

Uniwersytet Jana Kochanowskiego w Kielcach
Wydział Matematyczno-Przyrodniczy
Instytut Matematyki

Dr hab. prof. UJK GRZEGORZ ŁYSIK

Analiza Matematyczna II

Skrypt wykładów

Kielce, 2012.

1 Funkcje wielu zmiennych

1.1 Przestrzeń \mathbb{R}^n i jej podzbiory

1. Niech n będzie liczbą naturalną. Przestrzeń \mathbb{R}^n jest iloczynem kartezjańskim n egzemplarzy prostej rzeczywistej \mathbb{R} tzn. $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$. Punkt x należący do \mathbb{R}^n posiada n współrzędnych, czyli $x = (x_1, \dots, x_n)$. Punkty przestrzeni \mathbb{R}^n nazywamy też wektorami. Przestrzeń \mathbb{R}^n jest przestrzenią liniową z naturalnymi działaniami dodawania wektorów i mnożenia przez liczbę rzeczywistą

$$\begin{aligned}x + y &= (x_1 + y_1, \dots, x_n + y_n) \quad \text{dla } x, y \in \mathbb{R}^n; \\ \lambda x &= (\lambda x_1, \dots, \lambda x_n) \quad \text{dla } \lambda \in \mathbb{R}, x \in \mathbb{R}^n.\end{aligned}$$

W przestrzeni \mathbb{R}^n wprowadza się normę euklidesową l_2 wektora $x \in \mathbb{R}^n$ wzorem

$$\|x\| = \|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Norma euklidesowa spełnia następujące warunki

- 1. $\|x\| = 0$ wtedy i tylko wtedy, gdy $x = 0$;
- 2. $\|\lambda x\| = |\lambda| \|x\|$;
- 3. $\|x + y\| \leq \|x\| + \|y\|$.

Wykażemy, że norma euklidesowa spełnia warunek 3. W tym celu należy wykazać, że dla dowolnych $x_i, y_i \in \mathbb{R}, i = 1, \dots, n$ zachodzi nierówność

$$\sqrt{\sum_{i=1}^n (x_i + y_i)^2} \leq \sqrt{\sum_{i=1}^n x_i^2} + \sqrt{\sum_{i=1}^n y_i^2}. \quad (1)$$

Ponieważ wyrażenia podpierwiastkowe są nieujemne nierówność powyższa jest równoważna z nierównością

$$\sum_{i=1}^n (x_i + y_i)^2 \leq \left(\sqrt{\sum_{i=1}^n x_i^2} + \sqrt{\sum_{i=1}^n y_i^2} \right)^2 = \sum_{i=1}^n x_i^2 + 2 \sqrt{\sum_{i=1}^n x_i^2} \cdot \sqrt{\sum_{i=1}^n y_i^2} + \sum_{i=1}^n y_i^2. \quad (2)$$

Korzystając z tożsamości $\sum_{i=1}^n (x_i + y_i)^2 = \sum_{i=1}^n x_i^2 + 2 \sum_{i=1}^n x_i y_i + \sum_{i=1}^n y_i^2$, wystarczy zatem wykazać nierówność Cauchy'ego

$$\sum_{i=1}^n x_i y_i \leq \sqrt{\sum_{i=1}^n x_i^2} \cdot \sqrt{\sum_{i=1}^n y_i^2}. \quad (3)$$

W tym celu zdefiniujemy funkcję zmiennej rzeczywistej t

$$f(t) = \sum_{i=1}^n (x_i t + y_i)^2 = \left(\sum_{i=1}^n x_i^2 \right) t^2 + 2 \left(\sum_{i=1}^n x_i y_i \right) t + \sum_{i=1}^n y_i^2. \quad (4)$$

Ponieważ funkcja f jest nieujemną funkcją kwadratową jej wyróżnik musi być niedodatni, czyli

$$\left(\sum_{i=1}^n x_i y_i\right)^2 - \left(\sum_{i=1}^n x_i^2\right) \cdot \left(\sum_{i=1}^n y_i^2\right) \leq 0. \quad (5)$$

Stąd wynika nierówność (3).

W przestrzeni \mathbb{R}^n wprowadzamy metrykę euklidesową wzorem

$$\rho(x, y) = \|x - y\|_2 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad \text{dla } x, y \in \mathbb{R}^n$$

Metryka ρ jest nieujemną funkcją na iloczynie kartezjańskim $\mathbb{R}^n \times \mathbb{R}^n$ spełniającą warunki

- 1. $\rho(x, y) = 0$ wtedy i tylko wtedy, gdy $x = y$;
- 2. $\rho(x, y) = \rho(y, x)$;
- 3. $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$.

Ostatni warunek jest nazywany *nierównością trójkąta*. Wynika on z warunku 3. normy.

Użyteczne jest także wprowadzenie normy l_1 i normy l_∞ wektora $x \in \mathbb{R}^n$,

$$\begin{aligned} \|x\|_1 &= |x_1| + \dots + |x_n|, \\ \|x\|_\infty &= \max\{|x_1|, \dots, |x_n|\}. \end{aligned}$$

Pomiędzy normami $\|x\|_1$, $\|x\|_2$ i $\|x\|_\infty$ zachodzą nierówności

$$\begin{aligned} \|x\|_\infty &\leq \|x\|_2 \leq \|x\|_1; \\ \|x\|_1 &\leq \sqrt{n} \|x\|_2 \leq n \|x\|_\infty. \end{aligned}$$

Def. Otwartą kulą euklidesową o środku $\hat{x} \in \mathbb{R}^n$ i promieniu $r > 0$ nazywamy zbiór

$$B(\hat{x}, r) = \{x \in \mathbb{R}^n : \rho(x, \hat{x}) < r\}.$$

Def. Prostopadłościanem lub kostką nazywamy zbiór

$$P(a, b) = \{x \in \mathbb{R}^n : a_i < x_i < b_i, i = 1, \dots, n\}, \quad \text{gdzie } a < b.$$

Def. Sympleksem nazywamy zbiór

$$S(r) = \{x \in \mathbb{R}^n : x_1 > 0, \dots, x_n > 0, x_1 + \dots + x_n < r\}, \quad \text{gdzie } r > 0.$$

Def. Zbiór $\Omega \subset \mathbb{R}^n$ nazywamy otwartym jeśli dla dowolnego $x \in \Omega$ istnieje $r > 0$ takie, że $B(x, r) \subset \Omega$.

Zbiór $F \subset \mathbb{R}^n$ nazywamy domkniętym jeśli $\mathbb{R}^n \setminus F$ jest zbiorem otwartym.

Domknięciem zbioru $\Omega \subset \mathbb{R}^n$ nazywamy zbiór $\bar{\Omega} = \{x \in \mathbb{R}^n : \text{dla dowolnego } \varepsilon > 0 \text{ istnieje } y \in \Omega \text{ taki, że } \rho(x, y) < \varepsilon\}$.

Domknięcie $\bar{\Omega}$ zbioru Ω jest zbiorem domkniętym.

Wnętrzem zbioru $\Omega \subset \mathbb{R}^n$ nazywamy zbiór $int\Omega = \{x \in \Omega : \text{istnieje } r > 0 \text{ takie, że } B(x, r) \subset \Omega\}$.

Brzegiem zbioru $\Omega \subset \mathbb{R}^n$ nazywamy zbiór $\partial\Omega = \bar{\Omega} \setminus int\Omega$.

Zbiór $\Omega \subset \mathbb{R}^n$ nazywamy spójnym, jeśli dowolne dwa jego punkty można połączyć krzywą zawartą w Ω .

Zbiór otwarty i spójny nazywany obszarem.

Mówimy, że zbiór Ω jest ograniczony jeśli jest on zawarty w pewnej kuli.

1.2 Granica i ciągłość funkcji

Definicja 1.1 Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ oraz $\hat{x} \in \bar{\Omega} \setminus \Omega$. Mówimy, że granica funkcji f przy $x \rightarrow \hat{x}$ jest równa liczbie g jeśli dla każdego $\varepsilon > 0$ istnieje $\delta > 0$ taka, że jeśli $\|x - \hat{x}\| \leq \delta$, $x \in \Omega$, to $|f(x) - g| < \varepsilon$. Piszemy wówczas

$$\lim_{x \rightarrow \hat{x}} f(x) = g.$$

W przypadku funkcji wielu zmiennych definiuje się pojęcie granic iterowanych. Pojęcie to wprowadzimy dla funkcji dwóch zmiennych.

Definicja 1.2 Niech $(\hat{x}, \hat{y}) \in \mathbb{R}^2$ oraz f będzie zdefiniowana w zbiorze $\Omega = \{(x, y) \in \mathbb{R}^2 : |x - \hat{x}| < d_1, |y - \hat{y}| < d_2\} \setminus \{(\hat{x}, \hat{y})\}$ dla pewnych $d_1, d_2 > 0$. Załóżmy, że dla każdego $y \in \mathbb{R}$ takiego, że $0 < |y - \hat{y}| < d_2$ istnieje granica

$$\lim_{x \rightarrow \hat{x}} f(x, y) = g(y)$$

oraz, że istnieje granica

$$\lim_{y \rightarrow \hat{y}} g(y) = g.$$

Wówczas mówimy, że istnieje granica iterowana

$$\lim_{y \rightarrow \hat{y}} \lim_{x \rightarrow \hat{x}} f(x, y) = g.$$

Analogicznie definiujemy

$$\lim_{x \rightarrow \hat{x}} \lim_{y \rightarrow \hat{y}} f(x, y).$$

Okazuje się, że granice iterowane nie muszą być sobie równe nawet w przypadku, gdy obie istnieją.

Przykład 1.1 Niech

$$f(x, y) = \frac{|y|}{|x| + |y|} \quad \text{dla } (x, y) \neq (0, 0).$$

Wówczas

$$\lim_{y \rightarrow 0} \lim_{x \rightarrow 0} f(x, y) = 1 \neq 0 = \lim_{x \rightarrow 0} \lim_{y \rightarrow 0} f(x, y).$$

Twierdzenie 1.1 Niech $(\hat{x}, \hat{y}) \in \mathbb{R}^2$, $\Omega = \{(x, y) \in \mathbb{R}^2 : |x - \hat{x}| < d_1, |y - \hat{y}| < d_2\} \setminus \{(\hat{x}, \hat{y})\}$ dla pewnych $d_1, d_2 > 0$ oraz $f : \Omega \rightarrow \mathbb{R}$. Załóżmy, że istnieje granica

$$\lim_{(x, y) \rightarrow (\hat{x}, \hat{y})} f(x, y) = g.$$

Jeśli dla każdego ustalonego x takiego, że $0 < |x - \hat{x}| < d_1$ istnieje granica

$$\lim_{y \rightarrow \hat{y}} f(x, y) = g(x)$$

oraz dla każdego ustalonego y takiego, że $0 < |y - \hat{y}| < d_2$ istnieje granica

$$\lim_{x \rightarrow \hat{x}} f(x, y) = h(y),$$

to istnieją granice iterowane i są równe g .

Definicja 1.3 Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $\hat{x} \in \Omega$. Mówimy, że f jest ciągła w \hat{x} jeśli granica funkcji $f|_{\Omega \setminus \{\hat{x}\}}$ w punkcie \hat{x} jest równa $f(\hat{x})$.

Twierdzenie 1.2 Jeśli funkcje f i g są ciągłe w punkcie $\hat{x} \in \Omega$, to również funkcje $f \pm g$, $f \cdot g$ oraz f/g są ciągłe w \hat{x} i ile $g(\hat{x}) \neq 0$.

Twierdzenie 1.3 Jeśli f jest ciągła w \hat{x} oraz $f(\hat{x}) > 0$, to istnieje $\delta > 0$, taka, że $f(x) > 0$ dla $\|x - \hat{x}\| < \delta$.

Twierdzenie 1.4 (Weierstrassa.) Jeśli funkcja f jest ciągła na zbiorze zwartym $K \subset \mathbb{R}^n$, to f jest ograniczona na K i osiąga swoje kresy.

Definicja 1.4 Mówimy, że funkcja f jest jednostajnie ciągła na zbiorze D jeśli dla każdego $\varepsilon > 0$ istnieje $\delta > 0$ taka, że dla dowolnych $x, y \in D$ zachodzi implikacja: jeśli $\|x - y\| < \delta$, to $|f(x) - f(y)| < \varepsilon$.

Twierdzenie 1.5 *Jeśli f jest ciągła na zbiorze zwartym K , to jest jednostajnie ciągła na K .*

Definicja 1.5 *Niech funkcje $\varphi_1(t_1, \dots, t_k), \dots, \varphi_n(t_1, \dots, t_k)$ będą zdefiniowane na zbiorze $U \subset \mathbb{R}^k$. Niech $\Omega = \{x \in \mathbb{R}^n : \text{istnieje } t = (t_1, \dots, t_k) \in U \text{ takie, że } x_i = \varphi_i(t) \text{ dla } i = 1, \dots, n\}$ oraz niech $f : \Omega \rightarrow \mathbb{R}$ będzie funkcją na Ω . Wówczas definiujemy funkcję $u = f \circ \varphi : U \rightarrow \mathbb{R}$ będącą złożeniem funkcji f i $\varphi = (\varphi_1, \dots, \varphi_n)$ wzorem*

$$u(t) = f(\varphi_1(t), \dots, \varphi_n(t)).$$

Twierdzenie 1.6 *Jeśli funkcje $x_i = \varphi_i(t)$, $i = 1, \dots, n$, są ciągłe w punkcie \dot{t} oraz f jest ciągła w punkcie $\dot{x} = \varphi(\dot{t})$, to $f \circ \varphi$ jest ciągła w \dot{t} .*

Definicja 1.6 *Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$. Mówimy, że f spełnia warunek Lipschitza w punkcie $\dot{x} \in \Omega$ ze stałą Lipschitza $L < \infty$ jeśli istnieje $\delta > 0$ taka, że dla $x \in B(\dot{x}, \delta) \cap \Omega$ zachodzi*

$$\|f(x) - f(\dot{x})\| \leq L\|x - \dot{x}\|.$$

Twierdzenie 1.7 *Jeśli f spełnia warunek Lipschitza w punkcie $\dot{x} \in \Omega$, to f jest ciągła w \dot{x} . Ponadto, jeśli stała Lipschitza L nie zależy od $\dot{x} \in U \subset \Omega$, to f jest jednostajnie ciągła w U .*

Definicja 1.7 *Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$ oraz $\alpha \in (0, 1]$. Mówimy, że f jest hölderowsko ciągła w punkcie $\dot{x} \in \Omega$ z wykładnikiem α jeśli istnieją $L < \infty$, $\delta > 0$ takie, że dla $x \in B(\dot{x}, \delta) \cap \Omega$ zachodzi*

$$\|f(x) - f(\dot{x})\| \leq L\|x - \dot{x}\|^\alpha.$$

Twierdzenie 1.8 *Jeśli f jest hölderowsko ciągła w punkcie $\dot{x} \in \Omega$, to f jest ciągła w \dot{x} .*

Definicja 1.8 *Odwzorowanie $T : \mathbb{R}^n \rightarrow \mathbb{R}^k$ nazywamy liniowym jeśli dla dowolnych $x, y \in \mathbb{R}^n$ oraz $\alpha, \beta \in \mathbb{R}$ zachodzi*

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y).$$

Twierdzenie 1.9 *Odwzorowanie liniowe $T : \mathbb{R}^n \rightarrow \mathbb{R}^k$ jest jednostajnie ciągłe na \mathbb{R}^n .*

2 Rachunek różniczkowy rzędu pierwszego

2.1 Pochodne kierunkowe funkcji wielu zmiennych

Przypomnijmy definicję pochodnej funkcji jednej zmiennej.

Definicja 2.1 Niech f będzie funkcją określoną na przedziale otwartym $I = (a, b) \subset \mathbb{R}$ o wartościach rzeczywistych. Mówimy, że f ma pochodną w punkcie $\dot{x} \in I$ jeśli istnieje granica ilorazu różnicowego

$$\lim_{h \rightarrow 0} \frac{f(\dot{x} + h) - f(\dot{x})}{h} = f'(\dot{x}), \quad (6)$$

Funkcję $f : I \rightarrow \mathbb{R}$ nazywamy różniczkowalną w I jeśli f ma pochodną w każdym punkcie $\dot{x} \in I$.

W przypadku funkcji wielu zmiennych wyrażenie (6) nie ma sensu. Istotnie aby $f(\dot{x} + h)$ miało sens h powinno być wektorem, lecz nie jest zdefiniowana operacja dzielenia przez wektor. Można jednak zdefiniować pochodną funkcji w kierunku ustalonego wektora.

Definicja 2.2 Niech Ω będzie otwartym podzbiorem \mathbb{R}^n , $f : \Omega \rightarrow \mathbb{R}$, $\dot{x} \in \Omega$ oraz $v \in \mathbb{R}^n \setminus \{0\}$. Pochodną kierunkową funkcji f w kierunku wektora v w punkcie \dot{x} nazywamy granicę

$$\lim_{t \rightarrow 0} \frac{f(\dot{x} + tv) - f(\dot{x})}{t} =: L_v f(\dot{x})$$

o ile ta granica istnieje i jest skończona.

Zauważmy, że $L_v f(\dot{x})$ jest faktycznie pochodną funkcji jednej zmiennej $g(t) = f(\dot{x} + tv)$ w zerze.

Pochodną kierunkową w kierunku wektora e_i , $i = 1, \dots, n$, nazywamy pochodną cząstkową w kierunku e_i i oznaczamy

$$L_{e_i} f(\dot{x}) = \frac{\partial f}{\partial x_i}(\dot{x}) = f'_{x_i}(\dot{x}) = \lim_{t \rightarrow 0} \frac{f(\dot{x} + te_i) - f(\dot{x})}{t}.$$

Dla pochodnych cząstkowych stosuje się zwykle reguły różniczkowania. Jednak w przypadku funkcji wielu zmiennych z istnienia pochodnych kierunkowych nawet w całym zbiorze określoności funkcji i w każdym kierunku nie można wnioskować o ciągłości funkcji.

Przykład 2.1 Niech

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^4 + y^2} & \text{dla } (x, y) \neq (0, 0), \\ 0 & \text{dla } (x, y) = (0, 0). \end{cases}$$

Wówczas f posiada pochodne kierunkowe w kierunku dowolnego wektora lecz jest nieciągła w zerze.

Przykład 2.2 Niech

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{dla } (x, y) \neq (0, 0), \\ 0 & \text{dla } (x, y) = (0, 0). \end{cases}$$

Wówczas $f'_x(0, 0) = f'_y(0, 0) = 0$, lecz nie istnieją pochodne $L_v f(0, 0)$ dla $v \in (\mathbb{R} \setminus \{0\})^2$.

Przykład 2.3 Niech

$$f(x, y) = \sqrt{|xy|}$$

Wówczas f jest ciągła i ma pochodne kierunkowe w dowolnym punkcie w kierunku dowolnego wektora. Jeśli $v = (\cos \varphi, \sin \varphi)$, to

$$L_v f(0, 0) = \lim_{t \rightarrow 0} \frac{\sqrt{t^2 |\cos \varphi \sin \varphi|} - 0}{t} = \pm \sqrt{|\cos \varphi \sin \varphi|}.$$

Zatem $L_v f$ nie jest na ogół równa $v_1 \frac{\partial f}{\partial x} + v_2 \frac{\partial f}{\partial y}$.

Twierdzenie 2.1 Niech $v, w \in \mathbb{R}^n \setminus \{0\}$, $f : \Omega \rightarrow \mathbb{R}$ oraz $\hat{x} \in \Omega$. Załóżmy, że pochodna kierunkowa $L_v f$ istnieje w pewnym otoczeniu punktu \hat{x} i jest ciągła w \hat{x} . Wówczas jeśli istnieje pochodna kierunkowa $L_w f(\hat{x})$, to istnieje pochodna $L_{v+w} f(\hat{x})$ oraz

$$L_{v+w} f(\hat{x}) = L_v f(\hat{x}) + L_w f(\hat{x}).$$

Twierdzenie 2.2 Niech $f : \Omega \rightarrow \mathbb{R}$ oraz $\hat{x} \in \Omega$. Jeśli istnieją pochodne cząstkowe $\frac{\partial f}{\partial x_i}$ dla $i = 1, \dots, n$ w pewnym otoczeniu punktu \hat{x} oraz są ciągłe w \hat{x} , to dla $h \in \mathbb{R}^n$ takiego, że $\hat{x} + h \in \Omega$ zachodzi

$$f(\hat{x} + h) - f(\hat{x}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\hat{x}) \cdot h_i + \sum_{i=1}^n \alpha_i(h) \cdot h_i$$

przy czym $\lim_{h \rightarrow 0} \alpha_i(h) = 0$ dla $i = 1, \dots, n$.

Wniosek 2.1 Niech $f : \Omega \rightarrow \mathbb{R}$ oraz $\hat{x} \in \Omega$. Jeśli istnieją pochodne cząstkowe $\frac{\partial f}{\partial x_i}$ dla $i = 1, \dots, n$ w pewnym otoczeniu punktu \hat{x} oraz są ciągłe w \hat{x} , to f jest ciągła w \hat{x} .

Definicja 2.3 Niech Ω będzie otwartym podzbiorem \mathbb{R}^n , $f : \Omega \rightarrow \mathbb{R}$, $\hat{x} \in \Omega$. Jeśli istnieją pochodne cząstkowe $\frac{\partial f}{\partial x_i}(\hat{x})$ dla $i = 1, \dots, n$, to wektor

$$\left[\frac{\partial f}{\partial x_1}(\hat{x}), \dots, \frac{\partial f}{\partial x_n}(\hat{x}) \right]$$

nazywamy gradientem funkcji f w punkcie \hat{x} i oznaczamy $\text{grad} f(\hat{x})$ lub $\nabla f(\hat{x})$.

Tezę Twierdzenia 2.2 można sformułować następująco.

Wniosek 2.2 Jeśli istnieją pochodne cząstkowe $\frac{\partial f}{\partial x_i}$ dla $i = 1, \dots, n$ w pewnym otoczeniu punktu \hat{x} oraz są ciągłe w \hat{x} , to dla $h \in \mathbb{R}^n$ takiego, że $\hat{x} + h \in \Omega$ zachodzi

$$f(\hat{x} + h) - f(\hat{x}) = \text{grad}f(\hat{x}) \circ h + \alpha(h) \circ h,$$

gdzie \circ oznacza iloczyn skalarny, $\alpha(h) = (\alpha_1(h), \dots, \alpha_n(h))$ przy czym $\lim_{h \rightarrow 0} \alpha_i(h) = 0$ dla $i = 1, \dots, n$.

Wniosek 2.3 Niech $f : \Omega \rightarrow \mathbb{R}^k$, $\hat{x} \in \Omega$, Ω jest obszarem w \mathbb{R}^n . Jeśli istnieją pochodne cząstkowe $\frac{\partial f_j}{\partial x_i}$ dla $i = 1, \dots, n, j = 1, \dots, k$ w pewnym otoczeniu punktu \hat{x} oraz są ciągłe w \hat{x} , to dla $h \in \mathbb{R}^n$ takiego, że $\hat{x} + h \in \Omega$ zachodzi

$$f(\hat{x} + h) - f(\hat{x}) = \begin{pmatrix} \text{grad}f_1(\hat{x}) \\ \dots \\ \text{grad}f_k(\hat{x}) \end{pmatrix} h + \alpha(h)h,$$

gdzie $\alpha(h) = (\alpha^1(h), \dots, \alpha^k(h))^{\text{tr}}$ przy czym $\lim_{h \rightarrow 0} \alpha^i(h) = 0$ dla $i = 1, \dots, k$.

2.2 Różniczka odwzorowania

Definicja 2.4 Niech Ω będzie otwartym podzbiorem \mathbb{R}^n , $f : \Omega \rightarrow \mathbb{R}^k$, $\hat{x} \in \Omega$. Mówimy, że f jest odwzorowaniem różniczkowalnym w punkcie \hat{x} jeśli istnieje operator liniowy

$$L : \mathbb{R}^n \rightarrow \mathbb{R}^k$$

zwany różniczką funkcji f w punkcie \hat{x} i oznaczany $Df(\hat{x})$ (lub $df(\hat{x})$ jeśli $k = 1$) taki, że dla $h \in \mathbb{R}^n$ spełniającego $\hat{x} + h \in \Omega$ zachodzi

$$f(\hat{x} + h) = f(\hat{x}) + L(h) + \alpha(h)h \quad \text{przy czym} \quad \lim_{h \rightarrow 0} \alpha(h) = 0. \quad (7)$$

Uwaga. Warunek (7) jest równoważny warunkowi

$$\lim_{h \rightarrow 0} \frac{f(\hat{x} + h) - f(\hat{x}) - L(h)}{\|h\|} = 0. \quad (8)$$

Twierdzenie 2.3 Niech $f : \Omega \rightarrow \mathbb{R}$. Jeśli f jest różniczkowalna w \hat{x} , to dla każdego wektora $v \neq 0$ istnieje pochodna kierunkowa $L_v f(\hat{x})$ oraz

$$L_v f(\hat{x}) = df(\hat{x})(v).$$

W szczególności istnieją pochodne cząstkowe $\frac{\partial f}{\partial x_i}(\hat{x}) = df(\hat{x})(e_i)$, $i = 1, \dots, n$ oraz

$$df(\hat{x})(h) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\hat{x}) h_i = \text{grad}f(\hat{x}) \circ h \quad \text{dla} \quad h \in \mathbb{R}^n.$$

W przypadku odwzorowania różniczkowalnego o wartościach w \mathbb{R}^k

$$Df(\hat{x})(h) = \begin{pmatrix} \text{grad}f_1(\hat{x}) \\ \dots \\ \text{grad}f_k(\hat{x}) \end{pmatrix} h \quad \text{dla} \quad h \in \mathbb{R}^n.$$

Twierdzenie 2.4 Niech $f : \Omega \rightarrow \mathbb{R}$. Jeśli f jest różniczkowalna w \dot{x} , to f jest ciągła w \dot{x} .

Uwaga. Niech $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ oraz $v \in S^1$. Wówczas wektor v ma współrzędne $v = (\cos \alpha, \sin \alpha) = (\cos \alpha, \cos \beta)$, gdzie α (odpowiednio β) jest kątem pomiędzy v a osią OX (odpowiednio OY). Wówczas

$$L_v f(\dot{x}) = \frac{\partial f}{\partial x}(\dot{x}) \cos \alpha + \frac{\partial f}{\partial y}(\dot{x}) \cos \beta.$$

Analogicznie jeśli $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ oraz $v \in S^{n-1}$, to

$$L_v f(\dot{x}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\dot{x}) \cos \alpha_i,$$

gdzie α_i jest kątem pomiędzy v a osią OX_i , $i = 1, \dots, n$.

Przykład 2.4 Niech $F = (f_1, f_2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ będzie dane wzorem $F(x, y) = (x \cos y, x \sin y)$ Wówczas

$$\begin{aligned} \frac{\partial f_1}{\partial x} &= \cos y, & \frac{\partial f_1}{\partial y} &= -x \sin y, \\ \frac{\partial f_2}{\partial x} &= \sin y, & \frac{\partial f_2}{\partial y} &= x \cos y. \end{aligned}$$

Zatem dla $(h_1, h_2) \in \mathbb{R}^2$ mamy

$$DF(x, y) \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} \cos y & -x \sin y \\ \sin y & x \cos y \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} \cos y \cdot h_1 - x \sin y \cdot h_2 \\ \sin y \cdot h_1 + x \cos y \cdot h_2 \end{bmatrix}.$$

2.3 Reguły różniczkowania.

I Liniowość.

Różniczkowanie jest operacją liniową, tzn. Jeśli $F, G : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$ są różniczkowalne w \dot{x} oraz $\alpha, \beta \in \mathbb{R}$, to $\alpha F + \beta G$ jest różniczkowalne w \dot{x} oraz

$$D(\alpha F + \beta G)(\dot{x}) = \alpha DF(\dot{x}) + \beta DG(\dot{x}).$$

II Różniczka iloczynu.

Jeśli $\varphi : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ oraz $F : \Omega \rightarrow \mathbb{R}^k$ są różniczkowalne w \dot{x} , to $\varphi \cdot F : \Omega \rightarrow \mathbb{R}^k$ jest różniczkowalne w \dot{x} oraz

$$\begin{aligned} D(\varphi \cdot F)(\dot{x}) &= F(\dot{x}) \cdot d\varphi(\dot{x}) + \varphi(\dot{x}) \cdot Df(\dot{x}) \\ &= \begin{pmatrix} f_1(\dot{x}) \\ \dots \\ f_k(\dot{x}) \end{pmatrix} \cdot \text{grad}\varphi(\dot{x}) + \varphi(\dot{x}) \begin{pmatrix} \text{grad}f_1(\dot{x}) \\ \dots \\ \text{grad}f_k(\dot{x}) \end{pmatrix}. \end{aligned}$$

III Różniczka złożenia odwzorowań.

Niech $G : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ oraz $F : V \subset \mathbb{R}^m \rightarrow \mathbb{R}^k$. Załóżmy, że $\text{Im}G = \{y \in \mathbb{R}^m : \text{istnieje } x \in U : y = G(x)\} \subset V$. Wówczas możemy zdefiniować złożenie odwzorowań F i G wzorem

$$F \circ G(x) = F(G(x)) \quad \text{dla } x \in U.$$

Twierdzenie 2.5 *Przy powyższych oznaczeniach jeśli G jest różniczkowalne w punkcie $\hat{x} = (\hat{x}_1, \dots, \hat{x}_n) \in U$ oraz F jest różniczkowalne w punkcie $\hat{y} = (\hat{y}_1, \dots, \hat{y}_m) = G(\hat{x})$ (tzn. $\hat{y}_i = g_i(\hat{x})$ dla $i = 1, \dots, m$), to złożenie $F \circ G$ jest różniczkowalne w \hat{x} oraz*

$$D(F \circ G)(\hat{x}) = DF(G(\hat{x})) \circ DG(\hat{x}).$$

Zatem

$$\frac{\partial(F \circ G)_j}{\partial x_i}(\hat{x}) = \sum_{l=1}^m \frac{\partial F_j}{\partial l_i}(G_j(\hat{x})) \cdot \frac{\partial G_l}{\partial x_i}(\hat{x}), \quad i = 1, \dots, n, j = 1, \dots, k.$$

Wniosek 2.4 *Niezmienniczość pierwszej różniczki. Jeśli $G : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ jest różniczkowalne w $x \in U$ oraz $f : V \subset \mathbb{R}^n \rightarrow \mathbb{R}$ jest różniczkowalna w $y = G(x) \in \mathbb{R}^n$, to*

$$d(f \circ G)(x) = \sum_{i=1}^n f'_{y_i}(G_i(x)) \cdot dG_i(x).$$

Wniosek 2.5 *Różniczkowanie funkcji skalarnych złożonych. Niech $G = (g_1, \dots, g_n) : (a, b) \subset \mathbb{R} \rightarrow \mathbb{R}^n$ będzie różniczkowalne w $t \in (a, b)$ oraz $f : V \subset \mathbb{R}^n \rightarrow \mathbb{R}$ będzie różniczkowalna w $x = G(t) \in \mathbb{R}^n$. Wówczas $f \circ G : (a, b) \rightarrow \mathbb{R}$ jest różniczkowalna w t oraz*

$$(f \circ G)'(t) = \sum_{i=1}^n f'_{x_i}(G(t)) \cdot g'_i(t).$$

Definicja 2.5 *Niech $\Omega = \mathbb{R}^n \setminus \{0\}$ oraz $\lambda \in \mathbb{R}$. Funkcję $f : \Omega \rightarrow \mathbb{R}$ nazywamy (dotąd) jednorodną stopnia λ jeśli dla każdych $t \in \mathbb{R}_+$ i $x \in \Omega$ zachodzi*

$$f(tx_1, \dots, tx_n) = t^\lambda f(x_1, \dots, x_n).$$

Twierdzenie 2.6 *Jeśli f jest funkcją jednorodną stopnia λ i różniczkowalną w $\mathbb{R}^n \setminus \{0\}$, to*

$$x \circ \text{grad}f(x) = \lambda f(x) \quad \text{dla } x \in \mathbb{R}^n \setminus \{0\}.$$

Uwaga. Zachodzi też twierdzenie odwrotne, tzn. jeśli funkcja różniczkowalna $f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ spełnia $x \circ \text{grad}f(x) = \lambda f(x)$ dla $x \in \mathbb{R}^n \setminus \{0\}$, to f jest jednorodna stopnia λ .

2.4 Twierdzenie Lagrange'a o wartości średniej

Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ będzie funkcja różniczkowalna w punkcie $\dot{x} \in \Omega$ oraz $v \in \mathbb{R}^n \setminus \{0\}$. Wówczas

$$L_v f(\dot{x}) = df(\dot{x}) \cdot v = \text{grad}f(\dot{x}) \circ v.$$

Jeśli ograniczymy się do wektorów ze sfery jednostkowej $v \in S^{n-1}$, to można zapytać dla jakiego v , $L_v f(\dot{x})$ przyjmuje największą wartość. Oczywiście jeśli $\text{grad}f(\dot{x}) = 0$, to $L_v f(\dot{x}) = 0$ dla dowolnego v . Zatem możemy założyć, że $\text{grad}f(\dot{x}) \neq 0$. Wówczas na mocy nierówności Schwarz'a mamy

$$|L_v f(\dot{x})| = |\text{grad}f(\dot{x}) \circ v| \leq \|\text{grad}f(\dot{x})\| \cdot \|v\| \leq \|\text{grad}f(\dot{x})\|.$$

Ponadto, jeśli

$$v = \frac{\text{grad}f(\dot{x})}{\|\text{grad}f(\dot{x})\|},$$

to

$$L_v f(\dot{x}) = \text{grad}f(\dot{x}) \circ \frac{\text{grad}f(\dot{x})}{\|\text{grad}f(\dot{x})\|} = \frac{\|\text{grad}f(\dot{x})\|^2}{\|\text{grad}f(\dot{x})\|} = \|\text{grad}f(\dot{x})\|.$$

Zatem $L_v f(\dot{x})$ przyjmuje największą wartość dla $v = \frac{\text{grad}f(\dot{x})}{\|\text{grad}f(\dot{x})\|}$. Innymi słowami funkcja f najszybciej wzrasta w kierunku gradientu.

Twierdzenie 2.7 Lagrange'a o wartości średniej. *Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Jeśli f jest różniczkowalna w każdym punkcie odcinka $[a, b] \subset \Omega$, to istnieje punkt $c \in [a, b]$ taki, że*

$$f(b) - f(a) = df(c) \cdot (b - a).$$

Wniosek 2.6 *Jeśli $f : \Omega \rightarrow \mathbb{R}$ jest różniczkowalna w obszarze $\Omega \subset \mathbb{R}^n$ oraz $df(x) = 0$ dla $x \in \Omega$, to f jest stała.*

Uwaga. Twierdzenie Lagrange'a nie przenosi się dosłownie na funkcje o wartościach wektorowych. W tym przypadku mamy słabsze sformułowanie.

Twierdzenie 2.8 Lagrange'a o wartości średniej, wersja wektorowa. *Niech $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$. Jeśli F jest różniczkowalna w każdym punkcie odcinka $[a, b] \subset \Omega$, oraz $\|DF(x)\|_{L(\mathbb{R}^n, \mathbb{R}^k)} \leq M$ dla $x \in [a, b]$ to*

$$\|F(b) - F(a)\|_{\mathbb{R}^k} \leq M \|b - a\|_{\mathbb{R}^n}.$$

Definicja 2.6 *Odwzorowanie $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$ nazywamy odwzorowaniem (funkcją jeśli $k = 1$) klasy $C^1(\Omega)$ jeśli jego pochodne cząstkowe rzędu pierwszego $\frac{\partial f_j}{\partial x_i}$, $i = 1, \dots, n$, $j = 1, \dots, k$ są funkcjami ciągłymi w Ω .*

Wniosek 2.7 *Jeśli Ω jest obszarem w \mathbb{R}^n oraz $F : \Omega \rightarrow \mathbb{R}^k$ jest klasy $C^1(\Omega)$, to F spełnia warunek Lipschitza na każdym zwartym wypukłym podzbiornie $K \subset \Omega$.*

2.5 Przestrzeń styczna do wykresu funkcji

Definicja 2.7 Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Wykresem funkcji f nazywamy zbiór

$$\text{Gr}(f) = \{(x, y) \in \Omega \times \mathbb{R} : x \in \Omega, y = f(x)\}.$$

Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ będzie różniczkowalna w punkcie $\dot{x} \in \Omega$. Wówczas równanie

$$y = f(\dot{x}) + df(\dot{x}) \cdot (x - \dot{x})$$

wyznacza przestrzeń (prostą gdy $n = 1$, płaszczyznę gdy $n = 2$) do wykresu $\text{Gr}(f)$ funkcji f w punkcie \dot{x} .

Umiejętność wyznaczenia przestrzeni stycznej do wykresu jest przydatna do obliczania przybliżonej wartości funkcji. Istotnie jeśli $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ jest różniczkowalna w \dot{x} , to dla małych przyrostów argumentu Δx zachodzi

$$f(\dot{x} + \Delta x) = f(\dot{x}) + df(\dot{x}) \cdot \Delta x + \varepsilon(\Delta x)$$

przy czym

$$\frac{\varepsilon(\Delta x)}{\|\Delta x\|} \rightarrow 0 \quad \text{przy} \quad \Delta x \rightarrow 0.$$

Zatem $\varepsilon(\Delta x)$ jest małe w stosunku do Δx i można przyjąć, że

$$f(\dot{x} + \Delta x) \approx f(\dot{x}) + df(\dot{x}) \cdot \Delta x.$$

Przykład. Zbiornik ma kształt walca o wysokości $h_0 = 2 \text{ m}$ i średnicy $d_0 = 4 \text{ m}$, przy czym pomiary wykonano z dokładnością 1%. W jakich granicach może być rzeczywista objętość zbiornika i ile wynosi błąd względny.

Objętość walca wyraża się wzorem $V(h, d) = \frac{\pi}{4}hd^2$. W naszym przypadku $|h - h_0| \leq 0,2 =: \Delta h$, $|d - d_0| \leq 0,4 =: \Delta d$, $V(h_0, d_0) = 8\pi \text{ m}^3$. Zatem

$$V(h_0 + \Delta h, d_0 + \Delta d) - V(h_0, d_0) = \frac{\pi}{4}(d_0^2 \Delta h + 2h_0 d_0 \Delta d) = \frac{\pi}{4}(4^2 \cdot 0,2 + 2 \cdot 2 \cdot 4 \cdot 0,4) = 0,24\pi.$$

Błąd względny wynosi

$$\frac{V(h_0 + \Delta h, d_0 + \Delta d) - V(h_0, d_0)}{V(h_0, d_0)} = 0,03 = 3\%.$$

3 Rachunek różniczkowy drugiego rzędu

3.1 Pochodne kierunkowe drugiego rzędu

Definicja 3.1 Niech Ω będzie otwartym podzbiorem \mathbb{R}^n , $f : \Omega \rightarrow \mathbb{R}$, $\dot{x} \in \Omega$ oraz $v, w \in \mathbb{R}^n \setminus \{0\}$. Załóżmy, że pochodna kierunkowa $L_v f$ funkcji f w kierunku wektora v istnieje w otoczeniu U punktu \dot{x} . Wówczas odwzorowanie

$$U \ni x \mapsto L_v f(x)$$

jest funkcją zdefiniowaną w otoczeniu U punktu \dot{x} o wartościach w \mathbb{R} . Jeśli istnieje pochodna kierunkowa tej funkcji w kierunku wektora w w punkcie \dot{x} , to nazywamy ją pochodną kierunkową drugiego rzędu i oznaczamy

$$L_{wv}^2 f(\dot{x}) = L_w(L_v f)(\dot{x}).$$

Jeśli $v = e_i$, $w = e_j$, $i, j \in \{1, \dots, n\}$, są wersorami i -tej i j -tej osi współrzędnych, to drugą pochodną kierunkową $L_{e_j e_i}^2 f(\dot{x})$ nazywamy drugą pochodną cząstkową i oznaczamy

$$L_{e_j e_i}^2 f(\dot{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\dot{x}) = f''_{x_j x_i}(\dot{x}).$$

Jeśli $i = j$, to stosujemy też oznaczenie

$$L_{e_i e_i}^2 f(\dot{x}) = \frac{\partial^2 f}{\partial x_i^2}(\dot{x}) = f''_{x_i x_i}(\dot{x}).$$

Jeśli $i \neq j$, to $\frac{\partial^2 f}{\partial x_j \partial x_i}$ nazywamy też drugą pochodną cząstkową mieszaną.

Przykład 3.1 Niech $f(x, y) = x^a y^b$ dla $x > 0$, $y > 0$, gdzie $a, b \in \mathbb{R}$. Wówczas

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= ax^{a-1}y^b, & \frac{\partial f}{\partial y}(x, y) &= bx^a y^{b-1}; \\ \frac{\partial^2 f}{\partial x^2}(x, y) &= \frac{\partial}{\partial x}(ax^{a-1}y^b) = a(a-1)x^{a-2}y^b, \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) &= \frac{\partial}{\partial y}(ax^{a-1}y^b) = abx^{a-1}y^{b-1}, \\ \frac{\partial^2 f}{\partial x \partial y}(x, y) &= \frac{\partial}{\partial x}(bx^a y^{b-1}) = abx^{a-1}y^{b-1}, \\ \frac{\partial^2 f}{\partial y^2}(x, y) &= \frac{\partial}{\partial y}(bx^a y^{b-1}) = b(b-1)x^a y^{b-2}. \end{aligned}$$

Zauważmy, że $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$. Zatem jest uzasadnione przypuszczenie że tak jest w ogólnym przypadku. Niestety przypuszczenie to nie jest prawdziwe jak pokazuje przykład.

Przykład 3.2 Niech

$$f(x, y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & \text{dla } (x, y) \neq (0, 0), \\ 0 & \text{dla } (x, y) = (0, 0). \end{cases}$$

Wówczas

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

3.2 Druga różniczka

Definicja 3.2 Niech Ω będzie otwartym podzbiorem \mathbb{R}^n , $f : \Omega \rightarrow \mathbb{R}$, $\dot{x} \in \Omega$. Mówimy, że f jest dwukrotnie różniczkowalna w punkcie \dot{x} jeśli zachodzą warunki

1. f jest różniczkowalna w pewnym otoczeniu U punktu \dot{x} ;
2. Przy każdym ustalonym $h \in \mathbb{R}^n$ odwzorowanie

$$U \ni x \mapsto w_h(x) =: df(x) \cdot h \in \mathbb{R}$$

jest różniczkowalne w \dot{x} .

Wówczas różniczkę odwzorowania w_h nazywamy drugą różniczką f . Zatem dla $h, k \in \mathbb{R}^n$ mamy określone odwzorowanie

$$(k, h) \mapsto dw_h(\dot{x}) \cdot k =: d^2 f(\dot{x})(k, h)$$

Jest jasne, że powyższe odwzorowanie jest liniowe względem h oraz względem k . Zatem druga różniczka jest odwzorowaniem 2-liniowym na $\mathbb{R}^n \times \mathbb{R}^n$.

Jeśli f jest dwukrotnie różniczkowalna w każdym punkcie obszaru $\Omega \subset \mathbb{R}^n$, to $d^2 f$ jest odwzorowaniem z Ω w przestrzeń odwzorowań dwuliniowych na $\mathbb{R}^n \times \mathbb{R}^n$, tzn

$$d^2 f : \Omega \rightarrow L_2(\mathbb{R}^n \times \mathbb{R}^n).$$

Jeśli $f : (a, b) \subset \mathbb{R} \rightarrow \mathbb{R}$ jest dwukrotnie różniczkowalna w \dot{x} , to

$$d^2 f(\dot{x})(k, h) = f''(\dot{x}) \cdot k \cdot h.$$

Przykład 3.3 Niech $A : \mathbb{R}^n \rightarrow \mathbb{R}$ będzie funkcją liniową, tzn. $A(x) = A \cdot x$ dla $x \in \mathbb{R}^n$ i pewnej macierzy $A \in M(n \times 1)$. Wówczas $dA(x) = A$. Zatem dla $h \in \mathbb{R}^n$ odwzorowanie

$$\mathbb{R}^n \ni x \mapsto w_h(x) = dA(x) \cdot h = A \cdot h$$

nie zależy od x . Stąd $d^2 A = dw_h = 0$. Analogicznie jest w przypadku odwzorowania liniowego $A : \mathbb{R}^n \rightarrow \mathbb{R}^k$.

Przykład 3.4 Niech $A \in M(n \times n)$ oraz

$$f(x) = x^{\text{tr}} A x = \sum_{i,j=1}^n a_{ij} x_i x_j.$$

Wówczas

$$df(x)(h) = h^{\text{tr}}Ax + x^{\text{tr}}Ah \quad \text{dla } x \in \mathbb{R}^n, h \in \mathbb{R}^n.$$

Zatem przy ustalonym $h \in \mathbb{R}^n$ odwzorowanie

$$\mathbb{R}^n \ni x \mapsto w_h(x) = h^{\text{tr}}Ax + x^{\text{tr}}Ah$$

jest liniowe i jego różniczką jest

$$dw_h(x)(k) = h^{\text{tr}}Ak + k^{\text{tr}}Ah.$$

Czyli

$$d^2f(x)(k, h) = h^{\text{tr}}Ak + k^{\text{tr}}Ah \quad \text{dla } k \in \mathbb{R}^n, h \in \mathbb{R}^n.$$

W przypadku gdy macierz A jest symetryczna dostajemy

$$d^2f(x)(k, h) = 2h^{\text{tr}}Ak \quad \text{dla } k \in \mathbb{R}^n, h \in \mathbb{R}^n.$$

Twierdzenie 3.1 Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $\hat{x} \in \Omega$. Wówczas f jest dwukrotnie różniczkowalna w \hat{x} wtedy i tylko wtedy, gdy

1° f jest różniczkowalna w pewnym otoczeniu U punktu \hat{x} ;

2° Pochodne cząstkowe $\frac{\partial f}{\partial x_i}$, $i = 1, \dots, n$ są różniczkowalne w \hat{x} .

Twierdzenie 3.2 Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $\hat{x} \in \Omega$.

Jeśli f jest dwukrotnie różniczkowalna w \hat{x} , to

a) dla dowolnych $h, k \in \mathbb{R}^n$ istnieje

$$L_{kh}^2f(\hat{x}) = d^2f(\hat{x})(k, h);$$

b) Istnieją drugie pochodne cząstkowe $\frac{\partial^2 f}{\partial x_i \partial x_j}$, $i, j = 1, \dots, n$ oraz dla $h, k \in \mathbb{R}^n$,

$$d^2f(\hat{x})(k, h) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\hat{x}) h_i k_j.$$

Z twierdzenia 3.1 wynika

Wniosek 3.1 Jeśli f posiada w pewnym otoczeniu punktu \hat{x} ciągłe pochodne cząstkowe pierwszego rzędu oraz pochodne cząstkowe drugiego rzędu ciągłe w \hat{x} , to f jest dwukrotnie różniczkowalna w \hat{x} .

3.3 Symetria drugiej różniczki

Definicja 3.3 Odwzorowanie dwuliniowe $A : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ nazywamy symetrycznym, jeśli

$$A(x, y) = A(y, x) \quad \text{dla } x, y \in \mathbb{R}^n.$$

Twierdzenie 3.3 Schwarz o symetrii drugiej różniczki.

Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $\hat{x} \in \Omega$. Jeśli f jest dwukrotnie różniczkowalna w \hat{x} , to druga różniczka $d^2f(\hat{x})$ jest odwzorowaniem dwuliniowym symetrycznym, tzn.

$$L_{kh}^2f(\hat{x}) = L_{hk}^2f(\hat{x}).$$

W szczególności

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\hat{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\hat{x}) \quad \text{dla } i, j = 1, \dots, n.$$

3.4 Macierz Hessego

Niech f będzie dwukrotnie różniczkowalna w $\hat{x} \in \Omega$. Połóżmy

$$a_{ij} = \frac{\partial^2 f}{\partial x_j \partial x_i}(\hat{x}) \quad \text{dla } i, j = 1, \dots, n,$$

$$A = \left(a_{ij} \right)_{i,j=1}^n.$$

Wówczas macierz A jest symetryczna oraz na mocy wzoru z Twierdzenia 3.2

$$d^2 f(\hat{x})(k, h) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\hat{x}) h_i k_j = k^{\text{tr}} A h = h^{\text{tr}} A k.$$

Definicja 3.4 Formę kwadratową

$$\mathbb{R}^n \ni h \mapsto h^{\text{tr}} A h = d^2 f(\hat{x})(h, h) \in \mathbb{R}$$

nazywamy formą kwadratową Hessego, a odpowiadającą jej macierz A macierzą Hessego lub hesjanem funkcji f w punkcie \hat{x} .

Ślad hesjanu nazywamy operatorem Laplace'a

$$\text{Tr} A = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(x) = \Delta f(x).$$

3.5 Wzór Taylora drugiego rzędu.

Twierdzenie 3.4 Niech $f : \Omega \rightarrow \mathbb{R}$ będzie dwukrotnie różniczkowalna w $\hat{x} \in \Omega$. Wówczas dla $h \in \mathbb{R}^n$ takiego, że $\hat{x} + h \in \Omega$ zachodzi

$$f(\hat{x} + h) = f(\hat{x}) + df(\hat{x})h + \frac{1}{2}d^2 f(\hat{x})(h, h) + \|h\|^2 \psi(h),$$

gdzie ψ jest ciągła w zerze oraz $\psi(0) = 0$.

Definicja 3.5 Niech $f : \Omega \rightarrow \mathbb{R}$. Mówimy, że funkcja f jest klasy $C^2(\Omega)$ jeśli dla dowolnych ustalonych $h, k \in \mathbb{R}^n$ odwzorowanie

$$\Omega \ni x \mapsto d^2 f(x)(k, h) \in \mathbb{R}$$

jest ciągle lub równoważnie, gdy wszystkie pochodne cząstkowe drugiego rzędu $\frac{\partial^2 f}{\partial x_j \partial x_i}(\hat{x})$, $i, j = 1, \dots, n$, są ciągłe w Ω .

4 Rachunek różniczkowy k -tego rzędu

4.1 k -ta różniczka

Wzorując się na definicji drugiej różniczki można indukcyjnie zdefiniować różniczki wyższych rzędów.

Definicja 4.1 Niech Ω będzie otwartym podzbiorem \mathbb{R}^n , $f : \Omega \rightarrow \mathbb{R}$, $\dot{x} \in \Omega$ oraz $k \in \mathbb{N}$, $k \geq 2$. Mówimy, że f jest k -krotnie różniczkowalna w punkcie \dot{x} jeśli zachodzą warunki

1. f jest $(k - 1)$ -krotnie różniczkowalna w pewnym otoczeniu U punktu \dot{x} ;
2. Dla dowolnych ustalonych $h^{(1)}, \dots, h^{(k-1)} \in \mathbb{R}^n$ odwzorowanie

$$U \ni x \mapsto w_{h^{(1)}, \dots, h^{(k-1)}}(x) =: d^{k-1}f(x)(h^{(1)}, \dots, h^{(k-1)}) \in \mathbb{R}$$

jest różniczkowalne w \dot{x} .

Wówczas różniczkę odwzorowania $w_{h^{(1)}, \dots, h^{(k-1)}}$ nazywamy k -tą różniczką f . Zatem dla $h, h^{(1)}, \dots, h^{(k-1)} \in \mathbb{R}^n$ mamy określone odwzorowanie

$$(h, h^{(1)}, \dots, h^{(k-1)}) \mapsto dw_{h^{(1)}, \dots, h^{(k-1)}}(\dot{x}) \cdot h =: d^k f(\dot{x})(h, h^{(1)}, \dots, h^{(k-1)}).$$

k -ta różniczka jest odwzorowaniem k -liniowym na $\mathbb{R}^n \times \dots \times \mathbb{R}^n$ o wartościach w \mathbb{R} .

Twierdzenia dotyczące drugiej różniczki odpowiednio przenoszą się na przypadek k -tej różniczki.

Twierdzenie 4.1 Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $\dot{x} \in \Omega$. Wówczas f jest k -krotnie różniczkowalna w \dot{x} wtedy i tylko wtedy, gdy

- 1° f jest $(k - 1)$ -krotnie różniczkowalna w pewnym otoczeniu U punktu \dot{x} ;
- 2° Pochodne cząstkowe $(k - 1)$ rzędu są różniczkowalne w \dot{x} .

Twierdzenie 4.2 Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $\dot{x} \in \Omega$.

Jeśli f jest k -krotnie różniczkowalna w \dot{x} , to

- a) dla dowolnych wektorów $h^{(1)}, \dots, h^{(k)} \in \mathbb{R}^n$ istnieje k -ta pochodna kierunkowa

$$L_{h^{(k)}, \dots, h^{(1)}}^k f(\dot{x}) = d^k f(\dot{x})(h^{(1)}, \dots, h^{(k)});$$

- b) Istnieją pochodne cząstkowe rzędu k oraz dla $h^{(1)}, \dots, h^{(k)} \in \mathbb{R}^n$,

$$d^k f(\dot{x})(h^{(1)}, \dots, h^{(k)}) = \sum_{\alpha \in \mathbb{N}_0^k, \alpha_i \leq n} \frac{\partial^k f}{\partial x_{\alpha_1} \dots \partial x_{\alpha_k}}(\dot{x}) h_{\alpha_1}^{(1)} \dots h_{\alpha_k}^{(k)}.$$

W szczególności dla $k = 3$,

$$d^3 f(\dot{x})(h^{(1)}, h^{(2)}, h^{(3)}) = \sum_{\alpha_1, \alpha_2, \alpha_3=1}^n \frac{\partial^3 f}{\partial x_{\alpha_1} \partial x_{\alpha_2} \partial x_{\alpha_3}}(\dot{x}) h_{\alpha_1}^{(1)} h_{\alpha_2}^{(2)} h_{\alpha_3}^{(3)}.$$

Z twierdzenia 4.1 wynika

Wniosek 4.1 Jeśli f posiada w pewnym otoczeniu punktu \dot{x} ciągłe wszystkie pochodne cząstkowe rzędu $k - 1$ oraz pochodne cząstkowe rzędu k ciągłe w \dot{x} , to f jest k -krotnie różniczkowalna w \dot{x} .

4.2 Symetria k -tej różniczki

Twierdzenie 4.3 Schwarz'a o symetrii k -tej różniczki.

Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $\dot{x} \in \Omega$. Jeśli f jest k -krotnie różniczkowalna w \dot{x} , to k -ta różniczka $d^k f(\dot{x})$ jest odwzorowaniem k -liniowym symetrycznym, tzn.

$$d^k f(\dot{x})(h^{(1)}, \dots, h^{(k)}) = d^k f(\dot{x})(h^{(\sigma(1))}, \dots, h^{(\sigma(k))}),$$

gdzie $\sigma : \{1, \dots, k\} \rightarrow \{1, \dots, k\}$ jest dowolną permutacją. W szczególności, pochodne cząstkowe mieszane nie zależą od kolejności różniczkowania.

4.3 Wzór Taylora

Twierdzenie 4.4 Wzór Taylora.

Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ będzie k -krotnie różniczkowalna w każdym punkcie odcinka $[a, b] \subset \Omega$. Wówczas istnieje punkt $c \in [a, b]$ taki, że

$$f(b) = f(a) + \frac{df(a)}{1!}(b-a) + \dots + \frac{d^{k-1}f(a)}{(k-1)!}(b-a, \dots, b-a) + \frac{d^k f(c)}{k!}(b-a, \dots, b-a).$$

4.4 Funkcje klasy C^k

Definicja 4.2 Niech $f : \Omega \rightarrow \mathbb{R}$ oraz $k \in \mathbb{N}$. Mówimy, że funkcja f jest funkcją klasy $C^k(\Omega)$ jeśli jest ona k -krotnie różniczkowalna w każdym punkcie $x \in \Omega$ oraz dla dowolnych ustalonych $h^{(1)}, \dots, h^{(k)} \in \mathbb{R}^n$ odwzorowanie

$$\Omega \ni x \mapsto d^k f(\dot{x})(h^{(1)}, \dots, h^{(k)}) \in \mathbb{R}$$

jest ciągłe lub równoważnie, gdy wszystkie pochodne cząstkowe k -tego rzędu $\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(\dot{x})$, $i_j = 1, \dots, n$ dla $j = 1, \dots, k$, są ciągłe w Ω .

Definicja 4.3 Niech $f : \Omega \rightarrow \mathbb{R}$. Mówimy, że funkcja f jest funkcją klasy $C^\infty(\Omega)$ lub że jest funkcją gładką na Ω jeśli jest ona klasy $C^k(\Omega)$ dla dowolnego $k \in \mathbb{N}$ lub równoważnie, gdy wszystkie pochodne cząstkowe dowolnego rzędu są ciągłe w Ω .

Funkcjami gładkimi są wielomiany, funkcja wykładnicza, logarytmiczna, funkcje trygonometryczne, sumy, iloczyny funkcji gładkich, iloraz funkcji gładkich poza zerami mianownika, złożenia funkcji gładkich.

Definicja 4.4 Niech $f : \Omega \rightarrow \mathbb{R}$ będzie funkcją ciągłą. Nośnikiem funkcji f nazywamy domknięcie zbioru tych punktów $x \in \Omega$ takich, że $f(x) \neq 0$. Nośnik funkcji oznaczamy jako $\text{supp} f$, tak więc,

$$\text{supp} f = \overline{\{x \in \Omega : f(x) \neq 0\}}.$$

Przykład 4.1 *Niech*

$$g(t) = \begin{cases} e^{-1/t} & \text{dla } t > 0, \\ 0 & \text{dla } t \leq 0. \end{cases}$$

Wówczas f jest funkcją gładką na \mathbb{R} o nośniku $\overline{\mathbb{R}_+} = [0, \infty)$.

Dowód Jest jasne, że f jest klasy C^∞ na $\mathbb{R} \setminus \{0\}$. Następnie $f \in C^0(\mathbb{R})$, gdyż $\lim_{t \rightarrow 0^+} e^{-1/t} = 0$. Pozostaje wykazać, że dla dowolnego $k \in \mathbb{N}$ pochodna rzędu k jest funkcją ciągłą w zerze. W tym celu indukcyjnie dowodzi się, że

$$f^{(k)}(t) = e^{-1/t} \cdot W_{2k}(\frac{1}{t}) \quad \text{dla } t > 0,$$

gdzie W_{2k} jest pewnym wielomianem stopnia $2k$. W celu wykazania, że $f^{(k)}$ jest funkcją ciągłą w zerze wystarczy zatem wykazać następujący fakt.

Dla każdego $N \in \mathbb{N}$ istnieje stała $C_N < \infty$ taka, że

$$e^{-1/t} \leq C_N t^N \quad \text{dla } t > 0$$

lub równoważnie

$$x^N e^{-x} \leq C_N \quad \text{dla } x > 0.$$

Dowód tej nierówności można uzyskać badając przebieg zmienności funkcji $g(x) = x^N e^{-x}$, $x > 0$. Otóż funkcja ta jest rosnąca dla $0 < x < N$ i malejąca dla $x > N$.

W punkcie $x = N$ przyjmuje maksimum równe $N^N e^{-N} =: C_N$. \square

Podamy teraz przykład funkcji gładkiej na \mathbb{R}^n o nośniku równym kuli jednostkowej.

Przykład 4.2 *Niech*

$$g(t) = \begin{cases} e^{1/(t-1)} & \text{dla } t < 1, \\ 0 & \text{dla } t \geq 1 \end{cases}$$

oraz

$$f(x) = g(x^2) = g(x_1^2 + \dots + x_n^2) \quad \text{dla } x \in \mathbb{R}^n.$$

Wówczas f jest funkcją gładką na \mathbb{R} , której nośnikiem jest kula jednostkowa, tzn. $f(x) > 0$ dla $\|x\| < 1$ oraz $f(x) = 0$ dla $\|x\| \geq 1$.

Dowód. Jest jasne, że $\text{supp} f = \overline{B(0,1)}$. Indukcyjnie wykazuje się, że dowolna pochodna cząstkowa rzędu $k \in \mathbb{N}$ jest postaci

$$W(x_1, \dots, x_n, g'(x^2), \dots, g^{(k)}(x^2)),$$

gdzie W jest pewnym wielomianem. Zatem $f \in C^\infty(\mathbb{R}^n)$. \square

5 Ekstrema lokalne funkcji

Definicja 5.1 Niech Ω będzie otwartym podzbiorem \mathbb{R}^n , $f : \Omega \rightarrow \mathbb{R}$ oraz $\hat{x} \in \Omega$. Mówimy, że f posiada w punkcie \hat{x} *lokalne minimum (maksimum)* jeśli istnieje $\delta > 0$ taka, że

$$\begin{aligned} f(x) &\geq f(\hat{x}) \quad \text{dla } \|x - \hat{x}\| < \delta, \\ (f(x) &\leq f(\hat{x}) \quad \text{dla } \|x - \hat{x}\| < \delta). \end{aligned}$$

Lokalne minimum lub maksimum nazywamy *lokalnym ekstremum*.

5.1 Warunek konieczny 1-go rzędu

Twierdzenie 5.1 Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ oraz $\hat{x} \in \Omega$. Jeśli f posiada pochodne cząstkowe pierwszego rzędu w \hat{x} oraz \hat{x} jest punktem lokalnego ekstremum, to

$$\frac{\partial f}{\partial x_1}(\hat{x}) = \dots = \frac{\partial f}{\partial x_n}(\hat{x}).$$

Wniosek 5.1 Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ oraz $\hat{x} \in \Omega$. Jeśli f jest różniczkowalna w \hat{x} oraz \hat{x} jest punktem lokalnego ekstremum, to

$$df(\hat{x}) = 0.$$

Warto podkreślić, że warunek znikania pochodnych cząstkowych pierwszego rzędu jest tylko warunkiem koniecznym ekstremum lokalnego.

Przykład 5.1 Niech

$$f(x, y) = xy$$

Wówczas

$$f'_x(0, 0) = f'_y(0, 0) = 0,$$

ale f nie posiada ekstremum w punkcie $(0, 0)$.

Tym nie mniej warunek znikania pochodnych pierwszego rzędu pozwala ustalić punkty podejrzane o ekstrema.

Definicja 5.2 Niech Ω będzie otwartym podzbiorem \mathbb{R}^n , $f : \Omega \rightarrow \mathbb{R}$.

Punkt $\hat{x} \in \Omega$ nazywamy *punktem stacjonarnym (lub krytycznym)* funkcji f jeśli istnieją pochodne cząstkowe $\frac{\partial f}{\partial x_i}(\hat{x})$ dla $i = 1, \dots, n$ i są równe zeru.

Znajomość punktów stacjonarnych funkcji jest przydatna do wyznaczenia największej i najmniejszej wartości funkcji na zbiorze zwartym $K \subseteq \mathbb{R}^n$. Jak wiemy funkcja ciągła przyjmuje na zbiorze zwartym K swoje kresy. Zatem jeśli jest ona różniczkowalna we wnętrzu zbioru K , to wartości maksymalne i minimalne mogą być przyjmowane w punktach stacjonarnych wnętrza K lub na brzegu K .

Przykład 5.2 Niech

$$f(x, y) = 2x^2 - xy + y^2, \quad K = \{(x, y) : x^2 + y^2 \leq 1\}.$$

Wówczas $\text{int}K = \{(x, y) : x^2 + y^2 < 1\}$. Układem równań na punkty stacjonarne jest

$$\begin{aligned} \frac{\partial f}{\partial x} &= 4x - y = 0, \\ \frac{\partial f}{\partial y} &= -x + 2y = 0. \end{aligned}$$

Punktem stacjonarnym jest $O = (0, 0)$ oraz $f(0, 0) = 0$. W celu zbadania f na brzegu zbioru K zauważmy, że $\partial K = \{x^2 + y^2 = 1\} = \{(\cos \varphi, \sin \varphi) : 0 \leq \varphi \leq 2\pi\}$. Zatem

$$\begin{aligned} f(\cos \varphi, \sin \varphi) &= 2 \cos^2 \varphi - \cos \varphi \sin \varphi + \sin^2 \varphi \\ &= \cos^2 \varphi - \frac{1}{2} \sin 2\varphi + 1 =: g(\varphi). \end{aligned}$$

Liczmy pochodną funkcji g :

$$g'(\varphi) = -2 \cos \varphi \sin \varphi - \cos 2\varphi = -\sin 2\varphi - \cos 2\varphi.$$

$g'(\varphi) = 0$ wtedy i tylko wtedy, gdy $\varphi = \frac{3}{8}\pi$ lub $\varphi = \frac{7}{8}\pi$.

$g(\frac{3}{8}\pi) = \frac{1}{2}(3 - \sqrt{2}) > 0$, $g(\frac{7}{8}\pi) = \frac{1}{2}(3 + \sqrt{2})$. Zatem

$$\max_K f = f(\cos \frac{7}{8}\pi, \sin \frac{7}{8}\pi) = \frac{1}{2}(3 + \sqrt{2}), \quad \min_K f = f(0, 0) = 0. \quad \square$$

W przypadku gdy mamy wyznaczyć ekstrema funkcji na zbiorze domkniętym $F \subset \mathbb{R}^n$, poza punktami stacjonarnymi wnętrza zbioru F należy uwzględnić ekstrema funkcji na brzegu zbioru F oraz zachowanie się funkcji dla $F \ni x \rightarrow \infty$.

Przykład 5.3 Niech

$$f(x, y) = (x + y)e^{-x^2 - y^2}, \quad F = \{(x, y) : x \geq 0, y \geq 0\}.$$

Wówczas $\text{int}K = \{(x, y) : x > 0, y > 0\}$. Układem równań na punkty stacjonarne jest

$$\begin{aligned} \frac{\partial f}{\partial x} &= (-2x^2 - 2xy + 1)e^{-x^2 - y^2} = 0, \\ \frac{\partial f}{\partial y} &= (-2y^2 - 2xy + 1)e^{-x^2 - y^2} = 0. \end{aligned}$$

Punktem stacjonarnym we wnętrzu zbioru F jest $A = (\frac{1}{2}, \frac{1}{2})$ oraz $f(\frac{1}{2}, \frac{1}{2}) = e^{-1/2}$. W celu zbadania f na brzegu zbioru K zauważmy, że $\text{int}F = I_1 \cup I_2$, gdzie $I_1 = \{(x, 0) : x \geq 0\}$, $I_2 = \{(0, y) : y \geq 0\}$ Następnie

$$\max_{I_1} f = \max_{0 \leq x < \infty} xe^{-x^2} = f(\frac{\sqrt{2}}{2}, 0) = \frac{\sqrt{2}}{2}e^{-1/2}, \quad \min_{I_1} f = f(0, 0) = 0.$$

Analogicznie $\max_{I_2} f = f(0, \frac{\sqrt{2}}{2}) = \frac{\sqrt{2}}{2}e^{-1/2}$, $\min_{I_2} f = f(0, 0) = 0$.

W celu zbadania zachowania się funkcji f przy $(x, y) \ni F \rightarrow \infty$ połączmy $x = r \cos \varphi$, $y = r \sin \varphi$ dla $r > 0, 0 \leq \varphi \leq \pi/2$. Wówczas

$$f(r \cos \varphi, r \sin \varphi) = (\cos \varphi + \sin \varphi) r e^{-r^2} \rightarrow 0 \quad \text{gdy} \quad r \rightarrow \infty.$$

Zatem $\max_F f = \max(e^{-1/2}, \frac{\sqrt{2}}{2}e^{-1/2}) = e^{-1/2} = f(\frac{1}{2}, \frac{1}{2})$, $\min_F f = f(0, 0) = 0$. \square

Następny przykład pokazuje, że punktów stacjonarnych może być nieskończenie wiele.

Przykład 5.4 Niech

$$f(x, y) = xy^2, \quad \Omega = \mathbb{R}^2.$$

Wówczas rozwiązaniem układu równań na punkty stacjonarne

$$\begin{aligned} \frac{\partial f}{\partial x} &= y^2 = 0, \\ \frac{\partial f}{\partial y} &= 2xy = 0 \end{aligned}$$

jest cała prosta $\{y = 0\}$. Łatwo zauważyć, że w punktach $(x, 0)$, $x > 0$, funkcja f posiada minimum lokalne równe 0, natomiast w punktach $(x, 0)$, $x < 0$, funkcja f posiada maksimum lokalne równe 0. W punkcie $(0, 0)$ nie ma ani minimum ani maksimum lokalnego.

5.2 Forma kwadratowa Hessego

Warunek dostateczny na ekstremum funkcji – podobnie jak w przypadku 1-wymiarowym – można sformułować przy pomocy drugiej różniczki. Przypomnijmy, że druga różniczka funkcji to forma dwuliniowa

$$d^2 f(\hat{x})(k, h) = \sum_{i,j=1}^n \frac{\partial^2 f(\hat{x})}{\partial x_i \partial x_j} k_i h_j \quad \text{dla} \quad k, h \in \mathbb{R}^n.$$

Jej wartość na przekątnej $\{k = h\}$ nazywamy *formą kwadratową Hessego stowarzyszoną z $d^2 f(\hat{x})$* ,

$$H(h) = d^2 f(\hat{x})(h, h) = \sum_{i,j=1}^n \frac{\partial^2 f(\hat{x})}{\partial x_i \partial x_j} h_i h_j \quad \text{dla} \quad h \in \mathbb{R}^n.$$

Oznaczmy

$$A = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\hat{x}) \right)_{i,j=1}^n.$$

Wówczas, jeśli f jest dwukrotnie różniczkowalna w \hat{x} , to

$$H(h) = h^{\text{tr}} A h \quad \text{dla} \quad h \in \mathbb{R}^n.$$

Definicja 5.3 Formę kwadratową H na \mathbb{R}^n nazywamy *dodatnio* (odpowiednio, *ujemnie*) *określoną* jeśli $H(h) > 0$ (odpowiednio, $H(h) < 0$) dla każdego $h \in \mathbb{R}^n \setminus \{0\}$. Jeśli forma h przyjmuje zarówno wartości dodatnie jak i ujemne, to mówimy, że jest ona *nieokreślona*.

Zauważmy, że jeśli forma kwadratowa H jest dodatnio określona, to istnieje stała $M > 0$ taka, że

$$H(h) \geq Mh^2 \quad \text{dla } h \in \mathbb{R}^n.$$

Istotnie, jeśli H jest dodatnio określona, to przyjmuje wartości dodatnie na sferze jednostkowej i wobec zwartości sfery jednostkowej istnieje $M > 0$ takie, że

$$H(h) \geq M \quad \text{dla } h \in \mathbb{R}^n, \|h\| = 1.$$

Zatem

$$H\left(\frac{h}{\|h\|}\right) = \frac{1}{\|h\|^2} H(h) \geq M \quad \text{dla } h \in \mathbb{R}^n, h \neq 0.$$

Stąd

$$H(h) \geq M\|h\|^2 = Mh^2 \quad \text{dla } h \in \mathbb{R}^n. \quad \square$$

W przyszłości wykazemy, że dla macierzy symetrycznej $A \in M(n \times n)$ forma kwadratowa

$$H(h) = h^{\text{tr}} Ah \quad \text{dla } h \in \mathbb{R}^n$$

jest dodatnio określona, wtedy i tylko wtedy, gdy wszystkie wartości własne macierzy A są dodatnie.

Twierdzenie 5.2 Kryterium Sylwestera. *Niech $A \in M(n \times n)$ będzie macierzą symetryczną. Wówczas forma kwadratowa*

$$H(h) = h^{\text{tr}} Ah \quad \text{dla } h \in \mathbb{R}^n$$

jest dodatnio określona, wtedy i tylko wtedy, gdy wyznaczniki wszystkich minorów głównych są dodatnie, tzn.,

$$\det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1l} \\ a_{21} & a_{22} & \dots & a_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l1} & a_{l2} & \dots & a_{ll} \end{pmatrix} > 0 \quad \text{dla } l = 1, \dots, n.$$

Wniosek 5.2 Kryterium Sylwestera. *Niech $A \in M(n \times n)$ będzie macierzą symetryczną. Wówczas forma kwadratowa*

$$F(h) = h^{\text{tr}} Ah \quad \text{dla } h \in \mathbb{R}^n$$

jest ujemnie określona, wtedy i tylko wtedy, gdy

$$(-1)^l \det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1l} \\ a_{21} & a_{22} & \dots & a_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l1} & a_{l2} & \dots & a_{ll} \end{pmatrix} > 0 \quad \text{dla } l = 1, \dots, n.$$

5.3 Warunek dostateczny 2-go rzędu

Twierdzenie 5.3 Warunek dostateczny drugiego rzędu ekstremum lokalnego.

Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ oraz $\hat{x} \in \Omega$. Załóżmy, że f jest dwukrotnie różniczkowalna w \hat{x} , przy czym $df(\hat{x}) = 0$. Jeśli forma kwadratowa Hessego

$$H(h) = d^2f(\hat{x})(h, h) \quad \text{dla } h \in \mathbb{R}^n$$

jest dodatnio (odpowiednio, ujemnie) określona, to f przyjmuje w \hat{x} lokalne minimum (odpowiednio, lokalne maksimum).

Jeśli forma H jest nieokreślona, to f nie ma ekstremum w \hat{x} (\hat{x} jest wtedy punktem siodłowym funkcji f).

Twierdzenie 5.4 Warunek dostateczny drugiego rzędu ekstremum globalnego.

Niech $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ oraz $\hat{x} \in \Omega$. Załóżmy, że f jest klasy $C^2(\Omega)$ oraz $df(\hat{x}) = 0$. Jeśli forma kwadratowa Hessego

$$H(h) = d^2f(x)(h, h) \quad \text{dla } h \in \mathbb{R}^n$$

jest dodatnio (odpowiednio, ujemnie) określona dla każdego $x \in \Omega$, to f przyjmuje w \hat{x} ściśle minimum (odpowiednio, ściśle maksimum) globalne.

Przykład 5.5 Niech

$$f(x, y) = (1 + e^y) \cos x - ye^y, \quad \Omega = \mathbb{R}^2.$$

Wówczas rozwiązaniem układu równań na punkty stacjonarne

$$\begin{aligned} \frac{\partial f}{\partial x} &= -(1 + e^y) \sin x = 0, \\ \frac{\partial f}{\partial y} &= e^y(\cos x - 1 - y) = 0 \end{aligned}$$

są punkty $A_k = (2k\pi, 0)$, $k \in \mathbb{Z}$ oraz $B_k = ((2k + 1)\pi, -2)$, $k \in \mathbb{Z}$. Macierzą Hessego jest

$$H(x, y) = \begin{pmatrix} -(1 + e^y) \cos x & -e^y \sin x \\ -e^y \sin x & e^y(\cos x - 2 - y) \end{pmatrix}.$$
$$H(A_k) = \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix}, \quad H(B_k) = \begin{pmatrix} 1 + e^{-2} & 0 \\ 0 & -e^{-2} \end{pmatrix}$$

W punktach A_k mamy $H(A_k) < 0$. Zatem funkcja przybiera w tych punktach maksimum lokalne równe $f(2k\pi, 0) = 2$. W punktach B_k macierz Hessego jest nieokreślona, a zatem są to punkty siodłowe. \square

Przykład 5.6 Niech $A \in M(n \times n)$ będzie macierzą symetryczną, $B \in M(n \times 1)$ wektorem poziomym oraz $c \in \mathbb{R}$. Rozważmy funkcję kwadratową n -zmiennych

$$f(x) = x^{\text{tr}}Ax + Bx + c \quad \text{dla } x \in \mathbb{R}^n.$$

Jeśli $A = (a_{ij})_{i,j=1}^n$, $B = (b_1, \dots, b_n)$, to

$$\begin{aligned} f(x) &= \sum_{i,j=1}^n a_{ij}x_i x_j + \sum_{i=1}^n b_i x_i + c \\ &= x_1(a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n) \\ &\quad + x_2(a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n) \\ &\quad \vdots \\ &\quad + x_n(a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n) \\ &\quad + b_1x_1 + b_2x_2 + \dots + b_nx_n + c. \end{aligned}$$

Układem równań na punkty stacjonarne jest układ

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= 2(a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n) + b_1 = 0, \\ &\quad \vdots \\ \frac{\partial f}{\partial x_n} &= 2(a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n) + b_n = 0, \end{aligned}$$

czyli

$$\text{grad} f(x) = 2x^{\text{tr}}A + b = 0.$$

Jeśli macierz A jest odwracalna, to istnieje dokładnie jeden punkt stacjonarny

$$\hat{x} = -\frac{1}{2}A^{-1}b^{\text{tr}}.$$

Ponadto, $H(x) = 2A$. Zatem, jeśli $A > 0$, to \hat{x} jest globalnym, ścisłym minimum; jeśli $A < 0$, to \hat{x} jest globalnym, ścisłym maksimum; jeśli A jest nieokreślona, to \hat{x} jest punktem siodłowym. \square

5.4 Zastosowania

A. Odległość prostych

Prosta l w \mathbb{R}^n jest zadana przez punkt $a \in \mathbb{R}^n$ i wektor $v \in \mathbb{R}^n$. Zatem niech l_1, l_2 będą dwoma prostymi:

$$\begin{aligned} l_1 &= \{x \in \mathbb{R}^n : x = a + tv \text{ dla } t \in \mathbb{R}\}, \quad \text{gdzie } a \in \mathbb{R}^n, v \in \mathbb{R}^n, \\ l_2 &= \{y \in \mathbb{R}^n : y = b + sw \text{ dla } s \in \mathbb{R}\}, \quad \text{gdzie } b \in \mathbb{R}^n, w \in \mathbb{R}^n. \end{aligned}$$

Wówczas odległość prostej l_1 od prostej l_2 jest dana wzorem

$$\rho(l_1, l_2) = \min_{x \in l_1, y \in l_2} \|x - y\|.$$

Zatem

$$\rho^2(l_1, l_2) = \min_{(t,s) \in \mathbb{R}^2} \sum_{i=1}^n (a_i - b_i + tv_i - sw_i)^2.$$

Czyli trzeba znaleźć minimum funkcji kwadratowej

$$f(t, s) = \sum_{i=1}^n (a_i - b_i + tv_i - sw_i)^2.$$

B. Odległość prostej od hiperpłaszczyzny

Odległość punktu $p \in \mathbb{R}^n$ od $(n-1)$ -wymiarowej hiperpłaszczyzny Π zadanej równaniem

$$\Pi = \{x \in \mathbb{R}^n : b^{\text{tr}}x = c\}, \quad \text{gdzie } b \in \mathbb{R}^n, c \in \mathbb{R}$$

wyraża się wzorem

$$\rho^2(p, \Pi) = \min_{x \in \mathbb{R}^{n-1}} \left(\sum_{i=1}^{n-1} (p_i - x_i)^2 + \left(p_n - \frac{1}{b_n} \left(c - \sum_{i=1}^{n-1} b_i x_i \right) \right)^2 \right)$$

o ile $b_n \neq 0$.

Jeśli $n = 3$, $\Pi = \{b_1 x_1 + b_2 x_2 + b_3 x_3 = c\}$, przy czym $b_3 \neq 0$, to

$$\rho^2(p, \Pi) = \min_{(x_1, x_2) \in \mathbb{R}^2} \left((p_1 - x_1)^2 + (p_2 - x_2)^2 + \left(p_3 - \frac{c - b_1 x_1 - b_2 x_2}{b_3} \right)^2 \right).$$

C. Nierówność między średnią geometryczną i arytmetyczną

Zajmiemy się teraz problemem wyznaczenia najmniejszej wartości sumy n liczb dodatnich, których iloczyn jest ustalony.

Problem polega na znalezieniu minimum wyrażenia $x_1 + x_2 + \dots + x_n$ jeśli $x_1 > 0, x_2 > 0, \dots, x_n > 0$ oraz $x_1 \cdot x_2 \cdot \dots \cdot x_n = C, C > 0$. Z ostatniego warunku dostajemy

$$x_n = \frac{C}{x_1 \cdot x_2 \cdot \dots \cdot x_{n-1}}.$$

Zatem wystarczy znaleźć minimum funkcji

$$f(x_1, x_2, \dots, x_{n-1}) = x_1 + x_2 + \dots + x_{n-1} + \frac{C}{x_1 \cdot x_2 \cdot \dots \cdot x_{n-1}}$$

w obszarze $\mathbb{R}_+^{n-1} = \{x_1 > 0, x_2 > 0, \dots, x_{n-1} > 0\}$. Piszemy układ równań na punkt stacjonarny

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= 1 - \frac{C}{x_1^2 \cdot x_2 \cdot \dots \cdot x_{n-1}} = 0, \\ \frac{\partial f}{\partial x_2} &= 1 - \frac{C}{x_1 \cdot x_2^2 \cdot \dots \cdot x_{n-1}} = 0, \\ &\vdots \\ \frac{\partial f}{\partial x_{n-1}} &= 1 - \frac{C}{x_1 \cdot x_2 \cdot \dots \cdot x_{n-1}^2} = 0. \end{aligned}$$

Stąd $x_1 = x_2 = \dots = x_{n-1} = C^{1/n}$. Łatwo zauważyć, że jeśli $x_i \rightarrow 0$ lub $x_i \rightarrow \infty$ dla pewnego $i = 1, 2, \dots, n-1$, to $f(x_1, x_2, \dots, x_{n-1}) \rightarrow \infty$. Zatem punkt stacjonarny funkcji f jest punktem jej minimum globalnego, czyli

$$f(x_1, x_2, \dots, x_{n-1}) \geq f(C^{1/n}, C^{1/n}, \dots, C^{1/n}) = nC^{1/n} \quad \text{dla } (x_1, x_2, \dots, x_{n-1}) \in \mathbb{R}_+^{n-1}.$$

Stąd dostajemy nierówność pomiędzy średnimi arytmetyczną i geometryczną n liczb dodatnich

$$\frac{x_1 + x_2 + \dots + x_{n-1} + x_n}{n} \geq \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_{n-1} \cdot x_n}.$$

Korzystając z powyższej nierówności można rozwiązać następujące zadanie.

Zadanie 1. Znaleźć trójkąt o danym obwodzie $2p$, którego pole P jest największe. Oznaczając przez a, b, c boki trójkąta zadanie sprowadza się do wyznaczenia maksimum funkcji

$$P(a, b, c) = \sqrt{p(p-a)(p-b)(p-c)}$$

przy warunku $a + b + c = 2p$ lub równoważnie maksimum iloczynu

$$(p-a)(p-b)(p-c)$$

pod warunkiem, że suma $(p-a) + (p-b) + (p-c) = p$ jest stała.

Zadanie 2. Znaleźć $(n+1)$ -kąąt o największym polu P wpisany w koło o promieniu $R > 0$.

Oznaczmy przez $\varphi_1, \varphi_2, \dots, \varphi_{n+1}$ kąty środkowe $(n+1)$ -kąta. Wówczas $\varphi_1 + \varphi_2 + \dots + \varphi_{n+1} = 2\pi$. Pole $(n+1)$ -kąta wpisanego w koło o promieniu R wyraża się wzorem

$$P = \frac{1}{2}R^2(\sin \varphi_1 + \dots + \sin \varphi_{n+1}).$$

Zadanie sprowadza się do znalezienia maksimum funkcji

$$u(\varphi_1, \dots, \varphi_n) = \sin \varphi_1 + \dots + \sin \varphi_n - \sin(\varphi_1 + \dots + \varphi_n)$$

w zbiorze $F = \{\varphi_1 \geq 0, \dots, \varphi_n \geq 0, \varphi_1 + \dots + \varphi_n \leq 2\pi\}$. Punkt stacjonarny leżący we wnętrzu tego zbioru spełnia układ równań

$$\begin{aligned} \frac{\partial u}{\partial \varphi_1} &= \cos \varphi_1 - \cos(\varphi_1 + \dots + \varphi_n) = 0, \\ &\vdots \\ \frac{\partial u}{\partial \varphi_n} &= \cos \varphi_n - \cos(\varphi_1 + \dots + \varphi_n) = 0. \end{aligned}$$

Zatem $\varphi_1 = \dots = \varphi_n = \frac{2\pi}{n+1} = \varphi_{n+1}$ oraz $u_{max} = u(\frac{2\pi}{n+1}, \dots, \frac{2\pi}{n+1}) = (n+1) \sin \frac{2\pi}{n+1}$. Zauważmy jeszcze, że na ścianie $\{\varphi_n = 0\}$ zbioru F funkcja maksimum funkcji u wynosi $n \sin \frac{2\pi}{n} < (n+1) \sin \frac{2\pi}{n+1}$, gdyż funkcja $x \mapsto \frac{\sin(2\pi x)}{x}$ jest malejąca w przedziale $(0, \frac{1}{2})$.

D. Metoda najmniejszych kwadratów, regresja liniowa

Załóżmy, że na płaszczyźnie \mathbb{R}^2 danych jest n punktów

$$A_1 = (x_1, y_1), \dots, A_n = (x_n, y_n).$$

Zadanie *regresji liniowej* polega na znalezieniu prostej $y = ax + b$ leżącej najbliżej tych punktów w sensie, że suma kwadratów $\sum_{i=1}^n d_i^2$, gdzie $d_i = |ax_i + b - y_i|$, jest najmniejsza. Zadanie to ma duże zastosowanie w statystyce.

Celem jest znalezienie minimum funkcji kwadratowej

$$E(a, b) = \sum_{i=1}^n (ax_i + b - y_i)^2.$$

Układem równań na punkt stacjonarny jest

$$\begin{aligned} \frac{\partial E}{\partial a} &= 2 \sum_{i=1}^n (ax_i + b - y_i)x_i = 0, \\ \frac{\partial E}{\partial b} &= 2 \sum_{i=1}^n (ax_i + b - y_i) = 0. \end{aligned}$$

Zatem punkt stacjonarny spełnia układ

$$\begin{cases} a \cdot \sum_{i=1}^n x_i^2 + b \cdot \sum_{i=1}^n x_i &= \sum_{i=1}^n x_i y_i, \\ a \cdot \sum_{i=1}^n x_i + b \cdot n &= \sum_{i=1}^n y_i. \end{cases}$$

Zatem oznaczając $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ dostajemy

$$\begin{aligned} a &= \frac{n \cdot x \circ y - \left(\sum_{i=1}^n x_i\right) \left(\sum_{i=1}^n y_i\right)}{nx^2 - \left(\sum_{i=1}^n x_i\right)^2}, \\ b &= \frac{x^2 \cdot \sum_{i=1}^n y_i - \left(\sum_{i=1}^n x_i\right) \cdot x \circ y}{nx^2 - \left(\sum_{i=1}^n x_i\right)^2}. \end{aligned}$$

E. Regresja kwadratowa

Załóżmy, że na płaszczyźnie \mathbb{R}^2 danych jest n punktów

$$A_1 = (x_1, y_1), \dots, A_n = (x_n, y_n).$$

Zadanie *regresji kwadratowej* polega na znalezieniu paraboli $y = ax^2 + bx + c$ leżącej najbliżej tych punktów w sensie, że wyrażenie

$$E(a, b, c) = \sum_{i=1}^n (ax_i^2 + bx_i + c - y_i)^2$$

przybiera wartość najmniejszą. Układem równań na punkt stacjonarny jest

$$\begin{aligned}\frac{\partial E}{\partial a} &= 2 \sum_{i=1}^n (ax_i^2 + bx_i + c - y_i)x_i^2 = 0, \\ \frac{\partial E}{\partial b} &= 2 \sum_{i=1}^n (ax_i^2 + bx_i + c - y_i)x_i = 0, \\ \frac{\partial E}{\partial c} &= 2 \sum_{i=1}^n (ax_i^2 + bx_i + c - y_i) = 0.\end{aligned}$$

F. Regresja wykładnicza i logarytmiczna

Załóżmy, że na płaszczyźnie \mathbb{R}^2 danych jest n punktów

$$A_1 = (x_1, y_1), \dots, A_n = (x_n, y_n).$$

Zadanie *regresji wykładniczej* polega na znalezieniu funkcji wykładniczej $y = Be^{ax}$ leżącej najbliżej tych punktów w sensie, że wyrażenie

$$F(a, B) = \sum_{i=1}^n (Be^{ax_i} - y_i)^2$$

przybiera wartość najmniejszą. Ponieważ $\ln y = \ln(Be^{ax}) = ax + \ln B$ wystarczy znaleźć minimum funkcji

$$E(a, b) = \sum_{i=1}^n (ax_i + b - \ln y_i)^2, \quad \text{gdzie } b = \ln B.$$

Analogicznie definiuje się zadanie *regresji logarytmicznej*.

G. Programowanie liniowe

Zadanie *programowania liniowego* polega na znalezieniu ekstremów funkcji liniowej w zbiorze opisanym przez funkcje liniowe. Dla danego wektora $b \in \mathbb{R}^n$ oraz liczby $c \in \mathbb{R}$ znaleźć

$$\min_W (b^{\text{tr}}x + c) \quad \text{oraz} \quad \max_W (b^{\text{tr}}x + c),$$

gdzie

$$W = \{x \in \mathbb{R}^n : g_1(x) \leq 0, \dots, g_k(x) \leq 0\},$$

$g_i(x) = v_i^{\text{tr}}x + d_i$, $v_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$ dla $i = 1, \dots, k$.

Ponieważ funkcja $x \rightarrow b^{\text{tr}}x + c$ jest stała na hiperpłaszczyznach $b^{\text{tr}}x = \text{const}$, więc jej ekstrema są przyjmowane w punktach ekstremalnych zbioru W .

6 Zasada Banacha

6.1 Zwartość i zupełność

Definicja 6.1 Niech (X, ρ) będzie przestrzenią metryczną. Podzbiór $K \subset X$ nazywamy zwartym jeśli dowolny ciąg punktów zbioru K zawiera podciąg zbieżny do punktu zbioru K .

Fakt 1. Każdy zbiór zwarty jest domknięty.

Fakt 2. Domknięty podzbiór zbioru zwartego jest zwarty.

Fakt 3. Domknięty i ograniczony zbiór $K \subset \mathbb{R}^n$ jest zwarty.

Definicja 6.2 Niech (X, ρ) będzie przestrzenią metryczną. Mówimy, że ciąg $\{x_n\}_{n=1}^{\infty}$ punktów $x_n \in X$ spełnia warunek Cauchy'ego jeśli zachodzi

$$(C) \quad \text{dla dowolnego } \varepsilon > 0 \text{ istnieje } N \in \mathbb{N} \text{ takie, że} \\ \rho(x_n, x_m) < \varepsilon \quad \text{dla } n, m \geq N.$$

Lemat 6.1 *Każdy ciąg zbieżny w przestrzeni metrycznej (X, ρ) spełnia warunek Cauchy'ego.*

Lemat 6.2 *Każdy ciąg Cauchy'ego w przestrzeni metrycznej (X, ρ) jest ograniczony.*

Definicja 6.3 Niech (X, ρ) będzie przestrzenią metryczną. Mówimy, że przestrzeń (X, ρ) jest zupełna jeśli każdy ciąg Cauchy'ego elementów przestrzeni X jest zbieżny w X .

Przykład 6.1 *Przestrzeń $(0, 1)$ z metryką $\rho(x, y) = |x - y|$ nie jest zupełna. Podobnie \mathbb{Q} , gdzie \mathbb{Q} jest zbiorem liczb wymiernych, nie jest zupełna.*

Lemat 6.3 *Każda przestrzeń metryczna (X, ρ) zwarta jest zupełna.*

Lemat 6.4 *Jeśli w przestrzeni metrycznej (X, ρ) każda kula domknięta jest zwarta, to (X, ρ) jest zupełna.*

Ponieważ w przestrzeni \mathbb{R}^n kule domknięte są zwarte mamy

Wniosek 6.1 *Przestrzeń \mathbb{R}^n z metryką euklidesową jest zupełna*

Wniosek 6.2 *Domknięty podzbiór przestrzeni zupełnej jest przestrzenią zupełną.*

6.2 Zasada Banacha

Definicja 6.4 Niech (X, ρ) będzie przestrzenią metryczną.

Odwzorowanie $T : X \rightarrow X$ nazywamy *związającym* jeśli istnieje stała $0 < \lambda < 1$ taka, że dla dowolnych $x, y \in X$ zachodzi

$$\rho(T(x), T(y)) \leq \lambda \rho(x, y).$$

Łatwo zauważyć, że odwzorowanie związające jest ciągle.

Definicja 6.5 Punkt $\hat{x} \in X$ nazywamy punktem *stałym* odwzorowania $T : X \rightarrow X$ jeśli $T(\hat{x}) = \hat{x}$.

Przykład 6.2 Niech $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ będzie odwzorowaniem liniowym, tzn. $Tx = Ax$ dla pewnej macierzy $M \in M(n \times n)$. Wówczas punkt 0 jest punktem stałym odwzorowania T . T jest odwzorowaniem związającym wtedy i tylko wtedy, gdy moduły wartości własnych macierzy A są mniejsze od 1.

Przedstawimy teraz tzw. zasadę Banacha. Jest to jedno z najważniejszych twierdzeń analizy, wykorzystywane do dowodu istnienia rozwiązań wielu problemów matematycznych.

Twierdzenie 6.1 (Banacha o punkcie stałym.) *Niech T będzie odwzorowaniem związającym przestrzeni metrycznej, zupełnej (X, ρ) w siebie. Wówczas istnieje dokładnie jeden punkt stały odwzorowania T . Ponadto, jest wyznaczony jako granica ciągu $\{T^n x\}_{n=1}^{\infty}$, gdzie x jest dowolnym punktem przestrzeni X .*

Dowód. Weźmy dowolny punkt $x \in X$ i oznaczmy $d = \rho(x, Tx)$. Korzystając z definicji odwzorowania związającego oraz indukcji matematycznej łatwo zauważyć, że dla $n \in \mathbb{N}_0$ zachodzi

$$\rho(T^n x, T^{n+1} x) \leq \lambda^n d,$$

gdzie $0 < \lambda < 1$. Ponieważ szereg $\sum_{n=0}^{\infty} \lambda^n d$ jest zbieżny, więc ciąg $\{T^n x\}_{n=0}^{\infty}$ jest ciągiem Cauchy'ego w X . Wobec zupełności przestrzeni X istnieje granica

$$\lim_{n \rightarrow \infty} T^n x = \hat{x}.$$

Korzystając z ciągłości odwzorowania T dostajemy

$$T\hat{x} = T \lim_{n \rightarrow \infty} T^n x = \lim_{n \rightarrow \infty} T^{n+1} x = \hat{x}.$$

Zatem \hat{x} jest punktem stałym T . Jeśli \hat{y} też jest punktem stałym odwzorowania T , to ponieważ

$$\rho(\hat{x}, \hat{y}) = \rho(T\hat{x}, T\hat{y}) \leq \lambda \rho(\hat{x}, \hat{y})$$

oraz $\lambda < 1$, więc musi zachodzić $\rho(\hat{x}, \hat{y}) = 0$. Stąd $\hat{x} = \hat{y}$. \square

7 Odwracanie odwzorowań

Niech $f : (a, b) \rightarrow \mathbb{R}$ będzie funkcją klasy C^1 , której pochodna f' nigdzie nie znika na odcinku (a, b) . Ponieważ f' z założenia jest funkcją ciągłą, więc na mocy własności Darboux dla każdego $x \in (a, b)$ zachodzi $f'(x) > 0$ lub $f'(x) < 0$. Zatem f jest funkcją ściśle rosnącą lub ściśle malejącą, a więc różnowartościową. Stąd wnioskujemy, że istnieje funkcja odwrotna $f^{-1} : (f(a), f(b)) \rightarrow \mathbb{R}$, tzn. $f^{-1}(f(x)) = x$ dla każdego $x \in (a, b)$.

W przypadku odwzorowań wielowymiarowych $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ sytuacja jest bardziej złożona. Naturalnym odpowiednikiem warunku $f'(x) \neq 0$ jest nieosobliwość różniczki odwzorowania F , czyli warunek

$$JF(x) \neq 0 \quad \text{dla } x \in \Omega,$$

gdzie $JF = \det DF$ jest jakobianem odwzorowania F , czyli wyznacznikiem macierzy różniczki DF . Niestety dla $n \geq 2$ odwzorowanie o nieznikającym jakobianie nie musi być różnowartościowe.

Przykład 7.1 Niech $\Omega = \{(r, \varphi) : r > 0, \varphi \in \mathbb{R}\}$ oraz

$$F(r, \varphi) = (f_1(r, \varphi), f_2(r, \varphi)) := (r \cos \varphi, r \sin \varphi) \quad \text{dla } (r, \varphi) \in \Omega.$$

Wówczas F odwzorowuje Ω na $\mathbb{R}^2 \setminus \{(0, 0)\}$ oraz

$$JF(r, \varphi) = \det \begin{pmatrix} \frac{\partial f_1}{\partial r} & \frac{\partial f_1}{\partial \varphi} \\ \frac{\partial f_2}{\partial r} & \frac{\partial f_2}{\partial \varphi} \end{pmatrix} = \det \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} = r > 0 \quad \text{dla } (r, \varphi) \in \Omega.$$

Odwzorowanie F nie jest różnowartościowe, gdyż $F(r, \varphi + 2\pi) = F(r, \varphi)$. Zauważmy jednak, że dowolny punkt $(r_0, \varphi_0) \in \Omega$ posiada otoczenie otwarte $U = \{(r, \varphi) : r > 0, \varphi_0 - \pi < \varphi < \varphi_0 + \pi\}$, na którym F jest różnowartościowe. Dowolny punkt $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ można przedstawić w postaci $x = r \cos \varphi$, $y = r \sin \varphi$, gdzie $r = \sqrt{x^2 + y^2}$, $\varphi = \arctg \frac{y}{x}$. (r, φ) nazywamy współrzędnymi biegunowymi punktu (x, y) .

Definicja 7.1 Niech $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ będzie odwzorowaniem klasy C^1 . Mówimy, że F jest nieosobliwe w punkcie $\hat{x} \in \Omega$ jeśli $JF(\hat{x}) \neq 0$. Podobnie F jest nieosobliwe na zbiorze $U \subset \Omega$ jeśli jest nieosobliwe w każdym punkcie zbioru U .

Twierdzenie 7.1 O lokalnym odwracaniu odwzorowań.

Niech $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ będzie odwzorowaniem klasy C^1 oraz $\hat{x} \in \Omega$.

Jeśli F jest nieosobliwe w \hat{x} , to

a) $F(\Omega)$ jest otoczeniem punktu $\hat{y} = F(\hat{x})$;

b) Istnieje otoczenie U punktu \hat{x} takie, że $F|_U$ jest różnowartościowe.

Twierdzenie 7.2 O różniczkowalności odwzorowania odwrotnego.

Niech $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ będzie odwzorowaniem klasy $C^1(\Omega)$. Załóżmy, że F jest nieosobliwe na Ω . Wówczas

1. $F(\Omega)$ jest zbiorem otwartym;
2. Jeśli F jest różnowartościowe na Ω , to odwzorowanie odwrotne F^{-1} jest klasy C^1 oraz dla $y = F(x)$, $x \in \Omega$ zachodzi

$$DF^{-1}(y) = (DF(x))^{-1}, \quad JF^{-1}(y) = 1/JF(x).$$

Przykład 7.2 Niech $\Omega = \{(r, \varphi) : r > 0, -\pi < \varphi < \pi\}$ oraz

$$F(r, \varphi) = (f_1(r, \varphi), f_2(r, \varphi)) := (r \cos \varphi, r \sin \varphi) \quad \text{dla } (r, \varphi) \in \Omega.$$

Wówczas F odwzorowuje wzajemnie jednoznacznie Ω na $\mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\}$ oraz

$$JF(r, \varphi) = \det \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} = r > 0 \quad \text{dla } (r, \varphi) \in \Omega.$$

Z Twierdzenia 7.2 wynika, że odwzorowanie odwrotne $G = F^{-1}$ jest klasy C^1 . G można wyrazić w sposób jawny

$$G(x, y) = (g_1(x, y), g_2(x, y)) := (\sqrt{x^2 + y^2}, \operatorname{arctg} \frac{y}{x}).$$

Tym nie mniej różniczkę odwzorowania G można wyliczyć nie korzystając z jawnych wzorów na G . Otóż na podstawie punktu 2 Twierdzenia 7.2 mamy

$$\begin{aligned} DG(x, y) &= (DF)^{-1} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}^{-1} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{\sin \varphi}{r} & \frac{\cos \varphi}{r} \end{pmatrix} \\ &= \begin{pmatrix} \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \\ \frac{-y}{x^2 + y^2} & \frac{x}{x^2 + y^2} \end{pmatrix}. \end{aligned}$$

Stąd

$$dg_1(dx, dy) = \frac{xdx + ydy}{\sqrt{x^2 + y^2}}, \quad dg_2(dx, dy) = \frac{xdy - ydx}{x^2 + y^2}.$$

8 Dyfeomorfizmy

Z tematem odwracania odwzorowań wiąże się pojęcie dyfeomorfizmu.

Definicja 8.1 Odwzorowanie $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ nazywamy *dyfeomorfizmem zbioru Ω na obraz $F(\Omega)$* jeśli spełnione są trzy warunki

1. F jest klasy C^1 ,
2. F jest nieosobliwe i różnowartościowe,
3. odwzorowanie odwrotne $F^{-1} : F(\Omega) \rightarrow \mathbb{R}^n$ jest ciągłe.

Każdy dyfeomorfizm jest homeomorfizmem. Ponadto, z warunku nieosobliwości F wynika, że $m \geq n$. Z twierdzenia o różniczkowaniu złożenia odwzorowań wynika

Wniosek 8.1 *Jeśli $F : \Omega \subset \mathbb{R}^n \rightarrow U \subset \mathbb{R}^m$ jest dyfeomorfizmem Ω na U oraz $G : U \rightarrow \mathbb{R}^k$ jest dyfeomorfizmem U na $V = G(U)$, to $G \circ F$ jest dyfeomorfizmem Ω na V .*

Uwaga. Jeśli $m > n$, to z warunków nieosobliwości i różnowartościowości F nie wynika ciągłość odwzorowania odwrotnego. Pokazuje to poniższy przykład

Przykład 8.1 Niech $F : (-2\pi, 1) \subset \mathbb{R} \rightarrow \mathbb{R}^2$ będzie dane wzorem

$$F(t) = \begin{cases} (\cos t, \sin t) & \text{dla } -2\pi < t < 0, \\ (1, t) & \text{dla } 0 \leq t < 1. \end{cases}$$

Wówczas F jest odwzorowaniem klasy C^1 , jest nieosobliwe i różnowartościowe, lecz odwzorowanie odwrotne nie jest ciągłe.

Tym nie mniej z Twierdzenia 7.2 dostajemy

Wniosek 8.2 *Jeśli odwzorowanie $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ jest klasy C^1 oraz jest nieosobliwe i różnowartościowe, to jest ono dyfeomorfizmem Ω na $F(\Omega)$ oraz $F^{-1} : F(\Omega) \rightarrow \Omega$ jest dyfeomorfizmem.*

Podamy teraz serię ważnych przykładów dyfeomorfizmów.

Przykład 8.2 Niech

$$F(r, \varphi) = (r \cos \varphi, r \sin \varphi)$$

Wówczas F jest dyfeomorfizmem zbioru $\Omega = \{(r, \varphi) : r > 0, -\pi < \varphi < \pi\}$ na $\mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\}$. F nazywamy *dyfeomorfizmem biegunowym*.

Przykład 8.3 Niech

$$F(r, \varphi, z) = (r \cos \varphi, r \sin \varphi, z)$$

Wówczas F jest dyfeomorfizmem zbioru $\Omega = \{(r, \varphi, z) : r > 0, -\pi < \varphi < \pi, z \in \mathbb{R}\}$ na $\mathbb{R}^3 \setminus \{(x, 0, z) : x \leq 0\}$. F nazywamy *dyfeomorfizmem walcowym*.

Przykład 8.4 (Współrzędne sferyczne).

Współrzędnymi sferycznymi punktu $(x, y, z) \in \mathbb{R}^3$ nazywamy liczby r, ψ, φ takie, że

$$x = r \cos \psi \cos \varphi, \quad y = r \cos \psi \sin \varphi, \quad z = r \sin \psi.$$

Niech

$$F(r, \varphi, \psi) = (r \cos \psi \cos \varphi, r \cos \psi \sin \varphi, r \sin \psi)$$

Wówczas F jest dyfeomorfizmem zbioru $\Omega = \{(r, \varphi, \psi) : r > 0, -\pi < \varphi < \pi, -\frac{\pi}{2} < \psi < \frac{\pi}{2}\}$ na obraz.

Zauważmy, że F jest złożeniem dwóch dyfeomorfizmów walcowych $F = G \circ H$, gdzie

$$G(\rho, \varphi, z) = (\rho \cos \varphi, \rho \sin \varphi, z),$$

$$H(r, \varphi, \psi) = (r \cos \psi, \varphi, r \sin \psi).$$

Zatem F jest dyfeomorfizmem Ω na $\mathbb{R}^3 \setminus \{(x, 0, z) : x \leq 0\}$. Policzmy jeszcze macierz różniczki F . Na mocy twierdzenia o różniczce złożenia odwzorowań dostajemy

$$\begin{aligned} DF = DG \circ DH &= \begin{pmatrix} \cos \varphi & -\rho \sin \varphi & 0 \\ \sin \varphi & \rho \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \circ \begin{pmatrix} \cos \psi & 0 & -r \sin \psi \\ 0 & 1 & 0 \\ \sin \psi & 0 & r \cos \psi \end{pmatrix} \\ &\stackrel{\rho=r \cos \psi}{=} \begin{pmatrix} \cos \psi \cos \varphi & -r \cos \psi \sin \varphi & -r \sin \varphi \cos \psi \\ \cos \psi \sin \varphi & r \cos \psi \cos \varphi & -r \sin \psi \sin \varphi \\ \sin \psi & 0 & r \cos \psi \end{pmatrix}. \end{aligned}$$

Ponadto,

$$JF = JG \cdot JH = \rho \cdot r = r^2 \cos \psi.$$

Przykład 8.5 Niech $0 < a_1 < a_2$, $0 < b_1 < b_2$, $\alpha < \beta$ oraz

$$\Omega = \{(x, y) : x > 0, y > 0, a_1 x^\alpha < y < a_2 x^\alpha, b_1 x^\beta < y < b_2 x^\beta\}.$$

Jeśli oznaczymy $u = x^{-\alpha} y$, $v = x^{-\beta} y$, to $(x, y) \in \Omega$ wtedy i tylko wtedy, gdy $(u, v) \in P = \{(u, v) : a_1 < u < a_2, b_1 < v < b_2\}$. Niech

$$F(x, y) = (x^{-\alpha} y, x^{-\beta} y) \quad \text{dla } (x, y) \in \Omega.$$

F jest wzajemnie jednoznacznym odwzorowaniem Ω na P .

$$DF(x, y) = \begin{pmatrix} -\alpha x^{-\alpha-1} y & x^{-\alpha} \\ -\beta x^{-\beta-1} y & x^{-\beta} \end{pmatrix}$$

oraz

$$JF(x, y) = (\beta - \alpha) x^{-\alpha-\beta-1} y > 0 \quad \text{dla } (x, y) \in \Omega.$$

Zatem F jest dyfeomorfizmem Ω na P . Dyfeomorfizm odwrotny jest dany przez

$$F^{-1}(u, v) = \left(\left(\frac{u}{v} \right)^{\beta-\alpha}, \left(\frac{v^\beta}{u^\alpha} \right)^{1/(\beta-\alpha)} \right).$$

Przykład 8.6 Niech $\Omega = \{(x, y) : x \in I, 0 < y < f(x)\}$, gdzie $I = (a, b) \subset \mathbb{R}$ oraz niech $f : I \rightarrow \mathbb{R}_+$ jest funkcją klasy $C^1(I)$ o wartościach dodatnich. Połóżmy

$$\Psi(x, y) = \left(x, \frac{y}{f(x)} \right) \quad \text{dla } (x, y) \in \Omega.$$

Wówczas Ψ jest wzajemnie jednoznaczny odwzorowaniem Ω na $P = \{(u, v) : u \in I, 0 < v < 1\}$. Macierzą różniczki Ψ jest

$$D\Psi = \begin{pmatrix} 1 & 0 \\ \frac{-yf'(x)}{f^2(x)} & \frac{1}{f(x)} \end{pmatrix}$$

oraz

$$J\Psi(x, y) = \frac{1}{f(x)} > 0 \quad \text{dla } (x, y) \in \Omega.$$

Zatem Ψ jest dyfeomorfizmem Ω na P . Dyfeomorfizm odwrotny jest dany przez

$$\Psi^{-1}(u, v) = (u, vf(u)).$$

Przykład 8.7 (Symetria sferyczna).

Niech

$$I(x) = \frac{x}{\|x\|^2} \quad \text{dla } x \in \Omega := \mathbb{R}^n \setminus \{0\}.$$

Wówczas I jest wzajemnie jednoznaczny odwzorowaniem zbioru Ω na siebie oraz $I_{S^{n-1}} = Id$. Ponadto, $I^{-1} = I$ gdyż

$$I(I(x)) = \frac{\frac{x}{\|x\|^2}}{\left\| \frac{x}{\|x\|^2} \right\|^2} = x \quad \text{dla } x \in \Omega.$$

Zatem

$$DI(y) \circ DI(x) = Id \quad \text{dla } y = I(x) \in \Omega.$$

Czyli I jest nieosobliwe i jest to dyfeomorfizm $\mathbb{R}^n \setminus \{0\}$ na siebie.

Zadanie. Policzyc macierz różniczki odwzorowania I .

9 Funkcje uwikłane

Niech $F : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ będzie funkcją klasy C^1 . Rozważmy zbiór

$$S = \{(x, y) \in \Omega : F(x, y) = 0\}.$$

Niech $(\dot{x}, \dot{y}) \in S$. Będziemy chcieli wiedzieć kiedy w otoczeniu punktu (\dot{x}, \dot{y}) zbiór S jest wykresem funkcji $y = g(x)$ lub funkcji $x = h(y)$. Mówimy wówczas, że równanie $F(x, y) = 0$ określa w sposób uwikłany funkcję $y = g(x)$ lub funkcję $x = h(y)$.

Przykład 9.1 Niech $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ oraz niech $(\dot{x}, \dot{y}) \in S$.

Jeśli $-1 < \dot{x} < 1, \dot{y} > 0$, to funkcja $g(x) = \sqrt{1 - x^2}$ określona w otoczeniu punktu \dot{x} spełnia $g(\dot{x}) = \dot{y}$ oraz $x^2 + (g(x))^2 = 1$.

Podobnie, jeśli $-1 < \dot{x} < 1$ lecz $\dot{y} < 0$, to funkcja g dana przez $g(x) = -\sqrt{1 - x^2}$ spełnia $g(\dot{x}) = \dot{y}$ oraz $x^2 + (g(x))^2 = 1$.

Jeśli natomiast $\dot{y} = 0$, to funkcja $h(y) = \sqrt{1 - y^2}$ jeśli $\dot{x} = 1$ (odpowiednio $h(y) = -\sqrt{1 - y^2}$ jeśli $\dot{x} = -1$) zdefiniowana w otoczeniu punktu \dot{y} spełnia $(h(\dot{y}))^2 + \dot{y}^2 = 1$ oraz $h(\dot{y}) = \dot{x}$.

Twierdzenie 9.1 O funkcji uwikłanej.

Niech $F : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ będzie funkcją klasy C^1 ,

$$S = \{(x, y) \in \Omega : F(x, y) = 0\} \quad \text{oraz} \quad (\dot{x}, \dot{y}) \in S.$$

Jeśli

$$\frac{\partial F}{\partial y}(\dot{x}, \dot{y}) \neq 0,$$

to istnieją przedział $I \ni \dot{x}$ i funkcja $g : I \rightarrow \mathbb{R}$ klasy C^1 taka, że $g(\dot{x}) = \dot{y}$ oraz zbiór $S_1 = \{(x, g(x)) : x \in I\}$ jest zawarty w S i otwarty w S . Ponadto,

$$g'(x) = \frac{-\frac{\partial F}{\partial x}(x, g(x))}{\frac{\partial F}{\partial y}(x, g(x))} \quad \text{dla} \quad x \in I_1 \subset I.$$

Analogicznie, jeśli

$$\frac{\partial F}{\partial x}(\dot{x}, \dot{y}) \neq 0,$$

to istnieją przedział $J \ni \dot{y}$ i funkcja $h : J \rightarrow \mathbb{R}$ klasy C^1 taka, że $h(\dot{y}) = \dot{x}$ oraz zbiór $S_2 = \{(h(y), y) : y \in J\}$ jest zawarty w S i otwarty w S . Ponadto,

$$h'(y) = \frac{-\frac{\partial F}{\partial y}(h(y), y)}{\frac{\partial F}{\partial x}(h(y), y)} \quad \text{dla} \quad x \in J_1 \subset J.$$

Uogólnimy teraz twierdzenie o funkcji uwikłanej na wyższe wymiary. Niech $k, l, n \in \mathbb{N}$ oraz $k + l = n$. Współrzędne punktu $x \in \mathbb{R}^n$ podzielimy na dwie grupy $x = (x_1, \dots, x_l, x_{l+1}, \dots, x_n) = (x', y)$, gdzie $x' = (x_1, \dots, x_l)$, $y = (x_{l+1}, \dots, x_n) = (y_1, \dots, y_k)$. Niech $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$ będzie odwzorowaniem klasy $C^1(\Omega; \mathbb{R}^k)$, tzn. $F = (f_1, \dots, f_k)$, gdzie $f_i \in C^1(\Omega; \mathbb{R})$ dla $i = 1, \dots, k$. Jak wiemy różniczka DF odwzorowania F jest odwzorowaniem liniowym z \mathbb{R}^n do \mathbb{R}^k , którego macierzą jest

$$DF = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1} & \frac{\partial f_k}{\partial x_2} & \cdots & \frac{\partial f_k}{\partial x_n} \end{pmatrix}.$$

Wyznacznik macierzy kwadratowej otrzymanej z macierzy różniczki DF poprzez wybór kolumn o numerach p_1, p_2, \dots, p_k będziemy oznaczać

$$J_{p_1, p_2, \dots, p_k} F = \frac{\partial(f_1, f_2, \dots, f_k)}{\partial(x_{p_1} x_{p_2} \dots x_{p_k})} = \det \begin{pmatrix} \frac{\partial f_1}{\partial x_{p_1}} & \frac{\partial f_1}{\partial x_{p_2}} & \cdots & \frac{\partial f_1}{\partial x_{p_k}} \\ \frac{\partial f_2}{\partial x_{p_1}} & \frac{\partial f_2}{\partial x_{p_2}} & \cdots & \frac{\partial f_2}{\partial x_{p_k}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_{p_1}} & \frac{\partial f_k}{\partial x_{p_2}} & \cdots & \frac{\partial f_k}{\partial x_{p_k}} \end{pmatrix}$$

i nazywać jakobianem cząstkowym F .

Rozpatrzmy zbiór

$$\begin{aligned} S &= \{(x', y) \in \Omega : F(x', y) = 0\} \\ &= \{(x', y) \in \Omega : f_i(x', y) = 0 \text{ dla } i = 1, \dots, k\}. \end{aligned}$$

Twierdzenie 9.2 O funkcji uwikłanej, przypadek ogólny. *Przy powyższych założeniach jeśli $\hat{x} = (\hat{x}', \hat{y}) \in S$ oraz*

$$J_{l+1 \dots n} F(\hat{x}) \neq 0$$

to istnieje odwzorowanie $G = (g_1, \dots, g_k)$ klasy C^1 określone w pewnym obszarze $\Delta \subset \mathbb{R}^l$ o wartościach w \mathbb{R}^k takie, że $G(\hat{x}') = \hat{y}$, tzn.

$$\begin{aligned} g_1(\hat{x}_1, \dots, \hat{x}_l) &= \hat{x}_{l+1} = \hat{y}_1, \\ &\vdots \\ g_k(\hat{x}_1, \dots, \hat{x}_l) &= \hat{x}_{l+k} = \hat{y}_k \end{aligned}$$

dla $\hat{x}' \in \Delta$ oraz zbiór $S_1 = \{(x', g_1(x'), \dots, g_k(x')) : x' \in \Delta\}$ jest zawarty w S i otwarty w S . Ponadto,

$$DG(x') = -(D_{x'} F)^{-1}(x', G(x')) \circ (D_y F)(x', g(x')) \quad \text{dla } x' \in \Delta' \subset \Delta.$$

Przykład 9.2 Niech $F = (f_1, f_2) : \mathbb{R}^5 \rightarrow \mathbb{R}^2$, gdzie

$$\begin{aligned} f_1(x_1, x_2, y_1, y_2, y_3) &= e^{x_1} - x_1 y_2 + x_2 y_1 + y_3 - 7 = 0, \\ f_2(x_1, x_2, y_1, y_2, y_3) &= x_2^2 \cos x_1 - x_1 - x_1 y_3 + x_2 y_1 - 3 = 0 \end{aligned}$$

oraz

$$S = \{(x_1, x_2, y_1, y_2, y_3) \in \mathbb{R}^5 : F(x_1, x_2, y_1, y_2, y_3) = 0\}, \quad \dot{x} = (0, 1), \quad \dot{y} = (2, 3, 4).$$

Wówczas $(\dot{x}, \dot{y}) \in S$. Macierzą różniczki odwzorowania F jest

$$DF(x, y) = \begin{pmatrix} e^{x_1} - y_2 & y_1 & x_2 & -x_1 & 1 \\ -x_2^2 \sin x_1 - 1 - y_3 & 2x_2 \cos x_1 + y_1 & x_2 & 0 & -x_1 \end{pmatrix}.$$

Zatem

$$DF(\dot{x}, \dot{y}) = \begin{pmatrix} -2 & 2 & 1 & 0 & 1 \\ -5 & 4 & 1 & 0 & 0 \end{pmatrix}$$

oraz

$$J_{x_1 x_2} F(\dot{x}, \dot{y}) = \det \begin{pmatrix} -2 & 2 \\ -5 & 4 \end{pmatrix} = 2 \neq 0.$$

Na podstawie Twierdzenia 9.2 wnioskujemy, że w otoczeniu Δ punktu \dot{y} istnieje odwzorowanie $G = (g_1, g_2)$ klasy C^1 takie, że $F(g_1(y), g_2(y), y) \equiv 0$. Macierz różniczki odwzorowania G można wyznaczyć ze wzoru

$$DF = (D_x F) \circ DG + (D_y F) = 0.$$

Stąd $DG = -(D_x F)^{-1} \circ (D_y F)$. Zatem

$$\begin{aligned} DG(\dot{y}) &= - \begin{pmatrix} -2 & 2 \\ -5 & 4 \end{pmatrix}^{-1} \circ \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \\ &= - \begin{pmatrix} 2 & -1 \\ \frac{5}{2} & -1 \end{pmatrix} \circ \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & -2 \\ -\frac{3}{2} & 0 & -\frac{5}{2} \end{pmatrix}. \end{aligned}$$

Stąd

$$\begin{aligned} \frac{\partial g_1}{\partial y_1}(\dot{y}) &= -1, & \frac{\partial g_1}{\partial y_2}(\dot{y}) &= 0, & \frac{\partial g_1}{\partial y_3}(\dot{y}) &= -2, \\ \frac{\partial g_2}{\partial y_1}(\dot{y}) &= -\frac{3}{2}, & \frac{\partial g_2}{\partial y_2}(\dot{y}) &= 0, & \frac{\partial g_2}{\partial y_3}(\dot{y}) &= -\frac{5}{2}. \end{aligned}$$

10 Ekstrema warunkowe

Często zachodzi potrzeba znalezienia ekstremów funkcji rzeczywistej f na zbiorze opisanym przez ograniczenia równościowe

$$S = \{x \in \Omega \subset \mathbb{R}^n : g_1(x) = 0, \dots, g_k(x) = 0\}.$$

Mówimy wówczas, że mamy do czynienia z problemem znalezienia ekstremów warunkowych. Na przykład chcemy znaleźć minimum i maksimum funkcji liniowej

$$f(x) = a_1x_1 + \dots + a_nx_n$$

na $(n - 1)$ -wymiarowej sferze

$$S^{n-1} = \{x \in \mathbb{R}^n : g(x) = x_1^2 + \dots + x_n^2 - 1 = 0\}.$$

Definicja 10.1 Niech g_1, \dots, g_k będą funkcjami rzeczywistymi klasy $C^1(\Omega)$ oraz

$$S = \{x \in \Omega \subset \mathbb{R}^n : g_1(x) = 0, \dots, g_k(x) = 0\}.$$

Punkt $\hat{x} \in S$ nazywamy *punktem regularnym* zbioru S , jeśli gradienty $\nabla g_1(\hat{x}), \dots, \nabla g_k(\hat{x})$ są liniowo niezależnymi wektorami w \mathbb{R}^n (zatem $k \leq n$).

Przykład 10.1 Niech $S^{n-1} = \{x \in \mathbb{R}^n : g(x) = x_1^2 + \dots + x_n^2 - 1 = 0\}$ będzie sferą. Wówczas $\nabla g(x) = (2x_1, \dots, 2x_n) \neq 0$ dla $x \in S^{n-1}$. Zatem każdy punkt sfery S^{n-1} jest punktem regularnym.

Przykład 10.2 Niech $S = \{(x, y) \in \mathbb{R}^2 : g(x, y) = x^2 - y^3 = 0\}$. Wówczas $\nabla g(x, y) = (2x, 3y^2) \neq 0$ dla $(x, y) \in S \setminus \{(0, 0)\}$. Zatem $(0, 0)$ nie jest punktem regularnym zbioru S , a pozostałe punkty są regularne.

Definicja 10.2 Jeśli \hat{x} jest punktem regularnym zbioru S , to *przestrzenią styczną do S w punkcie \hat{x}* nazywamy zbiór

$$T_{\hat{x}}S = \{h \in \mathbb{R}^n : \nabla g_i(\hat{x})h = 0 \quad \text{dla } i = 1, \dots, k\}.$$

Czasami problem szukania ekstremów warunkowych można sprowadzić do szukania ekstremów funkcji na zbiorze otwartym. Na przykład jeśli zmienne (x_1, \dots, x_n) da się rozdzielić na dwie grupy $(x_1, \dots, x_l) = x'$ i $(x_{l+1}, \dots, x_n) = y$, $1 \leq l \leq n - 1$, w ten sposób, że układ równań

$$g_1(x', y) = 0, \dots, g_k(x', y) = 0$$

można rozwiązać globalnie, tzn. istnieją funkcje $h_j : U \subset \mathbb{R}^l \rightarrow \mathbb{R}$ dla $j = 1, \dots, k$, $k + l = n$ takie, że $g_i(x', h_1(x'), \dots, h_k(x')) = 0$, to problem znalezienia minimum

$$\min_{x \in \Omega} \{f(x) : g_1(x) = 0, \dots, g_k(x) = 0\}$$

można sprowadzić do problemu znalezienia minimum na zbiorze otwartym U

$$\min_{x' \in U} \{f(x', h_1(x'), \dots, h_k(x'))\}.$$

Przykład 10.3 Problem

$$\min_{(x_1, x_2) \in \mathbb{R}^2} \{x_1^2 + x_2^2 : x_1 + x_2 = 1\}$$

jest równoważny problemowi

$$\min_{x_1 \in \mathbb{R}} \{x_1^2 + (1 - x_1)^2\}.$$

Twierdzenie 10.1 (Warunek konieczny I rzędu ekstremum warunkowego).

Niech $f, g_1, \dots, g_k : \Omega \rightarrow \mathbb{R}$ będą funkcjami klasy C^1 na zbiorze otwartym $\Omega \subset \mathbb{R}^n$,

$$S = \{x \in \Omega : g_1(x) = 0, \dots, g_k(x) = 0\}.$$

Jeśli funkcja f ma w punkcie regularnym \hat{x} zbioru S ekstremum warunkowe, to istnieją liczby $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ takie, że

$$\nabla f(\hat{x}) + \sum_{i=1}^k \lambda_i \nabla g_i(\hat{x}) = 0. \quad (9)$$

Definicja 10.3 Liczby $\lambda_1, \dots, \lambda_k$ nazywamy współczynnikami Lagrange'a, a funkcję

$$L(x, \lambda) = f(x) + \sum_{i=1}^k \lambda_i \nabla g_i(x)$$

funkcją Lagrange'a.

Uwaga. Jeśli \hat{x} jest punktem regularnym zbioru S , to warunek (9) oznacza, że $\nabla_x L = 0$. Jednocześnie, mamy $\nabla_\lambda L = (g_1, \dots, g_k)$. Zatem jeśli funkcja f osiąga w punkcie regularnym zbioru S ekstremum lokalne, to dla pewnego $\lambda \in \mathbb{R}^k$ spełniony jest układ $(n + k)$ równań

$$\nabla L(\hat{x}, \lambda) = 0.$$

Przykład 10.4 Niech

$$f(x) = a_1 x_1 + \dots + a_n x_n, \quad \text{gdzie } (a_1, \dots, a_n) \neq (0, \dots, 0),$$
$$S^{n-1} = \{g(x) = 0\}, \quad \text{gdzie } g(x) = x_1^2 + \dots + x_n^2 - 1.$$

Wówczas

$$\nabla f(x) = (a_1, \dots, a_n), \quad \nabla g(x) = (2x_1, \dots, 2x_n) \neq 0 \quad \text{na } S.$$

Zatem każdy punkt sfery S^{n-1} jest regularny. Jeśli f ma ekstremum warunkowe w punkcie $\hat{x} \in S^{n-1}$, to istnieje $\lambda \in \mathbb{R}$ taka, że

$$(a_1, \dots, a_n) + 2\lambda(\hat{x}_1, \dots, \hat{x}_n) = 0.$$

Stąd

$$(\hat{x}_1, \dots, \hat{x}_n) = -\frac{1}{2\lambda}(a_1, \dots, a_n).$$

Lecz $\hat{x} \in S^{n-1}$, czyli $\hat{x}_1^2 + \dots + \hat{x}_n^2 = 1$. Zatem

$$\lambda = \pm \frac{1}{2\sqrt{a_1^2 + \dots + a_n^2}} \quad \text{oraz} \quad (\hat{x}_1, \dots, \hat{x}_n) = \pm \frac{(a_1, \dots, a_n)}{\sqrt{a_1^2 + \dots + a_n^2}}.$$

Minimum (odpowiednie, maksimum) funkcji f na sferze S^{n-1} wynosi $-\sqrt{a_1^2 + \dots + a_n^2}$ (odpowiednio, $\sqrt{a_1^2 + \dots + a_n^2}$).

Niech $A \in M(n \times n)$ będzie macierzą symetryczną oraz $H(h) = h^{\text{tr}}Ah$ formą kwadratową. W dowodzie kryterium Sylwestera skorzystaliśmy z faktu, że forma H jest dodatnio określona wtedy i tylko wtedy, gdy wszystkie wartości własne macierzy A są dodatnie. Obecnie możemy ten fakt udowodnić.

Przykład 10.5 Niech $A \in M(n \times n)$ będzie macierzą symetryczną oraz

$$H(x) = x^{\text{tr}}Ax = \sum_{i,j=1}^n a_{ij}x^i x^j \quad \text{dla } x \in \mathbb{R}^n.$$

Wyznamy ekstrema funkcji H na sferze $S^{n-1} = \{x \in \mathbb{R}^n : g(x) = 0\}$, gdzie $g(x) = x^2 - 1 = x_1^2 + \dots + x_n^2 - 1$. Ponieważ $\nabla g(x) = 2x \neq 0$ na S^{n-1} , więc każdy punkt sfery jest regularny. Połóżmy

$$L(x, \lambda) = x^{\text{tr}}Ax - \lambda(x^2 - 1).$$

Jeśli w punkcie $\hat{x} \in S^{n-1}$ funkcja H osiąga ekstremum warunkowe, to $\nabla L(\hat{x}, \lambda) = 0$ dla pewnego $\lambda \in \mathbb{R}$. Mamy

$$\begin{cases} \frac{\partial L}{\partial x_i} = 2 \sum_{j=1}^n a_{ij}x_j - 2\lambda x_i & \text{dla } i = 1, \dots, n, \\ \frac{\partial L}{\partial \lambda} = -g(x). \end{cases}$$

Zatem $(A - \lambda Id)\hat{x} = 0$. Aby ten układ miał rozwiązanie $\hat{x} \in S^{n-1}$, λ musi być wartością własną macierzy A . Ponieważ macierz A jest symetryczna jej wartości własne $\lambda_1, \dots, \lambda_n$ są rzeczywiste (tzn. $\text{spectrum} A = \{\lambda_1, \dots, \lambda_n\}$). Jeśli $Ax = \lambda_i x$ oraz $x^2 = 1$, to

$$H(x) = x^{\text{tr}}Ax = x^{\text{tr}}\lambda_i x = \lambda_i x^{\text{tr}}x = \lambda_i x^2 = \lambda_i.$$

Zatem wartości ekstremalne funkcji H na sferze S^{n-1} należą do zbioru $\{\lambda_1, \dots, \lambda_n\}$. Ponieważ dla $r > 0$ zachodzi

$$H\left(\frac{x}{r}\right) = \left(\frac{x}{r}\right)^{\text{tr}}A\frac{x}{r} = \frac{1}{r^2}x^{\text{tr}}Ax = \frac{1}{r^2}H(x).$$

wniosujemy stąd, że H jest dodatnio określona wtedy i tylko wtedy, gdy $\lambda_i > 0$ dla $i = 1, \dots, n$.

Twierdzenie 10.2 (Warunek dostateczny II rzędu minimum warunkowego).

Niech $f, g_1, \dots, g_k : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ będą funkcjami klasy C^2 ,

$$S = \{x \in \Omega \subset \mathbb{R}^n : g_1(x) = 0, \dots, g_k(x) = 0\}$$

oraz niech $\hat{x} \in S$ będzie punktem regularnym. Połóżmy

$$L(x, \lambda) = f(x) + \sum_{i=1}^k \lambda_i \nabla g_i(x)$$

Jeśli istnieje $\lambda = (\lambda_1, \dots, \lambda_k) \in \mathbb{R}$ taka, że

$$\nabla_x L(\hat{x}, \lambda) = 0 \tag{10}$$

oraz

$$d_x^2 L(\hat{x}, \lambda)(h, h) > 0 \tag{11}$$

dla każdego $\mathbb{R}^n \ni h \neq 0$ spełniającego $d(g_1, \dots, g_k)(\hat{x})h = 0$, to funkcja f ma w \hat{x} zbioru S ściśle minimum lokalne warunkowe.

Przykład 10.6 Znajdziemy ekstrema funkcji $f(x, y, z) = x + y + z$ na powierzchni $S = \{\frac{1}{x} + \frac{1}{y} + \frac{1}{z} = 1\}$. Konstruujemy funkcję Lagrange'a

$$L(x, y, z, \lambda) = x + y + z + \lambda \left(\frac{1}{x} + \frac{1}{y} + \frac{1}{z} - 1 \right),$$

a następnie liczymy jej pochodne cząstkowe pierwszego rzędu i przyrównujemy je do zera

$$\begin{aligned} \frac{\partial L}{\partial x} &= 1 - \frac{\lambda}{x^2} = 0, & \frac{\partial L}{\partial y} &= 1 - \frac{\lambda}{y^2} = 0, & \frac{\partial L}{\partial z} &= 1 - \frac{\lambda}{z^2} = 0, \\ \frac{\partial L}{\partial \lambda} &= \frac{1}{x} + \frac{1}{y} + \frac{1}{z} - 1 = 0. \end{aligned}$$

Aby ten układ miał rozwiązanie λ musi być dodatnie. Wówczas $x = \pm\sqrt{\lambda}$, $y = \pm\sqrt{\lambda}$, $z = \pm\sqrt{\lambda}$. Jeśli $x = y = z = \sqrt{\lambda}$, to $\lambda = 9$, a więc $x = y = z = 3$. Mamy zatem punkt stacjonarny $P_1 = (3, 3, 3)$. Jeśli $x = -\sqrt{\lambda}$, to $y = \sqrt{\lambda}$, $z = \sqrt{\lambda}$. Zatem $\lambda = 1$, $x = -1$, $y = z = 1$. Mamy zatem punkt stacjonarny $P_2 = (-1, 1, 1)$. Analogicznie dostajemy punkty stacjonarne $P_3 = (1, -1, 1)$ i $P_4 = (1, 1, -1)$. Liczymy teraz drugą różniczkę względem zmiennych (x, y, z) ,

$$d^2 L(h, h) = (h_1 \ h_2 \ h_3) \begin{pmatrix} \frac{2\lambda}{x^3} & 0 & 0 \\ 0 & \frac{2\lambda}{y^3} & 0 \\ 0 & 0 & \frac{2\lambda}{z^3} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = 2\lambda \left(\frac{h_1^2}{x^3} + \frac{h_2^2}{y^3} + \frac{h_3^2}{z^3} \right).$$

W punkcie P_1 mamy $d^2 L(P_1)(h, h) = \frac{18}{27}(h_1^2 + h_2^2 + h_3^2) > 0$. Zatem w P_1 jest funkcja f przyjmuje minimum warunkowe na S oraz $f(3, 3, 3) = 9$.

W punkcie P_2 mamy $d^2 L(P_2)(h, h) = 2(-h_1^2 + h_2^2 + h_3^2)$. Z warunku $dg(h) = 0$ w punkcie P_2 dostajemy $h_1 + h_2 + h_3 = 0$. Zatem $h_3 = -h_1 - h_2$, a więc $d^2 L(P_2)(h, h) = 2(-h_1^2 + h_2^2 + (h_1 + h_2)^2) = 4h_2(h_1 + h_2)$ przyjmuje zarówno wartości dodatnie jak i ujemne. Czyli w P_2 funkcja f ma punkt siodłowy na S . Podobnie w P_3 i P_4 funkcja f ma punkt siodłowy na S .

11 Rozmaitości

Definicja 11.1 Niech $n \leq m$. Zbiór $S \subset \mathbb{R}^m$ nazywamy n -wymiarowym płatem jeśli istnieje dyfeomorfizm $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ zbioru otwartego Ω na S , tzn. $F(\Omega) = S$. Dyfeomorfizm F nazywa się *parametryzacją* płata S . Płat 1-wymiarowy nazywamy *łukiem otwartym*, a płat 2-wymiarowy *powierzchnią otwartą*.

Przykład 11.1 Zbiór otwarty $\Omega \subset \mathbb{R}^n$ jest płatem n -wymiarowym. W tym przypadku dyfeomorfizm $F : \Omega \rightarrow \mathbb{R}^n$ jest dany przez $F(x) = x$ dla $x \in \Omega$. Wówczas $DF = \text{Id}$. Odwrotnie każdy płat n -wymiarowy w \mathbb{R}^n jest zbiorem otwartym.

Przykład 11.2 Niech $g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ będzie funkcją klasy C^1 . Wówczas wykres funkcji g czyli zbiór

$$S = \{(x, g(x)) : x \in \Omega\} \subset \mathbb{R}^{n+1}$$

jest płatem n -wymiarowym w \mathbb{R}^{n+1} . Istotnie jeśli położymy

$$F(x) = (x, g(x)) \quad \text{dla } x \in \Omega,$$

to F jest dyfeomorfizmem Ω na S , gdyż $DF = (\text{Id}, \nabla g)^{\text{tr}}$. Odwzorowanie odwrotne jest dane przez $F^{-1}(x, y) = x$ dla $x \in \Omega$, $y \in \mathbb{R}$ takich, że $y = g(x)$.

Analogicznie wykres odwzorowania $G : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$ czyli zbiór

$$S = \{(x, G(x)) : x \in \Omega\} \subset \mathbb{R}^{n+k}$$

jest płatem n -wymiarowym w \mathbb{R}^{n+k} .

Przykład 11.3 Niech

$$S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \setminus \{(-1, 0)\}.$$

Wówczas $S = F(\Omega)$, gdzie $\Omega = (-\pi, \pi)$, $F(\varphi) = (\cos \varphi, \sin \varphi)$ dla $\varphi \in \Omega$. Zatem S jest łukiem otwartym. Ciągłość odwzorowania odwrotnego F^{-1} wynika z Lematu 11.1. Zauważmy, że okrąg $S^1 = \{x^2 + y^2 = 1\}$ nie jest łukiem otwartym, gdyż jest on zbiorem zwartym i wobec tego nie może być obrazem zbioru otwartego $\Omega \subset \mathbb{R}$.

Lemat 11.1 Niech T, X będą przestrzeniami metrycznymi przy czym T jest zwarta oraz niech Ω będzie otwartym podzbiorem T . Załóżmy, że odwzorowanie $F : \Omega \rightarrow X$ rozszerza się do odwzorowania ciągłego $\tilde{F} : T \rightarrow X$ spełniającego warunek:

jeśli $\tilde{F}(t) = \tilde{F}(t')$ dla $t, t' \in T$, $t \neq t'$, to $t, t' \notin \Omega$.

Wówczas F jest homeomorfizmem Ω na $F(\Omega)$.

Przykład 11.4 Niech

$$\Omega = \{(\varphi, \psi) : -\pi < \varphi < \pi, -\frac{\pi}{2} < \psi < \frac{\pi}{2}\}$$

oraz $F : \Omega \rightarrow \mathbb{R}^3$,

$$F(\varphi, \psi) = (\cos \psi \cos \varphi, \cos \psi \sin \varphi, \sin \psi).$$

Wówczas $F(\Omega) = S^2 \setminus \{(x, 0, z) : x \leq 0, z \in \mathbb{R}\}$ jest płatem 2-wymiarowym, a F jest jego przedstawieniem parametrycznym. Ciągłość odwzorowania odwrotnego F^{-1} wynika z Lematem 11.1, w którym $T = \bar{\Omega}$, $X = \mathbb{R}^3$, \tilde{F} - przedłużenie F .

Zadanie. Wykazać, że powierzchnia walca

$$S = \{(x, y, z) : x^2 + y^2 = 1, 0 < z < 1\}$$

jest 2-wymiarowym płatem.

Zauważmy, że jeśli $F : \Omega \rightarrow \mathbb{R}^m$ jest przedstawieniem parametrycznym n -wymiarowego płata S oraz $G : \Omega_1 \rightarrow \Omega$ jest dyfeomorfizmem zbioru otwartego $\Omega_1 \subset \mathbb{R}^n$ na Ω , to złożenie $H = F \circ G$ jest też przedstawieniem parametrycznym płata S .

Zachodzi również twierdzenie odwrotne.

Twierdzenie 11.1 Niech $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ oraz $H : \Omega_1 \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ będą parametryzacjami płata n -wymiarowego $S \subset \mathbb{R}^m$. Wówczas złożenie $G = F^{-1} \circ H : \Omega_1 \rightarrow \Omega$ jest dyfeomorfizmem Ω_1 na Ω .

Pojęcie płata nie obejmuje ważnych zbiorów, które naturalnie pojawiają się w analizie. Przykładowo okrąg S^1 i sfera S^2 nie są płatami. Dlatego wprowadza się ogólniejsze pojęcie rozmaitości.

Definicja 11.2 Niech $n \leq m$. Zbiór $S \subset \mathbb{R}^m$ nazywamy n -wymiarową rozmaitością (klasy C^k), jeśli jest ona sumą pewnej rodziny n -wymiarowych płatów $S_j, j \in J$, które są zbiorami otwartymi względem S , tzn.

$$S = \bigcup_{j \in J} S_j,$$

gdzie $S_j = F(\Omega_j)$, F_j jest dyfeomorfizmem (klasy C^k) zbioru otwartego $\Omega_j \subset \mathbb{R}^n$ na S_j .

Mapą rozmaitości nazywamy parę (F, Ω) , gdzie $F : \Omega \rightarrow \mathbb{R}^m$ jest dyfeomorfizmem zbioru otwartego $\Omega \subset \mathbb{R}^n$ na zbiór $F(\Omega)$ zawarty w S i otwarty względem S .

Rodzinę map pokrywających rozmaitość S nazywamy atlasem rozmaitości.

Uwagi.

1. Dowolna rozmaitość $S \subset \mathbb{R}^n$ posiada atlas przeliczalny.
2. Jeśli rozmaitość S jest zwarta, to posiada atlas skończony.
3. Płat n -wymiarowy jest rozmaitością posiadającą atlas złożony z jednej mapy.

Przykład 11.5 Niech $S^1 = \{(x, y) : x^2 + y^2 = 1\}$ będzie okręgiem. Połóżmy

$$F_1(\varphi) = (\cos \varphi, \sin \varphi) \quad \text{dla } \varphi \in \Omega_1 = (-\pi, \pi) \subset \mathbb{R},$$

$$F_2(\varphi) = (\cos \varphi, \sin \varphi) \quad \text{dla } \varphi \in \Omega_2 = (0, 2\pi) \subset \mathbb{R}.$$

Wówczas F_1 i F_2 są dyfeomorfizmami zbiorów Ω_1 i Ω_2 odpowiednio na $S_1 = S^1 \setminus \{(-1, 0)\}$ i $S_2 = S^1 \setminus \{(1, 0)\}$. Zbiory S_1 i S_2 są otwarte względem S^1 oraz $S^1 = S_1 \cup S_2$. Zatem okrąg S^1 jest 1-wymiarową rozmaitością posiadającą atlas złożony z dwóch map.

Przykład 11.6 Niech $S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$ będzie sferą dwuwymiarową. Połóżmy

$$\Phi_{\pm}(u, v) = \left(\frac{2u}{1 + u^2 + v^2}, \frac{2v}{1 + u^2 + v^2}, \pm \frac{u^2 + v^2 - 1}{1 + u^2 + v^2} \right) \quad \text{dla } (u, v) \in \mathbb{R}^2.$$

Wówczas Φ_{\pm} jest dyfeomorfizmem \mathbb{R}^2 na $S_{\pm} = S^2 \setminus \{(0, 0, \pm 1)\}$. Zbiory S_+ i S_- są otwarte względem S^2 oraz $S^2 = S_+ \cup S_-$. Zatem okrąg S^2 jest 2-wymiarową rozmaitością posiadającą atlas złożony z dwóch map.

Definicja 11.3 Niech $S \subset \mathbb{R}^m$ będzie dowolnym niepustym zbiorem oraz $\hat{x} \in S$. Wektor $v \in \mathbb{R}^m$ nazywamy wektorem *stycznym do S w punkcie \hat{x}* jeśli istnieją ciąg punktów $x^i \in S$ zbieżny do \hat{x} oraz ciąg liczb dodatnich $a_i \in \mathbb{R}_+$ takie, że

$$v = \lim_{i \rightarrow \infty} a_i(x^i - \hat{x}).$$

Zbiór wszystkich wektorów stycznych do S w punkcie \hat{x} nazywamy *przestrzenią styczną do S w \hat{x}* i oznaczamy $T_{\hat{x}}S$.

Ponadto oznaczamy $TS = \bigcup_{x \in S} T_xS$ i nazywamy *przestrzenią styczną do S* .

Zauważmy, że wektor zerowy zawsze należy do $T_{\hat{x}}S$. Ponadto, jeśli $v \in T_{\hat{x}}S$ i $\lambda > 0$ to $\lambda v \in T_{\hat{x}}S$.

Jeśli S jest zbiorem otwartym w \mathbb{R}^m i $\hat{x} \in S$, to $T_{\hat{x}}S = \mathbb{R}^m$.

Przykład 11.7 Niech

$$S = \{(x, y) : y^2 = x^3\}, \quad (\hat{x}, \hat{y}) = (0, 0).$$

Wówczas $T_{(\hat{x}, \hat{y})}S = \overline{\mathbb{R}_+}$.

Uwaga. Jeśli S jest rozmaitością, to $T_{\hat{x}}S$ nazywamy *przestrzenią styczną*, a $\hat{x} + T_{\hat{x}}S$ *hiperpłaszczyzną styczną do S w \hat{x}* .

Twierdzenie 11.2 Niech $S \subset \mathbb{R}^m$ będzie n -wymiarową rozmaitością oraz $\hat{x} \in S$. Wówczas $T_{\hat{x}}S$ jest n -wymiarową podprzestrzenią liniową w \mathbb{R}^m , przy czym jeśli (F, Ω) jest mapą obejmującą punkt \hat{x} oraz $\hat{t} = F^{-1}(\hat{x})$, to

$$T_{\hat{x}}S = DF(\hat{t})(\mathbb{R}^n) = \{DF(\hat{t})h, h \in \mathbb{R}^n\}.$$

Twierdzenie 11.3 Niech $l + n = m$, $G \subset \mathbb{R}^m$ będzie zbiorem otwartym, $F : G \rightarrow \mathbb{R}^l$ odwzorowaniem klasy C^1 oraz $S = \{x \in G : F(x) = 0\}$ niepustym zbiorem.

Jeśli rząd $DF(x) = l$ dla $x \in S$, to

1°. S jest n -wymiarową rozmaitością;

2°. $T_{\hat{x}}S = \{v \in \mathbb{R}^m : DF(\hat{x})v = 0\} = \ker DF(\hat{x})$ dla $\hat{x} \in S$.

Zatem hiperpłaszczyzną styczną do S w \hat{x} jest zbiór

$$\{x \in \mathbb{R}^m : DF(\hat{x})(x - \hat{x}) = 0\}.$$

Przykład 11.8 Niech $g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ będzie klasy C^1 oraz niech $S = \{(x, g(x)) : x \in \Omega\} \subset \mathbb{R}^{n+1}$ będzie wykresem g . Wówczas

$$F(x) = (x, g(x)) \quad \text{dla } x \in \Omega,$$

jest F jest dyfeomorfizmem Ω na S . Jeśli $\hat{x} \in \Omega$, $\hat{y} = f(\hat{x})$, to $(\hat{x}, \hat{y}) \in S$ i przestrzenią styczną do S w (\hat{x}, \hat{y}) jest

$$T_{(\hat{x}, \hat{y})}S = \{(h, \nabla g(\hat{x}) \cdot h) : h \in \mathbb{R}^n\},$$

natomiast hiperpłaszczyzną styczną do S w (\hat{x}, \hat{y}) jest

$$\{(\hat{x} + h, \hat{y} + \nabla g(\hat{x})(\hat{x}) \cdot h) : h \in \mathbb{R}^n\} = \{(x, \hat{y} + \sum_{i=1}^n \frac{\partial g}{\partial x_i}(\hat{x})(x_i - \hat{x}_i)) : x \in \mathbb{R}^n\}.$$

Przykład 11.9 Niech

$$F(x) = x_1^2 + \dots + x_{n+1}^2 - 1 \quad \text{dla } x \in \mathbb{R}^{n+1}$$

oraz

$$S^n = \{x \in \mathbb{R}^{n+1} : F(x) = 0\}$$

będzie sferą n -wymiarową. Wówczas $DF(x) = \nabla F(x) = 2x$ dla $x \in S$. Spełnione są założenia Twierdzenia 11.3 przy $l = 1$, $m = n + 1$ i $G = \mathbb{R}^m$. Zatem S^n jest n -wymiarową rozmaitością oraz dla $\hat{x} \in S^n$,

$$T_{\hat{x}}S = \{x \in \mathbb{R}^{n+1} : \hat{x} \cdot x = 0\}.$$

Hiperpłaszczyzną styczną do S w \hat{x} jest

$$\{x \in \mathbb{R}^{n+1} : \hat{x} \cdot (x - \hat{x}) = 0\} = \{x \in \mathbb{R}^{n+1} : \hat{x} \cdot x = 1\}.$$

Przykład 11.10 Niech $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ oraz $S \subset \mathbb{R}^3$ będą zadane przez

$$F(x, y, z) = (x + y + z - 1, x^2 + y^2 + z^2 - 1),$$

$$S = \{(x, y, z) \in \mathbb{R}^3 : F(x, y, z) = 0\}.$$

Macierzą różniczki F jest

$$DF(x, y, z) = \begin{pmatrix} 1 & 1 & 1 \\ 2x & 2y & 2z \end{pmatrix}.$$

Ponieważ rząd $DF(x, y, z) = 2$ dla $(x, y, z) \in S$, S jest $3 - 2 = 1$ -wymiarową rozmaitością. Jeśli $p = (a, b, c) \in S$, to przestrzeń styczna jest prostą

$$T_pS = \{(x, y, z) : x + y + z = 0, ax + by + cz = 0\},$$

natomiast prosta styczna do S w p jest wyznaczona przez $\{(x, y, z) : x + y + z = 1, ax + by + cz = 1\}$.

12 Całka oznaczona Riemanna.

12.1 Definicja całki oznaczonej Riemanna.

Niech $I = [a, b] \subset \mathbb{R}$ będzie zwartym przedziałem oraz $f : I \rightarrow \mathbb{R}$ funkcją ograniczoną. Oznaczmy odpowiednio przez m i M kresy dolny i górny funkcji f na przedziale I , tzn.

$$m = \inf_{x \in I} f(x), \quad M = \sup_{x \in I} f(x).$$

W przedziale I wybierzmy rosnący ciąg punktów $\{x_i\}_{i=0}^k$ taki, że $a = x_0 < x_1 < \dots < x_k = b$. Oznaczmy $\Delta_i = x_i - x_{i-1}$ dla $i = 1, \dots, k$. Liczbę $\delta = \max_{i=1, \dots, k} \Delta_i$ nazywamy *średnicą podziału* $\Pi = \{x_1, \dots, x_k\}$ przedziału I . Dla $i = 1, \dots, k$ niech

$$m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}, \\ M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}.$$

Liczby s i S , gdzie

$$s = m_1 \Delta_1 + \dots + m_k \Delta_k, \\ S = M_1 \Delta_1 + \dots + M_k \Delta_k,$$

nazywamy odpowiednio *sumą dolną* i *sumą górną funkcji* f odpowiadającą podziałowi Π . Bezpośrednio z powyższej definicji otrzymujemy nierówności

$$m(b-a) \leq s \leq S \leq M(b-a).$$

Rozpatrzmy teraz ciąg podziałów $\{\Pi_n\}_{n=1}^\infty$ przedziału I . Niech δ_n będzie średnicą podziału Π_n , a s_n i S_n odpowiednio sumą dolną i górną funkcji f odpowiadającą podziałowi Π_n , $n = 1, \dots$. Ciąg $\{\Pi_n\}_{n=1}^\infty$ nazywamy *normalnym* jeśli $\lim_{n \rightarrow \infty} \delta_n = 0$.

Twierdzenie 12.1 Niech $f : I \rightarrow \mathbb{R}$ będzie funkcją ograniczoną. Wówczas dla dowolnego normalnego ciągu podziałów przedziału I istnieją skończone granice

$$s = \lim_{n \rightarrow \infty} s_n, \quad S = \lim_{n \rightarrow \infty} S_n$$

i nie zależą one od wyboru ciągu podziałów.

Definicja 12.1 Liczbę s nazywamy *całką dolną* funkcji f na przedziale I , a liczbę S nazywamy *całką górną* funkcji f na przedziale I . Stosujemy też oznaczenia

$$s = \int_a^b f(x) dx = \lim_{n \rightarrow \infty} s_n, \quad S = \int_a^b f(x) dx = \lim_{n \rightarrow \infty} S_n.$$

Definicja 12.2 Niech $f : I \rightarrow \mathbb{R}$ będzie funkcją ograniczoną. Mówimy, że f jest *całkowalna w sensie Riemanna* na przedziale I jeśli jej całka dolna na przedziale I jest równa jej całce górnej funkcji na przedziale I . Wówczas wspólną wartość tych całek nazywamy *całką oznaczoną Riemanna* funkcji f na przedziale I i oznaczamy

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

Bezpośrednio z definicji otrzymujemy nierówności

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

12.2 Warunki całkowalności

Z Twierdzenia 12.1 wynika

Wniosek 12.1 *Niech $f : I \rightarrow \mathbb{R}$ będzie funkcją ograniczoną. Jeśli dla dowolnego $\varepsilon > 0$ istnieje podział Π przedziału I taki, że $S - s \leq \varepsilon$, gdzie s i S są odpowiednio sumami dolną i górną funkcji f dla tego podziału, to f jest całkowalna na przedziale I .*

Twierdzenie 12.2 *Niech $f : I \rightarrow \mathbb{R}$ będzie funkcją ciągłą. Wówczas f jest całkowana w sensie Riemanna.*

Twierdzenie 12.3 *Niech $f : I \rightarrow \mathbb{R}$ będzie funkcją monotoniczną. Wówczas f jest całkowana w sensie Riemanna.*

Definicja 12.3 Niech $A \subset \mathbb{R}$. Mówimy, że A jest zbiorem *miary zero* jeśli dla dowolnego $\varepsilon > 0$ istnieje pokrycie zbioru A ciągiem przedziałów otwartych $\{(a_n, b_n)\}_{n=1}^{\infty}$ takie, że

$$\sum_{n=1}^{\infty} (b_n - a_n) < \varepsilon.$$

Każdy zbiór przeliczalny ma miarę zero. Zbiór Cantora jest przykładem zbioru nieprzeliczalnego o mierze zero.

Twierdzenie 12.4 (Lebesque'a.) *Niech $f : I \rightarrow \mathbb{R}$ będzie funkcją ograniczoną. Wówczas f jest całkowana w sensie Riemanna wtedy i tylko wtedy, gdy zbiór punktów nieciągłości f ma miarę zero.*

Wniosek 12.2

1. *Jeśli $f : I \rightarrow \mathbb{R}$ jest funkcją ograniczoną, całkowalną w sensie Riemanna, to całkowalne są też funkcje $|f|$ i f^2 .*
2. *Jeśli $f, g : I \rightarrow \mathbb{R}$ są funkcjami ograniczonymi, całkowalnymi w sensie Riemanna, to funkcja $f \cdot g$ też jest całkowalna.*

Pojęcie całki oznaczonej można też zdefiniować używając pojęcia ciągu aproksymacyjnego. Otóż jeśli w każdym odcinku $[x_{i-1}, x_i]$ podziału Π wybierzemy dowolnie punkt ξ_i , to sumę

$$\sigma = f(\xi_1)\Delta x_1 + \cdots + f(\xi_k)\Delta x_k$$

nazywamy *sumą przybliżoną*. Oczywiście między sumą dolną, sumą przybliżoną i sumą górną zachodzą nierówności

$$s < \sigma < S.$$

Ciągiem aproksymacyjnym nazywamy ciąg sum przybliżonych odpowiadający normalnemu ciągowi podziałów.

Twierdzenie 12.5 Niech $f : I \rightarrow \mathbb{R}$ będzie funkcją ograniczoną. Jeśli f jest całkowna na przedziale I , to dowolny ciąg aproksymacyjny jest zbieżny do $\int_a^b f(x)dx$. Odwrotnie, jeśli istnieje normalny ciąg podziałów taki, że wszystkie odpowiadające mu ciągi aproksymacyjne są zbieżne do tej samej granicy g , to f jest całkowna oraz $\int_a^b f(x)dx = g$.

12.3 Własności całki oznaczonej

Twierdzenie 12.6 Niech $f, g : I \rightarrow \mathbb{R}$ będą funkcjami całkownymi w sensie Riemanna na przedziale $[a, b]$ oraz $C \in \mathbb{R}$. Wówczas

1. $f + g$ i Cf są całkowne oraz

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

$$\int_a^b (Cf)(x)dx = C \int_a^b f(x)dx,$$

ozn. operacja $f \mapsto \int_a^b f(x)dx$ jest liniowa.

2. Jeśli $f(x) \leq g(x)$ dla $x \in I$, to

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

3. Jeśli $c \in (a, b)$, to całki $\int_a^c f(x)dx$ i $\int_c^b f(x)dx$ istnieją oraz

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

4. Funkcja $|f(x)|$ jest całkowna oraz

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

5. Jeśli f i g są funkcjami ciągłymi na $[a, b]$, $f(x) \leq g(x)$ dla $x \in [a, b]$ oraz $f(\dot{x}) < g(\dot{x})$ dla pewnego $\dot{x} \in [a, b]$, to

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

Twierdzenie 12.7 (O wartości średniej.) Jeśli f jest funkcją ciągłą w przedziale I , to istnieje punkt $\dot{x} \in I$ taki, że

$$f(\dot{x}) = \frac{1}{b-a} \int_a^b f(x)dx.$$

12.4 Podstawowy wzór rachunku całkowego

Twierdzenie 12.8 Niech $f : I \rightarrow \mathbb{R}$ będzie funkcją całkowalną w sensie Riemanna. Połóżmy

$$F(x) = \int_a^x f(x)dx \quad \text{dla } x \in I.$$

Wówczas F jest funkcją ciągłą na I . Ponadto F ma pochodną $F'(x)$ w każdym punkcie x ciągłości funkcji f oraz $F'(x) = f(x)$.

Funkcję F z powyższego twierdzenia nazywamy funkcją pierwotną dla f .

Wniosek 12.3 Jeśli $f : I \rightarrow \mathbb{R}$ jest ciągła oraz F jest jej funkcją pierwotną, to

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

Z powyższego wniosku oraz twierdzeń o całkowaniu przez części i przez podstawienie dostajemy

Twierdzenie 12.9 Jeśli funkcje f i g mają ciągłe pochodne w przedziale I , to

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

Twierdzenie 12.10 Niech $u : [a, b] \rightarrow [\alpha, \beta]$ oraz $f : [\alpha, \beta] \rightarrow \mathbb{R}$ będą funkcjami ciągłymi. Jeśli u ma ciągłą pochodną w przedziale $[a, b]$, to

$$\int_a^b f(u(x))u'(x)dx = \int_{u(a)}^{u(b)} f(u)du.$$

12.5 Zastosowania geometryczne całki oznaczonej

Całka oznaczona znajduje zastosowanie do obliczania pól obszarów na płaszczyźnie i objętości obszarów w przestrzeni.

Wniosek 12.4 Niech $f : I \rightarrow \mathbb{R}$ będzie funkcją ciągłą nieujemną. Wówczas pole obszaru D między wykresem funkcji $y = f(x)$, osią OX i prostymi $x = a$ oraz $x = b$ wyraża się wzorem

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

Wniosek 12.5 Niech $f, g : I \rightarrow \mathbb{R}$ będą funkcjami ciągłymi takimi, że $f(x) \leq g(x)$ dla $x \in I$. Wówczas pole obszaru D między wykresemi funkcji $y = f(x)$, $y = g(x)$ i prostymi $x = a$ oraz $x = b$ wyraża się wzorem

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

Całka oznaczona służy do obliczania objętość brył obrotowych.

Wniosek 12.6 Niech $f : I \rightarrow \mathbb{R}$ będzie funkcją ciągłą nieujemną. Wówczas objętość bryły V powstałej w wyniku obrotu obszaru ograniczonego wykresem funkcji $y = f(x)$ i prostymi OX , $x = a$ oraz $x = b$ dookoła osi OX wyraża się wzorem

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

Założmy teraz, że dane są funkcje ciągłe $x = x(t)$ i $y = y(t)$ określone dla $t \in I$. Niech $\Pi = \{a = t_0 < t_1, \dots, t_{k-1} < t_k = b\}$ będzie podziałem odcinka I . Wówczas punkty $A_i = (x(t_i), y(t_i))$, $i = 0, 1, \dots, k$, leżą na krzywej $\hat{\gamma} = \{(x, y) \in \mathbb{R}^2 : x = x(t), y = y(t)\}$ dla $t \in I$. Długość łamanej $A_0A_1 \cdots A_{k-1}A_k$ dana jest wzorem

$$d = \sum_{i=1}^k \sqrt{(x(t_i) - x(t_{i-1}))^2 + (y(t_i) - y(t_{i-1}))^2}.$$

Dla normalnego ciągu Π_n podziałów odcinka I oznaczmy przez d_n długość łamanej odpowiadającej podziałowi Π_n . Jeśli istnieje granica $\lim_{n \rightarrow \infty} d_n$ i nie zależy ona od wyboru normalnego ciągu podziałów, to krzywą $\hat{\gamma}$ nazywamy *prostowalną*, a granicę $|\hat{\gamma}| = \lim_{n \rightarrow \infty} d_n$ nazywamy *długością $\hat{\gamma}$* .

Wniosek 12.7 Niech $x, y : I \rightarrow \mathbb{R}$ będą funkcjami klasy C^1 na odcinku $x \in I$. Wówczas krzywa $\hat{\gamma} = \{(x, y) \in \mathbb{R}^2 : x = x(t), y = y(t)\}$ dla $t \in I$ jest prostowalna, a jej długość wyraża się wzorem

$$|\hat{\gamma}| = \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2} dx.$$

12.6 Całki niewłaściwe

Całkę oznaczoną Riemanna została zdefiniowana dla funkcji ograniczonej, określonej na odcinku zwartym $I = [a, b]$. Często zachodzi potrzeba zdefiniowania pojęcia całki dla funkcji określonej na przedziale niezwartym, lub też funkcji nieograniczonej. Rozważmy sytuację, że funkcja f jest określona na przedziale prawostronnie otwartym $[a, b)$, gdzie $a < b$ (b może być równe ∞). Jeśli dla dowolnego $\beta < b$ istnieje całka Riemanna

$$I(\beta) = \int_a^\beta f(x) dx$$

oraz istnieje granica (skończona lub nieskończona) $\lim_{\beta \rightarrow b} I(\beta)$, to tą granicę nazywamy *całką niewłaściwą* funkcji f na przedziale $[a, b)$.

Analogicznie określamy całkę niewłaściwą na przedziale lewostronnie otwartym $(a, b]$. Można również zdefiniować całkę na przedziale otwartym (a, b) . W tym celu wybieramy dowolny punkt $c \in (a, b)$ i definiujemy

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

o ile działanie dodawania w tym wzorze jest wykonalne, czyli nie jest to wyrażenie typu $\infty - \infty$.

Całki niewłaściwe są przydatne w badaniu zbieżności szeregów liczbowych.

Twierdzenie 12.11 (Kryterium całkowe zbieżności szeregów.) *Niech $f : [1, \infty) \rightarrow \mathbb{R}$ będzie funkcją ciągłą, nieujemną i nierosnącą. Wówczas szereg $\sum_{n=1}^{\infty} f(n)$ jest zbieżny, wtedy i tylko wtedy, gdy*

$$\int_1^{\infty} f(x)dx < \infty.$$

Przykład 12.1 Niech $f(x) = \frac{1}{x^\alpha}$ dla $x > 0$, gdzie $\alpha > 0$. Wówczas f jest funkcją ciągłą na $(0, \infty)$, nieujemną i malejącą. Dla $\alpha > 1$ mamy

$$\int_1^{\infty} f(x)dx = \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x^\alpha} dx = \lim_{b \rightarrow \infty} \frac{x^{-\alpha+1}}{-\alpha+1} \Big|_{x=1}^{x=b} = \lim_{b \rightarrow \infty} \frac{b^{-\alpha+1} - 1}{-\alpha+1} = \frac{1}{1-\alpha}.$$

Dla $\alpha = 1$ mamy

$$\int_1^{\infty} f(x)dx = \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x} dx = \lim_{b \rightarrow \infty} \ln x \Big|_{x=1}^{x=b} = \lim_{b \rightarrow \infty} \ln b = \infty.$$

Podobnie dla $0 < \alpha < 1$, $\int_1^{\infty} f(x)dx = \infty$.

Zatem szereg $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ jest zbieżny dla $\alpha > 1$ i rozbieżny dla $0 < \alpha \leq 1$.

12.7 Przechodzenie do granicy pod znakiem całki

Niech f_n będzie ciągiem funkcji całkownych w przedziale zwartym $I = [a, b]$ zbieżnym do funkcji f . Interesuje nas pytanie czy funkcja f jest całkowna i czy ciąg liczbowy $\int_a^b f_n(x)dx$ jest zbieżny do $\int_a^b f(x)dx$? Jak pokazuje następujący przykład w ogólnym przypadku odpowiedź na to pytanie jest negatywna.

Przykład 12.2 Niech

$$f_n(x) = 2nxe^{-nx^2} \quad \text{dla } x > 0.$$

Wówczas $\lim_{n \rightarrow \infty} f_n(x) = 0$ dla dowolnego $x > 0$. Tym nie mniej

$$\lim_{n \rightarrow \infty} \int_0^1 f_n(x)dx = \lim_{n \rightarrow \infty} -e^{-nx^2} \Big|_{x=0}^{x=1} = \lim_{n \rightarrow \infty} (1 - e^{-n}) = 1.$$

Powyższy przykład pokazuje, że ze zbieżności punktowej ciągu funkcji nie można wnioskować o zbieżności ciągu całek z tych funkcji.

Twierdzenie 12.12 *Niech $\{f_n\}_{n=1}^{\infty}$ będzie ciągiem funkcji całkownych na przedziale I zbieżnym jednostajnie na I do funkcji f . Wówczas f jest całkowna na I oraz dla dowolnego $\dot{x} \in I$ ciąg funkcji $\varphi_n(x) = \int_{\dot{x}}^x f_n(t)dt$ jest zbieżny jednostajnie do funkcji $\varphi(x) = \int_{\dot{x}}^x f(t)dt$ na I . W szczególności $\lim_{n \rightarrow \infty} \int_a^b f_n(t)dt = \int_a^b f(t)dt$.*

Wniosek 12.8 Niech $\{f_n\}_{n=1}^{\infty}$ będzie ciągiem funkcji całkownych na przedziale I oraz szereg $\sum_{n=1}^{\infty} f_n$ jest zbieżnym jednostajnie na I do funkcji S na I . Wówczas S jest całkowna na I oraz dla dowolnego $\dot{x} \in I$ szereg $\sum_{n=1}^{\infty} \int_{\dot{x}}^x f_n(t) dt$ jest zbieżny jednostajnie do funkcji $\varphi(x) = \int_{\dot{x}}^x S(t) dt$ na I .

Przykład 12.3 Rozważmy szereg $\sum_{n=1}^{\infty} nx^{n-1}$. Z kryterium d’Alamberta wynika, że szereg ten jest zbieżny dla $x \in (-1, 1)$. Oznaczmy jego sumę przez $S(x)$. Ponieważ szereg ten jest zbieżny jednostajnie na dowolnym domkniętym podprzedziale $[a, b] \subset (-1, 1)$, więc korzystając z Wniosku 12.7 dostajemy

$$\int_0^x S(t) dt = \int_0^x \sum_{n=1}^{\infty} nt^{n-1} dt = \sum_{n=1}^{\infty} \int_0^x nx^{n-1} dt = \sum_{n=1}^{\infty} x^n = \frac{x}{1-x}.$$

Stąd

$$S(x) = \left(\frac{x}{1-x} \right)' = \frac{1}{(1-x)^2} \quad \text{dla } x \in (-1, 1).$$

12.8 Całki zależne od parametru

Niech $U \subset \mathbb{R}^2$ będzie zbiorem otwartym, $[a, b] \times [c, d] \subset U$ oraz $f : U \rightarrow \mathbb{R}$. Załóżmy, że dla każdego $y \in [c, d]$ istnieje całka

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

Interesuje nas pytanie czy funkcja g jest różniczkowalna?

Twierdzenie 12.13 Jeśli w prostokącie $P = [a, b] \times [c, d]$ funkcja f jest ciągła i ma ciągłą pochodną cząstkową $\frac{\partial f}{\partial y}$, to funkcja

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

jest różniczkowalna dla $y \in (c, d)$ oraz

$$g'(y) = \int_a^b \frac{\partial f}{\partial y}(x, y) dx \quad \text{dla } y \in (c, d).$$

Całki zależne od parametru umożliwiają zdefiniowanie wielu nowych funkcji. Jedną z nich, niezwykle ważną w Analizie, jest funkcja Γ -Eulera. Określamy ją wzorem

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \quad \text{dla } x > 0.$$

Całka w tym wzorze jest całką niewłaściwą. Można wykazać, że jest ona zbieżna dla $x > 0$. Korzystając ze wzory na całkowanie przez części dostajemy wzór

$$\Gamma(x+1) = x\Gamma(x) \quad \text{dla } x > 0.$$

Ponieważ $\Gamma(1) = \int_0^\infty e^{-t} dt = 1$ wnioskujemy stąd, że $\Gamma(n) = (n - 1)!$ dla $n \in \mathbb{N}$. Zatem funkcja Γ jest przedłużeniem funkcji „silnia” na argumenty dodatnie. Okazuje się, że można ją też zdefiniować na płaszczyźnie zespolonej.

Inną ważną funkcją definiowaną przy pomocy całki z parametrami jest funkcja B -Eulera.

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt \quad \text{dla } x > 0, y > 0.$$

Związek pomiędzy funkcjami Γ i B wyraża wzór

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \quad \text{dla } x > 0, y > 0.$$

Wzór ten udowodnimy po nauczeniu się całkowania funkcji wielu zmiennych.

Zauważmy, że kładąc $t = \sin^2 \varphi$, mamy $1 - t = \cos^2 \varphi$, $dt = 2 \sin \varphi \cos \varphi$. Zatem

$$2 \int_0^{\pi/2} \sin^{2x-1} \varphi \cdot \cos^{2y-1} \varphi d\varphi = B(x, y).$$

13 Całki wielokrotne

13.1 Definicja i własności całki n -krotnej na przedziale

Przypomnijmy, że *przedziałem domkniętym* lub *kostką domkniętą* w przestrzeni n -wymiarowej \mathbb{R}^n nazywamy zbiór

$$P = P(a, b) = \{x \in \mathbb{R}^n : a_i \leq x_i \leq b_i, i = 1, \dots, n\}, \quad \text{gdzie } a < b.$$

Średnicą przedziału P nazywamy liczbę

$$\delta = \delta(P) = \sqrt{(b_1 - a_1)^2 + \dots + (b_n - a_n)^2}.$$

n -wymiarową miarą lub objętością przedziału P nazywamy iloczyn

$$|P| = (b_1 - a_1) \cdots (b_n - a_n).$$

Podziałem przedziału domkniętego P nazywamy rodzinę $\Pi = \{P_1, \dots, P_k\}$ przedziałów domkniętych P_i o rozłącznych wnętrzach

$$\text{int}P_i \cap \text{int}P_j \quad \text{dla } i \neq j$$

taką, że

$$P = \bigcup_{i=1}^k P_i.$$

Jeśli δ_i jest średnicą przedziału P_i w podziale Π przedziału P , to $\widehat{\delta} = \max_{i=1, \dots, k} \delta_i$ nazywamy *średnicą podziału* Π . Ciąg podziałów $\{\Pi_j\}_{j=1}^{\infty}$ przedziału P nazywamy *normalnym* jeśli $\lim_{j \rightarrow \infty} \widehat{\delta}_j = 0$ gdzie $\widehat{\delta}_j$ jest średnicą podziału Π_j .

Niech $f : P \rightarrow \mathbb{R}$ funkcją ograniczoną zdefiniowaną na przedziale domkniętym P , a $\Pi = \{P_1, \dots, P_k\}$ podziałem tego przedziału. Oznaczmy odpowiednio przez m i M kresy dolny i górny funkcji f na przedziale P , tzn.

$$m = \inf_{x \in P} f(x), \quad M = \sup_{x \in P} f(x).$$

Dla $i = 1, \dots, k$ niech

$$m_i = \inf\{f(x) : x \in P_i\}, \\ M_i = \sup\{f(x) : x \in P_i\}.$$

Liczby s i S , gdzie

$$s = m_1|P_1| + \dots + m_k|P_k|, \\ S = M_1|P_1| + \dots + M_k|P_k|,$$

nazywamy odpowiednio *sumą dolną* i *sumą górną* funkcji f odpowiadającą podziałowi Π . Bezpośrednio z powyższej definicji otrzymujemy nierówności

$$m(b-a)^{\mathbf{1}} \leq s \leq S \leq M(b-a)^{\mathbf{1}},$$

gdzie $(b-a)^{\mathbf{1}} = (b_1 - a_1) \cdots (b_n - a_n)$.

Rozpatrzmy teraz normalny ciąg podziałów $\{\Pi_j\}_{j=1}^{\infty}$ przedziału P . Niech $\widehat{\delta}_j$ będzie średnicą podziału Π_j , a s_j i S_j odpowiednio sumą dolną i górną funkcji f odpowiadającą podziałowi Π_j , $j = 1, 2, \dots$

Lemat 13.1 Niech $f : P \rightarrow \mathbb{R}$ będzie funkcją ograniczoną. Wówczas dla dowolnego normalnego ciągu podziałów przedziału P istnieją skończone granice

$$s = \lim_{j \rightarrow \infty} s_j, \quad S = \lim_{j \rightarrow \infty} S_j$$

i nie zależą one od wyboru normalnego ciągu podziałów.

Definicja 13.1 Liczbę s nazywamy *całką dolną* funkcji f na przedziale P , a liczbę S nazywamy *całką górną* funkcji f na przedziale P . Stosujemy też oznaczenia

$$s = \int_a^b f(x)dx = \lim_{j \rightarrow \infty} s_j, \quad S = \int_a^b f(x)dx = \lim_{j \rightarrow \infty} S_j.$$

Definicja 13.2 Niech $f : P \rightarrow \mathbb{R}$ będzie funkcją ograniczoną. Mówimy, że f jest *całkowalna w sensie Riemanna* na przedziale P jeśli jej całka dolna na przedziale P jest równa jej całce górnej funkcji na przedziale P . Wówczas wspólną wartość tych całek nazywamy *całką oznaczoną Riemanna* funkcji f na przedziale P i oznaczamy

$$\int_P f(x)dx = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \dots, x_n)dx_1 \dots dx_n.$$

Własności całki Riemanna.

1. Jeśli $f : P \rightarrow \mathbb{R}$ jest funkcją ciągłą na przedziale domkniętym P , to f jest całkowalna na P .

2. Jeśli $f, g : P \rightarrow \mathbb{R}$ są całkowalne na P i $c \in \mathbb{R}$, to $f + g$ i cf są całkowalne na P oraz

$$\int_P (f(x) + g(x))dx = \int_P f(x)dx + \int_P g(x)dx,$$

$$\int_P c \cdot f(x)dx = c \cdot \int_P f(x)dx.$$

Definicja 13.3 Niech $\Pi = \{P_1, \dots, P_k\}$ będzie podziałem przedziału P oraz niech $\xi_i \in P_i$ dla $i = 1, \dots, k$. Sumę

$$\sigma = f(\xi_1)|P_1| + \cdots + f(\xi_k)|P_k|$$

nazywamy *sumą przybliżoną*. Ciąg sum przybliżonych dla normalnego ciągu podziałów nazywamy *ciągami aproksymacyjnymi*.

Lemat 13.2 Funkcja $f : P \rightarrow \mathbb{R}$ jest całkowalna na przedziale P wtedy i tylko wtedy, gdy istnieje normalny ciąg podziałów taki, że wszystkie ciągi aproksymacyjne są zbieżne do tej samej granicy.

Interpretacja geometryczna całki podwójnej.

Niech $P = [a, b] \times [c, d] \subset \mathbb{R}^2$ oraz niech $f : P \rightarrow \mathbb{R}$ będzie funkcją ciągłą na P . Jeśli $f \geq 0$, to $\int_P f(x, y)dxdy$ jest objętością bryły G , gdzie

$$G = \{(x, y, z) \in \mathbb{R}^3 : a \leq x \leq b, c \leq y \leq d, 0 \leq z \leq f(x, y)\}.$$

13.2 Całka iterowana

Obliczanie całki wielokrotnej bezpośrednio z definicji jest zadaniem trudnym i praktycznie niewykonalnym. Przykładowo zadanie obliczenia całki

$$\iint_{[0,1]^2} xy dx dy$$

sprowadza się do znalezienia granicy przy $k \rightarrow \infty$ podwójnej sumy

$$\sum_{i,j=1}^k \frac{i \cdot j}{k^2} \cdot \frac{1}{k^2} = \frac{k(k+1)}{2k^2} \cdot \frac{k(k+1)}{2k^2} \rightarrow \frac{1}{4}.$$

Okazuje się, że całkę wielokrotną można sprowadzić do całki iterowanej, którą z kolei można obliczyć korzystając z całki nieoznaczonej jednej zmiennej. Pojęcie całki iterowanej wprowadzimy w przypadku funkcji dwóch zmiennych. W ogólnym przypadku pojęcie to można naturalnie uogólnić.

Definicja 13.4 Niech $P = [a, b] \times [c, d] \subset \mathbb{R}^2$ oraz niech $f : P \rightarrow \mathbb{R}$ będzie funkcją ograniczoną. Załóżmy, że dla każdego ustalonego $y \in [c, d]$ istnieje całka oznaczona Riemanna $\int_a^b f(x, y) dx$. Jeśli funkcja

$$[c, d] \ni y \mapsto g(y) = \int_a^b f(x, y) dx$$

jest całkowna w sensie Riemanna, to jej całkę na przedziale $[c, d]$ nazywamy *całką iterowaną* funkcji f i oznaczamy

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

Analogicznie definiujemy całkę iterowaną

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

W przypadku funkcji n zmiennych definiuje się $n!$ całek iterowanych.

Istnieją przykłady funkcji, dla których istnieją całki iterowane lecz nie istnieje całka podwójna. Tym nie mniej zachodzi

Twierdzenie 13.1 (Fubinięgo.) *Niech $P = [a, b] \times [c, d] \subset \mathbb{R}^2$ oraz niech $f : P \rightarrow \mathbb{R}$ będzie funkcją ciągłą. Wówczas istnieją całki iterowane i są równe całce podwójnej*

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

13.3 Całka wielokrotna na zbiorze ograniczonym

Pojęcie całki wielokrotnej można uogólnić na funkcje określone na zbiorach ograniczonych.

Definicja 13.5 Niech $D \subset \mathbb{R}^n$ będzie zbiorem ograniczonym oraz $f : P \rightarrow \mathbb{R}$ funkcją ograniczoną. Niech P będzie przedziałem zawierającym zbiór D . Wówczas określamy

$$\int_D f(x)dx = \int_P f_0(x)dx,$$

gdzie

$$f_0(x) = \begin{cases} f(x) & \text{dla } x \in D, \\ 0 & \text{dla } x \notin D. \end{cases}$$

Całki wielokrotne na *dobrych* zbiorach można wyznaczyć korzystając z całek iterowanych.

Definicja 13.6 Zbiór $D \subset \mathbb{R}^2$ nazywamy *normalnym względem osi OX* jeśli jest on postaci

$$D = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \varphi(x) \leq y \leq \psi(x)\},$$

gdzie φ, ψ są funkcjami ciągłymi na $[a, b]$ oraz $\varphi(x) \leq \psi(x)$ dla $x \in [a, b]$.

Analogicznie $D \subset \mathbb{R}^2$ nazywamy *normalnego względem osi OY* jeśli jest on postaci

$$D = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, \phi(y) \leq x \leq \chi(y)\},$$

gdzie ϕ, χ są funkcjami ciągłymi na $[c, d]$ oraz $\phi(y) \leq \chi(y)$ dla $y \in [c, d]$.

Twierdzenie 13.2 Jeśli $f : D \rightarrow \mathbb{R}$ będzie funkcją ciągłą na zbiorze $D \subset \mathbb{R}^2$ normalnym względem osi OX, to

$$\iint_D f(x, y)dx dy = \int_a^b \left[\int_{\varphi(x)}^{\psi(x)} f(x, y)dy \right] dx.$$

Analogicznie jeśli $f : D \rightarrow \mathbb{R}$ będzie funkcją ciągłą na zbiorze $D \subset \mathbb{R}^2$ normalnym względem osi OY, to

$$\iint_D f(x, y)dx dy = \int_c^d \left[\int_{\phi(y)}^{\chi(y)} f(x, y)dx \right] dy.$$

Zbiór $D \subset \mathbb{R}^2$ będący skończoną sumą zbiorów normalnych względem osi OX lub OY o rozłącznych wnętrzach nazywamy *zbiorem regularnym*.

Jeśli $D = D_1 \cup \dots \cup D_k$ jest zbiorem regularnym, to

$$\iint_D f(x, y)dx dy = \sum_{i=1}^k \iint_{D_i} f(x, y)dx dy$$

Analogicznie definiujemy zbiory regularne w \mathbb{R}^3 . Przykładowo zbiór $D \subset \mathbb{R}^3$ nazywamy *normalnym względem osi OX i płaszczyzny OXY* jeśli jest on postaci

$$D = \{(x, y, z) \in \mathbb{R}^3 : a \leq x \leq b, \varphi(x) \leq y \leq \psi(x), g(x, y) \leq z \leq h(x, y)\}.$$

Wówczas dla funkcji f ciągłej na D mamy

$$\iiint_D f(x, y, z) dx dy dz = \int_a^b \left(\int_{\varphi(x)}^{\psi(x)} \left(\int_{g(x, y)}^{h(x, y)} f(x, y, z) dz \right) dy \right) dx.$$

13.4 Zastosowania całek w geometrii

Pole obszaru płaskiego, regularnego $D \subset \mathbb{R}^2$ zadane jest wzorem

$$|D| = \iint_D 1 dx dy.$$

Środek ciężkości zbioru regularnego $D \subset \mathbb{R}^2$ ma współrzędne (ξ, η) , gdzie

$$\xi = \frac{1}{|D|} \iint_D x dx dy, \quad \eta = \frac{1}{|D|} \iint_D y dx dy.$$

Objętość bryły $V \subset \mathbb{R}^3$ zdefiniowanej jako

$$V = \{(x, y, z) \in \mathbb{R}^3 : g(x, y) \leq z \leq h(x, y), (x, y) \in D\},$$

gdzie $D \subset \mathbb{R}^2$ jest zbiorem regularnym, g i h funkcjami ciągłymi na D wyraża się wzorem

$$|V| = \iiint_V 1 dx dy dz = \iint_D (h(x, y) - g(x, y)) dx dy.$$

Niech $D \subset \{(x, y) : y > 0\}$ będzie zbiorem regularnym oraz V bryłą powstałą przez obrót zbioru D wokół osi OX . Wówczas

$$|V| = 2\pi \iint_D y dx dy.$$

Reguła Guldina. Objętość bryły obrotowej V powstałej w wyniku obrotu zbioru normalnego $D \subset \{(x, y) : y > 0\}$ dookoła osi OX jest równa iloczynowi pola zbioru D przez drogę, którą zatacza środek ciężkości zbioru D podczas tego obrotu, tzn.

$$|V| = 2\pi\eta|D|.$$

Przykład 13.1 Niech T będzie torusem powstałym w wyniku obrotu koła $K = \{(x, y, z) : (x - a)^2 + z^2 \leq r^2, y = 0\}$, $0 < r \leq a$, dookoła osi OZ . Ponieważ punkt ciężkości koła K ma współrzędne $(a, 0, 0)$, a jego pole wynosi πr^2 , więc z reguły Guldina dostajemy objętość torusa T ,

$$|T| = 2\pi \cdot a \cdot \pi r^2 = 2\pi^2 ar^2.$$

Twierdzenie 13.3 Pole powierzchni wykresu funkcji. Niech $D \subset \mathbb{R}^2$ będzie zbiorem regularnym, a $f : D \rightarrow \mathbb{R}$ funkcją klasy C^1 . Wówczas pole powierzchni wykresu funkcji f , t.j. pole powierzchni zbioru

$$S = \{(x, y, z) \in \mathbb{R}^3 : z = f(x, y), (x, y) \in D\}$$

wyraża się wzorem

$$|S| = \iint_D \sqrt{1 + \left(\frac{\partial f(x, y)}{\partial x}\right)^2 + \left(\frac{\partial f(x, y)}{\partial y}\right)^2} dx dy.$$

13.5 Zamiana zmiennych w całce wielokrotnej

Dla funkcji jednej zmiennej zachodzi wzór na całkowanie przez podstawienie.

Mianowicie, jeśli $\varphi : [a, b] \rightarrow [\alpha, \beta]$ oraz $f : [\alpha, \beta] \rightarrow \mathbb{R}$ są funkcjami ciągłymi oraz φ ma ciągłą pochodną w przedziale $[a, b]$, to

$$\int_a^b f(\varphi(x))\varphi'(x) dx = \int_{\varphi(a)}^{\varphi(b)} f(y) dy.$$

Uogólnimy teraz też wzór na przypadek całek wielokrotnych.

Twierdzenie 13.4 Niech Δ i D będą regularnymi zbiorami w \mathbb{R}^n oraz niech $\Phi = (\varphi_1, \dots, \varphi_n) : \Delta \rightarrow D$ będzie wzajemnie jednoznacznym odwzorowaniem klasy C^1 zbioru Δ na D (dyfeomorfizmem). Jeśli $f : D \rightarrow \mathbb{R}$ jest funkcją ciągłą, to

$$\int_D f(y) dy = \int_{\Delta} f(\Phi(x)) |J\Phi(x)| dx,$$

gdzie $J\Phi = \det \frac{D(\varphi_1, \dots, \varphi_n)}{D(x_1, \dots, x_n)}$ jest jakobianem odwzorowania Φ .

Dowód Twierdzenia 13.4 przeprowadzimy dla $n = 2$. Wówczas można je sformułować następująco.

Twierdzenie 13.5 Niech Δ i D będą regularnymi zbiorami w \mathbb{R}^2 oraz niech $\Phi = (\varphi, \psi) : \Delta \rightarrow D$ będzie wzajemnie jednoznacznym odwzorowaniem zbioru Δ na D , przy czym φ, ψ są klasy C^1 , $\varphi(u, v) = x$, $\psi(u, v) = y$ dla $(u, v) \in \Delta$.

Jeśli $f : D \rightarrow \mathbb{R}$ jest funkcją ciągłą, to

$$\iint_D f(x, y) dx dy = \iint_{\Delta} f(\varphi(u, v), \psi(u, v)) \left| \det \frac{D(\varphi, \psi)}{D(u, v)}(u, v) \right| du dv.$$

W wymiarze $n = 3$ Twierdzenie 13.4 przyjmuje postać.

Twierdzenie 13.6 Niech Δ i D będą regularnymi zbiorami w \mathbb{R}^3 oraz niech $\Phi = (\varphi, \psi, \chi) : \Delta \rightarrow D$ będzie wzajemnie jednoznacznym odwzorowaniem zbioru Δ na D , przy czym φ, ψ, χ są klasy C^1 , $\varphi(u, v, w) = x$, $\psi(u, v, w) = y$, $\chi(u, v, w) = z$ dla $(u, v, w) \in \Delta$. Jeśli $f : D \rightarrow \mathbb{R}$ jest funkcją ciągłą, to

$$\begin{aligned} & \iiint_D f(x, y, z) dx dy dz \\ &= \iiint_{\Delta} f(\varphi(u, v, w), \psi(u, v, w), \chi(u, v, w)) \left| \det \frac{D(\varphi, \psi, \chi)}{D(u, v, w)}(u, v, w) \right| du dv dw \end{aligned}$$

13.6 Przykłady

Przykład 13.2 Niech Δ i D będą regularnymi zbiorami w \mathbb{R}^n oraz niech $\Phi : \Delta \rightarrow D$ będzie wzajemnie jednoznacznym odwzorowaniem klasy C^1 zbioru Δ na D . Niech $f(x) = 1$ dla $x \in D$. Wówczas

$$|D| = \int_D 1 \, dy = \int_{\Delta} |J\Phi(x)| \, dx.$$

Liniowa zamiana zmiennych.

Niech $\Phi : \Delta \rightarrow D$ będzie wzajemnie jednoznacznym odwzorowaniem liniowym, tzn. $\Phi(x) = Ax$ dla pewnej nieosobliwej macierzy A . Wówczas macierzą różniczki Φ jest A czyli $J\Phi(x) = \det A$. Zatem dla funkcji f ciągłej na D mamy

$$\int_D f(y) \, dy = |\det A| \int_{\Delta} f(Ax) \, dx.$$

W szczególności, biorąc $f \equiv 1$ dostajemy

$$|D| = |\det A| \cdot |\Delta|.$$

Czyli odwzorowanie liniowe zmienia objętość obszaru proporcjonalnie do wyznacznika macierzy A .

Współrzędne biegunowe

Jeśli obszar $D \subset \mathbb{R}^2$ jest kołem (lub wycinkiem koła, lub pierścieniem), to do obliczenia całki po D stosujemy współrzędne biegunowe: $x = r \cos \varphi$, $y = r \sin \varphi$. Wówczas odwzorowanie $(r, \varphi) \rightarrow (x, y)$ przekształca prostokąt $\Delta = [0, R] \times [0, 2\pi]$ na koło $D = K(0, R) = \{x^2 + y^2 \leq R^2\}$. Jest to odwzorowanie wzajemnie jednoznaczne z $\text{Int}\Delta$ na $\text{Int}D \setminus \{(x, y) : x \geq 0, y = 0\}$. Ponieważ $\Delta \setminus \text{Int}\Delta$ ma miarę zero możemy stosować Twierdzenie 13.5.

Przykład 13.3 Policzmy całkę

$$\iint_D \frac{dx dy}{\sqrt{x^2 + y^2}}, \quad \text{gdzie } D = \{x^2 + y^2 \leq 1\}.$$

Ponieważ funkcja podcałkowa jest nieograniczona w zerze powyższa całka jest całką niewłaściwą, tzn. trzeba policzyć granicę

$$\lim_{a \rightarrow 0} \iint_{D_a} \frac{dx dy}{\sqrt{x^2 + y^2}}, \quad \text{gdzie } D_a = \{a^2 \leq x^2 + y^2 \leq 1\}, a > 0.$$

Stosując współrzędne biegunowe dostajemy

$$\begin{aligned} \lim_{a \rightarrow 0} \iint_{D_a} \frac{dx dy}{\sqrt{x^2 + y^2}} &= \lim_{a \rightarrow 0} \iint_{[a,1] \times [0,2\pi]} \frac{1}{r} \cdot r \, dr d\varphi \\ &= \lim_{a \rightarrow 0} \int_a^1 \left(\int_0^{2\pi} 1 d\varphi \right) dr = \lim_{a \rightarrow 0} 2\pi(1 - a) = 2\pi. \end{aligned}$$

Przykład 13.4 Policzmy całkę

$$I = \iint_{\mathbb{R}^2} e^{-x^2-y^2} dx dy.$$

Ponieważ całkujemy po obszarze nieograniczonym powyższa całka jest całką niewłaściwą, tzn. trzeba policzyć granicę

$$I = \lim_{R \rightarrow \infty} \iint_{D_R} e^{-x^2-y^2} dx dy, \quad \text{gdzie } D_R = \{x^2 + y^2 \leq R^2\}, \quad R > 0.$$

Stosując współrzędne biegunowe dostajemy

$$\begin{aligned} & \lim_{R \rightarrow \infty} \iint_{D_R} e^{-x^2-y^2} dx dy \\ &= \lim_{R \rightarrow \infty} \iint_{[0,R] \times [0,2\pi]} e^{-r^2} \cdot r dr d\varphi = \lim_{R \rightarrow \infty} \int_0^R \left(\int_0^{2\pi} e^{-r^2} r d\varphi \right) dr \\ &= \lim_{R \rightarrow \infty} \int_0^R 2\pi r e^{-r^2} dr = \lim_{R \rightarrow \infty} -\pi e^{-r^2} \Big|_{r=0}^{r=R} \\ &= \lim_{R \rightarrow \infty} (\pi - \pi e^{-R^2}) = \pi. \end{aligned}$$

Z drugiej strony na mocy twierdzenia Fubiniego

$$\begin{aligned} I &= \iint_{\mathbb{R}^2} e^{-x^2-y^2} dx dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-x^2} dx \right) e^{-y^2} dy \\ &= \int_{\mathbb{R}} e^{-x^2} dx \cdot \int_{\mathbb{R}} e^{-y^2} dy. \end{aligned}$$

Zatem wykazaliśmy ważny wzór

$$\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}.$$

Związek pomiędzy funkcjami Γ i B -Eulera.

Przypomnijmy, że funkcje Γ i B -Eulera były określone dla $x > 0$, $y > 0$ wzorami

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt.$$

Wyprowadzimy teraz związek pomiędzy tymi funkcjami. Korzystając z twierdzenia Fubiniego mamy

$$\Gamma(x)\Gamma(y) = \int_0^\infty t^{x-1} e^{-t} dt \cdot \int_0^\infty s^{y-1} e^{-s} ds = \int_0^\infty \int_0^\infty t^{x-1} s^{y-1} e^{-t-s} dt ds.$$

W całce podwójnej dokonajmy zamiany zmiennych $t = u$, $t + s = v$. Wówczas $(t, s) \in \mathbb{R}_+ \times \mathbb{R}_+$ wtedy i tylko wtedy, gdy $v \in \mathbb{R}_+$, $0 \leq u \leq v$ oraz $dt ds = du dv$. Zatem

$$\begin{aligned} \int_0^\infty \int_0^\infty t^{x-1} s^{y-1} e^{-t-s} dt ds &= \int_0^\infty \int_0^v u^{x-1} (v-u)^{y-1} e^{-v} du dv \\ &= \int_0^\infty \left(\int_0^v u^{x-1} (v-u)^{y-1} du \right) e^{-v} dv. \end{aligned}$$

Dokonajmy teraz zamiany zmiennych w całce wewnętrznej $u = tv$, $0 \leq t \leq 1$, $du = v dt$,

$$\begin{aligned} \int_0^v u^{x-1} (v-u)^{y-1} du &= \int_0^1 (tv)^{x-1} (v-tv)^{y-1} v dt \\ &= v^{x+y-1} \int_0^1 t^{x-1} (1-t)^{y-1} dt = v^{x+y-1} B(x, y). \end{aligned}$$

Zatem

$$\Gamma(x)\Gamma(y) = \int_0^\infty v^{x+y-1} B(x, y) e^{-v} dv = B(x, y)\Gamma(x+y).$$

Współrzędne walcowe.

Jeśli obszar $V \subset \mathbb{R}^3$ jest częścią walca lub stożka, to do liczenia całek wygodnie jest stosować współrzędne walcowe. $x = r \cos \varphi$, $y = r \sin \varphi$, $z = z$. Wówczas

$$|J(r, \varphi, z)| = \left| \det \begin{pmatrix} \cos \varphi & \sin \varphi & 0 \\ -r \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \right| = r.$$

Przykład 13.5 Policzmy całkę

$$\iiint_V (x^2 + y^2) dx dy dz,$$

gdzie V jest stożkiem ściętym $V = \{(x, y, z) : x^2 + y^2 \leq z^2, 0 < R_1 \leq z \leq R_2\}$. Przeciwobrazem V przy współrzędnych walcowych jest $\Delta = \{(r, \varphi, z) : 0 \leq r \leq z, 0 \leq \varphi \leq 2\pi, R_1 \leq z \leq R_2\}$. Zatem

$$\begin{aligned} \iiint_V (x^2 + y^2) dx dy dz &= \iiint_\Delta r^2 \cdot r dr d\varphi dz = \int_{R_1}^{R_2} \left(\int_0^{2\pi} \left(\int_0^z r^3 dr \right) d\varphi \right) dz \\ &= \int_{R_1}^{R_2} \left(\int_0^{2\pi} \frac{z^4}{4} d\varphi \right) dz = \int_{R_1}^{R_2} \frac{\pi z^4}{2} dz = \frac{\pi}{10} (R_2^5 - R_1^5). \end{aligned}$$

Współrzędne sferyczne.

Jeśli obszar $V \subset \mathbb{R}^3$ jest częścią kuli, to do liczenia całek stosujemy współrzędne sferyczne $x = r \cos \varphi \cos \psi$, $y = r \sin \varphi \cos \psi$, $z = r \sin \psi$. Wówczas

$$|J(r, \varphi, \psi)| = \left| \det \begin{pmatrix} \cos \varphi \cos \psi & \sin \varphi \cos \psi & \sin \psi \\ -r \sin \varphi \cos \psi & r \cos \varphi \cos \psi & 0 \\ -r \cos \varphi \sin \psi & -r \sin \varphi \sin \psi & r \cos \psi \end{pmatrix} \right| = r^2 \cos \psi.$$

Przykład 13.6 Policzmy objętość kuli $B(R)$ o promieniu R ,

$$|B(R)| = \iiint_{B(R)} 1 \, dx dy dz.$$

Przeciwbrazem $B(R)$ przy współrzędnych sferycznych jest $\Delta = \{(r, \varphi, \psi) : 0 \leq r \leq R, 0 \leq \varphi \leq 2\pi, -\pi/2 \leq \psi \leq \pi/2\}$. Zatem

$$\begin{aligned} |B(R)| &= \iiint_{\Delta} r^2 \cos \psi \, dr d\varphi d\psi = \int_0^R r^2 \, dr \cdot \int_0^{2\pi} d\varphi \cdot \int_{-\pi/2}^{\pi/2} \cos \psi \, d\psi \\ &= \frac{R^3}{3} \cdot 2\pi \cdot 2 = \frac{4}{3}R^3. \end{aligned}$$

Zauważmy, że przekształcenie liniowe $x = ax'$, $y = by'$, $z = cz'$ przekształca elipsoidę $E = \{\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1\}$ na kulę $B = \{x'^2 + y'^2 + z'^2 \leq 1\}$ oraz $dx dy dz = abc dx' dy' dz'$. Zatem $|E| = \frac{4}{3}\pi abc$.

Pole krzywoliniowego prostokąta.

Przykład 13.7 Obliczmy pole krzywoliniowego prostokąta D ograniczonego parabolami $y^2 = px$, $y^2 = qx$, $x^2 = ay$ i $x^2 = by$, gdzie $0 < p < q$, $0 < a < b$.

Zauważmy, że dowolny punkt $(x, y) \in D$ jest jednoznacznie wyznaczony przez parę $(u, v) \in \Delta = [p, q] \times [a, b]$, jako przecięcie parabol $y^2 = ux$ i $x^2 = vy$, mianowicie $x = \sqrt[3]{uv^2}$, $y = \sqrt[3]{u^2v}$. Zatem mamy dyfeomorfizm

$$\Delta \ni (u, v) \rightarrow \Phi(u, v) = (\sqrt[3]{uv^2}, \sqrt[3]{u^2v}) \in D.$$

Liczmy moduł jacobianu $J\Phi$,

$$|J\Phi(u, v)| = \left| \det \begin{pmatrix} \frac{1}{3}u^{-2/3}v^{2/3} & \frac{2}{3}u^{1/3}v^{-1/3} \\ \frac{2}{3}u^{-1/3}v^{1/3} & \frac{1}{3}u^{2/3}v^{-2/3} \end{pmatrix} \right| = \frac{1}{3}.$$

Zatem

$$|D| = \iint_D dx dy = \iint_{\Delta} \frac{1}{3} du dv = \frac{1}{3}(q-p)(b-a).$$

Analogicznie obliczamy pola krzywoliniowych prostokątów ograniczonych przez odcinki hiperbol, parabol, prostych itp.

14 Orientacja

14.1 Orientacja przestrzeni

Niech (e_1, \dots, e_n) , gdzie $e_i = (0, \dots, 1, \dots, 0)$, $i = 1, \dots, n$, będzie bazą standardową przestrzeni \mathbb{R}^n . Niech (f_1, \dots, f_n) będzie inną bazą przestrzeni \mathbb{R}^n , przy czym kolejność wektorów f_1, \dots, f_n jest istotna. Każdy wektor f_j można przedstawić w postaci liniowej kombinacji wektorów e_i , $i = 1, \dots, n$, tzn.

$$f_j = \sum_{i=1}^n a_{ij} e_i, \quad j = 1, \dots, n.$$

Utwórzmy macierz $A = (a_{ij})_{i,j=1,\dots,n}$. Mówimy, że baza (f_1, \dots, f_n) wyznacza *orientację* przestrzeni \mathbb{R}^n *zgodną* z bazą (e_1, \dots, e_n) (odpowiednio, *przeciwną*) jeśli $\text{Det}A > 0$ (odpowiednio, $\text{Det}A < 0$). Orientację wyznaczoną przez bazę (e_1, \dots, e_n) nazywa się też *orientacją dodatnią*, a orientację do niej przeciwną orientacją ujemną.

Przykład 14.1 Niech $f_1 = (2, 1)$, $f_2 = (1, 3)$. Wówczas (f_1, f_2) wyznacza orientację dodatnią \mathbb{R}^2 , natomiast (f_2, f_1) orientację ujemną.

14.2 Orientacja rozmaitości

Niech $S \subset \mathbb{R}^m$ będzie n -wymiarową rozmaitością, $n \leq m$. Wówczas dla każdego punktu $x \in S$ istnieje zbiór otwarty $\Omega \subset \mathbb{R}^n$ oraz dyfeomorfizm $\Phi : \Omega \rightarrow \mathbb{R}^m$ taki, że $\Phi(\Omega) \ni x$. Oznaczmy $a = \Phi^{-1}(x)$. Ponieważ rząd $D\Phi(a) = n$, więc przestrzenią styczną do S w punkcie x jest $T_x S = D\Phi(a)(\mathbb{R}^n)$. Przestrzeń $T_x S$ jest przestrzenią liniową, można więc określić w niej orientację. Mianowicie przez *dodatnią orientację* przestrzeni $T_x S$ rozumiemy orientację wyznaczoną przez układ wektorów (f_1, \dots, f_n) , gdzie $f_i = D\Phi(a)(e_i)$, $i = 1, \dots, n$.

Założmy, że w każdym punkcie $x \in S$ wybraliśmy orientację przestrzeni stycznej $T_x S$. Jeśli $x \in \Phi(\Omega)$ oraz dla dowolnego $y \in \Phi(\Omega)$ orientacja $T_y S$ jest zgodna z orientacją $T_x S$, to mówimy, że rozmaitość S ma wyznaczoną orientację. Rozmaitość dla której można wyznaczyć orientację nazywamy *orientowalną*. Jeśli S jest spójna, to można w niej wyznaczyć co najwyżej dwie orientacje.

Uwaga. Wstęga Möbiusa nie jest orientowalna.

Przypadki szczególne.

1. Zbiór otwarty $U \subset \mathbb{R}^n$ jest n -wymiarową rozmaitością. Ponadto $T_x U = \mathbb{R}^n$ dla $x \in U$. Jeśli w każdym punkcie $x \in U$ wybierzemy dodatnią (ujemną) orientację $T_x U$, to mówimy, że U jest *dodatnio (ujemnie) zorientowany*.

2. Każda krzywa gładka jest orientowalna i jej orientacja wyznacza kierunek obiegu krzywej.

3. Niech $M \subset \mathbb{R}^3$ będzie 2-wymiarową powierzchnią w \mathbb{R}^3 . Wówczas $T_x M = \mathbb{R}^2$ dla $x \in M$. Jeśli (u_x, v_x) jest bazą $T_x M$ zgodną z orientacją M , to wektor

$$n_x = \frac{u_x \times v_x}{\|u_x \times v_x\|}$$

nazywamy wektorem *zewnętrznym normalnym* do M w punkcie x . Wektor $n_x \in \mathbb{R}^3$ jest wektorem o długości 1 prostopadłym do wektorów u_x i v_x oraz trójka (u_x, v_x, n_x) tworzy bazę \mathbb{R}^3 . Mówimy, że (u_x, v_x) zadaje *dodatnią orientację* $T_x M$ jeśli (u_x, v_x, n_x) zadaje dodatnią orientację \mathbb{R}^3 .

Przykład 14.2 Niech $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ będzie funkcją klasy C^1 oraz M wykresem f , tzn.

$$M = \{(x, y, z) \in \mathbb{R}^3 : z = f(x, y) \text{ dla } (x, y) \in \Omega\}.$$

Odwzorowanie

$$\Phi(x, y) = (x, y, f(x, y)) \text{ dla } (x, y) \in \Omega$$

jest mapą płata M . Φ wyznacza dodatnią orientację M za pomocą pary wektorów $(u_{(x,y)}, v_{(x,y)})$, gdzie $u_{(x,y)} = D\Phi(x, y)(e_1)$, $v_{(x,y)} = D\Phi(x, y)(e_2)$. Ponieważ macierzą różniczkową Φ jest

$$D\Phi(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ f'_x(x, y) & f'_y(x, y) \end{pmatrix},$$

więc

$$\begin{aligned} u_{(x,y)} &= D\Phi(x, y)(e_1) = (1, 0, f'_x(x, y)), \\ v_{(x,y)} &= D\Phi(x, y)(e_2) = (0, 1, f'_y(x, y)). \end{aligned}$$

W celu wyznaczenia wektora normalnego zewnętrznego do M liczymy iloczyn wektorowy

$$\begin{aligned} u_{(x,y)} \times v_{(x,y)} &= \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 0 & f'_x(x, y) \\ 0 & 1 & f'_y(x, y) \end{vmatrix} = e_1 \cdot (-f'_x(x, y)) - e_2 \cdot (f'_y(x, y)) + e_3 \\ &= (-f'_x(x, y), -f'_y(x, y), 1). \end{aligned}$$

Zatem

$$n(x, y, f(x, y)) = \frac{(-f'_x(x, y), -f'_y(x, y), 1)}{\sqrt{1 + f'^2_x(x, y) + f'^2_y(x, y)}}.$$

14.3 Orientacja brzegu rozmaitości

Załóżmy, że S jest rozmaitością n -wymiarową oraz jej brzeg ∂S jest rozmaitością $(n-1)$ -wymiarową. Wówczas orientację ∂S można wyznaczyć korzystając z orientacji S . Opiszemy tę procedurę w kilku ważnych przypadkach.

Przypadek 1. Niech $U \subset \mathbb{R}^2$ będzie zbiorem otwartym, którego brzeg $\hat{\gamma} = \partial U$ jest krzywą zamkniętą. Wówczas na $\hat{\gamma}$ mamy orientację dodatnią jeśli podczas obiegu krzywej $\hat{\gamma}$ zbiór U leży po lewej stronie.

Przypadek 2. Niech $M \subset \mathbb{R}^3$ będzie 2-wymiarową powierzchnią z brzegiem, a jej brzeg $\hat{\gamma} = \partial M$ krzywą gładką. Wtedy dla każdego punktu $\dot{x} \in \hat{\gamma} = \partial M$ istnieje zbiór otwarty $\dot{x} \in U \subset M$ oraz homeomorfizm regularny

$$\Phi : K_+ = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1, y \geq 0\} \xrightarrow{na} U$$

taki, że $\Phi(0, 0) = \dot{x}$. Niech orientacja na M będzie zgodna z orientacją wyznaczoną przez mapę Φ . Połóżmy $f_1 = D\Phi(0, 0)(e_1)$, $f_2 = D\Phi(0, 0)(e_2)$. Wówczas para wektorów (f_1, f_2) zadaje orientację w \dot{x} , przy czym wektor f_1 jest styczny do $\hat{\gamma}$ w punkcie \dot{x} . Mówimy, że orientacje M i $\hat{\gamma}$ są zgodne jeśli f_1 wyznacza orientację na γ zgodną z pierwotną orientacją $\hat{\gamma}$.

Przypadek 3. Niech $U \subset \mathbb{R}^3$ będzie zbiorem otwartym, którego brzeg $M = \partial U$ jest 2-wymiarową powierzchnią. Niech orientacja $T_x M$ będzie zadana przez parę wektorów (u_x, v_x) oraz niech

$$n_x = \frac{u_x \times v_x}{\|u_x \times v_x\|}$$

będzie wektorem normalnym do M w punkcie x . Oznaczmy

$$x_\varepsilon = x + \varepsilon n_x.$$

Wówczas mówimy, że M jest *zorientowana dodatnio (ujemnie)* jeśli $x_\varepsilon \notin U$ ($x_\varepsilon \in U$) dla dostatecznie małych $\varepsilon > 0$. Innymi słowami $M = \partial U$ jest zorientowana dodatnio jeśli wektory n_x dla $x \in M$ są skierowane na zewnątrz U .

15 Całka krzywoliniowa

15.1 Krzywe w \mathbb{R}^n

Definicja 15.1 Niech $\gamma : [a, b] \rightarrow \mathbb{R}^n$ będzie funkcją ciągłą i różnowartościową w przedziałach $[a, b]$ i (a, b) . Wówczas zbiór

$$\hat{\gamma} = \{x \in \mathbb{R}^n : x = \gamma(t) \text{ dla pewnego } t \in [a, b]\}$$

jest krzywą zwartą i spójną, a γ jest parametryzacją $\hat{\gamma}$.

Krzywą $\hat{\gamma}$ nazywamy łukiem gładkim jeśli jest ona wyznaczona przez parametryzację $\gamma = (\gamma_1, \dots, \gamma_n) : [a, b] \rightarrow \mathbb{R}^n$ taką, że $\gamma_1, \dots, \gamma_n$ są funkcjami klasy C^1 oraz

$$\gamma_1'^2(t) + \dots + \gamma_n'^2(t) > 0$$

dla każdego $t \in (a, b)$. Orientację krzywej $\hat{\gamma}$ wyznacza wektor

$$\gamma'(t) = (\gamma_1'(t), \dots, \gamma_n'(t)).$$

Będziemy też dopuszczali przypadek, gdy krzywa $\hat{\gamma}$ jest sumą skończonej ilości łuków gładkich. Długością krzywej $\hat{\gamma}$ nazywamy całkę

$$d = d(\hat{\gamma}) = \int_a^b \sqrt{\gamma_1'^2(t) + \dots + \gamma_n'^2(t)} dt$$

15.2 Całka krzywoliniowa zorientowana

Niech $\mathbf{F} = (F_1, F_2, F_3)$ będzie polem sił określonym na obszarze $U \subset \mathbb{R}^3$ oraz $\hat{\gamma}$ krzywą zawartą w U . Naturalnym problemem jest potrzeba znalezienia wielkości pracy potrzebnej do przesunięcia zadanej masy wzdłuż krzywej $\hat{\gamma}$. Do rozwiązania tego problemu służy całka krzywoliniowa zorientowana.

Definicja 15.2 Niech $\gamma : [a, b] \rightarrow \mathbb{R}^n$ będzie parametryzacją krzywej $\hat{\gamma}$ oraz $\mathbf{F} = (F_1, \dots, F_n) : \hat{\gamma} \rightarrow \mathbb{R}^n$ polem wektorowym na $\hat{\gamma}$. Niech $\Pi = \{t_0, \dots, t_k\}$, gdzie $a = t_0 < t_1 < \dots < t_k = b$, będzie podziałem odcinka $[a, b]$. Wybierzmy dowolnie punkty $\theta_i \in [t_{i-1}, t_i]$ dla $i = 1, \dots, k$. Niech $\Delta x_{j,i} = \gamma_j(t_i) - \gamma_j(t_{i-1})$ będzie przyrostem γ_j w przedziale $[t_{i-1}, t_i]$, $j = 1, \dots, n$, $i = 1, \dots, k$. Wówczas sumę

$$\sigma = \sum_{i=1}^k F_1(\gamma(\theta_i)) \cdot \Delta x_{1,i} + \dots + F_n(\gamma(\theta_i)) \cdot \Delta x_{n,i}$$

nazywamy sumą przybliżoną. Następnie analogicznie do konstrukcji całki Riemanna funkcji jednej zmiennej definiujemy ciąg normalny podziałów oraz ciągi aproksymacyjne. Jeśli są one zbieżne do wspólnej granicy, to nazywamy ją całką zorientowaną pola \mathbf{F} wzdłuż krzywej $\hat{\gamma}$ oznaczamy

$$\int_{\hat{\gamma}} \mathbf{F}(x) dx = \int_{\hat{\gamma}} F_1(x) dx_1 + \dots + F_n(x) dx_n.$$

Zauważmy, że jeśli krzywa $\widehat{\gamma}$ jest sumą dwóch krzywych $\widehat{\gamma}_1 + \widehat{\gamma}_2$, to

$$\int_{\widehat{\gamma}} \mathbf{F}(x) dx = \int_{\widehat{\gamma}_1} \mathbf{F}(x) dx + \int_{\widehat{\gamma}_2} \mathbf{F}(x) dx.$$

W szczególności, jeśli $\widehat{\gamma}_2 = -\widehat{\gamma}_1$ tzn. $\gamma_2(t) = \gamma_1(-t)$ dla $t \in [a, b]$, to

$$\int_{-\widehat{\gamma}_1} \mathbf{F}(x) dx = - \int_{\widehat{\gamma}_1} \mathbf{F}(x) dx.$$

Uwaga. Wyrażenie $F_1(x)dx_1 + \dots + F_n(x)dx_n$ nazywane jest *1-formą*.

Okazuje się, że całka krzywoliniowa zorientowana zależy tylko od orientacji krzywej, a nie od jej parametryzacji. Ponadto można ją obliczyć korzystając z całki Riemanna.

Twierdzenie 15.1 Niech $\mathbf{F} = (F_1, \dots, F_n) : \widehat{\gamma} \rightarrow \mathbb{R}^n$ będzie ciągłym polem wektorowym na krzywej gładkiej $\widehat{\gamma}$ oraz niech $\gamma : [a, b] \rightarrow \mathbb{R}^n$ będzie parametryzacją $\widehat{\gamma}$ klasy C^1 . Wówczas całka zorientowana $\int_{\widehat{\gamma}} \mathbf{F} dx$ istnieje oraz

$$\begin{aligned} \int_{\widehat{\gamma}} \mathbf{F} dx &= \int_a^b \mathbf{F}(\gamma(t)) \circ \gamma'(t) dt \\ &= \int_a^b \left(F_1(\gamma(t)) \cdot \gamma'_1(t) + \dots + F_n(\gamma(t)) \cdot \gamma'_n(t) \right) dt. \end{aligned}$$

Ponadto $\int_{\widehat{\gamma}} \mathbf{F} dx$ zależy tylko od orientacji krzywej $\widehat{\gamma}$.

Wniosek 15.1 Jeśli $\widehat{\gamma} = \{(x(t), y(t)) : t \in [a, b]\}$ jest krzywą gładką w \mathbb{R}^2 oraz $\mathbf{F} = (P, Q)$ ciągłym polem wektorowym na $\widehat{\gamma}$, to

$$\int_{\widehat{\gamma}} P(x, y) dx + Q(x, y) dy = \int_a^b \left(P(x(t), y(t)) \cdot x'(t) + Q(x(t), y(t)) \cdot y'(t) \right) dt.$$

Wniosek 15.2 Jeśli $\widehat{\gamma} = \{(x(t), y(t), z(t)) : t \in [a, b]\}$ jest krzywą gładką w \mathbb{R}^3 oraz $\mathbf{F} = (P, Q, R)$ ciągłym polem wektorowym na $\widehat{\gamma}$, to

$$\begin{aligned} \int_{\widehat{\gamma}} P(x, y, z) dx + Q(x, y, z) dy + R(x, y, z) dz \\ = \int_a^b \left(P(x(t), y(t), z(t)) \cdot x'(t) + Q(x(t), y(t), z(t)) \cdot y'(t) + R(x(t), y(t), z(t)) \cdot z'(t) \right) dt. \end{aligned}$$

Przykład 15.1 Niech $\mathbf{F}(x, y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2} \right)$ będzie polem wektorowym na $\mathbb{R}^2 \setminus \{(0, 0)\}$ oraz $\widehat{\gamma} = \{(r \cos t, r \sin t), 0 \leq t \leq 2\pi\}$ okręgiem o promieniu $r > 0$. Wówczas

$$\begin{aligned} \int_{\widehat{\gamma}} \frac{x}{x^2+y^2} dy - \frac{y}{x^2+y^2} dx \\ = \int_0^{2\pi} \left\{ \frac{r \cos t}{(r \cos t)^2 + (r \sin t)^2} \cdot r \cos t + \frac{-r \sin t}{(r \cos t)^2 + (r \sin t)^2} \cdot (-r \sin t) \right\} dt \\ = \int_0^{2\pi} 1 dt = 2\pi. \end{aligned}$$

Zauważmy, że całka ta nie zależy od r .

15.3 Całka krzywoliniowa nieorientowana

Niech $f : \hat{\gamma} \rightarrow \mathbb{R}$ będzie funkcją rzeczywistą określoną na krzywej $\hat{\gamma}$. Jeśli f jest dodatnia, to można ją interpretować jako gęstość masy rozłożonej na krzywej $\hat{\gamma}$ i wówczas interesuje nas problem wyznaczenia masy całej krzywej. W tym celu wprowadzimy pojęcie całki krzywoliniowej nieorientowanej.

Definicja 15.3 Niech $\hat{\gamma} = \{\gamma(t) : t \in [a, b]\}$ będzie krzywą gładką oraz $f : \hat{\gamma} \rightarrow \mathbb{R}$ funkcją rzeczywistą. Niech $\Pi = \{t_0, \dots, t_k\}$, gdzie $a = t_0 < t_1 < \dots < t_k = b$ będzie podziałem odcinka $[a, b]$. Wybierzmy dowolnie punkty $\theta_i \in [t_{i-1}, t_i]$ dla $i = 1, \dots, k$. Niech Δs_i będzie długością łuku $(\gamma(t_{i-1}), \gamma(t_i))$. Wówczas

$$\Delta s_i = \int_{t_{i-1}}^{t_i} \sqrt{\gamma_1'^2(t) + \dots + \gamma_n'^2(t)} dt.$$

Definiujemy *sumą przybliżoną dla podziału Π* :

$$\sigma = \sum_{i=1}^k f(\gamma(\theta_i)) \Delta s_i.$$

Następnie dla normalnego ciągu podziałów (Π_j) oznaczamy przez σ_j ciąg sum przybliżonych. Jeśli granica ciągu σ_j nie zależy od wyboru normalnego ciągu podziałów oraz od wyboru punktów θ_i , to nazywamy ją *całką nieorientowaną* funkcji f po krzywej $\hat{\gamma}$ i oznaczamy

$$\int_{\hat{\gamma}} f(x) ds = \lim_{j \rightarrow \infty} \sigma_j.$$

Twierdzenie 15.2 Niech $f : \hat{\gamma} \rightarrow \mathbb{R}^n$ będzie funkcją ciągłą na krzywej gładkiej $\hat{\gamma}$ oraz niech $\gamma : [a, b] \rightarrow \mathbb{R}^n$ będzie parametryzacją $\hat{\gamma}$ klasy C^1 . Wówczas całka nieorientowana $\int_{\hat{\gamma}} f(x) ds$ istnieje oraz

$$\int_{\hat{\gamma}} f(x) ds = \int_a^b f(\gamma(t)) \sqrt{\gamma_1'^2(t) + \dots + \gamma_n'^2(t)} dt.$$

15.4 Związek całki zorientowanej z całką nieorientowaną

Niech $\hat{\gamma} = \{\gamma(t) : t \in [a, b]\}$ będzie krzywą gładką. Wówczas wektor

$$\gamma'(t) = (\gamma_1'(t), \dots, \gamma_n'(t))$$

jest styczny do $\hat{\gamma}$ i skierowany zgodnie z orientacją $\hat{\gamma}$. Jego długość wynosi

$$\|\gamma'(t)\| = \sqrt{\gamma_1'^2(t) + \dots + \gamma_n'^2(t)}.$$

Niech $\mathbf{F} : \hat{\gamma} \rightarrow \mathbb{R}^n$ będzie polem wektorowym na $\hat{\gamma}$. Oznaczmy przez $\alpha = \alpha(t)$ kąt pomiędzy wektorami $\mathbf{F}(t)$ a $\gamma'(t)$. Wówczas składowa F_s wektora \mathbf{F} styczna do $\hat{\gamma}$ wynosi

$$F_s = \|\mathbf{F}\| \cdot \cos \alpha.$$

Policzmy iloczyn skalarny

$$\mathbf{F}(\gamma(t)) \circ \gamma'(t) = \|\mathbf{F}(\gamma(t))\| \cdot \|\gamma'(t)\| \cdot \cos \alpha = F_s(\gamma(t)) \cdot \|\gamma'(t)\|.$$

Stąd

$$\begin{aligned} \int_{\hat{\gamma}} \mathbf{F}(x) dx &= \int_a^b \mathbf{F}(\gamma(t)) \circ \gamma'(t) dt \\ &= \int_a^b F_s(\gamma(t)) \cdot \|\gamma'(t)\| dt \\ &= \int_a^b F_s(\gamma(t)) \cdot \sqrt{\gamma_1'^2(t) + \dots + \gamma_n'^2(t)} dt \\ &= \int_{\hat{\gamma}} F_s(x) ds. \end{aligned}$$

Czyli

$$\int_{\hat{\gamma}} \mathbf{F}(x) dx = \int_{\hat{\gamma}} F_s(x) ds.$$

Wniosek 15.3 Jeśli pole wektorowe \mathbf{F} jest prostopadłe do krzywej $\hat{\gamma}$, to

$$\int_{\hat{\gamma}} \mathbf{F}(x) dx = 0.$$

15.5 Zastosowania całek krzywoliniowych

1. Jeśli $\mathbf{F} = (F_1, \dots, F_n)$ jest siłą działającą na punkt materialny, to pracę potrzebną do przesunięcia tego punktu wzdłuż krzywej $\hat{\gamma}$ obliczamy ze wzoru

$$\begin{aligned} W &= \int_{\hat{\gamma}} \mathbf{F}(x) dx = \int_{\hat{\gamma}} F_1(x) dx_1 + \dots + F_n(x) dx_n \\ &= \int_{\hat{\gamma}} F_s(x) ds. \end{aligned}$$

W szczególności, jeśli $\mathbf{F} \perp \hat{\gamma}$, tzn. siła \mathbf{F} jest prostopadła do $\hat{\gamma}$, to $W = 0$.

2. Niech $y = f(x)$, $a \leq x \leq b$ będzie nieujemną funkcją klasy C^1 . Wówczas pole powierzchni S powstałej w wyniku obrotu krzywej $\hat{\gamma} = \{(x, f(x)), x \in [a, b]\}$ dookoła osi OX wynosi

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

Ponieważ krzywa $\hat{\gamma}$ ma parametryzację $\hat{\gamma} = \{(x(t) = t, y(t) = f(t)), t \in [a, b]\}$, więc

$$|S| = 2\pi \int_a^b y(t) \sqrt{x'^2(t) + y'^2(t)} dt = 2\pi \int_{\hat{\gamma}} y ds.$$

Przypomnijmy, że środek ciężkości krzywej $\hat{\gamma}$ ma współrzędne (ξ, η) , gdzie $\xi = \frac{1}{|\hat{\gamma}|} \int_{\hat{\gamma}} x ds$, $\eta = \frac{1}{|\hat{\gamma}|} \int_{\hat{\gamma}} y ds$. Zatem otrzymujemy ponownie regułę Guldina

$$|S| = 2\pi \cdot \eta \cdot |\hat{\gamma}|.$$

15.6 Wzór Greena

Wzór Greena wyraża związek pomiędzy całką zorientowaną po krzywej zamkniętej (konturze), a całką podwójną po obszarze wewnątrz tej krzywej.

Twierdzenie 15.3 (Greena). *Niech $D \subset \mathbb{R}^2$ będzie zbiorem regularnym, którego brzeg $\hat{\gamma} = \partial D$ jest krzywą regularną zorientowaną dodatnio. Jeśli funkcje $P, Q : \bar{D} \rightarrow \mathbb{R}$ są klasy C^1 , to*

$$\oint_{\hat{\gamma}} P(x, y) dx + Q(x, y) dy = \iint_D \left(\frac{\partial Q(x, y)}{\partial x} - \frac{\partial P(x, y)}{\partial y} \right) dx dy.$$

Wniosek 15.4 *Niech $D \subset \mathbb{R}^2$ będzie zbiorem regularnym, którego brzeg $\hat{\gamma} = \partial D$ jest krzywą regularną zorientowaną dodatnio. Wówczas*

$$|D| = \oint_{\hat{\gamma}} x dy = - \oint_{\hat{\gamma}} y dx = \frac{1}{2} \oint_{\hat{\gamma}} (x dy - y dx).$$

Wniosek 15.5 *Niech $D \subset \mathbb{R}^2$ będzie zbiorem regularnym, którego brzeg $\hat{\gamma} = \partial D$ jest krzywą regularną zorientowaną dodatnio. Jeśli funkcje $P, Q : \bar{D} \rightarrow \mathbb{R}$ są klasy C^1 i spełniają warunek całkowności*

$$(C) \quad \frac{\partial Q(x, y)}{\partial x} = \frac{\partial P(x, y)}{\partial y} \quad \text{dla } (x, y) \in D,$$

to

$$\oint_{\hat{\gamma}} P(x, y) dx + Q(x, y) dy = 0.$$

Wniosek 15.6 *Niech $D \subset \mathbb{R}^2$ będzie obszarem spójnym i jednospójnym (tzn, zbiór $\mathbb{R}^2 \setminus D$ jest spójny). Jeśli funkcje $P, Q : \bar{D} \rightarrow \mathbb{R}$ są klasy C^1 i spełniają warunek całkowności (C), to dla dowolnej krzywej regularnej $\hat{\gamma} \subset D$ całka*

$$\int_{\hat{\gamma}} P(x, y) dx + Q(x, y) dy$$

zależy tylko od początku i końca tej krzywej.

Przy założeniach powyższego wniosku wybierając dowolnie punkt $(\dot{x}, \dot{y}) \in D$ możemy określić funkcję

$$U(x, y) = - \int_{(\dot{x}, \dot{y})}^{(x, y)} P(\xi, \eta) d\xi + Q(\xi, \eta) d\eta \quad \text{dla } (x, y) \in D.$$

Tak określoną funkcję $U : D \rightarrow \mathbb{R}$ nazywamy *potencjałem pola wektorowego* (P, Q) .
Zauważmy, że

$$\begin{aligned} \frac{\partial U(x, y)}{\partial x} &= -\lim_{h \rightarrow 0} \frac{1}{h} \left(\int_{(\hat{x}, \hat{y})}^{(x+h, y)} P(\xi, \eta) d\xi + Q(\xi, \eta) d\eta - \int_{(\hat{x}, \hat{y})}^{(x, y)} P(\xi, \eta) d\xi + Q(\xi, \eta) d\eta \right) \\ &= -\lim_{h \rightarrow 0} \frac{1}{h} \int_{(x, y)}^{(x+h, y)} P(\xi, \eta) d\xi + Q(\xi, \eta) d\eta \\ &= -\lim_{h \rightarrow 0} \frac{1}{h} \int_0^h P(x+t, y)(t)' dt \\ &= -P(x, y). \end{aligned}$$

Analogicznie dostajemy

$$\frac{\partial U(x, y)}{\partial y} = -Q(x, y).$$

Ponieważ P i Q są klasy C^1 , funkcja U jest klasy C^2 . Zatem jej pochodne mieszane są równe

$$\begin{aligned} \frac{\partial^2 U(x, y)}{\partial x \partial y} &= -\frac{\partial}{\partial x} Q(x, y) \\ &\parallel \\ \frac{\partial^2 U(x, y)}{\partial y \partial x} &= -\frac{\partial}{\partial y} P(x, y). \end{aligned}$$

Co jest zgodne z warunkiem (C).

15.7 Niezależność całki krzywoliniowej od drogi całkowania

Załóżmy teraz, że pole wektorowe $\mathbf{F} : D \rightarrow \mathbb{R}^n$ jest określone na spójnym i jedno-spójnym obszarze $D \subset \mathbb{R}^n$.

Definicja 15.4 Mówimy, że całka krzywoliniowa pola \mathbf{F} *nie zależy od drogi całkowania*, jeśli dla dowolnych krzywych regularnych $\hat{\gamma}_1$ i $\hat{\gamma}_2$ o tych samych końcach zachodzi

$$\int_{\hat{\gamma}_1} \mathbf{F}(x) dx = \int_{\hat{\gamma}_2} \mathbf{F}(x) dx.$$

Wówczas pole \mathbf{F} nazywamy *potencjalnym*, a jego *potencjał* wynosi

$$U(x) = \int_{\hat{x}}^x \mathbf{F}(\xi) d\xi.$$

Analogicznie jak w przypadku $n = 2$ wykazuje się, że

$$\frac{\partial U(x)}{\partial x_i} = -F_i(x) \quad \text{dla } i = 1, \dots, n.$$

Jeśli pole \mathbf{F} jest klasy C^1 , to potencjał U jest klasy C^2 oraz zachodzi

$$\begin{aligned} \frac{\partial^2 U(x)}{\partial x_j \partial x_i} &= -\frac{\partial}{\partial x_j} F_i(x) \\ &\parallel \\ \frac{\partial^2 U(x)}{\partial x_i \partial x_j} &= -\frac{\partial}{\partial x_i} F_j(x). \end{aligned}$$

Zatem pole \mathbf{F} spełnia warunek całkowalności

$$(C_n) \quad \frac{\partial}{\partial x_j} F_i(x) = \frac{\partial}{\partial x_i} F_j(x) \quad \text{dla } i, j = 1, \dots, n.$$

Uwaga. W przypadku pola $\mathbf{F} = [P, Q, R] : D \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ warunek całkowalności przyjmuje postać:

$$\begin{cases} \frac{\partial P(x,y,z)}{\partial y} = \frac{\partial Q(x,y,z)}{\partial x}, \\ \frac{\partial Q(x,y,z)}{\partial z} = \frac{\partial R(x,y,z)}{\partial y}, \\ \frac{\partial R(x,y,z)}{\partial x} = \frac{\partial P(x,y,z)}{\partial z} \end{cases} \quad \text{dla } (x, y, z) \in D.$$

Wykażemy później, że jeśli pole \mathbf{F} spełnia warunek całkowalności w obszarze jednospójnym, to jest ono potencjalne.

15.8 Całka krzywoliniowa funkcji o wartościach zespolonych

Niech $D \subset \mathbb{C} \simeq \mathbb{R}^2$ będzie obszarem na płaszczyźnie zespolonej oraz $\hat{\gamma}$ krzywą regularną w D . Niech $f : D \rightarrow \mathbb{C}$ będzie funkcją ciągłą o wartościach zespolonych. Oznaczmy $z = x + iy$, $f = u + iv$. Wówczas funkcje $u, v : D \rightarrow \mathbb{R}$ można traktować jako funkcje rzeczywiste na zbiorze $D \subset \mathbb{R}^2$. Definiujemy

$$\begin{aligned} \int_{\hat{\gamma}} f(z) dz &= \int_{\hat{\gamma}} (u(x, y) + iv(x, y)) (dx + idy) \\ &= \int_{\hat{\gamma}} (u(x, y) dx - v(x, y) dy) + i \int_{\hat{\gamma}} (v(x, y) dx + u(x, y) dy). \end{aligned}$$

Zauważmy, że warunek (C) niezależności całki od drogi całkowania przyjmuje postać

$$(CR) \quad \begin{cases} \frac{\partial u(x, y)}{\partial x} = \frac{\partial v(x, y)}{\partial y}, \\ \frac{\partial u(x, y)}{\partial y} = -\frac{\partial v(x, y)}{\partial x} \end{cases} \quad \text{dla } (x, y) \in D.$$

Równania (CR) nazywają się *układem Cauchy-Riemanna*.

Uwaga. Funkcja $f(z) = \frac{1}{z}$ spełnia równania (CR). Tym nie mniej $\int_{\partial B(0,R)} \frac{1}{z} dz = 2\pi i$. Zatem całka po konturze zamkniętym nie jest równa zero. Wynika to z faktu, że zbiór $B(0, R) \setminus \{0\}$ nie jest jednospójny.

16 Całki powierzchniowe

16.1 Całka powierzchniowa nieorientowana

Niech $D \subset \mathbb{R}^2$ będzie zbiorem regularnym oraz $f : D \rightarrow \mathbb{R}$ funkcją klasy C^1 . Wówczas wykres funkcji f czyli zbiór

$$S = \{(x, y, z) \in \mathbb{R}^3 : z = f(x, y), \quad (x, y) \in D\}$$

jest płatem regularnym. Jak wiemy pole płata S wyraża się wzorem

$$|S| = \iint_D \sqrt{1 + \left(\frac{\partial f(x, y)}{\partial x}\right)^2 + \left(\frac{\partial f(x, y)}{\partial y}\right)^2} dx dy.$$

Niech $F : S \rightarrow \mathbb{R}$ będzie funkcją ograniczoną. Podzielmy D na k zbiorów regularnych, tzn.

$$D = \bigcup_{i=1}^k D_i \quad \text{gdzie} \quad \text{int} D_i \cap \text{int} D_j = \emptyset \quad \text{dla} \quad i \neq j.$$

Oznaczmy przez Δ_i część wykresu funkcji f nad zbiorem D_i oraz przez $|\Delta S_i|$ pole ΔS_i , $i = 1, \dots, k$. Wybierzmy dowolnie punkt $A_i \in \Delta S_i$, $i = 1, \dots, k$ i utwórzmy sumę przybliżoną

$$\sigma = \sum_{i=1}^k F(A_i) |\Delta S_i|.$$

Następnie tworzymy normalny ciąg podziałów zbioru D , odpowiadający mu ciąg podziałów zbioru S oraz ciąg sum przybliżonych. Jeżeli wszystkie ciągi przybliżone są zbieżne do tej samej granicy to granicę tę nazywamy *całką powierzchniową nieorientowaną* funkcji F po płacie S i oznaczamy

$$\iint_S F(x, y, z) dS.$$

Zachodzi

Twierdzenie 16.1 Niech $f : D \rightarrow \mathbb{R}$ będzie funkcją klasy C^1 oraz

$$S = \{(x, y, z) \in \mathbb{R}^3 : z = f(x, y) \quad \text{dla} \quad (x, y) \in D\}$$

będzie płatem regularnym. Jeśli $F : S \rightarrow \mathbb{R}$ jest funkcją ciągłą, to

$$\iint_S F(x, y, z) dS = \iint_D F(x, y, f(x, y)) \sqrt{1 + f_x'^2(x, y) + f_y'^2(x, y)} dx dy.$$

Analogicznie definiuje się całkę powierzchniową nieorientowaną, w przypadku gdy płat S jest określony przez $S = \{(x, y, z) : y = f(x, z)\}$ lub $S = \{(x, y, z) : x = f(y, z)\}$.

W ogólnym przypadku zbiór $S \subset \mathbb{R}^3$ jest dwuwymiarowym płatem regularnym, jeśli istnieje zbiór regularny $D \subset \mathbb{R}^2$ i dyfeomorfizm $\Phi : D \rightarrow \mathbb{R}^3$ taki, że $\Phi(D) = S$. Jeśli $\Phi = (\varphi, \psi, \chi)$, gdzie $\varphi, \psi, \chi : D \rightarrow \mathbb{R}$ są funkcjami klasy C^1 , to zachodzi

Twierdzenie 16.2 Niech $S \subset \mathbb{R}^3$ będzie płatem dwuwymiarowym regularnym oraz $F : S \rightarrow \mathbb{R}$ funkcją ciągłą. Wówczas całka powierzchniowa niezorientowana funkcji F po płacie S wyraża się wzorem

$$\begin{aligned} & \iint_S F(x, y, z) dS \\ &= \iint_D F(\varphi(u, v), \psi(u, v), \chi(u, v)) \sqrt{J_1^2(u, v) + J_2^2(u, v) + J_3^2(u, v)} du dv, \end{aligned}$$

gdzie

$$J_1(u, v) = \begin{vmatrix} \frac{\partial \varphi}{\partial u} & \frac{\partial \varphi}{\partial v} \\ \frac{\partial \psi}{\partial u} & \frac{\partial \psi}{\partial v} \end{vmatrix} = C, \quad J_2(u, v) = \begin{vmatrix} \frac{\partial \varphi}{\partial u} & \frac{\partial \varphi}{\partial v} \\ \frac{\partial \chi}{\partial u} & \frac{\partial \chi}{\partial v} \end{vmatrix} = -B, \quad J_3(u, v) = \begin{vmatrix} \frac{\partial \psi}{\partial u} & \frac{\partial \psi}{\partial v} \\ \frac{\partial \chi}{\partial u} & \frac{\partial \chi}{\partial v} \end{vmatrix} = A.$$

Uwaga. Przypomnijmy, że macierzą różniczki odwzorowania Φ jest macierz

$$\begin{pmatrix} \frac{\partial \varphi}{\partial u} & \frac{\partial \varphi}{\partial v} \\ \frac{\partial \psi}{\partial u} & \frac{\partial \psi}{\partial v} \\ \frac{\partial \chi}{\partial u} & \frac{\partial \chi}{\partial v} \end{pmatrix}.$$

Jeśli $S \subset \mathbb{R}^3$ jest dwuwymiarową rozmaitością (powierzchnią) taką, że można ją rozłożyć na sumę płatów regularnych $S = S_1 \cup \dots \cup S_k$ o rozłącznych wnętrzach $\text{int}S_i \cap \text{int}S_j = \emptyset$ dla $i, j = 1, \dots, k$, $i \neq j$, to S nazywamy powierzchnią regularną i przez całkę po S rozumiemy sumę całek po płatach S_i , $i = 1, \dots, k$.

Interpretacja fizyczna

1. Jeśli na powierzchni S jest rozłożony ładunek elektryczny o gęstości $F(x, y, z)$, to całka $\iint_S F(x, y, z) dS$ wyraża całkowity ładunek na S .
2. Jeśli na powierzchni S jest rozłożona jest materia o gęstości $\rho(x, y, z)$, to całka $\iint_S \rho(x, y, z) dS$ wyraża całkowitą masę na S .
3. Jeśli na powierzchni S jest rozłożona jest materia o gęstości $\rho(x, y, z)$, to punkt O o masie jednostkowej jest przyciągany przez S siłą $\mathbf{F} = (F_x, F_y, F_z)$, gdzie

$$\begin{aligned} F_x &= \iint_S \rho(x, y, z) \frac{x}{(x^2 + y^2 + z^2)^{3/2}} dS, \\ F_y &= \iint_S \rho(x, y, z) \frac{y}{(x^2 + y^2 + z^2)^{3/2}} dS, \\ F_z &= \iint_S \rho(x, y, z) \frac{z}{(x^2 + y^2 + z^2)^{3/2}} dS. \end{aligned}$$

Przykład 16.1 Niech $S = \{x^2 + y^2 + z^2 = r^2, z \geq 0\}$ będzie jednorodną półsferą o stałej gęstości $\rho > 0$. Wówczas punkt O o masie jednostkowej jest przyciągany siłą $\mathbf{F} = (F_x, F_y, F_z)$. Wobec symetrii jest jasne, że $F_x = F_y = 0$. Natomiast

$$F_z = \iint_S \rho \cdot \frac{z}{(x^2 + y^2 + z^2)^{3/2}} dS.$$

W celu obliczenia tej całki zastosujemy współrzędne sferyczne. Otóż $S = \Phi(D)$, gdzie $D = [0 \leq \varphi \leq 2\pi] \times [0 \leq \psi \leq \pi/2]$ oraz

$$\Phi(\varphi, \psi) = (r \cos \varphi \cos \psi, r \sin \varphi \cos \psi, r \sin \psi) = (x, y, z).$$

Macierzą różniczki odwzorowania Φ jest macierz

$$\begin{pmatrix} -r \sin \varphi \cos \psi & -r \cos \varphi \sin \psi \\ r \cos \varphi \cos \psi & -\sin \varphi \sin \psi \\ 0 & r \cos \psi \end{pmatrix}.$$

Liczmy

$$J_1^2 + J_2^2 + J_3^2 = r^4 \sin^2 \psi \cos^2 \psi + r^4 \sin^2 \varphi \cos^4 \psi + r^4 \cos^2 \varphi \cos^4 \psi = r^4 \cos^2 \psi.$$

Zatem

$$\begin{aligned} F_z &= \rho \int_0^{2\pi} \int_0^{\pi/2} \frac{r \sin \psi}{r^3} \cdot r^2 \cos \psi d\psi d\varphi \\ &= 2\pi\rho \int_0^{\pi/2} \sin \psi \cos \psi d\psi = \pi\rho. \end{aligned}$$

Zauważmy że siła nie zależy od promienia półsfery (czy potrafisz to wytłumaczyć bez rachunków).

Stwierdzenie 16.1 Niech F będzie funkcją ciągłą w kuli domkniętej $\bar{B}(R)$. Wówczas

$$\iiint_{B(R)} F(x, y, z) dx dy dz = \int_0^R \left(\iint_{S_r} F(x, y, z) dS \right) dr,$$

gdzie $S_r = \partial B(r)$ jest sferą o promieniu r .

Dowód. Korzystając ze współrzędnych sferycznych i wykorzystując rachunki z Przykładu 16.1 dostajemy

$$\iint_{S_r} F(x, y, z) dS = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} G(r, \varphi, \psi) \cdot r^2 \cos \psi d\psi d\varphi,$$

gdzie

$$G(r, \varphi, \psi) = F(r \cos \varphi \cos \psi, r \sin \varphi \cos \psi, r \sin \psi).$$

Z drugiej strony

$$\iiint_{B(R)} F(x, y, z) dx dy dz = \int_0^R \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} G(r, \varphi, \psi) \cdot r^2 \cos \psi d\psi d\varphi dr.$$

Zatem na podstawie twierdzenia Fubinięgo dostajemy wyprowadzany wzór. \square

16.2 Całka powierzchniowa zorientowana

Motywacja.

Załóżmy, że przez powierzchnię $S \subset \mathbb{R}^3$ przepływa ciecz (gaz, strumień pola,...). W punkcie $(x, y, z) \in S$ prędkość przepływu cieczy wynosi $v(x, y, z) \in \mathbb{R}^3$. Oznaczmy przez n wektor normalny, zewnętrzny do S . Wówczas składowa prędkości prostopadła do S jest dana przez iloczyn skalarny $v \circ n$, a całkowita objętość przepływu cieczy przez S w jednostce czasu wynosi

$$V = \iint_S v \circ n \, dS.$$

Definicja 16.1 Niech $S \subset \mathbb{R}^3$ będzie płatem regularnym zorientowanym, a $n(x, y, z) \in \mathbb{R}^3$ wektorem normalnym zewnętrznym do S . Niech $\mathbf{F} : \rightarrow \mathbb{R}^3$ będzie ciągłym polem wektorowym na S . Przez *całkę powierzchniową zorientowaną pola \mathbf{F}* (strumieniem pola \mathbf{F}) przez powierzchnię S nazywamy całkę

$$\iint_S \mathbf{F}(x, y, z) \circ n(x, y, z) \, dS.$$

Załóżmy, że płat S jest zadany jako obraz zbioru otwartego $D \subset \mathbb{R}^2$ pod działaniem dyfeomorfizmu $\Phi = (\varphi, \psi, \chi) : D \rightarrow \mathbb{R}^3$, tzn.

$$S = \{(x, y, z) \in \mathbb{R}^3 : x = \varphi(u, v), y = \psi(u, v), z = \chi(u, v) \text{ dla } (u, v) \in D \subset \mathbb{R}^2\}.$$

Wówczas macierzą różniczki $D\Phi$ jest

$$\begin{pmatrix} \varphi'_u & \varphi'_v \\ \psi'_u & \psi'_v \\ \chi'_u & \chi'_v \end{pmatrix}.$$

Zatem przestrzeń styczna do S jest rozpięta na wektorach $[\varphi'_u, \psi'_u, \chi'_u] = f_1$ i $[\varphi'_v, \psi'_v, \chi'_v] = f_2$. Wektorem normalnym zewnętrznym jest

$$n = \frac{f_1 \times f_2}{\|f_1 \times f_2\|}.$$

Mamy

$$\begin{aligned} f_1 \times f_2 &= \begin{vmatrix} e_1 & e_2 & e_3 \\ \varphi'_u & \psi'_u & \chi'_u \\ \varphi'_v & \psi'_v & \chi'_v \end{vmatrix} = e_1 \cdot (\psi'_u \chi'_v - \psi'_v \chi'_u) - e_2 \cdot (\varphi'_u \chi'_v - \varphi'_v \chi'_u) + e_3 \cdot (\varphi'_u \psi'_v - \varphi'_v \psi'_u) \\ &= [A, B, C], \\ n &= \frac{[A, B, C]}{\sqrt{A^2 + B^2 + C^2}}. \end{aligned}$$

Ponadto

$$dS = \sqrt{A^2 + B^2 + C^2} \, du \, dv.$$

Zatem jeśli $\mathbf{F} = [P, Q, R]$, to

$$\begin{aligned} & \iint_S \mathbf{F}(x, y, z) \circ n(x, y, z) dS \\ &= \iint_D [P, Q, R] \circ \frac{[A, B, C]}{\sqrt{A^2 + B^2 + C^2}} \sqrt{A^2 + B^2 + C^2} dudv \\ &= \iint_D (PA + QB + RC) dudv. \end{aligned}$$

Uwaga. Całkę powierzchniową zorientowaną pola $\mathbf{F} = [P, Q, R]$ oznacza się również w postaci

$$\iint_S P dydz + Q dzdx + R dxdy.$$

Wyrażenie $P dydz + Q dzdx + R dxdy$ nazywa się *2-formą*. Wówczas jeśli

$$S = \{(x, y, z) \in \mathbb{R}^3 : x = x(u, v), y = y(u, v), z = z(u, v) \text{ dla } (u, v) \in D \subset \mathbb{R}^2\},$$

to

$$\begin{aligned} dydz &= \begin{vmatrix} y'_u & y'_v \\ z'_u & z'_v \end{vmatrix} dudv = A dudv, \\ dzdx &= \begin{vmatrix} z'_u & z'_v \\ x'_u & x'_v \end{vmatrix} dudv = B dudv, \\ dxdy &= \begin{vmatrix} x'_u & x'_v \\ y'_u & y'_v \end{vmatrix} dudv = C dudv. \end{aligned}$$

Jeśli powierzchnia S jest wykresem funkcji $f : D \rightarrow \mathbb{R}$, tzn.

$$S = \{(x, y, z) \in \mathbb{R}^3 : z = f(x, y) \text{ dla } (x, y) \in D \subset \mathbb{R}^2\},$$

to

$$n(x, y, z) = \frac{[-f'_x(x, y), -f'_y(x, y), 1]}{\sqrt{1 + f'^2_x(x, y) + f'^2_y(x, y)}}.$$

Zatem

$$\begin{aligned} & \iint_S P dydz + Q dzdx + R dxdy \\ &= \iint_D \left(-P(x, y, f(x, y))f'_x(x, y) - Q(x, y, f(x, y))f'_y(x, y) + R(x, y, f(x, y)) \right) dxdy. \end{aligned}$$

Przykład 16.2 Policzmy całkę zorientowaną

$$\iint_S x dydz + y dzdx + z dxdy,$$

gdzie S jest sferą $\{x^2 + y^2 + z^2 = R^2\}$ zorientowaną dodatnio. W tym celu stosujemy współrzędne sferyczne $x = R \cos \varphi \cos \psi$, $y = R \sin \varphi \cos \psi$, $z = R \sin \psi$, gdzie $0 \leq \varphi \leq 2\pi$, $-\pi/2 \leq \psi \leq \pi/2$. Wówczas

$$\begin{aligned} dydz &= \begin{vmatrix} R \cos \varphi \cos \psi & -R \sin \varphi \sin \psi \\ 0 & R \cos \psi \end{vmatrix} d\varphi d\psi = R^2 \cos \varphi \cos^2 \psi d\varphi d\psi, \\ dzdx &= \begin{vmatrix} 0 & R \cos \psi \\ -R \sin \varphi \cos \psi & -R \cos \varphi \sin \psi \end{vmatrix} d\varphi d\psi = R^2 \sin \varphi \cos^2 \psi d\varphi d\psi, \\ dxdy &= \begin{vmatrix} -R \sin \varphi \cos \psi & -R \cos \varphi \sin \psi \\ R \cos \varphi \cos \psi & -R \sin \varphi \sin \psi \end{vmatrix} d\varphi d\psi = R^2 \sin \psi \cos \psi d\varphi d\psi. \end{aligned}$$

Zatem

$$\begin{aligned} &\iint_S x dydz + y dzdx + z dxdy \\ &= \iint_{(0,2\pi) \times (-\pi/2, \pi/2)} \left(R \cos \varphi \cos \psi \cdot R^2 \cos \varphi \cos^2 \psi + R \sin \varphi \cos \psi \cdot R^2 \sin \varphi \cos^2 \psi \right. \\ &\quad \left. + R \sin \psi \cdot R^2 \sin \psi \cos \psi \right) d\varphi d\psi \\ &= \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} R^3 \cos \psi d\psi d\varphi = 4\pi R^3. \end{aligned}$$

16.3 Twierdzenie Gaussa-Ostrogradskiego

Następna twierdzenie wyraża związek pomiędzy całką powierzchniową zorientowaną a całką potrójną.

Twierdzenie 16.3 (Gaussa-Ostrogradskiego). *Niech $V \subset \mathbb{R}^3$ będzie zbiorem regularnym, którego brzeg $S = \partial V$ jest powierzchnią regularną zorientowaną dodatnio (zewnątrznie). Jeśli pole wektorowe $\mathbf{F} = (P, Q, R) : V \rightarrow \mathbb{R}^3$ jest klasy C^1 , to*

$$(GO) \quad \iint_S P dydz + Q dzdx + R dxdy = \iiint_V \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dxdydz.$$

Wniosek 16.1 *Jeśli $V \subset \mathbb{R}^3$ będzie zbiorem regularnym, którego brzeg $S = \partial V$ jest powierzchnią regularną zorientowaną dodatnio, to jego objętość wynosi*

$$|V| = \frac{1}{3} \iint_S x dydz + y dzdx + z dxdy.$$

Niech $\mathbf{F} = (P, Q, R)$ będzie polem wektorowym na V klasy C^1 . Wyrażenie

$$\operatorname{div} \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} : V \rightarrow \mathbb{R}$$

nazywamy *dywergencją pola* \mathbf{F} . Zatem twierdzenie Gaussa-Ostrogradskiego mówi, że strumień pola \mathbf{F} przez powierzchnię $S = \partial V$ jest równy całce objętościowej dywergencji pola \mathbf{F} . W szczególności jeśli $\operatorname{div} \mathbf{F} = 0$ w V , to strumień pola \mathbf{F} przez S wynosi zero. Takie pole \mathbf{F} nazywa się *beźródłowe*. Wzór Gaussa-Ostrogradskiego można też zapisać w postaci

$$\iiint_V \operatorname{div} \mathbf{F} \, dx dy dz = \iint_S \mathbf{F} \circ n \, dS,$$

gdzie $\mathbf{F} \circ n = \mathbf{F}_n$ jest składową normalną pola \mathbf{F} .

Przykład 16.3 Sprawdźmy wzór Gaussa-Ostrogradskiego pola $\mathbf{F} = (x^2, y^2, z^2)$ i kuli $V = B(R)$. Korzystając z rachunków przeprowadzonych w Przykładzie 16.2 dostajemy dla $S = \partial B(R)$

$$\begin{aligned} \iint_S \mathbf{F} \circ n \, dS &= \iint_S x^2 \, dy dz + y^2 \, dz dx + z^2 \, dx dy \\ &= \iint_{(0,2\pi) \times (-\pi/2, \pi/2)} \left((R \cos \varphi \cos \psi)^2 R^2 \cos \varphi \cos^2 \psi + (R \sin \varphi \cos \psi)^2 R^2 \sin \varphi \cos^2 \psi \right. \\ &\quad \left. + (R \sin \psi)^2 R^2 \sin \psi \cos \psi \right) d\varphi d\psi \\ &= R^4 \left\{ \int_0^{2\pi} \cos^3 \varphi \, d\varphi \int_{-\pi/2}^{\pi/2} \cos^4 \psi \, d\psi + \int_0^{2\pi} \sin^3 \varphi \, d\varphi \int_{-\pi/2}^{\pi/2} \cos^4 \psi \, d\psi \right. \\ &\quad \left. + \int_0^{2\pi} 1 \, d\varphi \int_{-\pi/2}^{\pi/2} \sin^3 \psi \cos \psi \, d\psi \right\} = 0. \end{aligned}$$

Z drugiej strony $\operatorname{div} \mathbf{F} = 2(x + y + z)$ jest funkcją nieparzystą względem płaszczyzny $x + y + z = 0$. Zatem

$$\iiint_{B(R)} 2(x + y + z) \, dx dy dz = 0. \quad \square$$

Niech $u : \Omega \rightarrow \mathbb{R}$ będzie funkcją klasy C^2 na obszarze $\Omega \subset \mathbb{R}^3$. Załóżmy, że dla każdego regularnego obszaru $V \subset \Omega$ strumień pola $\frac{\partial u}{\partial n} = \operatorname{gradu} \circ n = \left[\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z} \right] \circ n$ przez powierzchnię $S = \partial V$ znika. Wówczas ze wzoru (GO) dostajemy

$$0 = \iint_S \frac{\partial u}{\partial n} \, dS = \iint_S \operatorname{gradu} \circ n \, dS = \iiint_V \operatorname{div} \operatorname{gradu} \, dx dy dz.$$

Ponieważ powyższa równość zachodzi dla każdego regularnego obszaru $V \subset \mathbb{R}^3$, więc musi być $\operatorname{div} \operatorname{gradu} = 0$. Lecz

$$\operatorname{div} \operatorname{gradu} = \operatorname{div} \left[\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z} \right] = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} =: \Delta u.$$

Czyli funkcja u spełnia równanie Laplace'a $\Delta u = 0$. Funkcje spełniające równanie Laplace'a nazywamy harmonicznymi. Odwrotnie jeśli $u \in C^2$ jest funkcją harmoniczną, to $\iint_S \frac{\partial u}{\partial n} dS = 0$ dla każdej regularnej zamkniętej powierzchni S .

16.4 Twierdzenie Stokes'a

Twierdzenie Stokes'a wyraża związek pomiędzy całką zorientowaną po powierzchni $S \subset \mathbb{R}^3$ i całką krzywoliniową zorientowaną po brzegu $\hat{\gamma} = \partial S$ powierzchni S .

Twierdzenie 16.4 (Stokes'a.) *Niech krzywa regularna $\hat{\gamma}$ będzie brzegiem 2-wymiarowego płata regularnego $S \subset \mathbb{R}^3$, przy czym $\hat{\gamma}$ i S są zgodnie zorientowane. Jeśli pole wektorowe $\mathbf{F} = [P, Q, R]$ jest klasy C^1 w pewnym otoczeniu płata S , to*

$$\begin{aligned} & \oint_{\hat{\gamma}} P dx + Q dy + R dz \\ &= \iint_S \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) dydz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) dzdx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy. \end{aligned}$$

Uwaga. Zauważmy, że jeśli $S \subset \mathbb{R}^2$, to $R dz = 0$ i wzór Stokes'a redukuje się do wzoru Greena

$$\oint_{\hat{\gamma}} P dx + Q dy = \iint_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy.$$

Jeśli $\mathbf{F} = [P, Q, R]$ jest polem wektorowym klasy C^1 , to pole

$$\left[\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right] = \text{rot} \mathbf{F}$$

nazywa się *rotacją* lub *wirowością pola* \mathbf{F} . Jeśli \mathbf{F}_s jest składową pola \mathbf{F} styczną do $\hat{\gamma}$, to wzór Stokes'a można napisać w postaci

$$\oint_{\hat{\gamma}} \mathbf{F}_s ds = \iint_S (\text{rot} \mathbf{F}) \circ n dS.$$

Przykład 16.4

Zauważmy, że jeśli pole $\mathbf{F} : \Omega \rightarrow \mathbb{R}^3$ jest bezwirowe, tzn. $\text{rot} \mathbf{F} = 0$, to całka krzywoliniowa

$$\int_{\hat{\gamma}} P dx + Q dy + R dz$$

zależy tylko od początku i końca krzywej $\hat{\gamma}$ (o ile obszar Ω jest jednospójny). Zatem ustalając dowolnie punkt początkowy krzywej $\hat{\gamma}$ całka ta definiuje funkcję $U(x, y, z)$ końca krzywej $\hat{\gamma}$. Tak zdefiniowana funkcja U nazywana jest potencjałem pola \mathbf{F} .

Odwrotnie dla danej funkcji $U \in C^2(\Omega)$ jej gradient $\text{grad} U = \left(\frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}, \frac{\partial U}{\partial z} \right)$ wyznacza pole wektorowe $[P, Q, R] = \left[\frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}, \frac{\partial U}{\partial z} \right]$, którego rotacja znika. Mamy więc

Wniosek 16.2 Aby pole wektorowe $\mathbf{F} = [P, Q, R]$ zdefiniowane w obszarze jednorodnym $\Omega \subset \mathbb{R}^3$ było polem potencjalnym potrzeba i wystarcza, aby $\text{rot}\mathbf{F} = 0$. Wówczas praca takiego pola wzdłuż dowolnej krzywej zamkniętej wynosi 0.

Przykład 16.5 Niech $\Omega = \mathbb{R}^3 \setminus \{0\}$. (Jest to zbiór jednorodny w \mathbb{R}^3 .) Niech \mathbf{F} będzie polem sił wytworzonym przez siłę grawitacji masy jednostkowej położonej w początku układu współrzędnych. Wówczas

$$\mathbf{F} = \left(\frac{-x}{r^3}, \frac{-y}{r^3}, \frac{-z}{r^3} \right), \quad \text{gdzie } r = \sqrt{x^2 + y^2 + z^2}, \quad (|\mathbf{F}| = \frac{1}{r^2}).$$

Mamy

$$\begin{aligned} \frac{\partial Q}{\partial x} &= \frac{\partial}{\partial x} \left(\frac{-y}{(x^2 + y^2 + z^2)^{3/2}} \right) = \frac{3xy}{(x^2 + y^2 + z^2)^{5/2}} = \frac{\partial P}{\partial y}, \\ \frac{\partial R}{\partial y} &= \frac{3yz}{(x^2 + y^2 + z^2)^{5/2}} = \frac{\partial Q}{\partial z}, \quad \frac{\partial P}{\partial z} = \frac{3xz}{(x^2 + y^2 + z^2)^{5/2}} = \frac{\partial R}{\partial x}. \end{aligned}$$

Zatem pole \mathbf{F} jest bezwirowe. Aby znaleźć potencjał U pola \mathbf{F} należy znaleźć funkcję $U(x, y, z)$ taką, że

$$\frac{\partial U(x, y, z)}{\partial x} = \frac{-x}{r^3}, \quad \frac{\partial U(x, y, z)}{\partial y} = \frac{-y}{r^3}, \quad \frac{\partial U(x, y, z)}{\partial z} = \frac{-z}{r^3}.$$

Ustalmy $(\hat{x}, \hat{y}, \hat{z}) \in \mathbb{R}^3 \setminus \{0\}$, np. $(\hat{x}, \hat{y}, \hat{z}) = (0, 0, 1)$. Z pierwszego z tych równań dostajemy

$$U(x, y, z) = \int_0^x \frac{-\xi}{(\xi^2 + y^2 + z^2)^{3/2}} d\xi + V(y, z) = \frac{1}{(x^2 + y^2 + z^2)^{1/2}} - \frac{1}{(y^2 + z^2)^{1/2}} + V(y, z).$$

Następnie z drugiego równania mamy

$$\frac{\partial U(x, y, z)}{\partial y} = \frac{-y}{(x^2 + y^2 + z^2)^{3/2}} - \frac{-y}{(y^2 + z^2)^{3/2}} + \frac{\partial V(y, z)}{\partial y} = \frac{-y}{(x^2 + y^2 + z^2)^{3/2}}.$$

Zatem

$$\frac{\partial V(y, z)}{\partial y} = \frac{-y}{(y^2 + z^2)^{3/2}}.$$

Stąd

$$V(y, z) = \int_0^y \frac{-\eta}{(\eta^2 + z^2)^{3/2}} d\eta + W(z) = \frac{1}{(y^2 + z^2)^{1/2}} - \frac{1}{(z^2)^{1/2}} + W(z).$$

Teraz korzystając z trzeciego równania dostajemy

$$\frac{\partial U(x, y, z)}{\partial z} = \frac{-z}{(x^2 + y^2 + z^2)^{3/2}} - \frac{-z}{(y^2 + z^2)^{3/2}} + \frac{-z}{(y^2 + z^2)^{3/2}} - \frac{-z}{(z^2)^{3/2}} + \frac{\partial W(z)}{\partial z} = \frac{-z}{(x^2 + y^2 + z^2)^{3/2}}.$$

Zatem

$$\frac{\partial W(z)}{\partial z} = \frac{-z}{(z^2)^{3/2}}.$$

Stąd

$$W(z) = \int_1^z \frac{-\zeta}{(\zeta^2)^{3/2}} d\zeta = \frac{1}{(z^2)^{1/2}} - 1.$$

Ostatecznie $U(x, y, z) = (x^2 + y^2 + z^2)^{-1/2} - 1$. \square