8. Wykład. Pochodne wyższych rzędów i twierdzenie Schwarza o symetrii drugiej różniczki

8.1. **Pochodne cząstkowe wyższego rzędu.**

Definicja 17. Niech $f\colon G \subset \mathbb{R}^n \to \mathbb{R}^m$. Załóżmy, że zbiór tych $x \in G$ dla których istnieje pochodna cząstkowa $\frac{\partial f}{\partial x_j}(x)$ (dla danego $j \in \{1, ..., n\}$) jest niepusty. Wówczas jeśli istnieje pochodna cząstkowa odwzorowania $x \longmapsto \frac{\partial f}{\partial x_j}(x)$ w x_0 względem x_i ($i \in \{1, ..., n\}$) to nazywamy ją *drugą pochodną cząstkową* (lub *pochodną cząstkową drugiego rzędu*) w x_0 względem x_i i x_j, i oznaczamy przez $\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}(x_0)$, zaś przez $\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$ oznaczamy odwzorowanie $x \longmapsto \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$.

W podobny sposób definiujemy pochodne cząstkowe wyższych rzędów:

$$\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(x_0), \quad \frac{\partial^4 f}{\partial x_i \partial x_j \partial x_k \partial x_l}(x_0), \quad \ldots \quad \text{itd.}$$

8.2. **Różniczki wyższego rzędu.**

Idea:

Niech $f\colon G \overset{\text{otw.}}{\subset} \mathbb{R}^n \to \mathbb{R}^m$ i f różniczkowalna w G. Wtedy $Df(x_0) \in L(\mathbb{R}^n, \mathbb{R}^m) \quad \forall x_0 \in G$.
Wówczas $Df\colon G \overset{\text{otw.}}{\subset} \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^m)$. Jeśli Df jest różniczowalna w G, to

$$D^2 f(x_0) = D(Df)(x_0) \in L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m)) \backsimeq L_2(\mathbb{R}^n, \mathbb{R}^m) \quad \forall x_0 \in G.$$

Wówczas $D^2 f\colon G \overset{\text{otw.}}{\subset} \mathbb{R}^n \to L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m))$. Postępując tak dalej otrzymujemy różniczki wyższego rzędu.

Definicja 18 (indukcyjna różniczki k-tego rzędu)**.** Załóżmy, że:

(1) $f\colon G \overset{\text{otw.}}{\subset} \mathbb{R}^n \to \mathbb{R}^m$ ma $(k-1)$ różniczkę w x_0,
(2) dla dowolnych $X_1, ..., X_{k-1} \in \mathbb{R}^n$ odwzorowanie

$$x \longmapsto \omega_{X_1,....,X_{k-1}}(x) = D^{k-1} f(x) X_1 \ldots X_{k-1}$$

działające z G do $\mathbb{R}^m$ jest różniczkowalne w x_0.

Wtedy odwzorowanie $(X, X_1, ..., X_{k-1}) \longmapsto D\omega_{X_1,....,X_{k-1}}(x_0)X$ jest *k-tą różniczką odwzorowania f w punkcie x_0*, którą oznaczamy $D^k f(x_0)$.

Zatem $D^k f(x_0)\colon \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k \text{ razy}} \to \mathbb{R}^m$ jest odwzorowaniem k-liniowym, czyli $D^k f(x_0) \in L_k(\mathbb{R}^n, \mathbb{R}^m)$ – rodzina k-liniowych odwzorowań $\underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k} \to \mathbb{R}^n$.

Zauważmy, że $L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m)) \backsimeq L_2(\mathbb{R}^n, \mathbb{R}^m)$ i ogólnie

$$\underbrace{L(\mathbb{R}^n, L(, \mathbb{R}^n, L(\ldots, L(\mathbb{R}^n, \mathbb{R}^m)\ldots)}_{k \text{ razy}} \backsimeq L_k(\mathbb{R}^n, \mathbb{R}^m).$$

8.3. **Odwzorowania klasy C^k.**

Twierdzenie 15. *Niech $f\colon G \overset{\text{otw.}}{\subset} \mathbb{R}^n \to \mathbb{R}^m$, $x_0 \in G$. Odwzorowanie f jest k-krotnie różniczkowalne w x_0 ($k \geqslant 2$) wtedy i tylko wtedy, gdy są spełnione warunki:*

(1) f jest $(k-1)$-krotnie różniczkowalna w pewnym otoczeniu punktu x_0
(2) wszystkie pochodne cząstkowe $(k-1)$-ego rzędu odwzorowania f są różniczkowalne w punkcie x_0.

Wówczas

$$D^k f(x_0) h^{(1)} \ldots h^{(k)} = \sum_{i_1,\ldots,i_k=1}^{n} h^{(1)}_{i_1} \ldots h^{(k)}_{i_k} D_{i_1} \ldots D_{i_k} f(x_0)$$

Dowód:

Indukcja po $k = 2, 3, \dots$.

$\Leftarrow$ Niech f – funkcja $(k-1)$ różniczkowalna w każdym punkcie pewnego otoczenia V punktu x_0. Oznaczmy

$$w_{h^{(1)}\dots h^{(k-1)}}(x) = D^{k-1}f(x)h^{(1)}\dots h^{(k-1)} \quad \text{dla} \quad x \in V, \quad h^{(1)}\dots h^{(k-1)} \in \mathbb{R}^n. \tag{7}$$

Wówczas

$$w_{h^{(1)}\dots h^{(k-1)}}(x) = \sum_{i_1,\dots,i_{k-1}=1}^{n} h^{(1)}_{i_1}\dots h^{(k-1)}_{i_{k-1}} D_{i_1}\dots D_{i_{k-1}}f(x). \tag{8}$$

Ponieważ z 2) $D_{i_1}\dots D_{i_{k-1}}f(x)$ są różniczkowalne w x_0, to $w_{h^{(1)}\dots h^{(k-1)}}(x)$ też jest różniczkowalne w x_0 , a to oznacza, że $f(x)$ jest k-krotnie różniczkowalna w x_0 i dla każdego $h^{(k)} \in \mathbb{R}^n$ zachodzi $D^k f(x_0)h^{(1)}\dots h^{(k)} = \sum\limits_{i_1,\dots,i_k=1}^{n} h^{(1)}_{i_1}\dots h^{(k)}_{i_k} D_{i_1}\dots D_{i_k} f(x_0)$.

$\Rightarrow$ Jeśli f jest k-krotnie różniczkowalna w x_0, to $w_{h^{(1)}\dots h^{(k-1)}}(x)$ określone przez (7) jest różniczkowalne w x_0 i zachodzi wzór (8). Podstawiając $h^{(1)} = e_{i_1}$, ..., $h^{(k-1)} = e_{i_{k-1}}$ dostajemy

$$w_{h^{(1)}\dots h^{(k-1)}}(x) = D_{i_1}\dots D_{i_{k-1}}f(x),$$

z czego wynika różniczkowalność $D_{i_1}\dots D_{i_{k-1}}f$ w x_0. □

Z odpowiedniego twierdzenia rachunku różniczkowego pierwszego rzędu wynika od razu następujący warunek wystarczający k-krotnej różniczkowalności:

Twierdzenie 16. *Jeśli w pewnym otoczeniu punktu x_0 istnieją ciągłe wszystkie pochodne cząstkowe $k-1$ rzędu $D_{i_1}\dots D_{i_{k-1}}f(x)$ dla $i_1,\dots,i_{k-1} = 1,\dots,n$ oraz istnieją w pewnym otoczeniu tego punktu wszystkie k-tego rzędu pochodne cząstkowe $D_{i_1}\dots D_{i_k}f(x)$ i są one ciągłe w x_0, to f jest k-krotnie różniczkowalna w x_0.*

Definicja 19. Odwzorowanie $f\colon G \overset{\text{otw.}}{\subset} \mathbb{R}^n \to \mathbb{R}^m$ nazywamy *klasy* C^k jeśli jest ono k-krotnie różniczkowalne oraz dla dowolnych $h^{(1)},\dots,h^{(k)} \in \mathbb{R}^n$ odwzorowanie $x \longmapsto D^k f(x)h^{(1)}\dots h^{(k)}$ jest ciągłe dla $x \in G$.

8.4. **Twierdzenie Schwarza o symetrii drugiej różniczki.**

Definicja 20. Mówimy, że odwzorowanie 2-liniowe $A \in L_2(\mathbb{R}^n, \mathbb{R}^m)$ jest *symetryczne* jeśli

$$A(X,Y) = A(Y,X) \quad \forall X, Y \in \mathbb{R}^n.$$

Ogólnie mówimy, że odwzorowanie k-liniowe $A \in L_k(\mathbb{R}^n, \mathbb{R}^m)$ jest *symetryczne* jeśli

$$A(X_1,\dots,X_k) = A(X_{\sigma(1),\dots,\sigma(k)}) \quad \forall X_1,\dots,X_k \in \mathbb{R}^n \quad \forall\sigma \text{ – permutacji ciągu } \{1,\dots,k\}.$$

Twierdzenie 17 (Schwarza o symertii drugiej różniczki)**.** *Niech $f\colon G \overset{otw.}{\subset} \mathbb{R}^n \to \mathbb{R}^m$. Załóżmy, że istnieje druga różniczka odwzorowania f w $x_0 \in G$. Wtedy $D^2 f(x_0)$ jest odwzorowaniem dwuliniowym symetrycznym z $\mathbb{R}^n$ w $\mathbb{R}^m$ (tzn. $D^2 f(x_0) \in L_2(\mathbb{R}^n, \mathbb{R}^m)$).*

Dowód:

Bez zmniejszania ogólności możemy przyjąć, że $f\colon G_1 \overset{\text{otw.}}{\subset} \mathbb{R}^2 \to \mathbb{R}$ i wiemy, że $D^2 f(x_0)$ istnieje. Musimy wykazać, że

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0) \overset{?}{=} \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0).$$

Możemy też założyć, że $x_0 = (0,0)$ oraz, że $\frac{\partial f}{\partial x_1}(0,0) = \frac{\partial f}{\partial x_2}(0,0) = 0$, gdyż przypadek ogólny sprowadza się do tego przypadku przez podstawienie

$$(x_1, x_2) \longmapsto f(x_1,x_2) - x_1\frac{\partial f}{\partial x_1}(0,0) - x_2\frac{\partial f}{\partial x_2}(0,0).$$

Oznaczmy $a_{12} = \frac{\partial^2 f}{\partial x_1 \partial x_2}(0,0)$, $a_{21} = \frac{\partial^2 f}{\partial x_2 \partial x_1}(0,0)$, $a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(0,0)$ dla $i,j = 1,2$.

Ustalmy liczbę $a > 0$, tak by kwadrat $K = \{(x_1,x_2)\colon |x_1| < a, |x_2| < a\} \subset G_1$ oraz istnieją pochodne $\frac{\partial f}{\partial x_1}(x)$ i $\frac{\partial f}{\partial x_2}(x)$ dla $x \in K$.

Oznaczmy przez $g(t) = f(t,t) - f(0,t) - f(t,0) + f(0,0)$ dla $t \in (0,a)$ przyrost odwzorowania f na kwadracie $[0,t] \times [0,t]$.

Wykażemy, że istnieje granica $\lim_{t\to 0^+} \frac{1}{t^2} g(t)$ i jest ona równa a_{12} i jednocześnir a_{21}, co zakończy dowód.

Możemy założyć, że $a_{11} = a_{21} = 0$, gdyż przypadek dowolny sprowadzamy do tego szczególnego przez rozważenie odwzorowana

$$(x_1,x_2) \longmapsto f(x_1,x_2) - \frac{1}{2}x^2{}_1 a_{11} - x_1 x_2 a_{21}.$$

$\frac{\partial f}{\partial x_1}$ jest różniczkowalna w $(0,0)$, więc

$$\frac{\partial f}{\partial x_1}(x_1,x_2) - \frac{\partial f}{\partial x_1}(0,0) = D\frac{\partial f}{\partial x_1}(0,0)(x_1,x_2) + \|(x_1,x_2)\|\psi(x_1,x_2),$$

gdzie $\psi(0,0) = 0$ i ψ ciągłe w $(0,0)$. Zatem

$$\frac{\partial f}{\partial x_1}(x_1,x_2) = x_1 \frac{\partial^2 f}{\partial x_1^2}(0,0) + x_2 \frac{\partial^2 f}{\partial x_2 \partial x_1}(0,0) + \sqrt{x_1^2 + x_2^2}\cdot \psi(x_1 + x_2),$$

czyli

$$\frac{\partial f}{\partial x_1}(x_1,x_2) = \sqrt{x_1^2 + x_2^2}\cdot \psi(x_1,x_2), \quad (x_1,x_2) \in K. \tag{9}$$

Dla danego $\epsilon > 0$ istnieje $\delta > 0$ ($\delta \leqslant a$), że

$$|\psi(x_1,x_2)| < \epsilon \quad \text{dla} \quad |x_1| < \delta, \quad |x_2| < \delta. \tag{10}$$

Niech $t \in (0,\delta)$. Oznaczając $h(x_1) = f(x_1,t) - f(x_1,0)$ dla $0 \leqslant x_1 \leqslant t$ mamy $h(t) - h(0) = g(t)$,

$$h'(x_1) = \frac{\partial f}{\partial x_1}(x_1,t) - \frac{\partial f}{\partial x_1}(x_1,0) \overset{(9)}{=} \sqrt{x_1^2 + t^2}\cdot \psi(x_1,t) - x_1\psi(x_1,0).$$

Z (10)

$$|h'(x_1)| \leqslant \sqrt{x_1^2+t^2}\cdot |\psi(x_1,t)| + x_1|\psi(x_1,0)| \leqslant t\sqrt{2}\epsilon + t\epsilon = t(1+\sqrt{2})\epsilon, \quad 0 \leqslant x_1 \leqslant t.$$

Stąd

$$|h(t) - h(0)| \leqslant t \sup_{x_1 \in [0,t]} |h'(x_1)| \leqslant t^2(1+\sqrt{2})\epsilon,$$

czyli

$$\lim_{t\to 0^+} \frac{1}{t^2} g(t) = 0 = a_{21}.$$

Analogicznie wykazujemy, że $\lim_{t\to 0_+} \frac{1}{t^2} g(t) = 0 = a_{12}$. Stąd $a_{12} = a_{21}$, czyli

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0).$$

□

Twierdzenie Schwarza można uogólnić na k-te różniczki:

Twierdzenie 18 (Schwarza o symetrii k-tej różniczki)**.** *Jeśli $f\colon G \subset \mathbb{R}^n \to \mathbb{R}^m$ ma k-tą różniczkę w $x_0 \in G$, to $D^k f(x_0)$ jest odwzorowaniem k-liniowym symetrycznym przestrzeni $\mathbb{R}^n$ w $\mathbb{R}^m$.*