Rozdział 5

FUNKCJE WIELU ZMIENNYCH

Istniejące podręczniki

H. Margenau, G.M. Murphy: *Matematyka w fizyce i chemii*, rozdział 4, PWN, Warszawa 1962;
W. Kołodziej: *Analiza matematyczna*, rozdział 4, PWN, Warszawa 1978;
J.E. Marsden, A.J. Tromba *Vector Calculus*, Freeman and Company, New York 1988, rozdział 4.
Ryszard Rudnicki *Wykłady z analizy matematycznej*, PWN, Warszawa 2001, § 4.3 i 4.4.

5.1 Przestrzeń $\mathbb{R}^n$

Zbiór $\mathbb{R}^n$ jest **iloczynem kartezjańskim** n kopii zbioru $\mathbb{R}$, co się oznacza $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}$. Jego elementami są **ciągi n-elementowe** liczb rzeczywistych:

$$\mathbf{x} \in \mathbb{R}^n, \quad \mathbf{x} = \{x_1, \ldots, x_n\}, \quad x_k \in \mathbb{R}$$

zwane **wektorami** lub **punktami**. Umiemy je sobie wyobrazić na płaszczyźnie (wtedy $n = 2$) lub w przestrzeni trójwymiarowej (wtedy $n = 3$) jako odcinki zaczepione w jednym wspólnym punkcie, który nazywamy **początkiem układu współrzędnych**. Wtedy wektorem jest cały odcinek, a punktem jest drugi koniec odcinka (jeśli pierwszym końcem jest ten zaczepiony w początku układu).

W zbiorze $\mathbb{R}^n$ jest określone **dodawanie**: jeśli $\mathbf{x} = \{x_1, \ldots, x_n\}$, $\mathbf{y} = \{y_1, \ldots, y_n\}$, to

$$\mathbf{x} + \mathbf{y} = \{x_1 + y_1, \ldots, x_n + y_n\} \tag{5.1}$$

oraz **mnożenie przez skalary** $\alpha \in \mathbb{R}$:

$$\alpha\mathbf{x} = \{\alpha x_1, \ldots, \alpha x_n\}. \tag{5.2}$$

Jeśli te dwa działania spełniają pewne własności podane na wykładzie z algebry, to zbiór $\mathbb{R}^n$ nazywa się **przestrzenią wektorową**, **przestrzenią liniową** lub **lineałem**. Własności owe można sprawdzić na podstawie podanych definicji obu działań, więc $\mathbb{R}^n$ jest lineałem. W lineale $\mathbb{R}^n$ można wyróżnić n wektorów o wpółrzędnych składających się z zer i jedynek, przy czym w każdym z nich jest tylko jedna jedynka:

$$\mathbf{e}_1 = \{1, 0, 0, \ldots, 0, 0\}, \quad \mathbf{e}_2 = \{0, 1, 0, \ldots, 0, 0\}, \quad \ldots \quad , \mathbf{e}_n = \{0, 0, 0, \ldots, 0, 1\}. \tag{5.3}$$

Są one liniowo niezależne oraz tę własność, że dowolny wektor $\mathbf{x} \in \mathbb{R}^n$ da się zapisać jako kombinacja liniowa

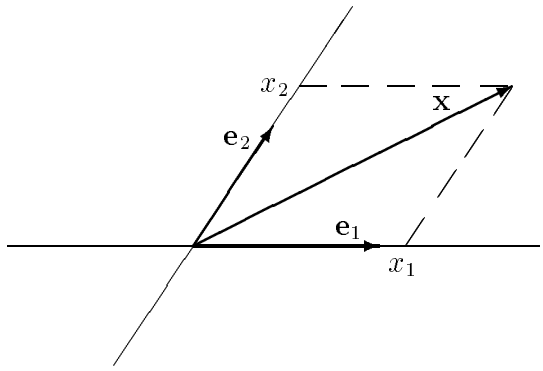
$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i. \tag{5.4}$$

Z tego powodu układ wektorów $\{\mathbf{e}_i\}$ jest **bazą**, a x_i są **współrzędnymi wektora** (lub **punktu**) $\mathbf{x}$ w tej bazie. Wzór (5.4) w trzech wymiarach przyjmuje postać

$$\{x_1, x_2, x_3\} = x_1 \{1, 0, 0\} + x_2 \{0, 1, 0\} + x_3 \{0, 0, 1\}.$$

Układ wektorów (5.3) nazywa się **bazą kanoniczną**.

Jeśli wprowadzić n prostych przechodzących przez początek układu współrzędnych, przy czym te proste nie leżą w jednej podprzestrzeni o wymiarze mniejszym niż n, a na każdej z tych prostych umieścimy po jednym wektorze $\mathbf{e}_i$, to nazywają się one **osiami układu współrzędnych**. Wtedy i-tą wpółrzędną wektora $\mathbf{x}$ znajduje się przez rzutowanie tego wektora na i-tą prostą równolegle do podprzestrzeni wyznaczonej przez pozostałe proste i znalezienie liczby x_i takiej, że ten rzut jest iloczynem $x_i\mathbf{e}_i$, patrz Rys. 5.1. Ujemna liczba x_i oznacza, że rzut wektora $\mathbf{x}$ na i-tą prostą ma zwrot przeciwny do $\mathbf{e}_i$.



Rysunek 5.1

W lineale $\mathbb{R}^n$ wprowadza się **iloczyn skalarny** wektorów

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i. \tag{5.5}$$

(Będę też nazywać go **skaloczynem**). Można sprawdzić, że tak określony iloczyn skalarny ma następujące własności:

1. $\forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$: $\quad \mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ – symetryczność,
2. $\forall\, \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$: $\quad \mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$ – rozdzielność względem dodawania,
3. $\forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,\ \alpha \in \mathbb{R}$: $\quad \mathbf{x} \cdot (\alpha \mathbf{y}) = \alpha(\mathbf{x} \cdot \mathbf{y})$ – jednorodność,
4. $\forall\, \mathbf{x} \in \mathbb{R}^n$: $\quad \mathbf{x} \cdot \mathbf{x} \geq 0$ – nieujemność,
5. $\mathbf{x} \cdot \mathbf{x} = 0 \quad \Rightarrow \quad \mathbf{x} = \mathbf{0}$ – niezwyrodniałość.

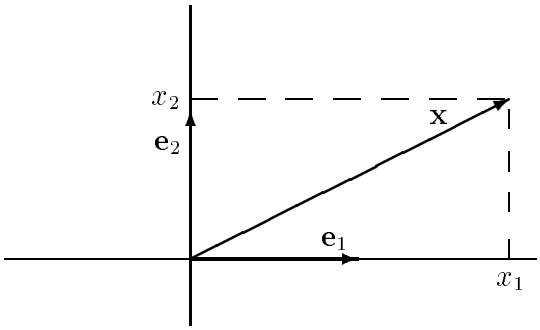
Za pomocą skaloczynu określa się **prostopadłość** dwóch wektorów:

$$\mathbf{x} \perp \mathbf{y} \quad \text{w.t.w. gdy} \quad \mathbf{x} \cdot \mathbf{y} = 0.$$

Można sprawdzić według definicji skaloczynu (5.5), że różne wektory bazowe (5.3) są prostopadłe $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ dla $i \neq j$, a każdy z nich ma kwadrat skalarny jeden: $\mathbf{e}_i \cdot \mathbf{e}_i = 1$. Obie te własności można zawrzeć w jednym wzorze: $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$, gdzie

$$\delta_{ij} = \begin{cases} 0 & \text{gdy} \quad i \neq j, \\ 1 & \text{gdy} \quad i = j. \end{cases}$$

nazywa się **deltą Kroneckera**. Stąd wynika, że po wprowadzeniu skaloczynu według wzoru (5.5) osie układu współrzędnych muszą być prostopadłe, a wybrane na nich wektory bazowe muszą być jednostkowe. W takim razie Rys. 5.1 powinien być teraz zastąpiony przez Rys. 5.2. [Leopold Kronecker (1823-1891) – niemiecki matematyk urodzony w Legnicy, autor prac z algebry i teorii liczb, zwolennik sprowadzenia matematyki do arytmetyki liczb naturalnych.]



Rysunek 5.2

Twierdzenie 1. Dla dowolnych $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ zachodzi nierówność

$$|\mathbf{x} \cdot \mathbf{y}| \leq (\mathbf{x} \cdot \mathbf{x})^{1/2} (\mathbf{y} \cdot \mathbf{y})^{1/2}$$

zwana **nierównością Schwarza**. [Hermann Amandus Schwarz (1843–1921) – matematyk niemiecki urodzony w Przewoźnikach koło Żar; główne jego prace są poświęcone analizie matematycznej i jej zastosowaniom w geometrii.]

D o w ó d . Z własności nieujemności skaloczynu dla dowolnych $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ i dowolnych $\alpha, \beta \in \mathbb{R}$ możemy napisać

$$(\alpha\mathbf{x} + \beta\mathbf{y}) \cdot (\alpha\mathbf{x} + \beta\mathbf{y}) \geq 0.$$

Korzystamy z jednorodności i rozdzielności względem dodawania:

$$\alpha^2(\mathbf{x} \cdot \mathbf{x}) + 2\alpha\beta(\mathbf{x} \cdot \mathbf{y}) + \beta^2(\mathbf{y} \cdot \mathbf{y}) \geq 0.$$

Teraz wybieramy $\alpha = \mathbf{y} \cdot \mathbf{y}$, $\beta = -\mathbf{x} \cdot \mathbf{y}$:

$$(\mathbf{y} \cdot \mathbf{y})^2(\mathbf{x} \cdot \mathbf{x}) - 2(\mathbf{y} \cdot \mathbf{y})(\mathbf{x} \cdot \mathbf{y})^2 + (\mathbf{x} \cdot \mathbf{y})^2(\mathbf{y} \cdot \mathbf{y}) \geq 0,$$

$$(\mathbf{y} \cdot \mathbf{y})^2(\mathbf{x} \cdot \mathbf{x}) - (\mathbf{y} \cdot \mathbf{y})(\mathbf{x} \cdot \mathbf{y})^2 \geq 0,$$

$$(\mathbf{y} \cdot \mathbf{y})(\mathbf{x} \cdot \mathbf{x}) - (\mathbf{x} \cdot \mathbf{y})^2 \geq 0,$$

$$(\mathbf{y} \cdot \mathbf{y})(\mathbf{x} \cdot \mathbf{x}) \geq (\mathbf{x} \cdot \mathbf{y})^2,$$

$$(\mathbf{x}\cdot\mathbf{y})^2 \leq (\mathbf{x}\cdot\mathbf{x})(\mathbf{y}\cdot\mathbf{y}).$$

Skoro pierwiastek jest funkcją rosnącą, to po spierwiastkowaniu zachowany jest zwrot nierówności:

$$|\mathbf{x}\cdot\mathbf{y}| \leq (\mathbf{x}\cdot\mathbf{x})^{1/2}(\mathbf{y}\cdot\mathbf{y})^{1/2}. \blacksquare$$

Po wprowadzeniu iloczynu skalarnego w lineale $\mathbb{R}^n$ definiuje się **normę skaloczynową**, która zadaje **długość** wektora $\mathbf{x}$:

$$\|\,\mathbf{x}\,\| = (\mathbf{x}\cdot\mathbf{x})^{1/2}. \tag{5.6}$$

Za pomocą tej normy można nierówność Schwarza zapisać inaczej

$$|\mathbf{x}\cdot\mathbf{y}| \leq \|\,\mathbf{x}\,\|\;\|\,\mathbf{y}\,\|\,. \tag{5.7}$$

Z wzorów (5.6) i (5.5) wynika, że norma wektora wyraża się przez jego współrzędne wzorem $\|\,\mathbf{x}\,\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$, a stąd wynika nierówność

$$\forall j \in \{1, 2, \ldots n\}: \quad |x_j| \leq \|\,\mathbf{x}\,\|\,. \tag{5.8}$$

Norma ma następujące własności

1. $\forall\,\alpha \in \mathbb{R},\ \mathbf{x} \in \mathbb{R}^n: \quad \|\,\alpha\mathbf{x}\,\| = |\alpha|\;\|\,\mathbf{x}\,\|$ – dodatnia jednorodność,
2. $\forall\,\mathbf{x} \in \mathbb{R}^n: \ \|\,\mathbf{x}\,\| \geq 0$ – nieujemność,
3. $\|\,\mathbf{x}\,\| = 0 \quad\Rightarrow\quad \mathbf{x} = \mathbf{0}$ – niezwyrodniałość,
4. $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \quad \|\,\mathbf{x}+\mathbf{y}\,\| \leq \|\,\mathbf{x}\,\| + \|\,\mathbf{y}\,\|$ – nierówność trójkąta.

Te wszystkie własności można wyprowadzić z podanych własności iloczynu skalarnego. My sprawdzimy jedynie własność 4. Korzytając z definicji normy (5.6) liczymy kwadrat lewej strony:

$$\|\,\mathbf{x}+\mathbf{y}\,\|^2 = (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y}) = (\mathbf{x}\cdot\mathbf{x}) + (\mathbf{y}\cdot\mathbf{y}) + 2\,(\mathbf{x}\cdot\mathbf{y}) = \|\,\mathbf{x}\,\|^2 + \|\,\mathbf{y}\,\|^2 + 2\,(\mathbf{x}\cdot\mathbf{y})$$

$$\leq \|\,\mathbf{x}\,\|^2 + \|\,\mathbf{y}\,\|^2 + 2\,|\mathbf{x}\cdot\mathbf{y}|.$$

Korzystamy z nierówności Schwarza (5.7):

$$\|\,\mathbf{x}+\mathbf{y}\,\|^2 \leq \|\,\mathbf{x}\,\|^2 + \|\,\mathbf{y}\,\|^2 + 2\;\|\,\mathbf{x}\,\|\;\|\,\mathbf{y}\,\| = (\|\,\mathbf{x}\,\| + \|\,\mathbf{y}\,\|)^2.$$

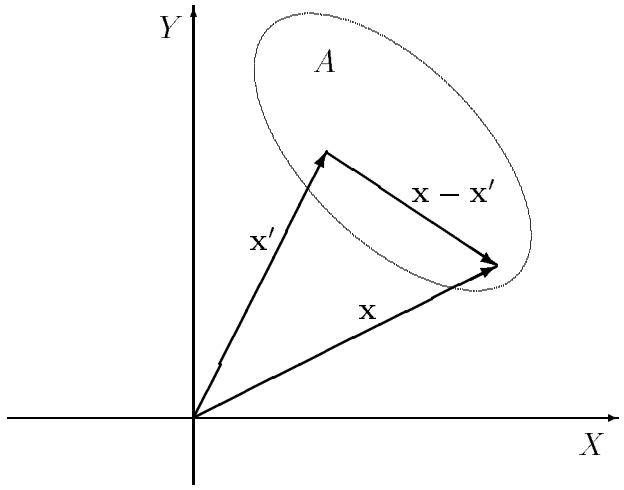
Skoro pierwiastek jest funkcją rosnącą, to po spierwiastkowaniu zachowany jest zwrot nierówności:

$$\|\,\mathbf{x}+\mathbf{y}\,\| \leq \|\,\mathbf{x}\,\| + \|\,\mathbf{y}\,\|\,. \blacksquare$$

Mając do dyspozycji pojęcie normy możemy wprowadzić wiele następnych pojęć. **Odległością** między punktami $\mathbf{x}$ i $\mathbf{y}$ nazywamy

$$d(\mathbf{x}, \mathbf{y}) = \|\,\mathbf{x}-\mathbf{y}\,\|\,.$$

Niech $A \subset \mathbb{R}^n$; liczbę $\operatorname{diam} A = \sup\{\|\,\mathbf{x}-\mathbf{x}'\,\|: \mathbf{x}, \mathbf{x}' \in A\}$ nazywamy **średnicą zbioru A**. Zbiór A nazywamy **ograniczonym**, jeśli jego średnica jest skończona. Taki zbiór jest ukazany na Rys. 5.3.



Rysunek 5.3

Zbiór

$$K(\mathbf{a}, r) = \{\mathbf{x} \in \mathbb{R}^n : \ ||\mathbf{x} - \mathbf{a}|| < r\} \tag{5.9}$$

nazywamy **niby-kulą otwartą o środku a i promieniu** r. Użyliśmy słowa niby-kula, bo w jednym wymiarze, czyli dla $n = 1$, jest to odcinek (otwarty o długości $2r$), w dwóch wymiarach - koło (bez okręgu), w trzech - kula (bez powierzchni), a w czterech nie ma na to słowa. Dla dowolnego n matematycy nazywają to po prostu kulą n-wymiarową. Zbiór

$$B(\mathbf{a}, r) = \{\mathbf{x} \in \mathbb{R}^n : \ ||\mathbf{x} - \mathbf{a}|| \leq r\} \tag{5.10}$$

nazywamy **niby-kulą domkniętą o środku a i promieniu r**.

Zbiór $U \subset \mathbb{R}^n$ nazywa się **otwarty**, jeśli

$$\forall \mathbf{x} \in U \quad \exists \delta > 0 : K(\mathbf{x}, \delta) \subset U. \tag{5.11}$$

Innymi słowy, zbiór jest otwarty, jeśli wraz z dowolnym punktem zawiera też pewną niby-kulę otwartą.

PRZYKŁAD 1. Sprawdzimy, że niby-kula K określona wzorem (5.9) jest zbiorem otwartym w sensie tej definicji. Weźmy dowolny punkt $\mathbf{x} \in K(\mathbf{a}, r)$. Na mocy (5.9) zachodzi nierówność $||\mathbf{x} - \mathbf{a}|| \ < r$. Wybieramy x

$$\delta = r - ||\mathbf{x} - \mathbf{a}||. \tag{5.12}$$

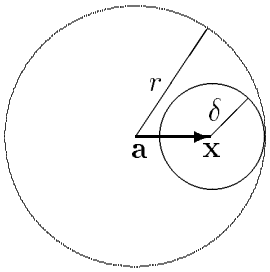
Rys. 5.4 ilustruje to w dwóch wymiarach. Twierdzimy, że zachodzi $K(\mathbf{x}, \delta) \subset K(\mathbf{a}, r)$. Weźmy punkt $\mathbf{x}' \in K(\mathbf{x}, \delta)$, co oznacza $||\mathbf{x}' - \mathbf{x}|| < \delta$. Z nierówności trójkąta mamy

$$||\mathbf{x}' - \mathbf{a}|| \leq ||\mathbf{x}' - \mathbf{x}|| + ||\mathbf{x} - \mathbf{a}|| < \delta + ||\mathbf{x} - \mathbf{a}||.$$

Korzystamy z (5.12):

$$||\mathbf{x}' - \mathbf{a}|| < r.$$

To oznacza, że $\mathbf{x}' \in K(\mathbf{a}, r)$. Pokazaliśmy wynikanie $\mathbf{x}' \in K(\mathbf{x}, \delta) \Rightarrow \mathbf{x}' \in K(\mathbf{a}, r)$, co dla dowolnego $\mathbf{x} \in K(\mathbf{a}, r)$ oznacza zawieranie zbiorów $K(\mathbf{x}, \delta) \subset K(\mathbf{a}, r)$. Zdanie (5.11) jest prawdziwe, więc niby-kula $K(\mathbf{a}, r)$ może być nazywana otwartą.



Rysunek 5.4

Niech $\langle a, b\rangle$ będzie przedziałem domkniętym w $\mathbb{R}$ i niech dana będzie funkcja $\boldsymbol{\sigma} : \langle a, b\rangle \to \mathbb{R}^n$. Zwykle będziemy zakładać, ze funkcja $\boldsymbol{\sigma}(t) = \{\sigma_1(t), \sigma_2(t), \ldots \sigma_n(t)\}$ ma wszystkie współrzędne różniczkowalne i ich pochodne są ciągłe. Zbiór

$$K = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \boldsymbol{\sigma}(t),\ t \in \langle a, b\rangle\} \tag{5.13}$$

nazywamy **krzywą** w $\mathbb{R}^n$, a funkcję $\boldsymbol{\sigma}$ nazywamy **parametryzacją** krzywej K. Można powiedzieć, że krzywa K to obraz funkcji $\boldsymbol{\sigma}$. W fizyce często posługujemy się takimi funkcjami $\mathbf{r} = \boldsymbol{\sigma}(t)$ opisującymi ruch jakiejś cząstki w naszej przestrzeni $\mathbb{R}^3$. W takim razie w terminologii fizycznej parametryzacja $\boldsymbol{\sigma}(t)$ to ruch, a krzywa K to tor tego ruchu.

Niech $\boldsymbol{\sigma} : \langle a, b\rangle \to \mathbb{R}^n$. Wektor $\mathbf{h} \in \mathbb{R}^n$ jest **granicą funkcji** $\boldsymbol{\sigma}$ w punkcie $c \in \langle a, b\rangle$, jeśli

$$\forall\ \epsilon > 0\ \ \exists\ \delta > 0 :\ 0 < |t - c| < \delta \Rightarrow ||\boldsymbol{\sigma}(t) - \mathbf{h}|| < \epsilon. \tag{5.14}$$

Piszemy wtedy

$$\lim_{t \to c} \boldsymbol{\sigma}(t) = \mathbf{h}.$$

Podane zdanie nosi nazwę **definicji Cauchy'ego**.

Funkcja $\boldsymbol{\sigma} : \langle a, b\rangle \to \mathbb{R}^n$ nazywa się **ciągła w punkcie** $c \in \langle a, b\rangle$, jeśli

$$\lim_{t \to c} \boldsymbol{\sigma}(t) = \boldsymbol{\sigma}(c). \tag{5.15}$$

Twierdzenie 2. Funkcja $\boldsymbol{\sigma} : \langle a, b\rangle \to \mathbb{R}^m$ jest ciągła w punkcie $c \in \langle a, b\rangle$ w.t.w. gdy każda współrzędna $\sigma_k : \langle a, b\rangle \to \mathbb{R}$ jest ciągła w punkcie c.

Dowód $\Rightarrow$. Załóżmy, że $\lim\limits_{t \to c} \boldsymbol{\sigma}(t) = \boldsymbol{\sigma}(c)$. Z definicji (5.14) to znaczy, że dla dowolnego $\epsilon > 0$ istnieje $\delta > 0$ taka, że dla $t \in (c - \delta, c + \delta)$ zachodzi $||\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}(c)|| < \epsilon$. Według wzoru (5.8) dla każdego $j \in \{1, 2, ..., n\}$ mamy $|\sigma_j(t) - \sigma_j(c)| \leq ||\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}(c)||$, więc dodając wszystkie wspomniane kwantyfikatory możemy napisać

$$\forall\ \epsilon > 0\ \ \exists\ \delta > 0 :\ 0 < |t - c| < \delta \Rightarrow \forall j \in \{1, \ldots, n\}\ |\sigma_j(t) - \sigma_j(c)| < \epsilon.$$

a to oznacza, że każda funkcja-współrzędna $\sigma_j(t)$ jest ciągław punkcie c. □

$\Leftarrow$. Załóżmy, że dla każdego $j \in \{1, 2, ..., n\}$ funkcja σ_j jest ciągła w punkcie c, tzn.

$$\forall j \in \{1, \ldots, n\}\ \forall\ \epsilon > 0\ \ \exists\ \delta_j > 0 :\ 0 < |t - c| < \delta_j \Rightarrow |\sigma_j(t) - \sigma_j(c)| < \epsilon.$$

Przyjęliśmy δ_j zależne od j, gdyż w założeniu jest ciągłość dla każdej j-tej współrzędnej z osobna. Jeśli wybierzemy $\delta = \min\{\delta_1, \delta_2, ..., \delta_n\}$, to będziemy mogli napisać zdanie z kwantyfikatorem zawierającym j na końcu:

$$\forall \epsilon > 0 \;\; \exists \delta > 0 : \; 0 < |t - c| < \delta \Rightarrow \forall j \in \{1, 2, \ldots, n\} : |\sigma_j(t) - \sigma_j(c)| < \frac{\epsilon}{\sqrt{n}}.$$

Z umieszczonej na końcu nierówności wynika taka

$$||\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}(c)|| = \sqrt{|\sigma_1(t) - \sigma_1(c)|^2 + |\sigma_2(t) - \sigma_2(c)|^2 + \ldots + |\sigma_n(t) - \sigma_n(c)|^2} <$$

$$< \sqrt{\frac{\epsilon^2}{n} + \frac{\epsilon^2}{n} + \ldots + \frac{\epsilon^2}{n}} = \sqrt{\epsilon^2} = \epsilon.$$

W takim razie mamy prawdziwe zdanie

$$\forall \epsilon > 0 \; \exists \delta > 0 : \; 0 < |t - c| < \delta \Rightarrow ||\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}(c)|| < \epsilon.$$

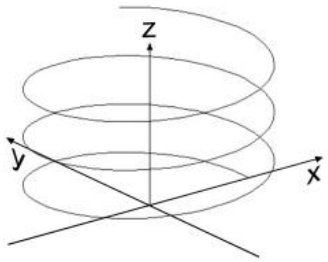
A to jest zdanie (5.14), więc funkcja $\boldsymbol{\sigma}$ jest ciągła w punkcie c. ■

PRZYKŁAD 2. Rozważmy funkcję $\boldsymbol{\sigma} : \langle 0, 4\pi \rangle \to \mathbb{R}^2$ zadaną dwoma wzorami $\sigma_1(t) = kt + a$, $\sigma_2(t) = r \sin t$. Pierwsza funkcja jest liniowa, a druga jest sinusem, więc otrzymana krzywa jest sinusoidą falującą wokół osi X_1, przeskalowaną na pierwszej osi przez czynnik k, a na drugiej przez czynnik r i przesuniętą w prawo o odcinek a. Skoro dziedzina jest przedziałem, a nie całą prostą, to z tej sinusoidy zostaną tylko dwa okresy. Gdybyśmy zmienili drugą funkcję na $\sigma_2(t) = b + r \sin t$, to wszystko trzeba przesunąć o stałą b równolegle do drugiej osi, więc sinusoida będzie falować wokół prostej $x_2 = b$. Obie funkcje σ_j są ciągłe, więc na mocy Twierdzenia 2 funkcja $\boldsymbol{\sigma}$ jest ciągła.

PRZYKŁAD 3. Rozważmy funkcję $\boldsymbol{\sigma} : \langle 0, 6\pi \rangle \to \mathbb{R}^2$ zadaną dwoma wzorami $\sigma_1(t) = p \cos t$, $\sigma_2(t) = r \sin t$. Obie funkcje są trygonometryczne, ich wartości zmieniają się odpowiednio w przedziałach $\langle -p, p \rangle$ oraz $\langle -r, r \rangle$, więc cała krzywa mieści się w prostokącie o bokach $2p$, $2r$. Sprawdzamy równość

$$\frac{x_1^2}{p^2} + \frac{x_2^2}{r^2} = 1,$$

więc stwierdzamy, że fukcja $\boldsymbol{\sigma}$ zadaje elipsę o środku w początku układu i półosiach p, r. Skoro argument t przebiega trzy okresy, to wybrana parametryzacja jest taka, że elipsa jest przebiegana trzy razy. Obie funkcje σ_i są ciągłe, więc funkcja $\boldsymbol{\sigma}$ jest ciągła.



Rysunek 5.4a

PRZYKŁAD 4. Rozważmy funkcję $\boldsymbol{\sigma} : \langle 0, 6\pi+\pi/4\rangle \to \mathbb{R}^3$ zadaną trzema wzorami $\sigma_1(t) = p\cos t$, $\sigma_2(t) = r\sin t$, $\sigma_3(t) = kt + c$. Dwie pierwsze współrzędne zakreślają elipsę w płaszczyźnie X_1, X_2, a jednocześnie trzecia współrzędna zmienia się liniowo. Mamy złożenie ruchu po elipsie z ruchem jednostajnym - to oznacza, że otrzymana krzywa trójwymiarowa jest eliptyczną linią śrubową o skoku $2k\pi$ obejmującą trochę więcej niż trzy skoki tej śruby i zaczynającą się od punktu $(p, 0, c)$ (zob Rys 5.4a). Gdybyśmy zmienili dwie pierwsze funkcje na $\sigma_1(t) = a + p\cos t$, $\sigma_2(t) = b + r\sin t$, to owa linia śrubowa byłaby nawinięta na linię równoległą do osi X_3 przesuniętą o wektor (a, b) na płaszczyźnie X_1, X_2. Dla $r = p$ eliptyczna linia śrubowa staje się zwykłą, czyli kołową linią śrubową. Wszystkie trzy funkcje σ_i są ciągłe, więc funkcja $\boldsymbol{\sigma}$ jest ciągła.

Podzbiór lineału $\mathbb{R}^n$ nazywamy **spójnym**, jeśli dowolne jego dwa punkty można połączyć krzywą ciągłą w całości zawartą w tym podzbiorze.

PRZYKŁAD 5. Przedział na prostej $\mathbb{R}$ jest zbiorem spójnym, dwa przedziały bez części wspólnej połączone w jeden zbiór dają zbiór niespójny. Zbiór spójny na prostej jest przedziałem albo półprostą, albo całą prostą.

PRZXYKŁAD 6. Płaszczyzna $\mathbb{R}^2$ jest zbiorem spójnym; płaszczyzna z usuniętą prostą nie jest zbiorem spójnym, bo dwóch punktów po różnych stronach tej prostej nie można połączyć krzywą ciągłą bez przecinania tej prostej. Przestrzeń $\mathbb{R}^3$ jest spójna, ale po usunięciu z niej płaszczyzny już nie jest zbiorem spójnym, bo krzywa ciągła łącząca punkty leżące po różnych stronach płaszczyzny musi przeciąć tę płaszczyznę.

Twierdzenie 3. Jeśli dwie funkcje $\boldsymbol{\sigma}$ i $\boldsymbol{\kappa}$ mają granicę w punkcie c, to ich suma i różnica mają granice w tym punkcie oraz zachodzą wzory

$$\lim_{t\to c}[\boldsymbol{\sigma}(t) + \boldsymbol{\kappa}(t)] = \lim_{t\to c}\boldsymbol{\sigma}(t) + \lim_{t\to c}\boldsymbol{\kappa}(t),$$
$$\lim_{t\to c}[\boldsymbol{\sigma}(t) - \boldsymbol{\kappa}(t)] = \lim_{t\to c}\boldsymbol{\sigma}(t) - \lim_{t\to c}\boldsymbol{\kappa}(t).$$

D o w ó d. Niech będą dane dwie funkcje $\boldsymbol{\sigma}$, $\boldsymbol{\kappa} : \langle a, b\rangle \to \mathbb{R}^n$ i niech

$$\lim_{t\to c}\boldsymbol{\sigma}(t) = \mathbf{h}, \quad \lim_{t\to c}\boldsymbol{\kappa}(t) = \mathbf{g}.$$

Według definicji (5.14) możemy napisać

$$\forall\ \epsilon > 0\ \ \exists\ \delta_\sigma > 0 :\ 0 < |t - c| < \delta_\sigma \Rightarrow ||\boldsymbol{\sigma}(t) - \mathbf{h}|| < \epsilon/2,$$

$$\forall\ \epsilon > 0\ \ \exists\ \delta_\kappa > 0 :\ 0 < |t - c| < \delta_\kappa \Rightarrow ||\boldsymbol{\kappa}(t) - \mathbf{g}|| < \epsilon/2.$$

Na mocy nierówności trójkąta mamy

$$||\,[\boldsymbol{\sigma}(t) + \boldsymbol{\kappa}(t)] - (\mathbf{h} + \mathbf{g})|| = ||\,[\boldsymbol{\sigma}(t) - \mathbf{h}] + [\boldsymbol{\kappa}(t)] - \mathbf{g}]\,|| \le ||\boldsymbol{\sigma}(t) - \mathbf{h}|| + ||\boldsymbol{\kappa}(t) - \mathbf{g}||.$$

Niech $\delta = \min\{\delta_\sigma, \delta_\kappa\}$. Wtedy dla $0 < |t - c| < \delta$ zachodzi

$$||\,[\boldsymbol{\sigma}(t) + \boldsymbol{\kappa}(t)] - (\mathbf{h} + \mathbf{g})|| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Dotychczasowe spostrzeżenia możemy połączyć w takie zdanie

$$\forall\ \epsilon > 0\ \ \exists\ \delta > 0:\ 0 < |t - c| < \delta \Rightarrow ||\,[\boldsymbol{\sigma}(t) + \boldsymbol{\kappa}(t)] - (\mathbf{h} + \mathbf{g})|| < \epsilon,$$

a według definicji (5.14) to oznacza, że

$$\lim_{t\to c}[\boldsymbol{\sigma}(t) + \boldsymbol{\kappa}(t)] = \mathbf{h} + \mathbf{g} = \lim_{t\to c}\boldsymbol{\sigma}(t) + \lim_{t\to c}\boldsymbol{\kappa}(t).$$

Dla różnicy granic dowód jest analogiczny. ■

Twierdzenie 4. Jeśli dwie funkcje: $f : \langle a, b\rangle \to \mathbb{R}$, $\boldsymbol{\sigma} : \langle a, b\rangle \to \mathbb{R}^n$ mają granice w punkcie $c \in \langle a, b\rangle$, to ich iloczyn ma granicę w tym punkcie i zachodzi wzór

$$\lim_{t\to c}[f(t)\boldsymbol{\sigma}(t)] = \lim_{t\to c} f(t) \cdot \lim_{t\to c}\boldsymbol{\sigma}(t).$$

D o w ó d pomijamy.

Niech $\boldsymbol{\sigma} : \langle a, b\rangle \to \mathbb{R}^n$. **Pochodną funkcji $\boldsymbol{\sigma}$ w punkcie** $c \in \langle a, b\rangle$ nazywamy granicę

$$\lim_{t\to c}\frac{\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}(c)}{t - c} = \lim_{h\to 0}\frac{\boldsymbol{\sigma}(c + h) - \boldsymbol{\sigma}(c)}{h}. \tag{5.16}$$

Pochodną w punkcie c zapisujemy na trzy sposoby:

$$\left[\frac{d\boldsymbol{\sigma}(t)}{dt}\right]_{t=c} = \frac{d\boldsymbol{\sigma}}{dt}(c) = \boldsymbol{\sigma}'(c).$$

Funkcję $\boldsymbol{\sigma}$ nazywamy **różniczkowalną w punkcie** c, jeśli istnieje jej pochodna w tym punkcie.

Lemat 1. Jeśli funkcja $\boldsymbol{\sigma}$ jest różniczkowalna w punkcie c to jest ciągła w tym punkcie.

D o w ó d z odpowiednimi zmianami taki sam, jak dla Lematu 1 w § 2.1.

Niezerowy wektor $\mathbf{v} = \boldsymbol{\sigma}'(c)$ ma interpretację **wektora stycznego do krzywej K w punkcie $\boldsymbol{\sigma}(c)$**. W interpretacji fizycznej jest to prędkość ruchu. Wektor pochodnej może też być funkcją. Ma ona współrzędne

$$\mathbf{v}(t) = \boldsymbol{\sigma}'(t) = \{\sigma_1'(t),\ \sigma_2'(t),\ \ldots,\ \sigma_n'(t)\}. \tag{5.17}$$

Krzywa zakreślana przez prędkość nazywa się **hodografem**.

PRZYKŁAD 7. Rozważmy funkcję $\boldsymbol{\sigma} : \langle 0, 6\pi\rangle \to \mathbb{R}^3$ zadaną wzorami $\sigma_1(t) = p\cos t$, $\sigma_2(t) = r\sin t$, $\sigma_3(t) = kt + c$, czyli funkcję z Przykładu 4 opisującą eliptyczną linię śrubową. Jej pochodna jest następującą funkcją

$$\boldsymbol{\sigma}'(t) = \{\sigma_1'(t),\ \sigma_2'(t),\ \sigma_3'(t)\} = \{-p\sin t,\ r\cos t,\ k\}.$$

Ostatnia współrzędna nie zależy od t, jest stała, więc otrzymana krzywa jest płaska – leży w płaszczyźnie $X_3 = k$. W dwóch pierwszych współrzędnych mamy funkcje trygonometryczne, więc hodograf jest elipsą o półosiach p, r tych samych, co w Przykładzie 4. Hodograf zaczyna się w punkcie $\{0, r, k\}$.

Twierdzenie 5. Niech funkcje $\boldsymbol{\sigma}$ i $\boldsymbol{\kappa}$ będą różniczkowalne w punkcie c. Pochodna ich sumy (bądź różnicy) w punkcie c jest sumą (bądź różnicą) ich pochodnych w tym punkcie.

D o w ó d z odpowiednimi zmianami taki sam, jak dla Twierdzenia 1 w § 2.1.

Twierdzenie 6. Niech będą dane dwie funkcje: $f : \langle a, b\rangle \to \mathbb{R}$, $\boldsymbol{\sigma} : \langle a, b\rangle \to \mathbb{R}^n$ różniczkowalne w punkcie $c \in \langle a, b\rangle$. Iloczyn tych funkcyj jest też różniczkowalny w punkcie c i zachodzi wzór

$$(f\boldsymbol{\sigma})'(c) = f'(c)\boldsymbol{\sigma}(c) + f(c)\boldsymbol{\sigma}'(c).$$

D o w ó d. Liczymy

$$\begin{aligned}
(f\boldsymbol{\sigma})'(c) &= \lim_{h\to 0} \frac{f(c+h)\boldsymbol{\sigma}(c+h) - f(c)\boldsymbol{\sigma}(c)}{h} \\
&= \lim_{h\to 0} \frac{[f(c+h)\boldsymbol{\sigma}(c+h) - f(c)\boldsymbol{\sigma}(c+h)] + [f(c)\boldsymbol{\sigma}(c+h) - f(c)\boldsymbol{\sigma}(c)]}{h} \\
&= \lim_{h\to 0} \frac{[f(c+h) - f(c)]\boldsymbol{\sigma}(c+h) + f(c)[\boldsymbol{\sigma}(c+h) - \boldsymbol{\sigma}(c)]}{h} \\
&= \lim_{h\to 0} \left[\frac{f(c+h) - f(c)}{h}\boldsymbol{\sigma}(c+h) + f(c)\frac{\boldsymbol{\sigma}(c+h) - \boldsymbol{\sigma}(c)}{h}\right].
\end{aligned}$$

Granica sumy jest sumą granic (Tw. 3), a granica iloczynu jest iloczynem granic (Tw. 4), więc

$$(f\boldsymbol{\sigma})'(c) = \lim_{h\to 0} \frac{f(c+h) - f(c)}{h} \lim_{h\to 0} \boldsymbol{\sigma}(c+h) + f(c) \lim_{h\to 0} \frac{\boldsymbol{\sigma}(c+h) - \boldsymbol{\sigma}(c)}{h}.$$

Funkcja $\boldsymbol{\sigma}$ z założenia jest różniczkowalna, więc na mocy Lematu 1 jest ciągła, stąd $\lim\limits_{h\to 0} \boldsymbol{\sigma}(c+h) = \boldsymbol{\sigma}(c)$ oraz

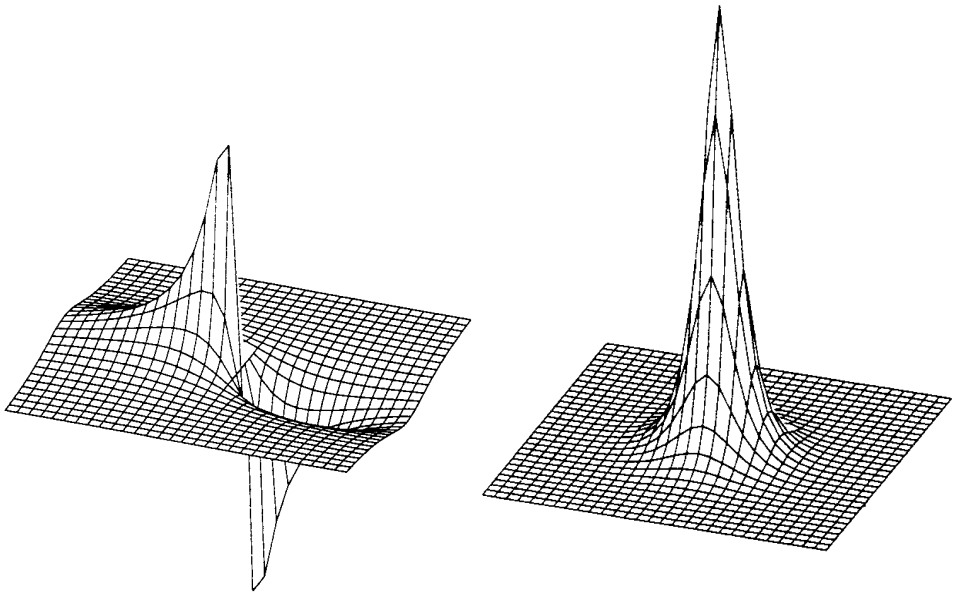
$$(f\boldsymbol{\sigma})'(c) = f'(c)\boldsymbol{\sigma}(c) + f(c)\boldsymbol{\sigma}'(c). \blacksquare$$

5.2 Funkcje o wartościach rzeczywistych

Przez **funkcję wielu zmiennych o wartościach rzeczywistych** i o dziedzinie A będziemy rozumieć odwzorowanie $f : A \to \mathbb{R}$, jeśli $A \subset \mathbb{R}^n$. Zbiór par uporządkowanych $(\mathbf{x}, f(\mathbf{x}))$ nazywa się **wykresem funkcji f**; jego oznaczenie

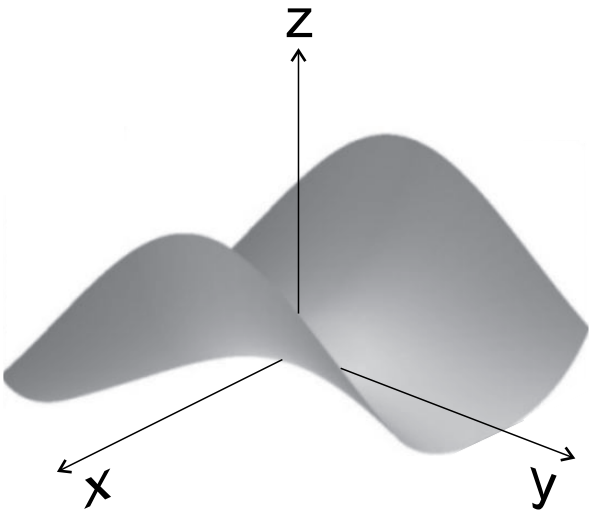
$$\operatorname{graf} f = \{(\mathbf{x}, f(\mathbf{x})) : \mathbf{x} \in A\}.$$

Wykres jest podzbiorem lineału $\mathbb{R}^{n+1}$. Tak więc dla sporządzenia wykresu funkcji jednej zmiennej potrzebna jest płaszczyzna, a dla funkcji dwóch zmiennych – przestrzeń trójwymiarowa. Na ogół dziedzinę funkcji dwóch zmiennych umieszcza się w płaszczyźnie (X, Y), a wartości funkcji – na osi Z. Wtedy wykresem funkcji jest jakaś powierzchnia (zob. Rys. 5.5, na którym ukazany jest wykres wzięty z publikacji fizycznej).



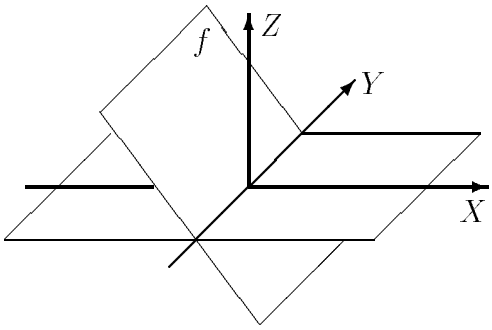
Rysunek 5.5

PRZYKŁAD 1. Wykresem funkcji $f(x,y) = x^2 - y^2$ jest powierzchnia siodłowa ukazana na Rys. 5.6. Przecięcie wykresu płasczyzną (X,Z) oznacza przyrównanie y do zera, co daje funkcję jednej zmiennej $f(x,0) = x^2$; jej wykresem jest parabola o ramionach zwrócnych ku górze. Przecięcie wykresu płaszczyzną (Y,Z) oznacza $x = 0$; wtedy wykresem funkcji jednej zmiennej $f(0,y) = -y^2$ jest parabola o ramionach zwróconych ku dołowi. Przy pewnej spostrzegawczości obie te parabole można zobaczyć na Rys. 5.6.

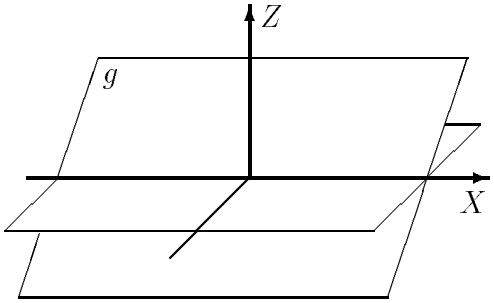


Rysunek 5.6

PRZYKŁAD 2. Wykresem funkcji $f(x,y) = -x$ jest płaszczyzna przechodząca przez oś Y i nachylona do płaszczyzny (X,Y) pod kątem 45^0 (zob. Rys. 5.7), a wykresem funkcji $g(x,y) = y$ jest płaszczyzna przechodząca przez oś X i też nachylona do płaszczyzny (X,Y) pod kątem 45^0 (zob. Rys. 5.8).



Rysunek 5.7



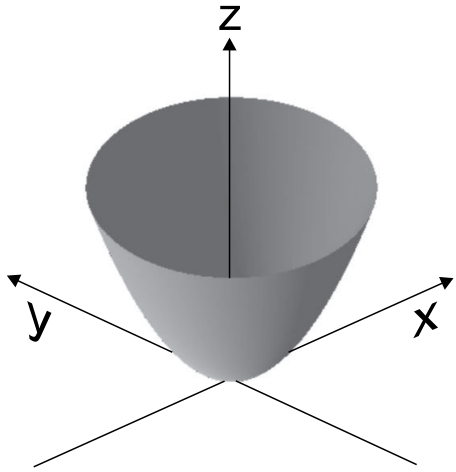
Rysunek 5.8

Jeśli $c \in \mathbb{R}$ jest ustaloną liczbą, to podzbiór dziedziny $L_c \subset D_f$ określony jako

$$L_c = \{\mathbf{x} \in D_f : f(\mathbf{x}) = c\}$$

nazywamy **poziomicą funkcji** f **o wartości** c. Zauważmy, że do wykreślania poziomic funkcji nie jest potrzebny lineał o wymiar wyższy, bo wykreślamy je w samej dziedzinie. Dla funkcji jednej zmiennej poziomicami są punkty, dla funkcji dwóch zmiennych – linie krzywe. Na przykład dla ukazanej na Rys. 5.7 funkcji f poziomicami są linie proste równoległe do osi Y, natomiast dla funkcji g z Rys. 5.8 poziomicami są proste równoległe do osi X. Gdyby wybierać poziomice w jednakowych odstępach wartości funkcji, np. $c_k = k$, $k \in \mathbb{Z}$, to odległości między poziomicami są miarą tego, jak stromy jest wykres. Dla podanych dwóch przykładów odległości między takimi poziomicami są stałe, co dowodzi tego, że wykresami są płaszczyzny.

PRZYKŁAD 3. Dla funkcji $h(x, y) = x^2 + y^2$ przekrój płaszczyzną (X, Z) o równaniu $y = 0$ daje funkcję jednej zmiennej $h(x, 0) = x^2$, której wykresem jest parabola. Jeśli napiszemy $h(\mathbf{r}) = ||\mathbf{r}||^2$, to zauważymy symetrię tej funkcji względem obrotów w płaszczyźnie (X, Y) wokół początku układu, a w przestrzeni wokół osi Z. Obracanie paraboli wokół jej osi symetrii daje **paraboloidę obrotową**, więc wykresem funkcji $h$jest paraboloida obrotowa, ukazaliśmy ją na Rys. 5.9. Poziomicami tej funkcji są okręgi wyznaczone przez równania $x^2 + y^2 = c$ o promieniach $r = \sqrt{x^2 + y^2} = \sqrt{c}$. Tym razem odstępy między poziomicami nie są stałe – maleją wraz ze wzrostem r, bo parabola jest coraz bardziej stroma w miarę oddalania się od jej wierzchołka.



Rysunek 5.9

PRZYKŁAD 4. Funkcja dana wzorem klamrowym

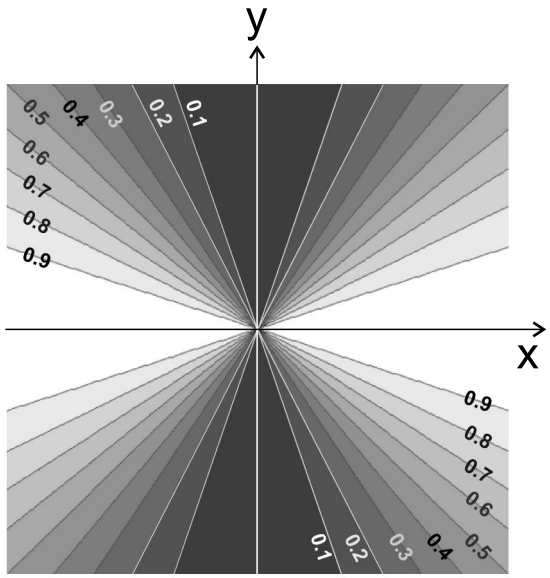
$$f(x,y) = \begin{cases} \frac{x^2}{x^2+y^2} & \text{gdy} \quad (x,y) \neq (0,0), \\ 0 & \text{gdy} \quad (x,y) = (0,0). \end{cases}$$

ma dziedzinę $D_f = \mathbb{R}^2$. Aby znaleźć jej poziomice, przyrównajmy wartość funkcji do stałej $c^2 \leq 1$ i przekształcajmy otrzymane równanie:

$$\frac{x^2}{x^2+y^2} = c^2, \qquad x^2 = c^2\left(x^2+y^2\right),$$

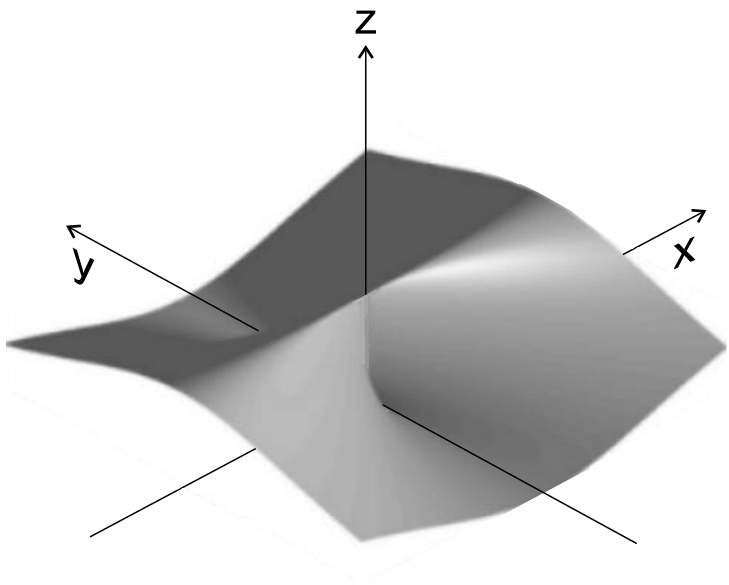
$$y^2 = \frac{1-c^2}{c^2}x^2, \qquad y = \pm\frac{\sqrt{1-c^2}}{|c|}\,x.$$

Widzimy, że poziomicami są proste przechodzące przez początek układu, patrz Rys. 5.10. Obszary między poziomicami są zacieniowane. Ciemniejszy odcień oznacza mniejszą wartość.



Rysunek 5.10

Omawiana funkcja ma wykres ukazany na Rys. 5.11. Chyba widać na nim, że miejscami stałej wartości są proste wychodzące z osi Z.

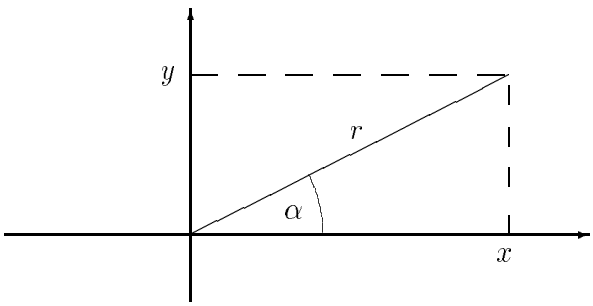


Rysunek 5.11

PRZYKŁAD 5. Dla opisania położenia punktu na płaszczyźnie oprócz współrzędnych kartezjańskich (x, y) można wprowadzić współrzędne biegunowe (r, α), a związek między nimi jest ukazany na Rys. 5.12 i opisany wzorami

$$x = r\cos\alpha, \quad y = r\sin\alpha. \tag{5.18}$$

Z tych wzorów widać, że $\frac{y}{x} = \operatorname{tg}\alpha$.



Rysunek 5.12

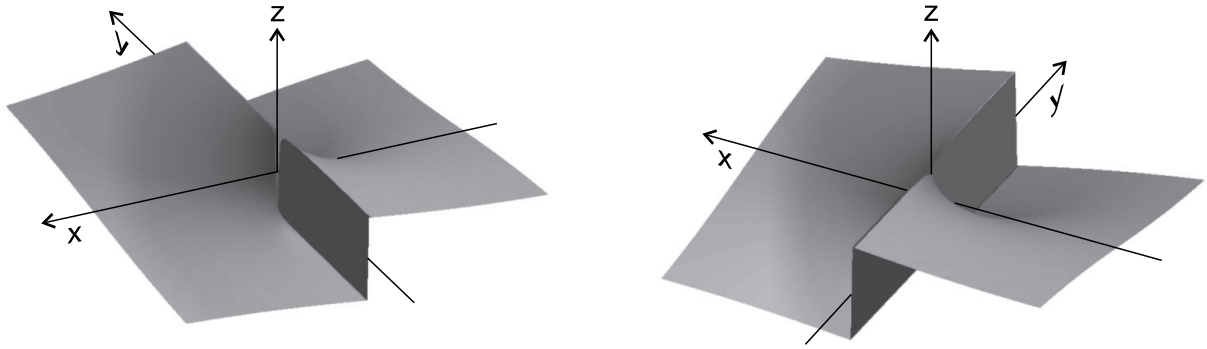
Rozważmy funkcję $\psi(x, y) = \operatorname{arc\,tg}(y/x)$. Z jej dziedziny trzeba usunąć punkty o współrzędnej $x = 0$, czyli oś Y, w takim razie $D_\psi = \mathbb{R}^2 \setminus \{(x, y) : x = 0\}$. Jej wykres ukazany jest na Rys. 5.13 w widoku z dwóch stron. Różnią się one obrotem o pewien kąt wokół osi Z. Skoro zachodzi $y/x = (-y)/(-x)$, to funkcja ψ jest symetryczna względem zmiany znaku obu współrzędnych $(x, y) \to (-x, -y) = -(x, y)$, czyli względem obrotu na płaszczyźnie (X, Y) wokół początku układu. Symetrię wykresu względem zamiany $\mathbf{r} \to -\mathbf{r}$ widać na Rys. 5.13.

Warto znaleźć poziomice tej funkcji. Uczynimy to wychodząc od równania $\operatorname{arc\,tg}(y/x) = \psi_0$ i znalezienia stąd zależności $y(x)$. Przekształcenia dają

$$\frac{y}{x} = \operatorname{tg}\psi_0,$$

$$y = (\operatorname{tg} \psi_0)\, x$$

Jest to równanie prostej przechodzącej przez początek układu, gdzie współczynnik przy x oznacza tangens kąta nachylenia prostej do osi X. Stąd mamy interpretację ψ_0 jako kąta między wektorem wodzącym do punktu (x, y) a osią X – funkcja ψ ma więc związek ze zwykłym kątem we współrzędnych biegunowych.



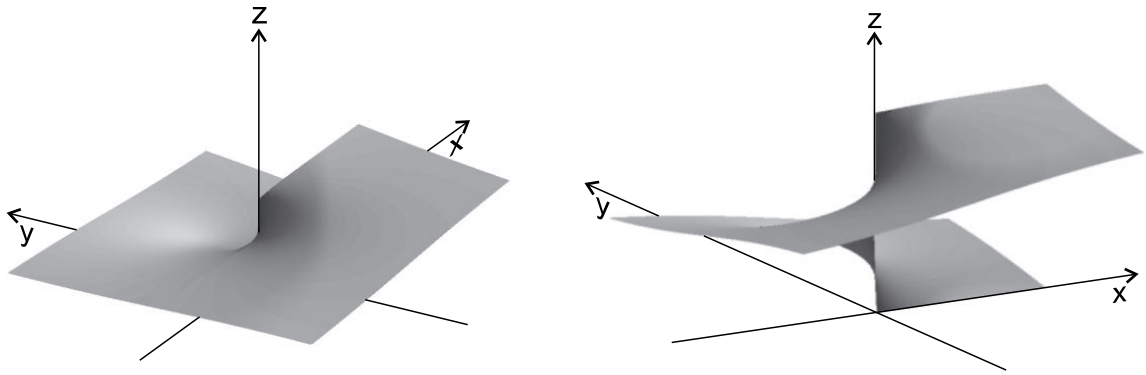
Rysunek 5.13

Nie jest to jednak dokładnie to samo. Kąt α we współrzędnych biegunowych zmienia się w przedziale $(0, 2\pi)$, zaś funkcja arcus tangens przyjmuje wartości w przedziale $(-\frac{\pi}{2}, \frac{\pi}{2})$. Poza tym ta funkcja jest symetryczna względem zmiany znaku obu współrzędnych $(x, y) \to (-x, -y)$, a kątowa współrzędna biegunowa nie ma tej własności.

Wyjściem z tych kłopotów jest określenie nowej funkcji za pomocą wzoru klamrowego:

$$\alpha(x, y) = \begin{cases} \operatorname{arc\,tg} \frac{y}{x} & \text{gdy} \quad x > 0,\ y \geq 0 \\ \pi + \operatorname{arc\,tg} \frac{y}{x} & \text{gdy} \quad x < 0,\ y \text{ dowolne}, \\ 2\pi + \operatorname{arc\,tg} \frac{y}{x} & \text{gdy} \quad x > 0,\ y \leq 0. \end{cases}$$

Wartość tej funkcji zmienia się od 0 do π, gdy przechodzić od dodatniej do ujemnej części osi X przy dodatnich wartościach y oraz zmienia się od π do 2π gdy posuwać się dalej od ujemnej do dodatniej części osi X przy ujemnych wartościach y. Tak właśnie zmienia się kątowa współrzędna biegunowa, co jest ukazane na Rys. 5.14 pod dwoma różnymi kątami widzenia. Jest to jeden zwój powierzchni śrubowej podobnej do części tkwiącej w środku maszynki do mielenia mięsa.



Rysunek 5.14

Dla funkcji trzech zmiennych nie jesteśmy w stanie narysować jej wykresu, bo do tego potrzebny jest lineał czterowymiarowy, natomiast możemy wskazywać jej poziomice. Na przykład dla funkcji $f(x,y,z) = x^2 + y^2 + z^2$ poziomicami wyznaczonymi przez równania $x^2+y^2+z^2 = c$ są sfery o promieniach $r = \sqrt{c}$. Te sfery są coraz gęstsze w miarę oddalania się od początku układu. W elektrostatyce poziomicami funkcji potencjału elektrycznego są **powierzchnie ekwipotencjalne**.

Mając pojęcie odległości w $\mathbb{R}^n$ możemy w analogii do funkcyj jednej zmiennej zdefiniować granicę funkcji w punkcie i ciągłość funkcji. **Definicja Cauchy'ego granicy funkcji**: funkcja $f : A \to \mathbb{R}$ dla $A \subset \mathbb{R}^n$ **ma granicę** b **w punkcie a**, jeśli

$$\forall \epsilon > 0 \;\; \exists \delta > 0 : \; 0 < \; ||\mathbf{x} - \mathbf{a}|| \; < \delta \Rightarrow |f(\mathbf{x}) - b\,| < \epsilon. \tag{5.19}$$

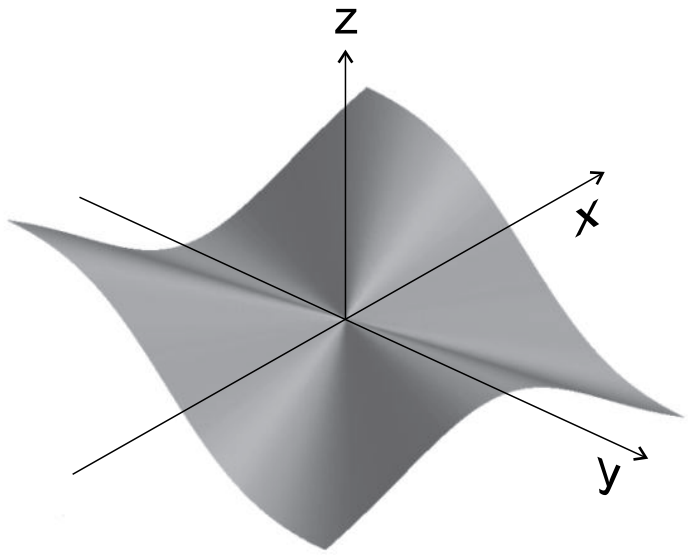
Piszemy wtedy

$$\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = b.$$

Widzimy, że jest to dostosowanie definicji Cauchy'ego granicy funkcji jednej zmiennej do przypadku wielu wymiarów.

PRZYKŁAD 6. Badamy, czy funkcja

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



Rysunek 5.15

ma granicę w zerze. Jej wykres jest ukazany na Rys. 5.15. Mamy nierówność $\frac{x^2}{x^2+y^2} \leq 1$, mnożymy ją stronami przez $|y|$:

$$\frac{x^2}{x^2+y^2}|y| \leq |y|,$$

skąd

$$\left|\frac{x^2 y}{x^2+y^2}\right| \leq |y|. \tag{5.20}$$

Jeśli $||\mathbf{r}-\mathbf{0}|| = \sqrt{x^2+y^2} < \delta$, to $|y| \leq \sqrt{x^2+y^2} < \delta$. Dla dowolnego $\epsilon > 0$ wystarczy dobrać $\delta = \epsilon$, a wtedy dzięki (5.20) mamy

$$|f(\mathbf{r}) - 0| = \left|\frac{x^2 y}{x^2+y^2}\right| \leq |y| < \epsilon.$$

Wykazaliśmy spełnienie zdania (5.19), więc $\lim\limits_{\mathbf{r}\to\mathbf{0}} f(\mathbf{r}) = 0$.

Twierdzenie 1. Niech $A \subset \mathbb{R}^n$. Jeśli funkcja $f : A \to \mathbb{R}$ ma granicę c w punkcie $\mathbf{x}_0$, to dla każdej krzywej ciągłej przechodzącej przez punkt $\mathbf{x}_0$ (tzn. o parametryzacji $\boldsymbol{\sigma}$ takiej, że $\boldsymbol{\sigma}(t_0) = \mathbf{x}_0$) funkcja złożona $f \circ \boldsymbol{\sigma}$ ma granicę c w punkcie t_0.

D o w ó d . Weźmy ciągłą krzywą K o parametryzacji $\boldsymbol{\sigma} : \langle a, b\rangle \to \mathbb{R}^n$ i o własności $\boldsymbol{\sigma}(t_0) = \mathbf{x}_0$ dla pewnego $t_0 \in \langle a, b\rangle$. Ciągłość tej krzywej z definicji oznacza:

$$\forall\ \epsilon_\sigma > 0\ \ \exists\ \delta_\sigma > 0 :\ 0 < |t - t_0| < \delta_\sigma \Rightarrow ||\boldsymbol{\sigma}(t) - \mathbf{x}_0|| < \epsilon_\sigma. \tag{5.21}$$

Zapiszmy fakt, że funkcja f ma granicę c w punkcie $\mathbf{x}_0$:

$$\forall \epsilon_f > 0\ \ \exists \delta_f > 0 :\ 0 < \ ||\mathbf{x} - \mathbf{x}_0|| \ < \delta_f \Rightarrow |f(\mathbf{x}) - c\,| < \epsilon_f. \tag{5.22}$$

Funkcja złożona $f \circ \boldsymbol{\sigma}$ jest funkcją jednej zmiennej i przeprowadza przedział $\langle a, b\rangle$ w $\mathbb{R}$, więc dla istnienia jej granicy w punkcie t_0 potrzebne jest takie zdanie:

$$\forall\ \epsilon > 0\ \ \exists\ \delta > 0 :\ 0 < |t - t_0| < \delta \Rightarrow |f(\boldsymbol{\sigma}(t)) - c| < \epsilon. \tag{5.23}$$

Korzystamy najpierw ze zdania (5.22) podstawiając do niego $\epsilon_f = \epsilon$. Wiemy z niego, że istnieje δ_f zapewniająca odpowiednie wynikanie. Następnie do zdania (5.21) podstawiamy $\epsilon_\sigma = \delta_f$. Wiemy z niego, że istnieje δ_σ zapewniająca odpowiednie wynikanie, które łącznie z poprzednim daje wynikanie

$$0 < |t - t_0| < \delta_\sigma \Rightarrow ||\boldsymbol{\sigma}(t) - \mathbf{x}_0|| < \epsilon_\sigma = \delta_f \Rightarrow |f(\boldsymbol{\sigma}(t)) - c\,| < \epsilon$$

potrzebne w zdaniu (5.23). A część z kwantyfikatorami też jest prawdziwa, bo już wcześniej zauważyliśmy, że do każdego ϵ istnieje $\delta = \delta_\sigma$. W takim razie całe zdanie (5.23) jest prawdziwe. ■

PRZYKŁAD 4a. Wiemy, że funkcja dana wzorem $f(x, y) = \frac{x^2}{x^2+y^2}$ ma dziedzinę $\mathbb{R}^2\backslash\{(0, 0)\}$. Jej wykres jest ukazany na Rys. 5.11. Badamy, czy ma ona granicę w zerze, czyli w punkcie $\mathbf{0}$=(0,0). Wybierzemy krzywą K o parametryzacji $\boldsymbol{\sigma}(t) = (0, t)$, czyli podążymy do zera przy $t \to 0$ po osi Y, wtedy funkcja jednej zmiennej $g(t) = (f \circ \boldsymbol{\sigma})(t) = f(0, t) = \frac{0}{t^2}$ przy $t \to 0$ ma granicę zero. Jeśli podążymy do zera po prostej $y = \alpha x$, czyli wybierzemy krzywą L o parametryzacji $\boldsymbol{\kappa}(t) = (t, \alpha t)$, to otrzymamy inną funkcję jednej zmiennej $h(t) = (f \circ \boldsymbol{\kappa})(t) = f(t, t) = t^2/(t^2 + \alpha^2 t^2) = \{\frac{1}{1+\alpha^2}\}$. Granicą tej funkcji przy $t \to 0$ jest $\frac{1}{1+\alpha^2}$. W ten sposób dla różnych $\alpha \in \mathbb{R}$ otrzymujemy różne liczby z przedziału $(0, 1\rangle$. Widzimy, że dla różnych krzywych przechodzących przez zero otrzymujemy różne granice wartości funkcji złożonej, teza Tw. 1 nie jest spełniona, więc funkcja f nie ma granicy w zerze. Rys. 5.11 to potwierdza, bo widać na nim, że punktowi (0,0) nie można przypisać jednej wartości, bo jest nad nim odcinek (0,1) osi Z.

Niech $A \subset \mathbb{R}^n$. Mówimy, że funkcja $f : A \to \mathbb{R}$ jest **ciągła w punkcie** $\mathbf{a} \in A$, jeśli

$$\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = f(\mathbf{a}),$$

czyli, gdy granica funkcji w punkcie jest równa wartości funkcji w tym punkcie.

PRZYKŁAD 6a. Funkcja

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

jest ciągła w zerze, bo sprawdziliśmy już, że jej granicą w zerze jest zero i ta granica jest równa wartości funkcji w zerze.

Funkcja z przykładu 4 nie jest ciągła, bo nie ma granicy w zerze, więc granica w zerze nie może być porównywana z wartością w zerze.

Niech $U \subset \mathbb{R}^n$ będzie zbiorem otwartym, niech $f : U \to \mathbb{R}$ będzie funkcją n zmiennych i niech $\mathbf{e}_j$ będzie jednym z wektorów bazowych (5.3). Oznaczamy

$$\frac{\partial f(\mathbf{x})}{\partial x_j} = \lim_{h\to 0} \frac{f(\mathbf{x}+h\mathbf{e}_j) - f(\mathbf{x})}{h}; \tag{5.24}$$

jeśli ta granica istnieje, to nazywamy ją **j-tą pochodną cząstkową funkcji** f albo **pochodną cząstkową funkcji** f **względem j-tej zmiennej**. Wyrażenie w liczniku po prawej stronie można zapisać inaczej

$$f(\mathbf{x}+h\mathbf{e}_j) - f(\mathbf{x}) = f(x_1, ..., x_{j-1}, x_j + h, x_{j+1}, ..., x_n) - f(x_1, ..., x_{j-1}, x_j, x_{j+1}, ..., x_n),$$

więc widzimy, że tę pochodną definiuje się jak pochodną funkcji jednej zmiennej x_j przy ustalonych pozostałych zmiennych.

PRZYKŁAD 7. Dla funkcji dwóch zmiennych $f(x,y) = \cos xy + x\cos y$ można rozważać dwie pochodne cząstkowe:

$$\frac{\partial f}{\partial x} = -y\sin xy + \cos y, \qquad \frac{\partial f}{\partial y} = -x\sin xy - x\sin y.$$

Pochodne te wyliczyliśmy korzystając z reguł różniczkowania funkcyj jednej zmiennej.

PRZYKŁAD 8. Przy funkcji dwóch zmiennych $f(x,y) = x^{1/3}y^{1/3}$ o dziedzinie $D_f = \mathbb{R}^2$ dwie pochodne cząstkowe w zerze dla ostrożności liczymy wprost z definicji:

$$\frac{\partial f(0,0)}{\partial x} = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{h^{1/3}\cdot 0 - 0}{h} = 0,$$

$$\frac{\partial f(0,0)}{\partial y} = \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} = \lim_{h\to 0} \frac{0\cdot h^{1/3} - 0}{h} = 0.$$

Na osi X dwie pochodne cząstkowe liczymy też z definicji:

$$\frac{\partial f(x_0,0)}{\partial x} = \lim_{h\to 0}\frac{f(x_0+h,0)-f(x_0,0)}{h} = \lim_{h\to 0}\frac{(x_0+h)^{1/3}\cdot 0+0}{h} = 0,$$

$$\frac{\partial f(x_0,0)}{\partial y} = \lim_{h\to 0}\frac{f(x_0,h)-f(x_0,0)}{h} = \lim_{h\to 0}\frac{x_0^{1/3}\cdot h^{1/3}}{h} = \lim_{h\to 0} x_0^{1/3}\cdot h^{-2/3},$$

a ta granica nie istnieje. Widzimy, że na osi X istnieje tylko pochodna cząstkowa względem x. Podobnie przekonamy się, że na osi Y istnieje tylko pochodna cząstkowa względem y. Poza osiami współrzędnych te pochodne istnieją i nawet są funkcjami ciągłymi, o czym wiemy z reguł różniczkowania funkcyj jednej zmiennej:

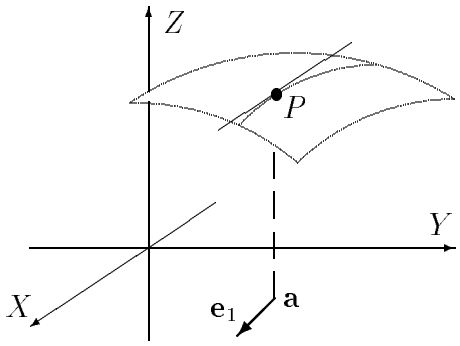
$$\frac{\partial f(x,y)}{\partial x} = \frac{1}{3}x^{-2/3}y^{1/3}, \quad \frac{\partial f(x,y)}{\partial y} = \frac{1}{3}x^{1/3}y^{-2/3}.$$

Inne oznaczenia na j-tą pochodną cząstkową funkcji $f : \mathbb{R}^n \to \mathbb{R}$:

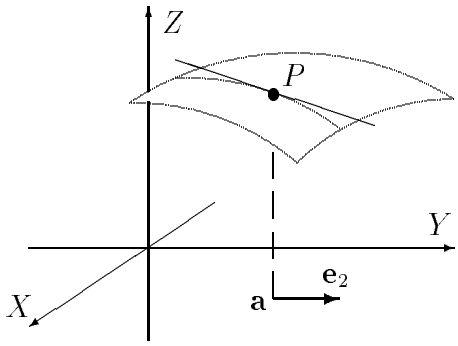
$$\partial_j f(\mathbf{x}), \quad \partial_{x_j} f(\mathbf{x}), \quad f'_{x_j}(\mathbf{x}).$$

Litera ∂ czytana jest „de kręcone". To oznaczenie wymyślono specjalnie po to, aby odróżnić ten rodzaj pochodnej od pochodnej funkcji jednej zmiennej, którą nazywa się **pochodną zwyczajną**.

W przypadku funkcji dwóch zmiennych dla interpretacji pochodnej cząstkowej względem x trzeba rozważać płaszczyznę równoległą do płaszczyzny (X,Z), przechodzącą przez punkt $P = \{\mathbf{a}, f(\mathbf{a})\}$. Przekrój tej płaszczyzny z wykresem funkcji wyznacza pewną krzywą, zob. Rys. 5.16. Pochodna cząstkowa $\partial_x f(\mathbf{a})$ jest tangensem kąta nachylenia stycznej do tej krzywej w punkcie P. Dla interpretacji pochodnej względem y trzeba rozważać płaszczyznę równoległą do płaszczyzny (Y,Z), przechodzącą przez punkt $P = \{\mathbf{a}, f(\mathbf{a})\}$. Przekrój tej płaszczyzny z wykresem funkcji wyznacza inną krzywą, zob. Rys. 5.17. Pochodna cząstkowa $\partial_x y(\mathbf{a})$ jest tangensem kąta nachylenia stycznej do tej krzywej w punkcie P.



Rysunek 5.16



Rysunek 5.17

Istnienie pochodnych cząstkowych w danym punkcie nie wystarcza do nazwania funkcji różniczkowalną w tym punkcie. Definicja tego pojęcia jest bardziej wymagająca. Niech $U \subset \mathbb{R}^n$ oraz $f : U \to \mathbb{R}$. Mówimy, że funkcja f jest **różniczkowalna w punkcie** $\mathbf{a} \in U$, jeśli dla wszystkich $j \in \{1, 2, \ldots, n\}$ istnieją pochodne cząstkowe $\partial_j f(\mathbf{a})$ oraz jeśli

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - \sum_{j=1}^n \partial_j f(\mathbf{a})\cdot(x_j - a_j)}{||\mathbf{x}-\mathbf{a}||} = 0. \tag{5.25}$$

Ta definicja jest wzięta z faktu, iż definicję pochodnej funkcji jednej zmiennej

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

można zapisać w sposób równoważny

$$\lim_{x \to a} \frac{f(x) - f(a) - f'(a)(x - a)}{x - a} = 0. \tag{5.26}$$

Jeśli funkcja jednej zmiennej f jest różniczkowalna w punkcie a, to różnicę $f(x) - f(a)$ dobrze przybliża iloczyn $f'(a)(x - a)$. Przybliża tak dobrze, że różnica między nimi dąży do zera szybciej, niż różnica argumentów $x - a$.

PRZYKŁAD 8a. Sprawdźmy, czy funkcja z Przykładu 8 jest różniczkowalna w zerze. Tym razem $\mathbf{x} = (x, y)$. Według definicji (5.25) mamy liczyć granicę

$$\lim_{\mathbf{r} \to \mathbf{0}} \frac{f(x,y) - f(0,0) - \left[\frac{\partial f}{\partial x}(0,0) \cdot (x - 0) + \frac{\partial f}{\partial y}(0,0) \cdot (y - 0)\right]}{\sqrt{(x-0)^2 + (y-0)^2}}$$

$$= \lim_{(x,y) \to (0,0)} \frac{x^{1/3}y^{1/3} - 0 - 0 \cdot x - 0 \cdot y}{\sqrt{x^2 + y^2}}.$$

Gdy dążymy do zera po osi X, otrzymujemy granicę

$$\lim_{x \to 0} \frac{x^{1/3} \cdot 0 - 0 - 0 \cdot x}{|x|} = 0.$$

Jeśli zaś podążamy po prostej $y = x$, to mamy granicę

$$\lim_{x \to 0} \frac{x^{1/3}x^{1/3} - 0 - 0 \cdot x - 0 \cdot x}{\sqrt{x^2 + x^2}} = \lim_{x \to 0} \frac{x^{2/3}}{\sqrt{2}|x|} = 2^{-1/2} \lim_{x \to 0} |x|^{-1/3},$$

a ta granica nie istnieje. W takim razie granica (5.25) nie istnieje, więc rozważana funkcja nie jest różniczkowalna w zerze. Istnienie dwóch pochodnych cząstkowych nie jest więc wystarczające dla różniczkowalności funkcji w danym punkcie.

Twierdzenie 2. Niech $U \subset \mathbb{R}^n$ oraz $f : U \to \mathbb{R}$. Jeśli pochodne cząstkowe $\partial_j f$ istnieją dla wszystkich $j \in \{1, \ldots, n\}$ i są ciągłe w punkcie $\mathbf{a}$, to funkcja f jest różniczkowalna w $\mathbf{a}$.

D o w o d u nie będzie.

Funkcja z Przykładu 8 nie ma ciągłych pochodnych cząstkowych w zerze, więc nic dziwnego, że nie jest tam różniczkowalna.

PRZYKŁAD 9. Niech $f(x, y) = xy^2 e^{xy}$. Ta funkcja ma za dziedzinę całą płaszczyznę (X, Y). Pochodne cząstkowe liczymy według reguł różniczkowania funkcyj jednej zmiennej:

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(xy^2)e^{xy} + xy^2 \frac{\partial}{\partial x}(e^{xy}) = y^2 e^{xy} + xy^2(ye^{xy}) = y^2(1 + xy)e^{xy},$$

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(xy^2)e^{xy} + xy^2 \frac{\partial}{\partial y}(e^{xy}) = 2xye^{xy} + xy^2(xe^{xy}) = xy(2 + xy)e^{xy}.$$

Pochodne te istnieją na całej dziedzinie i są ponadto funkcjami ciągłymi, zatem funkcja f jest różniczkowalna.

Teraz czas na definicję następnego pojęcia. Jeśli funkcja f jest różniczkowalna w punkcie $\mathbf{x}_0$, to układ pochodnych cząstkowych

$$\{\partial_1 f(\mathbf{x}_0),\ \ldots, \partial_n f(\mathbf{x}_0)\} = \nabla f(\mathbf{x}_0)$$

stanowi wektor w $\mathbb{R}^n$, który nazywamy **gradientem funkcji** f i oznaczamy $\nabla f = \operatorname{grad} f$. Trójkąt odwrócony wierzchołkiem na dół czytamy jako **nabla**. Ta nazwa wzięła się z nazwy strunowego instrumentu hinduskiego. Gradient zapisuje się również za pomocą wektorów bazowych:

$$\nabla f = \operatorname{grad} f = \sum_{j=1}^{n} \mathbf{e}_j \frac{\partial f}{\partial x_j}.$$

W przestrzeni trójwymiarowej ten wzór znany jest w postaci

$$\operatorname{grad} f = \mathbf{i}\frac{\partial f}{\partial x} + \mathbf{j}\frac{\partial f}{\partial y} + \mathbf{k}\frac{\partial f}{\partial z}$$

z trzema prostopadłymi wektorami jednostkowymi $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$.

We wzorze (5.25) rozpoznajemy współrzędne gradientu, więc możemy ten wzór zapisać za pomocą skaloczynu gradientu z wektorem $\mathbf{r} - \mathbf{a}$:

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - [\nabla f(\mathbf{a})]\cdot(\mathbf{x}-\mathbf{a})}{||\mathbf{x}-\mathbf{a}||} = 0. \tag{5.27}$$

Teraz lepiej widać podobieństwo do wzoru (5.26).

Niech będzie dana krzywa K o parametryzacji $\boldsymbol{\sigma} : \langle a, b\rangle \to \mathbb{R}^n$. Jeśli $U \subset \mathbb{R}^n$ i mamy funkcję $f : U \to \mathbb{R}$, a funkcja $\boldsymbol{\sigma}$ ma obraz zawarty w U, co można zapisać $K \subset U$, to można rozważać funkcję złożoną $f \circ \boldsymbol{\sigma} : \langle a, b\rangle \to \mathbb{R}$:

$$(f \circ \boldsymbol{\sigma})(t) = f(\boldsymbol{\sigma}(t)) = f(\sigma_1(t), \sigma_2(t),\ \ldots, \sigma_n(t)).$$

Funcja złożona jest funkcją jednej zmiennej. Przy liczeniu pochodnej zwyczajnej funkcji złożonej trzeba uwzględnić obecność zmiennej t w każdym z n argumentów funkcji f, więc wystąpi n iloczynów pochodnych funkcji zewnętrznej i wewnętrznej:

$$\frac{d(f\circ\boldsymbol{\sigma})}{dt} = \frac{\partial f}{\partial x_1}\frac{d\sigma_1}{dt} + \frac{\partial f}{\partial x_2}\frac{d\sigma_2}{dt} + \ \ldots\ + \frac{\partial f}{\partial x_n}\frac{d\sigma_n}{dt} = \sum_{j=1}^{n}\frac{\partial f}{\partial x_j}\frac{d\sigma_j}{dt}. \tag{5.28}$$

Ten wzór można zapisać krócej za pomocą gradientu i skaloczynu:

$$\frac{d(f\circ\boldsymbol{\sigma})}{dt} = \nabla f \cdot \frac{d\boldsymbol{\sigma}}{dt}. \tag{5.29}$$

Dwa ostatnie wzory są bardzo ważne w rachunkach przybliżonych, które matematycy uważają za nieścisłe. Jeśli równość (5.28) pomnożymy stronami przez dt, to otrzymamy różniczkę funkcji złożonej

$$d(f\circ\boldsymbol{\sigma}) = \frac{\partial f}{\partial x_1}d\sigma_1 + \frac{\partial f}{\partial x_2}d\sigma_2 + \ \ldots\ + \frac{\partial f}{\partial x_n}d\sigma_n.$$

Różniczka $d\sigma_j$ oznacza dowolnie mały przyrost współrzędnej x_j wektora wodzącego, więc można to zapisać inaczej:

$$df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \ \ldots\ + \frac{\partial f}{\partial x_n}dx_n.$$

Fizycy nazywają to wyrażenie **różniczką zupełną** funkcji wielu zmiennych i najczęściej wykorzystują w rachunku błędów w postaci

$$\Delta f = \frac{\partial f}{\partial x_1}\Delta x_1 + \frac{\partial f}{\partial x_2}\Delta x_2 + \ldots + \frac{\partial f}{\partial x_n}\Delta x_n.$$

Niech $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}$. **Pochodną dążnościową** funkcji f w punkcie $\mathbf{a}$ **z dążnością** $\mathbf{v}$ nazywamy

$$\lim_{h\to 0}\frac{f(\mathbf{a}+h\mathbf{v})-f(\mathbf{a})}{h}, \tag{5.30}$$

jeśli ta granica istnieje. Pochodną dążnościową w punkcie $\mathbf{a}$ oznaczamy $D_{\mathbf{v}}f(\mathbf{a})$. We wzorze (5.30) można zobaczyć złożenie funkcji f z parametryzacją pewnej szczególnej krzywej $\boldsymbol{\sigma}(t) = \mathbf{a} + t\mathbf{v}$. Ta krzywa jest linią prostą, a jeśli uznamy, że parametr t jest czasem, to ta parametryzacja odpowiada ruchowi jednostajnemu z prędkością $\mathbf{v}$. Wyrażenie $f \circ \boldsymbol{\sigma}$ to zamiana funkcji n zmiennych na funkcję jednej zmiennej. Pochodna dążnościowa to pochodna tej nowej funkcji w zerze:

$$D_{\mathbf{v}}f(\mathbf{a}) = \frac{d}{dt}f(\mathbf{a}+t\mathbf{v})|_{t=0}. \tag{5.31}$$

Można powiedzieć, że jest to szybkość zmiany wielkości f, jeśli cząstka przechodzi przez punkt $\mathbf{a}$ z prędkością $\mathbf{v}$. W tym zapisie widać zależność od wektora $\mathbf{v}$. Im większa będzie prędkość $\mathbf{v}$, tym szybciej będzie cząstka przechodzić przez punkt $\mathbf{a}$, więc tym większa będzie szybkość zmiany wielkości f, czyli większa będzie pochodna dążnościowa.

Lemat 1. Jeśli funkcja $f : U \to \mathbb{R}$ jest różniczkowalna w punkcie $\mathbf{a}$, to

$$D_{\mathbf{v}}f(\mathbf{a}) = \operatorname{grad} f(\mathbf{a}) \cdot \mathbf{v}.$$

D o w ó d. Pochodną dążnościową wyrażamy wzorem (5.31). Widzimy, że pod argumentem funkcji f stoi parametryzacja prostej $\boldsymbol{\sigma}(t) = \mathbf{a} + t\mathbf{v}$. Jej pochodna wynosi $d\boldsymbol{\sigma}/dt = \mathbf{v}$. Wystarczy to podstawić do wzoru (5.29)

$$D_{\mathbf{v}}f(\mathbf{a}) = \left.\frac{d(f\circ\boldsymbol{\sigma})}{dt}\right|_{t=0} = \operatorname{grad} f(\mathbf{a}) \cdot \left.\frac{d\boldsymbol{\sigma}}{dt}\right|_{t=0} = \operatorname{grad} f(\mathbf{a}) \cdot \mathbf{v}. \quad \blacksquare$$

Ten wynik można zapisać za pomocą nabli: $D_{\mathbf{v}}f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \mathbf{v}$. Iloczyn skalarny jest przemienny, więc można to zapisać jeszcze inaczej:

$$D_{\mathbf{v}}f(\mathbf{a}) = \mathbf{v} \cdot \nabla f(\mathbf{a}).$$

Z tego wzoru widać, że pochodna dążnościowa jest liniowa względem $\mathbf{v}$, tzn. **jednorodna**:

$$D_{\alpha\mathbf{v}}f = (\alpha\mathbf{v}) \cdot \nabla f = \alpha\,(\mathbf{v} \cdot \nabla f) = \alpha\, D_{\mathbf{v}}f$$

(tu potwierdza się to, co wspomnieliśmy wcześniej: im większa prędkość $\mathbf{v}$, tym większa pochodna dążnościowa) oraz **addytywna**:

$$D_{\mathbf{v}_+\mathbf{v}_2}f = (\mathbf{v}_1 + \mathbf{v}_2) \cdot \nabla f = \mathbf{v}_2 \cdot \nabla f + \mathbf{v}_1 \cdot \nabla f = D_{\mathbf{v}_1}f(\mathbf{a}) + D_{\mathbf{v}_2}f(\mathbf{a}).$$

Z wzoru na wartość iloczynu skalarnego mamy

$$\mathbf{v} \cdot \nabla f = ||\mathbf{v}|| \, ||\nabla f \, || \cos \alpha,$$

Przy ustalonych długościach czynników ten skaloczyn jest największy dla $\alpha = 0$. Jeśli wybierzemy wektor $\mathbf{v}$ o długości takiej samej jak gradient funkcji, to otrzymamy

$$\mathbf{v} \cdot \nabla f = ||\nabla f|| \, ||\nabla f \, || \cos \alpha,$$

Podstawmy za wektor $\mathbf{v}$ dokładnie gradient funkcji, wtedy $\alpha = 0$ oraz $\cos \alpha = 1$, skąd

$$D_{\nabla f} f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \nabla f(\mathbf{a}) = ||\nabla f(\mathbf{a})||^2.$$

To oznacza, że dla różnych wektorów $\mathbf{v}$ o długości takiej samej jak $\operatorname{grad} f$ pochodna dążnościowa jest największa dla $\mathbf{v} = \nabla f$. Stąd wyciągamy wniosek, że gradient jest wektorem, który pokazuje kierunek najszybszego wzrostu funkcji.

PRZYKŁAD 10. Niech $f(x, y) = x^2 - 4xy$. Znajdziemy gradient tej funkcji w punkcie $\mathbf{r}_1 = (3, -2)$:

$$\nabla f(x, y) = (2x - 4y, -4x).$$

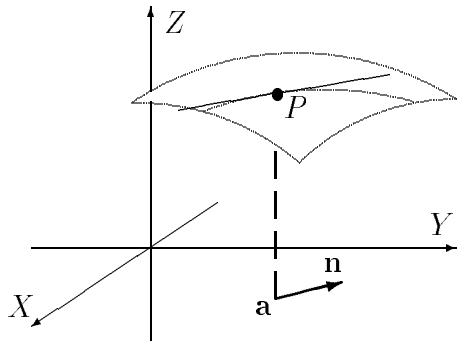
W takim razie we wskazanym punkcie $\nabla f(3, -2) = (6 + 8, -12) = (14, -12)$. Teraz wyliczymy pochodną dążnościową z wektorem łączącym punkty $\mathbf{r}_1$ i $\mathbf{r}_2 = (4, 1)$. Otóż $\mathbf{v} = \mathbf{r}_2 - \mathbf{r}_1 = (4 - 3, 1 - (-2)) = (1, 3)$, więc

$$\mathbf{v} \cdot \nabla f(3, -2) = (1, 3) \cdot (14, -12) = 14 - 36 = -22.$$

Jeśli wybierzemy $\mathbf{v} = \mathbf{n}$, gdzie $||\mathbf{n}|| = 1$, to pochodna dążnościowa $D_{\mathbf{n}} f$ nazywa się **pochodną kierunkową**. Pochodna cząstkowa jest jej szczególnym przypadkiem:

$$D_{\mathbf{e}_j} f = \mathbf{e}_j \cdot \nabla f = \frac{\partial f}{\partial x_j}.$$

Pochodna kierunkowa opisuje szybkość zmiany funkcji f w kierunku $\mathbf{n}$. Niech f będzie funkcją dwóch zmiennych . Weźmy płaszczyznę równoległą do wektorów $\mathbf{e}_3$ i $\mathbf{n}$ oraz przechodzącą przez punkt $P = \{\mathbf{a}, f(\mathbf{a})\}$. Przekrój tej płaszczyzny z wykresem funkcji wyznacza pewną krzywą, zob. Rys. 5.18. Pochodna kierunkowa $\mathbf{n} \cdot \nabla f(\mathbf{a})$ jest tangensem kąta nachylenia stycznej do tej krzywej w punkcie P.



Rysunek 5.18

PRZYKŁAD 10a. Dla tej samej funkcji, co w Przykładzie 10 znajdziemy pochodną kierunkową w kierunku wektora $\mathbf{v} = (1,3)$. Wprowadzamy wektor jednostkowy o potrzebnym kierunku:

$$\mathbf{n} = \frac{\mathbf{v}}{||\mathbf{v}||} = \frac{1}{\sqrt{1^2+3^2}}(1,3) = \frac{1}{\sqrt{10}}(1,3)$$

i obliczamy pochodną kierunkową

$$\mathbf{n}\cdot\nabla f(3,-2) = \frac{1}{\sqrt{10}}\mathbf{v}\cdot\nabla f(3,-2) = \frac{1}{\sqrt{10}}(-22) = -\frac{22}{\sqrt{10}} \approx -6,96.$$

Twierdzenie 3. Niech $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}$, niech S będzie poziomicą funkcji f i niech $\mathbf{x}_0 \in S$. Jeśli $\boldsymbol{\sigma}(t)$ jest krzywą leżącą na S taką, że $\boldsymbol{\sigma}(0) = \mathbf{x}_0$, to

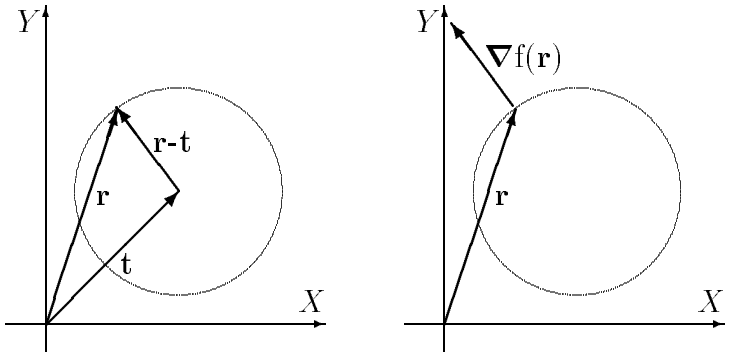
$$\mathbf{v}\cdot\nabla f(\mathbf{x}_0) = 0 \tag{5.32}$$

dla $\mathbf{v} = \boldsymbol{\sigma}'(0)$.

D o w ó d . Jeśli krzywa $\boldsymbol{\sigma}$ leży całkowicie na poziomicy S, to funkcja jednej zmiennej $(f\circ\boldsymbol{\sigma})(t)$ jest stała, więc jej pochodna jest zerem. Korzystamy ze wzoru (5.29):

$$0 = \frac{d}{dt}f(\boldsymbol{\sigma}(t))|_{t=0} = \nabla f\cdot\boldsymbol{\sigma}'(0) = \nabla f\cdot\mathbf{v}. \ \blacksquare$$

To twierdzenie jest bardzo ważne, bo mówi, że gradient jako wektor jest prostopadły do poziomic. Skoro krzywa $\boldsymbol{\sigma}$ leży na poziomicy, to wektor $\mathbf{v}$ jako styczny do krzywej jest też styczny do poziomicy. Skoro poza tym krzywa jest dowolna, to możemy uważać, że wektor $\mathbf{v}$ może być dowolnym wektorem stycznym do poziomicy. Równość (5.32) oznacza prostopadłość wektorów ∇f i $\mathbf{v}$. Gradient jest więc prostopadły do każdego wektora stycznego do poziomicy, a tak właśnie rozumie się prostopadłość do poziomicy.



Rysunek 5.19

PRZYKŁAD 11. Funkcja $f(x,y,z) = \frac{1}{2}[(x-a)^2+(y-b)^2+(z-c)^2]$ jako swoje poziomice ma sfery o wspólnym środku w punkcie $\mathbf{t} = (a,b,c)$. Na Rys. 5.19 – niestety płaskim – ukazaliśmy jedną taką poziomicę jako okrąg o środku w punkcie $\mathbf{t}$. Wyliczmy gradient tej funkcji

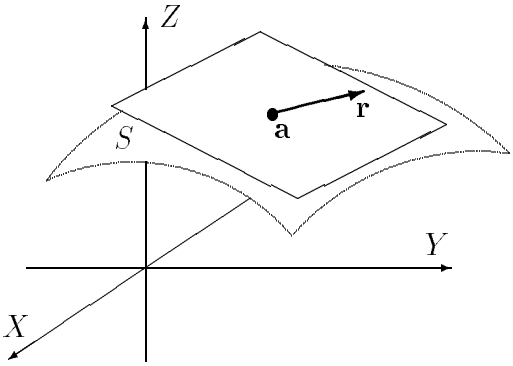
$$\nabla f(\mathbf{r}) = \frac{1}{2}[2(x-a),\ 2(y-b),\ 2(z-c)] = (x,y,z)-(a,b,c) = \mathbf{r}-\mathbf{t}.$$

Stwierdzamy, że gradient jest wektorem $\mathbf{r}-\mathbf{t}$. A ten wektor jest prostopadły do sfery w punkcie $\mathbf{r}$, bo jest to wektor łączący środek sfery z punktem $\mathbf{r}$. Aby ukazać go jako gradient, trzeba go zaczepić w punkcie $\mathbf{r}$. W miarę oddalania od wspólnego środka sfer gradienty są coraz większe, bo większy jest wektor $\mathbf{r}-\mathbf{t}$.

Niech $U \subset \mathbb{R}^3$ $f : U \to \mathbb{R}$. Wiemy, że dla takiej funkcji poziomicami są na ogół powierzchnie. Zbiór końców wszystkich wektorów stycznych do powierzchni w ustalonym punkcie stanowi płaszczyznę. Jeśli S jest poziomicą funkcji f oraz $\mathbf{a} \in S$, to na mocy tego, co już zauważyliśmy, wektor $\mathbf{r} - \mathbf{a}$ jest styczny do S w punkcie $\mathbf{a}$ w.t.w. gdy jest prostopadły do ∇f:

$$\nabla f(\mathbf{a}) \cdot (\mathbf{r} - \mathbf{a}) = 0. \tag{5.33}$$

Wektor $\mathbf{r}$ ma swój koniec na płaszczyźnie stycznej do S, zob. Rys. 5.20. W takim razie mamy prawo nazywać równość (5.33) **równaniem płaszczyzny stycznej do poziomicy** funkcji f w punkcie $\mathbf{a}$.



Rysunek 5.20

Niech $U \subset \mathbb{R}^n$ i niech funkcja $f : U \to \mathbb{R}$ będzie różniczkowalna w każdym punkcie $\mathbf{x} \in U$. Z definicji oznacza to, że istnieją pochodne cząstkowe $\partial_j f$ i one również są funkcjami z U w $\mathbb{R}$. Jeśli wszystkie pochodne cząstkowe są funkcjami ciągłymi, to mówimy, że funkcja f **jest klasy** C^1 i piszemy $f \in C^1$. Możemy rozważać z kolei ich pochodne:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \partial_i \partial_j f = f''_{x_j x_i},$$

które są **drugimi pochodnymi cząstkowymi funkcji** f albo **pochodnymi cząstkowymi drugiego rzędu**. Uwaga: najpierw liczona jest pochodna napisana bliżej funkcji f, czyli w tym wzorze względem x_j. Tych drugich pochodnych jest n^2. Jeśli wszystkie drugie pochodne są ciągłe, to mówimy, ze funkcja f **jest klasy** C^2.

Twierdzenie 4. Jeśli $U \subset \mathbb{R}^n$, a $f : U \to \mathbb{R}$ jest klasy C^2, to drugie pochodne mieszane są przemienne, tzn.

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

D o w ó d. Wprowadźmy oznaczenie

$$S(\Delta x_i, \Delta x_j) =$$

$$= f(x_1, \ldots, x_{i0} + \Delta x_i, \ldots, x_{j0} + \Delta x_j, \ldots, x_n) - f(x_1, \ldots, x_{i0}, \ldots, x_{j0} + \Delta x_j, \ldots, x_n)$$
$$- f(x_1, \ldots, x_{i0} + \Delta x_i, \ldots, x_{j0}, \ldots, x_n) + f(x_1, \ldots, x_{i0}, \ldots, x_{j0}, \ldots, x_n).$$

Jeśli x_{j0} i Δx_j są ustalone, to definiujemy fukcję jednej zmiennej

$$g(x_i) = f(x_1, \ldots, x_i, \ldots, x_{j0} + \Delta x_j, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_{j0}, \ldots, x_n). \tag{5.34}$$

Wtedy możemy napisać

$$S(\Delta x_i, \Delta x_j) = g(x_{i0} + \Delta x_i) - g(x_{i0}).$$

Z twierdzenia Lagrange'a o wartości pośredniej (Tw. 3 w § 2.2) mamy

$$g(x_{i0} + \Delta x_i) - g(x_{i0}) = g'(\tilde{x}_i)\Delta x_i, \qquad \text{dla} \quad \tilde{x}_i \in (x_{i0}, x_{i0} + \Delta x_i).$$

Wykorzystujemy wzór (5.34):

$$S(\Delta x_i, \Delta x_j) = \left[\frac{\partial f}{\partial x_i}(\,.\,.\,, \tilde{x}_i, \,\ldots, x_{j0} + \Delta x_j \ldots) - \frac{\partial f}{\partial x_i}(\ldots \tilde{x}_i \ldots, x_{j0} \ldots)\right] \Delta x_i.$$

Stosujemy drugi raz twierdzenie o wartości pośredniej:

$$S(\Delta x_i, \Delta x_j) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\,.\,.\,, \tilde{x}_i, \,\ldots, \tilde{x}_j \ldots)\Delta x_i \Delta x_j.$$

Stąd

$$\frac{S(\Delta x_i, \Delta x_j)}{\Delta x_i \Delta x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}(\,.\,.\,, \tilde{x}_i, \,\ldots, \tilde{x}_j \ldots).$$

Z założenia drugie pochodne cząstkowe są ciągłe, więc

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\,.\,.\,, x_{i0}, \,\ldots, x_{j0} \ldots) = \lim_{(\Delta x_i, \Delta x_j) \to (0,0)} \frac{S(\Delta x_i, \Delta x_j)}{\Delta x_i \Delta x_j}.$$

Jeśli inaczej określimy funkcję g we wzorze (5.34) z jednakowym argumentem x_j, to podobnie pokażemy, że

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\,.\,.\,, x_{i0}, \,\ldots, x_{j0} \ldots) = \lim_{(\Delta x_i, \Delta x_j) \to (0,0)} \frac{S(\Delta x_i, \Delta x_j)}{\Delta x_j \Delta x_i}.$$

Skoro mnożenie jest przemienne i prawe strony są równe, to lewe tak samo i otrzymujemy

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\,.\,.\,, x_{i0}, \,\ldots, x_{j0} \ldots) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\,.\,.\,, x_{i0}, \,\ldots, x_{j0} \ldots). \quad \blacksquare$$

5.3 Funkcje o wartościach wektorowych

Niech $m, n \in \mathbb{N}$ i $U \subset \mathbb{R}^n$. Przez **funkcję wielu zmiennych o wartościach wektorowych** rozumiemy funkcję $\mathbf{f} : U \to \mathbb{R}^m$.Będziemy też zapisywać

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \,\ldots, f_m(\mathbf{x}))\,,$$

gdzie **współrzędne** f_k są funkcjami n zmiennych o wartościach rzeczywistych (skalarnych).

Mając pojęcie odległości w $\mathbb{R}^n$ i w $\mathbb{R}^m$ możemy w analogii do funkcyj o wartościach rzeczywistych zdefiniować granicę funkcji w punkcie i ciągłość funkcji. Funkcja $\mathbf{f} : U \to \mathbb{R}^m$ dla $U \subset \mathbb{R}^n$ **ma granicę b w punkcie a**, jeśli

$$\forall \epsilon > 0 \;\; \exists\, \delta > 0 : \;\; 0 < \|\, \mathbf{x} - \mathbf{a} \,\| < \delta \Rightarrow \|\, \mathbf{f}(\mathbf{x}) - \mathbf{b} \,\| < \epsilon. \tag{5.35}$$

Piszemy wtedy

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{b}.$$

Widzimy, że jest to dostosowanie definicji Cauchy'ego granicy funkcji do przypadku wielu wymiarów.

Twierdzenie 1. Niech $A \subset \mathbb{R}^n$. Jeśli funkcja $\mathbf{f} : A \to \mathbb{R}^m$ ma granicę $\mathbf{b}$ w punkcie $\mathbf{x}_0$, to dla każdej krzywej ciągłej przechodzącej przez punkt $\mathbf{x}_0$ (tzn. o parametryzacji $\boldsymbol{\sigma}$ takiej, że $\boldsymbol{\sigma}(t_0) = \mathbf{x}_0$) funkcja złożona $\mathbf{f} \circ \boldsymbol{\sigma}$ ma granicę $\mathbf{b}$ w punkcie t_0.

D o w ó d jest powtórzeniem dowodu Tw. 1 z poprzedniego paragrafu z zamianą modułu na normę.

Lemat 1 (granica funkcji pomnożonej przez stałą). Niech $U \subset \mathbb{R}^n$, niech funkcja $\mathbf{f} : U \to \mathbb{R}^m$ ma granicę w punkcie $\mathbf{a}$ i niech $c \in \mathbb{R}$. Wtedy funkcja $\mathbf{h}(\mathbf{x}) = c\mathbf{f}(\mathbf{x})$ ma granicę w $\mathbf{a}$ oraz

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{h}(\mathbf{x}) = c \lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}).$$

D o w ó d. Dla $c = 0$ teza lematu jest w sposób oczywisty prawdziwa, bo funkcja $\mathbf{h}$ jest wtedy stała i równa zero. Niech więc $c \neq 0$. Mamy pokazać odpowiednik zdania (5.35) dla funkcji $\mathbf{h}$ przy założeniu prawdziwości samego zdania (5.35). Potrzebna nierówność $\| \mathbf{h} - c\mathbf{b} \| < \epsilon$ wynika z (5.35), gdy zamiast ϵ napisać tam $\epsilon/|c|$:

$$\| \mathbf{f}(\mathbf{x}) - \mathbf{b} \| < \frac{\epsilon}{|c|} \;\Rightarrow\; |c| \, \| \mathbf{f}(\mathbf{x}) - \mathbf{b} \| < \epsilon \;\Rightarrow\; \| c\mathbf{f}(\mathbf{x}) - c\mathbf{b} \| < \epsilon \;\Rightarrow\; \| \mathbf{h}(\mathbf{x}) - c\mathbf{b} \| < \epsilon.$$

Razem z kwantyfikatorami możemy więc napisać

$$\forall \epsilon > 0 \;\; \exists\, \delta > 0 : \;\; 0 < \| \mathbf{x} - \mathbf{a} \| < \delta \Rightarrow \| \mathbf{h}(\mathbf{x}) - c\mathbf{b} \| < \epsilon,$$

co oznacza, że funkcja $\mathbf{h}$ ma granicę w $\mathbf{a}$ oraz

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{h}(\mathbf{x}) = c\mathbf{b} = c \lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}). \;\blacksquare$$

Lemat 2 (granica sumy funkcyj). Niech $U \subset \mathbb{R}^n$, niech funkcje $\mathbf{f}, \mathbf{g} : U \to \mathbb{R}^m$ mają granice w punkcie $\mathbf{a}$. Wtedy funkcja $\mathbf{h}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})$ ma granicę w $\mathbf{a}$ oraz

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{h}(\mathbf{x}) = \lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) + \lim_{\mathbf{x}\to\mathbf{a}} \mathbf{g}(\mathbf{x}).$$

D o w ó d. Piszemy odpowiedniki zdania (5.35) dla funkcyj $\mathbf{f}$ i $\mathbf{g}$:

$$\forall \epsilon > 0 \;\; \exists\, \delta_1 > 0 : \;\; 0 < \| \mathbf{x} - \mathbf{a} \| < \delta_1 \;\Rightarrow \| \mathbf{f}(\mathbf{x}) - \mathbf{b} \| < \frac{\epsilon}{2}, \tag{5.36}$$

$$\forall \epsilon > 0 \;\; \exists\, \delta_2 > 0 : \;\; 0 < \| \mathbf{x} - \mathbf{a} \| < \delta_2 \;\Rightarrow \| \mathbf{g}(\mathbf{x}) - \mathbf{d} \| < \frac{\epsilon}{2}. \tag{5.37}$$

Dzięki nierówności trójkąta możemy napisać

$$\| \mathbf{h}(\mathbf{x}) - (\mathbf{b} + \mathbf{d}) \| = \| \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) - \mathbf{b} - \mathbf{d} \| = \| \mathbf{f}(\mathbf{x}) - \mathbf{b} + \mathbf{g}(\mathbf{x}) - \mathbf{d} \|$$

$$\leq \| \mathbf{f}(\mathbf{x}) - \mathbf{b} \| + \| \mathbf{g}(\mathbf{x}) - \mathbf{d} \| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

jeśli prawdziwe są oba zdania (5.36) i (5.37). A one są prawdziwe dla $\delta = \min\{\delta_1, \delta_2\}$. W takim razie razem z kwantyfikatorami można napisać

$$\forall \epsilon > 0 \ \ \exists\, \delta > 0 : \ 0 < \| \mathbf{x} - \mathbf{a} \| < \delta \ \Rightarrow \ \| \mathbf{h}(\mathbf{x}) - \mathbf{b} - \mathbf{d} \| < \epsilon,$$

co oznacza, że funkcja $\mathbf{h}$ ma granicę w $\mathbf{a}$ oraz

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{h}(\mathbf{x}) = \mathbf{b} + \mathbf{d} = \lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) + \lim_{\mathbf{x}\to\mathbf{a}} \mathbf{g}(\mathbf{x}). \ \blacksquare$$

Niech $U \subset \mathbb{R}^n$. Mówimy, że funkcja $\mathbf{f} : U \to \mathbb{R}^m$ jest **ciągła w punkcie** $\mathbf{a} \in U$, jeśli

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a}),$$

czyli, gdy granica funkcji w punkcie jest równa wartości funkcji w tym punkcie.

Twierdzenie 2. Funkcja $\mathbf{f} : U \to \mathbb{R}^m$ jest ciągła w punkcie $\mathbf{a} \in U$ w.t.w. gdy każda współrzędna $f_k : U \to \mathbb{R}$ jest ciągła w punkcie $\mathbf{a}$.

D o w ó d jest odpowiednią modyfikacją dowodu Tw. 2 z § 5.1.

W przypadku różniczkowania rozważanych w tym paragrafie funkcyj z $\mathbb{R}^n$ w $\mathbb{R}^m$ trzeba wprowadzić pochodne cząstkowe poszczególnych funkcyj-współrzędnych względem poszczególnych zmiennych

$$\frac{\partial f_k}{\partial x_j}, \quad k \in \{1, ..., m\}, \ j \in \{1, ..., n\},$$

które najwygodniej jest ułożyć w następującą macierz $m \times n$

$$D\mathbf{f}(\mathbf{a}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \frac{\partial f_1}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{a}) & \frac{\partial f_2}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_2}{\partial x_n}(\mathbf{a}) \\ \dots\dots & \dots\dots & \dots & \dots\dots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \frac{\partial f_m}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{pmatrix}, \tag{5.38}$$

którą będziemy nazywać **pochodną macierzową funkcji f**. Wtedy zapis $D\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a})$ oznacza mnożenie macierzy $D\mathbf{f}(\mathbf{a})$ przez wektor-kolumnę $(\mathbf{x} - \mathbf{a})^T$ o n elementach:

$$D\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \frac{\partial f_1}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{a}) & \frac{\partial f_2}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_2}{\partial x_n}(\mathbf{a}) \\ \dots\dots & \dots\dots & \dots & \dots\dots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \frac{\partial f_m}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{pmatrix} \begin{pmatrix} x_1 - a_1 \\ x_2 - a_2 \\ \dots\dots \\ x_n - a_n \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a})(x_1 - a_1) + \frac{\partial f_1}{\partial x_2}(\mathbf{a})(x_2 - a_2) + \ \dots \ + \frac{\partial f_1}{\partial x_n}(\mathbf{a})(x_n - a_n) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{a})(x_1 - a_1) + \frac{\partial f_2}{\partial x_2}(\mathbf{a})(x_2 - a_2) + \ \dots \ + \frac{\partial f_2}{\partial x_n}(\mathbf{a})(x_n - a_n) \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a})(x_1 - a_1) + \frac{\partial f_m}{\partial x_2}(\mathbf{a})(x_2 - a_2) + \ \dots \ + \frac{\partial f_m}{\partial x_n}(\mathbf{a})(x_n - a_n) \end{pmatrix},$$

co daje wektor-kolumnę o m elementach.

PRZYKŁAD 1. Niech funkcja $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^2$ będzie dana wzorem $\mathbf{f}(x, y, z) = (z\mathrm{e}^x, -y\mathrm{e}^z)$. Jest to funkcja ciągła, bo jej obie funkcje-współrzędne są funkcjami ciągłymi. Macierz jej pochodnych cząstkowych ma postać

$$D\mathbf{f}(x, y, z) = \begin{pmatrix} z\mathrm{e}^x & 0 & \mathrm{e}^x \\ 0 & -\mathrm{e}^z & -y\mathrm{e}^z \end{pmatrix}.$$

Te pochodne są są funkcjami ciągłymi, więc $\mathbf{f} \in C^1$.

Dla $m = 1$ nie piszemy tłustego f, bo to jest funkcja o wartościach rzeczywistych; w tym przypadku macierz Df jest jednowierszowa:

$$Df(\mathbf{a}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}),\ \frac{\partial f}{\partial x_2}(\mathbf{a}),\ \ldots,\ \frac{\partial f}{\partial x_m}(\mathbf{a})\right).$$

i dlatego pochodna macierzowa takiej funkcji pokrywa się z gradientem: $Df(\mathbf{a}) = \boldsymbol{\nabla} f(\mathbf{a})$.

Wspomniane wcześniej krzywe $\boldsymbol{\sigma} : \mathbb{R} \to \mathbb{R}^m$ są funkcjami rozważanego obecnie typu, chociaż jako funkcje jednej zmiennej nie mają pochodnych cząstkowych. Dla nich macierz (5.38) należałoby zapisać jako kolumnę

$$D\boldsymbol{\sigma}(a) = \begin{pmatrix} \frac{d\sigma_1}{dx}(a) \\ \\ \frac{d\sigma_2}{dx}(a) \\ \\ \cdots\cdots \\ \\ \frac{d\sigma_m}{dx}(a) \end{pmatrix} = \left(\frac{d\boldsymbol{\sigma}}{dx}(a)\right)^T.$$

Ten wektor w $\mathbb{R}^m$ można interpretować jako prędkość posuwania się po krzywej.

PRZYKŁAD 2. Niech funkcja $\boldsymbol{\sigma} : \mathbb{R} \to \mathbb{R}^3$ będzie dana wzorem $\boldsymbol{\sigma}(t) = (r\cos\omega t,\ r\sin\omega t,\ vt)$. Jest to funkcja ciągła. Opisuje ona **linię śrubową** (**helisę**) nawijającą się wokół trzeciej osi. Macierz jej pochodnych jest następującą kolumną

$$D\boldsymbol{\sigma}(t) = \begin{pmatrix} -r\omega\sin\omega t \\ r\omega\cos\omega t \\ v \end{pmatrix}.$$

Wyliczone pochodne są funkcjami ciągłymi, więc jest to funkcja klasy C^1.

Niech $U \subset \mathbb{R}^n$ oraz $\mathbf{f} : U \to \mathbb{R}^m$. Mówimy, że funkcja $\mathbf{f}$ jest**różniczkowalna w punkcie** $\mathbf{a} \in U$, jeśli dla wszystkich $j \in \{1, 2, \ldots, n\}, k \in \{1, \ldots, m\}$ istnieją pochodne cząstkowe $\partial_j f_k(\mathbf{a})$ oraz jeśli

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - D\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a})}{||\,\mathbf{x} - \mathbf{a}\,||} = 0. \tag{5.39}$$

Jest to z niewielką zmianą przepisanie wzoru (5.27) odpowiadającego definicji (5.25). Podobnie da się sformułować odpowiednik Tw. 2 z poprzedniego paragrafu.

Twierdzenie 3. Niech $U \subset \mathbb{R}^n$ oraz $\mathbf{f} : U \to \mathbb{R}^m$. Jeśli pochodne cząstkowe $\partial_j f_k$ istnieją dla wszystkich $j \in \{1, \dots, n\}$, $k \in \{1, \dots, m\}$ i są ciągłe w punkcie $\mathbf{a}$, to funkcja $\mathbf{f}$ jest różniczkowalna w $\mathbf{a}$.

D o w ó d można znaleźć w książce: J.E. Marsden, A.J. Tromba *Vector Calculus*, Freeman and Company, New York 1988, w § 2.7.

Lemat 3 (*pochodna funkcji pomnożonej przez stałą*). Niech $U \subset \mathbb{R}^n$, niech funkcja $\mathbf{f} : U \to \mathbb{R}^m$ będzie różniczkowalna w punkcie $\mathbf{a}$ i niech $c \in \mathbb{R}$. Wtedy funkcja $\mathbf{h}(\mathbf{x}) = c\mathbf{f}(\mathbf{x})$ jest różniczkowalna w $\mathbf{a}$ oraz

$$D\mathbf{h}(\mathbf{a}) = cD\mathbf{f}(\mathbf{a}).$$

D o w ó d. Mamy wykazać, że

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{a}) - cD\mathbf{f}(\mathbf{a})(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} = 0, \tag{5.40}$$

czyli że

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{c\mathbf{f}(\mathbf{x}) - c\mathbf{f}(\mathbf{a}) - cD\mathbf{f}(\mathbf{a})(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} = 0.$$

Stałą c można wyciągnąć przed ułamek dzięki własnościom dodawania wektorów, a potem dzięki Lematowi 1 przed znak granicy. W takim razie mamy otrzymać równość

$$c\lim_{\mathbf{x}\to\mathbf{a}} \frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - D\mathbf{f}(\mathbf{a})(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} = 0.$$

A ona jest spełniona na mocy założonej różniczkowalności funkcji $\mathbf{f}$. To oznacza, że równość (5.40) jest spełniona, więc funkcja $\mathbf{h}$ jest różniczkowalna i ma odpowiednią granicę. ■

Lemat 4 (*pochodna sumy*). Niech $U \subset \mathbb{R}^n$, niech funkcje $\mathbf{f}, \mathbf{g} : U \to \mathbb{R}^m$ będą różniczkowalne w punkcie $\mathbf{a}$. Wtedy funkcja $\mathbf{h}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})$ jest różniczkowalna w $\mathbf{a}$ oraz

$$D\mathbf{h}(\mathbf{a}) = D\mathbf{f}(\mathbf{a}) + D\mathbf{g}(\mathbf{a}).$$

D o w ó d. Przekształcamy wyrażenie na granicę, która powinna być zerem:

$$\begin{aligned}
&\lim_{\mathbf{x}\to\mathbf{a}} \frac{\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{a}) - [D\mathbf{f}(\mathbf{a}) + D\mathbf{g}(\mathbf{a})](\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} \\
&\quad = \lim_{\mathbf{x}\to\mathbf{a}} \frac{\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - \mathbf{g}(\mathbf{a}) - D\mathbf{f}(\mathbf{a})(\mathbf{x}-\mathbf{a}) - D\mathbf{g}(\mathbf{a})(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} \\
&\quad = \lim_{\mathbf{x}\to\mathbf{a}} \frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - D\mathbf{f}(\mathbf{a})(\mathbf{x}-\mathbf{a}) + \mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{a}) - D\mathbf{g}(\mathbf{a})](\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|}
\end{aligned}$$

Korzystamy z Lematu 2:

$$\begin{aligned}
&\lim_{\mathbf{x}\to\mathbf{a}} \frac{\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{a}) - [D\mathbf{f}(\mathbf{a}) + D\mathbf{g}(\mathbf{a})](\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} \\
&\quad = \lim_{\mathbf{x}\to\mathbf{a}} \frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - D\mathbf{f}(\mathbf{a})(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} + \lim_{\mathbf{x}\to\mathbf{a}} \frac{\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{a}) - D\mathbf{g}(\mathbf{a})(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|}.
\end{aligned}$$

Obie granice z założenia są zerami, więc liczona granica jest zerem. Skoro tak, to funkcja $\mathbf{h}$ jest różniczkowalna i zachodzi potrzebna równość. ■

Twierdzenie 4 (*pochodna funkcji złożonej*). Niech $U \subset \mathbb{R}^n$ oraz $V \subset \mathbb{R}^m$ będą zbiorami otwartymi. Niech będą dane funkcje $\mathbf{g} : U \to V$, $\mathbf{f} : V \to \mathbb{R}^p$, dzięki czemu określone jest złożenie $\mathbf{f} \circ \mathbf{g}$. Niech $\mathbf{g}$ będzie różniczkowalna w punkcie $\mathbf{a}$ zaś $\mathbf{f}$ w punkcie $\mathbf{b} = \mathbf{g}(\mathbf{a})$. Wtedy funkcja $\mathbf{f} \circ \mathbf{g}$ jest różniczkowalna w punkcie $\mathbf{a}$ oraz zachodzi równość

$$D(\mathbf{f} \circ \mathbf{g})(\mathbf{a}) = D\mathbf{f}(\mathbf{b}) \cdot D\mathbf{g}(\mathbf{a}), \tag{5.41}$$

gdzie kropka $\cdot$ oznacza mnożenie macierzy.

D o w ó d można znaleźć w książce: J.E. Marsden, A.J. Tromba *Vector Calculus*, Freeman and Company, New York 1988, w § 2.7.

Jeśli wprowadzimy oznaczenie $\mathbf{y} = \mathbf{g}(\mathbf{x})$ i uznamy, że $\mathbf{f}$ jest funkcją zmiennej $\mathbf{y} \in V$, to według wzoru (5.41) pochodne cząstkowe funkcji złożonej wyrażają się jako sumy

$$\frac{\partial(\mathbf{f} \circ \mathbf{g})_l}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial f_l}{\partial y_k} \frac{\partial g_k}{\partial x_j}. \tag{5.42}$$

Wzór (5.28) jest szczególnym przypadkiem tego wzoru, trzeba tylko zauważyć, że tam $p = 1$ oraz $\mathbf{g} = \boldsymbol{\sigma}$, a zmienne oznaczone są innymi literami.

PRZYKŁAD 3. Niech $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$ zaś $f : \mathbb{R}^m \to \mathbb{R}$ niech będzie dana wzorem $f(\mathbf{y}) = \sin(\mathbf{y} \cdot \mathbf{y}) = \sin(y_1^2 + y_2^2 + \ldots + y_m^2)$. Wtedy $h : \mathbb{R}^n \to \mathbb{R}$ jako złożenie $h = f \circ \mathbf{g}$ jest dana wzorem $h(\mathbf{x}) = \sin[\mathbf{g}(\mathbf{x}) \cdot \mathbf{g}(\mathbf{x})]$, gdyż podstawiliśmy $\mathbf{y} = \mathbf{g}(\mathbf{x})$. Pochodna macierzowa, czyli gradient funkcji f wynosi

$$\nabla f(\mathbf{y}) = (2y_1 \cos(\mathbf{y} \cdot \mathbf{y}), 2y_2 \cos(\mathbf{y} \cdot \mathbf{y}), \ldots, 2y_m \cos(\mathbf{y} \cdot \mathbf{y}))$$

$$= 2\cos(\mathbf{y} \cdot \mathbf{y})\ (y_1, y_2, \ldots, y_m) = 2\cos(\mathbf{y} \cdot \mathbf{y})\,\mathbf{y}.$$

W takim razie na podstawie (5.42) pochodną macierzową funkcji h możemy zapisać

$$Dh(\mathbf{x}) = \nabla f(\mathbf{y}) \cdot D\mathbf{g}(\mathbf{x}) = 2\cos[\mathbf{g}(\mathbf{x}) \cdot \mathbf{g}(\mathbf{x})]\ \mathbf{g}(\mathbf{x}) \cdot D\mathbf{g}(\mathbf{x}).$$

Jako iloczyn wiersza $\mathbf{y} = \mathbf{g}(\mathbf{x})$ o m elementach z macierzą prostokątną m na n $D\mathbf{g}(\mathbf{x})$ jest to macierz-wiersz o n elementach. I to dobrze, bo wartości funkcji h są rzeczywiste, więc jej pochodna macierzowa jest gradientem ∇h.

Twierdzenie 5 (*wzór Taylora*). Niech $U \subset \mathbb{R}^n$ i funkcja $f : U \to \mathbb{R}$ ma ciągłe pochodne cząstkowe trzeciego rzędu. Wtedy zachodzi

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\, h_i + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)\, h_i h_j + R_2(\mathbf{h}, \mathbf{x}_0),$$

gdzie $R_2(\mathbf{h}, \mathbf{x}_0)/||\mathbf{h}||^2 \to 0$ przy $\mathbf{h} \to \mathbf{0}$.

[Brook Taylor (1685–1731) – matematyk angielski, zajmował się analizą matematyczną i geometrią.]

D o w ó d. W wyrażeniu $f(\mathbf{x}_0 + t\mathbf{h})$ widzimy funkcję złożoną $(f \circ \boldsymbol{\sigma})$, gdzie $\boldsymbol{\sigma}(t) = \mathbf{x}_0 + t\mathbf{h}$. W tym wypadku $D\boldsymbol{\sigma}(t) = \frac{d\boldsymbol{\sigma}}{dt} = \mathbf{h}$. Stosujemy wzór (5.42) na pochodną funkcji złożonej:

$$\frac{d}{dt} f(\mathbf{x}_0 + t\mathbf{h}) = [Df(\mathbf{x}_0 + t\mathbf{h})] \cdot \mathbf{h} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\mathbf{h})\, h_i.$$

Całkujemy tę równość stronami względem t:

$$f(\mathbf{x}_0 + t\mathbf{h}) + C = \int \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\mathbf{h})\, h_i\, dt = \sum_{i=1}^{n} \int \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\mathbf{h})\, h_i\, dt.$$

Całkujemy przez części podstawiając $v' = 1,\ u = \frac{\partial f}{\partial x_i} h_i$; stąd $u' = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0 + t\mathbf{h})\, h_i h_j$ i wybieramy $v = t - 1$:

$$f(\mathbf{x}_0 + t\mathbf{h}) + C = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\mathbf{h})\, h_i (t-1) - \sum_{i,j=1}^{n} \int (t-1) \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0 + t\mathbf{h})\, h_i h_j\, dt.$$

Teraz liczymy całkę oznaczoną w granicach od $t = 0$ do $t = 1$:

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = -\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\, h_i (-1) + \sum_{i,j=1}^{n} \int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0 + t\mathbf{h})\, h_i h_j\, dt.$$

W ten sposób udowodniliśmy tożsamość

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\, h_i + R_1(\mathbf{h}, \mathbf{x}_0), \tag{5.43}$$

gdzie

$$R_1(\mathbf{h}, \mathbf{x}_0) = \sum_{i,j=1}^{n} \int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0 + t\mathbf{h})\, h_i h_j\, dt. \tag{5.44}$$

Całkujemy wyrażenie (5.44) przez części podstawiając

$$u = \frac{\partial^2 f}{\partial x_i \partial x_j}\, h_i h_j \quad \text{oraz} \quad v' = 1 - t.$$

Wówczas

$$u' = \sum_{k=1}^{n} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}\, h_i h_j h_k \quad \text{oraz} \quad v = -\frac{(t-1)^2}{2}$$

i otrzymujemy całkę oznaczoną

$$R_1(\mathbf{h}, \mathbf{x}_0) =$$

$$= -\sum_{i,j=1}^{n} \frac{(t-1)^2}{2} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0 + t\mathbf{h})\, h_i h_j \Big|_0^1 + \sum_{i,j,k=1}^{n} \int_0^1 \frac{(t-1)^2}{2} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\mathbf{x}_0 + t\mathbf{h})\, h_i h_j h_k\, dt.$$

$$= \sum_{i,j=1}^{n} \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)\, h_i h_j + \sum_{i,j,k=1}^{n} \int_0^1 \frac{(t-1)^2}{2} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\mathbf{x}_0 + t\mathbf{h})\, h_i h_j h_k\, dt.$$

W takim razie wykazaliśmy, że

$$f(\mathbf{x}_0+\mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\, h_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)\, h_i h_j + R_2(\mathbf{h},\mathbf{x}_0), \tag{5.45}$$

gdzie

$$R_2(\mathbf{h},\mathbf{x}_0) = \sum_{i,j,k=1}^{n} \left[\int_0^1 \frac{(t-1)^2}{2} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\mathbf{x}_0+t\mathbf{h})\, dt\right] h_i h_j h_k\,. \tag{5.46}$$

Szacujemy moduł tej reszty korzystając z nierówności trójkąta

$$|R_2(\mathbf{h},\mathbf{x}_0)| = \left|\sum_{i,j,k=1}^{n} \left[\int_0^1 \frac{(t-1)^2}{2} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\mathbf{x}_0+t\mathbf{h})\, dt\right] h_i h_j h_k\right|$$

$$\leq \sum_{i,j,k=1}^{n} \left|\int_0^1 \frac{(t-1)^2}{2} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\mathbf{x}_0+t\mathbf{h})\, dt\right| |h_i|\,|h_j|\,|h_k|.$$

Każda współrzędna h_j wektora $\mathbf{h}$ spełnia nierówność $|h_j| \leq ||\mathbf{h}||$, więc

$$|R_2(\mathbf{h},\mathbf{x}_0)| \leq \sum_{i,j,k=1}^{n} \left|\int_0^1 \frac{(t-1)^2}{2} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\mathbf{x}_0+t\mathbf{h})\, dt\right| ||\mathbf{h}||^3.$$

wobec tego istnieje stała

$$M = \sum_{i,j,k=1}^{n} \left|\int_0^1 \frac{(t-1)^2}{2} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\mathbf{x}_0+t\mathbf{h})\, dt\right| \geq 0$$

taka, że

$$|R_2(\mathbf{h},\mathbf{x}_0)| \leq M\, ||\mathbf{h}||^3.$$

Stąd wynika

$$\frac{|R_2(\mathbf{h},\mathbf{x}_0)|}{||\mathbf{h}||^2} \leq M\, ||\mathbf{h}|| \to 0 \quad \text{gdy} \quad \mathbf{h} \to 0.$$

Wzór (5.45) jest tym wzorem, jaki mieliśmy udowodnić, a ostatnia zbieżność jest tą, jaką należało wykazać. ■

Równość (5.43) nazywa się **wzorem Taylora pierwszego rzędu**, a (5.45) **wzorem Taylora drugiego rzędu**. Przez analogię można napisać **wzór Taylora trzeciego rzędu**:

$$f(\mathbf{x}_0+\mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\, h_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)\, h_i h_j$$

$$+ \frac{1}{3!}\sum_{i,j,k=1}^{n} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\mathbf{x}_0)\, h_i h_j h_k + R_3(\mathbf{h},\mathbf{x}_0), \tag{5.47}$$

i dalszych rzędów.

PRZYKŁAD 4. Znajdziemy wzór Taylora drugiego rzędu dla $f(x,y) = \sin(x+2y)$ wokół punktu $\mathbf{x}_0 = (0,0)$. Mamy $f(0,0) = 0$ tudzież

$$\partial_x f(x,y) = \cos(x+2y), \qquad \partial_y f(x,y) = 2\cos(x+2y),$$

$$\partial_x f(0,0) = \cos(0+2\cdot 0) = 1, \quad \partial_y f(0,0) = 2\cos(0+2\cdot 0) = 2,$$

$$\partial_x\partial_x f(0,0) = -\sin(0+2\cdot 0) = 0, \quad \partial_x\partial_y f(0,0) = -2\sin(0+2\cdot 0) = 0,$$

$$\partial_y\partial_y f(0,0) = -4\sin(0+2\cdot 0) = 0.$$

Zatem

$$f(\mathbf{0}+\mathbf{h}) = h_1 + 2h_2 + R_2(\mathbf{h},\mathbf{0}).$$

To twierdzenie można zastosować do przypadku $n = 1$, czyli do funkcji jednej zmiennej. Zmiana polega na tym, że nie ma tłustego druku, pochodne czątkowe zamieniają się na zwyczajne i nie ma sumowania względem róznych zmiennych, czyli względem i lub i,j. W takim razie wzór Taylora trzeciego rzędu dla funkcji jednej zmiennej ma postać

$$f(x_0+h) = f(_0) + \frac{df}{dx}(x_0)\,h + \frac{1}{2}\frac{d^2f}{dx^2}(x_0)\,h^2 + \frac{1}{6}\frac{d^3f}{dx^3}(x_0)\,h^3 + R_3(h,x_0),$$

gdzie $\lim\limits_{h\to 0} R_3(h,x_0)/h^3 = 0$. Wzór ten zgadza się z rozważanym w § 2.3 wzorem Taylora przy nieco innych oznaczeniach: $x_0 = a,\ x_0 + h = b,\ h = b - a$.

Niech $U \subset \mathbb{R}^n$ i $f : U \to \mathbb{R}$ jest daną funkcją. Punkt $\mathbf{a} \in U$ nazywa się **minimum lokalnym**, jeśli istnieje niby-kula otwarta $K(\mathbf{a},r)$ taka, że dla wszystkich $\mathbf{x} \in K(\mathbf{a},r)$ zachodzi $f(\mathbf{x}) \geq f(\mathbf{a})$. Jeśli ponadto zachodzi $f(\mathbf{x}) > f(\mathbf{a})$ dla $\mathbf{x} \neq \mathbf{a}$, to punkt $\mathbf{a}$ nazywa się **ścisłym minimum lokalnym**. Podobnie $\mathbf{b} \in U$ nazywa się **maksimum lokalnym**, jeśli istnieje niby-kula otwarta $K(\mathbf{b},r)$ taka, że dla wszystkich $\mathbf{x} \in K(\mathbf{b},r)$ zachodzi $f(\mathbf{x}) \leq f(\mathbf{b})$. Jeśli ponadto $f(\mathbf{x}) < f(\mathbf{b})$ dla $\mathbf{x} \neq \mathbf{b}$, to punkt $\mathbf{b}$ nazywa się **ścisłym maksimum lokalnym**. Mówimy, że punkt $\mathbf{a}$ jest **ekstremum lokalnym**, jeśli jest minimum lokalnym albo maksimum lokalnym. Jeśli w $\mathbf{x}_0$ zeruje się pochodna $Df(\mathbf{x}_0) = 0$, to punkt $\mathbf{x}_0$ nazywa się **punktem krytycznym**. Punkt krytyczny, który nie jest lokalnym ekstremum, nazywa się **punktem siodłowym**.

Twierdzenie 6. Jeśli $U \subset \mathbb{R}^n$ jest otwarty, funkcja $f : U \to \mathbb{R}$ jest różniczkowalna oraz $\mathbf{a} \in U$ jest lokalnym ekstremum, to $Df(\mathbf{a}) = 0$. Krócej: każde lokalne ekstremum jest punktem krytycznym.

D o w ó d . Załóżmy, że funkcja f osiąga w $\mathbf{a}$ ekstremum lokalne. Wtedy dla każdego $\mathbf{h} \in \mathbb{R}^n$ funkcja $g(t) = f(\mathbf{a}+t\mathbf{h})$ ma ekstremum lokalne w $t = 0$. Na mocy Tw. 1 z § 2.2 zachodzi $g'(0) = 0$. Z drugiej strony g jest funkcją złożoną $g = f \circ \boldsymbol{\sigma}$ dla $\boldsymbol{\sigma}(t) = \mathbf{a} + \mathbf{h}t$ oraz $D\boldsymbol{\sigma}(0) = d\boldsymbol{\sigma}/dt = \mathbf{h}$. Na mocy wzoru (5.29) mamy

$$g'(0) = [\nabla f(\mathbf{a})] \cdot \mathbf{h}.$$

W ten sposób $[\nabla f(\mathbf{a})] \cdot \mathbf{h} = 0$ dla dowolnych $\mathbf{h}$. Jedynym wektorem prostopadłym do wszystkich wektorów jest wektor zerowy, więc $\nabla f(\mathbf{a}) = 0$. ■

PRZYKŁAD 5. Znajdziemy punkty krytyczne funkcji $f(x,y) = x^2y + xy^2$. Różniczkowanie daje

$$\frac{\partial f}{\partial x} = 2xy + y^2, \qquad \frac{\partial f}{\partial y} = x^2 + 2xy.$$

Przyrównujemy te pochodne do zera

$$2xy + y^2 = 0, \qquad 2xy + x^2 = 0, \tag{5.48}$$

$$y^2 = -2xy, \qquad x^2 = -2xy.$$

Stąd $x^2 = y^2$, $x = \pm y$. Podstawienie $x = y$ do pierwszego z równań (5.48) daje

$$2y^2 + y^2 = 3y^2 = 0,$$

więc $y = 0$, a stąd $x = 0$. Kiedy $x = -y$, to samo równanie daje

$$-2y^2 + y^2 = -y^2 = 0,$$

więc $y = 0$, a stąd $x = 0$. W ten sposób jedynym punktem krytycznym jest początek układu $(0,0)$. Czy to jest ekstremum? Dla prostej $x = y$ dostajemy $f(x,x) = 2x^3$, a ta funkcja jednej zmiennej jest zarówno dodatnia jak i ujemna w pobliżu zera. Dlatego stwierdzamy, że punkt krytyczny $(0,0)$ nie jest ekstremum lokalnym – trzeba go nazwać punktem siodłowym. Ten przykład pokazuje, że w Twierdzeniu 6 wynikanie w drugą stronę nie jest prawdziwe.

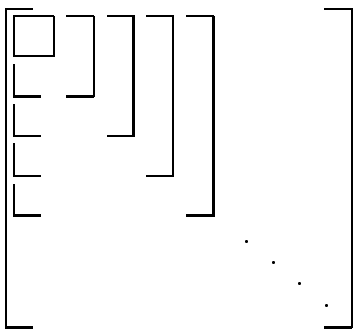
Niech $U \subset \mathbb{R}^n$, a funkcja $f : U \to \mathbb{R}$ ma wszystkie pochodne cząstkowe drugiego rzędu $f''_{x_i x_j}(\mathbf{a})$ w punkcie $\mathbf{a}$. **Macierzą Hessego** funkcji f w punkcie $\mathbf{a}$ jest macierz kwadratowa określona wzorem

$$Hf(\mathbf{a}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{a}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{a}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(\mathbf{a}) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{a}) \\ \dots\dots & \dots\dots & \dots & \dots\dots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{a}) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{a}) \end{pmatrix}, \tag{5.49}$$

[Ludwig Otto Hesse (1811-1874) – matematyk niemiecki, wprowadził swoją macierz drugich pochodnych w 1844 roku.] Macierz Hessego jest przeważnie stosowana w punktach krytycznych $\mathbf{x}_0 \in U$, dlatego pierwsze pochodne cząstkowe są tam zerami i wzór Taylora można zapisać

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \frac{1}{2}\sum_{i,j=1}^{n} Hf(\mathbf{x}_0)_{ij} h_i h_j + R_2(\mathbf{h}, \mathbf{x}_0).$$

W ten sposób w punkcie krytycznym macierz Hessego występuje w pierwszym nie stałym wyrazie w rozwinięciu Taylora danej funkcji. Wyznacznik macierzy Hessego nazywa się **wyznacznikiem Hessego** albo **hesjanem** funkcji f. Do następnego twierdzenia potrzebny jest hesjan i jego podwyznaczniki coraz mniejszego rzędu powstające przez odrzucenie ostatniego wiersza i ostatniej kolumny, zob. Rys. 7.21.



Rysunek 7.21

Twierdzenie 7. Niech $U \subset \mathbb{R}^n$, niech funkcja $f : U \to \mathbb{R}$ będzie klasy C^3 i niech $\mathbf{x}_0 \in U$ będzie punktem krytycznym tej funkcji. Jeśli element $(Hf)_{11}(\mathbf{x}_0)$ macierzy Hesego jest dodatni i wszystkie podwyznaczniki coraz wyższego rzędu leżące na głównej przekątnej aż do hesjanu są dodatnie, to $\mathbf{x}_0$ jest ścisłym minimum lokalnym funkcji f. Jeśli element $(Hf)_{11}$ macierzy Hessego jest ujemny, a podwyznaczniki coraz wyższego rzędu leżące na głównej przekątnej są na przemian dodatnie i ujemne, to $\mathbf{x}_0$ jest ścisłym maksimum lokalnym funkcji f.

D o w ó d (dla szczególnego przypadku $n = 2$) można znaleźć w książce: J.E. Marsden, A.J. Tromba *Vector Calculus*, Freeman and Company, New York 1988, w § 4.2.

PRZYKŁAD 6. Rozważmy funkcję $f : \mathbb{R}^2 \to \mathbb{R}$ daną wzorem $f(x, y) = x^2 - 2xy + 2y^2$. Jej pochodna macierzowa wynosi $\nabla f(x, y) = (2x - 2y, -2x + 4y)$. Szukamy punktów krytycznych przyrównując pochodną do zera:

$$2x - 2y = 0, \quad -2x + 4y = 0.$$

Z pierwszego równania mamy $y = x$; wstawiamy to do drugiego: $-2x + 4x = 0$, co daje $x = 0$. W takim razie też $y = 0$ i dostajemy tylko jeden punkt krytyczny $\mathbf{r}_0 = (0, 0)$.

Macierz Hessego ma postać

$$Hf(\mathbf{r}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x \partial x}(\mathbf{r}) & \frac{\partial^2 f}{\partial x \partial y}(\mathbf{r}) \\ \frac{\partial^2 f}{\partial y \partial x}(\mathbf{r}) & \frac{\partial^2 f}{\partial y \partial y}(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}.$$

Jej wyznacznik, czyli hesjan wynosi $\det H = 8 - 4 = 4$. Jako funkcja jest on stały, więc tyle samo wynosi w punkcie krytycznym. Jego podwyznacznik z głównej przekątnej to pierwszy element tej przekątnej wynoszący 2. Skoro oba te wyznaczniki są dodatnie, to na mocy Tw. 7 funkcja f ma lokalne minimum w początku układu współrzędnych. Widać to także z jej zapisania w innej postaci $f(x, y) = (x - y)^2 + y^2$, skąd $f(x, y) > 0$ dla $(x, y) \neq (0, 0)$

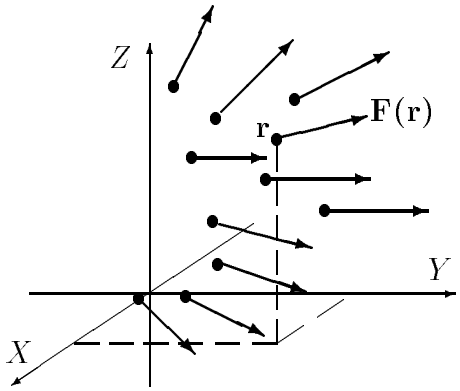
Weźmy funkcję $g : \mathbb{R}^2 \to \mathbb{R}$ przeciwną do poprzedniej: $g(x, y) = -x^2 + 2xy - 2y^2$. Możemy z góry powiedzieć, że jej jedyny punkt krytyczny jest taki sam, jak funkcji f, a z powodu jej przeciwnego znaku – w tym punkcie jest lokalne maksimum. Macierz Hessego tej funkcji

$$Hg(\mathbf{r}) = \begin{pmatrix} \frac{\partial^2 g}{\partial x \partial x}(\mathbf{r}) & \frac{\partial^2 g}{\partial x \partial y}(\mathbf{r}) \\ \frac{\partial^2 g}{\partial y \partial x}(\mathbf{r}) & \frac{\partial^2 g}{\partial y \partial y}(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 2 & -4 \end{pmatrix}$$

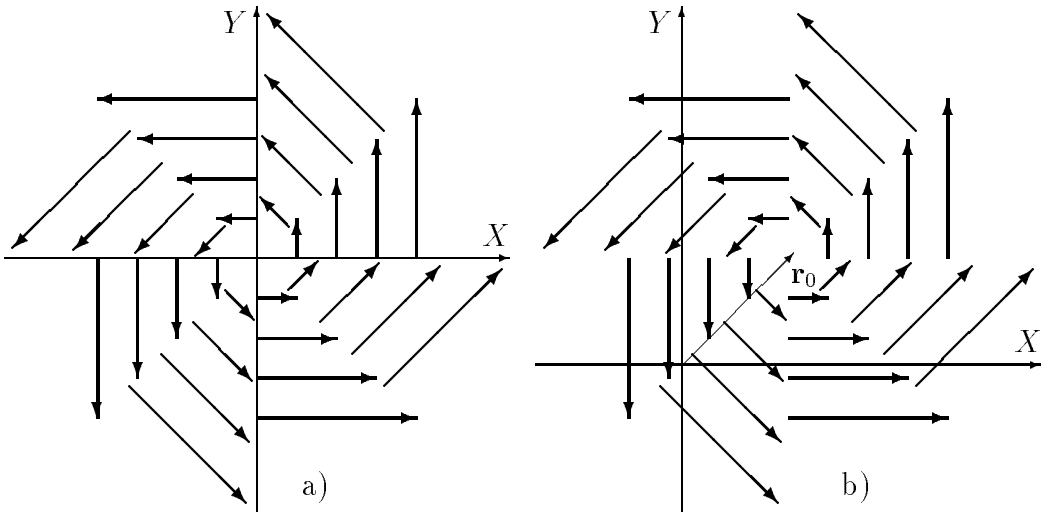
ma wyznacznik tego samego znaku: $\det Hg = 4$, ale pierwszy element głównej przekątnej jest ujemny. Według Tw. 7 funkcja g ma lokalne maksimum – zgodnie z pierwszym spostrzeżeniem.

5.4 Pola wektorowe

Polem wektorowym nazywamy funkcję $\mathbf{F}$, której dziedzina i przeciwdziedzina mają ten sam wymiar. A dokładniej dla $U \in \mathbb{R}^n$ mamy $\mathbf{F} : U \to \mathbb{R}^n$. Mówimy, że pole wektorowe każdemu punktowi $\mathbf{x} \in U$ przypisuje wektor $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^n$. Taką funkcję możemy zobrazować przypisując możliwie wielu punktom strzałki, zob. Rys. 5.21. Tym razem strzałka przedstawiająca wektor $\mathbf{F}(\mathbf{x})$ zaczepiana jest w punkcie $\mathbf{x}$ zamiast w początku układu, co jest normalnym zwyczajem dla wektorów. W tym kontekście funkcja $f : U \to \mathbb{R}$, która każdemu punktowi przypisuje liczbę rzeczywistą, nazywa się **polem skalarnym**. Fizycznym przykładem pola wektorowego jest natężenie pola elektrycznego, a pola skalarnego potencjał elektryczny.



Rysunek 5.21



Rysunek 5.22

PRZYKŁAD 1. Niech $\mathbf{V}(\mathbf{r})$ będzie polem prędkości punktów obracającej się bryły sztywnej, według wzoru

$$\mathbf{V}(x, y, z) = a(-y\mathbf{e}_1 + x\mathbf{e}_2) \tag{5.50}$$

z pewną stałą a o fizycznym wymiarze odwrotności czasu. Możemy sprawdzić, że w każdym punkcie $\mathbf{r} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$ jest ono prostopadłe do wektora wodzącego: $\mathbf{V} \cdot \mathbf{r} = 0$. Jego wartość rośnie liniowo wraz z wartością wektora $\mathbf{r}$. Jego ilustracją jest Rys. 5.22a wykonany dla płaszczyzny (X, Y). Pole $\mathbf{V}(\mathbf{r})$ nie zależy od zmiennej z, więc ma symetrię względem przeusunięć wzdłuż osi Z. To oznacza, że w każdej płaszczyźnie równległej do płaszczyzny (X, Y) wygląda tak samo, jak na Rys. 5.22a. Jest ono polem zakręconym o osi obracania pokrywającej się z osią Z. Dziedziną U, w której obowiązuje wzór (5.50), jest obszar zajęty przez bryłę sztywną.

Pole $\mathbf{V}'(\mathbf{r}) = \mathbf{V}(\mathbf{r} - \mathbf{r}_0)$, czyli dane wzorem

$$\mathbf{V}'(x, y, z) = a\,[-(y - y_0)\mathbf{e}_1 + (x - x_0)\mathbf{e}_2]$$

jest podobnym polem wirowym, ale „zakręconym" wokół innej osi, mianowicie równoległej do osi Z i przechodzącej przez koniec wektora $\mathbf{r}_0$. Jego wygląd na płaszczyźnie (X, Y) ukazany jest na Rys. 5.22b. Pole $\mathbf{V}'$ można zapisać inaczej

$$\mathbf{V}'(x, y, z) = a(-y\mathbf{e}_1 + x\mathbf{e}_2) - a(-y_0\mathbf{e}_1 + x_0\mathbf{e}_2) = \mathbf{V}(x, y, z) - \mathbf{V}_0, \qquad (5.51)$$

skąd widzimy, że różni się od pola $\mathbf{V}$ tylko o pewien wektor stały.

Rozważmy teraz pole wektorowe $\mathbf{F}$ klasy C^1 na $\mathbb{R}^3$:

$$\mathbf{F}(x, y, z) = F_1(x, y, z)\,\mathbf{e}_1 + F_2(x, y, z)\,\mathbf{e}_2 + F_3(x, y, z)\,\mathbf{e}_3.$$

Nowe pole wektorowe

$$\operatorname{rot}\mathbf{F} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\mathbf{e}_1 + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right)\mathbf{e}_2 + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\mathbf{e}_3 \qquad (5.52)$$

nazywamy **rotacją pola F**. Ten wzór jest łatwiejszy do zapamiętania, jeśli go przepisać z oznaczeniem operatorowym. Wprowadźmy **operator wektorowy nabla**:

$$\boldsymbol{\nabla} = \mathbf{e}_1\frac{\partial}{\partial x} + \mathbf{e}_2\frac{\partial}{\partial y} + \mathbf{e}_3\frac{\partial}{\partial z} = \mathbf{e}_1\frac{\partial}{\partial x_1} + \mathbf{e}_2\frac{\partial}{\partial x_2} + \mathbf{e}_3\frac{\partial}{\partial x_3}.$$

Słowo 'operator' oznacza, że on działa na funkcje o wartościach skalarnych lub wektorowych, aby je zróżniczkować. Jeśli spojrzymy na niego jak na wektor o współrzędnych $(\partial_x, \partial_y, \partial_z) = (\partial_1, \partial_2, \partial_3)$, to możemy rotację zapisać jako iloczyn wektorowy nabli i wektora $\mathbf{F}$ z użyciem symbolu wyznacznika

$$\boldsymbol{\nabla} \times \mathbf{F} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ F_1 & F_2 & F_3 \end{vmatrix} = (\partial_2 F_3 - \partial_3 F_2)\mathbf{e}_1 + (\partial_3 F_1 - \partial_1 F_3)\mathbf{e}_2 + (\partial_1 F_2 - \partial_2 F_1)\mathbf{e}_3. \qquad (5.53)$$

W ten sposób

$$\operatorname{rot}\mathbf{F} = \boldsymbol{\nabla} \times \mathbf{F}.$$

Zauważmy, że pole rot$\mathbf{F}$ jest klasy C^{k-1}, jeśli pole $\mathbf{F}$ jest klasy C^k.

Właściwie nabla występująca w zapisie gradientu też jest operatorem wektorowym, bo działając na pole skalarne f daje pole wektorowe $\boldsymbol{\nabla} f$ – odpowiada to mnożeniu wektora przez skalar, co daje wektor.

Twierdzenie 1. Niech $U \subset \mathbb{R}^3$. Dla dowolnej funkcji $f : U \to \mathbb{R}$ klasy C^2 zachodzi

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} f) = \mathbf{0},$$

tzn. rotacja gradientu jest wektorem zerowym.

D o w ó d. Do wzoru (5.52) należy podstawić $F_i = \partial f / \partial x_i$:

$$\mathrm{rot}(\mathrm{grad}) f = \left(\frac{\partial^2 f}{\partial x_2 \partial x_3} - \frac{\partial^2 f}{\partial x_3 \partial x_2} \right) \mathbf{e}_1 + \left(\frac{\partial^2 f}{\partial x_3 \partial x_1} - \frac{\partial^2 f}{\partial x_1 \partial x_3} \right) \mathbf{e}_2 + \left(\frac{\partial^2 f}{\partial x_1 \partial x_2} - \frac{\partial^2 f}{\partial x_2 \partial x_1} \right) \mathbf{e}_3.$$

Skoro drugie pochodne są ciągłe, to na mocy Tw. 1 z § 8.2 pochodne mieszane są przemienne, więc wyrażenia w nawiasach są zerami. Otrzymujemy wektor o zerowych współrzędnych, więc wektor zerowy. ■

PRZYKŁAD 1a. Wyliczmy rotację pola $\mathbf{V}$ z Przykładu 1.

$$\boldsymbol{\nabla} \times \mathbf{V} = (-\partial_3 V_2)\mathbf{e}_1 + (\partial_3 V_1)\mathbf{e}_2 + (\partial_1 V_2 - \partial_2 V_1)\mathbf{e}_3 = (\partial_1 V_2 - \partial_2 V_1)\mathbf{e}_3$$

$$= a(1+1)\mathbf{e}_3 = 2a\,\mathbf{e}_3.$$

Zauważamy, że rotacja dała pole jednorodne, tzn. niezależne od położenia, Jego kierunek jest równoległy do osi Z, czyli do osi obrotu płyty. To jest cecha charakterystyczna wprowadzonego odwzorowania – jest ono niezerowe wtedy, gdy pole wektorowe ma coś wspólnego z obracaniem, a jego kierunek jest równoległy do osi obrotu. Zresztą stąd jest nazwa tego operatora: rotacja.

Można zapytać, dlaczego rotacja tego pola nie wyróżnia jakoś osi Z, skoro to jest oś obrotu. Odpowiedzią na to jest rotacja pola $\mathbf{V}'$, która jest taka sama, jak rotacja pola $\mathbf{V}$, gdyż według wzoru (5.51) oba te pola różnią się o wektor stały, więc wszelkie ich pochodne muszą być jednakowe. Stąd

$$\boldsymbol{\nabla} \times \mathbf{V}' = 2a\,\mathbf{e}_3.$$

A osią obrotu dla pola $\mathbf{V}'$ jest prosta przechodząca przez koniec wektora $\mathbf{r}_0$. stały

PRZYKŁAD 2. Rozważmy pole wektorowe $\mathbf{v}$ w walcu o osi pokrywającej się z osią Z, ale z osią Z usuniętą z dziedziny:

$$\mathbf{v}(x,y,z) = A\,\frac{-y\mathbf{e}_1 + x\mathbf{e}_2}{\rho^2} = A\left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2} \right), \tag{5.54}$$

gdzie $\rho^2 = x^2 + y^2$. Zauważmy taki związek $\mathbf{v} = \frac{A}{a\rho^2}\mathbf{V}$ z polem z Przykładu 1. Jako równoległe do pola $\mathbf{V}$ to pole jest znowu styczne do okręgów o środkach na osi Z, ale jego wartość maleje wraz z oddalaniem się od osi, więc rysunek 5.22a już nie jest dla niego odpowiedni.

Rozważmy ciecz w jakimś przepływie ustalonym. Jeśli każdemu punktowi przypiszemy prędkość cieczy w tym punkcie, to otrzymamy pole prędkości $\mathbf{v}(\mathbf{r})$ cieczy. W takiej sytuacji warunek $\mathrm{rot}\,\mathbf{v} = 0$ oznacza fizycznie, że ciecz

jest wolna od wirów, albo jest **bezwirowa**. Uzasadnienie tego poglądu i użycie słowa „bezwirowa" zależy od twierdzenia Stokesa, które poznamy w następnym rozdziale. Pole (5.54) przybliża pole prędkości wody w ruchu okrężnym, jaki występuje w wiadrze wody z dziurą wywierconą w środku jego dna, czyli wtedy, gdy pojawia się wir. Wówczas prędkość jest coraz większa w miarę zbliżania się do osi wiru. Dlaczego jest napisane „przybliża"? Otóż dlatego, że w takim wiadrze prędkość nie jest idealnie styczna do okręgów o środkach na osi wiadra, lecz musi mieć niewielką składową w stronę osi.

Wyliczmy rotację tego pola:

$$\boldsymbol{\nabla} \times \mathbf{v} = (\partial_1 v_2 - \partial_2 v_1)\,\mathbf{e}_3 = \left(\frac{x^2+y^2-x\cdot 2x}{(x^2+y^2)^2} - \frac{-(x^2+y^2)+y\cdot 2y}{(x^2+y^2)^2}\right)\mathbf{e}_3$$

$$= \frac{-x^2+y^2+x^2-y^2}{(x^2+y^2)^2}\,\mathbf{e}_3 = \mathbf{0}.$$

Widzimy, że rotacja jest zerem w całej dziedzinie, czyli poza osią Z. Ten wynik oznacza, że nie ma innych wirów poza osią jednego wiru.

Innym operatorem działającym na pola wektorowe w $\mathbb{R}^3$ jest **dywergencja** określona jako

$$\operatorname{div}\mathbf{F} = \boldsymbol{\nabla}\cdot\mathbf{F} = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}. \tag{5.55}$$

Jest to iloczyn skalarny nabli i wektora $\mathbf{F}$. Zauważmy, że dywergencja jest polem skalarnym, więc w pewnym sensie jest czymś odwrotnym do gradientu. Następny przykład pomoże podać znaczenie fizyczne tego operatora.

PRZYKŁAD 3. Wyobraźmy sobie, że $\mathbf{F}$ jest polem prędkości gazu, danym przez

$$\mathbf{F}(\mathbf{r}) = a(x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3) = a\,\mathbf{r}.$$

Ten wzór oznacza, że gaz rozbiega się tym szybciej, im dalej jest od początku układu. Tak zachowują się galaktyki w naszym wszechświecie, a współczynnik a nazywa się **stałą Hubble'a**. Dla takiego pola $\mathbf{F}$:

$$\operatorname{div}\mathbf{F} = a(1+1+1) = 3a.$$

To oznacza, że gaz powiększa trzykrotnie swoją objętość na jednostkę czasu. Słowo dywergencja wzięło się z łacińskiego *divergare* = rozbiegać się. Zresztą jeszcze przed ostatnią wojną polscy matematycy zamiast „dywergencja" mówili „rozbieżność".

Pole

$$\mathbf{G}(\mathbf{r}) = \mathbf{F}(\mathbf{r}-\mathbf{r}_0) = a\,(\mathbf{r}-\mathbf{r}_0)$$

ma środek rozbiegania się w punkcie $\mathbf{r}_0$ zamiast w początku układu. Jednakże można je zapisać inaczej

$$\mathbf{G}(\mathbf{r}) = a\mathbf{r} - a\mathbf{r}_0 = \mathbf{F}(\mathbf{r}) - \mathbf{F}_0,$$

czyli różni się od pola $\mathbf{F}$ tylko o wektor stały, więc ma te same pochodne, czyli tę samą dywergencję, co pole $\mathbf{F}$:

$$\operatorname{div}\mathbf{G} = 3a.$$

Nic dziwnego więc, że pole $\mathbf{F}$ ma dywergencję jednakową w każdym punkcie.

Twierdzenie 2. Niech $U \subset \mathbb{R}^3$. Dla dowolnej funkcji $\mathbf{G} : U \to \mathbb{R}^3$ klasy C^2 zachodzi

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{G}) = \mathbf{0},$$

tzn. dywergencja rotacji jest zerem.

D o w ó d . Do wzoru (5.55) należy podstawić $F_1 = \partial_2 G_3 - \partial_3 G_2$ i podobnie dla F_2, F_3 z cykliczną zamianą wskaźników:

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{G}) = \frac{\partial}{\partial x_1}\left(\frac{\partial G_3}{\partial x_2} - \frac{\partial G_2}{\partial x_3}\right) + \frac{\partial}{\partial x_2}\left(\frac{\partial G_1}{\partial x_3} - \frac{\partial G_3}{\partial x_1}\right) + \frac{\partial}{\partial x_3}\left(\frac{\partial G_2}{\partial x_1} - \frac{\partial G_1}{\partial x_2}\right)$$

$$= \frac{\partial^2 G_3}{\partial x_1 \partial x_2} - \frac{\partial^2 G_2}{\partial x_1 \partial x_3} + \frac{\partial^2 G_1}{\partial x_2 \partial x_3} - \frac{\partial^2 G_3}{\partial x_2 \partial x_1} + \frac{\partial^2 G_2}{\partial x_3 \partial x_1} - \frac{\partial^2 G_1}{\partial x_3 \partial x_2} = 0,$$

gdyż pochodne mieszane są przemienne i upraszczają się parami. ■

PRZYKŁAD 3a. Na podstawie Twierdzenia 2 wnioskujemy, że pole $\mathbf{F}$ z Przykładu 3 nie może być rotacją innego pola wektorowego, bo wtedy miałoby zerową dywergencję.

Wprowadza się jeszcze **operator Laplace'a** zwany też **laplasjanem**:

$$\Delta = \frac{\partial^2}{\partial {x_1}^2} + \frac{\partial^2}{\partial {x_2}^2} + \frac{\partial^2}{\partial {x_3}^2},$$

który może działać zarówno na pola skalarne, jak i wektorowe, bo sam jest operatorem skalarnym. Natomiast w działaniu na pola skalarne można go zapisać jeszcze inaczej

$$\Delta f = \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} f) = \frac{\partial^2 f}{\partial {x_1}^2} + \frac{\partial^2 f}{\partial {x_2}^2} + \frac{\partial^2 f}{\partial {x_3}^2}.$$

[Pierre Simon Laplace (1749-1827) – francuski matematyk, fizyk i astronom. W wieku 26 lat miał stanowisko profesora w szkole artyleryjskiej w Paryżu. W okresie Rewolucji Francuskiej był ministrem spraw wewnętrznych. W astronomii zasłużył się dla mechaniki nieba, w matematyce dla teorii prawdopodobieństwa i równań różniczkowych cząstkowych, w fizyce dla teorii powszechnego ciążenia.]

Niech f będzie polem skalarnym, a $\mathbf{F}, \mathbf{G}$ – polami wektorowymi na $U \subset \mathbb{R}^3$. Poznane przez nas operatory różniczkowe zbudowane z nabli i laplasjanu spełniają następujące tożsamości

$$\operatorname{div}(f\mathbf{F}) = f \operatorname{div}\mathbf{F} + \mathbf{F} \cdot \operatorname{grad} f,$$

$$\operatorname{div}(\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot \operatorname{rot}\mathbf{F} - \mathbf{F} \cdot \operatorname{rot}\mathbf{G}, \tag{5.56}$$

$$\begin{aligned}
&\operatorname{rot}(f\mathbf{F}) = f \operatorname{rot}\mathbf{F} + (\operatorname{grad} f)\times\mathbf{F},\\
&\operatorname{rot}(\mathbf{F}\times\mathbf{G}) = \mathbf{F}\operatorname{div}\mathbf{G} - \mathbf{G}\operatorname{div}\mathbf{F} + (\mathbf{G}\cdot\boldsymbol{\nabla})\mathbf{F} - (\mathbf{F}\cdot\boldsymbol{\nabla})\mathbf{G}. \qquad (5.57)\\
&\operatorname{rot}(\operatorname{rot}\mathbf{F}) = \operatorname{grad}(\operatorname{div}\mathbf{F}) - \Delta\mathbf{F},
\end{aligned}$$

Udowodnienie większości z nich przenosimy na ćwiczenia, a teraz udowodnimy tożsamość (5.56). Lewa strona wynosi

$$\begin{aligned}
L &= \partial_1(\mathbf{F}\times\mathbf{G})_1 + \partial_2(\mathbf{F}\times\mathbf{G})_2 + \partial_2(\mathbf{F}\times\mathbf{G})_2\\
&= \partial_1(F_2G_3 - F_3G_2) + \partial_2(F_3G_1 - F_1G_3) + \partial_3(F_1G_2 - F_2G_1).
\end{aligned}$$

Prawa strona wynosi

$$\begin{aligned}
P &= G_1(\boldsymbol{\nabla}\times\mathbf{F})_1 + G_2(\boldsymbol{\nabla}\times\mathbf{F})_2 + G_3(\boldsymbol{\nabla}\times\mathbf{F})_3 - F_1(\boldsymbol{\nabla}\times\mathbf{G})_1 - F_2(\boldsymbol{\nabla}\times\mathbf{G})_2 - F_3(\boldsymbol{\nabla}\times\mathbf{G})_3\\
&= G_1(\partial_2F_3 - \partial_3F_2) + G_2(\partial_3F_1 - \partial_1F_3) + G_3(\partial_1F_2 - \partial_2F_1)\\
&\qquad -F_1(\partial_2G_3 - \partial_3G_2) - F_2(\partial_3G_1 - \partial_1G_3) - F_3(\partial_1G_2 - \partial_2G_1)\\
&= G_1\partial_2F_3 - G_1\partial_3F_2 + G_2\partial_3F_1 - G_2\partial_1F_3 + G_3\partial_1F_2 - G_3\partial_2F_1\\
&\qquad -F_1\partial_2G_3 + F_1\partial_3G_2 - F_2\partial_3G_1 + F_2\partial_1G_3) - F_3\partial_1G_2 + F_3\partial_2G_1.
\end{aligned}$$

Łączymy wyrazy z pochodnymi względem tych samych zmiennych:

$$\begin{aligned}
P &= (\partial_1F_2)G_3 + F_2(\partial_1G_3) - (\partial_1F_3)G_2 - F_3(\partial_1G_2)\\
&\quad +(\partial_2F_3)G_1 + F_3(\partial_2G_1) - (\partial_2F_1)G_3 - F_1(\partial_2G_3)\\
&\quad +(\partial_3F_1)G_2 + F_1(\partial_3G_2) - (\partial_3F_2)G_1 - F_2(\partial_3G_1).
\end{aligned}$$

Korzystamy z wzoru na pochodną iloczynu funkcyj

$$\begin{aligned}
P &= \partial_1(F_2G_3) - \partial_1(F_3G_2) + \partial_2(F_3G_1) - \partial_2(F_1G_3) + \partial_3(F_1G_2) - \partial_3(F_2G_1)\\
&= \partial_1(F_2G_3 - F_3G_2) + \partial_2(F_3G_1 - F_1G_3) + \partial_3(F_1G_2 - F_2G_1).
\end{aligned}$$

Zauważamy, że $L = P$.

Udowodnimy jeszcze tożsamość (5.57). Obie strony są wektorami. Wyliczmy piewrwszą współrzędną wektora po lewej stronie:

$$L_1 = \partial_2(\mathbf{F}\times\mathbf{G})_3 - \partial_3(\mathbf{F}\times\mathbf{G})_2 = \partial_2(F_1G_2 - F_2G_1) - \partial(F_3G_1 - F_1G_3).$$

Wykorzystujemy wzór Lebniza na pochodną iloczynu funkcyj

$$\begin{aligned}
L_1 &= G_2\partial_2F_1 + F_1\partial_2G_2 - G_1\partial_2F_2 - F_2\partial_1G_1 - G_1\partial_3F_3 - F_3\partial_3G_1 + G_3\partial_3F_1 + F_1\partial_3G_3\\
&= F_1\left(\partial_2G_2 + \partial_3G_3\right) - G_1\left(\partial_2F_2 + \partial_3F_3\right) + (G_2\partial_2 + G_3\partial_3)F_1 - (F_2\partial_2 + F_3\partial_3)G_1.
\end{aligned}$$

Liczymy pierwszą współrzędną prawej strony:

$$\begin{aligned}
P_1 &= F_1\operatorname{div}\mathbf{G} - G_1\operatorname{div}\mathbf{F} + (\mathbf{G}\cdot\boldsymbol{\nabla})F_1 - (\mathbf{F}\cdot\boldsymbol{\nabla})G_1 = F_1(\partial_1G_1 + \partial_2G_2 + \partial_3G_3)\\
&\quad -G_1(\partial_1F_1 + \partial_2F_2 + \partial_3F_3) + (G_1\partial_1 + G_2\partial_2 + G_3\partial_3)F_1 - (F_1\partial_1 + F_2\partial_2 + F_3\partial_3)G_1.
\end{aligned}$$

Wyrazy $F_1\partial_1G_1$ oraz $G_1\partial_1F_1$ występują po dwa razy z przeciwnymi znakami, więc się upraszczają i zostaje

$$P_1 = F_1\left(\partial_2G_2 + \partial_3G_3\right) - G_1\left(\partial_2F_2 + \partial_3F_3\right) + (G_2\partial_2 + G_3\partial_3)F_1 - (F_2\partial_2 + F_3\partial_3)G_1.$$

Zauważamy, że $L_1 = P_1$. Podobnie sprawdza się tożsamość dla pozostałych współrzędnych.