,$$

przy założeniu, że wnętrza prostokątów R_i są rozłączne pomiędzy sobą.

Uwaga. Prostokąty R_i, $i = 1, 2, \ldots, n$, nie muszą tworzyć podziału prostokąta R. Ale można rozdrobnić każdy z prostokątów R_i, aby uzyskać podział prostokąta R.

Twierdzenie 6.6 (Fubini)**.** *Załóżmy, że funkcja $f(x,y)$ jest ciągła na prostokącie $[a,b] \times [c,d]$. Wtedy*

$$\iint\limits_R f(x,y)\,dx\,dy = \int\limits_a^b \left(\int\limits_c^d f(x,y)\,dy\right)\,dx = \int\limits_c^d \left(\int\limits_a^b f(x,y)\,dx\right)\,dy.$$

Dowód. Rozważamy podziały $a = x_0 < x_1 < \ldots < x_n = b$ i $c = y_0 < y_1 < \ldots < y_m = d$.

$$\int_a^b \left(\int_c^d f(x,y)\,dy \right) dx = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} \left(\sum_{j=1}^m \int_{y_{j-1}}^{y_j} f(x,y)\,dy \right) dx$$
$$= \sum_{i=1}^n \sum_{j=1}^m \int_{x_{i-1}}^{x_i} \underbrace{\left(\int_{y_{j-1}}^{y_j} f(x,y)\,dy \right)}_{F_j(x)} dx = \sum_{i=1}^n \sum_{j=1}^m \int_{x_{i-1}}^{x_i} F_j(x)\,dx.$$

$F_j(x)$ jest funkcją ciągłą na $[x_{i-1}, x_i]$, co wynika z lematu poniżej.

Lemat 6.7. *Dla funkcji $f(x,y)$ ciągłej na $[a,b] \times [c,d]$ funkcja $F(x) = \int_c^d f(x,y)\,dy$ jest ciągła na $[a,b]$.*

Dowód lematu.

$$|F(x_1) - F(x_2)| = \left| \int_c^d f(x_1,y)\,dy - \int_c^d f(x_2,y)\,dy \right| \leqslant \int_c^d |f(x_1,y) - f(x_2,y)|\,dy.$$

Z jednostajnej ciągłości dla $\varepsilon > 0$ można znaleźć liczbę $\delta > 0$ taką, że

$$\|(x_1,y_1) - (x_2,y_2)\| < \delta \;\Rightarrow\; |f(x_1,y_1) - f(x_2,y_2)| < \frac{\varepsilon}{d-c}.$$

Wtedy dla $|x_1 - x_2| < \delta$ mamy $|f(x_1,y) - f(x_2,y)| < \frac{\varepsilon}{d-c}$. Ostatecznie

$$|F(x_1) - F(x_2)| \leqslant \frac{\varepsilon}{d-c}(d-c) = \varepsilon.$$

$\square$

Z twierdzenia o wartości średniej istnieją punkty ξ_{ij}, dla których

$$\int_{x_{i-1}}^{x_i} F_j(x)\,dx = F_j(\xi_{ij})\Delta x_i, \quad x_{i-1} \leqslant \xi_{ij} \leqslant x_i.$$

Dalej

$$F_j(\xi_{ij}) = \int_{y_{j-1}}^{y_j} f(\xi_{ij}, y)\,dy = f(\xi_{ij}, \eta_{ij})\Delta y_j, \quad y_{j-1} \leqslant \eta_{ij} \leqslant y_j,$$

dla pewnych punktów η_{ij}. Zatem

$$\int\limits_a^b \left(\int\limits_c^d f(x,y)\,dy \right)\,dx = \sum_{i=1}^n \sum_{j=1}^m f(\xi_{ij},\eta_{ij})\,\underbrace{\Delta x_i \Delta y_j}_{\Delta S_{ij}},$$

gdzie $S_{ij} = [x_{i-1}, x_i] \times [y_{j-1}.y_j]$. Punkt (ξ_{ij}, η_{ij}) leży w S_{ij}, stąd

$$L(\mathcal{P}, f) \leqslant \int\limits_a^b \left(\int\limits_c^d f(x,y)\,dy \right)\,dx \leqslant U(\mathcal{P}, f),$$

gdzie $\mathcal{P}$ jest podziałem wyznaczonym przez prostokąty S_{ij}. Ale

$$L(\mathcal{P}, f) \leqslant \iint\limits_R f(x,y)\,dx\,dy \leqslant U(\mathcal{P}, f).$$

Funkcja F jest całkowalna, więc można wybrać podział $\mathcal{P}$ taki, że $U(\mathcal{P}, f) - L(\mathcal{P}, f) < \varepsilon$. Wtedy

$$\left| \iint\limits_R f(x,y)\,dx\,dy - \int\limits_a^b \left(\int\limits_c^d f(x,y)\,dy \right)\,dx \right| < \varepsilon.$$

$\square$

Niekiedy będziemy musieli obliczać całki z funkcji nieciągłych, np. przy wyznaczaniu objętości brył, których podstawa nie jest prostokątem.

Przykład. Niech $f(x, y)$ będzie nieujemną funkcją ciągła określoną w kole $x^2 + y^2 \leqslant 1$. Chcemy obliczyć objętość obszaru pod wykresem. Wkładamy koło w kwadrat $[-1, 1] \times [-1, 1]$ i określamy funkcję

$$\tilde{f}(x,y) = \begin{cases} f(x,y) & x^2+y^2 \leqslant 1, \\ 0 & x^2+y^2 > 1. \end{cases}$$

Wtedy $V = \iint\limits_{[-1,1]\times[-1,1]} \tilde{f}(x,y)\,dx\,dy$.

Ogólnie, jeśli chcemy obliczyć całkę $\iint\limits_C f(x,y)\,dx\,dy$, gdzie $C \subset \mathbb{R}^2$, to wkładamy C w prostokąt o bokach równoległych do osi i obliczamy

$$\iint\limits_R f(x,y)\mathbb{1}_C(x,y)\,dx\,dy.$$

Pojawia się problem całkowalności funkcji $f(x,y)\mathbb{1}_C(x,y)$. Jeśli $\mathbb{1}_C(x,y)$ jest całkowalna a $f(x,y)$ jest ciągła, to iloczyn jest funkcją całkowalną, bo iloczyn funkcji całkowalnych jest całkowalny.

Przykład. $C = \{(x,y) \, : \, x^2+y^2 \leqslant 1.\}$. Funkcja $\mathbb{1}_C(x,y)$ jest nieciągła w punktach okręgu $x^2+y^2=1$. Ogólnie funkcja $\mathbb{1}_C(x,y)$ jest nieciągła na brzegu zbioru C oznaczanym symbolem ∂C.

Definicja 6.8. *Mówimy, że zbiór $A \subset \mathbb{R}^2$ ma miarę zero, jeśli dla dowolnej liczby $\varepsilon > 0$ istnieją prostokąty $\{R_n\}_{n=1}^\infty$ takie, że*

$$A \subset \bigcup_{n=1}^\infty R_n, \qquad \sum_{n=1}^\infty \Delta R_n < \varepsilon.$$

Przykłady.

(a) Punkt ma miarę zero. Skończony zbiór punktów ma miarę zero.

(b) Przeliczalny zbiór punktów ma miarę zero. W szczególności zbiór punktów w kwadracie $[0,1]^2$ o obu współrzędnych wymiernych ma miarę zero.

(c) Poziomy odcinek ma miarę zero. Również ukośny odcinek ma miarę zero.

(d) Zbiór punktów kwadratu $[0,1]^2$ o obu współrzędnych niewymiernych nie ma miary zero.

Twierdzenie 6.9. *Ograniczona funkcja na prostokącie jest całkowalna wtedy i tylko wtedy, gdy zbiór jej punktów nieciągłości ma miarę zero.*

Twierdzenie 6.10. *Niech f będzie funkcją całkowalną na prostokącie $R = [a,b]\times[c,d]$. Dla $a \leqslant x \leqslant b$ niech*

$$\underline{\int_c^d} f(x,y)\,dy = \mathcal{L}(x) \leqslant \mathcal{U}(x) = \overline{\int_c^d} f(x,y)\,dy.$$

Wtedy funkcje $\mathcal{L}(x)$ i $\mathcal{U}(x)$ są całkowalne na $[a,b]$ oraz

$$\iint_R f(x,y)\,dx\,dy = \int_a^b \mathcal{L}(x)\,dx = \int_a^b \mathcal{U}(x)\,dx.$$

Uwagi.

1. Jeśli funkcja $y \mapsto f(x,y)$ jest całkowalna na $[c,d]$ dla $a \leqslant x \leqslant b$, to
$$\iint\limits_R f(x,y)\,dx\,dy = \int\limits_a^b \left(\int\limits_c^d f(x,y)\,dy \right) dx.$$

2. Zamieniając rolami x i y i przyjmując, że funkcja $x \mapsto f(x,y)$ jest całkowalna na $[a,b]$ dla $c \leqslant y \leqslant d$, otrzymamy
$$\iint\limits_R f(x,y)\,dx\,dy = \int\limits_c^d \left(\int\limits_a^b f(x,y)\,dx \right) dy.$$

Dowód. Niech $\mathcal{P} = (\mathcal{P}_1, \mathcal{P}_2)$ będzie podziałem prostokąta R. Rozważmy jeden prostokąt podziału $S = S_1 \times S_2$. Mamy
$$m_S(f) = m_{S_1 \times S_2}(f) \leqslant m_{S_2}(f(x,\cdot)), \quad \text{dla } x \in S_1.$$

Zatem
$$\begin{aligned}\sum_{S_2 \in \mathcal{P}_2} m_{S_1 \times S_2}(f) \Delta S_2 &\leqslant \sum_{S_2 \in \mathcal{P}_2} m_{S_2}(f(x,\cdot)) \Delta S_2 \\ &= L(\mathcal{P}_2, f(x,\cdot)) \leqslant \underline{\int\limits_c^d} f(x,y)\,dy = \mathcal{L}(x), \quad \text{dla } x \in S_1.\end{aligned}$$

Po wzięciu kresu dolnego względem $x \in S_1$ otrzymujemy
$$\sum_{S_2 \in \mathcal{P}_2} m_{S_1 \times S_2}(f) \Delta S_2 \leqslant m_{S_1}(\mathcal{L}).$$

Zatem
$$L(\mathcal{P}, f) = \sum_{S_1 \in \mathcal{P}_1} \sum_{S_2 \in \mathcal{P}_2} m_{S_1 \times S_2}(f) \Delta S_1 \Delta S_2 \leqslant \sum_{S_1 \in \mathcal{P}_1} m_{S_1}(\mathcal{L}) \Delta S_1 = L(\mathcal{P}_1, \mathcal{L}).$$

Podobnie pokazujemy, że $U(\mathcal{P}, f) \geqslant U(\mathcal{P}_1, \mathcal{U})$. Reasumując otrzymujemy
$$L(\mathcal{P}, f) \leqslant L(\mathcal{P}_1, \mathcal{L}) \leqslant U(\mathcal{P}_1, \mathcal{L}) \leqslant U(\mathcal{P}_1, \mathcal{U}) \leqslant U(\mathcal{P}, f).$$

Z założenia $f(x,y)$ jest całkowalna. Stąd wynika, że $\mathcal{L}(x)$ jest całkowalna na $[a,b]$. Ponadto

$$L(\mathcal{P},f) \leqslant \iint\limits_R f(x,y)\,dx\,dy \leqslant U(\mathcal{P},f),$$

$$L(\mathcal{P},f) \leqslant \int\limits_a^b \mathcal{L}(x)\,dx \leqslant U(\mathcal{P},f).$$

Zatem $\iint\limits_R f(x,y)\,dx\,dy = \int\limits_a^b \mathcal{L}(x)\,dx.$ $\square$

Przykład. $D = \{(x,y) \ : \ x^2+y^2 \leqslant 1\}$. Znaleźć objętość obszaru pod wykresem funkcji $f(x,y) = 2x+y+5$ na D. Obliczamy

$$\begin{aligned}
\iint\limits_D f(x,y)\,dx\,dy &= \iint\limits_{[-1,1]^2} (2x+y+5)\mathbb{1}_D(x,y)\,dx\,dy \\
&= \int\limits_{-1}^{1}\left(\int\limits_{-1}^{1}(2x+y+5)\mathbb{1}_D(x,y)\,dy\right)dx = \int\limits_{-1}^{1}\int\limits_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}(2x+y+5)\,dy\,dx \\
&= 2\int\limits_{-1}^{1}(2x+5)\sqrt{1-x^2}\,dx = 10\int\limits_{-1}^{1}\sqrt{1-x^2}\,dx = 5\pi.
\end{aligned}$$

Twierdzenie 6.11. *Niech $y = f(x)$ będzie funkcją ciągłą na $[a,b]$. Wtedy wykres funkcji f ma miarę zero.*

Dowód. Ustalmy $\varepsilon > 0$. Można znaleźć liczbę naturalną N taką, że

$$|x-x'| < \frac{b-a}{N} \ \Rightarrow\ |f(x)-f(x')| < \frac{\varepsilon}{8(b-a)}.$$

Dzielimy przedział $[a,b]$ na N równych części punktami $a = x_0 < x_1 < \ldots < x_N = b$. Określmy $x_{-1} = a - \frac{b-a}{N}$, $x_{N+1} = b + \frac{b-a}{N}$. Każdy z punktów x przedziału $[a,b]$ leży w jednym z przedziałów (x_{i-1},x_{i+1}) dla $i = 1,2,\ldots,N$. Jeśli $x \in (x_{i-1},x_{i+1})$, to $|f(x)-f(x_i)| < \frac{\varepsilon}{8(b-a)}$. To oznacza, że

$$f(x) \in \left(f(x_i) - \frac{\varepsilon}{8(b-a)}, f(x_i) + \frac{\varepsilon}{8(b-a)}\right).$$

Zatem wykres jest zawarty w zbiorze

$$\bigcup_{i=1}^{N}(x_{i-1},x_{i+1})\times\left(f(x_i)-\tfrac{\varepsilon}{8(b-a)},f(x_i)+\tfrac{\varepsilon}{8(b-a)}\right).$$

Suma pól składników tego zbioru wynosi

$$\frac{2(b-a)}{N}\cdot\frac{\varepsilon}{4(b-a)}=\frac{\varepsilon}{2}.$$

$\square$

6.2.1 Obliczanie pól

Dla ograniczonego podzbioru $D\subset\mathbb{R}^2$ takiego, że ∂D ma miarę zero określamy

$$A(D)=\iint\limits_D dx\,dy=\iint\limits_R \mathbb{1}_D(x,y)\,dx\,dy,$$

gdzie R jest prostokatem zawierającym D. Niech $\mathcal{P}$ będzie podziałem prostokąta R. Wtedy

$$L(\mathcal{P},\mathbb{1}_D)=\sum_S m_S(\mathbb{1}_D)\Delta S,\qquad U(\mathcal{P},\mathbb{1}_D)=\sum_S M_S(\mathbb{1}_D)\Delta S.$$

Wielkość $L(\mathcal{P},\mathbb{1}_D)$ jest sumą pól prostokątów podziału całkowicie zawartych w D, natomiast $U(\mathcal{P},\mathbb{1}_D)$ jest sumą pól prostokątów podziału mających część wspólną z D. Polem wewnętrznym nazywamy kres górny liczb $L(\mathcal{P},\mathbb{1}_D)$ a polem zewnętrznym kres dolny liczb $U(\mathcal{P},\mathbb{1}_D)$. Mówimy, że obszar D ma pole, jeśli pole wewnętrzne jest równe polu zewnętrznemu. Obszar D ma pole wtedy i tylko wtedy, gdy ∂D ma miarę zero. Mówimy wtedy, że obszar jest mierzalny w sensie Jordana.

Twierdzenie 6.12. *Jeśli $f(x,y)$ jest funkcją ciągłą w prostokącie R i $D\subset R$ jest mierzalny w sensie Jordana, to całka*

$$\iint\limits_D f(x,y)\,dx\,dy$$

jest dobrze określona.

Dowód.
$$\iint\limits_D f(x,y)\,dx\,dy = \iint\limits_R f(x,y)\mathbb{1}_D(x,y)\,dx\,dy.$$

Funkcja $f(x,y)\mathbb{1}_D(x,y)$ jest nieciągła tylko w punktach ∂D. $\square$

Twierdzenie 6.13. *Niech D_1 i D_2 będą ograniczonymi rozłącznymi podzbiorami $\mathbb{R}^2$ mierzalnymi w sensie Jordana. Dla funkcji $f(x,y)$ ciągłej na $D_1 \cup D_2$ mamy*

$$\iint\limits_{D_1\cup D_2} f(x,y)\,dx\,dy = \iint\limits_{D_1} f(x,y)\,dx\,dy + \iint\limits_{D_2} f(x,y)\,dx\,dy.$$

Dowód. Wkładamy D_1 i D_2 w prostokąt R. Wtedy

$$\iint\limits_{D_1\cup D_2} f\,dx\,dy = \iint\limits_R f\,\mathbb{1}_{D_1\cup D_2}\,dx\,dy = \iint\limits_R f\,[\mathbb{1}_{D_1} + \mathbb{1}_{D_2}]\,dx\,dy$$
$$\iint\limits_R f\,\mathbb{1}_{D_1}\,dx\,dy + \iint\limits_R f\,\mathbb{1}_{D_2}\,dx\,dy = \iint\limits_{D_1} f\,dx\,dy + \iint\limits_{D_2} f\,dx\,dy.$$

$\square$

Przykład. Dwa boki równoległoboku D znajdują się na poziomach $y = c$ i $y = d$. Dolny bok mieści się pomiędzy $x = a$ i $x = b$ a górny pomiędzy a' i b' oraz $a' > a$. Wkładamy D w prostokąt $R = [a, b'] \times [c, d]$. Wtedy

$$A(D) = \iint\limits_R \mathbb{1}_D(x,y)\,dx\,dy = \int\limits_c^d \left(\int\limits_a^{b'} \mathbb{1}_D(x,y)\,dx \right) dy.$$

Przy ustalonej wartości y funkcja $\mathbb{1}_D(x,y)$ jest równa 1 na przedziale długości $b - a$. Zatem
$$A(D) = \int\limits_c^d (b-a)\,dy = (b-a)(d-c).$$

6.2.2 Zmiana kolejności całkowania

Rozważmy całkę iterowaną

$$\int_0^a \int_0^{\sqrt{a^2-x^2}} \sqrt{a^2-y^2}\,dy\,dx = \iint_D \sqrt{a^2-y^2}\,dx\,dy$$
$$= \int_0^a \int_0^{\sqrt{a^2-y^2}} \sqrt{a^2-y^2}\,dx\,dy = \int_0^a (a^2-y^2)\,dy = a^3 - \frac{a^3}{3} = \frac{2}{3}a^3.$$

Przy zmienionej kolejności całkowania obliczenia okazały się łatwiejsze. Podobnie

$$\int_1^2 \int_0^{\log x} (x-1)\sqrt{1+2e^y}\,dy\,dx = \int_0^{\log 2} \int_{e^y}^2 (x-1)\sqrt{1+2e^y}\,dx\,dy$$
$$= \int_0^{\log 2} \sqrt{1+2e^y}\,(2-e^y)\frac{1}{2}(1+e^y-1)\,dy.$$

W ostatniej całce wykonujemy podstawienie $u = 1+2e^y$. Wtedy

$$e^y = \frac{u-1}{2} \qquad du = 2e^y\,dy.$$

Otrzymujemy

$$\int_3^5 \sqrt{u}\left(2 - \frac{u-1}{2}\right)\frac{1}{4}\,du.$$

Definicja 6.14. *Obszar $D \subset \mathbb{R}^2$ nazywamy łukowo spójnym, jeśli dla dowolnych dwóch punktów (x_1, y_1) i (x_2, y_2) w D można znaleźć funkcję ciągłą $\varphi : [0,1] \to D$ taką, że $\varphi(0) = (x_1, y_1)$ oraz $\varphi(1) = (x_2, y_2)$.*

Twierdzenie 6.15 (o wartości średniej)**.** *Niech $f(x, y)$ będzie funkcją ciągłą na zwartym obszarze $D \subset \mathbb{R}^2$ mierzalnym w sensie Jordana i łukowo spójnym. Wtedy*

$$\iint_D f(x,y)\,dx\,dy = f(x_0, y_0)A(D)$$

dla pewnego punktu (x_0, y_0) w D.

Dowód. Mamy

$$m = \min_{(x,y)\in D} f(x,y) = f(x_1,y_1), \qquad M = \max_{(x,y)\in D} f(x,y) = f(x_2,y_2)$$

dla pewnych punktów (x_1,y_1) i (x_2,y_2) w D. Dalej

$$m\,A(D) \leqslant \iint\limits_D f(x,y)\,dx\,dy \leqslant M\,A(D).$$

Jeśli $A(D)=0$, to teza jest spełniona. Niech $A(D)>0$. Wtedy

$$f(x_1,y_1) = m \leqslant \underbrace{\frac{1}{A(D)}\iint\limits_D f(x,y)\,dx\,dy}_{\alpha} \leqslant M = f(x_2,y_2).$$

Niech φ będzie funkcją ciągłą taką, że $\varphi : [0,1] \to D$ oraz $\varphi(0) = (x_1,y_1)$, $\varphi(1) = (x_2,y_2)$. Rozważmy funkcję $g(t) = f(\varphi(t))$. Wtedy g jest funkcją ciągłą oraz $g(0) = f(x_1,y_1)$ i $g(1) = f(x_1,y_1)$. Ponadto $g(0) \leqslant \alpha \leqslant g(1)$. Z własności Darboux mamy $g(t_0) = \alpha$ dla pewnej wartości $0 \leqslant t_0 \leqslant 1$. Tzn. $f(\varphi(t_0)) = \alpha$ oraz $\varphi(t_0) = (x_0,y_0)$. □

6.2.3 Geometria odwzorowań z $\mathbb{R}^2$ w $\mathbb{R}^2$

Przykłady.

(a) Niech $D = [0,1]\times[0,2\pi)$. Określamy

$$T(r,\varphi) = (r\cos\varphi, r\sin\varphi).$$

T odwzorowuje prostokąt D w koło jednostkowe.

(b) $T(x,y) = \left(\frac{x+y}{2}, \frac{x-y}{2}\right)$, $D = [-1,1]\times[-1,1]$. T jest odwzorowaniem liniowym. Aby wyznaczyć obraz $T(D)$ wystarczy więc znaleźć obraz wierzchołków obszaru D. Można też wskazać warunki jakie muszą spełniać punkty z $T(D)$. Niech $u = \frac{x+y}{2}$ i $v = \frac{x-y}{2}$. Wtedy $x = u+v$ i $y = u-v$. Zatem punkt (u,v) spełnia $|u+v| \leqslant 1$ i $|u-v| \leqslant 1$. To oznacza, że $|u|+|v| \leqslant 1$.

6.3 Twierdzenie o zamianie zmiennych

Dane są dwa zbiory D i D^* w $\mathbb{R}^2$ i odwzorowanie $T : D^* \to D$ klasy C^1, różnowartościowe oraz $T(D^*) = D$. Zakładamy, że D i D^* są mierzalne w sensie Jordana. Chcemy wyrazić wielkość $\iint\limits_D f(x,y)\,dx\,dy$ jako całkę po zbiorze D^* z funkcji złożonej $f \circ T$.

Uwaga. Dla funkcji jednej zmiennej mamy

$$\int\limits_{\varphi(a)}^{\varphi(b)} f(x)\,dx = \int\limits_a^b f(\varphi(u))\varphi'(u)\,du, \qquad [a,b] \xrightarrow{\varphi} [\varphi(a),\varphi(b)].$$

Zaczniemy od przypadku $f \equiv 1$. Tzn. chcemy obliczyć $\iint\limits_D dx\,dy = A(D)$ za pomocą całki po obszarze D^* z funkcji 1 ewentualnie domnożonej przez jakąś funkcję zależną od T.

Wiemy, że jeśli T jest odwzorowaniem różniczkowalnym w (u_0, v_0), to dla odwzorowania liniowego $DT(u_0, v_0)$ zadanego macierzą

$$DT(u_0,v_0) = \begin{pmatrix} \dfrac{\partial x}{\partial u}(u_0,v_0) & \dfrac{\partial x}{\partial v}(u_0,v_0) \\ \dfrac{\partial y}{\partial u}(u_0,v_0) & \dfrac{\partial y}{\partial v}(u_0,v_0) \end{pmatrix}$$

mamy

$$T(u,v) \approx T(u_0,v_0) + DT(u_0,v_0)\begin{pmatrix}\Delta u \\ \Delta v\end{pmatrix} =: \widetilde{T}(u,v),$$

gdzie $\Delta u = u - u_0$ oraz $\Delta v = v - v_0$. Jeśli S jest małym prostokątem, o bokach równoległych do osi, wewnątrz D^*, którego dolnym lewym wierzchołkiem jest punkt (u_0, v_0), to obraz $T(S)$ jest w przybliżeniu równoległobokiem oraz

$$A(T(S)) \approx A(\widetilde{T}(S)) = |\det(DT(u_0,v_0))|\,A(S).$$

Przypuśćmy, że obszar D^* został włożony w prostokąt R, który następnie podzieliliśmy na małe prostokąty S_k. Rozważamy tylko prostokąty S_k całkowicie zawarte w D^*. Niech (u_k, v_k) oznacza lewy dolny wierzchołek prostokąta S_k. Wtedy

$$\iint\limits_D dx\,dy = A(D) = A(T(D^*)) \approx \sum_k |\det(DT(u_k,v_k))|\,A(S_k).$$

W granicy, gdy średnica podziału dąży do zera, otrzymamy

$$\iint\limits_D dx\,dy = \iint\limits_{D^*} |\det DT(u,v)|\,du\,dv.$$

Jakobianem odwzorowania T nazywamy wyznacznik

$$J_T(u,v) = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix}.$$

Oznaczmy $(x_k, y_k) = T(u_k, v_k)$. Mamy

$$\begin{aligned}\iint\limits_D f(x,y)\,dx\,dy &\approx \sum_k f(x_k,y_k)\,A(T(S_k)) \\ &\approx \sum_k f(x_k,y_k)|J_T(u_k,v_k))|\,A(S_k) \approx \iint\limits_{D^*} f(T(u,v))\,|J_T(u,v)|\,du\,dv.\end{aligned}$$

Ostatecznie otrzymujemy wzór

$$\iint\limits_D f(x,y)\,dx\,dy = \iint\limits_{D^*} f(T(u,v))\,|J_T(u,v)|\,du\,dv. \tag{6.1}$$

Przykład. Rozważmy całkę $\int\limits_{-\infty}^{\infty} e^{-x^2}\,dx$. Mamy

$$\begin{aligned}&\int\limits_{-\infty}^{\infty} e^{-x^2}\,dx = 2\int\limits_0^{\infty} e^{-x^2}\,dx \\ &= 2\left[\lim_{R\to\infty}\left(\int\limits_0^R e^{-x^2}\,dx\right)\left(\int\limits_0^R e^{-y^2}\,dy\right)\right]^{1/2} = 2\left[\lim_{R\to\infty}\iint\limits_{[0,R]^2} e^{-(x^2+y^2)}\,dx\,dy\right]^{1/2}\end{aligned}$$

Niech D_R oznacza część koła o środku w początku układu i promieniu R leżącą w pierwszej ćwiartce. Wtedy

$$\iint\limits_{D_R} e^{-(x^2+y^2)}\,dx\,dy \leqslant \iint\limits_{[0,R]^2} e^{-(x^2+y^2)}\,dx\,dy \leqslant \iint\limits_{D_{R\sqrt{2}}} e^{-(x^2+y^2)}\,dx\,dy.$$

Użyjemy współrzędnych biegunowych

$$x = r\cos\varphi,\ y = r\sin\varphi, \quad 0 \leqslant \varphi \leqslant \frac{\pi}{2},\ 0 \leqslant r \leqslant R.$$

Mamy

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \varphi} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \varphi} \end{vmatrix} = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r.$$

Prostokąt $[\delta, R] \times [0, \frac{\pi}{2}]$ jest przekształcony na ćwiartkę pierścienia

$$D_{\delta,R} = \{(x,y)\ :\ x, y \geqslant 0,\ \delta^2 \leqslant x^2 + y^2 \leqslant R^2\}.$$

Zatem

$$\iint\limits_{D_R} e^{-(x^2+y^2)}\,dx\,dy = \lim_{\delta\to 0^+} \iint\limits_{D_{\delta,R}} e^{-(x^2+y^2)}\,dx\,dy$$

$$= \lim_{\delta\to 0^+} \int\limits_{\delta}^{R}\int\limits_{0}^{\frac{\pi}{2}} e^{-r^2} r\,d\varphi\,dr = \frac{\pi}{2}\lim_{\delta\to 0^+}\int\limits_{\delta}^{R} e^{-r^2} r\,dr = \frac{\pi}{2}\int\limits_{0}^{R} e^{-r^2} r\,dr$$

$$= \frac{\pi}{2}\left(-\frac{1}{2}e^{-r^2}\Big|_0^R\right) = \frac{\pi}{4}(1 - e^{-R^2}) \underset{R\to\infty}{\longrightarrow} \frac{\pi}{4}.$$

W świetle poprzednich obliczeń otrzymujemy

$$\int\limits_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}.$$

Uwaga. Współrzędne biegunowe są użyteczne, gdy funkcja podcałkowa zawiera x^2+y^2 a obszar całkowania jest kołem lub fragmentem koła. Rozważmy całkę $\iint\limits_{D} \log(x^2+y^2)\,dx\,dy$, gdzie D jest wycinkiem koła opisanym przez warunki $a \leqslant r \leqslant b$ i $0 \leqslant \varphi \leqslant \frac{\pi}{2}$. Po zamianie zmiennych otrzymujemy

$$\int\limits_{a}^{b}\int\limits_{0}^{\frac{\pi}{2}} \log(r^2)\,r\,dr = \frac{\pi}{2}\int\limits_{a}^{b} \log(r^2)\,r\,dr.$$

Niekiedy warto użyć współrzędnych biegunowych mimo, że obszar nie jest "wygodny". Rozważmy całkę

$$\iint\limits_{[0,1]^2} \sqrt{x^2+y^2}\, dx\, dy.$$

Ze względu na symetrię mamy

$$\iint\limits_{[0,1]^2} \sqrt{x^2+y^2}\, dx\, dy = 2 \iint\limits_{\substack{[0,1]^2\\ y\leqslant x}} \sqrt{x^2+y^2}\, dx\, dy.$$

Mamy $0 \leqslant \varphi \leqslant \dfrac{\pi}{4}$ oraz $0 \leqslant r\cos\varphi \leqslant 1$. Tzn. $0 \leqslant r \leqslant \dfrac{1}{\cos\varphi}$. Otrzymujemy więc

$$\iint\limits_{[0,1]^2} \sqrt{x^2+y^2}\, dx\, dy = 2\int\limits_0^{\frac{\pi}{4}} \int\limits_0^{\frac{1}{\cos\varphi}} r^2\, dr\, d\varphi = \frac{2}{3}\int\limits_0^{\frac{\pi}{4}} \frac{d\varphi}{\cos^3\varphi} = \frac{2}{3}\int\limits_0^{\frac{\pi}{4}} \frac{\cos\varphi}{(1-\sin^2\varphi)^2}\, d\varphi.$$

W ostatniej całce po podstawieniu $u = \cos\varphi$ otrzymamy całkę z funkcji wymiernej.

7 Całki potrójne i wielokrotne

Przedziałem $R \subset \mathbb{R}^N$ nazywamy iloczyn kartezjański

$$R = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_N, b_N].$$

Objętością przedziału jest wielkość

$$\Delta R = (b_1 - a_1)(b_2 - a_2)\ldots(b_N - a_N).$$

Podział $\mathcal{P}$ przedziału R oznacza rodzinę podziałów $\mathcal{P} = (\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_N)$, gdzie $\mathcal{P}_i$ jest podziałem przedziału $[a_i, b_i]$ na k_i części. W ten sposób otrzymujemy podział R na $k_1k_2\ldots k_N$ części (podprzedziałów). Dla podprzedziału S określmy

$$m_S(f) = \inf_{x\in S} f(x), \qquad M_S(f) = \sup_{x\in S} f(x),$$

gdzie $f(x)$ jest funkcją ograniczoną na przedziale R. Sumy dolne, górne, całkę dolną i górną oraz całkę określamy tymi samymi wzorami co dla funkcji jednej i dwu zmiennych. Można podobnie udowodnić, że funkcje ciągłe są całkowalne.

Definicja 7.1. *Mówimy, że zbiór $A \subset \mathbb{R}^N$ jest miary zero, jeśli istnieje ciąg przedziałów S_n taki, że*

$$A \subset \bigcup_{n=1}^{\infty} S_n \qquad \sum_{n=1}^{\infty} \Delta S_n < \varepsilon,$$

dla dowolnie wcześniej ustalonej liczby dodatniej ε.

Twierdzenie 7.2. *Ograniczona funkcja f określona na przedziale $R \subset \mathbb{R}^N$ jest całkowalna wtedy i tylko wtedy, gdy zbiór jej punktów nieciągłości ma miarę zero.*

Twierdzenie 7.3 (Fubini)**.** *Niech $A \subset \mathbb{R}^N$ i $B \subset \mathbb{R}^M$ będą przedziałami. Załóżmy, że funkcja f określona na $A \times B \subset \mathbb{R}^N \times \mathbb{R}^M$*(8) *jest całkowalna. Dla $x \in A$ niech*

$$\mathcal{L}(x) = \underline{\int_B} f(x,y)\,dy, \qquad \mathcal{U}(x) = \overline{\int_B} f(x,y)\,dy.$$

Wtedy funkcje $\mathcal{L}(x)$ i $\mathcal{U}(x)$ są całkowalne na A oraz

$$\begin{aligned}\int_{A\times B} f(x,y)\,dx\,dy &= \int_A \mathcal{L}(x)\,dx = \int_A \mathcal{U}(x)\,dx \\ &= \int_A \left(\underline{\int_B} f(x,y)\,dy \right) dx = \int_A \left(\overline{\int_B} f(x,y)\,dy \right) dx.\end{aligned}$$

Jeśli funkcja $f(x,y)$ jest ciągła, to można pominąć znaki całki dolnej i górnej.

Przykład. Rozważmy funkcję trzech zmiennych $f(x,y,z)$ ciągłą na $R = [a_1,b_1] \times [a_2,b_2] \times [a_3,b_3]$. Określmy $A = [a_1,b_1] \times [a_3,b_3]$ i $B = [a_2,b_2]$. Wtedy

$$\begin{aligned}\iiint_R f(x,y,z)\,dx\,dy\,dz &= \iint_A \left(\int_{a_2}^{b_2} f(x,y,z)\,dy \right) dx\,dz \\ &= \int_{a_1}^{b_1} \left(\int_{a_3}^{b_3} \left(\int_{a_2}^{b_2} f(x,y,z)\,dy \right) dz \right) dx.\end{aligned}$$

---

8 Punkty z $\mathbb{R}^N \times \mathbb{R}^M$ będziemy oznaczać przez (x,y)

Mogliśmy zamienić całkę podwójną po A na całkę iterowaną bo całkowana funkcja zależy w sposób ciągły od x i z. Przy funkcji trzech zmiennych mamy sześć możliwości zamiany na całkę iterowaną.

Uwaga. Inny zapis całki iterowanej to:

$$\int_{a_1}^{b_1}\int_{a_3}^{b_3}\int_{a_2}^{b_2} f(x,y,z)\,dy\,dz\,dx = \int_{a_1}^{b_1} dx \int_{a_3}^{b_3} dz \int_{a_2}^{b_2} f(x,y,z)\,dy.$$

Twierdzenie 7.4. *Dla funkcji ciągłej φ określonej na przedziale $R \subset \mathbb{R}^{N-1}$ wykres funkcji φ, czyli zbiór $D = \{(x,\varphi(x)) \,:\, x \in R\}$ jest miary zero w $\mathbb{R}^N$.*

Przykład. Przesunięta podprzestrzeń $(N-1)$-wymiarowa w $\mathbb{R}^N$ ma miarę zero. Podprzestrzeń zadana jest wzorem

$$a_1x_1 + a_2x_2 + \ldots + a_Nx_N = b.$$

Mamy $a_k \neq 0$ dla pewnego k. Wtedy

$$x_k = \frac{b}{a_k} - \frac{1}{a_k}(a_1x_1 + \ldots + \cancel{a_kx_k} + \ldots + a_Nx_N)$$

opisuje wykres funkcji ciągłej na $\mathbb{R}^{N-1}$. Jeśli $D \subset \mathbb{R}^N$ nie jest przedziałem, to określamy

$$\int_D f(x)\,dx = \int_R f(x)\mathbb{1}_D(x)\,dx$$

dla przedziału R zawierającego D. Załóżmy, że $f(x)$ jest funkcją ciągłą. Wtedy funkcja $f(x)\mathbb{1}_D(x)$ może byc nieciągła tylko w punktach brzegu ∂D. Jeśli ∂D ma miarę zero, to $f(x)\mathbb{1}_D(x)$ jest całkowalna, np. gdy zbiór ∂D jest sumą kilku wykresów funkcji ciągłych $N-1$ zmiennych.

Przykład. W jest obszarem w $\mathbb{R}^3$ określonym przez warunki $x, y \geqslant 0$ oraz $x^2+y^2 \leqslant z \leqslant 2$. Chcemy obliczyć $\iiint\limits_W x\,dx\,dy\,dz$. Niech D będzie obszarem w płaszczyźnie (x, y) określonym przez $x, y \geqslant 0$ i $x^2 + y^2 \leqslant 2$. Wtedy

$$\iiint\limits_W x\,dx\,dy\,dz = \iint\limits_D dx\,dy \int\limits_{x^2+y^2}^{2} x\,dz = \iint\limits_D x(2 - x^2 - y^2)\,dx\,dy$$
$$= \int\limits_0^{\sqrt{2}} dx \int\limits_0^{\sqrt{2-x^2}} x(2 - x^2 - y^2)\,dy = \int\limits_0^{\sqrt{2}} \left[x(2 - x^2)^{3/2} - \frac{1}{3}x(2 - x^2)^{3/2}\right] dx$$
$$= \frac{2}{3}\int\limits_0^{\sqrt{2}} x(2 - x^2)^{3/2}\,dx = -\frac{2}{15}(2 - x^2)^{5/2}\Big|_0^{\sqrt{2}} = \frac{8}{15}\sqrt{2}.$$

Ścisłe uzasadnienie przejść do całek iterowanych jest następujące. Mamy

$$W \subset [0, \sqrt{2}] \times [0, \sqrt{2}] \times [0, 2] =: R.$$

$$\iiint\limits_W x\,dx\,dy\,dz = \iiint\limits_R x\mathbb{1}_W(x, y, z)\,dx\,dy\,dz = \iint\limits_{[0,\sqrt{2}]^2} dx\,dy \int\limits_0^2 x\mathbb{1}_W(x, y, z)\,dz$$
$$= \iint\limits_{[0,\sqrt{2}]^2} dx\,dy \int\limits_{x^2+y^2}^{2} x\mathbb{1}_D(x, y)\,dz = \iint\limits_{[0,\sqrt{2}]^2} x(2 - x^2 - y^2)\mathbb{1}_D(x, y)\,dx\,dy$$
$$= \int\limits_0^{\sqrt{2}} dx \int\limits_0^{\sqrt{2-x^2}} x(2 - x^2 - y^2)\,dy.$$

Twierdzenie 7.5 (o zamianie zmiennych)**.** *Niech D i D^* będą obszarami w $\mathbb{R}^n$. Załóżmy, że T jest odwzorowaniem różnowartościowym klasy C^1 takim, że $T(D^*) = D$. Wtedy dla funkcji $f(x)$ ciągłej (lub całkowalnej) określonej na D mamy*

$$\int\limits_D f(x)\,dx = \int\limits_{D^*} f(T(u))\,|J_T(u)|\,du,$$

gdzie $J_T(u)$ jest jakobianem odwzorowania T w punkcie u.

Uwaga. Dla u' blisko u mamy

$$T(u') \approx T(u) + DT(u)(u' - u),$$

czyli odwzorowanie T zachowuje się w przybliżeniu jak złożenie dwu przesunięć i przekształcenia liniowego o macierzy $DT(u)$. Przy takim przekształceniu objętość obrazu małego przedziału S obliczamy wzorem

$$\Delta T(S) \approx \Delta S\,|J_T(u)|, \quad \text{gdzie } u \in S.$$

Przykłady.

(a) $\iiint\limits_D e^{(x^2+y^2+z^2)^{3/2}}\,dx,\,dy\,dz$, gdzie D jest fragmentem kuli jednostkowej leżącym w pierwszym oktancie. Zastosujemy współrzędne sferyczne

$$\begin{aligned} x &= r\sin\varphi\cos\psi, \\ y &= r\sin\varphi\sin\psi, \\ z &= r\cos\varphi, \end{aligned} \tag{7.1}$$

gdzie $0 \leqslant \varphi,\ \psi \leqslant \frac{\pi}{2}$, $0 \leqslant r \leqslant 1$. Przyporządkowanie $(r,\varphi,\psi) \mapsto (x,y,z)$ określone wzorami wyżej nie jest różnowartościowe na $D^* = [0,1] \times [0, \frac{\pi}{2}] \times [0, \frac{\pi}{2}]$, ale staje się takie, gdy $r > 0$. Określamy $T(r,\varphi,\psi) = (x,y,z)$ wg wzorów (7.1). Mamy

$$|J_T(r,\varphi,\psi)| = r^2\sin\varphi.$$

Dalej

$$\iiint\limits_D e^{(x^2+y^2+z^2)^{3/2}}\,dx\,dy\,dz = \iiint\limits_{D^*} e^{r^3} r^2 \sin\varphi\,dr\,d\varphi\,d\psi$$

$$= \int_0^1 \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} r^2 e^{r^3} \sin\varphi\,d\varphi\,d\psi\,dr = \frac{\pi}{2}\int_0^1 r^2 e^{r^3}\,dr \int_0^{\frac{\pi}{2}} \sin\varphi\,d\varphi = \frac{\pi}{6} e^{r^3}\Big|_0^1 = \frac{\pi}{6}(e-1).$$

(b) Obliczymy objętość kuli $D = \{(x,y,z) : x^2+y^2+z^2 \leqslant R^2\}$. Mamy

$$V = \iiint\limits_V dx\,dy\,dz.$$

Przechodzimy do współrzędnych sferycznych.

$$V = \int_0^{\pi} d\varphi \int_0^{2\pi} d\psi \int_0^R r^2 \sin\varphi\,dr = 2\pi \int_0^R r^2\,dr \int_0^{\pi} \sin\varphi\,d\varphi = \frac{4}{3}\pi R^3.$$

Obliczenia nie są do końca ścisłe, bo współrzędne nie są jednoznaczne na pełnej kuli. Można je uściślić następująco. Rozważamy podzbiór kuli D_ε, dla $\varepsilon > 0$, określony warunkami

$$\varepsilon \leqslant r \leqslant R, \quad \varepsilon \leqslant \varphi \leqslant \pi - \varepsilon, \quad 0 \leqslant \psi \leqslant 2\pi - \varepsilon.$$

Wtedy

$$V = \lim_{\varepsilon\to 0^+} V_\varepsilon = \lim_{\varepsilon\to 0^+} \int\limits_{\varepsilon}^{\pi-\varepsilon} d\varphi \int\limits_0^{2\pi-\varepsilon} d\psi \int\limits_{\varepsilon}^{R} r^2 \sin\varphi \, dr.$$

Współrzędne cylindryczne określone są przez

$$\begin{aligned} x &= r\cos\varphi, \\ y &= r\sin\varphi, \\ z &= z, \end{aligned}$$

co oznacza, że w płaszczyźnie (x, y) przechodzimy do współrzędnych biegunowych. Wtedy

$$J = \frac{\partial(x,y,z)}{\partial(r,\varphi,z)} = \begin{vmatrix} \cos\varphi & -r\sin\varphi & 0 \\ \sin\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = r.$$

Przykład. $I = \int\limits_0^2 dx \int\limits_0^{\sqrt{2x-x^2}} dy \int\limits_0^a z\sqrt{x^2+y^2}\, dz$. Obszar całkowania względem x i y jest opisany warunkami $0 \leqslant x \leqslant 2$, $0 \leqslant y \leqslant \sqrt{2x - x^2}$. Po przekształceniu otrzymujemy $x^2 + y^2 \leqslant 2x$, $y \geqslant 0$. Rozpoznajemy górne półkole o promieniu 1 i środku w punkcie $(1, 0)$. Po przejściu do współrzędnych biegunowych otrzymujemy warunki $r \leqslant 2\cos\varphi$ oraz $0 \leqslant \varphi \leqslant \pi/2$. Zatem

$$\begin{aligned} I &= \int\limits_0^{\frac{\pi}{2}} d\varphi \int\limits_0^{2\cos\varphi} dr \int\limits_0^a zr^2\, dz = \frac{a^2}{2}\int\limits_0^{\frac{\pi}{2}} \frac{r^3}{3}\Big|_0^{2\cos\varphi} d\varphi = \frac{4a^2}{3}\int\limits_0^{\frac{\pi}{2}} \cos^3\varphi\, d\varphi \\ &= \frac{4a^2}{3}\int\limits_0^{\frac{\pi}{2}} (1-\sin^2\varphi)\cos\varphi\, d\varphi = \frac{4a^2}{3}\left(\sin\varphi - \frac{1}{3}\sin^3\varphi\right)\Big|_0^{\frac{\pi}{2}} = \frac{8a^2}{9}. \end{aligned}$$

Górne półkole sugeruje, że podstawienie

$$x = 1 + r\cos\varphi,\ y = r\sin\varphi, \quad \text{dla } 0 \leqslant r \leqslant 1,\ 0 \leqslant \varphi \leqslant \pi,$$

mogłoby być przydatne. Jednak po takim podstawieniu otrzymujemy "nieprzyjazną" całkę

$$I = \int_0^\pi d\varphi \int_0^1 dr \int_0^a zr\sqrt{r^2 + 2r\cos\varphi + 1}\, dz = \frac{a^2}{2}\int_0^\pi d\varphi \int_0^1 r\sqrt{r^2 + 2r\cos\varphi + 1}\, dr.$$

7.0.1 Środek masy

W punktach $P_1, P_2, \ldots, P_n$ umieszczamy masy $m_1, m_2, \ldots, m_n$. Środek masy P układu spełnia

$$\vec{OP} = \frac{\sum_{i=1}^n m_i \vec{OP_i}}{\sum_{i=1}^n m_i}.$$

Niech $P_i = (x_i, y_i, z_i)$, $m = \sum_i m_i$ oraz $P = (\overline{x}, \overline{y}, \overline{z})$. Wtedy

$$\overline{x} = \frac{1}{m}\sum_{i=1}^n m_i x_i, \quad \overline{y} = \frac{1}{m}\sum_{i=1}^n m_i y_i, \quad \overline{z} = \frac{1}{m}\sum_{i=1}^n m_i z_i.$$

Jeśli masa jest rozłożona w sposób ciągły w obszarze D z gęstością masy $\varrho(x, y, z)$ w punkcie (x, y, z), to środek masy wyraża się wzorem

$$\overline{x} = \frac{\iiint\limits_D x\varrho(x, y, z)\, dx\, dy\, dz}{\iiint\limits_D \varrho(x, y, z)\, dx\, dy\, dz}.$$

Podobnie wzory mamy dla współrzędnych $\overline{y}$ i $\overline{z}$.

Przykład. Znaleźć środek masy górnej półkuli o promieniu 1, czyli obszaru $x^2 + y^2 + z^2 \leqslant 1$, $z \geqslant 0$. Przyjmujemy stałą gęstość masy $\varrho \equiv 1$. Ze względu na symetrię obszaru środek masy ma współrzędne $(0, 0, \overline{z})$. Obliczamy

$$\overline{z} = \frac{3}{2\pi}\iiint\limits_D z\, dx\, dy\, dz = \frac{3}{2\pi}\int_0^{\frac{\pi}{2}} d\varphi \int_0^{2\pi} d\psi \int_0^1 r\cos\varphi\, r^2 \sin\varphi\, dr = 3 \cdot \frac{1}{4} \cdot \frac{1}{2} = \frac{3}{8}$$

7.0.2 Moment bezwładności

Rozważamy ciało D o gęstości masy $\varrho(x,y,z)$ w punkcie (x,y,z). Moment bezwładności względem osi x wyraża się wzorem

$$I_x = \iiint_D (y^2+z^2)\varrho(x,y,z)\,dx\,dy\,dz.$$

Podobnie określa się momenty I_y oraz I_z.

Przykład. Obliczyć moment bezwładności względem osi z obszaru pomiędzy paraboloidą $z = x^2+y^2$, cylindrem $x^2+y^2 = a^2$ oraz płaszczyzną $z=0$, przyjmując $\varrho \equiv 1$. Obszar opisany jest warunkami

$$0 \leqslant z \leqslant x^2+y^2 \leqslant a^2.$$

Użyjemy współrzędnych cylindrycznych. Wtedy

$$\begin{aligned} I_z = \iiint\limits_D (x^2+y^2)\,dx\,dy\,dz &= \int\limits_0^{2\pi} d\varphi \int\limits_0^{a^2} dz \int\limits_{\sqrt{z}}^{a} r^2\cdot r\,dr \\ &= 2\pi \int\limits_0^{a^2} \frac{1}{4}(a^4 - z^2)\,dz = \frac{\pi}{2}\left(a^6 - \frac{1}{3}a^6\right) = \frac{\pi}{3}a^6. \end{aligned}$$

7.0.3 Potencjał grawitacyjny

W punkcie (x,y,z) umieszczamy masę M. Siła oddziaływania na masę m umieszczoną w punkcie (x_1,y_1,z_1) jest gradientem potencjału

$$V(x_1,y_z,z_1) = \frac{GmM}{\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2}}.$$

Zakładamy, że masa jest rozmieszczona w obszarze D z gęstością $\varrho(x,y,z)$. Wtedy potencjał wyraża się wzorem

$$V(x_1,y_1,z_1) = \iiint\limits_D \frac{Gm\varrho(x,y,z)}{\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2}}\,dx\,dy\,dz.$$

Siła oddziaływania na masę m umieszczoną w punkcie (x_1,y_1,z_1) jest równa $\nabla V(x_1,y_1,z_1)$.

Przykład. Załóżmy, że D jest obszarem zawartym pomiędzy sferami

$$x^2 + y^2 + z^2 = r_1^2, \ x^2 + y^2 + z^2 = r_2^2,$$

gdzie $r_1 < r_2$. Przyjmujemy $\varrho \equiv 1$ oraz $m = 1$. Obliczymy wartość potencjału w punktach przestrzeni poza D. Ze względu na niezmienniczość na obroty względem początku układu wystarczy obliczyć $V(0,0,R)$. W obliczeniach użyjemy współrzędnych sferycznych.

$$\frac{1}{G}V(0,0,R) = \iiint\limits_D \frac{dx\,dy\,dz}{\sqrt{x^2+y^2+(z-R)^2}}$$
$$= \int\limits_0^\pi d\varphi \int\limits_0^{2\pi} d\psi \int\limits_{r_1}^{r_2} \frac{r^2 \sin\varphi}{\sqrt{r^2 - 2rR\cos\varphi + R^2}}\,dr = 2\pi \int\limits_{r_1}^{r_2} dr \int\limits_0^\pi \frac{r^2\sin\varphi}{\sqrt{r^2 - 2rR\cos\varphi + R^2}}\,d\varphi$$

W wewnętrznej całce stosujemy podstawienie

$$u = r^2 - 2rR\cos\varphi + R^2, \ du = 2rR\sin\varphi\,d\varphi.$$

$$\frac{1}{G}V(0,0,R) = \frac{\pi}{R}\int\limits_{r_1}^{r_2} r\,dr \int\limits_{(r-R)^2}^{(r+R)^2} \frac{du}{\sqrt{u}}\,du = \frac{2\pi}{R}\int\limits_{r_1}^{r_2} r[r+R-|r-R|]\,dr.$$

Załóżmy, że $R < r_1$. Wtedy

$$\frac{1}{G}V(0,0,R) = \frac{2\pi}{R}\int\limits_{r_1}^{r_2} 2Rr\,dr = 2\pi(r_2^2 - r_1^2).$$

Z kolei dla $R > r_2$ mamy

$$\frac{1}{G}V(0,0,R) = \frac{2\pi}{R}\int\limits_{r_1}^{r_2} 2r^2\,dr = \frac{4\pi}{3R}(r_2^3 - r_1^3).$$

Reasumując, wewnątrz obszaru potencjał jest stały (niezależny od R) zatem nie ma siły grawitacji. Z kolei na zewnątrz potencjał jest odwrotnie proporcjonalny do odległości punktu od początku układu. Zatem siła grawitacji jest odwrotnie proporcjonalna do kwadratu tej odległości.

8 Całki krzywoliniowe i powierzchniowe

8.1 Całka krzywoliniowa niezorientowana

Rozważamy funkcję $f : \mathbb{R}^3 \to \mathbb{R}$. Chcemy obliczyć całkę z funkcji $f(x,y,z)$ wzdłuż krzywej $\sigma : [a,b] \xrightarrow{1-1} \mathbb{R}^3$, $\sigma(t) = (x(t), y(t), z(t))$. Można myśleć, że obraz krzywej opisuje przewód, a wielkość $f(x,y,z)$ reprezentuje gęstość masy przewodu w punkcie (x,y,z). Chcemy, aby całka dała w wyniku całkowitą masę przewodu. Określamy całkę wzorem

$$
\begin{aligned}
\int\limits_{\sigma} f(x,y,z)\,ds &:= \int\limits_a^b f(\sigma(t))\|\sigma'(t)\|\,dt \\
&= \int\limits_a^b f(x(t),y(t),z(t))\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\,dt.
\end{aligned}
$$

Jeśli $\sigma(t)$ jest kawałkami klasy C^1 lub $f(\sigma(t))$ jest kawałkami ciągła, to rozbijamy przedział czasu na skończoną liczbę przedziałów, na których można zastosować powyższy wzór.

Przykład. Rozważmy spiralę $\sigma(t) = (\cos t, \sin t, t)$, $0 \leqslant t \leqslant 2\pi$. Niech $f(x,y,z) = x^2+y^2+z^2$. Mamy $\sigma'(t) = (-\sin t, \cos t, 1)$. Wtedy $\|\sigma'(t)\| = \sqrt{2}$. Zatem

$$
\int\limits_{\sigma} (x^2 + y^2 + z^2)\,ds = \int\limits_0^{2\pi} (1+t^2)\sqrt{2}\,dt = \sqrt{2}\left(2\pi + \frac{8}{3}\pi^3\right).
$$

8.1.1 Interpretacja całki

Podzielimy przedział czasu $[a,b]$ punktami $a = t_0 < t_1 < \ldots < t_n = b$. W ten sposób krzywa zostanie podzielona na n fragmentów. Przyjmujemy, że gęstość masy na danym fragmencie jest stała i wynosi $f(\sigma(s_i))$, gdzie $t_{i-1} \leqslant s_i \leqslant t_i$. Zakładamy, że długość fragmentu wynosi $\|\sigma(t_i) - \sigma(t_{i-1})\|$. Masa fragmentu wynosi w przybliżeniu

$$
m_i \approx f(\sigma(s_i))\,\|\sigma(t_i) - \sigma(t_{i-1})\|.
$$

Dalej z twierdzenia Lagrange'a mamy

$$\begin{aligned}
x(t_i) - x(t_{i-1}) &= x'(\alpha_i)\Delta t_i \approx x'(s_i)\Delta t_i,\\
y(t_i) - y(t_{i-1}) &= y'(\beta_i)\Delta t_i \approx y'(s_i)\Delta t_i,\\
z(t_i) - z(t_{i-1}) &= x'(\gamma_i)\Delta t_i \approx z'(s_i)\Delta t_i,
\end{aligned}$$

gdzie $t_{i-1} \leqslant \alpha_i, \beta_i, \gamma_i \leqslant t_i$. Całkowita masa krzywej wynosi w przybliżeniu

$$\sum_{i=1}^{n} f(\sigma(s_i))\sqrt{x'(s_i)^2 + y'(s_i)^2 + z'(s_i)^2}\,\Delta t_i \underset{n\to\infty}{\longrightarrow} \int_a^b f(\sigma(t))\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\,dt.$$

Sprawdzimy jeszcze, że różnica pomiędzy sumami

$$\sum_{i=1}^{n} f(\sigma(s_i))\sqrt{x'(\alpha_i)^2 + y'(\beta_i)^2 + z'(\gamma_i)^2}\,\Delta t_i$$

i sumą określoną wyżej dąży do zera, gdy $n \to \infty$. Skorzystamy z nierówności trójkąta

$$\left|\sqrt{a_1^2 + b_1^2 + c_1^2} - \sqrt{a_0^2 + b_0^2 + c_0^2}\right| \leqslant \sqrt{(a_1 - a_0)^2 + (b_1 - b_0)^2 + (c_1 - c_0)^2}.$$

Różnica miedzy sumami co do wartości bezwględnej nie przekracza

$$\begin{aligned}
&\sum_{i=1}^{n} |f(\sigma(s_i))| \left|\sqrt{x'(\alpha_i)^2 + y'(\beta_i)^2 + z'(\gamma_i)^2} - \sqrt{x'(s_i)^2 + y'(s_i)^2 + z'(s_i)^2}\right| \Delta t_i\\
\leqslant &\sum_{i=1}^{n} |f(\sigma(s_i))| \sqrt{[x'(\alpha_i) - x'(s_i)]^2 + [y'(\beta_i) - y'(s_i)]^2 + [z'(\gamma_i) - z'(s_i)]^2}\Delta t_i.
\end{aligned}$$

Funkcja f jest ograniczona na krzywej, np. $|f(\sigma(t))| \leqslant M$. Funkcje $x'(t)$, $y'(t)$ i $z'(t)$ są jednostajnie ciągłe. Możemy założyć, że przedział $[a, b]$ dzielimy na n równych części tak, aby oscylacja każdej z funkcji x', y' i z' była mniejsza niż $\varepsilon > 0$ na każdym podprzedziale podziału. Wtedy ostatnie wyrażenie jest mniejsze niż

$$\sum_{i=1}^{n} M\sqrt{\varepsilon^2 + \varepsilon^2 + \varepsilon^2}\,\Delta t_i = M(b - a)\sqrt{3}\varepsilon.$$

Jeśli rozważamy funkcję $f(x, y)$ dwu zmiennych i krzywą $\sigma : [a, b] \to \mathbb{R}^2$, to $\sigma(t) = (x(t), y(t))$ oraz

$$\int_\sigma f(x, y)\,ds = \int_a^b f(\sigma(t))\|\sigma'(t)\|\,dt = \int_a^b f(x(t), y(t))\sqrt{x'(t)^2 + y'(t)^2}\,dt.$$

Jeśli $f(x,y) \geqslant 0$, to całkę można interpretować jako pole powierzchni płotu, którego podstawą jest krzywa σ a wysokość w punkcie (x,y) wynosi $f(x,y)$.

Przykład. Ciocia Tomka Sawyera kazała wybiałkować płot z obu stron. Za każdy metr kwadratowy Tomek otrzymuje 2 \$. Płot opisany jest przez

$$\sigma(t) = (30\cos^3 t, 30\sin^3 t), \quad 0 \leqslant t \leqslant \pi, \quad f(x,y) = 1 + \frac{y}{3}.$$

Mamy

$$\sigma'(t) = 30(-3\cos^2 t\sin t, 3\sin^2 t\cos t), \quad \|\sigma'(t)\| = 90\sin t|\cos t|.$$

Powierzchnia płotu wynosi

$$90\int_0^{\pi}(1 + 10\sin^3 t)\sin t|\cos t|\,dt = 180\int_0^{\frac{\pi}{2}}(1 + 10\sin^3 t)\sin t\cos t\,dt$$

$$180\left[\frac{1}{2}\sin^2 t + 2\sin^5 t\right]\Bigg|_0^{\frac{\pi}{2}} = 180 \cdot \frac{5}{2} = 450.$$

Zarobek Tomka wyniesie zatem $450 \cdot 2 \cdot 2 = 1800$ \$.

8.2 Całka krzywoliniowa zorientowana

Niech $F(x,y,z)$ będzie polem sił w $\mathbb{R}^3$ (np. sił grawitacyjnych lub elektrycznych). Tzn. $F = (F_1, F_2, F_3)$. Załóżmy, że obiekt porusza się pod działaniem pola sił F wzdłuż krzywej $\sigma : [a,b] \to \mathbb{R}^3$. Przyjmijmy, że σ jest linią prostą i obiekt został przesunięty o wektor d. Załóżmy też, że pole sił jest stałe, tzn. $F(x,y,z)$ nie zależy od (x,y,z). Wykonana praca wynosi wtedy $\|F_d\|\|d\|$, gdzie F_d oznacza składową siły F równoległą do przesunięcia d. Niech α oznacza kąt pomiędzy F i d. Wtedy praca wynosi

$$\|F_d\|\,\|d\| = \|F\|\,\|d\|\cos\alpha = F \circ d.$$

Ogólnie, gdy σ nie jest linią prostą oraz $F(x,y,z)$ nie jest stałym polem sił, to dzielimy przedział $[a,b]$ na n równych części. Przyjmujemy, że fragment od $\sigma(t_{i-1})$ do $\sigma(t_i)$ jest odcinkiem i, że siła F jest stała na tym odcinku i wynosi $F(\sigma(s_i))$, gdzie $t_{i-1} \leqslant s_i \leqslant t_i$. Wykonana praca wynosi wtedy

$$W \approx \sum_{i=1}^{n} F(\sigma(s_i)) \circ [\sigma(t_i) - \sigma(t_{i-1})].$$

Dalej

$$\begin{aligned}\sigma(t_i)-\sigma(t_{i-1}) &= (x(t_i)-x(t_{i-1}), y(t_i)-y(t_{i-1}), z(t_i)-z(t_{i-1}))\\ &= (x'(\alpha_i), y'(\beta_i), z'(\gamma_i))\,\Delta t_i \approx (x'(s_i), y'(s_i), z'(s_i))\,\Delta t_i = \sigma'(s_i)\Delta t_i,\end{aligned}$$

gdzie $t_{i-1} \leqslant \alpha_i, \beta_i, \gamma_i \leqslant t_i$. Zatem

$$W \approx \sum_{i=1}^{n} F(\sigma(s_i)) \circ \sigma'(s_i)\Delta t_i \underset{n\to\infty}{\longrightarrow} \int\limits_a^b F(\sigma(t)) \circ \sigma'(t)\,dt.$$

Przyjmujemy więc

$$W = \int\limits_a^b F(\sigma(t)) \circ \sigma'(t)\,dt.$$

Tę wielkość nazywamy całką krzywoliniową zorientowaną. Stosuje się też inne oznaczenie na tę całkę

$$W = \int\limits_\sigma F \circ ds.$$

Uwaga. Praca jest równa całce z iloczynu skalarnego siły i wektora stycznego do krzywej, czyli wektora prędkości. Załóżmy, że $\sigma'(t) \neq 0$ dla $a \leqslant t \leqslant b$. Wtedy

$$T(t) = \frac{\sigma'(t)}{\|\sigma'(t)\|}$$

jest jednostkowym wektorem stycznym. Zatem

$$\begin{aligned}\int\limits_\sigma F \circ\, ds &= \int\limits_a^b F(\sigma(t)) \circ \sigma'(t)\,dt\\ &= \int\limits_a^b F(\sigma(t)) \circ T(t)\,\|\sigma'(t)\|\,dt = \int\limits_\sigma (F \circ T)\,ds. \qquad (8.1)\end{aligned}$$

Całka zorientowana jest zatem równa całce niezorientowanej z iloczynu skalarnego siły F z jednostkowym wektorem stycznym do σ.

Przykład. $\sigma(t) = (\sin t, \cos t, t)$ $0 \leqslant t \leqslant \pi$, oraz $F(x, y, z) = (x, y, z)$. Wtedy $\sigma'(t) = (\cos t, -\sin t, 1)$ oraz

$$\int\limits_\sigma F \circ ds = \int\limits_0^\pi (\sin t, \cos t, t) \circ (\cos t, -\sin t, 1)\,dt = \int\limits_0^\pi t\,dt = \frac{\pi^2}{2}.$$

Używamy też innych oznaczeń na całkę zorientowaną. Jeśli $F = (F_1, F_2, F_3)$ oraz $\sigma(t) = (x(t), y(t), z(t))$, to

$$\int_a^b F(\sigma(t)) \circ \sigma'(t)\, dt$$

$$= \int_a^b [F_1(x(t), y(t), z(t))\, x'(t) + F_2(x(t), y(t), z(t))\, y'(t) + F_3(x(t), y(t), z(t))\, z'(t)]\, dt$$

$$= \int_\sigma F_1\, dx + F_2\, dy + F_3\, dz.$$

W niektórych przypadkach całkę zorientowaną można obliczyć bez odwoływania się do definicji. Dotyczy to tzw. pól gradientowych.

Twierdzenie 8.1. *Jeśli $F(x, y, z) = \nabla f(x, y, z)$ dla funkcji $f(x, y, z)$ klasy C^1, to*

$$\int_\sigma F \circ ds = f(\sigma(b)) - f(\sigma(a)),$$

gdzie $\sigma : [a, b] \to \mathbb{R}^3$.

Dowód.

$$\int_\sigma F \circ ds = \int_a^b F(\sigma(t)) \circ \sigma'(t)\, dt = \int_a^b \nabla f(\sigma(t)) \circ \sigma'(t)\, dt$$

$$= \int_a^b \frac{d}{dt} f(\sigma(t))\, dt = f(\sigma(b)) - f(\sigma(a)).$$

□

Przykład. $\int_\sigma y\, dx + x\, dy$, gdzie $\sigma(t) = \left(\frac{t^4}{4}, \sin^3 \frac{\pi t}{2}, 0\right)$, dla $0 \leqslant t \leqslant 1$. Mamy $F = (y, x, 0)$. Wtedy $\nabla f = F$ dla $f(x, y, z) = xy$. Zatem

$$\int_\sigma y\, dx + x\, dy = xy\Big|_{(0,0,0)}^{(\frac{1}{4},1,0)} = \frac{1}{4}.$$

Uwaga. Nie każde pole wektorowe jest gradientem jakiejś funkcji. Załóżmy,

że
$$F = (F_1, F_2, F_3) = \nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right),$$
gdzie f jest klasy C^2. Wtedy
$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial z \partial x} = \frac{\partial^2 f}{\partial x \partial z}, \quad \frac{\partial^2 f}{\partial z \partial y} = \frac{\partial^2 f}{\partial y \partial z}.$$
Czyli
$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}, \quad \frac{\partial F_1}{\partial z} = \frac{\partial F_3}{\partial x}, \quad \frac{\partial F_2}{\partial z} = \frac{\partial F_3}{\partial y}.$$

Przykłady.

(a) $F(x, y, z) = (x^2 + yz, x + y, z)$. Mamy $\dfrac{\partial F_1}{\partial y} = z$, $\dfrac{\partial F_2}{\partial x} = 1$.

(b) $F(x, y, z) = (yz, zx, xy)$. Wtedy $F = \nabla(xyz)$.

(c) $F = -\dfrac{GMm}{r^3}(x, y, z)$. Dla $V(x, y, z) = \dfrac{GMm}{r}$ mamy $F = \nabla V$.

(d) $F = (0, 0, -mg)$. Dla $V = -mgz$ mamy $F = \nabla V$.

Definicja 8.2. *C nazywamy krzywą Jordana (simple curve) jeśli C jest obrazem odwzorowania $\sigma : [a, b] \to \mathbb{R}^3$ takiego, że σ jest kawałkami klasy C^1 oraz σ jest różnowartościowe. Tzn. C nie ma samoprzecięć. Punkty $\sigma(a)$ i $\sigma(b)$ nazywamy końcami krzywej C. Każda krzywa Jordana ma dwie orientacje. Krzywą Jordana z wybraną orientacją nazywamy zorientowaną krzywą Jordana.*

Definicja 8.3. *Zamkniętą krzywą Jordana nazywamy obraz przez odwzorowanie $\sigma : [a, b] \to \mathbb{R}^3$, kawałkami klasy C^1, gdzie σ jest różnowartościowe na $[a, b)$ oraz $\sigma(a) = \sigma(b)$.*

Przykład. $C = \{x^2 + y^2 = 1,\ z = 0\}$ jest okręgiem obieganym przeciwnie do wskazówek zegara (analogowego). Chcemy obliczyć $\displaystyle\int_C (y, 0, 0) \circ ds = \int_C y\, dx$.

Parametryzujemy C (zgodnie z orientacją)
$$\sigma(t) = (\cos t, \sin t, 0), \quad 0 \leqslant t \leqslant 2\pi.$$

Wtedy

$$\int_C y\,dx = \int_0^{2\pi} \sin t(-\sin t)\,dt = -\pi.$$

Trzeba uważać, aby parametryzacja σ była różnowartościowa i zgodna z orientacją krzywej C. Np. parametryzacja

$$\eta(t) = (\cos 2t, \sin 2t, 0) \quad 0 \leqslant t \leqslant 2\pi,$$

nie jest różnowartościowa. Okrąg C jest obiegany dwukrotnie. Z kolei

$$\eta(t) = (\sin t, \cos t, 0) \quad 0 \leqslant t \leqslant 2\pi,$$

jest parametryzacją niezgodną z orientacją okręgu C.

Dla zorientowanej krzywej Jordana C symbolem C^- oznaczymy tę samą krzywą, ale z przeciwną orientacją. Wtedy z (8.1) wynika

$$\int_{C^-} F \circ ds = -\int_C F \circ ds,$$

bo przy zmianie orientacji wektor T zmienia się na wektor przeciwny.

Jeśli C jest złożona z fragmentów $C_1, C_2, \ldots, C_n$, to parametryzujemy każdy fragment osobno. Obliczamy

$$\int_C F \circ ds = \int_{C_1} F \circ ds + \ldots + \int_{C_n} F \circ ds.$$

Przykład. C jest brzegiem kwadratu jednostkowego w pierwszej ćwiartce układu współrzędnych zorientowanego przeciwnie do wskazówek zegara (dodatnio). Kolejne boki kwadratu parametryzujemy przedziałem $[0, 1]$ następująco

$$t \mapsto (t, 0), \quad t \mapsto (1, t), \quad t \mapsto (1-t, 1), \quad t \mapsto (0, 1-t).$$

Wtedy

$$\int_C x^2\,dx + xy\,dy = \int_0^1 t^2\,dt + \int_0^1 t\,dt + \int_0^1 (1-t)^2(-1)\,dt = \frac{1}{2}.$$

9 Całki powierzchniowe

9.1 Powierzchnie w $\mathbb{R}^3$

Przykładem powierzchni jest wykres funkcji dwu zmiennych $z = f(x, y)$. Można zmienne zamienić rolami i otrzymać $x = h(y, z)$ lub $y = g(x, z)$.

Przykłady.

(a) $x = z - z^3$. W płaszczyźnie xz wykresem jest krzywa trzeciego stopnia. Do wykresu wraz punktem $(z - z^3, 0, z)$ należy też cała prosta $(z - z^3, y, z)$ równoległa do osi y. Wykres ma postać wygiętego nieskończonego arkusza papieru.

(b) Torus, czyli powierzchnia powstała przez obrót wokół osi z okręgu w płaszczyźnie yz, nie jest wykresem funkcji. Można tę powierzchnię podzielić na dwie części górną i dolną, które są wykresami funkcji zmiennych xy.

Definicja 9.1. *Powierzchnią sparametryzowaną nazywamy funkcję* $\Phi : D \to \mathbb{R}^3$, *gdzie* D *jest podzbiorem płaszczyzny. Powierzchnią nazywamy obraz* $S = \Phi(D)$. *Stosujemy zapis*

$$\Phi(u, v) = (x(u, v), y(u, v), z(u, v)), \quad (u, v) \in D.$$

Mówimy, że powierzchnia jest klasy C^1 *jeśli funkcje* x, y *i* z *są klasy* C^1.

Można myśleć, że odwzorowanie Φ skręca, wygina, rozciąga i ściska obszar D, aby otrzymać powierzchnię S.

9.2 Płaszczyzna styczna do powierzchni

Rozważamy odwzorowanie

$$t \to (x(t, v_0), y(t, v_0), z(t, v_0)) = \Phi(t, v_0),$$

gdzie (u_0, v_0) jest ustalonym punktem w D. To odwzorowanie opisuje krzywą w $\mathbb{R}^3$ leżącą w powierzchni S i przechodzącą w chwili $t = u_0$ przez punkt

$$\Phi(u_0, v_0) =: (x_0, y_0, z_0).$$

Wektorem stycznym do tej krzywej w punkcie (x_0, y_0, z_0) jest

$$T_u = \left(\frac{\partial x}{\partial u}(u_0, v_0), \frac{\partial y}{\partial u}(u_0, v_0), \frac{\partial z}{\partial u}(u_0, v_0)\right).$$

Podobnie rozpatrując krzywą $t \to \Phi(u_0, t)$ otrzymamy inny wektor styczny w punkcie (x_0, y_0, z_0)

$$T_v = \left(\frac{\partial x}{\partial v}(u_0, v_0), \frac{\partial y}{\partial v}(u_0, v_0), \frac{\partial z}{\partial v}(u_0, v_0)\right).$$

T_u i T_v są wektorami stycznymi w punkcie (x_0, y_0, z_0) do krzywych leżących w powierzchni S. Płaszczyzna rozpięta przez te wektory jest zatem styczna do powierzchni w tym punkcie. Wektorem normalnym do powierzchni w (x_0, y_0, z_0) nazywamy wektor $T_u \times T_v$.

Definicja 9.2. *Mówimy, że powierzchnia jest gładka w punkcie* $\Phi(u_0, v_0) = (x_0, y_0, z_0)$ *jeśli* $T_u \times T_v \neq 0$. *Intuicyjnie oznacza to, że punkt* (x_0, y_0, z_0) *nie leży na krawędzi ani też nie jest rogiem powierzchni.*

Przykład. $x = u\cos v$, $y = u\sin v$, $z = u$. Powierzchnia jest klasy C^1. Mamy

$$T_u = (\cos v, \sin v, 1), \qquad T_v = (-u\sin v, u\cos v, 0).$$

Wektory T_u i T_v są równoległe tylko, gdy $T_v = 0$. Tzn. $u = 0$. Zauważmy, że powierzchnia jest zapisana równaniem $x^2 + y^2 = z^2$, czyli opisuje dwa stożki stykające się w początku układu.

Przypuśćmy, że powierzchnia jest gładka w punkcie $(x_0, y_0, z_0) = \Phi(u_0, v_0)$. Wtedy równanie płaszczyzny stycznej ma postać

$$(x - x_0, y - y_0, z - z_0) \circ \eta = 0,$$

gdzie $\eta = T_u \times T_v$ obliczone w (u_0, v_0).

Uwaga.

$$(a_1, b_1, c_1) \times (a_2, b_2, c_2) = \left(\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}, \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix}, \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \right).$$

Otrzymany wektor jest prostopadły do (a_1, b_1, c_1) i (a_2, b_2, c_2). Rzeczywiście

$$a_1 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} + b_1 \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix} + c_1 \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} = 0.$$

Przykład. $x = u\cos v$, $y = u\sin v$, $z = u^2 + v^2$. Chcemy znaleźć punkty, w których płaszczyzna styczna jest dobrze określona. Mamy

$$T_u = (\cos v, \sin v, 2u), \qquad T_v = (-u\sin v, u\cos v, 2v).$$

$$\begin{aligned} T_u \times T_v &= \left(\begin{vmatrix} \sin v & 2u \\ u\cos v & 2v \end{vmatrix}, \begin{vmatrix} 2u & \cos v \\ 2v & -u\sin v \end{vmatrix}, \begin{vmatrix} \cos v & \sin v \\ -u\sin v & u\cos v \end{vmatrix} \right) \\ &= (2v\sin v - 2u^2\cos v, -2u^2\sin v - 2v\cos v,\ u). \end{aligned}$$

Zatem $T_u \times T_v = 0$ tylko, gdy $u = v = 0$. Przykładowo w punkcie $(u_0, v_0) = (1, 0)$ mamy

$$(x_0, y_0, z_0) = (1, 0, 1), \qquad T_u \times T_v \Big|_{\substack{u=1 \\ v=0}} = (-2, 0, 1).$$

Równanie płaszczyzny stycznej w punkcie $(1, 0, 1)$ ma zatem postać

$$-2(x-1) + (z-1) = 0$$

czyli po uproszczeniu

$$2x - z = 1.$$

Zauważmy, że

$$x^2 + y^2 = u^2, \qquad \frac{y}{x} = \operatorname{tg} v.$$

Zatem równanie powierzchni ma postać

$$z = x^2 + y^2 + \operatorname{arctg}^2\left(\frac{y}{x}\right).$$

Przypuśćmy, że powierzchnia jest wykresem funkcji $z = f(x, y)$, dla $(x, y) \in D$. Naturalną parametryzacją jest $x := x$, $y := y$ i $z = f(x, y)$. Wtedy

$$T_x = \left(1, 0, \frac{\partial f}{\partial x}\right), \quad T_y = \left(0, 1, \frac{\partial f}{\partial y}\right), \quad T_x \times T_y = \left(-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1\right).$$

Równanie płaszczyzny stycznej w punkcie (x_0, y_0, z_0), gdzie $z_0 = f(x_0, y_0)$ to

$$(x - x_0, y - y_0, z - z_0) \circ \left(-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1\right) = 0.$$

Po przekształceniu otrzymujemy

$$z - z_0 = \frac{\partial f}{\partial x}\,(x - x_0) + \frac{\partial f}{\partial y}\,(y - y_0),$$

gdzie pochodne cząstkowe obliczane są w punkcie (x_0, y_0).

Przykłady.

(a) Chcemy sparametryzować powierzchnię (hiperboloidę) o równaniu $x^2 + y^2 - z^2 = 1$. Dla ustalonej wartości z punkty (x, y) leżą na okręgu o środku w $(0, 0)$ i promieniu $\sqrt{z^2 + 1} = r \geqslant 1$. Możemy przyjąć, że

$$x = r\cos\varphi, \quad y = r\sin\varphi, \quad 0 \leqslant \varphi \leqslant 2\pi.$$

Mamy $r^2 - z^2 = 1$. Zatem możemy przyjąć

$$r = \cosh\psi, \quad z = \sinh\psi, \quad \psi \in \mathbb{R}.$$

Ostatecznie uzyskujemy

$$\begin{aligned} x &= \cosh\psi\cos\varphi, \\ y &= \cosh\psi\sin\varphi, \\ z &= \sinh\psi. \end{aligned}$$

(b) Torus uzyskujemy obracając wokół osi z okrąg o promieniu r. Równanie okręgu w płaszczyźnie (y, z) ma postać

$$y = R + r\cos\varphi, \quad z = r\sin\varphi, \quad 0 \leqslant \varphi \leqslant 2\pi.$$

Niech d oznacza odległość punktu (x, y, z) torusa od osi z czyli $d = \sqrt{x^2 + y^2}$. Otrzymamy więc

$$d = R + r\cos\varphi, \quad z = r\sin\varphi.$$

Zatem

$$\begin{aligned} x &= d\cos\psi = (R + r\cos\varphi)\cos\psi, \\ y &= d\sin\psi = (R + r\cos\varphi)\sin\psi, \\ z &= r\sin\varphi. \end{aligned}$$

9.3 Pole powierzchni w $\mathbb{R}^3$

Definicja 9.3. *Niech S będzie powierzchnią sparametryzowaną przez funkcję $\Phi : D \xrightarrow{1-1} \mathbb{R}^3$ klasy C^1, gdzie $D \subset \mathbb{R}^2$. Tzn. $S = \Phi(D)$. Polem powierzchni nazywamy liczbę*

$$A(S) = \iint_D \|T_u \times T_v\| \, du \, dv.$$

Wyjaśnienie. Przeanalizujemy sumy całkowe całki określającej $A(S)$. Załóżmy, że D jest prostokątem podzielonym na n^2 małych prostokątów R_{ij}. Wtedy

$$\Phi(D) = \bigcup_{i,j=1}^{n} \Phi(R_{ij}).$$

Prostokąty R_{ij} nie są rozłączne, bo mogą mieć wspólne boki. Ale część wspólna każdych dwu zbiorów postaci $\Phi(R_{ij})$ ma miarę zero. Zatem

$$A(S) = \sum_{i,j=1}^{n} A(\Phi(R_{ij})).$$

Rozważamy mały prostokąt R w płaszczyźnie (u, v) o lewym dolnym rogu w (u, v) a prawym górnym w $(u+\Delta u, v+\Delta v)$. Obraz $\Phi(R)$ jest w przybliżeniu równoległobokiem o bokach

$$\begin{aligned}
\Phi(u + \Delta u, v) - \Phi(u, v) \quad &\approx \quad \left(\frac{\partial x}{\partial u}(u, v), \frac{\partial y}{\partial u}(u, v), \frac{\partial z}{\partial u}(u, v)\right) \Delta u, \\
\Phi(u, v + \Delta v) - \Phi(u, v) \quad &\approx \quad \left(\frac{\partial x}{\partial v}(u, v), \frac{\partial y}{\partial v}(u, v), \frac{\partial z}{\partial v}(u, v)\right) \Delta v.
\end{aligned}$$

Czyli

$$\begin{aligned}
\Phi(u + \Delta u, v) - \Phi(u, v) \quad &\approx \quad T_u(u, v) \, \Delta u, \\
\Phi(u, v + \Delta v) - \Phi(u, v) \quad &\approx \quad T_v(u, v) \, \Delta v.
\end{aligned}$$

Pole równoległoboku wynosi $A(\varphi(R)) \approx \|T_u \times T_v\| \, \Delta u \Delta v$. Rzeczywiście niech $u = \dfrac{T_u \times T_v}{\|T_u \times T_v\|}$. Wtedy

$$A(\varphi(R)) \approx |\det(u, T_u \, \Delta u, T_v \, \Delta v)| = \|T_u \times T_v\| \, \Delta u \, \Delta v.$$

Ostatecznie

$$A(S) = \sum_{i,j=1}^{n} \|T_u \times T_v\|_{\substack{u=u_{i-1}\\ v=v_{j-1}}} \Delta u_i \, \Delta v_j \xrightarrow[n]{} \iint\limits_D \|T_u \times T_v\| \, du \, dv.$$

Przykłady.

(a) $D = \{(r,\theta) : 0 \leqslant \theta \leqslant 2\pi,\ 0 \leqslant r \leqslant 1\}$. Określamy $\Phi(r,\theta) = (x,y,z)$, gdzie

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = r.$$

Obliczamy

$$T_r = (\cos\theta, \sin\theta, 1), \quad T_\theta = (-r\sin\theta, r\cos\theta, 0).$$

Zatem

$$T_r \times T_\theta = (-r\cos\theta, -r\sin\theta, r), \quad \|T_r \times T_\theta\| = r\sqrt{2}.$$

Dla $S = \Phi(D)$ mamy więc

$$A(S) = \int_0^{2\pi}\int_0^1 \|T_r \times T_\theta\| \, dr \, d\theta = \int_0^{2\pi}\int_0^1 r\sqrt{2} \, dr \, d\theta = \pi\sqrt{2}.$$

(b) Helikoida jest opisana przez

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = \theta,$$

dla parametrów spełniających $0 \leqslant r \leqslant 1$, $0 \leqslant \theta \leqslant 2\pi$. Dla ustalonej wartości kąta θ otrzymujemy odcinek prostopadły do osi z łączący punkty $(0,0,\theta)$ i $(\cos\theta, \sin\theta, \theta)$. Powstała powierzchnia przypomina wałek do mielenia mięsa. Mamy

$$T_r = (\cos\theta, \sin\theta, 0), \qquad T_\theta = (-r\sin\theta, r\cos\theta, 1),$$

zatem

$$T_r \times T_\theta = (\sin\theta, -\cos\theta, r), \quad \|T_r \times T_\theta\| = \sqrt{r^2+1}.$$

Dalej

$$A(S) = \int_0^{2\pi}\int_0^1 \sqrt{r^2+1} \, dr \, d\theta$$

$$= 2\pi \left[\frac{1}{2} r\sqrt{r^2+1} + \frac{1}{2}\log(r + \sqrt{r^2+1})\right]\Bigg|_{r=0}^{r=1} = \pi\left[\sqrt{2} + \log(\sqrt{2}+1)\right].$$

Załóżmy, że powierzchnia jest wykresem funkcji $z = g(x,y)$, dla $(x,y) \in D$. Wtedy

$$T_x \times T_y = \left(-\frac{\partial g}{\partial x}, -\frac{\partial g}{\partial y}, 1\right),$$

zatem

$$A(S) = \iint\limits_D \sqrt{1 + \left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2}\, dx\, dy.$$

Przykład. Obliczymy pole półsfery o promieniu 1. Mamy

$$z = \sqrt{1 - x^2 - y^2}, \quad D = \{(x,y) \,:\, x^2 + y^2 \leqslant 1\}.$$

Niech

$$D_R = \{(x,y) \,:\, x^2 + y^2 \leqslant R^2\}, \quad R < 1.$$

Wtedy

$$\frac{\partial z}{\partial x} = \frac{-x}{\sqrt{1 - x^2 - y^2}}, \quad \frac{\partial z}{\partial y} = \frac{-y}{\sqrt{1 - x^2 - y^2}}.$$

Zatem

$$1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 = \frac{1}{1 - x^2 - y^2}.$$

Otrzymujemy

$$\begin{aligned} A(S_R) &= \iint\limits_{D_R} \frac{dx\, dy}{\sqrt{1 - x^2 - y^2}} = \int_0^{2\pi} \int_0^R \frac{r\, dr\, d\theta}{\sqrt{1 - r^2}} \\ &= 2\pi \left(-\sqrt{1 - r^2}\right)\Big|_{r=0}^{r=R} = 2\pi \left(1 - \sqrt{1 - R^2}\right) \xrightarrow[R \to 1^-]{} 2\pi. \end{aligned}$$

9.4 Całki powierzchniowe funkcji skalarnych (niezorientowane)

Rozważamy powierzchnię S sparametryzowaną za pomocą funkcji $\Phi : D \to \mathbb{R}^3$, $\Phi(D) = S$,

$$\Phi(u,v) = (x(u,v), y(u,v), z(u,v)).$$

Dla funkcji ciągłej $f(x,y,z)$ określonej na S definiujemy całkę wzorem

$$\iint\limits_S f(x,y,z)\,dS = \iint\limits_D f(x(u,v),y(u,v),z(u,v))\,\|T_u \times T_v\|\,du\,dv.$$

Po rozpisaniu otrzymujemy wyrażenie

$$\iint\limits_D f(x(u,v),y(u,v),z(u,v))\sqrt{\left(\frac{\partial(y,z)}{\partial(u,v)}\right)^2+\left(\frac{\partial(z,x)}{\partial(u,v)}\right)^2+\left(\frac{\partial(x,y)}{\partial(u,v)}\right)^2}\,du\,dv,$$

przy czym

$$\frac{\partial(f,g)}{\partial(u,v)} = \begin{vmatrix} \dfrac{\partial f}{\partial u} & \dfrac{\partial f}{\partial v} \\ \dfrac{\partial g}{\partial u} & \dfrac{\partial g}{\partial v} \end{vmatrix}.$$

9.4.1 Interpretacja całki powierzchniowej

Przypuśćmy, że funkcja $\varrho(x,y,z)$ opisuje gęstość masy powierzchni S w punkcie (x,y,z). Chcemy obliczyć całkowitą masę powierzchni. Załóżmy, że D jest prostokątem. Dzielimy D na n^2 mniejszych prostokątów D_{ij}. Oznaczmy $S_{ij} = \Phi(D_{ij})$. Symbol $A(S_{ij})$ oznacza pole powierzchni fragmentu S_{ij}. Dla dużych wartości n fragment S_{ij} jest "mały". Uznajemy, że gęstość masy na S_{ij} jest stała i wynosi $\varrho(\Phi(u_i,v_j))$, gdzie $(u_i,v_j) \in D_{ij}$ (np. (u_i,v_j) jest prawym górnym rogiem prostokąta D_{ij}). Całkowita masa wynosi w przybliżeniu

$$\begin{aligned}\sum_{ij,j=1}^{n} \varrho(\Phi(u_i,v_j))\,A(S_{ij}) &\approx \sum_{ij,j=1}^{n} \varrho(\Phi(u_i,v_j))\,\|T_u\times T_v\|\Big|_{\substack{u=u_i\\ v=v_j}}\,\Delta u_i\,\Delta v_j \\ &\underset{n}{\longrightarrow} \iint\limits_D \varrho(\Phi(u,v))\,\|T_u\times T_v\|\,du\,dv \\ &= \iint\limits_D \varrho(x(u,v),y(u,v),z(u,v))\,\|T_u\times T_v\|\,du\,dv.\end{aligned}$$

Jeśli $\varrho(x,y,z)\equiv 1$, to $A(S) = \displaystyle\iint\limits_S dS$.

Przykłady.

(a) Rozważamy funkcję $f(x,y,z) = \sqrt{x^2+y^2+1}$ i helikoidę S określoną przez

$$x = r\cos\varphi,\ y = r\sin\varphi,\ z = \theta, \qquad 0 \leqslant \theta \leqslant 2\pi,\ 0 \leqslant r \leqslant 1.$$

Wtedy $\|T_r \times T_\theta\| = \sqrt{r^2+1}$. Zatem

$$\iint_S \sqrt{x^2+y^2+1}\, dS = \int_0^{2\pi}\int_0^1 \sqrt{r^2+1}\sqrt{r^2+1}\, dr\, d\theta = 2\pi\left(\frac{1}{3}+1\right) = \frac{8\pi}{3}.$$

(b) $\iint_S z^2\, dS$, gdzie S jest sferą jednostkową $x^2+y^2+z^2=1$. Można użyć do obliczeń współrzędnych sferycznych. Inaczej, zauważamy, że

$$\iint_S z^2\, dS = \iint_S x^2\, dS = \iint_S y^2\, dS.$$

Zatem

$$\iint_S z^2\, dS = \frac{1}{3}\iint_S (x^2+y^2+z^2)\, dS = \frac{1}{3}\iint_S dS = \frac{1}{3}A(S) = \frac{4\pi}{3}.$$

Przypuśćmy, że powierzchnia S jest wykresem funkcji $z = g(x,y)$, dla $(x,y) \in D$. Wtedy

$$\iint_S f(x,y,z)\, dS = \iint_D f(x,y,g(x,y))\sqrt{1+\left(\frac{\partial g}{\partial x}\right)^2+\left(\frac{\partial g}{\partial y}\right)^2}\, dx\, dy. \tag{9.1}$$

Przykład. Powierzchnia S jest określona przez $z = x^2+y$ dla (x,y) z prostokąta D opisanego przez warunki $0 \leqslant x \leqslant 1$ i $-1 \leqslant y \leqslant 1$. Wtedy

$$\begin{aligned}\iint_S x\, dS &= \int_0^1\int_{-1}^1 x\sqrt{4x^2+2}\, dy\, dx = 2\sqrt{2}\int_0^1 x\sqrt{2x^2+1}\, dx \\ &= 2\sqrt{2}\,\frac{1}{6}(2x^2+1)^{3/2}\Big|_0^1 = \frac{\sqrt{2}}{3}[3\sqrt{3}-1].\end{aligned}$$

Rozważmy wykres $z = g(x,y)$. Równanie powierzchni ma postać

$$\Phi(x,y,z) := -g(x,y)+z = 0,$$

tzn. powierzchnia jest poziomicą funkcji Φ. Wiemy, że wektorem normalnym do powierzchni w punkcie (x, y, z) jest gradient funkcji Φ podzielony przez swoją długość, czyli wektor

$$n = \frac{\left(-\frac{\partial g}{\partial x}, -\frac{\partial g}{\partial y}, 1\right)}{\sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1}}.$$

Niech θ oznacza kąt pomiędzy wektorem n i wektorem $(0, 0, 1)$. Wtedy

$$\cos\theta = \frac{1}{\sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1}}.$$

Zatem

$$\iint\limits_S f(x, y, z)\, dS = \iint\limits_D f(x, y, g(x, y)) \frac{1}{\cos\theta}\, dx\, dy,$$

gdzie θ jest kątem pomiędzy wektorem normalnym i dodatnią półosią z.

Uwaga. Dla małego obszaru ΔA w płaszczyźnie (x, y) pole powierzchni fragmentu ΔS odpowiadającego ΔA wynosi w przybliżeniu

$$\text{Pole}(\Delta S) \approx \frac{\text{Pole}(\Delta A)}{\cos\theta}.$$

Przykład. Obliczyć $\iint\limits_S x\, dS$, gdzie S jest trójkątem o wierzchołkach $(1, 0, 0)$, $(0, 1, 0)$ i $(0, 0, 1)$. Równanie powierzchni to $x+y+z = 1$, czyli $z = 1-x-y$, dla (x, y) z trójkąta D w płaszczyźnie (x, y) opisanego przez $x, y \geqslant 0$ i $x+y \leqslant 1$. Mamy $n = \frac{1}{\sqrt{3}}(1, 1, 1)$. Zatem

$$\iint\limits_S x\, dS = \int\limits_0^1 dx \int\limits_0^{1-x} x\sqrt{3}\, dy.$$

Inaczej:

$$\iint\limits_S x\, dS = \frac{1}{3} \iint\limits_S \underbrace{(x+y+z)}_{1}\, dS = \frac{1}{3} A(S) = \frac{1}{3}\,\frac{\sqrt{3}}{4}\,(\sqrt{2})^2 = \frac{\sqrt{3}}{6}.$$

9.5 Całki powierzchniowe pól wektorowych (zorientowane)

Definicja 9.4. *Niech $F(x, y, z)$ będzie polem wektorowym określonym na powierzchni $S = \Phi(D)$, $\Phi : D \to \mathbb{R}^3$. Określamy całkę powierzchniową zorientowaną wzorem*

$$\iint\limits_{S_\Phi} F \circ dS = \iint\limits_{D} F \circ (T_u \times T_v)\, du\, dv,$$

gdzie w całce po prawej stronie $F = F(\Phi(u, v))$.

Uwaga. Możemy powiązać tę całkę z całką niezorientowaną. Załóżmy, że $T_u \times T_v \neq 0$. Wtedy dla $n = \dfrac{T_u \times T_v}{\|T_u \times T_v\|}$ mamy

$$\iint\limits_{D} F \circ (T_u \times T_v)\, du\, dv \;=\; \iint\limits_{D} (F \circ n)\, \|T_u \times T_v\|\, du\, dv \;=\; \iint\limits_{S} (F \circ n)\, dS.$$

Otrzymujemy więc

$$\iint\limits_{S_\Phi} F \circ dS = \iint\limits_{S} (F \circ n)\, dS.$$

Zwrot wektora normalnego n zależy od parametryzacji, nawet od kolejności zmiennych u i v, bo $T_u \times T_v = -(T_v \times T_u)$.

Przykład. Niech S będzie sferą jednostkową oraz $F(x, y, z) = (x, y, z)$. Użyjemy współrzędnych sferycznych.

$$\begin{aligned} x &= \sin\varphi\cos\psi, \\ y &= \sin\varphi\sin\psi, \\ z &= \cos\varphi. \end{aligned}$$

$$\begin{aligned} T_\varphi &= (\cos\varphi\cos\psi, \cos\varphi\sin\psi, -\sin\varphi), \\ T_\psi &= (-\sin\varphi\cos\psi, \sin\varphi\cos\psi, 0), \end{aligned}$$

zatem

$$T_\varphi \times T_\psi = (\sin^2\varphi\cos\psi, \sin^2\varphi\sin\psi, \cos\varphi\sin\varphi) = \sin\varphi\,(x, y, z).$$

Wektor normalny to

$$n = (x, y, z).$$

$$\iint\limits_{S_\Phi} F \circ dS = \iint\limits_{S} (n \circ n)\, dS = \iint\limits_{S} dS = 4\pi.$$

Definicja 9.5. *Powierzchnią zorientowaną nazywamy powierzchnię dwustronną, w której jedna strona została określona jako zewnętrzna (dodatnia) a druga jako wewnętrzna (ujemna).*

W każdym punkcie powierzchni mamy dwa wektory normalne n_1 i n_2, $n_2 = -n_1$. Załóżmy, że w każdym punkcie wybraliśmy jeden wektor normalny n tak, że wybrane wektory wskazują jedną stronę powierzchni. Niech $\Phi : D \to \mathbb{R}^3$ będzie parametryzacją powierzchni S. Wektor $T_u \times T_v$ jest prostopadły do powierzchni S w punkcie $\Phi(u, v)$. Zatem $T_u \times T_v = \lambda(u, v)\, n$, gdzie n jest wybranym wektorem normalnym w punkcie $\Phi(u, v)$. Jeśli $\lambda(u, v) > 0$ dla $(u, v) \in D$, to mówimy, że parametryzcja jest zgodna z orientacją. Jeśli $\lambda(u, v) < 0$ dla $(u, v) \in D$, to parametryzacja jest niezgodna z orientacją (jest przeciwna).

Niech S będzie wykresem funkcji $z = g(x, y)$. Domyślna orientacja jest wyznaczona przez

$$n = \frac{\left(-\frac{\partial g}{\partial x}, -\frac{\partial g}{\partial y}, 1\right)}{\sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1}}.$$

To oznacza, że górna część wykresu jest zewnętrzna.

Twierdzenie 9.6. *Niech S będzie powierzchnią zorientowaną, a Φ_1 i Φ_2 dwiema parametryzacjami gładkimi zachowującymi orientację. Wtedy*

$$\iint\limits_{S_{\Phi_1}} F \circ dS = \iint\limits_{S_{\Phi_2}} F \circ dS.$$

Jeśli Φ_1 zachowuje orientację, a Φ_2 zmienia orientację, to

$$\iint\limits_{S_{\Phi_1}} F \circ dS = -\iint\limits_{S_{\Phi_2}} F \circ dS.$$

Jeśli $f(x, y, z)$ jest funkcją ciągłą na S, to

$$\iint\limits_{S_{\Phi_1}} f\, dS = \iint\limits_{S_{\Phi_2}} f\, dS,$$

tzn. całka niezorientowana nie zależy od wyboru parametryzacji.

Dowód. Rozważamy dwie parametryzacje powierzchni S

$$\Phi_1 : D_1 \to \mathbb{R}^3 \quad \text{oraz} \quad \Phi_2 : D_2 \to \mathbb{R}^3, \quad \text{dla} \quad D_1, D_2 \subset \mathbb{R}^2.$$

Dla ustalonego punktu (x, y, z) powierzchni mamy

$$(x, y, z) = \Phi_1(u, v) = \Phi_2(u', v')$$

dla jedynych $(u, v) \in D_1$ oraz $(u', v') \in D_2$. Uzyskujemy w ten sposób odwzorowanie $g : D_1 \to D_2$

$$(u', v') = g(u, v).$$

Załóżmy, że g jest klasy C^1. Mamy

$$\Phi_1(u, v) = \Phi_2(u', v') = \Phi_2(g(u, v)).$$

Obliczamy macierz pochodnych obu stron.

$$D\Phi_1(u, v) = [T_u\ T_v] = D\Phi_2(\underbrace{g(u, v)}_{(u',v')})\, Dg(u, v) = [T_{u'}\ T_{v'}]\, Dg(u, v).$$

Lemat 9.7. *a i b są wektorami w $\mathbb{R}^3$ a M macierzą wymiaru 2×2. Jeśli $[c\ d] = [a\ b]\, M$, to $c \times d = \det M \cdot (a \times b)$.*

Dowód lematu. Niech $M = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$. Wtedy

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} = \begin{bmatrix} \alpha a + \gamma b & \beta a + \delta b \end{bmatrix} = \begin{bmatrix} c & d \end{bmatrix}.$$

$$\begin{aligned} c \times d = (\alpha a + \gamma b) \times (\beta a + \delta b) &= \alpha\delta\, a \times b + \gamma\beta\, b \times a \\ &= (\alpha\delta - \gamma\beta)\, a \times b = \det M \cdot (a \times b). \end{aligned}$$

□

Z lematu otrzymujemy

$$T_u \times T_v = \det Dg(u, v)\, T_{u'} \times T_{v'}.$$

Wykonujemy obliczenia stosując w trakcie podstawienie $(u', v') = g(u, v)$.

$$\begin{aligned}\iint_{S_{\Phi_2}} F \circ dS = \iint_{D_2} & F(\Phi_2(u', v')) \circ (T_{u'} \times T_{v'})\, du'\, dv' \\ &= \iint_{D_1} F(\Phi_2(g(u, v)) \circ (T_{u'} \times T_{v'})\ |\det Dg(u, v)|\ du\, dv.\end{aligned}$$

Załóżmy, że $\det Dg(u, v) > 0$ dla $(u, v) \in D_1$. Wtedy w wyniku otrzymujemy

$$\iint_{D_1} F(\Phi_1(u, v)) \circ (T_u \times T_v)\, du\, dv = \iint_{S_{\Phi_1}} F \circ dS.$$

Jeśli $\det Dg(u, v) < 0$ dla $(u, v) \in D_1$, to w wyniku dostaniemy

$$-\iint_{S_{\Phi_1}} F \circ dS.$$

Dalej w wyniku tego samego podstawienia mamy

$$\begin{aligned}\iint_{S_{\Phi_2}} f\, dS = \iint_{D_2} & f(\Phi_2(u', v'))\|T_{u'} \times T_{v'}\|\, du'\, dv' \\ &= \iint_{D_1} f(\Phi_1(u, v))\, \underbrace{\|T_{u'} \times T_{v'}\|\, |\det D_g(u, v)|}_{\|T_u \times T_v\|}\, du\, dv = \iint_{S_{\Phi_1}} f\, dS.\end{aligned}$$

$\square$

9.5.1 Interpretacja fizyczna całki powierzchniowej zorientowanej

Zbadamy sumy Riemanna całki

$$\iint_S F \circ dS = \iint_D F(\Phi(u, v)) \circ (T_u \times T_v)\, du\, dv.$$

Niech R będzie małym prostokątem leżącym w D o bokach Δu i Δv równoległych do osi współrzędnych. Lewy dolny róg prostokąta R oznaczymy przez (u, v). Obraz $\Phi(R)$ prostokąta jest w przybliżeniu równoległobokiem o bokach $T_u \Delta u$ i $T_v \Delta v$. Rozważmy wielkość

$$F(\Phi(u, v)) \circ (T_u \Delta u \times T_v \Delta v).$$

Ze wzoru

$$\det(a,b,c) = a \circ (b \times c), \quad \text{dla } a,b,c \in \mathbb{R}^3.$$

wynika, że jest to plus minus objętość równoległościanu rozpiętego przez wektory $F(\Phi(u,v))$, $T_u\Delta u$ i $T_v\Delta v$. Zakładamy, że powierzchnia S jest zorientowana i parametryzacja Φ jest zgodna z orientacją. Jeśli wektor F jest skierowany w stronę dodatnią powierzchni, to otrzymujemy objętość równoległościanu, a jeśli w stronę ujemną, to otrzymamy minus objętość równoległościanu.

Niech F oznacza prędkość przepływu jakiegoś płynu w punkcie $(x,y,z) = \Phi(u,v)$. Wtedy F wskazuje kierunek przepływu a liczba

$$|F(\Phi(u,v)) \circ (T_u\Delta u \times T_v\Delta v)|$$

mierzy ilość płynu jaki przepłynął przez fragment powierzchni $\Phi(R)$ w jednostce czasu. Ilość płynu jaka przepłynie w jednostce czasu jest równa zatem plus minus objętości równoległościanu rozpiętego przez $F(\Phi(u,v))$, $T_u\Delta u$ i $T_v\Delta v$. Znak zależy od tego, czy siła F jest skierowana na zewnątrz czy do wewnątrz powierzchni. Reasumując $F \circ (T_u \times T_v)\,\Delta u\,\Delta v$ jest prędkością przepływu na stronę zewnętrzną przez fragment powierzchni $\Phi(R)$. Podzielmy obszar D na małe prostokąty R_{ij}. Wtedy

$$\sum_{i,j=1}^{n} F \circ (T_u \times T_v)\Big|_{\substack{u=u_{i-1}\\ v=v_{j-1}}} \Delta u_i\,\Delta v_j$$

jest sumaryczną prędkością przepływu na zewnątrz powierzchni S. Ostatecznie całka

$$\iint_S F \circ dS$$

jest prędkością przepływu na zewnątrz powierzchni S.

Całka powierzchniowa służy do obliczania przepływu ciepła. Niech $T(x,y,z)$ oznacza temperaturę w punkcie (x,y,z). Rozważmy pole wektorowe

$$F = -k\nabla T = -k\left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z}\right),$$

gdzie k jest stałym współczynnikiem dodatnim, zależnym od ośrodka. Wtedy całka $\iint_S F \circ dS$ opisuje tempo przepływu ciepła na zewnątrz powierzchni S.

Przykład. $T(x,y,z) = x^2 + y^2 + z^2$, $S = \{(x,y,z) : x^2 + y^2 + z^2 = 1\}$. Załóżmy, że $k = 1$, czyli

$$F = -\nabla T = -2(x,y,z).$$

Wtedy $F \circ n = -2$ oraz

$$\iint\limits_S F \circ dS = \iint\limits_S (F \circ n)\, dS = -8\pi.$$

9.5.2 Całka powierzchniowa dla wykresów funkcji

Przypuśćmy, że S jest wykresem funkcji $z = g(x,y)$ dla $(x,y) \in D$. Stosujemy parametryzację

$$x := x, \ y := y, \ z = g(x,y).$$

Wtedy

$$T_x = \left(1, 0, \frac{\partial g}{\partial x}\right), \quad T_y = \left(1, 0, \frac{\partial g}{\partial y}\right), \quad T_x \times T_y = \left(-\frac{\partial g}{\partial x}, -\frac{\partial g}{\partial y}, 1\right).$$

Dla pola wektorowego $F = (P, Q, R)$ w $\mathbb{R}^3$ otrzymujemy

$$\iint\limits_S F \circ dS = \iint\limits_D F \circ (T_x \times T_y)\, dx\, dy = \iint\limits_D \left[-P\frac{\partial g}{\partial x} - Q\frac{\partial g}{\partial y} + R\right] dx\, dy, \quad (9.2)$$

przy czym funkcje P, Q i R są obliczone w $(x, y, g(x,y))$.

10 Wzór Greena

Twierdzenie podaje związek pomiędzy całką krzywoliniową zorientowaną, wzdłuż krzywej zamkniętej C w $\mathbb{R}^2$ a całką podwójną po obszarze D ograniczonym przez tę krzywą.

Definicja 10.1. *Obszar D nazywamy elementarnym typu I jeśli*

$$D = \{(x,y) : a \leqslant x \leqslant b, \ \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\},$$

gdzie φ_1 i φ_2 są funkcjami ciągłymi. D nazywamy elementarnym typu II jeśli

$$D = \{(x,y) : c \leqslant y \leqslant d, \ \psi_1(y) \leqslant x \leqslant \psi_2(y)\}.$$

D nazywamy obszarem elementarnym, jeśli jest jednocześnie elementarny typu I i typu II.

Brzeg każdego obszaru orientujemy przeciwnie do wskazówek zegara (analogowego).

Lemat 10.2. *Niech $P(x, y)$ będzie funkcją klasy C^1 na obszarze D typu I. Wtedy*

$$\int_C P\,dx = -\iint_D \frac{\partial P}{\partial y}\,dx\,dy,$$

gdzie C jest brzegiem obszaru D.

Uwaga. Przypuśćmy, że pole wektorowe F w $\mathbb{R}^3$ ma postać $F = (P, 0, 0)$. Wtedy

$$\int_C F \circ ds = \int_C P\,dx.$$

Dowód. Brzeg obszaru D składa się dwu odcinków pionowych odpowiadających $x = a$ i $x = b$ oraz z dwu fragmentów wykresu C_1^+ i C_2^- dla funkcji φ_1 i φ_2. Na odcinkach pionowych mamy $dx = 0$. Zatem

$$\int_C P\,dx = \int_{C_1^+} P\,dx - \int_{C_2^+} P\,dx.$$

$$\begin{aligned}\iint_D \frac{\partial P}{\partial y}\,dx\,dy &= \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} \frac{\partial P}{\partial y}\,dy\,dx = \int_a^b P(x, y)\Big|_{y=\varphi_1(x)}^{y=\varphi_2(x)}\,dx \\ &= \int_a^b P(x, \varphi_2(x))\,dx - \int_a^b P(x, \varphi_1(x))\,dx \\ &= \int_{C_2^+} P(x, y)\,dx - \int_{C_1^+} P(x, y)\,dx = -\int_C P\,dx.\end{aligned}$$

□

Zamieniając rolami P i Q oraz x i y otrzymamy

Lemat 10.3. *Niech $Q(x, y)$ będzie funkcją klasy C^1 na obszarze D typu II. Wtedy*

$$\int_C Q\,dy = \iint_D \frac{\partial Q}{\partial x}\,dx\,dy,$$

gdzie C jest brzegiem obszaru D.

Uwaga. Zmiana znaku z „−" na „+" wynika z tego, że zamieniając rolami x i y zmieniamy orientację.

Dowód. Brzeg obszaru D składa się dwu odcinków poziomych odpowiadających $y = c$ i $y = d$ oraz z dwu fragmentów wykresu C_1^- i C_2^+ dla funkcji ψ_1 i ψ_2. Na odcinkach poziomych mamy $dy = 0$. Zatem

$$\begin{aligned}\iint\limits_D \frac{\partial Q}{\partial x}\,dx\,dy &= \int\limits_c^d \int\limits_{\psi_1(y)}^{\psi_2(y)} \frac{\partial Q}{\partial x}\,dx\,dy = \int\limits_c^d Q(x,y)\Big|_{x=\psi_1(y)}^{x=\psi_2(y)}\,dy \\ &= \int\limits_c^d Q(\psi_2(y),y)\,dy - \int\limits_c^d Q(\psi_1(y),y)\,dy \\ &= \int\limits_{C_2^+} Q(x,y)\,dy + \int\limits_{C_1^-} Q(x,y)\,dy = \int\limits_C Q\,dy.\end{aligned}$$

□

Lematy dają w wyniku

Twierdzenie 10.4 (wzór Greena)**.** *Niech D będzie obszarem elementarnym z brzegiem C zorientowanym dodatnio. Niech $P(x,y)$ i $Q(x,y)$ będą funkcjami klasy C^1 określonymi na D. Wtedy*

$$\int\limits_C P(x,y)\,dx + Q(x,y)\,dy = \iint\limits_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)dx\,dy.$$

Uwaga. Wzór Greena jest prawdziwy dla obszarów, które można podzielić na kilka obszarów elementarnych. Przypuśćmy, że $D = D_1 \cup D_2$ oraz wnętrza obszarów D_1 i D_2 są rozłączne. Niech C_0 oznacza część wspólną brzegów obszarów ∂D_1 i ∂D_2. Wtedy

$$\begin{aligned}\iint\limits_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)dx\,dy &= \iint\limits_{D_1} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)dx\,dy + \iint\limits_{D_2} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)dx\,dy \\ &= \int\limits_{\partial D_1} P\,dx + Q\,dy + \int\limits_{\partial D_2} P\,dx + Q\,dy = \int\limits_{\partial D} P\,dx + Q\,dy,\end{aligned}$$

bo całki wzdłuż C_0 zniosą się.

Uwaga. Przypuśćmy, że funkcja P zeruje się na brzegu obszaru D, ale $\dfrac{\partial P}{\partial y}$ nie jest zerowa w D. Otrzymamy

$$\iint\limits_D \frac{\partial P}{\partial y}\,dx\,dy = -\int\limits_{\partial D} P(x,y)\,dx = 0.$$

Np. niech $D = \{(x,y)\ :\ x^2+y^2 \leqslant 1\}$ i $P(x,y) = 1-x^2-y^2$.

Twierdzenie 10.5. *Niech D będzie obszarem, dla którego można zastosować wzór Greena. Wtedy*

$$A(D) = \frac{1}{2}\int\limits_{\partial D}(x\,dy - y\,dx).$$

Dowód. Przyjmijmy $P(x,y) = -y$ oraz $Q(x,y) = x$. Mamy

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 2.$$

Zatem

$$\int\limits_{\partial D}(x\,dy - y\,dx) = \iint\limits_D 2\,dx\,dy = 2\,A(D).$$

$\square$

Przykład. Hipocykloida jest określona równaniem

$$x^{2/3} + y^{2/3} = a^{2/3}, \qquad a > 0.$$

Chcemy obliczyć pole obszaru ograniczonego przez tę krzywą. Zastosujemy parametryzację

$$x^{1/3} = a^{1/3}\cos\theta, \quad y^{1/3} = a^{1/3}\sin\theta, \qquad 0 \leqslant \theta \leqslant 2\pi.$$

Wtedy

$$x = a\cos^3\theta, \quad y = a\sin^3\theta.$$

Dalej

$$A(D) = \frac{1}{2}\int\limits_{\partial D}(x\,dy - y\,dx) = \frac{a^2}{2}\int\limits_0^{2\pi}[3\cos^3\theta\sin^2\theta\cos\theta + 3\sin^3\theta\cos^2\theta\sin\theta]\,d\theta$$

$$= \frac{3a^2}{2}\int\limits_0^{2\pi}\sin^2\theta\cos^2\theta\,d\theta = \frac{3a^2}{8}\int\limits_0^{2\pi}\sin^2 2\theta\,d\theta = \frac{3a}{16}\int\limits_0^{2\pi}[\sin^2 2\theta + \cos^2 2\theta]\,d\theta = \frac{3\pi}{8}a^2.$$

10.1 Rotacja

Dla pola wektorowego $F = (P, Q)$ na płaszczyźnie wielkość

$$\operatorname{curl} F = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$$

nazywamy rotacją. Wyobrażamy sobie, że $F(x, y)$ określa prędkość przepływu w punkcie (x, y). Załóżmy, że obiekt pod wpływem działania $F(x, y)$ przesuwa się w kierunku poziomym (równolegle do osi x) z punktu (x, y) o $\Delta x > 0$. Wtedy obiekt ten uzyska przyrost prędkości w górę (czyli w lewo) o

$$\frac{\partial Q}{\partial x}\Delta x \approx Q(x + \Delta x, y) - Q(x, y).$$

Czyli przyrost prędkości w górę na jednostkę przesunięcia w prawo wynosi $\frac{\partial Q}{\partial x}$. Tzn. wielkość $\frac{\partial Q}{\partial x}$ określa tendencję do obrotu obiektu w kierunku dodatnim (w lewo). Podobnie przy przesunięciu obiektu w górę o Δy z punktu (x, y) obiekt uzyskuje przyrost prędkości w prawo wynoszący

$$P(x, y + \Delta y) - P(x, y) \approx \frac{\partial P}{\partial y}(x, y)\,\Delta y.$$

Tzn. wielkość $\frac{\partial P}{\partial y}$ określa tendencję do skrętu w kierunku ujemnym. Reasumując wielkość $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$ określa wypadkową tendencję do skrętu obiektu w kierunku dodatnim.

11 Twierdzenie Stokesa

Twierdzenie Stokesa podaje związek pomiędzy całką krzywoliniową zorientowaną wzdłuż krzywej zamkniętej C w $\mathbb{R}^3$ a całką powierzchniową zorientowaną po powierzchni S, dla której krzywa C jest brzegiem, tzn. $C = \partial S$. Na brzegu wprowadzamy orientację zgodną z orientacją powierzchni, tzn. idąc wzdłuż brzegu z głową podniesioną w kierunku dodatniej strony powierzchni, powierzchnia znajduje się po naszej lewej stronie. Twierdzenie Stokesa przypomina twierdzenie Greena tyle, że powierzchnia S nie musi być płaska. Rozważymy przypadek, gdy S jest wykresem funkcji $z = g(x, y)$, dla $(x, y) \in D \subset \mathbb{R}^2$.

Twierdzenie 11.1 (wzór Stokesa)**.** *Niech S będzie zorientowaną powierzchnią będącą wykresem funkcji $z = g(x, y)$, $(x, y) \in D$, gdzie g jest klasy C^2. Zakładamy, że do obszaru D można zastosować wzór Greena. Wtedy*

$$\int_{\partial S} F \circ ds = \iint_S \mathrm{curl}\ F \circ dS,$$

gdzie $\mathrm{curl}\, F$ *jest polem wektorowym określonym dla* $F = (P, Q, R)$ *wzorem*

$$\mathrm{curl}\, F = \begin{vmatrix} e_1 & e_2 & e_3 \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right).$$

Uwaga. Jeśli $R \equiv 0$, oraz P i Q nie zależą od z, to $\mathrm{curl}\, F = (0, 0, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})$.

Dowód. Ze wzoru (9.2) mamy

$$\iint_S \mathrm{curl}\, F \circ dS$$
$$= \iint_D \left[-\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \frac{\partial z}{\partial x} - \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \frac{\partial z}{\partial y} + \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right] dx\, dy$$

Niech $\sigma(t) = (x(t), y(t))$ będzie parametryzacją brzegu ∂D, dla $a \leqslant t \leqslant b$. Wtedy

$$\eta(t) = (x(t), y(t), g(x(t), y(t)), \quad a \leqslant t \leqslant b$$

jest parametryzacją brzegu ∂S. Zatem

$$\begin{aligned}
\int\limits_{\partial S} F \circ ds &= \int\limits_a^b \left[P \frac{dx}{dt} + Q \frac{dy}{dt} + R \left(\frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt} \right) \right] dt \\
&= \int\limits_a^b \left[\left(P + R \frac{\partial z}{\partial x} \right) \frac{dx}{dt} + \left(Q + R \frac{\partial z}{\partial y} \right) \frac{dy}{dt} \right] \\
&= \int\limits_{\partial D} \left(P + R \frac{\partial z}{\partial x} \right) dx + \left(Q + R \frac{\partial z}{\partial y} \right) dy \\
&= \iint\limits_D \left[\frac{\partial Q}{\partial x} + \frac{\partial Q}{\partial z} \frac{\partial z}{\partial x} + \left(\frac{\partial R}{\partial x} + \frac{\partial R}{\partial z} \frac{\partial z}{\partial x} \right) \frac{\partial z}{\partial y} + R \frac{\partial^2 z}{\partial x \partial y} \right. \\
&\quad \left. - \frac{\partial P}{\partial y} - \frac{\partial P}{\partial z} \frac{\partial z}{\partial y} - \left(\frac{\partial R}{\partial y} + \frac{\partial R}{\partial z} \frac{\partial z}{\partial y} \right) \frac{\partial z}{\partial x} - R \frac{\partial^2 z}{\partial y \partial x} \right] dx\, dy \\
&= \iint\limits_S \operatorname{curl} F \circ dS
\end{aligned}$$

□

Przykłady.

(a) $F(x, y, z) = (ye^z, xe^z, xye^z)$. Przypuśćmy, że powierzchnia S spełnia warunki twierdzenia. Mamy

$$\operatorname{curl} F = \begin{vmatrix} e_1 & e_2 & e_3 \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ ye^z & xe^z & xye^z \end{vmatrix} = 0.$$

Zatem

$$\int\limits_{\partial S} F \circ ds = \iint\limits_S \operatorname{curl} F \circ dS = 0.$$

(b) C jest krzywą będącą przecięciem cylindra $x^2 + y^2 = 1$ oraz płaszczyzny $x + y + z = 1$. Orientacja krzywej wyznacza dodatnią orientację po zrzutowaniu na okrąg $x^2 + y^2 = 1$. C jest brzegiem powierzchni S, która jest wyznaczona przez wykres funkcji $z = 1 - x - y$ określonej na kole $x^2 + y^2 \leqslant 1$. Obliczamy

$$\int\limits_C -y^3\,dx + x^3\,dy - z^3\,dz = \iint\limits_S \operatorname{curl}(-y^3, x^3, -z^3) \circ dS.$$

Mamy

$$\operatorname{curl}(-y^3, x^3, -z^3) = \begin{vmatrix} e_1 & e_2 & e_3 \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ -y^3 & x^3 & -z^3 \end{vmatrix} = (0, 0, 3(x^2 + y^2)).$$

Zatem ze wzoru (9.2) otrzymujemy

$$\int\limits_C -y^3\,dx + x^3\,dy - z^3\,dz = \iint\limits_{x^2+y^2\leqslant 1} 3(x^2+y^2)\,dx\,dy = \int\limits_0^{2\pi}\int\limits_0^1 3r^3\,dr\,d\theta = 2\pi \cdot \frac{3}{4} = \frac{3\pi}{2}.$$

Wzór Stokesa jest prawdziwy dla powierzchni sparametryzowanych, a nie tylko dla wykresów funkcji $z = g(x, y)$, $(x, y) \in D$. Pewna komplikacja dotyczy brzegu ∂S, gdy S jest sparametryzowana.

Przykład. Rozważmy sferę jednostkową S i parametryzację przez współrzędne sferyczne $0 \leqslant \varphi \leqslant \pi$, $0 \leqslant \psi \leqslant 2\pi$. Parametryzacja nie jest różnowartościowa.

Twierdzenie 11.2 (wzór Stokesa)**.** *Niech S będzie powierzchnią sparametryzowaną przez $\Phi : D \to S$, gdzie D jest obszarem w $\mathbb{R}^2$, do którego można zastosować wzór Greena. Zakładamy, że odwzorowanie Φ jest klasy C^1 i różnowartościowe. Brzeg obszaru D orientujemy dodatnio. Wtedy brzegiem powierzchni S jest $\partial S = \Phi(\partial D)$. Wprowadzamy orientację na ∂S przenosząc ją z ∂D. Tzn., jeśli $\sigma(t)$ jest parametryzcją brzegu ∂D zgodną z orientacją, to $\eta(t) = \Phi(\sigma(t))$ jest parametryzacją ∂S. Wtedy*

$$\int\limits_{\partial S} F \circ ds = \iint\limits_S \mathrm{curlF} \circ dS.$$

Przykład. Niech S będzie powierzchnią klosza $x^2 + y^2 + (z - 1)^2 = 2$, $z \geqslant 0$. Brzegiem klosza jest okrąg $x^2 + y^2 = 1$, $z = 0$. Dla pola $F = (y, -x, e^{xz})$

chcemy obliczyć

$$\iint_S \operatorname{curl} F \circ dS = \int_{\partial S} F \circ ds.$$

Parametryzujemy ∂S poprzez $x = \cos t$, $y = \sin t$, $z = 0$ dla $0 \leqslant t \leqslant 2\pi$. Otrzymujemy w wyniku

$$\int_0^{2\pi} (-\sin^2 t - \cos^2 t)\, dt = -2\pi.$$

11.1 Interpretacja rotacji curl F

Wybierzmy wektor jednostkowy n i punkt P przestrzeni $\mathbb{R}^3$. Niech F będzie polem wektorowym w $\mathbb{R}^3$. Symbolem S_r oznaczamy koło o promieniu r i środku w P, prostopadłe do wektora n. Ze wzoru Stokesa mamy

$$\int_{\partial S_r} F \circ ds = \iint_{S_r} \operatorname{curl} F \circ dS = \iint_{S_r} (\operatorname{curl} F \circ n) dS = [\operatorname{curl} F(Q_r) \circ n]\, A(S_r),$$

gdzie Q_r jest pewnym punktem w S_r. Otrzymujemy więc

$$\operatorname{curl} F(Q_r) \circ n = \frac{1}{A(S_r)} \int_{\partial S_r} F \circ ds = \frac{1}{A(S_r)} \int_{\partial S_r} (F \circ T)\, ds,$$

gdzie T jest jednostkowym wektorem stycznym do krzywej. Przechodząc do granicy, gdy $r \to 0^+$ otrzymamy

$$\operatorname{curl} F(P) \circ n = \lim_{r \to 0^+} \frac{1}{A(S_r)} \int_{\partial S_r} (F \circ T)\, ds.$$

11.2 Interpretacja całki $\int_C (F \circ T)\, ds$ dla krzywej zamkniętej C

(a) Załóżmy, że pole F jest styczne do krzywej C w kierunku zgodnym z orientacją C. Wtedy $F \circ T > 0$. Zatem $\int_C (F \circ T)\, ds > 0$.

(b) Jeśli pole F jest styczne do krzywej C w kierunku przeciwnym, to $F \circ T < 0$. Zatem $\int_C (F \circ T)\, ds < 0$.

(c) Przypuśćmy, że pole F jest prostopadłe do C. Wtedy $\int\limits_C (F \circ T)\, ds = 0$.

Ogólnie wielkość $\int\limits_C (F \circ T)\, ds$ oznacza ilość płynu przepływającego w jednostce czasu, w kierunku dodatnim wokół krzywej C, jeśli F oznacza prędkość przepływu. Wielkość $\int\limits_C F \circ ds$ nazywamy cyrkulacją pola F wokół krzywej C. Zatem $\operatorname{curl} F(P) \circ n$ jest cyrkulacją pola na jednostkę powierzchni w punkcie P w płaszczyźnie prostopadłej do wektora n. Przy ustalonym punkcie P niech

$$n = \frac{\operatorname{curl} F(P)}{\|\operatorname{curl} F(P)\|},$$

przy założeniu, że $\operatorname{curl} F(P) \neq 0$. Wtedy wielkość $\operatorname{curl} F(P) \circ n$ jest największa z możliwych. Tzn. w płaszczyźnie prostopadłej do wektora $\operatorname{curl} F(P)$ mamy największą tendencję do cyrkulacji.

Uwaga. Wzór Stokesa można zapisać następująco, po przejściu do całek niezorientowanych.

$$\int\limits_{\partial S} (F \circ T)\, ds = \iint\limits_S (\operatorname{curl} F \circ n)\, dS.$$

12 Wzór Gaussa-Ostrogradskiego

Dla pola wektorowego klasy $F = (F_1, F_2, F_3)$ w $\mathbb{R}^3$ klasy C^1 dywergencją nazywamy funkcję

$$\operatorname{div} F = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \circ (F_1, F_2, F_3).$$

Twierdzenie Gaussa-Ostrogradskiego mówi, że przepływ pola F na zewnątrz zorientowanej powierzchni zamkniętej jest równy całce potrójnej z dywergencji pola F po obszarze ograniczonym przez powierzchnię S.

Definicja 12.1. *Mówimy, że obszar $\Omega \subset \mathbb{R}^3$ jest elementarny, jeśli ma jedną z trzech postaci:*

(a) $\Omega = \{(x, y, z) \,:\, (x, y) \in D_1,\ \varphi_1(x, y) \leqslant z \leqslant \varphi_2(x, y)\},$

(b) $\Omega = \{(x,y,z) : (x,z) \in D_2,\ \psi_1(x,z) \leqslant y \leqslant \psi_2(x,z)\}$,

(c) $\Omega = \{(x,y,z) : (y,z) \in D_3,\ \eta_1(y,z) \leqslant x \leqslant \eta_2(y,z)\}$,

gdzie funkcje φ_1, φ_2, ψ_1, ψ_1, η_1, η_2 są ciągłe, a obszary D_1, D_2 i D_3 takie jak we wzorze Greena. Obszar Ω jest elementarny w trzech kierunkach, jeśli ma każdą z tych postaci.

Przykłady. Prostopadłościan $[a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ i kula w $\mathbb{R}^3$ są elementarne w trzech kierunkach.

Powierzchnię zamkniętą orientujemy domyślnie tak, że zewnętrzna część jest dodatnia. Jeśli S składa się z kilku części S_1, S_2, ..., S_n, to

$$\iint\limits_S F \circ dS = \iint\limits_{S_1} F \circ dS + \ldots + \iint\limits_{S_n} F \circ dS.$$

Przykład. Niech S będzie brzegiem sześcianu $[-1,1]^3$. Zewnętrzne wektory normalne do poszczególnych ścian mają postać

$$\begin{aligned} n_1 &= (0,0,-1), & n_2 &= (0,0,1), \\ n_3 &= (0,-1,0), & n_4 &= (0,1,0), \\ n_5 &= (-1,0,0), & n_6 &= (1,0,0). \end{aligned}$$

Zatem

$$\begin{aligned} &\iint\limits_S F \circ dS = \iint\limits_S (F \circ n)\, dS = \sum_{i=1}^{6} \iint\limits_{S_i} (F \circ n_i)\, dS \\ &= -\iint\limits_{S_1} F_3\, dS + \iint\limits_{S_2} F_3\, dS - \iint\limits_{S_3} F_2\, dS + \iint\limits_{S_4} F_2\, dS - \iint\limits_{S_5} F_1\, dS + \iint\limits_{S_6} F_1\, dS. \end{aligned}$$

Twierdzenie 12.2 (wzór Gaussa Ostrogradskiego)**.** *Niech Ω będzie obszarem elementarnym w trzech kierunkach w $\mathbb{R}^3$. Brzeg $\partial\Omega$ orientujemy dodatnio. Wtedy dla pola wektorowego F klasy C^1 w $\mathbb{R}^3$ mamy*

$$\iint\limits_{\partial\Omega} F \circ dS = \iiint\limits_{\Omega} \operatorname{div} F\, dx\, dy\, dz.$$

Dowód. Mamy

$$\iiint\limits_{\Omega} \operatorname{div} F\, dx\, dy\, dz = \iiint\limits_{\Omega} \frac{\partial F_1}{\partial x}\, dx\, dy\, dz + \iiint\limits_{\Omega} \frac{\partial F_2}{\partial y}\, dx\, dy\, dz + \iiint\limits_{\Omega} \frac{\partial F_3}{\partial z}\, dx\, dy\, dz.$$

$$\iint\limits_{\partial\Omega} F \circ dS = \iint\limits_{\partial\Omega} (F \circ n)\, dS = \iint\limits_{\partial\Omega} F_1 n_1\, dS + \iint\limits_{\partial\Omega} F_2 n_2\, dS + \iint\limits_{\partial\Omega} F_3 n_3\, dS.$$

Wystarczy pokazać, że odpowiednie składniki są sobie równe.

Pokażemy, że

$$\iiint\limits_{\Omega} \frac{\partial F_3}{\partial z}\, dx\, dy\, dz = \iint\limits_{\partial\Omega} F_3 n_3\, dS.$$

Ω ma postać

$$\Omega = \{(x, y, z) : \varphi_1(x, y) \leqslant z \leqslant \varphi_2(x, y),\ (x, y) \in D_1\}.$$

$$\begin{aligned}\iiint\limits_{\Omega} \frac{\partial F_3}{\partial z}\, dx\, dy\, dz &= \iint\limits_{D_1} \int\limits_{\varphi_1(x,y)}^{\varphi_2(x,y)} \frac{\partial F_3}{\partial z}\, dz\ \ dx\, dy \\ &= \iint\limits_{D_1} [F_3(x, y, \varphi_2(x, y)) - F_3(x, y, \varphi_1(x, y))]\ dx\, dy \\ &= \iint\limits_{D_1} F_3(x, y, \varphi_2(x, y))\, dx\, dy - \iint\limits_{D_1} F_3(x, y, \varphi_1(x, y))\, dx\, dy.\end{aligned}$$

Brzeg obszaru Ω składa się z powierzchni dolnej S_1^- i górnej S_2^+ związanych z wykresami funkcji $z = \varphi_1(x, y)$ i $z = \varphi_2(x, y)$ dla $(x, y) \in D_1$ oraz z powierzchni pionowej pomiędzy tymi wykresami. Wektory normalne do powierzchni pionowej mają postać $(n_1, n_2, 0)$, tzn. $n_3 = 0$. Zatem

$$\iint\limits_{\partial\Omega} F_3 n_3\, dS = \iint\limits_{S_2^+} F_3 n_3\, dS - \iint\limits_{S_1^+} F_3 n_3\, dS.$$

Dla S_2 wektor normalny ma postać

$$n = \frac{\left(-\frac{\partial\varphi_2}{\partial x}, -\frac{\partial\varphi_2}{\partial y}, 1\right)}{\sqrt{\left(\frac{\partial\varphi_2}{\partial x}\right)^2 + \left(\frac{\partial\varphi_2}{\partial y}\right)^2 + 1}}.$$

Stąd na podstawie (9.1) otrzymujemy

$$\begin{aligned}
&\iint\limits_{S_2^+} F_3 n_3\, dS \\
&= \iint\limits_{D_1} F_3(x,y,\varphi_2(x,y)) \frac{1}{\sqrt{\left(\frac{\partial\varphi_2}{\partial x}\right)^2 + \left(\frac{\partial\varphi_2}{\partial y}\right)^2 + 1}} \sqrt{\left(\frac{\partial\varphi_2}{\partial x}\right)^2 + \left(\frac{\partial\varphi_2}{\partial y}\right)^2 + 1}\, dx\, dy \\
&\qquad = \iint\limits_{D_1} F_3(x,y,\varphi_2(x,y))\, dx\, dy.
\end{aligned}$$

Podobnie uzyskujemy

$$\iint\limits_{S_1^+} F_3 n_3\, dS = \iint\limits_{D_1} F_3(x,y,\varphi_1(x,y))\, dx\, dy.$$

$\square$

Przykłady.

(a) $F = (2x, y^2, z^2)$ oraz $S = \{x^2 + y^2 + z^2 = 1\}$. Niech B oznacza kulę jednostkową. Wtedy

$$\iint\limits_S F \circ dS = \iiint\limits_B (2+2y+2z)\, dx\, dy\, dz = \iiint\limits_B 2\, dx\, dy\, dz = 2\,\mathrm{vol}(B) = \frac{8\pi}{3}.$$

(b) Chcemy obliczyć całkę niezorientowaną $\iint\limits_S (x^2 + y + z)\, dS$, gdzie S jest sferą jednostkową. Mamy $n = (x,y,z)$. Niech $F = (x,1,1)$. Wtedy

$$\iint\limits_S (x^2 + y + z)\, dS = \iint\limits_S (x,1,1) \circ dS = \iiint\limits_B\, dx\, dy\, dz = \frac{4\pi}{3}.$$

12.1 Interpretacja fizyczna dywergencji

Rozważamy pole wektorowe F klasy C^1 w $\mathbb{R}^3$. Dla ustalonego punktu P niech B_r oznacza kulę o promieniu r i środku w punkcie P. Ze wzoru Gaussa-Ostrogradskiego mamy

$$\iiint\limits_{B_r} \mathrm{div}\, F\, dx\, dy\, dz = \int\limits_{\partial B_r} (F \circ n)\, dS = \int\limits_{\partial B_r} F \circ dS.$$

Z twierdzenia o wartości średniej mamy

$$\iiint_{B_r} \operatorname{div} F\, dx\, dy\, dz = \operatorname{div} F(Q_r) \cdot \operatorname{vol}(B_r)$$

dla pewnego punktu Q_r z B_r. Zatem

$$\operatorname{div} F(Q_r) = \frac{1}{\operatorname{vol}(B_r)} \int_{\partial B_r} F \circ dS.$$

Przechodząc do granicy $r \to 0^+$ otrzymujemy

$$\operatorname{div} F(P) = \lim_{r \to 0^+} \frac{1}{\operatorname{vol}(B_r)} \int_{\partial B_r} F \circ dS.$$

Wyrażenie

$$\frac{1}{\operatorname{vol}(B_r)} \int_{\partial B_r} F \circ dS$$

jest prędkością przepływu na zewnątrz sfery ∂B_r na jednostkę objętości. Jeśli $\operatorname{div} F(P) > 0$, to punkt P nazywamy źródłem. Jeśli $\operatorname{div} F(P) < 0$, to punkt P nazywamy odpływem.

12.2 Potencjały i funkcje harmoniczne

Rozważać będziemy funkcje dwu zmiennych. Przypuśćmy, że $F = \nabla V$, gdzie V jest funkcją dwu zmiennych klasy C^1. Niech C będzie krzywą łączącą punkt A z punktem B. Wtedy z Twierdzenia 8.1 otrzymujemy

$$\int_C F \circ ds = V(B) - V(A).$$

Oznaczmy $F = (P, Q)$. Powstaje problem jak stwierdzić, czy $F = \nabla V$ dla pewnego potencjału V. Szukamy funkcji V spełniającej

$$P = \frac{\partial V}{\partial x}, \qquad Q = \frac{\partial V}{\partial y}.$$

Załóżmy, że P i Q są klasy C^1. Tzn. potencjał V musiałby być klasy C^2. Wtedy

$$\frac{\partial P}{\partial y} = \frac{\partial^2 V}{\partial y \partial x} = \frac{\partial^2 V}{\partial x \partial y} = \frac{\partial Q}{\partial x}.$$

Powstaje następny problem, czy warunek

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

jest wystarczający dla istnienia potencjału V. Przypuśćmy, że warunek ten jest spełniony. Wtedy dla krzywej zamkniętej C otaczającej obszar D mamy

$$\int_C P\,dx + Q\,dy = \iint_D \left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) dx\,dy = 0.$$

Załóżmy, że pole wektorowe $F = (P, Q)$ spełnia $\dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x}$ w obszarze spójnym i jednospójnym D. Jednospójność oznacza, że dopełnienie D jest również spójne. Ustalmy punkt (x_0, y_0) w D. Dla punktu (x_1, y_1) z D określamy

$$V(x_1, y_1) = \int_{C_{x_1,y_1}} F \circ ds = \int_{C_{x_1,y_1}} P\,dx + Q\,dy,$$

gdzie C_{x_1,y_1} jest łamaną o skończonej liczbie poziomych i pionowych odcinków, łączącą (x_0, y_0) z (x_1, y_1). Całka nie zależy od wyboru łamanej C_{x_1,y_1}. Rzeczywiście niech $\widetilde{C}_{x_1,y_1}$ będzie inną taką łamaną. Łamane mogą przecinać się w kilku punktach. Rozważmy dwa kolejne punkty przecięcia i obszar D ograniczony przez łamane pomiędzy tym punktami. Niech C i $\widetilde{C}$ oznaczają odcinki łamanych C_{x_1,y_1} i $\widetilde{C}_{x_1,y_1}$ tworzące brzeg obszaru D. Wtedy $\partial D = C \cup \widetilde{C}^-$ lub $\partial D = C^- \cup \widetilde{C}$. Rozważymy pierwszy przypadek. Ze wzoru Greena mamy

$$0 = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy = \int_{\partial D} P\,dx + Q\,dy = \int_C P\,dx + Q\,dy - \int_{\widetilde{C}} P\,dx + Q\,dy.$$

Zatem

$$\int_C P\,dx + Q\,dy = \int_{\widetilde{C}} P\,dx + Q\,dy.$$

Poprzez zsumowanie otrzymamy

$$\int_{C_{x_1,y_1}} P\,dx + Q\,dy = \int_{\widetilde{C}_{x_1,y_1}} P\,dx + Q\,dy.$$

Pokażemy, że $\nabla V = F = (P, Q)$. Rozważamy punkty (x_1, y_1) oraz (x_1+h, y_1) dla małych wartości h. Możemy przyjąć, że $C_{x_1+h,y_1} = C_{x_1,y_1} \cup l_h$, gdzie l_h jest poziomym odcinkiem od (x_1, y_1) do $(x_1 + h, y_1)$. Odcinek ten możemy sparametryzować za pomocą

$$x = x_1 + t, \quad y = y_1,\ 0 \leqslant t \leqslant h.$$

Wtedy

$$\begin{aligned} V(x_1 + h, y_1) - V(x_1, y_1) &= \int\limits_{C_{x_1+h,y_1}} P\,dx + Q\,dy - \int\limits_{C_{x_1,y_1}} P\,dx + Q\,dy \\ &= \int\limits_{l_h} P\,dx + Q\,dy = \int\limits_0^h P(x_1 + t, y_1)\,dt. \end{aligned}$$

Zatem

$$\frac{V(x_1 + h, y_1) - V(x_1, y_1)}{h} = \frac{1}{h}\int\limits_0^h P(x_1 + t, y_1)\,dt \underset{h\to 0}{\longrightarrow} P(x_1, y_1).$$

To oznacza, że $\dfrac{\partial V}{\partial x} = P$. Podobnie pokazujemy, że $\dfrac{\partial V}{\partial y} = Q$. Mówimy, że

Definicja 12.3. *funkcja $u(x, y)$ klasy C^2 jest harmoniczna jeśli*

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Przypuśćmy, że funkcja harmoniczna $u(x, y)$ jest określona w obszarze spójnym i jednospójnym D. Oznaczmy

$$P = -\frac{\partial u}{\partial y}, \qquad Q = \frac{\partial u}{\partial x}.$$

Wtedy

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Z poprzedniego rozumowania istnieje potencjał $v(x, y)$ taki, że

$$\nabla v = (P, Q) = \left(-\frac{\partial u}{\partial y}, \frac{\partial u}{\partial x}\right),$$

tzn.

$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}, \qquad \frac{\partial v}{\partial y} = \frac{\partial u}{\partial x}. \tag{12.1}$$

Otrzymujemy

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = -\frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y \partial x} = 0,$$

czyli $v(x, y)$ jest też funkcją harmoniczną. Funkcję $v(x, y)$ określoną przez (12.1) nazywamy funkcją harmoniczną sprzężoną do funkcji $u(x, y)$.

Przykład. $u = x^2 - y^2$ i $2xy$ są sprzężonymi funkcjami harmonicznymi. Zauważmy, że dla $z = x + iy$ mamy

$$z^2 = (x + iy)^2 = x^2 - y^2 + i\,2xy.$$

12.3 Inny zapis całki $\iint\limits_S F \circ dS$

Niech $F = (P, Q, R)$ będzie polem wektorowym. Wtedy

$$\iint\limits_S F \circ dS = \iint\limits_S (F \circ n)\,dS = \iint\limits_S P\,n_1\,dS + \iint\limits_S Q\,n_2\,dS + \iint\limits_S R\,n_3\,dS.$$

Rozważmy mały fragment dS powierzchni S i rzut $dS_{x,y}$ tego fragmentu na płaszczyznę xy. Wtedy stosunek pól tych fragmentów zależy od trzeciej współrzędnej wektora normalnego do dS

$$n_3 \text{Pole}\,(dS) = \text{Pole}(dS_{x,y}).$$

Stąd

$$n_3\,dS = dx\,dy.$$

Podobnie

$$n_1\,dS = dy\,dz, \qquad n_2 dS = dz\,dx.$$

Stosuje się więc alternatywny zapis

$$\iint\limits_S F \circ dS = \iint\limits_S P\,dy\,dz + \iint\limits_S Q\,dz\,dx + \iint\limits_S R\,dx\,dy.$$