Definicja: Niech $\Omega \subset \mathbb{R}^n$ będzie zbiorem otwartym. Przekształcenie $F: \Omega \to \mathbb{R}^n$ nazywamy dyfeomorfizmem (klasy C^k), gdy

- F jest różniczkowalna (klasy C^k) na Ω
- $F(\Omega) \subset \mathbb{R}^n$ jest otwarty
- $F: \Omega \to F(\Omega)$ jest bijekcją

oraz • $F^{-1}: F(\Omega) \to \Omega$ jest różniczkowalna (klasy C^k)

W rzeczywistości drugi z warunków - że $F(\Omega)$ jest otwarty - wynika z pozostałych; jest to konsekwencja Twierdzenia Brouwera o zachowaniu obszaru:

[Jeżeli $\Omega \subset \mathbb{R}^n$ jest otwarty i $F: \Omega \to \mathbb{R}^n$ jest różnowartościowa i ciągła, to $F(\Omega)$ jest otwarty]

ale dowód tego twierdzenia wykracza poza zakres naszego wykładu.

Jeżeli istnieje dyfeomorfizm $F: A \to B = F(A)$, to mówimy, że zbiory A i B są dyfeomorficzne.

Wniosek z twierdzenia o funkcji odwrotnej: { i powyższego tw. Brouwera

Jeżeli $\Omega \subset \mathbb{R}^n$ jest otwarty, $F: \Omega \to \mathbb{R}^n$ jest różnowartościowa i dla każdego $x \in \Omega$ $DF(x)$ jest odwracalna, to F jest dyfeomorfizmem Ω na $F(\Omega)$.

Oczywista uwaga: Jeżeli $F: \Omega \to F(\Omega)$ jest dyfeomorfizmem to $F^{-1}: F(\Omega) \to \Omega$ też.

Zadanie: Wskaż (podaj wzór) dyfeomorfizmu

- koła $\{x \in \mathbb{R}^2 : \|x\| < 1\}$ i płaszczyzny $\mathbb{R}^2$
- koła $\{x \in \mathbb{R}^2 : \|x\| < 1\}$ i ćwiartki $\{x_1, x_2 > 0\} \subset \mathbb{R}^2$
- pasa $\{x_1 \in (0,1), x_2 \in \mathbb{R}\}$ i półpłaszczyzny $\{x_1 > 0, x_2 \in \mathbb{R}\}$

Wykaż, że ~~wnętrze ko~~ koło $\{\sqrt{x^2+y^2} < 1\} \subset \mathbb{R}^2$ i kwadrat $\{|x| + |y| \leq 1\} \subset \mathbb{R}^2$ są dyfeomorficzne.

Uwaga (przypomnienie?)

Przekształcenie $f: A \to B$ między przestrzeniami metrycznymi nazywamy homeomorfizmem, gdy jest ciągłą bijekcją i przekształcenie odwrotne $f^{-1}: B \to A$ też jest ciągłe. Oczywiście dyfeomorfizm jest homeomorfizmem, ale dyfeomorfizmy mamy zdefiniowane na otwartych podzbiorach $\mathbb{R}^n$ (ew. przestrzeni Banacha) bo takie funkcje umiemy różniczkować; homeomorfizmy mogą być określone na dowolnych podzbiorach $\mathbb{R}^n$.

Zadanie: Zbiór Cantora $C \subset [0,1]$ i jego kwadrat $C \times C \subset [0,1]^2$ są homeomorficzne.

Rozmaitości zanurzone

<u>Definicja</u>: Zbiór $M\subset\mathbb{R}^n$ nazywamy zanurzoną rozmaitością wymiaru m, klasy C^k, jeżeli dla każdego $p\in M$ istnieje (1) jego otoczenie $\mathcal{U}$ (tzn $\mathcal{U}\subset\mathbb{R}^n$ otwarty, $p\in\mathcal{U}$) ~~takie~~,

(2) zbiór otwarty $V\subset\mathbb{R}^m$

(3) przekształcenia $\Phi:\mathcal{U}\cap M\to V$ mapa

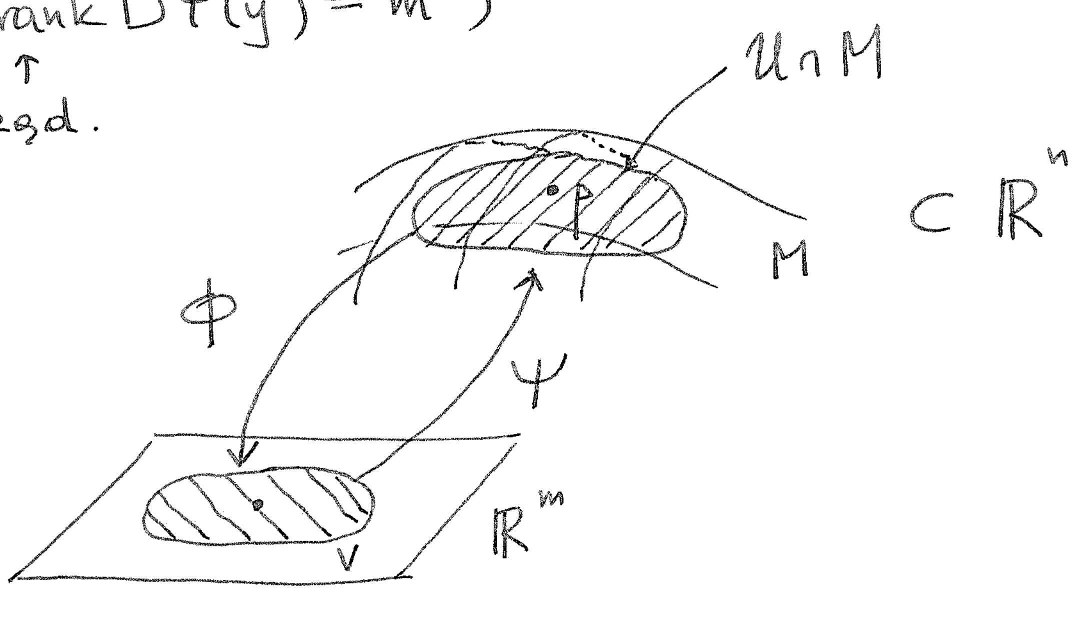
$\Psi: V\to\mathbb{R}^n$ lokalna parametryzacja

takie, że

$$\left.\begin{array}{l}\bullet\ \Phi \text{ jest ciągłe}\\ \bullet\ \Psi \text{ jest klasy } C^k\\ \bullet\ \Psi=\Phi^{-1}, \text{ czyli } \Psi\circ\Phi=\mathrm{id}_{\mathcal{U}\cap M}\\ \qquad\qquad\qquad\ \ \Phi\circ\Psi=\mathrm{id}_V\end{array}\right\}\Rightarrow \Phi \text{ jest homeomorfizmem}$$

oraz • $D\Psi(y)$ jest monomorfizmem liniowym we wszystkich punktach $y\in V$

(czyli $\underset{\substack{\uparrow\\ \text{rząd.}}}{\mathrm{rank}}\, D\Psi(y)=m$)



Mamy kluczowe twierdzenie o trzech równoważnych opisach rozmaitości zanurzonych.

Twierdzenie: Następujące warunki są równoważne.

(1) $M \subset \mathbb{R}^n$ jest zanurzoną podrozmaitością wymiaru m, klasy C^k (czyli każdy punkt $p \in M$ ma otoczenie $\mathcal{U}_1$, że ~~które~~ jego przecięcie z M jest płatem / dziedziną mapy / obrazem lok. parametryzacji)

(2) każdy punkt $p \in M$ ma otoczenie $\mathcal{U}_2$ takie, że $\mathcal{U}_2 \cap M$ jest wykresem funkcji klasy C^k m zmiennych, tzn. możemy wybrać m współrzędnych o numerach $1 \leq i_1 < i_2 < \ldots < i_m \leq n$ oraz zbiór otwarty $V_2 \subset \mathbb{R}^m$ i funkcję klasy C^k $f: V_2 \to \mathbb{R}^{n-m}$, że gdy pozostałe $n-m$ współrzędnych oznaczymy $\{j_1, \ldots, j_{n-m}\} = \{1, 2, \ldots, n\} \setminus \{i_1, \ldots, i_m\}$

to dla każdego $x \in \mathcal{U}_2 \cap M$ mamy

$$x_i = f_l(x_{i_1}, \ldots, x_{i_m}) \quad \text{gdy } i = j_l \text{ dla pewnego } l \in \{1, \ldots, n-m\}.$$

(3) każdy punkt $p \in M$ ma otoczenie $\mathcal{U}_3$ takie, że $\mathcal{U}_3 \cap M$ jest poziomicą regularną funkcji klasy C^k, tzn. istnieje funkcja $F: \mathcal{U}_3 \to \mathbb{R}^{n-m}$ taka, że

$$M \cap \mathcal{U}_3 = F^{-1}(\underset{\mathbb{R}^{n-m}}{0}) \quad \text{oraz} \quad \underbrace{\forall_{x \in M \cap \mathcal{U}_3} \; \underset{\substack{\text{macierz} \\ (n-m) \times n}}{DF(x)} \text{ ma rząd } n-m}_{\text{cyli 0 jest wartością regularną funkcji } f}$$

Dowód

(1) $\Rightarrow$ (2) Ustalmy $p \in M$. Wiemy, że istnieją $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$, $\Phi : U \cap M \to \mathbb{R}^m$, $\Psi : V \to \mathbb{R}^n$ takie, jak w definicji rozmaitości; niech $\Phi(p) = q \in V$.

Dla każdego $y \in V$ $D\Psi(y)$ jest monomorfizmem, w szczególności dla $y = q$ $D\Psi(q)$ jest monomorfizmem, czyli ma rząd m ($D\Psi(q)$ jest macierzą $n \times m$).

To znaczy, że macierz

$$D\Psi(q) = \begin{pmatrix} \frac{\partial \Psi_1(q)}{\partial y_1} & \frac{\partial \Psi_1(q)}{\partial y_2} & \cdots & \frac{\partial \Psi_1(q)}{\partial y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \Psi_n(q)}{\partial y_1} & \frac{\partial \Psi_n(q)}{\partial y_2} & \cdots & \frac{\partial \Psi_n(q)}{\partial y_m} \end{pmatrix}$$

ma niezerowy minor $m \times m$, czyli po wykreśleniu pewnych $n-m$ wierszy o numerach $1 \leq j_1 < j_2 < \ldots < j_{n-m} \leq n$ zostaną wiersze o numerach $1 \leq i_1 < i_2 < \ldots < i_m \leq n$ takie, że $\det \left(\frac{\partial \Psi_{i_l}}{\partial y_s} \right)_{l,s=1}^{m} \neq 0$,

a więc gdy przyjmiemy $\tilde{\Psi} = (\Psi_{i_1}, \ldots, \Psi_{i_m}) : V \to \mathbb{R}^m$, to $D\tilde{\Psi}(q)$ jest odwracalna. Z twierdzenia o funkcji odwrotnej istnieje takie otoczenie $\tilde{V} \subset V$ punktu q, że $\tilde{\Psi}$ jest ~~odwracalna~~ bijekcją $\tilde{V}$ na $\tilde{\Psi}(\tilde{V})$ (tzn. i odwrotna $\tilde{\Psi}^{-1} : \tilde{\Psi}(\tilde{V}) \to \tilde{V}$ też jest klasy C^k (oczywiście $\tilde{\Psi}$ jest klasy C^k, bo $\Psi \in C^k$ z założenia), a więc $\tilde{\Psi} : \tilde{V} \to \tilde{\Psi}(V)$ jest dyfeomorfizmem klasy C^k.

Mamy $\tilde{\Psi}\circ\tilde{\Psi}^{-1} = \mathrm{id}_{\tilde{\Psi}(\tilde{V})}$, $\tilde{\Psi}(\tilde{V})$ to otwarty podzbiór $\mathbb{R}^m$, czyli jeżeli $y=(y_1,\dots,y_m)\in\tilde{\Psi}(\tilde{V})$, to

$$\tilde{\Psi}\circ\tilde{\Psi}^{-1}(y) = (\Psi_{i_1}(\tilde{\Psi}^{-1}(y)), \Psi_{i_2}(\tilde{\Psi}^{-1}(y)), \dots, \Psi_{i_m}(\tilde{\Psi}^{-1}(y))) = (y_1,\dots,y_m)$$

Oznaczmy $\tilde{\Psi}(\tilde{V}) = V_2$ i rozważmy $\Psi\circ\tilde{\Psi}^{-1}: V_2 \to \mathbb{R}^n$.

$$\Psi\circ\tilde{\Psi}^{-1}(y_1,\dots,y_m) = (\Psi_1(\tilde{\Psi}^{-1}(y)), \Psi_2(\tilde{\Psi}^{-1}(y)), \dots, \Psi_n(\tilde{\Psi}^{-1}(y)))$$

i $\Psi_j(\tilde{\Psi}^{-1}(y)) = y_\ell$ gdy $j = i_\ell$

To znaczy, że $\Psi(\tilde{\Psi}^{-1}(V_2)) = \Psi(\tilde{V}) \subset \mathbb{R}^n$ jest wykresem funkcji $f: V_2 \to \mathbb{R}^{n-m}$, $f_\ell = \Psi_{j_\ell}\circ\tilde{\Psi}^{-1}$

Pozostaje ~~czysto~~ czysto techniczny problem: wskazać $\mathcal{U}_2$ t.ż. $\Psi(\tilde{V}) = \mathcal{U}_2\cap M$. Wiemy, że $\phi: \mathcal{U}\cap M \to V$ jest homeomorfizmem, więc $\phi^{-1}(\tilde{V}) = \Psi(\tilde{V})$ jest otwartym podzbiorem $\mathcal{U}\cap M$:

dla każdego $x\in\Psi(\tilde{V})$ istnieje $\varepsilon_x>0$ t.ż. $\underbrace{B(x,\varepsilon_x)\cap(\mathcal{U}\cap M)}$ jest zawarta w $\Psi(\tilde{V})$.

punkty $\mathcal{U}\cap M$ odległe od x o mniej niż ε_x

Rozważmy $\tilde{\mathcal{U}}_2 = \bigcup_{x\in\Psi(\tilde{V})} B(x,\varepsilon_x)$. Wtedy $\mathcal{U}_2$ jest zbiorem otwartym, $\Psi(\tilde{V}) \subset \tilde{\mathcal{U}}_2\cap M\cap\mathcal{U} \subset \Psi(\tilde{V})$, więc $\Psi(\tilde{V}) = M\cap(\mathcal{U}\cap\tilde{\mathcal{U}}_2)$

Wystarczy zatem wziąć $\mathcal{U}_2 = \mathcal{U}\cap\tilde{\mathcal{U}}_2$.

(2) $\Rightarrow$ (1) W tę stronę bardzo łatwo, bo wykres funkcji mamy w zasadzie od razu sparametryzowany:

bierzemy $V = V_2$, $\mathcal{U} = \mathcal{U}_2$

i dla $y = (y_1, \dots, y_m) \in V_2$ kładziemy

$$(\Psi(y))_i = \begin{cases} y_l & \text{gdy } i = i_l \\ f_l(y_1, \dots, y_m) & \text{gdy } i = j_l \end{cases}.$$

(2) $\Rightarrow$ (3) Niech $\mathcal{U}_3 = \mathcal{U}_2$ i $\big(F(x_1, \dots, x_n)\big)_l = x_{j_l} - f_l(x_{i_1}, \dots, x_{i_m})$

Wtedy $F(x) = \underset{\mathbb{R}^{n-m}}{0} \Leftrightarrow \big(F(x)\big)_l = 0$ dla $l = 1, 2, \dots, n-m$,

$$\Updownarrow$$

$x_{j_l} = f_l(x_{i_1}, \dots, x_{i_m})$, czyli $F^{-1}(0) = \mathcal{U}_2 \cap M$

Mamy $DF(x) = \left(\frac{\partial F(x)}{\partial x_1}, \dots, \frac{\partial F(x)}{\partial x_n}\right)$ i $\frac{\partial F(x)}{\partial x_{j_l}} = e_l$.

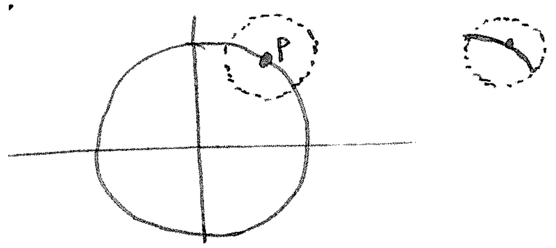
Stąd ~~wiersze~~ kolumny o numerach $j_1, j_2, \dots, j_{n-m}$ są liniowo niezależne, czyli $\operatorname{rank} DF(x) = n-m$; oczywiście $F \in C^k$ gdy $f \in C^k$.

(3) $\Rightarrow$ (2) to jest po prostu twierdzenie o funkcji uwikłanej, szczegóły zostawiam Państwu jako ćwiczenie.

Najprostszy nietrywialny przykład:

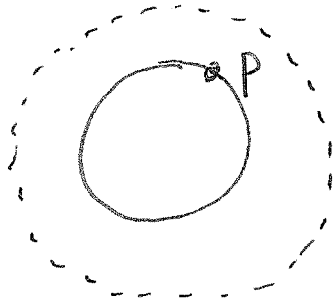
przyjrzyjmy się okręgowi jednostkowemu

$\mathbb{S}^1 \subset \mathbb{R}^2$.



gdy weźmę małe otoczenie $\mathcal{U}$ punktu $p \in \mathbb{S}^1$, to $\mathcal{U} \cap \mathbb{S}^1$ wygląda prawie jak odcinek.

Ale gdy weźmę duże otoczenie $\mathcal{U} \ni p$, to

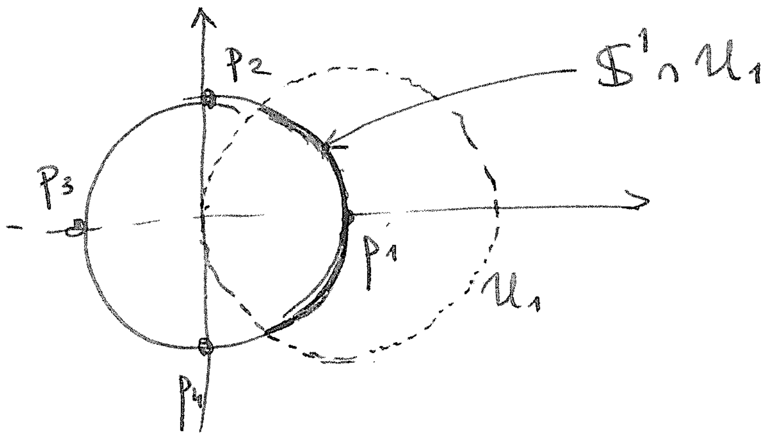


$\mathcal{U} \cap \mathbb{S}^1 = \mathbb{S}^1$ nie będzie homeomorficzne z odcinkiem, na przykład dlatego, że spójność zachowuje się przy przekształceniach ciągłych, więc jeżeli dwa zbiory są homeomorficzne, to albo oba są spójne, albo oba są niespójne. $\mathbb{S}^1$ i odcinek (otwarty) są wprawdzie oba spójne, ale gdy z każdego z nich wyjmiemy po jednym punkcie: $\mathbb{S}^1 \setminus \{x\}$, odcinek $\setminus \{\Phi(x)\}$, to $\mathbb{S}^1$ bez punktu będzie nadal spójne, a odcinek bez punktu już nie...

Inny powód: $\mathbb{S}^1 \subset \mathbb{R}^2$ jest zupełny, a odcinek (otwarty) nie, choć przekształcenie $\Phi : \mathbb{S}^1 \to V$ jest jednostajnie ciągłe, więc przekształca ciągi Cauchy'ego w ciągi Cauchy'ego.

Wykażemy na 2 sposoby, że $\mathbb{S}^1$ jest rozmaitością (klasy C^∞) wymiaru 1 w $\mathbb{R}^2$.

Jeszcze prosta uwaga: jeżeli dla pewnego $p \in M$ ustaliliśmy $\mathcal{U}, V, \Phi$ i Ψ ~~wybra~~ jak w definicji rozmaitości, to są one dobre dla każdego $x \in \mathcal{U} \cap M$



Niech $p_1 = (1,0)$, $\mathcal{U}_1 = B(p_1, 1)$, $\Phi_1(x,y) = y$, $\Psi_1(t) = (\sqrt{1-t^2}, t)$.
Łatwo można sprawdzić (proszę to zrobić!), że $V_1 = (-\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2})$ i $\mathcal{U}_1, \Phi_1, \Psi_1, V_1$ spełniają wszystkie warunki z definicji rozmaitości, w szczególności $D\Psi_1(t) = (-\frac{t}{\sqrt{1-t^2}}, 1)$ jest dla każdego $t \in V_1$ rzędu 1.

Dla p_2, p_3, p_4 definiujemy analogicznie

$\mathcal{U}_2 = B(p_2, 1)$, $\Phi_2(x,y) = x$, $\Psi_2(t) = (t, \sqrt{1-t^2})$.

$\mathcal{U}_3 = B(p_3, 1)$, $\Phi_3(x,y) = y$, $\Psi_3(t) = (-\sqrt{1-t^2}, t)$

$\mathcal{U}_4 = B(p_4, 1)$, $\Phi_4(x,y) = x$, $\Psi_4(t) = (t, -\sqrt{1-t^2})$

i $V_2 = V_3 = V_4 = (-\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2})$.

Zbiory $\mathcal{U}_1, \mathcal{U}_2, \mathcal{U}_3, \mathcal{U}_4$ pokrywają cały okrąg $\mathbb{S}^1$, więc dla każdego $p \in \mathbb{S}^1$ któryś z nich, wraz z zestawem Φ_i, Ψ_i, V_i, jest dobry.

<u>Definicja</u>: Zbiór map na rozmaitości, których dziedziny ją pokrywają, nazywamy <u>atlasem</u>

Dziedzinę pojedynczej mapy / obraz pojedynczej parametryzacji lokalnej nazywamy czasami <u>płatem</u>

Okrąg $\mathbb{S}^1$ możemy sparametryzować oszczędniej, używając tylko dwóch map, na przykład tak:

dla każdego $z \in \mathbb{C}\setminus\{0\}$ mamy funkcję $\mathrm{Arg}\, z \in [0, 2\pi)$ (nieciągłą na $\{\mathrm{Re}\, z \geq 0, \mathrm{Im}\, z = 0\}$), $z = |z| e^{i \mathrm{Arg}\, z}$, czyli $\mathrm{Arg}\, z$ = kąt między wektorem z a dodatnią półosią Ox.

$\mathbb{S}^1 \subset \mathbb{R}^2 \simeq \mathbb{C}$ to $\{e^{i\varphi} : \varphi \in [0, 2\pi)\}$

Bierzemy $U_1 = B(1,2) = \{z \in \mathbb{C} : |z-1| < 2\}$ $\left(p_1 = 1 \in \mathbb{C} \atop (1,0) \in \mathbb{R}^2\right)$

$$\Phi_1(z) = \begin{cases} \mathrm{Arg}\, z & : \mathrm{Im}\, z \geq 0 \\ \mathrm{Arg}\, z - 2\pi & : \mathrm{Im}\, z < 0 \end{cases}$$

(proszę sprawdzić, że Φ_1 jest ciągła i różnowartościowa na $U_1 \cap \mathbb{S}^1$).

$$V_1 = (-\pi, \pi), \quad \Psi_1(t) = e^{it} = (\cos t, \sin t)$$

I analogicznie, dla $p = -1 = (-1, 0)$,

$U_{-1} = B(-1, 2) = \{z \in \mathbb{C} : |z+1| < 2\}$

$\Phi_{-1} = \mathrm{Arg}\, z$, $V_{-1} = (0, 2\pi)$

$\Psi_{-1}(t) = e^{it} = (\cos t, \sin t)$.

Jak już wiemy, mamy 3 możliwe sposoby opisu rozmaitości różniczkowej $M \subset \mathbb{R}^n$ wymiaru m, klasy C^k, w otoczeniu punktu $p \in M$:

(1) opis parametryczny: dla pewnego otwartego $\mathcal{U}_1 \subset \mathbb{R}^n$, $p \in \mathcal{U}_1$, istnieje $V_1 \subset \mathbb{R}^m$ otwarty, parametryzacja $\Psi : V_1 \to \mathbb{R}^n$ taka, że Ψ jest klasy C^k, $\Psi(V_1) = M \cap \mathcal{U}_1$, $\Phi = \Psi^{-1} : M \cap \mathcal{U}_1 \to V_1$ jest homeomorfizmem, $D\Psi(y)$ jest monomorfizmem $\forall_{y \in V_1}$.

(2) opis przez wykres: dla pewnego otwartego $\mathcal{U}_2 \subset \mathbb{R}^n$, $p \in \mathcal{U}_2$ istnieje $V_2 \subset \mathbb{R}^m$ otwarty i $f : V_2 \to \mathbb{R}^{n-m}$ klasy C^k taka, że dla pewnych $1 \leq i_1 < i_2 < \dots < i_m \leq n$, $1 \leq j_1 < j_2 < \dots < j_{n-m} \leq n$ tż. $\{i_1, \dots, i_m, j_1, \dots, j_{n-m}\} = \{1, 2, \dots, n\}$ mamy
$\mathcal{U}_2 \cap M = \{x \in \mathbb{R}^n : (x_{i_1}, \dots, x_{i_m}) \in V_2, (x_{j_1}, \dots, x_{j_{n-m}}) = f(x_{i_1}, \dots, x_{i_m})\}$

(3) opis przez poziomicę: istnieje $\mathcal{U}_3 \subset \mathbb{R}^n$ otwarty, $p \in \mathcal{U}_3$, oraz $F : \mathcal{U}_3 \to \mathbb{R}^{n-m}$ klasy C^k takie, że $\mathcal{U}_3 \cap M = F^{-1}(0)$, i dla każdego $x \in \mathcal{U}_3 \cap M$ $DF(x)$ jest epimorfizmem liniowym.

Jak w każdym z tych przypadków wyznaczyć stożek styczny T_pM?

<u>Przypadek (1)</u>.

Jeżeli $0 \neq v \in T_pM$, to istnieje ciąg punktów $p_\ell \to p$, $p_\ell \in M \setminus \{p\}$, taki, że $\frac{p_\ell - p}{\|p_\ell - p\|} \longrightarrow \frac{v}{\|v\|}$. Bez straty ogólności możemy założyć, że $\forall_\ell$ $p_\ell \in \mathcal{U}_1 \cap M$; oznaczmy $q_\ell = \Phi(p_\ell)$, $q = \Phi(p)$. Skoro Φ jest homeomorfizmem, wiemy, że $q_\ell \neq q$ $\forall_\ell$.

$\frac{q_l - q}{\|q_l - q\|} \in \mathbb{S}^{m-1}$, a to jest zbiór zwarty, możemy więc z ciągu (p_l) wybrać podciąg (p_{l_s}) t.ż. $\frac{q_{l_s} - q}{\|q_{l_s} - q\|}$ jest zbieżny do jakiegoś $w \in \mathbb{S}^{m-1}$. Dla uproszczenia zapisu możemy przyjąć, że ciąg (p_l) ma już tę własność (tzn. $\frac{q_l - q}{\|q_l - q\|} \to w$).

Wtedy $\frac{v}{\|v\|} = \lim_{l\to\infty} \frac{p_l - p}{\|p_l - p\|} = \lim_{l\to\infty} \frac{\Psi(q_l) - \Psi(q)}{\|\Psi(q_l) - \Psi(q)\|} =$

$= \lim_{l\to\infty} \frac{D\Psi(q)(q_l - q) + R(q_l - q)}{\|\text{tu to samo, co w liczniku}\|} = \lim_{l\to\infty} \frac{D\Psi(q)\frac{q_l - q}{\|q_l - q\|} + R(q_l - q)/\|q_l - q\|}{\|\text{tu to samo, co w liczniku}\|} =$

$= \frac{D\Psi(q)w}{\|D\Psi(q)w\|} = D\Psi(q)\frac{w}{\|D\Psi(q)w\|}$

to nie jest zero, bo $D\Psi(q)$ jest monomorfizmem liniowym, a $w \in \mathbb{S}^{m-1}$, więc $w \neq 0$.

Stąd $v = D\Psi(q)\left(\frac{\|v\|}{\|D\Psi(q)w\|} \cdot w\right)$, więc $v \in D\Psi(q)(\mathbb{R}^m)$.

Tym samym $T_pM \subset D\Psi(q)(\mathbb{R}^m)$. Wykażemy, że to zawieranie jest w rzeczywistości równością. Niech bowiem $0 \neq v \in D\Psi(q)(\mathbb{R}^m)$, a więc $v = D\Psi(q)w$ dla pewnego $w \in \mathbb{R}^m$, $w \neq 0$.

Niech $q_l = q + \frac{1}{l}w$, $p_l = \Psi(q_l)$ (Dla dostatecznie dużych l $q_l \in V_1$).

Mamy $\lim_{l\to\infty} \frac{p_l - p}{\|p_l - p\|} = \lim_{l\to\infty} \frac{\Psi(q_l) - \Psi(q)}{\|\Psi(q_l) - \Psi(q)\|} = \lim_{l\to\infty} \frac{D\Psi(q)(q_l - q) + R(q_l - q)}{\|\text{to samo, co w liczniku}\|}$

$= \lim_{l\to\infty} \frac{D\Psi(q)(\frac{1}{k}w) + R(\frac{1}{k}w)}{\|\text{to samo, co w liczniku}\|} = \lim_{l\to\infty} \frac{D\Psi(q)w + \frac{R(\frac{1}{k}w)}{\|\frac{1}{k}w\|} \cdot \|w\|}{\|\text{to samo, co w liczniku}\|} =$

$= \frac{D\Psi(q)w}{\|D\Psi(q)w\|} = \frac{v}{\|v\|}$, skąd $v \in T_pM$, a więc $D\Psi(q)(\mathbb{R}^m) \subset T_pM$.

Wniosek: T_pM jest m-wymiarową podprzestrzenią liniową $\mathbb{R}^n$

Przypadek (2)

Tu łatwo: opis przy pomocy wykresu jest szczególnym przypadkiem parametryzacji, wystarczy więc

$\psi_{i_\ell}(x_1,\dots,x_m) = x_\ell$, $\psi_{j_\ell}(x_1,\dots,x_m) = f_\ell(x_1,\dots,x_m)$,

więc $T_pM = D\psi(q)(\mathbb{R}^m)$ jest, z dokładnością do poprzestawiania współrzędnych, wykresem $Df(q)$:

$$D(\psi_{i_1},\dots,\psi_{i_m})(q) = Id_{\mathbb{R}^m}, \quad D(\psi_{j_1},\dots,\psi_{j_{n-m}})(q) = Df(q).$$

Przypadek (3)

Wykażemy, że $T_pM = \ker DF(p)$. Wiemy, że $\dim \ker DF(p) = m$ (bo $DF(p)$ jest ~~epimorfizmem~~ liniowym $\mathbb{R}^n \to \mathbb{R}^{n-m}$), podobnie z punktu (1) wiemy, że $\dim T_pM = m$, wystarczy więc wykazać, że $T_pM \subset \ker DF(p)$.

Niech $0 \neq v \in T_pM$, a więc istnieją $p_\ell \in M$, $p_\ell \to p$, $p_\ell \neq p$ takie, że $\frac{p_\ell - p}{\|p_\ell - p\|} \to \frac{v}{\|v\|}$. Dla dost. dużych ℓ $p_\ell \in M \cap U_3$,

więc $F(p_\ell) = F(p) = 0$ i

$$0 = \frac{F(p_\ell) - F(p)}{\|p_\ell - p\|} = DF(p)\frac{p_\ell - p}{\|p_\ell - p\|} + \frac{R(p_\ell - p)}{\|p_\ell - p\|} \xrightarrow[\ell\to\infty]{} DF(p)\frac{v}{\|v\|} = \frac{1}{\|v\|} DF(p)v.$$

Stąd $DF(p)v = 0$, czyli $v \in \ker DF(p)$. $\square$.

Lemat: Niech $U \subset \mathbb{R}^n$ będzie otwarty, $A \subset U$ i załóżmy, że $f: U \to \mathbb{R}$ przyjmuje w $a \in A$ minimum lokalne na A, tzn. istnieje $\varepsilon > 0$ $\forall_{y \in A}$ $\|y-a\| < \varepsilon \Rightarrow f(y) \geq f(a)$.

Wówczas $\forall_{v \in T_a A}$ $\langle v, \nabla f(a) \rangle \geq 0$ $(*)$

Dowód Dla $v=0$ $(*)$ zachodzi, niech zatem $0 \neq v \in T_a A$. Istnieje ciąg (p_k) punktów $A \setminus \{a\}$ taki, że $p_k \to a$, $\frac{p_k - a}{\|p_k - a\|} \to \frac{v}{\|v\|}$

Dla dost. dużych k mamy $\|p_k - a\| < \varepsilon$, więc $f(p_k) - f(a) \geq 0$, zatem

$$0 \leq \frac{f(p_k)-f(a)}{\|p_k - a\|} = Df(a)\frac{p_k - a}{\|p_k - a\|} + \frac{R(p_k - a)}{\|p_k - a\|} \to \frac{Df(a)v}{\|v\|} = \frac{\langle \nabla f(a), v\rangle}{\|v\|}.$$

Stąd $\langle \nabla f(a), v \rangle \geq 0$.

~~Wniosek~~ **Uwaga**: Tak samo dowodzimy, że jeżeli f ma w a maksimum lokalne na A, to $\langle \nabla f(a), v \rangle \leq 0$ dla wszystkich $v \in T_a A$.

Wniosek: Jeżeli U, f, A są jak wyżej, $a \in A$, f przyjmuje w a ekstremum lokalne i dodatkowo wiemy, że $T_a A$ jest przestrzenią liniową, to $\nabla f(a) \perp T_a A$, tzn.

$\forall_{v \in T_a A}$ $\langle \nabla f(a), v \rangle = 0$.

Dowód: Załóżmy dla ustalenia uwagi, że f przyjmuje w a minimum lokalne na A. Ustalmy $v \in T_a A$; wiemy z Lematu, że $\langle \nabla f(a), v \rangle \geq 0$. Z drugiej strony $-v$ też należy do $T_a A$, więc $\underbrace{\langle \nabla f(a), -v \rangle}_{= -\langle \nabla f(a), v \rangle} \geq 0$ Stąd $\langle \nabla f(a), v \rangle = 0$.

Przypadek maksimum lokalnego idzie dokładnie tak samo.

Wniosek: Jeżeli $M \subset \mathbb{R}^n$ jest rozmaitością zanurzoną klasy C^1, $U \subset \mathbb{R}^n$ jest otwarty, $g \in C^1(U, \mathbb{R})$ przyjmuje w $p \in U \cap M$ ekstremum lokalne na M, to $\nabla g(p) \perp T_pM$, tzn. $\forall_{v \in T_pM} \; \langle \nabla g(p), v \rangle = 0$.

W sytuacji, gdy M dana jest jako poziomica, wniosek ten staje się ważnym twierdzeniem z nazwiskiem.

Ale najpierw trochę terminologii.

Podstawowym zagadnieniem teorii optymalizacji jest tak zwana optymalizacja z więzami: Szukamy minimum lub maksimum funkcji $f: U \to \mathbb{R}$ wśród punktów $x \in U$ spełniających dodatkowo warunek (więzy) $F(x) = 0$, gdzie $F = (F_1, \dots, F_m) \in C^1(U, \mathbb{R}^m)$, a więc szukamy ekstremów (lokalnych lub globalnych) funkcji f na zbiorze $A = F^{-1}(0)$.

Takie ekstrema (lokalne lub globalne) nazywamy ekstremum związanym (więzami $F_1(x) = F_2(x) = \dots = F_m(x) = 0$) lub warunkowym; gdy te ekstrema lokalne są właściwe, mówimy o ekstremach związanych właściwych.

I obiecane

Twierdzenie Lagrange'a o mnożnikach (warunek konieczny istnienia ekstremum związanego lokalnego).

Niech $U \subset \mathbb{R}^n$ będzie otwarty, załóżmy, że $g \in C^1(U, \mathbb{R})$, $F = (F_1, \dots, F_m) \in C^1(U, \mathbb{R}^m)$. Jeżeli g przyjmuje w punkcie $p \in U$ ekstremum związane (warunkami $F_1(x) = \dots = F_m(x) = 0$), zaś $DF(p)$ jest epimorfizmem liniowym, to istnieją $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ takie, że

$$\nabla g(p) = \sum_{i=1}^{m} \lambda_i \nabla F_i(p) \quad \text{(warunek Lagrange'a)}$$

Liczby $\lambda_1, \dots, \lambda_m$ nazywamy mnożnikami Lagrange'a, a warunek, że $DF(p)$ jest epimorfizmem liniowym nazywa się czasami warunkiem jakości więzów.

Ten ostatni warunek najczęściej wypowiada się nieco inaczej: to, że $DF(p)$ jest epimorfizmem oznacza, że macierz $DF(p) = (DF_1(p), \dots, DF_m(p))$ ma rząd maksymalny, czyli m (to jest macierz $n \times m$). W bazie standardowej jej kolumny to $\nabla F_1(p), \dots, \nabla F_m(p)$, więc warunek na rząd oznacza, że wektory $\nabla F_1(p), \dots, \nabla F_m(p)$ są liniowo niezależne.

Dowód tw. Lagrange'a.

~~Z wcześniejszych rozważań wiemy, że $\nabla g(p) \perp T_pM$~~,

Skoro $DF(p)$ jest epimorfizmem liniowym, a $F \in C^1(\mathcal{U}, \mathbb{R}^m)$, to istnieje $\underset{\text{otwarty}}{\tilde{\mathcal{U}} \subset \mathcal{U}}$, $p \in \tilde{\mathcal{U}}$, takie, że $DF(x)$ jest epimorfizmem liniowym dla wszystkich $x \in \tilde{\mathcal{U}}$, bo warunek na bycie epimorfizmem liniowym to nieznikanie pewnego minora (podwyznacznika) $m \times m$ macierzy $DF(x)$; jeżeli jest on niezerowy w p, to jest ~~też~~ różny od zera w pewnym otoczeniu p, bo jest funkcją ciągłą zmiennej x.

Stąd $M = F^{-1}(0) \cap \tilde{\mathcal{U}}$ jest rozmaitością różniczkową klasy C^1, wymiaru $n-m$, i z wcześniejszych rozważań wiemy, że $T_pM = \ker DF(p)$ jest podprzestrzenią liniową wymiaru $n-m$.

Jeżeli g ma w p ekstremum lokalne związane (a więc ekstremum lokalne na M), to $\nabla g(p) \perp T_pM$.

Zauważmy, że

- przestrzeń prostopadła do T_pM (czyli $(T_pM)^\perp$) jest $n-(n-m)=m$ wymiarowa

- jeżeli $v \in \ker DF(p) = T_pM$, to

$$\mathbb{R}^m \ni 0 = DF(p)v = (DF_1(p)v, DF_2(p)v, \dots, DF_m(p)v) =$$
$$= (\langle \nabla F_1(p), v\rangle, \dots, \langle \nabla F_m(p), v\rangle), \text{ więc}$$
$$v \perp \nabla F_i(p) \text{ dla } i = 1, 2, \dots, m.$$

Stąd $\nabla F_i(p) \in (T_pM)^\perp$.

- wektory $\nabla F_1(p), \dots, \nabla F_m(p)$ są liniowo niezależne i jest ich m

Stąd $(T_pM)^\perp = \operatorname{span}(\nabla F_1(p), \dots, \nabla F_m(p))$ i warunek $\nabla g(p) \in (T_pM)^\perp$ oznacza, że istnieją $\lambda_1, \dots, \lambda_m$ takie, że

$$\nabla g(p) = \lambda_1 \nabla F_1(p) + \lambda_2 \nabla F_2(p) + \dots + \lambda_m \nabla F_m(p)$$

czyli właśnie warunek Lagrange'a. $\square$.

<u>Twierdzenie</u> (warunki dostateczne istnienia ekstremum lokalnego związanego właściwego).

Niech $\mathcal{U} \subset \mathbb{R}^n$ będzie otwarty, $g \in C^2(\mathcal{U}, \mathbb{R})$, $F = (F_1, \dots, F_m) \in C^2(\mathcal{U}, \mathbb{R}^m)$, przypuśćmy, że dla pewnego $p \in M = F^{-1}(0)$ pochodna $DF(p)$ jest epimorfizmem liniowym (czyli $\nabla F_1(p), \dots, \nabla F_m(p)$ są liniowo niezależne) oraz $\nabla g(p) = \sum_{i=1}^m \lambda_i \nabla F_i(p)$ (czyli spełniony jest warunek Lagrange'a). Niech $L = g - \sum_{i=1}^m \lambda_i F_i \in C^2(\mathcal{U}, \mathbb{R})$.

Wówczas ① jeżeli $D^2L(p)|_{T_pM}$ jest dodatnio określona (tzn. $\forall_{0 \neq v \in T_pM}\ D^2L(p)(v,v) > 0$), to g ma w p właściwe minimum lokalne związane, ② jeżeli $D^2L(p)|_{T_pM}$ jest ujemnie określona, to g ma w p właściwe maksimum lokalne związane,

③ jeżeli $D^2L(p)|_{T_pM}$ ma wartości własne różnych znaków, a więc istnieją $v_1, v_2 \in T_pM$ takie, że $D^2L(p)(v_1,v_1) < 0 < D^2L(p)(v_2,v_2)$, to g nie ma w p ekstremum lokalnego związanego.

<u>Dowód</u>: ① Skoro $DF(p)$ jest epimorfizmem liniowym, to (jak w dowodzie tw. Lagrange'a) istnieje $\tilde{\mathcal{U}} \subset \mathcal{U}$ otwarte, $p \in \tilde{\mathcal{U}}$, takie, że $\tilde{\mathcal{U}} \cap M$ jest rozmaitością klasy C^2, wymiaru $n-m$. Istnieje zatem lokalna parametryzacja M w otoczeniu p: $V_1 \subset \mathbb{R}^{n-m}$, $\mathcal{U}_1 \subset \tilde{\mathcal{U}}$ otwarty, $p \in \mathcal{U}_1$, $\Phi: \mathcal{U}_1 \cap M \to V_1$ homeomorfizm, $\underbrace{\Psi = \Phi^{-1}: V_1 \to \mathcal{U}_1}_{\text{klasy } C^2}$ takie, że $D\Psi(y)$ jest monomorfizmem liniowym $\forall_{y \in V_1}$, w szczególności jeżeli $q = \Phi(p)$, to $D\Psi(q)$ jest monomorfizmem.
Oznaczmy $E = D\Psi(p)(S^{n-m-1})$. Jest to zbiór zwarty, $0 \notin E$, $E \subset T_pM = D\Psi(p)(\mathbb{R}^{n-m})$. Jeżeli $D^2L(p) > 0$ na T_pM, to istnieje $\alpha > 0$ takie, że $\forall_{v \in E}\ D^2L(p)(v,v) \geq \alpha$ (z tw. Weierstrassa o przyjmowaniu kresów).

Zauważmy teraz, że dla $x \in M$ mamy $L(x) = g(x)$, co więcej $\nabla L(p) = \nabla g(p) - \sum_{i=1}^{m} \lambda_i \nabla F_i(p) = 0$, więc L ma w p punkt krytyczny. Zapiszmy wzór Taylora stopnia 2 dla L w p:

$$L(x) = L(p) + \underbrace{DL(p)(x-p)}_{\langle \nabla L(p), x-p \rangle = 0} + \frac{1}{2} D^2L(p)(x-p, x-p) + R_2(x-p),$$

więc gdy $x \in M$,

$$g(x) - g(p) = L(x) - L(p) = \frac{1}{2} D^2 L(p)(x-p, x-p) + R_2(x-p). \quad (*)$$

Jeżeli $x \in \mathcal{U}_1 \cap M$, to oznaczmy $y = \Phi(x)$. Wtedy

$$x - p = \Psi(y) - \Psi(q) = D\Psi(q)(y-q) + r(y-q);$$

wstawiając to do $(*)$ otrzymujemy (dzieląc strony przez $\|x-p\|^2$)

$$\frac{g(x)-g(p)}{\|y-q\|^2} = \frac{1}{2} D^2 L(p)\left(\frac{x-p}{\|y-q\|}, \frac{x-p}{\|y-q\|}\right) + \frac{R_2(x-p)}{\|y-q\|^2} =$$

$$= \frac{1}{2} D^2 L(p)\left(D\Psi(q)\frac{y-q}{\|y-q\|} + \frac{r(y-q)}{\|y-q\|}, D\Psi(q)\frac{y-q}{\|y-q\|} + \frac{r(y-q)}{\|y-q\|}\right) +$$

$$+ \frac{R_2(x-p)}{\|x-p\|^2} \cdot \frac{\|D\Psi(q)(y-q) + r(y-q)\|^2}{\|y-q\|^2} = \left\{ \begin{array}{l} \text{korzystam z tego,} \\ \text{że } D^2L(p) \text{ jest formą} \\ \text{dwuliniową symetryczną} \end{array} \right\}$$

$$= \frac{1}{2} D^2 L(p)\left(D\Psi(q)\frac{y-q}{\|y-q\|}, D\Psi(q)\frac{y-q}{\|y-q\|}\right) + D^2 L(p)\left(D\Psi(q)\frac{y-q}{\|y-q\|} + \frac{1}{2}\frac{r(y-q)}{\|y-q\|}, \frac{r(y-q)}{\|y-q\|}\right)$$

$$+ \frac{R_2(x-p)}{\|x-p\|^2} \cdot \left\|D\Psi(q)\frac{y-q}{\|y-q\|} + \frac{r(y-q)}{\|y-q\|}\right\|^2 = A_1 + A_2 + A_3 = \otimes$$

$$A_1 = \frac{1}{2} D^2 L(p)\Big(\underbrace{D\Psi(q)\underbrace{\frac{y-q}{\|y-q\|}}_{\in S^{n-m-1}}}_{\in E}, \underbrace{D\Psi(q)\frac{y-q}{\|y-q\|}}_{\in E}\Big) \geq \frac{1}{2}\alpha > 0$$

$$|A_2| = \left|D^2 L(p)\left(D\Psi(q)\frac{y-q}{\|y-q\|} + \frac{1}{2}\frac{r(y-q)}{\|y-q\|}, \frac{r(y-q)}{\|y-q\|}\right)\right| \leq$$

$$\leq |||D^2 L(p)||| \cdot \Big\|\underbrace{D\Psi(q)\frac{y-q}{\|y-q\|}}_{\substack{\in E, \\ E \text{ zwarty, więc ograniczony}}} + \frac{1}{2}\underbrace{\frac{r(y-q)}{\|y-q\|}}_{\substack{\to 0 \\ \text{przy } y \to q}}\Big\| \, \Big\|\underbrace{\frac{r(y-q)}{\|y-q\|}}_{\substack{\to 0 \\ \text{przy } y \to q}}\Big\| \longrightarrow 0 \quad \text{przy } y \to q$$

Analogicznie, gdy $y \to q$, to $x = \Psi(y) \to \Psi(q) = p$, zatem

$$A_3 = \underbrace{\frac{R_2(x-p)}{\|x-p\|^2}}_{\substack{\downarrow \text{ przy } y\to q, \\ \text{więc } x\to p \\ 0}} \underbrace{\| \underbrace{D\Psi(q)\frac{y-q}{\|y-q\|}}_{\in E} + \underbrace{\frac{r(y-q)}{\|y-q\|}}_{\substack{\downarrow y\to q \\ 0}} \|^2}_{\text{a więc to jest ograniczone}} \longrightarrow 0 \text{ przy } y \to q.$$

Jeżeli więc tylko y jest dostatecznie blisko q (czyli x jest dostatecznie blisko p), to $|A_2| + |A_3| < \frac{1}{4}\alpha$, więc

$$\frac{g(x)-g(p)}{\|y-q\|^2} = \otimes \geqslant A_1 - |A_2| - |A_3| \geqslant \frac{1}{2}\alpha - \frac{1}{4}\alpha = \frac{1}{4}\alpha > 0,$$

skąd $g(x) > g(p)$ i g ma w p minimum lokalne właściwe związane.

Przypadek ② dowodzi się dokładnie tak samo (odwraca się część nierówności).

③ Niech $w_1 = D\Psi(q)^{-1}v_1$, $w_2 = D\Psi(q)^{-1}v_2$ i niech

$x_k = \Psi(q + \frac{1}{k}w_1)$, $z_k = \Psi(q + \frac{1}{k}w_2)$. $\Big\}$ dla k dost. dużych, by $q + \frac{1}{k}w_1 \in V_1$, $q + \frac{1}{k}w_2 \in V_1$

Mamy $x_k \to \Psi(q) = p$, $z_k \to \Psi(q) = p$.

$$g(x_k) - g(p) = L(x_k) - L(p) = \frac{1}{2}D^2L(p)(x_k - p, x_k - p) + R_2(x_k - p);$$

$$x_k - p = \Psi(q + \tfrac{1}{k}w_1) - \Psi(q) = D\Psi(q)\cdot\tfrac{1}{k}w_1 + r(\tfrac{1}{k}w_1) =$$
$$= \tfrac{1}{k}v_1 + r(\tfrac{1}{k}w_1), \text{ więc}$$

$$g(x_k) - g(p) = \frac{1}{2k^2}\Bigg[D^2L(p)(v_1, v_1) + 2D^2L\left(v_1 + \frac{r(\frac{1}{k}w_1)}{2\cdot\frac{1}{k}\|w_1\|}\cdot\|w_1\|, \frac{r(\frac{1}{k}w_1)}{\frac{1}{k}\|w_1\|}\right)$$
$$+ \frac{R_2(x_k-p)}{\|x_k-p\|^2}\cdot\frac{\|\frac{1}{k}v_1 + r(\frac{1}{k}w_1)\|^2}{\|\frac{1}{k}w_1\|^2}\cdot\|w_1\|^2\Bigg]$$

wewnątrz nawiasu [ ] pierwszy wyraz jest ujemny, a pozostałe dążą do zera przy $k \to \infty$, więc dla dużych k

$g(x_k) < g(p)$.

Analogicznie dowodzimy, że dla dużych k $g(z_k) > g(p)$.

Stąd g nie ma w p ekstremum lokalnego związanego. □

<u>Interpretacja mnożników Lagrange'a</u>

dla ustalenia uwagi, dla minimum analogicznie

Załóżmy, że $g \in C^1(\mathcal{U}, \mathbb{R})$ przyjmuje w $p_t \in \mathcal{U}$ maksimum związane warunkami $F(x) = (F_1(x), \dots, F_m(x)) = \underbrace{(t_1, \dots, t_m)}_{= t} \in \mathbb{R}^m$ (czyli maksimum na zbiorze $F^{-1}(t_1, \dots, t_m) = F^{-1}(t)$)

Oznaczmy $v(t) = g(p_t)$ i załóżmy, że v jest różniczkowalną funkcją t.

{ ten warunek jest spełniony w otoczeniu takiego $p \in \mathcal{U}$, że $DF(p)$ jest monomorfizmem, ale to dość złożona sprawa, odłożymy ją na wykład z geometrii różniczkowej.

Niech $H : \mathcal{U} \to \mathbb{R}$,

$$H(x) = g(x) - v(F(x)).$$

Oczywiście $H(p_t) = g(p_t) - v(F(p_t)) = g(p_t) - v(t) = 0$

i dla każdego $x \in \mathcal{U}$ ~~g(x)~~, $t = F(x)$,

$$H(x) = g(x) - v(F(x)) = g(x) - g(p_t) \le 0$$

bo $x, p_t \in F^{-1}(t)$ i g ma w p_t maksimum związane.

Stąd H ma w punkcie p_t maksimum lokalne, a więc, z tw. Fermata, $\nabla H(p_t) = 0$

$$0 = \nabla H(p_t) = \nabla g(p_t) - \langle \nabla v(\overbrace{F(p_t)}^{= t}), \nabla F(p_t) \rangle =$$

$$= \nabla g(p_t) - \sum_{i=1}^{m} \frac{\partial v}{\partial t_i}(t) \cdot \nabla F_i(p_t)$$

To znaczy, że w punkcie p_t spełniony jest warunek Lagrange'a z mnożnikami Lagrange'a

$$\lambda_i = \frac{\partial v}{\partial t_i}(t)$$

To daje ważną interpretację mnożników:

Wielkość mnożnika λ_i mówi, jak szybko się zmienia wartość ekstremum związanego warunkami

$$F_1(x) = t_1, \ldots, F_m(x) = t_m$$

gdy ruszamy parametrem t_m.